The University Of Sheffield.

Characterisation of TssA, a core component of the bacterial type VI secretion system

By:

Ruyue Sun

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

The University of Sheffield
Faculty of Medicine, Dentistry and Health
Department of Infection, Immunity & Cardiovascular Disease

Supervisor: Dr. Mark S. Thomas

29th July 2016

Abstract

Protein secretion in bacteria is a process of transporting selected protein molecules from the cytoplasm of a bacterial cell where they are synthesised, across the cell envelope to the cell exterior, which is an important mechanism in bacterial functioning, adaptation to the natural surrounding environment and survival. So far, nine protein secretion systems have been identified in Gram-negative bacteria, including the type VI secretion system (T6SS). The current model of the T6SS apparatus is based on evolutionary, structural and mechanistic similarities between several structural proteins of the T6SS and the gene products that make up the injection machinery of some contractile bacteriophages. The T6SS is composed of at least 13 core structural components TssA-TssM, among which TssA is the one of least characterised. In this study, two types of TssA, namely $TssA^{S}$ and $TssA^{EI}$, were analysed. They contain a conserved N-terminal region of approximately 60-80 amino acids but diverge with respect to their C-terminal regions. Both types of TssA were shown to be involved in similar interactions with many other T6SS core components both *in vivo* and *in vitro*, and are therefore proposed to be functional orthologues. Both types of TssA oligomerise into high molecular mass complexes. The domain organisation analysis suggested they both contain a NTD, MD and a small CTD which serves as the oligomerisation module for assembly of high molecular mass complexes. Both $TssA^{S}$ and $TssA^{EI}$ oligomerise into ring-like structures as visualised by TEM. However, a distinct $TssA^{S}_{CTD}$ ring is surrounded by disassociated $TssA^{S}_{NTD}$ whereas $TssA^{EI}$ has a broader thickness with an irregular outer edge and the regions corresponding to the NTD and CTD are not distinguishable. High resolution X-ray crystallographic analyses of $TssA^{S}$ and $TssA^{EI}$ domains were also presented, including $TssA^{S}_{NTD}$, $TssA^{S}_{CTD}$ and $TssA^{EI}_{MD-CTD}$. Overall, TssA is likely to locate within or close to the baseplate and probably undergoes conformational changes during T6SS firing.

Acknowledgements

First and foremost I would like to thank my supervisor Dr. Mark Thomas for his constant support and guidance throughout my PhD and in the writing of this thesis. I have learnt a lot during the entire time of my research.

I would like to thank all the group members I have been working with- Asma, Saeedah, Sayali, Helena, Jamie, Syakira, Aaron, Viktoriya, Daniel, Sarah, Chris and Richard for making my lab experience enriching and enjoyable.

I would also like to thank the collaborators who helped this project, Dr. Svetomir Tzokov and Prof. Per Bullough for the EM support, and Dr. Svetlana Sedelnikova, Hayley Owen, Sam Dix and Prof. David Rice for the crystallography analysis.

Thanks the following publishers for the permissions of figure reproduction: Macmillan Publishers Ltd: Nature Reviews Microbiology (Costa et al. 2015, doi:10.1038/nrmicro3456), Macmillan Publishers Ltd: Nature (Durand et al. 2015, doi:10.1038/nature14667; Taylor et al. 2016, doi:10.1038/nature17971; Zoued et al. 2016, doi:10.1038/nature17182), Philosophical Transactions of the Royal Society B (Cascales and Cambillau 2012, doi: 10.1098/rstb.2011.0209), and Cell Reports (Kube et al. 2014, doi: http://dx.doi.org/10.1016/j.celrep.2014.05.034).

Last, I am grateful for the continuous love, support and encouragement that my family have been giving me, without which I would not be able to achieve this.

List of Abbreviations

ABC - ATP-binding cassette
ACD - actin cross-linking domain
Act - *Aeromonas* cytotoxic enterotoxin
ADPRT - ADP-ribosyltransferase
AHL - acylhomoserine lactone
ATP - adenosine triphosphate
BACTH - bacterial adenylate cyclase two-hybrid
Bcc - *Burkholderia cepacia* complex
BcCV - *B. cenocepacia*-containing vacuole
BCESM - *B. cenocepacia* epidemic strain marker
BHI - brain-heart infusion
bp – base pairs
BSA – Bovine serum albumin
cAMP - cyclic AMP
CCI - *B. cenocepacia* genomic island
CF - cystic fibrosis
CIP - calf intestinal alkaline phosphatase
co-IP - co-Immunoprecipitation
CRP - cAMP receptor protein
CTD – C-terminal domain
CU - chaperone-usher
Da – daltons
D-BHI - dialysed brain-heart infusion
ddH_2O - double-distilled water
EAEC - enteroaggregative *Escherichia coli*
EDTA - ethylenediaminetetraacetic acid
EM - electron microscopy
ENP - extracellular nucleation-precipitation
ET12 - electrophoretic type-12
FHA - forkhead-associated
gp - gene product
H-NS - histone-like nucleoid associated
HRP - horseradish peroxidase
IAHP - IcmF-associated homologous protein
IgG - immunoglobulin G
IM - inner membrane
IMAC - immobilized metal ion affinity chromatography
IPTG - isopropyl β-D-1-thiogalactopyranoside
kDa - lilodaltons
LB - lysogeny broth
LPS - lipopolysaccharide

LPS - lipopolysaccharide
LS - light scattering
MBP - maltose binding protein
Mbp - megabase pair
MCS – multiple cloning site
MD – middle domain
MDa - megadaltons
MFP - membrane fusion protein
ml – millilitres
mM – millimolar
MRA - multi-reference alignment
Mu – Miller units
MW – molecular weight
MWCO - molecular weight cut-off
NTD – N-terminal domain
NTR – N-terminal region
OM - outer membrane
PAAR - proline-alanine-alanine-arginine repeat
PCR - polymerase chain reaction
PEI - polyethyleneimine
PG - peptidoglycan
POTRA - polypeptide transport associated
QELS - quasi-elastic light scattering
Rhs - recombination hot spot
RI- refractive index
SAXS - small angle X-ray scattering
SEC - size exclusion chromatography
SEC-MALS - size exclusion chromatography-multi-angle laser light scattering
SOE-PCR - splicing by overlap extension PCR
SRP - signal recognition particle
STPK - serine/threonine protein kinase
T1SS - Type I secretion system
T2SS - Type II secretion system
T3SS - Type III secretion system
T4SS - Type IV secretion system
T5SS - Type V secretion system
T6SS - Type VI secretion system
T7SS - Type VII secretion system
T8SS - Type VIII secretion system
T9SS - Type IX secretion system
TAA - trimeric autotransporter adhesion
Tae - type VI amidase effector
Tag - T6SS associated genes

Tai - type VI amidase immunity
TalT - TssA-like tag
Tat - twin-arginine translocation
TBS - Tris-buffered saline
TEC - T6SS effector chaperone
TEM - transmission electron microscopy
Tle - type VI lipase effectors
Tli - type VI lipase immunity
TLR5 - toll-like receptor 5
TMD – transmembrane domain
TPP - threonine phosphorylation pathway
TPS - two-partner secretion system
TTR - transthyretin
UPEC - uropathogenic *Escherichia coli*
v/v – volume/volume
WT - wild-type
w/v – weight/volume
X-gal - 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
γTuRC - γ-tubulin ring complex
Δ – deletion
µg – microgram
µl – microlitre

Table of contents

Chapter 1 Introduction

1.1 General introduction

Protein secretion in bacteria is a process of transporting selected protein molecules from the cytoplasm of a bacterial cell where are they synthesised, across the cell envelope to the cell exterior, which is an important mechanism in bacterial functioning, adaption the natural surrounding environment and survival. So far, nine secretion systems have been identified in the Gram-negative bacteria, including the type VI secretion system (T6SS). The T6SS is composed of at least 13 core structural components TssA-TssM, and it is encoded in at least 25% of all sequenced Gram-negative bacterial genomes, including pathogenic bacteria such as *Aeromonas hydrophila*, *Vibrio cholerae*, *Pseudomonas aeruginosa*, the *Burkholderia* species (Mougous et al. 2006; Pukatzki et al. 2006; Schell et al. 2007; Shalom et al. 2007).

One of the pathogenic species to possess a T6SS is *Burkholderia cenocepacia*, one of 20 closely related bacterial species that belongs to the *Burkholderia cepacia* complex (Bcc). Many Bcc bacterial species can causes disease in plants and are opportunistic human pathogens that lead to fatal infections in immunocompromised and cystic fibrosis patients. *A. hydrophila* is ubiquitously present in aquatic environments and food, and can cause diseases in fish as well as being a human pathogen in both immunocompromised and healthy people where its effects range from gastroenteritis and skin infections to systemic infections, such as bacteremia and necrotizing fasciitis. It was found that some virulence determents of *B. cenocepacia* and *A. hydrophila* are secreted in a T6SS-dependent manner.

During the last five years, significant insights into the mode of action of the T6SS have been achieved. It contains a membrane chamber complex spanning the cell envelope and a contractile phage tail-like structure. Despite this, the structure, localisation and potential role of some core components still have not been fully elucidated, including TssA.

1.2 The Bcc complex

1.2.1 Discovery and classification of the Bcc

The first member of the *Burkholderia* spp. to be isolated occurred in the early 1940s and was identified as a Gram-negative soil-dwelling plant pathogen that caused rot disease on onion bulbs. It was proposed to be named as *Pseudomonas cepacia* as it was thought to belong to the genus *Pseudomonas* with the species name '*cepacia*' indicating it was derived from onion (Burkholder 1950). With the advent of taxonomic framework over the years by rRNA-DNA hybridisation analyses and rRNA gene sequencing methods, the genus *Pseudomonas* was revised and *Pseudomonas cepacia* and six other species were transferred to a new genus *Burkholderia*, that belongs to β-proteobacteria class, while true *Pseudomonas* that belong to the γ-proteobacteria (Woese 1987; Yabuuchi et al. 1992). Subsequently, a seminal study of strains identified phenotypically as *Burkholderia cepacia* was carried out by Vandamme and colleagues (1997) that identified genetic diversity among these species and therefore divided these strains into nine genomovars (i.e. genetic species), referred to as the Bcc, including *B. cepacia* (genomovar I), *Burkholderia multivorans* (genomovar II), *Burkholderia cenocepacia* (genomovar III), *Burkholderia stabilis* (genomovar IV), *Burkholderia vietnamiensis* (genomovar V), *Burkholderia cepacia* genomovar VI, *Burkholderia ambifaria* (genomovar VII), *Burkholderia anthina* (genomovar VIII) and *Burkholderia pyrrocinia* (genomovar IX) (Vandamme et al. 1997; Vandamme et al. 2000; Coenye et al. 2001a; Coenye et al. 2001b; Vandamme et al. 2002; Vandamme et al. 2003) . The Bcc species generally have a rod-shaped flagella phenotype and inhabit in a range of places, including water, soil, plant rhizosphere, animals and humans. To date, the Bcc has risen to twenty closely related species including two recently included species *Burkholderia stagnalis* and *Burkholderia territorii* (De Smet et al. 2015). A summary of the taxonomy of the Bcc is shown in Figure 1.1.

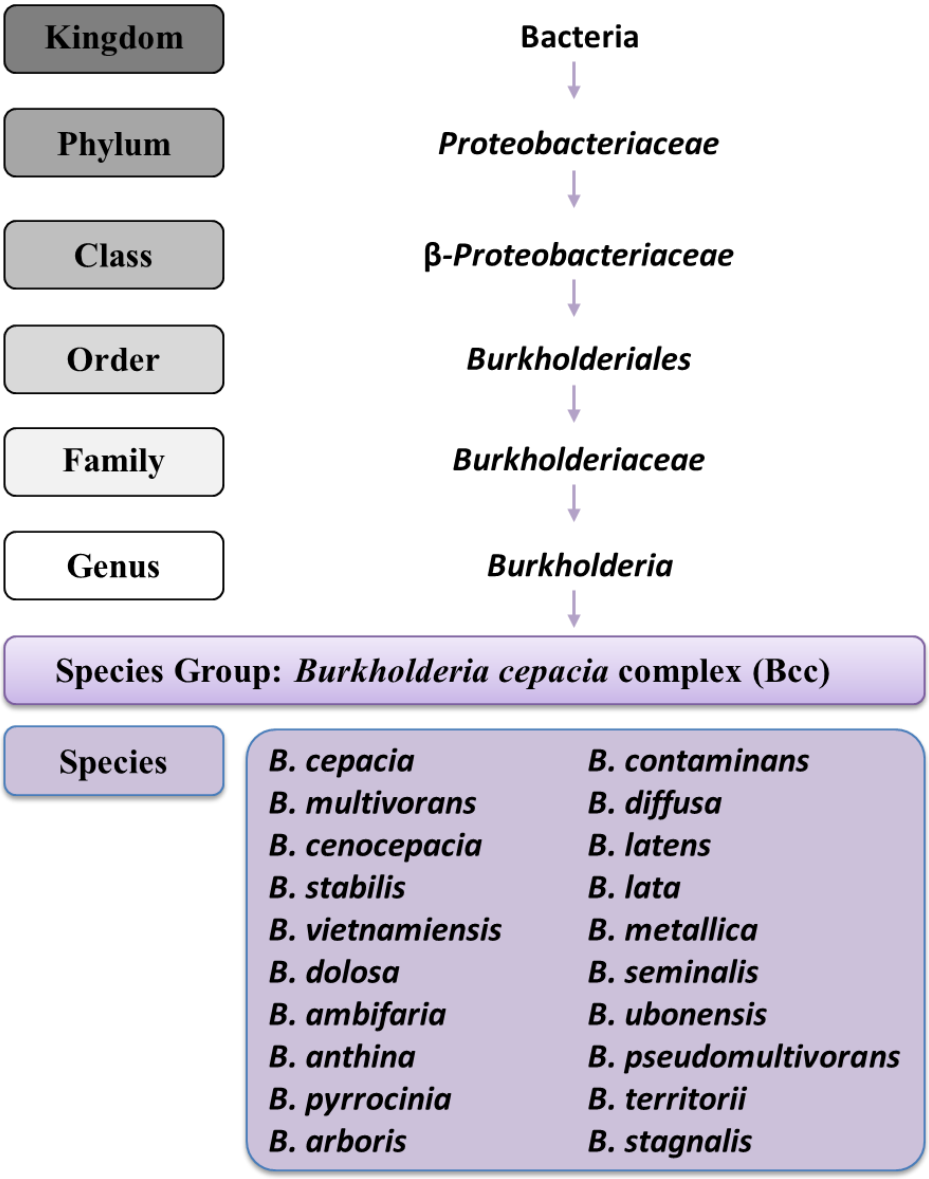


Figure 1.1 Taxonomy of the *Burkholderia cepacia* complex. Diagram listing the taxonomic classification of 20 bacterial species within the complex.

1.2.2 Genomics of the Bcc

The Bcc species examined so far contain large genomes ranging from 6-9 Mbp that are organised into multiple chromosomal replicons, and this huge genetic capacity underpins the versatility of the Bcc members in the natural environment and disease causation. In the past few years, great advances have been made in sequencing the whole genomes of the Bcc members, and this includes the human pathogenic *B. cenocepacia* strains J2315 and H111 (Holden et al. 2009) and the animal and plant pathogen *B. pseudomallei* K96243 (Holden et al. 2004) and the related species *B. thailandensis* E264 (Yu et al. 2006).

B. cenocepacia, as the most clinically important member of the Bcc of opportunistic pathogens, has triggered great interest. The genome of *B. cenocepacia* strain J2315, which was isolated from cystic fibrosis (CF) patients, belongs to the epidemic electrophoretic type-12 (ET12) lineage, and exhibits multi-drug resistance, has been sequenced. It has an 8.06 Mbp genome, organised in three circular chromosomes with sizes of 3.87 Mbp (named BCAL), 3.22 Mbp (BCAM) and 0.88 Mbp (BCAS) and a small plasmid of 92.7 kb, containing an average GC content of 66.9%. This genome contains approximately 7,300 genes encoding ~7,100 predicted proteins and it has also been shown that 21% of the *B. cenocepacia* J2315 genome is unique compared to other *B. cenocepacia* species highlighting its genomic plasticity (Holden et al. 2009). The genome also exhibits two highlighted features, the first is the relatively large number of insertion sequences, i.e. 79, which are small transposable mobile DNA sequences identified in bacteria that encode proteins only associated with their transposition activity (Mahillon and Chandler 1998). The other feature is the presence of multiple genomic islands, which are regions of genomic sequences that have a foreign origin as identified by their altered GC content, and probably arose from horizontal gene transfer.

One genomic island present in the *B. cenocepacia* genome, termed the *B. cenocepacia* island (CCI), has been identified as a bacterial pathogenicity island encoding genes involved in both virulence and metabolism (Hacker and Kaper 2000; Baldwin et al. 2004). The CCI also encodes a *B. cenocepacia* epidemic strain marker (BCESM) that was identified as a unique DNA region in virulent *B. cenocepacia* strains isolated from CF patients. Subsequently, studies suggested that BCESM is prevalent in all ET12 lineage strains and in several but not in all CF isolated *B. cenocepacia* strains. However, it has

been shown that BCESM containing strains are transmissible between patients and cause increased mortality in CF patients (Mahenthiralingam et al. 2001).

1.2.3 Environmental significance of the Bcc

Despite the fact that the Bcc can cause harmful effects on plants, such as sour skin disease in onion, several Bcc members are considered to be highly beneficial to the ecological systems in the natural environment. This includes significant agricultural uses. Many Bcc species have been shown to live as plant commensals in the rhizospheric soil of many commercially important plants and protect against fungal diseases as well as enhancing the yield, such as rice, wheat and maize (Parke and Gurian-Sherman 2001). The Bcc in the rhizosphere may facilitate root colonisation, biocontrol and promote plant growth. In addition, the Bcc members are also involved in controlling the spread of the soil-borne plant fungal disease, known as damping-off, which causes seed and seedling death by *Pythium* species. The Bcc members are supposed to achieve the biocontrol of damping-off disease by producing anti-fungal metabolites, such as antimicrobials and siderophores (microbial competition for iron). This biopesticidal (biocontrol agents) feature of Bcc members might supplant the use of traditional harmful chemical pesticides (Parke and Gurian-Sherman 2001). Apart from agricultural benefits, some Bcc members have also been shown to be able to degrade certain pollutants from groundwater and soil by using various complex carbon sources and therefore they may serve as bioremediation agents (Sangodkar et al. 1988; Krumme et al. 1993).

1.2.4 Clinical epidemiology of the Bcc

Aside from having many significant benefits on the natural environment, Bcc, as opportunistic pathogens, have significant clinical effects on humans who are immunocompromised and therefore not generally healthy individuals. CF patients are highly susceptible to Bcc lung infection. The clinical epidemiology of Bcc in CF patients was first recognised in the early 1980s when it was reported that ~18% of CF patients in a Canadian clinic were infected with Bcc with an 8% increase during the previous 10 years (Isles et al. 1984). However, variability in the clinical outcome was observed after

colonisation with a Bcc strain: some CF patients had chronic asymptomatic carriage, whereas some had accelerated deteriorating lung symptoms over months. A very small portion of Bcc infected CF patients caused acute deterioration in lung function with bacteraemia, necrotizing pneumonia, leucocytosis, and ultimately death in few weeks, known as cepacia syndrome (Isles et al. 1984; Zahariadis et al. 2003). Additionally, several Bcc bacteria were shown to be capable of transmission among patients, resulting in devastating infections (Lipuma et al. 1990; Govan et al. 1993). Moreover, the intrinsic resistance of Bcc bacteria to clinically available drugs leads to untreatable infections. So far, all Bcc species have been isolated from infected CF patients, including two recently described Bcc species *B. stagnalis* and *B. territorii*, and thereby suggesting the infection ability of the Bcc is not species dependent, although some species are more prevalent than others geographically and regionally (Mahenthiralingam et al. 2002; De Smet et al. 2015). The majority of isolates from CF patients in the US, Canada and Italy, belong to *B. cenocepacia* and *B. multivorans* species (Lipuma et al. 2001; Reik et al. 2005), and *B. contaminans* has been suggested to be the most frequently isolated from CF patients in Argentina and Spain (Martina et al. 2013; Medina-Pascual et al. 2015).

In consideration of the prevalence among CF patients, person-to-person transmission, the capability of causing life-threatening infections and the intrinsic resistance of Bcc members to multiple antibiotics, it highlights the risk of this pathogen for CF and other immunocompromised patients. Bcc members probably bear various common features that facilitate their infection to take place allowing them to function as highly problematic opportunistic pathogens. However, the molecular mechanisms which contribute to the virulence of the Bcc bacteria are still unclear.

1.2.5 Virulence determinants of the Bcc

Research on the virulence determinants may unveil the molecular mechanisms underlying Bcc pathogenicity that may allow the development of strategies, antibiotics or vaccines to combat Bcc infections. Bcc can produce a wide variety of virulence determinants that allow them to colonise the host efficiently, gaining sufficient nutrients while evading or supressing the host immunity system, and thereby being able to thrive as important opportunistic pathogens. Some of these common virulence determinants

include adhesins, haemolysin production, host cell invasion, intracellular survival, acquisition of iron, biofilm formation and protein secretion systems such as the type II, IV and VI secretion systems (Finlay and Falkow 1997; Aubert et al. 2008). Some of these virulence factors are regulated by quorum sensing systems, such as cell communication mediated acylhomoserine lactone (AHL)-dependent CepIR and CciIR systems and the expression of siderophores.

Haemolysin has been shown to be secreted by the Bcc species from both clinical and environmental isolates, including *B. cenocepacia*, which induces apoptosis and degranulation of mammalian phagocytes (Hutchison et al. 1998; Bevivino et al. 2002). In addition, the Bcc bacteria can produce four types of iron-chelating siderophores, i.e. ornibactin, pyochelin, cepabactin and cepaciachelin, which also contribute to the Bcc pathogenesis (Visser et al. 2004; Thomas 2007). Many studies have demonstrated invasion and intracellular survival of the Bcc bacteria in host cells. *B. cenocepacia* has been shown to invade and survive within A549 human alveolar epithelial cells *in vitro* (Cieri et al. 2002). Different mechanisms of invasion may be used by various Bcc species. *B. cenocepacia* was shown to form biofilms at the surface of epithelial cells followed by invasion and destruction of the cells which involved disruption of the glycocalyx and rearrangements of the actin cytoskeleton on the cell surface, whereas invasion by *B. stabilis* (genomovar IV) is biofilm-independent and penetrated the epithelium between cells (Schwab et al. 2002). Other than invading and survival in epithelial cells, *B. cenocepacia* and other Bcc members can also survive within phagocytes, such as macrophages, within *B. cenocepacia*-containing vacuoles (BcCVs) (Saini et al. 1999; Martin and Mohr 2000; Andrade and Valvano 2014).

Two most prevalent Bcc bacteria in CF patients, *B. cenocepacia* and *B. multivorans*, have been shown to form the most abundant biofilms among genomovars I to V *in vitro*, which is regulated by an AHL-quorum sensing system and maybe other signalling pathways (Conway et al. 2002). Biofilm formation may participate in the virulence of Bcc by facilitating invasion into the host or increasing resistance to antibiotics (Desai et al. 1998).

Bacterial surface structures such as lipopolysaccharide (LPS), flagella, cable pili and adhesins also play an essential role in the interaction with CF hosts. Similar to other LPS in other Gram-negative bacteria, the LPS of Bcc bacteria can induce a strong immune

response that contributes to host cell damage. However, it differs from that of other Gram-negative bacteria in having less phosphate in the core oligosaccharide and 4-amino-4-deoxyarabinose moieties bound to the phosphates of the lipid A backbone. These LPS modifications of the Bcc bacteria lower the anionic charge of the Bcc bacterial cell surface that implicating binding inhibition of cationic anti-microbial peptides and polymyxin (Cox and Wilkinson 1991; Shimomura et al. 2003). In addition, pili and an associated 22-kDa adhesin in *B. cenocepacia* have been shown to be involved in pathogenesis through adherence to host airway epithelial cells (Sajjan and Forstner 1993; Sajjan et al. 2000; Urban et al. 2005). Functional flagella are also required for cellular invasion, but are not essential for adherence by *B. cenocepacia* (Tomich et al. 2002). *B. cenocepacia* flagella were shown to contribute to virulence *in vivo* through a TLR5 (toll-like receptor 5)-mediated signalling pathway (Urban et al. 2004).

Secreted effectors of the Bcc bacteria by protein secretion systems are also important determents for the bacterial virulence and survival. Two zinc metalloproteases, ZmpA and ZmpB, are secreted by the type II secretion system (T2SS) of *B. cenocepacia*. These induce casein degradation *in vitro* and virulence *in vivo*, and gene expression of which is regulated by both the CepIR and CciIR quorum-sensing systems (Corbett et al. 2003; Kooi et al. 2005; Kooi et al. 2006). It is proposed that these metalloproteases may be involved in proteolytic degradation of proteins, such as fibronectin and collagen, leading to tissue damage and/or modulating host defences (Kooi et al. 2005). The type III secretion system (T3SS) has been shown to secrete virulence-associated protein, BscN, which is a potential pathogenicity determinant in immunocompromised and CF patients (Tomich et al. 2003). Trimeric autotransporter adhesins (TAAs), multimeric surface proteins secreted by the type V secretion system (T5SSa), have been shown to be involved in the virulence of the Bcc bacteria, and the epidemic strain *B. cenocepacia* J2315 was shown to encode several putative TAAs (Mil-Homens et al. 2010; Mil-Homens and Fialho 2011). The type VI secretion system (T6SS) has been shown to play a role in *B. cenocepacia* pathogenicity by actin rearrangement of macrophages upon infection through modulating the function of Rho GTPases that result in actin cytoskeletal defects (Rosales-Reyes et al. 2012a). TecA, a T6SS effector, has been shown to induce deamidation of a conserved asparagine in Rho and Rac1 GTPase which results in inactivation of GTPase and disruption of host actin cytoskeleton (Aubert et al. 2016). In addition, type VI secretion system is also suggested to facilitate the escape of

T2SS-dependent secreted proteins, ZmpA and ZmpB, from the BcCVs into the cytoplasm of the infected macrophages (Rosales-Reyes et al. 2012b).

1.3 *Aeromonas hydrophila*

A. hydrophila is a rod-shaped Gram-negative bacterium with polar flagella that generally inhabits aquatic environments, including residual water, ground water, lakes and marine water, as well as in food, such as fish and meat. They are human pathogens in both immunocompromised and healthy people that can produce a variety of virulence factors that causes diseases ranging from gastroenteritis and skin infections to systemic infections, such as bacteraemia and necrotizing fasciitis. *A. hydrophila* are commonly found in the isolates of skin and soft-tissue infected tsunami survivors that are proposed probably to be the result of being exposed to contaminated fresh water (Hiransuthikul et al. 2005). It is also a well-known causative agent of fish diseases, including motile aeromonad septicaemia, red sore disease and ulcerative infections (Joseph and Carnahan 1994). In addition, *A. hydrophila* is resistant to many common antibiotics, such as amoxicillin (Altwegg and Geiss 1989; Saavedra et al. 2004).

There are several identified pathogenic determinants of *A. hydrophila* including LPS, exopolysaccharide capsules (Merino et al. 1996; Zhang et al. 2002), the surface S-layer protein (Dooley and Trust 1988), flagella, amonabactin-mediated iron-binding siderophore system (Stintzi and Raymond 2000), the *Aeromonas* cytotoxic enterotoxin (Act) (Ferguson et al. 1997; Xu et al. 1998; Galindo et al. 2006) and extracellular hydrolytic enzymes such as lipases and nuclease. The cytotoxic enterotoxin Act is a substrate of the T2SS that has been demonstrated to be important in *Aeromonas*-mediated infection by inducing the production of inflammation mediators in macrophages and human intestinal epithelial cells. It has also been shown to induce upregulation of apoptosis-associated genes using a human intestinal epithelial cell line (Xu et al. 1998; Galindo et al. 2006). *A. hydrophila* also possesses functional T3SS and Type IV secretion system (T4SS) that directly deliver effectors into the targeted host cell cytoplasm. An *A. hydrophila* bifunctional T3SS-dependent effector, AexU, was shown to promote actin cytoskeleton depolymerisation and exhibit cytotoxic ADP-ribosyltransferase (ADPRT) activity to host cell proteins (Braun et al. 2002; Ring et

al. 2002; Sha et al. 2007; Vilches et al. 2008). Hcp, an *Aeromonas* T6SS-dependent effector, was shown to induce apoptosis following caspase 3 activation once it was translocated into the host cell cytosol (Suarez et al. 2008). It has also been shown that Hcp paralyses macrophages preventing phagocytosis (Suarez et al. 2010b).

1.4 Protein secretion systems in Gram-negative bacteria

Gram-negative bacteria employ various machineries for transport of proteins from the interior of the cell to its exterior. These secreted proteins could remain associated with the bacterial outer membrane (OM), be released into the extracelluar milieu or be directly injected into a target cell cytosol (either eukaryotic or proteolytic) (Bleves et al. 2010; Costa et al. 2015). These machines are defined as secretion systems, and have key roles in bacterial virulence and bacterial adaption to the environment (Bleves et al. 2010; Rego et al. 2010; Costa et al. 2015). Unlike Gram-positive bacteria, which possess on an inner membrane (IM) found in Gram-negative bacteria, there are two membranes (IM and OM). Therefore, for Gram-negative bacteria, secretion involves translocation across the IM, periplasmic space and the OM of the cell envelope.

A number of systems are responsible for protein secretion by Gram-negative bacteria (Figure 1.2). These can be divided into two categories: one-step transport secretion across both the IM and the OM, and two-step secretion where proteins cross the OM only but rely on the Sec translocon (secretory) or Tat (twin-arginine translocation) pathways for transport of proteins across the IM into the periplasmic space. Four secretion systems have been identified that utilize a one-step secretion mechanism, which are the type I secretion system (T1SS), T3SS, T4SS and T6SS. In contrast, the two-step secretion systems include the T2SS, T5SS, T7SS (chaperone-usher pathway), T8SS and T9SS, which can transfer proteins across the OM from the periplasm to extracellular space. In addition, majority of the one-step secretion and two-step secretion systems exclusively secrete unfolded or partially folded proteins, whereas the T2SS, T6SS and T7SS are able to secrete folded or partially folded proteins (Costa et al. 2015).

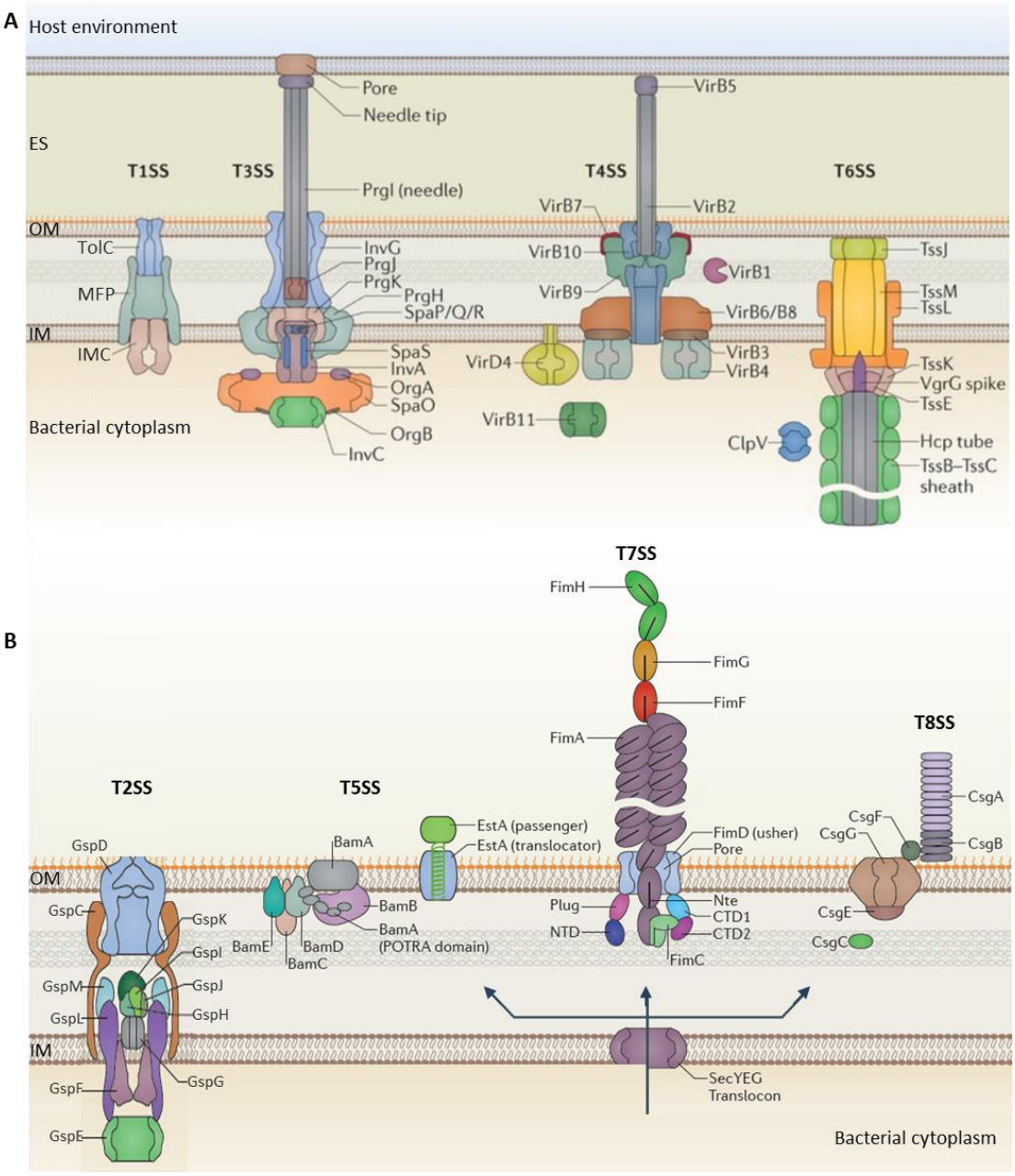


Figure 1.2 Structural organisation of the secretion systems of Gram-negative bacteria. Eight of the nine protein secretion systems are illustrated in this diagram. **A**. Secretion systems that span both inner and outer membranes and utilize a one-step secretion mechanism. ES, extracellular space. **B**. Secretion systems that possess two-step secretion mechanisms. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Microbiology (Costa et al. 2015), © 2015 (license number 3883041054835).

1.4.1 Sec and Tat pathways

In bacteria, many extracellular proteins are synthesised in the cytoplasm as preproteins with a signal peptide and are translocated through the IM into the periplasm using the Sec pathway or the Tat pathway. Proteins translocated across the IM by the Sec or Tat pathways have specific N-terminal signal sequences that guide the protein to the Sec or Tat system that located in the IM. The main functional difference between these two pathways is that the Sec pathway translocates polypeptides in an unfolded state, whereas the Tat pathway export proteins which are already folded (Palmer and Berks 2012). After translocation of the protein, the N-terminal signal sequence is cleavage off by the appropriate membrane-bound signal peptidase (Stathopoulos et al. 2000).

Proteins translocated across the IM that are recognized by the Sec-dependent pathway have a hydrophobic leader sequence of ~20-30 residues located at the N-terminus and are translocated in an unfolded state by co-translational or post-translational machineries (Papanikou et al. 2007). Nascent preproteins are targeted directly by the signal recognition particle (SRP) in a co-translational manner or the SecB chaperone in a post-translational manner (Randall and Hardy 2002; Luirink and Sinning 2004; Ullers et al. 2004). Although SRP is mainly employed by inner membrane proteins, there are exceptions of long, strongly hydrophobic signal peptides of nascent secretory proteins that are targeted by SRP (Schierle et al. 2003; Sijbrandi et al. 2003). The selection step for either SRP or SecB depends on the hydrophobicity and helicity of the emerging signal sequence, in which SRP binds the ribosome associated nascent protein chain when the emerging signal sequence displays a high level of hydrophobicity and helicity (Von Heijne 1985; Bruch et al. 1989; du Plessis et al. 2010). In both cases, the resulting SRP-preprotein and SecB-preprotein complexes are directed to the translocon complex SecYEG at the IM (Papanikou 2007). For SRP, this is achieved by docking to the membrane associated signal recognition particle receptor (FtsY), whereas for SecB, it is targeted by the cytosolic motor protein SecA (Papanikou et al. 2007; du Plessis et al. 2010). Although SecA does not contribute to the SRP-targeting route, when long hydrophilic segments are encountered, SecA is recruited to support their export (Neumann-Haefelin et al. 2000).

In the SecB-targeting route, SecA targets the preprotein-SecB binary complex and directs it to the SecYEG translocon (Schiebel et al. 1991). The binding of SecA to SecYEG and

ATP to SecA promotes the conformational change of SecA that leads to insertion of a hairpin loop of the signal sequence into the SecYEG translocon channel along with part of the mature domain of the preprotein. Therefore, SecA translocates ~2.5 kDa of the preprotein (Driessen et al. 1998). At the same time, SecB is dissociated from SecA (Engelman and Steitz 1981; Joly and Wickner 1993). Although this step only depends on ATP, the translocation rates can be enhanced by the proton-motive force (PMF), which is possibly involved in determining the correct orientation of the signal sequence within the channel (Economou 1998; du Plessis et al. 2010). Following translocation initiation, ATP hydrolysis releases one of the two SecA protomers from the preprotein and may weaken the binding affinity of the other SecA and SecYEG (du Plessis et al. 2010). SecA is then likely released from the SecYEG (Schiebel et al. 1991). Rebinding of SecA from the cytosolic pool to another region of the preprotein promotes the translocation of another approximately 2.5 kDa sequence through the channel (Schiebel et al. 1991; Van Der Wolk et al. 1997). Next binding of ATP to SecA would cause an additional translocation step, possibly another 5 kDa, and multiple rounds of rebinding and ATP binding and hydrolysis result in a stepwise translocation progress of the preprotein (du Plessis et al. 2010).

Furthermore, the adjoining SecDF membrane protein complex assists the later stages of protein translocation, which transiently interacts with the Sec translocon and is suggested to stimulate the stepwise translocation of the preprotein in a manner of PMF-dependent pulling on the translocating preprotein through the channel (Duong and Wickner 1997; Lycklama and Driessen 2012). Hence, translocation of the preproteins may depend on coupled forces: a driving force motored by SecA at the cytoplasm and a pulling force imposed by SecDF at the periplasm (Lycklama and Driessen 2012). Upon translocation through the IM and cleavage of signal peptides, the polypeptide chain folding initiates at the periplasmic side of the membrane (Nakamoto and Bardwell 2004; Mogensen and Otzen 2005; Gruber et al. 2006).

The Tat pathway is employed to export proteins which are already folded due to binding cofactor molecules and such proteins have hydrophobic N-terminal leader sequences that contain a conserved twin-arginine motif (Palmer and Berks 2012). In Gram-negative bacteria, protein transport using the Tat pathway is conducted by integral membrane proteins TatA, TatB and TatC, which can transport proteins with a range of diameters

(20-70 Å in *E. coli*) (Berks et al. 2000; Cline and McCaffery 2007). The Tat system is active and the energy source of it comes only from the PMF. In the system, TatB and TatC form an integral membrane complex consisting ~6-8 copies of each component (Bolhuis et al. 2001; Tarry et al. 2009). The transport cycles are initiated by the signal peptides of proteins binding to the TatBC complex (Cline and Mori 2001; Gérard and Cline 2007), in which the twin-arginine motif in the signal peptide is specifically recognized by a site in TatC and the remainder of the signal peptide and folded passenger domain are close to TatB (Alami et al. 2003; Gérard and Cline 2006; Maurer et al. 2010). Following this step, TatA protomers are recruited to the TatBC complex and polymerised into ring-like structures that require the PMF (Sargent et al. 2001; Gohlke et al. 2005). At this stage, the signal peptide is in contact with all the components of the resulting TatABC complex (Alami et al. 2003). The passenger domain of the substrate protein then translocates across the membrane through the polymerised TatA channel, and the signal peptide remains attached to the TatBC complex. When the passenger domain reaches the periplasmic side of the membrane, the signal peptide is then cleaved by a signal peptidase at the periplasmic side of the IM and TatA disassociates from the TatBC complex and depolymerises to free protomers (Palmer and Berks 2012).

1.4.2 Type I secretion system (T1SS)

The bacterial T1SS mediates the secretion of various protein substrates from the cytoplasm into the extracellular milieu, which are diverse in size and function, and are generally related to nutrient acquisition and pathogenesis (Kanonenberg et al. 2013). T1SS form a double membrane-spanning channel that consists of three major components, an ATP-binding cassette (ABC) transporter, a membrane fusion protein (MFP) and an outer membrane protein channel TolC (Figure 1.2). In the T1SS, the substrates are present in an unfolded state with a repeat glycine-rich motif at their carboxyl terminus that is recognised by the ABC transporter and are translocated from the cytoplasm across the IM into the periplasmic cavity of the MFP which is energized by ATP hydrolysis. Binding of the substrate to the ABC transporter enables the ABC-MFP complex to contact with the OM pore-forming complex (usually TolC), resulting in the unfolded substrates being released by TolC into the extracellular milieu (Holland et al. 2005; Kanonenberg et al. 2013; Costa et al. 2015).

1.4.3 Type II secretion system (T2SS)

Many Gram-negative bacteria, both pathogenic and non-pathogenic, employ the T2SS to secrete folded proteins from the periplasm, through the OM, into the extracellular environment. The T2SS substrates, which include hydrolysing enzymes and toxins, are first transported to the periplasm in the unfolded state by the Sec pathway or in the folded states by the Tat pathway (Nivaskumar and Francetic 2014). The T2SS consists of 12-15 components using the Gsp nomenclature that can be categorised into four groups: an IM platform, a pseudopilus, an OM complex and a cytoplasmic ATPase (Figure 1.2) (Korotkov et al. 2012; Costa et al. 2015). The IM platform is composed of at least four membrane proteins GspC, -F, -L and -M, that is connected to the OM complex through the binding of GspC to the periplasmic domains of GspD, a component of the OM complex (Korotkov et al. 2011; McLaughlin et al. 2012). The pseudopilus has structural similarities to the Type IV pilus and is comprised of GspG, -H, -I, -J and -K pseudopilins (Costa et al. 2015). The pseudopilins are proposed to be translocated into the periplasm in a Sec-dependent manner and their signal peptides are subsequently removed by GspO peptidase, except GspG, which remains inserted in the IM until it is extracted from the IM and polymerised into the pseudopilus (Korotkov et al. 2012; Costa et al. 2015). The OM complex is formed by the GspD oligomer and is termed the T2SS 'secretin', which is homologous to analogous components of the T3SS and Type IV pilus (Korotkov et al. 2012). GspE is a cytoplasmic ATPase that is recruited to the IM platform by interacting with the cytoplasmic regions of the IM complex components, GspL and -F (Py et al. 2001; Gray et al. 2011; Lu et al. 2013). The model of T2SS action, based on current understanding, is that the periplasmic region of GspC and/or the tip of the pseudopilus recruit the folded substrate from the periplasm (Costa et al. 2015). The binding of substrates stimulates the ATPase activity of GspE which is associated with the assembly of the pseudopilus on the periplasmic side of the IM platform. At the stage when the substrates and/or the tip of the growing pseudopilus contact with the secretin, the substrates are pushed out through the secretin channel by the growing pseudopilus and the channel is then blocked by the pseudopilus after the process (Hobbs and Mattick 1993; Shevchik et al. 1997; Korotkov et al. 2012; Costa et al. 2015).

1.4.4 Type III secretion system (T3SS)

The T3SS, termed the injectisome, is found in various pathogenic Gram-negative bacteria that target both plant and animal hosts (Tseng et al. 2009). The T3SS delivers unfolded bacterial effector proteins across the bacterial and host membranes into the cytoplasm of the target eukaryotic cells or inserted into its plasma membrane, resulting in modulation of host cell functions which facilitates bacterial invasion and colonization (Cornelis 2007; Buttner 2012). The T3SS consists of up to 25 proteins as a basal body that spans the IM, periplasm and OM and an external hollow needle structure protruding from the bacterial surface into the extracellular milieu (Figure 1.2). Seven major families of T3SS have been identified, two of which comprise Hrp1 and Hrp2 plant pathogens (Cornelis 2007; Tseng et al. 2009). The delivery of effector proteins requires chaperones to unfold the substrates during the export and direct the N-terminus of the substrate to the portal of the basal body. From here they are translocated via an uninterrupted conduit through the needle structure, with energy derived from T3SS ATPase (Yukihiro and Jorge 2005; Radics et al. 2014). The substrates are then translocated to the host cell membrane. Upon contact of the needle structure with the host cell, a pore is formed in the host cell membrane by three translocator proteins, named YopB, YopD and LcrV in *Yersinia* or PrpB, PopD and PcrV in *Pseudomonas*, the former two of which are hydrophobic and the latter one is hydrophilic (Mueller et al. 2008). The hydrophobic translocators anchor in the host cell membrane and form a translocation pore, which is facilitated by the hydrophilic translocator tip complex located at the distal end of the needle structure (Mueller et al. 2005). The T3SS is evolutionarily related to the flagellar protein export system, and they share functional and structural similarities (Desvaux et al. 2006).

1.4.5 Type IV secretion system (T4SS)

The T4SS functions as a one-step translocator of nucleoprotein or protein substrates from bacteria to the interior of bacteria or to eukaryotic cells. Based on the function of the T4SS, it can be categorised into three subfamilies. The first subfamily is employed for transferring plasmid DNA from one bacterium to another, which is defined as DNA conjugation, by exporting ssDNA (single-stranded DNA)-protein intermediates. This subfamily of T4SS includes the T4SSs encoded by *E. coli* pKM101 and R388 conjugative

plasmids, which facilitates the spread of plasmid-borne antibiotic resistance genes (Alvarez-Martinez and Christie 2009; Costa et al. 2015). The second subfamily is employed for secreting effector proteins to target eukaryotic cells, which is involved in bacterial pathogenesis, e.g. the A/B pertussis toxin of *Bordetella pertussis* (Burns 2003). This subfamily is present in the *Legionella pneumophila* Icm/Dot system which contains orhtologues of two T6SS components TssL (DotU) and TssM (IcmF) (Segal et al. 2005). The third subfamily of T4SS is responsible for the DNA release and uptake in bacteria that translocate DNA to or from the extracellular space, such as *Neisseria gonorrhoeae* (Ilangovan et al. 2015).

The best-characterised conjugative T4SS apparatus in the *E. coli* R388 plasmid is consisted of 12 components termed VirB1-11 and VirD4, and it assembles into a nanomachine complex that spans both membranes of Gram-negative bacteria (Figure 1.2) (Low et al. 2014). The T4SS contains an IM complex that is formed by association of Vir3, -6, and -8 with three ATPases VirB4, -11 and VirD4, which is connected to the core-OM complex constituted of VirB7, -9 and -10 through a central stalk structure of unknown composition. VirB10 is the hub of the translocation apparatus that inserts into both IM and OM and spans the entire periplasm which transmits the ATP-driven conformational change signal of cytoplasmic ATPase to the pore in the OM formed by it (Cascales and Christie 2004). A pilus structure consists of VirB2 and VirB5 that extends into the extracellular space with VirB5 forming the tip of the pilus. VirB1 functions as a lytic transglycosylase in the periplasm that degrades the peptidoglycan layer, which is proposed to be involved in pilus biogenesis (Baron et al. 1997; Zupan et al. 2007; Trokter et al. 2014). The exact mechanism for secreting unfolded protein substrates or conjugative DNA imtermediates remains to be elucidated. However, it has been proposed that T4SS switches between pilus biogenesis mode and a substrate translocation mode that might be controlled by VirB11 (Ripoll-Rozada et al. 2013; Trokter et al. 2014). In the pilus biogenesis mode, VirB11 associates with VirB4 that triggers the assembly of VirB2 to form a pilus. When the pilus tip binds to an unknown receptor on the target cell, VirB11 disassociates from VirB4 (Costa et al. 2015). In the substrate translocation mode, VirD4 is recruited to the translocation apparatus and VirB11 is subsequently bound to the VirD4, which facilitates the substrate translocation (Costa et al. 2015; Ilangovan et al. 2015).

1.4.6 Type V secretion system (T5SS)

The T5SS involves a two-step protein secretion mechanism that relies on the Sec pathway to translocate substrates across the IM, and the substrates can remain in an unfolded state or as semi-folded state to allow secretion through the OM (so called autotransporter) (Skillman et al. 2005; Costa et al. 2015). The proteins secreted using the T5SS may be involved in virulence, cell adhesion or biofilm formation (Leo et al. 2012). The T5SS does not require ATP hydrolysis or the PMF for substrate translocation, but instead, the energy comes from the folding of the passenger domain at the OM pore (Junker et al. 2009). There are five subclasses of the T5SS, termed T5SSa-e. The T5SSa sub-class contains classical autotransporters, to which many important virulence factors belong, such as IgA protease from *Neisseria meningitidis* (Mulks and Plaut 1978). The autotransporters consist of a secreted passenger domain and a transmembrane domain, known as translocator. The translocator domain is anchored into the OM as a β-barrel pore structure that allows the secretion of the passenger domain through it. T5SSb is also referred to as the two-partner secretion system (TPS), and substrates include filamentous haemagglutinin from *Bordetella pertussis* (Willems et al. 1994). T5SSa and T5SSb are closely related. However, in contrast to the T5SSa where the passenger and translocator components are produced as a single polypeptide, in T5SSb they are translated as two separate proteins referred to as TpsA and TpsB proteins, respectively. T5SSc is known as secreting trimeric autotransporter adhesions. T5SSd refers to as the fused two-partner secretion system, the passenger domain of which contains lipolytic activity that is cleaved autocatalytically after being autotransported. T5SSe is related to T5SSa but possesses an inverted order of passenger and transport domains allowing the passanger domain to be folded from its N- to C-terminus (Henderson et al. 2004; Leo et al. 2012).

The peptide precursors are directed to the Sec pathway by a signal peptide whereupon they are translocated into the periplasm. In the periplasm, precursors are maintained in the translocational state by periplasmic chaperones (e.g. Skp and SurA), which may also have a role in degrading the misfolded variants (Bodelón et al. 2009). Using T5SSa as an example, the C-terminal β-barrel of the precursor is subsequently targeted to the OM Bam complex, which facilities the insertion of all β-barrel OM proteins to form a pore in the

OM (Figure 1.2) (Knowles et al. 2009). The passenger domain of the precursor is then secreted through the OM from its C- to N-terminus into the extracellular milieu where the proteins are folded (Ieva and Bernstein 2009; Junker et al. 2009). However, in some cases, e.g. adhesins, the passenger domain is not cleaved by its peptidase domain but remains in contact with the OM (Costa et al. 2015).

Controversy still exists over the mechanism of secretion, in which the transfer of the passenger domain through the β-barrel is thought to occur by the interaction between the passenger domain and BamA POTRA (polypeptide transport associated) domains (Pavlova et al. 2013). Recent studies proposed that the translocation of the passenger domain may involve larger OM assemblies such as Bam and TAM (an alternative translocation and assembly module) complexes (Selkrig et al. 2012; Fabian et al. 2013; Pavlova et al. 2013).

1.4.7 Type VI secretion system (T6SS)

The T6SS is another secretion system employed by Gram-negative bacteria and constitutes a contractile phage tail-like apparatus which serves to translocate of effectors into the target prokaryotic or eukaryotic cells in a single step across the bacterial cell envelope (Figure 1.2). The T6SS gene clusters are present in one or more copies in a wide range of Gram-negative bacteria, including some pathogenic organisms such as *V. cholerae* and *P. aeruginosa*. The T6SS will be discussed in detail in Section 1.5.

1.4.8 Type VII secretion system (T7SS)

The T7SS present in the Gram-negative bacteria, corresponds to the chaperone-usher (CU) pathway, and is distinct from the ESX system (also known as T7SS) that is limited to mycobacteria (Simeone et al. 2008; Desvaux et al. 2009). The T7SS in Gram-negative bacteria is a two-step secretion system employed for assembly and secretion of polymeric cell surface appendages referred to as pili or fimbria in a wide range pathogenic Gram-negative bacteria, which mediate host recognition and attachment and are thereby employed for invasion and biofilm formation (Figure 1.2) (Wright et al. 2007). A well-studied pilus, termed the Type I pilus, is commonly used by uropathogenic

Escherichia coli (UPEC) strains for urinary tract colonisation and infection via invasion of human bladder epithelial cells (Martinez et al. 2000). The Type I pilus is composed of a rod structure formed by approximately one thousand copies of FimA, and a flexible tip structure comprised of FimF, -G and -H in which the adhesin FimH locates at the distal end of the pilus tip (Lillington et al. 2014). The pilus subunits are initially secreted by the Sec pathway into the periplasm in an unfolded state that are subsequently folded and stabilized by a periplasmic chaperone FimC (Hultgren et al. 1991; Vetsch et al. 2004). FimC not only functions as a typical chaperone in the periplasm by enhancing the efficiency of pilin to fold correctly, it also prevents premature polymerisation of pilin and stabilises them (Vetsch et al. 2004). Following which, the FimC-pilus subunit complexes are directed to the pilus assembly point at the OM, where FimD (also termed the usher) is located and forms a pore using its β-barrel domain in the OM (Nishiyama et al. 2003; Ng et al. 2004; Phan et al. 2011). The pilus subunits are then polymerised and secreted for the pilus assembly at the cell surface in a top to bottom fashion from the adhesin FimH to the rod component FimA (Thanassi et al. 2005).

1.4.9 Type VIII secretion system (T8SS)

Extracellular nucleation-precipitation (ENP) pathway, also termed the T8SS (Desvaus 2009), is an OM secretion pathway that is involved in the secretion and assembly of pili of the curli type on the surface of *Escherichia coli* and *Salmonella* spp. (Collinson et al. 1991; Collinson et al. 1996; Chapman et al. 2002). Curli are functional amyloid fibres that extend extracellularly from the bacterial OM and play a role in cell aggregation, bacterial adhesion and biofilm formation (Olsen et al. 1989; Vidal et al. 1998; Ryu et al. 2004; Kikuchi et al. 2005; Saldaña et al. 2009).

Curli formation involves two dedicated operons, *csgBAC* and *csgDEFG*, encoding GsgA-G subunits that rely on the Sec pathway to translocate across the IM into the periplasm (Figure 1.2). CsgD may function as a transcriptional activator of the *csgBAC* operon (Hammar et al. 1995). Curli fibres are comprised of the major curli subunit CsgA and its homologue GsgB, which differs from CsgA in the lack of conserved residues in R5, one of the five repeating units constituting the C-terminal amyloid core domain (named R1-R5), that are conserved in the other four repeating units. In fact, R5 of CsgB

contains residues that are essential for attaching CsgB to the cell surface (Wang et al. 2007; Hammer et al. 2012). Following secretion, CsgB nucleates the polymerisation of CsgA subunits into curli fibres (Hammar et al. 1996; Bian and Normark 1997; Chapman et al. 2002) Assembly and secretion of CsgA and CsgB through the OM is thought to involve an OM apparatus comprised of two soluble accessory subunits, chaperone-like GsgE and a nucleator protein CsgF, and an oligomeric lipoprotein CsgG located in the OM forming a hollow periplasmic porch that is in contact with the adaptor CsgE (Loferer et al. 1997; Chapman et al. 2002; Robinson et al. 2006; Nenninger et al. 2009; Nenninger et al. 2011). CsgE thereby forms a cage below the CsgG channel. CsgA and CsgB are translocated in an unfolded state through the CsgG secretion channel. The latest model proposes the unfolded substrates are entrapped in the CsgE-CsgG chamber complex, and the translocation of substrates to the cell surface through the CsgG secretion channel is a result of entropy gradient generated by local high concentration and conformational confinement of substrates in the chamber complex (Goyal et al. 2014). Currently, the function of CsgC in the T8SS is unknown.

1.4.10 Type IX secretion system (T9SS)

The Por secretion system, also referred to as the T9SS, is limited to members of the phylum *Bacteroidetes*. The T9SS is required for the gliding motility of some of the *Bacteroidetes* species by which they move along surfaces, such as *Flavobacterium johnsoniae* (McBride et al. 2009). Moreover, it is involved in the secretion of virulence factors, such as cysteine proteinases RgpA, RgpB and Kgp of the human periodontal pathogen *Porphyromonas gingivalis*, termed gingipains (Nakayama et al. 1995; Curtis et al. 1999). Proteins secreted by the T9SS possess conserved C-terminal domains that are proposed to be a common feature among the T9SS-secreted proteins (Seers et al. 2006; Sato et al. 2013; Veith et al. 2014). Secretion of these proteins involves the Sec pathway to translocate across the IM. There are at least ten subunits involved in the T9SS of *P. gingivalis*, which are PorK, -L, -M, -P, -Q, -U, -W, -X, -Y and Sov (Saiki and Konishi 2007; Sato et al. 2010). However, the structural organisation of these subunits is currently unknown and the mechanism by which substrates are translocated across the OM by this system remains a mystery.

1.5 Type VI secretion system (T6SS)

The T6SSs is one of the recently identified protein secretion systems. It exhibits high conservation and broad distribution in about a quarter of all Gram-negative bacterial species (Boyer et al. 2009). Many pathogens that possess a T6SS can cause serious infections in humans, e.g. *Vibrio cholerae, Yersinia pestis, Francisella tularensis, Burkholderia pseudomallei, Salmonella typhimurium* (Das et al. 2000; Folkesson et al. 2002; Gray et al. 2002; Das and Chaudhuri 2003; Nano et al. 2004; Shalom et al. 2007; Yen et al. 2008). The T6SS is a multiprotein complex spanning both IM and OM that translocates toxins into both prokaryotic and eukaryotic cells and thereby is involved in bacterial interactions and pathogenesis (Bladergroen et al. 2003; Pukatzki et al. 2007; Hood et al. 2010).

1.5.1 Discovery of the T6SS

The T6SS was named as a secretion system in 2006, although data from the mid-1990s had already suggested its existence (Pukatzki et al. 2006). In 1996, the haemolysin co-regulated protein (Hcp, a T6SS component named TssD) from *Vibrio cholerae* was identified as a secreted protein lacking the N-terminal hydrophobic signal peptide unlike other secreted proteins of this species (Williams et al. 1996). A year later, a chromosomal locus in *Rhizobium leguminosarum imp* was shown to negatively affect pea nodulation (Roest et al. 1997). In 1998, an infection-induced genetic locus of *Pseudomonas aeruginosa* (i.e. *ppkA*) encoding a serine/threonine protein kinase (STPK) was identified as being involved in virulence in neutropenic mice (Wang et al. 1998). Later on, the conserved gene clusters started to attract attention.

Pallen and colleagues observed that many species of pathogenic or symbiotic Gram-negative bacteria possess homologous gene clusters that included a gene encoding a forkhead-associated (FHA) domain, which is implicated in many bacterial processes, including regulation of cell shape, T3SS, and pathogenic and symbiotic host-bacterium interactions (Pallen et al. 2002). In the same year, a novel genomic island (*Salmonella enterica* centisome 7 genomic island, SCI) in *S. enterica* was demonstrated to be involved

in bacterial pathogenesis against eukaryotic cells, which included a conserved protein gene encoding SciS, a homologue of IcmF in *L. pneumophila*, a T4SS-associated component (Folkesson et al. 2002). In the following year, an *in silico* analysis identified an IcmF-associated homologous protein (IAHP) cluster in *V. cholerae* and nine other Gram-negative proteobacteria, in which a protein encoded by the gene cluster is a homologue of IcmF (Das and Chaudhuri 2003). In 2005, SciS was shown to control the intracellular replication and to be involved in virulence of *S. enterica* (Parsons and Heffron 2005). Later studies identified another gene in the conserved IAHP cluster, *icmH* (also termed *dotU*), which was also present in the *L. pneumophila* T4SS gene cluster. However, there was no evidence for the existence of a T4SS gene cluster in the IAHP-possessing *V. cholerae* strain V52. Subsequently, this conserved gene cluster that was responsible for protein secretion, including Hcp, and virulence, was proposed as the T6SS (Mougous et al. 2006; Pukatzki et al. 2006).

1.5.2 Genomic organization of T6SS gene clusters and its nomenclature

A variety of Gram-negative bacteria carry orthologues of the T6SS gene cluster (Figure 1.3). T6SS gene clusters comprise variable numbers of genes ranging from 16 in the *P. aeruginosa* HSI-2 gene cluster to 38 genes in *Serratia marcescens* (Murdoch et al. 2011; English et al. 2012); however, their gene order and orientation vary among species. A set of 13 'core genes' encoding components of the T6SS has been identified, known as *tssA-tssM* (type six secretion), leading to the proposal that the T6SS comprises 13 'core' components (Shalom et al. 2007; Zheng and Leung 2007; Boyer et al. 2009). The standardized 'Tss' nomenclature of the 13 core genes and the alternative nomenclatures proposed to describe these genes in a variety of bacterial species is shown in Table 1.1. A high level of co-occurrence of *tssJ*, *tssK*, *tssL* and *tssM* is observed in the majority of T6SS gene clusters, with six of the remaining nine genes co-occurring as groups of three, *tssB-tssC-tssD* and *tssE-tssF-tssG* (Boyer et al. 2009). Apart from these core genes, T6SS gene clusters commonly contain accessory genes, termed *tags* (T6SS associated genes). However, the function of many of these Tag proteins is currently unknown.

Many Gram-negative bacterial species possess one or two T6SSs. However, four to six clusters were found within a single genome in a few species, such as *Yersinia pestis* (four

copies) and *Burkholderia pseudomallei* (six copies) (Shalom et al. 2007). And in some cases, such as *P. aeruginosa* PAO1, it has been clearly shown that the three T6SS gene clusters did not arise by duplications, but instead they were acquired by horizontal gene transfer (Bingle et al. 2008). It is noticeable that T6SS clusters generally exist in pathogenicity islands in many bacterial species, i.e. the enteroaggregative *Escherichia coli* (EAEC) *pheU* island, *S. typhimurium* SCI (*Salmonella* centisome island), and the *P. aeruginosa* HSI (Hcp-secretion island) (Folkesson et al. 2002; Dudley et al. 2006; Mougous et al. 2006).

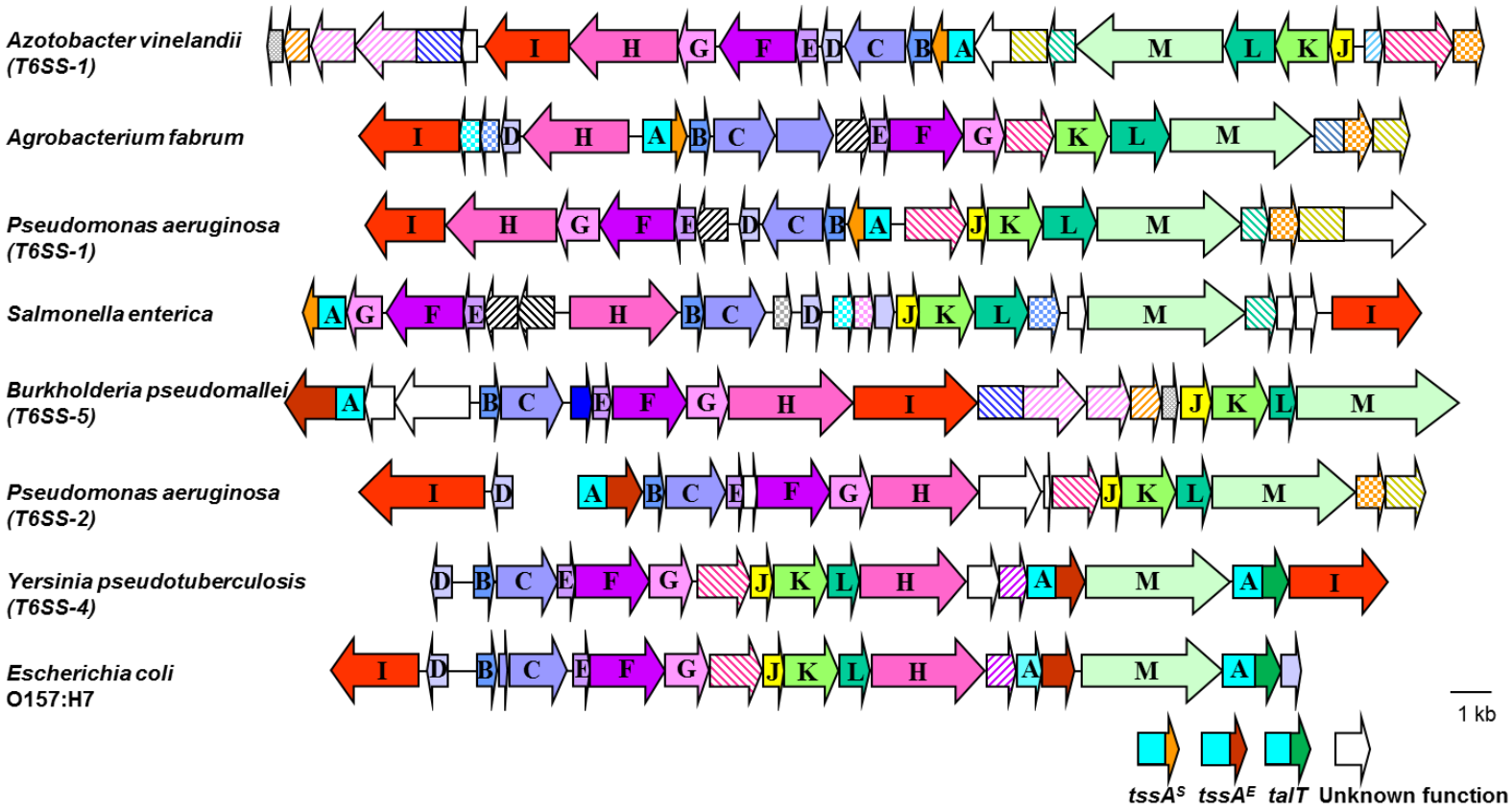


Figure 1.3 Orthologous T6SS gene clusters present in Gram-negative bacteria. Genes in each T6SS gene clusters are represented as block arrows. Direction of the arrows represents gene orientation. Orthologous genes are designated using the same colour. 'Core' T6SS genes with uniformly colours are labelled with single letters corresponding to relative *tss* genes; T6SS associated genes (*tags*) are represented as hatched arrows; genes with unknown function are shown as white arrows. *A. vinelandii* T6SS-1 is one of two *P. aeruginosa* T6SS-1 and T6SS-2 correspond to *P. aeruginosa* T6SS HSI-I and HSI-II (two of three T6SS gene clusters). *B. pseudomallei* T6SS-5 is the animal pathogenicity T6SS of *B. pseudomallei* (one of the six T6SS gene clusters), *Y. pseudotuberculosis* T6SS-4 is one of four T6SS gene clusters in this organism. Other Gram-negative bacteria presented here have a single T6SS gene cluster. Sun et al, submitted for publication.

Table 1.1 Standardisation of nomenclature for the T6SS

Standardised gene name	Corresponding gene name in other bacteria									
	S. typhimurium	EAEC(sciI)	*V. cholerae*	EAEC(sciII)	*E. tarda*	*P. aeruginosa* (HSI-I)	*P. aeruginosa* (HSI-II)	*R. leguminosarum*	*B. mallei*	*B. cenocepacia*
tssA	*sciA*	*sciA*	*vasJ*	*aaiJ*	*evpK*	*hsiA1*	*hsiA2*	*impA*	*bimE*	*bscE*
tssB	*sciH*	*sciH*	*vipA*	*aaiA*	*evpA*	*hsiB1*	*hsiB2*	*impB*	*tssA*	*bscL*
tssC	*sciI*	*sciI*	*vipB*	*aaiB*	*evpB*	*hsiC1*	*hsiC2*	*impC/D**	*tssB*	*bscK*
tssD	*sciK/M**	*sciD*	*hcp*	NP	*evpC*	*hcp1*	*hcpC*	NA	*hcp1*	*bscJ*
tssE	*sciD*	Ec042_4545	VCA0109	*aaiD*	*evpE*	*hsiF1*	*hsiF2*	*impF*	*tssC*	*bscI*
tssF	*sciC*	*sciC*	*vasA*	*aaiE*	*evpF*	*hsiG1*	*hsiG2*	*impG*	*tssD*	*bscH*
tssG	*sciB*	*sciB*	*vasB*	*aaiF*	*evpG*	*hsiH1*	*hsiH2*	*impH*	*tssE*	*bscG*
tssH	*sciG*	*sciG*	*vasG*	*aaiP*	*evpH*	*clpV1*	*clpV2*	NA	*clpV1*	*bscF*
tssI	*vrgS*	*vgrG*	*vgrG*	*aaiG*	*evpI*	*vgrG1*	*vgrG2*	NA	*vgrG1*	NP
tssJ	*sciN*	*sciN*	*vasD*	*aaiK*	*evpL*	*lip1*	*lip2*	NP	*tssJ*	*bscN*
tssK	*sciO*	*sciO*	*vasE*	*aaiL*	*evpM*	*hsiJ1*	*hsiJ2*	*impJ*	*tssK*	*bscO*
tssL	*sciP*	*sciP*	*vasF*	*aaiN*	*evpN*	*dotU1*	*dotU2*	*impK*	*tssL*	*bscP*
tssM	*sciS*	*sciS*	*vasK*	*aaiO*	*evpO*	*icmF1*	*icmF2*	*impL*	*icmF1*	*bscB_*

NA, no alternative name given; NC, gene is present at another locus outside the *tss* unit; *, two similar or identical orthologues are present in the same *T6SS* gene culster; EAEC, enteroaggregative *E. coli*. Adapted from (Shalom et al. 2007).

1.5.3 Structure and mechanism of action of the T6SS components

The current model of the T6SS apparatus is based on evolutionary, structural and mechanistic similarities between several structural proteins of the T6SS and the phage gene products (gp) that make up the injection machinery of some contractile bacteriophages (Figure 1.4). The T6SS consists of multiple subunits that are organised into a syringe-like structure containing two distinct sub-complexes, a membrane chamber complex that spans the cell envelope and anchors a contractile bacteriophage tail-like complex to the bacterial envelope. There are at least 13 genes present in the T6SS gene cluster required to assemble a functional T6SS (named *tssA-tssM*), and the gene products either function as structural components required for the proper assembly of the secretion apparatus and/or are secreted (Shalom et al. 2007; Zheng and Leung 2007; Aschtgen et al. 2008; Bingle et al. 2008; Cascales 2008). In some cases, T6SS require a Tag protein (TagL, TagN, TagP or TagW) which contains a peptidoglycan-binding (PG) domain to stabilise the apparatus in the cell envelope (Aschtgen et al. 2010b). The requirement for one of these Tag proteins occurs when TssL, a core component of the membrane complex, lacks the PG domain at its C-terminus.

In the current model of the T6SS (Figure 1.5), assembly initiates from the formation of the membrane chamber complex followed by assembly of a baseplate that serves as a platform for polymerisation of a tubular structure in the bacterial cytoplasm. The tubular structure is comprised of an inner tube tipped by a needle spike (tail spike) and wrapped by a contractile sheath structure. During the apparatus assembly, effectors delivered by the T6SS are loaded either into the inner tube or onto the tail spike. Contraction of the sheath pushes the inner tube and spike across the cell envelope through the membrane chamber complex that subsequently punctures the target host cell membrane and deliveries the effectors. After contraction, the sheath is disassembled and recycled by TssH, a dedicated ATPase encoded by the T6SS gene cluster (Bönemann et al. 2009; Basler and Mekalanos 2012; Basler et al. 2012; Kapitein et al. 2013).

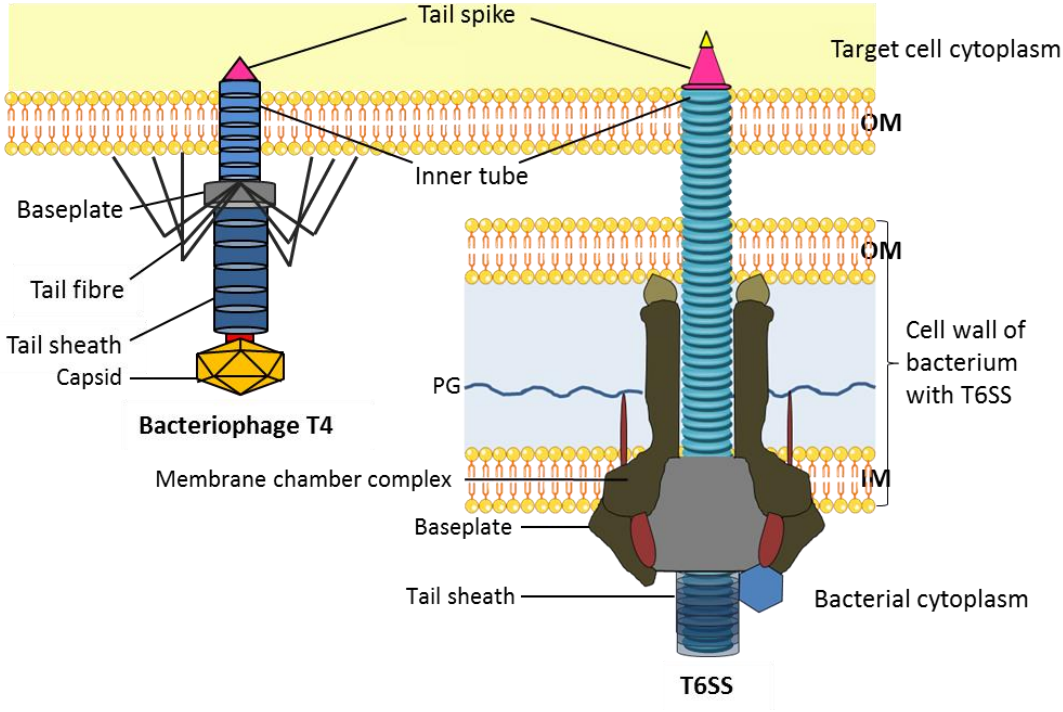


Figure 1.4 Similar injection mechanisms of bacteriophage T4 and T6SS. Contracted state of phage T4 and T6SS are illustrated and structurally similar components in bacteriophage T4 and T6SS are drawn in the same colour. OM, bacterial outer membrane; IM, bacterial inner membrane; PG, peptidoglycan.

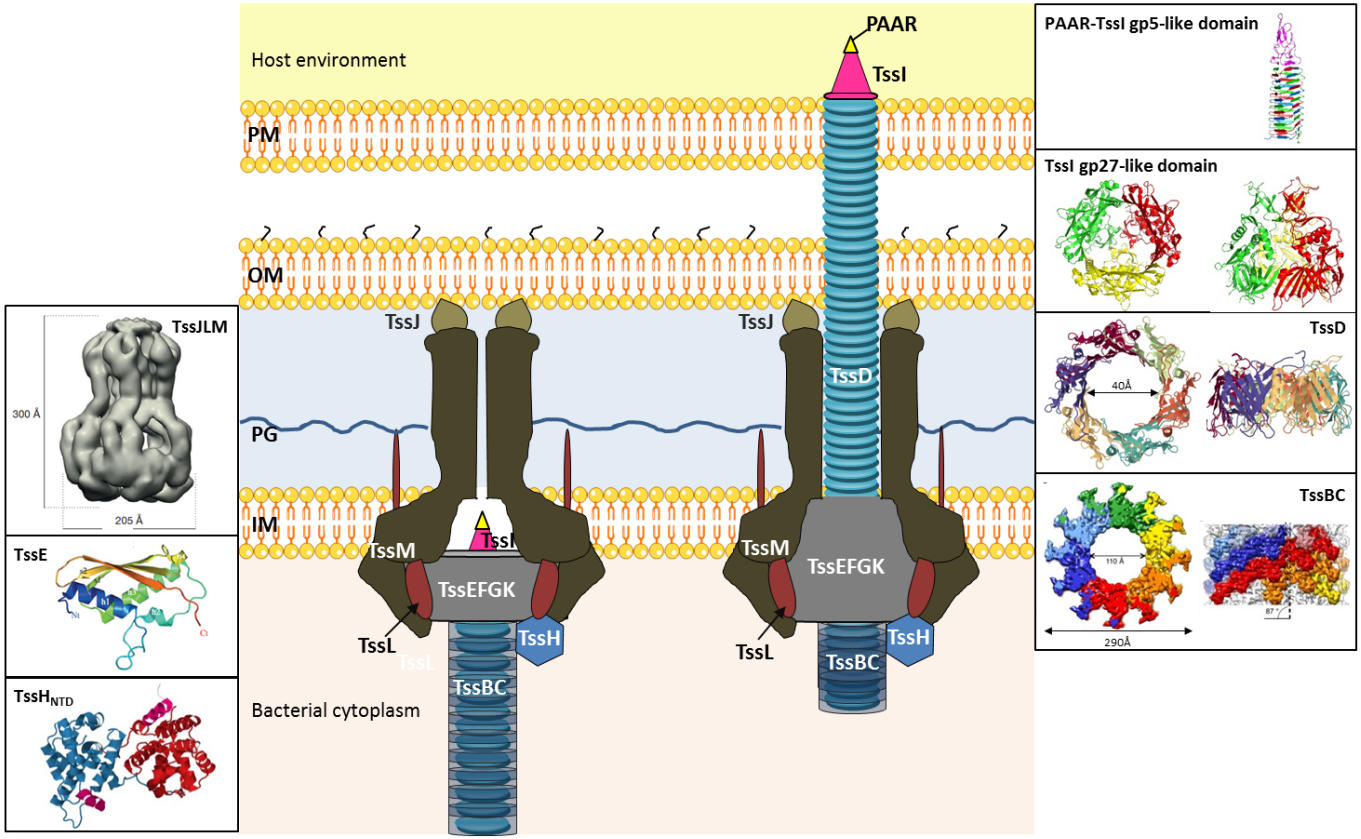


Figure 1.5 Diagrammatic representation of the T6SS apparatus. The diagram shows the extended state (left) and contracted state (right) of the T6SS apparatus. The apparatus is organised into a syringe-like structure that is embedded in the cell membrane by a membrane chamber sub-complex formed by TssJ, TssM and TssL, which spans the entire membrane and anchors a contractile bacteriophage tail-like complex to the bacterial envelope. The contractile phage tail-like complex is made up with a baseplate (TssE, TssF, TssG, TssK and a TssI spike sharpened with PAAR protein), a tail tube (TssD) and a tail sheath (TssB and TssC). The tail tube is surrounded by the cytoplasmically located tail sheath in an extended (resting) state which is then fired through the cell envelope by contraction of the sheath. TssJLM (11.6 Å negative-stain electron microscopy structure, reprinted by permission from Macmillan Publishers Ltd: Nature (Durand et al. 2015), © 2015 (license number 3883050177742; TssE (built model based on bacteriophage gp25, reprinted by permission from The Royal Society: Philosophical Transactions B (Cascales and Cambillau 2012) © 2012 (license number 3883090805309)); TssH$_{NTD}$ (PDB 3ZRJ (Pietrosiuk et al. 2011)); PAAR-TssI gp5-like domain (PDB 4JIW (Shneider et al. 2013)); TssI gp27-like domain (PDB 2P5Z (Leiman et al. 2009)); TssD (PDB 1Y12 (Mougous et al. 2006)); TssBC (6 Å cryo-electron microscopy structure, reprinted by permission from Cell Reports (Kube et al.) ©2014 (license number 3883170973305)). PM, plasma membrane; OM, bacterial outer membrane; IM, bacterial inner membrane; PG, peptidoglycan

1.5.3.1 The membrane chamber complex

A cell envelope anchored sub-complex is required to attach the contractile phage tail-like structure to the membrane and serve as a conduit to allow the passage of the tail tube/spike out of the bacterium upon sheath contraction. The membrane core complex is comprised of an OM lipoprotein TssJ and two IM embedded subunits TssL and TssM that assemble through a network of interactions between TssL and TssM, and TssM and TssJ (Aschtgen et al. 2010a; Felisberto-Rodrigues et al. 2011; Coulthurst 2013; Ho et al. 2014; Zoued et al. 2014; Durand et al. 2015).

The crystal structure of TssJ, TssL and C-terminal portion of the TssM periplasmic region on its own or with TssJ have been solved (Felisberto-Rodrigues et al. 2011; Rao et al. 2011; Durand et al. 2012b; Robb et al. 2012; Robb et al. 2013; Chang and Kim 2015; Durand et al. 2015). TssJ is anchored to the OM by the N-terminal lipid moiety and extends into the periplasm (Aschtgen et al. 2010a). TssL and TssM are homologues of T4SS components IcmH/DotU and IcmF, respectively. TssM is anchored in the IM by three transmembrane helices, two of which are located at the N-terminus and the third is located at the C-terminus of the cytoplasmic N-terminal domain which corresponds to approximately one third of the length of the protein. The remaining two thirds of TssM extend into the periplasm with the C-terminal end coming in contact with TssJ (Durand et al. 2015). There is a large cytoplasmic domain between the second and third transmembrane domain that carries Walker A and B motifs in a subset of TssM orthologues (Zheng and Leung 2007). The Walker B motif of TssM is required for ATPase activity (Ma et al. 2012). However, the Walker A motif of TssM orthologue in *E. tarda* is unnecessary for T6SS function, but it is required for T6SS function in *A. fabrum* by modifying the conformation of TssM and recruiting the tail tube subunit to the membrane complex (Zheng and Leung 2007; Ma et al. 2009b; Ma et al. 2012). The bulk of TssL is located in the cytoplasm and is anchored in the IM by a C-terminal transmembrane segment. The cytoplasmic domain forms a hook-like structure composed of two bundles of three α-helices, and is enriched in negatively charged residues. As it is conserved among TssL orthologues, it was suggested that it may recruit its partners (Robb et al. 2012). Several TssL orthologoues carry a C-terminal extension with a periplasmic PG motif which has been shown to be required for stabilising the membrane

complex in the cell envelope (Aschtgen et al. 2010a; Aschtgen et al. 2010b; Durand et al. 2012b). While the PG-binding motif is absent from TssL encoded by some T6SS clusters, an accessory protein is usually encoded by the T6SS gene cluster which carries this motif, e.g. TagL (Aschtgen et al. 2010a).

Recently, the structure of the TssJLM membrane complex of EAEC has been solved at ~12 Å by negative stain electron microscopy, which has a five-fold symmetry containing 10 copies of each core subunit forming an overall rocket-shape structure. The membrane complex spans the entire cell envelope and has a narrow inner channel of 15-20 Å that is suggested to exist during the resting (non-firing) state of the T6SS apparatus. It is proposed that the membrane complex undergoes conformational modifications to allow the passage of the tail spike and tube with ~110 Å in diameter in the contracted state. This is supported by a previous observation that TssM undergoes large conformational changes during secretion (Ma et al. 2012). Recent results have proposed that the assembly of the TssJLM membrane complex involves sequential addition of TssJ, TssM and then TssL, that then initiates T6SS apparatus formation towards the bacterial cytosol (Durand et al. 2015).

1.5.3.2 The contractile phage tail-like complex

The mechanism of the T6SS involves puncturing the target cell membrane in a manner that is similar to that of contractile bacteriophages, which have a cell-puncturing machinery to inject their DNA into the target bacterial cytosol. The minimal composition of a bacteriophage contractile injection system contains the conserved part of the wedges that are made up of the inner baseplate (corresponding to gene products, gp6, gp7, gp25 and gp53 in phage T4), a tail tube (gp19, gp48 and gp54) surrounded by the conserved part of the tail sheath (gp18) and a central spike complex (gp5, gp5.4 and gp27) at the tip of the tail tube, which act as a cell-puncturing apparatus to puncture the bacterial cell envelope (Kanamaru et al. 2002; Leiman et al. 2004; Kostyuchenko et al. 2005; Taylor et al. 2016). A large amount of data supports the model that the contractile phage tail-like complex of T6SS composed of analogous baseplate, tail tube and tail sheath components. The discovery of the structural similarities and/or bioinformatics conservation between some T6SS components and bacteriophage T4 proteins has been a

major breakthrough to progress the study of the T6SS (Leiman et al. 2009; Basler et al. 2012).

1.5.3.2.1 The T6SS baseplate

In contractile bacteriophage such as T4, the tail tube and sheath assemble on a platform termed the baseplate, which undergoes a large conformational change from a hexagon shape to a six-pointed star during contraction (Kostyuchenko et al. 2003; Leiman et al. 2004; Rossmann et al. 2004; Kostyuchenko et al. 2005; Leiman and Shneider 2012; Taylor et al. 2016). An analogous baseplate-like structure was observed by electron cryotomography of the T6SS in which a flared bell-shaped base connected the T6SS tail sheath to the membrane. This structure was observed to undergo a large conformational modification during sheath contraction (Basler et al. 2012). The bacteriophage T4 baseplate is comprised of at least 15 different proteins with variable copy numbers from 1 to 18 that are organised into a central spike complex, inner part, intermediate part and peripheral part (constituting the tail fibre) of the baseplate (Figure 1.6). The inner baseplate, comprising gp6, gp7, gp25 and gp53, is universally conserved in all the contractile injection systems. The T6SS core components TssF, TssG and TssE are orthologues of T4 gp6, gp7 and gp25, respectively. A gp53 orthologue is absent in the T6SS but may constitute a domain in TssG (Brunet et al. 2015; Taylor et al. 2016). In addition, the T6SS core components TssI and TssK are homologous to the T4 central spike components $(gp27\text{-}gp5)_3$ and a T4 baseplate peripheral component gp10, respectively (Leiman et al. 2009; Taylor et al. 2016). These five core components, TssE, -F, -G, -K and -I are believed to form the T6SS baseplate prior to be recruited to the membrane chamber complex (Leiman and Shneider 2012; Brunet et al. 2015). TssK is proposed to connect the baseplate to the membrane complex by interacting with the cytoplasmic region of TssL and TssM, which is stabilized by an additional interaction between TssM and the TssFG complex (Brunet et al. 2011; Durand et al. 2015).

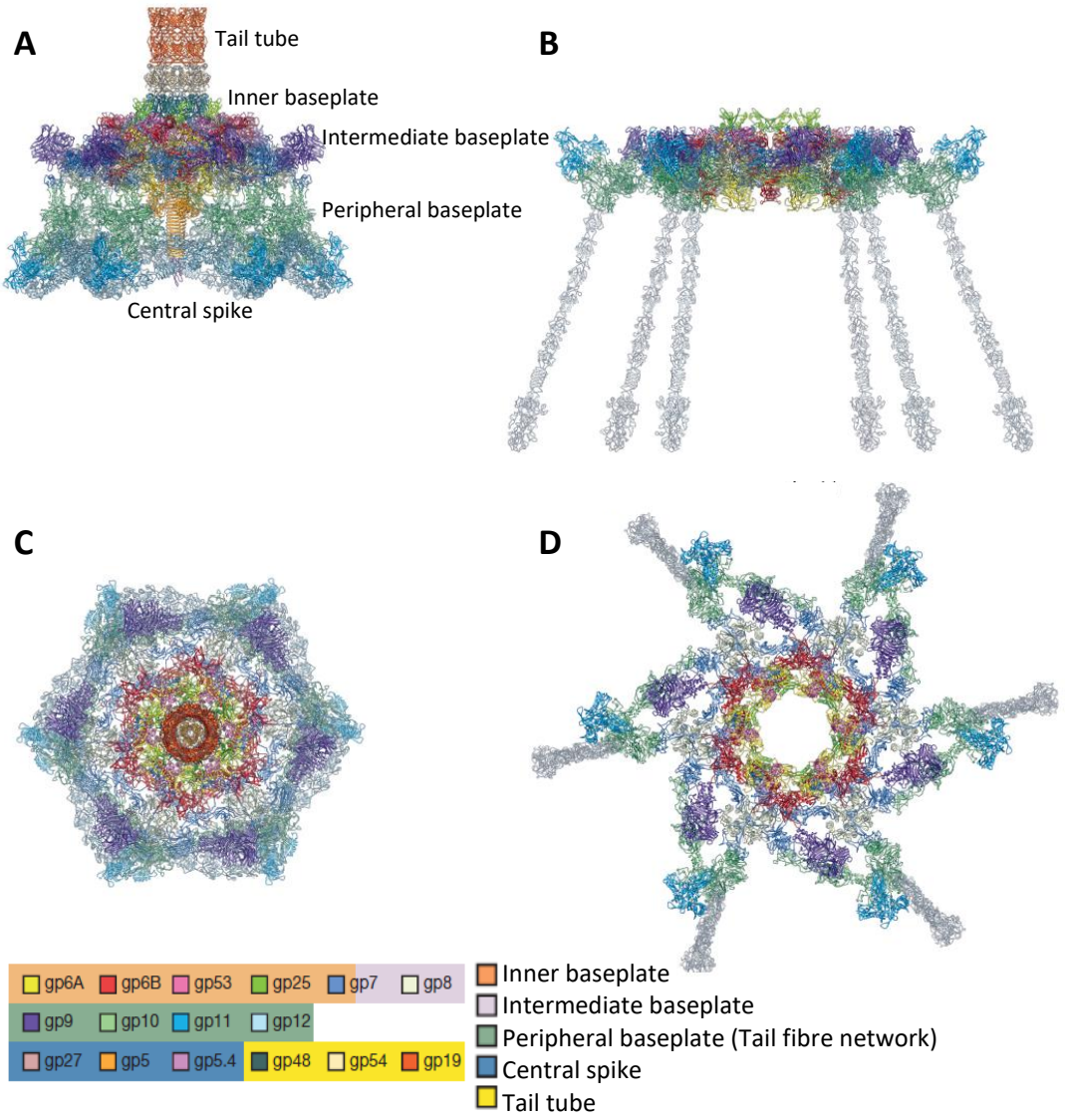


Figure 1.6 Atomic models of the bacteriophage T4 baseplate in extended and contracted states. Protein components are shown in colours as indicated. A and C are side and bottom views, respectively, of the baseplate when the contractile tail is in the extended state prior to attachment of the bacteriophage to the host cell wall; B and D are side and bottom views, respectively, of the contracted state that occurs after attachment to the host cell wall. Reprinted by permission from Macmillan Publishers Ltd: Nature (Taylor et al. 2016) © 2016 (license number 3883071436602).

1.5.3.2.1.1 TssE

TssE is a baseplate component of the T6SS, as it shares a high degree of similarity (~40%) with a bacteriophage T4 baseplate wedge component gp25 (Shalom et al. 2007; Bingle et al. 2008; Cascales 2008; Leiman et al. 2009; Lossi et al. 2011). It is proposed that gp25 triggers the polymerisation of tail sheath by interacting with the proximal coil row of the tail sheath subunit gp18 (Leiman and Shneider 2012). Currently, the crystal structure of TssE has not been solved, but a model of it was built based on the reported structure of gp25 due to the high degree of amino acid sequence conservation between them (Cascales and Cambillau 2012). A similar role for TssE in T6SS biogenesis is proposed based on the observations that TssE is present in the bacterial cytosol where it interacts with the tail sheath and is required for sheath polymerisation (Lossi et al. 2011; Basler et al. 2012; Brunet et al. 2013; Kapitein et al. 2013).

1.5.3.2.1.2 TssF and TssG

Currently, there is no structural information available for TssF and TssG. High levels of co-organisation between *tssE* and *tssF* (~87%), and *tssF* and *tssG* (~97%) in T6SS gene clusters was recognised, suggesting protein-protein interactions among the three encoded proteins (Boyer et al. 2009). Indeed, the interactions between TssF and TssG with each other and with TssE have been demonstrated. In addition, TssF and TssG have been shown to stabilise each other (Brunet et al. 2015). It has been proposed that TssF and TssG are T6SS baseplate components due to their homology to the conserved inner baseplate components of the phage T4 contractile injection system gp6 and gp7, respectively (Taylor et al. 2016). In agreement with the bacteriophage conserved inner baseplate structure constituted of gp25, gp6, gp7 and the tail spike complex $(gp27\text{-}gp5)_3$, the orthologous T6SS proteins TssE, TssF, TssG and TssI could be purified as a complex from enteroaggregative *E. coli* (EAEC) (Brunet et al. 2015).

Other than TssE, the TssFG complex also interacts with many other T6SS core components, such as TssC, -D, -K, -I and the cytoplasmic domain of TssM. TssF and TssG contacts with TssK and TssI require the pre-assembly of the TssFG complex. These results support the idea that the T6SS baseplate is constituted of TssE, TssF, TssG, TssK

and TssI (Brunet et al. 2015). In addition, TssF and TssG have been shown to play a role in proper assembly of the tail tube (Brunet et al. 2015).

1.5.3.2.1.3 TssK

Very little was known about TssK when this study started except its conserved genomic co-occurrence with genes encoding the membrane core complex, i.e. *tssJ*, *tssM* and *tssL* (Boyer et al. 2009). TssK was shown to co-precipitate with the TssBC tail sheath by interacting with TssC in 2012 (Basler et al. 2012). A year after, Zoued and colleagues showed TssK is a cytoplasmic protein oligomerising into a three-armed structure, and is required for tail tube/sheath assembly (Zoued et al. 2013; Brunet et al. 2015). Consistent with the co-occurrence of *tssK* with genes encoding the membrane complex, TssK was shown to interact with the cytoplasmic domain of TssL (membrane complex component) (Zoued et al. 2013). Apart from that, TssK was also demonstrated to interact with TssD (tail tube component), TssC (sheath component), TssF (baseplate component), TssG (baseplate component) and TssM (membrane complex component) (Zoued et al. 2013; English et al. 2014). Therefore, TssK is proposed to be a baseplate component based on its cellular localisation and interaction network with other T6SS core components (Zoued et al. 2013). More recently, TssK was shown to have a similar function to the T4 baseplate component gp10 in initiation of T6SS baseplate assembly (Taylor et al. 2016).

1.5.3.2.1.4 TssI

The trimeric secreted subunit TssI (also known as VgrG) shares structural similarities with the spike of bacteriophage T4 which is made up of a trimer of the gp27-gp5 heterodimer functioning as a cell-puncturing device. The bacteriophage T4 spike constitutes a β-prism formed by the C-terminal domain of gp5, an adaptor formed by gp27, and the N-terminal region of gp5 that connects to the tail tube (Kanamaru et al. 2002; Pukatzki et al. 2007; Leiman et al. 2009). TssI orthologues contain conserved N-terminal and central regions, in which the two conserved regions are structural homologues of the gp27 and gp5 proteins, respectively, but lack the lysozyme domain possessed by gp5 for lysing the cell wall (Leiman et al. 2009; Uchida et al. 2014). The

lysozyme domain of vibriophage KVP40, that belongs to the T4 phages, is also missing from gp5 and is proposed to be encoded somewhere else in the genome (Rossmann et al. 2004). Interestingly, two effectors secreted by the T6SS in *P. aeruginosa*, Tse1 and Tse3, exhibit peptidoglycan hydrolase activity that are delivered into the periplasm of the target bacteria (Russell et al. 2011). Therefore, in such cases, a lysozyme component is not necessary.

Analogous to the bacteriophage tail structure and assembly mechanism, the T6SS tail tube composed of TssD (see below) is proposed to assemble to the base of the trimeric TssI and is involved in the outgoing force of TssI as a result of tail sheath contraction. Puncture of the target cell is facilitated by the needle shaped β-helical domain of TssI (gp5-like region). This model is supported by the interdependent release of the tail tube component TssD and TssI in the culture supernatant, and the interaction of TssD with TssI and the fact that TssI is required for tail tube assembly (Pukatzki et al. 2007; Zheng and Leung 2007; Lin et al. 2013; Brunet et al. 2014). In many cases, TssI also carries one or more additional domains at the C-terminus of the gp5-like β-helix which may function as effector domains delivered into the target cell cytosol. These forms of TssI are termed 'specialised TssI' or 'evolved VgrG' (Pukatzki et al. 2007; Ma et al. 2009a; Pukatzki et al. 2009).

Recent studies have suggested that the T6SS spike may be sharpened by an additional protein with proline-alanine-alanine-arginine repeat, named PAAR, forming a sharp conical extension capping the needle end of TssI spike (Shneider et al. 2013). This is in agreement with the observation that many bacteriophage spikes are decorated with such a protein (gp5.4) that sharpens the spike as well (Browning et al. 2012). PAAR is therefore secreted by the T6SS and is also proposed to attach effectors to the TssI spike as many PAAR proteins possess putative effector domains as the N- or C-terminal extensions (Koskiniemi et al. 2013; Shneider et al. 2013).

1.5.3.2.2 The T6SS tail tube

The T6SS tail tube is constituted of TssD (Hcp), which is a small T6SS component that shares common evolutionary origin with the bacteriophage T4 tail tube protein gp19 and has structural similarity with the tail tube protein gpV of the bacteriophage λ (Leiman et

al. 2009; Pell et al. 2009a). TssD (Hcp) was initially discovered in *V. cholerae* as a secreted protein, and later it was also shown to be secreted in *P. aeruginosa* and is present in the sputum of *P. aeruginosa*-infected cystic fibrosis patients (Mougous et al. 2006). However, Zheng and colleagues demonstrated that the absence of TssD abolishes the activity of the T6SS (Zheng and Leung 2007), that led to a conclusion that TssD is not only a secreted protein, but also functions as a structural component of the T6SS.

TssD subunits assemble spontaneously into a hexameric ring structure *in vitro* with an outer diameter of 85-90 Å and an internal diameter of 35-40 Å (Mougous et al. 2006; Ballister et al. 2008), whereas it is observed to form a tubular structure by stacking hexamers on each other in a head-to-tail manner *in vivo* (Brunet et al. 2014). This suggests proper assembly of the tail tube requires other T6SS subunits. Later studies showed that the tail spike protein TssI located at the tip of the tail tube, along with other baseplate components, is required for the polymerisation of TssD hexamers (Brunet et al. 2014; Brunet et al. 2015). The TssD hexamers are polymerised in the orientation from the IM to the cytoplasm, but instead of directly contacting with the IM, it is connected to the IM by the T6SS baseplate complex, suggesting that formation of tail tube initiates from the baseplate (Basler et al. 2012). These results support the idea that the T6SS baseplate serves as a platform for tail tube assembly. In addition, the tubular structure is suggested to be wide enough to allow passage of small folded proteins, as well as unfolded or partly folded proteins (Mougous et al. 2006).

1.5.3.2.3 The T6SS tail sheath

In bacteriophage T4, the tail tube is wrapped by a contractile tail sheath that is a polymer of gp18 protein. In fact, the tail tube acts as a scaffold during sheath polymerisation (King 1968; Leiman et al. 2010). An analogous tail sheath structure is also found in the T6SS, almost perpendicular to the cell envelope and extends into the bacterial cytosol, in which two proteins, TssB and TssC, assemble the tail sheath. The sheath exists in two conformations, likely associated with the extended and contracted sheath structures in the cytoplasm. It has further been shown that the T6SS tail sheath exhibits a dynamic cycle between assembly, quick contraction, disassembly and re-assembly (Bönemann et al. 2009; Basler et al. 2012; Brunet et al. 2013; Kapitein et al. 2013; Zoued et al. 2013). This

provides further evidence of similarities between the contractile tail structures in both bacteriophage T4 and T6SS.

In the T6SS, co-occurrence of *tssB* and *tssC* genes are observed in all the T6SS gene clusters in the same orientation and the corresponding encoded proteins have been shown to interact and stabilise each other (de Bruin et al. 2007; Bönemann et al. 2009; Broms et al. 2009b; Aubert et al. 2010; Karna et al. 2010; Lin et al. 2013; Lossi et al. 2013; Zhang et al. 2013). TssB and TssC assemble spontaneously into a cylindrical complex *in vitro*, with six TssBC heterodimers that form a cogwheel-like structure exhibiting a 12-fold symmetry. The cylindrical complex formed by TssBC has an internal and external diameter of ~100 Å and ~300 Å, respectively, corresponding to the conformation of the contracted state (Bönemann et al. 2009; Basler et al. 2012; Lossi et al. 2013; Kube et al. 2014). By analogy to the bacteriophage tail sheath, the inner diameter of TssBC cylinder is sufficient to accommodate the TssD tail tube (~90 Å), suggesting that the tail sheath surrounds the tail tube in the T6SS. This is also supported by the extra density that is observed inside the extended TssBC sheath by cryo-tomography. Moreover, TssB interacts with TssD *in vivo* with low affinity which is probably due to the tail tube sliding within the sheath during contraction (Basler et al. 2012; Brunet et al. 2014). It has been demonstrated that TssD, along with three other T6SS core components TssE, TssM and TssK, is required for the TssBC polymerisation into extended sheaths, and TssD may act as a scaffold for the tail sheath assembly by forming pre-assembled complexes of TssBC-D which is being incorporated progressively. This is consistent with the observation that the contracted sheath assembled spontaneously by TssBC are rapidly dissociated by TssH (Basler et al. 2012; Kapitein et al. 2013; Zoued et al. 2013; Brunet et al. 2014; Zoued et al. 2014). Therefore, based on the structural similarities with the contractile bacteriophages, the mode of action of the T6SS is suggested to occur through contraction of the tail sheath which powers the tail tube toward the cell exterior or directly puncturing the target cells.

TssH (also known as ClpV) belongs to the AAA^+ family of ATPases which form hexamers and bind ATP through the conserved AAA domain and convert the energy of ATP hydrolysis into the unfolding of substrates. The ATPase activity of TssH has been shown to be essential for the T6SS in many species, such as *P. aeruginosa* and *V.*

cholerae (Mougous et al. 2006; Bönemann et al. 2009). In *V. cholerae*, *tssH* mutant cells retain ~10% of T6SS activity, leading to the idea that TssH is not involved in assembly or secretion, but in increasing the efficiency of the system. TssH has been shown to interact with and disassemble the contracted TssBC sheath in an ATP-dependent manner (Bönemann et al. 2009), and is required for the dynamic of TssBC sheath assembly/disassembly by recycling TssB and TssC after sheath contraction (Basler and Mekalanos 2012; Basler et al. 2012; Kapitein et al. 2013). The disassembly of the TssBC sheath is achieved by a specific interaction between the hydrophobic groove located at the N-terminal region of TssH with the protruding N-terminal α-helix of TssC in the contracted sheath that forms a complex to detach TssC from TssB (Pietrosiuk et al. 2011). Other than recycling the sheath subunits, TssH also plays a role in preventing the formation of non-functioning TssBC tubules in the cytoplasm (Kapitein et al. 2013).

1.5.3.2.4 TssA

Until recently, there was very little information reported in the published literature on TssA except that it is a cytoplasmic protein and also an interacting partner of TssK (Cascales and Cambillau 2012; Zoued et al. 2013). Primary sequence analysis and secondary structure predictions did not identify its homology to the bacteriophage components. TssA was designated as a core T6SS subunit based on the observation that a protein containing a conserved N-terminal region of approximately 60-80 amino acids (referred to as an ImpA-related N-terminal domain) is encoded in all T6SS gene clusters (Shalom et al. 2007). An amino acid sequence alignment of the N-terminal region of TssA-like proteins is shown in Figure 1.7. TssA-like proteins are divided into three types based on the diversity in their C-terminal regions, named $TssA^S$, $TssA^E$ and TalT (TssA-like tag). There is a weakly conserved block of 20-24 amino acids close to the N-termini in all TssA-like proteins. All T6SS gene clusters encode either $TssA^S$ or $TssA^E$ but never both, whereas TalT is only encoded in the $TssA^E$ type gene cluster in some species, such as *Y. pseudotuberculosis* (*T6SS-4*) and *A. hydrophila*. The organisation of the T6SS gene clusters in *B. cenocepacia* and *A. hydrophila* are as shown in Figure 1.8.

$TssA^S$ is classified according to its C-terminal similarity to the *S. typhimurium* TssA protein (SciA) (Figure 1.9). $TssA^S$ orthologues are generally 340-375 amino acids in

length. The secondary structure prediction using Psipred v3.3 suggests TssA^{S} orthologues are rich in α-helices and they fold into a large N-terminal region (~250 amino acids) containing the ImpA-related N-terminal domain close to the N-terminus, that is connected to a small conserved C-terminal domain (60-75 amino acids) by a long unstructured interdomain linker with variable sequences and lengths (Figure 1.10). The TssA orthologue from *B. cenocepacia* belongs to this group.

TssA^{E} orthologues are usually 460-540 amino acids in length and found in several bacteria that can cause enteric infections and are therefore referred to as the 'enteric class'. They differ from TssA^{S} by the nature of their C-terminal region. Similar to TssA^{S}, the C-terminal region of TssA^{E} orthologues is predicted to be mainly composed of α-helices but with 250-320 amino acids in length which is much larger than that of TssA^{S} (Figure 1.12). Some TssA^{E} orthologues contain a C-terminal extension of 20-40 amino acids, defined as the TssA^{EII} sub-class (i.e. Ec042_4540 from EAEC), and those without the C-terminal extension are referred to as TssA^{EI} (i.e. AHA1844 from *A. hydrophila*) (Figure 1.11).

TalT orthologues are generally 435-480 amino acids in length and contain a putative transmembrane domain in the middle of polypeptide which might anchor TalT into the IM (Figure 1.13). The C-terminal region of TalT is connected to its N-terminal region by a non-conserved random coil linker region of 60-70 amino acids (Figure 1.14). However, TalT is suggested to be an accessory protein as inactivation of the *talT* orthologue in *V. cholerae* (also known as *vasL*) did not abolish the T6SS function (Zheng et al. 2011). Predicted domain organisations of three classes of TssA are represented in Figure 1.15.

TssA^{S} from *B. cenocepacia* has been shown to self-associate into a ring-like structure and also interact with many other T6SS core components *in vivo*, i.e. TssD, -E, -F, -H, -I, -K and -L, by two colleagues in our group (Shastri 2011; Ahmad 2013). In addition, TssA^{S} was shown to be important for the function of the T6SS-1 in *B. cenocepacia* (Spiewak 2015). TssA^{EI} from *A. hydrophila* also self-associates and is involved in a similar network of interactions with other T6SS core components *in vivo* as TssA^{S}. However, unlike TssA^{S}, TssA^{EI} was not shown to interact with TssE, and the interaction between TssA^{EI} and TssF was not tested (Ahmad 2013). TssA^{EII} from EAEC has been demonstrated to play a role in proper formation of the tail tube (Brunet et al. 2015). More recently, great progress has been made in our understanding of the structure, localisation

and potential function of TssA^{EII} from EAEC. TssA^{EII} is assembled into a dodecameric six-pointed star structure that interacts with baseplate components, tail tube, tail sheath and the membrane complex (Figure 1.16) (Zoued et al. 2016). An assembly model of the T6SS apparatus is also proposed where TssA^{EII} is recruited to the membrane complex resulting in the baseplate recruitment and subsequent initiation of tail tube and sheath polymerisation during which TssA^{EII} remains in contact with the baseplate-distal end of the tail during elongation (Zoued et al. 2016).

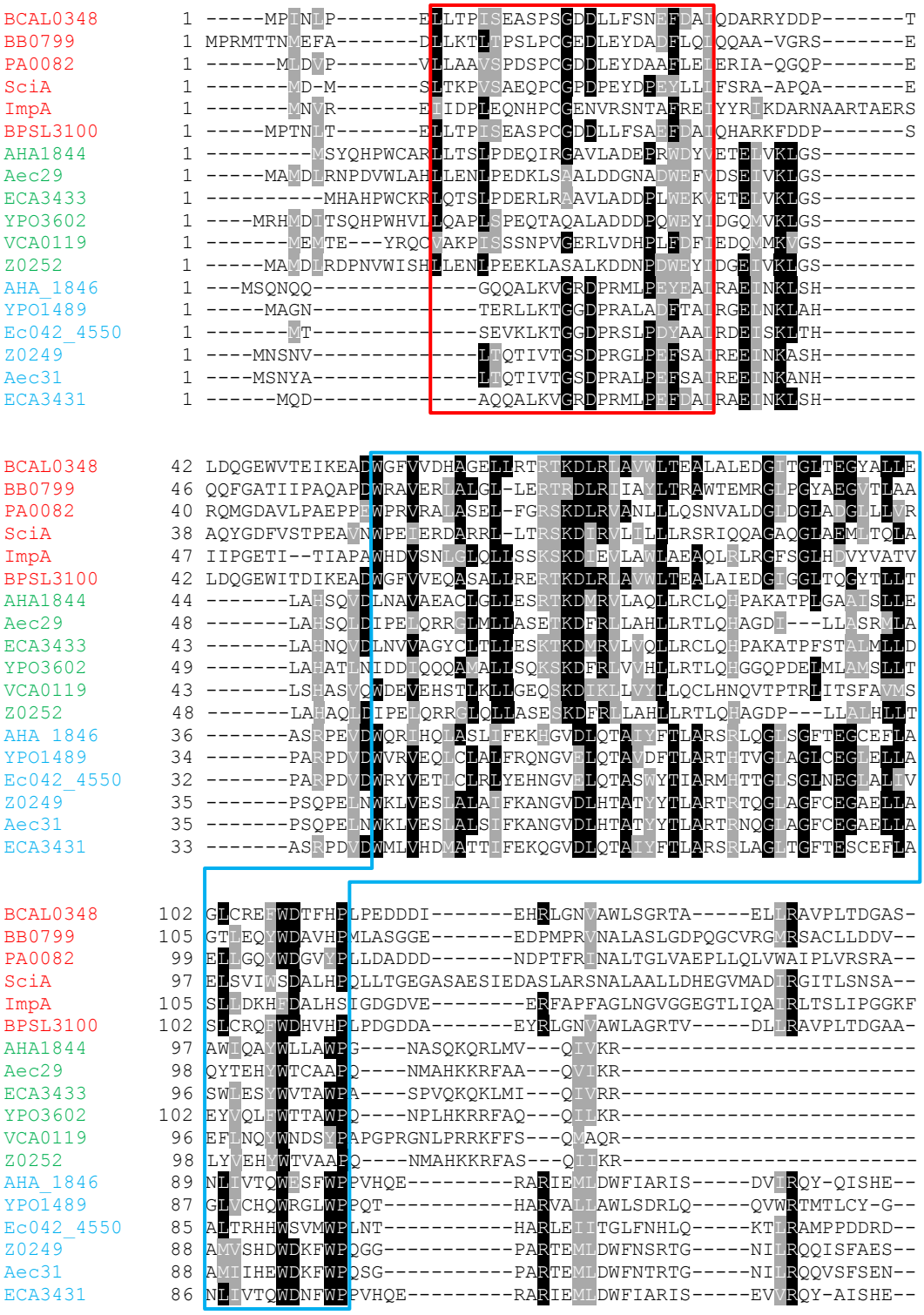


Figure 1.7 Amino acid sequence alignment of the N-terminal regions of TssA-like proteins from various Gram-negative bacteria.

Figure 1.7 Amino acid sequence alignment of the N-terminal regions of TssA-like proteins from various Gram-negative bacteria. Multiple sequence alignment was carried out between $TssA^S$ orthologues (red font) from *Burkholderia cenocepacia* strain J2315 (BCAL0348), *Bordetella bronchiseptica* strain RB50 (BB0799), *P. aeruginosa* strain PAO1 (PA0082), *S. enterica* serovar Typhimurium LT2 (SciA), *Rhizobium leguminosarum* bv. trifolii (ImpA) and *Burkholderia pseudomallei* strain K96243 (BPSL3100); $TssA^E$ orthologues (green font) from *Aeromonas hydrophila* strain ATCC 7966 (AHA_1844); *Escherichia coli* BEN2908 (Aec29), *Pectobacterium atrosepticum* strain SCRI1043 (ECA3433), *Yersinia pestis* CO92 (YPO3602), *Vibrio cholerae* serotype O1 (VC_A0119) and *Escherichia coli* O157:H7 (Z0252); and TalT orthologues (green font) from *Aeromonas hydrophila* strain ATCC 7966 (AHA_1846), *Yersinia pestis* CO92 (YPO1489), *Escherichia coli* 042 (Ec042_4550), *Escherichia coli* O157:H7 (Z0249), *Escherichia coli* BEN2908 (Aec31) and *Pectobacterium atrosepticum* SCRI1043 (ECA3431). Alignment was performed using Clustal Omega and conserved amino acids were shaded by Boxshade program. Amino acids that are identical at the same position in at least 50% of sequences are shown in white font and highlighted in black, whereas similar amino acids are highlighted in grey. The blue box encloses the amino acid sequence that is homologous in all TssA-like proteins (the ImpA-related N-terminal domain). The red box encloses the weakly conserved block in all three TssA-like proteins.

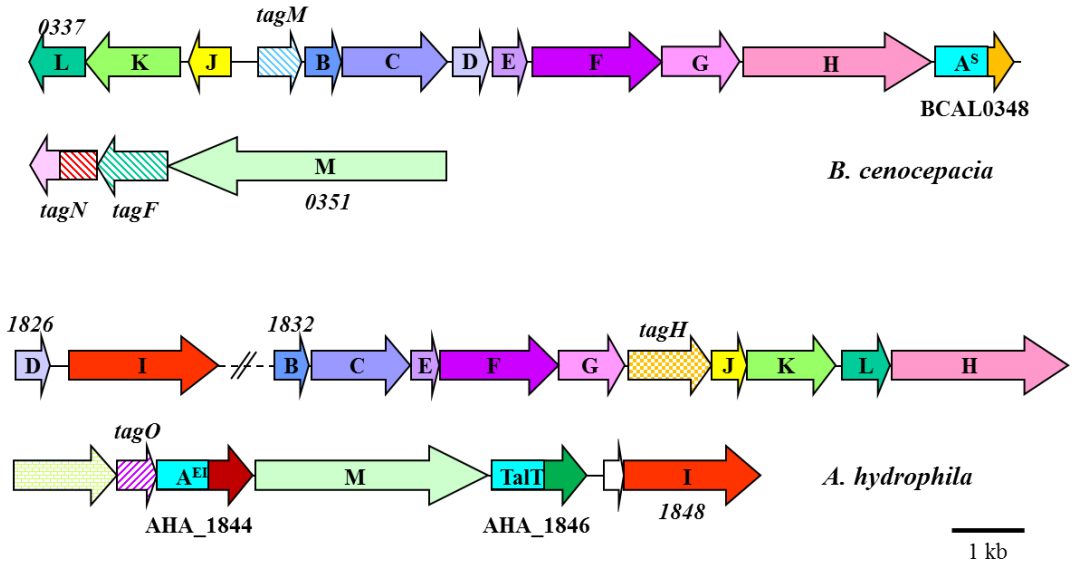
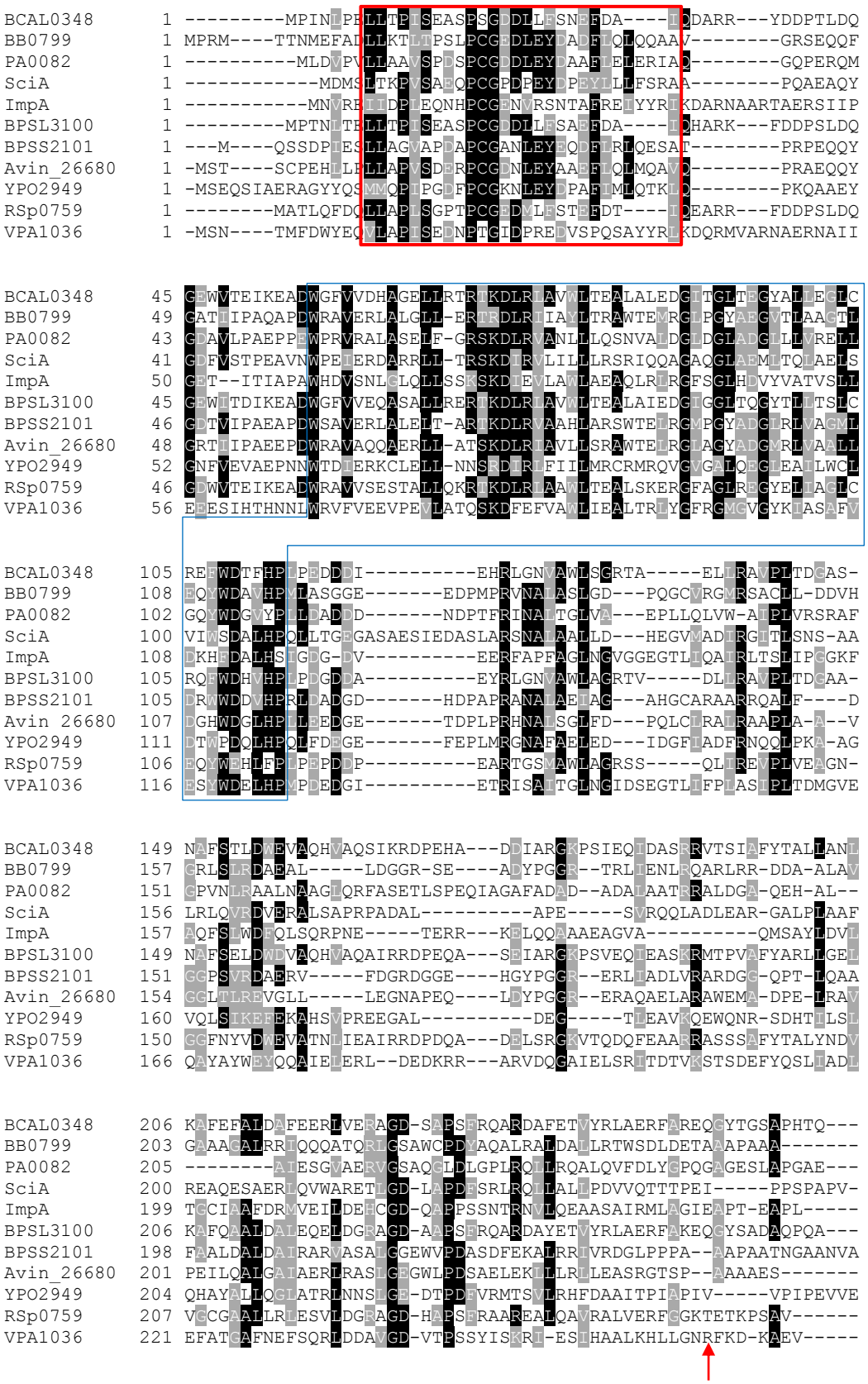


Figure 1.8 Organisation of the T6SS gene clusters in *B. cenocepacia and A. hydrophila*. Genes in each T6SS gene cluster are represented as block arrows. Direction of the arrows represents gene orientation. Orthologous genes are designated using the same colour. 'Core' T6SS genes with uniform colours are labelled with single letters corresponding to their *tss* nomenclature; T6SS associated genes (*tags*) are as indicated; hatched arrow in green encodes a σ^{54}-dependent activator; genes with unknown function are shown as white arrows. *B. cenocepacia* T6SS gene cluster encodes $TssA^{S}$; *A. hydrophila* contains two *tssA*-like genes, AHA_1844 and AHA_1846, flanking *tssM* that encode $TssA^{EI}$ and TalT, respectively.


BCAL0348
BB0799
PA0082
SciA
ImpA
BPSL3100
BPSS2101
Avin_26680
YPO2949
RSp0759
VPA1036

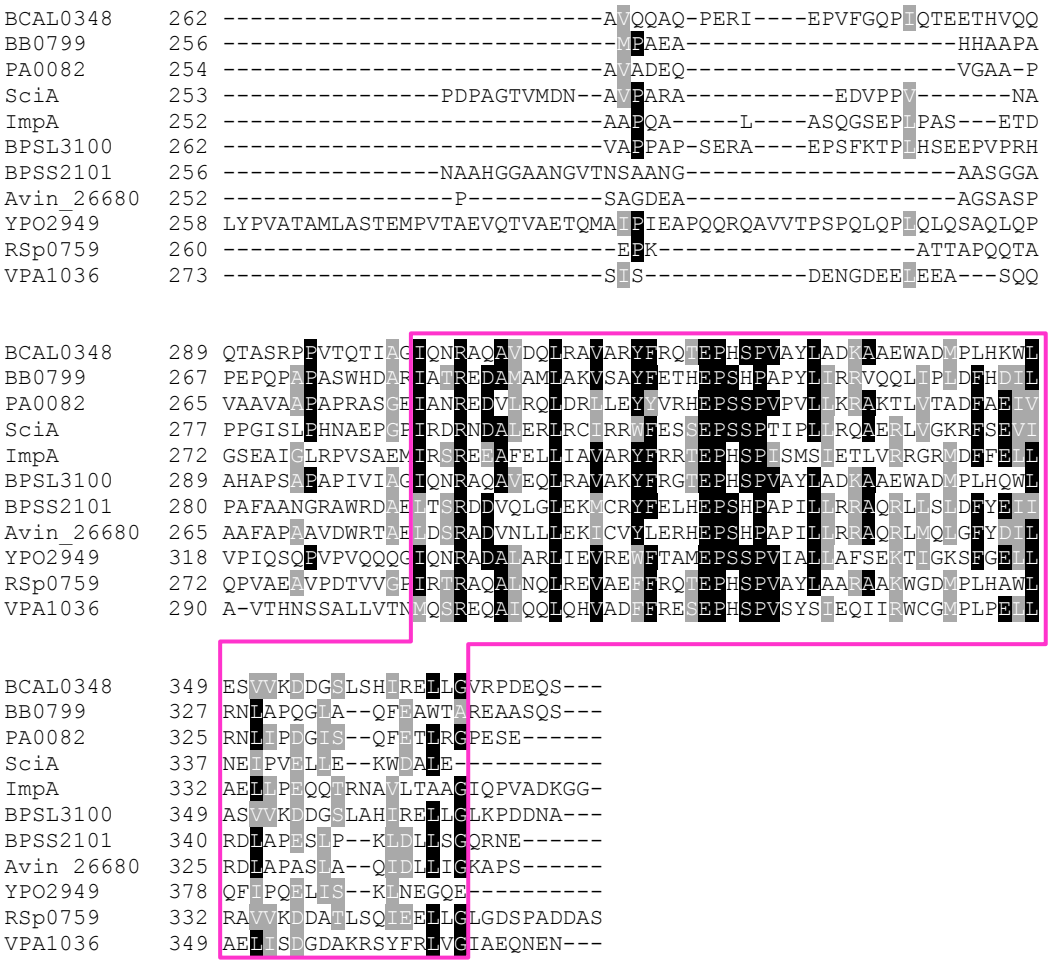


Figure 1.9 Amino acid alignment of TssA^S orthologues present in Gram-negative bacteria. Alignment was carried out between TssA^S orthologues present in *Burkholderia cenocepacia* strain J2315 (BCAL0348), *Bordetella bronchiseptica* strain RB50 (BB0799), *P. aeruginosa* strain PAO1 (PA0082), *S. enterica* serovar Typhimurium. LT2 (SciA), *Rhizobium leguminosarum* bv. trifolii (ImpA), *Burkholderia pseudomallei* strain K96243 (BPSL3100), *Burkholderia pseudomallei* strain K96243 (BPSS2101), *Azotobacter vinelandii* DJ. (Avin_26680), *Yersinia pestis* CO92 (YPO2949), *Ralstonia solanacearum* GMI1000 (RSp0759) and *Vibrio parahaemolyticus* RIMD 2210633 (VPA1036). Alignment was performed using ClustalW2 and conserved amino acids were shaded by the Boxshade program. Amino acids that are identical at the same position in ≥ 50% of sequences are shown in white font and highlighted in black, whereas similar amino acids are highlighted in grey. Red arrow indicates the start of the predicted possible inter-domain linker region. The red box encloses the weakly conserved block in all three TssA-like proteins; the blue box encloses the amino acid sequence that is homologous in all TssA-like proteins (the ImpA-related N-terminal domain); the magenta box encloses the conserved C-terminal region within the TssA^S orthologues.

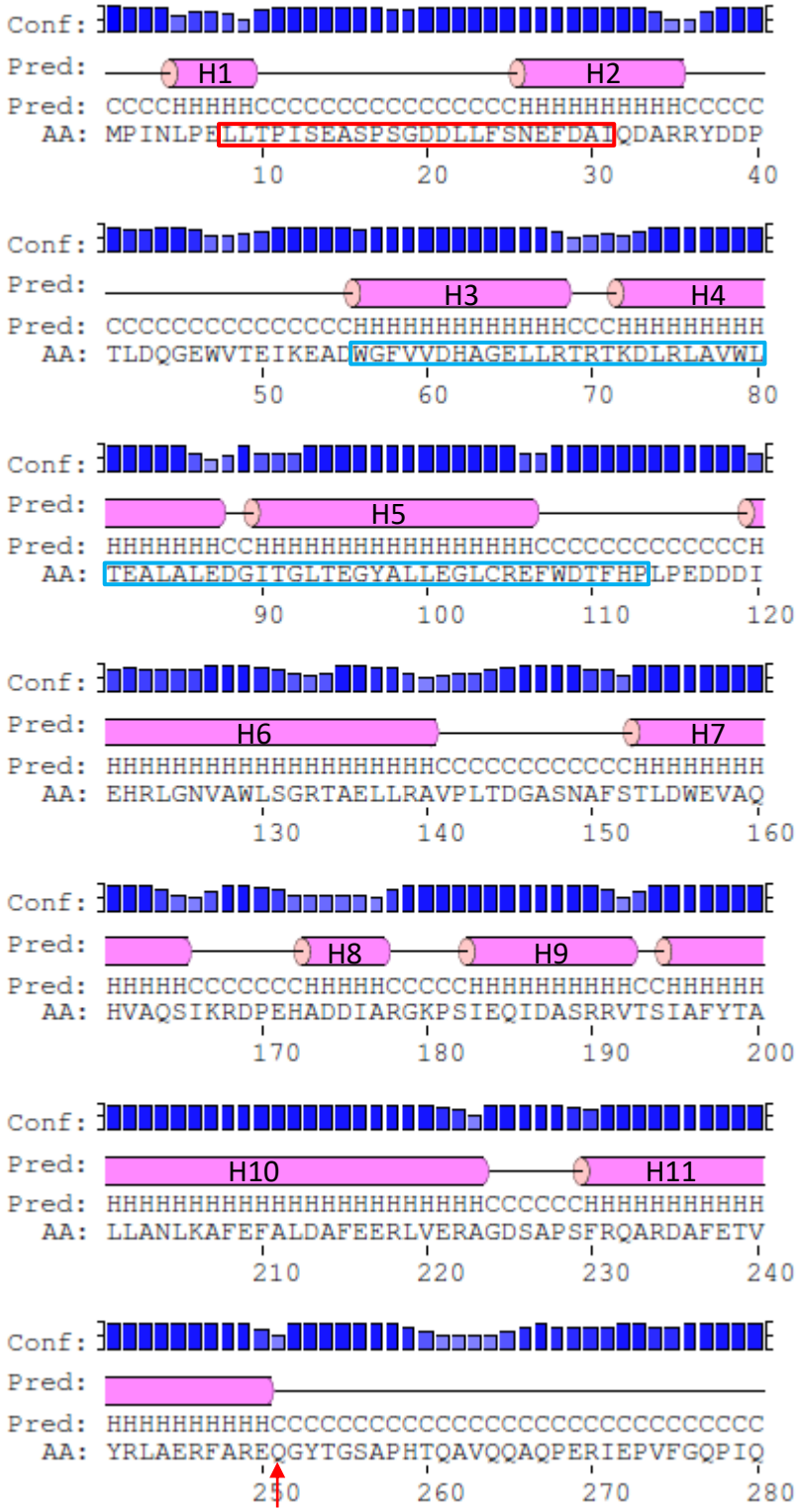
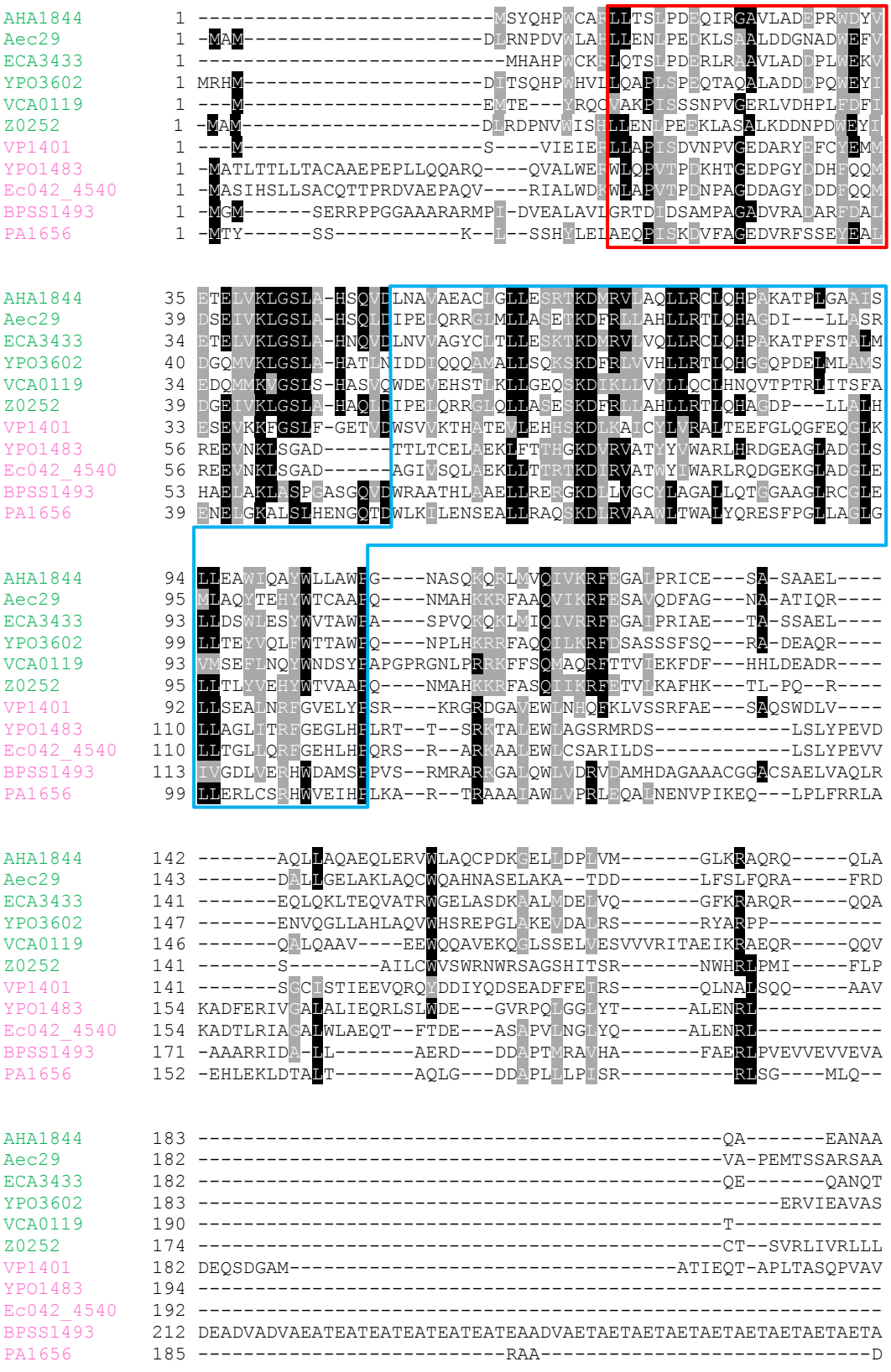
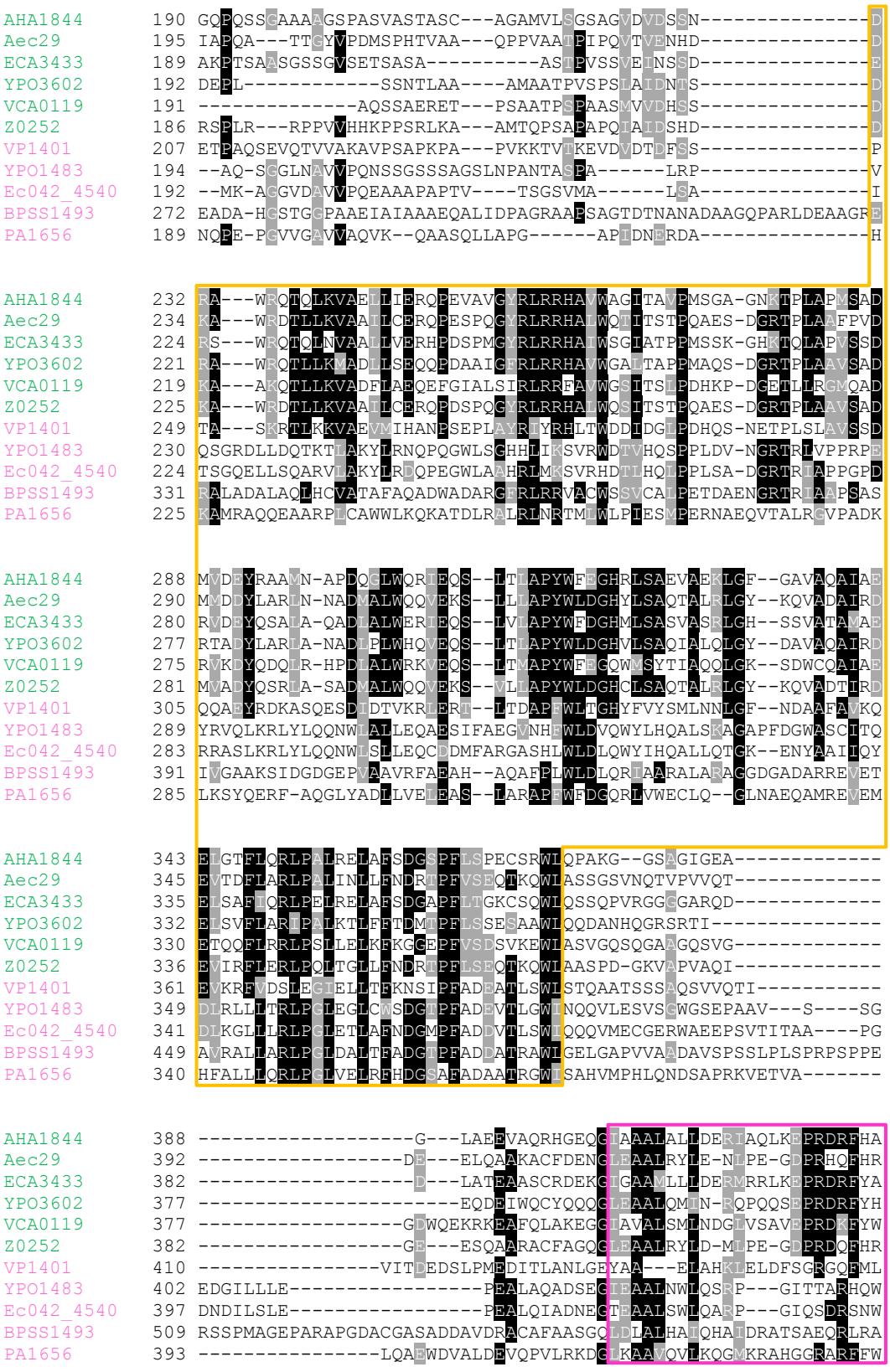

Conf:
Pred:
H1
H2
Pred: CCCCHHHHHCCCCCCCCCCCCCCCHHHHHHHHHHCCCCC
AA: MPINLPELLTPISEASPSGDDLLFSNEFDAIQDARRYDDP
10
20
30
40
Conf:
Pred:
H3
H4
Pred: CCCCCCCCCCCCCCCHHHHHHHHHHHHHCCCHHHHHHHHH
AA: TLDQGEWVTEIKEADWGFVVDHAGELLRTRTKDLRLAVWL
50
60
70
80
Conf:
Pred:
H5
Pred: HHHHHHHCCHHHHHHHHHHHHHHHHCCCCCCCCCCCCCH
AA: TEALALEDGITGLTEGYALLEGLCREFWDTFHPLPEDDDI
90
100
110
120
Conf:
Pred:
H6
H7
Pred: HHHHHHHHHHHHHHHHHHHHCCCCCCCCCCCCHHHHHHHH
AA: EHRLGNVAWLSGRTAELLRAVPLTDGASNAFSTLDWEVAQ
130
140
150
160
Conf:
Pred:
H8
H9
Pred: HHHHHCCCCCCCHHHHHCCCCCHHHHHHHHHHCCHHHHHH
AA: HVAQSIKRDPEHADDIARGKPSIEQIDASRRVTSIAFYTA
170
180
190
200
Conf:
Pred:
H10
H11
Pred: HHHHHHHHHHHHHHHHHHHHHHCCCCCCHHHHHHHHHH
AA: LLANLKAFEFALDAFEERLVERAGDSAPSFRQARDAFETV
210
220
230
240
Conf:
Pred:
Pred: HHHHHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCCCC
AA: YRLAERFAREQGYTGSAPHTQAVQQAQPERIEPVFGQPIQ
250
260
270
280

Figure 1.10 Secondary structure prediction of *B. cenocepacia* TssA^S (BCAL0348). Secondary structure was predicted using Psipred v3.3 and predicted helix numbers are indicated. The red box encloses weakly conserved block in all three TssA-like proteins; the blue box encloses the amino acid sequence that is homologous in all TssA-like proteins (the ImpA-related N-terminal domain); the magenta box encloses the conserved C-terminal region within the TssA^S orthologues. Red arrow indicates the start of the predicted inter-domain linker region.


AHA1844
Aec29
ECA3433
YPO3602
VCA0119
Z0252
VP1401
YPO1483
Ec042_4540
BPSS1493
PA1656


AHA1844 190 GQPQSSGAAAAGSPASVASTASC---AGAMVLSGSAGVDVDSSN----------D
Aec29 195 IAPQA---TTGYVPDMSPHTVAA---QPPVAATPIPQVTVENHD----------D
ECA3433 189 AKPTSAASGSSGVSETSASA----------ASTPVSSVEINSSD----------E
YPO3602 192 DEPL----------SSNTLAA----AMAATPVSPSLAIDNTS----------D
VCA0119 191 ----------AQSSAERET---PSAATPSPAASMVVDHSS----------D
Z0252 186 RSPLR---RPPVVHHKPPSRLKA---AMTQPSAPAPQIAIDSHD----------D
VP1401 207 ETPAQSEVQTVVAKAVPSAPKPA---PVKKTVIKEVDVDTDFSS----------P
YPO1483 194 --AQ-SGGLNAVVPQNSSGSSSAGSLNPANTASPA------LRP----------V
Ec042_4540 192 --MK-AGGVDAVVPQEAAAPAPTV----TSGSVMA------LSA----------I
BPSS1493 272 EADA-HGSTGGPAAEIAIAAAEQALIDPAGRAAPSAGTDTNANADAAGQPARLDEAAGRE
PA1656 189 NQPE-PGVVGAVVAQVK--QAASQLLAPG------APIDNERDA----------H
AHA1844 232 RA---WRQTQLKVAELLIERQPEVAVGYRLRRHAVWAGITAVPMSGA-GNKTPLAPMSAD
Aec29 234 KA---WRDTLLKVAAILCERQPESPQGYRLRRHALWQTITSTPQAES-DGRTPLAAFPVD
ECA3433 224 RS---WRQTQLNVAALLVERHPDSPMGYRLRRHAIWSGIATPPMSSK-GHKTQLAPVSSD
YPO3602 221 RA---WRQTLLKVADLLSEQQPDAAIGHRLRRHAVWGALTAPPMAQS-DGRTPLAAVSAD
VCA0119 219 KA---AKQTLLKVADFLAEQEFGIALSIRLRRFAVWGSITSLPDHKP-DGETLLRGMQAD
Z0252 225 KA---WRDTLLKVAAILCERQPDSPQGYRLRRHALWQSITSTPQAES-DGRTPLAAVSAD
VP1401 249 TA---SKRTLKKVAEVMIHANPSEPLAYRLYRHLTWDDIDGLPDHQS-NETPLSLAVSSD
YPO1483 230 QSGRDLLDQTKTLAKYLRNQPQGWLSGHHLIKSVRWDTVHQSPPLDV-NGRTRLVPPRPE
Ec042_4540 224 TSGQELLSQARVLAKYLRLQPEGWLAAHRLMKSVRHDTLHQLPPLSA-DGRTRLAPPGPD
BPSS1493 331 RALADALAQLHCVATAFAQADWADARGHRLRRVACWSSVCALPETDAENGRTRLAAPSAS
PA1656 225 KAMRAQQEAARPLCAWWLKQKATDLRALRLNRTMLWLPIESMPERNAEQVTALRGVPADK
AHA1844 288 MVDEYRAAMN-APDQGLWQRIEQS--LTLAPYWFEGHRLSAEVAEKLGF--GAVAQAIAE
Aec29 290 MMDDYLARLN-NADMALWQQVEKS--LLLAPYWLDGHYLSAQTALRLGY--KQVADAIRD
ECA3433 280 RVDEYQSALA-QADLALWERIEQS--LVLAPYWFDGHMLSASVASRLGH--SSVATAVAE
YPO3602 277 RTADYLARLA-NADLPLWHQVEQS--LTLAPYWLDGHVLSAQIALQLGY--DAVAQAIRD
VCA0119 275 RVKDYQDQLR-HPDLALWRKVEQS--LTMAPYWFEGQWMSYTIAQQLGK--SDWCQAIAE
Z0252 281 MVADYQSRLA-SADMALWQQVEKS--VLLAPYWLDGHCLSAQTALRLGY--KQVADTIRD
VP1401 305 QQAEYRDKASQESDLDTVKRLERI--LTDAPFWLTGHYFVYSMLNNLGF--NDAAFAVKQ
YPO1483 289 YRVQLKRLYLQQNWLALLEQAESIFAEGVNHEWLDVQWYLHQALSKAGAPFDGWASCITQ
Ec042_4540 283 RRASLKRLYLQQNWLSLLEQCDDMFARGASHLWLDLQWYIHQALLQTGK--ENYAAIIQY
BPSS1493 391 IVGAAKSIDGDGEPVAAVRFAEAH--AQAFPLWLDLQRIAARALARAGGDGADARREVET
PA1656 285 LKSYQERF-AQGLYADLLVELEAS--LARAPEWFDGQRLVWECLQ--GLNAEQAMREVEM
AHA1844 343 ELGTFLQRLPALRELAFSDGSPFLSPECSRWLQPAKG--GSAGIGEA----------
Aec29 345 EVTDFLARLPALINLLFNDRIPFVSEQTKQWLASSGSVNQTVPVVQT----------
ECA3433 335 ELSAFIQRLPELRELAFSDGAPFLTGKCSQWLQSSQPVRGGGGARQD----------
YPO3602 332 ELSVFLARIPALKTLFFTDMIPFLSSESAAWLQQDANHQGRSRTI----------
VCA0119 330 ETQQFLRRLPSLLELKFKGGEPFVSDSVKEWIASVGQSQGAAGQSVG----------
Z0252 336 EVIRFLERLPQLTGLLFNDRIPFLSEQTKQWIAASPD-GKVAPVAQI----------
VP1401 361 EVKRFVDSLEGIELLTFKNSIPFADEATLSWISTQAATSSSAQSVVQTI----------
YPO1483 349 DLRLLLTRLPGLEGLCWSDGIPFADEVTLGWINQQVLESVSGWGSEPAAV---S----SG
Ec042_4540 341 DLKGLLLRLPGLETLAFNDGMPFADDVTLSWIQQQVMECGERWAEEPSVTITAA----PG
BPSS1493 449 AVRALLARLPGLDALTFADGIPFADDATRAWIGELGAPVVAADAVSPSSLPLSPRPSPPE
PA1656 340 HFALLLQRLPGLVELRFHDGSAFADAATRGWISAHVMPHLQNDSAPRKVETVA-------
AHA1844 388 ----------G---LAEEVAQRHGEQGIAAALALLDERIAQLKEPRDRFHA
Aec29 392 ----------DE---ELQAAKACFDENGLEAALRYLE-NLPE-GDPRHQFHR
ECA3433 382 ----------D---LATEAASCRDEKGIGAAMLLLDERMRRLKEPRDREYA
YPO3602 377 ----------EQDEIWQCYQQQGLEAALQMIN-RQPQQSEPRDREYH
VCA0119 377 ----------GDWQEKRKEAFQLAKEGGIAVALSMLNDGLVSAVEPRDKEYW
Z0252 382 ----------GE---ESQAARACFAGQGLEAALRYLD-MLPE-GDPRDQFHR
VP1401 410 ----------VITDEDSLPMEDITLANLGEYAA---ELAHKLELDFSGRGQEML
YPO1483 402 EDGILLLE----------PEALAQADSEGIEAALNWLQSRP---GITTARHQW
Ec042_4540 397 DNDILSLE----------PEALQIADNEGTEAALSWLQARP---GIQSDRSNW
BPSS1493 509 RSSPMAGEPARAPGDACGASADDAVDRACAFAASGQLDLALHAIQHAIDRATSAEQRLRA
PA1656 393 ----------LQAEWDVALDEVQPVLRKDGLKAAVQVLKQGMKRAHGGRAREFW

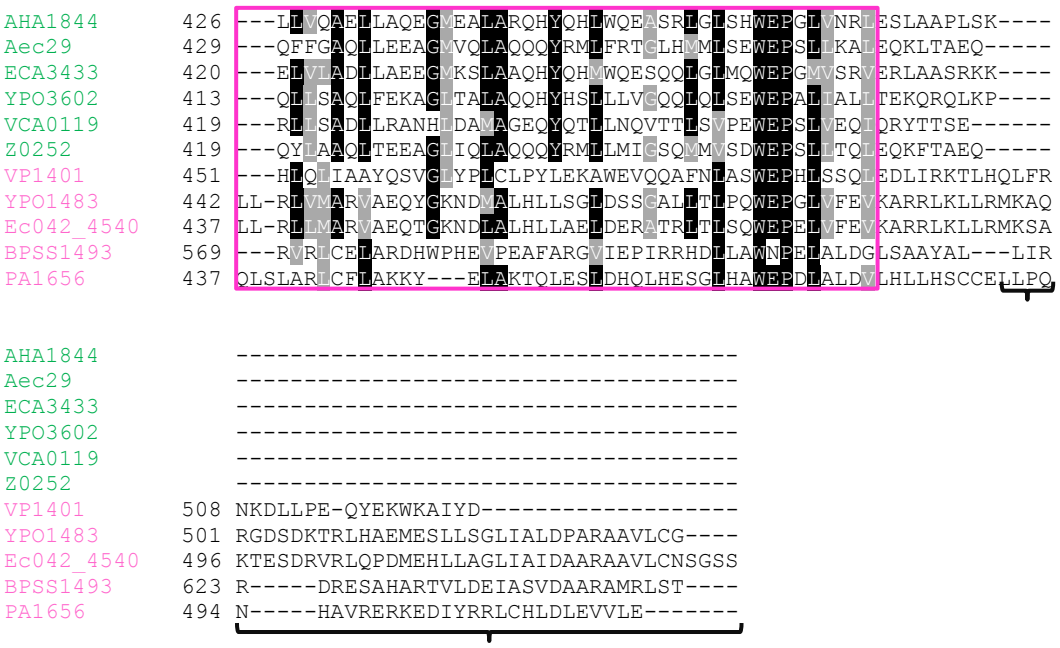


Figure 1.11 Amino acid alignment of TssA^{E} orthologues present in Gram-negative bacteria. Alignment was carried out between TssA^{E} orthologues present in *Aeromonas hydrophila* strain ATCC 7966 (AHA_1844); *Escherichia coli* (Aec29), *Pectobacterium atrosepticum* strain SCRI1043 (ECA3433), *Yersinia pestis* CO92 (YPO3602), *Vibrio cholerae* serotype O1 (VC_A0119), *Escherichia coli* O157:H7 (Z0252), *Vibrio parahaemolyticus* RIMD 2210633 (VP1401), *Yersinia pestis* CO92 (YPO1483), *Escherichia coli* 042 (Ec042_4540), *Burkholderia pseudomallei* K96243 (BPSS1493) and *Pseudomonas aeruginosa* PAO1 (PA1656). Alignment was performed using ClustalW2 and conserved amino acids were shaded by the Boxshade program. Amino acids that are identical at the same position in ≥ 50% of sequences are shown in white font and highlighted in black, whereas similar amino acids are highlighted in grey. The red box encloses the weakly conserved block in all three TssA-like proteins; the blue box encloses the amino acid sequence that is homologous in all TssA-like proteins (the ImpA-related N-terminal domain); the magenta box encloses the conserved C-terminal region within the TssA^{S} orthologues; the orange box encloses the conserved middle domain region within the TssA^{E} orthologues; the magenta box encloses the conserved C-terminal region within the TssA^{E} orthologues. The TssA^{EII} sub-class orthologues (magenta font) contain a C-terminal extension (black bracket). The TssA^{EI} sub-class orthologues are shown in green font.

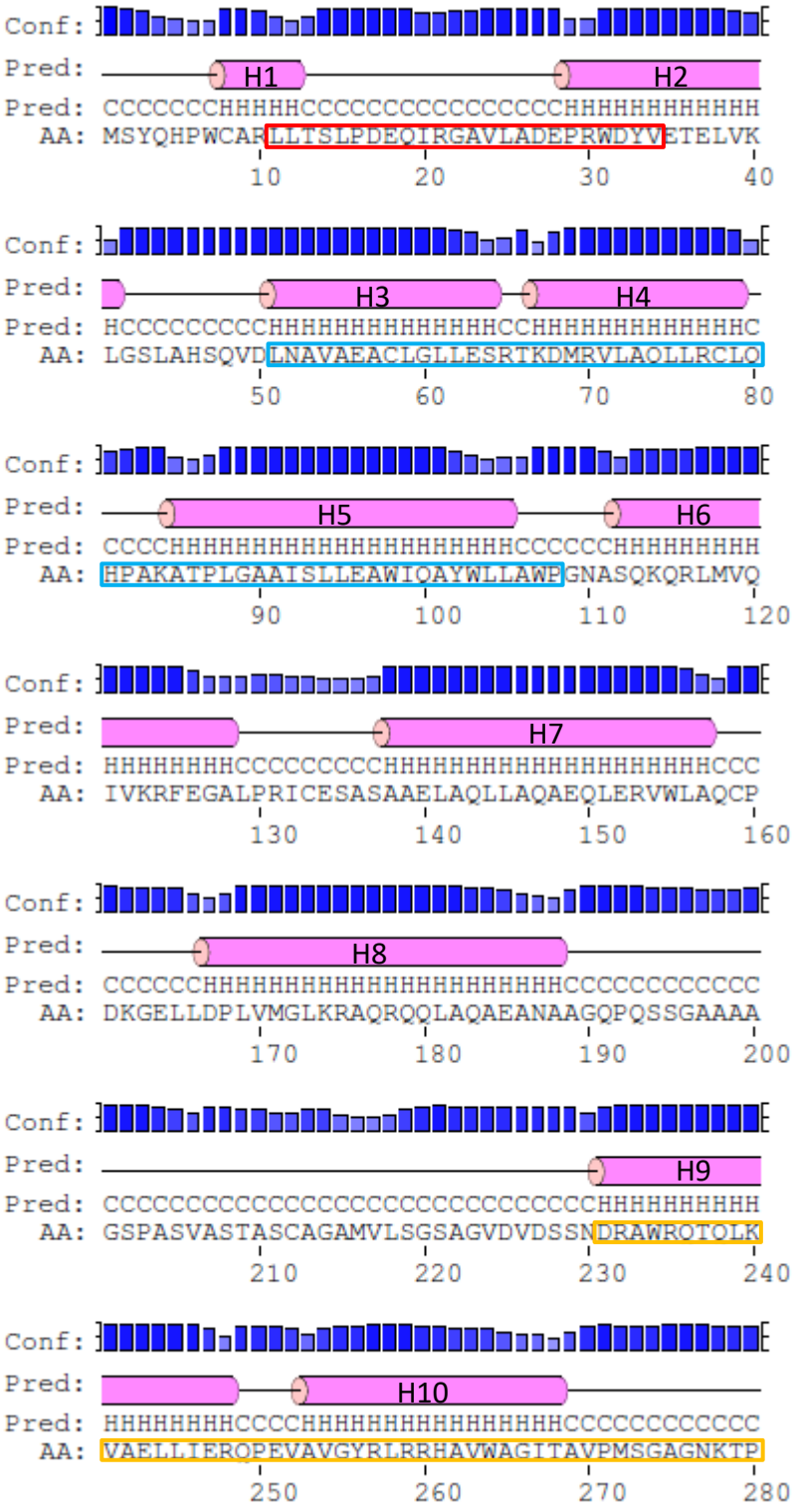

Conf:
Pred:
H1
H2
Pred: CCCCCCCHHHHHCCCCCCCCCCCCCCCHHHHHHHHHHHH
AA: MSYQHPWCARLLTSLPDEQIRGAVLADEPRWDYVETELVK
10
20
30
40
Conf:
Pred:
H3
H4
Pred: HCCCCCCCCCHHHHHHHHHHHHHHCCHHHHHHHHHHHHC
AA: LGSLAHSQVDLNAVAEACLGLLESRTKDMRVLAQLLRCLQ
50
60
70
80
Conf:
Pred:
H5
H6
Pred: CCCCHHHHHHHHHHHHHHHHHHHHCCCCCCHHHHHHHH
AA: HPAKATPLGAAISLLEAWIQAYWLLAWPGNASQKQRLMVQ
90
100
110
120
Conf:
Pred:
H7
Pred: HHHHHHHHCCCCCCCCCHHHHHHHHHHHHHHHHHHCCC
AA: IVKRFEGALPRICESASAAELAQLLAQAEQLERVWLAQCP
130
140
150
160
Conf:
Pred:
H8
Pred: CCCCCCHHHHHHHHHHHHHHHHHHHHHCCCCCCCCCCC
AA: DKGELLDPLVMGLKRAQRQQLAQAEANAAGQPQSSGAAAA
170
180
190
200
Conf:
Pred:
H9
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHHHHHHH
AA: GSPASVASTASCAGAMVLSGSAGVDVDSSNDRAWRQTQLK
210
220
230
240
Conf:
Pred:
H10
Pred: HHHHHHHHCCCCHHHHHHHHHHHHHHHCCCCCCCCCCC
AA: VAELLIERQPEVAVGYRLRRHAVWAGITAVPMSGAGNKTP
250
260
270
280

Figure 1.12 Secondary structure prediction of *A. hydrophila* TssA^{EI} (AHA_1844). Secondary structure was predicted using Psipred v3.3 and predicted helices are indicated. The red box encloses weakly conserved block in all three TssA-like proteins; the blue box encloses the amino acid sequence that is homologous in all TssA-like proteins (the ImpA-related N-terminal domain); the orange box encloses the conserved middle domain region within the TssA^{EI} orthologues; the magenta box encloses the conserved C-terminal region within the TssA^{EI} orthologues.

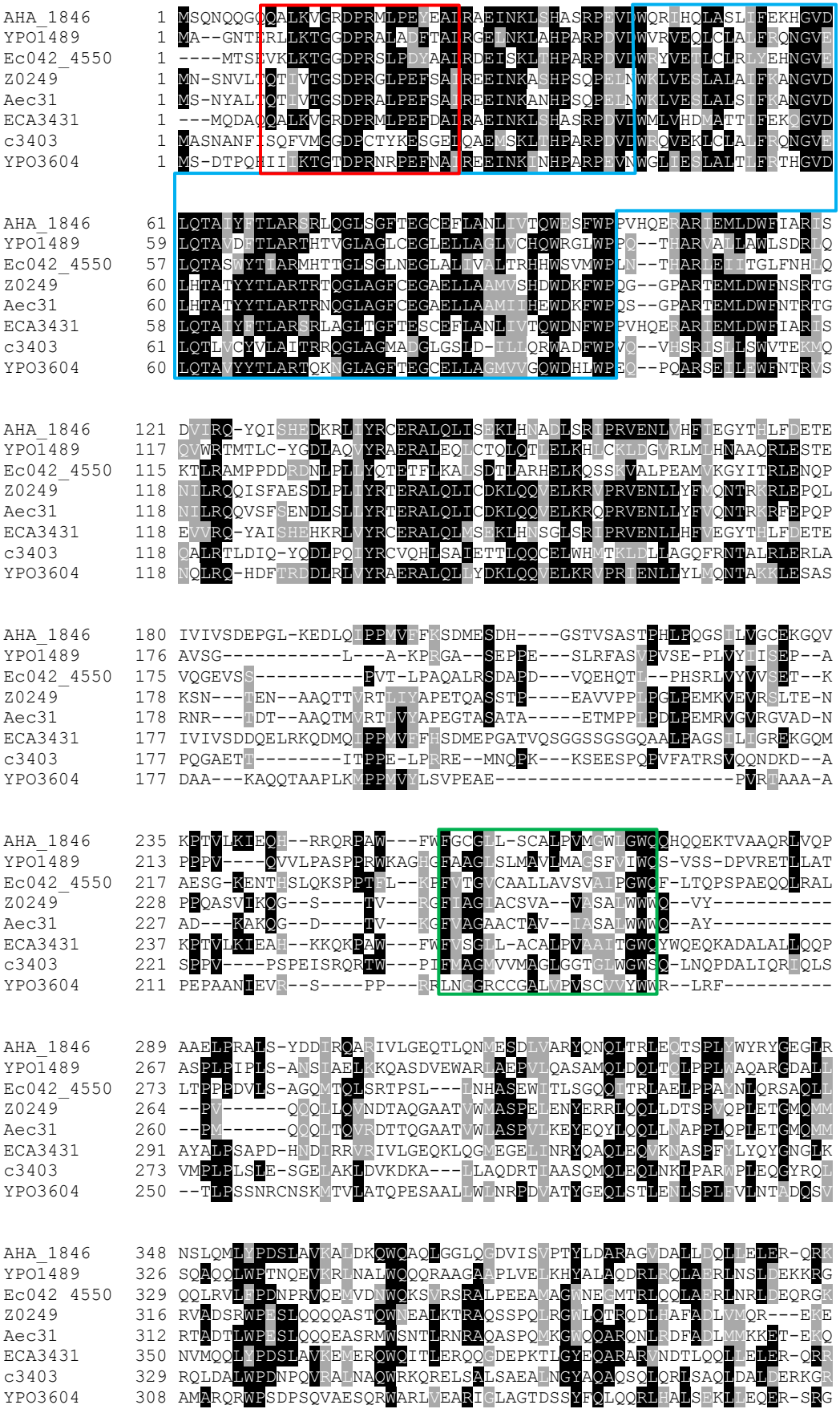

AHA_1846
YPO1489
Ec042_4550
Z0249
Aec31
ECA3431
c3403
YPO3604

Figure 1.13 Amino acid alignment of TalT orghologues present in Gram-negative bacteria. Alignment was carried out between TalT orthologues present in *Aeromonas hydrophila* strain ATCC 7966 (AHA_1846), *Yersinia pestis* CO92 (YPO1489), *Escherichia coli* 042 (Ec042_4550), *Escherichia coli* O157:H7 (Z0249), *Escherichia coli* S88 (Aec31), *Pectobacterium atrosepticum* SCRI1043 (ECA3431), *Burkholderia cenocepacia* (c3403) and *Yersinia pestis* CO92 (YPO3604). Alignment was performed using ClustalW2 and conserved amino acids were shaded by the Boxshade program. Amino acids that are identical at the same position in ≥ 50% of sequences are shown in white font and highlighted in black, whereas similar amino acids are highlighted in grey. The red box encloses weakly conserved block in all three TssA-like proteins; the blue box encloses the amino acid sequence that is homologous in all TssA-like proteins (the ImpA-related N-terminal domain); the green box encloses the putative transmembrane domain within the TalT orthologues.

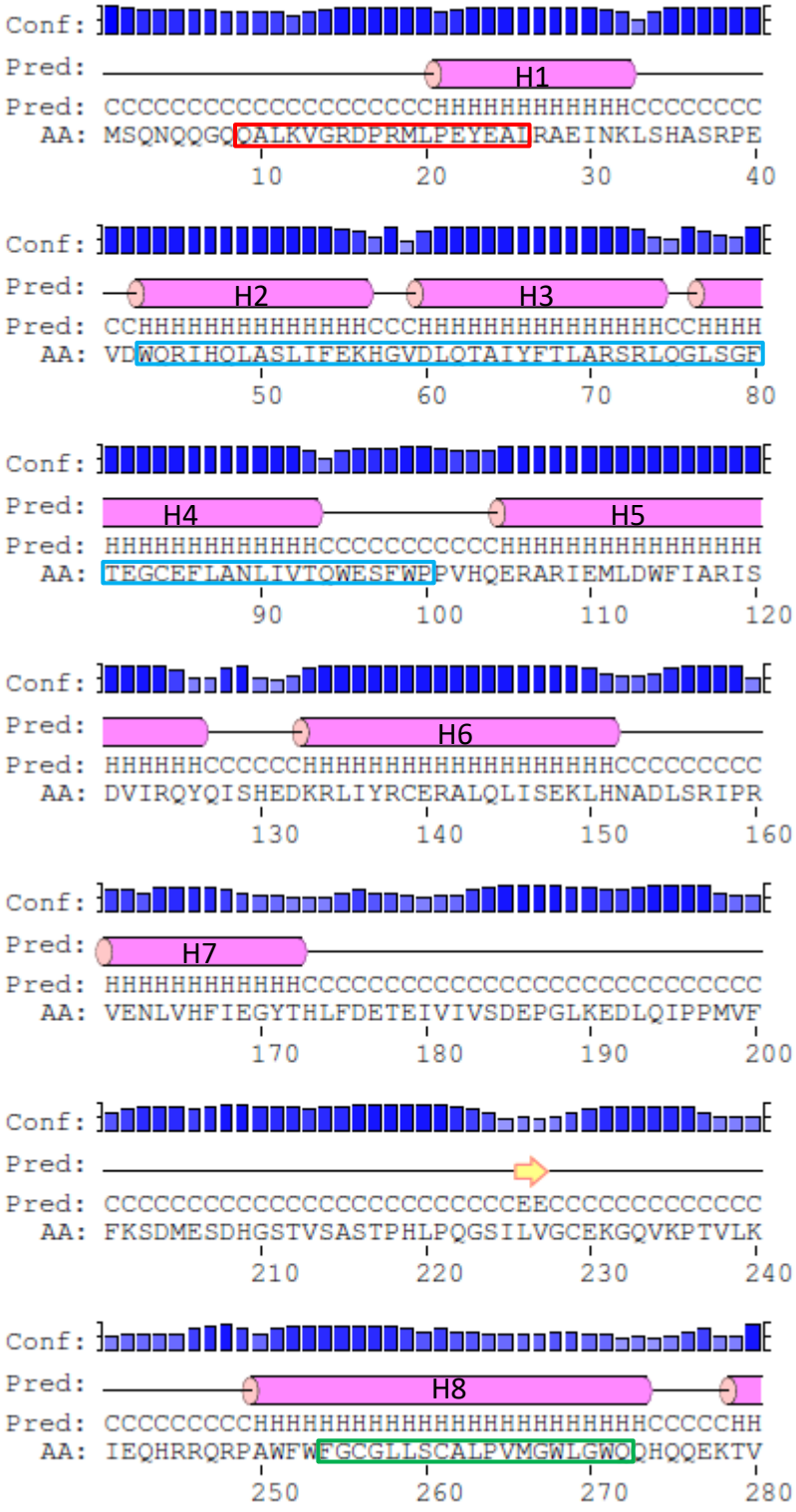

Conf:
Pred:
H1
Pred: CCCCCCCCCCCCCCCCCCCCHHHHHHHHHHHHCCCCCCCC
AA: MSQNQQGQALKVGRDPRMLPEYEAIRAEINKLSHASRPE
10
20
30
40
Conf:
Pred:
H2
H3
Pred: CCHHHHHHHHHHHHHHCCCHHHHHHHHHHHHHHCCHHHH
AA: VDWQRIHQLASLIFEKHGVDLQTAIYFTLARSRLQGLSGF
50
60
70
80
Conf:
Pred:
H4
H5
Pred: HHHHHHHHHHHHHCCCCCCCCCCCHHHHHHHHHHHHHHHH
AA: TEGCEFLANLIVTQWESFWPPVHQERARIEMLDWFIARIS
90
100
110
120
Conf:
Pred:
H6
Pred: HHHHHHCCCCCCHHHHHHHHHHHHHHHHHHCCCCCCCCC
AA: DVIRQYQISHEDKRLIYRCERALQLISEKLHNADLSRIPR
130
140
150
160
Conf:
Pred:
H7
Pred: HHHHHHHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCCC
AA: VENLVHFIEGYTHLFDETEIVIVSDEPGLKEDLQIPPMVF
170
180
190
200
Conf:
Pred:
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCEECCCCCCCCCCCCC
AA: FKSDMESDHGSTVSASTPHLPQGSILVGCEKGQVKPTVLK
210
220
230
240
Conf:
Pred:
H8
Pred: CCCCCCCCCHHHHHHHHHHHHHHHHHHHHHHHCCCCCHH
AA: IEQHRRQRPAWFWFGCGLLSCALPVMGWLGWQHQQEKTV
250
260
270
280

Figure 1.14 Secondary structure prediction of *A. hydrophila* TalT (AHA_1846). Secondary structure was predicted using Psipred v3.3 and predicted helices are indicated. The red box encloses weakly conserved block in all three TssA-like proteins; the blue box encloses the amino acid sequence that is homologous in all TssA-like proteins (the ImpA-related N-terminal domain); the green box encloses the putative transmembrane domain within the TalT orthologues.

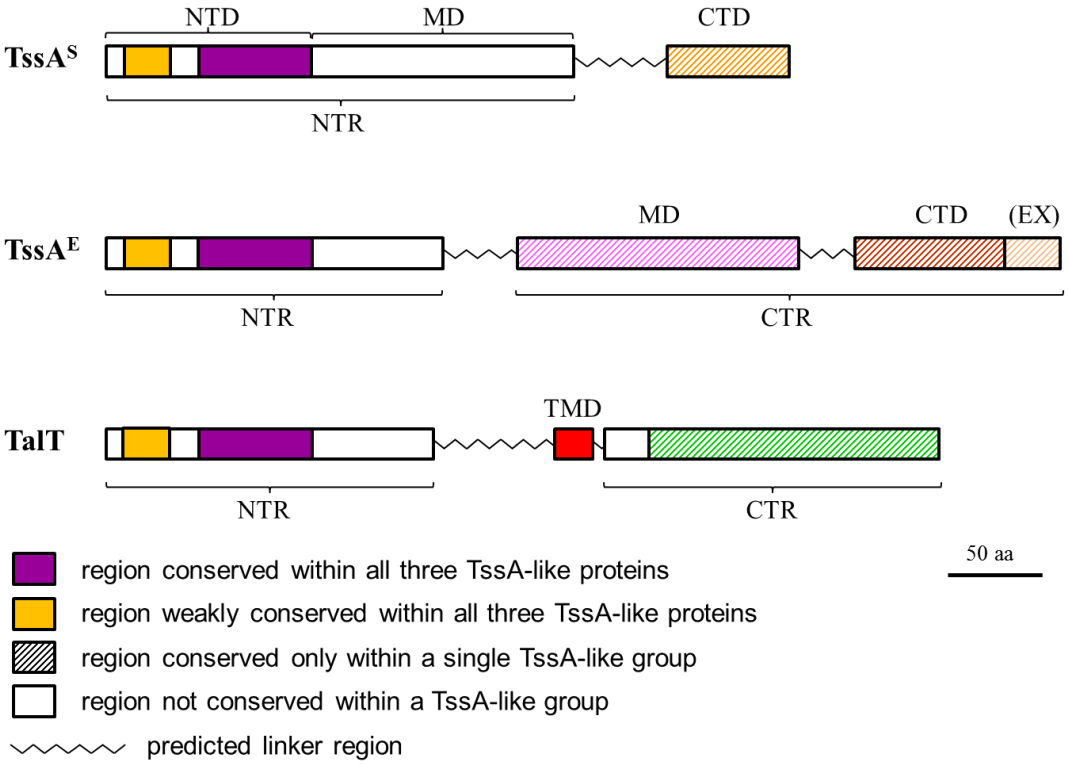


Figure 1.15 Predicted domain organisations of three TssA classes. Domain organisation was predicted based on amino acid sequence alignment and secondary structure prediction. NTD, N-terminal domain; MD, middle domain; CTD, C-terminal domain; NTR, N-terminal region; CTR, C-terminal region; EX, extension; TMD, transmembrane domain.

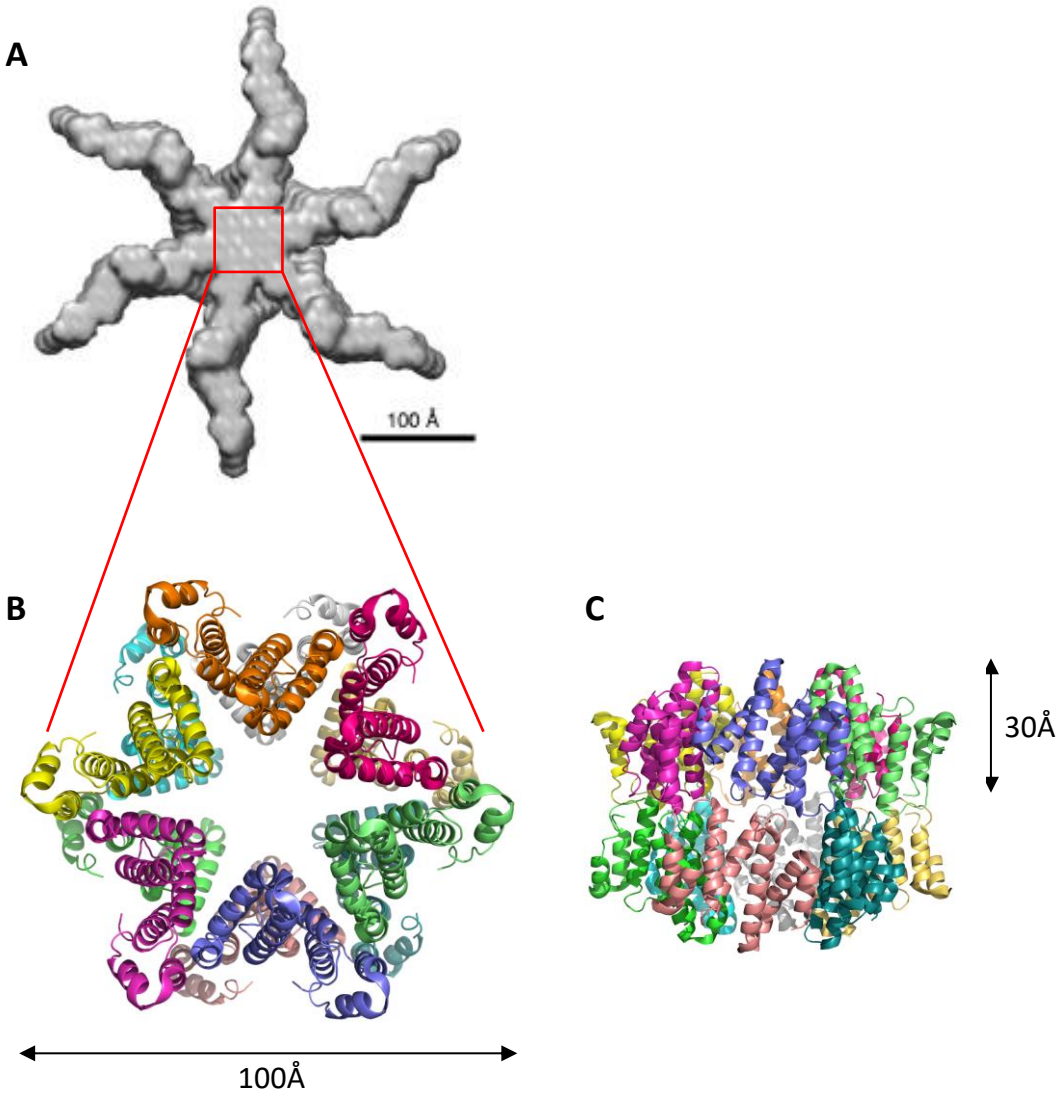


Figure 1.16 Structure of TssA^{EII} from *Enteroaggregative Escherichia coli*. **A**. Low resolution SAXS model of TssA^{EII}. Reprinted by permission from Macmillan Publishers Ltd: Nature (Zoued et al. 2016), © 2016 (license number 3883081120345). **B** and **C**, structure of TssA^{EII} C-terminal domain at 3.35 Å resolution (PDB 4YO5, (Zoued et al. 2016)). B, top view; C, side view.

1.5.4 Diverse biological functions of the T6SS

The major role of the T6SS was thought to be involved in virulence. However, subsequent studies have suggested diverse roles for the T6SS including modulation of bacterial competition (Hood et al. 2010; MacIntyre et al. 2010; Schwarz et al. 2010; Murdoch et al. 2011; Carruthers et al. 2013; Gueguen and Cascales 2013), biofilm formation (Enos-Berlage et al. 2005; Aschtgen et al. 2008; de Pace et al. 2010; Ma et al. 2014), quorum sensing (Weber et al. 2008), stress sensing and environmental adaptation (Ishikawa et al. 2012; Salomon et al. 2013).

1.5.4.1 Bacterial competition

The T6SS was initially shown to be involved in interbacterial competition following identification of anti-bacterial effectors secreted by *Pseudomonas aeruginosa* H1-T6SS into the neighbouring bacterial cells, named Tse1-3 (Hood et al. 2010). Subsequently, anti-bacterial effectors secreted by the T6SS present in a wide range of bacteria were reported, including the T6SS-1 of *Burkholderia thailandensis* (Schwarz et al. 2010), *V. cholerae* (MacIntyre et al. 2010), *Serratia marcescens* (Murdoch et al. 2011), *Citrobacter rodentium* (Gueguen and Cascales 2013), and *Acinetobacter baumanii* (Carruthers et al. 2013).

Beyond that, *in vivo* investigations using fluorescent microscopy have visualised the T6SS-modulated attack between adjacent bacterial cells (Basler and Mekalanos 2012; Leroux et al. 2012; Gerc et al. 2015). In addition, it has been demonstrated that T6SS can target bacterial cells from the same species (Dong et al. 2013; Unterweger et al. 2014; Whitney et al. 2014; Alcoforado Diniz and Coulthurst 2015). In short, the T6SS is engaged in bacterial competition, particularly towards gaining a survival advantage against other adjacent bacteria.

1.5.4.2 Pathogenesis

Several human bacterial pathogens utilise specific T6SSs to deliver effectors to target eukaryotic cells. T6SS-dependent cytotoxicity towards the unicellular eukaryote *Dictyostelium discoideum* and macrophages was discovered in *V. cholerae* (Pukatzki et al. 2006). This was later shown to be as a result of a secreted specialized TssI (VgrG1) spike

protein that carries an actin cross-linking domain (ACD) at the C-terminus which belongs to the toxins of the MARTX family. The ACD was delivered into the cytosol of macrophages where it targeted G-actin resulting in dysfunctioning macrophages via interference with the dynamics of the cell cytoskeleton (Pukatzki et al. 2007). The conserved T6SS-5 of *B. pseudomallei* and *B. mallei* was demonstrated to be required for virulence towards eukaryotic cells, in particular the formation of multi-nucleate giant cells (Schell et al. 2007; Burtnick et al. 2011; Schwarz et al. 2014; Toesca et al. 2014).

Although *P. aeruginosa* is mainly considered as an extracelluar pathogen, it has been shown to be an intracellular pathogen that can invade epithelial cells and macrophages (Fleiszig et al. 1995; Belon et al. 2015). An evolved TssI effector containing a putative zinc-metalloprotease domain (termed TssI-2b) delivered by the H2-T6SS of *P. aeruginosa* allows the internalization of the bacterium into epithelial cells by interaction with the γ-tubulin ring complex, a core component of the host microtubule network that facilitates nucleation of microtubules (Sana et al. 2015). Two phospholipase D effectors, PldA and PldB, delivered by the H2-T6SS and H3-T6SS of *P. aeruginosa*, respectively, were shown to be capable of targeting both prokaryotic and eukaryotic cells, i.e. they are trans-kingdom effectors. The anti-eukaryotic activity of these two effectors occurs through activation of the phosphatidylinositol 3-kinase/AKT-dependent signalling pathway for the subsequent actin rearrangement and protrusion formation in the human epithelial cells that facilitates bacterial internalisation (Jiang et al. 2014). More recently, another trans-kingdom effector, TplE (also known as Tle4), was identified to be secreted by the H2-T6SS of *P. aeruginosa*. TplE exhibits lipase and phospholipase A1 activities that are responsible for its anti-bacterial activity, and it also contains a eukaryotic PGAP1 (post-glycosylphosphatidylinositol attachment to proteins 1)-like domain for endoplasmic reticulum targeting which promotes autophagy in epithelial cells (Jiang et al. 2016).

1.5.5 Regulation of T6SS

As might be expected T6SS gene clusters are regulated for sensing modifications in the environmental conditions and generating responses accordingly. In agreement with its diverse functions, the T6SS gene clusters are not regulated in a unique manner. In general,

T6SS gene clusters are regulated in a bacterium-specific manner, i.e. at the transcriptional and post-transcriptional levels. In addition, a post-translational regulatory pathway has also been described.

1.5.5.1 Transcriptional and post-transcriptional regulation

The involvement of transcriptional and/or post-transcriptional regulation typically matches the role of the T6SS, including two-component regulatory systems, alternative sigma factors, histone-like nucleoid associated proteins (H-NS), quorum sensing, etc. An example of a two-component regulatory system is the VirA-VirG sensor kinase-response regulator which regulates the T6SS-5 gene cluster and genes for actin-based intracellular motility in *B. mallei* and its relatives (Schell et al. 2007). Another example of the two-component regulatory system, the AtsR-AtsT phosphorelay pathway, has been identified to be a major global regulator of *B. cenocepacia* pathogenicity (Aubert et al. 2008; Khodai-Kalaki et al. 2013). The alternative sigma factor σ^{54}-dependent transcriptional regulation is also involved in the T6SS regulation, such as in *V. cholerae*. In *V. cholerae*, σ^{54} (RpoN) and its activator VasH are crucial for its pathogenicity, in which VasH is encoded by the T6SS gene cluster and acts as an enhancer binding protein to promote transcription from σ^{54}-dependent T6SS promoters (Pukatzki et al. 2006; Ishikawa et al. 2009; Bernard et al. 2011; Kitaoka et al. 2011). Another commonly employed T6SS transcriptional regulation is by H-NS that controls the expression of several genes within T6SS gene clusters by binding to the AT-rich regions of some genes. In the HSI-II and -III clusters of *P. aeruginosa*, an H-NS-like protein, MvaT, has been shown to repress the transcription of *tssH*, *tssD* and *tssI* (Castang et al. 2008). In *V. cholerae*, the QS response regulator LuxO regulates the expression of the T6SS mediated by the transcriptional regulator HapR at low cell density, but conversely LuxO is inactivated at high cell density allowing the expression of a master regulator TsrA that represses the T6SS expression (Zheng et al. 2010).

Post-transcriptional regulatory pathways regulate the T6SS at the mRNA level. The virulence regulator RetS in *P. aeruginosa* represses the expression of the HSI-I T6SS gene cluster through modulation of the phosphorylation status of sensor kinase GacS in the GacS-GacA two-component system (Brencic and Lory 2009). Repressed GacS-GacA two-component system results in an increased expression of a translational regulator

RsmA through repressing the transcription of the small regulatory RNAs RsmY and RsmZ. RsmA is a post-transcriptional regulator which binds to and negatively regulates the translation of the target T6SS HSI-I gene associated mRNA by preventing ribosome binding for translation initiation (Brencic and Lory 2009).

1.5.5.2 Post-translational regulation

1.5.5.2.1 Threonine phosphorylation pathway (TPP)

Expression of T6SS gene clusters is also regulated by a post-translational phosphorylation-dependent pathway in some organisms, such as *P. aeruginosa*, facilitating a rapid response to specific stimuli. In the T6SS HSI-1 of *P. aeruginosa*, the TPP is activated in response to the surface growth, in which threonine kinase PpkA phosphorylates a FHA domain of Fha1 (TagE) that subsequently acts as a scaffold protein recruiting TssH to the T6SS apparatus (Mougous et al. 2007; Hsu et al. 2009; Silverman et al. 2011). Deactivation of the T6SS is secured by the dephosphorylation of Fha1 by the threonine phosphatase PppA (TagG) (Figure 1.17) (Mougous et al. 2007).

A recent study suggests PpkA phosphorylates TssL in *A. fabrum*, instead of Fha1, which facilitates a direct interaction between TssL and Fha1 resulting in the activation of the T6SS. However, the prevalence of this TssL phosphorylation-induced T6SS assembly in other species remains to be elucidated (Lin et al. 2014).

Activation of PpkA in *P. aeruginosa* requires several upstream regulatory proteins, including an OM lipoprotein TagQ which is required for localising a periplasmic protein TagR to the OM (Silverman et al. 2011; Casabona et al. 2013). TagR promotes polymerisation and activation of PpkA by interacting with the periplasmic domain of PpkA (Hsu et al. 2009). An IM ABC transporter complex comprised of TagT and TagS is also required for full activation of PpkA (Casabona et al. 2013).

More recently, an observation by time-lapsed fluorescence microscopy demonstrated that the T6SS HSI-1 of *P. aeruginosa* is activated in a response to the T6SS activity from another T6SS-possessing bacterium including adjacent sister cells, terming this phenomenon ‘T6SS duelling’ (Basler et al. 2013). Subsequently genetic analysis revealed this phenomenon is regulated by the TPP involving the TagQRST complex as a

result of an undefined signal related to the T6SS attack from another bacterium. It is further proposed that the TagQRST complex may directly sense the envelope disruption caused by exogenous T6SS attack (Basler et al. 2013). However, the T6SS duelling phenomenon is possibly restricted to a subset of T6SS-possessing organisms containing Fha1, in particular limited to *Pseudomonas* species only in which a full set of required regulatory components are present (Ho et al. 2014).

Apart from exogenous T6SS attack, membrane disruption caused by the T4SS-mediated mating pair formation pilus and some membrane disrupting antibiotics such as polymyxin B is also able to activate the T6SS HSI-1 of *P. aeruginosa* through the TPP (Ho et al. 2013).

1.5.5.2.2 Phosphorylation-independent pathway

The T6SS duelling activity mediated by the TPP is not a feature of all T6SSs. It is shown that the T6SS of *S. marcescens* is activated without having a T6SS-induced attack through cell contact (Gerc et al. 2015). Homologous to H1-T6SS of *P. aeruginosa*, the *S. marcescens* T6SS also has Fha, PpkA and PppA. Previous work has suggested the important role of PpkA for the T6SS regulation in *S. marcescens* (Fritsch et al. 2013). Instead of having a threonine kinase activity regulating the activation of the system, PpkA may respond to an 'off' signal and prevent the system from firing (Gerc et al. 2015). Similar observation has also been described in *V. cholerae* where activation of T6SS is not triggered by attack from a neighbouring bacteria, although this system does no possess orthologues of PpkA and PppA (Basler and Mekalanos 2012).

A phosphorylation-independent pathway involving a relatively conserved T6SS-associated protein TagF has been proposed. It is reported that TagF functions as a post-translational regulator that negatively regulates the T6SS HSI-I activity in *P. aeruginosa* without involvement of phosphorylated Fha1 and does not respond to surface growth (Silverman et al. 2011). Indeed, similar to the TPP, the TagF-mediated phosphorylation-independent pathway does require recruitment of TssH to the T6SS assembly. The recruitment involves unphosphoryated Fha1 through formation of a complex with TssH (Silverman et al. 2011). However, the involvement of other possible regulatory proteins in the TagF-mediated phosphorylation-independent pathway is

currently unclear. In some T6SS gene clusters, a TagF protein is encoded whereas no Fha-encoding gene is present, i.e. *B. cenocepacia* (Figure 1.8).

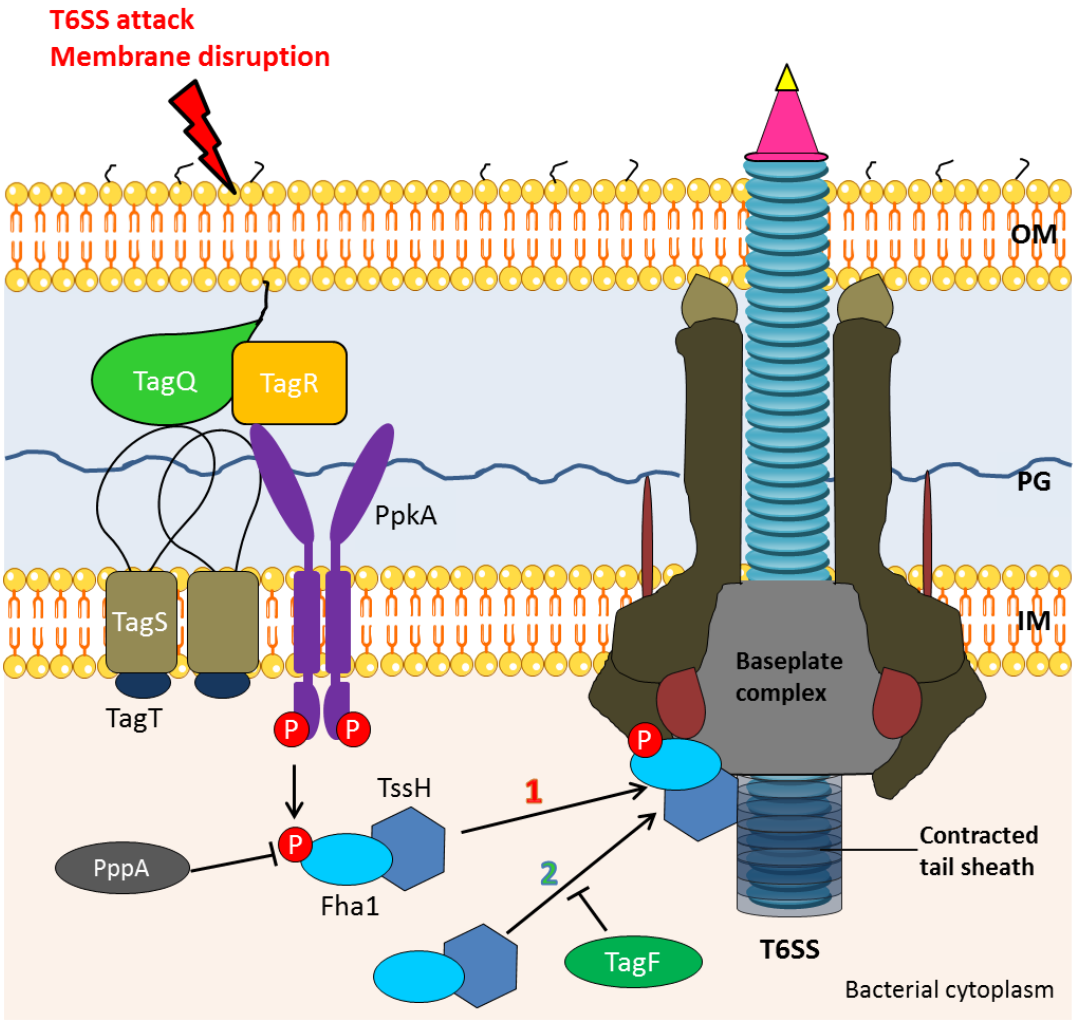


Figure 1.17 Model of two proposed post-translational pathways for regulation of the T6SS in *P. aeruginosa*. 1. Surface growth induced activation of the tyrosine phosphorylation pathway (TPP) and TagQRST system. This pathway may occur through membrane disruption. Stimulation of TPP results in the PpkA activation, phosphorylation of Fha1 (*p*-Fha1), TssH recruitment, followed by *p*-Fha1-TssH complex recruitment to the T6SS apparatus to positively regulate T6SS activity. 2. Phosphorylation-independent regulation pathway involving TagF. Activation of this pathway is through relief of inhibition of TagF. TagF negatively regulates the recruitment of unphosphorylated Fha1 in complex with TssH to the T6SS apparatus through an unknown signal. IM, inner membrane; PG, peptidoglycan; OM, outer membrane.

1.6 Effectors of the T6SS

Several secreted effectors of the T6SS have been identified and described over recent years. The diverse biological functions of the T6SS and targets are associated with a variety of effectors delivered by this secretion system.

1.6.1 Specialised TssI effectors

Early research on the T6SS effectors identified effector activity associated with some TssI proteins. These TssI-associated proteins commonly possess a C-terminal extension fused to the gp5-like β-prism which is delivered into the target cell cytosol after membrane perforation (Ma et al. 2009a). These C-terminal extensions comprised of a peptidoglycan binding (PG) domain in TssI-3 of *V. cholerae*, an actin cross-linking domain (ACD) in TssI-1 of *V. cholerae*, and a zinc-metalloprotease domain in TssI-2b of *P. aeruginosa* (Pukatzki et al. 2007; Pukatzki et al. 2009; Sana et al. 2015). These TssI proteins containing C-terminal extensions are termed specialised TssI proteins.

Further investigation also showed that TssI-1 can target eukaryotic cells by demonstrating actin-cross linking activity of TssI-1 that was inhibited both *in vitro* and *in vivo* in a *V. cholerae tssI*-1 mutant (Pukatzki et al. 2007). This actin-cross linking activity of TssI-1 has been shown to aggregate actin molecules and alter target cell morphology (Durand et al. 2012a). The peptidoglycan-binding region of *V. cholerae* TssI-3 has also been identified as exhibiting anti-bacterial properties (Brooks et al. 2013). The C-terminal region of TssI-2b of *P. aeruginosa* was also shown to target eukaryotic cells mediated by the H2-T6SS. It was shown to promote invasion of *P. aeruginosa* into epithelial cells as a result of specific interactions between TssI-2b and the host cell microtubule component, the γ-tubulin ring complex (Sana et al. 2015). Another characterised specialised TssI is the TssI-1 protein of *A. hydrophila* which contains a vegetative insecticidal protein domain in the C-terminal extension which exhibits ADP-ribosyltransferase activity that induces apoptosis of eukaryotic cells (Pukatzki et al. 2009; Suarez et al. 2010a).

1.6.2 Peptidoglycan-targeting effectors

PG, as a major component of bacterial cell walls, is a preferred target of anti-bacterial effectors. Hood and colleagues identified that H1-T6SS in *P. aeruginosa* mediated the secretion of three anti-bacterial effectors, named Tse1-3 (Hood et al. 2010). These effector proteins are paired with specific cognate immunity proteins i.e. Tsi1-3, respectively, which protects cells from self-toxicity. Tse1 and Tse3 were subsequently demonstrated to have peptidoglycan hydrolase activity acting as amidase and muramidase (glycosidase), respectively, upon being injected into the target bacterial periplasm by the T6SS (Russell et al. 2011).

There is growing evidence that there is a large family of PG-targeting effectors (and immunity proteins) that are present in several T6SS-containing bacteria and play a role in anti-bacterial activity, including amidase effector-immunity pairs referred to as Tae-Tai (type VI amidase effector/immunity) (Russell et al. 2012). The Tae family contains four subfamilies, named Tae1-4, exhibiting diverse bond cleavage specificity of PG molecules. Another PG-targeting family is the glycoside hydrolase effector-immunity pairs, known as Tge-Tgi (type VI glycoside effector/immunity), in which Tge1 and Tge3 function as muramidases whereas Tge2 is proposed to be a glucosaminidase (Whitney et al. 2013). Many of these PG-targeting effectors have been characterised, including an effector of the Tae1 family (Tse1/Tae1^{PA}) in *P. aeruginosa* (Hood et al. 2010; Russell et al. 2011), an effector of the Tae2 family (Tae2^{BT}) in *B. thailandensis* (Russell et al. 2012), effectors of Tae4 family (Ssp1/Tae4.1^{SM} and Ssp2/Tae4.2^{SM}) in *S. marcescens* (Murdoch et al. 2011; Fritsch et al. 2013), an effector of the Tge1family (Tse3/Tge1^{PA}) in *P. aeruginosa*, and an effector of the Tge2 family (Tge2^{PP}) in *P. protegens* (Hood et al. 2010; Whitney et al. 2013). More recently, the CTD of the specialised TssI-3 in *V. cholerae* has been demonstrated to have PG-targeting activity and categorised as a muramidase but it is unrelated to the Tge family (Brooks et al. 2013; Russell et al. 2014).

1.6.3 Membrane-targeting effectors

1.6.3.1 Phospholipase effectors

It has been recently identified that a superfamily of T6SS effectors exists exhibiting phospholipase activities, termed Tle (type VI lipase effectors) that target the bacterial membrane by hydrolysing membrane component phospholipids (Russell et al. 2013). The Tle superfamily can be divided into five subfamilies (Tle1-5) based on overall sequence homology. Members of Tle1-4 families exhibit a conserved GxSxG motif, and are phospholipase A and B derivatives, whereas Tle5 family members have dual HxKxxxxD motifs and have phospholipase D activity (Russell et al. 2013). Similar to the PG-targeting effectors, these membrane-targeting effectors also pair with specific cognate immunity proteins, named Tli1-5 (type VI lipase immunity 1-5), that are proposed to act in the periplasm (Russell et al. 2013).

Some effectors belong to the Tle superfamily have been characterised to be secreted into the target bacterial cells and cause cell lysis, including the Tle1 family effector $Tle1^{BT}$ in *B. thailandensis*, the Tle2 family effector $Tle2^{VC}$ (also known as TseL) in *V. cholerae*, and the Tle5 family effector PldA (also known as $Tle5^{PA}$) in *P. aeruginosa* (de Pace et al. 2010; Dong et al. 2013; Russell et al. 2013; Jiang et al. 2014). In addition, apart from anti-bacterial activity, $Tle2^{VC}$ in *V. cholerae* has been demonstrated to possess anti-eukaryotic predation activity that prevents it from being attenuated by eukaryotic cells (Dong et al. 2013).

1.6.3.2 Pore-forming effectors

Apart from phospholipase effectors, other membrane targeting effectors were also identified to be structural homologues of pore-forming colicins. This includes VasX of *V. cholerae* that exhibits anti-amoeboid properties (Miyata et al. 2011). It has been further proposed that VasX is recruited by an accessory protein VasW to the T6SS apparatus and it targets the bacterial IM from the periplasm following secretion (Miyata et al. 2013). *B. thailandensis* BTH_I2691, a substrate exported via the T6SS-1, also shares structural similarities with colicins suggesting it functions similarly to VasX (Russell et al. 2012). Both VasX and BTH_I2691 possess cognate immunity proteins (Russell et al. 2012; Miyata et al. 2013).

1.6.4 Nucleic acid-targeting effectors

Several T6SS effectors were demonstrated to be toxic in the cytoplasm of bacterial and eukaryotic cells and this is supported by the cytoplasmic localisation of their cognate immunity proteins. Some of these cytoplasmic-targeting effectors contain a nuclease domain fused to the C-terminus of recombination hot spot/tyrosine-aspartate (Rhs/YD)-repeats (GxxxRYxYDxxGRL(I/T). The Rhs/YD repeats form an extended β-sheet filamentous structure with a hollow shell that encapsulates and protects cells from damage by the toxins (Busby et al. 2013). Two Rhs/YD-repeat containing proteins RhsA and RhsB in *Dickeya dadantii* have been shown to be delivered in a T6SS-dependent manner and mediate growth inhibition of neighbouring cells. They both possess C-terminal nuclease domains that degrade DNA of the target bacterial cells (Koskiniemi et al. 2013). Rhs1 and Rhs2 in *S. marcescens* are also mostly involved in T6SS-dependent intraspecies competition (Alcoforado Diniz and Coulthurst 2015). Rhs2 comprises a C-terminal nuclease domain suggesting its DNase activity, whereas Rhs1 carries a novel anti-bacterial toxin domain, the cytoplasmic targets of which remain unclear (Alcoforado Diniz and Coulthurst 2015). Similarly, in most cases, the precise targets of these nucleic acid-targeting effectors have not been identified. Tse2 protein in *P. aeruginosa* H1-T6SS is toxic in the cytoplasm of both prokaryotic and eukaryotic cells and function as growth inhibitors towards other bacterial cells but not eukaryotes (Hood et al. 2010; Li et al. 2012). The target of Tse2 is currently unknown, but it is suggested to act as a nuclease effector based on structural homology predictions, although the experimental evidence is missing (Li et al. 2012).

1.6.5 Effector delivery

The molecular mechanism of effector delivery in a T6SS-dependent manner still remains elusive. It is suggested that the T6SS may employ several mechanisms to translocate effectors from the bacterial cytosol into the target host cells, as presented in Figure 1.18. It is also possible that multiple effectors are delivered simultaneously into the target host cells with one firing action of the T6SS apparatus (Shneider et al. 2013). Recent investigations highlighted two broad effector delivery mechanisms where effectors are

transported as either as specialised effectors (i.e. derivatives of T6SS subunits) or cargo effectors (free-standing effectors) which are generally mediated by TssI or TssD proteins. Small effectors may be delivered through/with the TssD tail tube, whereas large effectors may be translocated through binding to/as a TssI spike protein (Shneider et al. 2013).

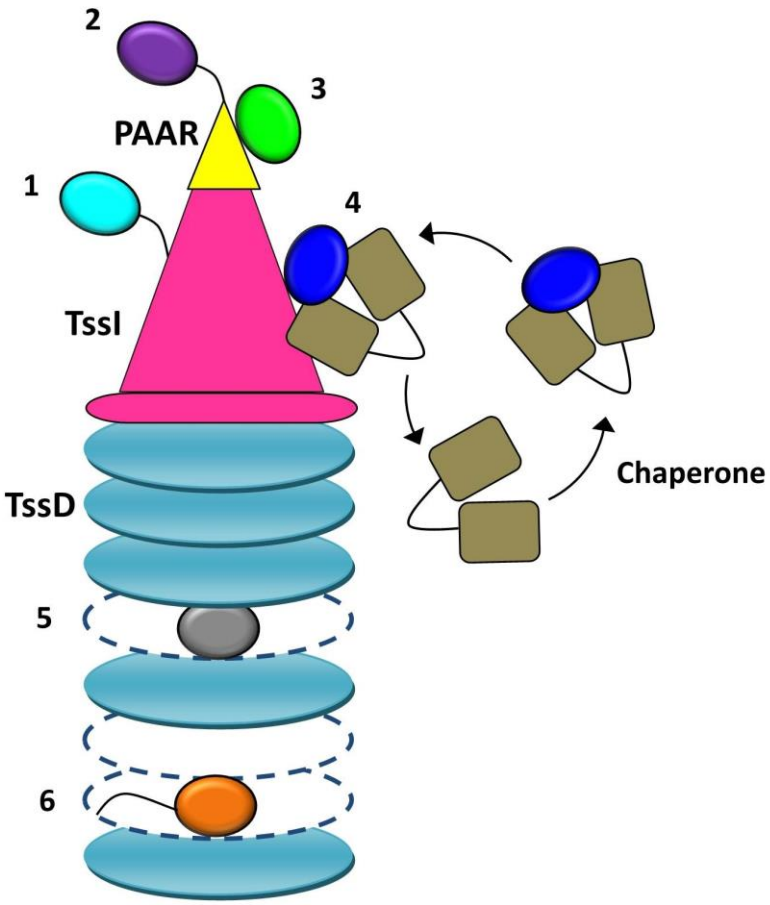


Figure 1.18 Model for T6SS effector delivery. Schematic drawing of six mechanisms by which effectors are predicted to be delivered by the T6SS. Mechanisms 1-4 are TssI-mediated effector delivery. Effectors delivered by (1) are specialised TssI proteins that carry C-terminal effector extensions, (2) specialised PAAR proteins that contain extension domains at the N- or C-termini, (3) cargo effectors that binds to a PAAR protein or the extension domains on a specialised PAAR protein, and (4) effectors that directly bind to the TssI protein or indirectly where the interaction is mediated by chaperone proteins. Chaperone proteins may not be secreted. Mechanisms 5 and 6 are TssD-mediated effector delivery. Effectors delivered by (5) are incorporated within the lumen of inner tube upon assembly of the T6SS apparatus, (6) specialised TssD proteins that carry C-terminal effector extensions.

1.6.5.1 TssI-mediated effector delivery

The TssI-mediated effector delivery mechanism includes specialised TssI proteins which carry C-terminal extension domains (Mechanism 1 in Figure 1.18) as previously discussed in Section 1.5.3.2.1.4. In addition, cargo effectors that bind to the TssI proteins to facilitate being transported may also be included in this category. In addition, the PAAR repeat proteins form conical structures that sharpen the TssI spike are also important for effector secretion and function.

The PAAR repeat proteins make up a diverse superfamily. Many PAAR proteins contain extension domains at the N- or C- termini, some of which are predicted to exhibit enzymatic activities, i.e. nuclease activities, that are toxic to target host cells (Mechanism 2 in Figure 1.18). These specialised PAAR proteins include RhsA and RhsB of *D. dadantii* (Koskiniemi et al. 2013). Rather than functioning directly as effectors, some of these extension domains of PAAR proteins, such as the transthyretin domain (TTR), might be binding partners of effectors that act as a connector or adaptor between TssI and effectors (Mechanism 3 in Figure 1.18) (Shneider et al. 2013). However, experimental evidence for direct interactions between PAAR proteins and effectors is lacking.

Recent studies have provided evidence for the presence of a superfamily of T6SS effector chaperone (TEC) proteins (also known as Tap, T6SS adaptor protein) that are essential for effector delivery via binding to TssI proteins. This superfamily comprises a highly conserved domain with unknown function, named DUF4123, and is generally encoded upstream of its cognate effector gene. The TEC protein in *V. cholerae* encoded by VC1417 is essential for the translocation of TseL ($Tle2^{VC}$), a T6SS-dependent effector involved in bacterial killing, by mediating the interaction between TseL and TssI-1 (Liang et al. 2015; Unterweger et al. 2015). Instead of being secreted with effectors into the target host cells, the TEC proteins are retained in the bacterial cell after mediating the interaction between effectors and TssI proteins, and it is suggested to be released once the effector has been loaded onto the TssI spike (Mechanism 4 in Figure 1.18) (Unterweger et al. 2015).

A TEC protein specific for the Rhs1 effector toxin in *S. marcescens*, named EagR1, has also been identified, which plays a role in Rhs1 secretion and Rhs1-dependent anti-bacterial activity (Alcoforado Diniz and Coulthurst 2015). EagR1 is not unique to *S.*

marcescens, it is homologous to proteins encoded immediately upstream of certain T6SS effectors in a variety of bacteria. These effectors includes an anti-bacterial PAAR-containing T6SS effector Tse6 (PA0093) of *P. aeruginosa* which has a TEC protein (PA0094) encoded by the gene located directly upstream of PA0093 (Hachani et al. 2014; Whitney et al. 2014).

1.6.5.2 TssD-mediated effector delivery

TssD-mediated effector delivery is the other broad effector delivery mechanism. TssD forms a hollow inner tube with 40 Å in diameter. This would potentially allow the passage of small folded, partially folded or unfolded effectors through the inner tube (Mougous et al. 2006; Ballister et al. 2008; Brunet et al. 2014). However, instead of translocation through the channel-like inner tube, it has been suggested that some small folded effectors, i.e. Tse1-4 secreted by *P. aeruginosa* H1-T6SS (Tse1 and -3 are PG hydrolases), directly interact with the inner side of the TssD hexameric ring structures and are stabilised by this interaction. Therefore, TssD serves as a chaperone protein for effectors, as well as incorporating the cargo effectors within the lumen of the inner tube upon assembly of the T6SS apparatus (Mechanism 5 in Figure 1.18) (Silverman et al. 2013; Whitney et al. 2014). The cargo effectors are then released into the target host cells along with TssD proteins upon firing action of the T6SS apparatus (Silverman et al. 2013). Evidence for this effector delivery model is also supported by other small effectors such as EvpP in *Edwardsiella tarda* (with unknown function) which was observed to be in association with TssD (Hcp) following secretion (Zheng and Leung 2007).

TssD may mediate T6SS effector delivery in the form of specialised TssD proteins (also known as evolved TssD), which contain a C-terminal enzymatic extension that was first identified in *S. enterica* subspecies *arizonae* in which TssD is fused to the N-terminus of a S-type pyocin and a HNH endonuclease domain (Mechanism 6 in Figure 1.18) (Blondel et al. 2009). However, so far, there is no experimental evidence to confirm the putative anti-bacterial activity of specialised TssD proteins as a functional T6SS subunit.

1.7 Aims and objectives

In order to investigate one of the least characterised T6SS core components, TssA, as well as two other core subunits TssE and TssK, the following aims and objectives are proposed:

1. Identify novel *in vivo* interactions between two types of TssA, i.e. $TssA^{S}$ (exemplified by the *B. cenocepacia* TssA subunit) and $TssA^{EI}$ (exemplified by the *A. hydrophila* TssA subunit) and other T6SS subunits to reveal cellular location of TssA in the T6SS apparatus

2. Quantify the efficiency of complementation for representative combinations of $TssA^{S}$ two-hybrid system fusions with other T6SS subunit fusions for validating the *in vivo* results

3. Provide biochemical evidence to support the *in vivo* protein-protein interaction results involving both $TssA^{S}$ and $TssA^{EI}$

4. Structurally characterise two types of TssA, i.e. $TssA^{S}$ and $TssA^{EI}$, for predication of the potencial role of TssA combined with its cellular location

5. Identify novel protein-protein interactions among other T6SS core components to reveal how T6SS is assembled

6. Structurally characterise TssE and TssK

Chapter 2 Materials and Methods

2.1 Bacterial strains and plasmids

Lists of bacterial strains and plasmids used in this study are shown in Appendices 1 and 2, respectively. The bacterial strains used in this study were streaked from -80 ℃ glycerol stocks (15% (v/v) glycerol) on appropriate agar plates (*B. cenocepacia* strains were streaked on M9 minimal agar plates with 0.5% glucose; *E. coli* strains were streaked on LB agar plates) and grown at temperatures between 30 ℃ and 37 ℃ for 16 to 120 hours dependent on optimal culturing conditions for the species and phenotype observation. After incubation, plates were sealed with parafilm and stored in the dark at room temperature for *B. cenocepacia* strains or 4 ℃ for *E. coli* strains for further strain maintenance.

Glycerol stocks were prepared for long-term strain maintenance. 0.7 ml of overnight culture grown in LB was added to 0.3 ml of sterile 50% glycerol in a cryogenic vial, to give a final concentration of ~15% glycerol. The vial was gently inverted several times and stored frozen at –80 ℃.

2.2 Bacteriological techniques

2.2.1 Bacteriological Media

2.2.1.1 Lysogeny broth (LB) broth

10 g Bacto™ Tryptone, 5 g yeast extract (Difco) and 10 g NaCl were dissolved in ddH_2O and the final volume was adjusted to 1 litre before being divided into 200 ml aliquots. The resultant broth was autoclaved immediately for 20 minutes at 120^{o}C/16 psi.

2.2.1.2 Brain-Heart Infusion (BHI)

37 g of BHI powder (BBL™) was dissolved in ddH_2O followed by adjusting the volume to 1 litre. The resultant BHI broth was then autoclaved immediately for 20 minutes at 120^{o}C /16 psi.

2.2.1.3 Dialysed Brain-Heart Infusion (D-BHI)

To reduce the interference of larger molecular weight proteins present in the BHI when analysing proteins in spent culture supernatants, D-BHI was used. For 1 litre of D-BHI, 18.5 g of BHI (BBL^{TM}) was dissolved in ddH_2O and made up to 50 ml, which was then dialysed against 500 ml of ddH_2O overnight at 4 ℃ using 12000-14000 MWCO dialysis tubing. Following this, the solution was sterilised by autoclaving for 20 minutes at 120^{o}C/16 psi.

2.2.1.4 Lennox broth + glucose

This corresponds to Lennox broth (Lennox 1955) with the addition of glucose. 10 g Bacto^{TM} Tryptone, 5 g yeast extract (Difco), 5 g NaCl and 2 g glucose were dissolved in ddH_2O and made up to 1 liter before dividing into 200 ml aliquots. The resultant broth was autoclaved immediately for 20 minutes at 120^{o}C/16 psi.

2.2.1.5 AUTO induction media

ZYM-5052 media (Studier 2005) was employed for inducing protein expression from vectors that employ the T7 promoter system without the need for IPTG when protein was not expressed in routine media, such as BHI.

2.2.1.6 LB agar

LB agar was made by adding 1.5% (w/v) bacteriological agar (Oxoid) into LB broth before autoclaving for 20 minutes at 120^{o}C/16 psi. The LB agar was allowed to cool down to ~50^{o}C, and appropriate antibiotics and other supplements were added prior to being poured into 90 mm Petri-dishes (25 ml per dish).

2.2.1.7 MacConkey maltose agar

40 g of MacConkey Agar base (Difco) was suspended in ddH_2O and made up to 1 liter. The resultant medium was sterilised by autoclaving for 20 minutes at 120^{o}C 16 psi. Sterilised glucose-free maltose 1% (w/v), 100 μg/ml ampicillin and 50 μg/ml kanamycin were added to the resultant medium once it had cooled to 50^{o}C for being

poured into 90 mm Petri-dishes.

2.2.1.8 M9 minimal salts agar

1 g D-glucose (VWR International) and 1.5% (w/v) bacteriological agar (Oxoid) were added to 180 ml of ddH_2O, and sterilised by autoclaving. After cooling to ~55^oC, 20 ml of sterilised 10x M9 salts were added, and filter-sterilised (0.22 μm) $CaCl_2$ and $MgSO_4$ solutions were added to a final concentration of 100 μM and 1 mM, respectively. The solution was poured into 90 mm Petri-dishes (25 ml per dish).

10x M9 salts:
60 g Na_2HPO_4, 30 g KH_2PO_4, 10 g NH_4Cl and 0.5 g NaCl were dissolved in 1 L of ddH_2O, autoclaved for 20 minutes and stored at room temperature.

2.2.2 Media supplements

Antibiotics added to media for selecting and maintaining plasmids included 100 μg/ml of ampicillin, 50 μg/ml of kanamycin, 25 μg/ml (for *E. coli*) or 100 μg/ml (for *B. cenocepacia*) of chloramphenicol and 10 μg/ml of tetracycline. All of the antibiotics were filter sterilised (0.22 μm) and stored at -20℃. Additional media supplements included 0.1-1 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG), 40 μg/ml of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and 1% (v/v) of maltose, which were filter sterilised (0.22 μm) if applicable.

2.3 Recombinant DNA techniques

2.3.1 Primers

To amplify DNA containing specific genes or gene fragments, primers were designed to anneal to a specific region of the gene. Primers used for cloning purposes had restriction sites and a GC clamp included in the 5' ends of the primers. Primers utilised in this study are shown in Appendix 3.

2.3.2 Genomic DNA preparation

Genomic DNA from *B. cenocepacia* strains was prepared as the DNA template for PCR reactions. A few colonies were selected from the agar plate which was streaked from glycerol stock, and resuspened in 200 µl of TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA) in a screw capped microcentrifuge tube, which was then boiled in water bath for 10 minutes. Following this, the solution was then centrifuged for 5 minutes at 15,000 x *g* and the supernatant was transferred into a fresh microcentrifuge tube.

2.3.3 Plasmid DNA preparation: E.Z.N.A.^{TM} Plasmid Miniprep Kit I

The E.Z.N.A.^{TM} Plasmid Miniprep Kit I was empolyed as per the manufacturer's guidelines to extract plasmid DNA from *E. coli* for sequencing (Section 2.3.14).

2.3.4 Plasmid DNA preparation: alkaline lysis with phenol-chloroform method

This procedure was carried out when the plasmid DNA was to be employed for DNA transformation. A single colony was isolated from a freshly streaked selective plate, and inoculated into 3 ml LB containing appropriate selective antibiotic, then grown overnight at 37^{o}C shaking incubator. On the following day, 1.5 ml of the overnight culture was transferred into a microcentrifuge tube followed by centrifugation at 13,000 x *g* for 5 minutes to harvest the bacterial cells. The supernatant was removed and discarded. The cell pellet was resuspended thoroughly in 100 µl of Solution I (Table 2.3), and incubated on ice for 5-10 minutes prior to adding 250 µl of Solution II (Table 2.3). The mixture was gently mixed to obtain a clear lysate, and incubated on ice for 5 minutes. Following this, 150 µl of Solution III (Table 2.3) was added and mixed immediately by inverting the microcentrifuge tube several times until a flocculent white precipitate formed, which contains precipitated proteins, chromosomal DNA and other cell debris. The suspension was incubated on ice for another 5 minutes followed by centrifugation at 13,000 x *g* for 5 minutes. The supernatant (~400 µl) was transferred by carefully aspirating it into a new microcentrifuge tube. An equal volume (400 µl) of phenol-chloroform-isoamyl alcohol (25:24:1, Fisher), was added and the mixture was briefly but thoroughly mixed on a vortex to extract any remaining protein, following

which it a centrifugation step was carried out at 13,000 x *g* for 5 minutes. The upper aqueous phase was kept, transferred into a new microcentrifuge tube, and mixed with 800 µl of 100% ethanol. The solution was left at room temperature for at least 5 minutes to allow nucleic acids to precipitate, and was then centrifuged at 13,000 x *g* for 5 minutes. The supernatant was removed and discarded, and the pellet was washed with 1 ml of 70% ethanol followed by centrifugation at 13,000 x *g* for 5 minutes. The supernatant was carefully removed and discarded, and the pellet was allowed to air dry for 30 minutes at room temperature. The purified plasmid DNA was then resuspended in 50 µl DNAse-free ddH_2O and stored at -20^oC.

Table 2.3 Composition of solutions used for plasmid DNA preparation

	Solution I	**Solution II**	**Solution III**
Ingredient	50 mM glucose 25 mM Tris-HCl (pH 8.0) 10 mM EDTA (pH 8.0)	0.2 M NaOH 1 % (w/v) SDS	3 M potassium acetate 2 M acetic acid
Autoclave	Yes	No	Yes
Storage	4^oC	RT	4^oC

Removal of RNA

In order to remove RNA present in the plasmid DNA preparation, which would interfere with viewing small DNA fragments released by restriction digestion in agarose gels, 1 µl of 1 mg/ml RNAse A was added to the preparation and this was incubated at 37^oC for 1 hour.

2.3.5 Agarose gel DNA electrophoresis

To prepare an agarose gel, high specification grade agarose for electrophoresis (Webscientific) was dissolved in 1x TAE buffer (40 mM Tris-acetate (pH 8.0), 1 mM EDTA) by boiling in a microwave oven until completely dissolved. A standard gel was 0.8% (w/v), however, a higher percentage gel (1-1.5%) was also used to resolve small DNA fragments <500 bp. After boiling, the agarose solution was cooled down to ~50^oC before being poured into the gel tray with comb inserted. When the gel had solidified, it was placed in a gel tank (BioRad) which was filled with the required amount of 1x TAE buffer to submerge the gel. DNA samples were mixed with 6x DNA loading dye (30% (v/v) glycerol, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF) and

loaded into the wells cast in the gel. An appropriate DNA ladder was also loaded into a well. The voltage applied to the system depended on the length of the gel tank. Following separation of the samples, the gel was immersed in a 5 μg/ml solution of ethidium bromide for 15 minutes to allow intercalation of the ethidium bromide into the DNA within the gel. The DNA fragments were then visualized on a UV transilluminator at 260 nm (Kodak EDAS 290 gel documentation system).

2.3.6 Polymerase Chain Reaction (PCR) for cloning

A proof-reading DNA polymerase (KOD hot start, Novagen) was used for amplification when the products were to be used for cloning. Genomic DNA (Section 2.3.2) or plasmid DNA (Sections 2.3.3 and 2.3.4) was used as template DNA. To set up the reaction, the following components were combined in a 0.5 μl PCR tube (Table 2.4).

Table 2.4 Composition of polymerase chain reaction for cloning

PCR components	Amount (μl)
Template DNA	3 μl (~100 ng plasmid DNA)
KOD 10x buffer	5
$MgSO_4$ (25 mM)	4
dNTP mix (2 mM each)	5
DMSO	2.5
Forward primer (10 μM)	3
Reverse primer (10 μM)	3
KOD hot start DNA polymerase (Merck Millipore)	0.5
ddH_2O	24

PCR regime

An initial denaturation step was carried out at 95^{o}C for 2 minutes. The following cycles was conducted and repeated 30 times.

- Denaturation at 95^{o}C for 30 seconds
- Annealing for 30 seconds. Temperature depended on the length and G+C content of the primers, which was calculated based on the formula [2x(A+T) + 4x(G+C)]-5^{o}C where A+T is the total number of A and T bases and G+C in the

total number of G and C bases in the annealed segment of the primer.

- Elongation (polymerisation) at 72^{o}C for 1 minute per kilobase of amplified DNA.

2.3.7 PCR screening for recombinant plasmids

Transformants containing recombinant plasmids were screened for the desired inserted DNA by PCR screening with Taq DNA polymerase. The PCR was carried out as follows (Table 2.5).

Table 2.5 Composition of polymerase chain reaction for screening recombinant plasmids

PCR components	Amount (µl)
Genomic DNA	a small mass of cells from the transformant colony
Taq 10x buffer	5
$MgCl_2$ (50 mM)	2
dNTP mix (10 mM each)	1
DMSO	2.5
Forward primer (10 µM)	3
Reverse primer (10 µM)	3
Taq DNA polymerase (Yorkshire Bioscience)	1
ddH_2O	32.5

The PCR regime was carried out without a hot start, and Taq DNA polymerase was added directly before the initial denaturation step.

2.3.8 DNA gel extraction

The Qiagen QIAquick Gel Extraction Kit was used to extract DNA restriction fragments or PCR products from electrophoresed agarose gels according to the manufacturer's guidelines.

2.3.9 DNA purification

To inactivate interfering enzymes, remove unwanted primers and exchange buffers resulting from PCR and restriction digestion, DNA samples were purified using a spin-column-based purification kit according to the manufacturer's instructions (Thermo Scientific/Qiagen/Macherey-Nagel).

2.3.10 Restriction endonuclease digestion

Restriction endonuclease digestions were carried out using enzymes and buffers (Promega/NEB). 5-20 µl (100-1000 ng of DNA) of PCR product or plasmid DNA, 5 µl of 10 x restriction enzyme compatible buffer, 0.5 µl of Bovine serum albumin (BSA, stock 10 mg/ml) and 1 µl of enzyme were mixed in a 1.5 ml microcentrifuge tube. The total volume was made up to 50 µl using ddH_2O. When double digestion was performed (i.e. cleavage of DNA with two enzymes), an appropriate buffer was chosen to obtain optimum efficiency for both enzymes (otherwise reactions were performed sequentially following a DNA purification step (Section 2.3.9)). Digestion mixtures were incubated at 37^oC waterbath for 2 hours and purified as described in Section 2.3.9.

2.3.11 DNA filling-in reaction using DNA polymerase I Klenow fragment

In order to generate blunt ends from the ends containing 5' overhangs, DNA filling-in reaction was conducted using the Klenow (large) fragment of *E. coli* DNA polymerase I by adding complementary bases at the 3' ends. 5 µl of dNTPs (2 mM each), 5 µl of 10× reaction buffer, 100-1000 ng of DNA and 1 µl of diluted DNA polymerase Klenow fragment (Promega) were combined. Klenow DNA Polymerase was freshly diluted 50 times in dilution buffer (20 mM Tris-HCl (pH 8.0), 50 mM NaCl, 0.1 mg/ml BSA, 50% glycerol and 1 mM EDTA) from the stock (2.6 mg/ml). The total reaction volume was made up to 50 µl with ddH_2O, and carried out at room temperature for 15 minutes followed by a purification step as described in Section 2.3.9.

2.3.12 DNA Ligation

Linearised plasmid DNA and DNA fragment of interest that were digested using

appropriate restriction enzymes were ligated using T4 DNA ligase (Promega). The molar ratio between DNA fragment of interest and plasmid DNA was 3:1 or 5:1, respectively, for directional or blunt cloning. Along with 1 µl of the ligase (1 u/µl) and 3 µl of 10x ligase buffer, the total volume was adjusted to 30 µl with ddH_2O. Vector control (no DNA fragment of interest or ligase) and ligation control (no DNA fragment of interest, but with ligase) were included with the ligation reaction. Reactions were incubated at room temperature overnight, and used to transform *E. coli* competent cells (Section 2.3.15.1).

2.3.13 Dephosphorylation of DNA 5' ends

To prevent self-ligation of vector DNA fragments by T4 DNA ligase, calf intestinal alkaline phosphatase (CIP) was employed which catalyzes the removal of 5' phosphate groups from DNA. This procedure promotes ligation between the inserted fragment and vector DNA molecules.

5.5 µl of CIP 10x reaction buffer was combined with 50 µl of restriction digested plasmid DNA without DNA purification. 1 µl of CIP (1 u/µl, Promega) was then added to the mixture and incubated at 37 ℃ for 30 minutes. Following this, another equivalent aliquot of CIP was added to the mixture and the incubation was extended for another 30 minutes. The CIP was then removed by DNA purification as described in Section 2.3.9.

2.3.14 DNA sequencing

To confirm that no mutations were introduced during PCR amplification, the recombinant plasmid constructs were sequenced with primers that annealed to vector DNA sequences flanking the cleaved DNA. The size and concentration of the plasmid DNA was confirmed by electrophoresis in a 0.8% agarose gel before being submissed to the Core Genomic Facility, Medical School, University of Sheffield for nucleotide sequence determination. Plasmid DNA used for sequencing was prepared as described in Section 2.3.3 and 10 µl of 100 ng/ µl was submitted.

2.3.15 Techniques for plasmid transfer

2.3.15.1 Transformation

2.3.15.1.1 Preparation of competent cells

Competent cells were prepared by Hanahan's method (Hanahan 1983). The bacterial strain was inoculated into 3-5 ml LB and cultured overnight at 37^{o}C in a shaking incubator. On the following day, 0.5 ml of the overnight culture was inoculated into 50 ml LB and allowed to grow to exponential phase (OD_{600} 0.3-0.5) in a 37^{o}C shaking incubator. Following this, the culture was chilled on ice for 15 minutes prior to harvesting the bacterial cells by centrifugation at 4,000 x *g* for 10 minutes at 4^{o}C. The supernatant was removed and discarded, and the cell pellet was resuspended in 16 ml of ice cold RF1 solution. The cell suspension was then inbated on ice for 30 minutes followed by centrifugation at 4,000 x *g* for 10 minutes at 4^{o}C. The supernatant was removed and discarded, and the cell pellet was resuspended in 4 ml of cold RF2 solution. The cell suspension was incubated on ice for 15 minutes prior to being divided into 0.5 ml aliquots. The competent cells were stored at -80^{o}C.

RF1 Solution:

100 mM KCl, 50 mM $MnCl_2.4H_2O$, 30 mM CH_3COOK, 10 mM $CaCl_2.2H_2O$ and 15% (v/v) glycerol were dissolved in ~750 ml ddH_2O, and the pH of the solution was adjusted to 5.8 using 0.2 M acetic acid. The final volume of the solution was then made up to 1 litre with ddH_2O. The solution was sterilised by autoclaving and stored at 4^{o}C.

RF2 solution:

Solution A contains 0.5 M MOPS (4-morpholinepropanesulfonic acid), pH 8.0; Solution B is comprised of 75 mM $CaCl_2.2H_2O$, 10 mM KCl and 15% (v/v) glycerol. The RF2 solution was made by mixing 0.2 ml of solution A and 9.8 ml of solution B on the day of use.

2.3.15.1.2 Heat-shock assisted Transformation

Frozen competent cells were stored at -80 ℃ and thawed on ice. Once defrosted, 100 µl of the competent cells were added into chilled DNA samples (1-2 µl of plamid DNA or

15 µl of a ligation reaction) and mixed gently. The mixture was incubated on ice for 30 minutes with agitation every 5 minutes, then transferred into a 42 ℃ waterbath for 2.5 minutes. The mixture was then chilled on ice for 5 minutes. 1 ml of LB was added to the mixture, and gently shaken before being incubated at 37 ℃ for 1 hour to allow the expression of antibiotic resistance genes present in the plasmid. 100 µl of the transformation mixture were spread onto LB, M9 or MacConkey agar plates containing appropriate antibiotic(s) to allow selection of the desired plasmid. A cell control which contained only competent cells (no plasmid DNA) was also spread on an antibiotic containing plate to check for contamination. Plates were incubated at 30 ℃ for MacConkey agar plates or 37 ℃ for LB and M9 plates, and the length of incubation varies among bacterial strains.

2.3.15.2 Conjugation

5 ml LB containing appropriate antibiotic(s) were inoculated with the donor strain (i.e. *E. coli*) containing the plasmid and the recipient strain (i.e. *B. cenocepacia*). The cultures were grown overnight in a 37 ℃ shaker incubator. 1 ml of each culture was harvested by centrifugation at 15,000 x *g* for 2 minutes in screw capped microcentrifuge tubes. The supernatant was removed and discarded, and the cell pellets were resuspended in 100 µl 0.85 % (w/v) sterile saline. Using sterile forceps, 0.45 µm nitrocellulose filter membrane (Millipore) were placed on LB agar plates. For the conjugation, 25 µl of donor and 25 µl of the recipient cultures were gently mixed in a screw capped microcentrifuge tube. As controls, 25 µl of the donor and recipient cultures were mixed separately with 25 µl 0.85 % (w/v) sterile saline in a screw capped microcentrifuge tube. Each of the cultures was pipetted directly onto separate filters and spread over the membrane using a pipette tip. When the filters had dried, the plates were inverted and incubated at 37 ℃ for 8-16 hours. The filters with bacterial growth were then transferred to sterile universal tubes containing 3 ml sterile 0.85 % (w/v) saline, and vortexed. 100 µl of the cultures and 10^{-1}, 10^{-2} and 10^{-3} dilutions were spread onto appropriate selection plates and incubated at 37 ℃ for 48-72 hours.

2.4 Recombinant protein overproduction and purification techniques

2.4.1 Protein expression

To determine the expression of recombinant proteins by a phage T7 promoter based expression plasmid, a fresh transformed colony of *E. coli* expression host strain containing the plasmid was inoculated in 50 ml of BHI (or Lennox broth with glucose for MBP fusion recombinant proteins) containing appropriate antibiotic(s) and grown at 22/30/37 ℃ in a shaking incubator until an OD_{600} of 0.5-0.8 was reached. The bacterial cell culture was then induced by IPTG to a final concentration of 0.1-1 mM for a further 2-3 hours. Before and after induction samples (1 ml) were collected for SDS-PAGE analysis. Samples were prepared by centrifugation at 13,000 x *g* for 3 minutes, and the supernatant was discarded and cell pellet was resuspended in a small volume of ddH_2O (30-45 µl) to which an equal volume of Sample Buffer (0.125 M Tris-HCl (pH 6.8), 4% SDS, 0.02% (w/v) bromophenol blue, 20% glycerol and 4-5% β-mercaptoethanol added freshly) was added and boiled for 10 minutes to denature and solubilise proteins.

To harvest the bacterial cell culture, the cells were incubated on ice for 15 minutes followed by centrifugation at 12,500 x *g* for 20 minutes at 4 ℃. The supernatant was removed and discarded, and the cell pellet was washed by Tris-buffered saline (TBS), followed by recentrifugation at 12,500 x *g* for 20 minutes at 4 ℃. The supernatant was removed and discarded, and the cell pellet was weighed and stored at -20 ℃.

2.4.2 Determination of recombinant protein solubility

The bacterial cell pellet was resuspended in lysis buffer (50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% (v/v) glycerol (if necessary) with or without 10 mM imidazole depending on whether IMAC was to be employed as the following step) at a ratio of 1:5 (1g cell pellet: 5 ml lysis buffer). Fresh lysozyme (~1 ml of 10 mg/ml ddH_2O dissolved lysozyme per 25 ml of the cell suspension) was added to the suspension and incubated at 4 ℃ on a roller for 30 minutes to lyse the bacterial cell wall. Following this, sodium deoxycholate was added to a final concentration of 500 µg/ml from a stock solution of 25 mg/ml (Sigma, ddH_2O dissolved) and serine protease inhibitor phenylmethanesulphonylfluoride (PMSF) was added to a final concentration of 25 µg/ml from a stock of 2.5 mg/ml (Applichem, 100% ethanol dissolved) and the

suspension was incubated at 4℃ for a further 20-30 minutes to lyse the cells and solubilise cellular and membrane components. The sample was then subjected to sonication for 30 seconds each time with 2-minute incubation on ice between each sonication until the viscosity of the solution was reduced (amplitude <21%, MSE soniprep 150 plus). 50 µl of crude sample was taken of the sonicated sample and combined with 50 µl of 2x sample buffer, to give the total protein fraction. The remainder of the crude protein suspension was centrifuged at 35,000 x *g* for 30 minutes at 4^{o}C. The supernatant was carefully transferred to a universal tube and 50 µl of sample was taken and combined with 50 µl of 2x Sample Buffer. The samples were boiled and analysed by SDS-PAGE. Overproduced proteins present in the supernatant of the cleared lysate were considered soluble and were ready to be subjected to chromatography.

2.4.3 PEI precipitation

5% polyethyleneimine (PEI) precipitation at pH 8.0 was applied to the clarified supernatant of lysed *E. coli* BL21(λDE3) cells containing overproduced recombinant protein to remove nucleic acids, followed by centrifugation at 40,000 x *g* for 30 minutes at 4^{o}C. The supernatant was carefully transferred to a clean universal tube.

2.4.4 Ammonium sulfate precipitation

Analytical ammonium sulfate cuts were performed on a small amount of aliquoted clarified supernatant of lysed *E. coli* BL21(λDE3) cells containing overproduced recombinant protein by adding increasing amounts of saturated ammonium sulfate solution to the lysate to achieve final concentrations of ammonium sulphate ranging from 20% to 70% to precipitate the protein of interest. Following centrifugation at 40,000 x *g* for 10 minutes, the clarified supernatant was transferred to a fresh tube and pellets were dissolved in 10-50 µl of buffers that the protein was in. The volume of clarified supernatant was determined and ammonium sulfate powder was added to increase the final concentration with 10% for each sample. Centrifugation step was repeated and both pellet and supernatant samples were collected for SDS-PAGE analysis. Therefore, samples were analysed with 20-30%, 30-40%, 40-50%, 50-60%, 60-70% and 70-80% ammonium sulphate precipitation. The initial ammonium sulfate

concentration that the recombinant protein was soluble but precipitated most of the recombinant protein after a 10% increase was applied to the remaining clarified cell lysate. The sample was incubated at 4^{o}C for 30 minutes on roller followed by centrifugation at 40,000 x *g* for 30 minutes and another step of ammonium sulphate precipitation with 10% concentration increase using the supernatant fraction. The precipitated protein was recovered by centrifugation at 40,000 x *g* for 30 minutes at 4^{o}C. The protein pellet was resuspended in a small volume of desired buffer, as appropriate for the subsequent purification step.

2.4.5 Immobilized metal affinity chromatography (IMAC)

The clarified supernatant of lysed *E. coli* BL21(λDE3) cells containing His-tagged overproduced recombinant protein was prepared as described in Section 2.4.2. A charged 1 ml HisTrap HP column (GE Healthcare) was manually equilibrated with binding buffer (50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole) using a syringe. The clarified protein lysate was passed through a syringe-driven 0.22 μM filter unit (Millipore) and manually loaded onto the column at a flow rate of ~1 ml/min. The flow-through was collected and re-loaded onto the column to maximise target protein binding if necessary depending on the protein binding affinity. Following this, the column was attached to a buffer-equilibrated ÄKTA purifier system (pump A, binding buffer (containing 10 mM imidazole); pump B, binding buffer containing 500 mM imidazole) and was washed with 10 column volumes (i.e. ~10 ml) of binding buffer to remove unbound or non-specifically bound proteins. To elute the bound His-tagged protein, a 10-500 mM imidazole gradient was applied over the column in a 30 ml volume with a 1 ml/min flow-rate. Eluted fractions were collected in 1 ml aliquots automatically using a fraction collector Frac-950 (GE Healthcare) and the elution process was monitored by UV absorption at 280 nm (the absorbance wavelength of tryptophan and tyrosine residues). Fractions corresponding to the protein elution peak on the UV trace was collected and analysed by SDS-PAGE.

2.4.6 Ion exchange chromatography

The clarified supernatant of lysed *E. coli* BL21 (λDE3) cells containing overproduced recombinant protein was prepared as described in Section 2.4.2 with minor alterations.

A 1 or 5 ml HiTrap DEAE Sepharose FF column (GE Healthcare) was equilibrated with binding buffer (50 mM Tris-HCl (pH 8.0)), and the clarified protein lysate was passed through a syringe-driven 0.22 μM filter unit (Millipore) and manually loaded onto the column at a flow rate of 1 ml/min. Following application of the sample, the column was attached to a buffer-equilibrated ÄKTA purifier system (pump A, binding buffer; pump B, binding buffer containing 500 mM NaCl) and was washed with 10 column volumes (i.e. ~10 ml) of binding buffer to remove unbound or non-specifically bound proteins. To elute the protein of interest, a 10-500 mM NaCl gradient was applied over the column in a 30 ml volume with a 1 ml/min flow-rate. Eluted fractions were collected in 1 ml aliquots automatically using a fraction collector Frac-950 (GE Healthcare) and the elution process was monitored by UV absorption at 280nm. Fractions corresponding to the protein elution peak on the UV trace was collected and analysed by SDS-PAGE.

2.4.7 Amylose affinity chromatography

To purify maltose-binding protein- (MBP-) fusion proteins, the clarified supernatant of lysed *E. coli* NEB express cells (New England Biolabs) containing overproduced recombinant protein was prepared as described in Section 2.4.2 with minor alterations. 12 ml of amylose resin was packed in micro bio-spin columns (BioRad) and equilibrated with column buffer (50 mM Tris-HCl (pH 8.0), 200 mM NaCl). The clarified cell lysate was loaded onto the column under gravity, the flow-through was retained and re-loaded onto the column to maximise target protein binding if necessary depending on the protein binding affinity. The column was washed with 2 column volumes (i.e. ~24 ml) of column buffer to remove unbound or non-specifically bound proteins. To elute the protein of interest, 3 column volumes (36 ml) of column buffer containing 10 mM maltose was applied onto the column under gravity. Eluted fractions were collected in 3 ml aliquots and analysed by Bradford assay to determine fractions containing the majority of the protein of interest. These fractions was then collected and analysed by SDS-PAGE.

2.4.8 Size exclusion chromatography (SEC)

Size exclusion chromatography was used to separate proteins based on their molecular weights. Superose 6 10/300 GL, Superose 12 10/300 GL, Superdex 200 HiLoad 16/600

and Superdex 200 10/300 GL columns supplied by GE healthcare were employed depending upon their appropriateness and availability. The column was attached to an ÄKTA purifier system and equilibrated with 50 mM Tris-HCl (pH 8.0), 500 mM NaCl with a flow-rate of 0.5 ml/min. Following equilibration, concentrated protein lysate was loaded into the column by injection into a superloop. A constant flow of 0.5 ml/min was applied across the column, and 0.5-2 ml fractions were collected. The elution of protein(s) from the column was monitored by the absorbance at 280 nm. The peak elution time/volume was recorded and used to determine the molecular weight based on a series of protein standards. Fractions corresponding to the protein elution peak on the UV trace were collected and analysed by SDS-PAGE.

In the case of estimation of the apparent molecular weight of proteins, a calibration kit (GE Healthcare) of protein standards with known molecular weights from low or high (13.7 to 2,000 kDa) was used for calibration of the gel filtration columns as per the manufacturer's instructions with slightly alterations. Standards were combined and dissolved in 1 ml equilibration buffer at a concentration of 1 mg/ml, followed by concentration into 100 µl using a 10 kDa cut-off Amicon Ultra-4 centrifugal filter unit (Millipore). The concentrated standards were loaded onto the gel filtration column and the elution time/volume monitored by UV absorbance was recorded for all the standards. Kav value for each standard was calculated (excluding Blue Dextran 2,000 kDa) using the equation Kav = (Ve-Vo)/(Vt-Vo), where Ve is elution volume, Vo is the column void volume determined by the point of elution of Blue Dextran (~2,000 kDa) and Vt is geometric column volume. The calibration plot was presented as Kav (X axis) versus logMW (Y axis). A semi-log regression curve was fitted to the data points, resulting in a line equation of $y=10^{(mx + c)}$. The peak elution volume of the protein of interest was determined and used to calculate its Kav value, which was then used to estimate its molecular mass using the calibration curve.

2.4.9 Buffer exchange

To remove interfering components within the buffer during or after protein purification steps, such as imidazole, NaCl or ammonium sulfate, protein samples were subjected to buffer exchange into the desired buffer using Zeba™ Desalt Spin columns or dialysis tubing (see below).

2.4.9.1 Zeba™ Desalt Spin columns

Zeba™ Desalt Spin columns (2 ml, Thermo Scientific) were used for buffer exchange of samples of small volume (200-700 µl). The bottom seal of the column was twisted off and the cap was loosened. The column was placed in a collection tube prior to centrifugation at 1,000 x *g* for 2 minutes at 4^{o}C to remove storage solution. 1 ml of the desired buffer was then added to the column followed by another centrifugation step. The equilibration step was repeated for an additional three times. The column was then placed in a new collection tube and samples were applied to the centre of the compact resin bed which was collected by another centrifugation step. The column was discarded after use.

2.4.9.2 Dialysis tubing

For larger volumes of protein samples (>2 ml), dialysis against the desired buffer using dialysis tubing was employed for buffer exchange if the subsequent step was not size exclusion chromatography. Dialysis tubing (MWCO 3 kDa) was cut to the appropriate size, soaked in ddH_2O for ~15 minutes and then equilibrated in the desired buffer. Protein samples were transferred into the dialysis tubing and the ends were clamped securely. After that, the dialysis tubing was placed in a beaker containing the desired buffer (100 times greater than the sample volume) with a magnetic rod stirring at 4^{o}C for 16 hours with an additional change of the buffer.

2.5 Native protein extraction and purification techniques

2.5.1 Secreted protein extraction from broth cultures

To extract proteins present in the broth culture supernatant, 20 ml D-BHI media (containing antibiotics if appropriate) was inoculated with an overnight culture to give an initial starting OD_{600} of 0.03. The culture was incubated at 37 ℃ with shaking until an OD_{600} of 0.6-0.8 was reached. Cultures were then centrifuged at 4,000 x *g* for 20 minutes at 4 ℃ to harvest the bacteria, following which the supernatant was filter sterilised (0.22 µm, Millipore). Proteins present in the supernatant were precipitated by DOC-TCA (Section 2.5.2). For cell-associated protein fractions, the subsequent cell pellet recovered from the above procedure was gently resuspended in 1 ml of 0.85%

(w/v) saline. 25 µl of this cell suspension sample was then combined with an equal volume of 2x Sample Buffer and boiled for 10 minutes to lyse the cells. Samples were then analysed by SDS-PAGE.

2.5.2 DOC-TCA protein precipitation

For protein precipitation of samples with low protein content, sodium deoxycholate solution (20 mg/ml stock, Sigma) solution was added to the filter sterilised supernatant to a final concentration of 0.2 mg/ml. The mixture was then chilled on ice for 30 minutes. 100% (w/v) chilled trichloroacetic acid (TCA) solution (Sigma) was then added to the supernatant to a final concentration of 10% (v/v) and the solution was incubated at -20 ℃ overnight. On the following day, the frozen suspension was thawed on ice and proteins were recovered by centrifugation at 4,000 x *g* for 1 hour at 4 ℃ in an Allegra X-22R benchtop refrigerated centrifuge. The supernatant was carefully removed and discarded. 2 ml of ice-cold acetone (Sigma) was added to the protein pellet, vortexed and incubated for 1 hour at -20 ℃. Following this, samples were centrifuged at 4,000 x *g* for 1 hour at 4 ℃. The supernatant was removed and discarded, and the pellet was allowed to air-dry for 30 minutes to remove residual acetone. The protein pellet was then resuspended in ddH_2O at a final volume of 2000-fold less than the initial culture supernatant volume. An equal volume of 2x sample buffer was added to the sample, 0.5-1.0 µl of 1M NaOH may need to be added to the solution to adjust the pH of the sample. The sample was then boiled for 10 minutes and analysed by SDS-PAGE.

2.6 Protein quantification and separation techniques

2.6.1 Determination of protein concentration

To quantify the concentration of protein in a solution, the Bradford assay was employed using a Bio-rad protein assay kit I. A series of protein standards (0.2, 0.4, 0.6, 0.8 and 1 mg/ml) was prepared using 10 mg/ml BSA (Promega) diluted in an appropriate buffer (such as 50 mM Tris-HCl (pH 8.0), 200 mM NaCl). The protein sample to be determined was also diluted in the same buffer in a final volume of 20 µl. The dye reagent was prepared by diluting 1 part Dye Reagent Concentrate (Bio-rad) with 4 parts ddH_2O and filter sterilised (0.22 µM, Millipore). 20 µl of protein standards or protein

sample was mixed with 1 ml of diluted dye reagent in the disposable cuvettes. After 5 minutes incubation at room temperature, the absorbance at 595 nm was measured using 1 ml of diluted dye reagent as reference. Protein samples with OD_{595} >0.9 (over the range of the spectrophotometer) were diluted further and re-tested with fresh dye reagent. The protein concentration (mg/ml, X axis) versus OD_{595} using protein standards was plotted and a standard curve was generated. The protein concentration to be determined was then calculated with consideration of the initial dilution of the protein sample.

2.6.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed using a Mini-PROTEAN-II or Tetra Cell system (Bio-Rad) with acrylamide/Bis-acrylamide 29:1 (40% solution/electrophoresis, Fisher). Protein samples were resuspended in 2x Sample Buffer and denatured by boiling at 100^{o}C for 10 minutes on a heat block. Samples were centrifuged at 15,000 x *g* for 10 minutes, and 10-15 µl of the supernatant were loaded onto a discontinuous gel consisting of a resolving gel and a stacking gel (Table 2.6) (Laemmli 1970). Protein molecular weight markers (E.Z-Run unstained/pre-stained *Rec* protein ladder, Fisher) were loaded next to the samples as a reference. The samples were electrophoresed in running buffer (25 mM Tris, 192 mM glycine, 1% (w/v) SDS) until the front of the dye reached the end of the gel. The gels were then rinsed with water and stained with 2.5% (w/v) Coomassie brilliant blue R250 in 45% (v/v) methanol, 45% (v/v) ddH_2O, 10% (v/v) glacial acetic acid for ~30 minutes depending on the colour development, and destained using 40% (v/v) methanol, 50% (v/v) ddH_2O, 10% (v/v) glacial acetic acid) until protein bands can be visualised.

Table 2.6 Recipe of SDS-PAGE for 4 gels

	Resolving gel			Stacking gel
Components	**10%**	**12%**	**15%**	**5%**
ddH_2O (ml)	9.6	8.6	7.1	6.3
1.5 M Tris-HCl (pH 8.8) (ml)	5	5	5	-
1.0 M Tris-HCl (pH 6.8) (ml)	-	-	-	1.25
Acrylamide mix (ml)	5	6	7.5	1.25
10% SDS (w/v) (µl)	200	200	200	100
10% Ammonium persulphate (w/v) (µl)	200	200	200	100
TEMED (µl)	10	10	10	5
Final volume (ml)	20	20	20	9.0

2.6.3 Western blotting

Protein samples were fractionated by SDS-PAGE with the E.Z-Run pre-stained *Rec* protein ladder (Fisher) as reference (Section 2.6.2). A 0.45 µm PVDF membrane (Millipore) was soaked in 100% methanol for 2 minutes and washed in ddH_2O for another 2 minutes with shaking. The SDS-PA gel, blotting paper, pad and PVDF membrane were soaked in transfer buffer (400 ml ddH_2O containing 2.9 g Tris and 1 g glycine) with 10% methanol for 15 minutes. The western blotting transfer cassette was assembled in the order from cathode to anode: pad, blotting paper, gel, PVDF membrane, blotting paper and pad. Excess buffer and air bubbles were pressed out of the blotting paper in order to obtain better transfer of protein from gel to membrane. The assembled cassette was placed in the transfer tank with transfer buffer containing 10% methanol and transferred at 100 V for 60 minutes. During the electrotransfer, an ice cassette was also placed in the transfer tank to keep the apparatus cool. Alternatively, Trans-Blot turbo mini PVDF transfer packs (0.2 µm, Bio-rad) were assembled with the SDS-polyacrylamide gel and electro-blotted on a Trans-Blot turbo transfer system (Bio-rad) according to the manufacturer's instructions. Following transfer, the membrane was blocked in TBS-T (137 mM NaCl, 25 mM Tris-HCl and 3 mM KCl (pH 8.0) containing 0.05% (v/v) Tween20) with 5% (w/v) milk powder at room temperature for 1 hour, and probed with the primary antibody at the desired dilution in TBS-T containing 5% milk powder at 4^oC overnight on a roller or shaker. The membrane was

then subjected to 4 times washes (10 minutes each) with TBS-T on a shaker and probed with the secondary antibody (as appropriate) linked to HRP diluted in TBS-T containing 5% (w/v) milk at room temperature for 1 hour. This was followed by another four 10-minute washes in TBS-T at which point the membrane was ready to apply the detection reagent (ECL, Biological Industries) as manufacturer's instructions, and chemiluminescence was detected using a Bio-rad molecular imager ChemiDoc^{TM} XRS+ with Image Lab software.

2.6.4 Antibodies

Antibodies and detection reagents used in this study are shown in Table 2.7. Apart from commercially available antibodies, custom polyclonal antibodies were also generated using purified proteins. Purified proteins were sent to Bioserve UK, University of Sheffield, for polyclonal antibody generation using rats (one rat per protein). According to the service licence, each rat can be given up to five immunizations with 100 μg of antigen. First bleed blood samples were collected after primary injection and antibody titre was estimated by western blotting. Different dilutions of first bleed blood samples were used as primary antibodies and commercial anti-rat antibody linked with HRP was used as the secondary antibody. Upon analysis of first bleed blood samples, the requirement for a boost injection and terminal bleed ten days later was determined.

Table 2.7 Antibodies used in this study

Antibody	Species	Working concentration	Company
His-probe linked with HRP	-	1:5,000	Thermo Scientific
Anti-VSVG polyclonal antibody	rabbit	1:5,000	Sigma-Aldrich
Anti-$TssA^S$ polyclonal antibody	rat	1:2,000	BioServe
Anti-TssK polyclonal antibody	rat	1:1,000	BioServe
Anti-rat antibody linked with HRP	goat	1:5,000	Southern Biotech
Anti-rabbit antibody linked with HRP	goat	1:5,000	Vector Labs
Anti-HA	rabbit	1:5,000	Cell Signaling
Anti-FLAG	rabbit	1:5,000	Sigma
Anti-TssD	rat	1:2,500	BioServe
Anti-His_6	mouse	1:5,000	BioLegend
Anti-MBP linked with HRP	mouse	1:5,000	Biolabs
Anti-$TssA^S$	rat	1:5,000	BioServe

2.7 Techniques for analysis of protein–protein interactions

2.7.1 Bacterial adenylate cyclase two hybrid assay (BACTH)

The BACTH system allows the study of protein-protein interactions *in vivo*. This method is based on reconstitution of adenylate cyclase activity (cyclic AMP (cAMP) production from ATP) in a *ΔcyaA E. coli* strain, i.e. BTH101, by the formation of an active adenylate cyclase domain from *Bordetella pertussis* CyaA. This occurs by interaction between two complementary fragments (T25 and T18) of CyaA (Karimova et al., 1998). In this system, various genes of interest were fused to the N- or C-terminus of the T25 and T18 fragments by cloning the corresponding DNA into the four BACTH vectors pKT25, pKNT25, pUT18 and pUT18C (Section 3.1). An *E. coli* strain BTH101 was transformed with two compatible recombinant fusion plasmids and spreaded on MacConkey agar containing 1% (w/v) maltose, 100 μg/ml ampicillin and 50 μg/ml kanamycin. The plates were incubated at 30 ℃ and the Mal^+ phenotype was observed after 5 nights' incubation. A strong Mal^+ phenotype corresponds to purple coloured colonies and a Mal^- phenotype is indicated by white colonies. A positive control was included involving co-introducing pUT18C-zip and pKT25-zip plasmids into BTH101, which yielded a strong positive Mal^+ phenotype after 5 nights' incubation.

2.7.2 β-galactosidase assay

β-galactosidase activity was measured to quantitate the degree of functional complementation between T18 and T25 hybrid proeins in the BACTH assay. This was carried out in liquid cultures using the method of Miller (Miller 1972; (Karimova et al. 1998). Two plasmids encoding T18 and T25 fusion proteins were co-transformed into BTH101 competent cells, and transformants were selected on LB agar plates supplemented with 100 μg/ml ampicillin and 50 μg/ml kanamycin. The transformation procedure was performed as described in Section 2.3.15.1. The next day, three independent colonies were inoculated into separate tubes containing 4 ml of LB containing ampicillin and kanamycin, and 0.5 mM of IPTG using toothpicks. Cultures were allowed to grow at 30^{o}C shaking incubator (~160 rpm) for 14-16 hours. The cultures were then chilled on ice for 15 minutes and 200 μl was diluted five fold in LB medium, and the OD_{600} was recorded. 50 μl of the diluted culture was used for the assay. 950 μl of Z buffer containing β-mercapthoethanol was added to each 10 cm glass test tube (1mm wall thickness), and this was performed in duplicate for each bacterial culture. 30 μl of chloroform was then shot directly onto the surface of the buffer, making a choloroform bead at the bottom of the tube. After that, 30 μl of 0.1% SDS was added gently to the side of each tube. Then, 50 μl of the diluted bacterial culture was added to the test tube in the same way as the SDS, and the test tubes were vortexed for 10 seconds. In parallel, two control tubes were also set up by adding 50 μl of sterile LB medium instead of the bacterial culture. The test tubes were then incubated in a 30^{o}C water bath for 15 minutes to equilibrate the temperature. 200 μl of ONPG (4 mg/ml dissolved in Z buffer containing β-mercapthoethanol) was added to each test tube at 30 seconds intervals to initiate the reaction. Each test tube was vortexed for 1 second before being returned into the water bath. After that, the development of yellow colour was monitored. Once the appropriate yellow colour was reached (OD_{420} between 0.2-0.8), the reaction was stopped by adding 500 μl of 1 M Na_2CO_3 (again at 30 seconds intervals). The test tubes were vortexed again for 1 second and kept at room temperature. The two control tubes were treated in the same way without yellow colour development. The time of initiating and stopping the assay were recorded accurately. The absorbance of each tube was measured at wavelengths 420 nm and 550 nm for the intensity of the yellow colour developed by hydrolysis of ONPG, and the light-scattering effect of the

cell debris. The controls were used as a blank. The following formula was used for the calculation of β-galactosidase activity:

$$Miller\ Units = \frac{\text{OD420} - 1.75(\text{OD550})}{\text{Time(min)} \times \text{ culture volume added (ml)} \times \text{ OD600}} \times 1000$$

Where,

OD_{420}: Absorbance of the yellow colour of ONPG at 420 nm

OD_{550}: Absorbance of the light scattering effect of the cell debris at 550 nm

Time: the reaction time (starting time ~ stopping time) in minutes

Volume of cells: the volume of the bacterial culture used for this assay in ml

OD_{600}: Absorbance of the cell density at 600 nm

The mean value of each duplicate pair was calculated, and the average value of each set of triplicate cultures was calculated for each protein-protein combination along with the standard deviation.

Z buffer is comprised of 16.1 g $Na_2HPO_4 \cdot 7H_2O$ or 8.53 g Na_2HPO_4, 5.5 g $NaH_2PO_4 \cdot H_2O$, 0.75 g KCl and 0.246 g $MgSO_4 \cdot 7H_2O$ which was dissolved in ddH_2O (< 1 litre), and the final volume was topped up to 1 litre. Z buffer was stored at 4°C. 0.27 ml of β-mercapthoethanol was added into 100 ml of Z buffer on the day of use.

2.7.3 Nickel affinity pull-down assay

The nickel affinity pull-down assay was used to determinate a physical interaction between proteins of interest *in vitro*. A histidine-tagged protein was used as bait, and untagged or proteins labelled with other epitope tags were used as prey. Proteins used for this assay were co-overproduced in the same cell from the same plasmid or two compatible plasmids as described in Section 2.4.1 with minor alterations. If proteins interact with each other, the prey protein should be retained on the nickel matrix along with the bait protein

Charged His-Buster Nickel Affinity Gel (Amocol Bioprocedure) was thoroughly resuspended and 200 µl were transferred to each microcentrifuge tube using 1000 µl, micropipette tips with the tapered end cut off to give an opening of ~3 mm. The nickel resins were washed three times with 1 ml of lysis buffer (50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole) followed by centrifugation at 13,000 x

g for 2 minutes to remove the ethanol. 300 µl of clarified protein lysate was incubated with washed and equilibrated nickel resins in a microcentrifuge tube for 2 hours at room temperature on a rotating wheel. Following this, the nickel resin was washed three times with 1 ml of lysis buffer. Elution of bound material was achieved by adding 100 µl of elution buffer (lysis buffer containing 500 mM imidazole) and incubated at room temperature for further 30 minutes on a rotating wheel. The mixture was then centrifuged at 5,000 x *g* for 2 minutes to isolate the resin, and the supernatant was transferred into a new microcentrifuge tube which was then mixed an equal volume of with 2x Sample Buffer for SDS-PAGE analysis.

2.7.4 Co-immunoprecipitation (co-IP)

The co-IP assay was also used to determine a physical interaction between two proteins of interest *in vitro*. Two methods were employed as described in Sections 2.7.4.1 and 2.7.4.2.

2.7.4.1 EZview^{TM} Red Protein A affinity gel

The affinity resin EZview^{TM} Red Protein A affinity gel (Sigma) contains *Staphylococcus aureus* Protein A covalently attached to crosslinked 4% agarose beads. Antibody that recognised one of the proteins of interest (TssX) was bound to the affinity resin, so that TssX can also be retained on the resin by forming complexes with antibody. If the other protein of interest (TssY) interacts with TssX, it will also be retained on the resin.

Proteins used for this assay were co-overproduced in the same cell from the same plasmid or two compatible plasmids as described in Section 2.4.1 with minor alterations. Bacterial cell lysates were prepared in lysis buffer (50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol), and 0.1-1.0 ml of which were incubated with 1-10 µl of antibody (anti-TssA^{S} polyclonal antibody, BioServe) on a rotating wheel for 1 hour at 4^{o}C to allow antibody-antigen complexes to form. In the meantime, the resin was thoroughly resuspended and 30 µl of the 50% suspension was aliquotted into microcentrifuge tubes using a 1000 µl micropipette tip with ~1 mm cut-off from the tip and kept on ice. The resin was then washed and equilibrated in 750 µl of lysis buffer and recovered by

centrifugation at 8,200 x *g* for 30 seconds. The supernatant was carefully removed and discarded, and the equilibration step was repeated three times. Once the sample of antibody-antigen complexes was prepared, it was incubated with the resins for 1 hour at 4^{o}C on a rotating wheel to allow the binding of antibody-antigen complexes to the resin. The suspension was then centrifuged at 8,200 x *g* for 30 seconds to recover the resin, and the supernatant was carefully removed. The resins were washed in 750 µl of lysis buffer for 5 minutes at 4^{o}C on a rotating wheel for three times. After discarding the supernatant of the final wash, the bound material was eluted from the resin by boiling with 25 µl of lysis buffer and 25 µl of 2x Sample Buffer for 5 minutes. This was followed by centrifugation at 8,200 x *g* for 30 seconds to pellet the resins. 10-20 µl of the supernatant was analysed by SDS-PAGE.

2.7.4.2 Anti-FLAG M2 affinity gel

Anti-FLAG M2 affinity gel (Sigma) is a purified murine IgG1 monoclonal antibody covalently attached to agarose by hydrazide linkage which can be used for immunoprecipitation of FLAG tagged fusion proteins. One of the proteins of interest was tagged with a FLAG epitope tag and the others were labelled with alternative epitope tags. The bacterial cell lysate was prepared in binding buffer containing 50 mM Tris-HCl (pH 7.4), 150/500 mM NaCl with EDTA-free protease inhibitor tablets (Roche) as described in Section 2.4.2 with minor alterations. The clarified cell lysate was transferred to a clean microcentrifuge tube and treated with 2.5 mM $MgCl_2$. After mixing gently by inverting the microcentrifuge tube several times, the mixture was centrifuged at 13,000 x *g* for 10 minutes at room temperature. The supernatant was retained, transferred to a clean microcentrifuge tube and mixed with 0.2-1% Tween® 20 for one hour at room temperature on a rotating wheel. The mixture was then clarified by re-centrifugation at 15,000 x *g* for 10 minutes at room temperature. The supernatant was then used for binding the affinity resin. In the meantime, 50 µl of the anti-FLAG M2 affinity resin suspension was aliquotted using a 1000 µl micropipette tip with ~1 mm cut-off from the tip into a clean microcentrifuge tube. All the steps were performed at 4^{o}C. The suspension was centrifuged at 7,000 x *g* for 30 seconds, and left for further 2 minutes to allow the surface of the resin pellet to become level at the bottom of the microcentrifuge tube. The supernatant was removed with a narrow-end pipette tip. The resin was then equilibrated in 1 ml of binding buffer, mixed gently by inverting several

times, and collected by centrifugation at 7,000 x *g* for 30 seconds. The supernatant was removed and discarded, and the equilibration step was repeated three times.

Cell lysate (200-1,000 µl) was then added to the equilibrated resin, and the final volume was topped up to 1 ml with binding buffer. The volume of cell lysate to be used depended on the expression level of the FLAG-tagged protein. The microcentrifuge tubes were loaded onto a rotating wheel for 2-3 hours at 4^{o}C for binding. After that, the mixture was centrifuged at 7,000 x *g* for 30 seconds, and the supernatant was removed. The resin was then washed in 1 ml of binding buffer and rotated for 5-15 mins before centrifugation at 7,000 x *g* for 30 seconds. The wash step was repeated two more times. After discarding the supernatant of the final wash, the bound material was eluted by boiling with 50 µl of sample buffer with or without 5% β-mercaptoethanol for 3 minutes. The mixture was then centrifuged at 7,000 x *g* for 30 seconds to pellet the resins. The supernatant (bound material) was transferred to a clean microcentrifuge tube and analysed by SDS-PAGE.

2.8 Techniques for protein structure analysis

2.8.1 Negative stain electron microscopy

The negative staining procedure for the protein samples was carried out using a uranyl formate stain. 0.75 % (w/v) uranyl formate stain was prepared by adding 37.5 mg uranyl formate powder to a small beaker and wrapped with foil for minimising the contact with light. 5 ml of ddH_2O was added and the suspension was boiled for 5 minutes with stirring. Following this, ~8 µl of 5 M NaOH was added to the suspension which was stirred for another 5 minutes. The uranyl formate solution was then filter sterilised using a 0.22 µm syringe filter unit, and stored in a foil wrapped universal tube at 4°C.

Carbon-coated grids were glow discharged freshly for 20-30 seconds using a high vacuum turbo carbon coater (Cressington 208carbon). Two drops of ddH_2O (~50 µl) and two drops of uranyl formate (~50 µl) were prepared on parafilm for each grid. Purified protein sample was diluted to 0.009-0.05 mg/ml and 3-5 µl of this was loaded onto the coated side of the discharged carbon grid and incubated for 1 minute. Excess protein sample was removed by blotting the grid on a clean filter paper. The grid was then subsequently washed twice in the ddH_2O drops and once in the uranyl formate

stain drop. The grid was blotted on a clean filter paper between each wash. This was followed by submerging the grid in the second drop of uranyl formate stain for 20 seconds before being blotted on a clean filter paper and dried using a small vacuum pump. The prepared carbon grid was examined by electron microscopy (CM200 FEG, Philips) and stored in a grid box at room temperature for future use.

2.8.2 X-ray Crystallography

Pre-made crystallization screens (Classics Suite, JCSG+, Proplex, PEGs, AMS04, PACT, Morpheus and MPD) were employed for identifying optimum conditions for crystal formation. Each screen was comprised of 96 different conditions, and the Hydra crystallization robot was used for sitting drops of purified protein (200 nl of 10 mg/ml) into 96-well screen plates. The plates were then sealed and incubated at 17^{o}C. The formation of crystals was checked every two days for the first two weeks and once every week afterward. Once crystals were identified in the screen, the conditions of crystallization were optimised by adjustment of buffer components, drop size and protein concentration. When suitably large crystals were obtained, they were mounted onto litho loops and then taken to the Diamond Light Source in Oxford for testing and data collection. Collaborators: S.E. Sedelnikova, H. Owen and S. Dix, Department of Molecular Biology and Biotechnology, University of Sheffield.

Chapter 3 Analysis of interactions between TssA and other T6SS subunits using two-hybrid and three-hybrid assays

3.1 Introduction

The T6SS has been proposed to consist of 13 core components known as TssA-TssM (Shalom et al. 2007). Consistent with this, a comprehensive mutagenesis study demonstrated that 13 out of 16 genes in the *Edwardsiella tarda* T6SS gene cluster encode subunits that are essential for exporting TssD (EvpC), TssI (EvpI) and a putative effector, EvpP (Zheng and Leung 2007). These genes are *evpA-evpC*, *evpE-evpI* and *evpK-evpO*, corresponding to *tssA-tssM*. To fulfil the interaction map of the T6SS core components, each of the 13 components of the *B. cenocepacia* T6SS were analyzed for their interactions with the other subunits using the bacterial adenylate cyclase two-hybrid system (BACTH) system (Shastri 2011; Ahmad 2013).

The BACTH system allows for *in vivo* screening of functional interactions between two proteins. It is based on the reconstitution of cAMP-producing components of the *Bordetella pertussis* adenylate cyclase enzyme, CyaA (Karimova et al. 1998). The enzyme contains two separate domains, T25 and T18 (Ladant 1988). When these two domains are physically separated, cAMP cannot be produced (Ladant and Ullmann 1999). When they are brought in close proximity to each other by fusion to interacting polypeptides, they become able to produce cAMP. In *E. coli*, cAMP then binds to the cAMP receptor protein, CRP (also known as the catabolite activator protein, CAP). The cAMP/CRP complex activates the expression of many catabolic operons, including genes of the *lac* and *mal* operons which are acquired for lactose and maltose catabolism. This facilitates the utilization of lactose or maltose as the unique carbon source by *E. coli* Δ*cya* (endogenous adenylate cyclase-deficient) host strains. This is distinguishable when the bacteria are grown on indicator media such as MacConkey-maltose agar (Figure 3.1) (Karimova et al. 1998). The bacterial *lacZ* gene encodes the enzyme β-galactosidase, which converts lactose into glucose and galactose. The efficiency of complementation between T25 and T18 can be quantified by measuring the β-galactosidase activities in liquid cultures.

To determine the interaction map of the T6SS core components, the individual subunits were analysed using the BACTH system by fusing each subunit to either the T25 or the T18 fragment of CyaA that is encoded by two compatible BACTH vectors (Shastri 2011; Ahmad 2013). T25 is encoded by pKT25 and pKNT25 plasmids (p15A derivatives), where multiple cloning sites (MCS) are available adjacent to the C- or N-terminal coding sequence of the T25 fragment, respectively, to allow for the insertion of the gene of interest in frame with the T25 coding sequence. They also contain a kanamycin resistance gene for selection of the plasmid. T18 fusions can be generated by cloning DNA fragments into pUT18 and pUT18C plasmids (ColE1 derivatives), which specify ampicillin resistance for selection, with the MCS located upstream and downstream of the T18 coding sequence, respectively (Figure 3.2). In the previous study, the *E. coli* Δ*cya* strain BTH101 was co-transformed with compatible pairs of plasmids encoding TssX-T18 or T18-TssX and TssY-T25 or T25-TssY fusion proteins. Transformants were selected on MacConkey-maltose agar with appropriate antibiotics and the maltose phenotype was observed after 120 hours incubation at 30 ℃ (Section 2.7.1). The combination of pKT25-*zip* and pUT18C-*zip* was used as a positive control, which contains the region coding for the leucine zipper region of the yeast protein GCN4 (Karimova et al. 1998). Co-transformation of BTH101 by this plasmid combination results in a strong Mal^+ phenotype. A number of negative controls were also included in the previous study appropriate for each pair of subunits being analyzed. The BACTH plasmid combinations and corresponding controls, along with the expected maltose phenotypes of these combinations, for a typical interaction analysis between TssA and another T6SS subunit (TssX), are shown in Table 3.1.

The results of the previous analysis revealed, among other things, that the TssA subunit interacted with many other T6SS subunits and may therefore play a central role in T6SS assembly and/or function. However, the amino acid sequence alignment of TssA from different Gram-negative bacteria showed that TssA can be classified into two groups according to the sequence divergence in the C-terminal region (Chapter 1). The two groups were named $TssA^S$ (exemplified by the *B. cenocepacia* TssA subunit) and $TssA^{EI}$ (exemplified by the *A. hydrophila* TssA subunit). Therefore, a representative of both groups of TssA was analysed in the earlier studies in order to address the hypothesis that $TssA^{EI}$ is a functional orthologue of $TssA^S$ (Ahmad 2013).

The aim of the work described in this chapter was to identify novel interactions between two types of TssA (i.e. $TssA^{S}$ and $TssA^{EI}$) and other T6SS subunits, and quantify the efficiency of complementation for representative combinations of $TssA^{S}$ two-hybrid system fusions with other T6SS subunit fusions. There were three reasons for doing this. (i) Due to the unstable maltose phenotype observed in colonies expressing certain two-hybrid fusion protein combinations, colonies exhibited a 'patchy' appearance, with segments of each colony showing a maltose-positive phenotype and the remaining areas of the colony consistent with a maltose-negative phenotype. (ii) Due to the higher sensitivity of the β-galactosidase assay, protein-protein interactions may be observed that were not detected using the less sensitive maltose-phenotype. (iii) The efficiency of CyaA complementation resulting from the interaction of TssA with another T6SS subunit can be quantitated. Moreover, the analysis also included some novel fusion protein combinations that were not investigated previously. For example, three-hybrid assays were performed to analyze whether the interaction between $TssA^{S}$ and another T6SS core component (TssX) could be enhanced or stabilized by the presence of a third T6SS subunit (TssY) that was shown to independently interact with TssX by BACTH analysis.

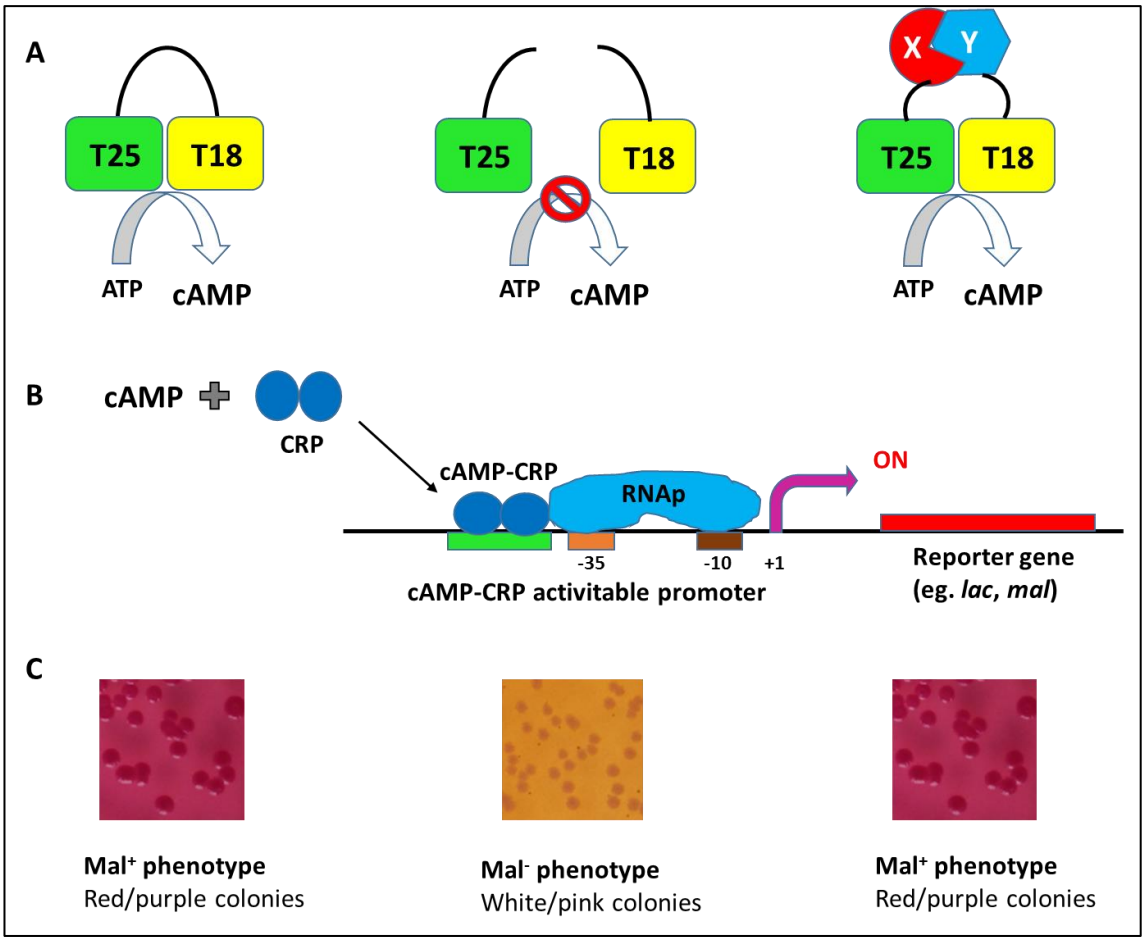


Figure 3.1 Principle of the BACTH system. A. Activation of the BACTH system in *E. coli* Δ*cyaA* host strains. T25 and T18 are two complementary fragments of the adenylate cyclase enzyme, CyaA, from *Bordetella pertussis*. They are not active when they are physically separated. When they are brought in close proximity to each other by fusion to interacting polypeptides X and Y, they become able to produce cAMP. **B**. Cyclic AMP then binds to the *E. coli* cAMP receptor protein (CRP). The cAMP/CRP complex turns on the expression of several resident genes, including genes of the *lac* and *mal* operons which are involved in lactose and maltose catabolism. This facilitates the utilization of lactose or maltose by the *E. coli* Δ*cya* host strain as the unique carbon source. This is distinguishable when the bacteria are grown on indicator media such as MacConkey-maltose agar. **C**. Maltose phenotypes that were observed from the BACTH assay corresponding to the activation of the BACTH system as shown in A.

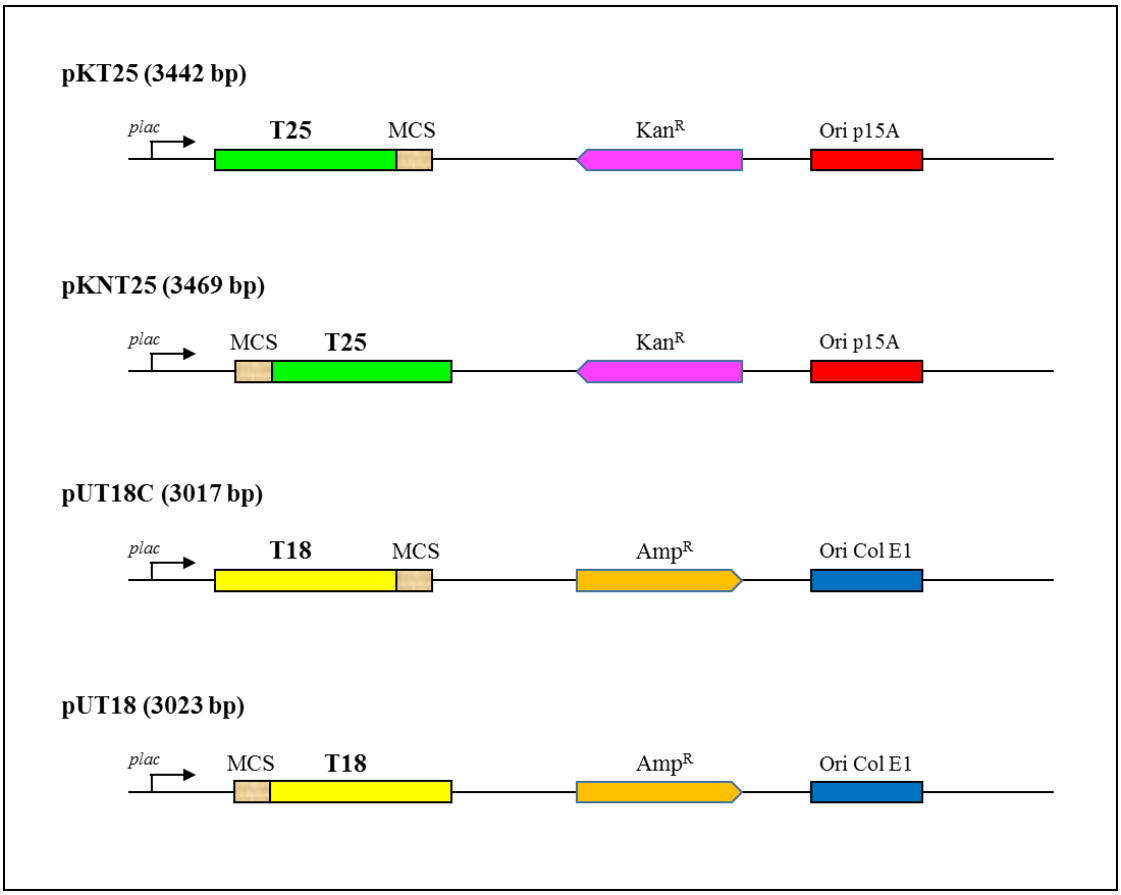


Figure 3.2 Schematic representation of the BACTH system plasmids. Plasmids are represented in linear form. pKT25 and pKNT25 plasmids encode the CyaA T25 fragment (in green); pUT18 and pUT18C plasmids encode the CyaA T18 fragment (in yellow). The location of the multiple cloning sites (MCS) is shown in each plasmid with a beige rectangle. Adapted from the BACTH system kit manual (Euromedex, 2000).

Table 3.1 Plasmid combinations used in the BACTH system for analysing protein-protein interactions between TssA and another T6SS core component

Plasmid 1	Plasmid 2	Expected Mal phenotype
A		
pKT25-*tssA*	pUT18-*tssX*	Unknown
pKT25-*tssA*	pUT18C-*tssX*	Unknown
pKNT25-*tssA*	pUT18-*tssX*	Unknown
pKNT25-*tssA*	pUT18C-*tssX*	Unknown
pUT18-*tssA*	pKT25-*tssX*	Unknown
pUT18C-*tssA*	pKT25-*tssX*	Unknown
pUT18-*tssA*	pKNT25-*tssX*	Unknown
pUT18C-*tssA*	pKNT25-*tssX*	Unknown
B		
pKT25-*tssA* or *tssX*	pUT18	Negative
pKT25-*tssA* or *tssX*	pUT18C	Negative
pKNT25-*tssA* or *tssX*	pUT18	Negative
pKNT25-*tssA* or *tssX*	pUT18C	Negative
pUT18-*tssA* or *tssX*	pKT25	Negative
pUT18C-*tssA* or *tssX*	pKT25	Negative
pUT18-*tssA* or *tssX*	pKNT25	Negative
pUT18C-*tssA* or *tssX*	pKNT25	Negative
pKT25-*zip*	pUT18-*zip*	Strong positive
C		
pKT25	pUT18	Negative
pKT25	pUT18C	Negative
pKNT25	pUT18	Negative
pKNT25	pUT18C	Negative

Note: *tssA* represents the entire *tssA* gene (either type); *tssX* represents a *tss* gene (T6SS core component) other than *tssA.*

A. Plasmid combinations used to probe protein-protein interactions.

B. Control plasmid combinations

C. Empty plasmid control combinations

3.2 Construction of plasmids encoding T6SS subunit-CyaA fragment fusions for two-hybrid analysis

3.2.1 Construction of pKNT25-*tssC*, pUT18-*tssC* and pUT18C-*tssC*

In order to investigate interactions between two types of TssA and TssC, three previously constructed BACTH plasmids encoding TssC fusions that were not made correctly were reconstructed. pKNT25-*tssC* and pUT18-*tssC* have an in-frame stop codon at the end of the *tssC* sequence, which stops subsequent translation of T25 and T18 mRNA located downstream of *tssC* and there is also a base substitution mutation that caused an amino acid change near the N-terminus of TssC. In addition, *tssC* was not cloned in the correct reading frame in the pUT18C plasmid due to an incorrectly designed forward primer. Therefore, these three plasmids were re-constructed in this study.

PCR amplification of *tssC* was carried out from chromosomal DNA of *B. cenocepacia* H111 with two pairs of primers Nterm-iotCfor and Nterm-tssCrev.NEW, and Cterminal-tssCfor.new and Cterm-iotCrev. The first pair of primers contains the restriction sites, *Hin*dIII and *Bam*HI in the forward and reverse primers, respectively, for the insertion of *tssC* (~1.5 kb) into pKNT25 and pUT18, whereas *Pst*I and *Xba*I were used in the second pair of primers for the insertion of *tssC* into pUT18C. In pUT18C, the MCS is located downstream of the T18 coding sequence, whereas pKNT25 and pUT18 plasmids have the MCS located upstream of the T25 and T18 coding sequences. Primers were designed to ensure the *tssC* was in the same reading frame as the T18/T25 coding sequences, and in the reverse primer used to amplify *tssC* for insertion into pKNT25 and pUT18, the stop codon of *tssC* was excluded. The plasmids were named as pKNT25-*tssC*, pUT18-*tssC* and pUT18C-*tssC* following DNA sequencing.

The pKNT25-*tssC* plasmid was then further modified by digestion with *Acc*65I followed by a DNA end filling-in reaction using DNA polymerase I Klenow fragment (Section 2.3.11) and religation of the plasmid in order to place the T25 coding sequence in the same reading frame as that of *tssC*. The integrity of the *tssC* sequence contained in the modified pKNT25-*tssC* plasmid was confirmed by DNA sequencing.

3.2.2 Construction of pKNT25-*tssG* and pUT18-*tssG*

In order to investigate interactions between two types of TssA and TssG, pKNT25-*tssG* and pUT18-*tssG* plasmids were reconstructed in which the *tssG* coding sequence contained an incorrect codon at the C-terminal coding end of the gene. PCR amplification of *tssG* was carried out from chromosomal DNA of *B. cenocepacia* H111 using primers N-iotGfrag1for and Nterminal-iotGrev2. In the reverse primer, the stop codon of the *tssG* gene was excluded to permit in-frame translation of T25 or T18 mRNA downstream of *tssG*. The amplified *tssG* DNA fragment of the expected size ~1.1 kb was digested with restriction enzymes *Hin*dIII and *Bam*HI that recognized sites in the forward and reverse primers and cloned into the corresponding restriction sites of pKNT25 and pUT18 plasmids in the plasmid MCS. The plasmids were named as pKNT25-*tssG* and pUT18-*tssG* following DNA sequencing.

3.2.3 Modification of pKNT25-$tssM_{NTD}$ and pUT18-$tssM_{NTD}$

In the previously constructed pKNT-$tssM_{NTD}$ and pUT18-$tssM_{NTD}$ plasmids which encodes the cytoplasmic domain of TssM, the $tssM_{NTD}$ sequence was not in the same reading frame as the downstream T18 and T25 coding sequences due to an incorrectly designed $tssM_{NTD}$ reverse primer. In order to investigate interactions between two types of TssA and $TssM_{NTD}$, it was decided to rectify this. These two defective plasmids were digested with *Acc*65I and a DNA filling-in reaction was carried out with DNA polymerase I Klenow fragment as descibed in Section 2.3.11, followed by religation of the plasmids. The desired recombinant plasmids were identified by DNA sequencing of the cloned DNA.

3.3 Analysis of interactions between $TssA^S$ and other T6SS subunits using two-hybrid assays

By using the bacterial two-hybrid system, $TssA^S$ of *B. cenocepacia* has been shown to self-interact and also to interact with many other T6SS core proteins; i.e. TssB, -D, -E, -F, -H, -I, -K and -L (Shastri 2011; Ahmad 2013). As some of the previous two-hybrid fusion plasmids were not correctly made, possible interactions between $TssA^S$ and other

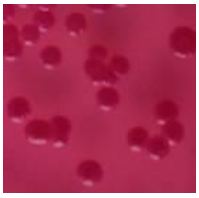
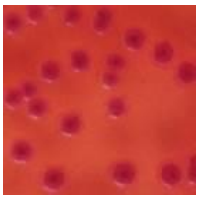
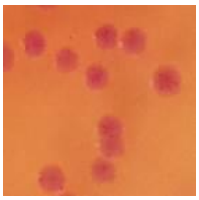
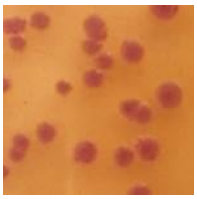
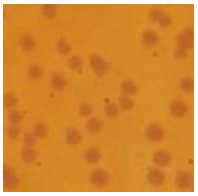
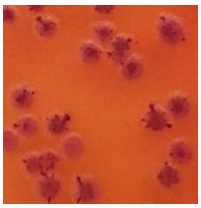
subunits (TssC, TssG and $TssM_{NTD}$) may not have been detected. Therefore, BACTH assays were repeated with the corrected plasmid constructs. In addition, the efficiency of the functional complementation between T18 and T25 for combinations of $TssA^S$ with all T6SS subunits (including TssA itself) was quantified by measuring the β-galactosidase activities of host cells growing in liquid culture (Karimova et al. 1998). To do this, particular compatible plasmid combinations encoding T25-$TssA^S$, $TssA^S$-T25, T18-$TssA^S$ and $TssA^S$-T18 fusion proteins with T25-TssX, TssX-T25, T18-TssX and TssX-T18 were introduced into the assay strain.

For each interaction between $TssA^S$ and a T6SS core subunit that was investigated by the BACTH assay, the degree of complementation of CyaA activity was also quantitated by performing a β-galactosidase assay on at least one of the fusion protein combinations. Based on the previous BACTH analysis of the interaction between $TssA^S$ and other Tss proteins, the priority choice of combinations for β-galactosidase assay were T25-TssX with T18-$TssA^S$, and T25-$TssA^S$ with T18-TssX, except for the interaction of $TssA^S$ with TssB and TssE, as in most of the cases the interactions between $TssA^S$ and other T6SS subunits were observed when the C-terminus of both of interacting T6SS subunits was free (Shastri 2011; Ahmad 2013). In situations where the maltose phenotype was negative and β-galactosidase activity was not detected with these combinations, TssX-T25 with T18-$TssA^S$, and $TssA^S$-T25 with T18-TssX combinations were assayed. In some cases, all possible compatible combinations were assayed. Table 3.2 describes the system used to classify the maltose phenotypes that were observed in the BACTH assay used in this study.

As evidence has been previously provided that $TssA^S$ potentially interacts with 8 other T6SS core proteins, $TssA^S$ is likely to be central for the T6SS assembly. Based on this, a subsequent effort was made to investigate the role of the $TssA^S$ N-terminal region and C-terminal domain in these interactions, including $TssA^S$ domain self-interaction. The domain organization of $TssA^S$ was predicted by multiple sequence alignment among $TssA^S$ orthologues present in other Gram-negative bacteria and the secondary structure prediction program Psipred (Section 1.5.3.2.4). Subsequently, the domain organization of $TssA^S$ was supported by proteolysis (Ahmad 2013 and Chapter 5). These analyses showed that $TssA^S$ contains a large N-terminal region that is partially homologous to TssA in other classes. It also contains a small C-terminal domain that is connected to

the N-terminal region by a long linker. DNA encoding predicted N-terminal region (residues 1-256) and C-terminal domain (residues 293-373) have been previously cloned into BACTH plasmids (Shastri 2011; Ahmad 2013). Where $TssA^S$ was shown to interact with another T6SS subunit by BACTH analysis, the role of the N-terminal region and the C-terminal domain of $TssA^S$ was also similarly investigated and the degree of complementation of CyaA activity was quantified. In addition, in this analysis, some three-hybrid assays were also included: specifically the interaction between $TssA^S$ and TssK in the presence of TssL, the interaction between $TssA^S$ and TssL in the presence of TssK, and the interaction between $TssA^S$ and TssF in the presence of TssG were tested.

Table 3.2 Description of the classification of the maltose phenotypes observed in the BACTH assay used in this study

Description of colony phenotype	Attributed rating [a]	
Very strong Mal$^+$ phenotype: dark red/purple colonies with purple halo around the colonies	+++++	
Strong Mal$^+$ phenotype: red/purple colonies	++++	
Moderate Mal$^+$ phenotype: moderate red/purple colonies	+++	
Weak Mal$^+$ phenotype: pink/red colonies	++	
Very weak Mal$^+$ phenotype: very weak pink/red colonies	+	
Negative maltose phenotype: white/yellow colonies	-	
Patchy phenotype: Colonies showed pink/red (Mal$^+$) colour with some red/purple sectors within each colony ("patches"). Patchy phenotype was combined with the 'cross' classification. Thus, the overall colour of the colonies was classified as crosses indicated above, and if patches were observed, an extra 'p' was added.	P	

[a] *E. coli* strain BTH101 was transformed with two compatible plasmids of the BACTH system and grown at 30^{o}C for 120 h on MacConkey-maltose agar with selective antibiotics. The maltose phenotype of the colonies was recorded and represented as crosses.

3.3.1 Analysis of TssA^S self-interaction

TssA^S has been shown to self-interact based on the Mal-positive phenotype of the indicator strain harbouring all four possible pairwise combinations of BACTH plasmids encoding TssA^S-CyaA domain fusion proteins. The combination of T25-TssA^S and T18-TssA^S fusion proteins was repeated in this study. After colonies were grown on maltose-MacConkey agar plates for 3 nights, the results showed that this combination gave a strong Mal$^+$ phenotype. As the C-terminus of TssA^S is free in both CyaA domain fusions, this result is consistent with previous observations demonstrating that TssA^S self-interacts by means of the C-terminal domain (Shastri 2011; Ahmad 2013).

To further confirm that self-interaction required TssA$^S_{CTD}$, the BACTH assay was also carried out between TssA^S and its domains, i.e. TssA$^S_{NTR}$ and TssA$^S_{CTD}$. All eight compatible combinations of TssA^S and TssA$^S_{NTR}$ fusion proteins were analysed using the BACTH assay. The results showed that three combinations out of eight gave rise to Mal$^+$ phenotypes. In each case, the C-terminus of TssA^S was free. The combination of TssA^S-T25 and T18-TssA$^S_{NTR}$ led to a negative maltose phenotype which is not consistent with previous observations (Ahmad 2013).

All eight compatible combinations of TssA^S- and TssA$^S_{CTD}$-CyaA fusions were also analysed, and it was observed that 6 out of 8 combinations gave rise to uniform deep red/purple colonies indicative of very strong Mal$^+$ phenotypes. However, the results were not consistent with the previous findings, where all 8 combinations showed a very strong Mal$^+$ phenotype (Ahmad 2013). The combination of TssA^S-T25 and TssA$^S_{CTD}$-T18 gave non-uniform coloured colonies; some colonies were light pink and some were red/purple. For the TssA^S-T18 and TssA$^S_{CTD}$-T25 combination, the majority of colonies were a weak pink colour containing a few dark red/purple areas. The reason for this variation in colony morphology is not clear but it suggested a very weak Mal$^+$ phenotype.

TssA^S domain-domain interactions were then investigated. All four pairwise combinations of TssA$^S_{CTD}$ fusion proteins were tested for TssA$^S_{CTD}$ self-interaction and all of them gave rise to uniform dark red/purple colonies corresponding to a very strong Mal$^+$ phenotype. This is in agreement with the previous results in which TssA$^S_{CTD}$ was indicated to self-interact (Ahmad 2013). The TssA$^S_{NTR}$-T25 and TssA$^S_{NTR}$-T18 combination was also tested to screen for TssA$^S_{NTR}$ self-interaction. However, a

negative maltose phenotype (white colonies) was observed for this combination. Therefore, $TssA^{S}_{NTR}$ can interact with $TssA^{S}$ but possibly not with $TssA^{S}_{NTR}$, although only a single combination of two-hybrid proteins was tested. These results support the proposal that $TssA^{S}$ self-interacts, and that its CTD is responsible for the oligomerisation (Shastri 2011; Ahmad 2013).

β-galactosidase assays were carried out for all the combinations that were analysed on MacConkey-maltose agar (Figure 3.3). A number of control combinations were also included as indicated in Table 3.1. The β-galactosidase activity of the Zip-Zip positive control was 6317 ± 190.3 Miller units (Mu); and all the negative controls gave rise to activities between 75-85 Mu. Consistent with the strong Mal^{+} phenotype generated by the $TssA^{S}$ self-interaction combination, T25-$TssA^{S}$ and T18-$TssA^{S}$, cells containing this two-hybrid combination exhibited a β-galactosidase activity of 5575.1 ± 220.1 Mu.

The three combinations of $TssA^{S}_{NTR}$ and $TssA^{S}$ that gave rise to a Mal^{+} phenotype, $TssA^{S}_{NTR}$-T25 and T18-$TssA^{S}$, T25-$TssA^{S}$ and T18-$TssA^{S}_{NTR}$, T25-$TssA^{S}_{NTR}$ and T18-$TssA^{S}$, gave the β-galactosidase activities of 3546.7 Mu, 1292.8 Mu and 321.8 Mu, respectively, whereas the other combinations that gave rise to a maltose-negative phenotype led to β-galactosidase activities that were similar to the negative controls. The results of the β-galactosidase assays on cells harbouring fusion protein combinations involving $TssA^{S}$ and $TssA^{S}_{CTD}$ were in agreement with the maltose phenotypes, except for two combinations, $TssA^{S}$-T25 with $TssA^{S}_{CTD}$-T18, and $TssA^{S}_{CTD}$-T25 with $TssA^{S}$-T18. These two combinations gave a patchy Mal^{+} phenotype, but gave β-galactosidase activities that were similar to the background level.

Regarding domain-domain interactions, the single combination of $TssA^{S}_{NTR}$ fusion proteins that was tested, $TssA^{S}_{NTR}$-T25 and $TssA^{S}_{NTR}$-T18, gave rise to a β-galactosidase activity that was slightly above the background, i.e. 118.7 ± 27.0 Mu (p = .012), although a maltose negative phenotype was observed. All four pairwise combinations of $TssA^{S}_{CTD}$ fusion proteins led to a very high β-galactosidase activity, consistent with the very strong maltose positive phenotype.

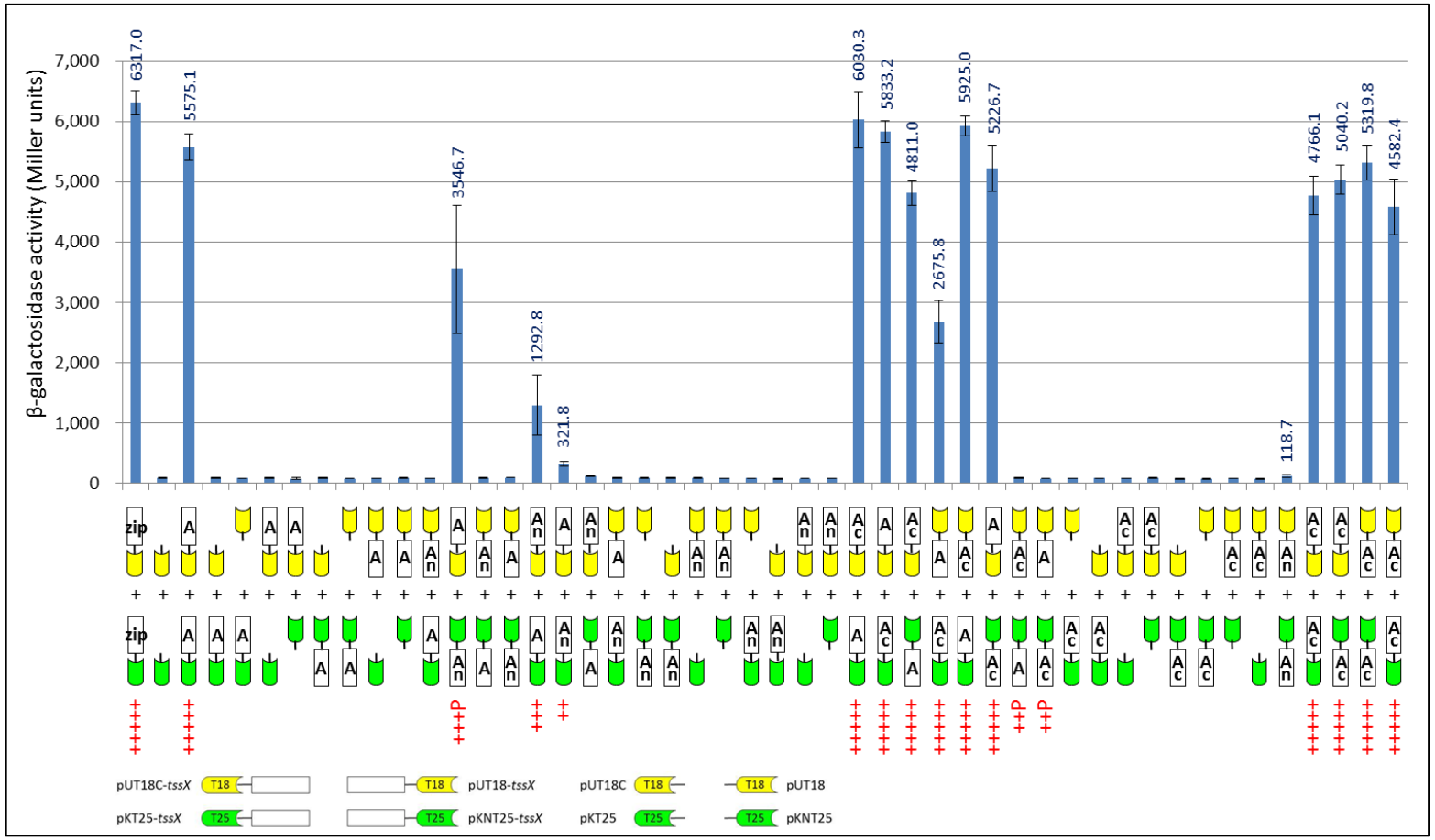


Figure 3.3 BACTH analysis of TssA^S self-interaction. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused to TssA^S and its domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2; 'P', patchy. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A' represents TssA^S, and 'An' and 'Ac' represent TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively.

3.3.2 Analysis of the interaction between $TssA^S$ and TssB

Two out of eight compatible plasmid combinations encoding $TssA^S$ and TssB CyaA fusions were assayed in the BACTH system. The combination of TssB-T25 and T18-$TssA^S$ gave rise to moderate red/purple colonies, recorded as a Mal^+ phenotype, whereas the $TssA^S$-T25 and T18-TssB combination gave rise to a negative maltose phenotype. The results are in agreement with the previous findings (Shastri 2011).

As a Mal^+ phenotype was observed in the combination in which the N-terminus of TssB and the C-terminus of $TssA^S$ was free, the corresponding TssB fusion was subsequently tested with $TssA^S_{CTD}$ and $TssA^S_{NTR}$ for identifying the domain of $TssA^S$ that interacts with TssB. Thus, $TssA^S_{NTR}$-T25, $TssA^S_{NTR}$-T18, T25-$TssA^S_{CTD}$ and T18-$TssA^S_{CTD}$ fusion proteins were tested with compatible TssB-T25 or TssB-T18 fusions. All of these four combinations yielded a negative maltose phenotype. However, one of these results is not consistent with the previously reported observation where the TssB-T25 and T18-$TssA^S_{CTD}$ combination was recorded as giving a very strong Mal^+ phenotype (Shastri 2011).

β-galactosidase assays were performed for all the combinations that were analysed on MacConkey-maltose agar. The results of the β-galactosidase assays were in agreement with the maltose phenotypes that were observed for the corresponding two-hybrid fusion protein combinations. Thus, the β-galactosidase activity of the TssB-T25 and T18-$TssA^S$ combination was 294.5 ± 24.6 Mu, that is statistically significantly higher than that of the negative controls ($p < .001$) and three-fold higher than the background level (80 Mu). For the remaining combinations, the β-galactosidase activities were similar to the background level that was measured in the negative controls (Figure 3.4). These results suggest that TssB interacts with $TssA^S$ and that both domains of $TssA^S$ are required for this interaction.

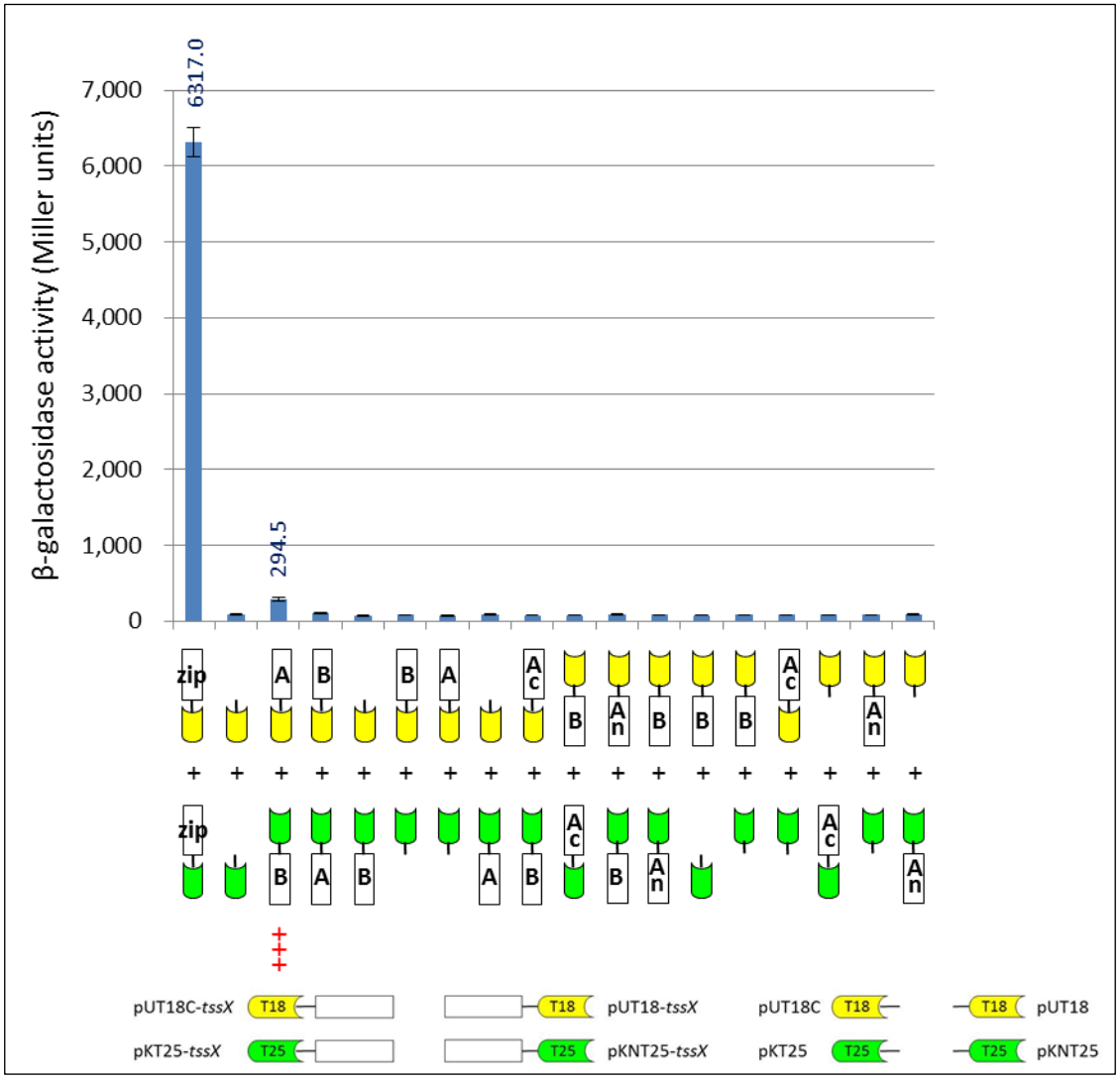


Figure 3.4 BACTH analysis of the interaction between TssA^S and TssB. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused to TssB, TssA^S and TssA^S domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'B', 'An' and 'Ac' represent TssA^S, TssB, TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively.

3.3.3 Analysis of the interaction between $TssA^S$ and TssC

It was previously observed that BACTH plasmids expressing TssC two-hybrid fusions in combination with $TssA^S$ fusions gave rise to a maltose negative phenotype for all of the eight pairwise combinations (Shastri 2011). However, it was subsequently shown that three of the four plasmids encoding TssC-CyaA domain fusions were defective. The reconstruction of pKNT25-*tssC*, pUT18-*tssC* and pUT18C-*tssC* plasmids was performed as described in Section 3.2. All eight pairwise BACTH combinations of $TssA^S$ and TssC were assayed with the rectified plasmids. One of the combinations in which the N-terminus of TssC and C-terminus of $TssA^S$ is free (i.e. TssC-T25 with T18-$TssA^S$), yielded a strong Mal^+ phenotype as judged by the formation of red/purple colonies. This is a new finding that suggests that there is an interaction between $TssA^S$ and TssC. The remaining combinations led to a maltose-negative phenotype. Subsequent investigation of the interaction between $TssA^S$ domains and fusions of TssC to the N-terminus of T25 or T18 was conducted by combining TssC-T25 or TssC-T18 with $TssA^S_{NTR}$ in three pairwise combinations of and with $TssA^S_{CTD}$ in three combinations. However, all of these six combinations gave rise to a maltose-negative phenotype.

β-galactosidase assays were performed for all the $TssA^S$ and TssC combinations that were screened on MacConkey-maltose agar. The combination of TssC-T25 and T18-$TssA^S$ gave a β-galactosidase activity of 1958.7 ± 115.9 Mu, in accordance with the maltose phenotype of colonies containing the corresponding pair of fusion proteins. All the remaining combinations gave β-galactosidase activities that were similar to the background level in agreement with the maltose phenotype observed (Figure 3.5). These results indicate that TssC interacts with $TssA^S$ and suggests that both domains of $TssA^S$ are required for this interaction.

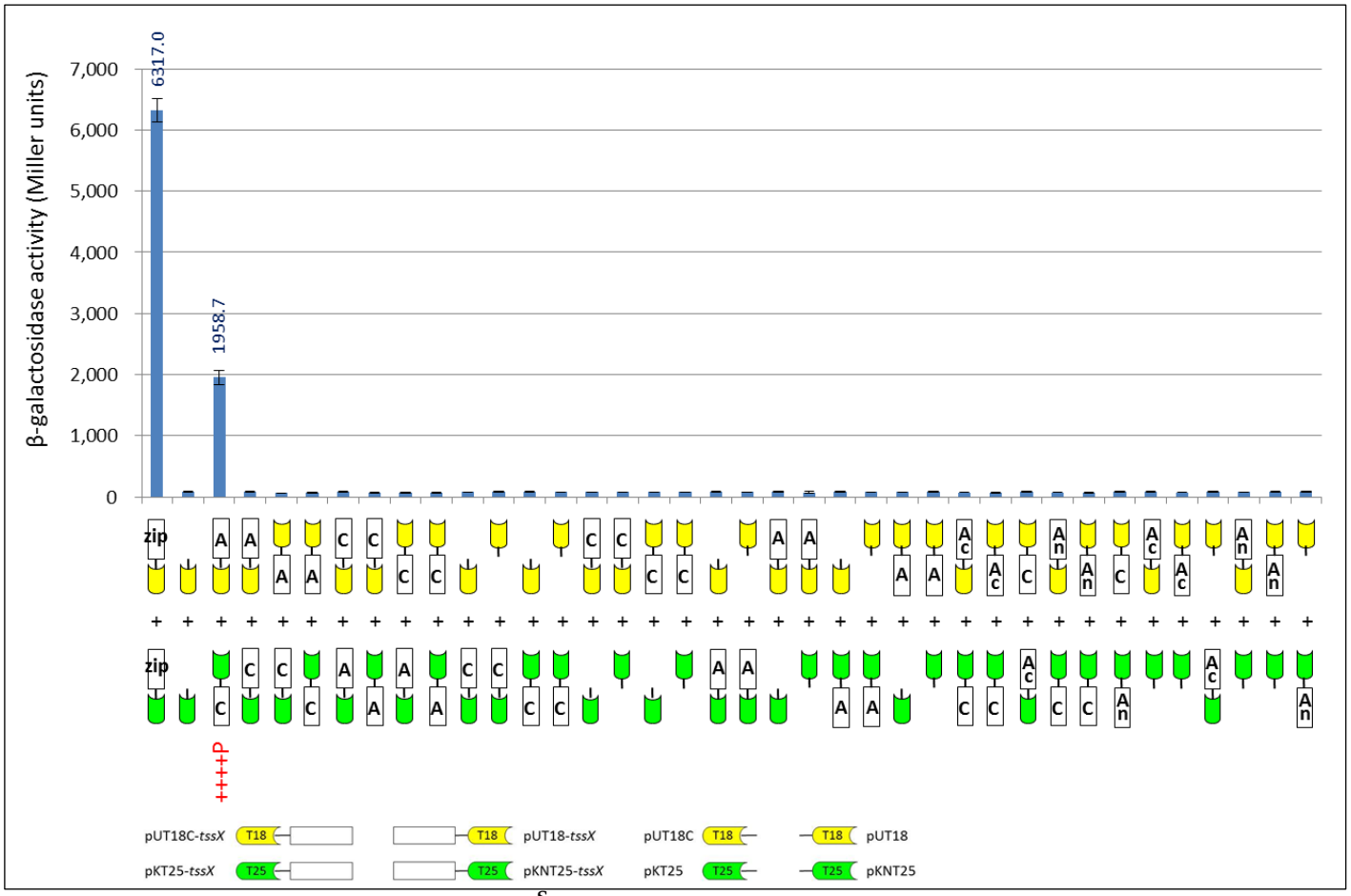


Figure 3.5 BACTH analysis of the interaction between TssA^S and TssC. Compatible BACTH plasmids encoding the CyaA T25 or T18 domain fused to TssC, TssA^S and TssA^S domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2; 'P', patchy. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'C', 'An' and 'Ac' represent TssA^S, TssC, TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively.

3.3.4 Analysis of the interaction between $TssA^S$ and TssD

Two out of eight compatible BACTH plasmid combinations encoding CyaA domain fusions to $TssA^S$ and TssD were investigated in the BACTH assay, i.e. T25-TssD with T18-$TssA^S$, and T25-$TssA^S$ with T18-TssD. Both combinations gave rise to a strong Mal^+ phenotype with red/purple colonies when the C-terminus of both $TssA^S$ and TssD were free. The result is consistent with previous findings within our group (Shastri 2011; Ahmad 2013).

As the C-terminus of TssD was free in the combinations with $TssA^S$ that yielded a Mal^+ phenotype, T25-TssD and T18-TssD were subsequently assayed with all compatible $TssA^S_{CTD}$ fusions and the two of the $TssA^S_{NTR}$ fusions in which the N-terminus of $TssA^S_{NTR}$ is free. The T25-$TssA^S_{CTD}$ and T18-TssD combination gave rise to pink coloured colonies with patchy dark red/purple dots within some individual colonies, whereas T25-TssD and T18-$TssA^S_{CTD}$ yielded a maltose-negative phenotype. When the N-terminus of $TssA^S_{CTD}$ was free in the combinations with pairwise T25-TssD or T18-TssD, one combination revealed a Mal^+ phenotype ($TssA^S_{CTD}$-T25 and T18-TssD); the other combination gave a negative maltose phenotype (T25-TssD and $TssA^S_{CTD}$-T18), which is in accordance with the previously reported observations (Shastri 2011), although the phenotype of the $TssA^S_{CTD}$-T25 and T18-TssD combination was not as strong as previously recorded (very strong, +++++p). In addition, the result of the combination of T25-TssD and $TssA^S_{CTD}$-T18 is not consistent with the previous observation (Shastri 2011). Both pairwise combinations of two-hybrid fusions in which the N-terminus of $TssA^S_{NTR}$ and the C-terminus of TssD are free led to a negative maltose phenotype.

β-galactosidase assays were performed for all the $TssA^S$ and TssD two-hybrid combinations that were assayed on MacConkey-maltose agar. The combinations of T25-TssD and T18-$TssA^S$, and T25-$TssA^S$ and T18-TssD gave β-galactosidase activities of 2880.3 ± 173.33 Mu and 3305.7 ± 193.72 Mu, respectively, which is in accordance with the maltose phenotype of colonies containing the corresponding pair of fusion proteins. The β-galactosidase activity of T25-$TssA^S_{CTD}$ with T18-TssD and $TssA^S_{CTD}$-T25 with T18-TssD were 145.1 ± 11.3 Mu ($p < .001$) and 473.3 ± 222.7 Mu, respectively, with the latter exhibiting a relatively high standard deviation. For the remaining combinations, the β-galactosidase activities were similar to the background

level that was measured in the negative controls (Figure 3.6). Overall, the results of β-galactosidase assays are in agreement with the maltose phenotypes observed. In conclusion, these results suggest that TssD interacts with $TssA^S$, and the C-terminal domain of $TssA^S$ is involved in the interaction with TssD.

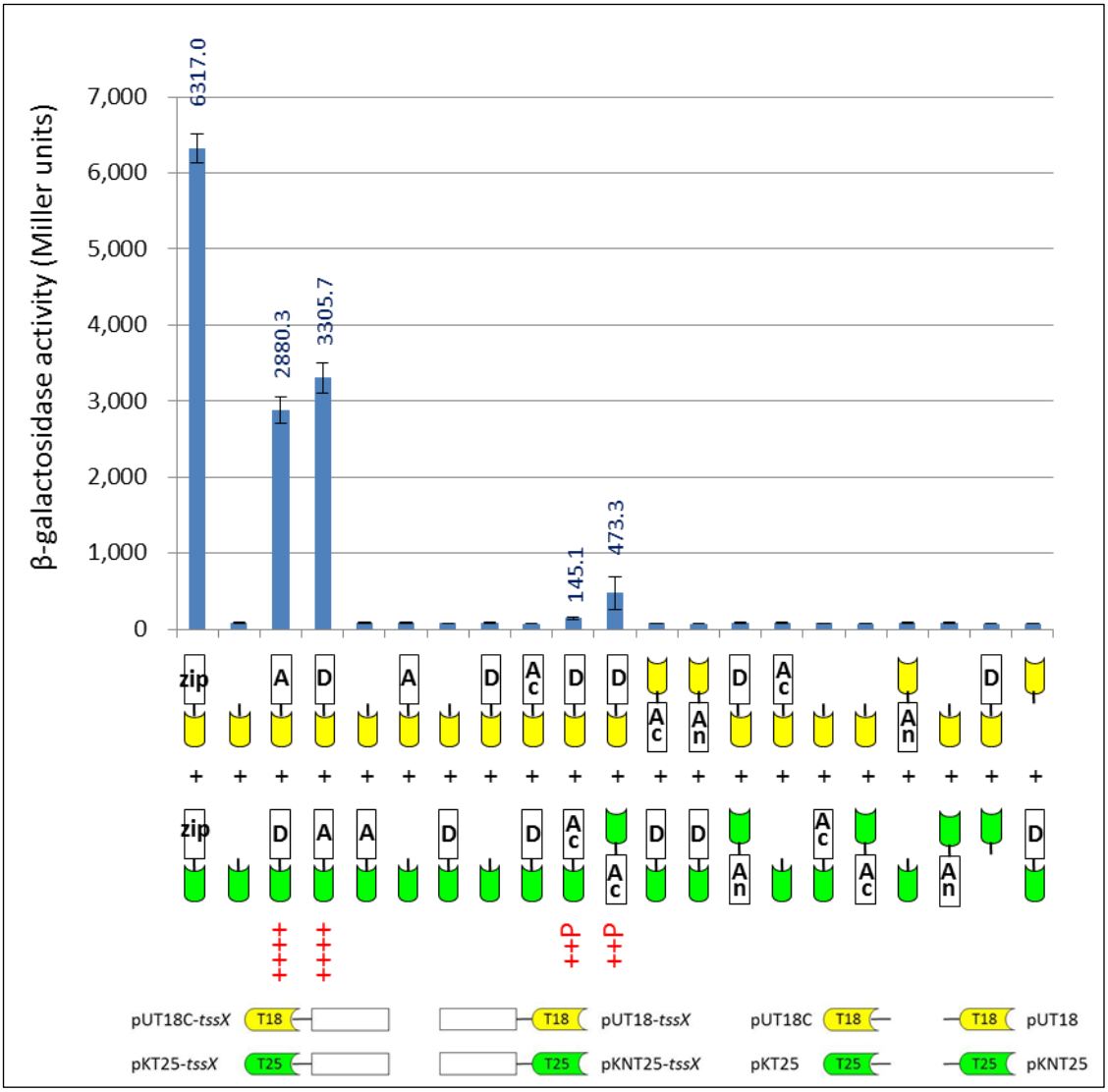


Figure 3.6 BACTH analysis of the interaction between TssA^S and TssD. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused to TssD, TssA^S and TssA^S domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2; 'P', patchy. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'D', 'An' and 'Ac' represent TssA^S, TssD, TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively.

3.3.5 Analysis of the interaction between TssA^S and TssE

Two out of eight possible compatible BACTH plasmid combinations encoding TssA^S and TssE fusion proteins were analysed by the BACTH assay, which were TssE-T25 with T18-TssA^S, and TssA^S-T25 with T18-TssE. The combination in which the N-terminus of TssE and C-terminus of TssA^S were free led to a patchy Mal$^+$ phenotype, whereas the other combination conferred a maltose-negative phenotype. The results are in accordance with the previous observations (Shastri 2011; Ahmad 2013). TssA^S domain fusions TssA$^S_{NTR}$-T25 or TssA$^S_{NTR}$-T18, in which TssA$^S_{NTR}$ is located at the N-terminus, and T25-TssA$^S_{CTD}$ or T18-TssA$^S_{CTD}$, in which TssA$^S_{CTD}$ is located at the C-terminus, were also assayed in combination with compatible TssE-T25 or TssE-T18 fusions. All of these four pairwise combinations yielded a negative maltose phenotype in agreement with previous observations (Shastri 2011; Ahmad 2013).

β-galactosidase activities were measured in cells harbouring all the TssA^S and TssE combinations that were assayed on MacConkey-maltose agar. The β-galactosidase activity of the combination of TssE-T25 and T18-TssA^S was 333.2 ± 43.7 Mu which is statistically significantly higher than the β-galactosidase activity measured in the negative controls ($p < .001$), whereas the β-galactosidase activities of all the remaining combinations were similar to the background level that was measured in the negative controls (Figure 3.7). The results of the β-galactosidase assays are in agreement with the maltose phenotypes and suggest that TssA^S interacts with TssE but may require the TssE N-terminus to be available to do so.

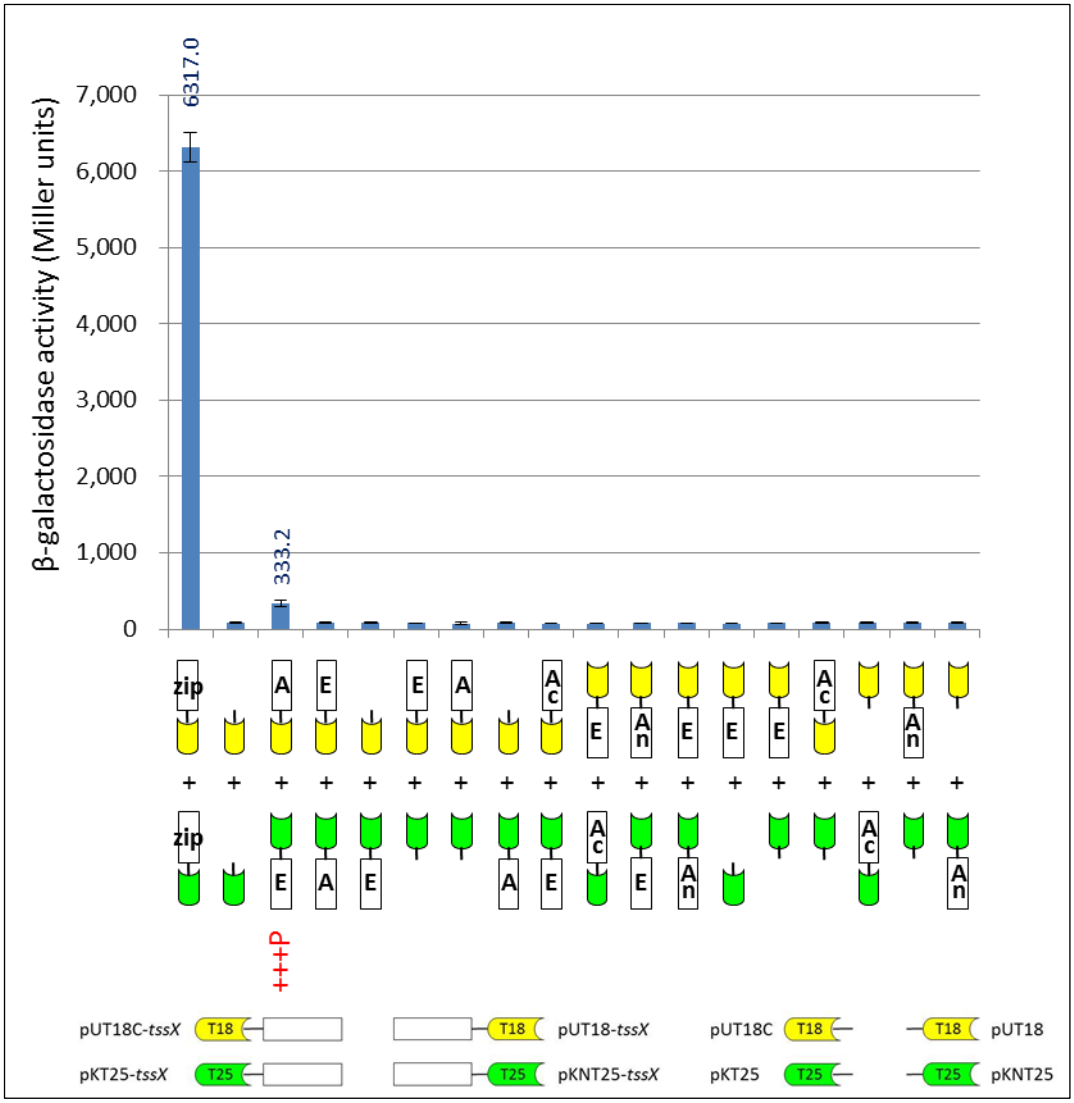


Figure 3.7 BACTH analysis of the interaction between TssA^{S} and TssE. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused to TssE, TssA^{S} and its domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2; 'P', patchy. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'E', 'An' and 'Ac' represent TssA^{S}, TssE, TssA$^{S}_{NTR}$ and TssA$^{S}_{CTD}$, respectively.

3.3.6 Analysis of the interaction between TssA^S and TssF

All eight possible pairwise combinations of TssA^S- and TssF-CyaA domain fusions were analysed in the BACTH assay. Both combinations of TssA^S and TssF fusion proteins where the C-terminus of both T6SS subunits is free, gave rise to red/purple patchy colonies, suggesting a Mal$^+$ phenotype. When TssF was fused at its C-terminus to T25 and combined with the T18-TssA^S fusion in which the C-terminus of TssA^S was free, a strong Mal$^+$ phenotype, with uniform dark red/purple colonies, was observed. All of the remaining five combinations led to a negative Mal phenotype. Thus, in all three fusion protein combinations that gave rise to a Mal$^+$ phenotype, TssA^S was located at the C-terminus of the hybrid protein.

Interestingly, one of the control combinations, pKNT25 and pUT18C-*tssF*, gave rise to pink/red colonies, whereas the corresponding combination with two fusion proteins (pKNT25-*tssA*^{S} and pUT18C-*tssF*) yielded a negative maltose phenotype. It was not previously recorded that this control showed a weak Mal$^+$ phenotype (Shastri 2011). The integrity of pKNT25 plasmid was confirmed by DNA sequencing, and this particular control combination was repeated a few times on MacConkey-maltose agar, but the result was still the same. Moreover, these results also disagree with previous observations which recorded TssA^S-T25 and T18-TssF as giving rise to a Mal$^+$ phenotype with patchy colonies (Shastri 2011; Ahmad 2013).

As TssA^S and TssF appear to interact based on the Mal phenotypes, five pairwise combinations between TssA$^S_{CTD}$ and TssF, and five combinations between TssA$^S_{NTR}$ and TssF were subsequently assayed for locating the possible region(s) of TssA^S that interacts with TssF. Two combinations of TssA$^S_{CTD}$ and TssF two-hybrid proteins yielded Mal$^+$ phenotypes, in which C-terminus of TssF was free in both cases with T25 fused at either N- or C-terminus of TssA$^S_{CTD}$. Despite the fact that a strong Mal$^+$ phenotype with uniform dark red/purple colonies was observed for the TssF-T25 and T18-TssA^S combination, the corresponding combination with TssA$^S_{CTD}$ in its normal context, i.e. TssF-T25 and T18-TssA$^S_{CTD}$, gave rise to a negative maltose phenotype. The combination of T25-TssA$^S_{NTR}$ with T18-TssF was the only combination of TssA$^S_{NTR}$ and TssF two-hybrid proteins that yielded a Mal$^+$ phenotype, although the colonies exhibited the patchy morphology.

β-galactosidase assays were performed for all the $TssA^S$ and TssF hybrid protein combinations that were assessed MacConkey-maltose agar (Figure 3.8). The combinations in which the C-terminus of both $TssA^S$ and TssF are free, gave β-galactosidase activities of 826.1 Mu (T25-TssF and T18-$TssA^S$) and 1012.5 Mu (T25-$TssA^S$ and T18-TssF), respectively. However, the standard deviations of both combinations were almost as high as the mean values of the Miller units that were calculated from 42 and 30 measurements, respectively. One hypothesis to explain the variation in the β-galactosidase activity is that it may be caused by unstable or transient interaction between these $TssA^S$ and TssF chimaeric proteins, and it may be related to the patchy Mal^+ phenotype that was observed for these two combinations. In contrast, the β-galactosidase activity measurements for the TssF-T25 with T18-$TssA^S$ combination, which gave rise to uniform Mal^+ colonies, was less variable (805.6 ± 86.6 Mu).

Consistent with the variable β-galactosidase activity of the combinations in which the C-terminus of TssF was free, $TssA^S_{CTD}$ fused to the N- or C-terminus of T25 in combination with T18-TssF also led to a variable β-galactosidase activity with mean values of 791.7 Mu and 1519.7 Mu, respectively. The $TssA^S_{CTD}$-T25 and T18-TssF combination exhibited less variance in β-galactosidase activity (1519.7 ± 559.4 Mu) compared to the other combination but it was high nevertheless. The only combination of $TssA^S_{NTR}$ and TssF that yielded a Mal^+ phenotype gave rise to a β-galactosidase activity of only 210.8 ± 142.3 Mu (p = .048), less than 3-fold higher than the background. For the remaining two-hybrid combinations, the β-galactosidase activities were similar to the background level that was measured in the negative controls (Figure 3.8). In conclusion, these results suggest that TssF interacts with $TssA^S$, and this interaction is likely to involve $TssA^S_{CTD}$. Given the very low β-galactosidase activity of the only combination of $TssA^S_{NTR}$ and TssF fusions that yielded a Mal^+ phenotype, a firm conclusion concerning the role of $TssA^S_{NTR}$ cannot be made.

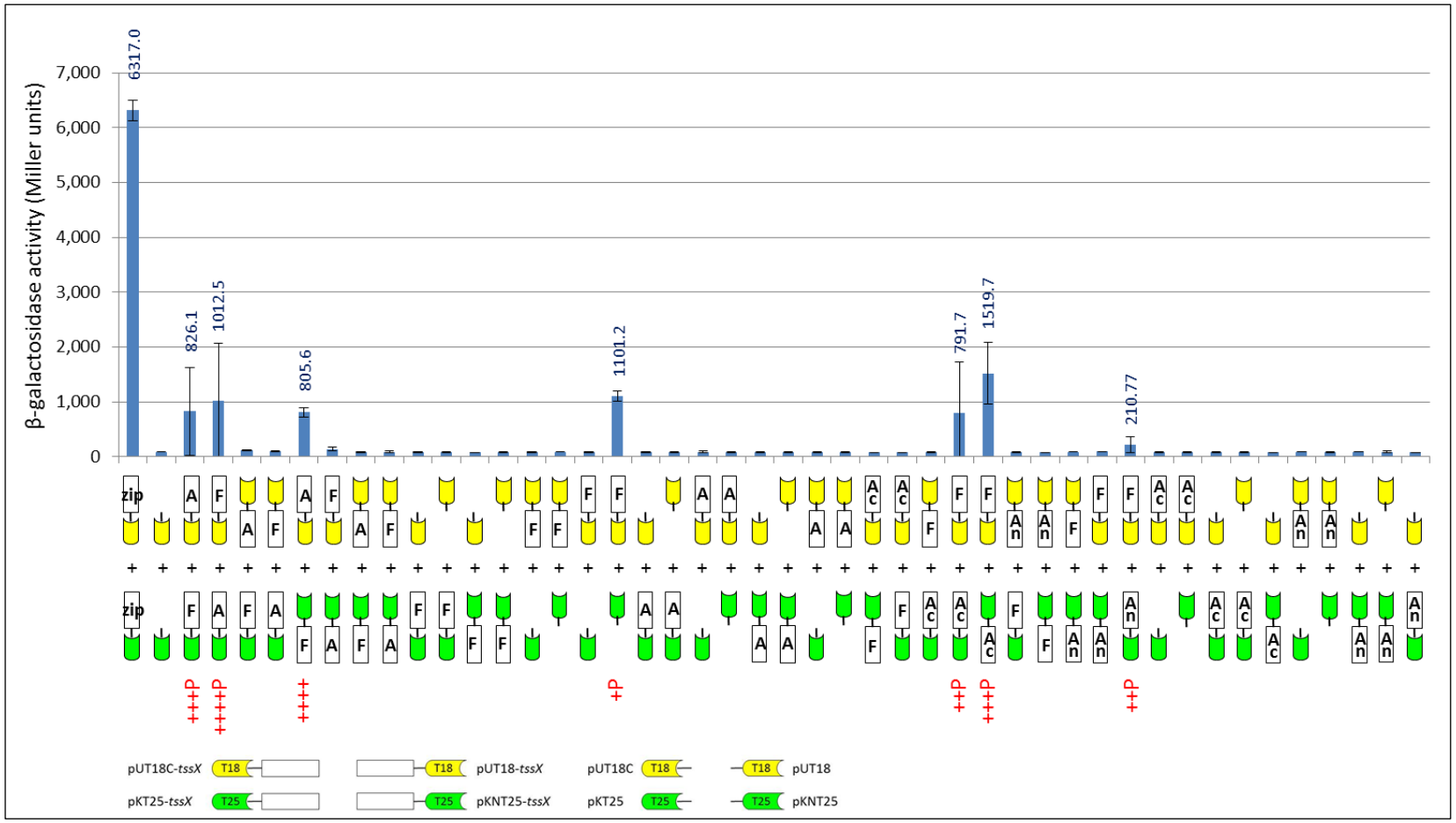


Figure 3.8 BACTH analysis of the interaction between TssA^S and TssF. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused to TssF, TssA^S and TssA^S domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2; 'P', patchy. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'F', 'An' and 'Ac' represent TssA^S, TssF, TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively.

3.3.7 Analysis of the interaction between $TssA^S$ and TssG

Four pairwise BACTH plasmid combinations encoding $TssA^S$ and TssG hybrid proteins were introduced into the BACTH assay strain and the resulting Mal phenotypes were recorded. In two of the combinations the C-terminus of both $TssA^S$ and TssG were free. The other two combinations were TssG-T25 with T18-$TssA^S$, and $TssA^S$-T25 with T18-TssG. All of these combinations gave rise to a negative maltose phenotype. The results are consistent with previous observations (Shastri 2011).

As the β-galactosidase assay is more sensitive than scoring the Mal phenotype, this assay was conducted for all the $TssA^S$ and TssG hybrid protein combinations that were analysed on MacConkey-maltose agar. However, the β-galactosidase activities of all the combinations were similar to the background level that was measured in the negative controls, which is in accordance with the observed maltose phenotypes (Figure 3.9). Therefore, these results suggest that TssG does not interact with $TssA^S$.

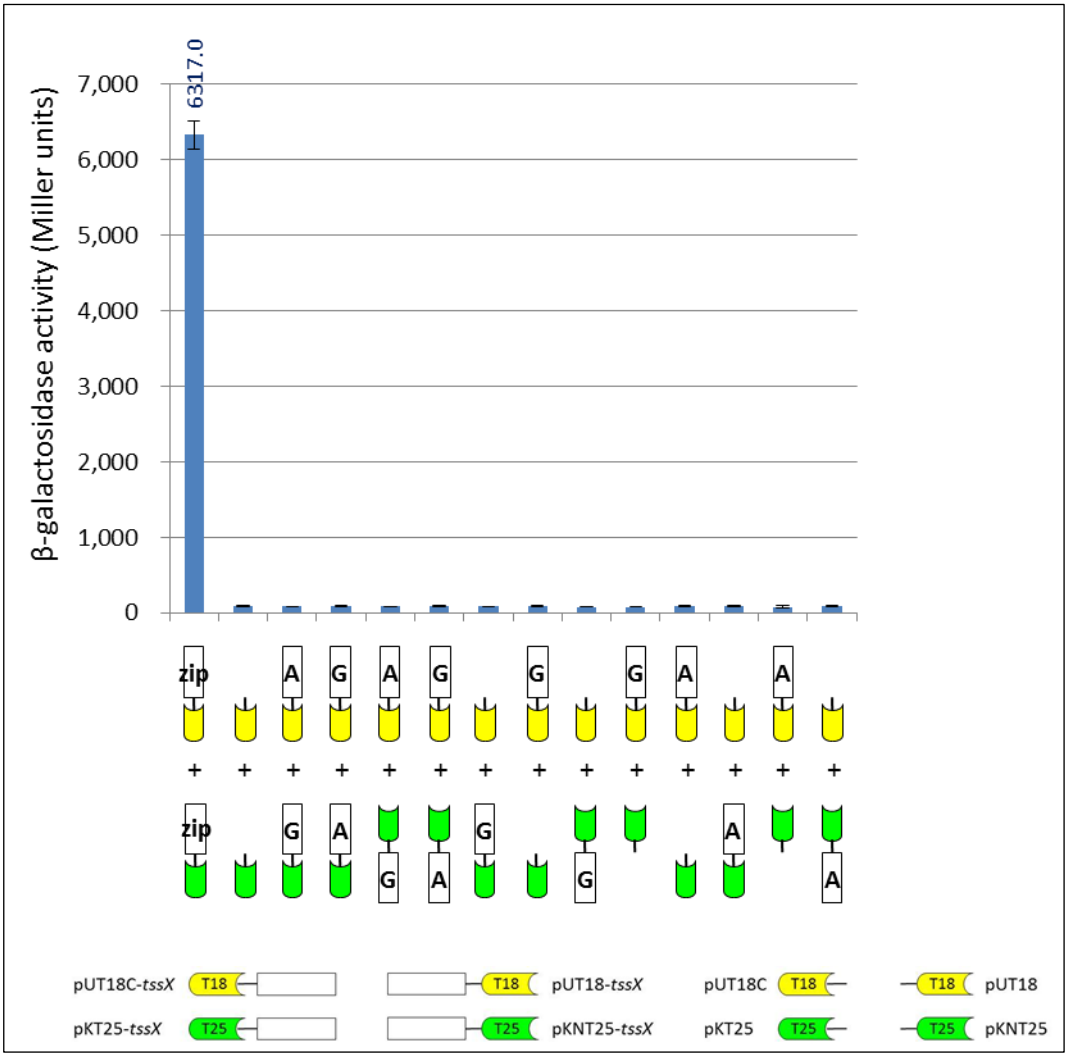


Figure 3.9 BACTH analysis of the interaction between TssA^S and TssG. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused to TssA^S and TssG were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A' and 'G' represent TssA^S and TssG, respectively.

3.3.8 Analysis of interactions between $TssA^S$ and TssH

Two out of eight possible compatible BACTH plasmid combinations encoding $TssA^S$ and TssH fusions to CyaA domains were assayed, which were T25-TssH and T18-$TssA^S$, and T25-$TssA^S$ and T18-TssH. Both combinations yielded a very strong Mal^+ phenotype with dark red/purple colonies. To identify the domain of $TssA^S$ that is involved in the interaction with TssH, all fusion proteins in which the C-terminus of $TssA^S_{CTD}$ and N-terminus of $TssA^S_{NTR}$ are free were combined with T25-TssH or T18-TssH as appropriate. Both combinations of $TssA^S_{CTD}$ with TssH led to a strong Mal^+ phenotype, whereas, both of the $TssA^S_{NTR}$ and TssH combinations yielded a maltose-negative phenotype. The results revealed that $TssA^S_{CTD}$ was responsible for the interaction with TssH, in agreement with previous observations (Ahmad 2013).

β-galactosidase assays were performed for all the $TssA^S$- and TssH-CyaA domain fusion combinations that were analysed on MacConkey-maltose agar (Figure 3.10). The two combinations in which the C-terminus of both $TssA^S$ and TssH are free gave β-galactosidase activities of approximately 2800 Mu with standard deviations that were 5-10% of the mean values. The β-galactosidase activity of the combination of T25-TssH with T18-$TssA^S_{CTD}$ was almost as high as that of the full-length $TssA^S$ and TssH two-hybrid combinations (2609.0 $\pm$ 105.7 Mu), whereas when TssH and $TssA^S_{CTD}$ were fused to the other termini of T18 and T25, respectively, the β-galactosidase activity decreased by ~60%, i.e. 1090.2 $\pm$ 42.7 Mu. This matched the weaker Mal^+ phenotype that was observed on MacConkey-maltose agar. The remaining combinations gave rise to β-galactosidase activities that were similar to the background level measured in the negative controls. In conclusion, these results suggest that TssH interacts with $TssA^S$, and the C-terminal domain of $TssA^S$ is involved in the interaction with TssH.

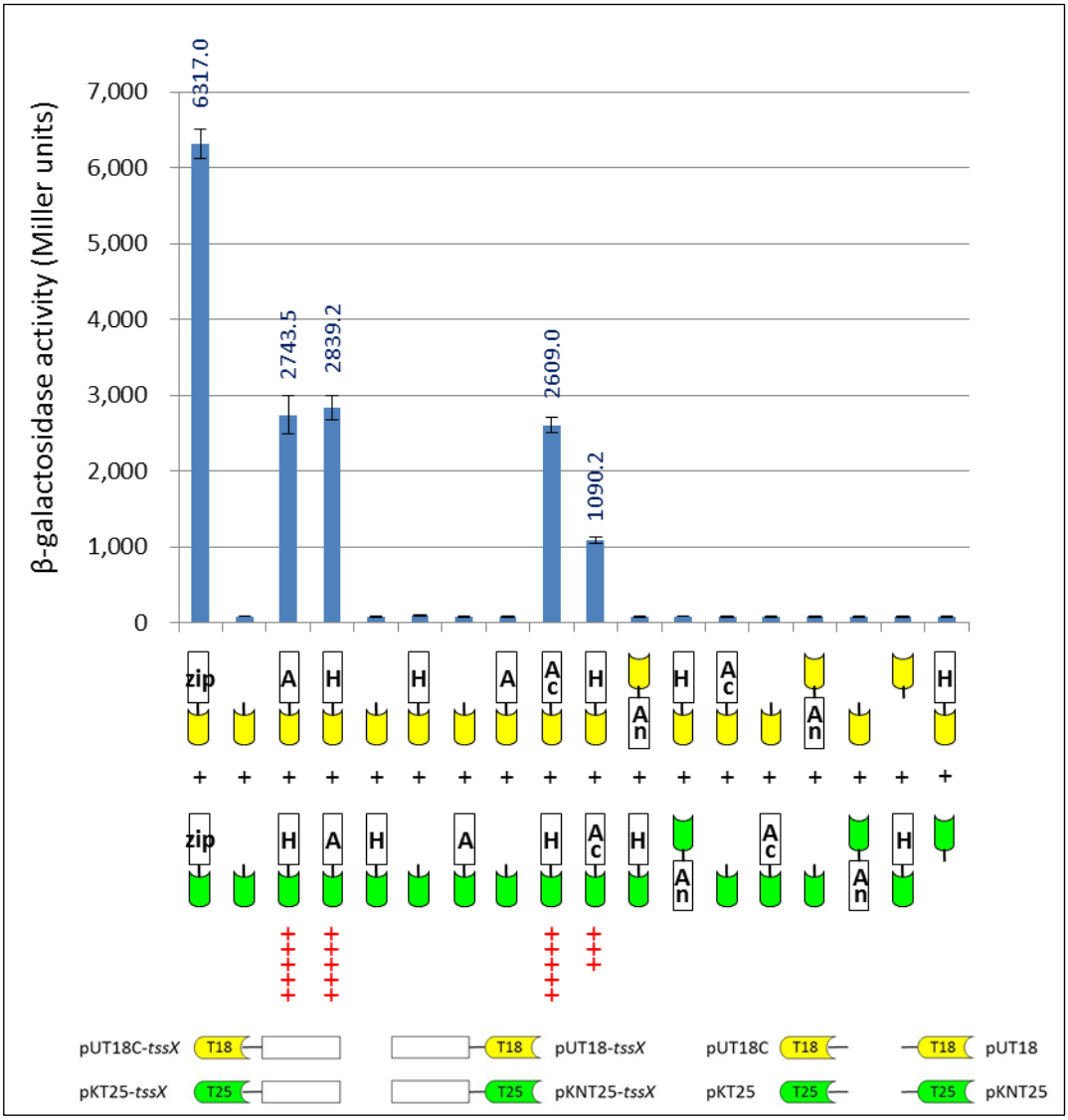


Figure 3.10 BACTH analysis of the interaction between TssA^S and TssH. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused to TssH, TssA^S and TssA^S domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'H', 'An' and 'Ac' represent TssA^S, TssH, TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively.

3.3.9 Analysis of interactions between $TssA^S$ and TssI

Two out of eight possible pairwise combinations of $TssA^S$ and TssI two-hybrid proteins were analysed by the BACTH assay, both of which gave the C-terminus of $TssA^S$ and TssI free. Both combinations yielded a very strong Mal^+ phenotype with dark red/purple colonies. The results are in agreement with previous observations (Ahmad 2013).

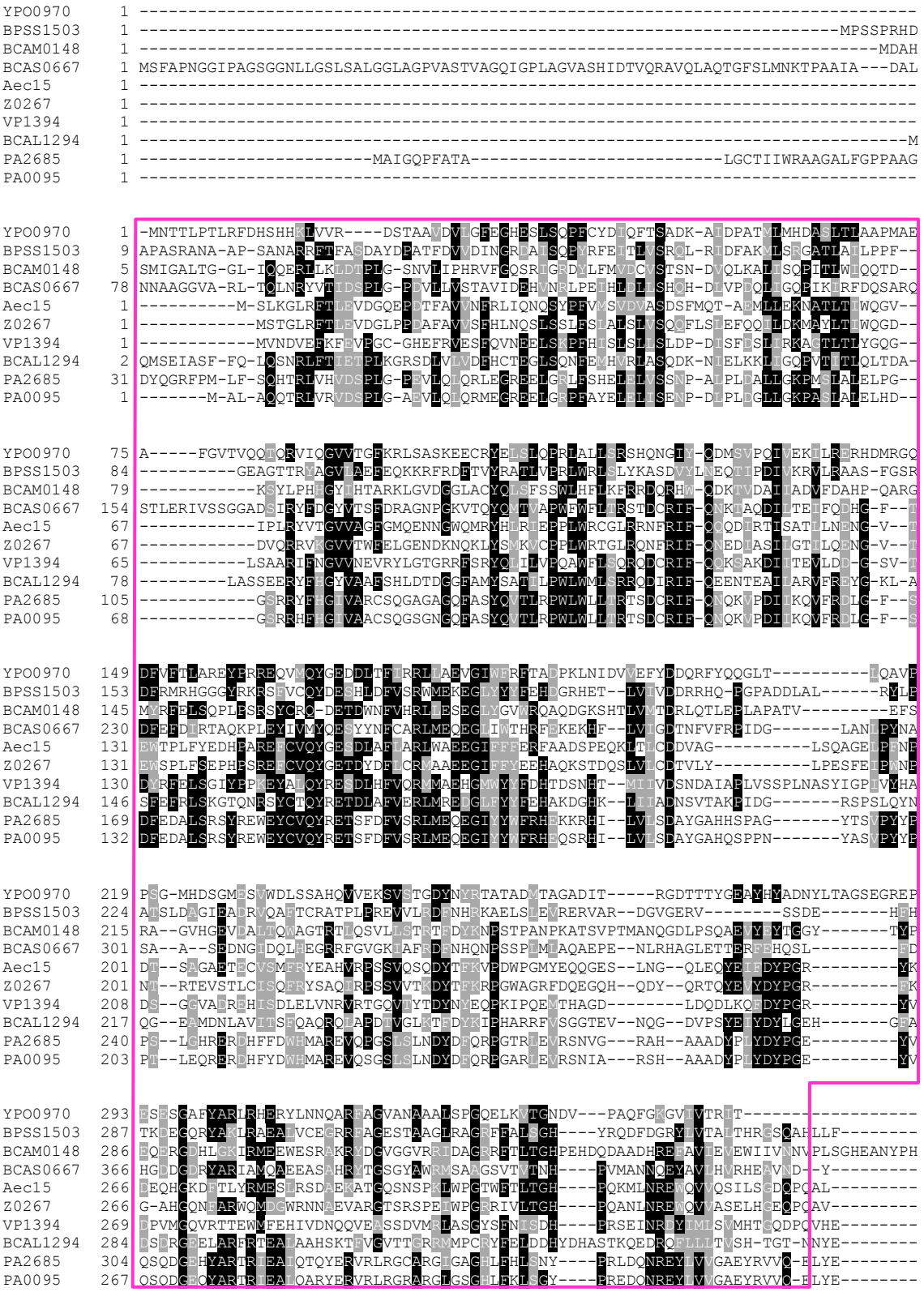
Many studies have shown that TssI forms a trimer, and it shares structural homology with the $(gp27)_3$-$(gp5)_3$ complex that makes up the bacteriophage T4 tail spike which functions as a puncturing device towards the target cell, permitting the delivery of the viral DNA (Kanamaru et al. 2002; Pukatzki et al. 2007; Leiman et al. 2009). It has been suggested that TssI locates on the top of the tail tube that is formed by stacked hexameric rings of TssD (Pukatzki et al. 2007; Hachani et al. 2011). Multiple sequence alignment of TssI amino acid sequences from orthologues present in various Gram-negative bacteria was carried out (Figure 3.11). The alignment indicates the core regions of TssI that correspond to the bacteriophage T4 gp5 and gp27 proteins, in agreement with a previously published study (Leiman et al. 2009). This alignment included three TssI proteins from *B. cenocepacia* J2315, including BCAM1048 (the BCAM0148 orthologue present in strain 715j was used in the BACTH assay). Therefore, the gp27- and gp5-homologous regions of BCAM0148 could be identified. To identify the region(s) of TssI that are involved in the interaction with $TssA^S$, DNA fragments encoding the gp27- and gp5-like domains of the BCAM0148-like TssI from *B. cenocepacia* 715j and the fused gp27gp5 region were cloned into pKT25 and pUT18C (Ahmad 2013). These two vectors were chosen as the two combinations of two-hybrid proteins in which the C-terminus of both $TssA^S$ and TssI are unfused had yielded a Mal^+ phenotype in the BACTH assay.

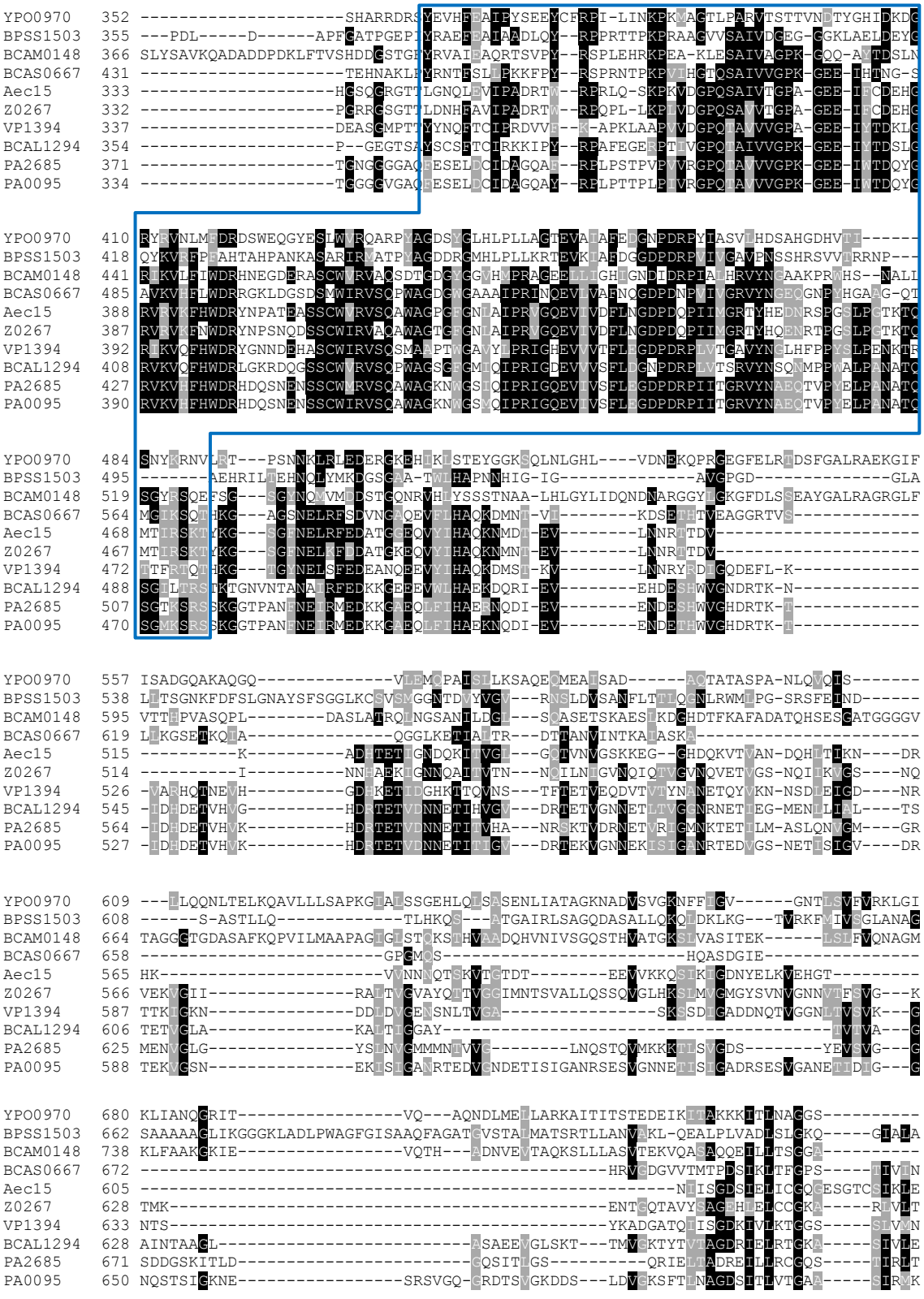
The BACTH assay was carried out for the gp27gp5-like fusion, and the separate gp27- and gp5-like regions of TssI in combination with the compatible T25-$TssA^S$ or T18-$TssA^S$ two-hybrid proteins. In addition, fusions of $TssA^S_{NTR}$ to the N-termini of T25 and T18 were assayed in combination with fusions in which the C-terminus of TssI was free, and fusions of $TssA^S_{CTD}$ to the C-termini of T25 and T18 were assayed in combination with two-hybrid proteins in which TssI, and the gp27gp5-, gp27- and gp5-like moieties of TssI ($TssI_{gp27gp5}$, $TssI_{gp27}$ and $TssI_{gp5}$) were fused to the C-termini of T25 and T18 (Figure 3.12).

The two-hybrid combinations containing the $TssI_{gp27gp5}$ and $TssA^S$ led to a very strong Mal^+ phenotype. However, when the individual $TssI_{gp27}$ and $TssI_{gp5}$ were tested, they gave rise to a negative maltose phenotype when combined with $TssA^S$. The results suggest that the entire gp27gp5-like region of TssI may be required for the interaction with $TssA^S$. Furthermore, in agreement with a previous report (Ahmad 2013), one of the $TssA^S_{CTD}$ and TssI combinations yielded a strong Mal^+ phenotype with red/purple colonies, whereas both $TssA^S_{NTR}$ and TssI combinations gave rise to a negative maltose phenotype. Interestingly, one of the control combinations, pKNT25 and pUT18C-*tssI*, gave rise to pink/red colonies, whereas the corresponding combination with two fusion proteins (pKNT25-$tssA^S_{NTR}$ and pUT18C-*tssI*) yielded a negative maltose phenotype. These results suggest the CTD of $TssA^S$ is responsible for the interaction of $TssA^S$ with TssI. Subsequently, potential interactions between $TssI_{gp27gp5}$, and $TssI_{gp27}$ and $TssI_{gp5}$ on the one hand, and $TssA^S_{CTD}$ on the other, were investigated. The results revealed that the only combination that yielded a Mal^+ phenotype was the combination of the T18-$TssI_{gp27gp5}$ fusion with T25-$TssA^S_{CTD}$ that gave rise to dark purple/red colonies. All of the remaining domain combinations gave rise to a maltose-negative phenotype. This result supported the hypothesis that the gp27gp5-like region of TssI and CTD region of $TssA^S$ are responsible for the interaction between their corresponding full-length proteins.

β-galactosidase assays were carried out for all the $TssA^S$ and TssI two-hybrid protein combinations that were assayed on MacConkey-maltose agar, including their corresponding domains. The β-galactosidase activity of the combination of T25-TssI with T18-$TssA^S$ was 1267.4 ± 17.9 Mu, whereas for T25-$TssA^S$ and T18-TssI, the β-galactosidase activity was two times higher, 2519.1 ± 121.5 Mu. The two $TssA^S$ and $TssI_{gp27gp5}$ two-hybrid combinations that were tested also gave high β-galactosidase activities which were 2281.7 ± 74.3 Mu and 1538.9 ± 97.5 Mu (Figure 3.12). The combination of T25-$TssA^S_{CTD}$ with T18-TssI gave rise to a β-galactosidase activity of 1952.2 ± 159.0 Mu, whereas the combination of T25-TssI with T18-$TssA^S_{CTD}$ gave rise to a similar β-galactosidase activity as the background controls. When pairwise combinations of the $TssI_{gp27gp5}$ and $TssA^S_{CTD}$ were assayed, one combination, T25-$TssA^S_{CTD}$ with T18-$TssI_{gp27gp5}$, gave rise to a β-galactosidase activity that was similar to that of T25-$TssA^S$ with either T18-TssI or T18-$TssI_{gp27gp5}$, i.e. 2396.3 ± 237.2 Mu, whereas the other combination, T25-$TssI_{gp27gp5}$ with T18-$TssA^S_{CTD}$, had a similar

activity to the background. The remaining combinations gave β-galactosidase activities that were similar to the background level that was measured in the negative controls, and the results of β-galactosidase assays are in agreement with the maltose phenotypes observed. Thus, the result supports the idea that TssI interacts with $TssA^{S}$, and this interaction occurs through the core TssI region and $TssA^{S}_{CTD}$.





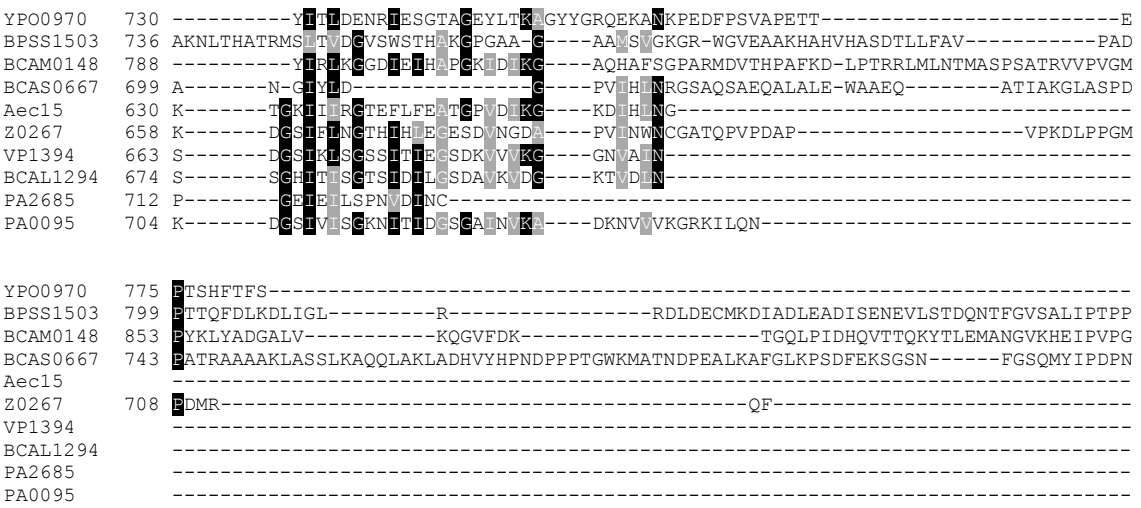


Figure 3.11 Amino acid sequence alignment of TssI proteins from various Gram-negative bacteria. Multiple sequence alignment was carried out between TssI from *Yersinia pestis* CO92 (YPO0970), *Burkholderia pseudomallei* K96243 (BPSS1503), *Burkholderia cenocepacia* J2315 (BCAM0148, BCAS0667 and BCAL1294), *Escherichia coli* S88 (Aec15), *Escherichia coli* serotype O157:H7 EDL933 (Z0267), *Vibrio parahaemolyticus* serotype O3:K6 RIMD 2210633 (VP1394) and *Pseudomonas aeruginosa* PAO1 (PA2685 and PA0095) using Clustal Omega. The conserved amino acids are shaded by Boxshade program. Amino acids that are identical at the same position in at least 50% of sequences are highlighted in black, whereas similar amino acids are highlighted in grey. gp27 and gp5 homology regions are enclosed with pink and blue boxes, respectively.

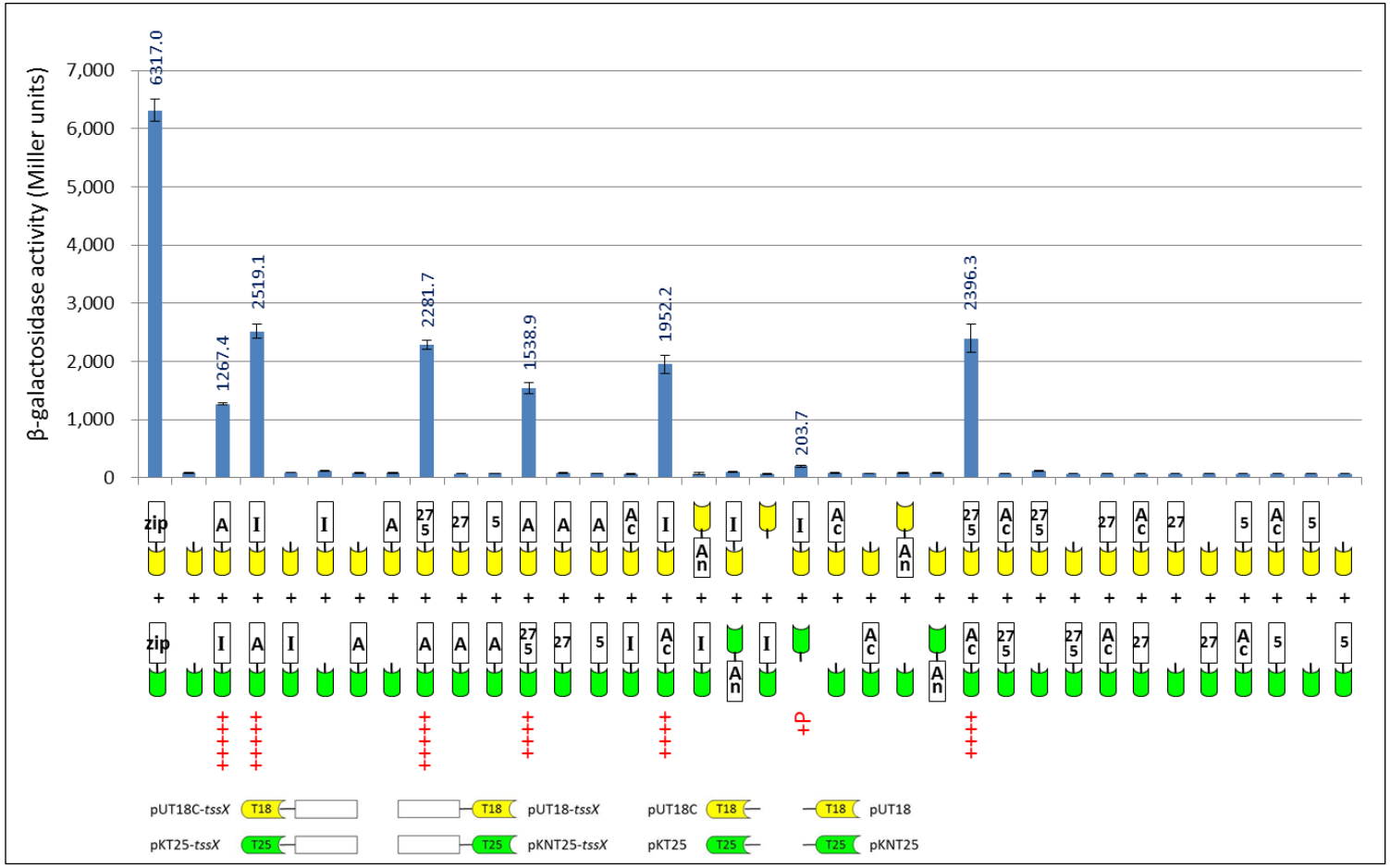


Figure 3.12 BACTH analysis of the interaction between TssA^S and TssI and their domains. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused of TssI, TssA^S and TssA^S domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2; 'P', patchy. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'I', 'An' and 'Ac' represent TssA^S, TssI, TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively. '27-5', '27' and '5' stand for the fused gp27gp5-, gp27- and gp5-like moieties of TssI, respectively.

3.3.10 Analysis of the interaction between $TssA^S$ and TssJ

Four combinations of $TssA^S$ and TssJ fusion proteins were assayed using the BACTH system. Only one combination, TssJ-T25 and T18-$TssA^S$, gave rise to a very weak Mal^+ phenotype with pink colonies. The remaining combinations generated colonies with a maltose-negative phenotype. The results are in conflict with the previous observation, where all the pairwise combinations of $TssA^S$ with TssJ were reported to give a negative maltose phenotype (Ahmad 2013). As TssJ-T25 and T18-$TssA^S$ yielded a weak Mal^+ phenotype, the fusion proteins in which the N-terminus of TssJ are free were subsequently tested in combination with $TssA^S_{CTD}$ fused to the C-terminus of T25 and T18, and $TssA^S_{NTR}$ fused to the N-terminus of T25 and T18. However, all of these four combinations led to a maltose-negative phenotype.

β-galactosidase assays were performed for all the $TssA^S$ and TssJ fusion protein combinations that were assayed on MacConkey-maltose agar (Figure 3.13). The β-galactosidase activity of the combination of TssJ-T25 with T18-$TssA^S$ was 203.1 ± 41.7 Mu. An unpaired t-test was carried out to further compare it with the background level of β-galactosidase activity that was measured in the control groups. This showed that the β-galactosidase activity in cells containing TssJ-T25 and T18-$TssA^S$ was statistically significantly higher than that of the control group (p=.001). The remaining combinations gave rise to β-galactosidase activities that were similar to the background level that was measured in the negative controls. As a low β-galactosidase activity (less than 3-fold higher than the background level) was measured for the only $TssA^S$ and TssJ fusion protein combination to yield a Mal^+ phenotype, there could be a weak interaction between $TssA^S$ and TssJ that is not physiologically relevant.

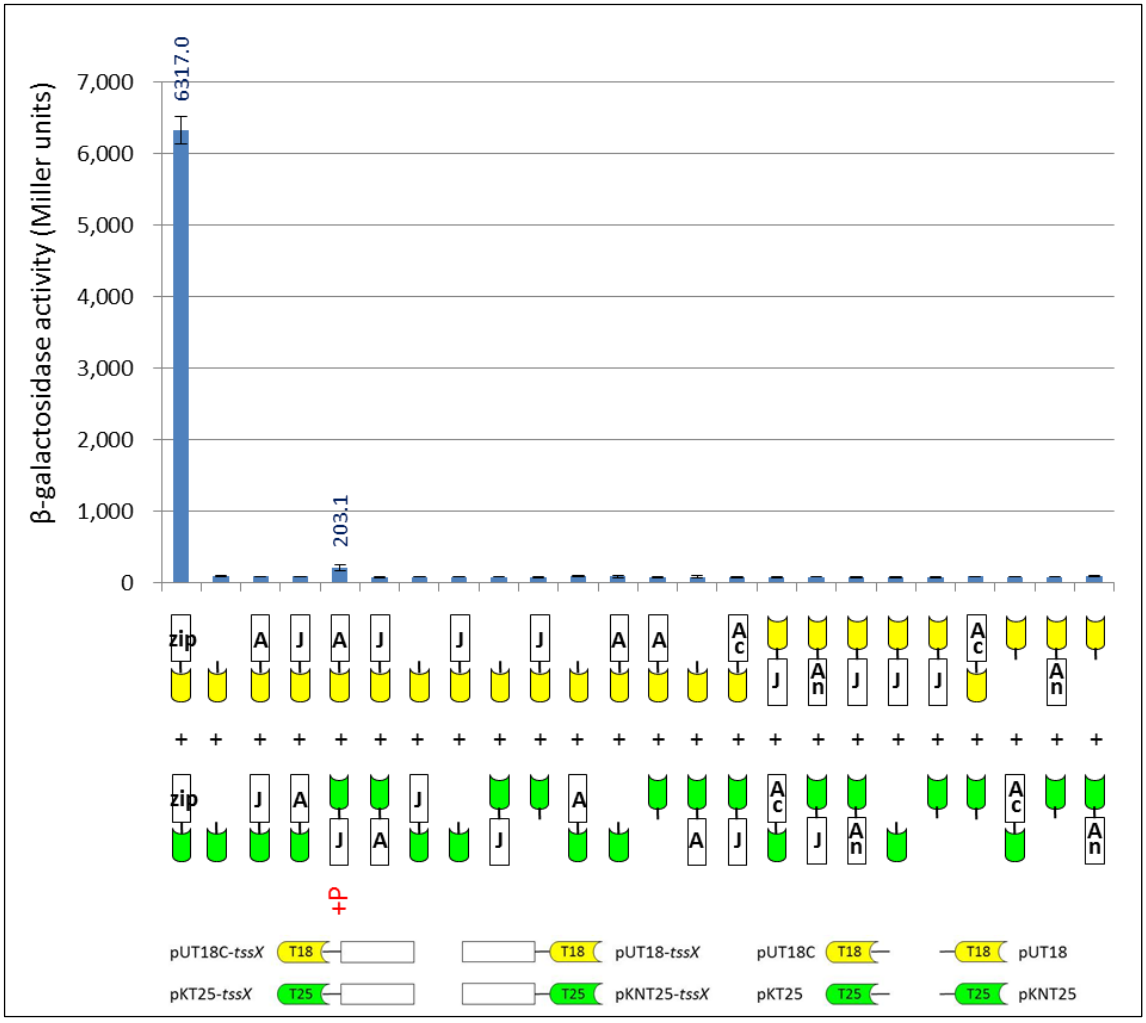


Figure 3.13 BACTH analysis of the interaction between TssA^{S} and TssJ. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused to TssJ, TssA^{S} and TssA^{S} domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2; 'P', patchy. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'J', 'An' and 'Ac' represent TssA^{S}, TssJ, TssA$^{S}_{NTR}$ and TssA$^{S}_{CTD}$, respectively.

3.3.11 Analysis of the interaction between $TssA^S$ and TssK

Two out of eight possible pairwise combinations of $TssA^S$ with TssK fusion proteins were assayed in the BACTH system, in which the C-termini of both $TssA^S$ and TssK were free. Both combinations yielded a Mal^+ phenotype with red/purple colonies. $TssA^S$ domains were then investigated in combination with TssK fused to the C-terminus of T25 or T18 to identify the domain of $TssA^S$ involved in the interaction with TssK. However, $TssA^S_{NTR}$-CyaA domain fusions with $TssA^S_{NTR}$ located at the N-terminus and CyaA domain-$TssA^S_{CTD}$ fusions with $TssA^S_{CTD}$ located at the C-terminus gave rise to a negative maltose phenotype when they were combined with T25- and T18-TssK fusions. The result is not consistent with a previous observation, where T25-$TssA^S_{CTD}$ in combination with T18-TssK was recorded as giving rise to a weak Mal^+ phenotype with patchy pattern (Ahmad 2013).

β-galactosidase assays were carried out for all the $TssA^S$ and TssK fusion protein combinations that were assessed on MacConkey-maltose agar (Figure 3.14). The β-galactosidase activity of both combinations of $TssA^S$ with TssK were similar, 2564.4 ± 147.5 Mu (T25-TssK and T18-$TssA^S$) and 2184.1 ± 53.5 Mu (T25-$TssA^S$ and T18-TssK), respectively. The β-galactosidase activities of the remaining combinations were similar to the background level that was measured in the negative controls. Therefore, the results of the β-galactosidase assays are in agreement with the maltose phenotypes observed. These results suggest that $TssA^S$ interacts with TssK and that both domains of $TssA^S$ are required for this interaction.

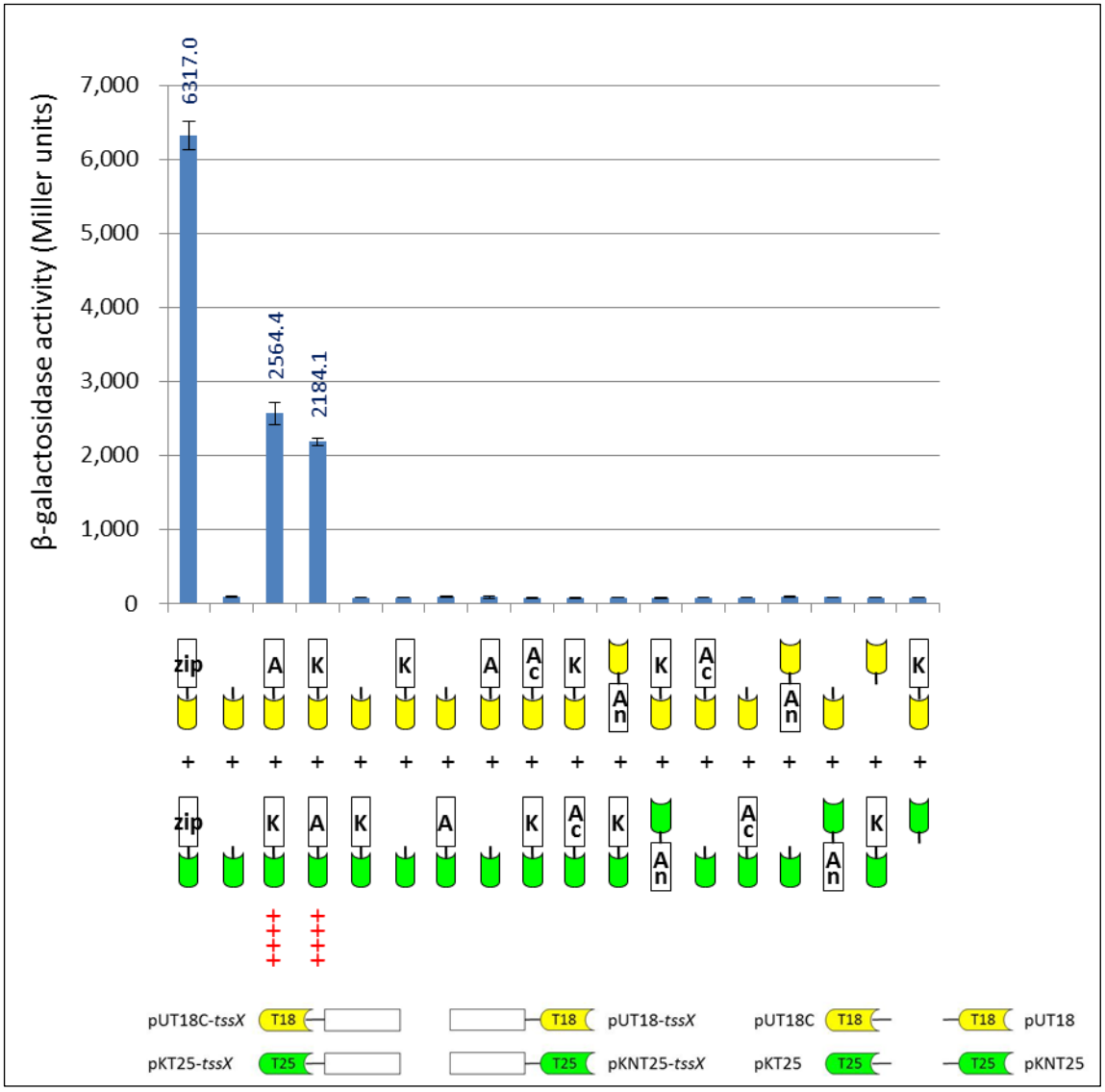


Figure 3.14 BACTH analysis of the interaction between TssA^S and TssK. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused to TssK, TssA^S and TssA^S domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'K', 'An' and 'Ac' represent TssA^S, TssK, TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively.

3.3.12 Analysis of the interaction between $TssA^S$ and TssL

Two out of eight possible compatible BACTH plasmid combinations encoding $TssA^S$ and TssL fusion proteins in which the C-terminus of both $TssA^S$ and TssL were free were analysed in the BACTH assay. Both combinations yielded a Mal^+ phenotype with moderate red/purple colonies. The interaction of $TssA^S$ domains with T25-TssL or T18-TssL was then investigated. The results showed that all the combinations tested, in which the N-terminus of $TssA^S_{NTR}$ and C-terminus $TssA^S_{CTD}$ were free, led to a maltose-negative phenotype. These results are in accordance with previous observations (Ahmad 2013).

β-galactosidase assays were performed for all the $TssA^S$ and TssL fusion protein combinations that were analysed on MacConkey-maltose agar (Figure 3.15). The two combinations in which the C-terminus of $TssA^S$ and TssL were free gave rise to β-galactosidase activities of 701.7 ± 155.7 Mu (T25-TssL and T18-$TssA^S$) and 1302.0 ± 50.0 Mu (T25-$TssA^S$ and T18-TssL), respectively. The β-galactosidase activities of the remaining combinations were similar to the background level that was measured in the negative controls. The results of the β-galactosidase assays are consistent with the maltose phenotypes observed, and suggest TssL interacts with $TssA^S$ and both domains of $TssA^S$ are required for this interaction.

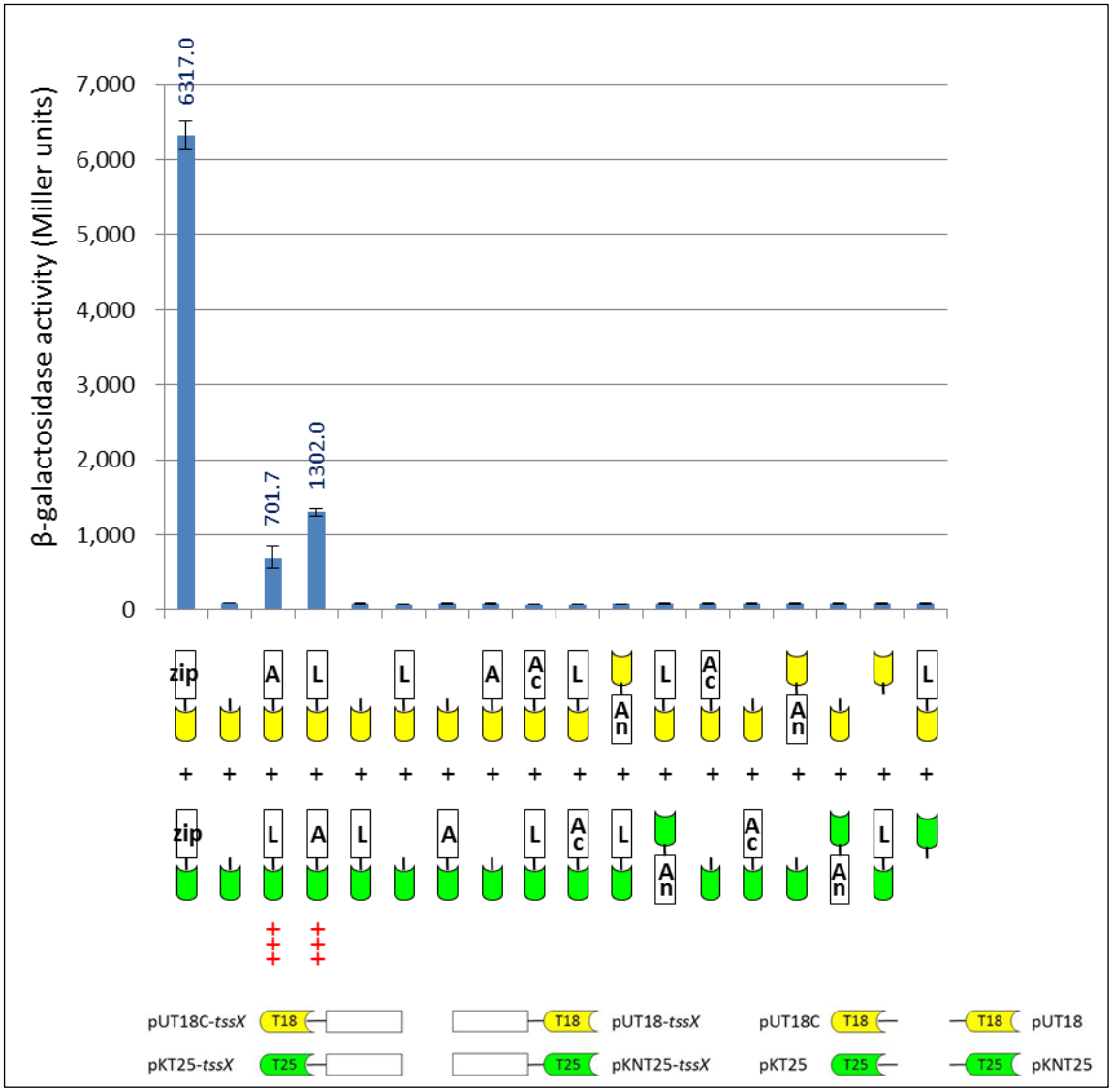


Figure 3.15 BACTH analysis of the interaction between TssA^S and TssL. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused to TssL, TssA^S and TssA^S domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'L', 'An' and 'Ac' represent TssA^S, TssL, TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively.

3.3.13 Analysis of the interaction between $TssA^S$ and TssM

TssM is anchored in the IM by three transmembrane helices, two of which are located at the N-terminus and the third is located at the C-terminus of the cytoplasmic N-terminal domain which corresponds to approximately one third of the length of the protein. The remaining two thirds of TssM extend into the periplasm with the C-terminal end coming in contact with TssJ (Durand et al. 2015). There is a large cytoplasmic domain between the second and third transmembrane domain that carries Walker A and B motifs in a subset of TssM orthologues (Zheng and Leung 2007). The Walker B motif of TssM has been demonstrated to be required for ATPase activity (Ma et al. 2012). As the BACTH system can only be used to screen for protein-protein interactions in the cytoplasm (as it depends on the presence of ATP), interactions may occur between $TssA^S$ and the periplasmic domain of TssM cannot be analysed using the native TssM protein. Therefore, to ensure that the C-terminal periplasmic domain of TssM was produced in the cytoplasm in the BACTH assay, DNA encoding *B. cenocepacia* TssM was divided into two fragments which were amplified separately for cloning into the BACTH system plasmids (Ahmad 2013). One DNA fragment encoded $TssM_{NTD}$, which is the cytoplasmic domain between transmembrane helices 2 and 3 (amino acids 58-445) and the other encoded $TssM_{CTD}$, which is the large periplasmic domain (amino acids 466-1314). DNA encoding the TMDs was omitted from these plasmids.

In the previous investigation, where all the pairwise combinations of $TssA^S$ and TssM domain two-hybrid fusions led to a negative maltose phenotype, it was subsequently determined that the experiments were conducted with defective pKNT25-$tssM_{NTD}$ and pUT18-$tssM_{NTD}$ plasmids (Ahmad 2013). The correction of pKNT25-$tssM_{NTD}$ and pUT18-$tssM_{NTD}$ plasmids was carried out as described in Section 3.2.3. All of the eight possible compatible BACTH plasmid combinations encoding $TssA^S$ and $TssM_{NTD}$ two-hybrid proteins were assayed, along with four combinations of $TssA^S$ with $TssM_{CTD}$. One combination in which the N-terminus of $TssM_{NTD}$ and the C-terminus of $TssA^S$ were free led to a Mal^+ phenotype with moderate red/purple colonies on maltose-MacConkey agar. This is a new finding that suggests an interaction between $TssA^S$ and $TssM_{NTD}$. All of the remaining combinations gave rise to a negative maltose phenotype. Further investigations were carried out involving three pairwise combinations of $TssA^S_{NTR}$ with $TssM_{NTD}$-T25 or $TssM_{NTD}$-T18 and three pairwise

combinations of $TssA^{S}_{CTD}$ with $TssM_{NTD}$ two-hybrid fusions. However, all of these combinations led to a maltose-negative phenotype.

β-galactosidase assays were performed for all the combinations of $TssA^{S}$ and $TssM_{NTD}$ two-hybrid proteins that were analysed on MacConkey-maltose agar (Figure 3.16). The combination of $TssM_{NTD}$-T25 with T18-$TssA^{S}$ gave a β-galactosidase activity of 544.1 ± 22.3 Mu, whereas the activities of all the remaining combinations were similar to the background level that was measured in the negative controls. Therefore, the results of the β-galactosidase assays are consistent with the maltose phenotypes that were observed on MacConkey-maltose agar. These results indicate that $TssA^{S}$ interacts with the cytoplasmic domain of TssM and that both domains of $TssA^{S}$ are required for this interaction.

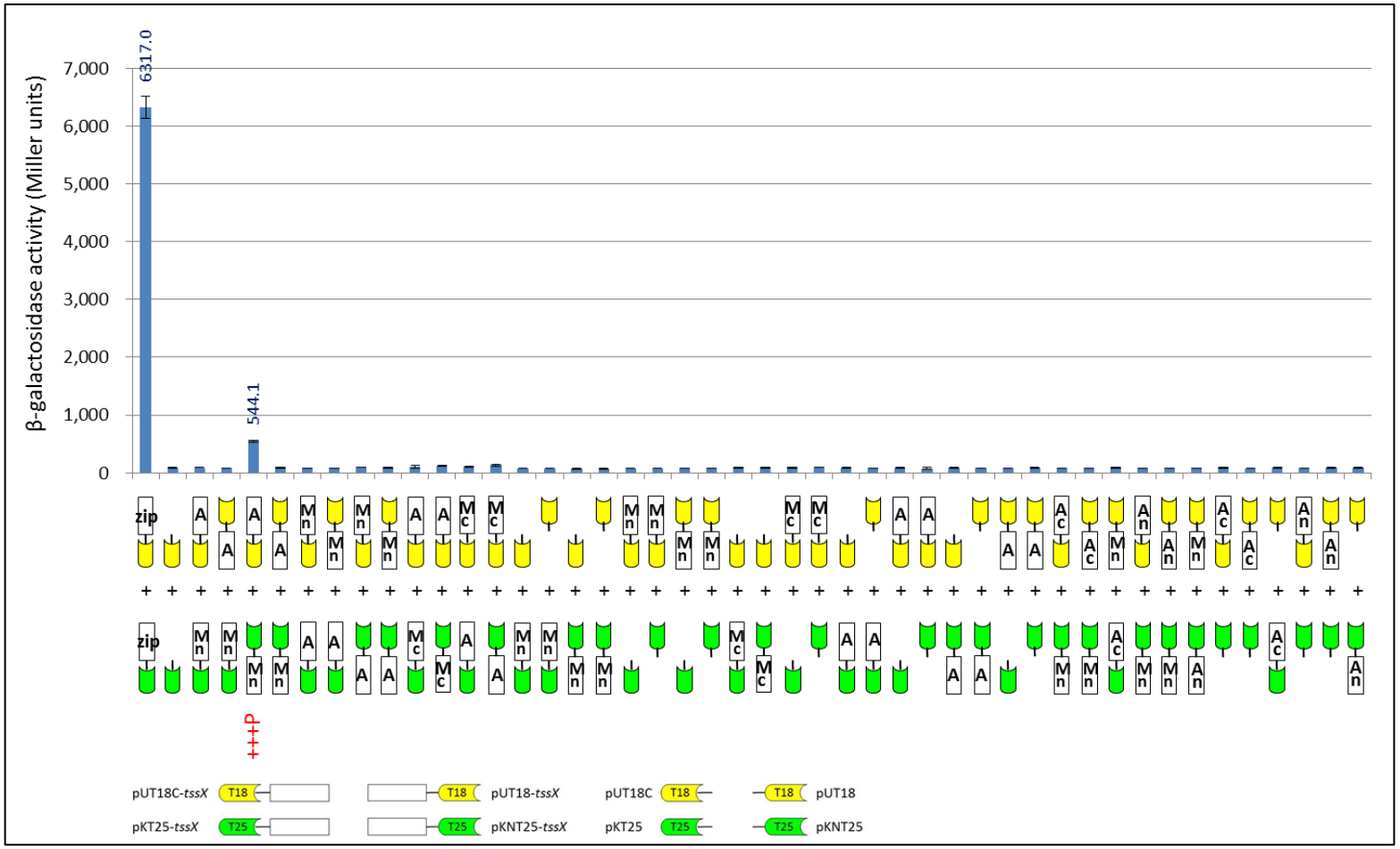


Figure 3.16 BACTH analysis of the interaction between TssA^S and TssM. Compatible BACTH plasmids encoding the CyaA T25 or T18 domains fused to TssM, TssA^S and TssA^S domains were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2; 'P', patchy. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'An' and 'Ac' represent TssA^S, TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively. 'Mn' and 'Mc' stand for TssM$_{NTD}$ and TssM$_{CTD}$, respectively.

3.4 Three-hybrid analysis of interactions between $TssA^S$ and other T6SS subunits

To investigate whether the presence of a third, unfused, T6SS subunit (TssY) can facilitate or enhance interactions between two other T6SS subunits in the two-hybrid assay, the DNA encoding the unfused subunit was introduced into a BACTH plasmid downstream of a pre-existing *cyaA-tssX* ORF fusion. This type of analysis has been referred to as a 'three-hybrid assay' although in fact there remain only two hybrid proteins in the assay along with a native protein (Karimova et al. 2005).

Three-hybrid plasmids were constructed based on the BACTH plasmid pUT18C, as generally *tssX* genes inserted into this plasmid were more often observed to give rise to a Mal^+ phenotype when they were combined with $TssA^S$ fusions. pUT18C is also a high copy number plasmid, so there is a higher probability of ensuring sufficient expression of the additional *tss* gene (*tssY*). Moreover, as pUT18C has the MCS located downstream of the T18 ORF, this facilitated the inclusion of the third gene (*tssY*) in the system by cloning it into the vector downstream of the T18-TssX fusion for constructing pUT18C-*tssX*-*tssY*. The forward primer for amplification of the third gene contained an additional stop codon to ensure no T18-TssX-TssY fusion protein was formed, and a Shine-Dalgarno sequence for *tssY* translation (Figure 3.17).

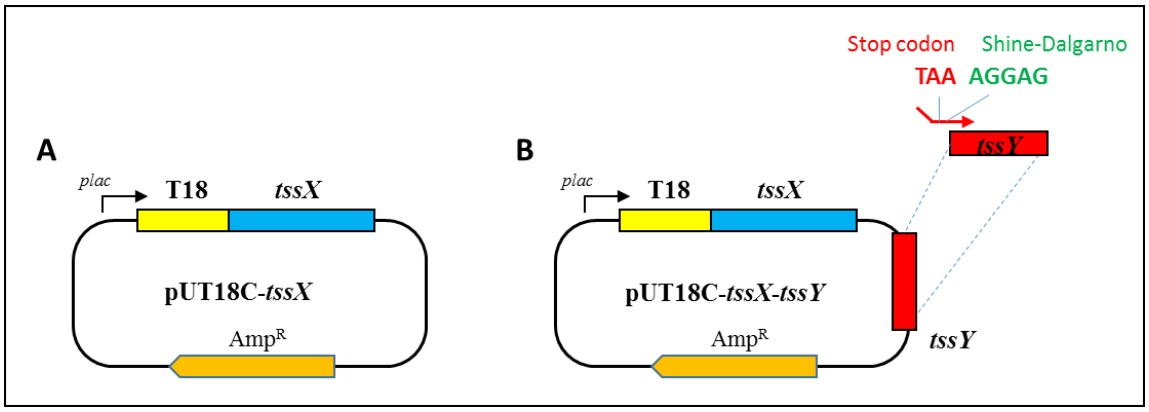


Figure 3.17 Construction of a two-cistron expression plasmid for use in a three-hybrid system. A. Schematic representation of *tssX* inserted pUT18C plasmid. DNA encoding TssX is located downstream of the T18 coding sequence that gives rise to a T18-TssX fusion protein. **B**. A second T6SS subunit gene (*tssY*) was inserted into the pUT18C-*tssX* plasmid using the remaining available restriction sites in the MCS region for the construction of pUT18C-*tssX*-*tssY*. The forward primer for amplifying *tssY* contained an additional stop codon to ensure no T18-TssX-TssY fusion protein was formed, and a Shine-Dalgarno sequence for *tssY* translation.

3.4.1 Construction of plasmids for three-hybrid assays

3.4.1.1 Construction of pUT18C-*tssK-tssL*

PCR amplification of *tssL* from *B. cenocepacia* strain H111 was performed with KOD DNA polymerase and a pair of *tssL*-specific primers which have *Kpn*I restriction sites in both forward and reverse primers (pUT18C-K-L.KpnI.for and pUT18C-K-L.KpnI.rev). A PCR product of the expected size (~600 bp) and pUT18C-*tssK* were then digested with *Kpn*I. pUT18C-*tssK* was subsequently treated with CIP in order to decrease the efficiency of self-ligation occurring (Section 2.3.13). Following ligation of the *tssL* PCR product to pUT18C-*tssK*, *E. coli* strain MC1061 was transformed with the ligation mixture and colonies were grown on LB agar plates containing ampicillin. Candidate recombinant plasmids in which the *tssL* DNA fragment was inserted downstream of *tssK* were identified by PCR screening with primers that were used to amplify *tssL* and followed by plasmid miniprep analysis to confirm that the plasmids contained the *tssL* inserts. The nucleotide sequence of the inserted gene was verified by DNA sequencing. The resulting plasmid was called pUT18C-*tssK*-*tssL*.

3.4.1.2 Construction of pUT18C-*tssL-tssK*

PCR amplification of *tssK* from *B. cenocepacia* strain H111 was carried out with a pair of primers that specified a *Kpn*I restriction site in the forward primer and a *Sma*I restriction site in the reverse primer (tssKtrihyfor and tssKtrihyrev). The amplified *tssK* DNA fragment of the expected size (~1.3 kb) was digested with *Kpn*I and *Sma*I, then ligated into *Kpn*I- and *Eco*53I-digested pUT18C-*tssL*. *E. coli* strain MC1061 was transformed with the ligation mixture and colonies were grown on LB agar plates containing ampicillin. Candidate recombinant plasmids containing *tssK* were identified by PCR screening with primers that were used to amplify *tssK* and followed by plasmid miniprep analysis to confirm that the plasmids contained the *tssK* inserts. The nucleotide sequence of the inserted gene was verified by DNA sequencing. The resulting plasmid was called pUT18C-*tssL-tssK*.

3.4.1.3 Construction of pUT18C-*tssF-tssG*

For constructing a plasmid containing both *tssF* and *tssG*, *tssG* was amplified from *B. cenocepacia* strain H111 with a pair of primers specifying a *Kpn*I restriction site in the forward primer and a *Sma*I site in the reverse primer (tssGtrihyfor and tssGtrihyrev). A PCR product of the expected size (~1.1 kb) was then digested with *Kpn*I and *Sma*I and ligated into *Kpn*I- and *Eco*53I-digested pUT18C-*tssF*. *E. coli* strain MC1061 was transformed with the ligation mixture and colonies were grown on LB agar plates containing ampicillin. Candidate recombinant plasmids containing *tssG* were identified by PCR screening with primers that were used to amplify *tssG* and followed by plasmid miniprep analysis to confirm that the plasmids contained the *tssG* inserts. The nucleotide sequence of the inserted gene was verified by DNA sequencing. The resulting plasmid was named pUT18C-*tssF-tssG*.

3.4.2 Three-hybrid analysis of interactions between TssA^S, TssF and TssG

The genes encoding TssF and TssG are found located next to each other in 97% of T6SS gene clusters suggesting that these protiens may interact with each other (Boyer et al. 2009; Coulthurst 2013; Brunet et al. 2015). Accordingly, a recent published report demonstrated that TssF and TssG interact and stabilize each other (English et al. 2014; Brunet et al. 2015). It was observed that the combination of TssA^S and TssF two-hybrid proteins gave rise to a Mal$^+$ phenotype when the C-terminus of TssA^S is free (Section 3.3.6), suggesting an interaction between them. However, when the C-terminus of TssF was also free, the β-galactosidase activity showed a wide variation from assay to assay, possibly suggesting that the interaction between these two proteins was unstable. The possibility that the presence of TssG can enhance or stabilise the interaction between TssA^S and TssF therefore was investigated by conducting a three-hybrid assay. The three-hybrid combination of TssA^S, TssF and TssG was carried out using the combination of pKT25-TssA^S with pUT18C-*tssF-tssG*, in which the C-terminus of TssA^S and TssF were free. The result showed that the combination of T25-TssA^S with T18-TssF and unfused TssG gave rise to patchy red/purple colonies, consistent with a Mal$^+$ phenotype. However, this was not an obviously stronger Mal$^+$ phenotype than that observed for the corresponding two-hybrid combination, i.e. T25-TssA^S with T18-TssF.

Subsequently, a three-hybrid assay was carried out between $TssA^{S}_{CTD}$ and TssF in the presence of TssG. The previous results showed that the combinations of T18-TssF and $TssA^{S}_{CTD}$ fused at either end of T25 yielded a Mal^{+} phenotype with red/purple colonies. However, the maltose phenotype was weaker in the three-hybrid assays when TssG was present. In fact, the combination of T25-$TssA^{S}_{CTD}$, T18-TssF and unfused TssG gave rise to a maltose-negative phenotype.

β-galactosidase assays were performed for all the plasmid combinations that were assayed on MacConkey-maltose agar (Figure 3.18). As reported earlier (Section 3.3.6), the β-galactosidase activity generated by the combination of T25-$TssA^{S}$ and T18-TssF was greater than ten-fold above the background, but it was variable, with a standard deviation that was almost as high as the mean value for the activity calculated from 30 measurements (1012.5 ± 1060.2 Mu). The three-hybrid combination of T25-$TssA^{S}$, T18-TssF and unfused TssG gave a β-galactosidase activity of 1496.2 ± 142.4 Mu. Although the mean β-galactosidase activities of the two-hybrid and three-hybrid combinations were not significantly different, the much lower standard deviation in the three-hybrid combination suggests the interaction between $TssA^{S}$ and TssF may be strengthened or stabilized by the presence of TssG.

Variable β-galactosidase activities were also obtained for the combination of T25-$TssA^{S}_{CTD}$ and T18-TssF. However, the β-galactosidase activity of its analogous three-hybrid combination, T25-$TssA^{S}_{CTD}$, T18-TssF and unfused TssG, was similar to the background level that was measured in the negative controls. Moreover, the β-galactosidase activity was significantly lower in the $TssA^{S}_{CTD}$-T25, T18-TssF and unfused TssG three-hybrid combination (163.7 ± 90.4 Mu) compared to its analogous two-hybrid combination (1519.7 ± 559.4 Mu). Therefore, whereas TssG appear to stabilise the interaction between $TssA^{S}$ and TssF, it either has the opposite effect on the $TssA^{S}_{CTD}$-TssF interaction or it may somehow interfere with the proximity of T18 and T25, possibly by steric hindrance.

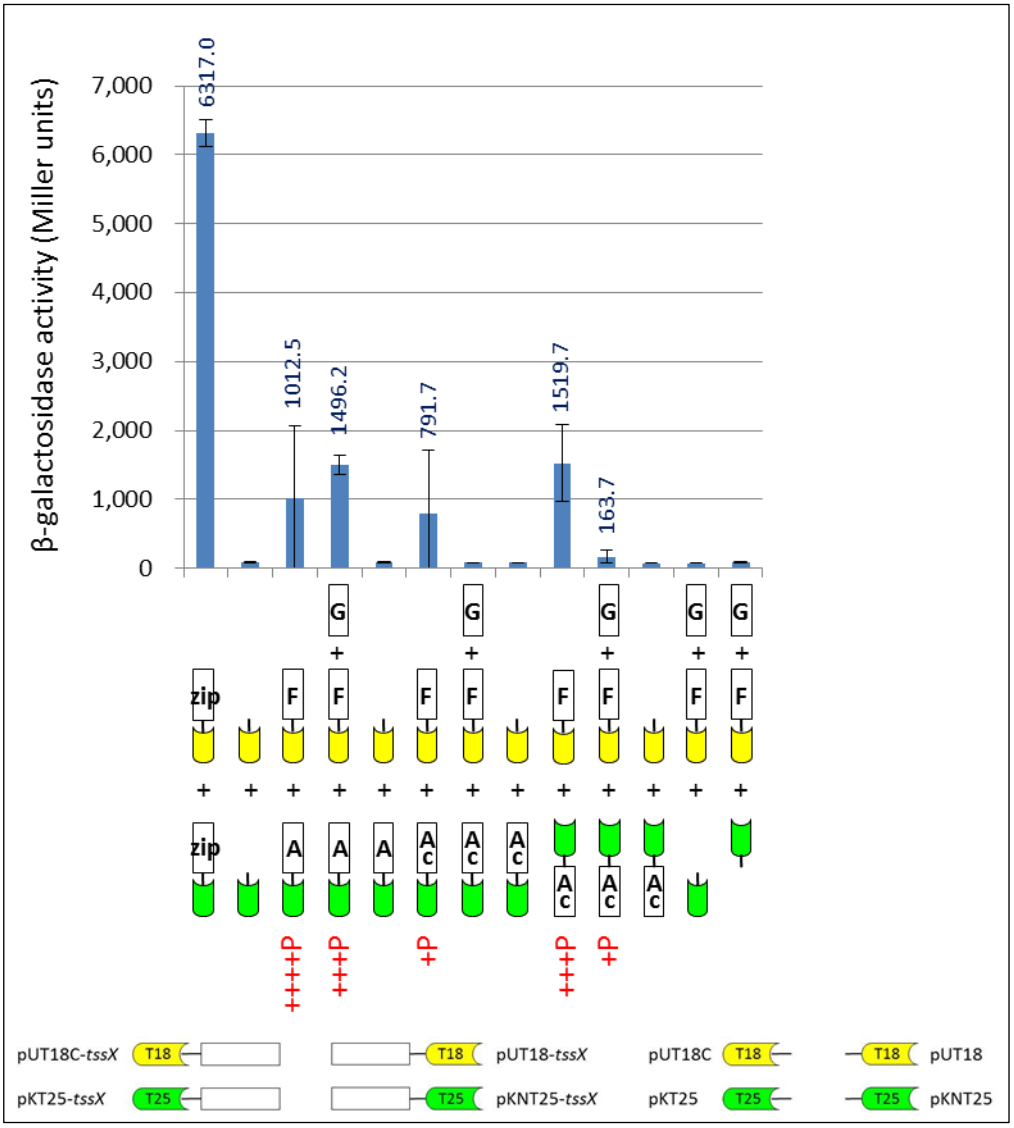


Figure 3.18 Three-hybrid analysis of the interaction between TssA^S and TssF in the presence of TssG. Compatible BACTH plasmids encoding T25 fusions to TssA^S or TssA$^S_{CTD}$ in combination with T18-TssF or T18-TssF and TssG were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2; 'P', patchy. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'F', 'G' and 'Ac' represent TssA^S, TssF, TssG and TssA$^S_{CTD}$, respectively.

3.4.3 Three-hybrid analysis of interactions between $TssA^S$, TssK and TssL

To investigate the effect of TssL on the interaction between $TssA^S$ and TssK, the two-hybrid protein combination of T25-$TssA^S$ with T18-TssK in which the C-terminus of both $TssA^S$ and TssK are free was assayed in the presence unfused TssL. As mentioned in Section 3.3.11, the combination of T25 or T18 fused to the N-termini of $TssA^S$ and TssK, respectively, yielded a Mal^+ phenotype with red/purple colonies. In the presence of TssL, this three-hybrid combination gave rise to dark red/purple colonies, indicating a stronger Mal^+ phenotype compared to that observed from the two-hybrid assay.

β-galactosidase assays were performed for the two combinations that were assayed on MacConkey-maltose agar for the analysis of the interaction between $TssA^S$ and TssK in the presence of TssL (Figure 3.19). The β-galactosidase activity of the three-hybrid combination of T25-TssA, T18-TssK and unfused TssL was 3397.2 ± 307.0 Mu, which was higher than for the same two-hybrid combination in the absence of TssL (2184.1 ± 53.5 Mu). The difference between these two combinations was statistically significant based on t-test ($p < .001$). The results were consistent with the maltose phenotypes and suggested the interaction of $TssA^S$ and TssK was enhanced by the presence of TssL.

To investigate the effect of TssK on the interaction between $TssA^S$ and TssL, the two-hybrid protein combination of T25-$TssA^S$ and T18-TssL in which the C-terminus of both $TssA^S$ and TssL are free was assayed in the presence of unfused TssK on MacConkey-maltose agar. As mentioned in Section 3.3.12, the two-hybrid combination yielded a Mal^+ phenotype with moderate red/purple colonies. However, in the three-hybrid assay, where TssK was also present, a negative maltose phenotype was observed.

The β-galactosidase activity generated by the combination of $TssA^S$ and TssL fusion proteins was 1302.0 ± 50.0 Mu (T25-$TssA^S$ and T18-TssL). However, the three-hybrid combination of T25-$TssA^S$ with T18-TssL and unfused TssK gave a β-galactosidase activity that was similar to the background level (95.7 ± 6.3 Mu). The results of β-galactosidase assay are consistent with the maltose phenotype observations, where the three-hybrid combination of T25-$TssA^S$ with T18-TssL and unfused TssK led to a negative maltose phenotype (Figure 3.19).

To confirm that pUT18C-*tssL-tssK* is actually able to produce T18-TssL, an analogous three-hybrid assay was performed with $TssM_{NTD}$, which interacts with TssL, replacing $TssA^S$ (Ahmad, 2013, Cascales use earlier reference). The results clearly show that pUT18C-*tssL-tssK* does produce TssL as it gives rise to a similar Mal^+ phenotype as pUT18C-*tssL* when in combination with either pKT25-$tssM_{NTD}$ or pKNT25-$tssM_{NTD}$ (Figure 3.19). β-galactosidase activity measurements also demonstrated that pUT18C-*tssL-tssK* was effective at providing T18-TssL as there was substantial complementation of CyaA function when combined with T25-$TssM_{NTD}$ or $TssM_{NTD}$-T25 (Figure 3.19). However, in both cases there was a statistically significant decrease in the β-galactosidase activities relative to the corresponding two-hybrid combinations that did not specify unfused TssK (p=.005 for pKT25-$tssM_{NTD}$ and pUT18C-*tssL-tssK*; p< .001 for pKNT25-$tssM_{NTD}$ and pUT18C-*tssL-tssK*). The decrease in CyaA complementation in the presence of TssK may be due to steric interference or a change in conformation of TssK and $TssM_{NTD}$ upon binding of TssK to the complex.

As TssK and TssL do not interact with the separate domains of $TssA^S$, a three-hybrid analysis involving $TssA^S$ domains was not performed.

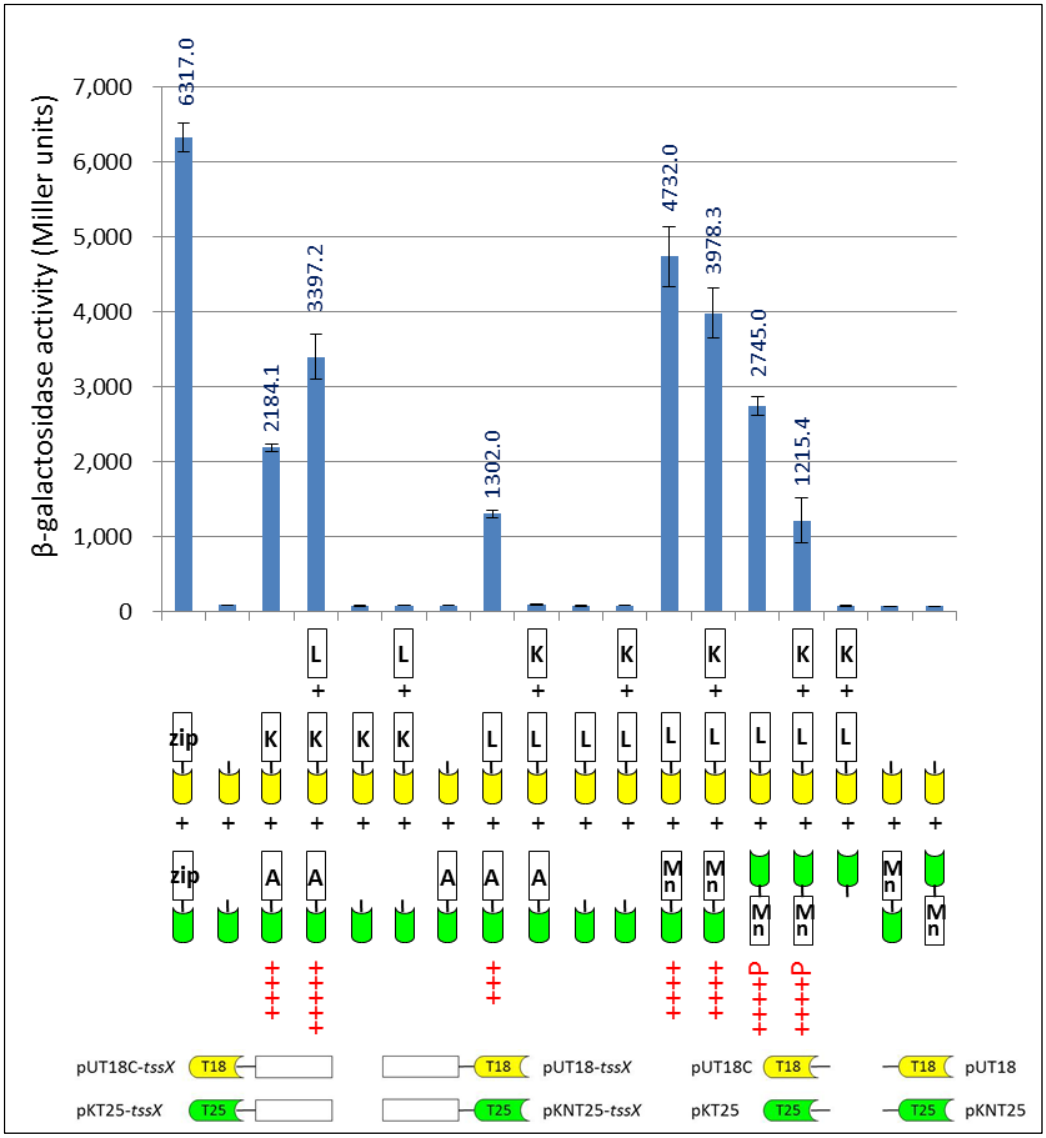


Figure 3.19 Three-hybrid analysis of the interaction between TssA^S, TssK and TssL. Compatible BACTH plasmids encoding T25-TssA^S in combination with T18-TssK, T18-TssK and TssL, T18-TssL, and T18-TssL and TssK were analysed using the BACTH system. β-galactosidase assay was employed to quantify the efficiency of functional complementation between the hybrid proteins. X-axis shows a diagrammatic representation of the combinations that were analysed. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated below the corresponding diagram as red crosses according to the classification described in Table 3.2; 'P', patchy. Negative maltose phenotypes were not indicated. Y-axis shows the β-galactosidase activity in Miller units and the bar represents the mean ± standard deviation obtained for at least three independent colonies. As a positive control, cells containing pKT25-*zip* and pUT18C-*zip* were used. As negative controls, cells containing empty vectors pKT25 and pUT18C, and cells containing one empty vector in combination with a compatible plasmid expressing a fusion protein were assayed. The background level of the β-galactosidase activity measured in the negative controls was 75-85 Mu. 'A', 'K', 'L' and 'Mn' represent TssA^S, TssK, TssL and TssM$_{NTD}$, respectively.

3.5 Analysis of interactions between $TssA^{EI}$ and other T6SS subunits using the BACTH system

To provide evidence that $TssA^{EI}$ is a functional orthologue of $TssA^{S}$, the BACTH assay was carried out on MacConkey-maltose agar for probing the interactions between $TssA^{EI}$ and other T6SS subunits. In this analysis, AHA1844 from *A. hydrophila* strain ATCC 7966 was employed as $TssA^{EI}$. DNA encoding $TssA^{EI}$ was cloned into the four BACTH plasmids generating T25 or T18 fused to $TssA^{EI}$ at either its N- or C-terminus (Ahmad 2013). The two-hybrid analysis of $TssA^{EI}$ self-interaction and its interactions with TssD, TssE, TssH, TssI, TssK, TssL and TssM have been carried out previously (Ahmad, 2013). The interactions between $TssA^{EI}$ with the remaining T6SS subunits, TssB, TssC, TssF, TssG and TssJ were investigated in this study with all eight possible pairwise BACTH plasmid combinations encoding $TssA^{EI}$ and the other core T6SS components fusions to CyaA domains. In addition, the interaction with $TssM_{NTD}$ was repeated with corrected plasmids as the previous studies were conducted with the pair of defective $TssM_{NTD}$ plasmids (Section 3.2.3).

The combination of TssB-T25 with T18-$TssA^{EI}$ gave rise to a Mal^{+} phenotype with weak red/purple colonies (Figure 3.20 A). One combination of $TssA^{EI}$ and TssC that gave the N-terminus of TssC and C-terminus of $TssA^{EI}$ free yielded a patchy Mal^{+} phenotype with weak red/purple colonies (Figure 3.20 B). Four combinations of $TssA^{EI}$ and TssF led to a Mal^{+} phenotype with patchy patterned colonies (Figure 3.20 C). All eight compatible BACTH combinations encoding $TssA^{EI}$ and TssG fusions to CyaA domains yielded a maltose-negative phenotype. Unlike $TssA^{S}$, where the combination of TssJ-T25 with T18-$TssA^{S}$ gave rise to a weak maltose-positive phenotype, all of these eight combinations yielded a maltose-negative phenotype (results not shown). All eight compatible BACTH plasmid combinations encoding $TssA^{EI}$ and $TssM_{NTD}$ fusions to CyaA domains yielded a maltose-negative phenotype (results not shown). The result is not in agreement with that of $TssA^{S}$, where the combination of $TssM_{NTD}$-T25 and T18-$TssA^{S}$ gave rise to a moderately strong Mal^{+} phenotype.

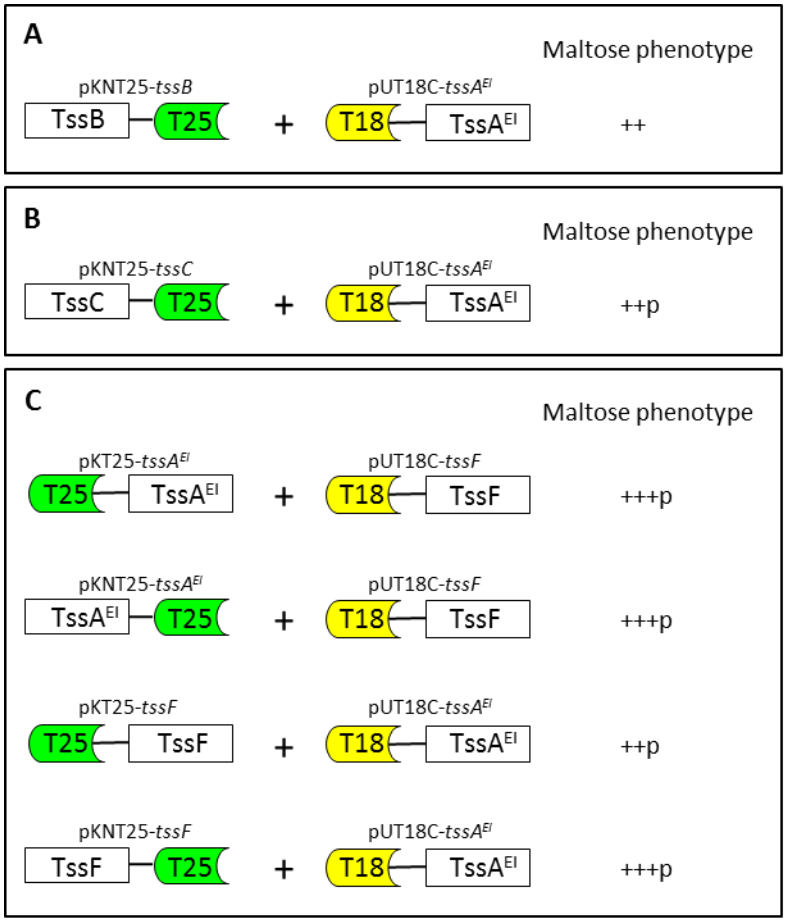


Figure 3.20 BACTH analysis of the interaction between TssA^{EI} and TssB, TssC and TssF on MacConkey-maltose agar. Compatible BACTH plasmids encoding the CyaA T25 or T18 fragments fused to TssA^{EI} and TssB/TssC/TssF were analysed using the BACTH system. Maltose phenotype was observed after 120 hours incubation. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated next to the corresponding diagram as crosses according to the classification described in Table 3.2; 'P', patchy; combinations that yielded a negative maltose phenotype were not shown. **A**. Combination of TssA^{EI} and TssB fusion proteins that gave rise to a maltose-positive phenotype. **B**. Combination of TssA^{EI} and TssC fusion proteins that gave rise to a maltose-positive phenotype. **C**. Combinations of TssA^{EI} and TssF fusion proteins that gave rise to a maltose-positive phenotype.

3.6 Discussion

3.6.1 Interactions between TssA^S and other T6SS subunits using two-hybrid and three-hybrid assays

So far, no studies on TssA^S have been published. The interactions between TssA^S and other T6SS subunits of *B. cenocepacia* were screened by the BACTH assay. Subsequent analysis was carried out to identify the domain(s) of TssA^S responsible for the interaction in situations where the full-length TssA^S showed an interaction with other T6SS subunits. Moreover, TssA^S was shown to form a ring-like structure (Ahmad 2013; Chapter 5) and was proposed to be a baseplate component (Ahmad 2013). Based on interaction analysis in this study, it is proposed that TssA^S locates within or close to the baseplate.

As the maltose phenotype that is observed on MacConkey-maltose agar lacks sensitivity and is not quantitative, β-galactosidase assays were carried out for obtaining quantitative data. For those BACTH combinations that gave a weak Mal$^+$ phenotype, a statistical analysis was also conducted for supporting data analysis. The results of the comprehensive interaction map of TssA^S in this chapter support the idea that TssA^S is located near the baseplate which contributes to our understanding of how the T6SS assembles. Figures 3.21 and 3.22 present a summary of the BACTH assay results for TssA^S and its domains with TssA^S-TssM.

It should be pointed out that in the TssA^S domain interaction analysis, if a T6SS subunit interacts with TssA^S through contacting the TssA$^S_{NTR}$, it may not be observed in a two-hybrid assay involving TssA$^S_{NTR}$ due to the structure of the oligomer that is formed by TssA^S. By using transmission electron microscopy (TEM), we have observed that TssA$^S_{CTD}$ forms a ring structure surrounded by a number of discrete protein projections formed by the NTR (Chapter 5 and Ahmad 2013). Although TssA$^S_{NTR}$ has been suggested to weakly self-associate by the β-galactosidase assay (Section 3.3.1), the weak self-association may be only stabilized in the full-length TssA^S following ring assembly. We hypothesise that the ring formation of TssA^S increases the local concentration of TssA$^S_{NTR}$ and this leads to its self-association. According to this hypothesis, without the TssA$^S_{CTD}$, the interaction of free TssA$^S_{NTR}$ with itself is so weak that it is unlikely to occur. Therefore, if the T6SS protein interacts with TssA$^S_{NTR}$

oligomers, the interaction is most likely only going to be observed for the full-length $TssA^S$.

3.6.1.1 Self-association of $TssA^S$

$TssA^S$ self-interaction was initially reported in our group by the BACTH system (Shastri 2011; Ahmad 2013). In this study, the results showed that the combination of T25-$TssA^S$ and T18-$TssA^S$ complemented CyaA activity nearly as efficiently as the Zip control. This supports the previous observations that suggest $TssA^S$ self-interacts (Shastri 2011; Ahmad 2013).

The interactions between $TssA^S$ and its domains, i.e. $TssA^S_{NTR}$ and $TssA^S_{CTD}$, were also assayed to identify the domain that is involved in the $TssA^S$ self-interaction. Three out of eight combinations between $TssA^S$ and $TssA^S_{NTR}$ yielded a Mal^+ phenotype, and in each case the C-terminus of $TssA^S$ was free. In addition, six out of eight combinations of $TssA^S$ and $TssA^S_{CTD}$ hybrid proteins gave rise to a very strong Mal^+ phenotype with β-galactosidase activities that were almost as high as that of $TssA^S$ self-interaction combinations. The remaining two combinations: $TssA^S$-T25 with $TssA^S_{CTD}$-T18, and $TssA^S_{CTD}$-T25 with $TssA^S$-T18, gave rise to colonies that exhibited variable Mal phenotypes.

The self-association of $TssA^S_{NTR}$ and $TssA^S_{CTD}$ were also investigated. Although $TssA^S_{NTR}$ can interact with $TssA^S$, no evidence for self-association of $TssA^S_{NTR}$ was obtained by observing the Mal phenotype in the BACTH system. The only combination in which $TssA^S_{NTR}$ was located at the N-terminus of both CyaA fusion proteins gave a β-galactosidase activity slightly higher than that of the negative controls, which was statistically significant (p=.012). This suggests $TssA^S_{NTR}$ can weakly self-associate *in vivo* that is agreement with our latter observations from electron microscopy (Chapter 5). In contrast, plasmids encoding $TssA^S_{CTD}$ showed very strong Mal^+ phenotypes in all four pairwise combinations, with the β-galactosidase activity nearly as high as that of the $TssA^S$ self-interaction combination. These results support the hypothesis that the C-terminal domain is responsible for the $TssA^S$ self-interaction (Shastri 2011; Ahmad 2013) and also provided quantitative data for validating observations on MacConkey-maltose agar.

3.6.1.2 Interaction of TssA^S with TssB and TssC

The *tssB* and *tssC* genes are always found together and in the same relative orientation in the T6SS gene clusters suggesting a functional relationship, and the interaction of TssB and TssC proteins has been shown in many bacterial species (Broms et al. 2009a; Aubert et al. 2010; Zhang et al. 2013). The *V. cholerae* TssB and TssC proteins (VipA/VipB) are found to form tubules that are proposed to act as the T6SS tail sheath (Mougous et al. 2006; Ballister et al. 2008; Bönemann et al. 2009; Lossi et al. 2013; Zhang et al. 2013). The tubular structure shares a similar organisation to the bacteriophage tail sheath although the individual subunits do not have amino acid sequence similarities with gp18, the phage T4 tail sheath subunit. Previous observations in our group have shown that *B. cenocepacia* TssA^S interacts with TssB, but does not interact with TssC by the BACTH assay (Shastri 2011). Plasmids expressing CyaA T25- and T18-TssC fusions were re-constructed in this study and the interactions between TssA^S and TssB, and TssC were reinvestigated.

The only combination of TssA^S with TssB that yielded a Mal$^+$ phenotype was TssB-T25 with T18-TssA^S, and the β-galactosidase activity of which was about three-fold higher than the background, suggesting an interaction between TssA^S and TssB. However, using the same TssB-T25 fusion, it was not possible to identify the domain of TssA^S that is involved in the interaction with TssB; all the combinations of TssB-T25, with TssA$^S_{NTR}$ or with TssA$^S_{CTD}$ yielded a maltose-negative phenotype and the β-galactosidase activities were similar to the background level. However, this is not consistent with the previous observation where the TssB-T25 and T18-TssA$^S_{CTD}$ combination was recorded as giving rise to a very strong Mal$^+$ phenotype (Shastri 2011).

A new finding of the interaction between TssA^S and TssC was reported in this study by carrying out the BACTH assay with rectified pKNT25-*tssC*, pUT18-*tssC* and pUT18C-*tssC* plasmids. A combination of CyaA fusion proteins in which the N-terminus of TssC and C-terminus of TssA^S were free yielded a strong Mal$^+$ phenotype and a β-galactosidase activity that was ~25-fold higher than the background. The previous observations with defective plasmids showed a negative maltose phenotype for all the combinations between TssA^S and TssC (Shastri 2011). However, no evidence for the interaction between TssA^S domains and TssC was obtained, as a

negative maltose phenotype was observed for all tested combinations and the β-galactosidase activity was similar to the background level.

As TssB and TssC tubules polymerise in the cytoplasm around the growing TssD tail tube during the T6SS assembly, it is reasonable to assume that the interactions of $TssA^S$ with TssB-C would likely occur in the cytoplasm if $TssA^S$ is proposed to locate within or close to the baseplate of T6SS. The interaction of TssB and TssC with $TssA^S$ may require contacts with both the N- and C-terminal domains of $TssA^S$ or they may require oligomerisation of $TssA^S_{NTR}$ as discussed earlier.

3.6.1.3 Interaction of $TssA^S$ with TssD

$TssA^S$ was suggested to interact with TssD based on the BACTH assay, where both possible combinations of fusion protein in which the C-terminus of both proteins are free yielded a Mal^+ phenotype. The result is consistent with previous findings (Shastri 2011; Ahmad 2013). The interaction between $TssA^S$ and TssD is suggested to occur through the C-terminal domain of $TssA^S$ as two out of four tested combinations of $TssA^S_{CTD}$ with TssD yielded a Mal^+ phenotype in the BACTH assay, although the β-galactosidase activities were less than two-fold higher than the background for one combination and six-fold for the other. No evidence of an interaction between $TssA^S_{NTR}$ and TssD was observed.

TssD forms a nanotube by stacked hexameric rings that initially assemble at the IM and can extend through the OM (Mougous et al. 2006; Ballister et al. 2008). It shares structural similarity to the lambda phage protein gpV and a common evolutionary origin with the bacteriophage T4 protein gp19, both of which assemble into the bacteriophage tail tube (Leiman et al. 2009; Pell et al. 2009a). The bacteriophage tail tube is driven through the baseplate during infection by contraction of tail sheath, which creates a channel for DNA injection from the capsid into the target bacterial cytosol (Leiman et al. 2004). Therefore, if $TssA^S$ locates within or close to the baseplate, it is possible for an interaction between the T6SS tail tube protein, TssD, and the ring-forming component of $TssA^S$, $TssA^S_{CTD}$, to occur.

3.6.1.4 Interaction of TssA^S with TssE

TssE has been shown to have a significant amino acid sequence similarity with the bacteriophage T4 baseplate component gp25 (Shalom et al. 2007; Leiman et al. 2009; Lossi et al. 2011). It is also predicted to behave in the same way as gp25 in the assembly of the T6SS baseplate by interacting with TssI (tail spike) in the resting state and TssD (tail tube) in the active state (Kostyuchenko et al. 2003; Kostyuchenko et al. 2005; Leiman et al. 2009; Leiman et al. 2010; Lossi et al. 2011).

One of the BACTH fusion protein combinations involving TssA^S and TssE (TssE-T25 and T18-TssA^S) gave rise to a patchy Mal$^+$ phenotype with a β-galactosidase activity that was approximately 4-fold higher than the background, suggesting an interaction between these two proteins. The patchy colony phenotype is in agreement with the previous observations (Shastri 2011; Ahmad 2013), but the β-galactosidase activity measurement provided more convincing evidence for an interaction between TssA^S and TssE. As TssE is proposed to be a T6SS baseplate component (Brunet and Cascales 2015), the interaction between TssA^S and TssE provided further evidence of the predicted the location of TssA^S. However, no evidence for an interaction between either of the TssA^S domains and TssE was observed. This agrees with the previous observations of Ahmad (2013) but is in contrast with the observation of Shastri (2011) who recorded that both compatible combinations of CyaA fusion proteins in which TssA$^S_{CTD}$ and TssE were located at the N-terminus yielded a strong Mal$^+$ phenotype.

3.6.1.5 Interaction of TssA^S with TssF and TssG

TssF and TssG are proposed to be T6SS baseplate components as they are homologous to proteins J and I, respectively, which are two baseplate components of phage P2 (Brunet et al. 2015). TssF and TssG have been reported to interact and stabilize each other, and are required for proper assembly of the T6SS tail tube that is formed by TssD (Brunet et al. 2015).

Three out of eight pairwise combinations of TssA^S and TssF fusion proteins that were tested in the BACTH system yielded a Mal$^+$ phenotype. The colony phenotype generated by the two fusion combinations where the C-terminus of both proteins is free was patchy patterned and the β-galactosidase activity measurements were highly

variable. The three-hybrid combination of T25-TssA^S with T18-TssF and TssG which investigated the interaction between TssA^S and TssF in the presence of TssG also yielded a Mal$^+$ phenotype but with more uniform coloured colonies and a more stable β-galactosidase activity. Although the mean β-galactosidase activities generated by these two-hybrid and three-hybrid combinations were not much different, the much lower standard deviation in the three-hybrid combination suggests the interaction between TssA^S and TssF may be strengthened or stabilized by the presence of TssG. In contrast, the two-hybrid combination in which the N-terminus of TssF is free gave uniform coloured colonies when it was combined with T18-TssA^S; it also showed less variable β-galactosidase activity compared to the other two-hybrid combinations of TssA^S with TssF. One hypothesis to account for the variability in the β-galactosidase activity is that it may be caused by unstable or transient interaction between these two proteins.

In conflict with previous observations in which the TssA^S-T25 and T18-TssF combination was observed to give rise to patchy maltose-positive colonies (Shastri 2011; Ahmad 2013), this combination led to a negative maltose phenotype and the β-galactosidase activity was similar to the background level. The cause of this discrepancy is unclear. One negative control combination, pKNT25 with pUT18C-*tssF*, i.e. T25 with T18-TssF, gave rise to a Mal$^+$ phenotype. This may be explained if TssF in the T18-TssF chimaeric protein was mis-folded and exhibited an affinity for the T25 component of CyaA when it was not fused, as all tested hybrid protein combinations of T18-TssF with C-terminus T25 fused TssA^S or its domains led to a negative maltose phenotype except with TssA$^S_{CTD}$-T25. Although the control problem of pKNT25 with pUT18C-*tssF* leading the Mal$^+$ phenotype observed for the combination of TssA$^S_{CTD}$-T25 with T18-TssF less reliable, another combination between TssA$^S_{CTD}$ and TssF hybrid proteins also yield Mal$^+$ phenotype, suggesting that the interaction between TssA^S and TssF is likely to involve TssA$^S_{CTD}$. However, the β-galactosidase activity was variable for these two combinations.

In the three-hybrid assay, the presence of TssG caused the β-galactosidase activity of the T18-TssF and TssA$^S_{CTD}$-T25, and T18-TssF and T25-TssA$^S_{CTD}$ combinations to decrease considerably and gave rise to a very weak or negative maltose phenotype. The inability of T18-TssF to complement CyaA function in combination with T25-TssA$^S_{CTD}$

or $TssA^{S}_{CTD}$-T25 when TssG is also present may be caused by steric hindrance or a large conformational change induced in TssF by the binding of TssG rather than decreased T18-TssF production. One combination between $TssA^{S}_{NTR}$ and TssF yielded a Mal^{+} phenotype, however, the β-galactosidase activity was less than 3-fold above the background. No evidence of an interaction between $TssA^{S}$ and TssG was observed, which is consistent with previous observations (Shastri 2011).

Therefore, the results suggest $TssA^{S}$ interacts with TssF, a baseplate component, and supports the proposed location of $TssA^{S}$ within or close to the T6SS baseplate.

3.6.1.6 Interaction of $TssA^{S}$ with TssH

TssH is an AAA+ ATPase. It forms a hexamer, binding ATP by the conserved AAA domains and converting the energy of ATP hydrolysis into the unfolding of substrates (proteins) (Schlieker et al. 2005; Mougous et al. 2006; Zheng and Leung 2007; Bönemann et al. 2009). The TssH chaperone has been shown to play a crucial role in T6SS assembly. It interacts with the two T6SS tail sheath proteins, TssB and TssC (Bönemann et al. 2009; Kapitein et al. 2013). TssH provides energy for recycling TssB-TssC tubules in their dynamic cycles of assembly, contraction and disassembly, and also prevents the formation of non-productive TssB-TssC tubules (Basler et al. 2012; Kapitein et al. 2013).

The two-hybrid fusion protein combinations in which the C-terminus of both $TssA^{S}$ and TssH are free yielded a very strong Mal^{+} phenotype with a β-galactosidase activity ~35 fold higher than the background level. Furthermore, the results revealed that $TssA^{S}_{CTD}$ is responsible for the interaction with TssH, in agreement with previous observations (Ahmad 2013). In support with the hypothesis of the location of $TssA^{S}$ within or close to the baseplate, the interaction between $TssA^{S}$ and TssH would likely to occur in the cytoplasm.

3.6.1.7 Interaction of $TssA^{S}$ with TssI

The TssI protein forms a trimer that is analogous to the $(gp27)_3$-$(gp5)_3$ tail spike complex of the T4 bacteriophage, which functions as a puncturing device towards the target cell, permitting the delivery of the viral DNA (Kanamaru et al. 2002; Pukatzki et

al. 2007; Leiman et al. 2009). Current understanding of the injection apparatus suggests that TssI is recruited along with other baseplate-like assembly platform components (TssE, TssF, TssG and TssK) by the TssJLM trans-membrane complex in the initial assembly step of the T6SS apparatus (Brunet et al. 2015). Polymerisation of the tail tube (TssD) and tail sheath (TssBC) is initiated afterwards, and TssI locates on the top of TssD tube which is employed by the bacteria to puncture the targeted host (Pukatzki et al. 2007; Hachani et al. 2011; Brunet et al. 2015).

$TssA^S$ interacts with TssI, as the two BACTH fusion combinations in which the C-terminus of both $TssA^S$ and TssI are free gave rise to very efficient complementation of CyaA activity. The conserved gp27- and gp5-like domains of TssI and the fused gp27gp5-like region were further analysed in combination with $TssA^S$. Both combinations of $TssA^S$ with the gp27gp5-like fusion yielded a very strong Mal^+ phenotype with a β-galactosidase activity that was about 31-fold higher than the background for one combination and nearly 16-fold higher for the other. However, when the components of the gp27gp5-like fusion protein were expressed separately as gp27- and gp5-like proteins, they yielded a maltose-negative phenotype in combination with $TssA^S$ and the β-galactosidase activity was similar to the background level. The results indicated the entire gp27gp5 like region of TssI was required for efficient interaction with $TssA^S$.

The combination of $TssA^S_{CTD}$ and TssI also gave a strong Mal^+ phenotype (~25-fold higher than the background levels), as did the combination of $TssA^S_{CTD}$ and the gp27gp5-like fusion protein (~30-fold above background levels of β-galactosidase activity). No interaction between $TssA^S_{CTD}$ and the gp27- and gp5-like moieties of TssI were observed. Thus, the results support the idea that TssI interacts with $TssA^S$, and this interaction occurs through the core TssI region and $TssA^S_{CTD}$. As TssI forms the tail spike of the T6SS that locates in the baseplate-like assembly platform during T6SS apparatus assembly, and it is possible for it to interact with the component locate within or close to the baseplate, i.e. $TssA^S$.

3.6.1.8 Interaction of $TssA^S$ with TssK

TssK is a trimeric protein located in the cytoplasm as shown in this study and elsewhere (Zoued et al. 2013), and has been further proposed to be a baseplate component (Brunet

et al. 2015). TssK is required for proper assembly of the T6SS tail tube, and is proposed to directly interact with TssL and TssM, along with the interaction of TssG with TssM, resulting in a network interaction connecting the tail of the contractile bacteriophage-like complex (i.e. TssB/TssC, TssD, TssE, TssF, TssG and TssI) with the trans-membrane complex (i.e. TssJ, TssL and TssM) (Zoued et al. 2013; Brunet et al. 2015).

$TssA^S$ was found to interact with TssK by the BACTH assay, consistent with previous observations (Ahmad 2013). The CyaA fusion protein combinations in which the C-terminus of both $TssA^S$ and TssK are free yielded a Mal^+ phenotype with β-galactosidase activities 27-31 fold higher than the background level. In the presence of TssL, the three-hybrid combination of T25-$TssA^S$ with T18-TssK and TssL led to a stronger Mal^+ colony phenotype compared to that of the corresponding two-hybrid assay and with a β-galactosidase activity that was ~43-fold higher than the background, which is statistically significantly higher than that in the absence of TssL. These results suggest that the interaction between $TssA^S$ and TssK is enhanced by the presence of TssL. However, no responsible domains of $TssA^S$ were found for the interaction with TssK. This is not consistent with a previous observation, where the combination of T25-$TssA^S_{CTD}$ with T18-TssK was recorded as giving rise to a patchy Mal^+ colony phenotype (Ahmad 2013). This could have been a false-positive result in the previous study. This specific combination was repeated and quantified by β-galactosidase assay, in which the activity was similar to the background level and consistent with the negative maltose phenotype observed from the BACTH assay. In agreement with the predicted location of $TssA^S$, the interaction between $TssA^S$ and TssK is expected.

3.6.1.9 Interaction of $TssA^S$ with TssJ, TssL and TssM

TssM and TssL locate at the IM by means of TMDs, and along with the OM lipoprotein TssJ (Aschtgen et al. 2008; Ma et al. 2009b; Felisberto-Rodrigues et al. 2011), they form a cell envelope chamber that anchors the T6SS in the membrane, which has overall five-fold symmetry and contains a large component in the cytoplasm. The complex extends into the periplasm by ten arches to form a dual-ring structure comprised of the carboxy-terminal domain of TssM and TssJ that is anchored in the OM by an acyl group (Durand et al. 2015).

There was no strong evidence provided for an interaction between $TssA^S$ and TssJ. In one combination of hybrid proteins, the BACTH assay yielded a very weak Mal^+ phenotype and the β-galactosidase activity was less than 3-fold higher than the background level. Although the β-galactosidase activity is statistically significantly higher compared to the background level (p=.001), given the very low β-galactosidase activity of the only combination of $TssA^S$ and TssJ fusions that yielded a weak Mal^+ phenotype, a firm conclusion concerning an interaction between $TssA^S$ and TssJ cannot be made.

$TssA^S$ and TssL are suggested to interact with each other; the two fusion combinations in which the C-terminus of both $TssA^S$ and TssL are free yielded a Mal^+ phenotype with β-galactosidase activities that were ~9-fold (T25-TssL with T18-$TssA^S$) and ~16-fold (T25-$TssA^S$ with T18-TssL) higher than the background, respectively. However, with TssK present in a three-hybrid assay, the maltose phenotype of the T25-$TssA^S$ with T18-TssL combination was negative and the β-galactosidase activity was decreased to background levels. This was not due to a lack of expression of the T18-TssL hybrid protein as pUT18C-*tssL-tssK* was effective at providing T18-TssL as there was substantial complementation of CyaA function when combined with T25-$TssM_{NTD}$ or $TssM_{NTD}$-T25. The decrease in CyaA complementation in the presence of TssK may be due to steric interference or a change in conformation of TssK and $TssM_{NTD}$ upon binding of TssK to the complex.

$TssA^S$ showed interaction with $TssM_{NTD}$ which is the cytoplasmic domain located between the second and third transmembrane helices (residues 58-445) but not with the periplasmic domain ($TssM_{CTD}$). This is a new finding that was observed after reconstruction of pKNT25-$tssM_{NTD}$ and pUT18-$tssM_{NTD}$ plasmids. Moreover, the results of the BACTH assay suggest both domains of $TssA^S$ are required for the interaction with TssL and $TssM_{NTD}$. The interaction of $TssA^S$ with the cytoplasmic components of the T6SS cell envelope chamber suggests $TssA^S$ locates close to the inner membrane, in agreement its predicted location, i.e. within or close to the baseplate.

Overall, the interaction of $TssA^S$ with the cytoplasmic components of the T6SS cell envelope chamber (TssL and $TssM_{NTD}$), along with the evidence of interacting with baseplate components (TssE, TssF, TssK and TssI) and components of the contractile

complex (i.e. TssB/TssC, TssD), supports the hypothesis that $TssA^S$ is likely to locate within or close to the baseplate.

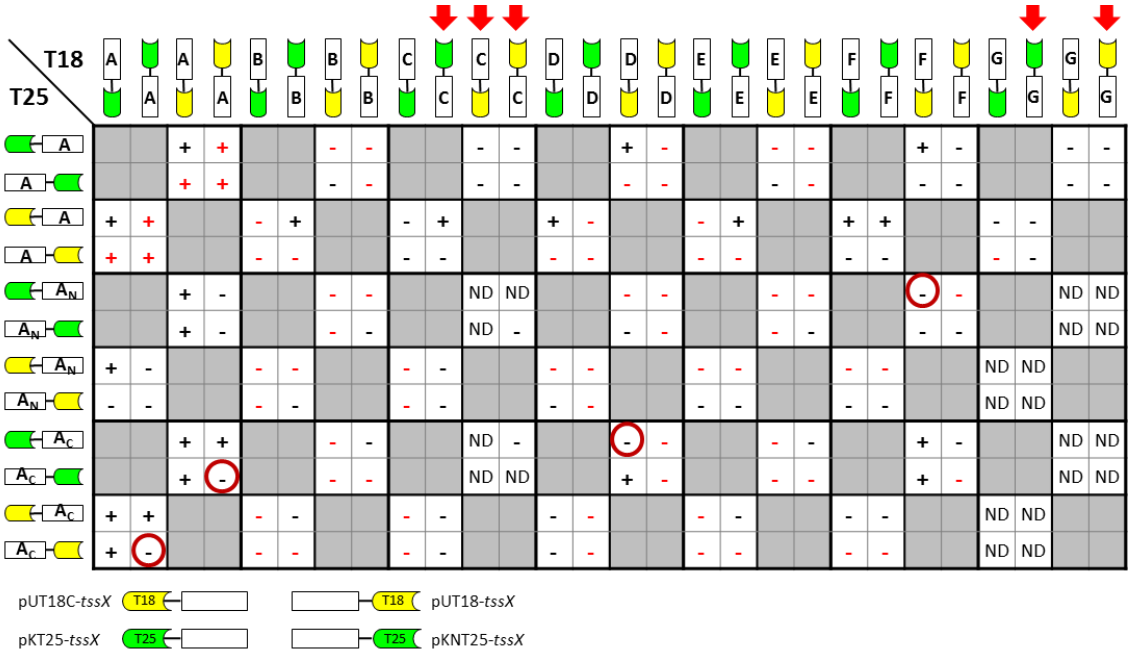


Figure 3.21 Summary of BACTH assay results for TssA^S and its domains with TssA^S-TssG. Diagrammatic representation of two-hybrid fusions are described underneath the grid. The CyaA-T6SS subunit fusions that were reconstructed in this study are indicated with red arrows. Fusion protein combinations shaded in grey are not possible due to plasmid incompatibility and antibiotic resistance markers. Results refer to the maltose phenotypes of each pairwise combination of hybrid protein tested. Results shown in red font are from previous studies (Shastri 2011; Ahmad 2013). '+', maltose-positive phenotype (Mal$^+$); '-', maltose-negative phenotype (Mal$^-$), 'ND', not done. Single letter in the rectangle represents corresponding TssX subunit; 'An' and 'Ac' represent TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively. Results of the fusion protein combinations that yield Mal$^+$ phenotype on maltose MacConkey agar but give β-galactosidase activity less than 3-fold above the background level are indicated with a dark red circle. (TssA^S domain-domain interaction results are not included).

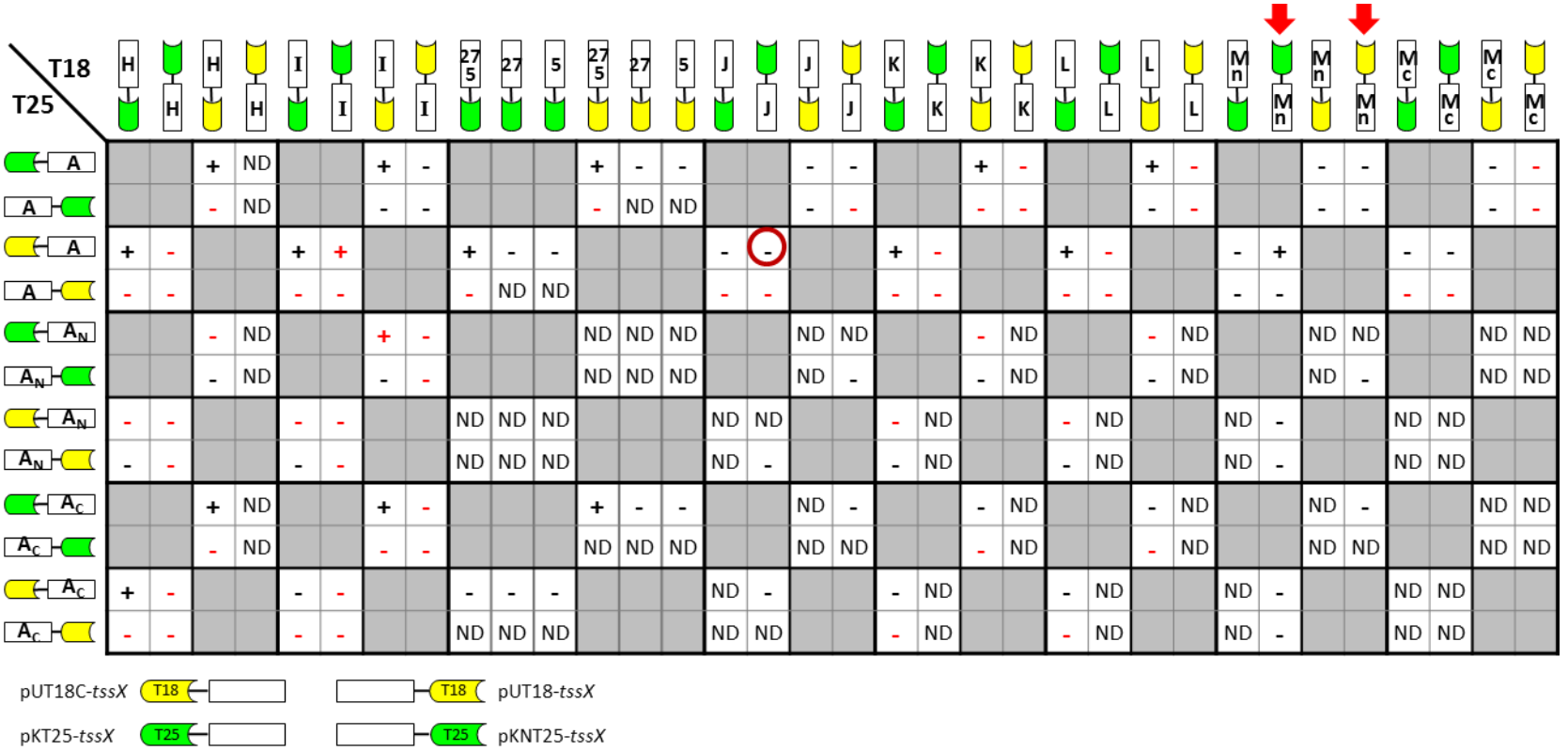


Figure 3.22 Summary of BACTH assay results for TssA^S and its domains with TssH-TssM. Diagrammatic representation of two-hybrid fusions are described underneath the grid. The CyaA-T6SS subunit fusions that were reconstructed in this study are indicated with red arrows. Fusion protein combinations shaded in grey are not possible due to plasmid incompatibility and antibiotic resistance markers. Results refer to the maltose phenotypes of each pairwise combination of hybrid protein tested. Results shown in red font are from previous studies (Shastri 2011; Ahmad 2013). '+', maltose-positive phenotype (Mal$^+$); '-', maltose-negative phenotype (Mal$^-$), 'ND', not done. Single letter in the rectangle represents corresponding TssX subunit; 'An' and 'Ac' represent TssA$^S_{NTR}$ and TssA$^S_{CTD}$, respectively. '27-5', '27' and '5' stand for the fused gp27gp5-, gp27- and gp5-like moieties of TssI, respectively. Results of the fusion protein combinations that yield Mal$^+$ phenotype on maltose MacConkey agar but give β-galactosidase activity less than 3-fold above the background level are indicated with a dark red circle.

3.6.2 Interactions between TssA^{EI} and other T6SS subunits

TssA^{S} was found to self-interact and also to interact with the majority of other T6SS subunits except TssG and possibly TssJ by the BACTH assays. TssA^{EI} was hypothesized to be a functional orthologue of TssA^{S} although the C-terminal regions of these two TssA classes do not share amino acid sequence similarity (Ahmad 2013). The self-interaction of TssA^{EI} from *A. hydrophila* has been shown previously in our group and it also interacts with TssD, -H, -I, -K and -L from *B. cenocepacia* (Ahmad 2013). However, no evidence for an interaction between TssA^{EI} and TssE or TssM$_{NTD}$ was observed (Ahmad 2013). This is not consistent with the observations from another study, in which TssA^{E} from enteroaggregative *E. coli* (EAEC) was shown to interact with TssE, but not with TssL from the same organism (Zoued et al. 2016).

In this study, the interactions between TssA^{EI} and the remaining T6SS subunits (TssB, -C, -F, -G and -J) of *B. cenocepacia* were investigated using the BACTH system, and the interaction with TssM$_{NTD}$ was re-evaluated using the reconstructed TssM$_{NTD}$ plasmids. The results showed that TssA^{EI} also interacts with TssB, TssC and TssF, but not with TssG, TssJ and TssM$_{NTD}$. The results are largely in agreement with Zoued and Cascales (2016), except for interactions with TssB and TssF, where no interactions of TssA^{EI} with TssB and TssF were observed.

In conclusion, TssA^{EI} from *A. hydrophila* interacts with the same T6SS subunits that TssA^{S} interacts with, except for TssE and TssM$_{NTD}$ as judged by the BACTH assay conducted on MacConkey-maltose agar (Ahmad 2013; this study). This could be due to the fact that the interacting surfaces of TssA^{S} and TssA^{EI} are different, as they were derived from different organisms. These results largely support our hypothesis that TssA^{EI} is a functional orthologue of TssA^{S} despite amino acid sequence divergence.

3.6.3 Limitations of the BACTH system

Although the BACTH system is a practical and powerful tool to screen for protein-protein interactions *in vivo*, when considering its limitations, one that has to be taken into account is the principle guiding recombinant protein expression. The two recombinant proteins are expressed from two compatible plasmids that have diverse

replication origins: low copy number plasmids (pKT25 and pKNT25) and high copy number plasmids (pUT18 and pUT18C), and these plasmids have the same promoter (wild type promoter of the lactose operon). These will result in different relative amounts of these recombinant proteins. As in some cases, the ratio between two proteins can be critical for their interaction. Therefore the interaction has to be tested in both pairwise combinations of T25 and T18 fusions. It is common for the interaction to be detected only in one of the two combinations (Battesti and Bouveret 2012). This problem can be solved by cloning both hybrid genes on the same plasmid, and also new plasmids with the same copy number have been designed by Jervis and Green (Jervis and Green 2007). Moreover, expressing two genes from the same plasmid also allows for co-translational assembly of interacting proteins (Shieh et al. 2015). Another potential problem is the possibility of false positive and false negative results. The proteins are fused to T18 or T25 domains, which might cause the protein to be mis-folded, or become unstable, or simply prevents the interaction with its partners or has intrinsic tendency to interact with any other proteins (so-called 'sticky protein') (Battesti and Bouveret 2012).

Some strong protein-protein interactions may yield a negative result in the BACTH assay if the T18 and T25 domains are not correctly positioned as illustrated in Figure 3.23. Moreover, the assay does not reflect the strength of the interaction, as a strong interaction between two proteins can still lead to a weak phenotype if T25 and T18 are not in close enough proximity. To measure the strength of a protein-protein interaction, other techniques are required. In addition, the maltose phenotype on MacConkey agar is not as sensitive at detecting an interaction as is monitoring the Lac phenotype on agar containing X-gal. Moreover, we frequently observed a patchy colony phenotype which was difficult to rationalise, and also made it difficult to score the phenotype. This was the main rationale for performing β-galactosidase assays on the two-hybrid and three-hybrid combinations described in this study.

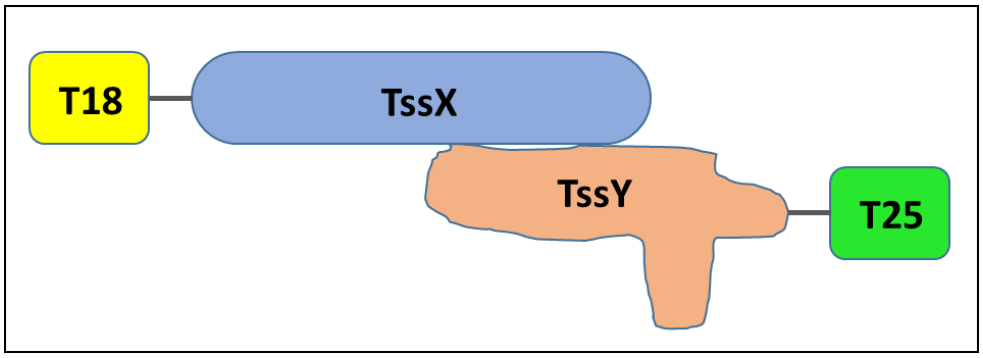


Figure 3.23 A schematic diagram illustrating how false-negative results may arise from two proteins that interact strongly. Interacting proteins TssX and TssY are fused to CyaA domains T18 and T25, respectively. In a BACTH assay carried out in an *E.coli* Δ*cyaA* host strain the complementation between T25 and T18 may be unsuccessful due to the distance separating them or steric hindrance by the interacting proteins.

Chapter 4 Biochemical analysis of interactions between TssA and other T6SS subunits

4.1 Introduction

By using the two- and three-hybrid analysis, both types of TssA ($TssA^S$ from *B. cenocepacia* and $TssA^E$ from *A. hydrophila*) were suggested to play an important role in the assembly of the T6SS as they interact with many other T6SS subunits, i.e. TssB, TssC, TssD, TssF, TssH, TssI, TssK and TssL. In addition, the interactions of $TssA^S$ with TssE and $TssM_{NTD}$ were also observed (Chapter 3). Although $TssA^{EI}$ does not appear to interact with TssE in the two hybrid assay (Ahmad 2013), the interaction between $TssA^E$ and TssE has been shown with subunits derived from enteroaggregative *E. coli* (EAEC) (Zoued et al. 2016). In order to provide biochemical evidence to support these *in vivo* results, nickel affinity pull-down assays and co-immunoprecipitation assays were employed. Due to their solubility and stability, the interactions between TssA and TssC, TssD, TssF, TssI core region, TssK, TssL and $TssM_{NTD}$ were conducted using these approashes. These experiments were carried out by co-expressing the genes encoding the two proteins of interest (TssA and another T6SS subunit) in the same *E. coli* cells using pACYCDuet or pETDuet plasmids (Figures 4.1 and 4.2). Both plasmids have two phage T7 promoters and each promoter has a *lac* operator for tight regulation and a MCS located downstream for cloning the genes for the co-expressed proteins. These two Duet vector are compatible as they were designed for overproduction of multiple proteins, although in this study the co-production of two or three proteins for an individual pull-down or co-immunoprecipitation experiment employed only one plasmid in most cases.

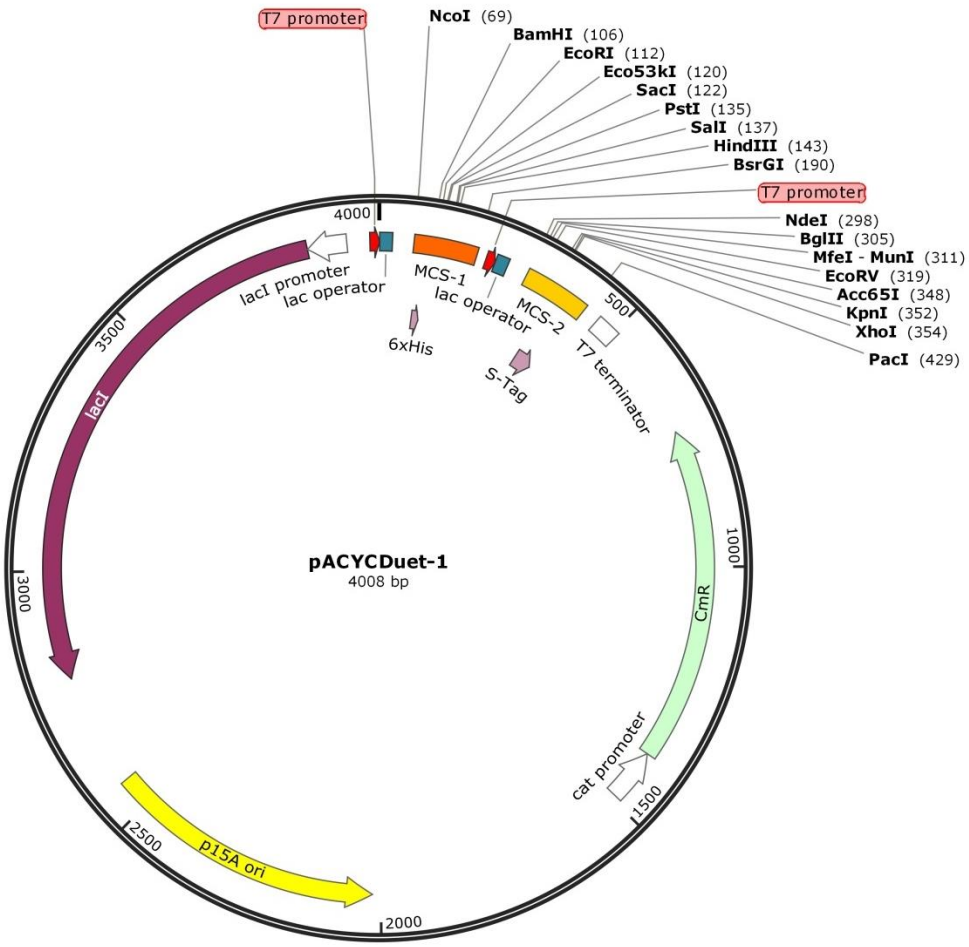


Figure 4.1 A representative map of pACYCDuet-1 vector that is designed for high level co-expression of two genes of interest. The pACYCDuet-1 vector contains two copies of the T7ϕ10 promoter (red block arrows), each followed by a *lac* operator (in blue), ribosome binding site and multiple cloning site (in orange and amber, respectively). It also carries the origin of replication of plasmid p15A, the *lacI* gene and chloramphenicol resistance gene. The plasmid has a hexa-histidine tag coding sequence (location 83-100) located between the *Nco*I and *Bam*HI restriction sites downstream of the first T7 promoter, and an S-tag coding sequence located between the *Xho*I and *Pac*I sites of the MCS for the second T7 promoter (location 366-410).

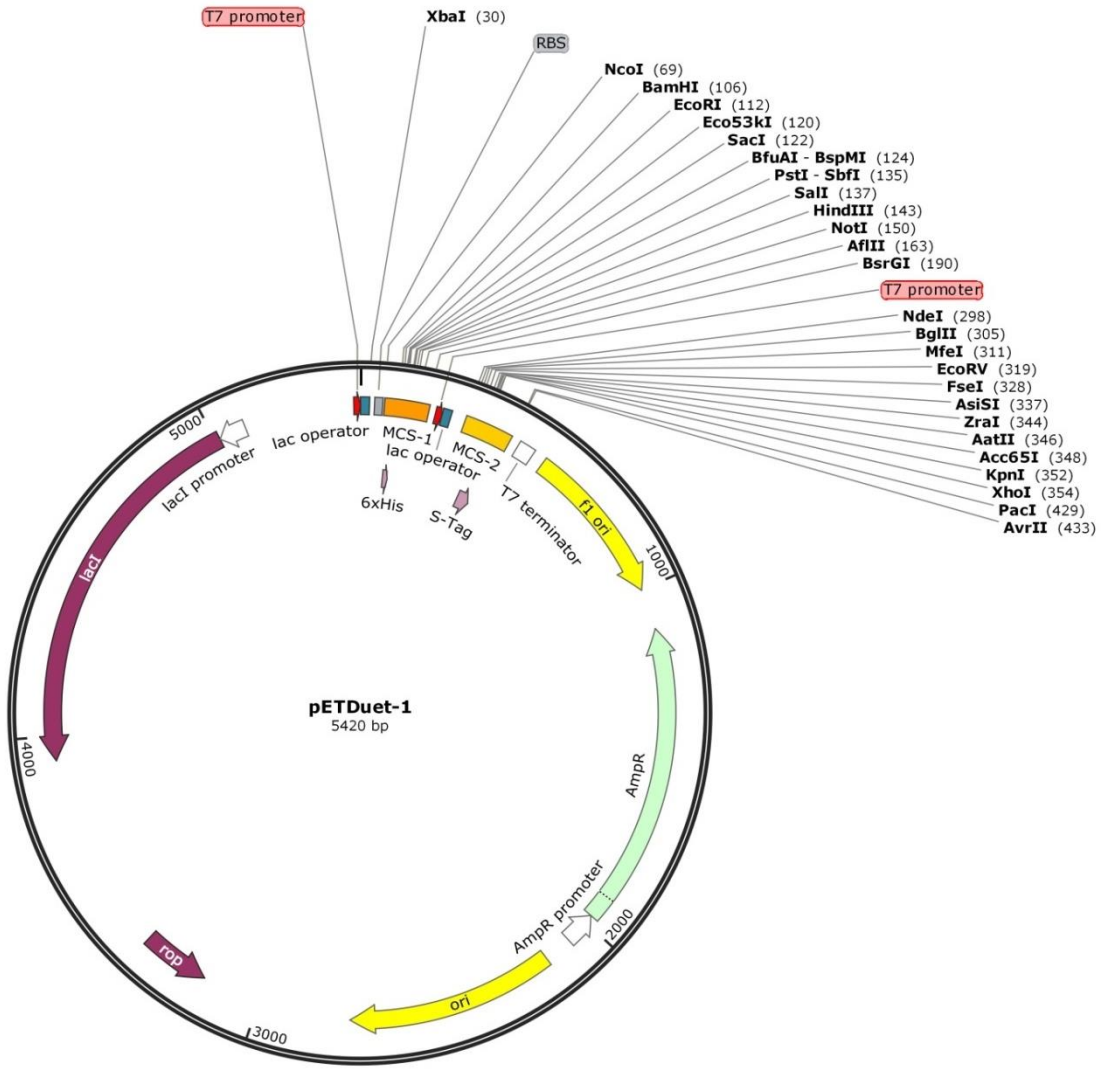


Figure 4.2 A representative map of pETDuet-1 vector that is designed for high level co-expression of two genes of interest. The pETDuet-1 vector contains two copies of the T7ϕ10 promoter (red block arrows), each followed by a *lac* operator (in blue), ribosome binding site and multiple cloning site (in orange and amber, respectively). Amp^R, ampicillin-resistance gene (*bla*). It also carries the origin of replication of plasmid ColE1, the *lacI* gene and ampicillin resistance gene. The plasmid has a hexa-histidine tag coding sequence (location 83-100) located between the *Nco*I and *Bam*HI restriction sites downstream of the first T7 promoter, and an S-tag coding sequence located between the *Xho*I and *Pac*I sites of the MCS for the second T7 promoter (location 366-410).

4.2 Nickel affinity pull-down assay of TssA^S with other T6SS subunits

4.2.1 Rationale for using $_{LinkerHis6}$.TssA^S in the nickel affinity pull-down assay

Pull-down experiments can be used to determine whether one protein interacts with another. For these experiments, one protein (the 'bait') contains an affinity tag that allows it to be recovered from the binding mixture. Often, a poly-histidine tag is used, which has a high affinity for Ni^{2+} ions. The interacting protein (the 'prey') can be identified by mass spectrometry, or in situations where a specific prey protein is being investigated, by western blotting using an antibody to the protein or to an eitope tag.

In the pull-down experiment where one protein contains a histidine tag, proteins that interact with the histidine-tagged protein will also be retained on the nickel affinity matrix. Then, they will be co-eluted with the bait protein by buffer containing a high concentration of imidazole. The retention of the interacting protein on the affinity matrix/beads is detected using specific antibodies to the native protein or to a non-native epitope that is appended to the prey protein using recombinant DNA techniques. In initial attempts to purify TssA^S, the protein was histidine-tagged at either the N-terminus (with 6 or 10 histidines) or at the C-terminus (with 10 histidines). In each case, however, only ~5% of the protein was retained on a nickel column (Ahmad 2013). It was thought that this may be due to the N- and C-terminus of TssA^S being sequestered in the centre of the folded TssA^S which might prevent the access of the histidine tag to nickel beads, as TssA^S forms a large complex (Chapter 5). The linker region of TssA^S that links its N-terminal and C-terminal domains is susceptibe to cleavage by proteases (Ahmad 2013), suggesting that it is exposed on the exterior of TssA^S. Hence, introduction of a histidine tag into the linker region is likely to lead to the tag being exposed, thereby facilitating its binding to the nickel sepharose. Furthermore, the length and amino acid sequence of this region is not conserved, and therefore it might tolerate insertion of a hexa-histidine tag without loss of function. For this reason, a plasmid was constructed (pACYCDuet-linkerHis$_6$.*tssA*^{S}) that overproduced TssA^S with a hexa-histidine tag located between amino acids 262 and 263 of TssA^S within the interdomain linker (J. Hall and M. Thomas, unpublished results). To confirm that insertion of the tag in the interdomain linker did not disrupt oligomerisation of TssA^S, the structure of the oligomer that is formed by linkerHis$_6$.TssA^S was revealed by

negative stain electron microscopy, where no detectable changes were observed in the overall geometry of the complex (Chapter 5).

As it was decided to use $_{\text{LinkerHis6}}$.TssA^{S} in the nickel pull-down experiments to first investigate the interaction of TssA^{S} with TssK and TssL, the effect of the presence of a histidine tag in the linker region of TssA^{S} on interactions between TssA^{S} and TssK, and between TssA^{S} and TssL was investigated using the BACTH assay. The two-hybrid interaction between TssA^{S} and TssK, or TssA^{S} and TssL, required the C-terminus of TssA^{S}, TssK and TssL to be free (Sections 3.3.11 and 3.3.12). Therefore, pKT25-linkerHis$_6$.*tssA*^{S} plasmid was constructed (J. Hall and M. Thomas, unpublished results). The BACTH assay results showed that the hexa-histidine tag in the linker region of TssA^{S} did not negatively affect the interaction between TssA^{S} and TssK, or TssA^{S} and TssL (Figure 4.3). Therefore, $_{\text{LinkerHis6}}$.TssA^{S} was used in the nickel affinity pull-down assay.

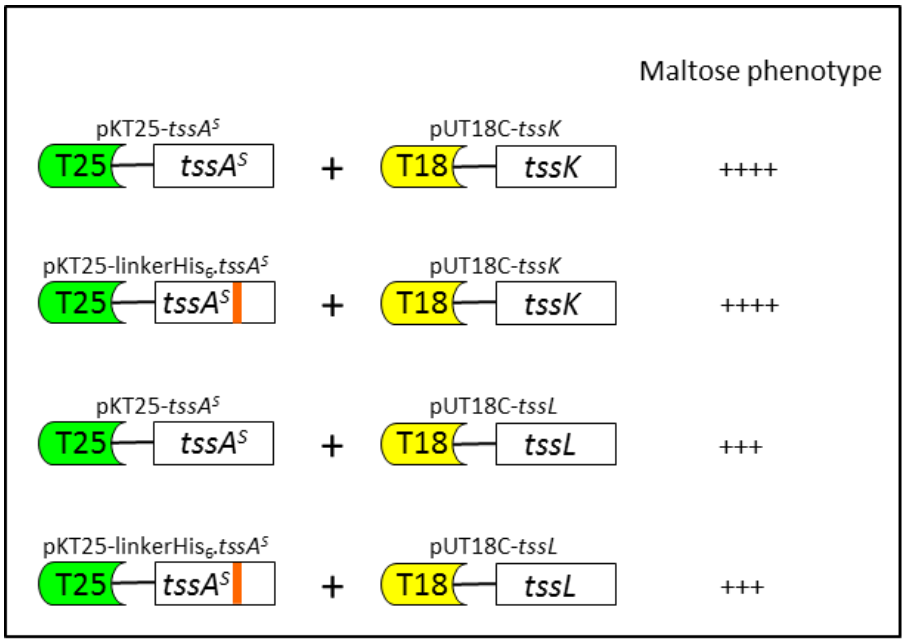


Figure 4.3 BACTH analysis of the effect of the presence of a histidine tag in the linker region of TssA^S on the interactions with TssK and TssL. A histidine tag was introduced between amino acids 262 and 263 of TssA^S within the non-conserved linker region connecting the NTR and CTD. Compatible BACTH plasmids encoding a T25-TssA^S or T25-$_{LinkerHis6}$.TssA^S fusion in combination with a T18-TssK or T18-TssL fusion were analysed using the BACTH system. The strength of the maltose phenotype observed from the BACTH assay is indicated by crosses according to the classification of the phenotype described in Table 3.2; All control combinations yielded negative maltose phenotypes (not shown).

4.2.2 Pull-down analysis of the interaction between TssA^S and TssK

To detect TssK in a nickel pull-down assay with His-tagged TssA^S, it was necessary to use antibodies to detect the interacting protein. A convenient way of doing this is to append an 'epitope tag' to the protein, to which an antibody is commercially available. Co-expression of both TssA^S and TssK proteins was conducted from pACYCDuet-1 vector. $_{\text{LinkerHis6}}$TssA^S has been previously cloned into pACYCDuet (pACYCDuet-linkerHis$_6$.*tssA*^{S}) between the *Nde*I and *Bgl*II restriction sites, under the control of the second T7 promoter (J. Hall and M. Thomas, unpublished results). Based on the BACTH results (Section 3.3.11), *B. cenocepacia* H111 *tssK* was amplified using a pair of primers with a VSV-g epitope tag coding sequence included at the 5' end of the forward primer (TssKforpET.NtermVSVGTag and TssK.BglII.rev). The amplified *tssK* DNA fragment of the expected size (~1.3 kb) was digested with restriction enzymes *Nco*I and *Bgl*II that recognized sites in the forward and reverse primers, respectively, and ligated into the *Nco*I and *Bam*HI digested pACYCDuet-linkerHis$_6$.*tssA*^{S} plasmid placing VSV-g.*tssK* under control of the first T7 promoter (Figure 4.4 A). The plasmid was called pACYCDuet-VSVg.*tssK*-linkerHis$_6$.*tssA*^{S}. In parallel, a control plasmid expressing only VSV-g tagged TssK was similarly constructed, i.e. pACYCDuet-VSVg.*tssK* (Figure 4.4 C). The nucleotide sequence of the inserted gene in both plasmids was verified by DNA sequencing.

Protein overexpression was carried out in *E. coli* strain BL21(λDE3) as described in Section 2.4.1. Following induction, both TssA^S and TssK were overproduced, although there was less TssA^S produced than TssK, and both proteins remained soluble. In the nickel affinity pull-down assay, the histidine tag allows TssA^S stick to the nickel sepharose, and if TssK interacts with TssA^S, TssK will be also retained. The results showed that a fraction of $_{\text{LinkerHis6}}$.TssA^S applied to the resin bound to nickel (~10%), which is higher than that previously observed for N- or C-terminal histidine tagged TssA^S (Figure 4.4 B). In addition, some $_{\text{VSVg}}$.TssK was retained on the resin and eluted with 500 mM imidazole along with the bound $_{\text{LinkerHis6}}$.TssA^S. However, a fraction of the $_{\text{VSVg}}$.TssK control was also, unexpectedly, observed to stick to nickel (Figure 4.4 D). Moreover, there was no obvious increase in the amount of TssK retained on the nickel sepharose when it was overproduced together with TssA^S. Therefore, the nickel pull-down assay did not confirm the interaction between TssA^S and TssK.

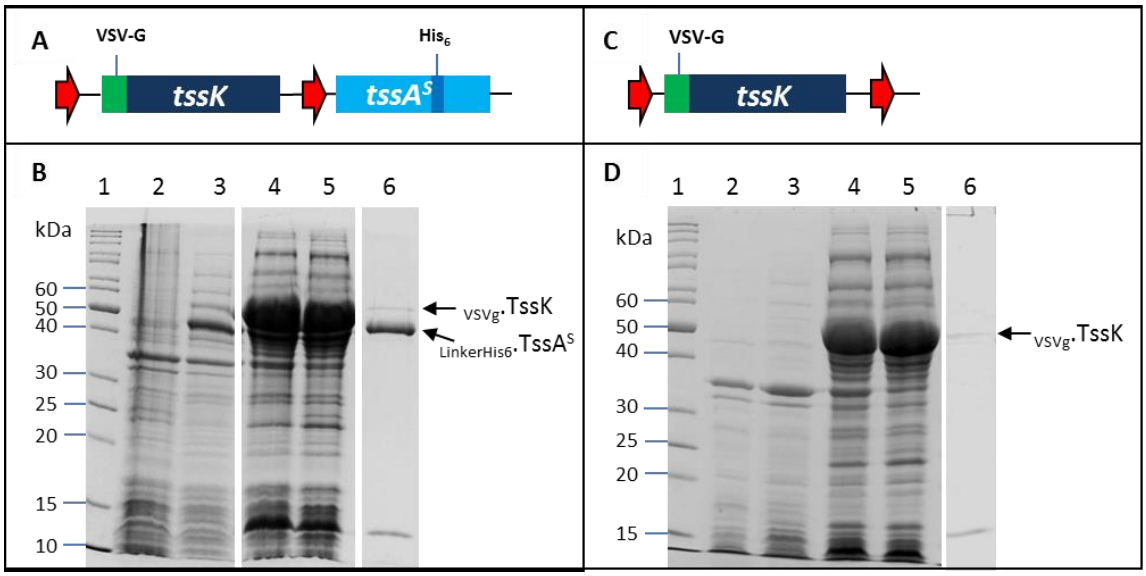


Figure 4.4 Nickel affinity pull-down analysis of the interaction between TssA^S and TssK. A. Schematic drawing of the arrangement of linkerHis$_6$.*tssA*^{S} and VSV-g tagged *tssK* genes in pACYCDuet-VSVg.*tssK*-linkerHis$_6$.*tssA*^{S}. The red arrows represent the two T7 promoters in the plasmid. **B**. A coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssK*-linkerHis$_6$.*tssA*^{S} in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 37^{o}C. Binding to the nickel sepharose was performed using the cell lysate soluble fraction in a buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole and elution was performed in the same buffer containing 500 mM imidazole. The arrows indicate the expected locations of TssA^S (~42 kDa) and TssK (~51 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction used for binding to nickel sepharose; lane 6, elution from nickel sepharose. **C**. Schematic drawing showing the location of VSV-g tagged *tssK* gene in pACYCDuet-VSVg.*tssK*. The red arrows represent the two T7 promoters in the plasmid. **D**. A coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssK* in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 37^{o}C. Binding to the nickel sepharose and elution were performed in the same condition as described above. The arrow indicates the expected location of TssK (~51 kDa) based on its MW. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction used for binding to nickel sepharose; lane 6, elution from nickel sepharose.

4.2.3 Pull-down analysis of the interaction between TssA^S and TssL

Co-expression of both TssA^S and TssL was achieved from pACYCDuet that contained the TssA^S and TssL coding sequences under the control of the two T7 promoters. In parallel, a control plasmid expressing only TssL was also constructed. According to the BACTH results, the interaction between TssA^S and TssL requires the C-terminus of both proteins to be free (Section 3.3.12). Therefore, *B. cenocepacia* H111 *tssL* was amplified with a pair of primers designed specifically to amplify the gene with a VSV-g tag coding sequence contained in the forward primer (TssLforpET.NtermVSVGTag and TssL.Rev). The amplified *tssL* DNA fragment of the expected size (~600 bp) was digested with restriction enzymes (*Nco*I and *Bgl*II) that recognized sites in the forward and reverse primers, respectively, and ligated into *Nco*I and *Bam*HI digested pACYCDuet-linkerHis$_6$.*tssA*^{S} and pACYCDuet for constructing pACYCDuet-VSVg.*tssL*-linkerHis$_6$.*tssA*^{S} (Figure 4.5 A) and pACYCDuet-VSVg.*tssL* (Figure 4.5 C), respectively. The nucleotide sequence of the inserted gene in both plasmids was verified by DNA sequencing.

Both TssA^S and TssL were overproduced and remained in the soluble fraction following cell lysis. In the nickel affinity pull-down assay, a fraction of $_{\text{LinkerHis6}}$.TssA^S bound to nickel as before, and was eluted with a high concentratin of imidazole. A small amount of $_{\text{VSVg}}$.TssL was also retained on the column and co-eluted with TssA^S (Figure 4.5 B). However, unexpectedly $_{\text{VSVg}}$.TssL expressed in the absence of $_{\text{LinkerHis6}}$.TssA^S also bound to the nickel resin (Figure 4.5 D). Given that there was no obvious increase in the amount of TssL retained on the nickel sepharose when it was co-overproduced with TssA^S, the interaction between TssA^S and TssL *in vitro* was not shown by the nickel affinity pull-down assay.

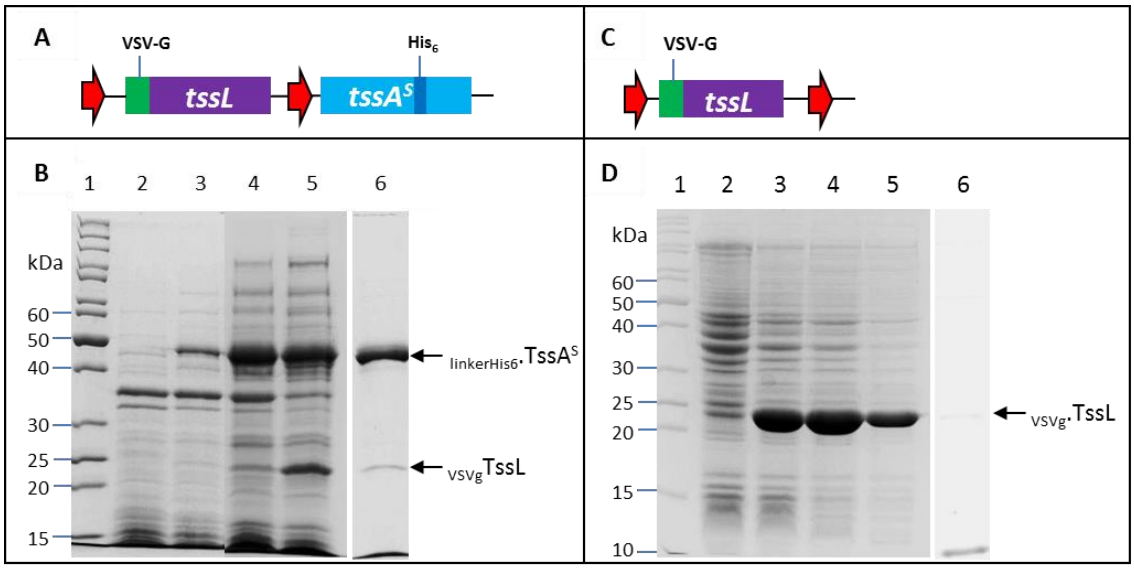


Figure 4.5 Nickel affinity pull-down analysis of the interaction between TssA^S and TssL. A. Schematic drawing showing organisation of the linkerHis$_6$.*tssA*^{S} and VSV-g tagged *tssL* into pACYCDuet-VSVg.*tssL*-linkerHis$_6$.*tssA*^{S}. The red arrows represent the two T7 promoters in the plasmid. **B**. A Coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssL*-linkerHis$_6$.*tssA*^{S} in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 37^{o}C. Binding to the nickel sepharose was performed using the cell lysate soluble fraction in a buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole and elution was performed in the same buffer containing 500 mM imidazole. The arrows indicate the expected locations of TssA^S (~42 kDa) and TssL (~24 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction used for binding to nickel sepharose; lane 6, elution from nickel sepharose. **C**. Schematic drawing showing the location of the VSV-g tagged *tssK* gene in pACYCDuet-VSVg.*tssK*. The red arrows represent the two T7 promoters in the plasmid. **D**. A Coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssK* in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 37^{o}C. Binding to the nickel sepharose and elution were performed in the same condition as described above. The arrow indicates the expected location of TssL based on its size (~24 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction used for binding to nickel sepharose; lane 6, elution from nickel sepharose.

4.2.4 Pull-down analysis of interactions between TssA^S, TssK and TssL

As the three-hybrid result suggested that the presence of TssL might stabilise the interaction between TssA^S and TssK (Section 3.4.3), it was decided to repeat the attempt to pull-down TssK and/or TssL by TssA^S in the presence of both prey proteins using the nickel affinity pull-down assay. Initially, two compatible expression plasmids pACYCDuet and pETDuet in combination were used to express these 3 proteins. However, the expression from the pETDuet plasmid dominated, and there was little or no expression from the pACYCDuet plasmid (result not shown). For this reason, it was decided to express TssA^S, TssK and TssL from the same plasmid, pACYCDuet, although it contains only two promoters.

The strategy for inserting three genes into pACYCDuet was to clone *tssL* gene into the pACYCDuet-VSVg.*tssK*-linkerHis$_6$.*tssA*^{S} plasmid that was already constructed. *B. cenocepacia* H111 *tssL* was amplified with a pair of primers designed to amplify the gene with a HA epitope tag coding sequence contained in the forward primer (TssL.MfeI.NcoI.HAtag.For and TssL.KpnI.Rev). The amplified *tssL* DNA fragment of the expected size (~600 bp), pACYCDuet and pACYCDuet-VSVg.*tssK*-linkerHis$_6$.*tssA*^{S} plasmids were digested with restriction enzymes *Mfe*I and *Kpn*I that recognized sites in the forward and reverse primers and the plasmid MCS. The forward *tssL* amplification primer contained an additional stop codon to ensure no TssA^S-TssL fusion protein was formed from very low level translation read-through of the TssA^S stop codon, and a Shine-Dalgarno sequence for *tssL* translation (Figure 4.6 A). Ligation was carried out to allow the insertion of the *tssL* DNA fragment into the two plasmids to construct pACYCDuet-HA.*tssL* and pACYCDuet-VSVg.*tssK*-linkerHis$_6$.*tssA*^{S}-HA.*tssL*. In the latter case, both *tssA* and *tssL* were under the control of the second T7 promoter of the pACYCDuet plasmid (Figure 4.6 B). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

Induction of gene expression from pACYCDuet-VSVg.*tssK*-linkerHis$_6$.*tssA*^{S}-HA.*tssL* in *E. coli* strain BL21(λDE3) cells resulted in successful overproduction of all three proteins (Figure 4.6 C). TssK was overproduced in a larger amount than TssA^S and TssL. All of these three proteins were present in the soluble fraction following cell lysis. In the nickel affinity pull-down assay, HA-tagged TssL was not retained on the nickel sepharose. However, as observed previously (Section 4.2.2), VSV-g epitope-tagged

TssK bound to the resin non-specifically, but even allowing for this, the assay did not show any obvious increased retention of TssK to the nickel sepharose when all three proteins were present (Figure 4.6 C). Therefore, interactions between $TssA^S$ and TssK, or $TssA^S$ and TssL were not demonstrated using this assay.

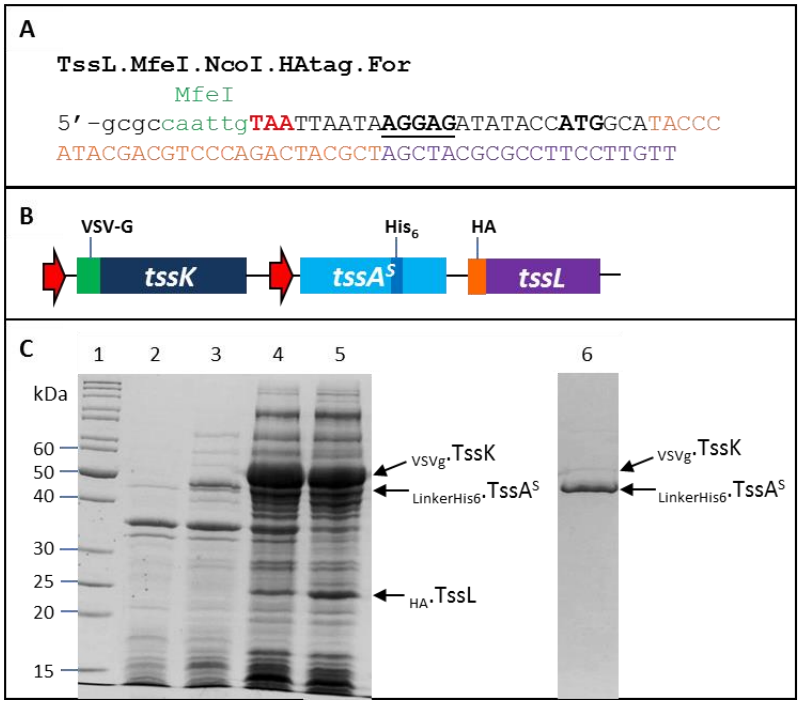


Figure 4.6 Nickel affinity pull-down analysis of the interactions between TssA^S and TssK, and TssA^S and TssL when all three proteins were present. A. Sequence of the forward primer used to amplify HA.*tssL* for the TssA^S, TssK and TssL tri-partite pull-down assay. *Mfe*I site and stop codon are shown in green and red, respectively. The Shine-Dalgano sequence introduced to allow for $_{HA}$TssL translation is underlined and in bold font. HA tag codons are shown in orange and the first 20 dodons of *tssL* are shown in purple. **B**. Schematic drawing showing arrangement of linkerHis$_6$.*tssA*^{S}, VSV-g tagged *tssK* and HA tagged *tssL* in pACYCDuet-VSVg.*tssK*-linkerHis$_6$.*tssA*^{S}-HA.*tssL*. The red arrows represent the two T7 promoters present in the plasmid. **C**. A Coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssK*-linkerHis$_6$.*tssA*^{S}-HA.*tssL* plasmid in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 37^{o}C. Binding to the nickel sepharose was performed using the cell lysate soluble fraction in a buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole and elution was performed in the same buffer containing 500 mM imidazole. The arrows indicate the expected locations of TssA^S (~42 kDa), TssK (~51 kDa) and TssL (~24 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, crude cell lysate from uninduced cells; lane 3, total protein from cells following induction; lane 4, cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction following induction used for binding to nickel sepharose; lane 6, elution from nickel sepharose. The pull-down of TssK control is shown in Figure 4.4D.

4.2.5 Pull-down analysis of the interaction between $TssA^S$ and $TssI_{BCAS0667}$

By using the BACTH assay, the interaction between $TssA^S$ and TssI has been demonstrated (Section 3.3.9). The *in vitro* interaction between these two proteins was investigated using a nickel affinity pull-down assay. To construct a plasmid for co-expressing both proteins, the TssI gene *B. cenocepacia* strain H111 corresponding to BCAS0667 of strain J2315 (i.e. I35_7843) was amplified with a pair of primers designed specifically to amplify the gene with a VSV-g epitope tag coding sequence contained in the forward primer (TssI0667.BspHI.VSVGTag.for and VgrGrev). Plasmid pET28a-His6.TssI served as the template. The amplified *tssI* DNA fragment of the expected size (~2.85 kb) was digested with *Bgl*II that recognized a site incorporated into the reverse primer whereas the N-terminal coding end of the *tssI* DNA fragment was left with a blunt end. pACYCDuet-LinkerHis$_6$.*tssA*^{S} was digested with *Nco*I that recognized a site in the first MCS followed by filling in the sticky ends by Klenow DNA polymerase to generate a blunt end (described as Section 2.3.11). Then the plasmid was further digested with *Bam*HI. Ligation was carried out to allow the insertion of the *tssI* DNA fragment into the plasmid for constructing pACYCDuet-VSVg.*tssI*-LinkerHis$_6$.*tssA*^{S} (the *Bgl*II-generated sticky end in the *tssI* DNA fragment is compatible with the *Bam*HI-generated sticky end in the vector) (Figure 4.7 A). The nucleotide sequence of the inserted gene was verified by DNA sequencing. A plasmid expressing only VSV-g tagged TssI as a pull-down control (pACYCDuet-VSVg.*tssI*) was also constructed by deleting *tssA*^{S} from pACYCDuet-VSVg.*tssI*-LinkerHis$_6$.*tssA*^{S} by digestion with *Bsr*GI and *Acc*65I followed by self-ligation of the plasmid (Figure 4.7 C).

Protein overexpression was carried out in *E. coli* strain BL21(λDE3) cells. A protein corresponding in size to VSV-g tagged TssI (102 kDa) was present in extracts of cells containing pACYCDuet-VSVg.*tssI* and remained soluble following induction (Figure 4.7 D). Co-expression of $_{VSVg}$.TssI and $_{LinkerHis6}$.$TssA^S$ was also successful in cells harbouring pACYCDuet-VSVg.*tssI*-linkerHis$_6$.*tssA*^{S} and both proteins remained soluble (Figure 4.7 B). In the nickel affinity pull-down assay, $_{LinkerHis6}$.$TssA^S$ bound to the nickel sepharose and $_{VSVg}$.TssI did not bind to the nickel sepharose non-specifically. A protein of 12 kDa was observed to bind to the nickel sepharose, which is lysozyme used in the cell lysis procedure. However, there was no visible amount of $_{VSVg}$.TssI retained on the

resin in the presence of $_{\text{LinkerHis6}}$.TssA^{S} (Figure 4.7). Therefore, the nickel pull-down assay did not provide supporting evidence for the interaction between TssA^{S} and TssI.

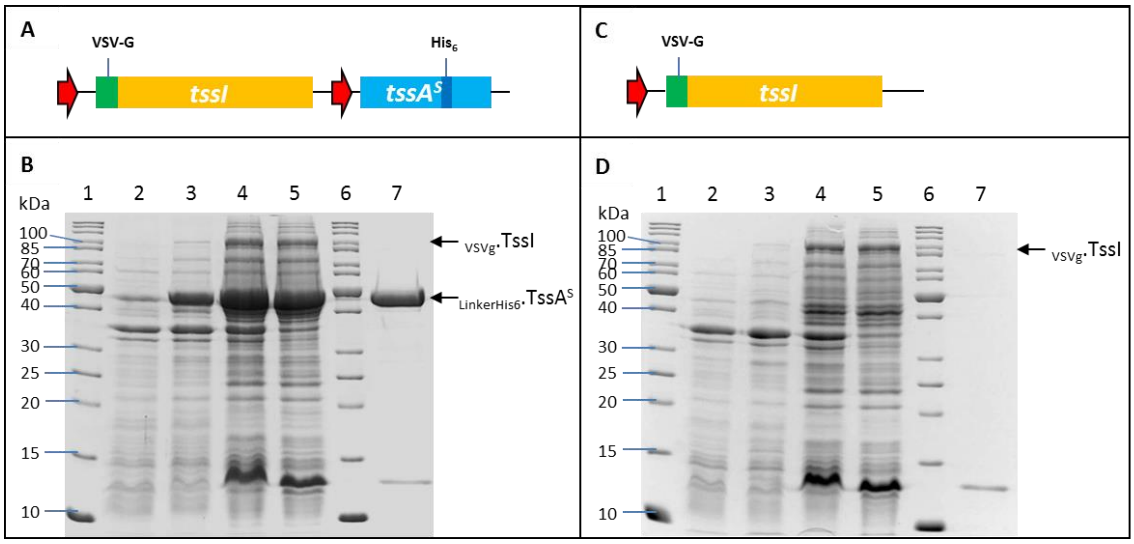


Figure 4.7 Nickel affinity pull-down assay of the interaction between TssA^S and TssI. A. Schematic drawing of the arrangement of VSV-g tagged *tssI* and LinkerHis$_6$.*tssA*^{S} in pACYCDuet-VSVg.*tssI*-LinkerHis$_6$.*tssA*^{S}. The red arrows represent the two T7 promoters in the plasmid. **B**. A coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssI*-LinkerHis$_6$.*tssA*^{S} in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37^{o}C. Binding to the nickel sepharose and elution were performed in the same condition as described above. The arrows indicate the expected locations of TssI (~102 kDa) and TssA^S (~42 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction; lane 5, soluble fraction of cell lysate following induction used for binding to nickel sepharose; lane 6, EZ-Run^{TM} Rec protein ladder (Fisher); lane 7, elution from nickel sepharose. **C.** Schematic drawing showing the location of VSV-g tagged *tssI* genes in pACYCDuet-VSVg.*tssI*. The red arrow represents the first T7 promoter in the plasmid, as the second T7 promoter was removed when deleting LinkerHis$_6$.*tssA*^{S} gene from pACYCDuet-VSVg.*tssI*-LinkerHis$_6$.*tssA*^{S}. **D**. A coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssI* in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37^{o}C. Binding to the nickel sepharose was performed using the cell lysate soluble fraction in a buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole and elution was performed in the same buffer containing 500 mM imidazole. The arrow indicates the expected location of TssI based on its size (~102 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction; lane 5, soluble fraction of cell lysate following induction used for binding to nickel sepharose; lane 6, EZ-Run^{TM} Rec protein ladder (Fisher); lane 7, elution from nickel sepharose. (The 12 kDa protein is lysozyme used in the cell lysis procedure)

4.3 Co-IP of TssA^S with other T6SS subunits using anti-TssA^S antibody

4.3.1 Generation of anti-TssA^S antibody

Due the possibility that the His-tag on TssA^S was masked by the binding of TssK and therefore only uncomplexed TssA^S would be retained by the nickel resin, it was decided to employ co-immunoprecipitation (co-IP) to demonstrate the interaction between TssA^S and other T6SS proteins. This should circumvent the problem of masking if a polyclonal antibody to TssA^S was used. Anti-TssA^S antibody was raised in rats at Bioserve UK (University of Sheffield) using native (i.e. untagged) TssA^S protein as described in Section 2.6.4. TssA^S was overproduced in *E. coli* BL21(λDE3) cells containing pET14b-*tssA*^{S} and purified by PEI precipitation, 30% ammonium sulphate precipitation and SEC (Section 5.2.1.2). The working concentration of anti-TssA^S antibody was tested prior to its use in the co-IP (Figure 4.8).

4.3.2 Co-IP analysis of interactions between TssA^S and TssK in the presence and absence of TssL

In the co-IP using anti-TssA^S antibody, a whole cell extract was prepared from *E. coli* cells containing single plasmid that expressed TssA^S with epitope-tagged TssK and with epitope-tagged TssK and TssL. The interaction between TssA^S and other T6SS proteins allows them to form a complex that will exist in the cell extract. Anti-TssA^S antibody was then added to the extract, forming a new complex by anti-TssA^S antibody binding to TssA^S. This protein-protein complex was then immobilized on protein A-sepharose beads (Sigma). Proteins that do not bind were removed by a series of washes. The complex was then eluted from the beads and dissociated by boiling in 2x sample buffer (Section 2.7.4.1). The material bound to the sepharose beads was then analysed by SDS-PAGE.

The interaction between TssA^S and $_{VSVg}$.TssK in the presence or absence of $_{HA}$.TssL was analysed by co-IP using the anti-TssA^S antibody. In the material that bound to the sepharose beads, $_{VSVg}$.TssK was present along with TssA^S whether $_{HA}$.TssL was present or not. However, $_{VSVg}$.TssK and $_{HA}$.TssL were found to non-specifically bind to the sepharose beads without TssA^S and the anti-TssA^S antibody (Figure 4.9). Other methods were tried to decrease the non-specific binding problem, e.g. increasing wash times and

blocking the sepharose beads by BSA. However, despite this, the problem was not resolved (results not shown).

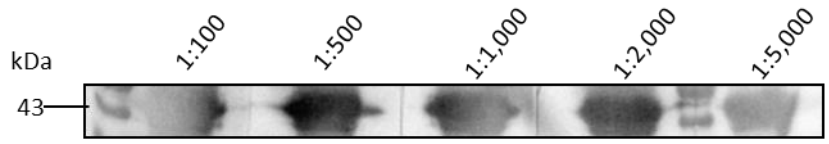


Figure 4.8 Western blotting showing the detection of TssA^S by anti-TssA^S antibody raised in rats. NativeTssA^S was overexpressed following induction of pET14b-*tssA*^{S} in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 37^{o}C. Soluble fraction of the cell lysate following induction was separated by electrophoresis in a 12% SDS-PA gel and immunodetected with anti-TssA^S antibody raised to purified TssA^S in rats. Primary anti-TssA^S antibody was diluted in different ratios as indicated, and 5,000x dilution of HRP-conjugated anti-rat secondary antibody was used in the pull-down experiments.

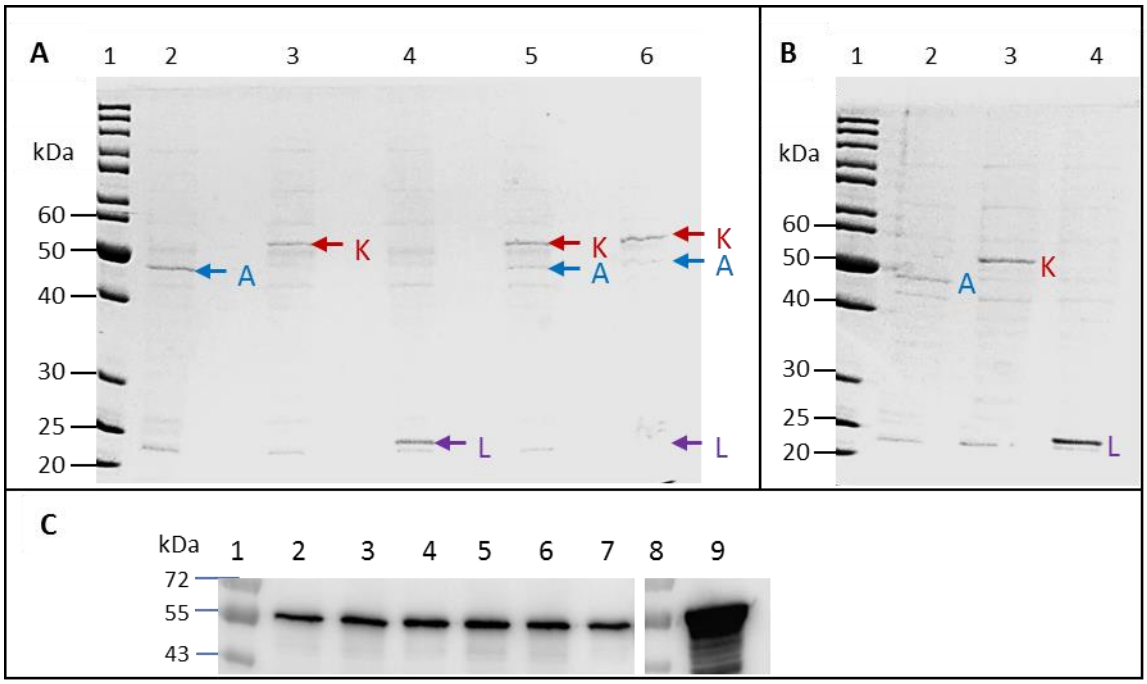


Figure 4.9 Co-IP of TssA^S and TssK in the presence or absence of TssL using anti-TssA^S antibody. Coomassie blue-stained 10% SDS-PA gels show the co-IP of TssA^S and $_{VSVg}$.TssK in the presence or absence of $_{HA}$.TssL using anti-TssA^S antibody and no-antibody control. Expression from pACYCDuet-VSVg.*tssK*-LinkerHis$_6$.*tssA*^{S}-HA.*tssL*, pACYCDuet-VSVg.*tssK*-LinkerHis$_6$.*tssA*^{S}, pACYCDuet-VSVg.*tssK*, pACYCDuet-LinkerHis$_6$.*tssA*^{S} and pACYCDuet-HA.*tssL* was induced in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37^{o}C. Soluble fraction of the cell lysates were used for co-IP in buffer containing 50 mM Tris-HCl (pH 8.0) and 50 mM NaCl. The immunoprecipitated (IP) material was separated in 12% SDS-PA gels. **A**. IP material bound to the anti-TssA^S antibody. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, IP from cells only expressing TssA^S (pACYCDuet-LinkerHis$_6$.*tssA*^{S}); lane 3, IP from cells expressing only TssK (pACYCDuet-VSVg.*tssK*); lane 4, IP from cells expressing only TssL (pACYCDuet-HA.*tssL*); lane 5, IP from cells co-expressing TssA^S and TssK (pACYCDuet-VSVg.*tssK*-LinkerHis$_6$.*tssA*^{S}); lane 6, IP from cells co-expressing TssA^S, TssK and TssL (pACYCDuet-VSVg.*tssK*-LinkerHis$_6$.*tssA*^{S}-HA.*tssL*). Proteins that were present in the immunoprecipitated material are indicated with coloured arrows. **B**. No-antibody control. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, IP from cells only expressing TssA^S; lane 3, IP from cells expressing only TssK; lane 4, IP from cells expressing only TssL. **C**. Western blotting of proteins that bound to the anti-TssA^S antibody, immuno-detected with anti-VSVg monoclonal antibody for $_{VSVg}$.TssK. Lane 1, EZ-Run™ Prestained Rec Protein Ladder (Fisher); lane 2, IP from cells expressing TssK that bound to the anti-TssA^S antibody; lane 3, IP from cells co-expressing TssA^S and TssK that bound to the anti-TssA^S antibody; lane 4, IP from cells co-expressing TssA^S, TssK and TssL that bound to the anti-TssA^S antibody; lane 5, IP from cells expressing TssK without antibody; lane 6, IP from cells co-expressing TssA^S and TssK without antibody; lane 7, IP from cells co-expressing TssA^S, TssK ang TssL without antibody; lane 8 EZ-Run™ Prestained Rec Protein Ladder (Fisher); lane 9, cell extracts of VSV-g tagged TssK overexpression following induction.

4.4 Co-IP of TssA^{S} with other T6SS subunits using anti-FLAG affinity gel

4.4.1 Co-IP analysis of interactions between TssA^{S} and other T6SS subunits using anti-FLAG affinity gel

The analysis of the interaction between TssA^{S} and other T6SS subunits were initially carried out by nickel affinity pull-down assay and co-IP using anti-TssA^{S} antibody. However, in both cases, there was a problem of non-specific binding of the prey proteins to the resins to which the bait specifically bound that was not able to be solved despite a lot of effort spent adjusting experimental conditions. Based on this, TssA^{S} was tagged by a different epitope tag, i.e. FLAG (DYKDDDDK), at its N-terminus and a commercial anti-FLAG M2 affinity gel (Sigma) was used for precipitating TssA^{S}. The anti-FLAG M2 affinity gel is an immunoglobulin G (IgG1) monoclonal antibody covalently attached to agarose by a hydrazide linkage. To carry out the co-IP, soluble fraction of the cell lysate following induction of $_{FLAG}$.TssA^{S} and potential interacting T6SS subunits (TssX) in *E. coli* was incubated with the anti-FLAG coupled beads, by which FLAG-tagged TssA^{S} can be recovered from the lysate along with the TssA^{S}-interacting proteins. A similar amount of TssX was applied to the resin in the absence of $_{FLAG}$.TssA^{S} as an experimental control. As an additional control, $_{FLAG}$.TssA^{S} was also immunoprecipitated in the absence of TssX. The unbound material was removed by a series of washes. The immunoprecipitated material was then eluted from the beads by boiling with sample buffer and analysed by SDS-PAGE and western blotting using antibodies that are specific to the epitope tags (Section 2.7.4.2.).

Co-expression of two or three proteins for co-IP analysis was achieved by cloning multiple *tss* genes into a single plasmid, as the expression of proteins from one plasmid dominated and little expression from the other was observed when expressing two or more proteins using a pair of compatible expression plasmids, i.e. pACYCDuet and pETDuet (results not shown). In most cases, pACYCDuet plasmid was used for protein co-expression (Figure 4.1). In general, the second Tss protein was tagged with a VSV-g epitope tag at either its N- or C-terminus based on the requirement for an unfused terminus as determined by the BACTH assay (Chapter 3). Maltose binding protein (MBP) was used as a negative control in the co-IP in combination with TssA^{S}.

4.4.1.1 Construction of a plasmid producing FLAG-tagged TssA^S and analysis of $_{FLAG}$.TssA^S overexpression

In order to construct a plasmid for overproducing $_{FLAG}$TssA^S as a control in the co-IP experiments, *B. cenocepacia* H111 *tssA*^{S} was amplified with a pair of primers designed specifically to amplify the gene with a FLAG tag coding sequence contained in the forward primer (TssA.NdeI.FLAG.for and TssA.BglII.rev). The amplified *tssA*^{S} DNA fragment of the expected size (~1.1 kb) and pACYCDuet plasmid were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites in the primers and the plasmid MCS. Ligation was carried out to allow the insertion of the *tssA*^{S} DNA fragment into pACYCDuet under the control of the second T7 promoter for construction of pACYCDuet-FLAG.*tssA*^{S} (Figure 4.10 A). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

To overproduce $_{FLAG}$.TssA^S, *E. coli* strain BL21(λDE3) cells containing pACYCDuet-FLAG.*tssA*^{S} were induced with IPTG. Upon SDS-PAGE analysis, an abundant protein of 42 kDa corresponding to $_{FLAG}$.TssA^S was observed in cell extracts (Figure 4.10 B). The overproduced $_{FLAG}$.TssA^S remained in the soluble fraction following cell lysis.

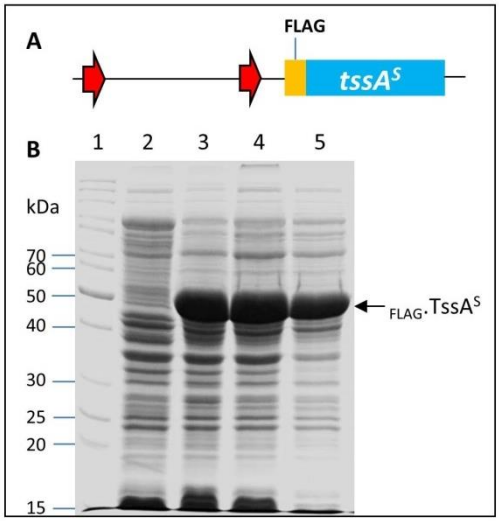


Figure 4.10 Overproduction and solubility of $_{FLAG}$.TssA^S. A. Schematic drawing showing the location of FLAG tagged *tssA*^{S} in pACYCDuet-FLAG.*tssA*^{S}. The red arrows represent the two T7 promoters in the plasmid. **B**. A coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-FLAG.*tssA*^{S} in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37^{o}C. The arrow indicates the expected location of TssA^S based on its size (~42 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

4.4.1.2 Co-IP analysis of the interaction between TssA^S and TssC

According to the result of the BACTH assay, TssA^S interacts with TssC when the N-terminus of TssC is free. Therefore, it was decided to introduce a VSV-g epitope tag to the C-terminus of TssC leaving the N-terminus free for the potential interaction with TssA^S. A pair of primers, tssC.NcoI.for and tssC.VSVg.BamHI.rev, were used for PCR amplification of *B. cenocepacia* H111 *tssC*, in which a VSV-g tag coding sequence was contained in the reverse primer. The plasmids for expressing TssC alone and for co-expression of TssA^S and TssC were constructed in parallel. The amplified *tssC* DNA fragment of the expected size (~1.5 kb), and pACYCDuet-1 and pACYCDuet-FLAG.*tssA*^{S} plasmids were digested with restriction enzymes *Nco*I and *Bam*HI that recognized sites in the primers and in the first MCS of the plasmids. Ligation was carried out to allow the insertion of the *tssC* DNA fragment into the two plasmids for constructing pACYCDuet-*tssC*.VSVg (Figure 4.11 A) and pACYCDuet-*tssC*.VSVg-FLAG.*tssA*^{S} (Figure 4.12 A). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

Protein overexpression was conducted in *E. coli* strain BL21(λDE3). Following induction of the T7 promoters on pACYCDuet-*tssC*.VSVg with 1 mM IPTG at 37^{o}C, a large amount of protein with the expected size of TssC (56 kDa) was produced, which was largely insoluble judged by SDS-PAGE (Figure 4.11 B). However, a very small fraction of TssC.$_{VSVg}$ was found to be soluble as demonstrated by immuno-detection with anti-VSV-g antibody (Figure 4.11 C). Other induction conditions, such as low temperature (30^{o}C and 22^{o}C) and lower concentrations of IPTG (0.5 mM and 0.1 mM), were tried in order to increase the fraction of soluble TssC.$_{VSVg}$. Despite these variations in inducing conditions, there was no obvious increase in the solubility of TssC.$_{VSVg}$ (an example is shown in Figure 4.11 D).

In the co-expression of $_{FLAG}$.TssA^S and TssC.$_{VSVg}$, TssC.$_{VSVg}$ was overproduced in a larger amount compared to $_{FLAG}$.TssA^S. Some $_{FLAG}$.TssA^S remained soluble, whereas again no visible amount of soluble TssC.$_{VSVg}$ was identified by SDS-PAGE (Figure 4.12 B). When the induction temperature was decreased to 22^{o}C and the concentration of IPTG to 0.5 mM, TssC.$_{VSVg}$ was still overproduced, and a larger fraction of overproduced TssC.$_{VSVg}$ was observed to be soluble (Figure 4.12 C). However, an even smaller amount of $_{FLAG}$.TssA^S was produced compared to the induction at 37^{o}C.

Overexpression of $_{FLAG}$.TssA^{S} in the absence of TssC is shown in Figure 4.10 (Section 4.4.1.1). Co-IP was conducted with the soluble fraction of the cell lysate following induction with 0.5 mM IPTG induction at 22^{o}C for TssC.$_{VSVg}$ overexpression alone and in combination with $_{FLAG}$.TssA^{S}. The results showed that TssC.$_{VSVg}$ did not bind to the anti-FLAG beads in the absence of $_{FLAG}$.TssA^{S}, whereas when it was co-expressed with $_{FLAG}$.TssA^{S}, TssC.$_{VSVg}$ was present in the immuno-precipitated material along with $_{FLAG}$.TssA^{S} suggesting a specific interaction between these two proteins (Figure 4.13).

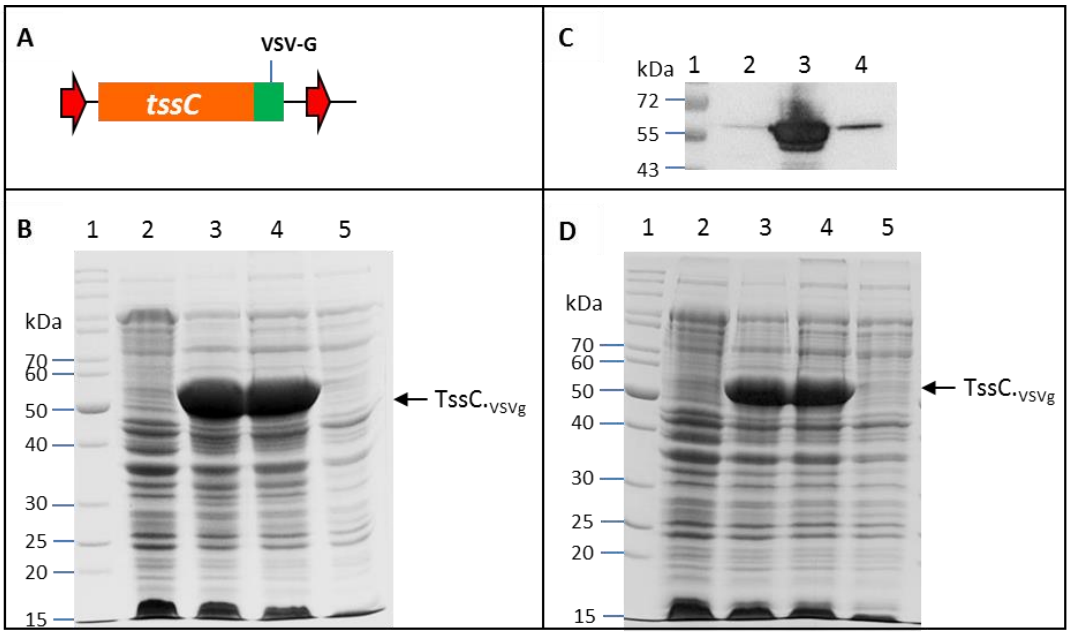


Figure 4.11 Expression and solubility of TssC in different induction conditions. A. Schematic drawing showing the location of VSV-g tagged *tssC* in pACYCDuet-*tssC*.VSVg. The red arrows represent the two T7 promoters in the plasmid. **B**. A coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-*tssC*.VSVg in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37°C. The arrow indicates the expected location of TssC based on its size (~56 kDa). Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **C**. Immuno-detection of TssC expression in *E. coli* strain BL21(λDE3) cells following 1mM IPTG induction at 37°C. Proteins in the cell lysate were fractionated by SDS-PAGE and, following western transfer, were probed with anti-VSVg mAb. Lane 1, EZ-Run™ Prestained Rec Protein Ladder; lane 2, total cell protein from uninduced cells; lane 3, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 4, soluble fraction of cell lysate following induction. **D**. A coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-*tssC*.VSVg in *E. coli* strain BL21(λDE3) with 0.5 mM IPTG at 22°C. The arrow indicates the expected location of TssC based on its size (~56 kDa). Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

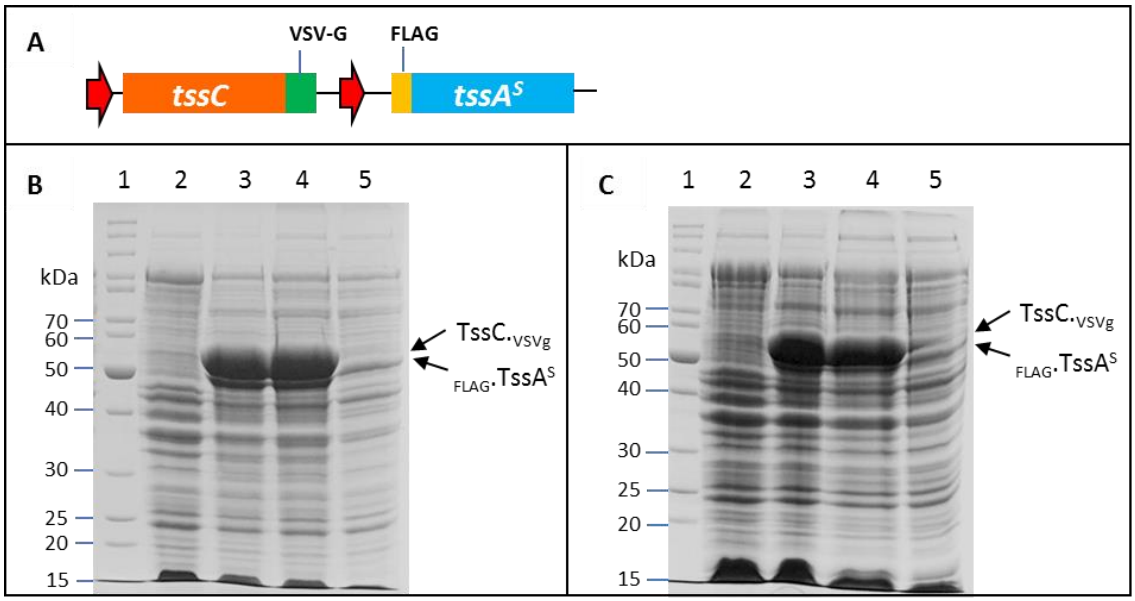


Figure 4.12 Co-expression of TssC and TssA^S. A. Schematic drawing showing the arrangement of VSV-g tagged *tssC* and FLAG-tagged *tssA*^{S} in pACYCDuet-*tssC*.VSVg-FLAG.*tssA*^{S}. The red arrows represent the two T7 promoters present in the plasmid. **B&C**. Coomassie blue-stained 10% SDS-PA gels showing the analysis of proteins following induction of pACYCDuet-*tssC*.VSVg-FLAG.*tssA*^{S} in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37^{o}C (**B**) or 0.5 mM IPTG at 22^{o}C (**C**). The arrows indicate the expected locations of TssC (~56 kDa) and TssA^S (~42 kDa). Lanes 1, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 2, total cell protein from uninduced cells; lanes 3, total cell protein from cells following induction; lanes 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lanes 5, soluble fraction of cell lysate following induction.

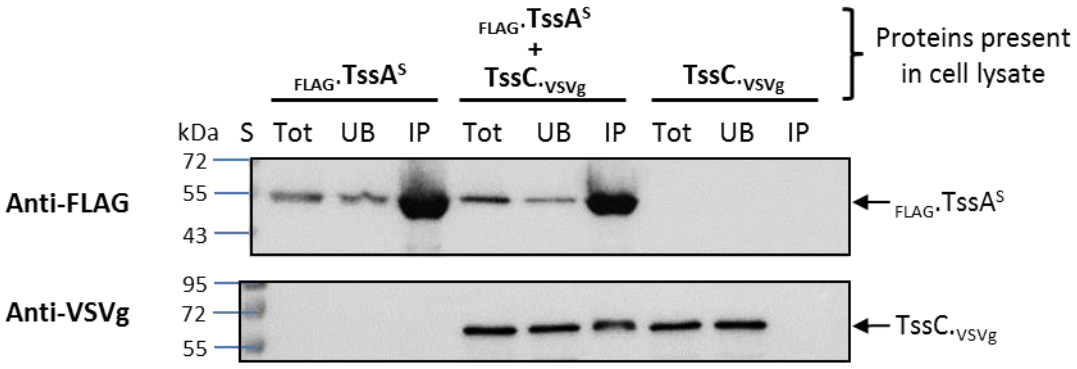


Figure 4.13 Co-IP analysis of the interaction between TssA^S and TssC. $_{FLAG}$.TssA^S and TssC.$_{VSVg}$ were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S}, pACYCDuet-*tssC*.VSVg and pACYCDuet-*tssC*.VSVg-FLAG.*tssA*^{S}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^S, upper panel) and anti-VSVg (TssC.$_{VSVg}$, lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssC (~56 kDa) and TssA^S (~42 kDa).

4.4.1.3 Co-IP analysis of the interaction between $TssA^S$ and TssD

For constructing a plasmid for co-expression of both $TssA^S$ and TssD, *B. cenocepacia tssD* was amplified from strain H111 using a pair of primers (tssD.forpACYC.NtermVSVgTag and tssD.BamHI.rev), in which a VSV-g tag coding sequence was contained in the forward primer. This is due to the BACTH results which showed that the interaction between $TssA^S$ and TssD needed the C-terminus of both proteins to be free. The plasmids for expressing TssD alone and for co-expression of $_{FLAG}TssA^S$ and $_{VSVg}$.TssD were constructed in parallel. The amplified *tssD* DNA fragment of the expected size (~500 bp), and plasmids pACYCDuet-1 and pACYCDuet-FLAG.$tssA^S$ were digested with restriction enzymes *Nco*I and *Bam*HI that recognized sites in the primers and in the first MCS of the plasmids. Ligation was carried out to allow the insertion of the *tssD* DNA fragment into plasmids for constructing pACYCDuet-VSVg.*tssD* (Figure 4.14 A) and pACYCDuet-VSVg.*tssD*-FLAG.$tssA^S$ (Figure 4.14 B). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

Following introduction of pACYCDuet-VSVg.*tssD* into BL21(λDE3) cells, the T7 promoters on the vector were induced with 1 mM IPTG at 37^{o}C. $_{VSVg}$.TssD was overproduced and remained soluble following cell lysis (Figure 4.14 C). The overexpression of $_{FLAG}TssA^S$ and $_{VSVg}$.TssD using pACYCDuet-VSVg.*tssD*-FLAG.$tssA^S$ was carried out in the same strain and induction conditions. Similar amounts of both proteins were produced in high abundance (Figure 4.14 D). However, much less $_{FLAG}TssA^S$ remained soluble under the same induction conditions compared to expressing $_{FLAG}TssA^S$ alone (compare Figure 4.10 with Figure 4.14 D). Co-IP was still carried out as the amount of soluble $_{FLAG}TssA^S$ was considered to be sufficient for conducting the experiment by increasing the amount of binding material. The results showed that VSV-g tagged TssD did not bind to the anti-FLAG coupled beads when it was produced alone. In contrast, when it was co-produced with $_{FLAG}TssA^S$, a small amount of $_{VSVg}$.TssD could be detected in the immunoprecipitated material, indicating a specific interaction between them (Figure 4.15). A protein of ~18 kDa was also detected by the anti-VSVg antibody which is most probably a degradation product of $_{VSVg}$.TssD.

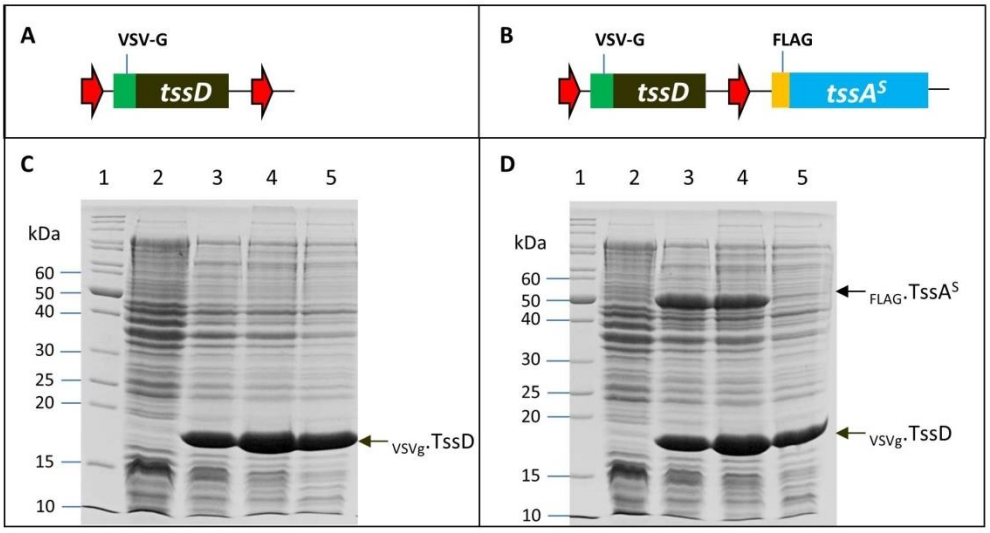


Figure 4.14 Overproduction and solubility of TssD in the absence and presence of with TssA^S. A. Schematic drawing showing the location of VSV-g tagged *tssD* in pACYCDuet-VSVg.*tssD*. The red arrows represent the two T7 promoters. **B.** Schematic drawing of the arrangement of VSV-g tagged *tssD* and FLAG-tagged *tssA*^{S} in pACYCDuet-VSVg.*tssD*-FLAG.*tssA*^{S}. The red arrows represent the two T7 promoters in the plasmid. **C**. A coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssD* in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37^{o}C. The arrow indicates the expected location of TssD based on its size (~20 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **D**. A coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssD*-FLAG.*tssA*^{S} in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37^{o}C. The arrows indicate the expected locations of TssA^S (~42 kDa) and TssD (~20 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

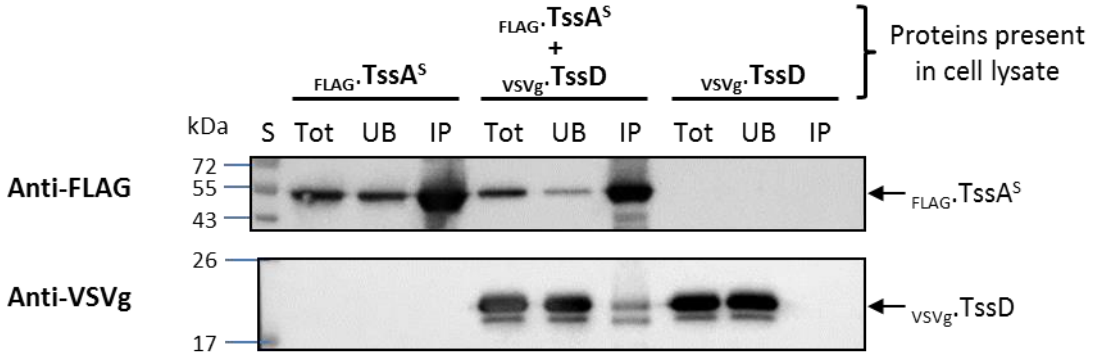


Figure 4.15 Co-IP analysis of the interaction between TssA^S and TssD. $_{FLAG}$.TssA^S and $_{VSVg}$.TssD were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S}, pACYCDuet-VSVg.*tssD* and pACYCDuet-VSVg.*tssD*-FLAG.*tssA*^{S}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^S, upper panel) and anti-VSVg ($_{VSVg}$.TssD, lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^S (~42 kDa) and TssD (~20 kDa).

4.4.1.4 Co-IP analysis of the interaction between $TssA^S$ and TssE

A DNA fragment containing *tssE* was cloned into pACYCDuet plasmid to construct pACYCDuet-His_6.*tssE* for TssE overproduction (Ahmad 2013). Overproduction of TssE was initially carried out by introducing pACYCDuet-His_6.*tssE* into BL21(λDE3) cells. As TssE has been shown to be completely insoluble when overproduced following induction with 1 mM IPTG at 37^{o}C (Ahmad 2013), modified induction conditions (i.e. 0.5 mM IPTG at 22^{o}C) were employed in order to improve the chances of solubilising it. Following induction, a large amount of $_{His6}$.TssE was produced. However, only a small fraction of $_{His6}$.TssE was found to be soluble based on SDS-PAGE analysis (Figure 4.16 A). Therefore, another plasmid, pACYCDuet-*tssE*.VSVg, was constructed by moving the position of the tag to the C-terminus of TssE and replacing it with the VSV-g epitope tag, according to the result of BACTH analysis, which showed $TssA^S$ interacts with TssE when the N-terminus of TssE is free (Section 3.3.5 and Shastri 2011). Following introduction of pACYCDuet-*tssE*.VSVg into BL21(λDE3) cells, the T7 promoters on the vector were induced with 0.5 mM IPTG at 22^{o}C. However, there was no TssE.$_{VSVg}$ overproduced (Figure 4.16 B).

Maltose binding protein (MBP) has been characterised to have excellent solubilisation properties and can be fused to other proteins that would otherwise be insoluble (Fox and Waugh, 2003). The pMAL-c5X plasmid is used for inducible cytoplasmic expression of MBP fusions in which a gene encoding an insoluble protein can be inserted into suitable restriction sites located downstream of the *malE* gene (MBP coding sequence) (Figure 4.17). The resulting fusion protein contains a poly-asparagine linker and a factor Xa cleavage site located between MBP and the fused insoluble protein. Fusion proteins are expressed from the strong *tac* promoter which is under control of the Lac repressor encoded by pMALc5X. Therefore, it can be used in a wide range of *E. coli* host strains. The vector was modified by introducing a hexa-histidine tag coding sequence into the N-terminal coding region of the *malE* gene (between codons 3 and 4) for overproducing $_{His6}$MBP fusions. The modified vector is named pMAL-c5X-His_6 (Mosby and Thomas, unpublished results) (Figure 4.17). pMAL-c5X-His_6 (ColE1 ori) is compatible with pACYCDuet-FLAG.*tssA*^{S} (p15A ori).

PCR amplification was previously carried out to obtain a *tssE* product with *Nde*I and *Bam*HI restriction sites that were inserted into pMAL-c5X-His_6 to generate

pMAL-c5X-His_6-*tssE* (Mosby and Thomas, unpublished results) (Figure 4.18 A). Following IPTG induction of the *tac* promoter on pMAL-c5X-His_6-*tssE* in BL21(λDE3) cells growing at at 30°C, $_{His6}$.MBP-TssE was produced in an abundant amount and remained in the soluble fraction following cell lysis (Figure 4.18 C). Co-expression of $_{His6}$.MBP-TssE and $_{FLAG}$.TssA^S was analysed by introducing pMAL-c5X-His_6-*tssE* and pACYCDuet-FLAG.*tssA*^{S} together into BL21(λDE3) cells (Figure 4.18 B). Following induction with IPTG, $_{His6}$.MBP-TssE was overproduced in a larger amount compared to $_{FLAG}$.TssA^S, and both proteins remained soluble (Figure 4.18 D). Overexpression of $_{FLAG}$.TssA^S in the absence of $_{His6}$.MBP-TssE is shown in Figure 4.10 (Section 4.4.1.1).

Co-IP was carried out for the interaction analysis between $_{FLAG}$.TssA^S and $_{His6}$.MBP-TssE. The results showed that $_{His6}$.MBP-TssE did not bind to the anti-FLAG coupled beads when it was produced alone. In contrast, when it was co-produced with $_{FLAG}$TssA^S, $_{His6}$.MBP-TssE was detected in the immunoprecipitated material (Figure 4.19).

To rule out the possibility that the interaction between $_{FLAG}$.TssA^S and $_{His6}$.MBP-TssE was a result of interaction between $_{FLAG}$.TssA^S and $_{His6}$.MBP, it was decided to carry out a co-IP experiment between $_{FLAG}$.TssA^S and $_{His6}$.MBP. Expression of both N-terminal FLAG-tagged TssA^S and Histidine tagged MBP proteins was analysed by introducing pMAL-c5X-His_6 and pACYCDuet-FLAG.*tssA*^{S} separately and together into BL21(λDE3) cells (Figure 4.20 A&B). Following induction with IPTG, both proteins were found to be overproduced and soluble, although the relative amount of $_{FLAG}$.TssA^S that was produced decreased in cells with pMAL-c5X-His_6 (Figures 4.20 D and 4.10). $_{His6}$.MBP was produced in abundant amounts in cells containing pMAL-c5X-His_6 alone (Figure 4.20 C) and with $_{FLAG}$.TssA^S (Figure 4.20 D). The soluble fraction of the cell lysate was then analysed by co-IP using anti-FLAG coupled beads. The results showed that $_{FLAG}$.TssA^S was captured by the resin as expected, and the $_{His6}$.MBP control did not show any non-specific binding to the resin (Figure 4.21). In the immuno-precipitated material from the cell lysate containing both $_{His6}$.MBP and $_{FLAG}$.TssA^S, $_{His6}$.MBP was not co-precipitated with $_{FLAG}$.TssA^S, suggesting $_{FLAG}$.TssA^S does not interact with $_{His6}$.MBP. Therefore, the co-IP result for the interaction between $_{FLAG}$.TssA^S and $_{His6}$.MBP-TssE indicated a specific interaction between the two proteins.

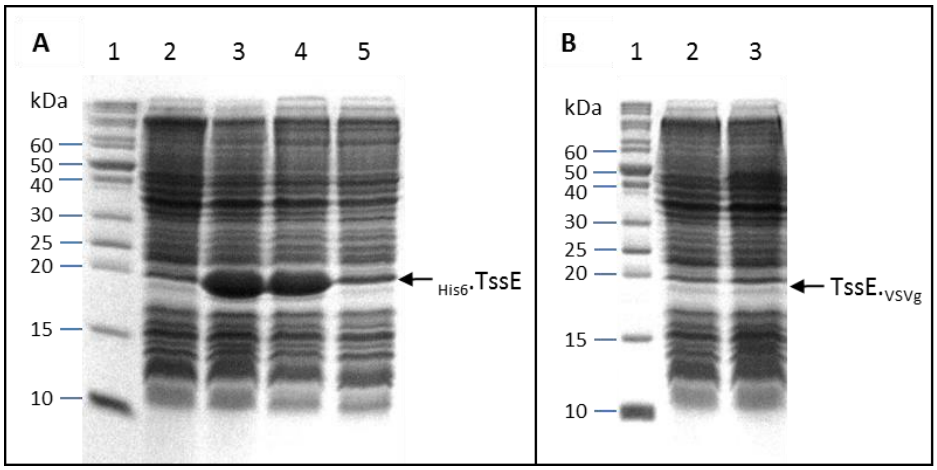


Figure 4.16 Overproduction and solubility of TssE. A. A Coomassie blue-stained 15% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-His_6.*tssE* in *E. coli* strain BL21(λDE3) with 0.5 mM IPTG at 22°C. Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B.** A Coomassie blue-stained 15% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-*tssE*.VSVg in *E. coli* strain BL21(λDE3) with 0.5 mM IPTG at 22°C. Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction. The arrows in A and B indicate the expected locations of TssE based on its size (~18 kDa).

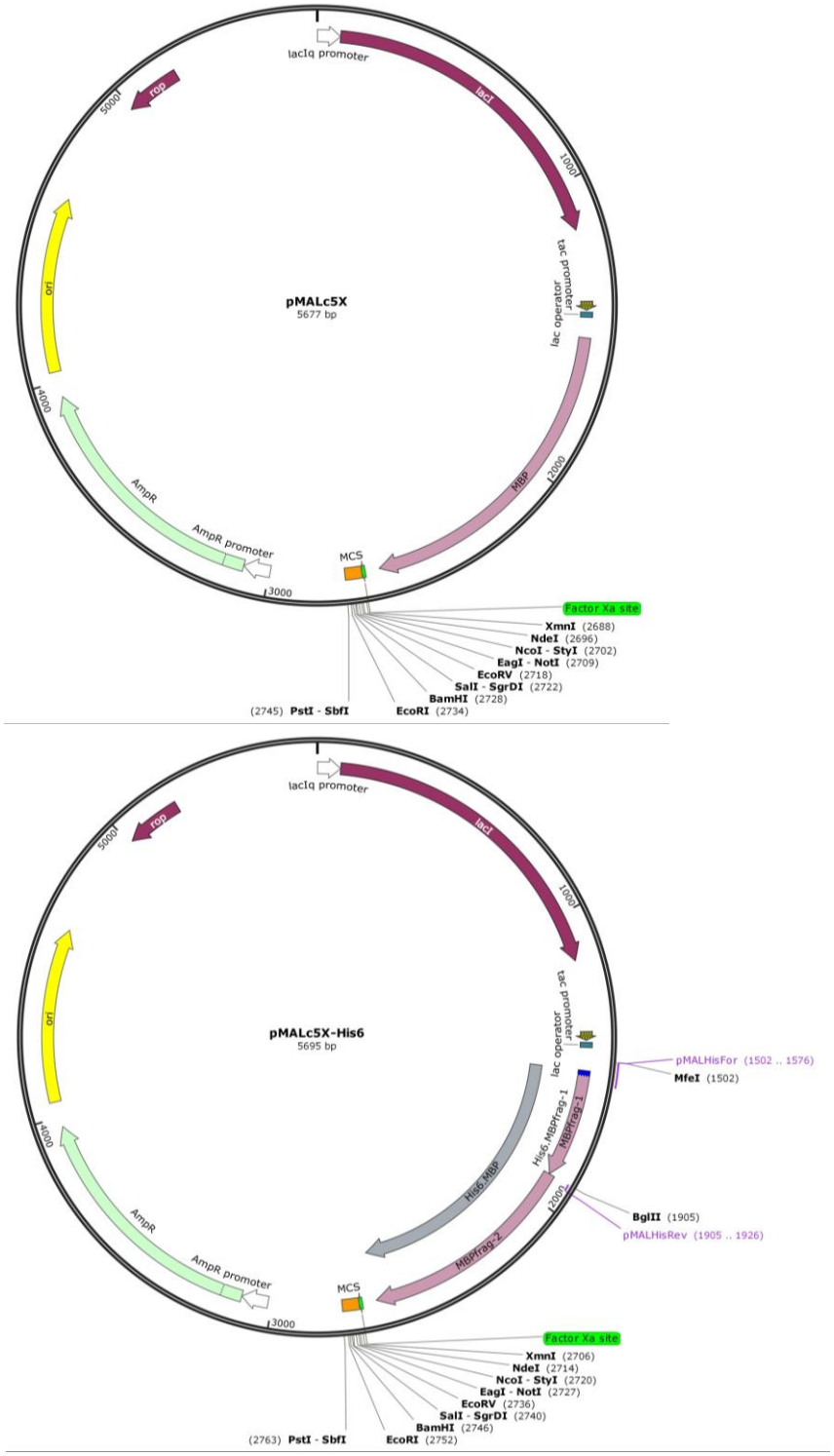


Figure 4.17 Schematic representations of pMALc5X and pMALc5X-His$_6$. Top, This vector contains the *malE* gene of *E. coli* that encodes maltose binding protein (MBP) with an exact deletion of the *malE* signal peptide coding sequence. This vector contains the *tac* promoter. The gene of interest is cloned downstream and in frame with of the *malE* gene where the MCS is located, resulting in the expression of an MBP fusion protein. **Bottom**, pMALc5X-His$_6$ vector that was derived from pMALc5X. A pair of primers (pMALHisFor and pMALHisRev) were used for amplifying a DNA fragment containing the MBP N-terminal coding sequence with a hexa-histidine tag coding sequence included in the forward primer. The amplified MBP fragment was digested with *Mfe*I and *Bgl*II, sites for which were included in the forward and reverse primers, respectively, and in the pMALc5X plasmid MCS. This vector has the same features as pMALc5X except that it codes for a hexa-histidine tagged MBP fusion protein (Mosby and Thomas, unpublished results). Ap^R, ampicillin-resistance gene (*bla*).

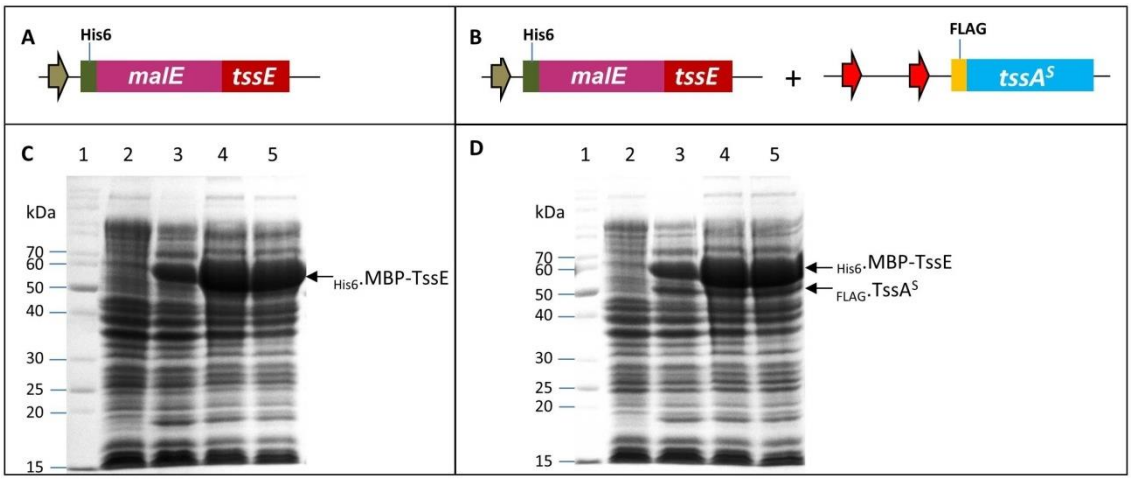


Figure 4.18 Overproduction and solubility of $_{His6}$.MBP-TssE in the absence and presence of TssA^S. A. Schematic drawing showing the location of His$_6$.*malE-tssE* in pMAL-c5X-His$_6$-*tssE*. The grey arrow represents the *tac* promoter. **B**. Schematic drawing of compatible pMAL-c5X-His$_6$-*tssE* and pACYCDuet-FLAG.*tssA*^{S} used for co-expression of $_{His6}$.MBP-TssE and $_{FLAG}$.TssA^S. **C**. A coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pMAL-c5X-His$_6$-*tssE* in *E. coli* strain BL21(λDE3) with 0.5 mM IPTG at 30^{o}C. The arrow indicates the expected location of $_{His6}$.MBP-TssE based on its size (~62 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **D**. A coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following co-expression of pMAL-c5X-His$_6$-*tssE* and pACYCDuet-FLAG.*tssA*^{S} in *E. coli* strain BL21(λDE3) following induction with 0.5 mM IPTG at 30^{o}C. The arrows indicate the expected locations of TssA^S (~42 kDa) and $_{His6}$.MBP-TssE (~62 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

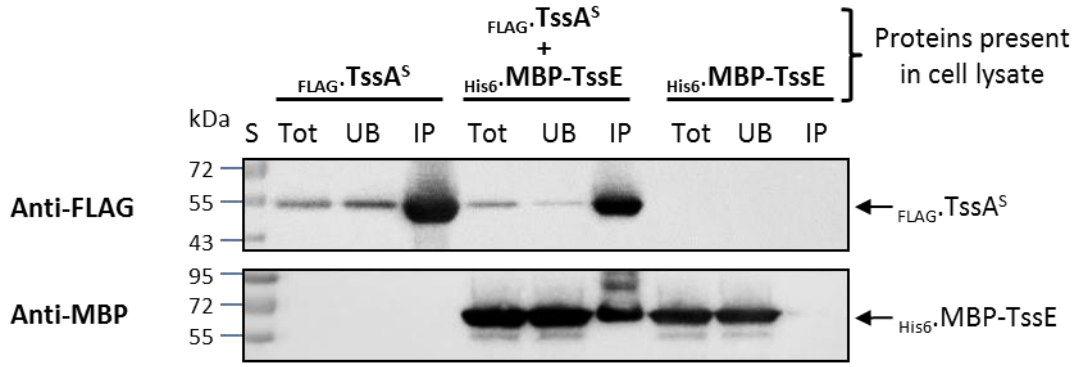


Figure 4.19 Co-IP analysis of the interaction between TssA^S and $_{His6}$.MBP-TssE. $_{FLAG}$.TssA^S and $_{His6}$.MBP-TssE were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S} and pMAL-c5X-His$_6$-*tssE* separately and together, accordingly, and induction with IPTG. Soluble fraction of the cell following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^S, upper panel) and anti-MBP ($_{His6}$.MBP-TssE, lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^S (~42 kDa) and $_{His6}$.MBP-TssE (~62 kDa).

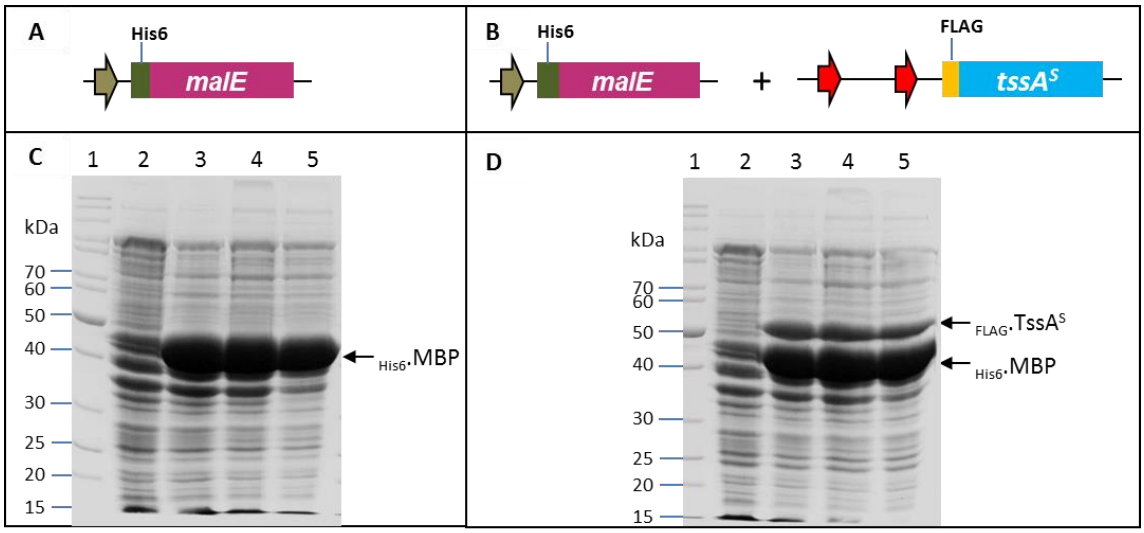


Figure 4.20 Overproduction and solubility of $_{His6}$.MBP in the absence and presence of TssA^S. A. Schematic drawing showing the location of His$_6$.*malE* in pMAL-c5X-His$_6$. The grey arrow represents the *tac* promoter. **B**. Schematic drawing of compatible pMAL-c5X-His$_6$ and pACYCDuet-FLAG.*tssA*^{S} used for co-expression of $_{His6}$.MBP and $_{FLAG}$.TssA^S. **C**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pMAL-c5X-His$_6$ in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37^{o}C. The arrow indicates the expected location of $_{His6}$.MBP based on its size (~45 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **D**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following co-expression of pMAL-c5X-His$_6$ and pACYCDuet-FLAG.*tssA*^{S} in *E. coli* strain BL21(λDE3) following induction with 1 mM IPTG at 37^{o}C. The arrows indicate the expected locations of TssA^S (~42 kDa) and $_{His6}$.MBP (~45 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

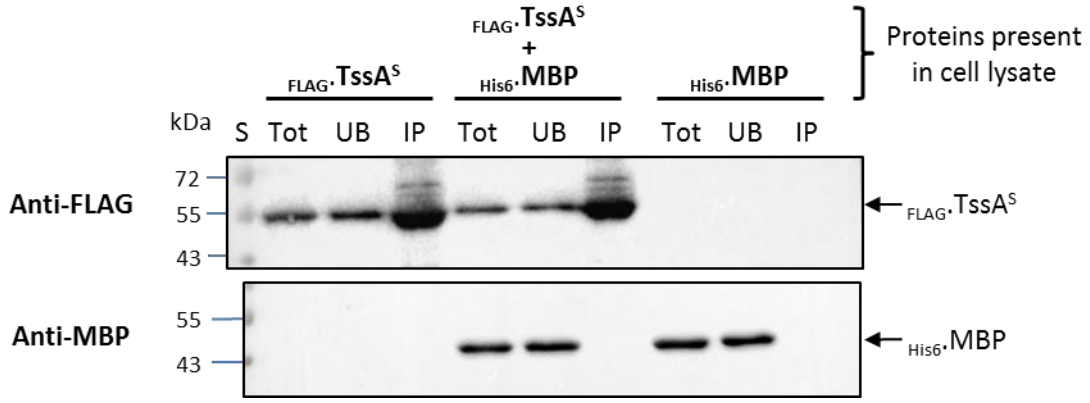


Figure 4.21 Co-IP analysis of the interaction between TssA^S and MBP. $_{FLAG}$.TssA^S and $_{His6}$.MBP were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S} and pMAL-c5X-His$_6$ separately and together, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^S, upper panel) and anti-MBP ($_{His6}$.MBP, lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expect locations of TssA^S (~42 kDa) and $_{His6}$.MBP (~45 kDa).

4.4.1.5 Co-IP analysis of the interaction between $TssA^S$ and TssF in the presence and absence of TssG

As the BACTH results suggested that the interaction between $TssA^S$ and TssF was stable when the N-terminus of TssF was free compared to when the C-terminus of TssF was free where the β-galactosidase activities were highly variable (Section 3.3.6), it was decided to place a HA tag at the C-terminus of TssF. The plasmids for expressing $TssF._{HA}$ alone and for co-expression of $TssA^S$ and $TssF._{HA}$ were constructed in parallel. *tssF* was amplified using primers tssF.NcoI.for and tssF.HA.HindIII.rev, in which a HA tag coding sequence was contained in the reverse primer. The amplified *tssF* DNA fragment of the expected size (~1.8 kb), pACYCDuet-1 and pACYCDuet-FLAG.$tssA^S$ plasmids were digested with restriction enzymes *Nco*I and *Hin*dIII that recognized sites in the primers and the MCS of the plasmids. Ligation was carried out to allow the insertion of the *tssF* DNA fragment into the plasmids downstream of the first T7 promoter for constructing pACYCDuet-*tssF*.HA and pACYCDuet-*tssF*.HA-FLAG.$tssA^S$, respectively (Figure 4.22 A and B). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

Protein overexpression was carried out in *E. coli* strain BL21(λDE3) cells containing pACYCDuet-*tssF*.HA and pACYCDuet-*tssF*.HA-FLAG.$tssA^S$. A small amount of $TssF._{HA}$ was overproduced in cells containing pACYCDuet-*tssF*.HA, and a small fraction of the overproduced protein remained soluble following cell lysis (Figure 4.22 C). Similar amounts of $TssA^S$ and $TssF._{HA}$ were overproduced in cells containing pACYCDuet-*tssF*.HA-FLAG.$tssA^S$. $TssA^S$ was found to be largely soluble. However, as observed previously for overproducing $TssF._{HA}$ alone, only a small fraction of $TssF._{HA}$ was soluble (Figure 4.22 D). Overexpression of $_{FLAG}.TssA^S$ in the absence of TssF is shown in Figure 4.10 (Section 4.4.1.1). The co-IP was carried out using anti-FLAG coupled beads, and the results showed that $TssF._{HA}$ did not bind to the anti-FLAG coupled beads when it was produced alone. In contrast, when it was co-produced with $_{FLAG}.TssA^S$, $TssF._{HA}$ was co-precipitated by anti-FLAG antibody. Therefore, the co-IP result for the interaction between $_{FLAG}.TssA^S$ and $TssF._{HA}$ indicated a specific interaction between them (Figure 4.23).

Although an interaction between TssF and TssG has not been observed in our group using the BACTH assay, these proteins have been found to interact and stabilize each

other by other workers (English et al. 2014; Brunet et al. 2015). As $TssA^S$ showed a specific interaction with $TssF._{HA}$, it was decided to analyse the interaction between $TssA^S$ and TssF in the presence of TssG by co-IP to investigate whether the TssF-TssG interaction could be detected.

Co-expression of $TssA^S$, TssF and TssG was conducted using two compatible plasmids, i.e. pACYCDuet and pETDuet, although when used together previously the expression from pETDuet plasmid dominated (Ahmad, Sun and Thomas, unpublished results). The HA tag was placed at the N-terminus of TssF to make the C-terminus available to interact with TssG as demonstrated by Brunet et al. (2015) and English et al. (2014). TssG was tagged with a hexa-histidine tag at its C-terminus. There is a natural Shine-Dalgarno sequence located upstream of the *tssG* gene for its translation and the stop codon for *tssF* overlaps the *tssG* open reading frame. For this reason, *tssF* and *tssG* were amplified together on the same DNA fragment from the *B. cenocepacia* H111 genome with the pair of primers tssF.NdeI.HA.for and tssG.Kpn.C-His.rev in which *Nde*I and *Kpn*I restriction sites were contained in the forward and reverse primers, respectively. The amplified *tssF-tssG* DNA fragment of the expected size (~2.9 kb) and pETDuet∆0-1 were digested with restriction enzymes *Nde*I and *Kpn*I that recognized sites in the primers and the second MCS of the plasmid. pETDuet∆0-1 lacks the *lac* operator sequences located downstream of each T7 promotor (Sun and Thomas, unpublished results). Ligation was carried out to allow the insertion of the *tssF-tssG* DNA fragment into the plasmid for constructing pETDuet∆0-HA.*tssF-tssG*.His_6. *E. coli* strain MC1061 was transformed with the ligation mixture and colonies were grown on LB plates containing ampcillin. The correct recombinant plasmids were identified by PCR screening, followed by plasmid miniprep and nucleotide sequence determination to confirm the plasmid contains the *tssF-tssG* insert.

Overexpression of TssF and TssG was perfomed in cells containing pETDuet∆0-HA.*tssF-tssG*.His_6 following induction of the T7 promotors of the plasmid with IPTG (Figure 4.24 A). There was a small amount of $_{HA}$.TssF overproduced, whereas no visible amount of $TssG._{His6}$ could be observed by gractionation of cell lysates in an SDS-PA gel (Figure 4.24 C). A small fraction of overproduced $_{HA}$.TssF remained in the soluble fraction of the cell lysate. To co-produce these three tagged proteins, *E. coli* strain BL21(λDE3) cells were co-transformed with

pETDuetΔ0-HA.*tssF-tssG*.His_6 and pACYCDuet-FLAG.*tssA*^{S} (Figure 4.24 B). Induction was carried out at both 37^{o}C and 30^{o}C, and both induction conditions gave the same result. The results showed that $_{FLAG}$.TssA^S was overexpressed and soluble. There was some leaky expression of $_{HA}$.TssF in the absence of induction presumably because the *lac* operators on pETDuet-1 were deleted. Following induction, $_{HA}$.TssF was overproduced and soluble. However, TssG.$_{His6}$ was produced in a smaller amount and was insoluble as determined by immuno-detection with anti-His_6 antibody (Figure 4.24 D). Co-IP of the interaction between $_{FLAG}$.TssA^S and $_{HA}$.TssF was still carried out without soluble TssG.$_{His6}$ present. The results showed that the N-terminal HA tagged TssF did not interact with $_{FLAG}$.TssA^S, as it was not recovered with $_{FLAG}$.TssA^S in the immuno-precipitated material by anti-FLAG antibody (Figure 4.25).

To rule out the possible interference of TssG with the interaction between TssA^S and $_{HA}$.TssF, pETDuetΔ0-HA.*tssF* was constructed by amplifying *tssF* with the tssF.NdeI.HA.for and Cterminal-iotFrev primers, in which a HA tag coding sequence was contained in the forward primer. The amplified *tssF* DNA fragment of the expected size (~1.8 kb) and pETDuetΔ0-1 plasmid were digested with restriction enzymes *Nde*I and *Acc*65I that recognized sites in the primers and the second MCS of the plasmid. Ligation was carried out to allow the insertion of the *tssF* DNA fragment into the plasmid downstream of the second T7 promotor (Figure 4.26 A). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

Overexpression of $_{HA}$.TssF was performed in cells containing pETDuetΔ0-HA.*tssF* following induction of the T7 promotors on the plasmid with IPTG (Figure 4.26 C). Overexpression of TssA^S and $_{HA}$.TssF was carried out in *E. coli* strain BL21(λDE3) cells containing both pETDuetΔ0-HA.*tssF* and pACYCDuet-FLAG.*tssA*^{S} by inducing the T7 promotor with 1 mM IPTG at 30^{o}C (Figure 4.26 B). Under this condition, $_{HA}$.TssF was overproduced to a similar degree as in cells containing only pETDuetΔ0-HA.*tssF*. However, no visible amount of TssA^S was observed on a Coomassie blue-stained SDS-PA gel (Figure 4.26 C). Co-IP was still carried out as it was expected that some TssA^S would be present. The western blotting confirmed the presence of TssA^S in the cell lysate following induction. However, the co-IP results were the same as observed when TssG was also present, i.e. it did not demonstrate an

interaction between TssA^{S} and $_{HA}$.TssF (Figure 4.27). Therefore, the results suggested that the N-terminus of TssF is involved in the interaction with TssA^{S}.

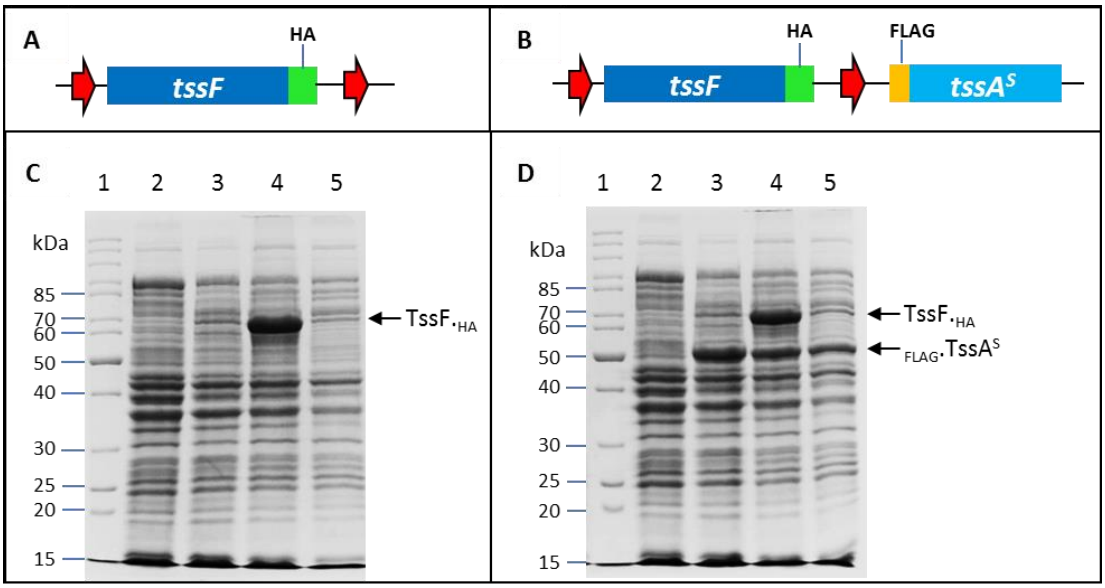


Figure 4.22 Overproduction and solubility of C-terminal HA-tagged TssF in the absence and presence of TssA^S. A. Schematic drawing showing the location of HA tagged *tssF* in pACYCDuet-*tssF*.HA. The red arrows represent the two T7 promoters present in the plasmid. **B.** Schematic drawing of the arrangement of HA tagged *tssF* and FLAG.*tssA*^{S} in pACYCDuet-*tssF*.HA-FLAG.*tssA*^{S}. The red arrows represent the two T7 promoters present in the plasmid. **C**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-*tssF*.HA in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 30^{o}C. The arrow indicates the expected location of TssF based on its size (~70 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **D**. A Coomassie blue-stained 10% SDS-PA gel showing the induction of pACYCDuet-*tssF*.HA-FLAG.*tssA*^{S} in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 30^{o}C. The arrows indicate the expected locations of TssA^S (~42 kDa) and TssF (~70 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

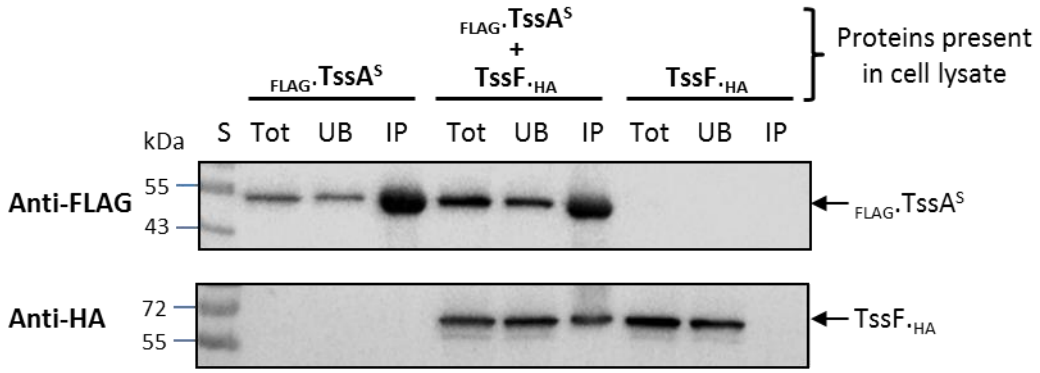


Figure 4.23 Co-IP analysis of the interaction between TssA^S and TssF.$_{HA}$. $_{FLAG}$.TssA^S and TssF.$_{HA}$ were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S}, pACYCDuet-*tssF*.HA and pACYCDuet-*tssF*.HA-FLAG.*tssA*^{S}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^S, upper panel) and anti-HA (TssF.$_{HA}$ lower panel) monoclonal antibodies monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssF (~70 kDa) and TssA^S (~42 kDa).

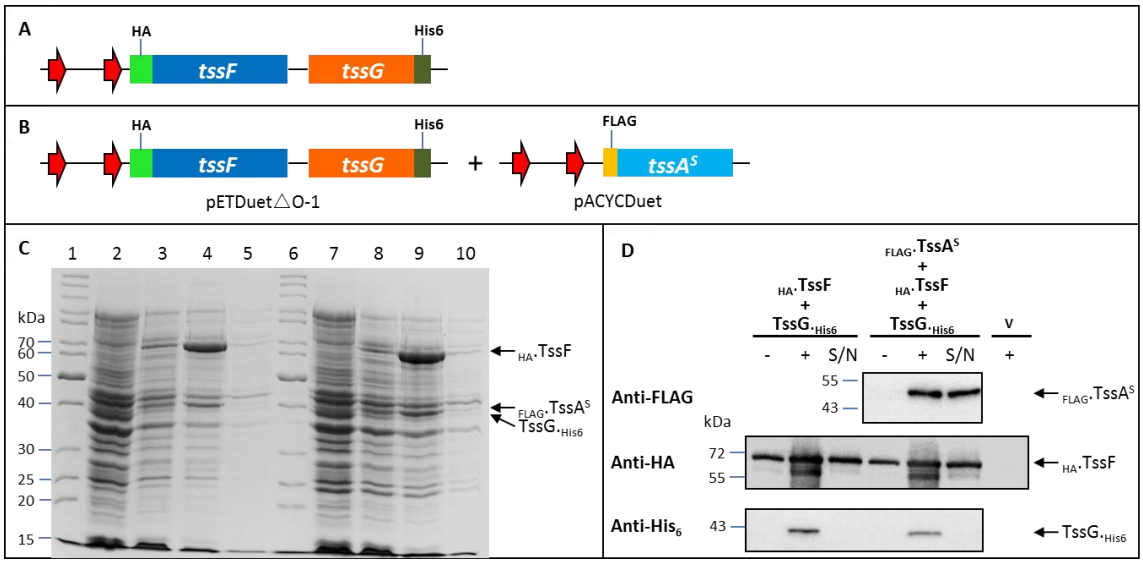


Figure 4.24 Co-overproduction and solubility of TssF and TssG in the presence and absence of TssA^S. A. Schematic drawing of the arrangement of N-terminal HA tagged *tssF* and C-terminal histidine tagged *tssG* in pETDuet∆O-HA.*tssF-tssG*.His$_6$. The red arrows represent the two T7 promoters present in the plasmid. **B.** Schematic drawing of compatible pETDuet∆O-HA.*tssF-tssG*.His$_6$ and pACYCDuet-FLAG.*tssA*^{S} used for co-expression of $_{FLAG}$.TssA^S, $_{HA}$.TssF and TssG.$_{His6}$. **C**. A Coomassie blue-stained 10% SDS-PA gel showing the induction of pETDuet∆O-HA.*tssF-tssG*.His$_6$, and co-expresssion of pETDuet∆O-HA.*tssF-tssG*.His$_6$ and pACYCDuet-FLAG.*tssA*^{S} plasmids in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 30^{o}C. The arrows indicate the expected locations of TssA^S (~42 kDa), TssF (~70 kDa) and TssG (~41 kDa). Lanes 1 and 6, EZ-Run^{TM} Rec protein ladder (Fisher); Lanes 2-5, induction of pETDuet∆O-HA.*tssF-tssG*.His$_6$ plasmid: lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. Lanes 7-10, pETDuet∆O-HA.*tssF-tssG*.His$_6$ and pACYCDuet-FLAG.*tssA*^{S} co-expression: lane 7, total cell protein from uninduced cells; lane 8, total cell protein from cells following induction; lane 9, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 10, soluble fraction of cell lysate following induction. **D**. Immuno-detection of co-expression of $_{FLAG}$.TssA^S, $_{HA}$.TssF and TssG.$_{His6}$ in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG induction at 30^{o}C. Cell lysate was fractionated by centrifugation, proteins were electrophoresed by SDS-PAGE and, following western transfer, were probed with anti-FLAG ($_{FLAG}$.TssA^S, upper panel), anti-HA ($_{HA}$.TssF middle panel) and anti-His$_6$ (TssG.$_{His6}$, lower panel) monoclonal antibodies. The arrows indicate the expected locations of TssA^S (~42 kDa), TssF (~70 kDa) and TssG (~41 kDa). '-', total cell protein from uninduced cells; '+', total cell protein from cells following induction; 'S/N', soluble fraction of cell lysate following induction. Proteins present in the cell lysates are as indicated; 'V', total cell protein from cells following induction of pETDuet∆O-1 plasmid (control).

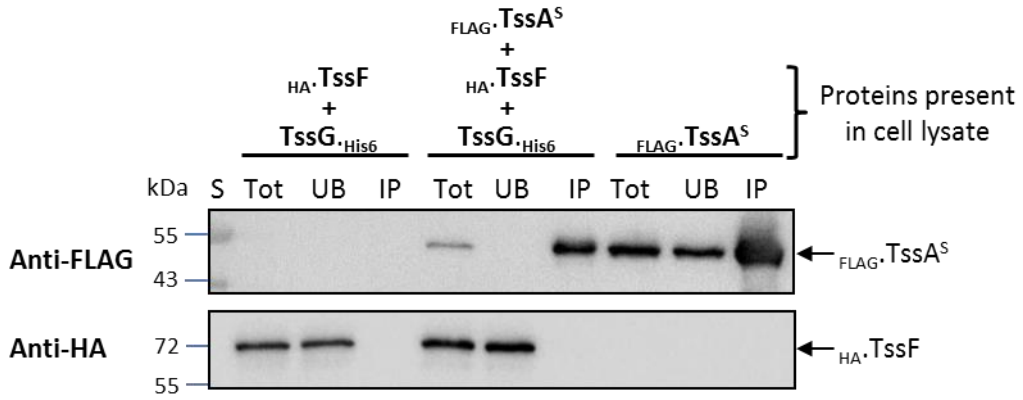


Figure 4.25 Co-IP analysis of the interaction between TssA^S and $_{HA}$.TssF in the presence of TssG. $_{FLAG}$.TssA^S, $_{HA}$.TssF and TssG were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S} and pETDuetΔO-HA.*tssF-tssG*.His$_6$ separately and together, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^S, upper panel) and anti-HA ($_{HA}$.TssF lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssF (~70 kDa) and TssA^S (~42 kDa).

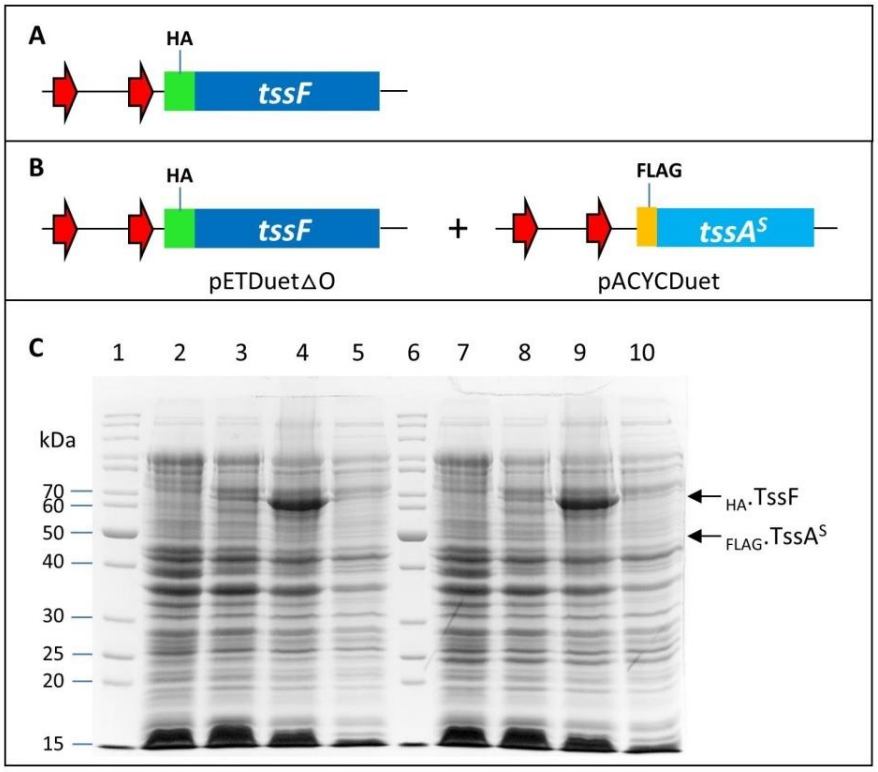


Figure 4.26 Co-overexpression and solubility of N-terminal HA-tagged TssF in the presence and absence of TssA^S. A. Schematic drawing showing the location of N-terminal HA tagged *tssF* in pETDuetΔO-HA.*tssF*. The red arrows represent the two T7 promoters present in the plasmid. **B.** Schematic drawing of compatible plasmids pETDuetΔO-HA.*tssF* and pACYCDuet-FLAG.*tssA*^{S} used for co-expression of $_{FLAG}$.TssA and $_{HA}$.TssF. **C**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pETDuetΔO-HA.*tssF* on its own, or with pACYCDuet-FLAG.*tssA*^{S} in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 30^{o}C. The arrows indicate the expected locations of TssF (~70 kDa) and TssA^S (~42 kDa). Lanes 1 and 6, EZ-Run^{TM} Rec protein ladder (Fisher); Lanes 2-5, induction of pETDuetΔO-HA.*tssF* plasmid: lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. Lanes 7-10, co-expression of pETDuetΔO-HA.*tssF* and pACYCDuet-FLAG.*tssA*^{S}: lane 7, total cell protein from uninduced cells; lane 8, total cell protein from cells following induction; lane 9, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 10, soluble fraction of cell lysate following induction.

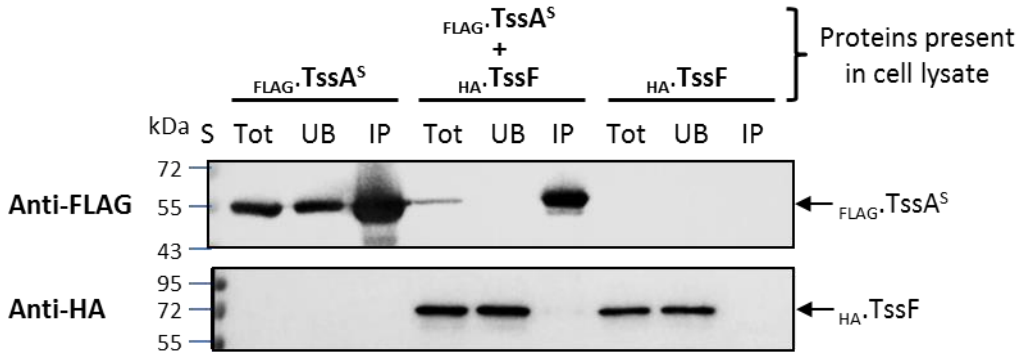


Figure 4.27 Co-IP analysis of the interaction between TssA^S and $_{HA}$.TssF in the absence of TssG. $_{FLAG}$.TssA^S and $_{HA}$.TssF were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S} and pETDuetΔO-HA.*tssF* separately and together, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^S, upper panel) and anti-HA ($_{HA}$.TssF lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssF (~70 kDa) and TssA^S (~42 kDa).

4.4.1.6 Co-IP analysis of the interaction between $TssA^S$ and $TssI_{gp27gp5}$

TssI forms a trimer, and it shares structural homology with the $(gp27)_3$-$(gp5)_3$ complex that makes up the bacteriophage T4 tail spike (Kanamaru et al. 2002; Pukatzki et al. 2007; Leiman et al. 2009). Multiple sequence alignment of TssI protein sequences from various Gram-negative bacterial species (Section 3.3.9) indicates the core regions of TssI that correspond to the bacteriophage T4 gp5 and gp27 proteins, in agreement with the study performed by Leiman (Leiman et al. 2009). Using the BACTH system, the fused gp27-gp5 region of TssI from strain 715j that corresponds to BCAM0148 from strain J2315 has been shown to be responsible for the interaction between $TssA^S$ and TssI (Section 3.3.9). Due to degradation of the full-length TssI protein that was observed previously during protein overproduction experiments in *E. coli* (result not shown), it was decided to use the core gp27gp5-like region of TssI in the co-IP assay with $TssA^S$ for providing the *in vitro* evidence for the interaction between $TssA^S$ and TssI.

B. cenocepacia $tssI_{gp27gp5}$ (encoding the gp27gp5-like core region of *B. cenocepacia* 715j orthologue of TssI BCAM0148) was amplified using a pair of primers, TssI.BspHI.VSVgtag.For and BCAM0148gp5.BglII.rev, in which a VSV-g tag coding sequence was contained in the forward primer. The amplified $tssI_{gp27gp5}$ DNA fragment of the expected size (~1.6 kb) was digested with restriction enzymes *Bsp*HI and *Bgl*II that recognized sites in the primers and ligated into *Nco*I and *Bam*HI digested pACYCDuet-1 DNA in order to construct pACYCDuet-VSVg.$tssI_{gp27gp5}$ in which $tssI_{gp27gp5}$ was located downstream of the first T7 promoter (Figure 4.28 A). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

To co-express of $TssA^S$ and $TssI_{gp27gp5}$, pACYCDuet-VSVg.$tssI_{gp27gp5}$-FLAG.$tssA^S$ was constructed. $tssA^S$ was amplified using primers TssA.NdeI.FLAG.for and TssA.BglII.rev, in which a FLAG tag coding sequence was contained in the forward primer. The amplified $tssA^S$ DNA fragment of the expected size (~1.1 kb) and pACYCDuet-VSVg.$tssI_{gp27gp5}$ were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites in the primers and the plasmid MCS. Ligation was carried out to allow insertion of the $tssA^S$ DNA fragment downstream of the second T7 promoter in the plasmid for constructing pACYCDuet-VSVg.$tssI_{gp27gp5}$-FLAG.$tssA^S$ (Figure 4.28 B).

The procedures for identifying the desired recombinant plasmid were as described above for constructing pACYCDuet-VSVg.*tssI*$_{gp27gp5}$.

Induction of the T7 promoters on pACYCDuet-VSVg.*tssI*$_{gp27gp5}$ and pACYCDuet-VSVg.*tssI*$_{gp27gp5}$-FLAG.*tssA*^{S} was conducted in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at both 30^{o}C and 37^{o}C. The chosen induction condition for the co-IP analysis was based on the large amount of soluble TssI$_{gp27gp5}$ that was produced at 30^{o}C (Figure 4.28 C). When TssI$_{gp27gp5}$ was co-expressed with TssA^{S} at 30^{o}C, there was less TssA^{S} overproduced compared with induction at 37^{o}C. When co-expressed at 37^{o}C, both proteins were overproduced in a reasonable amount with around 50% of TssI$_{gp27gp5}$ in the soluble fraction, and TssA^{S} remained soluble (Figure 4.28 D). Overproduction of $_{FLAG}$.TssA^{S} in the absence of TssI$_{gp27gp5}$ is shown in Figure 4.10 (Section 4.4.1.1).

In a preliminary co-IP assay with anti-FLAG coupled beads, TssI$_{gp27gp5}$ was non-specifically bound to the beads in TBS buffer containing 150 mM or 500 mM NaCl with 0.2% Tween 20 (results not shown). Therefore, a series of binding buffers with increasing concentration of Tween 20 from 0.5% to 2% in TBS was tested in order to identify suitable buffer conditions for solving the non-specific binding problem of TssI$_{gp27gp5}$. Moreover, β-mercaptoethanol was omitted from the buffer in the elution step in one of the buffer conditions. The co-IP assay was carried out as described in Section 2.7.4.2. Interestingly, it was observed that with increasing concentration of Tween 20 from 0.5% to 2%, the amount of non-specific binding of TssI$_{gp27gp5}$ to the anti-FLAG beads also increased. Tween 20 is a non-ionic detergent that decreases non-specific hydrophobic interactions that may occur in the co-IP by breaking protein-lipid associations. The observation that Tween 20 promotes binding of TssI$_{gp27gp5}$ to the anti-FLAG resin suggests the non-specific binding is not based on hydrophobic interaction, but rather it may be an ionic interaction.

In contrast, elution without β-mercaptoethanol in the buffer helped to decrease the amount of non-specific binding of TssI$_{gp27gp5}$ (Figure 4.29 A). Based on this observation, it was decided to perform the co-IP with or without the presence of 0.2% Tween 20, and in both conditions β-mercaptoethanol was omitted from the elution step. The result showed that there was no TssI$_{gp27gp5}$ non-specifically bound to the beads under both conditions, suggesting that the presence of β-mercaptoethanol in the elution step is related to the control problem of TssI$_{gp27gp5}$ (Figure 4.29 B). After identifying conditions

that prevented non-specific binding of $TssI_{gp27gp5}$ to the anti-FLAG resin, the co-IP of $TssA^S$ and $TssI_{gp27gp5}$ was performed using binding buffer without Tween 20 and without β-mercaptoethanol in the elution buffer. The results showed that whereas $TssI_{gp27gp5}$ did not bind to the beads non-specifically, when it was co-expressed with $TssA^S$, it was present in the immuno-precipitated material in a small amount. This result suggested $TssI_{gp27gp5}$ interacts with $TssA^S$ (Figure 4.30).

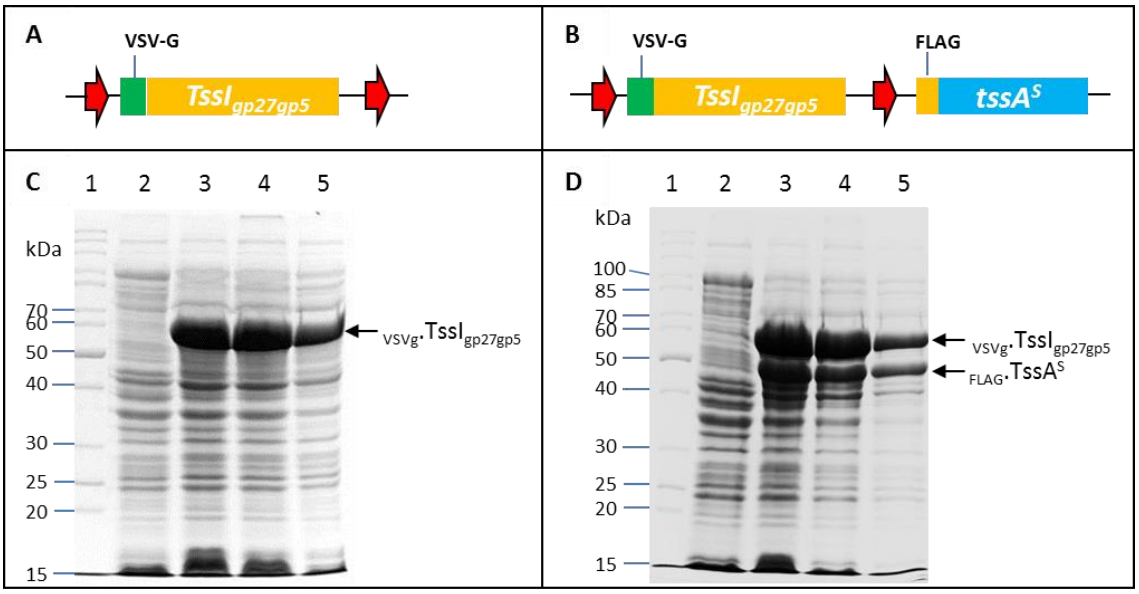


Figure 4.28 Overproduction and solubility of $TssI_{gp27gp5}$ in the presence and absence of $TssA^S$. A. Schematic drawing showing the location of VSV-g tagged *tssI.gp27gp5* in pACYCDuet-VSVg.*tssI.gp27gp5*. The red arrows represent the two T7 promoters present in the plasmid. **B.** Schematic drawing of the arrangement of VSV-g tagged *tssI.gp27gp5* and FLAG.*tssA*^{S} in pACYCDuet-VSVg.*tssI.gp27gp5*-FLAG.*tssA*^{S} plasmid. The red arrows represent the two T7 promoters present in the plasmid. **C**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssI.gp27gp5* in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 30^{o}C. The arrow indicates the expected location of based on its size $TssI_{gp27gp5}$ (~61 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **D**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssI.gp27gp5*-FLAG.*tssA*^{S} plasmid in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 37^{o}C. The arrows indicate the expected locations of $TssA^S$ (~42 kDa) and $TssI_{gp27gp5}$ (~61 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

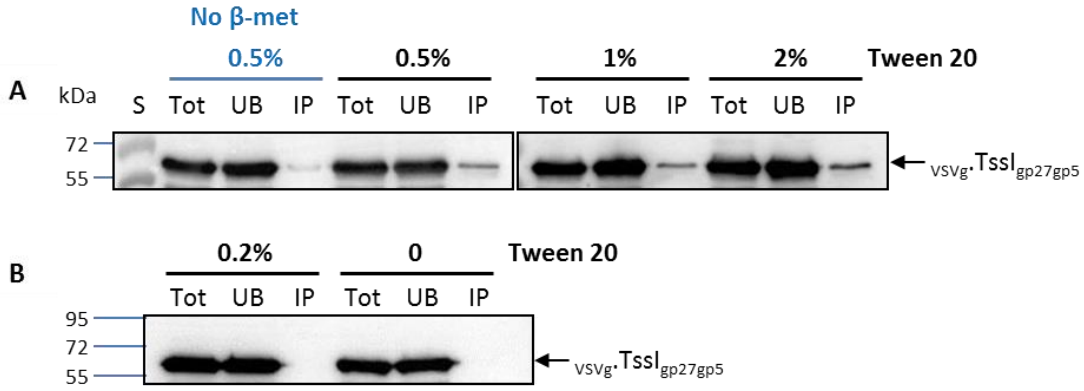


Figure 4.29 Optimisation of binding and elution buffers to prevent non-specific binding of $TssI_{gp27gp5}$ to the anti-FLAG coupled beads. Overproduction of $TssI_{gp27gp5}$ was achieved by the induction of pACYCDuet-VSVg.*tssI.gp27gp5* in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 30°C. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 500 mM NaCl and a series of Tween 20 concentrations, i.e. 0.5%, 1% and 2%. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and $_{VSVg}.TssI_{gp27gp5}$ was immuno-detected with anti-VSVg monoclonal antibody following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected location of $TssI_{gp27gp5}$ based on its size (~61 kDa). **A**. A series of Tween 20 concentrations as indicated. The set of samples in which β-met was omitted in the elution step is labelled in blue, otherwise, 5% β-met was included. **B**. Binding buffer with or without 0.2% Tween 20. β-met was omitted in the elution buffer. The arrows indicate the expected location of $TssI_{gp27gp5}$ based on its size (~61 kDa).

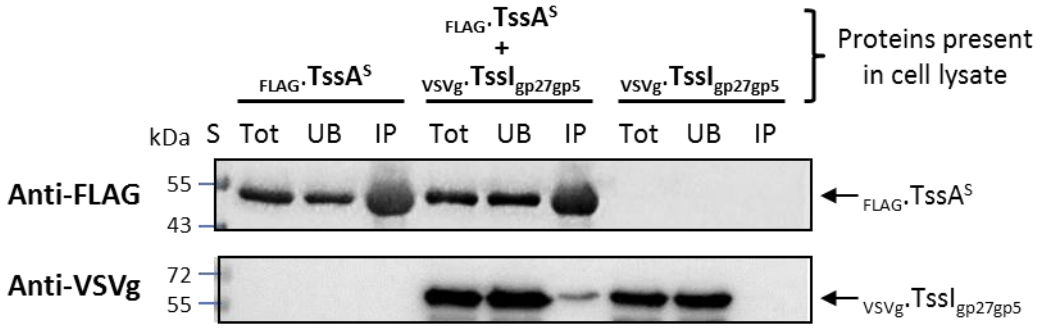


Figure 4.30 Co-IP analysis of the interaction between TssA^S and TssI$_{gp27gp5}$. $_{FLAG}$.TssA^S and $_{VSVg}$.TssI$_{gp27gp5}$ was expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S}, pACYCDuet-VSVg.*tssI.gp27gp5* and pACYCDuet-VSVg.*tssI.gp27gp5*-FLAG.*tssA*^{S}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4) and 500 mM NaCl. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^S, upper panel) and anti-VSVg ($_{VSVg}$.TssI$_{gp27gp5}$, lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssI$_{gp27gp5}$ (~61 kDa) and TssA^S (~42 kDa).

4.4.1.7 Co-IP analysis of the interaction between TssA^S and TssK

Using the BACTH assay, TssA^S and TssK were shown to interact when the C-terminus of both proteins were free (Section 3.3.11). A plasmid expressing only N-terminal VSVg-tagged TssK, pACYCDuet-VSVg.*tssK* (Figure 4.31 A), has been constructed as described in Section 4.2.2. In order to co-express both proteins from a single plasmid, it was decided to clone *tssA*^{S} into pACYCDuet-VSVg.*tssK*. A pair of primers, TssA.NdeI.FLAG.for and TssA.BglII.rev, was used for PCR amplification of *tssA*^{S}, in which a FLAG tag coding sequence was contained in the forward primer. The amplified *tssA*^{S} DNA fragment of the expected size (~1.1 kb) and pACYCDuet-VSVg.*tssK* were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites in the primers and the plasmid MCS. Ligation was carried out to allow the insertion of the *tssA*^{S} DNA fragment into the plasmid for constructing pACYCDuet-VSVg.*tssK*-FLAG.*tssA*^{S}, in which *tssK* is located downstream of the first T7 promoter and *tssA*^{S} is located downstream of the second (Figure 4.31 B). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

Following induction of *E. coli* stain BL21(λDE3) cells containing pACYCDuet-VSVg.*tssK*, TssK was overproduced and was found to be soluble (Figure 4.31 C). However, in the overexpression of both TssA^S and TssK, it was difficult to determine whether the overproduction of both proteins had occurred based on analysis by SDS-PAGE, as although the molecular weights of TssA^S and TssK differ by ~10 kDa, TssA^S migrates in gels according to a protein that is bigger than its actual size. Therefore, TssK and TssA^S co-migrate (Figure 4.31 D). Successful expression of both proteins was therefore checked by loading a smaller amount of sample on the SDS-PA gel. This showed that there were two proteins in the soluble fraction of the cell lysate that corresponded to the expected sizes of TssA^S and TssK, where TssK was approximately twice as abundant as TssA^S (Figure 4.31 E). Overproduction of $_{FLAG}$.TssA^S in the absence of TssK is shown in Figure 4.10 (Section 4.4.1.1). Co-IP was carried out for validating the BACTH results that showed an interaction between TssA^S and TssK. The results of the co-IP showed TssK did not bind to the anti-FLAG coupled beads in the absence of TssA^S. In contrast, when TssK was overexpressed with TssA^S, TssK was also present in the immuno-precipitated material supporting the BACTH assay result that indicated an interaction between these two proteins (Figure 4.32).

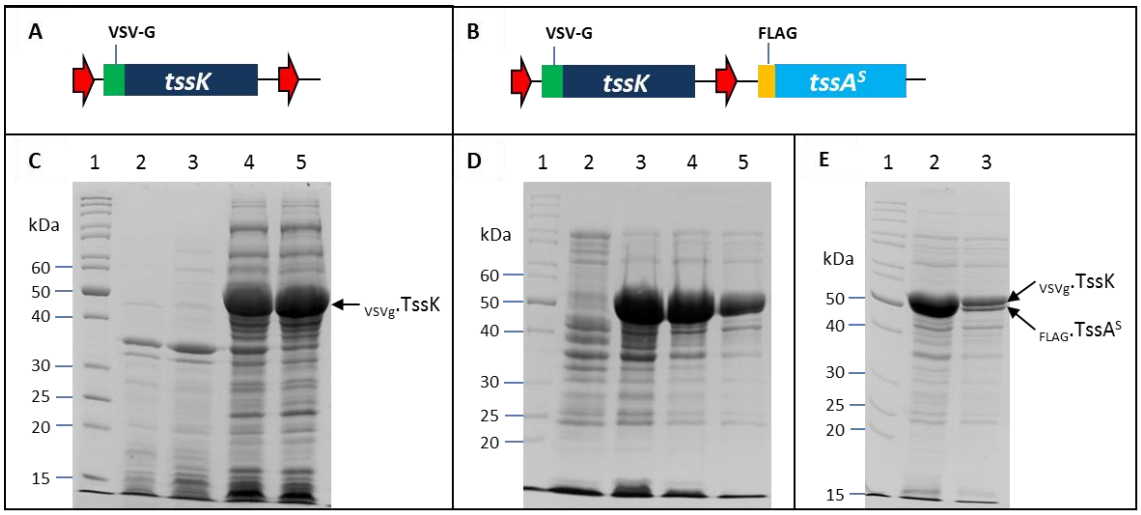


Figure 4.31 Overproduction and solubility of TssK in the presence and absence of TssA^S. A. Schematic drawing showing the location of VSV-g tagged *tssK* in pACYCDuet-VSVg.*tssK*. The red arrows represent the two T7 promoters present in the plasmid. **B.** Schematic drawing of the arrangement of FLAG tagged *tssA*^{S} and VSV-g tagged *tssK* in pACYCDuet-VSVg.*tssK*-FLAG.*tssA*^{S}. The red arrows represent the two T7 promoters present in the plasmid. **C**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssK* in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37^{o}C. The arrow indicates the expected location of TssK based on its size (~51 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **D**. A Coomassie blue-stained 10% SDS-PA gel showing the the analysis of proteins following induction of pACYCDuet-VSVg.*tssK*-FLAG.*tssA*^{S} plasmid in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 37^{o}C. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **E**. A Coomassie blue-stained 10% SDS-PA gel showing the induction of pACYCDuet-VSVg.*tssK*-FLAG.*tssA*^{S} plasmid in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 37^{o}C with one third of the amount of sample that was loaded on gel shown in D. The arrows indicate the expected locations of TssA^S (~42 kDa) and TssK (~51 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction.

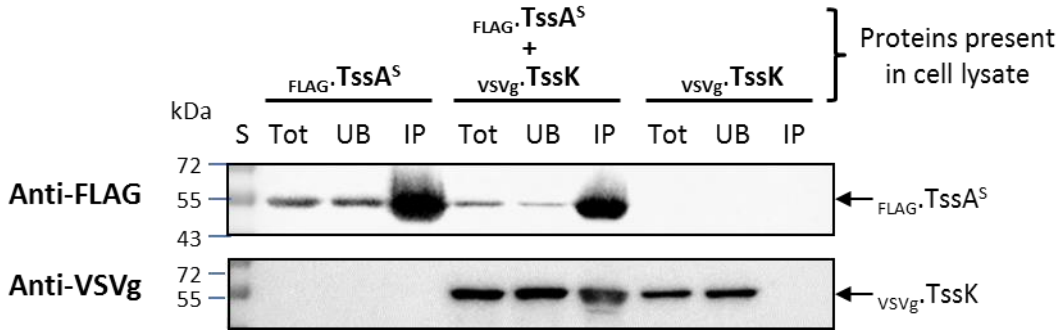


Figure 4.32 Co-IP analysis of the interaction between TssA^S and TssK. $_{FLAG}$.TssA^S and $_{VSVg}$.TssK were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S}, pACYCDuet-VSVg.*tssK* and pACYCDuet-VSVg.*tssK*-FLAG.*tssA*^{S}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^S, upper panel) and anti-VSVg ($_{VSVg}$.TssK lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^S (~42 kDa) and TssK (~51 kDa).

4.4.1.8 Co-IP analysis of the interaction between TssA^S and TssL

TssL was shown to interact with TssA^S by the BACTH assay when the C-terminal ends of both proteins were free (Section 3.3.12). A plasmid for expressing N-terminal VSVg-tagged TssL alone (pACYCDuet-VSVg.*tssL*) has been constructed as described in Section 4.2.3 (see also Figure 4.33 A). To construct a plasmid for co-expression of both proteins, *B. cenocepacia* H111 *tssL* was amplified using primers, TssLforpET.NtermVSVGTag and TssLRev, in which a VSV-g tag coding sequence was contained in the forward primer. The amplified *tssL* DNA fragment of the expected size (~600 bp) was digested with restriction enzymes *Nco*I and *Bgl*II that recognized sites in the primers, then ligated into *Nco*I and *Bam*HI digested pACYCDuet-FLAG.*tssA*^{S} plasmid for constructing pACYCDuet-VSVg.*tssL*-FLAG.*tssA*^{S}, in which *tssL* is located downstream of the first T7 promoter and *tssA*^{S} is located downstream of the second (Figure 4.33 B). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

Protein overproduction was conducted in *E. coli* strain BL21(λDE3). Overexpression of $_{VSVg}$.TssL from pACYCDuet-VSVg.*tssL* was successful, and a large amount of protein was overproduced with approximately 50% of it in the soluble fraction (Figure 4.33 C). The co-expression of TssL and TssA^S was also successful; both proteins were overproduced in similar amounts and remained soluble (Figure 4.33 D). Overproduction of $_{FLAG}$.TssA^S in the absence of TssL is shown in Figure 4.10 (Section 4.4.1.1). Co-IP was carried out using anti-FLAG coupled beads. The results showed that TssL did not bind to the beads in the absence of TssA^S. However, in the presence of TssA^S, some TssL was present in the immuno-precipitated material. This result supports the BACTH assay data suggesting there is an the interaction between TssA^S and TssL (Figure 4.34).

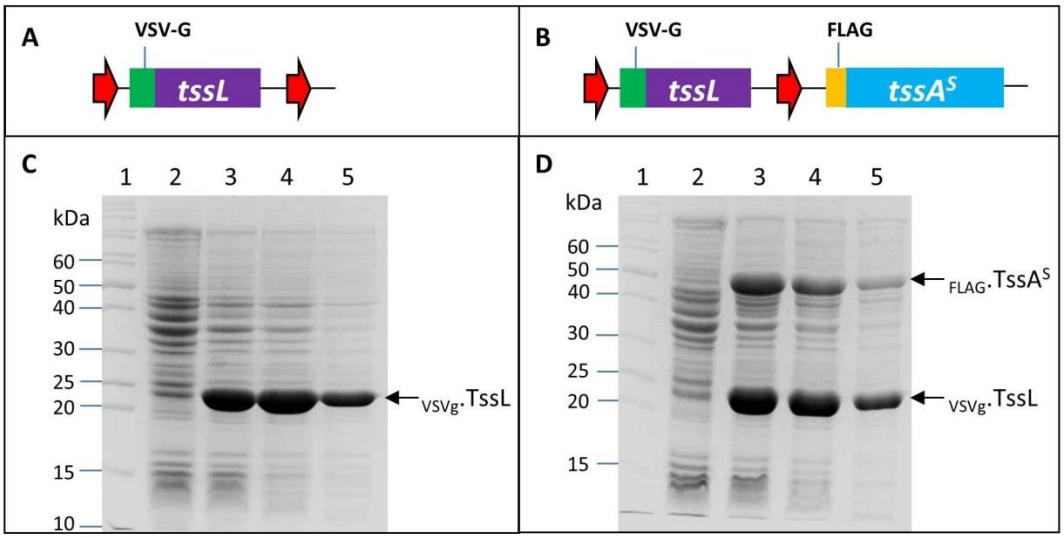


Figure 4.33 Overproduction and solubility of TssL in the presence and absence of TssA^S. A. Schematic drawing showing the location of VSV-g tagged *tssL* in pACYCDuet-VSVg.*tssL*. The red arrows represent the two T7 promoters present in the plasmid. **B.** Schematic drawing of the arrangement of VSV-g tagged *tssL* and FLAG.*tssA*^{S} in pACYCDuet-VSVg.*tssL*-FLAG.*tssA*^{S}. The red arrows represent the two T7 promoters present in the plasmid. **C**. A Coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssL* in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37^{o}C. The arrow indicates the expected location of TssL based on its size (~24 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **D**. A Coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssL*-FLAG.*tssA*^{S} plasmid in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 37^{o}C. The arrows indicate the expected locations of TssA^S (~42 kDa) and TssL (~24 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

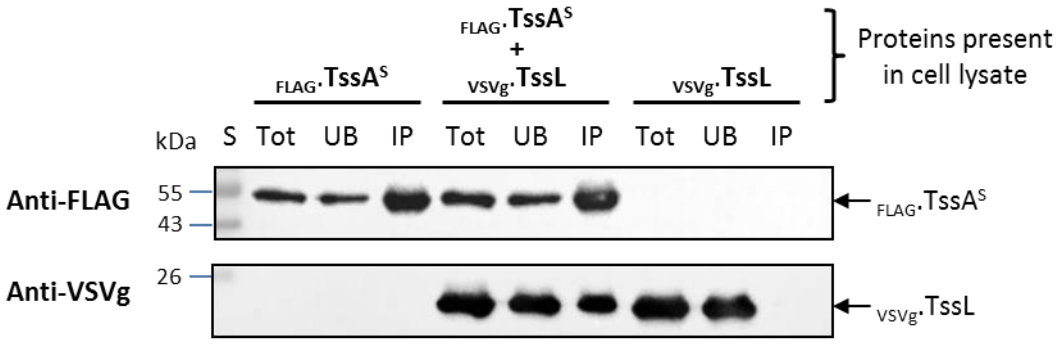


Figure 4.34 Co-IP analysis of the interaction between TssA^S and TssL. $_{FLAG}$.TssA^S and $_{VSVg}$.TssL were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S}, pACYCDuet-VSVg.*tssL* and pACYCDuet-VSVg.*tssL*-FLAG.*tssA*^{S}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 12% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^S, upper panel) and anti-VSVg ($_{VSVg}$.TssL lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^S (~42 kDa) and TssL (~24 kDa).

4.4.1.9 Co-IP analysis of the interaction between $TssA^S$ and $TssM_{NTD}$

$TssM_{NTD}$, the cytoplasmic domain located between the second and third transmembrane helices of TssM (amino acids 58-445 in BCAL0348), has been shown to interact with $TssA^S$ by the BACTH assay when the CyaA T25 fragment was fused at the C-terminus of $TssM_{NTD}$ (Section 3.3.13). The plasmids for expressing N-terminal VSVg-tagged TssM, alone and in combination with $_{FLAG}$.$TssA^S$, were constructed in parallel. *B. cenocepacia* H111 $tssM_{NTD}$ was amplified using primers, TssM1.VSVg.NcoI.for and TssM1.HindIII.rev, in which a VSV-g tag coding sequence was contained in the forward primer. Although the N-terminus of $TssM_{NTD}$ was unblocked in the two-hybrid protein that gave a positive result with $TssA^S$ in the BACTH assay, given that the VSVg epitope tag is a small tag and two closely spaced TMDs are normally located at the N-terminal end of $TssM_{NTD}$ in the native protein, the presence of the tag would not be expected to interfere with the interaction between $TssM_{NTD}$ with $TssA^S$. The amplified $tssM_{NTD}$ DNA fragment of the expected size (~1.17 kb), and plasmids pACYCDuet-1 and pACYCDuet-FLAG.$tssA^S$ plasmids were digested with restriction enzymes *Nco*I and *Hin*dIII that recognized sites in the primers and the upstream MCS of both plasmids. Ligation was carried out to allow the insertion of the $tssM_{NTD}$ DNA fragment into plasmids for constructing pACYCDuet-VSVg.$tssM_{NTD}$ and pACYCDuet-VSVg.$tssM_{NTD}$-FLAG.$tssA^S$. The nucleotide sequence of the inserted gene was verified by DNA sequencing. Diagrammatic representations of the two plasmids are shown in Figure 4.35 A and D.

Overproduction of $_{VSVg}$.$TssM_{NTD}$ alone was conducted in *E. coli* strain BL21(λDE3) cells containing pACYCDuet-VSVg.$tssM_{NTD}$. $_{VSVg}$.$TssM_{NTD}$ was successfully overproduced following induction with 1 mM IPTG at 37^{o}C. However, the protein was found in the insoluble fraction following cell lysis (result was not shown). A small fraction of $_{VSVg}$.$TssM_{NTD}$ became soluble with 1 mM IPTG induction carried out at lower temperature, i.e. 30^{o}C (result was not shown). After testing a few more induction conditions, the final condition chosen for $_{VSVg}$.$TssM_{NTD}$ induction employed 0.1 mM IPTG for cells growing at 22^{o}C (Figure 4.35 B). Although less protein was produced compared to higher induction temperatures, more soluble material was obtained in low temperature induction condition. An overproduced protein with the expected size of $_{VSVg}$.$TssM_{NTD}$ was confirmed to be $_{VSVg}$.$TssM_{NTD}$ by immuno-detection with anti-VSV-g epitope antibody (Figure 4.31 C). The result confirmed that there was a

reasonable amount of $_{VSVg}$.TssM$_{NTD}$ overproduced and it was partially soluble. Some degradation products of $_{VSVg}$.TssM$_{NTD}$ were also detected.

Co-production of $_{VSVg}$.TssM$_{NTD}$ and $_{FLAG}$.TssA^{S} was also conducted in *E. coli* strain BL21(λDE3) containing pACYCDuet-VSVg.*tssM$_{NTD}$*-FLAG.*tssA^{S}* using 0.1 mM IPTG induction at 22^{o}C. More $_{FLAG}$.TssA^{S} was overproduced than $_{VSVg}$.TssM$_{NTD}$ (Figure 4.35 E). Overexpression of $_{FLAG}$.TssA^{S} in the absence of TssM$_{NTD}$ is shown in Figure 4.10 (Section 4.4.1.1). Despite the small amount of $_{VSVg}$.TssM$_{NTD}$ that was soluble, co-IP was still carried out with the anti-FLAG coupled beads. The results showed that there were still some degradation products of $_{VSVg}$.TssM$_{NTD}$ present when it was co-expressed with $_{FLAG}$.TssA^{S} (Figure 4.36). The full-length $_{VSVg}$.TssM$_{NTD}$ and its degradation product did not bind to the beads in the absence of $_{FLAG}$.TssA^{S}. However, when both $_{VSVg}$.TssM$_{NTD}$ and $_{FLAG}$.TssA^{S} were present in the cell lysate, $_{VSVg}$.TssM$_{NTD}$ was also immuno-detected in the precipitated material indicating an interaction between them. The largest degradation product of $_{VSVg}$.TssM$_{NTD}$ also interacted with $_{FLAG}$.TssA^{S}, but the interaction between the degraded $_{VSVg}$.TssM$_{NTD}$ and $_{FLAG}$.TssA^{S} was weak.

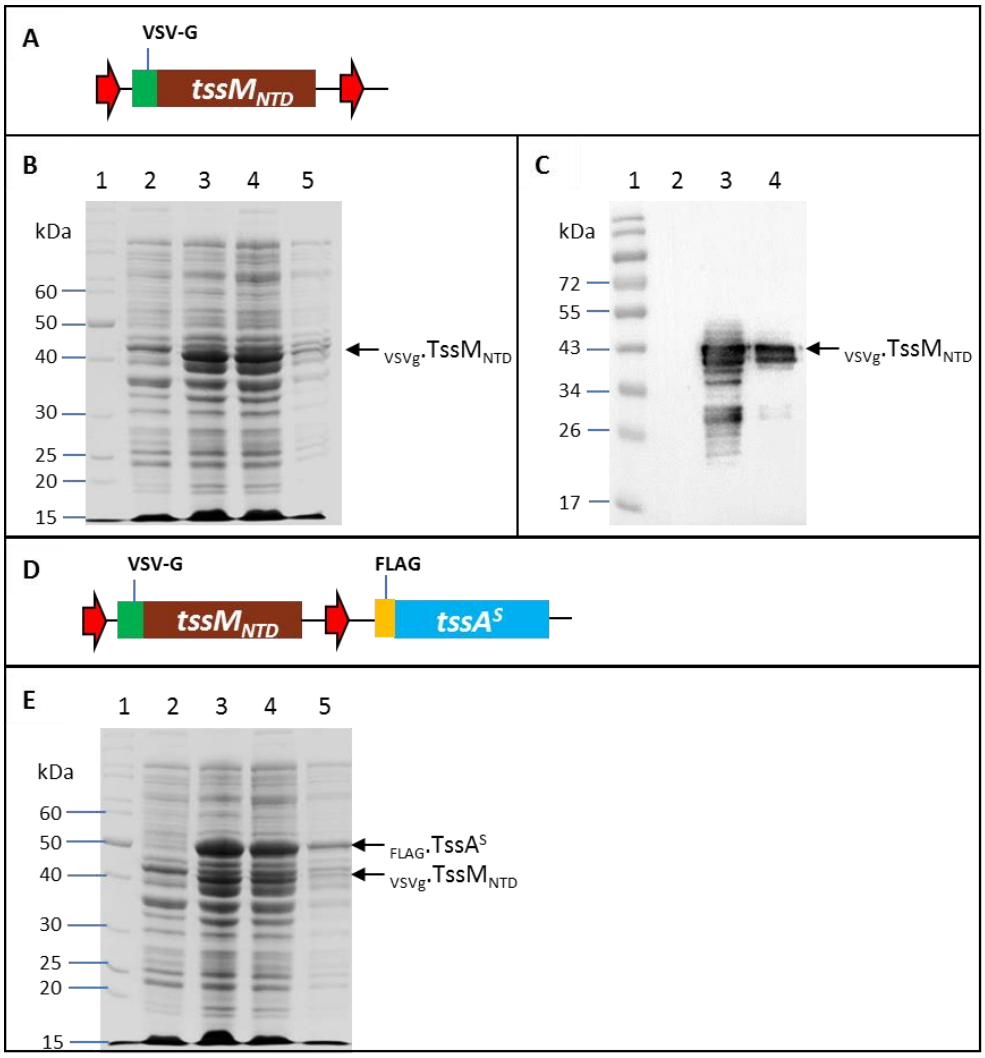


Figure 4.35 Overproduction and solubility of $TssM_{NTD}$, in the presence and absence of $TssA^S$. A. Schematic drawing showing the location of VSV-g tagged *tssM*$_{NTD}$ in pACYCDuet-VSVg.*tssM*$_{NTD}$. The red arrows represent the two T7 promoters present in the plasmid. **B**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssM*$_{NTD}$ in *E. coli* strain BL21(λDE3) with 0.1 mM IPTG at 22°C. Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **C**. Immuno-detection of $TssM_{NTD}$ solubility in *E. coli* strain BL21(λDE3) cells following 0.1 mM IPTG induction at 22°C with anti-VSVg mAb. The arrows in B and C indicate the expected locations of $TssM_{NTD}$ based on its size (~45 kDa). Lane 1, EZ-Run™ Prestained Rec Protein Ladder; lane 2, total cell protein from uninduced cells; lane 3, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 4, soluble fraction of cell lysate following induction. **D.** Schematic drawing of the arrangement of VSV-g tagged *tssM*$_{NTD}$ and FLAG.*tssA*^{S} in pACYCDuet-VSVg.*tssM*$_{NTD}$-FLAG.*tssA*^{S}. The red arrows represent the two T7 promoters present in the plasmid. **E**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssM*$_{NTD}$-FLAG.*tssA*^{S} plasmid in *E. coli* strain BL21(λDE3) cells with 0.1 mM IPTG at 22°C. The arrows indicate the expected locations of $TssA^S$ (~42 kDa) and $TssM_{NTD}$ (~45 kDa). Lane 1, EZ-Run™ Prestained Rec Protein Ladder; lane 2, total cell protein from uninduced cells; lane 3, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 4, soluble fraction of cell lysate following induction.

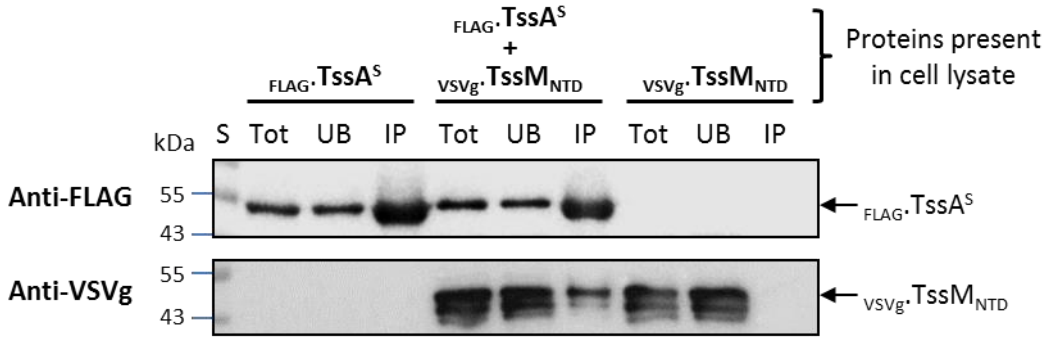


Figure 4.36 Co-IP analysis of the interaction between $TssA^S$ and $TssM_{NTD}$. $_{FLAG}.TssA^S$ and $_{VSVg}.TssM_{NTD}$ were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S}, pACYCDuet-VSVg.*tssM*$_{NTD}$ and pACYCDuet-VSVg.*tssM*$_{NTD}$-FLAG.*tssA*^{S}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 500 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}.TssA^S$, upper panel) and anti-VSVg ($_{VSVg}.TssM_{NTD}$ lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of $TssA^S$ (~42 kDa) and $TssM_{NTD}$ (~45 kDa).

4.4.2 Co-IP analysis of interactions between $TssA^{EI}$ and other T6SS subunits using anti-FLAG affinity gel

Using the BACTH system, *A. hydrophila* $TssA^{EI}$ has been shown to interact with *B. cenocepecia* TssB, TssC, TssF in this study and with five other T6SS subunits (TssD, TssH, TssI, TssK and TssL) in a previous study (Ahmad 2013), but not with TssE or $TssM_{NTD}$. The *in vivo* interactions between $TssA^{S}$ and other T6SS subunits, i.e. TssC, TssD, TssE, TssF, $TssI_{gp27gp5}$, TssK, TssL and $TssM_{NTD}$ have been corroborated *in vitro* by using co-IP with anti-FLAG affinity gel (Section 4.1.1). As $TssA^{EI}$ is predicted to be a functional orthologue of $TssA^{S}$, it was decided to provide biochemical evidence for the *in vivo* interactions between $TssA^{E}$ and the $TssA^{S}$-interacting T6SS subunits to demonstrate the similarities of these two proteins with the respect to their functions.

To do this, similarly to $TssA^{S}$, $TssA^{EI}$ was tagged by a FLAG tag as its N-terminus and anti-FLAG affinity gel was employed for precipitating $TssA^{EI}$. In general, the other T6SS subunit (TssX) was tagged with a VSV-g epitope tag. Co-expression of two proteins ($TssA^{EI}$ and TssX) for co-IP analysis was achieved by cloning multiple *tss* genes into a single plasmid, pACYCDuet, except for TssE and TssF, where plasmids containing *tssE/tssF* which are compatible with *tssA*^{EI}-containing pACYCDuet were used. To carry out the co-IP, soluble fraction of the cell lysate following induction containing $_{FLAG}$.$TssA^{EI}$ and potential interacting T6SS subunits (TssX) was incubated with the anti-FLAG coupled beads, by which FLAG-tagged $TssA^{E}$ can be recovered from the lysate along with the $TssA^{EI}$-interacting proteins. A similar amount of TssX was applied to the resin in the absence of $_{FLAG}$.$TssA^{EI}$ as an experimental control. As another control, $_{FLAG}$.$TssA^{EI}$ was also immunoprecipitated in the absence of TssX. The unbound material was removed by a series of washes. The immunoprecipitated material was then eluted from the beads by boiling in sample buffer and analysed by SDS-PAGE and western blotting using antibodies that are specific to the epitope tags (Section 2.7.4.2).

4.4.2.1 Construction of a plasmid producing FLAG-tagged TssA^{EI} and analysis of $_{FLAG}$.TssA^{EI} overexpression

A. hydrophila tssA^{EI} was amplified using primers, AHA1844.NdeI.FLAG.for and AHA1844.BglII.rev, in which a FLAG tag coding sequence was contained in the forward primer. The amplified *tssA*^{EI} DNA fragment of the expected size (~1.4 kb) and pACYCDuet-1 plasmid were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites in the primers and the plasmid MCS located downstream of the second T7 promoter. Ligation was carried out to allow the insertion of the *tssA*^{EI} DNA fragment into plasmid for constructing pACYCDuet-FLAG.*tssA*^{EI} (Figure 4.37 A). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

TssA^{EI} overproduction was conducted in *E. coli* strain BL21(λDE3) under different induction conditions. Large amounts of TssA^{EI} protein were produced when the induction was performed at both 37^{o}C and 30^{o}C. However, most of the TssA^{EI} that was produced at the higher temperature was insoluble, whereas, when the induction was carried out at 30^{o}C, approximately 50% of TssA^{EI} became soluble (Figure 4.37 B).

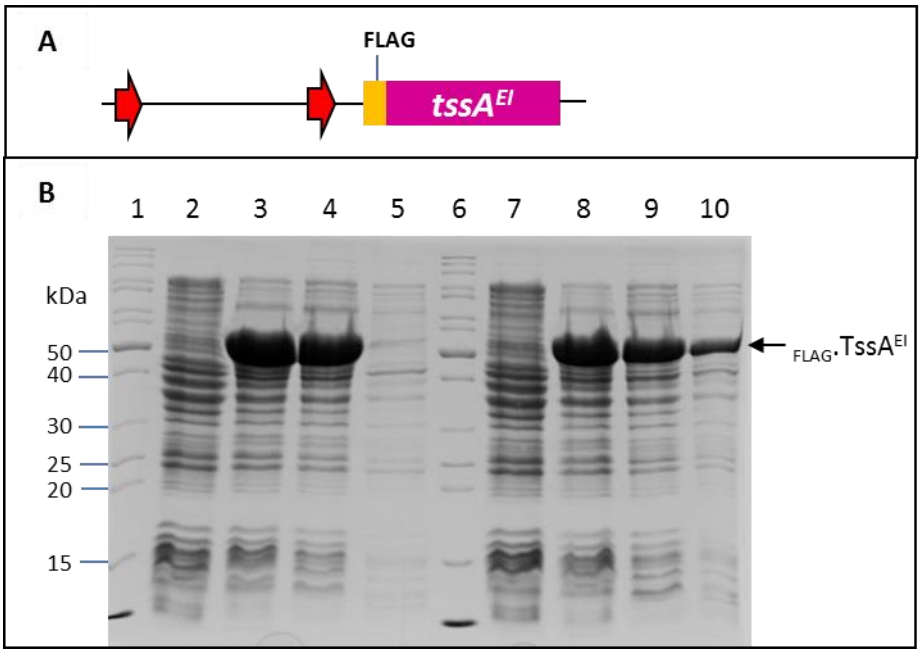


Figure 4.37 Overproduction and solubility of $_{FLAG}$.TssA^{EI}. A. Schematic drawing showing the location of FLAG tagged *tssA*^{EI} in pACYCDuet-FLAG.*tssA*^{EI}. The red arrows represent the two T7 promoters present in the plasmid. **B**. A Coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-FLAG.*tssA*^{EI} in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37/30^{o}C. The arrow indicates the expected location of TssA^{EI} based on its size (~53 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 2-5, induction at 37^{o}C: lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction; Lane 6, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 7-10, induction at 30^{o}C: lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

4.4.2.2 Co-IP analysis of the interaction between TssA^{EI} and TssC

To construct a plasmid for co-expression of both TssA^{EI} and TssC, *tssA^{EI}* was amplified using primers, AHA1844.NdeI.FLAG.for and AHA1844.KpnI.rev, in which a FLAG tag coding sequence was contained in the forward primer. The amplified *tssA^{EI}* DNA fragment of the expected size (~1.4 kb) and pACYCDuet-*tssC*.VSVg were digested with restriction enzymes *Nde*I and *Acc*65I that recognized sites in the primers and the MCS located downstream of the second T7 promoter in the plasmid. Ligation was carried out to allow the insertion of the *tssA^{EI}* DNA fragment into the plasmid for constructing pACYCDuet-*tssC*.VSVg-FLAG.*tssA^{EI}*. The nucleotide sequence of the inserted gene was verified by DNA sequencing. A diagrammatic representation of the TssC.$_{VSVg}$ and $_{FLAG}$.TssA^{EI} co-expression plasmid is shown in Figure 4.38 A.

Protein production was carried out in *E. coli* strain BL21(λDE3) by inducting the T7 promoters on pACYCDuet-*tssC*.VSVg-FLAG.*tssA^{EI}* with 0.5 mM IPTG at 30^{o}C after attempts in different induction conditions. Both proteins were overexpressed very well. Although the molecular weights of these two proteins are similar, two thick protein bands could be identified on a Coomassie blue-stained SDS-PA gel (Figure 4.38 B). However, there was not much TssC protein present in the soluble cell lysate based on the previous efforts to increase the fraction of soluble TssC by varying the induction condition (Section 4.4.1.2). Nevertheless, co-IP was carried out as the amount of soluble TssC was considered to be immuno-detectable. Overexpression of $_{FLAG}$.TssA^{EI} in the absence of TssC and overexpression of TssC in the absence of $_{FLAG}$.TssA^{EI} are shown in Figure 4.37 (Section 4.4.2.1) and Figure 4.11 (Section 4.4.1.2), respectively. The co-IP results showed that TssC did not bind to the anti-FLAG coupled beads in the absence of TssA^{EI}, whereas when it was co-expressed with TssA^{EI}, it was co-immuno-precipitated by the antibody, suggesting that there is a specific interaction between these two proteins (Figure 4.39). This is consistent with the *in vivo* results obtained by the BACTH assay, where a combination of TssA^{EI} and TssC two-hybrid proteins, TssC-T25 with T18-TssA^{EI}, yielded a maltose positive phenotype (Section 3.5.2).

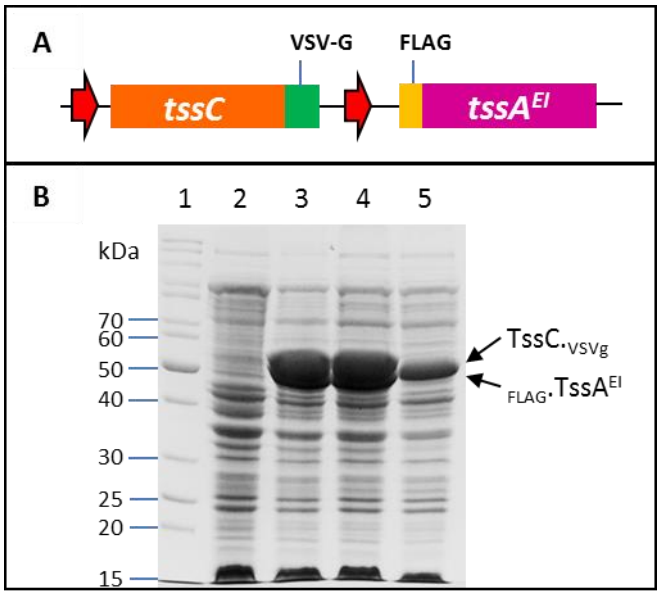


Figure 4.38 Co-expression and solubility of TssA^{EI} and TssC. A. Schematic drawing of the arrangement of FLAG tagged *tssA*^{EI} and *tssC*.VSVg in pACYCDuet-*tssC*.VSVg-FLAG.*tssA*^{EI}. The red arrows represent the two T7 promoters present in the plasmid. **B**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-*tssC*.VSVg-FLAG.*tssA*^{EI} plasmid in *E. coli* strain BL21(λDE3) with 0.5 mM IPTG at 30^{o}C. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssC (~56 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

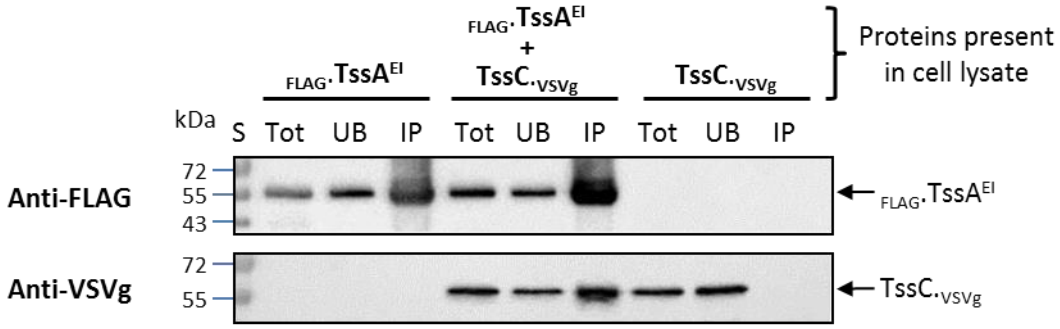


Figure 4.39 Co-IP analysis of the interaction between TssA^{EI} and TssC. $_{FLAG}$.TssA^{EI} and TssC.$_{VSVg}$ were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{EI}, pACYCDuet-*tssC*.VSVg and pACYCDuet-*tssC*.VSVg-FLAG.*tssA*^{EI}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^{EI}, upper panel) and anti-VSVg (TssC.$_{VSVg}$ lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssC (~56 kDa).

4.4.2.3 Co-IP analysis of the interaction between TssA^{EI} and TssD

The strategy for co-expression of both TssA^{EI} and TssD was similar to that for TssA^{E} and TssC, which was to insert *tssA^{EI}* into the second MCS of pACYCDuet that already contained the *tssD* gene inserted in the first MCS (i.e. pACYCDuet-VSVg.*tssD*). However, on this occasion, a different pair of primers was employed for amplifying *tssA^{EI}* (AHA1844.BglII.FLAG.for and AHA1844.MunI.rev) which contained different restriction sites, *Bgl*II and *Mfe*I, in the forward and reverse primer, respectively. A FLAG tag coding sequence was also included in the forward primer for introducing a FLAG epitope tag at the N-terminus of TssA^{EI}. The amplified *tssA^{EI}* DNA fragment of the expected size (~1.4 kb) and pACYCDuet-VSVg.*tssD* were digested with restriction enzymes *Bgl*II and *Mfe*I and ligated for constructing pACYCDuet-VSVg.*tssD*-FLAG.*tssA^{EI}*. The nucleotide sequence of the inserted gene was verified by DNA sequencing. A diagrammatic representation of the $_{VSVg}$.TssD and $_{FLAG}$.TssA^{EI} co-expression plasmid is shown in Figure 4.40 A.

Overproduction of both proteins was carried out in *E. coli* BL21(λDE3) containing pACYCDuet-VSVg.*tssD*-FLAG.*tssA^{EI}* by induction of the T7 promoters with 0.5 mM IPTG at 30^{o}C. Both proteins were overproduced very well and present in the soluble fraction of the cell lysate (Figure 4.40 B). Overexpression of $_{FLAG}$.TssA^{EI} in the absence of $_{VSVg}$.TssD and overexpression of $_{VSVg}$.TssD in the absence of $_{FLAG}$.TssA^{EI} are shown in Figure 4.37 (Section 4.4.2.1) and Figure 4.14C (Section 4.4.1.3), respectively. Co-IP was carried out using anti-FLAG coupled beads. The results showed that $_{VSVg}$.TssD was not present in the immuno-precipitated material whether $_{FLAG}$.TssA^{EI} was present or not, suggesting $_{VSVg}$.TssD does not interact with TssA^{EI} (Figure 4.41). However, the *in vivo* result from the BACTH assay suggested $_{FLAG}$.TssA^{EI} and $_{VSVg}$.TssD interact with each other (Ahmad 2013).

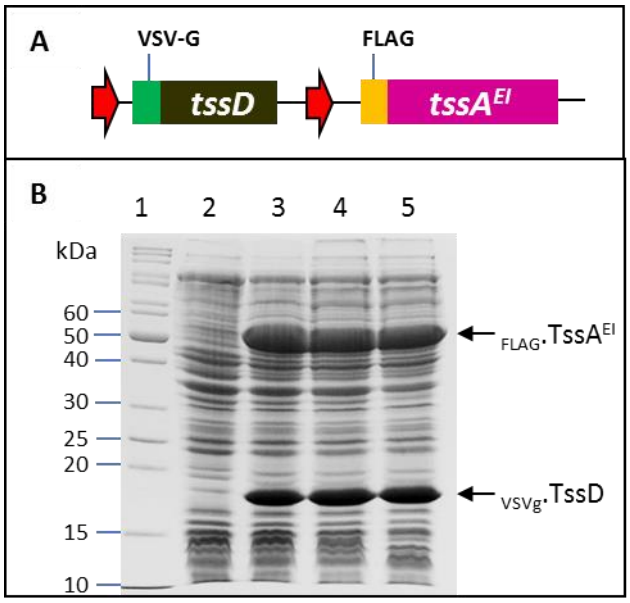


Figure 4.40 Co-overexpression and solubility of TssA^{EI} and TssD. A. Schematic drawing of the arrangement of FLAG tagged *tssA*^{EI} and VSVg.*tssD* in pACYCDuet-VSVg.*tssD*-FLAG.*tssA*^{EI}. The red arrows represent the two T7 promoters present in the plasmid. **D**. A Coomassie blue-stained 12% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssD*-FLAG.*tssA*^{EI} plasmid in *E. coli* strain BL21(λDE3) with 0.5 mM IPTG at 30^{o}C. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssD (~20 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

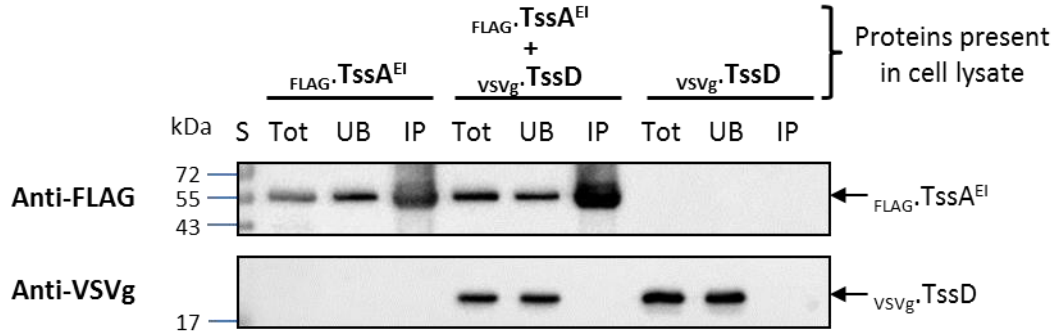


Figure 4.41 Co-IP analysis of the interaction between TssA^{EI} and TssD. $_{FLAG}$.TssA^{EI} and $_{VSVg}$.TssD were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{EI}, pACYCDuet-VSVg.*tssD* and pACYCDuet-VSVg.*tssD*-FLAG.*tssA*^{EI}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^{EI}, upper panel) and anti-VSVg ($_{VSVg}$.TssD lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssD (~20 kDa).

4.4.2.4 Co-IP analysis of the interaction between $TssA^{EI}$ and TssE

As TssE was observed almost insoluble (Section 4.4.1.4), MBP-TssE fusion was engaged for analysing the interaction between $TssA^{E}$ and TssE. Co-expression of $_{His6}$.MBP-TssE and $_{FLAG}$.$TssA^{EI}$ was carried out by introducing pMAL-c5X-His_6-*tssE* and pACYCDuet-FLAG.$tssA^{EI}$ together into strain BL21(λDE3) cells (Figure 4.42 A). Following induction with IPTG, $_{His6}$.MBP-TssE and $_{FLAG}$.$TssA^{EI}$ was overproduced in similar amount, and both proteins remained in the soluble fraction following cell lysis (Figure 4.42 B). Overexpression of $_{FLAG}$.$TssA^{EI}$ in the absence of $_{His6}$.MBP-TssE and overexpression of $_{His6}$.MBP-TssE in the absence of $_{FLAG}$.$TssA^{E}$ are shown in Figure 4.37 (Section 4.4.2.1) and Figure 4.18C (Section 4.4.1.4), respectively.

Co-IP was carried out for the interaction analysis between $_{FLAG}$.$TssA^{EI}$ and $_{His6}$.MBP-TssE. The results showed that $_{His6}$.MBP-TssE did not bind to the anti-FLAG coupled beads when it was produced in the absence of $_{FLAG}$.$TssA^{E}$. In contrast, in the presence of $_{FLAG}$.$TssA^{EI}$, $_{His6}$.MBP-TssE was present in the immunoprecipitated material, suggesting a specific interaction between these two proteins (Figure 4.43). However, the *in vitro* result is not consistent with the *in vivo* result from the BACTH assay, in which there was no indication for the interaction between $_{FLAG}$.$TssA^{EI}$ and TssE (Ahmad 2013).

To rule out the possibility that the interaction between $_{FLAG}$.$TssA^{EI}$ and $_{His6}$.MBP-TssE was a result of interaction between $_{FLAG}$.$TssA^{EI}$ and $_{His6}$.MBP, it was decided to carry out a co-IP experiment between $_{FLAG}$.$TssA^{EI}$ and $_{His6}$.MBP. Co-expression of both proteins was carried out by induction of the T7 and *tac* promoters present on pACYCDuet-FLAG.$tssA^{EI}$ and pMALC5x-His_6 in BL21(λDE3) cells growing at 30^{o}C (Figure 4.44 A). Both proteins were expressed very well in similar amount and remained in the soluble fraction following cell lysis (Figure 4.44 B). Overexpression of $_{His6}$.MBP in the absence of $_{FLAG}$.$TssA^{EI}$ and overexpression of $_{FLAG}$.$TssA^{EI}$ in the absence of $_{His6}$.MBP are shown in Figure 4.20 C (Section 4.4.1.4) and Figure 4.37 (Section 4.4.2.1), respectively.

Co-IP was carried out with anti-FLAG coupled beads that precipitate $_{FLAG}$.$TssA^{EI}$. The results showed that MBP did not bind to the beads non-specifically and there was also no MBP present in the immuno-precipitated material when it was co-expressed with

$TssA^{E}$. The result suggested $_{FLAG}$.$TssA^{EI}$ does not interact with $_{His6}$.MBP (Figure 4.45). Therefore, the co-IP result for the interaction between $_{FLAG}$.$TssA^{EI}$ and $_{His6}$.MBP-TssE indicated a specific interaction between the two proteins.

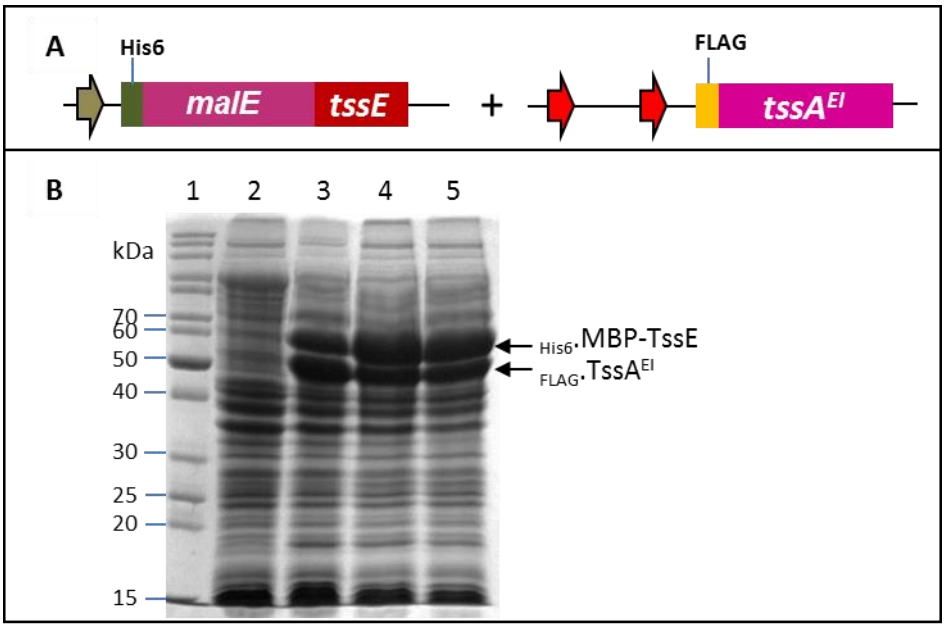


Figure 4.42 Co-overexpression and solubility of $_{His6}$.MBP-TssE and TssA^{EI}. A. Schematic drawing of compatible pMAL-c5X-His$_6$-*tssE* and pACYCDuet-FLAG.*tssA*^{EI} used for co-expression of $_{His6}$.MBP-TssE and TssA^{EI}. The grey arrow represents the *tac* promoter in pMAL-c5X-His$_6$-*tssE*. The red arrows represent the two T7 promoters present in pACYCDuet-FLAG.*tssA*^{EI}. **B**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following co-expression of pMAL-c5X-His$_6$-*tssE* and pACYCDuet-FLAG.*tssA*^{EI} in *E. coli* strain BL21(λDE3) following induction with 0.5 mM IPTG at 30^{o}C. The arrows indicate the expected locations of $_{His6}$.MBP-TssE (~62 kDa) and TssA^{EI} (~53 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

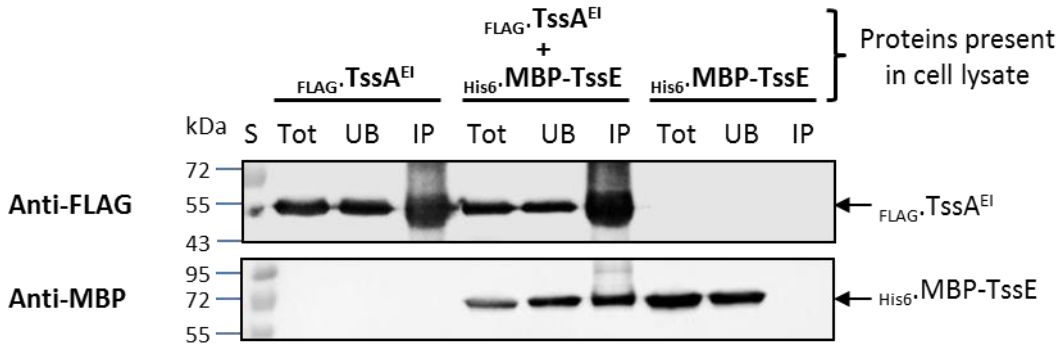


Figure 4.43 Co-IP analysis of the interaction between TssA^{EI} and $_{His6}$.MBP-TssE. $_{FLAG}$.TssA^{EI} and $_{His6}$.MBP-TssE were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{EI} and pMAL-c5X-His$_6$-*tssE* separately and together, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^{EI}, upper panel) and anti-MBP ($_{His6}$.MBP-TssE, lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expect locations of TssA^{EI} (~53 kDa) and $_{His6}$.MBP-TssE (~62 kDa).

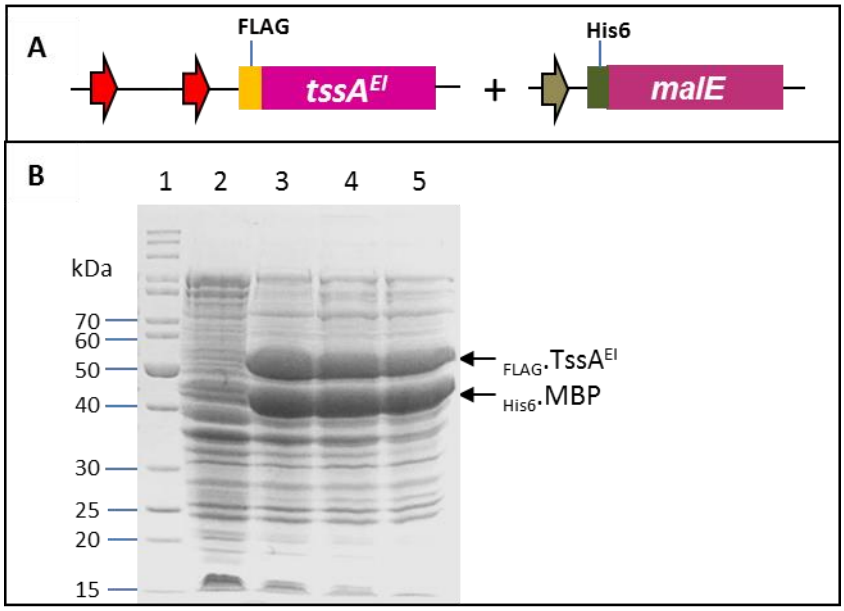


Figure 4.44 Co-overproduction and solubility of MBP and TssA^{EI}. A. Schematic drawing of compatible pMAL-c5X-His$_6$ and pACYCDuet-FLAG.*tssA^{EI}* used for co-expression of $_{His6}$.MBP and $_{FLAG}$.TssA^{EI}. **B**. A Coomassie blue-stained 10% SDS-PA gel showing the co-expression of pMAL-c5X-His$_6$ and pACYCDuet-FLAG.*tssA^{EI}* in *E. coli* strain BL21(λDE3) following induction with 1 mM IPTG at 30^{o}C. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and $_{His6}$.MBP (~45 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

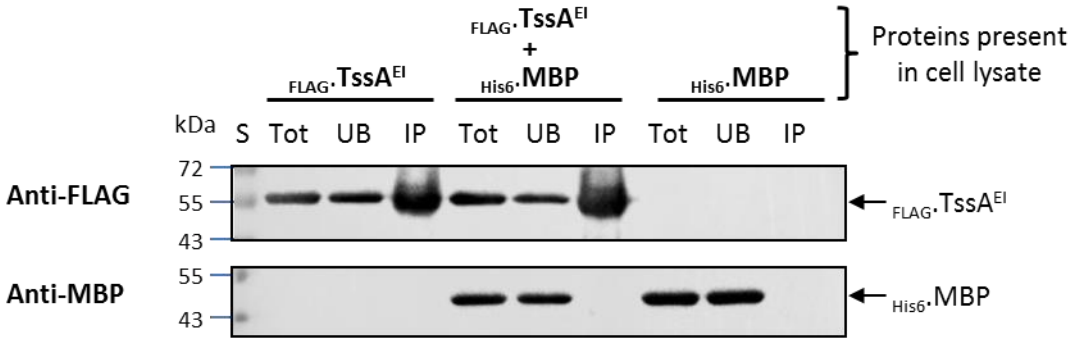


Figure 4.45 Co-IP analysis of the interaction between TssA^{EI} and MBP. $_{FLAG}$.TssA^{EI} and $_{His6}$.MBP were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{EI} and pMAL-c5X-His$_6$ separately and together , accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^{EI}, upper panel) and anti-MBP ($_{His6}$.MBP, lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and $_{His6}$.MBP (~45 kDa).

4.4.2.5 Co-IP analysis of the interaction between TssA^{EI} and TssF

The BACTH assay showed an interaction between *A. hydrophelia* TssA^{EI} and *B. cenocepacia* TssF (Section 3.5.3), and the interaction between *B. cenocepacia* TssA^{S} and TssF has been demonstrated both *in vivo* and *in vitro* (Section 4.4.1.5). Therefore, it was decided to investigate the *in vitro* interaction between TssA^{E} and TssF by co-IP using anti-FLAG coupled beads. The strategy for co-expression of both TssA^{EI} and TssF was similar to that employed for TssA^{EI} and TssC, which was to insert *tssA^{EI}* into the second MCS of pACYCDuet that already had the *tssF* gene inserted at the first MCS (pACYCDuet-*tssF*.HA). The same pair of primers was employed for amplifying *tssA^{EI}* which contained *Nde*I and *Acc*65I in the forward and reverse primer, respectively. A FLAG tag coding sequence was also included in the forward primer for introducing a FLAG epitope tag at the N-terminus of TssA^{EI}. Ligation, selection and identification of the correct recombinant plasmid, pACYCDuet-*tssF*.HA-FLAG.*tssA^{EI}*, was performed as described previously for constructing pACYCDuet-*tssC*.VSVg-FLAG.*tssA^{EI}* (Section 4.4.2.2).

C-terminal HA tagged TssF was used for the investigation of the interaction of TssF with TssA^{EI}, as it was shown earlier that the N-terminus of TssF may be involved in the interaction with TssA^{S} (Section 4.4.1.5). Co-expression of TssA^{EI} and TssF was carried out in *E. coli* BL21(λDE3) with pACYCDuet-*tssF*.HA-FLAG.*tssA^{EI}* (Figure 4.46 A). Both proteins were found to be overproduced. However, only a very small amount of TssF and less than 40% of TssA^{EI} remained in the soluble fraction of the cell lysate (Figure 4.46 B). Therefore, it suggests that TssF reduces the solubility of TssA^{EI}. Overexpression of $_{FLAG}$.TssA^{EI} in the absence of TssF.$_{HA}$ and overexpression of TssF.$_{HA}$ in the absence of $_{FLAG}$.TssA^{EI} are shown in Figure 4.37 (Section 4.4.2.1) and Figure 4.22 C (Section 4.4.1.5), respectively. Co-IP was carried out using anti-FLAG coupled beads. The result showed that TssF did not bind to the beads non-specifically in the absence of TssA^{EI} (Figure 4.47). In contrast, when it was co-expressed with $_{FLAG}$.TssA^{EI}, TssF was present in the immuno-precipitated material, suggesting an interaction between these two proteins.

N-terminal HA tagged TssF was also investigated for its interaction with $_{FLAG}$.TssA^{EI} although it was not shown to interact with TssA^{S}. Co-expression of both $_{HA}$.TssF and $_{FLAG}$.TssA^{EI} was conducted in BL21(λDE3) containing compatible pETDuet-HA.*tssF*

and pACYCDuet-FLAG.*tssA*^{EI} (Figure 4.48 A). Both proteins were overproduced and only a small amount of TssF and approximately half of $_{FLAG}$.TssA^{EI} remained in the soluble fraction following cell lysis (Figure 4.48 B). Overexpression of $_{HA}$.TssF in the absence of $_{FLAG}$.TssA^{EI} is shown in Figure 4.26 C (Section 4.4.1.5). The co-IP result showed that the N-terminal HA tagged TssF was precipitated in the presence of $_{FLAG}$.TssA^{EI}, but not in the absence of $_{FLAG}$.TssA^{EI} (Figure 4.49). Therefore, the presence of HA epitope tag at either the N- or C-terminus of TssF did not interfere with the interaction of this protein with TssA^{EI}.

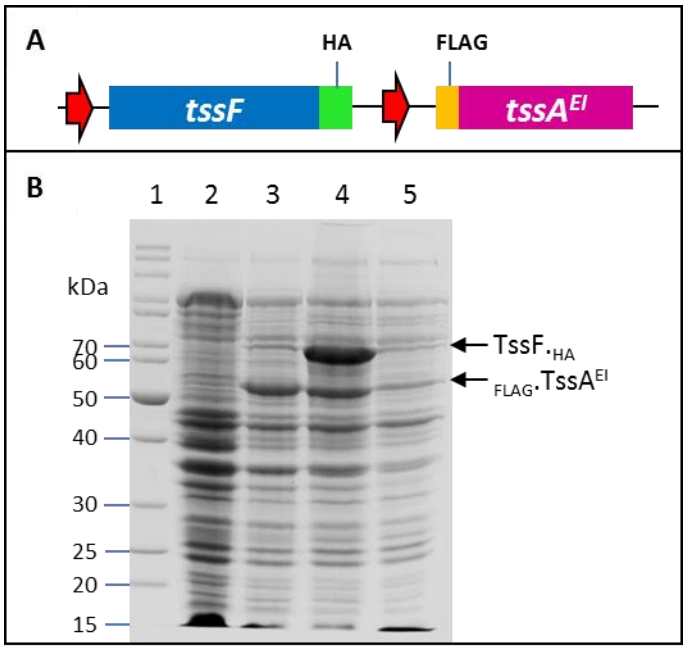


Figure 4.46 Co-overexpression and solubility of TssA^{EI} and TssF.$_{HA}$. A. Schematic drawing of the arrangement of FLAG tagged *tssA*^{EI} and *tssF*.HA in pACYCDuet-*tssF*.HA-FLAG.*tssA*^{EI}. The red arrows represent the two T7 promoters present in the plasmid. **B**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-*tssF*.HA-FLAG.*tssA*^{EI} plasmid in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 30^{o}C. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssF (~70 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

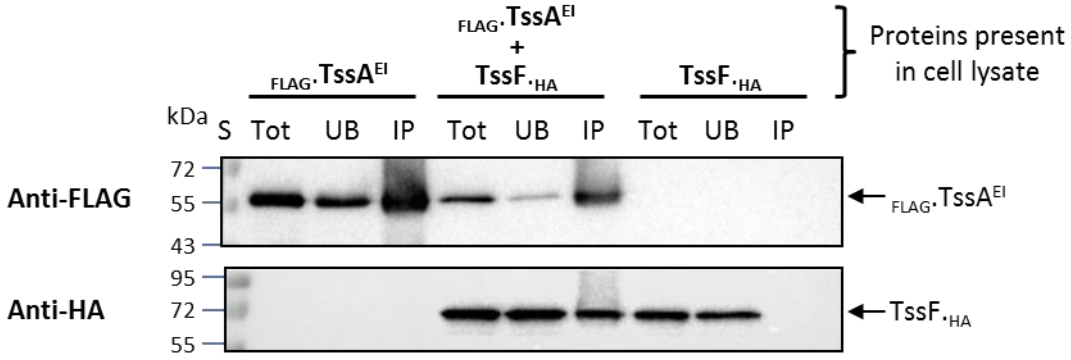


Figure 4.47 Co-IP analysis of the interaction between TssA^{EI} and C-terminal HA tagged. $_{FLAG}$.TssA^{EI} and TssF.$_{HA}$ were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{S}, pACYCDuet-*tssF*.HA and pACYCDuet-*tssF*.HA-FLAG.*tssA*^{EI}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^{EI}, upper panel) and anti-HA (TssF.$_{HA}$ lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssF (~70 kDa).

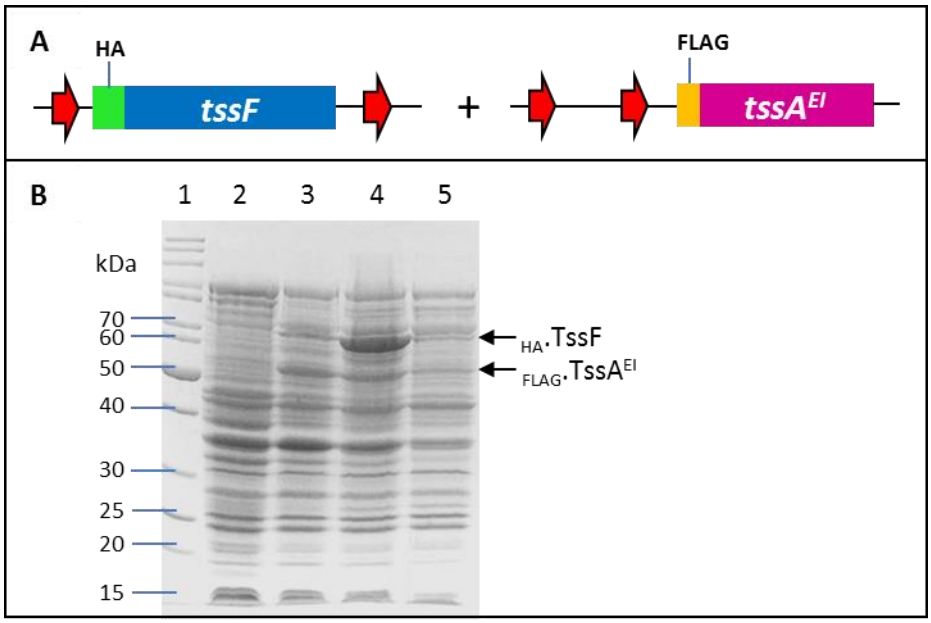


Figure 4.48 Co-overexpression and solubility of TssA^{EI} and N-terminal HA tagged TssF. A. Schematic drawing of co-expression of $_{FLAG}$.TssA^{EI} and $_{HA}$.TssF from two compatible plasmids pETDuet-HA.*tssF* and pACYCDuet-FLAG.*tssA^{EI}*. The red arrows represent two T7 promoters present in both plasmids. **B**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of the T7 promoters of *E. coli* strain BL21(λDE3) containing both pETDuet-HA.*tssF* and pACYCDuet-FLAG.*tssA^{EI}* plasmids with 1 mM IPTG at 30^{o}C. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssF (~70 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

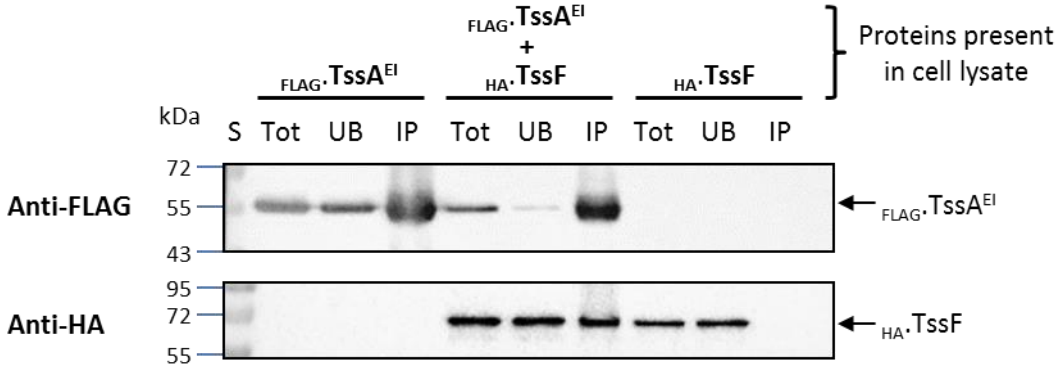


Figure 4.49 Co-IP analysis of the interaction between TssA^{EI} and N-terminal HA tagged TssF. $_{FLAG}$.TssA^{EI} and $_{HA}$.TssF were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pETDuet-HA.*tssF* and pACYCDuet-FLAG.*tssA*^{EI} separately and together, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^{EI}, upper panel) and anti-HA ($_{HA}$.TssF lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssF (~70 kDa).

4.4.2.6 Co-IP analysis of the interaction between $TssA^{EI}$ and $TssI_{gp27gp5}$

$TssA^{EI}$ has been shown to interact with $TssI_{gp27gp5}$ by the BACTH assay (Ahmad 2013), and the interaction between $TssA^{S}$ and $TssI_{gp27gp5}$ has been demonstrated both *in vivo* and *in vitro* in this study. To validate the *in vivo* result for the interaction between $TssA^{EI}$ and $TssI_{gp27gp5}$, a plasmid for co-expression of both proteins was constructed. $tssA^{EI}$ was amplified using primers, AHA1844.NdeI.FLAG.for and AHA1844.BglII.rev, in which a FLAG tag coding sequence was contained in the forward primer. The amplified $tssA^{EI}$ DNA fragment of the expected size (~1.4 kb) and pACYCDuet-VSVg.$tssI_{gp27gp5}$ (construction of this plasmid as described in Section 4.4.1.6) were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites in the primers and the second MCS of the plasmid. Ligation was carried out to allow the insertion of the $tssA^{EI}$ DNA fragment into plasmid for constructing pACYCDuet-VSVg.$tssI_{gp27gp5}$-FLAG.$tssA^{EI}$. The nucleotide sequence of the inserted gene was verified by DNA sequencing. A diagrammatic representation of the $_{VSVg}$.$TssI_{gp27gp5}$ and $_{FLAG}$.$TssA^{EI}$ co-expression plasmid is shown in Figure 4.50 A.

Protein overexpression was carried out in *E. coli* strain BL21(λDE3) cells by induction of the T7 promoters present on pACYCDuet-VSVg.$tssI_{gp27gp5}$-FLAG.$tssA^{EI}$ with 0.5 mM IPTG at 30^{o}C. Both proteins were overproduced in similarly large amounts, and were shown to present in the soluble fraction following cell lysis (Figure 4.50 B). Overexpression of $_{FLAG}$.$TssA^{EI}$ in the absence of $_{VSVg}$.$TssI_{gp27gp5}$ and overexpression of $_{VSVg}$.$TssI_{gp27gp5}$ in the absence of $_{FLAG}$.$TssA^{EI}$ are shown in Figure 4.37 (Section 4.4.2.1) and Figure 4.28C (Section 4.4.1.6), respectively. Co-IP was performed using anti-FLAG coupled beads. The results showed that the $_{VSVg}$.$TssI_{gp27gp5}$ did not bind to the beads in the absence of $_{FLAG}$.$TssA^{EI}$. In contrast, in the presence of $_{FLAG}$.$TssA^{EI}$, $_{VSVg}$.$TssI_{gp27gp5}$ was present in the immuno-precipitated material (Figure 4.51). This result suggested that $TssA^{EI}$ interacts with the core region of TssI.

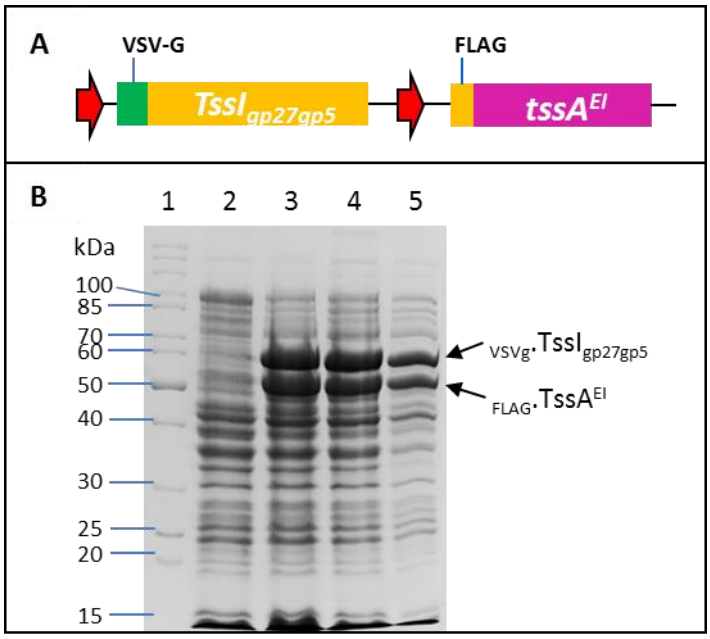


Figure 4.50 Co-overexpression and solubility of TssA^{EI} and TssI$_{gp27gp5}$. A. Schematic drawing of the arrangement of FLAG tagged *tssA^{EI}* and VSVg.*tssI.gp27gp5* in pACYCDuet-VSVg.*tssI.gp27gp5*-FLAG.*tssA^{EI}*. The red arrows represent the two T7 promoters present in the plasmid. **B**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssI.gp27gp5*-FLAG.*tssA^{EI}* in *E. coli* strain BL21(λDE3) with 0.1 mM IPTG at 30^{o}C. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssI$_{gp27gp5}$ (~61 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

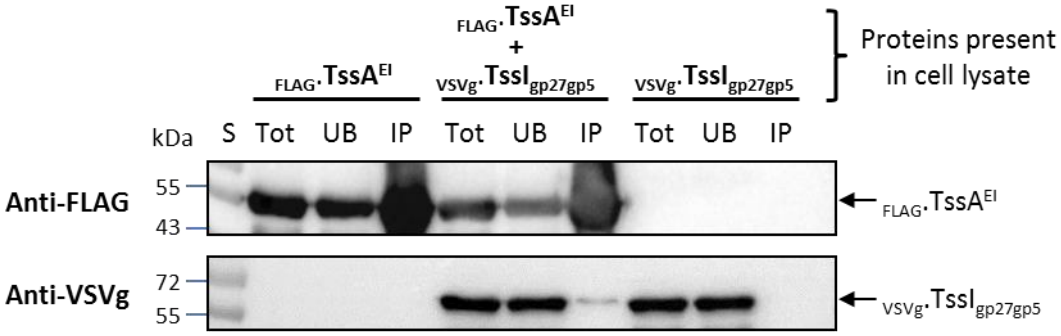


Figure 4.51 Co-IP analysis of the interaction between TssA^{EI} and TssI$_{gp27gp5}$. $_{FLAG}$.TssA^{EI} and $_{VSVg}$.TssI$_{gp27gp5}$ were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA^{EI}*, pACYCDuet-VSVg.*tssI.gp27gp5* and pACYCDuet-VSVg.*tssI.gp27gp5*-FLAG.*tssA^{EI}*, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4) and 500 mM NaCl. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^{EI}, upper panel) and anti-VSVg ($_{VSVg}$.TssI$_{gp27gp5}$ lower panel) monoclonal antibodies following transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssI$_{gp27gp5}$ (~61 kDa).

4.4.2.7 Co-IP analysis of the interaction between TssA^{EI} and TssK

The strategy employed for co-expression of both TssA^{EI} and TssK was similar to that used for TssA^{EI} and TssI$_{gp27gp5}$, which was to insert the *tssA^{E}* ORF into the second MCS of pACYCDuet that already contained the *tssK* gene inserted in the first MCS (pACYCDuet-VSVg.*tssK*, constructed as described in Section 4.2.2). The same pair of primers was employed for amplifying *tssA^{EI}* which contained *Nde*I and *Bgl*II restriction sites in the forward and reverse primer, respectively, and a FLAG tag coding sequence in the forward primer. Ligation, selection and identification of the desired recombinant plasmid, pACYCDuet-VSVg.*tssK*-FLAG.*tssA^{EI}* (Figure 4.52 A) were performed as described previously for constructing pACYCDuet-VSVg.*tssI$_{gp27gp5}$*-FLAG.*tssA^{EI}* (Section 4.4.2.6).

Co-expression of TssA^{EI} and TssK was conducted in *E. coli* strain BL21(λDE3) containing pACYCDuet-VSVg.*tssK*-FLAG.*tssA^{EI}* following induction of the T7 promoters with 0.5 mM IPTG at 30^{o}C. A large amount of protein of ~50 kDa was produced as determined by SDS-PAGE (Figure 4.52 B). However, as the molecular weights of TssA^{EI} (~53 kDa) and TssK (~51 kDa) are almost the same, it was difficult to establish which protein was overproduced or whether both of them were present. Therefore, the induced total protein from cells containing pACYCDuet-VSVg.*tssK*-FLAG.*tssA^{EI}* was subjected to western blotting and probed with both anti-FLAG and anti-VSVg monoclonal antibodies for detecting the presence of $_{FLAG}$.TssA^{EI} and $_{VSVg}$.TssK, respectively. The result confirmed the overproduction of both proteins (Figure 4.52 C and D). Overexpression of $_{FLAG}$.TssA^{EI} in the absence of $_{VSVg}$.TssK and overexpression of $_{VSVg}$.TssK in the absence of $_{FLAG}$.TssA^{EI} are shown in Figure 4.37 (Section 4.4.2.1) and Figure 4.31C (Section 4.4.1.7), respectively. Co-IP was carried out using anti-FLAG coupled beads. The results showed that $_{VSVg}$.TssK was co-precipitated with $_{FLAG}$.TssA^{EI} as it was present in the immuno-precipitated material along with TssA^{EI} and it was not bound to the beads in the absence of TssA^{EI} (Figure 4.53). The results suggested an interaction between these two proteins and support the observations from the BACTH assay where TssA^{EI} was shown to interact with TssK *in vivo* (Ahmad 2013).

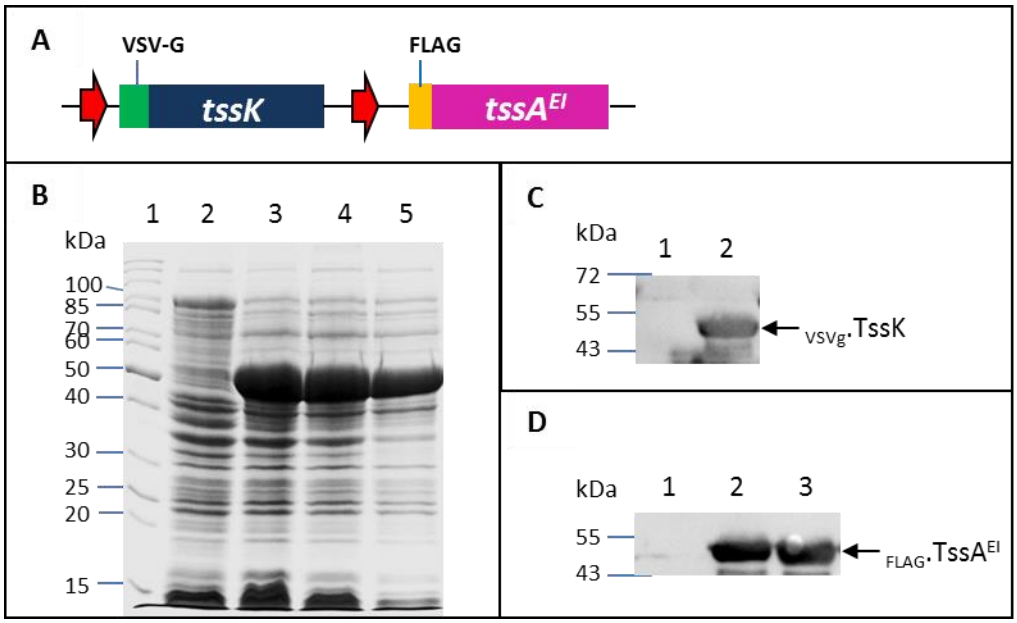


Figure 4.52 Co-overexpression and solubility of TssA^{EI} and TssK. A. Schematic drawing of the arrangement of FLAG tagged *tssA^{EI}* and VSVg.*tssK* in pACYCDuet-VSVg.*tssK*-FLAG.*tssA^{EI}*. The red arrows represent the two T7 promoters present in the plasmid. **B**. A coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssK*-FLAG.*tssA^{EI}* in *E. coli* strain BL21(λDE3) with 0.5 mM IPTG at 30^{o}C. The expected locations of TssA^{EI} and TssK are ~53 and ~51 kDa, respectively. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **C**. Immuno-detection of TssK. Cell lysate following induction centrifugation, proteins were electrophoresed by SDS-PAGE, and following western transfer, were probed with anti-VSVg mAb. Lane 1, total protein from BL21(λDE3) containing pACYCDuet-FLAG.*tssA^{EI}* following induction; lane 2, total protein from BL21(λDE3) containing pACYCDuet-VSVg.*tssK*-FLAG.*tssA^{EI}* following induction. The arrow indicates the expected location of TssK based on its size (~51 kDa). **D**. Immuno-detection of TssA^{E} expression. Cell lysate following induction centrifugation, proteins were electrophoresed by SDS-PAGE, and following western transfer, were probed with anti-FLAG mAb. The arrow indicates the expected location of TssA^{EI} based on its size (~53 kDa). Lane 1, total protein from BL21(λDE3) containing pACYCDuet-VSVg.*tssK* following induction; lane 2, total protein from BL21(λDE3) containing pACYCDuet-FLAG.*tssA^{EI}* following induction; lane 3, total protein from BL21(λDE3) containing pACYCDuet-VSVg.*tssK*-FLAG.*tssA^{EI}* following induction.

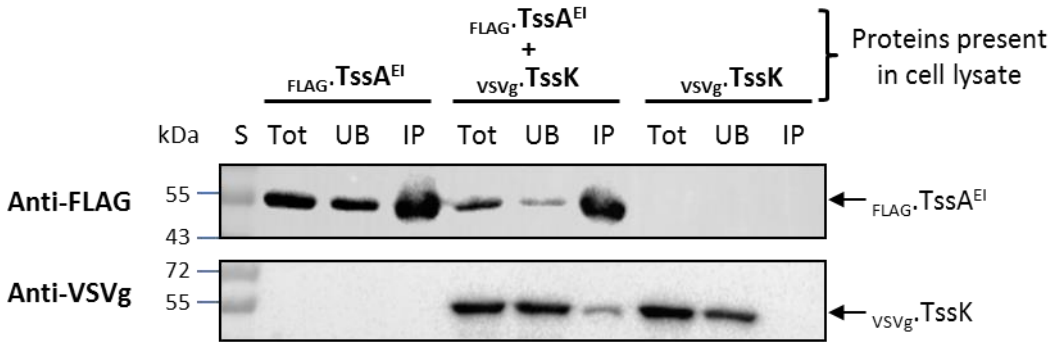


Figure 4.53 Co-IP analysis of the interaction between TssA^{EI} and TssK. $_{FLAG}$.TssA^{EI} and $_{VSVg}$.TssK were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{EI}, pACYCDuet-VSVg.*tssK* and pACYCDuet-VSVg.*tssK*-FLAG.*tssA*^{EI}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^{EI}, upper panel) and anti-VSVg ($_{VSVg}$.TssD lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssK (~51 kDa).

4.4.2.8 Co-IP analysis of the interaction between $TssA^{EI}$ and TssL

To construct a plasmid for co-expression of both $TssA^{EI}$ and TssL, *tssA*^{EI} was amplified using primers, AHA1844.BglII.FLAG.for and AHA1844.KpnI.rev, in which a FLAG tag coding sequence was contained in the forward primer. The amplified *tssA*^{EI} DNA fragment of the expected size (~1.4 kb) and the previously constructed plasmid pACYCDuet-VSVg.*tssL* (Section 4.4.1.8) were digested with restriction enzymes *Bgl*II and *Kpn*I that recognized sites in the primers and the second MCS of the plasmid. Ligation was carried out to insert the *tssA*^{EI} DNA fragment into the plasmid for constructing pACYCDuet-VSVg.*tssL*-FLAG.*tssA*^{EI} (Figure 4.54 A). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

Co-expression of both proteins was achieved in BL21(λDE3) following induction of the T7 promoters present on pACYCDuet-VSVg.*tssL*-FLAG.*tssA*^{EI} with 1 mM IPTG at 30^{o}C. Both proteins were produced in large amounts and remained in the soluble fraction following cell lysis (Figure 4.54 B). Overexpression of $_{FLAG}$.$TssA^{EI}$ in the absence of $_{VSVg}$.TssL and overexpression of $_{VSVg}$.TssL in the absence of $_{FLAG}$.$TssA^{EI}$ are shown in Figure 4.37 (Section 4.4.2.1) and Figure 4.33 C (Section 4.4.1.8), respectively. Co-IP was carried out using anti-FLAG affinity gel. The results showed that TssL did not bind to the anti-FLAG coupled beads in the absence of $TssA^{E}$. In contrast, when it was co-expressed with $TssA^{EI}$, TssL was detected in the immuno-precipitated material along with $TssA^{EI}$, indicating an interaction between these two proteins (Figure 4.48). The result is consistent with the observations from the BACTH assay on MacConkey-maltose agar where $TssA^{EI}$ and TssL were shown to interact with each other (Ahmad 2013).

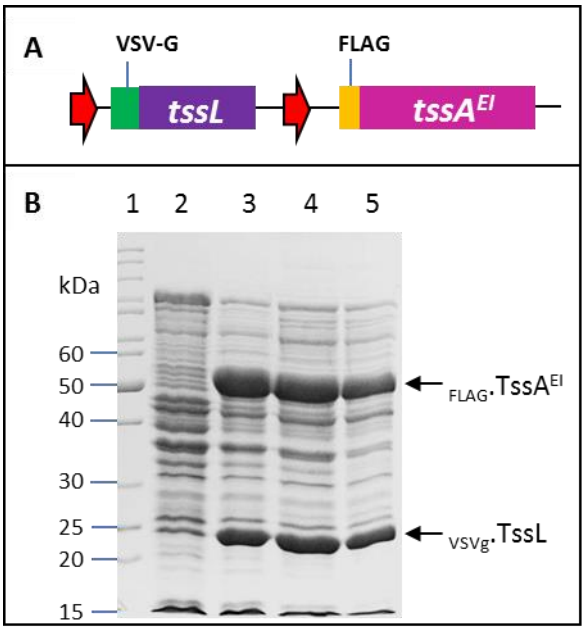


Figure 4.54 Co-overexpression and solubility of TssA^{EI} and TssL. A. Schematic drawing of the arrangement of FLAG tagged *tssA^{EI}* and VSVg.*tssL* in pACYCDuet-VSVg.*tssL*-FLAG.*tssA^{EI}*. The red arrows represent the two T7 promoters present in the plasmid. **D**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssL*-FLAG.*tssA^{EI}* in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 30^{o}C. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssL (~24 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

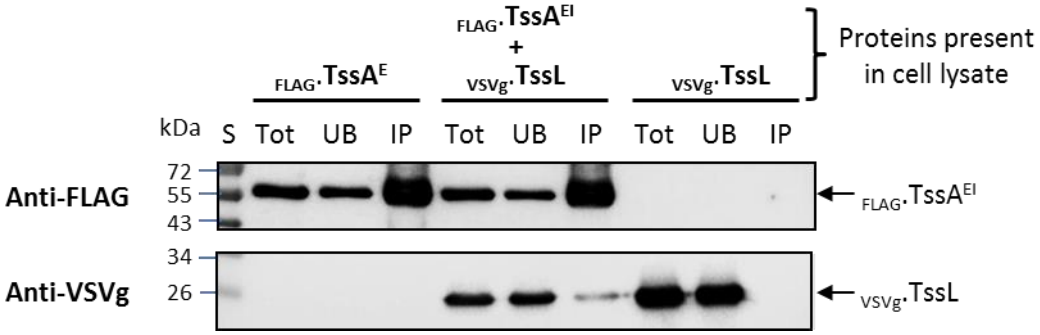


Figure 4.55 Co-IP analysis of the interaction between TssA^{EI} and TssL. $_{FLAG}$.TssA^{EI} and $_{VSVg}$.TssL were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{EI}, pACYCDuet-VSVg.*tssL* and pACYCDuet-VSVg.*tssL*-FLAG.*tssA*^{EI}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 500 mM NaCl and 0.2% Tween 20 The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 12% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}$.TssA^{EI}, upper panel) and anti-VSVg ($_{VSVg}$.TssL lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssL (~24 kDa).

4.4.2.9 Co-IP analysis of the interaction between $TssA^{EI}$ and $TssM_{NTD}$

The strategy for co-expressing $TssA^{EI}$ and $TssM_{NTD}$ was similar to that employed for $TssA^{EI}$ and $TssI_{gp27gp5}$ co-expression, which was to insert *tssA*^{EI} into the pACYCDuet vector that already contained the *tssM*$_{NTD}$ in the downstream of the first T7 promoter (pACYCDuet-VSVg.*tssM*$_{NTD}$, constructed as described in Section 4.4.1.9). The same pair of primers was employed for amplifying *tssA*^{EI} which contained *Nde*I and *Bgl*II in the forward and reverse primers, respectively, and a FLAG tag coding sequence in the forward primer for introducing a FLAG epitope tag at the N-terminus of *tssA*^{EI}. Ligation, selection and identification of the required recombinant plasmid, pACYCDuet-VSVg.*tssM*$_{NTD}$-FLAG.*tssA*^{EI} (Figure 4.56 A) were performed as described previously for constructing pACYCDuet-VSVg.*tssI*$_{gp27gp5}$-FLAG.*tssA*^{EI} (Section 4.4.2.6).

Co-expression of both proteins was carried out in *E. coli* strain BL21(λDE3) following induction of the T7 promoters present on pACYCDuet-VSVg.*tssM*$_{NTD}$-FLAG.*tssA*^{EI} with 0.1 mM IPTG at 22^{o}C, as induction with higher IPTG concentrations and temperatures produced smaller amounts of soluble $_{VSVg}$.$TssM_{NTD}$ (results not shown). Both proteins were overproduced in small amounts, and a smaller amount of $_{VSVg}$.$TssM_{NTD}$ was produced compared to that of $TssA^{EI}$ (Figure 4.56 B). Both proteins remained soluble following cell lysis. Overexpression of $_{FLAG}$.$TssA^{EI}$ in the absence of $_{VSVg}$.$TssM_{NTD}$ and overexpression of $_{VSVg}$.$TssM_{NTD}$ in the absence of $_{FLAG}$.$TssA^{EI}$ are shown in Figure 4.37 (Section 4.4.2.1) and Figure 4.35C (Section 4.4.1.9), respectively. Co-IP was carried out using anti-FLAG coupled beads. The results showed that there was a degradation product of $_{VSVg}$.$TssM_{NTD}$ around 40 kDa ($TssM_{NTD}$ is ~45 kDa). $_{VSVg}$.$TssM_{NTD}$ did not bind to the anti-FLAG coupled beads in the absence of $_{FLAG}$.$TssA^{EI}$, whereas there was a detectable amount of $_{VSVg}$.$TssM_{NTD}$ present in the immunoprecipitated material when it was co-expressed with $_{FLAG}$.$TssA^{EI}$ (Figure 4.57). Interestingly, the putative $_{VSVg}$.$TssM_{NTD}$ degradation product was not present in the co-immunoprecipitate, suggesting that part or the entire $TssA^{EI}$ interaction region is absent. The result suggested a specific interaction between $TssA^{EI}$ and $TssM_{NTD}$. However, the result is not consistent with the BACTH assay observations, where the maltose phenotype of all the pairwise combinations between $TssA^{EI}$ and $TssM_{NTD}$ was negative (Section 3.5.6).

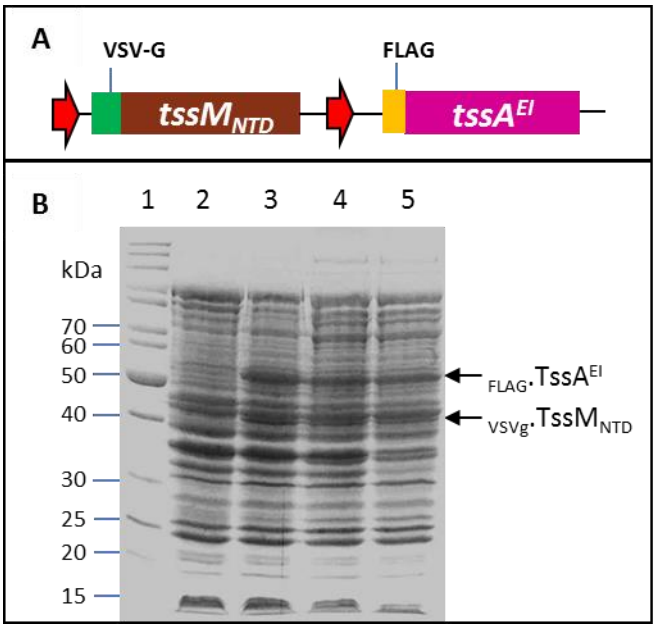


Figure 4.56 Co-overexpression and solubility of TssA^{EI} and TssM$_{NTD}$. A. Schematic drawing of the arrangement of FLAG tagged *tssA^{EI}* and VSVg.*tssM$_{NTD}$* in pACYCDuet-VSVg.*tssM$_{NTD}$*-FLAG.*tssA^{EI}*. The red arrows represent the two T7 promoters present in the plasmid. **B**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssM$_{NTD}$*-FLAG.*tssA^{EI}* plasmid in *E. coli* strain BL21(λDE3) cells with 0.1 mM IPTG at 22^{o}C. The arrows indicate the expected locations of TssA^{EI} (~53 kDa) and TssM$_{NTD}$ (~45 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

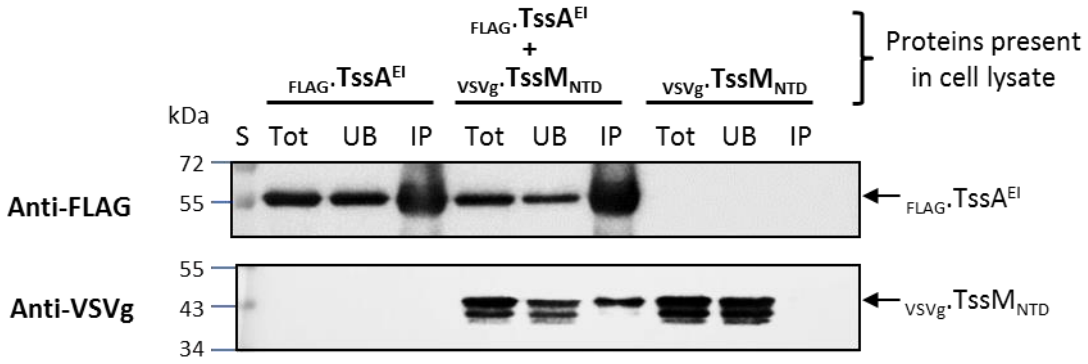


Figure 4.57 Co-IP analysis of the interaction between $TssA^{EI}$ and $TssM_{NTD}$. $_{FLAG}.TssA^{EI}$ and $_{VSVg}.TssM_{NTD}$ were expressed separately and together following transformation of *E. coli* strain BL21(λDE3) with pACYCDuet-FLAG.*tssA*^{EI}, pACYCDuet-VSVg.*tssM*$_{NTD}$ and pACYCDuet-VSVg.*tssM*$_{NTD}$-FLAG.*tssA*^{EI}, accordingly, and induction with IPTG. Soluble fraction of the cell lysates following induction of protein expression were immunoprecipitated with anti-FLAG-coupled beads in buffer containing 50 mM Tris-HCl (pH 7.4), 500 mM NaCl and 0.2% Tween 20. The total soluble material (Tot), unbound (UB) and the immunoprecipitated (IP) material were separated by electrophoresis in 10% SDS-PA gels and immuno-detected with anti-FLAG ($_{FLAG}.TssA^{EI}$, upper panel) and anti-VSVg ($_{VSVg}.TssM_{NTD}$ lower panel) monoclonal antibodies following western transfer. S, EZ-Run™ Prestained Rec Protein Ladder. The arrows indicate the expected locations of $TssA^{EI}$ (~53 kDa) and $TssM_{NTD}$ (~45 kDa).

4.5 Electron microscopy analysis of the interaction between $TssA^S$ and the core region of TssI

The interaction between $TssA^S$ and $TssI_{gp27gp5}$ has been demonstrated in BACTH assays and co-IP assays. $TssA^S$ was proposed to be a T6SS baseplate component based on the fact that it forms a ring-like structure and interacts with other baseplate subunits, components of the tail tube, and tail-spike, and the cell envelope anchoring subunits (Ahmad 2013). TssI is a trimeric protein that forms the tail spike sitting at the top of the tail tube consisting of stacked rings of TssD (Pukatzki et al. 2007; Hachani et al. 2011). The spike protein complex of phage T4 $(gp27gp5)_3$ forms the hub of the baseplate prior to DNA injection (Leiman and Shneider 2012). Therefore, the complex formed by $TssA^S$ and TssI may involve insertion of TssI into the ring-like structure formed by $TssA^S$. This may be observable by electron microscopy where TssI would occupy the space inside the $TssA^S$ ring. According to a recently proposed model, $TssA^E$ acts as a tail terminator protein (Zoued et al. 2016). However, in this model TssI may also be located within the TssA ring prior to tail tube and tail sheath polymerisation. To test this hypothesis, it was decided to purify $TssA^S$-TssI complexes by a pull-down in which $TssA^S$ N-terminally histidine tagged for purification using immobilized metal ion affinity chromatography (IMAC).

To co-express both proteins from a single plasmid (pACYCDuet), PCR amplification of $tssA^S$ was carried out from chromosomal DNA of *B. cenocepacia* H111 using the pair of primers TssA.NdeI.His10.for and pET14b-iotArev. The amplified $His_{10}.tssA^S$ DNA fragment of the expected size (~1.1 kb) was digested with restriction enzymes *Nde*I and *Bam*HI that recognized sites in the forward and reverse primers, respectively, and ligated between the *Nde*I and *Bgl*II sites located downstream of the second T7 promoter of pACYCDuet-VSVg.$tssI_{gp27gp5}$ (Section 4.4.1.6) for constructing pACYCDuet-VSVg.$tssI_{gp27gp5}$-$His_{10}.tssA^S$ (Figure 4.58 A). The nucleotide sequence of the inserted gene was verified by DNA sequencing.

Overexpression of N-terminal VSVg-tagged core TssI and deca-histidine-tagged $TssA^S$ from pACYCDuet-VSVg.$tssI_{gp27gp5}$-$His_{10}.tssA^S$ was conducted in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37°C. More $_{His10.}TssA^S$ was produced than $TssI_{gp27gp5}$ (Figure 4.58 B). Purification of $_{His10.}TssA^S$ from the clarified cell lysate was carried out using IMAC. A fraction of $TssA^S$ will be bound by the nickel resin and $TssI_{gp27gp5}$ was

also expectd to be retained as it has been shown to interact with $TssA^S$ by the BACTH assay and co-IP (Sections 3.3.9 and 4.4.1.6). The soluble fraction of the cell lysate following induction was applied to nickel-sepharose (1ml, GE Healthcare) and incubated with gentle mixing for 1.5 hours at 4^{o}C to allow proteins to interact. The nickel-sepharose was then washed with 20 volumes of binding buffer to remove non-specifically bound proteins and followed by elution with a 100, 150, 200, 300 and 500 mM imidazole step gradient. A fraction size of 4 ml was collected for each concentration of imidazole, which were analysed by SDS-PAGE (Figure 4.59 A). The result showed that a small fraction of $_{His10}$.$TssA^S$ (5-10%) was bound to the nickel-sepharose. It has been reported previously that histidine tagged $TssA^S$ is poorly bound to the resin during IMAC (Ahmad 2013). In contrast, the majority of $_{VSVg}$.$TssI_{gp27gp5}$ was retained on the column. However, in the elution fractions, $_{VSVg}$.$TssI_{gp27gp5}$ was mainly present in the 100 mM imidazole fraction, whereas $TssA^S$ was mainly eluted in the 200 mM imidazole fraction. In the fraction eluted with 150 mM imidazole, both proteins were present in a similar amount. The presence of $_{VSVg}$.$TssI_{gp27gp5}$ in the 150 mM imidazole elution fraction was confirmed by western blotting with immuno-detection by anti-VSVg monoclonal antibody (Figure 4.59 B). However, there was no $TssI_{gp27gp5}$ present in the 300 mM and 500 mM imidazole fractions analysed by western blotting (result not shown) even though there were visible amounts of $TssA^S$ on a coomassie blue-stained SDS-PA gel. This observation suggests the presence of $_{VSVg}$.$TssI_{gp27gp5}$ in the elution fractions was likely to be due to non-specific binding to the nickel-sepharose. The 150 mM imidazole elution fraction was further analysed by negative stain transmission electron microscopy (TEM) to determine whether any $TssA^S$-TssI complexes had formed (Department of Molecular Biology and Biotechnology, University of Sheffield). The result showed that $TssA^S$ formed the ring-like structure as expected. In some cases, high density material was present inside the rings. However, given that $TssI_{gp27gp5}$ is a trimer with molecular weight ~180 kDa, this is not consistent with the size of the particles located inside some of the $TssA^S$ rings, i.e. the high density particles were too small to correspond to a trimer of the core TssI subunit (Figure 4.59 C).

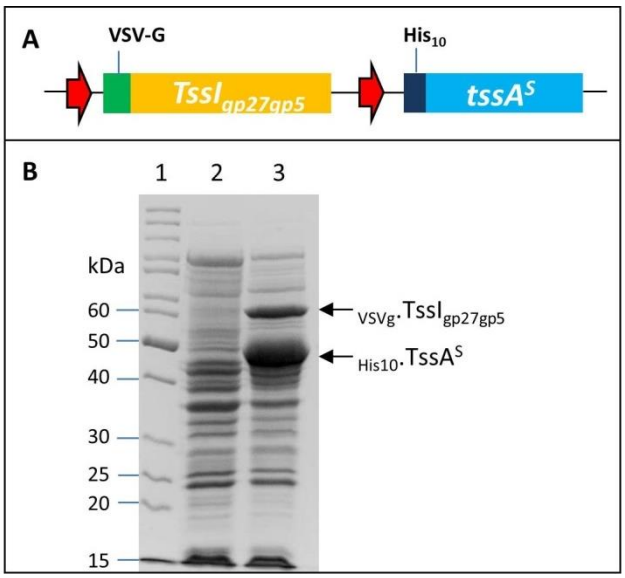


Figure 4.58 Co-expression of $_{His10}$.TssA^S and $_{VSVg}$.TssI$_{gp27gp5}$. A. Schematic drawing of the arrangement of His$_{10}$.*tssA*^{S} and VSVg.*tssI.gp27gp5* in pACYCDuet-VSVg.*tssI.gp27gp5*-His$_{10}$.*tssA*^{S}. The red arrows represent the two T7 promoters present in the plasmid. **B**. A Coomassie blue-stained 10% SDS-PA gel showing the analysis of proteins following induction of pACYCDuet-VSVg.*tssI.gp27gp5*-His$_{10}$.*tssA*^{S} in *E. coli* strain BL21 (λDE3) cells with 1 mM IPTG at 37^{o}C. The arrows indicate the expected locations of TssA^S (~42 kDa) and TssI$_{gp27gp5}$ (~60 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction.

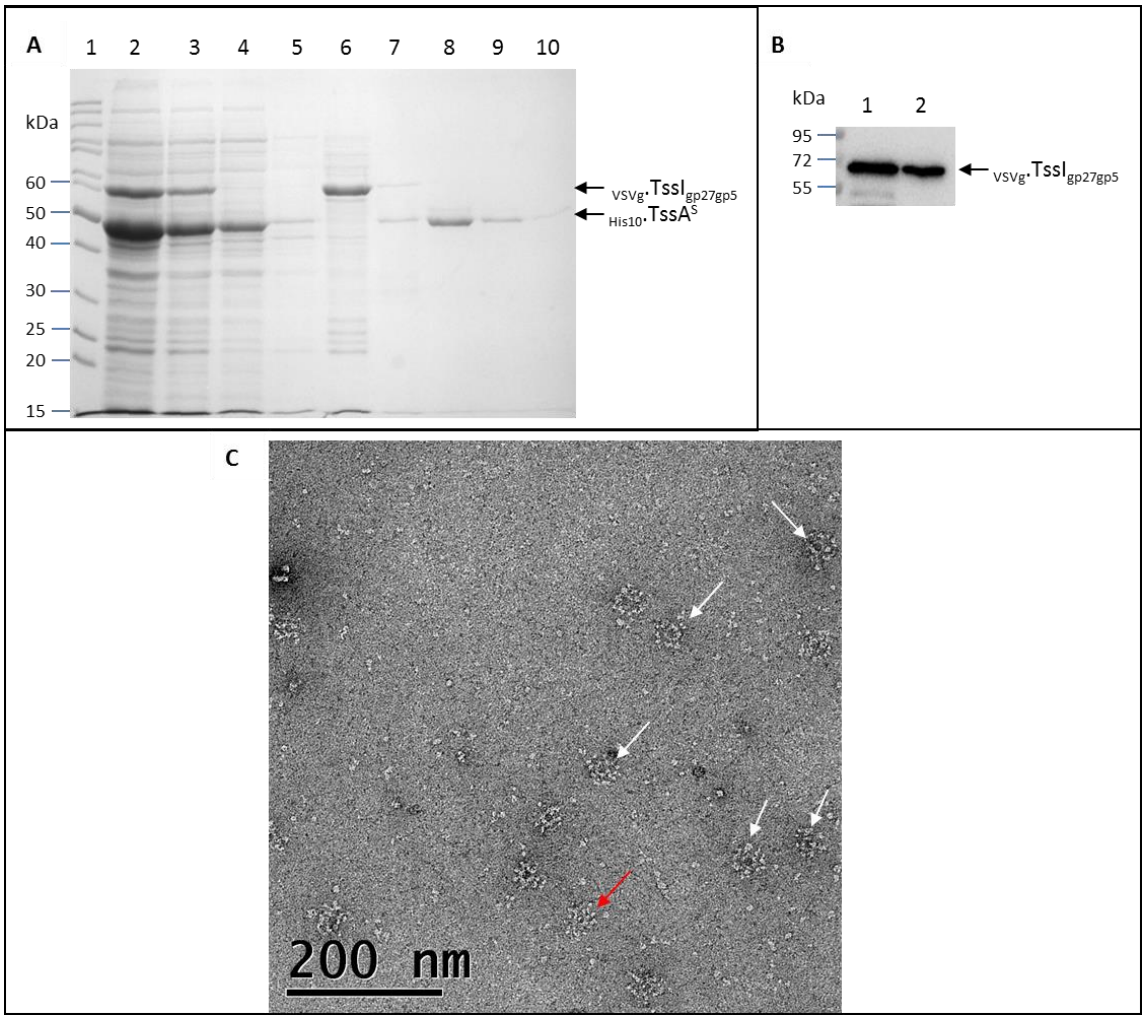


Figure 4.59 Co-purification of $TssA^S$ and $TssI_{gp27gp5}$ using IMAC and analysis by TEM. A. Coomassie blue-stained 10% SDS-PA gel showing the analysis of attempted co-purification of $_{His10}$.$TssA^S$ and $_{VSVg}$.$TssI_{gp27gp5}$ using IMAC. Proteins were overexpressed in *E. coli* strain BL21(λDE3) containing pACYCDuet-VSVg.*tssI.gp27gp5*-His_{10}.$tssA^S$ by induction with 1 mM IPTG at 37°C. The soluble fraction of the cell lysate was applied to the nickel-sepharose in buffer containing 50 mM Tris-HCl (pH 7.4) and 500 mM NaCl. Following washing, elution was performed in the same buffer with increasing concentrations of imidazole upto 500 mM. The arrows indicate the expected locations of $TssA^S$ (~42 kDa) and $TssI_{gp27gp5}$ (~60 kDa). Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 3, soluble fraction of cell lysate following induction; lane 4, flow-through (unbound material) of nickel-sepharose column; lane 5, wash of nickel-sepharose column with 20 volumes of binding buffer; lanes 6-10, elution fractions with increasing concentrations of imidazole: 100, 150, 200, 300 and 500 mM. **B**. Immuno-detection of $_{VSVg}$.$TssI_{gp27gp5}$ from attempted co-purification with $_{His10}TssA^S$ by IMAC with anti-VSVg monoclonal antibody. The arrow indicates the expected location of $TssI_{gp27gp5}$ based on its size (~60 kDa). Lane 1, soluble fraction of cell lysate containing pACYCDuet-VSVg.*tssI.gp27gp5*-His_{10}.$tssA^S$ following induction; lane 2, elution fraction of IMAC with 150 mM imidazole. **C**. Elution fraction of IMAC with 150 mM imidazole was analysed by negative staining electron microscopy with 0.75% uranyl formate. White arrows point to representative ring-like structures formed by $TssA^S$ (or potencially with $TssI_{gp27gp5}$ present pointed by red arrow).

4.6 Discussion

Using the BACTH assays, both types of TssA ($TssA^S$ from *B. cenocepacia* and $TssA^{EI}$ from *A. hydrophila*) were suggested to play an important role in the assembly of the T6SS as they interact with many other T6SS subunits from *B. cenocepacia*, i.e. TssB, -C, -D, -F, -H, -I, -K and -L. In addition, interactions of $TssA^S$ with TssE and $TssM_{NTD}$ were also observed (Chapter 3). Although $TssA^{EI}$ does not appear to interact with TssE in the two hybrid assay (Ahmad 2013), the interaction between $TssA^E$ and TssE has been shown when both subunits were derived from the same T6SS, i.e. T6SS-1 of enteroaggregative *E. coli* (EAEC) (Zoued et al. 2016). Nickel affinity pull-down and co-immunoprecipitation assays were employed in order to provide biochemical evidence to support the *in vivo* results suggesting interactions between both types of TssA and TssC, -D, -F, -I, -K and -L. The *in vitro* interactions of both types of TssA with TssE and $TssM_{NTD}$ were also investigated. Although the maltose-phenotypes of the combinations of the tail sheath protein TssB with $TssA^S$ or $TssA^{EI}$ were weak, in contrast, much stronger maltose phenotypes were observed for combinations of $TssA^S$ with the other tail sheath protein TssC. Therefore, the interactions between TssB and $TssA^S$/$TssA^{EI}$ *in vitro* were not tested. TssH is an AAA^+ chaperone that has been shown to play a crucial role in T6SS assembly, and it interacts with almost all of the other T6SS subunits based on the BACTH assay (Jones 2012; Ahmad 2013). Therefore, its interactions with $TssA^S$ and $TssA^{EI}$ *in vitro* were also not investigated. A summary of *in vivo* and *in vitro* interactions of $TssA^S$ and $TssA^{EI}$ with other T6SS subunits is shown in Table 4.1.

4.6.1 Nickel affinity pull-down assays did not demonstrate interactions between $TssA^S$ and TssI, TssK or TssL

By using a nickel affinity pull-down assay, the interactions between $TssA^S$ and TssI, TssK, and TssL were analysed. $_{LinkerHis6}$.$TssA^S$ was used as the bait in the assay, since it has been shown previously that low retention of $TssA^S$ on nickel sepharose occurs when it was histidine-tagged at either the N-terminus (with 6 or 10 histidines) or at the C-terminus (with 10 histidines), and the flexible interdomain linker region of $TssA^S$ is predicted to be exposed on the exterior of the protein as it is sensitive to protease (Ahmad, 2013). The histidine tag that was introduced into the linker region of $TssA^S$ did

slightly increase the retention of $TssA^S$ on the nickel sepharose (~10%) compared to the N- or C-terminal tagged $TssA^S$. However, the candidate prey proteins that were tagged by a VSV-g epitope tag were observed to bind to the nickel sepharose non-specifically. Moreover, there was no obvious increase in the amount of the prey protein retained on the nickel sepharose when the bait protein $TssA^S$ was present. Therefore, no evidence for the interaction between $TssA^S$ and TssI, TssK, and TssL was obtained from the nickel affinity pull-down assays. This might be a result of the influence of pH on protein-protein interactions due to alteration of the electrostatic properties of the protein surfaces (Dumetz et al. 2008).

It has been previously shown that $TssA^S$ and TssK interact with TssL by BACTH assay (Ahmad 2013). This raised the possibility that the $TssA^S$-TssK interaction may be more stable in the presence of TssL. In support of this idea, the presence of TssL gave rise to a stronger maltose phenotype for the interaction between $TssA^S$ and TssK, compared to the maltose phenotype without TssL (Section 3.4.3). A pull-down assay was therefore conducted to investigate whether the interaction between $TssA^S$ and TssK could be enhanced by the presence of TssL. However, the nickel pull-down assay did not show any increased retention of TssK or TssL when all three proteins were present. This may be because of the possibility that the histidine tag on $TssA^S$ was masked or sequestered as a result of interacting with TssK or TssL. This possibility can be overcome by co-IP with a polyclonal antibody againt $TssA^S$.

4.6.2 Co-IP using anti-$TssA^S$ antibody failed to reveal the interactions between $TssA^S$ and TssK or TssL when all three proteins were present

Due the possibility that the Histidine-tag on $TssA^S$ was masked by the binding of TssK, attempt was made to carry out co-IP with polyclonal $TssA^S$ antiserum to demonstrate an interaction between native $TssA^S$ and TssK. However, $TssA^S$, TssK and TssL were all found to non-specifically bind to the protein A sepharose resin in the absence of the anti-$TssA^S$ antibody. In order to decrease the non-specific binding problem, modifications of the protocol were made by extending wash times and blocking the sepharose resin by BSA which can reduce the non-specific hydrophobic binding of proteins. Despite the fact that hydrophobic interactions are important in epitope-antibody binding, such interactions also facilitate non-specific binding to the

matrix. Most proteins have some extent of hydrophobicity because of the non-polar side chains of several amino acids. However, despite the use of BSA, the problem was not resolved.

4.6.3 Co-IP using anti-FLAG affinity gel demonstrated specific interactions between both types of TssA and many other T6SS subunits

As the previous two approaches did not demonstrate an interaction between $TssA^{S}$ and other T6SS subunits, co-IP with anti-FLAG coupled beads was carried out by tagging the N-terminus of both types of TssA with the FLAG tag. Other T6SS subunits were also tagged by epitope tags other than the FLAG tag for immuno-detection by commercially available antibodies in western blotting. Both proteins were co-expressed from *E. coli* BL21(λDE3) cells containing a single plasmid in all cases, except for the co-expression of TssA and TssE, and TssA and TssF, where two compatible plasmids were used. The protein expression and solubility were tested in various induction conditions for those subunits with low solubility. By using the anti-FLAG coupled beads, $TssA^{S}$ was shown to interact with TssC (tail sheath), TssD (tail tube), TssE (baseplate component), TssF (baseplate component), $TssI_{gp27gp5}$ (the core region of the tail spike protein), TssK (baseplate component), TssL (component of cell evolope chamber) and $TssM_{NTD}$ (component of cell evolope chamber), that is consistent with the findings from the BACTH assay.

$TssA^{EI}$ was demonstrated to interact with all the $TssA^{S}$-interacting subunits *in vitro* except TssD (this study), although $TssA^{EI}$ and TssD have been previously shown to give rise to a positive maltose phenotype *in vivo* by the BACTH assay (Ahmad 2013). In addition, the *in vitro* interaction between EAEC-derived $TssA^{E}$ and TssD has been demonstrated by surface plasmon resonance (Zoued et al. 2016). As the $TssA^{EI}$ and TssD subunits tested in this study were derived from different organisms (*A. hydrophila* and *B. cenocepacia*), it possible that these two components do not interact as $TssA^{S}$ and TssD due to differences in interacting surfaces of $TssA^{S}$ and $TssA^{EI}$. In addition, TssD is a chaperone for T6SS-effectors and so may be T6SS-specific. However, if this was the case, the *in vivo* interaction between $TssA^{EI}$ and TssD could be a false positive result or an experimental error of the BACTH assay.

Although no *in vivo* interactions of $TssA^{EI}$ with TssE and $TssM_{NTD}$ were suggested by the BACTH assay (Ahmad 2013; this study), their *in vitro* interactions were subsequently demonstrated by co-IP. This result is in agreement with a recent study in which the interaction between $TssA^{E}$ and $TssM_{NTD}$ is not detecable *in vivo* but detectable *in vitro* in the presence of TssL (Zoued et al. 2016). In addition, the interaction of $TssA^{E}$ with TssE was shown both *in vivo* and *vitro* (Zoued et al. 2016). The undetectable *in vivo* interactions between $TssA^{EI}$ and TssE or $TssM_{NTD}$ could be because the assay involving scoring the maltose phenotype lacks sensitivity. Another possibility is that the very large 'tags' involved with the BACTH assay (T18 and T25) may interfere with the interaction whereas the smaller epitope tag used for the co-IP assay for $TssA^{EI}$ and $TssM_{NTD}$ may not interfere. However, an interaction between $TssA^{S}$ and $TssM_{NTD}$ was detected by the BACTH assay where the CyaA T25 or T18 tags did not interfere, and the *in vitro* interaction between $TssA^{EI}$ and TssE is detectable although TssE was fused to MBP. A possible explanation could be due to differences in the MD and CTDs of $TssA^{S}$ and $TssA^{EI}$ and the location of the CyaA fragments may interfere with one interaction with $TssM_{NTD}$ and not the other.

The *in vitro* interaction between $TssA^{S}$ and TssF was only detected when the HA tag was fused to the C-terminus of TssF. Although the BACTH assay observations on MacConkey-maltose agar did not indicate which end of TssF is involved in the interaction with $TssA^{S}$, the results from the β-galactosidase assay suggested the interaction between $TssA^{S}$ and TssF was more stable when the N-terminus of TssF was free. A tentative conclusion was drawn that the N-terminus of TssF is involved in the interaction with $TssA^{S}$. Indeed, in the TssF domain analysis with $TssA^{S}$, the N-terminal domain of TssF was shown to interact with $TssA^{S}$ by the BACTH assay on MacConkey-maltose agar (Shastri 2011). However, the location of the epitope tag on TssF did not appear to markedly influence the interaction with $TssA^{EI}$ in the co-IP assay. $TssA^{EI}$ interacted with TssF *in vitro* irrespective of whether the N- or C- terminus was tagged. This is consistent with the BACTH assay observations on MacConkey-maltose agar where the interaction between $TssA^{EI}$ and TssF did not indicate the resquirement for a particular end of TssF to be free. This may be due to differences between $TssA^{S}$ and $TssA^{EI}$ in the conformation they adopt in the interaction with TssF, which resulted in the small HA epitope tag not affecting the interaction between $TssA^{EI}$ and TssF. The β-galactosidase assay can be carried out for the combinations of $TssA^{EI}$ with TssF to

determine whether a stable β-galactosidase activity can be obtained for one or more of the two-hybrid combinations that yielded a positive maltose phenotype.

In conclusion, based on the analysis, $TssA^{S}$ and $TssA^{EI}$ behaved similarly in interacting with respect to their interactions with other T6SS subunits *in vitro*, i.e. TssC, TssE, TssF, TssI, TssK, TssL, $TssM_{NTD}$. In addition, $TssA^{S}$ interacts with one other T6SS subunit (i.e. TssD) that $TssA^{EI}$ does not appear to interact with based on the *in vitro* analysis.

Table 4.1 Summary table of *in vivo* and *in vitro* interactions of $TssA^{S}$ and $TssA^{EI}$ with other T6SS subunits

	$TssA^{S}$ *in vivo* (i.e. BACTH assay)	$TssA^{S}$ *in vitro* (i.e. co-IP assay)	$TssA^{EI}$ *in vivo* (i.e. BACTH assay)	$TssA^{EI}$ *in vitro* (i.e. co-IP assay)
TssB	✓	ND	✓	ND
TssC	✓	✓	✓	✓
TssD	✓	✓	✓	X
TssE	✓	✓ (MBP fusion)	X	✓ (MBP fusion)
TssF	✓	N-terminal HA tagged X	✓	N-terminal HA tagged ✓
		C-terminal HA tagged ✓		C-terminal HA tagged ✓
TssG	X	ND	X	ND
TssH	✓	ND	✓	ND
TssI	✓	✓ (gp27gp5-like core region)	✓	✓ (gp27gp5-like core region)
TssJ	X	ND	X	ND
TssK	✓	✓	✓	✓
TssL	✓	✓	✓	✓
$TssM_{NTD}$	✓	✓	X	✓
$TssM_{CTD}$	X	ND	X	ND

$TssA^{S}$ from *B. cenocepacia* and $TssA^{EI}$ from *A. hydrophila* were employed in this study. All other subunits were from *B. cenocepacia*. '✓' interaction detected; 'X' interaction not detected; 'ND' not done. The results labelled in red are from a previous study (Ahmad 2013).

Chapter 5 Characterisation of TssA^{S}

5.1 Introduction

TssA^S is proposed to play an essential role in the assembly of the T6SS, as it has been demonstrated to interact with many other T6SS subunits by the bacterial two- and three-hybrid analysis (Chapter 3). These *in vivo* interactions between TssA^S and other T6SS subunits were subsequently confirmed *in vitro* (Chapter 4). Although evidence has been provided by Shastri (2011) and Ahmad (2013) that TssA^S is able to oligomerise into large complexes, its structure and function are still unknown. In this chapter, an investigation into the structure of TssA^S is presented and discussed.

5.2 Structural analysis of TssA^S

5.2.1 Overproduction and purification of TssA^S

Both linker hexa-histidine tagged TssA^S and untagged TssA^S were used for facilitating these studies, due to the fact that the N- or C-terminal histidine tagged TssA^S derivatives were not efficiently retained on a HisTrap column during IMAC (~5% retention), as previously discussed in Section 4.2. It was thought that this may be due to the N- and C-terminus of TssA^S being sequestered in the centre of the folded TssA^S which might prevent the access of the histidine tag to nickel beads. The linker region of TssA^S that links its N-terminal and C-terminal domains is susceptible to cleavage by proteases (Ahmad 2013), suggesting that it is exposed on the exterior of TssA^S. Hence, introduction of a histidine tag into the linker region is likely to lead to the tag being exposed, thereby facilitating its binding to the nickel sepharose. Furthermore, the length and amino acid sequence of this region is not conserved among TssA^S orthologues, and therefore it might tolerate insertion of a hexa-histidine tag without loss of function. For this reason, a plasmid was constructed (pACYCDuet-linkerHis$_6$.*tssA*^{S}) that overproduced TssA^S with a hexa-histidine tag located between amino acids 262 and 263 of TssA^S which is predicted to be the interdomain linker region (Hall and Thomas, unpublished results). Introducing a hexa-histidine tag in the linker region of TssA^S did not cause conformational change of TssA^S overall structure based on TEM analysis (Section 5.2.2) and did not negatively interfere with its interaction with other tested T6SS subunits as judged by BACTH assay, i.e. TssK and TssL (Section 4.2.1).

5.2.1.1 Overproduction and purification of $_{LinkerHis6}$.TssA^{S}

$_{LinkerHis6}$.TssA^{S} was overproduced from *E. coli* strain BL21(λDE3) cells containing pACYCDuet-linkerHis$_6$.*tssA*^{S}. Insertion of six histidine codons in the linker region of TssA^{S} was achieved by SOE-PCR (Splicing by Overlap Extension PCR), following which the amplified product was digested with *Nde*I and *Bam*HI restriction enzymes and ligated into pACYCDuet between the *Nde*I and *Bgl*II restriction sites which generated a hybrid *Bam*HI/*Bgl*II site (Hall and Thomas, unpublished results). Following induction with IPTG as described in Section 2.4.1, a protein with the expected size of $_{linkerHis6}$.TssA^{S} was overproduced as observed by SDS-PAGE (Figure 5.1 A). The overproduced protein was found largely in the soluble fraction following cell lysis based on SDS-PAGE analysis (Figure 5.1 B).

A large scale bacterial cell culture of *E. coli* strain BL21(λDE3) cells containing pACYCDuet-linkerHis$_6$.*tssA*^{S} was prepared. Purification of $_{LinkerHis6}$.TssA^{S} was achieved by immobilized metal ion affinity chromatography (IMAC). Soluble cell lysate containing overproduced $_{linkerHis6}$.TssA^{S} was applied onto a HisTrap HP column. The flow-through was recycled by re-applying onto the column to maximize the binding of $_{LinkerHis6}$.TssA^{S}. Upon SDS-PAGE analysis, it was apparent that there was no significant decrease in the amount of $_{LinkerHis6}$.TssA^{S} in the flow-through compared to that present in the loaded material (Figure 5.1 B). This may be due to abundant material was used for purification and its poorly retention (~10%) by the nickel resin (Section 4.2.2). Moreover, the recycle step did not facilitate increased $_{LinkerHis6}$.TssA^{S} binding to the column. The column was then washed with 20 volumes of the binding buffer before the elution step. $_{LinkerHis6}$.TssA^{S} was eluted in the range 167-233 mM imidazole with a single peak observed according to the UV trace at 280 nm (result not shown). The peak fractions were judged to be ≥80% pure by SDS-PAGE analysis.

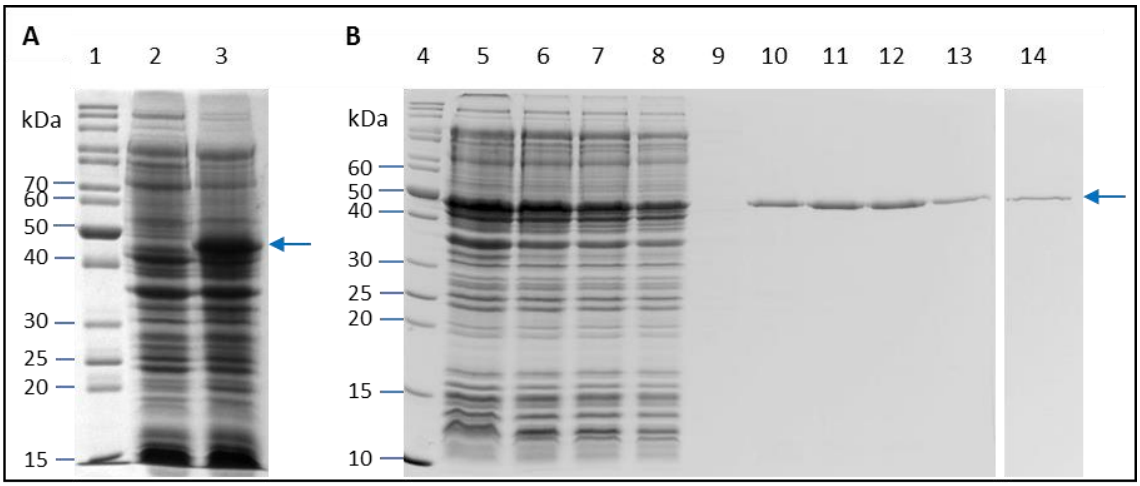


Figure 5.1 Overexpression, solubility and purification of $_{LinkerHis6}$.TssA^{S}. Coomassie blue-stained 10% (**A**) and 12% (**B**) SDS-PA gels showing the induction of pACYCDuet-linkerHis$_6$.*tssA*^{S} in *E. coli* strain BL21 (λDE3) cells with 1 mM IPTG at 30^{o}C. Following induction, the soluble fraction of the cell lysate was subjected to IMAC using a HisTrap HP column (1 ml, GE Healthcare) in a buffer containing 50 mM Tris-HCl, 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. The arrow indicates the expected locations of $_{LinkerHis6}$.TssA^{S} (~42 kDa) based on its MW. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, EZ-Run^{TM} Rec protein ladder (Fisher); lane 5, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 6, soluble fraction of cell lysate following induction; lane 7, flow-through_1 from HisTrap column; lane 8, flow-through_2 from HisTrap column; lane 9, wash of protein-loaded HisTrap column before elution step; lanes 10-14, HisTrap fractions corresponding to the peak of the UV trace at 280 nm.

5.2.1.2 Overproduction and purification of native TssA^S

Native TssA^S was produced in *E. coli* strain BL21(λDE3) cells containing plasmid pET14b-*tssA*^{S}. pET14b-*tssA*^{S} was constructed by cloning *tssA*^{S} from *B. cenocepacia* strain H111 into pET14b between the *Xba*I and *Bam*HI restriction sites, resulting in removal of the histidine tag and thrombin cleavage site coding sequences from the plasmid (Ahmad 2013). Following induction of the phage T7 promoter contained on the plasmid, a protein with the expected size of TssA^S (~42 kDa) was overproduced upon analysis by SDS-PAGE, and TssA^S remained in the soluble fraction following cell lysis (Figure 5.2 A). Bacterial cell culture of BL21(λDE3) cells containing pET14b-*tssA*^{S} was scaled up from 50 ml for the solubility analysis to 1 litre for TssA^S purification.

To purify TssA^S, nucleic acids were removed from the clarified cell lysate by precipitating with 5% polyethyleneimine (PEI) at pH 8.0. After clarification, ammonium sulfate was added to the supernatant to a final concentration of 30% to precipitate TssA^S (Section 2.4.3). The precipitated protein was recovered by centrifugation, and the protein pellet was resuspended in a small volume of buffer containing 50 mM Tris-HCl (pH 8.0) and 50 mM NaCl, where upon it was subjected to SEC on a Superose 6 GL column equilibrated with the same buffer (Section 2.4.8). As the BACTH assay results suggested TssA^S self-interacts (Section 3.3.1), the usage of a size exclusion chromatography also facilities an estimation of the number of monomers in the TssA^S oligomer (although this is only likely to be accurate for globular complexes). According to the elution profile, the UV trace at 280 nm shows a single peak eluting at ~13.75 minute (Figure 5.2 B), which corresponds to a protein size about 1,500 kDa calculated using the calibration curve. The results indicated that TssA^S is able to form a large complex (~36 subunits), and this is consistent with a previous observation (Ahmad 2013). The purity of TssA^S was judged to be >95% upon SDS-PAGE analysis (Figure 5.2 C). A more accurate size determination of TssA^S was carried out by SEC-MALS in Section 5.2.3.

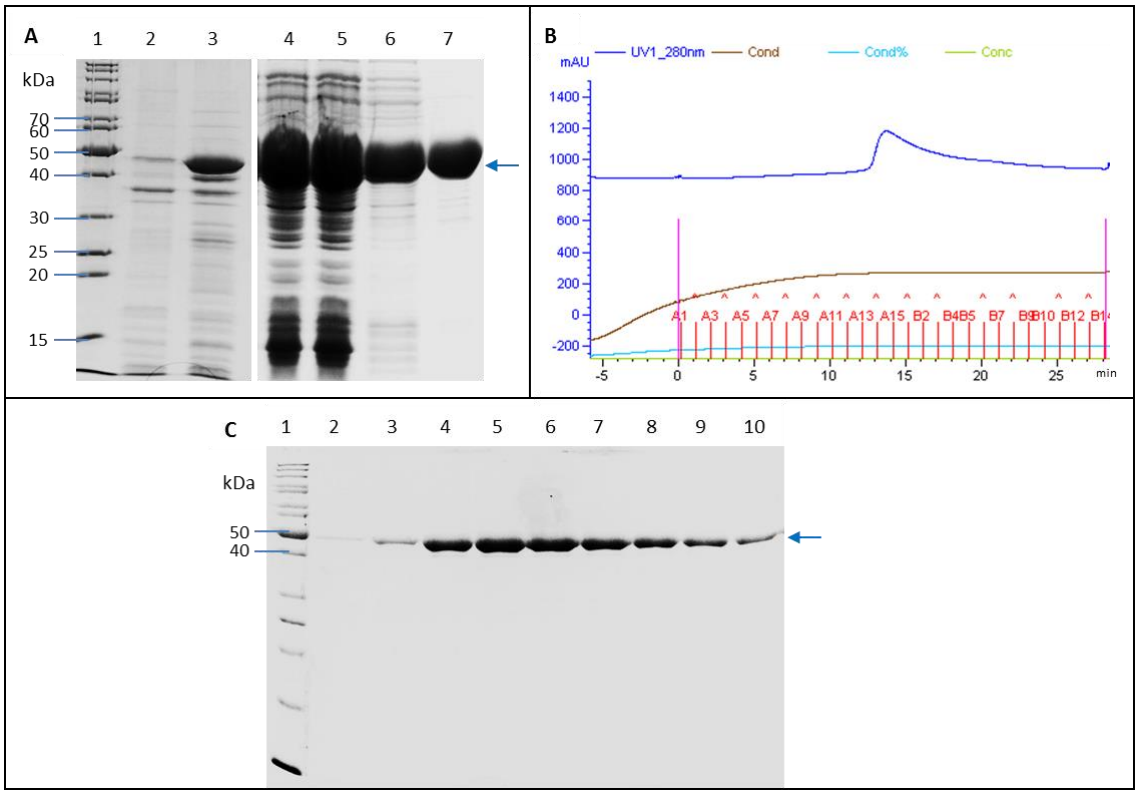


Figure 5.2 Overproduction, solubility and purification of untagged TssA^{S}. TssA^{S} synthesis was induced from pET14b-*tssA*^{S} in *E. coli* strain BL21 (λDE3) with 1 mM IPTG at 37^{o}C. Following induction, the soluble fraction of the cell lysate containing TssA^{S} was subjected to PEI precipitation, 30% ammonium sulphate precipitation and SEC to purify TssA^{S}. **A**. Overexpression and solubility of TssA^{S} followed by PEI precipitation and ammonium sulphate precipitation analysed by electrophoresis in 12% SDS-PA gels. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction; lane 6, supernatant after PEI precipitation; lane 7, resuspended pellet after 30% ammonium sulphate precipitation. Arrow indicates the expected location of TssA^{S} based on its molecular weight (~42 kDa). **B**. Elution profile of TssA^{S} on Superose 6 GL. The UV trace at 280 nm is shown in blue. **C**. Purification of TssA^{S} by SEC on a Superose 6 GL column analysed by electrophoresis in a 12% SDS-PA gel. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 2-10, peak fractions from the Superose 6 GL column according to the UV trace at 280 nm. Arrow indicates the expected location of TssA^{S} based on its molecular weight (~42 kDa).

5.2.2 Transmission electron microscopy of TssA^S

The large size of the oligomer that TssA^S forms suggested that structural information may be obtained from TEM. The TssA^S complex was shown to from ring-like structures with discrete protein projections surrounded in a previous study (Ahmad 2013; this study). However, the number of the outer projections was not consistent among individual particles, and a variation in the diameter and shape of some rings was also observed (Ahmad 2013; this study). Native TssA^S was purified as described in Section 5.2.1.2, and optimal conditions for visualising the native TssA^S were attempted in four different buffer conditions: A, 25 mM Tris-HCl (pH 7.4), 150 mM NaCl and 20% glycerol; B, 25 mM Tris-HCl (pH 7.4) and 150 mM NaCl; C, 25 mM Tris-HCl (pH 6.8) and 25 mM NaCl; D, 25 mM Tris-HCl (pH 6.8) and 500 mM NaCl. However, there was no improvement in stabilising the outer discrete projections of the complex (results not shown). Therefore, in an attempt to gain more information regarding the structure of TssA^S, it was decided to average a large numbers of images.

As it was desirable to avoid that a precipitation step that was included in the purification of untagged TssA^S, it was decided to purify $_{\text{LinkerHis6}}$.TssA^S by IMAC as described in Section 5.2.1.1, following which the sample was exchanged into buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. The sample was then analysed by TEM at Department of Molecular Biology and Biotechnology, University of Sheffield. The TssA^S complex was observed to form ring-like structures containing a number of outer-projections (Figure 5.3). Averaging of TssA^S was performed with purified $_{\text{LinkerHis6}}$.TssA^S. Drop of 0.013 mg/ml protein was loaded onto freshly glow-discharged grids and stained with 0.75% uranyl formate stain (Section 2.8.1) before being analysed by TEM (CM200 FEG, Philips). Averaging was carried out with 1,208 images of single particles. Imagic-5 software was used for data analysis in collaboration with Dr. Svetomir Tzokov (Department of Molecular Biology and Biotechnology, University of Sheffield). Alignments of single particles are based on cross correlation. Direct alignment was carried including translational and rotational modes after filtering and normalisation of the images. The averaging clearly showed the central ring of $_{\text{LinkerHis6}}$.TssA^S, however, the disassociated projections around it did not have a regular distribution (Figure 5.4 A) after four steps of summing and aligning against the total sum. The rotational symmetry of $_{\text{LinkerHis6}}$.TssA^S was analysed in order to identify specific symmetries that it may have. The approach is based on correlation

between the rotated total average with its unrotated self. Two- to 20-fold symmetries were tested; however, the correlation coefficient results did not indicate a specific symmetry that had a higher score compared to the others (results not shown). Therefore, the symmetry of $_{LinkerHis6}$.TssA^{S} could not be identified using test rotational symmetry.

In addition, alignment by classification was also performed for 1,208 single particles of $_{LinkerHis6}$.TssA^{S}, which is a method that obtains classes by aligning the particles to references generated from the data set itself by multivariate statistical analysis (MSA) classification. Multi-reference alignment (MRA) of the particles was then performed against the selected class averages. The mirror symmetry option was used to cover for particles laying on the carbon film in opposite orientations. Two main classes of $_{LinkerHis6}$.TssA^{S} were found and MRA was carried out based on the references generated by these two classes and the mirrored references of them (Figure 5.4 B). However, no symmetry information of $_{LinkerHis6}$.TssA^{S} was gained after five steps of MRA (Figure 5.4 C). Because of using the mirrored references, some mirror symmetry can be seen in the final averages, but this is an artefact, result of the random orientation of the densities in the outside ring. The size estimation was conducted on total averages. The dimension of the TssA^{S} ring-like structure is ~450 Å, and the interior and exterior diameters of the central ring are 110-115 Å and 170-190 Å, respectively.

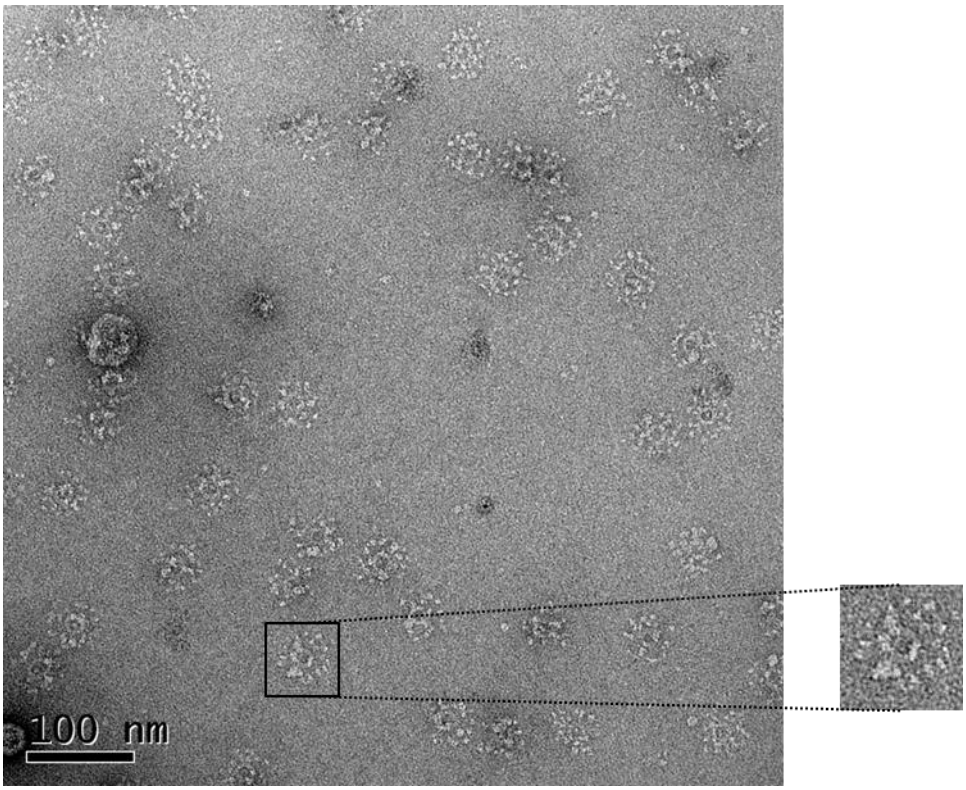


Figure 5.3 Negative stain transmission electron micrograph of $_{LinkerHis6}$.TssA^{S}. $_{LinkerHis6}$.TssA^{S} (0.013 mg/ml) in buffer containing 50 mM Tris-Cl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole was loaded onto freshly glow-discharged grids and stained with 0.75% uranyl formate stain before being analysed by TEM (CM200 FEG, Philips).

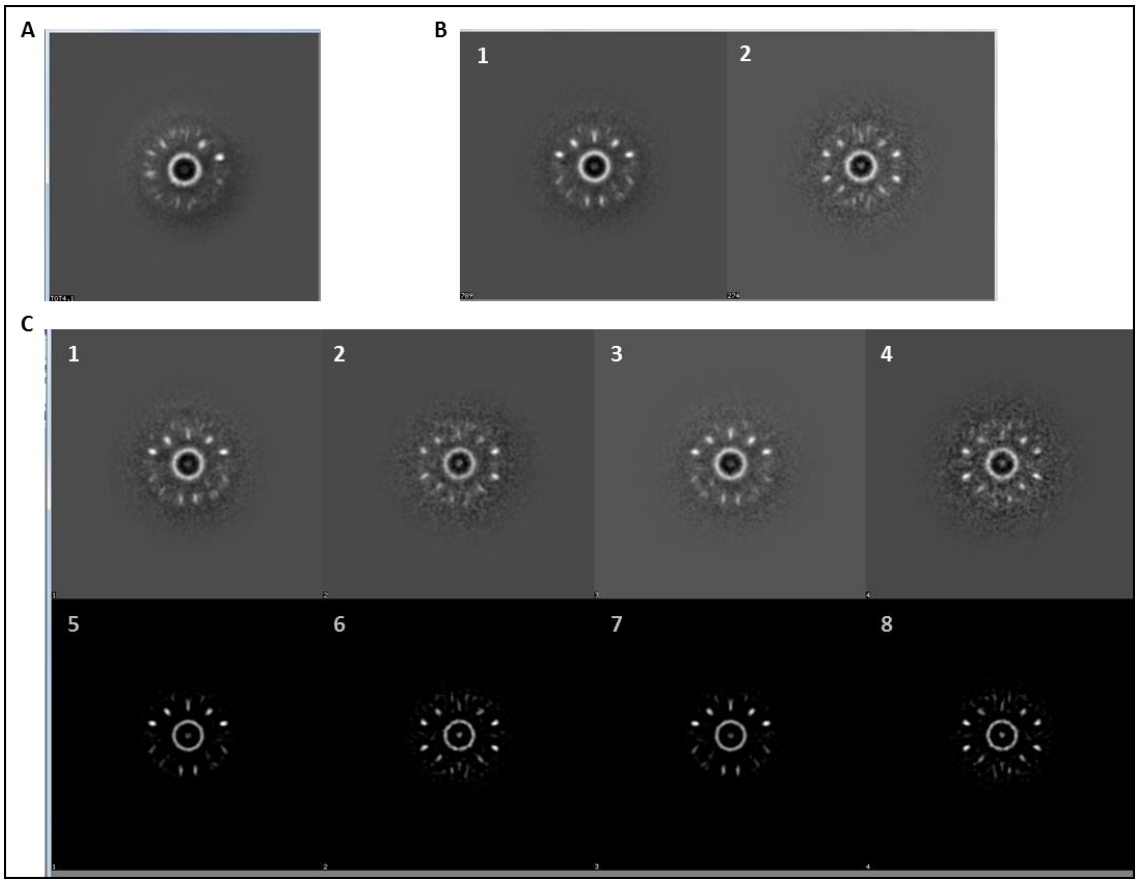


Figure 5.4 Averaging of $_{LinkerHis6}$.TssA^{S} single particles. 1,208 single particles of $_{LinkerHis6}$.TssA^{S} were collected following negative stain transmission electron microscopy. Averaging was carried out using Imagic-5 software. **A**. Total sum of $_{LinkerHis6}$.TssA^{S} after four steps of direct alignment. **B**. Two main classes of $_{LinkerHis6}$.TssA^{S} were identified by MSA. 1. Class_1 has 1 fold symmetry (no symmetry); 2. Class_2 has 2-fold symmetry. **C**. Average results of MRA based on the two main classes of $_{LinkerHis6}$TssA^{S} and the mirrored references that were identified by MSA. 1-4, the averages of $_{LinkerHis6}$.TssA^{S} after MRA; 5-8 are the references used in MRA; 5 and 7 are from Class_1, they are a pair of mirrored references; 6 and 8 are from Class_2, they are a pair of mirrored references.

5.2.3 Molecular weight estimation of TssA^S by SEC-MALS

Protein molecular weight (MW) estimation by size exclusion chromatography (SEC) is based on the assumption that the protein is globular. As the ring-shaped complex formed by TssA^S contains empty space, this will introduce a degree of inaccuracy into the SEC estimation as this technique measures hydrodynamic volume of the protein (Stulik et al. 2003). Therefore, size exclusion chromatography-multi-angle laser light scattering (SEC-MALS) was carried out at University of York Bioscience Technology Facility for the purified native (i.e. untagged) TssA^S, which gives a more accurate (within 2-5%) estimation of the molecular mass of non-globular proteins.

A Superose 6 10/300 GL (G.E. Healthcare) column was used for this procedure. The peak analysis of TssA^S was based on the elution profile using Astra software which defined four peak regions: Peak 1 (17.52-19.70 minute), Peak 2 (16.17-22.61 minute), Peak 3 (15.13-15.59 minute) and Peak 4 (14.76-16.08 minute) (Figure 5.5 A). It is noticeable that the defined Peak 1 and Peak 2 overlapped, with Peak 2 having a broader range in elution time. This is also the case for the Peak 3 and Peak 4, with Peak 4 having the broader range. The elution profile indicated that there was some aggregate material in the void volume of the column at ~14 minute then a minor peak at ~15.5 minute and a major peak at ~18.5 minute. Integration of the RI traces indicates there was ~59 µg of protein in the major peak (Peak 2) and ~2.8 µg (Peak 4, ~5% of the material) in the minor peak. The light scattering trace showed a peak for aggregate material excluded from the column material; however, this was only a small amount as it showed little or no RI or UV response.

The MW is determined by a Zimm fit procedure at each point then averages calculated over the peak region (Figure 5.5 B). The MW determined for each peak from Peak 1 to 4 is 1,238 kDa, 1,255 kDa, 3,727 kDa and 4,413 kDa, respectively. The aggregate material that was excluded from the column void volume was only a small fraction of the total amount of protein (little or no RI or UV response); however, the MW estimation of Peak 4 included the aggregate material. This will decrease the accuracy of the MW estimation by giving rise to the steep drop in the MW estimate curve. Following this, there was a minor peak (Peak 3) amounting to ~5% of the material. The estimated MW for this peak was ~3.7 MDa but this was significantly affected by the tail of the large aggregates and therefore was likely to be overestimated. The major peak (Peak 1) corresponded to most

of the material from the area under the RI trace, which was also consistent with the UV trace, implying that the peak was protein and not significantly contaminated with nucleic acid. The MW curve increased sharply at the beginning of Peak 2 because of tailing from the earlier peaks but then flattened to a quite consistent value of 1,238 kDa. This suggested that the material is homogeneous (or has a narrow distribution of sizes). There was no evident material of lower MW. As the $TssA^{S}$ polypeptide lacking the N-terminal methionine has a calculated MW of 41.65 kDa, this corresponds to 30 monomers in the complex.

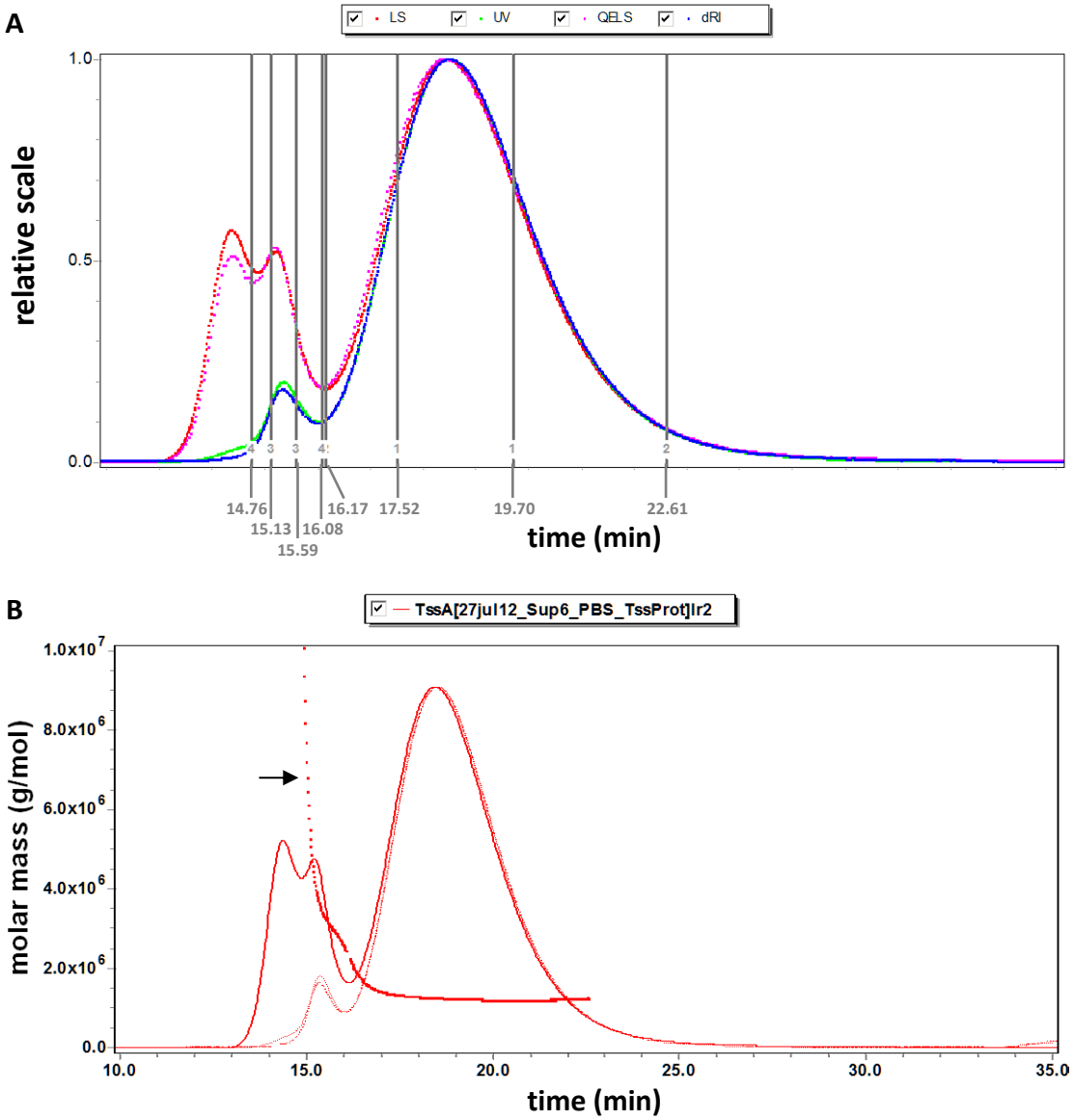


Figure 5.5 SEC-MALL analysis of TssA^S. A. TssA^S elution profile and peak positions. The normalised signals for light scattering (LS) are shown in red, refractive index (RI) in blue, UV absorption (280 nm) in green and quasi-elastic light scattering (QELS) in magenta. The peak areas chosen for analysis are indicated by numbered vertical grey lines that defined by Astra software. The X-axis shows elution time (minutes) and Y-axis shows relative scale. **B**. TssA^S molar mass plot. The MW is determined by a Zimm fit procedure at each point. Solid line represents LS trace. Dashed and dotted traces represent RI and UV responses, respectively. Molecular weight estimate curve is shown in heavier dots (pointed with black arrow). The X-axis is the elution time (minutes) and the Y-axis is the MW scale. Figures provided by University of York Bioscience Technology Facility.

5.2.4 Proteolytic analysis of TssA^S

A preparation of TssA^S, purified as described in Section 5.2.1.2, was found to be cleaved by unknown protease(s) after incubation at room temperature for 48 hours and at 37^{o}C for 8 hours, followed by further 16 hours incubation at room temperature. In order to provide further evidence for the domain organisation predicted by the bioinformatics analysis (Figures 1.9 and 1.10), the location(s) that is susceptible to proteolytic cleavage was analysed which would be exposed on the exterior of the protein. The sample was applied onto Superose 6 GL. The elution profile and Bradford assay indicated that there were two peaks containing proteins eluting at 12.02 ml (Peak 1) and 17.16 ml (Peak 2), respectively (Figure 5.6 A). The protein material in Peak 1 was eluted in the column void volume (13.5 ml). Although beyond this point, thyroglobulin (600 kDa) runs as twice bigger protein calculated using the calibration plot for Superose 6 made on basis of proteins from 6 to 400 kDa, it suggests the large complex of the material in Peak 1 could be the full-length TssA^S. The latter peak (Peak 2) corresponds to a protein apparent MW of 30 kDa.

Fractions corresponding to the two peaks were analysed by SDS-PAGE (Figure 5.6 B). The results showed that the fractions of Peak 1 contained a small amount of full-length TssA^S based on the migration of the protein in SDS-PA gels. However, most of the material in these fractions migrated as five protein bands of low molecular weight ranging from 6 to 12 kDa. A typical fraction (No. 10) was analysed by mass spectrometry. The results showed that there were eight proteins present, and all of them had the same C-terminal end as TssA^S, but had different N-terminal ends that corresponded to positions within the predicted inter-domain linker region of TssA^S (Figure 5.6 C). The longest peptide contained residues 271-373 of TssA^S. Therefore, all of the eight proteins are consistent with the C-terminal domain of TssA^S and they must be present as aggregates or assembled into ring-like oligomers in order to have eluted in the void volume. The location of the proteolytic cleavages provided further evidence for the domain organisation predicted on the bioinformatics analysis.

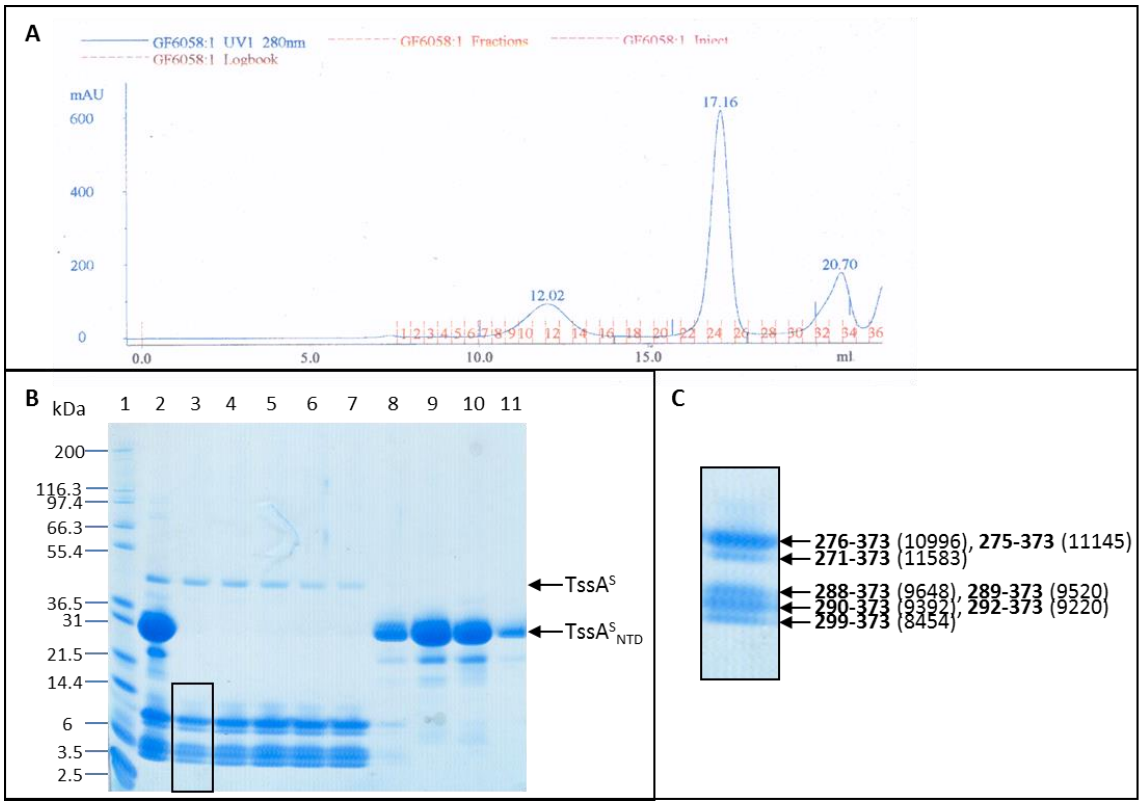


Figure 5.6 'Natural' proteolysis of TssA^S. A. SEC elution profile at 280 nm of proteolytic digest of TssA^S on Superose-6 GL. **B**. 4-12% SDS-PAGE (Novex) analysis of the fractions corresponding to the eluted protein peaks. Lane 1, protein markers (Novex); lane 2, proteolytically digested prep of TssA^S; lanes 3-7, fractions 10-14 corresponding to the Peak 1 of SEC (~12.02 mls); lanes 8-11, fractions 23, 25, 27 and 29 corresponding to the Peak 2 of SEC (~17.16 mls). **C**. Enlargement of gel region enclosed in black rectangle in Figure B showing proteins that were analysed by mass spectrometry. The position of peptide sequences obtained from MS corresponding to the amino acid sequence of TssA^S are shown in black bold font, and the calculated MWs are shown in brackets.

5.2.5 Crystallography trials with TssA^S

To obtain structural information of TssA^S, native TssA^S was purified as stated in Section 5.2.1.2. Optimum conditions for crystal formation were discovered by Qiagen screens described in Section 2.8.2. TssA^S had shown possible crystal formation in one of the 500 different conditions. However, when trying to optimise the conditions, no crystals were obtained (H. Owen and S. E. Sedelnikova, personal communication). Therefore, analytical TssA^S crystals were not obtained. It was hypothesised that this was due to the presence of the flexible interdomain linker of TssA^S which would affect the crystal formation. Therefore, attempt to identify for conditions TssA^S crystal formation conditions were discontinued, and it was decided to put more effort into determining the structure of TssA^S domains.

5.3 Analysis of TssA^S domains

To facilitate a structural analysis of TssA^S, the individual domains were characterized. The domain boundaries were proposed by multiple sequence alignment among TssA^S orthologues present in other Gram-negative bacteria (Figure 1.9) and the secondary structure prediction by Psipred (Figure 1.10), and was further supported by analysis of proteolytic degradation products (Ahmad 2013) and Section 5.2.4. Based on these analyses, TssA^S contains a large N-terminal region (residues 1-250) part of which (W56-P113) is conserved in other classes including TalT. It also contains a small C-terminal domain (residues 303-373) that is connected to the N-terminal region by a long linker of variable length and sequence.

5.3.1 Overproduction, solubility and purification of TssA$^S_{NTR}$

Overexpression of $_{His6}$.TssA$^S_{NTR}$ was carried out in *E. coli* strain BL21(λDE3) cells containing plasmid pET14b-His$_6$.*tssA*$^S_{NTR}$. pET14b-His$_6$.*tssA*$^S_{NTR}$ was constructed by cloning *tssA*$^S_{NTR}$ into the T7 promoter-driven pET14b plasmid between the *Nde*I and *Bam*HI restriction sites, which resulted in a hexa-histidine tag encoded by the vector located at the N-terminus of TssA$^S_{NTR}$. The plasmid had been previously used to generate antibodies to TssA$^S_{NTR}$ following removal of the tag with thrombin (Ahmad 2013).

Following induction, a large amount of protein of the expected size (~29 kDa) was observed by SDS-PAGE (Figure 5.7 A). It was also noticed that there was some leaky expression of $_{His6}$.TssA$^{S}_{NTR}$ in the uninduced total cells. $_{His6}$.TssA$^{S}_{NTR}$ was found to remain in the soluble fraction after cell lysis.

Purification of $_{His6}$.TssA$^{S}_{NTR}$ was performed by IMAC on a HisTrap column (Section 2.4.5). The binding step of $_{His6}$.TssA$^{S}_{NTR}$ to nickel-sepharose was repeated with the flow-through from the initial binding step. Upon SDS-PAGE analysis, $_{His6}$.TssA$^{S}_{NTR}$ was observed to bind to the HisTrap column very efficiently in the initial binding step (Figure 5.7 B). Bond $_{His6}$.TssA$^{S}_{NTR}$ was significantly eluted in the range 83-167 mM imidazole. Although the eluted $_{His6}$.TssA$^{S}_{NTR}$ was adjudged to be >80% pure, an additional SEC step was applied, which facilitates further purification and analysis of its oligomerisation status. The peak fractions from the IMAC were concentrated (MWCO 10 kDa, Vivaspin centrifugal concentrator) before being loaded onto a Superose 12 column. The elution profile monitored by UV at 280 nm showed that $_{His6}$.TssA$^{S}_{NTR}$ eluted from the column at 26.45 minute (Figure 5.7 C and D), corresponding to an apparent MW of 28.71 kDa calculated using the standard curve plotted by calibrating the column with proteins of known MW. This is in agreement with the monomeric MW of the protein, which is 29 kDa. Therefore, consistent with a previous observation (Ahmad 2013), the result suggests $_{His6}$.TssA$^{S}_{NTR}$ is a monomer.

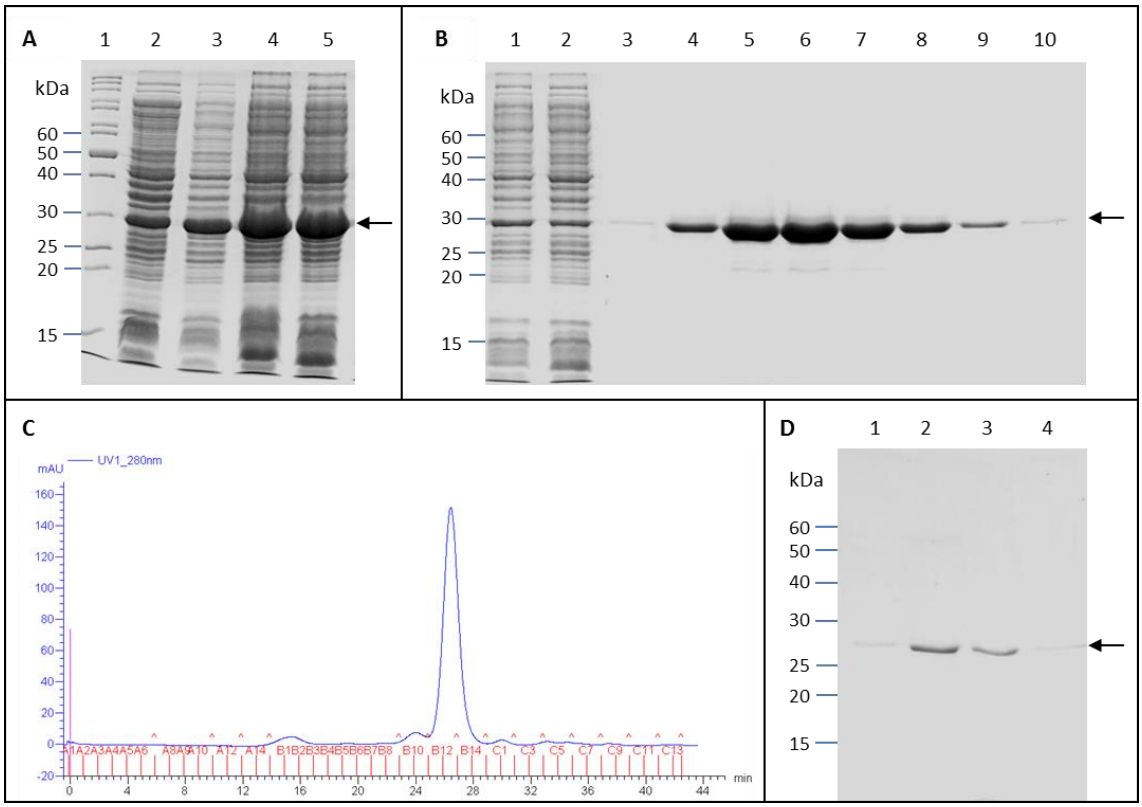


Figure 5.7 Overproduction, solubility and purification of $_{His6}$.TssA$^{S}_{NTR}$. $_{His6}$.TssA$^{S}_{NTR}$ synthesis was induced from pET14b-His$_6$.*tssA*$^{S}_{NTR}$ in *E. coli* strain BL21 (λDE3) cells with 1 mM IPTG at 37°C. Soluble fraction of the cell lysate following induction was subjected to IMAC on a HisTrap HP column (GE Healthcare) in a buffer containing 50 mM Tris-HCl, 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. The peak elution fractions was combined, concentrated and subjected to SEC on Superose 12 column in a buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. **A**. 12% SDS-PAGE analysis of overexpression and solubility of $_{His6}$.TssA$^{S}_{NTR}$. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. Purification of $_{His6}$.TssA$^{S}_{NTR}$ by IMAC with a 10-500 mM imidazole gradient. Lane 1, flow-through_1 from HisTrap column; lane 2, flow-through_2 from HisTrap column; lanes 3-10, HisTrap fractions corresponding to the protein peak based on the UV trace at 280 nm. **C**. Elution profile of $_{His6}$.TssA$^{S}_{NTR}$ from Superose 12 column. **D**. Purification and oligomerisation state estimation of $_{His6}$.TssA$^{S}_{NTR}$ with Superose 12 column. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); Lanes 2-8, peak protein elution fractions (B12-B15) according to the of the UV trace at 280 nm. Arrows in A, B and D indicate the expected location of $_{His6}$.TssA$^{S}_{NTR}$ based on its MW (~29 kDa).

5.3.2 X-ray structure determination of $TssA^{S}_{NTR}$

$TssA^{S}_{NTR}$ was subjected to X-ray crystallography analysis in order to obtain high resulution structural information. N-terminal hexa-histidine tagged $TssA^{S}_{NTR}$ was overexpressed and purified as stated in Section 5.3.1. Optimizing conditions for crystal formation was performed in the Department of Molecular Biology and Biotechnology as described in Section 2.8.2. Crystals were obtained from two conditions, i.e. (i) 0.16 M calcium acetate, 0.08 M sodium cacodylate (pH 6.5), 14.4% PEG8000 and 20% glycerol, and (ii) 0.1 M calcium acetate, 0.1 M MES (pH 6.0), containing 15% PEG400. The structure of $TssA^{S}_{NTR}$ was solved at 1.8 Å from one of the crystals obtained from the first buffer condition using Iodine-SAD (single wavelength anomalous dispersion). The structural work was carried out by Hayley Owen, Department of Molecular Biology and Biotechnology, University of Sheffield.

The structure of $TssA^{S}_{NTR}$ is largely composed of 11 alpha-helices (Figure 5.8 A). It is organised into two domains containing compact groups of helices ($TssA^{S}_{NTD}$ (residues 1-106) and $TssA^{S}_{MD}$ (residues 120-254)) that are connected to each other by an interdomain linker of 13 amino acids. $TssA^{S}_{NTD}$ has five helices (H1-H5) that form a helical bundle (Figure 5.8 B). $TssA^{S}_{NTD}$ contains most of the predicted ImpA-rel_N domain, that is conserved in all TssA-like proteins and corresponds to W56-P113 (Figure 1.7). The linker region that connects NTD and MD extends from F107 to D119. P113, that marks the C-terminal end of the conserved ImpA-rel_N region, located in the middle of the linker between $TssA^{S}_{NTD}$ and $TssA^{S}_{MD}$. Therefore, the conserved ImpA-rel_N-like region is composed of H3-H5 of $TssA^{S}_{NTD}$ and the first seven amino acids of the NTD-MD interdomain linker.

$TssA^{S}_{MD}$ contains six alpha-helices (H6-H11), which is organised into two subdomains $TssA^{S}_{MD.SD1}$ (H6, H10 and H11) and $TssA^{S}_{MD.SD2}$ (H7-H9). $TssA^{S}_{MD}$ has an overall boot-like shape with SD2 as the foot (Figure 5.8 C). The three SD1 helices (H6, H10 and H11) are long and contain 21, 29 and 20 amino acids, respectively, that form a bundle of anti-parallel helices. There is a sharp kink between L124 and G125 of H6, resulting in H6a and H6b. The three helices of SD2 are inserted between the first long helix (H6) and the other two long helices (H10 and H11) that are vertical to the $TssA^{S}_{MD.SD1}$ helical bundle. These helices are short and contain 17, 5 and 10 amino acids (H7-H9, respectively).

Prior to determination of the $TssA^{S}_{NTR}$ structure, Psipred was used for the secondary structure prediction (Figure 1.10). The Psipred prediction is in very good agreement with the solved structure of $TssA^{S}_{NTR}$ regarding the location of the alpha helices (Figure 5.8 D).

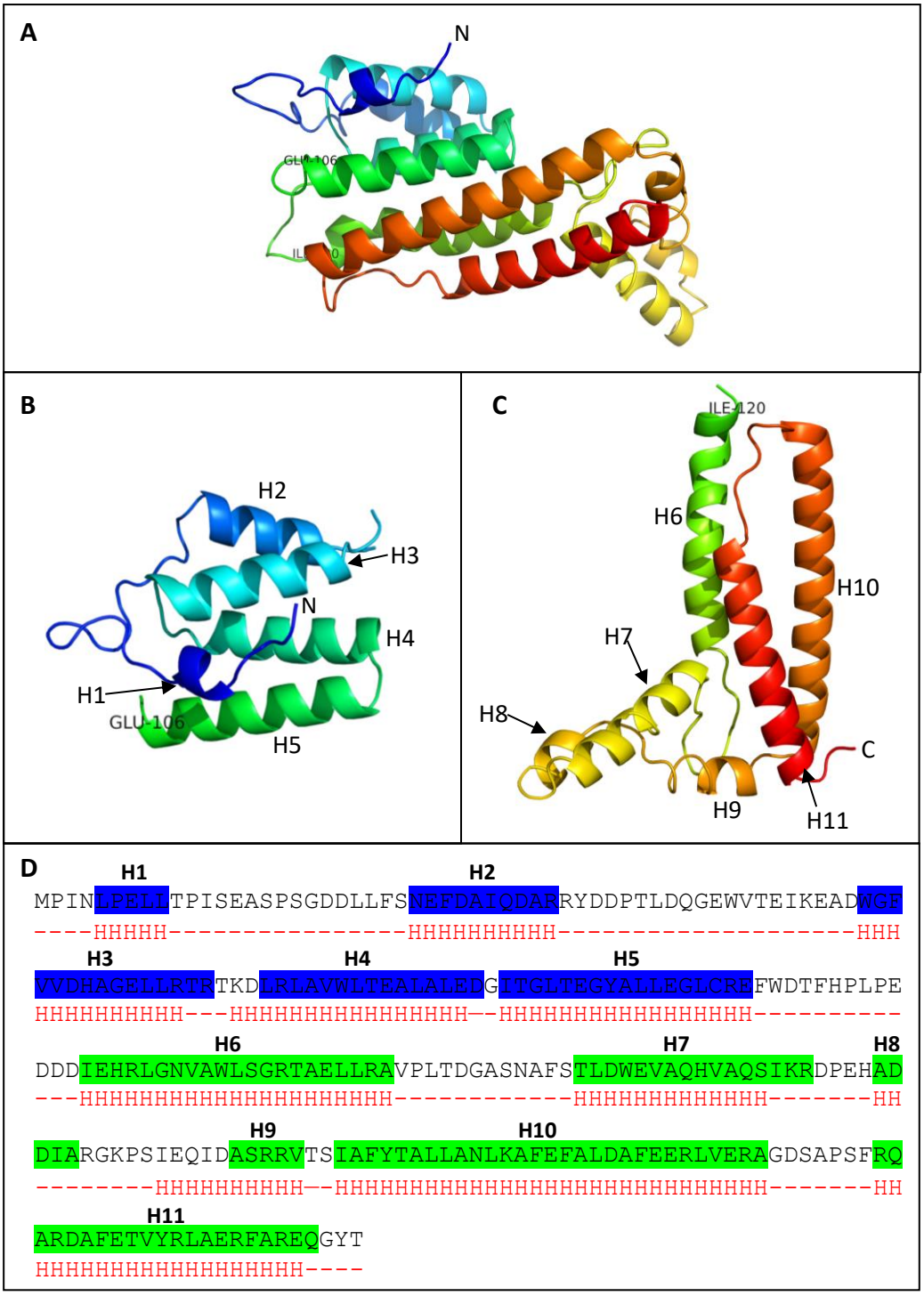


Figure 5.8 Structure of TssA$^{S}_{NTR}$. A. The overall structure of TssA$^{S}_{NTR}$. Chain is coloured in rainbow from blue to red, corresponding to N- and C-terminus of the structure. The histidine tag and thrombin cleavage site residues encoded by vector have been masked. **B**. The structure of TssA$^{S}_{NTD}$ (residues 1-106). **C**. The structure of TssA$^{S}_{MD}$ (residues 120-254). **D**. Amino acid sequence of TssA$^{S}_{NTR}$. Residues 1-106 constituting TssA$^{S}_{NTD}$ are coloured in blue and residues 120-254 constituting TssA$^{S}_{MD}$ are coloured in green. The helix numbers are labelled above the sequences in black bold font. Dashes in red are Psipred predicted coil; 'H' in red are Psipred predicted helices. Residues M1 to T254 were resolved in the structure.

5.3.3 Overproduction and purification of $TssA^{S}_{CTD}$

$TssA^{S}_{CTD}$ was fused to different affinity tags for facilitating purification and this includes MBP which allows solubilisation of the protein. Plasmid construction, and overproduction, solubility and purification of $TssA^{S}_{CTD}$ variants are described below. A summary for $TssA^{S}_{CTD}$ derivatives are shown in Table 5.1.

5.3.3.1 Overproduction and solubility of $_{His6}.TssA^{S}_{CTD}$ derivatives

DNA encoding H111 *B. cenocepacia* $TssA^{S}_{CTD}$ (residues 294-373) was cloned into the T7 promoter plasmid pET14b resulting in addition of a hexa-histidine tag at the N-terminus of $TssA^{S}_{CTD}$, i.e. $_{His6}.TssA^{S}_{CTD}$V1a (Shastri 2011). Another version of $TssA^{S}_{CTD}$ (named $_{His6}.TssA^{S}_{CTD}$V2a, residues 299-373) with 5 amino acids trimmed from the N-terminus of $_{His6}.TssA^{S}_{CTD}$V1a was also constructed in pET14b plasmid as difficulties in solubilising $_{His6}.TssA^{S}_{CTD}$V1a (Ahmad 2013). Although the predicted location of the N-terminal boundary of $TssA^{S}_{CTD}$ varies according to amino acid alignment prediction (I303) (Figure 1.9), PsiPred secondary structure prediction (R306) (Figure 1.10), and $TssA^{S}$ proteolytic cleavage analysis (T299) (Section 5.2.4), and it was thought that removing these amino acids from the predicted linker region would not affect the integrity of the CTD and might help in solubilising it. In this section, the production and solubility of the two versions of $_{His6}.TssA^{S}_{CTD}$ are described.

Three *E. coli* host strains were used for expressing both versions of $_{His6}.TssA^{S}_{CTD}$, i.e. BL21(λDE3), C41(λDE3) and C43(λDE3). Cells containing pET14b-His_6.$tssA^{S}_{CTD}$V1a and pET14b-His_6.$tssA^{S}_{CTD}$V2a were induced with 1 mM IPTG at both 30°C and 37°C. The results showed that a reasonable amount of protein with expected size of $_{His6}.TssA^{S}_{CTD}$V1a (~10.6 kDa) was produced under three conditions, i.e. C41(λDE3) host strain with induction at both 37°C and 30°C, and C43 host strain with induction at 37°C (Figure 5.9 A). However, the overproduced $_{His6}.TssA^{S}_{CTD}$V1a was insoluble following cell lysis (results not shown). There was no visible amount of $_{His6}.TssA^{S}_{CTD}$V2a overproduced under any of the tested conditions (Figure 5.9 B).

Auto-induction medium ZYM-5052 and a lower induction temperature were employed in an attempt obtain soluble $_{His6}.TssA^{S}_{CTD}$V1a and for solving the overexpression problem of $_{His6}.TssA^{S}_{CTD}$V2a. BL21(λDE3)/C41(λDE3)/C43(λDE3) cells were transformed with

both versions of pET14b-His$_6$.*tssA*$^{S}_{CTD}$ and the cells were grown at 22^{o}C for 16 hours containing ampicillin. However, no expression of either version of $_{His6}$.TssA$^{S}_{CTD}$ was observed (results not shown).

Due to the difficulties in overproducing $_{His6}$.TssA$^{S}_{CTD}$V2a from C41(λDE3), C43(λDE3) and BL21(λDE3) host strains, another *E. coli* strain, Lemo21(DE3), was employed. It is a T7 RNA polymerase-based protein overexpression host strain which facilitates identifying the optimal conditions for protein overexpression in single strain by modulating T7 RNA polymerase activity precisely through its natural inhibitor T7 lysozyme that is expressed from the extremely well titratable rhamnose promoter (Lemo 21 manual). To sample different expression levels, various concentrations of L-rhamnose were added to the medium, i.e. 0, 100, 250, 500, 750, 1,000 and 2,000 μM. Lemo21 cells containing plasmid pET14b-His$_6$.*tssA*$^{S}_{CTD}$V2a were induced with 400 μM IPTG in BHI medium at both 30^{o}C and 20^{o}C. However, no expression of $_{His6}$.TssA$^{S}_{CTD}$V2a was observed under all these conditions (Figure 5.10).

After all the efforts in overproducing $_{His6}$.TssA$^{S}_{CTD}$V2a from pET14b-His$_6$.*tssA*$^{S}_{CTD}$V2a, the sequence of the plasmid was rechecked. The results indicated that this plasmid was constructed incorrectly, as the T7 promoter of the pET14b vector was missing. This was probably caused by cutting pET14b and the *tssA*$^{S}_{CTD}$V2a amplicon with *Bgl*II and *Bam*HI restriction enzymes during cloning rather than *Nde*I and *Bam*HI. *Bgl*II will not cut the amplicon but there is a *Bgl*II restriction site located upstream of the T7 promoter in pET14b. It was decided not to reconstruct pET14b-His$_6$.*tssA*$^{S}_{CTD}$V2a but to try expressing TssA$^{S}_{CTD}$ fused to a solubility tag.

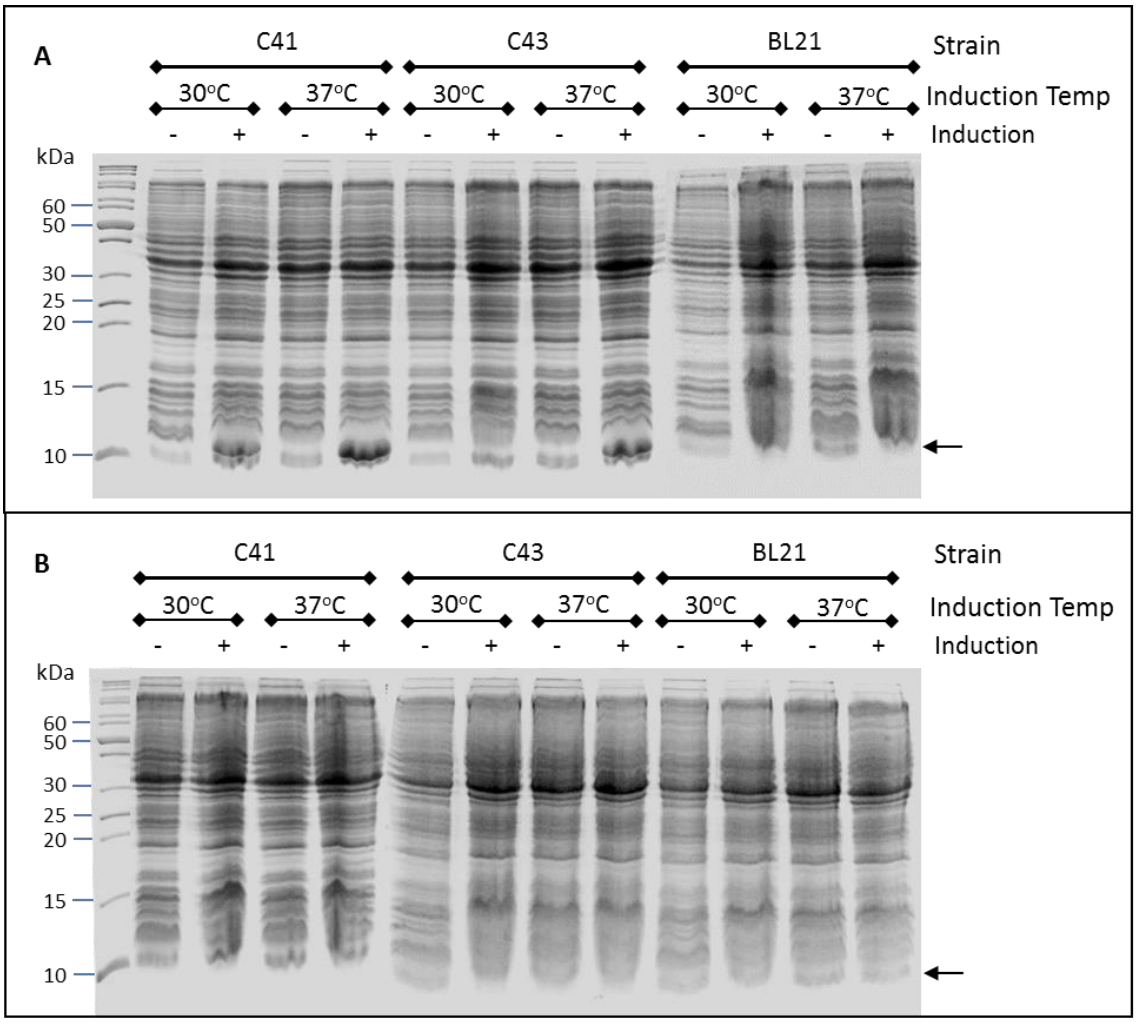


Figure 5.9 Overproduction of $_{His6}$.TssA$^{S}_{CTD}$V1a and $_{His6}$.TssA$^{S}_{CTD}$V2a in different host strains and various induction conditions. Coomassie blue-stained 15% SDS-PA gels showing the induction of both versions of pET14b-His$_6$.*tssA*$^{S}_{CTD}$ in *E. coli* strain BL21 (λDE3) cells with 1 mM IPTG at different induction temperatures, i.e. 30°C or 37°C. The protein reference markers used is EZ-Run™ Rec protein ladder (Fisher). Host strains and induction temperatures are as indicated. -, no induction; +, induction for 3 hours with 1 mM IPTG. The black arrows indicate the expected locations of $_{His6}$.TssA$^{S}_{CTD}$V1a and $_{His6}$.TssA$^{S}_{CTD}$V2a based on the MW (~10.6 kDa). **A**. Overproduction of $_{His6}$.TssA$^{S}_{CTD}$V1a. **B**. Overproduction of $_{His6}$.TssA$^{S}_{CTD}$V2a.

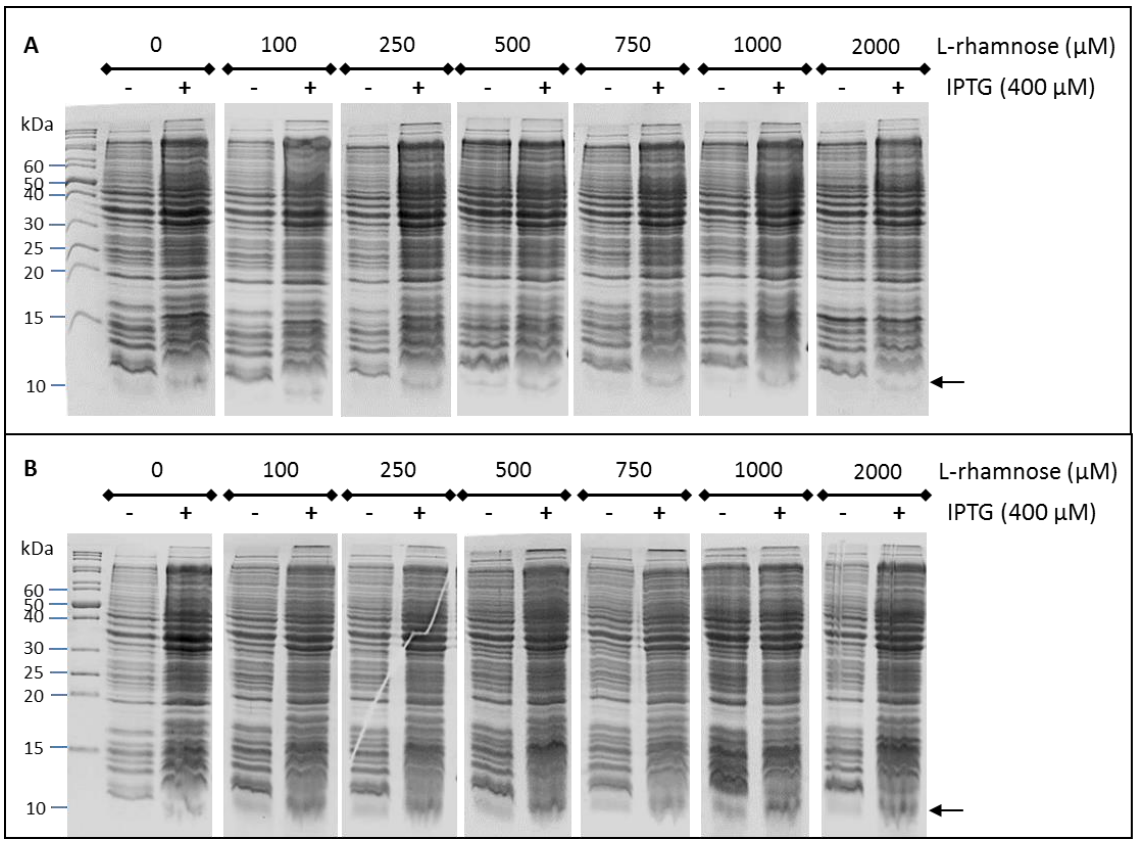


Figure 5.10 Overproduction of $_{His6}$.TssA$^{S}_{CTD}$V2a in *E. coli* Lemo21 host strain in the presence of various levels of L-rhamnose. Coomassie blue-stained 15% SDS-PA gels showing the induction of pET14b-His$_6$.*tssA*$^{S}_{CTD}$V2a in *E. coli* strain BL21 (λDE3) cells with 400 µM IPTG and different concentrations of L-rhamonose at different induction temperatures (30/20^{o}C). The reference used is EZ-Run^{TM} Rec protein ladder (Fisher). The concentrations of L-rhamonose present in the culture and induction temperatures are as indicated. -, no induction; +, induction for 3 hours with IPTG. Arrows indicate the expected locations of $_{His6}$.TssA$^{S}_{CTD}$V2a based on its MW (~10.6 kDa). **A**. Overproduction of $_{His6}$.TssA$^{S}_{CTD}$V2a at 30^{o}C. **B**. Overproduction of $_{His6}$.TssA$^{S}_{CTD}$V2a at 20^{o}C.

5.3.3.2 Purification of TssA$^{S}_{CTD}$ from MBP-TssA$^{S}_{CTD}$ fusion proteins

As TssA$^{S}_{CTD}$V1a was insoluble, an attempt was made to solubilise it by fusing TssA$^{S}_{CTD}$ to maltose-binding protein (MBP), which has been shown to have excellent solubilisation properties (Fox and Waugh 2003). The rationale for its ability to act as a solubility enhancer protein is currently unknown. However, it was proposed to function as a molecular chaperone for the fused protein (Richarme and Caldas 1997). In addition, MBP is able to facilitate the purification of the MBP fusion proteins by acting as a natural affinity tag, as it binds tightly to amylose (Kellermann and Ferenci 1982; Kapust and Waugh 1999).

The pMAL-c5X plasmid is used for inducible cytoplasmic expression of maltose-binding protein (MBP) fusions with a Factor Xa cleavage site located C-terminal to MBP, in which the cloned gene can be inserted into vector downstream of the *malE* gene (MBP coding sequence) (Figure 4.10). This results in the presence of a Factor Xa cleavage site between MBP and fused target protein, facilitating subsequent cleavage of the overproduced MBP-fusion protein into MBP and the target protein with the factor Xa protease (di Guan et al. 1988; Maina et al. 1988).

5.3.3.2.1 Construction and purification of MBP-TssA$^{S}_{CTD}$ fusion proteins

Analysis of MBP-tagged *B. cenocepacia* TssA$^{S}_{CTD}$ (named TssA$^{S}_{CTD}$V1b, residues 294-373) has been carried out in a previous study, which showed that TssA$^{S}_{CTD}$V1b remained soluble in its oligomerized form following proteolytic removal of MBP by factor Xa (Dr. Mark Thomas, personal communication). TssA$^{S}_{CTD}$V1b was overexpressed with MBP fused at the N-terminus in NEB Express cells containing plasmid pMAL-c5X-*tssA*$^{S}_{CTD}$V1b and the purified protein was subjected to crystallisation trials following factor Xa cleavage of purified MBP-TssA$^{S}_{CTD}$V1b. However, the collected crystals were fragile. For this reason, another two versions of pMAL-c5X-*tssA*$^{S}_{CTD}$V1b were constructed for expressing MBP-TssA$^{S}_{CTD}$ fusions, which have 5 or 9 amino acids trimmed from the N-terminus of TssA$^{S}_{CTD}$V1b resulting in inclusion of shorter region of the inter-domain linker of TssA^{S}, i.e. one fusion protein contains residues 299-373 (named TssA$^{S}_{CTD}$V2b) and the other contains residues 303-373 (named TssA$^{S}_{CTD}$V3b). There are 4 amino acids (isoleucine, serine, histidine and

methionine) added to the N-termini of TssA$^{S}_{CTD}$ variants encoded by pMAL-c5X once cleaved with factor Xa. DNA encoding both TssA$^{S}_{CTD}$V2b and TssA$^{S}_{CTD}$V3b were separately cloned into pMAL-c5X between the *Nde*I and *Bam*HI restriction sites following PCR amplifications with two pairs of primers, i.e. pET14b-ACTDfor2 and pET14b-iotArev for TssA$^{S}_{CTD}$V2b, and pET14b-ACTDfor3 and pET14b-iotArev for TssA$^{S}_{CTD}$V3b. The desired plasmids were identified by PCR screening and DNA sequencing following transformation of *E. coli* MC1061. *E. coli* strain NEB Express cells containing pMAL-c5X-*tssA*$^{S}_{CTD}$(V2b/V3b) plasmids were cultured in Lennox broth with 0.2% glucose comprising ampicillin, followed by induction with IPTG at 37^{o}C as described in Section 2.4.1. Both MBP-TssA$^{S}_{CTD}$V2b and MBP-TssA$^{S}_{CTD}$V3b proteins were overexpressed in large amounts, and >70% of the fusion proteins remained soluble following cell lysis (Figure 5.11).

The purification steps of TssA$^{S}_{CTD}$V2b and V3b from MBP-TssA$^{S}_{CTD}$V2b or V3b involved amylose affinity chromatography, factor Xa protease cleavage and SEC. Meanwhile, MBP-TssA$^{S}_{CTD}$V2b was also purified for crystallisation trails. Soluble cell fractions were loaded onto an amylose column (~12 mls), and chromatography was performed under gravity. Elution of bound protein was performed with buffer containing 10 mM maltose. Ten 3 ml fractions were collected. The following steps for purification of TssA$^{S}_{CTD}$V2b and TssA$^{S}_{CTD}$V3b from their corresponding MBP fusion proteins are described separately due to modifications imposed by the different efficacy of factor Xa protease in each case (Sections 5.3.3.2.2 and 5.3.3.2.3).

The unbound of MBP-TssA$^{S}_{CTD}$V2b collected in the flow-through from the amylose column was recycled by re-ran on an equilibrated amylose column and used for purification of MBP-TssA$^{S}_{CTD}$V2b with an additional SEC step (Figure 5.12 A). The concentrated material of amylose column fractions was then chromatographed on a Superdex 200 HiLoad 16/600 column. The elution profile showed a single peak at 54.81 ml monitored by UV trace at 280 nm (Figure 5.12 B). Peak fractions (7-11) were combined, concentrated and analysed by SDS-PAGE. The pooled eluted MBP-TssA$^{S}_{CTD}$V2b was judged to be >90% pure by SDS-PAGE analysis. The sample was then subjected to buffer exchange into 5 mM Tris-HCl (pH 8.0) containing 50 mM NaCl using a Zeba column (Thermo) for crystallisation trials.

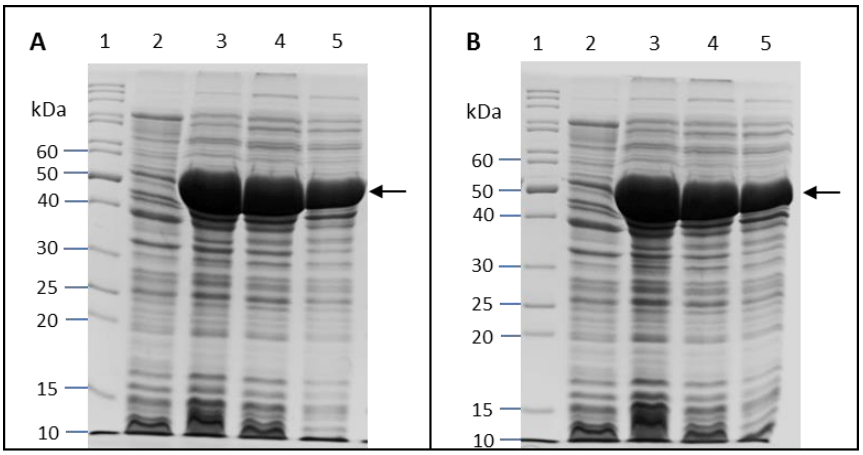


Figure 5.11 Overproduction and solubility of MBP-TssA$^{S}_{CTD}$V2b and MBP-TssA$^{S}_{CTD}$V3b. Coomassie blue-stained 12% SDS-PA gels showing the induction of pMAL-c5X-*tssA*$^{S}_{CTD}$V2b/V3b with 0.3 mM IPTG at 37^{o}C. The arrows indicate the expected locations of MBP-TssA$^{S}_{CTD}$V2b (~51.3 kDa) and MBP-TssA$^{S}_{CTD}$V3b (~51 kDa) based on their MWs. **A**. MBP-TssA$^{S}_{CTD}$V2b. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. MBP-TssA$^{S}_{CTD}$V3b. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction; lane 5, soluble fraction of cell lysate following induction.

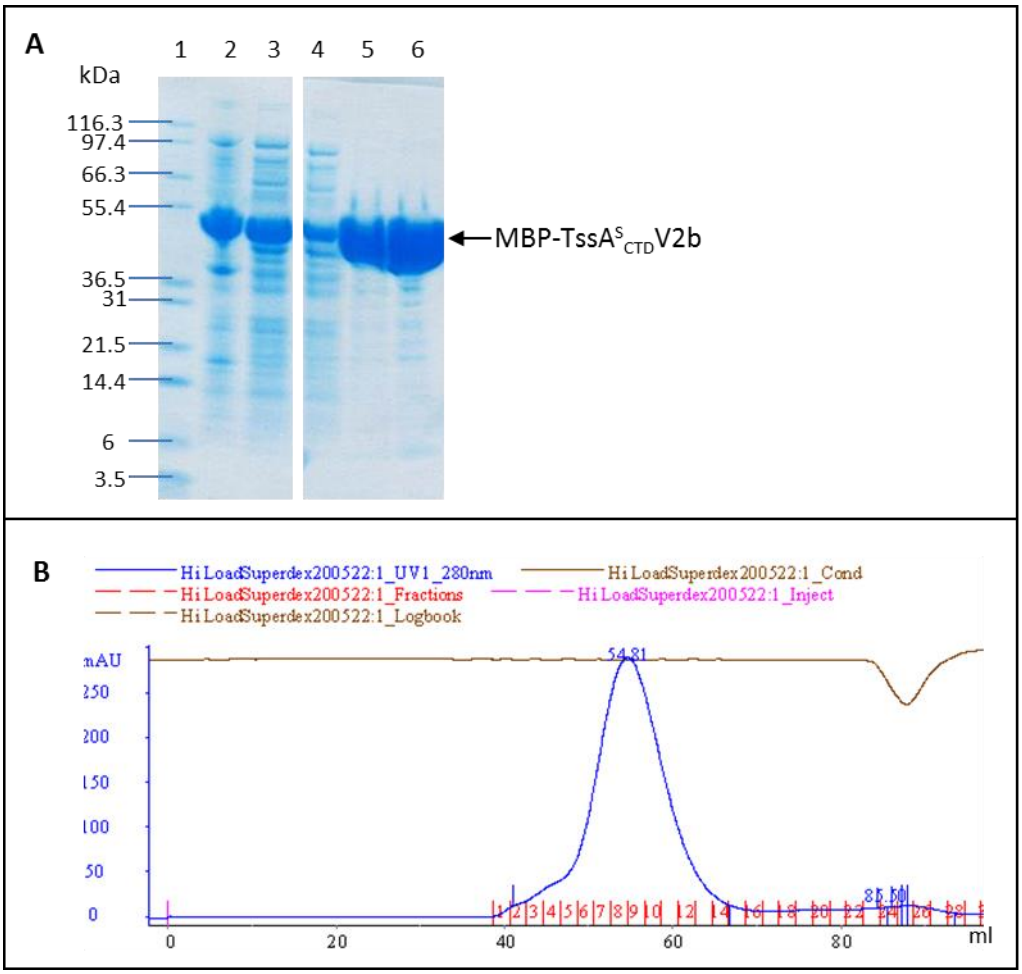


Figure 5.12 Purification of MBP-TssA$^{S}_{CTD}$V2b. A. 4-12% gradient SDS-PA gels showing the induction of *E. coli* NEB Express cells containing pMAL-c5X-*tssA*$^{S}_{CTD}$V2b with 0.3 mM IPTG at 37^{o}C. Soluble cell lysate following induction containing MBP-TssA$^{S}_{CTD}$V2b were subjected to amylose affinity chromatography in a buffer containing 50 mM Tris-HCl (pH 8.0) and 200 mM NaCl. Elution was performed in buffer containing 10 mM maltose. Elution fractions containing MBP-TssA$^{S}_{CTD}$V2b was subjected to SEC on a Superdex 200 HiLoad 16/600 column in a buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. The arrow indicates the expected location of MBP-TssA$^{S}_{CTD}$V2b based on its MW (~51.3 kDa). Lane 1, protein marker (Novex); lane 2, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 3, soluble fraction of cell lysate following induction; lane 4, unbound material from amylose column; lane 6, bound material from amylose column that was specifically eluted with maltose and concentrated as SEC load; lane 5, concentrated eluted peak fractions from SEC. **B.** Elution profile of MBP-TssA$^{S}_{CTD}$V2b from SEC monitored by UV at 280 nm.

5.3.3.2.2 Purification of TssA$^{S}_{CTD}$V2b from MBP-TssA$^{S}_{CTD}$V2b fusion protein

The protein content in the peak fractions (3-5) from the amylose column containing MBP-TssA$^{S}_{CTD}$V2b was assessed by Bradford assay, combined and concentrated to a small volume (7 mg/ml) using a centrifugal concentrator (MWCO 30 kDa). Under these conditions no aggregation material was observed. The concentrated MBP-TssA$^{S}_{CTD}$V2b was digested with factor Xa protease (Biolabs) at a final concentration of ~60 μg/ml in the presence of 2 mM $CaCl_2$ at room temperature overnight. However, the digestion was not completely effective as a small amount of undigested MBP-TssA$^{S}_{CTD}$V2b (~10%) was observed upon analysis by SDS-PAGE (Figure 5.13 A). The digestion mixture was then loaded onto a Superdex 200 HiLoad 16/600 column for separating MBP and TssA$^{S}_{CTD}$V2b. The elution profile showed two peaks at 68.25 ml and 84.58 ml that were detected at UV_{280nm} (Figure 5.13 B). The more slowly migrating peak corresponds to the monomeric molecular weight of MBP (~42 kDa); whereas the earlier peak corresponds to material with an apparent MW of 315 kDa. The fractions corresponding to the earlier peak (fractions 14-18) were pooled and concentrated (MWCO 30 kDa), followed by SDS-PAGE analysis. The majority of the material in this peak corresponded to the MW of monomeric TssA$^{S}_{CTD}$V2b, but there were also small amounts of MBP and MBP-TssA$^{S}_{CTD}$V2b present. As the protein content was predominantly TssA$^{S}_{CTD}$V2b (>80%), the sample was subjected to buffer exchange into 5 mM Tris-HCl (pH 8.0) and 50 mM NaCl and subjected to crystallisation trials.

5.3.3.2.3 Purification of TssA$^{S}_{CTD}$V3b from MBP-TssA$^{S}_{CTD}$V3b fusion protein

Fractions of amylose column (2-5) containing MBP-TssA$^{S}_{CTD}$V3b were combined and concentrated (12 mg/ml) using centrifugal concentrator (MWCO 10 kDa). The flow-through from the amylose column indicated that the unbound material constituted ~30% of MBP-TssA$^{S}_{CTD}$V3b based on SDS-PAGE analysis (Figure 5.14 A). The unbound material was recycled and re-loaded onto calibrated amylose column. The elution of the recycled material was performed as described previously, and the concentrated samples from both elutions were combined and cleaved by factor Xa protease (~25 μg/ml) in the presence of 2 mM $CaCl_2$ at room temperature overnight. The digested samples were loaded onto a Superdex 200 HiLoad 16/600 column for separating

MBP and TssA$^{S}_{CTD}$V3b. The elution profile showed two peaks at 61.83 ml and 86.71 ml (Figure 5.14 B). The later peak corresponds to the MW of MBP (~42 kDa). However, the fast running peak was eluted earlier than the elution of TssA$^{S}_{CTD}$ alone (~68 ml for TssA$^{S}_{CTD}$V2b). Upon SDS-PAGE analysis, this peak was found to contain both MBP-TssA$^{S}_{CTD}$V3b and TssA$^{S}_{CTD}$V3b proteins, suggesting the digestion of MBP-TssA$^{S}_{CTD}$V3b by factor Xa was incomplete.

A further digestion step was therefore carried out. Fractions corresponding to the peak of 61.83 ml were collected and concentrated into small volume (~1.5 ml) and then digested with factor Xa (~40 μg/ml) with 2 mM $CaCl_2$ at room temperature overnight. The digestion mixture was further incubated at 37^{o}C for 4 hours on the following day before being applied onto an amylose column to remove MBP and any undigested MBP-TssA$^{S}_{CTD}$V3b proteins which will remain bound to the resin. Ten fractions of 1.5 ml each, containing unbound materials were collected. The bound material was eluted by column buffer containing 10 mM maltose. SDS-PAGE analysis showed that there was still a small amount of undigested protein remaining in the unbound material (Figure 5.14 C). The vast majority of MBP and undigested MBP-TssA$^{S}_{CTD}$V3b bound to the amylose column, which was eluted with maltose. The unbound material of amylose column containing TssA$^{S}_{CTD}$V3b was judged to be >80% pure, and was used for crystallisation trails upon buffer exchange into 5 mM Tris-HCl (pH 8.0) containing 50 mM NaCl using a Zeba column (Thermo) for crystallisation trials.

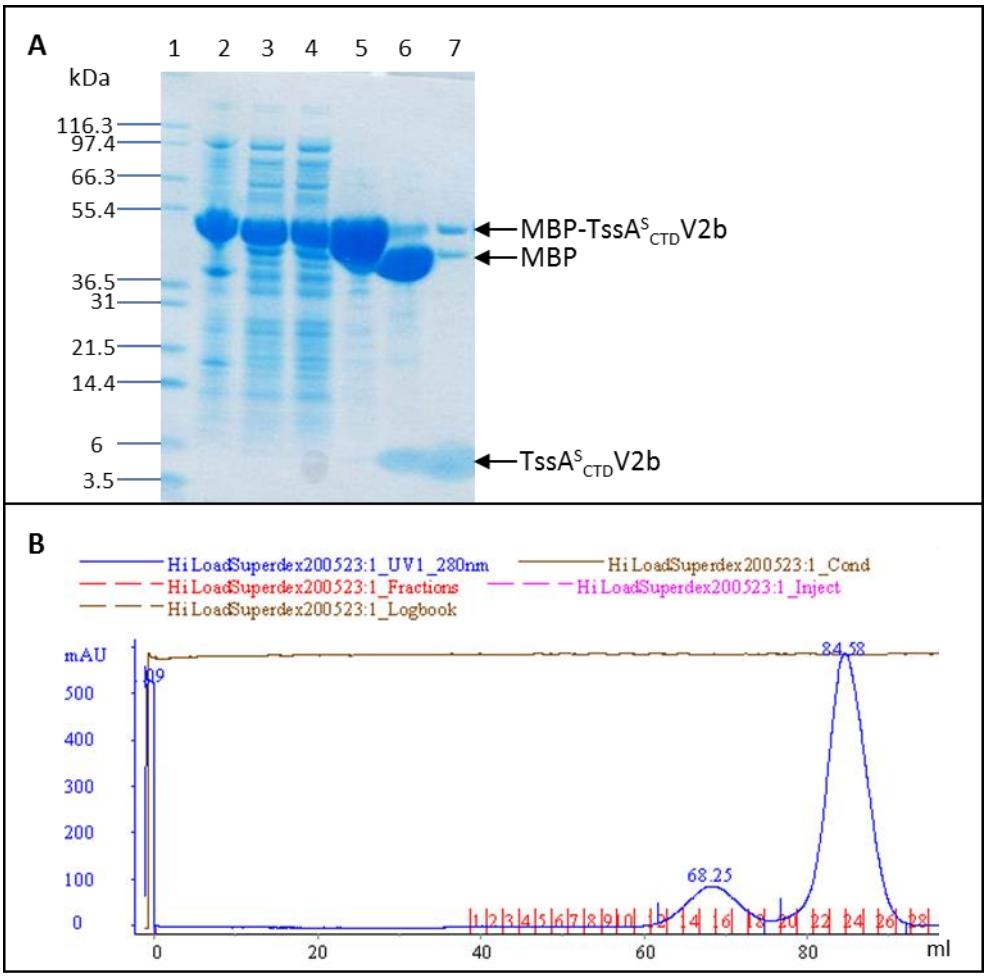


Figure 5.13 Purification of TssA$^{S}_{CTD}$V2b by factor Xa protease cleavage of MBP-TssA$^{S}_{CTD}$V2b. A. 4-12% gradient SDS-PA gel (Novex) showing the induction of *E. coli* NEB Express cells containing pMAL-c5X-*tssA*$^{S}_{CTD}$V2b with 0.3 mM IPTG at 37°C. Soluble cell lysate containing MBP-TssA$^{S}_{CTD}$V2b following induction was subjected to amylose affinity chromatography, followed by treatment with factor Xa protease cleavage and SEC on a Superdex 200 HiLoad 16/600 column. The arrows indicate the expected locations of the expected sizes of MBP-TssA$^{S}_{CTD}$V2b (~51.3 kDa), MBP (~42.5 kDa) and TssA$^{S}_{CTD}$V2b (~8.9 kDa), respectively. Lane 1, protein marker (Novex); lane 2, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 3, soluble fraction of cell lysate following induction; lane 4, unbound material from amylose column; lane 5, concentrated MBP-TssA$^{S}_{CTD}$V2b material that bound to amylose column and was specifically eluted with maltose before protease cleavage; lane 6, concentrated MBP-TssA$^{S}_{CTD}$V2b material after factor Xa cleavage; lane 7, combined peak fractions of SEC that had been concentrated and buffer exchanged for crystallisation trial. **B**. Elution profile of factor Xa-cleaved MBP-TssA$^{S}_{CTD}$V2b from SEC monitored by UV at 280 nm.

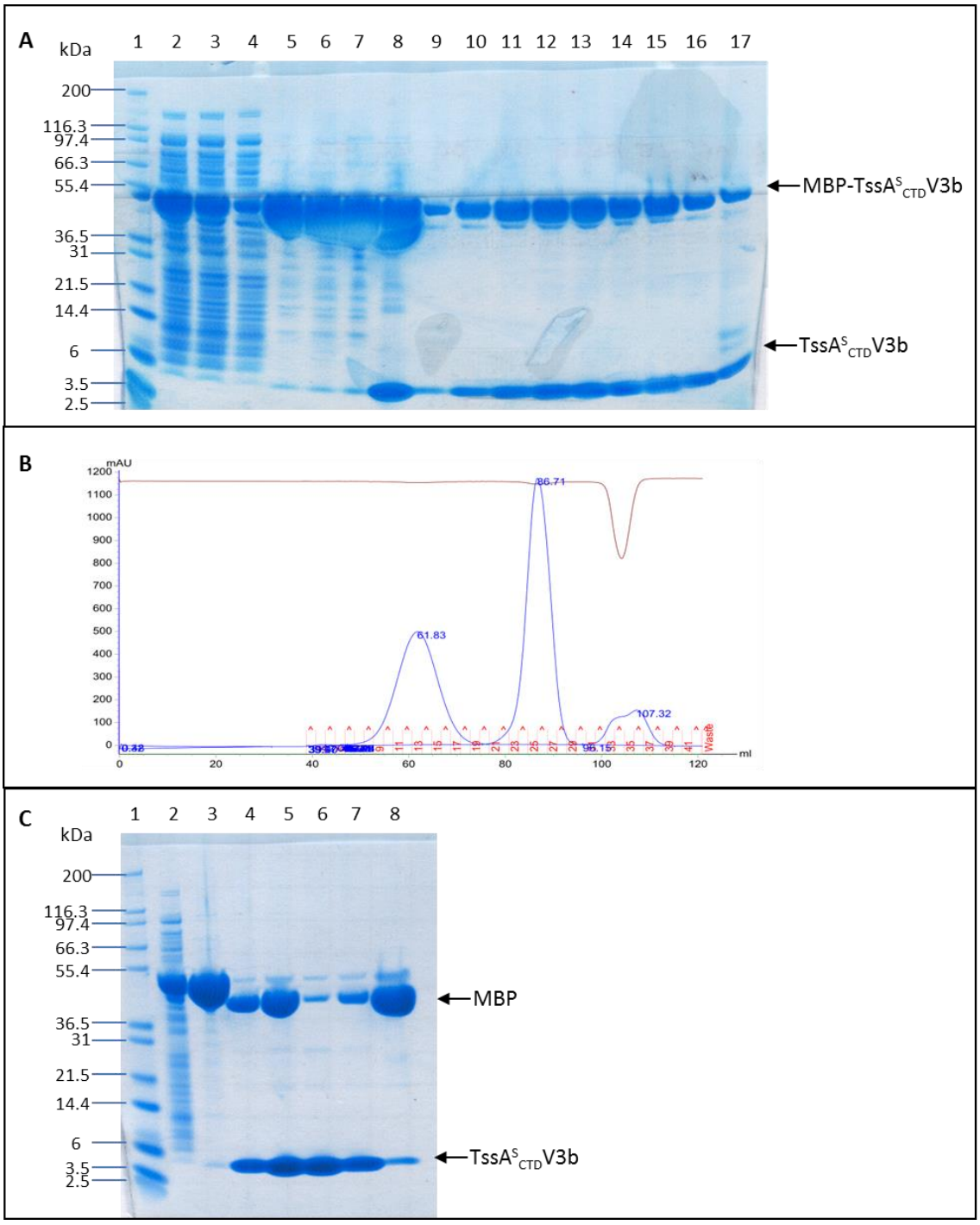


Figure 5.14 Purification of TssA$^{S}_{CTD}$V3b by factor Xa protease cleavage of MBP-TssA$^{S}_{CTD}$V3b.

Figure 5.14 Purification of TssA$^{S}_{CTD}$V3b by factor Xa protease cleavage of MBP-TssA$^{S}_{CTD}$V3b. MBP-TssA$^{S}_{CTD}$V3b synthesis was induced from pMAL-c5X-*tssA*$^{S}_{CTD}$V3b in *E. coli* NEB Express cells with 0.3 mM IPTG at 37^{o}C. Following induction, the soluble fraction of the cell lysate was subjected to amylose affinity chromatography and SEC on a Superdex 200 HiLoad 16/600 column, followed by a further step of amylose affinity chromatography for removing MBP and undigested MBP-TssA$^{S}_{CTD}$V3b protein. **A**. 4-12% gradient SDS-PA gel (Novex) showing purification of MBP-TssA$^{S}_{CTD}$V3b by amylose affinity chromatography and SEC and first treatment with factor Xa protease. Arrows indicate the expected locations of MBP-TssA$^{S}_{CTD}$V3b (~51 kDa) and TssA$^{S}_{CTD}$V3b (~8.6 kDa), respectively. Lane 1, protein marker (Novex); lane 2, soluble fraction of cell lysate following induction; lane 3, unbound material (flow-through) from first load of amylose column; lane 4, unbound material (flow-through) from second load of amylose column with recycled material (from unbound of first load); lane 5, bound material of amylose column from first load eluted with maltose; lane 6, bound material of amylose column from second load with recycled material eluted with maltose; lane 7, concentrated MBP-TssA$^{S}_{CTD}$V3b material before protease cleavage; lane 8, concentrated MBP-TssA$^{S}_{CTD}$V3b material after factor Xa cleavage; lanes 9-17, SEC fractions corresponding to the eluted protein peak at 61.83 ml monitored by UV at 280 nm. **B**. Elution profile of factor Xa-cleaved MBP-TssA$^{S}_{CTD}$V3b from SEC. **C**. 4-12% gradient SDS-PA gel (Novex) showing purification of TssA$^{S}_{CTD}$V3b following longer incubation with factor Xa protease. Arrows indicate the expected locations of MBP (~42.5 kDa) and TssA$^{S}_{CTD}$V3b (~8.6 kDa), respectively. Lane 1, protein marker (Novex); lane 2, soluble fraction of cell lysate following induction; lane 3, concentrated MBP-TssA$^{S}_{CTD}$V3b material from SEC fractions corresponding to the eluted protein peak at 61.83 ml monitored by UV at 280 nm (before factor Xa protease cleavage); lane 4, concentrated MBP-TssA$^{S}_{CTD}$V3b material after factor Xa cleavage; lane 5, load materials on amylose column; lane 6, peak fractions of the unbound material (TssA$^{S}_{CTD}$V3b) from amylose column that subjected to crystallisation trials; lane7, side peak fractions of unbound material from amylose column; lane 8, bound materials from amylose column (MBP and MBP-TssA$^{S}_{CTD}$V3b) was specifically eluted with maltose.

5.3.3.3 TssA$^{S}_{CTD}$V3c from protease cleaved $_{His6.linkerXa}$.TssA^{S}

As the crystals obtained from TssA$^{S}_{CTD}$V3b following proteolytic removal of the MBP tag were fragile (results not shown), an alternative method of obtaining soluble TssA$^{S}_{CTD}$ was considered. In this alternative approach, TssA$^{S}_{NTR}$ would serve as the solubility tag so that assembly of the TssA$^{S}_{CTD}$ would occur in its native context. To remove TssA$^{S}_{NTR}$, a factor Xa site would be introduced into the flexible linker region of TssA^{S} at a position so as to maintain the same N-terminal sequence of TssA$^{S}_{CTD}$ derived from MBP-TssA$^{S}_{CTD}$V3b, including or excluding the four pMALc5X vector-encoded amino acids at the N-terminus (Figure 5.28), named TssA$^{S}_{CTD}$V3c and TssA$^{S}_{CTD}$V3d, respectively.

To do this, SOE-PCR was carried out by separate amplification of two DNA fragments encoding amino acid sequences on either side of the insertion site where the factor Xa coding sequence was going to be introduced between amino acids G302 and I303 of TssA^{S}. Relative to the insertion site, the N-terminal coding region of *tssA*^{S}, together with an N-terminal His-tag coding sequence, was regarded as fragment A; the C-terminal coding region of *tssA*^{S} (encoding TssA$^{S}_{CTD}$V3c) was regarded as fragment B. Amplification of fragment A and fragment B involved 2 pairs of primers, pET14b-iotAfor and TssAlinkerXa1rev, and TssAlinkerXa1for and pET14b-iotArev, respectively. Primers of TssAlinkerXa1for and TssAlinkerXa1rev include complementary 5' tail with 24 non-template bases that contained the coding sequence for a factor Xa site and the same four pMALc5X vector-encoded amino acids that are present in TssA$^{S}_{CTD}$V3b (i.e. ISHM). The resulting PCR products were electrophoresed in a 1% agarose gel and purified. In order to fuse the two PCR fragments, another round of PCR was performed using pET14b-iotAfor and pET14b-iotArev primers and the two fragments as template. The products of the second round PCR were digested with *Nde*I and *Bam*HI restriction enzymes which recognized restriction sites in the pET14b-iotAfor and pET14b-iotArev primers, and ligated into pET14b plasmid that was digested with the same restriction enzymes.

E. coli strain MC1061 was transformed with the ligation mixture and colonies were grown on LB plates containing ampicillin. The correct recombinant plasmids were identified by PCR screening followed by plasmid miniprep to confirm the plasmid contained the linkerXa1.*tssA*^{S} insert. The nucleotide sequence of the inserted gene was

verified by DNA sequencing. As this cloning procedure resulted in addition of a hexa-histidine tag coding sequence to the N-terminal coding sequence of linkerXa1.*tssA*^{S} the resulting plasmid was called pET14b-His_6.linkerXa1.*tssA*^{S}.

The procedure for construction of pET14b-His_6.linkerXa2.*tssA*^{S} was intended to be the same but using primers TssAlinkerXa2for substituted for TssAlinkerXa1for, and TssAlinkerXa2rev substituted for TssAlinkerXa1rev. These primers incorporate a factor Xa cleavage site in the TssA^{S} interdomain linker but not the 4 additional amino acids present in $_{His6.linkerXa1}$.TssA^{S}. However, as a *Bam*HI restriction site was inadvertently introduced into the TssAlinkerXa2for and TssAlinkerXa2rev primers at the junction of the factor Xa site and TssA$^{S}_{CTD}$V3d coding sequences during their design, pET14b-His_6.linkerXa2.*tssA*^{S} was not made. The following work refers pET14b-His_6.linkerXa1.*tssA*^{S} as pET14b-His_6.linkerXa.*tssA*^{S}.

To overproduce $_{His6.linkerXa}$.TssA^{S}, *E. coli* strain BL21(λDE3) containing pET14b-His_6.linkerXa.*tssA*^{S} was induced at 37^{o}C with 1 mM IPTG. $_{His6.linkerXa}$.TssA^{S} was overproduced as evidenced by a large amount of a polypeptide migrating at ~45 kDa, and ~60% of the overexpressed $_{His6.linkerXa}$.TssA^{S} remained in the soluble fraction following cell lysis (Figure 5.15). The purification of $_{His6.linkerXa}$.TssA^{S} was achieved by IMAC (HisTrap HP). $_{His6.linkerXa}$.TssA^{S} bound to the nickel resin very efficiently, and was eluted from the column at ~300 mM imidazole as monitored by UV at 280 nm. Peak elution fractions (14-17) from the HisTrap column were combined and concentrated to 2 ml and buffer exchanged into 50 mM Tris-HCl (pH 8.0) containing 200 mM NaCl. $_{His6.linkerXa}$.TssA^{S} was adjudged to be over 70% pure, but there was still some appreciable contaminating low molecular weight proteins present (Figure 5.16 A). The HisTrap purified $_{His6.linkerXa}$.TssA^{S} was then digested with factor Xa protease (~50 μg/ml) in the presence of 2 mM $CaCl_2$ at room temperature for 30 hours. The majority of $_{His6.linkerXa}$.TssA^{S} was digested into $_{His6}$TssA$^{S}_{NTR_linker}$ and TssA$^{S}_{CTD}$V3c. However, there was still a small amount of undigested $_{His6.linkerXa}$.TssA^{S} remaining (<10% of the total). The digestion mixture was then loaded onto a Superdex 200 HiLoad 16/600 column for separating $_{His6}$.TssA$^{S}_{NTR_linker}$ and TssA$^{S}_{CTD}$V3c. The elution profile of SEC showed two peaks at 68.49 ml and 82.24 ml (Figure 5.16 B), corresponding to TssA$^{S}_{CTD}$V3c oligomers and $_{His6}$TssA$^{S}_{NTR_linker}$ monomers, respectively. Fractions corresponding to TssA$^{S}_{CTD}$V3c were combined and concentrated using a centrifugal concentrator (MWCO

30 kDa), and judged to be >85% pure by SDS-PAGE (Figure 5.16 A). The sample was then used for crystallisation trails upon buffer exchange into 5 mM Tris-HCl (pH 8.0) containing 50 mM NaCl using a Zeba column (Thermo) for crystallisation trials.

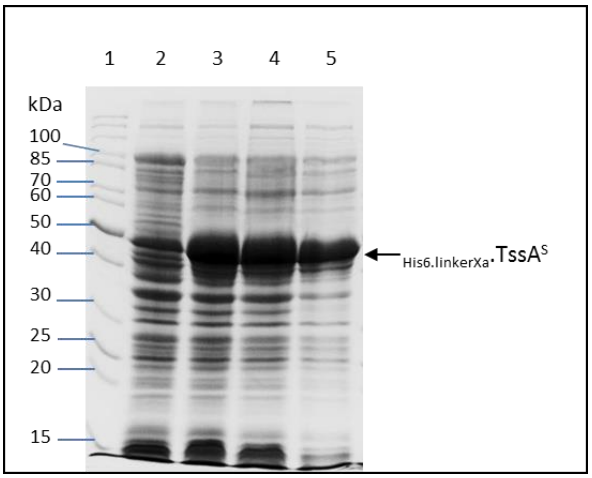


Figure 5.15 Overexpression and solubility of $_{His6.linkerXa}$.TssA^{S}. Coomassie blue-stained 10% SDS-PA gel showing the induction of pET14b-His_6.linkerXa.*tssA*^{S} in *E. coli* strain BL21 (λDE3) cells with 1 mM IPTG at 37^{o}C. The arrow indicates the expected location of $_{His6.linkerXa}$.TssA^{S} based on its MW (~45 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

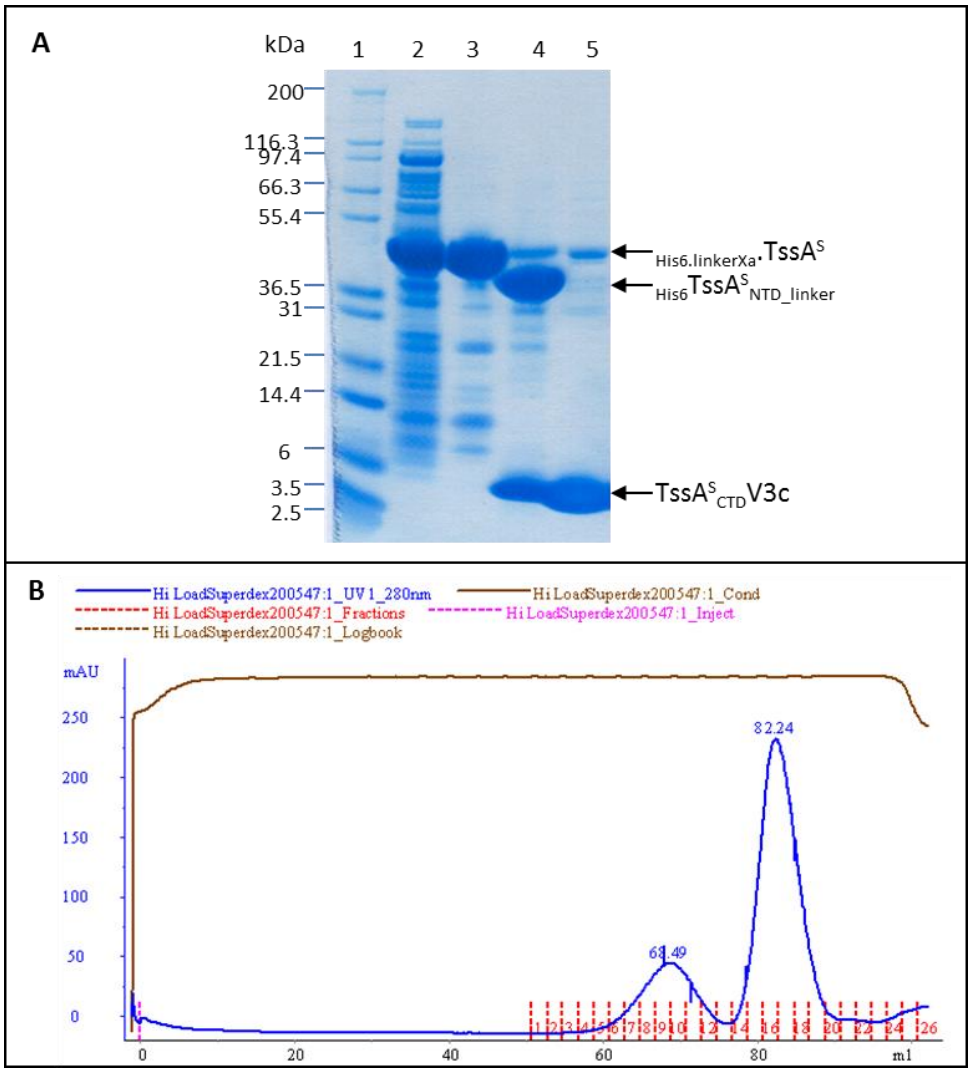


Figure 5.16 Purification of $TssA^{S}_{CTD}$V3c (residues 303-373) by factor Xa protease cleavage of $_{His6.linkerXa}$.$TssA^{S}$. *E. coli* strain BL21(λDE3) cells containing pET14b-His_6.linkerXa.*tssA*^{S} were induced with 1 mM IPTG at 37°C. The soluble fraction of the cell lysate containing $_{His6.linkerXa}$.$TssA^{S}$ was subjected to IMAC in a buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. Elution fractions containing $_{His6.linkerXa}$.$TssA^{S}$ were combined and digested with factor Xa. The digested sample was subjected to SEC in a buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. **A**. 4-12% gradient SDS-PA gel (Novex) showing the purification steps of $TssA^{S}_{CTD}$V3c. The arrows indicate the expected locations of $_{His6.linkerXa}$.$TssA^{S}$ (~45 kDa), $_{His6}TssA^{S}_{NTR\_linker}$ (~36 kDa) and $TssA^{S}_{CTD}$V3c (~8.6 kDa), respectively. Lane 1, protein marker (Novex); lane 2, soluble fraction of cell lysate following induction; lane 3, concentrated $_{His6.linkerXa}$.$TssA^{S}$ material that bound to a HisTrap column (before protease cleavage); lane 4, concentrated $_{His6.linkerXa}$.$TssA^{S}$ material after factor Xa cleavage (SEC load); lane 5, combined, concentrated peak fractions of SEC containing $TssA^{S}_{CTD}$V3c. **B**. Elution profile of separation of $_{His6}TssA^{S}_{NTR\_linker}$ and $TssA^{S}_{CTD}$V3c by SEC.

5.3.4 Transmission electron microscopy of $TssA^{S}_{CTD}$

$TssA^{S}_{CTD}$ has been shown to form a ring structure without discrete protein projections surrounded in a previous study (Ahmad 2013), suggesting the N-terminal region of $TssA^{S}$ is responsible for the discrete protein projections. A preparation of purified native $TssA^{S}$ was found to be cleaved by unknown protease(s) as described in Section 5.2.4. Following purification by SEC on Superose-6 GL, peak fractions eluting at 12.02 ml were identified as containing a small amount of full-length $TssA^{S}$, but mainly $TssA^{S}_{CTD}$ with various lengths of the inter-domain linker. A typical fraction (No.14) was buffer exchanged into 5 mM Tris-HCl (pH 8.0) and 50 mM NaCl and stored at 4^{o}C for 30 days. The sample was then subjected to analysis by TEM following centrifugation to remove any aggregates, although no visible aggregates were observed. The results showed that the ring structures typically formed by $TssA^{S}_{CTD}$ are well defined without any discrete protein projections (Figure 5.17).

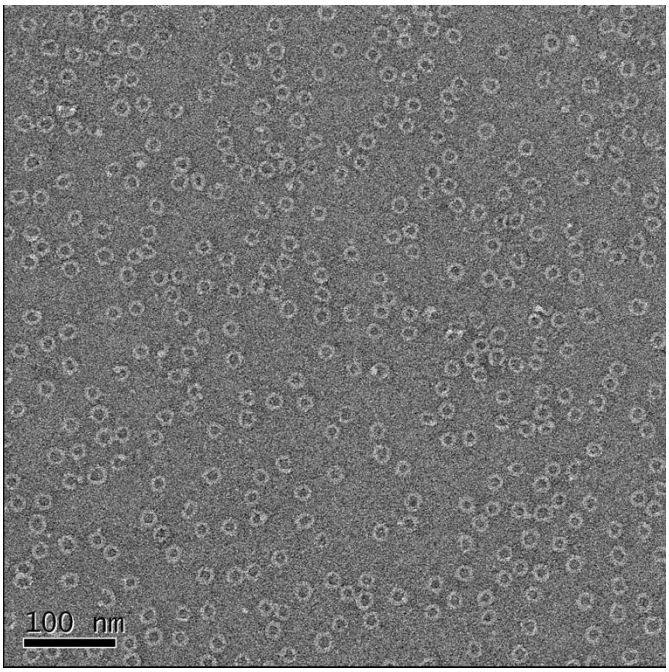


Figure 5.17 Negative stain TEM of $TssA^{S}_{CTD}$. $TssA^{S}_{CTD}$ was released from natural proteolytically digested native $TssA^{S}$ and separated from $TssA^{S}_{NTR}$ by SEC on Superose 6 GL. The sample, mainly containing $TssA^{S}_{CTD}$ (0.015 mg/ml), in buffer containing 5 mM Tris-HCl (pH 8.0) and 50 mM NaCl, was loaded onto freshly glow-discharged grids and stained with 0.75% uranyl formate stain before being analysed by TEM (CM200 FEG, Philips).

5.3.5 Molecular weight estimation of TssA$^{S}_{CTD}$ by SEC-MALS

As TssA$^{S}_{CTD}$ was observed to form a ring structure, the size estimation by SEC would not be accurate, as it includes the empty space of the ring. Therefore, SEC-MALS analysis was performed for a more accurate MW estimation of TssA$^{S}_{CTD}$ (University of York Bioscience Technology Facility).

TssA$^{S}_{CTD}$V1b (residues 294-373) was used for SEC-MALS analysis. TssA$^{S}_{CTD}$V1b was overexpressed with MBP fused at the N-terminus in NEB Express cells containing plasmid pMAL-c5X-*tssA*$^{S}_{CTD}$V1b. Purification of TssA$^{S}_{CTD}$V1b was achieved by amylose affinity chromatography, followed by factor Xa cleavage and SEC for separation of MBP and TssA$^{S}_{CTD}$V1b (results not shown).

In the SEC-MALS analysis, there was a small amount of very high MW material eluting about 15 minutes in the column void volume (Superdex 200 10/300 GL), then some high MW material in a separate slower moving peak, but both of these fractions only constitute a small amount of the total protein by weight. The major peak of TssA$^{S}_{CTD}$V1b eluted at 21 minute and there are no further peaks (Figure 5.18 A). The peak analysis of TssA$^{S}_{CTD}$V1b was based on the elution profile using Astra software which defined the peak regions into three, which are Peak 1 (20.35-21.47 minute), Peak 2 (18.93-23.25 minute) and Peak 3 (16.18-18.48 minute). Peak 2 comprised Peak 1, which has a broader range in the elution time. Integration of the RI traces indicated ~162 µg of TssA$^{S}_{CTD}$V1b was present in Peak 2, and the Peak 3 contained ~0.7 µg of material.

The MW was determined by a Zimm fit procedure at each point then averages were calculated over the peak region (Figure 5.18 B). The slight downward slope in the MW curve for TssA$^{S}_{CTD}$V1b was not significant. The MW determined for Peak 1 and Peak 2 were consistent, which was 282 kDa. There was no indication of any material with a smaller MW. Given a monomer MW of 9.4 kDa, it suggests TssA$^{S}_{CTD}$V1b comprises 30 subunits. This result is precise in agreement with the oligomerisation estimation obtained from SEC-MALS of the full-length TssA^{S}, which suggests TssA^{S} also contains 30 subunits (Section 5.2.3).

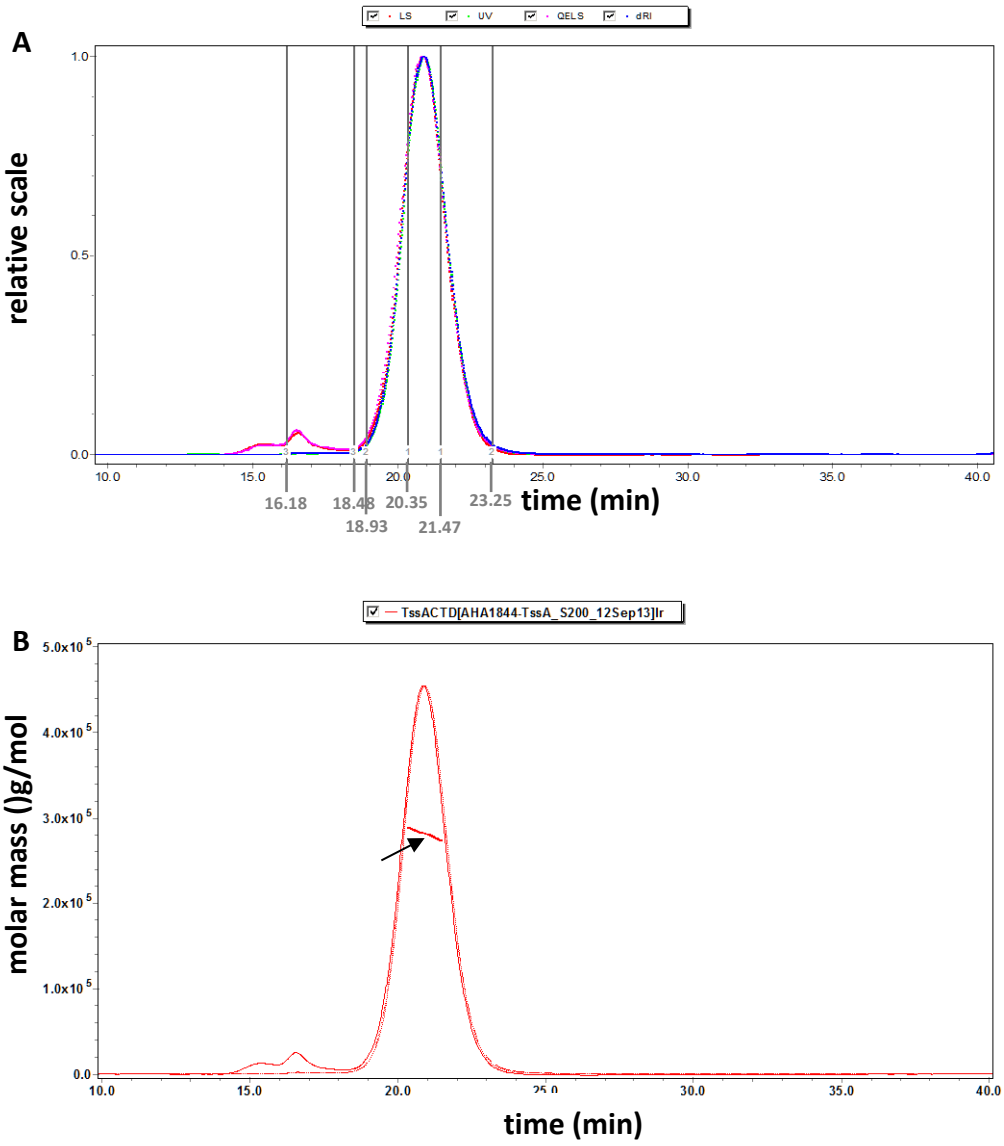


Figure 5.18 SEC-MALS analysis of TssA$^{S}_{CTD}$V1b for molecular weight estimation. **A**. TssA$^{S}_{CTD}$ elution profile and peak positions. The normalised signals for light scattering (LS) are shown in red, refractive index (RI) in blue, UV absorption (280 nm) in green and quasi-elastic light scattering (QELS) in magenta. The peak regions chosen for analysis are indicated by numbered vertical grey lines that were defined by Astra software. The X-axis shows elution time in minutes and Y-axis shows relative scale. **B**. TssA$^{S}_{CTD}$ molar mass plot. The MW was determined by a Zimm fit procedure at each point. Solid line represents LS trace. Dashed and dotted traces represent RI and UV responses, respectively. Molecular weight estimate curve is shown in heavier dots (pointed with a black arrow). The X-axis is the elution time in minutes and the Y-axis is the MW scale. Figures provided by University of York Bioscience Technology Facility.

5.3.6 Analysis of the role of individual helices in oligomerization of TssA$^{S}_{CTD}$

During TssA$^{S}_{CTD}$V3b crystallisation trials, a type of crystal diffracted, from which it was possible to obtain structural information. However, when the structure was solved it became apparent that it was a fragment of TssA$^{S}_{CTD}$V3b, probably due to proteolytic cleavage of the purified full-length TssA$^{S}_{CTD}$, which occurred during crystallisation. Solving the structure revealed that this fragment only contained the region from I303 to W347 corresponding to the first two helices (H12 and H13) of the four helices of TssA$^{S}_{CTD}$ that are predicted by Psipred (Figure 1.10). In good agreement with the Psipred prediction, in the crystal structure R306 to T323 and P328 to D341 were folded into alpha-helices that corresponded to H12 and H13, respectively. According to this finding, it was hypothesised that the presence of one of the two or both H12 and H13 in TssA$^{S}_{CTD}$ enables the domain to dimerize, whereas H14 and H15 allows the formation of a more complex structure – a 30-mer in the case of TssA$^{S}_{CTD}$. The structure is presented and discussed in more detail in Section 5.3.7.1.

5.3.6.1 Construction of plasmids to investigate the role of individual helices in TssA$^{S}_{CTD}$ oligomerisation

In order to test the above hypothesis, three plasmid constructs expressing C-terminally truncated TssA$^{S}_{CTD}$ derivatives were created – one expressing only H12 (pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD}$V3e.$_{H12}$), another expressing both H12 and H13 (pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD}$V3e.$_{H12\text{-}H13}$) and the third expressing H12, H13 and H14 (pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD}$V3e.$_{H12\text{-}H14}$). As a control, a plasmid expressing the TssA$^{S}_{CTD}$ residues 303-373 (pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD}$V3e) was also constructed in pMAL-c5X, which encodes the same region of TssA$^{S}_{CTD}$ as TssA$^{S}_{CTD}$V3b. The proteins were fused to MBP for enhancing solubility, although the solubility of the partial TssA$^{S}_{CTD}$ proteins was currently unknown. In addition, it facilities easier purification of the small truncated proteins. A hexa-histidine tag was introduced in between the factor Xa cleavage site and the TssA$^{S}_{CTD}$ proteins (Figure 5.28).

PCR amplification of *B. cenocepacia* H111 DNA encoding TssA$^{S}_{CTD}$V3e and the three truncated TssA$^{S}_{CTD}$ polypeptides was carried out using the same forward primer (pACYC-tssA.CTDfor) but different reverse primers for the different C-terminal endpoints of TssA$^{S}_{CTD}$, i.e. pACYC-tssA.CTDH1rev2 for *tssA*$^{S}_{CTD}$.$_{H12}$,

pACYC-tssA.CTDH2rev2 for *tssA*$^{S}_{CTD.H12\text{-}H13}$, pACYC-tssA.CTDH3rev2 for *tssA*$^{S}_{CTD.H12\text{-}H14}$ and pET14b-iotArev for *tssA*$^{S}_{CTD}$. The 5' non-annealing tail of the forward primer included the coding sequencing for a hexa-histidine tag. The amplified *tssA*$^{S}_{CTD}$ DNA fragments of the expected sizes and the pMAL-c5X plasmid were digested with restriction enzymes *Nde*I and *Bam*HI that recognized sites in the forward and four reverse primers and in the plasmid MCS. Ligation was carried out to allow the insertion of the *tssA*$^{S}_{CTD}$ DNA fragments into plasmids for generating pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD.H12}$, pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD.H12\text{-}H13}$, pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD.H12\text{-}H14}$ and pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD}$. *E. coli* strain MC1061 was transformed with the ligation mixtures and colonies were grown on LB plates containing ampicillin. The correct recombinant plasmids were identified by PCR screening followed by plasmid miniprep to confirm the plasmid contains the *tssA*$^{S}_{CTD}$ inserts. The nucleotide sequence of the inserted gene was verified by DNA sequencing.

5.3.6.2 Overproduction and purification of C-terminally truncated TssA$^{S}_{CTD}$ proteins

5.3.6.2.1 MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$

E. coli NEB Express cells containing pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD}$V3e.$_{H12}$ were induced with 0.3 mM IPTG at 37^{o}C. Following induction, a large amount of polypeptide of the expected size for MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ was observed. Upon cell lysis, ~50% of the protein remained in the soluble fraction (Figure 5.19 A). IMAC on a HisTrap HP column was used as the initial purification step. Almost all of the MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ material bound to the column, and there was no visible amount of unbound MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ in the washes based on SDS-PAGE analysis (Figure 5.19 B). MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ was then eluted from the column in the range 83-183 mM imidazole followed by SEC on a Superose 12 column to estimate the oligomerisation state of the protein. The elution profile from the Superose 12 column showed a single peak at 24.66 minute monitored by UV at 280 nm (Figure 5.19 C), corresponding material with an apparent MW of 47 kDa. The peak fractions were analysed on SDS-PA gels, and corresponded to the MW of monomeric MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ ~47 kDa

(Figure 5.19 D). Therefore, the single helix derivative MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ exists as a monomer.

5.3.6.2.2 MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$

Overproduction of the predicted two-helix derivative was conducted as stated in the above section (5.3.6.2.1). MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ was overproduced evidenced by the presence of a large amount of a polypeptide migrating at the expected location of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ (~49 kDa) analysed by SDS-PAGE. Following cell lysis, ~60% of overproduced MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ remained in the soluble fraction (Figure 5.20 A). MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ was purified by IMAC on a HisTrap HP column. The majority of the MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ bound to the column, and there was no visible amount of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ present in the wash fraction as judged by SDS-PAGE analysis. The bound material was then eluted from the column in the range at 83-167 mM imidazole (Figure 5.20 B), which was then further purified by SEC on a Superose 12 column. The elution profile from the Superose 12 column showed a single peak at 21.91 minute monitored by UV at 280 nm (Figure 5.20 C), corresponding to a protein of an apparent MW of 108.43 kDa. The peak fractions were analysed by SDS-PAGE, and this showed a protein corresponding to the monomeric size of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ (Figure 5.20 D). Therefore, given the monomer MW is 49.1 kDa, MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ is a dimer. As MBP is known to be a monomer, $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ is proposed to be responsible for the dimeric structure, which is consistent with the dimeric structure obtained from the crystal of the TssA$^{S}_{CTD}$V3b degradation product (Section 5.3.7.1).

5.3.6.2.3 MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$

Overproduction of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ was conducted as stated in the Section 5.3.6.2.1. Following induction, MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ was overproduced in a huge amount and ~50% of the overproduced protein remained soluble following cell lysis (Figure 5.21 A). MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ was purified by IMAC on a HisTrap HP column was used and ~75% of the MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ was retained on the

column. A small amount of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ remained unbound and was present in the wash fraction was evidenced by SDS-PAGE analysis. The bound material was eluted from the column in the range 83-183 mM imidazole (Figure 5.21 B). Fractions containing MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ were combined and subjected to SEC on a Superose 12 column. The elution profile from the Superose 12 column showed a single peak at 19.60 minute monitored by UV at 280 nm (Figure 5.21 C), corresponding to a protein of an apparent MW of 213.2 kDa. The peak fractions were analysed on a SDS-PA gel, and contained a protein corresponding to the monomeric size of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ (Figure 5.21 D). Therefore, given the monomer MW is 50.2 kDa, MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ is estimated to be a tetramer.

As MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ forms a dimer and MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ is estimated to be a tetramer, the presence of H14 may be responsible for dimerisation of the H12 and H13 dimers into tetramers while H15 may allow assembly of tetramers into the TssA$^{S}_{CTD}$ ring.

5.3.6.2.4 $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$

As MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ is estimated to be a tetramer by SEC, to rule out any possible artefacts caused by MBP (although it exists as a monomer), $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ was purified by proteolytic removal of MBP from MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$. Protein overexpression was performed as described in Section 5.3.6.2.3, and soluble fraction following cell lysis was applied onto an amylose column equilibrated with column buffer (50 mM Tris-HCl (pH 8.0), 200 mM NaCl). MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ bound to column efficiently and there was only a very small amount of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ was present in the unbound material (Figure 5.22 A). Bound material was eluted from the column with column buffer containing 10 mM maltose. The protein content of fractions containing MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ were determined by Bradford assay and combined, concentrated (MWCO 30 kDa), and subjected to cleavage by factor Xa protease (~100 μg/ml) in the presence of 2 mM $CaCl_2$ at room temperature overnight. The digestion of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ was more efficient than occurred previously for MBP-TssA$^{S}_{CTD}$V2b (Section 5.3.3.2.2) and MBP-TssA$^{S}_{CTD}$V3b (Section 5.3.3.2.3) as a

higher concentration of factor Xa was used. For separating MBP and $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12\text{-}H14}$, the digested sample was then applied onto a HisTrap column. $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12\text{-}H14}$ was eluted from the column in the range 315-420 mM imidazole. Eluted fractions containing $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12\text{-}H14}$ were combined and injected onto a Superdex 200 GL column for determining the MW of it. The elution profile showed a single peak at 80.74 ml detected by UV at 280 nm (Figure 5.22 B), corresponding to a protein with an apparent MW of 30 kDa calculated using the standard calibration curve. The peak fractions were analysed by SDS-PAGE (Figure 5.22 A) which showed that they mainly contained polypeptides corresponding to the monomeric size of $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12\text{-}H14}$ (~7.9 kDa), which was then used for crystallisation trials upon buffer exchange into 5 mM Tris-HCl (pH 8.0) containing 50 mM NaCl using a Zeba column (Thermo). Therefore, given the monomer MW of $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12\text{-}H14}$ is 7.9 kDa, the results indicate $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12\text{-}H14}$ occurs as a tetramer, and the presence of MBP in the fusion protein did not contribute to the tetrameric structure.

5.3.6.2.5 MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e

As a control for the truncated MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e derivatives, i.e. to show the MBP-$_{His6}$ tag did not interfere with folding and assembly of TssA$^{S}_{CTD}$, protein overproduction was conducted as stated in the Section 5.3.6.2.1. Following induction, MBP-$_{His6}$.TssA$^{S}_{CTD}$ was overproduced in a huge amount and ~50% of the overproduced protein remained soluble following cell lysis (Figure 5.23 A). MBP-$_{His6}$TssA$^{S}_{CTD}$ was purified by IMAC on a HisTrap HP column. However, unlike the truncated derivatives, only small amount of protein (~5%) was bound to the column. The bound material was eluted from the column in the range 200-250 mM imidazole as judged by analysis in SDS-PA gels (Figure 5.23 B). As only a small amount of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e was able to be purified by IMAC, SEC was not performed. Previous results have shown that TssA$^{S}_{CTD}$ can oligomerise into a ring containing ~30 monomers, and it is quite likely the poor binding to the nickel resin is caused by its complex oligomer structure. To confirm that MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e was able to oligomerise into a ring structure, the purified MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e was subjected to analysis by TEM in the following section.

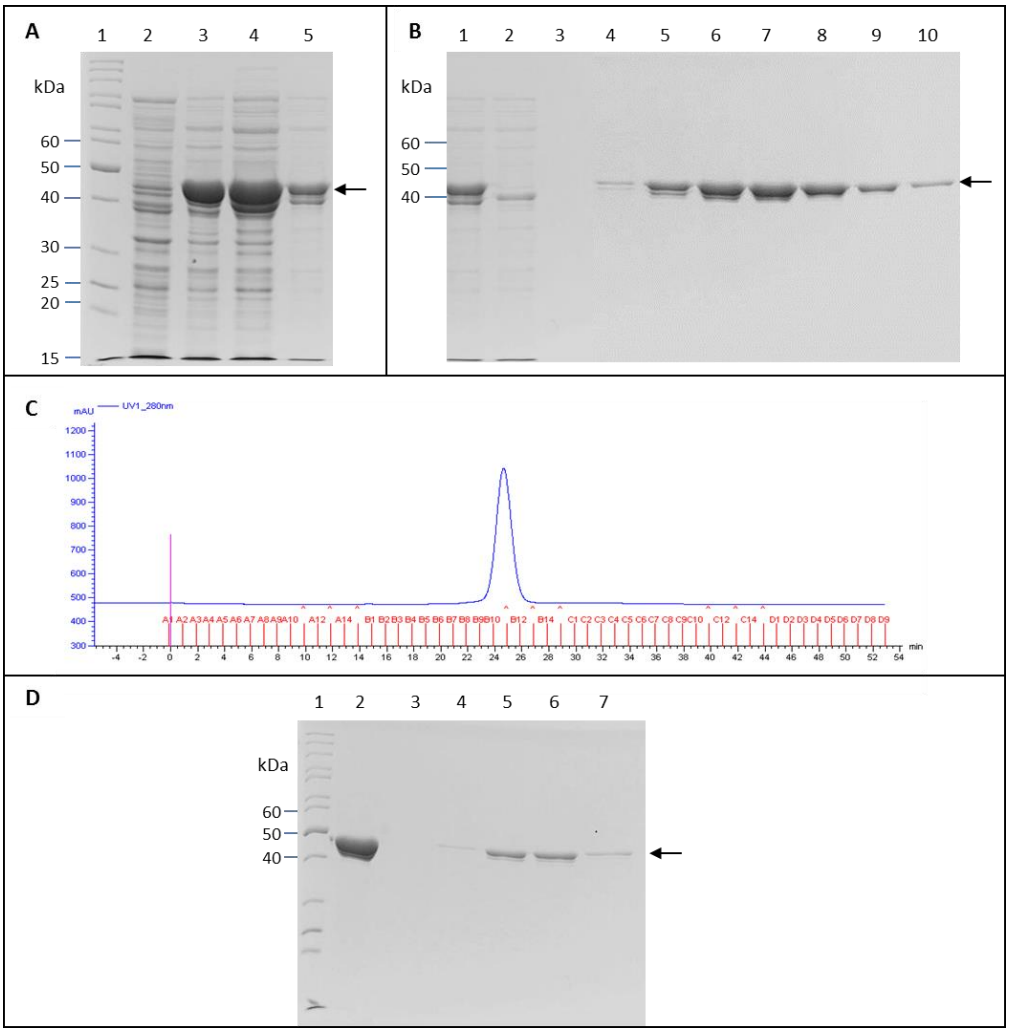


Figure 5.19 Overproduction and purification of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$. MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ synthesis was induced from pMAL-c5X-His$_6$.*tssA$^{S}_{CTD}$*V3e.$_{H12}$ in *E. coli* strain NEB Express with 0.3 mM IPTG at 37^{o}C. Following induction, the soluble fraction of the cell lysate was applied onto a HisTrap HP column (1ml) in a buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. Elution fractions containing MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ were combined, concentrated and subjected to SEC on a Superose 12 column in a buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. **A**. Overexpression and solubility of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ analysed by electrophoresis in a 10% SDS-PA gel. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. 10% SDS-PAGE analysis of purification of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ by IMAC. Lane 1, soluble fraction of cell lysate following induction; lane 2, flow-through from HisTrap column; lane 3, wash of protein loaded on HisTrap column before elution step; lanes 4-10, HisTrap fractions corresponding to the peak of the UV trace at 280 nm. **C**. Elution profile of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ during SEC. **D**. 10% SDS-PAGE analysis of SEC of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, concentrated MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ following IMAC (SEC load); lanes 3-7, elution fractions of SEC corresponding to the peak of the UV trace at 280 nm. Arrows in A, B and D indicate the expected locations of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12}$ based on its MW (~47 kDa).

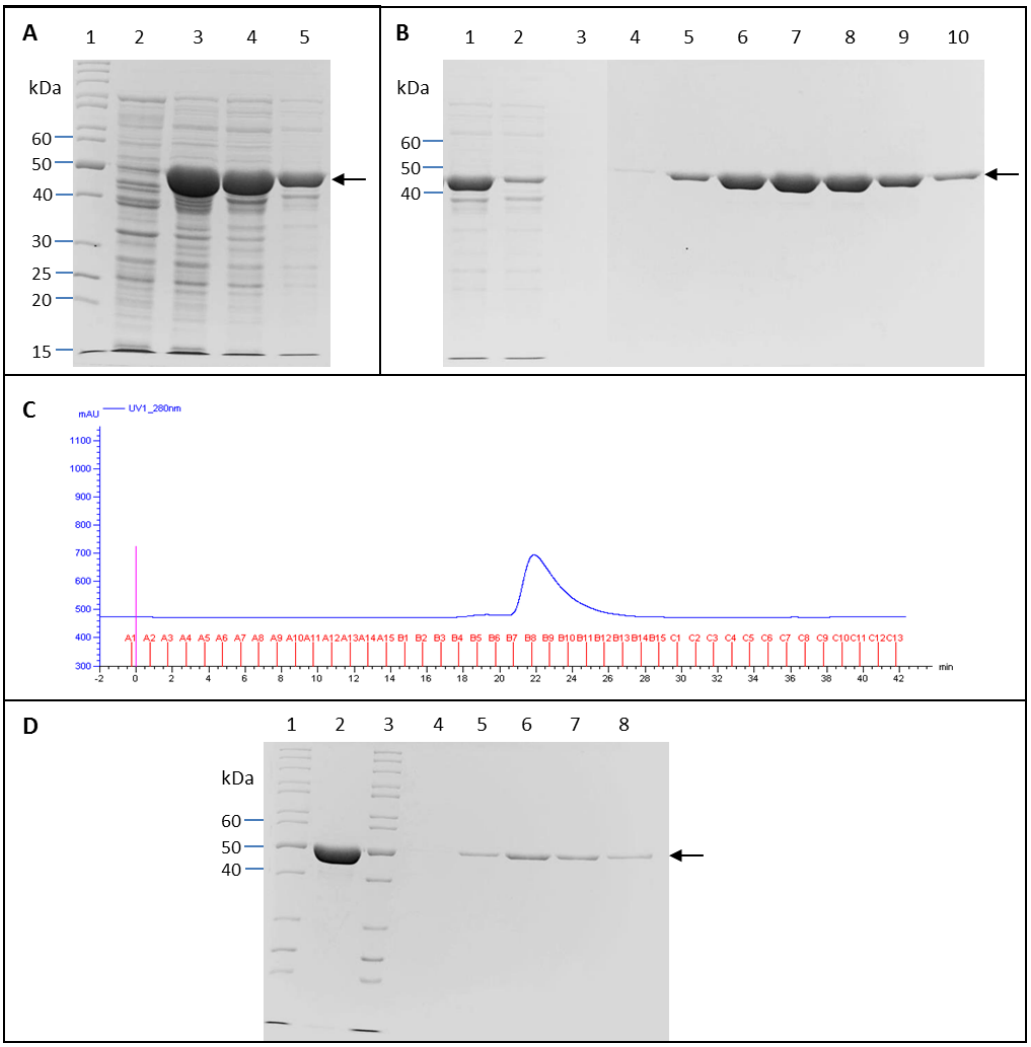


Figure 5.20 Overproduction and purification of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$. MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ synthesis was induced from pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD}$V3e.$_{H12-H13}$ in *E. coli* strain NEB Express with 0.3 mM IPTG at 37°C. Following induction, the soluble fraction of the cell lysate was applied onto a HisTrap HP column (1 ml) in a buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. Elution fractions containing MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ were combined, concentrated and subjected to SEC on a Superose 12 column in a buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. **A**. Overexpression and solubility of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ analysed by electrophoresis in a 10% SDS-PA gel. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. 10% SDS-PAGE analysis of purification of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ by IMAC. Lane 1, soluble fraction of cell lysate following induction; lane 2, flow-through from HisTrap column; lane 3, wash of protein loaded on HisTrap column before elution step; lanes 4-10, HisTrap fractions corresponding to the peak of the UV trace at 280 nm. **C**. Elution profile of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ during SEC. **D**. 10% SDS-PAGE analysis of SEC of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, concentrated MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ following IMAC (SEC load); lane 3, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 4-8, elution fractions of SEC corresponding to the peak of the UV trace at 280 nm. Arrows in A, B and D indicate the expected locations of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H13}$ based on its MW (~49 kDa).

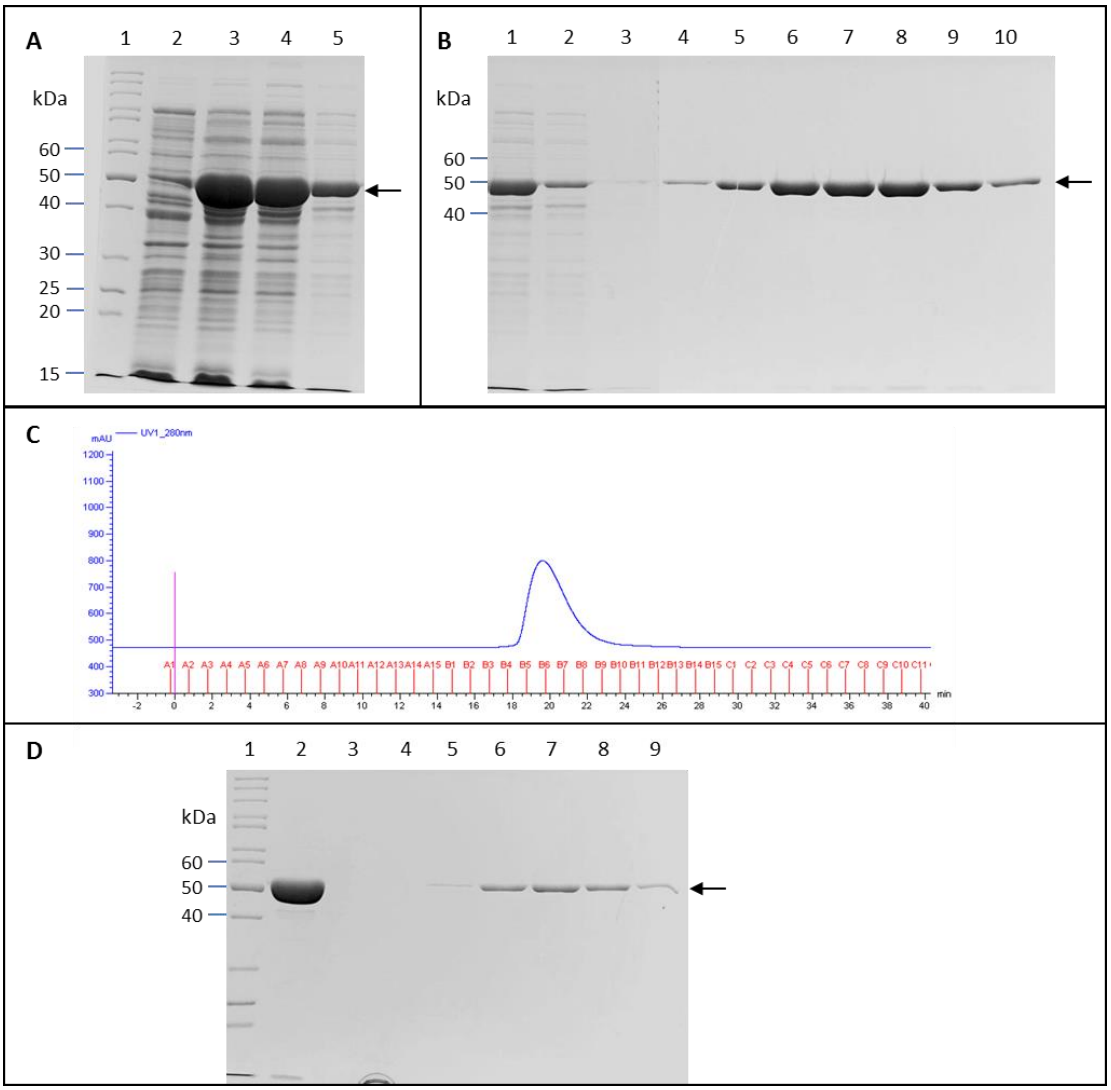


Figure 5.21 Overproduction and purification of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$. MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ synthesis was induced from pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD}$V3e.$_{H12-H14}$ in *E. coli* strain NEB Express with 0.3 mM IPTG at 37^{o}C. Following induction, the soluble fraction of the cell lysate was applied onto a HisTrap HP column (1 ml) in a buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. Elution fractions containing MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ were combined, concentrated and subjected to SEC on a Superose 12 column in a buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. **A**. Overexpression and solubility of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ analysed by electrophoresis in a 10% SDS-PA gel. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. 10% SDS-PAGE analysis of purification of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ by IMAC. Lane 1, soluble fraction of cell lysate following induction; lane 2, flow-through from HisTrap column; lane 3, wash of protein loaded on HisTrap column before elution step; lanes 4-10, HisTrap fractions corresponding to the peak of the UV trace at 280 nm. **C**. Elution profile of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ during SEC. **D**. 10% SDS-PAGE analysis of SEC of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, concentrated MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ following IMAC (SEC load); lanes 3-9, elution fractions of SEC corresponding to the peak of the UV trace at 280 nm. Arrows in A, B and D indicate the expected locations of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ based on its MW (~50 kDa).

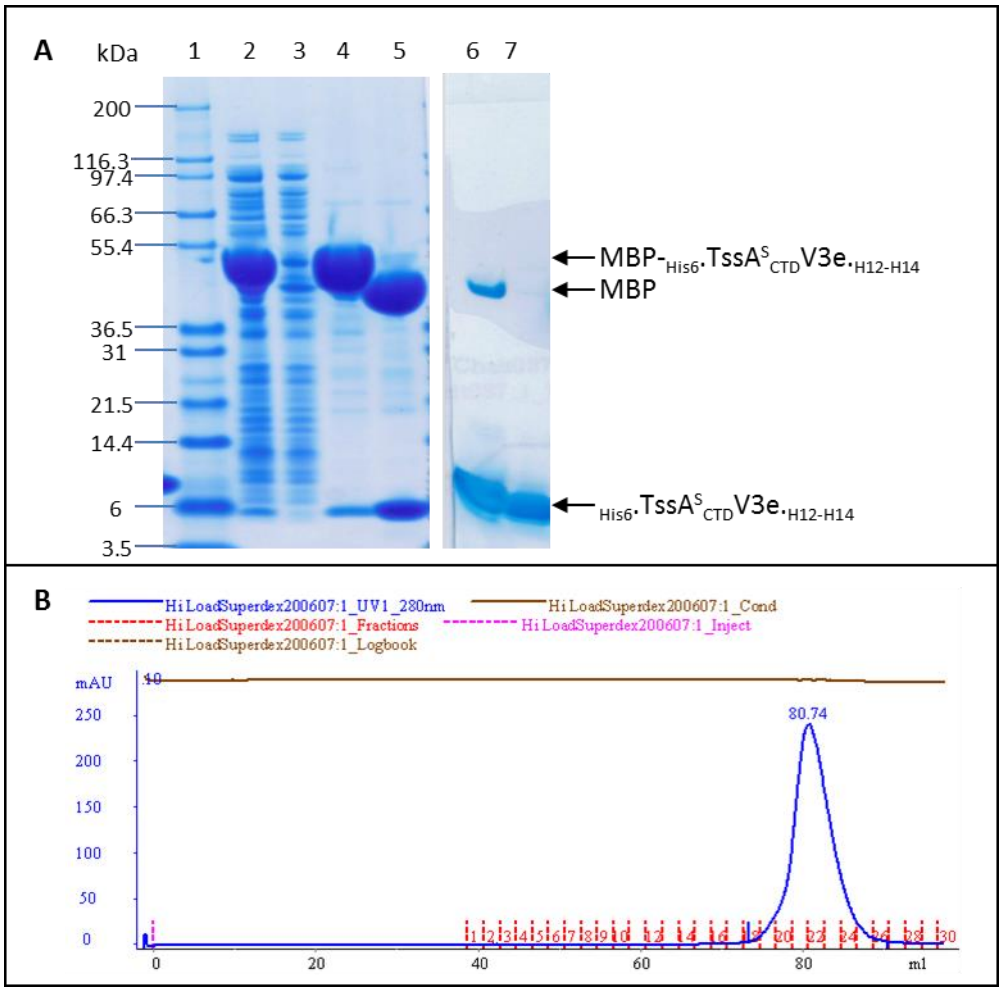


Figure 5.22 Purification of $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ by factor Xa cleavage of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$. MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ synthesis was induced from pMAL-c5X-His$_6$.*tssA*$^{S}_{CTD}$V3e.$_{H12-H14}$ in *E. coli* strain NEB Express with 0.3 mM IPTG at 37°C. Soluble fraction of the cell lysate containing MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ was subjected to amylose affinity chromatography followed factor Xa cleavage. NaCl concentration was adjusted from 0.2 to 0.5 M before the following purification step on a HisTrap column. Proteolytically released $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ was purified from the MBP tag by IMAC (1 ml HisTrap) in a buffer containing 50 mM Tris-HCl (pH 8.0) and 0.5 M NaCl. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. Fractions containing $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ were combined, concentrated and subjected to SEC on a Superose 12 column in a buffer containing 50 mM Tris-HCl (pH 8.0) and 0.5 M NaCl. **A**. 4-12% gradient SDS-PA gel (Novex) showing the purification of $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ from factor Xa-cleaved MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$. The arrows indicate the expected locations of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ (~50 kDa), MBP (~42.5 kDa) and $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ (~7.9 kDa), respectively. Lane 1, protein marker (Novex); lane 2, soluble fraction of cell lysate following induction; lane 3, unbound material from amylose column; lane 4, bound material of amylose column that specifically eluted with maltose (before factor Xa cleavage); lane 5, after factor Xa cleavage; lane 6, combined eluted peak fractions of IMAC (SEC load); lane 7, combined eluted peak fractions of SEC. **B**. Elution profile of $_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ during SEC.

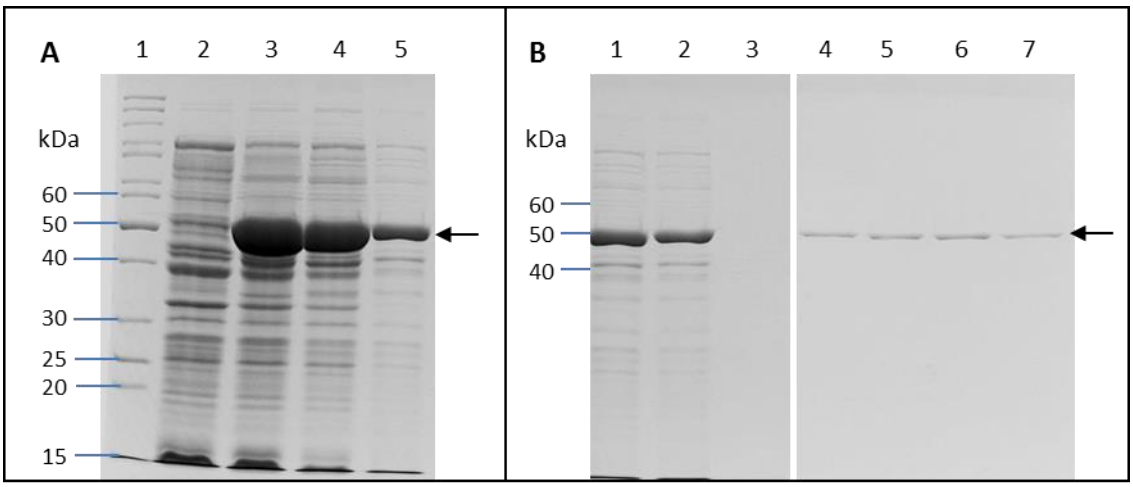


Figure 5.23 Overproduction and purification of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e. Coomassie blue-stained 10% SDS-PA gels showing the induction of pMAL-c5X-His$_6$.*tssA$^{S}_{CTD}$*V3e in *E. coli* strain NEB express with 0.3 mM IPTG at 37^{o}C. Following induction, the soluble fraction of the cell lysate was applied onto a HisTrap HP column (1 ml) in a buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. The arrows indicate the expected locations of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e based on its MW (~52 kDa). **A**. Overexpression and solubility of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. Purification of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e by IMAC. Lane 1, soluble fraction of cell lysate following induction; lane 2, flow-through from HisTrap column; lane 3, wash of protein loaded on HisTrap column before elution step; lanes 4-7, HisTrap fractions corresponding to the peak of the UV trace at 280 nm.

5.3.6.3 TEM analysis of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12\text{-}H14}$ and MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e

Negative strain TEM was carried out to compare the difference between the MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12\text{-}H14}$ tetramer and MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e. MBP fusions of both proteins were purified as described in Section 5.3.6.2, and subjected to TEM. Samples were prepared as described in Section 2.8.1. The results showed MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12\text{-}H14}$ was observed as amorphous particles, whereas MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e still formed rings (Figure 5.24), but with an uneven external edge contributed by MBP. This provided further evidence of the crucial function of Helix 15 in the formation of the higher order ring structures of TssA^{S}.

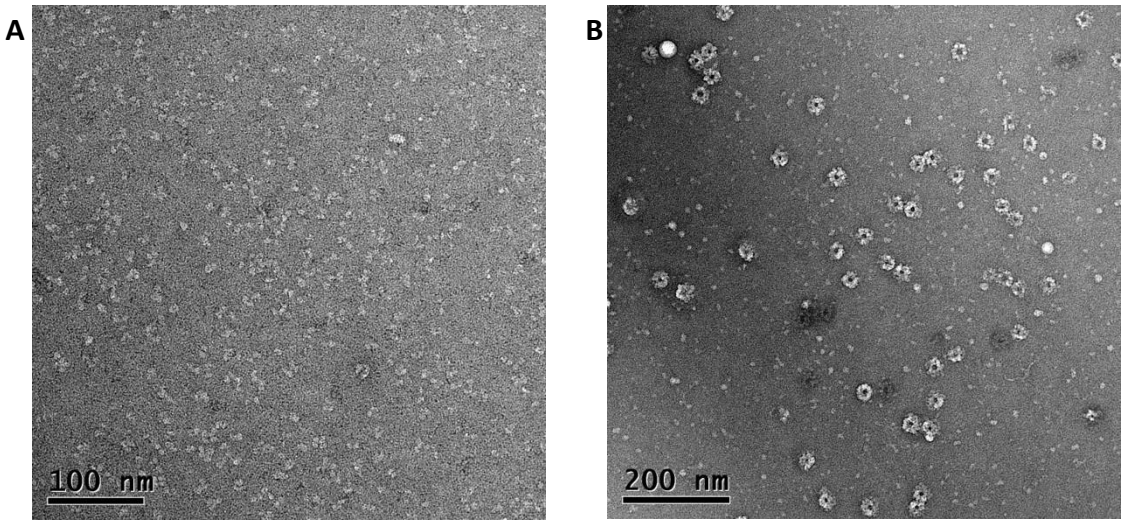


Figure 5.24 Negative stain TEM of MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ and MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e. A. MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e.$_{H12-H14}$ (0.028 mg/ml) in buffer containing 50 mM Tris-HCl pH 8.0 and 500 mM NaCl. **B**. MBP-$_{His6}$.TssA$^{S}_{CTD}$V3e (0.025 mg/ml, right side) in buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole were loaded onto freshly glow-discharged grids and stained with 0.75% uranyl formate stain before being analysed on TEM (CM200, Philips).

5.3.7 X-ray structure determination of $TssA^{S}_{CTD}$

5.3.7.1 Structure of $TssA^{S}_{CTD\cdot H12\text{-}H13}$

While attempting to solve the structure of $TssA^{S}_{CTD}$, a type of crystal was obtained resulting from proteolysis of $TssA^{S}_{CTD}$V3b in which only Helix 12 and Helix 13 were present in 0.1 M citric acid, 20% PEG6000 (pH 5.0). The crystal structure of $TssA^{S}_{CTD.H12\text{-}H13}$ (I303-W347) was solved at 1.78 Å by seleno-methionine-incorporated protein using multiple-wavelength anomalous dispersion (MAD). The structural work was carried out by Hayley Owen, Department of Molecular Biology and Biotechnology, University of Sheffield. It was found that the regions R306-T323 and P328-D341 are folded into alpha-helices that corresponded to H12 and H13, respectively, that was predicted by Psipred (Figure 1.10). The two helices are anti-parallel and connected by a loop. The loop is comprised of four residues (EPHS) that includes two of the four amino acids (EP) that are conserved in all $TssA^{S}$ proteins (Figure 1.9). The structure is dimeric, and the two chains are orientated in a head to tail position (Figure 5.25).

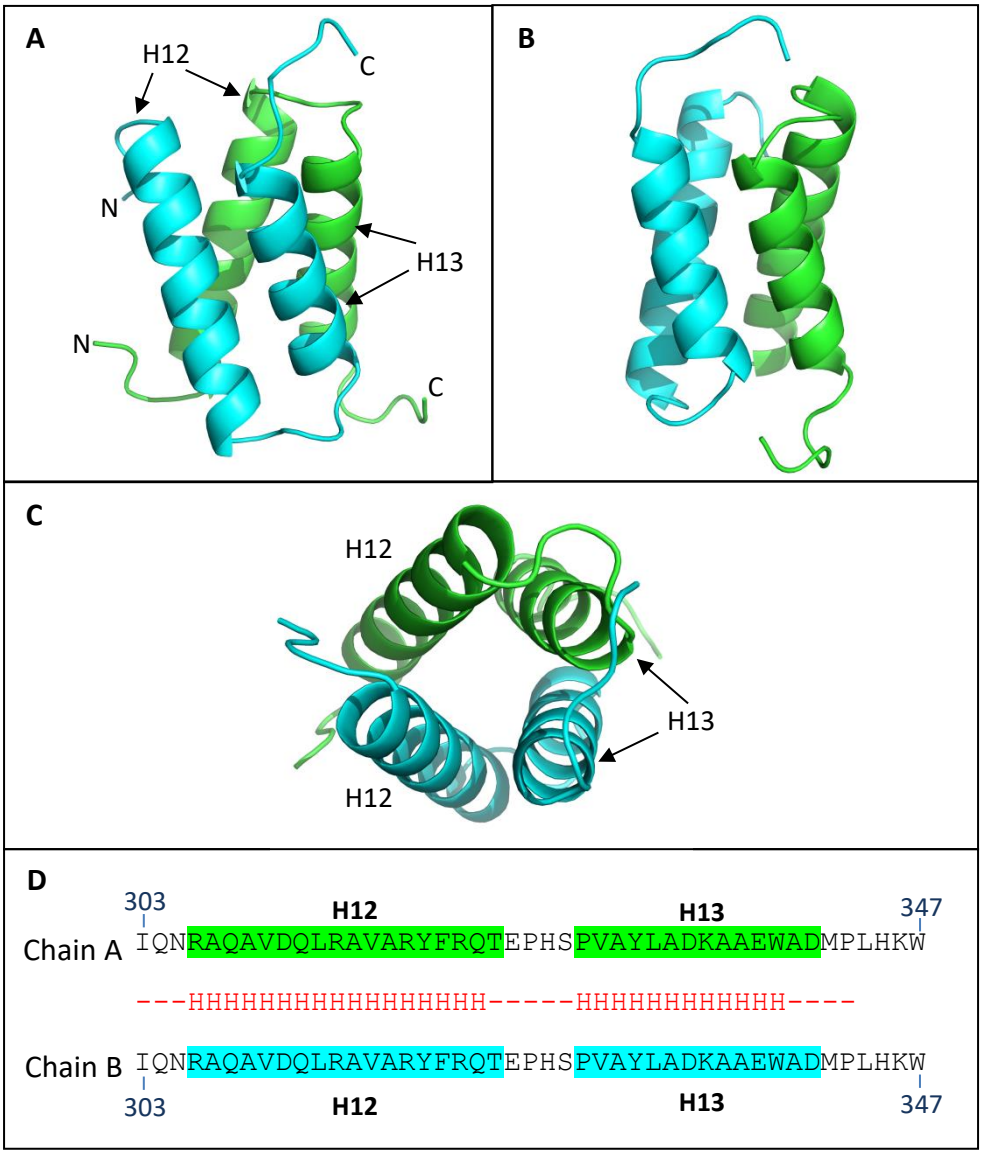


Figure 5.25 Structure of TssA$^{S}_{CTD\cdot H12\text{-}H13}$. A and B. Side views of the structure of the TssA$^{S}_{CTD\cdot H12\text{-}H13}$ dimer. **C**. Top view of the structure of the TssA$^{S}_{CTD\cdot H12\text{-}H13}$ dimer. **D**. Amino acid sequence of the chains and comparison with Psipred prediction. The two chains are labelled with different colours. The helix numbers are labelled above the sequences in bold black font with the amino acids of helices are highlighted in green and cyan. Dashes shown in red are Psipred predicted coil; 'H' in red are Psipred predicted helices. Residues I303 to W347 were resolved in the structure.

5.3.7.2 Structure of $TssA^{S}_{CTD.H12-H14}$

$TssA^{S}_{CTD}$ was predicted to be organised into four alpha-helices by Psipred (Figure 1.10), $TssA^{S}_{CTD.H12-H14}$ was estimated to be a tetramer by SEC as an MBP-$TssA^{S}_{CTD.H12-H14}$ fusion protein and in the free form following proteolytic removal of MBP (Sections 5.3.6.2.3 and 5.3.6.2.4), an attempt at solving the crystal structure of $TssA^{S}_{CTD.H12-H14}$ was also made. $_{His6}.TssA^{S}_{CTD}V3e._{H12-H14}$ was purified by proteolytic removal of MBP from MBP-$_{His6}.TssA^{S}_{CTD}V3e._{H12-H14}$ as described in Section 5.3.6.2.4. A crystal of $TssA^{S}_{CTD.H12-H14}$ was formed in 0.18 M tri-ammonium citrate and 20% (w/v) PEG 3350, and the structure of $TssA^{S}_{CTD.H12-H14}$ was solved at 2.35 Å by molecular replacement using the previous $TssA^{S}_{CTD.H12-H13}$ structure model (Figure 5.26 A and B). The structural work was carried out by Hayley Owen, Department of Molecular Biology and Biotechnology, University of Sheffield. $TssA^{S}_{CTD.H12-H14}$ has a dimeric structure rather than tetrameric. Residues 303-358 were observed to fold into three alpha helices as predicted by Psipred. However, there were some differences in the precise location of the helices (Figure 5.26 C). H12 comprises residues R306-T323 in both chains in agreement with the structure obtained for $TssA^{S}_{CTD.H12-H13}$. However, the length of H13 and H14 differs in chains A and B. Moreover, H13 of Chain B was one amino acid shorter than in the structure of the $TssA^{S}_{CTD}$ degradation product ($TssA^{S}_{CTD.H12-H13}$). H13 in Chain A is comprised of P328-D341, which contains 1 amino acid (D341) longer than that in Chain B. In addition, the length of H14 was different in the two chains by the presence of more ordered amino acids in one chain compared to the other, i.e. H14 in Chain A (residues L344-V352) is 4 amino acids shorter than that in Chain B (P343 and D355). The organisation of Helices 12 and 13 in both chains are consistent with that of dimeric $TssA^{S}_{CTD.H12-H13}$, i.e. the two helices are anti-parallel, connected by a loop. Although the length of H14 differs in 2 chains, it is perpendicular to H13 in the corresponding chain, forming a projection.

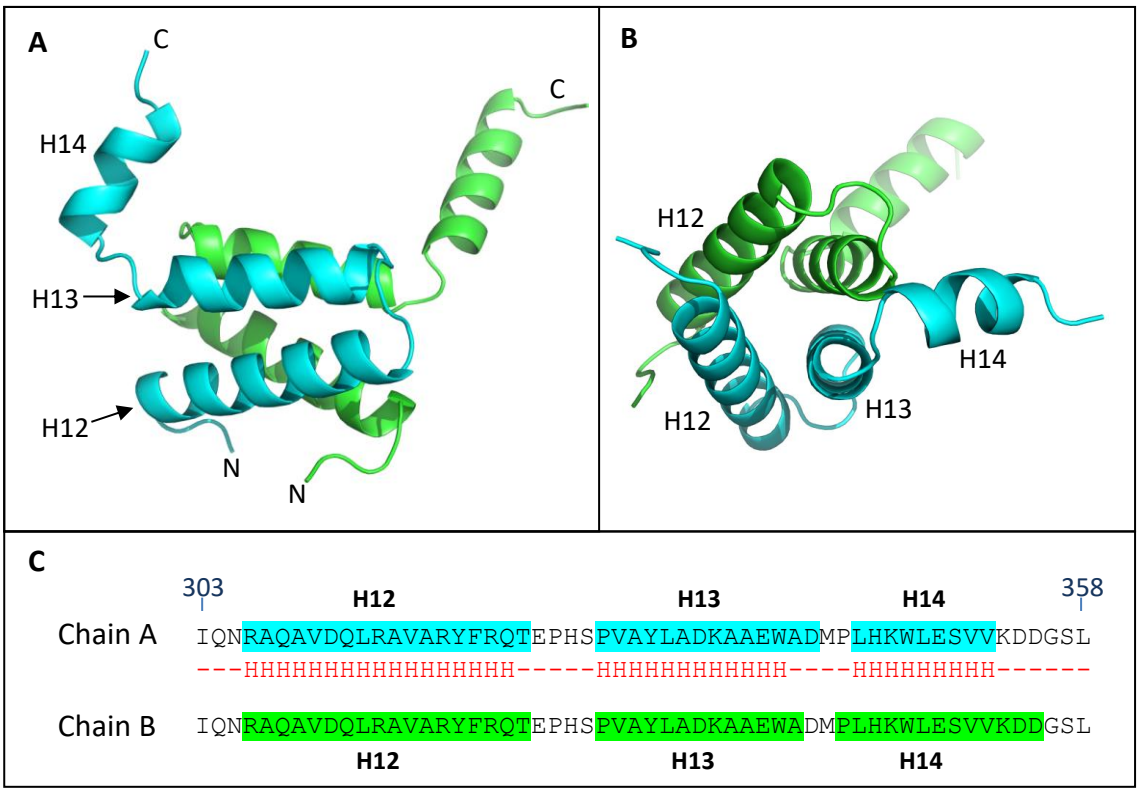


Figure 5.26 Structure of TssA$^{S}_{CTD\cdot H12\text{-}H14}$. A. Side view of the structure of the TssA$^{S}_{CTD\cdot H12\text{-}H14}$ dimer. **B**. Top view of the structure of the TssA$^{S}_{CTD\cdot H12\text{-}H14}$ dimer. Chain A and Chain B are shown in cyan and green, respectively, in panels A and B. **C.** Comparison of helix lengths in chains A and B of TssA$^{S}_{CTD\cdot H12\text{-}H14}$. Amino acid sequence of the chains are highlighted in cyan for Chain A and green for Chain B. The helix numbers are labelled above the sequences in black bold font. Dashes in red are Psipred predicted coil; 'H' in red are Psipred predicted helices. Residues I303 to L358 were resolved in the structure.

5.3.7.3 Structure of TssA$^{S}_{CTD}$

The crystal structure of TssA$^{S}_{CTD}$ was determined following proteolytic removal of MBP from MBP-TssA$^{S}_{CTD}$V3b. The purification of TssA$^{S}_{CTD}$V3b (I303-S373) was performed as described in Section 5.3.3.2.3. Crystal was obtained in 1 M ammonium sulfate, 2% PEG3350 and 0.1M Bis-Tris (pH 6.0). The structure of TssA$^{S}_{CTD}$ was solved by seleno-methionine single wavelength anomalous dispersion (SAD), and has a 3.03 Å resolution. The structural work was carried out by Hayley Owen, Department of Molecular Biology and Biotechnology, University of Sheffield. In the structure, there are 32 subunits forming a dual ring layer structure with 16-fold symmetry in the A axis dimension and two-fold symmetry in B and C axes dimensions. The inner and outer diameters are ~110 Å and ~200 Å, respectively, with ~20 Å in depth (Figure 5.27 A and B). The two layers of the ring possess opposite orientations. Each monomer contains four alpha helices, H12 (R306-T323), H13 (P328-D341), H14 (L344-V352) and H15 (D355-L365). The residues contained in H12, H13 and H14 are consistent with those of Chain A of the dimeric TssA$^{S}_{CTD}$ H12-H14 structure (Section 5.3.7.2). However, H12 in four of the 32 chains are longer than others, which contain six extra amino acids, SHM generated from the pMALc5X vector-encoded amino acids, I303, Q304 and N305. The remaining H13-H15 helices in these four chains are consistent with other.

In the TssA$^{S}_{CTD}$ structure, each monomer is involved in two pairs of hydrophobic interactions. H12 and H13 of Chain A interact with H12 and H13 of adjacent Chain B in the other layer of the ring; whereas H14 and H15 of Chain B interact with that of adjacent Chain C in the same layer of ring as Chain A. The residues involved in hydrophobic interactions are A309, L313, V316 and F320 from H12, V329, A333, A336 and A340 from H13, L344, W347, L348, V351 and V352 from H14, and I361 and L365 from H15 (Figure 5.27 C and D). The inter-docking interaction among chains contributes to the overall dual ring structure of the full-length TssA$^{S}_{CTD}$. The C-termini of the subunit domains are projecting up or downwards perpendicular to the plane of the ring. The overall structure of TssA$^{S}_{CTD}$ supports the prediction based on the TEM that the N-terminal domain of TssA^{S} locates on the exterior of the TssA$^{S}_{CTD}$ ring structure.

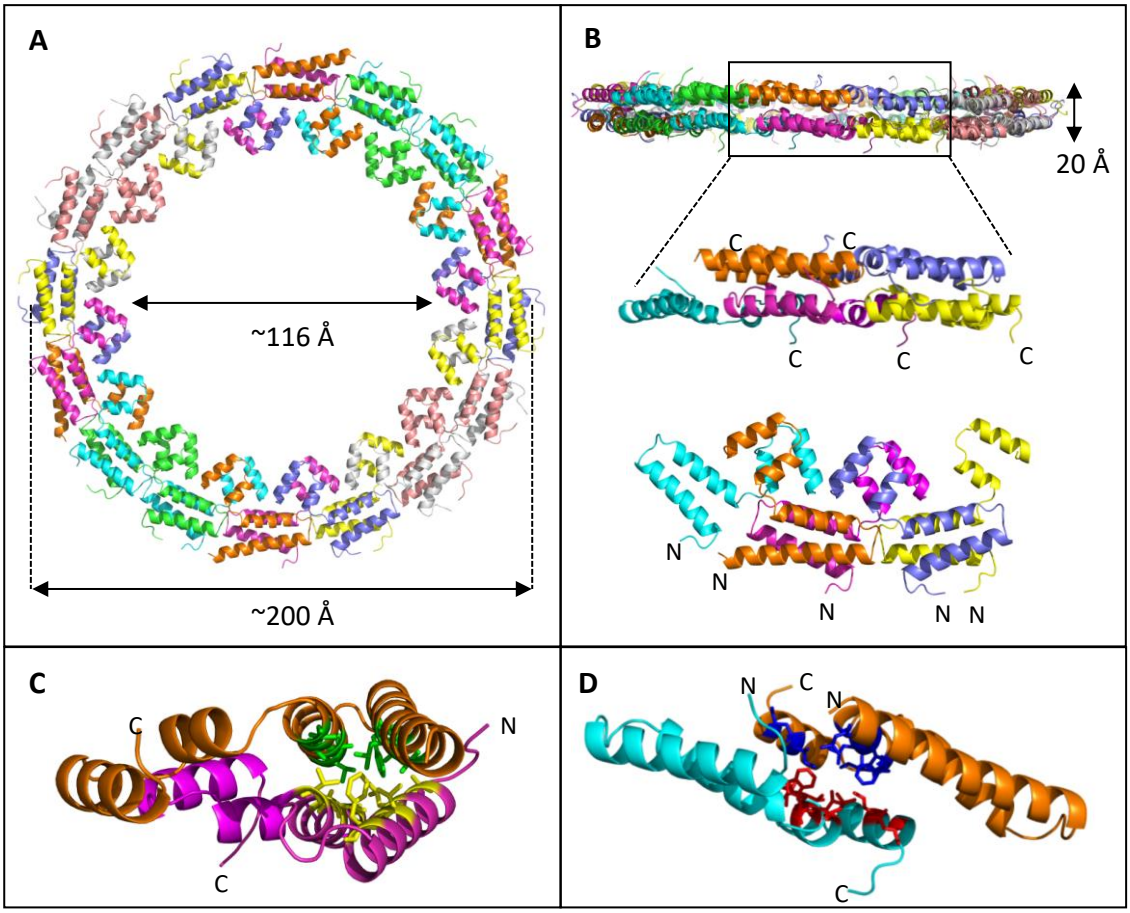


Figure 5.27 Structure of TssA$^{S}_{CTD}$. A. Top view of the TssA$^{S}_{CTD}$ structure. Dimensions are as indicated. **B**. Side view of the TssA$^{S}_{CTD}$ structure. The depth is as indicated. The black rectangle encloses chains that are shown enlarged in the two images underneath in side and top views, respectively. **C**. Hydrophobic interaction between H12 and H13 of adjacent chains. Chain A is shown in magenta and Chain B is coloured in orange. Residues involved in hydrophobic interactions of H12 and H13 are shown in yellow in Chain A and green in Chain B. **D**. Hydrophobic interaction between H14 and H15 of adjacent chains. Chain B is coloured in orange and Chain C is shown in cyan. Residues involved in hydrophobic interactions of H14 and H15 are shown in blue in Chain B and red in Chain C. Residues I303 to P369 were resolved in the structure.

5.4 Discussion

TssA^{S} plays an essential role in the assembly of the T6SS as it has been demonstrated to interact with many other T6SS subunits by the bacterial two- and three-hybrid analysis (Chapter 3). These *in vivo* interactions between TssA^{S} and other T6SS subunits were subsequently confirmed *in vitro* (Chapter 4). Although evidence of that TssA^{S} is able to oligomerise into a large complex has been provided in previous studies (Shastri 2011; Ahmad 2013), its structure is still unknown.

TssA^{S}, together with TssA^{E} is classified as two major groups of TssA based on our bioinformatics analysis according to the sequence divergence in the C-terminal region. A representative of TssA^{S} derived from the *Burkholderia cenocepacia* was subjected a structural analysis as reported in this chapter.

5.4.1 Domain organisation of TssA^{S}

The domain boundaries were proposed by multiple sequence alignment among TssA^{S} orthologues present in other Gram-negative bacteria (Figure 1.9) and the secondary structure prediction by Psipred (Figure 1.10), which suggested that TssA^{S} folds into a NTR of ~250 amino acids and a smaller conserved C-terminal region CTD of ~70 amino acids. Both NTR and CTD are rich in alpha-helices and are connected by a non-structured interdomain linker of 53-56 amino acids. The multiple sequence alignment shows part of NTR (W56-P113) is also conserved in other TssA-like proteins including TalT. The conserved region (W56-P113) is annotated as ImpA-rel_N domain. Furthermore, the region C-terminal to P113 until the end of the predicted alpha helical region (E250) is not even conserved at the amino acid sequence level in the TssA^{S} class.

The domain organisation of TssA^{S} was further supported by analysis of proteolytic degradation products (Ahmad 2013) and Section 5.2.4. The locations of the proteolytic cleavages largely occurred in the predicted interdomain linker region that connects NTR and CTD based on multiple sequence alignment and Psipred secondary structure prediction. In addition, the X-ray crystallographic results of TssA$^{S}_{NTR}$ and TssA$^{S}_{CTD}$ suggest that NTR is folded into two domains, i.e. NTD (H1-H5, residues L5-E106) and MD (H6-H11, residues I120-E250), with the MD being further organised into two

subdomains, i.e. $TssA^{S}_{MD.SD1}$ (H6, H10 and H11) and $TssA^{S}_{MD.SD2}$ (H7-H9). An additional four alpha helices (H12-H15) comprise its CTD.

Therefore, based on these analyses, $TssA^{S}$ contains a NTD (L5-E106) and a MD (I120-E250) that is connected to a small CTD (I303-S373) by a long interdomain linker (Q251-G302).

5.4.2 $TssA^{S}$ oligomerises into a high molecular mass complex with its CTD serves as the oligomerisation module

In agreement with the BACTH results showing that $TssA^{S}$ self-associates (Chapter 3), $TssA^{S}$ was observed to oligomerise into a large complex by SEC about 1.5 MDa (Ahmad 2013; this study). The results of BACTH analysis also suggested that $TssA^{S}_{CTD}$ is responsible for the oligomerisation of $TssA^{S}$ (Ahmad 2013; Chapter 3) and this observation was supported by later investigations into the oligomerisation ability of $TssA^{S}_{NTR}$ and $TssA^{S}_{CTD}$. $TssA^{S}_{NTR}$ exists as a monomer, whereas, $TssA^{S}_{CTD}$ was shown to be a multimer.

The $TssA^{S}$ complex was visualised by TEM to form ring-like structures in the dimension of ~450 Å with a number apparently discrete protein projections surrounding the central ring which has an internal diameter of 110-115 Å with the outer diameter measuring 170-190 Å. Unfortunately, it was not possible to address the number of surrounding discrete projections after averaging 1,200 particles of $TssA^{S}$, and also no symmetry information was obtained. In agreement with a previous study (Ahmad 2013), $TssA^{S}_{CTD}$ was shown to form ring structures without outer projections. These results indicate that these discrete protein projections are formed by $TssA^{S}_{NTR}$.

As SEC estimates the protein MW based on the assumption that the protein is globular (i.e. it measures the hydrodynamic volume), SEC-MALS analysis was carried out for $TssA^{S}$ and $TssA^{S}_{CTD}$, which gives a more accurate estimation of the molecular mass of non-globular proteins, i.e. it will exclude the empty space of the central hole. SEC-MALS combines a sensitive light scattering detector to detect the scattering off the axis of the light beam caused by interaction with the molecules, and a refractive index detector for determining an accurate concentration of the solute. The different refractive index of solute depends on the molar mass of large particles on the scattering angle, and also an

excess scattering on the concentration of solute. The light scattering signal from aggregates will show angular dependence. Therefore, upon determining accurate concentration of aggregates, the scattering angle can be empolyed to estimate its size (Wyatt 1991). SEC-MALLS gives consistent results for $TssA^S$ and $TssA^S_{CTD}$, suggesting $TssA^S$ comprises 30 subunits, and $TssA^S_{CTD}$ is necessary and sufficient for the $TssA^S$ multimer formation.

Further investigations into the role of individual Psipred predicted helices in $TssA^S_{CTD}$ oligomerisation using SEC suggests that the first helix (H12) of $TssA^S_{CTD}$ exists as a monomer, and H13 allows $TssA^S_{CTD.H12\text{-}H13}$ dimerise. In addition, H14 is responsible for dimerisation of the H12 and H13 dimers into tetramers while H15 allows assembly of tetramers into the $TssA^S_{CTD}$ ring. The essential role of H15 in assembly of the $TssA^S_{CTD}$ ring was also demonstrated by TEM, in which MBP-$TssA^S_{CTD}$ fusion protein forms ring structures, whereas without H15 in the fusion protein ring structures are replaced by amorphous particles. Although the buffers conditions used for visualising these two proteins are slightly different in the aspects of NaCl concentration and presence or absence of glycerol, the NaCl concentration and usage of glycerol have been tested for not having negative effects on the ring structure formation of $TssA^S$ earlier in this chapter (Section 5.2.2).

As $TssA^S$ contains 30 subunits and the presence of 30 MBP tags surrounding the $TssA^S_{CTD}$ ring in the MBP-$TssA^S_{CTD}$ fusion protein gave rise to broader rings with an irregular outer edge in EM images instead of a well-defined inner ring surrounded by a number apparently discrete projections in the native $TssA^S$, this resulted in an assumption that $TssA^S_{NTR}$ may weakly self-associate, as MBP exists as monomer. Evidence can also be found by the BACTH system that $TssA^S_{NTR}$ can interact with $TssA^S$, but not for the self-association of $TssA^S_{NTR}$ (Section 3.3.1). It is possible that the ring formation of $TssA^S_{CTD}$ increases the local concentration of $TssA^S_{NTR}$ and this leads to its self-association. According to this hypothesis, without the $TssA^S_{CTD}$, the interaction of free $TssA^S_{NTR}$ with itself is so weak that it is unlikely to occur expect at very high concentration. Therefore, NTR is arrayed and tethered to the $TssA^S_{CTD}$ by the interdomain linker.

5.4.3 X-ray crystallographic analysis of $TssA^{S}_{NTR}$ and $TssA^{S}_{CTD}$

X-ray crystallographic analysis of $TssA^{S}_{NTR}$ revealed that it is monomeric and contains 11 alpha helices. Moreover, the structure shows that the NTR is folded into two domains, i.e. NTD (H1-H5) and MD (H6-H11), with the latter further organised into two subdomains.

The $TssA^{S}_{CTD}$ structure was solved (residues I303-S373) and demonstrated to form a dual-ring with 16-fold symmetry, containing 32 subunits (two more than predicted by SEC-MALS). The dimensions of the ring are largely consistent with that obtained from TEM of the $TssA^{S}$, the inner dimension of which is large enough to accommodate the tail tube (85-90 Å). The structure may correspond to the extended state of the T6SS apparatus. However, how the lumen of $TssA^{S}_{CTD}$ can be closed or undergoes conformational changes in the contracted state, if this is what happens, is still unclear. The structures of H12-H13, and H12-H14 of $TssA^{S}_{CTD}$ were solved separately and both were shown to exist as dimers. The dimeric structure of $TssA^{S}_{CTD.H12-H14}$ was obtained instead of the estimated tetrameric structure (in solution) and this may be because the assembly beyond the dimer with the additional H14 is not biologically relevant, or alternatively the crystal form is not biologically relevant. In addition, the projection formed by H14 may cause an inaccuracy of the molecular weight estimation by SEC. Therefore, SEC-MALS may be employed as a useful method to give a more precise estimation of the oligomerisation state of $TssA^{S}_{CTD.H12-H14}$ in solution.

As $TssA^{S}_{CTD}$ was shown to be 30-mer in solution but forms a 32-mer in the crystal structure, this is reminiscent of bacteriophage HK97 gp6, which forms an oligomeric ring consisting of 13 subunits in the crystal structure (middle ring of HK97 connector) but is proposed to be a 12-mer ring when assembled into a phage that is homologous to SPP1 gp15 (Cardarelli et al. 2010). The assembly of gp6 into a 13-mer in the absence of other phage components may occur due to the possibility that the 13-mer is a lower energy structure than the 12-mer that assembles in the phage (Cardarelli et al. 2010).

It has been determined that the tail tube protein TssD assembles into hexameric rings from diverse bacteria that possess a T6SS which has an identical symmetry with the phage baseplate (Mougous et al. 2006; Jobichen et al. 2010; Osipiuk et al. 2011; Douzi et al. 2014). However, the TssJLM cell envelope chamber complex has been shown to have a five-fold symmetry resolved at 12 Å resolution by EM (Durand et al. 2015). In addition,

the crystal structure of 16-fold symmetric $TssA^{S}_{CTD}$ may not reflect the situation in solution or a low-energy structure was formed during crystallisation in the absence of other T6SS components. Co-crystallisation of $TssA^{S}$ and other T6SS components may help to solve the potential inconsistency. However, how the inconsistency of the symmetry among T6SS components allows them to be assembled successfully and functionally remains to be elucidated.

5.4.4 Location of $TssA^{S}$ in the T6SS

As the current model of T6SS is considered to be similar to bacteriophage with contractile tails where both have a cell puncturing machinery, the location of $TssA^{S}$ in the overall T6SS model is worth being discussed. $TssA^{S}$ has been shown to interact with cytoplasmic T6SS components TssK, TssL and $TssM_{NTD}$ both *in vivo* and *vitro* (Chapters 3 and 4), suggesting part or all of $TssA^{S}$ is located in the cytoplasm.

TEM showed $TssA^{S}_{CTD}$ forms a ring structure surrounded by $TssA^{S}_{NTR}$ that is tethered to the ring by the interdomain linker of $TssA^{S}$. This is supported by the $TssA^{S}_{CTD}$ structure where the N-terminal amino acids of each CTD face outwards. The $TssA^{S}_{CTD}$ ring has an internal diameter of ~110 Å, which is strikingly similar in dimension to the central channel of the TssBC tail sheath (100-110 Å) (Bönemann et al. 2009; Lossi et al. 2013; Kube et al. 2014). In addition, $TssA^{S}$ has been shown to interact with both TssB and TssC *in vivo* by the BACTH system (Chapter 3), and *in vitro* evidence for the interaction between $TssA^{S}$ and TssC was also obtained (Chapter 4) (*in vitro* interaction between $TssA^{S}$ and TssB was not tested). As the outer diameter of the $TssA^{S}_{CTD}$ ring structure is ~200 Å which cannot be accommodated inside the TssBC tail sheath channel, $TssA^{S}$ may therefore locate at either end of the TssBC tail sheath.

$TssA^{S}$ is also shown to interact with the T6SS tail tube (TssD) both *in vitro* and *in vivo* (this study). The central lumen of $TssA^{S}$ is sufficient in size to accommodate a hexameric TssD ring (85-90 Å) (Mougous et al. 2006; Ballister et al. 2008). TssD, together with the TssBC tail sheath polymerises dynamically in the cytoplasm, and it is suggested that the TssD hexamers are used as a scaffold for polymerisation of the TssBC tail sheath (Mougous et al. 2006; Ballister et al. 2008; Brunet et al. 2014). However, the role of $TssA^{S}$ in polymerisation of TssD, and TssBC is still unknown.

Apart from TssK, $TssA^S$ is also shown to interact with three other baseplate components, TssE, TssF and the core-region of TssI both *in vivo* and *in vitro*. The T6SS baseplate has been shown to be anchored in the inner membrane by a network interaction of components that constitute the cell envelope chamber and baseplate, including contacts between TssK and TssM, TssK and TssL, and TssM and TssG (Zoued et al. 2013; Brunet et al. 2015). As $TssA^S$ interacts with components of the cell envelope chamber complex that is located at the IM ($TssM_{NTD}$ and TssL), the baseplate (TssE, TssF, TssK and TssI) and the tail sheath (TssB and TssC), this supports the earlier hypothesis that $TssA^S$ is likely to locate within or close to the baseplate.

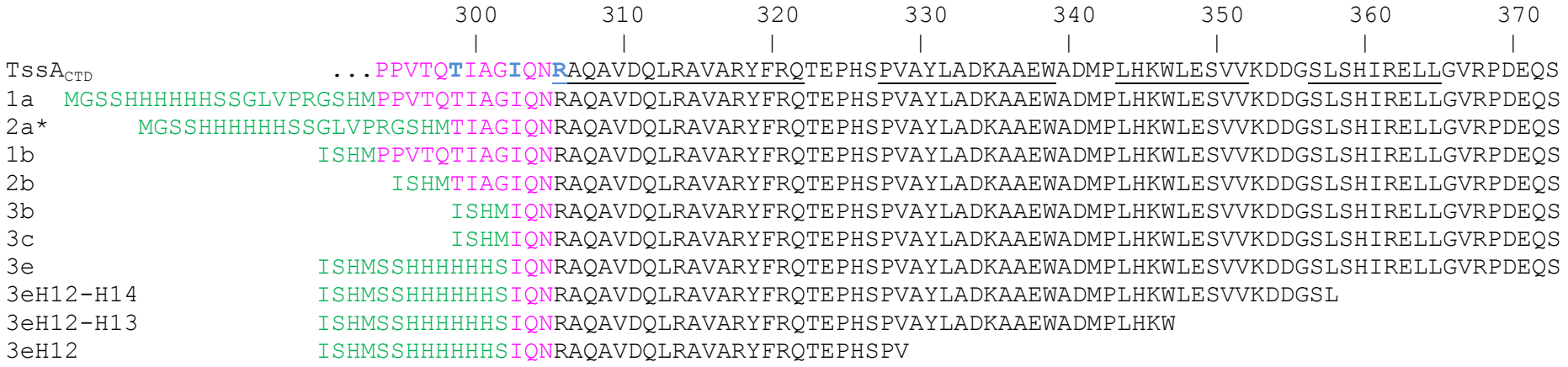


Figure 5.28 TssA$^{S}_{CTD}$ derivatives used in this study. The *B. cenocepacia* TssA$^{S}_{CTD}$ amino acid sequence is shown at the top of the figure along with 12 amino acids of the interdomain linker (magenta) as predicted by Psipred v3.3. Amino acids that are predicted to form alpha-helices are underlined. The locations of N-terminal amino acid of TssA$^{S}_{CTD}$ based on proteolytic cleavage (T299), amino acid sequence alignment (I303) and Psipred v3.3 (R306) are shown in bold blue font. Sequences below correspond to TssA$^{S}_{CTD}$ derivatives discussed in this work, including non-native tag sequences (green font). The designations on the left refer to suffixes used to describe each derivative. The number refers to the N-terminus of TssA$^{S}_{CTD}$ (1, P294; 2, T299; 3, I303), while the lower case letter refers to the plasmid source and the presence or absence of a tag (a, pET14b expressing His-tagged TssA$^{S}_{CTD}$ derivatives; b, pMAL-c5X expressing MBP-TssA$^{S}_{CTD}$ derivatives; c, pET14b expressing His-tagged TssA^{S} derivatives with a factor Xa cleavage site and ISHM tag located in the interdomain linker; d, pET14b expressing His-tagged TssA^{S} derivatives with only a factor Xa cleavage site located in the interdomain linker; e, pMAL-c5X expressing His-tagged TssA$^{S}_{CTD}$ and truncated derivatives thereof). *, constructed incorrectly (see text for details).

Table 5.1 Summary table of TssA^S protein overexpression

Protein	Residues	Plasmid	Host strain and induction condition[a]	Overexpression and solubility
$_{LinkerHis6}$.TssA^S	1-373	pACYCDuet-linkerHis$_6$.*tssA*^{S}	BL21(λDE3), 1 mM IPTG at 30^{o}C	Overproduced and soluble
Native TssA^S	1-373	pET14b-*tssA*^{S}	BL21(λDE3), 1 mM IPTG at 37^{o}C	Overproduced and soluble
$_{His6}$.TssA$^S_{NTR}$	1-254	pET14b-His$_6$.*tssA*$^S_{NTR}$	BL21(λDE3), 1 mM IPTG at 37^{o}C	Overproduced and soluble
$_{His6}$.TssA$^S_{CTD}$V1a	294-373	pET14b-His$_6$.*tssA*$^S_{CTD}$V1a	C41(λDE3), 1mM IPTG at 30/37^{o}C expressed C43(λDE3), 1mM IPTG at 37^{o}C expressed	Overproduced but insoluble
$_{His6}$.TssA$^S_{CTD}$V2a[b]	299-373	pET14b-His$_6$.*tssA*$^S_{CTD}$V2a	Defective construct	Defective construct
TssA$^S_{CTD}$V2b^C	299-373	pMAL-c5X-*tssA*$^S_{CTD}$V2b	NEB Express, 0.3 mM IPTG at 37^{o}C	Overproduced and soluble
TssA$^S_{CTD}$V3b^C	303-373	pMAL-c5X-*tssA*$^S_{CTD}$V3b	NEB Express, 0.3 mM IPTG at 37^{o}C	Overproduced and soluble
MBP-TssA$^S_{CTD}$V2b	299-373	pMAL-c5X-*tssA*$^S_{CTD}$V2b	NEB Express, 0.3 mM IPTG at 37^{o}C	Overproduced and soluble
TssA$^S_{CTD}$V3c^C	303-373	pET14b-His$_6$.linkerXa.*tssA*^{S}	BL21(λDE3), 1 mM IPTG at 37^{o}C	Overproduced and largely soluble
MBP-$_{His6}$.TssA$^S_{CTD}$V3e.$_{H12}$	303-329	pMAL-c5X-His$_6$.*tssA*$^S_{CTD}$V3e.$_{H12}$	NEB Express, 0.3 mM IPTG at 37^{o}C	Overproduced and soluble
MBP-$_{His6}$.TssA$^S_{CTD}$V3e.$_{H12-H13}$	303-347	pMAL-c5X-His$_6$.*tssA*$^S_{CTD}$V3e.$_{H12-H13}$	NEB Express, 0.3 mM IPTG at 37^{o}C	Overproduced and soluble
MBP-$_{His6}$.TssA$^S_{CTD}$V3e.$_{H12-H14}$	303-358	pMAL-c5X-His$_6$.*tssA*$^S_{CTD}$V3e.$_{H12-H14}$	NEB Express, 0.3 mM IPTG at 37^{o}C	Overproduced and soluble
$_{His6}$.TssA$^S_{CTD}$V3e.$_{H12-H14}$	303-358	pMAL-c5X-His$_6$.*tssA*$^S_{CTD}$V3e.$_{H12-H14}$	NEB Express, 0.3 mM IPTG at 37^{o}C	Overproduced and soluble
MBP-$_{His6}$.TssA$^S_{CTD}$V3e	303-373	pMAL-c5X-His$_6$.*tssA*$^S_{CTD}$V3e	NEB Express, 0.3 mM IPTG at 37^{o}C	Overproduced and soluble

[a], not all the tested induction conditions are listed. Full information is provided in the relevant section.
[b], constructed incorrectly, see text for detail
[c], proteins purified from overproduced MBP fusion proteins or protease cleaved protein

Chapter 6 Characterisation of $TssA^{EI}$

6.1 Introduction

TssA^{EI} has been shown to self-interact both *in vivo* and *in vitro* in a previous study (Ahmad 2013). It was also shown to involve in similar interactions with many other T6SS subunits as TssA^{S} by the bacterial two-hybrid analysis (Chapter 3), and the majority of these *in vivo* interactions between TssA^{EI} and other T6SS subunits were subsequently confirmed *in vitro* (Chapter 4). TssA^{EI} is therefore proposed to have functional similarity to TssA^{S}. In this chapter, a structural analysis of TssA^{EI} is carried out and discussed. *A. hydrophila* AHA1844 was used as a representative of the TssA^{EI} class of TssA proteins for this analysis.

6.2 Structural analysis of TssA^{EI}

6.2.1 Overproduction and purification of TssA^{EI}

$_{His6}$.TssA^{EI} was overexpressed from BL21(λDE3) cells containing pACYCDuet-His$_6$.*tssA^{EI}* plasmid, in which *tssA^{E}* was amplified from *A. hydrophila* ATCC 7966 and cloned into a T7 promoter driven pACYCDuet plasmid between *Bam*HI and *Hin*dIII restriction sites (Ahmad 2013). This resulted in addition of a hexa-histidine tag coded by the vector to the N-terminus of TssA^{EI}. Induction of protein synthesis was carried out in cells growing at 30^{o}C with 1 mM IPTG. $_{His6}$.TssA^{E} was overproduced evidenced by the presence of a large amount protein of the expected size (53.3 kDa). $_{His6}$.TssA^{EI} remained in the soluble fraction upon cell lysis as judged by SDS-PAGE analysis (Figure 6.1 A).

Overproduced $_{His6}$.TssA^{EI} was purified by IMAC on a HisTrap HP column (1ml) with an elution gradient of increasing concentration of imidazole up to 500 mM in the column buffer (50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol). As abundant material was used for purification initially, a large amount of $_{His6}$.TssA^{EI} remained in the unbound material even after recycling the flow-through onto the column (Figure 6.1 B). The bound $_{His6}$.TssA^{EI} was eluted from the column in the range 100-200 mM imidazole monitored by the UV trace at 280 nm. The peak fractions were collected and concentrated (MWCO 30 kDa, Vivaspin centrifugal concentrator) before being loaded onto Superose 6 GL for further purification and estimation of the oligomerisation state.

The elution profile of $_{His6}$.TssA^{EI} on Superose 6 showed a single peak at 19.55 minute corresponding to a protein of apparent MW of 726 kDa calculated using the standard calibration curve (Figure 6.1 C and D). As the monomeric MW of $_{His6}$.TssA^{EI} is ~53 kDa, this indicates $_{His6}$.TssA^{EI} is able to oligomerise into a large complex which for a globular protein would be estimated to contain ~14 monomers.

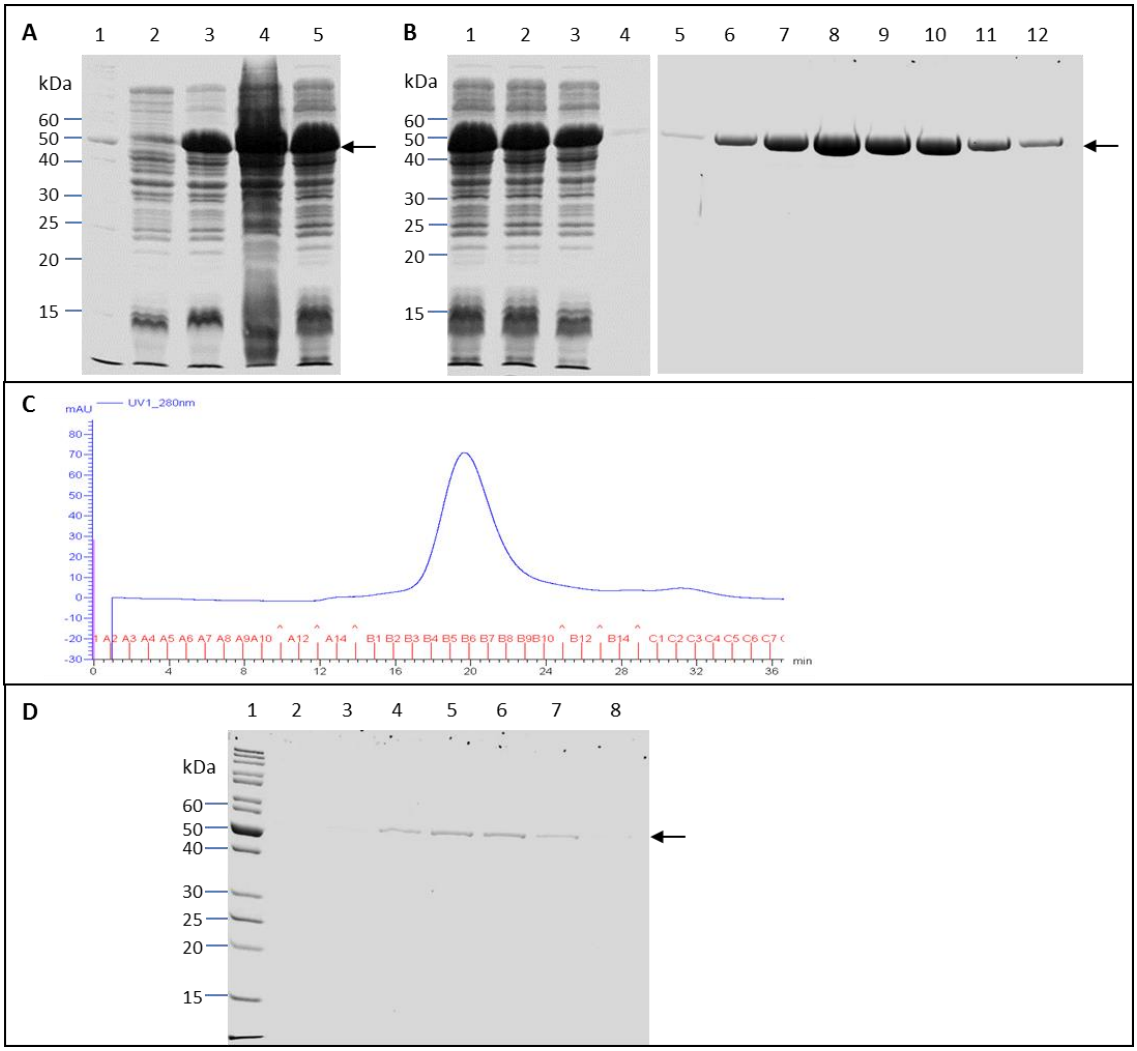
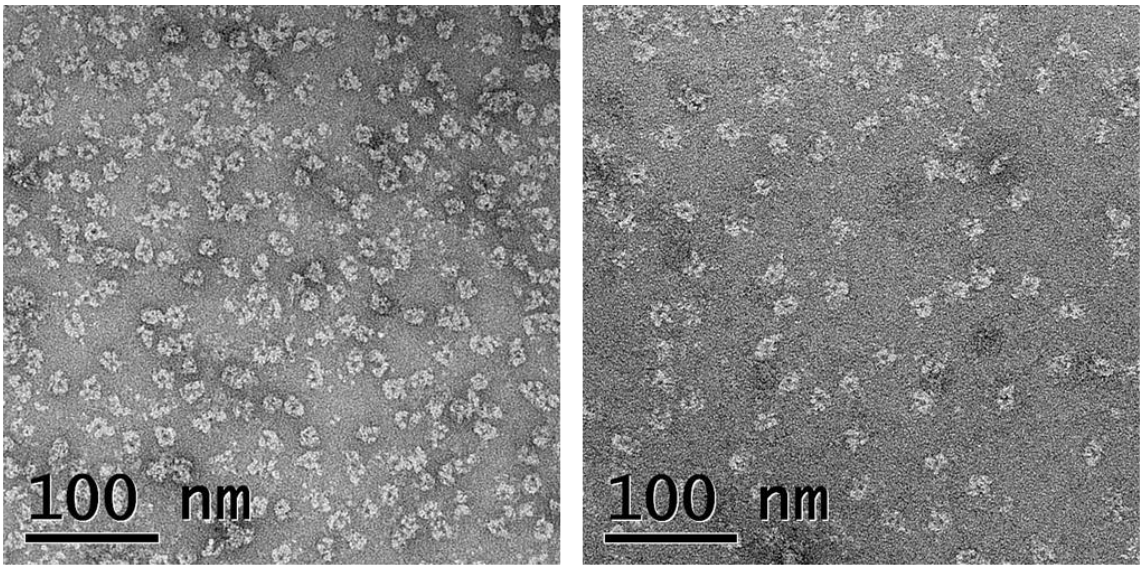


Figure 6.1 Overproduction, solubility and purification of $_{His6}$.TssA^{EI}. $_{His6}$.TssA^{EI} synthesis was induced from pACYCDuet-His$_6$.*tssA^{EI}* in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 30°C. Following induction, the soluble fraction of the cell lysate was subjected to IMAC (HisTrap) in the column buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. This was followed by SEC (Superose 6 GL) in the same column buffer (10 mM imidazole). **A**. Overproduction and solubility of $_{His6}$.TssA^{EI} analysed by electrophoresis in 12% SDS-PA gels. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. Purification of $_{His6}$.TssA^{EI} by IMAC analysed by electrophoresis in 12% SDS-PA gels. Lane 1, soluble fraction of cell lysate following induction; lane 2, flow-through_1 from HisTrap column; lane 3, flow-through_2 from HisTrap column; lane 4, wash of protein-loaded HisTrap column before elution step; lanes 5-12, HisTrap fractions corresponding to the peak of the UV trace at 280 nm. **C**. Elution profile of $_{His6}$.TssA^{EI} during SEC. **D**. 12% SDS-PAGE analysis of SEC of $_{His6}$.TssA^{EI}. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 2-8, peak fractions of SEC corresponding to the peak of the UV trace at 280 nm. Arrows in A, B and D indicate the expected location of $_{His6}$.TssA^{EI} based on its MW (~53.3 kDa).

6.2.2 Transmission electron microscopy of $_{His6}$.TssA^{EI}

Since $_{His6}$.TssA^{EI} is able to oligomerise into a large complex, it was expected to form a ring-like structure similar as TssA^{S}, $_{His6}$.TssA^{EI} was purified by IMAC as described in the above section and subjected to TEM analysis. Attempts were made to optimise the TEM conditions for $_{His6}$.TssA^{EI} by preparing the pure protein in six different buffers as listed in Figure 6.2. Among these, Buffers B and D were optimum for observing the $_{His6}$.TssA^{EI} structure (Figure 6.2). Although there were some protein aggregates, $_{His6}$.TssA^{EI} was observed to have a ring-like structure, containing a distinct central hole with an irregular overall shape. The results of other buffer conditions are not shown. The inner diameter of the central hole formed by the ring-like structure will be determined in the domain analysis in Section 6.3.2.6. The exterior size estimation was carried out by averaging the radius of 30 single particles, and the result suggested it has a diameter of ~230 Å, which is approximately half the size of TssA^{S} (~450 Å).



Buffer A: 25 mM Tris-HCl, 150 mM NaCl, 20% glycerol, pH 7.4
Buffer B: 25 mM Tris-HCl, 150 mM NaCl, pH 7.4
Buffer C: 25 mM Tris-HCl, 150 mM NaCl, 0.5 M each of glutamic and arginine, pH 7.4
Buffer D: 25 mM Tris-HCl, 150 mM NaCl, 20% glycerol, 0.5 M each of glutamic and arginine, pH 7.4
Buffer E: 25 mM Tris-HCl, 150 mM NaCl, pH 6.8
Buffer F: 25 mM Tris-HCl, 500 mM NaCl, pH 7.4

Figure 6.2 Negative stain TEM of $_{His6}$.TssA^{EI}. N-terminal hexa-histidine tagged TssA^{EI} (0.06 mg/ml) in Buffer B (left side) and Buffer D (right side) were loaded onto freshly glow-discharged grids and stained with 0.75% uranyl formate stain before being analysed on TEM (CM200, Philips).

6.2.3 Molecular weight estimation of $_{His6}$.TssA^{EI} by SEC-MALS

As $_{His6}$.TssA^{EI} was observed to form a circular structure, a more accurate molecular weight estimation was carried out by SEC-MALS. The technique employed SEC on a Superdex 200 10/300 GL (G.E. Healthcare) column was applied. The peak analysis of TssA^{EI} was based on the elution profile using Astra software which defined peak regions: Peak 1 (18.72-19.46 minute), Peak 2 (17.87-21.81 minute), Peak 3 (16.02-16.67 minute) and Peak 4 (14.77-17.76 minute) (Figure 6.3 A). It is noticeable that the defined Peak 1 and Peak 2 overlapped, with Peak 2 having a broader range in elution time. This is also the case for the Peak 3 and Peak 4, with Peak 4 having the broader range. As a small the region of Peak 2 and Peak 4 overlaps, the division of the peaks was arbitrary. The $_{His6}$.TssA^{EI} elution profile indicated that there was a significant amount of material of high MW (Peak 4, ~34.4% of the material) as evidenced by the RI/UV traces close to the void volume. The tail of this large MW peak passes under a subsequent peak (Peak regions 1 and 2) and so may skew the MW estimates for those peaks upward. There was some material in the RI/UV at 34 minute but this had no LS response above baseline and must be low MW material, and it is not consistent with expectations for monomeric TssA^{EI}.

The MW was determined by a Zimm fit procedure at each point then averages calculated over the peak region (Figure 6.3 B). The MWs determined for individual peak from Peak 1 to 4 are 623 kDa, 681 kDa, 5,867 kDa and 5,101 kDa, respectively. The data showed that the large MW peak (Peak 4) was a mixture of sizes from >10 MDa down to about 1 MDa. The Peak 3 had an average of ~6 MDa but even this relatively narrow region was still quite heterogeneous. The later peak (Peak 2) MW estimate was affected by the tail of the earlier one but there was a minimum at about 600 kDa. The material in the Peak 2 was more homogeneous (the MW curve flattens out) but its actual MW would be slightly lower than 600 kDa. As the $_{His6}$.TssA^{EI} polypeptide lacking the N-terminal methionine has a calculated MW of 53.3 kDa, this corresponds to 11.25 monomers in the complex if the MW is considered as 600 kDa. However, as stated above, the MW estimation for $_{His6}$.TssA^{EI} is bigger than the actual MW. Therefore, the precise oligomerisation state, i.e. how many subunits TssA^{EI} contains, will need to be further investigated. The MW estimation of the

oligomerisation domains of $TssA^{EI}$ (Sections 6.3.2.7 and 6.3.4.5) will facilitate a more accurate oligomerisation determination for $TssA^{EI}$.

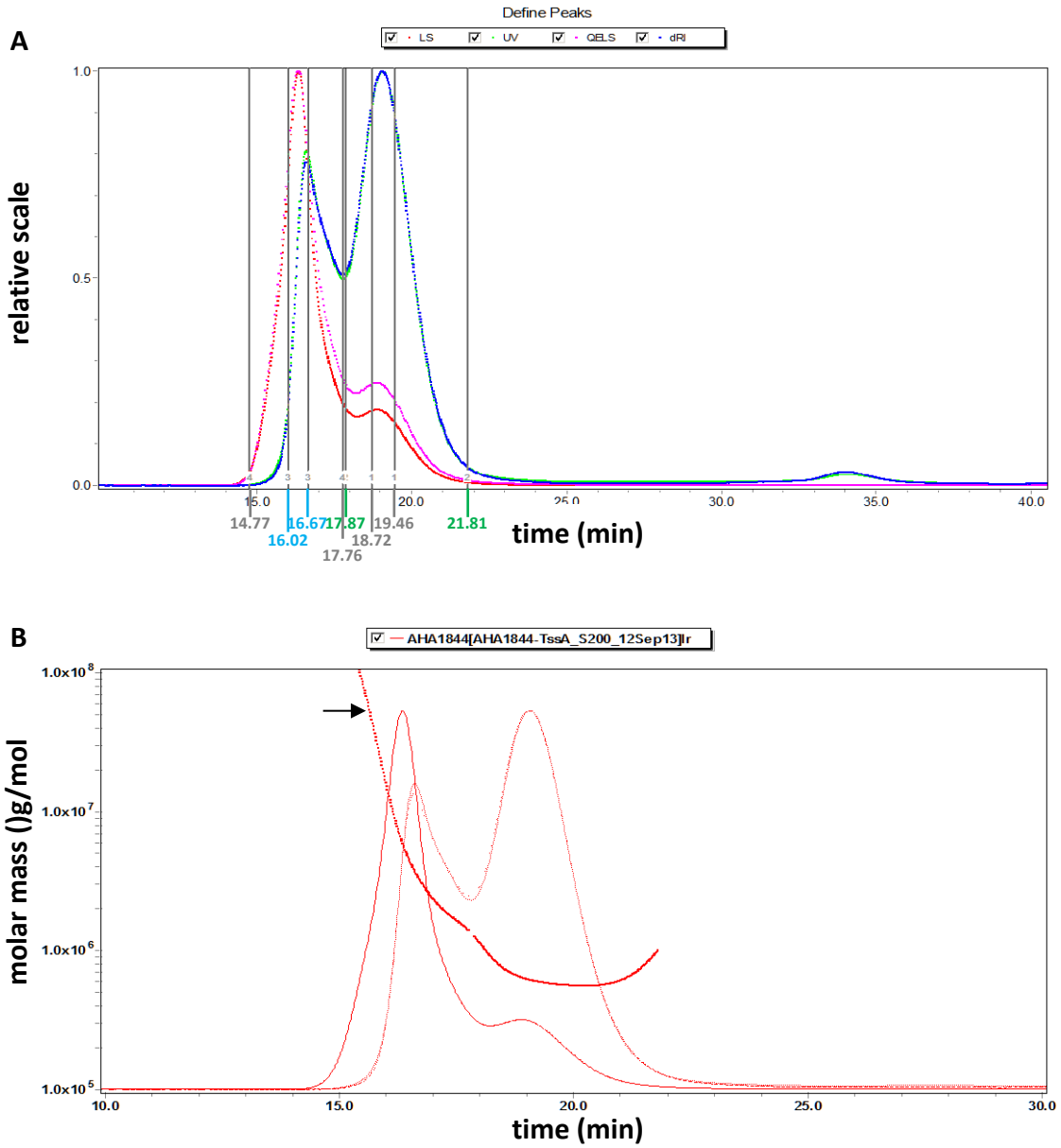


Figure 6.3 SEC-MALS analysis of TssA^{EI}. A. TssA^{EI} elution profile and peak positions. The normalised signals for light scattering (LS) are shown in red, refractive index (RI) in blue, UV absorption (280 nm) in green and quasi-elastic light scattering (QELS) in magenta. The peak regions chosen for analysis are indicated by numbered vertical grey lines that defined by Astra software. The X-axis shows elution time in minutes and Y-axis shows relative scale. **B**. TssA^{EI} molar mass plot. The MW is determined by a Zimm fit procedure at each point. Solid line represents LS trace. Dashed and dotted traces represent RI and UV responses, respectively. Molecular weight estimate curve is shown in heavier dots (pointed with black arrows). The X-axis is the elution time in minutes and the Y-axis is the MW scale. Figures provided by University of York Bioscience Technology Facility.

6.2.4 X-ray crystallography of $TssA^{EI}$

$_{His6}$.$TssA^{EI}$ was purified by IMAC and SEC as described in Section 6.2.1. The purified $_{His6}$.$TssA^{EI}$ was then subjected to X-ray crystallography trials. However, no hint of crystal formation was obtained for $_{His6}$.$TssA^{EI}$. Crystal trials were eventually discontinued, due to time constraints, and a change of strategy to obtain crystal structures from $_{His6}$.$TssA^{EI}$ domains was initiated.

6.2.5 Proteolytic analysis of $TssA^{EI}$

Limited proteolysis of $TssA^{EI}$ was carried out to identify domain boundaries based on the increased susceptibility of exposed or unfolded regions (loops and interdomain linkers) to cleavage by proteases.

6.2.5.1 Peptide analysis by mass spectrometry

$_{His6}$.$TssA^{EI}$ was overproduced in *E. coli* BL21(λDE3) cells, and purified by IMAC (HisTrap) and SEC (Superdex 200). The purification steps can refer as described in Section 6.2.1, although a Superose 6 GL column was used during SEC. $_{His6}$.$TssA^{EI}$ eluted from the Superdex 200 column at ~61.7 ml, corresponding to a protein of apparent MW of 850 kDa. The purified $_{His6}$.$TssA^{EI}$ was prepared at 3.7 mg/ml in buffer containing 175 mM NaCl, 16.5 mM Tris-HCl (pH 8.0) and subjected to limited cleavage by a variety of proteases, i.e. *S. aureus* V8 Glu-C endopeptidase (Sigma), Gly-C endopeptidase (Protogen) and Trypsin type I (Sigma). The proteases were prepared at a concentration of 1 mg/ml in water, and the reaction was carried out in 20 µl of $_{His6}$.$TssA^{EI}$ by adding 1 µl of the protease solution. The mixtures were incubated for 90 minutes at room temperature and reactions were stopped by adding 20 µl of 10% acetic acid. During endopeptidase digestion, samples were taken at different time points (0, 5, 30 and 90 minutes). The digested samples were analysed on SDS-PA gels (Figure 6.4 A). It was observed that a polypeptide migrating at ~31 kDa was present in the preparation before protease digestion, which may be a natural degradation product of the full-length $_{His6}$.$TssA^{EI}$ that occurred during or after purification. The Gly-C endopeptidase digested $_{His6}$.$TssA^{EI}$ more efficiently and there was no visible amount of full-length protein remaining at the longer time points. In contrast, there was a small

amount of full-length $_{His6}$.TssA^{EI} present in the Glu-C and tryptic endopeptidase digestions. The digested samples were then sent for mass spectrometry (MS) analysis (Figure 6.4 B).

The MS result of Glu-C endopeptidase digestion products showed three major peptides, which are 31,171 Da (suggested residues 200-492 of $_{His6}$.TssA^{EI}), 22,128 Da (suggested residues 2-199 of $_{His6}$.TssA^{EI}) and 9,653 Da (suggested residues 407-492 of $_{His6}$.TssA^{EI}). There was also a small amount of peptides with a lower MW detected (<5,300 Da) in the digested sample. These small peptides were not identical to TssA^{E}, and they may have arisen from protein contaminants. The digestion after E199 released two of the major peptides. Although it was expected to find sets of peptides in the mixture which would add up to the entire sequence of $_{His6}$.TssA^{E}, there was no peptide corresponding to residues 200-406 of $_{His6}$.TssA^{EI} (21,537 kDa) as a result of having a peptide corresponding to residues 407-492 of $_{His6}$.TssA^{EI} (9,653 kDa). This could be because these peptides were unstable and were digested into smaller fragments.

The Gly-C endopeptidase digestion gave rise to several large peptides with identity to $_{His6}$.TssA^{EI} and one peptide of ~10 kDa but no small peptides in the analysis. From the MS analysis, these were 28,212 Da (residues 235-492 of $_{His6}$.TssA^{EI}), 25,085 Da (residues 2-234 of $_{His6}$.TssA^{EI}), 23,079 Da (residues 280-492 of $_{His6}$.TssA^{EI}), 21,928 Da (residues 203-412 of $_{His6}$.TssA^{EI}) and 10,851 Da (residues 393-492 of $_{His6}$.TssA^{EI}). The latter three peptides occurred due to cleavage at positions located one or two amino acid(s) N-terminal to a glycine but after an alanine, which was unexpected. The first two peptides listed above, which were generated by digestion after G220 (G234 of $_{His6}$.TssA^{EI}), are joinable and represent the entire $_{His6}$.TssA^{EI} residues (Figure 6.4 B).

The trypsin endopeptidase digestion resulted in generation of five peptides with identity to $_{His6}$.TssA^{EI}, which are 21,359 Da (residues 2-192 of $_{His6}$.TssA^{EI}), 10,852 Da (residues 393-492 of $_{His6}$.TssA^{EI}), 9,198 Da (residues 411-492 of $_{His6}$.TssA^{EI}), 3,241 Da (residues 99-128 of $_{His6}$.TssA^{EI}) and 1,281 Da (residues 177-188 of $_{His6}$.TssA^{EI}). None of these were joinable (Figure 6.4 B). There were few peptides that could not be matched to $_{His6}$.TssA^{EI} (results not shown).

Although it was expected that peptides coving the entire length of $_{His6}$.TssA^{EI} would be present in the digested samples, the lack of many expected peptides may be explained as

during preparation of the sample for MS in formic acid and methanol, and during first step of analysis on reverse phase column some of the peptides precipitated or absorbed on the beads. The presence of the contaminating peptides could be explained by the impurity of the original protein or they may have originated from the protease itself.

6.2.5.2 Peptide analysis by N-terminal sequence analysis

$_{His6}$.TssA^{EI} was overproduced in *E. coli* BL21(λDE3) cells, and purified by sequential IMAC (HisTrap HP) and anion exchange chromatography (HiTrap Q HP) in an attempt to improve its purity. The purification on the HisTrap column was performed as described in Section 6.2.1. The peak elution fractions from the HisTrap column were combined for further purification on a HiTrap Q column. However, there was no significant improvement in purity of $_{His6}$.TssA^{EI} compared with a second step using SEC (result not shown). The purified $_{His6}$.TssA^{EI} was prepared at 3.7 mg/ml in buffer containing 175 mM NaCl and 16.5 mM Tris-HCl (pH 8.0), and then subjected cleavage by *S. aureus* V8 Glu-C endopeptidase (Sigma), Gly-C endopeptidase (Protogen) and Trypsin type I (Sigma) proteases. The digestion procedures were carried out as described in Section 6.2.5.1 for 100 minutes. Samples were then subjected to SDS-PAGE analysis (result not shown) and blotted on a PVDF membrane before submitting for N-terminal sequence analysis.

The N-terminal sequence analysis results suggested natural proteolysis digestion had occurred after the ion exchange step. Three peptides resulted from natural proteolysis, the N-terminus of two of them were very close to each other, A188 and G190. Given the MW of these two peptides are about 31 kDa, the peptides are predicted to extend to the C-terminal end of TssA^{E}. The third peptide started at G383. The Glu-C digestion gave three peptides that had identity with sequences within $_{His6}$.TssA^{EI}, two of which contained the N-terminus of $_{His6}$.TssA^{EI} and the other one started from A200. Gly-C digested $_{His6}$.TssA^{EI} resulted in two peptide peptides with identity to $_{His6}$.TssA^{EI}, the N-terminal ends of both peptides corresponded to S221. The trypsin digestion generated five peptides that corresponded to sequences within $_{His6}$.TssA^{EI}, two of them comprised the N-terminal terminus of $_{His6}$.TssA^{EI}. The other three peptides released by trypsin

endopeptidase digestion had N-termini corresponding to V224, G379 and H397, respectively (Figure 6.4 B).

Overall, most of the protease cleavage sites that were identified by both MS and N-terminal sequence analysis occurred in the predicted linker regions according to both amino acid sequence alignment and secondary structure prediction (Psipred v3.3) analyses (Figures 1.11 and 1.12). The amino acid sequence alignment of $TssA^{E}$ orthologues present in other Gram-negative bacteria predicts two non-conserved regions, i.e. P130-N230 and Q375-G401. The secondary structure prediction by Psipred largely supports the amino acid sequence alignment, except two predicted helices (H7 and H8) in the region of A138-A188 of native $TssA^{EI}$ (Figure 1.12). The results from limited proteolytic analysis largely occurred in Q158-N230 and Q375-G401 regions. Therefore, combine the results of these analyses, similarly to $TssA^{S}$, $TssA^{EI}$ contains a large N-terminal domain (M1-A157), a middle domain (D231-L374) and a small C-terminal domain (I402-K478).

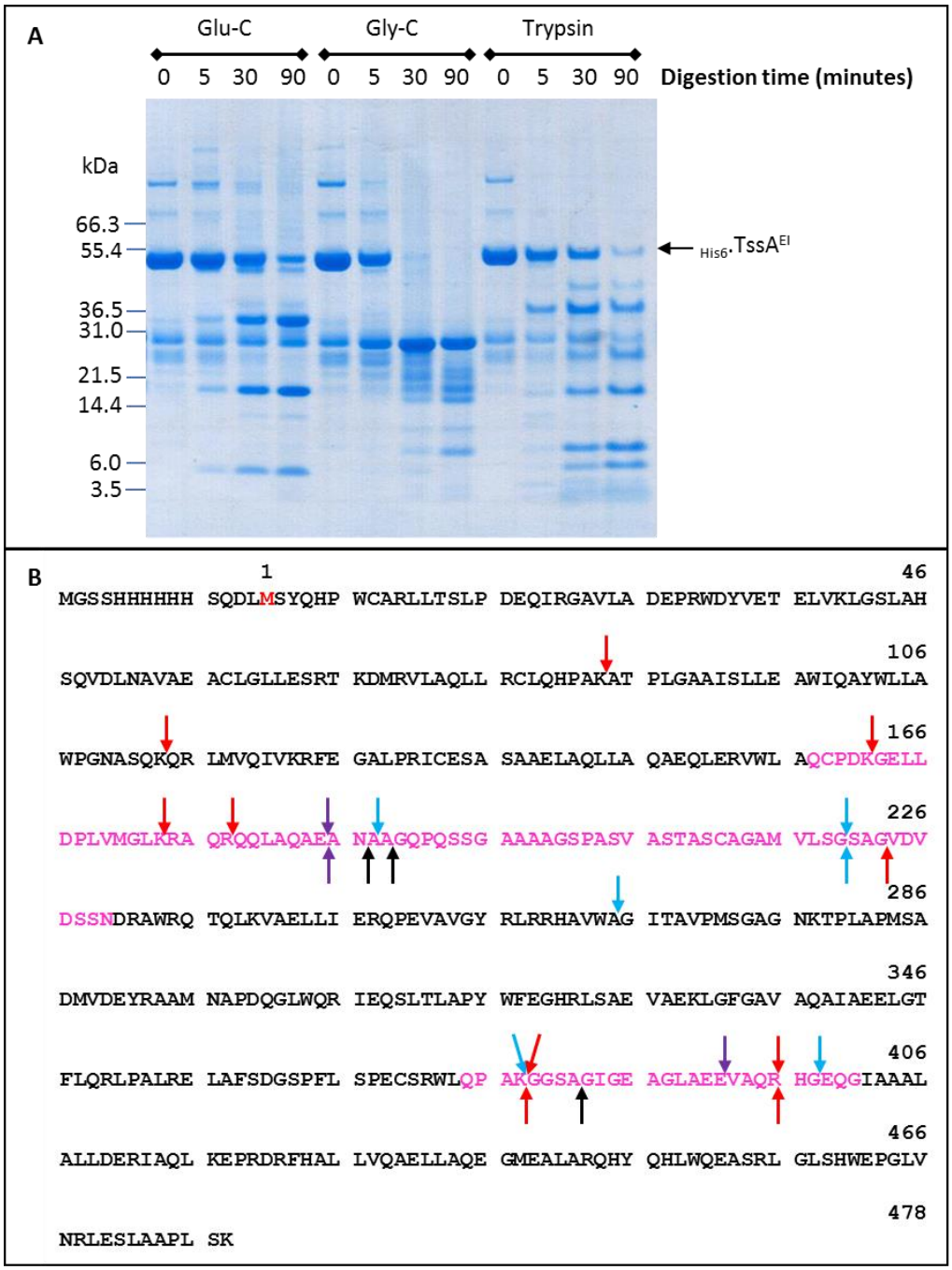


Figure 6.4 Limited proteolysis of $_{His6}$.TssA^{EI}. $_{His6}$.TssA^{EI} was purified by two different approaches, which were IMAC combined with SEC for the MS analysis of peptides, and IMAC combined with anion exchange chromatography for N-terminal analysis of peptides. *S. aureus* V8 Glu-C endopeptidase (Sigma), Gly-C endopeptidase (Protogen) and Trypsin type I (Sigma) at a concentration of 1 mg/ml were used for digestion of purified $_{His6}$.TssA^{EI} (3.7 mg/ml). **A**. 4-12% gradient SDS-PA gel (Novex) showing the endopeptidase digestion of $_{His6}$.TssA^{EI} that was purified by IMAC and SEC. Digestion time and types of endopeptidase are labelled above the gel. The arrow indicates the expected location of $_{His6}$.TssA^{E} based on its MW (~53.3 kDa). **B**. Amino acid sequence of $_{His6}$.TssA^{EI} showing the bioinformatically predicted linker regions in magenta font with proteolytic cleavage sites indicated by vertical arrows in different colours: purple, Glu-C cleavage; blue, Gly-C cleavage; red, trypsin cleavage. Down arrows indicate the cleavage positions that were identified by MS; up arrows show the locations that were identified by N-terminal sequence analysis. The initiating methionine of native TssA^{EI} is shown in red font.

6.3 Structural analysis of TssA^{EI} domains

In order to facilitate the structural analysis of TssA^{EI}, the individual domains were characterized. The domain boundaries were defined by multiple sequence alignment among TssA^{EI} orthologues present in other Gram-negative bacteria and the secondary structure prediction by Psipred, and was further supported by limited proteolysis. TssA^{EI} contains a large N-terminual domain (M1-A157) that is homologous to a corresponding region in other TssA-like proteins. It also contains a middle domain (D231-L374) and a small C-terminal domain (I402-K478). The individual domains of TssA^{EI} were overproduced and purified as described below. Different versions of each domain were generated due to various reasons such as expression problem, protein solubility, hints from crystallisation trials, trimming domain boundaries, etc., which will be specified in the relevant section. A summary of TssA^{EI} domain derivatives are shown in Figures 6.38, 6.39, 6.40 and 6.41. Table 6.1 shows a summary of protein overexpression in this chapter.

6.3.1 Overproduction and purification of TssA$^{EI}_{NTD}$ derivatives

6.3.1.1 $_{His6}$.TssA$^{EI}_{NTD}$

DNA encoding *A. hydrophila* strain ATCC 7966 AHA1844 (TssA^{E}) residues 1-229 was amplified using primers pACYC-AHA1844.for and AHA1844NTDrev. The amplified *tssA*$^{EI}_{NTD}$ DNA fragment of the expected size (~735 bp) was digested with restriction enzymes *Bgl*II and *Kpn*I that recognized sites in the non-annealing regions of the forward and reverse primers and the product was ligated to *Bam*HI and *Kpn*I digested pACYCDuet-1 plasmid. This resulted in placement of a hexa-histidine tag coding sequence in frame with the N-terminal coding sequence of *tssA*$^{EI}_{NTD}$. *E. coli* strain MC1061 was transformed with the ligation mixture and colonies were grown on LB plates containing chloramphenicol. The correct recombinant plasmids were identified by PCR screening and followed by plasmid miniprep to confirm the plasmids contain the *tssA*$^{E}_{NTD}$ inserts. The nucleotide sequence of the inserted gene in the plasmids was verified by DNA sequencing and the resulting plasmid was called pACYCDuet-His$_6$.*tssA*$^{EI}_{NTD}$.

$_{His6}$.TssA$^{EI}_{NTD}$ was overproduced in *E. coli* strain BL21(λDE3) cells containing pACYCDuet-His$_6$.*tssA*$^{EI}_{NTD}$ as evidenced by a large amount of a polypeptide migrating at the expected size of $_{His6}$.TssA$^{EI}_{NTD}$, ~26 kDa (Figure 6.5 A). The overproduced $_{His6}$.TssA$^{EI}_{NTD}$ remained in the soluble fraction following cell lysis and was purified by IMAC on a HisTrap HP column. $_{His6}$.TssA$^{EI}_{NTD}$ bound to nickel-sepharose with high efficiency, and was eluted in the range 50-183 mM imidazole (Figure 6.5 B). Peak fractions containing $_{His6}$.TssA$^{EI}_{NTD}$, as monitored by UV at 280 nm were combined and concentrated (MWCO 10 kDa) before being injected onto Superose 12 as an additional purification step and to estimate the oligomerise status of the protein. The elution profile showed a single peak at 26.38 minute (Figure 6.5 C), which corresponded to a protein of apparent MW ~29.3 kDa calculated using the calibration curve. SDS-PAGE analysis showed that the peak fractions contained $_{His6}$.TssA$^{EI}_{NTD}$ (Figure 6.5 D). Given that the monomeric MW of $_{His6}$.TssA$^{EI}_{NTD}$ is 26 kDa, it exists as a monomer in isolation. Different preparation of purified $_{His6}$.TssA$^{EI}_{NTD}$ was used for crystallisation trials.

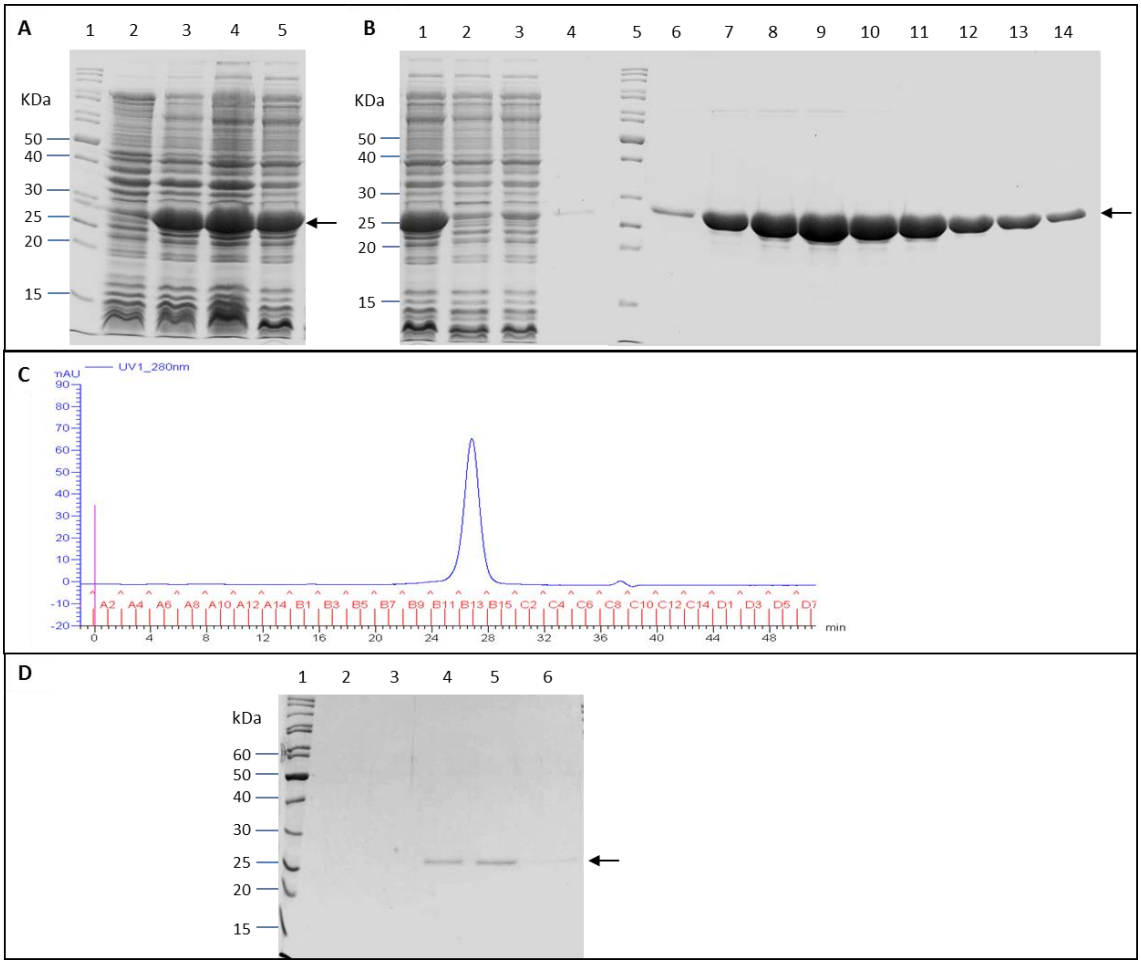


Figure 6.5 Overproduction, solubility and purification of $_{His6}$.TssA$^{EI}_{NTD}$. $_{His6}$.TssA$^{EI}_{NTD}$ synthesis was induced from pACYCDuet-His$_6$.*tssA*$^{E}_{NTD}$ in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 30^{o}C. Following induction, the soluble fraction of the cell lysate was subjected to IMAC on a HisTrap HP column (1ml, GE Healthcare) in the column buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. Fractions containing $_{His6}$.TssA$^{EI}_{NTD}$ was further purified by SEC on a Superose 12 column in buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. **A**. Overexpression and solubility of $_{His6}$.TssA$^{EI}_{NTD}$ analysed by electrophoresis in a 12% SDS-PA gel. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. 12% SDS-PAGE analysis of purification of $_{His6}$.TssA$^{EI}_{NTD}$ by IMAC. Lane 1, soluble fraction of cell lysate following induction; lane 2, flow-through_1 from HisTrap column; lane 3, flow-through_2 from HisTrap column; lane 4, wash of protein-loaded HisTrap before elution step; lane 5, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 6-14, HisTrap fractions corresponding to the peak of the UV trace at 280 nm. **C**. Elution profile of $_{His6}$.TssA$^{EI}_{NTD}$ on Superose 12. **D**. Purification of $_{His6}$.TssA$^{EI}_{NTD}$ by SEC on a Superose 12 column. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 2-6, peak fractions from the Superose 12 column according to the UV trace at 280 nm. Arrows in A, B and D indicate the expected location of $_{His6}$.TssA$^{EI}_{NTD}$ based on its MW (~26 kDa).

6.3.1.2 TssA$^{EI}_{NTD}$V2

As no hint of crystal formation was obtained for $_{His6}$.TssA$^{EI}_{NTD}$ and $_{His6}$.TssA$^{EI}_{NTD}$ contains a region that is predicted to form a long interdomain linker (Q179-S229) that may inhibit crystal formation, it was decided to construct a plasmid specifying a shorter version of TssA$^{EI}_{NTD}$ lacking this region. DNA encoding residues M1-R178 of AHA1844 was amplified using primers AHA1844.NdeI.for and AHA1844NTDrev2. The amplified *tssA*$^{EI}_{NTD}$V2 DNA fragment of the expected size (~534 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites in the amplicon incorporated by the forward and reverse primers and the plasmid MCS. Ligation, screening for correct recombinant plasmids of pACYCDuet-*tssA*$^{EI}_{NTD}$V2 and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1.

TssA$^{EI}_{NTD}$V2 (untagged) was overexpressed in a large amount in *E. coli* strain BL21(λDE3) cells containing pACYCDuet-*tssA*$^{EI}_{NTD}$V2, but it was not soluble with induction at 37^{o}C. However, when induction was carried out at 30^{o}C, the protein remained soluble (Figure 6.6 A). Following induction at 30^{o}C, TssA$^{EI}_{NTD}$V2 was purified by anion exchange chromatography on a HiTrap DEAE column (G.E. Healthcare). However, it did not bind to the column at pH 8.0 (results not shown). The protein was then purified by ammonium sulphate precipitation. Analytical ammonium sulphate cuts ranging from 1.5-3.5 M were carried out and analysed on SDS-PA gel (Figure 6.6 B). The results showed that 2.0-2.5 M precipitated the majority of TssA$^{EI}_{NTD}$V2. Subsequently, 1.5 M ammonium sulphate was performed prior to 2.2 M, which was used to precipitate the protein during large scale purification. The precipitated material was recovered by centrifugation and dissolved in 50 mM Tris-HCl (pH 8.0). The material was then subjected to SEC on Superdex 200 10/300 GL for further purification. TssA$^{EI}_{NTD}$ eluted from the column at 89.17 ml corresponding to a protein MW of ~20 kDa and was verified as TssA$^{EI}_{NTD}$V2 by SDS-PAGE (Figure 6.6 C and D). Therefore, TssA$^{EI}_{NTD}$V2 also occurs as a monomer. The other two peaks at 44.16 and 72.23 ml contained contaminations. The peak fractions of TssA$^{EI}_{NTD}$V2 from SEC (26 and 27) were combined and concentrated to 11 mg/ml, buffer exchanged into 5 mM Tris-HCl (pH 8.0) and 50 mM NaCl for crystallisation trials.

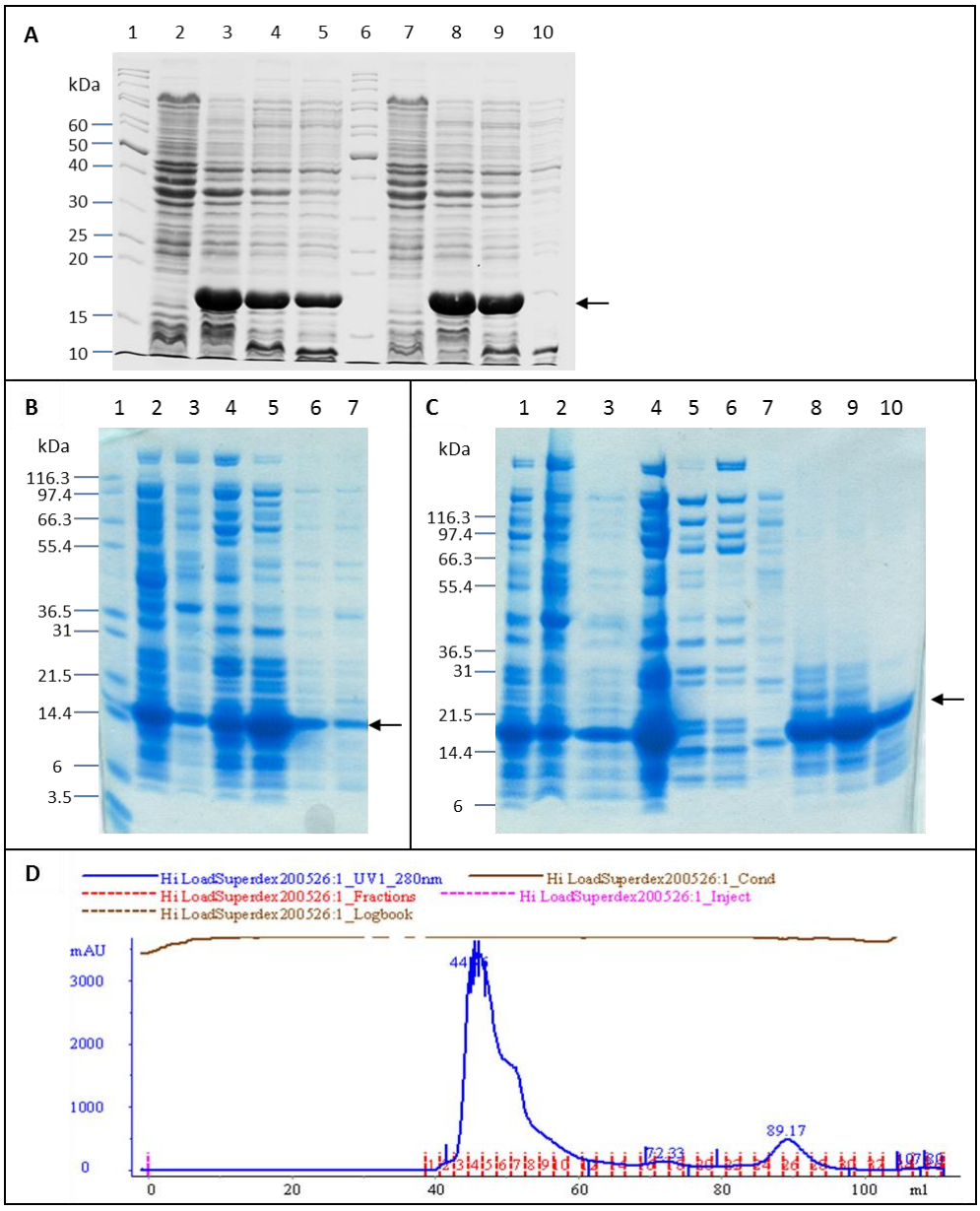


Figure 6.6 Overproduction, solubility and purification of TssA$^{EI}_{NTD}$V2.

Figure 6.6 Overproduction, solubility and purification of TssA$^{EI}_{NTD}$V2. TssA$^{EI}_{NTD}$V2 synthesis was induced from pACYCDuet-*tssA$^{EI}_{NTD}$*V2 in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 30^{o}C. Following induction, the soluble fraction of the cell lysate was purified by a 2.2 M ammonium sulphate precipitation step followed by SEC on Superdex 200 10/300 GL in buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. **A**. Coomassie blue-stained 12% SDS-PA gel showing the overproduction and solubility of TssA$^{E}_{NTD}$V2 with different induction conditions. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction at 30^{o}C; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction at 30^{o}C and cell lysis; lane 5, soluble fraction of cell lysate following induction at 30^{o}C; lane 6, EZ-Run^{TM} Rec protein ladder (Fisher); lane 7, total cell protein from uninduced cells; lane 8, total cell protein from cells following induction at 37^{o}C; lane 9, crude cell lysate containing both insoluble and soluble proteins following induction at 37^{o}C and cell lysis; lane 10, soluble fraction of cell lysate following induction at 37^{o}C. **B**. 4-12% gradient SDS-PA gel (Novex) showing the analytical ammonium sulphate cuts of TssA$^{EI}_{NTD}$V2. Lane 1, protein marker (Novex); lane 2, soluble fraction of cell lysate following induction; lane 3, 1.5 M ammonium sulphate precipitation; lane 4, 2.0 M ammonium sulphate precipitation; lane 5, 2.5 M ammonium sulphate precipitation; lane 6, 3.0 M ammonium sulphate precipitation; lane 7, 3.5 M ammonium sulphate precipitation. **C**. 4-12% gradient SDS-PA gel (Novex) showing the purification of TssA$^{EI}_{NTD}$V2. Lane 1, soluble fraction of cell lysate following induction; lane 2, 1.5 M ammonium sulphate precipitation; lane 3, supernatant of 2.2 M ammonium sulphate precipitation; lane 4, material that was loaded onto Superdex 200; lanes 5-10, SEC elution fractions No. 5, 7, 17, 25, 26 and 28, respectively. **D**. Elution profile of TssA$^{EI}_{NTD}$V2 on Superdex 200 column. Arrows in A, B and C indicate expected location of TssA$^{EI}_{NTD}$V2 based on its MW(~20 kDa).

6.3.1.3 TssA$^{EI}_{NTD}$V3

As no hint of crystal formation was obtained for TssA$^{EI}_{NTD}$V2, generation of TssA$^{EI}_{NTD}$ derivatives was considered. The Psipred predicted Helix 8 contains a proline (P168) which casts doubt on the ability of this region to adopt a helical conformation. Moreover, the proteolytic analysis also showed few cleavages occurred in the predicted H8. Therefore, it was decided to construct a plasmid specifying a shorter version of TssA$^{EI}_{NTD}$ (i.e. TssA$^{E}_{NTD}$V3) lacking predicted H8 for increasing the chance of obtaining the crystal structure. DNA encoding M1-Q158 of AHA1844 was amplified using primers AHA1844.NdeI.for and AHA1844NTDrev3. The amplified *tssA*$^{EI}_{NTD}$V3 DNA fragment of the expected size (~470 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites in the amplicon, that were added by the forward and reverse primers, and the plasmid MCS. Ligation, screening for correct recombinant plasmids (pACYCDuet-*tssA*$^{EI}_{NTD}$V3) and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1.

Overproduction of TssA$^{EI}_{NTD}$V3 (untagged) was carried out in *E. coli* strain BL21(λDE3) cells. Different induction conditions were tested as TssA$^{E}_{NTD}$V3 was overproduced but found to be insoluble with 1 mM IPTG induction at both 37^{o}C and 30^{o}C (results not shown). Attempts were also made to solubilise it by inducing with lower concentration of IPTG (0.1 mM) at both temperatures. However, TssA$^{EI}_{NTD}$V3 remained insoluble upon cell lysis (Figure 6.7). TssA$^{EI}_{NTD}$V3 was overproduced in a large amount under all the tested conditions.

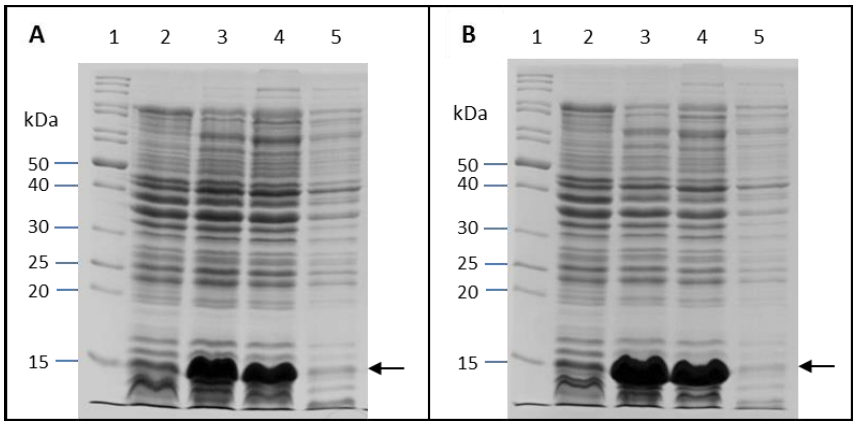


Figure 6.7 Overproduction and solubility of TssA$^{EI}_{NTD}$V3. Coomassie blue-stained 12% SDS-PA gels showing the induction of pACYCDuet-*tssA*$^{EI}_{NTD}$V3 in *E. coli* strain BL21(λDE3) cells with 0.1 mM IPTG at 30^{o}C (**A**) and 37^{o}C (**B**). The black arrows indicate the expected locations of TssA$^{EI}_{NTD}$V3 based on its MW (~17.5 kDa). For both **A** & **B**: Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

6.3.1.4 X-ray crystallography of $TssA^{EI}_{NTD}$

Purified $_{His6}$.$TssA^{EI}_{NTD}$ and $TssA^{EI}_{NTD}$V2 were subjected to crystallisation trials. However, so far, no hints of crystals have been obtained for either version of $TssA^{EI}_{NTD}$.

6.3.2 Overproduction and purification of $TssA^{EI}_{MD-CTD}$ derivatives

6.3.2.1 $_{His6}$.$TssA^{EI}_{MD-CTD}$

The N-terminal histidine tagged $TssA^{EI}_{MD-CTD}$ was constructed for facilitating purification. DNA encoding AHA1844 residues G223-K478 was amplified using primers pACYC-AHA1844.CTDfor and pACYC-AHA1844.CTDrev. The amplified $tssA^{EI}_{MD-CTD}$ DNA fragment of the expected size (~813 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Bam*HI and *Hin*dIII that recognized sites introduced into the amplicon by the forward and reverse primers and in the plasmid MCS. This resulted in fusion of the $tssA^{EI}_{MD-CTD}$ coding sequence downstream and in frame with the coding sequence of the hexa-histidine tag in the vector. Ligation, screening for correct recombinant plasmids of pACYCDuet-His_6.$tssA^{EI}_{MD-CTD}$ and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1.

$_{His6}$.$TssA^{EI}_{MD-CTD}$ was overproduced in *E. coli* strain BL21(λDE3) cells containing pACYCDuet-His_6.$tssA^{EI}_{MD-CTD}$. The protein was successfully overexpressed and remained soluble following cell lysis as concluded by the large amount of protein migrating at the expected size of 31 kDa in the soluble fraction of the cell lysate (Figure 6.8 A). Purification was carried out by IMAC on a HisTrap HP column. The unbound material was recycled for maximising the chance of $_{His6}$.$TssA^{EI}_{MD-CTD}$ binding to the resin. As a large amount of material was used for the purification, there was some unbound $_{His6}$.$TssA^{EI}_{MD-CTD}$ and the recycling step did not further enhance the protein binding (Figure 6.8 B). $_{His6}$.$TssA^{EI}_{MD-CTD}$ was eluted from the HisTrap column in the range 120-250 mM imidazole. Although the purity of $_{His6}$.$TssA^{EI}_{MD-CTD}$ was very high after IMAC, a SEC step on a Superose 12 column was employed for further purification and MW estimation. The elution profile of Superose 12 showed a single peak at 17.36

minute monitored by UV_{280}, consistent with a protein of apparent MW of ~410 kDa. As the monomeric MW of $_{His6}$.TssA$^{EI}_{MD\text{-}CTD}$ is ~31 kDa, $_{His6}$.TssA$^{EI}_{MD\text{-}CTD}$ is able to form a multimer, and the results suggest that TssA$^{EI}_{MD\text{-}CTD}$ is responsible for the formation of the large complex of the full-length TssA^{EI}. The purified $_{His6}$.TssA$^{EI}_{MD\text{-}CTD}$ was subjected to crystallisation trials upon buffer exchange into 5 mM Tris-HCl (pH 8.0) containing 50 mM NaCl using a Zeba column (Section 2.4.9.1).

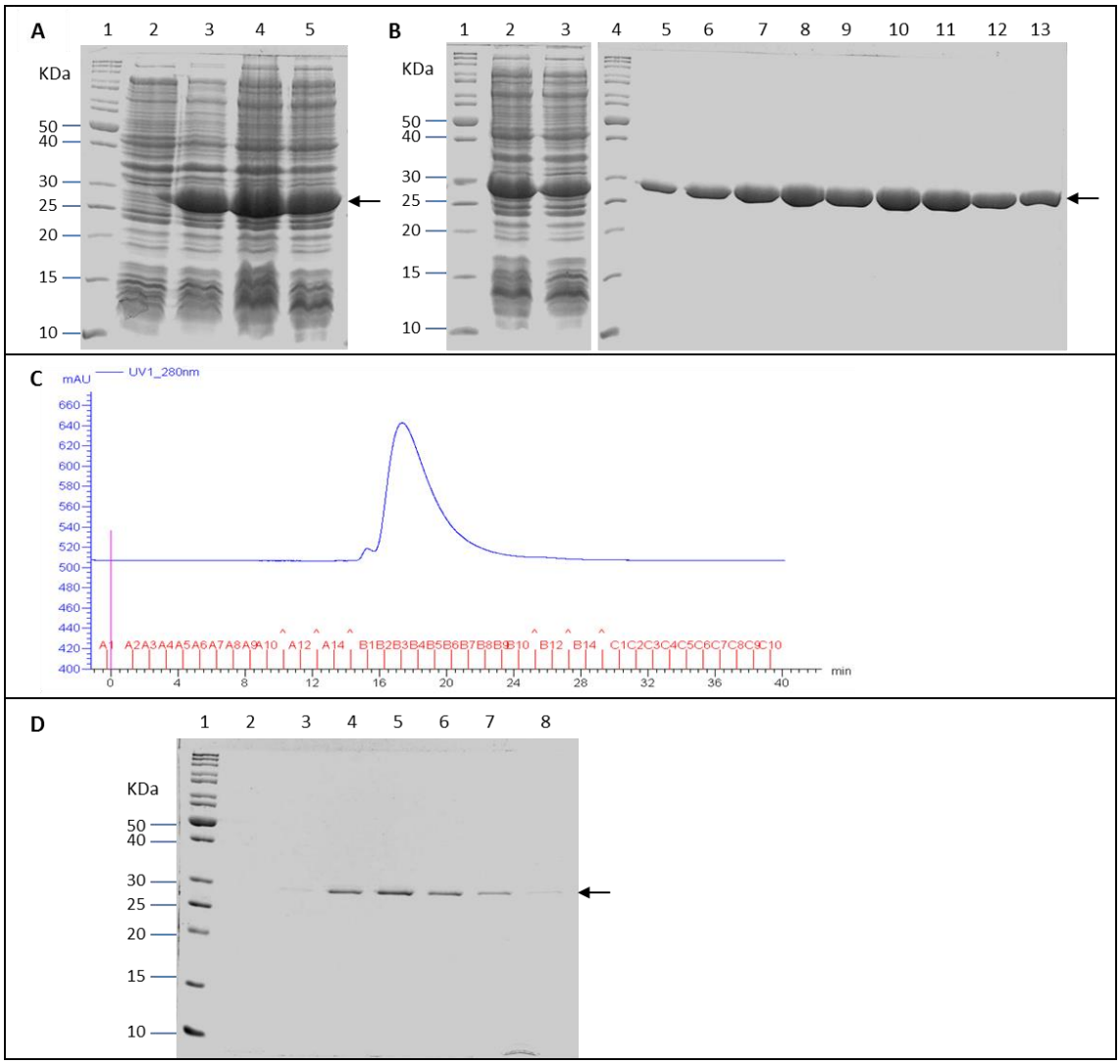


Figure 6.8 Overproduction, solubility and purification of $_{His6}$.TssA$^{EI}_{MD\text{-}CTD}$. $_{His6}$.TssA$^{EI}_{MD\text{-}CTD}$ synthesis was induced from pACYCDuet-His$_6$.*tssA*$^{EI}_{MD\text{-}CTD}$ in *E. coli* strain BL21(λDE3) cells with 1 mM IPTG at 37°C. Following induction, the soluble fraction of the cell lysate was subjected to IMAC on a HisTrap HP column (1 ml, GE Healthcare) in buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. Flow-through was recycled and re-loaded onto the HisTrap. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. The peak elution fractions as monitored by UV at 280 nm were combined, concentrated and subjected to SEC on Superose 12 in buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. **A**. Overexpression and solubility of $_{His6}$.TssA$^{EI}_{MD\text{-}CTD}$ analysed by electrophoresis in a 12% SDS-PA gel. Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. 12% SDS-PAGE analysis of purification of $_{His6}$.TssA$^{EI}_{MD\text{-}CTD}$ by IMAC. Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, flow-through_1 from HisTrap column; lane 3, flow-through_1 from HisTrap column; lane 4, EZ-Run™ Rec protein ladder (Fisher); lanes 5-13, HisTrap fractions corresponding to the peak of the UV trace at 280 nm. **C**. Elution profile of $_{His6}$.TssA$^{EI}_{MD\text{-}CTD}$ during SEC. **D**. Purification of $_{His6}$.TssA$^{EI}_{MD\text{-}CTD}$ by SEC on a Superose 12 column analysed by electrophoresis in a 12% SDS-PA gel. Lane 1, EZ-Run™ Rec protein ladder (Fisher); lanes 2-8, peak fractions from Superose 12 column corresponding to the peak of the UV trace at 280 nm. Arrows in A, B and D indicate the expected locations of $_{His6}$.TssA$^{E}_{MD\text{-}CTD}$ based on its MW (~31 kDa).

6.3.2.2 TssA$^{EI}_{MD\text{-}CTD}$V2

It was decided to generate an untagged TssA$^{EI}_{MD\text{-}CTD}$ derivative for crystallisation trial. DNA encoding residues S228-K478 of AHA1844 was amplified using primers AHA1844.CTD.NdeI.for2 and AHA1844.BglII.rev. The amplified *tssA*$^{EI}_{MD\text{-}CTD}$ DNA fragment (~756 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites in the vector MCS and which were also incorporated into the PCR product in the forward and reverse primers. Ligation, screening for correct recombinant plasmids referred to as pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V2, and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1.

E. coli strain BL21(λDE3) cells containing pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V2 was employed for TssA$^{EI}_{MD\text{-}CTD}$V2 overproduction. TssA$^{EI}_{MD\text{-}CTD}$V2 was overproduced in a huge amount during growth at 37^{o}C with a range of IPTG inducer concentrations from 1 to 0.1 mM. However, no TssA$^{EI}_{MD\text{-}CTD}$V2 remained soluble following induction with 1 mM IPTG and only ~5% remained soluble with 0.1 mM IPTG induction (Figure 6.9 A and B). Therefore, induction was carried out at a lower temperature (30^{o}C) in an attempt to increase the protein solubility. TssA$^{EI}_{MD\text{-}CTD}$V2 was still overproduced in a large amount but less than that obtained at 37^{o}C. Following cell lysis, TssA$^{EI}_{MD\text{-}CTD}$V2 remained soluble with 1 or 0.1 mM IPTG induction (Figure 6.9 C and D). The overproduced TssA$^{EI}_{MD\text{-}CTD}$V2 with 1 mM IPTG induction at 30^{o}C was purified as described below.

Purification of TssA$^{EI}_{MD\text{-}CTD}$V2 was conducted by anion exchange chromatography on a HiTrap DEAE column. The majority of TssA$^{EI}_{MD\text{-}CTD}$V2 bound to the column, and was eluted from the column in the range 200-270 mM NaCl. However, the purity of TssA$^{EI}_{MD\text{-}CTD}$V2 was not significantly improved (Figure 6.10 A). Therefore, the elution fractions of DEAE column containing TssA$^{EI}_{MD\text{-}CTD}$V2 were combined and concentrated (MWCO 10 kDa). It was noticeable by SDS-PAGE analysis that some of the TssA$^{EI}_{MD\text{-}CTD}$V2 was lost during the concentration step although a MWCO 10 kDa centrifugal concentrator had been used. The concentrated sample was loaded onto Superdex 200 HiLoad 16/600 column. The elution profile showed two peaks at 65.05 and 81.57 ml (Figure 6.10 B). After being analysed by electrophoresis, TssA$^{EI}_{MD\text{-}CTD}$V2 was confirmed to be present in the 65.05 ml peak fraction, consistent with an apparent

MW of 340 kDa. The peak fractions from SEC (17-19) were judged to be 80% pure. However, during the final concentration step, $TssA^{EI}_{MD\text{-}CTD}V2$ appeared to aggregate at a concentration of 2 mg/ml and resulted in a dramatic loss of protein. Therefore, the final preparation looks much less pure than the fractions that were obtained from SEC (Figure 6.10 A). Due to solubility problem of $TssA^{EI}_{MD\text{-}CTD}V2$, the purification was not repeated.

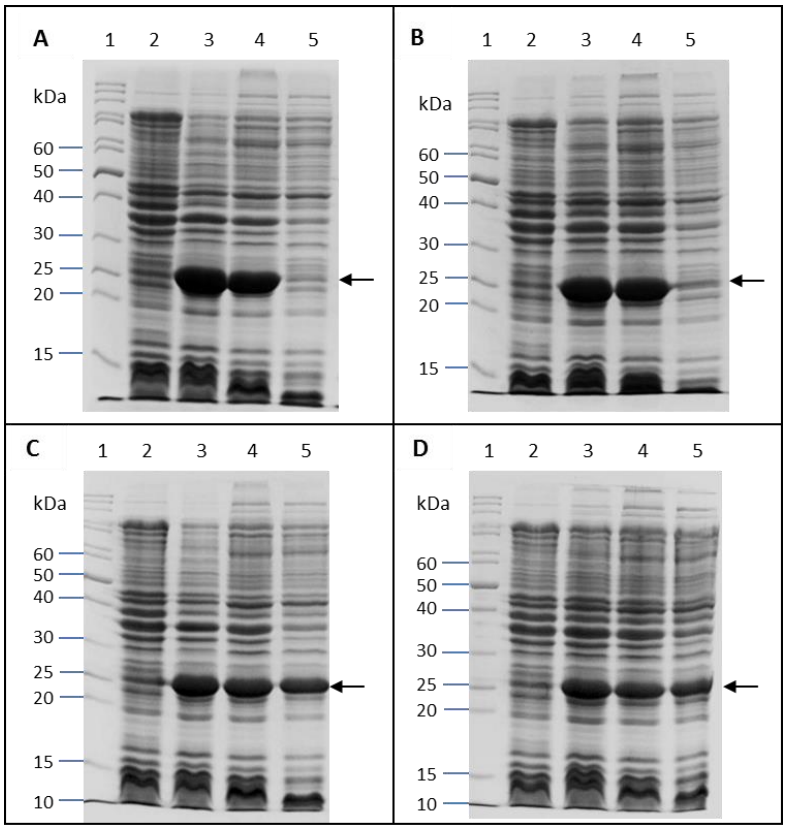


Figure 6.9 Overproduction and solubility of TssA$^{EI}_{MD\text{-}CTD}$V2. Coomassie blue-stained 12% SDS-PA gels showing the induction of pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V2 in *E. coli* strain BL21(λDE3) cells under different conditions: (**A**) 1 mM IPTG at 37^{o}C; (**B**) 0.1 mM IPTG at 37^{o}C; (**C**) 1 mM IPTG at 30^{o}C; (**D**) 0.1 mM IPTG at 30 ^{o}C. The arrows indicate the expected location of TssA$^{EI}_{MD\text{-}CTD}$V2 based on its size (~28.7 kDa). **A**-**D**: Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

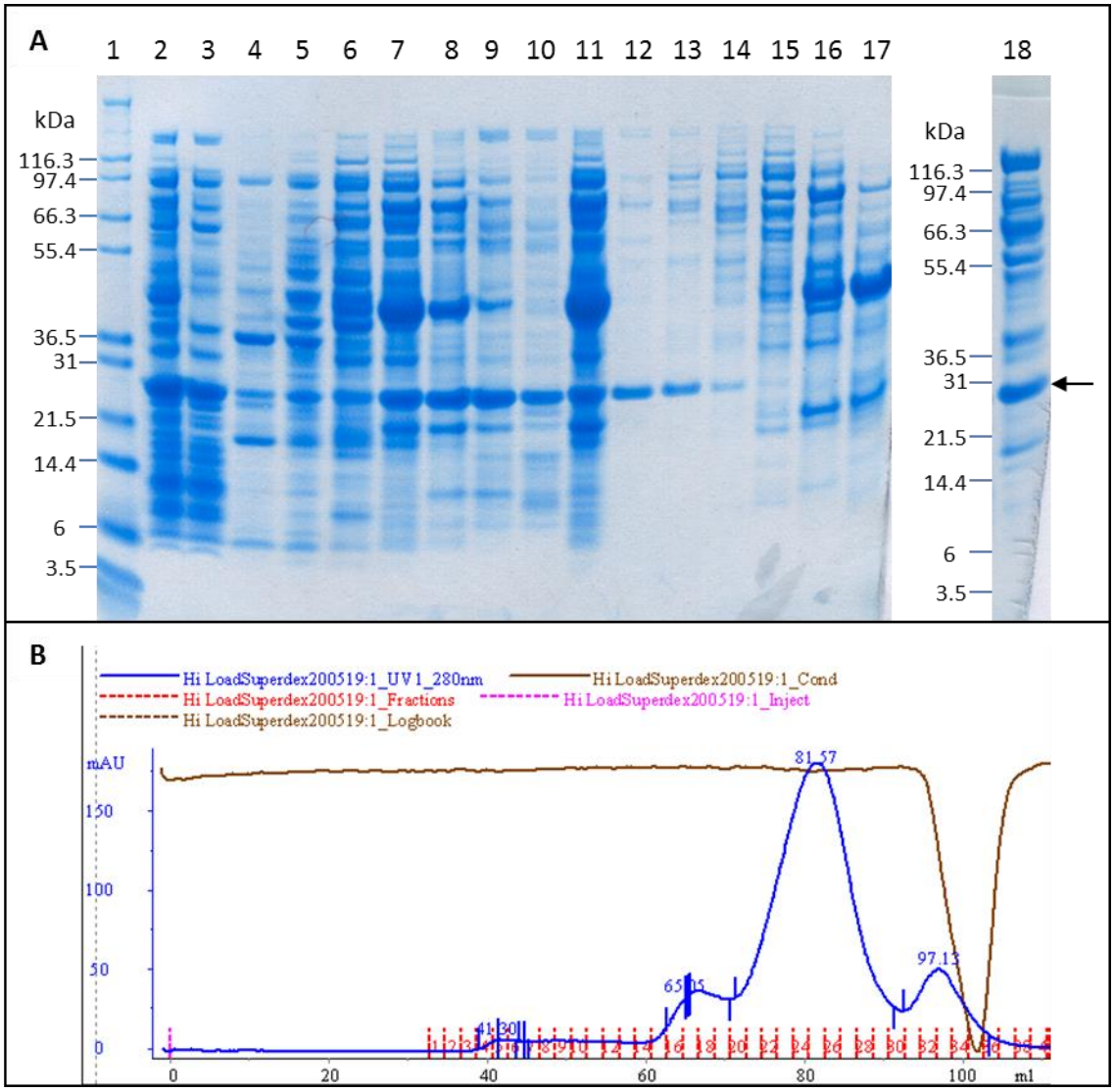


Figure 6.10 Purification of TssA$^{EI}_{MD\text{-}CTD}$V2. TssA$^{EI}_{MD\text{-}CTD}$V2 was overproduced in *E. coli* strain BL21(λDE3) with 1 mM IPTG induction at 30°C. The protein was then purified by anion exchange chromatography on a HiTrap DEAE column in buffer containing 50 mM Tris-HCl (pH 8.0). Elution was performed in the same buffer containing an increasing gradient concentration of NaCl up to 500 mM. Fractions containing TssA$^{EI}_{MD\text{-}CTD}$V2 were combined, concentrated and subjected to SEC on a Superdex 200 HiLoad 16/600 column in buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. **A**. 4-12% gradient SDS-PA gel (Novex) showing the purification of TssA$^{EI}_{MD\text{-}CTD}$V2. The arrow indicates the expected location of TssA$^{EI}_{MD\text{-}CTD}$V2 based on its MW (~28.7 kDa). Lane 1, protein marker (Novex); lane 2, soluble fraction of cell lysate following induction; lane 3, unbound material from DEAE; lanes 4-10, DEAE elution fractions (No.11, 13, 15, 17, 19, 21 and 23); lane 11, SEC load; lanes 12-17, SEC elution fractions (No.17, 19, 21, 25, 27 and 29); lane 18, concentrated peak elution fractions of SEC. **B**. Elution profile of TssA$^{EI}_{MD\text{-}CTD}$V2 during SEC monitored by UV at 280 nm.

6.3.2.3 TssA$^{EI}_{MD\text{-}CTD}$V3

As the purification of TssA$^{EI}_{MD\text{-}CTD}$V2 was hindered by its low solubility, a shorter version of native TssA$^{E}_{MD\text{-}CTD}$ was constructed for crystallisation trials, called TssA$^{EI}_{MD\text{-}CTD}$V3, which is 57 amino acids shorter than TssA$^{EI}_{MD\text{-}CTD}$V2 at its N-terminus (Figure 6.39). The trimmed region comprises two helices H9 and H10 (residues D231-R248 and A253-T268) predicted by Psipred (Figure 1.12). As they are predicted to be connected to the following helix, H11 (residues A286-M296), by a 17 amino acid unconstructed region, they were removed. Therefore, TssA$^{EI}_{MD\text{-}CTD}$V3 starts from H11. DNA encoding residues M284-K478 of AHA1844 was amplified using primers AHA1844.CTD.NdeI.for3 and AHA1844.BglII.rev. The amplified *tssA*$^{EI}_{MD\text{-}CTD}$V3 DNA fragment of the expected size (~588 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Nde*I and *Bgl*II that in the vector MCS and which were also incorporated into the PCR product in the forward and reverse primers. Ligation, screening for correct recombinant plasmids referred to as pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V3, and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1.

E. coli strain BL21(λDE3) cells containing pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V3 were induced under different conditions involving a range of IPTG concentration and different temperatures. However, there was no visible amount of TssA$^{EI}_{MD\text{-}CTD}$V3 with expected size of ~22 kDa was observed upon SDS-PAGE analysis (Figure 6.11 A). Other *E. coli* strains, i.e. C41(λDE3) and C43(λDE3) were also employed for protein overexpression. However, these alternative host strains did not result in overproduction of TssA$^{EI}_{MD\text{-}CTD}$V3 (Figure 6.11 B and C).

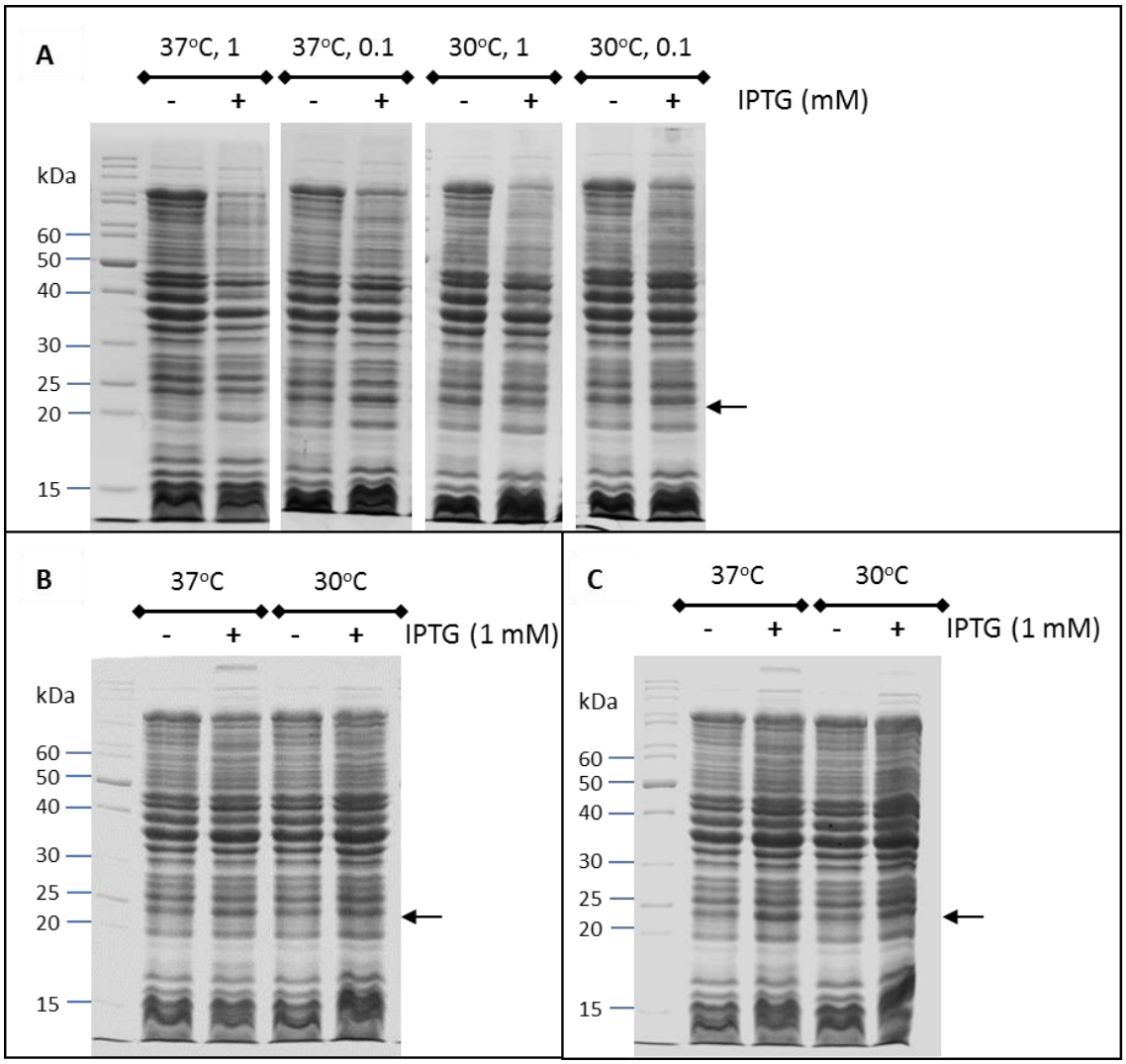
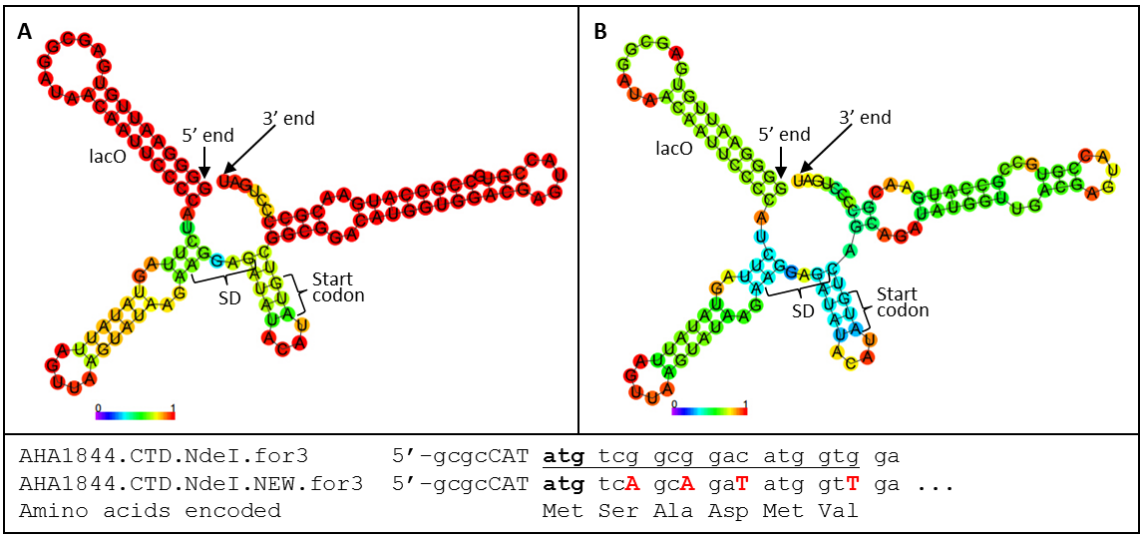


Figure 6.11 Overexpression of TssA$^{EI}_{MD\text{-}CTD}$V3. Coomassie blue-stained 12% SDS-PA gels showing the induction of pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V3 in *E. coli* strains BL21(λDE3)/C41(λDE3)/C43(λDE3) under different conditions. Arrows indicate the expected size of TssA$^{EI}_{MD\text{-}CTD}$V3 based on its predicted MW (~22 kDa). **A**. *E. coli* strain BL21(λDE3) cells. Induction temperatures (37/30°C) and concentrations of IPTG (1/0.1 mM) are labelled above the gels. **B**. *E. coli* strain C41(λDE3) cells. Induction temperatures (37/30°C) and concentrations of IPTG (1 mM) are labelled above the gel. **C**. *E. coli* strain C43(λDE3) cells. Induction temperatures (37/30°C) and concentrations of IPTG (1 mM) are labelled above the gel.

6.3.2.4 TssA$^{EI}_{MD\text{-}CTD}$V4

As it was not possible to overproduce TssA$^{EI}_{MD\text{-}CTD}$V3, the RNA folding status of the vector coded untranslated region (UTR) and the first several codons encoding the N-terminal amino acid of the TssA$^{EI}_{MD\text{-}CTD}$V3 sequence was analysed on the RNAfold WebServer. The results showed that there is a very high probability that Shine-Dalgarno sequence and translation initiation codon of TssA$^{EI}_{MD\text{-}CTD}$V3 are sequestered in an RNA hairpin structure (Figure 6.12 A), which may block translation. Therefore, a new forward primer, AHA1844.CTD.NdeI.NEW.for3, was designed in an attempt to break the hairpin structure by changing some early *tssA*$^{EI}_{MD\text{-}CTD}$V3 codons in the primer AHA1844.CTD.NdeI.for3 which was used for PCR amplification without changing the amino acids they encode. The new primer was checked for RNA folding status before used for cloning, and the result indicated that the possibility of forming a hairpin structure was reduced (Figure 6.12 B).

The modified forward primer was used for plasmid construction of pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V4 instead of AHA1844.CTD.NdeI.for3. Although there is no changes in the amino acids it encodes, a new name, TssA$^{EI}_{MD\text{-}CTD}$V4, was given to distinguish it from TssA$^{E}_{MD\text{-}CTD}$V3. PCR amplification and cloning for pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V4 construction was analogous to that of pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V3 (Section 6.3.3.3). TssA$^{EI}_{MD\text{-}CTD}$V4 was overexpressed from *E. coli* strain BL21(λDE3) cells at both 30^{o}C and 37^{o}C with 1 mM IPTG induction as concluded by the presence of a large amount of polypeptide with the expected size of TssA$^{EI}_{MD\text{-}CTD}$V4 (~22 kDa). However, there was no visible amount of soluble TssA$^{EI}_{MD\text{-}CTD}$V4 following cell lysis identified on SDS-PA gels (Figure 6.13 A). Attempts were made to solubilise it by reducing the concentration of IPTG used for induction (0.5 mM and 0.1 mM), however, the protein remained in the insoluble fraction. Different *E. coli* strains, C41(λDE3) and C43(λDE3), were tested in an attempt to improve the expression and solubility of TssA$^{EI}_{MD\text{-}CTD}$V4. The protein was overproduced in small amount in C41(λDE3) cells, but remained insoluble (Figure 6.13 B). There was no expression of TssA$^{EI}_{MD\text{-}CTD}$V4 from C43(λDE3) cells (Figure 6.13 C).



AHA1844.CTD.NdeI.for3	5′-gcgcCAT	**atg** tcg gcg gac atg gtg ga
AHA1844.CTD.NdeI.NEW.for3	5′-gcgcCAT	**atg** tcA gcA gaT atg gtT ga ...
Amino acids encoded		Met Ser Ala Asp Met Val

Figure 6.12 Comparison of RNA folding status of pACYCDuet-derived UTR and N-terminal 17 amino acids of sequences encoding TssA$^{EI}_{MD\text{-}CTD}$V3 and TssA$^{EI}_{MD\text{-}CTD}$V4. RNA folding status of the vector untranslated region beginning at the start of the T7 promoter-directed transcript (5' end) and first 17 codons specifying the N-terminus TssA$^{EI}_{MD\text{-}CTD}$V3/V4 (M284-D300 of native TssA^{E}) were analysed on the RNAfold WebServer. **A**, RNA folding status of TssA$^{EI}_{MD\text{-}CTD}$V3 showing first 17 codons of *tssA*$^{EI}_{MD\text{-}CTD}$; **B**, RNA folding status of TssA$^{EI}_{MD\text{-}CTD}$V4 showing first 17 codons of *tssA*$^{E}_{MD\text{-}CTD}$. The possibility of forming a hairpin structure is indicated at the bottom of the figures, i.e. purple=0, red=1. The sequences at the bottom are the two forward primers that were used for PCR amplification and cloning of *tssA*$^{EI}_{MD\text{-}CTD}$V3/V4. The complete sequence of AHA1844.CTD.NdeI.for3 primer is shown and the corresponding region of AHA1844.CTD.NdeI.NEW.for3 primer is shown with the annealing region omitted. The codons labelled in bold red font in the AHA1844.CTD.NdeI.NEW.for3 primer are those bases that were changed for breaking the hairpin structure.

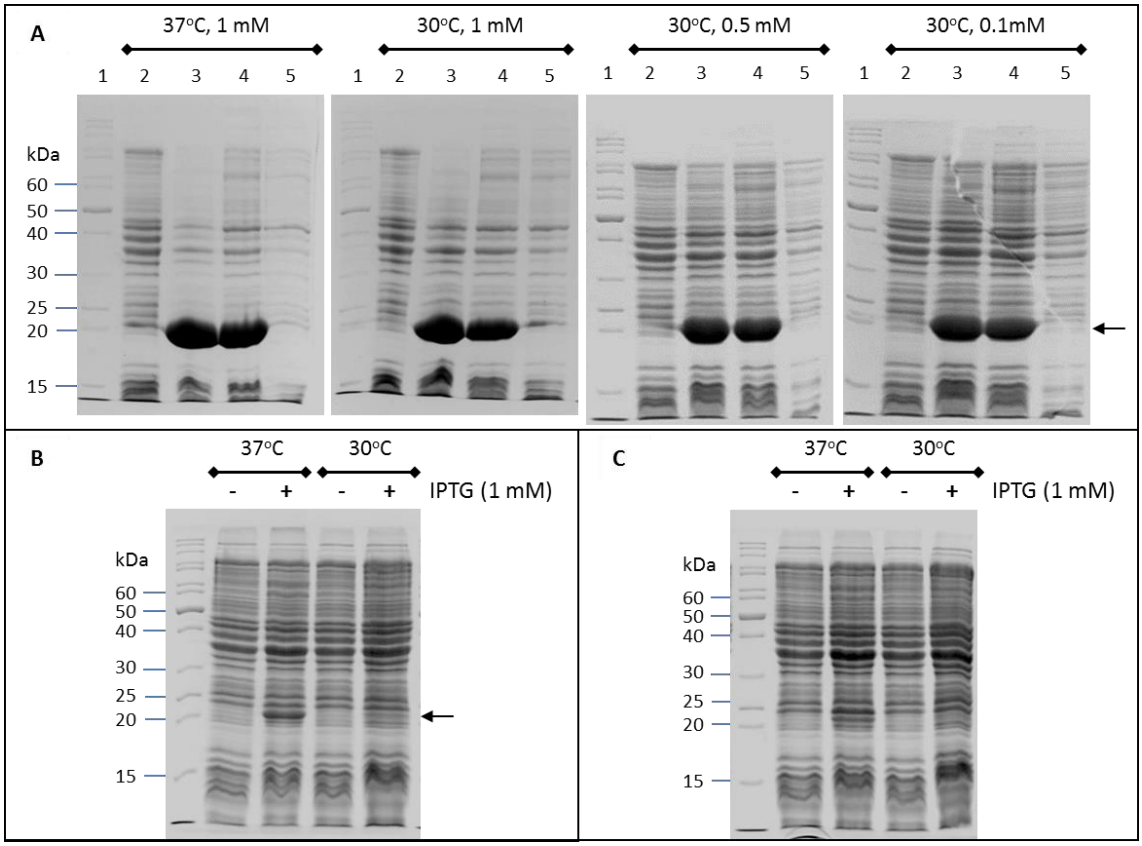


Figure 6.13 Overproduction and solubility of TssA$^{EI}_{MD\text{-}CTD}$V4. Coomassie blue-stained 12% SDS-PA gels showing the induction of pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V4 in *E. coli* strains BL21(λDE3)/C41(λDE3)/C43(λDE3) under different conditions. Induction temperatures and concentrations of IPTG are as labelled. The arrows indicate the expected locations TssA$^{EI}_{MD\text{-}CTD}$V4 based on its size (~22 kDa). **A**. Induction of pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V4 in *E. coli* strain BL21(λDE3). Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. Induction of pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V4 in *E. coli* strain C41(λDE3). **C**. Induction of pACYCDuet-*tssA*$^{EI}_{MD\text{-}CTD}$V4 in *E. coli* strain C43(λDE3). B and C, '-', total cell protein from uninduced cells; '+', total cell protein from cells following induction.

6.3.2.5 TssA$^{EI}_{MD\text{-}CTD}$V5

As it was not possible to obtain soluble TssA$^{EI}_{MD\text{-}CTD}$V4, an attempt was made to obtain the soluble protein by cleaving the MD-CTD region from native TssA^{E}, which is soluble. This can be achieved by introducing a factor Xa cleavage site (IEGR) into the flexible linker region of TssA^{EI} between P280 and L281, which can then be cleaved by factor Xa protease resulting in a TssA$^{EI}_{MD\text{-}CTD}$V5. TssA$^{EI}_{MD\text{-}CTD}$V5 has an extra three bases of TssA^{EI} (L281, A282 and P283) at the N-terminus compared to TssA$^{EI}_{MD\text{-}CTD}$V4. To do this, SOE-PCR was carried out by amplification of two fragments of *tssA*^{EI} coding sequences on either side of the insertion site (P280) where the factor Xa site was going to be introduced. Based on the insertion site, the N-terminal coding region of *tssA*^{EI} was regarded as fragment A; the C-terminal coding region of *tssA*^{EI} (encoding TssA$^{EI}_{MD\text{-}CTD}$V5) was regarded as fragment B. The PCR amplification to generate TssA^{EI} containing a factor Xa site (named $_{linkerXa}$.TssA^{EI}) comprised two sequential PCR steps. The PCR for amplifying fragment A and fragment B involved 2 pairs of primers, pACYC-AHA1844.for and AHA1844linkerXarev; AHA1844linkerXafor and pACYC-AHA1844.rev, respectively. Primers AHA1844linkerXafor and AHA1844linkerXarev include 25 bases that anneals to *tssA*^{EI} and 12 complementary bases coding for the factor Xa cleavage site. The resulting PCR products were electrophoresed in a 1% agarose gel and purified. In order to fuse the two PCR fragments (fragment A ~860 bp and fragment B ~620 bp), another round of PCR was performed using pACYC-AHA1844.for and pACYC-AHA1844.rev primers with the two fragments as template. The product of the second round PCR, linkerXa.*tssA*^{EI} (~1.45 kb), was digested with *Bgl*II and *Hin*dIII which recognized restriction sites incorporated by the pACYC-AHA1844.for and pACYC-AHA1844.rev primers, and ligated into pACYCDuet that was digested with *Bam*HI and *Hin*dIII. This resulted in insertion of the linkerXa.*tssA*^{EI} coding sequence in the same open reading frame as the pACYCDuet histidine tag codons thereby adding a hexa-histidine tag at the N-terminus of $_{linkerXa}$.TssA^{EI}. Ligation, screening for correct recombinant plasmids and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1. The plasmid was called pACYCDuet-His$_6$.linkerXa.*tssA*^{EI}.

$_{His6.linkerXa}$.TssA^{EI} was overproduced in a huge amount in *E. coli* strain BL21(λDE3) containing pACYCDuet-His$_6$.linkerXa.*tssA*^{EI} following induction with 1 mM IPTG at

37^{o}C. However, the protein was almost insoluble (~5% solubility) upon SDS-PAGE analysis (Figure 6.14 A). Inductions with lower temperature (30^{o}C) with lower concentration of IPTG (0.5 mM and 0.1 mM) were carried out in an attempt to increase the fraction of soluble $_{His6.linkerXa}$.TssA^{EI}. However, under these conditions, although $_{His6.linkerXa}$.TssA^{EI} was expressed very well, there was no increase in its solubility. *E. coli* strains C41(λDE3) and C43(λDE3) containing pACYCDuet-His$_6$.linkerXa.*tssA*^{EI} were also used for expressing soluble $_{His6.linkerXa}$.TssA^{EI}. However, there was no protein overproduced from these two strains (Figure 6.14 B and C).

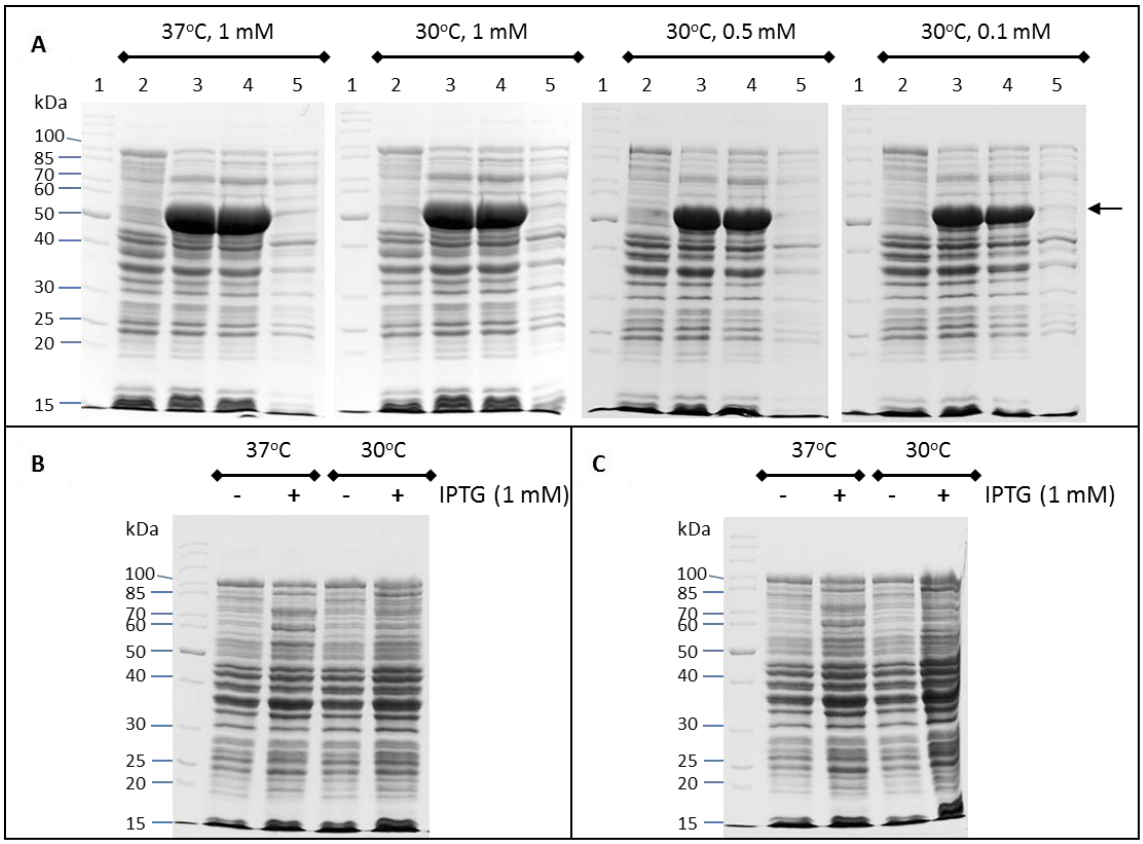


Figure 6.14 Overproduction and solubility of $_{His6.linkerXa}$.TssA^{EI}. Coomassie blue-stained 10% SDS-PA gels showing the induction of pACYCDuet-His$_6$.linkerXa.*tssA*^{EI} in *E. coli* strains BL21(λDE3)/C41(λDE3)/C43(λDE3) under different conditions. Induction temperatures and concentrations of IPTG are labelled. The arrows indicate expected locations of $_{His6.linkerXa}$.TssA^{EI} based on its size (~52 kDa). **A**. Induction of pACYCDuet-His$_6$.linkerXa.*tssA*^{EI} in *E. coli* strain BL21(λDE3). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. Induction of pACYCDuet-His$_6$.linkerXa.*tssA*^{EI} in *E. coli* strain C41(λDE3). **C**. Induction of pACYCDuet-His$_6$.linkerXa.*tssA*^{E} in *E. coli* strain C43(λDE3). B and C, '-', total cell protein from uninduced cells; '+', total cell protein from cells following induction.

6.3.2.6 Transmission electron microscopy of $TssA^{EI}_{MD\text{-}CTD}$

Two of the five plasmids designed to overproduce a $TssA^{EI}_{MD\text{-}CTD}$ derivative produced large amounts of soluble protein. The two proteins were $_{His6}.TssA^{EI}_{MD\text{-}CTD}$ $TssA^{EI}_{MD\text{-}CTD}V2$ (Sections 6.3.3.1 and 6.3.3.2). Both proteins were observed by SEC to be able to oligomerise into a complex. However, $_{His6}.TssA^{E}_{MD\text{-}CTD}$ was more soluble than $TssA^{EI}_{MD\text{-}CTD}V2$. Therefore, $_{His6}.TssA^{EI}_{MD\text{-}CTD}$ was analysed by TEM to obtain additional structural information. The results from the TEM showed that $_{His6}.TssA^{EI}_{MD\text{-}CTD}$ was able to form rings but they were not circular, more like a five-pointed star (Figure 6.15). In order to confirm the 'five-pointed' star shape averaging of $TssA^{EI}_{MD\text{-}CTD}$ single particles was performed.

Averaging was carried out based on 633 single particles. Alignment of single particles was based on cross correlation. Direct alignment was carried out including translational and rotational modes after filtering and normalisation of the images. The five pointed star shape of the protein complex became clear after six iterations of summing and aligning against the total sum (Figure 6.16 A). After 12 steps of summing and aligning against the total sum, those single particles that would not align properly with the reference were removed and 494 particles remained for further analysis. The images of the particles that would not align properly were determined by analysis of the alignment procedure, and were selected as the particles that would not have a best alignment position during 20 iterations within a single alignment step. In total, 20 steps of summing and aligning against the total sum were performed.

Test rotational symmetry of $_{His6}.TssA^{EI}_{MD\text{-}CTD}$ was conducted for confirming the specific symmetries that $_{His6}.TssA^{EI}_{MD\text{-}CTD}$ may have, although a five-pointed star shape has become clear. The rotational symmetry test was based on correlation between the rotated total average with its unrotated self. Two- to eight-fold symmetry were tested, and the results showed that five-fold symmetry gave the highest score of correlation coefficients (90.8%), which is significantly higher than other folds of symmetries (Figure 6.16 B). The rotational symmetry test provided further evidence of the five-pointed star shape of $_{His6}.TssA^{EI}_{MD\text{-}CTD}$. Based on the averaging results, five-fold symmetry was imposed on the averages after the 20 steps of summing and aligning against the total sum (Figure 6.16 C). The size estimation was conducted on the total

average, and the results showed the $TssA^{EI}_{MD\text{-}CTD}$ ring has an exterior of ~160 Å and interior of ~36 Å.

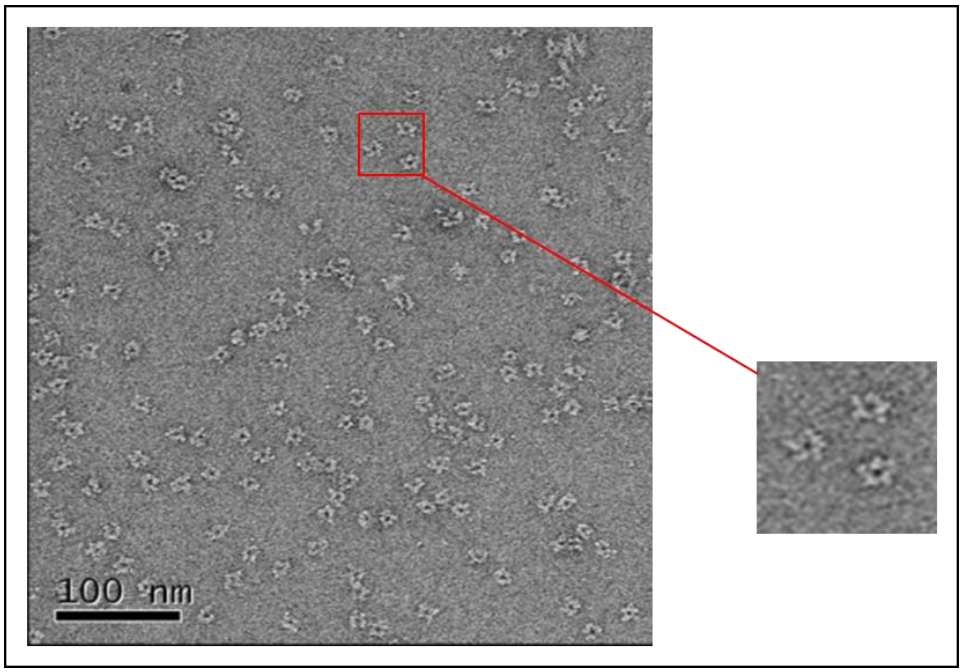


Figure 6.15 Negative stain transmission electron micrograph of $_{His6}$.TssA$^{EI}_{MD-CTD}$. $_{His6}$.TssA$^{EI}_{MD-CTD}$ (0.025 mg/ml) in buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole was loaded onto freshly glow-discharged grids and stained with 0.75% uranyl formate stain before being analysed by TEM (CM200 FEG, Philips).

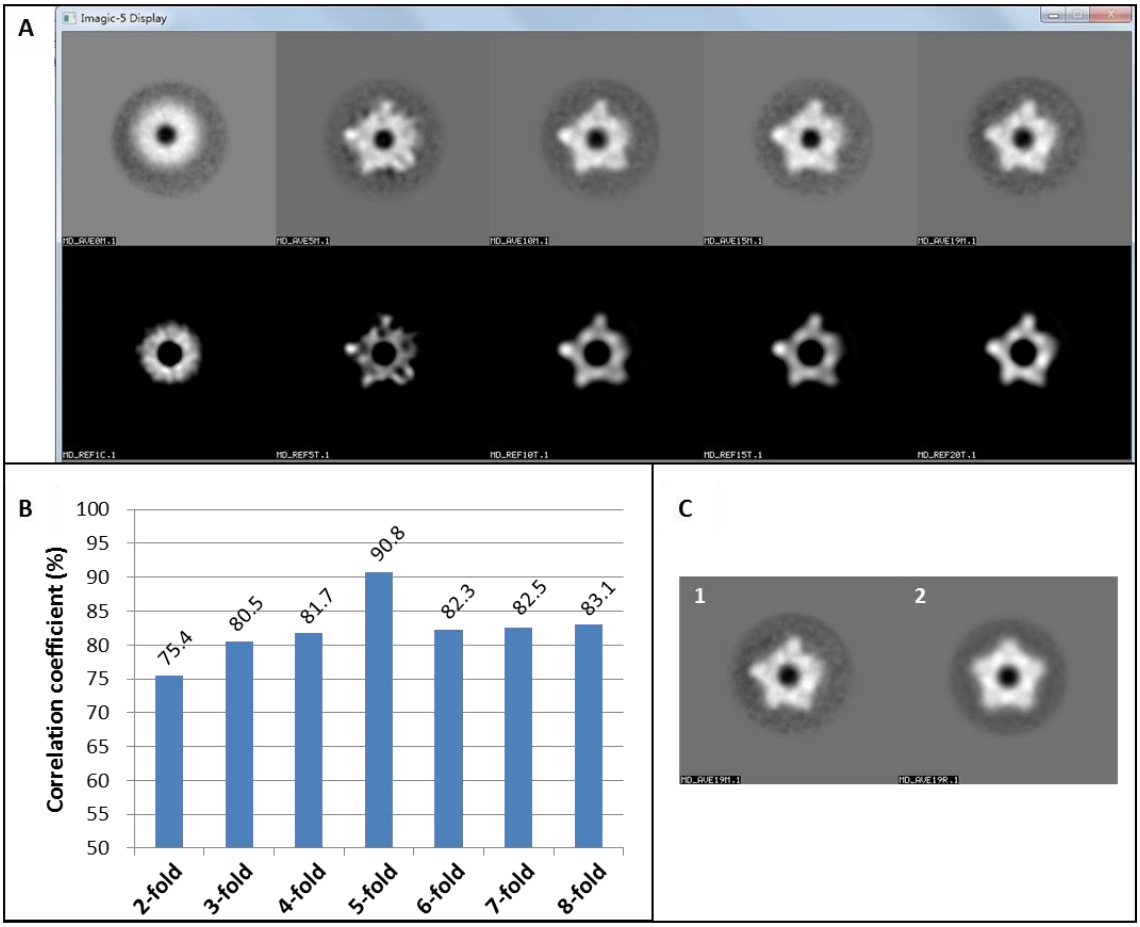


Figure 6.16 Averaging of $_{His6}$.TssA$^{EI}_{MD-CTD}$ single particles. 633 single particles of $_{His6}$.TssA$^{EI}_{MD-CTD}$ were collected following negative stain transmission electron microscopy. Averaging was carried out using Imagic-5 software. After 12 steps of summing and aligning against the total sum, those single particles that would not align properly with the reference were removed (see text) and 494 particles were retained for further analysis. Total of 20 steps of direct alignment were carried out. **A**. Total sum of $_{His6}$.TssA$^{EI}_{MD-CTD}$ averaging. Upper panel, total sum of $_{His6}$.TssA$^{EI}_{MD-CTD}$ after certain iterations of direct alignment, from left to right: 1, 6, 11, 16, 20 iterations; lower panel, filtered total sum corresponding to the images above them. **B**. Test rotational symmetry of $_{His6}$.TssA$^{EI}_{MD-CTD}$. The X-axis shows the specific symmetry that was tested, and Y-axis shows the correlation coefficient (%) of all the single particles that rotate the single particle for a specific angle according to the fold of symmetry, then correlates the rotated particle with its unrotated self. **C**. Five-fold symmetry applied on the final average after 20 steps of direct alignment. 1, before imposing symmetry; 2, after imposing symmetry.

6.3.2.7 Molecular weight estimation of $_{His6}$.TssA$^{EI}_{MD-CTD}$ by SEC-MALS

As $_{His6}$.TssA$^{EI}_{MD-CTD}$ was observed to have a star shape with a central hole structure by TEM, the size estimation by SEC would not be accurate, as it includes the empty space of the hole in the middle. Therefore, SEC-MALS analysis was performed with purified $_{His6}$.TssA$^{EI}_{MD-CTD}$ for a more accurate MW estimation.

The elution profile of $_{His6}$.TssA$^{EI}_{MD-CTD}$ from the Superdex 200 10/300 GL column showed there was some aggregate material at ~15.5 minute but the low response in the RI and UV indicated that this is negligible (Figure 6.17 A). There was a minor peak at ~18.5 minute containing only a small amount of material (<1 µg) and a major peak at 21.5 minute containing ~129 µg of protein. There was no indication of any lower MW material. The peak analysis of $_{His6}$.TssA$^{EI}_{MD-CTD}$ was based on the elution profile using Astra software which defined the peak regions into three: Peak 1 (21.02-21.94 minute), Peak 2 (19.87-23.63 minute) and Peak 3 (17.69-19.63 minute). The defined Peak 2 comprised Peak 1, which had a broader range in the elution time.

The MW was determined by a Zimm fit procedure at each point then averages calculated over the peak region (Figure 6.17 B). The MW estimate for the major peak was 294 kDa. For the minor peak at 15.5 minute the amount was ~0.7 µg. The amount under the aggregate peak at ~18.5 minute was even smaller. As the monomer MW of $_{His6}$.TssA$^{EI}_{MD-CTD}$ is 29.54 kDa, this suggests TssA$^{EI}_{MD-CTD}$ is able to form a decamer. The MW estimation of TssA$^{EI}_{MD-CTD}$ is more precise than that of the full-length TssA^{EI}, where the MW is likely to be overestimated due to influence of the aggregation material, and therefore the oligomerisation estimation for MD-CTD is likely to be more accurate than that of TssA^{EI}.

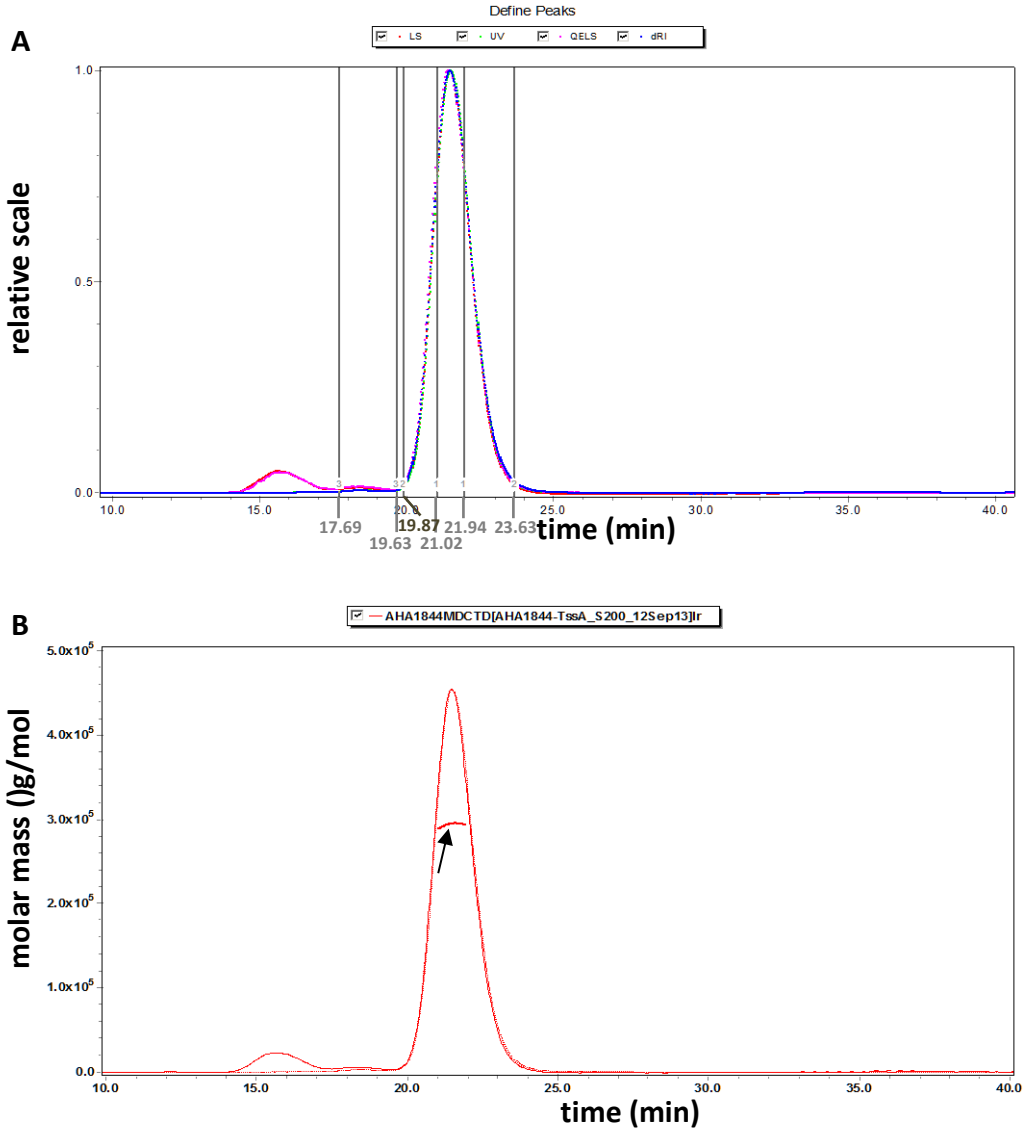


Figure 6.17 SEC-MALS analysis of $_{His6}$.TssA$^{EI}_{MD-CTD}$. A. $_{His6}$.TssA$^{EI}_{MD-CTD}$ elution profile and peak positions. The normalised signals for light scattering (LS) are shown in red, refractive index (RI) in blue, UV absorption (280 nm) in green and quasi-elastic light scattering (QELS) in magenta. The peak regions chosen for analysis are indicated by numbered vertical grey lines that defined by Astra software. The X-axis shows elution time in minutes and Y-axis shows relative scale. **B**. $_{His6}$.TssA$^{EI}_{MD-CTD}$ molar mass plot. The MW is determined by a Zimm fit procedure at each point. Solid line represents LS trace. Dashed and dotted traces represent RI and UV responses, respectively. Molecular weight estimate curve is shown in heavier dots (pointed with black arrows). The X-axis is the elution time in minutes and the Y-axis is the MW scale. Figures provided by University of York Bioscience Technology Facility.

6.3.2.8 X-ray structure determination of $TssA^{EI}_{MD\text{-}CTD}$

Although two versions of soluble $TssA^{EI}_{MD\text{-}CTD}$ were obtained, only $_{His6}$.$TssA^{EI}_{MD\text{-}CTD}$ was used for crystallisation trials due to aggregation of $TssA^{EI}_{MD\text{-}CTD}$V2. A crystal of $_{His6}$.$TssA^{EI}_{MD\text{-}CTD}$ was obtained in buffer containing 0.2 M sodium acetate, 0.1 M sodium citrate and 10 % (w/v) PEG 4000 (pH 5.5). The structure was solved at 3.25 Å resolution by molecular replacement using the $TssA^{E}_{MD}$ structure model. The detail of the $TssA^{EI}_{MD}$ and $TssA^{EI}_{MD\text{-}CTD}$ structures will be presented and discussed in Sections 6.3.3.5 and 6.3.4.7, respectively.

6.3.3 Overproduction and purification of $TssA^{EI}_{MD}$ derivatives

As the results suggest the $TssA^{EI}_{MD\text{-}CTD}$ is responsible for the oligomerisation of $TssA^{EI}$, further analysis was carried out to identify the region/domain of $TssA^{EI}$ that is responsible for oligomerisation. Furthermore, it was speculate that it may be possible to obtain structural data for the predicted individual domains. $TssA^{EI}_{MD}$ which was predicted by bioinformatics analysis and further verified by limited proteolysis was analysed. A histidine-tagged $TssA^{EI}_{MD}$ and the other three native $TssA^{EI}_{MD}$ derivatives containing different lengths of amino acids were constructed.

6.3.3.1 $_{His6}$.$TssA^{EI}_{MD}$

To construct a hexa-histidine-tagged $TssA^{EI}_{MD}$ for facilitating protein purification, DNA encoding AHA1844 residues G223-A387 was amplified using primers pACYC-AHA1844.CTDfor and pACYC-AHA1844.CTDSD1rev. The amplified $tssA^{EI}_{MD}$ DNA fragment of the expected size (~540 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Bam*HI and *Hin*dIII that recognized sites incorporated in the PCR product by the forward and reverse primers and the plasmid MCS. This in a hexa-histidine tag encoded by the vector located at the N-terminus of $TssA^{EI}_{MD}$. Ligation, screening for correct recombinant plasmids and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1. The plasmid was called pACYCDuet-His_6.$tssA^{EI}_{MD}$.

$_{His6}$.TssA$^{EI}_{MD}$ was overproduced very well in *E. coli* strain BL21(λDE3) following induction at 37 ^{o}C with 1 mM IPTG as evidenced by the large amount of a polypeptide that migrated at the expected size of ~19.4 kDa. However, only small amount of protein was present in the soluble fraction (5-10%). Therefore, induction with lower temperature (30^{o}C) and lower concentration of IPTG were tested for solubilising the protein. The results showed that $_{His6}$.TssA$^{E}_{MD}$ was overexpressed well at 30^{o}C with a range of IPTG concentration 0.01-1 mM, and the protein remained soluble (Figure 6.18). It was not overproduced with 0.01 mM IPTG induction at 30^{o}C.

The overproduced soluble $_{His6}$.TssA$^{EI}_{MD}$ from the 0.1 mM IPTG induction at 30^{o}C was purified by IMAC on a HisTrap HP column. $_{His6}$.TssA$^{EI}_{MD}$ bound to the nickel resin very well, and was eluted by a gradient concentration of imidazole in the range 117-233 mM imidazole (Figure 6.19 A). The fractions containing $_{His6}$.TssA$^{EI}_{MD}$ were combined and concentrated (MWCO 10 kDa, Vivaspin centrifugal concentrator) for further purification and oligomerisation estimation by SEC on a Superose 12 column. $_{His6}$.TssA$^{EI}_{MD}$ was eluted from the column at 25.61 minute corresponding to an apparent MW of ~48 kDa (Figure 6.19 B and C). Given the the monomeric MW is 19.4 kDa, $_{His6}$.TssA$^{EI}_{MD}$ is estimated to occur as a dimer. The purified $_{His6}$.TssA$^{EI}_{MD}$ was subjected to crystallisation trials upon buffer exchanged into 5 mM Tris-HCl (pH 8.0) containing 50 mM NaCl.

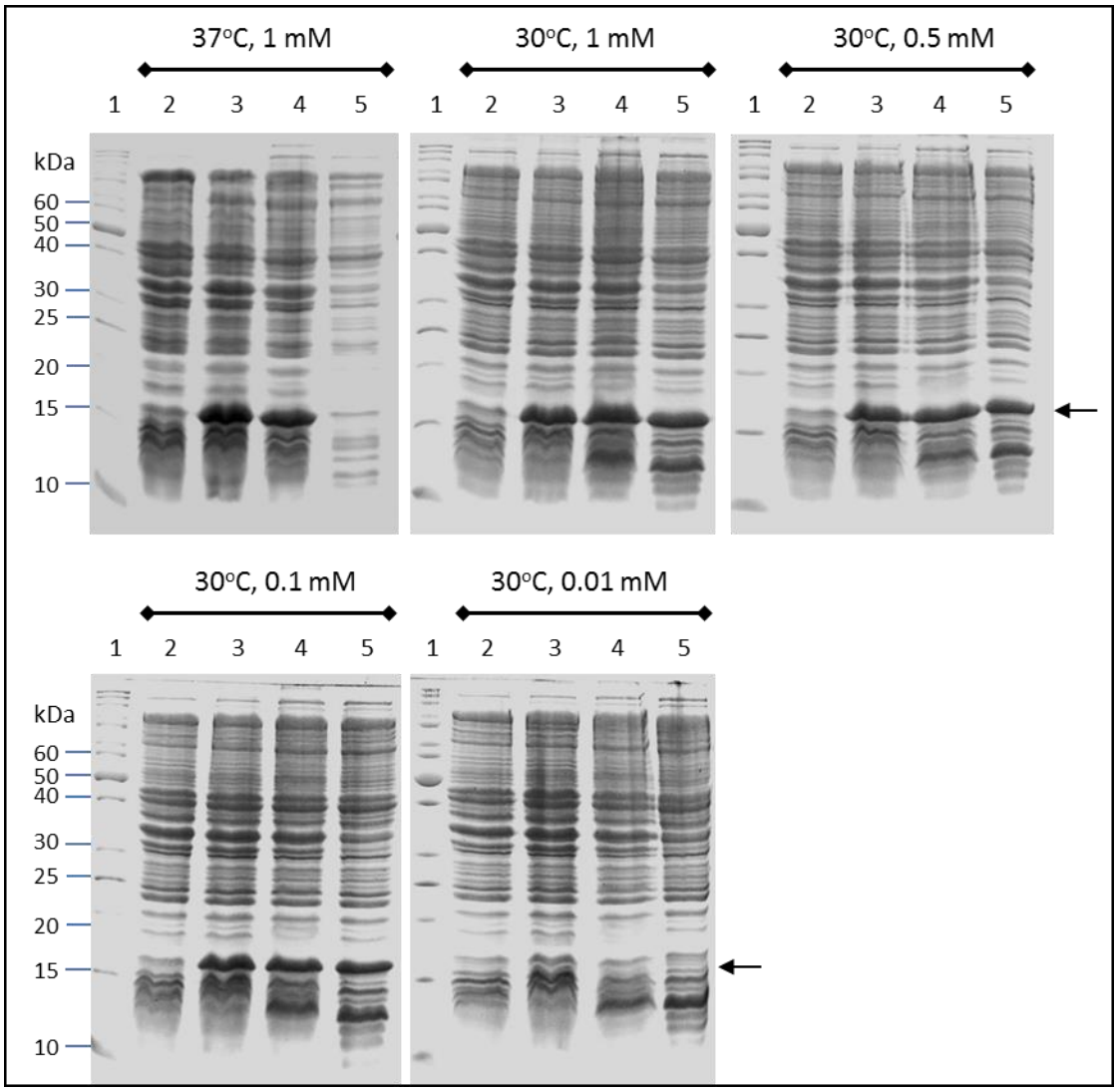


Figure 6.18 Overproduction and solubility of $_{His6}$.TssA$^{EI}_{MD}$. Coomassie blue-stained 12% SDS-PA gels showing the induction of pACYCDuet-His$_6$.*tssA*$^{EI}_{MD}$ in *E. coli* strain BL21(λDE3) under different conditions. Induction temperatures and concentrations of IPTG are as indicated. The arrows indicate the expected locations of $_{His6}$.TssA$^{EI}_{MD}$ based on its size (~19.4 kDa). Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

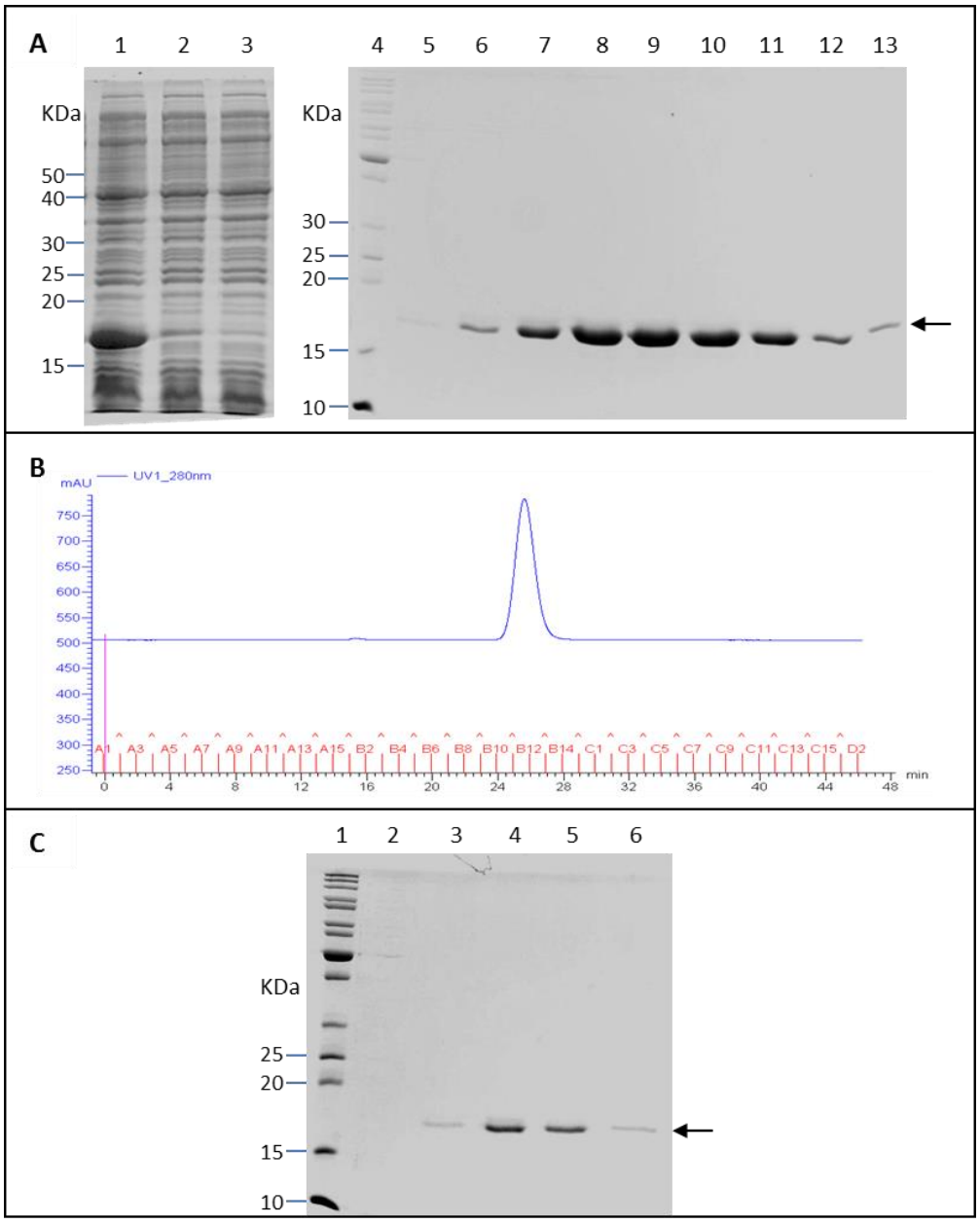


Figure 6.19 Purification of $_{His6}$.TssA$^{EI}_{MD}$. $_{His6}$.TssA$^{EI}_{MD}$ synthesis was induced from pACYCDuet-His$_6$.*tssA*$^{EI}_{MD}$ in *E. coli* strain BL21(λDE3) cells with 0.1 mM IPTG at 30°C. Soluble fraction of the cell lysate following induction was subjected to IMAC on a HisTrap HP column (1 ml) in the buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. Fractions containing $_{His6}$.TssA$^{EI}_{MD}$ were combined, concentrated and subjected to SEC on a Superose 12 column in a buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol. **A**. Purification of $_{His6}$.TssA$^{EI}_{MD}$ by IMAC analysed by electrophoresis in 12% SDS-PA gels. Lane 1, soluble fraction of cell lysate following induction; lane 2, flow-through_1 from HisTrap column; lane 3, flow-through_2 from HisTrap column; lane 4, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 5-13, elution fractions of HisTrap column corresponding to the peak of the UV trace at 280 nm. **B**. Elution profile of $_{His6}$.TssA$^{EI}_{MD}$ during SEC. **C**. 15% SDS-PAGE analysis of SEC of $_{His6}$.TssA$^{EI}_{MD}$. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 2-6, elution fractions of SEC corresponding to the peak of the UV trace at 280 nm. Arrows in A and C indicate the expected location of $_{His6}$.TssA$^{EI}_{MD}$ based on its size (~19.4 kDa).

6.3.3.2 TssA$^{EI}_{MD}$V2

It was decided to generate an untagged TssA$^{EI}_{MD}$ derivative for facilitating crystallisation trials, DNA encoding AHA1844 residues S228-Q375 was amplified using primers AHA1844.CTD.NdeI.for2 and AHA1844.CTDSD1.BglII.rev2, which also removed the 12 amino acids of predicted interdomain linker at the C-terminus of $_{His6}$.TssA$^{EI}_{MD}$. The amplified *tssA*$^{EI}_{MD}$ DNA fragment (named *tssA*$^{EI}_{MD}$V2) of the expected size (~444 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites incorporated in the PCR product by the forward and reverse primers and the plasmid MCS. Ligation, screening for correct recombinant plasmids and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1. The plasmid was called pACYCDuet-*tssA*$^{EI}_{MD}$V2.

E. coli strain BL21(λDE3) were used for overexpressing TssA$^{EI}_{MD}$V2. TssA$^{EI}_{MD}$V2 was overproduced in a reasonable amount with both 0.5 and 0.1 mM IPTG induction at 37^{o}C based on SDS-PAGE analysis (~16.7 kDa). However, the protein was largely insoluble (Figure 6.20). Attempts were made to solublise the protein by carrying out the induction at 30^{o}C with 0.5 mM IPTG. However, there was no TssA$^{EI}_{MD}$V2 overproduced at 30^{o}C.

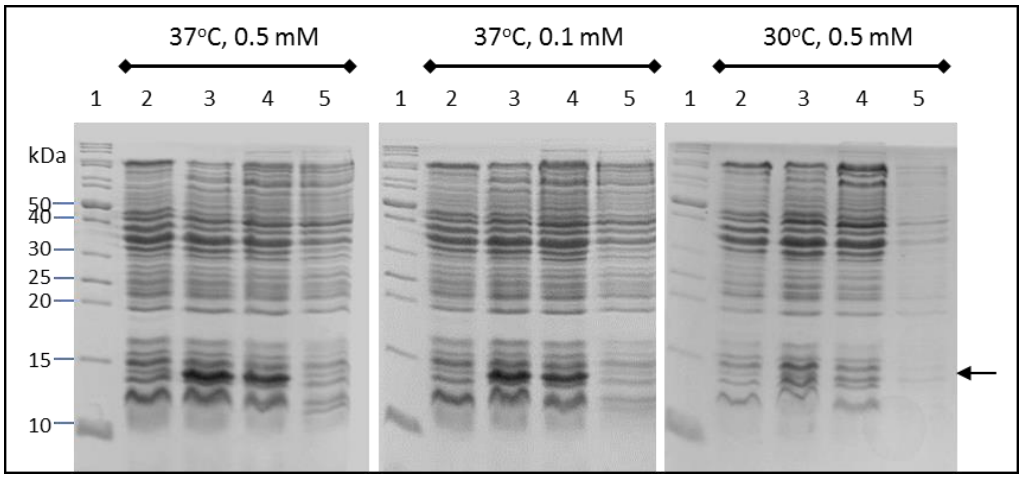


Figure 6.20 Overproduction and solubility of TssA$^{EI}_{MD}$V2. Coomassie blue-stained 15% SDS-PA gels showing the induction of pACYCDuet-*tssA*$^{EI}_{MD}$V2 in *E. coli* strain BL21(λDE3) under different conditions. Induction temperatures and concentrations of IPTG are as indicated. The arrows indicate the expected locations of TssA$^{EI}_{MD}$V2 based on its size (~16.7 kDa). Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction.

6.3.3.3 TssA$^{EI}_{MD}$V3

As it was not possible to overproduce soluble TssA$^{EI}_{MD}$V2, a shorter version of native TssA$^{EI}_{MD}$ was constructed for crystallisation trials, called TssA$^{EI}_{MD}$V3, which is 57 amino acids shorter than TssA$^{EI}_{MD}$V2 at its N-terminus (Figure 6.40). The trimmed region comprises two helices H9 and H10 (residues D231-R248 and A253-T268) predicted by Psipred (Figure 1.12). As they are predicted to be connected to the following helix, H11 (residues A286-M296), by a 17 amino acid unconstructed region, they were removed. Therefore, TssA$^{EI}_{MD}$V3 starts from H11 but has the same endpoints as TssA$^{EI}_{MD}$V2 (i.e. Q375).

DNA encoding residues M284-Q375 of AHA1844 was amplified using primers AHA1844.CTD.NdeI.for3 and AHA1844.CTDSD1.BglII.rev2. The amplified *tssA*$^{EI}_{MD}$V3 DNA fragment of the expected size (~276 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Nde*I and *Bgl*II that in the vector MCS and which were also incorporated into the PCR product in the forward and reverse primers. Ligation, screening for correct recombinant plasmids and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1. The plasmid was called pACYCDuet-*tssA*$^{EI}_{MD}$V3.

E. coli strain BL21(λDE3) was used for overexpressing TssA$^{EI}_{MD}$V3. However, no protein of the expected size of TssA$^{EI}_{MD}$V3 (~10.2 kDa) was observed to be overexpressed with 1 mM IPTG induction at either 30^{o}C or 37^{o}C based on SDS-PAGE analysis (Figure 6.21). The overexpression problem of TssA$^{EI}_{MD}$V3 may be caused by RNA folding sequestering the translation initiation region, which was the problem that impeded expression of TssA$^{EI}_{MD\text{-}CTD}$V3 (Section 6.3.2.3).

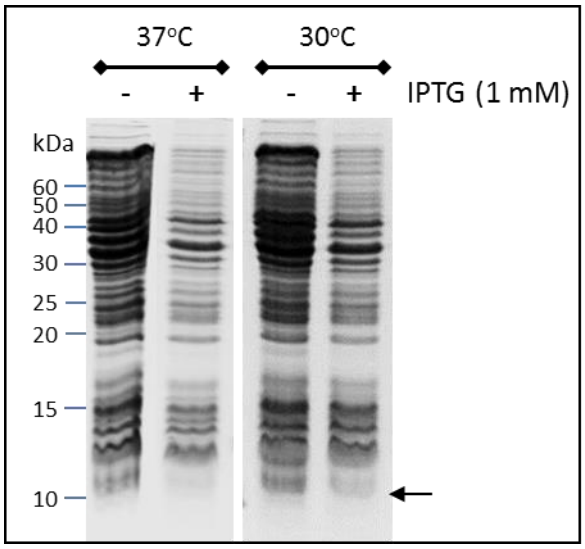


Figure 6.21 Overproduction of TssA$^{EI}_{MD}$V3. Coomassie blue-stained 15% SDS-PA gels showing the induction of pACYCDuet-*tssA*$^{EI}_{MD}$V3 in *E. coli* strain BL21(λDE3) under different conditions as indicated. Induction temperatures and concentrations of IPTG are as indicated. The arrow indicates the expected location of TssA$^{EI}_{MD}$V3 based on its size (~10.2 kDa). EZ-Run™ Rec protein ladder (Fisher) was used as protein reference. '-', total cell protein from uninduced cells; '+', total cell protein from cells following induction.

6.3.3.4 TssA$^{EI}_{MD}$V4

As it was not possible to overproduce TssA$^{EI}_{MD}$V3, it was considered to be due to an RNA hairpin forming at the beginning of the translated sequence. A new forward primer, AHA1844.CTD.NdeI.NEW.for3, was designed in an attempt to break the hairpin structure by changing some codons at the 5' end of *tssA*$^{EI}_{MD}$V3 in the primer AHA1844.CTD.NdeI.for3 which was used for PCR amplification without changing the amino acids they encode. The modified version of the MD was called TssA$^{EI}_{MD}$V4 although its amino acid sequence is unchanged relative to TssA$^{EI}_{MD}$V3.

DNA encoding residues M284-Q375 of AHA1844 was amplified using primers AHA1844.CTD.NdeI.NEW.for3 and AHA1844.CTDSD1.BglII.rev2. The amplified *tssA*$^{EI}_{MD}$V4 DNA fragment of the expected size (~276 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Nde*I and *Bgl*II that in the vector MCS and which were also incorporated into the PCR product in the forward and reverse primers. Ligation, screening for correct recombinant plasmids and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1. The plasmid was called pACYCDuet-*tssA*$^{EI}_{MD}$V4.

E. coli strain BL21(λDE3) was used for overexpressing TssA$^{EI}_{MD}$V4. However, no TssA$^{EI}_{MD}$V4 was overproduced with 1 mM IPTG induction at both 30^{o}C and 37^{o}C upon SDS-PAGE analysis (Figure 6.22). Other *E. coli* strains C41(λDE3) and C43(λDE3) were also used in attempts to overproduce the protein. Nevertheless, it was still not possible to overproduce TssA$^{EI}_{MD}$V4. This was unexpected as the same forward primer was used to successfully express TssA$^{EI}_{MD\text{-}CTD}$V4.

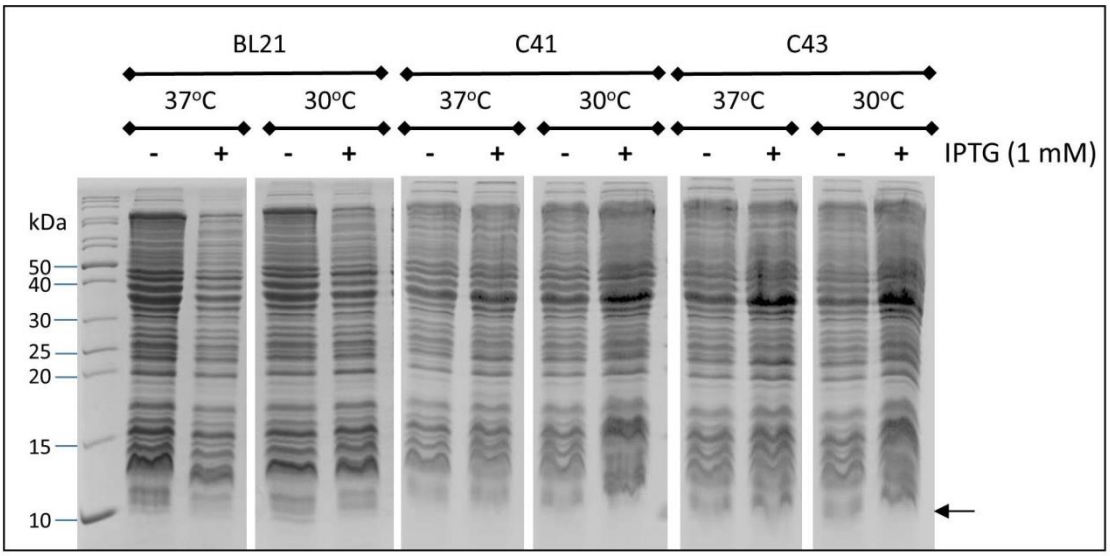


Figure 6.22 Overproduction of TssA$^{EI}_{MD}$V4. Coomassie blue-stained 15% SDS-PA gels showing the induction of pACYCDuet-*tssA*$^{EI}_{MD}$V4 in *E. coli* strains BL21(λDE3)/C41(λDE3)/C43(λDE3) under different conditions. Host strains, induction temperatures and concentrations of IPTG are as indicated. Arrow indicates the expected location of TssA$^{E}_{MD}$V4 based on its size (~10.2 kDa). '-', total cell protein from uninduced cells; '+', total cell protein from cells following induction.

6.3.3.5 X-ray crystallography of $TssA^{EI}_{MD}$

As $_{His6}$.$TssA^{EI}_{MD}$ (G223-A387) was the only successfully expressed soluble form of the MD, it was subjected to crystallisation trials. $_{His6}$.$TssA^{EI}_{MD}$ formed crystals in three tested conditions. The structure of it was solved from a crystal grown in 0.04 M potassium phosphate (monobasic), 16 % (w/v) PEG 8000 and 20 % (v/v) glycerol. The structure was solved at 2.1 Å resolution by experimental phasing.

The structure of $TssA^{EI}_{MD}$ forms a dimeric alpha helical bundle, and is composed of eight alpha-helices in each chain, i.e. H1 (D231-R248), H2 (A253-W264), H3 (A286-M296), H4 (Q301-L313), H5 (F318-L331), H6 (G334-R350), H7 (P352-E356) and H8 (P368-W373) (Figure 6.23 A and B). The dimer forms through hydrophobic interactions between side chains in amino acid residues W234 and L245 in H1 and V254 and L258 in H2 with each monomer in inverted orientation with respect to the other (Figure 6.23 C).

Prior to determination of the $TssA^{EI}_{MD}$ structure, Psipred v3.3 was used for the secondary structure prediction (Figure 1.12). The Psipred prediction is largely in agreement within the solved structure, except for an extra helix (H7) which is present in the solved structure but is absent according to the Psipred prediction (Figure 6.23 D).

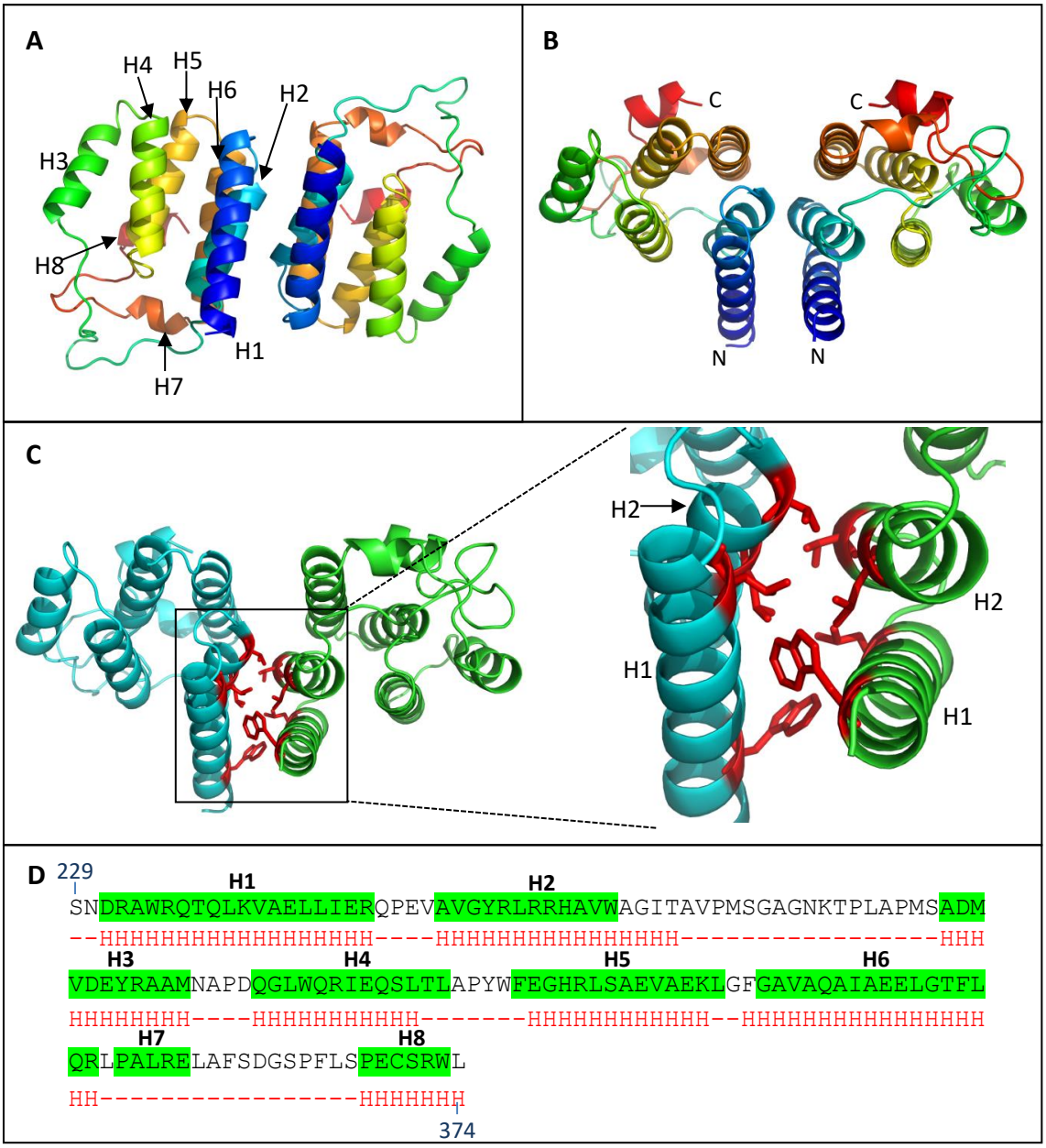


Figure 6.23 Structure of TssA$^{EI}_{MD}$. A. Side view of the structure of the TssA$^{EI}_{MD}$ dimer. **B**. Top view of the structure of the TssA$^{EI}_{MD}$ dimer. **C**. Hydrophobic interaction between the two monomers. Two chains are coloured in cyan and green. Residues involved in hydrophobic interactions of H1 and H2 are shown in red in both chains. The black rectangle encloses the interacting regions that are shown enlarged in the image on the right. **D**. Amino acid sequence and location of alpha-helices of one of the TssA$^{EI}_{MD}$ chains compared to the Psipred prediction. The helix numbers are labelled above the sequences in black bold font with the amino acids of the actual helices are highlighted in green. Dashes in red are Psipred predicted coil; 'H' in red are Psipred predicted helices. Residues S229 to L374 were resolved in the structure. The structural work was carried out by Sam Dix, Department of Molecular Biology and Biotechnology, University of Sheffield.

6.3.4 Overproduction and purification of $TssA^{EI}_{CTD}$ derivatives

6.3.4.1 $_{His6}$.$TssA^{EI}_{CTD}$

As it was shown that $TssA^{EI}_{MD}$ is able to dimerise, the oligomerisation of $TssA^{EI}_{MD-CTD}$ into a decamer was then hypothesized to be caused by the ability of its CTD to oligomerise. In order to test the hypothesis, DNA encoding AHA1844 residues S381-K478 was amplified using primers pACYC-AHA1844.CTDSD2for and pACYC-AHA1844.rev. The amplified $tssA^{EI}_{CTD}$ DNA fragment (~339 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Bam*HI and *Hin*dIII that recognized sites incorporated in the PCR product by the forward and reverse primers and the plasmid MCS. This resulted in a hexa-histidine tag encoded by the vector located at the N-terminus of $TssA^{EI}_{CTD}$. Ligation, screening for correct recombinant plasmids and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1. The plasmid was called pACYCDuet-His_6.$tssA^{EI}_{CTD}$.

$_{His6}$.$TssA^{EI}_{CTD}$ was overproduced in a huge amount from *E. coli* strain BL21(λDE3) cells containing pACYCDuet-His_6.$tssA^{EI}_{CTD}$ following induction, and ~50% of the overproduced protein remained soluble following cell lysis as judged by SDS-PAGE (Figure 6.24 A). $_{His6}$.$TssA^{EI}_{CTD}$ was purified by IMAC on a HisTrap HP column. The flow-through was recycled onto the column for maximising the chance of protein binding. $_{His6}$.$TssA^{EI}_{CTD}$ bound to nickel resin very well, and was eluted from the column in the range 133-250 mM imidazole (Figure 6.24 B). The elution fractions containing $_{His6}$.$TssA^{EI}_{CTD}$ were combined, concentrated (MWCO 10 kDa) and subjected to SEC on a Superose 12 column for MW estimation. The elution profile of SEC shows a single peak at 21.42 minute, consistent with a protein MW of ~125 kDa (Figure 6.24 C). Upon SDS-PAGE analysis, the peak fractions were found to contain $_{His6}$.$TssA^{EI}_{CTD}$ based on its migration (Figure 6.24 D). Given the monomer MW of $_{His6}$.$TssA^{EI}_{CTD}$ is 12.4 kDa, $_{His6}$.$TssA^{EI}_{CTD}$ is estimated to form a decamer, which is in agreement with the decameric complex that $TssA^{EI}_{MD-CTD}$ is capable of forming. Therefore, $TssA^{E}_{CTD}$ is responsible for the decamerisation of the MD-CTD and is consistent with the large complex formed by full-length $TssA^{EI}$. The results support the idea that the SEC-MALS estimation for the oligomerisation status of full-length $TssA^{EI}$ is a slightly overestimation (Section 6.2.3), and that the true oligomerisation status is a decamer. The purified protein was

buffer exchanged into 5 mM Tris-HCl (pH 8.0) containing 50 mM NaCl using a Zeba column for crystallisation trials as described in Section 2.4.9.1.

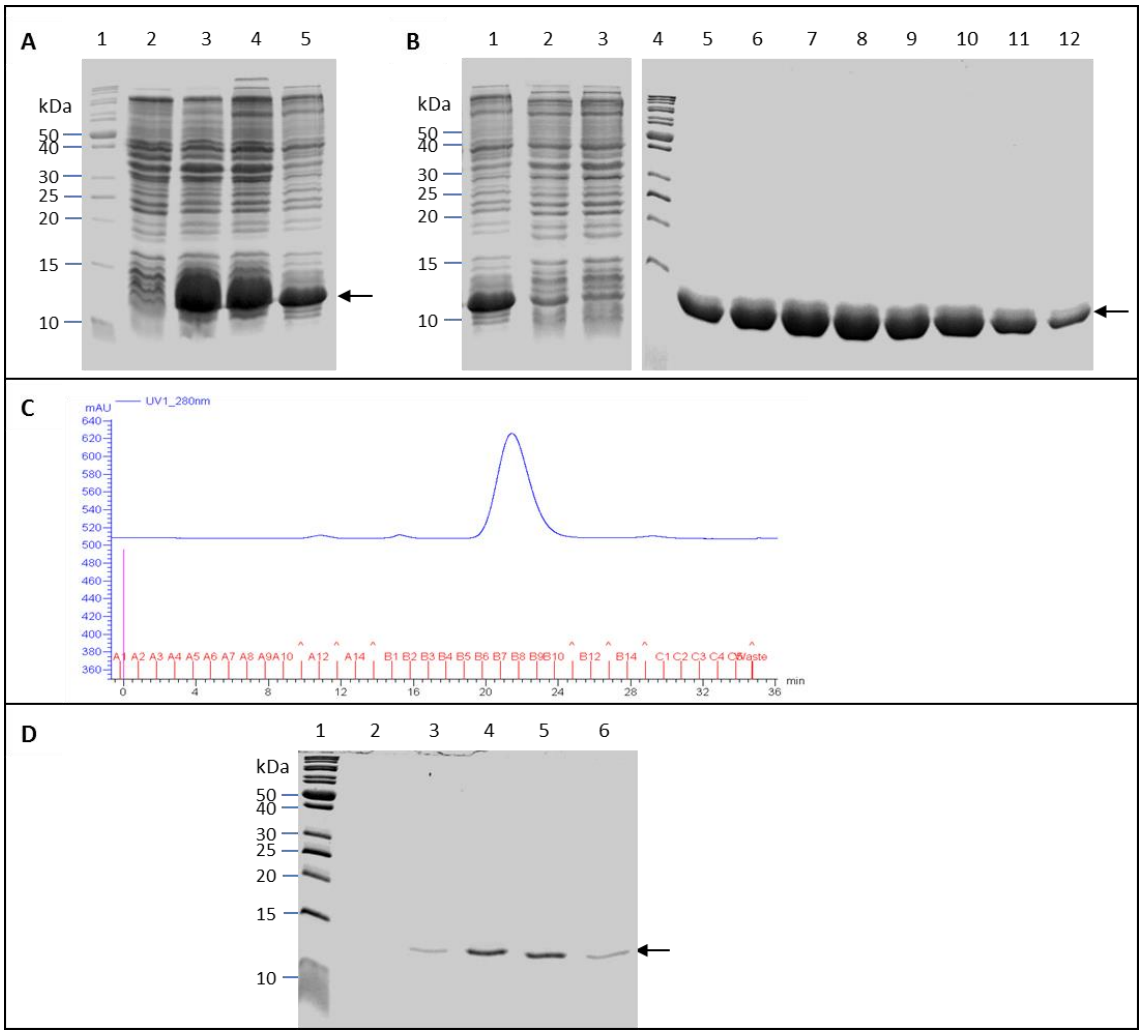


Figure 6.24 Overproduction, solubility and purification of $_{His6}$.TssA$^{EI}_{CTD}$. $_{His6}$.TssA$^{EI}_{CTD}$ synthesis was induced from pACYCDuet-His$_6$.*tssA$^{EI}_{CTD}$* in *E. coli* strain BL21(λDE3) with 1 mM IPTG at 37°C. Following induction, the soluble fraction of the cell lysate was subjected to IMAC on a HisTrap HP column in buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. Fractions containing $_{His6}$.TssA$^{EI}_{CTD}$ were combined, concentrated and subjected to SEC on a Superose 12 column in buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl and 10% glycerol. **A**. Overexpression and solubility of $_{His6}$.TssA$^{EI}_{CTD}$ analysed by electrophoresis in a 15% SDS-PA gel. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. 15% SDS-PAGE analysis of purification of $_{His6}$.TssA$^{EI}_{CTD}$ by IMAC. lane 1, soluble fraction of cell lysate following induction; lane 2, flow-through_1 from HisTrap column; lane 3, flow-through_2 from HisTrap column; lane 4, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 5-12, HisTrap fractions corresponding to the peak of the UV trace at 280 nm. **C**. Elution profile of $_{His6}$.TssA$^{EI}_{CTD}$ during SEC. **D**. 15% SDS-PAGE analysis of SEC of $_{His6}$.TssA$^{EI}_{CTD}$. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 2-6 elution fractions of SEC corresponding to the peak of the UV trace at 280 nm. Arrows in A, B and D indicate the expected locations of $_{His6}$.TssA$^{EI}_{CTD}$ based on its size (~12.4 kDa).

6.3.4.2 TssA$^{EI}_{CTD}$V2

$_{His6}$.TssA$^{EI}_{CTD}$ includes Helix 16 (residues 389-399) predicted by Psipred (Figure 1.12). However, it is located in a region that is not conserved in amino acid sequence alignments and may be part of the linker connecting MD to CTD. Moreover, this region is sensitive to protease cleavage in limited proteolysis analysis. Therefore, it was decided to make an untagged TssA$^{EI}_{CTD}$ derivative (named TssA$^{EI}_{CTD}$V2) in which most of H16 was omitted. DNA encoding residues H397-K478 of AHA1844 was amplified using primers AHA1844.CTDSD2.NdeI.for2 and AHA1844.BglII.rev, which removed 16 TssA^{EI} amino acids at the N-terminus of $_{His6}$.TssA$^{E}_{CTD}$. The amplified *tssA*$^{E}_{CTD}$V2 DNA fragment (~252 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites incorporated in the PCR product by the forward and reverse primers and the plasmid MCS. Ligation, screening for correct recombinant plasmids (termed pACYCDuet-*tssA*$^{EI}_{CTD}$V2) and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1. However, despite a lot of effort adjusting PCR conditions, digestion buffers and ligation conditions, it was not possible to obtain the correct construct. The optimised construct of pACYCDuet-*tssA*$^{EI}_{CTD}$V2 obtained contained sequence verified *tssA*$^{EI}_{CTD}$V2, however, it was not in the correct reading frame. Therefore, *tssA*$^{E}_{CTD}$V2 was released from the defective pACYCDuet-*tssA*$^{EI}_{CTD}$V2 by digesting with restriction enzymes *Nde*I and *Bgl*II, and was subsequently cloned into pETDuet plasmid that recognized the same sites. The plasmid was called pETDuet-*tssA*$^{EI}_{CTD}$V2.

E. coli strain BL21(λDE3) cells containing pETDuet-*tssA*$^{EI}_{CTD}$V2 was used for overexpressing TssA$^{EI}_{CTD}$V2. However, no TssA$^{EI}_{CTD}$V2 was overproduced with 1 mM IPTG induction at both 30^{o}C and 37^{o}C (Figure 6.25). Other *E. coli* strains C41(λDE3) and C43(λDE3) were also used in an attempt to overproduce the protein but this was still unsuccessful.

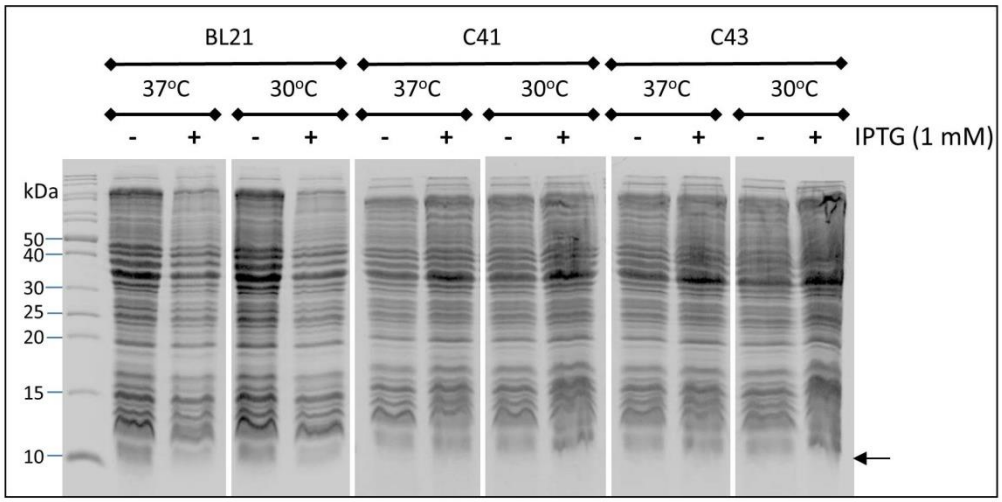


Figure 6.25 Overproduction of TssA$^{EI}_{CTD}$V2. Coomassie blue-stained 15% SDS-PA gels showing the induction of pETDuet-*tssA*$^{EI}_{CTD}$V2 in cells of *E. coli* strains BL21(λDE3)/C41(λDE3)/C43(λDE3) under different conditions as indicated. Host strains, induction temperatures and concentrations of IPTG are as indicated above the gels. Arrow indicates expected location of TssA$^{EI}_{CTD}$V2 based on its size (~9.2 kDa). '-', total cell protein from uninduced cells; '+', total cell protein from cells following induction.

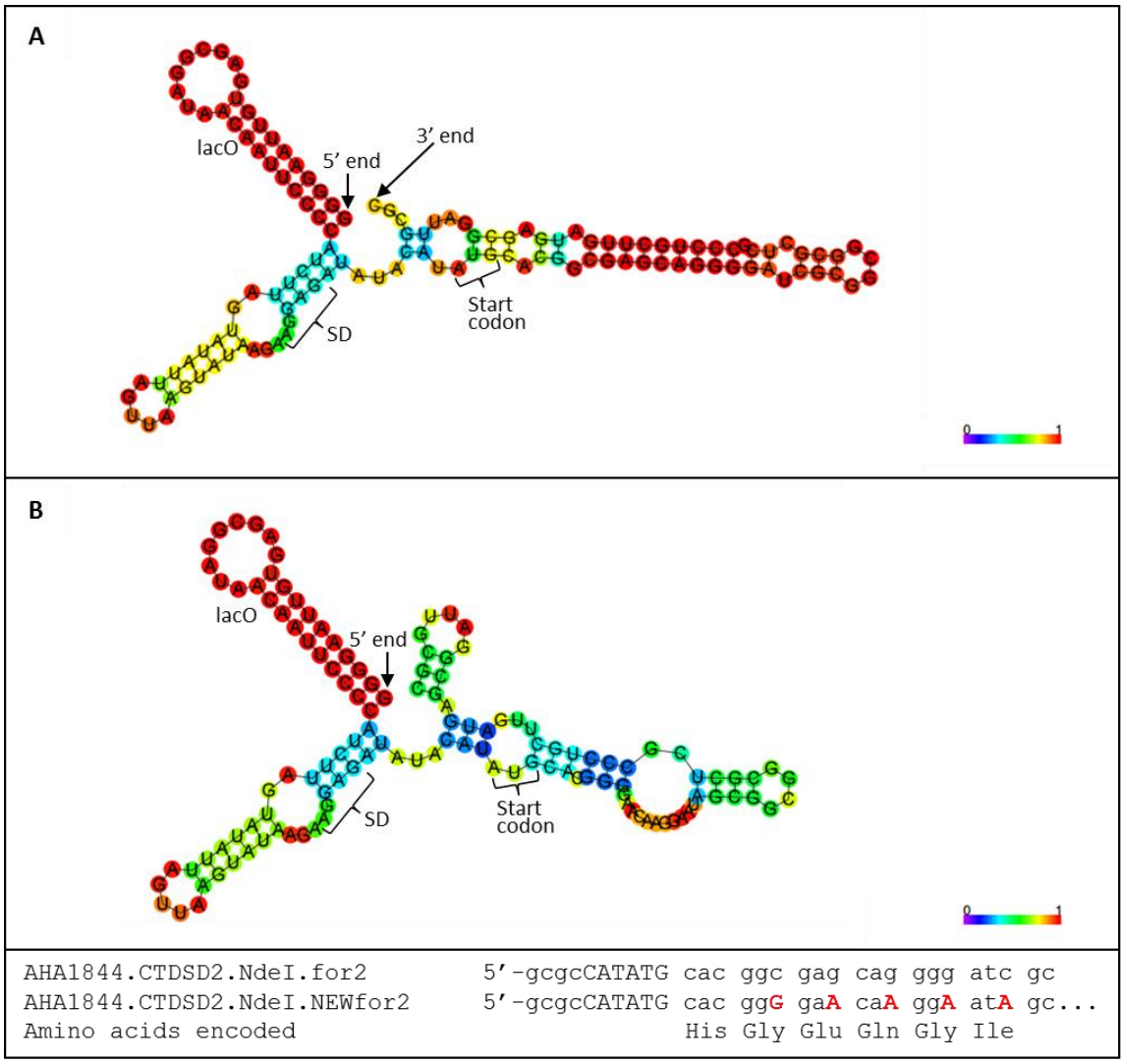
6.3.4.3 TssA$^{EI}_{CTD}$V3

As there was no expression of TssA$^{EI}_{CTD}$V2, the RNA folding status of its coding sequence adjacent to the translation initiation region present on the vector was checked by RNAfold WebServer. This showed that there was high potential to form a hairpin structure in the region downstream of the initiating AUG codon (Figure 6.26). Therefore, a new forward primer, AHA1844.CTDSD2.NdeI.NEWfor2, was designed to break the RNA hairpin without causing changes in the encoded amino acid sequence. Although the encoded amino acid sequence was the same as TssA$^{EI}_{CTD}$V2, a new name was given for the protein (TssA$^{EI}_{CTD}$V3) to distinguish it from the protein coded by pETDuet-*tssA*$^{EI}_{CTD}$V2. As the vector coded untranslated regions (UTR) for the second T7 promoter of both pACYCDuet and pETDuet vectors, where the gene was intended to be cloned, have the same sequence, pACYCDuet was used for inserting *tssA*$^{EI}_{CTD}$V3. PCR amplification was performed using AHA1844.CTDSD2.NdeI.NEWfor2 and AHA1844.BglII.rev primers. The amplified *tssA*$^{EI}_{CTD}$V3 DNA fragment (~252 bp) and pACYCDuet-1 plasmid were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites incorporated in the PCR product by the forward and reverse primers and the plasmid MCS. Ligation, screening for correct recombinant plasmids (termed pACYCDuet-*tssA*$^{EI}_{CTD}$V3) and verification of the nucleotide sequence of the inserted gene were carried out as described in Section 6.3.1.1.

TssA$^{EI}_{CTD}$V3 was overproduced in *E. coli* strain BL21(λDE3) cells containing pACYCDuet-*tssA*$^{EI}_{CTD}$V3 following induction as concluded by the presence of a large amount of polypeptide migrating with the expected size of TssA$^{EI}_{CTD}$V3 (9.2 kDa). The overproduced TssA$^{EI}_{CTD}$V3 remained in the soluble fraction upon cell lysis (Figure 6.27 A). To purify it, anion exchange chromatography was performed on a HiTrap DEAE column. Under these conditions TssA$^{EI}_{CTD}$V3 bound very well to the ion exchange resin, and the bound material was eluted from the column in the range 175-375 mM NaCl. However, there were still some contaminating proteins identified on SDS-PA gels (Figure 6.27 B). Therefore, the peak elution fractions of DEAE column were combined and concentrated (MWCO 100 kDa, as $_{His6}$.TssA$^{EI}_{CTD}$ has been shown to form a large complex (Section 6.3.4.1)) and the concentrated material was then loaded onto Superdex 200 10/300 GL for further purification by SEC. According to the elution profile, there were three protein peaks at 45.87, 68.03 and 77.19 ml (Figure 6.27 C).

Upon electrophoresis analysis, the peak at 68.03 ml contained the majority of $TssA^{EI}_{CTD}V3$ (Figure 6.27 D). The later peak (77.19 ml), which slightly overlapped with the earlier peak at 68.03 ml also comprised a small portion of $TssA^{EI}_{CTD}V3$ but mainly contained contaminants with MW at 40-50 kDa. The fractions corresponding to the $TssA^{EI}_{CTD}V3$ protein peak of 68.03 ml was estimated to be 60-70% pure and were combined and concentrated, subjected to buffer exchanged into 5 mM Tris-HCl (pH 8.0) containing 50 mM NaCl using a Zeba column for crystallisation trials (Figure 6.27 E).

As the purity after SEC was still not very high, with contaminants of MW ~100 kDa which is very similar to the MW of the $TssA^{EI}_{CTD}V3$ oligomer, the side peak fractions of 68.03 ml from Superdex 200 10/300 GL were also collected and further purified by anion exchange chromatography on a HiTrap Q HP column. $TssA^{E}_{CTD}V3$ bound to the HiTrap Q resin in buffer containing MES (pH 6.5) and was eluted in the range 222-361 mM NaCl (Figure 6.27 F). However, the purity of $TssA^{EI}_{CTD}V3$ was not significantly improved. Due to time constraints, no other purification methods were attempted.




```
AHA1844.CTDSD2.NdeI.for2        5′-gcgcCATATG cac ggc gag cag ggg atc gc
AHA1844.CTDSD2.NdeI.NEWfor2     5′-gcgcCATATG cac ggG gaA caA ggA atA gc...
Amino acids encoded                           His Gly Glu Gln Gly Ile
```

Figure 6.26 Comparison of RNA folding status of UTR and N-terminal 18 amino acids of sequences encoding TssA$^{EI}_{CTD}$V2 and TssA$^{EI}_{CTD}$V3. RNA folding status of the vector untranslated region beginning at the start of the T7 promoter-directed transcript (5' end) and first 18 codons specifying the N-terminus TssA$^{EI}_{CTD}$V2/V3 (H397-A414 of native TssA^{EI}) were analysed on the RNAfold WebServer. **A**, RNA folding status of TssA$^{EI}_{CTD}$V2 showing first 18 codons of *tssA*$^{EI}_{CTD}$; **B**, RNA folding status of TssA$^{EI}_{CTD}$V3 showing first 18 codons of *tssA*$^{EI}_{CTD}$. The possibility of forming a hairpin structure is indicated at the bottom of the figures, i.e. purple=0, red=1. The sequences at the bottom are the two forward primers that were used for PCR amplification and cloning of *tssA*$^{EI}_{CTD}$V2/V3. The complete sequence AHA1844.CTDSD2.NdeI.for2 primer is shown and the corresponding region of AHA1844.CTDSD2.NdeI.NEWfor2 primer is shown with the annealing region omitted. The codons labelled in bold red font in the AHA1844.CTDSD2.NdeI.NEWfor2 primer are those bases that were changed for breaking the hairpin structure.

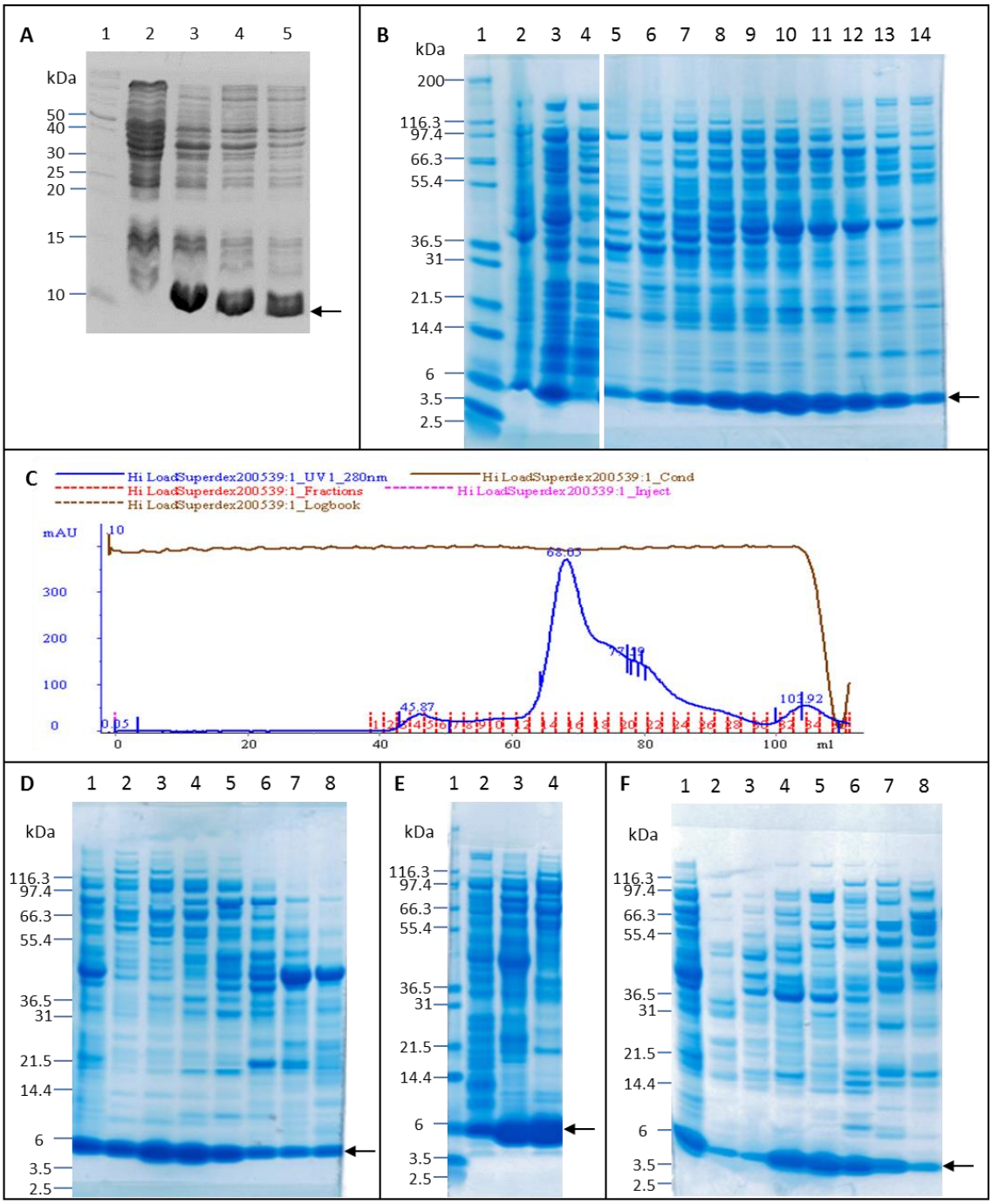


Figure 6.27 Overproduction and purification of TssA$^{EI}_{CTD}$V3.

Figure 6.27 Overproduction and purification of TssA$^{EI}_{CTD}$V3. TssA$^{EI}_{CTD}$V3 synthesis was induced from pACYCDuet-*tssA*$^{EI}_{CTD}$V3 in *E. coli* strain BL21 (λDE3) with 1 mM IPTG at 37^{o}C. Following induction, the soluble fraction of the cell lysate was subjected to anion exchange chromatography using a HiTrap DEAE FF column in buffer containing 50 mM Tris-HCl (pH 8.0). The bound material was eluted with an increasing gradient concentration of NaCl up to 0.5 M. Peak elution fractions were then combined, concentrated and subjected to SEC on a Superdex 200 10/300 column in buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. Peak fractions were combined and buffer exchanged before using for crystallisation trials. Side peak elution fractions of SEC were further purified by anion exchange chromatography on a HiTrap Q column in buffer containing 20 mM MES (pH 6.5). The bound material was eluted with an increasing gradient concentration of NaCl up to 0.5 M. **A**. Coomassie blue-stained 15% SDS-PA gel showing the overproduction and solubility of TssA$^{EI}_{CTD}$V3. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **B**. 4-12% gradient SDS-PA gels (Novex) showing the purification of TssA$^{EI}_{CTD}$V3 on a HiTrap DEAE column. Lane 1, protein maker (Novex); lane 2, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 3, soluble fraction of cell lysate following induction; lane 4, unbound material from DEAE column; lanes 5-14, elution fractions from DEAE column corresponding to the peak of the UV trace at 280 nm. **C**. Elution profile of TssA$^{EI}_{CTD}$V3 during SEC. **D**. 4-12% gradient SDS-PA gels (Novex) showing the purification of TssA$^{EI}_{CTD}$V3 by SEC. Lanes 1-5, fractions from Superdex 200 column eluting at 64-70 ml (peak at 68.05 ml); lanes 6-8, fractions from Superdex 200 column eluting at 70.1-82 ml. **E**. 4-12% gradient SDS-PA gel (Novex) showing the combined elution fractions during purification of TssA$^{E}_{CTD}$V3. Lane 1, protein marker (Novex); lane 2, soluble fraction of cell lysate following induction; lane 3, combined peak elution fractions from DEAE column; lane 4, combined peak elution fraction from Superdex 200 column. **F**. 4-12% gradient SDS-PA gel (Novex) showing the purification of TssA$^{EI}_{CTD}$V3 on HiTrap Q HP column. Lane 1, combined side peak fractions of TssA$^{EI}_{CTD}$V3 from Superdex 200 column (HiTrap Q load); lanes 2-8, elution fractions from HiTrap Q column corresponding to the peak of the UV trace at 280 nm.

6.3.4.4 Transmission electron microscopy of $TssA^{EI}_{CTD}$

$TssA^{EI}_{CTD}$ was observed by SEC to be able to oligomerise into multimers. To obtain structural information on $TssA^{EI}_{CTD,\ His6}$.$TssA^{E}_{CTD}$ was purified as described in Section 6.3.4.1, and subjected to TEM. $TssA^{E}_{CTD}$ was observed to form a five-pointed star with a central pore (Figure 6.28), but the five-pointed star shape was not as obvious as that of $TssA^{EI}_{MD\text{-}CTD}$, suggesting the MD locates to the periphery of the complex and extends the arms of the 'star'.

In order to confirm the features of $TssA^{EI}_{CTD}$, averaging of $TssA^{EI}_{CTD}$ images was performed. Purified $TssA^{EI}_{CTD}$ (0.03 mg/ml) was loaded onto freshly glow-discharged grids and stained with 0.75% uranyl formate before being analysed by TEM (CM200 FEG, Philips). Averaging was carried out using 707 single particles. Imagic-5 software was used for the image analysis. Alignments of single particles were based on cross correlation. Direct alignment was carried out including translational and rotational modes after filtering and normalisation of the images. The shape of five outer projections with a central hole of $TssA^{EI}_{CTD}$ complex converged after six steps of summing and alignment against the total sum (Figure 6.29 A).

Test of the rotational symmetry of $_{His6}$.$TssA^{EI}_{CTD}$ was conducted for confirming the specific symmetries that $TssA^{EI}_{CTD}$ may have, although a shape with 5 outer projections had already become clear. The rotational symmetry test was based on correlation between the rotated total average with its unrotated self. Two- to 12-fold symmetries were tested, and the results showed that five-fold symmetry gave the highest correlation coefficient score (81.7%) compared to all the other tested symmetries (Figure 6.29 B). The rotational symmetry test provided further evidence of the five-pointed star shape of $TssA^{EI}_{CTD}$. Based on the averaging results, five-fold symmetry was imposed on the averages after six steps of summing and aligning against the total sum (Figure 6.29 C). The size estimation was conducted on the total average and average after five-fold symmetrisation, and the results showed the $TssA^{EI}_{CTD}$ star shape has an exterior diameter of 100-120 Å and the diameter of the central pore is 36-40 Å. The interior diameter is consistent with that of $TssA^{EI}_{MD\text{-}CTD}$ (36 Å). The symmetrised $TssA^{EI}$ CTD and MD-CTD final averages were superimposed in a single image showing that the MD locates at the exterior of the $TssA^{EI}_{CTD}$ star shape in two possible positions (Figure 6.30).

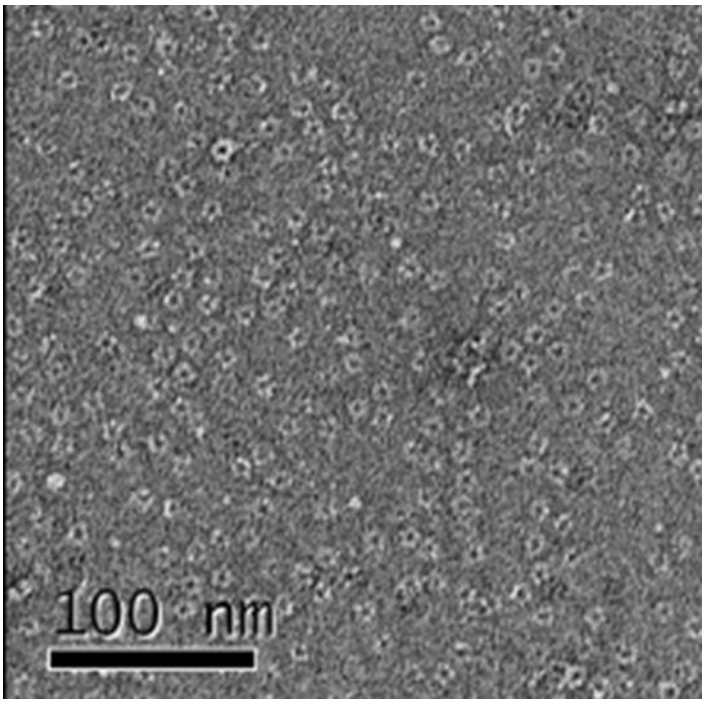


Figure 6.28 Negative stain transmission electron micrograph of $TssA^{EI}_{CTD}$. $TssA^{EI}_{CTD}$ (0.03 mg/ml) in buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole was loaded onto freshly glow-discharged grids and stained with 0.75% uranyl formate stain before being analysed by TEM (CM200 FEG, Philips).

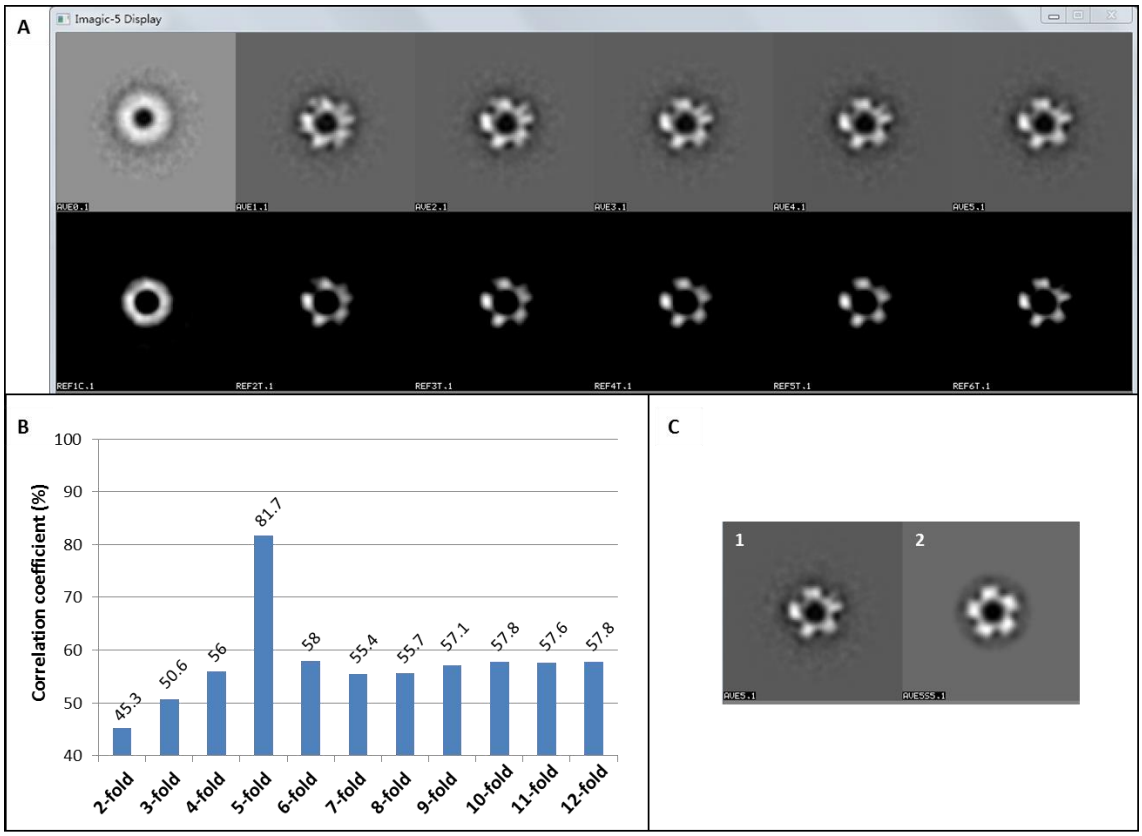


Figure 6.29 Averaging of $_{His6}$.TssA$^{EI}_{CTD}$ single particles. 707 single particles of $_{His6}$.TssA$^{EI}_{CTD}$ were collected following negative stain transmission electron microscopy. Averaging was carried out using Imagic-5 software. Six steps of direct alignment were run until results converged. **A**. Total sum of $_{His6}$.TssA$^{EI}_{CTD}$. Upper panel, total sum of $_{His6}$.TssA$^{EI}_{CTD}$ after iterations of direct alignment, from left to right: 1-6 iterations; lower panel, filtered total sum corresponding to the images above them. **B**. Test rotational symmetry of $_{His6}$.TssA$^{EI}_{CTD}$. The X-axis shows the specific symmetry that was tested, and Y-axis shows the correlation coefficient (%) of all the single particles that rotate the single particle for a specific angle according to the fold of symmetry, then correlates the rotated particle with its unrotated self. **C**. Five-fold symmetry applied on the final average. 1, before imposing symmetry; 2, after imposing symmetry.

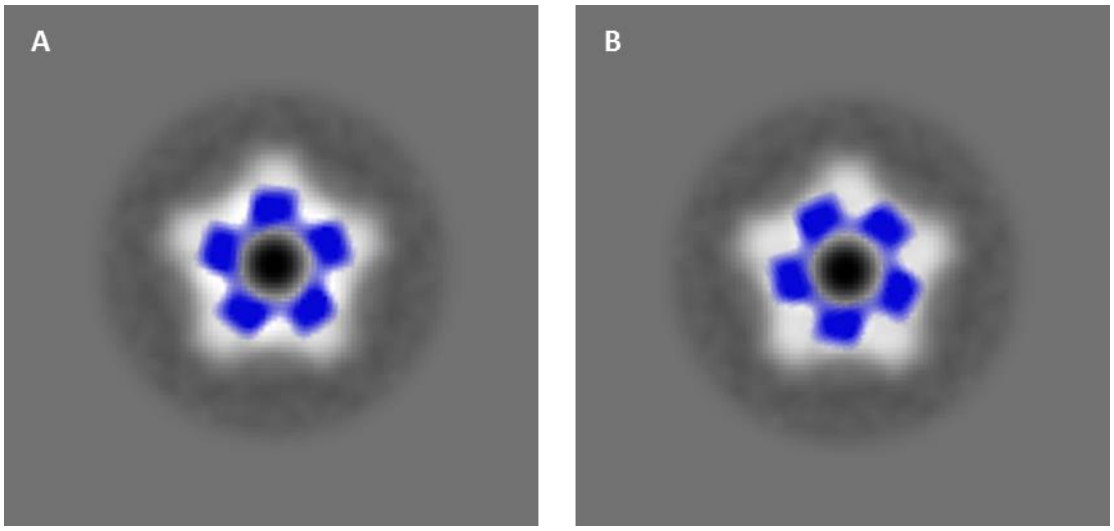


Figure 6.30 Merged images of symmetrised $TssA^{EI}_{MD\text{-}CTD}$ and $TssA^{EI}_{CTD}$. Single particles of $TssA^{EI}_{MD\text{-}CTD}$ and $TssA^{EI}_{CTD}$ were stained with 0.75% uranyl formate before analysis by TEM (CM200 FEG, Philips). Superimposed final averages of $TssA^{EI}_{MD\text{-}CTD}$ (in white) and $TssA^{EI}_{CTD}$ (in blue) with five-fold symmetry imposed. Two possible positions of MD locating at the exterior of the CTD are shown in A and B. A, MD extends each arm of the CTD star; B, MD creates new 'arms' between each adjacent pair of CTD arms.

6.3.4.5 Molecular weight estimation of $_{His6}$.TssA$^{EI}_{CTD}$ by SEC-MALS

As TssA$^{EI}_{CTD}$ was observed to form five-pointed star conformation with a central pore, the size estimation by SEC would not be accurate as it is not a globular protein and in addition the volume includes the empty space of the ring. Therefore, SEC-MALS was performed for a more accurate MW estimation of TssA$^{EI}_{CTD}$, purification was conducted as described in Section 6.3.4.1.

SEC-MALS was performed using a Superdex 200 10/300 GL column. According to the elution profile of TssA$^{EI}_{CTD}$, there is a tiny amount of aggregate material in the void volume of the column (~15 minute) followed by the major peak (~26.8 minute) and then a small amount of material in RI and UV traces but negligible LS at 37 minutes (Figure 6.31 A). Integration of the RI trace indicated ~154 µg protein in the major peak that is consistent with the amount injected. The peak analysis of TssA$^{EI}_{CTD}$ was based on the elution profile using Astra software which defined the peak region into two, which are Peak 1 (26.34-27.24 minute) and Peak 2 (25.21-28.64 minute). The defined Peak 2 comprises Peak 1, which has a broader range in the elution time.

The MW was determined by a Zimm fit procedure at each point, then averages calculated over the peak region (Figure 6.31 B). There is only one peak in the elution profile. The estimated MW for the peak was 122 kDa. There is a slight upward slope on the MW plot but this is probably not significant. There was no indication of any material with a smaller MW. Given a monomer MW of 12.22 kDa, the result suggests that TssA$^{EI}_{CTD}$ comprises 10 subunits. The result of SEC-MALS of TssA$^{EI}_{CTD}$ is consistent with that obtained from SEC. Hence, this provided further evidence that TssA$^{EI}_{CTD}$ forms a decamer, and supports the idea that TssA$^{EI}_{CTD}$ is responsible for the decamerisation of the full-length TssA^{EI}.

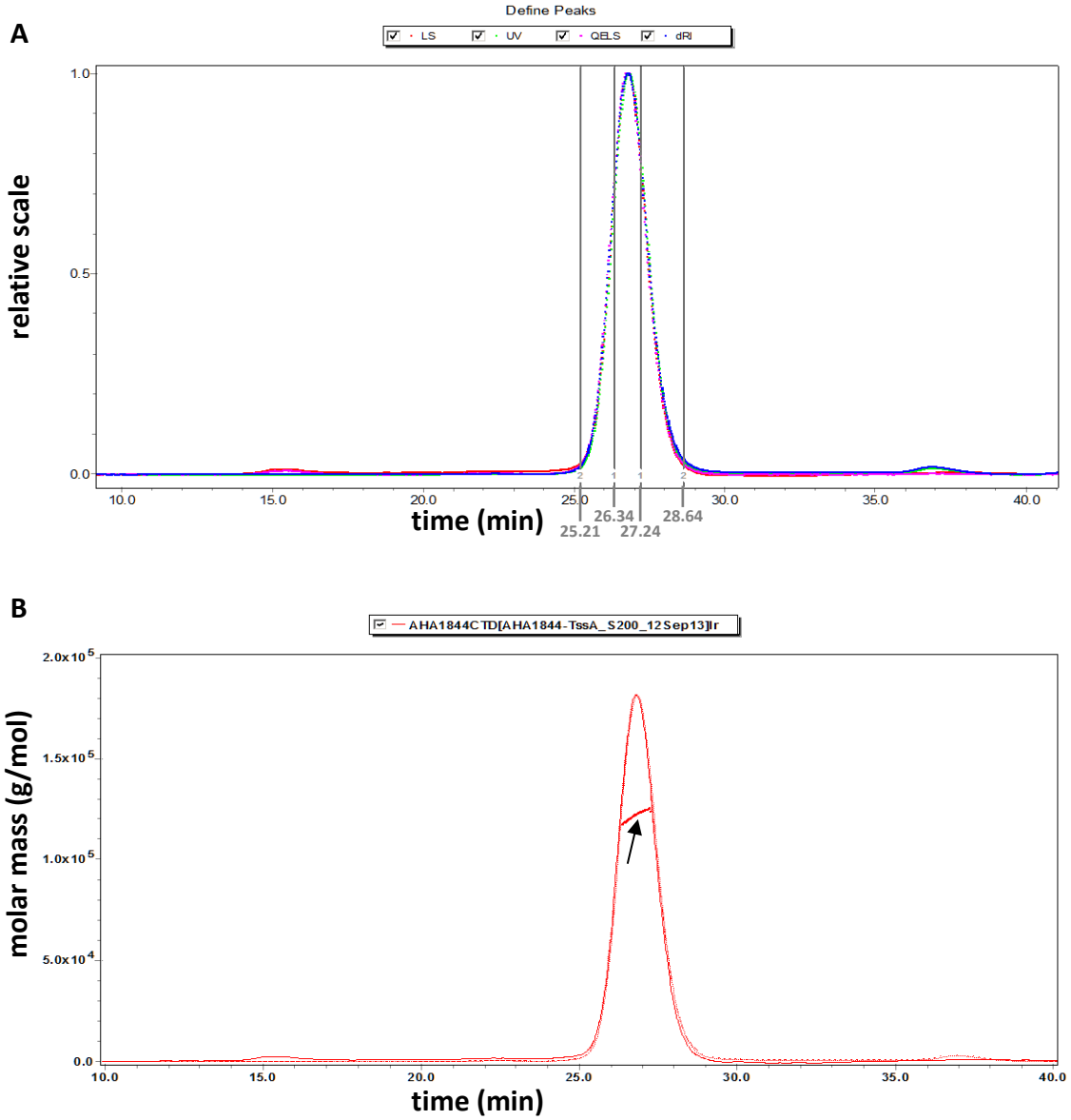


Figure 6.31 SEC-MALS analysis of $_{His6}$.TssA$^{EI}_{CTD}$. A. $_{His6}$.TssA$^{EI}_{CTD}$ elution profile and peak positions. The normalised signals for light scattering (LS) are shown in red, refractive index (RI) in blue, UV absorption (280 nm) in green and quasi-elastic light scattering (QELS) in magenta. The peak regions chosen for analysis are indicated by numbered vertical grey lines that defined by Astra software. The X-axis shows elution time in minutes and Y-axis shows relative scale. **B.** $_{His6}$.TssA$^{EI}_{CTD}$ molar mass plot. The MW is determined by a Zimm fit procedure at each point. Solid line represents LS trace. Dashed and dotted traces represent RI and UV responses, respectively. Molecular weight estimate curve is shown in heavier dots (pointed with black arrows). The X-axis is the elution time in minutes and the Y-axis is the MW scale. Figures provided by University of York Bioscience Technology Facility.

6.3.4.6 Role of H20 in oligomerisation of $TssA^{EI}_{CTD}$

According to the secondary structure prediction of $TssA^{EI}$ by Psipred, $TssA^{EI}_{CTD}$ consisted of five helices (H16-H20) (Figure 1.12). As it was shown that the C-terminal helix (H15) of $TssA^{S}$ is responsible for the large complex formation of $TssA^{S}$, a hypothesis was formulated similarly where the C-terminal helix of $TssA^{EI}$ (H20) may be responsible for the decamerisation of $TssA^{EI}$.

6.3.4.6.1 Construction of a plasmid for $_{His6}.TssA^{EI}_{CTD.H16\text{-}H19}$ overproduction

In order to test the hypothesis, a plasmid construct expressing partial $TssA^{EI}_{CTD}$ (residues S381-P463 of AHA1844) was created, which expressed the helices 16-19 of $TssA^{EI}$. As a control, plasmid pACYCDuet-His_6.$tssA^{E}_{CTD}$, expressing the full-length $TssA^{EI}_{CTD}$ was used which was constructed as described in Section 6.3.4.1. The construction of pACYCDuet-His_6.$tssA^{EI}_{CTD}$.H16-H19 was achieved by deleting the coding sequence of H20 from pACYCDuet-His_6.$tssA^{EI}_{CTD}$ by digesting the plasmid with *Sma*I and *Pac*I. A DNA polymerise filling-in reaction was carried out as described in Section 2.3.11 for generating blunt ends before the ligation step was performed. Following transformation of *E. coli* MC1061, plasmid minipreps were performed to identify recombinant plasmids of pACYCDuet-His_6.$tssA^{EI}_{CTD}$.H16-H19. Deletion of H20 was verified by DNA sequencing.

6.3.4.6.2 Overproduction and purification of $_{His6}.TssA^{EI}_{CTD.H16\text{-}H19}$

$_{His6}.TssA^{EI}_{CTD.H16\text{-}H19}$ was overproduced in *E. coli* strain BL21(λDE3) cells containing pACYCDuet-His_6.$tssA^{EI}_{CTD}$.H16-H19 as evidenced by the presence of a large amount of polypeptide migrating at the expected size of ~10.7 kDa in SDS-PA gels (Figure 6.32 A). The protein remained in the soluble fraction following cell lysis, and was subjected to IMAC on a HisTrap HP column for purification. $_{His6}.TssA^{EI}_{CTD.H16\text{-}H19}$ bound to the nickel resin very well, and the bound material was eluted in the range 222-289 mM imidazole (Figure 6.32 B). Peak elution fractions were combined, concentrated (MWCO 10 kDa, Vivaspin centrifugal concentrator) and subjected to further analysis by SEC on

a Superdex 200 GL column for oligomerisation analysis. The elution profile showed a major peak at 16.84 ml that contained $_{His6}$.TssA$^{EI}_{CTD.H16\text{-}H19}$ (Figure 6.32 C). The peak at 18.84 ml consistents with a protein MW less than 8.3 kDa. The peak at 16.84 ml contained the vast majority of the loaded material, corresponding to a protein MW of 24 kDa calculated using the calibration curve. Given the monomeric MW of $_{His6}$.TssA$^{EI}_{CTD.H16\text{-}H19}$ is 10.7 kDa, this indicates that it forms a dimer. Therefore, the region encoding the predicted C-terminal helix (H20) of TssA$^{EI}_{CTD}$ is crucial for the decamer structure of TssA$^{EI}_{CTD}$ and TssA^{EI}.

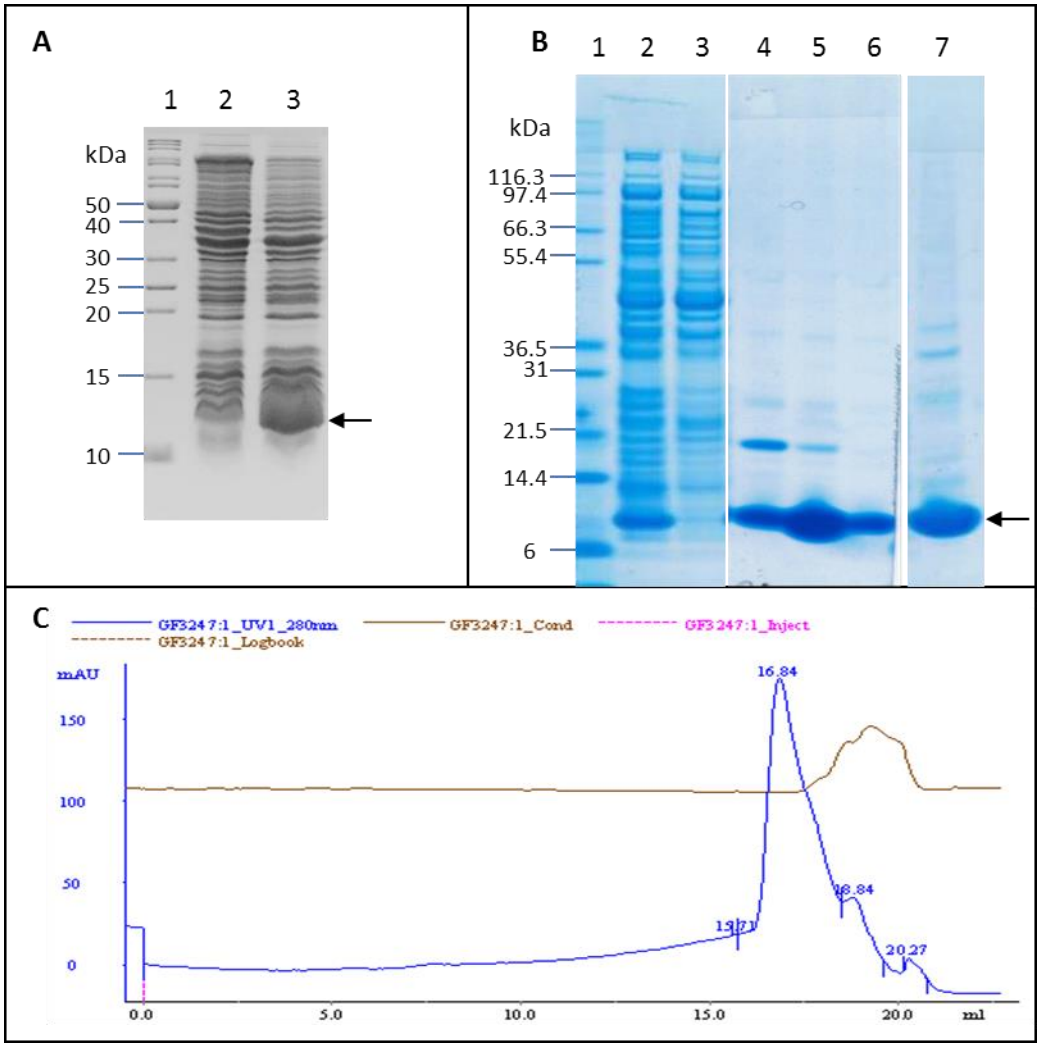


Figure 6.32 Overproduction, solubility and purification of $_{His6}$.TssA$^{EI}_{CTD.H16-H19}$. $_{His6}$.TssA$^{EI}_{CTD.H16-H19}$ synthesis was induced from pACYCDuet-His$_6$.*tssA*$^{EI}_{CTD.H16-H19}$ in *E. coli* strain BL21 (λDE3) with 1 mM IPTG at 37°C. Following induction, the soluble fraction of the cell lysate was purified by IMAC on a HisTrap HP column in buffer containing 50 mM Tris-HCl (pH 8.0) and 0.5 M NaCl. The bound material was then eluted with an increasing gradient concentration of imidazole up to 500 mM. Peak elution fractions were combined, concentrated and subjected to SEC on a Superdex 200 GL column in buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. **A**. Coomassie blue-stained 15% SDS-PA gel showing the overproduction of $_{His6}$.TssA$^{EI}_{CTD.H16-H19}$. Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction. **B**. 4-12% gradient SDS-PA gels (Novex) showing the purification of $_{His6}$.TssA$^{EI}_{CTD.H16-H19}$ on a HisTrap HP column. Lane 1, protein marker (Novex); Lane 2, soluble fraction of cell lysate following induction; lane 3, unbound material from HisTrap; lanes 4-6, elution fractions corresponding to the peak of the UV trace at 280 nm; lane 7, concentrated peak elution fraction from HisTrap column containing purified $_{His6}$.TssA$^{EI}_{CTD.H16-H19}$. **C**. Elution profile of $_{His6}$.TssA$^{EI}_{CTD.H16-H19}$ during SEC.

6.3.4.7 X-ray crystallography of $TssA^{EI}_{CTD}$

$_{His6}.TssA^{EI}_{CTD}$ (containing predicted helices H16-H20) and $TssA^{EI}_{CTD}$V3 (containing predicted helices H17-H20) were purified as described in Sections 6.3.4.1 and 6.3.4.3, and used for crystallisation trials. A crystal of $_{His6}.TssA^{EI}_{CTD}$ was obtained in one condition. However, it did not diffract well. The crystallographic analysis of $TssA^{EI}_{CTD}$ is still in progress with more crystals being grown and used for data collection.

The CTD structure presented in this study was obtained by masking the MD from the $TssA^{EI}_{MD-CTD}$ structure which was solved at 3.25 Å resolution by molecular replacement using the previous $TssA^{EI}_{MD}$ structure model (Section 6.3.3.5). The structural work was carried out by Sam Dix, Department of Molecular Biology and Biotechnology, University of Sheffield. In the overall structure, $TssA^{EI}_{CTD}$ forms a decamer organised as a pentameric star which is in agreement with the observation from EM (Section 6.3.4.4) where five-pointed stars with a central pore were observed (Figure 6.33 A-C). The $TssA^{EI}_{CTD}$ monomer contains five alpha-helices: L389-E399 (H9), A403-A414 (H10), P419-E436 (H11), L441-L456 (H12) and P463-S471 (H13) (helices numbering continued from the high resolution structure of MD) with H9-H12 fomring an anti-parallel helical bundle. The ten chains of the $TssA^{EI}_{CTD}$ structure are packed in a single layer with an inner lumen diameter of ~30 Å and an outer dimension of ~110 Å, which is largely consistent with the measurement from EM. The depth of it is ~30 Å. Each arm of the star is composed of two monomers that interact via hydrophobic interactions between H10-H12 of two adjacent chains. Three consecutive highly conserved amino acids within $TssA^{EI}$ orthologues, WEP, mainly locate in the loop between H12 and H13, and allow the association between two adjacent arms by hydrophobic burying between W461 of one chain with F423, L456, L458 and L465 of the other chain, and ionic interactions between the helical faces of the interacting chains. Thus, the WEP motif contributes to the inter-docking of arms in the pentameric star (Figure 6.33 D). The C-termini of the subunit domains are projecting in the same orientation perpendicular to the plane of the pentameric star; whereas, the N-termini of the $TssA^{EI}_{CTD}$ locate on the exterior of the pentameric star. The Psipred prediction (Figure 1.12) is in good agreement with the structure of $TssA^{EI}_{CTD}$ regarding the location of the alpha helices (Figure 6.33 E) where H9-H13 correspond to H16-H20 of the predicted secondary structure.

In the $TssA^{EI}_{MD-CTD}$ structure, the MD dimers are proposed to associate with the N-termini of the CTDs belonging to two adjacent arms connected by the WEP motif, generating new arms between each pair of CTD arms. As the interdomain linker (~13 amino acids) connecting the MD and CTD is not resolved in the structure, this model is based on the proximity of MD to CTD in the current $TssA^{EI}_{MD-CTD}$ structure (Figure 6.34 A and B). In the structure, MD dimers locate at different positions relative to associated CTDs. Moreover, the distance between the MD and CTD varies among chains between 12.6-31.3 Å (Figure 6.34 C and D), but within the distance of 13 amino acids (~45.5 Å in a β strand). This variation is probably due to the flexibility of the interdomain linker connecting these two domains. The MD in the $TssA^{EI}_{MD-CTD}$ structure also dimerises in the same manner as that observed from the $TssA^{EI}_{MD}$ structure (Section 6.3.3.5).

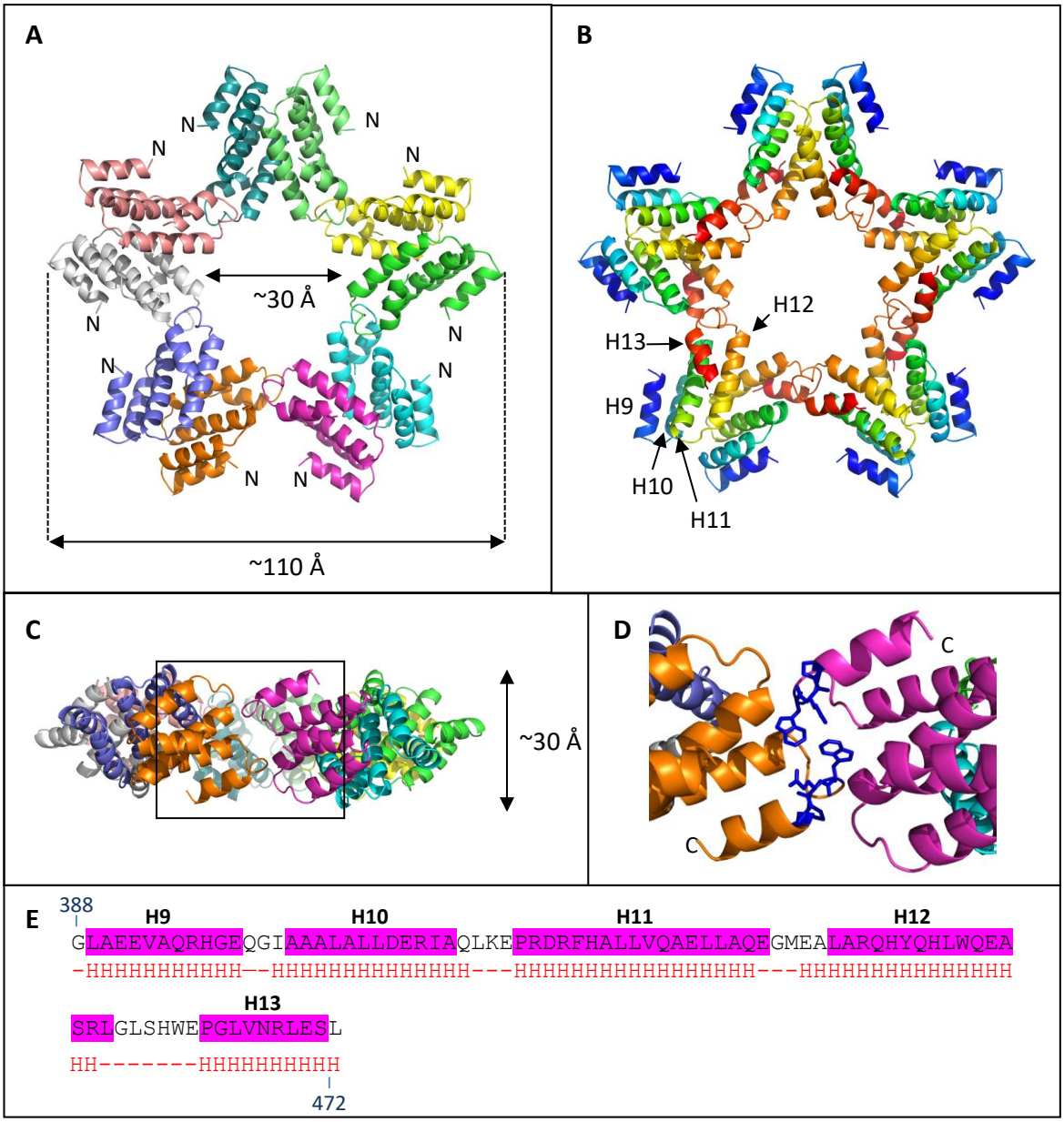


Figure 6.33 Structure of TssA$^{EI}_{CTD}$. Structure of TssA$^{EI}_{CTD}$ is shown by masking MD from the TssA$^{E}_{MD-CTD}$ structure. **A**. Top view of the TssA$^{EI}_{CTD}$ structure. Chains are shown in different colours. Dimensions are as indicated. **B**. Top view of the TssA$^{EI}_{CTD}$ structure. Each chain is coloured in rainbow from blue to red, corresponding to N- and C-terminus of the structure. **C**. Side view of the TssA$^{EI}_{CTD}$ structure. The depth is as indicated. The black rectangle encloses helices that are shown enlarged in Figure D. **D**. Conserved WEP motif for inter-docking the penameric TssA$^{EI}_{CTD}$ structure. The WEP amino acids are shown in blue. **E**. Amino acid sequence and helix organisation of one of the TssA$^{EI}_{CTD}$ chains compared to Psipred prediction. The helix numbers are labelled above the sequences in black bold font with the amino acids of helices highlighted in magenta. Dashes in red are Psipred predicted coil; 'H' in red are Psipred predicted helices. Residues G388-L472 were resolved in the structure.

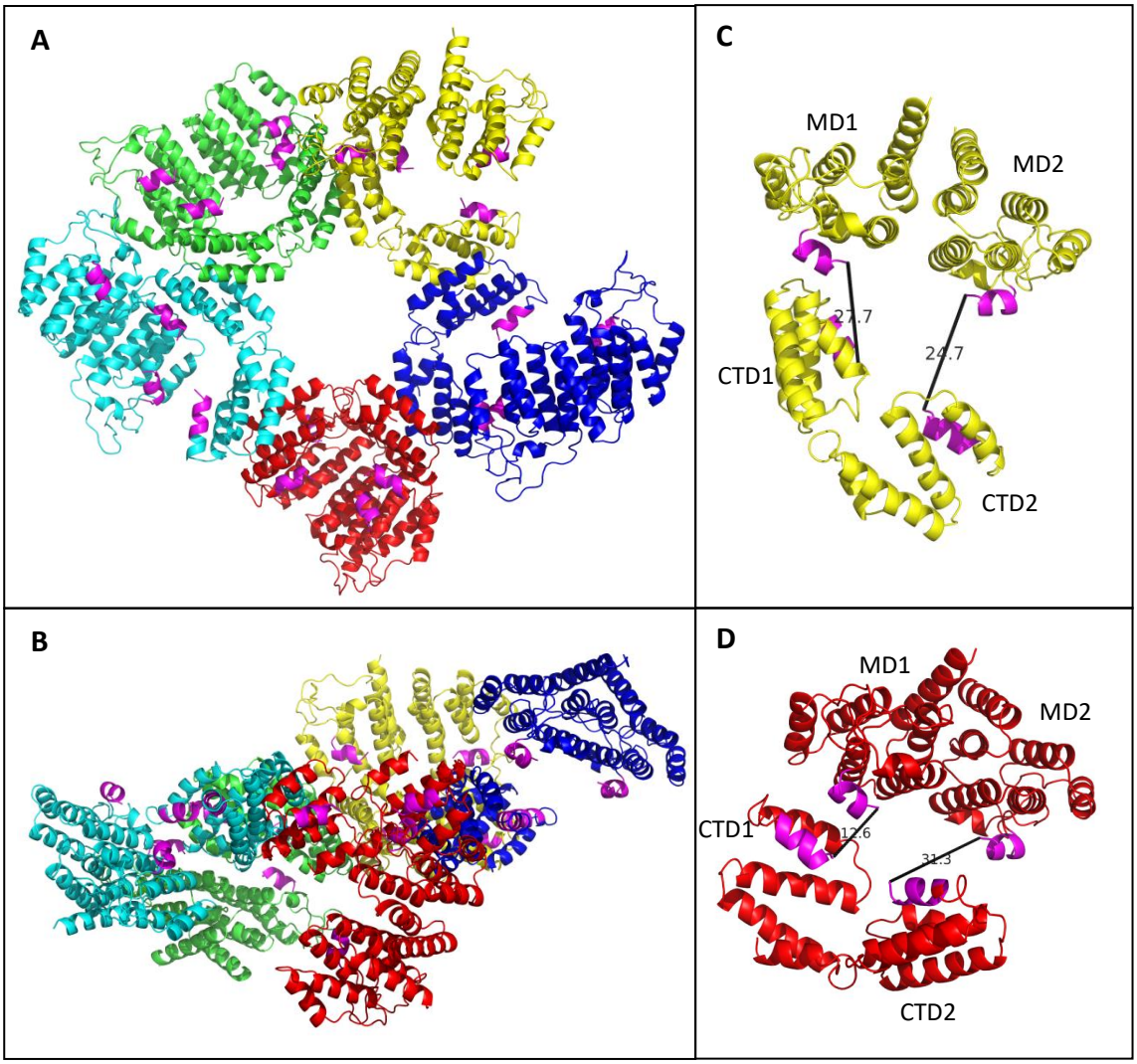


Figure 6.34 Structure of TssA$^{EI}_{MD-CTD}$. A. Top view of the TssA$^{EI}_{MD-CTD}$ structure. **B**. Side view of the TssA$^{EI}_{MD-CTD}$ structure with the CTD decameric star in the same plane. **C** and **D**, Distance between MD and CTD are shown in TssA$^{EI}_{MD-CTD}$ dimers as indicated in Å. The last helices in MD and CTD are coloured in magenta. Residues N230-L472 were resolved in the structure (residues G223-K478 were present in $_{His6}$.TssA$^{EI}_{MD-CTD}$). The interdomain linker connecting the MD and CTD is not resolved in the structure.

6.4 Analysis of T6SS activity in a *B. cenocepacia* H111 *tssA*^{S} mutant containing TssA^{EI} and TssA^{EI}-TssA^{S} hybrid proteins

TssA^{S} has been shown to be required for T6SS function as deletion of *tssA*^{S} in *B. cenocepacia* results in loss of secretion. T6SS activity can be restored to the mutant through genetic complementation with a plasmid expressing TssA^{S}, i.e. pBBR1MCS-*tssA*^{S} (Spiewak 2015). The *B. cenocepacia* H111-*tssA*^{S}::Tp mutant strain was constructed by insertional inactivation of *tssA*^{S} with a trimethoprim resistance cassette followed by introducing the mutated gene into the T6SS locus by allelic replacement (Spiewak 2015). As TssA^{S} and TssA^{EI} share similarities in their N-terminal domains and interact with similar T6SS core components, the ability of hybrid proteins containing fused TssA^{S} and TssA^{EI} domains in restoring the T6SS function in the H111-*tssA*^{S}::Tp mutant strain was investigated. One fusion protein (named Hybrid 1) contained the conserved ImpA-rel_N domain region from TssA^{E} (L51-P108) fused to TssA^{S} lacking its conserved ImpA-rel_N domain region (W56-P113), and the other (named Hybrid 2) included the conserved ImpA-rel_N domain region, the remainder of the predicted N-terminal helical region and the predicted linker region from TssA^{EI} fused to TssA$^{S}_{CTD}$ (Figure 6.35). In parallel, full-length TssA^{E} was also included in the genetic complementation assay. TssA^{S} sequences were derived from BCAL0348 and TssA^{EI} sequences were derived from AHA1844.

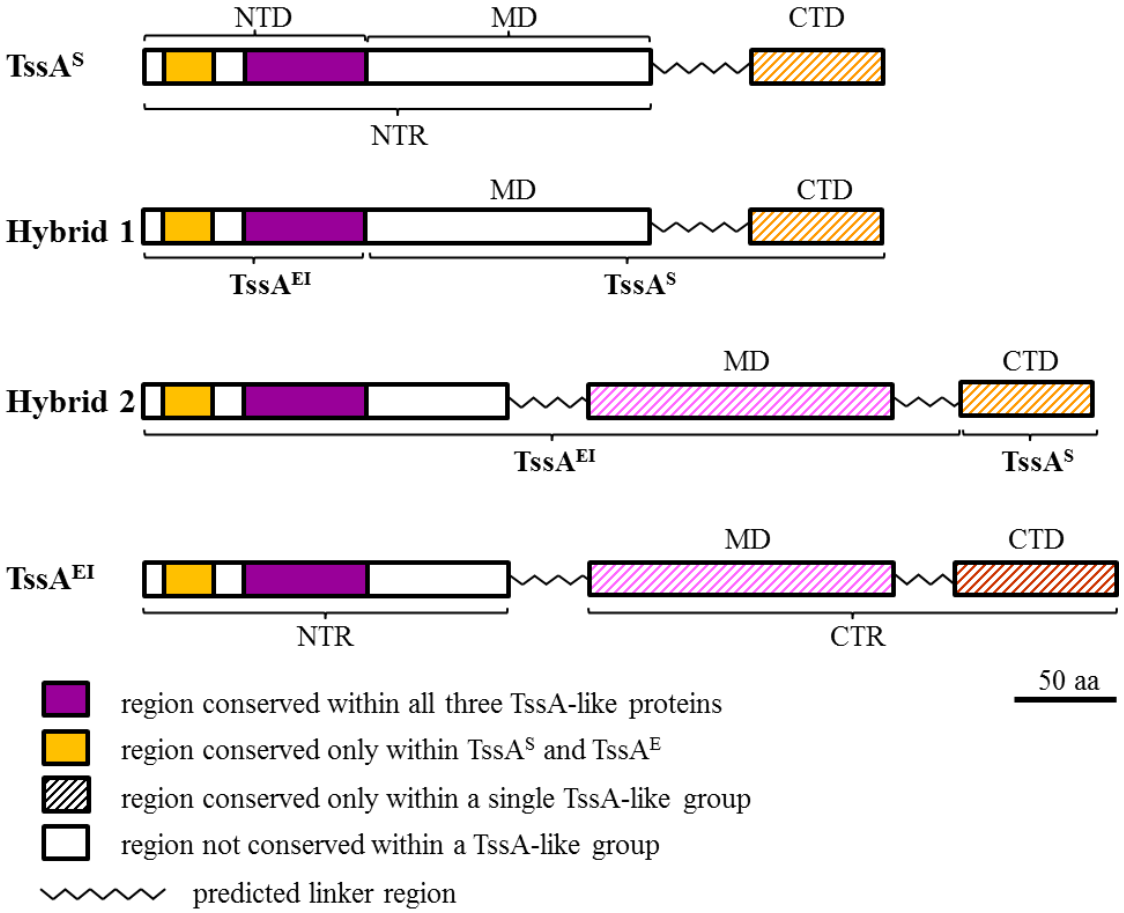


Figure 6.35 Schematic drawing of the TssA^{EI}-TssA^{S} hybrid proteins. NTD, N-terminal domain; MD, middle domain; CTD, C-terminal domain; NTR, N-terminal region; CTR, C-terminal region.

6.4.1 Construction of complementation plasmids

For genetic complementation of the H111-*tssA*^{S}::Tp mutant strain, a broad host-range plasmid, pBBR1MCS, containing the coding sequences of Hybrid 1, Hybrid 2 and *tssA*^{EI} genes were constructed separately. In order to improve the gene expression, a phage T7 Shine-Dalgarno sequence was included in the forward primer (pBBR1MCS.AHA1844.KpnI.for) used for PCR amplification. The pBBR1MCS plasmid has a *lac* promoter and specifies resistance to chloramphenicol.

To construct pBBR1MCS-Hybrid 1, SOE-PCR was carried out by separate amplification of two DNA fragments encoding the N-terminal region of TssA^{EI} from M1 to P108, that incorporates the conserved ImpA-rel_N domain region from TssA^{EI} (regarded as fragment A), and amino acid sequences L114-S373 of TssA^{S} which lacks its conserved ImpA-rel_N domain region (regarded as fragment B). Amplification of fragment A and fragment B involved 2 pairs of primers, pBBR1MCS.AHA1844.KpnI.for and pBBR1MCS.AHA1844.NTD.SD1.inter_rev, and pBBR1MCS.TssA.NTD.SD2.inter_for and TssA.HindIII.rev, respectively. Primers pBBR1MCS.AHA1844.NTD.SD1.inter_rev and pBBR1MCS.TssA.NTD.SD2.inter_for include 40 complementary template bases allowing fusion of fragments A and B in the second PCR. Two previously constructed plasmids pUT18C-*tssA*^{S} and pUT18C-*tssA*^{EI} were used as the source of *tssA*^{S} and *tssA*^{EI}, respectively. Following electrophoresis in a 1% agarose gel and gel purification, fragments A and B were fused by another round of PCR using pBBR1MCS.AHA1844.KpnI.for and TssA.HindIII.rev primers with the two DNA fragments as template. The products of the second round PCR were digested with *Acc*65I and *Hin*dIII restriction enzymes which recognized restriction sites in the primers, and ligated into pBBR1MCS plasmid that was digested with the same restriction enzymes. *E. coli* strain JM83 was transformed with the ligation mixture and colonies were grown on LB plates containing chloramphenicol (25 μg/ml), X-gal (40 μg/ml) and IPTG (100 μM). Colonies containing recombinant plasmids were identified by white appearance. Following plasmid miniprep, the nucleotide sequence of recombinant plasmids containing the Hybrid 1 gene was verified by DNA sequencing.

The procedure for construction of pBBR1MCS-Hybrid 2 was intended to be the same but using primers pBBR1MCS.AHA1844.NTD.inter_rev substituted for pBBR1MCS.AHA1844.NTD.SD1.inter_rev, and pBBR1MCS.TssA.CTD.interfor

substituted for pBBR1MCS.TssA.NTD.SD2.inter_for, which resulted in amplification of two DNA fragments encoding the ImpA-rel_N domain and the remainder of the predicted N-terminal helical region and the predicted linker region from $TssA^{EI}$ (M1-N230) fused to $TssA^{S}_{CTD}$ (I303-S373).

The pBBR1MCS-*tssA*^{EI} plasmid was also constructed. The PCR product of *tssA*^{EI} following amplification using primers pBBR1MCS.AHA1844.KpnI.for and pACYC-AHA1844.rev primers were digested with *Acc*65I and *Hin*dIII restriction enzymes which recognized restriction sites in the primers, and ligated into pBBR1MCS plasmid that was digested with the same restriction enzymes. Selection and confirmation of the correct recombinant plasmid were carried out as stated above for pBBR1MCS-Hybrid 1.

6.4.2 Complementation of TssD secretion in *B. cenocepacia* H111 *tssA*^{S}-deficient mutant strain

To investigate whether the absence of secreted TssD in the H111-*tssA*^{S}::Tp mutant strain can be restored by genetic complementation with genes encoding hybrid TssA subunits that contain N-terminal sequences derived from $TssA^{EI}$, pBBR1MCS-Hybrid 1, pBBR1MCS-Hybrid 2 and pBBR1MCS-*tssA*^{EI} along with the pBBR1MCS empty vector and pBBR1MCS-*tssA*^{S} were transferred separately to the H111-*tssA*^{S}::Tp mutant strain by conjugation using *E. coli* S17-1(λpir) as a donor (Section 2.3.15.2). Extracellular proteins present in liquid cultures of the *B. cenocepacia* exconjugants were extracted as described in Section 2.5 and the presence of TssD was analysed by SDS-PAGE followed by immuno-detection (western blotting) using anti-TssD antibody. The results were compared to the wild-type and *tssA*^{S}::Tp mutant strains lacking a plasmid.

The results showed that the secretion of TssD was not restored in the H111-*tssA*^{S}::Tp mutant by genetic complementation with pBBR1MCS-Hybrid 1, pBBR1MCS-Hybrid 2 or pBBR1MCS-*tssA*^{EI} (Figure 6.36). Although a small amount of TssD was present in the culture supernatant when complementing of the *tssA*^{S}::Tp mutant by the pBBR1-based plasmids (pBBR1MCS-Hybrid 1, pBBR1MCS-Hybrid 2 or pBBR1MCS-*tssA*^{EI}), the level of secreted TssD was similar to that of complementing with the pBBR1MCS empty vector. Consistent with a previous study (Spiewak 2015),

TssD secretion by the *tssA*^{S}::Tp mutant complemented with pBBR1MCS-*tssA*^{S} was restored to a similar level compared to that of the WT strain.

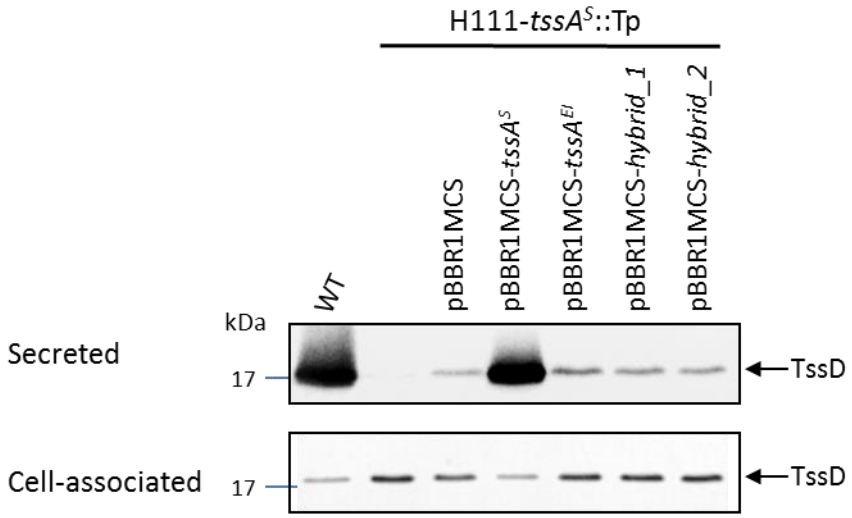


Figure 6.36 Genetic complementation of *B. cenocepacia* H111-*tssA*^{S}::Tp mutant by pBBR1MCS-based plasmids expressing hybrid TssA proteins and TssA^{EI}. pBBR1MCS-Hybrid 1, pBBR1MCS-Hybrid 2, pBBR1MCS-*tssA*^{EI} and pBBR1MCS-*tssA*^{S} complementation plasmids and pBBR1MCS empty vector were transferred into the *B. cenocepacia* H111-*tssA*^{S}::Tp mutant by conjugation. Extracellular proteins were then extracted from late exponential grown liquid cultures by DOC-TCA precipitation, electrophoresed in 12% SDS-PA gels and immuno-detected with custom anti-TssD polyclonal antibody for the presence of TssD. WT and *tssA*^{S} mutant strains were also included as controls. Cell-associated proteins were also analysed. EZ-Run™ Prestained Rec Protein Ladder was used as protein marker. The arrows indicate the expected locations of TssD (~18.4 kDa) based on its size.

6.5 Discussion

TssA^{EI} derived from *A. hydrophila* (AHA1844) was shown to be involved in interactions with many other *B. cenocepacia* T6SS subunits similar to TssA^{S} by the bacterial two-hybrid analysis (Chapter 3), and the majority of these interactions were subsequently confirmed *in vitro* (Chapter 4). On the basis alone, TssA^{EI} is proposed to have functional similarity to TssA^{S}. However, while this work was in progress, the structure and function of EAEC TssA^{E} (Ec042_4540) was determined (Zoued et al. 2016). Ec042_4540 is in the same class of TssA-like proteins as AHA1844 but belongs to a different TssA^{E} subclass (i.e. TssA^{EII}). TssA^{EII} subclass orthologues have a C-terminal extension (~40 amino acids), the majority of which are predicted to be comprised of two alpha-helices by Psipred (results not shown). Moreover, AHA1844 forms a decamer whereas Ec042_4540 assembles into a dodecamer. In the following discussion, TssA^{EI} and TssA^{EII} refer to *A. hydrophila* TssA^{E} and EAEC TssA^{E}, respectively.

6.5.1 Domain organisation of TssA^{EI}

In general, TssA^{EI} orthologues are 470-530 amino acids in length, which is larger than TssA^{S} (340-375 amino acids). TssA^{EI} orthologues are also distinguished from TssA^{S} by distinct sequences of the CTD, although the CTDs from the TssA^{EI} subclass are similar in length with those from TssA^{S} orthologues. The multiple sequence alignment shows there is a conserved region in all TssA-like proteins of ~60 amino acids in length corresponding to L51-P108 in *A. hydrophila* TssA^{EI} (AHA1844) which is reffered to as an ImpA-related N-terminal domain.

The amino acid sequence alignment of TssA^{EI} orthologues present in other Gram-negative bacteria predicts two non-conserved regions, i.e. P130-N230 and Q375-G401 of TssA^{EI} which may correspond to interdomain linkers. The secondary structure prediction by Psipred largely supports the amino acid sequence alignment, except two predicted helices (H7 and H8) occur in the A138-A188 region of TssA^{EI}. The results from limited proteolytic analysis showed susceptibility to cleavage as largely occurring in Q158-N230 and Q375-G401 regions of TssA^{EI}. Therefore, the results of these analyses indicated that similarly to TssA^{S}, TssA^{EI} contains a large

N-terminal domain (M1-A157), a middle domain (D231-L374) and a small C-terminal domain (I402-K478).

The domain organisation of EAEC TssA^{EII} analysed by Zoued et al. (2016) is in agreement with that of *A. hydrophila* TssA^{EI} in this study, where the NTD1 (residues M1-L220) of EAEC TssA^{EII} corresponds to *A. hydrophila* TssA$^{EI}_{NTD}$ and is separated from NTD2 by an interdomain linker. NTD2 (residues S221-M377) of EAEC TssA^{EII} corresponds to the MD of *A. hydrophila* TssA^{EI}, and CTD of EAEC TssA^{EII} corresponds to the *A. hydrophila* TssA$^{EI}_{CTD}$ with a 15 amino acid interdomain linker between MD and CTD. However, the CTD of EAEC TssA^{EII} is 54 amino acids longer than that of *A. hydrophila* TssA^{EI}.

6.5.2 TssA^{EI} oligomerises into a high molecular mass complex with its CTD serving as the oligomerisation module

TssA^{EI} has been shown to self-interact *in vivo* by BACTH assay, and indicated to form a large complex of over ~2 MDa by SEC (Ahmad 2013). This observation is supported by analysis in this study, where TssA^{EI} was also shown to be a multimer but with a much smaller estimated MW by SEC, ~700 kDa (monomeric TssA^{EI} is ~52 kDa). A more accurate molecular weight estimation by SEC-MALS suggested TssA^{EI} is slightly less than 600 kDa, but this is likely to be overestimated due to the influence of aggregation material. Later investigations into the oligomerisation ability of TssA$^{EI}_{NTD}$, TssA$^{EI}_{MD}$ and TssA$^{EI}_{CTD}$ by SEC indicated NTD exists as a monomer, whereas, MD and CTD form dimers and decamers, respectively. SEC-MALS results suggest a consistent oligomerisation estimation for TssA$^{EI}_{MD-CTD}$ and TssA$^{EI}_{CTD}$, i.e. both of them exist as decamers. Therefore, TssA$^{EI}_{CTD}$ is responsible for the decamerisation of the MD-CTD resulting in the large complex formed by full-length TssA^{EI}. The results support the idea that the SEC-MALS estimation for the oligomerisation status of full-length TssA^{EI} (11 protomers per oligomer) was slightly overestimated, and that the true oligomerisation status is a decamer. In addition, similar to the role of the C-terminal helix (H15) of TssA^{S} which is responsible for the large complex formation of TssA^{S}, the C-terminal helix of TssA^{EI} (H20, numbering in Psipred) is shown to be responsible for the decamerisation of TssA^{EI}.

The $TssA^{EI}$ complex was visualised by TEM to form a ring-like structure containing a distinct central hole similar as $TssA^{S}$, but the ring-like structure of $TssA^{EI}$ is not surrounded by discrete $TssA^{S}_{NTD}$ projections as was $TssA^{S}$. Moreover, instead of a well-defined inner ring as in the case of $TssA^{S}_{CTD}$, the ring-like structure of $TssA^{EI}$ has a broader thickness with an irregular outer edge. The overall dimension of the $TssA^{EI}$ ring-like structure is 232 Å, which is about half of that of the $TssA^{S}$ complex (~450 Å). Without NTD, $TssA^{EI}_{MD-CTD}$ was observed not to form rings five-pointed stars with a central lumen. Investigation of the $TssA^{EI}_{CTD}$ complex shows the five-pointed star shape was not as obvious as that of $TssA^{E}_{MD-CTD}$, suggesting the MD locates to the periphery of the complex and extends the arms of the 'star'. $TssA^{EI}_{MD-CTD}$ and $TssA^{EI}_{CTD}$ complexes have a 5-fold symmetry.

Although NTD2 of EAEC $TssA^{EII}$ was shown to be a dimer, in agreement with the dimeric MD of *A. hydrophila* $TssA^{EI}$, EAEC $TssA^{EII}$ and its CTD exist as dodecamers, not decamers (Zoued et al. 2016). The inconsistency of the oligomerisation state of these two $TssA^{E}$ subclasses may be caused by the C-terminal extension of EAEC $TssA^{EII}$. Consistent with *A. hydrophila* $TssA^{EI}$, the CTD of EAEC $TssA^{EII}$ is also responsible for the dodecamerisation of the full length protein (Zoued et al. 2016). Moreover, the EAEC $TssA^{EII}$ complex is observed to have a starfish-shape by EM, but with 6 arms and ~300 Å dimension (Zoued et al. 2016).

6.5.3 X-ray crystallographic analysis of $TssA^{EI}_{MD}$ and $TssA^{EI}_{CTD}$

The structures of $TssA^{EI}_{MD}$ and $TssA^{EI}_{MD-CTD}$ were solved at 2.1 Å and 3.25 Å resolution, respectively. However, no crystal hints of isolated $TssA^{EI}_{NTD}$ and $TssA^{EI}_{CTD}$ have been gained so far. Consistent with the EM observation that the five-pointed star shape is more obvious in the $TssA^{EI}_{MD-CTD}$ than that of the $TssA^{EI}_{CTD}$, $TssA^{EI}_{CTD}$ forms an overall pentameric star with the MD locating to the periphery of the complex. The dimeric MDs most likely create new arms between each pair of CTD arms, supporting model B of the merged EM images of symmetrised $TssA^{EI}_{MD-CTD}$ and $TssA^{EI}_{CTD}$ (Figure 6.30 B). $TssA^{EI}_{MD}$ also dimerises when it was crystallised on its own, and shares high structural similarity with the NTD2 domain of EAEC $TssA^{EII}$ (Figure 6.37 A), which is in agreement with the conserved amino acid sequence present in the corresponding

regions of $TssA^{EI}$ and EAEC $TssA^{EII}$ (Figure 1.11). In the crystal structure of $TssA^{EI}_{MD\text{-}CTD}$, the dimeric MD extended from the CTD at different angles most probably due to the conformational flexibility of the interdomain linker (~13 amino acids) present between these two domains. However, how MD can be crystallised along with CTD when they are connected by a flexible linker is still a question to be answered.

In the $TssA^{EI}_{MD\text{-}CTD}$ structure, $TssA^{EI}_{CTD}$ packs into a single layer, the dimension of which is largely in agreement with the measurement from EM. In addition, the dimensions of $TssA^{EI}_{CTD}$ are similar to that of EAEC $TssA^{EII}_{CTD}$ in terms of its outer and inner dimensions. However, the packing pattern of $TssA^{EI}_{CTD}$ is different from that of $TssA^{EII}_{CTD}$ which features two stacked hexaflexagons ~30 Å thick (Figure 6.37 B) (Zoued et al. 2016). This is likely due to the presence of two additional helices in the C-terminal extension of the $TssA^{EII}_{CTD}$ orthologues (Figure 1.11) locating to the exterior of the complex. The conserved WEP motif in all $TssA^{E}$ orthologues mainly locates in the loop between H12 and H13 in $TssA^{E}_{CTD}$. However, whereas in $TssA^{EI}$ it contributes to the inter-docking of arms in the pentameric star, in $TssA^{EII}_{CTD}$ it is involved in intra-ring interactions that allow the stacking of hexamers (Zoued et al. 2016).

The inner lumen of $TssA^{EI}_{CTD}$ is ~30 Å, similar to that of $TssA^{EII}_{CTD}$, corresponding to the primed state (extended sheath). It has been proposed that $TssA^{EII}_{CTD}$ undergoes large conformational changes in order to open a large central channel of ~90 Å to accommodate the tail tube (composed of stacked hexameric TssD rings (85-90 Å)) by displacing the six wedges to the exterior (Zoued et al. 2016). Similarly, helical reorientation of the $TssA^{EI}_{CTD}$ pentameric star may result in formation of a ring with a lumen of increased circumference (90-100 Å). On the other hand, no structural similarity between $TssA^{EI}_{CTD}$ and $TssA^{S}_{CTD}$ has been identified. They have different packing patterns and dimensions, and appear to correspond to different states of the T6SS apparatus. More crystals of $TssA^{E}_{MD\text{-}CTD}$ and $TssA^{EI}_{CTD}$ can be grown and analysed to compare with the current structure regarding their packing pattern and possibly a structure of the open (firing) form of the molecule can also be obtained. In addition, the difference between orthologues of these two $TssA^{E}$ subclasses needs to be further investigated and whether the observations from two representative strains discussed here are common features of the different subclasses also needs to be

experimentally addressed. Indeed, it may also be possible that other subclasses of TssA^{E} may exist with different subunit number and packing.

6.5.4 Location and function prediction of TssA^{EI} in the T6SS

As the current model for the mechanism of the T6SS is considered to be similar to that of DNA injection by contractile bacteriophage in that they both have a cell puncturing machinery, the location of TssA^{EI} in the overall T6SS model is worth being discussed. EAEC TssA^{EII} has recently been shown to have a dynamic role in T6SS assembly (Zoued et al. 2016). It is recruited to the TssJLM membrane chamber complex at the early stage of the T6SS assembly, initiating and guiding the polymerisation of the TssD tail tube and the TssBC tail sheath probably by incorporating TssD and TssBC building blocks (Zoued et al. 2016). EAEC TssA^{EII} locates within the baseplate initially but remains at the baseplate-distal end of the tubules during elongation and is released after sheath contraction (Zoued et al. 2016). EAEC TssA^{EII} is proposed to be analogous to the bacteriophage T4 gp15/gp3 tail terminator complex as they have very similar dynamics and dimensions (Pell et al. 2009b; Fokine et al. 2013; Zoued et al. 2016). The tail terminator complex of bacteriophage T4 adopts conformational changes, i.e. closed and open, associated with the extended and contracted sheath conformations (Leiman and Shneider 2012). The closed central channel of the terminator complex prevents DNA being released from the capsid in the extended tail (Leiman and Shneider 2012). Similarly, TssA^{EI} is likely capable to bear conformational change (i.e. open and closed) during extended and contracted sheath conformations. Although there is no capsid-like structure in the T6SS, the conformational modifications of TssA^{EI} is probably related to incorporation of TssD or TssBC building blocks during tail polymerisation. In addition, TssA^{EI} is involved in similar interactions with many other T6SS subunits that EAEC TssA^{EII} interact with, and this supports the dynamics and role in guiding tail tube/sheath elongation of *A. hydrophila* TssA^{EI}.

The introduction into a *B. cenocepacia tssA*^{S}-deficient mutant strain of two pBBR1-based plasmids encoding hybrid domain fusions and another plasmid encoding TssA^{EI} did not successfully restore TssD secretion (a hallmark of an active T6SS). Although there was a small amount of TssD secreted in each case, this was similar to

that of the H111 *tssA*^{S}-deficient mutant complemented with the pBBR1MCS empty vector. Therefore, no convincing conclusion can be drawn, but it is most likely that the T6SS activity was not restored. One possibility for the lack of complementation is due to the fact that TssA^{EI} did not interact with *B. cenocepacia* TssD *in vitro* (Section 4.4.2.3). However, as this interaction involves TssA$_{CTD}$, and the CTD present in the hybrid proteins was derived from TssA^{S} this may not be the reason. Another possibility could be that these genes were not expressed sufficiently from the pBBR1-based plasmids although a Shine-Dalgarno sequence has been incorporated in few bases upstream of the *tssA*^{EI} start codon. The expression of TssA^{EI} can be investigated in an *A. hydrophila tssA*^{EI} mutant, however, it has not been constructed. Also, the ability of the two hybrid fusion proteins to form ring-like oligomers still needs to be experimentally investigated. Although TssA^{EI} may be an orthologue of TssA^{S}, the difference in its 5-fold symmetry from 16-fold symmetric TssA^{S} could be another reason for the failure of TssA^{EI} to restore the T6SS activity in the *tssA*^{S}-deficient mutant following introduction of pBBR1MCS-*tssA*^{EI}.

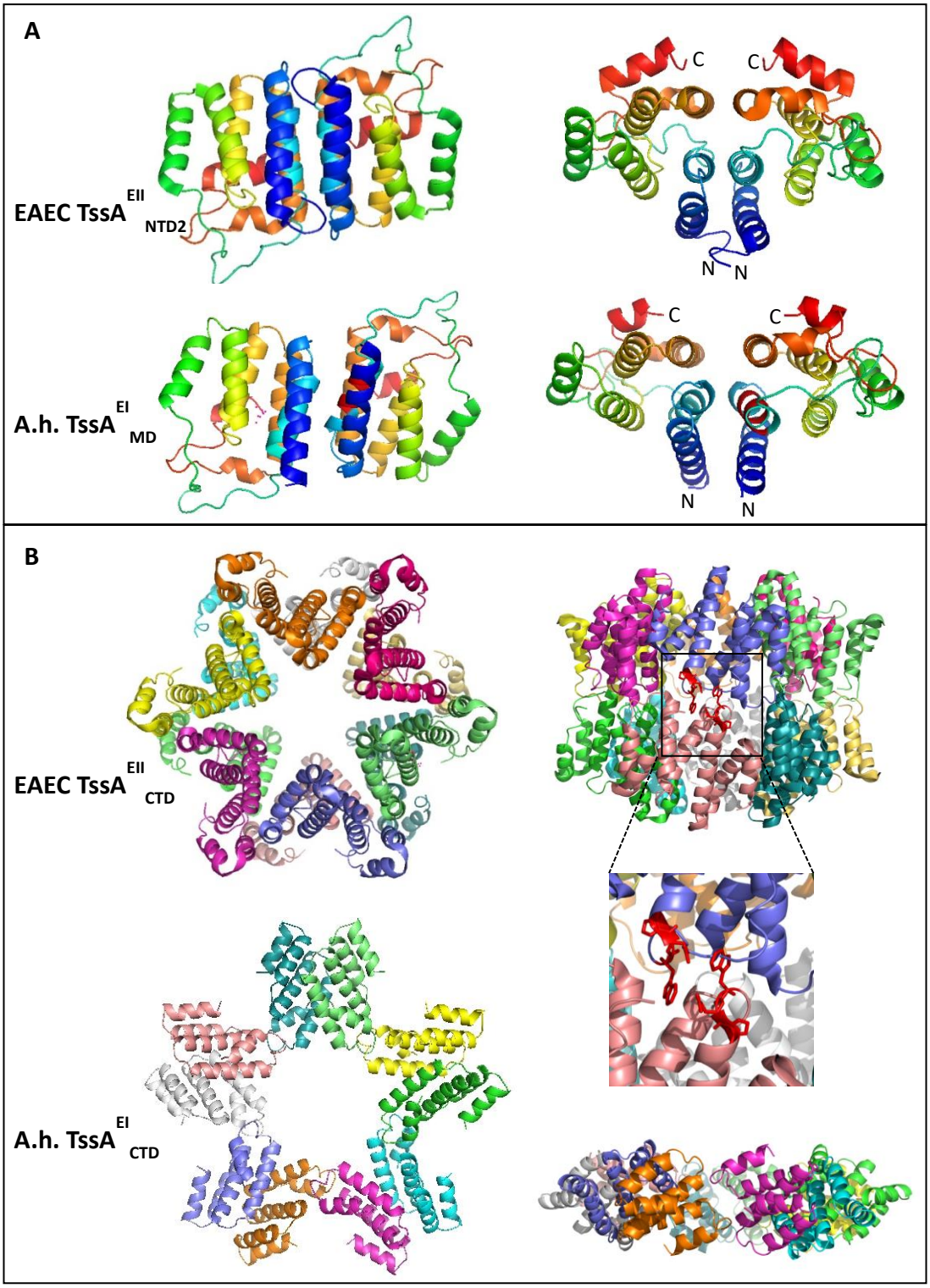


Figure 6.37 Structures of EAEC TssA^{EII} NTD2 and CTD. A. The structure of TssA$^{EII}_{NTD2}$ compared to TssA$^{EI}_{MD}$ (lower panel). Chains are coloured in rainbow from blue to red, corresponding to N- and C-terminus of the structure. **B**. Top (left panel) and side (right panel) views of the TssA$^{EII}_{CTD}$ structure compared to TssA$^{EI}_{CTD}$. Chains are in different colours. The WEP motif of TssA$^{EII}_{CTD}$ structure is shown in red sticks. Black rectangle encloses chains that are shown enlarged underneath. PDB 4YO3 and 4YO5 (Zoued et al. 2016).

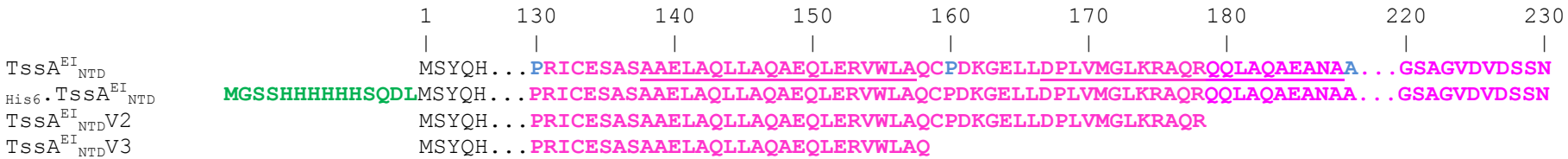


Figure 6.38 TssA$^{EI}_{NTD}$ derivatives used in this study. The amino acid sequence of the *A. hydrophila* TssA$^{EI}_{NTD}$ and the NTD-MD interdomain linker (M1-N230) is shown at the top of the figure with two regions (P6-L129 and G190-S219) omitted (represented by ...). The 72 amino acids of the interdomain linker as predicted by amino acid sequence alignment is shown in magenta font. The locations of the N-terminus of the predicted interdomain linker based on amino acid sequence alignment (P130), proteolytic cleavage (C159) and Psipred v3.3 secondary structure prediction (A189) are shown in bold blue font. Amino acids that are predicted to form α-helices according to Psipred are underlined. Sequences below correspond to TssA$^{E}_{NTD}$ derivatives expressed in this study, including non-native tag sequences where applicable (green font).

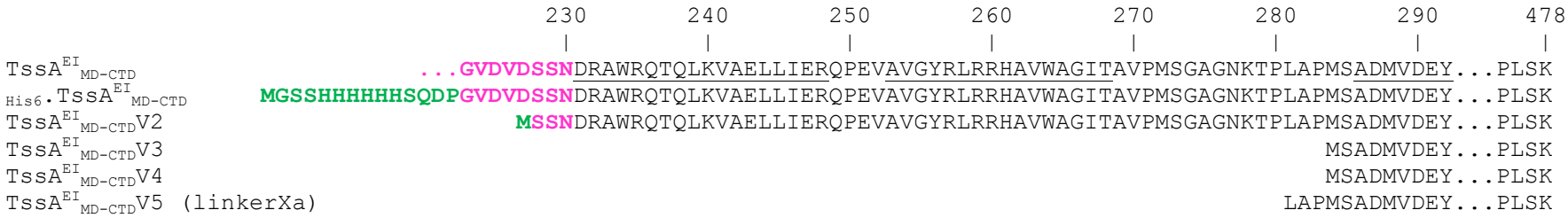


Figure 6.39 TssA$^{EI}_{MD\text{-}CTD}$ derivatives used in this study. The *A. hydrophila* TssA$^{EI}_{MD\text{-}CTD}$ amino acid sequence (D231-K478) is shown at the top of the figure with the R293-A474 region omitted (represented by ...). The C-terminal end of the 72 amino acid NTD-MD interdomain linker as predicted by amino acid sequence alignment, Psipred v3.3 secondary structure prediction and proteolytic cleavage (consistent prediction obtained) is also shown in magenta font. Amino acids that are predicted to form α-helices according to Psipred are underlined. Sequences below correspond to TssA$^{E}_{MD\text{-}CTD}$ derivatives expressed in this study, including non-native tag sequences where applicable (green font).

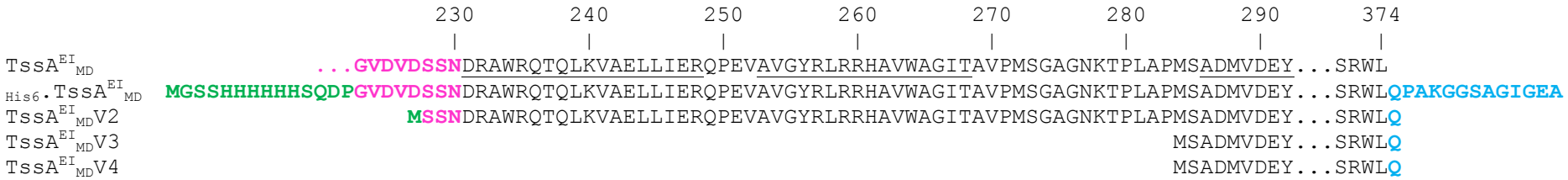


Figure 6.40 TssA$^{EI}_{MD}$ derivatives used in this study. The *A. hydrophila* TssA$^{EI}_{MD}$ amino acid sequence (D231-L374) is shown at the top of the figure with the R293-C370 region omitted (represented by ...). The C-terminal end of the NTD-MD interdomain linker (magenta), and the N-terminal end of the MD-CTD interdomain linker (blue) as predicted by amino acid sequence alignment, Psipred v3.3 secondary structure prediction and proteolytic cleavage are also shown. Amino acids that are predicted to form α-helices according to Psipred are underlined (not all predicted helices are shown). Sequences below correspond to TssA$^{EI}_{MD}$ derivatives expressed in this study, including non-native tag sequences where applicable (green font).

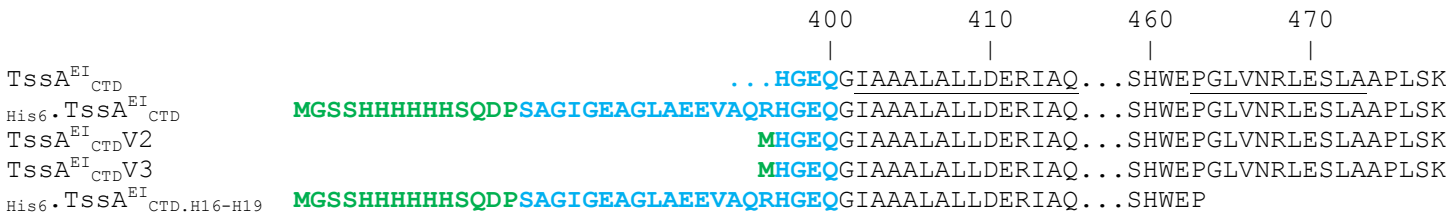


Figure 6.41 TssA$^{EI}_{CTD}$ derivatives used in this study. The *A. hydrophila* TssA$^{EI}_{CTD}$ amino acid sequence (I402-K478) is shown at the top of the figure with the L416-L458 r`egion omitted (represented by ...). The C-terminal end of the MD-CTD interdomain linker as predicted by amino acid sequence alignment, Psipred v3.3 secondary structure prediction and proteolytic cleavage is shown in blue font. Amino acids that are predicted to form α-helices according to Psipred are underlined. Sequences below correspond to TssA$^{EI}_{CTD}$ derivatives expressed in this study, including non-native tag sequences where applicable (green font).

Table 6.1 Summary table of TssA^{E} and TssA^{E} domain protein overexpression in *E. coli*

Protein	Residues	Plasmid	Host strain and induction condition[a]	Overexpression and solubility
$_{His6}$.TssA^{EI}	1-478	pACYCDuet-His$_6$.*tssA^{EI}*	BL21(λDE3), 30^{o}C with 1 mM IPTG	Overproduced and soluble
$_{His6}$.TssA$^{EI}_{NTD}$	1-229	pACYCDuet-His$_6$.*tssA$^{EI}_{NTD}$*	BL21(λDE3), 1 mM IPTG at 30^{o}C	Overproduced and soluble
TssA$^{EI}_{NTD}$V2	1-178	pACYCDuet-*tssA$^{EI}_{NTD}$*V2	BL21(λDE3), 1 mM IPTG at 30^{o}C	Overproduced and soluble
TssA$^{EI}_{NTD}$V3	1-158	pACYCDuet-*tssA$^{EI}_{NTD}$*V3	BL21(λDE3), ≥0.1 mM at 30^{o}C	Overproduced but insoluble
$_{His6}$.TssA$^{EI}_{MD-CTD}$	223-478	pACYCDuet-His$_6$.*tssA$^{EI}_{MD-CTD}$*	BL21(λDE3), 1 mM IPTG at 37^{o}C	Overproduced and soluble
TssA$^{EI}_{MD-CTD}$V2	228-478	pACYCDuet-*tssA$^{EI}_{MD-CTD}$*V2	BL21(λDE3), 1 or 0.1 mM IPTG at 30^{o}C	Overproduced and soluble
TssA$^{EI}_{MD-CTD}$V3	284-478	pACYCDuet-*tssA$^{EI}_{MD-CTD}$*V3	BL21(λDE3), C41(λDE3), C43(λDE3)	No overproduction
TssA$^{EI}_{MD-CTD}$V4	284-478	pACYCDuet-*tssA$^{EI}_{MD-CTD}$*V4	BL21(λDE3), 1 mM IPTG at 30/37^{o}C 0.5/0.1 mM 30^{o}C	Overproduced but insoluble
			C41(λDE3) 1 mM IPTG at 37^{o}C	Overproduced but insoluble
			C43(λDE3) 1 mM IPTG at 30/37^{o}C	no expression
TssA$^{EI}_{MD-CTD}$V5 ($_{His6}$.TssA$^{EI}_{linkerXa}$)	281-478	pACYCDuet-His$_6$.linkerXa.*tssA^{EI}*	BL21(λDE3) ≥0.1 mM IPTG at 30^{o}C	Overproduced but insoluble
			C41(λDE3)/C43(λDE3) 1 mM IPTG at 30/37^{o}C	no expression
$_{His6}$.TssA$^{EI}_{MD}$	223-387	pACYCDuet-His$_6$.*tssA$^{EI}_{MD}$*	BL21(λDE3), 0.1-1 mM IPTG at 30^{o}C	Overproduced and soluble
TssA$^{EI}_{MD}$V2	228-375	pACYCDuet-*tssA$^{EI}_{MD}$*V2	BL21(λDE3), 0.1/0.5 mM IPTG at 37^{o}C	Overproduced but insoluble
			BL21(λDE3) 0.5 mM IPTG at 30^{o}C	no expression
TssA$^{EI}_{MD}$V3	284-375	pACYCDuet-*tssA$^{EI}_{MD}$*V3	BL21(λDE3) 1 mM IPTG at 30/37^{o}C	no expression

TssA$^{EI}_{MD}$V4	284-375	pACYCDuet-*tssA*$^{EI}_{MD}$V4	BL21(λDE3)/C41(λDE3)/C43(λDE3) 1 mM IPTG at 30/37^{o}C	no expression
$_{His6}$.TssA$^{EI}_{CTD}$	381-478	pACYCDuet-His$_6$.*tssA*$^{EI}_{CTD}$	BL21(λDE3) 1 mM IPTG at 37^{o}C	Overproduced and partially soluble
TssA$^{EI}_{CTD}$V2	397-478	pETDuet-*tssA*$^{EI}_{CTD}$V2	BL21(λDE3)/C41(λDE3)/C43(λDE3) 1 mM IPTG 30/37^{o}C	no expression
TssA$^{EI}_{CTD}$V3	397-478	pACYCDuet-*tssA*$^{EI}_{CTD}$V3	BL21(λDE3) 1 mM IPTG at 37^{o}C	Overproduced and soluble
$_{His6}$.TssA$^{EI}_{CTD.H16-H19}$	381-463	pACYCDuet-His$_6$.*tssA*$^{EI}_{CTD}$.H16-H19	BL21(λDE3) 1 mM IPTG at 37^{o}C	Overproduced and soluble

[a], not all the tested induction conditions are listed. Full information is provided in the relevant section.

Chapter 7 Two hybrid analysis of interactions between different T6SS subunits (excluding TssA)

7.1 Introduction

The analysis of interacting network among T6SS subunits provides information as tofacilitate reveal how the T6SS assembles and functions. Some of the interaction analyses have been carried out by two previous group members (Shastri 2011; Ahmad 2013) and evidence for interactions among T6SS subunits can also be found in many other publications. As some of the two-hybrid plasmids used in the previous studies (Shastri 2011; Ahmad 2013) were defective, the interaction analysis involving defective two-hybrid plasmids was repeated with rectified plasmid constructs (rectification as described in Section 3.2) and the three-hybrid analysis of interactions between TssE and TssF in the presence of TssG was also assayed.

7.2 Analysis of interactions between TssC and all the other T6SS subunits

Three previously constructed two-hybrid plasmids encoding TssC fusions, i.e. pKNT25-*tssC*, pUT18-*tssC* and pUT18C-*tssC* (Shastri 2011) were not made correctly. These three plasmids were reconstructed as described in Section 3.2.1. The interaction analysis between TssC and all the other T6SS subunits except $TssA^S$ (the latter is described in Section 3.3.3) was conducted using BACTH assay with three rectified two-hybrid plasmids and the analysis involving pKT25-*tssC* was repeated. The results showed that seven out of eight possible pairwise combinations of TssC and TssB gave rise to a maltose positive phenotype, suggesting an interaction between them (Figure 7.1 A). In addition, the combination of T25-TssH with T18-TssC also yielded a positive maltose phenotype (Figure 7.1 B). All the remaining combinations gave rise to a negative maltose phenotype (results not shown). However, this is not consistent with a previous study (Shastri 2011), where a positive maltose phenotype was observed between T25-TssC and TssF-T18.

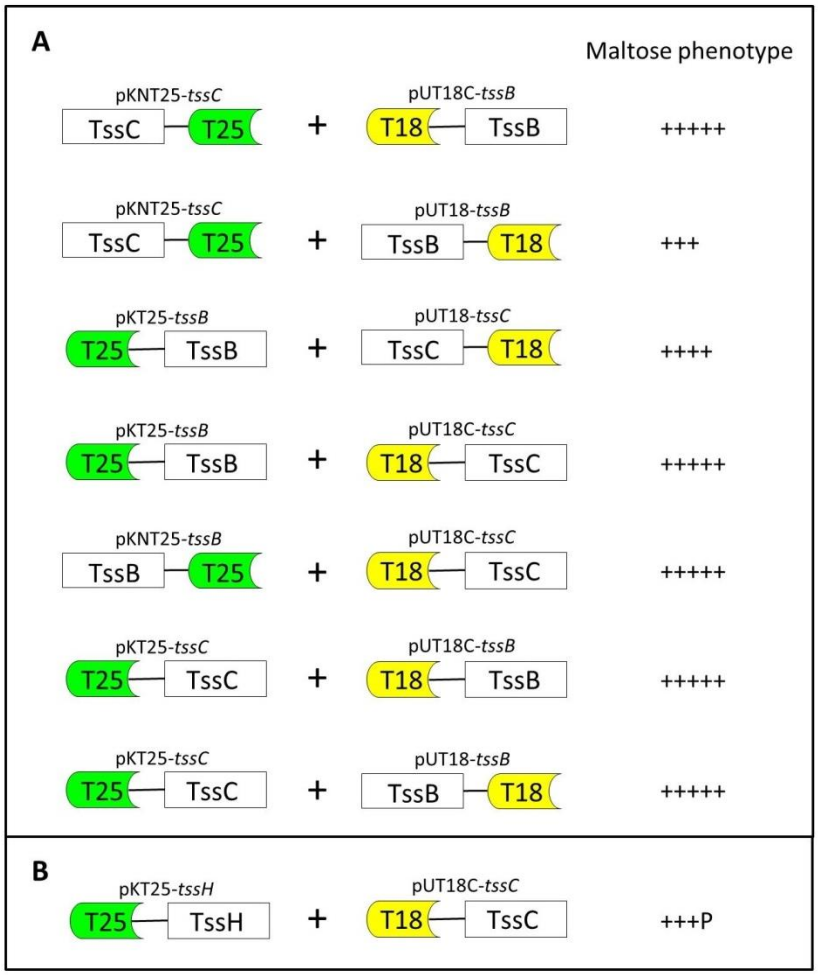


Figure 7.1 BACTH analysis of the interactions between TssC and TssB, TssH on MacConkey-maltose agar. Compatible BACTH plasmids encoding the CyaA T25 or T18 fragments fused to TssC and TssB/TssH were introduced into *E. coli* strain BTH101 and transformants were grown on MacConkey-maltose agar at 30℃. Maltose phenotype was observed after 120 hours incubation. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated next to the corresponding diagram as crosses according to the classification described in Table 3.2; 'P', patchy; combinations that yielded a negative maltose phenotype were not shown. **A**. Combinations of TssC and TssB fusion proteins that gave rise to a maltose-positive phenotype. **B**. Combination of TssC and TssH fusion proteins that gave rise to a maltose-positive phenotype.

7.3 Analysis of interactions between TssE and TssF in the presence or absence of TssG

In the bacterial genome that possesses T6SS, high levels of co-organisation between *tssE* and *tssF*, and *tssF* and *tssG* genes were demonstrated in 87% and 97% of the T6SS gene clusters (Boyer et al. 2009; Brunet et al. 2015). Therefore, the interaction between TssE and TssF in the presence and absence of TssG were investigated.

All eight possible pairwise combinations between TssE and TssF were assayed using the BACTH system. The result showed that one of the combinations, TssE-T25 and T18-TssF gave rise to a maltose positive phenotype (Figure 7.2) and all the remaining combinations showed a negative maltose phenotype. However, one of the control combinations, pKNT25 and pUT18C-*tssF*, gave rise to pink/red colonies as described in Section 3.3.6. This causes the positive maltose phenotype observed for the combination of TssE-T25 with T18-TssF is less reliable. In addition, inconsistency with a previous study (Shastri 2011) where a positive maltose phenotype was also observed for another three combinations, T25-TssE and T18-TssF, T25-TssF and T18-TssE, and T25-TssF and TssE-T18. Therefore, there was no strong evidence provided for an interaction between TssE and TssF.

The interaction analysis between TssE and TssF was also carried out in the presence of TssG (three-hybrid assay). The construction of the three-hybrid plasmid pUT18C-*tssF*-*tssG* was described in Section 3.4.1.3. However, the TssE-T25 and T18-TssF-TssG yielded a negative maltose phenotype (result not shown).

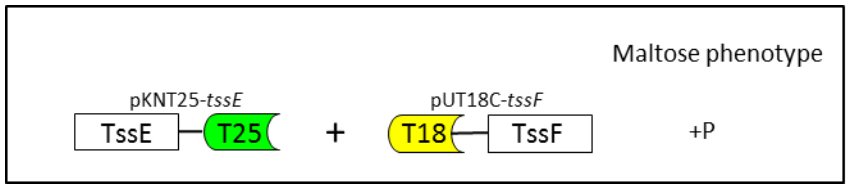


Figure 7.2 BACTH analysis of the interactions between TssE and TssF on MacConkey-maltose agar. Compatible BACTH plasmids encoding the CyaA T25 or T18 fragments fused to TssE and TssF were introduced into *E. coli* strain BTH101 and transformants were grown on MacConkey-maltose agar at 30℃. Maltose phenotype was observed after 120 hours incubation. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated next to the corresponding diagram as crosses according to the classification described in Table 3.2; 'P', patchy; combinations that yielded a negative maltose phenotype were not shown.

7.4 Analysis of interactions between TssG and all the other T6SS subunits

As pKNT25-*tssG* and pUT18-*tssG* two-hybrid constructs were defective, reconstruction of the plasmids was carried out as described in Section 3.2.2. The interactions of TssG-T25 and TssG-T18 with all the other T6SS subunits except $TssA^S$ (the latter results are presented in Section 3.3.7) were repeated with rectified constructs. The interactions between T25 fused to the N- or C-terminus of TssG with TssH-T18 were not assayed as the pUT18-*tssH* plasmid was not constructed. Apart from these combinations, four other combinations for analysing the interaction between TssF and TssG were also assayed, i.e. T25-TssG with T18 fused to the N- or C-terminus of TssF, and T18-TssG with T25 fused to the N- or C-terminus of TssF. However, no interactions between TssG and other T6SS subunits were detected by the BACTH assay (results not shown).

7.5 Analysis of interactions of TssI (and the TssI core region) with TssB, TssE, TssF, TssJ, TssK and TssL

TssI has been shown to interact with TssA, TssB, TssD, TssE, TssF, TssH, TssJ, TssK and TssL in a previous study by BACTH assay (Ahmad 2013). The interactions with TssA, TssD and TssH were subsequently shown to occur through the gp27gp5-like moiety of TssI ($TssI_{gp27gp5}$) but interactions with other subunits were not tested. The interactions between $TssI_{gp27gp5}$ and TssB, TssE, TssF, TssJ, TssK and TssL were assayed in this study and the interactions of these subunits with TssI were repeated for comparison.

The $TssI_{gp27gp5}$ two-hybrid plasmids used were pKT25-$tssI_{gp27gp5}$ and pUT18C-$tssI_{gp27gp5}$ in pairwise combination with BACTH plasmids encoding T18 or T25 fused to T6SS subunits. The corresponding two-hybrid plasmids pKT25-*tssI* and pUT18C-*tssI* were also included. The results showed that one combination of TssI and TssB hybrid proteins gave rise to a positive maltose phenotype (Figure 7.3 A), which was TssB-T25 and T18-TssI. This is not consistent with the previous record, where a maltose positive phenotype was also observed for the combination of T25-TssB and T18-TssI (Ahmad 2013). Moreover, two combinations of $TssI_{gp27gp5}$ and TssB fusion proteins gave rise to

a positive maltose phenotype, which were T18-TssI$_{gp27gp5}$ in combination with T25 fused to the N- or C-terminus of TssB.

One of the combinations of TssI and TssE hybrid proteins yielded a positive maltose phenotype, TssE-T25 and T18-TssI, and two combinations between TssI$_{gp27gp5}$ and TssE also gave rise to a positive maltose phenotype, which were T18-TssI$_{gp27gp5}$ with T25 fused to the N- or C-terminus of TssE (Figure 7.3 B). Three out of four combinations between TssI and TssF yielded a positive maltose phenotype, except the combination of T25-TssI and TssF-T18 (Figure 7.4). The result of the combination of T25-TssI and T18-TssF is not consistent with a previous record, where a negative maltose phenotype was observed (Ahmad 2013). A positive maltose phenotype was also observed for three combinations of TssI$_{gp27gp5}$ and TssF hybrid proteins, except T25-TssI$_{gp27gp5}$ and TssF-T18. One of the four assayed combinations of TssI and TssJ hybrid proteins yielded a positive maltose phenotype with pink/red colonies, TssJ-T25 and T18-TssI, which is consistent with the record of a previous study (Ahmad 2013). However, as mentioned in the Section 3.3.9 that a control combination of pKNT25 with pUT18C-tssI gave rise to pink/red colonies and this is the only combination of TssI with TssJ that yielded a positive maltose phenotype. Therefore, there was no convincing evidence for the interaction between these two proteins (results not shown).

Three combinations of TssI and TssK hybrid proteins gave rise to a maltose positive phenotype, except the combination of T25-TssI with TssK-T18 (Figure 7.5 A). However, a negative maltose phenotype was observed for the combination of TssK-T25 with T18-TssI in a previous study (Ahmad 2013). Consistent with this, three combinations of TssI$_{gp27gp5}$ and TssK hybrid proteins yielded a positive maltose phenotype, except T25-TssI$_{gp27gp5}$ with TssK-T18, where a negative maltose phenotype was observed. Two combinations of TssI with TssL hybrid proteins yielded a positive maltose phenotype, which were T25-TssL with T18-TssI, and T25-TssI with T18-TssL (Figure 7.5 B), and this is consistent with the previous study (Ahmad 2013). Three out of four combinations of TssI$_{gp27gp5}$ and TssL hybrid proteins gave rise to a positive maltose phenotype with the exception being T25-TssI$_{gp27gp5}$ and TssL-T18.

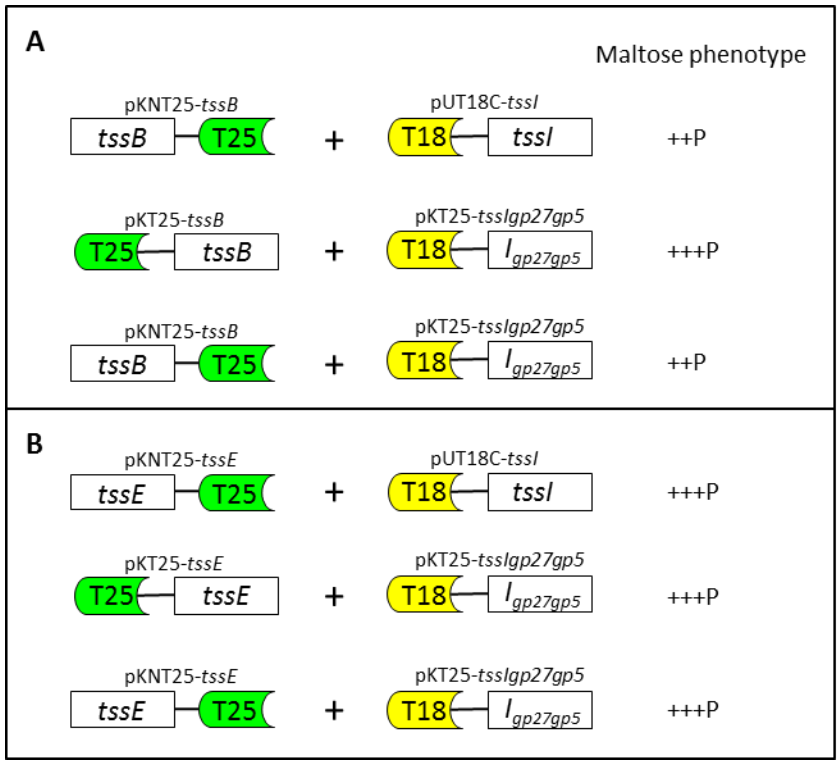


Figure 7.3 BACTH analysis of the interactions between TssI or the $TssI_{gp27gp5}$ moiety and TssB or TssE on MacConkey-maltose agar. Compatible BACTH plasmids encoding the CyaA T25 or T18 fragments fused to TssI or the $TssI_{gp27gp5}$ moiety and TssB/TssE were introduced into *E. coli* strain BTH101 and transformants were grown on MacConkey-maltose agar at 30 ℃. Maltose phenotype was observed after 120 hours incubation. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated next to the corresponding diagram as crosses according to the classification described in Table 3.2; 'P', patchy; combinations that yielded a negative maltose phenotype were not shown. **A**. Combinations of TssB and TssI or the $TssI_{gp27gp5}$ moiety fusion proteins that gave rise to a maltose-positive phenotype. **B**. Combinations of TssE and TssI or the $TssI_{gp27gp5}$ moiety fusion proteins that gave rise to a maltose-positive phenotype.

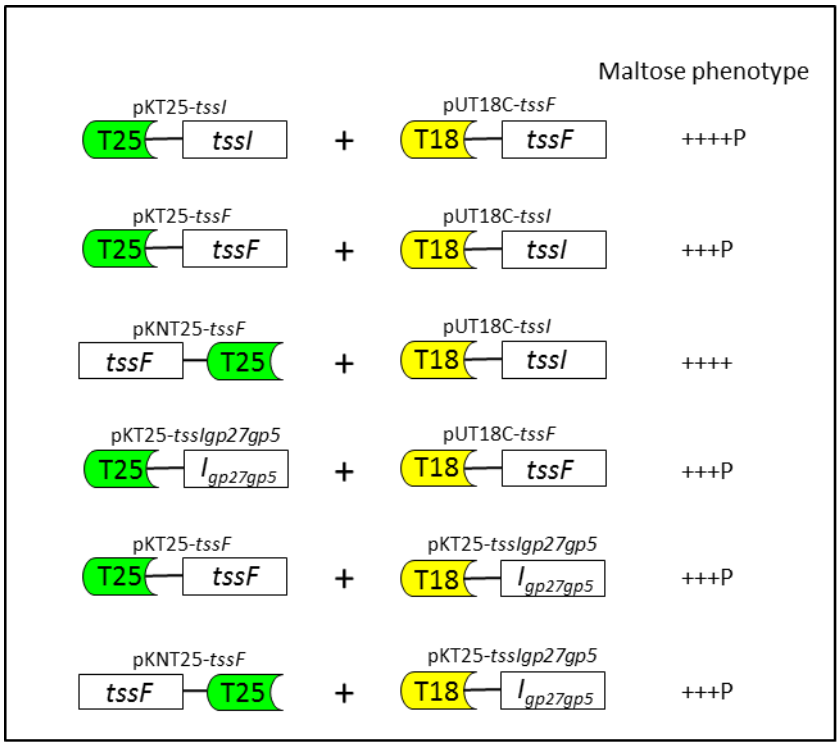


Figure 7.4 BACTH analysis of the interactions between TssI or the $TssI_{gp27gp5}$ moiety and TssF on MacConkey-maltose agar. Compatible BACTH plasmids encoding the CyaA T25 or T18 fragments fused to TssI or the $TssI_{gp27gp5}$ moiety and TssF were introduced into *E. coli* strain BTH101 and transformants were grown on MacConkey-maltose agar at 30 ℃. Maltose phenotype was observed after 120 hours incubation. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated next to the corresponding diagram as crosses according to the classification described in Table 3.2; 'P', patchy; combinations that yielded a negative maltose phenotype were not shown

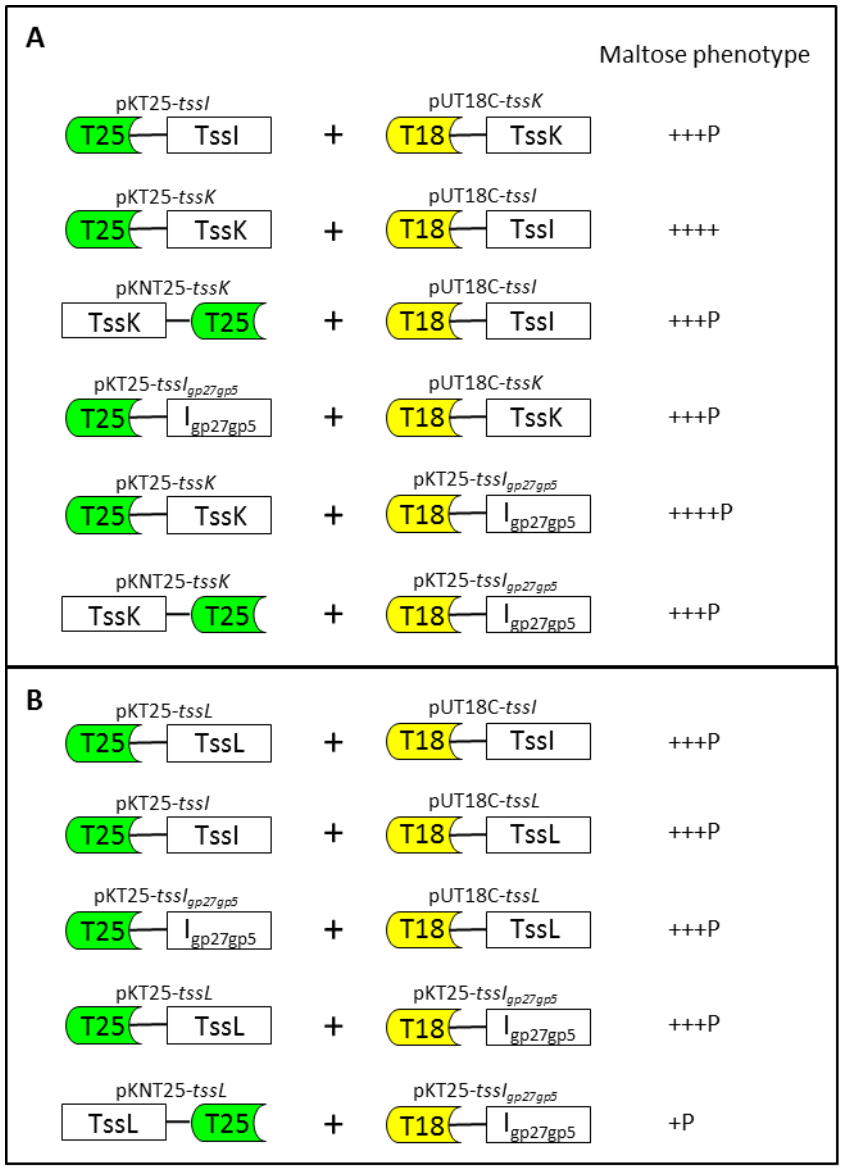


Figure 7.5 BACTH analysis of the interactions between TssI or the $TssI_{gp27gp5}$ moiety and TssK or TssL on MacConkey-maltose agar. Compatible BACTH plasmids encoding the CyaA T25 or T18 fragments fused to TssI or the $TssI_{gp27gp5}$ moiety and TssK/TssL were introduced into *E. coli* strain BTH101 and transformants were grown on MacConkey-maltose agar at 30 ℃. Maltose phenotype was observed after 120 hours incubation. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated next to the corresponding diagram as crosses according to the classification described in Table 3.2; 'P', patchy; combinations that yielded a negative maltose phenotype were not shown. **A**. Combinations of TssK and TssI or the $TssI_{gp27gp5}$ moiety fusion proteins that gave rise to a maltose-positive phenotype. **B**. Combination of TssL and TssI or the $TssI_{gp27gp5}$ moiety of TssI fusion proteins that gave rise to a maltose-positive phenotype.

7.6 Analysis of interactions between $TssM_{NTD}$ and all other T6SS subunits except TssA

As pKNT25-*tssM*$_{NTD}$ and pUT18-*tssM*$_{NTD}$ two-hybrid constructs were defective, they were rectified as described in Section 3.2.3. The interactions between $TssM_{NTD}$-T25 and $TssM_{NTD}$-T18 with other T6SS subunits except $TssA^{S}$ (Section 3.3.13) were then repeated with the correct plasmids. The interactions between T25 fused to the N- or C-terminus of $TssM_{NTD}$ with TssH-T18 were not assayed as the pUT18-*tssH* two-hybrid plasmid was not constructed. The results showed that two combinations of plasmids gave rise to a positive maltose phenotype which encoded $TssM_{NTD}$-T25 and T18-TssL, and T25-TssL and $TssM_{NTD}$-T18 (Figure 7.6). All the remaining combinations yielded a negative maltose phenotype (results not shown). Another combination between $TssM_{NTD}$ and TssL hybrid proteins was also assayed, which was T25-$TssM_{NTD}$ and T18-TssL. This combination resulted in a positive maltose phenotype, and this is consistent with a previous study (Ahmad 2013). The β-glactosidase assay was carried out for two combinations between $TssM_{NTD}$ and TssL (T25-$TssM_{NTD}$ and T18-TssL, and $TssM_{NTD}$-T25 and T18-TssL) and the data was shown in Section 3.4.3.

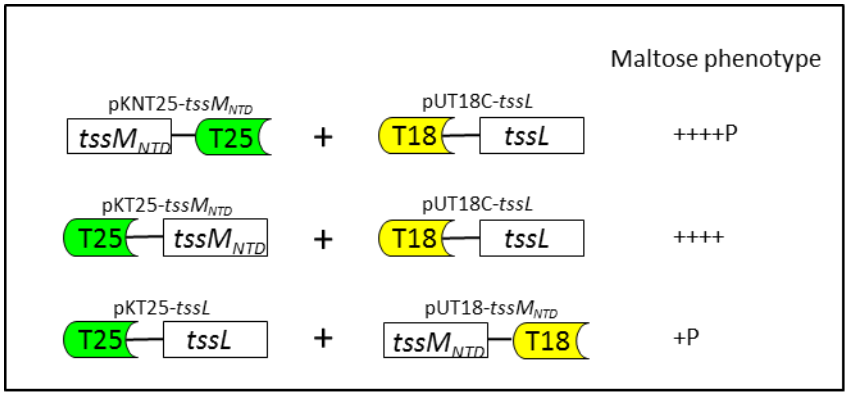


Figure 7.6 BACTH analysis of the interactions between $TssM_{NTD}$ and all the other T6SS subunits (except TssA) on MacConkey-maltose agar. Compatible BACTH plasmids encoding the CyaA T25 or T18 fragments fused to a T6SS subunit (except TssA) and $TssM_{NTD}$-T25 or $TssM_{NTD}$-T18, including a combination of T25-$TssM_{NTD}$ and T18-TssL were introduced into *E. coli* strain BTH101 and transformants were grown on MacConkey-maltose agar at 30 ℃. Maltose phenotype was observed after 120 hours incubation. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated next to the corresponding diagram as crosses according to the classification described in Table 3.2; 'P', patchy; combinations that yielded a negative maltose phenotype were not shown.

7.7 Analysis of homo-oligomerisation of TssC, TssF, TssG and $TssM_{NTD}$

The possible self-interaction of TssC, TssF, TssG and $TssM_{NTD}$ were analysed using the BACTH assay. Previous studies have assessed the self-interaction of these proteins with defective two-hybrid constructs of TssC, TssG and $TssM_{NTD}$. Therefore, the analysis was repeated with corrected two-hybrid plasmids in this study. In addition, the self-interaction of TssF was also repeated. All four possible pairwise self-combinations of each subunit were assayed, except $TssM_{NTD}$, where the combination of T25-$TssM_{NTD}$ and T18-$TssM_{NTD}$ was not included in the BACTH assay.

The results showed that all four TssC self-combinations gave rise to a negative maltose phenotype (results not shown). In contrast, three self-combinations of TssF showed a positive maltose phenotype, except the combination of TssF-T25 and TssF-T18 hybrid proteins (Figure 7.7 A). The combinations of TssG-T25 and T18-TssG, and T25-TssG and TssG-T18 hybrid proteins gave rise to a positive maltose phenotype (Figure 7.7 B), but not the other two combinations. Two combinations of $TssM_{NTD}$ yielded a positive maltose phenotype, i.e. $TssM_{NTD}$-T25 and T18-$TssM_{NTD}$, and T25-$TssM_{NTD}$ and $TssM_{NTD}$-T18 (Figure 7.7 C). Hence, the results suggest TssF, TssG and $TssM_{NTD}$ are able to self-interact to form higher order complexes.

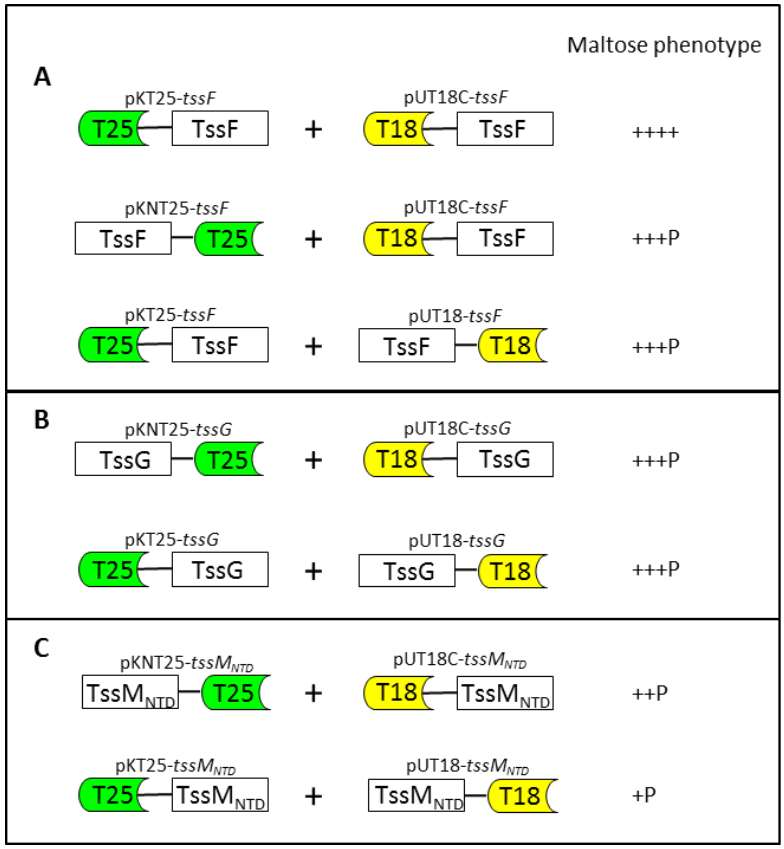


Figure 7.7 BACTH analysis of the self-interactions of TssF, TssG and $TssM_{NTD}$ on MacConkey-maltose agar. Compatible BACTH plasmids encoding the CyaA T25 or T18 fragments fused to TssF, TssG or $TssM_{NTD}$ were introduced into *E. coli* strain BTH101 and transformants were grown on MacConkey-maltose agar at 30 ℃. Maltose phenotype was observed after 120 hours incubation. For combinations that gave rise to a maltose-positive phenotype, the degree of the phenotype is indicated next to the corresponding diagram as crosses according to the classification described in Table 3.2; 'P', patchy; combinations that yielded a negative maltose phenotype were not shown. **A**. Combinations of TssF fusion proteins that yielded a Mal^+ phenotype. **B**. Combinations of TssG fusion proteins that yielded a Mal^+ phenotype. **C**. Combinations of $TssM_{NTD}$ fusion proteins that yielded a Mal^+ phenotype.

7.8 Discussion

Two-hybrid combinations in the protein interaction analysis that involved defective two-hybrid plasmids encoding TssC, TssG and $TssM_{NTD}$ were repeated with reconstructed or rectified plasmids, and the three-hybrid analysis of interactions between TssE and TssF in the presence of TssG was also assayed in this study. In addition, homo-oligomerisation of TssC, TssF, TssG and $TssM_{NTD}$ was also analysed. A summary of the BACTH assay results for TssA-TssL, $TssM_{NTD}$ and $TssM_{CTD}$ are presented in Figure 7.8. The tail sheath protein, TssC, was shown to interact with the other tail sheath protein, TssB, and no interactions with other T6SS subunits were identified by the BACTH system. The observation is in agreement with previous findings (Bönemann et al. 2009; Lossi et al. 2013; Zhang et al. 2013).

High levels of co-organisation of the *tssE*, *tssF* and *tssG* genes within T6SS gene clusters suggest an interaction between the products of these three genes and, in fact, TssE, TssF and TssG have been shown to be the baseplate components of T6SS (Boyer et al. 2009; Brunet et al. 2015). A weak positive maltose phenotype was observed for one of the combinations of TssE with TssF in this study. However, due to the problematic plasmid control combination of pKNT25 and pUT18C-*tssF*, there was no strong evidence provided for an interaction between TssE and TssF. This result is consistent with the two-hybrid analysis of Brunet and colleagues (Brunet et al. 2015), where no interaction between TssE and TssF was identified *in vivo*. Nevertheless, an interaction between these two proteins *in vitro* was demonstrated by co-IP by this group.

In support of previous observations where an interaction between TssM and TssL has been shown (Ma et al. 2009a; Aschtgen et al. 2010a; Felisberto-Rodrigues et al. 2011), interactions between $TssM_{NTD}$ (the cytoplasmic region of TssM) and TssL was also demonstrated in this study, evidenced by three different two-hybrid combinations that yielded a maltose-positive phenotype.

The tail spike protein, TssI, gave rise to a maltose-positive phenotype in combination with TssB (tail sheath protein), TssE (baseplate component), TssF (baseplate component), TssK (baseplate component) and the cytoplasmic domain of TssL (inner

membrane protein of the transmembrane complex). The interaction between TssI and TssF in the presence of TssG was previously reported (Brunet et al. 2015), and other interactions between TssI and TssB, TssE, TssK or TssL were previously recorded by Ahmad (2013). It is expected the tail spike protein TssI would interact with one of the tail sheath proteins, such as TssB, in the extended state, or transiently during in the injection state. The interactions between TssI and the baseplate components, TssE, TssF and TssK, are most likely to occur in the extended state as well, when the T6SS apparatus has been assembled and is ready for injection. In addition, it is more likely a transient interaction between TssI and TssL could happen in the extended state, where TssL, together with TssM and TssJ forms a trans-membrane channel to allow the passage of the tail tube TssD with TssI spike sitting on the top of it. Moreover, the results of this study also suggested the interactions between TssI and other T6SS subunits occurred through the gp27gp5-like moiety of TssI, which is functionally equivalent to the bacteriophage T4 tail spike.

TssF and TssG were observed to homo-oligomerise, evidenced by three different combinations of TssF hybrid proteins and two combinations of TssG hybrid proteins that yielded a positive maltose phenotype in the BACTH assay. These results are consistent with a previously reported study (Brunet et al. 2015). In addition, two combinations of $TssM_{NTD}$ hybrid proteins gave rise to positive-maltose phenotype, suggesting $TssM_{NTD}$ is able to oligomerise.

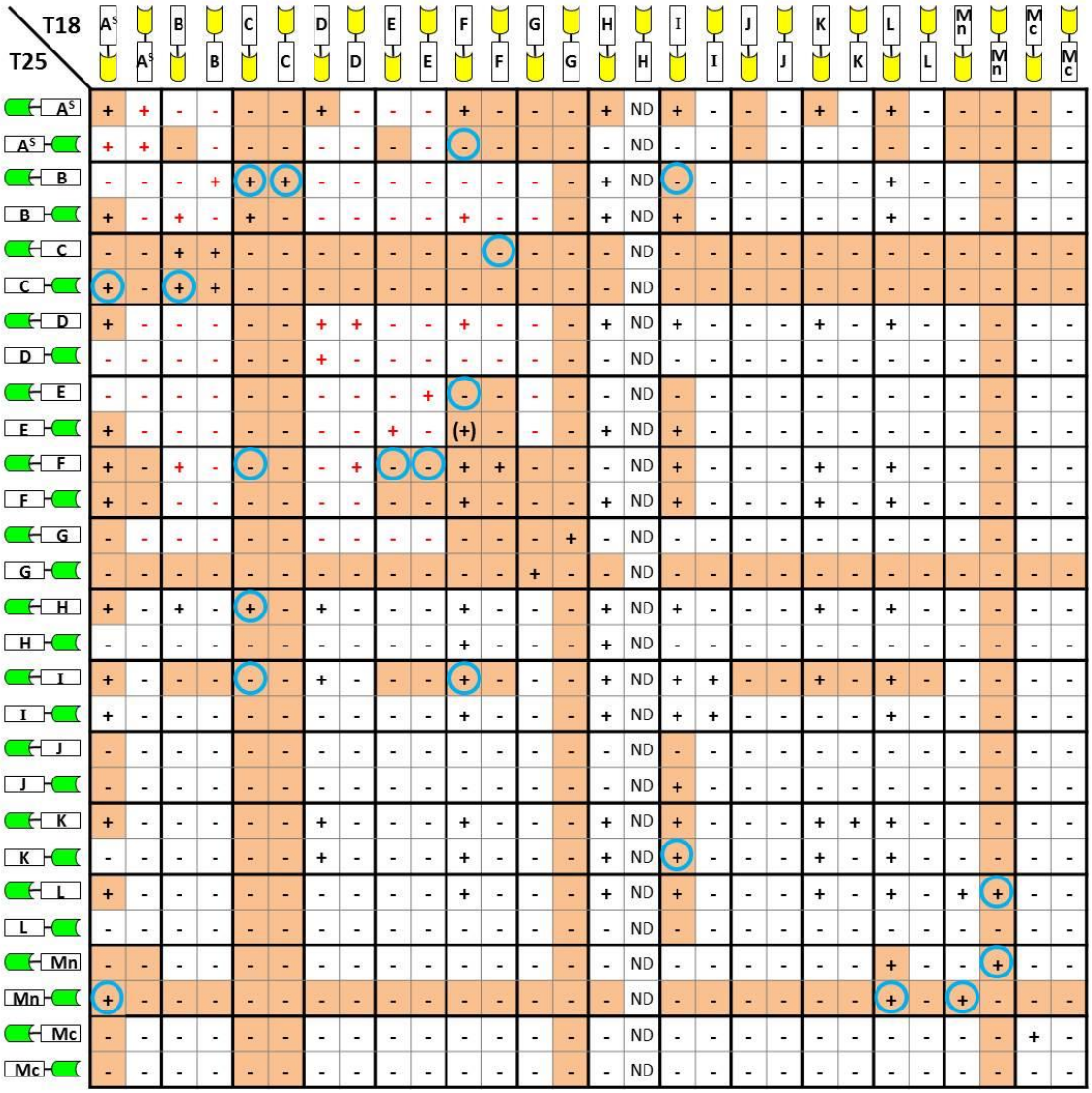


Figure 7.8 Summary of BACTH assay results for TssA-TssL, $TssM_{NTD}$ and $TssM_{CTD}$. Diagrammatic representation of two-hybrid fusions are described underneath the table. Results highlighted in faint red are from this study. Those in red fount are from Shastri (2011) and those in black font are from Ahmad (2013). '+', maltose-positive phenotype (Mal^+); '-', maltose-negative phenotype (Mal^-), 'ND', not done (pUT18-tssH was not constructed). Results of this study that are not consistent with previously reported results (Shastri 2011; Ahmad 2013) are indicated with a blue circle. Observations from this study were included in the table for conflictive results. Result in brackets indicates no strong evidence suggesting an interaction between these two proteins.

Chapter 8 Characterisation of TssE and TssK

8.1 Introduction

TssE has been reported to have a significant amino acid sequence similarity (~40%) with bacteriophage T4 gp25 (Shalom et al. 2007; Leiman et al. 2009; Lossi et al. 2011), which is a constituent of the bacteriophage baseplate and it was proposed to have a lysozyme-like activity (Kostyuchenko et al. 2003; Yap et al. 2010). Although the reported gp25 structure from *Geobacter sulfurreducens* facilitates the prediction of the structure of TssE (Cascales and Cambillau 2012), its structure has not been solved. In addition, when work on this investigation started, TssK had not been characterised. Therefore, it was decided to investigate into the structures of TssE and TssK.

8.2 Analysis of TssE

8.2.1 Construction of a plasmid for overproduction of a MBP-TssE fusion protein

As TssE is largely insoluble when overproduced with an N-terminal hexa-histidine tag from pACYCDuet-His_6.*tssE* in *E. coli* and no overproduction of C-terminal VSVg-tagged TssE was observed from pACYCDuet-*tssE*.VSVg (Section 4.4.1.4), pMAL-c5X-His_6-*tssE* was constructed which encodes a $_{His6}$.MBP-TssE fusion for solubilising TssE (Mosby and Thomas, unpublished results). This protein contains a hexahistidine tag between amino acids 3 and 4 of MBP and a factor Xa cleavage site is located between the MBP and TssE. Therefore, TssE can be released from the fusion protein following factor Xa cleavage.

8.2.2 Overproduction and purification of TssE from $_{His6}$.MBP-TssE fusion protein

To overproduce $_{His6}$.MBP-TssE, the *tac* promoter on pMAL-c5X-His_6-*tssE* in NEB Express cells was induced with 0.3 mM IPTG at 30^{o}C. Following induction, a protein with the expected size of $_{His6}$.MBP-TssE (~62 kDa) was overproduced and remained in the soluble fraction following cell lysis as observed by SDS-PAGE (Figure 8.1 A). However, the amount of overproduced $_{His6}$.MBP-TssE in NEB Express cells was less than that obtained in *E. coli* BL21(λDE3) (Section 4.4.1.4, Figure 4.18 C).

To purify it, IMAC on Ni-ABT was employed. However, the $_{His6}$.MBP-TssE did not bind to the column, probably due to the presence of EDTA in the binding buffer (Figure 8.1 B). The unbound material was then collected and diluted to reduce the NaCl concentration from 500 mM in the IMAC binding buffer to 200 mM prior to being subjected to the amylose affinity chromatography. $_{His6}$.MBP-TssE bound to the amylose column efficiently and the bound material was then eluted with buffer containing 10 mM maltose. Fractions containing $_{His6}$.MBP-TssE were combined and the protein was concentrated using a Vivaspin centrifugal concentrator (MWCO 30 kDa), followed by digestion with factor Xa protease at a final concentration of ~42 μg/ml in the presence of 2 mM $CaCl_2$ at room temperature overnight. However, during cleavage, the released TssE precipitated as observed by SDS-PAGE analysis. Insoluble TssE was recovered by centrifugation and attempts were made to solublise it by resuspending the pellet in sequential buffers, i.e. (i) 1 M NaCl and 50 mM Tris-HCl pH 8.0, (ii) 1 M NaAc (pH 4.6), (iii) 0.8 M HEPES (pH 7.6) and 1.6 M urea and (iv) 0.4 M NaCl, 50 mM Tris-HCl (pH 8.0) and 1.8% dodecyl maltoside (DDM), in which a centrifugation step was included among different solublising conditions. Following centrifugation, the remaining insoluble material was dissolved in 8 M urea and subjected to SDS-PAGE analysis. However, TssE was largely present in the insoluble 8 M urea fraction.

In another purification procedure, IMAC was omitted and purification was initiated with amylose affinity chromatography (Figure 8.2 A). $_{His6}$.MBP-TssE that eluted from the amylose column was subjected to SEC on a Superdex 200 column. The elution profile of $_{His6}$.MBP-TssE from Superdex 200 column showed two peaks at 14.12 and 15.23 ml detected by UV_{280}, corresponding to an apparent protein MW of 66 kDa and 44 kDa, respectively (Figure 8.2 B). Upon SDS-PAGE analysis, the earlier peak corresponded to $_{His6}$.MBP-TssE (~62 kDa) and the later peak was consistent with the MW of MBP (~45 kDa). Moreover, two additional polyeptides were observed which migrated according to molecular weights of ~20 kDa and ~5 kDa (Figure 8.2 A). The former corresponded in size to TssE. The results suggested it was likely that a proportion of the $_{His6}$.MBP-TssE fusion was being cleaved in the polyasparagine linker region or adjacent factor Xa cleavage site prior to being subjected to Factor Xa protease. This observation also suggests that TssE exists as a monomer is solution. The elution fractions containing a mixture of $_{His6}$.MBP-TssE fusion and potential MBP were mixed with arginine and glutamic acid at a final concentration of 0.5 M each for preventing protein aggregation

and precipitation before being concentrated (MWCO 50 kDa, 9 mg/ml) and cleaved with factor Xa protease at a final concentration of ~50 μg/ml in the presence of 2 mM $CaCl_2$ at room temperature overnight. The digestion was effective. Although the presence of arginine and glutamic acid increased the proportion of TssE that remained in the soluble fraction, the majority of TssE (~80%) was still present in the insoluble fraction. Therefore it was unsuccessful to obtain large amount of soluble material of TssE for further purification.

As a large amount of $_{His6}$.MBP-TssE was present in the cell lysate used for purification, there was still a lot of fusion protein present in the unbound fraction following amylose column (Figure 8.2 A). Therefore, the unbound material from amylose chromatography was recycled onto the column and subjected to SEC for obtaining pure $_{His6}$.MBP-TssE. The procedures were as described previously. $_{His6}$.MBP-TssE (7.5 mg/ml) was used for crystallisation trials upon buffer exchange into 5 mM Tris-HCl (pH 8.0) containing 50 mM NaCl using a Zeba column (Thermo) for crystallisation trials. However, there was no crystal hints observed.

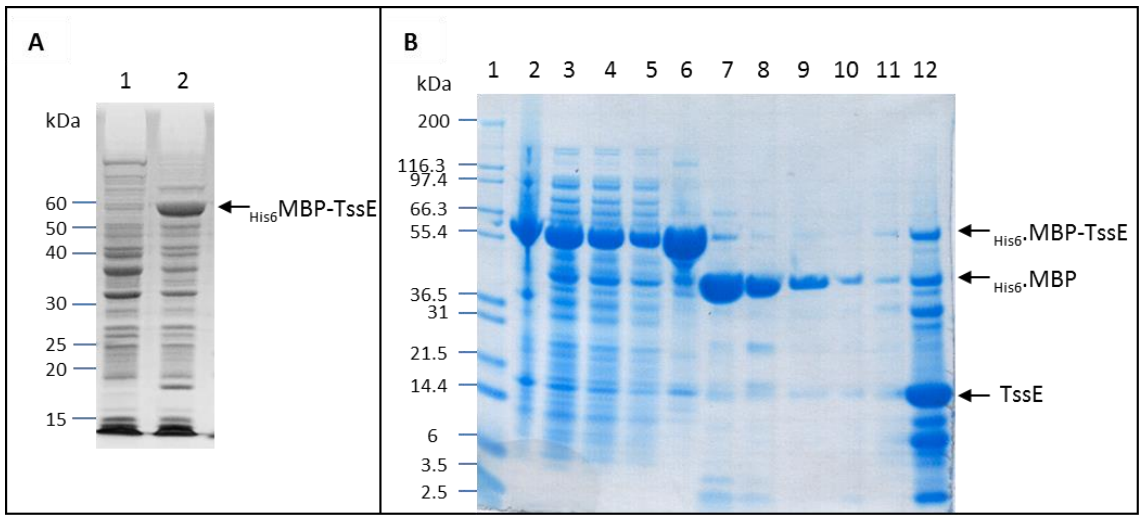


Figure 8.1 Solubility and purification of TssE following overproduction and cleavage of $_{His6}$.MBP-TssE by factor Xa. $_{His6}$.MBP-TssE was overproduced from *E. coli* strain NEB Express cells containing pMAL-c5X-His$_6$-*tssE* with 0.3 mM IPTG induction at 37^{o}C. **A**. A Coomassie blue-stained 12% SDS-PA gel showing the analysis of $_{His6}$MBP-TssE fusion protein overproduction. Lane 1, total cell protein from uninduced cells; lane 2, total cell protein from cells following induction. The arrow indicates the expected location of $_{His6}$.MBP-TssE based on its size (~62 kDa). **B**. 4-12% gradient SDS-PA gel (Novex) showing the solubility of $_{His6}$MBP-TssE and purification of TssE from MBP fusion by factor Xa cleavage. Following induction, the soluble fraction of the cell lysate was subjected to IMAC in a buffer containing 50 mM Tris-HCl (pH 8.0), 500 mM NaCl and 1 mM EDTA, followed by amylose chromatography in a buffer containing 50 mM Tris-HCl (pH 8.0) and 200 mM NaCl. Elution was performed in the same buffer containing 10 mM maltose. The elution fraction containing $_{His6}$.MBP-TssE was subjected to factor Xa cleavage. Lane 1, protein marker (Novex); lane 2, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 3, soluble fraction of cell lysate following induction; lane 4, unbound material (flow-through) from nickel agarose chromatography; lane 5, unbound material (flow-through) from amylose column; lane 6, bound material of amylose column (i.e. following elution with maltose); lane 7, soluble fraction following factor Xa cleavage; lanes 8-12, solubilisation of TssE under different buffer conditions: lane 8, 1 M NaCl and 50 mM Tris-HCl (pH 8.0); lane 9, 1 M NaAc (pH 4.6); lane 10, 0.8 M HEPES (pH 7.6) and 1.6 M uera; lane 11, 0.4 M NaCl, 50 mM Tris-HCl (pH 8.0) and 1.8% DDM; lane 12, pellet dissolved in 8 M urea. The arrows indicate the expected locations of $_{His6}$.MBP-TssE (~62 kDa), $_{His6}$.MBP (~45 kDa) and TssE (~18 kDa).

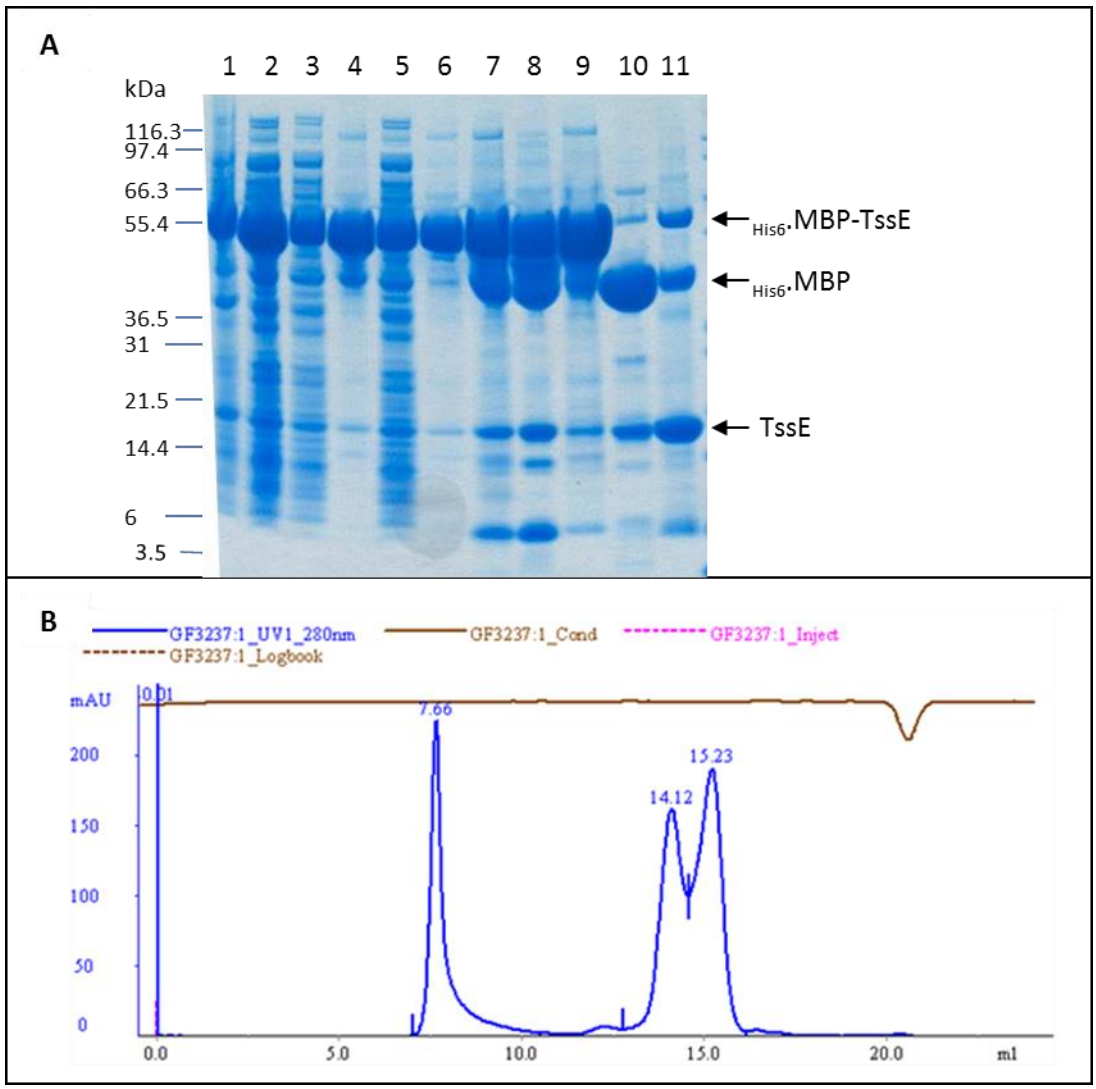


Figure 8.2 Solubility of TssE in the presence of arginine and glutamic acid and purification of $_{His6}$.MBP-TssE. $_{His6}$.MBP-TssE was overproduced from *E. coli* strain NEB Express cells containing pMAL-c5X-His$_6$-*tssE* with 0.3 mM IPTG induction at 37^{o}C. Following induction, the soluble fraction of the cell lysate was subjected to amylose chromatography in a buffer containing 50 mM Tris-HCl (pH 8.0) and 200 mM NaCl. Elution was performed in buffer containing 10 mM maltose. The elution fractions containing $_{His6}$.MBP-TssE were combined and subjected to SEC on a Superdex 200 10/300 GL column in buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl. $_{His6}$.MBP-TssE was then subjected to factor Xa cleavage in the presence of arginine and glutamic acid (0.5 M each), and another preparation of $_{His6}$.MBP-TssE was subjected to crystallization trials. **A**. 4-12% gradient SDS-PA gel (Novex) showing purification of $_{His6}$.MBP-TssE, factor Xa cleavage and solubilisation of TssE in the presence of arginine and glutamic acid. Lane 1, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 2, soluble fraction of cell lysate following induction; lane 3, unbound material_1 from amylose column; lane 4, bound material_1 from amylose column; lane 5, unbound material_2 from amylose column (from recycled unbound material_1); lane 6, bound material_2 from amylose column; lane 7, peak elution fractions round 1 from SEC; lane 8, peak elution fractions round 2 from SEC; lane 9, material subjected to factor Xa protease cleavage in the presence of arginine and glutamic acid; lane 10, soluble fraction following protease cleavage; lane 11, insoluble material following protease cleavage. The arrows indicate the expected locations of $_{His6}$.MBP-TssE (~62 kDa), $_{His6}$.MBP (~45 kDa) and TssE (~18 kDa). **B**. Elution profile of $_{His6}$.MBP-TssE during SEC.

8.3 Analysis of TssK

8.3.1 Analysis of N-terminal hexahistidine tagged TssK

8.3.1.1 Construction of a plasmid for overproduction of $_{His6}$.TssK

To overproduce TssK with a hexa-histidine tag at its N-terminus to facilitate its purification, pACYCDuet-His_6.*tssK* was constructed. *B. cenocepacia* H111 *tssK* was released from pUT18C-*tssK* BACTH plasmid by digesting with *Bam*HI and *Acc*65I, and ligated to pACYCDuet-1 digested with *Bam*HI and *Bsr*GI. Therefore, *tssK* was under the control of the first T7 promoter of the plasmid, and the encoded TssK protein possessed a hexa-histidine tag at its N-terminus that was encoded by the plasmid. *E. coli* strain MC1061 was transformed with the ligation mixtures and colonies were grown on LB plates containing chloramphenicol. The correct recombinant plasmids were identified by PCR screening and followed by plasmid miniprep and nucleotide sequence analysis to confirm the plasmid contained the *tssK* insert. A diagrammatic representation of the $_{His6}$.TssK overproduction plasmid is shown in Figure 8.3 A.

8.3.1.2 Overproduction and purification of $_{His6}$.TssK

Following introduction of pACYCDuet-His_6.*tssK* into BL21(λDE3) cells, the T7 promoters present on the plasmid were induced with 1 mM IPTG at 37^{o}C. $_{His6}$.TssK was overproduced and remained in the soluble fraction after cell lysis as judged by SDS-PAGE analysis (Figure 8.3 B).

IMAC was employed to purify $_{His6}$.TssK. The soluble fraction of the cell lysate containing overproduced $_{His6}$.TssK was applied onto a HisTrap column. The flow-through was recycled by re-applying onto the column to maximize the binding of $_{His6}$.TssK. Upon SDS-PAGE analysis, it was apparent that there was no significant decrease in the amount of $_{His6}$.TssK in the flow-through compared to that present in the loaded material, and recycle step did not improve the protein binding to the column (Figure 8.3 C). This may have been due to the large amount of TssK that was used for purification that have saturated the column. The elution of $_{His6}$.TssK from HisTrap was monitored by UV at 280 nm, and a single peak was observed corresponding to protein that was eluted with buffer containing approximately 200 mM imidazole. Elution

fractions corresponding to the peak were combined and judged to be >90% pure. Following concentration to 7 mg/ml (MWCO 30 kDa, Vivaspin centrifugal concentrator), the partially purified protein was subjected to SEC on a Superose 6 column for molecular size estimation. The UV trace at 280 nm showed a single peak eluting at ~26 minute, corresponding to a protein size of 150-200 kDa calculated using the calibration curve of known proteins. Given the monomer MW of $_{His6}$.TssK is ~51 kDa, TssK was estimated to be a trimer or tetramer.

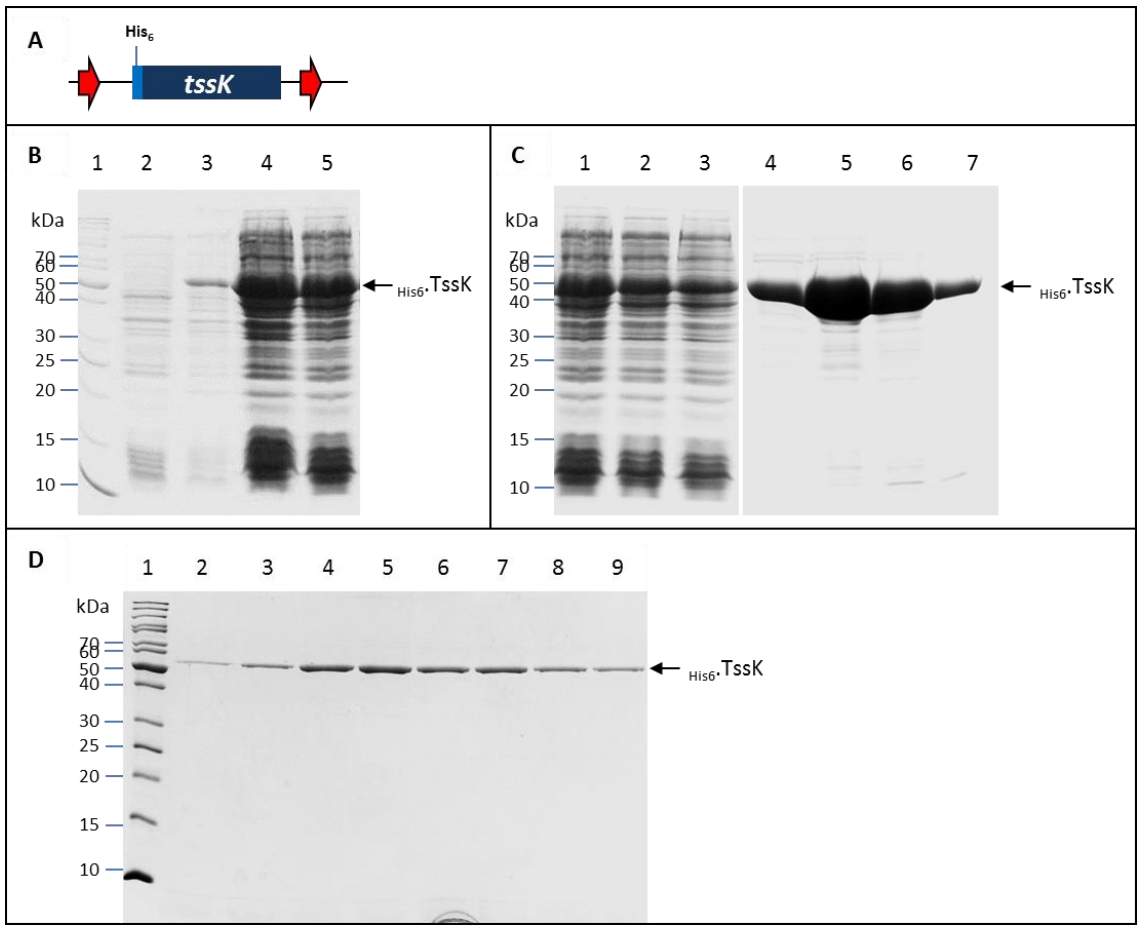


Figure 8.3 Overproduction, solubility and purification of $_{His6}$.TssK. Induction of pACYCDuet-His$_6$.*tssK* in *E. coli* strain BL21 (λDE3) cells with 1 mM IPTG at 37^{o}C. Following induction, the soluble fraction of the cell lysate was applied onto a HisTrap HP column (1 ml, GE Healthcare) in a buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10% glycerol and 10 mM imidazole. Elution was performed in the same buffer containing an increasing gradient concentration of imidazole up to 500 mM. Fractions containing $_{His6}$.TssK were combined, concentrated and subjected to SEC on a Superose 6 column in a buffer containing 50 mM Tris-HCl (pH 8.0) and 50 mM NaCl. **A**. Schematic drawing showing the location of *tssK* in pACYCDuet-His$_6$.*tssK*. The red arrows represent the two T7 promoters present in the plasmid. **B**. Coomassie blue-stained 12% SDS-PA gel showing the overproduction and solubility of $_{His6}$.TssK. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction; lane 5, soluble fraction of cell lysate following induction. **C**. Coomassie blue-stained 12% SDS-PA gels showing the purification of $_{His6}$.TssK by IMAC. Lane 1, soluble fraction of cell lysate following induction; lanes 2 and 3, flow-through from HisTrap column following application of clarified lysate; lanes 4-7, peak elution fractions corresponding to the UV tace at 280 nm. **D.** Coomassie blue-stained 12% SDS-PA gel showing the purification of $_{His6}$.TssK by SEC. Lane 1, EZ-Run^{TM} Rec protein ladder (Fisher); lanes 2-9, peak elution fractions from Superose 6 column corresponding to the UV trace at 280 nm. The arrows in B-D indicate the expected location of TssK based on its size (~51 kDa).

8.3.1.3 SEC-MALS of TssK

As SEC suggested that TssK was either a trimer or a tetramer (Section 8.3.1.2), SEC-MALS analysis was carried out for a more accurate MW estimation of $_{His6}$.TssK (purification of $_{His6}$.TssK as described in Section 8.3.1.2).

In the SEC-MALS analysis, there was a very small amount of aggregated material eluted in the column void volume (Superose 6 10/300 GL). TssK was eluted as a single peak with traces from all detectors well overlaid (Figure 8.4 A). The peak analysis of TssK was based on the elution profile using Astra software which defined the peak region into two, which are Peak 1 (31.26-32.36 minute) and Peak 2 (30.19-33.70 minute). Peak 2 included Peak 1, which has a broader range in the elution time. Integration of the RI traces indicates ~33 µg of TssK was present in Peak 2.

The MW was determined by a Zimm fit procedure at each point then averages calculated over the peak region (Figure 8.4 B). The slightly downward slope in the MW curve for TssK was not significant. The MW determined for Peak 1 and Peak 2 were 155 kDa and 151 kDa, respectively. There was no indication of any material with a MW of 200 kDa. Given a monomer MW of $_{His6}$.TssK is 52 kDa, the plots are entirely consistent with a trimer and not a tetramer.

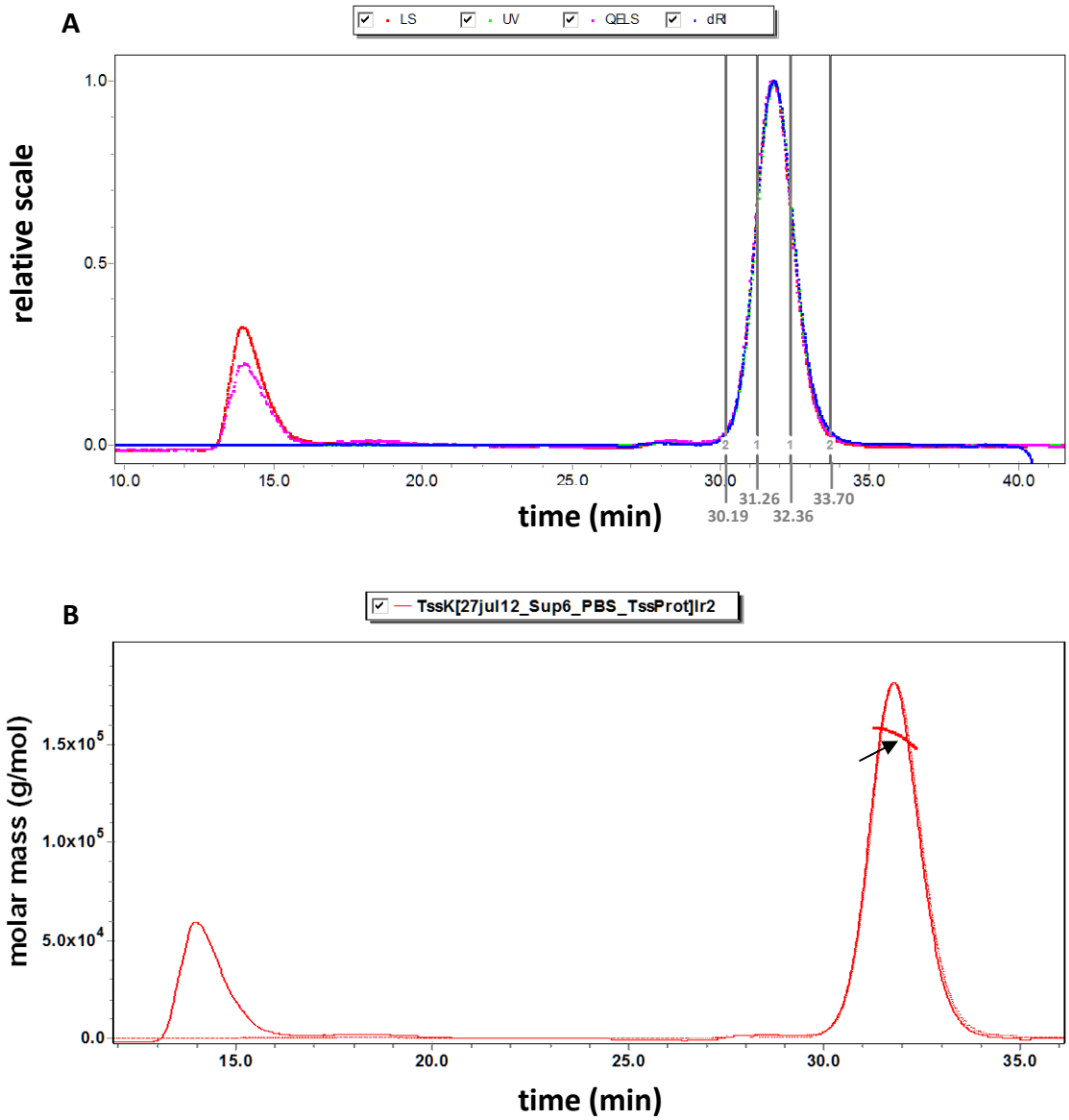


Figure 8.4 SEC-MALS analysis of $_{His6}$.TssK for molecular weight estimation. A. TssK elution profile and peak positions. The normalised signals for light scattering (LS) are shown in red, refractive index (RI) in blue, UV absorption (280 nm) in green and quasi-elastic light scattering (QELS) in magenta. The peak areas chosen for analysis are indicated by numbered vertical grey lines that defined by were Astra software. The X-axis shows elution time in minutes and Y-axis shows relative scale. **B**. TssK molar mass plot. The MW was determined by a Zimm fit procedure at each point. Solid line represents LS trace. Dashed and dotted traces represent RI and UV responses, respectively. Molecular weight estimate curve is shown in heavier dots (indicated with a black arrow). The X-axis is the elution time in minutes and the Y-axis is the MW scale. Figures provided by University of York Bioscience Technology Facility.

8.3.2 Native TssK

8.3.2.1 Construction of a plasmid for overproduction of native TssK

To construct a plasmid expressing native TssK for crystallography, *B. cenocepacia* H111 *tssK* was amplified using primers TssK.NdeI.for and TssK.BglII.rev. The amplified *tssK* DNA fragment of the expected size (~1.35 kb) and pACYCDuet were digested with restriction enzymes *Nde*I and *Bgl*II that recognized sites in the primers and in the second MCS of the plasmid. Ligation was carried out to insert the *tssK* DNA fragment into pACYCDuet for constructing pACYCDuet-*tssK*. *E. coli* strain MC1061 was transformed with the ligation mixture and colonies were grown on LB plates with chloramphenicol. The correct recombinant plasmids were identified by PCR screening and followed by plasmid miniprep and agarose gel electrophoresis to confirm the plasmid contains an insert of the expected size. The nucleotide sequence of the inserted gene was verified by DNA sequencing. A diagrammatic representation of the TssK overexpression plasmid is shown in Figure 8.5 A.

8.3.2.2 Overproduction and purification of native TssK

Native TssK was overproduced in *E. coli* strain BL21 (λDE3) containing pACYCDuet-*tssK* following induction of the T7 promoters present on the plasmid and the protein remained in the soluble fraction following cell lysis (Figure 8.5 B). To purify TssK, the soluble fraction of the cell lysate was subjected to anion exchange chromatography on a DEAE column. TssK bound to the column very well and was eluted from the column in the range 0.2-0.3 M NaCl (Figure 8.5 C). Fractions containing TssK were combined and subjected to ammonium sulphate precipitation (final concentration 1.4 M). Precipitated protein was recovered by centrifugation, dissolved in a small volume of buffer containing 50 mM Tris-HCl (pH 8.0) and 500 mM NaCl, and then subjected to SEC on a Superdex 200 column equilibrated in the same buffer. The elution profile of SEC at UV_{280} showed a single peak at 62.86 ml, and was verified to contain TssK by SDS-PAGE analysis. The peak fractions were combined and concentrated to 20 mg/ml (Viva spin centrifugal concentrator MWCO 30 kDa), and the protein was used for crystallisation trials upon buffer exchange into 5 mM Tris-HCl (pH 8.0) containing 50 mM NaCl using a Zeba column (Thermo).

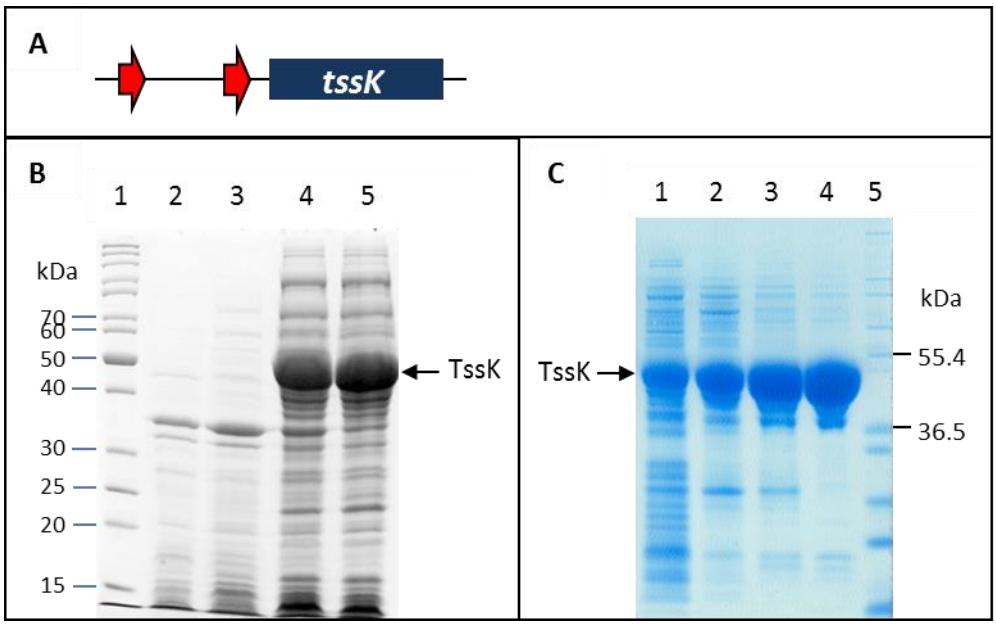


Figure 8.5 Overproduction, solubility and purification of native TssK. A. Schematic drawing showing the location of *tssK* in pACYCDuet-*tssK*. The red arrows represent the two T7 promoters present in the plasmid. **B**. Overproduction and solubility of native TssK. A Coomassie blue-stained 12% SDS-PA gel shows the analysis of proteins following induction of pACYCDuet-*tssK* in *E. coli* strain BL21 (λDE3) cells with 1 mM IPTG at 37°C. The arrow indicates the expected location of TssK based on its size (~50 kDa). Lane 1, EZ-Run™ Rec protein ladder (Fisher); lane 2, total cell protein from uninduced cells; lane 3, total cell protein from cells following induction; lane 4, crude cell lysate containing both insoluble and soluble proteins following induction and cell lysis; lane 5, soluble fraction of cell lysate following induction. **C**. 4-12% gradient SDS-PA gel (Novex) showing the purification of native TssK. Soluble fraction of cell lysate containing overproduced TssK was subjected to anion exchange chromatography on a 5 ml DEAE column (GE Healthcare) in buffer containing 50 mM Tris-HCl (pH 8.0). Elution was performed in the same buffer containing an increasing gradient concentration of NaCl up to 500 mM. The material was then subjected to 1.4 M ammonium sulfate precipitation and SEC on a Superdex 200 column. The arrow indicates the expected location of TssK based on its size (~50 kDa). Lane 1, soluble fraction of cell lysate following induction; lane 2, combined elution fractions from DEAE column containing TssK; lane 3, resuspended protein pellet after 1.4 M ammonium sulfate precipitation; lane 4, combined peak elution fractions from Superdex 200 column; lane 5, protein marker (Novex).

8.3.2.3 X-ray crystallography of TssK

To obtain structural information on TssK, native TssK was purified as described in Section 8.3.2.2. Optimum conditions for crystal formation were identified by Qiagen screens (Section 2.8.2). A crystal was formed in one of the screened conditions, i.e. 0.1 M MES (pH 6.5) and 1.5 M NaCl after a few months. However, the crystal was very thin, and did not survive in the cryo conditions. Two different shapes of crystals formed. Data collected from both, but structures couldn't be solved with the data was collected. No other crystals were obtained afterwards.

8.3.3 Generation of anti-TssK antibody

Anti-TssK antibody was raised in rats at Bioserve UK (University of Sheffield) using $_{His6}$.TssK (Section 2.6.4). $_{His6}$.TssK was purified as described in Section 8.3.1.2. First bleed blood samples were collected and used for immuno-detecting TssK that was overproduced from pACYCDuet-*tssK* in *E. coli* strain BL21 (λDE3). Different dilutions of the serum ranging from 100 to 5,000 times dilution were used as primary antibodies and commercial anti-rat antibody was used as the secondary antibody. The result showed that the first bleed blood serum contained a low titre of anti-TssK antibody (Figure 8.6 A). Therefore, a boost injection was conducted. The final bleed blood serum gave rise to clear TssK-cross reacting bands with serum dilution varying from 1 in 100 to 1 in 5,000 (Figure 8.6 B).

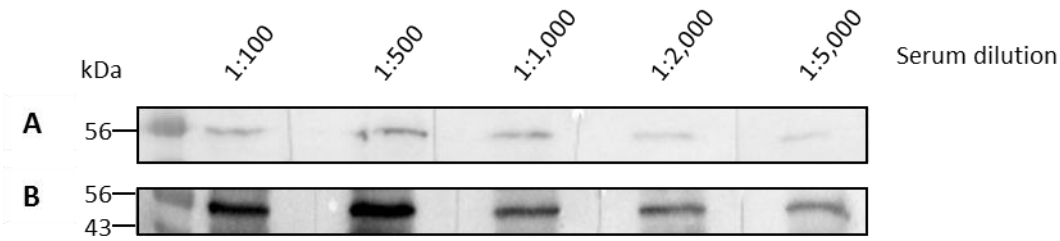


Figure 8.6 Western blots showing the detection of TssK by anti-TssK antibody raised in rats. Native TssK was overproduced following induction of pACYCDuet-*tssK* in *E. coli* strain BL21 (λDE3) with 1 mM IPTG at 37^{o}C. The soluble fraction of the cell lysate following induction was separated by electrophoresis in a 12% SDS-PA gel and immunodetected with anti-TssK antibody raised to purified $_{His6}$.TssK in rats. Primary anti-TssK antibody was diluted in different ratios as indicated, and 5,000x dilution of HRP-conjugated anti-rat secondary antibody was used for all the samples. **A**. Immuno-detection of TssK using primary bleed blood serum. **B**. Immuno-detection of TssK using final bleed blood serum.

8.4 Discussion

TssE has been reported to have a significant amino acid sequence similarity (~40%) to bacteriophage T4 gp25 (Shalom et al. 2007; Leiman et al. 2009; Lossi et al. 2011). Gp25 is a constituent of the bacteriophage baseplate and was proposed to have a lysozyme-like activity (Kostyuchenko et al. 2003; Yap et al. 2010). The function of gp25 is to interact with the tail spike $(gp27\text{-}gp5)_3$ complex in the extended state or the tail tube (gp19) when the injection apparatus is in the contracted state (Kostyuchenko et al. 2003; Kostyuchenko et al. 2005; Leiman et al. 2010). The presence of a gp25 homologous protein, TssE, in the T6SS suggests it may contribute in the assembly of T6SS baseplate by interacting with the trail spike (TssI) and trail tube (TssD) (Leiman et al. 2009; Lossi et al. 2011; Cascales and Cambillau 2012). It was subsequently confirmed by Brunet et al. (2015) that TssE is a baseplate component of T6SS. Although the reported gp25 structure from *Geobacter sulfurreducens* facilitates the prediction of the structure of TssE (Lossi et al. 2011; Cascales and Cambillau 2012), its structure has not been solved.

In the bacteriophage T4 baseplate, there is no evidence of self-association of gp25 in solution (Yap et al. 2010). There are six copies of gp25 in the bacteriophage T4 baseplate through network interactions with other baseplate components, i.e. gp6 and gp53 (Yap et al. 2010; Leiman and Shneider 2012). TssE has been shown to self-associate by the BACTH assay in a previous study (Shastri 2011). However, in this study, an MBP-TssE fusion protein was shown to be a monomer in solution. It is possible that the oligomerisation of TssE needs other T6SS baseplate subunits to present, such as TssF and/or TssG, as high levels of co-organisation of the *tssE*, *tssF* and *tssG* genes within T6SS gene clusters were observed and the interactions among these proteins have been demonstrated (Brunet et al. 2015; this study). In addition, TssF and TssG have been shown to be able to self-associate (Brunet et al. 2015; this study) and have similarities with bacteriophage T4 baseplate proteins gp6 and gp7, respectively (Taylor et al. 2016).

TssE was largely present in the insoluble fraction in solution when it was overproduced on its own without solubilizing tags. MBP has been shown to have excellent solubilisation properties (Fox and Waugh 2003), which was employed for solublising TssE. When TssE was fused to MBP, the fusion protein did exhibit excellent solubility. However, it was not possible to obtain large amount of soluble TssE upon the MBP tagged being cleaved. Moreover, despite a lot of effort on increasing its solubility using

various chemicals, including arginine and glutamic acid, which have been shown to be effective in preventing protein aggregation and precipitation (Golovanov et al. 2004), TssE was still largely insoluble. This avenue of investigation was eventually discontinued, due to time constraints.

While the work was in progress, a couple of studies on TssK were published. TssK derived from EAEC was shown to assemble into trimers that exhibit a three-armed shape by TEM and small angle X-ray scattering (SAXS) analysis (Zoued et al. 2013). The result is consistent with the observations in this study in which TssK is shown to be a trimer by SEC-MALS analysis. In addition, TssK was proposed to be a baseplate component and has been shown to interact with tail sheath (TssC), tail tube (TssD), cytoplasmic domains of inner membrane proteins TssL and TssM which form a trans-membrane chamber along with TssJ, and TssA (Zoued et al. 2013; Zoued et al. 2016; this study). Unfortunately, the crystal structure of TssK was not obtained.

Chapter 9 Final discussion

9.1 General discussion

The overall aim of this study was to characterise two types of TssA, i.e. $TssA^{S}$ and $TssA^{EI}$, a core structural component of the T6SS, and to identify and verify novel interactions between both types of TssA and other T6SS subunits *in vivo* and *in vitro*. This would help to determine the location of TssA in the T6SS complex and thereby indicate its potential role. $TssA^{S}$ and $TssA^{E}$ contain a conserved N-terminal region of approximately 60-80 amino acids but diverse in the nature of their C-terminal regions. TssA is present in at least one subunit encoded in all T6SS gene clusters. All T6SS gene clusters encode either $TssA^{S}$ or $TssA^{E}$ but never both. $TssA^{EI}$ is a sub-class of $TssA^{E}$ orthologues which lacks the C-terminal extension of 20-40 amino acids. In agreement with previous studies (Shastri 2011; Ahmad 2013), $TssA^{S}$ from *B. cenocepacia* and $TssA^{EI}$ from *A. hydrophila* are suggested to play an important role in the assembly of the T6SS as they interact with many other T6SS subunits from *B. cenocepacia*, i.e. TssB, -C, -D, -F, -H, -I, -K and -L using the BACTH assays. In addition, *in vivo* interactions of $TssA^{S}$ with TssE and $TssM_{NTD}$ were also observed. Although A.h. $TssA^{EI}$ does not appear to interact with B.c. TssE *in vivo* (Ahmad 2013), an interaction between $TssA^{EII}$ and TssE has been shown when both subunits were derived from the same T6SS, i.e. T6SS-1 of enteroaggregative *E. coli* (EAEC) (Zoued et al. 2016). As the maltose phenotype that is observed on MacConkey-maltose agar lacks sensitivity and is not quantitative, β-galactosidase assays were carried out for obtaining quantitative data. The results largely supported the maltose phenotypes that were observed on MacConkey-maltose agar. In addition, biochemical evidence was provided to support these *in vivo* results which showed $TssA^{S}$ and $TssA^{EI}$ behaved similarly with respect to their interactions with other T6SS subunits *in vitro*, i.e. TssC, TssE, TssF, TssI, TssK, TssL, $TssM_{NTD}$. $TssA^{S}$ interacts with one additional T6SS subunit (i.e. TssD) that $TssA^{EI}$ does not appear to interact with.

$TssA^{S}$ and $TssA^{EI}$ were also structurally characterised in this study. Both types of TssA oligomerise into high molecular mass complexes. They were observed to form ring-like structures as visualised by TEM. However, a distinct $TssA^{S}_{CTD}$ ring is surrounded by disassociated $TssA^{S}_{NTD}$ projections whereas $TssA^{EI}$ has a broader thickness with an irregular outer edge and the regions corresponding to the NTD and CTD are not distinguishable. In addition, the diameter of $TssA^{EI}$ is half that of $TssA^{S}$. The domain organisation analysis of $TssA^{S}$ and $TssA^{EI}$ suggested they both contain a NTD, MD and a small CTD which serves as the oligomerisation module in both types of TssA. The $TssA^{S}$

oligomer was estimated to be composed of 30 subunits by SEC-MALS analysis, whereas the $TssA^{EI}$ complex oligomerises as a decamer.

Instead of having 30 subunits as estimated by SEC-MALLS, the crystal structure of $TssA^{S}_{CTD}$ showed that it contains 32 subunits organised into a dual-layered ring with 16 subunits in each layer, whereas $TssA^{EI}_{CTD}$ is packed into a single layer with 10 subunits. Although both complexes have a central pore, the sizes of the pores are different and the individual domains are structurally dissimilar. For example, there are four alpha-helices in $TssA^{S}_{CTD}$ instead of five in $TssA^{EI}_{CTD}$ and the helices are organised completely differently (Figure 9.1). In addition, the diameter of the $TssA^{EI}_{CTD}$ decamer is half that of the $TssA^{S}_{CTD}$ ring complex.

While this work was in progress, the structure of EAEC $TssA^{EII}_{CTD}$ (Ec042_4540) was solved in its 'closed' state (Zoued et al. 2016), which belongs to a different $TssA^{E}$ subclass ($TssA^{EII}$) than *A. hydrophila* $TssA^{EI}$ (AHA1844). Although the overall size of $TssA^{EI}_{CTD}$ is largely in agreement with that of $TssA^{EII}_{CTD}$ in terms of its outer and inner dimensions, the packing patterns of them are different. For example, $TssA^{EII}_{CTD}$ features two stacked hexaflexagons (Figure 6.37 B) unlike the single layer that is formed by $TssA^{EI}_{CTD}$ (Zoued et al. 2016). Moreover, the conserved WEP motif within the $TssA^{E}$ orthologues is involved in the inter-ring interaction in $TssA^{EI}_{CTD}$ but intra-ring interaction in $TssA^{EII}_{CTD}$ (Figures 6.33 D and 6.37 B). The difference in the CTD structures between these two sub-classes is possibly due to the presence of two additional helices in the C-terminal extension among the orthologues belonging to the $TssA^{EII}_{CTD}$ sub-class locating to the exterior of the complex. Although both $TssA^{S}_{CTD}$ and $TssA^{EII}_{CTD}$ monomers assemble bilayered complexes, there is no similarity between these two structures in terms of their packing.

It has been determined that the tail tube protein TssD assembles into hexameric rings from diverse bacteria that possess a T6SS (Mougous et al. 2006; Jobichen et al. 2010; Osipiuk et al. 2011; Douzi et al. 2014). However, the cell envelope chamber TssJLM complex has been shown to have a five-fold symmetry (Durand et al. 2015). Interestingly, $TssA^{EI}_{CTD}$ also possesses a five-fold symmetry, consistent with that of the cell envelope chamber complex from a species that having a subclass EII TssA. However, how the inconsistency of the symmetry among T6SS components allows them to be assembled successfully and functionally remains to be elucidated. It might be the case that the

crystal structure of the CTD of these three TssA proteins may not reflect the situation in solution or a low-energy structure was formed during crystallisation in the absence of other T6SS components. Co-crystallisation of TssA with other T6SS components may help to solve the potential inconsistency.

The current model of the T6SS is proposed based on evolutionary, structural and mechanistic similarities between several structural proteins of the T6SS and the gene products (gp) that make up the injection machinery of some contractile bacteriophages. $TssA^{EII}$ is proposed to be analogous to the bacteriophage T4 gp15/gp3 tail terminator complex as they have very similar dynamics and dimensions (Pell et al. 2009b; Fokine et al. 2013). The tail terminator complex of bacteriophage T4 adopts conformational changes, i.e. closed and open, associated with the extended and contracted sheath conformations (Leiman and Shneider 2012). The diameters of the pores formed by the CTD complexes of $TssA^{E}$ and $TssA^{S}$ are 30 Å and 116 Å, respectively. Similar to $TssA^{EII}$, $TssA^{EI}$ is proposed to undergo a conformational change, i.e. closed and open state, to allow the incorporation of TssD hexamers during tube polymerisation. However, how and whether $TssA^{S}$ would undergo similar conformational changes is still unclear.

Like phage T4 gp15/gp3, EAEC $TssA^{EII}$ has been shown to have a dynamic role in T6SS assembly (Zoued et al. 2016). It is recruited to the T6SS membrane chamber complex TssJLM at the early stage of the T6SS assembly, initiating and guiding the polymerisation of the TssD tail tube and TssBC tail sheath, probably by incorporating TssD and TssBC building blocks (Zoued et al. 2016). EAEC $TssA^{EII}$ locates within the baseplate initially and remains at the baseplate-distal end of the tubules during elongation and is released after sheath contraction (Zoued et al. 2016). $TssA^{S}$ and $TssA^{EI}$ share similar NTDs probably due to the fact that they are involved in similar interactions with other T6SS components, including components of the cell envelope chamber complex that is located at the IM ($TssM_{NTD}$ and TssL), the baseplate (TssE, TssF, TssK and TssI) and the tail sheath (TssB and TssC), and may both initially associate with the T6SS at the same position, i.e. at the baseplate. However, rather than remaining at the baseplate-distal end of the contractile tube during its polymerisation, as $TssA^{EI}$ is proposed to, $TssA^{S}$ may stay associated with the baseplate during the apparatus assembly (Figure 9.2). The dynamics of $TssA^{S}$ and $TssA^{EI}$ during the T6SS assembly still needs to be experimentally addressed.

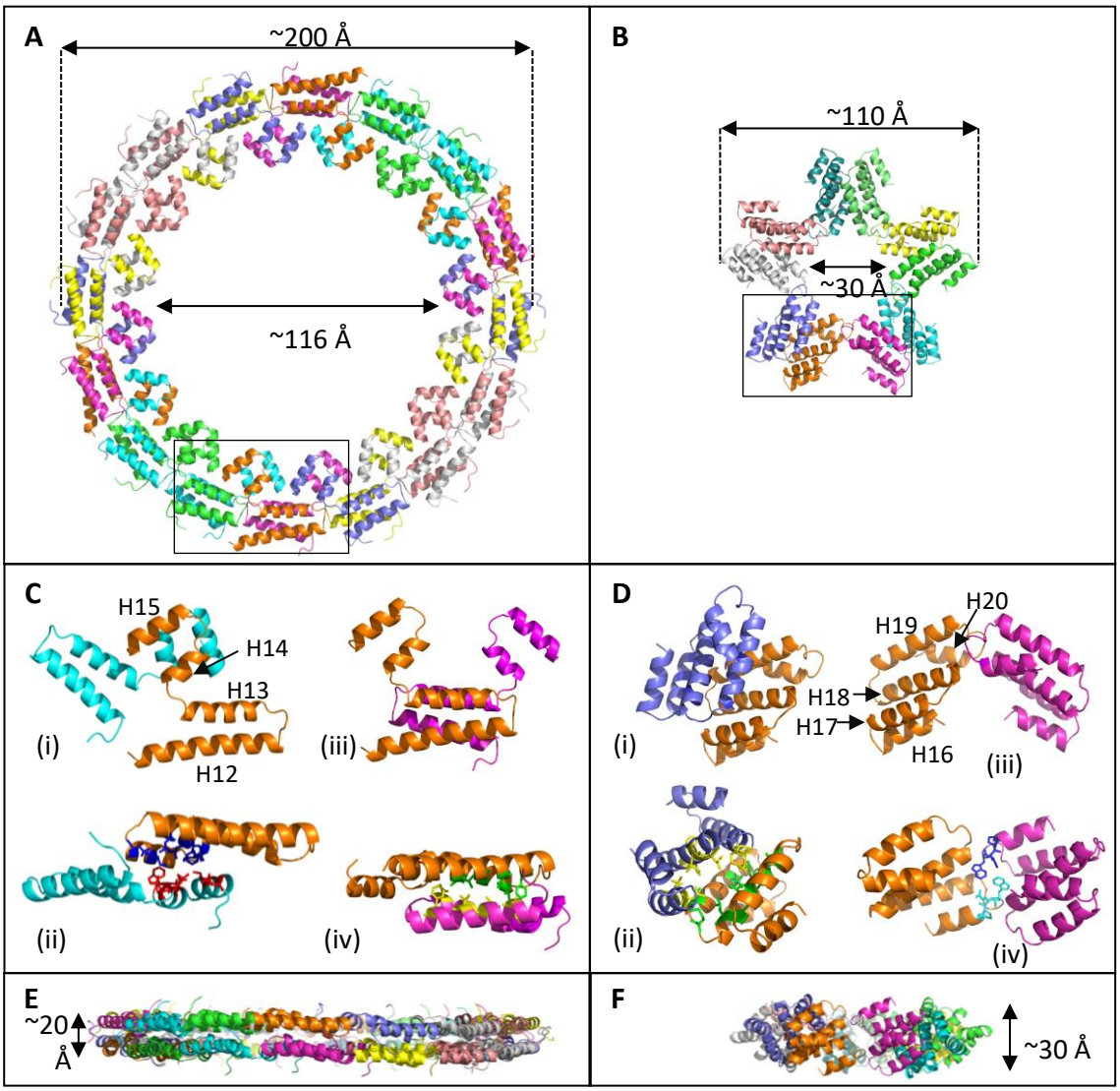


Figure 9.1 Structures of $TssA^{S}_{CTD}$ and $TssA^{EI}_{CTD}$. Subunit chains are shown in different colours. **A**. Top view of $TssA^{S}_{CTD}$. **B**. Top view of $TssA^{EI}_{CTD}$. **C**. Hydrophobic interactions between adjacent chains of $TssA^{S}_{CTD}$ for its ring packing. Side-chains of residues involved in hydrophobic interactions are shown as sticks using different colours. (i) and (ii), hydrophobic interactions between H14 and H15 of adjacent chains in different layers; (iii) and (iv), hydrophobic interactions between H12 and H13 of adjacent chains in different layers. **D**. Interactions between adjacent chains of $TssA^{EI}_{CTD}$ for its ring packing. Residues involved in interactions are shown as sticks using different colours. (i) and (ii), hydrophobic interactions between H17-H19 within two adjacent chains constituting an 'arm' of the pentameric star; (iii) and (iv), conserved WEP motif for inter-docking of adjacent arms of the star-shaped complex. **E**. Side view of $TssA^{S}_{CTD}$. **F**. Side view of $TssA^{EI}_{CTD}$.

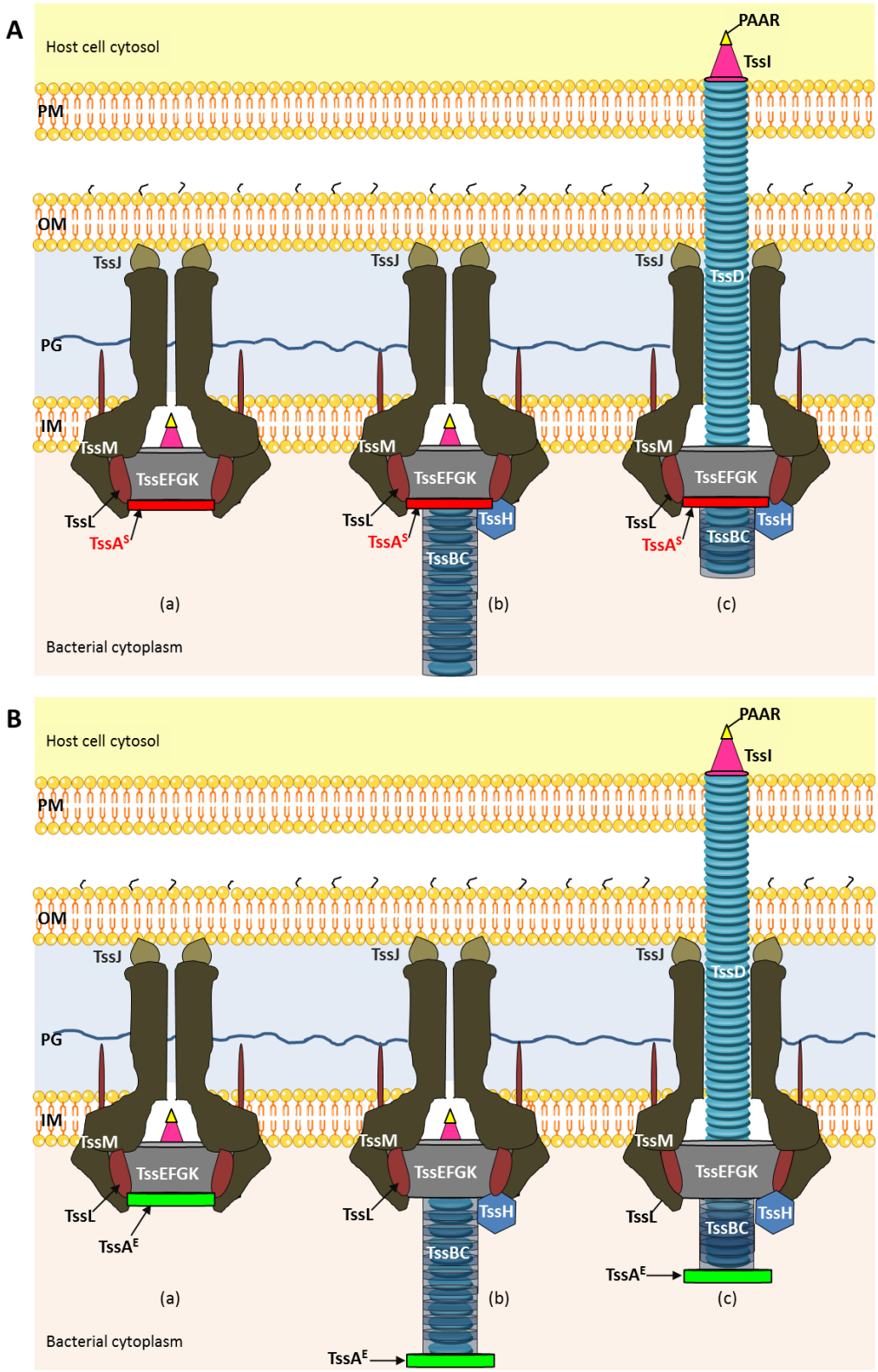


Figure 9.2 Diagrammatic representation of the T6SS apparatus. The diagram shows the proposed $TssA^S$ (**A**) and $TssA^E$ (**B**) positions in the T6SS. (a) Assembly of the T6SS membrane complex and baseplate; (b) fully assembled T6SS primed for firing (extended sheath); (c) fully assembled T6SS post-firing (contracted sheath). The predicted locations of TssA are as indicated. PM, plasma membrane of target eukaryotic cell; OM, outer membrane; PG, peptidoglycan; IM, inner membrane.

9.2 Future directions

Although this study provided detailed insights into the structures of one of the least characterised T6SS components, i.e. TssA, there are still several areas that can be further explored to fully elucidate the T6SS. Two subclasses of $TssA^{E}$ protein pack in different patterns. Orthologues belonging to these two $TssA^{E}$ subclasses need to be further investigated to determine whether the different packing patterns observed from the two representative examples are common features of the different subclasses. Moreover, how the inconsistency of the symmetries among T6SS components allows them to be assembled successfully and functionally remains to be elucidated. Co-purification and crystallisation of TssA with additional T6SS components may help to solve the potential inconsistency. In addition, as shown in Chapter 6, the introduction of three pBBR1-based plasmids by containing two $TssA^{EI}$-$TssA^{S}$ hybrid proteins and $TssA^{EI}$ did not successfully restore TssD secretion to a *B. cenocepacia* H111 *tssA*^{S}-deficient mutant strain. It is possibly due to the TssD specificity, i.e. TssD from the same species (i.e. *A. hydrophila*) is required for a functional T6SS, as $TssA^{EI}$ did not interact with *B. cenocepacia* TssD *in vitro*. Therefore, the interaction analysis between $TssA^{EI}$ and TssD from *A. hydrophila* should be analysed. Apart from that, an *A. hydrophila tssA*^{EI}-deficient mutant strain should also be constructed for investigating the T6SS activity, and the ability of $TssA^{EI}$ as well as $TssA^{S}$ in restoring the T6SS can also be carried out (complementation). Moreover, the ability of the two $TssA^{EI}$-$TssA^{S}$ hybrid proteins to form ring-like oligomers still needs to be experimentally investigated. Finally, the dynamics of $TssA^{S}$ and $TssA^{EI}$ during T6SS assembly also needs to be analysed, possibly by fluorescence lifetime imaging.

Appendices

Appendix 1 Bacterial strains used in this study

Strain	Genotype	Source
Escherichia coli strains		
MC1061	*hsdR araD*139 Δ(*ara-leu*)$_{7697}$ Δ*lacX74 galU galK rpsL* (Sm^R)	(Casadaban and Cohen 1980)
JM83	*F- ara Δ(lac-proAB) rpsL [Φ80, lacZΔM15] thi* (Sm^R)	(Yanisch-Perron et al. 1985)
S17-1(λpir)	*thi proA hsdR recA* RP4-2-*tet*::Mu-1 *kan*::Tn*7* λpir (Tp^R, Sm^R)	(Simon et al. 1983)
BL21(λDE3)	F$^-$ *omp*T *hsd*S$_B$ ($r_B^- m_B^+$) *dcm gal* λDE3	(Studier and Moffatt 1986)
C41(λDE3)	Spontaneous mutant of BL21(λDE3)	(Miroux and Walker 1996)
C43(λDE3)	Spontaneous mutant of BL21(λDE3)	(Miroux and Walker 1996)
NEB express	*fhuA2 [lon] ompT gal sulA11 R(mcr-73::miniTn10--Tet^S)2 [dcm] R(zgb-210::Tn10--Tet^S) endA1 Δ(mcrC-mrr)114::IS10.*	Enhanced BL21 derivative generated by New England Biolabs.
BTH101	F$^-$ *cya-99 araD139 galE15 galK16 hsdR2 mcrA1 mcrB1 rpsL1* (Sm^R)	(Karimova et al. 1998)
Burkholderia cenocepacia strains		
H111	CF isolate from a patient at the Medizinische Hochschule Hannover (1988)	(Romling et al. 1994)
H111-*tssA*^{S}::Tp	H111 mutant with *dfrB2* (Tp^R) cassette inserted into *tssA* gene in reverse orientation to *tssA*.	(Shastri 2011)
J2315	CF isolate from patient in Edinburgh Hospital, ET12 lineage	(Isles et al. 1984; Holden et al. 2009)
***Aeromonas hydrophila* strain**		
ATCC 7966	*Aeromonas hydrophila* subsp. *hydrophila*, isolate from tin of milk with a fishy odour	(Seshadri et al. 2006)

Abbreviations: Tp^R trimethoprim-resistant; Sm^R, streptomycin resistant

Appendix 2 Plasmids used in this study

Plasmid	Relevant characteristics	Source or reference
Vectors		
pBBR1MCS1	Broad-host range cloning vector, pBBR1-replicon, lacZα MCS (Cm^R)	(Kovach et al. 1994)
pBBR1MCS2	Broad host range cloning vector, containing pBluescript II KS-*lacZα* MCS., Tra^-, Mob^+ (Km^R)	(Kovach et al. 1995)
pET14b	T7 promoter-based expression vector; *E. coli* specific for creating fusion proteins with N-terminal Hexa-histidine tags (Ap^R)	(Rosenberg et al. 1987)
pET28a	T7 promoter-based expression vector; *E. coli* specific for creating fusion proteins with N or C-terminal Hexa-histidine tags (Km^R)	Novagen
pACYCDuet-1	Dual T7 promoter expression vector, P15A origin of replication (Cm^R)	Novagen
pETDuet-1	Dual T7 promoter expression vector, the pBR322-derived ColE1 origin of replication (Ap^R)	Novagen
pETDuetΔO-1	pETDuet-1 vector lacking the *lac* operator sequences located downstream of each T7 promotor (Ap^R)	This study
pMAL-c5X	*E. coli* vector containing ampicillin resistance genes (*bla*) and *malE* gene (encoding MBP) under control of the *tac* promoter, MCS located at 3' end of *malE* (Ap^R)	New England Biolabs
pMAL-c5X-His_6	pMAL-c5X vector modified with a hexa-histidine tag coding sequence cloned into the N-terminal coding region of the *malE* gene between the *Mfe*I and *Bgl*II sites (Ap^R)	Mosby and Thomas, unpublished results
pKT25	BACTH vector, encodes the T25 fragment (amino acids 1–224 of CyaA) with a MCS at the C-terminal coding end of T25 (Km^R)	(Karimova et al. 1998)
pKNT25	BACTH vector, encodes the T25 fragment (amino acids 1–224 of CyaA) with a MCS at the N-terminal	(Karimova et al. 1998)

	coding end of T25 (Km^R)	
pUT18	BACTH vector, encodes the T18 fragment (amino acids 225–399 of CyaA) with a MCS at the N-terminal coding end of T18 (Amp^R)	(Karimova et al. 1998)
pUT18C	BACTH vector, encodes the T18 fragment (amino acids 225–399 of CyaA) with a MCS at the C-terminal coding end of T18 (Amp^R)	(Karimova et al. 1998)
Bacterial adenylate cyclase two-hybrid plasmids		
pKT25-*tssA*^{S}	pKT25 containing *tssA*^{S} of *B. cenocepacia*, between the *Pst*I and *Xba*I sites (Km^R)	Shastri 2011
pKNT25-*tssA*^{S}	pKNT25 containing *tssA*^{S} cloned between the *Hin*dIII and *Bam*HI sites (Km^R)	Shastri 2011
pUT18C-*tssA*^{S}	pUT18C containing *tssA*^{S} cloned between the *Pst*I and *Xba*I sites (Amp^R)	Shastri 2011
pUT18-*tssA*^{S}	pUT18 containing *tssA*^{S} cloned between the *Hin*dIII and *Bam*HI sites (Amp^R)	Shastri 2011
pKT25-linkerHis_6.*tssA*^{S}	pKT25 containing linkerHis_6.*tssA*^{S} of *B. cenocepacia* cloned between the *Nde*I and *Bam*HI/*Bgl*II sites (Km^R)	Hall and Thomas, unpublished results
pKT25-*tssA*$^S_{NTD}$	pKNT25 encoding N-terminal domain (residues 1-256) of $TssA^S$ cloned between the *Hin*dIII and *Bam*HI sites (Km^R)	Shastri 2011
pKNT25-*tssA*$^S_{NTD}$	pKNT25 encoding N-terminal domain (residues 1-256) of $TssA^S$ cloned between the *Hin*dIII and *Bam*HI sites (Km^R)	Shastri 2011
pUT18C-*tssA*$^S_{NTD}$	pUT18C encoding N-terminal domain (residues 1-256) of $TssA^S$ cloned between the *Pst*I and *Bam*HI sites (Ap^R)	Shastri 2011
pUT18-*tssA*$^S_{NTD}$	pUT18 encoding N-terminal domain (residues 1-256) of $TssA^S$ cloned between the *Hin*dIII and *Bam*HI sites (Ap^R)	Shastri 2011
pKT25-*tssA*$^S_{CTD}$	pKT25 encoding C-terminal	Ahamd 2013

	domain (residues 294-373) of $TssA^S$ cloned between the *Hin*dIII and *Bam*HI sites (Km^R)	
pKNT25-$tssA^S_{CTD}$	pKNT25 encoding C-terminal domain (residues 294-373) of $TssA^S$ cloned between the *Hin*dIII and *Bam*HI sites (Km^R)	Shastri 2011
pUT18C-$tssA^S_{CTD}$	pUT18C encoding C-terminal domain (residues 294-373) of $TssA^S$ cloned between the *Pst*I and *Bam*HI sites (Ap^R)	Shastri 2011
pUT18-$tssA^S_{CTD}$	pUT18 encoding C-terminal domain (residues 294-373) of $TssA^S$ cloned between the *Hin*dIII and *Bam*HI sites (Ap^R)	Shastri 2011
pKT25-*tssB*	pKT25 containing *tssB* cloned between the *Xba*I and *Eco*RI sites (Km^R)	Shastri 2011
pKNT25-*tssB*	pKNT25 containing *tssB* cloned between the *Sph*I and *Xba*I sites (Km^R)	Shastri 2011
pUT18C-*tssB*	pUT18C containing *tssB* cloned between the *Xba*I and *Eco*RI sites (Ap^R)	Shastri 2011
pUT18-*tssB*	pUT18 containing *tssB* cloned between the *Sph*I and *Xba*I sites (Ap^R)	Shastri 2011
pKT25-*tssC*	pKT25 containing *tssC* cloned between the *Pst*I and *Xba*I sites (Km^R)	Shastri 2011
pKNT25-*tssC*	pKNT25 containing *tssC* cloned between the *Hin*dIII and *Bam*HI sites (Km^R)	This study
pUT18C-*tssC*	pUT18C containing *tssC* cloned between the *Pst*I and *Xba*I sites (Ap^R)	This study
pUT18-*tssC*	pUT18C containing *tssC* cloned between the *Hin*dIII and *Bam*HI sites (Ap^R)	This study
pKT25-*tssD*	pKT25 containing *tssD* cloned between the *Pst*I and *Xba*I sites (Km^R)	Shastri 2011
pUT18C-*tssD*	pUT18C containing *tssD* cloned between the *Pst*I and *Xba*I sites	Shastri 2011

	(Ap^R)	
pKT25-*tssE*	pKT25 containing *tssE* cloned between the *Xba*I and *Eco*RI sites (Ap^R)	Shastri 2011
pKNT25-*tssE*	pKNT25 containing *tssE* cloned between the *Hin*dIII and *Bam*HI sites (Km^R)	Shastri 2011
pUT18C-*tssE*	pUT18C containing *tssE* cloned between the *Xba*I and *Eco*RI sites (Ap^R)	Shastri 2011
pUT18-*tssE*	pUT18 containing *tssE* cloned between the *Hin*dIII and *Bam*HI sites (Ap^R)	Shastri 2011
pKT25-*tssF*	pKT25 containing *tssF* cloned between the *Xba*I and *Kpn*I sites (Km^R)	Shastri 2011
pKNT25-*tssF*	pKNT25 containing *tssF* cloned between the *Hin*dIII and *Xba*I sites (Km^R)	Shastri 2011
pUT18C-*tssF*	pUT18C containing *tssF* cloned between the *Xba*I and *Kpn*I sites (Ap^R)	Shastri 2011
pUT18-*tssF*	pUT18 containing *tssF* cloned between the *Hin*dIII and *Xba*I sites (Ap^R)	Shastri 2011
pKT25-*tssG*	pKT25 containing *tssG* cloned between the *Xba*I and *Eco*RI sites (Km^R)	Shastri 2011
pKNT25-*tssG*	pKNT25 containing *tssG* cloned between the *Hin*dIII and *Bam*HI sites (Km^R)	This study
pUT18C-*tssG*	pUT18C containing *tssG* cloned between the *Xba*I and *Eco*RI sites (Ap^R)	Shastri 2011
pUT18-*tssG*	pUT18 containing *tssG* cloned between the *Hin*dIII and *Bam*HI sites (Ap^R)	This study
pKT25-*tssH*	pKT25 containing *tssH* cloned between the *Xba*I and *Kpn*I sites (Km^R)	Ahmad 2013
pUT18C-*tssH*	pUT18C containing *tssH* cloned between the *Xba*I and *Kpn*I sites (Ap^R)	Ahmad 2013
pKNT25-*tssI*	pKNT25 containing *tssI* cloned	Ahmad 2013

	between the *Xba*I and *Sma*I sites (Km^R)	
pUT18C-*tssI*	pUT18C containing *tssI* cloned between the *Xba*I and *Sma*I sites (Ap^R)	Ahmad 2013
pKT25-$tssI_{gp27gp5}$	pKNT25 containing $tssI_{gp27gp5}$ cloned between the *Xba*I and *Bgl*II/*Bam*HI sites (Km^R)	Ahmad 2013
pUT18C-$tssI_{gp27gp5}$	pUT18C containing $tssI_{gp27gp5}$ cloned between the *Xba*I and *Bgl*II/*Bam*HI sites (Ap^R)	Ahmad 2013
pKT25-$tssI_{gp27}$	pKNT25 containing $tssI_{gp27}$ cloned between the *Xba*I and *Bgl*II/*Bam*HI sites (Km^R)	Ahmad 2013
pUT18C-$tssI_{gp27}$	pUT18C containing $tssI_{gp27}$ cloned between the *Xba*I and *Bgl*II/*Bam*HI sites (Ap^R)	Ahmad 2013
pKT25-$tssI_{gp5}$	pKNT25 containing $tssI_{gp5}$ cloned between the *Xba*I and *Bgl*II/*Bam*HI sites (Km^R)	Ahmad 2013
pUT18C-$tssI_{gp5}$	pUT18C containing $tssI_{gp5}$ cloned between the *Xba*I and *Bgl*II/*Bam*HI sites (Ap^R)	Ahmad 2013
pKT25-*tssJ*	pKT25 containing *tssJ* cloned between the *Pst*I and *Bam*HI sites (Km^R)	Ahmad 2013
pKNT25-*tssJ*	pKNT25 containing *tssJ* cloned between the *Pst*I and *Bam*HI sites (Km^R)	Ahmad 2013
pUT18C-*tssJ*	pUT18C containing *tssJ* cloned between the *Pst*I and *Bam*HI sites (Ap^R)	Ahmad 2013
pUT18-*tssJ*	pUT18 containing *tssJ* cloned between the *Pst*I and *Bam*HI sites (Ap^R)	Ahmad 2013
pKT25-*tssK*	pKT25 containing *tssK* cloned between the *Bam*HI and *Kpn*I sites (Km^R)	Ahmad 2013
pKNT25-*tssK*	pKNT25 containing *tssK* cloned between the *Bam*HI and *Kpn*I sites (Km^R)	Ahmad 2013
pUT18C-*tssK*	pUT18C containing *tssK* of *B. cenocepacia* cloned between the *Bam*HI and *Kpn*I sites (Ap^R)	Ahmad 2013

pUT18-*tssK*	pUT18 containing *tssK* of *B. cenocepacia* cloned between the *Bam*HI and *Kpn*I sites (Ap^R)	Ahmad 2013
pKT25-*tssL*	pKT25 containing *tssL* cloned between the *Xba*I and *Kpn*I sites (Km^R)	Ahmad 2013
pKNT25-*tssL*	pKNT25 containing *tssL* cloned between the *Xba*I and *Kpn*I sites (Km^R)	Ahmad 2013
pUT18C-*tssL*	pUT18C containing *tssL* of *B. cenocepacia* cloned between the *Xba*I and *Kpn*I sites (Ap^R)	Ahmad 2013
pUT18-*tssL*	pUT18 containing *tssL* of *B. cenocepacia* cloned between the *Xba*I and *Kpn*I sites (Ap^R)	Ahmad 2013
pKT25-$tssM_{NTD}$	pKT25 containing $tssM_{NTD}$ cloned between the *Xba*I and *Kpn*I sites (Km^R)	Ahmad 2013
pKNT25-$tssM_{NTD}$	pKNT25 containing $tssM_{NTD}$ cloned between the *Xba*I and *Kpn*I sites (Km^R)	This study
pUT18C-$tssM_{NTD}$	pUT18C containing $tssM_{NTD}$ cloned between the *Xba*I and *Kpn*I sites (Ap^R)	Ahmad 2013
pUT18-$tssM_{NTD}$	pUT18 containing $tssM_{NTD}$ cloned between the *Xba*I and *Kpn*I sites (Ap^R)	This study
pKT25-$tssM_{CTD}$	pKT25 containing $tssM_{CTD}$ cloned between the *Bam*HI and *Kpn*I sites (Km^R)	Ahmad 2013
pKNT25-$tssM_{CTD}$	pKNT25 containing $tssM_{CTD}$ cloned between the *Bam*HI and *Kpn*I sites (Km^R)	Ahmad 2013
pUT18C-$tssM_{CTD}$	pUT18C containing $tssM_{CTD}$ cloned between the *Bam*HI and *Kpn*I sites (Ap^R)	Ahmad 2013
pUT18C-*tssK*-*tssL*	pUT18C containing *tssK* and *tssL* cloned between the *Bam*HI and *Kpn*I sites for *tssK*. *tssL* was cloned as a fragment into the *Kpn*I restriction site of phosphatase-treated pUT18C-*tssK* vector (Ap^R)	This study
pUT18C-*tssL*-*tssK*	pUT18C containing *tssL* and *tssK*	This study

	cloned between the *Xba*I and *Kpn*I sites, and the *Kpn*I and *Sma*I/*Eco*53I sites, respectively (Ap^R)	
pUT18C-*tssF-tssG*	pUT18C containing *tssF* and *tssG* cloned between the *Xba*I and *Kpn*I sites, and the *Kpn*I and *Sma*I/*Eco*53I sites, respectively (Ap^R)	This study
pKT25-$tssA^E$	pKT25 containing $tssA^E$ cloned between the *Xba*I and *Kpn*I sites (Km^R)	Ahmad 2013
pKNT25-$tssA^E$	pKNT25 containing $tssA^E$ cloned between the *Xba*I and *Kpn*I sites (Km^R)	Ahmad 2013
pUT18C-$tssA^E$	pUT18C containing $tssA^E$ cloned between the *Xba*I and *Kpn*I sites (Ap^R)	Ahmad 2013
pUT18-$tssA^E$	pUT18 containing $tssA^E$ cloned between the *Xba*I and *Kpn*I sites (Ap^R)	Ahmad 2013
pKT25-zip	A derivative of pKT25 in which the leucine zipper is genetically fused in frame to the T25 fragment (Km^R)	(Karimova et al. 1998)
pUT18C-zip	A derivative of pUT18C in which the leucine zipper of GCN4 is genetically fused in-frame to the T18 fragment (Ap^R)	(Karimova et al. 1998)
Protein expression plasmids		
pACYCDuet-linkerHis$_6$.$tssA^S$	pACYCDuet-1 containing DNA fragment encoding $TssA^S$ with a hexa-histidine tag in the linker region (between residues 262 and 263) cloned between the *Nde*I and *Bam*HI/*Bgl*II sites (Cm^R)	Hall and Thomas, unpublished results
pACYCDuet-VSVg.*tssK*	pACYCDuet-1 containing DNA fragment encoding N-terminal VSVg-tagged TssK cloned between the *Nco*I and *Bgl*II/BamHI sites (Cm^R)	This study
pACYCDuet-VSVg.*tssK*-linkerHis$_6$.$tssA^S$	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssK cloned between the *Nco*I and *Bgl*II/BamHI sites and	This study

	TssA^{S} with a hexa-histidine tag in the linker region (between residues 262 and 263) cloned between the *Nde*I and *Bam*HI/*Bgl*II sites (Cm^{R})	
pACYCDuet-VSVg.*tssL*	pACYCDuet-1 containing DNA fragment encoding N-terminal VSVg-tagged TssL cloned between the *Nco*I and *Bgl*II/BamHI sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssL*-linkerHis$_6$.*tssA*^{S}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg tagged TssK cloned between the *Nco*I and *Bgl*II/BamHI sites and TssA^{S} with a hexa-histidine tag in the linker region (between residues 262 and 263) cloned between the *Nde*I and *Bam*HI/*Bgl*II sites (Cm^{R})	This study
pACYCDuet-HA.*tssL*	pACYCDuet-1 containing DNA fragment encoding N-terminal HA-tagged TssL cloned between the *Mfe*I and *Kpn*I sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssK*-linkerHis$_6$.*tssA*^{S}-HA.*tssL*	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssK cloned between the *Nco*I and *Bgl*II/BamHI sites, TssA^{S} with a hexa-histidine tag in the linker region (residues 262 and 263) cloned between *Nde*I and *Bam*HI/*Bgl*II sites and N-terminal HA-tagged TssL cloned between the *Mfe*I and *Kpn*I sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssI*	pACYCDuet-1 containing DNA fragment encoding N-terminal VSVg-tagged TssI cloned between the *Nco*I (blunt end generated, see text for detail) and *Bgl*II/BamHI sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssI*-LinkerHis$_6$.*tssA*^{S}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssI cloned between the *Nco*I (blunt end generated, see text for detail) and *Bgl*II/BamHI sites and TssA^{S} with a hexa-histidine tag in the linker region (residues 262 and 263) cloned between the *Nde*I and *Bam*HI/*Bgl*II sites (Cm^{R})	This study

pACYCDuet-FLAG.*tssA*^{S}	pACYCDuet-1 containing DNA fragment encoding N-terminal FLAG-tagged TssA^{S} cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-*tssC*.VSVg	pACYCDuet-1 containing DNA fragment encoding C-terminal VSVg-tagged TssC cloned between the *Nco*I and *Bam*HI sites (Cm^{R})	This study
pACYCDuet-*tssC*.VSVg-FLAG.*tssA*^{S}	pACYCDuet-1 containing DNA fragments encoding C-terminal VSVg-tagged TssC cloned between the *Nco*I and *Bam*HI sites and N-terminal FLAG-tagged TssA^{S} cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssD*	pACYCDuet-1 containing DNA fragment encoding N-terminal VSVg-tagged TssD cloned between the *Nco*I and *Bam*HI sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssD*-FLAG.*tssA*^{S}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssD cloned between the *Nco*I and *Bam*HI sites and N-terminal FLAG-tagged TssA^{S} cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-His$_6$.*tssE*	pACYCDuet-1 DNA fragment encoding N-terminal hexa-histidine tagged *tssE* cloned between the *Bam*HI and *Hin*dIII sites (Cm^{R})	Ahmad 2013
pACYCDuet-*tssE*.VSVg	pACYCDuet-1 DNA fragment encoding C-terminal VSVg-tagged *tssE* cloned between the *Nco*I and *Bgl*II sites (Cm^{R})	Thomas, unpublished data
pMAL-c5X-His$_6$-*tssE*	pMAL-c5X-His$_6$ containing DNA fragment encoding TssE cloned between the *Nde*I and *Bam*HI sites (Amp^{R})	Mosby and Thomas, unpublished results
pACYCDuet-*tssF*.HA	pACYCDuet-1 containing DNA fragment encoding C-terminal HA-tagged TssF cloned between the *Nco*I and *Hin*dIII sites (Cm^{R})	This study
pACYCDuet-*tssF*.HA-FLAG.*tssA*^{S}	pACYCDuet-1 containing DNA fragments encoding C-terminal	This study

	HA-tagged TssF cloned between the *Nco*I and *Hin*dIII sites and N-terminal FLAG tagged TssA^{S} cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	
pETDuet∆O-HA.*tssF*	pETDuet-1 vector lacking the *lac* operator sequences located downstream of each T7 promotor containing DNA fragment encoding N-terminal HA-tagged TssF between the *Nde*I and *Acc*65I sites (Amp^{R})	This study
pETDuet∆O-HA.*tssF-tssG*.His$_6$	pETDuet-1 vector lacking the *lac* operator sequences located downstream of each T7 promotor containing DNA fragments encoding N-terminal HA-tagged TssF and C-terminal hexa-histidine tagged TssG between the *Nde*I and *Kpn*I sites (Amp^{R})	This study
pACYCDuet-VSVg.*tssI*$_{gp27gp5}$	pACYCDuet-1 containing DNA fragment encoding N-terminal VSVg tagged TssI$_{gp27gp5}$ cloned between the *Bsp*HI/*Nco*I and *Bgl*II/*Bam*HI sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssI*$_{gp27gp5}$-FLAG.*tssA*^{S}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssI$_{gp27gp5}$ cloned between the *Bsp*HI/*Nco*I and *Bgl*II/*Bam*HI sites and N-terminal FLAG-tagged TssA^{S} cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssK*-FLAG.*tssA*^{S}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssK cloned between the *Nco*I and *Bgl*II/BamHI sites and N-terminal FLAG-tagged TssA^{S} cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssL*-FLAG.*tssA*^{S}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssK cloned between the *Nco*I and *Bgl*II/BamHI sites and N-terminal FLAG-tagged TssA^{S} cloned between the *Nde*I and *Bgl*II	This study

	sites (Cm^R)	
pACYCDuet-VSVg.*tssM*$_{NTD}$	pACYCDuet-1 containing DNA fragment encoding N-terminal VSVg tagged $TssM_{NTD}$ (residues 58-445) cloned between the *Nco*I and *Hin*dIII sites (Cm^R)	This study
pACYCDuet-VSVg.*tssM*$_{NTD}$-FLAG.*tssA*^{S}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged $TssM_{NTD}$ (residues 58-445) cloned between the *Nco*I and *Hin*dIII sites and N-terminal FLAG-tagged $TssA^S$ cloned between the *Nde*I and *Bgl*II sites (Cm^R)	This study
pACYCDuet-FLAG.*tssA*^{E}	pACYCDuet-1 containing DNA fragment encoding N-terminal FLAG-tagged $TssA^E$ cloned between the *Nde*I and *Bgl*II sites (Cm^R)	This study
pACYCDuet-*tssC*.VSVg-FLAG.*tssA*^{E}	pACYCDuet-1 containing DNA fragments encoding C-terminal VSVg-tagged TssC cloned between the *Nco*I and *Bam*HI sites and N-terminal FLAG-tagged $TssA^S$ cloned between the *Nde*I and *Acc*65I sites (Cm^R)	This study
pACYCDuet-VSVg.*tssD*-FLAG.*tssA*^{E}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssD cloned between the *Nco*I and *Bam*HI sites and N-terminal FLAG-tagged $TssA^S$ cloned between the *Bgl*II and *Mfe*I sites (Cm^R)	This study
pACYCDuet-*tssF*.HA-FLAG.*tssA*^{E}	pACYCDuet-1 containing DNA fragments encoding C-terminal HA tagged TssF cloned between the *Nco*I and *Hin*dIII sites and N-terminal FLAG tagged $TssA^S$ cloned between the *Nde*I and *Acc*65I sites (Cm^R)	This study
pACYCDuet-VSVg.*tssI*$_{gp27gp5}$-FLAG.*tssA*^{E}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged $TssI_{gp27gp5}$ cloned between the *Bsp*HI/*Nco*I and *Bgl*II/*Bam*HI sites and N-terminal FLAG-tagged $TssA^E$ cloned between the *Nde*I and *Bgl*II sites (Cm^R)	This study

pACYCDuet-VSVg.*tssK*-FLAG.*tssA*^{E}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssK cloned between the *Nco*I and *Bgl*II/BamHI sites and N-terminal FLAG-tagged TssA^{E} cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssL*-FLAG.*tssA*^{E}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssK cloned between the *Nco*I and *Bgl*II/BamHI sites and N-terminal FLAG-tagged TssA^{E} cloned between the *Bgl*II and *Kpn*I sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssM*$_{NTD}$-FLAG.*tssA*^{E}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssM$_{NTD}$ (residues 58-445) cloned between the *Nco*I and *Hin*dIII sites and N-terminal FLAG-tagged TssA^{E} cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-VSVg.*tssI*$_{gp27gp5}$-His$_{10}$.*tssA*^{S}	pACYCDuet-1 containing DNA fragments encoding N-terminal VSVg-tagged TssI$_{gp27gp5}$ cloned between the *Bsp*HI/*Nco*I and *Bgl*II/*Bam*HI sites and N-terminal deca-histidine tagged TssA^{S} cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pET14b-*tssA*^{S}	pET14b containing DNA fragment encoding TssA^{S} cloned between the *Nde*I and *Bam*HI sites (Ap^{R})	Shastri 2011
pET14b-His$_{6}$.*tssA*$^{S}_{NTR}$	pET14b containing DNA fragment encoding N-terminal hexa-histidine tagged TssA$^{S}_{NTR}$ (residues 1-254) cloned between the *Nde*I and sites *Bam*HI (Ap^{R})	Ahmad 2013
pET14b-His$_{6}$.*tssA*$^{S}_{CTD}$V1a	pET14b containing DNA fragment encoding N-terminal hexa-histidine tagged TssA$^{S}_{CTD}$V1a (residues 294-373) cloned between the *Nde*I and *Bam*HI sites (Ap^{R})	Shastri 2011
pET14b-His$_{6}$.*tssA*$^{S}_{CTD}$V2a	pET14b containing DNA fragment encoding N-terminal hexa-histidine	Ahmad 2013

	tagged $TssA^{S}_{CTD}$V2a (residues 299-373) cloned between the *Nde*I and *Bam*HI sites (Ap^{R})	
pMAL-c5X-*tssA*$^{S}_{CTD}$V2b	pMALc5X containing a DNA fragment encoding $TssA^{S}_{CTD}$ (residues 299-373) cloned between the *Nde*I and *Bam*HI sites (Ap^{R})	This study
pMAL-c5X-*tssA*$^{S}_{CTD}$V3b	pMALc5X containing a DNA fragment encoding $TssA^{S}_{CTD}$ (residues 303-373) cloned between the *Nde*I and *Bam*HI sites (Ap^{R})	This study
pET14b-His_6.linkerXa.*tssA*^{S}	pET14b containing DNA fragment encoding N-terminal hexa-histidine tagged $TssA^{S}$ containing a factor Xa site between G302 and I303 cloned between the *Nde*I and *Bam*HI sites (Ap^{R})	This study
pMAL-c5X-His_6.*tssA*$^{S}_{CTD}$V3e.H12	pMAL-c5X containing DNA fragment encoding N-terminal hexa-histidine tagged $TssA^{S}_{CTD}$V3e.$_{H12}$ (residues 303-329) cloned between the *Nde*I and *Bam*HI sites (Ap^{R})	This study
pMAL-c5X-His_6.*tssA*$^{S}_{CTD}$V3e.H12-H13	pMAL-c5X containing DNA fragment encoding N-terminal hexa-histidine tagged $TssA^{S}_{CTD}$V3e.$_{H12\text{-}H13}$ (residues 303-347) cloned between the *Nde*I and *Bam*HI sites (Ap^{R})	This study
pMAL-c5X-His_6.*tssA*$^{S}_{CTD}$V3e.H12-H14	pMAL-c5X containing DNA fragment encoding N-terminal hexa-histidine tagged $TssA^{S}_{CTD}$V3e.$_{H12\text{-}H14}$ (residues 303-358) cloned between the *Nde*I and *Bam*HI sites (Ap^{R})	This study
pMAL-c5X-His_6.*tssA*$^{S}_{CTD}$V3e	pMAL-c5X containing DNA fragment encoding N-terminal hexa-histidine tagged $TssA^{S}_{CTD}$V3e (residues 303-373) cloned between the *Nde*I and *Bam*HI sites (Ap^{R})	This study
pACYCDuet-His_6.*tssA*^{E}	pACYCDuet-1 containing DNA fragment encoding N-terminal hexa-histidine tagged $TssA^{E}$ cloned between the *Bam*HI and *Hin*dIII sites (Cm^{R})	Ahmad 2013
pACYCDuet-His_6.*tssA*$^{E}_{NTD}$	pACYCDuet-1 containing DNA fragment encoding N-terminal	This study

	hexa-histidine tagged TssA$^{E}_{NTD}$ (residues 1-229) cloned between the *Bgl*II/*Bam*HI and *Kpn*I sites (Cm^{R})	
pACYCDuet-*tssA*$^{E}_{NTD}$V2	pACYCDuet-1 containing DNA fragment encoding TssA$^{E}_{NTD}$V2 (residues 1-178) cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-*tssA*$^{E}_{NTD}$V3	pACYCDuet-1 containing DNA fragment encoding TssA$^{E}_{NTD}$V3 (residues 1-158) cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-His$_6$.*tssA*$^{E}_{MD-CTD}$	pACYCDuet-1 containing DNA fragment encoding N-terminal hexa-histidine tagged TssA$^{E}_{MD-CTD}$ (residues 223-478) cloned between the *Bam*HI and *Hin*dIII sites (Cm^{R})	This study
pACYCDuet-*tssA*$^{E}_{MD-CTD}$V2	pACYCDuet-1 containing DNA fragment encoding TssA$^{E}_{MD-CTD}$V2 (residues 228-478) cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-*tssA*$^{E}_{MD-CTD}$V3	pACYCDuet-1 containing DNA fragment encoding TssA$^{E}_{MD-CTD}$V3 (residues 284-478) cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-*tssA*$^{E}_{MD-CTD}$V4	pACYCDuet-1 containing DNA fragment encoding TssA$^{E}_{MD-CTD}$V4 (residues 284-478) cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-His$_6$.linkerXa.*tssA*^{E}	pACYCDuet-1 containing DNA fragment encoding N-terminal hexa-histidine tagged TssA^{E} containing a factor Xa site between P280 and L281 cloned between the *Bgl*II/ *Bam*HI and *Hin*dIII sites (Cm^{R})	This study
pACYCDuet-His$_6$.*tssA*$^{E}_{MD}$	pACYCDuet-1 containing DNA fragment encoding N-terminal hexa-histidine tagged TssA$^{E}_{MD}$ (residues 223-387) cloned between the *Bam*HI and *Hin*dIII sites (Cm^{R})	This study
pACYCDuet-*tssA*$^{E}_{MD}$V2	pACYCDuet-1 containing DNA fragment encoding TssA$^{E}_{MD}$V2 (residues 227-375) cloned between the *Nde*I and *Bgl*II sites (Cm^{R})	This study
pACYCDuet-*tssA*$^{E}_{MD}$V3	pACYCDuet-1 containing DNA	This study

	fragment encoding $TssA^{E}_{MD}V3$ (residues 284-375) cloned between the *Nde*I and *Bgl*II sites (Cm^R)	
pACYCDuet-$tssA^{E}_{MD}V4$	pACYCDuet-1 containing DNA fragment encoding $TssA^{E}_{MD}V4$ (residues 284-375) cloned between the *Nde*I and *Bgl*II sites (Cm^R)	This study
pACYCDuet-$His_6.tssA^{E}_{CTD}$	pACYCDuet-1 containing DNA fragment encoding N-terminal hexa-histidine tagged $TssA^{E}_{CTD}$ (residues 381-478) cloned between the *Bam*HI and *Hin*dIII sites (Cm^R)	This study
pETDuet-$tssA^{E}_{CTD}V2$	pETDuet-1 containing DNA fragment encoding $TssA^{E}_{CTD}V2$ (residues 397-478) cloned between the *Nde*I and *Bgl*II sites (Amp^R)	This study
pACYCDuet-$tssA^{E}_{CTD}V3$	pACYCDuet-1 containing DNA fragment encoding $TssA^{E}_{CTD}V3$ (residues 397-478) cloned between the *Nde*I and *Bgl*II sites (Cm^R)	This study
pACYCDuet-$His_6.tssA^{E}_{CTD.H16-H19}$	pACYCDuet-1 containing DNA fragment encoding $TssA^{E}_{CTD.H16-H19}$ (residues 381-463) (Cm^R) (See text for detail)	This study
pACYCDuet-$His_6.tssK$	pACYCDuet-1 containing DNA fragment encoding N-terminal hexa-histidine tagged TssK cloned between the *Bam*HI and *Acc*65I/*Bsr*GI sites (Cm^R)	This study
pACYCDuet-*tssK*	pACYCDuet-1 containing DNA fragment encoding TssK cloned between the *Nde*I and *Bgl*II sites (Cm^R)	This study
T6SS genetic complementation plasmids		
pBBR1MCS-$tssA^{S}$	pBBR1MCS-1 containing $tssA^{S}$ from *B. cenocepacia* H111 between the *Hin*dIII and *Bam*HI sites, orientated in same direction as *lac* promoter (Cm^R)	Shastri 2011
pBBR1MCS-Hybrid 1	pBBR1MCS-1 containing a DNA fragment encoding a hybrid TssA composed of M1-P108 of $TssA^{E}$ from *A. hydrophila* ATCC 7966 fused to L114-S373 of $TssA^{E}$ from *B.*	This study

	cenocepacia cloned between the *Acc*65I and *Hin*dIII sites, orientated in same direction as *lac* promoter (Cm^R)	
pBBR1MCS-Hybrid 2	pBBR1MCS-1 containing a DNA fragment encoding a hybrid TssAcomposed of M1-N230 of $TssA^E$ from *A. hydrophila* ATCC 7966 fused to I303-S373 of $TssA^S$ from *B. cenocepacia* cloned between the *Acc*65I and *Hin*dIII sites, orientated in same direction as *lac* promoter (Cm^R)	This study
pBBR1MCS-$tssA^E$	pBBR1MCS-1 containing $tssA^E$ from *A. hydrophila* ATCC 7966 cloned between the *Acc*65I and *Hin*dIII sites, orientated in same direction as *lac* promoter (Cm^R)	This study

Abbreviations: Km^R, kanamycin resistant; Ap^R, ampicillin resistant; Cm^R, chloremphenicol resistant

Appendix 3 Primers used in this study

Primer name	Sequence 5'> 3'	Restriction site in 5' extension
BACTH primers		
Nterm-iotCfor	GCGCAAGCTTGCAGATGAACCAGCAAACGGCT	*Hind*III
Nterm-tssCrev.NEW	GCGCGGATCCTTGAATTTGCCGGTTTGGGCA	*Bam*HI
Cterminal-tssCfor.new	GCGCCTGCAGAATGAACCAGCAAACGGCT	*Pst*I
Cterm-iotCrev	GCGCTCTAGACTCGCTTATGAATTTGCCGG	*Xba*I
N-iotGfrag1for	GCGCAAGCTTGATGCAAGCCCCGAACCGGCG	*Hin*dIII
Nterminal-iotGrev2	GCGCGGATCCTCGTGAATCACGTGCAGTTCGTA	*Bam*HI
pUT18C-K-L.KpnI.for	GCGCGGTACCTAAAAAAGGAGATATACCATGAGCTACGCGCCTTCCTT	*Kpn*I
pUT18C-K-L.KpnI.rev	GCGCGGTACCTTACTTCAAGCGGTGCGCGATCT	*Kpn*I
tssKtrihyfor	GCGCGGTACCTAAAAAAGGAGATATACC ATGAGTTATTCGGCCAAGGT	*Kpn*I
tssKtrihyrev	GCGCCCCGGGTCATGATGTGACCGCGATCA	*Sma*I
tssGtrihyfor	GCGCGGTACCATAAAAAAGGAGATATACCATGCAAGCCCCGAACCGGCG	*Kpn*I
tssGtrihyrev	GCGCCCCGGGGGGTCTCAGTGAATCACGTG	*Sma*I
Protein expression plasmid primers[a]		
TssKforpET.NtermVSVGTag	GCGCCCATGGCA**TACACTGATATCGAAATGAACCGCCTGGGTAAG**ATGAGTTATTCGGCCAAGGT	*Nco*I
TssK.BglII.rev	GCGCAGATCTTCATGATGTGACCGCGATCA	*Bgl*II
TssLforpET.NtermVSVGTag	GCGCCCATGGCA**TACACTGATATCGAAATGAACCGCCTGGGTAAG**ATGAGCTACGCGCCTTCCTT	*Nco*I
TssL.Rev	GCGCAGATCTTTACTTCAAGCGGTGCGCGAT	*Bgl*II

TssL.MfeI.NcoI.HAtag.For	GCGCCAATTGTAATTAATAAGGAGA TATACCATGGCA**TACCCATACGACG TCCCAGACTACGCT**AGCTACGCGC CTTCCTTGTT	*Mfe*I and *Nco*I
TssL.KpnI.Rev	GCGCGGTACCTTACTTCAAGCGGTG CGCGATCT	*Kpn*I
TssI0667.BspHI.VSVGTag.for	GCGCTCATGACA**TACACTGATATCG AAATGAACCGCCTGGGTAAG**ATGA ACAAAACGCCGGCCGC	*Bsp*HI
VgrGrev	GCGCAGATCTACTCCAACGAGTGCG GCAAT	*Bgl*II
TssA.NdeI.FLAG.for	GCGCCATATG**GATTATAAAGACGA CGATGATAAA**CCGATCAATCTCCCC GAGCT	*Nde*I
TssA.BglII.rev	GCGCAGATCTTGCGTTTACGACTGCT CGTC	*Bgl*II
tssC.NcoI.for	GCGCCCATGGCAAACCAGCAAACGG CTGC	*Nco*I
tssC.VSVg.BamHI.rev	GCGCGGATCCTTA**CTTACCCAGGCG GTTCATTTCGATATCAGTGTA**TGA ATTTGCCGGTTTGG	*Bam*HI
tssD.forpACYC.NtermVSVgTag	GCGCCCATGGCA**TACACTGATATCG AAATGAACCGCCTGGGTAAG**ATGT TACATATG	*Nco*I
tssD.BamHI.rev	GCGCGGATCCTTAGACCGCGTAGGT CTTGT	*Bam*HI
tssF.NcoI.for	GCGCCCATGGAAGAATTGCTGCCGT ATTA	*Nco*I
tssF.HA.HindIII.rev	GCGCAAGCTTCTA**AGCGTAGTCTG GGACGTCGTATGGGTA**CGCCAGGA TCGATTCGC	*Hin*dIII
tssF.NdeI.HA.for	GCGCCATATG**TACCCATACGACGT CCCAGACTACGCT**GAAGAATTGCT GCCGTATTA	*Nde*I
tssG.Kpn.C-His.rev	GCGCGGTACCTCA**ATGATGGTGAT GGTGATG**GTGAATCACGTGCAGTTC GT	*Kpn*I
Cterminal-iotFrev	GCGCGGTACCATCGACTACGCCAGG ATCGA	*Acc*65I
TssI.BspHI.VSVgtag.For	GCGCTCATGAGC**TACACTGATATCG AAATGAACCGCCTGGGTAAG**ATGG ATGCGCACAGCATGAT	*Bsp*HI
BCAM0148gp5.BglII.rev	GCGCAGATCTTTAAAACTCCTGGGA CCGGTAGC	*Bgl*II

TssLforpET.NtermVSVGTag	GCGCCCATGGCA**TACACTGATATCG AAATGAACCGCCTGGGTAAG**ATGA GCTACGCGCCTTCCTT	*Nco*I
TssLRev	GCGCAGATCTTTACTTCAAGCGGTG CGCGAT	*Bgl*II
TssM1.VSVg.NcoI.for	GCGCCCATGGCA**TACACTGATATCG AAATGAACCGCCTGGGTAAG**GTGA AGCGTGCGAACCAGCA	*Nco*I
TssM1.HindIII.rev	GCGCAAGCTTTTAGAAGGTCGCGTA GCGCAGCC	*Hin*dIII
AHA1844.NdeI.FLAG.for	GCGCCATATG**GATTATAAAGACGA CGATGATAAA**AGCTATCAACACCCC TGGTG	*Nde*I
AHA1844.BglII.rev	GCGCAGATCTTCATTTCGACAACGG CGCCG	*Bgl*II
AHA1844.KpnI.rev	GCGCGGTACCTCATTTCGACAACGG CGCCG	*Kpn*I
AHA1844.BglII.FLAG.for	GCGCAGATCTAATG**GATTATAAAGA CGACGATGATAAA**AGCTATCAACA CCCCTGGTG	*Bgl*II
AHA1844.MunI.rev	GCGCCAATTGTCATTTCGACAACGG CGCCG	*Mfe*I
TssA.NdeI.His10.for	GCGCCATATG**CATCACCATCACCAT CATCACCATCACCAT**CCGATCAATC TCCCCGAGCT	*Nde*I
pET14b-iotArev	GCGCGGATCCTGCGTTTACGACTGCT CGTC	*Bam*HI
pET14b-ACTDfor2	GCGCCATATGACGATCGCCGGCATC CAGAA	*Nde*I
pET14b-ACTDfor3	GCGCCATATGATCCAGAACCGTGCG CAGGC	*Nde*I
TssAlinkerXa1for	ATCGAGGGAAGGATTTCACACATGA TCCAGAACCGTGCGCAGGC	-
TssAlinkerXa1rev	CATGTGTGAAATCCTTCCCTCGATGC CGGCGATCGTCTGCGTCA	-
TssAlinkerXa2for	GCCGGCATCGAGGGAAGGATCCAGA ACCGTGCGCAGGC	-
TssAlinkerXa2rev	CTGGATCCTTCCCTCGATGCCGGCGA TCGTCTGCGTCA	-
pET14b-iotAfor	GCGCCATATGCCGATCAATCTCCCC GA	*Nde*I
pACYC-tssA.CTDfor	GCGCCATATGAGCAGCCATCATCAT CATCATCACAGCATCCAGAACCGTG CGCAGGC	*Nde*I

pACYC-tssA.CTDH1rev2	GCGCGGATCCTTAGACCGGGCTGTG CGGCTCGG	*Bam*HI
pACYC-tssA.CTDH2rev2	GCGCGGATCCTTACCACTTGTGCAG CGGCATGT	*Bam*HI
pACYC-tssA.CTDH3rev2	GCGCGGATCCTTACAGCGAGCCGTC GTCCTTCA	*Bam*HI
pACYC-AHA1844.for	GCGCAGATCTAATGAGCTATCAACA CCCCTG	*Bgl*II
AHA1844NTDrev	GCGCGGTACCTTAGCTGGAACTGTC GACGTCGACG	*Kpn*I
AHA1844.NdeI.for	GCGCCATATGAGCTATCAACACCCC TG	*Nde*I
AHA1844NTDrev2	GCGCAGATCTTTAGCGCTGGGCCCG CTTCAGCC	*Bgl*II
AHA1844NTDrev3	GCGCAGATCTTTACTGCGCCAGCCA GACCCGCT	*Bgl*II
pACYC-AHA1844.CTDfor	GCGCGGATCCAGGCGTCGACGTCGA CAGTTC	*Bam*HI
pACYC-AHA1844.CTDrev	GCGCAAGCTTTCATTTCGACAACGG CGCCG	*Hin*dIII
AHA1844.CTD.NdeI.for2	GCGCCATATGAGTTCCAACGATCGT GCCTG	*Nde*I
AHA1844.CTD.NdeI.for3	GCGCCATATGTCGGCGGACATGGTG GA	*Nde*I
AHA1844.CTD.NdeI.NEW.for3	GCGCCATATGTCAGCAGATATGGTT GAC	*Nde*I
AHA1844linkerXarev	GCCAGCCTTCCCTCGATCGGGGTTTT ATTGCCCGCCC	-
AHA1844linkerXafor	AACCCCGATCGAGGGAAGGCTGGCG CCCATGTCGGCGGA	-
pACYC-AHA1844.rev	GCGCAAGCTTTCATTTCGACAACGG CGCCG	*Hin*dIII
pACYC-AHA1844.CTDSD1rev	GCGCAAGCTTTTAAGCCTCGCCAAT GCCCGCACT	*Hin*dIII
AHA1844.CTDSD1.BglII.rev2	GCGCAGATCTTTATTGCAGCCAGCG GCTGCACT	*Bgl*II
pACYC-AHA1844.CTDSD2for	GCGCGGATCCGAGTGCGGGCATTGG C GAGGC	*Bam*HI
AHA1844.CTDSD2.NdeI.for2	GCGCCATATGCACGGCGAGCAGGGG ATCGC	*Nde*I
AHA1844.CTDSD2.NdeI.NEWfor2	GCGCCATATGCACGGGGAACAAGGA ATAGCGGCGGCGCTCGCCCTGCTTG	*Nde*I
TssK.NdeI.for	GCGCCATATGAGTTATTCGGCCAAG GT	*Nde*I

TssK.BglII.rev	GCGCAGATCTTCATGATGTGACCGC GATCA	*Bgl*II
T6SS genetic complemtation plasmids primers		
pBBR1MCS.AHA1844.KpnI .for	GCGCGGTACCATAAAAAAGGAGATA TACCATGAGCTATCAACACCCCTG	*Kpn*I
pBBR1MCS.AHA1844.NTD .SD1.inter_rev	ATGTCGTCGTCCTCGGGCAGGGGCC AGGCCAGCAGCCAGT	-
pBBR1MCS.TssA.NTD.SD2 .inter_for	ACTGGCTGCTGGCCTGGCCCCTGCCC GAGGACGACGACAT	-
TssA.HindIII.rev	GCGCAAGCTTTGCGTTTACGACTGCT CGTC	*Hin*dIII
pBBR1MCS.AHA1844.NTD .inter_rev	GCCTGCGCACGGTTCTGGATGTTGG AACTGTCGACGTCGA	-
pBBR1MCS.TssA.CTD.inter for	TCGACGTCGACAGTTCCAACATCCA GAACCGTGCGCAGGC	-

[a] Epitope tag coding sequences are shown in bold font.

Appendix 4 Nucleotide sequences of *tssA*^{S} from *B. cenocepacia* H111 and *tssA*^{E} from *A. hydrophila* ATCC 7966

A. *tssA*^{S} (*B. cenocepacia* H111)

ATGccgatcaatctccccgagctgctgacgccgatcagcgaggcgtcgcccagcggcgacgacctgctgt
tctcgaacgaattcgacgcgatccaggacgcgcggcgctacgacgacccgacgctcgaccagggcgaatg
ggtgaccgagatcaaggaggccgactggggcttcgtcgtcgaccatgcgggcgagctgctgcgcacccgc
acgaaggacctgcggctcgccgtgtggctgaccgaggcgctcgcgctcgaggacggcatcaccggcctca
ccgaaggctatgcgctgctcgagggcctgtgccgcgagttctgggacaccttccacccgctgcccgagga
cgacgacatcgagcaccggctcggcaacgtcgcatggctttccggtcgcacggccgagctgctgcgcgcg
gtgccgctgaccgacggtgcatcgaatgcgttcagcacgctcgactgggaagtcgcgcagcacgtcgcgc
agtcgatcaagcgcgaccccgagcatgcggacgacatcgcgcgcggcaagccgtcgatcgagcagatcga
cgcgtcgcggcgcgtgacgtcgatcgcgttctacacggcgctgctggcgaacctgaaggcgttcgaattc
gcgctcgatgcgttcgaggagcggctcgtcgagcgcgcgggcgattcggcgccgagcttccggcaggcgc
gcgacgcgttcgagaccgtgtaccggctcgccgagcgcttcgcgcgcgaacagggctataccggcagcgc
gccgcacacgcaggcggcgcagcaggcccagcccgagcgcatcgagccggtgttcggccaaccgatccag
accgaggagactcacgtgcagcagcagaccgcttcgcgtcctccggtgacgcagacgatcgccggcatcc
agaaccgtgcgcaggccgtcgaccagttgcgtgcggtggcgcgctatttccgtcagaccgagccgcacag
cccggtcgcgtatctcgccgacaaggcggccgaatgggccgacatgccgctgcacaagtggctcgagagc
gtcgtgaaggacgacggctcgctgtcgcatattcgcgaactgctgggcgtgcggcccgacgagcagtcgT
AA

B. *tssA*^{E} (*A. hydrophila* ATCC 7966)

ATGagctatcaacacccctggtgtgcacgcctgctcaccagcctgccggatgaacagataagaggcgccg
tgctggccgacgagccgcgctgggactatgtggagaccgagctggtcaagctgggatccctcgctcacag
ccaggtcgatctcaatgcggtggccgaggcctgcctcggtttgctggagagtcgcaccaaggacatgcgg
gtgctggcccagttgctgcgctgcctgcagcaccctgccaaggcgaccccgctgggggcggcgatcagcc
tgctggaggcctggatccaggcttactggctgctggcctggcccggcaatgccagccagaaacagcggct
gatggtgcagatcgtcaagcgcttcgaaggggccctgccgcgcatttgcgagagcgcttcggcggccgag
ctggcccagctgctggcccaggccgagcagctggagcgggtctggctggcgcagtgtcccgacaaggggg
agctgctcgacccgctggtgatggggctgaagcgggcccagcgccagcagctggcacaggccgaggccaa
tgccgccgggcagccccagagcagcggcgcggcggctgcgggctccccggcctcggtggcgtcgaccgcc
agttgtgccggcgccatggtactcagcggcagtgctggcgtcgacgtcgacagttccaacgatcgtgcct
ggcgtcagacccagctcaaggtggcggagctcctcatcgagcgtcagcccgaggtggcggtaggctatcg
gctgcgtcgtcacgccgtctgggccgggatcacggcggtccccatgagcggggcgggcaataaaaccccg
ctggcgcccatgtcggcggacatggtggacgagtaccgtgccgccatgaacgcccctgatcaggggctct
ggcagcgcatcgaacagagcctgaccctggcaccctactggttcgaggggcacaggctttcggccgaggt
ggccgagaaactcggctttggggcggtggcccaggcgatcgccgaggagctcgggacttttctgcagcgc
ctgccggccctgcgggaactcgccttcagcgacggctcgccgtttctctccccgagtgcagccgctggc
tgcaaccggccaagggcggcagtgcgggcattggcgaggcggggctggccgaggaggtcgcccagcgcca
cggcgagcaggggatcgcggcggcgctcgccctgcttgatgagcggattgcgcagttgaaagagccaaga
gatcgcttccacgccttgctggtgcaggcggagctgttggcgcaggaaggcatggaggcgcttgcgcgcc
agcactatcaacacttgtggcaagaggccagtcgcctcgggttgtcgcactgggagcccgggctggtcaa
ccgcctggagagcctggcggcgccgttgtcgaaaTGA

References

Ahmad, A. 2013. "Protein-protein interactions in the bacterial type VI secretion system." In Department of Infection and Immunity: University of Sheffield.

Alami, M., I. Lüke, S. Deitermann, G. Eisner, H. G. Koch, J. Brunner and M. Müller. 2003. "Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli*." Molecular Cell 12: 937-946.

Alcoforado Diniz, J. and S. J. Coulthurst. 2015. "Intraspecies competition in *Serratia marcescens* is mediated by Type VI-secreted Rhs effectors and a conserved effector-associated accessory protein." Journal of Bacteriology 197: 2350-2360.

Altwegg, M. and H. K. Geiss. 1989. "*Aeromonas* as a human pathogen." Critical Reviews in Microbiology 16: 253-286.

Alvarez-Martinez, C. E. and P. J. Christie. 2009. "Biological diversity of prokaryotic type IV secretion systems." Microbiology and Molecular Biology Reviews 73: 775-808.

Andrade, A. and M. A. Valvano. 2014. "A *Burkholderia cenocepacia* gene encoding a non-functional tyrosine phosphatase is required for the delayed maturation of the bacteria-containing vacuoles in macrophages." Microbiology 160: 1332-1345.

Aschtgen, M. S., C. S. Bernard, S. De Bentzmann, R. Lloubes and E. Cascales. 2008. "SciN is an outer membrane lipoprotein required for type VI Secretion in *Enteroaggregative Escherichia coli*." Journal of Bacteriology 190: 7523-7531.

Aschtgen, M. S., M. Gavioli, A. Dessen, R. Lloubes and E. Cascales. 2010a. "The SciZ protein anchors the enteroaggregative *Escherichia coli* type VI secretion system to the cell wall." Molecular Microbiology 75: 886-899.

Aschtgen, M. S., M. S. Thomas and E. Cascales. 2010b. "Anchoring the type VI secretion system to the peptidoglycan TssL, TagL, TagP, ... what else?" Virulence 1: 535-540.

Aubert, D. F., R. S. Flannagan and M. A. Valvano. 2008. "A novel sensor kinase-response regulator hybrid controls biofilm formation and type VI secretion system activity in Burkholderia cenocepacia." Infection and Immunity 76: 1979-1991.

Aubert, D. F., H. Xu, J. Yang, X. Shi, W. Gao, L. Li, F. Bisaro, S. Chen, M. A. Valvano and F. Shao. 2016. "A *Burkholderia* type VI effector deamidates Rho GTPases to activate the pyrin inflammasome and trigger inflammation." Cell Host & Microbe 19: 664-674.

Aubert, D., D. K. MacDonald and M. A. Valvano. 2010. "BcsKC is an essential protein for the type VI secretion system activity in *Burkholderia cenocepacia* that formsan outer membrane complex with BcsLB." Journal of Biological Chemistry 285: 35988-35998.

Bönemann, G., A. Pietrosiuk, A. Diemand, H. Zentgraf and A. Mogk. 2009. "Remodelling of VipA/VipB tubules by ClpV-mediated threading is crucial for type VI protein secretion." EMBO Journal 28: 315-325.

Baldwin, A., E. Mahenthiralingam, J. Parkhill and P. A. Sokol. 2004. "The *Burkholderia cepacia* epidemic strain marker is part of a novel genomic island encoding both virulence and metabolism-associated genes in *Burkholderia cenocepacia*." Infection and Immunity 72: 1537-1547.

Ballister, E. R., A. H. Lai, R. N. Zuckermann, Y. Cheng and J. D. Mougous. 2008. "In vitro self-assembly from a simple protein of tailorable nanotubes building block." Proceedings of the National Academy of Sciences of the United States of America 105: 3733-3738.

Baron, C., M. Llosa, P. C. Zambryski and S. Zhou. 1997. "VirB1, a component of the T-complex transfer machinery of *Agrobacterium tumefaciens*, is processed to a C-terminal secreted product, virB1." Journal of Bacteriology 179: 1203-1210.

Basler, M., B.T. Ho and J. J. Mekalanos. 2013. "Tit-for-Tat: type VI Secretion system counterattack during bacterial cell-cell interactions." Cell 152: 884-894.

Basler, M. and J. J. Mekalanos. 2012. "Type 6 secretion dynamics within and between bacterial cells." Science 337: 815.

Basler, M., M. Pilhofer, G. P. Henderson, G. J. Jensen and J. J. Mekalanos. 2012. "Type VI secretion requires a dynamic contractile phage tail-like structure." Nature 483: 182-186.

Battesti, A. and E. Bouveret. 2012. "The bacterial two-hybrid system based on adenylate cyclase reconstitution in *Escherichia coli*." Methods 58: 325-334.

Belon, C, C. Soscia, A. Bernut, A. Laubier, S. Bleves and A-B, Blanc-Potard. 2015. "A macrophage subversion factor is shared by intracellular and extracellular pathogens." PLoS Pathogens 11: e1004969.

Berks, B. C., F. Sargent and T. Palmer. 2000. "The Tat protein export pathway." Molecular Microbiology 35: 260-274.

Bernard, C. S., Y. R. Brunet, M. Gavioli, R. Lloubes and E. Cascales. 2011. "Regulation of type VI secretion gene clusters by sigma54 and cognate enhancer binding proteins." Journal of Bacteriology 193: 2158-2167.

Bevivino, A., C. Dalmastri, S. Tabacchioni, L. Chiarini, M. L. Belli, S. Piana, A. Materazzo, P. Vandamme and G. Manno. 2002. "*Burkholderia cepacia* complex bacteria from clinical and environmental sources in Italy: genomovar status and distribution of traits related to virulence and transmissibility." Journal of Clinical Microbiology 40: 846.

Bian, Z. and S. Normark. 1997. "Nucleator function of CsgB for the assembly of adhesive surface organelles in *Escherichia coli*." EMBO Journal 16: 5827-5836.

Bingle, L. E. , C. M. Bailey and M. J. Pallen. 2008. "Type VI secretion: a beginner's guide." Current Opinion in Microbiology 11: 3-8.

Bladergroen, M. R., K. Badelt and H. P. Spaink. 2003. "Infection-blocking genes of a symbiotic Rhizobium leguminosarum strain that are involved in temperature-dependent protein secretion." Molecular Plant-Microbe Interactions 16: 53-64.

Bleves, S., V. Viarre, R. Salacha, G. P. Michel, A. Filloux and R. Voulhoux. 2010. "Protein secretion systems in *Pseudomonas aeruginosa*: a wealth of pathogenic weapons." International Journal of Medical Microbiology 300: 534-543.

Blondel, C. J., J. C. Jimenez, I. Contreras and C. A. Santiviago. 2009. "Comparative genomic analysis uncovers 3 novel loci encoding type six secretion systems differentially distributed in Salmonella serotypes." BMC Genomics 10: 354.

Bodelón, G., E. Marín and L. A. Fernández. 2009. "Role of periplasmic chaperones and BamA (YaeT/Omp85) in folding and secretion of intimin from enteropathogenic *Escherichia coli* strains." Journal of Bacteriology 191: 5169.

Bolhuis, A., J. E. Mathers, J. D. Thomas, C. M. Barrett and C. Robinson. 2001. "TatB and TatC form a functional and structural unit of the twin-arginine translocase from *Escherichia coli*." The Journal of Biological Chemistry 276: 20213.

Boyer, F., G. Fichant, J. Berthod, Y. Vandenbrouck and I. Attree. 2009. "Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources?" BMC Genomics 10: 104.

Braun, M., K. Stuber, Y. Schlatter, T. Wahli, P. Kuhnert and J. Frey. 2002. "Characterization of an ADP-ribosyltransferase toxin (AexT) from Aeromonas *salmonicida* subsp. *salmonicida*." Journal of Bacteriology 184: 1851-1858.

Brencic, A. and S. Lory. 2009. "Determination of the regulon and identification of novel mRNA targets of *Pseudomonas aeruginosa* RsmA." Molecular Microbiology 72: 612-632.

Broms, J. E., M. Lavander and A. Sjöstedt. 2009a. "A conserved α-helix essential for a type VI secretion-like system of *Francisella tularensis*." Journal of Bacteriology 191: 2431-2446.

Broms, J. E., M. Lavander and A. Sjostedt. 2009b. "A conserved alpha-helix essential for a type VI secretion-like system of *Francisella tularensis*." Journal of Bacteriology 191: 2431-2446.

Brooks, T. M., D. Unterweger, V. Bachmann, B. Kostiuk and S. Pukatzki. 2013. "Lytic activity of the *Vibrio cholerae* type VI secretion toxin VgrG-3 is inhibited by the antitoxin TsaB." The Journal of Biological Chemistry 288: 7618.

Browning, C., M. M Shneider, V. D. Bowman, D. Schwarzer and P. G. Leiman. 2012. "Phage pierces the host cell membrane with the iron-loaded spike." Structure 20: 326-339.

Bruch, M. D., C. J. McKnight and L. M. Gierasch. 1989. "Helix formation and stability in a signal sequence." Biochemistry 28: 8554-8561.

Brunet, Y. R., C. S. Bernard, M. Gavioli, R. Lloubès and E. Cascales. 2011. "An epigenetic switch involving overlapping Fur and DNA methylation optimizes expression of a type VI secretion gene cluster (epigenetic switch controlling EAEC sci1 expression)." PLoS Genetics 7: e1002205.

Brunet, Y. R., L. Espinosa, S. Harchouni, T. Mignot and E. Cascales. 2013. "Imaging type VI secretion-mediated bacterial killing." Cell Reports 3: 36-41.

Brunet, Y. R., J. Hénin, H. Celia and E. Cascales. 2014. "Type VI secretion and bacteriophage tail tubes share a common assembly pathway." EMBO Reports 15: 315-321.

Brunet, Y. R., A. Zoued, F. Boyer, B. Douzi and E. Cascales. 2015. "The type VI secretion TssEFGK-VgrG phage-like baseplate is recruited to the TssJLM membrane complex via multiple contacts and serves as assembly platform for tail tube/sheath polymerization (T6SS phage-like assembly platform)." PLoS Genetics 11: e1005545.

Burkholder, W. H. 1950. "Sour skin, a bacterial rot of onion bulbs." Phytopathology 40: 115-117.

Burns, D. L. 2003. "Type IV transporters of pathogenic bacteria." Current Opinion in Microbiology 6: 29-34.

Burtnick, M. N., P. J. Brett, S. V. Harding, S. A. Ngugi, W. J. Ribot, N. Chantratita, A. Scorpio, T. S. Milne, R. E. Dean, D. L. Fritz, S. J. Peacock, J. L. Prior, T. P. Atkins and D. Deshazer. 2011. "The cluster 1 type VI secretion system is a major virulence determinant in *Burkholderia pseudomallei*." Infection and Immunity 79: 1512.

Busby, J. N., S. Panjikar, M. J. Landsberg, M. R. Hurst and J. S. Lott. 2013. "The BC component of ABC toxins is an RHS-repeat-containing protein encapsulation device." Nature 501: 547-550.

Buttner, D. 2012. "Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal- pathogenic bacteria." Microbiology and Molecular Biology Reviews 76: 262-310.

Cardarelli, L., R. Lam, A. Tuite, L. A. Baker, P. D. Sadowski, D. R. Radford, J. L. Rubinstein, K. P. Battaile, N. Chirgadze, K. L. Maxwell and A. R. Davidson. 2010. "The crystal structure of bacteriophage HK97 gp6: defining a large family of head-tail connector proteins." Journal of Molecular Biology 395: 754-768.

Carruthers, M. D., P.A. Nicholson, E. N. Tracy and R. S. Munson. 2013. "Acinetobacter baumannii utilizes a type VI secretion system for bacterial competition." PLoS One 8: e59388.

Casabona, M. G., J. M. Silverman, K. M. Sall, F. Boyer, Y. Couté, J. Poirel, D. Grunwald, J. D. Mougous, S. Elsen and I. Attree. 2013. "An ABC transporter and an outer membrane lipoprotein participate in posttranslational activation of type VI secretion in *Pseudomonas aeruginosa*." Environmental Microbiology 15: 471-486.

Casadaban, M. J. and S. N. Cohen. 1980. "Analysis of gene-control signals by dna-fusion and cloning in *Escherichia coli*." Journal of Molecular Biology 138: 179-207.

Cascales, E. 2008. "The type VI secretion toolkit." EMBO Reports 9: 735-741.

Cascales, E. and C. Cambillau. 2012. "Structural biology of type VI secretion systems." Philosophical Transactions of the Royal Society B-Biological Sciences 367: 1102-1111.

Cascales, E. and P. J. Christie. 2004. "Agrobacterium VirB10, an ATP energy sensor required for type IV secretion." Proceedings of the National Academy of Sciences of the United States of America 101: 17228-17233.

Castang, S., H. R. McManus, K. H. Turner and S. L. Dove. 2008. "H-NS family members function coordinately in an opportunistic pathogen." Proceedings of the National Academy of Sciences of the United States of America 105: 18947.

Chang, J. H. and Y. G. Kim. 2015. "Crystal structure of the bacterial type VI secretion system component TssL from *Vibrio cholerae*." The Journal of Microbiology 53: 32-37.

Chapman, M. R., L. S. Robinson, J. S. Pinkner, R. Roth, J. Heuser, M. Hammar, S. Normark and S. J. Hultgren. 2002. "Role of *Escherichia coli* curli operons in directing amyloid fiber formation." Science 295: 851-855.

Cieri, M. V., N. Mayer-Hamblett, A. Griffith and J. L. Burns. 2002. "Correlation between an in vitro invasion assay and a murine model of *Burkholderia cepacia* lung infection." Infection and Immunity 70: 1081.

Cline, K. and M. McCaffery. 2007. "Evidence for a dynamic and transient pathway through the TAT protein transport machinery." EMBO Journal 26: 3039-3049.

Cline, K. and H. Mori. 2001. "Thylakoid DeltapH- dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport." Journal of Cell Biology 154: 719.

Coenye, T., J. J. LiPuma, D. Henry, B. Hoste, K. Vandemeulebroecke, M. Gillis, D. P. Speert and P. Vandamme. 2001a. "*Burkholderia cepacia* genomovar VI, a new member of the *Burkholderia cepacia* complex isolated from cystic fibrosis patients." International Journal of Systematic and Evolutionary Microbiology 51: 271-279.

Coenye, T., E. Mahenthiralingam, D. Henry, J. J. LiPuma, S. Laevens, M. Gillis, D. P. Speert and P. Vandamme. 2001b. "*Burkholderia ambifaria* sp. nov., a novel member of the *Burkholderia cepacia* complex including biocontrol and cystic fibrosis-related isolates." International Journal of Systematic and Evolutionary Microbiology 51: 1481-1490.

Collinson, S., L. Emoedy, K. H. Mueller, T. Trust and W. Kay. 1991. "Purification and characterization of thin, aggregative fimbriae from *Salmonella enteritidis*." Journal of Bacteriology 173: 4773-4781.

Collinson, S. K., S. C. Clouthier, J. L. Doran, P. A. Banser and W. W. Kay. 1996. "Salmonella enteritidis agfBAC operon encoding thin, aggregative fimbriae." Journal of Bacteriology 178: 662-667.

Conway, B. A., V. Venu and D. P. Speert. 2002. "Biofilm formation and acyl homoserine lactone production in the *Burkholderia cepacia* complex." Journal of Bacteriology 184: 5678.

Corbett, C. R., M. N. Burtnick, C. Kooi, D. E. Woods and P. A. Sokol. 2003. "An extracellular zinc metalloprotease gene of *Burkholderia cepacia*." Microbiology 149: 2263.

Cornelis, G. R. 2007. "The type-III secretion injectisome." International Journal of Medical Microbiology 297: 28-28.

Costa, T. R., C. Felisberto-Rodrigues, A. Meir, M. S. Prevost, A. Redzej, M. Trokter and G. Waksman. 2015. "Secretion systems in Gram-negative bacteria: structural and mechanistic insights." Nature Reviews Microbiology 13: 343-359.

Coulthurst, S. J. 2013. "The type VI secretion system - a widespread and versatile cell targeting system." Research in Microbiology 164: 640-654.

Cox, A. D. and S. G. Wilkinson. 1991. "Ionizing groups in lipopolysaccharides of Pseudomonas cepacia in relation to antibiotic resistance." Molecular Microbiology 5: 641-646.

Curtis, M. A., H. K. Kuramitsu, M. Lantz, F. L. Macrina, K. Nakayama, J. Potempa, E. C. Reynolds and J. Aduse-opoku. 1999. "Molecular genetics and nomenclature of proteases of *Porphyromonas gingivalis*." Journal of Periodontal Research 34: 464-472.

Das, S., A. Chakrabortty, R. Banerjee, S. Roychoudhury and K. Chaudhuri. 2000. "Comparison of global transcription responses allows identification of *Vibrio cholerae* genes differentially expressed following infection." FEMS Microbiology Letters 190: 87-91.

Das, S. and K. Chaudhuri. 2003. "Identification of a unique IAHP (IcmF Associated Homologous Proteins) cluster in *Vibrio cholerae* and other proteobacteria through In silico analysis." In Silico Biology: 287-300.

de Bruin, O. M., J. S. Ludu and F. E. Nano. 2007. "The *Francisella* pathogenicity island protein IglA localizes to the bacterial cytoplasm and is needed for intracellular growth." BMC Microbiology 7: 1.

de Pace, F., G. Nakazato, A. Pacheco, J. B. de Paiva, V. Sperandio and W. D. Da Silveira. 2010. "The type VI secretion system plays a role in type 1 fimbria expression and pathogenesis of an avian pathogenic *Escherichia coli* strain." Infection and Immunity 78: 4990.

De Smet, B., M. Mayo, C. Peeters, J. E. Zlosnik, T. Spilker, T. J. Hird, J. J. LiPuma, T. J. Kidd, M. Kaestli, J. L. Ginther, D. M. Wagner, P. Keim, S. C. Bell, J. A. Jacobs, B. J. Currie and P. Vandamme. 2015. "*Burkholderia stagnalis* sp. nov. and *Burkholderia territorii* sp. nov., two novel *Burkholderia cepacia* complex species from environmental and human sources." International Journal of Systematic and Evolutionary Microbiology 65: 2265-2271.

Desai, M., T. Bhler, P. H. Weller and M. R. W. Brown. 1998. "Increasing resistance of planktonic and biofilm cultures of *Burkholderia cepacia* to ciprofloxacin and ceftazidime during exponential growth." Journal of Antimicrobial Chemotherapy 42: 153-160.

Desvaux, M., M. Hébraud, I. R. Henderson and M. J. Pallen. 2006. "Type III secretion: what's in a name?" Trends in Microbiology 14: 157-160.

Desvaux, M., M. Hébraud, R. Talon and I. R. Henderson. 2009. "Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue." Trends in Microbiology 17: 139-145.

di Guan, C., P. Li, P. D. Riggs and H. Inouye. 1988. "Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein." Gene 67: 21-30.

Dong, T. G., B. T. Ho, D. R. Yoder-Himes and J. J. Mekalanos. 2013. "Identification of T6SS-dependent effector and immunity proteins by Tn-seq in *Vibrio cholerae*." Proceedings of the National Academy of Sciences of the United States of America 110: 2623.

Dooley, J. S. and T. J. Trust. 1988. "Surface protein composition of *Aeromonas hydrophila* strains virulent for fish: identification of a surface array protein." Journal of Bacteriology 170: 499.

Douzi, B., S. Spinelli, S. Blangy, A. Roussel, E. Durand, Y. R. Brunet, E. Cascales and C. Cambillau. 2014. "Crystal structure and self-interaction of the type VI secretion tail-tube protein from enteroaggregative *Escherichia coli*." PLoS One 9: e86918.

Driessen, A. J., P. Fekkes and J. P. van Der Wolk. 1998. "The Sec system." Current Opinion in Microbiology 1: 216-222.

du Plessis, D. J. F., N. Nouwen and A. J. M. Driessen. 2010. "The Sec translocase." BBA - Biomembranes 1808: 851-865.

Dudley, E. G., N. R. Thomson, J. Parkhill, N. P. Morin and J. P. Nataro. 2006. "Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative *Escherichia coli*." Molecular Microbiology 61: 1267-82.

Dumetz, A. C., A. M. Chockla, E. W. Kaler and A. M. Lenhoff. 2008. "Effects of pH on protein-protein interactions and implications for protein phase behavior." Biochimica et Biophysica Acta 1784: 600-610.

Duong, F. and W. Wickner. 1997. "Distinct catalytic roles of the SecYE, SecG and SecDFyajC subunits of preprotein translocase holoenzyme." EMBO Journal 16: 2756-2768.

Durand, E., E. Derrez, Gi. Audoly, S. Spinelli, M. Ortiz-Lombardia, D. Raoult, E. Cascales and C. Cambillau. 2012a. "Crystal structure of the VgrG1 actin cross-linking domain of the *Vibrio cholerae* type VI secretion system." The Journal of Biological Chemistry 287: 38190.

Durand, E., V. S. Nguyen, A. Zoued, L. Logger, G. Péhau-Arnaudet, M. S. Aschtgen, S. Spinelli, A. Desmyter, B. Bardiaux, A. Dujeancourt, A. Roussel, C. Cambillau, E. Cascales and R. Fronzes. 2015. "Biogenesis and structure of a type VI secretion membrane core complex." Nature 523: 555.

Durand, E., A. Zoued, S. Spinelli, P. J. H. Watson, M. S. Aschtgen, L. Journet, C. Cambillau and E. Cascales. 2012b. "Structural characterization and oligomerization of the TssL protein, a component shared by bacterial type VI and type IVb secretion systems." Journal of Biological Chemistry 287: 14157-14168.

Economou, A. 1998. "Bacterial preprotein translocase: mechanism and conformational dynamics of a processive enzyme." Molecular Microbiology 27: 511-518.

Engelman, D. M. and T. A. Steitz. 1981. "The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis." Cell 23: 411-422.

English, G., O. Byron, F. R. Cianfanelli, A. R. Prescott and S. J. Coulthurst. 2014. "Biochemical analysis of TssK, a core component of the bacterial type VI secretion system, reveals distinct oligomeric states of TssK and identifies a TssK-TssFG subcomplex." Biochemical Journal 461: 291.

English, G., K. Trunk, V. A. Rao, V. Srikannathasan, W. N. Hunter and S. J. Coulthurst. 2012. "New secreted toxins and immunity proteins encoded within the type VI secretion system gene cluster of *Serratia marcescens*." Molecular Microbiology 86: 921-936.

Enos-Berlage, J. L., Z. T. Guvener, C. E. Keenan and L. L. McCarter. 2005. "Genetic determinants of biofilm development of opaque and translucent *Vibrio parahaemolyticus*." Molecular Microbiology 55: 1160-1182.

Fabian, G., Z. Franziska, P. J. Roman, M. B. Björn, H. Sebastian and M. Timm. 2013. "The structural basis of autotransporter translocation by TamA." Nature Structural & Molecular Biology 20: 1318.

Felisberto-Rodrigues, C., E. Durand, M. S. Aschtgen, S. Blangy, M. Ortiz-Lombardia, B. Douzi, C. Cambillau and E. Cascales. 2011. "Towards a structural comprehension of bacterial type VI secretion systems: characterization of the TssJ-TssM complex of an *Escherichia coli* pathovar." PLoS Pathogens 7: e1002386.

Ferguson, M. R., X. J. Xu, C. W. Houston, J. W. Peterson, D. H. Coppenhaver, V. L. Popov and A. K. Chopra. 1997. "Hyperproduction, purification, and mechanism of action of the cytotoxic enterotoxin produced by *Aeromonas hydrophila*." Infection and Immunity 65: 4299.

Finlay, B. B. and S. Falkow. 1997. "Common themes in microbial pathogenicity revisited." Microbiology and Molecular Biology Reviews 61: 136-169.

Fleiszig, S. M. J., T. S. Zaidi and G. B. Pier. 1995. "*Pseudomonas aeruginosa* invasion of and multiplication within corneal epithelial cells *in vitro*." Infection and Immunity 63: 4072–4077.

Fokine, A., Z. Zhang, S. Kanamaru, V. D. Bowman, A. A. Aksyuk, F. Arisaka, V. B. Rao and M. G. Rossmann. 2013. "The molecular architecture of the bacteriophage T4 neck." Journal of Molecular Biology 425: 1731-1744.

Folkesson, A., S. Lofdahl and S. Normark. 2002. "The *Salmonella enterica* subspecies I specific centisome 7 genomic island encodes novel protein families present in bacteria living in close contact with eukaryotic cells." Research in Microbiology 153: 537-545.

Fox, J. D. and D. S. Waugh. 2003. "Maltose-binding protein as a solubility enhancer." Methods in Molecular Biology 205: 99-117.

Fritsch, M., K. Trunk, J. Diniz, M. Guo, M. Trost and S. J. Coulthurst. 2013. "Proteomic identification of novel secreted antibacterial toxins of the *Serratia marcescens* type VI secretion system." Molecular and Cellular Proteomics 12:2735-2749.

Gérard, F. and K. Cline. 2006. "Efficient twin arginine translocation (Tat) pathway transport of a precursor protein covalently anchored to its initial cpTatC binding site." Journal of Biological Chemistry 281: 6130-6135.

Gérard, F. and K. Cline. 2007. "The thylakoid proton gradient promotes an advanced stage of signal peptide binding deep within the Tat pathway receptor complex." Journal of Biological Chemistry 282: 5263-5272.

Galindo, C. L., C. Gutierrez Jr and A. K. Chopra. 2006. "Potential involvement of galectin- 3 and SNAP23 in *Aeromonas hydrophila* cytotoxic enterotoxin- induced host cell apoptosis." Microbial Pathogenesis 40: 56-68.

Gerc, A. J., A. Diepold, K. Trunk, M. Porter, C. Rickman, J. P. Armitage, N. R. Stanley-Wall and S. J. Coulthurst. 2015. "Visualization of the serratia type VI secretion system reveals unprovoked attacks and dynamic assembly." Cell Reports 12: 2131-2142.

Gohlke, U., L. Pullan, C. A. McDevitt, I. Porcelli, E. De Leeuw, T. Palmer, H. R. Saibil, B. C. Berks and W. T. Wickner. 2005. "The TatA component of the twin- arginine protein transport system forms channel complexes of variable diameter." Proceedings of the National Academy of Sciences of the United States of America 102: 10482-10486.

Golovanov, A. P., G. M. Hautbergue, S. A. Wilson and L. Y. Lian. 2004. "A simple method for improving protein solubility and long-term stability." Journal of the American Chemical Society 126: 8933-8939.

Govan, J. R. W., C. J. Doherty, J. W. Nelson, P. H. Brown, A. P. Greening, J. Maddison, M. Dodd and A. K. Webb. 1993. "Evidence for transmission of *Pseudomonas cepacia* by social contact in cystic fibrosis." The Lancet 342: 15-19.

Goyal, P., P. V. Krasteva, N. Van Gerven, F. Gubellini, I. Van den Broeck, A. Troupiotis-Tsailaki, W. Jonckheere, G. Pehau-Arnaudet, J. S. Pinkner, M. R. Chapman, S. J. Hultgren, S. Howorka, R. Fronzes and H. Remaut. 2014. "Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG." Nature 516: 250-253.

Gray, C. G., S. C. Cowley, K. K. M. Cheung and F. E. Nano. 2002. "The identification of five genetic loci of *Francisella novicida* associated with intracellular growth." FEMS Microbiology Letters 215: 53-56.

Gray, M. D., M. Bagdasarian, W. G. Hol and M. Sandkvist. 2011. "In vivo cross- linking of EpsG to EpsL suggests a role for EpsL as an ATPase-pseudopilin coupling protein in the type II secretion system of *Vibrio cholerae*." Molecular Microbiology 79: 786.

Gruber, C. W., M. Čemažar, B. Heras, J. L. Martin and D. J. Craik. 2006. "Protein disulfide isomerase: the structure of oxidative folding." Trends in Biochemical Sciences 31: 455-464.

Gueguen, E. and E. Cascales. 2013. "Promoter swapping unveils the role of the *Citrobacter rodentium* CTS1 type VI secretion system in interbacterial competition." Applied and Environmental Microbiology 79: 32.

Hachani, A., L. P. Allsopp, Y. Oduko and A. Filloux. 2014. "The VgrG proteins are "à la carte" delivery systems for bacterial type VI effectors." The Journal of Biological Chemistry 289: 17872.

Hachani, A., N. S. Lossi, A. Hamilton, C. Jones, S. Bleves, D. Albesa-Jove and A. Filloux. 2011. "Type VI Secretion System in *Pseudomonas aeruginosa* secretion and multimerization of VgrG proteins." The Journal of Biological Chemistry 286: 12317-12327.

Hacker, J. and J. B. Kaper. 2000. "Pathogenicity islands and the evolution of microbes." Annual Review of Microbiology 54: 641-679.

Hammar, M., A. Arnqvist, Z. Bian, A. Olsén and S. Normark. 1995. "Expression of two csg operons is required for production of fibronectin- and Congo red- binding curli polymers in *Escherichia coli* K-12." Molecular Microbiology 18: 661-670.

Hammar, M., Z. Bian and S. Normark. 1996. "Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*." Proceedings of the National Academy of Sciences of the United States of America 93: 6562-6566.

Hammer, N. D., B. A. McGuffie, Y. Zhou, M. P. Badtke, A. A. Reinke, K. Brännström, J. E. Gestwicki, A. Olofsson, F. Almqvist and M. R. Chapman. 2012. "The C- terminal repeating units of CsgB direct bacterial functional amyloid nucleation." Journal of Molecular Biology 422: 376-389.

Henderson, I. R., F. Navarro-Garcia, M. Desvaux, R. C. Fernandez, Ala and D. Aldeen. 2004. "Type V protein secretion pathway: the autotransporter story." Microbiology and Molecular Biology Reviews 68: 692-744.

Hiransuthikul, N., W. Tantisiriwat, K. Lertutsahakul, A. Vibhagool and P. Boonma. 2005. "Skin and softtissue infections among tsunami survivors in southern Thailand." Clinical Infectious Diseases 41: e93-e96.

Ho, B. T., M. Basler and J. J. Mekalanos. 2013. "Type 6 secretion system- mediated immunity to type 4 secretion system-mediated gene transfer." Science 342: 250-253.

Ho, B. T., T. G. Dong and J. J. Mekalanos. 2014. "A view to a Kill: the bacterial type VI secretion system." Cell Host & Microbe 15: 9-21.

Hobbs, M. and J. S. Mattick. 1993. "Common components in the assembly of type 4 fimbriae, DNA transfer systems, filamentous phage and protein-secretion apparatus: a general system for the formation of surface-associated protein complexes." Molecular Microbiology 10: 233-243.

Holden, M. T. , H. M. Seth-Smith, L. C. Crossman, M. Sebaihia, S. D. Bentley, A. M. Cerdeno-Tarraga, N. R. Thomson, N. Bason, M. A. Quail, S. Sharp, I. Cherevach, C. Churcher, I. Goodhead, H. Hauser, N. Holroyd, K. Mungall, P. Scott, D. Walker, B. White, H. Rose, P. Iversen, D. Mil-Homens, E. P. C. Rocha, Am Fialho, A. Baldwin, C. Dowson, B. G. Barrell, J. R. Govan, P. Vandamme, Ca Hart, E. Mahenthiralingam and J. Parkhill. 2009. "The genome of *Burkholderia cenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients." Journal of Bacteriology 191: 261-277.

Holden, M. T., R. W. Titball, S. J. Peacock, A. M. Cerdeno-Tarraga, T. Atkins, L. C. Crossman, T. Pitt, C. Churcher, K. Mungall, S. D. Bentley, M. Sebaihia, N. R. Thomson, N. Bason, I. R. Beacham, K. Brooks, K. A. Brown, N. F. Brown, G. I. Challis, I. Cherevach, T. Chillingworth, A. Cronin, B. Crossett, P. Davis, D. Deshazer, T. Feltwell, A. Fraser, Z. Hance, H. Hauser, S. Holroyd, K. Jagels, Ke Keith, M. Maddison, S. Moule, C. Price, M. A. Quail, E. Rabbinowitsch, K. Rutherford, M. Sanders, M. Simmonds, S. Songsivilai, K. Stevens, S. Tumapa, M. Vesaratchavest, S. Whitehead, C. Yeats, B. G. Barrell, P. C. Oyston and J. Parkhill. 2004. "Genornic plasticity of the causative agent of melioidosis, *Burkholderia pseudomallei*." Proceedings of the National Academy of Sciences of the United States of America 101: 14240-14245.

Holland, I. B., L. Schmitt and J. Young. 2005. "Type 1 protein secretion in bacteria, the ABC-transporter dependent pathway (review)." Molecular Membrane Biology 22: 29-39.

Hood, R. D., P. Singh, F. Hsu, T. Guevener, M. A. Carl, R. R. Trinidad, J. M. Silverman, B. B. Ohlson, K. G. Hicks, R. L. Plemel, M. Li, S. Schwarz, W. Y. Wang, A. J. Merz, D. R. Goodlett and J. D. Mougous. 2010. "A type VI secretion system of *Pseudomonas aeruginosa* targets, a toxin to bacteria." Cell Host & Microbe 7: 25-37.

Hsu, F., S. Schwarz and J. D. Mougous. 2009. "TagR promotes PpkA-catalysed type VI secretion activation in *Pseudomonas aeruginosa*." Molecular Microbiology 72: 1111-1125.

Hultgren, S. J., S. Normark and S. N. Abraham. 1991. "Chaperone-assisted assembly and molecular architecture of adhesive pili." Annual Review of Microbiology 45: 383-415.

Hutchison, M. L., I. R. Poxton and J. R. Govan. 1998. "*Burkholderia cepacia* produces a hemolysin that is capable of inducing apoptosis and degranulation of mammalian phagocytes." Infection and Immunity 66: 2033-2039.

Ieva, R. and H. D. Bernstein. 2009. "Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane." Proceedings of the National Academy of Sciences of the United States of America 106: 19120.

Ilangovan, A., S. Connery and G. Waksman. 2015. "Structural biology of the Gram-negative bacterial conjugation systems." Trends in Microbiology 23: 301-310.

Ishikawa, T., P. K. Rompikuntal, B. Lindmark, D. L. Milton and S. N. Wai. 2009. "Quorum sensing regulation of the two hcp alleles in *Vibrio cholerae* O1 Strains (Hcp in Serotype O1 *V. cholerae*)." PLoS One 4: e6734.

Ishikawa, T., D. Sabharwal, J. Bröms, D. L. Milton, A. Sjöstedt, B. E. Uhlin and S. N. Wai. 2012. "Pathoadaptive conditional regulation of the type VI secretion system in *Vibrio cholerae* O1 strains." Infection and Immunity 80: 575-584.

Isles, A., I. Maclusky, M. Corey, R. Gold, C. Prober, P. Fleming and H. Levison. 1984. "*Pseudomonas cepacia* infection in cystic fibrosis: an emerging problem." The Journal of Pediatrics 104: 206-210.

Jervis, A. J. and J. Green. 2007. "In vivo demonstration of FNR dimers in response to lower O(2) availability." Journal of Bacteriology 189: 2930-2932.

Jiang, F., N. R. Waterfield, J. Yang, G. Yang and Q. Jin. 2014. "A *Pseudomonas aeruginosa* type VI secretion phospholipase D effector targets both prokaryotic and eukaryotic Cells." Cell Host & Microbe 15: 600-610.

Jiang, F., X. Wang, B. Wang, L. Chen, Z. Zhao, N. R. Waterfield, G. Yang and Q. Jin. 2016. "The *Pseudomonas aeruginos*a type VI secretion PGAP1-like effector induces host autophagy by activating endoplasmic reticulum stress." Cell Reports 16: 1502-1509.

Jobichen, C., S. Chakraborty, M. Li, J. Zheng, L. Joseph, Y. K. Mok, K. Y. Leung and J. Sivaraman. 2010. "Structural basis for the secretion of EvpC: a key type VI secretion system protein from *Edwardsiella tarda*." PLoS One 5: e12910.

Joly, J. C. and W. Wickner. 1993. "The SecA and SecY subunits of translocase are the nearest neighbors of a translocating preprotein, shielding it from phospholipids." The EMBO journal 12: 255.

Jones, R. 2012. "Characterisation of putative type VI secretion system effector proteins from burkholderia cenocepacia." In Department of Infection and Immunity University of Sheffield.

Joseph, S. W. and A. Carnahan. 1994. "The isolation, identification, and systematics of the motile *Aeromonas* species." Annual Review of Fish Diseases 4: 315-343.

Junker, M., R. N. Besingi and P. L. Clark. 2009. "Vectorial transport and folding of an autotransporter virulence protein during outer membrane secretion." Molecular Microbiology 71: 1323-1332.

Kanamaru, S., P. G. Leiman, V. A. Kostyuchenko, P. R. Chipman, V. V. Mesyanzhinov, F. Arisaka and M. G. Rossmann. 2002. "Structure of the cell-puncturing device of bacteriophage T4." Nature 415: 553-557.

Kanonenberg, K., C. K. W. Schwarz and L. Schmitt. 2013. "Type I secretion systems - a story of appendices." Research in Microbiology 164: 596-604.

Kapitein, N., G. Boenemann, A. Pietrosiuk, F. Seyffer, I. Hausser, J. K. Locker and A. Mogk. 2013. "ClpV recycles VipA/VipB tubules and prevents non-productive tubule formation to ensure efficient type VI protein secretion." Molecular Microbiology 87(5): 1013-1028.

Kapust, R. B. and D. S. Waugh. 1999. "*Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused." Protein Science 8: 1668-1674.

Karimova, G., N. Dautin and D. Ladant. 2005. "Interaction network among *Escherichia coli* membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis." Journal of Bacteriology 187: 2233-2243.

Karimova, G., J. Pidoux, A. Ullmann and D. Ladant. 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: 5752-5756.

Karna, S. L., X. Zogaj, J. Barker, J. Seshu, S. Dove and K. E. Klose. 2010. "A bacterial two-hybrid system that utilizes Gateway cloning for rapid screening of protein-protein interactions." BioTechniques 49: 831-833.

Kellermann, O. K. and T. Ferenci. 1982. "Maltose-binding protein from *Escherichia coli*." Methods Enzymol 90: 459-463.

Khodai-Kalaki, M., D. F. Aubert and M. A. Valvano. 2013. "Characterization of the AtsR hybrid sensor kinase phosphorelay pathway and identification of its response regulator in Burkholderia cenocepacia." The Journal of Biological Chemistry 288: 30473.

Kikuchi, T., Y. Mizunoe, A. Takade, S. Naito and S. Yoshida. 2005. "Curli fibers are required for development of biofilm architecture in Escherichia coli K-12 and enhance bacterial adherence to human uroepithelial cells." Microbiology and Immunology 49: 875-884.

King, J. 1968. "Assembly of the tail of bacteriophage T4." Journal of Molecular Biology 32: 231.

Kitaoka, M., S. T. Miyata, T. M. Brooks, D. Unterweger and S. Pukatzki. 2011. "VasH is a transcriptional regulator of the type VI secretion system functional in endemic and pandemic *Vibrio cholerae*." Journal of Bacteriology 193: 6471.

Knowles, T. J., A. Scott-Tucker, M. Overduin and I. R. Henderson. 2009. "Membrane protein architects: the role of the BAM complex in outer membrane protein assembly." Nature Reviews Microbiology 7: 206-214.

Kooi, C., C. R. Corbett, P. A. Sokol and P. A. Corbett. 2005. "Functional analysis of the *Burkholderia cenocepacia* ZmpA metalloprotease." Journal of Bacteriology 187: 4421-4429.

Kooi, C., B. Subsin, R. Chen, B. Pohorelic and P. A. Sokol. 2006. "*Burkholderia cenocepacia* ZmpB is a broad- specificity zinc metalloprotease involved in virulence." Infection and Immunity 74: 4083-4093.

Korotkov, K., T. Johnson, M. Jobling, J. Pruneda, E. Pardon, A. Heroux, S. Turley, J. Steyaert, R. Holmes and et al. 2011. "Structural and Functional Studies on the Interaction of GspC and GspD in the type ll secretion system." PLoS Pathogens 7: e1002228.

Korotkov, K. V., M. Sandkvist and W. G. Hol. 2012. "The type II secretion system: biogenesis, molecular architecture and mechanism." Nature Reviews Microbiology 10: 336-351.

Koskiniemi, S., J. G. Lamoureux, K. C. Nikolakakis, C. Kint de Roodenbeke, M. D. Kaplan, D. A. Low and C. S. Hayes. 2013. "Rhs proteins from diverse bacteria mediate intercellular competition." Proceedings of the National Academy of Sciences of the United States of America 110: 7032.

Kostyuchenko, V. A., P. R. Chipman, P. G. Leiman, F. Arisaka, V. V. Mesyanzhinov and M. G. Rossmann. 2005. "The tail structure of bacteriophage T4 and its mechanism of contraction." Nature Structural & Molecular Biology 12: 810-813.

Kostyuchenko, V. A., P. G. Leiman, P. R. Chipman, S. Kanamaru, M. J. van Raaij, F. Arisaka, V. V. Mesyanzhinov and M. G. Rossmann. 2003. "Three-dimensional structure of bacteriophage T4 baseplate." Nature Structural & Molecular Biology 10: 688-693.

Kovach, M. E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop, 2nd and K. M. Peterson. 1995. "Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes." Gene 166: 175-176.

Kovach, M. E., R. W. Phillips, P. H. Elzer, R. M. Roop, 2nd and K. M. Peterson. 1994. "pBBR1MCS: a broad-host-range cloning vector." BioTechniques 16: 800-802.

Krumme, M. L., K. N. Timmis and D. F. Dwyer. 1993. "Degradation of trichloroethylene by *Pseudomonas cepacia* G4 and the constitutive mutant strain G4 5223 PR1 in aquifer microcosms." Applied and Environmental Microbiology 59: 2746-2749.

Kube, S., N. Kapitein, T. Zimniak, F. Herzog, A. Mogk and P. Wendler. 2014. "Structure of the VipA/B type VI secretion complex suggests a contraction-state-specific recycling mechanism." Cell Reports 8: 20-30.

Ladant, D. 1988. "Interaction of *Bordetella pertussis* adenylate cyclase with calmodulin. Identification of two separated calmodulin-binding domains." The Journal of Biological Chemistry 263: 2612-2618.

Ladant, D. and A. Ullmann. 1999. "*Bordetella pertussis* adenylate cyclase: a toxin with multiple talents." Trends in Microbiology 7: 172-176.

Laemmli, U. K. 1970. "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature 227: 680-685.

Leiman, P. G., F. Arisaka, M. J. van Raaij, V.A. Kostyuchenko, A. A. Aksyuk, S. Kanamaru and M. G. Rossmann. 2010. "Morphogenesis of the T4 tail and tail fibers." Virology Journal 7: 355.

Leiman, P. G., M. Basler, U. A. Ramagopal, J. B. Bonanno, J. M. Sauder, S. Pukatzki, S. K. Burley, S. C. Almo and J. J. Mekalanos. 2009. "Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin." Proceedings of the National Academy of Sciences of the United States of America 106: 4154-4159.

Leiman, P. G., P. R. Chipman, V. A. Kostyuchenko, V. V. Mesyanzhinov and M. G. Rossmann. 2004. "Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host." Cell 118: 419-429.

Leiman, P. G. and M. M. Shneider. 2012. "Contractile tail machines of bacteriophages." Advances in Experimental Medicine and Biology 726: 93-114.

Lennox, E. S. 1955. "Transduction of linked genetic characters of the host by bacteriophage P1." Virology 1: 190-206.

Leo, J. C., I. Grin and D. Linke. 2012. "Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane." Philosophical Transactions of the Royal Society B 367: 1088-1101.

Leroux, M., J. A. De Leon, N. J. Kuwada, A. B. Russell, D. Pinto-Santini, R. D. Hood, D. M. Agnello, S. M. Robertson, P. A. Wiggins and J. D. Mougous. 2012. "Quantitative single-cell characterization of bacterial interactions reveals type VI secretion is a double-edged sword." Proceedings of the National Academy of Sciences of the United States of America 109: 19804.

Li, M., I. Le Trong, M. A. Carl, E. T. Larson, S. Chou, J. A. De Leon, S. L. Dove, R. E. Stenkamp and J. Mougous. 2012. "Structural basis for type VI secretion effector recognition by a cognate immunity protein." PLoS Pathogens 8: e1002613.

Liang, X., R. Moore, M. Wilton, M. J. Q. Wong, L. Lam and T. G. Dong. 2015. "Identification of divergent type VI secretion effectors using a conserved chaperone domain." Proceedings of the National Academy of Sciences of the United States of America 112: 9106.

Lillington, J., S. Geibel and G. Waksman. 2014. "Biogenesis and adhesion of type 1 and P pili." BBA -general subjects 1840: 2783-2793.

Lin, J. S., L. S. Ma and E. M. Lai. 2013. "Systematic dissection of the *Agrobacterium* type VI secretion system reveals machinery and secreted components for subcomplex formation." PLoS One 8: e67647.

Lin, J. S., H. H. Wu, P. H. Hsu, L. S. Ma, Y. Y. Pang, M. D. Tsai, E. M. Lai and J. D. Mougous. 2014. "Fha interaction with phosphothreonine of TssL activates type VI secretion in *Agrobacterium tumefaciens*." PLoS Pathogens 10: e1003991.

Lipuma, J. J., S. E. Dasen, T. L. Stull, D. W. Nielson and R. C. Stern. 1990. "Person-to-person transmission of *Pseudomonas cepacia* between patients with cystic fibrosis." The Lancet 336: 1094-1096.

Lipuma, J. J., T. Spilker, L. H. Gill, P. W. Campbell, L. Liu and E. Mahenthiralingam. 2001. "Disproportionate distribution of *Burkholderia cepacia* complex species and transmissibility markers in cystic fibrosis." American journal of respiratory and critical care medicine 164: 92.

Loferer, H., M. Hammar and S. Normark. 1997. "Availability of the fibre subunit CsgA and the nucleator protein CsgB during assembly of fibronectin-binding curli is limited by the intracellular concentration of the novel lipoprotein CsgG." Molecular Microbiology 26: 11-23.

Lossi, N. S., R. Dajani, P. Freemont and A. Filloux. 2011. "Structure-function analysis of HsiF, a gp25-like component of the type VI secretion system, in *Pseudomonas aeruginosa*." Microbiology 157: 3292-305.

Lossi, N. S., E. Manoli, A. Förster, R. Dajani, T. Pape, P. Freemont and A. Filloux. 2013. "The HsiB1C1 (TssB-TssC) complex of the *Pseudomonas aeruginosa* type VI secretion system forms a bacteriophage tail sheathlike structure." The Journal of Biological Chemistry 288: 7536.

Low, H. H., F. Gubellini, A. Rivera-Calzada, N. Braun, S. Connery, A. Dujeancourt, F. Lu, A. Redzej, R. Fronzes, E. V. Orlova and G. Waksman. 2014. "Structure of a type IV secretion system." Nature 508: 550.

Lu, C., S. Turley, S.T. Marionni, Y. J. Park, K. K. Lee, M. Patrick, R. Shah, M. Sandkvist, M. F. Bush and W. G. Hol. 2013. "Hexamers of the Type II Secretion ATPase GspE from *Vibrio cholerae* with Increased ATPase Activity." Structure 21: 1707-1717.

Luirink, J. and I. Sinning. 2004. "SRP-mediated protein targeting: structure and function revisited." BBA - Molecular Cell Research 1694: 17-35.

Lycklama, J. A. and A. J. Driessen. 2012. "The bacterial Sec-translocase: structure and mechanism." Philosophical transactions of the Royal Society of London 367: 1016-1028.

Ma, A. T., S. McAuley, S. Pukatzki and J. J. Mekalanos. 2009a. "Translocation of a *Vibrio cholerae* type VI secretion effector requires bacterial endocytosis by host cells." Cell Host & Microbe 5: 234-243.

Ma, J., Y. Bao, M. Sun, W. Dong, Z. Pan, W. Zhang, C. Lu and H. Yao. 2014. "Two functional type VI secretion systems in avian pathogenic *Escherichia coli* are involved in different pathogenic pathways." Infection and Immunity 82: 3867.

Ma, L. S., J. . Lin and E. M. Lai. 2009b. "An IcmF family protein, ImpL(M), is an integral inner membrane protein interacting with ImpK(L), and its Walker A motif is required for type VI secretion system-mediated Hcp secretion in *Agrobacterium tumefaciens*." Journal of Bacteriology 191: 4316-4329.

Ma, L. S., F. Narberhaus and E. M. Lai. 2012. "IcmF family protein TssM exhibits ATPase activity and energizes type VI secretion." The Journal of Biological Chemistry 287: 15610.

MacIntyre, D. L., S. T. Miyata, M. Kitaoka and S. Pukatzki. 2010. "The *Vibrio cholerae* type VI secretion system displays antimicrobial properties." Proceedings of the National Academy of Sciences of the United States of America 107: 19520-19524.

Mahenthiralingam, E., A. Baldwin and P. Vandamme. 2002. "*Burkholderia cepacia* complex infection in patients with cystic fibrosis." Journal of Medical Microbiology 51: 533-538.

Mahenthiralingam, E., P. Vandamme, M. E. Campbell, D. A. Henry, A. M. Gravelle, L. T. Wong, A. G. Davidson, P. G. Wilcox, B. Nakielna and D. P. Speert. 2001. "Infection with *Burkholderia cepacia* complex genomovars in patients with cystic fibrosis: virulent transmissible strains of genomovar III can replace *Burkholderia multivorans*." Clinical Infectious Diseases 33: 1469-1475.

Mahillon, J. and M. Chandler. 1998. "Insertion Sequences." Microbiology and Molecular Biology Reviews 62: 725-774.

Maina, C. V., P. D. Riggs, A. G. Grandea, B. E. Slatko, L. S. Moran, J. A. Tagliamonte, L. A. McReynolds and C. D. Guan. 1988. "An *Escherichia coli* vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein." Gene 74: 365-373.

Martin, D. W. and C. D. Mohr. 2000. "Invasion and intracellular survival of *Burkholderia cepacia*." Infection and Immunity 68: 24.

Martina, P., M. Bettiol, C. Vescina, P. Montanaro, M. C. Mannino, C. I. Prieto, C. Vay, D. Naumann, J. Schmitt, O. Yantorno, A. Lagares and A. Bosch. 2013. "Genetic diversity of *Burkholderia contaminans* isolates from cystic fibrosis patients in Argentina." Journal of Clinical Microbiology 51: 339.

Martinez, J. J., M. A. Mulvey, J. D. Schilling, J. S. Pinkner and S. J. Hultgren. 2000. "Type 1 pilus-mediated bacterial invasion of bladder epithelial cells." EMBO Journal 19: 2803-2812.

Maurer, C., S. Panahandeh, A. C. Jungkamp, M. Moser and M. Müller. 2010. "TatB functions as an oligomeric binding site for folded tat precursor proteins." Molecular Biology of the Cell 21: 4151-4161.

McBride, M. J., G. Xie, E. C. Martens, A. Lapidus, B. Henrissat, R. G. Rhodes, E. Goltsman, W. Wang, J. Xu, D. W. Hunnicutt, A. M. Staroscik, T. R. Hoover, Y. Q. Cheng and J. L. Stein. 2009. "Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis." Applied and Environmental Microbiology 75:6864.

McLaughlin, L. S., R. J. Haft and K. T. Forest. 2012. "Structural insights into the type II secretion nanomachine." Current Opinion in Structural Biology 22:208-216.

Medina-Pascual, M. J., S. Valdezate, G. Carrasco, P. Villalon, N. Garrido and J. A. Saez-Nieto. 2015. "Increase in isolation of *Burkholderia contaminans* from Spanish patients with cystic fibrosis." Clinical Microbiology and Infection 21:150-156.

Merino, S., X. Rubires, A. Aguillar, J. F. Guillot and J. M. Tomás. 1996. "The role of the O-antigen lipopolysaccharide on the colonization in vivo of the germfree chicken gut by *Aeromonas hydrophila* serogroup O:34." Microbial Pathogenesis 20:325-333.

Mil-Homens, D. and A. M. Fialho. 2011. "Trimeric autotransporter adhesins in members of the *Burkholderia cepacia* complex: a multifunctional family of proteins implicated in virulence." Frontiers in Cellular and Infection Microbiology 1:13.

Mil-Homens, D., A. M. Fialho and E. P. Rocha. 2010. "Genome-wide analysis of DNA repeats in *Burkholderia cenocepacia* J2315 identifies a novel adhesin-like gene unique to epidemic-associated strains of the ET-12 lineage." Microbiology 156:1084-1096.

Miroux, B. and J. E. Walker. 1996. "Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels." Journal of Molecular Biology 260:289-298.

Miyata, S. T., M. Kitaoka, T. M. Brooks, S. B. McAuley and S. Pukatzki. 2011. "*Vibrio cholerae* requires the type VI secretion system virulence factor VasX to kill *Dictyostelium discoideum*." Infection and Immunity 79:2941.

Miyata, S. T., D. Unterweger, S. P. Rudko and S. Pukatzki. 2013. "Dual expression profile of type VI secretion system immunity genes protects pandemic *Vibrio cholerae* (*V. cholerae* T6SS regulation of immunity genes)." PLoS Pathogens 9:e1003752.

Mogensen, J. E. and D. E. Otzen. 2005. "Interactions between folding factors and bacterial outer membrane proteins." Molecular Microbiology 57:326-346.

Mougous, J. D., M. E. Cuff, S. Raunser, A. Shen, M. Zhou, C. A. Gifford, A. L. Goodman, G. Joachimiak, C. L. Ordonez, S. Lory, T. Walz, A. Joachimiak and J. J. Mekalanos. 2006. "A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus." Science 312:1526-1530.

Mougous, J. D., C. A. Gifford, T. L. Ramsdell and J. J. Mekalanos. 2007. "Threonine phosphorylation post-translationally regulates protein secretion in *Pseudomonas aeruginosa*." Nature Cell Biology 9:797.

Mueller, C. A., P. Broz and G. R. Cornelis. 2008. "The type III secretion system tip complex and translocon." Molecular Microbiology 68: 1085-1095.

Mueller, C. A., P. Broz, S. A. Müller, P. Ringler, F. Erne-Brand, I. Sorg, M. Kuhn, A. Engel and G. R. Cornelis. 2005. "The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles." Science 310: 674.

Mulks, M. H. and A. G. Plaut. 1978. "IgA protease production as a characteristic distinguishing pathogenic from harmless *neisseriaceae*." The New England Journal of Medicine 299: 973-976.

Murdoch, S. L., K. Trunk, G. English, M. J. Fritsch, E. Pourkarimi and S. J. Coulthurst. 2011. "The opportunistic pathogen Serratia marcescens utilizes type VI secretion to target bacterial competitors." Journal of Bacteriology 193: 6057.

Nakamoto, H. and J. C. A. Bardwell. 2004. "Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm." BBA - Molecular Cell Research 1694: 111-119.

Nakayama, K., T. Kadowaki, K. Okamoto and K. Yamamoto. 1995. "Construction and characterization of arginine-specific cysteine proteinase (Arg-gingipain)-deficient mutants of Porphyromonas gingivalis. Evidence for significant contribution of Arg-gingipain to virulence." The Journal of Biological Chemistry 270: 23619.

Nano, F. E., N. Zhang, S. C. Cowley, K. E. Klose, K. K. M. Cheung, M. J. Roberts, J. S. Ludu, G. W. Letendre, A. I. Meierovics, G. Stephens and K. L. Elkins. 2004. "A *Francisella tularensis* pathogenicity island required for intramacrophage growth." Journal of Bacteriology 186: 6430-6436.

Nenninger, A. A., L. S. Robinson, N. D. Hammer, E. A. Epstein, M. P. Badtke, S. J. Hultgren and M. R. Chapman. 2011. "CsgE is a curli secretion specificity factor that prevents amyloid fibre aggregation." Molecular Microbiology 81: 486.

Nenninger, A. A., L. S. Robinson and S. J. Hultgren. 2009. "Localized and efficient curli nucleation requires the chaperone-like amyloid assembly protein CsgF." Proceedings of the National Academy of Sciences of the United States of America 106: 900.

Neumann-Haefelin, C., U. Schäfer, M. Müller and H. G. Koch. 2000. "SRP-dependent co-translational targeting and SecA-dependent translocation analyzed as individual steps in the export of a bacterial protein." EMBO Journal 19: 6419-6426.

Ng, T. W., L. Akman, M. Osisami and D. G. Thanassi. 2004. "The usher N terminus is the initial targeting site for chaperone-subunit complexes and participates in subsequent pilus biogenesis events" Journal of Bacteriology 186: 5321-5331.

Nishiyama, M., M. Vetsch, C. Puorger, I. Jelesarov and R Glockshuber. 2003. "Identification and characterization of the chaperone-subunit complex-binding domain from the type 1 pilus assembly platform FimD." Journal of Molecular Biology 330: 513-525.

Nivaskumar, M. and O. Francetic. 2014. "Type II secretion system: a magic beanstalk or a protein escalator." BBA - Molecular Cell Research 1843: 1568-1577.

Olsen, A., A. Jonsson and S. Normark. 1989. "Fibronectin binding mediated by a novel class of surface organelles on *Escherichia coli*." Nature 338: 652-655.

Osipiuk, J., X. Xu, H. Cui, A. Savchenko, A. Edwards and A. Joachimiak. 2011. "Crystal structure of secretory protein Hcp3 from *Pseudomonas aeruginosa*." Journal of Structural and Functional Genomics 12: 21-26.

Pallen, M. J., R. Chaudhuri and A. Khan. 2002. "Bacterial FHA domains: neglected players in the phospho-threonine signalling game?" Trends in Microbiology 10: 556-563.

Palmer, T. and B. C. Berks. 2012. "The twin-arginine translocation (Tat) protein export pathway." Nature Reviews Microbiology 10: 483-496.

Papanikou, E., S. Karamanou and A. Economou. 2007. "Bacterial protein secretion through the translocase nanomachine." Nature Reviews Microbiology 5: 839-851.

Parke, J. L. and D. Gurian-Sherman. 2001. "Diversity of the *Burkholderia cepacia* complex and implications for risk assessment of biological control strains." Annual Review of Phytopathology 39: 225-258.

Parsons, D. A. and F. Heffron. 2005. "sciS, an icmF homolog in *Salmonella enterica* serovar *typhimurium*, limits intracellular replication and decreases virulence." Infection and Immunity 73: 4338-4345.

Pavlova, O., J. H. Peterson, R. Ieva and H. D. Bernstein. 2013. "Mechanistic link between β barrel assembly and the initiation of autotransporter secretion." Proceedings of the National Academy of Sciences of the United States of America 110: E938.

Pell, L. G., V. Kanelis, L. W. Donaldson, P. L. Howell and A. R. Davidson. 2009a. "The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system." Proceedings of the National Academy of Sciences of the United States of America 106: 4160-4165.

Pell, L. G., A. Liu, L. Edmonds, L. W. Donaldson, P. L. Howell and A. R. Davidson. 2009b. "The X-ray crystal structure of the phage lambda tail terminator protein reveals the biologically relevant hexameric ring structure and demonstrates a conserved mechanism of tail termination among diverse long-tailed phages." Journal of Molecular Biology 389: 938-951.

Phan, G., H. Remaut, T. Wang, W. J. Allen, K. F. Pirker, A. Lebedev, N. S. Henderson, S. Geibel, E. Volkan, J. Yan, M. B. Kunze, J. S. Pinkner, B. Ford, C. W. Kay, H. Li, S. J. Hultgren, D. G. Thanassi and G. Waksman. 2011. "Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate." Nature 474: 49.

Pietrosiuk, A., E. D. Lenherr, S. Falk, G. Boenemann, J. Kopp, H. Zentgraf, I. Sinning and A. Mogk. 2011. "Molecular basis for the unique role of the AAA(+) chaperone ClpV in type VI protein secretion." Journal of Biological Chemistry 286: 30010-30021.

Pukatzki, S., A. T. Ma, A.T. Revel, D. Sturtevant and J. J. Mekalanos. 2007. "Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin." Proceedings of the National Academy of Sciences of the United States of America 104: 15508-15513.

Pukatzki, S., A. T. Ma, D. Sturtevant, B. Krastins, D. Sarracino, W. C. Nelson, J. F. Heidelberg and J. J. Mekalanos. 2006. "Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the Dictyostelium host model system." Proceedings of the National Academy of Sciences of the United States of America 103: 1528-1533.

Pukatzki, S., S. B. McAuley and S. T. Miyata. 2009. "The type VI secretion system: translocation of effectors and effector-domains." Current Opinion in Microbiology 12: 11-17.

Py, B., L. Loiseau and F. Barras. 2001. "An inner membrane platform in the type II secretion machinery of Gram-negative bacteria." EMBO Reports 2: 244-248.

Radics, J., L. Konigsmaier and T. C. Marlovits. 2014. "Structure of a pathogenic type 3 secretion system in action." Nature Structural & Molecular Biology 21: 82-87.

Randall, L. L. and S. J. S. Hardy. 2002. "SecB, one small chaperone in the complex milieu of the cell." Cellular and Molecular Life Sciences 59: 1617-1623.

Rao, V. A., S.M. Shepherd, G. English, S. J. Coulthurst and W. N. Hunter. 2011. "The structure of *Serratia marcescens* Lip, a membrane-bound component of the type VI secretion system." Acta Crystallographica Section D 67: 1065-1072.

Rego, A. T., V. Chandran and G. Waksman. 2010. "Two-step and one-step secretion mechanisms in Gram-negative bacteria: contrasting the type IV secretion system and the chaperone-usher pathway of pilus biogenesis." Biochemical Journal 425: 475-488.

Reik, R., T. Spilker and J. J. Lipuma. 2005. "Distribution of *Burkholderia cepacia* complex species among isolates recovered from persons with or without cystic fibrosis." Journal of Clinical Microbiology 43: 2926.

Richarme, G. and T. D. Caldas. 1997. "Chaperone properties of the bacterial periplasmic substrate-binding proteins." The Journal of Biological Chemistry 272: 15607-15612.

Ring, A., J. S. Braun, J. Pohl, V. Nizet, W. Stremmel and J. L. Shenep. 2002. "Group B streptococcal beta-hemolysin induces mortality and liver injury in experimental sepsis." The Journal of Infectious Diseases 185: 1745-1753.

Ripoll-Rozada, J., S. Zunzunegui, F. de La Cruz, I. Arechaga and E. Cabezón. 2013. "Functional interactions of VirB11 traffic ATPases with VirB4 and VirD4 molecular motors in type IV secretion systems." Journal of Bacteriology 195: 4195.

Robb, C. S., M. Assmus, F. E. Nano and A. B. Boraston. 2013. "Structure of the T6SS lipoprotein TssJ1 from *Pseudomonas aeruginosa*." Acta Crystallographica Section F 69: 607-610.

Robb, C. S., F. e. Nano and A. B. Boraston. 2012. "The structure of the conserved Type Six Secretion Protein TssL (DotU) from *Francisella novicida*." Journal of Molecular Biology 419: 277-283.

Robinson, L. S., E. M. Ashman, S. J. Hultgren and M. R. Chapman. 2006. "Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein." Molecular Microbiology 59: 870-881.

Roest, H. P., I. H. Mulders, H. P. Spaink, C. A. Wijffelman and B. J. Lugtenberg. 1997. "A *Rhizobium leguminosarum* biovar trifolii locus not localized on the sym plasmid hinders effective nodulation on plants of the pea cross-inoculation group." Molecular Plant-Microbe Interactions 10: 938-941.

Romling, U., B. Fiedler, J. Bosshammer, D. Grothues, J. Greipel, H. Vonderhardt and B. Tummler. 1994. "Epidemiology of chronic *Pseudomonas aeruginosa* infections in cystic-fibrosis." Journal of Infectious Diseases 170: 1616-1621.

Rosales-Reyes, R., D. F. Aubert, J. S. Tolman, A. O. Amer and M. A. Valvano. 2012a. "*Burkholderia cenocepacia* type VI secretion system mediates escape of type II secreted proteins into the cytoplasm of infected macrophages (T6SS-mediated escape of T2SS secreted proteins)." PLoS One 7: e41726.

Rosales-Reyes, R., A. M. Skeldon, D. F. Aubert and M. A. Valvano. 2012b. "The Type VI secretion system of *Burkholderia cenocepacia* affects multiple Rho family GTPases disrupting the actin cytoskeleton and the assembly of NADPH oxidase complex in macrophages." Cellular Microbiology 14: 255-273.

Rosenberg, A. H., B. N. Lade, D. S. Chui, S. W. Lin, J. J. Dunn and F. W. Studier. 1987. "Vectors for selective expression of cloned DNAs by T7 RNA polymerase." Gene 56: 125-135.

Rossmann, M. G., V. V. Mesyanzhinov, F. Arisaka and P. G. Leiman. 2004. "The bacteriophage T4 DNA injection machine." Current Opinion in Structural Biology 14: 171-180.

Russell, A. B., R. D. Hood, N. K. Bui, M. LeRoux, W. Vollmer and J. D. Mougous. 2011. "Type VI secretion delivers bacteriolytic effectors to target cells." Nature 475: 343-347.

Russell, A. B., M. LeRoux, K. Hathazi, D. M. Agnello, T. Ishikawa, P. A. Wiggins, S. N. Wai and J. D. Mougous. 2013. "Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors." Nature 496: 508-512.

Russell, A. B., S. B. Peterson and J. D. Mougous. 2014. "Type VI secretion system effectors: poisons with a purpose." Nature Reviews Microbiology 12: 137-148.

Russell, A. B., P. Singh, M. Brittnacher, Nk Bui, R. D. Hood, M. A. Carl, D. M. Agnello, S. Schwarz, Dr Goodlett, W. Vollmer and J. D. Mougous. 2012. "A widespread bacterial type VI secretion effector superfamily identified using a heuristic approach." Cell Host & Microbe 11: 538-549.

Ryu, J. H., H. Kim, J. F. Frank and L. R. Beuchat. 2004. "Attachment and biofilm formation on stainless steel by *Escherichia coli* O157: H7 as affected by curli production." Letters in Applied Microbiology 39: 359-362.

Saavedra, M. J., S. Guedes-Novais, A. Alves, P. Rema, M. Tacao, A. Correia and A. Martinez-Murcia. 2004. "Resistance to beta-lactam antibiotics in *Aeromonas hydrophila* isolated from rainbow trout (Oncorhynchus mykiss)." International Microbiology 7: 207-211.

Saiki, K. and K. Konishi. 2007. "Identification of a *Porphyromonas gingivalis* novel protein sov required for the secretion of gingipains." Microbiology and Immunology 51: 483-491.

Saini, L. S., S. B. Galsworthy, M. A. John and M. A. Valvano. 1999. "Intracellular survival of Burkholderia cepacia complex isolates in the presence of macrophage cell activation." Microbiology 145: 3465-3475.

Sajjan, U. S. and J. F. Forstner. 1993. "Role of a 22-kilodalton pilin protein in binding of *Pseudomonas cepacia* to buccal epithelial cells." Infection and Immunity 61: 3157.

Sajjan, U., Y. Wu, G. Kent and J. Forstner. 2000. "Preferential adherence of cable-piliated *Burkholderia cepacia* to respiratory epithelia of CF knockout mice and human cystic fibrosis lung explants." Journal of Medical Microbiology 49: 875.

Saldaña, Z., . Xicohtencatl-cortes, F. Avelino, A. D. Phillips, J. B. Kaper, J. L. Puente and J. A. Girón. 2009. "Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing *Escherichia coli* and identification of Fis as a negative regulator of curli." Environmental Microbiology 11: 992-1006.

Salomon, D., H. Gonzalez, B. Updegraff and K. Orth. 2013. "*Vibrio parahaemolyticus* type VI secretion system 1 is activated in marine conditions to target bacteria, and is differentially regulated from system 2." PLoS One 8:e61086.

Sana, T. G., C. Soscia, A. Hachani, R. Voulhoux, S. Bleves, C. Baumann, K. L. Bennett, G. Superti-Furga, A. Merdes, T. Rattei, C. Jones and A. Filloux. 2015. "Internalization of Pseudomonas aeruginosa strain PAO1 into epithelial cells is promoted by interaction of a T6SS effector with the microtubule network." mBio 6: e00712-15.

Sangodkar, U. M., P. J. Chapman and A. M. Chakrabarty. 1988. "Cloning, physical mapping and expression of chromosomal genes specifying degradation of the herbicide 2,4,5-T by Pseudomonas cepacia AC1100." Gene 71: 267-277.

Sargent, F., U. Gohlke, E. De Leeuw, N. R. Stanley, T. Palmer, H. R. Saibil and B. C. Berks. 2001. "Purified components of the Escherichia coli Tat protein transport system form a double-layered ring structure." European Journal of Biochemistry 268: 3361-3367.

Sato, K., M. Naito, H. Yukitake, H. Hirakawa, M. Shoji, M. J. McBride, R. G. Rhodes and K. Nakayama. 2010. "A protein secretion system linked to bacteroidete gliding motility and pathogenesis." Proceedings of the National Academy of Sciences of the United States of America 107: 276.

Sato, K., H. Yukitake, Y. Narita, M. Shoji, M. Naito and K. Nakayama. 2013. "Identification of P orphyromonas gingivalis proteins secreted by the P or secretion system." FEMS Microbiology Letters 338: 68-76.

Schell, M. A., R. L. Ulrich, W. J. Ribot, E. E. Brueggemann, H. B. Hines, D. Chen, L. Lipscomb, H. S. Kim, J. Mrazek, W. C. Nierman and D. DeShazer. 2007. "Type VI secretion is a major virulence determinant in *Burkholderia mallei*." Molecular Microbiology 64: 1466-1485.

Schiebel, E., A. J. Driessen, F. U. Hartl and W. Wickner. 1991. "Δμ H+ and ATP function at different steps of the catalytic cycle of preprotein translocase." Cell 64: 927-939.

Schierle, C. F., M. Berkmen, D. Huber, D. Boyd, J. Beckwith and C. Kumamoto. 2003. "The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the *Escherichia coli* periplasm via the signal recognition particle pathway." Journal of Bacteriology 185: 5706-5713.

Schlieker, C., H. Zentgraf, P. Dersch and A. Mogk. 2005. "ClpV, a unique Hsp100/Clp member of pathogenic proteobacteria." Biological Chemistry 386: 1115-1127.

Schwab, U., M. Leigh, C. Ribeiro, J. Yankaskas, K. Burns, P. Gilligan, P. Sokol and R. Boucher. 2002. "Patterns of epithelial cell invasion by different species of the *Burkholderia cepacia* complex in well-differentiated human airway epithelia." Infection and Immunity 70: 4547.

Schwarz, S., P. Singh, J. D. Robertson, M. Leroux, S. J. Skerrett, D. R. Goodlett, T. E. West and J. D. Mougous. 2014. "VgrG-5 is a Burkholderia type VI secretion system-exported protein required for multinucleated giant cell formation and virulence." Infection and Immunity 82: 1445.

Schwarz, S., T. E. West, F. Boyer, W. C. Chiang, M. A. Carl, R. D. Hood, L. Rohmer, T. Tolker-Nielsen, S. J. Skerrett and J. D. Mougous. 2010. "*Burkholderia* type VI secretion systems have distinct roles in eukaryotic and bacterial cell interactions (*B. thai* T6SSs critical for host and interspecies bacterial interactions)." PLoS Pathogens 6: e1001068.

Seers, C. A., N. Slakeski, P. D. Veith, T. Nikolof, Y. Y. Chen, S. G. Dashper and E. C. Reynolds. 2006. "The RgpB C-terminal domain has a role in attachment of RgpB to the outer membrane and belongs to a novel C-terminal-domain family found in *Porphyromonas gingivalis*." Journal of Bacteriology 188: 6376.

Segal, G., M. Feldman and T. Zusman. 2005. "The Icm/Dot type-IV secretion systems of *Legionella pneumophila* and *Coxiella burnetii*." FEMS Microbiology Reviews 29: 65-81.

Selkrig, J., K. osbahi, C. T. Webb, M. J. Belousoff, A. J. Perry, T. J. Wells, F. Morris, D. L. Leyton, M. Totsika, M. D. Phan, N. Celik, M. Kelly, C. Oates, E. L. Hartland, R. M. Robins-Browne, S. H. Ramarathinam, A. W Purcell, M. A. Schembri, R. A. Strugnell, I. R. Henderson, D. Walker and T. Lithgow. 2012. "Discovery of an archetypal protein transport system in bacterial outer membranes." Nature Structural & Molecular Biology 19:506.

Seshadri, R., S. W. Joseph, A. K. Chopra, J. Sha, J. Shaw, J. Graf, D. Haft, M. Wu, Q. Ren, M. J. Rosovitz, R. Madupu, L. Tallon, M. Kim, S. Jin, H. Vuong, O. C. Stine, A. Ali, A. J. Horneman and J. F. Heidelberg. 2006. "Genome sequence of *Aeromonas hydrophila* ATCC 7966T: jack of all trades." Journal of Bacteriology 188: 8272-8282.

Sha, J. , S. F. Wang, G. Suarez, J. C. Sierra, A. A. Fadl, T. E. Erova, S. M. Foltz, B. K. Khajanchi, A. Silver, J. Graf, C. H. Schein and A.K. Chopra. 2007. "Further characterization of a type III secretion system (T3SS) and of a new effector protein from a clinical isolate of *Aeromonas hydrophila*-Part I." Microbial Pathogenesis 43: 127-146.

Shalom, G., J. G. Shaw and M. S. Thomas. 2007. "In vivo expression technology identifies a type VI secretion system locus in *Burkholderia pseudomallei* that is induced upon invasion of macrophages." Microbiology 153: 2689-2699.

Shastri, S. 2011. "Characterisation of the Type VI secretion system of Burkholderia cenocepacia." In Department of Infection and Immunity: University of Sheffield.

Shevchik, V. E., J. Robert-Baudouy and G. Condemine. 1997. "Specific interaction between OutD, an *Erwinia chrysanthemi* outer membrane protein of the general secretory pathway, and secreted proteins." EMBO Journal 16: 3007-3016.

Shieh, Y. W., P. Minguez, P. Bork, J. J. Auburger, D. L. Guilbride, G. Kramer and B. Bukau. 2015. "Operon structure and cotranslational subunit association direct protein assembly in bacteria." Science 350: 678-680.

Shimomura, H., M. Matsuura, S. Saito, Y. Hirai, Y. Isshiki and K. Kawahara. 2003. "Unusual interaction of a lipopolysaccharide isolated from *Burkholderia cepacia* with polymyxin B." Infection and Immunity 71: 5225-5230.

Shneider, M. M., S. A. Buth, B. T. Ho, M. Basler, J. J. Mekalanos and P. G. Leiman. 2013. "PAAR-repeat proteins sharpen and diversify the type VI secretion system spike." Nature 500: 350-353.

Sijbrandi, R., M. L. Urbanus, C. M. Ten Hagen-Jongman, B. Oudega, B. R. Otto, J. Luirink and H. D. Bernstein. 2003. "Signal recognition particle (SRP)-mediated targeting and Sec-dependent translocation of an extracellular *Escherichia coli* protein." The Journal of Biological Chemistry 278: 4654-4659.

Silverman, J. M., L. S. Austin, F. Hsu, K. G. Hicks, R. D. Hood and J. D. Mougous. 2011. "Separate inputs modulate phosphorylation -dependent and -independent type VI secretion activation." Molecular Microbiology 82: 1277-1290.

Silverman, J. M., D. M. Agnello, H. Zheng, B. T. Andrews, M. Li, C. E. Catalano, T. Gonen and J. D. Mougous. 2013. "Haemolysin coregulated protein is an exported receptor and chaperone of type VI secretion substrates." Molecular Cell 51: 584-593.

Simeone, R., D. Bottai and R. Brosch. 2008. "ESX/type VII secretion systems and their role in host-pathogen interaction." Current Opinion in Microbiology 12: 4-10.

Simon, R., U. Priefer and A. Puhler. 1983. "A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram-negative bacteria." Nature Biotechnology 1: 784-791.

Skillman, K. M., T. J. Barnard, J.H. Peterson, R. Ghirlando and H. D. Bernstein. 2005. "Efficient secretion of a folded protein domain by a monomeric bacterial autotransporter." Molecular Microbiology 58: 945-958.

Spiewak, H. L. 2015. "Identifying novel secreted effectors of the type VI protein secretion system in Burkholderia cenocepacia." In Department of Infection, Immunity & Cardiovascular Diseases: University of Sheffield.

Stathopoulos, C., D. R. Hendrixson, D. G. Thanassi, S. J. Hultgren, J. W. St Geme and R. Curtiss. 2000. "Secretion of virulence determinants by the general secretory pathway in Gram-negative pathogens: an evolving story." Microbes and Infection 2: 1061-1072.

Stintzi, A. and K. N. Raymond. 2000. "Amonabactin-mediated iron acquisition from transferrin and lactoferrin by *Aeromonas hydrophila* : direct measurement of individual microscopic rate constants." Journal of Biological Inorganic Chemistry 5: 57-66.

Studier, F. W. and B. A. Moffatt. 1986. "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." Journal of Molecular Biology 189: 113-130.

Studier, F. W. 2005 "Protein production by auto-induction in high-density shaking cultures." Protein Expression and Purification 41: 207-234.

Stulik, K., V. Pacakova and M. Ticha. 2003. "Some potentialities and drawbacks of contemporary size-exclusion chromatography." Journal of Biochemical and Biophysical Methods 56: 1-13.

Suarez, G., J. C. Sierra, T. E. Erova, J. Sha, A. J. Horneman and A. K. Chopra. 2010a. "A type VI secretion system effector protein, VgrG1, from *Aeromonas hydrophila* that induces host cell toxicity by ADP ribosylation of actin." Journal of Bacteriology 192: 155-168.

Suarez, G., J. C. Sierra, M. L. Kirtley and A. K. Chopra. 2010b. "Role of Hcp, a type 6 secretion system effector, of *Aeromonas hydrophila* in modulating activation of host immune cells." Microbiology 156: 3678-3688.

Suarez, G., J. C. Sierra, J. Sha, S. F. Wang, T. E. Erova, A. A. Fadl, S. M. Foltz, A. J. Horneman and A. K. Chopra. 2008. "Molecular characterization of a functional type VI secretion system from a clinical isolate of *Aeromonas hydrophila*." Microbial Pathogenesis 44: 344-361.

Tarry, M. J., E. Schäfer, S. Chen, G. Buchanan, N. P. Greene, S. M. Lea, T. Palmer, H. R. Saibil and B. C. Berks. 2009. "Structural analysis of substrate binding by the TatBC component of the twin-arginine protein transport system." Proceedings of the National Academy of Sciences of the United States of America 106: 13284.

Taylor, N. M., N. S. Prokhorov, R. C. Guerrero-Ferreira, M. M. Shneider, C. Browning, K. N. Goldie, H. Stahlberg and P. G. Leiman. 2016. "Structure of the T4 baseplate and its function in triggering sheath contraction." Nature 533: 346-352.

Thanassi, D. G., C. Stathopoulos, A. Karkal and H. Li. 2005. "Protein secretion in the absence of ATP: the autotransporter, two-partner secretion and chaperone/usher pathways of Gram-negative bacteria (review)." Molecular Membrane Biology 22:63-72.

Thomas, M. S. 2007. "Iron acquisition mechanisms of the *Burkholderia cepacia* complex." Biometals 20: 431-452.

Toesca, I. J., C. T. French and J. F. Miller. 2014. "The type VI secretion system spike protein VgrG5 mediates membrane fusion during intercellular spread by *pseudomallei* group *Burkholderia* species." Infection and Immunity 82: 1436.

Tomich, M., A. Griffith, C. A. Herfst, J. Burns and C. Mohr. 2003. "Attenuated virulence of a *Burkholderia cepacia* type III secretion mutant in a murine model of infection." Infection and Immunity 71: 1405-1415.

Tomich, M., C. A. Herfst, J. Golden and C. Mohr. 2002. "Role of flagella in host cell invasion by *Burkholderia cepacia*." Infection and Immunity 70:1799-1806.

Trokter, M., C. Felisberto-Rodrigues, P. J. Christie and G. Waksman. 2014. "Recent advances in the structural and molecular biology of type IV secretion systems." Current Opinion in Structural Biology 27: 16-23.

Tseng, T. T., B. M. Tyler and J. C. Setubal. 2009. "Protein secretion systems in bacterial-host associations, and their description in the gene ontology." BMC Microbiology 9: S2.

Uchida, K., P. G. Leiman, F. Arisaka and S. Kanamaru. 2014. "Structure and properties of the C-terminal β-helical domain of VgrG protein from O157." The Journal of Biochemistry 155: 173-182.

Ullers, R. S., J. Luirink, N. Harms, F. Schwager, C. Georgopoulos, P. Genevaux and L. L. Randall. 2004. "SecB is a bona fide generalized chaperone in *Escherichia coli*." Proceedings of the National Academy of Sciences of the United States of America 101: 7583-7588.

Unterweger, D., B. Kostiuk, R. Ötjengerdes, A. Wilton, L. Diaz-Satizabal and S. Pukatzki. 2015. "Chimeric adaptor proteins translocate diverse type VI secretion system effectors in *Vibrio cholerae*." EMBO Journal 34: 2198-2210.

Unterweger, D., S. T. Miyata, V. Bachmann, T. M. Brooks, T. Mullins, B. Kostiuk, D. Provenzano and S. Pukatzki. 2014. "The *Vibrio cholerae* type VI secretion system employs diverse effector modules for intraspecific competition." Nature Communications 5: 3549.

Urban, T. A., J. B. Goldberg, J. F. Forstner and U. S. Sajjan. 2005. "Cable pili and the 22-kilodalton adhesin are required for *Burkholderia cenocepacia* binding to and transmigration across the squamous epithelium." Infection and Immunity 73: 5426.

Urban, T., A. Griffith, A. Torok, M. Smolkin, J. Burns and J. Goldberg. 2004. "Contribution of *Burkholderia cenocepacia* flagella to infectivity and inflammation." Infection and Immunity 72: 5126-5134.

Van Der Wolk, J. P., J. G. De Wit and A. J. Driessen. 1997. "The catalytic cycle of the *Escherichia coli* SecA ATPase comprises two distinct preprotein translocation events." EMBO Journal 16: 7297-7304.

Vandamme, P., D. Henry, T. Coenye, S. Nzula, M. Vancanneyt, J. J. LiPuma, D. P. Speert, J. R. Govan and E. Mahenthiralingam. 2002. "*Burkholderia anthina* sp. nov. and *Burkholderia pyrrocinia*, two additional *Burkholderia cepacia* complex bacteria, may confound results of new molecular diagnostic tools." FEMS Immunology and Medical Microbiology 33: 143-149.

Vandamme, P., B. Holmes, T. Coenye, J. Goris, E. Mahenthiralingam, J. J. Lipuma and J. R. Govan. 2003. "*Burkholderia cenocepacia* sp. nov. -a new twist to an old story." Research in Microbiology 154: 91-96.

Vandamme, P., B. Holmes, M. Vancanneyt, T. Coenye, B. Hoste, R. Coopman, H. Revets, S. Lauwers, M. Gillis, K. Kersters and J. R. Govan. 1997. "Occurrence of multiple genomovars of *Burkholderia cepacia* in cystic fibrosis patients and proposal of *Burkholderia multivorans* sp. nov." International Journal of Systematic and Evolutionary Microbiology 47: 1188-1200.

Vandamme, P., E. Mahenthiralingam, B. Holmes, T. Coenye, B. Hoste, P. De Vos, D. Henry and D. P. Speert. 2000. "Identification and population structure of *Burkholderia stabilis* sp. nov. (formerly *Burkholderia cepacia* genomovar IV)." Journal of Clinical Microbiology 38: 1042.

Veith, P. D., Y. Y. Chen, D. G. Gorasia, D. Chen, M. D. Glew, N. M. Brien-Simpson, J. D. Cecil, J. A. Holden and E. C. Reynolds. 2014. "Porphyromonas gingivalis outer membrane vesicles exclusively contain outer membrane and periplasmic proteins and carry a cargo enriched with virulence factors." Journal of Proteome Research 13: 2420.

Vetsch, M., C. Puorger, T. Spirig, U. Grauschopf, E. U. Weber-Ban and R. Glockshuber. 2004. "Pilus chaperones represent a new type of protein-folding catalyst." Nature 431: 329-333.

Vidal, O., R. Longin, C. Prigent-Combaret, C. Dorel, M. Hooreman and P. Lejeune. 1998. "Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new ompR allele that increases curli expression." Journal of Bacteriology 180: 2442-2449.

Vilches, S., M. Wilhelms, H. B. Yu, K. Y. Leung, J. M. Tomás and S. Merino. 2008. "*Aeromonas hydrophila* AH-3 AexT is an ADP-ribosylating toxin secreted through the type III secretion system." Microbial Pathogenesis 44: 1-12.

Visser, M. B., S. Majumdar, E. Hani and P. A. Sokol. 2004. "Importance of the ornibactin and pyochelin siderophore transport systems in *Burkholderia cenocepacia* lung infections." Infection and Immunity 72: 2850-2857.

Von Heijne, G. 1985. "Signal sequences. The limits of variation." Journal of Molecular Biology 184: 99.

Wang, J., H. Yang, A. Mushegian and S. Jin. 1998. "A novel serine/threonine protein kinase homologue of *Pseudomonas aeruginosa* is specifically inducible within the host infection site and is required for full virulence in neutropenic mice." Journal of Bacteriology 180: 6764-6768.

Wang, X., D. R. Smith, J. W. Jones and M. R. Chapman. 2007. "In vitro polymerization of a functional *Escherichia coli* amyloid protein." The Journal of Biological Chemistry 282: 3713.

Weber, B., A. Croxatto, C. Chen and .Debra L. Milton. 2008. "RpoS induces expression of the *Vibrio anguillarum* quorum-sensing regulator VanT." Microbiology 154: 767-780.

Whitney, J. C., C. M. Beck, Y. A. Goo, A. B. Russell, B. N. Harding, J. A. De Leon, D. A. Cunningham, B. Q. Tran, D. A. Low, D. R. Goodlett, C. S. Hayes and J. D. Mougous. 2014. "Genetically distinct pathways guide effector export through the type VI secretion system." Molecular Microbiology 92: 529-42.

Whitney, J. C., S. Chou, A. B. Russell, J. Biboy, T. E. Gardiner, M. A. Ferrin, M. Brittnacher, W. Vollmer and J. D. Mougous. 2013. "Identification, structure, and function of a novel type VI secretion peptidoglycan glycoside hydrolase effector-immunity pair." The Journal of Biological Chemistry 288: 26616.

Willems, R. J., C. Geuijen, H. G. Heide, G. Renauld, P. Berlin, W. M. Akker, C. Locht and F. R. Mooi. 1994. "Mutational analysis of the *Bordetella pertussis* fim/fha gene cluster: identification of a gene with sequence similarities to haemolysin accessory genes involved in export of FHA." Molecular Microbiology 11: 337-347.

Williams, S. G., L. T. Varcoe, S. R. Attridge and P. A. Manning. 1996. "*Vibrio cholerae* Hcp, a secreted protein coregulated with HlyA." Infection and Immunity 64:283-289.

Woese, C. R. 1987. "Bacterial evolution." Microbiological Reviews 51: 221-271.

Wright, K. J., P. C. Seed and S. J. Hultgren. 2007. "Development of intracellular bacterial communities of uropathogenic *Escherichia coli* depends on type 1 pili." Cellular Microbiology 9: 2230-2241.

Wyatt, P. J. 1991. "Combined differential light scattering with various liquid chromatography separation techniques." Biochemical Society Transactions 19: 485.

Xu, X. J., M. Ferguson, V. Popov, C. Houston, J. Peterson and A. Chopra. 1998. "Role of a cytotoxic enterotoxin in *Aeromonas*-mediated infections: development of transposon and isogenic mutants." Infection and Immunity 66: 3501-3509.

Yabuuchi, E., Y. Kosako, H. Oyaizu, I. Yano, H. Hotta, Y. Hashimoto, T. Ezaki and M. Arakawa. 1992. "Proposal of *Burkholderia gen*. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species *Burkholderia cepacia* (Palleroni and Holmes 1981) comb. nov." Microbiology and Immunology 36: 1251-1275.

Yanisch-Perron, C., J. Vieira and J. Messing. 1985. "Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors." Gene 33: 103-119.

Yap, M. L., K. Mio, P. G. Leiman, S. Kanamaru and F. Arisaka. 2010. "The baseplate wedges of bacteriophage T4 spontaneously assemble into hubless baseplate-like structure in vitro." Journal of Molecular Biology 395: 349-360.

Yen, Y. T., M. Bhattacharya and C. Stathopoulos. 2008. "Genome-wide in silico mapping of the secretome in pathogenic *Yersinia pestis* KIM." FEMS Microbiology Letters 279: 56-63.

Yu, Y., H. S. Kim, H. H. Chua, C. H. Lin, S. H. Sim, D. Lin, A. Derr, R. Engels, D. Deshazer, B. Birren, W. C. Nierman and P. Tan. 2006. "Genomic patterns of pathogen evolution revealed by comparison of *Burkholderia pseudomallei*, the causative agent of melioidosis, to avirulent B*urkholderia thailandensis*." BMC Microbiology 6: 46.

Yukihiro, A. and E. G. Jorge. 2005. "Chaperone release and unfolding of substrates in type III secretion." Nature 437: 911.

Zahariadis, G., M. H. Levy and J. L. Burns. 2003. "Cepacia-like syndrome caused by *Burkholderia multivorans*." Canadian Journal of Infectious Diseases and Medical Microbiology 14: 123.

Zhang, X. Y., Y. R. Brunet, L. Logger, L. Journet, E. Cascales, B. Douzi and C. Cambillau. 2013. "Dissection of the TssB-TssC interface during type VI secretion sheath complex formation." PLoS One 8: e81074.

Zhang, Y. L., E. Arakawa and K. Y. Leung. 2002. "Novel *Aeromonas hydrophila* PPD134/91 genes involved in O-antigen and capsule biosynthesis." Infection and Immunity 70: 2326-2335.

Zheng, J., B. Ho and J. J. Mekalanos. 2011. "Genetic analysis of anti-amoebae and anti-bacterial activities of the type VI secretion system in Vibrio cholerae." PLoS One 6: e23876.

Zheng, J. and K. Y. Leung. 2007. "Dissection of a type VI secretion system in *Edwardsiella tarda*." Molecular Microbiology 66: 1192-1206.

Zheng, J., O. S. Shin, D. E. Cameron and J. J. Mekalanos. 2010. "Quorum sensing and a global regulator TsrA control expression of type VI secretion and virulence in *Vibrio cholerae*." Proceedings of the National Academy of Sciences of the United States of America 107: 21128.

Zoued, A., Y. R. Brunet, E. Durand, M. S. Aschtgen, L. Logger, B. Douzi, L. Journet, C. Cambillau and E. Cascales. 2014. "Architecture and assembly of the type VI secretion system." BBA - Molecular Cell Research 1843: 1664-1673.

Zoued, A., E. Durand, C. Bebeacua, Y. R. Brunet, B. Douzi, C. Cambillau, E. Cascales and L. Journet. 2013. "TssK is a trimeric cytoplasmic protein interacting with components of both phage-like and membrane anchoring complexes of the type VI secretion system." The Journal of Biological Chemistry 288: 27031.

Zoued, A., E. Durand, Y. R. Brunet, S. Spinelli, B. Douzi, M. Guzzo, N. Flaugnatti, P. Legrand, L. Journet, R. Fronzes, T. Mignot, C. Cambillau and E. Cascales. 2016. "Priming and polymerization of a bacterial contractile tail structure." Nature 531: 59-63.

Zupan, J., C. A. Hackworth, J. Aguilar, D. Ward and P. Zambryski. 2007. "VirB1* promotes T-pilus formation in the vir-type IV secretion system of *Agrobacterium tumefaciens*." Journal of Bacteriology 189: 6551-6563.