Author:
Winter, Ash

Title:
The molecular basis of HicA dependent binding of HicB to DNA

The molecular basis of HicA dependent binding of HicB to DNA

Ashley James Winter

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of PhD in Chemistry in the Faculty of Science

School of Chemistry

September 2018

Word Count: 56,795

Abstract

Toxin-antitoxin (TA) systems are typically encoded as bicistronic operons in a variety of bacterial species and archaea and act as response effectors to a variety of intra and extra-cellular stressors. TA members are classed into 6 generic systems and type 2 TA systems are able to directly autoregulate their own expression via a series of complex interactions between the protein toxin, the protein antitoxin and the DNA upstream of the encoding genes. This ensures that in a non-stress environment, the protein antitoxin is in excess of the protein toxin resulting in neutralisation. During a state of stress, the intracellular toxin concentration accumulates and the toxin is able to exert its toxicity, typically resulting in bacteriostasis. This mechanism has been implicated in the generation of antibiotic tolerant (persister) cells, which can result in chronic infections. The HicAB system is one of the prototypic type 2 TA systems found in several bacteria and archaea, yet remains underinvestigated when compared to other TA systems. This thesis reports structural studies on the *Burkholderia pseudomallei* HicAB TA pair and associated DNA interactions. HicB forms a tetramer and this can interact with DNA upstream of the bicistronic *hicAB* operon to form a tight binding complex. HicA and HicB form a hetero-octameric complex that involves structural re-organisation of the carboxy-terminal (DNA binding) region of HicB. HicA therefore has a profound impact on binding of HicB to DNA sequences upstream of *hicAB*, in a stoichiometric-dependent way. At low ratios of HicA:HicB there is no effect on DNA binding, but at higher ratios the affinity for DNA declines co-operatively, driving dissociation of the HicA:HicB:DNA complex. This provides the first example of the molecular mechanisms by which HicA de-represses a HicB-DNA complex and is a distinct mechanism when compared to other TA family members investigated to date

Acknowledgements

Firstly, I would like to thank my supervisor Professor Matt Crump for all his guidance and support throughout my PhD, particularly during the last year. His patience has been extrodinary and I have grown into a more diligent worker over the years. Thanks also go to my secondary supervisor Professor Richard Titball and our collaborators at the University of Exeter (Dr Michail Isupov and Professor Nicholas Harmer), who have all offered invaluable guidance and support, including numerous scientific discussions over the years. A special mention to Dr Chris Williams, who guided me in various techniques including cloning, protein expression and purification, and to this day continues to teach me. Thanks should also go to Paul for sharing almost all of the PhD journey with me and providing great conversation (and friendship) both in and out of work.

I have been fortunate to work and collaborate with a great number of people at the University of Bristol and thanks must go to them: Ryan, Matt G, Paul, Alice B, Alice P, Erik, Tom, Lina, Claudio, Luoyi, Olly, Mark Dillingham, Guto, Tim, Kris and all BAG members throughout the years for their company. Thank you to all BSc and MSci students (Hannah, Mariya, Dan, Phil and George) who have been part of this project over several years.

Thank you to the BBSRC who funded the project and the SWBio DTP which I am fortunate to have been a part of for the past four years.

To my parents, thank you for always supporting me and being receptive to a phone call throughout the years. The encouragement and patience you have had with me over the years has been extraordinary. This piece of work would not be possible without your persistence and sacrifice to allow me to follow this pathway into Science.

Finally, the last acknowledgement must go to my partner Ella. Thank you for your eternal support throughout the last 8 years and I know I would not be at this point without you. Words cannot express how grateful I am (and you know I'm bad at putting thoughts on paper). I'm sure you will now enjoy a brief respite from my constant nattering on toxin-antitoxin systems. Thank you for being part of every step of this journey from the start of College to this point and I can't wait to take the next steps with you in our journey through life.

Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE:.....

Contents

Abbreviations

5'UTR	5' Untranslated region
AbrB	AbrB superfamily
AmpR	Ampicillin resistance
AP	Aprotinin
AR	Antibiotic resistance
ASEC	Analytical size exclusion chromatography
ASU	Asymmetric unit
AUC	Analytical ultracentrifugation
CA	Carbonic anhydrase
cAMP	Cyclic AMP
CO	Conalbumin
CRP	cAMP receptor protein
CRP-S	Sxy-activated promoter
C(s)	Continuous size distribution
CT	C-terminal
DBD	DNA binding domain
DNA	Deoxy ribonucleic acid
dsRBD	Double strand RNA binding domain
EMSA	Electophoretic mobility shift assay
ESI	Electrospray ionisation
ESMS	Electrospray mass spectrometry
ExPEC	Extraintestinal pathogenic *Escherichia coli*
FA	Fluorescence anisotropy
HCOONa	Sodium acetate
HEX	Hexachlorofluorescein
$HicA_{EC}$	*Escherichia coli* HicA
$HicA_{SP}$	*Streptococcus pneumoniae* HicA
$HicB_{EC}$	*Escherichia coli* HicB
$HicB_{SP}$	*Streptococcus pneumoniae* HicB
hip	High persistence
HSQC	Heteronuclear single quantum coherence
HTH	Helix turn helix
ITC	Isothermal titration calorimetry
IMAC	Immobilised metal affinity chromatography
K_{av}	Partial coefficient
KanR	Kanamycin resistance
LacZ	β-galactosidase
LB	Luria broth
MIC	Minimum inhibitory concentration
MIR	Multiple isomorphic replacement
MPD	2-methyl-2,4-pentanediol
MR	Molecular replacement
mRNA	Messenger ribonucleic acid
Mw_{app}	Apparent molecular weight
m/z	Mass:charge ratio
NaOAC	Sodium formate
NCS	Non crystallographic symmetry
NCM	Non-canonical motif

NH_4Cl	Ammonium chloride
NMR	Nuclear magnetic resonance
NT	N-terminal
OV	Ovalbumin
PEG	Polyethylene glycol
PDB	Protein dataBank
P(r)	Pair wise distribution
RA	Ribonuclease A
RNA	Ribonucleic acid
RBS	Ribosome binding site
RHH	Ribbon helix helix
RMSD	Root mean square deviation
RPM	Revolutions per minute
SAD	Single wavelength anomalous dispersion
SAXS	Single angle X-ray scattering
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEAUC	Sedimentation equilibrium analytical ultracentrifugation
Sec	Secretory dependent pathway
SEC	Size exclusion chromatography
SEC-MALS	SEC-Multiangle light scattering
SeMet	Selenomethionine
S/N	Supernatant
sRNA	Small RNA
SPR	Surface plasmon resonance
SVAUC	Sedimentation velocity analytical ultracentrifugation
TA	Toxin-antitoxin
TCEP	Tris (2-carboxyethyl) phosphine
tRNA	Transfer RNA
tm-RNA	Transfer-messenger RNA
TSS	Transcriptional start site
Ve	Volume elution
WHO	World Health Organisation
$\chi 2$	Normalised discrepancy Function

1. Introduction

1.1. Antibiotic resistance

Antibiotic resistance (AR) is a widespread phenomenon enabling the survival and proliferation of microbes when challenged with antibiotics.[1] AR is a considerable health and economic burden[2] and a global surveillance report has revealed high rates of resistance amongst bacteria commonly observed across all regions studied.[3] In 2017, it was reported that the current pipeline of antibiotics in development (Phase 1-3) were insufficient to stem the rise of antibiotic resistance.[4] Of the fifty-one antibiotics within the pipeline, most were simple modifications of existing antibiotics rather than novel antibiotic classes. Despite the best efforts of the WHO, antibiotic misuse by the public and indiscriminative use within agriculture have propagated the rise of AR. In 2018, it was reported that sub-minimum inhibitory concentration (MIC) levels of streptomycin can induce the generation of resistant *Salmonella enterica* over several hundred generations, with resistance conferred by a combination of mutations within several genes associated with aminoglycoside resistance (*gidB, cyoB, nuoG, trkH and znuA)* that confer a high level of resistance.[5] The ability of bacteria to resist antibiotic treatments has enabled bacterial infections to persist and proliferate within the host.

1.2. Chronic infections

Bacterial infections initially manifest in an acute variant form, typically cleared by the immune system. In some cases, the immune system does not completely eradicate this infection with a resultant sub-population of bacteria acting as a reservoir for future relapses. This gives rise to chronic infections, defined as the establishment of a persistent pathogenic bacterial presence.[6, 7] Establishment can be explained by a range of factors, including evasion from the host immune system, the formation of biofilms and adaption to an intracellular environment.[8] Each establishment factor causes a reduced efficacy of antimicrobial agents. Chronic infections often initially colonize at one location[7, 9-12], but can spread throughout the host.[9, 11] While modes of evasion, establishment and pathogenesis vary, it is clear that all

chronic infections share one common feature; they propagate from a subpopulation of cells that are not processed and eradicated by the immune system. Understanding the mechanisms of cellular survival and the resulting persistence of pathogenic bacterial infections are vital to efforts to combat chronic infections.

1.3. Persister cells

Persister cells, referred to as dormant/antibiotic tolerant cells were first described in 1944 by Joseph Bigger.[13] These are generated by a stochastic mechanism, where cells switch between a non-growing dormant antibiotic tolerant (persister) state and a non-dormant antibiotic susceptible state (Figure 1.1). The precise details of this switching mechanism are still unclear but population heterogeneity is believed to be a crucial factor.[14] The creation of a subpopulation of persister cells, typically 1% of the total population, enables the bacteria to survive a lethal stress induced for example by an antibiotic. Persistence is deemed to be a phenotypic characteristic, rather than genotypic as daughter cells do not inherit this trait. As persisters can survive numerous stressor challenges (for example, nutritional depravation or pH alteration), persisters are hypothesized to act as a failsafe priming mechanism to prevent the complete eradication of a bacterial population from a unknown future stressor.[15-17] It is therefore likely that this response to a broad range of stress factors evolved in preference to encoding individual responses for each stressor.

In 2017, experimental evidence suggested that antibiotic tolerance preceded resistance to ampicillin by several generations. The survival advantage conveyed by antibiotic tolerance enabled the remaining *Escherichia coli* to accumulate mutations in the promoter of AmpC, a β-lactamase, to a point where resistance was achieved.[18] Simulations revealed that although similar mutations accumulated in populations of antibiotic susceptible cells, these populations were destroyed prior to reaching a state of antibiotic resistance. Antibiotic tolerant cells can therefore provide a safe haven for the accumulation of mutations, increasing the likelihood that resistance is achieved. The pathway of tolerance to resistance may provide a key checkpoint in combating antibiotic resistant cells.

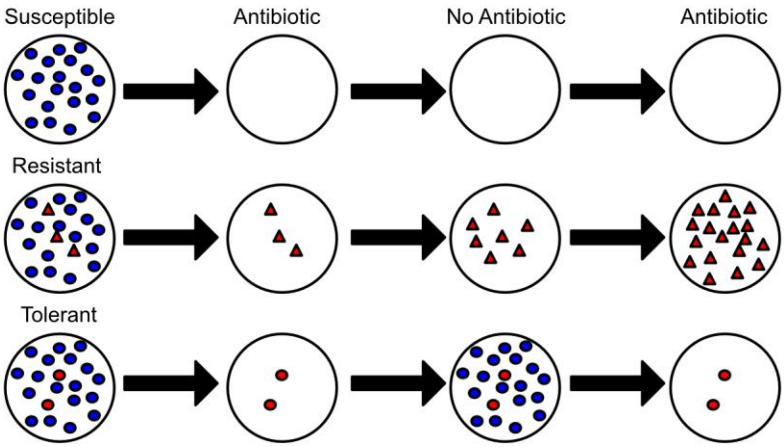


Figure 1.1. Overview of antibiotic tolerance. Antibiotic susceptible cells (blue circles) do not survive an antibiotic challenge. Resistance cells (red triangles) survive an antibiotic challenge due to a genotypic variation. These are able to replicate to ensure the next generation of cells are all antibiotic resistant. Antibiotic tolerant cells (red circles) are ~1% of the population which survive an antibiotic challenge due to phenotypic variation- resulting in a dormant cellular state. Due to the phenotypic nature of antibiotic tolerant cells- the next generation of cells consist of two distinct populations of antibiotic tolerant cells and antibiotic susceptible cells. The cycle then repeats upon further antibiotic challenges.

The underlying molecular mechanisms of persistence mediated by persister cells remained unknown[14], despite their ability to confer antibiotic tolerance within a range of bacterial biofilms.[19-27] Toxin-antitoxin (TA) systems, initially referred to as addiction modules, were first described by Gerdes in 1986.[28] These two component systems are found within numerous bacterial chromosomes and extrachromosomal plasmids.[29, 30] TA systems all share three common features observed in addiction modules.[31-35] First, each system is comprised of a toxin moiety that results in cell death[28, 36] or bacteriostasis[37, 38] and an "antidote" antitoxin that reverses toxicity. Secondly, each antitoxin is labile in contrast to their stable toxin enabling programmed cell death/stasis in response to antibiotics.[32] Lastly, antitoxins appear to autoregulate the toxin-antitoxin operon or cause plasmid addiction.

In 1983, an extensive screening study sought to identify genes upregulated in *E. coli* persister cells.[39] Mutations within the locus of a bicistronic operon (*hip*AB) enhanced persisters up to 10,000 fold and led to the identification of the first persister related gene: *hipA*.[40] This operon resulted in a toxic phenotype (cell death) in the absence of *hipB*.[40-42] HipA and HipB were later shown to form a stable complex that was able to repress the transcriptional activity of the *hipAB* promoter.[43] *hipAB* was proposed to encode a TA system due to the presence of the three archetypal features outlined above.[44] TA systems had been implicated in programmed cell death, for the majority of a population of cells, in response to antibiotics prior to 2004[30, 45, 46] but only the HipAB system had been linked to persistence.

In 2004, Keren et al. investigated the gene expression profile of *E. coli* persister cells which revealed a range of persister related genes upregulated in response to stress including several TA systems (*dinJ, mazEF* and *relBE*).[47] One toxin, RelE was determined to induce the formation of persister cells.[47] This study was the first to link experimentally confirmed TA systems to persistence. Concurrent to this, microfluidic experiments[48] confirmed two types of persistence within *E. coli*: Type 1 persister cells, generated within the stationary phase, were typified by the HipA7 mutant (a mutant within the *hipA* locus) and Type 2 persister cells, typified by the HipQ mutant (a mutant within an uncharacterised locus within *E. coli*, whose genes are unknown[49, 50]) were continuously generated through growth, albeit at a slower rate than type 1 persister cells. Overexpression of the HipA7 mutant enabled the survival of a sub-population of bacteria in response to ampicillin. These cells were viable and able to resume growth upon removal of antibiotic and this elegant study was the first to display the generation of persister cells in real time. Shortly after both of these studies were reported, HipAB was designated as a TA system[51-54] and multiple TA systems encoded in *E. coli* have been correlated with the persister phenotype. [55]

Since these two studies, TA systems have been implicated in the generation of multidrug tolerant persister cells in extraintestinal pathogenic *E. coli* (ExPEC)[56], and *Salmonella typhimurium*.[57]. Within ExPEC, three TA systems promoted bacterial infections of the bladder and upregulated persister cell formation in a murine model. TA upregulation has also been linked to persister cell formation within biofilms observed in *Burkholderia cenocepacia*[58], *E. coli*[59], , *Staphylococcus aureus*[60, 61] and *Streptococcus pneumoniae*.[62]

Helaine et al. found that *Salmonella* persisters were induced by vacuolar acidification and nutritional deprivation, two stressors that caused the upregulation of fourteen TA systems and the persister phenotype after internalisation by macrophages. Some persister cells remained viable after phagocytosis by naïve macrophages, providing a localised reservoir to promote chronic infections.[57, 63]

Current knowledge therefore implicates TA systems in the generation of multidrug tolerant persister cells and bacterial pathogenesis.[64] More recently, however, the view that TA systems directly induce persister cell formation has become controversial[65, 66 67] and the role of persister cell formation needs to be reviewed following several key retractions of model studies in *E. coli*.[68, 69]

1.4. Toxin-antitoxin systems

TA systems are conserved through bacteria and archaea acting as response elements that react to stressors including acidic pH, amino acid/nutrient starvation and antibiotics.[70-72] Activation of toxins result in programmed cell death or stasis, depending on the nature of the toxin and the level of its expression.[51] Extensive research throughout the last twenty years have implicated TA systems in chronic infections (persister cells), plasmid maintenance (Programmed cell death) and biofilm formation.[51, 73, 74] The stringent[75] and SOS response[76, 77], have also been implicated in the activation and regulation of several TA systems.[72, 78-83] The actions of certain toxins have been shown to induce the SOS response[84-86], resulting in a feedback loop to further activate several other TA systems. Both the stringent and SOS response ensure that cells respond to nutritional depravation and DNA damage via hardwired stress response systems in a concerted manner to induce cell dormancy via the activation of multiple TA systems, each acting on an individual cellular target. This results in a rapid switch containing several layers of multiplicity to preserve the bacterial population under antibiotic/environmental stressors.

At the present time (2018), individual TA members are classified into six distinct systems (Figure 1.2) based on the nature of the toxin and antitoxin components as well as the mode of toxin neutralisation.

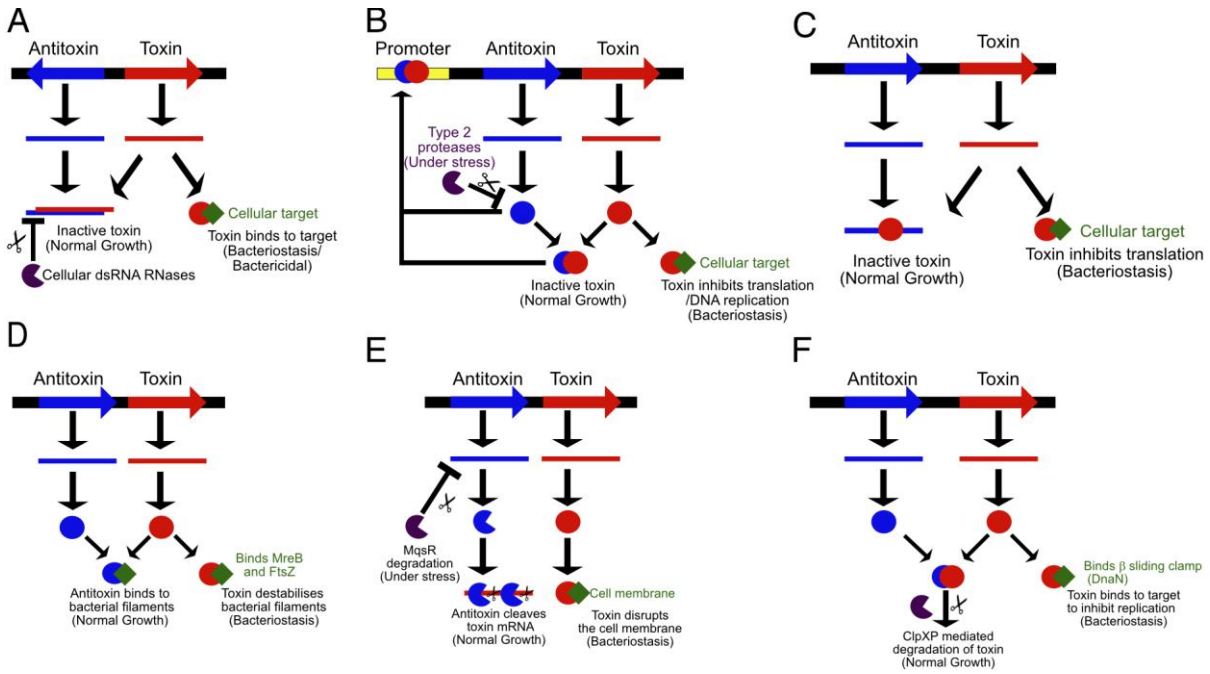


Figure 1.2. An overview of the six known TA systems and how each antitoxin (blue) interacts with its respective toxin (red). (A) Type 1 TA systems involve a direct interaction between the mRNA that encodes the toxin and antitoxin RNA to cause toxin neutralisation. (B) Type 2 systems are expressed from a bicistronic locus where a direct protein-protein interaction between the toxin and antitoxin is required for neutralisation. (C) Type 3 TA systems involves a protein toxin inhibited by an RNA antitoxin. (D) Type 4 systems involves the antitoxin shielding the cellular targets of the toxin by directly binding to the cellular target. (E) In the Type 5 system, the antitoxin (endonuclease) actively cleaves the toxin mRNA to prevent toxicity (F) The type 6 system SocAB binds the DnaN sliding clamp to inhibit DNA replication. Under normal cellular conditions, the toxin is degraded by SocA mediated ClpP degradation.

Type 1 systems (Figure 1.2A) are either arranged as overlapping gene pairs or divergent gene pairs. Each system consists of a protein toxin and a RNA antitoxin.[28, 72, 81, 87-96] The toxins are typically short hydrophobic proteins that insert into the cell membrane and induce depolarisation, in turn decreasing the proton motive force and the intracellular ATP concentration (Hok, TisB, IbsC and ShoB).[88, 97, 98] The small RNA (sRNA) antitoxins typically act *in trans* to inhibit toxin messenger RNA (mRNA) by forming dsRNA complexes.[94] These dsRNA complexes are subsequently cleaved by cellular RNases, preventing toxin translation and downstream toxicity.[94]

Type 2 TA systems have been extensively probed with various type 2 TA systems identified due to their organisation within bicistronic operons. In contrast to other systems, direct protein-protein interactions are formed between the toxin and cognate antitoxin to neutralise toxicity (Figure 1.2B)[99] and are covered in more detail below**.**

Type 3 systems (Figure 1.2C) utilise sRNA molecules as antitoxins to directly neutralise

protein toxins endoribonucleases resulting in plasmid addiction and phage protection[100-103]. This family therefore represents an intermediate class between type 1 and type 2 systems.[102, 104-106]

Type 4 systems consist of a bicistronic operon where the protein antitoxin directly interacts with the cellular target of the toxin to neutralise toxicity, rather than through the formation of a toxin-antitoxin complex (Figure 1.2D).[107, 108]

To date there is only one member of the type 5 system: GhoST (Figure 1.2E). The *ghoST* operon, found in *E. coli* encodes a membrane lytic protein that causes the formation of ghost cells, by permeabilising the membrane, similar to the type 1 toxin TisB.[109, 110] In comparison to other systems, the antitoxin GhoS acts as an endoribonuclease that cleaves ghoT mRNA- a novel mechanism with respect to current antitoxins.

Finally, the type 6 system SocAB from *Caulobacter crescentus* is similar to a type 2 system, consisting of a bicistronic operon and the formation of a protein-protein complex, SocAB. Unusually, the toxin SocB inhibits DNA replication by interfering with the β-sliding clamp (DnaN) (Figure 1.2F) and in contrast to type 2 systems the toxin, SocB rather than the antitoxin (SocA) is degraded preventing SocB mediated toxicity (cell death). SocA, the antitoxin is an adaptor for ClpXP through the formation of a SocAB-ClpXP complex which mediates continuous degradation of SocB.[111]

There are however, and likely to increasingly be, exceptions to these general classifications (Appendix Table 1 and 2).

1.5. Type 2 TA Systems

The structure and function relationships have been characterised for many type 2 TA systems and are summarised in Table 1.1. Type 2 toxins have been shown to have mRNase, kinase, acetylase and DNA gyrase inhibitory activities leading to mRNA cleavage, phosphorylation of translational machinery[99, 112, 113] (for example EF-Tu[114]), translational inhibition via the acetylation of charged tRNAs and inhibition of DNA replication[63, 115-117] respectively.

The type 2 toxins are neutralised by complexation with an antitoxin protein and the resultant complexes have been shown to negatively autoregulate their own transcription by binding to operator regions with their 5'untranslated region (5'UTR) preventing operon transcription by RNA polymerase. Antitoxins are labile and readily degraded by proteases such as Lon or Clp.[64, 118-120] Therefore, under normal growth conditions, the operon is continuously transcribed to ensure that a steady stream of antitoxin is available to neutralise its cognate toxin. During cellular stress such as nutritional starvation or antibiotic presence, the intracellular concentration of toxin rises to eventually exceed that of the less stable antitoxin as global translation within the cell is reduced. Insufficient antitoxin levels enable a population of free toxin to exert toxicity resulting in cell dormancy.

The high-resolution three-dimensional structures of at least one component of 18 Type 2 systems are known (Table 1.1). Figure 1.3-1.5 provides the details of the crystallised complexes of type 2 TA systems and individual components (toxin/antitoxin) for systems whose TA complexes are unknown. Within systems that share a specific function (i.e, DNA replication inhibition), there are structural variations for both the toxin and antitoxin, highlighting variability within the type 2 TA subclass.

TA system	Toxin	Function	Structure	Antitoxin	Structure	Reference
AtaAT	AtaT	Acetylates Met-tRNA^{fMet}	N/A	AtaA	N/A	115
Axe-Txe	Txe	mRNA endoribonuclease	N/A	Axe	N/A	121, 122
CcdAB	CcdB	DNA gyrase inhibitor	Dimer (X-ray crystallography)	CcdA	Monomer (NMR)	123-126
DinJ-YafQ	YafQ	Sequence specific mRNA endoribonuclease	Monomer (X-ray crystallography)	DinJ	Dimer (X-ray crystallography)	83, 127
Hha-TomB	Hha	Activates lytic prophage genes	Monomer (NMR)	TomB	Monomer (X-ray crystallography)	128-130
HicAB	HicA	mRNA endoribonuclease	Monomer (NMR)	HicB	Tetramer (X-ray crystallography)	131 132
HigBA	HigB	Translation dependent mRNA endoribonuclease	Monomer (X-ray crystallography)	HigA	Dimer (X-ray crystallography)	83, 133, 134
HipAB	HipA	Phosphorylates Glu-tRNA-ligase	Monomer (X-ray crystallography)	HipB	Dimer (X-ray crystallography)	135-138
FicTA	FicT	DNA gyrase and topoisomerase IV inhibitor	Monomer (X-ray crystallography)	FicA	Monomer (X-ray crystallography)	139-141
FitAB	FitB	Endonuclease	Monomer (X-ray crystallography)	FitA	Monomer (X-ray crystallography	142
GmvAT	GmvT	Acetylates tRNAs	Unknown	GmtA	N/A	143
KacAT	KacT	Acetylates tRNAs	Dimer (X-ray crystallography)	KacA	N/A	116
Kid-Kis	Kid	mRNA endonuclease	Dimer (X-ray crystallography)	Kis	N/A	144
MazEF	MazF	Sequence specific mRNA endoribonuclease	Dimer (X-ray crystallography)	MazE	Dimer (X-ray crystallography)	145, 146
MqsRA	MqsR	Sequence specific mRNA endoribonuclease	Monomer (X-ray crystallography)	MqsA	Dimer (X-ray crystallography/ NMR)	147-149
ParDE	ParE	DNA gyrase inhibitor	Dimer (X-ray crystallography)	ParD	Dimer (X-ray crystallography/ NMR)	150, 151
PezAT	PezT	Phosphorylates UDP-N-acetyl-D-glucosamine	Dimer (X-ray crystallography)	PezA	Dimer (X-ray crystallography)	152
Phd-Doc	Doc	Phosphorylates EF-Tu	Monomer (X-ray crystallography)	Phd	Dimer (X-ray crystallography)	133, 153
TacTA	TacT	Acetylates elongator tRNAs	Dimer (X-ray crystallography)	TacA	N/A	63, 154
RelBE	RelE	mRNA endoribonuclease	Monomer (X-ray crystallography)	RelB	Dimer (X-ray crystallography)	155-157
VapBC	VapC	tRNA endonuclease	Monomer (X-ray crystallography)	VapB	Dimer (X-ray crystallography)	158, 159
YefM-YoeB	YoeB	mRNA endoribonuclease	Monomer (X-ray crystallography)	YefM	Dimer (X-ray crystallography)	160
Zeta/ Epsilon	Zeta	Phosphorylates UDP-N-acetyl-D-glucosamine	Dimer (X-ray crystallography)	Epsilon	Dimer (X-ray crystallography)	161, 162

Table 1.1. A summary of characterised type 2 TA systems. The function of the toxin and the known structural information on the toxin and antitoxin components where present for each system are highlighted. Three-dimensional structures are highlighted in Figure 1.3-1.5.

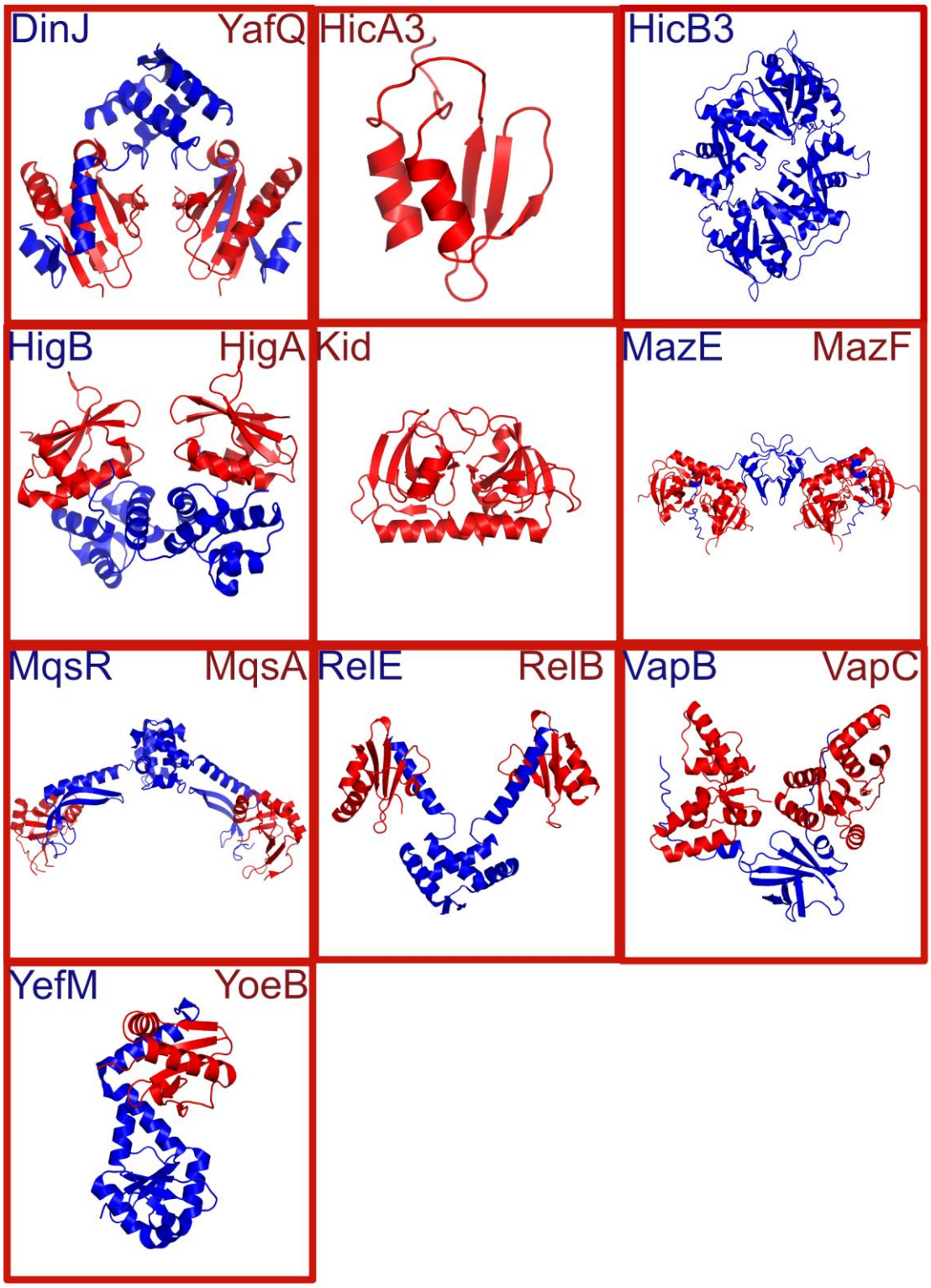


Figure 1.3. Three-dimensional structures of type 2 toxin-antitoxin systems. Structures of current TA systems whose toxin exhibits an endoribonuclease activity. Solved structures of the TA complexes, where available, are presented with the toxin and antitoxin highlighted in red and blue respectively. PDB codes: 4Q2U (DinJ/YafQ), 4C26 (HicA), 4P7D (HicB), 4MCT (HigBA), 1M1F (Kid), 1UB4 (MazEF), 3H12 and 3GN5 (MqsRA), 4FXE (RelBE), 3TND (VapBC), 2A6Q (YefM/YoeB).

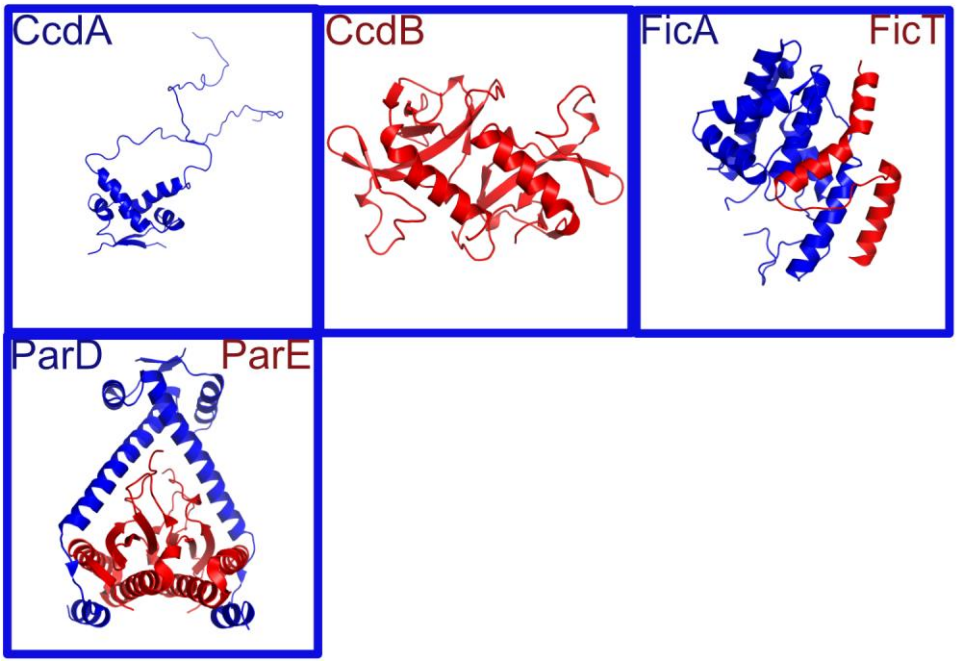


Figure 1.4. Three-dimensional structures of type 2 toxin-antitoxin systems. Structures of current TA systems whose toxin inhibits DNA replication (blue outline). Solved structures of the TA complexes, where available, are presented with the toxin and antitoxin highlighted in red and blue respectively. PDB codes 2ADL (CcdA), 1VUB (CcdB), 5JFZ (FicTA) and 3KXE (ParDE).

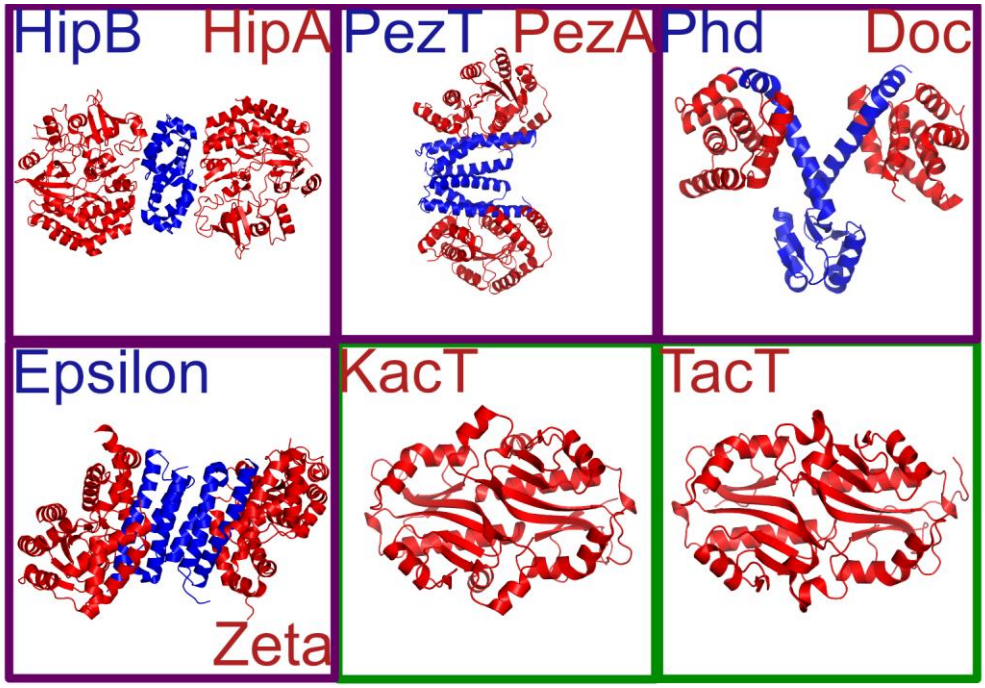


Figure 1.5. Three-dimensional structures of type 2 toxin-antitoxin systems. Structures of toxin-antitoxin complexes where the toxin is a kinase (purple outline) or an acetylase (green outline). Toxins and antitoxins are highlighted in red and blue respectively. PDB codes: 2WIU (HipAB), 2P5T (PezAT), 3KH2 (Phd/Doc), 1GVN (Zeta/Epsilon), 5XUN (KacT) and 5FVJ (TacT).

Endoribonucleases can be classified based on their mechanism of action. These are either dependent on the presence of ribosomes (translation dependent endoribonucleases): HigB, RelE, YafQ and YoeB and those that are not (translation independent endoribonucleass): HicA, Kid, MazF, MqsR, Txe and VapC.[113] Although each translation dependent endonuclease exerts cleavage within the ribosomal A site, the specificity of each varies. For example, HigB and YafQ recognise specific sequences respectively cleaving at AAA/ACA codons and AAA(G/A) sites.[163-166] In contrast, both RelE and YoeB display broad sequence specificity, with cleavage immediately seen at the 5' end of mRNA.[167-171]

The translation independent endonuclease (MazF and MqsR family members) recognise and cleave specific mRNA sequences[71, 172-176] and recently the *M. tuberculosis* MazF-mt9 alongside VapBC family members encoded in several bacteria have been shown to primarily cleave tRNA, although exceptions have been reported.[177-182] Kid specifically cleaves UA(A/C/U) triplets and inhibits translation as well as a ColE1 RNA II primer leading to inhibition of DNA replication. Kid therefore disrupts two cellular mechanisms to induce dormancy.[183-185] Txe exhibits a broad specificity and in discriminatively cleaves three nucleotides after the start AUG codon.[186] The specificity of HicA remains unclear and studies so far have observed mRNA cleavage but a recognition site has not been determined.[132, 187]

The broad specificity within the endoribonuclease toxins allows a concerted action in response to stress to inhibit translation at multiple levels resulting in dormancy. The multiplicity of type 2 systems within bacteria suggests that inhibition of one is insufficient to prevent cell dormancy due to the global network of endoribonucleases.[113]

The molecular mechanisms within toxins that act as inhibitors of DNA gyrase, kinases and acetylases also varies. Both CcdB and ParE bind the GyrA subunit of DNA gyrase to inhibit DNA replication[188-192] whereas FicT adenylates the GyrB subunit of DNA gyrase to inhibit DNA replication.[141] Within toxins that are kinases, Doc phosphorylates Ef-Tu to prevent the formation of the ternary (Ef-Tu-GTP-tRNA) complex.[114, 193] HipA inhibits Glu-tRNA-synthetase by phosphorylation to cause an accumulation of uncharged tRNAs and as a result both these toxins inhibit translation.[194] Other kinases, Zeta and its homolog PezT phosphorylate Uridine diphosphate-N-acetylglucosamine and inhibit cell wall synthesis resulting in cell lysis.[162] The molecular mechanism of acetylase toxins has only been uncovered in Ata[115] and TacT[63], both toxins acetylate the amino group of an amino acid

attached to charged tRNAs to prevent translation via inhibition of an initiator complex (Ata) and peptide bond formation (TacT).

1.6. Conditional Cooperativity

Transcription of prokaryotic genes is controlled by an upstream promoter region which enables binding of RNA polymerase in close proximity to a transcription start site (TSS, typically ATG/GTG).[195] Bacterial promoters often encode two consensus sequences typically 10 and 35 bp upstream of the TSS: the -10 Pribnow box (TATAAT)[196] and the -35 hexameric sequence (TTGACA).[197] The promoter region also contains the ribosome binding site (RBS), referred to as the Shine-Dalgarno sequence (consensus: AGGAGG[198]) typically between the -10 Pribnow box and the TSS.

Both the antitoxin and toxin-antitoxin complexes of type 2 TA systems are capable of binding to palindromic, imperfect-palindromic and inverted repeat sequences typically within the promoter region of their own bicistronic operon and are referred to as operator sites (Table 1.2). These often overlap with the -10 and -35 sequences or RBS to prevent transcription and therefore translation of the TA system A generalised model for the autoregulatory function of TA systems is as follows. The antitoxin alone typically binds weakly to the operator site, with enhanced binding upon complexation with a toxin, typically due to dimerization of adjacent DNA binding domains. The activity of these toxin-antitoxin complexes vary depending on the ratio of toxin:antitoxin and this mechanism has been termed conditional cooperativity (Figure 1.6).

Under basal cell conditions, the antitoxin is typically in excess compared to the toxin and weakly binds to the operator site.[199] This alone is not sufficient to result in full repression of the TA operon. Formation of a toxin-antitoxin complex results in the full repression of this operon, negating the expression of further toxin or antitoxin. Equimolar ratios of toxin:antitoxin are reached due to the proteolytic degradation of antitoxins by proteases.[51, 64, 120] Further proteolytic degradation leads to a higher toxin:antitoxin ratio, this reaches a critical point and further toxin modules bind to lower affinity sites on the antitoxin resulting in a higher order complex. In several systems the operator site cannot occupy two higher order complexes adjacent to each other, so these dissociate from DNA.[153, 155] The result of operon de-repression is transcriptional reactivation of the operon, resulting in expression of

the antitoxin and restoration of toxin:antitoxin back to basal levels. These models have been established using a variety of structural and biophysical techniques including X-ray crystallography, cryo electron microscopy and NMR.

Conditional cooperativity therefore represents a mode of bistable switching between two population states via transcriptional regulation and this mechanism has been proposed through mathematical modelling to result in bistability within bacteria to generate two populations of cells: normal growing antitbiotic susceptible and antibiotic tolerant persister cells.[200-202] This mechanism would enable cells to switch into a persistent state as the intracellular concentration of toxin increases with respect to the antitoxin while also ensuring the process is reversible, allowing cells to switch back into a normal growth phase when the toxin reaches an excess.

Conditionality and the mechanism of a bistable switch is not unique to TA systems and the mechanism of transcriptional regulation through protein-protein (toxin-antitoxin) interactions (Figure 1.6A) shares several characteristics to the bistable switches within the lactococcal phage TP901-1 (Figure 1.6B)[203, 204] and the σ/anti-σ systems (Figure 1.6C).[205, 206] σ/anti-σ systems are are bicistronic, where the anti-σ factor is synthesized in excess of the σ factor and is proteolytically degraded within the cell.[199, 206-208] Formation of a σ-anti-σ protein-protein complex sterically prevents formation of an intact RNA polymerase holo-enzyme and therefore blocks the transcriptional regulation mediated by σ factors.[206, 209]

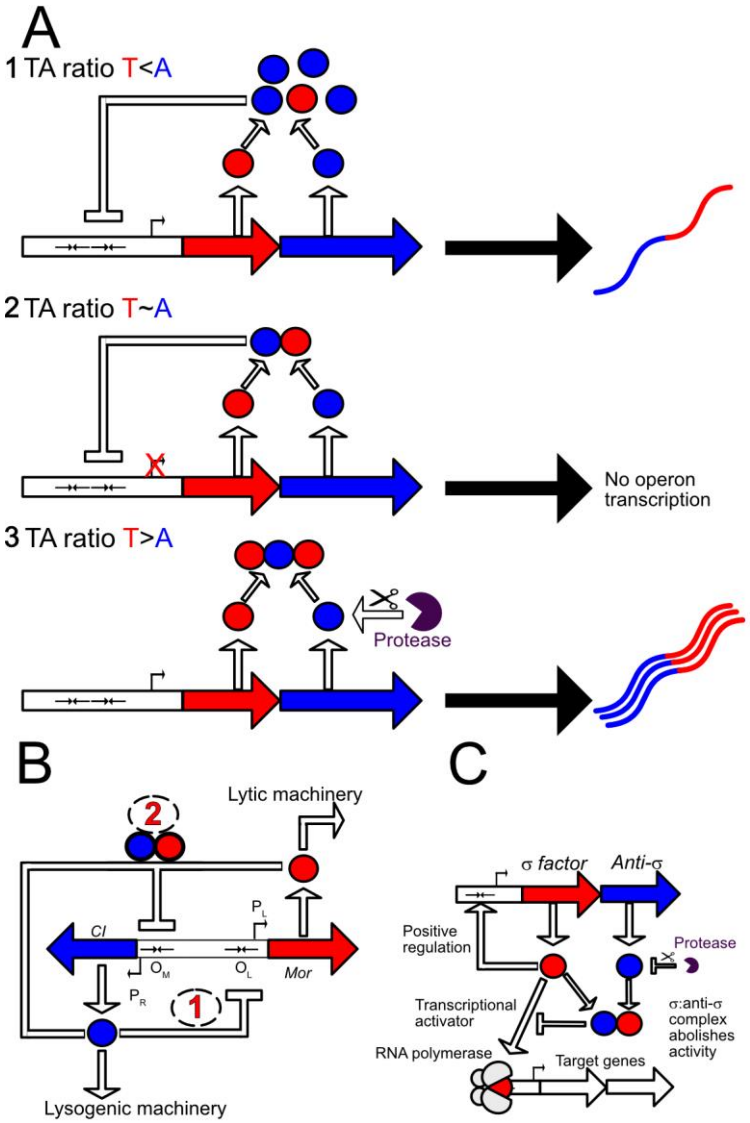


Figure 1.6. Commonly viewed model of conditional cooperativity and other bistable switches that regulate transcription via protein-protein interactions. (A) Conditional cooperativity in Toxin-Antitoxin systems. 1: When the ratio of toxin (red): antitoxin (blue) is low (antitoxin is in excess of the toxin), the operon is expressed as the antitoxin cannot achieve transcriptional repression. 2: Up to approximately equimolar concentrations of both toxin and antitoxin, formation of a TA complex results in tight binding to the promoter region and full repression of the operon. 3: At high toxin: antitoxin ratios, a result of the proteolytic degradation of antitoxins over time, a higher order protein complex is formed that is incapable of transcriptional autoregulation. (B) The TP901-1 phage system. Transcriptional regulation by Mor and CI proteins determines whether the temperate phage enters a lytic or lysogenic life-cycle. (1) The CI protein (blue) is capable of binding to 3 operator sequences upstream of the mor gene, termed the O_L site, preventing expression of mor from the P_L promoter. This ensures the phage enters the lysogenic cycle rather than the lytic cycle through the downstream actions of the CI protein.[203, 204] (2) Expression of both *cI* and *mor* genes enables the formation of a CI:Mor heterotetramer which is capable of binding the O_M operator causing transcriptional repression of the cI gene by blocking the P_R promoter. This enables the phage to enter the lytic cycle rather than the lysogenic cycle through the downstream actions of the Mor protein (red). (C) σ/anti-σ systems. These are similar to type 2 TA systems and are organised as bicistronic operators expressed from a single promoter. Each σ factor (red) is responsible for directly regulating a subset of genes by binding to the core subunits of RNA polymerase (grey) to initiate transcription. σ factors have been shown to regulate their own transcription via a positive feedback mechanism. The anti- σ factor (blue) specifically regulates its cognate σ factor by forming a tight protein-protein complex, preventing the transcription of downstream genes via RNA polymerase. Like type 2 antitoxins, anti- σ factors are regulated by proteases in order to switch on the transcription of specific genes.[206, 209]

While the precise nature of the DNA bound complex leading to conditional cooperativity often differs, the majority of TA systems adhere to this phenomenon. Exceptions include DinJ/YafQ and MqsRA (Table 1.2). Structural descriptions of co-repression and de-repression have been reported for Phd-Doc[210] and RelBE.[155] Structures have been hypothesized for both CcdAB[126] and VapBC[211] for full de-repressed states but these have not been experimentally validated.

TA system	Co-repression observation	Initial complex	De-repression observation	Higher order complex
Axe-Txe	EMSA [121]	N/A	N/A	N/A
CcdAB	EMSA [212]	$CcdA_2$-$CCdB_2$	EMSA [212]	$CcdB_2$-$CcdA_2$-$CcdB_2$ [126]
DinJ-YafQ	No [83, 213]	YafQ-$DinJ_2$-YafQ [83]	No	N/A
HicAB	No [214]	$HicA3_2HicB3_4$	EMSA [214]	N/A
HigBA	EMSA [215]	HigB-$HigA_2$-HigB [215, 216]	Upon disruption of HigA dimerization [134, 215]	Predicted to be HigBA heterodimer.
HipAB	No [217]	$HipB_2$-$HipA_2$-$HipB_2$ [137]	N/A	N/A
FitAB	Fluorescence polarization [218]	$FitA_4B_4$ [142]	N/A	N/A
Kid-Kis	EMSA [219]	$(Kid_2$-$Kis_2)_n$ oligomers	EMSA [183]	Kid_2-Kis_2-Kid_2-Kis_2, Kis_2-Kid_2-Kis_2-Kid_2-Kis_2 [219]
MazEF	ITC and EMSA [220, 34, 221]	$MazF_2$-$MazE_2$-$MazF_2$ [145]	EMSA [220, 222]	N/A
MqsRA	No [147, 148]	N/A	EMSA [223]	Predicted MqsRA heterotetramer
ParDE	EMSA [224, 225]	Assumed $ParD_2$-$ParE_2$	Yes [224]	N/A
PezAT (Zeta/Epsilon)	LacZ assays [152]	Assumed $PezA_2$-$PezT_2$	N/A	N/A
Phd-Doc	EMSA [226]	Phd_2-Doc-Phd_2	EMSA and ITC [210, 227]	Doc-Phd_2-Doc
RelBE	ITC [156]	$(RelB_2$-$RelE)_2$ [155, 228]	EMSA and modelling [155, 228]	$RelB_2$-$RelE_2$
TacTA	EMSA [154]	N/A	N/A	N/A
VapBC	SPR and EMSA [211, 229]	$VapB_4C_4$ [229]	EMSA [211]	$VapB_2$-$VapC_2$-$VapB_2$ [211, 230]
YefM-YoeB	EMSA [231]	Predicted $YefM_2$-YoeB	N/A	N/A

Table 1.2. Overview of Conditional cooperativity in TA systems. A list of TA systems where co-repression has been observed and the TA complex determined to be responsible for this. Where applicable the observation of de-repression and the higher order complex responsible is also noted. N/A: Not applicable

The majority of type 2 systems studied (9/13) bind to palindromic sites (Table 1.3). In contrast DinJ, HicA3B3 and RelBE recognise imperfect repeat sequences while FitAB/VapBC in different bacterial species have been shown to recognise inverted sequences/imperfect palindromic sequences that overlap the -10 site. Several reports have detailed high resolution structures of TA complexes and TA components bound to DNA. Each antitoxin binds DNA through defined DNA-binding motifs: AbrB superfamily[232], Helix-turn-helix (HTH)[233], Ribbon-helix-helix (RHH)[234] or a non-canonical motif (NCM).[160]

TA System	Host Organism	Binding motif	Operator sequence	TA-DNA structure	Reference
CcdAB	*E. coli*	RHH	5'-GTATAC-3' (1)	$CcdA_2$-DNA (A_2-B_2-A_2)-DNA	126, 212
DinJ-YafQ	*E. coli*	RHH	Site1: 5'-TGTTGX$_6$TGAGCTX$_3$-3' (2) Site2: 5'-GCTGAAX$_4$TATACAX$_2$-3'(2)	$DinJ_2$-DNA (YafQ-$DinJ_2$-YafQ)-DNA	83
FitAB	*Neisseria gonorhoeae*	RHH	Site1: 5'-TGCTATCA-3' (3) Site2: 5'-TGATAGCA-3' (3)	($FitA_4FitB_4$)-DNA	218, 142
HicA3B3	*Yersinia pestis*	RHH	Site1:5'-TGGGTATTACACCTA-3'(2) Site2: 5'-TAGGTTATGATACCCA-3' (2)	Predicted as $HicB3_4$-DNA	132
HipAB	*E. coli*	HTH	5'-TATCCX$_8$GGATA-3' (1)	($HipA_2HipB_2$)-DNA	43, 235
	Shewanella oneidensis	HTH	5'-GTGTAX$_6$TACAC-3' (1)	($HipA_2HipB_2$)-DNA	217
MazEF	*E. coli*	AbrB	5'-ATATA-3' (4)	$MazE_2$-DNA	220
MqsRA	*E. coli*	HTH	Site1: 5'-TAACCT-3' (4) Site2: 5'-AGGTTA-3' (4)	$MqsA_2$-DNA	148
ParDE	*E. .coli* plasmid RK2	RHH	5'CACATATGTG-3' (4)	Modelled as $ParD_2$-DNA	236, 150
Phd-Doc	*E. coli* phage P1	NCM	Site1: 5'-TGTGTACACA-3' (4) Site2 5'-CGAGTACACG-3' (1)	Predicted as $(PhD_2)_2$-DNA	226
RelBE	*E. coli*	RHH	Site1: 5'-TTGTAATGACAT-3' (2) Site2: 5'-TTGTAATTACAA-3' (2)	Predicted as $(RelB_2E)_2$-DNA	228, 155, 156
VapBC	*Shigella flexneri*	AbrB	Site1: 5'-CTATATGX$_4$CTGTATAG-3' (2) Site2: 5'-CATATAGX$_5$CTGTATA-3' (3)	Predicted $VapB_4C_4$ –DNA	159
VapBC	*Rickettsia felis*	AbrB	Site1: 5'-TAATATATACT-3' (1) Site2: 5'-TAATATATACT-3' (1)	$VapB_4C_4$-DNA	229
VapBC	*C. crescentus*	AbrB	Site1: 5'-CGTATATACG-3' (1) Site2: 5'ATATTGACGG-3' (1)	$VapB_4C_4$-DNA	230
YoeB-YefM	*E. coli*	NCM	5'-TGTACA-3' (4)	Predicted $YefM_2$-$YoeB_2$/YefM-$YoeB_2$	231
Zeta/ Epsilon	*Streptococcus pyrogenes*	HTH	N/A	N/A	161

Table 1.3. Summary of TA systems implicated in DNA binding and structures indicated where present. The host organism, DNA binding motif: AbrB superfamily (AbrB), Helix-turn-helix (HTH), Ribbon-helix-helix (RHH), or a non-canonical motif (NCM) are highlighted for each TA family. The operator sequence that the antitoxin binds to and the specific oligomeric state of a toxin-antitoxin-DNA structure are presented. X refers to any nucleotide. (1): imperfect palindrome, (2): imperfect inverted repeat, (3): perfect inverted repeat, (4): perfect palindrome.

Six type 2 antitoxins (CcdB, DinJ, FitA, HicB3, ParD and RelB) contain a RHH domain with structural/functional data known for all except HicB3. Each RHH inserts into the major groove of DNA via a β sheet and mediates protein-DNA interactions via polar/charged residues[124, 126, 150, 156, 224, 237] (Figure 1.7). Mutation of these residues reduces DNA binding of each antitoxin and autoregulation to their specific operon.[83, 124, 218, 228]

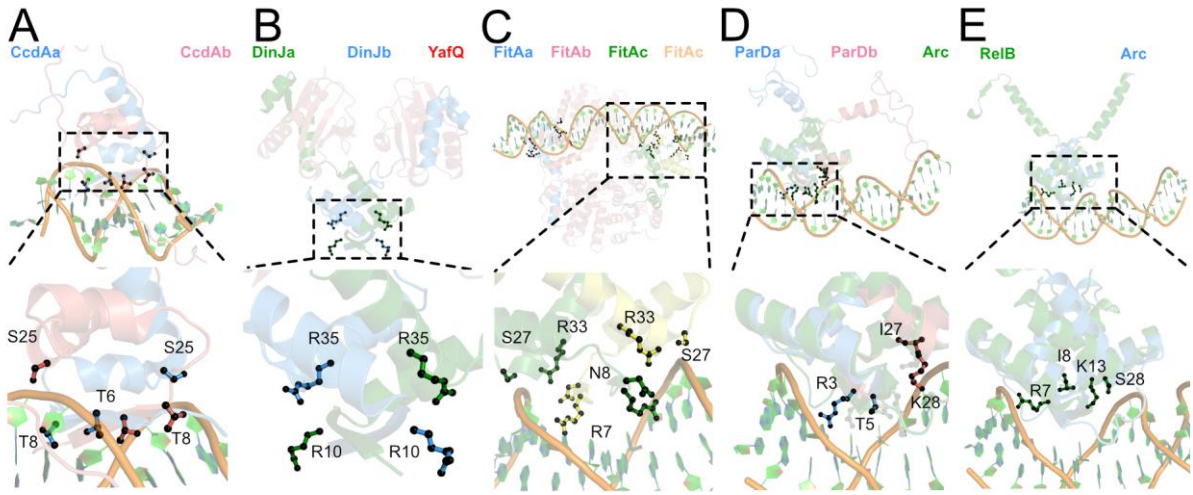


Figure 1.7. Antitoxins containing ribbon-helix-helix domains and their antitoxin-DNA structures. (A) Structure of CcdA-DNA (PDB: 2H3A), where amino acids interacting with DNA: T6, T8 and S25 are highlighted according to their specific subunit (red/blue). (B) Structure of DinJ-YafQ (PDB: 4Q2U), amino acids within DinJ implicated in DNA binding: R10 and R35 are highlighting according to their specific subunit (blue/green). (C) Structure of FitAB-DNA (PDB:2H10) where amino acids that contact DNA: R7, N8, S27, R33 from each FitB subunit (green/yellow) are highlighted. (D) Model of ParD-DNA interaction (PDB: 2AN7) based on the homologous Arc-DNA structure (blue) from *Bacillus subtilis* (PDB:1PAR). Amino acids R3, T5, I27 and K28 from adjacent subunits of ParD (Blue/Red) are predicted to bind DNA due to their equivalent positions in Arc. (E) Model of RelB-DNA interaction (PDB: 4FXE) based on the homologous Arc-DNA structure (blue) from bacteriophage P22 (PDB: 1BDT) where R7, I8, K13 and S28 of RelB (green) project into the major groove.

1.7. Burkholderia pseudomallei

One bacterium that is an excellent case study in relation to persistence mediated chronic infections is *Burkholderia pseudomallei*.[238, 239] This motile Gram negative saprophyte[240], exists in at least 3 sources (soil, water and zoonotic) and is the causative agent of melioidosis.[241-244] Melioidosis exists in both acute and chronic forms and has been extensively reviewed.[239, 245, 246] A recent worldwide analysis of melioidosis[247] indicated that South Asia, in particular Northeast (NE) Thailand and Northern Australia were deemed to be endemic (Figure 1.8). Sporadic cases have also been reported in Central[248], South America[249, 250] and Africa[251], where both the temperature and soil have been predicted to be suitable for the long term survival the organism.[247]

It is estimated that worldwide there are approximately 165,000 reported cases of melioidosis a year with an estimated 89,000 resulting in death.[247] Misdiagnosis is common as the initial symptoms (chronic pneumonia and skin lesions) are mimicked by tuberculosis.[252] Both pneumonia and pulmonary melioidosis have been widely observed for both acute and chronic melioidosis[253] with the development of systematic sepsis upon transition to chronic infection, often preceding death. Current specialised tests take more than seven days to confirm bacteraemic melioidosis and if the patient suffers from the chronic variant they often succumb to sepsis before confirmation. It is likely that the true numbers of cases/year are under-reported and the severity of melioidosis is higher than estimated.

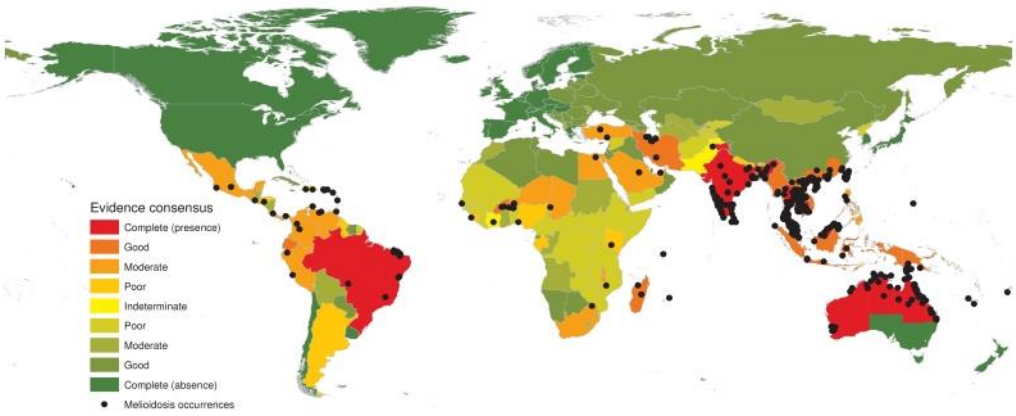


Figure 1.8. Endemic locations of melioidosis and suitable environments for establishment of *B. pseudomallei*. Geographical locations of melioidosis cases from 1910-2014 indicating a high prevalence in South Asia. Image directly taken from Figure 1 with the author's permission. [247] The evidence consensus for countries is color coded: red indicates the presence of melioidosis with complete consensus descending to dark green representing the absence of melioidosis with complete consensus.

Common routes of infection include the absorption of bacteria through skin abrasions and inhalation from contaminated sources. High risk groups within South Asia include rice paddy farmers, the most common occupation in NE Thailand. These farmers repeatedly work in flooded fields with no protective footwear and commonly show signs of repeated foot trauma. One study reported that 85% of patients with melioidosis were rice paddy farmers.[254]

Within endemic areas, antibodies against *B. psuedomallei* are widespread (60%) in children belonging to rice farming communities[255], hinting that the bacteria has adapted to survive the immune response. *B. psuedomallei* is inherently resistant to antibiotics[256, 257], and the current treatment for melioidosis is a course of ceftazidime/meropenem followed by trimethoprim-sulfamethoxazole for several months.[258]

Numerous case studies on melioidosis have occurred and a selection of the larger studies are summarised in Table 1.4 that highlight mortality rates and risk factors.[259-264] For each study sepsis shock was a good indicator of melioidosis and typically the cause of death. The 20 year study in Darwin, Northern Australia is to date the most comprehensive melioidosis study. This first suggested that *B. pseudomallei* should be viewed as an opportunistic pathogen and identified many health risk factors associated with melioidosis including diabetes, hazardous alcohol use, chronic lung disease and chronic renal disease, with a resultant 15-20% mortality rate.[259] Similar risk factors were seen in the other case studies with diabetes repeatedly highlighted as the highest risk factor. Rainfall was also correlated as an environmental risk factor for melioidosis[259-261, 264-266] and patients with known risk factors were twice as likely to develop bacteraemic pneumonia in seasons with heavy rainfall (>125 mm) resulting in a 2.5 fold increase in the melioidosis mortality rate.[267]

B. pseudomallei is therefore a potent pathogen that gives rise to a severe chronic infection (melioidosis). Experimental observations have recently reported its spread to the Caribbean, Central America and Mexico.[248] It is clear that the global prevalence of melioidosis is under-reported in several countries where *B. pseudomallei* is known to be endemic and further investigation on *B. pseudomallei* is crucial.

Variable	Australia[259]	India[261]	Malaysia[262]	Singapore[263]	Thailand[254]	Thailand[260]
No of cases	540	95	145	372	423	2,243
Incidence	19.6	N/A	16.4	1.7	4.4	12.7
Median age (year)	49	50	50	55	45	49
Male:Female ratio	2.2:1	2:1	3:1	4.5:1	1.4:1	1.4:1
Mortality rate (%)						
Overall	19	9.5	34	40	44	42.6
Bacteremic cases (%)	55	38.9	50	55	60	N/A
Pneumoniae [Deaths] (%)	16 [49]	32.6 [22.5]	42 [65]	N/A	N/A	N/A
Sepsis [Deaths] (%)	21 [50]	22 [40.9]	7.6 [42.9]	59.7 [55.4]	N/A	N/A
Risk factors (%)						
Diabetes mellitus	39	75.8	57	57	20	46
Heavy Alcohol use	39	13.6	N/A	N/A	12.3	N/A
Chronic lung disease	26	9.5	2.8	3.8	5.9	N/A
Chronic renal disease	12	21	2.8	5.6	20.1	N/A
Cases in the monsoon season	81	71.6	49.6	N/A	N/A	N/A

Table 1.4. Comparative epidemiology of melioidosis across 5 case studies. Risk factors are included where present (N/A) if not reported.

1.8. HicAB

The HicAB system was first identified in *Haemophilus influenzae* in 1998[268] and in several bacteria and archaea in 2006.[269] HicA was predicted to have a dsRNA binding fold, while HicB protein members contained a degraded nonactive RNase H fold, fused to a DNA binding domain (DBD). HicA was first characterised in *E. coli*[187] where it was described as a toxin with the ability to cleave both mRNA and tm-RNA, reducing the rate of translation within cells. Transcription of HicB prevented RNA cleavage, providing evidence that HicB was an antitoxin. Like other type 2 TA systems, the *hicAB* locus is regulated by stress, upregulation of this operon was observed during amino acid starvation and from degradation of HicB by Lon protease.[187] HicA from *B. pseudomallei* induced the formation of persister cells when overexpressed in *E. coli* and treated with ciprofloxacin and ceftazidime. Mutagenesis studies identified two key residues governing HicA toxicity: G22 and H24 that are located in the β1-β2 loop and are highly conserved (97%) amongst 75 HicA homologues. A H24A mutant of HicA was non-toxic, did not induce persister formation and could be recombinantly expressed in high yield in *E. coli* (10 mg/L). The three-dimensional structure was characterised by NMR exhibiting a dsRNA binding domain-like fold (dsRBD) αββββα, comprised of a triple stranded β sheet, two α helices and a net positive charge (Figure 1.9).[131] HicA-H24A was able to bind RNA but did not degrade it. The structure of the homologue HicA3, HicB3 and a truncated HicA3B3-NT complex from *Yersinia pestis* has also been crystallised.[132] HicA3 shared a conserved dsRNA binding domain (Figure 1.9A) and equivalent catalytic residues: G26 and H28, which were sequestered in a HicA3B3 complex (Figure 1.9B). Mutation of this histidine in HicA3 negated toxicity and HicA3 was deemed to act as an mRNase in vitro.

HicB3 was shown to autoregulate the *hicA3B3* operon by binding to operator regions of DNA to prevent its transcription, similar to other type 2 systems. HicB3 (Figure 1.9C) adopts a ring type tetramer with dimerization interfaces seen between adjacent subunits (blue, red, green, yellow) via their N-terminal and C-terminal domains. A crystal structure of a truncated complex: HicA3B3-NT indicated that HicA3 molecule bound the HicB3 N-terminal domain with a 1:1 stoichiometry (Figure 1.9D). The entire C-terminal domain was removed using subtilisin A prior to crystallisation. This complex is therefore incomplete with regards to the full length HicB and was not a sufficient model to explain complexation. The observed

stoichiometry (1:1) was extended to the full length HicB3 and a $HicA3_4$-$HicB3_4$ model was proposed.[132]

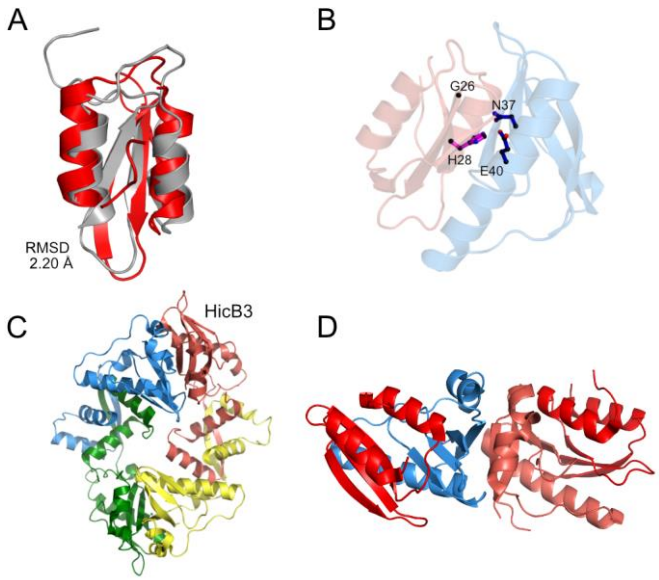


Figure 1.9. Structures of the HicAB family. (A) Structure of HicA (PDB: 4C26-grey) superimposed onto HicA3 (PDB:4P78-red), RMSD: 2.20 Å. (B) Sequestration of G26 and H28 in HicA3HicB3-NT model. (C) The crystal structure of HicB3 (PDB:4P7D) with individual subunits highlighted in blue, pink, green and yellow. (D) The $HicA3_2HicB3\text{-}NT_2$ dimer reported HicB subunits are highlighted in blue and pink and HicA in red.

The *E. coli hicAB* operon system was shown to be under the regulation of Sxy and cAMP receptor proteins.[270, 271] Turnball and Gerdes[214] provided an explaination of the regulation of this operon via two distinct promoters. The first promoter (P1) is under the control of both Sxy and cAMP receptor proteins and is located within the 5'UTR resulting in the transcription of both HicA and HicB. The second promoter (P2) is located in the operator region and is under the control of HicB alone with HicA shown not to act as a co-repressor by EMSA or LacZ assays. HicB binding to P2 to prevented transcription of the *hicAB* operon whereas an excess of HicA destabilised the HicB-promoter complex, thereby re-activating *hicAB* transcription.

Surprisingly, only HicB is transcribed from this promoter (P2), rather than both HicA and HicB. This dual promoter mechanism is an efficient strategy to ensure that levels of both HicA and HicB within the cell are controlled. An excess of HicA results in the upregulation of HicB, which in turn neutralises HicA (Figure 1.10). It is unknown how HicA destabilises the HicB-DNA complex.[214]

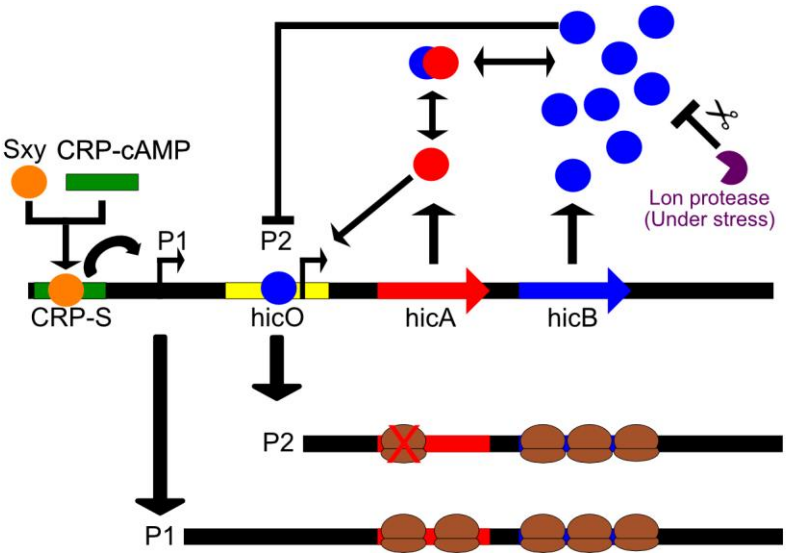


Figure 1.10. *hicAB* regulation within *E. coli*. HicB binds to the second promoter region (P2) to repress the *hicAB* operon. An excess of HicA destabilises this complex, causing transcription of HicB to restore the intracellular concentration of HicB to neutralise excess HicA. CRP, HicA, HicB, Ribosomes and Sxy are respectively highlighted in green, red, blue, brown and orange. Adapted from[214].

1.9. Aims

TA systems have long been implicated in the formation of persister cells, despite this there is a lack of structural and functional information on the HicAB system encoded within in many bacteria and archaea. One HicAB system is found in *B. pseudomallei* with prior work confirming the formation of persisters induced by HicA and the first structure of HicA.[131] Information on *B. pseudomallei* HicB is unknown and yet to be investigated, including structure, toxin sequestration and DNA binding properties. In this work the aim was to characterise the HicAB system from *B. pseudomallei*. This was achieved by:

- Characterising and solving the high-resolution structure of HicB using a variety of structural and biophysical techniques
- Solving the first complete high-resolution structure of a HicAB complex and determining whether multiple complexes could be formed and how this is achieved.
- Determining the factors governing DNA binding of HicB and HicAB and the identification of a possible target site for DNA binding in *B. pseudomallei.*
- Proposing a model for the regulation of the HicAB TA system and determining the precise mode of conditional cooperativity.

2. Structural studies of HicB

HicB is an antitoxin that neutralises the activity of the toxin HicA and can prevent the formation of persister cells.[131] Although structural studies of of HicAB TA systems have been reported, several of these have drawbacks. Deposited structures included the NMR structure of HicA from *B. pseudomallei* (PDB:4C26), a tetrameric crystal structure of the homologous antitoxin HicB3 from *Y. pestis* (PDB:4P7D) and an incomplete structure of the associated complex with the toxin HicA (HicA3B3, PDB:4P78, 28% sequence identity).[132] During the writing of this work a HicAB complex (PDB: 5YRZ) from *Streptococcus pneumoniae* was also reported but lacks the uncomplexed HicB for comparison.[272] There is no structural or functional data for *B. pseudomallei* HicB with respect to the mechanism of toxin sequestration and DNA recognition. More generally the molecular mechanism by which HicAB family members regulate their own operon is also unknown. Structural studies on *B. pseudomallei* HicB and extension to the complex of HicAB will shed light on the molecular mechanisms associated with this TA system that will be translatable to other members of the HicAB family encoded within major bacteria and archaeal clades.[187, 269] The work presented in this chapter details the recombinant expression, purification, crystallisation and subsequent structural analysis of HicB.

2.1. Expression and purification of full length HicB

A pET26b-HicB construct containing a N-terminal signal secretion sequence (PelB), His_6-tag and enterokinase cleavage sequence upstream of the HicB gene was provided by Dr Aaron Butt (University of Exeter) (Appendix Table 3). HicB was expressed in *E. coli* T7 cells. During expression the PelB sequence exported HicB to the periplasm via the general secretory (Sec) pathway.[273] During translocation, the PelB secretion sequence was cleaved by Signal Peptidase 1 at the AMA sequence.[274]

HicB was purified by immobilised metal affinity chromatography (IMAC) and size exclusion chromatography (SEC) before cleavage with enterokinase (Appendix Figure 1.A). Proteolytic cleavage of the N-terminal His_6-tag was unsuccessful and a maximum of 50% conversion of His_6-HicB to HicB was observed after systematic variation of reaction conditions (buffers, temperature, protein concentration and reaction time). It became clear that proteolytic cleavage of His_6-HicB occurred at several sites and multiple products were observed via SDS-PAGE (Appendix Figure 1B). In light of this, the tagged variant (His_6-HicB) was concentrated to 350 μM (where the tetrameric concentration [$HicB_T$] was 87.5 μM) and screened for crystallisation conditions against 384 commercially available Molecular Dimension Screens (Morpheus, Proplex, JCSG+ and Structure Screen I+II) using the sitting drop vapour diffusion technique. No positive hits were observed.

As N-terminal His_6-tags and their associated cleavage sites are predicted to be disordered in solution, their presence might be sub-optimal for crystallisation of the protein. To remedy this full length *hic*B was amplified from *B. pseudomallei* K96243 genomic DNA (HicB_FL Primer 1 and 2, Appendix Table 4) and subcloned into the pOPINE vector (pOPINE-HicB_FL).[275] The pOPINE vector encodes a C-terminal His_6-tag, but with no cleavage sequence present (Appendix Table 3). Although the tag is still likely to be disordered, the overall appended sequence was shorter which might have less impact on crystallisation but facilitate protein purification.

HicB_FL was expressed using *E. coli* T7 express cells (NEB) and purified by IMAC and SEC in a final yield of 7.5 mg/L (Figure 2.1A, B). SDS-PAGE showed the presence of monomeric HicB_FL (expected mass: 16,202 Da) Analytical SEC suggested that isolated HicB_FL existed as a single folded tetrameric species in solution (Figure 2.1C). This was further supported by native mass spectrometry which showed the presence of dimeric (blue) and tetrameric (red) charge states (Figure 2.1D). Deconvolution yielded a signal corresponding to the tetramer (Appendix Figure 2) with good mass agreement (observed: 64,800 Da, expected: 64,808 Da) given the rather poor signal to noise ratio.

Purified HicB_FL (500 μM, $HicB_T$ =125 μM) was screened against 384 commercially available Molecular Dimension screens (Morpheus, Proplex, JCSG+ and Structure Screen I+II) using the sitting drop vapour diffusion technique. Crystals were observed in 0.1M NaOAc pH 4.6, 2 M HCOONa after six days and no further hits were observed in any of the other assayed conditions. Crystals were subsequently optimised via hanging drop vapour

diffusion to yield diffraction quality crystals that could be individually looped, before cryoprotection in 20% (v/v) glycerol and data collection at the Diamond Light Source synchrotron. It was noted that the crystals were delicate and most cracked during looping or freezing. The best crystals of HicB_FL only diffracted to 7 Å and could not be used for structural determination. Further optimisation of crystallisation and cryoprotectant conditions (MPD, PEG 200, 400 etc) only improved the diffraction limit to 5 Å.

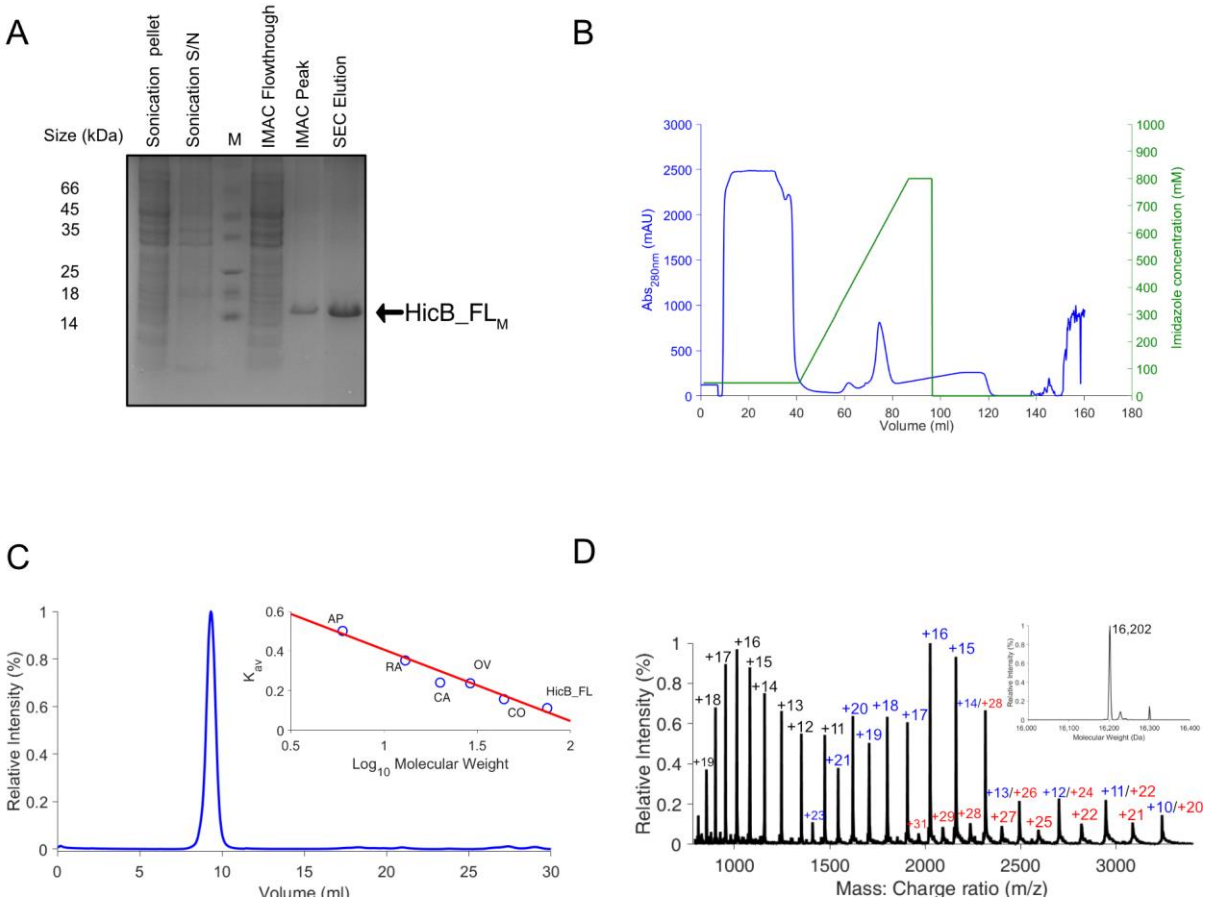


Figure 2.1. Initial characterisation of HicB_FL. (A) SDS-PAGE following the expression and purification of HicB_FL from post induction to isolation via IMAC and SEC. HicB_FL$_M$ refers to the monomeric state of HicB_FL. (B) IMAC of HicB_FL eluting as a single peak. (C) SEC of HicB_FL eluting as a single oligomeric species on an analytical 10/300 GL Superdex 75 column. (inset) calibration curve using known standards: aprotinin (AP, 6.5 kDa), ribonuclease A (RA, 13.7 kDa), carbonic anhydrase (CA, 29 kDa), ovalbumin (OV, 42.7 kDa) and conalbumin (CO, 76 kDa). (D) Native mass spectrometry of HicB_FL. Peaks spanning 750 to 3400 m/z were assigned based on their charge state: monomeric (black), dimeric (blue) or tetrameric (red). Inset: De-convoluted spectrum of the monomeric charge states within the m/z envelope shown to determine the monomeric mass of HicB_FL, observed: 16,201 Da, expected: 16,202 Da.

2.2. Expression and purification of HicB

Attempts to isolate and purify untagged HicB_FL from a variety of vectors (pET151, pOPINE and pOPINF) were unsuccessful: HicB bearing an N-terminal His_6-tag (encoded within pET151 and pOPINF vectors) was isolated by IMAC and SEC as previously described however, His_6-HicB and untagged-HicB could not be separated via IMAC following cleavage of the His_6-tag despite the use of different enzymatic cleavage sites within pOPINF and pET151 vectors (thrombin and TEV protease respectively). Untagged HicB (encoded in a pOPINE vector) could not be expressed to a high yield (1 mg/5 L) despite optimation of the expression protocol and this construct was deemed to be impractical for crystallisation experiments.The pOPINE HicB construct with the simple C-terminal His_6-tag (^{139}KHHHHHH^{145}) remained the most promising candidate for crystallisation optimisation. Bioinformatic analysis via I-Tasser [276] and Xtal Pred [277] predicted a series of disordered amino acids ^{135}VRHA^{138} at the C-terminal of HicB_FL immediately preceeding the His_6-tag. Therefore a construct lacking this stretch of amino acids was amplified from genomic DNA (Appendix Table 4, HicB Primer 1 and 2) and subsequently subcloned into pOPINE (pOPINE-HicB), to minimise flexibility at the C-terminus and improve crystal formation.

This truncated construct herein termed HicB was expressed as a soluble protein using T7 Express cells (NEB) (Figure 2.2) and this was purified by IMAC and SEC. SDS-PAGE revealed a single band corresponding to monomeric HicB and the mass confirmed by denatured electrospray mass spectrometry (ESMS) (observed, 15,738 Da: expected mass, 15,738 Da) (Appendix Figure 3). Removal of ^{135}VRHA^{138}, residues located at the C-terminus of HicB_FL, did not impact the formation of tetrameric HicB as judged by both SEC and native mass spectrometry; the latter displayed an equivalent profile of dimeric and tetrameric charge states observed for HicB_FL. Deconvolution of the tetrameric charge states yielded a single species (observed: 62,924 Da, expected: 62,952 Da). The final yield was 25 mg/ L, a significant improvement over HicB_FL.

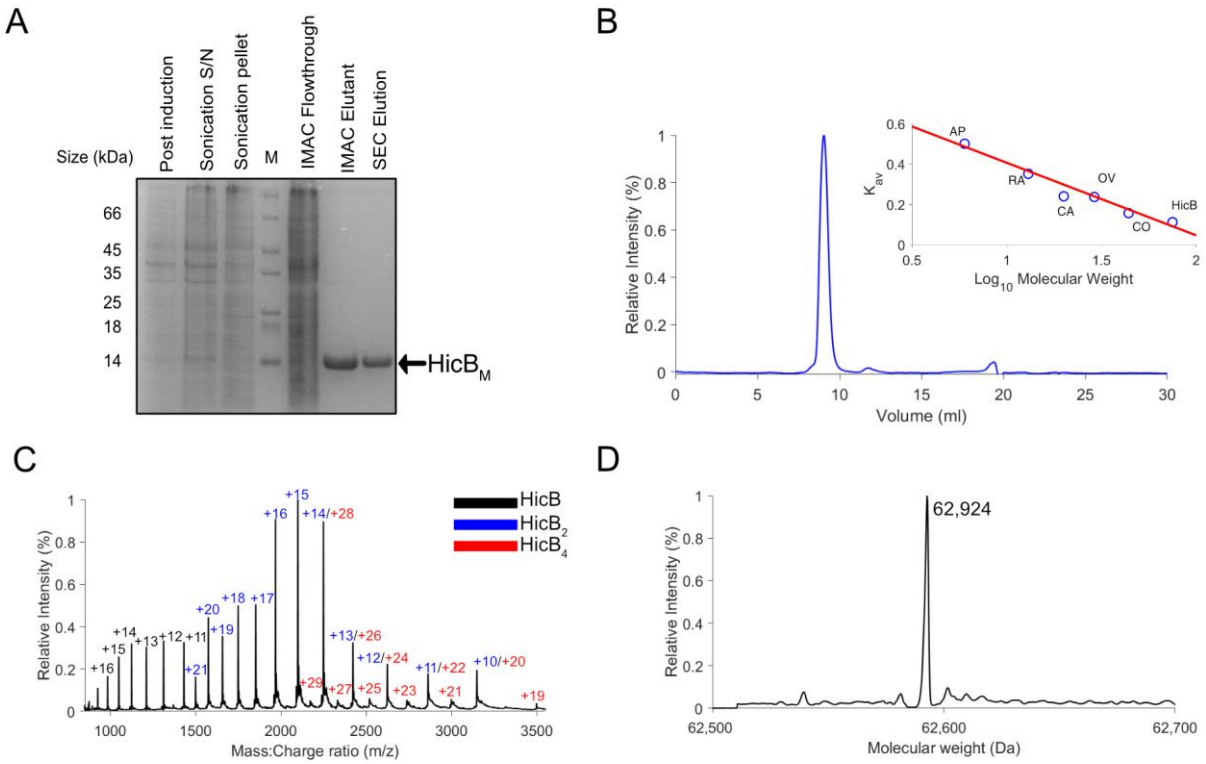


Figure 2.2. Purification and initial characterisation of HicB. (A) SDS-PAGE following the expression and purification of HicB from post induction to isolation via IMAC and SEC. $HicB_M$ refers to monomeric HicB. (B) SEC of HicB_FL eluting as a single oligomeric species on an analytical 10/300 GL Superdex 75 column. (inset) Calibration curve using known standards: aprotinin (AP), ribonuclease A (RA), carbonic anhydrase (CA), ovalbumin (OV) and conalbumin (CO). (C) Native mass spectrometry of HicB. Peaks were assigned based on their charge state: monomeric (black), dimeric (blue) or tetrameric (red). (D) De-convoluted spectrum of the tetrameric charge states shown in (C), yielding tetrameric HicB, observed: 62,924 Da, expected: 62,952 Da.

2.3. Crystallisation of HicB

HicB (600 μM, $HicB_T$ = 150 μM) crystallised in identical conditions to HicB_FL (0.1 M NaOAc, pH 4.6, 2 M HCOONa). Crystals were looped and cryoprotected in 40% glycerol (v/v) before flash freezing. Analysis of the diffraction quality of crystals at Diamond Light Source was hindered as a full 360 ° dataset could not be obtained before the crystal shattered which was attributed to poor cryoprotection and cracking during freezing. To reduce the damage caused during flash freezing, crystals were grown in a gradient of glycerol (0-6%) to minimise handling and exposure time during looping and freezing. Only crystals grown in the original crystallisation condition supplemented with 4% (v/v) glycerol were free of visible structural deformations and these were directly frozen for diffraction analysis.

A dataset of 2,000 images was collected at a wavelength of 0.9763 Å, exposure time: 0.02 s, oscillation: 0.15 ° and beam flux: 100% with cryogenic temperatures maintained through data acquisition (Table 2.1). Autoprocessing by the Diamond Pipeline (Fast DP) suggested a resolution limit of 2.4 Å, a substantial improvement compared to the HicB_FL crystals (best 5 Å). HicB (1,500 μM, $HicB_T$ = 375 μM) also formed crystals in a second set of conditions via hanging drop vapour diffusion: 0.1 M NaOAc, pH 4.6, 8% (w/v) PEG 4000, 10% (v/v) glycerol. It was not determined if HicB_FL crystallised in this condition. Crystals were cryoprotected in 40% glycerol (v/v) before being flash frozen and data collection occurred at Diamond Light Source synchrotron (Table 2.1). Autoprocessing (Fast DP) suggested a resolution of 1.95 Å, an improvement over the previously collected dataset. Both datasets were processed as described in the Materials and Methods.[278, 279] Numerous attempts to solve the phase problem and therefore the structure via molecular replacement (MR) in Phenix[280, 281] using *ab initio* models[276] and the homologue HicB3 (PDB:4P7D) to calculate initial phases and then refine the solution were unsuccessful for both datasets of HicB, despite the high resolution of the data. At this time the HicB3 structure was the only homologous crystal structure of HicB available (28% sequence identity) and it was not possible to determine if a single subunit of HicB retained the fold observed within HicB3. In order to solve the structure of HicB, a combination of selenomethionine (SeMet) labelling and truncation of HicB to an N-terminal domain (only residues 1-85) were employed in parallel to solve the phase problem and obtain suitable search models for molecular replacement.

	0.1M NAOAc, pH 4.6, 2 M HCOONa, 4% glycerol	**0.1 M NAOAc, pH 4.6 8% (w/v) PEG 4000, 10% (v/v) glycerol**
Cryoprotectant	N/A	40% (v/v) glycerol
Number of images	2,000	828
Wavelength (Å)	0.9763	0.9763
Exposure time (s)	0.02	0.057
Oscillation (°)	0.15	0.15
Beam flux (%)	100	20
FastDP resolution limit (Å)	2.4	1.95

Table 2.1. Table of parameters associated with data collection at Diamond Light Source for crystals of HicB obtained in different crystallisation conditions.

2.4. Selenomethionine labelling of HicB

Solving crystal structures by multiple isomorphic replacement (MIR) has been widely documented.[282-285] Incorporation of heavy atoms and more recently selenium via selenomethionine (SeMet) substitution into protein structures are both approaches undertaken to solve the phase problem of novel protein datasets when MR fails. For selenomethione substitution the general consensus is that the target protein should contain one methionine per fifty amino acid residues in order to maximise the success of solving the phase problem via this method.[286] As the native HicB sequence only contains one methionine residue located at the N-terminus, and an additional methionine residue incorporated by subcloning into the pOPINE vector (Position 0), additional methionine residues had to be introduced. To this end bioinformatics searches were undertaken to identify suitable candidate positions to mutate to a methionine residue. Previous work by McLachlan[287] had shown that either leucine or isoleucine were the most favourable amino acids to conservatively mutate to methionine whilst maintaining the native state of the protein.[288-290] To identify suitable mutation sites the sequence of HicB was aligned to *Y. pestis* HicB3 and ten homologous sequences from different organisms that contained >70% sequence identity to HicB (identified by Clustal Omega)[291] (Figure 2.3).

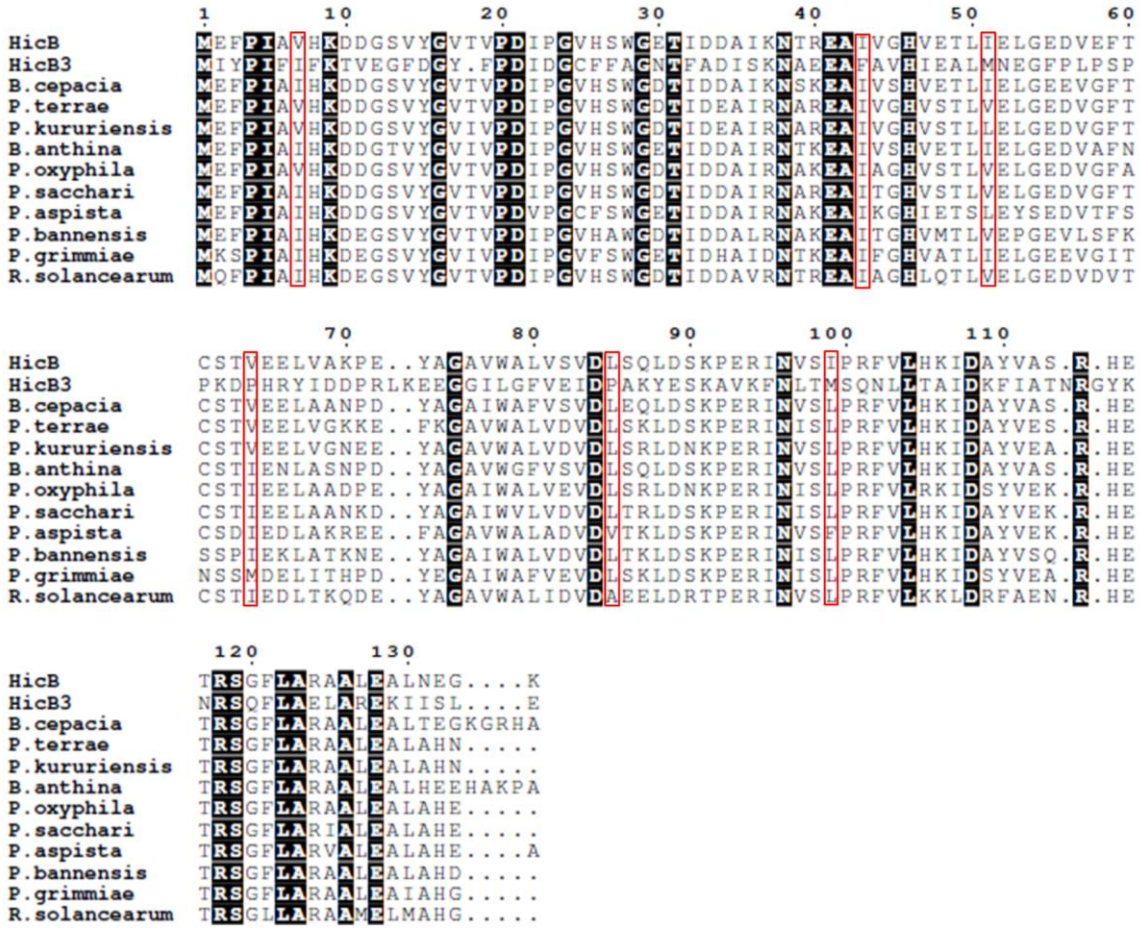


Figure 2.3. HicB (*B. pseudomallei)* aligned to HicB homologues from several species with a sequence identity >70%. Species aligned were**:** *Burkholderia cepacia* (88%), *Paraburkholderia terrae* (87%), *Paraburkolderia kururiensis* (86%), *Burkholderia anthina* (83%), *Paraburkholderia oxyphila* (82%), *Paraburkholderia sacchari* (82%), *Pandoraea aspista* (75%), *Paraburkholderia bannensis* (74%), *Paraburkholderia grimmiae* (74%) and *Ralstonia solancearum* (71%). The methionine residue incorporated by the pOPINE vector and the C-terminal His_6-tag were not included in the bioinformatics search. The figure was made using Espirit after alignments with Clustal Omega.[291, 292] Potential mutation sights to engineer in a methionine residue (Position 7, 43, 51, 64, 85 and 99) are highlighted in red.

Five initial amino acid residues (V7, I43, I51, V64, L85 and I99, red boxes in Figure 2.3) were determined to be potential candidates for mutation as these positions showed variation in the hydrophobic side-chains and therefore a conservative mutation might be tolerated. As two amino acids were required between residues 50 and 138, positions I51 and I99 were selected. Both of these residues were present as methionines in HicB3 (PDB: 4P7D) and were known to pack inside the hydrophobic core of the structure and therefore might be tolerated in HicB. Importantly they were also judged not to be involved in dimerization or HicA binding.

Three mutants: I51M, I99M and an I51M/I99M double mutant (pOPINE_DM) were constructed by whole plasmid PCR mutagenesis using mutagenic primers (Materials and Methods, Appendix Table 4). Both I51M and I99M mutants were individually expressed in an identical manner to HicB. Each mutant remained soluble and retained a tetrameric organisation (Figure 2.4). Denatured ESMS confirmed the mutation of isoleucine to methionine (+18 Da) within individual subunits of each mutant, when compared to HicB (Figure 2.4). Native mass spectrometry reaffirmed the incorporation of four additional methionine residues within the tetramer (Appendix Figure 4).

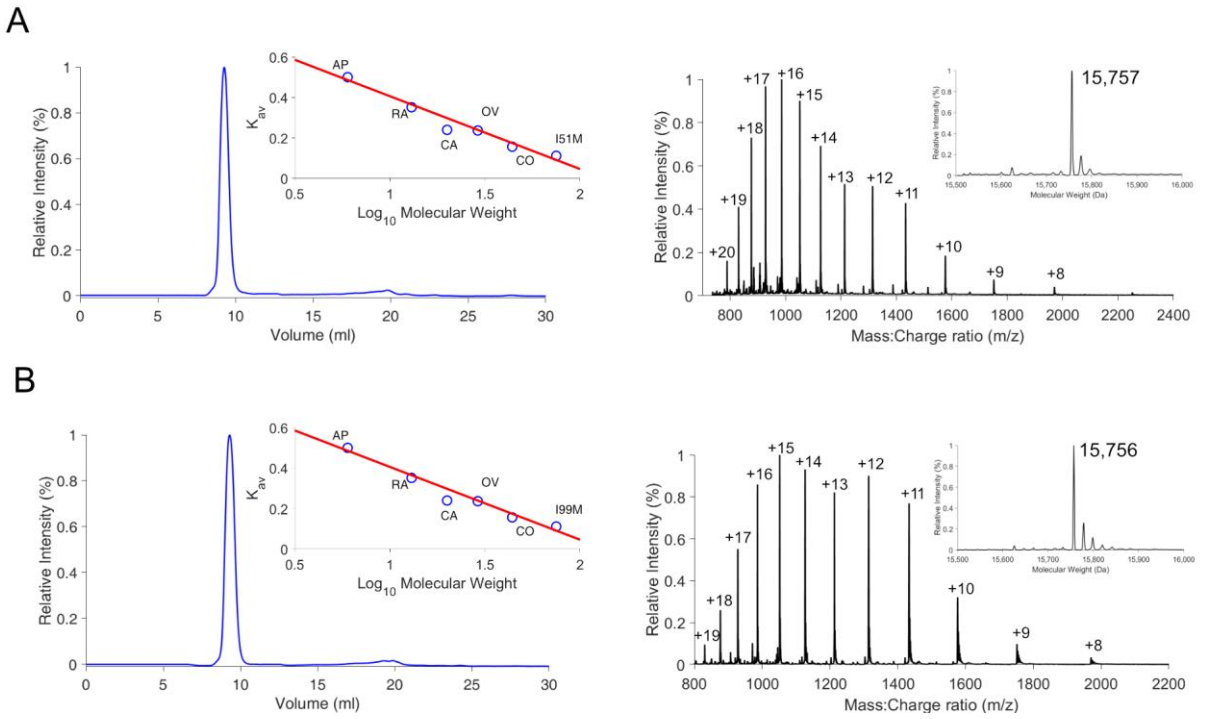


Figure 2.4. Purification of methionine mutants I51M and I99M via SEC and ESMS. (A) The I51M SEC profile eluted as a single species similar to the wild type. Denatured ESMS showed a similar distribution of m/z charge states to the wild type, with deconvolution (inset) resulting in a single species (observed mass: 15,757 Da, expected: 15,757 Da). (B) The I99M SEC profile eluted as a single species. Denatured ESMS showed a similar distribution of m/z charge states to the wild type, with deconvolution (inset) resulting in a single species (observed mass: 15,756 Da, expected: 15,757 Da). The expected mass for HicB is 15,738 Da. For both SEC traces, the inset contains the calibration curve using known standards: aprotinin (AP), ribonuclease A (RA), carbonic Anhydrase (CA), ovalbumin (OV) and conalbumin (CO).

The double mutant containing I51M/I99M (HicB_DM) was expressed in T7 Express cells, as a soluble protein with a final yield of 20 mg/L (Figure 2.5). Denatured ESMS confirmed the presence of a single species and the mutation of both isoleucine residues to methionine (Figure 2.5C, D, observed: 15,776 Da, expected: 15,774 Da). Furthermore native mass spectrometry showed a tetramer with a mass increase of + 147 Da (expected + 144 Da

increase), corresponding to the introduction of eight methionine residues when compared to the wild type (Appendix Figure 5) and confirmed that the mutations did not disrupt tetramer formation. This construct was therefore judged to be suitable for SeMet labelling.

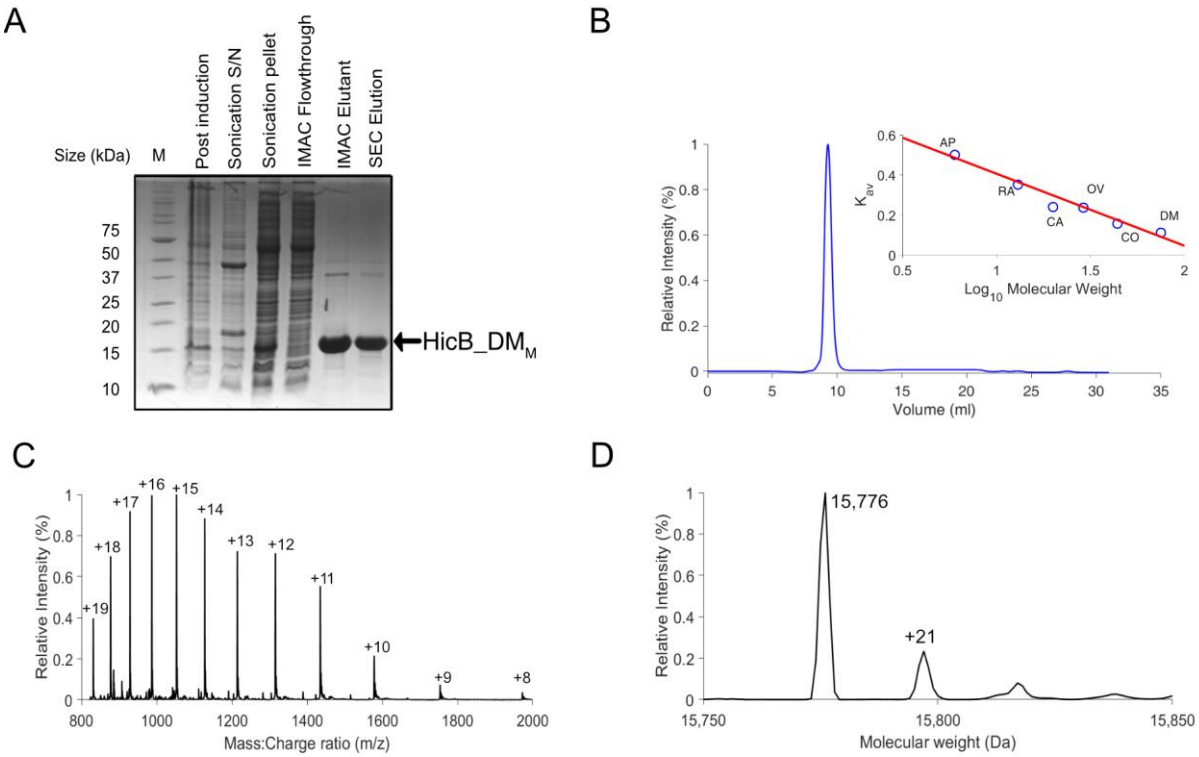


Figure 2.5. Initial characterisation of HicB_DM. (A) SDS-PAGE following the expression and purification of HicB_DM from post induction to isolation via IMAC and SEC. HicB_DM$_M$ refers to the monomeric state of HicB_DM. (B) SEC of HicB_DM eluting as a single oligomeric species, similar to the wild type protein, (inset) calibration curve using known standards: aprotinin (AP), ribonuclease A (RA), carbonic Anhydrase (CA), ovalbumin (OV) and conalbumin (CO). (C) Denatured mass/charge spectrum of HicB_DM. Peaks were assigned based on the monomeric charge state (black). (D) De-convolution of the m/z envelope, yielding monomeric HicB_DM (observed mass: 15,776 Da, expected mass: 15,774 Da). The expected mass for HicB is 15,738 Da.

pOPINE_DM was expressed in B834 (DE3) cells (NEB) and grown in the presence of SeMet (Materials and Methods) before purification.[293, 294] 2 mM TCEP was added to buffers throughout purification to prevent oxidation of the selenium atom.[295] Oxidation has been linked to radiation damage of the selenomethionine side chains during data collection that may hinder the phasing power and subsequent structural determination. Addition of selenium did not severely affect the expression or purification of HicB_DM, which was isolated with >90% purity with a final yield of 12 mg/L as judged via SDS-PAGE and SEC (Figure 2.6). SEC confirmed the purified protein formed a tetramer in solution.

Denatured ESMS confirmed the successful labelling of methionine residues with SeMet (Figure 2.6). Deconvolution of the resultant m/z profile of labelled HicB_DM resulted in an observed monomeric mass of 15,961 Da (Expected mass: 15,961.6 Da), an increase of 187 Da when compared to the expected monomeric mass of the double mutant: 15,774 Da. Incorporation of four SeMet atoms would result in an expected molecular weight increase of + 187.6 Da, indicating 100% successful incorporation SeMet at each methionine residue within one subunit at positions 0, 1, 51 and 99. Analytical SEC further confirmed that both HicB_DM and the SeMet labelled variant retained the same elution profile compared to the wild type (Appendix Figure 6).

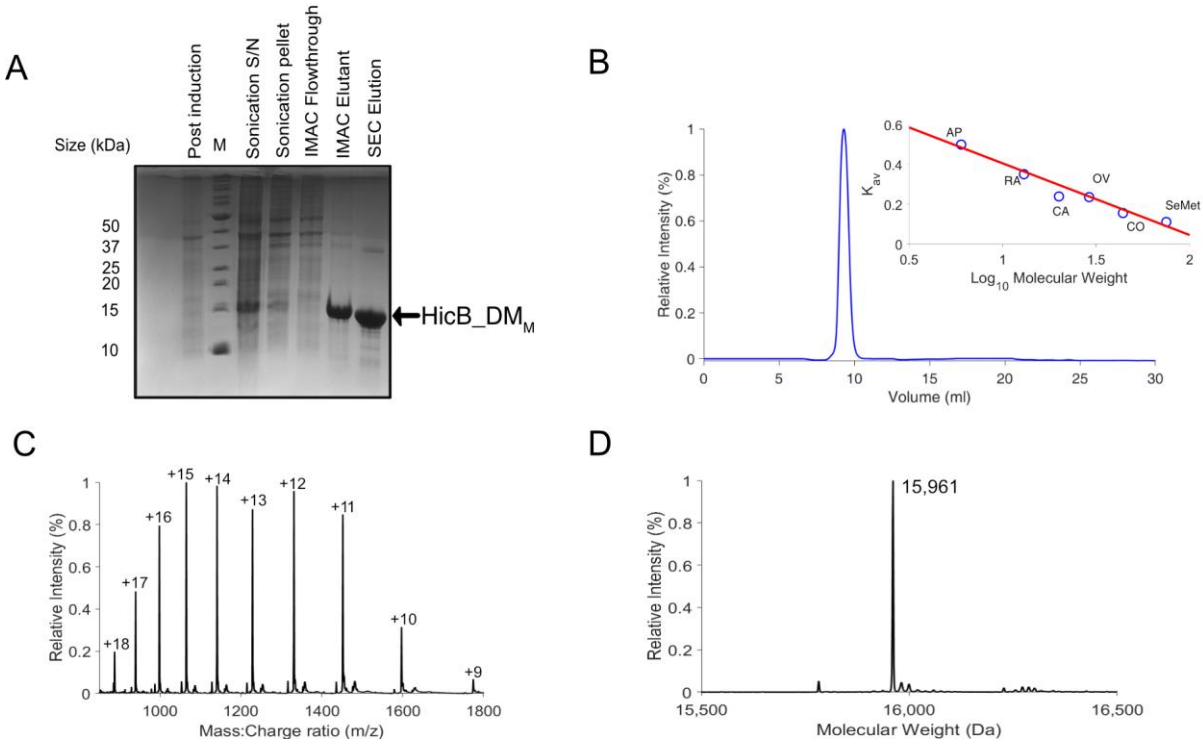
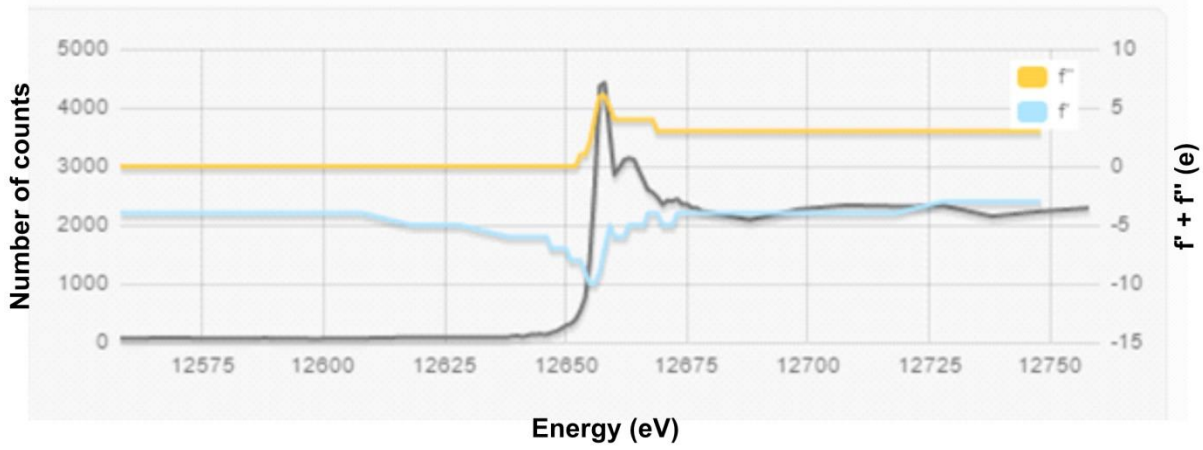


Figure 2.6: Initial characterisation of SeMet labelled HicB_DM. (A) SDS-PAGE following the expression and purification of SeMet labelled HicB_DM from post induction to isolation via IMAC and SEC. HicB_DM$_M$ refers to the monomeric state of HicB_DM. (B) SEC of SeMet labelled HicB_DM eluting as a single oligomeric species. inset calibration curve using known standards: Aprotinin (AP), Ribonuclease A (RA), Carbonic Anhydrase (CA), Ovalbumin (OV) and Conalbumin (CO). (C) Denatured ESMS of SeMet labelled HicB_DM. Peaks were assigned based on monomeric charge state (black). (D) De-convoluted spectrum of the envelope shown in (C), yielding monomeric HicB_DM labelled with SeMet. (observed mass: 15,961 Da, expected mass: 15,961.6 Da)

2.4.1. Crystallisation of SeMet HicB_DM and diffraction analysis

Crystals of SeMet labelled HicB (600 μM, $HicB_T$ = 150 μM) formed in 0.1 M NaOAc, pH 4.6, 2 M HCOONa, 11% (v/v) glycerol and were flash frozen and analysed at Diamond Light Source. To determine the appropriate wavelength for a single wavelength anomalous dispersion (SAD) experiment, fluorescence scans centred at the selenium K-edge were undertaken to determine the anomalous scattering factor components f' and f''. Results of this are summarised in Figure 2.7. A dataset of 1,800 images was subsequently collected at a wavelength corresponding to the fluorescence peak (0.9790 Å), exposure time: 0.1 s, oscillation: 0.2 °, beam flux: 100% with cryogenic temperatures maintained during acquisition. Auto-processing by Fast DP suggested the resolution of the dataset was ≈ 3 Å.



	Energy (eV)	f'(e)	f''(e)
Peak	12,658	6.76	-7.13
Inflection	12,655	4.14	-10.86

Figure 2.7. Fluorescence scan of HicB_DM-SeMet crystals. The point of inflection and absorbance are shown by the black line. X-axis: X-ray energy (eV), Y-axis: Anomalous scattering factors. Fluorescence scan parameters (The anomalous scattering signal of the crystal: f' and f'') are presented.

Structural determination was carried out with assistance from Dr Chris Williams (University of Bristol) using the Crank2[296] pipeline in the CCP4 jsCoFE interface.[297] Briefly, after collection, data were processed into $P2_12_12_1$[278, 279] before input into the Crank2 pipeline. Here, a freeR set was generated (SF Tools) and the Se structure factor |FA| was estimated using SHELXC.[298] 16 heavy atom sites were subsequently identified after 5000 trials in SHELXD[298], before phasing[298], hand determination[299] and density modification via Parrot[300] before refinement.[301] Autobuild runs (Buccaneer)[302] yielded a partial monomeric model which refined[301] to a R/R_{free} value of 0.327/0.390 (Table 2.2). Manual rebuilding and

refinement in Phenix and Coot[303] was unable, however, to reduce the R/R_{free} value, or build in multiple subunits of HicB within the asymmetric unit (ASU).

2.5. The truncated N-terminal domain construct of HicB

Investigation of the structural arrangement of the homologue HicB3 (PDB:4P7D) revealed two domains (Figure 2.8A). The N-terminal domain adopts a β1β2β3α1α2α3β4 fold (residues 1-90) that is connected to a RHH motif (residues 93-134) via a disordered hinge region. Previous attempts to crystallise the *Y. pestis* HicA3B3 complex were only successful after the addition of the protease subtilisin A immediately prior to crystallisation, resulting in a partial complex consisting of the N-terminal of HicB3 complexed to HicA3 ($HicA3_2$-HicB3-NT_2) (PDB:4P78).[132] Proteolytic cleavage occurred within the hinge region resulting in the loss of the C-terminal domain. Crystallisation of this resulting partial complex demonstrated however that the N-terminal domain of HicB3 formed a discrete, folded unit amenable to crystallisation.

The secondary structure of HicB was predicted by PSIPRED[304] to be composed of an N-terminal domain (residues 1-85) and a C-terminal domain (residues 94-134) that displays a RHH motif linked via a long loop (residues 86-93) (Figure 2.8B). These secondary structure elements mirrored those observed in *Y. pestis* HicB3 (β1β2β3α1α2α3β4α4β5α5α6) and by inference would be likely to have a similar N-terminal domain, which was amenable to crystallisation.

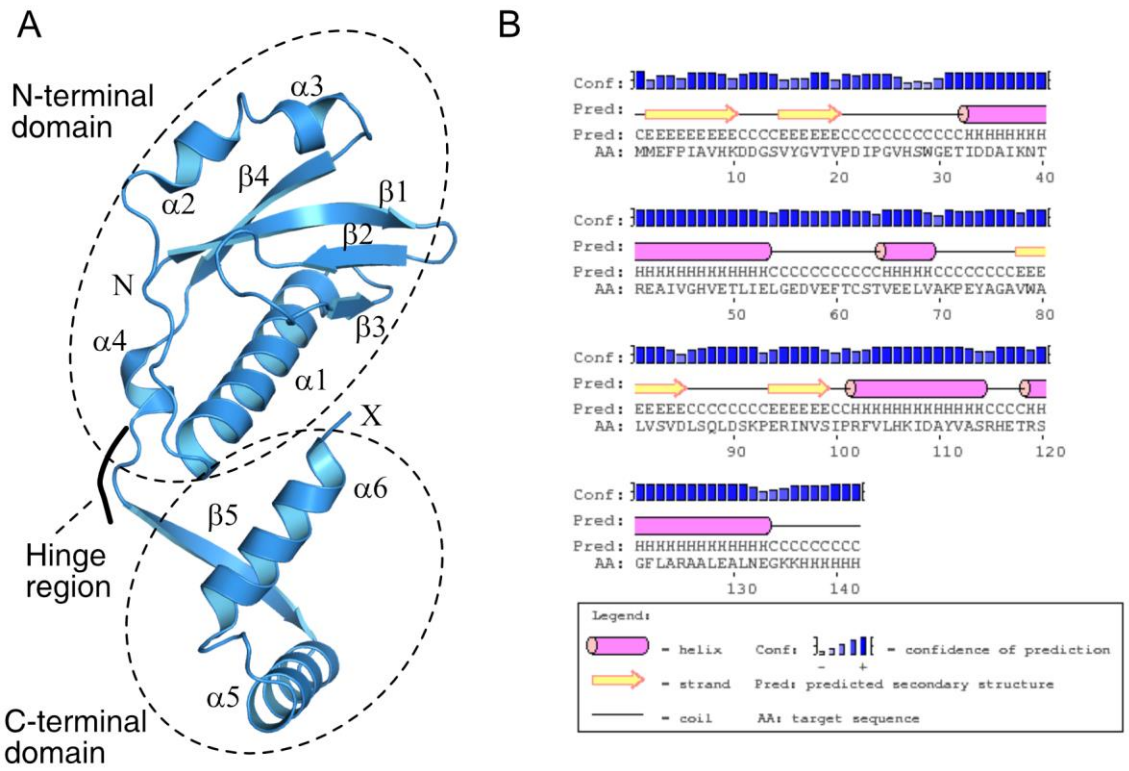


Figure 2.8. Domain identification of HicB. (A) Structural investigation of a HicB3 subunit indicating a compact N-terminal domain and a C-terminal domain bearing a RHH motif. (B) PSIPRED result for HicB. Results indicate that the C-terminal domain covers residues E93-E132, with a disordered loop region separating the N-terminal and C-terminal domain (L85-P92). Based on this result, HicB was truncated back to L85 to remove both the C-terminal domain and the loop region to isolate the ordered N-terminal domain of HicB. X refers to the location of the His_6-tag.

An N-terminal construct consisting of HicB residues 1-85 was cloned from *B. pseudomallei* K96243 genomic DNA (HicB-NT Primer 1 and 2, Appendix table 4) and sub-cloned into the pOPINE vector (pOPINE-HicB-NT). HicB-NT was expressed in T7 Express cells and purified as per HicB in a final yield of 15 mg/L (Figure 2.9A). Initial purification via IMAC (Figure 2.9A), yielded a major band by SDS-PAGE at ~10 kDa, corresponding to monomeric HicB-NT (calculated mass 10,133 Da) but this co-purified with several higher molecular mass impurities. HicB-NT was further purified via SEC, where it eluted as a dimeric species. Denatured ESMS gave an observed mass of 10,133 Da (Appendix figure 7) and native mass spectrometry yielded charge states originating from a mixture of monomeric and dimeric species (Figure 2.9C). Deconvolution of these m/z envelopes yielded an observed dimeric mass of 20,261 Da (expected mass: 20,266 Da).

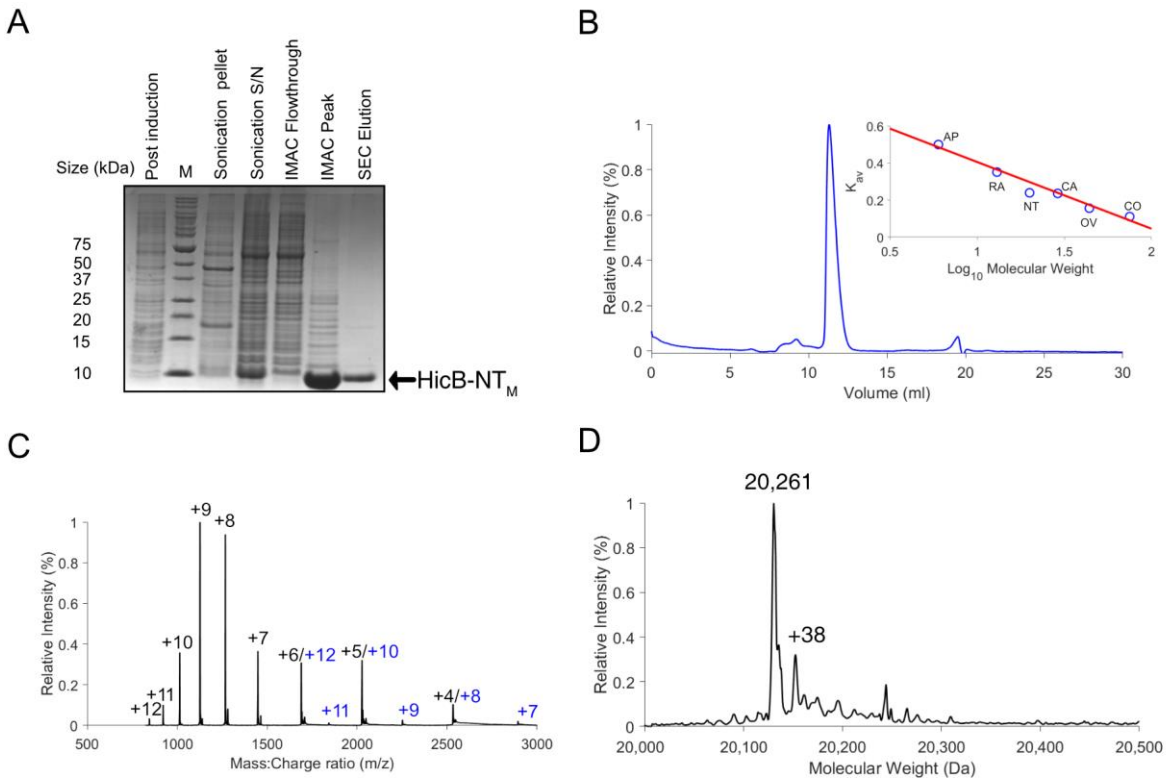


Figure 2.9. Initial characterisation of HicB-NT (A) SDS-PAGE following the expression and purification of HicB-NT from post induction to isolation via IMAC and SEC. HicB-NT$_M$ refers to the mononomeric state of HicB-NT. (B) SEC of HicB-NT eluting as a single oligomeric species on a Superdex 75 10/300 GL column (inset) The calibration curve using known standards: aprotinin (AP, 6.5 kDa), ribonuclease A (RA, 13.7 kDa), carbonic anhydrase (CA, 29 kDa), ovalbumin (OV, 42.7 kDa) and conalbumin (CO, 76 kDa). (C) Native mass spectrometry of HicB-NT. Peaks were assigned based on monomeric charge state (black) or dimeric (blue). (D) De-convoluted spectrum of the envelope shown in (C), yielding dimeric HicB-NT (observed mass: 20,261 Da, expected: 20,266 Da).

2.5.1. Crystallisation of HicB-NT

Purified HicB-NT (1,000 µM, where the concentration of dimer in solution [HicB-NT$_D$] was 500 µM) was screened through a total of 384 conditions (Morpheus, JSCG +, Structure Screen I and II and PACT Premier) as previously described. Successful initial crystallisation hits were observed in 0.1 M NaOAc, pH 4.6, 8% (w/v) PEG 4000 using hanging drop vapour diffusion supplemented with 10% (w/v) glycerol as a cryoprotectant. Crystals were subsequently flash frozen and data collected at the Diamond Light Source synchrotron. A dataset of 1,800 images were collected at a wavelength of 0.9795 Å, exposure: 0.1 s, oscillation: 0.2 °, beam flux: 100% with cryogenic temperatures maintained during acquisition. Autoprocessing pipelines at Diamond Light Source indicated that the crystals diffracted to 1.56 Å.

Structure determination was performed with assistance from Dr Chris Williams (University of Bristol) (an iterative process). The diffraction data were processed into $P2_12_12_1$[278, 279],

before MR was undertaken using a partial monomeric search model from the selenomethionine dataset (Figure 2.10) as a search model within Phaser-MR (Phenix)[280, 281] and the Autobuild pipeline.[305] Iterative rounds of model refinement[306] and manual model building were carried out using Phenix and Coot[303] to build 4 monomers within the ASU. The final structure was validated with MolProbity[307] and the PDB_REDO[308] web server. Statistics are given in Table 2.2 (Located at the end of this chapter).

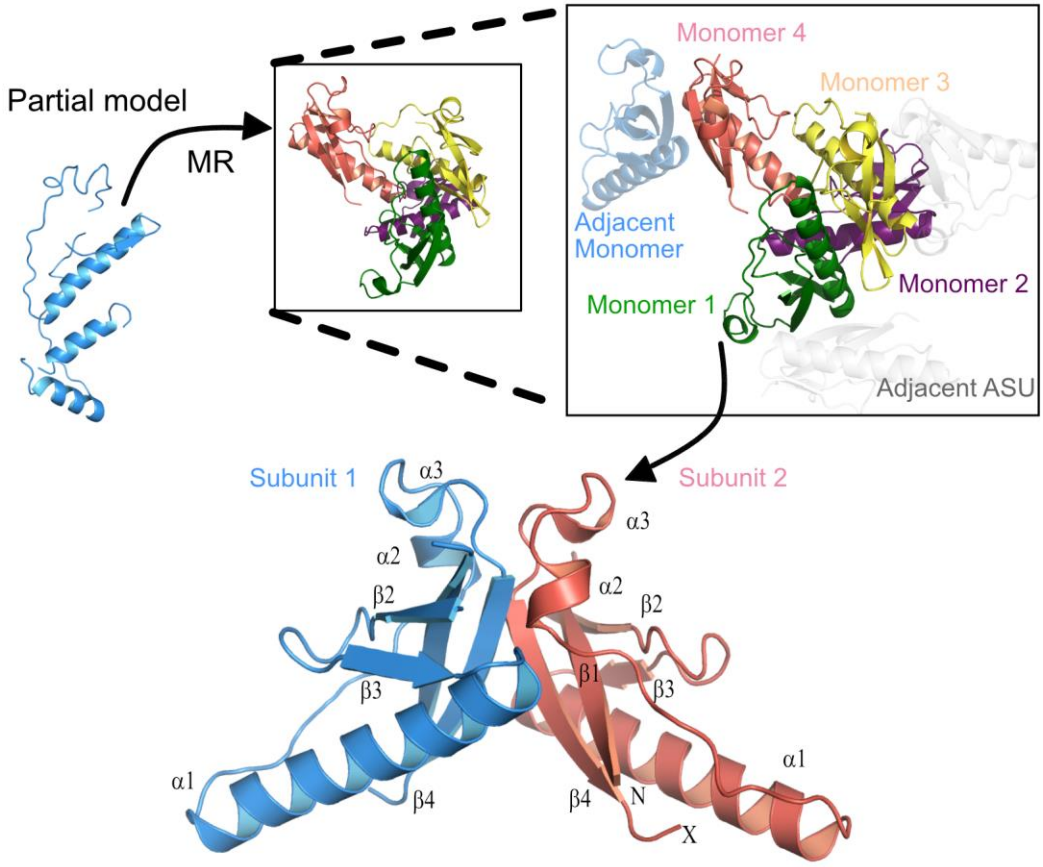


Figure 2.10. Structural determination of HicB-NT. The partial model derived from a SeMet labelled HicB dataset was used as a MR model to determine the 4 monomers of HicB-NT within the ASU (black box) labelled red, green, blue and yellow. Neither of these formed the expected dimeric organisation; however dimers were formed across the ASU interface for each monomer, resulting in the formation of a further four dimer interfaces (expanded black box). For clarity only 3 dimer interfaces within the adjacent ASU are shown here with one highlighted in blue (adjacent monomer) and the others grey (adjacent ASU). (Bottom) One dimer of HicB-NT with a similar interface to HicB3-NT is highlighted in blue and red. Secondary structure elements are highlighted and X refers to the location of the His_6-tag.

The apparent ASU containing four molecules of HicB-NT, was determined to be an artefact of crystallisation by PISA[309] (Appendix Table 5). Investigation of the neighbouring ASU within the crystal revealed the formation of dimeric interfaces across the ASU (Highlighted in blue for subunit 1 and grey for other monomers) - resulting in the formation of four dimers

(Figure 2.10). PISA analysis suggested these latter dimeric interactions had a higher stability in solution than the four molecule ASU assembly observed (ΔG_{diss}: 1.5 v -7.6 kcal/mol) (Appendix Table 6).HicB-NT forms a dimer in solution with each monomer adopting an antiparallel β1β2β3α1α2α3β4 fold, where the central helix α1 (I32-L53) lies within a cleft formed by the β1β2β3 strands within a single subunit (Figure 2.11A). The β4 strand (V77-S82) and the α2 helix (V64-L67) of an individual HicB-NT subunit contact an adjacent subunit forming the dimerization interface (Figure 2.11B). This interface consists of a network of hydrophobic interactions (Figure 2.11C) including those at the N-terminus of the α1 helix (I32, I36), the α2 helix (V64) and the β4 strand (V77-L80). This hydrophobic patch is supplemented by hydrogen bonds (E2-D33, I32-E65, V64-E33 and E65-D63) and one electrostatic interaction (E2-K37) burying a total of 1, 300 Å^2 surface area (from a total of 10, 500 Å^2 over the complete dimer).[309, 310] Convincing electron density could not be observed for the vector encoded C-terminal His_6-tag and loop 1 in any of the four subunits and these were not built into the model.

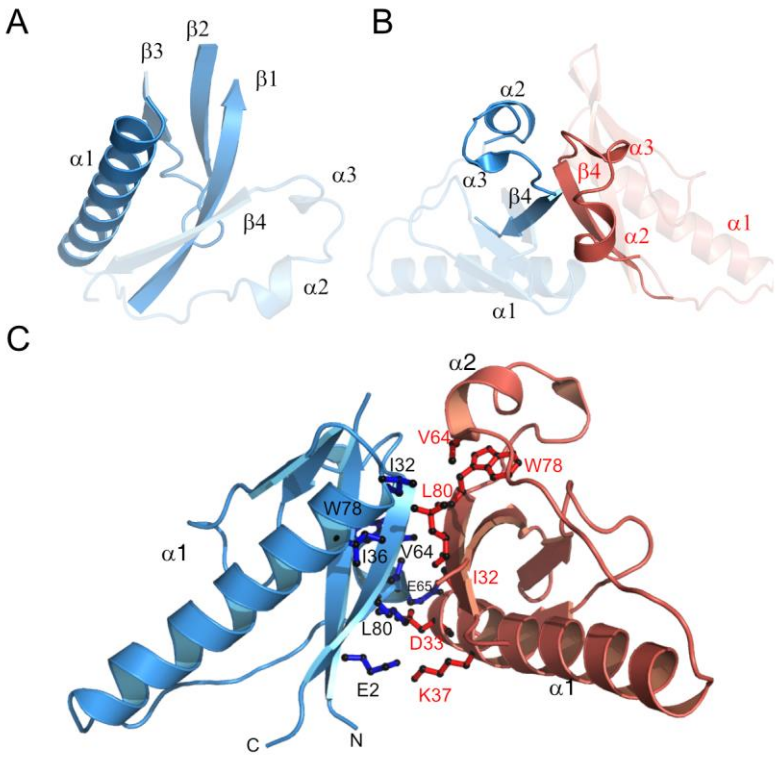


Figure 2.11. Hydrophobic dimerization site of the N-terminal domains of HicB. (A) The α1 helix lies within a cleft formed by the β1β2β3 sheet within a single subunit (blue). (B) Orientation of the dimerization interface between adjacent HicB-NT subunits (red and blue) where the β4 strand is the central element. (C) Full overview of the hydrophobic dimerization interface with residues labelled for subunit 1 (black) and 2 (red).

2.5.1.1. Structural comparison with the N-terminal domain of HicB3

Comparison of HicB-NT and *Yersinia* HicB3-NT (PDB: 4P78) (Figure 2.12) reveal the structural similarities between the two proteins (RMSD: 2.02 Å) despite their low sequence homology (28%). Both proteins possess an identical fold (β1β2β3α1α2α3β4), with only subtle differences. In particular the α2 and α3 of HicB3-NT is almost a continuous alpha helical element with three residues separating the two secondary structure elements. In contrast within HicB-NT there are five residues that separate the α2 and α3 helix. There is gross conservation of the hydrophobic dimerization domain with secondary structure elements conserved, although the specific hydrophobes vary, burying 1, 900 $Å^2$ of the total surface area (11, 300 $Å^2$) within HicB3-NT.

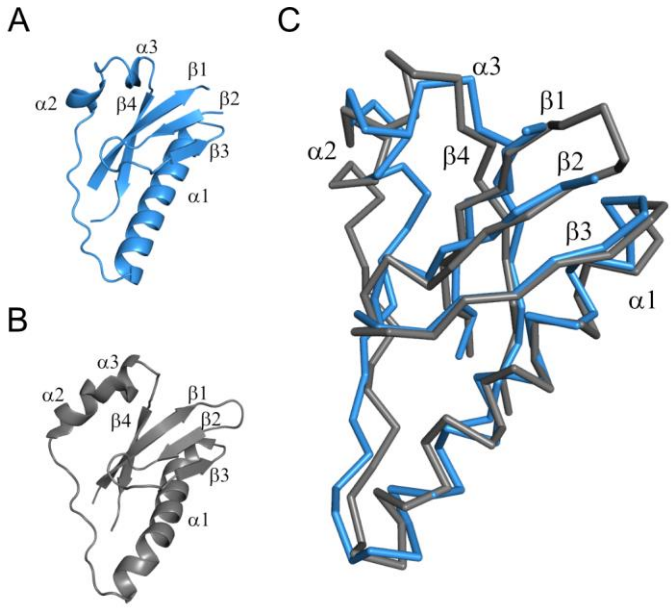


Figure 2.12. Superimposition of the N-terminal domain of HicB3 (grey, PDB: 4P78) onto HicB_NT (blue) for a single subunit. (A) Cartoon representation of the subunit organisation of a HicB subunit (blue).(B) Cartoon representation of the subunit organisation of a HicB3 subunit (grey). (C) Ribbon superimposition of HicB and HicB3 (RMSD: 2.02 Å).

2.6. Structural determination of HicB

With the high resolution HicB-NT data, the previously collected HicB dataset was solved at a resolution of 1.85 Å (Table 2.2). After processing the collected diffraction images into the $P4_1$ space group[278, 279], the ASU was determined to contain four HicB subunits by the Matthews coefficient. The dimeric HicB-NT model was used as a search model for MR in Phenix[280, 281] to identify dimers of the N-terminal domain. Subsequent iterative rounds of

model building and refinement in Coot[303] and Phenix[306] built the disordered hinge region (S85-E93) and the following C-terminal domain (R94-K132) for one subunit into the electron density. A single subunit of HicB (Subunit 1, Figure 2.13) was then used as a search model to place all four subunits of HicB within the ASU. Manual rebuilding into electron density and subsequent and refinement built the N-terminal and C-terminal interfaces between adjacent subunits resulting in the formation of a ring type tetramer, the biological unit of HicB. The final structure at 1.85 Å was validated with MolProbity[307] and the PDB_REDO web server.[308] This structure is subsequently referred to as crystal form A. The tetrameric model of HicB was then used as a search model for MR on the 2.4 Å dataset and this structure of HicB is referred to as crystal form B. Interestingly placement of the C-terminal domains differed in this dataset and the correct placement was achieved by successive rounds of manual building and refinement. The autobuild pipeline in Phenix[305] could not be used as it incorrectly placed the C-terminal α5 and α6 helices into the electron density, so a ring type tetramer was not formed. Statistics for both datasets are presented in Table 2.2.

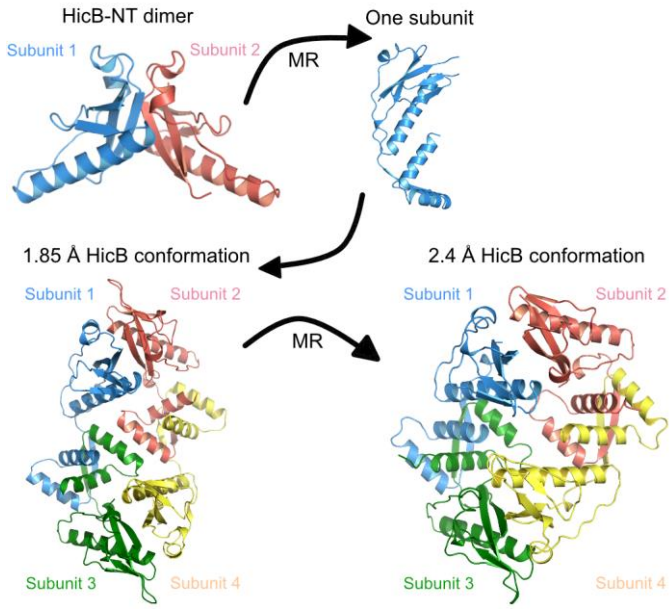


Figure 2.13. Structural determination of the two crystal forms of HicB. The dimeric HicB-NT structure was used as a MR search model to determine the complete subunit of HicB within the ASU of the 1.85 Å dataset. One HicB subunit was then used as a MR search model to determine the four subunit tetramer within the ASU (Crystal form A). The tetrameric crystal form A of HicB was used as a MR search model to determine the tetrameric structure of HicB at 2.4 Å, under different crystallisation conditions (referred to as crystal form B). A 2 fold symmetry is observed within each tetramer as subunits 1 (blue) and 4 (yellow) are equivalent as are subunits 2 (pink) and 3 (green).

Within both crystal forms, HicB existed as a ring-type tetramer, with 2 fold symmetry resembling a dimer of dimers (Figure 2.14A). Each subunit is comprised of an N-terminal domain (M1-L85), a linker region (S86-E93) and a C-terminal domain (R94-K132) (Figure 2.14B). The N-terminal domain retains the β1β2β3α1α2α3β4 fold seen in the HicB-NT structure and the N-terminal dimerization interface observed between adjacent subunits of HicB-NT is conserved within HicB (Figure 2.14B). The C-terminal domain contains a RHH motif (β5α5α6) that forms a strand-swapped dimer formed between adjacent partner C-terminal domains (Figure 2.14C). Dimerization interfaces at the N/C terminal domain of each subunit stabilises the topology of the ring-type tetramer. PISA analysis suggested that the observed tetrameric assembly of HicB within crystal form A was stable in solution based on the positive value of Standard Free Energy of dissociation into nearest stable assemblies (ΔG_diss: 12.4 kcal/mole) (Appendix Table 7).[309] The ring type tetramer interface was suggested to be unique among HicB family members, with only minor similarity to a putative *Lactobacillus plantarum* RNA binding protein (PDB: 3KWR) which resembled the HicB-NT structure (Appendix Table 8). DALI also calculated structural similarity with other HicB family members, but not non-HicB type 2 antitoxins (Appendix Table 9).[311] Instead HicB resembled several RNA binding proteins and endoribonuclease/helicases, consistent with the observation that HicB family members contain a inactive RNase H fold and a DNA binding domain.[269] Furthermore, DALI calculated that the C-terminal domain was highly similar to RHH domain containing proteins: Arc (PDB: 1BAZ)[312], CopG (PDB: 2CPG)[313], DinJ (PDB: 4Q2U)[83], MazF-4 (PDB: 5XE3)[314], ORF Omega (PDB: 2BNZ)[315], PutA (PDB: 2AY0)[316], RelB (PDB: 4FXE)[155] and VapB26 (PDB: 5X3T)[317] (Appendix Table 10).

For the tetrameric conformation of HicB within crystal form A, compelling electron density was not observed for the vector-encoded His_6-tag of each subunit and these were omitted from the final model. All loops and the α2-α3 helix were solvent exposed within the ASU, exhibiting no packing interactions with adjacent subunits.

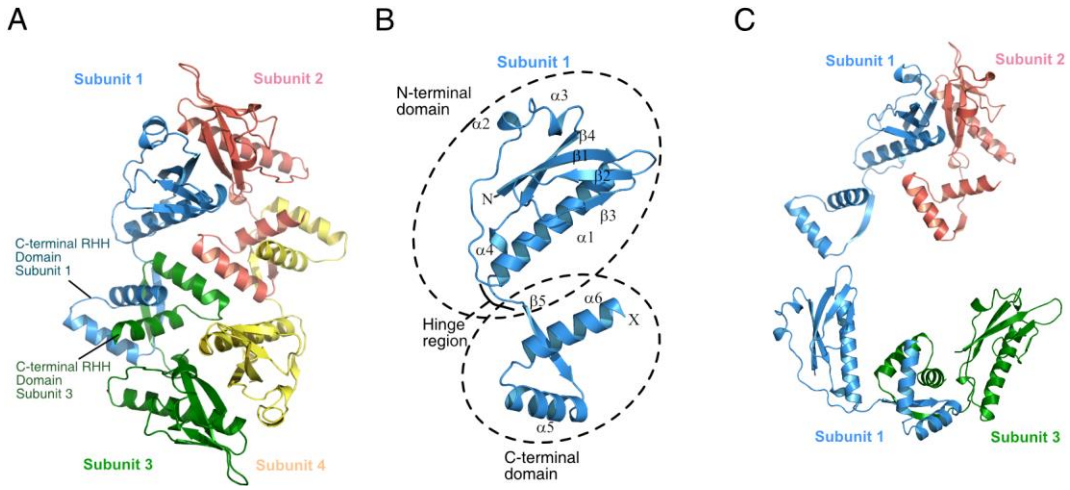


Figure 2.14. Crystal form A of HicB (1.85 Å dataset). (A) The tetrameric structure of HicB coloured by subunit, the dimerization interfaces are formed by both the N-terminal and C-terminal domains across adjacent subunits. (B) Single subunit of HicB with secondary structure elements labelled. (C) Non-symmetrical dimer interfaces between subunits 1 and 2 compared to subunits 1 and 3.

The hydrophobic N-terminal dimerization interface between adjacent subunits is equivalent to the interface observed between adjacent HicB-NT subunits and is not affected by the presence of the C-domain. The C-terminal dimerization interface between adjacent C-terminal domains is comprised of a hydrophobic network, supplemented by hydrogen bonds and tethering electrostatic interactions. Hydrophobic residues within the β5 strand (I95, V97, I99), α5 helix (F102, V103, L104, I107, Y110), and the α6 helix (F121, A122, A123, L127) form a network with corresponding C-terminal domain residues on an adjacent subunit-resulting in a hydrophobic patch (Figure 2.15A). R94 and N96 form side chain-side chain hydrogen bonds across adjacent β5 strands (Figure 2.15B), while electrostatic interactions between E93-R101 and E93-H105 tether the hinge region of one subunit to the α5 helix of an adjacent subunit (Figure 2.15C). Together these bury a total of 3, 200 $Å^2$ surface area (from a total of 9, 100 $Å^2$ over the complete dimer).[309, 310]

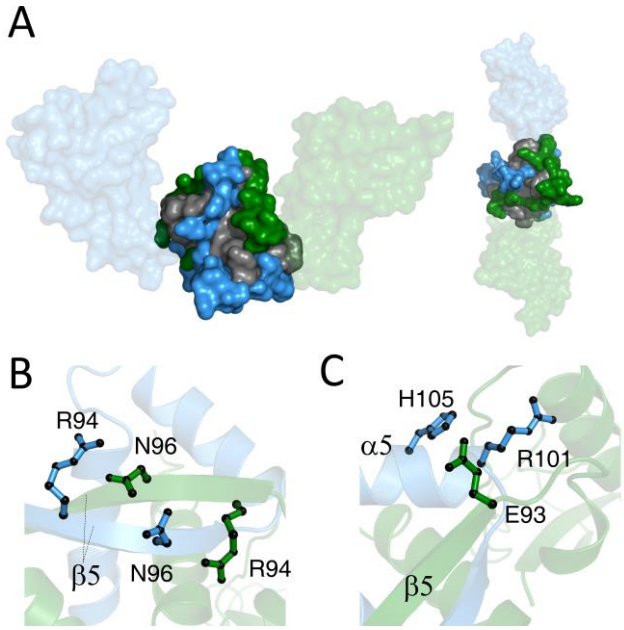


Figure 2.15. Interactions at the C-terminal dimerization interface. (A) Surface depiction of the hydrophobic network at the C-terminal domain. Hydrophobes within each individual subunit (I95, V97, I99, P100, F102, V103, L104, I107, Y110, F121, L122, A123, A125, A126 and L127) implicated in the interface are highlighted in grey. For each image, Subunit 1 and 3 are highlighted blue and green respectively to depict the dimerisation interface. An equivalent interface is found between subunit 2 and 4. (B) Hydrogen bonds between adjacent β5 strands. (C) Electrostatic interactions between the hinge region preceding the β5 strand (E93) and the α5 helix (R101, H105).

The orientation between the N-terminal and C-terminal domains of HicB can adopt one of two different configurations where the N-terminal α1 helices lie either parallel to the C-terminal domains (subunits 1 and 4) or perpendicular and subsequently are solvent exposed (subunits 2 and 3) (Figure 2.16A). This non-symmetric arrangement leads to two distinct sets of hydrophobic and electrostatic interactions between adjacent N- and C-terminal domains (Figure 2.16B). In HicB subunits, when the α1 helix lies parallel to the C-terminal domain (Subunit 1 and Subunit 4), the interface between adjacent subunits is dominated by electrostatic interactions between the N-terminal domain: E48, E52 (α1) and D89 (hinge region) of a subunit and residues of the α5 helix (R101 and H105) of and adjacent subunit (Figure 2.16D, box 1), as highlighted for subunits 1 and 3. Surprisingly no hydrophobic interactions are seen at this interface (Figure 2.16D, box 2). If the α1 helix is surface exposed, for example within subunit 2 and 3, the interface is dominated by hydrophobic interactions (Figure 2.16D, box 3) and no electrostatic interactions are observed (Figure 2.16D, box 4). Residues in and around the α1 helix and the hinge region (I51, V57, F59, L85, L88) form a hydrophobic network with the α5 helix (P100 and F102) of an adjacent subunit, as observed between subunit 2 and 4 (Figure 2.16D, box 3). These interactions tether the C-

terminal domains in a parallel arrangement at the base of the structure. The interface between pairs of adjacent C-terminal domains is comprised of A109, A112 and H115, resulting in a minimal interface between the two (Figure 2.16C).

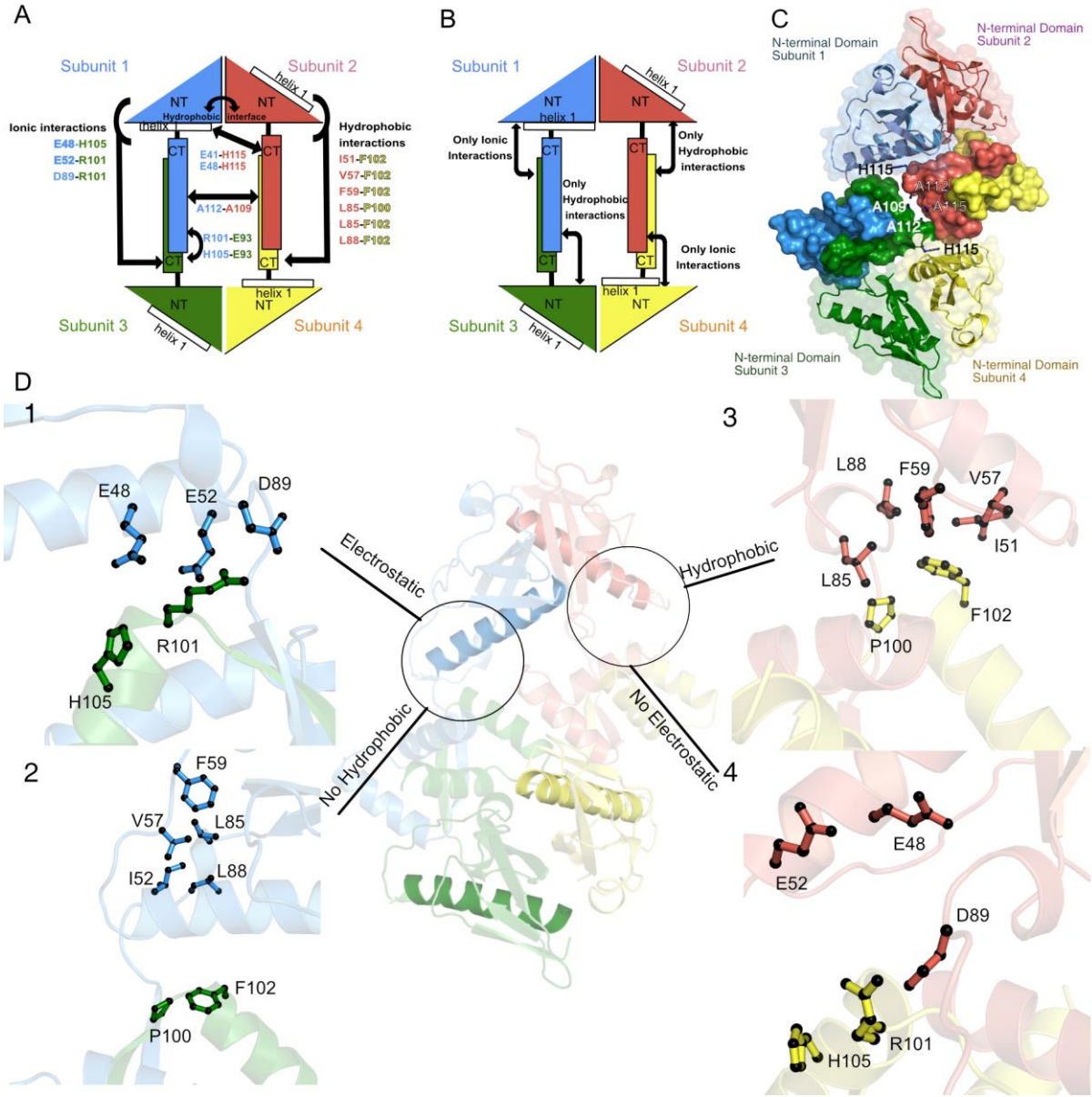


Figure 2.16. Tethering interactions between adjacent N-terminal and C-terminal domains of HicB within the crystal form A (1.85 Å dataset). (A) Cartoon representation of the hydrophobic and electrostatic interactions between adjacent C-terminal domains and N-terminal/hinge region. (B) Cartoon representation of the location of hydrophobic and electrostatic sites, dependent on the placement of the α1 helix. (C) HicB tetramer with the C-terminal domains rendered with a solid surface and residues at the interfaces between C-terminal pairs annotated (D) Interaction sites across the tetramer. Box 1 highlights the electrostatic interface (Residues E48, E52, D89, R101 and H105) between adjacent subunits and box 2 shows the absence of the hydrophobic site (Residues I51, V57, F59, L85, L88, P100 and F102) between adjacent subunits when the α1 helix lies parallel to the C-terminal domain. Box 3 and 4 show the respective presence of the hydrophobic interface and absence of the electrostatic interface between adjacent subunits 2 and 4 when the α1 helix is surface exposed. For clarity only the interfaces at subunit 1 and 2 are highlighted, but the equivalent sites are present at the α1 helix of subunits 3 and 4.

The solution of crystal form B at 2.4 Å was determined using crystal form A of HicB as a search model (Figure 2.17). The experimental data did not contain compelling electron density for the side chains of residues H7-Y15 (loop 1) or the vector encoded His_6 tag and these were omitted from the final model.

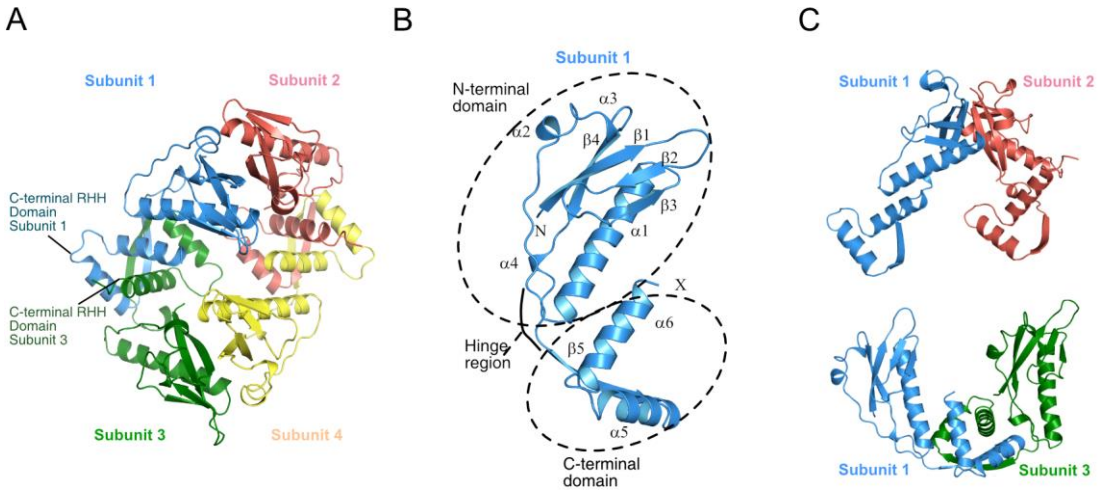


Figure 2.17. Crystal form B of HicB (2.4 Å dataset). (A) The tetrameric structure of HicB coloured by subunit, the dimerization interfaces are formed by both the N-terminal and C-terminal across adjacent subunits. (B) Single subunit of HicB with secondary structure elements labelled. (C) Non-symmetrical dimer interfaces between subunits 1 and 2 compared to subunits 1 and 3.

Surprisingly, despite the conservation of the ring type tetramer, however, the two structures were non-identical. The conformation of crystal form B (2.4 Å) is more compact (73 x 78 Å) than crystal form A (1.85 Å), (63 x 99 Å) with respect to horizontal and vertical axis. The structure of crystal form A appeared to be elongated in comparison to the conformation of crystal form B. Both are now referred to as the compact (2.4 Å) and elongated (1.85 Å) form of HicB respectively. The placement and orientation of the C-terminal domain differs by ~ 5°, resulting in a 9 Å displacement of the C-domain for the elongated conformation (Figure 2.18A). Superimposition of the two structures results in a good agreement for the N-terminal domain and a single subunit (RMSD: 1.8 Å over the carbon backbone) but not the dimer (Figure 2.18 B, RMSD: 4.2 Å) or tetramer (Figure 2.18C, RMSD: 29.45 Å). Conformational shifts within the C-domain are propagated to the tetramer via the strand-swapped dimers resulting in gross variation between the pairs of dimeric C-terminal domains. Strand-swapping has been observed in various proteins to propagate the formation of higher order structurers and give rise to functional units.[318]

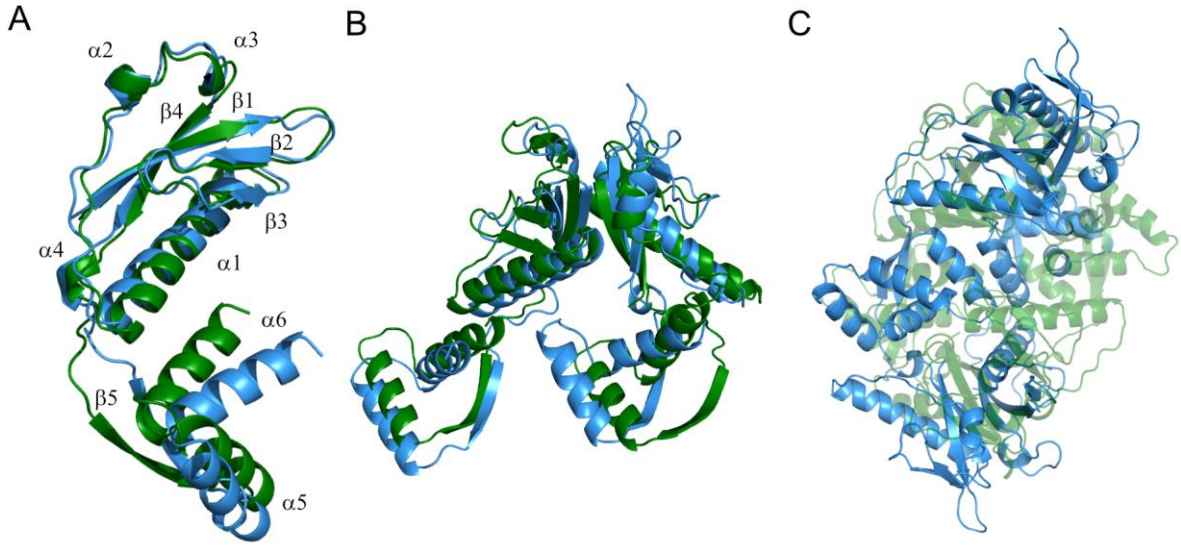


Figure 2.18. Superimposition of the compact and elongated conformations of HicB. (A) Cartoon superimposition of individual subunits of HicB (RMSD: 1.8 Å). (B) Cartoon superimposition of adjacent HicB N-terminal dimers (RMSD: 4.8 Å) (C) Cartoon superimposition of HicB tetramers. The compact and elongated conformations are highlighted in blue and green respectively and all superimpositions were made over the carbon backbone.

The variation between the two structures is explained by an increased number of inter-subunit interactions between adjacent subunits in the compact conformation (Figure 2.19) and a number of additional internal interactions within subunits when compared to the elongated conformation.

A myriad of hydrophobic and electrostatic interactions are conserved within the compact conformation of HicB, dependent on the placement of the α1 helix (Figure 2.19), as previously observed for the elongated conformation of HicB (Figure 2.16). The interfaces between adjacent subunits where the α1 helix lies parallel to the C-terminal domain are conserved within both conformations of HicB. When the α1 helix is surface exposed, the interface is dominated by hydrophobic interactions (Figure 2.19C, box 3). However, the orientation of the α1- α2 loop enables D56, E58 and D89 (hinge) to form new electrostatic interactions with residues in the α5 helix (R101, H105 and K106) of an adjacent subunit, as highlighted for subunits 2 and 4 (Figure 2.19C, box 4). These interactions are not observed in the elongated conformation of HicB and their formation brings the α1 helix closer to the α5 helix resulting in a more compact ring type tetramer.

Additional internal cation-π interactions (Y110-R114 and Y110-R124) between the α4 and α5 helix within the C-terminal domain of the compact form contribute to its rigid nature.

These are not present in the elongated conformation, although it is likely that the C-terminal domains remain dynamic in solution and crystallisation may have captured two conformations of HicB. However, this does not preclude further conformations that might be adopted in the presence of HicA or DNA.

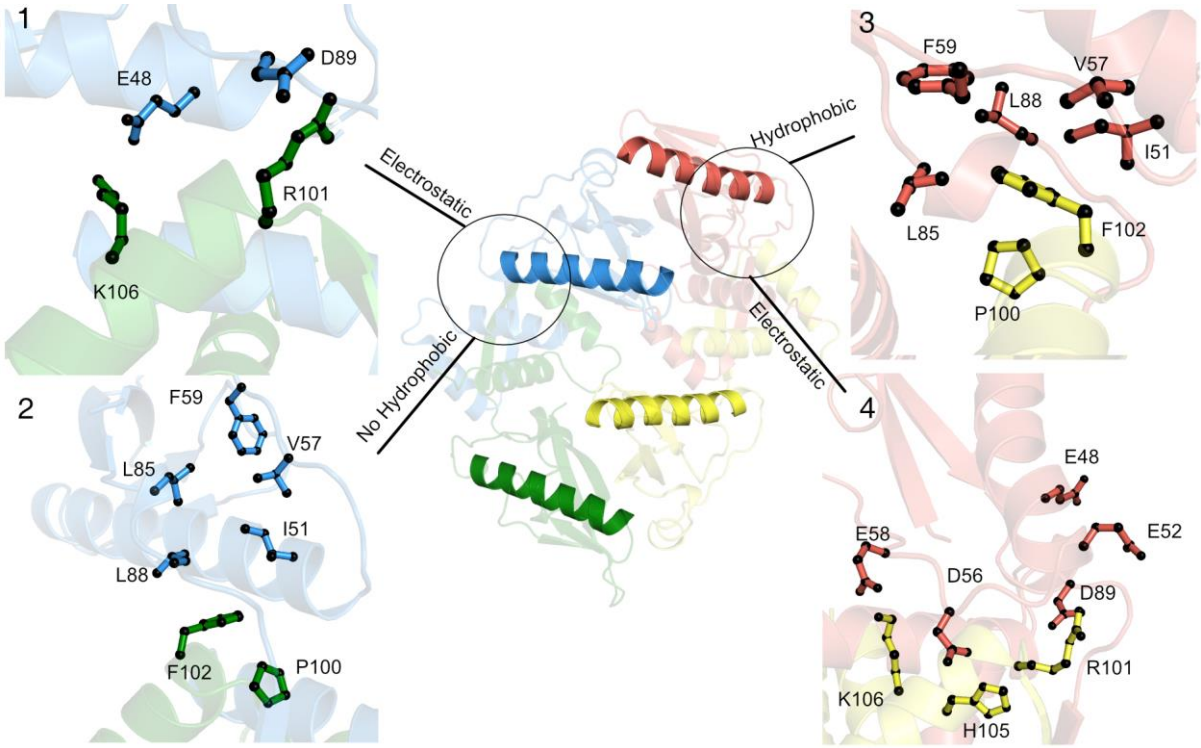


Figure 2.19. Tethering interactions between adjacent N-terminal and C-terminal domains of the compact conformation of HicB. Interaction sites across the tetramer. Box 1 highlights the electrostatic interface (residues E48, D89, R101 and K106) between adjacent subunits when the α1 helix lies parallel to the C-terminal domain. Box 2 indicates the absence of a hydrophobic interface residues I51, V57, F59, L85, L88, P100 and F102) between adjacent subunits 1 and 4. Box 3 and 4 show the respective presence of the hydrophobic interface and absence of the electrostatic interface between adjacent subunits 2 and 4 when the α1 helix is surface exposed. Additional electrostatic interactions (D56-H105, E58-K106 and D89-R101).

SAXS was used to determine the *ab initio* shape envelope of HicB ($HicB_T$ = 125 μM) in solution (Materials and Methods).[319-324] This envelope was used to further confirm the tetrameric organisation of HicB and determine which crystal form of HicB best represented the dynamic solution structure of HicB and therefore the native state (Figure 2.20, Appendix Figure 8, Appendix Table 11). The three-dimensional structures of HicB were also quantitatively analysed to determine whether the computed scattering profile of either structure resembled the experimental profile of HicB observed in solution. The calculated Normalised Discrepancy Function (χ^2) reported (analogous to the R factor) was calculated to determine the fit between the experimental profile and crystal structure.[325] A χ^2 value of 1

indicates an ideal fit to the experimental data and it is generally accepted that a χ^2 value greater than 3 reflects a poor fit to the experimental data (Robert Rambo Personal Communication).[326, 327] The elongated conformation (crystal form A) shows a good agreement to the *ab initio* shape envelope, with a χ^2 value of 1.98 between the crystal structure and the raw scattering data. In comparison the compact conformation of HicB (crystal form B) is a poor fit with a χ^2 value of 9. From this data, it was proposed that the elongated conformation solved at 1.85 Å was a better model to describe the shape of HicB in solution.

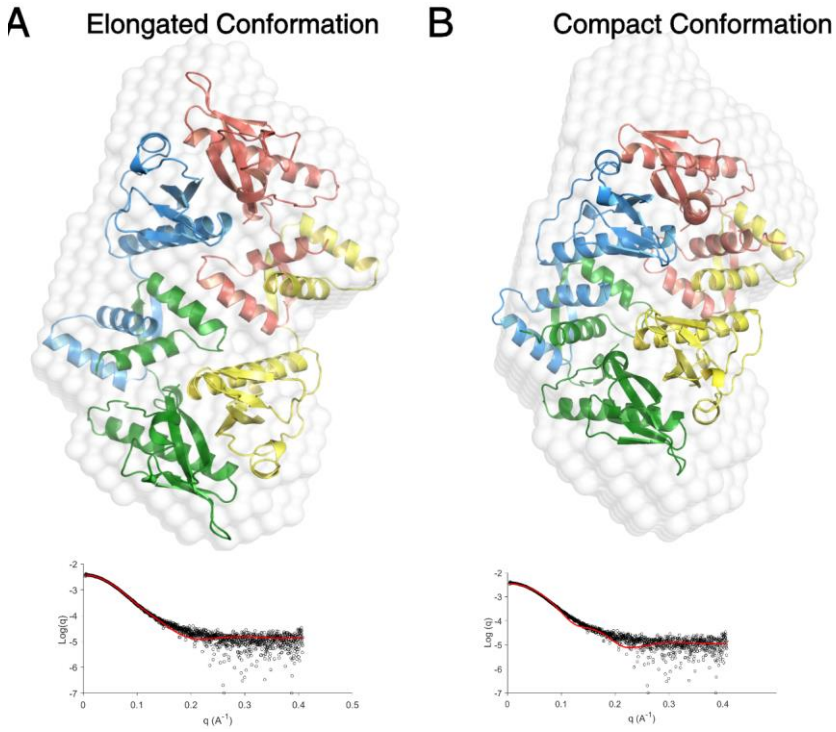


Figure 2.20. Small angle X-ray scattering of HicB. (A) *Ab initio* modelling of the structure of the elongated conformation of HicB into the shape envelope of HicB (white). The FoXS profile of the proposed scattering profile for the crystal structure (red) against the experimental raw scattering data is underneath (χ^2= 1.98) (B) *Ab initio* modelling of the structure of the compact conformation of HicB into the shape envelope of HicB (white). The FoXS profile of the proposed scattering profile for the crystal structure (red) against the experimental raw scattering data (black) (χ^2=9).

2.6.1. Structural comparisons to HicB3

As previously discussed the homologue HicB3 (PDB:4P7D) exists as a ring type tetramer and the gross topology is shared between the two distant homologues despite their sequence identity (28%) with the orientation of α1 helices conserved within individual subunits of HicB3.[132] Superimposition of a subunit of the elongated conformation of HicB to HicB3 (Figure 2.21A, RMSD: 3.6 Å) showed a good agreement at the N-terminal domain. However, subunits differ at the C-terminal domain, with particular deviation of the β5 strand and α6 helix. This C-terminal domain of HicB3 represents yet another conformation that the RHH motif within the HicB family members can adopt, reinforcing its dynamic nature in the absence of DNA/HicA. The topology of the overall ring type tetramer of HicB3 varies from the elongated conformation of HicB (RMSD: 4.5 Å), predominantly due to the different orientations of the C-terminal domains (Figure 2.21B).

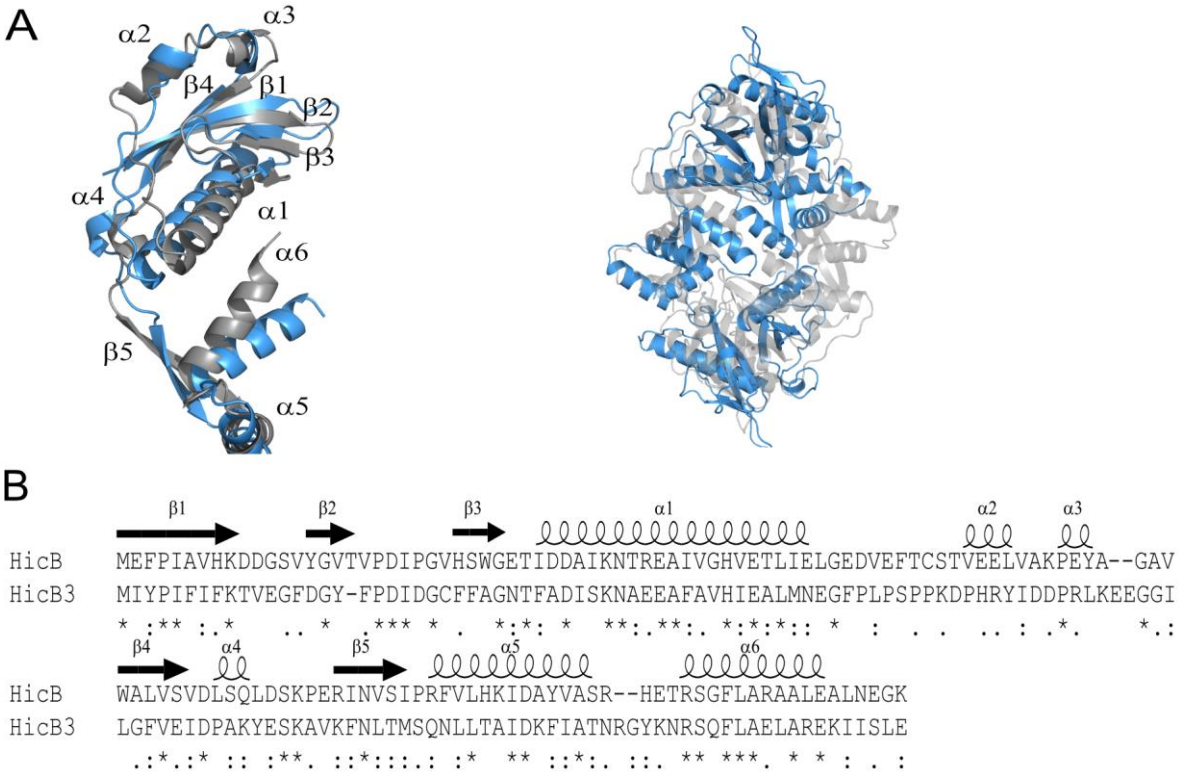


Figure 2.21. Superimposition of the elongated conformation of HicB and HicB3. (A) Cartoon representation of the superimposition of individual subunits of HicB (blue) and HicB3 (grey), (RMSD: 3.6 Å) and their tetrameric form (RMSD: 4.5 Å). The vector encoded His_6-tag was not observed in either structure. (B) Sequence of *B. pseudomallei* HicB and comparison to the structural homologue, HicB3 from *Y. pestis*. The vector encoded His_6-tag for HicB and HicB3 is not shown. The secondary structure elements from HicB are displayed at the top of the alignment, symbols indicate a conserved residue (*), conservative mutation (:) and a semi-conservative mutation (.).

The largest difference between HicB and HicB3, is the presence of a longer α4 helix in HicB3 (^{85}DPAKYE^{90}) compared to HicB (^{85}LSQ^{87}). The HicB3 α4 helix within the hinge region makes two internal interactions to the α1 helix (Y89-E39 and Y89-F42) within individual subunits that are not present in HicB. These additional interactions may tether the hinge region to the N-terminal domain, reducing the flexibility of this region and therefore the subsequent C-terminal domain (Figure 2.22A). Additional cation-π interactions are observed between adjacent C-terminal domains (F96-R121, F111-K132 and Y118-K132) (Figure 2.22B). These additional interactions strengthen the dimerization interface between adjacent C-terminal domains. Both F111 and K132 are conserved in HicB but poor electron density was observed for K132 and the side-chain orientation could not be determined. F96 and Y118 are replaced by I95 and K106 respectively within the β5 strand and the α4-α5 loop.

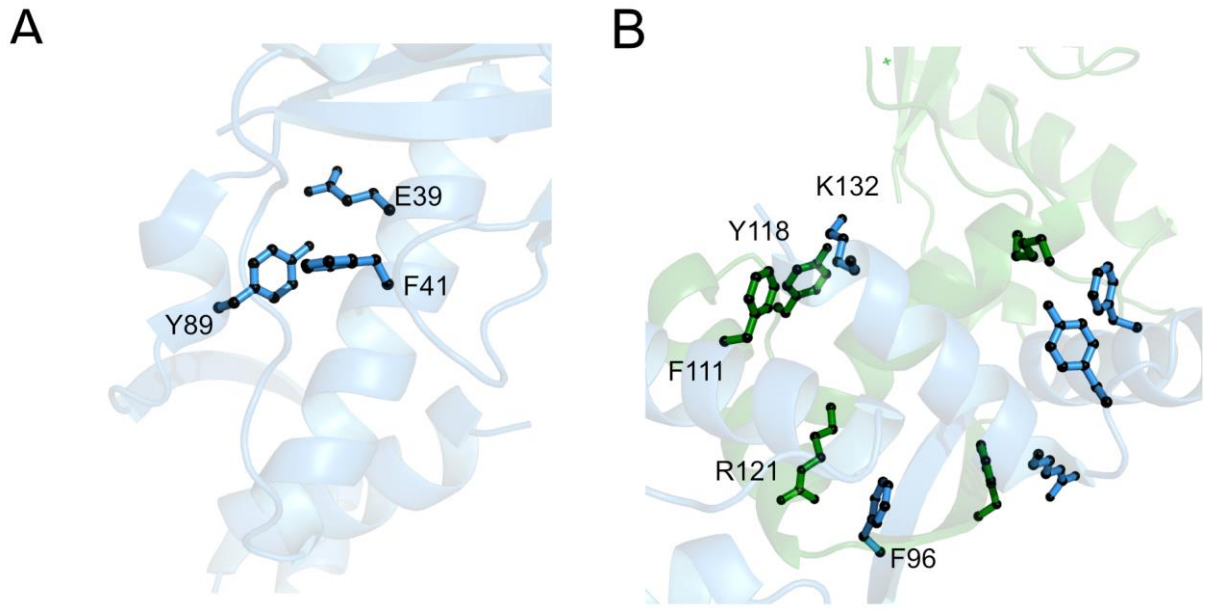


Figure 2.22. Aromatic interactions within HicB3. (A) Anchoring interactions between the α4 helix (Y89) and the α1 helix (E39 and F41). (B) Cation-π interactions between adjacent C-terminal domains (F96-R121, F111-K132 and Y118-K132) stabilise the hydrophobic interface that drives dimerization. For clarity only one set of interactions are labelled. Adjacent HicB3 subunits are labelled blue and green respectively.

While not discussed in the original publication of HicB3[132], similar sets of tethering interactions are seen between adjacent subunits (Figure 2.23). Again, like those identified in HicB, the network of interactions are dependent on the orientation of the α1 helix and whether this lies parallel to the C-terminal domain (electrostatic network) or perpendicular to it (hydrophobic network).

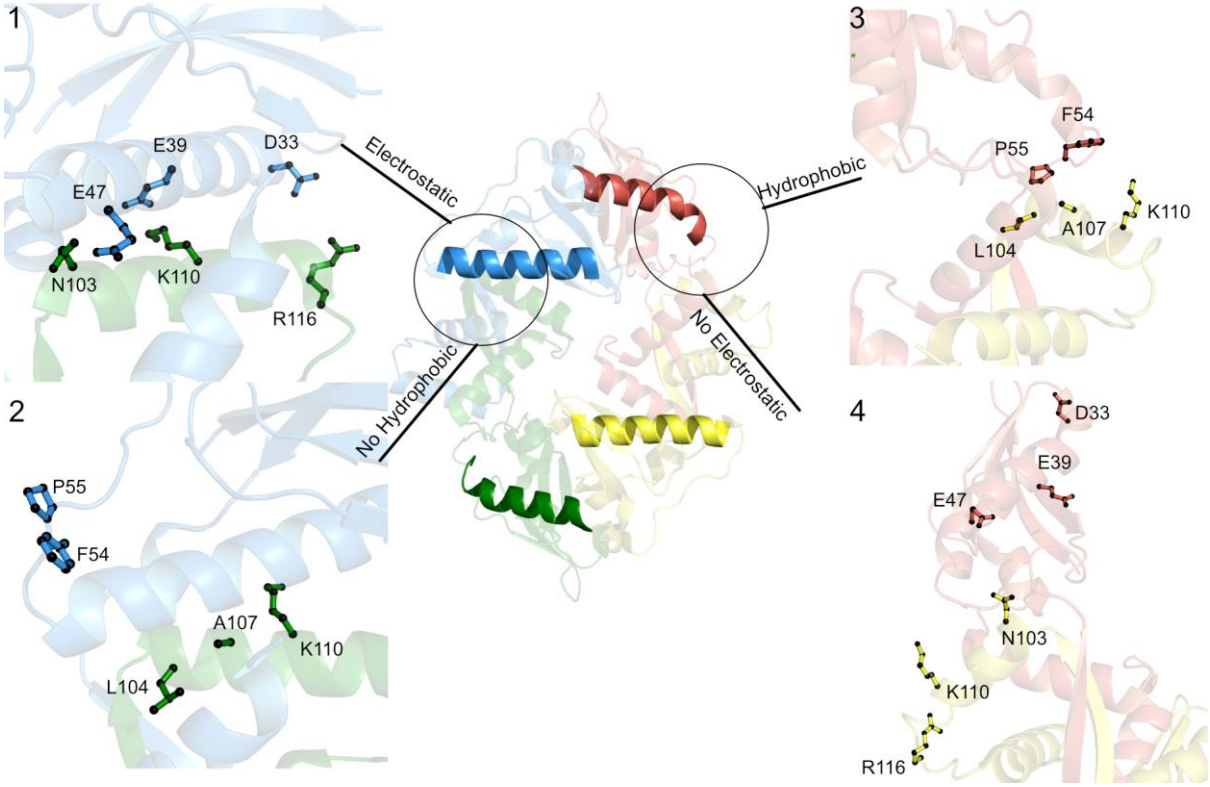


Figure 2.23. HicB3 interaction sites across the tetramer. Interaction sites across the tetramer. Box 1 highlights the electrostatic interface (residues D33, E39, K110 and R116) between adjacent subunits 1 and 3. An additional hydrogen bond is seen at the C-terminus of the α1 helix (E47) with subunit 3 (N103). Box 2 shows the absence of the hydrophobic site (residues F54, P55, L104 and A107) between adjacent subunits, where the α1 helix lies parallel to the C-terminal domain. Box 3 and 4 show the respective presence of the hydrophobic interface and absence of the electrostatic interface between adjacent subunits when the α1 helix is surface exposed highlighted for subunit 2 and 4.

The electrostatic interface consists of interactions between the α1 helix (D33 and E39) of HicB3 and the α5 helix (K100 and R116) of an adjacent subunit, as highlighted between subunit 1 and 3 (Figure 2.23, box 1). A supplementary hydrogen bond is formed between the C-terminal of the α1 helix (E47) and the adjacent α5 helix (N103) that further stabilises this site (Figure 2.23, box 1). The hydrophobic network consists of interactions between the α1-α2 loop (P54 and F55) with residues in and around the α5 helix (L104 and A107) of an adjacent subunit, as highlighted between subunit 2 and 4 (Figure 2.23, box 3). Fewer hydrophobic contacts are present by comparison to the equivalent HicB site (Figure 2.16, box 2) although there is an additional cation-π interaction between F55 and K110. Aromatic amino acids responsible for the hydrophobic interface in HicB (P100 and F102) lie on the α5 helix interacting with residues within an adjacent α1 helix (I51, V57, F59, L85, L88). In HicB3 these equivalent aromatic residues (F54 and P55) are found in the α1-α2 loop and interact with hydrophobic residues around the α5 helix of an adjacent subunit. Similar to

HicB, if the electrostatic interface is present the hydrophobic interface is not present (Figure 2.23, box 2) and vice versa (Figure 2.23, box 4). Conservation of the tethering interface, despite the low homology between HicB and HicB3 (28%) suggests a conserved mechanism of spatial regulation of the tetramer between HicB family members. This may ensure that the HicA and DNA binding sites are surface exposed and correctly orientated for optimal binding.

2.7. Summary

This chapter consists of the expression and purification of several HicB constructs and the structural determination of 3 HicB variants: HicB-NT dimer (1.56 Å), compact (2.4 Å) and elongated (1.85 Å) tetramer of HicB.

Attempts to solve HicB by MR were initially unsuccessful and could only be solved via a combination of heavy atom phasing through SeMet labelling and the 1.56 Å structure of the HicB N-terminal domain. Subsequent MR solved the tetrameric conformation of HicB from two independent datasets at 2.4 Å and 1.85 Å. Each structure displayed an altered conformation with respect to placement of the C-terminal domain within subunits. Adjacent C-terminal domains dimerise to form intact RHH domains. These are predicted to be flexible in the absence of DNA and there are minimal interactions between individual N/C terminal domains and between adjacent RHH domains.

The ring type tetramer was stabilised by a myriad of tethering interactions: hydrophobic and electrostatic, dependent on the position of the α1 helix, within individual subunits. Structural comparisons of HicB to the *Yersinia* HicB3 homologue suggested that these tethering interactions may be conserved in HicB family members

Project	HicB SeMet	HicB-NT	HicB	HicB
Wavelength (Å)	0.9790	0.9763	0.9763	0.9763
Resolution range (Å)	59.85-2.73 (2.86-2.73)	54.30 - 1.56 (1.62-1.56)	59.62-2.38 (2.465-2.38)	39.42 - 1.85 (1.92 - 1.85)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P4_1$	$P\ 4_1$
Unit cell	58.7 59.9 172.8Å 90 90 90°	63.5 76.7 76.9Å 90 90 90°	63.1 63.1 181.2 Å 90 90 90°	62.6 62.6 173.5Å 90 90 90°
Multiplicity	12.7 (13.1)	13.2 (13.0)	5.7 (4.9)	4.7 (4.7)
Completeness (%)	99.3 (99.3)	100.0 (100.0)	99.0 (98.0)	99.0 (99.0)
Mean I/σ(I)	16.0 (1.5)	29.9 (2.1)	15.2 (1.9)	9.84 (0.40)
Wilson B-factor ($Å^2$)	70	40	70	50
R_{meas}	0.11 (2.09)	0.04 (1.26)	0.06 (0.87)	0.07 (3.96)
$CC_{1/2}$	0.998 (0.621)	1.000 (0.805)	1.00 (0.946)	0.998 (0.129)
DelAnom correlation between half-sets	0.493 (-0.03)	-	-	-
R/R_{free} **for partial model**	0.327/0.390	-	-	-
Reflections used in refinement	-	53900	28098	55828
R_{work}	-	0.200 (0.3972)	0.221 (0.3359)	0.202 (0.4480)
R_{free}	-	0.233 (0.4530)	0.269 (0.3717)	0.236 (0.4517)
Number of protein atoms	-	2502	4088	4327
RMS (bonds) (Å)	-	0.013	0.009	0.008
RMS (angles) (°)	-	1.14	1.06	0.86
Ramachandran favoured (%)		100	95	97
Ramachandran allowed (%)		0.32	3	2.5
Ramachandran outliers (%)		0	1.9	0.2
Rotamer outliers (%)		0.7	2.5	1.2
Average B-factors ($Å^2$)				
Protein	-	40	90	70
Solvent	-	50		50

Table 2.2. Refinement statistics for HicB SeMet, HicB-NT and HicB at resolutions of 2.4 and 1.85 Å. Values in parenthesis correspond to the highest resolution shell.

3. Structural studies on HicAB

This chapter reports the complete three-dimensional crystallographic structure of the HicAB complex from *B. pseudomallei.* This now provides the first example of an intact HicAB complex and comparison with free HicB that extends our current knowledge beyond that provided by the partial HicA3-HicB3 complex.[132] The truncation of the C-terminal domains in the hetero-tetramer reported by Bibi-Triki et al. left numerous unanswered questions surrounding the stoichiometry of the HicA3/HicB3 interaction, how HicA3 would be sterically accommodated in the complete complex and how HicA3 interacts with HicB3 to modulate DNA interactions and repression/de-repression of the *hicA3B3* operon.[132] During the writing of this work a second complete HicAB complex from *S. pneumoniae* was reported (PDB:5YRZ)[272] but was crystallised in a distinctly different conformation than observed for *B. pseudomallei*, raising further important questions surrounding HicAB structure, dynamics and functional diversity between family members.

3.1. Expression and purification of HicA

A pET26-b HicA construct containing HicA bearing a H24A mutation (Provided by Dr Aaron Butt, University of Exeter) was expressed using T7 Express cells (NEB) and purified by IMAC.[131] During purification, it was apparent that previously reported yields of HicA could not be replicated. To improve yields, a codon optimised HicA-H24A gene was synthesized and cloned into a pET151-D/TOPO vector (ThermoFisher™) to generate pET151-HicA. This vector encodes an N-terminal His_6-tag followed by a TEV cleavage sequence (ENLYFQ) and the HicA-H24A gene. The gene product is herein called HicA. The TEV cleavage site was selected, due to the in-house availability of His_6-TEV, via a pET9d-His_6-TEV construct (kindly provided by Dr Matthew Goodwin, University of Bristol).

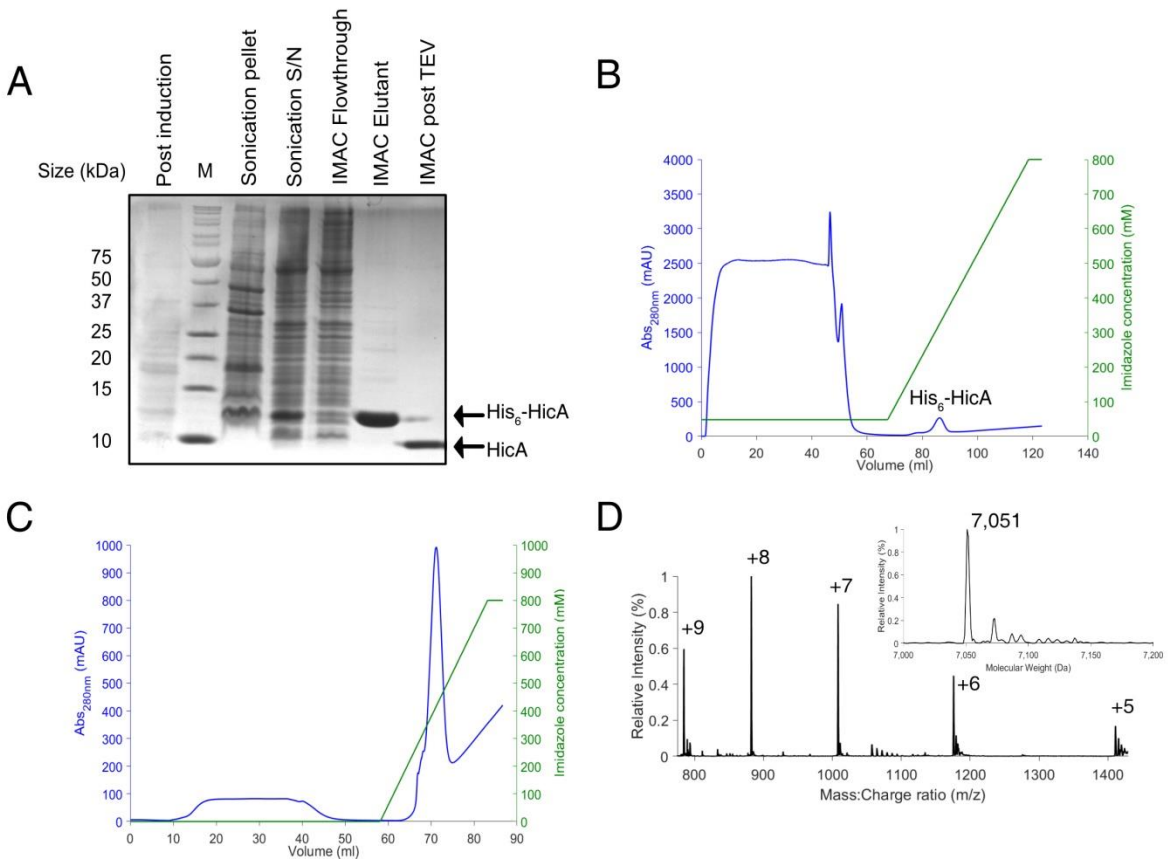


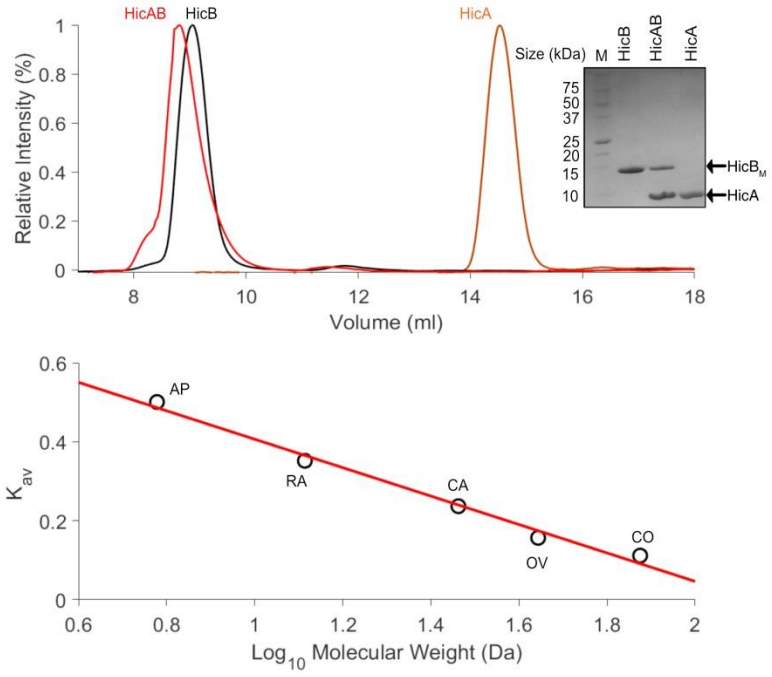
Figure 3.1. Expression and purification of HicA. (A) SDS-PAGE following the expression and purification of HicA from post induction to isolation via IMAC. Prior to TEV cleavage, observed mass ~13 kDa, expected mass = 10,204 Da and post TEV cleavage, observed mass ~10 kDa, expected mass = 7,052 Da. (B) IMAC of HicA eluting as a single peak. (C) IMAC of HicA post TEV cleavage. Untagged HicA eluted during the flow-through (15-45 ml). (D) Denatured ESMS of HicA. m/z signals spanning 780-1,400 were assigned based on their charge state. Inset: De-convoluted spectrum of the m/z envelope to determine the monomeric mass of HicA.

HicA was expressed as soluble protein in *E. coli* T7 Express cells (NEB) and purified by IMAC (Figure 3.1A, B). SDS-PAGE revealed a single band corresponding to monomeric His_6-HicA after IMAC and the mass confirmed by ESMS (observed: 10,203 Da, expected: 10,204 Da) (Appendix Figure 9). Proteolytic cleavage (TEV) and negative purification (IMAC, Figure 3.1C) yielded cleaved HicA as judged by SDS-PAGE (Figure 3.1A) and confirmed by ESMS (Observed mass: 7,051 Da, expected: 7,052 Da) (Figure 3.1D). A small proportion of His_6-HicA was present after negative purification via IMAC (Figure 3.1A) and could not be removed by subsequent purification steps (repeated negative IMAC steps and downstream SEC). HicA was deemed to be ~ 90% pure with a final yield of 1.5 mg/L. Purification of HicA prepared using M9 minimal media supplemented with $^{15}NH_4Cl$ yielded a uniformly ^{15}N labelled sample for study by two-dimensional heteronuclear NMR. The ^{1}H-

^{15}N HSQC spectrum of HicA acquired by Professor Matt Crump and Dr Chris Williams was similar to a ^{1}H-^{15}N HSQC acquired previously using ^{15}N labelled protein prepared from the pET26-b HicA construct and from which a full solution structure was determined (Appendix Figure 10).[131] However residues I7, D44, L45, K55 and S56 which were previously assigned in the pET26-b HicA spectrum could not be assigned for HicA. As the pH (7.5) and temperature (25 °C) remained constant it is likely that the unassigned peaks are close in proximity to those that were assigned in the pET26-b HicA spectrum and small chemical shift perturbations have hindered the assignment of these residues for HicA. Moreover, both linker regions of the proteins (HicA: GIFPFT and pET26-b HicA: DRWGSEL) upstream of the start methionine could not be assigned and partly account for the disparity between the two spectra. Despite the disparity between the two spectra, the current preparation was deemed to contain folded protein.

3.2. Assaying complex formation between HicA and HicB

Prior to crystallisation trials, analytical SEC (ASEC) was used to investigate the oligomeric state of HicB and its interaction with HicA. A Superdex 75 10/300 GL analytical column (AS75) was pre-calibrated with proteins spanning 6-75 kDa including aprotinin, carbonic anhydrase, ovalbum, conalbumin and blue dextran (Figure 3.2, Appendix figure 11, Materials and Methods, equation (1)). The peak elution volume (V_e) of each sample allowed the calculation of the apparent molecular weight (Mw_{app}) and inferred the oligomeric state of each sample. The peak width was used to calculate the range of molecular weight values encompassed by each sample. HicB (50 μM, $HicB_T$ = 12.5 μM) eluted as a single species at an Mw_{app} of 65.6 kDa (52.3-76.5) with parenthesis indicative of the molecular weight range (kDa) encompassed by the peak. Similarly, HicA (50 μM) eluted as a single species corresponding to a Mw_{app} of 8.2 kDa (6.5-9.8 kDa). Both proteins are in agreement with their expected molecular weights (63.0 and 7.1 kDa).



Sample	Mw (Da)	Mw_{app} (kDa)	Mw range (kDa)	$Log_{10}Mw$	V_e (ml)	K_{av}
HicA	7,053	8.3	6.5-9.8	0.92	14.5	0.44
$HicB_T$	62,952	65.6	52.3-76.5	1.82	9.03	0.11
$HicAB_T$	77,058	71.9	46.8-95.0	1.86	8.79	0.10

Figure 3.2. AS75 profiles of $HicB_T$, $HicAB_T$ and HicA. (Top): SEC profiles for all reported constructs and the corresponding SDS-PAGE of eluted samples. Arrows refer to monomeric HicB ($HicB_M$) and HicA. (Middle): Calibration curve using known standards: aprotinin (AP, 6.5 kDa), ribonuclease A (RA, 13.7 kDa), carbonic anhydrase (CA, 29 kDa), ovalbumin (OV, 42.7 kDa) and conalbumin (CO, 76 kDa) used to determine the Mw_{app} for each analysed protein. (Bottom): Summary of results for each analysed sample reporting the expected molecular weight (Mw), Mw_{app}, Range of molecular weight values encompassed by each peak, Log_{10} Mw, V_e (Volume of elution) and K_{av} (Partition coefficient) determined by the calibration curve and equation (1).

Complexation with HicA resulted in a complete loss of the HicA peak and a new peak eluted at a higher Mw_{app} than HicB alone at (71.9 kDa) (46.8 - 95.0 kDa, Figure 3.2A). SDS-PAGE analysis confirmed this new peak contained both HicA and HicB thereby confirming the proteins had formed a complex. The Mw_{app} of HicAB corresponds to 0.93 x the expected molecular weight of a $HicA_2HicB_4$ complex (77.1 kDa) and the profile also encompasses $HicA_3HicB_4$ (84.1 kDa) and $HicA_4HicB_4$ (91.2 kDa). In addition, the profile of HicAB displayed a shoulder spanning 8.05- 8.55 ml corresponding to a molecular weight of 78.6 – 95.0 kDa. ASEC therefore confirmed that complexation had occurred, but it was not possible to definitively assign an oligomeric state. SEC analysis of the homologous HicA3B3 suggested that a $HicA3_2HicB3_4$ complex was the major species formed in solution which is in close agreement with the predominant species observed here.[132]

3.3. Crystallisation of HicAB

The HicAB complex was screened using a variety of crystallisation concentrations (500-2000 µM) and different preformed stoichiometric states against commercially available Molecular Dimensions screens (Morpheus, Proplex, JSCG+ and Structure Screen I + II).

Small crystals of preformed $HicA_2HicB_4$ (HicA: HicB ratio = 250:500 µM) formed in 0.1 M MES, pH 6.5, 0.2 M NH_4SO_4, 30% (v/v) PEG 5000 MME, but crystals were fragile and could not be efficiently looped. A preformed complex 1:1 ratio of HicAB (500:500 µM, $HicA_4HicB_4$) formed longer thin crystals in the same conditions that could be looped. Only crystals free of physical deformities were looped in 25% (v/v) glycerol and directly frozen for diffraction analysis. A dataset of 1,800 images was collected at wavelength 0.9795 Å, exposure time: 0.06 s, oscillation: 0.2 ° and beam flux 100%, with cryogenic temperatures maintained throughout data collection.

Autoprocessing by the FastDP pipeline suggested a resolution of 2.75 Å. Data were processed as described in Materials and Methods into the space group: P 2 with an ASU containing seven molecules of HicA/HicB using the Matthews Coefficient. Attempts to solve the dataset using MR with existing HicB crystal structures and the NMR structure of HicA (PDB: 4C26) as search models were however unsuccessful.[131, 280, 281] Searches for individual N- or C-terminal domains of HicB were also unsuccessful. One N-terminal domain of HicB could be identified but it was not possible to manually build adjacent subunits into the surrounding electron density. No MR solution refined below an R_{free} of ~ 50%.

Rotation of individual domains was not initially accounted for in any search models used in MR or within attempts to manually rebuild subunits of the tetramer. The structural determination of free HicB had however, revealed that the C-terminal domains could adopt different orientations and might also reorient upon HicA binding. Using this notion, the structure was solved by Dr Michail Isupov/Professor Nic Harmer (University of Exeter) using the C-terminal strand-swapped dimer of the elongated conformation of HicB as a search model for MR in MOLREP[328] within CCP4[329] (Figure 3.3A). Two C-terminal dimers were then found within the HicAB dataset with rotation peaks 1 (5.74 σ) and 5 (4.45 σ). These C-terminal dimers were used as fixed models and four N-terminal domains of HicB could only be located when 200 rotation peaks were used in the translation search (Peaks 5, 10, 162 and 4). The resultant HicB model was refined to R/R_{free} 0.408/0.422 at 2.5 Å before

four copies of HicA (PDB:4C26) were positioned into the resultant electron density using the spherically averaged phased translation function and phased translation function implemented in MOLREP.[330] The four-fold NCS averaging in DM[300] was used for phase improvement of HicAB structure with masks and NCS operators calculated separately for HicA and each domain of HicB. Density modification phases were input for phase refinement[331] in REFMAC5.[301] Resulting electron density maps allowed building of the hinge regions of HicB monomers in Coot.[303] Manual rebuilding and refinement further improved the R_{free} value associated with the model before the final structure was validated with MolProbity[307] and the PDB_REDO web server.[308] Statistics are shown in Table 3.1.

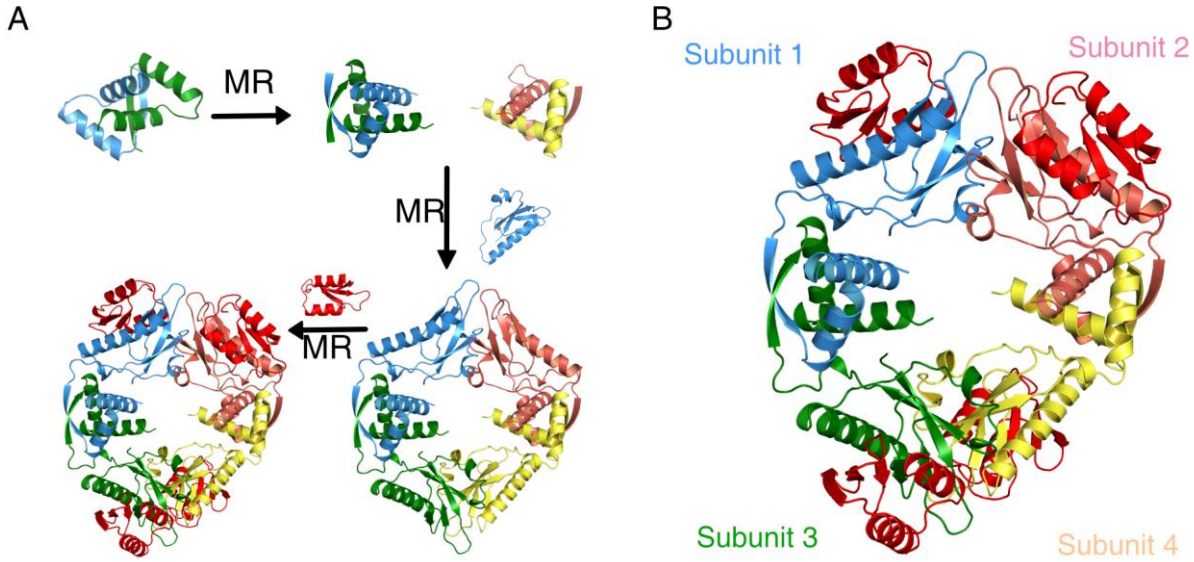


Figure 3.3. The HicAB hetero-octameric complex. (A) The strand-swapped C-terminal domain of HicB was used as a search model to identify electron density corresponding to these domains. The four N-terminal domains of HicB could then be placed by MR. Finally, four HicA subunits (PDB: 4C26) were built into the remaining electron density. (B) Cartoon representation of the HicAB hetero-octamer where HicA (red) binds to the N-terminal domain of each HicB subunit (blue, pink, green and yellow).

The structure of the HicAB complex (Figure 3.3B) clarified the difficulties encountered during initial MR trials. Firstly, the initial space group assignment (P 2) was incorrect and was re-assigned as P 2_1 with an ASU of 8 molecules (four HicA and four HicB chains).The $P2_1$ spacegroup was not searched for during MR trials. Secondly, the C-terminal domains are rotated by 90° with respect to their initial orientation in HicB (Figure 2.14), but as the core of the strand-swapped dimers remained intact, this ultimately aided the molecular replacement.

The asymmetric unit of the HicAB complex is arranged as a symmetrical hetero-octamer with four identical HicA interaction sites at the N-terminal domains of HicB (Figure 3.3B). Here the positively charged face of the β-sheet of HicA binds to the α1 helix of HicB. The conversion of HicB tetramer to the symmetric HicAB hetero-octamer requires extensive conformational rearrangements that generate a more open conformation (Figure 3.4A). These conformational rearrangements are most marked for subunits 1 (blue) and 4 (yellow) of HicB that undergo a domain rearrangement to flip the α1 helix from a sequestered conformation (α1 helix tethered to a C-terminal domain via electrostatic interactions) into an exposed conformation required for HicA binding, in the process breaking these electrostatic interactions. Conversely helix α1 is already partially solvent exposed in subunits 2 (pink) and 3 (green) of HicB and undergoes a smaller 2.4 Å lateral displacement upon HicA binding (Figure. 3.4A)

For the rearrangements in subunits 2 and 3 of HicB, displacement of the α1 helix is accompanied by reshuffling of the hydrophobic interactions within the N-terminal domain. Reshuffling of these elements reduces the length of the longest axis of the symmetrical conformation of HicB by 14 Å compared to the unbound conformation of HicB. The most significant hydrophobic switch arises from L85 (HicB subunits 2 and 3) that rotates 60°, significantly reducing the tethering interaction to the C-domains of subunit 4 and 1 respectively (P100 and F102) (Appendix Figure 12). Similarly, L88 rotates into an empty pocket created by the 2.8 Å shift of E48 (HicB) that forms a new salt-bridge with R19 of HicA and breaking the L88-F102 interaction between adjacent HicB subunits.

The loss of these packing interactions allows the C-terminal domains to rotate outwards in HicAB (Figure 3.4A) without any significant internal rearrangements (hydrophobic packing is 80% conserved within the C-domains between HicB and HicAB): L85 and L88 are drawn into a hydrophobic network (I22, I43, V47, V57 and V83) within the N-terminal domain, stabilising the open conformation (Figure 3.4B). L85 acts as a direct replacement for F59 (α1-α2 loop) which mediates several internal hydrophobic interactions within the N-terminal domain of each subunit (V47, V57, L85 and L88) in the free form of HicB (Appendix Figure 12) and these are disrupted upon α1 helix displacement. Residues from the C-terminal helix (R124 and L127) now form new intra-subunit interactions with residues F59, D84, S86 and Q87 from the N-terminal domain (Appendix figure 13). The hydrophobes P100 and F102 (having lost interaction with L85 and L88) now form inter-subunit interactions with N-

domain residues I51, V57, F59 and P92 (Figure 3.4B). Supplementary electrostatic interactions between adjacent subunits: D89-R101 and E58-K106 form additional inter-subunit interactions to stabilise the symmetrical conformation of HicB.

PISA confirmed that the hetero-octameric state of HicAB represented a biological interface and was not a crystal artefact (Appendix Table 12).[309] A ΔG_{diss} of 8.8 kcal/mole suggested that the hetero-octamer was stable in solution with 32% of the solvent accessible area buried upon formation of the hetero-octamer. PISA suggested that this organisation of the tetramer in the absence of HicA was unlikely to be stable in solution. This implies that the symmetrical tetramer of HicB is a consequence of HicA binding, rather than a conformation adopted by free HicB alone in solution.

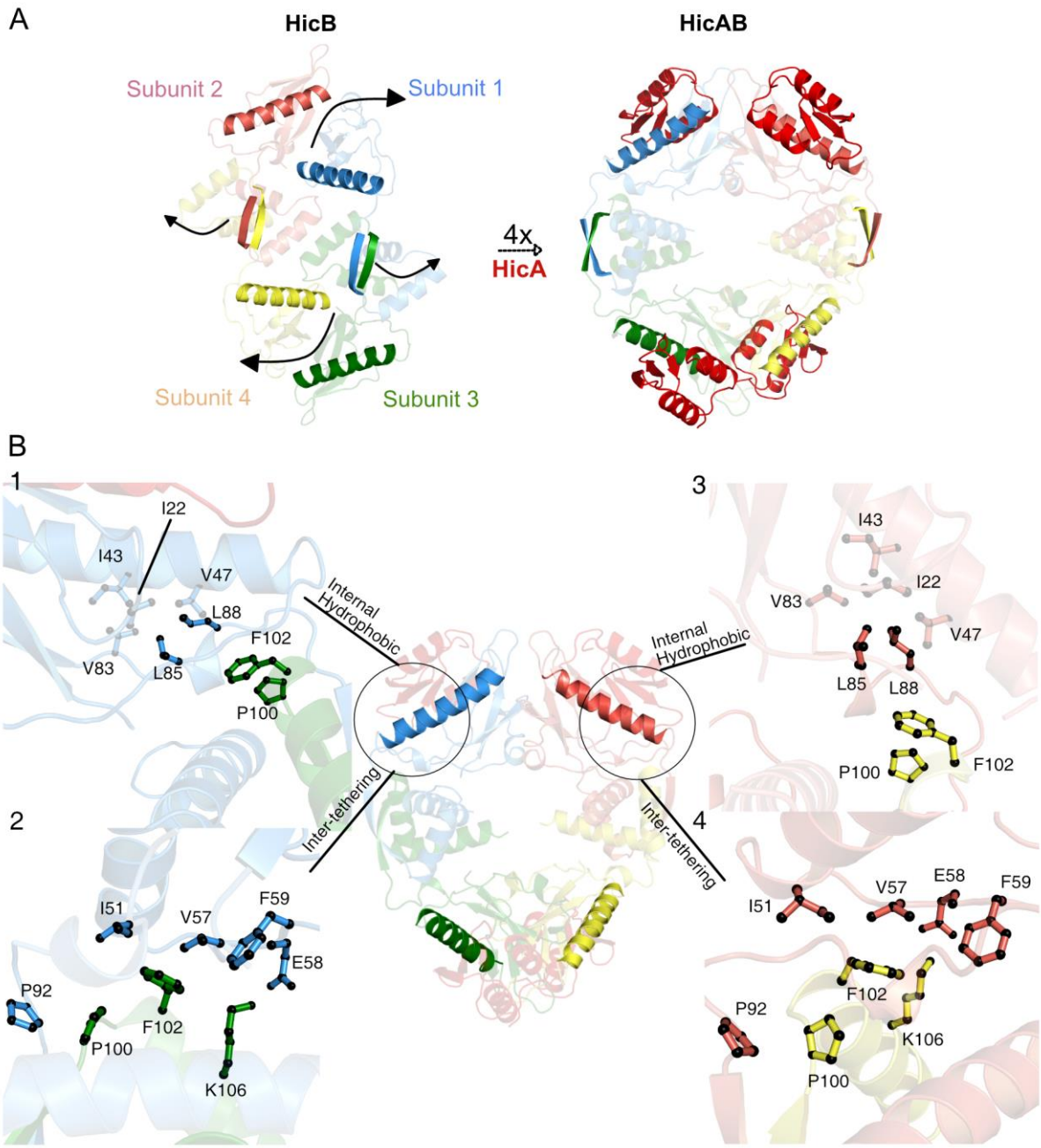


Figure 3.4. Key interactions in the HicAB hetero-octameric complex. (A) Conformation rearrangements induced by HicA binding and forming the complex HicAB. (left) HicA (red) interacts with the α1 helix of each subunit of HicB (highlighted blue, pink, green and yellow) in the unbound conformation to form the hetero-octameric HicAB complex (right), with β5 strands of HicB also highlighted to illustrate their rotation upon complexation. (B) Cartoon representation of new interaction sites across the HicB tetramer. Box 1 highlights the new intra-subunit hydrophobic network formed due to the rotation of L85 and L88 to interact with I22, I43, V47 and V83 within subunit 1 of HicB (blue). Box 2 indicates new inter-subunit interactions between adjacent subunits (1 and 3) of HicB (blue and green). P100 and F102 form hydrophobic contacts to I51, V57, F59 and P92, while E58 interacts with K106 (For clarity the D89-R101 interaction is not shown). Due to the symmetrical nature of HicB within the hetero-octamer, these interactions sites are conserved between subunits of HicB. This is highlighted for HicB subunit 2 (box 3 and 4).

Analysis of the surface charge distribution for HicB revealed a clustering of positive charge on each of the RHH C-terminal domains (Figure 3.5). In the free form the C-domains are oriented in such a way as these clusters align to form an extended positively charged patch. When viewed from a 90° rotated angle this patch can be seen to encompass an obvious groove at its centre. In the HicAB structure however, these domains reorient to move these clusters of charges to opposite sides of the HicAB complex, breaking the positively charged patch apart (Figure 3.5).

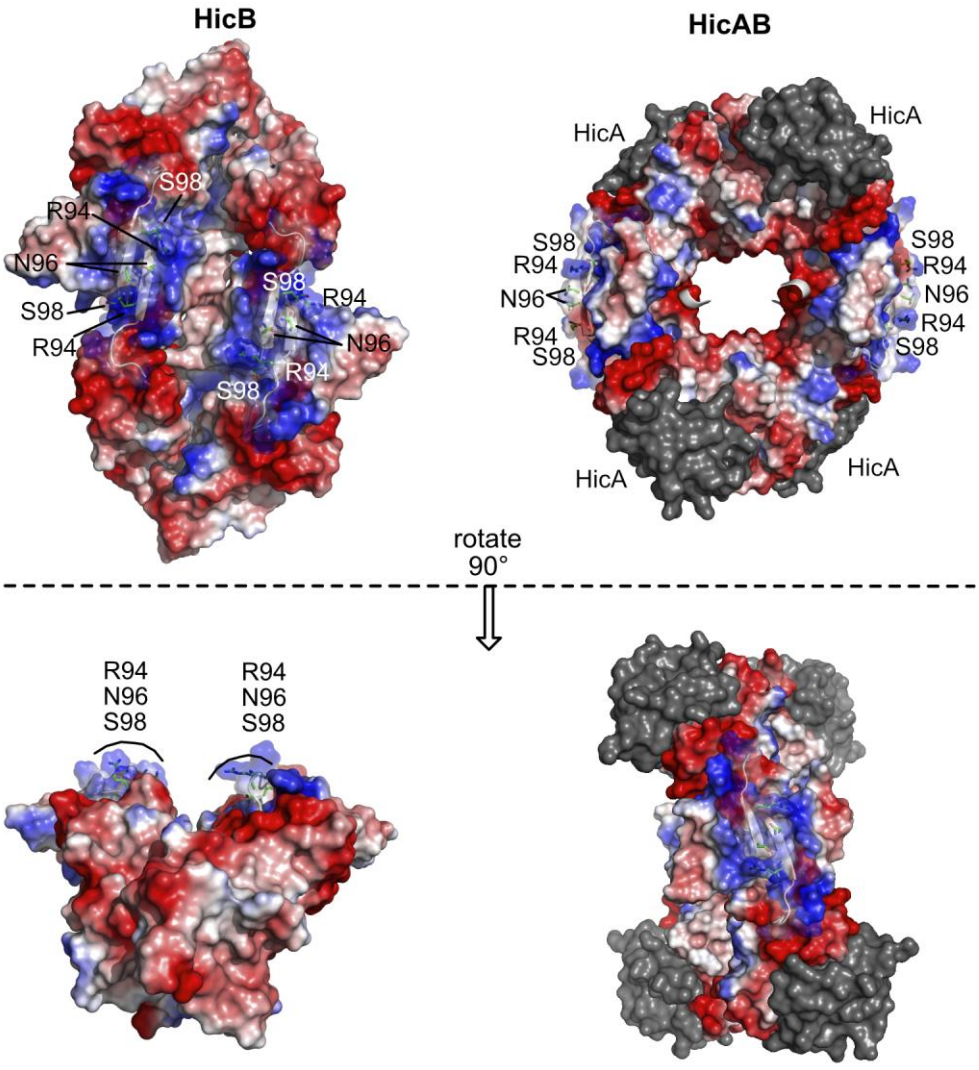


Figure 3.5. Surface charge distribution of HicA and HicAB. (Top): Surface representation of HicB showing clustering of positive charges on one face mapped to R94, N96 and S98 of the C-terminal domain of HicB. (Bottom): Surface representation of HicAB highlighting perturbation of the positively charged patch of the C-terminal domain (R94, N96 and S98) due to rotation of the two RHH domains. HicA is represented in grey.

HicA retains its dsRBD-like fold (RMSD: 1.22 Å, Appendix Figure 14) and binds the α1 helix of HicB via the positively charged β-sheet and α2 helix (Figure 3.6). Several electrostatic interactions R19-E48, R19-E52, K28-E55 tether the β-sheet (HicA) to the C-terminus of the α1 helix (HicB) (Figure 3.6A). This electrostatic interface is supplemented by hydrogen bonds (S22-W28 and T37-H46). P39 and P41 within the β2-β3 loop of HicA stack against the β3 strand of HicB (W28) via CH-π interactions[332] and further hydrophobic residues of the HicA β-sheet interact with the a1 helix of HicB (V18-L53, A23-A42, L35-L50). HicA forms one tethering interaction outside of this region between the α2 helix (K55) and the α5 helix (E72) of HicB. This interaction site buries 900 Å^2 of the total surface area (13,500 Å^2) and the interactions between the two proteins are summarised in Appendix Table 13. The symmetry of the hetero-octamer conserved this interaction site at each HicB subunit. PISA determined that the HicAB interface was only structurally similar to other HicAB family members (Appendix Table 14) [309] and there was no observed similarity to other members of the PDB (July 2018).

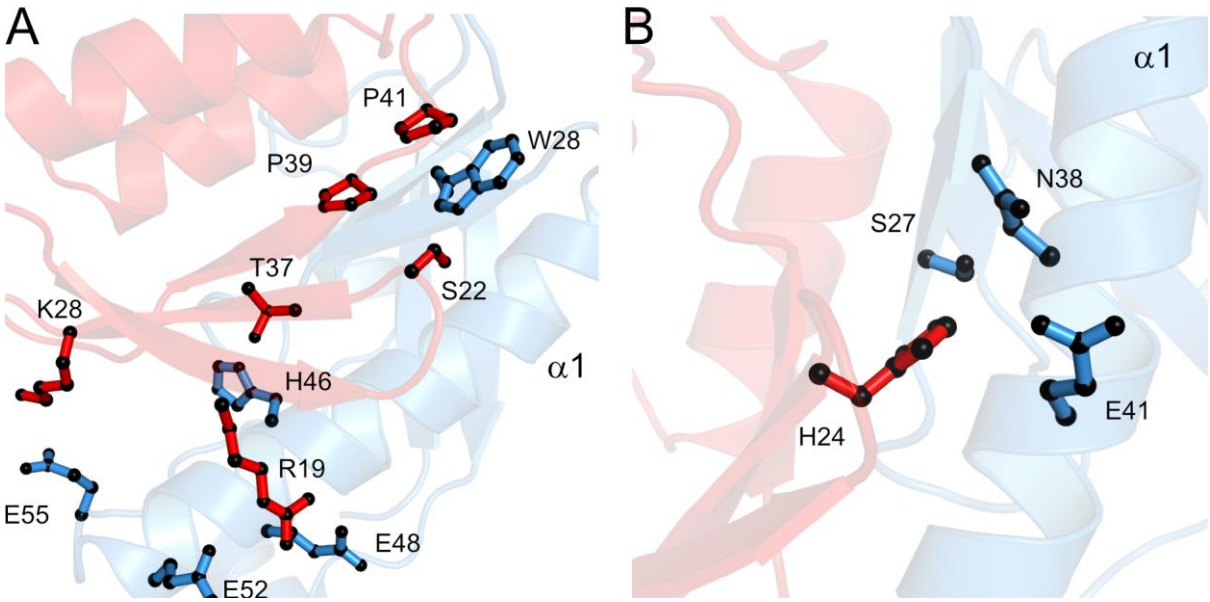


Figure 3.6. Interaction site between HicA and HicB. (A) HicA (red) interacts with the α1 helix of HicB primarily through the three-stranded β-sheet and connection loops bearing hydrophobic, hydrogen and electrostatic interactions. (B) The functionally critical residue H24 (H24A in the crystal structure) is predicted to project into the polar pocket of HicB formed by S27, N38 and E41. Key residues rendered as sticks.

3.3.1. Sequestration of HicA functionally relevant residues

Within canonical dsRBD proteins, there are three distinct sites of interaction which form interactions with the minor groove of RNA (α1 helix and the β1-β2 loop) and the major groove (α2). [333] HicA does not contain the canonical motifs associated with these sites (E8, a GpxH motif and a KKxAK motif) within the α1, β1-β2 loop and the α2 helix.[333] Instead HicA bears a ^{22}GSH^{24} motif which has been verified as the site of endoribonuclease activity for *Y. pestis* and *S. pneumoniae* HicA although not yet confirmed in *B. pseudomallei*. The motif is present in other HicA family members (Appendix Figure 15) with the exception of a GSNH motif in *E. coli* HicA. E8 appears to be substituted by E7 within *B. pseudomallei, E. coli* and *Y. pestis* HicA but is not present within *S. pneumoniae* HicA. The KKxAK motif within the α2 helix of canonical dsRBD proteins (Appendix figure 15) may be equivalent to the ^{51}R/KxxxK^{55} motif at the α2 helix, found within these HicA family members. Residue K51 within *Burkholderia* HicA is typically conserved as either an arginine/lysine at this position while residue K55 within is conserved throughout HicA family members.[131]

The HicA-HicB interactions identified here bury the functionally critical residues (G22 and H24) within the β1-β2 loop (Figure 3.6B). A similar mechanism of complexation has been observed in the homologous dsRBD (αββββα) HicA3B3 and HicAB$_{SP}$ systems for *Y. pestis* and *S. pneumonaie* (Figure 3.7, Figure 3.8). Superimposition of the HicAB interaction site to HicA3B3-NT revealed a conserved mechanism of interaction: (Figure 3.7A, RMSD: 3.58 Å over the carbon backbone). Functional critical residues for HicA (G22 and H24) share equivalent positions with those of HicA3 (G26 and H28). Both H24 and H28 are buried by the conserved polar pocket of N38/N37 and E41/E40 of HicB/HicB3 (Figure 3.7B). Several electrostatic interactions within HicAB: R19-E48, K28-E55 and K55-E72 are conserved within HicA3B3 as R21-E46, R18-E51 and K51-E75 (Figure 3.7C, Appendix Table 15). Several hydrophobic and hydrogen bond interactions (L35-L50, L35-L50 and T37-H46) are also conserved (Figure 3.7D, E).

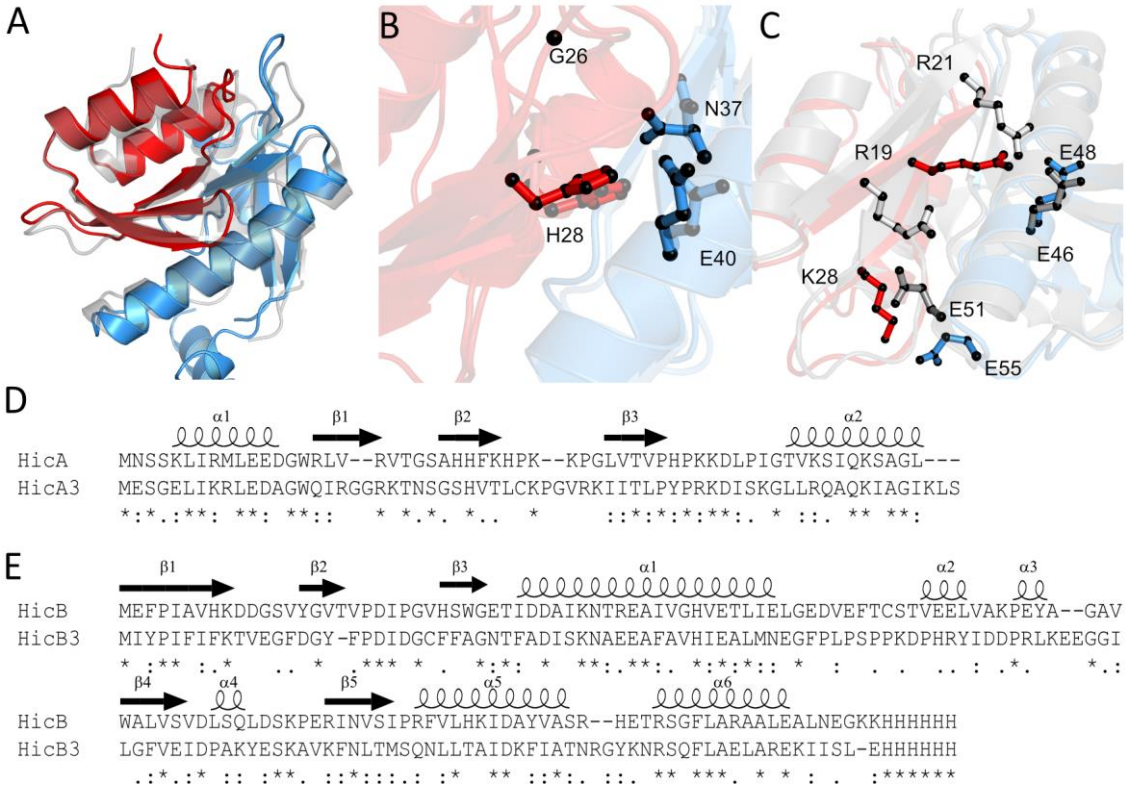


Figure 3.7. Comparison of the HicAB crystal structure to HicA3B3. (A) Superimposition of HicA (red) HicB (blue) to HicA3 (Light grey) HicB3 (Dark grey) (PDB:4P78). (B) Sequestration of G26 and H28 in HicA3 (light grey) HicB3 (dark grey) by the polar pocket N37 and E40. Sequestration of G22 and H24 in HicA (red) by the polar pocket of N38 and E41 of HicB (blue) is highlighted and transparent for comparison. (C) Conservation of electrostatic residues seen at the interaction site of $HicA_1$ (red) $HicB_1$ (blue) and $HicA3_1$ (Light grey) $HicB3_1$ (Dark grey). (D) Sequence of *B. pseudomallei* HicA and comparison to the closest structural homologue, HicA3 from *Y. pestis*. (E) Sequence of *B. pseudomallei* HicB and comparison to the closest structural homologue, HicB3 from *Y. pestis*. Symbols indicate a conserved residue (*), conservative mutation (:) and a semi-conservative mutation (.).

$HicA_{SP}$ also binds in a similar conformation to HicA and buries the catalytic H36 in a polar pocket formed by T33 and E47 ($HicB_{SP}$) (Figure 3.8A, B) and supplementary electrostatic and hydrophobic interactions between $HicA_{SP}$ and $HicB_{SP}$ are conserved (Figure 3.8 C, Appendix Table 16, Appendix-Figure 16). The $HicAB_{SP}$ interaction site is roughly conserved (85%) at each $HicB_{SP}$ subunit despite the difference in subunit orientation (Appendix Table 16). In comparison to *B. pseudomallei* HicAB, multiple interactions tether the α2 helix of $HicA_{SP}$ to $HicB_{SP}$ via the β1-β2 loop, α2 helix and α2-β4 loop through a series of hydrophobic, hydrogen and electrostatic interactions that are not conserved within *B. pseudomallei* HicAB (Appendix Table 16). Only one interaction is observed between the α2 helix of HicA and HicB (K55-E72), whereas $HicAB_{SP}$ forms fifteen interactions at this site. The α2 helix seems to form interaction sites in place of the β1 strand of $HicB_{SP}$ which is

much shorter (^{27}IK^{28}) than *Burkholderia* HicA (^{16}RLVR^{19}) and does not form any interactions with $HicB_{SP}$ whereas HicAB forms three key interactions at this site. As these sites are not present, the $HicAB_{SP}$ interaction site is laterally shifted when compared to HicAB with respect to the placement of the secondary structure elements, positioning α2 in an optimum position to form multiple contacts with $HicB_{SP}$.

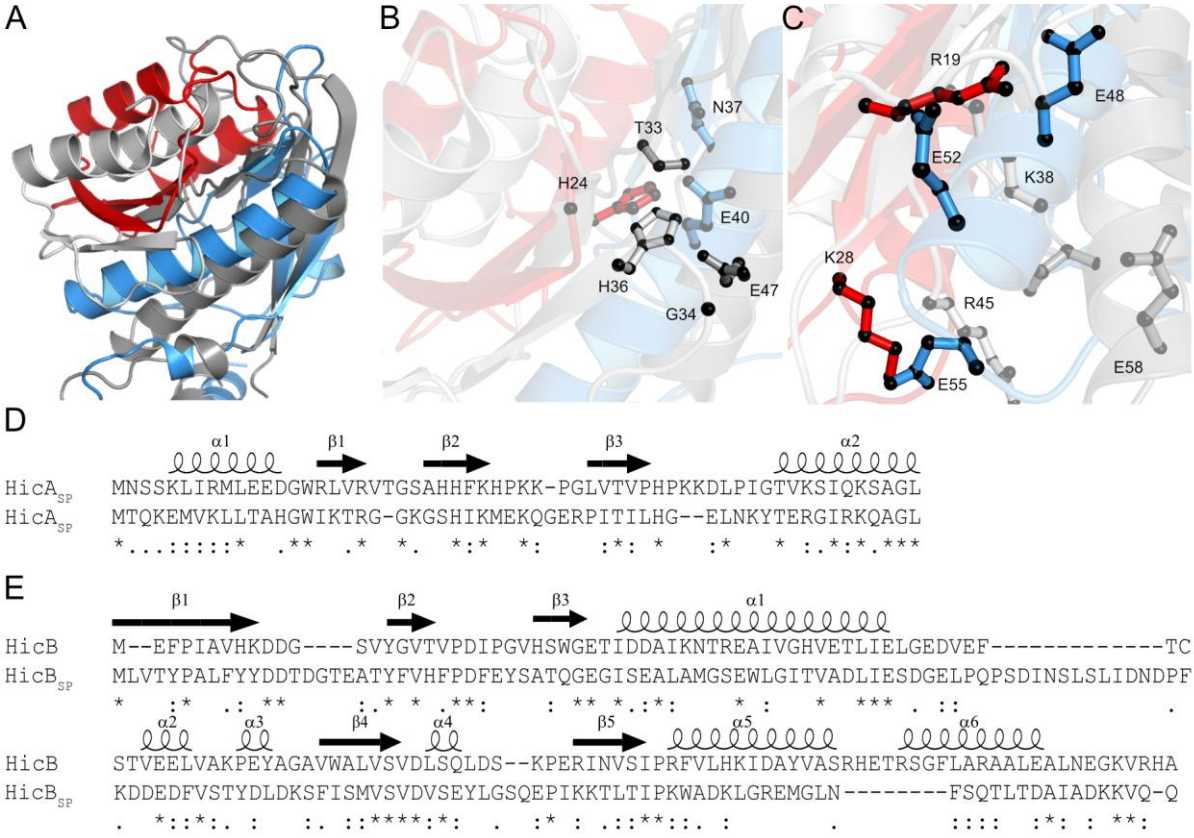


Figure 3.8. Comparison of the HicAB and $HicAB_{SP}$ interaction site. (A) Superimposition of HicA (red)-HicB (blue) to $HicA_{SP}$ (Light grey)-$HicB_{SP}$ (Dark grey) (PDB:5YRZ). (B) Sequestration of G34 and H36 in $HicA_{SP}$ by a polar pocket of $HicB_{SP}$ (T33 and E47). Sequestration of HicA G22/H24 is also shown. For clarity HicA/HicB are transparent for comparison. (C) Conservation of electrostatic residues seen at the interaction site of HicAB and $HicAB_{SP}$. K38-D55/E58, R45-D55 and K56-D12/D15 are equivalent within HicAB (R19-E48/E52, K28-E55 and K43-E31). (D) Sequence of *B. pseudomallei* HicA and comparison to the structural homologue $HicA_{SP}$ from *S. pneumoniae.* (E) Sequence of *B. pseudomallei* HicB and comparison to the structural homologue $HicB_{SP}$ from *S. pneumoniae.* Symbols indicate a conserved residue (*), conservative mutation (:) and a semi-conservative mutation (.).

Further alignment of *Burkholderia* HicB to 51 cognate antitoxins to toxins that bear a high homology to HicA (Butt 2014, Appendix Figure 17) revealed a high degree of conservation within the α1 helix of HicB members (Figure 3.9). Residues involved in the HicAB interaction site (N38, E41, A42, H46, E48, L50 and E52) are highly conserved. The α1 helix is therefore likely to be the main interaction site for HicA within HicB family members due to its role in burying the functionally critical histidine for each studied system to date.

Figure 3.9. A Web logo[334] generated graphic depicting the conservation of amino acids at the a1 helix of HicB when aligned to 51 homologous HicB sequences using Clustal Omega[291]. The amino acid sequence of the α1 helix of HicB is highlighted underneath. Residues of HicB that form interactions with HicA are underlined and show strong conservation among HicB family members.

Several other type 2 toxins also degrade mRNA and these toxins are classified as either translation independent or dependent endoribonucleases with respect to their mechanism of action. Translation independent endoribonucleases, which includes HicA[187], MazF[145] and VapC[182] (Figure 3.10A) appear to neutralise in a similar mechanism where the formation of a toxin-antitoxin complex directly occludes the functional residues of each toxin. Despite the similarity in neutralisation within translation independent endoribonucleases, HicA only displays structural homology to other HicA family members and not to non-HicA type 2 toxins (Appendix Table 17).

The mechanism of direct occlusion of functional residues does not apply to all type 2 TA systems, this is not surprising as there is considerable structural variation between different TA members.[112, 335, 336] The majority of translation dependent toxins (Figure 3.10B), with the exception of YafQ which displays direct occlusion of functional residues[83], are neutralised by other mechanisms, which prevent ribosome association as observed in HigB[165] or disrupt the active site via helix displacement (RelB[155] and YefM[160]).

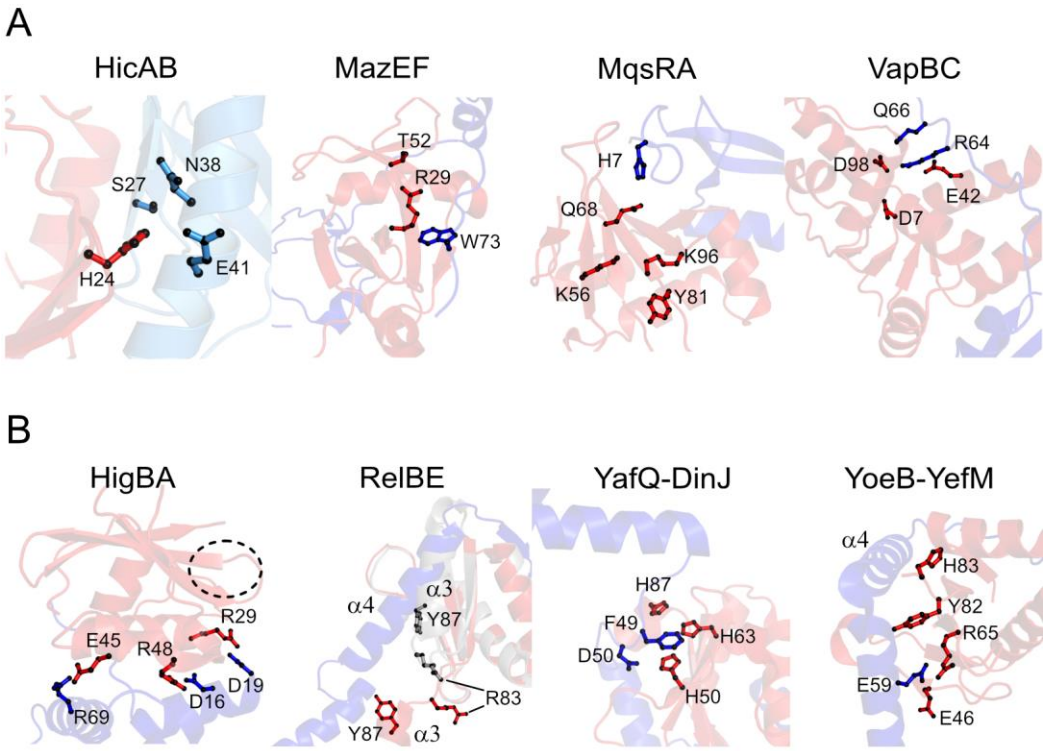


Figure 3.10. Neutralisation of the toxin catalytic site by their cognate antitoxins. (A) Neutralisation of translation independent toxins HicA (4C26), MazF (1UB4), MqsR (3HI2) and VapC (3TND). For each complex, the toxin and antitoxin are highlighted red and blue respectively. Note for VapC (3TND) the toxin cleaves tRNA not mRNA. (B) Neutralisation of translation dependent toxins HigB (4MCT), RelE (4FXE), YafQ (4Q2U) and YoeB (2A6Q). For HigB, the catalytic active site is circled. Within RelB, the unbound state is highlighted in grey to show α3 helix displacement. For each complex, residues that constitute the catatlytic site of the toxin (red) and the residues that neutralise them within the antitoxin (blue) are highlighted as sticks.

3.3.2. Structural comparisons to HicA3B3

The HicA3B3 complex ($HicA3_2$-HicB3-NT_2) is an incomplete structure (PDB: 4P78) due to the removal of the C-terminal domains of HicB3[132]. The resultant HicA3B3-NT structure exists as a hetero-tetramer with a 1:1 stoichiometry also observed for the *Burkholderia* HicAB complex (Figure 3.11). HicA3 displayed a dsRBD fold similar to HicA (RMSD: 2.2 Å, Appendix Figure 14).

Superimposition of the HicA3B3-NT heterotetramer to the full length HicB3 structure suggested that the orientation of HicB3 N-terminal domains in the heterotetramer were identical to the subunit orientation seen within the HicB3 tetramer (Figure 3.11). Steric clashes were seen at two interaction sites between HicA3 molecules and the C-terminal domains of HicB3 (Subunit 1 and 4) in a $HicA3_4HicB3_4$ model (Figure 3.11). Due to the presence of steric clashes, it was assumed that a 1:1 stoichiometry ($HicA3_4HicB3_4$ complex)

could not be adopted by the full length HicB3. The model of a $HicA3_4HicB3_4$ complex was discarded as biophysical evidence (SEC-MALS) indicated that a single complex ($HicA3_2HicB3_4$) was formed rather than a $HicA3_4HicB3_4$ complex. Within HicB3, two α1 helices (Subunits 2 and 3) are solvent exposed in a similar manner to HicB and it is presumed that HicA3 binds at these sites to form a $HicA3_2HicB3_4$ complex.[132]

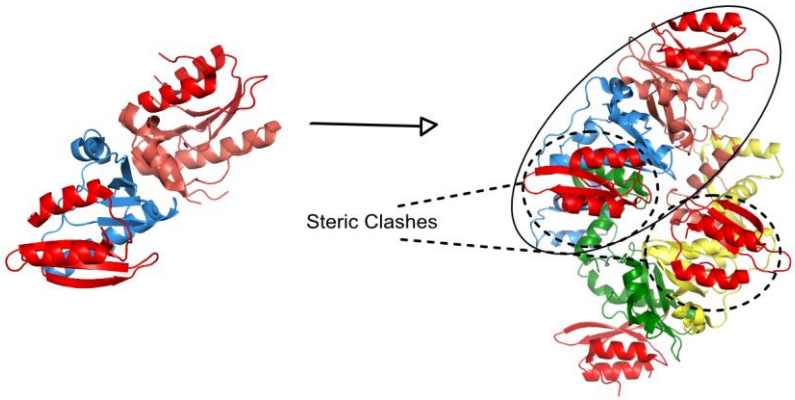


Figure 3.11. Steric clashes in a $HicA3_4HicB3_4$ complex. Extension of the $HicA3_2HicB3_2$ structure with C-terminal domains modelled. This revealed steric clashes between HicA3 monomers and the C-terminal domains of subunit 1 (blue) and 4 (yellow) of HicB3 as the N-terminal domains of HicB3 within HicA3B3 are not correctly rearranged as observed in HicAB.

Within the HicAB crystal structure, HicA binding caused a lateral displacement of α1 helices within HicB subunits 2 and 3 resulting in a 90 ° rotation of the C-terminal domains, actively preventing steric clashes in a HicAB complex (1:1 stoichiometry) through the formation of a symmetrical HicB tetramer through extensive domain reorganisation. Steric clashes observed in the $HicA3_4HicB3_4$ model at the C-terminal domain may be prevented by the conversion of HicB3 to a symmetrical tetramer similar to HicAB. Given the structural similarities between the proteins, this mechanism is quite probable. Presumably Bibi-Triki and colleagues could not appropriately model a $HicA3_4HicB3_4$ complex was because they had not accounted for rotation of the C-terminal domains or were not aware that this could occur.

3.3.3. Structural comparison of *B. pseudomallei* HicAB and *S. pneumoniae* HicAB

Given the steric clashes expected in a hetero-octamer that would arise without re-arrangement of the HicB tetramer, it was expected that a similar conformational change would have been reported for *S. pneumoniae* HicAB ($HicAB_{SP}$).[272] The full structure of HicAB, however, shares some but importantly not all structural characteristics with the $HicAB_{SP}$ complex (Figure 3.12A, B). $HicAB_{SP}$ is a hetero-octamer and the $HicB_{SP}$ subunits contain two domains, an N-terminal domain ($^{I}\beta1\beta2\beta3\alpha1\alpha2\beta4^{103}$) and a C-terminal domain

(^{114}β5α6α7^{147}) separated by a hinge region (V104-I114) encompassing a α5 helix. Comparison of the HicAB to $HicAB_{SP}$ reveal the structural similarities over the Cα backbone, particularly evident within HicA (RMSD = 1.5 Å, Appendix Figure 14), including a shortened β1 strand in $HicA_{SP}$. $HicB_{SP}$ lacks a α3 helix and instead displays a large loop region between the α2 helix and β4 strand (S73-Y90) (Figure 3.12C). The authors noted that no interactions were made between the C-terminal domain of $HicB_{SP}$ and the rest of the $HicAB_{SP}$ complex, in contrast to the tethering interactions observed within HicB. Despite a similar organisation between individual subunits of HicB (RMSD: 2.8 Å over the Cα backbone) the structural organisation of the $HicB_{SP}$ tetramer within $HicB_{SP}$ resembles HicB in an *unbound* form (RMSD = 63 Å, Figure 3.12D), rather than the symmetrical conformation observed in the HicAB hetero-octamer. $HicA_{SP}$ binds at two subunits (Subunit 1, blue and Subunit 4, yellow) whose α1 helices project towards the C-terminal domains but no steric clashes are observed at these sites, due to the long hinge region between the N-terminal and C-terminal domain that can actually accommodate two $HicA_{SP}$ molecules.

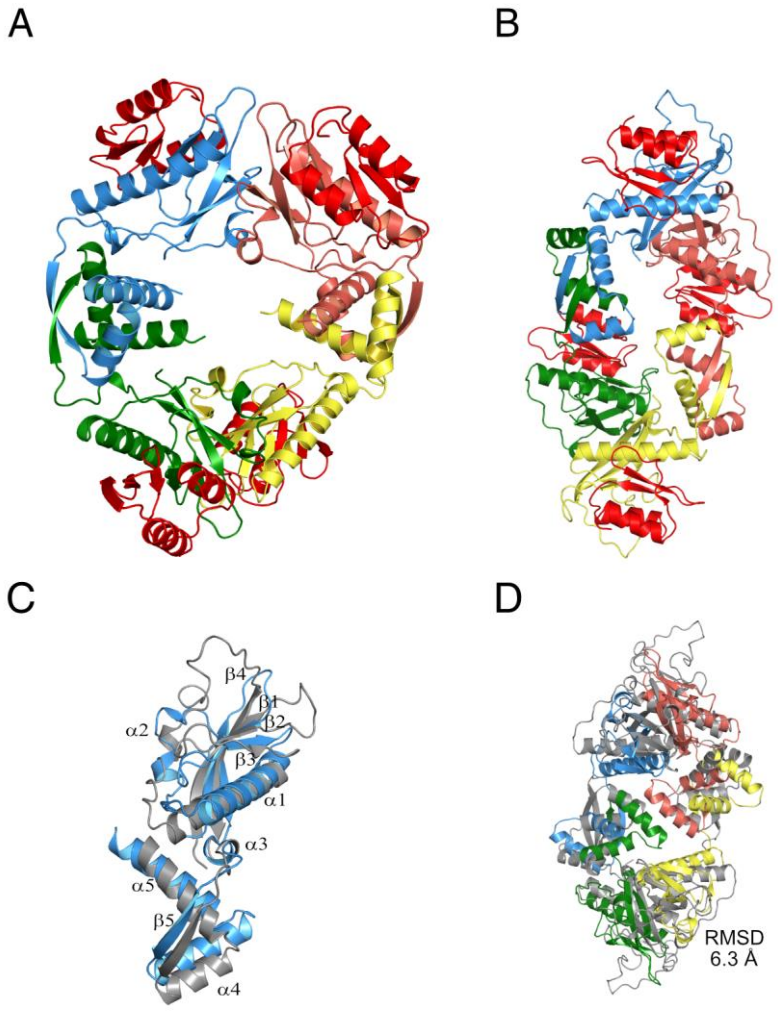


Figure 3.12. Structural Comparison of HicAB and $HicAB_{SP}$. (A) *B.pseudomallei* hetero-octamer of HicAB where HicB subunits are arranged in a symmetrical manner. (B) *S. pneumoniae* HicAB hetero-octamer (PDB: 5YRZ) where C-terminal domains lie parallel to each other. For both structures HicB subunits are highlighted in blue/light red/ green/ yellow and HicA subunits in red. (C) Superimposition of $HicB_{SP}$ (dark grey) to HicB (blue) over the carbon backbone (RMSD: 2.8 Å). (D) Superimposition of the $HicB_{SP}$ tetramer (dark grey) to the elongated conformation of HicB over the Cα backbone (RMSD: 6.3 Å).

The C-terminal strand-swapped dimers of $HicB_{SP}$ within $HicAB_{SP,}$ are organised parallel to each other, rather than facing inwards as observed in HicB, within HicAB (Figure 3.13). Inspection of the surface exposed charge suggests that a concerted patch of positive charge at the C-terminal domain was visible despite complexation with $HicA_{SP}$ (Figure 3.13A). Within HicB, this patch had been disrupted due to C-terminal domain rotation upon HicA binding (Figure 3.13B). As the structure of $HicB_{SP}$ in the unbound form is unknown, it is not possible to determine precisely which rearrangements have occurred as a consequence of $HicA_{SP}$ binding.

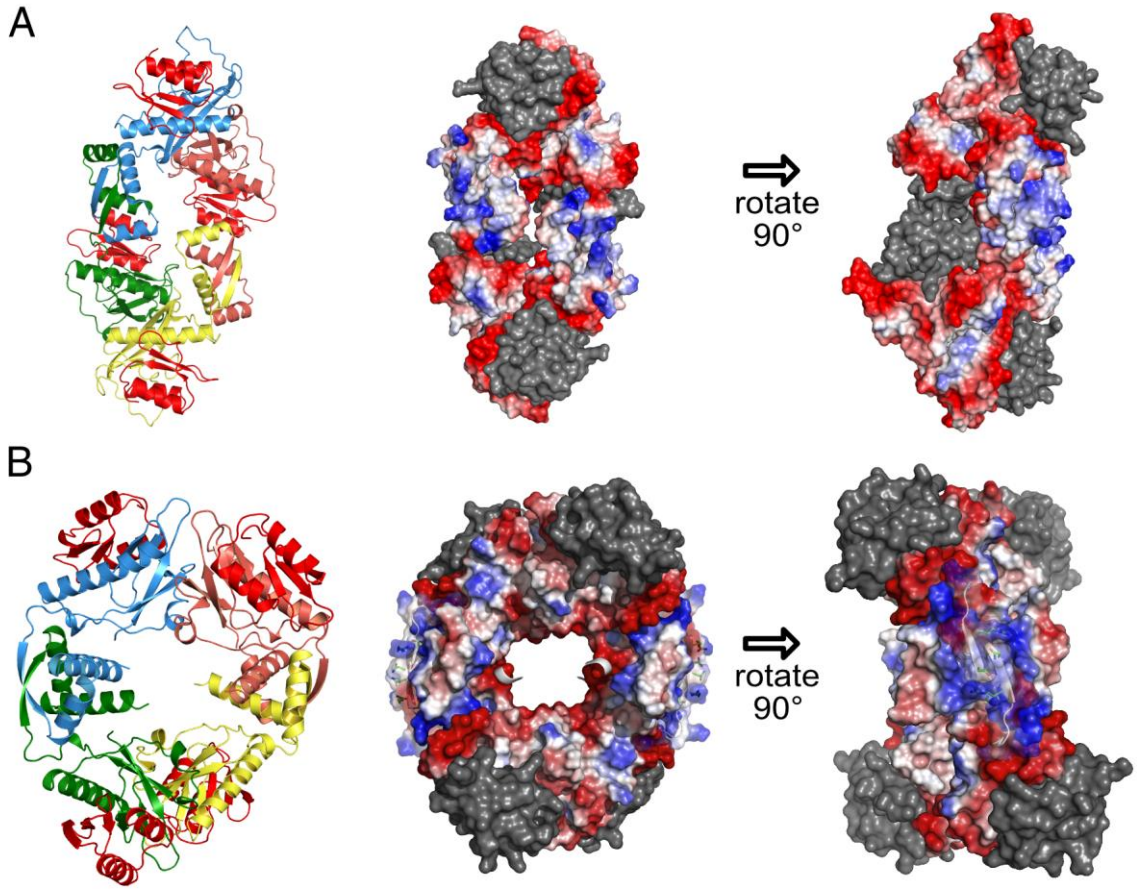


Figure 3.13. Charge distribution of $HicAB_{SP}$ and HicAB. (A) Cartoon and surface representation of $HicAB_{SP}$ showing clustering of positive charge on one face mapped to the C-terminal RHH domains of $HicB_{SP}$. (B) Cartoon and surface representation of HicAB highlighting the positively charged patch mapped to the C-terminal RHH domain HicAB. For both images, regions of positive, negative and no charge are coloured blue, red and white respectively. $HicA_{SP}$/HicA is represented as grey to emphasize the surface charge of $HicB_{SP}$/HicB.

3.4. Biophysical techniques

To further probe the $HicA_4HicB_4$ stoichiometry observed by crystallography, HicAB (preformed at a 1:1 ratio of HicA:HicB, HicA (500 μM): HicB (500 μM)) was analysed by SAXS (Figure 3.14, Appendix Figure 18, Appendix Table 18). The bulk of the resultant *ab initio* shape envelope appeared to be derived from the symmetrical conformation of HicB (Figure 3.14A, χ^2=2.89) rather than the unbound conformation (Figure 3.13B, χ^2=8.94) (Figure 2.20). This suggested the symmetrical conformation of HicAB is not simply a crystal artefact and formation of the symmetrical HicB tetramer can be observed in solution.

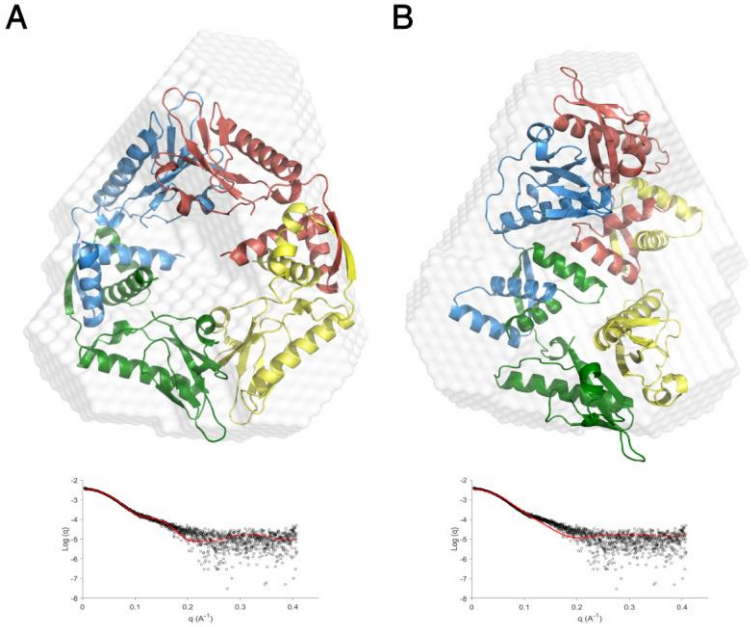


Figure 3.14. SAXS shape envelope of HicAB. (A) *Ab initio* modelling of the heterotetrameric HicB component of the HicAB complex into the shape envelope of HicAB (white). The FoXS profile of the proposed scattering profile for a HicAB heterotetramer (red) against the experimental raw scattering data (black) (χ^2=2.89). (B) *Ab initio* modelling of the HicB tetramer into the shape envelope of HicAB (white). The FoXS profile of the proposed scattering profile for tetrameric HicB (red) against the experimental raw scattering data (black) (χ^2=8.94).

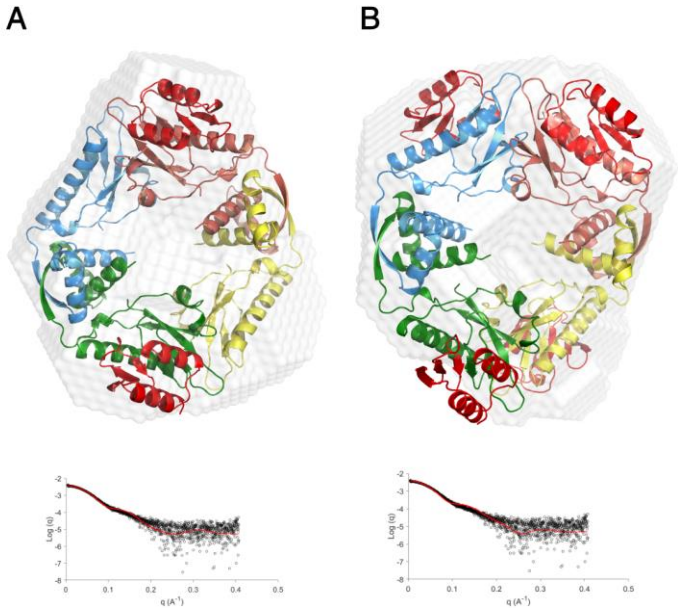


Figure 3.15. Distinguishing between complexes of HicAB. (A) *Ab initio* modelling of the $HicA_2HicB_4$ complex into the shape envelope of HicAB (white). The FoXS profile of the proposed scattering profile for the $HicA_2HicB_4$ heterotetramer (red) against the experimental raw scattering data (black) (χ^2=2.45). (B) *Ab initio* modelling of the $HicA_4HicB_4$ complex into the shape envelope of HicAB (white). The FoXS profile of the proposed scattering profile for tetrameric $HicA_4HicB_4$ (red) against the experimental raw scattering data (black) (χ^2=8.23).

SAXS shape envelope suggested that the $HicA_2HicB_4$ complex formed by removing two HicA molecules from the HicAB crystal structure (Figure 3.15A, χ^2=2.45) was formed over the $HicA_4HicB_4$ complex (Figure 3.15B, χ^2=8.23). Attempts to fit the $HicA_4HicB_4$ complex were poor due to a single HicA molecule failing to fit within the shape envelope, despite the apparent fit of a structure where all α1 helices are surface exposed. This resulted in the high observed χ^2 (8.23) for this complex. The $HicAB_{SP}$ conformation did not agree with the experimental SAXS data (Appendix Figure 19).

SVAUC indicated the presence of a single oligomeric species for both HicB (Figure 3.16A, Materials and Methods, equation (2-6)) and HicAB (preformed at a HicA: HicB ratio = 0.5:1) (Figure 3.16B) with calculated molecular weights of 62.4 kDa and 68 kDa equivalent to 0.99 x tetrameric HicB (62.9 kDa) and 0.97 x $HicA_1HicB_4$ (70 kDa). As SVAUC reports on an inferred molecular weight based off experimentally derived Sedimentation and Diffusion values [337-339], SEAUC was performed for HicAB (preformed at a HicA: HicB ratio = 0.5:1, Materials and Methods) as this directly reports on the molecular mass (Figure 3.16C).[340, 341] Modelling a single species of HicAB (equation (7)), inferred by SVAUC, reported a calculated mass of 77.6 kDa (95% confidence interval + 1.3 – 1.4 kDa) equivalent to 1.07 x $HicA_2HicB_4$ (77 kDa).

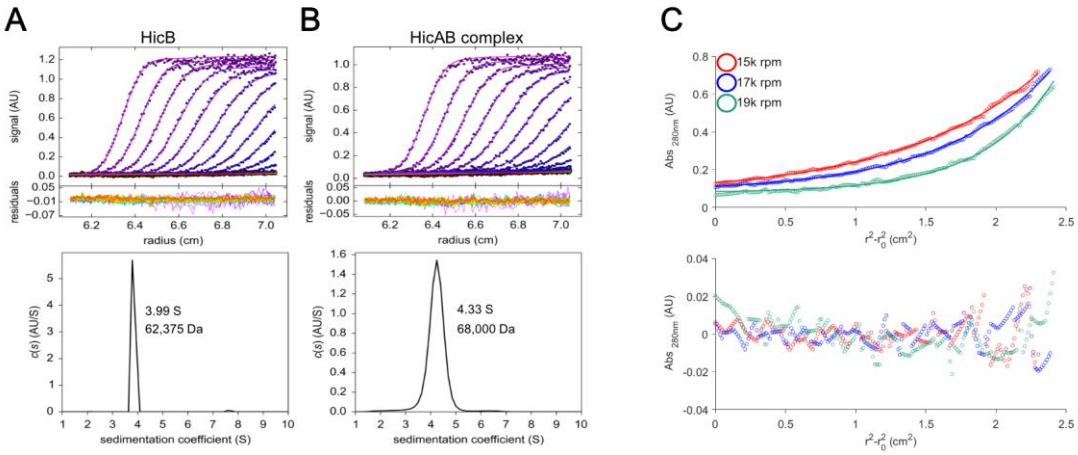


Figure 3.16. AUC of HicB and HicAB. (A) SV AUC for HicB (70 μM) and $HicA_2HicB_4$ (50 μM) preformed at a HicA: HicB ratio of 0.5:1. Experimental scans (symbols) and best fit models (solid lines) for HicB at 50,000 rpm, 20 °C. For clarity only every 4th scan is shown. The lower panel shows best fit residuals of the plotted scans. (Bottom) Continuous size (C(s)) distribution fit for HicB.[338, 341] The peak S value in this fit is 3.99 S with a frictional ratio of 1.31 and a molecular weight of 62,375 Da. (B) Experimental scans (symbols) and best fit models (solid lines) for HicAB at 50,000 rpm, 20 °C. For clarity only every 4th scan is shown. The lower panel shows best fit residuals of the plotted scans. (Bottom) Continuous size (C(s)) distribution fit for HicAB. The peak S value in this fit is 4.33 S with a frictional ratio of 1.286 and a molecular weight of 68,000 Da. The RMSD for both for HicB and for HicAB was 0.009. (C) SEAUC for $HicA_2HicB_4$ (20 μM), preformed at a HicA:HicB ratio of 0.5:1**.** The experimental data (top, open circles) and fits (lines) to a single ideal species (equation (7), residuals (bottom) between the experimental and filled data points are seen for 15k rpm (red), 17k rpm (blue) and 19k rpm (green). Data was recorded at 20 °C, 280 nm.

The $HicA_4HicB_4$ hetero-octamer was observed by native mass spectrometry (Materials and Methods, Figure 3.17)[342, 343], as well as free HicB components and sub-stoichiometric complexes of HicAB ($HicB_2$, $HicB_4$, $HicA_1$-$HicB_4$ and $HicA_2$-$HicB_4$).

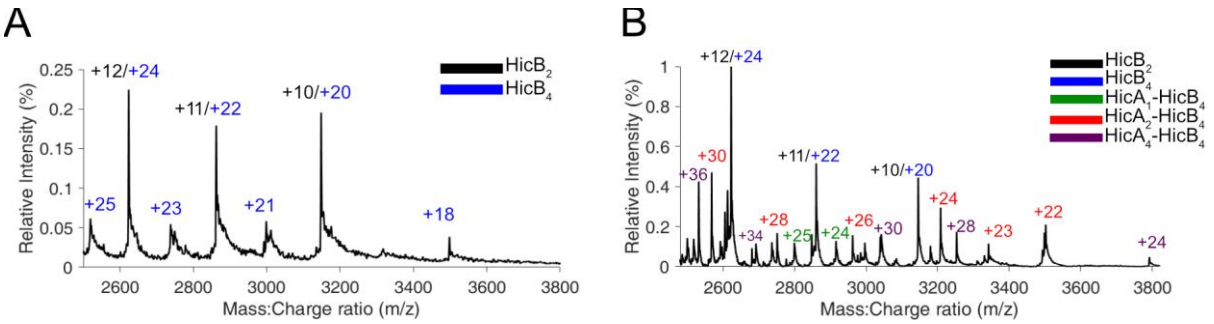


Figure 3.17. Native mass spectrometry of $HicA_4HicB_4$. Native mass spectrometry of HicB within the region 2500-3800 m/z with oligomeric states: dimeric (black) and tetrameric (blue) highlighted based on their charge state value. (D) Native mass spectrometry of HicAB (preformed at a ratio of HicA: HicB = 1:1) within the region 2450-3850 m/z with oligomeric states of $HicB_2$ (black), $HicB_4$ (blue) $HicA_1$-$HicB_4$ (green), $HicA_2$-$HicB_4$ (red) and $HicA_4$-$HicB_4$ (purple) highlighted.

In agreement with SEC-MALS data for HicA3B3 reported by Bibi-Triki ($HicA3_2HicB3_4$)[132], SEAUC confirms that *Burkholderia* HicAB forms a $HicA_2$-$HicB_4$ complex with SEC-SAXS suggesting that HicB adopts a symmetrical conformation within this complex, where each α1 helix is surface exposed. The $HicA_4$-$HicB_4$ hetero-octamer is observed within native mass spectrometry suggesting it is a significant conformation and not simply a crystal artefact. It is hypothesized that initial binding of HicA forms a $HicA_2HicB_4$ complex, further exposing two α1 helices (Subunits 1 and 4). Exposure of these sites ensures HicB can fully neutralise HicA at equimolar concentrations. However, in vivo it is unlikely to be as simple as this. HicB is flexible and can adopt several conformations in the absence of DNA. The dual ability to bind both DNA and HicA invokes further complexity to the system.

3.5. Summary

Formation of the symmetrical HicAB complex, requires the reorganisation of tethering interactions between adjacent HicB subunits. The electrostatic interactions that tethered the α1 helix of subunit 1 and 4 to the adjacent C-terminal domain within the unbound form are broken due to the interaction with HicA due to subunit 1 and 4 flipping their α1 helix to an exposed orientation. The consequence of HicA interaction is the rotation of the C-terminal domains to prevent steric clashes between HicA and HicB subunits. Steric clashes had previously been noted between *Yersinia* HicA3 and HicB3 in a proposed $HicA3_4HicB3_4$ complex. As tethering interactions and the orientation of α1 helices within the HicB/HicB3 tetramer are conserved, it is hypothesized that this phenomenon of C-terminal rotation upon toxin complexation might be a general feature of HicAB yet surprisingly the *Streptococcus* HicAB complex did not capture this and a complex resembling a compact conformation was reported with HicA being accommodated by longer connecting loops between the N- and C-domains. However, the absence of tethering interactions in the crystal structure of HicAB suggests that this form is a snapshot within an array of different HicAB conformations as these free C-terminal domains will flex in solution, possibly to an open state observed for *Burkholderia* HicAB.

	HicAB
Wavelength (Å)	0.9795
Resolution range (Å)	34.02 – 2.49 (2.58-2.49)
Space group	$P2_1$
Unit cell	85.1 72.2 85.3Å 90 90.1 90°
Multiplicity	6.8 (6.8)
Completeness (%)	99.9 (99.8)
Mean I/σ(I)	14.1 (1.4)
Wilson B-factor (Å^2)	70
R_{meas}	0.12 (1.23)
$CC_{1/2}$	0.998 (0.692)
Reflections used in refinement	35696
R_{work}	0.181 (0.330)
R_{free}	0.220 (0.375)
Number of protein atoms	6272
RMS (bonds) (Å)	0.034
RMS (angles) (°)	1.71
Ramachandran favoured (%)	98
Ramachandran allowed (%)	1.5
Ramachandran outliers (%)	0.37
Rotamer outliers (%)	5.7
Average B-factors (Å^2)	
Protein	60
Solvent	50

Table 3.1. Final refinement statistics for HicAB. Parenthesis refer to statistics for the high resolution shell

4. DNA binding of HicB

TA systems are able to bind sequences of DNA within their promoter regions via DNA binding domains and this has been investigated for HicB homologues in *E. coli*, *S. pneumoniae* and *Y. pestis*.[131, 132, 272] However, there is no structural or functional data on DNA binding properties of *B. pseudomallei* HicB and the presence or identity of the DNA binding motif(s) upstream of this *hicAB* operon have not been established. This chapter focuses on the identification of a DNA binding site for HicB upstream of the *hicAB* operon and the structure-function relationships for this protein-DNA interaction.

4.1. Identification of the DNA binding site

A series of oligomers and electrophoretic mobility shift assays (EMSA) were used to determine if HicB bound to the intergenic region (279 bp) immediately upstream of the *hicAB* operon. The intergenic region was initially dissected into 48 bp double stranded fragments to individually investigate the -10 and -35 promoter elements and the sequences encompassing these. Bioinformatic searches for palindromic sequences/inverted repeats revealed that the 0-48 bp fragment was the most promising candidate when compared to the other five fragments analysed (Table 4.1).[344, 345] This fragment contained a likely -10 TATAAT Pribnow box[346] (Table 4.1, Red, underlined) and a stretch of repeated nucleotides ([24]ACACACAA[17]). Where possible a complete potential binding site was maintained to prevent accidental disruption of an intact binding site during fragmentation of the intergenic region.

DNA Segment	Sequence
0-48 5’	5′-GATCGTGATTGGATGTGTATAATTACACACAAGACATTCGGGGGAGCT-3′
0-48 3’	5′-CTAGCACTAACCTACACATATTAATGTGTGTTCTGTAAGCCCCCTCGA-3′
48-96 5’	5′-ATAGGGGCGAAACAATGTGAAAATACGCACGGCTACACAAAACTTGAG-3′
48-96 3’	5′-TATCCCCGCTTTGTTACACTTTTATGCGTGCCGATGTGTTTTGAACAC-3′
96-144 5’	5′-TCGAGGCCGCGCGCGATGCTTCAGTCTTGCCCAGCGGACGGGATAAAA-3′
96-144 3’	5′-AGCTCCGGCGCGCGCTACGAAGTCAGAACGGGTCGCCTGCCCTATTTT-3′
144-192 5’	5′-GCCGCCCGCTGGCGGCAGATTACGACGAGTGGCTACACACGAGGAATG-3′
144-192 3’	5′-CGGCGGGCGACCGCCGTCTAATGCTGCTCACCGATGTGTGCTCCTTAC-3′
192-240 5’	5′-GTGCCGAATAAGGGTAACTATCCTGTCAATGTTGACGGGCAAGAGGTG-3′
192-240 3’	5′-CACGGCTTATTCCCATTGATAGGACAGTTACAACTGCCCGTTCTCCAG-3′
240-279 5’	5′-GCCGAATCCTCAGTGGAAAATCCGCGAGGA-3′
240-279 3’	5′-CGGCTTAGGAGTCACCTTTTAGGCGCTCCT-3′

Table 4.1. Sequences of oligonucleotides used to determine the HicB binding site. The predicted -10 (Red underlined) and -35 promoter elements (blue, underlined) within the 0-48 and 48-96 bp double stranded DNA fragments are highlighted.

A clear shift in DNA migration upon the addition of tetrameric HicB ($HicB_T$), indicative of a protein-DNA complex, was only observed for the 0-48 bp DNA fragment and not for the other five fragments analysed (Figure 4.1A). EMSA was performed with $HicB_T$ (3.2 µM) and each DNA fragment (3.0 µM) and complete binding of 0-48 bp DNA under these conditions suggested a 1:1 stoichiometry (assuming intact tetramer). For each fragment, a weak band was observed with a slower migration time than the free DNA, irrespective of the presence of $HicB_T$ (Figure 4.1A, all lanes). It is hypothesized that upon formation of the double stranded DNA fragments (see Materials and Methods), a small proportion of free 5' and 3' DNA strands mis-annealed to form a higher order molecular weight complex, which would explain the presence of this band. This was observed for several other sequences described in this chapter (see Appendix Table 4, where sequences that displayed this additional band within EMSA assays are marked with an asterix (*)). Subsequent dissection of the 0-48 bp region initially revealed that the first forty nucleotides were sufficient for complete DNA binding (Figure 4.1B). Further dissection then determined that the DNA binding motif that HicB bound was located within a fragment spanning 17-36 bp upstream of the *hicAB* operon (Figure 4.1C). This DNA fragment also appeared to display an additional DNA band at high DNA concentrations (Appendix Figure 20) as observed for other DNA fragments.

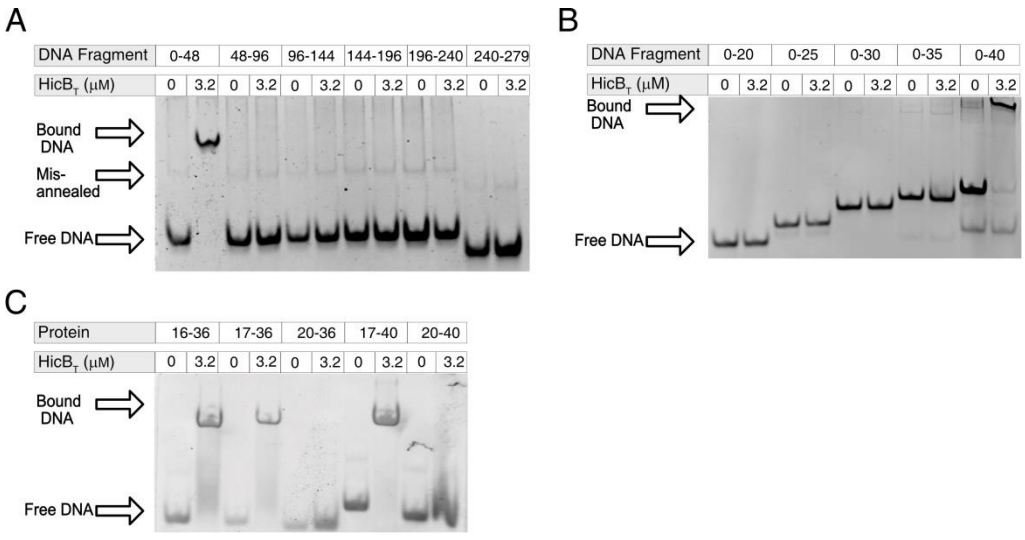


Figure 4.1. EMSA of DNA binding to HicB. (A) EMSA investigating individual fragments of the intergenic region of DNA upstream of the *hicAB* operon. (B) EMSA probing the DNA binding site of $HicB_T$ within the 0-48 bp region. Within the DNA fragment 0-40 binding experiment the presence of unannealed single stranded DNA is observed at a lower migration than the free dsDNA. (C) EMSA probing the DNA binding site to identify the sequence of interest as 17-36 bp.

This sequence 5'-^{36}ATGTGTATAATTACACACAA^{17}-3' contained an inverted repeat sequence at two sites: S1 spanning ^{24}ACACA^{20} and S2 ^{35}TGTGT^{31} (Figure 4.2A). Removal of the proximal 5' adenine (36) or the 3' ^{19}CAA^{17} abolished binding. Hereon in the hypothesized promoter site for the *hicAB* operon will be referred to as S1-2 DNA. This inverted repeat overlaps a predicted -10 sequence (^{31}TATAAT^{26}) and a CRP binding site^{344} (^{37}GATGTGTA^{30}) (Figure 4.2). Binding of $HicB_T$ to S1-2 DNA was dependent on the presence of both intact S1 and S2 sites: individual base substitutions within either S1: ACACA to AGGGA (termed MS1) or S2: TGTGT to TCCCT (MS2) abolished binding of $HicB_T$, when compared to the native S1-S2 sequence (Figure 4.2B). This finding was reinforced by earlier assays (Figure 4.1B) as DNA sequences bearing only the S1 sequence (0-30/35) or the S2 sequence (20-36/40) did not bind HicB at the concentrations tested. HicB-NT, which lacks the RHH motif, did not bind S1-2 DNA at concentrations where full complexation was observed for $HicB_T$ confirming that the C-terminal domains were necessary for DNA binding (Figure 4.2C).

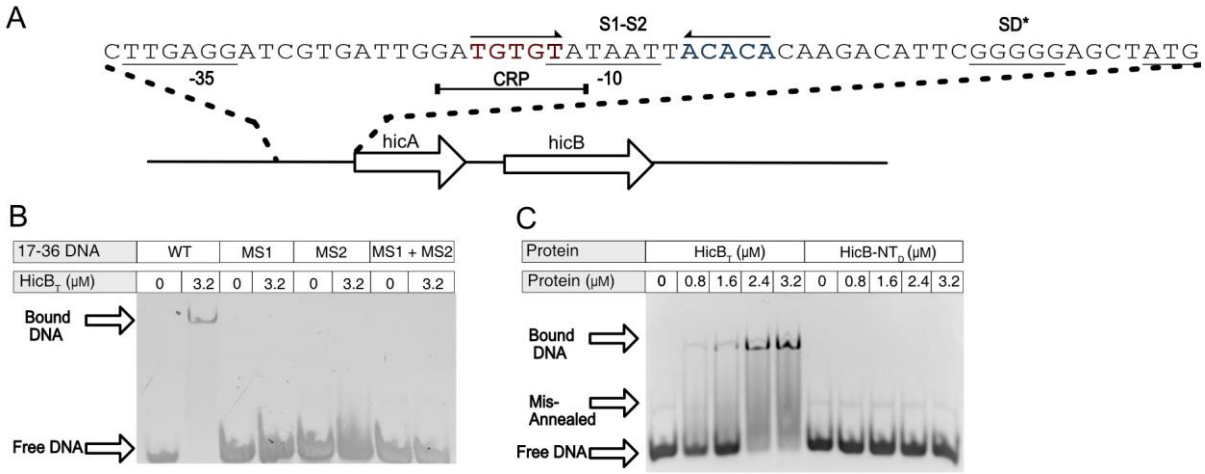


Figure 4.2. Overview of HicB binding to S1-2. (A) Overview of the *hicAB* operon, the inverted repeat sequence within the upstream region of *hicAB* that binds DNA: S1 and S2 are highlighted in red and blue. The putative -10, -35, CRP binding site and SD sequence (SD*) are also highlighted. (B) Mutations within S1-2 (2 μM) at either S1 or S2 abolishes binding to HicB (3.2 μM). (C) EMSA of $HicB_T$ and HicB-NT_D (0-3.2 μM) binding to 2 μM S1-2. DNA self-dimers were present in each lane regardless of the presence of HicB-NT_D.

4.2. Comparison to other HicAB operator sequences

The *Y. pestis* homologue HicB3 and the HicA3B3 complex ($HicA3_2HicB3_4$ confirmed by SEC-MALS) were reported to bind two operator regions of DNA (BS1 and BS2) immediately upstream of the Shine Dalgarno sequence (SD) resulting in transcriptional

inhibition of the *phicA3B3* operon (Figure 4.3A).[132] Each operator contains a 15 bp inverted repeat bearing two half sites: GGTA-N5-CCTA (BS1) and ATCC-N6-ATGG (BS2) (Figure 4.3A, underlined in red). Individual mutations of either BS1 or BS2 were tolerated however mutation of both sites prevented DNA binding by HicB3 and transcriptional repression was not observed. BS2, which overlaps the -10 promoter region was subsequently confirmed as the main operator site for *hicA3B3* transcription.[132]

Based on EMSA experiments (Fig 4a[132]) Bibi-Triki et al. showed that both operator sites are individually occupied by HicB3 tetramers in a concentration dependent manner, indicating that there is broad specificity within the HicB3 RHH domains as they recognise two different sequences in BS1 and BS2, adding a layer of complexity to transcriptional regulation. There is a possibility that $HicB3_T$ binds exclusively to the sequence GGTA-N5-CCTA, found in BS1 and the reverse complement of BS2 as ATCC-N6-ATGG, and Bibi-Triki et al. suggest that each $HicB3_T$ recognises a TRGGTRT half-site, suggesting that there must be sequence plasticity in the sequence recognition by $HicB3_T$. In comparison, the sequences of S1 and S2 found in the *B. pseudomallei* putative promoter site for *hicAB* are equivalent; the reverse complement of TGTGT is ACACA. It is predicted that each RHH strand-swapped dimer, of which there are two in HicB, each individually bind to a single site (S1 and S2). This is to ensure that the HicB tetramer is coordinated in a tight complex and is highly specific to the S1-2 sequence. The organisation of S1-S2 is similar to the BS2 site of *phicA3B3*, the main operator site for *hicA3B3* expression.[132] There is no equivalent of BS1 within the upstream region of *B. pseudomallei hicAB*, and the S1-2 site is the only site where HicB binding is observed. The upstream intergenic region (279 bp) does not appear contain any other operator sites, based on EMSA observations (Figure 4.1A).

The *E. coli* immediate intergenic region upstream of *hicAB* contains two promoter sites: P1 and P2 (Figure 4.3B).[214] P1 is under the control of both CRP-cAMP and Sxy resulting in transcription of the full *hicAB* operon. In contrast P2 contains one inverted repeat sequence termed *hic0.* This overlaps the -35 promoter element within *pHicAB* and is specifically regulated by HicB resulting in transcriptional repression of the *hicAB* operon. Mutation of the inverted repeat (TAAACA-N4-TGTTTACTT to T**GG**ACA-N4-TG**GG**ACTT) abolished HicB binding and therefore transcriptional repression of the operon. *hic0* therefore shows similarities to the *Burkholderia hicAB* S1-2 site.[214]

S. pneumoniae HicB was reported to recognise a 28 bp sequence upstream of the *hicAB* operator (Figure 4.3C). However the specific sub-sites that $HicB_{SP}$ binds and mutational studies on this DNA sequence have not been reported. Bioinformatic analysis suggests that this sequence contains an imperfect inverted repeat sequence (Table 4.2), which overlaps a predicted -10 Pribnow box[344], although it is uncertain if this DNA sequence controls transcription of the *hicAB*$_{SP}$ operon or if multiple operator sequence regulate this operon.[272]

Organism	Operator sequence
Burkholderia	TGTGT-NNNNNN-ACACA
Escherichia	TAAACA-NNNN-TGTTTA
Streptococcus(*)	TAATAG-(N17)-CTTTA
Yersinia (BS1)	GGTA-NNNNN-CCTA
Yersinia (BS2)	ATCC-NNNNNN-ATGG

Table 4.2. The operator site for each HicB family member interaction site with sites of interest highlighted. (*) The half sites for *S. pneumoniae* are predicted by bioinformatics analysis and these have not been experimentally confirmed via EMSA/LacZ assays.

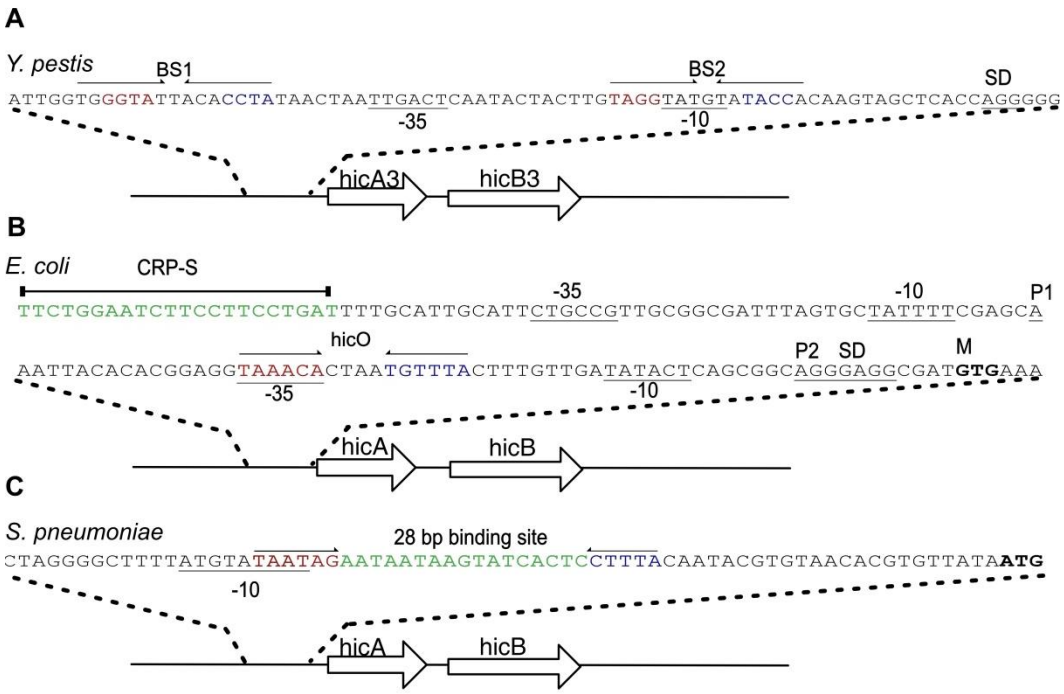


Figure 4.3. Genetic organisation of HicA3B3, HicAB and HicAB$_{SP}$ from *Y. pestis, E. coli* and *S. pneumoniae.* (A) The *Y. pestis hicA3B3* operon with the elements within *phicA3B3*: the inverted repeats BS1 and BS2 are highlighted (arrows) with the conserved GGTA (red) and CCTA (blue) sequences within the TRGGTRT half sites highlighted. The -10, -35 and the SD sequence are highlighted (underlined). (B) The *E. coli hicAB* operon with two promoters P1, controlled by CRP-S (CRP-cAMP and Sxy)(green) and P2, controlled by HicB binding via the *hicO* inverted repeat sequence. Half sites within the inverted repeat (hicO) are marked by arrows and individually highlighted (red/blue). The -10 and -35 sequence for each promoter (P1 and P2) are underlined and the SD sequence and initiator codon of HicA are highlighted. (C) The *S. pneumoniae hicAB* putative operator site. The 28 bp binding site is highlighted (green). The putative inverted repear half sites are highlighted (red/blue, arrows). The putative -10 sequence is highlighted (underlined) and the ATG start codon of HicA$_{SP}$.

4.3. Fluorescence anisotropy K_d determination

Anisotropy is the ratio of the polarised components of light: parallel and perpendicular to the excitation plane (I_{II} and $I_{\perp}$) to the total intensity (I_{tot}) (Materials and Methods, Equation (8)). Free fluorescent ligands have a low anisotropy value as they rapidly tumble, emitting randomly polarised light.[347] DNA binding results in a higher order complex, with a greater molecular volume when compared to the free ligand. The tumbling of the fluorescent ligand decreases as it acquires the larger correlation time of the slowly tumbling protein and light is predominantly emitted along one axis this leads to the emission of polarised light, resulting in a high anisotropy value.[348]

$$r = \frac{I_{II}\text{-}I_{\perp}}{I_{tot}} = \frac{I_{II}\text{-}I_{\perp}}{I_{II}+2I_{\perp}} \qquad (8)$$

Fluorescence anisotropy was used to quantify the binding interaction between HicB and fluorescently labelled DNA (HEX-S1-2) and an adapted variant of the Morrison tight binding equation was used to extract the dissociation constant K_d [349-352] (Materials and Methods, Equation (9)).

$$Y = (A_{DP} - A_D)\left(\frac{[X] + K_d + n \pm \sqrt{([X] + n + K_d)^2 - (4[X]n)}}{2n}\right) + A_D \qquad (9)$$

Where Y is the anisotropy value, A_D is the response (anisotropy) in the absence of protein [X], while A_{DP} is the maximal response (anisotropy) of a protein-DNA complex. n refers to the stoichiometry of the reaction which was assumed to be a single site mode of action.

$HicB_T$ bound DNA with a calculated K_d of 3 ± 0.4 nM (95% confidence interval = 2. - 4 nM) (Figure 4.4.A). This K_d value should be regarded as an upper limit because the binding is too tight to accurately measure at the probe concentration used in experiments (7.5 nM). Simulations of K_d values (0.75 – 3 nM) for HicB-HEX-S1-2 revealed no discernible difference between the simulated dose response curves aside from the 0.75 nM simulated curve, which did not fit the experimental data (Figure 4.4.B). This K_d was much tighter than the equivalent K_d values of 300 nM and 8 μM reported for the HicAB systems within *E. coli* and *S. pneumoniae*.[132, 272] The stoichiometry was deemed to be one ($HicB_T$ binds to one

HEX-S1-2 DNA) based on simulation of the stoichiometry value. When the K_d was fixed at its experimentally determined value (3 nM), the stoichiometry was approximated as 0.973 ± 0.04 using equation (9). This provided evidence that one tetramer binds both S1 and S2 sites simultaneously, as previously hypothesized and is in agreement with similar observations reported in *S. pneumoniae* and *Y. pestis*.[132, 272] As observed by EMSA, increasing concentrations of HicB-NT_D does not result in a substantial change in anisotropy, indicative of DNA binding, even at concentrations where $HicB_T$ elicited a maximum response (150 nM) (Figure 4.4C).

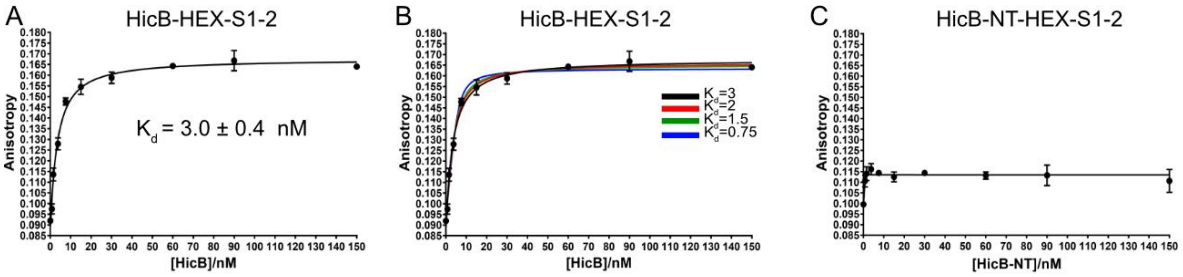


Figure 4.4. Fluorescence anisotropy experiments of HicB and HicB-NT to HEX-S1-S2. (A) Quantification of HicB binding to HEX-S1-2. Samples contained 7.5 nM HEX-S1-2 in DNA binding buffer. The proportion of HEX-S1-2 bound by increasing concentration of HicB was followed (n=1). (B) The K_d of HicB-HEX-S1-2 were simulated to determine appropriate values that described the experimental data seen in panel A. (C) Quantification of HicB-NT binding to HEX-S1-2. The proportion of HEX-S1-2 bound by varying concentrations of HicB-NT was followed (n=1). For all experiments three independent fits were fit to equation (9) and the mean value is plotted with error bars representing the SEM. Standard errors of K_d values were calculated in GraphPad Prism.

4.4. Probing the HicB DNA binding site

To determine how HicB may bind DNA, the exposed surface charge of *B. psuedomallei* HicB was investigated (Figure 4.5A, B). A patch of positively charged amino acids were identified on one face of the tetramer, corresponding to the RHH motif of each subunit and included the surface exposed residues (R94, N96 and S98) of the strand-swapped β-sheet (β5, ^{94}RINVSI^{99}).

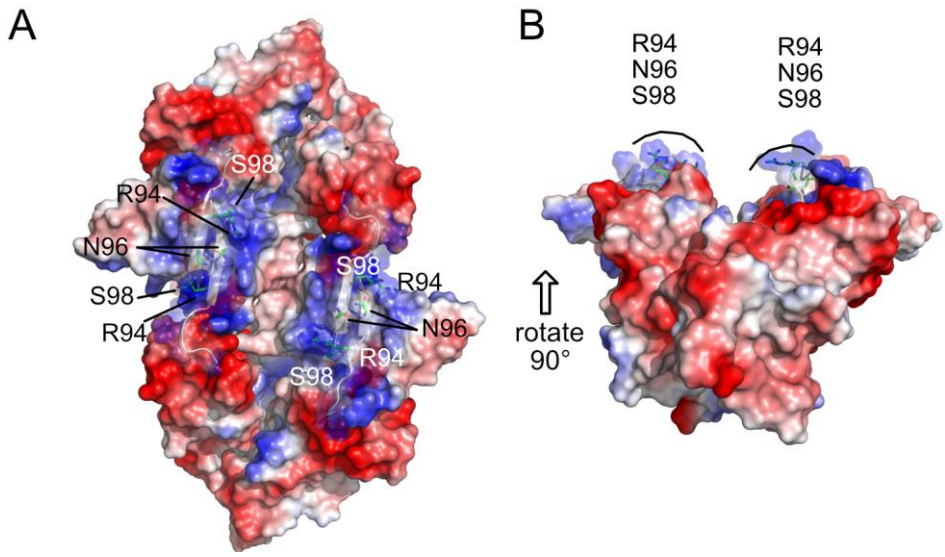


Figure 4.5. Analysis of the surface charge distribution of HicB. (A) Positively charged patch (blue) of HicB formed by R94, N96 and S98 in the β5 strand. This β-sheet lies flat across the structure due to the parallel nature of the C-terminal strand swapped dimers. (B) 90 ° rotation to further highlight the patch of positive charge.

Superimposition of HicB to structurally similar well characterised proteins that bear an RHH domain (and antitoxins that contain a RHH motif) (Appendix Table 10) indicated the structural conservation of R94, N96 and S98 (Appendix Figure 21). The equivalent amino acids encoded in these structurally similar proteins have been demonstrated to form protein–DNA interactions.[83, 142, 228, 312, 313, 315, 316] Equivalent amino acids to R94, N96 and S98 were observed in *Yersinia* HicB3 (K95, N97 and T99), but not *Escherichia* HicB, which encodes a HTH rather than a RHH, where no sequence conservation was observed (Appendix Figure 22). Residues within these HicB homologues have not been mutated to investigate their effect on DNA binding.

Alignment of HicB to 44 homologous HicB antitoxins (Appendix Figure 23) revealed conservation of R94, N96 and S98 within the β5 strand of the C-terminal domain within HicB family members (Figure 4.6).

Figure 4.6. A Web logo[334] generated graphic depicting the conservation of amino acids at the β5 strand of HicB when aligned to 44 homologous HicB sequences using Clustal Omega[291]. The amino acid sequence of the β5 strand of HicB is highlighted underneath to indicate the strong conservation of polar and hydrophobic residues among HicB family members.

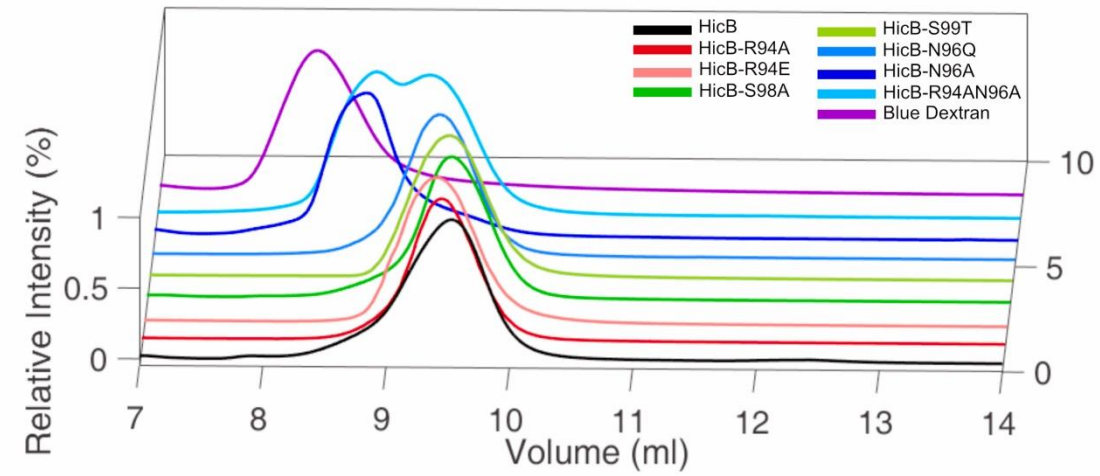
To investigate the role of these residues further, genes containing HicB encoding specific alanine mutations (R94A, N96A, R94N96A and S98A) and semi-conservative mutants (R94E, N96Q and S98T) were designed. Alanine mutants and N96Q were synthesized by ThermoFisher™ and subcloned into the pOPINE vector (Mutant Primer 1 and 2, Appendix Table 4). R94E and S98T constructs were generated by PCR site-directed mutagenesis from pOPINE-HicB-R94/S98A constructs (R94E and S98T primer 1 and 2, Appendix Table 4). All mutants were individually expressed in T7 express cells (NEB) and were purified by IMAC and SEC, with the following yields: R94A (20 mg/L), R94E (25 mg/L) N96A (3 mg/L), N96Q (23 mg/L) R94AN96A (2.5 mg/L), S98A (20 mg/L) and S98T (22 mg/L).

Each purified HicB mutant (Appendix Figure 24, 25) was soluble and denatured ESMS confirmed the presence of each mutation (Appendix Figure 26). Deconvolution of each m/z envelope revealed a single species for each protein: R94A (observed 15,654, expected: 15,653 Da), R94E (observed: 15,711, expected: 15,711 Da) N96A (observed: 15,694, expected: 15,695 Da), N96Q (observed: 15,751, expected: 15,752 Da), R94AN96A (observed: 15,609, expected: 15,610 Da), S99A (observed 15,721 Da, expected: 15,722 Da) and S99T (observed 15,751 Da, expected: 15,752 Da).

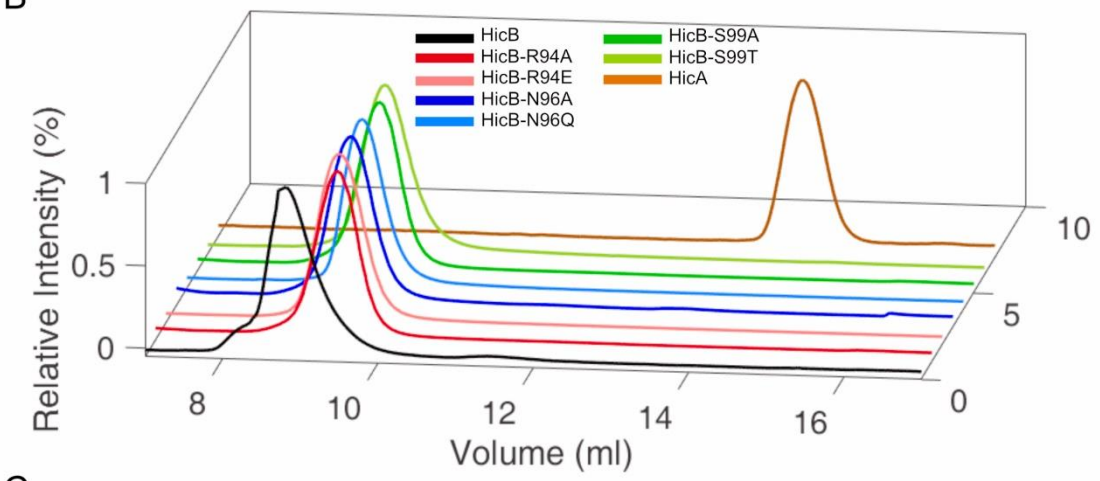
Native mass spectrometry confirmed that each mutant formed a tetrameric species (Appendix Figure 27, 28). Interestingly, additional charge states were observed between 3,000-4,500 m/z for the N96A and R94AN96A mutants. These could not be assigned as monomeric, dimeric, tetrameric or octameric charge states and are believed to be aggregates formed in ammonium acetate during sample preparation. These additional states were not observed in R94A, R94E or N96Q suggesting that conservative changes have less impact on the structure of the tetramer.

ASEC (AS75) was also used to determine the oligomeric state of each mutant (50 μM, [$HicB_T$] was 12.5 μM) (Figure 4.7). With the exception of N96A and R94AN96A, each mutant gave a similar profile to the HicB tetramer (Figure 4.7A, black). N96A (Figure 4.7A, blue) eluted as a single species with a Mw_{app} of 85.9 kDa (42.7-104.6 kDa) and the elution profile overlapped the profile of blue dextran (Figure 4.7A, purple) which is indicative of a species larger than the separation size of the column (75 kDa). R94AN96A (Figure 4.7A, cyan) eluted as two species, one with a Mw_{app} of 71.6 kDa (66.3-95.5), which overlapped the blue dextran profile and the other with an Mw_{app} of 60.4 kDa (40.9-65.4), which overlapped the profile of HicB. The dual species associated with R94AN96A were not observed by SEC on a preperative S75 column during initial purification, although it is likely that this column lacked the resolution to distinguish the two peaks. Complexation of each mutant with HicA at a 1:1 ratio (50 μM HicB ($HicB_T$ =12.5 μM):50 μM HicA) determined that each mutant retained the ability to bind HicA due to the complete loss of the HicA peak, previously observed for the HicAB complex, although a peak shift for each HicB mutant was not observed (Figure 4.7B). This suggested that the N-terminal domain remained intact and any structural deviations, as observed for N96A and R94AN96A, were confined to the C-terminal domain.

A



B



C

Sample	Mw (Da)	Mw_{app} (kDa)	Mw range (kDa)	$Log_{10}Mw$	V_e (ml)	K_{av}
HicB	62,952	65.6	52.3-76.5	1.82	9.03	0.11
R94A	62,612	56.3	43.2-77.2	1.75	9.43	0.14
R94E	62,952	65.6	52.3-76. 5	1.82	9.03	0.11
N96A	62,780	85. 9	47.7-104.6	1.93	8.32	0.07
N96Q	63,008	57.5	47.08-69.3	1.76	9.38	0.13
R94AN96A	62,436	71.6/60.4	66.3-95.5/ 40.9-65.4	1.85/1.78	8.8/ 9.25	0.10/ 0.13
S98A	62,888	55.5	40.9-81.1	1.85	9.47	0.10
S98T	63,008	56.5	46.6-68.9	1.75	9.44	0.14

Figure 4.7. Analytical size exclusion profiles of HicB DNA binding mutants. (A) SEC profiles for all HicB mutants: HicB (black), R94A (red), R94E (pink), N96A (blue), N96Q (light blue), R94AN96A (cyan), S99A (green) and S99T (light green). Blue dextran (purple) is shown to indicate the void volume of the column (B) SEC profiles of HicB mutants complexed with HicA. Traces are highlighted as per panel A with HicA (brown) added. (C) Summary of results for each analysed sample reporting the expected molecular weight (Mw), Mw_{app}, Range of molecular weight values encompassed by each peak, Log_{10} Mw, V_e (Volume of elution) and K_{av} (Partition coefficient) determined by the calibration curve. The calibration curve can be seen in Figure 3.2B.

The DNA binding ability of each mutant was assayed using EMSA (Figure 4.8). To minimise the presence of concentration-dependent aggregation, protein samples used within this assay were not concentrated following SEC. R94A, N96A, R94A/N96A and S98A were unable to bind S1-2 (2 μM) at concentrations where full complexation of S1-2 by $HicB_T$ was observed (3.2 μM) (Figure 4.8A-D). Increasing the protein concentration of R94A and S98A to 250 μM also did not result in a band shift (Appendix Figure 29). N96A and R94AN96A could not be assayed at these concentrations due to the low yields associated with their purification. R94E and N96Q did not bind S1-2 (2 μM) at concentrations where a HicB-S1-2 complex was observed ($HicB_T$ = 3.2 μM) (Figure 4.8E, F). In contrast S98T caused notable DNA smearing at concentrations greater than 1.6 μM with a faint band seen at 3.2 μM (Figure 4.8G, circled). The smearing suggested the ability to bind DNA, albeit at a lower affinity, forming protein-DNA complexes that were labile and dissociated during electrophoresis.[353, 354]

At high concentrations (250 μM) smearing was also observed for N96Q, but no discrete band could be observed for either N96Q or S98T (Figure 4.8H). R94E showed no smearing and therefore any observable DNA interactions, consistent with the complete charge reversal at this site.

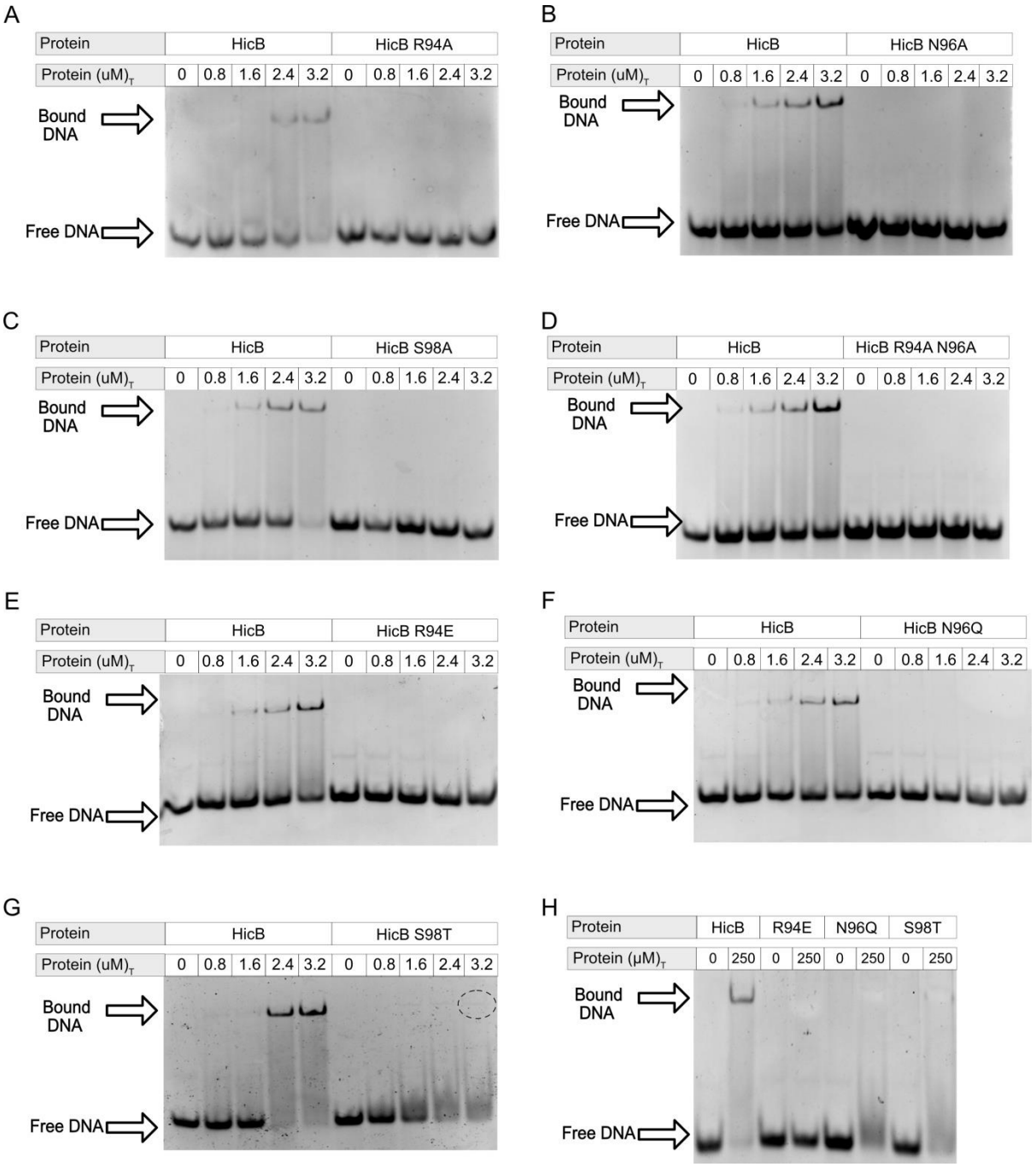


Figure 4.8. EMSA of HicB mutants involved in the DNA binding site. (A) EMSA of $HicB_T$ and $HicB_T$-R94A (0-3.2 μM) binding to 2 μM S1-2. (B) EMSA of $HicB_T$ and $HicB_T$-N96A (0-3.2 μM) binding to 2 μM S1-2. (C) EMSA of $HicB_T$ and $HicB_T$-S98A (0-3.2 μM) binding to 2 μM S1-2. (D) EMSA of $HicB_T$ and $HicB_T$-R94AN96A (0-3.2 μM) binding to 2 μM S1-2. (E) EMSA of $HicB_T$ and $HicB_T$-R94E (0-3.2 μM) binding to 2 μM S1-2. (F) EMSA of $HicB_T$ and $HicB_T$-N96Q (0-3.2 μM) binding to 2 μM S1-2. (G) EMSA of $HicB_T$ and $HicB_T$-S98T (0-3.2 μM) binding to 2 μM S1-2. (H) EMSA of $HicB_T$ and each semi conservative mutant (250 μM) against 2 μM S1-2.

To quantitatively determine the effect of individual mutations to the β5 strand, FA was used to determine K_d values for each semi-conservative mutant (Figure 4.9). Alanine mutants were not investigated as no binding was observed even at 250 μM for R94A and S98A. A K_d value

could not be determined for R94E as there was no observed increase in anisotropy upon increasing concentrations of R94E (0-1000 nM) (Figure 4.9A). N96Q gave a K_d value of 296 ± 32 nM (95% confidence interval = 231 – 376 nM), a ~ 100-fold reduction compared to HicB (Figure 4.9B) whereas S98T displayed a K_d of 94 ± 8 nM (95% confidence interval = 79.-112 nM), an ~ 30-fold reduction when compared to HicB (Figure 4.9C). This correlated with EMSA observations as S98T displayed DNA binding at lower concentrations (3.2 μM) than N96Q, although as HicB DNA mutants do not appear to tightly bind S1-2 DNA, these K_d values are likely to be underestimation of the true value.

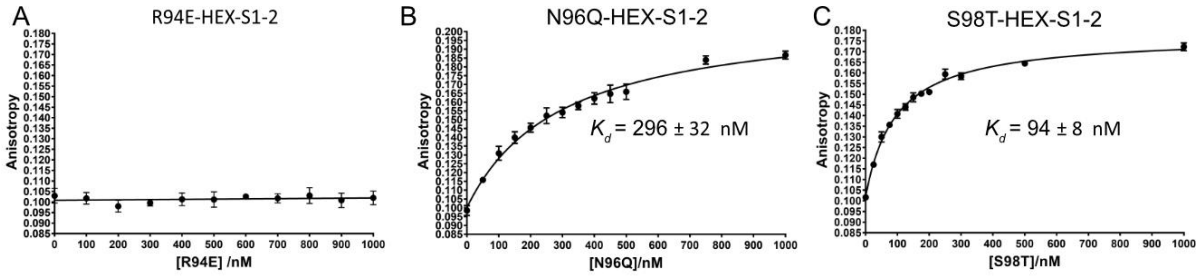


Figure 4.9. Fluorescence anisotropy experiments of HicB-R94E, N96Q and S98T. (A) Quantification of R94E binding to HEX-S1-2. (B) Quantification of N96Q binding to HEX-S1-2. (C) Quantification of S98T binding to HEX-S1-2. For each experiment, for 3 independent repeats (n =1). Data were fit to equation (9). The mean value is plotted with error bars representing the SEM. Standard errors of K_d values were calculated in GraphPad Prism.

The loss of DNA binding observed for HicB constructs containing N96 alanine mutants might arise mainly from structural effects on HicB. Within the strand-swapped dimer of HicB adjacent N96 residues are tethered together via a side chain-side chain hydrogen bond and are capable of forming a hydrogen bond with the side chain of R94 (Figure 2.15). Removal of N96 may diminish this network and the additional removal of R94 (R94AN96A) may further destabilize HicB, perhaps explaining the much lower yield associated with the R94AN96A mutant due to possible misfolding and instability. Closer analysis of the β5 strand reveals that R94, N96 and S98 (Figure 4.10A, B, grey) are flanked by hydrophobic residues (Figure 4.10, A, B, red). Mutation of either R94 or N96 to an alanine residue, or both, would increase the hydrophobicity of this region (Figure 4.10C). In the case of the double mutant (R94A, N96A), introduction of an alanine at both positions leads to a distinct hydrophobic patch across the β5 strand spanning 12 Å (Figure 4.10D), where S98 is the only remaining polar residue. Such a large patch could promote structural rearrangements to bury the surface exposed hydrophobic patch or create a sticky patch leading to aggregation. In comparison these hydrophobic patches are unlikely to be present in the R94E, N96Q and S98T mutants.

SAXS confirmed the presence of shape perturbations for alanine mutants but not semi-conservative mutants, when compared to HicB (Appendix Figure 30-32, Appendix Table 19-20). High-resolution structures of each mutant are required to confirm this hypothesis.[355-357] Each mutant was screened at 500 μM ($HicB_T$ = 125 μM) and 1,000 μM ($HicB_T$ = 250 μM) against commercially available Molecular Dimensions screens (Morpheus, Proplex, JSCG+ and Structure Screen I + II). To date no crystals have formed in any conditions for any alanine mutants and it appears their surface entropy differs from HicB (Appendix Figure 33). Crystals of R94E, N96Q and S98T formed in identical conditions to HicB (0.1 M NaOAc, pH 4.6 8% (w/v) PEG 4000, 10% (v/v) glycerol), but these have not yet been analysed due to a lack of beamline time.

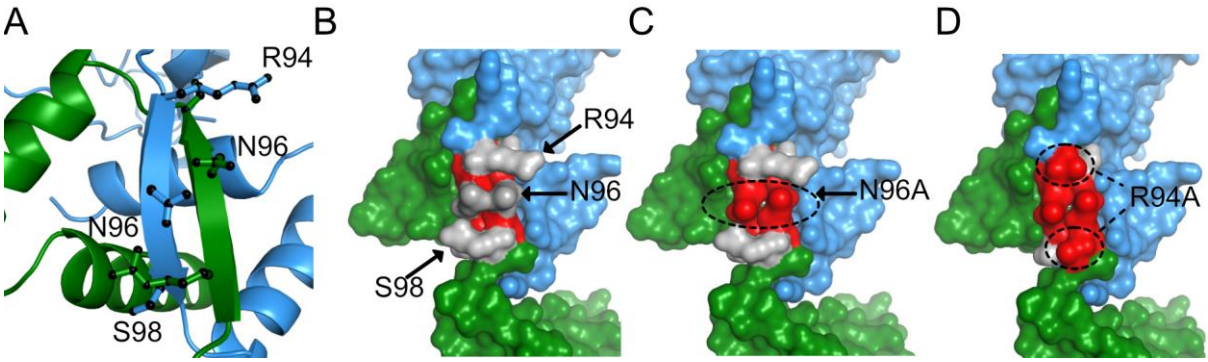


Figure 4.10. Hydrophobic analysis of HicB. (A) Location of the targeted amino acids: R94, N96 and S98 in the β5 ribbon-helix-helix. (B) Surface of the β5 strand with R94, N96 and S98 highlighted in grey while the hydrophobic residues (I95, V97 and I99) are highlighted in red. (C) The hydrophobic patch produced when N96 is mutated to alanine. (D) Hydrophobic patch now surface exposed when both R94 and N96 are mutated to alanine.

4.5. HicB-DNA modelling

SEC-SAXS was first used to investigate the HicB-S1-2 complex, preformed at a 1:1.6 ratio ($HicB_T$: 300 μM: S1-2: 500 μM) (Appendix Table 21, Appendix Figure 34) but initial *ab initio* modelling of a HicB-S1-2 complex was unsuccessful (data not shown). However, an *ab initio* shape envelope was obtained for free S1-2 DNA (Figure 4.11A). A S1-2 DNA model was built by 3D dart[358] and superimposed onto the *ab initio* shape envelope which indicated that free S1-2 DNA was bent and the distance between the S1 and S2 sites was 37/38 Å within the 5' and 3' strand (Figure 4.11A). Within free HicB, the distance between the adjacent RHH domains is 32 Å and this distance increases by 44 Å in the HicAB complex (Figure 4.11B, C), suggesting that only free HicB can bind S1-2.

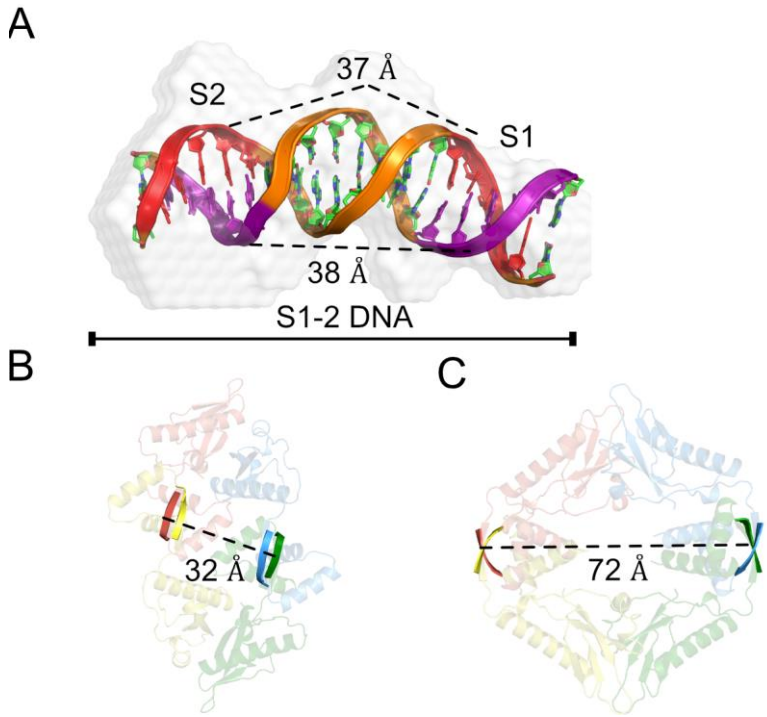
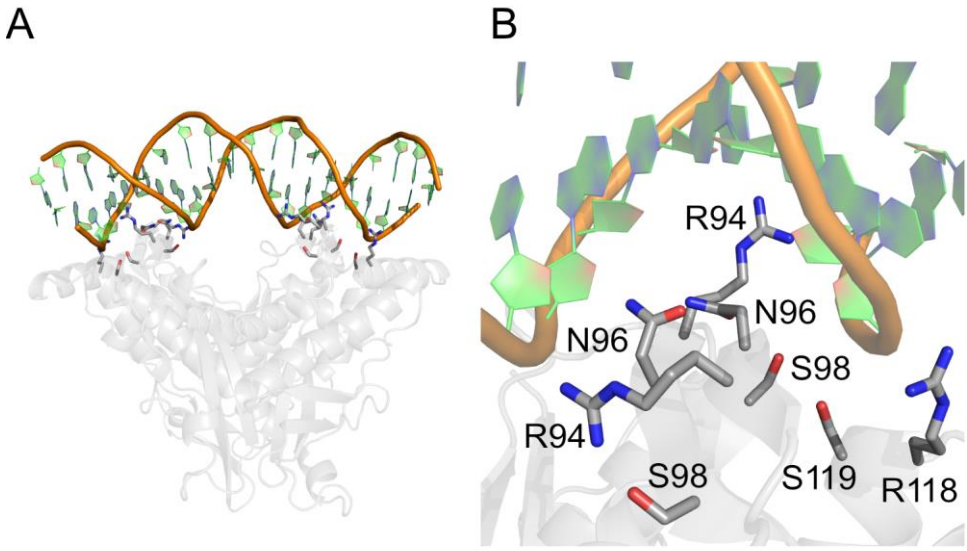


Figure 4.11. Distance between DNA binding sites. (A) Model of S1-2 DNA superimposed onto an *ab inito* shape envelope of S1-2 DNA. Here the half sites (S1/S2) of the inverted repeat are separated by 33/37 Å. (B) Cartoon representation of free HicB highlighting the distance between adjacent RHH domains. (C) Cartoon representation of HicAB highlighting the distance between adjacent RHH domains. For clarity S1 and S2 sequences on the 5' and 3' strand are highlighted in red and purple.

Following discussion with Professor John McGeehan (University of Portsmouth), a preformed $HicB_T$-S1-2 complex (1:1.2 stoichiometry) was screened at a variety of concentrations ($HicB_T$ = 125, 250 and 375 μM) using S1-2 DNA and a DNA variant that contained an overhang (Overhang Primer 1 and Overhang Primer 2) against commercially available Molecular Dimensions screens (Morpheus, Proplex, JSCG+ and Structure Screen I + II) and Hampton Research (Natrix HT) and Qiagen (Nucleix) screens. Small crystals were obtained in a number of conditions, but the best only diffracted to 8 Å and further optimisation is ongoing. MR was not attempted and it has not been fully determined if crystals contain a protein-DNA complex or just DNA.

In the absence of a HicB-DNA crystal structure, *in silico* molecular docking using the High Ambiguity Driven protein-protein DOCKing algorithm (HADDOCK)[359] was used to investigate the structure of the HicB-S1-2 complex and the interface between HicB and S1-2. S1-2 models were built using the 3D-DART server [358] before modelling onto the crystallised elongated conformation of HicB. Active residues of HicB were designated as R94, N96 and S98, inferred from EMSA and FA experiments. S1-2 active residues were designated as

^{24}ACACA^{20} and ^{35}TGTGT^{31}. In this HicB-S1-2 model (Figure 4.12), individual RHH domains bind to ^{24}ACACA^{20} and ^{35}TGTGT^{31} which lies flat across the base of HicB. A model of the interaction site highlights the R94, N96 and S98 project into the major groove and suggests that R118 and S119 form interactions with the DNA backbone (Figure 4.12). This model was a good fit to the SAXS experimental profile of HicB-S1-2 (χ^2 =2.10) and the resultant *ab initio* shape envelope (Appendix Figure 35).



C

	Cluster 6	Cluster 3
Haddock score	-173.8 ± 17.4	-162.5 ± 6.8
Cluster score	8	12
RMSD (Å)	0.6 ± 0.3	14.4 ± 0.1
Van der Waals energy	-75.9 ± 5.0	-78.5 ± 5.1
Electrostatic energy	-513.8 ± 48.2	-410.3 ± 13.1
Desolvation energy	-10.8 ± 6.3	-20.5 ± 9.9
Restraints Violation energy	157.3 ± 52.27	185.2 ± 30.40
Buried Surface Area	2124.0 ± 45.3	2165.6 ± 41.1.
Z score	-1.5	-0.9

Figure 4.12. HADDOCK modelling of HicB-S1-2. (A) Cartoon representation of the complex where DNA lies across the two RHH domains of HicB. (B) 90° rotation to show the antiparallel β-sheet of the RHH domain projects into the major groove of TGTGT and ACACA. (C) Statistics shown for the HADDOCK clusters generated from simulation of HicB and S1-2 generated by 3D-DART. Clusters are ranked by their HADDOCK score, RMSD to the input models and Z-score.

4.6. Conservation of the HicB-DNA binding site

The DNA binding ability of HicB has been investigated in both *Y. pestis* and *E. coli*[132, 214], however specific amino acids implicated in this interaction have not been determined. Recently the $HicB_{SP}$-DNA interaction site was investigated by NMR and modelled by HADDOCK, where they also reported that DNA binds across the RHH domains of HicB in a similar manner to the HicB-S1-2 model (Figure 4.13).[272] Here they report two interaction sites for DNA. The first: located within the β5 strand of the RHH domain (I114, K115, K116 and T117) are directly homologous to DNA binding residues in *Burkholderia* HicB and are predicted to project into the DNA major groove (Figure 4.13). The second within the α5 helix (F135 and Q137, ^{134}NFSQ^{137}) anchors the protein to the DNA phosphate backbone. The observations of the importance of specific β5 residues reinforce the trend observed in several HicB family members (Figure 4.6), where a proximal basic residue succeeded by several polar amino acids are conserved across HicB species. F135 and Q137 were not conserved across homologues to *B. pseudomallei* HicB (Appendix Figure 23) however; the HicB α6 helix residues R118 and S119 were conserved throughout homologous sequences (Figure 4.13). HADDOCK modelling implies that they constitute an equivalent secondary DNA binding site of HicB (Figure 4.12).

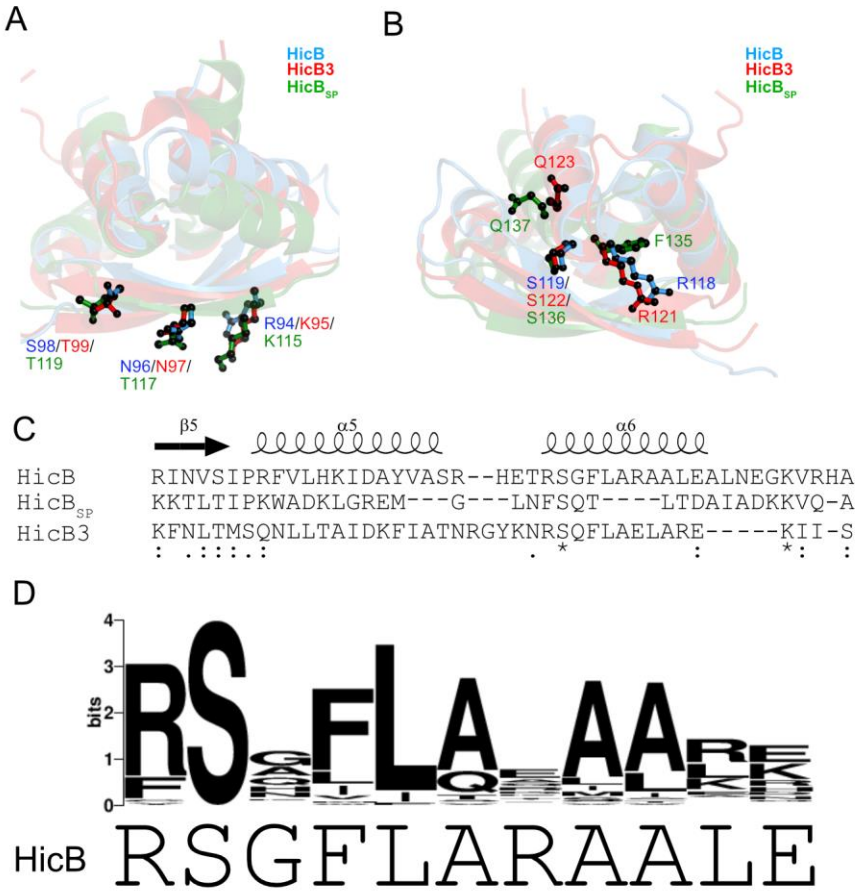


Figure 4.13. Comparison of HicB, HicB$_{SP}$ and HicB3. (A) Superimposition of the RHH domain of each HicB family member with residues of the central β strand involved in binding highlighted. (B) Residues of the α2 helix of HicB$_{SP}$ experimentally determined to bind DNA and the close proximity residues of HicB and HicB3 that could mediate interactions with the DNA backbone. (C) Sequence of *B. pseudomallei* HicB RHH domain and comparison to HicB$_{SP}$ and HicB3. Symbols indicate a conserved residue (*), conservative mutation (:) and a semi-conservative mutation (.). (D) Graphical representation of the amino acid residues in 44 homologous HicB sequences following sequence alignment corresponding to the α6 helix of HicB. The amino acid sequence of the α6 strand of HicB is depicted underneath.

Full verification of the orientation at which DNA binds HicB and whether the a6 helix of HicB constitutes a second binding site awaits a high-resolution structure of HicB-S1-2, however crystallisation trials to date have been unsuccessful (best diffraction 8 Å) and further optimisation is ongoing. Despite the differences between *Burkholderia* and *Streptococcus* HicAB complexes, modelling suggests that they both contact DNA in a similar manner. It is possible that this is a general mechanism for HicB family members that encode a RHH domain.

4.7. Summary

The experimental data depicted in this chapter provides a high degree of certainty that R94, N96 and S98 are functionally important for the formation of a stable HicB-DNA complex and form the "active site" of the DNA binding site. It is clear that alanine mutants (R94A, N96A and S98A) are not capable of binding DNA in comparison to their semi-conservative counterparts (R94E, N96Q and S98T). It appears that a single alanine mutation in each subunit (x4) may have a small effect to the tertiary structure at the local site of DNA binding. In comparison the presence of semi-conservative mutants has a minimal impact on the tertiary structure of $HicB_T$ and the reduction in DNA binding observed within EMSA and FA experiments are due to a knockdown of function rather than local structural perturbations. The specific nucleotides of the DNA binding site upstream of the *hicAB* operon has been determined as an inverted sequence (TGTGT-N6-ACACA), found within the first 48 nucleotides upstream from the proposed transcriptional start site. Data from sequence alignments, structures and mutagenesis were then used to *in silico* model the HicB-DNA complex.

5. DNA binding of HicAB

The precise molecular details of HicAB-DNA interactions are not understood. *Y. pestis* HicAB ($HicA3_2HicB3_4$) has been shown to bind DNA to the same extent as HicB3 alone via EMSA experiments and an excess of HicA3 alleviated repression of *phicA3B3* in vivo.[132] Turnbull and Gerdes then demonstrated using LacZ assays that *E.coli* HicB ($HicB_{EC}$) is not dependent on *E. coli* HicA ($HicA_{EC}$) for full transcriptional repression of the *hicAB* operon and when in excess, $HicA_{EC}$ acted as a de-repressor. This is in contrast to the behaviour of other type 2 toxins, which typically function as co-repressors.[112] It was acknowledged that they could only speculate on the molecular mechanism by which HicA destabilises the HicB-DNA complex and suggested that the mechanism of regulation may be similar to MqsRA where the antitoxin is competed off DNA, due to overlapping binding sites for DNA and toxin.[214]

This chapter probes the interaction of *B. pseudomallei* HicAB with S1-2 DNA and the role of HicA in modulating the HicB-DNA interaction.

5.1. HicA de-represses HicB-S1-2

Initially a simple EMSA was used to monitor the effect of HicA on a preformed HicB-DNA complex. $HicB_T$ (3.2 μM) and S1-2 (2 μM) were mixed and formation of the HicB-S1-2 complex confirmed by EMSA (Figure 5.1A, lanes 1 and 2). HicA was then titrated into solutions of this complex to give final HicA concentrations ranging from 0.8 – 50.8 μM. At ratios of $HicB_T$: HicA <1:1 a bandshift corresponding to the HicB-S1-2 complex persisted although a sizeable fraction of the HicB-S1-2 complex appeared to dissociate. At concentrations of HicA that equal or exceed the concentration of $HicB_T$ and where all four HicA binding sites are potentially occupied (HicA ≥ 12.8 μM), complete dissociation of S1-2 DNA from the HicB-S1-2 complex was observed. Qualitatively excess HicA appears to act as a de-repressor in a similar fashion to the $HicA_{EC}$ driven de-repression of the $HicB_{EC}$-DNA complex[214] and is hypothesized to be a consequence of the conversion of HicB to the $HicA_4HicB_4$ complex. HicA alone was unable to bind S1-2 (Figure 5.1B) and it was therefore assumed that dissociation of S1-2 was a direct consequence of complexation between HicA

and HicB. In the lower concentration ranges, the addition of HicA (0.8 - 6.4 μM) may lead to the formation of initial $HicA_1HicB_4$/$HicA_2HicB_4$ complexes which significantly populated a non-DNA bound state. Complete quantification of unbound and bound populations of S1-2 and HicB-S1-2 was not possible due to evident smearing on EMSA gels upon complex formation, seen for other type 2 TA systems when bound to DNA.[212, 219, 220, 224, 360]

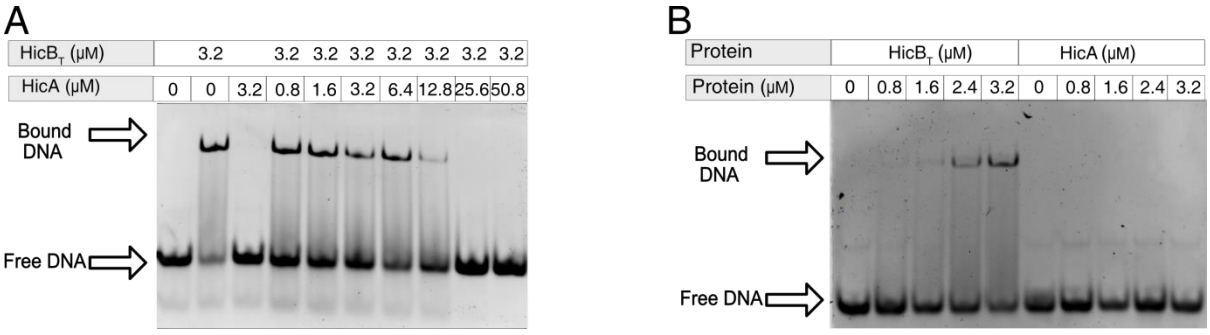


Figure 5.1. HicA de-repression of HicB-S1-S2 DNA. (A) EMSA of varying HicA concentrations (0.8 μM – 50.8 μM) titrated into $HicB_T$ (3.2 μM) resulting in the complete de-repression of HicB-S1-2 binding upon an excess of HicA (25.6 μM). (B) EMSA indicated that HicA does not bind to S1-S2 at concentrations assayed for HicB.

Fluorescence anisotropy was used to further quantify these observations. A preformed complex of HicAB (HicA:HicB ratio = 0.5:1), was titrated (0-150 nM) into HEX-S1-2 (Figure 5.2A). The calculated K_d was 1.9 ± 0.2 nM (95% confidence interval = 1.54-2.2 nM).

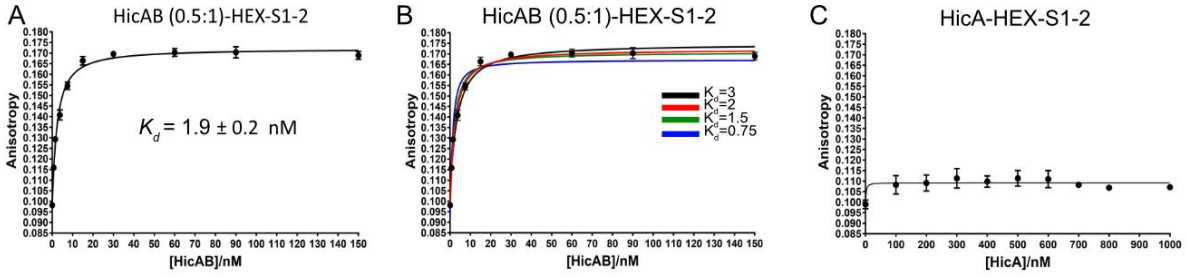


Figure 5.2. Fluorescence anisotropy experiments of HicAB (<1:1, HicA:HicB = 0.5:1) to HEX-S1-S2. (A) Quantification of HicAB (preformed at a ratio of HicA: HicB = 0.5:1) binding to HEX-S1-2. Samples contained 7.5 nM HEX-S1-2 in DNA binding buffer. The proportion of HEX-S1-2 bound by increasing concentration of HicAB was followed (n=1). Data of three independent repeats were fit to equation (9). (B) The K_d of HicAB was simulated to determine appropriate values that described the experimental data seen in panel A. (C) Quantification of HicA binding to HEX-S1-2. The proportion of HEX-S1-2 bound by varying concentrations of HicA was followed (n=1). Again 3 independent repeats were fit to equation (9). For each experiment the mean value is plotted with error bars representing the SEM. Standard errors of K_d values were calculated in GraphPad Prism.

Simulation of K_d values (0.75-3 nM) for HicAB-HEX-S1-2 (Figure 5.2B-C) revealed no discernible difference between K_d values of 1.5-3 nM, although it was clear that a K_d of 0.75

nM did not fit the data. In accordance with experimental error associated with the experiment, there is no change in the affinity of HicB for S1-2 upon HicA binding at sub-stoichiometric levels (HicB-S1-2 K_d = 3.0 ± 0.4 nM). The stoichiometry remained at 1 (One $HicAB_T$ binds to 1 HEX-S1-2 DNA) based on floating the stoichiometry value and constraining the K_d to 1.9 nM and is similar to $HicB_T$ alone. Similar to HicB-HEX-S1-2 assays, the K_d value is lower than the probe concentration (7.5 nM) and this value should be regarded as an upper limit for the K_d value. A K_d for HicA-HEX-S1-2 could not be obtained indicating that HicA does not bind HEX-S1-2, in agreement with EMSA observations (Figure 5.2C).

A preformed complex of HicAB where HicA was in excess (HicA:HicB =1.2:1) was titrated into HEX-S1-2, resulting in a calculated K_d of 0.4 μM ± 0.05 μM (95% confidence interval = 49-1000 nM) with binding at 150 nM only reaching ~ 30% saturation (Figure 5.3A), when compared to the data for HicB alone (Figure 4.4).

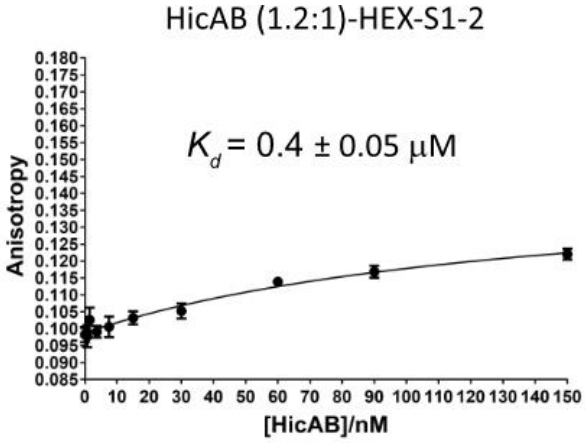


Figure 5.3. Quantification of HEX-S1-2 binding to HicAB (>1:1, HicA: HicB = 1.2:1) for 3 independent repeats (*n*=1). Data were fit to equation (9). The mean value is plotted with error bars representing the SEM. Standard errors of K_d values were calculated in GraphPad Prism.

To confirm if HicA was actively driving the dissociation of the HicB-S1-2 complex through the formation of a HicAB complex, two titration experiments were performed where HicA was titrated into a preformed HicB-S1-2 complex. Firstly, HicA was titrated into a preformed a $HicB_T$ (50 nM)-S1-2 complex (7.5 nM), corresponding to a ~ 95% HicB-S1-2 bound state (based on previous FA assays) (Figure 5.4A). No change in anisotropy was observed at concentrations of HicA below 200 nM (<1:1 HicA:HicB ratio). Once the titration of HicA exceeded 200 nM, a rapid decrease in anisotropy was observed, with approximately 0% HEX-S1-2 bound to HicB after the addition of 700 nM HicA. This experiment apparently

suggested that HicA concentrations approaching a 1:1 stoichiometry with HicB did not have any effect on DNA binding which did not agree with the EMSA titration experiment above. As the HicB-S1-2 complex was 95% bound, there would be a population of free HicB that HicA may preferentially interact with upon titration rather than directly interacting with HicB bound to S1-2 DNA. These experimental parameters were therefore not ideal for investigating the direct effect of HicA on HicB-S1-2 complexes.

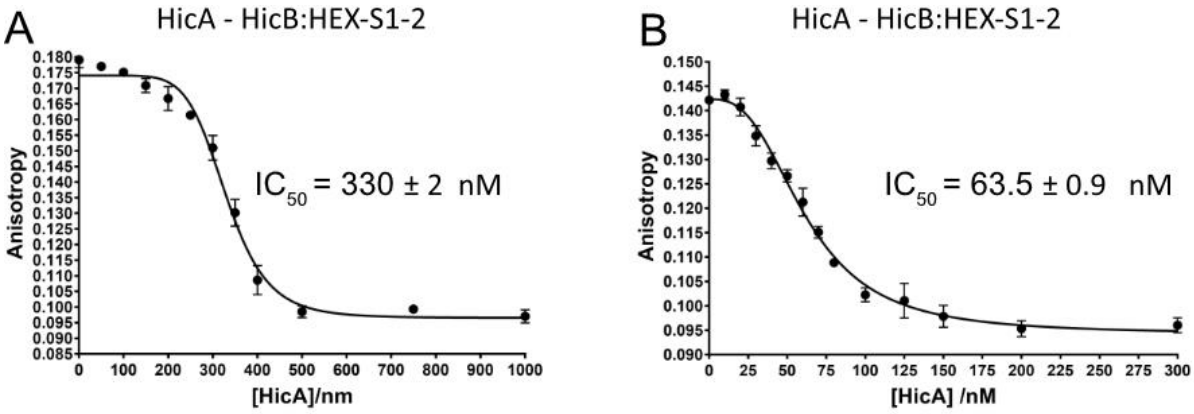


Figure 5.4. Fluorescence anisotropy experiments titrating HicA into HicB-HEX-S1-2. (A) Quantification of HicA binding to a preformed $HicB_T$ (50 nM)-HEX-S1-2 complex (7.5 nM), $HicB_T$ refers to the tetramer concentration of HicB, ($HicB_M$ = 200 nM). (B) Quantification of HicA binding to a preformed $HicB_T$ (10 nM)-HEX-S1-2 complex (7.5 nM). Here $HicB_M$ = 40 nM. For each titration experiment, the proportion of substrate displaced by increasing concentrations of HicA was calculated for three independent repeats via equation (10). For each experiment the mean value is plotted with error bars representing the SEM. Standard errors of K_d values were calculated in GraphPad Prism.

The competition experiment was repeated with $HicB_T$ (10 nM) and HEX-S1-S2 (7.5 nM), corresponding to a ~ 80% protein-DNA complex[1] (Figure 5.4B). Titration of HicA approaching a 1:1 stoichiometry of HicA:HicB (0-40 nM) and >40 nM caused HEX-S1-2 to raidly dissociate from HicB and full dissociation was observed between 200-300 nM HicA. Modelling of this data to equation (10) determined the IC_{50} as 63.5 ± 0.9 nM (95% confidence interval = 57-71 nM), indicating that an excess of HicA is required to drive

[1] Work by Huang[361], determined that competition assays should use a protein-DNA complex between 50-80% bound, so that titration of the competitor directly affects the bound population of protein, rather than the free population of protein. The concentration of protein that corresponds to a 80% protein-DNA complex was determined from previous HicB-S1-2 FA experiments using 7.5 nM HEX-S1-2 DNA.

dissociation of HicB from a HicB-HEX-S1-2 complex at 80% bound. The calculated Hill coefficient was 2.64 ± 0.33. Data for the 95% bound HicB-S1-2 competition assay gave an IC_{50} and hill coefficient value of 330 ± 1.8 nM (95% confidence interval = 317-343 nM) and 7.1. This IC_{50} should be taken with caution due to the excess of free HicB in solution, which resulted in an approximately constant anisotropy value until an equimolar ratio of HicA:HicB (200 nM HicA: 200 nM HicB) was reached. Further titration of HicA resulted in a sharp decrease in anisotropy which resulted in the artificially high Hill coefficient value of > 7.

A K_i value for the HicA dissociation of HicB from DNA can be modelled through a variety of equations (Table 5.1). The Cheng Prussof equation[362] (Equation 11) can be used, however this equation relies on the assumption that free inhibitor equals the total inhibitor present, thus neglecting the possibility of bound and free inhibitor components. Due to this simplification, the K_i is often overestimated by 10 fold. Munson and Rodbard[363] applied an exact correction (y_0) to this equation by accounting for the ratio of bound:free ligand (Equation 12) for cases where the Cheng Prussof equation is not appropriate. A K_i prediction server BotBD[364] can also predict the K_i through a series of equations (13-17) based on the free concentration of the ligand, rather than the ratio of bound:free ligand used in the Munson and Rodbard correction.

	80% bound ($HicB_T$ 10 nM)
Cheng Prussof	18.1 ± 1.88 nM
Munson and Rodbard	10.5 ±1.30 nM
BotDB K_i	12.9 ±1.46 nM

Table 5.1. List of calculated K_i values dependent on a range of equations (Cheng Prussof, Munson and Rodbard correction and BotDB webserver). This was calculated using the IC_{50} value calculated from the 80% HicB-S1-2 titration.

The calculated K_i for 80% bound ($HicB_T$: 10 nM) varied depending on the method used. Both the Munson and Rodbard correction and BotDB calculation were in close agreement for the experimental K_i value (10.5/12.9 nM). In comparison the Cheng-Prussof equation resulted in a larger K_i value (18.1 nM), suggesting an overestimation of the K_i value.

One aspect that neither of these equations takes into account is cooperativity within the system. The HicAB system does not appear to adhere to a simple single site competition model and calculations have not considered the presence of intermediate HicAB-DNA complexes or cooperativity between HicA and HicB. We cannot currently discount the

importance of possible intermediate HicAB complexes and further work such as the determination of high-resolution three-dimensional structures are needed to investigate their role in S1-2 binding. Determination of K_i values using these equations without accounting for cooperativity may be inadequate and the precise apparent K_i value varies on the specific mode of analysis used (Table 5.1). Rather than rely on K_i values to quantitatively describe the dissociation of HicB-DNA by HicA, the IC_{50} value obtained at 80% bound (63.5 ± 0.9 nM), rather than 95% bound (330 ± 2 nM) is a more reliable value to describe the HicA driven dissociation of the HicB-S1-2 complex.

A single $HicA_4HicB_4$ complex cannot simultaneously bind both sites (S1 and S2) and HicA therefore actively drives the dissociation of HicB from S1-S2. Due to the fact that both RHH domains are intact within the hetero-octameric HicAB complex (Figure 5.5), albeit 70 Å apart, it is possible that HicAB could bind to either S1 or S2 through a single RHH domain (Subunit 1-3 or Subunit 2-4). However, EMSA evidence previously determined that both S1 and S2 sites need to be occupied by one tetramer (or two individual RHH domains), to result in tight binding (Figure 4.1B, 4.2C). Two independent hetero-octameric HicAB complexes cannot individually bind to S1 and S2, as the R_g of one $HicA_4HicB_4$ complex is 31.90 Å (Appendix Table 18), there would be large steric clashes between two $HicA_4HicB_4$ molecules and S1-2 DNA.

De-repression of HicB-DNA binding has been observed in 3 separate HicAB family members. Both *B. pseudomallei* and *Y. pestis* HicB moieties contain a RHH domain, in contrast to *E. coli* HicB which contains a HTH domain. It is hypothesized that the mechanism of complex formation: rotation of DNA binding domains in response to HicA might be a generalised event for HicAB family members irrespective of their specific DNA binding domain.

Crystal structures of HicB and HicAB reported in this work shed light on the mechanism of de-repression. As determined in chapter 4, residues R94, N96 and S98 are critical for S1-2 binding by HicB and they project a positive surface of charge optimal for DNA binding (Figure 5.5A, B). Addition of HicA (Figure 5.5C, D, grey) to form a symmetrical $HicA_4HicB_4$ complex causes a 90 ° rotation of the RHH domains and disrupts the patch of positive surface charge observed in the unbound conformation of HicB. This is believed to be an allosteric mechanism of regulating DNA binding through induced conformational changes. This does not follow the same mechanism of de-repression as observed for MqsA,

speculated by Turnbull and Gerdes as the mechanism for HicAB de-repression.[214] The HicA and DNA binding sites of HicB are on separate binding domains and do not overlap, as opposed to MqsA.[223] EMSA experiments here are in agreement with observations by Turnbull and Gerdes that an excess of $HicA_{EC}$ causes de-repression of the $HicB_{EC}$-DNA complex[214], however additional FA experiments suggest that de-repression occurs almost immediately upon the addition of HicA, albeit at a low level until an equimolar ratio of HicA:HicB is reached. Experimental data provided within this chapter implies that formation of this hetero-octamer actively prevents HicB binding to S1-2 DNA. However, further functional and structural studies on other HicAB pairs are required to determine how generalisable this mechanism of de-repression is.

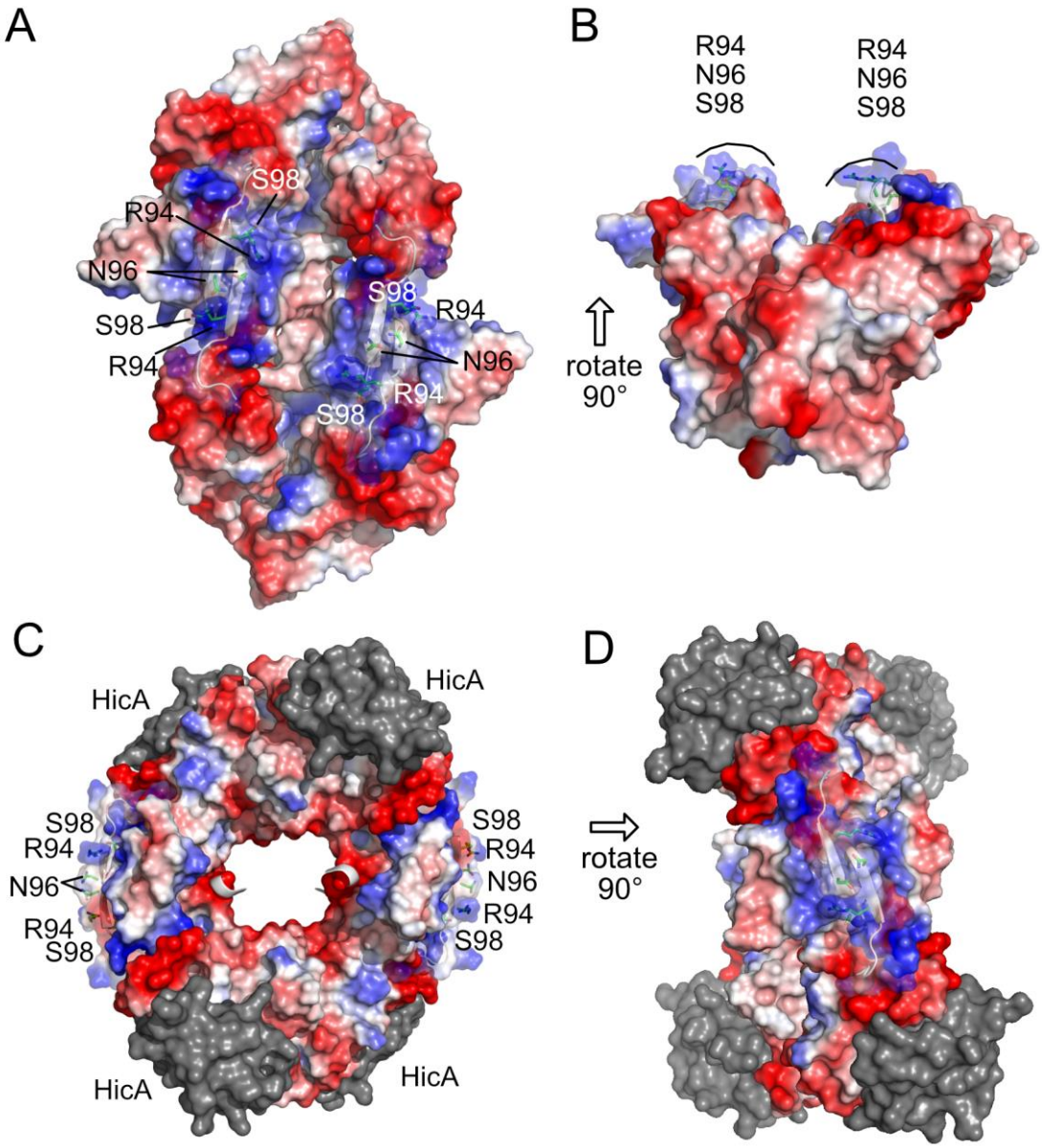


Figure 5.5. Surface net charge of HicB and HicAB. (A) Surface representation of HicB showing clustering of positive charges on one face mapped to R94, N96 and S98 of the C-terminal domain of HicB. (B) A 90° rotation of the structure in panel (A). (C) Surface representation of HicAB highlighting perturbation of the positively charged patch of the C-terminal domain (R94, N96 and S98) due to rotation of the two RHH domains. HicA is represented as grey to emphasize the surface charge of HicB. (D) A 90 ° rotation of the structure in panel (C). For each image, positive charge, negative charge and neutral charge are highlighted in blue, red and white.

5.2. Comparisons to $HicAB_{SP}$

The *S. pneumoniae* HicAB hetero-octameric complex recently reported[272] does not resemble the open conformation of HicAB reported here (Chapter 3). As the three-dimensional structure of free $HicB_{SP}$ has not been reported, it is unknown whether conformational changes regarding the C-terminal RHH domains occur upon formation of a $HicAB_{SP}$ complex.

However, the $HicAB_{SP}$ complex has a positively charged DNA binding patch (Figure 3.13) and $HicA_{SP}$ appeared to induce a modest increase of $HicB_{SP}$:DNA binding affinity (K_d 8 μM for $HicB_{SP}$:DNA, 4 μM $HicAB_{SP}$:DNA), inferring its activity as a co-repressor rather than a de-repressor. This is in stark contrast to work reported here and within *E. coli* and *Y. pestis* where HicA did not increase the affinity of HicB to DNA. However, as a titration experiment of $HicA_{SP}$ with $HicB_{SP}$-DNA was not reported, it is not possible to conclude whether an excess of $HicA_{SP}$ causes de-repression for this system.

6. Conditional Cooperativity

A generalised model for the transcriptional regulation of most type 2 TA systems has been introduced in Chapter 1 (Appendix Table 22). It is termed conditional cooperativity and is thought to be the mechanism by which cell populations undergo bistable switching between normal (antibiotic susceptible) and persistent (antibiotic tolerant) states.[200-202] This regulation is typically dependent on two factors: first, the presence of several adjacent operator sites within the regulatory region upstream of the toxin-antitoxin operon and secondly the ability of the toxin to act as a co-repressor at a low toxin:antitoxin ratio and a de-repressor at high toxin-antitoxin ratios within the cell. A common misconception is that conditional cooperativity is a consequence of the toxin:antitoxin ratios simply changing within the cell. As type 2 antitoxins are synthesized at a greater rate than their cognate toxin, under normal growth conditions the toxin:antitoxin ratio is low.[199] Proteolytic degradation of antitoxins over time[120] coupled with the high stability of toxins eventually establishes a high intracellular toxin:antitoxin ratio, even if the specific TA system does not adhere to conditional cooperativity.

6.1. Molecular mechanisms of Conditional Cooperativity

Full conditional cooperativity has been observed and demonstrated in CcdAB, Kid-Kis, MazEF, ParDE, Phd-Doc, RelBE and VapBC systems where multiple operator sites are present (Appendix table 22). Apart from ParD, Kid and VapB, the antitoxins contain an intriniscally disordered region (IDR) and the toxin-antitoxin complex, aside from ParDE, displays an alternating array of toxin and antitoxin modules.

In 2014 Garcia-Pino and Loris proposed three distinct mechanisms to explain conditional cooperativity.[365] These include steric exclusion between non-repressing TA complexes, low-high affinity switches of toxin-antitoxin interaction sites and allosteric communication between antitoxin domains. The RelBE systems adheres to the first mechanism (steric exclusion) whereby a high affinity antitoxin-toxin-DNA complex is competitively dissociated by an excess of the toxin RelE (Figure 6.1A).[155] In contrast Phd-Doc[227] (Figure. 6.1B) and CcdAB[125] both display a switch between the toxin using both high and low affinity binding

sites on the antitoxin to exclusively occupying high affinity binding sites that results in conditional cooperativity. Phd (antitoxin) is also allosterically regulated by Doc to form a structured DNA binding domain upon the formation of the initial Phd-Doc complex.[210, 227]

The HicAB system described here falls into a class of TA systems including DinJ/YafQ[83] (Figure. 6.1C) and MqsRA[223](Figure 6.1D) where the antitoxin can bind an individual operator site and where far less is known about the mechanism of de-repression. It has been proposed that in a subset of these systems, where conditional cooperativity operates (e.g. in DinJ/YafQ it does not) it does so in a simplified form in that an excess of toxin is still effectively countered at the transcriptional level.[365] Although the MqsRA operator contains two binding sites, MqsA is able to independently bind to one of these and the toxin MqsR solely acts as a de-repressor to the MqsA-DNA interaction via a steric mechanism, rather than a co-repressor.[223] For HicAB, sub-stochiometric levels of HicA have no discernible effect on DNA binding and an excess of HicA causes dissociation of the HicB-DNA complex through large scale conformational reorganisation of HicB to form a hetero-octameric complex (Figure 6E) rather than via overlapping DNA/toxin binding sites as observed for MqsRA.[223] This allosteric regulation of interdomain interactions has been proposed as a mechanism for this single site conditional cooperativity, but not demonstrated until now.

While the two endpoints of HicB in a DNA binding and non-binding form have been determined, further intermediate states are likely to exist (Figure 6.1E). Indeed we have observed $HicA_1$-$HicB_4$ and $HicA_2$-$HicB_4$ complexes by native mass spectrometry and AUC (Chapter 3) and the *Y. pestis* $HicA3_2$-$HicB3_4$ complex has previously been observed by size exclusion chromatography multi-angle light scattering (SEC-MALS)[132]

It is proposed that conversion of HicB to the hetero-octameric HicAB complex may follow one of two possible pathways (Figure 6E).

In (1), binding of sub-stoichiometric levels of HicA can be accommodated by to N-terminal binding sites on HicB (Subunits 2 and 3) that are predicted to be surface exposed and require no conformational rearrangement of HicB. This would have little or no effect on the thermodynamic equilibrium of HicB-DNA binding and is in agreement with the DNA binding asays reported under these conditions which show little change in the K_d (3 nM v 1.9 nM). However in accordance with previous studies we were unable to confirm the formation of a stable $HicA_{1/2}$-$HicB_T$-DNA complex during EMSA competition experiments.[214] Further

titration of HicA does however compete with DNA for HicB binding. Four molecules of HicA can be accommodated by HicB, but only when HicB adopts an open conformation that does not bind DNA and promotes highly cooperative binding of HicA.

The equilibrium therefore will be influenced by the affinity of DNA for HicB and the stabilisation of the closed conformation of HicB versus the affinity of HicA for HicB and stabilisation of the open, non-binding DNA binding form of HicB.

In pathway (2), association of two HicA molecules might stabilise the open non-DNA binding form of HicB and exposure of the additional toxin binding domains is the basis for binding of further toxin. This model seems unlikely given that HicA:HicB ratios of 0.5:1 still showed significant tight DNA binding, suggesting at ratios of HicA:HicB <1:1, the open non-DNA binding form of HicB is not significantly populated. However, at this stage, we cannot exclude this pathway or the possibility that there is a slight equilibrium between the two forms of HicB upon binding of two HicA molecules.

Pathway (1) appears to the most plausible mechanism based on the structural and biochemical data presented. The essentially cooperative mechanism is similar to related systems[148], with the exception that de-repression of HicB-DNA binding is highly cooperative at high concentrations of toxin. However, further studies are required to determine the intervening mechanistic steps that underpin the dissociative mechanism of HicB:DNA in the presence of varying concentrations of HicA as well as potential structural and functional diversity amongst the HicAB family.

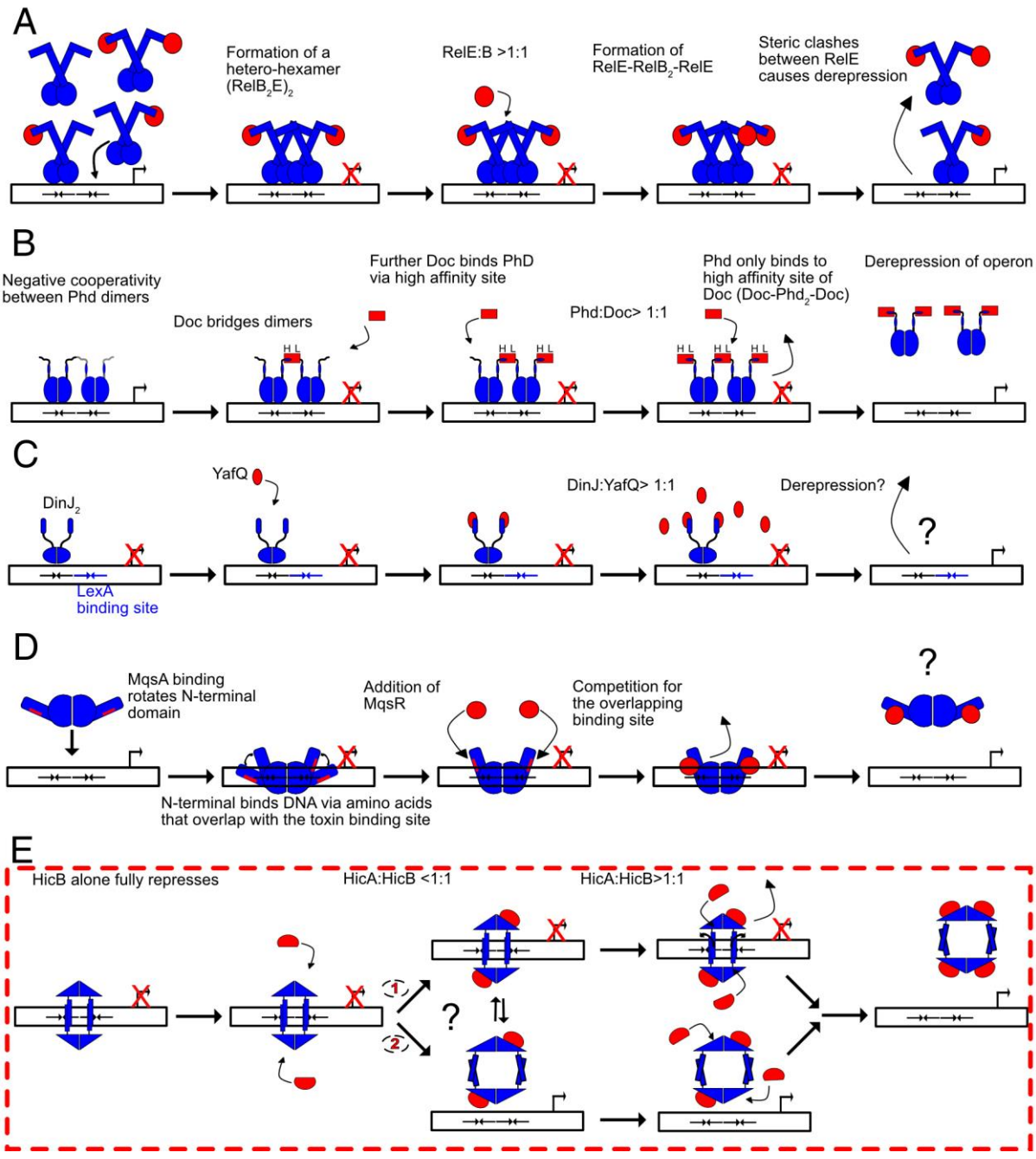

A
Formation of a hetero-hexamer ($RelB_2E)_2$
RelE:B >1:1
Formation of RelE-$RelB_2$-RelE
Steric clashes between RelE causes derepression
B
Negative cooperativity between Phd dimers
Doc bridges dimers
Further Doc binds PhD via high affinity site
Phd:Doc> 1:1
Phd only binds to high affinity site of Doc (Doc-Phd_2-Doc)
Derepression of operon
C
$DinJ_2$
YafQ
LexA binding site
DinJ:YafQ> 1:1
Derepression?
?
D
MqsA binding rotates N-terminal domain
Addition of MqsR
Competition for the overlapping binding site
?
N-terminal binds DNA via amino acids that overlap with the toxin binding site
E
HicB alone fully represses
HicA:HicB <1:1
HicA:HicB>1:1
1
2
?

Figure. 6.1. Overview of de-repression models for RelBE, Phd-Doc, DinJ-YafQ, MqsRA and HicAB. For each panel, the toxin is highlighted in red and the antitoxin in blue. (A) The $RelB_2$ dimer or $RelB_2E$ complex can bind to one of the two adjacent DNA operator sites, but it is proposed that the formation of a W-shaped hetero-hexameric complex $(RelB_2E)_2$ may occupy both adjacent DNA operator sites simultaneously to confer full transcriptional repression of the *relBE* operon. Excess toxin (RelE) binds a second site of a RelB dimer forming a rigid $RelB_2E_2$ hetero-tetramer, two hetero-tetramers cannot simultaneously bind both operator sites due to steric hindrance resulting in transcriptional de-repression. Transcription of the *relBE* operon returns RelB and RelE to stoichiometric levels (~1:1). (B) Doc forms a hetero-pentameric complex with Phd (Phd_2-Doc-Phd_2) by binding to Phd_2 via low (L) and high (H) affinity sites to confer full repression of the *phd-doc* operon by binding two operator sites. Doc allosterically regulates Phd to form a structured DNA binding domain to ensure full transcription repression only occurs upon formation of a Phd-Doc complex. An excess of Doc preferentially binds Phd solely through H sites resulting in the formation of a rigid heterotetramer (Doc-Phd_2-Doc) that cannot occupy both operator sites due to steric clashes between adjacent heterotetramers. In contrast both DinJ-YafQ (C) and MqsRA (D) do not follow the model of conditional co-operativity, as both toxins act as de-repressors rather than co-repressors. (C) $DinJ_2$ fully represses its operator via a single palindromic site. Addition of YafQ forms a YafQ-$DinJ_2$-YafQ hetero tetrameric complex, but an excess of YafQ does not result in de-repression of the DinJ/YafQ-DNA complex and the de-repression mechanism is unknown. It has been demonstrated that LexA can bind to the second operator site and repress transcription in vivo, but it is uncertain if this occurs in vitro. (D) Likewise, MqsA fully represses in the absence of MqsR. MqsR competes with an overlapping DNA binding site of MqsA and formation of a proposed hetero-tetrameric MqsRA complex results in de-repression (the published MqsRA complex is a partial model). MqsA exclusively binds to MqsR or DNA it cannot bind both simultaneously. (E) Like MqsA and DinJ, HicB alone results in saturation of the palindromic sequences (S1-2). I hypothesize that binding of HicA is likely to result in two possible pathways. (1) HicA binds the surface exposed α1 helices of subunit 2 and 3 to form an intermediate $HicA_2HicB_4$ complex that does not result in an increase of affinity to S1-2. At concentrations of HicA > HicB, there is binding of a further two HicA molecules that results in the 90° rotation of the ribbon-helix-helix motifs prevents binding to the palindromic sequences and dissociation of HicB from DNA. (2). Binding of 2 HicA molecules causes rotation to the symmetrical HicAB complex. At concentrations of HicA>HicB, additional HicA molecules bind to stabilise the symmetrical HicAB complex. The intermediate steps of this pathway are unknown and either route, or an equilibrium between the two cannot be discounted as of present.

6.1.1. Reverse gene order and non-canonical transcriptional regulation

The canonical gene organisation of type 2 TA systems is where the antitoxin gene precedes the toxin gene within the bicistronic operon. With the exception of DinJ-YafQ[83], these TA systems adhere to conditional cooperativity. The DinJ-YafQ system is somewhat distinct as the promoter region contains LexA controlled operator site immediately adjacent to the operator site that DinJ exclusively binds (Figure 6.1B).

However, the HicAB, HigBA, HipAB and MqsRA systems are organised in reverse (toxin gene followed by antitoxin gene). Evidence to date suggests that individual type 2 TA systems with a reverse gene organisation are transcriptionally regulated by novel mechanisms rather than conditional cooperativity. This was suggested by Loris[365] but to date there has been no review or a generalised link between a reverse gene organisation and a novel transcriptional mechanism. This is in part due to the fact that the molecular mechanisms of de-repression have been understudied and only recently discussed at a molecular level for

HicAB (This work[214]) and MqsRA.[223] The mechanisms for HigBA and HipAB are unknown.[217,365-367]

V. cholera chromosomally encoded HigBA does not adhere to conditional cooperativity and instead adheres to a postulated novel anti-cooperativity mechanism, the details of which are unpublished. It appears that HigA binding to HigB-DNA causes negative cooperativity resulting in transcriptional de-repression at high toxin-antitoxin ratios.[366-368] In direct contrast the *P. vulgaris* HigBA system encoded on the Rts1 plasmid appears to adhere to conditional cooperativity, as the toxin HigA is a co-repressor.[215, 369, 370] The same observation has been observed for the *A. baumanni* HigBA system encoded on the PAB120 plasmid.[216] This implies that the gene organisation within the bicistronic operon is likely not the sole determinant for regulation by conditional cooperativity or a separate mechanism. There is a possibility based on findings on the HigBA systems that the genetic environment, where the system is located (chromosomally/ plasmid encoded) may play a factor in the regulation of the systems.

It is likely that the HipAB system of *E. coli* and *S. oneidensis* do not adhere to conditional cooperativity as the toxin does not act as a co-repressor.[43, 217] Recently Schumacher[137] determined that the toxin HipA was able to form direct interactions with DNA, a novel interaction not observed in other type 2 TA systems. It has not yet been determined if transcriptional de-repression occurs for this system.

An interesting observation links the regulation of type 2 TA systems via conditional cooperativity to the formation of an alternating T-AT complex (Appendix Table 22). For type 2 TA systems that do not form an alternating complex (DinJ-YafQ, HicAB, HigBA, HipAB, FitAB, MqsRA, ParDE and VapBC), with the exception of the ParDE and VapBC systems, hallmarks of conditional cooperativity (co-repression and de-repression) have not been demonstrated. However, *S. enterica* VapBC displayed conditional cooperativity at a toxin:antitoxin ratio of 10:1 in vitro and it is uncertain if this ratio is physiologically attainable.[211]

If TA complexes do not form an alternating pattern of T-AT complexes as observed for RelBE (Figure 6.1A), Phd-Doc (Figure 6.1B) and CcdAB, but circular closed complexes[365](Like HicAB, HipAB and FitAB/VapBC, Appendix Table 22) then it is plausible that these act at a single operator site and are independent from any adjacent

operator regions due to their separation. No steric occlusion is observed between TA complexes bound at adjacent operator sites for these circular complexes. Within the initial TA complexes formed at physiological toxin:antitoxin ratios, all toxin binding sites appear to be occupied, with the exception of *S. enterica* VapBC[2]. The molecular mechanisms for de-repression of these closed complexes and therefore transcriptional control of these TA systems, including those with a reverse gene organisation, are not yet fully understood. Further work is required to investigate the novel transcriptional regulation in other TA systems which do not seem to adhere to conditional cooperativity.

[2] Modelling of *S. enterica* VapBC to *N. gonorrhoea* FitAB (PDB:2H10) suggests that VapBC initially forms a hetero-octamer ($VapB_4C_4$) and upon an excess of VapC (10:1) forms the higher order complex $VapC_2$-VapB-$VapC_2$ [211]. It is uncertain if other VapBC family members adhere to this proposed mechanism.

7. Conclusions and future work

7.1. Conclusions

This work presents for the first time the model of a molecular mechanism of the stoichiometric HicA dependence of HicB binding to DNA. At low ratios of HicA:HicB there is no effect on DNA binding, but at higher ratios the affinity for DNA declines co-operatively, driving dissociation of the HicA:HicB:DNA complex. This has not yet been confirmed at a transcriptional level for *Burkholderia* HicAB, but the phenomenon of HicA driven de-repression been confirmed at a transcriptional level for *E. coli* HicAB[214]. The molecular mechanisms within this system distinctly differ from traditional conditional cooperativity observed in other TA systems and instead appear to represent a novel mode of regulation, distinct from MqsRA[223] and DinJ-YafQ[83] TA systems which also do not adhere to conditional cooperativity.

Three-dimensional crystal structures of HicB and HicAB support this molecular mechanism as HicAB complexation results in the rotation of the C-terminal RHH domains of HicB into a conformation where they cannot bind both sites of the inverted repeat of DNA (S1 and S2). The active site for DNA binding by HicB has been probed and the main interaction site was experimentally determined to consist of residues within the β5 strand (R94, N96 and S98) binding to two repeats [24]ACACA[20] and [35]TGTGT[31] of DNA upstream of the *hicAB* operon. A three-dimensional structure of HicB-DNA is unknown, *ab initio* and *in silico* modelling support a structure where DNA lies flat across the two RHH domains, in agreement with modelling data for the *Streptococcus* HicB-DNA complex.[272]

7.2. Future work

Future work requires the findings presented here to be confirmed at a transcriptional level by a series of LacZ assays, which have been performed for both *E. coli* and *Y. pestis* HicAB systems. Initial crystallisation hits of HicB-S1-2 should be further optimised to obtain a three-dimensional structure and reinforce modelling presented for both *B. pseudomallei* and

S. pneumoniae HicAB systems. A three-dimensional structure would confirm whether the α6 helix within the RHH domain represents a secondary DNA-binding site and R118 and S119 residues within the α6 helix should be mutated and investigated by EMSA and FA to obtain suitable evidence on their involvement and importance in DNA binding. At this stage however, the possibility of multiple DNA binding mechanisms of HicB and HicAB cannot be discounted.

There is no information regarding accurate K_d values on the HicA-HicB interaction for *B. pseudomallei* and other HicAB family members. ITC for example might provide data that provides a fuller thermodynamic description of the full envelope of mechanistic steps associated with HicA dependent dissociation of HicB from DNA, including details on the behaviour of intermediary HicAB/DNA complexes.

The mechanism by which HicA family members binds and cleaves mRNA is poorly understood when compared to other type 2 mRNAses. As the catalytic residue implicated in toxicity has been identified for *B. pseudomallei, E. coli, S. pneumoniae* and *Y. pestis* HicA toxins, structural and functional analysis would yield crucial details regarding their mechanism of action.

Further studies on HicAB family members in other bacteria and archae will enhance knowledge of the system, especially with respect to the structural and functional diversity between systems. The ultimate goal is to understand how the full complement of TA systems within an organism behave. To date this has not been fully elucidated, but evidence and the general concensus within the current literature suggests that TA systems form a highly regulated network that can collectively compute some output state of the cell (presumed to be persistence) from a network of inputs composed of both internally and externally generated stimuli. To date only initial primitive investiations (single TA deletions) have occurred and this is partly due to the emergence of new TA systems on an annual basis. Proteomics and transcriptomics at a whole cell level and studies on the role of TA systems within a whole cell population (for example, how individual cells interact within biofilms) may begin to unravel this network.

8. Materials and Methods

Sources of materials

Laboratory reagents and were purchased from Sigma Aldrich, Thermo Fisher Scientific and Merck Millipore. Isotopically enriched $^{15}NH_4Cl$ was purchased from Goss Scientific. Competent cells (T7 Express and NEB 5-α) were purchased from New England Biolabs or Merck Millipore (Novagen B834 (DE3) and Novagen BL21 (DE3) cells). Kits used throughout the work were the GenElute Plasmid Miniprep Kit, GenElute Gel Extraction kit (Both Sigma Aldrich), KOD Hot Start DNA Polymerase (Merck Millipore) and the PureLink PCR Purification kit (Thermo Fisher Scientific). All enzymes used were purchased from Thermo Fisher Scientific.

Buffers and solutions

Buffer	Procedure	Composition
50x TAE	Agarose gel electrophoresis	242 g Trizma base, 57 ml 100% glacial acetic acid, 100 ml of 0.5 M EDTA, up to 1 L dH_2O
10x TAE	EMSA	400 mM Trizma base, 25 mM EDTA, pH 7.8 (with acetic acid)
Buffer A	IMAC	25 mM Tris-HCl, 0.5 M NaCl, 10% glycerol, pH 8.0
Buffer B	IMAC	25 mM Tris-HCl, 0.5 M NaCl, 10% glycerol, 800 mM imidazole, pH 8.0
Buffer C	SEC	25 mM Tris-HCl, 150 mM NaCl, pH 7.5
Buffer D	SEC	10 mM Na_2HPO_4, 2.2 mM KH_2PO_4, 137 mM NaCl, pH 7.5
Buffer E	EMSA/FA	25 mM Tris-HCL, 100 mM NaCl, 10% glycerol, pH 7.5
Anode buffer	SDS-PAGE	200 mM Trizma base, pH 8.9
Cathode buffer	SDS-PAGE	100 mM Trizma base, 3.5 mM SDS, 100 mM Tricine, pH 8.25
Gel buffer	SDS-PAGE	3 M Trizma base, 10 mM SDS, pH 8.45
Loading buffer	SDS-PAGE	780 mM Trizma base, 142 mM β-mercaptoethanol, 35 mM SDS, 10% glycerol, 140 mM bromophenol blue, pH 6.8
10% Separating solution	SDS-PAGE	2.5 ml 40% acrylamide, 2.8 ml dH_2O, 3.33 ml Gel buffer, 50 µl 10% AMPS, 20 µl TEMED
4% Stacking solution	SDS-PAGE	0.5 ml 40% acrylamide, 3.3 ml dH_2O, 1.25 ml Gel buffer, 50 µl 10% AMPS, 20 µl TEMED
Coomassie stain	SDS-PAGE	2.5 g Commassie Blue (G250), 450 ml ethanol, 100 ml glacial acetic acid, 450 ml dH_2O
Solution A	ESMS	50/50 acetonitrile: dH_2O
Solution B	ESMS	0.1% trifluoroacetic acid (TFA)
Solution C	ESMS	5% methanol, 0.1% TFA
Solution D	ESMS	80% acetonitrile, 0.1% TFA

Table 8.1. Buffers and recipes used throughout the work used for all protein constructs. Where appropriate the buffer pH was adjusted using 1 M NaOH or 1 M HCl.

Media	**Ingredients**
Luria Broth (LB)	Tryptone (10 g/L), NaCl (10 g/L) and yeast extract (5 g/L)
M9 Minimal media pH 7.5	KH_2PO_4 (3 g/L), Na_2HPO_4 (1.8 g/L) and NaCl (0.5 g/L).
SelenoMet™ medium base	21.6 g in 1L dH_2O
Additives for media	
1000 x Trace Metal Elements	$CaCl_2$ (20 mM), $CoCl_2.6H_2O$ (2 mM), $CuCl_2.2H_2O$ (2 mM), H_3BO_3, HCL (60 mM), $MnCl_2.4H_2O$ (10 mM), $Na_2MoO_4.2H_2O$ (2 mM), $Na_2SeO_3.5H_2O$ (2 mM), $NiCl_2.6H_2O$ (2 mM), $ZnSO_4.7H_2O$ (2 mM)

Table 8.2. Media and additive recipes. For 1000 x Trace Metal elements, parenthesis correspond to the working concentration in solution.[371]

Sterile technique

Where applicable, both media and solutions were sterilised by autoclaving at 121 °, 151 psi for 15 minutes. Solutions that could not be autoclaved (Antibiotics and glucose solutions) were filter sterilised via a 0.2 μM filter (Thermo Fisher), while buffer solutions were filter sterilised via a 0.2 μM cellulse nitrate filter (Whatman, GE Healthcare) and degassed prior to their use.

Preparation of LB Agar plates

LB Agar plates containing carbenicillin (100 μg/ml) or kanamycin (100 μg/ml) were prepared as previously described.[372]

Agarose gel electrophoresis

Agarose gels were prepared in 1x TAE buffer as previously described[372] with the addition of 4 μl Midori Green (NIPPON Genetics). PCR samples (2 μl) were mixed with 2 μl of 5x loading dye (Bioline) before loading into wells. Gels were run at 150 V for 30 mins before visualising under UV light.

Preparation of HicA and HicB plasmids.

Preparation of pOPINE-HicB_FL

PCR amplification: Primers were designed in-house and synthesized (IDT Technologies™) (Appendix Table 4, HicB_FL Primer 1 and 2). *hicB* was amplified from K96243 genomic DNA with KOD Hot Start DNA Polymerase (Table 8.3) using a peqSTAR 96x Universal Gradient (VWR) 0.2 ml PCR tubes were heated at 95 °C, 3 minutes, then 30 cycles of denaturation (95 °C, 30 seconds), annealing (temperature gradient: 38-50 °C, 15 seconds) and extension (70 °C, 30 seconds). A final extension of 70 °C for 10 minutes was applied before storing at 4 °C DNA amplification was confirmed by agarose gel electrophoresis. After confirming the amplification of the *hicB* gene, samples containing the gene were further purified by a PCR Purification Kit according to the manufacturer's protocol.

Component	Volume (μl)	Final concentration
10x buffer KOD Hot start DNA polymerase	5	1X
25mM $MgSO_4$	3	1.5 mM
2mM dNTPs	5	0.2 mM
Filter sterilised dH_2O	32	N/A
5' Primer	1.5	0.3 μM
3' Primer	1.5	0.3 μM
Template DNA	1	N/A
KOD DNA polymerase	1	0.02 Units/μl
Total volume	50	

Table 8.3. Reagents for a single (50 μl) PCR reaction using KOD DNA polymerase. Specific primers outlined in Appendix Table 4 were used to amplify each construct

Preparation of pOPINE vector: 50 μl pOPINE vector (100 ng/μl, provided by Dr Marisa Till, University of Bristol) was digested using NcoI and PmeI (5 μl each) and 10 μl 10x FastDigest Green Buffer (Thermo Scientific) overnight (37 °C). The digest reaction was analysed by 1% agarose gel before visualisation and isolation of the linearised vector using a scalpel under UV light. This was then isolated using the GeneElute gel extraction kit following the manufacturers protocol.

Ligation of *hicB* into pOPINE: A 10 μl ligation reaction mixture contained 1 μl digested pOPINE plasmid, 3 μl gene insert (*hicB*), 4 μl dH_2O and 2 μl 5x In-fusion® HD enzyme premix.[275] This was incubated at 50 °C for 20 minutes before 2.5 μl of this mixture was added to 50 μl NEB 5-α cells. Cells were incubated on ice (30 mins), before a heat shock step (42 °C, 30 seconds) and 200 μl LB media was added before incubation at 37 °C, 200 rpm for 1 hour. Cells were plated on carbenicillin-LB agar plates containing 400 μg/ml X-gal and 100 mM IPTG for blue/white colony screening. Plates were incubated overnight (37 °C) before storing at 4 °C.

pOPINE-HicB_FL culture growth: A single colony of NEB 5-a cells containing the full length *hicB* gene (white colonies) was added to 10 ml LB media containing 100 μg/ml carbenicillin. Cells were grown for 16 hours (37 °C, 200 rpm), before centrifugation (5,000 x g, 10 mins). The plasmid was extracted using the GenElute Plasmid Miniprep Kit according to the manufacturer's protocol and sequenced by GeneWiz™.

This procedure outlined the protocol for the cloning of the full length *hicB* gene into pOPINE vector (pOPINE-HicB_FL). This exact protocol was used to generate pOPINE-HicB and pOPINE-HicB_NT vectors using respective primers (Appendix Table 4, HicB Primer 1 and 2, HicB-NT Primer 1 and 2).

Preparation of HicB methionine mutants

Whole plasmid mutagenesis was used to introduce methionine point mutations into the pOPINE-HicB plasmid. This involves an adapted procedure of PCR amplification (outline above, Table 8.3) with the addition of 5% DMSO to the reaction mixture and the use of mutagenic primers (I99M Primer 1 and 2). 0.2 ml PCR tubes were heated at 95 °C, 5 minutes, then 30 cycles of denaturation (95 °C, 30 seconds), annealing (temperature gradient: 57-63 °C, 15 seconds) and extension (70 °C, 6.5 minutes) then a final extension period as previously described. Following confirmation that the plasmid had been amplified (agarose gel electrophoresis), DNA samples were pooled and digested with DpnI overnight (50 μl product, 5 μl DpnI, 5 μl 10 x Tango Buffer) to remove parental template DNA before purifying using the PCR Purification kit. 4 μl plasmid was transformed into 50 μl NEB-5a cells as outlined above, plasmids were then isolated for individual colonies and sent for sequencing. Sequencing confirmed that one set of primers (I99M Primer 1 and 2) had generated the pOPINE-I51M, pOPINE-I99M and pOPINE-DM vectors.

Preparation of HicB DNA mutants

The nucleotide sequence of HicB bearing specific individual mutations (R94A, N96A, R94A/N96A N96Q, and S98A) was codon optimised for *E. coli*, then synthesized and sub-cloned into a pMA-T vector via SacI/KpnI recognition sites by Thermo Fisher Scientific as whole plasmid PCR mutagenesis to generate these mutants were unsucessful. The resultant freeze-dried vector (5 μg) was resuspended in 50 μl dH_2O. The plasmid was miniprepped and then digested using KpnI and SacI as previously described. The digested DNA product was then isolated as previously described.

Isolated HicB genes were amplified with Mutant Primer 1 and 2 by PCR as previously described (Table 8.3) and the resultant product was further purified and subcloned into digest pOPINE vector as following the same procedure detailed earlier for HicB_FL. Sequencing confirmed the presence of each gene in the pOPINE vector. The gene had to be digested out of the pMA-T vector as PCR amplification of the *hicB* gene using the pMA-T vector as a template resulted in a misprimed PCR product that did not contain the correct sequence of the *hicB* gene bearing a specific mutation.

pOPINE-R94E and pOPINE-S98T vectors were generated by whole plasmid mutagenesis by Philip Marsh (University of Bristol) using the pOPINE-R94A and pOPINE-S98A vectors as template plasmid and R94E Primer 1/2 and S98T Primer 1/2. The protocol used was identical to that previously described for generation of methionine mutants, except the annealing temperature was modified to (47-52 °C).

Preparation of HicA.

The nucleotide sequence of *Burkholderia hicA* was codon optimised for *E. coli*, then synthesized and subcloned into a pET151-D/TOPO vector by Thermo Fisher Scientific. The resultant plasmid was isolated as previously described.

Plasmid	Property	Source
pET26-b-HicA	pET26-b bearing a NT PelB sequence, His_6-tag, Enterokinase cleavage site and *hicA* under a T7 promoter, KanR	Aaron Butt
pET151-HicA.6His	pET151 D/TOPO bearing *hicA* with a cleavable NT His_6-tag under a T7 promoter, AmpR	Thermos Scientific®
pET26-b HicB	pET26-b bearing a NT PelB sequence, His_6-tag, Enterokinase cleavage site and *hicB* under a T7 promoter, KanR	Aaron Butt
pOPINE-HicB_FL	pOPINE bearing the full length *hicB* gene preceding a His_6-tag under a T7 promoter, AmpR	This work
pOPINE-HicB	pOPINE bearing *hicB* with a CT deletion of 4 amino acids preceding a His_6-tag under a T7 promoter, AmpR	This work
pOPINE-I51M	pOPINE-HicB containing a I51M mutation	This work
pOPINE-I99M	pOPINE-HicB containing a I99M mutation	This work
pOPINE-DM	pOPINE-HicB containing both I51M and I99M mutations	This work
pOPINE-HicB_NT	pOPINE bearing a truncated *hicB* with a CT deletion of 54 amino acids preceding a His_6-tag under a T7 promoter, AmpR	This work
pOPINE-R94A	pOPINE-HicB containing a R94A mutation	This work
pOPINE-R94E	pOPINE-HicB containing a R94E mutation	This work
pOPINE-N96A	pOPINE-HicB containing a N96A mutation	This work
pOPINE-N96Q	pOPINE-HicB containing a N96Q mutation	This work
pOPINE-R94A/N96A	pOPINE-HicB containing both R94A and N96A mutations	This work
pOPINE-S98A	pOPINE-HicB containing a S98A mutation	This work
pOPINE-S98T	pOPINE-HicB containing a S98T mutation	This work

Table 8.4. List of plasmids generated during this work. AmpR and KanR refer to ampicillin/kanamycin resistance

Protein Expression in *E. coli*.

Expression of HicA and HicB constructs.

1 μl of the appropriate plasmid (Table 8.4) was transformed into 25 μl T7 Express cells and plated on LB-Agar plates containing an appropriate antibiotic. Overnight pre-cultures: A single colony was added to 200 ml LB containing 100 μg/ml of appropriate antibiotic and incubated overnight (37 °C, 200 rpm). Overnight pre-cultures were diluted 1:100 in 1.5 L LB media supplemented with 100 μg/ml of appropriate antibiotic and incubated (37 °C, 150 rpm) until approximately OD_{600nm} = 0.7. Cells were induced with 0.5 mM IPTG (HicA) or 0.3 mM IPTG (HicB) and the temperature was reduced to 25 °C and incubated overnight (16 hours, 25 C, 150 rpm). Due to the low yield associated, HicA was grown in 5 L batches.

Expression of isotopically labelled HicA

1 μl of the pET151-D/TOPO plasmid was transformed into 25 μl Novagen BL21 (DE3) cells following the protocol detailed above. Overnight Pre-culture: A single colony was added to 200 ml LB media supplemented with 100 μg/ml carbenicillin and incubated overnight (16 hours, 37 °C, 200 rpm). Successive overnight pre-cultures were diluted 1:100 in 4 L LB media supplemented with 100 mg/ml Carbenicillin and incubated (37 °C, 150 rpm) until OD_{600nm} = 2.0. Cells were harvested (5,000 x g, 10 mins) and then washed by resuspending the cell pellets in 200 ml M9 minimal media and centrifuging (5,000 x g, 10 mins). The cell pellet was then resuspended in 1 L M9 minimal media supplemented with 100 μg/ml carbenicillin, 1x trace metal elements, 1 mM $MgSO_4$, 0.7 mM glucose, 13.5 mM glycerol and 1g/L $^{15}NH_4Cl$. Cells were equilibrated for 1 hour at 25 °C before induction with 0.5 mM IPTG, for 16 hours at 150 rpm.

Expression of SeMet labelled pOPINE-DM

1 μl of pOPINE-DM was transformed into 25 μl B834 (DE3) cells as detailed above and used to inoculate 100 ml SelenoMet™ medium base supplemented with 40 mg/L methionine, SelenoMet™ Nutrient Mix (Molecular Dimensions) and 100 μg/ml carbenicillin and incubated overnight (37 °C, 200 rpm). After incubation overnight at 37 °C, the cells were harvested and the pellet was washed in 200 ml dH_2O three times through resuspension and

centrifugation (5,000 x g, 10 mins) before resuspension in 1L pre-heated and pre-aerated (37 °C, 150 rpm, 1 hour) SelenoMet™ medium base supplemented with 100 μg/ml carbenicillin, SelenoMet™ nutrient mix and 40 mg/L selenomethionine (Molecular Dimensions). Cells were incubated (37 °C, 150 rpm) until an OD_{600nm} = 0.5, where an additional 20 mg/L selenomethionine was added and the temperature was reduced to 25 °C. Cells were allowed to equilibrate to 25 °C for 1 hour and then induced with 0.5 mM IPTG and left to incubate overnight (25 °C, 16 hours, 150 rpm).

Protein extraction

After 16 hours, cells were harvested by centrifugation (5,000 x g, 10 mins) and resuspended in 30-40 ml Buffer A (Table 8.1), then immediately purified or stored at -20 °C. Resuspended pellets were then sonicated in a VCX130 Vibra-Cell, (Sonics) for 15 mins: pulse on (5 s), pulse off (5 s) before centrifugation (17,000 x g, 30 mins). The resulting supernatant was collected and further purified.

pET26-b-HicB purification

The cell lysate containing His_6-tagged pET26-b-HicB was purified by IMAC on a 5 ml HiTrap™ column (GE Healthcare) pre-loaded with 10 ml 100 mM $NiSO_4$ which was connected to a GE Healthcare Äkta FPLC and equilibrated in Buffer A. Following sample loading, the column was washed with 6% buffer B (5 column volumes) before the protein was eluted with an imidazole gradient (6-100% Buffer B) over 10 column volumes. The protein was collected in 1 ml fractions usinga GE Healthcare fraction collector (Frac-950) and analysed by SDS-PAGE. Fractions containing the protein of interest were pooled and destalted into buffer A using a HiPrep™ 26/10 Desalting column (GE Healthcare). The His_6-tag was removed using bovine enterokinase (Sigma, 1 unit per 1 μg protein) at 16 °C for up to 16 hours before analysis via SDS-PAGE As cleavage was unsuccessful, isolated pET26-b HicB following IMAC was injected onto a HiLoad™ 26/60 Superdex 75 Prep grade column (GE Healthcare) that had been equilibrated in Buffer C before the protein was concentrated by ultracentrifugation to 7 mg/ml (GE Healthcare Vivaspin, MWCO 10 kDa, 5,000 x g).

pOPINE-HicB purification

All pOPINE HicB, HicB mutants and N-terminal HicB constructs were purified using the same procedure. Briefly, cell lysates containing pOPINE HicB were purified by Ni^{2+} IMAC following the procedure above and immediately further purified using a HiLoad™ 26/60 Superdex 75 Prep grade column running Buffer C before concentration by ultracentrifugation to 20-30 mg/ml (GE Healthcare Vivaspin, MWCO 10 kDa, 5,000 x g). The purification protocol for SeMet labelled HicB-DM was identical except for the addition of 2 mM TCEP to all purification buffers.

HicA purification

The cell lysate containing HicA was purified by IMAC as detailed for pET26-b HicB with minor changes. Following IMAC, the protein was desalted into buffer A and cleaved with TEV protease (1:100) for 16 hours at 4 °C. The cleavage reaction was then passed through a Ni^{2+} IMAC to remove the His_6-tagged HicA and His_6-TEV before protein (HicA) was desalted into buffer C using a HiPrep™ 26/10 desalting column and concentrated by ultracentrifugation to 4 mg/ml (GE Healthcare Vivaspin, MWCO 3 kDa, 5,000 x g).

Purification of ^{15}N HicA was identical except after negative purification, untagged HicA was exchanged into Buffer D for NMR analysis.

Protein Detection

SDS-PAGE

10% Tris-Tricine SDS-gels were prepared following a previously describing protocol.[373] Samples for analysis were prepared in a 1:1 (v/v) ratio with loading buffer before heating (100 ° C, 10 mins). The gel was run for 100 V, 1.5 hours and then stained in Coomassie brilliant blue stain for 1 hour. Gels were destained in H_2O through heating and agitating before visualisation on an E-gel imager (Thermo Fisher Scientific).

Denatured Electrospray Mass Spectrometry

Ziptips™ (Merck Millepore) were used to prepare protein solutions for mass spectrometry (typically at a 1 mg/ml concentration) using solutions A, B, C and D (Table 8.1) and multiple aspirations (10 μl). Briefly, 2 aspirations of solution A were used to wet the Ziptip, before equilibration in 0.1% TFA with 5 aspirations of solution B. The protein was bound to the matrix with 15 aspirations before washing with 15 aspirations of solution C. The final protein sample was eluted with a final aspiration with solution D. Samples were analysed by ESMS on the Synapt G2-Si (Waters), which is coupled to a Triverse Nanomate (Advion), (an automated chip based spraying device), with the source operating in the positive ion mode. Samples were sprayed using a capillary voltage of 1.5 kV. Data was acquired over a 500 - 4000 m/z range for 1 minute, before data analysis was performed with MassLynx 4.1 (Waters).

Native Mass Spectrometry

20 μM samples of HicB and HicAB were dialysed into 100 mM ammonium acetate (AmAc) [342] using 3.5 kDa Slide-A-lyzers (ThermoFisher) for 4 h prior to analysis on a Synapt G2-Si using the parameters described above. To acquire drift spectra, the ion mobility cell was filled with nitrogen gas at 0.5 mbar. Parameters used during data acquisition were adapted from those reported by [374]: sample cone: 80 V, bias voltage: 35 V, trap collision energy: 10 V, transfer collision energy: 5 V, ion wave velocity: 400 m/s, wave height: 5 V. The TOF tube vacuum pressure was maintained at 9.5 $x10^{-7}$ mbar. Data was acquired over a 500- 4000 m/z range for 10 minutes, before data analysis was performed with MassLynx 4.1. For analysis of the HicAB complex, proteins were mixed at equimolar ratios (HicA:HicB = 1:1, 20 μM:20 μM ($HicA_4HicB_4$)) for 1 hour at 16 °C prior to dialysis.

Protein quantification

The DeNovix DS-11 UV-Vis Spectrophotometer was used to measure the absorbance of protein at 280 nm. The concentration of protein was calculated by the Beer-Lambert Law and the extinction coefficient of the protein was calculated from the amino acid sequence[375] using ProtParam[376]: pET26-b HicB (16,960 M^{-1} cm^{-1}, 280 nm, H_2O), HicB_FL (16,202 M^{-1} cm^{-1}, 280 nm, H_2O), HicB and all mutants (15,470 M^{-1} cm^{-1}, 280 nm, H_2O), His_6-HicA (6,990 M^{-1}

cm^{-1}, 280 nm, H_2O) and HicA (5,500 M^{-1} cm^{-1}, 280 nm, H_2O). Triplicate microvolume measurements were used to assay the accuracy of the reported value.

Crystallisation and data collection

pET26-b-HicB (350 μM, $HicB_T$ = 87.5 μM), HicB_FL (500 μM, $HicB_T$ =125 μM), HicB (600 and 1,500 μM, ($HicB_T$ = 150 and 375 μM)), HicB-NT (1,000 μM, ($HicB\text{-}NT_D$ = 500 μM)), HicAB (preformed at a 1:1 ratio, HicA (500 μM): HicB (500 μM) for 1 hour at 16 °C) and HicB-S1-2, preformed at a 1:1.2 ratio ($HicB_T$:S1-2 = 125 μM:150 μM, 250 μM:300 μM and 375 μM:450 μM) for 1 hour at 16 °C were crystallised via sitting drop vapour diffusion at 293 K with 0.5 μl protein to 0.5 μl reservoir solution (50 μl reservoir) using a ArtRobbins Crystal Phoenix™ crystallisation robot and 96 well MRC2 well crystallisation plates (Hampton Research). The sparse screens used are described in the text. Initial sparse matrix hits were optimised via hanging drop vapour diffusion in 24 well VDX plates (Hampton Research) by varying the temperature, ratio of protein: crystallisation buffer and the pH and salt/precipitant of the crystallisation buffer. Individual crystals were looped using appropriately sized litholoops (Molecular Dimensions) and cryoprotected in up to 40% (v/v) glycerol before flash freezing in liquid N_2 for data collection using PILATUS detectors at Diamond Light Source beamlines (I03 and I04).

Diffraction images were processed and integrated using XDS[278] and Xia2[279] and further analysed using Phenix[281] and CCP4[329] as described in the text where exact methods are described. The following is an overview of each procedure. SeMet labelled HicB-DM: The Crank2 pipeline within the jsCoFe interface[297] was used to yield a partial model from the dataset. HicB-NT: Phaser-MR[280] was used for MR using the model derived from SeMet labelled HicB-DM and the Autobuild pipeline[305]. Manual model building and further refinement was undertaken using Coot[303] and Phenix. HicB: Phaser MR using the HicB-NT structure as a search model and further manual model building and refinement built the HicB tetramer. HicAB: MOLREP[328] within CCP4 was used for molecular replacement using the C-terminal domain dimers of HicB as a search model, before manual rebuilding and MOLREP was used to rebuild the N-terminal domains of HicB and place the HicA molecules (PDB:4C26) into the electron density before refinement using REFMAC5[301] as described in Chapter 3. Final models were refined using Molprobity[307] and PDB_Redo sever[308] and then analysed using Pymol[377].

Biophysical techniques

Analytical Size Exclusion Chromatography

A Superdex 75 10/300 GL pre-packed column (GE Healthcare) was equilibrate with Buffer C prior to use and calibrated using 1 mg/ml of Aprotinin, Carbonic Anhydrase, Conalbumin, Ovalbumin, Ribonuclease A and Blue Dextran (GE Healthcare) in Buffer C to generate a calibration curve as previously described[378] (Equation (1)). Protein samples of HicA, $HicB_T$ constructs (50 μM) and a HicAB complex preformed at a 1:1 ratio (50 μM each, where $HicB_T$ = 12.5 μM) and incubated at 25 °C (1 hour) were analysed.

$$K_{av} = \frac{V_e - V_o}{V_t - V_o} \quad (1)$$

Here the particle coefficient value (K_{av}), used to plot the calibration curve is calculated through the elution volume of the sample (V_e), the total column volume (V_t) and the void volume of the column (V_0).

Analytical Ultracentrifugation

All analytical centrifugation experiments were carried out by myself and Hannah Crocker/Mariya Gromova with the aid of Dr Guto Rhys and Dr Antony Burton (Both University of Bristol).

Sedimentation Velocity Analytical Ultracentrifugation

Sedimentation velocity analytical ultracentrifugation (SVAUC) experiments were carried out at 20 °C in a Beckman Beckman Optima XL-A analytical centrifuge (An-60 Ti rotor). 410 μl solutions at OD_{280nm} = 1 (70 μM $HicB_T$, 50 μM HicAB, preformed at a 0.5:1 HicA: HicB ratio ($HicA_2HicB_4$)) and 420 μl Buffer C were loaded ino a two-channel aluminium sedimentation velocity cell containing quartz windows. Absorbance scans were recorded across a radial range (5.85-7.25 cm) at five minute intervals during centrifugation (50,000 rpm) until a total of 100 scans. The resultant absorbance scans were fitted to a continuous c(s) distribution model in SEDFIT[337] with 95% confidence interval by modelling the sedimentation distribution profile to the Lamm equation[379] (2) to determine the sedimentation coefficient distribution (c(s)). Within the Lamm equation the Diffusion coefficient (*D*) and

Sedimentation coefficient (*S*) are described by equations (3-6)[337, 338, 341] and are used to compute the molecular weight (Equation 6). This fitting procedure fit the baseline, meniscus, frictional coefficient (f/f_0), systematic time-invariant and radial-invariant noise were fitted. RMSD values obtained for the fit of samples were 0.009 for both HicB and HicAB. Parameters for fitting: Buffer density/viscosity and the partial specific volume ($\bar{\upsilon}$) of the analysed particle were calculated using SEDNTERP (http://sednterp.unh.edu/).[380, 381] Images were made using Gussi.[382]

$$\frac{\partial c}{\partial t} = D\left[\frac{\delta^2 c}{\partial r^2} + \frac{1}{r}\frac{\partial c}{\partial r}\right] - S\omega^2\left[r\frac{\partial c}{\partial r} + 2c\right] \tag{2}$$

The concentration of the particle (*c*) at a given time (*t*) and radius (*r*) is described by its diffusion coefficient (*D*) as well as the sedimentation coefficient (*S*).

$$\frac{M(1-\bar{\upsilon}p)}{N\left(\frac{f}{f_0}\right)} = \frac{\mu}{\omega^2 r} = s \tag{3}$$

Here $M(1-\bar{\upsilon}p)$ is the effective buoyant molecular mass (*Mb*) and is described by the partial specific volume ($\bar{\upsilon}$) and the solvent density (*p*). *N*: avogadros constant, μ: velocity, $\omega^2 r$: centrifugal force, f/f_0: frictional coefficient.

$$D = \frac{RT}{N\left(\frac{f}{f_0}\right)} \tag{4}$$

Here *D* is related to the frictional coefficient by gas constant (*R*) and absolute temperature (*T*).

$$\frac{dr/dt}{\omega^2 r} = \frac{Mb}{f} \tag{5}$$

S can also be related to the frictional coefficient.

$$\frac{S}{D}=\frac{M(1-\bar{v}p)}{RT} \quad (6)$$

Both *S* and *D* can be related to each other through the Svedberg Equation (6), allowing the indirect determination of the molecular mass[339].

Sedimentation Equilibration Analytical Ultracentrifugation.

Sedimentation equilibrium analytical ultracentrifugation (SEAUC) experiments were conducted at 20 °C in a Beckman-Optima XL-I analytical ultracentrifuge (An-60 Ti rotor).A 110 μl solution at OD_{280nm} = 0.4 (20 μM HicAB pre-formed at a 0.5:1 HicA:HicB ratio ($HicA_2HicB_4$)) and 120 μl buffer C were loaded into a six-channel aluminium sedimentation equilibrium cell containing quartz windows (three protein channels and three reference channels). Samples were radially equilibrated for 16 hours before the samples were centrifuged in the range of 9,000-21,000 rpm, with the absorbance at 280 nm recorded for each channel at each centrifugation speed across a radial range (5.8-7.3 cm). The resultant data were fit to a single ideal species model using Ultrascan II[383], evaluating the molecular mass as a function of the radial position (*r*) within the ultracentrifuge cell (Equation (7)).[340, 341] 95% confidence intervals were determined by Monte Carlo analysis based on 10,000 iterations based on a least-squares statistical analysis.[384]

$$C(r)=C(r_0)\exp\left(\frac{Mb\omega^2\left(r^2-r_0^2\right)}{2RT}\right) \quad (7)$$

Here $C(r)$ is the radial concentration, $C(r_0)$ is the concentration at a designated arbitrary reference distance (r_0).[340, 341]

Electrophoretic Gel Mobility Shift Assay

Oligonucleotides used in EMSA assays (Appendix Table 4) were resuspended to 100 μM in Buffer E and annealed together to form a dsDNA complex (50 μM) by mixing together at 95 °C and left to anneal overnight at 4 °C. Serial dilutions of $HicB_T$ constructs (including DNA mutants) and HicB-NT_D were prepared to give a final concentration between 0.8-3.2 μM in Buffer E. HicA was prepared to give a final concentration of 0.8-50.8 μM in buffer E. For determination of the DNA binding region, 3.2 μM $HicB_T$ was equilibrated at room

temperature with specific DNA fragments described in the text (3 μM) for 30 minutes in Buffer E and loaded onto a 10% Tris-acetate-EDTA (TAE) gel (prepared as previously described[385]) in 1xTAE buffer[385] and run at 100 V (1 hour, 16 C) before staining in ethidium bromide in 1x TAE buffer and visualised under UV using an E-gel Imager. For stepwise binding between the 0-20 bp and 0-40 bp fragments and 16-36 bp to 20-40 bp fragments, the mixture (2 μM DNA, 3.2 μM $HicB_T$) was equilibrated in an identical manner and then loaded onto a 15% TAE gel and analysed in the same manner. Competition assay: A preformed HicB-S1-2 complex (3.2 μM $HicB_T$, 2 μM S1-2 in Buffer E) was challenged with increasing concentrations of HicA (0.8-50.8 μM) and equilibrated for 30 minutes prior to analysis on a 15% TAE gel.

Fluorescence Anisotropy

Fluorescence anisotropy experiments were carried out in collaboration with Dr Oliver Wilkinson and Prof Mark Dillingham (Both University of Bristol).

Fluorescence equilibrium titrations were performed (Jobin Yvon Fluorolog, Horiba Scientific) to determine the dissociation constants (K_d) for HicB, R94E, N96Q, S98T, HicAB, HicA and HicB-NT binding at 20 °C to S1-2 (5'-ATGTGTATAATTACACACAA-3') which harboured the fluorescence dye Hexafluorescein (HEX) at the 5' end (7.5 nM). The excitation wavelength was set at 530 nm and the emission intensity set at 554 nm. Slit widths were set to 5.0 nm for both excitation and emission. Following sample equilibration for 2 minutes, 3 data points with an integration time of 0.5 s were collected for each titration point. Anisotropy was determined by equation (8).[348]

$$r = \frac{I_{II} - I_{\perp}}{I_{tot}} = \frac{I_{II} - I_{\perp}}{I_{II} + 2I_{\perp}} \qquad (8)$$

Titrations were performed in Buffer E in a total volume of 200 μl in a 10 x 2 mm quartz cuvette (Hellma Analytics). HicB, preformed HicAB and HicB-NT were titrated into HEX-S1-2 up to 150 nM. R94E, N96Q, S98T and HicA were titrated until an end point of 1000 nM. For competition assays, a preformed $HicB_T$ (10 nM)-HEX-S1-2 (7.5 nM) complex (80% bound was formed before the addition of HicA to a final concentration of 300 nM or $HicB_T$ (50 nM)-HEX-S1-2 (7.5 nM) complex (95% bound) was preformed before HicA titration (0-1000 nM).

Data fitting

Data were fitted to a single-component binding equation to determine the dissociation binding constant K_d using GraphPad Prism 7.04: equation (9).[349, 350, 352, 386]

$$Y= (A_{DP} - A_D)\left(\frac{[X]+ K_d + n \pm \sqrt{([X]+ n + K_d)^2 - (4[X]n)}}{2n}\right) + A_D \qquad (9)$$

Where A_{DP} is the anisotropy of the fully bound HicB-S1-2 complex, A_D is the anisotropy of unbound S1-2, n is the stoichiometry and [X] is the concentration of protein. For competition assays involving HicA, data were fitted to a four-component inhibition equation to determine the IC_{50} value via equation (10).

$$A= A_D + \frac{(A_{DP} - A_D)}{1 + \frac{[HA]^H}{IC_{50}{}^H}} \qquad (10)$$

Where A is the measured anisotropy, A_D is the anisotropy of free HEX-S1-2, A_{DP} is the anisotropy of the saturated HicB-HEX-S1-2 complex, [HA] is the concentration of HicA, H is the Hill slope.

Ki values were calculated from the Cheng-Prussof equation[362] (11), the Munson derivatisation to the Cheng-Prussof equation[363] (12) and the BotDB web server[364] (13-17).

$$K_i = \frac{IC_{50}}{1 + \frac{D_T}{K_d}} \qquad (11)$$

Where *Ki* refers to the inhibitory binding constant

$$K_i = \frac{IC_{50}}{\frac{1+D_T(y_0+1)}{2*K_d(y_0+2)} + y_0} - K_d * \left(\frac{y_0}{y_0+2}\right) \quad (12)$$

Where y_0 refers to the ratio of bound: free ligand

$$P_0 = \frac{[-(K_d + D_T - P) + ((K_d + D_T - P])^2 + 4PK_d)^{0.5}]}{2} \quad (13)$$

Where P refers to the total protein concentration and P_0 is the free protein at 0% inhibition.

$$PL_0 = P - P_0 \quad (14)$$

Where PL_0 refers to the initial protein-ligand complex

$$PL_{50} = \frac{PL_0}{2} \quad (15)$$

Wheree PL_{50} refers to the bound complex at 50% inhibition

$$L_{50} = L - PL_{50} \quad (16)$$

Here L_{50} refers to the concentration of ligand at 50% inhibition

$$K_i = \frac{IC_{50}}{\frac{L_{50}}{K_d} + \frac{P_0}{K_d} + 1} \quad (17)$$

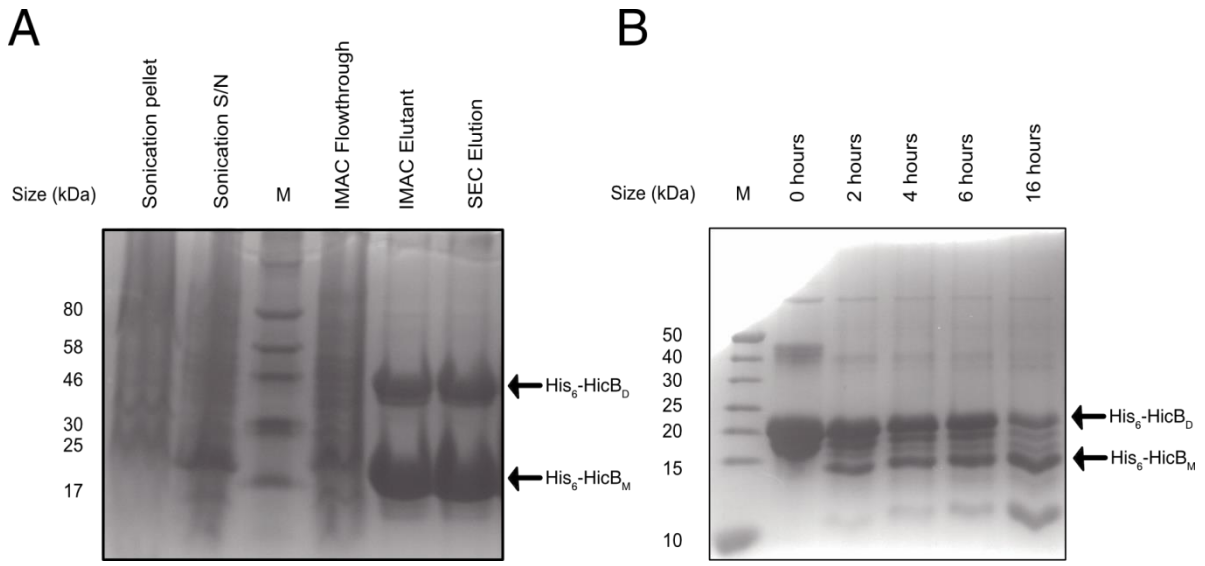
Nuclear Magnetic Resonance

250 μM ^{15}N HicA in Buffer E supplemented with 10% D_2O was transferred to a 3 mm NMR tube and a 2D ^{1}H-^{15}N HSQC was acquired at 25 °C using a Varian VNMRS equipped with a cryoprobe, operating at 600 MHz by Dr Chris Williams (University of Bristol). NMR data was processed within NMRPipe[387] and spectra analysed in CcpNmr Analysis version 2.1.5.[388] The spectrum was assigned by myself using the previous pET26-b HicA assignments.[131]

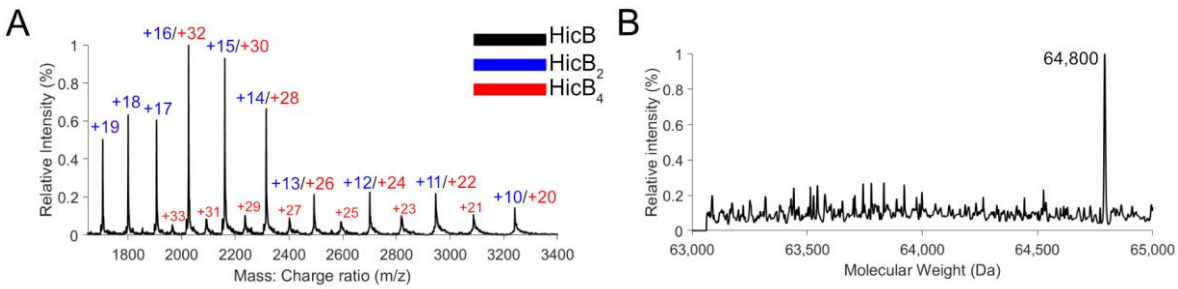
Small angle X-ray scattering

In-line SEC-SAXS for HicB, HicAB and HicB-S1-2 were collected at Beamline 21, Diamond Light Source using an Agilent 1200 HPLC and 2.4 mL Superdex S200 column (GE Healthcare). 50 μL of protein at an appropriate concentration: 500 μM of each HicB variant ($HicB_T$ = 125 μM), HicAB (preformed at a 1:1 ratio, 500 μM HicA: 500 μM HicB ($HicA_4HicB_4$)) and HicB-S1-2 (preformed at a 1:1.6 ratio, 300 μM $HicB_T$: 500 μM S1-2) were loaded (0.04 mL/min) onto a Superdex 200 column (GE Healthcare) running (Buffer C). Frames were collected at 3 seconds per frame at 25 °C and X-ray scattering was recorded (Pilatus 2M detector) with a fixed camera length of 4.014 m, at 12.4 keV. Angular range q data were collected between 0.006-0.042 Å^{-1}. ScÅtter (www.bioisis.net) determined the radius of gyration (R_g), the maximum particle dimension (D_{max}) and the pair distribution function (P(r)).[323, 389] *Ab initio* bead density shape envelope models for each dataset were generated by DAMMIF[390], averaging over twenty three independent runs using the program DAMAVER[391], before a single DAMMIN refinement run.[392] *Ab initio* bead density shape envelope models were superimposed to three dimensional structures of proteins using SUPCOMB.[393] FoXS[325] was used to quantitatively compare the calculated X-ray scattering of crystal structures with the experimental scattering profile of each protein.

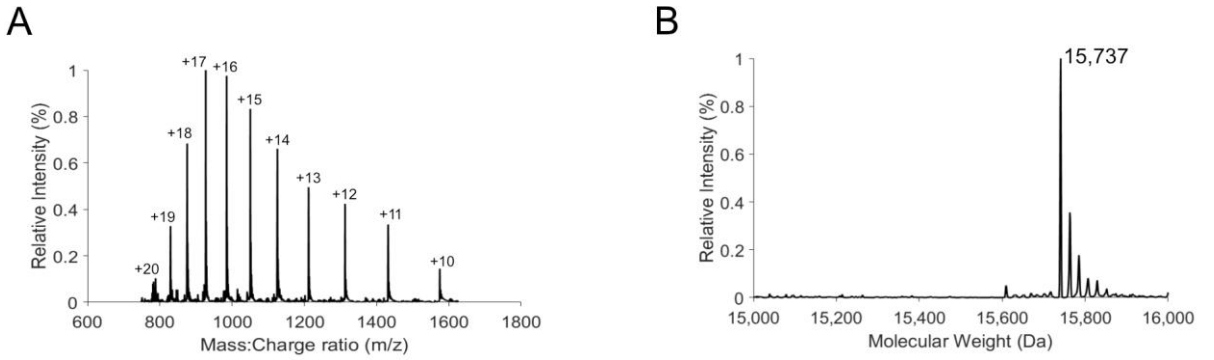
9. Appendix



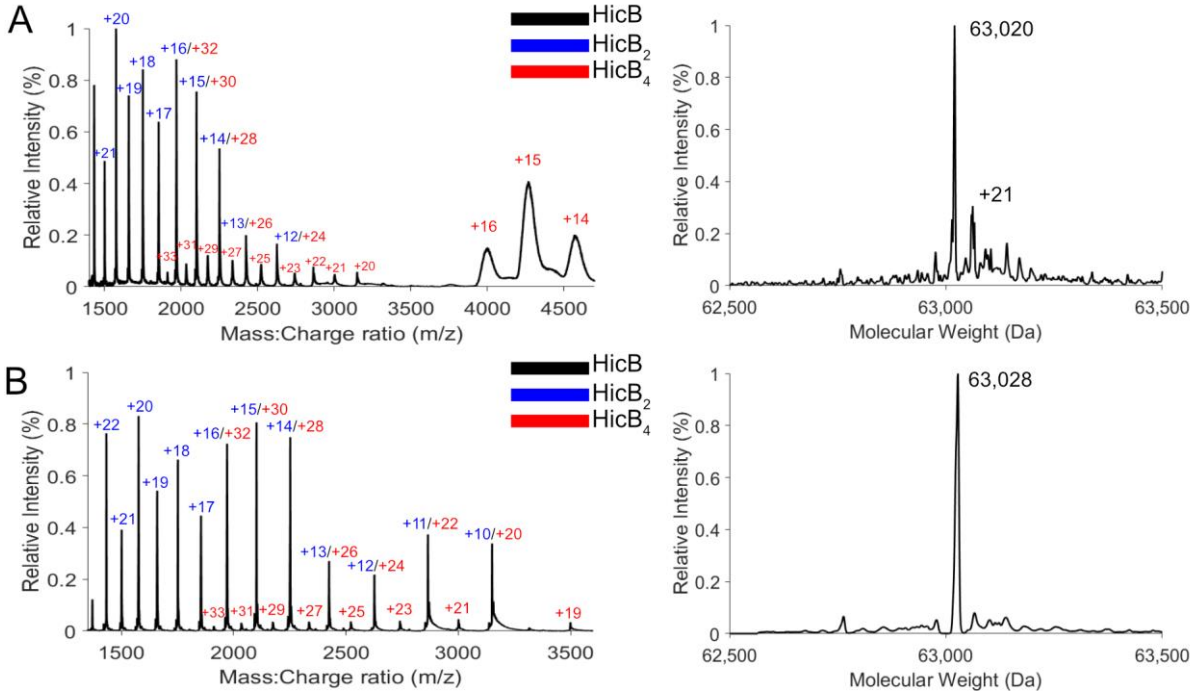
Appendix Figure 1. Purification of pET26-b HicB. (A) SDS-PAGE following the expression and purification of pET26-b HicB from sonication to isolation via SEC. (B) Enterokinase trials to cleave His_6-HicB to HicB that did not result in a specific conversion and instead generated multiple species. Arrows refer to the position of monomeric His_6-HicB (His_6-$HicB_M$) and dimeric His_6-HicB (His_6-$HicB_D$).



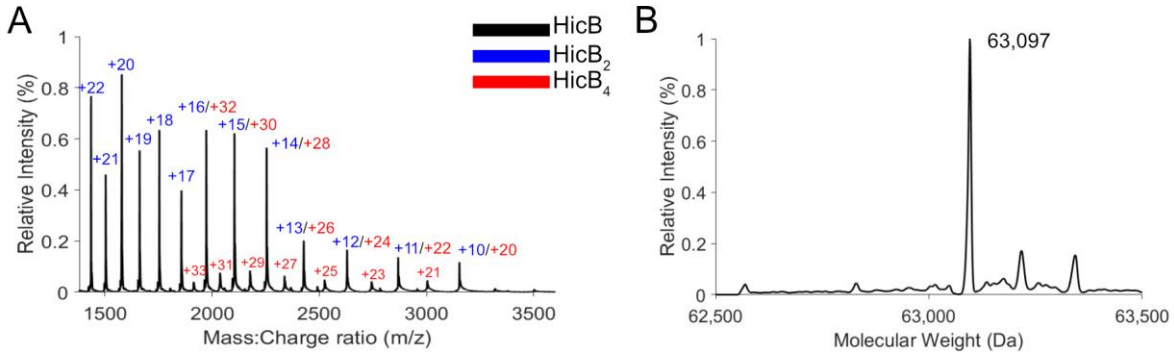
Appendix Figure 2. Native mass spectrometry of HicB_FL. (A) Mass: charge ratio spectra of HicB_FL spanning 1,650- 3,400 mz. Peaks were assigned based on their charge state: dimeric (blue) or tetrameric (red). (B) De-convolution spectrum of the m/z envelope shown in (A) to determine the tetrameric mass of HicB_FL: Observed mass: 64,800 Da (Expected mass: 64,809 Da).



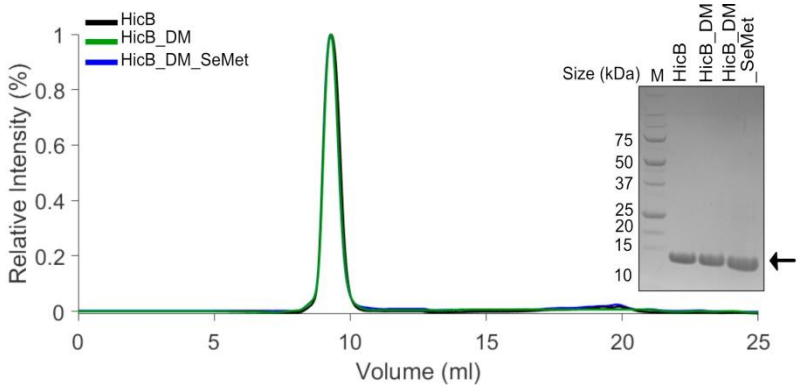
Appendix Figure 3. Denatured ESMS of HicB. (A) Denatured mass/charge spectrum of HicB. Peaks were assigned based on the monomeric charge state (black). (B) De-convolution of the m/z envelope, yielding monomeric HicB: Observed mass: 15,737 Da (Expected mass: 15,738 Da).



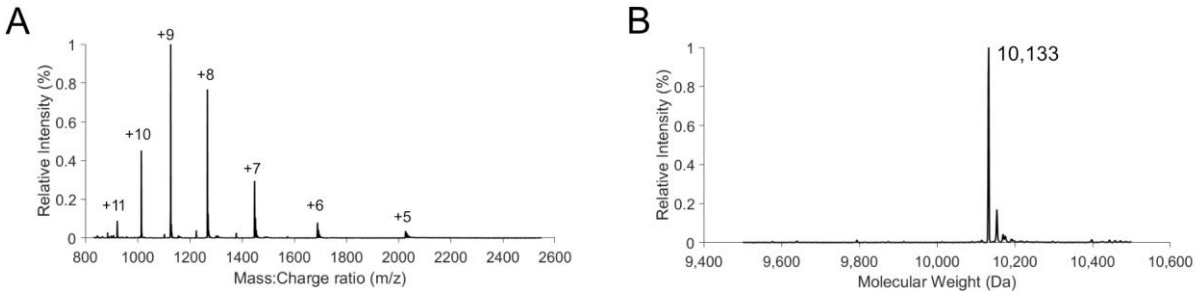
Appendix Figure 4. Native mass spectrometry of methionine mutants I51M and I99M. (A) Native mass spectrometry of I51M showed a similar distribution of m/z charge states to HicB, with deconvolution resulting in a single species (observed mass: 63,020 Da, expected mass: 63,027 Da). (B) Native mass spectrometry of I99M displayed a similar distribution of m/z charge states to HicB, with deconvolution resulting in a single species (observed mass: 63,028 Da, expected mass 63.027 Da). For both native mass spectrometry profiles, a series of m/z charge states of dimeric (blue) and tetrameric (red) states of each mutant are seen.



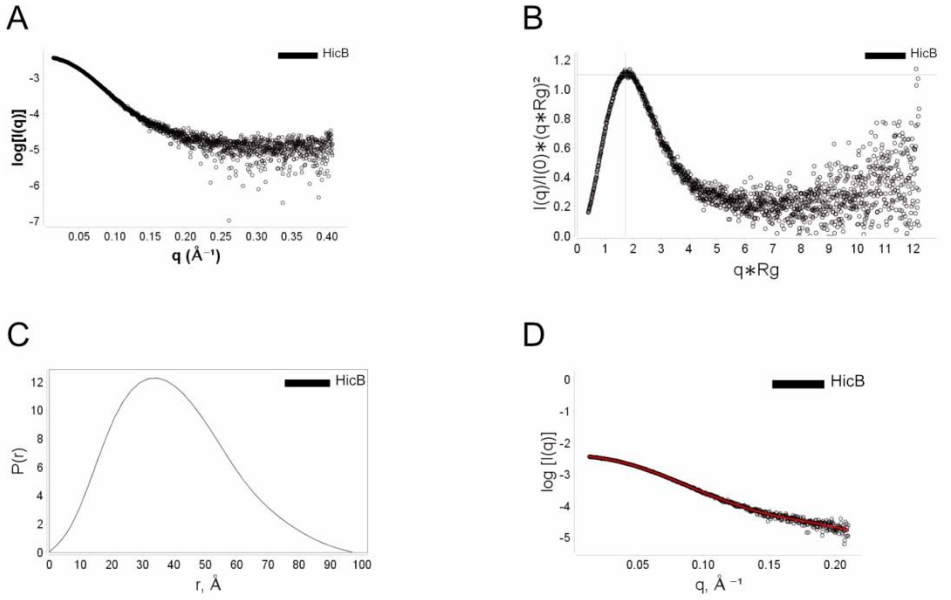
Appendix Figure 5. Native mass spectrometry of HicB_DM. (A) Native mass spectrometry of HicB_DM displayed a similar distribution of m/z charge states to HicB. (B) Deconvolution of the m/z envelope resulting in a single species (observed mass: 63,097 Da, expected mass: 63.099 Da). A series of m/z charge states of dimeric (blue) and tetrameric (red) states of HicB_DM are observed.



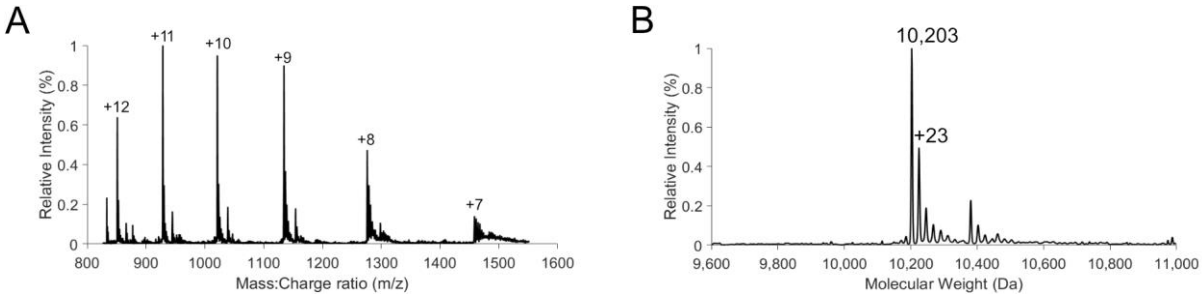
Appendix Figure 6. Analytical SEC profiles of $HicB_T$, HicB_DM and HicB_DM_SeMet. SEC profiles for each construct and an SDS-PAGE of eluted samples. The arrow refers to the band corresponding to monomeric HicB for each construct.



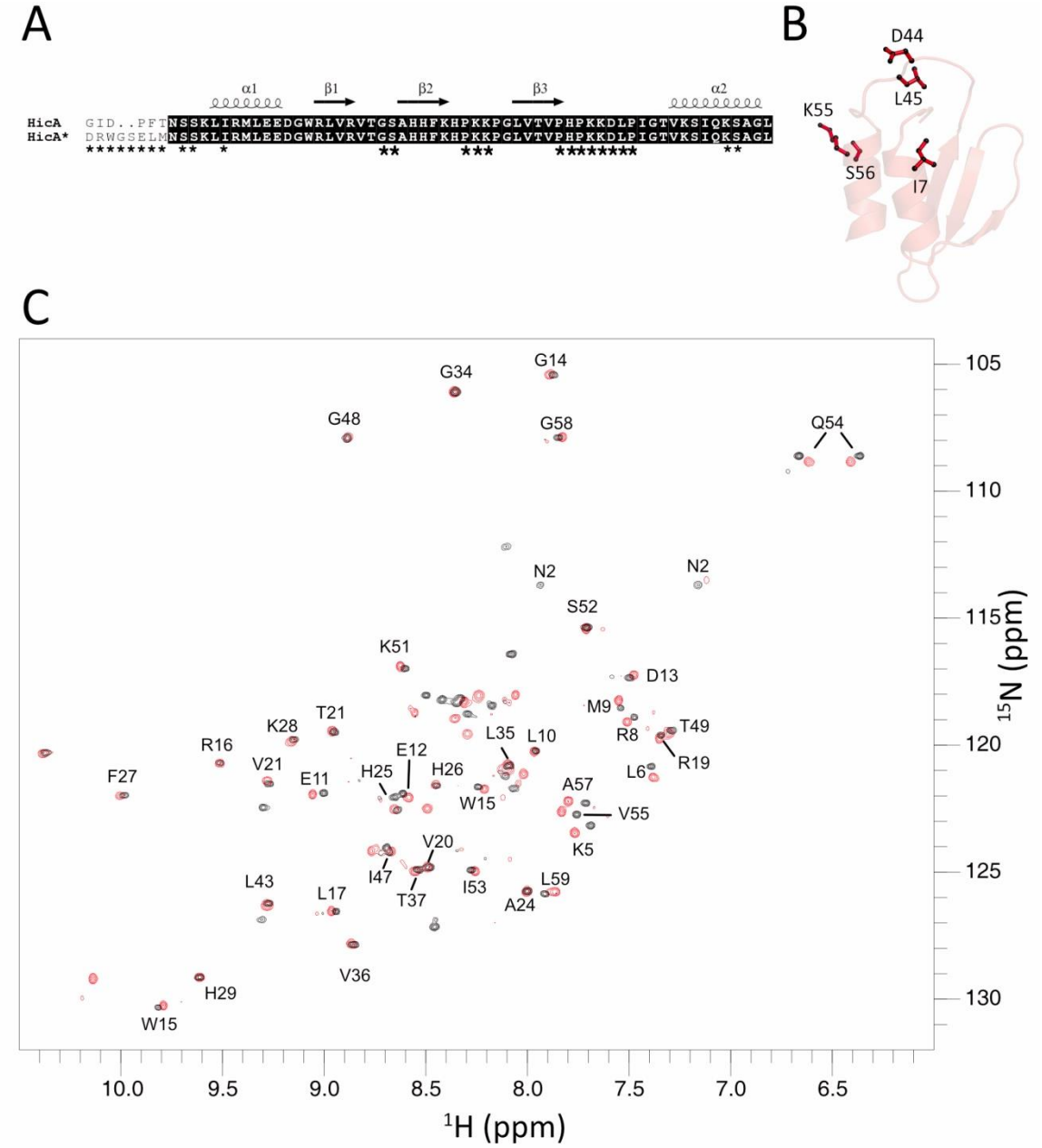
Appendix Figure 7. Denatured ESI mass spectrometry of HicB-NT. (A) Denatured m/z spectrum of HicB-NT. Peaks were assigned based on the monomeric charge state (black). (B) De-convolution of the m/z envelope, yielding monomeric HicB_NT: Observed mass: 10,133 Da (Expected mass: 10,133 Da).



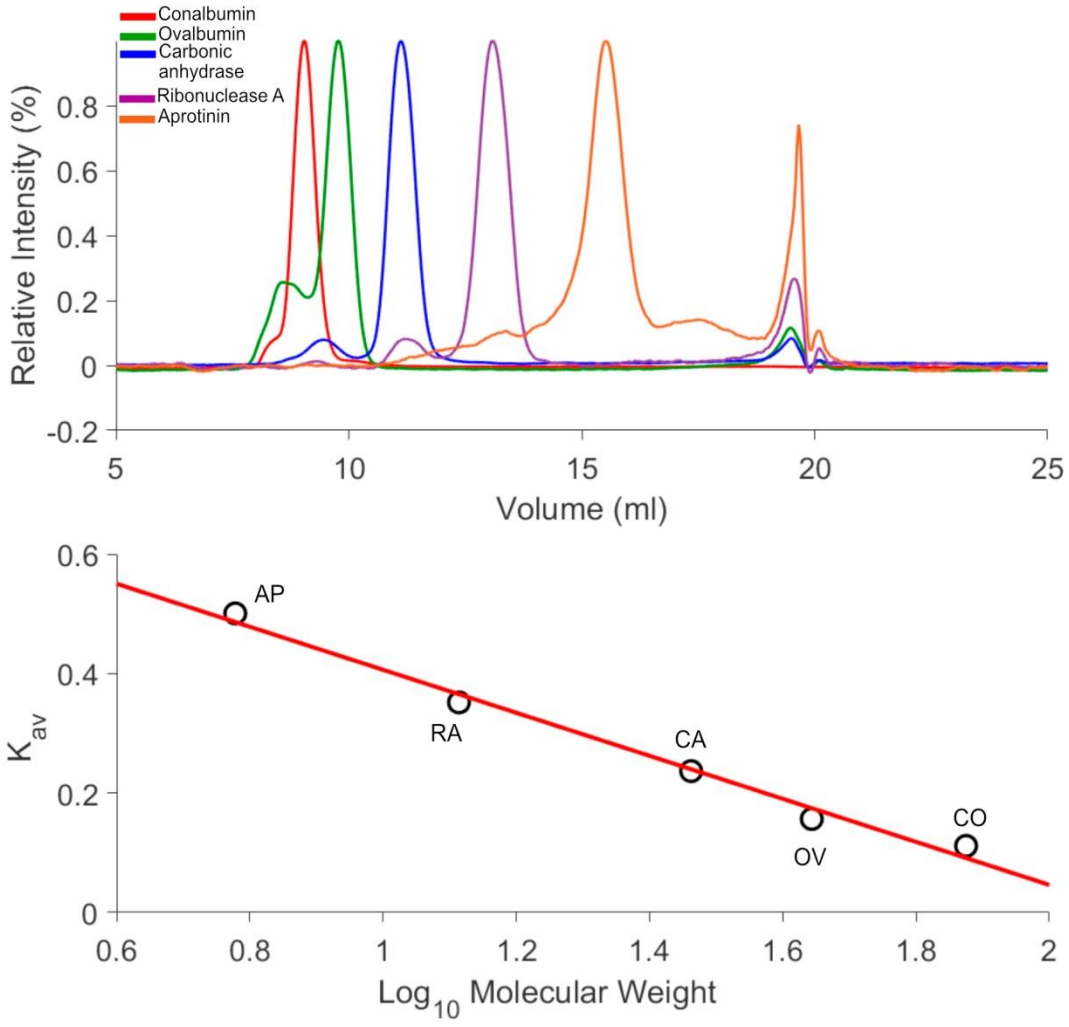
Appendix Figure 8. SAXS of HicB (A) Log_{10} SAXS intensity versus scattering vector (q). (B) Dimensionless Kratky plot. Cross-hair marks the Guinier-Kratky plot (1.732, 1.1), the main peak for folded globular particles. (C) Pair-distance, P(r) distribution function. Maximum dimension, d_{max} is the largest non-negative value that supports a smooth distribution function. (D) Fit of the parameters computed by the P(r) distribution (red line) to the raw scattering data (green).



Appendix Figure 9. Denatured ESI mass spectrometry of His_6-HicA. (A) Denatured m/z spectrum of His_6-HicA. Peaks were assigned based on the monomeric charge state (black). (B) De-convolution of the m/z envelope, yielding monomeric His_6-HicA: Observed mass: 10,203 Da (Expected mass: 10,204 Da).

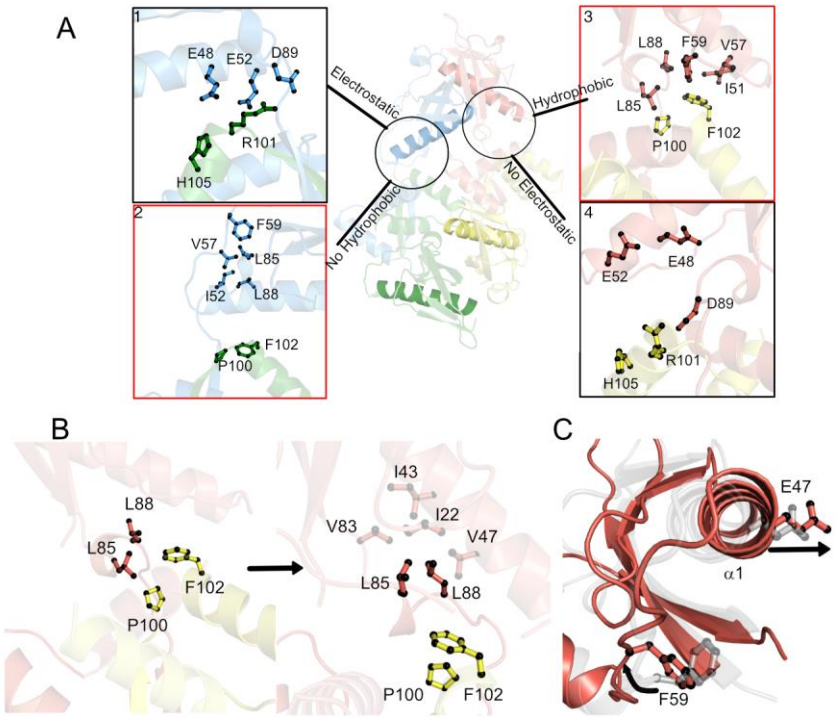


Appendix Figure 10. HSQC of HicA. (A) Sequence alignment between the pET151-HicA construct (HicA) and the pET26b-HicA construct (HicA*). 100% conservation is seen by the presence of a black box. * underneath residues correlate to those that could not be assigned in the ^{1}H-^{15}N HSQC. (B) Cartoon representation of HicA. Residues that could not be assigned within the ^{1}H-^{15}N HSQC spectra of HicA that were assigned in pET26-b HicA are highlighted. (C) ^{1}H-^{15}N HSQC of HicA (black) overlayed onto the previous acquired dataset of pET26-b HicA red).[131]

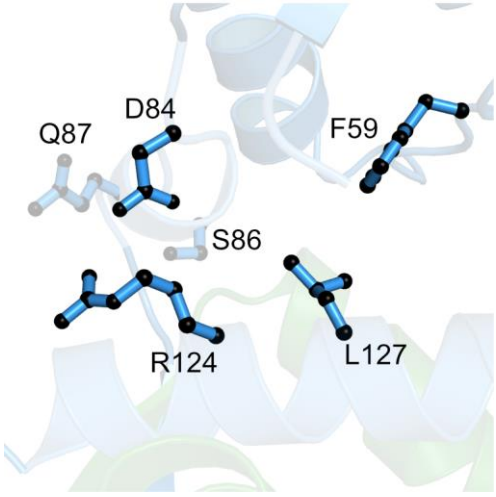


Sample	Mw (kDa)	$Log_{10}Mw$	V_e (ml)	K_{av}
Blue dextran	2000	6.3	7.14	0.00
Conalbumin	75	1.87	9.02	0.11
Ovalbumin	44	1.64	9.78	0.16
Carbonic anhydrase	29	1.46	11.13	0.24
Ribonuclease A	13	1.11	13.07	0.35
Aprotinin	6	0.78	15.58	0.50

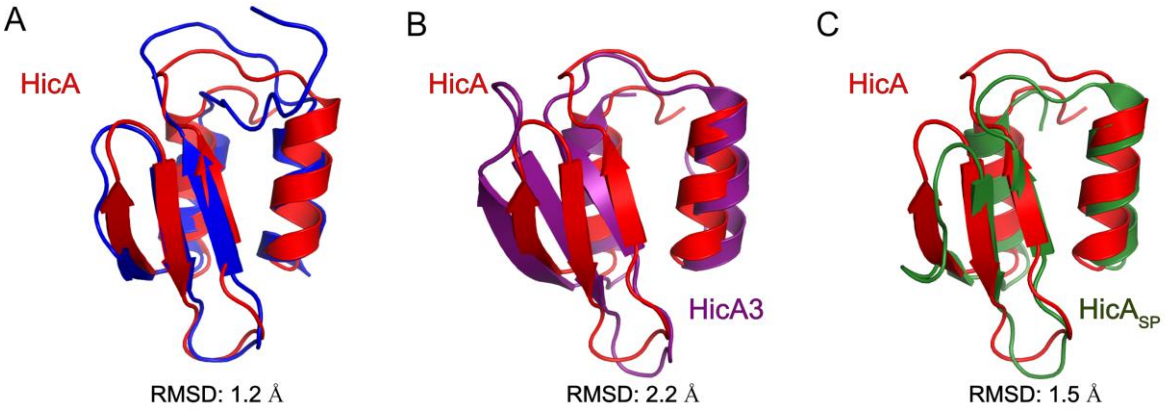
Appendix Figure 11. Analytical SEC profiles of calibrants. (Top) SEC profiles for all reported constructs. (Middle): Calibration curve using known standards: Aprotinin (AP), Ribonuclease A (RA), Carbonic Anhydrase (CA), Ovalbumin (OV) and Conalbumin (CO) used to determine the Mw_{app} for analysed proteins. (Bottom) Summary of results for each calibrant reporting the, Log_{10} Mw, V_e and calculated K_{av} value.



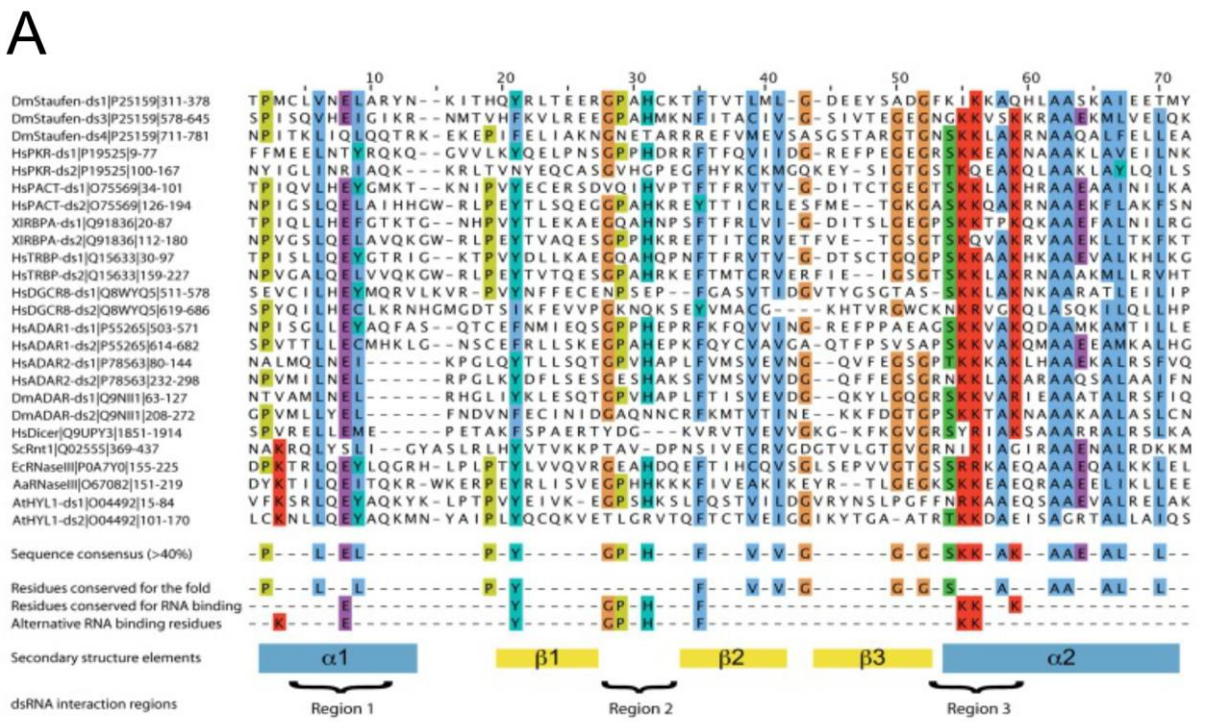
Appendix Figure 12. Disruption of the hydrophobic tethering interaction across adjacent subunits induced in HicB upon binding HicA. (A) Interaction sites across the tetramer with electrostatic and hydrophobic sites highlighted in black and red boxes. Box 1 highlights the electrostatic interaction site (Residues E48, E52, D89, R101 and H105) and box 2 shows the absence of the hydrophobic site (Residues I51, V57, F59, L85, L88, P100 and F102) between adjacent subunits when the α1 helix lies parallel to the C-terminal domain. Box 3 and 4 shows the respective presence of the hydrophobic interface and absence of the electrostatic interface between adjacent subunits 2 and 4 when the α1 helix is surface exposed. Subunits are highlighted blue, red, green and yellow respectively for subunits 1-4. (B) Rotation of L85 and L88 to maintain the hydrophobic core. Box1 highlights the positions of L85 and L88 in the unbound form for subunit 2, forming tethering interactions with P100 and F102 of subunit 4 (yellow). Box 2 highlights the rotation of L85 and L88 of subunit 2 away from P100 and F102 to interact with residues maintaining the hydrophobic core: I22, I43, V47 and V83. Due to the interaction of HicA at each subunit, these changes are mimicked in subunit 3 where the α1 helix is orientated in an exposed manner. (C) Comparison of the unbound (grey) and bound (red) orientations of subunit 2. Lateral displacement (2.4 Å) of the α1 helix results in a lateral shift of E47 and an upward shift of F59 of the α1-α2 loop.



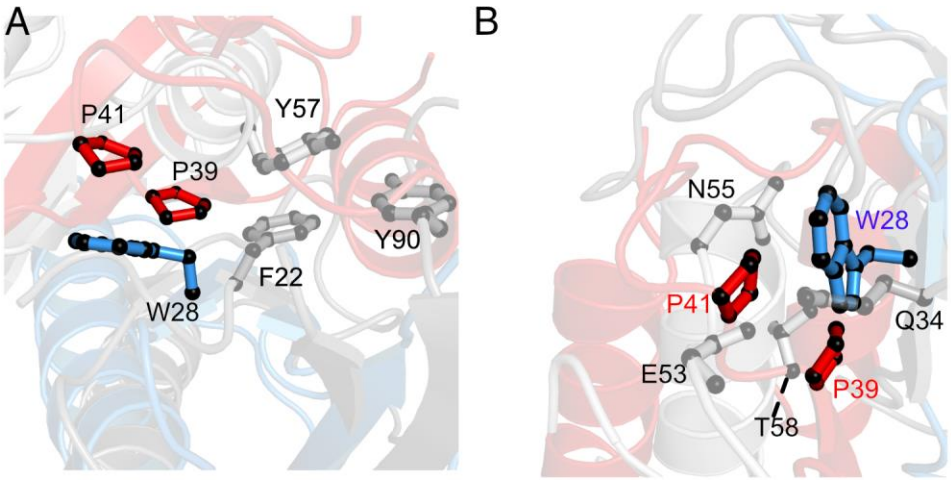
Appendix Figure 13. Intrasubunit interactions between individual N-terminal and C-terminal domains within subunit 1 of HicAB. R124 forms an electrostatic interaction with D84, whose position is stabilised by hydrogen bonds to S86 and Q87. L127 forms a hydrophobic interaction with F59. Due to the symmetrical nature of HicB within the HicAB complex, each set of intersubunit and intrasbunit interactions are conserved throughout HicB.



Appendix Figure 14. Conservation of the dsRBD fold within HicA family members. (A) Superimposition of the solution state of *Burkholderia* HicA (PDB: 4C26, blue) onto the crystalline conformation of HicA (red) with an RMSD of 1.2 Å. (B) Superimposition of the crystalline conformation of HicA (red) onto the crystalline conformation of *Yersinia* HicA3 (PDB: 4P78, purple), RMSD: 2.2 Å. (C) Superimposition of the crystalline conformation of HicA (red) onto the crystalline formation of *Streptoccocus* HicA (PDB: 5YRZ, green), RMSD: 1.5 Å. HicA (red), HicA3 and HicA$_{SP}$ are representative of their conformations when bound to their cognate HicB antitoxin.



Appendix Figure 15. Conservation of residues within dsRBD proteins. (A) Multiple sequence alignment of various dsRBD within proteins encoded in several organisms: *Homo sapiens* (Hs), *Drosophila melanogaster* (Dm), *Saccharomyces cerevisiae* (Sc), *Xenopus leavis* (Xl), *Arabidopsis thaliana* (At), *E. coli* (Ec) and *Aquifex aeolicus* (Aa). The name of each protein, their UniProt accession code and the sequence rnge corresponding to the location of thhe dsRBD in the protein are given. Sequences are highlighted according to their amino acid conservation (>40%) and thhe resultant consensus sequence that describes these dsRBDs is depicted underneath. The three distinct regions of interaction for dsRBDS (with RNA) and the consensus second structure elements for the dsRBDs are highlighted below the consensus sequence. Taken directly from Figure 1[333] with the authors direct permission. (B) Sequence alignment of HicA to HicA family members whose have been experimentally investigated: *Yersinia* HicA, *E. coli* HicA (HicA$_{EC}$) and *Streptococcus* HicA (HicA$_{SP}$). Symbols indicate a conserved residue (*), conservative mutation (:) and a semi-conservative mutation (.).



Appendix Figure 16. Conservation of the hydrophobic patch of P39/P41-W28 between HicAB and $HicAB_{SP}$. (A) The hydrophobic patch between the β3 strand of HicB (blue) and HicA (red) (P39/P41-W28) is conserved within $HicAB_{SP}$. Y57 (α2) within $HicA_{SP}$ (light grey) forms several hydrophobic interactions with F22, F80, F86, Y90 (all α2-β4 loop) of $HicB_{SP}$ (dark grey). For clarity only interactions between Y57 and F22/Y90 are shown. (B) This hydrophobic patch within $HicAB_{SP}$ is supplemented by a hydrogen bond network between the α2 helix of $HicA_{SP}$ (E53, N55 and T58) interacting with the β3 strand of $HicB_{SP}$ (Q34). The side chains of residues are represented as sticks.

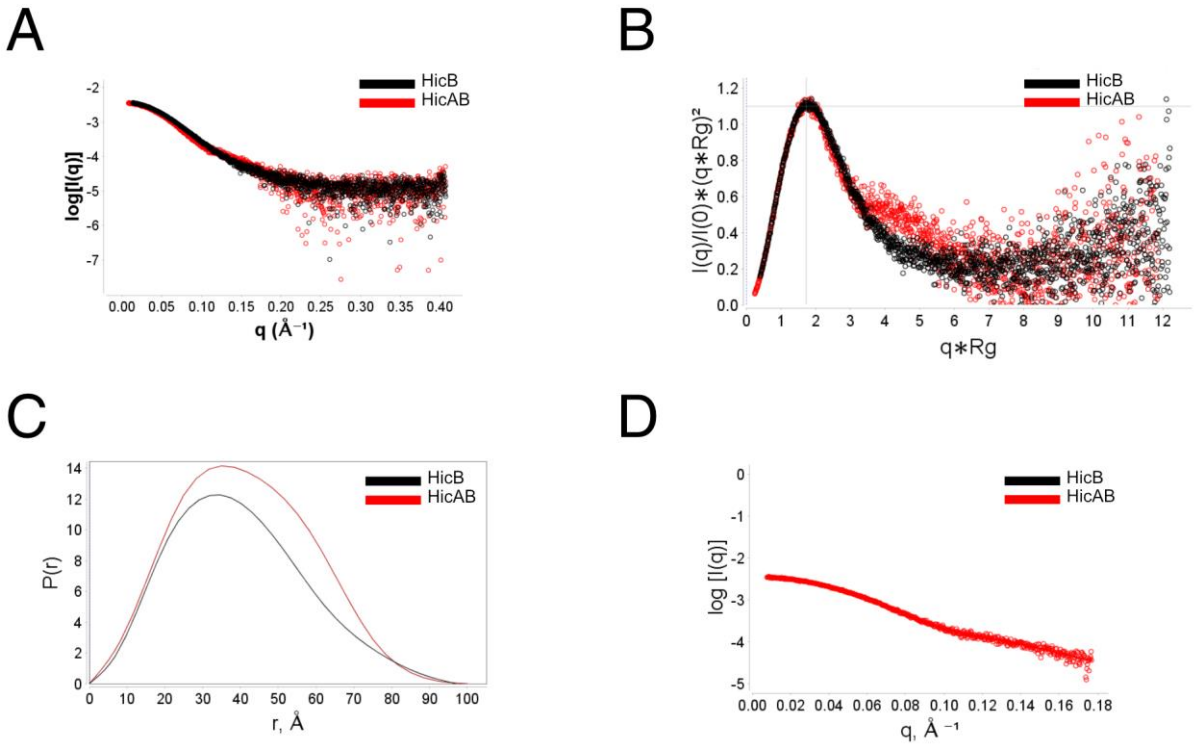
```
Burkholderia_pseudomallei                ------MEFPIAVH-K-DDGSVYGVTVPDI----PGVHSWGETIDDAIKNTREAIVGHVE  48
Burkholderia_ambifaria                   ------MRFFVAVD-K-QEGTDYGVVVPDF----PGCNSWGTTLEQALENTQEAIEGWVE  48
Pseudomonas_Syringae                     ------MKYPICIE-WGDETTAFGIQIPDI----PGAITAGDTFEEAHAAAVEIAHIML   49
Burkholderia_thailandensis_E264          ------MEFPIAVH-K-DDGSVYGVTVPDI----PGVHSWGETIDDAIKNTREAIVGHVE  48
Microcystis_aeruginosa_NIELS-843         -----MITYLATVH-KDN-HSDYGVQFYDF----PGCISAGETIEEAKKMATEALKGHIS  49
Polaromonas_napthalenivorans_CJ2         ------MRYPITIE-PGNDTTAWGVVVPDL----PGCFSAGDTLEEAMIQAEDAVTAWIE  49
Neisseria_gonorrhoaea_FA1090             ------MFIPAALH-K-DEHSAYGVTIPDL----PGCFSCGDTVEEAVANARSAAYMHID  48
Gluconobacter_oxydans_621H               ------MRYPVVIE-RGSETTAFGVVFPDL----PGCFSAGDTLDEALSCAEEAAAAWID  49
Rhodopseudomonas_palustris_HaA2          --MSATSGYLALVY-KDR-DTSYGVAFPDV----PGCISAGDTFEQAIDNAAEALGGHLA  52
Pseudomonas_fluorescens_SBW25            ------MLYPIAIS-TGDNDHAWGVEVPDI----PGCYSAGDDLDDAMAMAREAIEGHFE  49
Roseiflexus_castenholzii_DSM_13941       ------MHYPVVIH-KDT-ESDYGVTVPDL----PGCFSAGETLDEALQEAIEAIECHLE  48
Carboxydothermus_hydrogenoformas_Z-2901  -MYKDLYVFPAIFD-YAE--DGISVEFPDL----PGCLTCGDNTEEALKNAKEALELHLY  52
Mesorhizobium_loti_MAFF303099            -----MKQYIGLIH-KDA-GSDFGVSFPDF----PGVITAGKDLDDARSMAEEALTLHIE  49
Bordetella_parapertussis_12822           ------MLYPIHVS-K-DEGSAYGAAFPDF----PGCFAAADELQDLPRAAQEAVEAHFY  48
Clostridium_perfringens_13               MKKKDTYIFPAIVT-TES--DGITITFPDL----EGCITCAYSDDEIMKVSKEALGLHLY  53
Yersinia_pestis_biovar_Medievalis_91001  -----MAIYPAYVH-VDNDGSASG-YFPDV----KGCIFAINAGEDLFAEASSALDAHFE  49
Haemophilus_somnus_129PT                 ------MLYPIAIE-PGDETHAFGVIVPDI----PGCHSAGDTLEEAYANAKEAIKGHLE  49
Nitrobacter_hamburgensis_X14             -----MRQFIALIH-KDA-DSDYGVSFPDL----PGCVTAGTTLDEARDMAAEALALHLE  49
Serratia_proteamaculans_568              ------MFYPAYVH-SEHDGSASG-FFPDV----PGCYFAGDTLDRAFEDAKSALDAHFE  48
Yersinia_pestis_C092                     -----MAIYPAYVH-VDNDGSASG-YFPDV----KGCIFAINAGEDLFAEASSALDAHFE  49
Yersinia_pestis_Antiqua_JGI              -----MAIYPAYVH-VDNDGSASG-YFPDV----KGCIFAINAGEDLFAEASSALDAHFE  49
Yersinia_pestis_KIM                      -----MAIYPAYVH-VDNDGSASG-YFPDV----KGCIFAINAGEDLFAEASSALDAHFE  49
Yersinia_pestis_Nepa1516                 -----MAIYPAYVH-VDNDGSASG-YFPDV----KGCIFAINAGEDLFAEASSALDAHFE  49
Caulobacter_crescentus_CB15              -----MADYIALIH-KDA-DSAFGVSFPDL----PGCISVGETLEEARANAEEALALHVE  49
Finegolida_magna_ATCC_29328              --MKNDILYPVVMS-KED--GRYFAQCYDY----EEIYAEGDTIEEALKNIKEVFGLIFY  51
Haemophilus_somnus_2336                  ------MLYPIAIE-PGDETHAFGVIVPDI----PGCHSAGDTLEEAYANAKEAIKGHLE  49
Legionella_pneumophila_Lens              ------MRYPVVIH-KDE-HSDFGVIVPDI----PGCYSSGSTYDEALTNVIEAIECHL   48
Escherichia_coli_UMN026                  ------MFYPAYIH-SDPDGSASG-FFPDV----PGCYFAGDTLDNAFQDARSALVAHFE  48
Clostridium_botulinum_B_Eklund_17B       MNAKDMYVFPAIFT-YDN--DGISIEFPDL----PGCLSCADTTDEAIKMAKEALALHLY  53
Treponema_denticola_ATCC_35405           MKKQERFFYPAIFT-YNG-KKEIAVVFPDL-----GCATSGKDEADALLSARELLGCVLC  53
Rhodospirillum_rubum_ATCC_11170          ----MNYTYAVVIE-TDD-LGGFSAFFPDL----PGCVGAGDSVEACFQDALDALALHLA  50
Bartonella_tribocorum_cIP105476          -----MKRFFALVH-KDE-DSAFGVQFPDF----EGLFSASDEEENLIINATEALQLY--  47
Nostoc_sp._PCC7120                       ------MRYAVVIE-KGE--TSYGAYVPDL----PGCVAVGETLEEVKQLITEAIEFHIE  47
Maricaulis_maris_MCS10                   -----MPNFYAIVT-QEG-DSAHGIHFPDL----PGCFSAADTLDEVPAKAREALSLW--  47
Pelotomaculum_thermopropionicum          ---MDKYIFPAIFE-VGE-IKGYCVTFPDL----PGCITEGNTLEEALQMAREALELHLY  51
Leptotrichia_goodfellowii_F0264          -MTKDSYLYPAIFK-YGE--DGITITFPDL----PGCISCGKNDEEALYMARDVLGGWMY  52
Clostridium_kluyveri_DSM_555             ---MDKYMYPALFE-AYE-DGGYTVSFPDL----PGCITEGDSLTEALTMAKEALELFLW  51
Candidatus_methanoregula_boonei_6A8      -----MVSYTVIIE-TGK--HNCSAYCPDL----PGVIATGRTQEETFENMKSAITFHLE  48
Methanospirillum_hungatei_JF-1           -----MYRFLIVIE-KAE--TNYAAYSPDL----PGCVATGKTREEAEEQMHEAIAFHIE  48
Alkalilimnicola_ehrlichei_MLHE-1         ------MKYAIVIE-KAG--QNYSAYVPDL----PGCVATGDSIPEAEALIREAIEWHLA  47
Syntrophomonas_wolfei_subsp.             ---MHKLTYFAVFE-PVA--TGYSVYFPDI----PGCVSYGEDFEEAQKQAADALGLHLY  50
Geobacter_uraniireducens_Rf4             -----MYRFLVVIE-KTD--TGFSAYSPDL----PGCIATGLTREETEKNMHEAIEFHIH  48
Bartonella_henselae_Houston-1            -----MKRFFALVH-KDE-DSAFGVQFPDF----EGLFSAADEEENLIINATEALQLY--  47
Shewanella_baltica_OS195                 ------MLFMVGIETPADETEAFGIIVPVFEKLGYGCFSATDSQEEILFKAKEAILLMAE  54
Haemophlus_influenzae_PittEE             ------MIFTVGVETPENENQAYGMIVPALCQLDYGCFSGADDVDDLLPMVTEAITMMLE  54
Salinispora_arenicola_CNS-205            -----MNAYVVIIE-RAE-DGGYGAWSPDL----PGCVALGETRREALAEMRAAIAFHLE  49
Haemophilus_influenza_86_028NP           ------MLFTIGIETPTNENEAYGITVPALFTEEYSCFSAADTLEEIPTQVTDAIHSILE  54
Rhodobacter_sphaeroides_ATCC_17029       -----MHYFTAIVH-KDP-NSAYGLTFPDL----PGCFAAADDWSEIPAKAAEALDLW--  47
Vibrio vulnificus YJ016                  ------MWFTLGVETPKDKDSSYGIVVPALCNDRYSCYSASDSSKEIAAHSAEAIESILE  54
8Escherichia coli 0157:H7 str. EC41      ------MFFSVGVETPKDDHTAYGITVPAFDRFDFGCVSAADTQSEIPVMAREAILAIVE  54
Thermus thermophiles HB27                ----MRRRYRVVVE-RDE-EGYFVAHVPEL-----HAHTQAQSFEELLRRLQEAIAVSLE  49
Yersinia_pseudotuberculosis_IP_31758     ------MIYPIFIF-KTVEG--FDGYFPDI----DGCFFAGNTFADISKNAEEAFAVHIE  47
Yersinia_pestis_angola                   ------MIYPIFIF-KTVEG--FDGYFPDI----DGCFFAGNTFADISKNAEEAFAVHIE  47
                                                                .

Burkholderia_pseudomallei                TLIELG-----EDVEFTCSTVEE-LVAK--PEYAGA--VWA-------LVSVD----LSQ  87
Burkholderia_ambifaria                   TSVEAG-----DEVRFEPTSFEE-LRDN--PDYAGS--FWA-------VVDID----PSK  87
Pseudomonas_Syringae                     EIAASG-----GSIPKVGTVAEH-AKN---PDFSGM--GWG-------MLEID----VTP  87
Burkholderia_thailandensis_E264          TLIELG-----DDVEFTCSTVEE-LVAK--PEYAGA--VWA-------LVSVD----LSQ  87
Microcystis_aeruginosa_NIELS-843         FMLADG-----DEIPTPSTLETI-L-TD--ADHQDA--I---------AFLLI----EVS  85
Polaromonas_napthalenivorans_CJ2         VALDAG-----QDIPAPSHIEAL-RTAH--PEFDGW--LWA-------LVKVD----PAM  88
Neisseria_gonorrhoaea_FA1090             GMIEDG-----GFKNLAVSSIAD-LSQE--PDYHGA--TWV-------MIEID----PAK  87
Gluconobacter oxydans 621H               ATLDAG-----EAVPSPSSLDAI-Q-SN--PDYQGW--TVG-------IVTID----PAI  87
Rhodopseudomonas palustris HaA2          LLRADG-----DPIPQPRSLEQL-R-DD--PEFVAE--AADA-----VVAFVR----PQA  92
Pseudomonas_fluorescens_SBW25            ILAEDG-----APIPSAQKVTLH-AAN---PQYAGC--TWA-------VVDID----VTK  87
Roseiflexus_castenholzii_DSM_13941       GLLADG-----EPIPTPKPVELH-Q-SN--PDYAGG--IWA-------FVTVD----VTK  86
Carboxydothermus_hydrogenoformas_Z-2901  GMEKDN-----EPIPEPTPIDKIKI--E--P---NQ--VLV-------LVEAWMPLVRSE  91
Mesorhizobium_loti_MAFF303099            GLVEDG-----EAIPEPSSLQAV-M-AD--ADNRDG--V---------AILVA----AKT  85
Bordetella_parapertussis_12822           G---ET-----ERIPAPSAPEAW-AND---EDYQGG--YWM-------MVDID----LSK  83
Clostridium_perfringens_13               GLETDEEIEGKELIPVPSKLNDLKL--E--K---NQ--ATT-------LVEVYMPVIRQV  97
Yersinia_pestis_biovar_Medievalis_91001  ALVSEG-----IEIPEAHDMPYH-VYRNPCDYADGG--QWY-------NVNID----MSK  90
Haemophilus_somnus_129PT                 LLVEMG-----EEVPLPTSINNH-RNN---PDFTEYDMFFG-------FVDVD----ITH  89
Nitrobacter_hamburgensis_X14             GMAEDG-----EGVPEPASLEEV-M-AD--PVNRDG--V---------AVLVA----APA  85
Serratia_proteamaculans_568              LLSEDN-----QSIPRPQAVSFH-LAQES-DTLNGG--QWL-------LVDIN----MDK  88
Yersinia pestis C092                     ALVSEG-----IEIPEAHDMPYH-VYRNPCDYADGG--QWY-------NVNID----MSK  90
Yersinia pestis Antiqua JGI              ALVSEG-----IEIPEAHDMPYH-VYRNPCDYADGG--QWY-------NVNID----MSK  90
Yersinia pestis KIM                      ALVSEG-----IEIPEAHDMPYH-VYRNPCDYADGG--QWY-------NVNID----MSK  90
Yersinia_pestis_Nepa1516                 ALVSEG-----IEIPEAHDMPYH-VYRNPCDYADGG--QWY-------NVNID----MSK  90
Caulobacter_crescentus_CB15              GMIEDG-----ETLPAPSALDEV-M-AA--PGNRDG--V---------VILVP----LKT  85
Finegolida_magna_ATCC_29328              DFEEDE-------YKKASNIEDIMKELK--E---GE--FVV-------YVNVWLPYEFSK  90
Haemophilus_somnus_2336                  LLVEMG-----EEVPLPTSINNH-RQN---PDFTEYDMFFG-------FVDVD----ITH  89
Legionella pneumophila Lens              GLLIDN-----EPLPVGTTIDRW-I-ND--EEFQGG--VWA-------FVDID----LSQ  86
Escherichia_coli_UMN026                  TLCEMN-----EELPLPGSVETH-LVQRA-QDFIGG--QWL-------LVDIN----MNQ  88
Clostridium_botulinum_B_Eklund_17B       GMEEDN-----ESIPKDTPINNLTL--L--E---NQ--IPM-------LIEVYMPLYRTA  92
Treponema_denticola_ATCC_35405           GLEDDG-----EPIPKASHLSNIKT--G--K---NE--KSV-------LIDVYLPTYRM   92
Rhodospirillum_rubum_ATCC_11170          GMIEDG-----DSLPAPTPLSDV-E-VD--PDIQVA--AIL-------L---A----SVP  85
Bartonella_tribocorum_cIP105476          --CEDM-----DTVPVPLKFE------E--VIQQKA--VKKALSEGAFLIQVP----FIE  86
Nostoc_sp._PCC7120                       GMIVDG-----LPIPQPTSMTHE-VEVL--M--------------------------S-  71
```

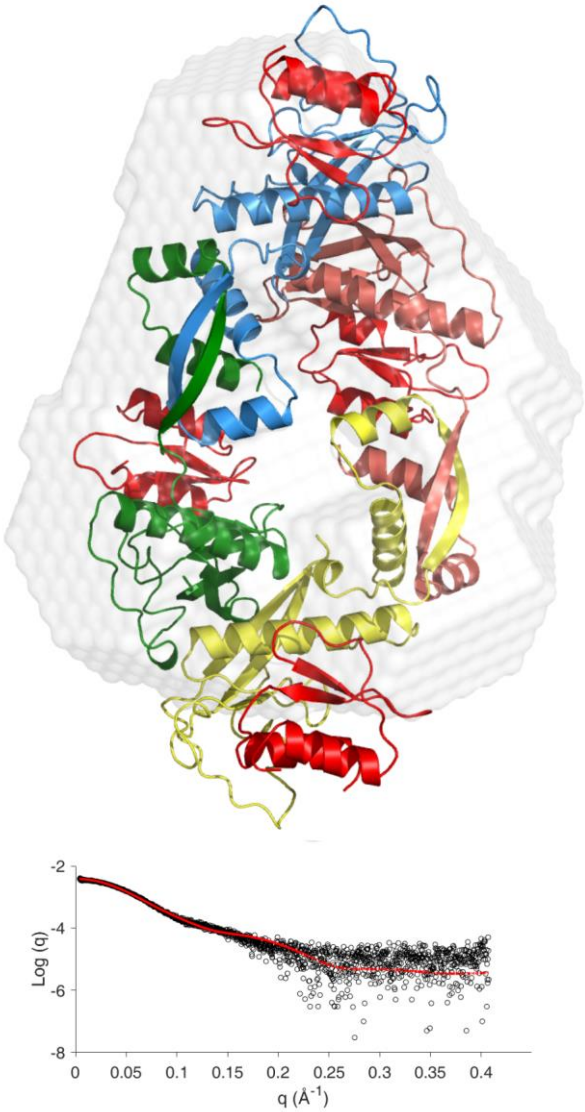
```
Maricaulis_maris_MCS10                  --FEDE-----AMIA-PSSLADV-H-EK--IRNDLG--YIDLMSSGSFLMSIP----VVS  89
Pelotomaculum_thermopropionicum         GMEEDE-----DEIPSPTPPERVDV--P--E---KG--FVA-------IVEARMPLIRDR  90
Leptotrichia_goodfellowii_F0264         QIERAK-----EKIPKSSSLNMINL--N--F---DE--KVL-------LIDVWMPSVRKS  91
Clostridium_kluyveri_DSM_555            NMEDDN-----EEIPEPTPPEKIET--K--K---GN--FIV-------PVEADMALIRAK  90
Candidatus_methanoregula_boonei_6A8     GLLEDN-----QEIPTPVSKAKT-IRIS--R---------------------------KH  73
Methanospirillum_hungatei_JF-1          GMKEDG-----LPIPQACSSASW-VLV-------------------------------  69
Alkalilimnicola_ehrlichei_MLHE-1        GLREDG-----APIPQPSSHVEY-VDVA-------------------------------  69
Syntrophomonas_wolfei_subsp.            GMEKDG-----DEIPTPSKVPQVDTETA--P---GY--IIS-------PVSIFPSIVRNE  91
Geobacter_uraniireducens_Rf4            GLLEDH-----LSVPESSSFAEF-MAVA-------------------------------  70
Bartonella_henselae_Houston-1           --CEDM-----DTVPVPLKFE------E--VIQQKA--VKKALSEGAFLIQVP----FIE  86
Shewanella_baltica_OS195                EVINDG-----HL-IDALNEGYRDYRAT--H--PEFE-QWV-------AVEVP----LEA  92
Haemophlus_influenzae_PittEE            AMVED-----GFDLTTLKDKGVTHYKAD--PEYADFD-TWL-------LVDVD----ISE  95
Salinispora_arenicola_CNS-205           GLREAG-----EPVPAPKAVGVD-L-IA--A---------------------------  71
Haemophilus_influenza_86_028NP          MMFED-----GIDINELQDKGYRHYQTQ--EDFNYCD-TWL-------LLDVD----ISA  95
Rhodobacter_sphaeroides_ATCC_17029      --FEDE-----PAVE-PASIDAI-Q-SR--E----D--VRKALAEGASLMTFA----YVP  85
Vibrio_vulnificus_YJ016                 EMLAD-----GFDVLGLEDQGALAYRED--EAYRFCD-VWL-------QIEID----VSR  95
Escherichia_coli_0157:H7_str._EC41      EMVLSG----SYSVDDIHDDGCLTYAAN--QDYSHCD-SWF-------VIDVD----LSE  96
Thermus_thermophiles_HB27               EERAEV-----VGLEGALEIEAA-----------------------------------  67
Yersinia_pseudotuberculosis_IP_31758    ALMNEG-----FPLPSPPKDPHR-YIDDPRLKEEGG--ILG-------FVEID----PAK  88
Yersinia_pestis_angola                  ALMNEG-----FPLPSPPKDPHR-YIDDPRLKEEGG--ILG-------FVEID----PAK  88


Burkholderia_pseudomallei               LDSKPERINVSIPRFVLHKIDAYVAS-RH--ETRSGFLARAALEALNEGKVRHA  138
Burkholderia_ambifaria                  FDMKPERVNISLPRFVLTQIDAFAEA-HK--ETRSGFLVRAALEAIRHEQTAV-  137
Pseudomonas_Syringae                    YLGKTEKVNVTLPGFVIRQIDRYVRD-HS-IKSRSTFLADAALEKLGRA-----  134
Burkholderia_thailandensis_E264         LDSKPERINVSIPRFVLHKIDAYVAS-RH--ETRSGFLARAALEALNEGKVRHA  138
Microcystis_aeruginosa_NIELS-843        EAILNHQE-LCVKN---------------------------------------  98
Polaromonas_napthalenivorans_CJ2        LDDTVERVNISLPRRVLRRLDARART-AG--ETRSGFIARMAVEGREPAAHA--  137
Neisseria_gonorrhoaea_FA1090            ISRQQIRFNVSWPQYLLDRVDEYTSA-NH--ETRSGFLAKAALLTMNQA-----  133
Gluconobacter_oxydans_621H              LDDRVERANISLPRRVLERLDALARA-SR--ESRSGMIAAMTLQAKLTPDTKSR  138
Rhodopseudomonas_palustris_HaA2         ERAAAE-----------------------------------------------  98
Pseudomonas_fluorescens_SBW25           YLGKAQKLNITLPGYLLNRIDEYVLH-HPEEKSRSGFLASAALKVLQQEH----  136
Roseiflexus_castenholzii_DSM_13941      LSGKTRRINITLPERVLHLVNKYASE-RG--ESRSGLITQAVLEYIGSRETLAR  137
Carboxydothermus_hydrogenoformas_Z-2901 MDNKAVKKTLTLPKWLNDL----AEK-KK--INFSRVLQQALKEQLGIKEREI-  137
Mesorhizobium_loti_MAFF303099           EARKAVRINVTLPDDVLRRIDAFAEA-HG--YTRSGFLAKAAEKAMELESA---  133
Bordetella_parapertussis_12822          VNTKAVRLNISLPENLVHRIDEEAKA-R--RLSRSAFLAMAAEHEMADA-----  129
Clostridium_perfringens_13              LDNKAVKKTVTIPNWLDVQ----AKK-YE--INFSQLLQESIRNLLQL------  138
Yersinia_pestis_biovar_Medievalis_91001 YDGKVERINVTLPHRLIHQIDTIVKV-RPEYASRSNFLAEAARKELQKLA----  139
Haemophilus_somnus_129PT                LLGKAEKINITMPAYLIKRIDEFVAT-HKEYKSRSGFLAQIAADKILSD-----  137
Nitrobacter_hamburgensis_X14            PAVKSVRINVTLPADVLEQIDRHAAS-EG--FTRSGFLAHAAKKALAA------  130
Serratia_proteamaculans_568             FDGRAERINITLPHRLLNRIDSMVQQ-HPGYGSRSAFLAAAARNELLNAG----  137
Yersinia_pestis_C092                    YDGKVERINVTLPHRLIHQIDTIVKV-RPEYASRSNFLAEAARKELQKLA----  139
Yersinia_pestis_Antiqua_JGI             YDGKVERINVTLPHRLIHQIDTIVKV-RPEYASRSNFLAEAARKELQKLA----  139
Yersinia_pestis_KIM                     YDGKVERINVTLPHRLIHQIDTIVKV-RPEYASRSNFLAEAARKELQKLA----  139
Yersinia_pestis_Nepa1516                YDGKVERINVTLPHRLIHQIDTIVKV-RPEYASRSNFLAEAARKELQKLA----  139
Caulobacter_crescentus_CB15             VQPRTVRVNITLPETTLREIDAFAET-YG--YTRSGLLARAAKQMIDAA-----  131
Finegolida_magna_ATCC_29328             VKVVNTKKTLTIPNHLNML----GQY-KN--INFSKILTEALEKELNVK-----  132
Haemophilus_somnus_2336                 LLGKAEKINITMPAYLIKRIDEFVAT-HKEYKSRSGFLAQIAADKILSA-----  137
Legionella_pneumophila_Lens             ISGKAKRINITMPERILSLIDLYAKN-HA-IKNRSSFLADAALSYMESHK----  134
Escherichia_coli_UMN026                 FEGRAERINITMPKRLLNKIDTYVRN-NPDYANRSAFLAEAARRVLPGV-----  136
Clostridium_botulinum_B_Eklund_17B      IENQSIKKTLTIPQWLNKL----AEK-NE--INFSQILQAALKEQLGIHSTLNK  139
Treponema_denticola_ATCC_35405          QKTRSVNRTVTLPAWLNAL----ASE-HN--INCSQLLQESLKKQFQTVLQTR-  138
Rhodospirillum_rubum_ATCC_11170         VTGRTMRVNITMDATLVAAIDGVT-------SNRSAWLADAALRALRSNR----  128
Bartonella_tribocorum_cIP105476         NDSEVVRTNISIERGLLRAIDNCAQE-RG--LTRSAFLATAARHELNI------  131
Nostoc_sp._PCC7120                      -----------------------------------------------------  71
Maricaulis_maris_MCS10                  MTGRTVKANLTMDAGLLQAIDETARE-RG--ITRSAFLADLARREIAA------  134
Pelotomaculum_thermopropionicum         MAKKAVKKTVTLPKWLNDL----AEK-EK--VNFSQLLQTAIKEKLHIRENH--  135
Leptotrichia_goodfellowii_F0264         IRNKAVKKTLTIPQWLNER----AIE-KN--LNFSHILQEALKEELGIK-----  133
Clostridium_kluyveri_DSM_555            MNNKTVNTTVTMPQWLKYQ----AEK-KK--INFSQVLQAALKEQLNIKN----  133
Candidatus_methanoregula_boonei_6A8     LCETVPA----------------------------------------------  80
Methanospirillum_hungatei_JF-1          -----------------------------------------------------  69
Alkalilimnicola_ehrlichei_MLHE-1        -----------------------------------------------------  69
Syntrophomonas_wolfei_subsp.            LDNRAVKTNLTIPAWLKEI----AEA-QG--VNYSKIFQSALMDYLGIKTK---  135
Geobacter_uraniireducens_Rf4            -----------------------------------------------------  70
Bartonella_henselae_Houston-1           NDSEVVRTNISIERGLLRAIDNCAQE-RG--LTRSAFLATAARHELNI------  131
Shewanella_baltica_OS195                LKAKQKRLNITLPESQLARIDSYVEF-HSEFKDRSDFLSKAADKLMYNADLSR-  144
Haemophlus_influenzae_PittEE            YLGKKQRINVSLPEYLLTRIDRRVAAMGNYYKDRSHFLANAAHRELHAHSDKEM  149
Salinispora_arenicola_CNS-205           -----------------------------------------------------  71
Haemophilus_influenza_86_028NP          YQGKRHRINISLPEYLIKRIDSRVAS-NPIYKDRSHFLAIASQKELRQ------  142
Rhodobacter_sphaeroides_ATCC_17029      ADGKLERVNVSMDRGLLRAIDATASA-RG--MTRSAFLAAVARREIAGF-----  131
Vibrio_vulnificus_YJ016                 YSGKSQRVNISLPESLLAVIDRKVEN-SDLYRDRSHFLAIAARNELC-------  141
Escherichia_coli_0157:H7_str._EC41      IEGKQQRINIALPDVLIRRIDGFVRESGGVYRDRSHFLAQAARHELSYK-----  145
Thermus_thermophiles_HB27               -----------------------------------------------------  67
Yersinia_pseudotuberculosis_IP_31758    YESKAVKFNLTMSQNLLTAIDKFIAT-NRGYKNRSQFLAELAREKIIS------  135
Yersinia_pestis_angola                  YESKAVKFNLTMSQNLLTAIDKFIAT-NRGYKNRSQFLAELAREKIIS------  136
```

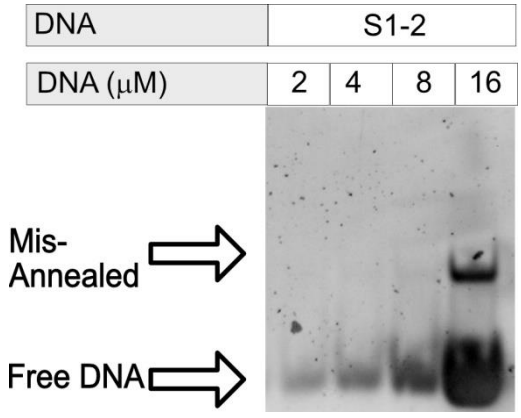
Appendix Figure 17. **Alignment of HicB to 51 homologous sequences across different bacterial strains.** Sequences were aligned using Clustal W[291] and conservation at specific amino acid positions are shown: conserved (*), conservative mutation (:) and a semi-conservative (.).



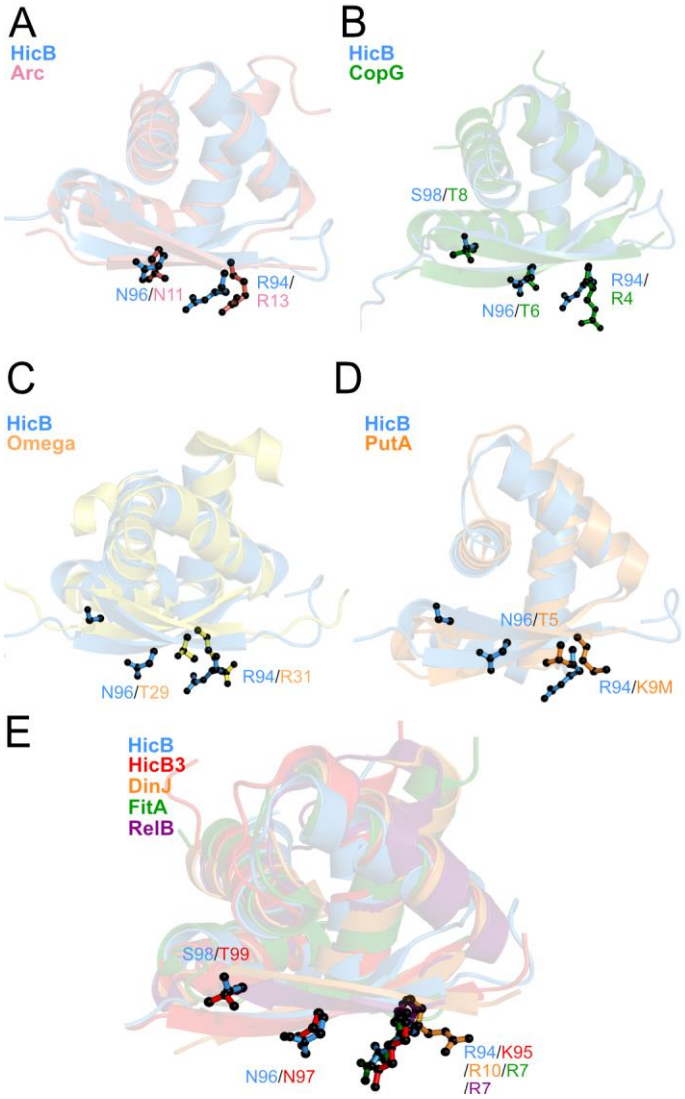
Appendix Figure 18. SAXS of HicAB (A) Log_{10} SAXS intensity versus scattering vector (q). (B) Dimensionless Kratky plot. Cross-hair marks the Guinier-Kratky plot (1.732, 1.1), the main peak for folded globular particles. (C) Pair-distance, P(r) distribution function. Maximum dimension, d_{max} is the largest non-negative value that supports a smooth distribution function. (D) Fit of the parameters computed by the P(r) distribution (red line) to the raw scattering data (red). Within each figure (apart from (D)) HicB and HicAB are highlighted in black and red.



Appendix Figure 19. *Ab initio* modelling of the shape envelope of the HicAB SAXS profile with the crystal structure of *S.pneumoniae* HicAB complex. *Ab initio* modelling of the hetero-octameric state of $HicAB_{SP}$ into the shape envelope of HicAB (white). (Underneath) The FoXS profile of the proposed scattering profile for a $HicAB_{SP}$ hetero-octamer (red) against the experimental raw scattering data (black) (χ^2=3.34).



Appendix Figure 20. Gel-shift analysis investigating if mis-annealed higher order states of S1-2 DNA form at higher concentrations of DNA.



Appendix Figure 21. Superimposition of structures with a high structural similarity to the RHH domain of HicB with functional residues of high importance highlighted. (A) Superimposition of HicB (Blue) with Arc (PDB: 1BAZ, red). (B) CopG (PDB:2CPG, green) superimposed onto HicB. (C) Omega (PDB: 2BNZ, yellow) superimposed to HicB. (D) PutA (PDB:2AY0, Orange) superimposed with HicB. (E) Superimposition of HicB with RHH containing antitoxins: DinJ (PDB: 4Q2U, orange), HicB3 (PDB: 4P7D, red), FitA (PDB: 2H10, green) and RelB (PDB: 4FXE, purple).

```
HicB            R--IN-------VSIPRFVLHKIDAYVASRHE--T-RSGF--LARAALEA
HicB3           AVKFN-------LTMSQNLLTAIDKFIATNRGYKN-RSQF--LAELARE-
HicB_EC         SE-ITQQELARRIGKPKQ---EITRLFNLHHA--T-KIDAVQLAAKALG-
HicB_SP         IK-KT-------LTIPKWA-DKLGREMGLNFS--QTLTDAI-----ADK-
                    .       :  .:     :   .  .                *

HicB            LNEGKVRH---A
HicB3           ----KIIS----
HicB_EC         ----KELSLVMV
HicB_SP         ----KVQ----A
                    *
```

Appendix Figure 22. Conservation between the C-terminal domains of *Burkholderia, Yersinia, Escherichia* and *Streptococcus* HicB. Symbols indicate a conserved residue (*), conservative mutation (:) and a semi-conservative mutation (.).

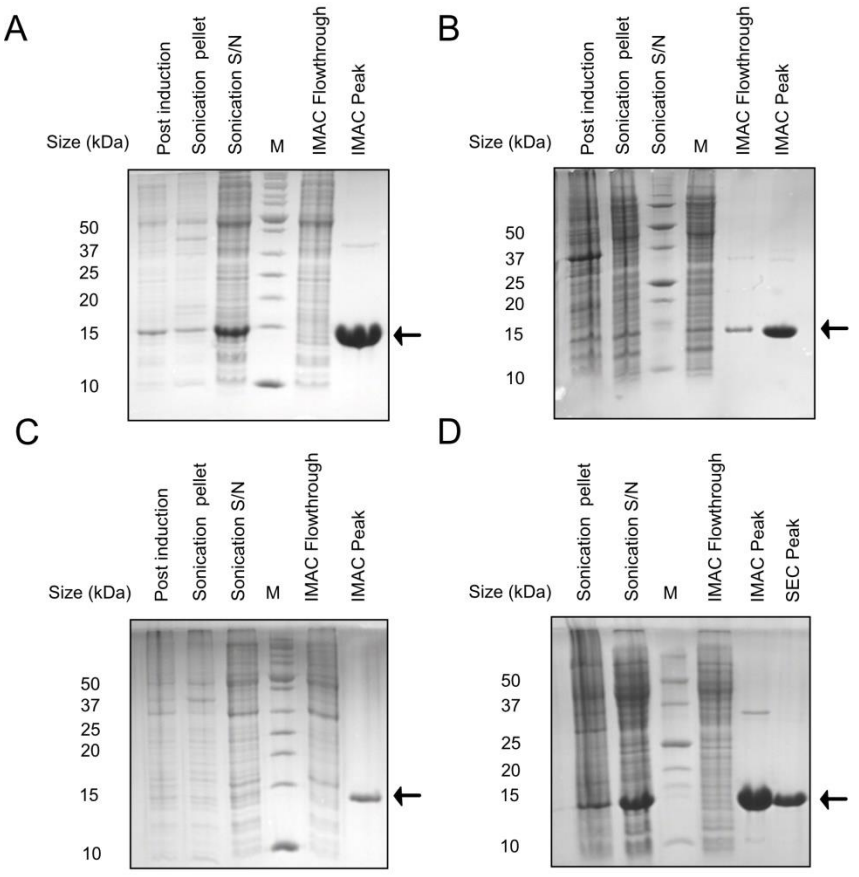
```
Burkholderia_pseudomallei                ------MEFPIAV-HKDD-GSVYGVTVPDIP----GVHSWGETIDDAIKNTREAIVGHVE  48
Burkholderia_ambifaria                   ------MRFFVAV-DKQE-GTDYGVVVPDFP----GCNSWGTTLEQALENTQEAIEGWVE  48
Pseudomonas Syringae                     ------MKYPICI-EWGDETTAFGIQIPDIP----GAITAGDTFEEAHAAAVEIAHIMLQ  49
Burkholderia thailandensis E264          ------MEFPIAV-HKDD-GSVYGVTVPDIP----GVHSWGETIDDAIKNTREAIVGHVE  48
Polaromonas napthalenivorans CJ2         ------MRYPITI-EPGNDTTAWGVVVPDLP----GCFSAGDTLEEAMIQAEDAVTAWIE  49
Neisseria_gonorrhoaea_FA1090             ------MFIPAAL-HKDE-HSAYGVTIPDLP----GCFSCGDTVEEAVANARSAAYMHID  48
Gluconobacter_oxydans_621H               ------MRYPVVI-ERGSETTAFGVVFPDLP----GCFSAGDTLDEALSCAEEAAAAWID  49
Pseudomonas fluorescens SBW25            ------MLYPIAI-STGDNDHAWGVEVPDIP----GCYSAGDDLDDAMAMAREAIEGHFE  49
Roseiflexus castenholzii DSM 13941       ------MHYPVVI-HKDT-ESDYGVTVPDLP----GCFSAGETLDEALQEAIEAIECHLE  48
Carboxydothermus hydrogenoformas Z-2901  -MYKDLYVFPAIF-DYAE--DGISVEFPDLP----GCLTCGDNTEEALKNAKEALELHLY  52
Mesorhizobium_loti_MAFF303099            -----MKQYIGLI-HKDA-GSDFGVSFPDFP----GVITAGKDLDDARSMAEEALTLHIE  49
Bordetella_parapertussis_12822           ------MLYPIHV-SK-DEGSAYGAAFPDFP----GCFAAADELQDLPRAAQEAVEAHFY  48
Clostridium_perfringens_13               MKKKDTYIFPAIV-TTES--DGITITFPDLE----GCITCAYSDDEIMKVSKEALGLHLY  53
Yersinia_pestis_biovar_Medievalis_91001  -----MAIYPAYV-HVDNDGSASG-YFPDVK----GCIFAINAGEDLFAEASSALDAHFE  49
Haemophilus_somnus_129PT                 ------MLYPIAI-EPGDETHAFGVIVPDIP----GCHSAGDTLEEAYANAKEAIKGHLE  49
Nitrobacter hamburgensis X14             -----MRQFIALI-HKDA-DSDYGVSFPDLP----GCVTAGTTLDEARDMAAEALALHLE  49
Serratia proteamaculans 568              ------MFYPAYV-HSEHDGSASG-FFPDVP----GCYFAGDTLDRAFEDAKSALDAHFE  48
Yersinia pestis C092                     -----MAIYPAYV-HVDNDGSASG-YFPDVK----GCIFAINAGEDLFAEASSALDAHFE  49
Yersinia_pestis_Antiqua_JGI              -----MAIYPAYV-HVDNDGSASG-YFPDVK----GCIFAINAGEDLFAEASSALDAHFE  49
Yersinia_pestis_KIM                      -----MAIYPAYV-HVDNDGSASG-YFPDVK----GCIFAINAGEDLFAEASSALDAHFE  49
Yersinia_pestis_Nepa1516                 -----MAIYPAYV-HVDNDGSASG-YFPDVK----GCIFAINAGEDLFAEASSALDAHFE  49
Caulobacter crescentus CB15              -----MADYIALI-HKDA-DSAFGVSFPDLP----GCISVGETLEEARANAEEALALHVE  49
Finegolida magna ATCC 29328              --MKNDILYPVVM-SKED--GRYFAQCYDYE----EIYAEGDTIEEALKNIKEVFGLIFY  51
Haemophilus_somnus_2336                  ------MLYPIAI-EPGDETHAFGVIVPDIP----GCHSAGDTLEEAYANAKEAIKGHLE  49
Legionella_pneumophila_Lens              ------MRYPVVI-HKDE-HSDFGVIVPDIP----GCYSSGSTYDEALTNVIEAIECHLE  48
Escherichia_coli_UMN026                  ------MFYPAYI-HSDPDGSASG-FFPDVP----GCYFAGDTLDNAFQDARSALVAHFE  48
Clostridium_botulinum_B_Eklund_17B       MNAKDMYVFPAIF-TYDN--DGISIEFPDLP----GCLSCADTTDEAIKMAKEALALHLY  53
Treponema_denticola_ATCC_35405           MKKQERFFYPAIF-TYNG-KKEIAVVFPDL-----GCATSGKDEADALLSARELLGCVLC  53
Rhodospirillum rubum ATCC 11170          ----MNYTYAVVI-ETDD-LGGFSAFFPDLP----GCVGAGDSVEACFQDALDALALHLA  50
Bartonella tribocorum cIP105476          -----MKRFFALV-HKDE-DSAFGVQFPDFE----GLFSASDEEENLIINATEALQLYCE  49
Maricaulis_maris_MCS10                   -----MPNFYAIV-TQEG-DSAHGIHFPDLP----GCFSAADTLDEVPAKAREALSLWFE  49
Pelotomaculum_thermopropionicum          ---MDKYIFPAIF-EVGE-IKGYCVTFPDLP----GCITEGNTLEEALQMAREALELHLY  51
Leptotrichia_goodfellowii_F0264          -MTKDSYLYPAIF-KYGE--DGITITFPDLP----GCISCGKNDEEALYMARDVLGGWMY  52
Clostridium_kluyveri_DSM_555             ---MDKYMYPALF-EAYE-DGGYTVSFPDLP----GCITEGDSLTEALTMAKEALELFLW  51
Syntrophomonas wolfei subsp.             ---MHKLTYFAVF-EPVA--TGYSVYFPDIP----GCVSYGEDFEEAQKQAADALGLHLY  50
Bartonella henselae Houston-1            -----MKRFFALV-HKDE-DSAFGVQFPDFE----GLFSAADEEENLIINATEALQLYCE  49
Shewanella_baltica_OS195                 ------MLFMVGIETPADETEAFGIIVPVFEKLGYGCFSATDSQEEILFKAKEAILLMAE  54
Haemophlus_influenzae_PittEE             ------MIFTVGVETPENENQAYGMIVPALCQLDYGCFSGADDVDDLLPMVTEAITMMLE  54
Haemophilus_influenza_86_028NP           ------MLFTIGIETPTNENEAYGITVPALFTEEYSCFSAADTLEEIPTQVTDAIHSILE  54
Rhodobacter_sphaeroides_ATCC_17029       -----MHYFTAIV-HKDP-NSAYGLTFPDLP----GCFAAADDWSEIPAKAAEALDLWFE  49
Vibrio_vulnificus_YJ016                  ------MWFTLGVETPKDKDSSYGIVVPALCNDRYSCYSASDSSKEIAAHSAEAIESILE  54
Escherichia_coli_0157:H7_str._EC41       ------MFFSVGVETPKDDHTAYGITVPAFDRFDFGCVSAADTQSEIPVMAREAILAIVE  54
Yersinia pseudotuberculosis IP 31758     ------MIYPIFI-FKTVEG--FDGYFPDID----GCFFAGNTFADISKNAEEAFAVHIE  47
Yersinia_pestis_angola                   ------MIYPIFI-FKTVEG--FDGYFPDID----GCFFAGNTFADISKNAEEAFAVHIE  47
                                                                  .                                .

Burkholderia_pseudomallei                TLIELG-----ED-VEFTCSTVE-ELV-AK---PEYAGA--V-------WALVSVDLSQL  88
Burkholderia_ambifaria                   TSVEAG-----DE-VRFEPTSFE-ELR-DN---PDYAGS--F-------WAVVDIDPSKF  88
Pseudomonas Syringae                     EIAASG-----GS-IP-KVGTVA-EHA-KN---PDFSGM--G-------WGMLEIDVTPY  88
Burkholderia_thailandensis_E264          TLIELG-----DD-VEFTCSTVE-ELV-AK---PEYAGA--V-------WALVSVDLSQL  88
Polaromonas_napthalenivorans_CJ2         VALDAG-----QD-IP-APSHIE-ALRTAH---PEFDGW--L-------WALVKVDPAML  89
Neisseria_gonorrhoaea_FA1090             GMIEDG-----GF-KNLAVSSIA-DLS-QE---PDYHGA--T-------WVMIEIDPAKI  88
Gluconobacter_oxydans_621H               ATLDAG-----EA-VP-SPSSLD-AIQ-SN---PDYQGW--T-------VGIVTIDPAIL  88
Pseudomonas_fluorescens_SBW25            ILAEDG-----AP-IP-SAQKVT-LHA-AN---PQYAGC--T-------WAVVDIDVTKY  88
Roseiflexus_castenholzii_DSM_13941       GLLADG-----EP-IP-TPKPVE-LHQ-SN---PDYAGG--I-------WAFVTVDVTKL  87
Carboxydothermus hydrogenoformas Z-2901  GMEKDN-----EP-IP-EPTPID-KIKI--------EPN--QVLVLVEAWM--PLVRSEM  92
Mesorhizobium_loti_MAFF303099            GLVEDG-----EA-IP-EPSSLQ-AVM-AD---ADNRDG--V-------AILVAAK--TE  86
Bordetella_parapertussis_12822           G---ET-----ER-IP-APSAPE-AWA-ND---EDYQGG--Y-------WMMVDIDLSKV  84
Clostridium_perfringens_13               GLETDEEIEGKEL-IP-VPSKLN-DLKL--------EKN--QATTLVEVYM--PVIRQVL  98
Yersinia_pestis_biovar_Medievalis_91001  ALVSEG-----IE-IP-EAHDMP-YHV-YRNPCDYADGG--Q-------WYNVNIDMSKY  91
Haemophilus somnus 129PT                 LLVEMG-----EE-VP-LPTSIN-NHR-NN---PDFTEYDMF-------FGFVDVDITHL  90
Nitrobacter hamburgensis X14             GMAEDG-----EG-VP-EPASLE-EVM-AD---PVNRDG--V-------AVLVAAP--AP  86
Serratia proteamaculans 568              LLSEDN-----QS-IP-RPQAVS-FHL-AQES-DTLNGG--Q-------WLLVDINMDKF  89
Yersinia_pestis_C092                     ALVSEG-----IE-IP-EAHDMP-YHV-YRNPCDYADGG--Q-------WYNVNIDMSKY  91
Yersinia_pestis_Antiqua_JGI              ALVSEG-----IE-IP-EAHDMP-YHV-YRNPCDYADGG--Q-------WYNVNIDMSKY  91
Yersinia_pestis_KIM                      ALVSEG-----IE-IP-EAHDMP-YHV-YRNPCDYADGG--Q-------WYNVNIDMSKY  91
Yersinia_pestis_Nepa1516                 ALVSEG-----IE-IP-EAHDMP-YHV-YRNPCDYADGG--Q-------WYNVNIDMSKY  91
Caulobacter_crescentus_CB15              GMIEDG-----ET-LP-APSALD-EVM-AA---PGNRDG--V-------VILVPLK--TV  86
Finegolida_magna_ATCC_29328              DFEEDE---------YK-KASNIE-DIMKEL------KEG--EFVVYVNVWL--PYEFSKV  91
Haemophilus_somnus_2336                  LLVEMG-----EE-VP-LPTSIN-NHR-QN---PDFTEYDMF-------FGFVDVDITHL  90
Legionella_pneumophila_Lens              GLLIDN-----EP-LP-VGTTID-RWI-ND---EEFQGG--V-------WAFVDIDLSQI  87
Escherichia_coli_UMN026                  TLCEMN-----EE-LP-LPGSVE-THL-VQRA-QDFIGG--Q-------WLLVDINMNQF  89
Clostridium_botulinum_B_Eklund_17B       GMEEDN-----ES-IP-KDTPIN-NLTL--------LEN--QIPMLIEVYM--PLYRTAI  93
Treponema_denticola_ATCC_35405           GLEDDG-----EP-IP-KASHLS-NIKT--------GKN--EKSVLIDVYL--PTYRMMQ  93
Rhodospirillum rubum ATCC 11170          GMIEDG-----DS-LP-APTPLS-DVE-VD---PDIQVA--A-------ILLASV---PV  86
Bartonella_tribocorum_cIP105476          D----M-----DT-VP-VPLKFE------E---VIQQKA--VKKALSEGAFLIQVPFIEN  87
Maricaulis_maris_MCS10                   D----E-----AM-IA--PSSLA-DVH-EK---IRNDLG--YIDLMSSGSFLMSIPVVSM  90
Pelotomaculum_thermopropionicum          GMEEDE-----DE-IP-SPTPPE-RVDV--------PEK--GFVAIVEARM--PLIRDRM  91
Leptotrichia_goodfellowii_F0264          QIERAK-----EK-IP-KSSSLN-MINL--------NFD--EKVLLIDVWM--PSVRKSI  92
Clostridium_kluyveri_DSM_555             NMEDDN-----EE-IP-EPTPPE-KIET--------KKG--NFIVPVEADM--ALIRAKM  91
Syntrophomonas_wolfei_subsp.             GMEKDG-----DE-IP-TPSKVP-QVDTET------APG--YIISPVSIFP--SIVRNEL  92
Bartonella_henselae_Houston-1            D----M-----DT-VP-VPLKFE------E---VIQQKA--VKKALSEGAFLIQVPFIEN  87
Shewanella_baltica_OS195                 EVINDG------HL-I-DALNEGYRDYRAT---H--PEF--------EQWVAVEVPLEAL  93
Haemophlus_influenzae_PittEE             AMVED------GFDLT-TLKDKGVTHYKAD---PEYADF--------DTWLLVDVDISEY  96
Haemophilus_influenza_86_028NP           MMFED------GIDIN-ELQDKGYRHYQTQ---EDFNYC--------DTWLLLDVDISAY  96
Rhodobacter_sphaeroides_ATCC_17029       D----E-----PA-VE--PASID-AIQ-SR---E----D--VRKALAEGASLMTFAYVPA  86
```

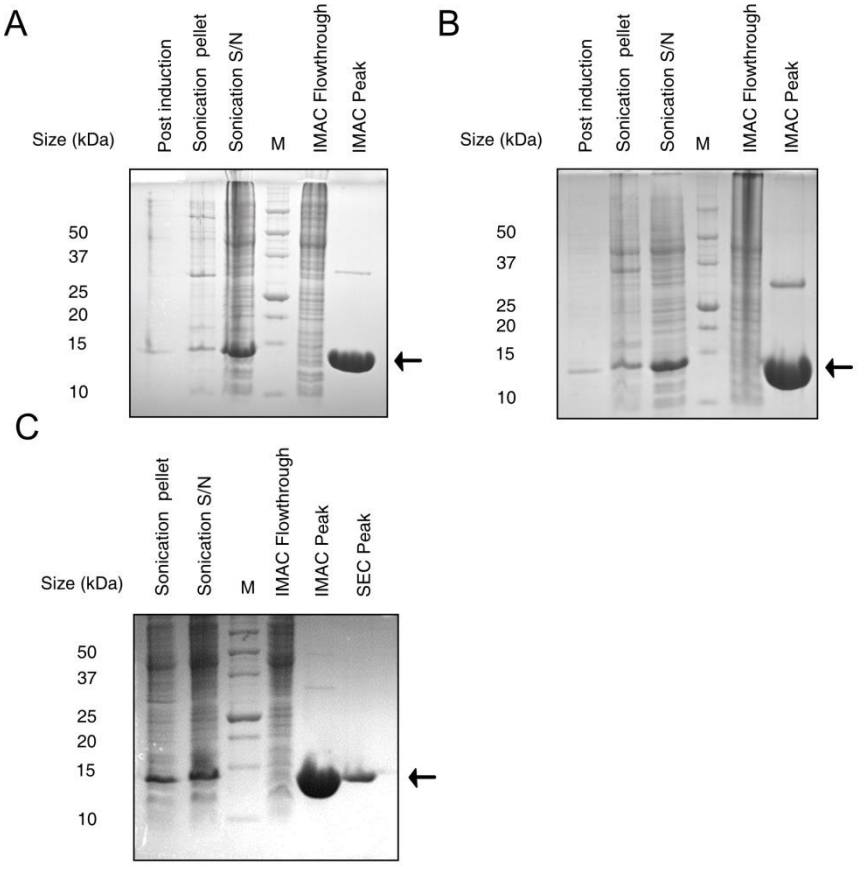
```
Vibrio_vulnificus_YJ016                 EMLAD------GFDVL-GLEDQGALAYRED---EAYRFC--------DVWLQIEIDVSRY  96
Escherichia_coli_0157:H7_str._EC41      EMVLSG-----SYSVD-DIHDDGCLTYAAN---QDYSHC--------DSWFVIDVDLSEI  97
Yersinia_pseudotuberculosis_IP_31758    ALMNEG-----FP-LP-SPPKDP-HRY-IDDPRLKEEGG--I-------LGFVEIDPAKY  89
Yersinia_pestis_angola                  ALMNEG-----FP-LP-SPPKDP-HRY-IDDPRLKEEGG--I-------LGFVEIDPAKY  89


Burkholderia_pseudomallei               DSKPERINVSIPRFVLHKIDAYVAS-RH--ETRSGFLARAALEALNEGKVRHA   138
Burkholderia_ambifaria                  DMKPERVNISLPRFVLTQIDAFAEA-HK--ETRSGFLVRAALEAIRHEQTAV-   137
Pseudomonas_Syringae                    LGKTEKVNVTLPGFVIRQIDRYVRD-HS-IKSRSTFLADAALEKLGRA-----   134
Burkholderia_thailandensis_E264         DSKPERINVSIPRFVLHKIDAYVAS-RH--ETRSGFLARAALEALNEGKVRHA   138
Polaromonas_napthalenivorans_CJ2        DDTVERVNISLPRRVLRRLDARART-AG--ETRSGFIARMAVEGREPAAHA--   137
Neisseria_gonorrhoaea_FA1090            SRQQIRFNVSWPQYLLDRVDEYTSA-NH--ETRSGFLAKAALLTMNQA-----   133
Gluconobacter_oxydans_621H              DDRVERANISLPRRVLERLDALARA-SR--ESRSGMIAAMTLQAKLTPDTKSR   138
Pseudomonas_fluorescens_SBW25           LGKAQKLNITLPGYLLNRIDEYVLH-HPEEKSRSGFLASAALKVLQQEH----   136
Roseiflexus_castenholzii_DSM_13941      SGKTRRINITLPERVLHLVNKYASE-RG--ESRSGLITQAVLEYIGSRETLAR   137
Carboxydothermus_hydrogenoformas_Z-2901 DNKAVKKTLTLPKWLNDL----AEK-KK--INFSRVLQQALKEQLGIKEREI-   137
Mesorhizobium_loti_MAFF303099           ARKAVRINVTLPDDVLRRIDAFAEA-HG--YTRSGFLAKAAEKAMELESA---   133
Bordetella_parapertussis_12822          NTKAVRLNISLPENLVHRIDEEAKA-R--RLSRSAFLAMAAEHEMADA-----   129
Clostridium_perfringens_13              DNKAVKKTVTIPNWLDVQ----AKK-YE--INFSQLLQESIRNLLQL------   138
Yersinia_pestis_biovar_Medievalis_91001 DGKVERINVTLPHRLIHQIDTIVKV-RPEYASRSNFLAEAARKELQKLA----   139
Haemophilus_somnus_129PT                LGKAEKINITMPAYLIKRIDEFVAT-HKEYKSRSGFLAQIAADKILSD-----   137
Nitrobacter_hamburgensis_X14            AVKSVRINVTLPADVLEQIDRHAAS-EG--FTRSGFLAHAAKKALAA------   130
Serratia_proteamaculans_568             DGRAERINITLPHRLLNRIDSMVQQ-HPGYGSRSAFLAAAARNELLNAG----   137
Yersinia_pestis_C092                    DGKVERINVTLPHRLIHQIDTIVKV-RPEYASRSNFLAEAARKELQKLA----   139
Yersinia_pestis_Antiqua_JGI             DGKVERINVTLPHRLIHQIDTIVKV-RPEYASRSNFLAEAARKELQKLA----   139
Yersinia_pestis_KIM                     DGKVERINVTLPHRLIHQIDTIVKV-RPEYASRSNFLAEAARKELQKLA----   139
Yersinia_pestis_Nepal516                DGKVERINVTLPHRLIHQIDTIVKV-RPEYASRSNFLAEAARKELQKLA----   139
Caulobacter_crescentus_CB1              QPRTVRVNITLPETTLREIDAFAET-YG--YTRSGLLARAAKQMIDAA-----   131
Finegolida_magna_ATCC_2932              KVVNTKKTLTIPNHLNML----GQY-KN--INFSKILTEALEKELNVK-----   132
Haemophilus_somnus_2336                 LGKAEKINITMPAYLIKRIDEFVAT-HKEYKSRSGFLAQIAADKILSA-----   137
Legionella_pneumophila_Len              SGKAKRINITMPERILSLIDLYAKN-HA-IKNRSSFLADAALSYMESHK----   134
Escherichia_coli_UMN026                 EGRAERINITMPKRLLNKIDTYVRN-NPDYANRSAFLAEAARRVLPGV-----   136
Clostridium_botulinum_B_Eklund_17B      ENQSIKKTLTIPQWLNKL----AEK-NE--INFSQILQAALKEQLGIHSTLNK   139
Treponema_denticola_ATCC_35405          KTRSVNRTVTLPAWLNAL----ASE-HN--INCSQLLQESLKKQFQTVLQTR-   138
Rhodospirillum_rubum_ATCC_1117          TGRTMRVNITMDATLVAAIDGVT-------SNRSAWLADAALRALRSNR----   128
Bartonella_tribocorum_cIP10547          DSEVVRTNISIERGLLRAIDNCAQE-RG--LTRSAFLATAARHELNI------   131
Maricaulis_maris_MCS10                  TGRTVKANLTMDAGLLQAIDETARE-RG--ITRSAFLADLARREIAA------   134
Pelotomaculum_thermopropionicum         AKKAVKKTVTLPKWLNDL----AEK-EK--VNFSQLLQTAIKEKLHIRENH--   135
Leptotrichia_goodfellowii_F0264         RNKAVKKTLTIPQWLNER----AIE-KN--LNFSHILQEALKEELGIK-----   133
Clostridium_kluyveri_DSM_555            NNKTVNTTVTMPQWLKYQ----AEK-KK--INFSQVLQAALKEQLNIKN----   133
Syntrophomonas_wolfei_subsp             DNRAVKTNLTIPAWLKEI----AEA-QG--VNYSKIFQSALMDYLGIKTK---   135
Bartonella_henselae_Houston-1           DSEVVRTNISIERGLLRAIDNCAQE-RG--LTRSAFLATAARHELNI------   131
Shewanella_baltica_OS195                KAKQKRLNITLPESQLARIDSYVEF-HSEFKDRSDFLSKAADKLMYNADLSR-   144
Haemophlus_influenzae_PittEE            LGKKQRINVSLPEYLLTRIDRRVAAMGNYYKDRSHFLANAAHRELHAHSDKEM   149
Haemophilus_influenza_86_028NP          QGKRHRINISLPEYLIKRIDSRVAS-NPIYKDRSHFLAIASQKELRQ------   142
Rhodobacter_sphaeroides_ATCC_17029      DGKLERVNVSMDRGLLRAIDATASA-RG--MTRSAFLAAVARREIAGF-----   131
Vibrio_vulnificus_YJ016                 SGKSQRVNISLPESLLAVIDRKVEN-SDLYRDRSHFLAIAARNELC-------   141
Escherichia_coli_0157:H7_str._EC41      EGKQQRINIALPDVLIRRIDGFVRESGGVYRDRSHFLAQAARHELSYK-----   145
Yersinia_pseudotuberculosis_IP_31758    ESKAVKFNLTMSQNLLTAIDKFIAT-NRGYKNRSQFLAELAREKIIS------   135
Yersinia_pestis_angola                  ESKAVKFNLTMSQNLLTAIDKFIAT-NRGYKNRSQFLAELAREKIIS------   135
                                             . .::                      *  :
```

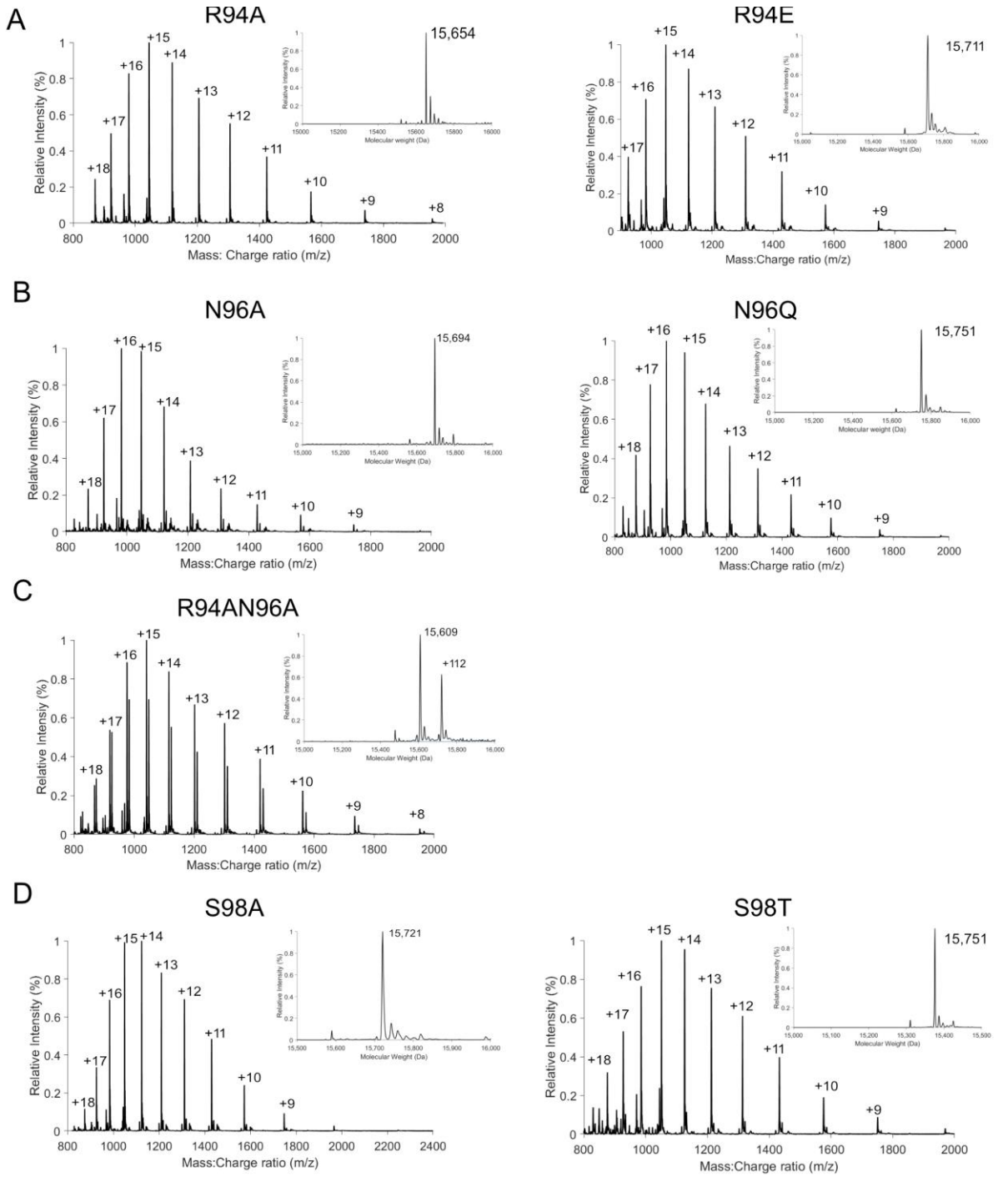
Appendix Figure 23. **Alignment of HicB to 44 homologous sequences across different bacterial strains.** Sequences were aligned using ClustalW [291] and conservation at specific amino acid positions is highlighted: conserved (*), conservative (:) and semi-conservative (.).



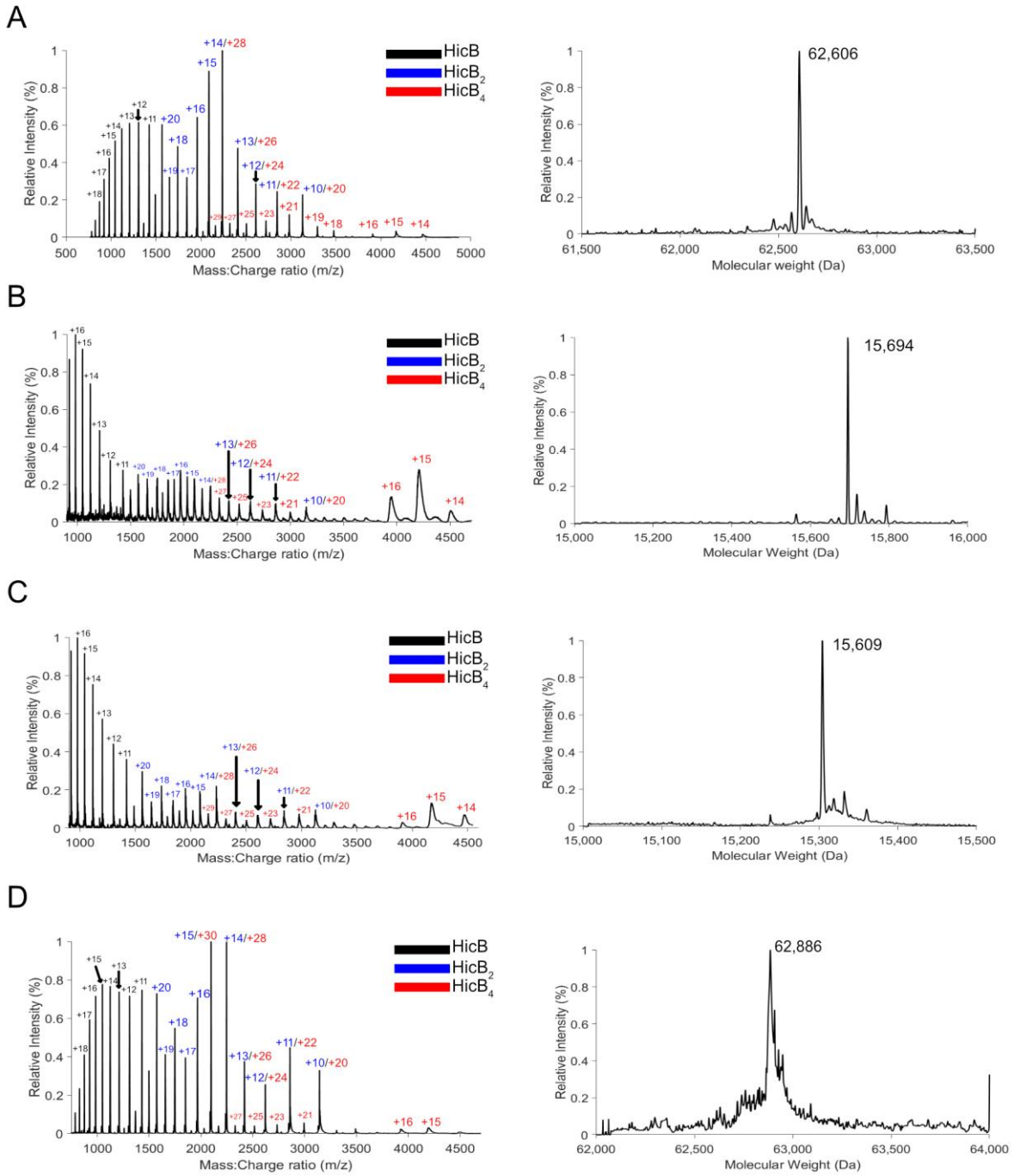
Appendix Figure 24. SDS-PAGE following the purification of all HicB alanine mutants via IMAC and SEC. (A) R94A. (B) N96A. (C) R94AN96A. (D) S99A. Arrows refer to the band corresponding to monomeric HicB for each mutant.



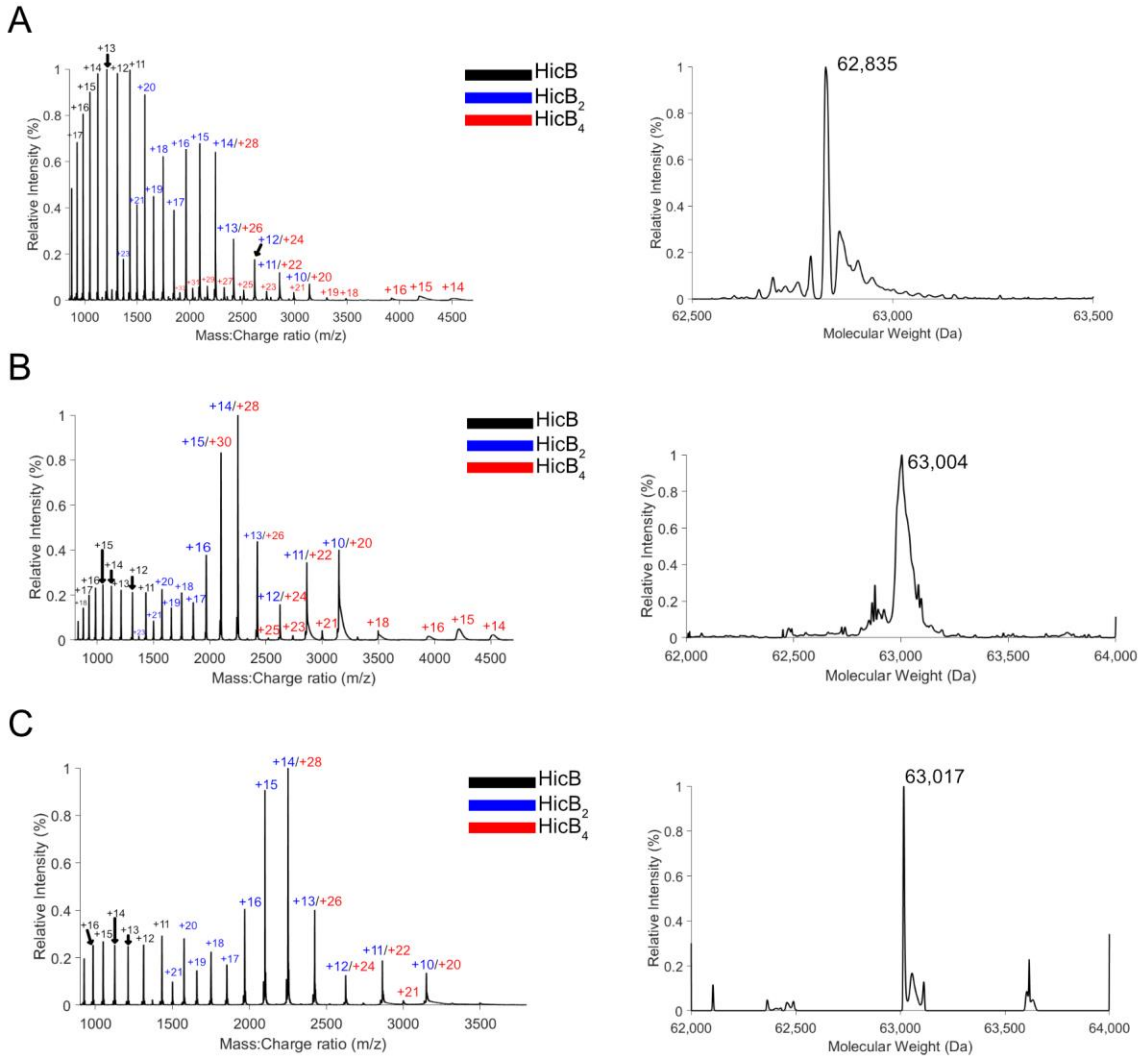
Appendix Figure 25. SDS-PAGE following the purification of all HicB alanine mutants via IMAC and SEC. (A) R94E. (B) N96Q (C) S99T. Arrows refer to the band corresponding to monomeric HicB for each mutant.



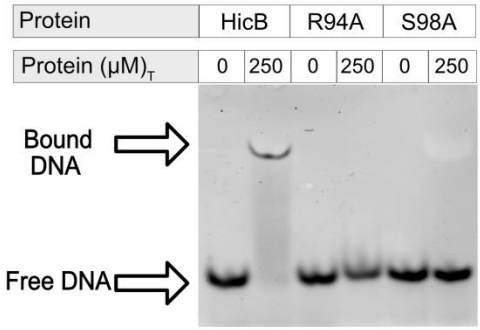
Appendix Figure 26. Denatured ESMS of all HicB mutants. (A) (left) Denatured ESMS of R94A showed a similar distribution of m/z charge states to HicB, with deconvolution resulting in a single (observed mass: 15,654, expected: 15,653 Da). (right) Denatured ESMS of R94E indicated a single species (observed mass: 15,711, expected: 15,711 Da). (B) (left) Denatured ESMS of N96A indicated a single species (observed mass: 15,694, expected: 15,695 Da). (right) Denatured ESMS of N96Q indicated a single species (observed mass: 15,751, expected: 15,752 Da). (C) Denatured ESMS of R94AN96A indicated a single species (observed mass: 15,609, expected: 15,610 Da). An additional mass adduct (+122 Da) was observed for R94AN96A which was not present after SEC (D) (left) Denatured ESMS of S98A indicated a single species (observed mass: 15,721 Da, expected mass: 15,722 Da). (right) Denatured ESMS of S98T indicated a single species (observed mass: 15,751, expected: 15,752 Da). Each mutant showed a similar distribution of m/z charge states to HicB.



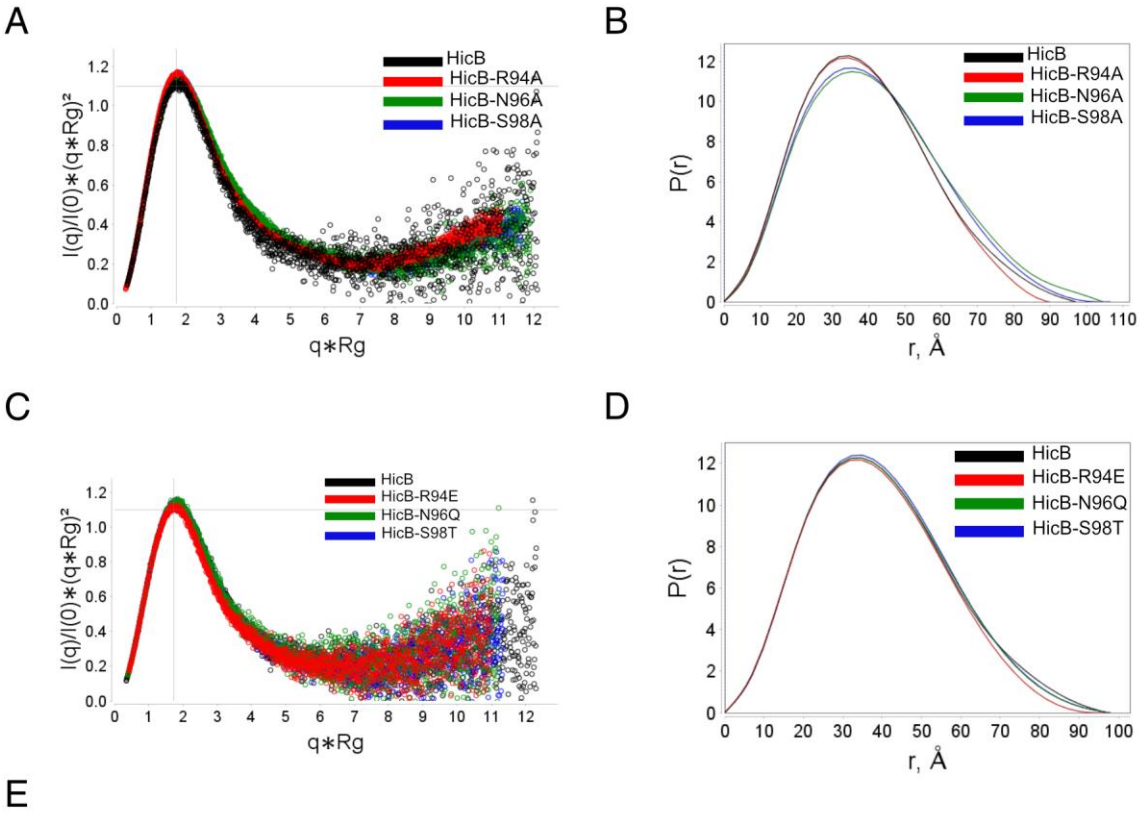
Appendix Figure 27. Native mass spectrometry of HicB alanine mutants: R94A (A), N96A (B), R94AN96A (C) and S98A (D). Peaks were assigned based on their charge state: monomeric (black), dimeric (blue) or tetrameric (red). Deconvolution of each spectrum where shown and tetrameric species for R94A (observed mass: 62,606, expected: 62,612 Da) and S98A (observed mass: 62,886, expected: 62,888 Da). The m/z envelopes of N96A and R94AN96A could not be deconvoluted into accurate tetrameric species, although charge states attributed to tetrameric species are observed.



Appendix Figure 28. Native mass spectrometry of HicB alanine mutants: R94E (A), N96Q (B) and S98T (C). Peaks were assigned based on their charge state: monomeric (black), dimeric (blue) or tetrameric (red). Deconvolution of each spectrum showed tetrameric species for R94A (observed mass: 62,835 expected: 62,844 Da), N96Q (observed mass: 63,004, expected 63,008 Da and S98T (observed mass: 63,017, expected: 62,008 Da).

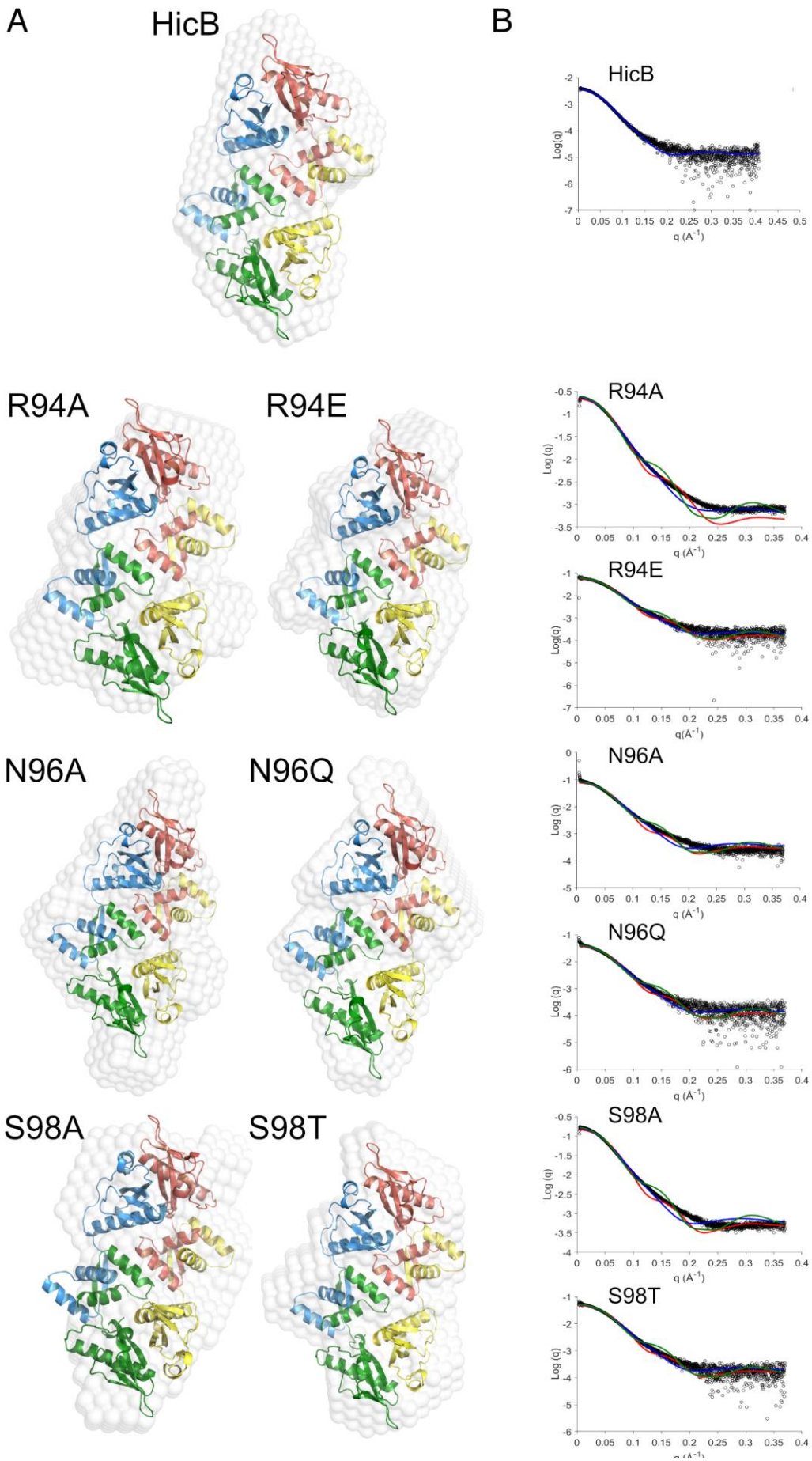


Appendix Figure 29. Gel-shift assay of $HicB_T$ and R94A/S98A mutants at 250 μM to investigate their binding against 2 μM S1-2.



	HicB	R94A	R94E	N96A	N96Q	S98A	S98T
Start	28	72	49	62	42	64	62
End	750	1000	1000	1000	1000	1000	1000
Io Real	3.78E-3	2.39E-1	5.89E-2	8.80E-2	3.98E-2	1.66E-1	5.20E-2
Io Reciprocal	3.82E-3	2.36E-1	5.98E-2	9.15E-2	4.08E-2	1.72E-1	5.35E-2
Rg Real (Å)	30.12	29.79	29.54	32.21	30.02	31.75	30.11
Rg Reciprocal (Å)	30.67	29.51	29.76	32.24	30.82	31.99	30.85
R average (Å)	39.7	38.7	38.4	41.7	38.9	41.1	39.1
Dmax (Å)	97	96	97	104.5	97.5	106.5	98
χ^2	1.94	0.55	0.95	1.46	0.45	1.87	1.51

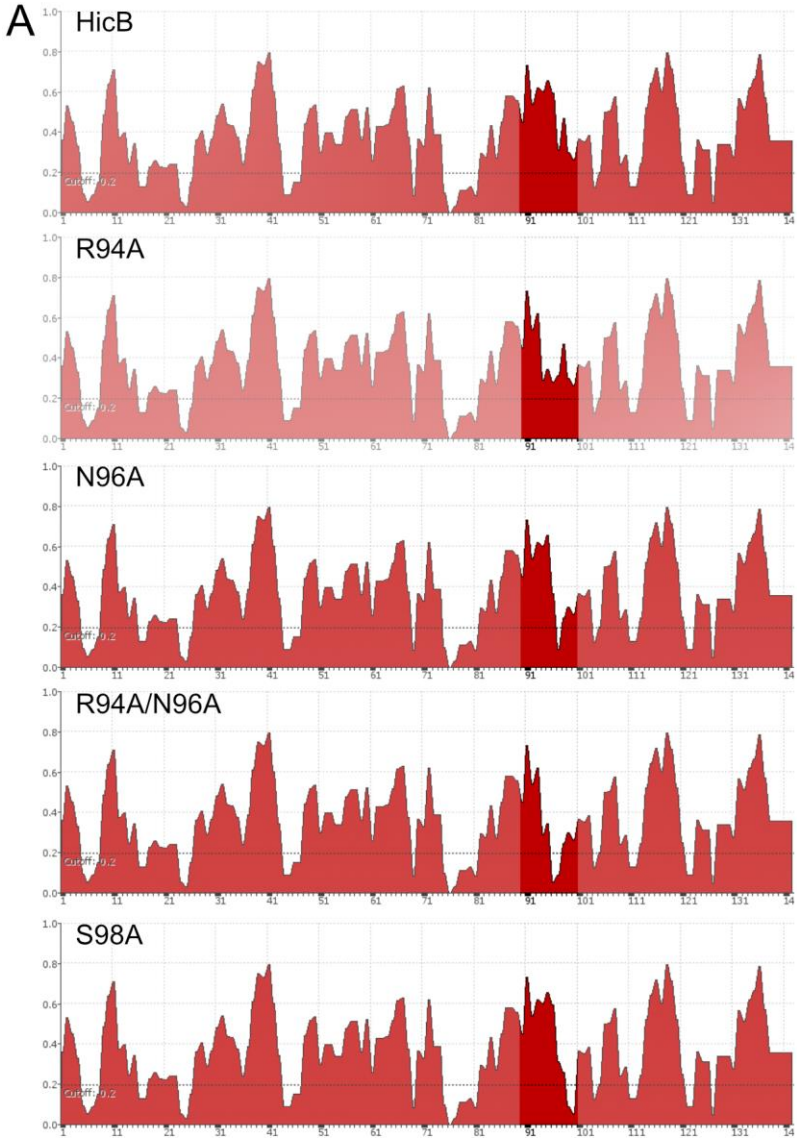
Appendix Figure 30. SAXS analysis of the HicB binding mutants. (A-B) SEC-SAXS of HicB, R94A, N96A and S98A. (A) Dimensionless Kratky plot for each protein. Cross hairs marks the Guinier-Kratky plot (1.732, 1.1)- the main peak for globular particles. (B) Pair-distance P(r) distribution function for each protein. The maximum particle dimension (dmax) is the largest non-negative value that supports a smooth distribution function (x axis intercept). (C-D) SEC-SAXS of HicB, R94E, N96Q and S98T. (C) Dimensionless Kratky and (D) P(r) distribution of each mutant. (E) Parameters obtained from the P(r) distribution fitting for HicB and DNA binding mutants. For each particle both the Io (Maximum scattering intensity) and Rg are calculated in the real and reciprocal domain, while the R average, Dmax and resulting χ^2 value of the fit are reported.

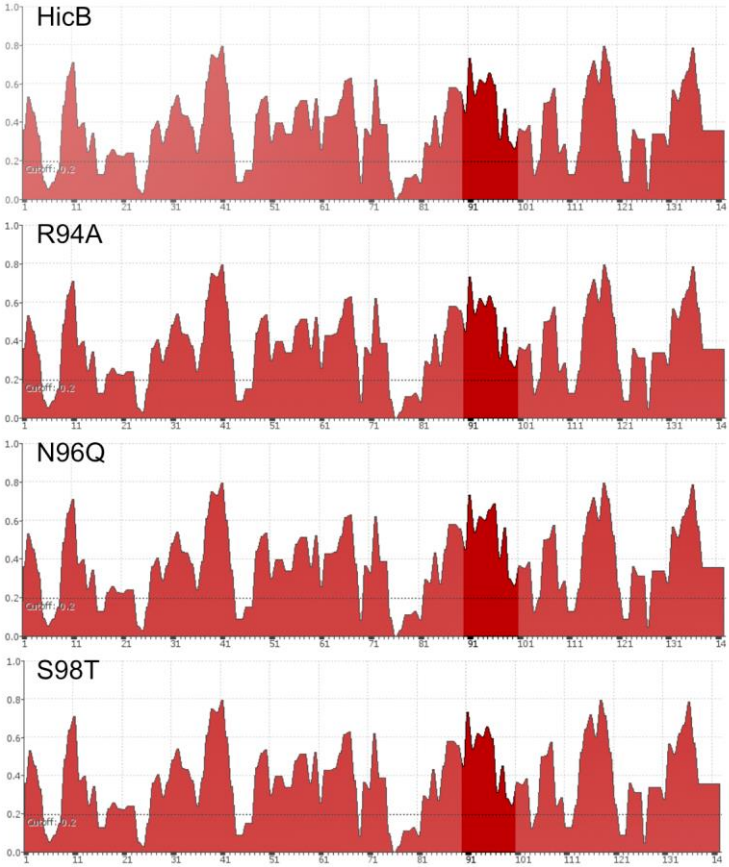

A
HicB
B
HicB
R94A
R94E
R94A
R94E
N96A
N96Q
N96A
N96Q
S98A
S98T
S98A
S98T
Log(q)
Log (q)
q (Å^{-1})

Appendix Figure 31. *Ab initio* modelling of each shape envelope. (left) Superimposition of the elongated conformation of HicB onto the resultant calculated shape envelope (white) for each mutant. (right) Quantitative analysis of the fit of each HicB conformation: compact (red), elongated (blue) and symmetrical HicB (green) to the raw scattering data of each curve (black). χ^2 values are reported in Appendix Figure 30.

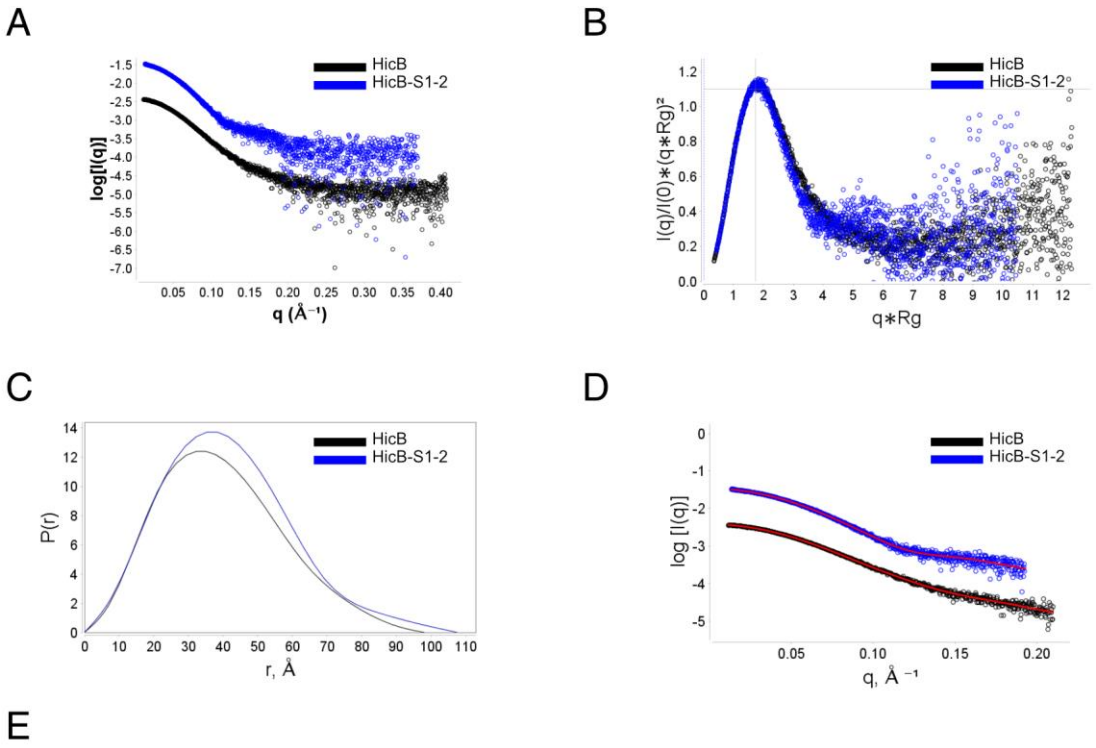
	χ^2		
	Compact	**Elongated**	**Symmetrical**
HicB	9	1.99	7.96
R94A	278.56	36.72	328.70
R94E	13.25	2.34	17.56
N96A	153.01	55.05	51.41
N96Q	12.11	2.62	10.69
S98A	370.56	114.92	136.89
S98T	15.84	2.89	13.91

Appendix Figure 32. χ^2 values for each HicB conformation against the raw scattering data of each analysed mutant to determine the validity of *ab initio* modelling and if mutants retained the same scattering characteristics as HicB.



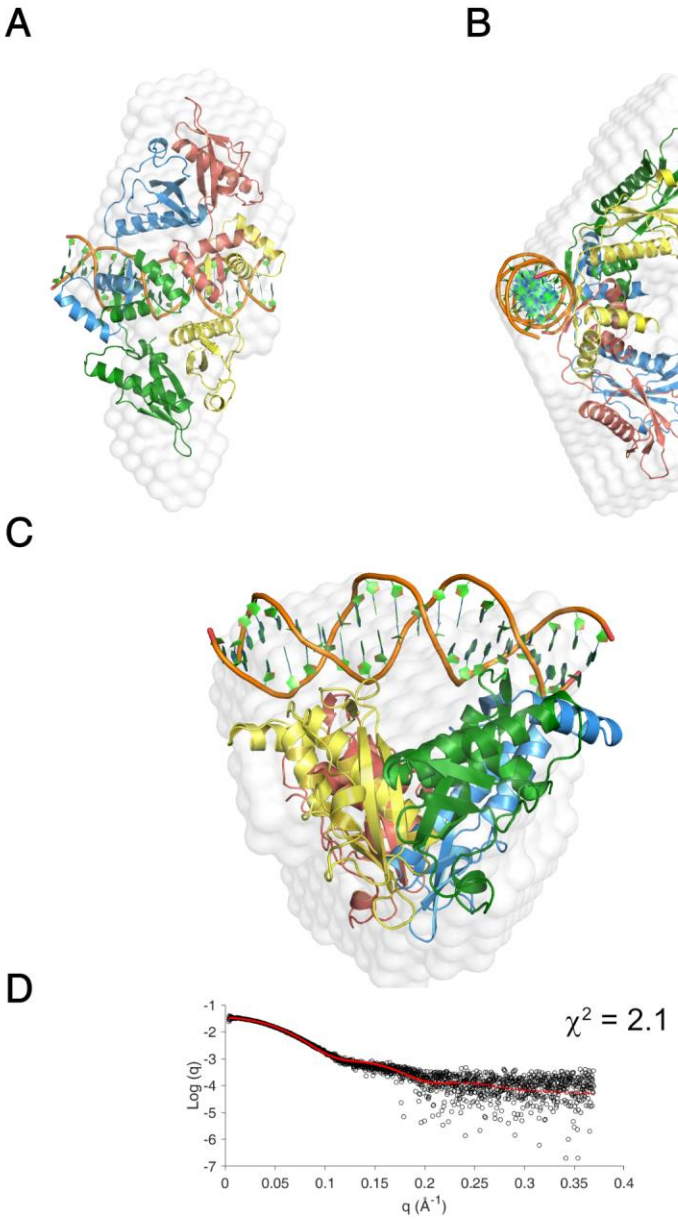


Appendix Figure 33. Surface entropy of all HicB mutant. (A) Normalised sidechain entropy average at each residue position as given by Sternberg, computed in a three residue window for one subunit of each HicB mutant compared to HicB as computed by the SERp server.[394, 395] Residues with entropy average above 0.2 contribute to the overall score. The axis is normalised so alanine = 0 and lysine = 1. The region encompassing the β5 strand is highlighted.



	HicB	HicB-S1-2
Start	28	45
End	750	850
Io Real	3.78E-3	3.44E-2
Io Reciprocal	3.82E-3	3.42E-2
Rg Real (Å)	30.12	31.39
Rg Reciprocal (Å)	30.67	31.74
R average (Å)	39.7	40.7
Dmax (Å)	97	107.5
χ^2	1.94	1.26

Appendix Figure 34. SAXS of HicB-S1-2. (A) Log_{10} SAXS intensity versus scattering vector (q). (B) Dimensionless Kratky plot. Cross-hair marks the Guinier-Kratky plot (1.732, 1.1), the main peak for folded globular particles. (C) Pair-distance, P(r) distribution function. Maximum dimension, d_{max} is the largest non-negative value that supports a smooth distribution function. (D) Fit of the parameters computed by the P(r) distribution (red line) to the raw scattering data (green). Within each figure (apart from (D)) HicB and HicB-S1-2 are highlighted in black and red. (E) Parameters obtained from the P(r) distribution fitting for HicB and HicB-S1-2. For each particle both the Io (Maximum scattering intensity) and Rg are calculated in the real and reciprocal domain, while the R average, Dmax and resulting χ^2 value of the fit are reported.



Appendix Figure 35. *Ab initio* modelling of the shape envelope of the HicB-S1-2 SAXS profile. (A) Front view of the HicB-S1-2 HADDOCK model superimposed on the shape envelope of HicB-S1-2 (white). (B) 90 ° view of Panel (A). (C) Side on view of the HicB-S1-2 HADDOCK model on the shape envelope of HicB-S1-2 (white). (D) Quantitative analysis of the fit of the HicB-S1-2 HADDOCK model (red) to the raw scattering data of the HicB-S1-2 profile(black).

9. **Appendix**

TA subtype	Exception	Reference
Type 1		
LdrD	Bactericidal phenotype with observed nucleoid condensation	91
BsrG	Bactericidal phenotype (cell lysis): Delocalises MreB (Actin homologue)	396
RalR	Not a short hydrophobic peptide, exerts DNase activity	397
SymE	Not a short hydrophobic peptide, exerts mRNase activity	398
Type 4		
AbiEii	Plasmid stabilisation and phage resistance similar to type 3 systems	399, 400

Appendix Table 1. Non-canonical TA systems. The TA subtype (1-6) and the non-canonical feature associated with a specific TA system is detailed.

	Cross talk between TA systems	Reference
Type 1	*txpA/RatA* and the type II *mazEF* system exhibit transcriptional cross talk	401
Type 2		
CcdAB	Bacteria that encode multiple ccdAB operons display cross talk via protein-protein interactions	402
MqsRA	Directly regulates the type 5 GhoST system	403
RelBE	Bacteria that encode multiple relBE operons display cross talk via protein-protein interactions	404
RelBE	*relBE* operon is transcriptionally upregulated by MazF, MqsR, HicA and HipA	405
RelBE	RelE transcriptionally activates *mqsRA, dinJ-yafQ, hicAB and yefm-yoeb*	406, 70, 407
VapBC	*M. tuberculosis* encodes 43 VapBC family members and displays cross talk via protein-protein interactions	
Type 4		
CbtA, YafW and YfiJ	Each antitoxin interacts with FtsZ and MreB in a protective manner (protein-protein) to prevent toxicity mediated by their cognate toxins.	108

Appendix Table 2. Cross talk between TA systems. The TA subtype (1-6) and type of cross talk observed (transcription/protein-protein)

Construct	Protein sequence
pET26-b HicA	MKYLLPTAAAGLLLLAAQP**AMA**MGGS**HHHHHH**GMASMTGGQQMGRDLY**DDDDK**DRTGSELMNSSKLIRMLEED GWRLVRVTGSAHHFKHPKKPGLVTVPHPKKDLPIGTVKSIQKSAGL
HicA	M**HHHHHH**GKPIPNPLLGLDST**ENLYFQ**GIDPFTNSSKLIRMLEEDGWRLVRVTGSAHHFKHPKKPGLVTVPHP KKDLPIGTVKSIQKSAGL
pET26-b HicB	MKYLLPTAAAGLLLLAAQP**AMA**MGGS**HHHHHH**GMASMTGGQQMGRDLYDDDDKDRTGSELMEFPIAVHKDDGS VYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEFTCSTVEELVAKPEYAGAVWALVSVDLS QLDSKPERINVSIPRFVLHKIDAYVASRHETRSGFLARAALEALNEGKVRHA
HicB_FL	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPERINVSIPRFVLHKIDAYVASRHETRSGFLARAALEALNEGKVRHA**KHHHHHH**
HicB	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPERINVSIPRFVLHKIDAYVASRHETRSGFLARAALEALNEGK**KHHHHHH**
HicB_I51M	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETL**M**ELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPERINVSIPRFVLHKIDAYVASRHETRSGFLARAALEALNEGK**KHHHHHH**
HicB_I99M	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPERINVS**M**PRFVLHKIDAYVASRHETRSGFLARAALEALNEGK**KHHHHHH**
HicB_DM	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETL**M**ELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPERINVS**M**PRFVLHKIDAYVASRHETRSGFLARAALEALNEGK**KHHHHHH**
HicB_R94A	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPE**A**INVSIPRFVLHKIDAYVASRHETRSGFLARAALEALNEGK**KHHHHHH**
HicB_R94E	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPE**E**INVSIPRFVLHKIDAYVASRHETRSGFLARAALEALNEGK**KHHHHHH**
HicB_N96A	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPERI**A**VSIPRFVLHKIDAYVASRHETRSGFLARAALEALNEGK**KHHHHHH**
HicB_N96Q	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPERI**Q**VSIPRFVLHKIDAYVASRHETRSGFLARAALEALNEGK**KHHHHHH**
HicB_R94A/N 96A	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPE**A**I**A**VSIPRFVLHKIDAYVASRHETRSGFLARAALEALNEGK**KHHHHHH**
HicB_S98A	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPERINV**A**IPRFVLHKIDAYVASRHETRSGFLARAALEALNEGK**KHHHHHH**
HicB_S98T	**M**MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEFTCSTVEELVAKPE YAGAVWALVSVDLSQLDSKPERINV**T**IPRFVLHKIDAYVASRHETRSGFLARAALEALNEGK**KHHHHHH**
HicB_NT	MEFPIAVHKDDGSVYGVTVPDIPGVHSWGETIDDAIKNTREAIVGHVETLIELGEDVEF**KHHHHHH**

Appendix Table 3. List of protein sequences used in this project. His_6-tags (Bold), and protease cleavage sequences (Bold, underlined), mutation sites (bold) and additional methionine residues incorporated during cloning (Bold) are highlighted.

Oligonucleotides		
0-48 Primer 1 (*)	IDT	gatcgtgattggatgtgtataattacacacaagacattcgggggagct
0-48 Primer 2 (*)	IDT	agctcccccgaatgtcttgtgtgtaattatacacatccaatcacgatc
48-96 Primer 1 (*)	IDT	ataggggcgaaacaatgtgaaaatacgcacggctacacaaaacttgag
48-96 Primer 2 (*)	IDT	ctcaagttttgtgtagccgtgcgtattttcacattgtttcgcccctat
96-144 Primer 1 (*)	IDT	tcgaggccgcgcgcgatgcttcagtcttgcccagcggacgggataaaa
96-144 Primer 2 (*)	IDT	ttttatcccgtccgctgggcaagactgaagcatcgcgcgcggcctcga
144-196 Primer 1 (*)	IDT	gccgcccgctggcggcagattacgacgagtggctacacacgaggaatg
144-196 Primer 2 (*)	IDT	cattcctcgtgtgtagccactcgtcgtaatctgccgccagcgggcggc
192-240 Primer 1 (*)	IDT	gtgccgaataagggtaactatcctgtcaatgttgacgggcaagaggtg
192-240 Primer 2 (*)	IDT	cacctcttgcccgtcaacattgacaggatagttacccttattcggcac
240-279 Primer 1 (*)	IDT	gccgaatcctcagtggaaaatccgcgaggactttagcac
240-279 Primer 2 (*)	IDT	gtgctaaagtcctcgcggattttccactgaggattcggc
0-20 Primer 1	IDT	acaagacattcgggggagct
0-20 Primer 2	IDT	agctcccccgaatgtcttgt
0-25 Primer 1	IDT	tacacacaagacattcgggggagct
0-25 Primer 2	IDT	agctcccccgaatgtcttgtgtgta
0-30 Primer 1	IDT	ataattacacacaagacattcgggggagct
0-30 Primer 2	IDT	agctcccccgaatgtcttgtgtgtaattat
0-35 Primer 1	IDT	tgtgtataattacacacaagacattcgggggagct
0-35 Primer 2	IDT	agctcccccgaatgtcttgtgtgtaattatacaca
0-40 Primer 1	IDT	ttggatgtgtataattacacacaagacattcgggggga
0-40 Primer 2	IDT	agctcccccgaatgtcttgtgtgtaattatacacatc
16-36 Primer 1 (*)	IDT	atgtgtataattacacacaag
16-36 Primer 2 (*)	IDT	cttgtgtgtaattatacacat
17-36 Primer 1 (*)	IDT	atgtgtataattacacacaa
17-36 Primer 2 (*)	IDT	ttgtgtgtaattatacacat
17-40 Primer 1 (*)	IDT	ttggatgtgtataattacacacaa
17-40 Primer 2 (*)	IDT	ttgtgtgtaattatacacatccaa
20-36 Primer 1 (*)	IDT	atgtgtataattacaca
20-36 Primer 2 (*)	IDT	tgtgtaattatacacat
20-40 Primer 1 (*)	IDT	ttggatgtgtataattacaca
20-40 Primer 2 (*)	IDT	tgtgtaattatacacatccaa
MS1 Primer 1	IDT	atgtgtataattagggacaa
MS1 Primer 2	IDT	ttgtccctaattatacacat
MS2 Primer 1	IDT	atccctataattacacacaa
MS2 Primer 2	IDT	ttgtgtgtaattatagggat
MS1MS2 Primer 1	IDT	atccctataattagggacaa
MS1MS2 Primer 2	IDT	ttgtccctaattatagggat
Overhang Primer 1	IDT	tatgtgtataattacacacaa
Overhang Primer 2	IDT	attgtgtgtaattatacacat
HicB_FL Primer 1	IDT	aggagatataccatgatggaatttcccatcgcagt
HicB_FL Primer 2	IDT	gtgatggtgatgttttgcgtgcctaactttgccttc
HicB Primer 1	IDT	aggagatataccatgatggaatttcccatcgca
HicB Primer 2	IDT	gtgatggtgatgtttttttgccttcattaagtgcctc
HicB-NT Primer 1	IDT	aggagatataccatggaatttcccatcgcagtg
HicB-NT Primer 2	IDT	gtgatggtgatgtttaagatcaacgctgacgag
I51M Primer 1	IDT	catgtagagacattgatggagcttggagaagat
I51M Primer 2	IDT	atcttctccaagctccatcaatgtctctacatg
I99M Primer 1	IDT	cggatcaatgtgagtatgcctcgcttcgtgctg
I99M Primer 2	IDT	cagcacgaagcgaggcatactcacattgatccg
Mutant Primer 1	IDT	aggagatataccatgatggaatttccgattgccgtgcat
Mutant Primer 2	IDT	gcactggaagcactgaatgaaggtaaaaaacatcaccatcac
R94E Primer 1	IDT	aaaccggaagaaattaatgttagc
R94E Primer 2	IDT	gctaacattaatttcttccggttt
S98T Primer 1	IDT	cgtattaatgttaccattccgcgtttt
S98T Primer 2	IDT	aaaacgcggaatggtaacattaatacg

Appendix Table 4. List of primers used throughout the project (*) indidicates DNA that may form mis-annealed higher order DNA fragments.

Monomer	N_{at}	N_{res}	S_{at}	S_{res}	Area (Å^2)	ΔG (kcal/mol)
A	615	80	406	75	5200	-74
B	655	86	413	81	53004	-79
C	628	82	383	77	50000	-77
V	646	85	423	80	5500	-76

Non-stable crystal assemblies							
Split No	**Size**	**Type**	**ASA (Å^2)**	**BSA (Å^2)**	**ΔG_diss (kcal/mol)**	**ΔG0 (kcal/mol)**	**Composition**
	2	1	9300	1230	0.0	0.0	AB
1	1	2	5000	0.0	0.0	0.0	C
	1	2	5400.1	0.0	0.0	0.0	V
	4	26	17000	4215	-8	-2	ABCV

Appendix table 5. PISA analysis of HicB_NT ASU. Monomers were investigated for their total number of atoms (N_{at}), residues (N_{res}), total number of surface atoms (S_{at}) and residues (S_{res}). The total surface area (Area) and solvation energy of folding (ΔG) were calculated. Assemblies (Crystal splits) were investigated with the Accessible Solvent Area (ASA), Buried Surface Area (BSA), Standard Free Energy of dissociation into nearest stable assemblies (ΔG_diss), Standard Free Energy of dissociation into monomeric units (ΔG0) and Composition of assembly reported. It was reported that the 4-chain assembly in the ASU was not stable in solution. Chain A, B, C and V correspond to Subunit 2, Monomer 1, Monomer 2 and Monomer 3 in Figure 2.10.

Monomer	N_{at}	N_{res}	S_{at}	S_{res}	Area (Å^2)	ΔG (kcal/mol)
C	629	82	387	77	5000	-77.0
D	646	85	426	80	5500	-77

Metastable crystal assemblies							
Split No	**Size**	**Type**	**ASA (Å^2)**	**BSA (Å^2)**	**ΔG_diss (kcal/mol)**	**ΔG0 (kcal/mol)**	**Composition**
1	2	1	9200	1300	2	2	CD

Appendix table 6. PISA analysis of a HicB_NT dimer. Monomers were investigated for their total number of atoms (N_{at}), residues (N_{res}), total number of surface atoms (S_{at}) and residues (S_{res}). The total surface area (Area) and solvation energy of folding (ΔG) were calculated. Assemblies (Crystal splits) were investigated with the Accessible Solvent Area (ASA), Buried Surface Area (BSA), Standard Free Energy of dissociation into nearest stable assemblies (ΔG_diss), Standard Free Energy of dissociation into monomeric units (ΔG0) and Composition of assembly reported. It was reported that the dimer had a higher stability than the ASU assembly of HicB_NT, however this dimer organisation may also not be stable in solution. Chain C and D refer to Subunit 1 and Subunit 2 in Figure 2.10.

Monomer	N_{at}	N_{res}	S_{at}	S_{res}	Area (Å^2)	ΔG (kcal/mol)
A	1019	132	724	128	9300	-106
B	1019	132	719	128	9300	-107
C	1041	135	729	130	9700	-109
D	1041	135	725	129	9700	-107

Stable crystal assemblies							
Split No	**Size**	**Type**	**ASA (Å^2)**	**BSA (Å^2)**	**ΔG_diss (kcal/mol)**	**ΔG0 (kcal/mol)**	**Composition**
1	4	1	27000	11000	13	14.0	ABCD
2	2	2	15300	3700	22.0	11.0	BC
	2	2	15300	3700	22	11	AD
Metastable crystal assemblies							
3	2	3	17000	1600	0.6	1	AC
	2	3	17000	1600	0.0	0.0	BD

Appendix Table 7. PISA analysis of HicB tetramer observed within the 1.85 Å dataset. Monomers were investigated for their total number of atoms (N_{at}), residues (N_{res}), total number of surface atoms (S_{at}) and residues (S_{res}). The total surface area (Area) and solvation energy of folding (ΔG) were calculated. Assemblies (Crystal splits) were investigated with the Accessible Solvent Area (ASA), Buried Surface Area (BSA), Standard Free Energy of dissociation into nearest stable assemblies (ΔG_diss), Standard Free Energy of dissociation into monomeric units (ΔG0) and Composition of assembly reported. It was reported that the tetramer was stable in solution. Chains A, B, C and D refer to subunits 2, 3, 1 and 4 in figure 2.16.

Entry	Interface Number	mmSize	Q score	Seq Id	Interface area, Å^2	ΔG kcal/mol	CSS
4P7D	1	4	0.395	0.273	2100	-30	0.844
4P78	3	4	0.369	0.268	900	-22	0.931
5YRZ	3	4	0.284	0.194	600	-8	0.268
3KWR	2	4	0.269	0.096	500	-9	0.815

Appendix Table 8. Structural conservation of the HicB interfaces within the PDB. The PISA server reports the PDB entry, interface number, mmSize (number of macromolecular monomers within the assembly), Q score (relates from 0 to 1 for unrelated-identical proteins), Sequence identity (Seq ID), interface area, ΔG and Complexation significance score (CSS, relates from 0 to 1 as the interface relevance increases). HicB3 (PDB:4P7D), HicA3-HicB3 (PDB:4P78), HicAB (PDB:5YRZ) and a uncharacterized RNA binding protein from *Lactobacillus plantarum* (PDB: 3KWR) were reported.

Chain	Z score	RMSD	Identity	Description
4P7D-A	11.8	10.9	25	Antitoxin HicB3 from *Y. pestis*
5YRZ-A	9.9	2.6	23	Antitoxin HicB from *S. pneumonaie*
3KWR-B	10	3.7	15	Putative RNA binding protein
2DSY-C	6.1	2.2	21	Hypothetical protein TTHA0281
3K6Q-A	5.0	2.3	14	Putative ligand binding protein from *Syntrophomonas Wolfei*
1ZBT-A	4.5	10	13	Peptide chain release factor
6CI7-F	3.9	14.3	8	YCAO
5CFF-F	3.7	2.2	10	Miranda/Staufen dsRBD5 complex
3VYY-A	3.6	3.5	10	ATP-Dependent RNA helicase A
4WYQ-B	3.5	2.6	8	Endoribonuclease DICER
2MDR-A	3.5	2.2	4	Double-stranded RNA-specific adenosine deaminase
2LTR-A	3.4	2.6	7	Protein RDE-4
3P1X-B	3.4	2.0	8	Interleukin enhancer binding factor 3
3ADL-A	3.3	2.4	14	RISC-loading complex subunit TARBP2
1WHQ-A	3.1	2.9	11	RNA helicase A

Appendix Table 9. The DALI server search result against a single subunit of HicB. The PDB chain, Z score, RMSD, sequence identity and description are highlighted. The Z score relates to the structural similarity between proteins. Proteins with a Z score below 3 were not included in this as this represented a low structural similarity to HicB.

Structure	Z score	RMSD (Å)	Identity	Description
Arc (PDB:1BAZ)	5.9	1.3	15	Arc repressor
CopG (PDB: 2CPG)	6.1	1.3	13	Transcriptional repressor CopG
DinJ (PDB: 4Q2U)	5.2	2.7	13	Antitoxin DinJ
MazF-4 (PDB:5XE3)	4.9	1.6	8	Endoribonuclease MazF
ORF Omega (PDB: 2BNZ)	5.7	1.5	15	ORF Omega
PutA (PDB: 2AY0)	5.9	1.8	13	Bifunctional PutA protein
RelB (PDB: 4FXE)	4.0	2.7	18	Antitoxin RelB
VapB26 (PDB: 5X3T)	4.6	1.3	10	Antitoxin VapB26

Appendix Table 10. The DALI server search result against the RHH domain of HicB. Structures which displayed a RMSD between 1.3-2.7 Å are shown, with their Z score, sequence identity and description shown. Fully characterised proteins, implicated in DNA binding are shown.

Data-collection parameters	
Instrument	SEC-SAXS at B21 Diamond Light Source
SEC column	Supedex 200 (GE Healthcare)
Temperature	25°C
q range ($Å^{-1}$)	0.006-0.40
Structural parameters	HicB
I(0) (cm^{-1}) [from P(r)]	3.78E-3 ± 8.9E-6
Rg (Å) [from P(r)]	30.12 ± 0.24
I(0) (cm^{-1}) (from Guinier)	3.82E-3 ± 8.9E-6
Dmax (Å)	97
Rg (Å) (from Guinier)	30.67 ± 0.37
Porod volume estimate ($Å^{-3}$)	142,088
$\chi 2$	1.94
Dry volume calculated from sequence ($Å^{-3}$)	76,104
Molecular-mass determination	
Partial specific volume ($cm^3\ g^{-1}$)	0.7352
Contrast ($\Delta\rho$ x 10^{30} cm^{-2})	3.047
Molecular mass M_r [from I(0)]	67000
Calculated M_r from sequence (kDa)	62,952
From Porod Volume [Vporod/1.7] (Da)	83,581
Software employed	
Primary data reduction	GDA (Diamond Light Source)
Data processing	Scatter
Ab initio analysis	DAMMIN
Validation and averaging	DAMAVER
Rigid-body modelling	N/A
Computation of model intensities	FoXS
Three-dimensional graphics representations	Pymol

Appendix table 11. Statistics and parameters obtained via SAXS experiments of HicB

Monomer	N_{at}	N_{res}	S_{at}	S_{res}	Area (Å^2)	ΔG (kcal/mol)
A	1019	132	724	128	9300	-106
B	1019	132	719	128	9300	-107
C	1041	135	729	130	9700	-109
D	1041	135	725	129	9700	-107
E	479	61	297	56	4000	-55
F	466	60	296	56	4000	-53
G	472	60	294	56	4000	-55
H	461	59	305	57	4200	-50

Stable crystal assemblies							
Split No	**Size**	**Type**	**ASA (Å^2)**	**BSA (Å^2)**	**ΔG_diss (kcal/mol)**	**ΔG0 (kcal/mol)**	**Composition**
1	8	1	36200	17600	9	74	ABCDEFGH
2	4	2	19500	700	9	33	ADFH
	4	2	19200	7600	9	33	BCEG
3	2	3	11500	1900	5	5	CG
	2	3	11700	1800	5	5	DH
	2	3	11600	1800	5	5	AF
	2	3	11600	1900	4	4	BE
4	4	4	27400	10000	10	57	ABCD
Metastable crystal assemblies							
5	4	5	22000	5000	-1	-1	BDEH
	4	5	22000	5000	-1	-1	ACFG
Marginally stable Assemblies							
6	4	4	27000	10000	10	57	ABCD
7	2	6	17000.6	1200	0	0	BD
	2	6	18000.5	1200	0	0	AC

Appendix Table 12. PISA analysis of HicAB hetero-octamer. Monomers were investigated for their total number of atoms (N_{at}), residues (N_{res}), total number of surface atoms (S_{at}) and residues (S_{res}). The total surface area (Area) and solvation energy of folding (ΔG) were calculated. Assemblies (Crystal splits) were investigated with the Accessible Solvent Area (ASA), Buried Surface Area (BSA), Standard Free Energy of dissociation into nearest stable assemblies (ΔG_diss), Standard Free Energy of dissociation into monomeric units (ΔG0) and Composition of assembly reported. It was reported that the tetramer was stable in solution. Chains A, B,C,D refer to Subunit 4, 1, 3 and 2 of HicB, while chains F, E, G and H refer to HicA moieties bound to each respective subunit.

Interaction	HicA-HicB
Hydrophobic	
	V18-L53
	A24-A42
	L35-L50
	L35-L53
	P39-W28
	P41-W28
	L45-V14
	L45-W28
	P46-V14
Hydrogen bonds	
	S23-W28
	T37-H46
Electrostatic	
	R19-E48
	R19-E52
	K28-E55
	K43-E30

Appendix Table 13. List of interactions between HicA and HicB. Residues that have direct equivalents within the HicA3B3 interaction site are underlined.

Entry	Interface Number	mmSize	Q score	Seq Id	Interface area, Å^2	ΔG kcal/mol	CSS
5YRZ	6	4	0.782	0.246	1114.5	-7.9	0.468
4P78	2	4	0.710	0.328	1179.3	-7.5	0.608

Appendix Table 14. Structural conservation of the HicAB interface within the PDB. mmsize indicates the multimeric state of the most probable macromolecular assembly in the PDB entry, Q score is a structure alignment score (0-1, where 1 is an identical structure), Sequence identity (fraction), the total accessible surface area, DG upon formation of the interface and the Complexation Significance Score (CSS), which determines how significant an assembly is (0-1).

Interaction	HicA3-HicB3
Hydrophobic	
	I39-A48
	I39-L49
	I39-F54
	I40-A27
	I40-F25
	L42-F25
	L53-F25
	A57-F25
	I60-F25
Hydrogen bonds	
	N24-D33
	N27-N37
	H28-N37
	T41-H45
	Q56-R71
Electrostatic	
	R18-E52
	R21-347
	H28-E40
	K38-D22
	R46-E12
	R46-D15
	R46-D33
	K51-E75
	R55-E75
	K59-E74

Appendix Table 15. List of interactions between HicA3 and HicB3.

Interaction	$HicA_{SP}$-$HicB_{SP}$
Hydrophobic	
	L50-M44
	Y57-Y10
	Y57-F22
	Y57-F80
	Y57-F86
	Y57-Y90
Hydrogen bonds	
	S35-E47
	H36-T33
	H36-E47
	K38-D55
	E53-Q34
	N55-Q34
	Y57-D12
	Y57-T89
	T58-Q34
	D83-R60
Electrostatic	
	K33-E103
	R30-D145
	H36-E47
	K38-D55
	K38-E58
	R45-D55
	K56-D12
	K56-D15
	R60-D83
	R60-D85
	K64-D83

Appendix Table 16. List of interactions for $HicAB_{SP}$. Residues that have direct equivalents within the HicAB interaction site are underlined. Residues that are not conserved at each $HicAB_{SP}$ interaction site are highlighted in bold.

Chain	Z score	RMSD	Identity	Description
4P78-D	6.4	2.1	33	Toxin HicA3 from *Y. pestis*
1WHZ-A	5.9	2.1	32	Hypothetical protein
5YRZ-B	5.8	2.3	26	Toxin HicA
3P8A-B	4.0	3.9	6	Uncharacterised protein from *S. aureus*
5HCW-A	3.6	2.7	10	Glyoxalase/ Bleomycine resistance protein/ dioxygena
2P25-A	3.5	2.9	15	Glyoxalase family protein
3FCD-B	3.3	2.4	8	Lyase
3HDP-A	3.2	3.0	8	Glyoxalase-I
5TE2-B	3.2	3.5	13	Adensylmethionine-8-Amino-7-oxononanoate aminotransferase

Appendix Table 17. The DALI server search result against a monomer of HicA. The PDB chain, Z score, RMSD, sequence identity and description are highlighted. The Z score relates to the structural similarity between proteins. Proteins with a Z score below 3 were not included in this table as this represented a low structural similarity to HicA.

Data-collection parameters	
Instrument	SEC-SAXS at B21 Diamond Light Source
SEC column	Supedex 200 (GE Healthcare)
Temperature	25°C
q range ($Å^{-1}$)	0.006-0.40
Structural parameters	HicAB
I(0) (cm^{-1}) [from P(r)]	3.49E-6 ± 3.7E-5
Rg (Å) [from P(r)]	31.31 ± 0.21
I(0) (cm^{-1}) (from Guinier)	3.63E-3 ± 9.2E-6
Dmax (Å)	100
Rg (Å) (from Guinier)	31.90 ± 0.47
Porod volume estimate ($Å^{-3}$)	283,640
$\chi 2$	0.75
Dry volume calculated from sequence ($Å^{-3}$)	110.149
Molecular-mass determination	
Partial specific volume (cm^3 g^{-1})	0.7402
Contrast ($\Delta\rho$ x 10^{30} cm^{-2})	3.047
Molecular mass M_r [from I(0)]	74,000
Calculated M_r from sequence (kDa)	91,164
From Porod Volume [Vporod/1.7] (Da)	166,847
Software employed	
Primary data reduction	GDA (Diamond Light Source)
Data processing	Scatter
Ab initio analysis	DAMMIN
Validation and averaging	DAMAVER
Rigid-body modelling	N/A
Computation of model intensities	FoXS
Three-dimensional graphics representations	Pymol

Appendix table 18. Statistics and parameters obtained via SAXS experiments of HicAB.

Data-collection parameters			
Instrument	SEC-SAXS at B21 Diamond Light Source		
Temperature	25°C		
q range ($Å^{-1}$)	0.006-0.40		
Structural parameters	R94A	N96A	S98A
I(0) (cm^{-1}) [from P(r)]	2.27E-1 ±1.2E-3	8.8E-2 ±7.7E-4	1.66E-1 ±1.1E-3
Rg (Å) [from P(r)]	29.74±0.15	32.21±0.16	31.75±0.15
I(0) (cm^{-1}) (from Guinier)	2.36E-1 ±1.1E-4	9.15E-2 ±1.1E-4	1.72E-1 ±9.6E-5
Dmax (Å)	96	104.5	106.5
Rg (Å) (from Guinier)	29.51 ±0.08	33.24 ±0.22	31.99 ±0.09
Porod volume estimate ($Å^{-3}$)	134002	15649	142002
$\chi 2$	0.55	1.46	1.87
Dry volume calculated from sequence ($Å^{-3}$)	75,693	75,896	76,027
Molecular-mass determination			
Partial specific volume (cm^3 g^{-1})	0.738384	0.738384	0.738384
Contrast ($\Delta\rho$ x 10^{30} cm^{-2})	3.047	3.047	3.047
Molecular mass M_r [from I(0)]	65,000	75,000	74,000
Calculated M_r from sequence (kDa)	62,612	62,780	62,888
From Porod Volume [Vporod/1.7] (Da)	78,824	92,022	83,530
Software employed			
Primary data reduction	GDA (Diamond Light Source)		
Data processing	Scatter		
Ab initio analysis	DAMMIN		
Validation and averaging	DAMAVER		
Rigid-body modelling	N/A		
Computation of model intensities	FoXS		
Three-dimensional graphics representations	Pymol		

Appendix table 19. **Statistics and parameters obtained via SAXS experiments of HicB-R94A, N96A and S98A.**

Data-collection parameters			
Instrument	SEC-SAXS at B21 Diamond Light Source		
Temperature	25°C		
q range ($Å^{-1}$)	0.006-0.40		
Structural parameters	R94E	N96Q	S98T
I(0) (cm^{-1}) [from P(r)]	5.59E-2 ±7.3E-4	3.98E-2 ±4.3E-4	5.20E-2 ±8.2E-4
Rg (Å) [from P(r)]	29.54±0.25	30.02±0.21	30.11±0.25
I(0) (cm^{-1}) (from Guinier)	5.98E-2 ±9.1E-5	4.07E-2 ±7.9E-5	5.35E-2 ±9.5E-5
Dmax (Å)	97	97.5	98
Rg (Å) (from Guinier)	29.76 ±0.25	30.82 ±0.35	30.85 ±0.30
Porod volume estimate ($Å^{-3}$)	133,369	136,319	138,869
$\chi 2$	0.95	0.45	1.51
Dry volume calculated from sequence ($Å^{-3}$)	75,973	76,172	76,172
Molecular-mass determination			
Partial specific volume ($cm^3\ g^{-1}$)	0.738384	0.738384	0.738384
Contrast ($\Delta\rho$ x 10^{30} cm^{-2})	3.047	3.047	3.047
Molecular mass M_r [from I(0)]	66,000	67,000	67,000
Calculated M_r from sequence (kDa)	62,844	63,008	63,008
From Porod Volume [Vporod/1.7] (Da)	78,452	80,187	81,687
Software employed			
Primary data reduction	GDA (Diamond Light Source)		
Data processing	Scatter		
Ab initio analysis	DAMMIN		
Validation and averaging	DAMAVER		
Rigid-body modelling	N/A		
Computation of model intensities	FoXS		
Three-dimensional graphics representations	Pymol		

Appendix table 20. **Statistics and parameters obtained via SAXS experiments of HicB-R94E, N96Q and S98T.**

Data-collection parameters	
Instrument	SEC-SAXS at B21 Diamond Light Source
SEC column	Supedex 200 (GE Healthcare)
Temperature	25°C
q range ($Å^{-1}$)	0.006-0.38
Structural parameters	HicB-DNA
I(0) (cm^{-1}) [from P(r)]	3.43E-2 ± 4.3E-4
Rg (Å) [from P(r)]	31.39 ± 0.26
I(0) (cm^{-1}) (from Guinier)	3.42E-2 ± 8.2E-5
Dmax (Å)	107.5
Rg (Å) (from Guinier)	31.74 ± 0.41
Porod volume estimate ($Å^{-3}$)	280399
$\chi 2$	1.26
Dry volume calculated from sequence ($Å^{-3}$)	86,104
Molecular-mass determination	
Partial specific volume (cm^3 g^{-1})	0.7402
Contrast ($\Delta\rho$ x 10^{30} cm^{-2})	3.047
Molecular mass M_r [from I(0)]	78,000
Calculated M_r from sequence (kDa)	75,181
From Porod Volume [Vporod/1.7] (Da)	164,940
Software employed	
Primary data reduction	GDA (Diamond Light Source)
Data processing	Scatter
Ab initio analysis	DAMMIN
Validation and averaging	DAMAVER
Rigid-body modelling	N/A
Computation of model intensities	FoXS
Three-dimensional graphics representations	Pymol

Appendix table 21. Statistics and parameters obtained via SAXS experiments of HicB-S1-2.

9. **Appendix**

TA System	Bacteria	Plasmid/ Chromosome	No of operator sites	IDR within the antitoxin	Alternating T:AT complex	Co-repression observed	De-repression observed	Reference
Axe-Txe	*Enterococcus faecium*	Plasmid PRum	2 palindromes	N/A	N/A	Yes	N/A	121
CcdAB	*E. coli*	Plasmid F	8 palindromes	Yes	B_2-A_2-B_2-A_2-B_2	Yes	Yes B_2-A_2-B_2	85, 125, 126
DinJ/YafQ	*E. coli*	Chromosome	1 inverted repeat	No	No	No	No	83
HicAB	*E. coli*	Chromosome	1 palindrome	No	N/A-assumed no	No	Yes	214
HicAB	*B. pseudomallei*	Chromosome	1 palindrome	No	No Enclosed	No	Yes	This study
HicAB	*Y. pestis*	Chromosome	2 inverted repeats	No	N/A-assumed no	No	Yes (not published)	132
HicAB	*S. pneumoniae*	Chromosome	1 palindrome[1]	No	No Enclosed	Yes	N/A	272
HigBA	*Vibrio cholerae*	Chromosome	1 inverted repeat[2]	Yes[3] (Hadzi 2017)	No (Hadzi 2017)	No (Budde 2007)	Yes (Hadzi, unpublished results)	366-368
HigBA	*Proteus vulgaris*	Plasmid Rts1	2 inverted repeats	N/A	No (Schureck)	Yes (Tian 1996, 2001)	N/A	369, 370, 215
HigBA	*Mycobacterium tuberuclosis*	Chromosome	1 palindrome	N/A	N/A-assumed no	N/A	N/A	408
HigBA	*Acinectobacter baumanni*	Plasmid	N/A	N/A	N/A-assumed no	Yes	N/A	216
HipAB	*E. coli*	Chromosome	4 operator sites	Yes[3]	Yes	No	N/A	43, 135, 137
HipAB	*Shewanella oneidensis*	Chromosome	4 operator sites	N/A	No Enclosed	No	No	217
FitAB	*Neisseria gonorrhoae*	Chromosome	2 inverted repeats	No	No Enclosed	Yes	No	142, 218
Kid-Kis	*E. coli*	Plasmid 1	2 operator sites	N/A	Yes	Yes	Yes	219
MazEF	*E. coli*	Chromosome	3 operator sites	Yes	Yes	Yes	Yes (10:1) ratio	34, 145, 220
MqsRA	*E. coli*	Chromosome	2 operator sites	No	No	No	Yes	148, 223
ParDE	*E. coli*	Plasmid RK2	3 operator sites	No	No	Yes	Yes	224, 150 151
PezAT	*S. pneumoniae*	Chromosome	1 palindromic sequence	N/A	N/A	Yes	No	152
Phd-Doc	Bacteriophage P1	Genome	2 palindromic sequences	Yes	Yes	Yes	Yes	210, 227, 360
RelBE	*E. coli*	Chromosome	2 palindromes	Yes[3]	Modelled as so	Yes	Yes	156, 155, 228
VapBC	*Rickettsia felis*	Chromosome	2 palindromes	N/A	No	Yes	No	229
VapBC	*Salmonella enterica*	Chromosome	2 operator sites	N/A	No	Yes	Yes (10:1 ratio)	211
VapBC	*Shigella flexneri*	Chromosome	2 operator sites	N/A	No	N/A	N/A	159
VapBC	*Caulobacter crescentus*	Chromosome	2 operator sites	N/A	No	N/A	N/A	230
YefM-YoeB	*E. coli*	Chromosome	2 operators	Yes	N/A	Yes	No (At 8:1 ratio)	160, 64

Appendix Table 22. List of toxin-antitoxin systems within bacteria where conditional cooperativity has been investigated (August 2018). The location of the genes (Chromosome/Plasmid), number of operators within the regulatory promoter region, the presence or absence of an intrinsically disordered domain (IDR) within the antitoxin, whether the complex forms an alternating array of toxin:antitoxin T:AT modules and if toxins have been observed to act as co-repressors and de-repressors. [1] indicates the uncertainty of the specific location of the operator sites. [2] Operator sites that have not been experimentally confirmed. [3] Presence of a IDR at the toxin binding site not the DNA binding site.

Protein	$\bar{\upsilon}$ (g.ml^{-1})
HicA	0.75383
HicB	0.73840
$HicA_2$-$HicB_4$	0.74305

Appendix Table 23. Partial specific volume for each protein component analysed in both SVAUC and SEAUC. These values were calculated via SEDNERPT[380, 381] based on the molecular weight of the species and the buffer density.

10. References

1. Tenover F.C. Mechanisms of antimicrobial resistance in bacteria. *Am J Infect Control*. 2006;**34**:S3-S10.

2. WHO. Antimicrobial Resistance: Global report on Surveillance 2014. Accessed May 2016 Available from: http://apps.who.int/iris/bitstream/10665/112647/1/WHO_HSE_PED_AIP_2014.2_eng.pdf?ua=1.

3. WHO. Antimcrobial Resistance 2018. Accessed April 2018 Available from: http://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance.

4. WHO. Antimicrobial agents in clinical developments. An analysis of the antimicrobial clinical development pipeline, including tuberculosis 2017. Accessed April 2018 Available from: http://www.who.int/medicines/areas/rational_use/antibacterial_agents_clinical_development/en/.

5. Wistrand-Yuen E., Knopp M., Hjort K., Koskiniemi S., Berg O.G., Andersson D.I. Evolution of high-level resistance during low-level antibiotic exposure. *Nat Commun*. 2018;**9**:12.

6. Young D., Hussell T., Dougan G. Chronic bacterial infections: living with unwanted guests. *Nat Immunol*. 2002;**3**:1026-32.

7. Grant S.S., Hung D.T. Persistent bacterial infections, antibiotic tolerance, and the oxidative stress response. *Virulence*. 2013;**4**:273-83.

8. Costerton J.W., Stewart P.S., Greenberg E.P. Bacterial biofilms: A common cause of persistent infections. *Science*. 1999;**284**:1318-22.

9. Monack D.M., Mueller A., Falkow S. Persistent bacterial infections: The interface of the pathogen and the host immune system. *Nat Rev Microbiol*. 2004;**2**:747-65.

10. Gomez J.E., McKinney J.D. *M. tuberculosis* persistence, latency, and drug tolerance. *Tuberculosis*. 2004;**84**:29-44.

11. Lyczak J.B., Cannon C.L., Pier G.B. Lung infections associated with cystic fibrosis. *Clin Microbiol Rev*. 2002;**15**:194-222.

12. Parsonnet J., Friedman G.D., Vandersteen D.P., Chang Y., Vogelman J.H., Orentreich N., et al. *Helicobacter pylori* infection and the risk of gastric-carcinoma. *N Engl J Med*. 1991;**325**:1127-31.

13. Bigger J.W. Treatment of staphylococcal infections with penicillin - By intermittent sterilisation. *Lancet*. 1944;**2**:497-500.

14. Lewis K. Programmed death in bacteria. *Microbiol Mol Biol Rev*. 2000;**64**:503-14.

15. Kotte O., Volkmer B., Radzikowski J.L., Heinemann M. Phenotypic bistability in *Escherichia coli*'s central carbon metabolism. *Mol Syst Biol*. 2014;**10**:11.

16. Veening J.W., Stewart E.J., Berngruber T.W., Taddei F., Kuipers O.P., Hamoen L.W. Bet-hedging and epigenetic inheritance in bacterial cell development. *Proc Natl Acad Sci U S A*. 2008;**105**:4393-8.

17. Verstraeten N., Knapen W.J., Kint C.I., Liebens V., Van den Bergh B., Dewachter L., et al. Obg and Membrane Depolarization Are Part of a Microbial Bet-Hedging Strategy that Leads to Antibiotic Tolerance. *Mol Cell*. 2015;**59**:9-21.

18. Levin-Reisman I., Ronin, I., Gefen O., Braniss, I., Shoresh, N., Balaban N.Q. Antibiotic tolerance facilitates the evolution of antibiotic resistance. *Science*. 2017;**355**:826-830.

19. Björkman J., Nagaev I., Berg O.G., Hughes D., Andersson D.I. Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance. *Science*. 2000;**287**:1479-82.

20. Wilson S.G., Sanders C.C. Selection and chatacterization of strains of *Staphylococcus aureus* displaying unusual resistance to aminoglycosides. *Antimicrob Agents Chemother*. 1976;**10**:519-25.

21. von Eiff C., Heilmann C., Proctor R.A., Woltz C., Peters G., Götz F. A site-directed *Staphylococcus aureus* hemB mutant is a small-colony variant which persists intracellularly. *J Bacteriol*. 1997;**179**:4706-12.

22. Lewis K. Riddle of biofilm resistance. *Antimicrob Agents Chemother*. 2001;**45**:999-1007.

23. Brooun A., Liu S.H., Lewis K. A dose-response study of antibiotic resistance in *Pseudomonas aeruginosa* biofilms. *Antimicrob Agents Chemother*. 2000;**44**:640-6.

24. Ashby M.J., Neale J.E., Knott S.J., Critchley I.A. Effect of antibiotics on non-growing planktonic cells and biofilms of *Escherichia coli*. *J Antimicrob Chemother*. 1994;**33**:443-52.

25. Muli F.W., Struthers J.K. The growth of *Gardnerella vaginalis* and *Lactobacillus acidophilus* in Sorbarod biofilms. *J Med Microbiol*. 1998;**47**:401-5.

26. Spoering A.L., Lewis K. Biofilms and planktonic cells of *Pseudomonas aeruginosa* have similar resistance to killing by antimicrobials. *J Bacteriol*. 2001;**183**:6746-51.

27. Wolfson J.S., Hooper D.C., Shih D.J., McHugh G.L., Swartz M.N. Isolation and characterization of an *Escherichia coli* strain exhibiting partial tolerance to quinolones. *Antimicrob Agents Chemother*. 1989;**33**:705-9.

28. Gerdes K., Bech F.W., Jorgensen S.T., Lobnerolesen A., Rasmussen P.B., Atlung T., 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*. 1986;**5**:2023-9.

29. Hayes F. Toxins-antitoxins: Plasmid maintenance, programmed cell death, and cell cycle arrest. *Science*. 2003;**301**:1496-9.

30. Engelberg-Kulka H., Glaser G. Addiction modules and programmed cell death and antideath in bacterial cultures. *Annu Rev Microbiol*. 1999;**53**:43-70.

31. 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*. 1993;**233**:414-28.

32. Van Melderen L., Bernard P., Couturier N. Lon-dependent proteolysis of CcdA is the key control for activation of CcdB in plasmid-free segregant bacteria. *Mol Microbiol*. 1994;**11**:1151-7.

33. Tsuchimoto S., Ohtsubo H., Ohtsubo E. Two genes, pemK and pemI, responsible for stable maintenance of resistance plasmid R100. *J Bacteriol*. 1988;**170**:1461-6.

34. 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. *Proc Natl Acad Sci U S A*. 1996;**93**:6059-63.

35. Gotfredsen M., Gerdes K. The *Escherichia coli* relBE genes belong to a new toxin-antitoxin gene family. *Mol Microbiol*. 1998;**29**:1065-76.

36. 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*. 1994;**237**:35-51.

37. Jaffé, A., Ogura T., Hiraga S. Effects of the ccd function of the F plasmid on bacterial growth. *J Bacteriol*. 1985;**163**:841-9.

38. 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*. 1992;**174**:4205-11.

39. 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*. 1983;**155**:768-75.

40. Moyed H.S., Broderick S.H. Molecular cloning and expression of hipA, a gene of *Escherichia coli* K-12 that afects frequency of persistence after inhibition of murein synthesis. *J Bacteriol*. 1986;**166**:399-403.

41. Scherrer R., Moyed H.S. Conditional impairment of cell division and altered lethality in hipA mutants of *Escherichia coli* K-12. *J Bacteriol*. 1988;**170**:3321-6.

42. Black D.S., Kelly A.J., Mardis M.J., Moyed H.S. Structure and organization of hip, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *J Bacteriol*. 1991;**173**:5732-9.

43. Black D.S., Irwin B., Moyed H.S. Autoregulation of hip, an operon that affects lethality due to inhibition of peptidoglycan or DNA synthesis. *J Bacteriol*. 1994;**176**:4081-91.

44. Falla T.J., Chopra I. Stabilization of Rhizobium symbiosis plasmids. *Microbiology*. 1999;**145**:515-6.

45. Jensen R.B., Gerdes K. Programmed cell death in bacteria: proteic plasmid stabilization systems. *Mol Microbiol*. 1995;**17**:205-10.

46. Sat B., Hazan R., Fisher T., Khaner H., Glaser G., Engelberg-Kulka H. Programmed cell death in *Escherichia coli*: Some antibiotics can trigger mazEF lethality. *J Bacteriol*. 2001;**183**:2041-5.

47. Keren I., Shah D., Spoering A., Kaldalu N., Lewis K. Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *J Bacteriol*. 2004;**186**:8172-80.

48. Balaban N.Q., Merrin J., Chait R., Kowalik L., Leibler S. Bacterial persistence as a phenotypic switch. *Science*. 2004;**305**:1622-5.

49. Gefen O., Balaban N.Q. The importance of being persistent: heterogeneity of bacterial populations under antibiotic stress. *Fems Microbiol Rev*. 2009;**33**:704-17.

50. 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*. 2013;**195**:647-57.

51. Gerdes K., Christensen S.K., Lobner-Olesen A. Prokaryotic toxin-antitoxin stress response loci. *Nat Rev Microbiol*. 2005;**3**:371-82.

52. 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*. 2006;**188**:3826-36.

53. 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*. 2006;**188**:3494-7.

54. Correia F.F., D'Onofrio A., Rejtar T., Li L.Y., Karger B.L., Makarova K., et al. Kinase activity of overexpressed HipA is required for growth arrest and multidrug tolerance in *Escherichia coli*. *J Bacteriol*. 2006;**188**:8360-7.

55. Wu N., He L., Cui P., Wang W.J., Yuan Y.H., Liu S., et al. Ranking of persister genes in the same *Escherichia coli* genetic background demonstrates varying importance of individual persister genes in tolerance to different antibiotics. *Front Microbiol*. 2015;**6**:11.

56. Norton J.P., Mulvey M.A. Toxin-Antitoxin Systems Are Important for Niche-Specific Colonization and Stress Resistance of Uropathogenic *Escherichia coli*. *PLoS Pathog*. 2012;**8**:13.

57. Helaine S., Cheverton A.M., Watson K.G., Faure L.M., Matthews S.A., Holden D.W. Internalization of *Salmonella* by Macrophages Induces Formation of Nonreplicating Persisters. *Science*. 2014;**343**:204-8.

58. Van Acker H., Sass A., Dhondt I., Nelis H.J., Coenye T. Involvement of toxin-antitoxin modules in *Burkholderia cenocepacia* biofilm persistence. *Pathog Dis*. 2014;**71**:326-35.

59. Harrison J.J., Wade W.D., Akierman S., Vacchi-Suzzi C., Stremick C.A., Turner R.J., et al. The Chromosomal Toxin Gene yafQ Is a Determinant of Multidrug Tolerance for *Escherichia coli* Growing in a Biofilm. *Antimicrob Agents Chemother*. 2009;**53**:2253-8.

60. Donegan N.P., Thompson E.T., Fu Z.B., Cheung A.L. Proteolytic Regulation of Toxin-Antitoxin Systems by ClpPC in *Staphylococcus aureus*. *J Bacteriol*. 2010;**192**:1416-22.

61. Conlon B.P. Staphylococcus aureus chronic and relapsing infections: Evidence of a role for persister cells An investigation of persister cells, their formation and their role in *S. aureus* disease. *Bioessays*. 2014;**36**:991-6.

62. Chan W.T., Domenech M., Moreno-Cordóba I., Navarro-Martinez V., Nieto C., Moscoso M., et al. The *Streptococcus pneumoniae* yefM-yoeB and relBE Toxin-Antitoxin Operons Participate in Oxidative Stress and Biofilm Formation. *Toxins*. 2018;**10**.

63. Cheverton A.M., Gollan B., Przydacz M., Wong C.T., Mylona A., Hare S.A., et al. A *Salmonella* Toxin Promotes Persister Formation through Acetylation of tRNA. *Mol Cell*. 2016;**63**:86-96.

64. Kędzierska B., Hayes F. Emerging Roles of Toxin-Antitoxin Modules in Bacterial Pathogenesis. *Molecules*. 2016;**21**:25.

65. Song S., Wood T.K. Post-segregational Killing and Phage Inhibition Are Not Mediated by Cell Death Through Toxin/Antitoxin Systems. *Front Microbiol*. 2018;**9**:6.

66. Goormaghtigh F., Fraikin N., Putrinš M., Hallaert T., Hauryliuk V., Garcia-Pino A., et al. Reassessing the role of type II Toxin-Antitoxin systems in formation of Escherichia coli type II persister cells. *MBio*. 2018;**12**:e00640-18.

67. Harms A., Fino C., Sorensen M.A., Semsey S., Gerdes K. Prophages and Growth Dynamics Confound Experimental Results with Antibiotic-Tolerant Persister Cells. *Mbio*. 2017;**8**:18.

68. Maisonneuve E., Shakespeare L.J., Jorgensen M.G., Gerdes K. Bacterial persistence by RNA endonucleases. *Proc Natl Acad Sci U S A*. 2011;**108**:13206-11.

69. Maisonneuve E., Castro-Camargo M., Gerdes K. RETRACTION: (p)ppGpp Controls Bacterial Persistence by Stochastic Induction of Toxin-Antitoxin Activity (Retraction of Vol 154, Pg 1140, 2013). *Cell*. 2018;**172**:1135-.

70. Gupta A., Venkataraman B., Vasudevan M., Bankar K.G. Co-expression network analysis of toxin-antitoxin loci in *Mycobacterium tuberculosis* reveals key modulators of cellular stress. *Sci Rep*. 2017;**7**:14.

71. Christensen-Dalsgaard M., Jørgensen M.G., Gerdes K. Three new RelE-homologous mRNA interferases of *Escherichia coli* differentially induced by environmental stresses. *Mol Microbiol*. 2010;**75**:333-48.

72. Dörr T., Vulić M., Lewis K. Ciprofloxacin Causes Persister Formation by Inducing the TisB toxin in *Escherichia coli*. *PLoS Biol*. 2010;**8**:8.

73. Karimi S., Ghafourian S., Kalani M.T., Jalilian F.A., Hemati S., Sadeghifard N. Association Between Toxin-Antitoxin Systems and Biofilm Formation. *Jundishapur J Microbiol*. 2015;**8**:6.

74. Thisted T., Gerdes K. Mechanism of post-segregational killing by the hok/sok system of plasmid R1 Sok antisense RNA regulates hok gene exxpression indirectly through the overlapping mok gene. *J Mol Biol*. 1992;**223**:41-54.

75. Cashel M., Gentry D.R., Hernandez V.I., Winella D. The Stringent Response. 2 ed. Neidhardt FC, Curtiss IR, Ingraham JL, Lin CY, Low KB, Magasanik B, et al., editors. Washington DC: American Society for Microbiology Press; 1996.

76. Radman M. SOS repair hypothesis: phemenology of an inducible DNA repair which is accompanied by mutagenesis. *Basic Life Sciences*. 1975; **5A**:355-67.

77. Little J.W., Mount D.W. The SOS regulatory system of *Escherichia coli*. *Cell*. 1982;**29**:11-22.

78. 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*. 2003;**50**:1199-213.

79. Mosteller R.D., Kwan S.F. Isolation of relaxed-control mutants of *Escherichia coli* K-12 which are sensitive to glucose starvation. *Biochem Biophys Res Commun*. 1976;**69**:325-32.

80. Diderichsen B., Fiil N.P., Lavallé R. Genetics of the relB locus in *Escherichia coli*. *J Bacteriol*. 1977;**131**:30-3.

81. Kawano M., Aravind L., Storz G. An antisense RNA controls synthesis of an SOS-induced toxin evolved from an antitoxin. *Mol Microbiol*. 2007;**64**:738-54.

82. Dörr T., Lewis K., Vulić M. SOS Response Induces Persistence to Fluoroquinolones in *Escherichia coli*. *PLoS Genet*. 2009;**5**:9.

83. Ruangprasert A., Maehigashi T., Miles S.J., Giridharan N., Liu J.X., Dunham C.M. Mechanisms of Toxin Inhibition and Transcriptional Repression by *Escherichia coli* DinJ-YafQ. *J Biol Chem*. 2014;**289**:20559-69.

84. Bernard P., Couturier M. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J Mol Biol*. 1992;**226**:735-45.

85. Dao-Thi M.H., Van Melderen L., De Genst E., Afif H., Buts L., Wyns L., et al. Molecular basis of gyrase poisoning by the addiction toxin CcdB. *J Mol Biol*. 2005;**348**:1091-102.

86. Yuan J., Yamaichi Y., Waldor M.K. The Three Vibrio cholerae Chromosome II-Encoded ParE Toxins Degrade Chromosome I following Loss of Chromosome II. *J Bacteriol*. 2011;**193**:611-9.

87. Gerdes K., Larsen J.E.L., Molin S. Stable inheritance of plasmid R1 requires two different loci. *J Bacteriol*. 1985;**161**:292-8.

88. Fozo E.M., Hemm M.R., Storz G. Small Toxic Proteins and the Antisense RNAs That Repress Them. *Microbiol Mol Biol Rev*. 2008;**72**:579-89.

89. Fozo E.M., Makarova K.S., Shabalina S.A., Yutin N., Koonin E.V., Storz G. Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. *Nucleic Acids Res*. 2010;**38**:3743-59.

90. Han K., Kim K.S., Bak G., Park H., Lee Y. Recognition and discrimination of target mRNAs by Sib RNAs, a cis-encoded sRNA family. *Nucleic Acids Res*. 2010;**38**:5851-66.

91. 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*. 2002;**45**:333-49.

92. Silvaggi J.M., Perkins J.B., Losick R. Small untranslated RNA antitoxin in *Bacillus subtilis*. *J Bacteriol*. 2005;**187**:6641-50.

93. Jahn N., Preis H., Wiedemann C., Brantl S. BsrG/SR4 from *Bacillus subtilis*- the first temperature-dependent type I toxin-antitoxin system. *Mol Microbiol*. 2012;**83**:579-98.

94. Brantl S., Jahn N. sRNAs in bacterial type I and type III toxin-antitoxin systems. *Fems Microbiol Rev*. 2015;**39**:413-27.

95. Meißner C., Jahn N., Brantl S. In Vitro Characterization of the Type I Toxin-Antitoxin System bsrE/SR5 from *Bacillus subtilis*. *J Biol Chem*. 2016;**291**:560-71.

96. Kodama T., Matsubayashi T., Yanagihara T., Komoto H., Ara K., Ozaki K., et al. A Novel Small Protein of *Bacillus subtilis* Involved in Spore Germination and Spore Coat Assembly. *Biosci Biotechnol Biochem*. 2011;**75**:1119-28.

97. 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*. 1990;**4**:1807-18.

98. Steinbrecher T., Prock S., Reichert J., Wadhwani P., Zimpfer B., Bürck J., et al. Peptide-Lipid Interactions of the Stress-Response Peptide TisB That Induces Bacterial Persistence. *Biophys J*. 2012;**103**:1460-9.

99. Yamaguchi Y., Inouye M. Regulation of growth and death in *Escherichia coli* by toxin-antitoxin systems. *Nat Rev Microbiol*. 2011;**9**:779-90.

100. Samson J.E., Bélanger M., Moineau S. Effect of the Abortive Infection Mechanism and Type III Toxin/Antitoxin System AbiQ on the Lytic Cycle of *Lactococcus lactis* Phages. *J Bacteriol*. 2013;**195**:3947-56.

101. Bélanger M., Moineau S. Mutational Analysis of the Antitoxin in the Lactococcal Type III Toxin-Antitoxin System AbiQ. *Appl Environ Microbiol*. 2015;**81**:3848-55.

102. Short F.L., Pei X.Y., Blower T.R., Ong S.L., Fineran P.C., Luisi B.F., 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*. 2013;**110**:E241-E9.

103. 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*. 2018;**8**:10.

104. Blower T.R., Short F.L., Rao F., Mizuguchi K., Pei X.Y., Fineran P.C., et al. Identification and classification of bacterial Type III toxin-antitoxin systems encoded in chromosomal and plasmid genomes. *Nucleic Acids Res*. 2012;**40**:6158-73.

105. Wen J., Fozo E.M. sRNA Antitoxins: More than One Way to Repress a Toxin. *Toxins*. 2014;**6**:2310-35.

106. Rao F., Short F.L., Voss J.E., Blower T.R., Orme A.L., Whittaker T.E., 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*. 2015;**43**:9529-40.

107. 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*. 2012;**84**:979-89.

108. Wen Z.L., Wang P.X., Sun C.L., Guo Y.X., Wang X.X. Interaction of Type IV Toxin/Antitoxin Systems in Cryptic Prophages of *Escherichia coli* K-12. *Toxins*. 2017;**9**:14.

109. Wang X.X., Lord D.M., Cheng H.Y., Osbourne D.O., Hong S.H., Sanchez-Torres V., et al. A new type V toxin-antitoxin system where mRNA for toxin GhoT is cleaved by antitoxin GhoS. *Nat Chem Biol*. 2012;**8**:855-61.

110. Kim J.S., Schantz A.B., Song S., Kumar M., Wood T.K. GhoT of the GhoT/GhoS toxin/antitoxin system damages lipid membranes by forming transient pores. *Biochem Biophys Res Commun*. 2018;**497**:467-72.

111. Aakre C.D., Phung T.N., Huang D., Laub M.T. A Bacterial Toxin Inhibits DNA Replication Elongation through a Direct Interaction with the ß Sliding Clamp. *Mol Cell*. 2013;**52**:617-28.

112. Page R., Peti W. Toxin-antitoxin systems in bacterial growth arrest and persistence. *Nat Chem Biol*. 2016;**12**:208-14.

113. Harms A., Brodersen D.E., Mitarai N., Gerdes K. Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. *Mol Cell*. 2018;**70**:768-84.

114. Castro-Roa D., Garcia-Pino A., De Gieter S., van Nuland N.A.J., Loris R., Zenkin N. The Fic protein Doc uses an inverted substrate to phosphorylate and inactivate EF-Tu. *Nat Chem Biol*. 2013;**9**:811-7.

115. Jurėnas D., Chatterjee S., Konijnenberg A., Sobott F., Droogmans L., Garcia-Pino A., et al. AtaT blocks translation initiation by N-acetylation of the initiator tRNA(fMet). *Nat Chem Biol*. 2017;**13**:640-6.

116. Qian H.L., Yao Q.Q., Tai C., Deng Z.X., Gan J.H., Ou H.Y. Identification and characterization of acetyltransferase-type toxin-antitoxin locus in *Klebsiella pneumoniae*. *Mol Microbiol*. 2018;**108**:336-49.

117. Yeo C.C. GNAT toxins of bacterial toxin-antitoxin systems: acetylation of charged tRNAs to inhibit translation. *Mol Microbiol*. 2018;**108**:331-5.

118. Gerdes K., Maisonneuve E. Bacterial Persistence and Toxin-Antitoxin Loci. *Annu Rev Microbio*l.2012;**66**:103-23.

119. Schuster C.F., Bertram R. Toxin-antitoxin systems are ubiquitous and versatile modulators of prokaryotic cell fate. *FEMS Microbiol Lett*. 2013;**340**:73-85.

120. Muthuramalingam M., White J.C., Bourne C.R. Toxin-Antitoxin Modules Are Pliable Switches Activated by Multiple Protease Pathways. *Toxins*. 2016;**8**:16.

121. Boss L., Labudda L., Węgrzyn G., Hayes F., Kędzierska B. The Axe-Txe Complex of *Enterococcus faecium* Presents a Multilayered Mode of Toxin-Antitoxin Gene Expression Regulation. *PLoS One*. 2013;**8**:13.

122. Grady R., Hayes F. Axe-Txe, a broad-spectrum proteic toxin-antitoxin system specified by a multidrug-resistant, clinical isolate of *Enterococcus faecium*. *Mol Microbiol*. 2003;**47**:1419-32.

123. Loris R., Dao-Thi M.H., Bahassi E.M., Van Melderen L., Poortmans F., Liddington R., et al. Crystal structure of CcdB, a topoisomerase poison from *E coli*. *J Mol Biol*. 1999;**285**:1667-77.

124. Madl T., Van Melderen L., Mine N., Respondek M., Oberer M., Keller W., et al. Structural basis for nucleic acid and toxin recognition of the bacterial antitoxin CcdA. *J Mol Biol*. 2006;**364**:170-85.

125. De Jonge N., Garcia-Pino A., Buts L., Haesaerts S., Charlier D., Zangger K., et al. Rejuvenation of CcdB-Poisoned Gyrase by an Intrinsically Disordered Protein Domain. *Mol Cell*. 2009;**35**:154-63.

126. Vandervelde A., Drobnak I., Hadži S., Sterckx Y.G.J., Welte T., De Greve H., et al. Molecular mechanism governing ratio-dependent transcription regulation in the ccdAB operon. *Nucleic Acids Res*. 2017;**45**:2937-50.

127. Liang Y.J., Gao Z.Q., Wang F., Zhang Y.L., Dong Y.H., Liu Q.S. Structural and Functional Characterization of *Escherichia coli* Toxin-Antitoxin Complex DinJ-YafQ. *J Biol Chem*. 2014;**289**:21191-202.

128. Marimon O., Teixeira J.M.C., Cordeiro T.N., Soo V.W.C., Wood T.L., Mayzel M., et al. An oxygen-sensitive toxin-antitoxin system. *Nat Commun*. 2016;**7**:10.

129. Yee A., Chang X.Q., Pineda-Lucena A., Wu B., Semesi A., Le B., et al. An NMR approach to structural proteomics. *Proc Natl Acad Sci U S A*. 2002;**99**:1825-30.

130. 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*. 2008;**3**:15.

131. Butt A., Higman V.A., Williams C., Crump M.P., Hemsley C.M., Harmer N., et al. The HicA toxin from *Burkholderia pseudomallei* has a role in persister cell formation. *Biochem J*. 2014;**459**:333-44.

132. Bibi-Triki S., Li de la Sierra-Gallay I., Lazar N., Leroy A., Van Tilbeurgh H., Sebbane F., et al. Functional and Structural Analysis of HicA3-HicB3, a Novel Toxin-Antitoxin System of *Yersinia pestis*. *J Bacteriol*. 2014;**196**:3712-23.

133. Arbing M.A., Handelman S.K., Kuzin A.P., Verdon G., Wang C., Su M., et al. Crystal Structures of Phd-Doc, HigA, and YeeU Establish Multiple Evolutionary Links between Microbial Growth-Regulating Toxin-Antitoxin Systems. *Structure*. 2010;**18**:996-1010.

134. Yang J.S., Zhou K., Liu P., Dong Y.H., Gao Z.Q., Zhan J.J., et al. Structural insight into the *E. coli* HigBA. *Biochem Biophys Res Commun*. 2016;**478**:1521-7.

135. Schumacher M.A., Piro K.M., Xu W., Hansen S., Lewis K., Brennan R.G. Molecular Mechanisms of HipA-Mediated Multidrug Tolerance and Its Neutralization by HipB. *Science*. 2009;**323**:396-401.

136. Schumacher M.A., Min J.K., Link T.M., Guan Z.Q., Xu W.J., Ahn Y.H., et al. Role of Unusual P Loop Ejection and Autophosphorylation in HipA-Mediated Persistence and Multidrug Tolerance. *Cell Reports*. 2012;**2**:518-25.

137. Schumacher M.A., Balani P., Min J.K., Chinnam N.B., Hansen S., Vulić M., et al. HipBA-promoter structures reveal the basis of heritable multidrug tolerance. *Nature*. 2015;**524**:59-U108.

138. Evdokimov A., Voznesensky I., Fennell K., Anderson M., Smith J.F., Fisher D.A. New kinase regulation mechanism found in HipBA: a bacterial persistence switch. *Acta Crystallogr Sect D-Biol Crystallogr*. 2009;**65**:875-9.

139. Stanger F.V., Harms A., Dehio C., Schirmer T. Crystal Structure of the *Escherichia coli* Fic Toxin-Like Protein in Complex with Its Cognate Antitoxin. *PLoS One*. 2016;**11**:21.

140. Engel P., Goepfert A., Stanger F.V., Harms A., Schmidt A., Schirmer T., et al. Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins. *Nature*. 2012;**482**:107-10.

141. Harms A., Stanger F.V., Scheu P.D., de Jong I.G., Goepfert A., Glatter T., et al. Adenylylation of Gyrase and Topo IV by FicT Toxins Disrupts Bacterial DNA Topology. *Cell Reports*. 2015;**12**:1497-507.

142. Mattison K., Wilbur J.S., So M., Brennan R.G. Structure of FitAB from *Neisseria gonorrhoeae* bound to DNA reveals a tetramer of toxin-antitoxin heterodimers containing pin domains and ribbon-helix-helix motifs. *J Biol Chem*. 2006;**281**:37942-51.

143. McVicker G., Tang C.M. Deletion of toxin-antitoxin systems in the evolution of *Shigella sonnei* as a host-adapted pathogen. *Nat Microbiol*. 2017;**2**:8.

144. Hargreaves D., Santos-Sierra S., Giraldo R., Sabariegos-Jareño R., de la Cueva-Méndez G., Boelens R., et al. Structural and functional analysis of the Kid toxin protein from *E coli* plasmid R1. *Structure*. 2002;**10**:1425-33.

145. Kamada K., Hanaoka F., Burley S.K. Crystal structure of the MazE/MazF complex: Molecular bases of antidote-toxin recognition. *Mol Cell*. 2003;**11**:875-84.

146. Simanshu D.K., Yamaguchi Y., Park J.H., Inouye M., Patel D.J. Structural Basis of mRNA Recognition and Cleavage by Toxin MazF and Its Regulation by Antitoxin MazE in *Bacillus subtilis*. *Mol Cell*. 2013;**52**:447-58.

147. Brown B.L., Grigoriu S., Kim Y., Arruda J.M., Davenport A., Wood T.K., 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*. 2009;**5**:15.

148. Brown B.L., Wood T.K., Peti W., Page R. Structure of the *Escherichia coli* Antitoxin MqsA (YgiT/b3021) Bound to Its Gene Promoter Reveals Extensive Domain Rearrangements and the Specificity of Transcriptional Regulation. *J Biol Chem*. 2011;**286**:2285-96.

149. Papadopoulos E., Collet J.F., Vukojević V., Billeter M., Holmgren A., Gräslund A., et al. Solution structure and biophysical properties of MqsA, a Zn-containing antitoxin from *Escherichia coli*. *BBA-Proteins Proteomics*. 2012;**1824**:1401-8.

150. 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*. 2007;**16**:1676-88.

151. Dalton K.M., Crosson S. A Conserved Mode of Protein Recognition and Binding in a ParD-ParE Toxin-Antitoxin Complex. *Biochemistry*. 2010;**49**:2205-15.

152. Khoo S.K., Loll B., Chan W.T., Shoeman R.L., Ngoo L., Yeo C.C., et al. Molecular and structural characterization of the PezAT chromosomal toxin-antitoxin system of the human pathogen *Streptococcus pneumoniae*. *J Biol Chem*. 2007;**282**:19606-18.

153. Garcia-Pino A., Christensen-Dalsgaard M., Wyns L., Yarmolinsky M., Magnuson R.D., Gerdes K., et al. Doc of Prophage P1 Is Inhibited by Its Antitoxin Partner Phd through Fold Complementation. *J Biol Chem*. 2008;**283**:30821-7.

154. VanDrisse C.M., Parks A.R., Escalante-Semerena J.C. A Toxin Involved in *Salmonella* Persistence Regulates Its Activity by Acetylating Its Cognate Antitoxin, a Modification Reversed by CobB Sirtuin Deacetylase. *Mbio*. 2017;**8**:14.

155. Bøggild A., Sofos N., Andersen K.R., Feddersen A., Easter A.D., Passmore L.A., et al. The Crystal Structure of the Intact *E. coli* RelBE Toxin-Antitoxin Complex Provides the Structural Basis for Conditional Cooperativity. *Structure*. 2012;**20**:1641-8.

156. Li G.Y., Zhang Y., Inouye M., Ikura M. Structural mechanism of transcriptional autorepression of the *Escherichia coli* RelB/RelE antitoxin/toxin module. *J Mol Biol*. 2008;**380**:107-19.

157. Takagi H., Kakuta Y., Okada T., Yao M., Tanaka I., Kimura M. Crystal structure of archaeal toxin-antitoxin RelE-RelB complex with implications for toxin activity and antitoxin effects. *Nat Struct Mol Biol*. 2005;**12**:327-31.

158. Miallau L., Faller M., Chiang J., Arbing M., Guo F., Cascio D., et al. Structure and Proposed Activity of a Member of the VapBC Family of Toxin-Antitoxin Systems VapBC-5 from *Mycobacterium tuberculosis*. *J Biol Chem*. 2009;**284**:276-83.

159. Dienemann C., Bøggild A., Winther K.S., Gerdes K., Brodersen D.E. Crystal Structure of the VapBC Toxin-Antitoxin Complex from *Shigella flexneri* Reveals a Hetero-Octameric DNA-Binding Assembly. *J Mol Biol*. 2011;**414**:713-22.

160. Kamada K., Hanaoka F. Conformational change in the catalytic site of the ribonuclease YoeB toxin by YefM antitoxin. *Mol Cell*. 2005;**19**:497-509.

161. Meinhart A., Alonso J.C., Sträter N., Saenger W. Crystal structure of the plasmid maintenance system epsilon/zeta: Functional mechanism of toxin zeta and inactivation by epsilon(2)zeta(2) complex formation. *Proc Natl Acad Sci U S A*. 2003;**100**:1661-6.

162. 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*. 2011;**9**:12.

163. Prysak M.H., Mozdzierz C.J., Cook A.M., Zhu L., Zhang Y.L., Inouye M., 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*. 2009;**71**:1071-87.

164. Hurley J.M., Woychik N.A. Bacterial Toxin HigB Associates with Ribosomes and Mediates Translation-dependent mRNA Cleavage at A-rich Sites. *J Biol Chem*. 2009;**284**:18605-13.

165. Schureck M.A., Dunkle J.A., Maehigashi T., Miles S.J., Dunham C.M. Defining the mRNA recognition signature of a bacterial toxin protein. *Proc Natl Acad Sci U S A*. 2015;**112**:13862-7.

166. Schureck M.A., Repack A., Miles S.J., Marquez J., Dunham C.M. Mechanism of endonuclease cleavage by the HigB toxin. *Nucleic Acids Res*. 2016;**44**:7944-53.

167. Zhang Y.L., Zhang J.J., Hara H., Kato I., Inouye M. Insights into the mRNA cleavage mechanism by MazF, an mRNA interferase. *J Biol Chem*. 2005;**280**:3143-50.

168. Pedersen K., Zavialov A.V., Pavlov M.Y., Elf J., Gerdes K., Ehrenberg M. The bacterial toxin RelE displays codon-specific cleavage of rnRNAs in the ribosomal A site. *Cell*. 2003;**112**:131-40.

169. Hurley J.M., Cruz J.W., Ouyang M., Woychik N.A. Bacterial Toxin RelE Mediates Frequent Codon-independent mRNA Cleavage from the 5 ' End of Coding Regions in Vivo. *J Biol Chem*. 2011;**286**:14770-8.

170. Dunican B.F., Hiller D.A., Strobel S.A. Transition State Charge Stabilization and Acid Base Catalysis of mRNA Cleavage by the Endoribonuclease ReIE. *Biochemistry*. 2015;**54**:7048-57.

171. Hwang J.Y., Buskirk A.R. A ribosome profiling study of mRNA cleavage by the endonuclease RelE. *Nucleic Acids Res*. 2017;**45**:327-36.

172. 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*. 2009;**284**:28746-53.

173. Zhang J.J., Zhang Y.L., Zhu L., Suzuki M., Inouye M. Interference of mRNA function by sequence-specific endoribonuclease PemK. *J Biol Chem*. 2004;**279**:20678-84.

174. Chowdhury N., Kwan B.W., McGibbon L.C., Babitzke P., Wood T.K. Toxin MqsR cleaves single-stranded mRNA with various 5' ends. *MicrobiologyOpen*. 2016;**5**:370-7.

175. Zhu L., Inoue K., Yoshizumi S., Kobayashi H., Zhang Y.L., Ouyang M., et al. Staphylococcus aureus MazF Specifically Cleaves a Pentad Sequence, UACAU, Which Is Unusually Abundant in the mRNA for Pathogenic Adhesive Factor SraP. *J Bacteriol*. 2009;**191**:3248-55.

176. Pellegrini O., Mathy N., Gogos A., Shapiro L., Condon C. The *Bacillus subtilis* ydcDE operon encodes an endoribonuclease of the MazF/PemK family and its inhibitor. *Mol Microbiol*. 2005;**56**:1139-48.

177. Schifano J.M., Cruz J.W., Vvedenskaya I.O., Edifor R., Ouyang M., Husson R.N., et al. tRNA is a new target for cleavage by a MazF toxin. *Nucleic Acids Res*. 2016;**44**:1256-70.

178. Winther K., Tree J.J., Tollervey D., Gerdes K. VapCs of *Mycobacterium tuberculosis* cleave RNAs essential for translationl. *Nucleic Acids Res*. 2016;**44**:9860-71.

179. Winther K.S., Gerdes K. Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proc Natl Acad Sci U S A*. 2011;**108**:7403-7.

180. Lopes A.P.Y., Lopes L.M., Fraga T.R., Chura-Chambi R.M., Sanson A.L., Cheng E., et al. VapC from the Leptospiral VapBC Toxin-Antitoxin Module Displays Ribonuclease Activity on the Initiator tRNA. *PLoS One*. 2014;**9**:11.

181. Arcus V.L., McKenzie J.L., Robson J., Cook G.M. The PIN-domain ribonucleases and the prokaryotic VapBC toxin-antitoxin array. *Protein Eng Des Sel*. 2011;**24**:33-40.

182. McKenzie J.L., Robson J., Berney M., Smith T.C., Ruthe A., Gardner P.P., et al. A VapBC Toxin-Antitoxin Module Is a Posttranscriptional Regulator of Metabolic Flux in Mycobacteria. *J Bacteriol*. 2012;**194**:2189-204.

183. Ruiz-Echevarría M.J., G. G.-G., Sabraigeos-Jareño R., Díaz-Oreas R. Kid, a small protein of the parD stability system of plasmid R1, is an inhibitor of DNA replication acting at the initiation of DNA synthesis. *J Mol Biol*. 1995;**247**:568-77.

184. Potrykus K., Santos S., Lemonnier M., Díaz-Orejas R., Wegrzyn G. Differential effects of Kid toxin on two modes of replication of lambdoid plasmids suggest that this toxin acts before, but not after, the assembly of the replication complex. *Microbiology*. 2002;**148**:2489-95.

185. Muñoz-Gómez A.J., Lemonnier M., Santos-Sierra S., Berzal-Herranz A., Díaz-Orejas R. RNase/anti-RNase activities of the bacterial parD toxin-antitoxin system. *J Bacteriol*. 2005;**187**:3151-7.

186. Halvorsen E.M., Williams J.J., Bhimani A.J., Billings E.A., Hergenrother P.J. Txe, an endoribonuclease of the enterococcal Axe-Txe toxin-antitoxin system, cleaves mRNA and inhibits protein synthesis. *Microbiology*. 2011;**157**:387-97.

187. 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*. 2009;**191**:1191-9.

188. Critchlow S.E., Odea M.H., Howells A.J., Couturier M., Gellert M., Maxwell A. The interaction of the F plasmid killer protein, CcdB, with DNA gyrase: Induction of DNA cleavage and blocking of transcription. *J Mol Biol*. 1997;**273**:826-39.

189. Smith A.B., Maxwell A. A strand-passage conformation of DNA gyrase is required to allow the bacterial toxin, CcdB, to access its binding site. *Nucleic Acids Res*. 2006;**34**:4667-76.

190. 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*. 2002;**44**:971-9.

191. Yuan J., Sterckx Y., Mitchenall L.A., Maxwell A., Loris R., Waldor M.K. Vibrio cholerae ParE2 Poisons DNA Gyrase via a Mechanism Distinct from Other Gyrase Inhibitors. *J Biol Chem*. 2010;**285**:40397-408.

192. Collin F., Karkare S., Maxwell A. Exploiting bacterial DNA gyrase as a drug target: current state and perspectives. *Appl Microbiol Biotechnol*. 2011;**92**:479-97.

193. Cruz J.W., Rothenbacher F.P., Maehigashi T., Lane W.S., Dunham C.M., Woychik N.A. Doc Toxin Is a Kinase That Inactivates Elongation Factor Tu. *J Biol Chem*. 2014;**289**:7788-98.

194. Germain E., Castro-Roa D., Zenkin N., Gerdes K. Molecular Mechanism of Bacterial Persistence by HipA. *Mol Cell*. 2013;**52**:248-54.

195. Alberts B., Johnson A., Lewis J., Raff M., Roberts K., Walter P. Molecular Biology of The Cell Fifth Edition. New York: Garland Science; 2008.

196. Pribnow D. Bacteriophage T7 early promoters: nucleotide sequences of two RNA polymerase binding sites. *J Mol Biol*. 1975;**99**:419-43.

197. Harley C.B., Reynolds R.P. Analysis of *Escherichia coli* promoter sequences. *Nucleic Acids Res*. 1987;**15**:2343-61.

198. Shine J., Dalgarno L. Identical 3'-terminal octanucleotide sequence in 18S ribosomal ribonucleic acid from different eukaryotes. A proposed role for this sequence in the recognition of terminator codons. *Biochem J*. 1974;**141**:609-15.

199. Li G.W., Burkhardt D., Gross C., Weissman J.S. Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. *Cell*. 2014;**157**:624-35.

200. 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*. 2013;**9**:17.

201. Cataudella I., Sneppen K., Gerdes K., Mitarai N. Conditional Cooperativity of Toxin - Antitoxin Regulation Can Mediate Bistability between Growth and Dormancy. *PLoS Comput Biol*. 2013;**9**:9.

202. Cataudella I., Trusina A., Sneppen K., Gerdes K., Mitarai N. Conditional cooperativity in toxin-antitoxin regulation prevents random toxin activation and promotes fast translational recovery. *Nucleic Acids Res*. 2012;**40**:6424-34.

203. Nakanishi H., Pedersen M., Alsing A.K., Sneppen K. Modeling of the Genetic Switch of Bacteriophage TP901-1: A Heteromer of CI and MOR Ensures Robust Bistability. *J Mol Biol*. 2009;**394**:15-28.

204. Pedersen M., Hammer K. The Role of MOR and the CI Operator Sites on the Genetic Switch of the Temperate Bacteriophage TP901-1. *J Mol Biol*. 2008;**384**:577-89.

205. Chen D., Arkin A.P. Sequestration-based bistability enables tuning of the switching boundaries and design of a latch. *Mol Syst Biol*. 2012;**8**:7.

206. Paget M.S. Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. *Biomolecules*. 2015;**5**:1245-65.

207. Hughes K.T., Mathee K. The anti-sigma factors. *Annu Rev Microbiol*. 1998;**52**:231-86.

208. Trevino-Quintañilla L.G., Freyre-González J.A., Martínez-Flores I. Anti-Sigma Factors in *E coli*: Common Regulatory Mechanisms Controlling Sigma Factors Availability. *Curr Genomics*. 2013;**14**:378-87.

209. Ades S.E., Connolly L.E., Alba B.M., Gross C.A. The *Escherichia coli* sigma(E)-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-sigma factor. *Genes Dev*. 1999;**13**:2449-61.

210. Garcia-Pino A., De Gieter S., Talavera A., De Greve H., Efremov R.G., Loris R. An intrinsically disordered entropic switch determines allostery in Phd-Doc regulation. *Nat Chem Biol*. 2016;**12**:490-6.

211. Winther K.S., Gerdes K. Regulation of Enteric vapBC Transcription: Induction by VapC Toxin Dimer-Breaking. *Nucleic Acids Res*. 2012;**40**:4347-57.

212. Afif H., Allali N., Couturier M., Van Melderen L. The ratio between CcdA and CcdB modulates the transcriptional repression of the ccd poison-antidote system. *Mol Microbiol*. 2001;**41**:73-82.

213. Armalyté J., Jurėnaitė M., Beinoravičiūtė G., Teišerskas J., Sužiedėienė E. Characterization of *Escherichia coli* dinJ-yafQ Toxin-Antitoxin System Using Insights from Mutagenesis Data. *J Bacteriol*. 2012;**194**:1523-32.

214. Turnbull K.J., Gerdes K. HicA toxin of *Escherichia coli* derepresses hicAB transcription to selectively produce HicB antitoxin. *Mol Microbiol*. 2017;**104**:781-92.

215. Schureck M.A., Maehigashi T., Miles S.J., Marquez J., Cho S.E., Erdman R., et al. Structure of the *Proteus vulgaris* HigB-(HigA)(2)-HigB Toxin-Antitoxin Complex. *J Biol Chem*. 2014;**289**:1060-70.

216. Armalyté J., Jurėnas D., Krasauskas R., Čepauskas A., Sužiedėlienė E. The higBA Toxin-Antitoxin Module From the Opportunistic Pathogen *Acinetobacter baumannii* - Regulation, Activity, and Evolution. *Front Microbiol*. 2018;**9**:13.

217. Wen Y.R., Behiels E., Felix J., Elegheert J., Vergauwen B., Devreese B., et al. The bacterial antitoxin HipB establishes a ternary complex with operator DNA and phosphorylated toxin HipA to regulate bacterial persistence. *Nucleic Acids Res*. 2014;**42**:10134-47.

218. Wilbur J.S., Chivers P.T., Mattison K., Potter L., Brennan R.G., So M. *Neisseria gonorrhoeae* FitA interacts with FitB to bind DNA through its ribbon-helix-helix motif. *Biochemistry*. 2005;**44**:12515-24.

219. Monti M.C., Hernandez-Arriaga A.M., Kamphuis M.B., Lopez-Villarejo J., Heck A.J.R., Boelens R., 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*. 2007;**35**:1737-49.

220. Zorzini V., Buts L., Sleutel M., Garcia-Pino A., Talavera A., Haesaerts S., et al. Structural and biophysical characterization of *Staphylococcus aureus* SaMazF shows conservation of functional dynamics. *Nucleic Acids Res*. 2014;**42**:6709-25.

221. Marianovsky I., Aizenman E., Engelberg-Kulka H., Glaser G. The regulation of the *Escherichia coli* mazEF promoter involves an unusual alternating palindrome. *J Biol Chem*. 2001;**276**:5975-84.

222. Tripathi A., Dewan P.C., Siddique S.A., Varadarajan R. MazF-induced Growth Inhibition and Persister Generation in *Escherichia coli*. *J Biol Chem*. 2014;**289**:4191-205.

223. Brown B.L., Lord D.M., Grigoriu 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*. 2013;**288**:1286-94.

224. Johnson E.P., Strom A.R., Helinski D.R. Plasmid RK2 toxin protein ParE: Purification and interaction with the ParD antitoxin protein. *J Bacteriol*. 1996;**178**:1420-9.

225. Davis T.L., Helinski D.R., Roberts R.C. Transcription and autoregulation of the stabilizing functions of broad-host-range plasmid RK2 in *Escherichia coli*, *Agrobacterium tumefaciens* and *Pseudomonas aeruginosa*. *Mol Microbiol*. 1992;**6**:1981-94.

226. Gazit E., Sauer R.T. Stability and DNA binding of the Phd protein of the phage P1 plasmid addiction system. *J Biol Chem*. 1999;**274**:2652-7.

227. Garcia-Pino A., Balasubramanian S., Wyns L., Gazit E., De Greve H., Magnuson R.D., et al. Allostery and Intrinsic Disorder Mediate Transcription Regulation by Conditional Cooperativity. *Cell*. 2010;**142**:101-11.

228. 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*. 2009;**394**:183-96.

229. Maté M.J., Vincentelli R., Foos N., Raoult D., Cambillau C., Ortiz-Lombardía M. Crystal structure of the DNA-bound VapBC2 antitoxin/toxin pair from *Rickettsia felis*. *Nucleic Acids Res*. 2012;**40**:3245-58.

230. Bendtsen K.L., Brodersen D.E. Higher-order structure in bacterial VapBC toxin-antitoxin complexes. *Subcell Biochem*. 2017:381-412.

231. Kędzierska B., Lian L.Y., Hayes F. Toxin-antitoxin regulation: bimodal interaction of YefM-YoeB with paired DNA palindromes exerts transcriptional autorepression. *Nucleic Acids Res*. 2007;**35**:325-39.

232. Sullivan D.M., Bobay B.G., Kojetin D.J., Thompson R.J., Rance M., Strauch M.A., et al. Insights into the Nature of DNA Binding of AbrB-like Transcription Factors. *Structure*. 2008;**16**:1702-13.

233. Aravind L., Anantharaman V., Balaji S., Babu M.M., Iyer L.M. The many faces of the helix-turn-helix domain: Transcription regulation and beyond. *Fems Microbiol Rev*. 2005;**29**:231-62.

234. Schreiter E.R., Drennan C.L. Ribbon-helix-helix transcription factors: variations on a theme. *Nat Rev Microbiol*. 2007;**5**:710-20.

235. Lin C.Y., Awano N., Masuda H., Park J.H., Inouye M. Transcriptional Repressor HipB Regulates the Multiple Promoters in *Escherichia coli*. *J Mol Microbiol Biotechnol*. 2013;**23**:440-7.

236. Roberts R.C., Spangler C., Helinski D.R. Characteristics and signficance of DNA-binding activity of plasmid stabilization protein ParD from the broad host range plasmid RK2. *J Biol Chem*. 1993;**268**:27109-17.

237. Overgaard M., Borch J., Jørgensen M.G., Gerdes K. Messenger RNA interferase RelE controls relBE transcription by conditional cooperativity. *Mol Microbiol*. 2008;**69**:841-57.

238. Inglis T.J.J., Sagripanti J.L. Environmental factors that affect the survival and persistence of *Burkholderia pseudomallei*. *Appl Environ Microbiol*. 2006;**72**:6865-75.

239. Lewis E.R.G., Torres A.G. The art of persistence-the secrets to *Burkholderia* chronic infections. *Pathog Dis*. 2016;**74**:11.

240. Whitmore A. On the bacteriology of an infective disease occurring in Rangoon. *Br Med J*. 1912;**1912**:1306-8.

241. Whitmore A. An Account of a Ganders-like Disease ocurring in Rangoon. *J Hyg*. 1913;**13**:1-34.

242. Thomas A.D., Forbesfaulkner J., Parker M. Isolation of *pseudomonas-pseudomallei* from clay lasers at defined depths. *Am J Epidemiol*. 1979;**110**:515-21.

243. Thomas A.D., Forbesfaulkner J.C. Persistence of *Pseudomonas pseudomallei* in soil. *Aust Vet J*. 1981;**57**:535-6.

244. Thomas A.D., Norton J.H., Forbesfaulkner J.C., Woodland G. Mellioidosis in an intensive piggery. *Aust Vet J*. 1981;**57**:144-5.

245. Dance D.A.B. Melioidosis- The tip of the iceberg? *Clin Microbiol Rev*. 1991;**4**:52-60.

246. Cheng A.C., Currie B.J. Melioidosis: Epidemiology, pathophysiology, and management. *Clin Microbiol Rev*. 2005;**18**:383-416.

247. Limmathurotsakul D., Golding N., Dance D.A.B., Messina J.P., Pigott D.M., Moyes C.L., et al. Predicted global distribution of *Burkholderia pseudomallei* and burden of melioidosis. *Nat Microbiol*. 2016;**1**:5.

248. Sanchez-Villamil J.I., Torres A.G. Melioidosis in Mexico, Central America, and the Caribbean. *Trop Med Infect Dis*. 2018;**3**:24.

249. Miralles I.S., Maciel Mdo C., Angelo M.R., Gondini M.m., Frota L.H., dos Reis C.M., et al. *Burkholderia pseudomallei*: a case report of a human infection in Ceará, Brazil. *Rev Inst Med Trop Sao Paulo*. 2004;**46**:51-4.

250. Rolim D.B., Vilar D., Sousa A.Q., Miralles I.S., de Oliveira D.C.A., Harnett G., et al. Melioidosis, northeastern Brazil. *Emerg Infect Dis*. 2005;**11**:1458-60.

251. Katangwe T., Purcell J., Bar-Zeev N., Denis B., Montgomery J., Alaerts M., et al. Human Melioidosis, Malawi, 2011. *Emerg Infect Dis*. 2013;**19**:981-4.

252. Kunnathurparambil S.G., Sathar S.A., Tank D.C., Sressh S., Mukunda M., Narayanan P., et al. Splenic abscess due to chronic meloidosis in a patient previously misdiagnosed as tuberculosis. *Ann Gastroenterol*. 2013;**26**:77-9.

253. Currie B.J., Fisher D.A., Anstey N.M., Jacups S.P. Melioidosis: acute and chronic disease, relapse and re-activation. *Trans Roy Soc Trop Med Hyg*. 2000;**94**:301-4.

254. Suputtamongkol Y., Chaowagul W., Chetchotisakd P., Lertpatanasuwun N., Intaranongpai S., Ruchutrakool T., et al. Risk factors for melioidosis and bacteremic melioidosis. *Clin Infect Dis*. 1999;**29**:408-13.

255. Wuthiekanun V., Chierakul W., Langa S., Chaowagul W., Panpitpat C., Saipan P., et al. Short report: Development of antibodies to *Burkholderia pseudomallei* during childhood in melioidosis-endemic northeast Thailand. *Am J Trop Med Hyg*. 2006;**74**:1074-5.

256. Wiersinga W.J., Currie B.J., Peacock S.J. Melioidosis. *N Engl J Med*. 2012;**367**:1035-44.

257. Nierman W.C., Yu Y., Losada L. The In vitro Antibiotic Tolerant Persister Population in *Burkholderia pseudomallei* is Altered by Environmental Factors. *Front Microbiol*. 2015;**6**:9.

258. Limmathurotsakul D., Peacock S.J. Melioidosis: a clinical overview. *Br Med Bull*. 2011;**99**:125-39.

259. Currie B.J., Ward L., Cheng A.C. The Epidemiology and Clinical Spectrum of Melioidosis: 540 Cases from the 20 Year Darwin Prospective Study. *Plos Neglect Trop Dis*. 2010;**4**:11.

260. Limmathurotsakul D., Wongratanacheewin S., Teerawattanasook N., Wongsuvan G., Chaisuksant S., Chetchotisakd P., et al. Increasing Incidence of Human Melioidosis in Northeast Thailand. *Am J Trop Med Hyg*. 2010;**82**:1113-7.

261. Vidyalakshmi K., Lipika S., Vishal S., Damodard S., Chakrapani M. Emerging clinico-epidemiological trends in melioidosis: analysis of 95 cases from western coastal India. *Int J Infect Dis*. 2012;**16**:E491-E7.

262. Hassan M.R.A., Pani S.P., Peng N.P., Voralu K., Vijayalakshmi N., Mehanderkar R., et al. Incidence, risk factors and clinical epidemiology of melioidosis: a complex socio-ecological emerging infectious disease in the Alor Setar region of Kedah, Malaysia. *BMC Infect Dis*. 2010;**10**:8.

263. Heng B.H., Goh K.T., Yap E.H., Loh H., Yeo M. Epidemiological surveillance of melioidosis in Singapore. *Ann Acad Med Singapore*. 1998;**27**:478-84.

264. Suputtamongkol Y., Hall A.J., Dance D.A.B., Chaowagul W., Rajchanuvong A., Smith M.D., et al. The epidemiology of melioidosis in Ubon Ratchatani, northeast Thailand. *Int J Epidemiol*. 1994;**23**:1082-90.

265. Currie B., Howard D., Nguyen V.T., Withnall K., Merianos A. The 1990-1991 outbreak of melioidosis in the Northern Territory off Australia: clinical aspects. *Southeast Asian J Trop Med Public Health*. 1993;**24**:436-43.

266. Liu X., Pang L., Sim S.H., Goh K.T., Ravikumar S., Win M.S., et al. Association of Melioidosis Incidence with Rainfall and Humidity, Singapore, 2003-2012. *Emerg Infect Dis*. 2015;**21**:159-62.

267. Currie B.J., Jacups S.P. Intensity of rainfall and severity of melioidosis, Australia. *Emerg Infect Dis*. 2003;**9**:1538-42.

268. Mhlanga-Mutangadura T., Morlin G., Smith A.L., Eisenstark A., Golomb M. Evolution of the major pilus gene cluster of Haemophilus influenzae. *J Bacteriol*. 1998;**180**:4693-703.

269. Makarova K.S., Grishin N.V., Koonin E.V. The HicAB cassette, a putative novel, RNA-targeting toxin-antitoxin system in archaea and bacteria. *Bioinformatics*. 2006;**22**:2581-4.

270. Sinha S., Cameron A.D.S., Redfield R.J. Sxy Induces a CRP-S Regulon in *Escherichia coli*. *J Bacteriol*. 2009;**191**:5180-95.

271. Jaskolska M., Gerdes K. CRP-dependent Positive Autoregulation and Proteolytic Degradation Regulate Competence Activator Sxy of *Escherichia coli*. *Mol Microbiol*. 2015;**95**:833-45.

272. Kim D.H., Kang S.M., Park S.J., Jin C., Yoon H.J., Lee B.J. Functional insights into the *Streptococcus pneumoniae* HicBA toxin-antitoxin system based on a structural study. *Nucleic Acids Res*. 2018;**46**:6371-86.

273. Yoon S.H., Kim S.K., Kim J.F. Secretory production of recombinant proteins in *Escherichia coli*. *Recent Pat Biotechnol*. 2010;**4**:23-9.

274. Auclair S.M., Bhanu M.K., Kendall D.A. Signal peptidase I: Cleaving the way to mature proteins. *Protein Sci*. 2012;**21**:13-25.

275. Berrow N.S., Alderton D., Sainsbury S., Nettleship J., Assenberg R., Rahman N., et al. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res*. 2007;**35**:12.

276. Yang J.Y., Yan R.X., Roy A., Xu D., Poisson J., Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods*. 2015;**12**:7-8.

277. Slabinski L., Jaroszewski L., Rychlewski L., Wilson I.A., Lesley S.A., Godzik A. XtalPred: a web server for prediction of protein crystallizability. *Bioinformatics*. 2007;**23**:3403-5.

278. Kabsch W. XDS. *Acta Crystallogr Sect D-Biol Crystallogr*. 2010;**66**:125-32.

279. Winter G. xia2: an expert system for macromolecular crystallography data reduction. *J Appl Crystallogr*. 2010;**43**:186-90.

280. McCoy A.J., Grosse-Kunstleve R.W., Adams P.D., Winn M.D., Storoni L.C., Read R.J. Phaser crystallographic software. *J Appl Crystallogr*. 2007;**40**:658-74.

281. Adams P.D., Afonine P.V., Bunkóczi G., Chen V.B., Davis I.W., Echols N., et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D-Biol Crystallogr*. 2010;**66**:213-21.

282. Taylor G. Introduction to phasing. *Acta Crystallogr D Biol Crystallogr*. 2010;**66**:325-38.

283. Perutz M.F. Isomorphous replacement and phase determination in non-centrosymmetric space groups. *Acta Crystallographica*. 1956;**9**:867-73.

284. Kendrew J.C., Bodo G., Dintzis H.M., Parrish R.G., Wyckoff H., Phillips D.C. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature*. 1958;**181**:662-6.

285. Hendrickson W.A., Horton J.R., Lemaster D.M. Selenomthionyl proteins produced or analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *Embo J*. 1990;**9**:1665-72.

286. Dayhoff M.O., Schwartz R.M., Orcut B.M. A model of evolutionary change in proteins. Dayhoff M.O, editor. Washington, D.C: National Biomedical Research foundation; 1978. 345-52.

287. McLachlan A.D. Tests for comparing related amino-acid sequences. Cytochrome c and cytochrome c 551. *J Mol Biol*. 1971;**61**:409-24.

288. Jones D.T., Taylor W.R., Thornton J.M. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*. 1992;**8**:275-82.

289. Gonnet G.H., Cohen M.A., Benner S.A. Exhaustive matching of the entire protein sequence database. *Science*. 1992;**256**:1443-5.

290. Lim W.A., Farruggio D.C., Sauer R.T. Structural and energetic consequences of disruptive mutations in a protein core. *Biochemistry*. 1992;**31**:4324-33.

291. Goujon M., McWilliam H., Li W.Z., Valentin F., Squizzato S., Paern J., et al. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res*. 2010;**38**:W695-W9.

292. Robert X., Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. 2014;**42**:W320-W4.

293. Doublié S. Production of Selenomethionyl Protein in Prokaryotic and Eukaryotic Expresson Systems. *Methods Mol Biol*. 2007;**363**:91-108.

294. Leahy D.J., Hendrickson W.A., Aukhil I., Erickson H.P. Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein. *Science*. 1992;**258**:987-91.

295. Holton J.M. XANES measurements of the rate of radiation damage to selenomethionine side chains. *J Synchrot Radiat*. 2007;**14**:51-72.

296. Skubák P., Pannu N.S. Automatic protein structure solution from weak X-ray data. *Nat Commun*. 2013;**4**:6.

297. Krissinel E., Uski V., Lebedev A., Winn M., Ballard C. Distributed computing for macromolecular crystallography. *Acta Crystallogr Sect D-Struct Biol*. 2018;**74**:143-51.

298. Sheldrick G.M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr Sect D-Biol Crystallogr*. 2010;**66**:479-85.

299. Abrahams J.P., Leslie A.G.W. Methods used in the structure determination of bovine mitochondrial F-1 ATPase. *Acta Crystallogr Sect D-Biol Crystallogr*. 1996;**52**:30-42.

300. Cowtan K. Recent developments in classical density modification. *Acta Crystallogr Sect D-Biol Crystallogr*. 2010;**66**:470-8.

301. Murshudov G.N., Skubák P., Lebedev A.A., Pannu N.S., Steiner R.A., Nicholls R.A., et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr Sect D-Biol Crystallogr*. 2011;**67**:355-67.

302. Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr Sect D-Biol Crystallogr*. 2006;**62**:1002-11.

303. Emsley P., Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr Sect D-Biol Crystallogr*. 2004;**60**:2126-32.

304. McGuffin L.J., Bryson K., Jones D.T. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000;**16**:404-5.

305. Terwilliger T.C., Grosse-Kunstleve R.W., Afonine P.V., Moriarty N.W., Zwart P.H., Hung L.W., et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr Sect D-Biol Crystallogr*. 2008;**64**:61-9.

306. Afonine P.V., Grosse-Kunstleve R.W., Echols N., Headd J.J., Moriarty N.W., Mustyakimov M., et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr Sect D-Biol Crystallogr*. 2012;**68**:352-67.

307. Chen V.B., Arendall W.B., Headd J.J., Keedy D.A., Immormino R.M., Kapral G.J., et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr Sect D-Biol Crystallogr*. 2010;**66**:12-21.

308. Joosten R.P., Long F., Murshudov G.N., Perrakis A. The PDB_REDO server for macromolecular structure model optimization. *IUCrJ*. 2014;**1**:213-20.

309. Krissinel E., Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 2007;**372**:774-97.

310. Tina K.G., Bhadra R., Srinivasan N. PIC: Protein Interactions Calculator. *Nucleic Acids Res*. 2007;**35**:W473-W6.

311. Holm L., Laakso L.M. Dali server update. *Nucleic Acids Res*. 2016;**44**:W351-W5.

312. Schildbach J.F., Karzai A.W., Raumann B.E., Sauer R.T. Origins of DNA-binding specificity: Role of protein contacts with the DNA backbone. *Proc Natl Acad Sci U S A*. 1999;**96**:811-7.

313. Gomis-Ruth F.X., Sola M., Acebo P., Parraga A., Guasch A., Eritja R., et al. The structure of plasmid-encoded transcriptional repressor CopG unliganded and bound to its operator. *Embo J*. 1998;**17**:7404-15.

314. Ahn D.H., Lee K.Y., Lee S.J., Park S.J., Yoon H.J., Kim S.J., et al. Structural analyses of the MazEF4 toxin-antitoxin pair in *Mycobacterium tuberculosis* provide evidence for a unique extracellular death factor. *J Biol Chem*. 2017;**292**:18832-47.

315. Weihofen W.A., Cicek A., Pratto F., Alonso J.C., Saenger W. Structures of omega repressors bound to direct and inverted DNA repeats explain modulation of transcription. *Nucleic Acids Res*. 2006;**34**:1450-8.

316. Larson J.D., Jenkins J.L., Schuermann J.P., Zhou Y.Z., Becker D.F., Tanner J.J. Crystal structures of the DNA-binding domain of *Escherichia coli* proline utilization A flavoprotein and analysis of the role of Lys9 in DNA recognition. *Protein Sci*. 2006;**15**:2630-41.

317. Kang S.M., Kim D.H., Lee K.Y., Park S.J., Yoon H.J., Lee S.J., et al. Functional details of the *Mycobacterium tuberculosis* VapBC26 toxin-antitoxin system based on a structural study: insights into unique binding and antibiotic peptides. *Nucleic Acids Res*. 2017;**45**:8564-80.

318. Liu Y., Eisenberg D. 3D domain swapping: As domains continue to swap. *Protein Sci*. 2002;**11**:1285-99.

319. Putnam C.D., Hammel M., Hura G.L., Tainer J.A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys*. 2007;**40**:191-285.

320. Putnam D.K., Lowe Jr E.W., Meiler J. Reconstruction of SAXS Profiles from Protein Structures. *Comput Struct Biotechnol J*. 2013;**8**:e20138006.

321. Trewhella J. Small-angle scattering and 3D structure interpretation. *Curr Opin Struct Biol*. 2016;**40**:1-7.

322. Vestergaard B. Analysis of biostructural changes, dynamics, and interactions - Small-angle X-ray scattering to the rescue. *Arch Biochem Biophys*. 2016;**602**:69-79.

323. Glatter O. New method for evaluation of small-angle scattering data. *J Appl Crystallogr*. 1977;**10**:415-21.

324. Svergun D.I., Koch M.H.J. Small-angle scattering studies of biological macromolecules in solution. *Rep Prog Phys*. 2003;**66**:1735-82.

325. Schneidman-Duhovny D., Hammel M., Tainer J.A., Sali A. Accurate SAXS Profile Computation and its Assessment by Contrast Variation Experiments. *Biophys J*. 2013;**105**:962-74.

326. Rambo R.P., Tainer J.A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature*. 2013;**496**:477-81.

327. Trewhella J., Duff A.P., Durand D., Gabel F., Guss J.M., Hendrickson W.A., et al. 2017 publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution: an update. *Acta Crystallogr Sect D-Struct Biol*. 2017;**73**:710-28.

328. Vagin A., Teplyakov A. Molecular replacement with MOLREP. *Acta Crystallogr Sect D-Biol Crystallogr*. 2010;**66**:22-5.

329. Winn M.D., Ballard C.C., Cowtan K.D., Dodson E.J., Emsley P., Evans P.R., et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D-Biol Crystallogr*. 2011;**67**:235-42.

330. Vagin A.A., Isupov M.N. Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr Sect D-Biol Crystallogr*. 2001;**57**:1451-6.

331. Pannu N.S., Murshudov G.N., Dodson E.J., Read R.J. Incorporation of prior phase information strengthens maximum-likelihood structure refinement. *Acta Crystallogr Sect D-Biol Crystallogr*. 1998;**54**:1285-94.

332. Baker E.G., Williams C., Hudson K.L., Bartlett G.J., Heal J.W., Goff K.L.P., et al. Engineering protein stability with atomic precision in a monomeric miniprotein. *Nat Chem Biol*. 2017;**13**:764-70.

333. Masliah G., Barraud P., Allain F.H.T. RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. *Cell Mol Life Sci*. 2013;**70**:1875-95.

334. Crooks G.E., Hon G., Chandonia J.M., Brenner S.E. WebLogo: A sequence logo generator. *Genome Res*. 2004;**14**:1188-90.

335. Hall A.M.J., Gollan B., Helaine S. Toxin-antitoxin systems: reversible toxicity. *Curr Opin Microbiol*. 2017;**36**:102-10.

336. Lobato-Márquez D., Díaz-Orejas R., García-Del Portillo F. Toxin-antitoxins and bacterial virulence. *FEMS Microbiol Rev*. 2016;**40**:592-609.

337. Schuck P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys J*. 2000;**78**:1606-19.

338. Schuck P. Analytical ultracentrifugation as a tool for studying protein interactions. *Biophys Rev*. 2013;**5**:159-71.

339. Fujita H. Foundations of ultracentrifugal analysis. Fujita H, editor. New York: J. Wiley & Sons; 1975.

340. Fujita H. Mathematical theory of sedimentation analysis. Fujita H, editor. New York: Academic Press; 1962.

341. Zhao H., Brautigam C.A., Ghirlando R., Schuck P. Overview of current methods in sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation. *Curr Protoc Protein Sci*. 2013;**71**:20.12.1–20.12.49.

342. Leney A.C., Heck A.J.R. Native Mass Spectrometry: What is in the Name? *J Am Soc Mass Spectrom*. 2017;**28**:5-13.

343. Heck A.J.R. Native mass spectrometry: a bridge between interactomics and structural biology. *Nat Methods*. 2008;**5**:927-33.

344. Brázda V., Kolomaznik J., Lýsek J., Hároníková L., Coufal J., Šťastnÿ J. Palindrome analyser - A new web-based server for predicting and evaluating inverted repeats in nucleotide sequences. *Biochem Biophys Res Commun*. 2016;**478**:1739-45.

345. Medina-Rivera A., Defrance M., Sand O., Herrmann C., Castro-Mondragon J.A., Delerce J., et al. RSAT 2015: Regulatory Sequence Analysis Tools. *Nucleic Acids Res*. 2015;**43**:W50-W6.

346. Solovyev V., Salamov A. Automatic Annotation of Microbial Genomes and Meagenomic Sequences. In Metagenomics and it's applications in Agriculture, Biomedicine and Environmental Sciences. Li R.W, editor. New York: Nova Science Publishers; 2011.

347. Cheow L.F., Viswanathan R., Chin C.S., Jennifer N., Jones R.C., Guccione E., et al. Multiplexed Analysis of Protein-Ligand Interactions by Fluorescence Anisotropy in a Microfluidic Platform. *Anal Chem*. 2014;**86**:9901-8.

348. Checovich W.J., Bolger R.E., Burke T. Fluoresence polarization- a new tool for cell and molecular biology. *Nature*. 1995;**375**:254-6.

349. Williams J.W., Morrison J.F. The kinetics of reversible tight-binding inhibition. *Methods Enzymol*. 1979;**63**:437-67.

350. Kuzmic P., Elrod K.C., Cregar L.M., Sideris S., Rai R., Janc J.W. High-throughput screening of enzyme inhibitors: Simultaneous determination of tight-binding inhibition constants and enzyme concentration. *Anal Biochem*. 2000;**286**:45-50.

351. Murphy D.J. Determination of accurate K-I values for tight-binding enzyme inhibitors: an in silico study of experimental error and assay design. *Anal Biochem*. 2004;**327**:61-7.

352. Greco W.R., Hakala M.T. Evaluation of methods for estimating the dissociation constant of tight binding enzyme inhibitors. *J Biol Chem*. 1979;**254**:2104-9.

353. Laniel M.A., Béliveau A., Guérin S.L. Electrophoretic Mobility Shift Assays for the Analysis of DNA-Protein Interactions. *Methods Mol Biol*. 2001;**148**:13-30.

354. Sidorova N.Y., Hung S., Rau D.C. Stabilizing labile DNA-protein complexes in polyacrylamide gels. *Electrophoresis*. 2010;**31**:648-53.

355. Abrusan G., Marsh J.A. Alpha Helices Are More Robust to Mutations than Beta Strands. *PLoS Comput Biol*. 2016;**12**:16.

356. Matthews B.W. Genetic and structural analysis of the protein stability problem. *Biochemistry*. 1987;**26**:6885-8.

357. Spolar R.S., Ha J.H., Record M.T. Hydrophobic effect in protein folding and other noncovalent processes invoving proteins. *Proc Natl Acad Sci U S A*. 1989;**86**:8382-5.

358. van Dijk M., Bonvin A. 3D-DART: a DNA structure modelling server. *Nucleic Acids Res*. 2009;**37**:W235-W9.

359. van Zundert G.C.P., Rodrigues J., Trellet M., Schmitz C., Kastritis P.L., Karaca E., et al. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol*. 2016;**428**:720-5.

360. Magnuson R., Yarmolinsky M.B. Corepression of the P1 addiction operon by Phd and Doc. *J Bacteriol*. 1998;**180**:6342-51.

361. Huang X.Y. Fluorescence polarization competition assay: The range of resolvable inhibitor potency is limited by the affinity of the fluorescent ligand. *J Biomol Screen*. 2003;**8**:34-8.

362. Cheng Y., Prusoff W.H. Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem Pharmacol*. 1973;**22**:3099-108.

363. Munson P.J., Rodbard D. An exact correction to the Cheng-Prusoff correction. *J Recept Res*. 1988;**8**:533-46.

364. Cer R.Z., Mudunuri U., Stephens R., Lebeda F.J. IC50-to-Ki: a web-based tool for converting IC50 to K-i values for inhibitors of enzyme activity and ligand binding. *Nucleic Acids Res*. 2009;**37**:W441-W5.

365. Loris R., Garcia-Pino A. Disorder- and Dynamics-Based Regulatory Mechanisms in Toxin-Antitoxin Modules. *Chem Rev*. 2014;**114**:6933-47.

366. Budde P.P., Davis B.M., Yuan J., Waldor M.K. Characterization of a higBA toxin-antitoxin locus in *Vibrio cholerae*. *J Bacteriol*. 2007;**189**:491-500.

367. Hadzi S., Garcia-Pino A., Haesaerts S., Jurenas D., Gerdes K., Lah J., et al. Ribosome-dependent *Vibrio cholerae* mRNAse HigB2 is regulated by a beta-strand sliding mechanism. *Nucleic Acids Res*. 2017;**45**:4972-83.

368. Hadži S. Structural and thermodynamic basis of transcription regulation of the *hig*BA2 toxin-antitoxin module from *Vibrio cholerae*. Vrije Universiteit Brussel2017.

369. Tian Q.B., Hayashi T., Murata T., Terawaki Y. Gene product identification and promoter analysis of hig locus of plasmid Rts1. *Biochem Biophys Res Commun*. 1996;**225**:679-84.

370. Tian Q.B., Ohnishi M., Murata T., Nakayama K., Terawaki Y., Hayashi T. Specific protein-DNA and protein-protein interaction in the hig gene system, a plasmid-borne proteic killer gene system of plasmid Rts1. *Plasmid*. 2001;**45**:63-74.

371. Studier F.W. Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif*. 2005;**41**:207-34.

372. Sambrook J., Fritsch E.F., Maniatis T. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory; 1989.

373. Laemmli U.K. Cleavage of Structural Proteins during Assembly of Head of Bacteriophage-T4. *Nature*. 1970;**227**:680-5.

374. Wen Y.R., Sobott F., Devreese B. ATP and autophosphorylation driven conformational changes of HipA kinase revealed by ion mobility and crosslinking mass spectrometry. *Anal Bioanal Chem*. 2016;**408**:5925-33.

375. Pace C.N., Vajdos F., Fee L., Grimsley G., Gray T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci*. 1995;**4**:2411-23.

376. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., et al. Protein Identification and Analysis Tools on the ExPASy Server. Walker, J.M, editor. New York: Humana Press; 2005.

377. Schrõdinger L. The PyMOL Molecular Graphics System, Version 1.3 [

378. Tayyab S., Qamar S., Islam M. Size exclusion chromatography and size exclusion hplc of proteins. *Biochem Educ*. 1991;**19**:149-52.

379. Lamm O. The theory and method of ultra centrifuging. *Z Phys Chem A-Chem Thermodyn Kinet Elektrochem Eigenschlehre*. 1929;**143**:177-90.

380. Laue T.M., Shah B.D., Ridgeway T.M., Pelletier S.L. Computer-aided interpretation of analytical sedimentation data for proteins. Harding, S.E., Rowe, A.J and Horton, J.C, editors. Cambridge: Royal Society of Chemistry; 1992.

381. Perkins S.J. Protein volume and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *Eur J Biochem*. 1986;**157**:169-80.

382. Brautigam C.A. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. In: Cole JL, editor. Analytical Ultracentrifugation. Methods in Enzymology. 562. San Diego: Elsevier Academic Press Inc; 2015. p. 109-33.

383. Demeler B. Ultrascan, A comprehensive data analysis software package for analytical ultracentrifugation experiments. Modern analytical ultracentrifugation: Techniques and Methods. UK: Royal Society of Chemistry; 2005.

384. Johnson M.L., Faunt L.M. Parameter estimation by least squares methods. *Method Enzymol*. 1992;**210**:1-37.

385. Hellman L.M., Fried M.G. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc*. 2007;**2**:1849-61.

386. Copeland R.A. Kinetics of single substrate enzyme reactions. Enzymes: A practical introduction to structure, mechanism and data analysis. 2nd edition ed. New York: John Wiley & Sons; 2000.

387. Delaglio F., Grzesiek S., Vuister G.W., Zhu G., Pfeifer J., Bax A. NMRPIPE- A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR*. 1995;**6**:277-93.

388. Vranken W.F., Boucher W., Stevens T.J., Fogh R.H., Pajon A., Llinas P., et al. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins*. 2005;**59**:687-96.

389. Svergun D.I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr*. 1992;**25**:495-503.

390. Franke D., Svergun D.I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr*. 2009;**42**:342-6.

391. Volkov V.V., Svergun D.I. Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Crystallogr*. 2003;**36**:860-4.

392. Svergun D.I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J*. 1999;**76**:2879-86.

393. Kozin M.B., Svergun D.I. Automated matching of high- and low-resolution structural models. *J Appl Crystallogr*. 2001;**34**:33-41.

394. Goldschmidt L., Cooper D.R., Derewenda Z.S., Eisenberg D. Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. *Protein Sci*. 2007;**16**:1569-76.

395. Sternberg M.J.E., Chickos J.S. Protein Side-Chain Conformational Entropy Derived from Fusion Data - Comparison with Other Empirical Scales. *Protein Eng*. 1994;**7**:149-55.

396. Jahn N., Brantl S., Strahl H. Against the mainstream: the membrane-associated type I toxin BsrG from *Bacillus subtilis* interferes with cell envelope biosynthesis without increasing membrane permeability. *Mol Microbiol*. 2015;**98**:651-66.

397. Guo Y.X., Quiroga C., Chen Q., McAnulty M.J., Benedik M.J., Wood T.K., et al. RalR (a DNase) and RalA (a small RNA) form a type I toxin-antitoxin system in Escherichia coli. *Nucleic Acids Res*. 2014;**42**:6448-62.

398. Kawano M. Divergently overlapping cis-encoded antisense RNA regulating toxin-antitoxin systems from E-coli hok/sok, ldr/rdl, symE/symR. *RNA Biol*. 2012;**9**:1520-7.

399. 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*. 2014;**42**:4590-605.

400. Hampton H.G., Jackson S.A., Fagerlund R.D., Vogel A.I.M., Dy R.L., Blower T.R., 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*. 2018;**430**:1141-56.

401. Wessner F., Lacoux C., Goeders N., d'Hérouel A.F., Matos R., Serror P., et al. Regulatory crosstalk between type I and type II toxin-antitoxin systems in the human pathogen Enterococcus faecalis. *RNA Biol*. 2015;**12**:1099-108.

402. Yang M., Gao C.H., Wang Y., Zhang H., He Z.G. Characterization of the Interaction and Cross-Regulation of Three Mycobacterium tuberculosis RelBE Modules. *PLoS One*. 2010;**5**:12.

403. Kwan B.W., Lord D.M., Peti W., Page R., Benedik M.J., Wood T.K. The MqsR/MqsA toxin/antitoxin system protects *Escherichia coli* during bile acid stress. *Environ Microbiol*. 2015;**17**:3168-81.

404. Wilbaux M., Mine N., Guerout 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*. 2007;**189**:2712-9.

405. Kasari V., Mets T., Tenson T., Kaldalu N. Transcriptional cross-activation between toxin-antitoxin systems of *Escherichia coli*. *BMC Microbiol*. 2013;**13**:13.

406. Ahidjo B.A., Kuhnert D., McKenzie J.L., Machowski E.E., Gordhan B.G., Arcus V., et al. VapC Toxins from *Mycobacterium tuberculosis* Are Ribonucleases that Differentially Inhibit Growth and Are Neutralized by Cognate VapB Antitoxins. *PLoS One*. 2011;**6**:14.

407. Sala A., Bordes P., Genevaux P. Multiple Toxin-Antitoxin Systems in *Mycobacterium tuberculosis*. *Toxins*. 2014;**6**:1002-20.

408. Fivian-Hughes A.S., Davis E.O. Analyzing the Regulatory Role of the HigA Antitoxin within *Mycobacterium tuberculosis*. *J Bacteriol*. 2010;**192**:4348-56.