THE UNIVERSITY *of* EDINBURGH

Protein-Protein Interactions of the Cold Shock Protein CspE of *Salmonella* Typhimurium

Peter J Gwynne

PhD Thesis

University of Edinburgh - 2015

THE UNIVERSITY *of* EDINBURGH

Declaration

I declare that all of the work presented herein is my own, except where stated, and that the document was written by myself.

Signed _____

Peter J Gwynne

September 2014

Acknowledgements

Thanks, firstly, to Maurice and to Garry for their countless contributions. The advice and assistance of Flavia de lima Alves and Juri Rappsilber was invaluable for the mass spectrometry. Sander Granneman and the Tollervey lab were of great help with the HTP purifications. Everyone on the 6th, 7th, and 8th floors of Darwin provided virtually endless chat and advice, some of which might have been about science. The Edinburgh Revue were a constant and valuable bad influence, and signing up for Bright Club remains one of my better decisions. I'm grateful for the indulgence of everyone at home while I refuse to grow up and get a real job. Long may it continue.

Table of Contents

Abstract

Despite their name, a number of the cold shock proteins are expressed during normal growth, and not just during cold shock, in several species. The function of these constitutively expressed CspA paralogues is unclear. In *Salmonella* Typhimurium (a major worldwide cause of gastrointestinal disease) they have been linked to various stress responses and the establishment of virulence. Study of the cold shock proteins as gene regulators is therefore of great interest, and they also have potential as targets for antimicrobial development.

CspE in *Salmonella* Typhimurium is constitutively expressed during normal growth. In order to determine its function, attempts were made to identify the interactions it forms with other cellular proteins. Initially, a proteomic investigation attempted to identify proteins which complex with CspE by *in vivo* cross-linking and affinity purification followed by mass spectrometry. Although no defined complex was consistently identified, the results suggested a handful of proteins which might interact with CspE in a weak or transient manner. These proteins included many from the nucleoid and ribosomal entry site, hinting at CspE's cellular localisation.

In order to investigate these transient interactions, a bacterial two-hybrid system was employed. Interactions between CspE and HupA, a nucleoid protein identified in the proteomic analysis, were probed, as were interactions between CspE and CsdA, an RNA helicase thought to function co-operatively with CspE. The two-hybrid system also allowed investigation of CspE dimerisation, which has been reported *in vitro* but not investigated *in vivo* until this study. CspE appears not to interact significantly with either HupA, CsdA, or itself at 37^{o}C.

Finally in a further attempt to identify interactions of CspE, a genomic library was created to test CspE interactions by two-hybrid assay with random peptides derived from the whole *Salmonella* genome. The library was successfully created and screened for evidence of interaction, and revealed an association between CspE and a transcriptional repressor, DeoT. DeoT is a repressor of several genes for catabolic processes, suggesting a role for CspE in the regulation of central metabolism. The findings of this work present a number of novel discoveries and several interesting opportunities for further studies.

Figures

Tables

Abbreviations

(p)ppGpp	Guanosine tetraphosphate or pentaphosphate
aa	Amino acid (residues)
AEBSF	4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride
Amp	Ampicillin (carbenicillin)
BACTH	Bacterial adenylate cyclase two-hybrid
BLAST	Basic local alignment search tool
BS3	bis(sulfosuccinimidyl) suberate
Cam	Chloramphenicol
cAMP	Cyclic adenosine monophosphate
CAP	cAMP-activated protein
c-di-AMP	Cyclic di-adenosine monophosphate
ChIP-seq	Chromatin immunoprecipitation and sequencing
CIP	Cold induced protein
CLMS	Cross-linking mass spectrometry
CSD	Cold shock domain
Csp	Cold shock protein
DH5α	*Escherichia coli* strain DH5α
DHM1	*Escherichia coli* strain DHM1
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deooxynucleotide triphosphase
DSS	Disuccinimidyl suberate
FRET	Fluorescent resonance energy transfer
GI	Gastrointestinal
GTP	Guanosine triphosphate
HRP	Horseradish peroxidase
HTH	Helix-turn-helix
HTP	Histidine tandem purification
ID	Infective dose

IF	Initiation factor
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
LB	Lysogeny Broth
LC MS/MS	Liquid chromatography tandem mass spectrometry
LT2	*Salmonella* Typhimurium strain LT2
MCS	Multiple cloning site
mRNA	Messenger RNA
MS	Mass spectrometry
NAD	Nicotinamide adenine dinucleotide
NADPH	Nicotinamide adenine dinucleotide phosphate
Nal	Nalidixic acid
NCBI	National Centre for Biotechnology Information
NEB	New England Biolabs
nt	Nucleotide
OB	Oligonucleotide/oligosaccharide binding
OD_{600}	Optical density at 600 nm
ONPG	Ortho-Nitrophenyl-β-galactoside
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PNPase	Polynucleotide phosphorylase
PPI	Protein-protein interaction
qPCR	Quantitative PCR
RBS	Ribosome binding site
RNA	Ribonucleic acid
RNAP	RNA polymerase
RNP	Ribonucleotide-protein binding site
RPM	Revolutions per minute
rRNA	Ribosomal RNA
SDS	Sodium dodecylsulfate

SIF	*Salmonella*-induced filament
SL1344	*Salmonella* Typhimurium strain SL1344
SPI	*Salmonella* pathogenicity island
sRNA	Small RNA
ssDNA	Single-stranded DNA
ssRNA	Single-stranded RNA
SVC	*Salmonella*-containing vacuole
T18	18 kDa fragment of adenylate cyclase
T25	25 kDa fragment of adenylate cyclase
TAP	Tandem affinity purification
TCA	Tricarboxylic acid
TEV	Tobacco Etch Virus
TFA	Trifluoroacetic acid
TNF	Tumour necrosis factor
tRNA	Transfer RNA
TST	Tris-salt-Tween
UMP	Uridine monophosphate
UTP	Uridine triphosphate
UTR	Untranslated region
wHTH	Winged helix-turn-helix
WT	Wild-type
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Chapter 1

Introduction

Salmonella enterica serovar Typhimurium

The *Salmonella* genus contains two species, *Salmonella enterica* and *Salmonella bongori*. *S. enterica* contains seven subspecies and over 2,000 serotypes classified by flagellar and lipopolysaccharide antigens. Serotypes exhibit a range of host specificities, with some highly adaptable and some specific to a single species (Fierer & Guiney 2001). Serotypes of the subspecies *enterica* include *S. enterica enterica* Typhi, the causative agent of typhoid disease, and *S. enterica enterica* Typhimurium, a major foodborne pathogen. The subspecies is abbreviated *S.* Typhimurium (Tindall et al. 2005).

Isolated from pigs in the 1880s (Schultz 2008), salmonellae are short, motile, Gram-negative rods, up to 2.5 μm in length. As pathogens of the gastrointestinal (GI) tract they are facultatively anaerobic and grow optimally at 37-41^{o}C (Holt et al. 2000). The laboratory strain SL1344 was initially isolated from a cow and made a histidine auxotroph as a pathogenic model for vaccine development (Hoiseth & Stocker 1981). The most recently published genome of *S.* Typhimurium SL1344 is 4.88 mbp in length, with a GC content of 52% and 4,446 predicted protein coding genes (Kröger et al. 2012). Three plasmids are carried: pSLT, pCol1B9, and pRSF1010. pSLT carries virulence genes (García et al. 2014), while pCol1B9 drives conjugation in the gut (Stecher et al. 2012). The function of pRSF1010 has not been studied.

Salmonella infection is generally considered as either typhoidal (systemic) or non-typhoidal (salmonellosis or gastrointestinal infection). Although a less severe illness, non-typhoidal salmonellosis remains a major health burden, particularly in the developing world. Precise statistics are hard to gather due to issues of surveillance and diagnosis, but it is estimated that 94 million cases and 155,000 deaths result from *Salmonella* infection each year. The majority of infections (80 million) are thought to be foodborne (Majowicz et al. 2010). Typhoidal salmonellosis, by comparison, infects around 23 million people per year and results in 200,000 deaths (Crump et al. 2004); again, infections are typically foodborne. salmonellae account for the majority (over half) of foodborne disease outbreaks in England and Wales (Gormley et al. 2011).

Pathogenicity in *S*. Typhimurium is substantially determined by the presence of at least five pathogenicity islands (Marcus et al. 2000). Around 4% of the *S.* Typhimurium genome (~200 genes) is thought to be given over to virulence factors, reflecting a complicated intracellular infective cycle (Bowe et al. 1998). Although antimicrobial therapy is usually effective, numerous drug-resistant strains have been identified (Tamang et al. 2007; Kingsley et al. 2009). *Salmonella* has an unusual redundancy of metabolic pathways and can utilise a wide range of nutrients from the host, which makes development of antimicrobials difficult (Becker et al. 2006). The search for novel target pathways for antimicrobial development is therefore of critical importance

Salmonella Typhimurium in the food industry

Salmonella ssp. outbreaks are traditionally associated with meat products and poultry in particular. Incidence of *Salmonella* contamination of chicken ranges from 43% in China (Yang et al. 2014) to 11% in the USA (Mazengia et al. 2014). Despite this association, *Salmonella ssp*. can be carried on virtually any food, and are the most common pathogens isolated from fruit and vegetables (Heaton & Jones 2008). Although the frequency of outbreaks in the UK is falling, they remain common: in England and Wales there were 321 cases in 2011, and were 52 outbreaks from 2000-2011 (Harker et al. 2014). *S*. Typhimurium, along with *S*. Enteritidis, is consistently one of the most commonly isolated serotypes of *Salmonella* enterica (Jackson et al. 2013; Pulido-Landínez et al. 2013). For research purposes, *S*. Typhimurium has the advantage of a well-characterised model strain (SL1344) which retains pathogenicity determinants.

The reason for *Salmonella*'s common isolation from meat products is its ability to survive in a wide range of hosts. Commensal infection of cattle (Van Kessel et al. 2012) and chickens (Humphrey 2006) is often asymptomatic, making detection difficult. Meanwhile, widespread use of prophylactic antibiotics in farming, particularly in the USA, increases the prevalence of antibiotic resistance (Gyles 2008). Food animals are not the only zoonotic reservoir for salmonellae. Cases of cross-infection from

domestic (Dipineto et al., 2014) and wild (Lawson et al. 2014) animals are frequently reported.

A major factor in *Salmonella* pathogenicity is its ability to survive in a wide range of exacting conditions. Outside the host dessication, starvation and extremes of temperature are encountered. Inside the host pH fluctuates, innate antimicrobials are encountered, and the immune system subjects infecting agents to oxidative attack (Foster & Spector 1995). During food processing salmonellae are able to survive dessication, heating, chilling, and osmotic stress. Such is the importance of extra-host survival that some antimicrobial resistance plasmids and other virulence genes are believed to impact pathogenicity by facilitating various stress responses (Paytubi et al. 2014).

Salmonella Typhimurium infection

Salmonella Typhimurium is introduced to a host by the faecal-oral route and undergoes a complex intracellular infective cycle reliant on the action of a number of pathogenicity islands, of which those named 1 and 2 are the best studied. Initial infection of the intestinal mucosa is followed by invasion and replication inside intestinal epithelia and macrophages. After replication inside intestinal cells, bacterial cells can become systemic by invading immune cells in the blood and lymph. Systemic infection is common in the mouse model, but less so in humans. The infective pathway is illustrated in Figure 1.1.

Having passed the acidic environment of the stomach, *S.* Typhimurium virulence factors contained in *Salmonella* pathogenicity island (SPI) 2 confer a competitive advantage over the resident gut microbiota. By inducing inflammation SPI-2 genes promote production of tetrathionate, which can be used in fermentation by *Salmonella ssp.* but not by most other commensals (Winter et al. 2010). Once established in the gut, association with the epithelia of the small intestine is mediated by fimbriae and flagella (amongst others) (Dibb-Fuller et al. 1999).

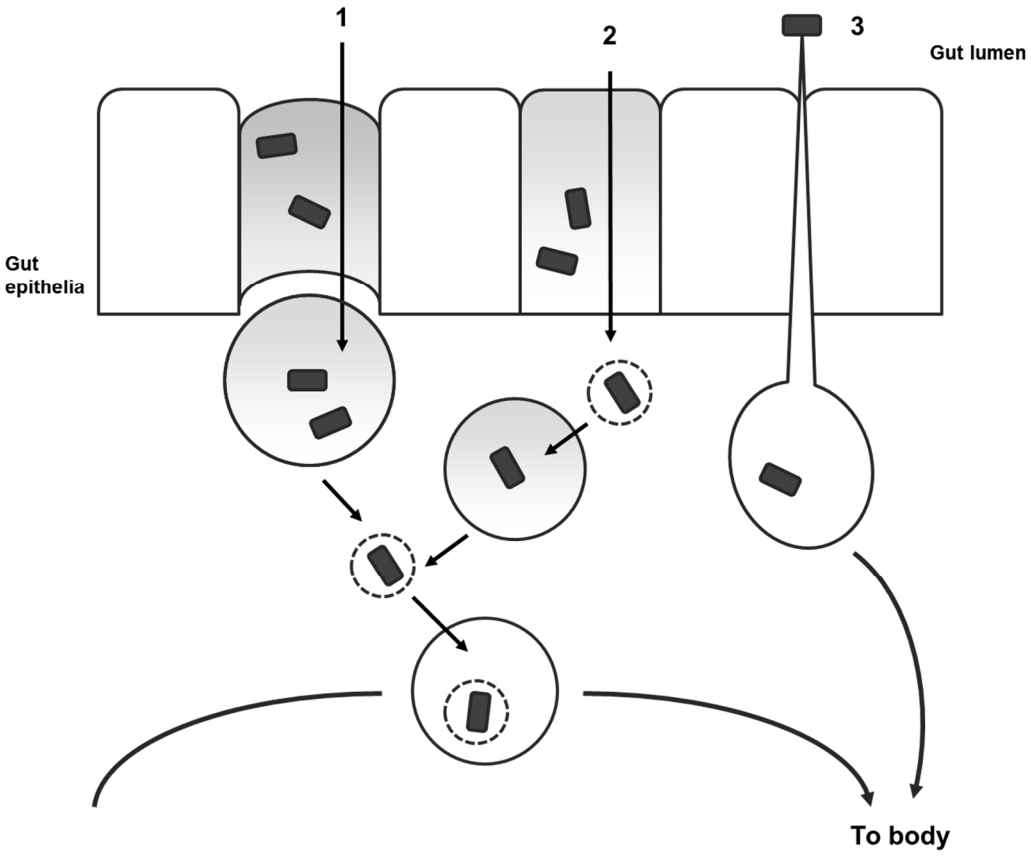


Figure 1.1. *Infective routes of* Salmonella *Typhimurium; cells in which apoptosis is induced shaded grey. First (and preferred) route is through M cells which present infecting bacteria directly to macrophages. Second route is through invasion and lysis of epithelial cells. Apoptotic bodies resulting from epithelial apoptosis are internalised by intestinal macrophages.* S. *Typhimurium replicates inside intestinal macrophages, causes apoptosis, and is taken up by systemic macrophages which transport bacterial cells throughout the body. The third route is mediated by dendritic cells, which can reach into the gut lumen and internalise infecting cells, before carrying them around the body.*

Invasion of the host cells is carried out largely by genes carried on SPI-1, including those encoding excreted effector proteins and a type-three secretion system. Effectors include the actin-binding SipA, which promotes actin polymerisation at the point of bacterial attachment (Zhou et al. 1999) and SopD, which drives vesicle formation (Bakowski et al. 2007). Having been internalised, *S*. Typhimurium downregulates inflammation to limit the immune response of the host epithelial cells (Fu & Galán 1999). *S*. Typhimurium preferentially invades M-cells (specialised immune cells in the intestinal wall) and are subsequently transcytosed directly into intestinal macrophages in a process which kills the M cell (Jones et al. 1994).

Within the host cell, infecting bacteria survive within a modified endocytic compartment termed the *Salmonella*-containing vacuole (SVC) (Uchiya et al. 1999). Formation of this compartment is dependent on SPI-2 genes, the expression of which is induced by low magnesium and phosphate levels encountered in the host cell (Deiwick et al. 1999). The integrity of the SCV is maintained by secreted proteins which induce actin remodelling, preventing completion of the endocytic pathway (Uchiya et al. 1999). The SCV is positioned close to the nucleus and the Golgi body, from where it matures into a stable structure which facilitates replication (Deiwick et al. 2006).

The SCV matures into a tubular structure termed the *Salmonella*-induced filament (SIF), again dependent on the function of SPI-2. The function of the SIF is not clear but may help to overcome nutrient starvation (Rajashekar et al. 2008). Having created an SIF, invading cells protect it from the host immune response. Reactive oxygen species are evaded by SPI-2 genes which prevent the localisation of NADPH oxidase to the SCV (Vazquez-Torres 2000). A similar disruption of the localisation of inducible nitric oxide synthase protects from killing by reactive nitrogen (Chakravortty et al. 2002). SPI-2 genes are induced under the control of SsrB, which responds to high NO; presumably a signal of localisation inside the macrophage (Husain et al. 2010, p.203). In addition, numerous systems exist to detoxify reactive nitrogen and oxygen species, which are described in detail below.

Apoptosis of endothelial cells is promoted by the secretion of a pair of proteins from the SIF. Expressed from the virulence plasmid pSLT, SpvB causes actin remodelling and cell death (Kurita et al. 2003). SlrP interacts with the apoptotic

regulator thioredoxin (Bernal-Bayard & Ramos-Morales 2009) and with DnaJ, a protein folding chaperone. This second interaction induces cell death by accumulation of misfolded proteins (Bernal-Bayard et al. 2010). In intestinal macrophages, apoptosis is induced by various proteins of SPI-1 (Velden et al. 2000).

After initial intestinal infection, in susceptible hosts, *S*. Typhimurium cells are spread throughout the body by the action of a number of phagocytic immune cells. Such cells are recruited as part of the inflammatory response and take up infecting bacteria from the apoptotic bodies of epithelia and macrophages (Fabrega & Vila 2013). In systemic macrophages, a delayed apoptosis is mediated by SPI-2 and facilitates the spread of infecting cells around the body (Velden et al. 2000). Antigen-presenting dendritic cells are highly mobile and carry infecting cells to the spleen and hence the rest of the body (Jantsch et al. 2011). *Salmonella* prevents the presentation of antigens by the ubiquitination of MHC-II complexes. This is a process mediated by a secreted protein of SPI-2, SsvA, but the mechanism is unknown (Lapaque et al. 2009). An additional pathway of entry to the host is through dendritic cells, which can reach into the intestinal lumen and take up infecting cells directly (Rescigno et al. 2001).

Salmonella Stress Responses

Salmonella has a diverse life cycle, moving between the environment and the host. In the environment, diverse stresses are encountered including nutrient starvation, extremes of temperature and pH, osmotic shock, desiccation, and predation. *S*. Typhimurium survives in soil, freshwater, saltwater, on surfaces, and a in variety of plant and animal reservoirs (Winfield & Groisman 2003). Once a host has been entered a further series of stresses are encountered in the complex life cycle described above. The innate immune system challenges incoming bacteria with acid in the stomach and bile in the small intestine where, in addition, oxygen is limited and osmotic shock is suffered (Álvarez-Ordóñez et al. 2011). As discussed above, during intracellular infection oxidative attack and nutrient starvation are encountered. Environmental stresses which arise during the *S*. Typhimurium life cycle are outlined in Figure 1.2. The responses to each environmental stress, and others not mentioned here, are complex and varied. Some of the main aspects of *S*. Typhimurium response to a number of environmental challenges are outlined below.

A recent study assessed the contribution of *S*. Typhimurium genes to intestinal colonisation in several important animal species (cow, pig, chicken and mouse) using a collection of 7,700 random mutants (Chaudhuri et al. 2013). Given the factors described above, it is unsurprising that many stress response genes were found to contribute to colonisation in at least one species. These included those for effectors of acid adaptation, oxygen limitation and starvation as well as regulators of numerous stress responses such as the sigma factors *rpoS*, *rpoN* and *rpoE*.

Although *Salmonella* species are generally well-studied, in some instances systems (the starvation and heat shock responses, for example) are better studied in *E. coli*. Some illustrative examples are included here owing to the general similarity of the two organisms, with research conducted in *E. coli* indicated. The two organisms are not identical, however, and examples exist of paralagous proteins performing different functions (Winfield & Groisman 2004). Therefore, some caution must be exercised when drawing direct comparisons between the two species.

Acid tolerance

The pH of stomach acid can reach as low as 1.5, being chiefly maintained by hydrochloric acid but with the additional presence of organic acids such as lactic and acetic acid (Álvarez-Ordóñez et al. 2011). In the food industry, acid is a common preservative (pickles are typically pH 3-4 (Lu et al. 2013)). Indeed, exposure to moderate acid (such as that of fruit juice) during initial growth can help to protect *S.* Typhimurium from subsequent exposure to the stronger acid of the stomach (Yuk & Schneider 2006).

Often, the pathogenicity of an organism (as defined by minimum infective dose (ID)) is directly related to its ability to survive stomach acid. The ID of enteric pathogens correlates directly to their ability to survive acidic conditions: *Vibrio cholerae* (ID 10^9) is more susceptible to acid than *Shigella flexneri* (ID 10^2) (Audia et al. 2001). While the infective dose of nontyphoidal *Salmonella* is typically around 10^5 cells, when ingested with a food which protects against acid, the ID can be as low as 100 cells. This is a common phenomenon in fatty foods (Waterman & Small 1998). Acid tolerance arises from the induction of a number of systems including efflux pumps, chaperones, and repair proteins.

Basal pH in *Salmonella* is maintained by a series of proton efflux pumps which preserve an intracellular pH of 7.6-7.8 (Álvarez-Ordóñez et al. 2011). However, when extracellular pH falls, an additional pair of systems are induced which export protons by creating derivatives of lysine (Morita et al. 2006) and arginine (Kieboom & Abee 2006). Both systems consist of two proteins. The first (CadAB and AdiA for lys and arg respectively) is a decarboxylase which consumes a proton in producing an amino acid derivative (cadaverine from lysine and agmatine from arginine). The second (CadC or AdiC) is an antiporter which exchanges an intracellular derivative for an extracellular amino acid, thereby exporting a proton. Addition of lysine or arginine to growth medium significantly improves acid tolerance of *S.* Typhimurium (Álvarez-Ordóñez et al. 2010).

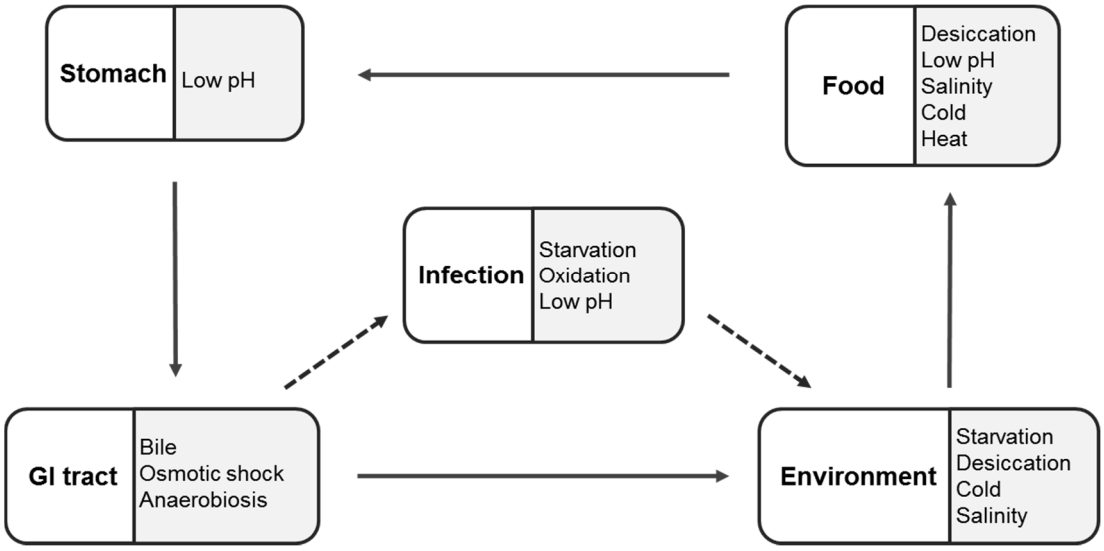


Figure 1.2. *Stages of the* S. *Typhimurium life cycle, and stress factors encountered at each.GI = gastrointestinal.*

Numerous gene regulators are known to coordinate the acid shock response including RpoS (Soo Lee et al. 1995), PhoPQ (Bearson et al. 1998), and Fur (Foster & Hall 1992). The exact tolerance genes expressed are dependent on growth media, pH, and growth phase. It is known that around fifty acid shock genes are induced around pH 3 in *S*. Typhimurium, including several regulators of stress responses such as *rpoS* (Audia et al. 2001). RpoS is an alternative sigma factor which changes the specificity of RNA polymerase to alter transcription. It responds to a wide range of environmental factors, as well as being induced in stationary phase and after a reduction in growth rate. It is estimated that RpoS controls expression of around 10% of the *Salmonella* genome (Battesti et al. 2011). In the acid response, as well as induction of DNA repair proteins, some lipid modifying enzymes are expressed in response to RpoS activity and drive alterations in membrane composition (in favour of saturated and cyclic fatty acids (Kim et al. 2005).

However, many of the genes induced are not essential for acid tolerance, reflecting the fact that as well as being encountered in the stomach, low pH is also encountered inside the host cell and is thus a signal of internalisation. Numerous acid-activated genes, particularly those under PhoPQ control, are virulence genes upregulated in response to the low pH of the *Salmonella*-containing vacuole (Martin-Orozco et al. 2006). Therefore, acid sensing and tolerance facilitate not only survival and passage through the gastrointestinal (GI) tract, but also virulence. The OmpR/EnvZ two-component sensing system is active in response to a range of stresses including the low pH and magnesium starvation in the SCV. As well as activating acid shock genes (Bang et al. 2002), OmpR controls the expression of several genes in the *Salmonella* pathogenicity islands. The supercoiling state of the chromosome likely facilitates OmpR (the transcription factor element of the two-component system) access to both sets of promoters (Quinn et al. 2014).

Osmotic shock and desiccation

Having passed the acidic stomach, enteric pathogens enter the small intestine; an environment usually rich in nutrients and therefore with a high salinity and an osmolar pressure equivalent to 0.3 M NaCl (Sleator & Hill 2002). Osmotic stress is resisted in proteobacteria by the accumulation of a number of low molecular weight compounds whose presence in high concentration is not disruptive to cell functions,

but maintains turgor pressure. Hence, they are termed compatible solutes and include proline and trehalose (Sleator & Hill 2002). A further condition of limited water availability is desiccation, which is an important means of food preservation. *Salmonella ssp.* are known to survive for long periods on dried or powdered food, even at high temperatures due to the protective effect of the solid medium (Finn, Condell, et al. 2013).

The initial bacterial response to a medium of high salt content is the rapid influx of potassium ions, presumably to balance osmotic pressure and prevent lysis (Csonka et al. 1994). It appears that accumulation of K^+ has a dual function, however. A concurrent synthesis of glutamate leads to the formation of potassium glutamate, which also acts (in *E. coli*) as a second messenger to stimulate expression of various genes through RpoS (Lee & Gralla 2004). RpoS activates transcription at promoters containing an osmotic shock element at which an RNA polymerase is constitutively bound but inactive until the binding of potassium glutamate.

Several operons are strongly upregulated during osmotic shock: one producing a potassium import system, with others contributing the accumulation of compatible solutes. These include genes coding for a ProP, transport system for glycine-betaine and proline, and an enzyme involved in trehalose synthesis (Balaji et al. 2005). ProP was initially described as a proline transporter but is now known to also transport glycine-betaine, the major compatible solute in *S.* Typhimurium (Cairney et al. 1985). Two other transporters for the uptake of compatible solvents are known: OsmU (Frossard et al. 2012) and ProU (Stirling et al. 1989). Trehalose (synthesised by OtsAB) prevents damage to proteins and lipids by replacing solvating water molecules. Although induced in the small intestine, *otsAB* are not essential for virulence (Howells et al. 2002).

Regulation of osmotic shock genes involves the alternative sigma factors RpoS and RpoE; Δ*rpoSE* single and double mutants are more sensitive to osmotic shock (McMeechan et al. 2007). An additional level of control may derive from negative supercoiling of the chromosome, which promotes transcription of osmotic shock genes in *E. coli* (Cheung et al. 2003). This effect is possibly mediated by RpoS, which is also regulated by supercoiling. The supercoiling state of chromosomal and plasmid DNA regulates a host of stress response genes in *Salmonella*, and is altered by factors such as microaerobic growth, growth phase, temperature, and pH (Cameron et al. 2011).

Although certain enzymes required for maintenance of supercoiling are known it is not clear whether the conformational changes are a biological response to, or a physical effect of, changing environmental conditions. They may result from a combination of both.

Desiccation is an extreme form of osmotic shock where little or no water is available. Survival of *Salmonella ssp.* after desiccation, particularly in low temperatures, appears to be virtually indefinite; on dried almonds stored at 4°C, there was no reduction in cell count after 550 days (Uesugi et al. 2006). Survival during desiccation is dependent on the same compatible solute transporters as function during osmotic shock, namely ProP, ProU, and OsmU. RpoE is also required, and the transcriptome of desiccated cells reflects a state of energy conservation. Motility genes are downregulated, whilst glucose and fatty acid catabolism are promoted to supply the production of trehalose (Finn, Händler, et al. 2013). On rehydration, the cells return to normal activity. Long-term survival appears to depend on low cellular activity. In peanut oil, *S*. Enteritidis was found to exist in a dormant state in which only 5% of the genome is transcribed (compared to 78% in rich medium) (Deng et al. 2012).

Anaerobiosis

In conditions of oxygen limitation, such as in the lumen of the small intestine, *Salmonella ssp.* is able to grow anaerobically by utilising alternative electron acceptors. The switch between aerobic and anaerobic growth is mediated by a two-component system ArcAB (Evans et al. 2011), which senses extracellular oxygen, and the DNA-binding Fnr (coded by *oxrA* in *Salmonella*), which reponds to intracellular oxygen (Fink et al. 2007). Transcriptomic data from mutants of both systems suggest their function is broadly the same in *S*. Typhimurium as in *E. coli*. Namely, the pair repress genes of oxidative phosphorylation and the TCA cycle whilst stimulating expression of genes for the production and utilisation of lactate. Many (120) genes are regulated in the same manner by both Fnr and ArcAB. (Evans et al. 2011; Fink et al. 2007).

In addition to metabolic genes that could be expected, the regulators of anaerobiosis also control a number of genes involved in motility. This supports the observation that chemotaxis towards electron acceptors enhances growth anaerobically (Rivera-Chávez et al. 2013). The availability and utilisation of electron

acceptors in the gut appears a large determinant of virulence in *S.* Typhimurium. Virulence factors enhance production of tetrathionate by host epithelia, giving *Salmonella* a competitive advantage over resident commensal species unable to utilise the sulfur compound (Winter et al. 2010). Phage-mediated horizontal transfer of genes to enhance production of another alternative electron acceptor, nitrate, in the host has been demonstrated to increase *Salmonella* virulence (Lopez et al. 2012).

Oxidative stress

Macrophages kill most infecting *Salmonella* cells within hours with reactive oxygen and nitrogen species (produced by NADPH oxidase and by inducible nitric oxide synthase). Reactive oxygen makes the greater contribution to initial killing, while reactive nitrogen is largely responsible for long-term limitation of growth (Vazquez-Torres et al. 2000).

Reactive oxygen produced by NADPH oxidase takes the form of the superoxide radical O^-, which dismutates to form hydrogen peroxide H_2O_2. The chief mechanism of H_2O_2 toxicity is its reactivity with iron, which liberates the metal from iron-sulfur clusters in essential enzymes such as those of the electron transport chain (Fang 2011). Reactions with iron or other heavy metals reforms the superoxide form of oxygen, which can react to reform H_2O_2 or interact with DNA directly, causing damage such as single or double stranded breaks or point mutations (Lloyd et al. 1998).

The activity of the nitric oxide radical NO is largely bacteriostatic rather than bactericidal. As NO can also react with iron clusters the electron transport chain is inhibited, leading to growth arrest. DNA synthesis and repair is stalled by a reaction of NO with the catalytic site of ribonucleotide reductase, an enzyme involved in the synthesis of deoxyribonucleotides (Henard & Vazquez-Torres 2011).

Reactive nitrogen and oxygen both lead to activation of the bacterial starvation response, which involves a downregulation of translational apparatus (Bourret et al. 2008). Induction of the starvation response is likely to be a product of damage to amino acid synthases, resulting in low availability of amino acids, and disruption of the TCA cycle (Park et al. 2011; Richardson et al. 2011) DskA is a regulator of the stringent response but also mediates defence against oxidative damage. Metabolic pathways under DskA control produce NADH and NADPH, which supply reducing power to

counteract reactive oxygen (Henard et al. 2010). Similarly, amino acid synthetic pathways under DskA control contribute to detoxification of reactive nitrogen (Henard & Vázquez-Torres 2012).

A pair of *Salmonella* virulence genes prevents the localisation of NADPH oxidase (Vazquez-Torres 2000) and inducible nitric oxide synthase (Chakravortty et al. 2002) to the SCV. In addition, numerous systems are in place to detoxify reactive oxygen and nitrogen in the cell. Low molecular weight thiol compounds (the most abundant being reduced glutathione) scavenge reactive nitrogen or oxygen, yielding various higher products. Constitutively produced, these low molecular weight compounds are the cell's first protection against oxidative or nitrosative stress (Song et al. 2013). *Salmonella* also employs a host of inducible detoxifying enzymes including three catalases and three peroxidases (Hebrard et al. 2009; Horst et al. 2010), all of which contribute to virulence. In response to nitrogen, the regulator NsrR promotes expression of several detoxifying proteins (Karlinsey et al. 2012), the most important of which is Hmp. Hmp is a flavohaemoglobin which reduces NO to NO_3^- with the concurrent oxidation of NADPH (Bang et al. 2006).

Starvation

The infective cycle of *S.* Typhimurium and other foodborne pathogens necessitates of survival outside the host, sometimes for long periods. As well as responses specific to a single starvation condition (such as C, N, P or specific metals) general starvation-induced genes (induced by limited availability of more than one nutrient) can be induced (Foster & Spector 1995). The importance of extra-host survival is such that virulence and antibiotic resistance plasmids are thought to contribute to pathogenesis by also helping adaptation to starvation conditions (Paytubi et al. 2014). Starvation responses also impact directly on infection, where certain micronutrients are limited by the innate immune system. *Salmonella* overcomes zinc sequestration by neutrophils, for instance, by expressing a high-affinity zinc transporter, thus helping it out-compete commensal bacteria in the gut (Liu et al. 2012).

The bacterial starvation response is mediated by a pair of second messenger molecules, guanosine tetraphosphate and pentaphosphate, collectively abbreviated

(p)ppGpp. (p)ppGpp signalling leads to repression of motility and division, as well as reduced synthesis of proteins, ribosomes and phospholipids. Meanwhile, amino acid synthesis, carbon metabolism, and virulence genes are promoted (Kanjee et al. 2012). The essential function is to balance protein synthesis and replication with nutrient availability. Cellular concentrations of the second messenger (p)ppGpp are regulated by a pair of proteins, a synthetase RelA and a synthase-hydrolase SpoT. RelA detects ribosomes stalled due to uncharged tRNA while SpoT is sensitive to fatty acid or carbon starvation, amongst others. They combine to regulate (p)ppGpp levels (Battesti & Bouveret 2006).

(p)ppGpp interacts with RNA polymerase to effect transcriptional change, but can also interact directly with effector proteins to inhibit processes such as DNA replication (in *B. subtilis* (Wang et al. 2007)). DksA is often described as a cofactor of (p)ppGpp binding to RNA polymerase but is also a regulator in its own right; the two function independently as well as in combination (Song et al. 2010). DksA also functions outwith the stringent response, including in the oxidative stress response during *Salmonella* infection (Henard et al. 2010). Given its presence in many pathogens (and evidence suggesting co-regulation of virulence genes) but absence in humans, (p)ppGpp and the stringent response have been suggested as antibiotic targets (Godfrey et al. 2002).

As with many stress responses, the alternative sigma factors RpoE and RpoS are also involved in regulating the starvation response (McMeechan et al. 2007). In addition, the function of housekeeping RpoD sigma factor is repressed by the action of an RNA antagonist (which binds to the RpoD-RNAP complex and inhibits transcription) and an anti-sigma factor Rsd (which sequesters RpoD) (in *E. coli*; (Sharma & Chatterji 2010)). Sigma factor regulons can comprise hundreds of genes and they are an efficient mechanism for the induction of widespread transcriptional change in a cell. RpoS (σ^{38}) in particular plays a role in the regulation of many stress responses, as described above, as well as in virulence (Kazmierczak et al. 2005).

Heat shock

Clearly a major advantage for foodborne pathogens is an ability to survive the high temperatures encountered during cooking and food processing. The principal impediment suffered by bacteria at high temperature is the aggregation of misfolded proteins. As such, effectors of the heat shock response (HSR) include protein folding chaperones (Craig et al. 1994) (GroELS and DnaKJ being the most studied), proteases (Meyer & Baker 2011), and small heat shock proteins which stabilise misfolded proteins before re-folding (Tomoyasu et al. 2003; Veinger et al. 1998).

Although essential at all temperatures, *S*. Typhimurium's GroESL is also upregulated during heat shock due to an RNA thermometer in its mRNA. A hairpin structure occludes the Shine-Dalgarno site at 37°C but is melted at higher temperatures, making it's translation more efficient (Cimdins et al. 2013). A similar secondary structure accounts in part for the increase in translation of *rpoH* (the heat shock sigma factor) at high temperatures in *E. coli* (Morita et al. 1999). The sigma factor is also regulated postranslationally. RpoH is signalled for degradation (by DnaKJ) at 37°C but the binding affinity of the chaperone complex is decreased at higher temperatures due to a structural rearrangement, leading to less proteolysis and higher RpoH availability (Chakraborty et al. 2014).

Although the greater part of the HSR is the preservation of mature proteins, a number of HSR effectors are involved in maintaining synthesis of new proteins, which is destabilised at high temperatures. Numerous aspects of protein synthesis are modulated (in *E. coli*) by heat shock proteins whose functions include increasing ribosome stability, preventing translational stalling, and methionine synthesis (Rasouly & Ron 2009). Another aspect to the heat shock response is the preservation of membrane integrity; 25% of the genes upregulated at high temperature are membrane-bound, including the protease FtsH. A disproportionate number (~60%) of proteins which induce RpoH and the heat shock response are also located in the membrane (Nonaka et al. 2006).

The Bacterial Cold Shock Response

Mesophiles such as *Salmonella* survive over a range of temperatures; in the context of the infective cycle, low temperatures are usually encountered after exit from the host. For foodborne pathogens, survival and replication at low temperatures is of added importance in the context of refrigeration during food preparation and storage. *S.* Typhimurium is able to grow efficiently at 8°C (Smadi et al. 2012); European Union regulations recommend refrigeration temperatures of "8°C or below" (EC regulation 852/2004).

Under laboratory conditions, cold shock is usually induced by a downshift in temperature from 37°C to between 5 and 15°C. In general, two phases of cold adaptation are observed. In the initial acclimation phase, growth stalls while cold induced proteins are preferentially expressed (or their presence is otherwise elevated) and contribute to the adaptation of the cell to replication at the lower temperature. During the subsequent adapted phase, expression of many cold induced proteins falls (but remains elevated compared to growth at 37°C) and growth is able to resume, albeit slower than usual (Barria et al. 2013). Expression of cold shock proteins, other cellular proteins, and cell growth is described in Figure 1.3.

In *E. coli*, around thirty cold induced proteins (CIP) are known, although not all of their functions are clear (they are listed in Table 1.1). They are sometimes grouped into two general classes. Class I CIP are strongly and exclusively produced upon cold shock whereas Class II are expressed during normal growth but upregulated during cold shock (Phadtare et al. 1999). The principal function of CIP are to facilitate translation at low temperatures, although other functions include the maintenance of membrane fluidity, preservation of protein folding, and condensation of the chromosome. Although much of the study of cold shock has been in *E. coli*, *S.* Typhimurium possesses many of the same effectors.

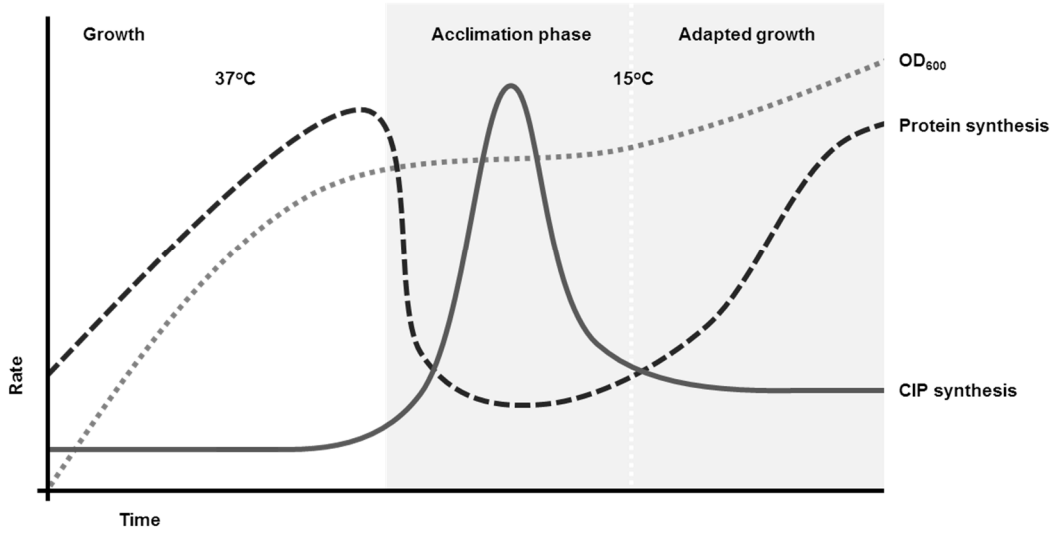


Figure 1.3. *Rate of growth and synthesis of cold induced proteins (CIP) and other cellular proteins before (white) and after cold shock (grey). An example growth curve (OD_{600}) is shown as a dotted line, protein synthesis dashed, CIP synthesis solid. Adapted from Barria et al., 2013.*

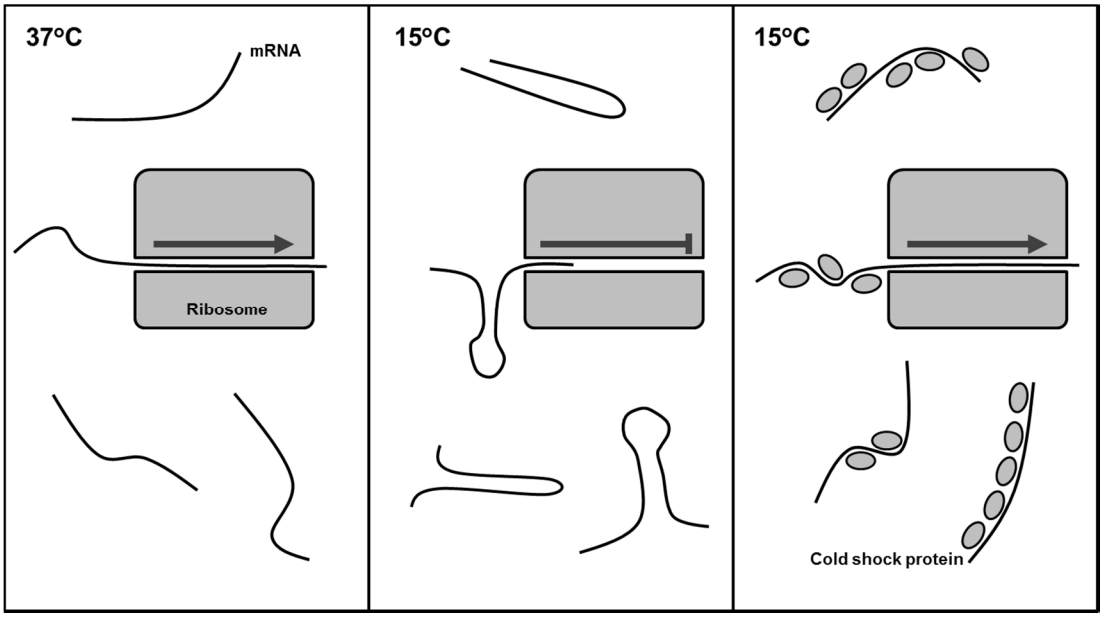


Figure 1.4. *Translational block is relieved by cold induced proteins. At 15°C, mRNA can form stabilised secondary structures which impair translation by occluding Shine-Dalgarno sequences and preventing procession through the ribosome (middle). Cold induced proteins facilitate translation by relieving such secondary structures (right).*

Gene	Function	Reference
aceE	Pyruvate dehydrogenase, decarboxylase	Gualerzi et al. (2003)
aceF	Pyruvate dehydrogenase, dihydrolipoamide acetyltransferase	Gualerzi et al. (2003)
cspA	Cold-inducible RNA chaperone and anti-terminator; transcriptional enhancer	Gualerzi et al. (2003)
cspB	CspA paralogue	Gualerzi et al. (2003)
cspE	RNA chaperone; transcriptional antitermination (CspA paralogue)	Gualerzi et al. (2003)
*cspG**	CspA paralogue	Gualerzi et al. (2003)
*cspI**	CspA paralogue	Gualerzi et al. (2003)
csdA	ATP-dependent RNA helicase, facilitates translation of mRNAs with 5′ secondary structures	Jones et al.(1996); Moll et al.(2002)
dnaA	DNA binding and replication initiator, global transcription regulator	Gualerzi et al. (2003)
gyrA	DNA gyrase, subunit A; DNA binding/cleaving/rejoining subunit of gyrase	Gualerzi et al. (2003)
hns	Nucleoid protein, transcriptional repressor, repressor supercoiling	Gualerzi et al. (2003)
hscA	DnaK-like chaperone	Lelivelt & Kawula (1995)
hscB	DnaJ-like co-chaperone for HscA	Lelivelt & Kawula (1995)
hupB	Nucleoid protein, DNA supercoiling	Giangrossi et al.(2002)
infA	Protein chain initiation factor IF1, translation initiation	Gualerzi et al. (2003)
infB	Protein chain initiation factor IF2, translation initiation, fMet-tRNA binding, protein chaperone	Gualerzi et al. (2003)
infC	Protein chain initiation factor IF3, translation initiation, stimulates mRNA translation	Gualerzi et al. (2003)
lpxP	Lipid A synthesis; cold-inducible	Vorachek-Warren et al. (2002)
nusA	Transcription termination/ antitermination/elongation L factor	Bae et al. (2000)
otsA	Trehalose phosphate synthase; cold- and heat-induced,	Kandror et al.(2002)

otsB	Trehalose phosphate phosphatase; cold- and heat-induced,	Kandror et al.(2002)
pnp	3′–5′ exoribonuclease; component of RNA degradosome; cold shock protein required for growth at low temperatures	Yamanaka & Inouye (2001)
rnr	3′−5′ exonucleases; increases 10-fold in cold shock	Cairrão et al. (2003)
rbfA	Ribosome-binding factor required for efficient processing of 16S rRNA; cold shock adaptation protein	Gualerzi et al. (2003)
recA	General recombination and DNA repair; induction of the SOS response	Gualerzi et al.(2003)
tig	Protein-folding chaperone, multiple stress protein, ribosome-binding	Kandror & Goldberg (1997)
*ves**	Cold- and stress-inducible protein, function unknown	Yamada et al.(2002)
yfiA	Protein Y, associated with 30S ribosomal subunit, inhibits translation	Di Pietro et al. (2013)

Table 1.1. *Known cold-induced proteins of* E. coli*, their gene names and functions. Adapted from Barria et al., 2013. * denotes proteins without a direct homologue in* S. *Typhimurium SL1344. All others listed possess a direct homologue in SL1344.*

Cold shock and the lipid bilayer

At low temperatures, the usually fluid phospholipid cell membranes become more ordered, due to short-range interactions between hydrocarbon chains of the lipids (Cevc 1991). This causes the cell membranes to become harder and more rigid, with many effects on the integrity and function of the membrane and its components. Numerous cold shock genes are induced following membrane gelling and function to restore fluidity by modification of constituent lipids.

The rigidified membrane is detected in *Bacillus subtilis* by a well-studied two-component system DesKR. The membrane-bound domain of the sensor kinase DesK appears to detect the fluidity of the bilayer, leading to phosphorylation of the cognate transcription factor DesR when the membrane rigidifies (Hunger et al. 2004). DesR binds to promoter regions upstream of, amongst others, a desaturase gene *des*. Expression of *des* is repressed by the presence in the membrane of unsaturated fatty acids, suggesting the mechanism of repression once fluidity has been restored (Aguilar et al. 2001). No analogous system has been described in *E. coli* or *Salmonella* (neither of which possess direct homologues of DesKR), but both possess a number of two-component systems. The CpxPAR system (composing an inner membrane-bound kinase CpxA, a response regulator CpxR and a periplasmic accessory protein CpxP) senses a number of membrane and periplasmic stresses, and is a possible candidate for temperature sensing (Hunke et al. 2012). The mechanism of its activation is not yet known.

During normal growth, the inner membrane of salmonellae is composed mostly of phosphatidylethanolamine harbouring lipid groups of 16 or 17 carbons in length (Oyofo et al. 1989). At low temperature, general changes in bacterial membranes include increasing desaturation and cis-trans isomerisation, both of which serve to spatially separate lipid chains (Shivaji & Prakash 2010). In normal growth, the major saturated fatty acid present is palmitate (16:0) while the most abundant unsaturated fatty acids are palmitoleate (16:1) and cis-vaccinate (18:1). After a shift to 15^{o}C the percentage of unsaturated fatty acids in the membrane rises from 20% to 53% (Cronan 1975). In addition, cis-vaccenate is accumulated in place of palmiteoleate. This accumulation is very rapid and apparently independent of transcription or translation (Garwin & Cronan 1980), indicating that the effector is present in the cell even at

higher temperatures. The effector enzyme concerned (β-ketoacyl-ACP synthase, coded by *fabF*) is more active at low temperatures than high (Garwin et al. 1980).

The outer membrane's outer leaflet consists largely of lipid A, upon which lipopolysaccharide is assembled. At low temperatures, the multiple fatty acid chains of lipid A are modified to increase the percentage presence of palmitoleate (16:1) at the expense of laurate (12:0). The change is mediated by an acetyltransferase (LpxP) induced strongly after cold shock (Carty et al. 1999). While this change is not essential for viability at low temperatures (as determined by knockout of *lpxP*), it does contribute to maintaining membrane integrity (Vorachek-Warren et al. 2002). A similar change is observed in *Salmonella* (Wollenweber et al. 1983), which also carries *lpxP*.

Cold shock and protein folding

Being a process driven by thermodynamics, protein folding is affected by low temperatures as much as by high. At low temperatures the structural rearrangements required for refolding occur more slowly and are further impeded due to increased solvent viscosity, slowing the rate of folding. Additionally, the burying of hydrophobic residues is an endothermic process and thus proceeds slower at low temperatures (Piette et al. 2011). At high temperatures (in yeast), protein folding chaperones co-function with compatible solvents to preserve protein integrity (Lee & Goldberg 1998). A similar situation appears to occur at low temperatures, with trehalose synthesis enzymes being induced in *E. coli* (Kandror et al. 2002) and other compatible solvents in *Listeria monocytogenes* (Bayles & Wilkinson 2000). Accumulation of trehalose is essential below 4^{o}C but begins at 16^{o}C, apparently as a precaution in case of further temperature decrease (Kandror et al. 2002). It may also function to preserve membrane fluidity and to protect against the formation of ice crystals at freezing temperatures.

Protein folding chaperones are also induced by cold shock. A pair of chaperones analogous to DnaKJ (named HscA and HscB) are expressed at low temperatures, and appear to have the same function of co-translational protein folding (Lelivelt & Kawula 1995). Trigger factor (TF) is another protein folding chaperone which associates to the ribosome and prevents or reverses the early misfolding of nascent proteins. As well as large N- and C-terminal domains which interact with substrates, TF has a domain with

peptidyl-prolyl isomerase activity (Hoffmann et al. 2012). Specific functional adaptations of TF in psychrophiles suggests that defects in protein folding are different at low and high temperatures (Godin-Roulling et al. 2014). At low temperatures, TF's unfolding activity is reduced in favour of an increased activity in suppression of premature cotranslational folding.

Protein aggregation appears to play a role in sensing temperature and gene regulation also. A DNA-binding gene regulator, TlpA, polymerises in a temperature-dependent manner due to a coiled-coil region which unfolds as temperature increases. Without the functional coiled-coil dimerisation domain, TlpA's DNA-binding ability is lost (Hurme et al. 1997). TlpA is a repressor which becomes inactive at higher temperatures, suggesting that it may contribute to the expression of virulence genes upon entry to the (warmer) host environment. It is expressed from a virulence plasmid (Hurme et al. 1996). Although TlpA may not be related to the cold shock response, the mechanism of temperature-dependent oligomerisation is an interesting one. H-NS controls the expression of a number of stress and virulence genes, and also exhibits temperature-dependent oligomerisation (Ono et al. 2005). H-NS also has a role in chromosome condensation by regulating several genes involved in DNA supercoiling.

Cold shock and DNA topology

It has long been known that transcription in prokaryotes is directly influenced by the supercoiling state of the template DNA (Pruss & Drlica 1989). Supercoiling of genomic and plasmid DNA is maintained by DNA gyrase (which introduces negative supercoils) and DNA topoisomerases (which relax them). The state of supercoiling changes according to environmental conditions such as pH, osmolarity, and temperature (Drlica 1992). In particular, cold shock leads to transiently increased formation of negative supercoils in *E. coli* DNA; the opposite effect is observed during heat shock (Mizushima et al. 1997). The change in DNA topology is observed during the acclimation phase but returns to normal thereafter. This effect is dependent on the DNA gyrase subunit GyrA, which is induced at low temperature (Gualerzi 2003), although not as a direct response to the cell's supercoiling state. Rather, *gyrA* transcription appears to be promoted at low temperatures by a cold shock protein, CspA (Jones et al. 1992).

The transience of supercoiling suggests that it serves a regulatory, rather than a protective, function. Indeed, the specific expression of several stress response pathways is dependent on DNA topology (Aldridge et al. 2013; Weinstein-Fischer et al. 2000). Microarray studies in the cyanobacterium *Synechocystis* have identified specific genes of the cold shock response regulated by supercoiling (Prakash et al. 2009). One such gene is the desaturase *desB*, which acts to restore membrane fluidity (Los 2004). A possible mechanism for the positive effect of supercoiling on gene expression is a phenomenon known as stress-induced duplex destabilisation, in which the induction of supercoils causes tension in the DNA duplex, lowering the energy required to form the open complex of transcription (Hatfield & Benham 2002). In some instances, the association of DNA-binding proteins is also promoted by supercoiling.

The DNA-binding protein HU has two alternative subunits, HupA and HupB, which can form heterodimers or homodimers. HU binds to and preserves supercoiling in bacterial cells, acting in a manner similar to eukaryotic histone proteins (Guo & Adhya 2007). As a supercoiling factor, and in line with the duplex destabilisation theory, HU stimulates bacterial transcription (Morales et al. 2002). HupB is a cold shock protein and at low temperatures HupB homodimers and A/B heterodimers predominate over HupA homodimers (Giangrossi et al. 2002). The functional effect of this change is not clear. The three alternative forms all appear to bind DNA with little specificity and at 37^{o}C HupB is the least important (Prieto et al. 2012). Many species preserve only one HU protein subunit, which is usually HupB, suggesting the major function of HU may be during cold shock.

Cold shock and translation

At low temperatures, translation is impaired for a number of reasons. The principal cause is the tendency of mRNA to form stabilised secondary structures at low temperatures, as described in Figure 1.4. In addition, ribosome maturation is impaired at low temperatures (likely due to a similar effect in rRNA), as is turnover of RNA. Given that chloramphenicol (which inhibits the bacterial ribosome) induces synthesis of cold shock proteins (Jiang et al. 1993), it is likely that the translational block is a signal for the beginning of the cold shock response. A drop in the rate of translation can be directly observed after cold shock where single ribosomes accumulate at the

expense of polysomes, the usual structures seen during active translation (Jones & Inouye 1996).

Secondary structures in mRNA impair translation by prohibiting translocation through the ribosome, and also by masking ribosome binding sites and start codons in double-stranded regions. Given that the efficiency of translational initiation is almost entirely determined by accessibility of the Shine-Dalgarno sequence (Smit & Duin 1990), occlusion of such sites leads to a huge drop in translation rates. It is thus unsurprising that a number of cold induced proteins display RNA chaperone or helicase activity. The cold shock proteins bind to double stranded RNA regions, promote unwinding, and maintain mRNA as single strands by binding along their length (Phadtare 2005). The functions of the cold shock proteins are discussed in detail below. Another protein with some similar activities is the cold shock RNA helicase CsdA, which has been shown to promote availability of the *rpoS* mRNA ribosome binding site (RBS) in *E. coli*. (Resch et al. 2010a).

CsdA also functions in a modified RNA degrading complex at low temperatures (Prud'homme-Généreux et al. 2004), illustrating the importance of RNA turnover in the cold. Again, the formation of secondary structures in RNA inhibits their degradation. RNAse R has helicase as well as RNase activity (Awano et al. 2010), and is highly induced during cold shock. It is the only RNase able to overcome a *csdA* deletion (Awano et al. 2007), highlighting the importance of helicase activity at low temperatures. RNA degradation also helps to regulate the cold shock response, with PNPase being responsible for the selective degradation of cold shock protein mRNA at the end of the acclimation phase (Yamanaka & Inouye 2001b). PNPase is likely important elsewhere in the cold shock response as well, being essential for growth at low temperature (Awano et al. 2008).

RNA folding is also of great importance in the development of functional ribosomes (in *E. coli*). CsdA in particular is known to have a role in ribosome maturation, assisting the association of proteins S1 and S2 to the 30S subunit (Moll et al. 2002) and co-purifying with a 50S subunit precursor (Charollais et al. 2004). The ribosome-binding protein RbfA is required for correct genesis of the 16S rRNA at low temperatures (Xia et al. 2003). It is believed that RbfA aids the maturation of a 16S loop required for full 30S function but which forms defectively at low temperature (Datta et al. 2007). It is suggested that the transcription antiterminator NusA also assists

ribosome development by ensuring the correct fold in nascent rRNA (Bubunenko et al. 2013). It is likely this role, in addition to an increased requirement for antitermination at low temperatures, that accounts for NusA's cold shock induction (Bae et al. 2000). The translation initiation factors IF1, 2, and 3 are all induced in the cold, with IF3 in particular apparently selectively facilitating translation of the cold shock proteins (Gualerzi 2003). IF1 facilitates the dissociation of 70S ribosomes (which are more stable at low temperatures) to increase the pool of free 30S subunits for translation initiation (Giangrossi et al. 2007). In *E. coli*, IF3 appears to perform a similar function whilst introducing a bias towards cold shock mRNAs during the initiation of translation through an unknown mechanism (Giuliodori et al. 2007).

As well as the action of IF3, preferential translation of cold shock mRNAs is accounted for by *cis* and *trans* regulatory elements. *cspA* mRNA is a well-studied example of regulation in cis, and is discussed in detail below. The mRNA forms different structures in and out of cold shock, with the cold shock structure being translated with greater efficiency (Giuliodori et al. 2010). The regulation of the *E. coli* alternative sigma factor RpoS at low temperature is a good example of regulation in *trans*. At low temperature *rpoS* mRNA forms a secondary structure in which the RBS is hidden in a hairpin. A combination of CsdA's helicase activity, the chaperone Hfq, and the sRNA DsrA together alter the secondary structure in such a way as to make the RBS available (Hammerle et al. 2013). As is the case in many stress conditions, RpoS regulates a large number of cold shock genes at the transcriptional level (White-Ziegler et al. 2008).

The Cold Shock Protein Family

Ubiquity of the cold shock proteins

Having been initially identified in *E. coli* as CS 7.4 (Jones et al. 1992), cold shock protein A has given it's name to to an enormous array of cold shock proteins, called the CspA family. CspA of *E. coli* remains one of the best studied members, along with CspB of *Bacillus subtilis*. Paralogues have been found in many other mesophiles (Bresolin et al. 2006; Katzif et al. 2005) and in psychrophiles (Hankins et al. 2007), as might be expected given their function in low temperature adaptation. However, cold shock proteins have also been identified in thermophiles (*Thermotoga maritima* (Welker et al. 1999)) and appear to be present in human commensals which maintain constant temperature (*Bacteroides thetaiotaomicron* (Xu et al. 2003)) as well as species with minimal genomes (*Buchnera aphidicola* (Ham et al. 2003)). The family, indeed, appears near universal.

As well as the occurrence of direct CspA paralogues in Archaea (Giaquinto et al. 2007), plants and animals both possess proteins containing the cold shock domain (CSD), a domain with highly similar structure to the bacterial Csps. In plants, the CSD proteins appear to have similar functions to their bacterial counterparts, being expressed both during low temperature stress and also during normal growth (Chaikam & Karlson 2010). In eukaryotes CSDs are involved in many DNA and RNA processes with their versatility being facilitated by their combination with a wide array of other domains of other functions (Mihailovich et al. 2010).

The sequences and structures of cold shock proteins across bacterial species are highly conserved, as illustrated in Figure 1.5. The number of paralogues present varies greatly across species: *E. coli* K12 has nine (Xia et al. 2001), *B. subtilis* three (Weber, Volkov, et al. 2001) and many species (including the other major foodborne pathogen *Campylobacter* (Gundogdu et al. 2007)) have only one. In species with multiple CspA paralogues, there appears to be a degree of genetic redundancy. In *E. coli* it is necessary to delete four paralogues before a phenotype can be observed (Xia et al. 2001). In a single deletion, the function of the missing paralogue is complemented by those which remain (with the exception of CspD, which appears more diverged in function and sequence). Despite this redundancy, the persistence of multiple copies

and their different expression patterns suggests a specificity and divergence of function in organisms where more than one paralogue is found.

The family's nomenclature and functions do not necessarily correlate across species. For instance, although CspE appears to have the same activity in *E. coli* and *S. enterica*, CspC of *B. subtilis* has different functions to CspC in either species (being closer in expression and activity to *E. coli*'s CspA). Therefore, the bacterial cold shock proteins are best considered as members of a homologous family, rather than as direct homologues of each other. In the following, the functions of the Csp family will be discussed, rather than those of individual members.

Cold shock proteins at low temperature

The principal activity for cold shock proteins is believed to be the melting of RNA secondary structures which are stabilised at low temperatures (Rinnenthal et al. 2010). Deletion of multiple cold shock genes in *E. coli* led to cold-sensitive phenotypes, whilst the deletion of a single Csp caused a complementary overexpression of those that remained. The nine paralogues in *E. coli* exhibit a degree of redundancy; four deletions (of CspA, B, E, and G) are needed to impair cold adaptation, and growth at 15°C can be restored by the expression *in trans* of any single paralogue save CspD (Xia et al. 2001). The observation that the Csp knockouts can be complemented by the nucleotide-binding S1 domain of PNPase (Xia et al. 2001) (which shares a similar structure to the cold shock proteins) supports the idea that RNA chaperoning is their principal role at low temperatures. It is likely that the cold shock response is triggered, at least in part, by translational arrest. Other stresses that induce the same arrest such as starvation (Fraser et al. 2006) and the antimicrobial chloramphenicol (Jiang et al. 1993) both induce CspA expression. Localisation of cold shock proteins to cytosolic spaces near the nucleoid correlates with occurrence of translation (Weber, Volkov, et al. 2001).

However, as well as regulating at the posttranscriptional level, it has been suggested that the cold shock proteins may control the expression of a number of genes directly. A transcriptomic approach identified genes upregulated during the cold shock response and found that deletion of cold shock proteins repressed transcription of genes otherwise expressed during the initial acclimation phase (Phadtare & Inouye

2004). This supports the observation that the Csps are most important during the initial adaptation phase, where their expression is highest (Hankins et al. 2007). There are two mechanisms by which gene regulation at the level of transcription can occur: firstly, by transcription antitermination and secondly by binding at promoter sites to stimulate transcription.

The Csps prevent the termination of transcription by the formation of stem-loops in the nascent RNA (Bae et al. 2000). By a mechanism detailed below, a Csp melts the secondary structures which form in RNA as it is being transcribed, thereby preventing the stalling or termination of transcription. This function serves to upregulate the expression of a gene, and accounts for the observed upregulation of (amongst others) cold-induced genes such as *nusA*, *infB*, *rbfA*, and *pnp*. In *E. coli*, deletions of *cspE* and *cspC* identified a number of genes which, while not cold inducible, required a cold shock protein for basal expression at low temperatures (Phadtare et al. 2006). The method used was transcriptomic, so the requirement for a Csp appears to be at the transcriptional rather than the translational level, suggesting that basal transcription is affected by low temperatures. Presumably, this effect is due to the stabilisation of secondary structures in nascent RNA which the cold shock proteins relieve.

CspA has been shown *in vitro* to bind at promoter sites for two cold-induced genes, *hns* (A. Brandi et al. 1994) and *gyrA* (Jones et al. 1992). In both cases it is assumed that binding serves to promote transcription, although a mechanism is yet to be identified. The promoter region of *gyrA* contains a Y-box sequence which is suggested as a recognition site for CspA, which has a preference *in vitro* for the sequence (Schröder et al. 1995). Despite this assertion, it has been demonstrated that the Y-box is not required for binding to or transcription of several cold shock genes (M. Giangrossi et al. 2001; Goldenberg et al. 1997). The Y-box (ATTGG) is similar to the sequence bound by the eukaryotic Csp homologues Y-box proteins, although it should be considered that the Y-box proteins have other domains which may be responsible for conferring binding specificity. Various preferred DNA and RNA sequences have been described for cold shock proteins *in vitro* (and will be discussed below), but none have been demonstrated to have *in vivo* importance.

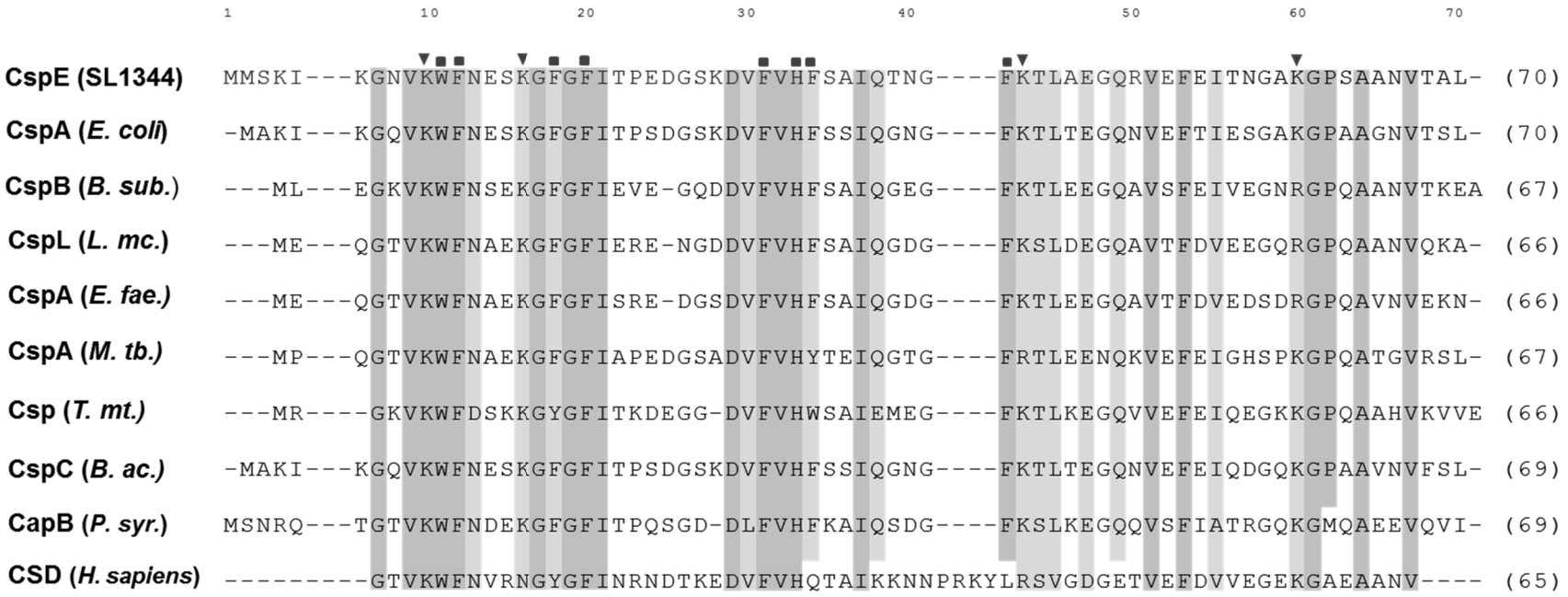


Figure 1.5. *Alignment of cold shock protein homologues in various species:* S. *Typhimurium* SL1344, Escherichia coli, Bacillus subtilis, Listeria monocytogenes, Enterococcus faecalis, Mycobacterium tuberculosis, Thermotoga maritime, Buchnera aphidicola, Pseudomonas syringae, *and a human cold shock domain. Sequences obtained from the Uniprot database and aligned using clustalW. Identical residues are highlighted dark grey, similar residues light grey. Conserved residues believed to be important for polynucleotide binding indicated above: Conserved aromatic residues represented* (■)*, basic residues represented* (▼).

Functions of cold shock proteins at higher temperatures

The role of the cold shock proteins is not limited to low temperature adaptation, despite their name. Numerous Csps are expressed during normal growth, including CspC and CspE at 37°C (in both *E. coli* (Hankins et al. 2007) and *S. enterica* (Hutchinson 2005)). CspA, although initially described as being cold-induced (Jiang et al. 1993) is also highly expressed during exponential phase (Brandi et al. 1999). Moreover Csps are essential during stationary phase in *B. subtilis* (Graumann & Marahiel 1999) and *Caulobacter crescentus* (Balhesteros et al. 2010). Constitutively expressed Csps have been described localised in the nucleoid (Giangrossi et al., 2001), the cytosol (Weber, Volkov, et al. 2001) and at the cell surface (Michaux et al. 2013).

The function of the cold shock proteins at this higher temperature is not as immediately obvious as at lower temperatures. RNA secondary structures are inherently less stable at higher temperatures (Rinnenthal et al. 2010), making the RNA chaperone activity of the Csps less important. One exception is in the heat shock response which, like the cold shock response, is transient and rapidly switched off even if temperatures remain high. One of the mechanisms of swift downregulation relies on CspC (In *E. coli*). Several heat shock mRNAs (including that of the chaperone GroEL) have lower stability at high temperatures, but are stabilised by CspC (Shenhar et al. 2009). A negative feedback loop is responsible for the deactivation of the heat shock response. As GroEL is translated, it specifically targets CspC for protease degredation, resulting in destabilisation of heat shock mRNAs and deactivation of the heat shock response (Lenz & Ron 2014). The interaction of CspC and GroEL during heat shock is described in Figure 1.6.

Transcriptional antitermination has already been discussed. Although constitutive antitermination (ie. facilitating basal transcription) is required less at higher temperatures due to lower frequency of secondary structures, some genes use termination/antitermination as a regulatory mechanism. A number of CspC/E responsive genes in *E. coli* are predicted to contain rho-independent termination sites, suggesting that the mechanism of their regulation is related to antitermination activity (Phadtare et al. 2006).

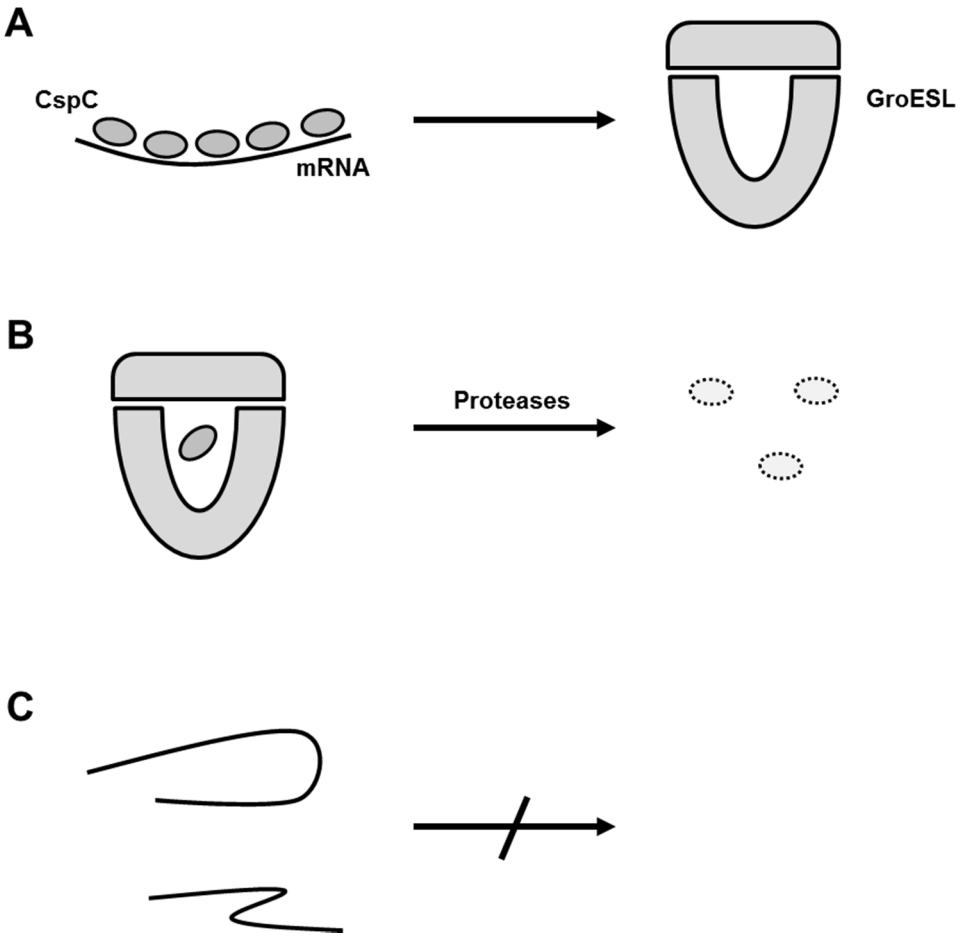


Figure 1.6. *Interaction of CspC and GroESL during heat shock adaptation. At high temperatures mRNA coding for various heat shock proteins, including GroEL, are stabilised by CspC (A). GroEL specifically targets CspC degradation by heat shock proteases, resulting in a drop in CspC levels after heat shock (B). In the absence of CspC heat shock mRNAs are unstable and not translated, resulting in the switching off of the heat shock response.*

A number of genes are known to be regulated by CspE through means other than antitermination. CspE represses the transcription of CspA during growth at 37°C by binding the nascent *cspA* mRNA and promoting transcriptional pause by blocking the passage of RNA polymerase (Bae et al. 1999). Expression of a suite of stress response genes is under CspE regulation through the stationary phase sigma factor RpoS and the stress protein UspA (Hankins et al. 2007). CspE (and/or CspC) has been shown to regulate *rpoS* post-transcriptionally by binding to *rpoS* mRNA and increasing its stability and longevity in the cell (Shenhar et al. 2012). It is believed this stabilisation occurs through a physical protection of the mRNA from degradation by RNAses. CspE is known to bind RNA and inhibit the action of the nucleases polynucleotide phosphorylase and RNAse E (Feng 2001).

The regulation of stress response mediators such as RpoS and UspA implicates the cold shock proteins in the response to a vast array of stress responses including osmotic, oxidative and heat shock as well as carbon starvation (Battesti et al. 2011). In *Listeria monocytogenes*, it was determined that, while all Csp paralogues can function during adaptation to both osmotic and cold shock, CspA predominated during cold and CspD during osmotic shock (Schmid et al. 2009). In addition to being induced by low temperatures, CspA of *Staphylococcus aureus* was linked to a number of phenotypes during normal growth including pigmentation and resistance to both antimicrobials (Duval et al. 2010) and antimicrobial peptides (Katzif et al. 2003). In the same organism, hydrogen peroxide and certain antimicrobials were found to induce expression of CspC to a greater extent than falling temperatures (Chanda et al. 2009).

Several Csp paralogues appear to play a role in chromosome maintenance. CspE and CspC were initially described as recovering a mutant deficient in chromosome segregation during division (Yamanaka et al. 1994). CspE has since been shown to protect cells from chromosomal decondensation by camphor (Hu et al. 1996). It has been shown *in vitro* that *E. coli*'s CspE can both bind ssDNA and dimerise, which is a suggested mechanism for chromosome condensation (Johnston et al. 2006), although this is yet to be demonstrated *in vivo*. In *E. coli* CspD is known to inhibit DNA replication (Yamanaka & Inouye 1997), presumably by binding to ssDNA and preventing the passage of DNA polymerase, although it is more divergent in sequence than the other Csp paralogues. The transcription of *cspD* fits this function; it is low in exponential phase, where cells are growing, and higher in stationary phase and in starvation conditions where less replication is expected (Czapski & Trun 2014)

The requirement for low-temperature expression of the Csps is explicable by virtue of the stabilised RNA secondary structures which occur. Clearly they have regulatory functions during normal growth, however. Our current understanding of the constitutively expressed Csps is fragmented and lacks a unifying feature. It seems unlikely that the Csps are direct and specific regulators of all the pathways and genes described; the array of pathways in which the cold shock proteins appear involved might suggest some common regulatory element yet to be elucidated. The Csps might have a role in a small molecule signalling pathway such as that of (p)ppGpp, which have an enormous range of cellular functions, or a specific protein partner may direct, or be directed by, the Csps. The identified targets of cold shock proteins may possess a common sequence or structural motif, either in their DNA or mRNA. Whichever the case, no such higher regulators have yet been identified.

Cold shock proteins and virulence

The myriad roles of the cold shock proteins link them to virulence in a number of ways. As well as their roles in various stress responses, in foodborne pathogens adaptation to low temperature (ie. during food processing and refrigeration) is essential for survival. It is unsurprising then that cold shock proteins have been linked to virulence and pathogenesis in a number of species. Csps are upregulated during infection in *Vibrio parahaemolyticus* (Nydam et al. 2014) and *L. monocytogenes* (Camejo et al. 2009) and have been linked to virulence in *Enterococcus* (Michaux et al. 2012) and *Brucella* (Wang et al. 2014). Persistence, a major factor in recurrent infections, is influenced by a toxin/antitoxin system through CspD in *E. coli* (Kim & Wood 2010) and by the cold shock proteins in *Salmonella* (Shrimpton 2011).

The infective pathway of *L. monocytogenes* is a particularly good example, and extensively studied with respect to the cold shock proteins. Csp mutant strains were both less able to invade host cells and less able to survive once internalised. The loss of survival in the macrophage is likely explained by a similar weakening of the oxidative stress response (Hankins et al. 2007). The discussed role of Csps in activation of the stress response factor RpoS (which has a regulatory role in the oxidative stress response (Chiang & Schellhorn 2012)) could be the reason for this observation.

In addition, a major haemolysin of *L. monocytogenes* is under the control of Csps; Csp mutants were deficient in haemolysis. The regulatory mechanism was believed to be based in RNA turnover: in the absence of Csps, the half-life of the haemolysin transcript was reduced by two-thirds (Schärer et al. 2013). A number of other proteins involved in RNA turnover have been linked to virulence, including Hfq (Christiansen et al. 2004) and RNAse R (Erova et al. 2008).

Whilst it is unsurprising to find RNA–modulating proteins involved in gene regulation, the case of the cold shock proteins is unique due to their common presence in multiple copies. *L. monocytogenes* has three cold shock proteins (CspA, B, D) but each does not have an equal effect on the virulence phenotypes described (Hankins et al. 2007). If their exclusive function was that of non-specific RNA chaperones, one copy of the gene would be sufficient. In order to regulate gene expression, a non-specific chaperone (with no intrinsic means of switching on or off) must be regulated in turn by another means. For example, CspE upregulates RpoS by preventing the degredation of its mRNA, but unless removed by some other factor, the *rpoS* transcript would never be degraded. Some external controlling factor of CspE is required, either at RNA or protein level, to restore normal cellular function. The means by which this higher regulation of the cold shock proteins is achieved is currently unknown.

Regulation of the cold shock proteins

The first characterisation of the Csp family was as proteins induced at low temperature, although many have since been discovered to respond to other stimuli. The cold shock proteins can be considered in three groups: those expressed exclusively in response to low temperature, those expressed constitutively, and those expressed in response to other stimuli.

Regulation of Csp expression at low temperatures is chiefly at the post-transcriptional level. The mRNA of *E. coli*'s CspA has a long (160 nucleotides) 5'-untranslated region (UTR). The 5'-UTR has two alternate conformations, which are temperature-dependent. At 37^{o}C, both the Shine-Dalgarno sequence and the start codon are hidden in double-stranded regions. At 15^{o}C, when a pseudoknot stabilises an alternative structure, these features are both accessible (Giuliodori et al. 2010). As well as altering translation initiation, an alternative method of regulating expression is

through mRNA stability. The 5'-UTR of CspE also has a role in low temperature activation, despite being much shorter than that of CspA (at 42 nucleotides). In this instance, the 5'UTR contains an RNAse E cleavage site which, although exposed at 37°C, folds into an inaccessible stem-loop at 15°C (Uppal et al. 2008). Neither of these observations can provide an explanation for the induction of Csps under conditions other than cold shock, however. Numerous aspects of Csp regulation outwith cold shock remain unknown.

Transcriptional regulation of the Csps has also been observed. Transcription of *cspD* is promoted by the starvation-responsive second messenger (p)ppGpp in *C. crescentus* (da Silva et al. 2010) and *E. coli* (Yamanaka & Inouye 1997), likely as a response to carbon starvation. cAMP-activated protein (CAP), which responds to energy levels in the cell and regulates carbon metabolism, can also stimulate *cspD* expression (Uppal et al. 2014). Expression of CspE is also linked to energy availability through CAP (Uppal et al. 2011), as is that of CspA through the little-understood second messenger c-di-AMP (L. Zhang et al. 2013). Both CspD (Kim et al. 2010) and CspE (Hu et al. 2011) are under the transcriptional control of toxin/antitoxin systems. In both cases, an antitoxin mRNA binds at the *csp* promoter site and represses transcription.

There is also evidence of post-translational regulation of Csp protein levels. CspD is degraded by Lon protease (itself a regulator of a number of stress responses) in response to growth rate (Langklotz & Narberhaus 2011). Availability of CspD correlates to its function as an inhibitor of DNA replication. It is degraded fastest in exponential phase (where replication of DNA is rapidly occurring) and slowest in stationary phase (where DNA replication is not required). The discussed role of CspC in the heat shock response is regulated by degradation by the protease GroEL. Degradation is specific to CspC; CspE, also expressed constitutively in *E. coli*, does not interact with GroEL (Lenz & Ron 2014).

Cold shock protein structure

A number of cold shock protein structures have been solved by crystallography or NMR. These include the well-studied CspA of *E. coli* (hereafter, *Ec*CspA) (Schindelin et al. 1994), CspB of *B. subtilis* (*Bs*CspB) (Schindelin et al. 1993), and CspE from *S.* Typhimurium (Morgan et al. 2009). All three share a highly similar fold, being composed of five β-strands arranged into a compact β-barrel structure. Disordered loops separate the strands, with the largest being loop 4, located between β3 and β4. The principal structural feature of the small proteins (CspE of *S.* Typhimurium is 70 aa and 7.5 kDa) is a well-conserved nucleotide binding surface composed of basic and aromatic residues. The binding surface is illustrated in Figure 1.7.

This binding surface resembles the well characterised oligonucleotide/oligosaccharide binding (OB) fold found in many prokaryotic and eukaryotic proteins, usually as a domain in a larger quaternary structure (Bochkarev & Bochkareva 2004). Similarly the cold shock protein structure has given its name to the cold shock domain, found in plants and higher eukaryotes as well as bacteria. In eukaryotes, the cold shock domain is largely found between two other domains in structures called the Y-box binding proteins. The other domains guide polynucleotide binding and mediate interactions with other proteins. (Chaikam & Karlson 2010).

The quaternary structure of the cold shock proteins remains unclear. *Bs*CspB was purified and crystallised as both a dimer and a monomer, with the dimerisation interface believed to be at the opposite face to the nucleotide binding surface (Schindelin et al. 1993). This dimerisation surface is illustrated in Figure 1.6. Curiously, *Ec*CspA (62% identical to *Bs*CspB) purified only as a monomer (Schindelin et al. 1994). CspE (66% identical to *Bs*CspB and 81% identical to *Ec*CspA) was also found to dimerise during the purification process, but this dimerisation was observed to occur at the nucleotide binding face (Morgan et al. 2009). Given the importance of nucleotide binding to all known cellular functions of the cold shock proteins, it seems unlikely that a dimer thus formed would be functional *in vivo*. Csp proteins from *B. subtilis* have been crystallised binding polynucleotides in both dimeric (Max et al. 2007) and monomeric (Hankins et al. 2007) forms, perhaps illustrating the vagaries of crystallography. *In vivo* evidence for the dimerisation of any cold shock proteins is yet to be found.

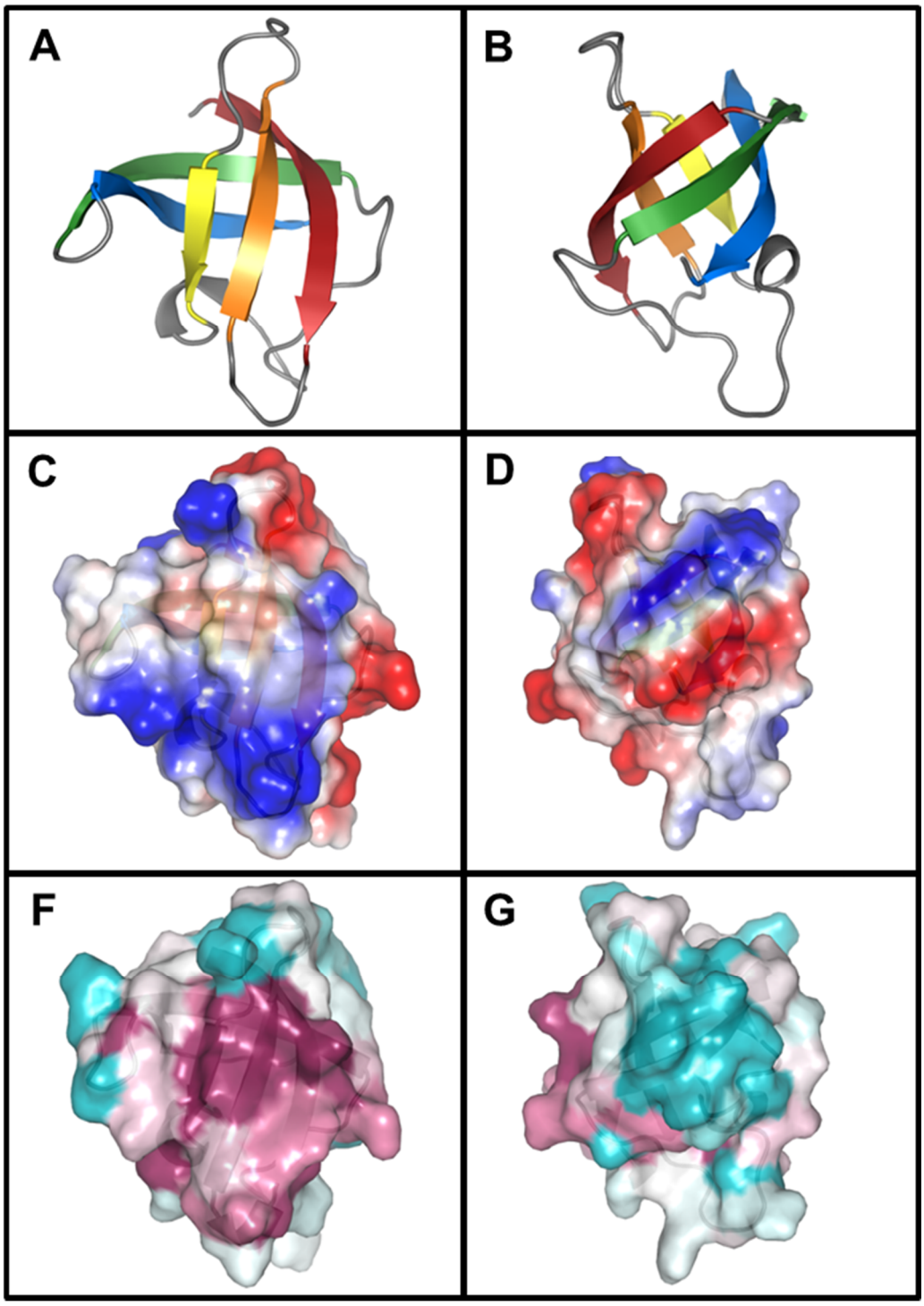


Figure 1.7. *Crystal structures of CspE from* Salmonella *Typhimurium.A and B show the fold of CspE's 5 β strands. Below are molecular surfaces showing the basic polynucleotide binding surface (C) and the proposed dimerisation interface on the opposite surface (D).C and D show basic residues blue and acidic residues red. F and G surfaces are coloured by conservation (based on the aligned paralogues shown in Fig 1.5), showing high conservation (purple) in the RNA-binding surface and more variability (blue) in the opposite surface. PDB ID: 3I2Z*

Cold shock protein interactions with polynucleotides

The initial characterisation of cold shock protein (Csp) function was as RNA chaperones capable of binding to and melting secondary structures in RNA (Jiang et al. 1997). Binding to ssDNA has also been widely reported *in vitro* (Johnston et al. 2006; Zeeb et al. 2006) and may be a function of the apparently nonspecific nature of Csp interaction with polynucleotide substrates. Although various binding preferences have been reported (for polyT and polyC ssDNA (Lopez & Makhatadze 2000), for the cold shock promoter sequence ATTGG (Graumann et al. 1997; Morgan et al. 2009), or for AT rich regions (Phadtare & Inouye 1999)), *in vivo* evidence for substrate specificity is lacking. Similarly, there is little *in vitro* or *in vivo* evidence for sequence specificity in RNA binding.

Numerous amino acid residues have been identified as important for nucleotide binding, largely centred around two motifs termed RNP1 (W11, F18, F20) and RNP2 (F31, H33, F34) (RNP = ribonucleoprotein site). Residue numbers given refer to *S.* Typhimurium CspE: other paralogues exhibit different numbering, but the same structures. These motifs contain aromatic amino acids which are believed to intercalate with bases of DNA or RNA (Phadtare & Severinov 2010). Various other aromatic residues have been identified by sequential mutagenesis (Schröder et al. 1995) and NMR mapping (Zeeb & Balbach 2003). In addition, a number of residues with positively charged side chains have been shown to be important and are believed to be responsible for an initial electrostatic attraction which is supported by the intercalating aromatic residues (Phadtare et al. 2002). This two-stage binding, in which initial backbone contacts are made by basic residues and reinforced by stacking of aromatics, is similar to that of the homologous OB-fold proteins discussed above (Bochkarev & Bochkareva 2004). Residues known to be involved in polynucleotide binding are highlighted in Figure 1.5. The nucleotide binding surface of CspE is illustrated in Figure 1.8

The processes of binding to a polynucleotide and melting secondary structures appear to be discreet, however. The mechanism of melting stem-loops has been studied using *E. coli* CspE. The sequence of events involved have been delineated by mutation of F18 and F31, which produce mutants able to bind RNA, but not to remove secondary structures (Phadtare et al. 2002). It appears that the mechanism of melting involves first the intercalation of F18 and F31 which allow the subsequent intercalation of H33,

the residue which completes the unwinding (Phadtare et al. 2004). Melting is energy-independent and is initiated by the binding of a single CspE monomer to the junction between single- and double-stranded substrate (ie. the junction of stem and loop, or the base of a stem). Subsequent cooperative binding of further CspE proteins on the single stranded region is believed to apply a strain to the stem and force unwinding (Phadtare 2005). The minimum length of the single-stranded section of substrate for unwinding is 4 nt; this is likely reflective of the fact that a single Csp binds to 6 nt (Lopez et al. 1999). A substrate with a single-stranded overhang of 4 nt can be unwound by a single CspE monomer.

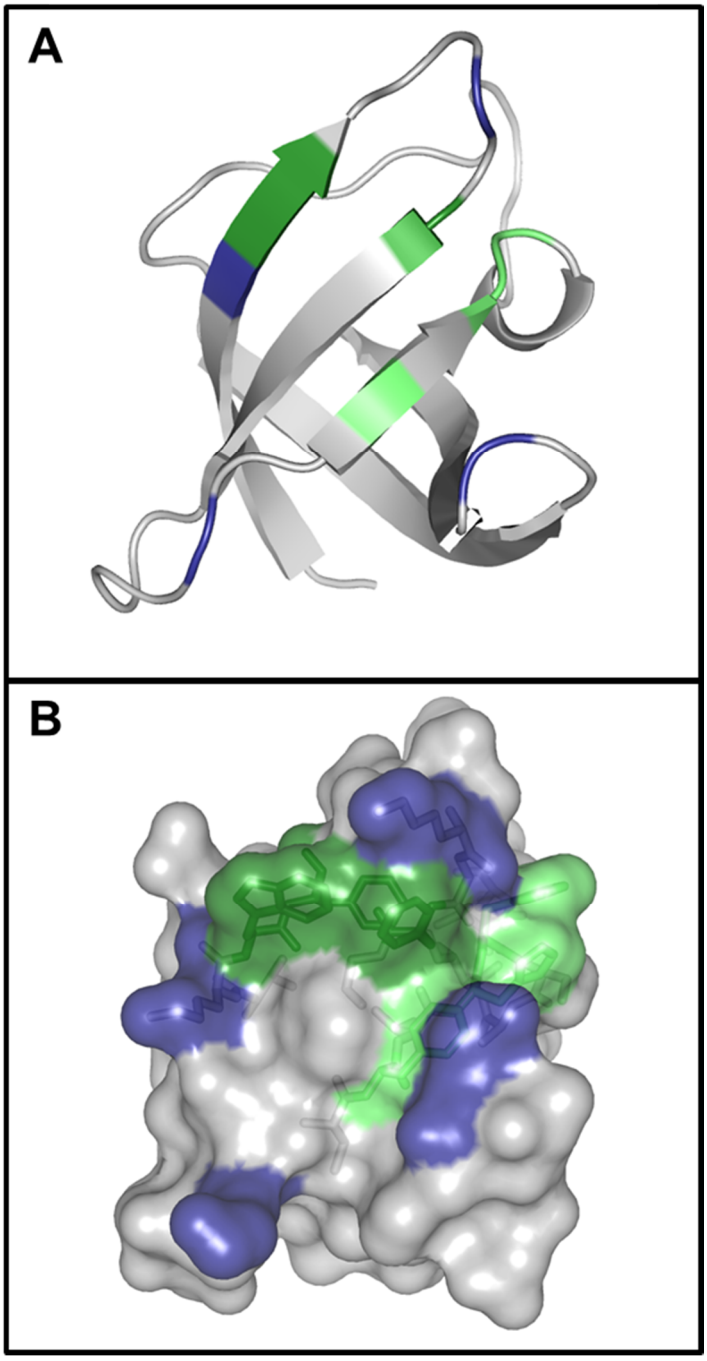


Figure 1.8. *Polynucleotide binding surface of CspE shown as a ribbon (A) and molecular surface, with key residues shown below (B). Intercalating aromatic residues of the RNP 1 and 2 sites are shown dark and light green respectively. Positively charged lysine residues, which make initial contact with the nucleotide backbone, shown blue. Polynucleotide chain binds diagonally, approximately top right to bottom left.PDB ID: 3I2Z*

Cold shock proteins of *Salmonella* Typhimurium

S. enterica Typhimurium possesses six cold shock proteins named (after their homologues in *E. coli*) CspA, B, C, D, E, and H (Kröger et al. 2012). An additional cold shock protein is found in a strain of *S. enterica* Enteritidis. It is coded for by a small plasmid associated with phage resistance, although the cold shock protein ORF is not required for the resistance phenotype (Rychlik et al. 2001). It is considerably less conserved than the chromosomal copies, with a highest percentage identity of 31.8% with CspD although both RNP sites are conserved. The function of this additional Csp is unknown.

Table 1.2 describes the protein sequence identities of the six Csps in *S.* Typhimurium. As might be expected, the two constitutively expressed paralogues CspE and CspC are the most similar at 84.1% identity, amounting to 12 divergent amino acids between the two. The most divergent is CspH, the least characterised paralogue. Also shown are the identities the *S.* Typhimurium Csps share with their *E. coli* counterparts. With the exception of CspB, the *E. coli* and *S.* Typhimurium paralogues are highly similar, and two (CspA and CspC) are identical. This is useful for the study of the family in *Salmonella*, as it is likely that the proteins share the same functions in both species. *E. coli* and *S.* Typhimurium CspE have two residues which differ: *S.* Typhimurium's has an additional methionine at the N-terminus, and a nonconservative substitution near the C-terminus (I68T). Figure 1.9 shows the computed evolutionary divergence of the CspA paralogues in *S.* Typhimurium.

	CspA	**CspB**	**CspC**	**CspD**	**CspE**	**CspH**
CspA	100.0 (to *E. coli*)	65.2	69.7	50.0	71.2	47.0
CspB	65.2	67.1 (to *E. coli*)	69.6	52.2	71.4	45.7
CspC	69.7	69.56	100.0 (to *E. coli*)	47.8	84.1	43.5
CspD	50.0	52.2	47.8	91.8 (to *E. coli*)	47.8	29.9
CspE	71.2	71.4	84.1	47.8	98.6 (to *E. coli*)	48.6
CspH	47.0	45.7	43.5	29.9	48.6	81.4 (to *E. coli*)

Table 1.2. *Protein sequence identity matrix of* S. *Typhimurium cold shock proteins. Highest identity is between the two constitutively expressed homologues, CspC and CspE. CspD and CspH are the most diverged from the remainder. Shown darker is each homologue's identity to their counterpart in* E. coli*. All have high identity with their* E. coli *paralogue.*

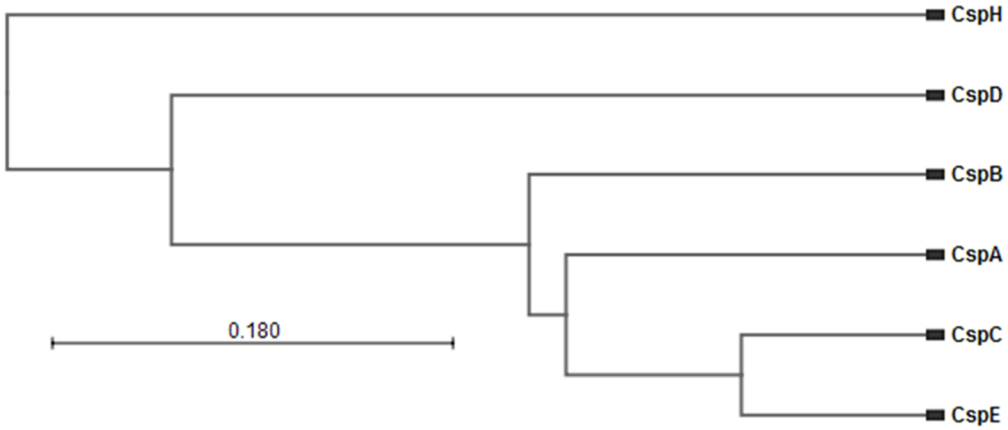


Figure 1.9. *Unrooted phylogram showing conservation of the 6 Csp proteins of S. Typhimurium SL1344. Distance is reflective of difference between protein sequences: CspH is the most diverged of the 6, with CspC and CspE the most similar. Scale bar represents distance for 0.18 substitutions per residue (total width is equivalent to the number of substitutions between CspH and CspE (22)). Phylogram drawn using CLC sequence viewer 7.0.*

Cold-induced paralogues

As discussed, expression of CspA is strongly induced by cold shock (Hankins et al. 2007) and in exponential phase (Brandi et al. 1999). Translation (but not transcription) of CspA is increased in conditions of high nutrient availability (Yamanaka & Inouye 2001a). Although the mechanism for this nutrient sensitivity is unknown, cold shock induction is likely to be largely due to the preferential translation of the *cspA* mRNA at low temperatures (Giuliodori et al. 2010). The induction of cold shock proteins at low temperatures appears to be linked to a block in translation, as translation-inhibiting antimicrobials also produce elevated Csp translation (Etchegaray & Inouye 1999). It has been suggested that translational block also accounts for CspA induction in exponential phase and other conditions which limit translation (Brandi & Pon 2012). The regulation of *cspA* mRNA by its thermosensing 5'-UTR has already been discussed. Additionally, it is thought that *cspA* is more sensitive to RNAse E degradation at higher temperatures, reducing its expression outside of cold shock (Fang et al. 1997).

Given its extreme induction at low temperature and other translation-limiting conditions, it is likely that the primary function of CspA is to facilitate translation by resolving RNA secondary structures. However, CspA also contributes to the cold shock response by a direct upregulation of at least two cold shock genes, *hns* (A Brandi et al. 1994) and *gyrA* (Jones et al. 1992) at the transcriptional level. It is not known whether the effect observed is the result of antiterminator activity or interaction with the promoter sequences.

CspB is the second cold-induced paralogue in *S.* Typhimurium, although it appears to have a slightly different pattern of expression. It was found in *E. coli* that while CspA is expressed strongly between 24 and 10°C, CspB was present only below 20°C, and maximally at 15°C (Etchegaray et al. 1996). It should be noted, however, that CspB is the paralogue with most divergence between *E. coli* and *S.* Typhimurium (Table 1.2), and therefore may not share the same function in both. In *S.* Typhimurium CspB accumulates below 20°C and maximally at 10°C. Stability of the *cspB* transcript appears to be much greater at lower temperatures, suggesting that an alternative low-temperature conformation protects it from RNAse degradation in the cold (Craig et al. 1998). The function of CspB has not been independently studied; it is widely assumed to be similar to that of CspA.

Constitutively expressed paralogues

There are two Csp paralogues expressed constitutively in *S.* Typhimurium. CspC and CspE appear to share many of the same functions, and a single knockout rarely produces a phenotype. For this reason, much of the research done into them has grouped them together (Hutchinson 2005; Hankins et al. 2007; Phadtare et al. 2006). It is therefore difficult to separate their functions, and likely that many of them are the same. There are differences, however. Only CspC is involved in the heat shock response (Lenz & Ron 2014) whilst only CspE has been linked to chromosome condensation (Hu et al. 1996). They do not appear to restore cold growth to a *csp* null mutant to the same degree (Hutchinson 2005). This specificity of function arises despite the fact that both CspC and CspE share similar patterns of expression both at 37 and 15°C (Czapski & Trun 2014), have the same substrate (ssDNA or RNA), and are only 12 amino acids different in primary sequence (of which differences, 9 are conservative).

In *E.coli*, CspE transcription is constant through log and stationary phase but is induced after a cold shock, especially in minimal medium (Czapski & Trun 2014). CspC follows a similar pattern, although is less expressed in stationary phase. *cspE* mRNA stability is greater at low temperatures due to the occlusion of an RNAse E substrate site by an alternative low temperature conformation (Uppal et al. 2008). CspE is therefore one of the most ubiquitous of Csps, functioning both in and out of cold shock. As well as the roles in translation and transcription antitermination discussed, CspC and CspE have a role in many stress responses by regulating RpoS, the stationary phase sigma factor (Shenhar et al. 2009). The *rpoS* mRNA secondary structure on which CspE acts is conserved in *S.* Typhimurium.

in vitro, CspE has been reported to bind ssDNA and dimerise (Johnston et al. 2006); this is a suggested mechanism of chromosome condensation, either at low or high temperatures. Uniquely among the *S.* Typhimurium Csps, *cspE* is expressed as part of an operon with *pagP* (formerly *crcA*) and *crcB*. The three genes appear to be co-operatively involved in chromosome condensation (Hu et al. 1996), although the mechanism remains unknown. PagP is an outer membrane protein involved in lipid modification (Bishop 2005) whilst CrcB is predicted to be a fluoride exporter (Baker et al. 2012).

Both CspC and CspE are upregulated slightly after cold shock (Holden 2000), and regulate a suite of genes at low temperatures. Whilst many such genes can respond

to any of the Csp paralogues, a small number appear to be specific to CspC and E (Phadtare et al. 2006).

Nutrient-sensitive paralogues

CspD is functionally the most divergent of the *S*. Typhimurium paralogues, being largely described (in *E. coli*) as an inhibitor of DNA replication rather than as an RNA chaperone (Yamanaka & Inouye 1997). Accordingly, it is found in the nucleoid (Giangrossi et al., 2001). Although one of the most diverged (Figure 1.4) CspD does preserve the nucleotide-binding features described above, with only a single conservative mutation (F33Y) in an RNP site. Despite this, it is the only *E. coli* paralogue incapable of rescuing a quadruple deletion strain in the cold, suggesting it is functionally different to the others (Xia et al. 2001). In *S*. Typhimurium, CspD is upregulated in starvation conditions (Holden 2000), which likely correlates to its function in slowing cell growth.

Owing to its role in inhibiting DNA replication, CspD is sometimes referred to as a toxin gene (Kim et al. 2010; Uppal et al. 2014) and indeed overexpression is lethal (Yamanaka & Inouye 1997). It appears to be regulated at transcription by various signalling molecules including c-di-AMP (Uppal et al. 2014) and (p)ppGpp (Yamanaka & Inouye 1997) and by a protease in a growth-phase dependent manner (Langklotz & Narberhaus 2011). CspD, along with a toxin-antitoxin system, were found to be involved in the formation of persister cells (Kim & Wood 2010).

CspH is upregulated in response to nutrient abundance (Holden 2000), apparently by a mechanism involving DNA gyrase and the DNA-binding protein Fis (Kim et al. 2001). Little is known about its function, although it lacks many of the features believed to be important in nucleotide binding. Of the RNP sites, only one aromatic residue (H32) is conserved, with the remainder undergoing non-conservative mutations. This suggests that if CspH does bind polynucleotides, the mechanism is different from that described for the other Csp paralogues. Its presence in numerous species, however, might suggest that there is a role for the cold shock proteins other than nucleotide binding.

Protein-Protein interactions of CspE

Clearly the functions of the cold shock proteins are not fully understood, especially during normal growth. A particular curiosity is the presence of multiple cold shock proteins in many species, and their diversity of function which is maintained by expression patterns despite the fact that most share the same basic properties. Many species express multiple copies in situations where only a single copy is required, such as CspC and E during stationary phase at 37°C. Despite this, different functions are observed for CspC and CspE.

The mechanism by which this specificity of function is achieved remains unclear. It seems unlikely to be structural, given the high conservation of the cold shock protein paralogues and the polynucleotide binding site in particular. The polynucleotide binding and melting site appears to have little inherent specificity, yet it has been shown that the *in vivo* RNA substrates differ between the different paralogues (McGibbon 2013).

If specificity of function is not due to the intrinsic properties of each paralogue, it must be the result of some other factor acting to direct their activity. It is possible that the varied functions of the cold shock proteins are derived from, or mediated by, other proteins. The eukaryotic Csp orthologues are part of multidomain proteins, with other domains guiding specificity. This suggests some further mediation of their interactions is required. There has been little study of the Csp proteins *in vivo*: much of what is known of their function is based on phenotypes of mutants or studies of RNA. Large datasets have tended to be transcriptomic rather than proteomic (Czapski & Trun 2014; Phadtare et al. 2006), despite the known role of posttranscriptional regulation on activity of the cold shock proteins.

Of the proteins themselves, much is known of their structures and polynucleotide binding properties thanks to extensive *in vitro* study. *in vivo*, the localisations of CspA and CspD have been identified (Giangrossi et al., 2001) and homologues in *B. subtilis* have been found concurrent with newly synthesised mRNA (Weber, Volkov, et al. 2001). Direct evidence of protein-protein interactions between a cold shock protein and another also arose in *B. subtilis*, in which CspB and C were found to interact with a pair of cold-induced DNA helicases (Hunger et al. 2006),

supporting the possibility that protein-protein interactions are important for Csp function.

The protein interactions of any protein are indicative of its localisation and function. In the case of CspE, these properties are assumed but not yet confirmed. Many aspects of CspE's function remain unknown: the origin of substrate selection, posttranslational regulation, and the mechanism (or existence) of a function in chromosome condensation. Fundamental information such as whether the protein dimerises *in vivo* remains unknown. Identification of the interacting partners of CspE would reveal details of its function at 37^{o}C, how substrate specificity is conferred, and perhaps the means by which the Csps regulate (and are regulated by) cellular processes.

In addition to basic knowledge of cell function, identified interactions are potential targets for antimicrobial drugs. The cold shock proteins are known to be required for pathogenesis in some species and are essential in others. In *S.* Typhimurium, Csp mutants are deficient in pathogenesis (MPG; unpublished data) and persistence (Shrimpton, 2011). CspE in particular appears to contribute to virulence. Therefore, artificial inhibition of CspE's function may prove an effective means of limiting infection. Although the cold shock proteins are not a traditional drug target (lacking an active site or binding cleft) any identified protein-protein interactions of CspE would be a possible avenue for antimicrobial development.

Protein-protein interactions as drug targets

Traditionally protein-protein interactions (PPI) were considered difficult to design small molecule inhibitors for, in comparison to enzyme active sites. This is because the interfaces tend to be large and flat, giving few pockets or grooves in which a small molecule could bind (Wells & McClendon 2007). PPI are generally dependent on two properties of an interacting surface: hydrophobicity and electrostatic potential. Exposed hydrophobic residues are entropically unfavourable, and result in a large release of free energy when buried. Hydrophobic residues tend to be highly concentrated in patches which are mutually buried during interaction (Lijnzaad & Argos 1997). Electrostatic attractions between complimentary charged residues on

interacting surfaces contribute to an interaction, which are often stabilised by hydrogen bonding (Xu et al. 1997).

Traditional drug design has focused on targeting sites which bind endogenous small molecules such as enzyme active sites or nucleotide exchange sites (Hopkins & Groom 2002). Consider, for example, the antimicrobial β-lactams, which target cell wall synthetic enzymes (Bayles 2000) or tetracycline, which prevents the access of aminoacyl tRNA to the ribosome (Semenkov et al. 1982). The endogenous small molecule ligands of such proteins provide a model upon which an inhibitory drug can be based.

However, advances in crystallography and structural biology have led to growing appreciation of interaction "hot spots"; interacting residues on the surface of proteins which, although covering a small area, account for a significant amount of the free energy change involved in the interaction (Moreira et al. 2007). The most common residues found in hot spots are tryptophan, arginine, and tyrosine (Bogan & Thorn 1998). Tyrosine and tryptophan are both aromatic and contribute to hydrophobic burying, whilst arginine's three amino groups can form multiple hydrogen bonds. Hot spots occupy small surface areas and are often found in clusters (Keskin et al. 2005).

Whilst an entire interface would be difficult to target with a small molecule, hot spots offer a potential drug target. Once the structure of a hot spot is known, various computational methods can be used to identify small molecules or peptides that might inhibit an interaction. One computational method examined five cancer-associated proteins which had proven difficult to design drugs against by traditional means, and found fourteen potential inhibition sites, along with small molecule inhibitors of interest (Seco et al. 2009).

One example in bacteria is the identification of inhibitors of the ZipA-FtsZ interaction which precedes cell division. Identification of three hydrophobic residues which formed an interaction hot spot on ZipA (Mosyak et al. 2000) allowed the surface to be targeted with a library of small molecules, whose binding affinities were examined by NMR, a fluorescence-based binding assay, and crystallography (Tsao et al. 2006). A number of compounds were found which bound to ZipA's interacting hot spot, with their structures resembling the combined side chains of the complementary

interacting residues of FtsZ. Although their affinities were lower than optimal for therapeutic potential, the method is an exciting one.

More success has arisen in the study of various oncogenic pathways, where a number of compounds have been identified which effectively kill lymphoma cells (Bruncko et al. 2007). Interestingly, the possibilities of PPI inhibitors are not limited to disruption of a hydrophobic binding surface. TNF-α is a trimeric cytokine protein and the target of a number of drugs treating rheumatoid arthritis. Where most current drugs target TNF-α either indirectly (through other pathway enzymes) or through unknown mechanisms, one potential drug has been shown to bind and disrupt trimer formation, instead promoting formation of a non-functional dimer (He et al. 2005). An alternative mechanism was also demonstrated for the inactivation of nitric oxide synthase (which is linked to various autoimmune diseases). A small molecule bound at the active site of the monomer allosterically inhibits the formation of active dimers (McMillan et al. 2000).

Networks of protein-protein interactions typically classify proteins as either highly-connected (hub) or lowly-connected. A phenomenon known as the centrality-lethality rule suggests that hub proteins, which affect a wide variety of other proteins and processes, make the best potential drug targets (Hankins et al. 2007). An added advantage of this approach is that, being central to cell function, these proteins are less likely to mutate in such a way as to rapidly produce drug resistance.

Possible interactions of CspE

The structure of CspE (Morgan et al. 2009) reveals a number of possible sites at which it could form intermolecular interactions. The most striking may be the large disordered loop composed of residues 34-47. Disordered proteins or domains are prone to interactions due to a number of factors. Disordered regions display structural elasticity, enabling them to interact with numerous targets, and allowing rearrangement during interaction, or binding-induced folding into an ordered state (Tompa 2002). CspE's major loop contains five hydrophobic residues and one with positive charge, any of which could be potential sites for interactions. It is easy to imagine a mechanism in which burying of the hydrophobic residues is supported by a hydrogen bond formed by the positive residue (K42). CspE also has a surface-exposed

hydrophobic patch composed of residues I4, V29, and I54 which may be a second possible site of interaction.

The cold shock proteins, as outlined above, have many roles within the cell and are connected (directly or indirectly) with a wide range of processes including several associated with virulence. Inhibition of Csps, as has been demonstrated in various knockout studies (Graumann et al. 1997; Hankins et al. 2007; Phadtare et al. 2006) has wide-ranging consequences and is often lethal. Although their redundancy means that any inhibited individual is likely to be complemented by another paralogue, their structural similarity makes any identified drug unlikely to be specific to a single family member. The protein-protein interaction of cold shock proteins are therefore of potential interest as drug targets.

Aims of this project

Although their function during cold shock is well understood, the roles of cold shock proteins outside of that stress condition are unknown. Elucidation of the protein-protein interactions of CspE should expand the current understanding of the function and regulation of the cold shock proteins during normal growth. In addition, any identified interactions might be suitable targets for further study with the aim of drug design.

The present study, therefore, aimed to identify the interactions of CspE. Initial experiments were based in proteomics, and involved *in vivo* cross-linking followed by mass spectrometry in order to determine which proteins CspE associates with in the cell. This method was aimed at identifying proteins which complex with CspE. A subsequent approach was genomic in nature, and more suited to detection of transient or weak interactions. It employed a two-hybrid system to explore CspE's interaction with random peptides derived from genomic DNA fragments. In addition, the formation of dimers or oligomers by CspE was assayed *in vivo* for the first time, providing insights into the function of the protein.

Chapter 2

Materials and Methods

Bacterial Strains and Media

Strain	Genotype	Origin
Salmonella enterica serovar Typhimurium SL1344	*his*$^-$	(Hoiseth & Stocker 1981)
Escherichia coli DH5α	F$^-$ Φ80 *lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK–, mK+) phoA supE44 λ– thi-1 gyrA96 relA1*	(Taylor et al. 1993)
Escherichia coli DHM1	F- *cya-854 recA1 endA1 gyrA96 (nal^R) thi1 hsdR17 Δ spoT1 rfbD1 glnV44*	(Karimova et al., 1998)
S. Typhimurium SL1344 cspE-HTP	*his*$^-$ cspE-HTP cat (cam^R)	(Woodall 2010)

Table 2.1. *Bacterial strains used in this study, their genotype, and origin.*

All strains (see Table 2.1) were grown, unless otherwise stated, in lysogeny broth (LB) broth comprised of 10 g L^{-1} bacto-tryptone, 10 g L^{-1} NaCl, and 5 g L^{-1} yeast extract. For solid media agar was added to 1.5%. Routine growth of liquid cultures was performed by inoculating 5 mL of LB broth and incubating overnight at 37^{o}C with shaking at 200 RPM. Growth on solid media used 20 mL of LB-agar in 9 cm vented petri dishes, which were incubated at 37^{o}C after inoculation. Appropriate antibiotics were added to solid and liquid media as per Table 2.2. All supplements were sterilised by passage through a 0.22 μm filter. For long term storage of strains, glycerol was added to overnight cultures to 10% before storage at -80^{o}C. For carbon utilisation assays, MacConkey agar (Oxoid) was supplemented with appropriate antibiotics and maltose to 1%.

Additive	**Stock Concentration (solvent)**	**Working Concentration**
Carbenecillin (amp)	50 mg mL^{-1} (water)	50 µg mL^{-1}
Kanamycin (kan)	50 mg mL^{-1} (water)	50 µg mL^{-1}
Chloramphenicol (cam)	34 mg mL^{-1} (ethanol)	34 µg mL^{-1}
Tetracycline (tet)	5 mg mL^{-1} (ethanol)	5 µg mL^{-1}
Nalidixic acid (nal)	30 mg mL^{-1} (0.3 M NaOH)	30 µg mL^{-1}
IPTG	100 mg mL^{-1} (water)	100 µg mL^{-1}
X-gal	40 mg mL^{-1} (DMSO)	40 µg mL^{-1}

Table 2.2. Antibiotics and other additives to growth media. IPTG = Isopropyl β-D-1-thiogalactopyranoside. X-gal = 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside. DMSO = dimethyl sulfoxide.

Plasmids and Primers

Name	Function	Features	Source
pTOF24-cspE-HTP	Addition of HTP purification tag to chromosomal *cspE*	Temperature-dependent integration into chromosome (cam^R)	(Woodall 2010)
pUT18	Bacterial two hybrid experiment	T18 domain of *cya*. MCS at N-terminus	(Karimova et al. 1998)
pUT18C	Bacterial two hybrid experiment	T18 domain of *cya*. MCS at C-terminus	(Karimova et al. 1998)
pKNT25	Bacterial two hybrid experiment	T25 domain of *cya*. MCS at N-terminus	(Karimova et al. 1998)
pKT25	Bacterial two hybrid experiment	T25 domain of *cya*. MCS at C-terminus	(Karimova et al. 1998)
pUT18C-zip	Bacterial two hybrid experiment positive control	T18 domain fused to GCN4 leucine zipper motif	(Karimova et al. 1998)
pKT25-zip	Bacterial two hybrid experiment positive control	T25 domain fused to GCN4 leucine zipper motif	(Karimova et al. 1998)
pKT25-*cspE*	Bacterial two hybrid experiment	T25 domain fused to CspE's N-terminus	This study
pKNT25-*cspE*	Bacterial two hybrid experiment	T25 domain fused to CspE's C-terminus	This study
pUT18C-*cspE*	Bacterial two hybrid experiment	T18 domain fused to CspE's N-terminus	This study

pUT18-*cspE*	Bacterial two hybrid experiment	T18 domain fused to CspE's C-terminus	This study
pUT18C-*hupA*	Bacterial two hybrid experiment	T18 domain fused to HupA's N-terminus	This study
pUT18-*hupA*	Bacterial two hybrid experiment	T18 domain fused to HupA's C-terminus	This study
pUT18C-*csdA*	Bacterial two hybrid experiment	T18 domain fused to CsdA's N-terminus	This study
pUT18-*csdA*	Bacterial two hybrid experiment	T18 domain fused to CsdA's C-terminus	This study

Table 2.3 Plasmids used in this study, their functions and origins. GNC4 = *Saccharomyces cerevisiae* transcription factor GCN4

Primer Name	Sequence 5' - 3'	Function
cspE BACTH F	AAAAAAGGATCCCATGTCTAA GATTAAAGGTAACG	Amplification of *cspE* for ligation into BACTH vectors
cspE BACTH-stop R	TTTTTTGAATTCGCTTACAGA GCAGTTACG	Amplification of *cspE* for ligation into pKT25/pUT18C
cspE BACTH-no stop R	TTTTTTGAATTCGCCAGAGCA GTTACGTTTGC	Amplification of *cspE* for ligation into pKNT25/pUT18
hupA BACTH F	AAAAAAGGATCCCATGAACAA GACTCAACTGATT	Amplification of *hupA* for ligation into BACTH vectors
hupA BACTH-stop R	TTTTTTGAATTCTTACTTAACT GCGTCTTTCAG	Amplification of *hupA* for ligation into pKT25/pUT18C
hupA BACTH-no stop R	TTTTTTGAATTCGCCTTAACT GCGTCTTTCAGAG	Amplification of *hupA* for ligation into pKNT25/ pUT18
csdA BACTH F	AAAAAAGGATCCCATGGCTG AATTCGAAACCAC	Amplification of *csdA* for ligation into BACTH vectors
csdA BACTH-stop R	TTTTTTGAGCTCTTACGCATC ACCGCCGAA	Amplification of *csdA* for ligation into pKT25/pUT18C
csdA BACTH F-no stop R	TTTTTTGAGCTCCGCATCACC GCCGAAACG	Amplification of *csdA* for ligation into pKNT25/pUT18
pKT25 confirm F	GCGGATATCGACATGTTC	Amplification of pKT25 MCS
pKT25 confirm R	ATTAAGTTGGGTAACGCC	Amplification of pKT25 MCS
pKNT25 confirm F	GGCTCGTATGTTGTGTGGAA	Amplification of pKNT25 MCS

pKNT25 confirm R	ATAGTCAAGCCGCTCTTTCG	Amplification of pKNT25 MCS
pUT18 confirm F	ATGCTTCCGGCTCGTATGTT	Amplification of pUT18 MCS
pUT18 confirm R	CCGTATGCACGGTTCTCGTA	Amplification of pUT18 MCS
pUT18C confirm F	ATGTACTGGAAACGGTGC	Amplification of pUT18C MCS
pUT18C confirm R	GCAGATTGTACTGAGAGTGC	Amplification of pUT18C MCS
pKT25 sequencing	CAGTCACGACGTTGTAAAAC	Sequencing pKT25 constructs
pKNT25 sequencing	ATGTTGTGTGGAATTGTGAG	Sequencing pKNT25 constructs
pUT18 sequencing F	TTACACTTTATGCTTCCGGC	Sequencing pUT18 constructs
pUT18 sequencing R	CGATATTCATGTCGCCGTCG	Sequencing pUT18 constructs
pUT18C sequencing F	GGCATCAGAGCAGATTGTAC	Sequencing pUT18C constructs
pUT18C sequencing R	CGGACGTTCGAAGTTCTCG	Sequencing pUT18C constructs

Table 2.4. Primers used in this study, their sequences and applications. MCS = multiple cloning site

Methods

DNA Handling

Specific DNA fragments were amplified by the polymerase chain reaction (PCR) using Phusion® polymerase purchased from New England Biolabs (NEB) and the following reaction mixture: 10 µL buffer (as supplied), 1.5 µL dimethyl sulfoxide (DMSO), 1 µL dNTP mix (10 mM dATP, dTTP, dCTP, dGTP (Fisher Scientific)), 1 µL forward and reverse primer (10 µM stock), 1 µL template DNA (prepared plasmid or genomic DNA), 0.5 µL Phusion polymerase. Water was added to 50 µL and PCRs were performed in 25 µL reactions. For colony PCR, 1 µL water was substituted for DNA template. The following standard PCR program was used: 5 minute denaturation at 98°C followed by 34 cycles of 30 seconds annealing (56°C), elongation of 15 seconds per 0.5 kb product length (72°C) and 30 second denaturation at 98°C. Final elongation was 5 minutes at 72°C. PCR products were stored at 4°C overnight or -20°C for longer storage.

Plasmids were prepared from overnight cultures using standard protocols for mini- or midi-prep kits from Qiagen and stored at -20°C in water. Genomic DNA was prepared using Purelink genomic DNA mini kit from Invitrogen and stored at -20°C in water.

Restriction digests were performed using enzymes and buffers as supplied, incubated with DNA at 37°C unless otherwise stated. Before ligation, plasmids were dephosphorylated by addition of 1 µL antarctic phosphatase and 2 µL buffer and incubation at 37°C for 30 minutes. Restriction enzymes and phosphatase were either inactivated by heating to 65°C for 20 minutes or removed from the reaction using a PCR cleanup kit (Qiagen). Unless otherwise stated, ligations were performed using a 3:1 ratio of vector to insert with 1 µL T4 DNA ligase and 1 µL buffer either for 1 hour at room temperature or at 16°C overnight. All enzymes and buffers were purchased from New England Biolabs.

For transformations of SL1344 and *E. coli*, cells were made electrocompetent by growing to exponential phase by diluting 1 mL of overnight culture into 100 mL of LB (supplemented with antibiotics as appropriate) and incubating at 37°C with shaking at 200 RPM until an OD_{600} of 0.4 – 0.5 was achieved. Cells were collected and washed

three times (using a Beckman J2-21 centrifuge (rotor JA 14) at 12,000 RPM (22,000 xg) for 15 minutes) in ice-cold 10 % glycerol before being resuspended in the same and snap frozen in a dry ice and ethanol bath in 50 µL aliquots. Electroporations were performed using a Biorad Gene Pulser (200 Ω, 3.0 µFD, 1.5 kV, and a pulse of 0.5 seconds). Transformed cells were recovered in 1 mL super optimal broth with catabolite repression (SOC) medium (20.0 g L^{-1} tryptone, 5.0 g L^{-1} yeast extract, 4.8 g L^{-1} $MgSO_4$, 3.6 g L^{-1} glucose, 0.5 g L^{-1} NaCl, 0.2 g L^{-1} KCl) for 1 hour in a water bath at 37°C before being plated onto LB with appropriate antibiotics.

DNA was visualised on agarose gels made at 0.8 % for fragments 1 kb or greater and 1.0 % for fragments less than 1 kb. Unless otherwise stated, 5 µL of PCR product or 2 µL of plasmid was added to 2 µL of 6x loading dye (NEB) and run alongside 1 µL of the appropriate size markers (NEB). Gels were poured with 0.005% (0.5 µL per 10mL) SYBR® Safe gel stain (Life Technologies) and visualised under ultraviolet light.

Bioinformatics

Alignments were performed using ClustalW2 (Larkin et al. 2007) with the default settings and slow alignment speed. The Basic Local Alignment Search Tool (BLAST) (Altschul et al. 1990) was used to identify DNA and protein sequences. Images of protein structures were created using PyMol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC). CLC Viewer 7 was used to draw phylograms (by neighbour joining with the Jukes-Cantor distance measure and 100 bootstrapping replicates). Calculations were performed and graphs drawn using MS Excel 2010. DNA sequences were visualised and plasmid maps drawn using SnapGene Viewer.

Cross Linking Mass Spectrometry

Cross-linking Time Course

SL1344 containing the *cspE*-HTP fusion was grown to exponential phase in LB-cam (OD_{600} = 0.45) and formaldehyde was added directly to the growth medium to 1% v/v. The mixture was incubated for up to 30 minutes, after which cells were collected at 14,000 RPM (18,000 xg) in a Philip Harris 1-15 benchtop centrifuge. Crude lysates were prepared by mixing cells with 50 µL SDS (sodium dodecyl sulfate) loading dye (Invitrogen) and boiling for 10 minutes. 10 µL of each time point sample was loaded onto an SDS-PAGE (polyacrylamide gel electrophoresis) gel for visualisation.

One-dimensional gel electrophoresis of proteins was performed in gels measuring 70 x 85 x 0.5 mm with a final separating gel concentration of 12.5% (w/v) acrylamide and a stacking gel concentration of 4% (w/v) acrylamide. Separating gel was buffered to pH 8.8 with 1.5 M Tris-HCl. Stacking gel was buffered to pH 6.8 with 0.5 M Tris-HCl. Gels were run in a BioRad Mini Protean II gel tank for 45 minutes at 180 V; the running buffer was 15 g L^{-1} Tris, 72 g L^{-1} glycine, 5 g L^{-1} SDS. Coomassie staining was performed by agitating gels for 20 minutes with staining solution (40% methanol, 10% acetic acid, 0.1% Commassie Brilliant Blue) before transforming to destaining solution (40% methanol, 10% acetic acid) and agitating until bands were clearly visible. Prestained protein ladders were purchased from New England Biolabs.

After electrophoresis, proteins were transferred to nitrocellulose (Whatman Protran 0.45µm pore size) using a Trans-blot semi-dry transblotter (BioRad) at 0.2 A for 90 minutes in a transfer buffer containing 3 g L^{-1} Tris, 14.4 g L^{-1} glycine and 20% (v/v) methanol. Blots prepared thus were placed in blocking solution (1% Marvel milk powder, 0.05% Tween-20 in PBS) for 30 minutes or overnight. Blots were washed three times in PBS-Tween (0.05% Tween-20) before agitating with antibody solution (1% Marvel milk powder, 0.05% Tween-20 in PBS) for 2 hours. Antibody concentration for peroxidase-conjugated anti-peroxidase (Sigma) was 1 in 10,000. Following two washes with PBS-Tween, blots were imaged using Immobilon Western Chemiluminescent HRP substrate detection kit (Millipore). Luminescence was detected using a dark box and camera (SRX-101A; Konica) and the software LabWorks 4.

Protein-Protein Interactions of CspE

Cell Preparation and Lysis

Using the vector pTOF24-*cspE*-HTP an *S.* Typhimurium SL1344 strain was previously made in which the native CspE is tagged with a histidine tandem purification tag (Woodall 2010). This strain was grown in 4.5 L LB-cam at 42°C, with shaking, until OD_{600} of 0.45 was reached. Formaldehyde was added to a final concentration of 1% and cells were incubated for 5 minutes before being collected in a Beckman J2-21 centrifuge (rotor JA 14) at 12000 RPM (22,000 xg) for 20 minutes at 4°C. The supernatant was discarded and the pellets washed twice in 200 ml PBS.

Pelleted cell were lysed by addition of three volumes of 0.5 mm zirconia-silica beads (Biospec) and one volume of lysis buffer (50 mM Tris-HCl pH 7.8, 150 mM NaCl, 5 mM β-mercaptoethanol, 5 µM pepstatin, 5 µM leupeptin, 5 µM 4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF)). Cells were vortexed for 1 minute and placed on ice for 1 minute, alternating, for 10 minutes. Centrifugation at 4,000 RPM (2115 xg) in an MSE Mistral 1000 collected beads, and the supernatant was decanted into microfuge tubes for a second centrifugation at 14,000 RPM (18,000 xg) in a Philip Harris 1-15 benchtop centrifuge to collect cell debris. The supernatant lysate was collected for purification of CspE-HTP complexes.

HTP Purification

Cell lysate (1 mL) was loaded onto 150 µL IgG sepharose resin (GE Healthcare) equilibrated with 5 mL lysis buffer and washed twice with 750 µL TST (50 mM Tris-HCl, 150 mM NaCl, 0.05% Tween-20, pH 7.6) and twice with 750 µL 0.5 M CH_3COOH (pH 3.4). Cross-linked CspE complexes were eluted with homemade GST-TEV (Tobacco etch virus) protease, kindly supplied by the lab of David Tollervey (Granneman et al. 2009). 20 U TEV protease was incubated on the column in 250 µL equilibration buffer for 4 hours at 18°C and overnight at 4 °C.

TEV elutate (250 µL) was tumbled with 50 µL pre-equilibrated Ni-NTA magnetic agarose beads (Quiagen) for 1 hour at 4°C. Beads were washed twice with 200 µL wash buffer I (50 mM Tris-HCl (pH 7.8), 300 mM NaCl, 10 mM imidazole, 5 mM β –mercaptoethanol, 0.1% Tween-20) and twice with 200 µL wash buffer II (50 mM Tris-HCl (pH 7.8), 50 mM NaCl, 10mM imidazole, 5 mM β –mercaptoethanol, 0.1%

Tween-20). Cross-linked complexes were eluted in 50 µL elution buffer (50 mM Tris-HCl (pH 7.8), 50 mM NaCl, 300mM imidazole, 5 mM β –mercaptoethanol, 0.1 % Tween-20). Visualisation of purification steps was performed by SDS-PAGE and Western blot as described above. Antibiody concentration of peroxidase-conjugated anti-his_6 (Santa Cruz) was 1 in 1000.

Preparation of Protein Complexes for Mass Spectrometry

HTP eluate was run briefly on a precast SDS-PAGE gel (NuPage® 4-12% Bis-Tris (Invitrogen) with NuPage® MES SDS running buffer (Invitrogen) and stained with Coomassie blue as before. A band containing the entire eluate was excised and subjected to in-gel trypsin digestion as per standard protocols (Shevchenko et al. 1996). Briefly, the gel was dehydrated with acetonitrile before proteins were reduced (10 mM dithiothreitol, 30 minutes at 37°C), alkylated (55 mM iodoacetamide, 20 minutes room tempurature in the dark), and digested with 13 ng μL^{-1} trypsin (proteomics grade, Sigma) overnight at 37°C.

Digestion media was acidified with 0.1% trifluoroacetic acid (TFA) and added to StageTips which were used to prepare samples for mass spectrometric analysis as described previously (Rappsilber et al. 2003). Peptides were eluted from StageTips in 20 µL 80% acetonitrile 0.1% TFA and concentrated to 4 µL (Eppendorf Concentrator 5301). Samples were diluted to 5 µL with 0.1% TFA for analysis by liquid chromatography tandem mass spectrometry (LC MS/MS).

Mass Spectrometric Analysis of Peptides

LC MS/MS was performed using a Velos LTQ-Orbitrap mass spectrometer (Thermo Scientific) coupled on-line to a NanoAQUITY UPLC (Waters). The analytical column was injected with with a self-assembled particle frit as reported in (Ishihama et al. 2002). A spray emitter (75 µm ID, 8 µm opening, 300 mm length) was packed with C18 material (ReproSil-pur C18-AQ 3 µm; Dr Maisch, GmbH) using an air pressure pump (Proxeon Biosystems).

Mobile phase A was 0.1% formic acid in water. Mobile phase B was 0.1% formic acid in acetonitrile; the two were run in a gradient over 90 minutes. Peptides were

loaded onto the column at a rate of 0.6 µL min^{-1} and eluted at a rate of 0.3 µL min^{-1} according to the following gradient: 1 – 5% buffer B, 1 min. 6 – 32% buffer B, 79 min. 33 – 76% buffer B, 11 min. 77 – 85% buffer B, 1 min. Fourier transform mass spectra were recorded at 60,000 resolution and the twenty most intense peaks were selected in the ion trap for MS2. The ion trap used normal scan and wideband activation to fill 5 x 10^5 ions for MS1, and 1 x 10^4 ions for MS2. Maximum fill time was 100 ms and dynamic exclusion was for 60 s.

Mass searches were performed against a database of *S*. Typhimurium SL1344 protein sequences (Kröger et al. 2012) using MASCOT v2.5. Search parameters were: MS accuracy 6ppm, MS/MS accuracy 0.6 Da, enzyme trypsin, allowed number of missed cleavages 2, fixed modification carbamidomethylation on cysteine, variable modification oxidation on methionine. Proteins found twice or more in the experimental (cross-linked) samples but not in the control (untreated) samples were deemed significant.

Bacterial Two Hybrid System

Construction of Two-Hybrid System

The bacterial adenylate cyclase two-hybrid (BACTH) test kit was purchased from Euromedex and follows the work described in (Karimova et al. 1998). The kit comprises a host strain, DHM1, and plasmids described in Table 2.3. Plasmids were transformed into, maintained in, and prepared from *E. coli* DH5α as previously described. A positive control strain was made by transforming electrocompetent DHM1 with pKT25-zip and pUT18C-zip. For a negative control, DHM1 was transformed with the empty plasmids pKT25 and pUT18C.

cspE was cloned into all four BACTH vectors and *hupA* was cloned into pUT18 and pUT18C as described previously. *cspE* was PCR amplified using primer *cspE BACTH F* paired with either *cspE BACTH-stop R* (including *cspE's* stop codon: for cloning into pKT25, pUT18C) or *cspE BACTH-no stop R* (with *cspE's* stop codon removed: for cloning into pKNT25, pUT18). *hupA* was amplified using *hupA BACTH F*

and either *hupA BACTH-stop R* (cloning into pUT18C) or *hupA BACTH-no stop R* (cloning into pUT18). In all instances, PCR fragments and were digested with EcoR1 and BamH1 at 37°C for one hour while plasmid DNA was similarly digested and dephosphorylated as described. Insert and vector were ligated and transformed into DH5α by electroporation followed by recovery and selection on kan (pKT25 and pKNT25) or amp (pUT18 and pUT18C).

To confirm correct constructs, 600 ng plasmid DNA was sent to DNA Sequencing and Services (Dundee University) along with sequencing primers (3.2 μM) described in Table 2.4. Sanger sequencing returned reads of 900-1300 bp which were used to confirm insertion, orientation, and reading frame of constructs.

Interaction assays

To perform two-hybrid tests, the plasmids were co-transformed into DHM1 in twelve combinations as shown in Table 2.5. Transformants were picked and streaked onto LB-agar supplemented with 40 μg mL^{-1} chromogenic lactose analogue X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside), 0.5 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside), nal, amp, and kan. Positive and negative controls were streaked concurrently, and all were incubated at 37°C overnight, or until blue colouration had developed in the positive control.

β-galactosidase activity was quantified by a colourimetric assay based on hydrolysis of ortho-Nitrophenyl-β-galactoside (ONPG) (Sambrook et al. 1989). Cells were grown to exponential phase (OD_{600} ~ 0.4) or stationary phase (overnight culture) as previously described. Cells were collected by centrifugation for 10 minutes at 14,000 RPM (18,000 xg) in a Philip Harris 1-15 benchtop centrifuge. Supernatant was discarded and cells were resuspended in Z buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1mM $MgSO_4$) so as to concentrate cells 5x. At this point, optical density was taken for subsequent calculations. 30 μL chloroform and 30 μL 0.1% SDS were added to cell suspension, which was then incubated at 37°C and 200 RPM for 30 minutes to permeabilise cells.

100 μL permeabilised cell suspension was added to 900 μL Z buffer and incubated at 28°C for 5 minutes. 200 μL 0.1% ONPG in phosphate buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4), pre-warmed to 28°C, was added to each reaction. Once a

yellow colour developed the reaction was stopped by addition of 500 µL 1 M Na_2CO_3. OD_{420} was recorded, as well as OD_{550} and the duration of the reaction in minutes. Assays were performed in triplicate, and for each sample three replicates were measured and averaged.

Strain	T25 construct (plasmid)	T18 construct (plasmid)
ENHC	T25-CspE (pKT25)	HupA-T18 (pUT18)
ENHN	T25-CspE (pKT25)	T18-HupA (pUT18C)
ENEC	T25-CspE (pKT25)	CspE-T18 (pUT18)
ENEC	T25-CspE (pKT25)	T18-CspE (pUT18C)
ECHC	CspE-T25 (pKNT25)	HupA-T18 (pUT18)
ECHN	CspE-T25 (pKNT25)	T18-HupA (pUT18C)
ECEC	CspE-T25 (pKNT25)	CspE-T18 (pUT18)
ECEN	CspE-T25 (pKNT25)	T18-CspE (pUT18C)
ENDC	T25-CspE (pKT25)	CsdA-T18 (pUT18)
ENDN	T25-CspE (pKT25)	T18-CsdA (pUT18C)
ECDC	CspE-T25 (pKNT25)	CsdA-T18 (pUT18)
EDDN	CspE-T25 (pKNT25)	T18-CsdA (pUT18C)

Table 2.5. *Names of, protein components of, and expression vectors contained in strains to assay interactions by bacterial two-hybrid.*

Genomic Library Preparation

A pre-existing *cya*::*tetR* mutation was transduced from LT2 into SL1344 to provide the source of genomic DNA for library generation. *cya*::*tet* LT2 was grown overnight in 10 mL LB-tet, to which culture 100 µL p22 lysate (~ 10^7 pfu) was added

before a further overnight incubation. The resultant lysate was clarified by centrifugation at 4000 RPM (2115 xg) in an MSE Mistral 1000. 200 µL chloroform was added to the supernatant and left at room temperature for 1 hour in order to lyse remaining cells. Chloroform and cell debris were collected by a repeat of the centrifugation step, and the chloroform incubation was repeated.

A transduction reaction was set up containing 200 µL early stationary phase SL1344 (OD_{600} ~ 0.6), 45 µL LB, and 5 µL phage lysate. The transduction was incubated at 37°C, with shaking, for 30 minutes. Cells were collected and washed twice by centrifugation (14,000 RPM (18,000 xg) in a Philip Harris 1-15 benchtop centrifuge) in 1 mL LB-10 mM trisodium citrate. Transductants were identified by plating the reaction onto LB-tet, 10 mM trisodium citrate. In order to confirm transduction of the *cya*::*tet* mutation, cells were subcultured onto MacConkey-maltose agar.

Genomic DNA was prepared from 10 mL SL1344 *cya*::*tet* overnight culture (LB-tet) using an Invitrogen PureLink® genomic DNA mini kit according to the manufacturer's instructions. Genomic DNA was eluted with two 100 µL volumes of water. gDNA was digested by Sau3a (New England Biolabs) in a volume 100 µL at 37°C for 15 minutes or at room temperature for 5 minutes.

Digests were run on a 0.8% agarose gel in order to gel extract fragments of either 500-1500 or 1500-3000 bp. Gel extractions were performed using the Gel Extraction Spin Kit (Qiagen); fragments were eluted in 30 µL water. Vector pUT18C was digested with BamH1 (New England Biolabs) at 37°C for 30 minutes. Both vector and insert were quantified using a NanoDrop 2000 (Thermo) to allow a 3:1 ratio of insert : vector to be used for ligations. Ligations were incubated at 15°C overnight.

2 µL aliquots of ligation mix were transformed into 50 µL electrocompetent DHM1 + pKT25-*cspE*, made electrocompetent as described previously. 950 µL SOC medium was added and transformants were incubated at 37°C for 1 hour before plating onto LB-agar supplemented with amp, kan, nal, IPTG and X-gal. 900 µL was spread onto a single 150 mm Petri dish using glass beads. To the remaining 100 µL, 400 µL SOC medium was added to facilitate spreading on a second 150 mm Petri dish. Plates were incubated overnight, or until colonies began to turn blue, at 37°C. Colony counts were taken from 100 µL plates and multiplied by 10 to give approximate colony numbers.

Genomic Library Screen

Colonies were picked and screened by PCR (primers *pUT18C confirm F*, *pUT18C confirm R*) to determine size range of inserts and number of colonies containing empty pUT18C. Colonies developing blue pigment were subcultured twice. Where pigmentation persisted, cells were grown in LB-amp, kan, nal overnight and the plasmid mixture (pKT25-*cspE* and pUT18C containing an unknown fragment) was isolated by miniprep. Plasmid mixture was transformed into electrocompetent DHM1 as before. Transformants were split in two and plated onto LB-amp, kan or LB-amp alone. From transformants on the LB-amp plates overnights were set up to miniprep, thereby isolating the pUT18C plasmid alone. This was transformed again into fresh DHM1 with empty pKT25. DHM1 containing pKT25-*cspE*, pUT18C-unknown and pKT25 + pUT18C-unknown were streaked onto LB-agar + carb, kan, nal, IPTG and X-gal to assay for interaction by colour change.

Colonies which appeared to have a significant and CspE-specific interaction were assayed for β-galactosidase activity (as previously described) in cells containing the unknown pUT18C construct and either empty pKT25 or pKT25-*cspE*. Gram staining was performed to confirm absence of contaminating species. Gram stains were performed by heat-fixing liquid cultures to slides and staining for 1 minute sequentially with crystal violet, iodine solution, and basic fuschin. Interacting protein were identified by sequencing of the pUT18C vector containing the unknown insert.

Chapter 3

Proteomic Investigation of CspE Interactions

Proteomic Investigation of CspE Interactions

The cold shock proteins are well conserved across many species, and the six homologues present in *Salmonella* Typhimurium SL1344 have a high sequence identity of up to 84% (between CspE and CspC). Despite this the six homologues have different functions. They respond to different environmental stimuli, and each appears to bind a different set of RNA substrates. It is possible, therefore, that these activities are modulated *in trans* by other cellular proteins.

Little is known of such modulating proteins, although it has been shown that an interaction between cold-induced helicases CshA and CshB and the cold shock protein CspB is essential for cold adaptation in *Bacillus subtilis* (Hunger et al. 2006). In *E. coli*, the multiprotein RNA processing complex known as the degradosome has an alternative cold-shock form in which the RNA helicase component RlhB is replaced by a cold shock helicase CsdA (Prud'homme-Généreux et al. 2004), suggesting that some higher form of organisation is required for RNA metabolism at low temperatures.

Being regulated in response to various environmental signals, sometimes simultaneously, stress adaptation pathways are often controlled by regulatory complexes which facilitate co-ordination of many inputs. For example, in several bacterial species, the function of sigma factor B (σ^B) is regulated by a series of interactions between an anti-sigma factor and a number of anti-sigma factor repressors. These repressors, and ultimately the σ^B-dependent stress response are under the indirect control of a multiprotein complex known as the stressosome (Pané-Farré et al. 2005). The stressosome senses an array of energetic (carbon, phosphate status) and environmental (acid, salt, temperature stress) signals but processes all of them into a single response; the activation of σ^B (Akbar et al. 2001). Structural studies have suggested that the organisation of the stressosome complex facilitates this versatility: each sensor domain appears to be connected to a single, central, phosphatase responsible for initiating the cascade to release σ^B (Marles-Wright et al. 2008).

An initial bioinformatic investigation into CspE interactions was performed using a number of programmes and databases. The STRING database (Franceschini et al. 2013) predicts protein interactions by various methods including coexpression, gene positions, and conservation. Probed with CspE, STRING suggests interactions with

CrcB and PagP (previously discussed) based on their conserved expression as an operon, and with the initiation factor InfA, whose gene is also often found close to *cspE*. Interaction with the RNA polymerase core enzyme is suggested based on CspE's involvement in gene regulation, and is consistent CspE's antiterminator function. The protein annotation server PredictProtein (Yachdav et al. 2014) predicted a number of possible interaction sites located in the CspE's unstructured loops.

To investigate the possibility that CspE functions as part of a complex a cross-linking mass spectrometry (CLMS) technique was used to fix, purify, and identify proteins co-localised with CspE. *in vivo* cross-linking covalently captures intermolecular interactions, allowing them to be preserved through several purification stages before identification of unknown interactors by mass spectrometry. A workflow is shown in Figure 3.1A. Although initially developed for the study of a nuclear ribonucleoprotein complex in yeast (Neubauer et al., 1997) the method can be applied to any protein complex that can be biochemically purified, as long as a sequenced and annotated reference genome is available. For example, a study of the protein interactions of glyceraldehyde-3-phosphate dehydrogenase in *Escherichia coli* implicated the metabolic protein in other functions such as DNA repair and quorum sensing (Ferreira et al. 2013).

This method is highly sensitive and adaptable, although it lacks the high throughput of, for example, a two-hybrid or phage display library. Similar to co-immunoprecipitation CLMS allows the detection of *in vivo* interactions, thereby preserving the correct stoichiometry of interactions and retaining other cellular components that may modify interactions *in trans*. This factor is particularly advantageous given the function of CspE as an RNA chaperone. Performing the cross-linking step *in vivo* also eliminates the risk of losing a complex which can become unstable after cell lysis.

In brief, the method involves *in vivo* cross-linking of cellular components before purification and identification of purified proteins by mass spectrometry. Cross-linking with formaldehyde forms covalent bonds and allows purification of entire complexes of interacting proteins even under strenuous purification conditions. Such experiments are designed in order to determine whether CspE functions as part of a wider complex and, if so, which other proteins are involved.

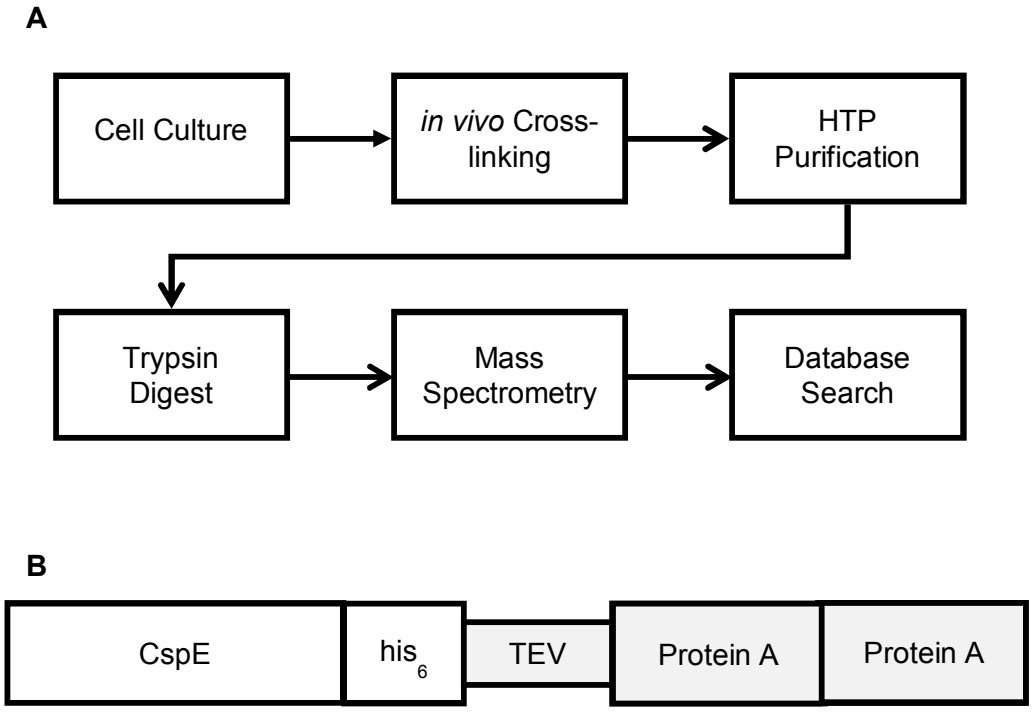


Figure 3.1: *Schematics of (A) the workflow for cross-linking mass spectrometric identification of interacting proteins and (B) the CspE-HTP (histidine tandem purification) construct.Protein A moieties and TEV site, shaded grey, are cleaved during purification. TEV = Tobacco Etch Virus protease cleavage site.*

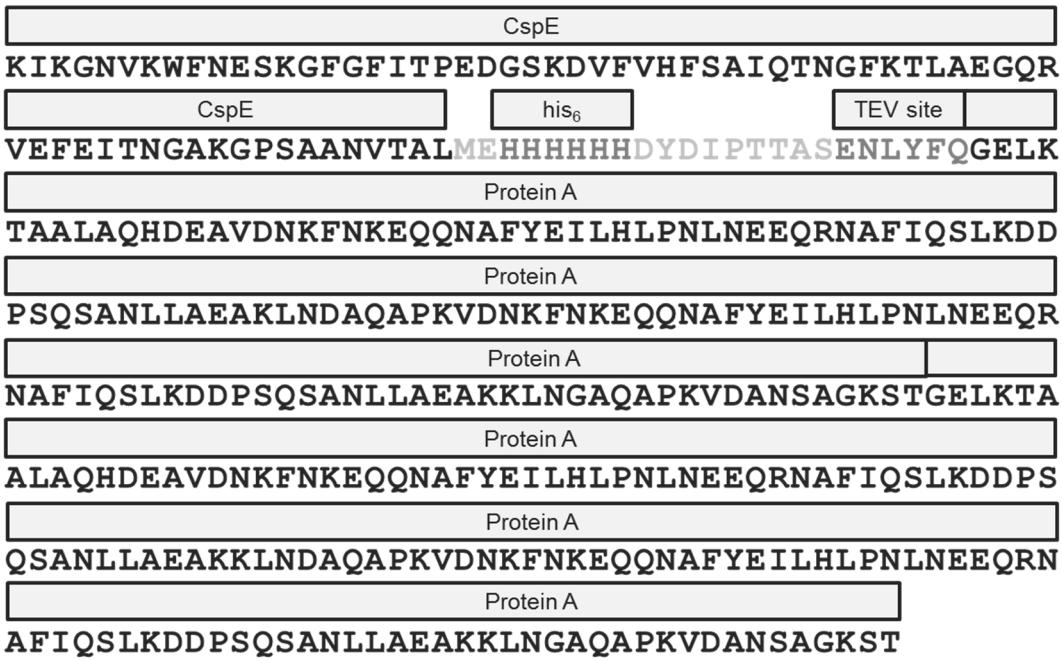


Figure 3.2. *Amino acid sequence of the translated CspE-HTP purification construct showing CspE and the his_6 tag separated from Protein A ZZ moieties by a Tobacco Etch Virus (TEV) protease cleavage site. CspE and Protein A amino acid residues black, his_6 tag and TEV cleavage site residues dark grey, linker residues light grey.*

Results

CspE-HTP Purification Construct

A modified TAP tag consisting of a his_6 tag and two protein A ZZ domains (derived from *Staphylococcus aureus*) separated by a tobacco etch virus (TEV) protease cleavage site (Figure 3.1B) was used for cross-linking. The construct was based on one previously constructed (Granneman et al. 2009) and was created in pTOF24, a pKO3 derivative allowing chromosomal integration (Merlin et al. 2002). pTOF24-cspE-HTP, a temperature-sensitive suicide vector, was integrated into the chromosome of *S.* Typhimurium SL1344 by homologous recombination with the chromosomal copy of *cspE*, thereby ensuring the construct is expressed under *cspE*'s native promoter (Woodall 2010). Recombination is temperature dependent and only occurs at 42°C, whereupon it can be selected for by simultaneous integration of a chloramphenicol resistance cassette. This expression system avoids oversaturation of the experimental cells with CspE (such as would occur with a plasmid-expressed purification construct) and helps minimise the formation of non-significant cross-links. The CspE-HTP construct has been used extensively and previously demonstrated not to interfere with the function of CspE (McGibbon 2013; Woodall 2010). The fusion protein retains RNA binding ability (McGibbon 2013).

in vivo Formaldehyde Cross-linking

An enormous variety of chemical cross-linking reagents are available, offering a wide range of properties. Broadly, they all consist of two reactive groups separated by a spacer. The composition of the reactive groups determines which molecular species are cross-linked *in vivo*, and specific chemical groups (such as protein side chains) can be targeted. Heterobifunctional cross-linkers feature a different reactive group at each end of the spacer while homobifunctional cross-linkers have two identical groups. In addition to functional groups, the length of the spacers vary and this factor determines the maximum distance between cross-linked groups (Paramelle et al. 2013). A commonly used cross-linker for structural studies is bis(sulfosuccinimidyl) suberate (BS3), a homobifunctional linker which reacts with amine groups. Although widely used *in vitro* (Darii et al. 2014; Płociński et al. 2014), BS3 is charged and therefore cannot cross a cell membrane. This renders it unsuitable for *in vivo* work.

Disuccinimidyl suberate (DSS) is a membrane-permeable analogue of BS3, but requires an incubation time of 30 minutes to 2 hours for cross-linking (Herzog et al. 2012; Santos et al. 2012); this presents a significant delay for the *in vivo* system utilised here. This compound is especially problematic given that DSS is not water-soluble and so would require addition of an organic solvent (dimethyl sulfoxide or dimethyl formamide) to the cell culture, which may disrupt cell viability and function independently of the cross-linking reaction.

Formaldehyde is one of the shortest available cross-linkers, with a length of ~2.5Å (CspE, by comparison, is around 30Å across). The experimental effect of this is to allow only tightly interacting proteins to be cross-linked, reducing the instance of non-significant interactions (ie. those without physiological relevance; or cross-links formed aberrantly) being captured. Formaldehyde has the added advantage in this case of reacting with a number of amino acid side chains (arginine, asparagine, glutamine, histidine, tryptophan, and tyrosine, as well as the amino terminus (Metz et al. 2004)), ensuring that all proteins (including small ones, such as CspE itself) can be cross-linked. Formaldehyde is commonly used in cross-linking mass spectrometry applications as it does not impair the function of mass spectrometers. It is also suited to capturing dynamic interactions as it passes the cell membrane rapidly and reacts within a few minutes: this is critical for avoiding artificial stress conditions which might be induced by longer reactions (as a result of, for example, jamming of the ribosome (Sutherland et al. 2008)).

In order to determine the optimal duration for formaldehyde cross-linking, a time course experiment was carried out. An ideal cross-linking reaction captures all significant interactions with a minimum of non-significant cell material. Crude lysates were prepared by boiling cell pellets and visualised by SDS-PAGE. Separated proteins were transferred to nitrocellulose by Western blotting. Blots were probed using peroxidase-conjugated anti-peroxidase antibody, to which the ZZ domains of protein A bind. As protein A binds at the constant chain of an antibody, the specificity of the variable chain is irrelevant here.

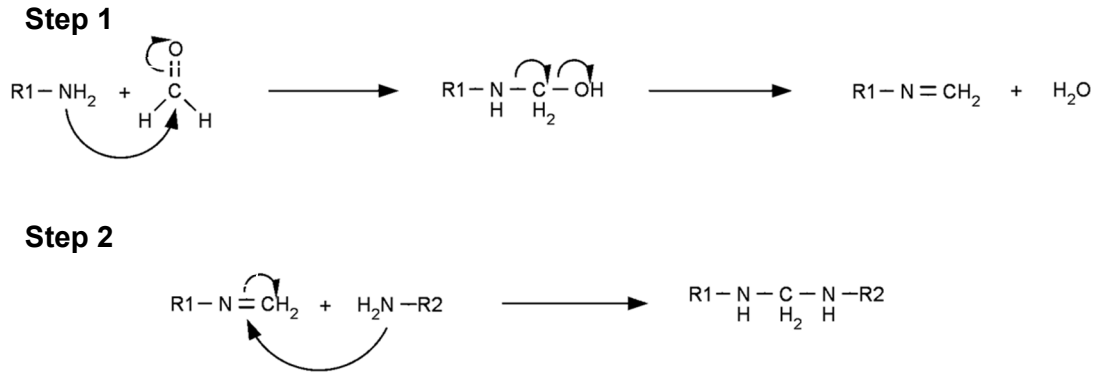


Figure 3.3.. *Reaction scheme of cross-linking by formaldehyde (CH_2O). A protein-derived nucleophilic group (here NH_2) attacks formaldehyde, producing an unstable methanol (CH_2OH) derivative. After dehydration, an imine (N=C) remains (step 1). A second protein-derived nucleophile attacks the imine and forms a second bond, leaving the two proteins (R1 and R2) linked by the planar methylene bridge.*

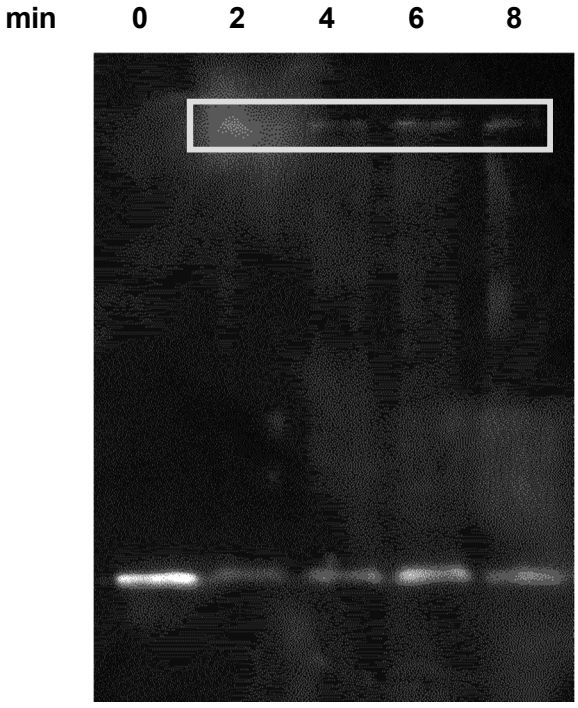


Figure 3.4. *Time course of formaldehyde cross-linking. Western blot against protein A domains of CspE-HTP (peroxidase-conjugated antiperoxidase antibody, 1:10,000 dilution) showing much of the material remaining unlinked but, over time, some larger complexes visible as a smear or retained in the wells of the gel (boxed). Antibody signal was detected with chemiluminescent peroxidase substrate.*

During cross-linking, formaldehyde (CH_2O) is subjected to attack by a peptide-derived nucleophilic moiety (derived from lysine, tryptophan, or cysteine side chains as well as protein N-termini) to yield a carbinol derivative ($R\text{-}CH_2\text{-}OH$), which undergoes dehydration to leave the protein tagged with a methylene group ($R{=}CH_2$). This derivative is subject to a second nucelophilic attack by another peptide-derived moiety, either from the same protein or a different one, producing a methylene bridge which links the two proteins ($R\text{-}CH_2\text{-}R$) (Toews et al. 2008). A reaction scheme is shown in Figure 3.4. Results of the time course are shown in Figure 3.3. The majority of the CspE-HTP construct was not cross-linked at 8 minutes although it can be seen that, after 2 minutes, cross-linked material starts to appear. Separate experiments (not shown here) showed that total cross-linking was not achieved even at 30 minutes.

Based on the time course, a 5 minute reaction time was chosen, corresponding to a minimum time at which cross-links would be reliably formed. Although longer cross-linking times would increase the number of interactions captured and purified, the potential significance of those interactions would be reduced, owing to the increased risk of non-specific binding being captured. A shorter cross-linking duration ensures that only the strongest or most common interactions are captured, increasing the likelihood that identified interactions exhibit a significant physiological function.

Protein-Protein Interactions of CspE

Purification of complexes

Following disruption of cells using silicate beads, cross-linked proteins were purified by affinity purification. The previously described histidine tandem purification (HTP) construct is covalently attached to the C-terminus of CspE (Woodall 2010) as illustrated in Figure 3.2. The tag is a tandem affinity purification (TAP) tag derivative composed of a protein A group separated from a polyhistidine (his_6) tag by a protease cleavage site. Cross-linked CspE complexes are purified by two sequential affinity purification steps. The first uses immunoglobulin G (IgG) resin to bind Protein A ZZ domains, and removes the majority of non-cross-linked cell material. Protease cleavage allows the liberation of CspE (and its cross-linked partners) but leaves the ZZ domains bound to the IgG resin. CspE complexes are further purified using the his_6 tag exposed by cleavage of the ZZ domains, which has affinity to nickel (Ni^{2+}-NTA) resin. This second affinity purification removes remaining non-cross-linked proteins along with any remaining Protein A moieties. Use of a two-stage purification method ensures high sample purity at the expense of some loss of yield, making the highly sensitive LTQ-Orbitrap LC-MS/MS system an ideal platform for subsequent mass spectrometric analysis. The full procedure is outlined in Chapter 2.

Figure 3.4 illustrates the purification process visualised by silver-stained polyacrylamide gel and Western blot. Methods for both are described in Chapter 2. Figure 3.4A shows the elimination of proteins not cross-linked to CspE-HTP from the cell lysate (lane 1). Most material is removed in the IgG flow through and washes (lanes 2 and 3 respectively) although, as can be seen in the TEV eluate (4) some contaminating proteins remain. This illustrates the necessity of a two-step purification. These proteins are removed during the Ni^{2+} purification steps (lanes 5 and 6) leaving only the desired cross-linked CspE complexes in the final eluate (lane 8, 60 kDa and above). As expected, a large amount of CspE was not cross-linked and is seen as a band of ~7 kDa in the final eluate.

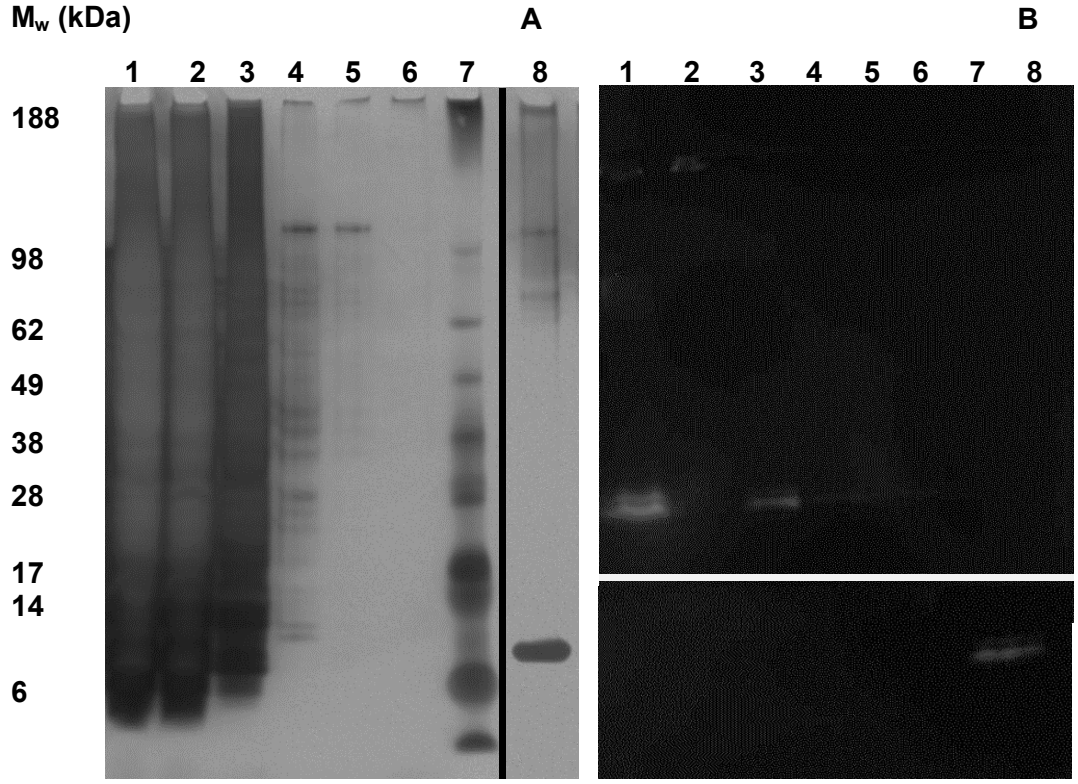


Figure 3.5. *Composite image showing purification of cross-linked CspE complexes.*

(A) silver stained polyacrylamide gel showing two-stage purification. Lane 1: lysate. Lane 2: IgG resin flow through. Lane 3: IgG resin wash,Lane 4: eluate following TEV cleavage. Lane 5: Ni^{2+}-NTA bead flow through. Lane 6: Ni^{2+}-NTA bead wash. Lane 7: size markers (represented masses shown left). Lane 8: final eluted complexes.

(B) top, Western blot against protein A of the same samples (peroxidase-conjugated antiperoxidase antibody, 1:10,000 dilution). Protein A tagged CspE (~26 kDa) is present in the cell lysate (B1) and faintly in the wash (B3). (B) bottom, Western blot against his_6 tag (peroxidase-conjugated anti-his_6 antibody, 1:1000 dilution): tag is only visible in the unlinked portion of the final eluate (~7 kDa, B8).

S*. Typhimurium SL1344 was grown to OD_{600} 0.4 in 4.5 L LB broth. Cells were cross-linked with formaldehyde and lysed by vortexing with silicate beads. HTP purification was performed as described, with samples collected and run on an SDS-PAGE gel (NuPage® 4-12% Bis-Tris (Invitrogen) with NuPage® MES SDS running buffer (Invitrogen) at 200V. Proteins were transferred to nitrocellulose membrane and incubated with horseradish peroxidase-conjugated antibody, detected by chemiluminescent peroxidase substrate.*

Figure 3.5B shows two Western blots against the protein A portion of the tag (top) and the his_6 portion (bottom). The bright band seen at ~28 kDa in the anti-protein A blot corresponds to the size of the entire tagged CspE protein without any cross-links. Also seen in lane 1 (the cell lysate) is cross-linked material, visible as a smear of higher masses and some material retained in the well. Complexes retained in the well must be >188 kDa (the size marker which remains in the well) although are likely to be much larger – they may be large aggregates formed as a by-product of the cross-linking reaction. Such aggregates are expected to form randomly from many cellular proteins during the cross-linking reaction due to the high activity and low specificity of formaldehyde as a cross-linker. Most of these larger aggregates should flow through the IgG column (in lane 2); their size may destabilise interactions with the IgG resin, or protein A moieties may be hidden by cross-linked proteins. It is unlikely that these lost complexes contain any significant or otherwise undetected interacting partners.

A little CspE-HTP is lost in the first wash step (lane 3), likely meaning that some cross-linked complexes are also lost. However, vigorous washing helps to ensure the purity of the samples going forward into mass spectrometry and thereby improve the reliability of the results. A small amount of CspE-HTP survives the TEV cleavage. Although the cleavage should leave protein A bound to the IgG resin and therefore undetectable in this blot, a little appears to have been retained in the TEV eluate (4), although it is lost in the Ni^{2+} resin flow through. As expected, no protein A is detected in the Ni^{2+} wash or eluate (5 and 7).

The lower blot in Figure 3.4B shows the presence of the his_6 tag. This tag is only detectable in the unlinked portion of the final eluate, being at other times hidden by protein A or remaining bound to the Ni^{2+} resin. Cross-linked complexes are shown to be present in the final eluate by silver staining (lane 7) but are undetected in the Western blot due to low concentration or occlusion of the his_6 epitope by cross-linked proteins. Purified complexes (of which only a small sample was run for the gels above) were prepared for mass spectrometry by immobilising in a pre-cast protein gel as described in Chapter 2.

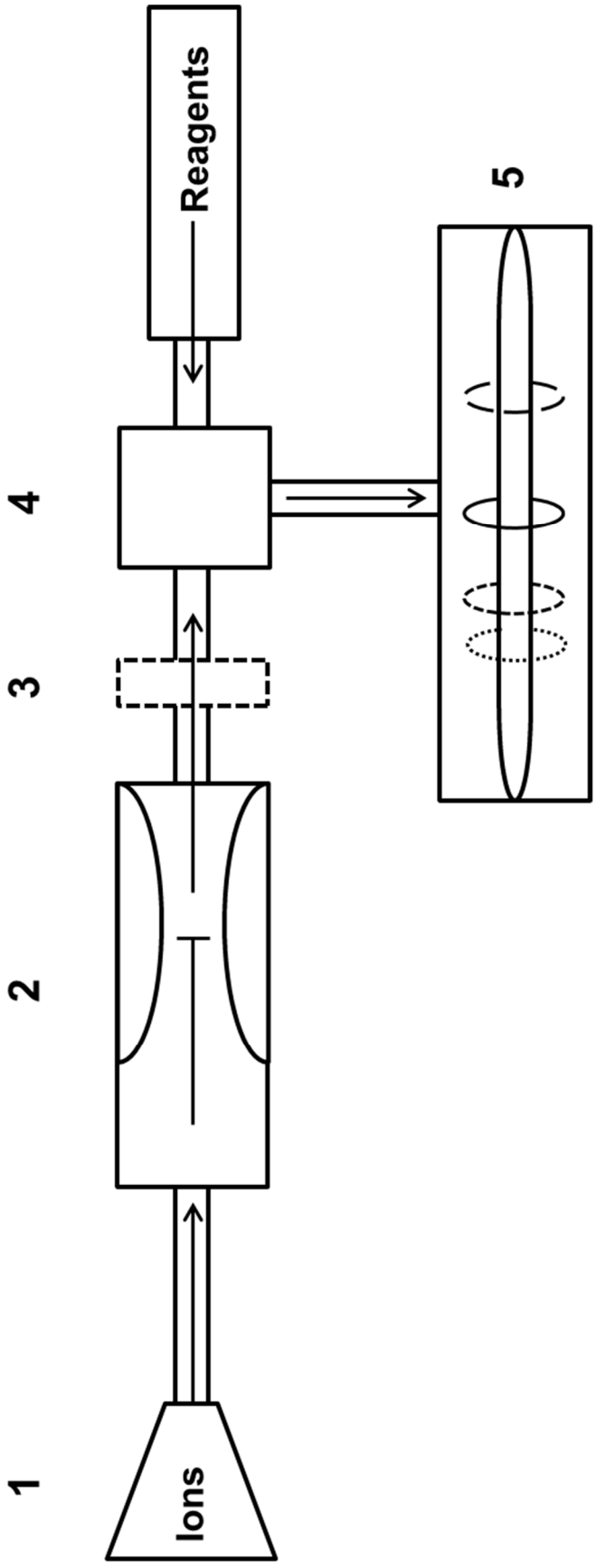


Figure 3.6. *Schematic of the LTQ-Orbitrap. Ions separated during liquid chromatography are subjected to electrospray ionisation (1) and injected into the first MS step in the linear trap quadrupole (LTQ; 2). After the first MS ions are filtered (3) and passed via the C-trap (4) into the Orbitrap (5). In the Orbitrap, ions orbit around a central electrode, where their mass is calculated from the frequency and speed of oscillation.*

Mass Spectrometry

Purified protein complexes were immobilised in polyacrylamide gel, subjected to an overnight in-gel trypsin digest, and analysed using a Velos LTQ-Orbitrap mass spectrometer (Thermofisher Scientific). Figure 3.3 shows a schematic of the LTQ-Orbitrap. Resultant masses were used to probe a database of *Salmonella* proteins derived from the genome of strain SL1344 (Kröger et al. 2012). A control sample prepared identically but for the omission of the cross-linking step allows identification of background proteins erroneously purified due to, for example, abundance or aberrant interaction with purification materials. Proteins detected in the control are subtracted from those detected in the experimental samples, thereby leaving only the significant cross-linked interactors.

Analysis of complexes was carried out with the assistance of F. de lima Alves (Rappsilber Lab, Wellcome Trust Centre for Cell Biology, Edinburgh) using an Orbitrap tandem mass spectrometer as described in Chapter 2. Three repeats were performed, the full details of which are presented in Appendix A. Interacting proteins were determined by subtracting peptides identified in the control sample from those identified in the cross-linked sample, thus eliminating background hits. Data from the three repeats is summarised in Figure 3.5. In addition to the proteins shown, CspE was identified in each control and experimental sample, confirming success of the purification.

Full mass spectrometry conditions are given in Chapter 2. In brief, the method involved two mass spectrometric steps with the first filtering out low abundance ions allowing better resolution of the desired ions in the second step. Proteins identified by two or more unique peptides in the experimental (cross-linked) samples but not in the control (untreated) samples were deemed significant. Although the Orbitrap mass spectrometer does enumerate the number of peptides detected for each protein the number does not necessarily relate to abundance. Many other factors including the protein size, number and accessibility of trypsin cleavage sites, and the properties of the peptide ions themselves can affect the number of peptides detected. As such, the figures are of limited use and are not included here, although they are given in Appendix A.

Interactors identified in Sample 1

Name	Function	Name	Function
AceF	Pyruvate dehydrogenase	GltA	Citrate synthase
ValS	Valyl-tRNA ligase	PhoP	Transcriptional regulator
FabF	Fatty acid synthesis	YdhD	Hypothetical
LeuS	Leucyl-tRNA ligase	MetG	Methionyl-tRNA ligase
Pnp	Polynucleotide phosphorylase	GuaC	GMP reductase
Gnd	6-P guconate dehydrogenase	DeoC	Deoxyribose-phosphate aldolase
TalB	Transaldolase	PepQ	Proline dipeptidase
Tsf	Elongation factor Ts	GlmM	Phosphoglucosamine mutase
LysS	Lysly-tRNA ligase	PrlC	Oligopeptidase
IcdA	Isocitrate dehydrogenase	CarB	Carbamoyl-phosphate synthase
PykF	Pyruvate kinase	Ssb	ssDNA-binding protein
Udp	Uridine phosphorylase	SfcA	Malate oxidoreductase
GapA	G3P dehydrogenase	Upp	Uracil phosphoribosyltransferase
DnaN	DNA polymerase	RpsR	30S ribosomal
ArgS	Arginyl-tRNA ligase	YebC	Hypothetical
DeoB	Phosphopentomutase	IleS	Isoleucyl-tRNA ligase
GlnA	Glutamine synthetase	PrfC	Peptide release factor
AspC	Aspartate aminotransferase	TrxA	Thioredoxin
TsaA	Alkyl hydroperoxide reductase	TopA	DNA topoisomerase
TktA	Transketolase	IlvD	Dihydroxyacid dehydratase
PurA	Aadenylosuccinate synthetase	RfbG	CDP-glucose 4,6-dehydratase
SecB	Protein export	HupA	DNA binding protein
SucD	Succinyl-CoA synthetase	PepD	Dipeptidase
AhpC	Alkyl hydroperoxide reductase	SL1344_1223	Oxidoreductase
DapD	Succinyl transferase	Pgm	Phosphoglucomutase
TrxB	Tioredoxin reductase	NagD	Dephosphorylase
Fba	Fructose 1,6-biP aldolase	GalF	Uridylyltransferase
PyrH	Uridylate kinase	YecO	Hypothetical
Rnb	Exoribonuclease II	Rl18	50S ribosomal
NuoG	NADH dehydrogenase		

Interactors identified in Sample 2

Name	Function
Tig	Trigger Factor
SlyD	Peptidyl-prolyl isomerase
RplV	50S ribosomal
GroES	Protein folding chaperone
RplI	50S ribosomal
RpsJ	30S ribosomal
GapA	G3P dehydrogenase
RpsO	30S ribosomal
Tsf	Elongation factor Ts
RpsN	30S ribosomal
Rpl32	50S ribosomal
ValS	Valyl-tRNA ligase
HupA	DNA binding protein
RpsR	30S ribosomal
FtsK	Cell division

Interactors identified in Sample 3

Name	Function
RnE	Ribonuclease E
YqjL	Hypothetical
GyrB	DNA gyrase
LepA	Elongation factor
CbpA	DNA-binding
Mrp	Fe/S cluster assembly

Figure 3.7. *Tables showing interactions found by three repeats of cross-linking mass spectrometry experiment. Proteins related to translation (ribosomal proteins and tRNA ligases) highlighted light grey, those associated with the chromosome highlighted dark grey. n proteins identified in sample 1 = 59, sample 2 = 15, sample 3 = 6. Full tables of results are included in Appendix A.*

Across three repeats 80 proteins were detected unique to the experimental samples. Full tables of results are included in Appendix A. As expected, many of the most abundant cellular proteins were detected in both the control and experimental samples, and were thus discounted. The three repeats yielded a total of 75 possible interacting proteins. None were common between all three samples, although 5 (6.7%) were identified twice. These 5 are listed in Table 3.1.

Interactors identified in Samples 1 and 2	
Name	**Function**
GapA	G3P dehydrogenase
HupA	DNA binding protein
RpsR	30S ribosomal
Tsf	Elongation factor Ts
ValS	Valyl-tRNA ligase

Table 3.1. *Table showing the names and functions of the 5 common interactions identified in repeats 1 and 2.*

Discussion

Cross-linking mass spectrometry was utilised to determine whether CspE functions as part of a multi-protein complex at 37°C and, if so, what the other components of the complex are. When successful this method identifies, along with the target protein, a small number of proteins consistently cross-linked and purified, which are taken to be co-functional with it.

in vivo cross-linking, affinity purification, and mass spectrometry were all carried out successfully. Interactions identified are outlined in Figure 3.7 and Table 3.1, and discussed below. Although no clear complex partners were consistently identified across all 3 samples, a number of possible interactors from two of the three samples support current models of CspE's function in translation. In addition, a number of co-purified proteins originate in the nucleoid, providing potential evidence for an additional function of CspE around the chromosome.

The three samples showed a high degree of variability, however, and the results are difficult to interpret. Growth conditions were constant across all repeats, as were the cross-linking, purification and mass spectrometry steps. The observed variation is consistent with a function of CspE modulated not by the action of a single complex but by more transient interactions with a wider range of partners. Such transient interactions might be inconsistently captured and identified by cross-linking mass spectrometry.

Identified Interactors of CspE

A number of interacting partners of CspE were identified (80 in total, representing around 1.8% of the total *Salmonella* Typhimurium proteome). Being identified in two of the three repeats, five possible interactors seem the most likely to be significant. Those five, as listed in Table 3.1, are GapA, HupA, RpsR, ValS, and Tsf.

GapA, an NAD^+-dependent glyceraldehyde-3-phosphate dehydrogenase, is involved in the early stages of glycolysis (Wang et al. 2013). Although a metabolic gene, it has been implicated in stress responses such as recovery from heat shock in *Salmonella* (Kobayashi et al. 2005) and in protection against oxidative stress in *Lactobacillus* (Rochat et al. 2012). In *Staphylococcus*, GapA has been linked to

pathogenesis (Purves et al. 2010). Highly conserved and apparently essential even under conditions where glycolysis is not required, it has been suggested that GapA (and a handful of other glycolytic enzymes) have important non-enzymatic functions. These include ATP synthesis, translation, and regulating transcription (Ferreira et al. 2013; Kim & Dang 2005). In addition, energy metabolism has been shown to link to CspE regulation via the cyclic AMP receptor protein CAP, which activates the *cspE* promoter (Uppal et al. 2011). It is not obvious, however, why GapA might interact directly with the RNA chaperone CspE.

CspE and translation

One of CspE's known functions is in facilitating translation. This makes it is unsurprising to find an association between CspE and RspR (the ribosomal S18 subunit) or Tsf (elongation factor Tu), both being located close to the ribosomal entry site (Schmeing et al. 2009). Other translation-related proteins are highlighted in Figure 3.7. The association between CspE and the ribosome has been noted in *B. subtilis* to be dependent on translation (Weber, Volkov, et al. 2001) and complementation of a *cspBD* double deletion strain by the translation initiation factor IF1 (again in *B. subtilis*) suggests some overlap in function (Weber, Beckering, et al. 2001). A localisation close to the ribosome, as evidenced by the findings herein, supports the idea that CspE is involved in translation. Although CspE's association with Tsf and RpsR fits with the model in which CspE facilitates translation, neither are candidates for components of a higher complex. The elongation factor Ef-Tu has a specific function (delivering aminoacyl tRNA to the ribosome) requiring dimerisation with a partner Ef-Ts (Kawashima et al. 1996). This dimerisation may well be impaired by significant interactions with another protein, preventing the correct physiological function of Ef-Tu. Functioning as part of the ribosome, RpsR is unlikely to have another role elsewhere.

The intrinsic secondary structures of tRNA molecules may well prompt CpsE binding, and the tRNA may act as a scaffold in promoting CspE-Tsf interaction. This same factor may account for the observed interaction between CspE and valyl-tRNA ligase (ValS), although it could be expected that if CspE does associate with tRNAs other aminoacyl tRNA ligases would also have been identified during this experiment. ValS was the only such enzyme identified twice, although sample 2 identified a further 5

tRNA ligases. Other investigations in our lab into the RNA-binding activity of cold shock proteins revealed that many tRNAs associate with CspE, although the abundance of tRNA in the cell makes it difficult to be certain of the significance of such associations (McGibbon 2013).

CspE and the nucleoid

The nucleoid protein HupA forms one half of the DNA-binding dimer HU, which exists as heterodimers containing HupA and HupB, although the composition appears variable. HU preferentially binds supercoiled DNA, inducing bending, and also to structural abnormalities such as nicks and junctions where it acts as a stabiliser of bent DNA. HU, by bending DNA as well as by stabilising and inducing supercoiling, contributes to compression of the bacterial chromosome (Dame 2005). In exponential phase, HupA-HupA homodimers dominate, while in stationary phase HupA-HupB heterodimers are more common. Following a cold shock, translation of HupA is decreased (due to destabilisation of mRNA transcripts) while that of HupB increases, leading to increased numbers of either heterodimers or HupB-HupB homodimers (Giangrossi et al. 2002). Given that the experiment described was carried out in exponential phase, the presence of HupA and absence of HupB is as expected.

Association of CspE and HupA suggests a localisation of CspE to the nucleoid (other nucleoid-associated proteins are highlighted in Figure 3.7). It is known that CspE associates to nascent mRNA (Hanna & Liu 1998), a function which would presumably place CspE close to the chromosome (with RNA polymerase (Cabrera & Jin 2003)). In addition, it has been reported that CspD is located in the nucleoid (M Giangrossi et al. 2001) although the precise subcellular location of CspE has not been previously determined. CspE has long been reported to play a role in chromosome condensation, however (Hu et al. 1996), which suggests some role around the nucleoid although the mechanism remains unknown.

HupA binds genomic DNA in double-stranded regions but also has a high affinity for single-stranded DNA, which arises in the chromosome due to mismatches or in structures such as forks and overhangs (Kamashev et al. 2008). Although traditionally considered an RNA chaperone, the cold shock proteins have been demonstrated to bind ssDNA *in vitro* (Jiang et al. 1997). It is possible that, rather than

interacting, CspE and HupA simply share the same substrate (in ssDNA), which would account for their co-localisation. The single-stranded binding protein SSB was also purified in sample 1. Similarly, the role of HupA in initiating DNA replication by promoting strand opening (thereby forming a single-stranded region) (Chodavarapu et al. 2008) may explain the co-purification of CspE and DnaN, part of the DNA polymerase processivity factor.

Co-localisation or interaction between CspE and the nucleoid protein HU has not been reported previously, and although this finding obviously requires further validation and characterisation, it is an intriguing avenue to explore. However, many of the proteins purified from the nucleoid can be accounted for by common binding sites at ssDNA, suggesting an association with CspE may be coincidental. Whilst it may be that CspE has a physiological role in binding ssDNA in the nucleoid, possibly contributing to or preserving chromosome condensation, it may also be that ssDNA binding is a product of the non-specific nucleic acid binding surface of the cold shock proteins. In either case, the role of the cold shock proteins in the nucleoid is little understood, either at optimal or low temperatures.

Transient Protein-Protein Interactions

Transient protein-protein interactions are generally distinguished from permanent interactions by their lower affinity and shorter lifetime. However, the term covers a wide range of attractions and thus they are often grouped into strong and weak transient interactions (Ozbabacan et al. 2011). Interactions are specific either to a surface structure or a peptide sequence; they occur between two surfaces (domain-domain) or between a surface and a recognised linear motif (domain-peptide). Transient interactions are known to play an important role in a number of signalling and regulatory processes, systems which depend on dynamism and change.

Transient intermolecular interactions may facilitate diverse functions such as ligand binding (e.g. DNA binding proteins) or exchange (e.g. GTP exchange factors), conformational change, postranslational modification, or change in stability (Nooren & Thornton 2003). These functions are clearly vital for many processes, and so therefore are the interactions that underpin them. Weak interactions are not merely found in the absence of a strong interaction; they can fulfil functions which a stronger interaction

would prohibit. In the example of Che, the chemotaxis-regulating two-component system, weak interactions between kinase CheA and regulators CheB and CheY allows competition between the two regulators, producing a more sensitive response than stronger interaction would accommodate (Li et al. 1995). In the case studied here it is possible, then, that weak transient interactions may implicate CspE in a greater range of functions (or subject it to a greater range of modulators) than involvement in a single complex.

The previously mentioned single-stranded DNA binding protein SSB coordinates a range of functions by recruiting DNA modulating proteins to ssDNA. SSB is a coordinator of DNA replication through DnaG (Naue et al. 2013), recombination through the primase PriA (Cadman & McGlynn 2004), and degredation by the ssDNA-degrading exonuclease 1 (Exo1) (Genschel et al. 2000). This range of function is facilitated by the disordered C-terminal tails of SSB, which functions as a homotetramer. The amphipathic tails (8 aa in length) contain a series of acidic aspartate residues and three hydrophobic residues. The resulting amphipathy and intrinsic disorder of the tail promotes interaction with a wide range of proteins (Shereda et al. 2008). As discussed in Chapter 1, there are a number of sites on CspE which could function as non-specific interaction sites.

Measured protein-protein interactions contained in a database of thermodynamic data (Kumar & Gromiha 2006) display a range of dissociation constants from 10^4 (between cytochrome C and cytochrome C peroxidase, which exchange reducing factors during electron transport (Pelletier & Kraut 1992)) to 10^{15} (in the barnase-barstar toxin-antitoxin system (Buckle et al. 1994)). The transient interactions discussed here have dissociation constants around 10^5 - 10^6, in the cases of SSB - PriA and SSB - Exo1 (Genschel et al. 2000).

The idea of a molecular scaffold which facilitates interactions between proteins with otherwise low affinity is attractive in the case of CspE. Many scaffolds are proteins but RNA can also display scaffolding activity (Spitale et al. 2011; Tsai et al. 2010). As discussed, ssDNA would also be a suitable candidate in the case of CspE. It is easy to imagine an RNA molecule (nascent, free or bound to a ribosome) acting as a scaffold around which many nucleic acid binding proteins (such as those discussed above) are found. Elevated local concentrations would facilitate interactions between proteins, even in the absence of an obvious interacting interface. Such scaffolds can take a

number of forms including, for example, a ligand stabilising interactions within a transporter complex (Vigonsky et al. 2013) and a misfolded protein stabilising a heat shock chaperone (Bepperling et al. 2012). The second example is particularly interesting as the transient nature of its complex formation appears to facilitate a non-specific activity. Thus, transient interactions formed around a scaffold are not only possible but also potentially valuable in allowing a generic, rather than a specific, function. A possible function for CspE is as a non-specific RNA chaperone; transient interactions may facilitate this function.

Issues of Reproducibility

Previous high-throughput cross-linking mass spectrometry experiments have shown a surprisingly high level of variation. Two such examples were published in 2005 (Butland et al. 2005) and 2006 (Arifuzzaman et al. 2006), both with the aim of identifying as many protein-protein interactions as possible in *E. coli*. The 2005 study expressed TAP- or SPA- (sequential peptide affinity) tagged proteins under their native promoter, whereas the 2006 study overexpressed his_6 tagged proteins from a plasmid vector. The two were otherwise similar: *E. coli* K12 was grown to exponential phase and chemically lysed, proteins were affinity purified and extracted from SDS-PAGE gels before being identified by mass spectrometry. Neither study utilised a cross-linking agent. The two identified 521 proteins in common, with 5030 interactions occurring between them in 2005 and 3088 in 2006 (a total of 8118 interactions identified). Of those interactions described, however, only 218 were found in both studies, representing 2.8% of the total interactions discovered.

This discrepancy highlights the importance of several decisions regarding the experimental approach for such techniques. The 2005 study placed purification tags at the C terminus, where the 2006 study tagged proteins at the N-terminus. The former used two-step purification, thereby increasing purity at the expense of yield, where the latter used a single purification step. One study retained native stoichiometry by utilising the native promoters where the other overexpressed target proteins. Overexpression overcomes the limits of instrument sensitivity but also elevates the risk of false positives or abnormal cell behaviour. Clearly these differences can (and did) have a huge impact on the data generated, and raises the question of whether the difference in the results is due to differences in the biology of the system or simply the

method chosen. With the addition of cross-linking, a further variable is introduced with respect to the duration of cross linking and the balance between amount of cross-linked material formed and ensuring the significance of the interactions captured. Ideally each variable (location of tag, protein expression cross-linking times, purification method) would be investigated experimentally, as would other biological variables such as growth phase or exposure to stress conditions such as low temperature. Unfortunately the complexity of the method used in the present study prohibits extensive optimisation, especially in the absence of a clearly identified interacting partner.

Conclusions

Clearly, the described cross-linking mass spectrometry method suffers a lack of reproducibility. The three experimental repeats overlap by only 7% and one repeat has a completely discreet set of hits from the other two. Cross-linking mass spectrometry of a stable complex could be expected to yield a much more consistent set of data. It is likely, therefore, that CspE does not interact with other proteins in a stable and defined complex at 37^{o}C. Rather, it may interact with a wide array of proteins in a weak or transient manner, such as would not be reliably captured by the short period of covalent cross-linking used here. An alternative method, better suited to capturing transient protein-protein interactions, may be more appropriate. Studies carried out in Chapters 4 and 5 utilise a two-hybrid system using growing cells over a longer time period, which should allow detection of these weaker interactions.

A number of potential interactions were identified, however, with HupA in particular being an interesting possible binding partner of CspE. Confirmation and further characterisation of this interaction may reveal a novel function for CspE.

Chapter 4

Further Studies of CspE Interactors

Further Studies of CspE Interactors

Previous experiments used a proteomic method to examine the possibility of CspE involvement in a multiprotein complex. A small number of proteins were identified which might interact with CspE, although the results were not conclusive. As a test case the interaction between CspE and HupA, which was suggested by the cross-linking mass spectrometry experiment of Chapter 3, will be examined further here. As it was not consistently identified by CLMS, the potential interaction must be validated by another method. The bacterial adenylate cyclase two-hybrid system will be used here to examine the interactions of CspE in greater detail. The system allows quantitative assessment of the capacity of CspE and HupA to interact, and will also allow investigation of CspE's interaction with other proteins. For example, auto-association of CspE is an unknown aspect of its function *in vivo*, although it has been reported *in vitro*. An interaction between CspE and a cold-shock helicase has been reported at low temperatures in *Bacillus subtilis*, and can also be examined using this method. Additionally, the two-hybrid system described below can be further expanded into a genome-wide screen for interactions of CspE, should such a development be required.

Initial experiments will focus on the potential interaction between HupA and CspE, which was suggested by CLMS but requires further study owing to the inconclusive results from that experiment. The co-purification of CspE and HupA may have been the result of a direct interaction, but may also have been the result of a co-localisation around a common substrate. In addition, CspE oligomerisation will be studied. Dimerisation of CspE has been reported *in vitro* (Johnston et al. 2006; Morgan et al. 2009) but not *in vivo*. Further, the cold shock proteins are proposed to bind co-operatively along the length of an RNA substrate (Phadtare 2005), raising the possibility of oligomerisation *in vivo*.

The bacterial adenylate cyclase two-hybrid (BACTH) system was first described in 1998 (Karimova et al. 1998). The system is based on the two fragments of the catalytic domain of the *Bordetella pertussis* adenylate cyclase toxin (T25 and T18, being 25 and 18 kDa respectively). When fused to two interacting proteins, the two fragments are brought to sufficient proximity to reconstitute a functional adenylate

cyclase, resulting in production of cyclic AMP (cAMP) and activation of a number of reporter genes. cAMP levels are assayed in an adenylate cyclase mutant (Δ*cyaA*) *Escherichia coli* strain, DHM1. Typical assays for adenylate cyclase activity are based on expression of β-galactosidase, which is dependent on cAMP levels in the cell, being under control of the catabolite activated protein CAP (Jacob & Monod 1961). β-galactosidase expression can be assayed by growth on minimal medium with lactose as a sole carbon source, hydrolysis of X-gal on solid medium, or hydrolysis of ONPG by cell lysates. A strong interaction between the proteins fused to T25 and T18 produces more cAMP and therefore higher expression of β-galactosidase. The assay is outlined in Figure 4.1.

The BACTH system has been used extensively since its first description. The first (and still most common) use was for the investigation of interacting domains of a single protein such as, in the initial application, tyrosyl-tRNA synthase (Karimova et al. 1998)(Hankins et al. 2007)[181]. However, the method has also been used for diverse applications such as studies of dimerisation (Al-Bassam et al. 2014), regulatory interactions (Hryckowian et al. 2014) and the delineation of complex structures (X. Y. Zhang et al. 2013).

One of the key advantages of BACTH over other two-hybrid systems is that the signal is transduced by a freely-diffusing second messenger (cAMP), where many systems require the interaction to occur at the site of transcription initiation. Most yeast two-hybrid systems, for example, fuse interacting partners to the DNA binding and activating domains of the transcription factor Gal4 (Fields & Song 1989). While this produces efficient activation of transcription, it requires both interacting partners to diffuse freely and be able to access the nucleus. In addition, the compact nucleus might prohibit access to other components required for an interaction, such as other proteins or cellular structures.

The first bacterial two-hybrid system followed a similar methodology, with the interacting proteins being fused to a DNA-binding transcriptional activator protein and the α-subunit of RNA polymerase (Dove et al. 1997). In this instance, in addition to the problems discussed, the system suffers a level of basal transcription, resulting in false positives. The system also requires RNA polymerase to function with a protein fused to its α-subunit, and for the interaction assayed to be weak enough for RNA polymerase to break free during transcription initiation.

BACTH allows detection of interactions which occur at sites spatially distinct from the nucleoid and unable to move around the cell, such as those of membrane-bound proteins. For example, the system has been used to probe the interactions in bacterial two-component systems (Scheu et al. 2012). This factor is particularly important given that some cold shock proteins have been observed to localise outside the nucleoid, at the cell poles (M Giangrossi et al. 2001).

The adenylate cyclase of *Bordetella pertussis* (Cya) has been extensively studied due to its role in the virulence of the bacterium, which causes whooping cough. Secreted as an unfolded polypeptide, the protein contains a catalytic domain of around 400 amino acids which is only folded correctly in the presence of an external protein, the eukaryotic calmodulin. Inside the eukaryotic cell, activated by calmodulin, *B. pertussis* adenylate cyclase changes host behaviour by altering cAMP levels (Vojtova et al. 2006). The capacity of the two catalytic domain subunits (T18 and T25) to be reconstituted thus is exploited in the BACTH assay, wherein two interacting proteins are fused to the catalytic domains and assume the function of eukaryotic calmodulin. The cAMP produced by active Cya allosterically activates the transcription factor cAMP-activated protein (CAP), which regulates expression of a range of genes by interacting with RNA polymerase and a number of CAP-dependent promoters (Harman 2001). Among those regulated are the *lac* and *mal* operons, either of which can be assayed easily. Utilising the *lac* operon, the BACTH system is highly sensitive and in the Δ*cyaA E. coli* host strain false positives should be minimal.

The BACTH assay indicates the potential for interactions between two partners. Both partners are overexpressed and other regulators or cofactors may be absent, so the system does not mirror genuine *in vivo* conditions directly. A false negative result may arise if an interaction requires some cofactor or other structure absent in the *E. coli* BACTH host. However, *E. coli* and *S.* Typhimurium are highly similar, so it is likely any such required cell components would be preserved in *E. coli*. Even a positive result may not be physiologically relevant at in all conditions. Other factors such as expression patterns, subcellular localisation, and the stoichiometry of an interaction must be considered.

Equally, the constructs need not necessarily retain full function in order for an interaction to be detected. As long as the structural aspects which drive interaction (such as exposed aromatic residues or hydrophobic patches) are preserved, an

interaction should still occur and be detected even if that interaction is not fulfilling its biological function. The constructs are highly expressed from their plasmid vectors, and so will be in excess in the cell and as such deletions of the chromosomal gene copies are not required. The *E. coli* chromosomal homologues of the proteins studied will still be present, although the abundance of the constructs improves the chance of detecting interactions. The two constructs must be essentially in direct contact to elicit significant adenylate cyclase activity; the linkers between inserted protein and Cya fragment are only 10-15 aa in length.

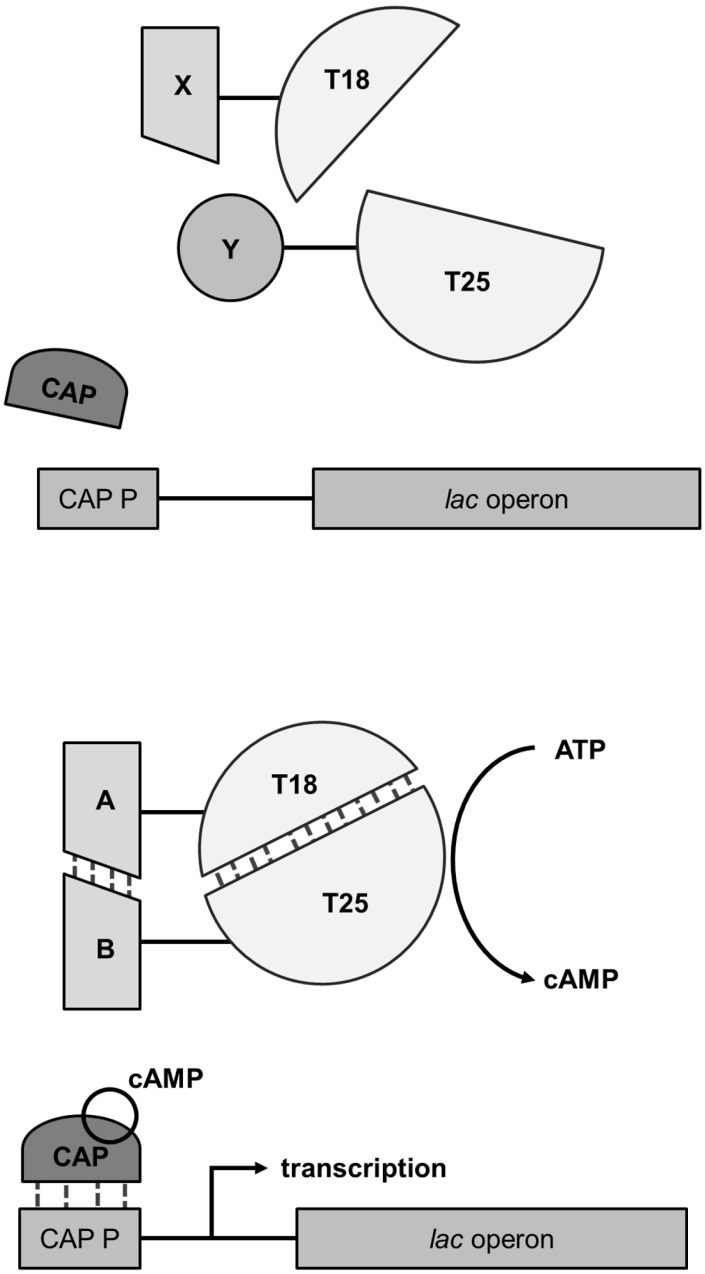


Figure 4.1 *Detection of interacting proteins by bacterial two-hybrid. Where proteins interact (A and B, bottom) adenylate cyclase (T18 and T25) is reconstituted and produces cAMP. cAMP binds to catabolite activating protein (CAP) which, activated, binds the CAP promoter (CAP P) and facilitates transcription of the* lac *operon. Where proteins do not interact (X and Y, top), no cAMP is produced and therefore no transcription occurs.*

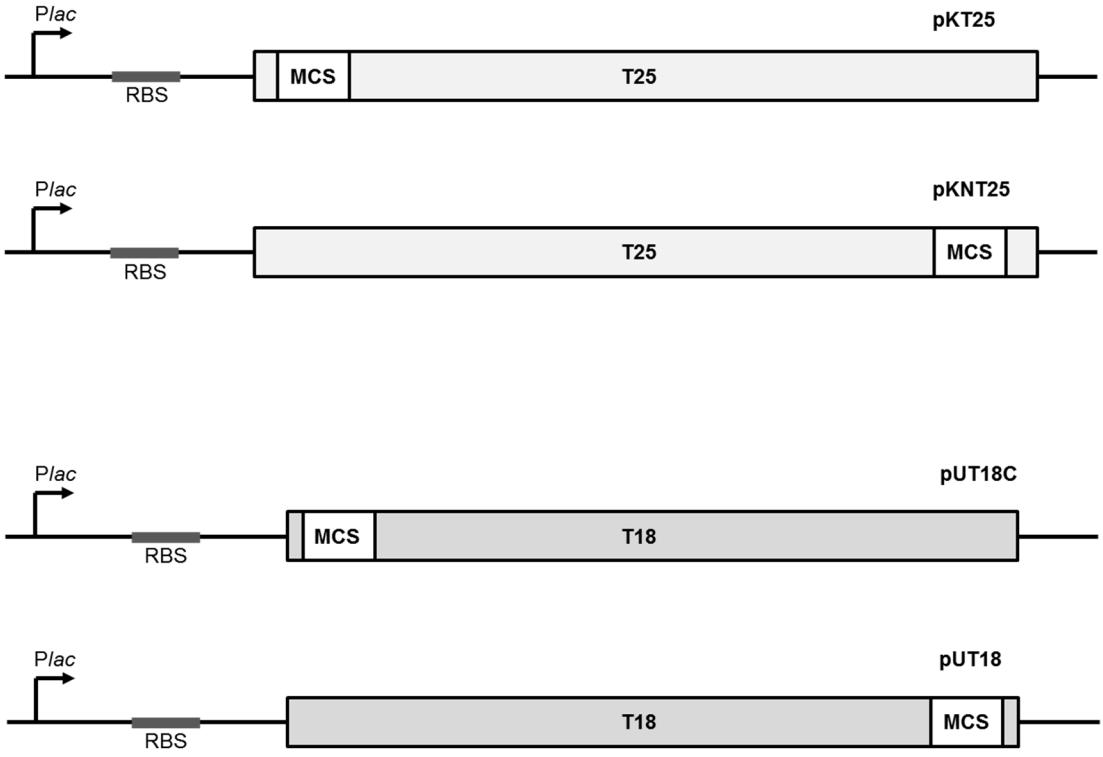


Figure 4.2. *The four vectors used in the bacterial adenylate cyclase system.All four Cya constructs are expressed under the* lac *promoter, P*lac. *Multiple cloning sites (MCS) are at the C-terminus of the* cya *fragments (T25 and T18) in vectors pKT25 and pUT18C or the N-terminus in vectors pKNT25 and pUT18. The MCS are located inside the* cya *coding sequences, so the* cya *start and stop codons initiate and halt translation. Ribosome binding site (RBS) also shown.*

Results

Construction of BACTH vectors containing CspE and HupA

Initially, six constructs were made. Four constructs tagged CspE with the *cya* T18 and T25 fragments at the N- and C-termini, while HupA was tagged with the T18 *cya* fragment at the N- and C-termini. Open reading frames containing *cspE* and *hupA* were amplified from the *S*. Typhimurium SL1344 chromosome using primers *cspE BACTH F* and *hupA BACTH F* primers. For proteins tagged at the N-terminus (ie. into vectors pKT25 and pUT18C), reverse primers *cspE BACTH-stop R* and *hupA BACTH-stop R* were used to amplify the target gene. For proteins tagged at the C-terminus (vectors pKNT25 and pUT18), the native stop codons of *cspE* and *hupA* were omitted using the primers *cspE BACTH-no stop R* and *hupA BACTH-no stop R*. Primer sequences are given in Chapter 2, whilst vectors are described in Figure 4.2.

Amplified fragments and relevant vectors were digested with BamH1 and EcoR1 before ligation and transformation into DH5α by electroporation as described in Chapter 2. Colonies recovered after an overnight incubation were subcultured into overnight cultures for plasmid isolation by miniprep. Minipreps were screened by PCR for presence of the desired insert. PCRs indicating presence of inserts are shown in Figures 4.3 and 4.4. Primer pairs *pKT25 confirm*, *pKNT25 confirm*, *pUT18 confirm*, and *pUT18C confirm* amplify across the plasmid MCS, showing a band of increased size where an insert is present. Having confirmed inserts by PCR, these inserts were sequenced by Sanger sequencing using the primers *pKT25 sequencing*, *pKNT25 sequencing*, *pUT18 sequencing*, and *pUT18C sequencing*; these primers flank the insert sequences. Sequence data is shown in Appendix B.

BACTH fusion plasmids were co-transformed into the BACTH host strain DHM1. A total of eight test strains were created to assay interactions between CspE and HupA and oligomerisation of CspE. The numerous combinations of N- and C-terminal tags reduces the chance of Cya tags disrupting the structure of either protein. The eight strains created are detailed in Table 4.1. Test strains are named according to the proteins assayed and at which terminus they are tagged; first the T25 *cya* construct is given, then the T18. For example, strain ENHC contains CspE tagged with T25 at the N-terminus (using vector pKT25) and HupA tagged with T18 at the C terminus (using vector pUT18).

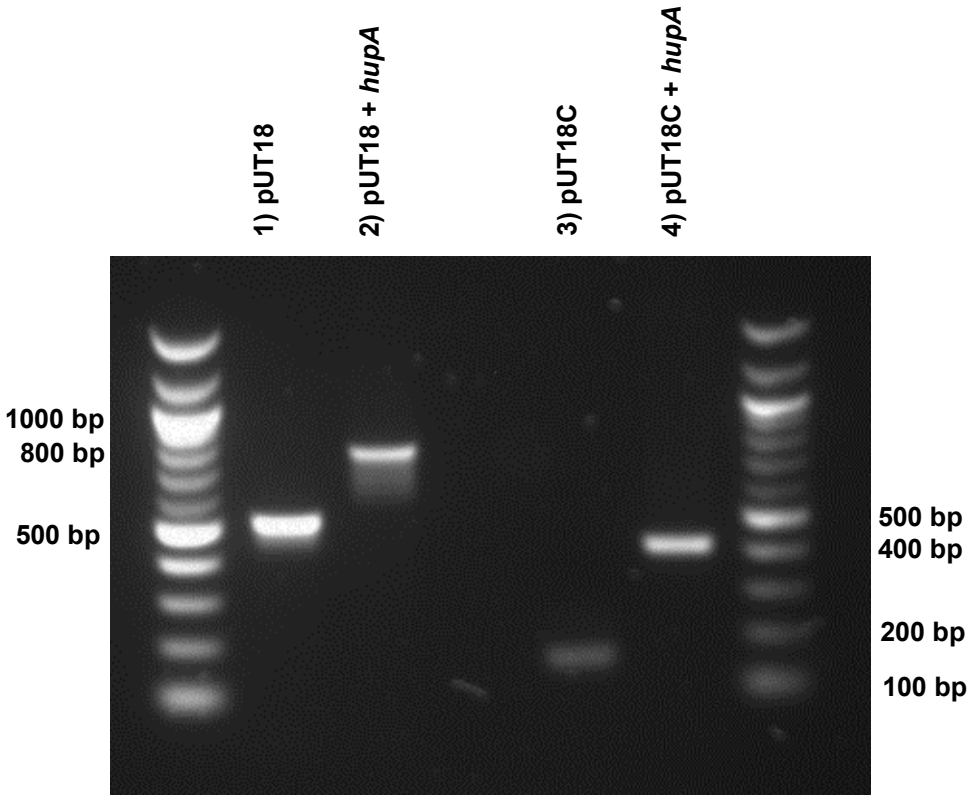


Figure 4.3. *Construction of pUT18-*hupA*, pUT18C-*hupA*. Both vectors amplified across the cloning site by PCR: gene insert is 270 bp in length. Empty pUT18 (1) produces a PCR product just over 500 bp, rising to around 800 bp when* hupA *is inserted (2). Empty pUT18C (3) amplifies a product of 150 bp, rising to a little over 400 bp when* hupA *is inserted (4). PCR products visualised on a 1% gel stained with SYBR® Safe, detected by ultraviolet light.*

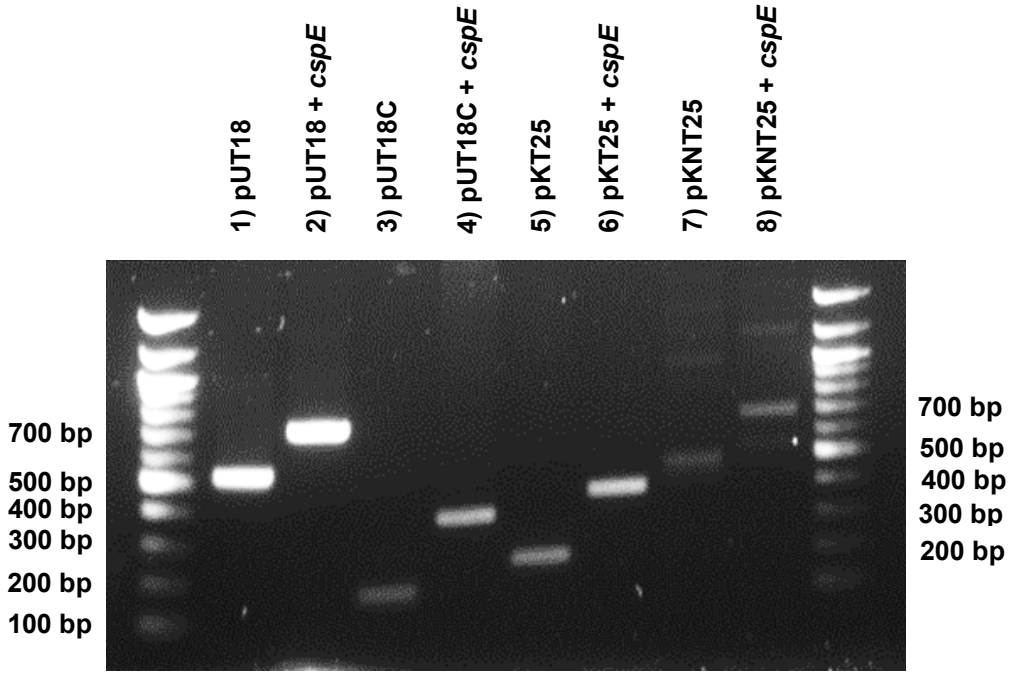


Figure 4.4. *Insertiontion of* cspE *into pKT25, pKNT25, pUT18, pUT18C.Vectors screened across the cloning site by PCR: gene insert is 210 bp in length. Empty pUT18 (1) produces a PCR product just over 500 bp, rising to around 700 bp when* cspE *is inserted (2). Empty pUT18C (3) amplifies a product of 150 bp, rising to ~ 350 bp when* cspE *is inserted (4). Empty pKT25 (5) amplifies around 200 bp, with cspE (6) around 400 bp. pKNT23 runs at ~500 bp when empty (7), and ~700 bp with* cspE *(8). PCR products visualised on a 1% gel stained with SYBR® Safe, detected by ultraviolet light.*

Strain	T25 construct (plasmid)	T18 construct (plasmid)
ENHC	T25-CspE (pKT25)	HupA-T18 (pUT18)
ENHN	T25-CspE (pKT25)	T18-HupA (pUT18C)
ENEC	T25-CspE (pKT25)	CspE-T18 (pUT18)
ENEN	T25-CspE (pKT25)	T18-CspE (pUT18C)
ECHC	CspE-T25 (pKNT25)	HupA-T18 (pUT18)
ECHN	CspE-T25 (pKNT25)	T18-HupA (pUT18C)
ECEC	CspE-T25 (pKNT25)	CspE-T18 (pUT18)
ECEN	CspE-T25 (pKNT25)	T18-CspE (pUT18C)
ENDC	T25-CspE (pKT25)	CsdA-T18 (pUT18)
ENDN	T25-CspE (pKT25)	T18-CsdA (pUT18C)
ECDC	CspE-T25 (pKNT25)	CsdA-T18 (pUT18)
EDDN	CspE-T25 (pKNT25)	T18-CsdA (pUT18C)

Table 4.1. *Names of, protein components of, and expression vectors contained in strains to assay interactions by bacterial two-hybrid.*

Strains were first assayed for interaction on solid medium using the chromogenic lactose analogue X-gal, which is cleaved by β-galactosidase to give a blue-coloured product. Expression of the adenylate cyclase fusions is under control of the *lac* promoter, induced by IPTG. The native chromosomal *lac* operon is not activated by IPTG because of the requirement for the additional presence of cAMP, which is absent in the *cyaA*- host.

To allow quantitation of β-galactosidase activity, a Miller assay was used as described in Chapter 2. Results are displayed in Figures 4.5-8. In order to determine the most reliable method of assaying β-galactosidase activity, Miller assays were performed both in stationary (from overnight cultures) and exponential phase (OD_{600} = 0.4). No significant difference was found between the two methods, as illustrated by the data illustrated in Figures 4.5 and 4.7. Although the absolute values obtained in stationary phase were lower than those in exponential, the trends described remained the same.

The method used for quantitating the activity of β-galactosidase in samples is essentially a standard Miller assay (Sambrook et al. 1989); exact conditions are described in Chapter 2. Cells are grown to the correct optical density, pelleted, and resuspended in assay buffer. Lysis is achieved by addition of chloroform and SDS and lysates are clarified to remove debris by a second centrifugation. The β-galactosidase activity of lysates is determined by addition of a chromogenic substrate ONPG, which is cleaved to yield a yellow colour. Miller units, which are an arbitrary measure of β-galactosidase activity, are calculated from the density of yellow colour (OD_{420}) and the incubation time taken for colour to arise. Corrections are made for culture density (OD_{600}) and scatter due to cell debris (1.75 x OD_{550}). The equation is as follows, where *t* is incubation time in minutes and *d* is the dilution factor of the lysate.

$$\text{Miller units} = 1000 \times \frac{OD_{420} - (1.75 \times OD_{550})}{t \times d \times OD_{600}}$$

The definition of a positive result varies from study to study. The authors of the original paper (Karimova et al. 2005) suggest an interaction producing 4 to 5 times the β-galactosidase activity of the negative control. Others have suggested 3 times the negative control (Thanikkal et al. 2012) or any interaction with a significantly different value to the negative control (Li et al. 2010). The positive control consists of two halves of a *Saccharomyces* leucine zipper transcription factor, which exhibit exceptionally

high binding affinity. It is rare to obtain β-galactosidase activity approaching that of the positive control. One study of a two-component system (which could be expected to interact with high affinity) observed activity at 61% of that of the positive control (Scheu et al. 2012).

It should be noted that due to the CspE and HupA adenylate cyclase fusions being expressed from plasmids, rather than their native promoters, the effect of growth phase on expression is minimal. If there were growth-phase dependent differences in interactions to be observed in the native *S*. Typhimurium, they would likely not be observed in this system. The BACTH system assays whether two proteins are capable of interacting but cannot account for factors that might modify such interaction in the native host, such as other modulators (other proteins or nucleic acids) or changes in expression. However, given the similarity between *E. coli* and *S*. Typhimurium (their CspEs differ in only two amino acids) it might be expected that any such third-party modulators present in the host could cross-function with the *S*. Typhimurium CspE construct. The non-specific nature of CspE's polynucleotide binding is of benefit here; the constructs should bind to *E. coli* RNA and ssDNA. Thus, if an interaction requires a polynucleotide substrate or scaffold, it should still be observed.

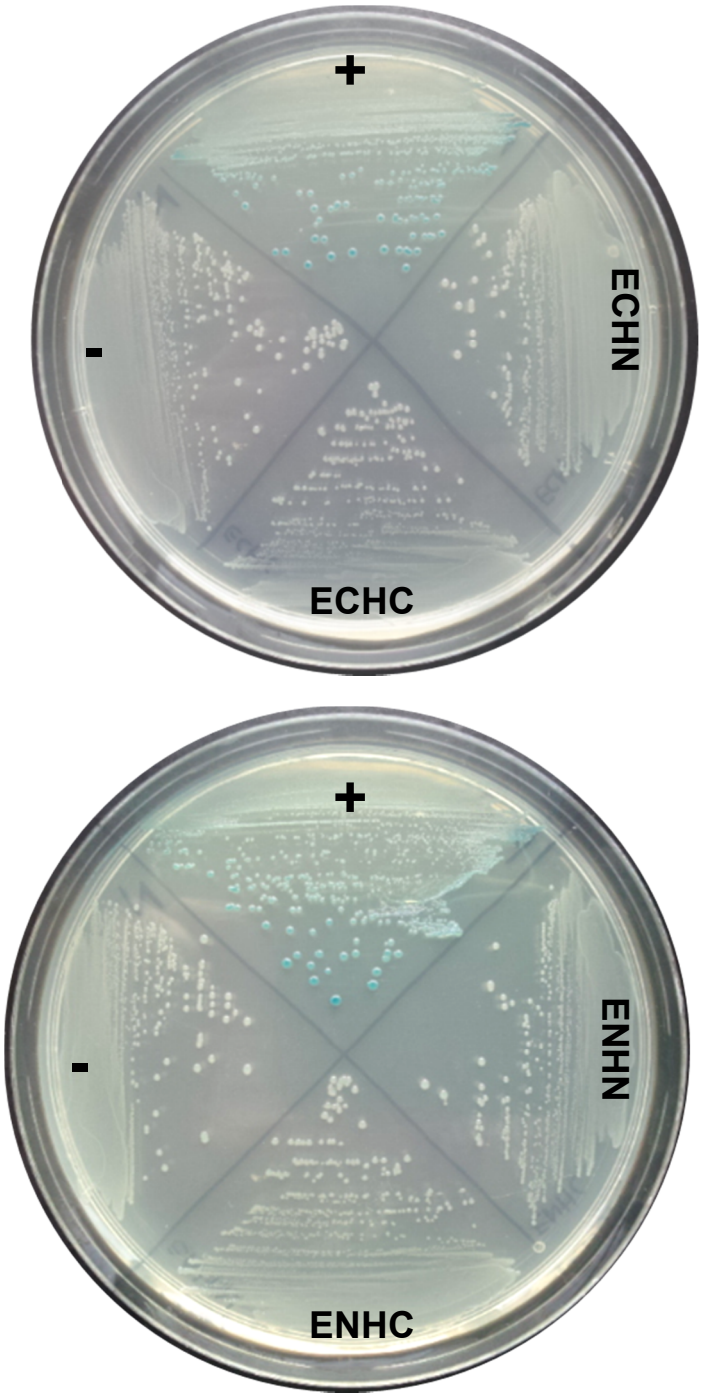


Figure 4.5. *Qualitative colourimetric assay for interaction between CspE and HupA on LB-X-gal plates grown overnight at 37°C. Positive control (+) demonstrates colour change to blue as a result of X-gal cleavage. Strains ECHC, ECHN, ENHC, ENHN show no evidence of CspE-HupA interaction, remaining white; the same as the negative control strain (-).*

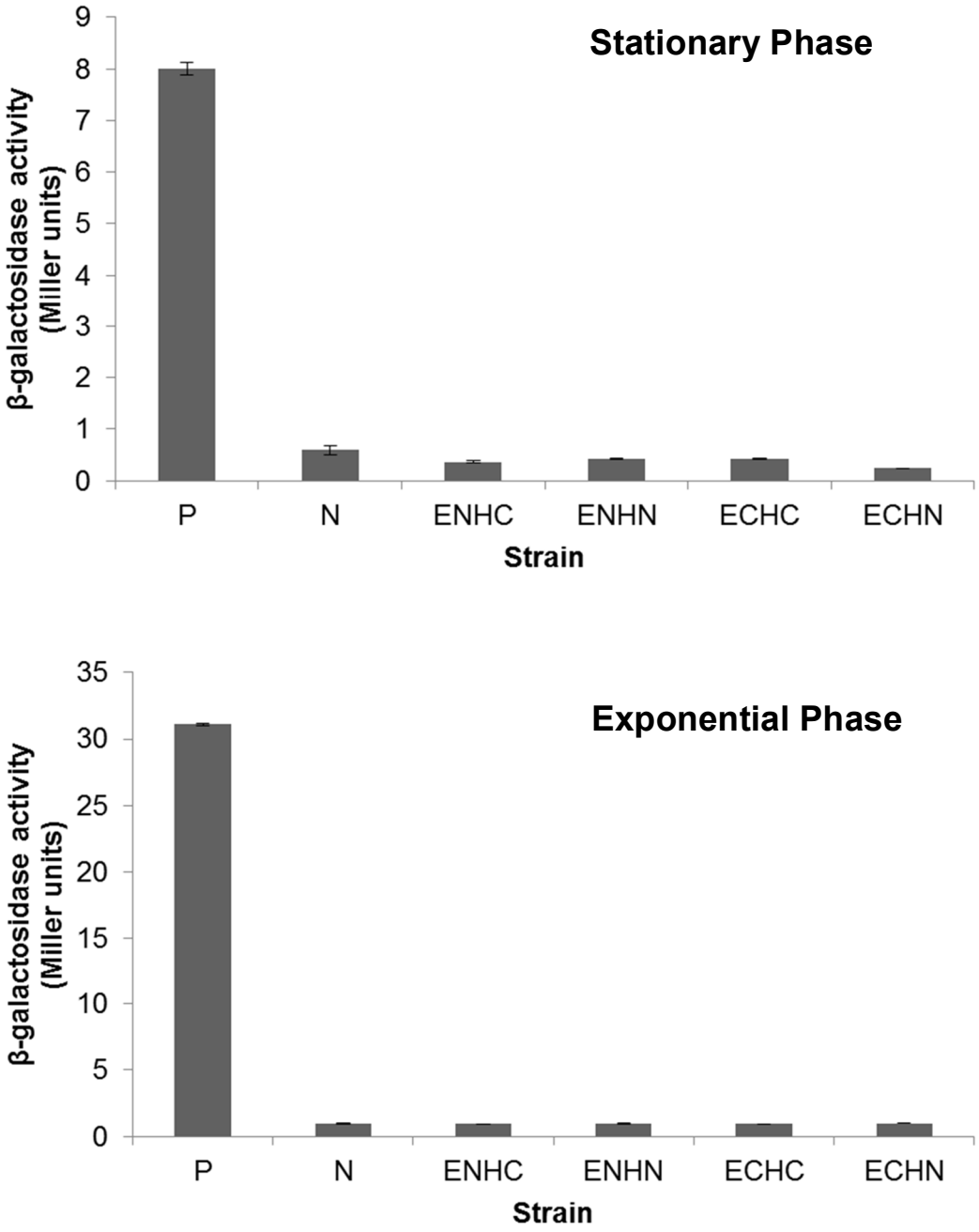


Figure 4.6. *Quantitative assay of β-galactosidase activity (given in Miller units) in strains ECHC, ECHN, ENHC, ENHN, assaying for a CspE-HupA interaction. All four test strains show similar levels of β-galactosidase activity to the negative control (N) in both stationary (top) and exponential phase (bottom). Positive control (P) is around 8-fold higher. Error bars represent standard deviation from three samples prepared identically. Graphs shown are representative of three experimental repeats.*

Further Studies of CspE Interactors

HupA does not interact with CspE at 37°C

Quantitative and qualitative assays of β-galactosidase activity were performed in order to determine the presence or absence of an interaction between CspE and HupA. Both plate and Miller assays showed the same result. Miller assays were performed using stationary and exponential phase cells in order to determine the most reliable time at which to take readings; the two both produced the same results. No evidence was found to support an interaction between CspE and HupA at 37°C. Results from the β-galactosidase assays testing this interaction are shown in Figures 4.5 and 4.6.

CspE functions as a monomer at 37°C

Both CspE dimerisation (Johnston et al. 2006) and polymerisation (Phadtare 2005) have been reported *in vitro*, but not *in vivo*. In the first case, it is suggested that dimerisation is a mechanism of chromosome condensation; CspE binds chromosomal ssDNA and contributes to condensation by dimerising. This model is derived from the observation that overexpression of CspE protects against camphor-induced chromosome decondensation (Hu et al. 1996). No evidence was found to indicate the formation of CspE oligomers in the conditions assayed. The results in Figures 4.7 and 4.8 suggest CspE is present as a monomer at 37°C, indicating that the proposed chromosome condensation function is not a significant CspE activity under the conditions examined here. Alternatively, the observed protection may not have been a direct activity of CspE, and could have been the result of a CspE-mediated gene regulatory event.

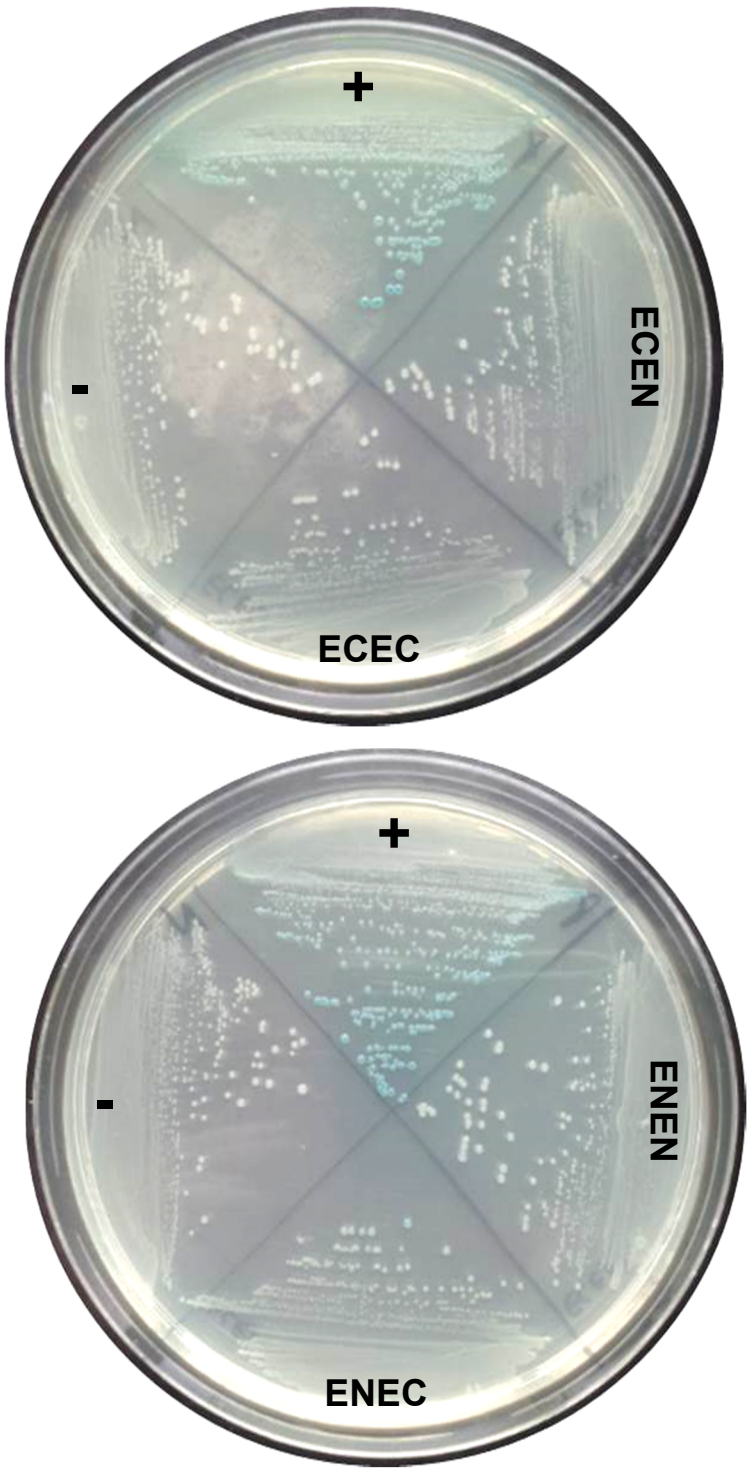


Figure 4.7. *Qualitative colourimetric assay for CspE oligomerisation on LB-X-gal plates grown overnight at 37°C. Positive control (+) demonstrates colour change to blue as a result of X-gal cleavage. Strains ECEC, ECEN, ENEC, ENEN show no evidence of CspE oligomerisation, remaining white; the same as the negative control strain (-).*

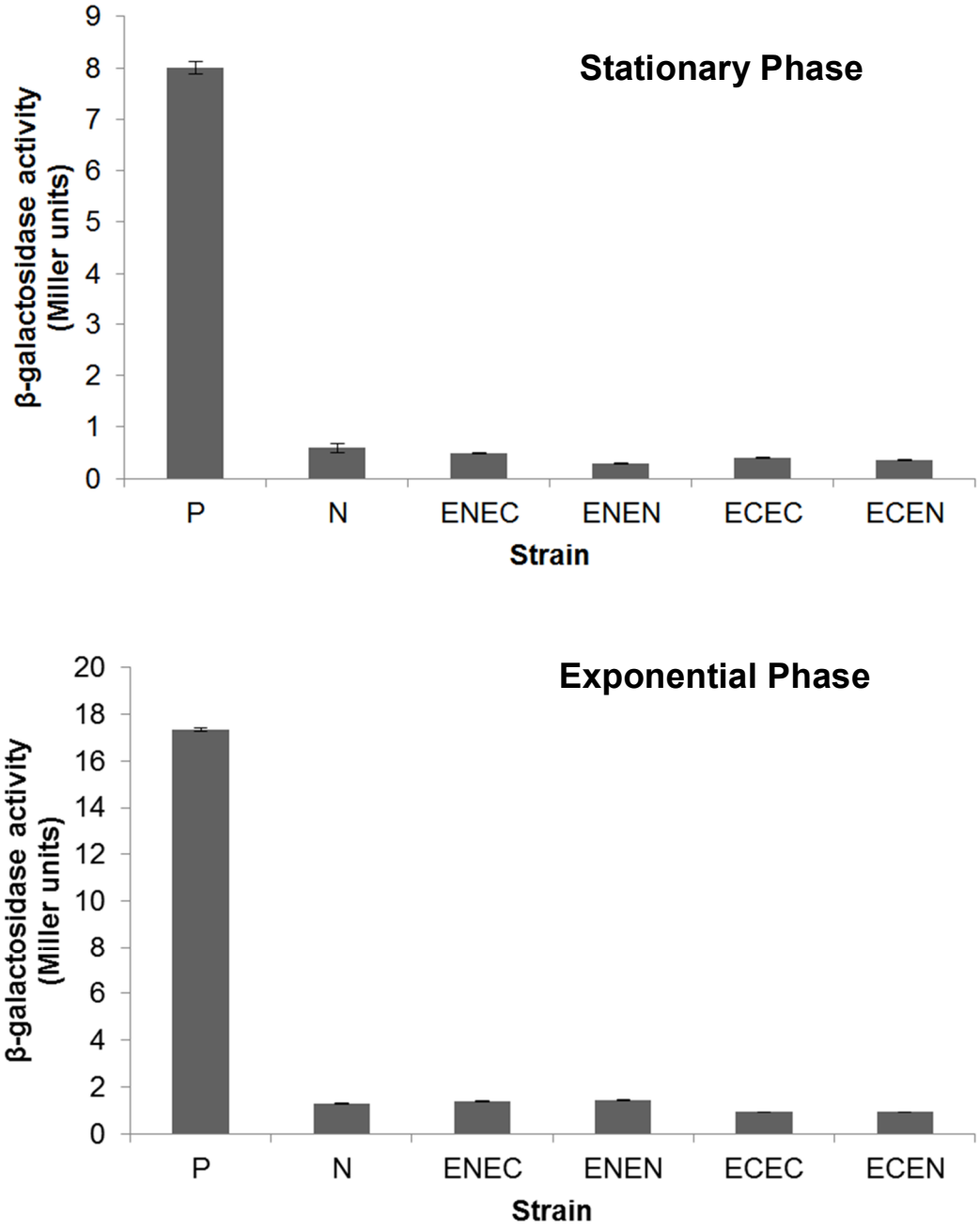


Figure 4.8. *Quantitative assay of β-galactosidase activity (given in Miller units) in strains ENEC, ENEN, ECEC, ECEN, assaying for CspE oligomerisation. All four test strains show similar levels of β-galactosidase activity to the negative control (N) in both stationary (top) and exponential phase (bottom). Positive control (P) is around 8-fold higher. Error bars represent standard deviation from three samples prepared identically. Graphs shown are representative of three experimental repeats.*

Investigation of CspE – CsdA interaction

The cold shock helicase CsdA (also called DeaD) is part of the DEAD-box family of RNA helicases, named for a characteristic amino acid motif. The ATP-dependent helicase has a wide array of functions (mostly described in *E. coli*), both in cold shock and normal cell growth. Induced during cold shock, the helicase has specific functions in ribosome maturation (Charollais et al. 2004) and expression of *rpoS* (Resch et al. 2010b) as well as a general function in RNA turnover as part of a modified cold-shock RNA degrading complex (Prud'homme-Généreux et al. 2004). CsdA's function during normal growth is still being characterised but a recent study of gene expression suggested regulation of genes in pathways as diverse as heat shock, carbon metabolism, cell division, and manganese homeostasis (Vakulskas et al. 2014).

Like most of the other DEAD-box helicases, CsdA has many functions but apparently little substrate specificity *in vitro*; it is believed specificity *in vivo* is derived from regulatory protein interactions (Iost & Dreyfus 2006). Indeed, at low temperatures, the *Bacillus subtilis* paralogues CshA and CshB appear to interact directly with the cold shock protein CspB (Hunger et al. 2006). Interactions of CsdA are therefore of interest both in and out of cold shock. In order to determine whether *Salmonella*'s CspE and CsdA interact , CsdA was cloned into the two-hybrid vectors pUT18 and pUT18C as before (Figure 4.9). Strains ENDC, ENDN, ECDC, and ECDN were created, as described in Table 4.1. Subsequent growth and β-galactosidase assays are presented in Figure 4.10. No evidence for interaction was observed between CspE and CsdA at 37^{o}C.

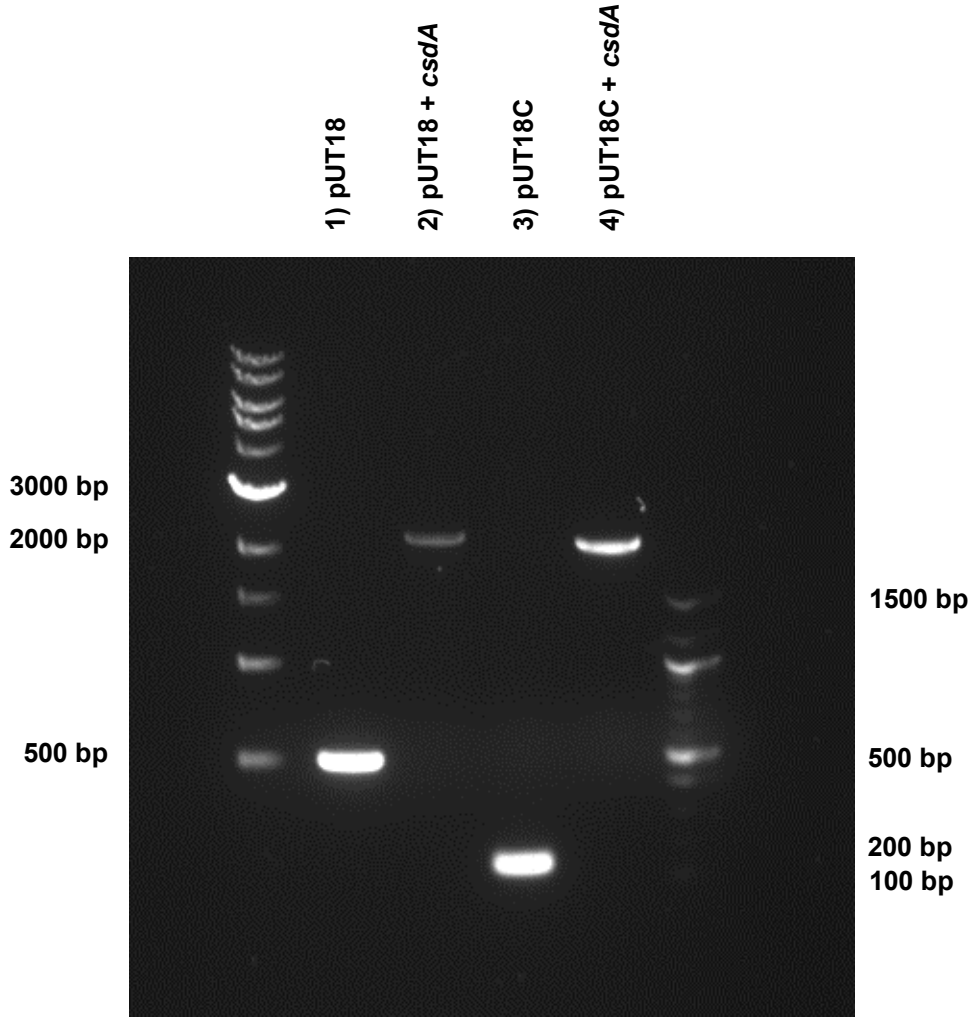


Figure 4.9. *Construction of pUT18-*csdA*, pUT18C-*csdA*. Both vectors screened across the cloning site by PCR: gene insert is 1850 bp in length. Empty pUT18 (1) produces a PCR product just over 500 bp, rising to over 2000 bp when* csdA *is inserted (2). Empty pUT18C (3) amplifies a product of 150 bp, rising to 2000 bp when* csdA *is inserted (4). PCR products visualised on a 0.8% gel stained with SYBR® Safe, detected by ultraviolet light.*

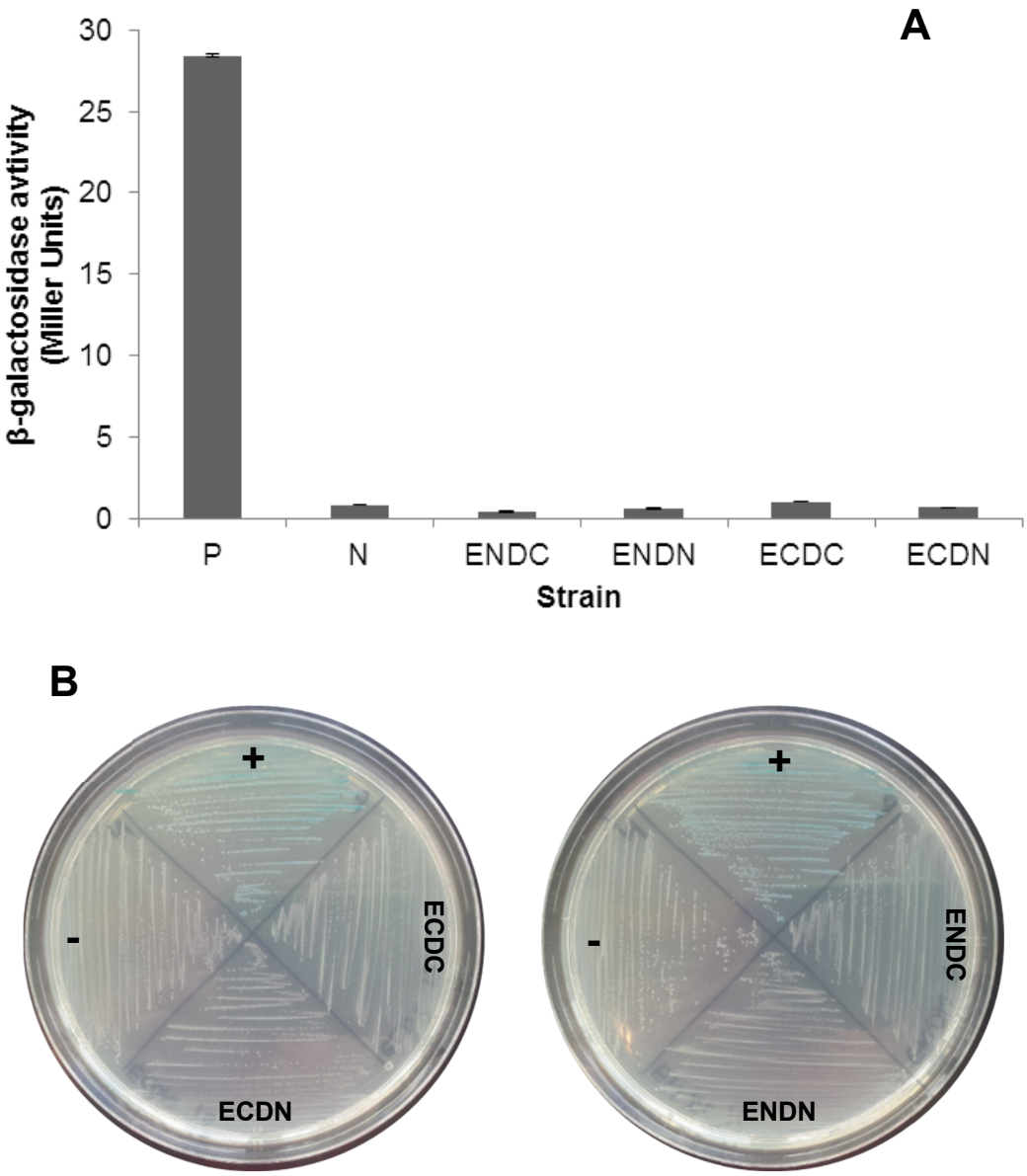


Figure 4.10. A) *Quantitative assay of β-galactosidase activity (given in Miller units) in strains ENDC, ENDN, ECDC, ECDN, assaying for CspE oligomerisation. All four test strains show similar levels of β-galactosidase activity to the negative control (N) in stationary phase. Positive control (P) is significantly higher. Error bars represent standard deviation from three samples prepared identically. Graph representative of three experimental repeats. B) Plate assays CspE – CsdA interaction, performed as previously. Strains ENDC, ENDN, ECDC, ECDN show no development of blue pigment.*

Further Studies of CspE Interactors

Interactions at low temperature

The protein-protein interactions of CspE at 37°C are unknown and are interesting avenues of study. The functions of the cold shock proteins at low temperatures, although better understood, are not yet fully characterised. The systems created here allow the opportunity to investigate CspE oligomerisation, as well as its interaction with HupA and CsdA, at low temperatures.

Again, the assays presented are not strictly representative of the functions of the relevant proteins during the cold shock response of *S.* Typhimurium. Being overexpressed and outside of their usual host, the determinants of interaction are different in this system. Rather, the following experiments will determine whether the intrinsic affinity of the proteins assayed is increased at low temperature. This could be the result of a temperature-dependent structural rearrangement (such as observed for TlpA, discussed in Chapter 1 (Hurme et al. 1997)) or perhaps a structural change in an RNA or ssDNA substrate which mediates interaction. This could be a result of, for example, the stabilisation of an mRNA secondary structure or the alteration of chromosomal supercoiling at low temperature.

The effect of low temperature was assayed by growing each of the 12 test strains to stationary phase (overnight at 37°C) before shifting them to 10°C for a further night of growth. *E. coli* can grow at 10°C but exhibits a longer lag phase (Kim et al. 2014), hence the need to start cultures at 37°C. Given that growth can be observed, the transcription and translation of β-galactosidase required to detect an interaction is known to occur at 10°C, albeit likely at a slower rate. The incubation at lower temperatures of ~18 hours is long enough to allow for expression of β-galactosidase, as evidenced by the signals detected in the positive control.

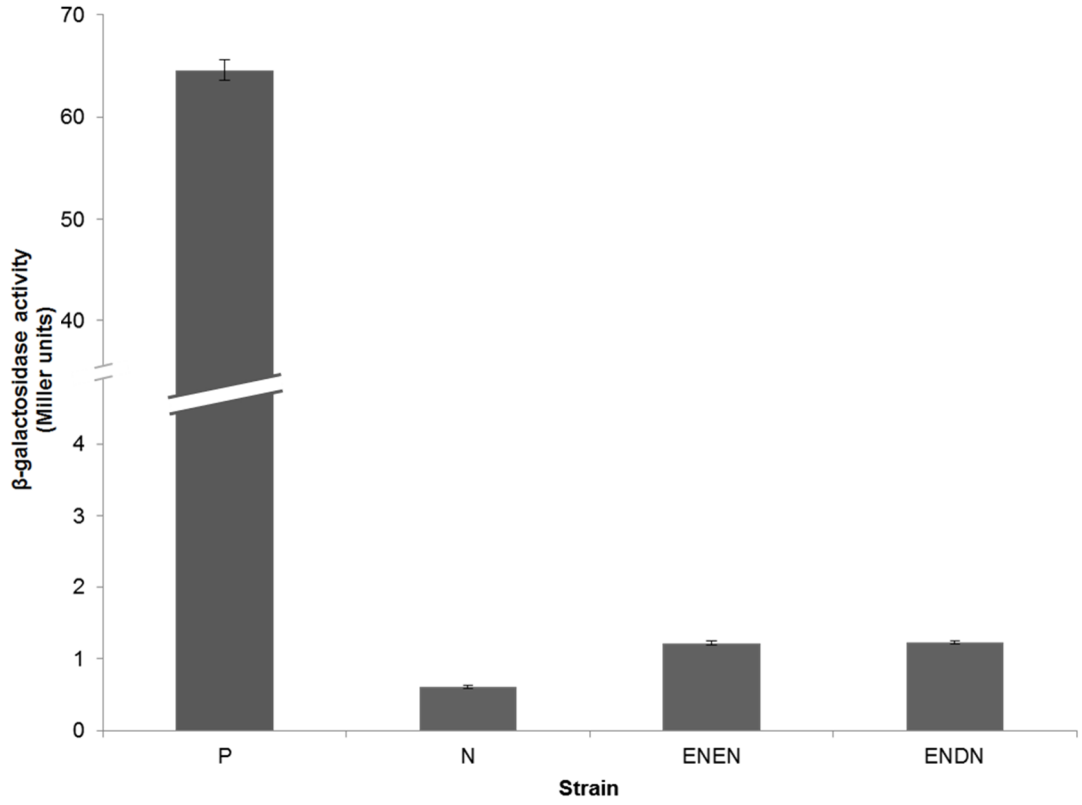


Figure 4.11. *Assay of β-galactosidase activity in strains ENEN and ENDN grown at low temperatures (10°C).The two tested strains give around twice the β-galactosidase activity of the negative control (1.2 Miller units to 0.6), while the positive control is considerably higher. Positive control broken here,to allow better comparison of the lower values. The true positive control value is shown above (64.5 Miller units, with a standard deviation of 1 unit). Data presented representative of six repeats, with error bars showing standard deviation across the three replicates of the repeat presented*.

All 12 strains were assayed in this manner. Although most remained negative, two (ENEN and ENDN) displayed a slight increase in interaction at low temperatures. Across six repeats, ENEN produced β-galactosidase at an average rate 1.6 times that of the negative control, and ENDN at 2.0 times the negative control. Both are modest increases, but statistically significant (as determined by a 1-tailed T test in MS Excel 2010). For ENEN, T stat was -2.35, while T critical was 1.94. For ENDN, T stat was -2.31 while T critical was 2.02. As the differences in β-galactosidase activity were low, and in order to ensure reproducibility, each of these assays was repeated 6 times rather than 3. A third strain, ECHC, also suggested a slight interaction but after sextuplet repeats this did not prove to be significant. A graph showing the β-galactosidase activity of these two strains is shown in Figure 4.10.

Discussion

CspE and HupA

HupA was selected for further study based on its co-purification with CspE in the previous cross-linking mass spectrometry experiment. As the results of that experiment were inconclusive, further work was required to confirm or rule out an interaction between CspE and HupA. Although there is no previous evidence for a direct interaction between the two proteins, they both share an ability to bind ssDNA (Johnston et al. 2006; Kamashev et al. 2008), raising the possibility of an interaction being mediated by this shared substrate. However, the observed co-purification may also have arisen simply due to their proximity on this shared substrate. In order to determine which of these two possibilities is the case *in vivo*, the potential for direct interaction between HupA and CspE was probed using BACTH. No evidence for interaction was detected at either 37°C or 10°C.

HU forms homo- or heterodimers *in vivo*, with the subunit composition altering according to growth phase and temperature, as discussed in Chapter 3. However, it is likely that the quaternary structure during DNA binding is more complex. HU has been crystallised as an octamer, composed of four dimers arranged in a number of possible conformations (Guo & Adhya 2007). These structures are both complex and compact, with DNA being tightly coiled around the exterior surface, as illustrated in Figure 4.10. Given the complexity of the structure, it is likely that the addition of the adenylate cyclase T18 fragment would disrupt either the protein-protein or protein-DNA interactions required for correct HU function. While this would not impact on cell function or viability (the DHM1 host retains chromosomal copies of the *hupA* and *hupB* proteins), it may limit the chances of detecting an interaction. HupA was tagged with T18 at both the C and N termini, but both may impede function. A C-terminal tag could disrupt protein-protein interactions at the core of the HU octamer, whereas an N-terminal tag may interfere with DNA-binding. Although the image presented in Figure 4.10 shows binding of dsDNA, ssDNA is proposed to share the same binding surface (Kamashev et al. 2008).

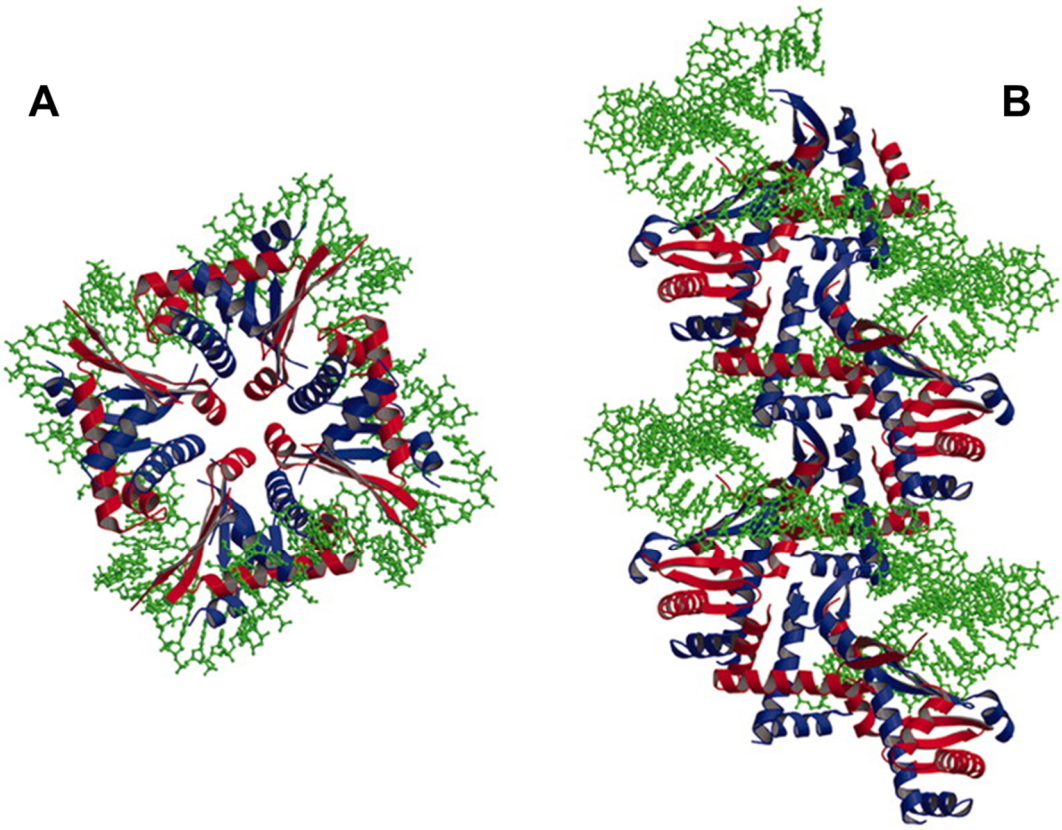


Figure 4.12. *Complex octameric structure of the DNA binding protein HU shown from the top (A) and the side (B). HU subunits HupA and HupB shown in red and blue, DNA molecule in green. C-termini of both subnits are located toward the middle of the structure, with N-termini outwards, towards the DNA-binding surface. Images adapted from models generated in Guo and Adhya, 2007.*

Given that HupA functions exclusively as part of the HU dimer, an *in vivo* interaction between CspE and HupA would likely be dependent on the correct formation of that dimer. The loss of dimerisation (or octamerisation) may well result in the loss of CspE interaction. Another possibility from the CLMS experiment is a co-localisation of CspE and HupA to a common substrate. If this is the case, then loss of HU's ability to bind DNA would result in a loss of that apparent interaction.

Chromosomal *hupA* (HU-α) is downregulated at low temperatures in favour of expression of *hupB* (HU-β) (Giangrossi et al. 2002), although the plasmid-expressed HupA construct tested here will be unaffected. The functional differences between the α and β HU subunits is unclear although the subunits are regulated differently (Bahloul et al. 2001; Giangrossi et al. 2002). The selective expression of HupB at low temperature suggests some functional discrepancy, however, and if the HupB were to interact with CspE, that interaction would not be detected here. The cross-linking mass spectrometry experiment only identified HupA, likely because the cross-linking was carried out at 37°C. A possibility of HupB interaction with CspE at low temperature remains.

This second experiment (Figures 4.5, 4.6) provided no additional data to support the CspE-HupA interaction suggested by the first (Chapter 3). CspE does not appear to interact with the HU subunit HupA. An association with the complete HU protein remains possible, as it may have been missed by the BACTH system due to impairment of formation and function of the higher structure. Therefore, the co-localisation of CspE and HU to ssDNA substrates remains a possibility.

The results from the previous cross-linking mass spectrometry experiment did not reveal any clear interacting proteins, but suggested a localisation of CspE around the nucleoid and the ribosome (Figure 3.7). The evidence here supports the theory that the CspE-HupA association is circumstantial rather than specific, as no strong interaction was detected between the two. Other hits from the library may therefore be accounted for by co-localisation in crowded sites such as the ribosomal entry site rather than from specific interactions.

Further Studies of CspE Interactors

CspE and chromosome condensation

One of the initial characterisations of CspE was as a contributor to chromosome condensation, which gives a phenotype of resistance to camphor toxicity (Hu et al. 1996). It was later proposed that CspE achieves this condensation effect by binding single stranded regions of chromosomal DNA and dimerising (Johnston et al. 2006), although no *in vivo* evidence has yet supported that theory. The results presented here suggest that CspE does not dimerise *in vivo* at 37°C, suggesting that another mechanism must account for chromosome condensation. The same mechanism may be achieved by CspE's binding to another protein not assayed here, or CspE may contribute to condensation as a monomer. Interestingly, the protective effect of CspE against camphor was only observed in the presence of two other genes co-transcribed with *cspE*: *pagP* and *crcB*. Both encode membrane proteins, however, with little obvious role in the nucleoid (Hu et al. 1996).

Condensation is a result of specific DNA-binding proteins and molecular crowding effects (ie. pressure exerted by the density of proteins and other macromolecules in the cell (de Vries 2010)). Thus, the protective effect of CspE, PagP, and CrcB could simply be a result of their overproduction in the experiment, although that would not necessarily account for the need for all three proteins. The presence of globular proteins alone is not usually sufficient for condensation (de Vries 2010) and most current models of condensation require formation of some sort of bridging dimer, whether along the length of one DNA strand or across two separate strands (Broedersz et al. 2014). This suggests CspE interaction with another protein may be required, but that protein is neither HupA or CspE itself, as determined here. Interaction with another Csp paralogue (CspC and CspD are both expressed in stationary phase (Czapski & Trun 2014)) is possible, and would not have been detected by this experiment.

CspE as an antiterminator

The current model for CspE-mediated melting of RNA secondary structures involves an initial binding of a single protein to an RNA stem-loop junction followed by unwinding of the structure by subsequent binding of monomers (Phadtare 2005). While not a true polymerisation, this mechanism would be expected to result in CspE-CspE interactions as many copies coat the unfolding RNA (a single CspE molecule

covers only around 6 nucleotides). Again, this is not supported (at 37°C) by the evidence presented. mRNA secondary structures can be expected to be less common at 37°C than at low temperatures, so it may be the case that CspE's unwinding function is less important at higher temperatures.

If nucleic acid unwinding is considered a more important function of CspE at low temperatures than at 37°C, then other functions must be more prominent at the higher temperature. A further activity of the cold shock proteins is transcription antitermination (Bae et al. 2000). This function is known to be dependent on the ability of CspE to melt RNA secondary structures (Phadtare 2001), suggesting a mechanism wherein CspE binds to secondary structures formed by terminators and melts them, allowing progression of RNA polymerase. Studies of the CspE melting mechanism suggested that although stem-loops with a loop of 8-61 nt required multiple copies of CspE to unwind, those with loops of 4 nt could be melted by a single CspE monomer. Given that a majority of rho-independent terminators feature loops of 4 nt and stems of 5-7 nt (Lesnik et al. 2001), the antitermination function of CspE is likely to be largely performed by monomeric protein. The results presented suggest that the major function of CspE at 37°C could be that of a transcriptional antiterminator, although more work is needed to confirm this.

Deletion of *cspE* in *E. coli* revealed a number of genes which are dependent on that gene for expression during growth at 37°C. Further studies using a mutant deficient in RNA melting, but not binding, demonstrated that the melting activity (and therefore presumably transcription antitermination) is the means by which these genes are regulated (Phadtare et al. 2006). Transcriptional antitermination is thus a means of gene regulation during normal growth as well as during cold shock. The results presented here suggest that it may be this function, rather than non-specific RNA chaperoning, which predominates at 37°C.

As an exclusively prokaryotic process, transcription antitermination is a potential target for inhibition by antimicrobials. The antimicrobial bicyclomycin acts as an inhibitor of the antiterminator rho (Zwiefka et al. 1993). Bismuth salts, which are widely used to treat gastric and peptic ulcers, also prevent rho function and target a broad range of pathogens (Brogan et al. 2005). Being an enzymatic process, Rho-dependent antitermination may be easier to target than rho-independent antitermination, but targeting protein-RNA interactions is a possibility. Inhibitors of

protein-RNA interaction are being studied as antiviral therapies (Alfadhli et al. 2013; Ellenbecker et al. 2012).

CspE at low temperature

Assaying for CspE oligomerisation, a small but significant interaction was detected in strain ENEN. Using BACTH, it is impossible to determine the stoichiometry of such an interaction, but dimerisation of CspE has previously been reported (Morgan et al. 2009). A significant dimer formation would be expected to result in greater β-galactosidase activity in the assay strains, however. One similar experiment testing dimerisation of Fnr reported a tenfold increase in β-galactosidase activity over a negative control (Jervis & Green 2007). The detection of a weak interaction is perhaps consistent with a co-localisation along the length of an RNA molecule, or a stretch of ssDNA. Such an aggregation of CspE might bring the monomers into close enough proximity to produce a weak adenylate cyclase activity. The absence of such an activity at 37°C could be explained by the lower frequency of RNA secondary structure formation at the higher temperature.

It may be significant that both CspE constructs for which an interaction was detected were N-terminally tagged (pUT28C-*cspE*, pKT25-*cspE*). Although the C- and N-termini of CspE are located close to each other (on the opposite face to the nucleotide binding surface) it is possible that addition of a tag to the C-terminus has a destabilising effect not seen with an N-terminal tag. It should be noted that a tag at either terminus is likely to leave the nucleotide binding surface unaffected, and so RNA and DNA binding ought not (in principal) be affected.

As well as binding RNA, Csp proteins *in vitro* can also bind ssDNA (Johnston et al. 2006). Thus, ssDNA in the nucleoid might be another substrate around which CspE monomers might associate. Strand opening is an endothermic process and therefore less efficient at low temperature, and part of the cold shock response is dedicated to facilitating strand opening (Hatfield & Benham 2002). A possible reason for CspE binding to ssDNA at low temperature might be to prevent reannealing once the double helix has been opened for transcription or DNA replication. CspA has been reported to facilitate *hns* transcription at low temperature by facilitating or maintaining strand opening (M. Giangrossi et al. 2001). It is presumed in that paper that CspE directly

promotes RNA polymerase binding although the data also fits a model where some other factor promotes opening which is subsequently maintained by CspE, facilitating RNA polymerase binding and transcription.

CspE and CsdA

Examination of the CspE-CsdA interaction revealed a similar situation to that of the CspE-CspE autoassociation. Again, no interaction was observed at 37°C with evidence for a slight interaction being seen at 10°C. Given that the role of CsdA at low temperatures is broadly the same as that of the cold shock proteins (unwinding RNA secondary structures) it is possible that the modest interaction observed is again the result of localisation around a common substrate. CsdA forms part of the RNA-degrading apparatus at low temperature (Prud'homme-Généreux et al. 2004), an apparatus with which CspE has been demonstrated to interact functionally at higher temperatures by repressing digestion of certain transcripts (Feng 2001). For example, CspE upregulates *rpoS* by impeding RNAse R digestion of the mRNA during stress conditions (Liu et al. 2007). CsdA upregulates *rpoS* during cold shock by unwinding a secondary structure to facilitate translation (Resch et al. 2010b); it is reasonable to assume CspE and the other cold shock proteins may have similar activity in the cold.

At 37°C, CsdA facilitates translation of more than 30 genes by interaction with their mRNAs (Vakulskas et al. 2014) and plays a role in the maturation of the 50S ribosomal subunit (Peil et al. 2008). If, as suggested above, CspE functions mainly in the nucleoid as an antiterminator at 37°C then CspE and CsdA would be unlikely to either co-localise or interact functionally, explaining the lack of β-galactosidase activity at that higher temperature. In contrast to the earlier work (Hunger et al. 2006), another investigation in *B. subtilis* of the interactions of a DEAD-box helicase found no evidence for interaction with Csp proteins during normal growth (Lehnik-Habrink et al. 2013).

Research using the Gram-positive model *B. subtilis* has studied the interaction between CspB and a pair of DEAD-box helicases (CshA and B) analogous to CsdA of *E. coli* and *S.* Typhimurium (Hunger et al. 2006). The study found that both helicases localise near to the nucleoid in a manner dependent on active transcription, a finding which mirrors the localisation of the cold shock proteins (Weber, Volkov, et al. 2001). Using fluorescence resonance energy transfer, interaction of CshA and CshB with CspB

was demonstrated at low temperatures. In the assayed strain, fluorescence emission was 28% higher than that of a negative control, which was deemed to be a significant and strong interaction. Although that increase (in fluorescence) is lower than the one reported here (in β-galactosidase activity), direct comparison of the two methods is difficult. The interaction observed was also determined to be dependent on active transcription. FRET detects interactions between proteins up to 10 nm apart (Willemse et al., 2011), which may be a greater distance than detectable by BACTH, although the detection distance for BACTH is not known precisely.

The authors propose a model wherein the DEAD-box helicase unwinds mRNA secondary structures, which are then coated by Csp proteins to prevent refolding (Hunger et al. 2006). This model would fit the weak interactions observed here, as it would bring CspE proximate to CsdA without requiring the formation of a tight interacting surface, which would generate a strong β-galactosidase signal. CspE constructs have been made which are deficient in nucleotide binding and unwinding (McGibbon 2013; Phadtare et al. 2002). The models suggested for CspE-CspE and CspE-CsdA interaction could be confirmed by repeating the low-temperature assays with a CspE protein unable to bind DNA. If the model were correct, the weak interactions observed here would not be observed in the absence of DNA binding. The main focus of this work, however, remains the function of CspE at 37^{o}C.

Conclusions

The bacterial adenylate cyclase two-hybrid system was used to probe CspE oligomerisation as well as possible CspE interactions with HupA and CsdA. At 37^{o}C, no evidence of interaction for any of these combinations was detected. At 10^{o}C, a slight but statistically significant interaction was detected between two CspE constructs, and CspE and CsdA constructs. Interaction between CspE and HupA was not detected at either temperature, although this may have been due to a perturbation of HupA function with the presence of the adenylate cyclase tags.

The observed weak interactions between CspE-CspE and CspE-CsdA fit a model wherein, at low temperatures, the cold shock proteins coat mRNA molecules in order to prevent the formation of secondary structures. CsdA would be proximate and facilitate the unwinding of complex structures. This would account for the weak interactions as

the proteins would be brought into close proximity without forming the tight bipartite interactions which would yield a strong signal using BACTH.

CspE appears to exist as a monomer at 37°C. This finding contradicts previous *in vitro* work which suggested CspE may be a functional dimer (Johnston et al. 2006). One function of the cold shock proteins correlates to a monomeric population; transcription antitermination. Thus, it is proposed that the primary function of CspE at 37°C may be that of a housekeeping antiterminator. It is also a possibility that CspE is capable of dimerising, but is prevented from doing so *in vivo* (under the conditions tested) by an unknown factor.

While the observations described above concur with current models for CspE function at low temperature, they offer little development into the question of cold shock protein specificity at 37°C. The substrate specificity demonstrated by the cold shock proteins (McGibbon 2013) remains unexplained and the possibility remains that it is derived from interactions with an unknown protein. In order to identify proteins with which CspE is capable of interacting *in vivo*, a genomic library will be incorporated into the BACTH system described above. The library will also serve as a *de facto* validation of the findings from the previous CLMS experiment. If any of the proteins identified therein are genuine interactors, they should be revealed by the screening of a genomic library.

Chapter 5

Genomic Investigation of CspE Interactions

Genome-wide Identification of CspE Interactors

The two-hybrid studies presented in the previous chapter suggest that CspE does not interact with the proteins tested (HupA and CsdA) at 37°C. Nor does it appear to interact with itself. Although there was some evidence for weak interactions of CspE at lower temperatures, the focus of this study is on the function of the cold shock proteins at 37°C, where they are potential targets for antimicrobial development. Therefore, for a comprehensive search for CspE-interacting proteins, a whole-genome library of *S.* Typhimurium proteins was constructed and screened for interaction with CspE using the bacterial adenylate cyclase two-hybrid system.

The library screening method is essentially the same as that used previously, but with unknown fragments of chromosomal DNA inserted into the pUT18C vector. Random protein fragments from *S.* Typhimurium are therefore tagged with T18 and tested for interaction with CspE. Previous genomic libraries have been created with *E. coli* (Domain et al. 2007; Karimova et al. 2009) and *Mycobacterium tuberculosis* (Klepp et al. 2009) DNA. In *E. coli*, the library screening method suggested the function of a previously uncharacterised protein YmgF, which was found to be required for localisation of the cell division machinery (Karimova et al. 2009). In *M. tuberculosis*, interactions were identified to help guide study of a virulence-associated protein whose function was unknown (Klepp et al. 2009).

Genomic DNA fragments were inserted into pUT18C, which has the MCS towards the C-terminus of the T18 Cya fragment (see Figure 4.2). This ensures that any stop codons contained within the inserted fragment will not cause arrest of translation before the 18C tag is translated. Given the results of the previous experiments with the CspE-T25 constructs, the *cspE*-pKT25 plasmid was chosen as the vector for *cspE*. This construct tags CspE at the N-terminus, and was the CspE construct for which the two slight interactions were observed in Chapter 4.

There are numerous possibilities for proteins which may interact with CspE. Eukaryotic Y-box proteins feature a cold shock domain with other domains at either terminus, which guide function of the CSD (Lyabin et al. 2014). A BLAST search of human and mouse Y-box protein 1 against the SL1344 genome shows that the C-terminal domain has homology to elongation factor P, suggesting a possible co-function of a Csp and Ef-P. Any of the proteins discovered to co-purify with CspE in Chapter 3's

cross-linking mass spectrometry experiment may be expected to interact here; in this case, the library method would serve as *de facto* validation of those findings. PagP and CrcB, which seem to co-function with CspE in chromosome condensation (Hu et al. 1996), are candidates for interaction. The library method outlined below is a means of screening each of these many possibilities, as well as all other *Salmonella* proteins. It is more efficient, and therefore more likely to identify novel interactions, than the cloning and testing of individual proteins, as performed in Chapter 4.

The genome of *S*. Typhimurium SL1344 is predicted to encode 4446 proteins ranging from 21 amino acids to 5559, with a mean length of 312 aa (Kröger et al. 2012). The mean is distorted by a small number of very large hypothetical proteins, however, so the median protein length of 269 aa may be a better representation of the average. A length distribution histogram (see Figure 5.1) shows the majority of proteins grouped between 100-500 aa. Libraries containing two distinct groups of inserts were therefore created. A library of 500 - 1500 bp fragments (called the 500 bp library) covers the length of most protein-coding genes, which are less than 1500 bp or 500 aa. A second library containing larger fragments of 1500 - 4000 bp (the 1500 bp library) should cover the majority of larger proteins, up to 1300 aa. 12 predicted proteins in SL1344 are larger than 1320 aa (Kröger et al. 2012), and would be unlikely to be included as complete proteins in this screen (although they may be present as fragments). The two largest proteins (at 3824 and 5559 aa) are predicted hypothetical proteins, and may or may not actually be translated *in vivo*. The largest characterised protein is the host colonisation factor ShdA at 2039 aa.

The genomic libraries were created from genomic DNA extracts by performing a partial digest with the restriction enzyme Sau3A1. Sau3A1 has a 4 bp recognition site and thus cuts more frequently than most cloning enzymes (theoretically every 256 bp), producing a complete digest of the chromosomal DNA. Fragments generated by Sau3A1 can be ligated into vector digested with BamH1, which leaves the same complimentary ends as Sau3A1. After ligation and transformation colonies containing inserts are directly screened for interaction on the recovery medium, which contains X-gal. As before, colonies develop blue pigmentation if two interacting proteins (or partial proteins) are expressed. Colonies are routinely tested to rule out two possible reasons for false positives. Firstly, transformation of plasmids containing the putatively interacting proteins into fresh host cells confirms pigmentation has not developed due to a host mutation. This step also confirms that the blue colony is not that of a

contaminant. Secondly, the genomic fragment is transformed into host cells along with empty pKT25. This ensures that an observed interaction is specific to CspE, rather than a non-specific interaction or one specific to the T25 fragment of Cya. One possible source of false positives is the cloning of a functional adenylate cyclase protein from SL1344 genomic DNA. To avoid this possibility, a Δ*cyaA* SL1344 mutant was created as a source of genomic DNA for use in library construction.

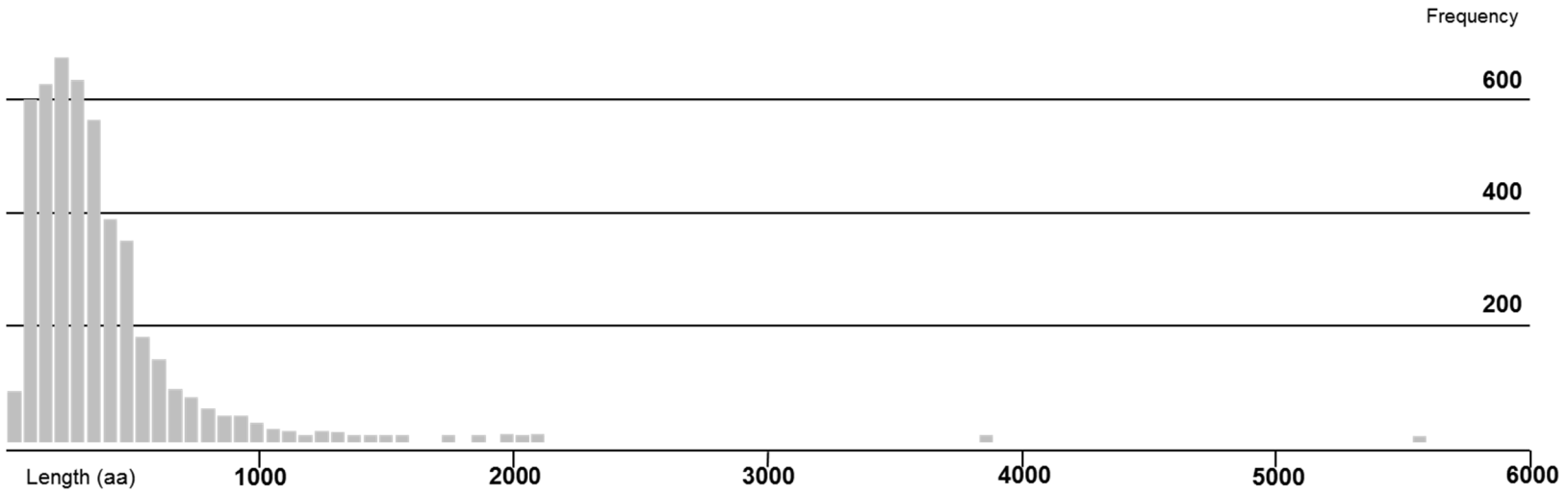


Figure 5.1. *Histogram of gene length distribution in S. Typhimurium SL1344. Majority of the 4446 chromosomally-coded proteins are 100-500 aa in length. Highest three bars represent 181-240 aa (n=681), 241-300 (n=644), and 121-180 aa (n = 635). Histogram modified from that presented by the NCBI database based on the most recent SL1344 genome sequence (Kröger et al., 2012).*

Results

Creation of ΔcyaA SL1344

Given that the BACTH system relies on reconstitution of adenylate cyclase, the introduction of a complete or partial *cyaA* gene as a part of a genomic library would be likely to result in false positives. In order to avoid this possibility, an existing *cyaA*::Tn10 mutant (Singer et al. 1989) was transduced from *Salmonella enterica* LT2 to SL1344 using bacteriophage P22; the method is found in Chapter 2. The Tn10 transposon encodes tetracycline resistance, allowing easy identification of transductants. Substitution of *cyaA* is confirmed by the loss of ability to ferment maltose, as visualised on MacConkey-maltose agar plates. The *cyaA* deletion strain is shown in Figure 5.2.

Generation of an SL1344 genomic library

To produce random fragments for a genomic library, Sau3a was used to digest genomic DNA prepared from Δ*cyaA* SL1344. Two genomic libraries were prepared, composed of fragments from 0.5-1.5 kb and 1.5-4.0 kb respectively (as detailed in Chapter 2). The upper limit of 4.0 kb was chosen for technical reasons. In the presence of smaller fragments, the ligation of longer fragments is less likely, so fragments above 4.0 kb would be unlikely to be found in the library. In addition, the need for extra agarose gel material to resolve the fragments (see Figure 5.2) would lower the efficiency of the gel extraction method. The digest to yield the shorter fragments was performed for 10 minutes at 37^{o}C. To prepare the longer fragments, the digest was performed for 5 minutes at room temperature, in order to slow down Sau3a activity. Results of the digest and DNA fragments extracted are shown in Figure 5.3.

After ligation and transformation as detailed in Chapter 2, transformants were grown overnight on LB-agar containing carbenicillin (selecting for presence of pUT18 containing genomic fragments), kanamycin (selecting for pKT25-*cspE*) and nalidixic acid (selecting for the Δ*cya* DHM1 host strain). To screen resultant colonies for interaction, IPTG was also added to induce expression from both plasmids, and the chromogenic β-galactosidase substrate X-Gal was added in indicate *lacZ* induction. Colony counts from each library are given in Table 5.1.

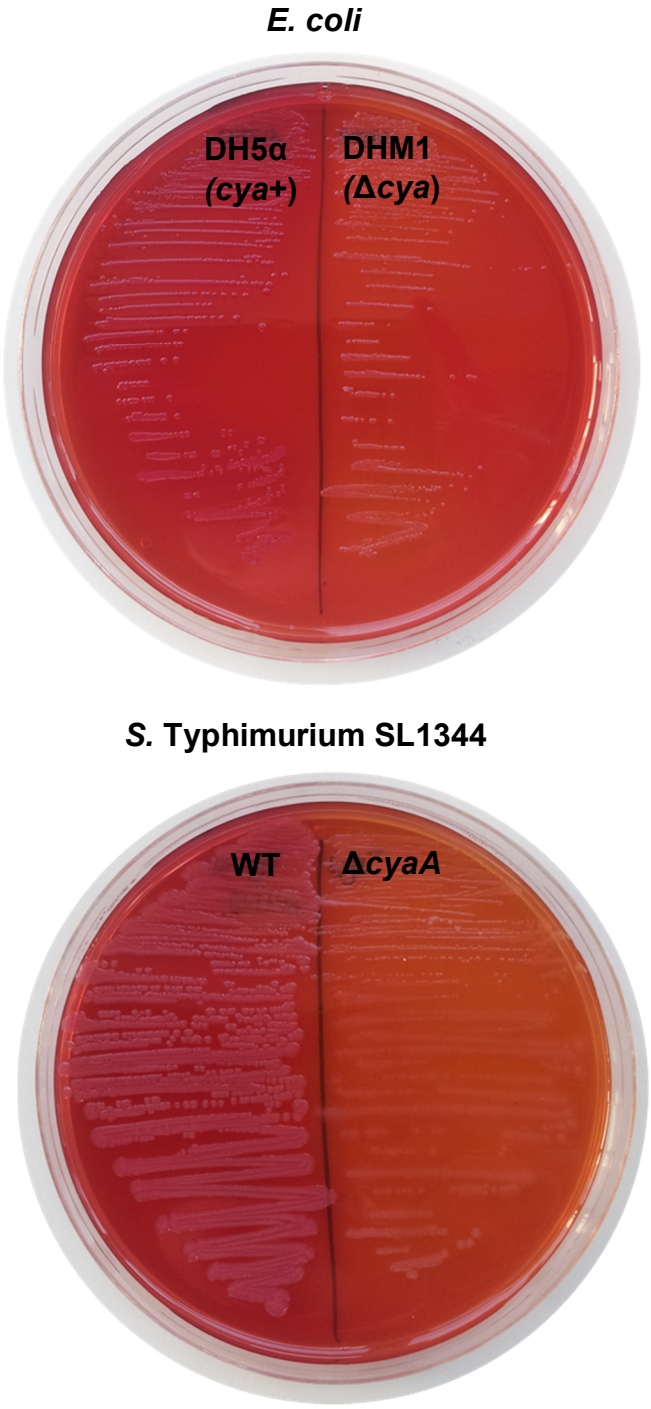


Figure 5.2. *Deletion of* cyaA *from SL1344. MacConkey-maltose plates indicate function of adenylate cyclase by pH change following fermentation of maltose.* cya^{+} *strains DH5α and SL1344 ferment maltose, resulting in lowered pH of media and red colony pigmentation. DHM1 and the created Δ*cya *SL1344 strain are unable to ferment maltose as cAMP is required for induction of the* mal *operon. These two strains ferment peptone, producing ammonia which raises the pH of the media and produces white colonies.*

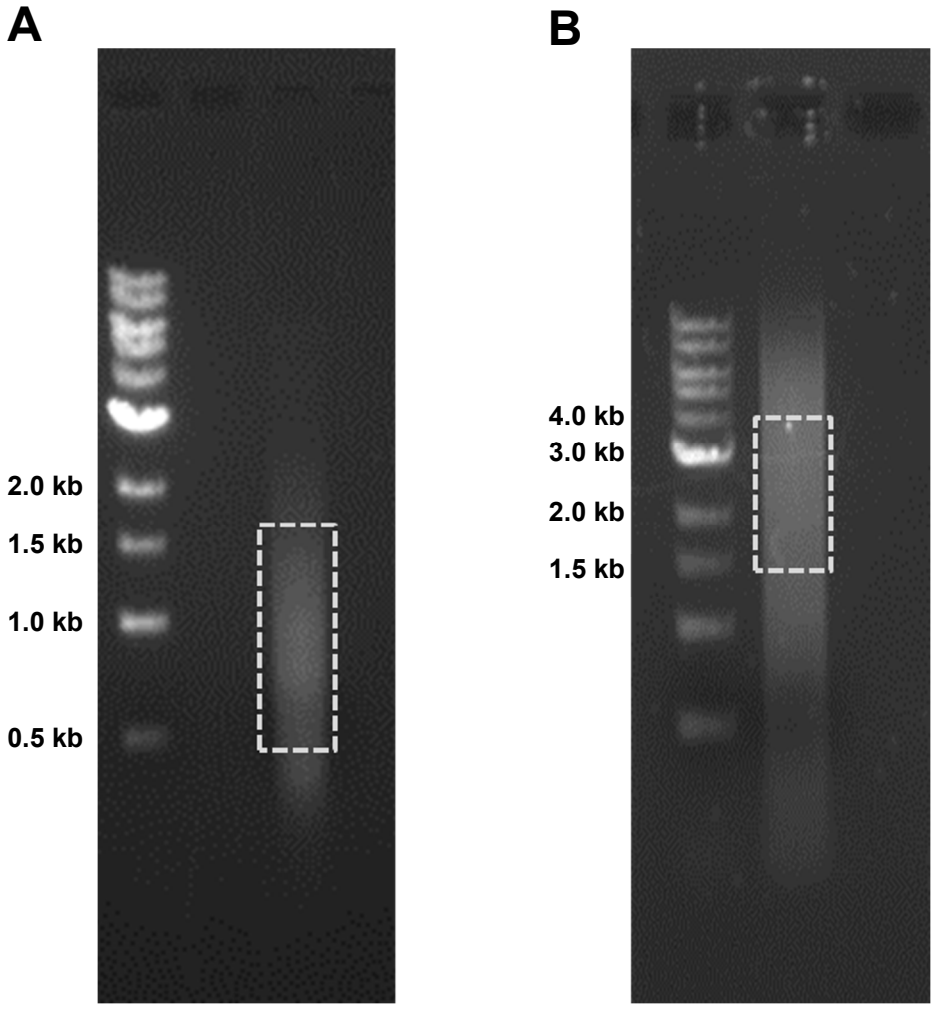


Figure 5.3. *Gel extractions from Sau3a digests. Dashed boxes represent regions extracted from 0.8% agarose gels to yield fragments 500 – 1500 bp (A) and 1500 – 4000 bp (B).*

Genomic Investigation

In order to determine both the insert size of the library and the number of empty plasmids transformed, a colony PCR screen was performed. The results are shown in Figures 5.4 and 5.5, which show representative colonies from 54 screened in each library. Screens confirmed that insert sizes were as expected, being concentrated around 500-1000 bp in the case of the 500 bp library and 1500 – 3000 bp in the 1500 bp library. The screens also allowed enumeration of the number of colonies containing pUT18 which had religated without an insert. This was infrequent in both libraries: 7.4% in both cases, or four out of 54 colonies screened. Calculations to determine coverage assume the minimum size for all inserts (450 or 1500 bp) where actually many will be larger. Thus, the estimates of coverage are likely to be underestimates.

Library	Transformant Colonies	Insert Size (bp)	% Empty	Genome coverage	P coverage (%)
500 bp	44,020	450-1350	7.4	3.60 x	97.3
1500 bp	30,230	1500-3000	7.4	8.23 x	99.7

Table 5.1: *Statistics pertaining to 500 and 1500 bp libraries. Insert sizes and % empty vectors transformed based on results of colony PCR screen. Genome coverage is calculated by dividing the product of insert size(taken to be 500 bp) and number of colonies (less those containing empty vectors without an insert) by the genome size of SL1344 (5.1 mb). Probability P of including a given fragment of DNA in the library is calculated by the Clarke-Carbon equation (Clarke and Carbon, 1976), described below.*

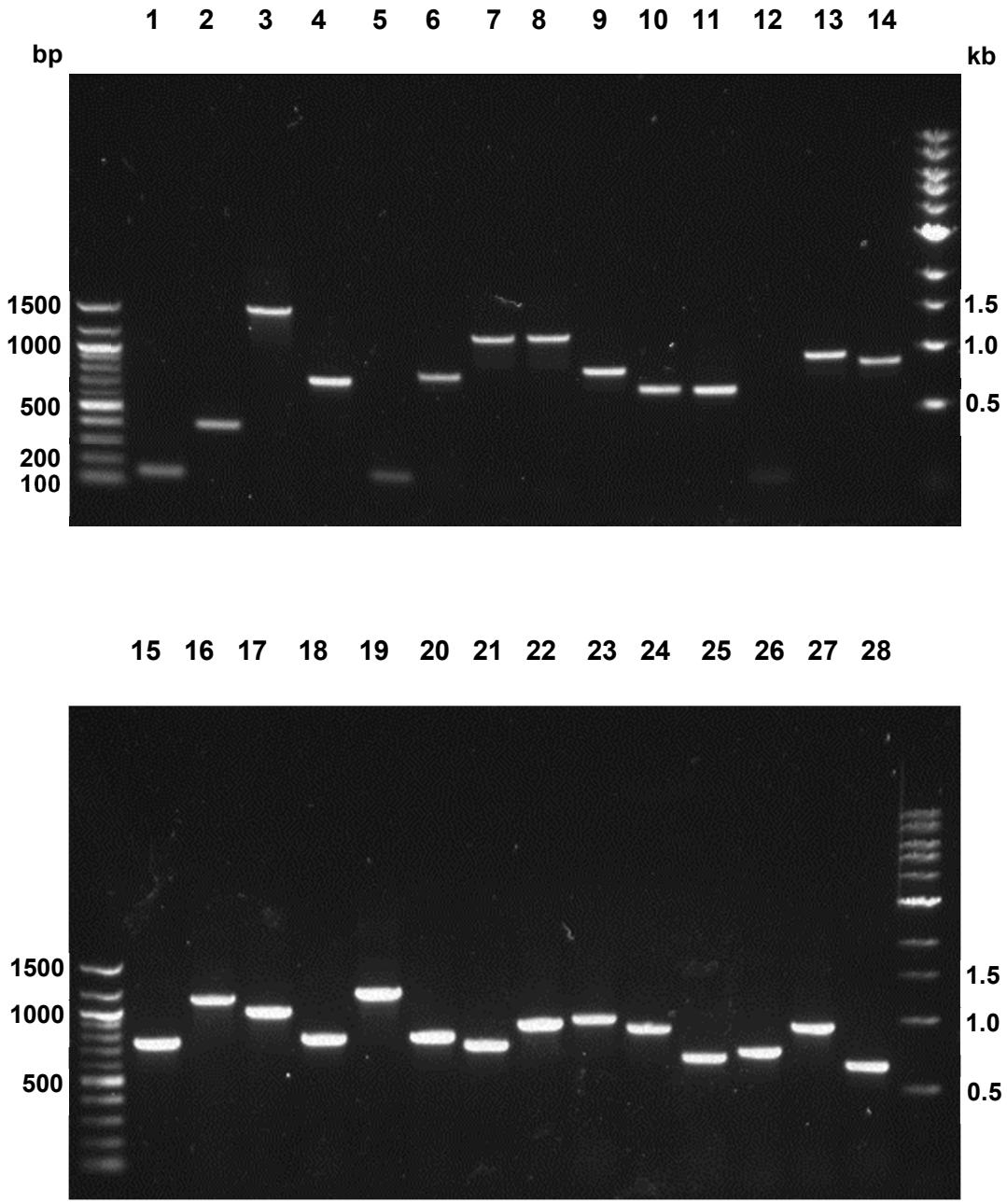


Figure 5.4. *PCR screen of colonies making up the 500 bp library. PCR primers amplify across the multiple cloning site of pUT18C; empty vector produces a product of ~150 bp (lane 1) Products from colonies containing transformed library (all other lanes) range from 600 bp to 1500 bp, equating to inserts 450 bp – 1350 bp in length. The majority of inserts are grouped between 500 and 1000 bp. Two of the 27 screened colonies (lanes 5 and 12) show plasmid without an insert, suggesting 7.4% of the total recovered colonies contain empty vector.*

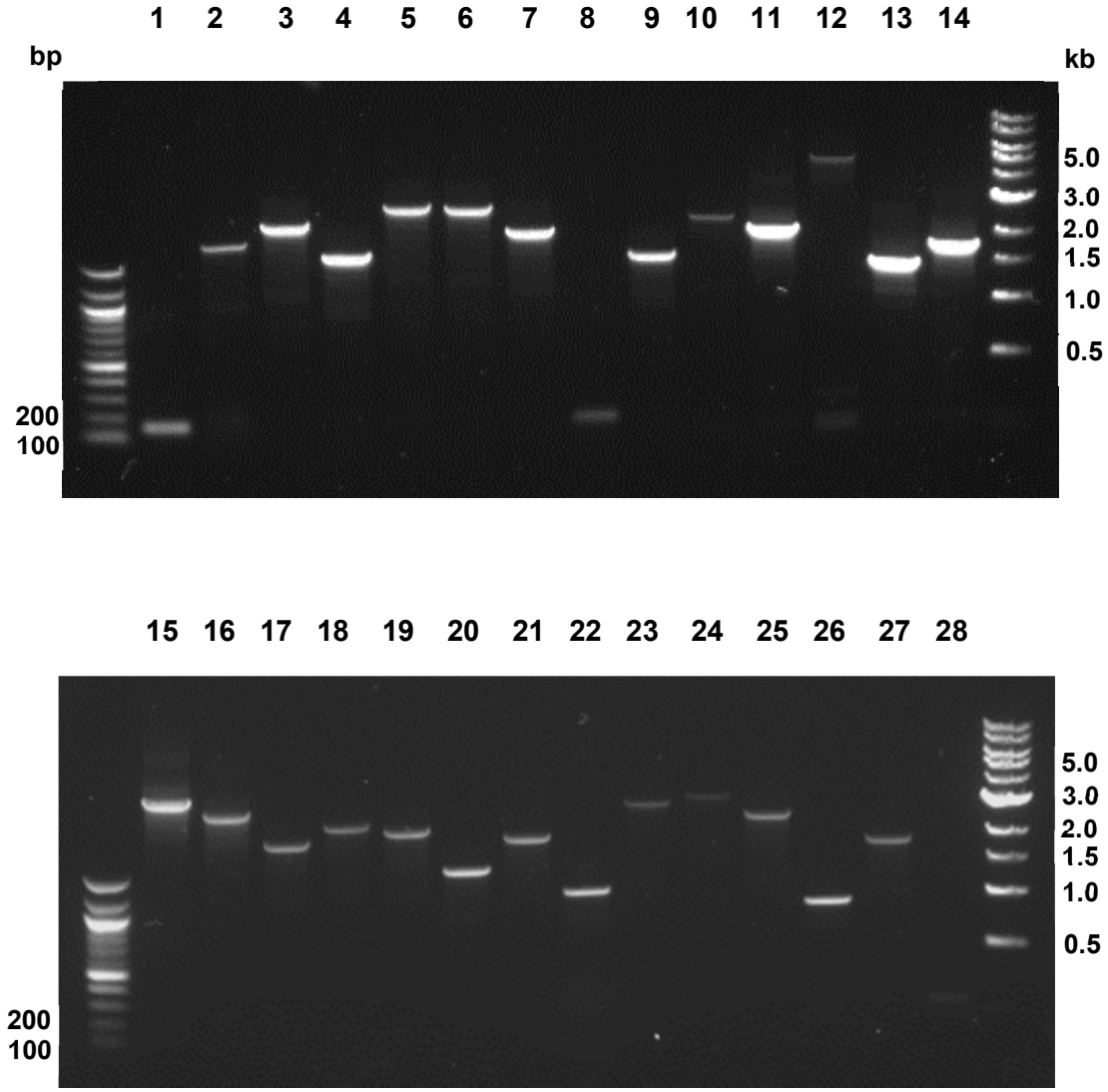


Figure 5.5. *PCR screen of colonies making up the 1500 bp library. PCR primers amplify across the multiple cloning site of pUT18C; empty vector produces a product of ~150 bp (lane 1) Products from colonies containing inserts (all other lanes) range from 1.5 kb to 4 kb, equating to inserts 1.35 kb – 3.85 kb in length. The majority of inserts are grouped between 1.5 – 3.0 kb. Of 27 colonies screened, two (lanes 8 and 28) showed a band consistent with empty vector, suggesting that 7.4% of the total recovered colonies contain vector without an insert.*

Protein-Protein Interactions of CspE

To calculate coverage of the genome, the Clarke-Carbon equation (Clarke & Carbon 1976) was used. The equation expresses the probability of any given sequence being included in a library of random fragments, assuming the fragment is of equal or lesser size than that of the inserts. The equation uses the number of unique clones, the fragment size, and the size of the genome. The Clarke-Carbon equation is as follows, where n = number of colonies required, P = probability, b = genome size and a = fragment size.

$$n = \frac{\ln (1-P)}{\ln (1-1/(b/a))}$$

To reach a 95% probability of including a fragment in the SL1344 genome (5.2 mb) with fragments of 500 bp, 30, 555 clones are required. Using fragments of 1500 bp 10, 184 clones are required. Creation of both the 500 bp and 1500 bp libraries was successful, with theoretical coverage of 97.3% and 99.7% respectively. It is worth considering that in order to reach a 100% probability of including a point in the library, an infinite number of clones would be needed.

Identifying interactions in the libraries

Initially, potential interactions are easily identified due to the development of blue pigmentation in a colony. The colour change observed using X-gal is dependent on growth time (eventually, all colonies develop colouring), so pigmented colonies were subcultured from the initial transformant plates to be sure of interaction. Any colony which maintained blue colouring after subculturing was a candidate for further investigation. Subcultured colonies were grown alongside positive and negative controls as comparisons of pigmentation. If colouration developed before that of the negative control, the test was positive.

Those that retained pigmentation were grown in LB broth (supplemented with ampicillin, kanamycin, and nalidixic acid) overnight and miniprepped to extract plasmid DNA. The DNA extracted contains both pKT25-*cspE* and pUT18 containing an

unknown fragment. In order to rule out the possibility that pigmentation had arisen due to host mutation or transformation of a contaminant, the DNA mixture was transformed into fresh electrocompetent DHM1 host cells. Those that retained the blue pigmentation in fresh cells are therefore confirmed to produce the pigmentation as a result of an interaction between the T18 and T25 constructs.

In order to confirm the specificity of the interaction, the unknown fragment was transformed into fresh DHM1 along with an untagged T25 on pKT25. In order to isolate the unknown fragment on pUT18, the previously isolated plasmid mix was transformed into DHM1 and plated onto ampicillin only; thereby recovering only pUT18 transformants. These transformants were miniprepped again, and the plasmid co-transformed with pKT25. Resultant colonies are either blue, indicating that an interaction occurs without the presence of CspE (therefore the interaction is non-specific, or specific to the T25 portion of Cya) or white, indicating that the presence of CspE is required for interaction (and therefore that the interaction is specific to CspE). This screening process is outlined in Figure 5.6.

Confirmation of positive interactions

In total, 47 genomic fragments were screened for interaction with CspE based on colour change on solid medium: 27 from the 500 bp library and 20 from the 1500 bp library. Of these, 31 produced blue pigmentation even when co-transformed with empty pKT25 vector, suggesting the interaction of the translated T18-fragment fusion is not specific to CspE. The fragment could either non-specifically interact with many cellular proteins (for instance, if the fragment were translated to a protein chaperone) or could interact specifically with adenylate cyclase. The fragment carried could also code for another activator of *lacZ* transcription, independent of cAMP. 15 of the plasmids yielded no colour change when transformed into fresh cells, suggesting that the initial blue pigmentation was either due to a mutation in the host or due to a contaminating species in the initial library.

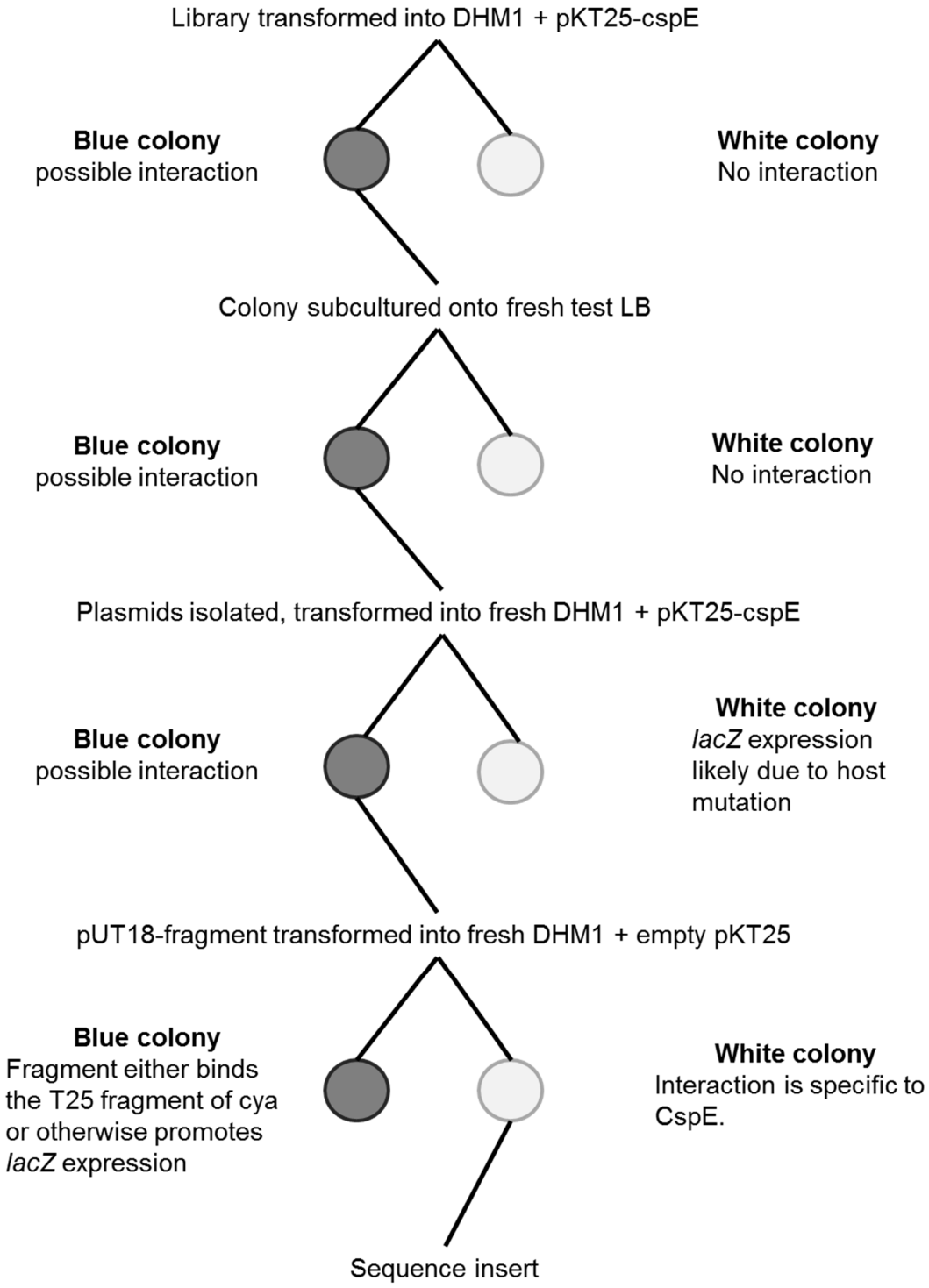


Figure 5.6. *Flow diagram of colony screening process. Colonies containing an insert which interacts with CspE are blue through each step, except when transformed with empty pKT25.Blue colonies indicate an interaction occurring, white colonies imply no interaction*

Genomic Investigation

One genomic fragment yielded blue colonies when transformed into fresh host cells, meaning that interaction was not due to host mutation, and white when transformed with empty pKT25, meaning that the interaction was specific to CspE. In order to confirm and quantify the strength of the interaction, β-galactosidase assays were carried out on the fresh host cells (F), and those containing empty vector (E).

The tested strains were Gram stained in order to confirm that β-galactosidase activity was not a result of a contaminating species. The presence of contaminants is unlikely given that three antibiotics (nal, amp, and kan) were present in all growth media. A contaminant would have to either possess resistance to all three antibiotics or be resistant to nalidixic acid and naturally competent, in order to be transformed with the plasmid mixture carrying the other two resistance genes. Common naturally competent bacteria include *Bacillus* (Kumpfmüller et al. 2013), *Neisseria* (Berry et al. 2013) and *Streptococcus* (Lecomte et al. 2014). However, any of these contaminants would be easily identified by a Gram stain and by colony morphology. Gram staining confirmed that the tested strains were all Gram negative rods.

One strain, designated 1500.P, gave a strong β-galactosidase signal in the fresh host cells, and little signal in the host with empty vector. This is indicative of a strong interaction specific to CspE. β-galactosidase activity in fresh (F) cells was a median of 58% that of the positive control and around 130 times the negative control. The β-galactosidase activity of the empty (E) cells was only 5% that of the F cells, but still higher than the negative. This suggests a degree of non-specific interaction occuring with the T25 portion of Cya, but a much greater interaction occurring with CspE-T25. Strain 1500.P has a slight slow growth phenotype, visible both on plates (where colonies are smaller) and in liquid culture (where overnights reach a lower OD_{600}). Based on the preparation of cells for β-galactosidase assay, strain 1500.P grew at around a quarter the rate of the positive control (ie. 25% of the maximum growth rate in liquid culture). A further two genomic fragments, which produced marginal results on the plate assay, were also tested but their interactions proved to be non-specific when quantitated. Neither showed a significant difference between F and E strains.

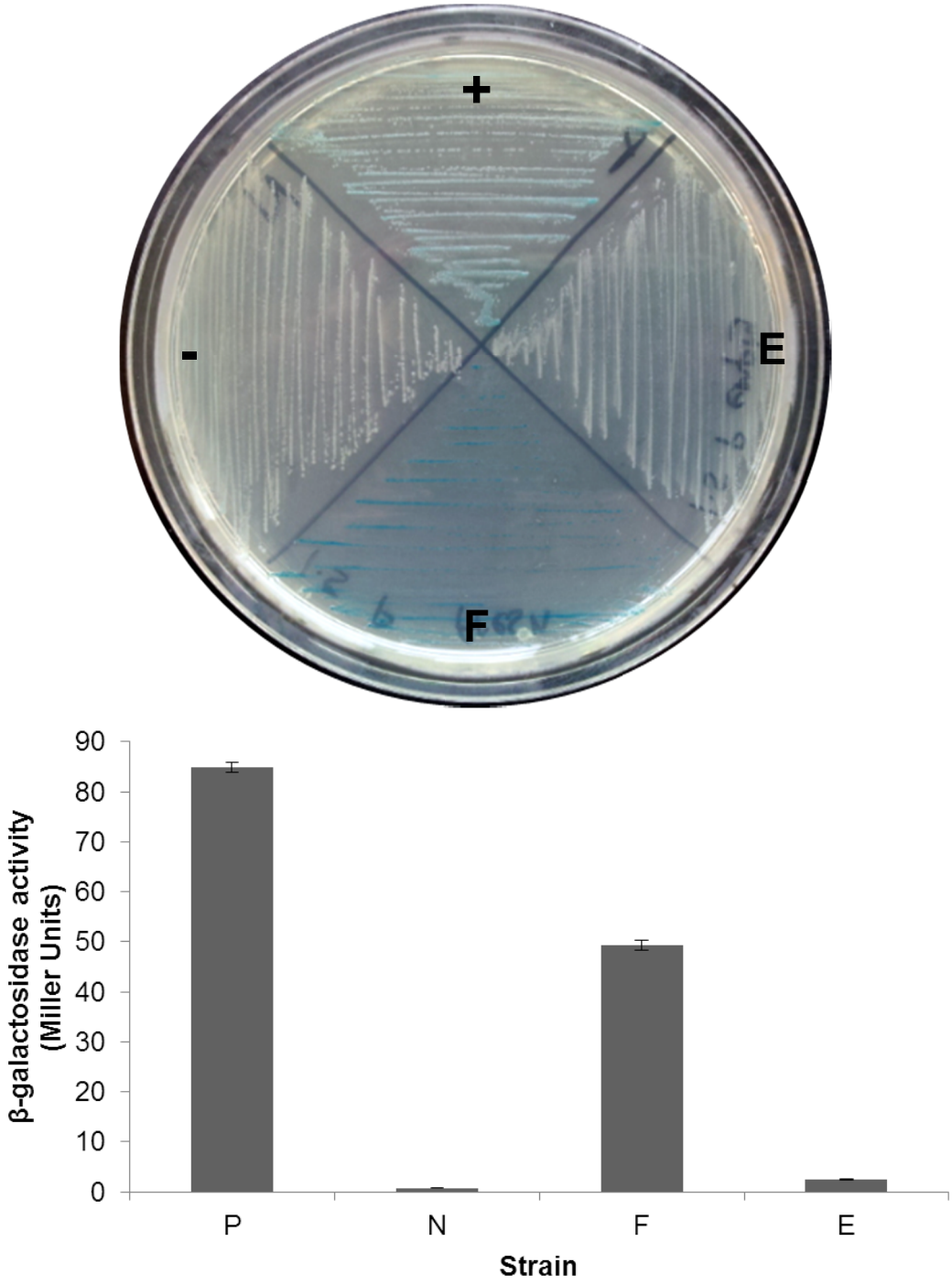


Figure 5.7. *Specific interaction between protein coded by unknown fragment 1500.P and CspE. Top, qualitative plate assay of cleavage of X-gal by β-galactosidase on solid media: F cells (CspE-T25 + 1500.P-T18) turn blue where E cells (T25 + 1500.P-T18) do not. Bottom, quantitation of β-galactosidase activity by Miller assay. F cells have much higher activity than both negative control (N, -) and E cells. Positive control (P, +) included for comparison.Graph representative of three repeats. Error bars represent standard deviation of three samples prepared identically. Miller units calculated using equation detailed in Chapter 4. Assays carried out as described in Chapter 2..*

Genomic Investigation

The pUT18C plasmid from strain 1500.P was sequenced (using primers *pUT18C sequencing F* and *R* (see Table 2.4)) in order to determine the genomic DNA fragment it carried. The sequencing showed an insert of >1500 bp which, in frame to the T18 fragment 5'- of the insert, extends for 705 bp before a translational stop codon. Thus, the gene fusion translates as T18 followed by a 235 aa SL1344 protein fragment. The protein fragment was translated *in silico* and identified by BLAST as a portion of a transcriptional regulator, DeoT.

DeoT was previously identified in *E. coli* and *Salmonella* Typhimurium as a repressor of transcription belonging to the DeoR family (Elgrably-Weiss et al. 2006). The DeoR family of regulators typically feature two domains: an N-terminal DNA-binding helix-turn-helix and a C-terminal effector binding and oligomerisation domain (Jones et al. 2014). The genomic fragment identified here includes all but the initial part of the N-terminal domain (ie. it lacks the first 15 aa from the N-terminus of DeoT). The protein translation of the construct sequenced is shown in Figure 5.8, as well as the domain architecture of DeoT.

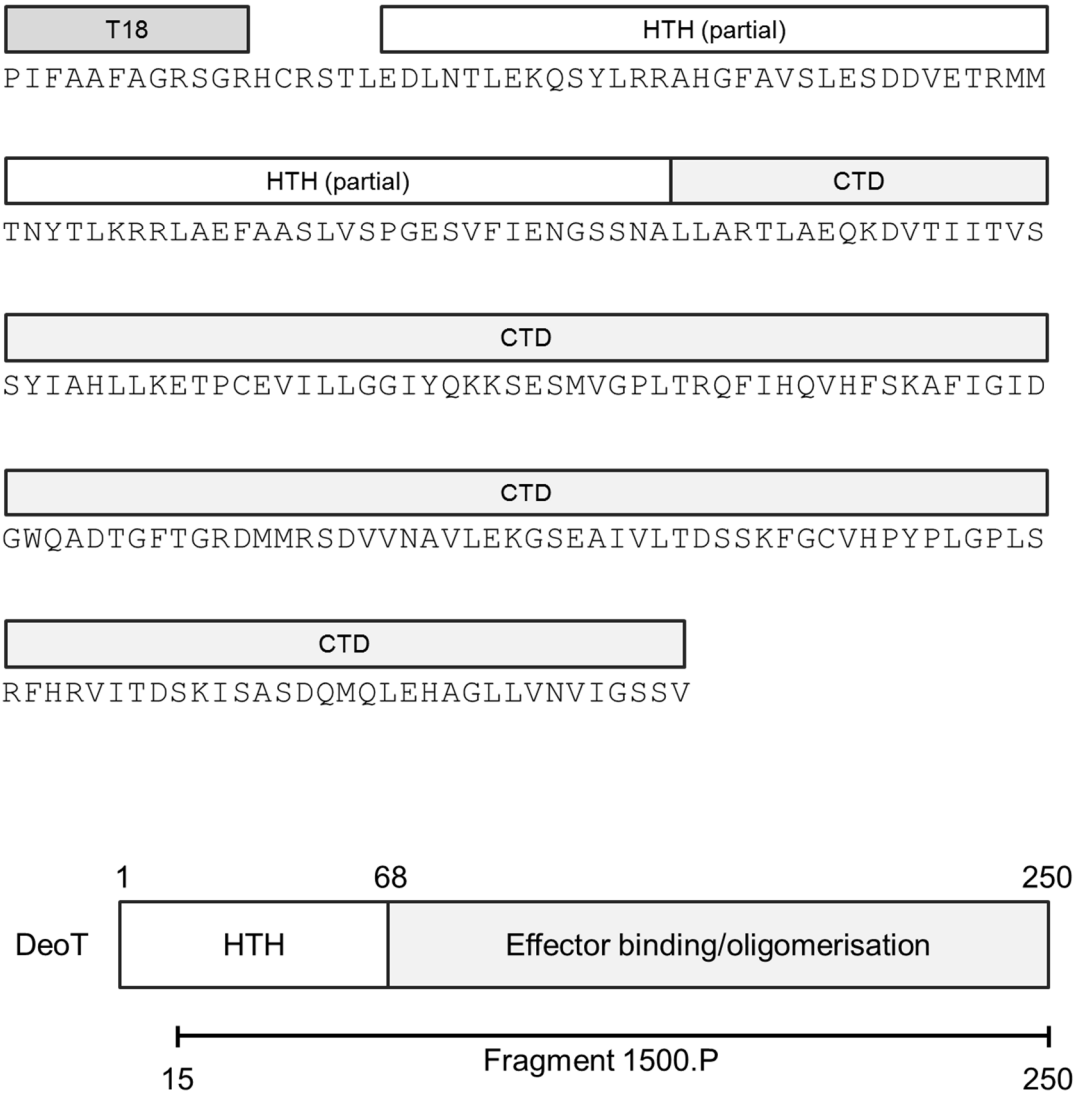


Figure 5.8. *Translated protein coded for by the genomic fragment 1500.P. Top, the translated fusion protein showing the end of the T18 tag and the translated portion of DeoT, which includes part of the N-terminal helix-turn-helix domain (HTH) and the complete C-terminal effector binding/oligomerisation domain (CTD). Below, the domain architecture of complete DeoT with the region covered by the fusion protein indicated.*

Discussion

Two libraries composed of fragments 500-1500 bp and 1500-4000 bp were created from SL1344 genomic DNA. Both were screened for production of proteins interacting with CspE. Most of the protein constructs screened either exhibited no interaction or did not interact specifically with CspE, but one genomic fragment produced a protein fusion which demonstrated a strong and specific interaction. The protein was a transcriptional regulator, DeoT.

DeoT, a DeoR-type transcriptional regulator

The SL1344-derived interacting protein construct had 100% sequence identity to a protein of *S*. Typhimurium strain LT2 characterised as DeoT, a transcriptional regulator of the DeoR family (Elgrably-Weiss et al. 2006). The *Salmonella* DeoT is 82% identical to that of *E. coli*, and therefore may well be functional in the *E. coli* host. The DeoR family are DNA-binding repressors of transcription. DeoR itself represses transcription by cooperatively binding multiple operator sites upstream of the regulated *deo* operon, inducing formation of a DNA loop which inhibits the initiation of transcription (Amouyal et al. 1989). The *deo* operon encodes a number of genes required deoxyribonucleotide transport and synthesis (Dandanell et al. 1991).

The dissociation of DeoR-family proteins from their operator regions is highly influenced by the presence of an effector molecule, usually a sugar substrate. Most of the DeoR-family repressors regulate a single gene or operon, usually related to sugar metabolism, with specific derepression by a related inducer molecule. Examples include the glycerol (*glp)* operon regulator GlpR (Escapa et al. 2013), the glucitol repressor GutR (Yamada & Saier 1988), and the fucose-reponsive regulator FucR (Hooper et al. 1999). The repressor GlpR, for example, dissociates from the operator region when glycerol-3-phosphate is bound to its C-terminal domain (Zeng et al. 1996). The fucose regulator is unusual in that a single regulator-inducer combination regulates two separate operons. In the presence of fucose, a secreted inducer of fucose production by the host is repressed, while fucose utilisation genes are induced. In the absence of fucose, the reverse occurs (Hooper et al. 1999). The DeoR family are generally very similar, being 240-260 amino acids in length and featuring an N-

terminal DNA-binding helix-turn-helix domain and a C-terminal effector binding and oligomerisation domain (Elgrably-Weiss et al. 2006).

DeoT is unusual in that it appears to regulate a number of unrelated genes at distinct and distal locations. Initial studies in *E. coli* into the regulator identified up to 22 genes apparently regulated by DeoT by random insertion a *lacZ* reporter gene into the chromosome. Genes which responded to varying DeoT levels were identified by β-galactosidase assays. Only four were studied in detail, however. These four included two maltose utilisation operons (*malEFG* and *malK-lamB*), a peptidase (*pepN*) and the fatty acid β-oxidation complex (*fadBA*) (Elgrably-Weiss et al. 2006). All four were repressed at the level of transcription, in line with the current model of DeoR-type repressor function. However, some of the genes apparently regulated by DeoT have other regulatory mechanisms described. The *mal* operons are subject to regulation by a regulator MalT, which is sensitive to a number of stimuli including various maltose compounds, cAMP, and the products of the *mal* operons (Reimann & Wolfe 2011). *fadBA* is regulated by the repressor FadR, whose repression is relieved by the presence of long-chain fatty acids (Fujita et al. 2007). *pepN* expression is promoted under a number of environmental conditions (Gharbi et al. 1985) although the mechanisms of induction are unknown. That DeoT affects a number of target genes, combined with the fact that these genes appear to have alternative regulatory mechanisms, suggests that the effect of DeoT may be indirect. In addition, no common sequence motif was found in the DeoT-regulated promoter regions, lessening the chance of a common recognition site as a unifying factor. Therefore, DeoT's repressive activity could be mediated through another regulator, or at a level away from the initiation of transcription. No co-repressor has been described to date for DeoT or any other DeoR-family protein.

Interaction between DeoT and CspE

An interaction between CspE, an antiterminator and RNA chaperone, and DeoT, a DNA-binding transcriptional regulator, is unexpected. However, an indication of the nature of the interaction might be found in DeoT's helix-turn-helix domain. Interpro, an *in silico* predictor of domain architecture and family membership (Jones et al. 2014), classifies DeoT's HTH as a 'winged' HTH (wHTH), a variant of the structure of which some examples have been demonstrated to bind RNA as well as DNA (Aravind et al. 2005). The function and fold of the domain in the absence of its initial 15 aa (as in the

fragment assayed here) is unknown, however. The best studied wHTH domains include the selenocysteine incorporation factor SelB, which binds mRNA hairpins (Yoshizawa et al. 2005) and the eukaryotic La motif, which is found in a number of multidomain proteins and binds mRNA with a preference for a 3'-UUU_{OH} (Dong et al. 2004). This specificity appears to be contributed to by another domain, the RNA recognition motif. A similar case is that of the protein CvfB of *Staphylococcus aureus*, which has a wHTH domain and binds ssRNA (Matsumoto et al. 2010). RNA binding is dependent on the presence of two domains, the wHTH domain and an S1 domain; a structure with which the cold shock proteins share considerable similarity (Xia et al. 2001).

The wHTH domain of DeoT might be involved in RNA binding, therefore. Such a function would deviate from that described for other DeoR-family proteins, but might account for the association with CspE and for the diversity of targets upon which DeoT seems to act. A 5'- untranslated region (UTR) would be an obvious RNA scaffold about which CspE and DeoT could interact, given the evidence for Csps binding to such structures (McGibbon 2013). However, the majority of transcription start sites of *S.* Typhimurium have been mapped (Kröger et al. 2012) and neither the DeoT-regulated operons, nor the higher regulators *malT* and *fadR*, appear to possess 5'UTRs. *deoT* itself possesses an leader region of 298 bp which is annotated in the original description as coding for a small protein of 70 aa, DeoL (Elgrably-Weiss et al. 2006). Determination of the transcription start site (Kröger et al. 2012) suggests that translation of the leader is impossible, however (the transcription start site is 1 bp downstream of the proposed ATG start codon in SL1344). Therefore, *deoT* mRNA appears to have a 5'-UTR of around 300 nt, which could present a possible susbstrate for CspE. This may not account for a CspE-DeoT interaction, however, as autoregulation of DeoR-family regulators (which would be the presumed model of action) has not been described. Interaction of CspE with *deoT's* 5'-UTR would also likely not account for CspE's direct interaction with DeoT.

A possible association around a nascent RNA in transcription complex with RNA polymerase is an attractive possibility, especially given CspE's antiterminator activity discussed in Chapter 4. Presumably, if this were the case, DeoT's wHTH domain would have to display some specificity for mRNAs of the regulated genes discussed above. The specificity may be structural, as in the case of SelB (Yoshizawa et al. 2005) or sequence based, as in the case of La (Dong et al. 2004).

Protein-Protein Interactions of CspE

Another possibility is an interaction of CspE and DeoT at a single-stranded region of DNA. However, as the mechanism of a HTH binding to DNA does not require strand opening (Beamer & Pabo 1992) ssDNA seems an unlikely substrate. Although examples exist in which strand separation is contributed to by a wHTH domain, these instances tend to require involvement of several other domains, such as in RNA helicases (Büttner et al. 2007; He et al. 2010). However, co-association of a cold shock protein and a transcription factor might account for previous *in vitro* observations which suggest a CspA function as a promoter of transcription (A Brandi et al. 1994; Jones et al. 1992) despite it lacking dsDNA binding capacity. A possible model could be one in which CspE binds to a transcriptional repressor and stimulates gene expression by antirepression.

There exist a small but expanding group of helix-turn-helix proteins which have a role in mediating protein-protein, as well as protein-DNA interactions. Most of the examples are eukaryotic (Harami et al. 2013), but present interesting models of the diversity of HTH functions. In *E. coli*, RecQ possesses a HTH-like domain which is required for DNA unwinding but also for formation of tetramers, which have an alternative function to that of the RecQ monomer (Lucic et al. 2011). Similarly, RecQ interactions with the DNA-binding protein SSB are mediated by a wHTH domain (Shereda et al. 2009). It is possible then that DeoT has alternative functions as a monomer and with CspE bound. Protein-protein interactions seem to be a secondary function of characterised HTH proteins, however, and they still retain DNA-binding function, so the interaction with DeoT is still likely to be based around a polynucleotide substrate.

Function of the DeoT-CspE interaction

Given the general function of DeoR (Short & Singer 1984) and other members of the DeoR family (Campos et al. 2004; Schweizer et al. 1985), it is likely that DeoT is a repressor of transcription. That being the case, a functional interaction with CspE would have the cold shock protein functioning as either a co-repressor or an anti-repressor. Although the results of the BACTH experiments (which indicate only the occurrence of an interaction) do not help to answer this question, the slow growth phenotype observed during interaction might. Slow growth was observed only in the presence of both CspE and DeoT ('F' strain), and not DeoT alone ('E' strain). The

reduction in growth is therefore a result of the interaction of the two proteins, rather than simply the overexpression of DeoT. If DeoT is a constitutive repressor, then it is unlikely that over-repression (such as would occur if CspE is a co-repressor) would lead to toxicity. On the other hand, if CspE is an anti-repressor, then DeoT's usual function would be disrupted by the interaction during the BACTH experiment and its target genes would be aberrantly overexpressed. Overexpression of DeoT targets involved in fatty acid and peptide degradation enzymes might increase their activity to toxic levels. As well as wasting resources in their unnecessary expression, they might aberrantly degrade vital cell components. Although chromosomal DeoT would still be expressed in the host, the abundance of overexpressed CspE construct is likely to overcome the minimal amounts of chromosomal DeoT present, negating its activity.

The general model for the DeoR family of protein regulators is that they are constitutively expressed and inhibit transcription until their repression is relieved, usually by an effector molecule relating to the pathway which is regulated. This is usually a substrate of the regulated pathway, for example L-ascorbate for UlaR (Garces et al. 2008). DeoT, however, appears to regulate several genes with diverse substrates (sugars, fatty acids, and peptides) and therefore it seems unlikely that derepression can be mediated by a substrate. In addition, both the maltose operons and the *fadBA* operon have known regulators which are specific to their substrates (Fujita et al. 2007; Reimann & Wolfe 2011). Although no direct signalling pathway has been described for *pepN*, it appears to be under the control of several systems (Gharbi et al. 1985).

Perhaps, then, the function of DeoT is to provide a second level of regulation, in conjunction with an anti-repressor CspE. Expression of CspE is upregulated during starvation conditions (Czapski & Trun 2014), indicating CspE has a role in the response to low nutrient availability. Full expression of the DeoT target operons could require two conditions: the presence of their substrate molecules (fatty acids or maltose) and the requirement for additional energy sources, as determined by the (starvation-related) presence of CspE. CspE is present in a basal level in the cell at 37^{o}C, allowing some expression of the target operons, but elevated CspE levels would be required for high expression. Such a system would be similar to that of the *lac* operon, in which two conditions are required for expression (Jacob & Monod 1961). Firstly, presence of the substrate is indicated by derepression of LacI, and secondly the requirement for lactose metabolism is signalled by falling energy levels indicated by CAP-cAMP.

Protein-Protein Interactions of CspE

DeoT was not identified as a possible interactor of CspE during the proteomic screen described in Chapter 2. There are a number of reasons an interaction might have been missed, however. Expression of DeoT is likely to be low (repression of an operon in theory only requires the binding of a single functional DeoT unit), and an interaction with CspE would therefore be difficult to detect. In addition, the CLMS experiment was performed with cells growing under optimal conditions. If the interaction between DeoT and CspE is related to starvation, it may not have occurred under the conditions tested during the first experiment. Using the BACTH system, with both proteins overexpressed, the interaction would occur regardless of environmental or physiological factors.

Models for CspE antagonism of DeoT

The most obvious model for CspE-mediated derepression of DeoT is one in which CspE takes the functional role of the small-molecule effectors of other DeoR-family proteins. These small molecules often function by moderating interactions between the regulator subunits. DeoR is a functional octamer (Mortensen et al. 1989) and, moreover, binds distal operator sites before interacting with itself to form an inhibitory loop of DNA (Amouyal et al. 1989). Binding of an effector molecule, therefore, could destabilise DeoR function at two stages.

This model would be similar to those described for other DeoR-family proteins, such as UlaR. UlaR binds DNA as a tetramer to induce loop formation, until the presence of L-ascorbate in the cell relieves repression. This is due to an alteration of UlaR quaternary structure in the presence of L-ascorbate from a DNA-binding tetramer to a dimer free in solution (Garces et al. 2008). A similar model, in which CspE destabilises DeoT binding to an operator, is shown in Figure 5.9.

A similar mechanism has a more refined regulatory role in the *gal* regulon of *E. coli*, where GalR regulates genes for galactose anabolism (eg for use during glycosylation reactions) and catabolism (for use of galactose as an energy source). In the presence of low cellular D-galactose, the DNA loop transiently opens due to destabilisation of GalR tetramers. In the presence of high D-galactose, the loop is more consistently relaxed and an additional CAP-cAMP promoter allows full expression of

galactose utilisation genes. For DeoT, varying levels of CspE could similarly determine the expression of the target genes.

However, the main failing of this model is the apparent absence of any recurring DNA operator sequence upstream of the target genes. The original description of DeoT in *E. coli* was unable to find any common motif at the promoter regions of the *malEFG*, *malK-lamB*, *fadBA* or *pepN* operons (Elgrably-Weiss et al. 2006). There are two possible explanations for this finding. Firstly, DeoT may not act directly on these targets but rather might affect transcription of another regulator common to each of them. If this is the case, that regulator was not identified by the original study. However, there were a number of genes (18) without a sufficiently strong DeoT-dependent response to merit further investigation, of which a master regulator (a modulator of cAMP, say, or (p)ppGpp) could have been one. Were DeoT to act as a repressor of such a global regulator, though, it might be expected that more than 22 genes would have been found to be responsive to its inactivation.

Using the motif searching tool MEME (Bailey et al. 2009) the possible presence of a regulatory motif 500 bp upstream of the *fadBA*, *pepN*, *malK-lamB* and *malEFG* transcription start sites was investigated. No motif was apparent in SL1344, as was the case in *E. coli* (Elgrably-Weiss et al. 2006). Another search including the upstream region of *malT* (the common regulator of the *mal* operons) revealed a pair of possible operator motifs, but on further investigation neither is likely to be genuine. The first (CTGGCCG) is present 1754 times in the SL1344 genome, and as such is too common to relate to a specific regulator. Another motif (TCWGCCSWGC) was present 72 times in the SL1344 genome, but largely within coding sequences, rather than upstream of start sites.

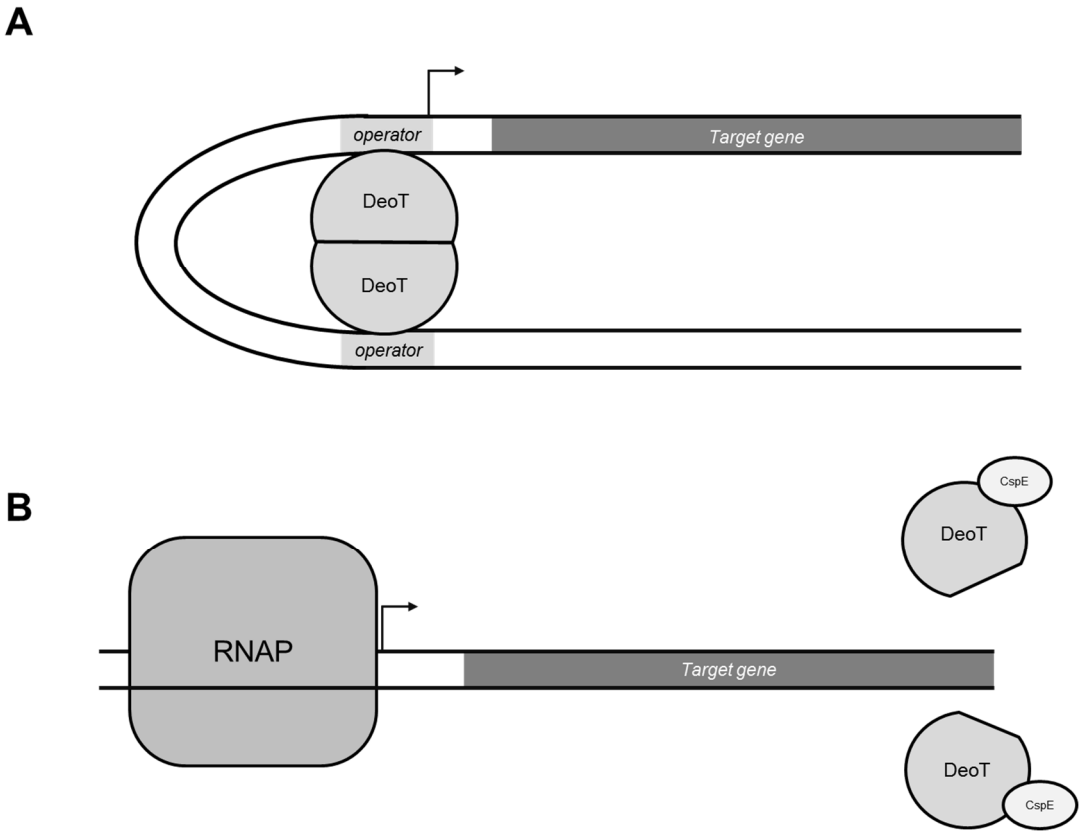


Figure 5.9. *CspE derepression of DeoT by disruption of operator binding. A) DeoT binds to an operator regions, forming a DNA loop near the transcription start site of a target gene (represented by an arrow), and thus prevents RNA polymerase (RNAP) transcription. DeoT may bind as a dimer, tetramer, or another structure: a dimer is shown here. B) CspE binds to DeoT and destabilises its interaction with the operator region. RNAP is able to initiate transcription. DeoT dissociates and may be held in an inactive conformation by the presence of CspE.*

An alternative explanation is a function of DeoT away from the initiation of transcription. CspE has a known function as an antiterminator protein, as discussed in Chapter 4, and it may interact with DeoT on a nascent RNA. Although such a model would differ from those described for other DeoR-family regulators, it would account for the multiple targets of DeoT and the lack of an apparent operator sequence, as well as DeoT's association with CspE. DeoT could recognise a structural motif present in nascent RNA, which are difficult to predict computationally. *pepN*, *fadBA*, and *malT* (the common regulator of *malEFG* and *malK-lamB*) are all predicted to contain rho-independent terminator sites (predicted by FindTerm (Solovyev & Salamov 2011)), and are thus possible substrates for CspE. A possible mechanism for DeoT-mediated antitermination by CspE is presented in Figure 5.10.

This guided antitermination model would have DeoT recognising its regulated RNAs, perhaps at the structural motif generated by terminator sequences, and associating with them during transcription. In the absence of CspE the terminator sequence would cause premature arrest of transcription and repression of the gene. With elevated CspE present, DeoT's affinity for CspE would guide its antitermination activity. CspE melting of the terminator structure would cause DeoT to dissociate, as the structure it recognises is lost, and would allow transcription to occur. There is precedent for interaction of multiple proteins at an antiterminator site in the λ phage Nus system in *E. coli*, where NusA interacts with the λ protein N (amongst others) to guide recognition of an RNA sequence (Prasch et al. 2009; Prasch et al. 2006).

Although there is precedent for the binding of RNA secondary structures by a helix-turn-helix domain (Yoshizawa et al. 2005), the mechanism described in Figure 5.8 would be a departure from those known for other proteins of the DeoR family. Either of the models for the precise interaction of DeoT and CspE require significant further work to confirm. Indeed, while it seems reasonable for CspE to be functioning as an antagonist of DeoT, that too requires experimental validation.

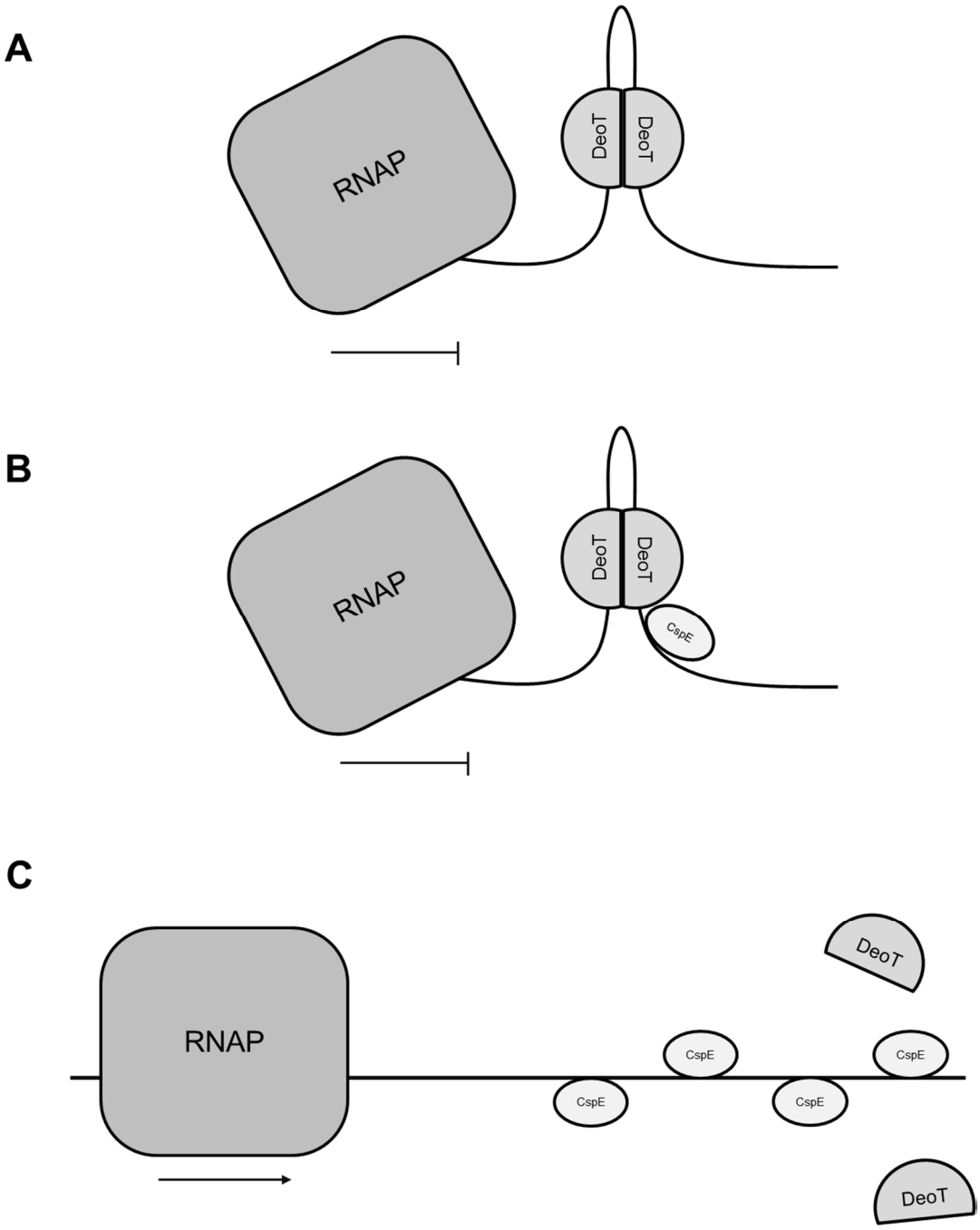


Figure 5.10. *Guided antitermination of DeoT target genes by CspE. A) Terminator sequence in DeoT target mRNA prevents RNA polymerase (RNAP) completing transcription. B) DeoT, recognising (for example) the terminator's secondary structure, binds and promotes association of CspE. C) CspE induces unwinding of the terminator sequence, with two effects: DeoT dissociates and RNAP can complete transcription.*

Future studies of DeoT-CspE

Time constraints have prevented more extensive characterisation of the CspE-DeoT interaction, leaving the nature and mechanism of their presumed co-regulation uncertain. Two possible models are presented: one in which DeoT binds DNA at an unknown operator sequence, and another in which DeoT binds nascent RNA. Determining which of the two is likely to be correct experimentally could be simply achieved by examining DeoT binding to dsDNA and ssRNA, for instance by isothermal titration calorimetry or gel retardation assays.

CspE may act antagonistically or synergistically with DeoT. This would be easily assayed by quantitation of DeoT target mRNA levels in a CspE knockout, by Northern blot or quantitative PCR. Generation of a reporter gene fusion (*lacZ*, for instance, or a luciferase) to one of the regulated operons would also allow easy quantitation. In the absence of CspE, if it were an antagonist, target mRNA levels would be expected to be lower than the wild type. If CspE were a co-repressor, target mRNAs would increase with its absence. These data would be complemented by similar experiments using an overexpressed CspE *in trans*, where the reverse results would be expected.

If DeoT functions in a manner similar to the other members of the DeoR family, then the major unknown factor is that of the operator sequence. If not a common sequence found upstream of DeoT's four described targets, then the operator must be elsewhere. This suggests another regulon of DeoT, which represses the four known targets indirectly. Other targets of DeoT could be determined as in the original paper, by altered expression following DeoT knockout (Elgrably-Weiss et al. 2006), or more directly using a method such as ChIP-seq (chromatin immunoprecipitation and sequencing). In the latter method, DeoT would be co-purified with whatever DNA fragments it binds *in vivo* (if such an association exists), which are identified by sequencing. This would determine the operator sequences of DeoT repression. If there is a single, principal, gene target of DeoT then the involvement of CspE would be intriguing. Where the other DeoR-type regulators sense small molecule effectors, DeoT would appear to respond to the presence of an entire protein. This would suggest regulation not by extrinsic factors such as environmental nutrients but by some intrinsic aspect of cell function, such as growth phase or a stress response.

If DeoT emerges as an RNA-binding protein, the issue of specificity is again important. A structural or sequence specificity could be identified by the same ITC or

gel retardation experiments using RNA with an intrinsic secondary structure. Once determined, would allow prediction of other genes under DeoT-CspE dual control. The software MEME (Bailey et al. 2009) found no common sequence motifs in the coding sequences of *fadBA*, *malK-lamB*, *malEFG*, *pepN* or *malT* , suggesting the specificity of DeoT might be to a secondary structure, rather than a sequence. Such structures would be harder to predict *in silico*, meaning an *in vivo* method might be more suited to discovery of unknown DeoT-CspE regulated genes. A transcriptomic analysis of mRNA in *deoT* and *cspE* mutants would be an obvious approach to assess their impact on the whole genome.

Exploitation of DeoT-CspE interaction

The apparent toxicity of the DeoT-CspE interaction is an interesting aspect of the discovery, and raises the possibility of further study towards drug design (although clearly much work remains to be done). A tightly-binding analogue of CspE, which leaves DeoT permanently derepressed might extend the slow growth phenotype observed here into a fully lethal effect. In order to design such an analogue, the interaction site of DeoT and CspE must first be mapped. Variations of DeoR-family proteins which do not respond to their effector molecules have been engineered (Ray & Larson 2004) for studying their functions, raising the possibility that the reverse effect could also be achieved.

Such investigation of interacting sites can be achieved with modifications of the BACTH system. Systematic deletion of sections of each protein could be used; this method determined the dimerisation region of the curved DNA binding protein CbpA. A subsequent alanine scan of the identified region revealed the exact residues involved in interaction (Cosgriff et al. 2010). An alternative approach would be to employ a random mutagenesis screen of DeoT, CspE, or both to find important interaction sites, as used to determine interaction sites between acyl carrier proteins and SpoT (Angelini et al. 2012). Using either method, the effect of mutation on interaction would be easily assayed (whether it be weakened or completely prevented) using the standard methods described here and in Chapter 4.

A detailed understanding of the interaction hot spots between DeoT and CspE would allow the rational design of a mimetic drug. Equally, it would be possible to

screen a library of likely compounds. A high-throughput library screen has been used to find inhibitors of a protein-protein interaction for antimicrobial development (Paschos et al. 2011), and could presumably be put to the same effect in the search for an interacting analogue. Again, the BACTH system is easily modified for use in multiwell plates (Battesti and Bouveret, 2012) and would present an easy assay for investigating modifiers of interaction. If a likely site for protein-protein interaction was identified on CspE, it may be the case that other proteins are also able to interact with that site, suggesting another avenue for further investigation.

Success and developments of the genomic library

The nature of a genomic library is such that it is impossible to be certain of absolute coverage. However, the calculations performed in Table 5.1 suggest that the libraries generated have a strong probability of comprehensive coverage. That only a single interacting partner of CspE was discovered seems unusual and it is possible that other partners have been missed, as the probability of library coverage cannot reach 100%. Given that the other whole-cell approach presented in Chapter 3 failed to identify any strong interactions, the possibility must be considered that factors other than protein-protein interactions are responsible for guiding CspE's function *in vivo*.

The library of fusion proteins created here is a useful tool for further study of CspE interactions. The experiment described here investigated the interactions of CspE on rich media at 37^{o}C, but it would be simple to assay other conditions to see whether CspE's interactions might change with environmental conditions. A repeat at low temperature or, given the link to alternative energy sources identified here, on various growth media, would be obvious initial options. The screening process could be made easier by plating the library onto minimal media supplemented with lactose. This would allow the growth of only colonies carrying interacting proteins (or fragments), and make identification of such interactions easier. Due to the need for colony counts in order to assess library coverage, this approach was not possible during the generation of the library but could be applied to subsequent experiments.

The library was constructed from (Δ*cyaA*) SL1344 genomic DNA, and thus was unlikely to include DNA fragments from the three plasmids carried by the strain (Kröger et al. 2012). As these plasmids include a number of virulence determinants, it

would be valuable to assess their interaction with CspE. The plasmids express a combined 209 proteins, making cloning of individuals impractical. A small library, generated from miniprep DNA and prepared as described here, would be a simple way to screen the plasmids. To cover the plasmids with 95% probability of inclusion of a given sequence, only 400 (for a 1500 bp library) or 1500 (for a 500 bp library) colonies would be required (Clarke & Carbon 1976).

Conclusions

A genomic library of S. Typhimurium DNA was successfully created and used to screen fusion proteins for interaction with CspE; one such interaction was discovered between CspE and a transcriptional repressor DeoT. Many aspects of the interaction remain unknown, but it seems likely that CspE acts as an anti-repressor. A number of mechanisms by which this effect could be mediated are possible, and two models are presented here. The first, an operator-binding model, is in line with current theories of the function of DeoR-like proteins, but lacks an apparent operator site which is crucial for activity. The second, a guided antitermination model, would be a novel function of the DeoR family but fits the available evidence.

Much further work remains to be carried out. This includes further characterisation of the observed DeoT-CspE interaction in order to determine which (if either) of the two proposed models is correct. Additionally, a slow growth phenotype was observed where the interaction was present; this raises the possibility of studying the precise details of the interaction as a possible target for drug design. The library generated herein is a valuable research tool, and could be used to study the protein-protein interactions of CspE under various conditions.

Chapter 6

Concluding Discussion

Discussion

Cold shock protein E of *Salmonella* Typhimurium is a curiosity: a cold shock protein expressed constitutively, an apparently non-specific chaperone with a unique set of substrates, and a near ubiquitous protein whose function remains unclear. Despite being chiefly characterised as a mediator of the cold shock response, the cold shock proteins are found in species with no requirement for such a response and even in mesophiles are expressed during growth outside the cold. Their function in such instances remains elusive. They may have specific regulatory roles yet to be discovered, or they may be acting in the same manner as at lower temperatures, as general RNA chaperones and antiterminators.

Despite the characterisation of the family at low temperature, several aspects of the cold shock protein's function during normal growth remain unknown; not least of these is their principal function at that higher temperature. In addition, their means of substrate selection and their formation of quaternary structures are unknown. The studies described here aimed to expand knowledge of the function of a cold shock protein, CspE, outwith growth at low temperatures. Investigation of the protein-protein interactions of CspE was an interesting avenue to explore during this initial characterisation. Such studies can indicate a protein's cellular location, functional partners, and regulatory mechanisms as well as offering broader insights into function.

Proteomic analysis of CspE interactions

The initial attempt to identify protein interacting partners of CspE was based on a proteomic method, in which whole cells were subjected to chemical cross-linking in order to fix noncovalent interactions. This facilitated affinity purification of protein complexes before identification of their components by mass spectrometry. The method can be powerful for identifying the members of protein complexes, where the subunits are reliably found in close proximity to each other, and form tight interactions.

CspE appears not to form such interactions, however, as no such complexes were consistently purified. This suggests that CspE does not function as part of a defined multiprotein complex, but rather may form a range of different, low-frequency, transient interactions. The proteins identified as co-purifying with CspE gave an

indication of the subcellular location of CspE, however. The identified proteins were enriched for many found in two discreet locations: in the nucleoid, and around the ribosomal entry site. Such a finding is in line with current theories of CspE's function as an RNA chaperone during transcription and translation, but does little to expand on knowledge of its activity.

Although a small number of potential CspE interactors were identified by mass spectrometry, the inconclusive results required further validation. In order to investigate the possible association between CspE and HupA, a bacterial two-hybrid system was employed to assay the potential for interaction between the two. Despite the possibility of their interacting around an ssDNA scaffold, the two-hybrid assay did not reveal any further evidence of interaction. With consideration of that result, the remaining results from the cross-linking mass spectrometry experiment were deemed unlikely to be significant and were not investigated further.

Investigation of CspE oligomerisation

The two-hybrid system offered the opportunity to investigate CspE dimerisation (or higher oligomerisation) *in vivo*. Dimerisation of CspE has been reported previously *in vitro* (Johnston et al. 2006; Morgan et al. 2009) but never observed *in vivo*. The system employed here was unable to detect any evidence of CspE dimerisation at 37°C, suggesting the apparent dimer formation observed *in vitro* might not occur *in vivo*. However, a modest interaction between CspE monomers was detected at lower temperatures, which may be consistent with a co-localisation along an RNA substrate.

At 37°C, CspE appears to function as a monomer. If the interaction observed during cold shock is due to CspE's chaperoning activity, then that activity was not observed at 37°C. Similarly, the lack of observed dimerisation seems to rule out a role in chromosome condensation, which has also been investigated as a function of the cold shock proteins (Hu et al. 1996). One function of CspE which could be effected by monomeric protein is that of transcription antitermination. It is proposed, therefore, that the principal function of CspE during growth at higher temperatures is that of a transcription antiterminator.

Protein-Protein Interactions of CspE

Also assayed was a potential interaction between CspE and a cold-shock RNA helicase, CsdA. This investigation was based on an observation from *Bacillus subtilis* that homologues of CsdA and CspE were found by fluorescence resonance energy transfer to interact at low temperatures. Like CspE, CsdA is also expressed at all temperatures (rather than exclusively during cold shock) and so interaction was assayed at both 37°C and 10°C. Again like CspE, no interaction was detected at the higher temperature, although a weak signal was detected at the lower. The weak interaction might again be consistent with a common substrate at lower temperatures, with the two proteins being proposed to co-operate in melting RNA secondary structures in the cold (Hunger et al. 2006).

Genomic analysis of CspE interactions

Given that the proteomic approach to identifying protein interactions met with limited success, an alternative approach was attempted. Using the same bacterial two-hybrid system, a library of random peptides derived from genomic DNA was assayed for interaction with CspE. Two libraries comprised of different sizes of DNA inserts were screened and revealed a single strong interaction, between CspE and a transcriptional regulator DeoT.

The assumed model for DeoT's function prior to this study was based on that of similar proteins in the DeoR family, although it deviates from them in regulating a number of genes (where most of the family repress a single operon). The regulation of several genes by DeoT, and its interaction with CspE, suggests that its function may not be exactly the same as the rest of the DeoR family. Two models have been proposed here based on the available evidence, but further work is required to determine which, if either, of the two is correct. In the first, DeoT-mediated repression of target genes (achieved by binding at an operator sequence and inhibiting RNA polymerase association) is antagonised by CspE, allowing transcription. In the second, DeoT acts as a recruitment factor to a nascent RNA during transcription and promotes CspE-mediated antitermination. The first model lacks an apparent operator sequence, which was not identified either here or in previous studies (Elgrably-Weiss et al. 2006), while the second model is a major departure from known functions of the DeoR-family proteins, but does fit the available evidence.

Interaction of CspE with DeoT

CspE's interaction with DeoT suggest a novel function for the cold shock proteins in gene regulation. DeoT is not a well characterised protein, however, and its interaction with CspE perhaps raises more questions than it answers. DeoT appears to be an atypical example of the DeoR family of repressors, apparently controlling the expression of multiple physically discreet genetic systems (Elgrably-Weiss et al. 2006). The genes regulated have a general function of facilitating catabolism of non-optimal (or unusual) sources of energy for central metabolism. An interesting aspect of the interaction of the two proteins is that it appears to be toxic; a strain overexpressing both proteins suffered a slow growth phenotype.

Based on this slow growth phenotype, it is suspected that CspE is an antagonist of DeoT-mediated gene repression. Based on limited knowledge of the nature of the interaction and the precise function of DeoT, the suggested mechanisms for their co-function are speculative, and require significant further work to confirm. DeoT's regulation of catabolic processes might suggest a function for CspE in the response to low availability of nutrients. In *Caulobacter*, expression of a cold shock protein is upregulated by (p)ppGpp, the signalling molecule of the starvation response, while CspE's expression is increased in minimal medium in *E. coli* (Czapski & Trun 2014). cAMP, a molecule abundant when cellular energy availability is low, is another regulator of CspE (Uppal et al. 2011)

One possible model for CspE's function during starvation is one in which it is a co-regulator of genes for scavenging of alternative energy sources. Starvation is unlikely to have a significant impact on RNA structures, so the chaperone activity of CspE would not account for its upregulation there. An alternative function, such as in regulating gene expression, may. CspE's role in the response to low nutrient availability or a fall in energy status might be to signal requirement for expression of alternative catabolic pathways. A mediatory activity for catabolic genes would account for the observed interaction between CspE and DeoT. A similar model occurs in the *lac* operon, where both presence of and requirement for lactose are necessary for expression. Such a model is presented in Figure 6.1. A role in regulation of central metabolism may correlate to CspE's increase in expression in minimal medium, whereas in rich medium expression is lower.

The function of CspE at 37°C

Observations have been made which can be used to develop understanding of CspE's function during growth at normal temperatures. That the protein appears to localise to the nucleoid and ribosomes is in line with current models of its function, and supports the idea that its chief activities are those of an RNA chaperone and antiterminator. The discovery that CspE is apparently unable to dimerise *in vivo* answers a fundamental question of its biology, and will be of benefit to future studies.

That CspE is active as a monomer allows inference of its function. While it is likely to function as a chaperone of RNA (as a virtue of its ubiquity and non-specific binding capacity) it does not appear that this function is of great significance at 37°C. If chaperone activity occurs, it is infrequent enough not to be detected by the system employed. The weak interaction between CspE monomers at low temperatures is likely reflective of increased chaperone activity when RNA structures are more common. This theory is supported by a similar weak interaction observed between CspE and CsdA, another protein which co-localises to RNA substrates (Hunger et al. 2006). Absence of these weak signals at higher temperatures suggests another function must predominate. Additionally, a speculated role of CspE in chromosome condensation (Johnston et al. 2006) does not appear to be supported by the evidence here, as all current models of chromosome condensation require dimerisation of the condensing protein factors (de Vries 2010).

One function which is consistent with an active CspE monomer is that of transcription antitermination. Although long terminator sequences might require the binding of more than one CspE monomer, the majority of Rho-independent terminator sequences are 4 bp (Lesnik et al. 2001) and could be melted by monomeric CspE. An antitermination activity would account for constitutive expression, but would additionally require some factor to give specificity to its activity. Without a modulating factor, and being constitutively expressed, CspE would likely prevent termination indiscriminately, rendering the entire termination-antitermination system redundant. Although CspE has structural binding specificity, most terminators share a common stem-loop structure.

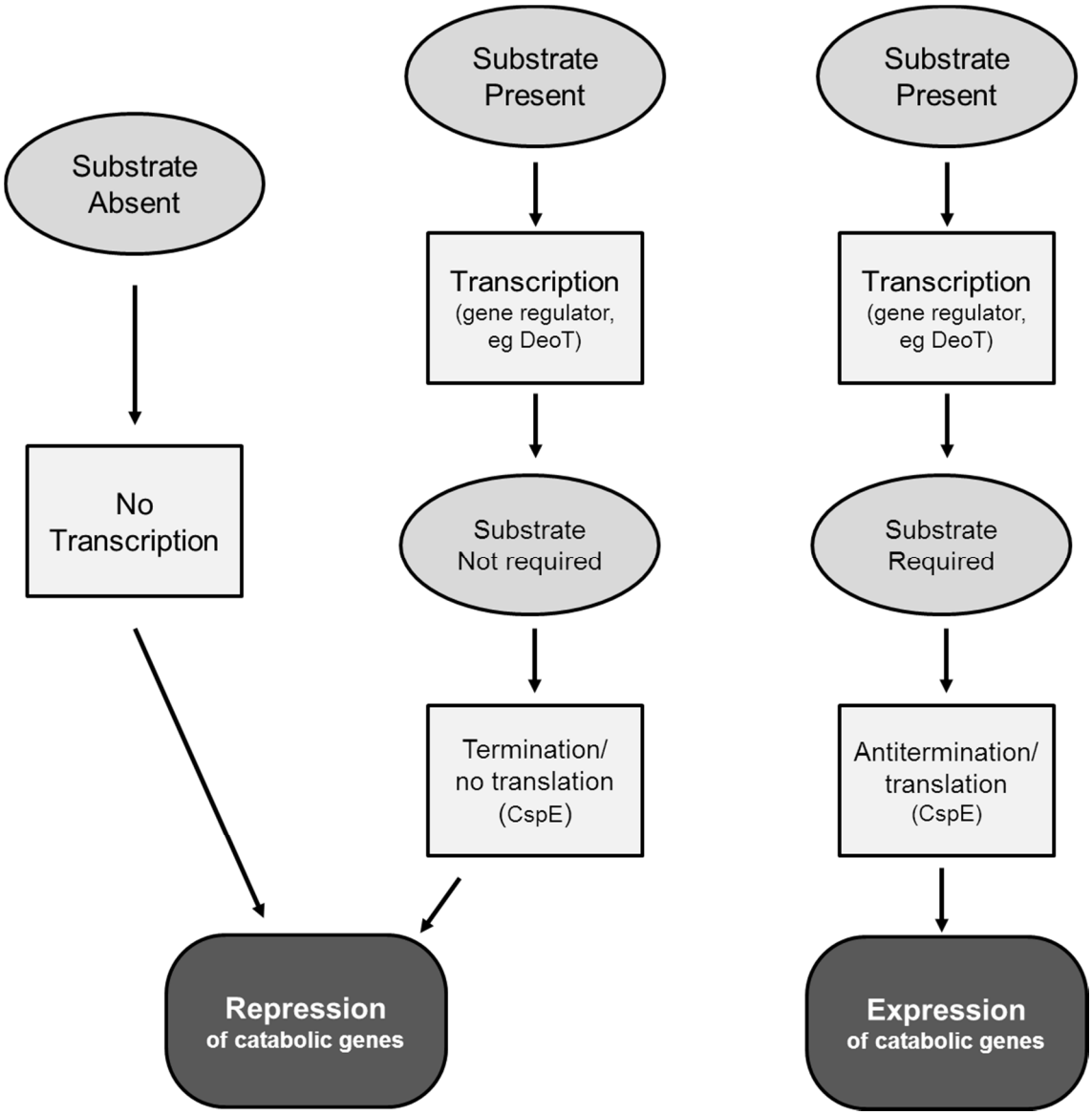


Figure 6.1. *Flow diagram of possible cooperation between gene regulators of specific catabolism genes (such as the maltose utilisation operons) and CspE, which may function as a mediator of the starvation response.*

Alternative modulators of CspE function

The second great motivation behind this study was to investigate the nature of CspE's observed selection of RNA substrates, given that the proteins themselves appear to have little sequence specificity. A number of possible cellular factors can impart specificity, as demonstrated by the antitermination systems of bacteria. The termination factor Rho is moderated by riboswitches (Proshkin et al., 2014), whose alterations of secondary structure can facilitate or prohibit binding of Rho. Another common mediator of termination is tRNA, which adopts alternative conformations depending on aminoacyl binding. These alternative states can stimulate antitermination, for example of animoacyl tRNA synthases when the levels of charged tRNA fall in a cell (Gutiérrez-Preciado et al. 2009).

Protein antiterminators can be regulated by interactions with other proteins. Many are two-component systems, which leave the antiterminator inactive in the absence of an extracellular signal. BglG is an activator of genes for β-glucoside metabolism, which is activated by phosphorylation in the presence of such substrates by its cognate membrane-bound sensor, BglF (Amster-Choder 2005). Signalling molecules can also moderate antiterminator activity. As a terminator protein, PyrR regulates synthesis of pyramidine nucleotides in response to availability of uridine. Where UTP is prevalent in the cell, PyrR is inactivated and cannot bind RNA, thus transcription continues. Where UMP (the product of the regulated operon) is present, PyrR is activated and transcription is attenuated.

Given that only a single protein interactor of CspE was identified herein, it is possible that the library was not comprehensive, and that some other protein regulators were missed by the library screen. CspE's antitermination activity could be regulated by a number of other factors, however. It is known to bind a number of small RNAs (McGibbon 2013), which could inhibit its antitermination activity by occupation of the single RNA binding site. Bacteria frequently utilise sRNA as mediators of stress responses (Hoe et al. 2013), so their expression changes with the environment and they could regulate CspE accordingly. Alternatively, the expression of CspE might regulate its activity. As well as being upregulated in the cold, CspE expression is controlled in response to growth phase and nutrient availability (Czapski & Trun 2014). Although the cold shock proteins seem to lack any small molecule binding sites through which they might be regulated directly, their transcription is regulated by the

second messengers cAMP and (p)ppGpp, as discussed above. This could be one mechanism by which the cold shock proteins are expressed in response to environmental signals.

Multiple activities of CspE

Despite the apparent presence of largely monomeric CspE at 37^{o}C, it is probable that antitermination is not its only function at that temperature. Being a non-specific binder of RNA, CspE is likely to have some role in translation. Indeed, global identification of CspE-bound RNA found hundreds of targets (McGibbon 2013) and although some contained termination sites not all did, and Csps also bound at other sites. This suggests that antitermination is not the exclusive function of CspE. Other roles may be performed by monomeric CspE, including that of an RNA chaperone. In the absence of a secondary structure, binding of CspE monomers would be less targeted, and protein units would bind randomly along the length of an RNA; this would account for a loss of the weak interaction signal at higher temperatures. At lower temperatures, a secondary structure would guide co-operative binding and closer association of CspE monomers.

CspE coupling of transcription and translation during normal growth is another activity which fits the available evidence, then. The protein may function as a general chaperone, binding to mRNA and preventing folding to maintain translation rates. Equally, cold shock proteins regulate RNA turnover and translation by binding at 5'-UTRs. They could modulate translation by altering accessibility of critical sites such as translation start codons and RNAse degradation sites, as fully discussed in Chapter 1. An involvement in translation might also link to the starvation response. The starvation response is linked to translation by RelA, which produces (p)ppGpp in response to stalled ribosomes which arise when tRNAs are uncharged. As discussed in Chapter 1, many factors which cause a translational block induce the expression of cold shock proteins, including low temperature, chloramphenicol treatment, and starvation. The control over catabolic genes proposed in Figure 6.1 could incorporate cold shock protein control at the level of either transcription or translation, or both.

CspE's potential as a drug target

Although the study of the cold shock proteins offers interesting insights into the molecular biology of gene regulation, there is another benefit to determination of their activity. The Csp proteins are connected to virulence in numerous species, and *Salmonella* strains lacking the cold shock proteins are highly attenuated in mice. This raises the possibility of their targeting as pathways for antimicrobial development. Although the findings detailed in this study are the very beginning of a drug development pathway, they may be interesting to study further. Antitermination is an exclusively prokaryotic process, and therefore would be an obvious mechanism to inhibit therapeutically. It is possible that attenuation in *csp* null strains is a manifestation of aberrant termination of transcription, and this would be an interesting avenue for further studies. Similarly, the apparent toxicity of the CspE-DeoT interaction might suggest a potential target for further investigation. Although many questions concerning the mechanism and function of the interaction, if it reduces cell viability then further study would be of great interest. Other proteins may be identified which interact with CspE: it could be that the observed toxicity of the CspE-DeoT interaction is indirect, and caused by occupation of CspE's interaction site by DeoT and prevention of other cellular activities.

Further studies

Immediate further work might focus on a more complete validation and characterisation of the CspE-DeoT interaction. The first priority will be to confirm the interaction using an alternative method; a pull-down experiment would be an obvious option. This would require the overexpression and purification of each interacting partner, either of which could be immobilised on an affinity column and used to selectively purify the other from a whole-cell extract. If each protein is able to specifically bind and purify the other, the interaction observed by two-hybrid analysis would be validated. The apparent slow-growth phenotype observed during BACTH experiments should be properly studied by simple growth curves. Having obtained purified CspE and DeoT, other methods such as surface plasmon resonance and isothermal titration calorimetry could be employed to further investigate the interaction.

Having validated the interaction, the subsequent challenge would be in its characterisation. Two models have been presented for possible function of the CspE-DeoT interaction, but both require experimental validation. The determination of DeoT's nucleotide binding preference (ie. for DNA or RNA) would be a useful first step, and should guide further study into the frequency, longevity, specificity, and activity of the interaction. DeoT's substrate specificity could be determined by ITC with DNA and RNA oligonucleotides to examine binding preference. That the regulatory targets of DeoT have been previously determined (in *E. coli*, at least) should be of help to further experiments. Northern blot or quantitative PCR of these targets would allow measurement of the interaction's effect on transcription, and thus confirmation of the antirepressor activity proposed here for CspE.

In the longer term, studies could investigate two aspects of cold shock protein function discussed above: their possible function as a signal of starvation, and their potential as targets for antimicrobial inhibition. Both would be significant undertakings. The role of the Csps in starvation could perhaps be investigated by comparative transcriptomics of wild type and Δ*csp* strains in and out of starvation conditions, while the response of *cspE* itself to specific signals of starvation such as (p)ppGpp could be quantitated at the RNA level by qPCR or at the protein level by quantitative Western blot. The development of small-molecule antimicrobials can take two approaches. Rational design would require mapping of the interaction site of DeoT and CspE, by systematic mutation or co-crystallisation. A screen for small-molecule modulators of the interaction could be assayed easily using the β-galactosidase readout of the BACTH method. General toxicity of small molecules could be assayed by growth inhibition in the wild-type, with specific Csp-related pathways being targeted by comparison of mutant and wild-type strains.

Conclusions

The studies described here were designed with the aim of expanding on current knowledge of several aspects of the functions of the cold shock proteins during normal growth. Analysis of CspE's potential to interact with itself suggested that it functions *in vivo* as a monomer, indicating that its principal function at 37°C is that of a transcriptional antiterminator, although other activities (including mRNA chaperoning) probably also occur concurrently. Similar analyses at low temperature were consistent with function as an RNA chaperone during cold shock, in line with current theories.

A whole-genome screen of interacting proteins suggested a possible means by which specificity for CspE's RNA substrates is conferred, although much work remains in order to clarify these findings. However, the observed interaction between CspE and DeoT reveals a number of possible novel roles for both proteins. It implicates CspE in gene regulation during the starvation reponse, and might indicate a new function as a direct regulator of transcription. This interaction can form the initial part of an interesting new body of work aimed at characterising the interaction of CspE and DeoT, CspE's role in the starvation response, and possible development of the cold shock proteins as targets for antimicrobial therapy.

References

Aguilar, P.S. et al., 2001. Molecular basis of thermosensing: a two-component signal transduction thermometer in Bacillus subtilis. *The EMBO Journal*, 20(7), pp.1681–1691.

Akbar, S. et al., 2001. New Family of Regulators in the Environmental Signaling Pathway Which Activates the General Stress Transcription Factor çB of Bacillus subtilis. *Journal of Bacteriology*, 183(4), pp.1329–1338.

Aldridge, M. et al., 2013. A novel bifunctional histone protein in Streptomyces: a candidate for structural coupling between DNA conformation and transcription during development and stress? *Nucleic Acids Research*, 41(9), pp.4813–4824.

Alfadhli, A. et al., 2013. Analysis of Small Molecule Ligands Targeting the HIV-1 Matrix Protein-RNA Binding Site. *Journal of Biological Chemistry*, 288(1), pp.666–676.

Altschul, S.F. et al., 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), pp.403–410.

Álvarez-Ordóñez, A. et al., 2010. Arginine and lysine decarboxylases and the Acid Tolerance Response of Salmonella Typhimurium. *International Journal of Food Microbiology*, 136(3), pp.278–282.

Álvarez-Ordóñez, A. et al., 2011. Salmonella spp. survival strategies within the host gastrointestinal tract. *Microbiology*, 157(12), pp.3268–3281.

Amouyal, M. et al., 1989. Single and double loop formation when deoR repressor binds to its natural operator sites. *Cell*, 58(3), pp.545–551.

Amster-Choder, O., 2005. The bgl sensory system: a transmembrane signaling pathway controlling transcriptional antitermination. *Current Opinion in Microbiology*, 8(2), pp.127–134.

Angelini, S., My, L. & Bouveret, E., 2012. Disrupting the Acyl Carrier Protein/SpoT Interaction In Vivo: Identification of ACP Residues Involved in the Interaction and Consequence on Growth. *PLoS ONE*, 7(4).

Aravind, L. et al., 2005. The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiology Reviews*, 29(2), pp.231–262.

Arifuzzaman, M. et al., 2006. Large-scale identification of protein–protein interaction of Escherichia coli K-12. *Genome Research*, 16(5), pp.686–691.

Audia, J.P., Webb, C.C. & Foster, J.W., 2001. Breaking through the acid barrier: An orchestrated response to proton stress by enteric bacteria. *International Journal of Medical Microbiology*, 291(2), pp.97–106.

Awano, N. et al., 2007. Complementation analysis of the cold-sensitive phenotype of the Escherichia coli csdA deletion strain. *Journal of Bacteriology*, 189(16), pp.5808–5815.

Awano, N. et al., 2010. Escherichia coli RNase R Has Dual Activities, Helicase and RNase. *Journal of Bacteriology*, 192(5), pp.1344–1352.

Awano, N., Inouye, M. & Phadtare, S., 2008. RNase Activity of Polynucleotide Phosphorylase Is Critical at Low Temperature in Escherichia coli and Is Complemented by RNase II. *Journal of Bacteriology*, 190(17), pp.5924–5933.

References

Bae, W. et al., 1999. Characterization of Escherichia coli cspE, whose product negatively regulates transcription of cspA, the gene for the major cold shock protein. *Molecular Microbiology*, 31(5), pp.1429–1441.

Bae, W. et al., 2000. Escherichia coli CspA-family RNA chaperones are transcription antiterminators. *Proceedings of the National Academy of Sciences of the United States of America*, 97(14), pp.7784–7789.

Bahloul, A., Boubrik, F. & Rouviere-Yaniv, J., 2001. Roles of Escherichia coli histone-like protein HU in DNA replication:HU-beta suppresses the thermosensitivity of dnaA46ts. *Biochimie*, 83(2), pp.219–229.

Bailey, T.L. et al., 2009. MEME Suite: tools for motif discovery and searching. *Nucleic Acids Research*, 37(suppl 2), pp.W202–W208.

Baker, J.L. et al., 2012. Widespread Genetic Switches and Toxicity Resistance Proteins for Fluoride. *Science*, 335(6065), pp.233–235.

Bakowski, M.A. et al., 2007. SopD acts cooperatively with SopB during Salmonella enterica serovar Typhimurium invasion. *Cellular Microbiology*, 9(12), pp.2839–2855.

Balaji, B. et al., 2005. Timing of Induction of Osmotically Controlled Genes in Salmonella enterica Serovar Typhimurium, Determined with Quantitative Real-Time Reverse Transcription-PCR. *Applied and Environmental Microbiology*, 71(12), pp.8273–8283.

Balhesteros, H. et al., 2010. CspC and CspD are essential for Caulobacter crescentus stationary phase survival. *Archives of Microbiology*, 192(9), pp.747–758.

Bang, I.S. et al., 2002. Autoinduction of the ompR response regulator by acid shock and control of the Salmonella enterica acid tolerance response. *Molecular Microbiology*, 44(5), pp.1235–1250.

Bang, I.-S. et al., 2006. Maintenance of Nitric Oxide and Redox Homeostasis by the Salmonella Flavohemoglobin Hmp. *Journal of Biological Chemistry*, 281(38), pp.28039–28047.

Barria, C., Malecki, M. & Arraiano, C.M., 2013. Bacterial adaptation to cold. *Microbiology*, 159(Pt 12), pp.2437–2443.

Al-Bassam, M.M. et al., 2014. Response Regulator Heterodimer Formation Controls a Key Stage in Streptomyces Development. *PLoS Genetics*, 10(8).

Battesti, A. & Bouveret, E., 2006. Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism. *Molecular Microbiology*, 62(4), pp.1048–1063.

Battesti, A. & Bouveret, E., 2012. The bacterial two-hybrid system based on adenylate cyclase reconstitution in Escherichia coli. *Methods*. 58(4) pp.325-334

Battesti, A., Majdalani, N. & Gottesman, S., 2011. The RpoS-Mediated General Stress Response in Escherichia coli. *Annual Review of Microbiology*, 65(1), pp.189–213.

Bayles, D.O. & Wilkinson, B.J., 2000. Osmoprotectants and cryoprotectants for Listeria monocytogenes. *Letters in Applied Microbiology*, 30(1), pp.23–27.

Bayles, K.W., 2000. The bactericidal action of penicillin: new clues to an unsolved mystery. *Trends in Microbiology*, 8(6), pp.274–278.

Beamer, L.J. & Pabo, C.O., 1992. Refined 1.8 A crystal structure of the lambda repressor-operator complex. *Journal of Molecular Biology*, 227(1), pp.177–196.

Bearson, B.L., Wilson, L. & Foster, J.W., 1998. A Low pH-Inducible, PhoPQ-Dependent Acid Tolerance Response Protects Salmonella typhimurium against Inorganic Acid Stress. *Journal of Bacteriology*, 180(9), pp.2409–2417.

Becker, D. et al., 2006. Robust Salmonella metabolism limits possibilities for new antimicrobials. *Nature*, 440(7082), pp.303–307.

Bepperling, A. et al., 2012. Alternative bacterial two-component small heat shock protein systems. *Proceedings of the National Academy of Sciences of the United States of America*, 109(50), pp.20407–20412.

Bernal-Bayard, J., Cardenal-Muñoz, E. & Ramos-Morales, F., 2010. The Salmonella type III secretion effector, salmonella leucine-rich repeat protein (SlrP), targets the human chaperone ERdj3. *The Journal of Biological Chemistry*, 285(21), pp.16360–16368.

Bernal-Bayard, J. & Ramos-Morales, F., 2009. Salmonella type III secretion effector SlrP is an E3 ubiquitin ligase for mammalian thioredoxin. *The Journal of Biological Chemistry*, 284(40), pp.27587–27595.

Berry, J.-L. et al., 2013. Functional Analysis of the Interdependence between DNA Uptake Sequence and Its Cognate ComP Receptor during Natural Transformation in Neisseria Species. *PLoS Genet*, 9(12), p.e1004014.

Bishop, R.E., 2005. The lipid A palmitoyltransferase PagP: molecular mechanisms and role in bacterial pathogenesis. *Molecular Microbiology*, 57(4), pp.900–912.

Bochkarev, A. & Bochkareva, E., 2004. From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Current Opinion in Structural Biology*, 14(1), pp.36–42.

Bogan, A.A. & Thorn, K.S., 1998. Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology*, 280(1), pp.1–9.

Bourret, T.J. et al., 2008. Nitric Oxide Antagonizes the Acid Tolerance Response that Protects Salmonella against Innate Gastric Defenses. *PLoS ONE*, 3(3), p.e1833.

Bowe, F. et al., 1998. At Least Four Percent of the Salmonella typhimurium Genome Is Required for Fatal Infection of Mice. *Infection and Immunity*, 66(7), pp.3372–3377.

Brandi, A. et al., 1999. Massive presence of the Escherichia coli “majorcold-shock protein” CspA under non-stress conditions. *EMBO J*, 18(6), pp.1653–1659.

Brandi, A., Pon, C. & Gualerzi, C., 1994. Interaction of the main cold shock protein CS7.4 (CspA) of Escherichia coli with the promoter region of hns. *Biochimie*, 76(10-11), pp.1090–1098.

Brandi, A. & Pon, C.L., 2012. Expression of Escherichia coli cspA during early exponential growth at 37 °C. *Gene*, 492(2), pp.382–388.

Brandi, A., Pon, C.L. & Gualerzi, C.O., 1994. Interaction of the main cold shock protein CS7.4 (CspA) of Escherichia coli with the promoter region of hns. *Biochimie*, 76(10-11), pp.1090–1098.

References

Bresolin, G. et al., 2006. Transcriptional analysis of long-term adaptation of Yersinia enterocolitica to low-temperature growth. *Journal of Bacteriology*, 188(8), pp.2945–2958.

Broedersz, C.P. et al., 2014. Condensation and localization of the partitioning protein ParB on the bacterial chromosome. *Proceedings of the National Academy of Sciences*, 111(24), pp.8809–8814.

Brogan, A.P. et al., 2005. Bismuth–dithiol inhibition of the Escherichia coli rho transcription termination factor. *Journal of Inorganic Biochemistry*, 99(3), pp.841–851.

Bruncko, M. et al., 2007. Studies Leading to Potent, Dual Inhibitors of Bcl-2 and Bcl-xL. *Journal of Medicinal Chemistry*, 50(4), pp.641–662.

Bubunenko, M. et al., 2013. Nus transcription elongation factors and RNase III modulate small ribosome subunit biogenesis in Escherichia coli. *Molecular Microbiology*, 87(2), pp.382–393.

Buckle, A.M., Schreiber, G. & Fersht, A.R., 1994. Protein-protein recognition: crystal structural analysis of a barnase-barstar complex at 2.0-A resolution. *Biochemistry*, 33(30), pp.8878–8889.

Butland, G. et al., 2005. Interaction network containing conserved and essential protein complexes in Escherichia coli. *Nature*, 433(7025), pp.531–537.

Büttner, K., Nehring, S. & Hopfner, K.-P., 2007. Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nature Structural & Molecular Biology*, 14(7), pp.647–652.

Cabrera, J.E. & Jin, D.J., 2003. The distribution of RNA polymerase in Escherichia coli is dynamic and sensitive to environmental cues. *Molecular Microbiology*, 50(5), pp.1493–1505.

Cadman, C.J. & McGlynn, P., 2004. PriA helicase and SSB interact physically and functionally. *Nucleic Acids Research*, 32(21), pp.6378–6387.

Cairney, J., Booth, I.R. & Higgins, C.F., 1985. Salmonella typhimurium proP gene encodes a transport system for the osmoprotectant betaine. *Journal of Bacteriology*, 164(3), pp.1218–1223.

Camejo, A. et al., 2009. In Vivo Transcriptional Profiling of Listeria monocytogenes and Mutagenesis Identify New Virulence Factors Involved in Infection. *PLoS Pathog*, 5(5), p.e1000449.

Cameron, A.D.S., Stoebel, D.M. & Dorman, C.J., 2011. DNA supercoiling is differentially regulated by environmental factors and FIS in Escherichia coli and Salmonella enterica. *Molecular Microbiology*, 80(1), pp.85–101.

Campos, E. et al., 2004. Regulation of expression of the divergent ulaG and ulaABCDEF operons involved in LaAscorbate dissimilation in Escherichia coli. *Journal of Bacteriology*, 186(6), pp.1720–1728.

Carty, S.M., Sreekumar, K.R. & Raetz, C.R.H., 1999. Effect of Cold Shock on Lipid A Biosynthesis inEscherichia coli. *Journal of Biological Chemistry*, 274(14), pp.9677 –9685.

Cevc, G., 1991. How membrane chain-melting phase-transition temperature is affected by the lipid chain asymmetry and degree of unsaturation: an effective chain-length model. *Biochemistry*, 30(29), pp.7186–7193.

Chaikam, V. & Karlson, D.T., 2010. Comparison of structure, function and regulation of plant cold shock domain proteins to bacterial and animal cold shock domain proteins. *BMB reports*, 43(1), pp.1–8.

Chakraborty, A. et al., 2014. Conformational Adaptation in the E. coli Sigma 32 Protein in Response to Heat Shock. *The Journal of Physical Chemistry B*, 118(18), pp.4793–4802.

Chakravortty, D., Hansen-Wester, I. & Hensel, M., 2002. Salmonella Pathogenicity Island 2 Mediates Protection of Intracellular Salmonella from Reactive Nitrogen Intermediates. *The Journal of Experimental Medicine*, 195(9), pp.1155 –1166.

Chanda, P.K. et al., 2009. Antibiotics, arsenate and H2O2 induce the promoter of Staphylococcus aureus cspC gene more strongly than cold. *Journal of Basic Microbiology*, 49(2), pp.205–211.

Charollais, J., Dreyfus, M. & Iost, I., 2004. CsdA, a cold-shock RNA helicase from Escherichia coli, is involved in the biogenesis of 50S ribosomal subunit. *Nucleic Acids Research*, 32(9), pp.2751–2759.

Chaudhuri, R.R. et al., 2013. Comprehensive Assignment of Roles for Salmonella Typhimurium Genes in Intestinal Colonization of Food-Producing Animals. *PLoS Genet*, 9(4), p.e1003456.

Cheung, K.J. et al., 2003. A microarray-based antibiotic screen identifies a regulatory role for supercoiling in the osmotic stress response of Escherichia coli. *Genome Research*, 13(2), pp.206–215.

Chiang, S.M. & Schellhorn, H.E., 2012. Regulators of oxidative stress response genes in Escherichia coli and their functional conservation in bacteria. *Archives of Biochemistry and Biophysics*, 525(2), pp.161–169.

Chodavarapu, S. et al., 2008. Escherichia coli DnaA interacts with HU in initiation at the E. coli replication origin. *Molecular Microbiology*, 67(4), pp.781–792.

Christiansen, J.K. et al., 2004. The RNA-Binding Protein Hfq of Listeria monocytogenes: Role in Stress Tolerance and Virulence. *Journal of Bacteriology*, 186(11), pp.3355–3362.

Cimdins, A. et al., 2013. Differential control of Salmonella heat shock operons by structured mRNAs. *Molecular Microbiology*, 89(4), pp.715–731.

Clarke, L. & Carbon, J., 1976. A colony bank containing synthetic Col El hybrid plasmids representative of the entire E. coli genome. *Cell*, 9(1), pp.91–99.

Cosgriff, S. et al., 2010. Dimerization and DNA-dependent aggregation of the Escherichia coli nucleoid protein and chaperone CbpA. *Molecular Microbiology*, 77(5), pp.1289–1300.

Craig, E.A., Weissman, J.S. & Horwich, A.L., 1994. Heat shock proteins and molecular chaperones: Mediators of protein conformation and turnover in the cell. *Cell*, 78(3), pp.365–372.

Craig, J.E. et al., 1998. Expression of the cold-shock gene cspB in Salmonella typhimurium occurs below a threshold temperature. *Microbiology*, 144(3), pp.697–704.

Cronan, J.E., 1975. Thermal regulation of the membrane lipid composition of Escherichia coli. Evidence for the direct control of fatty acid synthesis. *Journal of Biological Chemistry*, 250(17), pp.7074–7077.

References

Crump, J.A., Luby, S.P. & Mintz, E.D., 2004. The global burden of typhoid fever. *Bulletin of the World Health Organization*, 82(5), pp.346–353.

Csonka, L.N. et al., 1994. The accumulation of glutamate is necessary for optimal growth of Salmonella typhimurium in media of high osmolality but not induction of the proU operon. *Journal of Bacteriology*, 176(20), pp.6324–6333.

Czapski, T.R. & Trun, N., 2014. Expression of csp genes in E. coli K-12 in defined rich and defined minimal media during normal growth, and after cold-shock. *Gene*.

Dame, R.T., 2005. The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Molecular Microbiology*, 56(4), pp.858–870.

Dandanell, G., Norris, K. & Hammer, K., 1991. Long-Distance deoR Regulation of Gene Expression in Escherichia colia. *Annals of the New York Academy of Sciences*, 646(1), pp.19–30.

Darii, E. et al., 2014. Structural studies of the sBBI/trypsin non-covalent complex using covalent modification and mass spectrometry. *Rapid communications in mass spectrometry: RCM*, 28(5), pp.413–429.

Datta, P.P. et al., 2007. Structural Aspects of RbfA Action during Small Ribosomal Subunit Assembly. *Molecular Cell*, 28(3), pp.434–445.

De Vries, R., 2010. DNA condensation in bacteria: Interplay between macromolecular crowding and nucleoid proteins. *Biochimie*, 92(12), pp.1715–1721.

Deiwick, J. et al., 1999. Environmental regulation of Salmonella pathogenicity island 2 gene expression. *Molecular Microbiology*, 31(6), pp.1759–1773.

Deiwick, J. et al., 2006. The translocated Salmonella effector proteins SseF and SseG interact and are required to establish an intracellular replication niche. *Infection and Immunity*, 74(12), pp.6965–6972.

Deng, X., Li, Z. & Zhang, W., 2012. Transcriptome sequencing of Salmonella enterica serovar Enteritidis under desiccation and starvation stress in peanut oil. *Food Microbiology*, 30(1), pp.311–315.

Dibb-Fuller, M.P. et al., 1999. Fimbriae- and flagella-mediated association with and invasion of cultured epithelial cells by Salmonella enteritidis. *Microbiology*, 145 (Pt 5), pp.1023–1031.

Domain, F., Bina, X.R. & Levy, S.B., 2007. Transketolase A, an enzyme in central metabolism, derepresses the marRAB multiple antibiotic resistance operon of Escherichia coli by interaction with MarR. *Molecular Microbiology*, 66(2), pp.383–394.

Dong, G. et al., 2004. Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch. *The EMBO Journal*, 23(5), pp.1000–1007.

Dove, S.L., Joung, J.K. & Hochschild, A., 1997. Activation of prokaryotic transcription through arbitrary protein–protein contacts. *Nature*, 386(6625), pp.627–630.

Drlica, K., 1992. Control of bacterial DNA supercoiling. *Molecular Microbiology*, 6(4), pp.425–433.

Duval, B.D. et al., 2010. Altered Growth, Pigmentation, and Antimicrobial Susceptibility Properties of Staphylococcus aureus Due to Loss of the Major Cold Shock Gene cspB. *Antimicrobial Agents and Chemotherapy*, 54(6), pp.2283–2290.

Elgrably-Weiss, M. et al., 2006. DeoT, a DeoR-type transcriptional regulator of multiple target genes. *FEMS microbiology letters*, 254(1), pp.141–148.

Ellenbecker, M., Lanchy, J.-M. & Lodmell, J.S., 2012. Identification of Rift Valley Fever Virus Nucleocapsid Protein-RNA Binding Inhibitors Using a High-Throughput Screening Assay. *Journal of Biomolecular Screening*, 17(8), pp.1062–1070.

Erova, T.E. et al., 2008. Cold Shock Exoribonuclease R (VacB) Is Involved in Aeromonas hydrophila Pathogenesis. *Journal of Bacteriology*, 190(10), pp.3467–3474.

Escapa, I.F. et al., 2013. The role of GlpR repressor in Pseudomonas putida KT2440 growth and PHA production from glycerol. *Environmental Microbiology*, 15(1), pp.93–110.

Etchegaray, J.P. & Inouye, M., 1999. CspA, CspB, and CspG, major cold shock proteins of Escherichia coli, are induced at low temperature under conditions that completely block protein synthesis. *Journal of Bacteriology*, 181(6), pp.1827–1830.

Etchegaray, J.-P., Jones, P.G. & Inouye, M., 1996. Differential thermoregulation of two highly homologous cold-shock genes, cspA and cspB, of Escherichia coli. *Genes to Cells*, 1(2), pp.171–178.

Evans, M.R. et al., 2011. Analysis of the ArcA regulon in anaerobically grown Salmonella enterica sv. Typhimurium. *BMC Microbiology*, 11(1), p.58.

Fabrega, A. & Vila, J., 2013. Salmonella enterica Serovar Typhimurium Skills To Succeed in the Host: Virulence and Regulation. *Clinical Microbiology Reviews*, 26(2), pp.308–341.

Fang, F.C., 2011. Antimicrobial Actions of Reactive Oxygen Species. *mBio*, 2(5), pp.e00141–11.

Fang, L. et al., 1997. Promoter-independent cold-shock induction of cspA and its derepression at 37°C by mRNA stabilization. *Molecular Microbiology*, 23(2), pp.355–364.

Feng, Y., 2001. Escherichia coli Poly(A)-binding Proteins That Interact with Components of Degradosomes or Impede RNA Decay Mediated by Polynucleotide Phosphorylase and RNase E. *Journal of Biological Chemistry*, 276(34), pp.31651–31656.

Ferreira, E. et al., 2013. Protein interaction studies point to new functions for Escherichia coli glyceraldehyde-3-phosphate dehydrogenase. *Research in Microbiology*, 164(2), pp.145–154.

Fields, S. & Song, O., 1989. A novel genetic system to detect protein-protein interactions. *Nature*, 340(6230), pp.245–246.

Fierer, J. & Guiney, D.G., 2001. Diverse virulence traits underlying different clinical outcomes of Salmonella infection. *Journal of Clinical Investigation*, 107(7), pp.775–780.

Fink, R.C. et al., 2007. FNR Is a Global Regulator of Virulence and Anaerobic Metabolism in Salmonella enterica Serovar Typhimurium (ATCC 14028s). *Journal of Bacteriology*, 189(6), pp.2262–2273.

Finn, S., Condell, O., et al., 2013. Mechanisms of survival, responses and sources of Salmonella in low-moisture environments. *Frontiers in Microbiology*, 4.

References

Finn, S., Händler, K., et al., 2013. ProP Is Required for the Survival of Desiccated Salmonella enterica Serovar Typhimurium Cells on a Stainless Steel Surface. *Applied and Environmental Microbiology*, 79(14), pp.4376–4384.

Foster, J.W. & Hall, H.K., 1992. Effect of Salmonella typhimurium ferric uptake regulator (fur) mutations on iron- and pH-regulated protein synthesis. *Journal of Bacteriology*, 174(13), pp.4317–4323.

Foster, J.W. & Spector, M.P., 1995. How Salmonella Survive Against the Odds. *Annual Review of Microbiology*, 49(1), pp.145–174.

Franceschini, A. et al., 2013. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Research*, 41, pp.D808–815.

Fraser, K.R. et al., 2006. Global effects of homocysteine on transcription in Escherichia coli: induction of the gene for the major cold-shock protein, CspA. *Microbiology*, 152(Pt 8), pp.2221–2231.

Frossard, S.M. et al., 2012. Identification of a Third Osmoprotectant Transport System, the OsmU System, in Salmonella enterica. *Journal of Bacteriology*, 194(15), pp.3861–3871.

Fujita, Y., Matsuoka, H. & Hirooka, K., 2007. Regulation of fatty acid metabolism in bacteria. *Molecular Microbiology*, 66(4), pp.829–839.

Fu, Y. & Galán, J.E., 1999. A salmonella protein antagonizes Rac-1 and Cdc42 to mediate host-cell recovery after bacterial invasion. *Nature*, 401(6750), pp.293–297.

Garces, F. et al., 2008. Quaternary structural transitions in the DeoR-type repressor UlaR control transcriptional readout from the L-ascorbate utilization regulon in Escherichia coli. *Biochemistry*, 47(44), pp.11424–11433.

García, P. et al., 2014. Diversity of plasmids encoding virulence and resistance functions in Salmonella enterica subsp. enterica serovar Typhimurium monophasic variant 4,[5],12:i:- strains circulating in Europe. *PloS One*, 9(2), p.e89635.

Garwin, J.L. & Cronan, J.E., 1980. Thermal modulation of fatty acid synthesis in Escherichia coli does not involve de novo enzyme synthesis. *Journal of Bacteriology*, 141(3), pp.1457–1459.

Garwin, J.L., Klages, A.L. & Cronan, J.E., 1980. Structural, enzymatic, and genetic studies of beta-ketoacyl-acyl carrier protein synthases I and II of Escherichia coli. *Journal of Biological Chemistry*, 255(24), pp.11949–11956.

Genschel, J., Curth, U. & Urbanke, C., 2000. Interaction of E. coli single-stranded DNA binding protein (SSB) with exonuclease I. The carboxy-terminus of SSB is the recognition site for the nuclease. *Biological Chemistry*, 381(3), pp.183–192.

Gharbi, S. et al., 1985. Multiple controls exerted on in vivo expression of the pepN gene in Escherichia coli: studies with pepN-lacZ operon and protein fusion strains. *Journal of Bacteriology*, 163(3), pp.1191–1195.

Giangrossi, M. et al., 2007. Cold-shock-induced de novo transcription and translation of infA and role of IF1 during cold adaptation. *Molecular Microbiology*, 64(3), pp.807–821.

Giangrossi, M. et al., 2001. Different in vivo localization of the Escherichia coli proteins CspD and CspA. *FEMS Microbiology Letters*, 202(2), pp.171–176.

Giangrossi, M. et al., 2002. Selective expression of the β-subunit of nucleoid-associated protein HU during cold shock in Escherichia coli. *Molecular Microbiology*, 44(1), pp.205–216.

Giangrossi, M., Gualerzi, C.O. & Pon, C.L., 2001. Mutagenesis of the downstream region of the Escherichia coli hns promoter. *Biochimie*, 83(2), pp.251–259.

Giaquinto, L. et al., 2007. Structure and Function of Cold Shock Proteins in Archaea. *Journal of Bacteriology*, 189(15), pp.5738–5748.

Giuliodori, A.M. et al., 2007. Cold-stress-induced de novo expression of infC and role of IF3 in cold-shock translational bias. *RNA*, 13(8), pp.1355–1365.

Giuliodori, A.M. et al., 2010. The cspA mRNA Is a Thermosensor that Modulates Translation of the Cold-Shock Protein CspA. *Molecular Cell*, 37(1), pp.21–33.

Godfrey, H.P., Bugrysheva, J.V. & Cabello, F.C., 2002. The role of the stringent response in the pathogenesis of bacterial infections. *Trends in Microbiology*, 10(8), pp.349–351.

Godin-Roulling, A. et al., 2014. Functional adaptations of the bacterial chaperone trigger factor to extreme environmental temperatures. *Environmental Microbiology*, p.n/a–n/a.

Goldenberg, D. et al., 1997. Role of Escherichia coli cspA promoter sequences and adaptation of translational apparatus in the cold shock response. *Molecular and General Genetics* , 256(3), pp.282–290.

Gormley, F.J. et al., 2011. A 17-year review of foodborne outbreaks: describing the continuing decline in England and Wales (1992-2008). *Epidemiology and Infection*, 139(5), pp.688–699.

Granneman, S. et al., 2009. Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proceedings of the National Academy of Sciences*, 106(24), pp.9613–9618.

Graumann, P. et al., 1997. A family of cold shock proteins in Bacillus subtilis is essential for cellular growth and for efficient protein synthesis at optimal and low temperatures. *Molecular Microbiology*, 25(04), pp.741–756.

Graumann, P.L. & Marahiel, M.A., 1999. Cold shock proteins CspB and CspC are major stationary-phase-induced proteins in Bacillus subtilis. *Archives of Microbiology*, 171(2), pp.135–138.

Gualerzi, C., 2003. Transcriptional and Post-transcriptional Control of Cold-shock Genes. *Journal of Molecular Biology*, 331(3), pp.527–539.

Gundogdu, O. et al., 2007. Re-annotation and re-analysis of the Campylobacter jejuni NCTC11168 genome sequence. *BMC genomics*, 8, p.162.

Guo, F. & Adhya, S., 2007. Spiral structure of Escherichia coli HUαβ provides foundation for DNA supercoiling. *Proceedings of the National Academy of Sciences*, 104(11), pp.4309–4314.

Gutiérrez-Preciado, A. et al., 2009. Biochemical Features and Functional Implications of the RNA-Based T-Box Regulatory Mechanism. *Microbiology and Molecular Biology Reviews*, 73(1), pp.36–61.

Gyles, C.L., 2008. Antimicrobial resistance in selected bacteria from poultry. *Animal Health Research Reviews*, 9(Special Issue 02), pp.149–158.

References

Hammerle, H. et al., 2013. Duplex formation between the sRNA DsrA and rpoS mRNA is not sufficient for efficient RpoS synthesis at low temperature. *RNA Biology*, 10(12), pp.1834–1841.

Ham, R.C.H.J. van et al., 2003. Reductive genome evolution in Buchnera aphidicola. *Proceedings of the National Academy of Sciences*, 100(2), pp.581–586.

Hankins, J.S. et al., 2007. Role of RNA Structure and Susceptibility to RNase E in Regulation of a Cold Shock mRNA, cspA mRNA. *Journal of Bacteriology*, 189(12), pp.4353–4358.

Hanna, M.M. & Liu, K., 1998. Nascent RNA in transcription complexes interacts with CspE, a small protein in E. coli implicated in chromatin condensation. *Journal of Molecular Biology*, 282(2), pp.227–239.

Harami, G.M., Gyimesi, M. & Kovács, M., 2013. From keys to bulldozers: expanding roles for winged helix domains in nucleic-acid-binding proteins. *Trends in Biochemical Sciences*, 38(7), pp.364–371.

Harker, K.S. et al., 2014. National outbreaks of Salmonella infection in the UK, 2000–2011. *Epidemiology & Infection*, 142(03), pp.601–607.

Harman, J.G., 2001. Allosteric regulation of the cAMP receptor protein. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1547(1), pp.1–17.

Hatfield, G.W. & Benham, C.J., 2002. DNA topology-mediated control of global gene expression in Escherichia coli. *Annual Review of Genetics*, 36, pp.175–203.

Heaton, J. c. & Jones, K., 2008. Microbial contamination of fruit and vegetables and the behaviour of enteropathogens in the phyllosphere: a review. *Journal of Applied Microbiology*, 104(3), pp.613–626.

Hebrard, M. et al., 2009. Redundant Hydrogen Peroxide Scavengers Contribute to Salmonella Virulence and Oxidative Stress Resistance. *Journal of Bacteriology*, 191(14), pp.4605–4614.

He, M.M. et al., 2005. Small-Molecule Inhibition of TNF-α. *Science*, 310(5750), pp.1022–1025.

Henard, C.A. et al., 2010. Control of Redox Balance by the Stringent Response Regulatory Protein Promotes Antioxidant Defenses of Salmonella. *Journal of Biological Chemistry*, 285(47), pp.36785–36793.

Henard, C.A. & Vázquez-Torres, A., 2012. DksA-Dependent Resistance of Salmonella enterica Serovar Typhimurium against the Antimicrobial Activity of Inducible Nitric Oxide Synthase. *Infection and Immunity*, 80(4), pp.1373–1380.

Henard, C.A. & Vazquez-Torres, A., 2011. Nitric oxide and Salmonella pathogenesis. *Frontiers in Microbiology*, 2, p.84.

Herzog, F. et al., 2012. Structural Probing of a Protein Phosphatase 2A Network by Chemical Cross-Linking and Mass Spectrometry. *Science*, 337(6100), pp.1348–1352.

He, Y., Andersen, G.R. & Nielsen, K.H., 2010. Structural basis for the function of DEAH helicases. *EMBO reports*, 11(3), pp.180–186.

Hoe, C.-H. et al., 2013. Bacterial sRNAs: Regulation in stress. *International Journal of Medical Microbiology*, 303(5), pp.217–229.

Hoffmann, A. et al., 2012. Concerted Action of the Ribosome and the Associated Chaperone Trigger Factor Confines Nascent Polypeptide Folding. *Molecular Cell*, 48(1), pp.63–74.

Hoiseth, S.K. & Stocker, B.A., 1981. Aromatic-dependent Salmonella typhimurium are non-virulent and effective as live vaccines. *Nature*, 291(5812), pp.238–239.

Holden, N., 2000. *The Cold Shock Response of Salmonella enterica serovar Typhumirium.* PhD Thesis. University of Edinburgh.

Holt, J.G. et al., 2000. *Bergey's Manual of Determinative Bacteriology* 9th ed., Williams and Wilkins.

Hooper, L.V. et al., 1999. A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. *Proceedings of the National Academy of Sciences of the United States of America*, 96(17), pp.9833–9838.

Hopkins, A.L. & Groom, C.R., 2002. The druggable genome. *Nature Reviews Drug Discovery*, 1(9), pp.727–730.

Horst, S.A. et al., 2010. Thiol Peroxidase Protects Salmonella enterica from Hydrogen Peroxide Stress In Vitro and Facilitates Intracellular Growth. *Journal of Bacteriology*, 192(11), pp.2929–2932.

Howells, A.M. et al., 2002. Role of trehalose biosynthesis in environmental survival and virulence of Salmonella enterica serovar typhimurium. *Research in Microbiology*, 153(5), pp.281–287.

Hryckowian, A.J. et al., 2014. IraL Is an RssB Anti-adaptor That Stabilizes RpoS during Logarithmic Phase Growth in Escherichia coli and Shigella. *mBio*, 5(3).

Hu, K.H. et al., 1996. Overproduction of three genes leads to camphor resistance and chromosome condensation in Escherichia coli. *Genetics*, 143(4), pp.1521–1532.

Humphrey, T., 2006. Are happy chickens safer chickens? Poultry welfare and disease susceptibility 1. *British Poultry Science*, 47(4), pp.379–391.

Hunger, K. et al., 2006. Cold-induced putative DEAD box RNA helicases CshA and CshB are essential for cold adaptation and interact with cold shock protein B in Bacillus subtilis. *Journal of Bacteriology*, 188(1), pp.240–248.

Hunger, K., Beckering, C.L. & Marahiel, M.A., 2004. Genetic evidence for the temperature-sensing ability of the membrane domain of the Bacillus subtilis histidine kinase DesK. *FEMS Microbiology Letters*, 230(1), pp.41–46.

Hunke, S., Keller, R. & Müller, V.S., 2012. Signal integration by the Cpx-envelope stress system. *FEMS Microbiology Letters*, 326(1), pp.12–22.

Hurme, R. et al., 1997. A Proteinaceous Gene Regulatory Thermometer in Salmonella. *Cell*, 90(1), pp.55–64.

Hurme, R. et al., 1996. DNA Binding Exerted by a Bacterial Gene Regulator with an Extensive Coiled-coil Domain. *Journal of Biological Chemistry*, 271(21), pp.12626–12631.

Husain, M. et al., 2010. Redox sensor SsrB Cys203 enhances Salmonella fitness against nitric oxide generated in the host immune response to oral infection. *Proceedings of the National Academy of Sciences*, 107(32), pp.14396–14401.

Hutchinson, I.W., 2005. *Characterisation of the Cold Shock Response of Salmonella enterica sv. Typhimurium - Genomic and Genetic analysis*. PhD Thesis, University of Edinburgh.

Hu, Y., Benedik, M.J. & Wood, T.K., 2011. Antitoxin DinJ influences the general stress response through transcript stabilizer CspE. *Environmental Microbiology*. 14(3) pp.669-679.

Iost, I. & Dreyfus, M., 2006. DEAD-box RNA helicases in Escherichia coli. *Nucleic Acids Research*, 34(15), pp.4189–4197.

Ishihama, Y. et al., 2002. Microcolumns with self-assembled particle frits for proteomics. *Journal of Chromatography A*, 979(1–2), pp.233–239.

Jackson, B.R. et al., 2013. Outbreak-associated Salmonella enterica Serotypes and Food Commodities, United States, 1998–2008. *Emerging Infectious Diseases*, 19(8), pp.1239–1244.

Jacob, F. & Monod, J., 1961. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology*, 3, pp.318–356.

Jantsch, J., Chikkaballi, D. & Hensel, M., 2011. Cellular aspects of immunity to intracellular Salmonella enterica. *Immunological Reviews*, 240(1), pp.185–195.

Jervis, A.J. & Green, J., 2007. In Vivo Demonstration of FNR Dimers in Response to Lower O_2 Availability. *Journal of Bacteriology*, 189(7), pp.2930–2932.

Jiang, W., Hou, Y. & Inouye, M., 1997. CspA, the Major Cold-shock Protein of Escherichia coli, Is an RNA Chaperone. *Journal of Biological Chemistry*, 272(1), pp.196–202.

Jiang, W., Jones, P. & Inouye, M., 1993. Chloramphenicol induces the transcription of the major cold shock gene of Escherichia coli, cspA. *Journal of Bacteriology*, 175(18), pp.5824–5828.

Johnston, D. et al., 2006. Specificity of DNA binding and dimerization by CspE from Escherichia coli. *The Journal of Biological Chemistry*, 281(52), pp.40208–40215.

Jones, B.D., Ghori, N. & Falkow, S., 1994. Salmonella typhimurium initiates murine infection by penetrating and destroying the specialized epithelial M cells of the Peyer's patches. *The Journal of Experimental Medicine*, 180(1), pp.15–23.

Jones, P. et al., 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 30(9), pp.1236–1240.

Jones, P.G. et al., 1992. DNA gyrase, CS7.4, and the cold shock response in Escherichia coli. *Journal of Bacteriology*, 174(18), pp.5798–5802.

Jones, P.G. & Inouye, M., 1996. RbfA, a 30S ribosomal binding factor, is a cold-shock protein whose absence triggers the cold-shock response. *Molecular Microbiology*, 21(6), pp.1207–1218.

Kamashev, D. et al., 2008. HU binds and folds single-stranded DNA. *Nucleic Acids Research*, 36(3), pp.1026–1036.

Kandror, O., DeLeon, A. & Goldberg, A.L., 2002. Trehalose synthesis is induced upon exposure of Escherichia coli to cold and is essential for viability at low temperatures. *Proceedings of the National Academy of Sciences of the United States of America*, 99(15), pp.9727–9732.

Kanjee, U., Ogata, K. & Houry, W.A., 2012. Direct binding targets of the stringent response alarmone (p)ppGpp. *Molecular Microbiology*, 85(6), pp.1029–1043.

Karimova, G. et al., 1998. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 95(10), pp.5752–5756.

Karimova, G., Dautin, N. & Ladant, D., 2005. Interaction network among Escherichia coli membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis. *Journal of Bacteriology*, 187(7), pp.2233–2243.

Karimova, G., Robichon, C. & Ladant, D., 2009. Characterization of YmgF, a 72-Residue Inner Membrane Protein That Associates with the Escherichia coli Cell Division Machinery. *Journal of Bacteriology*, 191(1), pp.333–346.

Karlinsey, J.E. et al., 2012. The NsrR regulon in nitrosative stress resistance of Salmonella enterica serovar Typhimurium. *Molecular Microbiology*, 85(6), pp.1179–1193.

Katzif, S. et al., 2005. CspA regulates pigment production in Staphylococcus aureus through a SigB-dependent mechanism. *Journal of Bacteriology*, 187(23), pp.8181–8184.

Katzif, S. et al., 2003. The major cold shock gene, cspA, is involved in the susceptibility of Staphylococcus aureus to an antimicrobial peptide of human cathepsin G. *Infection and Immunity*, 71(8), pp.4304–4312.

Kawashima, T. et al., 1996. The structure of the Escherichia coli EF-Tu· EF-Ts complex at 2.5 Å resolution. *Nature*, 379(6565), pp.511–518.

Kazmierczak, M.J., Wiedmann, M. & Boor, K.J., 2005. Alternative Sigma Factors and Their Roles in Bacterial Virulence. *Microbiology and Molecular Biology Reviews*, 69(4), pp.527–543.

Keskin, O., Ma, B. & Nussinov, R., 2005. Hot Regions in Protein–Protein Interactions: The Organization and Contribution of Structurally Conserved Hot Spot Residues. *Journal of Molecular Biology*, 345(5), pp.1281–1294.

Van Kessel, J.A.S. et al., 2012. Dynamics of Salmonella serotype shifts in an endemically infected dairy herd. *Foodborne Pathogens and Disease*, 9(4), pp.319–324.

Kieboom, J. & Abee, T., 2006. Arginine-Dependent Acid Resistance in Salmonella enterica Serovar Typhimurium. *Journal of Bacteriology*, 188(15), pp.5650–5653.

Kim, B.H. et al., 2001. Expression of cspH, Encoding the Cold Shock Protein in Salmonella enterica Serovar Typhimurium UK-1. *Journal of Bacteriology*, 183(19), pp.5580–5588.

Kim, B.H. et al., 2005. The formation of cyclopropane fatty acids in Salmonella enterica serovar Typhimurium. *Microbiology*, 151(1), pp.209–218.

Kim, J. & Dang, C.V., 2005. Multifaceted roles of glycolytic enzymes. *Trends in Biochemical Sciences*, 30(3), pp.142–150.

Kim, K. et al., 2014. Kinetic Behavior of Escherichia coli on Various Cheeses under Constant and Dynamic Temperature. *Asian-Australasian Journal of Animal Sciences*, 27(7), pp.1013–1018.

Kim, Y. et al., 2010. Escherichia coli toxin/antitoxin pair MqsR/MqsA regulate toxin CspD. *Environmental Microbiology*, 12(5), pp.1105–1121.

Kim, Y. & Wood, T.K., 2010. Toxins Hha and CspD and small RNA regulator Hfq are involved in persister cell formation through MqsR in Escherichia coli. *Biochemical and Biophysical Research Communications*, 391(1), pp.209–213.

Kingsley, R.A. et al., 2009. Epidemic multiple drug resistant Salmonella Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Research*, 19(12), pp.2279 –2287.

Klepp, L.I. et al., 2009. Identification of two proteins that interact with the Erp virulence factor from Mycobacterium tuberculosis by using the bacterial two-hybrid system. *BMC Molecular Biology*, 10(1), p.3.

Kobayashi, H. et al., 2005. Identification of factors involved in recovery of heat-injured Salmonella Enteritidis. *Journal of Food Protection*, 68(5), pp.932–941.

Kröger, C. et al., 2012. The transcriptional landscape and small RNAs of Salmonella enterica serovar Typhimurium. *Proceedings of the National Academy of Sciences*, 109(20) pp.1277-1286.

Kumar, M.D.S. & Gromiha, M.M., 2006. PINT: Protein–protein Interactions Thermodynamic Database. *Nucleic Acids Research*, 34(suppl 1), pp.D195–D198.

Kumpfmüller, J., Kabisch, J. & Schweder, T., 2013. An optimized technique for rapid genome modifications of Bacillus subtilis. *Journal of Microbiological Methods*, 95(3), pp.350–352.

Kurita, A. et al., 2003. Intracellular expression of the Salmonella plasmid virulence protein, SpvB, causes apoptotic cell death in eukaryotic cells. *Microbial Pathogenesis*, 35(1), pp.43–48.

Langklotz, S. & Narberhaus, F., 2011. The Escherichia coli replication inhibitor CspD is subject to growth-regulated degradation by the Lon protease. *Molecular Microbiology*, 80(5), pp.1313–1325.

Lapaque, N. et al., 2009. Salmonella regulates polyubiquitination and surface expression of MHC class II antigens. *Proceedings of the National Academy of Sciences*, 106(33), pp.14052 –14057.

Larkin, M.A. et al., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23(21), pp.2947–2948.

Lawson, B. et al., 2014. Epidemiological evidence that garden birds are a source of human salmonellosis in England and Wales. *PloS One*, 9(2), p.e88968.

Lecomte, X. et al., 2014. The naturally competent strain Streptococcus thermophilus LMD-9 as a new tool to anchor heterologous proteins on the cell surface. *Microbial Cell Factories*, 13(1), p.82.

Lee, D.H. & Goldberg, A.L., 1998. Proteasome Inhibitors Cause Induction of Heat Shock Proteins and Trehalose, Which Together Confer Thermotolerance inSaccharomyces cerevisiae. *Molecular and Cellular Biology*, 18(1), pp.30–38.

Lee, S.J. & Gralla, J.D., 2004. Osmo-Regulation of Bacterial Transcription via Poised RNA Polymerase. *Molecular Cell*, 14(2), pp.153–162.

Lehnik-Habrink, M. et al., 2013. DEAD-Box RNA Helicases in Bacillus subtilis Have Multiple Functions and Act Independently from Each Other. *Journal of Bacteriology*, 195(3), pp.534–544.

Lelivelt, M.J. & Kawula, T.H., 1995. Hsc66, an Hsp70 homolog in Escherichia coli, is induced by cold shock but not by heat shock. *Journal of Bacteriology*, 177(17), pp.4900–4907.

Lenz, G. & Ron, E.Z., 2014. Novel Interaction between the Major Bacterial Heat Shock Chaperone (GroESL) and an RNA Chaperone (CspC). *Journal of Molecular Biology*, 426(2), pp.460–466.

Lesnik, E.A. et al., 2001. Prediction of rho-independent transcriptional terminators in Escherichia coli. *Nucleic Acids Research*, 29(17), pp.3583–3594.

Li, H. et al., 2010. DmsD, a Tat system specific chaperone, interacts with other general chaperones and proteins involved in the molybdenum cofactor biosynthesis. *Biochimica Et Biophysica Acta*, 1804(6), pp.1301–1309.

Li, J. et al., 1995. Response Regulators CheB and CheY Exhibit Competitive Binding to the Kinase CheA. *Biochemistry*, 34(45), pp.14626–14636.

Lijnzaad, P. & Argos, P., 1997. Hydrophobic patches on protein subunit interfaces: Characteristics and prediction. *Proteins: Structure, Function, and Bioinformatics*, 28(3), pp.333–343.

Liu, J.Z. et al., 2012. Zinc sequestration by the neutrophil protein calprotectin enhances Salmonella growth in the inflamed gut. *Cell Host & Microbe*, 11(3), pp.227–239.

Liu, W.-T. et al., 2007. Role of the universal stress protein UspA of Salmonella in growth arrest, stress and virulence. *Microbial Pathogenesis*, 42(1), pp.2–10.

Lloyd, D.R., Carmichael, P.L. & Phillips, D.H., 1998. Comparison of the Formation of 8-Hydroxy-2‘-deoxyguanosine and Single- and Double-Strand Breaks in DNA Mediated by Fenton Reactions. *Chemical Research in Toxicology*, 11(5), pp.420–427.

Lopez, C.A. et al., 2012. Phage-Mediated Acquisition of a Type III Secreted Effector Protein Boosts Growth of Salmonella by Nitrate Respiration. *mBio*, 3(3), pp.e00143–12.

Lopez, M.M. & Makhatadze, G.I., 2000. Major cold shock proteins, CspA from Escherichia coli and CspB from Bacillus subtilis, interact differently with single-stranded DNA templates. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1479(1–2), pp.196–202.

Lopez, M.M., Yutani, K. & Makhatadze, G.I., 1999. Interactions of the Major Cold Shock Protein of Bacillus subtilis CspB with Single-stranded DNA Templates of Different Base Composition. *Journal of Biological Chemistry*, 274(47), pp.33601–33608.

Los, D.A., 2004. The effect of low-temperature-induced DNA supercoiling on the expression of the desaturase genes in synechocystis. *Cellular and Molecular Biology,* 50(5), pp.605–612.

Lucic, B. et al., 2011. A prominent β-hairpin structure in the winged-helix domain of RECQ1 is required for DNA unwinding and oligomer formation. *Nucleic Acids Research*, 39(5), pp.1703–1717.

Lu, H.J., Breidt, F. & Pérez-Díaz, I., 2013. Development of an effective treatment for A 5-log reduction of Escherichia coli in refrigerated pickle products. *Journal of Food Science*, 78(2), pp.M264–269.

Lyabin, D.N., Eliseeva, I.A. & Ovchinnikov, L.P., 2014. YB-1 protein: functions and regulation. *Wiley Interdisciplinary Reviews: RNA*, 5(1), pp.95–110.

Majowicz, S.E. et al., 2010. The Global Burden of Nontyphoidal Salmonella Gastroenteritis. *Clinical Infectious Diseases*, 50(6), pp.882–889.

Marcus, S.L. et al., 2000. Salmonella pathogenicity islands: big virulence in small packages. *Microbes and Infection*, 2(2), pp.145–156.

Marles-Wright, J. et al., 2008. Molecular Architecture of the "Stressosome," a Signal Integration and Transduction Hub. *Science*, 322(5898), pp.92–96.

Martin-Orozco, N. et al., 2006. Visualization of Vacuolar Acidification-induced Transcription of Genes of Pathogens inside Macrophages. *Molecular Biology of the Cell*, 17(1), pp.498–510.

Matsumoto, Y. et al., 2010. Structure of a Virulence Regulatory Factor CvfB Reveals a Novel Winged Helix RNA Binding Module. *Structure*, 18(4), pp.537–547.

Max, K.E.A. et al., 2007. Common mode of DNA binding to cold shock domains. *FEBS Journal*, 274(5), pp.1265–1279.

Mazengia, E. et al., 2014. Prevalence, concentrations, and antibiotic sensitivities of Salmonella serovars in poultry from retail establishments in Seattle, Washington. *Journal of Food Protection*, 77(6), pp.885–893.

McGibbon, L., 2013. *The RNA Interactome of Cold Shock Proteins, CspA and CspE, in Salmonella Typhimurium*. PhD Thesis. University of Edinburgh.

McMeechan, A. et al., 2007. Role of the alternative sigma factors σE and σS in survival of Salmonella enterica serovar Typhimurium during starvation, refrigeration and osmotic shock. *Microbiology*, 153(1), pp.263–269.

McMillan, K. et al., 2000. Allosteric inhibitors of inducible nitric oxide synthase dimerization discovered via combinatorial chemistry. *Proceedings of the National Academy of Sciences*, 97(4), pp.1506–1511.

Merlin, C., McAteer, S. & Masters, M., 2002. Tools for characterization of Escherichia coli genes of unknown function. *Journal of Bacteriology*, 184(16), pp.4573–4581.

Metz, B. et al., 2004. Identification of Formaldehyde-induced Modifications in Proteins: Reactions with Model Peptides. *Journal of Biological Chemistry*, 279(8), pp.6235–6243.

Meyer, A.S. & Baker, T.A., 2011. Proteolysis in the Escherichia coli heat shock response: a player at many levels. *Current Opinion in Microbiology*, 14(2), pp.194–199.

Michaux, C. et al., 2013. Cold-shock RNA-binding protein CspR is also exposed to the surface of Enterococcus faecalis. *Microbiology*, 159(Pt 10), pp.2153–2161.

Michaux, C. et al., 2012. CspR, a Cold Shock RNA-Binding Protein Involved in the Long-Term Survival and the Virulence of Enterococcus faecalis. *Journal of Bacteriology*, 194(24), pp.6900–6908.

Mihailovich, M. et al., 2010. Eukaryotic cold shock domain proteins: highly versatile regulators of gene expression. *BioEssays*, 32(2), pp.109–118.

Mizushima, T. et al., 1997. Increase in negative supercoiling of plasmid DNA in Escherichia coli exposed to cold shock. *Molecular Microbiology*, 23(2), pp.381–386.

Moll, I. et al., 2002. Effects of ribosomal proteins S1, S2 and the DeaD/CsdA DEAD-box helicase on translation of leaderless and canonical mRNAs in Escherichia coli. *Molecular Microbiology*, 44(5), pp.1387–1396.

Morales, P., Rouviere-Yaniv, J. & Dreyfus, M., 2002. The Histone-Like Protein HU Does Not Obstruct Movement of T7 RNA Polymerase in Escherichia coli Cells but Stimulates Its Activity. *Journal of Bacteriology*, 184(6), pp.1565–1570.

Moreira, I.S., Fernandes, P.A. & Ramos, M.J., 2007. Hot spots—A review of the protein–protein interface determinant amino-acid residues. *Proteins: Structure, Function, and Bioinformatics*, 68(4), pp.803–812.

Morgan, H.P. et al., 2009. Crystallization and X-ray structure of cold-shock protein E from *Salmonella typhimurium*. *Acta Crystallographica Section F Structural Biology and Crystallization Communications*, 65(12), pp.1240–1245.

Morita, M. et al., 2006. Characterization of lysine decarboxylase-negative strains of Salmonella enterica serovar Enteritidis disseminated in Japan. *FEMS Immunology & Medical Microbiology*, 46(3), pp.381–385.

Morita, M. et al., 1999. Heat-Induced Synthesis of ç32 inEscherichia coli: Structural and Functional Dissection ofrpoH mRNA Secondary Structure. *Journal of Bacteriology*, 181(2), pp.401–410.

Mortensen, L., Dandanell, G. & Hammer, K., 1989. Purification and characterization of the deoR repressor of Escherichia coli. *The EMBO Journal*, 8(1), pp.325–331.

Mosyak, L. et al., 2000. The bacterial cell-division protein ZipA and its interaction with an FtsZ fragment revealed by X-ray crystallography. *The EMBO journal*, 19(13), pp.3179–3191.

Naue, N. et al., 2013. The helicase-binding domain of Escherichia coli DnaG primase interacts with the highly conserved C-terminal region of single-stranded DNA-binding protein. *Nucleic Acids Research*, 41(8), pp.4507–4517.

Nonaka, G. et al., 2006. Regulon and promoter analysis of the E. coli heat-shock factor, σ32, reveals a multifaceted cellular response to heat stress. *Genes & Development*, 20(13), pp.1776–1789.

Nooren, I.M.A. & Thornton, J.M., 2003. Structural Characterisation and Functional Significance of Transient Protein–Protein Interactions. *Journal of Molecular Biology*, 325(5), pp.991–1018.

Nydam, S.D., Shah, D.H. & Call, D.R., 2014. Transcriptome analysis of Vibrio parahaemolyticus in type III secretion system 1 inducing conditions. *Frontiers in Cellular and Infection Microbiology*, 4(1).

Ono, S. et al., 2005. H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. *Biochemical Journal*, 391(2), p.203.

References

Oyofo, B.A. et al., 1989. Lipid Content of Salmonella Typhimurium. *The Journal of General and Applied Microbiology*, 35(5), pp.333–342.

Ozbabacan, S.E.A. et al., 2011. Transient protein-protein interactions. *Protein engineering, design & selection: PEDS*, 24(9), pp.635–648.

Pané-Farré, J., Lewis, R.J. & Stülke, J., 2005. The RsbRST stress module in bacteria: a signalling system that may interact with different output modules. *Journal of Molecular Microbiology and Biotechnology*, 9(2), pp.65–76.

Paramelle, D. et al., 2013. Chemical cross-linkers for protein structure studies by mass spectrometry. *Proteomics*, 13(3-4), pp.438–456.

Park, Y.M. et al., 2011. Nitrosative stress causes amino acid auxotrophy in hmp mutant Salmonella Typhimurium. *Microbiology and Immunology*, 55(10), pp.743–747.

Paschos, A. et al., 2011. An in vivo high-throughput screening approach targeting the type IV secretion system component VirB8 identified inhibitors of Brucella abortus 2308 proliferation. *Infection and Immunity*, 79(3), pp.1033–1043.

Paytubi, S. et al., 2014. A novel role for antibiotic resistance plasmids in facilitating Salmonella adaptation to non-host environments. *Environmental Microbiology*, 16(4), pp.950–962.

Peil, L., Virumäe, K. & Remme, J., 2008. Ribosome assembly in Escherichia coli strains lacking the RNA helicase DeaD/CsdA or DbpA. *FEBS Journal*, 275(15), pp.3772–3782.

Pelletier, H. & Kraut, J., 1992. Crystal structure of a complex between electron transfer partners, cytochrome c peroxidase and cytochrome c. *Science*, 258(5089), pp.1748–1755.

Phadtare, S. et al., 2006. Analysis of Escherichia coli Global Gene Expression Profiles in Response to Overexpression and Deletion of CspC and CspE. *Journal of Bacteriology*, 188(7), pp.2521–2527.

Phadtare, S., 2005. Nucleic acid melting by Escherichia coli CspE. *Nucleic Acids Research*, 33(17), pp.5583–5590.

Phadtare, S., 2001. The Nucleic Acid Melting Activity of Escherichia coli CspE Is Critical for Transcription Antitermination and Cold Acclimation of Cells. *Journal of Biological Chemistry*, 277(9), pp.7239–7245.

Phadtare, S. et al., 2002. Three Amino Acids in Escherichia coli CspE Surface-exposed Aromatic Patch Are Critical for Nucleic Acid Melting Activity Leading to Transcription Antitermination and Cold Acclimation of Cells. *Journal of Biological Chemistry*, 277(48), pp.46706 –46711.

Phadtare, S., Alsina, J. & Inouye, M., 1999. Cold-shock response and cold-shock proteins. *Current Opinion in Microbiology*, 2(2), pp.175–180.

Phadtare, S. & Inouye, M., 2004. Genome-Wide Transcriptional Analysis of the Cold Shock Response in Wild-Type and Cold-Sensitive, Quadruple-csp-Deletion Strains of Escherichia coli. *Journal of Bacteriology*, 186(20), pp.7007–7014.

Phadtare, S. & Inouye, M., 1999. Sequence-selective interactions with RNA by CspB, CspC and CspE, members of the CspA family of Escherichia coli. *Molecular Microbiology*, 33(5), pp.1004–1014.

Phadtare, S., Inouye, M. & Severinov, K., 2004. The Mechanism of Nucleic Acid Melting by a CspA Family Protein. *Journal of Molecular Biology*, 337(1), pp.147–155.

Phadtare, S. & Severinov, K., 2010. RNA remodeling and gene regulation by cold shock proteins. , 7(6), pp.788–795.

Piette, F., Struvay, C. & Feller, G., 2011. The protein folding challenge in psychrophiles: facts and current issues. *Environmental Microbiology*, 13(8), pp.1924–1933.

Płociński, P. et al., 2014. Identification of Protein Partners in Mycobacteria Using a Single-Step Affinity Purification Method. *PLoS ONE*, 9(3), p.e91380.

Prakash, J.S.S. et al., 2009. DNA supercoiling regulates the stress-inducible expression of genes in the cyanobacterium Synechocystis. *Molecular BioSystems*, 5(12), p.1904.

Prasch, S. et al., 2006. Interaction of the Intrinsically Unstructured Phage λ N Protein with Escherichia coli NusA. *Biochemistry*, 45(14), pp.4542–4549.

Prasch, S. et al., 2009. RNA-binding specificity of E. coli NusA. *Nucleic Acids Research*, 37(14), pp.4736–4742.

Prieto, A.I. et al., 2012. Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in Escherichia coli K12. *Nucleic Acids Research*, 40(8), pp.3524–3537.

Proshkin, S., Mironov, A. & Nudler, E., 2014. Riboswitches in regulation of Rho-dependent transcription termination. *Biochimica et Biophysica Acta*, 1839(10) pp.974-977

Prud'homme-Généreux, A. et al., 2004. Physical and functional interactions among RNase E, polynucleotide phosphorylase and the cold-shock protein, CsdA: evidence for a "cold shock degradosome. *Molecular Microbiology*, 54(5), pp.1409–1421.

Pruss, G.J. & Drlica, K., 1989. DNA supercoiling and prokaryotic transcription. *Cell*, 56(4), pp.521–523.

Pulido-Landínez, M. et al., 2013. Serotype and Antimicrobial Resistance Patterns of Salmonella Isolates from Commercial Birds and Poultry Environment in Mississippi. *Avian Diseases*, 58(1), pp.64–70.

Purves, J. et al., 2010. Comparison of the regulation, metabolic functions, and roles in virulence of the glyceraldehyde-3-phosphate dehydrogenase homologues gapA and gapB in Staphylococcus aureus. *Infection and immunity*, 78(12), pp.5223–5232.

Quinn, H.J., Cameron, A.D.S. & Dorman, C.J., 2014. Bacterial Regulon Evolution: Distinct Responses and Roles for the Identical OmpR Proteins of Salmonella Typhimurium and Escherichia coli in the Acid Stress Response. *PLoS Genet*, 10(3), p.e1004215.

Rajashekar, R. et al., 2008. Dynamic Remodeling of the Endosomal System During Formation of Salmonella-Induced Filaments by Intracellular Salmonella enterica. *Traffic*, 9(12), pp.2100–2116.

Rappsilber, J., Ishihama, Y. & Mann, M., 2003. Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption/Ionization, Nanoelectrospray, and LC/MS Sample Pretreatment in Proteomics. *Analytical Chemistry*, 75(3), pp.663–670.

Rasouly, A. & Ron, E.Z., 2009. Interplay between the heat shock response and translation in Escherichia coli. *Research in Microbiology*, 160(4), pp.288–296.

Ray, W.K. & Larson, T.J., 2004. Application of AgaR repressor and dominant repressor variants for verification of a gene cluster involved in N-acetylgalactosamine metabolism in Escherichia coli K-12. *Molecular Microbiology*, 51(3), pp.813–826.

Reimann, S.A. & Wolfe, A.J., 2011. Exposure to Glycolytic Carbon Sources Reveals a Novel Layer of Regulation for the MalT Regulon. *International Journal of Microbiology*, 2011, p.e107023.

Resch, A. et al., 2010a. Requirement of the CsdA DEAD-box helicase for low temperature riboregulation of rpoS mRNA. *RNA biology*, 7(6), pp.796–802.

Rescigno, M. et al., 2001. Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nature Immunology*, 2(4), pp.361–367.

Richardson, A.R. et al., 2011. Multiple targets of nitric oxide in the tricarboxylic acid cycle of Salmonella enterica serovar typhimurium. *Cell Host & Microbe*, 10(1), pp.33–43.

Rinnenthal, J. et al., 2010. Direct observation of the temperature-induced melting process of the Salmonella fourU RNA thermometer at base-pair resolution. *Nucleic Acids Research*, 38(11), pp.3834–3847.

Rivera-Chávez, F. et al., 2013. Salmonella Uses Energy Taxis to Benefit from Intestinal Inflammation. *PLoS Pathog*, 9(4), p.e1003267.

Rochat, T. et al., 2012. Proteomic analysis of spontaneous mutants of Lactococcus lactis: Involvement of GAPDH and arginine deiminase pathway in H_2O_2 resistance. *Proteomics*, 12(11), pp.1792–1805.

Rychlik, I. et al., 2001. Low-Molecular-Weight Plasmid of Salmonella enterica Serovar Enteritidis Codes for Retron Reverse Transcriptase and Influences Phage Resistance. *Journal of Bacteriology*, 183(9), pp.2852–2858.

Sambrook, J., Fritsch, E.F. & Maniatis, T., 1989. *Molecular Cloning: A Laboratory Manual* 2nd ed., Cold Spring Harbor Laboratory Press.

Santos, A.M. et al., 2012. FERM domain interaction with myosin negatively regulates FAK in cardiomyocyte hypertrophy. *Nature Chemical Biology*, 8(1), pp.102–110.

Schärer, K., Stephan, R. & Tasara, T., 2013. Cold shock proteins contribute to the regulation of listeriolysin O production in Listeria monocytogenes. *Foodborne Pathogens and Disease*, 10(12), pp.1023–1029.

Scheu, P.D. et al., 2012. CitA/CitB Two-Component System Regulating Citrate Fermentation in Escherichia coli and Its Relation to the DcuS/DcuR System In Vivo. *Journal of Bacteriology*, 194(3), pp.636–645.

Schindelin, H. et al., 1994. Crystal structure of CspA, the major cold shock protein of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 91(11), pp.5119–5123.

Schindelin, H., Marahiel, M.A. & Heinemann, U., 1993. Universal nucleic acid-binding domain revealed by crystal structure of the B. subtilis major cold-shock protein. *Nature*, 364(6433), pp.164–168.

Schmeing, T.M. et al., 2009. The Crystal Structure of the Ribosome Bound to EF-Tu and Aminoacyl-tRNA. *Science*, 326(5953), pp.688–694.

Schmid, B. et al., 2009. Role of Cold Shock Proteins in Growth of Listeria monocytogenes under Cold and Osmotic Stress Conditions. *Applied and Environmental Microbiology*, 75(6), pp.1621–1627.

Schröder, K. et al., 1995. Mutational analysis of the putative nucleic acid-binding surface of the cold-shock domain, CspB, revealed an essential role of aromatic and basic residues in binding of single-stranded DNA containing the Y-box motif. *Molecular Microbiology*, 16(4), pp.699–708.

Schultz, M., 2008. Theobald Smith. *Emerging Infectious Diseases*, 14(12), pp.1940–1942.

Schweizer, H., Boos, W. & Larson, T.J., 1985. Repressor for the sn-glycerol-3-phosphate regulon of Escherichia coli K-12: cloning of the glpR gene and identification of its product. *Journal of Bacteriology*, 161(2), pp.563–566.

Seco, J., Luque, F.J. & Barril, X., 2009. Binding site detection and druggability index from first principles. *Journal of Medicinal Chemistry*, 52(8), pp.2363–2371.

Semenkov, Y.P. et al., 1982. Kinetic aspects of tetracycline action on the acceptor (A) site of Escherichia coli ribosomes. *FEBS Letters*, 144(1), pp.125–129.

Sharma, U.K. & Chatterji, D., 2010. Transcriptional switching in Escherichia coli during stress and starvation by modulation of σ70 activity. *FEMS Microbiology Reviews*, 34(5), pp.646–657.

Shenhar, Y. et al., 2009. Adaptation of Escherichi coli to elevated temperatures involves a change in stability of heat shock gene transcripts. *Environmental Microbiology*, 11(12), pp.2989–2997.

Shenhar, Y., Biran, D. & Ron, E.Z., 2012. Resistance to environmental stress requires the RNA chaperones CspC and CspE. *Environmental Microbiology Reports*, 4(5), pp.532–539.

Shereda, R.D. et al., 2009. Identification of the SSB Binding Site on E. coli RecQ Reveals a Conserved Surface for Binding SSB's C Terminus. *Journal of Molecular Biology*, 386(3), pp.612–625.

Shereda, R.D. et al., 2008. SSB as an Organizer/Mobilizer of Genome Maintenance Complexes. *Critical Reviews in Biochemistry and Molecular Biology*, 43(5), pp.289–318.

Shevchenko, A. et al., 1996. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Analytical Chemistry*, 68(5), pp.850–858.

Shivaji, S. & Prakash, J.S.S., 2010. How do bacteria sense and respond to low temperature? *Archives of Microbiology*, 192(2), pp.85–95.

Short, S.A. & Singer, J.T., 1984. Studies on deo operon regulation in Escherichia coli: cloning and expression of the deoR structural gene. *Gene*, 31(1-3), pp.205–211.

Shrimpton, S.E., 2011. *Antibiotic persistence in Salmonella enterica serovar Typhimurium : involvement of the CspA paralogues*. PhD Thesis, University of Edinburgh.

Da Silva, C.A.P.T. et al., 2010. SpdR, a response regulator required for stationary-phase induction of Caulobacter crescentus cspD. *Journal of Bacteriology*, 192(22), pp.5991–6000.

References

Singer, M. et al., 1989. A collection of strains containing genetically linked alternating antibiotic resistance elements for genetic mapping of Escherichia coli. *Microbiological Reviews*, 53(1), pp.1–24.

Sleator, R.D. & Hill, C., 2002. Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiology Reviews*, 26(1), pp.49–71.

Smadi, H. et al., 2012. Growth and inactivation of Salmonella at low refrigerated storage temperatures and thermal inactivation on raw chicken meat and laboratory media: Mixed effect meta-analysis. *Journal of Epidemiology and Global Health*, 2(4), pp.165–179.

Smit, M.H. de & Duin, J. van, 1990. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proceedings of the National Academy of Sciences*, 87(19), pp.7668–7672.

Solovyev, V. & Salamov, A., 2011. Automatic Annotation of Microbial Genomes and Metagenomic Sequences. In *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies*. Nova Science.

Song, M. et al., 2013. Low-molecular-weight thiol-dependent antioxidant and antinitrosative defences in Salmonella pathogenesis. *Molecular Microbiology*, 87(3), pp.609–622.

Song, M. et al., 2010. ppGpp-mediated stationary phase induction of the genes encoded by horizontally acquired pathogenicity islands and cob/pdu locus in Salmonella enterica serovar Typhimurium. *Journal of Microbiology*, 48(1), pp.89–95.

Soo Lee, I. et al., 1995. The stationary-phase sigma factor σS (RpoS) is required for a sustained acid tolerance response in virulent Salmonella typhimurium. *Molecular Microbiology*, 17(1), pp.155–167.

Spitale, R.C., Tsai, M.-C. & Chang, H.Y., 2011. RNA templating the epigenome: long noncoding RNAs as molecular scaffolds. *Epigenetics,* 6(5), pp.539–543.

Stecher, B. et al., 2012. Gut inflammation can boost horizontal gene transfer between pathogenic and commensal Enterobacteriaceae. *Proceedings of the National Academy of Sciences*, 109(4), pp.1269–1274.

Stirling, D.A. et al., 1989. Molecular characterization of the proU loci of Salmonella typhimurium and Escherichia coli encoding osmoregulated glycine betaine transport systems. *Molecular Microbiology*, 3(8), pp.1025–1038.

Sutherland, B.W., Toews, J. & Kast, J., 2008. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein–protein interactions. *Journal of Mass Spectrometry*, 43(6), pp.699–715.

Tamang, M. et al., 2007. Emergence of multidrug-resistant Salmonella enterica serovar Typhi associated with a class 1 integron carrying the dfrA7 gene cassette in Nepal. *International Journal of Antimicrobial Agents*, 30(4), pp.330–335.

Taylor, R.G., Walker, D.C. & McInnes, R.R., 1993. E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Research*, 21(7), pp.1677–1678.

Thanikkal, E.J., Mangu, J.C.K. & Francis, M.S., 2012. Interactions of the CpxA sensor kinase and cognate CpxR response regulator from Yersinia pseudotuberculosis. *BMC Research Notes*, 5, p.536.

Tindall, B.J. et al., 2005. Nomenclature and taxonomy of the genus Salmonella. *International Journal of Systematic and Evolutionary Microbiology*, 55(1), pp.521 –524.

Toews, J. et al., 2008. Mass spectrometric identification of formaldehyde-induced peptide modifications under in vivo protein cross-linking conditions. *Analytica Chimica Acta*, 618(2), pp.168–183.

Tomoyasu, T. et al., 2003. A new heat shock gene, AgsA, which encodes a small chaperone involved in suppressing protein aggregation in Salmonella enterica serovar typhimurium. *Journal of Bacteriology*, 185(21), pp.6331–6339.

Tompa, P., 2002. Intrinsically unstructured proteins. *Trends in Biochemical Sciences*, 27(10), pp.527–533.

Tsai, M.-C. et al., 2010. Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science*, 329(5992), pp.689–693.

Tsao, D.H.H. et al., 2006. Discovery of novel inhibitors of the ZipA/FtsZ complex by NMR fragment screening coupled with structure-based design. *Bioorganic & Medicinal Chemistry*, 14(23), pp.7953–7961.

Uchiya, K. et al., 1999. A Salmonella virulence protein that inhibits cellular trafficking. *The EMBO Journal*, 18(14), pp.3924–3933.

Uesugi, A.R., Danyluk, M.D. & Harris, L.J., 2006. Survival of Salmonella enteritidis phage type 30 on inoculated almonds stored at -20, 4, 23, and 35°C. *Journal of Food Protection*, 69(8), pp.1851–1857.

Uppal, S. et al., 2011. Cyclic AMP Receptor Protein Regulates cspE, an Early Cold-Inducible Gene, in Escherichia coli. *Journal of Bacteriology*, 193(22), pp.6142–6151.

Uppal, S., Rao Akkipeddi, V.S.N. & Jawali, N., 2008. Posttranscriptional regulation of cspE in Escherichia coli: involvement of the short 5'-untranslated region. *FEMS Microbiology Letters*, 279(1), pp.83–91.

Uppal, S., Shetty, D.M. & Jawali, N., 2014. Cyclic AMP receptor protein regulates cspD, a bacterial toxin gene, in Escherichia coli. *Journal of Bacteriology*, 196(8), pp.1569–1577.

Vakulskas, C.A. et al., 2014. Global effects of the DEAD-box RNA helicase DeaD (CsdA) on gene expression over a broad range of temperatures. *Molecular Microbiology*, 92(5), pp.945–958.

Vazquez-Torres, A. et al., 2000. Antimicrobial Actions of the Nadph Phagocyte Oxidase and Inducible Nitric Oxide Synthase in Experimental Salmonellosis. I. Effects on Microbial Killing by Activated Peritoneal Macrophages in Vitro. *The Journal of Experimental Medicine*, 192(2), pp.227–236.

Vazquez-Torres, A., 2000. Salmonella Pathogenicity Island 2-Dependent Evasion of the Phagocyte NADPH Oxidase. *Science*, 287(5458), pp.1655–1658.

References

Veinger, L. et al., 1998. The Small Heat-shock Protein IbpB from Escherichia coli Stabilizes Stress-denatured Proteins for Subsequent Refolding by a Multichaperone Network. *Journal of Biological Chemistry*, 273(18), pp.11032–11037.

Velden, A.W.M. van der et al., 2000. Salmonella Pathogenicity Island 1-Independent Induction of Apoptosis in Infected Macrophages bySalmonella enterica Serotype Typhimurium. *Infection and Immunity*, 68(10), pp.5702–5709.

Vigonsky, E., Ovcharenko, E. & Lewinson, O., 2013. Two molybdate/tungstate ABC transporters that interact very differently with their substrate binding proteins. *Proceedings of the National Academy of Sciences*, 110(14) pp.5440-5445.

Vojtova, J., Kamanova, J. & Sebo, P., 2006. Bordetella adenylate cyclase toxin: a swift saboteur of host defense. *Current Opinion in Microbiology*, 9(1), pp.69–75.

Vorachek-Warren, M.K. et al., 2002. An Escherichia coli Mutant Lacking the Cold Shock-induced Palmitoleoyltransferase of Lipid A Biosynthesis. *Journal of Biological Chemistry*, 277(16), pp.14186 –14193.

Wang, J.D., Sanders, G.M. & Grossman, A.D., 2007. Nutritional Control of Elongation of DNA Replication by (p)ppGpp. *Cell*, 128(5), pp.865–875.

Wang, Y., San, K.-Y. & Bennett, G.N., 2013. Improvement of NADPH bioavailability in Escherichia coli by replacing NAD(+)-dependent glyceraldehyde-3-phosphate dehydrogenase GapA with NADP (+)-dependent GapB from Bacillus subtilis and addition of NAD kinase. *Journal of Industrial Microbiology & Biotechnology*, 40(12), pp.1449–1460.

Wang, Z., Wang, S. & Wu, Q., 2014. Cold shock protein A plays an important role in the stress adaptation and virulence of Brucella melitensis. *FEMS microbiology letters*, 354(1), pp.27–36.

Waterman, S.R. & Small, P.L.C., 1998. Acid-Sensitive Enteric Pathogens Are Protected from Killing under Extremely Acidic Conditions of pH 2.5 when They Are Inoculated onto Certain Solid Food Sources. *Applied and Environmental Microbiology*, 64(10), pp.3882–3886.

Weber, M.H.W., Volkov, A.V., et al., 2001. Localization of Cold Shock Proteins to Cytosolic Spaces Surrounding Nucleoids in Bacillus subtilis Depends on Active Transcription. *Journal of Bacteriology*, 183(21), pp.6435–6443.

Weber, M.H.W., Beckering, C.L. & Marahiel, M.A., 2001. Complementation of Cold Shock Proteins by Translation Initiation Factor IF1 In Vivo. *Journal of Bacteriology*, 183(24), pp.7381–7386.

Weinstein-Fischer, D., Elgrably-weiss, M. & Altuvia, S., 2000. Escherichia coli response to hydrogen peroxide: a role for DNA supercoiling, Topoisomerase I and Fis. *Molecular Microbiology*, 35(6), pp.1413–1420.

Welker, C. et al., 1999. Cloning, overexpression, purification, and physicochemical characterization of a cold shock protein homolog from the hyperthermophilic bacterium Thermotoga maritima. *Protein Science*, 8(2), pp.394–403.

Wells, J.A. & McClendon, C.L., 2007. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature*, 450(7172), pp.1001–1009.

White-Ziegler, C.A. et al., 2008. Low temperature (23 °C) increases expression of biofilm-, cold-shock- and RpoS-dependent genes in Escherichia coli K-12. *Microbiology*, 154(1), pp.148–166.

Winfield, M.D. & Groisman, E.A., 2004. Phenotypic differences between Salmonella and Escherichia coli resulting from the disparate regulation of homologous genes. *Proceedings of the National Academy of Sciences*, 101(49), pp.17162–17167.

Winfield, M.D. & Groisman, E.A., 2003. Role of Nonhost Environments in the Lifestyles of Salmonella and Escherichia coli. *Applied and Environmental Microbiology*, 69(7), pp.3687–3694.

Winter, S.E. et al., 2010. Gut inflammation provides a respiratory electron acceptor for Salmonella. *Nature*, 467(7314), pp.426–429.

Wollenweber, H.-W. et al., 1983. Fatty Acid in Lipopolysaccharides of Salmonella Species Grown at Low Temperature. *European Journal of Biochemistry*, 130(1), pp.167–171.

Woodall, K., 2010. *The Cold Shock Response of Salmonella enterica serovar typhimurium; the Involvement of the CspA Paralogues*. PhD Thesis. University of Edinburgh.

Xia, B. et al., 2003. The role of RbfA in 16S rRNA processing and cell growth at low temperature in Escherichia coli. *Journal of Molecular Biology*, 332(3), pp.575–584.

Xia, B., Ke, H. & Inouye, M., 2001. Acquirement of cold sensitivity by quadruple deletion of the cspA family and its suppression by PNPase S1 domain in Escherichia coli. *Molecular Microbiology*, 40(1), pp.179–188.

Xu, D., Lin, S.L. & Nussinov, R., 1997. Protein binding versus protein folding: the role of hydrophilic bridges in protein associations. *Journal of Molecular Biology*, 265(1), pp.68–84.

Xu, J. et al., 2003. A Genomic View of the Human-Bacteroides thetaiotaomicron Symbiosis. *Science*, 299(5615), pp.2074–2076.

Yachdav, G. et al., 2014. PredictProtein--an open resource for online prediction of protein structural and functional features. *Nucleic Acids Research*, 42(Web Server issue), pp.W337–343.

Yamada, M. & Saier, M.H., 1988. Positive and negative regulators for glucitol (gut) operon expression in Escherichia coli. *Journal of Molecular Biology*, 203(3), pp.569–583.

Yamanaka, K. et al., 1994. Cloning, sequencing, and characterization of multicopy suppressors of a mukB mutation in Escherichia coli. *Molecular Microbiology*, 13(2), pp.301–312.

Yamanaka, K. & Inouye, M., 1997. Growth-phase-dependent expression of cspD, encoding a member of the CspA family in Escherichia coli. *Journal of Bacteriology*, 179(16), pp.5126–5130.

Yamanaka, K. & Inouye, M., 2001a. Induction of CspA, an E. coli major cold-shock protein, upon nutritional upshift at 37°C. *Genes to Cells*, 6(4), pp.279–290.

Yamanaka, K. & Inouye, M., 2001b. Selective mRNA Degradation by Polynucleotide Phosphorylase in Cold Shock Adaptation in Escherichia coli. *Journal of Bacteriology*, 183(9), pp.2808–2816.

References

Yang, B. et al., 2014. Counts, serotypes, and antimicrobial resistance of Salmonella isolates on retail raw poultry in the People's Republic of China. *Journal of Food Protection*, 77(6), pp.894–902.

Yoshizawa, S. et al., 2005. Structural basis for mRNA recognition by elongation factor SelB. *Nature Structural & Molecular Biology*, 12(2), pp.198–203.

Yuk, H.G. & Schneider, K.R., 2006. Adaptation of Salmonella spp. in juice stored under refrigerated and room temperature enhances acid resistance to simulated gastric fluid. *Food Microbiology*, 23(7), pp.694–700.

Zeeb, M. et al., 2006. Recognition of T-rich single-stranded DNA by the cold shock protein Bs-CspB in solution. *Nucleic Acids Research*, 34(16), pp.4561–4571.

Zeeb, M. & Balbach, J., 2003. Single-stranded DNA binding of the cold-shock protein CspB from Bacillus subtilis: NMR mapping and mutational characterization. *Protein Science,* 12(1), pp.112–123.

Zeng, G., Ye, S. & Larson, T.J., 1996. Repressor for the sn-glycerol 3-phosphate regulon of Escherichia coli K-12: primary structure and identification of the DNA-binding domain. *Journal of Bacteriology*, 178(24), pp.7080–7089.

Zhang, L., Li, W. & He, Z.-G., 2013. DarR, a TetR-like transcriptional factor, is a cyclic di-AMP-responsive repressor in Mycobacterium smegmatis. *The Journal of Biological Chemistry*, 288(5), pp.3085–3096.

Zhang, X.Y. et al., 2013. Dissection of the TssB-TssC Interface during Type VI Secretion Sheath Complex Formation. *PLoS ONE*, 8(11).

Zhou, D., Mooseker, M.S. & Galán, J.E., 1999. Role of the S. typhimurium actin-binding protein SipA in bacterial internalization. *Science*, 283(5410), pp.2092–2095.

Zwiefka, A., Kohn, H. & Widger, W.R., 1993. Transcription termination factor rho: The site of bicyclomycin inhibition in Escherichia coli. *Biochemistry*, 32(14), pp.3564–3570.

Appendices

Appendix A: Mass Spectrometry

Tables show gene and protein name for each protein detected using LTQ-Orbitrap in each of three experimental (cross-linked) repeats, ordered by abundance (peptides detected). Amount of the same protein detected in the control (without cross-linker) also shown. Those detected only in the experimental samples indicated bold.

Repeat 1

Protein [description]	Peptides detected	Peptides in control
[gene=groEL] [protein=GroEL protein]	56	27
[gene=rpoC] [protein=DNA-directed RNA polymerase, beta~-subunit]	46	92
[gene=tufA] [protein=elongation factor Tu]	44	45
[gene=rpoB] [protein=DNA-directed RNA polymerase, beta-subunit]	37	78
[gene=fusA] [protein=elongation factor G] [protein_id=CBW19508.1]	32	11
[gene=clpB] [protein=ClpB protein (heat shock protein f84.1)]	32	22
[gene=rplB] [protein=50S ribosomal subunit protein L2]	31	38
[gene=aceE] [protein=pyruvate dehydrogenase E1 component]	25	51
[gene=dnaK] [protein=Chaperone protein dnaK]	22	22
[gene=accC] [protein=biotin carboxylase]	21	21
[gene=rpsA] [protein=30S ribosomal protein S1]	21	16
[gene=pta] [protein=phosphate acetyltransferase]	20	19
[gene=lon] [protein=Lon protease]	20	55
[gene=lpdA] [protein=dihydrolipoamide dehydrogenase]	19	17
[gene=tig] [protein=trigger factor]	18	4
[gene=atpA] [protein=ATP synthase alpha subunit]	18	19
[gene=aceF] [protein=pyruvate dehydrogenase]	18	0
[gene=valS] [protein=valyl-tRNA synthetase]	18	0
[gene=rplN] [protein=50S ribosomal subunit protein L14]	17	17
[gene=alaS] [protein=Alanyl-tRNA synthetase]	16	15
[gene=atpD] [protein=ATP synthase beta subunit]	16	21
[gene=ptsI] [protein=phosphoenolpyruvate-protein phosphotransferase]	16	14
[gene=fabI] [protein=enoyl-[acyl-carrier-protein] reductase (NADH)]	16	14
[gene=prs] [protein=Ribose-phosphate pyrophosphokinase]	15	17
[gene=rpsB] [protein=30S ribosomal protein S2]	15	18
[gene=rho] [protein=transcription termination factor]	15	10
[gene=rpsD] [protein=30S ribosomal subunit protein S4]	15	16
[gene=glpD] [protein=aerobic glycerol-3-phosphate dehydrogenase]	15	8
[gene=glpK] [protein=glycerol kinase]	14	22
[gene=fabB] [protein=3-oxoacyl-[acyl-carrier-protein] synthase I]	14	0
[gene=pflB] [protein=formate acetyltransferase 1]	14	9
[gene=aspA] [protein=aspartate ammonia-lyase]	13	3

[gene=rpsE] [protein=30S ribosomal subunit protein S5]	12	13
[gene=htpG] [protein=heat shock protein HtpG]	12	4
[gene=sucA] [protein=2-oxoglutarate dehydrogenase E1 component]	12	36
[gene=iscS] [protein=Cysteine desulfurase]	12	9
[gene=adh] [protein=alcohol dehydrogenase]	11	22
[gene=deoD] [protein=purine nucleoside phosphorylase]	11	8
[gene=pez] [protein=DNA-directed RNA polymerase alpha chain]	11	14
[gene=ackA] [protein=acetate kinase]	10	8
[gene=SL1344_1723] [protein=hypothetical glutamate dehydrogenase]	10	9
[gene=asnCa] [protein=asparaginyl-tRNA synthetase]	10	7
[gene=glyS] [protein=glycine-tRNA synthetase, beta subunit]	10	11
[gene=rpsC] [protein=30S ribosomal subunit protein S3]	10	9
[gene=hslU] [protein=heat shock protein]	10	8
[gene=typA] [protein=GTP-binding protein]	10	15
[gene=leuS] [protein=leucyl-tRNA synthetase]	10	0
[gene=fabG] [protein=3-oxoacyl-[acyl-carrier protein] reductase]	9	8
[gene=deaD] [protein=ATP-dependent RNA helicase (dead-box protein)]	9	14
[gene=tyrS] [protein=tyrosyl-tRNA synthetase]	9	3
[gene=hemL] [protein=Glutamate-1-semialdehyde 2,1-aminomutase]	9	8
[gene=rplC] [protein=50S ribosomal subunit protein L3]	9	12
[gene=infB] [protein=protein chain initiation factor 2]	9	14
[gene=rplQ] [protein=50S ribosomal subunit protein L17]	9	13
[gene=pnp] [protein=polynucleotide phosphorylase]	9	0
[gene=metK] [protein=S-adenosylmethionine synthetase]	8	7
[gene=groES] [protein=GroES protein]	8	0
[gene=pheT] [protein=Phenylalanyl-tRNA synthetase beta chain]	8	8
[gene=rfaE] [protein=ADP-heptose synthase]	8	9
[gene=eno] [protein=Enolase]	8	2
[gene=pykA] [protein=pyruvate kinase A]	8	4
[gene=ychF] [protein=hypothetical ATP/GTP-binding protein]	8	12
[gene=gnd] [protein=6-phosphogluconate dehydrogenase, decarboxylating]	8	0
[gene=talB] [protein=transaldolase B]	8	0
[gene=nrdA] [protein=ribonucleoside-diphosphate reductase 1 alpha chain]	8	15
[gene=tsf] [protein=Elongation factor Ts]	8	0
[gene=sdaA] [protein=L-serine deaminase 1]	8	7
[gene=mreB] [protein=rod shape-determining protein]	8	8
[gene=rplV] [protein=50s ribosomal protein l22]	8	10
[gene=lysS] [protein=Lysyl-tRNA synthetase]	8	0
[gene=icdA] [protein=isocitrate dehydrogenase]	8	0
[gene=gyrB] [protein=DNA gyrase subunit B]	8	18
[gene=pykF] [protein=pyruvate kinase]	7	0
[gene=sucC] [protein=succinyl-CoA synthetase beta chain]	7	6
[gene=rplF] [protein=50S ribosomal subunit protein L6]	7	3
[gene=udp] [protein=uridine phosphorylase]	7	0

[gene=rpsM] [protein=30S ribosomal subunit protein S13]	7	9
[gene=gcpE] [protein=4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase]	7	8
[gene=pyrG] [protein=CTP synthase] [protein_id=CBW19031.1]	7	10
[gene=mdh] [protein=malate dehydrogenase]	7	4
[gene=yjjK] [protein=conserved hypothetical ABC transporter]	7	8
[gene=gapA] [protein=glyceraldehyde 3-phosphate dehydrogenase A]	7	0
[gene=dnaN] [protein=DNA polymerase III beta-subunit]	7	0
[gene=rpsL] [protein=30S ribosomal subunit protein S12]	7	7
[gene=rplE] [protein=50S ribosomal subunit protein L5]	7	16
[gene=thrS] [protein=Threonyl-tRNA synthetase]	7	6
[gene=yciL] [protein=hypothetical pseudouridine synthase]	7	6
[gene=pfkA] [protein=6-phosphofructokinase]	6	6
[gene=rplK] [protein=50S ribosomal subunit protein L11]	6	4
[gene=rfaD] [protein=ADP-L-Glycero-D-mannoheptose-6-epimerase]	6	3
[gene=minD] [protein=septum site determining protein]	6	9
[gene=yhgF] [protein=hypothetical transcription accessory protein]	6	8
[gene=rplO] [protein=50S ribosomal subunit protein L15]	6	10
[gene=argS] [protein=Arginyl-tRNA synthetase]	6	0
[gene=fabF] [protein=3-oxoacyl-[acyl-carrier-protein] synthase II]	6	4
[gene=rplJ] [protein=50S ribosomal subunit protein L10]	6	5
[gene=clpX] [protein=ATP-dependent clp protease ATP-binding subunit ClpX]	6	6
[gene=lctD] [protein=hypothetical L-lactate dehydrogenase]	6	13
[gene=rpsJ] [protein=30S ribosomal subunit protein S10]	6	2
[gene=rpl19] [protein=50S ribosomal protein L19]	6	4
[gene=deoB] [protein=phosphopentomutase]	6	0
[gene=glnA] [protein=glutamine synthetase]	6	0
[gene=aspC] [protein=aspartate aminotransferase]	6	0
[gene=tsaA] [protein=alkyl hydroperoxide reductase]	6	0
[gene=tktA] [protein=transketolase]	6	0
[gene=kdsA] [protein=2-dehydro-3-deoxyphosphooctonate aldolase]	6	6
[gene=rpmB] [protein=50S ribosomal subunit protein L28]	6	8
[gene=purA] [protein=adenylosuccinate synthetase]	6	0
[gene=secB] [protein=protein-export protein SecB]	5	0
[gene=rplI] [protein=50s ribosomal subunit protein L9]	5	3
[gene=tpx] [protein=Probable thiol peroxidase]	5	3
[gene=glk] [protein=Glucokinase]	5	2
[gene=rpsK] [protein=30S ribosomal subunit protein S11]	5	8
[gene=murA] [protein=UDP-N-acetylglucosamine 1-carboxyvinyltransferase]	5	3
[gene=serS] [protein=seryl-tRNA synthetase]	5	3
[gene=fumA] [protein=Fumarate hydratase class I, aerobic]	5	3
[gene=glyQ] [protein=glycine-tRNA synthetase, alpha subunit]	5	4
[gene=udhA] [protein=possible pyridine nucleotide-disulphide oxidoreductase]	5	4
[gene=recA] [protein=RecA protein]	5	13

[gene=sdaB] [protein=L-serine dehydratase 2 (L-serine deaminase 2)]	5	6
[gene=sucD] [protein=succinyl-CoA synthetase alpha chain]	5	0
[gene=ahpC] [protein=alkyl hydroperoxide reductase c22 protein]	5	0
[gene=rfbK] [protein=phosphomannomutase]	5	0
[gene=rfbH] [protein=dehydratase]	4	5
[gene=dapD] [protein=tetrahydropyridine succinyltransferase]	4	0
[gene=ucpA] [protein=hypothetical oxidoreductase]	4	4
[gene=fabH] [protein=3-oxoacyl-[acyl-carrier-protein] synthase III]	4	2
[gene=trxB] [protein=thioredoxin reductase]	4	0
[gene=fba] [protein=fructose 1,6-bisphosphate aldolase]	4	0
[gene=pyrH] [protein=Uridylate kinase]	4	0
[gene=phoL] [protein=PhoH-like ATP-binding protein]	4	5
[gene=rnb] [protein=Exoribonuclease II]	4	0
[gene=nuoF] [protein=NADH dehydrogenase I chain F]	4	3
[gene=glmS] [protein=glucosamine--fructose-6-phosphate aminotransferase]	4	4
[gene=rpsH] [protein=30S ribosomal subunit protein S8]	4	7
[gene=cspE] [protein=cold shock-like protein cspE]	4	5
[gene=nuoG] [protein=NADH dehydrogenase I chain G]	4	0
[gene=gltA] [protein=citrate synthase]	4	0
[gene=pheS] [protein=phenylalanyl-tRNA synthetase alpha chain]	4	0
[gene=pmbA] [protein=PmbA protein]	4	5
[gene=phoP] [protein=transcriptional regulatory protein PhoP]	4	0
[gene=crp] [protein=cyclic AMP receptor protein]	4	8
[gene=ahpF] [protein=alkyl hydroperoxide reductase F52A protein (subunit F)]	4	5
[gene=ydhD] [protein=conserved hypothetical protein]	4	0
[gene=hsdM] [protein=type I restriction enzyme]	4	8
[gene=metG] [protein=Methionyl-tRNA synthetase]	4	0
[gene=guaC] [protein=GMP reductase]	4	0
[gene=glmU] [protein=UDP-N-acetylglucosamine pyrophosphorylase]	3	3
[gene=rplA] [protein=50S ribosomal subunit protein L1]	3	5
[gene=deoC] [protein=deoxyribose-phosphate aldolase]	3	0
[gene=pepQ] [protein=proline dipeptidase]	3	0
[gene=yciK] [protein=hypothetical oxidoreductase]	3	3
[gene=rpsG] [protein=30S ribosomal subunit protein S7]	3	4
[gene=pgk] [protein=phosphoglycerate kinase]	3	3
[gene=proS] [protein=prolyl-tRNA synthetase]	3	5
[gene=yfiF] [protein=hypothetical RNA methyltransferase]	3	3
[gene=glmM] [protein=PGM/PMM-family protein]	3	0
[gene=aroB] [protein=3-dehydroquinate synthase]	3	2
[gene=mukB] [protein=cell division protein]	3	13
[gene=hisS] [protein=Histidyl-tRNA synthetase]	3	3
[gene=prlC] [protein=oligopeptidase A]	3	0
[gene=carB] [protein=carbamoyl-phosphate synthase large chain]	3	0

[gene=ppk] [protein=Polyphosphate kinase]	3	3
[gene=ssb] [protein=single-strand DNA-binding protein]	3	0
[gene=sfcA] [protein=malate oxidoreductase]	3	0
[gene=upp] [protein=uracil phosphoribosyltransferase]	3	0
[gene=rpsR] [protein=30s ribosomal subunit protein S18]	3	0
[gene=yebC] [protein=hypothetical protein]	3	0
[gene=SL1344_3662] [protein=hypothetical racemase]	3	3
[gene=glpA] [protein=anaerobic glycerol-3-phosphate dehydrogenase]	3	12
[gene=ileS] [protein=isoleucyl-tRNA synthetase]	3	0
[gene=parC] [protein=topoisomerase IV subunit A]	3	8
[gene=prfC] [protein=peptide chain release factor 3]	3	0
[gene=trxA] [protein=thioredoxin]	3	0
[gene=topA] [protein=DNA topoisomerase I, omega protein I]	2	0
[gene=ilvD] [protein=dihydroxyacid dehydratase]	2	0
[gene=rfbG] [protein=CDP-glucose 4,6-dehydratase]	2	0
[gene=hupA] [protein=histone like DNA-binding protein HU-alpha]	2	0
[gene=hepA] [protein=probable ATP-dependent helicase HepA]	2	5
[gene=ftsZ] [protein=cell division protein FtsZ]	2	2
[gene=pepD] [protein=aminoacyl-histidine dipeptidase precursor]	2	0
[gene=rplP] [protein=50S ribosomal subunit protein L16]	2	6
[gene=SL1344_1223] [protein=hypothetical oxidoreductase]	2	0
[gene=pgm] [protein=phosphoglucomutase]	2	0
[gene=nagD] [protein=NagD protein]	2	0
[gene=galF] [protein=UTP-glucose-1-phosphate uridylyltransferase]	2	0
[gene=yjeA] [protein=lysyl-tRNA synthetase]	2	2
[gene=yecO] [protein=conserved hypothetical protein]	2	0
[gene=rl18] [protein=L18 Ribosomal Protein]	2	0

Repeat 2

Protein [description]	Peptides detected	Peptides in Control
[gene=tufA] [protein=elongation factor Tu]	39	49
[gene=rpoC] [protein=DNA-directed RNA polymerase, beta~-subunit]	39	130
[gene=rpoB] [protein=DNA-directed RNA polymerase, beta-subunit]	35	130
[gene=aceE] [protein=pyruvate dehydrogenase E1 component]	35	92
[gene=groEL] [protein=GroEL protein]	28	40
[gene=lon] [protein=Lon protease]	26	78
[gene=fusA] [protein=elongation factor G]	23	24
[gene=dnaK] [protein=Chaperone protein dnaK]	21	53
[gene=rpsA] [protein=30S ribosomal protein S1]	15	38
[gene=pez] [protein=DNA-directed RNA polymerase alpha chain]	14	29
[gene=rplB] [protein=50S ribosomal subunit protein L2]	12	27
[gene=rpsB] [protein=30S ribosomal protein S2]	12	24
[gene=htpG] [protein=heat shock protein HtpG]	11	6

[gene=rplE] [protein=50S ribosomal subunit protein L5]	11	25
[gene=ackA] [protein=acetate kinase]	10	15
[gene=iscS] [protein=Cysteine desulfurase]	10	28
[gene=lpdA] [protein=dihydrolipoamide dehydrogenase]	9	28
[gene=rplN] [protein=50S ribosomal subunit protein L14]	8	13
[gene=clpB] [protein=ClpB protein (heat shock protein f84.1)]	8	45
[gene=rpsD] [protein=30S ribosomal subunit protein S4]	8	18
[gene=rpsC] [protein=30S ribosomal subunit protein S3]	8	13
[gene=atpA] [protein=ATP synthase alpha subunit]	8	30
[gene=glpK] [protein=glycerol kinase]	7	33
[gene=atpD] [protein=ATP synthase beta subunit]	7	30
[gene=cspE] [protein=cold shock-like protein cspE]	6	5
[gene=prs] [protein=Ribose-phosphate pyrophosphokinase]	6	17
[gene=tig] [protein=trigger factor]	6	0
[gene=fabB] [protein=3-oxoacyl-[acyl-carrier-protein] synthase I]	6	8
[gene=accC] [protein=biotin carboxylase]	6	21
[gene=sucA] [protein=2-oxoglutarate dehydrogenase E1 component]	6	53
[gene=infB] [protein=protein chain initiation factor 2]	6	36
[gene=rho] [protein=transcription termination factor]	6	17
[gene=ptsI] [protein=phosphoenolpyruvate-protein phosphotransferase]	5	28
[gene=recA] [protein=RecA protein]	5	18
[gene=glpD] [protein=aerobic glycerol-3-phosphate dehydrogenase]	5	28
[gene=rpsM] [protein=30S ribosomal subunit protein S13]	5	10
[gene=typA] [protein=GTP-binding protein]	5	24
[gene=adh] [protein=alcohol dehydrogenase]	5	59
[gene=sdaA] [protein=L-serine deaminase 1]	5	17
[gene=pta] [protein=phosphate acetyltransferase]	5	30
[gene=deoD] [protein=purine nucleoside phosphorylase]	5	9
[gene=asnCa] [protein=asparaginyl-tRNA synthetase]	5	26
[gene=rplM] [protein=50S ribosomal subunit protein L13]	5	10
[gene=metK] [protein=S-adenosylmethionine synthetase]	4	13
[gene=aspS] [protein=Aspartyl-tRNA synthetase]	4	15
[gene=rpsK] [protein=30S ribosomal subunit protein S11]	4	6
[gene=fabF] [protein=3-oxoacyl-[acyl-carrier-protein] synthase II]	4	11
[gene=slyD] [protein= peptidyl-prolyl cis-trans isomerase]	4	0
[gene=pyrG] [protein=CTP synthase] [protein_id=CBW19031.1]	4	22
[gene=deaD] [protein=ATP-dependent RNA helicase	4	25
[gene=aceF] [protein= component (E2) of pyruvate dehydrogenase]	4	12
[gene=fabG] [protein=3-oxoacyl-[acyl-carrier protein] reductase]	4	6
[gene=rplO] [protein=50S ribosomal subunit protein L15]	4	7
[gene=rpmB] [protein=50S ribosomal subunit protein L28]	4	8
[gene=rplV] [protein=50s ribosomal protein l22]	4	0
[gene=groES] [protein=GroES protein]	3	0
[gene=rplC] [protein=50S ribosomal subunit protein L3]	3	13
[gene=rplI] [protein=50s ribosomal subunit protein L9]	3	0
[gene=eno] [protein=Enolase]	3	6
[gene=rpsJ] [protein=30S ribosomal subunit protein S10]	3	0

[gene=rplD] [protein=50S ribosomal subunit protein L4]	3	11
[gene=alaS] [protein=Alanyl-tRNA synthetase]	3	33
[gene=gapA] [protein=glyceraldehyde 3-phosphate dehydrogenase]	3	0
[gene=rpsL] [protein=30S ribosomal subunit protein S12]	3	8
[gene=tsaA] [protein=alkyl hydroperoxide reductase]	3	6
[gene=rpsS] [protein=30S ribosomal subunit protein S19]	3	8
[gene=pflB] [protein=formate acetyltransferase 1]	3	27
[gene=rpl20] [protein=50S ribosomal protein L20]	3	5
[gene=purA] [protein=adenylosuccinate synthetase]	3	9
[gene=fabI] [protein=enoyl-[acyl-carrier-protein] reductase (NADH)]	3	20
[gene=rpsO] [protein=30S ribosomal subunit protein S15]	3	0
[gene=tsf] [protein=Elongation factor Ts]	3	0
[gene=pheT] [protein=Phenylalanyl-tRNA synthetase beta chain]	3	25
[gene=rpsN] [protein=30S ribosomal subunit protein S14]	3	0
[gene=pykF] [protein=pyruvate kinase]	3	5
[gene=rpl32] [protein=50S ribosomal protein L32]	2	0
[gene=rdgC] [protein=recombination associated protein RdgC]	2	12
[gene=glmS] [protein=glucosamine--fructose-6-phosphate aminotransferase]	2	11
[gene=rplA] [protein=50S ribosomal subunit protein L1]	2	7
[gene=rplF] [protein=50S ribosomal subunit protein L6]	2	5
[gene=valS] [protein=valyl-tRNA synthetase]	2	0
[gene=hupA] [protein=histone like DNA-binding protein HU-alpha]	2	0
[gene=rpsR] [protein=30s ribosomal subunit protein S18]	2	0
[gene=aspA] [protein=aspartate ammonia-lyase]	2	7
[gene=rplQ] [protein=50S ribosomal subunit protein L17]	2	10
[gene=mreB] [protein=rod shape-determining protein]	2	19
[gene=rpsG] [protein=30S ribosomal subunit protein S7]	2	6
[gene=rplK] [protein=50S ribosomal subunit protein L11]	2	5
[gene=yciL] [protein=hypothetical pseudouridine synthase]	2	17
[gene=secB] [protein=protein-export protein SecB] [	2	4
[gene=crp] [protein=cyclic AMP receptor]	2	14
[gene=clpX] [protein=ATP-dependent clp]	2	11
[gene=ftsK] [protein=cell division protein FtsK]	2	0
[gene=kbl] [protein=2-amino-3-ketobutyrate coenzyme A ligase]	2	8

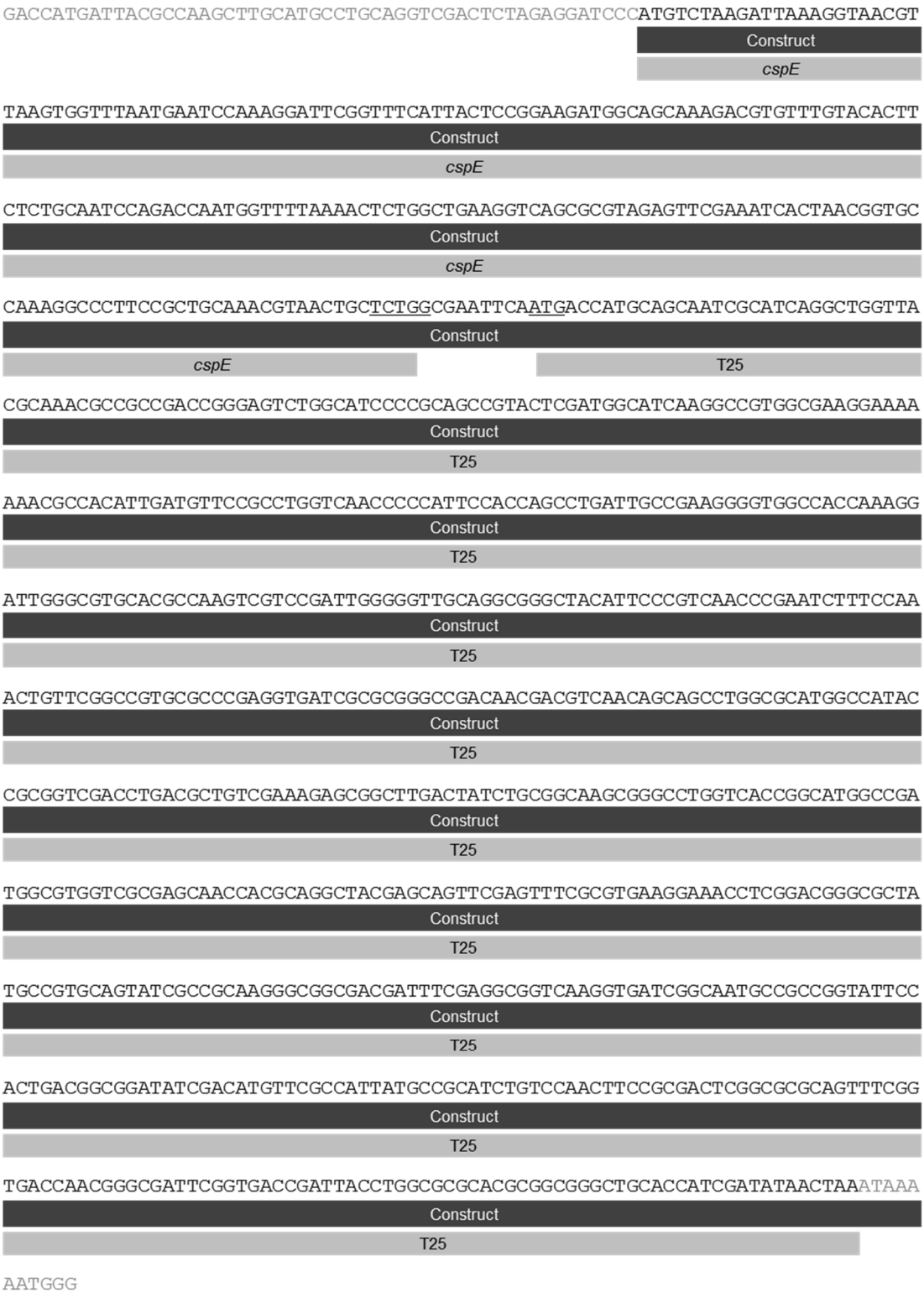
Protein-Protein Interactions of CspE

Repeat 3

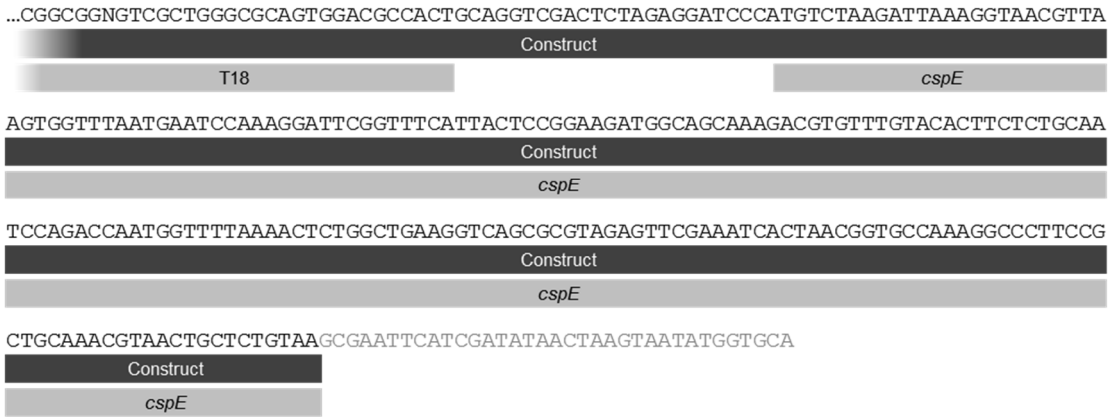
Protein [description]	Peptides detected	Peptides in Control
[gene=tufA] [protein=elongation factor Tu]	20	27
[gene=rplB] [protein=50S ribosomal subunit protein L2]	12	10
[gene=rne] [protein=ribonuclease E]	9	0
[gene=hfq] [protein=host factor-I protein(HF-I)]	8	4
[gene=cspE] [protein=cold shock-like protein cspE]	6	6
[gene=rplQ] [protein=50S ribosomal subunit protein L17]	6	4
[gene=rplN] [protein=50S ribosomal subunit protein L14]	6	4
[gene=rpsD] [protein=30S ribosomal subunit protein S4]	6	4
[gene=rpsM] [protein=30S ribosomal subunit protein S13]	5	4
[gene=rpsB] [protein=30S ribosomal protein S2]	5	5
[gene=rplM] [protein=50S ribosomal subunit protein L13]	5	3
[gene=rpsO] [protein=30S ribosomal subunit protein S15]	5	4
[gene=rnt] [protein=ribonuclease T]	4	2
[gene=crp] [protein=cyclic AMP receptor protein,catabolite gene	4	4
[gene=yqjl] [protein=conserved hypothetical protein]	4	0
[gene=rpl20] [protein=50S ribosomal protein L20]	4	4
[gene=clpX] [protein=ATP-dependent clp protease ATP-binding subunit	4	3
[gene=slyD] [protein=FKBP-type peptidyl-prolyl cis-trans isomerase]	3	3
[gene=rpsK] [protein=30S ribosomal subunit protein S11]	3	2
[gene=clpB] [protein=ClpB protein (heat shock protein f84.1)]	3	8
[gene=ptsI] [protein=phosphoenolpyruvate-protein phosphotransferase]	3	1
[gene=gyrB] [protein=DNA gyrase subunit B]	3	0
[gene=ychF] [protein=hypothetical ATP/GTP-binding protein]	3	2
[gene=rplU] [protein=50S ribosomal subunit protein L21]	3	1
[gene=rplE] [protein=50S ribosomal subunit protein L5]	3	5
[gene=rpsC] [protein=30S ribosomal subunit protein S3]	2	10
[gene=recA] [protein=RecA protein]	2	1
[gene=groEL] [protein=GroEL protein]	2	14
[gene=lipA] [protein=lipoic acid synthetase]	2	1
[gene=rpsR] [protein=30s ribosomal subunit protein S18]	2	2
[gene=lepA] [protein=GTP-binding protein LepA]	2	0
[gene=rplO] [protein=50S ribosomal subunit protein L15]	2	3
[gene=rplC] [protein=50S ribosomal subunit protein L3]	2	1
[gene=cbpA] [protein=curved DNA-binding protein]	2	0
[gene=rpsL] [protein=30S ribosomal subunit protein S12]	2	1
[gene=iscS] [protein=Cysteine desulfurase]	2	3
[gene=rplA] [protein=50S ribosomal subunit protein L1]	2	4
[gene=fusA] [protein=elongation factor G]	2	18
[gene=mrp] [protein=conserved hypothetical protein]	2	0
[gene=rpsS] [protein=30S ribosomal subunit protein S19]	2	1

Appendix B: Construct Sequences

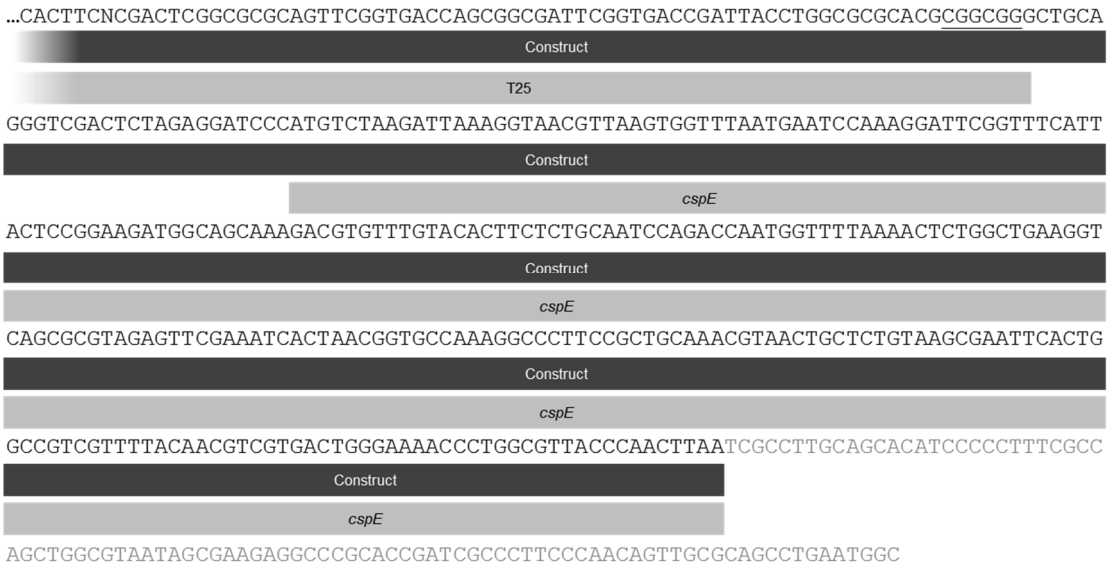
pKT25-cspE



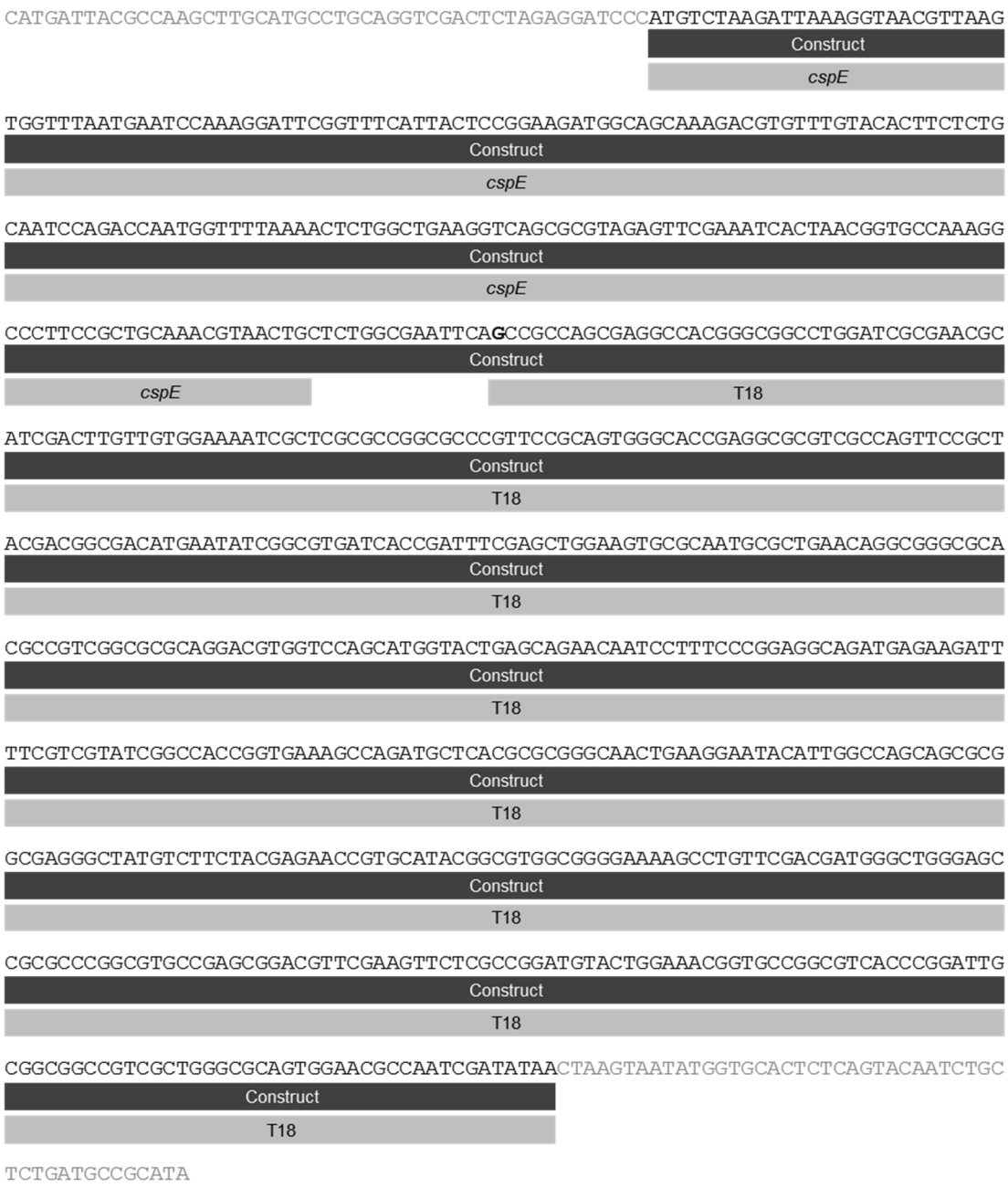
pKNT25-cspE



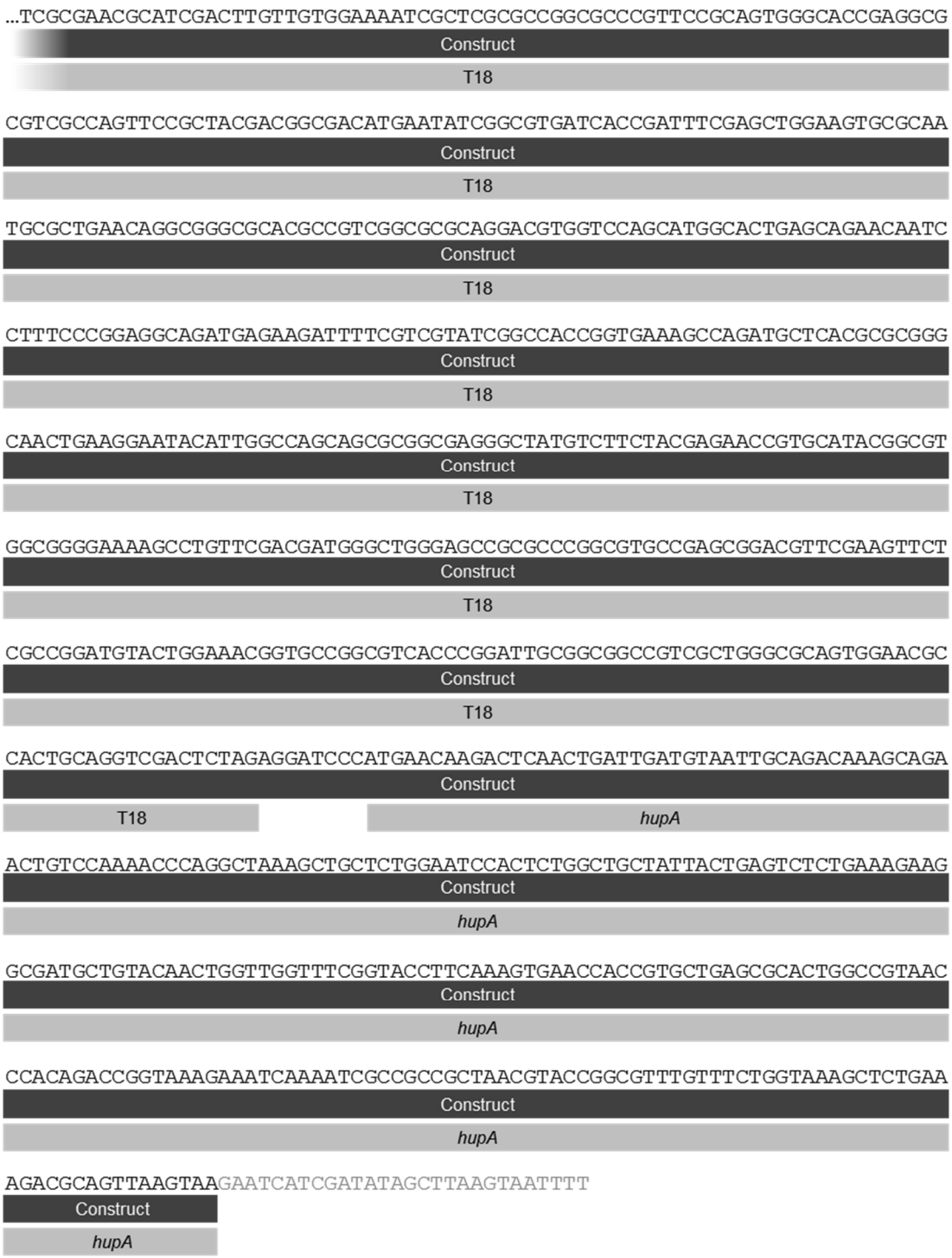
pUT18C-cspE



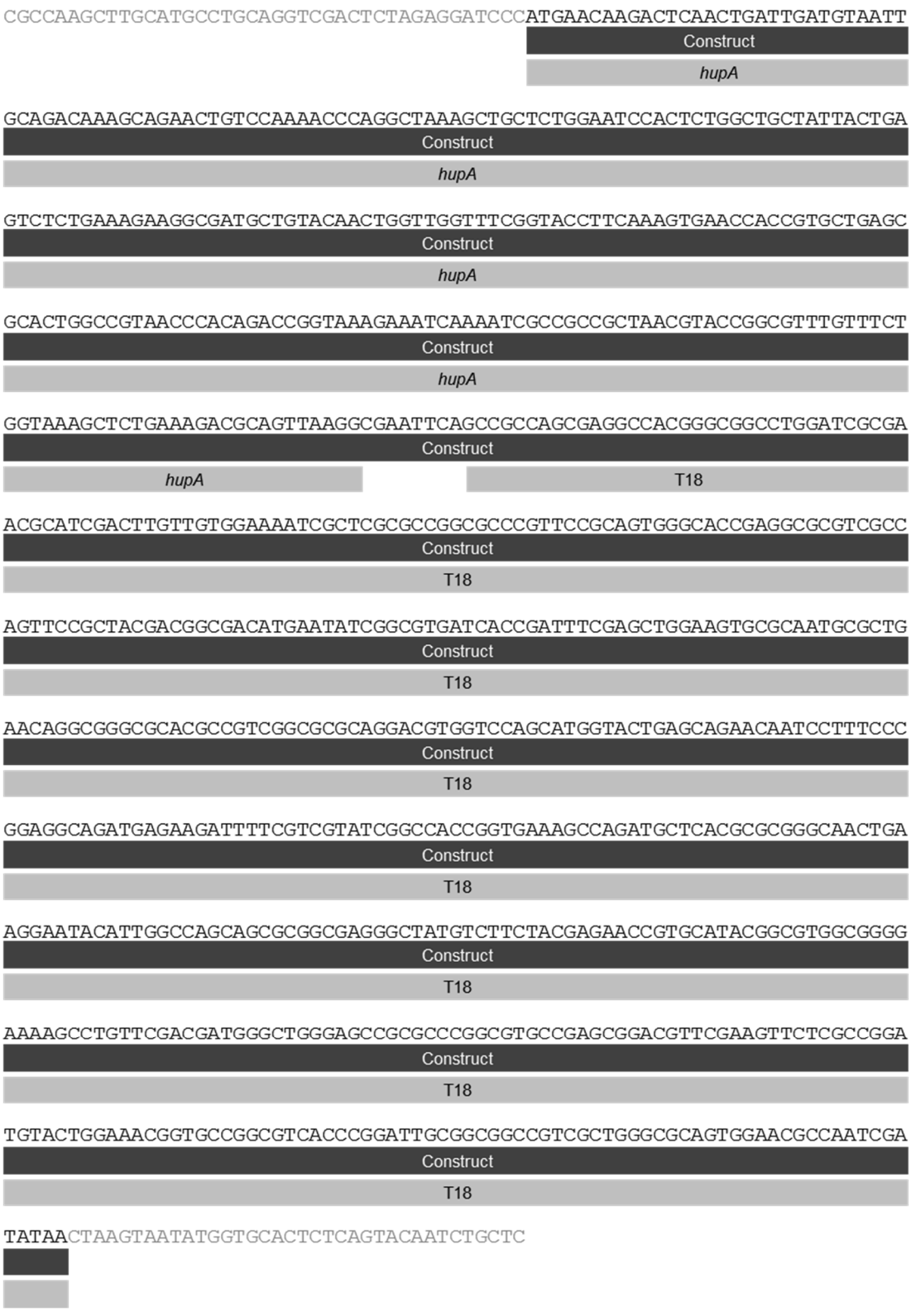
pUT18-cspE


CATGATTACGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCATGTCTAAGATTAAAGGTAACGTTAAG
Construct
cspE
TGGTTTAATGAATCCAAAGGATTCGGTTTCATTACTCCGGAAGATGGCAGCAAAGACGTGTTTGTACACTTCTCTG
Construct
cspE
CAATCCAGACCAATGGTTTTAAAACTCTGGCTGAAGGTCAGCGCGTAGAGTTCGAAATCACTAACGGTGCCAAAGG
Construct
cspE
CCCTTCCGCTGCAAACGTAACTGCTCTGGCGAATTCAGCCGCCAGCGAGGCCACGGGCGGCCTGGATCGCGAACGC
Construct
cspE
T18
ATCGACTTGTTGTGGAAAATCGCTCGCGCCGGCGCCCGTTCCGCAGTGGGCACCGAGGCGCGTCGCCAGTTCCGCT
Construct
T18
ACGACGGCGACATGAATATCGGCGTGATCACCGATTTCGAGCTGGAAGTGCGCAATGCGCTGAACAGGCGGGCGCA
Construct
T18
CGCCGTCGGCGCGCAGGACGTGGTCCAGCATGGTACTGAGCAGAACAATCCTTTCCCGGAGGCAGATGAGAAGATT
Construct
T18
TTCGTCGTATCGGCCACCGGTGAAAGCCAGATGCTCACGCGCGGGCAACTGAAGGAATACATTGGCCAGCAGCGCG
Construct
T18
GCGAGGGCTATGTCTTCTACGAGAACCGTGCATACGGCGTGGCGGGGAAAAGCCTGTTCGACGATGGGCTGGGAGC
Construct
T18
CGCGCCCGGCGTGCCGAGCGGACGTTCGAAGTTCTCGCCGGATGTACTGGAAACGGTGCCGGCGTCACCCGGATTG
Construct
T18
CGGCGGCCGTCGCTGGGCGCAGTGGAACGCCAATCGATATAACTAAGTAATATGGTGCACTCTCAGTACAATCTGC
Construct
T18
TCTGATGCCGCATA

pUT18C-hupA


...TCGCGAACGCATCGACTTGTTGTGGAAAATCGCTCGCGCCGGCGCCCGTTCCGCAGTGGGCACCGAGGCG
Construct
T18
CGTCGCCAGTTCCGCTACGACGGCGACATGAATATCGGCGTGATCACCGATTTCGAGCTGGAAGTGCGCAA
Construct
T18
TGCGCTGAACAGGCGGGCGCACGCCGTCGGCGCGCAGGACGTGGTCCAGCATGGCACTGAGCAGAACAATC
Construct
T18
CTTTCCCGGAGGCAGATGAGAAGATTTTCGTCGTATCGGCCACCGGTGAAAGCCAGATGCTCACGCGCGGG
Construct
T18
CAACTGAAGGAATACATTGGCCAGCAGCGCGGCGAGGGCTATGTCTTCTACGAGAACCGTGCATACGGCGT
Construct
T18
GGCGGGGAAAAGCCTGTTCGACGATGGGCTGGGAGCCGCGCCCGGCGTGCCGAGCGGACGTTCGAAGTTCT
Construct
T18
CGCCGGATGTACTGGAAACGGTGCCGGCGTCACCCGGATTGCGGCGGCCGTCGCTGGGCGCAGTGGAACGC
Construct
T18
CACTGCAGGTCGACTCTAGAGGATCCCATGAACAAGACTCAACTGATTGATGTAATTGCAGACAAAGCAGA
Construct
T18
hupA
ACTGTCCAAAACCCAGGCTAAAGCTGCTCTGGAATCCACTCTGGCTGCTATTACTGAGTCTCTGAAAGAAG
Construct
hupA
GCGATGCTGTACAACTGGTTGGTTTCGGTACCTTCAAAGTGAACCACCGTGCTGAGCGCACTGGCCGTAAC
Construct
hupA
CCACAGACCGGTAAAGAAATCAAAATCGCCGCCGCTAACGTACCGGCGTTTGTTTCTGGTAAAGCTCTGAA
Construct
hupA
AGACGCAGTTAAGTAAGAATCATCGATATAGCTTAAGTAATTTT
Construct
hupA

pUT18-hupA


CGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCATGAACAAGACTCAACTGATTGATGTAATT
Construct
hupA
GCAGACAAAGCAGAACTGTCCAAAACCCAGGCTAAAGCTGCTCTGGAATCCACTCTGGCTGCTATTACTGA
Construct
hupA
GTCTCTGAAAGAAGGCGATGCTGTACAACTGGTTGGTTTCGGTACCTTCAAAGTGAACCACCGTGCTGAGC
Construct
hupA
GCACTGGCCGTAACCCACAGACCGGTAAAGAAATCAAAATCGCCGCCGCTAACGTACCGGCGTTTGTTTCT
Construct
hupA
GGTAAAGCTCTGAAAGACGCAGTTAAGGCGAATTCAGCCGCCAGCGAGGCCACGGGCGGCCTGGATCGCGA
Construct
hupA
T18
ACGCATCGACTTGTTGTGGAAAATCGCTCGCGCCGGCGCCCGTTCCGCAGTGGGCACCGAGGCGCGTCGCC
Construct
T18
AGTTCCGCTACGACGGCGACATGAATATCGGCGTGATCACCGATTTCGAGCTGGAAGTGCGCAATGCGCTG
Construct
T18
AACAGGCGGGCGCACGCCGTCGGCGCGCAGGACGTGGTCCAGCATGGTACTGAGCAGAACAATCCTTTCCC
Construct
T18
GGAGGCAGATGAGAAGATTTTCGTCGTATCGGCCACCGGTGAAAGCCAGATGCTCACGCGCGGGCAACTGA
Construct
T18
AGGAATACATTGGCCAGCAGCGCGGCGAGGGCTATGTCTTCTACGAGAACCGTGCATACGGCGTGGCGGGG
Construct
T18
AAAAGCCTGTTCGACGATGGGCTGGGAGCCGCGCCCGGCGTGCCGAGCGGACGTTCGAAGTTCTCGCCGGA
Construct
T18
TGTACTGGAAACGGTGCCGGCGTCACCCGGATTGCGGCGGCCGTCGCTGGGCGCAGTGGAACGCCAATCGA
Construct
T18
TATAACTAAGTAATATGGTGCACTCTCAGTACAATCTGCTC