Recombinant Expression and Characterisation of the Colicin N Immunity Protein

Daria Stroukova

Thesis submitted for the degree of

Doctor of Philosophy

Institute for Cell and Molecular Biosciences

September 2015

Abstract

Some *Escherichia coli* express plasmid encoded toxins called colicins to eliminate ecological competitors. Colicins are a diverse group of toxins, usually divided into three groups: nucleases, pore-formers and peptidoglycan synthesis inhibitors. In addition, each colicinogenic plasmid codes for an immunity protein - a self-protection mechanism against its own toxin. Immunity proteins are highly specific to their cognate toxins but it is not clearly known how immunity proteins to pore-forming colicins achieve protection.

This work focuses on the immunity protein for the smallest pore-forming colicin, colicin N. To investigate *in vitro* how the Colicin N immunity protein (CNI) neutralises the toxin, different overexpression and purification methods were tested. The fusion CNI-3C-HALO7-6His yielded the most protein, when expressed in C41 *E. coli* cells at 37 °C overnight in Terrific Broth. The fusion protein increases resistance to Colicin N dramatically and is therefore folded and localised correctly. Decyl β-D-maltopyranoside is the most effective detergent to solubilise and stabilise the protein fusion. To investigate if CNI inactivates ColN by binding to it via hydrophobic α-helical interactions, the protein's helical regions were defined *in silico* and swapped with the respective parts of the Colicin A immunity protein, CNI's closest homologue, which provides immunity to Colicin A. The N-terminal region of CNI, including the first transmembrane helix, is not necessary for specific CNI-Colicin N interaction but the other helices are crucial for protein stability and function. CNI can be expressed in an active form fused to GFP for microscopy. Epifluorescent and TIRF microscopy revealed that CNI is localised to the cell periphery and appears to be very evenly distributed in the cell membrane. Colicin N addition does not affect CNI distribution. Single molecule tracking in TIRF microscopy showed that the diffusion rate of CNI does not change when ColN is added.

Future work can pursue either of the three approaches used here to study CNI using *in vitro* structural analysis, site-directed mutagenesis of the active site or *in vivo* interaction of CNI with other *E. coli* inner membrane proteins.

I dedicate this work to my mother, Irina Stroukova, the most courageous woman I know and always my best friend and idol, and my late grandfather, Naum Shveisky, whose dream of becoming a biochemist I fulfilled.

What is the most resilient parasite? Bacteria? A virus? An intestinal worm? An idea. Resilient... highly contagious. Once an idea has taken hold of the brain it's almost impossible to eradicate. An idea that is fully formed - fully understood - that sticks; right in there somewhere*.*

(Cobb, Inception)

Acknowledgements

I would like to thank Professor Jeremy Lakey for the opportunity to work in his lab and his advice and support during my project as well as for reviewing this thesis. Jeremy, it is your never ceasing optimism and calmness in the most hopeless of situations which inspired and convinced me to finish this project. I thank the BBSRC for funding this project, which allowed me to gain further scientific expertise and academic knowledge as well as to grow as a person and develop transferable skills for all walks of life.

I would like to thank Dr Henrik Strahl for providing the strain MC1000, the plasmid pBAD322-WALP23-eGFP and his patience and dedication in helping me with epifluorescent and TIRF microscopy and data analysis at Newcastle University as well as for reviewing part of this thesis. A special thanks goes to my collaborators and helpers at the Research Campus Harwell, namely Dr Isabel DeMoraes and Mr James Birch at MPL for their help in detergent screening, Dr Ray Owens, Mrs Heather Rada, Dr Louise Bird at OPPF for their help with expression system optimisation and Dr Marisa Martin-Fernandez, Dr Stephen Webb and Dr Daniel Rolfe at CLF for their help with single particle tracking and data analysis as well as reviewing part of this thesis.

I cannot possibly thank Dr Helen “world’s best technician” Waller enough for her help with everything – science and life, her patient teaching and explaining and the care and support she provided me with over the years as well as all the Geordie phrases I have learned from this local lass. I thank Dr Christopher Johnson for his invaluable support at the beginning of my PhD and for re-sequencing and synthesising CNI. I thank my colleagues and dear friends Dr Wanatchaporn Arunmanee, Dr Javier Abellon Ruiz, Dr Hannah Alfonsa and Dr Yakup Ulusu as well as all other members of the Lakey lab for supporting me scientifically and keeping me sane through all those difficult days, late evenings and sunny weekends in the lab and for all the dinners, coffees, biscuits, cakes, tears and laughs we shared.

Finally, I would like to thank my family and friends who have always believed in me, supported and advised me from near and far and gave me the strength, courage and confidence to complete this project.

Contents

Table of Figures

Table of tables

List of abbreviations

Abbreviation	Explanation
%	percent
°C	degrees Celsius
µg	microgram
µl	microlitre
µM	micromolar
AEBSF	4-(2-Aminoethyl)benzenesulfonyl fluoride hydrochloride
APS	Ammonium persulfate
bp	base pair
CAI	Colicin A Immunity protein
CCCP	Carbonyl cyanide m-chlorophenyl hydrazone
CCCP	Carbonyl cyanide m-chlorophenyl hydrazone
CD	Circular Dichroism
CMC	critical micelle concentration
CNI	Colicin N Immunity protein
CNI-sfGFP-6His	fusion protein of CNI and superfolder GFP with a C-terminal polyhistidine tag
ColA	Colicin A
ColA P-domain	Poreforming domain of Colicin A
ColA R-domain	Receptor-binding domain of Colicin A
ColA T-domain	Translocation domain of Colicin A
ColB	Colicin B
ColE1	Colicin E1
ColE1	Colicin E1
ColN	Colicin N
ColN P-domain	Pore-forming domain of Colicin N
ColN R-domain	Receptor-binding domain of Colicin N
ColN T-domain	Translocation domain of Colicin N
DDM	n-Dodecyl-β-maltoside detergent
DEPC	diethyl-pyrocarbonate
dH2O	distilled H2O
DM	n-decyl-β-maltoside detergent
DNA	deoxyribonucleic acid
dNTPs	deoxyribonuleotide triphosphates
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
eGFP	enhanced green fluorescent protein
FSEC	fluorescence size exclusion chromatography
g	gram
GFP	Green fluoroescent protein
GST	Glutathione S-transferase fusion protein
h	hours
HALO7	Halogenase fusion protein
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His-tag	polyhistidine tag

IMAC	immobilised metal ion afficinity chromatography
IPTG	isopropyl ß-D-1-thiogalactopyranoside
kDa	kilodalton
l	litre
LB	lysogeny broth
LC-MS	Liquid Chromatography Mass Spectrometry
LDAO	Lauryldimethylamine N-oxide
LPS	lipopolysacharide
MBP	maltose binding protein
MIC	Minimal inhibitory concentration
min	minutes
ml	millilitre
mM	millimolar
MSD	mean square displacement
MYSB	mysB protein
NaP	Mixture of monosodium and disodium phosphate to adjust pH
nm	nanometer
NTA	Nitrilotriacetic acid
OD600 nm	optical density at 600 nm
OG	n -octyl-β- D -glucoside
PAGE	poly-acrylamide gel electrophoresis
Para	arabinose inducible promoter
PB	power broth
PBS	phosphate buffered saline
Pcni	natural promoter for CNI
PCR	polymerase chain reaction
pH	potential of hydrogen
SDS	Sodium dodecyl sulfate
secs	seconds
sfGFP	superfolder green fluorescent protein
SPR	surface plasmon resonance
SS	signal sequence
SUMO	Small Ubiquitin-like Modifier protein
TB	terrific broth
TCEP	tris(2-carboxyethyl)phosphine
TCS	thrombin cleavage site
TEMED	N,N,N',N' - tetramethylethylenediamine
TEV protease	Tobacco etch virus protease
TIRF	total internal reflection fluorescence microscopy
Tris	Tris(hydroxymethyl)aminomethane
U	unit
UV	ultraviolet
WT	wild type

Chapter 1. Introduction

Antimicrobial agents are produced by both eukaryotic and prokaryotic organisms and can be broadly divided into two types, the classic enzymatically produced small molecule antibiotics and polypeptides produced by ribosomes. Each type differs in mode of action as well as ecological purpose (Chao and Levin, 1981; Riley and Wertz, 2002a). While small molecules antibiotics usually require micromolar concentrations to be effective against a broad range of targets, polypeptide antimicrobials are mostly targeted at members of the same species and are lethal at nanomolar concentrations. Antimicrobial peptides are a common weapon used by microbes to kill their ecological competitors, crucially influencing microbial population dynamics (Riley, 1993; Gordon *et al.*, 1998; Riley, 1998; Czaran *et al.*, 2002; Kerr *et al.*, 2002; Riley and Wertz, 2002a; Riley *et al.*, 2012). They can maintain biodiversity within a population by preventing one species or strain from taking over, giving a strain the opportunity to invade new environments or preventing the invasion of alien strains. However, the production of antimicrobial compounds only provides a competitive advantage if the antibiotic is highly efficient and able to provide an advantage great enough to justify its production cost. Therefore, polypeptide antimicrobials tend to be more prominent in nutrient rich environments and a weapon of choice when invading an environment (Brown *et al.*, 2009). This advantage is easier to achieve in an environment where resources are not distributed randomly and where toxin release directly results in nutritional gain. In addition, an equally efficient self-defence mechanism must ensure the bacteriocin producer's own safety (Alonso *et al.*, 2000).

Antimicrobial peptides produced by bacteria against bacteria are called bacteriocins and can be divided into two groups based on their producers – Gram positive and Gram negative bacteria (Gordon and O'Brien, 2006; Heng et al., 2007). Bacteriocins from Gram positive bacteria include the heat-stable lantibiotics and have become a source of interest because they can be efficient against multidrug-resistant pathogens such as methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant enterococcus (VRE) (Breukink and de Kruijff, 1999; Cotter et al., 2005; Guinane et al., 2005; Cotter et al., 2006; Diep et al., 2006; Nes et al., 2007; Cotter et al., 2013; Mathur et al., 2015).

Gram negative bacteriocin producers are mainly found amongst the Enterobacteriaceae, which produce antibacterial proteins (30–80 kDa) as well as smaller peptides (1 - 10 kDa), named microcins. Both larger bacteriocins and microcins are produced in response to stress, but while larger bacteriocin production is associated with DNA damage, the microcin production can be triggered by nutrient depletion (Rebuffat, 2011). The most prominently studied bacteriocins from Gram negative bacteria are Colicins, produced by 30–50% of the strains isolated from human hosts (Pugsley, 1984c; Pugsley, 1984b; James *et al.*, 1987; Pugsley, 1987b; Benedetti *et al.*, 1991b; Cascales *et al.*, 2007; Barneoud-Arnoulet *et al.*, 2010; Lloubes *et al.*, 2012; Kim *et al.*, 2014). Other well-studied bacteriocins from Gram negative bacteria include the pyocins produced in *Pseudomonas aeruginosa* (Michel-Briand and Baysse, 2002).

1.1 Potential human uses for colicins

Colicins are an example of highly effective bacteriocins produced by *Escherichia coli* (*E. coli*) to kill related strains. Apart from being a valuable weapon in microbial warfare, colicins can be useful antimicrobial agents for human purposes. Bacteriocins have been and continue to be an excellent source of probiotics and antibiotics (Breukink and de Kruijff, 1999; Joerger, 2003; Audisio *et al.*, 2005; Corr *et al.*, 2007).

Schulz *et al.* (2015) reported that Colicin M alone and in combination with Colicin E7 was highly effective in reducing pathogenic *E. coli load*, such as enterohemorrhagic *Escherichia coli* (EHEC), on animal feedstock and human food and, therefore, could be an alternative to heat inactivation. EHEC is responsible for thousands of bacterial enteric infections worldwide, which can lead to hospitalisation and death. These colicins can be produced cost effectively in edible plants but commercial production still requires regulatory approval. In contrast to small molecule antibiotics colicins have a narrow target range, can be deactivated through heat and digestion and are bactericidal in much lower doses, suggesting that resistance would develop slower. Their use would not be harmful to the beneficial gut flora. As well as making feedstock and food more safe, colicinogenic non-pathogenic *E. coli* can also be applied onto medical equipment, like catheters, as a kind of pro-biotic. Trautner *et al.* (2005) showed that a colicinogenic *E. coli* K-12 strain prevents the population of urinary catheters by sensitive uropathogenic *E. coli.*

Colicins could also have medical applications. Brown *et al.* (2012) suggests the use of Colicin-like pyocins on *Pseudomonas aeruginosa* biofilms based on preliminary data in the *Galleria mellonella* caterpillar model. Biofilms are difficult to treat with conventional antibiotics but play a role in common diseases like chronic rhinosinusitis, bacterial endocarditis, *Pseudomonas aeruginosa* lung infection in patients with cystic fibrosis, recurrent urinary tract infections and Crohn's disease. Brown *et al.* (2015) investigated the use of Colicins E1, E3, E9, and D as an alternative to conventional antibiotics to treat adherent-invasive *Escherichia coli* (AIEC), which are associated with bacterial dysbiosis in Crohn's disease patients.

Considering the many beneficial roles for bacteriocins in agriculture and medicine further efforts are being made to discover new bacteriocins using bioinformatics (Kemperman et al., 2003; Dirix et al., 2004; de Jong et al., 2006) and classical methods (Martirani et al., 2002; Riley et al., 2012; Balciunas et al., 2013; Mathur et al., 2015; Woraprayote et al., 2015).

1.2 Colicin evolution

In natural isolates, colicinogenic *E. coli* strains make up 10% - 50% of the population, while strains immune to colicins are more abundant, ranging from 50% - 98% (Cascales *et al.*, 2007). It is more costly to produce both the toxin and an immunity mechanism, rather than just the immunity mechanism (Riley, 1993). Populations of sensitive cells are usually small when toxin producers are present and gaining immunity/resistance but losing the toxin production mechanism seems most advantageous for a natural population (Riley, 1993; Gordon *et al.*, 1998), but without the toxin the plasmid has less of a selective advantage.

Colicin evolution has mainly followed two paths, creating either pore-forming or enzymatic colicins (Cascales *et al.*, 2007). Enzymatic colicins are composed of nucleases, Colicin D and Colicins E2-E9 (Wallis *et al.*, 1992; Giffard *et al.*, 1997; Kuhlmann *et al.*, 2000; Kleanthous, 2010; Papadakos *et al.*, 2011; Housden and Kleanthous, 2012), and Colicin M, a phosphodiesterase which inhibits murein synthesis by degrading peptidoglycan lipid intermediate II (Braun, 1974; Braun *et al.*, 1974; Braun *et al.*, 2002; Barneoud-Arnoulet *et al.*, 2010; Barreteau *et al.*, 2010; Barreteau *et al.*, 2012). To date, the known pore-forming colicins are ColA, ColB,

ColE1, ColIa, ColIb, ColK, ColN, ColU, ColV, Col5 and Col10 (Cascales *et al.*, 2007; Kim *et al.*, 2014).

Colicin relationships and evolution have been analysed by comparing DNA and protein sequences, identifying DNA sequence polymorphisms in natural isolates as well as by experimental evolution and mathematical modelling (Cascales *et al.*, 2007). It is proposed that to diversify, pore-forming colicins have mainly relied on recombination and domain shuffling, while nuclease colicins have mainly evolved through single base mutations. Within their group, colicins generally share a very high sequence similarity, e.g. 50% to 97% DNA sequence similarity in nuclease colicins (Cascales *et al.*, 2007), with the greatest differences lying in the immunity region. This suggests a two-step evolution model. First mutations in the immunity protein occur, broadening immunity and providing a survival advantage. Then mutations in the colicin gene change the killing mechanism, producing a new kind of colicin. Further mutations create a group of closely related colicins and immunity proteins (Riley, 1993). The creation of new pore-forming colicins by recombination depends on their existing abundance because it requires a certain variety and concentration of templates to occur together in the same cell. Enzymatic colicins appear to evolve by point mutation whereas homologous recombination appears to be more common in pore-formers (Riley and Wertz, 2002b).

1.3 Colicin operons

Colicins are mostly coded for on plasmids allowing for gene exchange and genome flexibility within a population. In close proximity to the gene for the each colicin are usually the genes for the cognate immunity protein and the lysis gene, which is responsible for the release of colicins from their producer cells. For nuclease colicins, all three genes are arranged together in an operon regulated by the same promoter (Cascales *et al.*, 2007). Colicin M and its immunity protein are coded for on the same plasmid but in reverse orientation (Olschlager and Braun, 1987a). To date no lysis protein has been identified for ColM (Thumm *et al.*, 1988). For pore-forming colicins, the colicin and the lysis gene are arranged on one strand in the same direction, while the immunity gene runs on the other strand in the opposite direction and is regulated by a separate promoter.

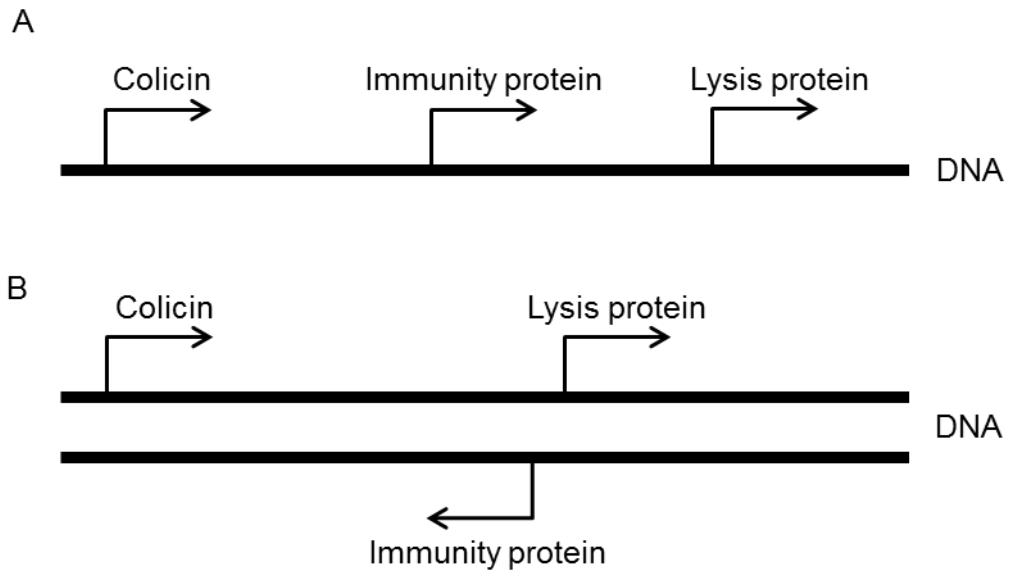


Figure 1: Relative gene locations for colicins, their immunity proteins and lysis genes for (A) Nuclease and (B) Pore-forming colicins. Arrows indicate relative transcription start positions and orientations. The immunity proteins to nuclease colicins are transcribed immediately behind the colicin. The immunity proteins to pore-forming colicins are transcribed from an independent promoter. Diagram adapted from Cascales et al., 2007.

1.4 Colicin binding and translocation

In order to kill, colicins must reach their target environment. ColM and pore-forming colicins translocate through the periplasmic space of target cells until they reach the inner membrane. ColM is active in the periplasm just above the cytoplasmic membrane (Barreteau *et al.*, 2012), while pore-forming colicins form voltage-gated ion channels in the inner membrane to kill cells (Schein *et al.*, 1978; Tokuda and Konisky, 1978b; Weiss and Luria, 1978). Nuclease colicins cross the inner membrane to exert their activity on RNAs and DNA in the cytosol (Papadakos *et al.*, 2011). Colicins are dependent on particular receptor proteins in the outer membrane and particular translocation systems in the periplasm (Kleanthous, 2010; de Zamaroczy and Chauleau, 2011; Kim *et al.*, 2014). Colicin receptors are usually porins or porin-like β-barrelled outer membrane proteins which normally regulate ion traffic, e.g. OmpF, BtuB, Cir, FepA, Tsx and FhuA. Based on the translocation system they use, colicins can be divided into two groups. Group A, dependent on the Tol system, comprises ColA, Col E1-9, ColK, ColN and Col10 (uses TolC and TonB), and Group B is composed of ColB, ColIa, ColIb, ColM, ColU, ColV and Col5 and is dependent on the Ton system (Figure 2).

Figure 2 includes the main components of the Tol (consisting of Pal (peptidoglycan associated lipoprotein), TolB, TolA, TolQ and TolR) and the Ton (consisting of TonB, ExbB and ExbD) systems (Kleanthous, 2010). The key components of the two systems are TonB and TolA, which span the periplasm and are coupled to the proton motive force of the inner membrane through their inner-membrane interaction partners ExbB and ExbD, and TolQ and TolR, respectively. The TolQ–TolR–TolA complex is the functional unit of the Tol system, while TolB is the receptor which engages this unit for potential ligands. Pal could be considered an off-switch for the Tol system. The Ton system passes on energy for proton motive force-driven nutrients uptake through outer-membrane transporters such BtuB and FepA (Kleanthous, 2010). The Tol-Pal proteins are, partly due to their elongated structure, assumed to play an important role in connecting and stabilising the inner and outer membrane of Gram negative bacteria, in particular when they are recruited to the septum during cell division.

Together with the outer membrane receptors, such as outer membrane porins or LPS, both translocation systems are essential for colicin transport through the periplasmic space as cells become resistant to colicins when the required translocation proteins are deleted. However, the cell membranes also become more sensitive to detergents like SDS and possibly other kinds of stress (Ridley and Lakey, 2015) and it is suggested that the outer membrane becomes more permeable. Through their interaction with the Tol-Pal and the Ton-Exb systems colicins and bacteriophages use the energy provided by the proton motive force for translocation (Kleanthous, 2010). Colicin N is the smallest known pore-forming colicin (Pugsley, 1987a; Pugsley, 1988; Wilmsen *et al.*, 1990; Vetter *et al.*, 1998) and is special in the way that it uses LPS and only one porin, OmpF, as a receptor (Jakes, 2014; Johnson *et al.*, 2014). Like other A-type colicins, ColN uses the Tol system for translocation through the periplasm but it uses a different TolA binding site than ColA (Penfold *et al.*, 2012; Ridley and Lakey, 2015). Colicin N is the only one so far whose disordered translocation domain is bacteriocidal independently from the rest of the protein (Johnson *et al.*, 2013).

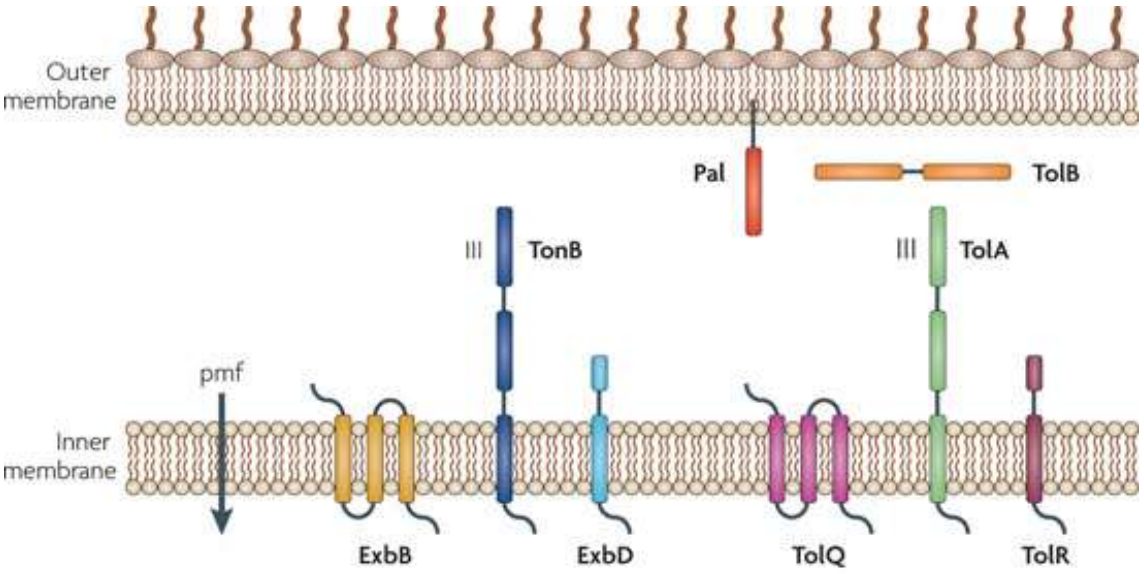


Figure 2: Outline showing the components of the Tol (right: Pal, TolB, TolA, TolQ and TolR) and the Ton (left: TonB, ExbB and ExbD) systems taken from Kleanthous (2010). Both systems depend on the proton motive force (pmf).

1.5 Colicin structure

All colicins share a common general primary structure composed of three domains, the translocation domain (T-domain) at the N-terminus, the receptor binding domain (R-domain) in the middle and the cytotoxic domain (C-domain) at the C-terminus (Figure 3). Each domain varies in size, sequence and structure for each colicin but the greatest sequence similarity is often found in the cytotoxic domain within each type of colicin (Parker *et al.*, 1992).

In nuclease proteins, the cytotoxic domain targets DNA and different kinds of RNAs, interfering with gene expression and protein synthesis (Osborne *et al.*, 1996; Carr *et al.*, 2000; Graille *et al.*, 2004; Papadakos *et al.*, 2011). In ColM, the enzymatic domain degrades the peptidoglycan lipid intermediates I and II, thereby inhibiting peptidoglycan synthesis (El Ghachi *et al.*, 2006). In pore-forming colicins the cytotoxic domain is the pore-forming domain (P-domain), which kills by permeabilising the energised inner membrane (Guihard *et al.*, 1993).

Figure 3: General colicin structure composed of the translocation domain at the N-terminus, the receptor binding domain in the middle and the cytotoxic domain at the C-terminus.

1.5.1 Pore – forming domains

Pore-forming colicins are further divided into A-Type, with the type protein ColA (Lloubes *et al.*, 1986; Espesset *et al.*, 1994; Nardi *et al.*, 2001b; Padmavathi and Steinhoff, 2008; Honigmann *et al.*, 2012; Penfold *et al.*, 2012), and E1-Type, with the type protein ColE1 based on the structure of their P-domain (Cramer *et al.*, 1992; Zakharov *et al.*, 1998a; Zakharov *et al.*, 1998c; Zakharov and Cramer, 1999; Griko *et al.*, 2001; Erukova *et al.*, 2004; Sobko *et al.*, 2004b; Sobko *et al.*, 2005; Sobko *et al.*, 2006a; Sobko *et al.*, 2006b; Ho and Merrill, 2009; Smajs *et al.*, 2010; Ho *et al.*, 2011; Ho and Merrill, 2011; Prieto and Lazaridis, 2011; Ho *et al.*, 2013).

P-domain structures have been solved for A-Type ColA (Parker *et al.*, 1989) and ColN (Vetter *et al.*, 1998) and E1-Type ColIa (Wiener *et al.*, 1997), ColE1 (Elkins *et al.*, 1997) and ColB (Hilsenbeck *et al.*, 2004). They are very similar in sequence (Figure 5) and structure (Figure 6), leading to the assumption that their mechanism of action is also similar. P-domains consist of 10 α-helices, arranged so that two hydrophobic helices form a hydrophobic core surrounded by 8 amphipathic helices. This arrangement makes the P-domain water soluble until it reaches the inner membrane of the target cell, where it inserts and converts to a membrane protein (Lakey *et al.*, 1991; Vandergoot *et al.*, 1991; Lakey *et al.*, 1993). Schein *et al.* (1978) showed that purified, aqueous colicin can bind pure lipid membranes and form voltage dependent channels without any additional target cell proteins. Interestingly, a similar architecture has been observed in apoptosis regulator proteins such as Bcl and Bax (Olschlager and Braun, 1987b; Muchmore *et al.*, 1996; Aisenbrey *et al.*, 2007; Aisenbrey *et al.*, 2008; Nedelkina *et al.*, 2008) (Figure 4). This is intriguing since they operate on the membranes of mitochondria which are evolutionary related to Gram-negative bacteria.

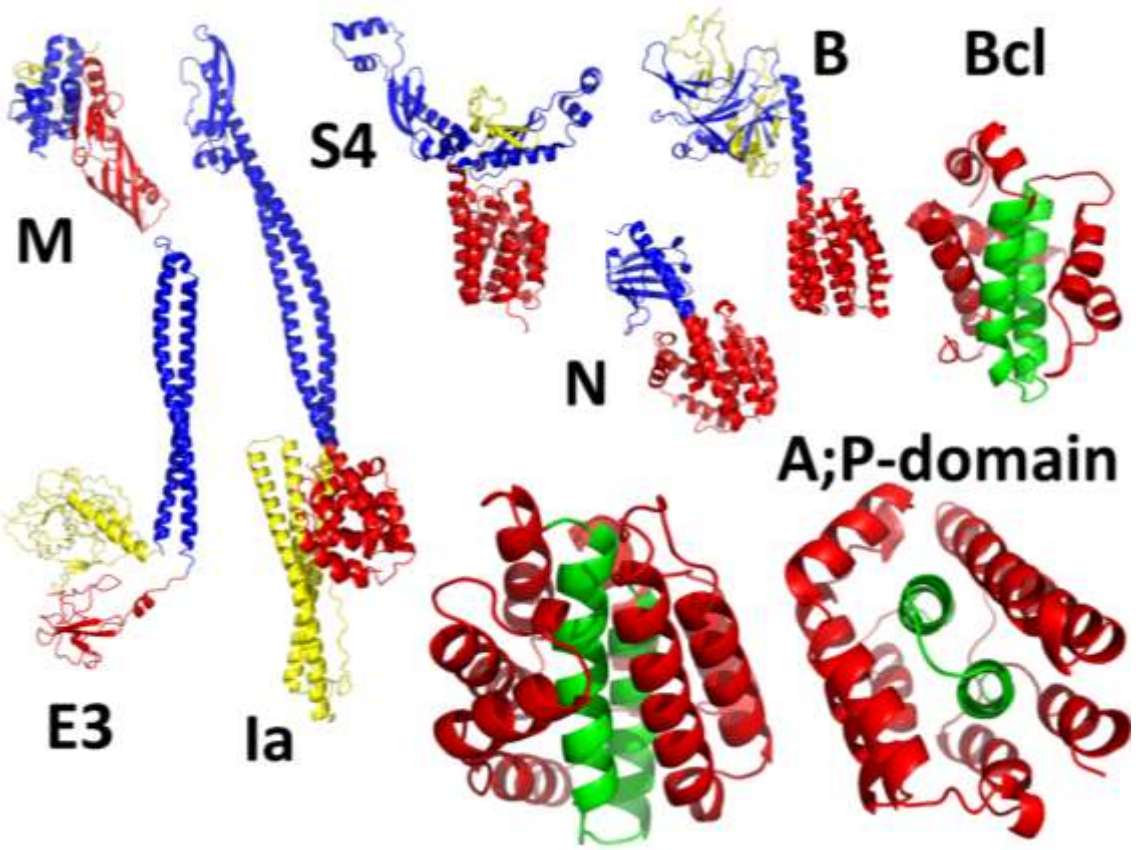


Figure 4: Pore-forming colicins and related structures. (PDB references in parenthesis). For the pore-forming colicins Ia (1CII) 7 (Wiener *et al.*, 1997), N (1A87) (Vetter *et al.*, 1998), B (1RH1) (Hilsenbeck *et al.*, 2004) & S4 (3FEW) (Arnold *et al.*, 2009), the helical pore-forming domain is shown in red, the receptor binding domain in blue and the translocation domain (which may be incomplete due to disorder e.g. colicin N) in yellow. The colicins E3 (3EIP) (Soelaiman *et al.*, 2001) and M (3DA4) (Zeth *et al.*, 2008) which target RNA and peptidoglycan respectively are shown with the same colour scheme revealing the dissimilar toxic domains. The first pore-forming domain to be defined, Colicin A (1COL) (Parker *et al.*, 1989), is shown enlarged in two orientations with the hydrophobic helical hairpin in green. Finally, the structurally homologous Bcl-XL (1LXL) (Muchmore *et al.*, 1996) apoptosis regulator protein is shown with the same colour scheme as colicin A.

```
S4              SMNRDRIQSDVLNKAAEVISDIGNKVGDYLGDAYKSLAREIADDVKNFQGKTIRSYDDAM 60
ColA            SMNRDRIQSDVLNKTAEVISDIGNKVGDYLGDAYKSLAREIADDVKNFQGKTIRSYDDAM 60
ColU            -----NDEKAVLTKASEIIISVGDKAGEYLGDKYKVLSREIADNIKNFQGKTIRSYDEAM 55
ColY            -----NDEKAVLTKASEIIISVGDKVGEYLGDKYKALSREIAGNIKNFQGKTIRSYDEAI 55
ColB            KKEQYNDEKAVLTKTSEVIINVGDKAGEYLGDKYRTLSREIAENIRNFQGKTIRSYDEAM 60
ColN            ----------ALLKASELVSGMGDKLGEYLGVKYKNLAKEVANDIKNFHGRNIRSYNEAM 50
                           * *  *     * * * ***  *  *  * *    ** *  ****  *

S4              ASLNKVLSNPGFKFNRADSDALANVWRSIDAQDMANKLGNISKAFKFADVVMKVEKVREK 120
ColA            ASLNKVLSNPGFKFNRADSDALANVWRSIDAQDMANKLGNISKAFKFADVVMKVEKVREK 120
ColU            ASVNKLMANPDLKINAADRDAIVNAWKAFDAEDMGNKFAALGKTFKAADYVMKANNVREK 115
ColY            ASVNKLMANPDLKINAADRDVIVNAWKAFDAEDMGNKFAALGKTFKAADYVMKANNVREK 115
ColB            SSINKLIENPNLKINATDKEAIVNAWKAFDAKDMGNKFASLGKTFKAADYAIKANNIREK 120
ColN            ASLNKVLANPKMKVNKSDKDAIVNAWKQVNAKDMANKIGNLGKAFKVADLAIKVEKIREK 110
                 * **   **  * *  *     * *    * ** **     * ** **   *    ***

S4              SIEGYETGNWGPLMLEVESWVLSGIASAVALGVFSATLGAYALSLGAPAIAVGIVGILLA 180
ColA            SIEGYETGNWGPLMLEVESWVLSGIASAVALGVFSATLGAYALSLGAPAIAVGIVGILLA 180
ColU            SIEGYQTGNWGPLMLEIESWVLSGIASAVALSFFSAIFGTFAMLGVFSTSLAGILAVILA 175
ColY            SIEGYQTGNWGPLMREVESWVVSGIASAVALAIFSATLGAYLLAVGASAAVVGIIGIIIA 175
ColB            SIEGYQTGNWGPLMLEIESWVLSGIASTVALGLFSTIAGSALLAVGTPPVVVGIMGVFVA 180
ColN            SIEGYNTGNWGPLLLEVESWIIGGVVAGVAISLFGAVL-SFLPISGLAVTALGVIGIMTI 169
                ***** *******  * ***   *    **   *                 *

S4              AVVGALLDDKFADALNKEI 199
ColA            AVVGALIDDKFADALNNEI 199
ColU            GLVGALIDDNFVDKLNNEI 194
ColY            SFIGALIDDKFIDRLNNEI 194
ColB            AVVGVLIDDKFADVLNNEI 199
ColN            SYLSSFIDANRVSNINNII 188
                       *       *  *
```

Figure 5: Amino acid alignment of A-Type Colicin P-domains. Similarities are indicated with a star. The hydrophobic hair-pin, including helices 8 and 9 and the connecting turn (underlined) are highlighted yellow.

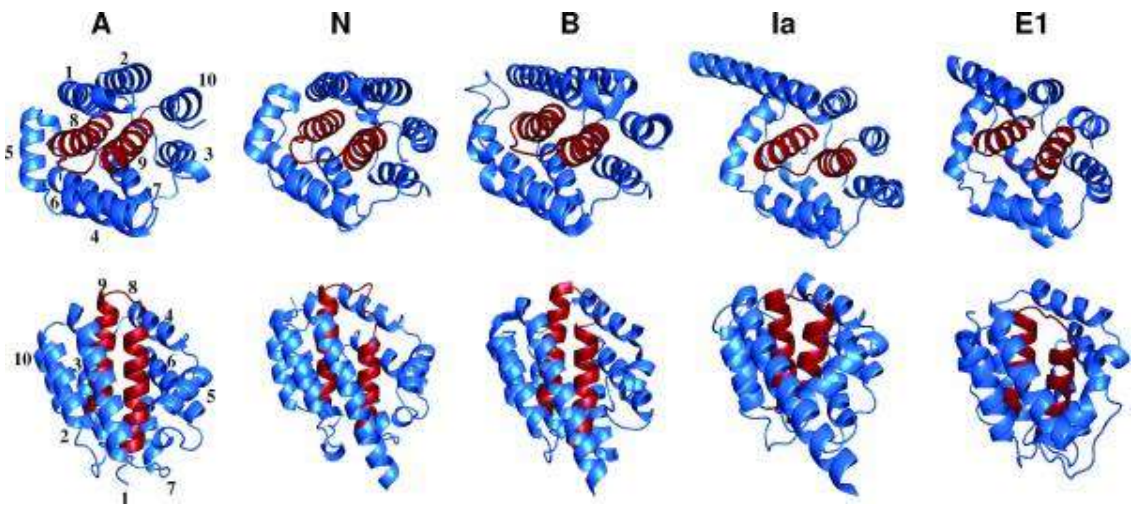


Figure 6: P-domains of colicins A, N, B, Ia and E1 taken from Cascales *et al.* (2007). The P-domains are structurally very similar, showing hydrophobic helices (red) surrounded by 8 amphipathic helices (blue).

How the P-domain inserts into the inner membrane and forms a voltage dependent channel is not entirely clear but several models have been proposed. The umbrella model suggests that the hydrophobic helices 8 and 9 (modelled on ColA and coloured red in Figure 6, highlighted yellow

3 in Figure 5) insert into the inner membrane while the remaining helices spread out on top of the inner membrane, like spokes of an umbrella (Parker *et al.*, 1989). A different model suggests that only helices 1 and 2 move away from the hydrophobic core, while the rest of the molecule sinks into the membrane, including helices 8 and 9 (Lakey *et al.*, 1993; Duche *et al.*, 1994). FRET and quenching experiments with E1 (Zakharov *et al.*, 1998a; Zakharov *et al.*, 1998b; Zakharov *et al.*, 1998c; Lindeberg *et al.*, 2000) have proposed a modification of the umbrella model. Supposedly, the molecule rearranges and unfolds before inserting helices 8 and 9 into the inner membrane, while the amphipathic helices remain on the membrane surface in a spiral-like orientation. The models agree that the structure must be flexible to allow insertion of helices 8 and 9 and channel formation (Zakharov and Cramer, 2002) and that this is a general model for protein-membrane interactions (Vandergoot *et al.*, 1991). Colicin A and B seem to interact with the inner membrane via a pH sensitive molten-globule state, losing the tertiary structure but maintaining the secondary structure (Vandergoot *et al.*, 1991; Evans *et al.*, 1996a). This form of acidic unfolding does not occur in colicin E1 (Schendel and Cramer, 1994), and Colicin N (Evans *et al.*, 1996a).

Despite discrepancies between the insertion models, the interactions of the hydrophobic hairpin with the lipid tails of the inner membrane remain undisputed and are supported by *in vivo* experiments on E1 (Song *et al.*, 1991) and *in vitro* experiments on ColIa (Kienker *et al.*, 1997). The interhelical turn between helix 8 and 9 of ColIa is exposed to the cytosol in the channel's open state and in at least some closed states.

Several studies using domain swapping or deletion have attempted to determine the P-domain parts essential for channel formation. Helices 1-3 of ColE1 and ColA can be removed without losing activity and some parts of helices 4 and 5 might not be crucial (Collarini *et al.*, 1987; Baty *et al.*, 1990; Lindeberg *et al.*, 2000; Nardi *et al.*, 2001b). However, some truncated P-domains form channels with different properties. For ColIa and A, helices 2-5 were determined as unnecessary (Qiu *et al.*, 1996; Slatin *et al.*, 2004).

1.6 The Colicin Immunity Proteins

Cells carrying colicinogenic plasmids need an effective and efficient protection mechanism against their own toxin. This protection is provided by the highly specific immunity proteins which interact with the cytotoxic domain. Similarly to colicins, immunity proteins can be separated into three groups: immunity proteins to nuclease, peptidoglycan synthesis inhibitor and pore-forming colicins (Figure 7). Immunity proteins to pore-forming colicins can be separated into two types based on sequence homology and predicted structure. E1-type immunity proteins (E1, 5, K, 10, Ia, Ib) have three transmembrane helices, while A-type immunity proteins (A, B, N, U, Y, S4) have four. For E1-type colicin immunity proteins, the N-termini are in the cytoplasm while the C-termini are in the periplasm. For A-type colicin immunity proteins both termini are in the cytoplasm. Figure 7 shows an example of suggested topography in the *E. coli* inner membrane using Colicin 5 and Colicin A immunity proteins as models for E1-type and A-type immunity proteins.

Immunity proteins of nuclease colicins are small proteins of ca. 10 kDa coded directly after them on the plasmid (Figure 1) and so are regulated by the same promoter and can attach to them as soon as the colicins are produced (Cascales *et al.*, 2007). Enzymes are usually inactivated through an inhibitor binding to the active site or catalytic site. Nuclease colicins are unusual because, while some are inactivated by the immunity protein binding to the substrate binding site, e.g tRNase-specific immunity proteins like Im5 and ImD (Graille *et al.*, 2004; Luna-Chavez *et al.*, 2006), others, e.g. DNase colicins E2, E7, E9 or rRNase colicins E3, E4, E6 (Kleanthous *et al.*, 1999; Kleanthous and Walker, 2001; Cheng *et al.*, 2002), are inactivated by the immunity protein binding to an exosite, which lies upstream and adjacent to the active site. Exosite inhibitors can inactivate a protein by modifying the active site allosterically or through steric or electrostatic hindrance. Kleanthous and Walker (2001) suggest that the binding of nucleic acid to the enzyme active site is prevented through electrostatic repulsion by a negatively charged immunity protein as well as steric hindrance which does not allow a large molecule such as DNA or RNA to bind. Some evidence suggests that the immunity proteins attach to the colicins adjacent and upstream of the substrate binding site because this part of the colicin leaves the ribosome first, so the immunity proteins are able to inactivate the colicins before they have even left the ribosome completely. Free immunity proteins in the cytoplasm protect the cell against incoming nuclease colicins. Immunity proteins to nuclease

colicins protect the colicinogenic cells against cognate colicins with binding affinities in the femtomolar range but detach at the cell surface shortly after the colicin binds to its high affinity outer membrane receptor (Vankemmelbeke *et al.*, 2009; Farrance *et al.*, 2013).

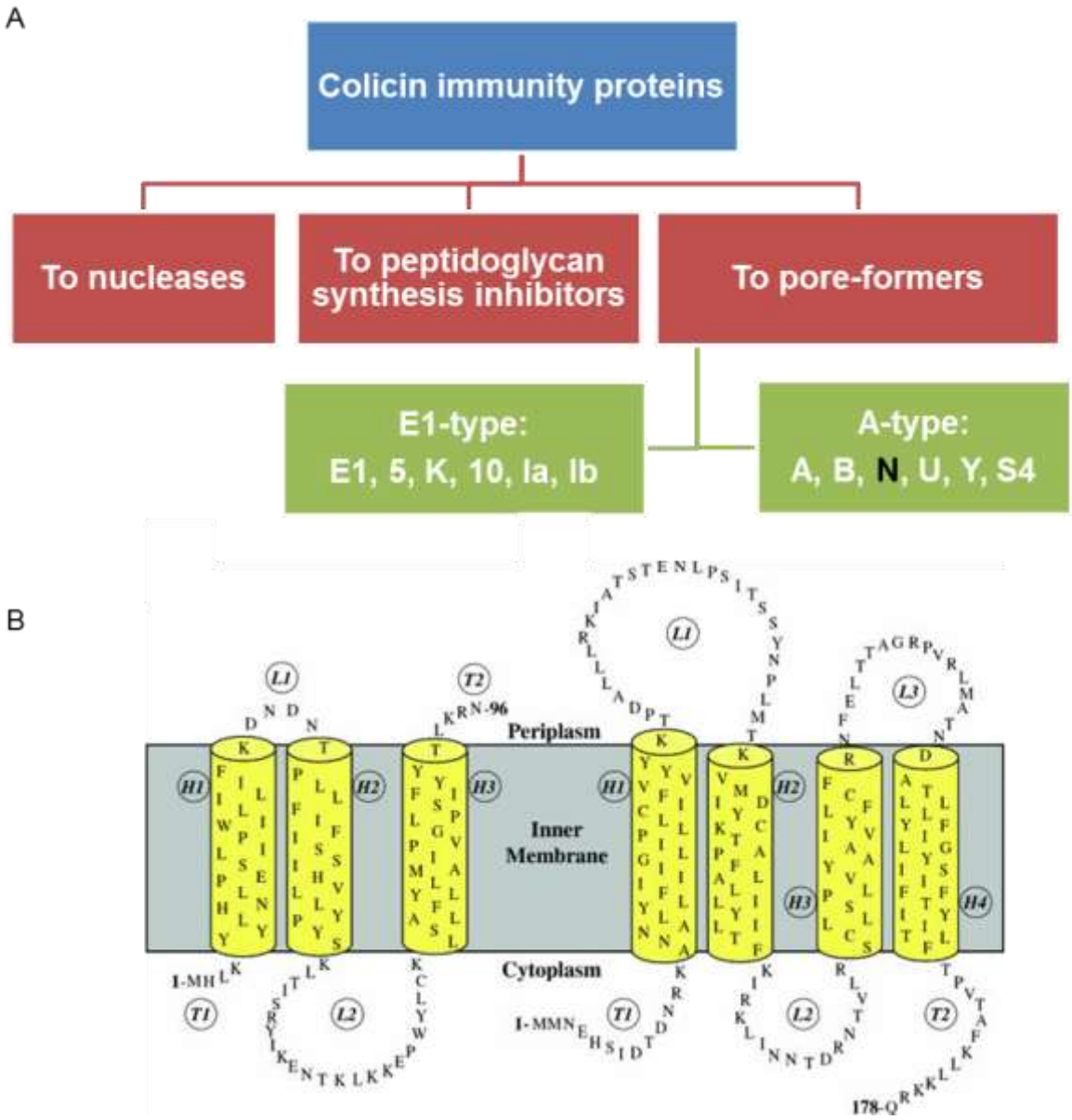


Figure 7: Colicin immunity proteins can are separated into three groups based on the function of their cognate colicins. (A) There are immunity proteins to nuclease, peptidoglycan synthesis inhibitor and pore-forming colicins. Immunity proteins to pore-forming colicins can be separated into two types based on sequence homology and predicted structure. E1-type immunity proteins (E1, 5, K, 10, Ia, Ib) have 3 transmembrane domains, while A-type immunity proteins (A, B, N, U, Y, S4) have 4. (B) Topographical models of Colicin 5 (left) and Colicin A (right) immunity proteins as examples for likely E1-type and A-type immunity protein insertion into the *Escherichia coli* inner membrane. Loops (L), termini (T) and helices (H) are numbered. Diagramm reprinted from Cascales *et al.*, 2007.

Colicin M is not released together with its immunity protein, CMI, but Colicin M immunity protein was found anchored to the cytoplasmic membrane in the periplasm, which is the site of ColM action (Olschlager and Braun, 1987a). Gerard *et al.* (2011) published an X-ray structure of CMI showing that it is composed of 4 β-strands and 4 α-helices and also determined important residues involved in CMI stability and activity using site directed mutagenesis.

1.6.1 Immunity proteins to pore forming colicins

The immunity proteins of pore-forming proteins are located in the inner membrane (Goldman *et al.*, 1985; Geli *et al.*, 1988) and allow cells to live at a colicin concentration of 10^4 to 10^7 times greater than is needed to kill sensitive cells (Cascales *et al.*, 2007). They are regulated by their own promoter and, based on evidence for colicin A immunity protein (CAI) (Lloubes *et al.*, 1986), are expressed constitutively. A low number of immunity proteins to pore-forming colicins seems to protect cells against high numbers of exogenous colicins (Song and Cramer, 1991). Immunity proteins may also interact with inner membrane components of the translocation machinery (Figure 2) (Song and Cramer, 1991; Fridd *et al.*, 2002). For CNI, these may be TolA, Q and R (Lloubes *et al.*, 2001). However, in studies where T- and R-domains have been swapped between colicins depending on different translocation systems in order to send the P-domain through a different translocation system, the immunity protein cognate to the P-domain still neutralised killing (Zhang and Cramer, 1993). This indicates that the translocation system is not interacting specifically with the immunity protein. TolR/Q are homologs of ExbB and ExbD of the Ton system and so it can be assumed that the mechanism is similar (Figure 2).

Based on sequence similarity, immunity proteins to pore forming colicins can be divided into A-type (A, B, N and U) and E1-type (E1, 5, K, 10, Ia and Ib), which correlates with the two groups of P-domains and their differently sized hydrophobic helices. The A–type is predicted to have 4 with both ends in the cytoplasm (Geli *et al.*, 1989a) (Pilsl *et al.*, 1998) and the E1-type to have 3 transmembrane domains (Song and Cramer, 1991); the N – terminus is located in the cytoplasm and the C-terminus in the periplasm (Song and Cramer, 1991; Pilsl and Braun, 1995). The topology of CNI is unknown largely because the published sequence (Pugsley, 1988) has an error which alters the amino acid sequence and truncates the protein (see

Chapter 3 for details). The full length sequence predicts 4 transmembrane sequences, expected of A-type immunity proteins.

Domain swapping experiments in Colicins Ia and Ib (Mankovich *et al.*, 1986; Benedetti *et al.*, 1991a), as well as A and E1 (Bishop *et al.*, 1985), have shown that the immunity proteins interact exclusively with the P-domain and neutralise its lethal action through an, as yet, unknown mechanism. Espesset *et al.* (1994) have shown that cytoplasmically expressed P-domain can kill when directed from the cytoplasm to the inner membrane but can be inhibited by an immunity protein. Allegedly, lateral diffusion is enough to ensure rapid neutralisation of the pore forming domain (Geli and Lazdunski, 1992; Zhang and Cramer, 1993).

Although the P- domains have a great sequence and structural (Figure 5 and Figure 6, Parker *et al.* (1992)) similarity, the immunity proteins are specific to their cognate colicin. This specificity of immunity proteins to their cognate colicins has been used in sequence swapping (Geli and Lazdunski, 1992; Pilsl and Braun, 1995; Lindeberg and Cramer, 2001) and mutagenesis studies (Zhang and Cramer, 1993) attempting to determine the P-domain parts involved in interacting with the immunity protein. This has been difficult because the P-domain's structural orientation during insertion and channel formation has not been shown beyond doubt. Also, specificity is not strictly confined to the cognate colicin as there is some evidence of cross-immunity between closely related colicins. For example, the Col10 immunity protein provides some resistance against ColE1, but the ColE1 immunity protein provides no measurable resistance against Col10 (Lindeberg and Cramer, 2001). The immunity regions are located in areas of high sequence divergence in P-domains which would support the two step evolutionary model (Figure 5). This is supported by a gradual shift in immunity specificity from one colicin to the other, if specificity determining residues are exchanged one at a time.

The immunity proteins to ColE1 and ColA have been best studied and can be considered case studies for their type of immunity protein.

1.6.2 E1-type colicin immunity proteins

Zhang and Cramer (1993) mapped the main areas of interaction between ColE1 and Colicin E1 immunity protein to residues V441-W460 and A471-A488, with residues Ala474, Ser477, His440, Phe443 and Gly444 of ColE1 being key. Mutants in these

showed reduced lethality against sensitive cells but were able to kill immune cells. Sobko *et al.* (2009) identified His440 as the residue responsible for channel formation. The five identified residues are located in helices 6 and 7 which are believed to be on top of the inner membrane when the channel is in a closed state but their conformation may change once the channel is open (Vandergoot *et al.*, 1991; Evans *et al.*, 1996a; Lakey and Slatin, 2001; Sobko *et al.*, 2009). Colicin E1 immunity protein is, therefore, likely to interact with ColE1 in an open state (Zhang and Cramer, 1993). Furthermore, the authors disprove a previous assumption (Song and Cramer, 1991) that charged domains of the immunity protein, i.e. the termini and cytosolic and periplasmic loops, are involved in interaction with ColE1. Instead, they suggest that the interaction may be mediated by salt bridges, hydrogen bonds or specific packing and interactions of non-polar residues between the hydrophobic helices of both molecules.

Lindeberg and Cramer (2001) confirmed residues 419-501 on ColE1 as an important region for interaction. However, this study focused on finding specificity determinants which make Colicin E1 immunity protein specific to its colicin rather than residues crucial for its activity. It suggests that 448, 470, 472 and 474 are the specificity determinants. Unsurprisingly, they lie in the same region as the residues suggested by Zhang and Cramer (1993), but interestingly, they are different ones.

1.6.3 A-type colicin immunity proteins

For A-type colicins, the hydrophobic helical hairpin of the P-domain has been identified as the region of interaction with the immunity protein (Espesset *et al.*, 1996; Pilsl *et al.*, 1998; Nardi *et al.*, 2001a). For interaction to occur, only this region needs to be inserted into the membrane (Espesset *et al.*, 1996; Nardi *et al.*, 2001a) and, in contrast to the E-type colicins, the opened state is not required.

Zhang *et al.* (2010) give the most recent description of Colicin A immunity protein (CAI) structure and interaction with ColA. They show that CAI exists as a dimer *in vivo,* which is formed through disulphide bonds between helix 3 of each molecule. CAI has 4 cysteine residues. It is proposed that the other 3 residues form intramolecular bonds between either helix 1 and 2 or helix 2 and 3 to stabilise the molecule. All bonds are broken when ColA is added, either through external full-length ColA or co-expression of the P-domain. Furthermore, Zhang *et al.* (2010) confirms that channel activity is not required for interaction, because an inactive P-

domain, which can insert the hydrophobic helical hairpin into the inner membrane but cannot form a channel, also breaks the inter- and intramolecular disulphide bonds. This indicates that the immunity protein binds to ColA while the channel is in a closed state and might prevent opening.

Smajs *et al.* (2006) investigated the CYI interaction site with CYA using domain swapping, followed by more precise site-directed mutagenesis of CUI residues into CYI residues in combination with a crude spot test assay. Colicin U is inactivated by CUI and CYI, but Colicin Y is only inactivated by CYI, not CUI. They obtained mutants of CUI that could inactivate both ColU and ColY, constituting gain of function mutations. The amino acids responsible for interaction with ColY are S104, S107, F110, A112 on helix 3 and A159 on helix 4 of CYI and are located close to the cytoplasmic membrane where they could interact with helix 8 and the hydrophobic hairpin. Interestingly, at least three mutations, in residues C104S, T107S & I159A are needed to confer immunity to CYI, while single mutations cause no change in specificity. The fact that the mutant CUI (CS104S, TS107S, IF110F, AA112V, IA159A) is able to inactivate both ColY and ColU suggests that conserved residues are responsible for binding while variable residues define specificity. This is similar to the observations made by for ColE1 (Zhang and Cramer, 1993; Lindeberg and Cramer, 2001).

Subsequently, Smajs *et al.* (2008) mapped the ColY residues which interact with CYI by using ColY/ColU hybrids, ColY with point mutations and a crude spot test assay. ColY residues I578, T582 and Y586 located in helix 8 and V590 located in the hydrophobic hairpin were important for the interaction. This was validated by mutating the corresponding residues of ColB to the residues of ColY and so establishing immunity to this ColB mutant in a strain expressing CYI.

Using the CYI mutants Smajs *et al.* (2006) in combination with ColY mutated in the above residues, Smajs *et al.* (2008) were able to identify the possible interaction pairs between ColY and CYI as I578 – S107, T582 – S107, T582 – S104 and Y586 – S104, which means that these residues are located on helix 8 and the hydrophobic hairpin of ColY interacts with residues on helix 3 and 4 of CYI. These pairings and a simulation model based on lowest energy structures and conformations led to the conclusion that helix-helix interactions play a crucial role in the inactivation of ColY by CYI. Due to the high level of amino acid and structure conservation in these residues

across A-type colicins, it is proposed that the general interaction mechanism between the colicin and its cognate immunity protein is also conserved.

The studies of CAI and Colicin E1 immunity protein together with the investigation of channel formation (Sobko *et al.*, 2009; Sobko *et al.*, 2010) have led to the proposal of two different models of immunity protein action (Figure 8).

E1-type immunity proteins are suggested to interact with colicin molecules when the channel is in an open state conformation because the assumed structural arrangement of both proteins in the membrane does not support interaction when the channel is in a closed state. A-type immunity proteins interact with their colicins while the channel is in a closed state. Both models of interaction are heavily based on the assumed structure of colicin molecules in the membrane and may only be validated when experiments showing the colicin structure beyond reasonable doubt have been conducted.

All successful studies of immunity proteins to pore-forming colicins have so far been *in vivo*. *In vitro* data is very sparse due to the hydrophobicity of immunity proteins and the fact that they are integral membrane proteins. Even where overexpression and purification have been successful (Mankovich *et al.*, 1986; Geli *et al.*, 1989b; Shirabe *et al.*, 1993; Taylor *et al.*, 2000), a demonstration of activity or even interaction with purified colicins has been unsuccessful (Geli *et al.*, 1989b).

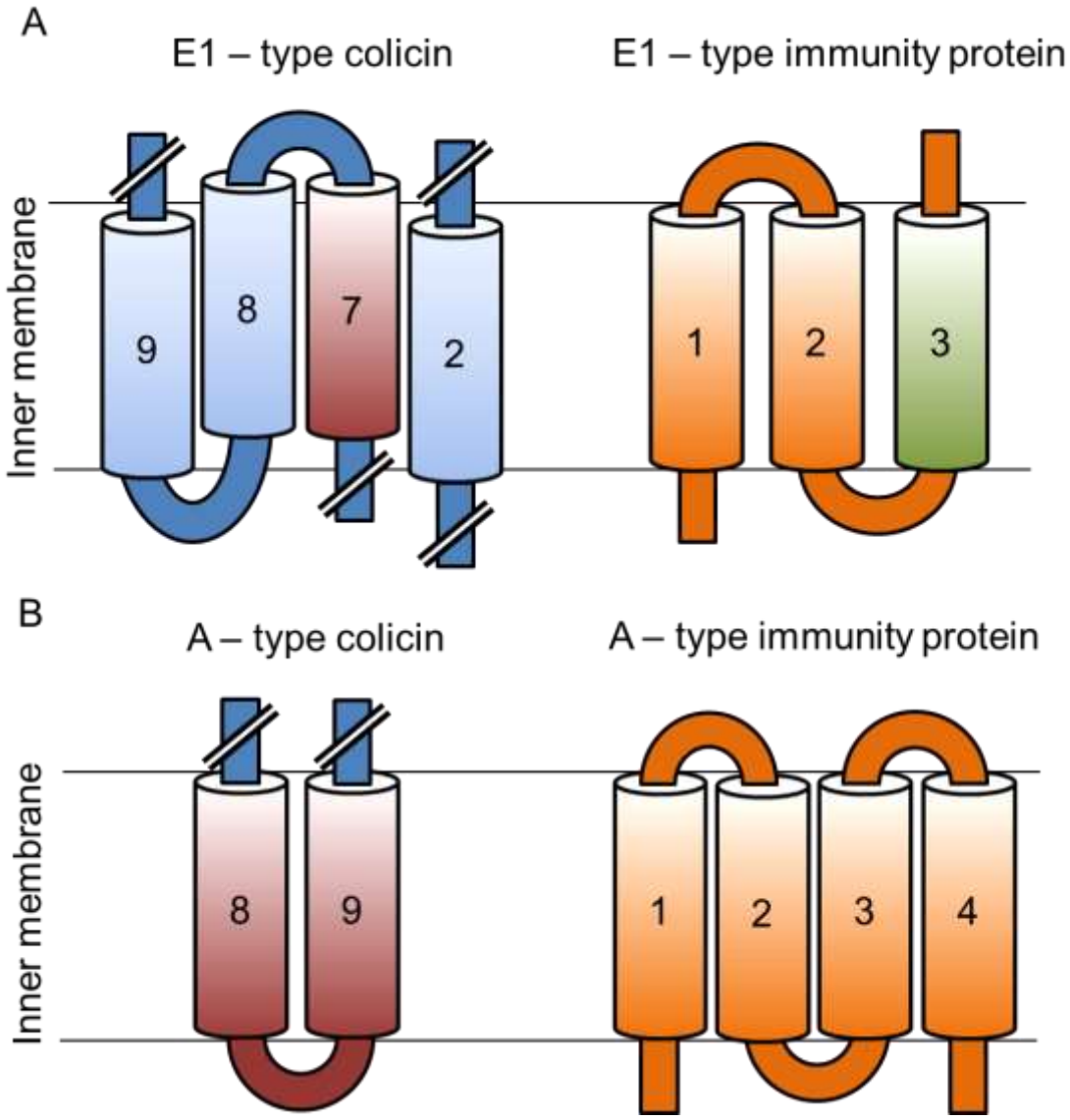


Figure 8: Models representing transmembrane helices of E1 - and A - type colicins and their immunity proteins based on Zhang and Cramer (1993); Bohme *et al.* (2009); Sobko *et al.* (2009); Zhang *et al.* (2010); Honigmann *et al.* (2012). Cytosolic and periplasmic domains have been omitted for the colicins. (A) Helix 7 (red) of E1 - type colicins probably interacts with helix 3 (green) of the E1 – type immunity protein, when the channel is in an open state. (B) Helices 8 and 9 of A-type colicins and in some cases also the connecting turn between these two helices interact with A - type immunity proteins, when the channel is in a closed state. The specific site of interaction on the immunity protein is unknown.

1.7 Project Aims:

To the best of my knowledge, the only attempt to characterise the immunity protein to Colicin N is mentioned by Pugsley (1988) but it has since remained unstudied. The aim of this project was to characterise the Colicin N immunity protein (CNI), *in vivo* and *in vitro* with a broad range of techniques in case one way would prove more successful than another. Although it is known that immunity proteins neutralise colicins by interacting with them, it is unclear how the colicins are inactivated with such remarkable speed and efficiency considering that the number of immunity protein molecules per cell seems to be low (Cascales *et al*., 2007).

This work aimed to develop and optimise an overexpression and purification procedure for CNI in preparation for biophysical and structural studies, determine CNI location in the inner membrane and whether Colicin N attack changes it and find the protein- protein interaction site with ColN on CNI using chimeric proteins with CNI's closest homologue Colicin A immunity protein (CAI).

Large parts of this introduction, including Figure 5 and Figure 8, are part of Stroukova and Lakey, (2015, in press; see appendix).

Chapter 2. Materials and Methods

All chemicals were purchased from Melford Laboratories Ltd unless otherwise stated.

2.1 Molecular Biology

2.1.1 Gene synthesis and sequencing

The full CNI gene sequence was obtained by re-sequencing the plasmid pCHAP4 (Pugsley, 1988) at GATC Biotech Ltd., The London BioScience Innovation Centre, 2 Royal College Street, London, UK. The new sequence was used to synthesise the constructs *Para*-6His-CNI, *Para*-CNI-6His and *Para*-CNI-GFP-6His construct in a pBAD322 vector at GeneArt® (Life technologies Ltd., 3 Fountain Drive, Inchinnan Business Park, Paisley PA4 9RF, UK). *Para* is an arabinose inducible promoter (Cronan, 2006).

2.1.2 Vectors

Table 1 summarises all vectors used in this work and their relevant genotypes. Vectors based on pBAD322 and pET3a confer resistance to 100 µg/ml ampicillin. The pWALDO vector is based on pET28 (+) and confers resistance to 50 µg/ml kanamycin (Drew *et al.*, 2006). The pOPIN CD vectors confer resistance to 50 µg/ml spectinomycin. Enhanced GFP is abbreviated eGFP (Patterson *et al.*, 1997) and superfolder GFP is abbreviated sfGFP (Pedelacq *et al.*, 2006). Gene expression from pBAD322 can be regulated with arabinose and glucose (Guzman *et al.*, 1995; Cronan, 2006). The pBAD322 copy number is temperature sensitive (Cronan, 2006). All cultures using this plasmid were grown at 37 °C. Expression from plasmids containing the T7 promoter was induced with 1 mM IPTG. P_{CNI} is the natural promoter from the natural pCHAP4 vector.

Arabinose and IPTG can both be used to regulate the expression of genes from an operon. IPTG can be used to activate the *lac* operon. When no IPTG is present the *lac* operon repressor lacI binds to the operator and prevents the binding of the RNA polymerase to the operator region. The gene is not transcribed and the protein of interest is not expressed. When IPTG is added it binds allosterically to the repressor lacI and it can no longer bind to the operator and block transcription, so the gene is translated and the protein is expressed. The arabinose operon is regulated in a similar way. In the absence of arabinose, a dimer of the regulator protein AraC binds

DNA in a loop so that RNA polymerase cannot bind to the DNA. When arabinose is present, it binds the repressor AraC and prevents its interaction with DNA so RNA polymerase can bind and transcribe the protein of interest. Both operons also need high cAMP levels for protein expression. When cAMP levels are reduced, for example, through the addition of high levels of glucose, gene expression is also repressed. A feature of the lac operon is the "leaky" promoter, meaning that a low level of transcription still takes place even without the inducers IPTG or lactose. This can be problematic if the protein to be expressed is already toxic in low levels or its expression needs to be regulated very tightly for a different reason, e.g. even a small amount of protein has a big effect on an *in vivo* phenotype as is the case for the colicin immunity proteins. In these cases, using the arabinose regulated operon is advantageous because it can be much more tightly regulated by using arabinose as an inducer or by using glucose as a catabolite inhibitor.

Table 1: Plasmids used in this thesis.

Vector name	Relevant genotype	Source
pBAD322	*bla rop araC Para*	Cronan (2006)
pBAD322-CNI-sfGFP	*bla Para*	Lakey group
pBAD322-CNI-FLAG-TCS-sfGFP	*bla Para,* thrombin cleavage site	This work
pBAD-P_{cni}-CNI-GFP		This work
pET3a-ColN	*bla T7*	Lakey group
pISA-ColA	*bla T7*	Gokce and Lakey (2003)
pCHAP4	natural plasmid	Pugsley (1988)
pET3a-TolA-GFP		Lakey group
pBAD-WALP23-GFP		Strahl (Newcastle Uni), Nyholm *et al.* (2007)
pWALDO-CNI-TEV-eGFP	*kanR T7*	Based on pET28(+), Drew *et al.* (2006)
pOPINCDE	*specR T7*	Oxford Protein Production vector suite https://www.oppf.rc-harwell.ac.uk/OPPF/protocols/cloning.jsp These vectors are available through NCBI under the same name.
pOPINCDF	*specR T7*	
pOPINS3C	*bla T7*	
pOPINMSYB	*bla T7*	
pOPINCDJ	*specR T7*	
pOPINE-3C-eGFP	*bla T7*	
pOPINE-3C-HALO	*bla T7*	
pOPINCDM	*specR T7*	
pOPINO	*bla T7*	
pOPINP	*bla T7*	
pOPINDsbA	*bla T7*	
pOPINMalE	*bla T7*	
pMHT238Δ – SuperTEV - 6His	*kanR*	MPL, Blommel and Fox (2007)

2.1.3 Polymerase chain reactions for site-directed mutagenesis and in preparation for InFusion cloning.

All PCRs were carried out using the Biorad T100 Thermal cycler. KOD HotStart DNA Polymerase (Merck MilliPore) was used for all PCRs according to Merck MilliPore instructions. Each reaction contained 1x reaction buffer, 3 µl $MgSO_4$, 150 µM dNTPs, 1.5 µl 10 pM forward primer, 1.5 µl 10 pM reverse primer, 20 ng template, 1 µl KOD HotStart Polymerase and nuclease free water up to 50 µl total reaction volume. The PCRs consisted of an initial 2 min at 95 °C, followed by 20 cycles of: 30 sec at 95 °C, 10 sec at 55 °C, 30 sec per kb (for vectors), 10 sec per kb (for inserts) 70 °C elongating and a final elongation step of 5 min at 70 °C. Sizes of amplification products and primers are detailed in Table 2 and Table 3.

After site-directed mutagenesis where nucleotides were changed for the cysteine113 to serine mutation or added for the thrombin cleavage site or FLAG-tag, constructs were treated with 1 µl DpnI (Promega) for 1 h at 37 °C and transformed into TOP 10 cells using the heat shock protocol (2.2.3, Table 2).

After construct amplification for InFusion cloning (ClonTech, 2.1.4, Table 3), PCR products were sized with gel electrophoresis in 1% agarose gels in TBE (For 1l of 10x TBE: 108 g Tris, pH = 8, 55 g boric acid, 40 ml of 0.5 M ETDA), cut out from the gel and purified with a Qiagen PCR clean up kit.

For the insertion of the amplified CAI gene into pOPINE-3C-HALO, the vector was linearised with restriction enzymes NcoI and PmeI in CutSmart® buffer (NEB) at 37°C for 20 min. Then it was purified like the amplification products with agarose gel electrophoresis and a Qiagen PCR clean up kit.

Table 2: Site directed mutagenesis: modifications and primer sequences.

Modification	construct	Modification sequence	Primers
Thrombin cleavage site insertion to give pBAD322-CNI-TCS-sfGFP (plasmid size ca. 6.3 kb)	pBAD322-CNI-sfGFP	Leu-Val-Pro-Arg-Gly-Ser **CTG GTT CCG CGT GGC TCC**	Thr_F: 5'-GAATACTCCGTCGACAC**CTGGTTCCGCGTGGCTCC**ATGAGCAAAGGTGAAG -3' Thr_R: 5'-CTTCACCTTTGCTCAT**GGAGCCACGCGGAACCAG**GTGTCGACGGAGTATTC -3'
Flag-tag insertion to give pBAD322-CNI-FLAG-TCS-sfGFP (plasmid size ca. 6.3 kb)	pBAD322-CNI-TCS-sfGFP	Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys **gat tac aaa gat gat gat gat aaa**	FlagF: 5'-ggaatactccgtcgacac**gattacaaagatgatgatgataaa**CTGGTTCCGCGTGGC -3' (upper case letters: beginning of thrombin cleavage site) FlagR: 5' -GCCACGCGGAACCAG**tttatcatcatcatctttgtaatc**gtgtcgacggagtattcc – 3' (upper case letters: beginning of thrombin cleavage site)
C113S mutagenesis (plasmid size ca. 6.6 kb)	pWALDO-CNI-TEV-eGFP	Cys113 to Ser **tgt** to **tca**	C113S_F: 5'- ccatgatatac**tca**ttcttgctaactgac – 3' C113S_R: 5'- gtcagttagcaagaa**tga**gtatatcatgg – 3'

Table 3: Amplified inserts for InFusion cloning: insert sizes and primer sequences.

Amplified construct	source	size	Primers
Helix 1 CAI Helix 2 CAI Helix 3 CAI Helix 4 CAI	CAI from pISAColA	H1: 162 bp H2: 114 bp H3: 150 bp H4: 129 bp	Halo_H1_fwd: 5' – caaaggagatataccATGATGAATGAACAC -3' (lower case is part of vector back bones) H1_rev: 5' – TGCGATGATTTTTTCAAGTAAAGCGTCTGG – 3' H2_fwd: 5'–TTTCATTCATTAAACAACCCATTAATGACA-3' H2_rev: 5' – TTTTAACTCTAATTCTTTAAAGGTTAGGAT – 3' H3_fwd: 5' –ATATTCACACACAAAATCAGAAAATTAATC -3' H3_rev: 5' – TGTTTTTGATGACAATGTTAACTCAAAATT – 3' H4_fwd: 5' –GTATTAATGTCAAAAAATGACGCAACACTA-3' H4_rev: 5' – TCCTCCCCTGGTAAATAATTTAAATGCAGT –3' (upper case is *cai* gene sequence)
CNI-HALO7-6His vector linearisation for insertion of Helix 1 CAI Helix 2 CAI Helix 3 CAI Helix 4 CAI	CNI-HALO7-6His	6.4 kb	H1vFwd: 5' – GAAAAAATCATCGCATACCTATCC – 3' Halo_H1vRev: 5'- ggtatatctcctttgattg – 3' (lower case is part of vector back bones) H2vFwd: 5' – GAATTAGAGTTAAAACCAAAGTCG – 3' H2vRev: 5'- GTTTAATGAATGAAATCCTGGTAG – 3' H3vFwd: 5' – TTGTCATCAAAAACATTTGTATTA – 3' H3vRev: 5'- TTTGTGTGTGAATATGAATAAGCT – 3' H4vFwd: 5' –TTTACCAGGGGAGGAATACTCCGT – 3' H4vRev: 5'- TTTTGACATTAATACAAATGTTTT – 3' (upper case is *cni* gene sequence)
Natural promoter of CNI (P_{cni})	pCHAP4	230 bp	<u>Insert amplification:</u> CNIproFwd: 5' – CTGATTCGTTACCAAGGATCCTTTGATCTT – 3' CNIproRev_corrected: 5' – GTCTTTTATATCCATGTTGCTTTCTTTGGATGG – 3' <u>Vector linearisation:</u> pBAD322_rev: 5' – TTGGTAACGAATCAGACAATTGACG - 3' pBAD322_fwd: 5' – ATGGATATAAAAGACAGAAATAAGATATCA – 3'
CAI gene for cloning into pOPINE-3C-HALO-6His	pISAColA	537 bp, 20.5 kDa	Halo_H1_fwd: 5' – caaaggagatataccATGATGAATGAACAC -3' (lower case is part of vector back bones) CAIrev: 5'-cagaacttccagtttCTGCCTTTTTTTTAATAATTTAAATGC-3' (lower case is start of 3C protease cleavage site)

2.1.4 In-Fusion® HD cloning

In-Fusion® HD cloning is a commercial cloning technology (Clontech® Laboratories, Inc.) which allows seamless cloning without the introduction of additional nucleotides into plasmids and was particularly useful here in replacing the helices of CNI with homologous helices from the Colicin A immunity protein (CAI). The basic principle involves the amplification of the insert with overhangs that match the destination site in the host plasmid and a linearisation of the host vector by restriction enzymes or PCR amplification. Amplified plasmid and linearised vector are mixed with the kit reagent and incubated for 15 min at 50 °C before 5 µl are used to transform competent cells using the heat shock method (2.2.3).

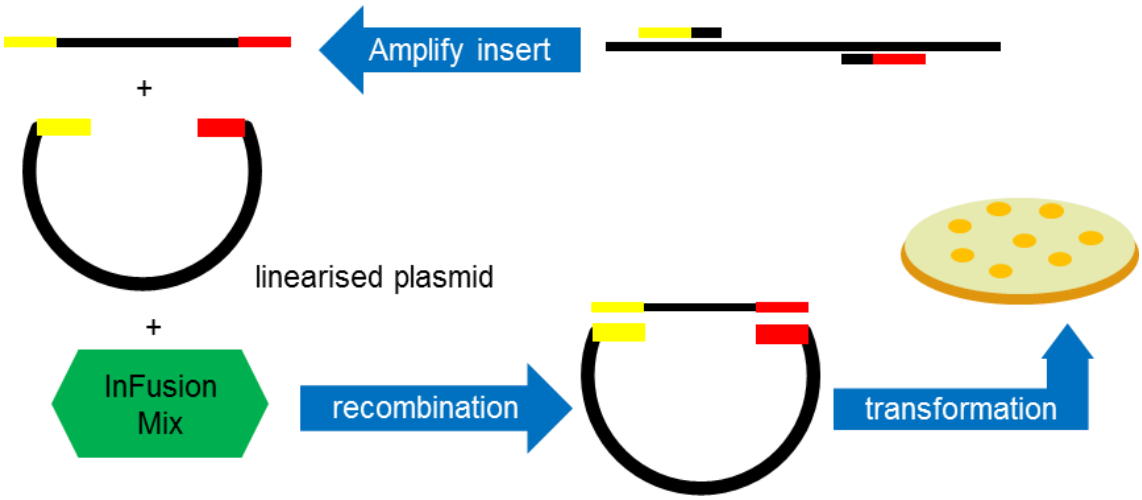


Figure 9: In-Fusion® HD cloning procedure. The amplified insert with overhangs that match the destination site in the host plasmid and a linearised host vector are mixed with the kit reagent and incubated for 15 min at 50 °C and transformed into competent cells. The InFusion mix is supplied by Clontech® Laboratories, Inc.

2.2 Microbiology

2.2.1 Bacterial strains

The following *E. coli* strains were used for plasmid amplification, protein overexpression and fluorescent microscopy.

TOP 10 One Shot

TOP 10 (Invitrogen) cells were used for amplification of all plasmids used. The genotype is F$^-$ *mcrA Δ(mrr*$^-$*hsdRMS*$^-$*mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara, leu)7697 galU galK rpsL (StrR) endA1 nupG.*

BL21-AI™ One Shot

BL21-AI™ (Invitrogen) have the genotype F$^-$ *omp*T *hsdSB* ($r_B^-m_B^-$) *gal dcm ara*B::T7RNAP-*tetA.* They were used for overexpression of ColN, ColA, TEV protease, N-terminally and C-terminally polyhistidine-tagged CNI (6His-CNI and CNI-6His), CNI-TCS-sfGFP-6His. BL21-AI are BL21 cells which do not utilise arabinose as a nutrient and therefore gene expression can be induced using arabinose.

Lemo21, C41, C43 and Rosetta™2(DE3) strains

Lemo21, C41, C43 and Rosetta™2(DE3) *E. coli* strains were used in a protein production screen to determine the strain yielding the most CNI.

Lemo21 cells (New England Biolabs) are BL21 (DE3) cells containing the Lemo System™, which is maintained by chloramphenicol selection. The genotype is *fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS/ pLemo(*Cam^R*)* *λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5* pLemo = pACYC184-*PrhaBAD-lysY*

The *E. coli* strain C41 (Cambridge Bioscience Limited) was originally characterised by Miroux and Walker (1996) as particularly suitable to produce toxic proteins (Dumon-Seignovert *et al.*, 2004). It is derived from BL21 (DE3). The genotype is F$^-$ *omp*T *gal dcm* hsdSB(rB$^-$ mB$^-$)(DE3).

The strain C43(DE3) (Cambridge bioscience) was derived from C41(DE3) (Miroux and Walker, 1996) by selecting for resistance to a different toxic protein and can express a different set of toxic proteins to C41(DE3). The genotype is also F$^-$ *omp*T

gal dcm hsdSB(rB⁻ mB⁻) (DE3). C41(DE3) and C43(DE3) can tolerate expression of toxic proteins because the mutations in the T7 polymerase slow down protein production.

Rosetta™2(DE3) cells (Novagen) are genotype F⁻*ompT hsdSB(rB⁻mB⁻) gal dcm* (DE3) pRARE2 (Cam^R). Rosetta™2(DE3) cells are derived from BL21 cells but grow slower which can improve membrane protein production.

MC1000 for microscopy

E.coli MC1000 cells (Casadaban and Cohen, 1980) were used for epifluorescent and TIRF microscopy as well as single molecule tracking.

The genotype is *Δ(araA,leu)7697 [araD139]B/r Δ(codB-lacI)3 galK16 galE15(GalS) λ-e14- relA1 rpsL150*(strR) *spoT1 spoT1.* This strain cannot utilise arabinose as a nutrient, which is useful when used in combination with arabinose inducible expression systems such as pBAD322.

JC207

The *E. coli* strain JC207 is a TolA knock out strain (ΔTolA) and was used here for overexpression of TolA-GFP in preparation for fluorescent microscopy. JC207 has the genotype metB, lacI, lky-207 (Lazzaroni *et al.*, 1986; Ridley and Lakey, 2015).

2.2.2 Growth media

Lysogeny broth (LB)

Unless otherwise stated, all cultures were grown in LB medium (Melford Ltd.). LB contains 10 g/l casein digest peptone, 5 g/l yeast extract, 10 g/l NaCl and was prepared as per manufacturer's instructions. To prepare LB agar plates 2% agar were added to LB before autoclaving.

Minimal media (M9)

Minimal media was used when we wished to detect fluorescence in liquid cultures or to grow cells for fluorescent microscopy because it has a lower background fluorescence in comparison to coloured media like LB and TB. M9 minimal media was prepared in house. Components: 1X M9 salts (24 mM Na_2HPO_4, 10 mM KH_2PO_4, 4.3 mM NaCl, 9.4 mM NH_4Cl), 1 mM thiamine hydrochloride, 0.4 % glycerol, 0.2 % casamino acids, 2 mM $MgSO_4$, 0.1mM $CaCl_2$.

Terrific Broth

Terrific broth was used for overexpression of CNI fusion proteins. To prepare, 12 g tryptone, 24 g yeast extract (powder, Melford Ltd.) and 4 ml glycerol were mixed with 900 ml deionised water and autoclaved. Separately, a 10X potassium phosphate solution (0.17 M KH_2PO_4 and 0.72 M K_2HPO_4, pH = 7) was prepared and autoclaved. The sterile 10X potassium phosphate solution was mixed with the tryptone-yeast-glycerol solution in a 1:9 ratio.

2.2.3 Escherichia coli heat shock transformation

50 ng of plasmid DNA were added to 50 µl chemically competent cells and chilled on ice for 30 min, then heat-shocked at 42 °C for 45 sec and cooled for 2 min on ice. 250 µl of LB were added to the cells and they were incubated for recovery for 1 h at 37 °C, gently shaking. Cells were spread on LB agar plates containing the appropriate antibiotic for transformant selection. Plates were incubated overnight at 37 °C. Selector antibiotics were used at 100 µg/ml ampicillin, 100 µg/ml carbenicillin, 30 µg/ml kanamycin, 30 µg/ml spectinomycin, 35 µg/ml chloramphenicol. When two selector antibiotics were used, the given concentration of both was halved.

2.2.4 Para-CNI-sfGFP-6His expression toxicity assay

The effect of arabinose addition on culture viability was monitored using a plate reader (FLUOstar OPTIMA, BMG Labtech) in Costar 96 well plates. Overnight cultures of BL21-AI cells transformed with the empty vector or pBAD322-CNI-sfGFP-6His were used to inoculate 150 µl media in each well to a final cell count of ca. $4x10^5$. The media was composed of LB and 100 µg/ml ampicillin and arabinose concentrations of 0 %, 0.02 %, 0.05 %, 0.1 %, 0.2 %, 0.5 %, 1 % and 2 % were added at the beginning of culture growth

2.2.5 Assaying CNI activity by determining the minimal inhibitory concentration (MIC) of ColN to sensitive and immune cells

CNI activity was assessed through its ability to immunise otherwise sensitive cells against ColN, allowing them to grow at ColN concentrations sensitive cells cannot. The MIC of ColN was assayed using a plate reader (FLUOstar OPTIMA, BMG Labtech). Cells were grown in flasks to an $OD_{600\ nm}$ = 1 and diluted to ca. $4x10^5$ cells

in 150 µl media in each well of a Costar 96 well plate before 1 pM, 10 pM, 0.1 nM, 1 nM, 10 nM, 0.1 µM, 1 µM, 10 µM ColN and no colicin as a control were added. Each concentration was assayed in triplicate. Cultures overexpressing CNI were induced with 0.2 % arabinose from the start of culture growth in flasks as well as in the plate reader. Growth was monitored for 24 h at 37 °C. This procedure was adapted from my Masters of Research project.

2.2.6 Assaying CNI activity by spot test assay

Flask cultures with 50 ml LB and 100 µg/ml Ampicillin were inoculated from an overnight culture. Protein production was not induced but solely relied on the leaky T7 promoter which allows enough gene expression to compare immune and sensitive cells. When the culture $OD_{600\ nm}$ = 0.5, 50 µl from the flask culture were mixed with 5 ml LB and 100 µg/ml Ampicillin and 5 ml sloppy agar (in 100ml: 1 g Tryptone, 0.5 g NaCl, 0.6 g agarose; heated to 60 °C in a water bath). The mixture was layered onto 2 % LB agar plates containing 100 µg/ml Ampicillin and allowed to set and dry in the incubator (37 °C) for 1 h. Colicin N or A were spotted in 2 µl drops onto the set agar and allowed to dry. Plates were incubated at 37 °C overnight. The concentration of Colicin N and A used varied depending on the cell type and protein expressed and are directly indicated in the figures of the Results chapters.

2.2.7 Assaying CNI activity with the cell survival assay in a plate reader

A modification of the cell survival assay used in Cavard and Lazdunski (1981) and Cavard (1994) was used to compare the viable cell count of sensitive and immune cultures following treatment with ColN and ColA. This method is a less time and resource consuming alternative to the traditional colony counting method, where following antibiotic treatment of liquid cultures the remaining live cell count is established through counting colony forming units on agar plates. Viability in this assay is measured through culture tubidity or optical density at 600 nm. The plate ready laser (FLUOstar OPTIMA, BMG Labtech) needs a certain level of turbidity before regarding an optical density as not zero and recording growth. Therefore, the lag time seen in the viability graphs is not the time it takes for growth of viable cells to start again but the time it takes for this growth to cause enough turbidity to be measurable by the plate reader. Therefore, the time it takes for a culture to reach a certain turbidity is dependent on the size of the inoculum, which in turn is determined by how many of a given amount of cells were killed through the colicin addition. Since

Colicin N or A concentrations and optical density of starting cultures are the same, a difference in lag time is directly corelated to how potent the colicin is or how immune the cells are.

From overnight cultures 50 ml LB flask cultures were inoculated in a 1:100 ratio and incubated at 37 °C until $OD_{600\ nm}$ = 0.5, shaking well to allow aeration. Cultures were diluted 1:1 with 10 mM sodium phosphate buffer, 300 mM NaCl, pH = 7.4 and chilled on ice for 30 min. Colicin N or Colicin A was dispensed into a 96 well plate (flat bottom, Costar) and mixed with chilled cells 1:1, resulting in final approximate culture $OD_{600\ nm}$ = 0.125. Colicin N and A concentrations vary depending on the experiment and are indicated in Results for each case. Cells were incubated with Colicin A or N for 30 min at 37 °C, while constantly shaking, then diluted 1:100 into a new 96 well plate (flat bottom, Costar) and incubated for at least 18 h at 37 °C. The $OD_{600\ nm}$ was measured every 10 min by the plate reader FLUOstar OPTIMA (BMG Labtech).

2.2.8 Epifluorescence and Total Internal Reflection Fluorescence Microscopy (TIRF)

The pBAD322 vector (Cronan, 2006) is a low copy number plasmid and has an arabinose inducible promoter (*Para*) which can be used to regulate gene expression depending on the amount of arabinose added and is therefore useful for fluorescent microscopy. Very low levels of expression are often required to avoid saturation of the fluorescence signal. Also, we wished to observe CNI at as close to natural levels as possible

<u>CNI-sfGFP-6His expression titration by simultaneous and delayed metabolite inhibition with glucose or fucose</u>

Glucose and fucose can be used to repress protein expression from the arabinose inducible *Para* promoter. To test how the addition of different glucose and fucose concentrations affects the CNI-sfGFP-6His expression levels, MC1000 cells transformed with pBAD322-*Para*-CNI-sfGFP-6His were grown in 50 ml LB and 100 µg/ml ampicillin flask cultures until $OD_{600\ nm}$ = 0.5 and then aliquoted into a 96 well plate (Costar). CNI-sfGFP-6His expression was induced with 15 mM (13 mM = 0.2 % arabinose) and inhibited at the same time through metabolite inhibition with 30 mM, 15 mM, 1.5 mM, 300 µM, 150 µM, 30 µM, 15 µM, 3 µM or 1.5 µM glucose and fucose. Relative fluorescence intensity was monitored at 510 nm in a FLUOstar plate reader OPTIMA (BMG Labtech) for 24 h at 37 °C.

When CNI-sfGFP-6His expression is simultaneously induced with arabinose and repressed with either glucose or fucose, the protein production is either on or off depending on the amount of repressor added and it is not possible to only reduce the expression. To achieve a reduced expression level, MC1000 cells were grown in 50 ml LB and 100 µg/ml ampicillin flask cultures until $OD_{600\ nm}$ = 0.5. CNI-sfGFP-6HIS expression was then induced with 0.2 % arabinose and repression by 0.2% glucose addition was delayed by 10 sec, 20 sec or 30 sec. Fluorescence intensity was observed with a Nikon Eclipse Ti-E inverted epifluorescence microscope.

Epifluorescent and TIRF microscopy carried out in Newcastle

MC1000 cells were used for epifluorescent microscopy. They were transformed with vectors pBAD322-*Para*-CNI-sfGFP-6His, pBAD322-P_{CNI}-CNI-sfGFP-6His, pET3c-TolA-eGFP-6His, WALP23-eGFP-6His or the empty vector (pBAD322) control. WALP23 belongs to a group of artificial peptides composed of a stretch of tryptophans, alanines, leucines and prolines which readily fold into artificial transmembrane helices and insert into lipid bilayers, such as the *E. coli* inner membrane. This man-made, artificial transmembrane helix has no known biological function and was originally designed to study lipid membrane fluidity and the way a generic transmembrane protein moves inside the lipid bilayer (Nyholm *et al.* 2007). Here, a version of WALP23 fused to eGFP has been used so that it can be observed in vivo using fluorescence microscopy. In this thesis, it represents a generic transmembrane protein which has no interaction with pore-forming colicins, immunity proteins or the Tol-Pal system and which has no specific localisation pattern in the E.coli inner membrane. It serves to determine if any potential effects on protein diffusion rates tested here are due to specific interactions between Colicin N and its immunity protein or if they are general effects on membrane protein movement or bilayer fluidity.

Overnight cultures were diluted 1:1000, and grown in flask culture with M9 minimal media until $OD_{600\ nm}$ = 0.5. Where an arabinose inducible promoter was used, cultures were induced with 0.2 % arabinose for 10 sec, then repressed with 0.2 % glucose. Where the CNI natural promoter (P_{CNI}) was used, cultures were not induced. For the treatment of cells with ColN or ColA, 200 µl aliquots from the flask cultures were taken and a sub-lethal concentration of Colicin N or Colicin A, both at 100 pM, was added for 10 min at 37 °C, while the tube was gently shaking. A drop of 0.5 µl of

each culture was fixed on 1.2 % agarose on microscope slides and observed using a Nikon Eclipse Ti-E inverted epifluorescence microscope. Cells were immobilised using 1.2% agarose by sandwiching the cells between the agarose and a thin coverslip (width: 1 mm; Figure 10). The agarose was prepared by dissolving it in water using a microwave, cooling it to 60 °C and then applying 500 µl to a Teflon coated glass slide (C. A. Hendley (Essex) Ltd.). The agar was flattened using a second glass slide and allowed to cool and solidify. The top glass was removed and 0.5 µl of the cell culture was applied onto agar. The drop was allowed to dry out and a coverslip was added for imaging. This procedure was adapted from my Masters of Research project.

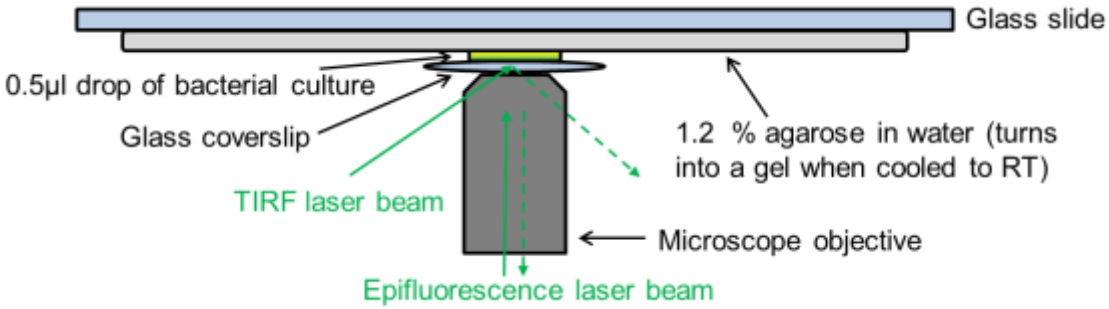


Figure 10: TIRF set up. Cells were immobilised using 1.2% agarose by sandwiching the cells between the agarose and a thin coverslip. The glass slide was inverted and placed onto the objective. The laser beam is totally internally reflected, causing an evanescent wave just above the cover slide which allows viewing of the bacterial membrane in TIRF mode.

Two different measuring techniques were used to answer two different questions. Firstly, the homogeneity of each cell's fluorescence was measured to see if addition of Colicin N would change the otherwise even distribution of CNI-sfGFP-6His into concentrated and possibly localised clusters. Cells were imaged at 30 ms per frame. Using the freely available image processing software Image J[1] (NIH, USA), the fluorescence intensity was measured as a transection across each cell and plotted against the length of the cell.

To determine if ColN addition would change the diffusion rate of CNI, a method called total internal reflection fluorescence with continuous bleaching (TIR-CP) was used (Slade *et al.*, 2009; Strahl *et al.*, 2014), which is similar to fluorescence recovery after photo bleaching (FRAP). The model to quantify the bleaching kinetics is defined in the appendix. Cells were imaged over time at 100 ms per frame, deliberately

[1] http://rsb.info.nih.gov/ij/

strongly bleaching the fluorescent signal. Each recording formed a stack of images over 20 secs. Using the image processing software Image J, all images in the stack were treated for background subtraction and then the overall signal intensity of each image was measured and plotted against time. The fluorescence intensity decays over time as more fluorophores are bleached than can be replenished. The fluorescence intensity decay was expressed in percentage, normalising the initial fluorescence intensity at the start of the measurement as 100 %. This allows the comparison of repeats of the same kind of condition, even if the fields of view do not contain the same amount of cells and do not have the same absolute intensity.

Single Molecule Tracking at CLF, Harwell

Single molecule tracking was carried out at Central Laser Facility, Harwell, UK, in collaboration with the Martin-Fernandes group. The main goal of this collaboration was to establish if the diffusion rate of CNI changes *in vivo* upon ColN addition. Originally, it was intended to co-localise CNI and ColN, so ColN and ColA, as the negative control, were fluorescently labelled (see 2.3.7). The proteins TolA-eGFP-6His, WALP23-eGFP-6His and CNI-sfGFP-6His were expressed in MC1000 cells. Flask cultures with 20 ml M9 minimal media and 100 µg/ml ampicillin were inoculated in a 1:20 ratio with overnight cultures and induced with 1 mM IPTG (for TolA-eGFP-6His) or 0.2 % arabinose (for WALP23-eGFP-6His and CNI-sfGFP-6His) at an $OD_{600\ nm}$ = 0.1. The culture expressing CNI-sfGFP-6His was inhibited with 0.2 % glucose 10 sec after induction. Cultures were then grown to an $OD_{600\ nm}$ = 1. Aliquots (200 µl) were taken one at a time and mixed with ColN or ColA to give final concentrations of 40 nM, 400 nM and 4 µM and incubated with shaking for 10 min at 37 °C, before 0.5 µl drops were fixed on 1.2 % agarose in water as previously in Newcastle (Figure 10) and imaged straight away. Due to the laser light and the nutrient poor environment of the agarose these samples can only be imaged for a limited time and so each sample was prepared separately and imaged immediately.

Imaging conditions were optimised based on the best conditions for CNI-sfGFP-6His and then remained consistent. Laser intensity at 488 nm was 6.5 mW, while it was 4.43 mW at 642 nm. The microscope used was a Zeiss Axio Observer and the lasers used were an Omicron LightHUB equipped with a PHOXX 488 nm continuous wave and a PHOXX 640 nm continuous wave laser. The laser has a circular area of 100 µm. Consecutive images were taken at 60 ms per frame. Each data set needed an

extensive amount of bleaching to reduce background fluorescence and reduce the amount of fluorescent molecules to an extent where single molecules were visible. This was dependent on the protein imaged as all three protein fusions have different characteristics. Image processing and single molecule tracking was carried out using specialist in house software (Rolfe *et al.*, 2011) by Dr Daniel Rolfe, CLF, Harwell, UK. For a summary of the procedure, see Results Chapter 5.

2.3 Biochemistry

2.3.1 Protein purification using immobilised metal affinity chromatography (IMAC)

ColN, ColA and ColN P-domain protein purification using a Nickel ion affinity chromatography

The same purification protocol was used to purify ColN, ColA and ColN P-domain. Plasmids coding for ColN, ColA or ColN P-domain were transformed into BL21-AI cells using the *E. coli* heat shock method (2.2.3). A single colony from a selection plate was used to inoculate a 50 ml LB with 100 µg/ml Ampicillin overnight culture, which was grown at 37 °C, well aerated. Flasks with 500 ml 2XLB and 100 µg/ml Ampicillin were inoculated with overnight cultures in a 1:100 ratio and incubated at 37 °C, shaking until $OD_{600\ nm}$ = 0.7-0.8 and then induced with 2 % arabinose concentration for max 3 h. Cells were harvested by centrifuging at 11300 x g (JA-10 rotor, Beckman Coulter) for 10 min. The supernatant was discarded and the pellets resuspended in a cell disruption buffer which also serves as the column loading buffer (50 mM NaP (a mixture of monosodium and disodium phosphate to achieve the appropriate pH), pH=7.4, 300 mM NaCl, 10 mM imidazole, proteinase inhibitor cocktail (100 mg/ml AEBSF, 0.5 mg/ml aprotinin, 100 mg/ml benzamidine, 1mg/ml Pepstatin A, 1 mg/ml Leupeptin), 100 mg DNAse, 100 mg RNAse, 100 mg lysozyme). Cells were disrupted with a Constant Systems continuous cell disruptor (model: TS2/40/AA/AA, Constant Systems Ltd.) and the unbroken cells as well as the cell membrane were spun down at 140531 g (TI-70 rotor, Beckman Coulter) for 30 min. The pellets were discarded and the supernatant applied to a 2 ml resin self-poured Ni^{+}-NTA column (Qiagen). The column was washed with 20x column volume (≈ 40 ml) wash buffer (50 mM NaP, 300 mM NaCl, 25 mM Imidazole, pH=7.4,). The proteins were eluted with elution buffer (50 mM NaP, 300 mM NaCl, 350 mM imidazole, pH=7.4). After dialysis against a phosphate buffer (50 mM NaP, 300 mM NaCl, pH=7.4), aliquots were frozen at -20°C.

CNI purification

An optimised purification protocol for CNI was one of the key goals of this work and the evidence to support changes in the procedure, e.g. SDS-PAGE gels with purification fractions, is covered in detail in Chapter 3. During the optimisation

process the expression vector, expression cells and buffer composition, including the detergent, varied but the general procedure remained the same. The overall procedure has been adapted from Bird *et al.* (2015) and is given below while the parameters changed are summarised in Table 4 for clarity.

Host cells were transformed with a vector using the heat-shock protocol (2.2.3). A single colony from the selection plates was used to inoculate 50 ml LB with the appropriate antibiotic (depending on vector) and grown overnight at 37 °C, well aerated. Ten flasks with 500 ml of growth medium and the appropriate antibiotic (depending on vector) were inoculated with the overnight culture in a 1:100 ratio and incubated at 37 °C, shaking, until $OD_{600\ nm}$ = 0.8 - 1 and then induced with 0.2 % arabinose concentration or 1 mM IPTG (depending on vector). Initially, cultures were grown overnight at 37 °C and harvested the next morning. For cultures with the pOPIN vectors during and after visiting OPPF, the growth temperature was reduced to 18 °C for growth overnight.

Cells were harvested by centrifuging at 11300 x g for 15 min, resuspended in cell disruption buffer and lysed with a Constant Systems continuous cell disruptor (model: TS2/40/AA/AA, Constant Systems Ltd.). Unbroken cells were spun down at 11300 x g for 15 min and the supernatant was ultra-centrifuged at 140531 x g for 1 h at 4 °C. Pellets were homogenised using a glass homogeniser in chilled column loading buffer containing the appropriate detergents. The homogenised resuspension was slowly agitated on a tube roller mixer (Stuart) overnight at 4 °C to extract the membrane protein. In this thesis, this process is referred to as protein solubilisation. The cell debris was spun down at 140531 x g for 1 h at 4 °C and the protein was purified using a self-poured 2 ml NTA-Ni^{+} column (Novagen), by loading the solubilised protein onto the column, washing the column with the wash buffer and eluting with the elution buffer.

Table 4: Key changes carried out during CNI purification optimisation. While the general purification steps remained the same, the vector construct, host strain, growth media and purification buffer composition, including the detergents, varied.

Vector construct	Host strain	Growth media	Purification buffer	Detergents
6His-CNI	BL21-AI	LB	Cell disruption: 50 mM NaP, 300 mM NaCl, 10 mM imidazole, pH = 7.4. Detergent added for column loading, wash and elution buffer. Imidazole concentration was increased to 25 mM for wash buffer and 300 mM for elution.	0.5% w/v n-dodecyl-β-D-maltoside (DDM), Genapol–X080, Triton-X100 or N-Lauroylsarcosine were used for the entire purification procedure.
CNI-6His	BL21-AI	LB	Cell disruption: 50 mM NaP, 300 mM NaCl, 10 mM imidazole, pH = 7.4. Detergent added for column loading, wash and elution buffer. Imidazole concentration was increased to 25 mM for wash buffer and 300 mM for elution.	0.5 % n-dodecyl-β-D-maltoside (DDM) were used for the entire purification procedure.
pBAD332-CNI-FLAG-TCS-sfGFP-6His	BL21-AI	LB	Cell disruption: 50 mM NaP, 300 mM NaCl, 10 mM imidazole, pH = 7.4. Detergent added for column loading, wash and elution buffer. Imidazole concentration was increased to 25 mM for wash buffer and 300 mM for elution. 476 µM DTT and 2.4 mM $CaCl_2$ were added for thrombin cleavage. Dialyses was carried out against Tris (pH = 7.4), 300 mM NaCl, 10 % glycerol.	2% n-decyl-β-D-maltoside (DM) was used in Newcastle. At MPL 1% DDM, DM, Cymal 6, Triton-X100, LDAO or lauroylsarcosine were used for solubilisation. For purification, 0.09 % Cymal 6 was used (equal to 3x the critical micelle concentration).
pWALDO-CNI-TEVCS-GFP-8His pWALDO-CNI^{C113S}-TEVCS-GFP-8His	C41	LB	Cell disruption: 50 mM NaP, 300 mM NaCl, 10 mM imidazole, 10 % glycerol, pH = 7.4. Detergent added for column loading, wash and elution buffer. Imidazole concentration was increased to 25 mM for wash buffer and 300 mM for elution.	For solubilisation 1 % Cymal 6 and for IMAC 0.09 % Cymal 6 was used.
pOPINCDE, pOPINCDF, pOPINS3C, pOPINMSYB ,pOPINCDJ, pOPINCDM pOPINE-3C-eGFP, pOPINE-3C-HALO	Lemo21, C41, C43, Rosetta™ 2(DE3)	TBO or PB	Cell disruption: 50 mM NaP, 300 mM NaCl, 10 mM imidazole, 10 % glycerol, pH = 7.4. Detergent added for column loading, wash and elution buffer. Imidazole concentration was increased to 25 mM for wash buffer and 300 mM for elution.	1 % DM was used for solubilisation and was reduced to 0.1% for IMAC purification.
pOPINE-3C-HALO (Expression scale-up in Newcastle)	C41	Terrific Broth	Cell disruption: 50 mM NaP, 300 mM NaCl, 10 mM imidazole, pH = 7.4. Detergent added for column loading, wash and elution buffer. Imidazole concentration was increased to 25 mM for wash buffer and 300 mM for elution. The NaP buffer was used for initial scale up but it was decided to switch to a Tris buffer in preparation for 3C cleavage. The final buffer composition is shown in Table 5.	For solubilisation 1% Tween and 1% DM were used. The concentrations were reduced to 0.1 % for IMAC purification. The Tween was removed in the in the final formulation of the Tris buffer.

Table 5: Optimised buffer composition for CNI-Halo7-6His purification from C41 cells, grown in Terrific Broth medium.

Cell disruption buffer	Protein extraction buffer	Column Wash buffer	Protein elution buffer	Dialysis/Desalt buffer
50 mM Tris-HCl, pH = 7.5	50 mM Tris-HCl, pH = 7.5	50 mM Tris-HCl, pH = 7.5	50 mM Tris-HCl, pH = 7	50 mM Tris-HCl, pH = 7
1 M NaCl	1 M NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
50 mM Imidazole	50 mM Imidazole	100 mM Imidazole	350 mM Imidazole	
Proteinase inhibitor (100 mg/ml AEBSF, 0.5 mg/ml aprotinin, 100 mg/ml benzamidine, 1mg/ml Pepstatin A, 1 mg/ml Leupeptin), DNAse, RNAse, lysozyme				
			1 mM EDTA	1 mM EDTA
			1 mM DTT	1 mM DTT
	1 % DM	0.1% DM	0.1% DM	0.1% DM
	10% glycerol	10% glycerol	10% glycerol	10% glycerol

TEV protease purification

The protocol for TEV purification as well as the plasmid were provided by MPL and are an adaptation of the protocol used by Blommel and Fox (2007). The TEV protease version purified here is the SuperTEV and is expressed from the plasmid pMHT238Δ, which provides kanamycin resistance. It was transformed into Rosetta™2(DE3) cells using the heat shock protocol (2.2.3) and grown in LB at 37 °C. When $OD_{600\ nm}$ = 0.5, the growth temperature was reduced to 25 °C for overnight growth and the protein expression was induced with 0.4 mM IPTG.

Cells were harvested by centrifuging at 11300 x g (JA-10 rotor, Beckman Coulter) for 10 min. The pellets resuspended in a cell disruption buffer which also serves as the column loading buffer (1 x PBS, 150 mM NaCl, 50 mM imidazole, protease inhibitor (100 mg/ml AEBSF, 0.5 mg/ml aprotinin, 100 mg/ml benzamidine, 1mg/ml pepstatin A, 1 mg/ml leupeptin, DNase I, RNase, 0.3 mM TCEP, pH = 7.5). Cells were disrupted with a Constant Systems continous cell disruptor (model: TS2/40/AA/AA, Constant Systems Ltd.) and the unbroken cells as well as the cell membrane were spun down at 140 531 g (TI-70 rotor, Beckman Coulter) for 30 min. The pellets were

discarded and the supernatant applied to a 5 ml self-poured Ni^{+}-NTA column (Novagen). The column was washed with 20x column volume (≈ 40 ml) wash buffer (20 mM Tris pH = 7.5, 300 mM NaCl, 50 mM imidazole, 30 % glycerol, 0.3 mM TCEP, pH = 7.4). The proteins were eluted with elution buffer (20 mM Tris pH =7.5, 300 mM NaCl, 250 mM imidazole, 30 % glycerol, 0.3 mM TCEP). After dialysis against a Tris buffer (20 mM Tris pH = 7.5, 300 mM NaCl, 50 % glycerol, 0.3 mM TCEP), the protein concentration was measured using the Nanodrop 1000 (ThermoScientific) (ε_{280} = 32 770 l mol-1 cm-1, MW = 28.62 kDa) and 1 mg/ml aliquots were flash-frozen and stored at -80 °C. SuperTEV as well as TEV tends to precipitate heavily, therefore, large amounts of glycerol are required and the protease needs centrifuging immediately before use.

2.3.2 Sodium dodecyl polyacrylamide gel electrophoresis (SDS-PAGE)

For SDS-PAGE, samples were mixed 1:1 with 2X SDS loading buffer (0.125 M TrisHCl, pH = 6.8, 15 % glycerol, 5 mM NaEDTA. 2 % SDS, 0.1 % bromophenol blue, 1 % β-mercaptoethanol) and loaded onto 10 %, 12 % or 15 % acrylamide SDS-PAGE gels. The running buffer was 25 mM Tris, 192 mM glycine, 0.1 % SDS. Electrophoresis as carried out for 1 h at 200 V. Gels were stained with Coomassie stain (10 % acetic acid, 10 % propan-2-ol, Coomassie Blue) for 1 h and de-stained (solution: 10 % acetic acid, 10 % propan-2-ol) overnight, slowly agitated. In-gel fluorescence images were taken with the Typhoon™ FLA 9500 biomolecular imager (GE Healthcare).

Table 6: Composition of Tris gels.

	10 %	**12 %**	**15 %**
Resolving gel	**9 ml**		
1.5 M Tris, 14 mM SDS (pH=6.8) (ml)	2.25		
40 % acrylamide/bis- acrylamide (ml)	2.25	2.7	3.38
diWater (ml)	4.43	3.98	3.3
10 % APS (µl)	50		
TEMED (µl)	20		
Stacking gel (4%)	**4 ml**		
1.5 M Tris, 14 mM SDS (pH=6.8) (ml)	1		
40 % acrylamide/bis- acrylamide (ml)	350		
diWater (ml)	2.59		
10 % APS (µl)	50		
TEMED (µl)	10		

2.3.3 Western blotting

Western blotting was performed using a wet blotting system (BioRad) for 2 h at 350 mA using a Tris-glycine transfer buffer (20 mM Tris base, 150 mM glycine, 0.5 % SDS, 20 % isopropanol) onto a nitrocellulose membrane. Following overnight blocking with 5 % skimmed milk (dried power resuspended in 20 mM Tris-HCl, 140 mM NaCl, 0.1% Tween-20, pH = 7.5 (TBST)), membranes were incubated for 2 h with 1 % milk in TBST and the anti-polyhistidine alkaline phosphatase conjugated antibody (Life Technologies, Anti-His (C-term)/AP Ab). Blots washed with TBST 3 x 5 min and were revealed using SIGMAFAST™ BCIP®/NBT as substrate for the conjugated alkaline phosphatase.

2.3.4 Size exclusion chromatography

Size exclusion chromatography was carried out using a Superose 12 10/300 GL (GE Healthcare) column in 50 mM NaP pH = 7.4, 300 mM NaCl, 10 mM imidazole, 0.1 % Tween20 and 0.1 % DM using an Äkta Purifier system.

2.3.5 Thrombin digest

A thrombin digest was carried out after CNI-FLAG-TCS-sfGFP-6His purification according to the manufacturer's instructions. 1U of thrombin (Novogen, restriction grade) will cleave 1 mg of protein in optimal conditions. Since detergent was added to the protein sample which could inhibit thrombin, 4U were used to cleave 200 µg of CNI-FLAG-TCS-sfGFP-6His. Thrombin was added directly to the dialysed protein and incubated at 20 °C overnight. In preparation for SDS-PAGE, 1% acetic acid (final concentration) was added to a sample of the cleaved and uncleaved protein to denature the protein structure.

2.3.6 Protease cleavage with 3C (PreScission Protease) cleavage

A commercially available version of the viral protease 3C (PreScission Protease, GE Healthcare) was used to cleave the CNI-HALO7-6His fusion. It specifically cleaves between the Gln and Gly residues of the recognition sequence of LeuGluValLeuPheGlnGlyPro. One unit will cleave > 90% of 100 µg of a test GST-fusion protein in 50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH = 7.0 at 5 °C for 16 h. As the cleavage buffer was slightly modified (Table 5) protease cleavage was optimised by testing different protease amounts and different cleavage temperatures. To optimise protease amount 8.6 µg fusion protein was mixed with 1

U, 2 U, 3 U, 4 U, 5 U of PreScission Protease and incubated at 4 °C for 16 h. To optimise cleavage temperature, 1 U of protease was mixed with 11.2 µg and incubated for 16 h at 4 °C, 20 °C and 37 °C. The cleavage results were analysed using SDS-PAGE.

2.3.7 Protein labelling

ColN and ColA were fluorescently labelled for (fluorescent microscopy experiments at CLF, Harwell) with the CF™640R fluorescent dye (Biotium, Inc.) according to the manufacturer's instructions. ColN was labelled via a maleimide reaction in either one or two cysteine residues at positions 3 or 94 and 179 (Figure 11), while ColA was labelled with CF™640R succinimidyl ester via its primary amines, mostly lysine residues. The dye absorbs at 280 nm as well as around 600 nm and 640 nm. The final concentrations of labelled proteins was 4.2 mg/ml (91.91 mM) for S3C, 7 mg/ml (163.28 µM) for V94C/Q179C and 2.1 mg/ml for ColA.

>gi|116070|sp|P08083.1|CEAN_ECOLX RecName: Full=Colicin-N

MGSNGADNAHNNAFGGGKNPGIGNTSGAGSNGSASSNRGNSNGWSWSNKPHKNDGFHSDGSYH ITFHGDNNSKPKPGGNSGNRGNNGDGASAKVGEITITPDNSKPGRYISSNPEYSLLAKLIDAESIKGT EVYTFHTRKGQYVKVTVPDSNIDKMRVDYVNWKGPKYNNKLVKRFVSQFLLFRKEEKEKNEKEALL KASELVSGMGDKLGEYLGVKYKNVAKEVANDIKNFHGRNIRSYNEAMASLNKVLANPKMKVNKSD KDAIVNAWKQVNAKDMANKIGNLGKAFKVADLAIKVEKIREKSIEGYNTGNWGPLLLEVESWIIGGV VAGVAISLFGAVLSFLPISGLAVTALGVIGIMTISYLSSFIDANRVSNINNIISSVIR

Figure 11: Cysteine mutations in ColN at positions 3, 94 and 179. Highlighted residues have been mutated to cysteines prior to this project. S3C is single mutant in the ColN T-domain. V94C/Q179C is a double mutant in the ColN R-domain.

2.4 Biophysical methods

2.4.1 Liquid chromatography–mass spectrometry (LC-MS)

The CNI-HALO7-6His samples were prepared for LC-MS using the final CNI-HALO7-6His purification protocol with the Tris buffer outlined in 2.3.1. The final concentration of protein provided was 1.12 mg/ml. The LC-MS measurement and the data analysis was carried out by the LC-MS specialist Dr Joseph Gray, Newcastle University, as a service provided by Pinnacle Lab[2].

2.4.2 Circular Dichroism

To open up the helix-bundle structure of ColN P-domain in preparation for the pull-down assays (2.4.3), ColN P-domain, purified using the procedure outlined in 2.3.1, was incubated in a Tris buffer (10 mM Tris, 300 mM NaCl, pH=7.4) with and without 0.1 % DM for 24 h at 4 °C. Near-UV Circular Dichroism was used to investigate whether the tertiary structure had changed due to the added DM (Dover *et al.*, 2000). The spectra were measured on a CD Spectrometer Jasco J-810. Parameters: band width of 2 nm, response time: 4 sec, measurement range: 320-250 nm, data pitch: 0.5 nm, scanning speed 20 nm/min, cell length: 1 cm, temperature: 25 °C, protein concentration: 3.8453×10^{-5} mol/l. The traces are an average of 9 accumulations.

2.4.3 CNI pull-down assay

A pull-down assay was used to assess protein-protein interaction between CNI and Colicin N. CNI-HALO7-6His was bound via its Halo-tag to HaloLink resin (Promega) according to the manufacturer's instructions by calibrating the HaloLink resin with a Tris buffer (50 mM Tris-HCl, pH = 7, 300 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1 % DM, 10 % glycerol) and incubating the resin with CNI-HALO7-6His for 30 min at room temperature. Excess protein was washed off with Tris buffer and the resin was incubated with ColN, ColN P-domain, ColA and ColA P-domain in Tris buffers containing either 1.6 mM DM, 25 mM OG or 8 mM SDS at 4 °C overnight. Excess protein was washed off and to establish what remained bound to the beads due to protein-protein interaction, the beads were resuspended in protein loading buffer (0.125 M Tris HCl, pH = 6.8, 15 % glycerol, 5 mM NaEDTA. 2 % SDS, 0.1 %

[2] http://www.ncl.ac.uk/camb/pinnacle/

bromophenol blue, 1 % β-mercaptoethanol) and heat-denatured at 90 °C for 10 min. Washes and “pulled-down” proteins were analysed using SDS-PAGE.

2.4.4 Surface Plasmon Resonance (SPR)

In preparation for SPR, the proteins to be used here were purified with the protocols outlined in (2.3.1). For CNI, the CNI-Halo7-6His final purification protocol with Tris buffer was used. All proteins were buffer exchanged with a PD-10 column (GE Healthcare) against a HEPES buffer (10 mM HEPES pH = 7.5, 150 mM NaCl, 50 µM EDTA, different concentrations of detergents DM, OG or SDS) as used in Johnson *et al.* (2014) and incubated at 37 °C for 1 h to open up the P-domain structure with the detergent. The chip used for SPR was a GE Healthcare NTA Sensor Chip. The chip matrix consists of carboxymethylated dextran pre-immobilized with nitrilotriacetic acid (NTA). His-tagged proteins can be immobilised through nickel ions which bind the histidine imidazole rings as well as NTA.

Before the SPR binding experiments could be carried out, the polyhistidine-tag on ColN, ColA, ColN P-domain and ColA P-domain had to be modified with diethylpyrocarbonate (DEPC; Sigma-Aldrich) to prevent these proteins from binding to the NTA chip. The modification method was adapted from Lundblad and Noyes, 1984, chapter 9. A specific amount of DEPC was added to the protein solution and incubated at room temperature for 20 min, before quenching with an excess of 5 mM imidazole. It was unknown how much DEPC would be enough to modify the imidazole ring on the His-tag, but not disturb the protein structure and prevent potential protein-protein interaction. Therefore, different molar ratios were investigated and SPR was used to determine if the modified protein would still bind to the NTA-Ni^{+} chip.

Here, DEPC was added in molar ratios between DEPC and histidine molecules. All proteins were diluted to a final concentration of 15 µM. The molar concentration of histidine was calculated based on the protein concentration and the amount of histidines per molecule. For example, for a polyhistidine-tag with 6 histidines and a protein concentration of 15 µM, the solution had a histidine concentration of 6 x 15 = 90 µM. If DEPC was added in a 1:5 ratio, the final concentration of DEPC in the solution had to be 5 x 90 = 450 µM. ColN P-domain has 7 His, ColN has 13 His, ColA P-domain has 9 His and ColA has 13, all including the 6His-tag which is common to all these proteins. DEPC modifies the imidazole ring of histidine residues and cannot

discriminate between the histidine in the tag and the histidine in the main protein. The progress of modification was observed by measuring protein absorption at 240 nm and 280 nm with a UV spectrophotometer (UV-1800 Shimadzu). An optimisation experiment, testing if the modified protein still bound to the chip, was carried out with ColN P-domain, where DEPC was added in a ratio of 1:1 based on the 6 histidines in the tag only and based on all 7 histidines ratios of 1:1, 1:5, 1:10 and 1:25. To confirm 1:5 ratio as the optimal ratio, a further run was carried out with ratios 1:1, 1:2, 1:3, 1:4, 1:5, 1:6 based on all 7 histidines.

The SPR was carried out on a BIAcore X100 at a flow rate of 5 µl/min. The NTA chip has two chambers, one for measuring and one for reference to detect any non-specific binding. The chip is cleaned with 0.5 % SDS, 2 M NaCl, HEPES regeneration buffer (10 mM HEPES, 150 mM $NaCl_2$, 500 mM EDTA, 5 mM $CaCl_2$, pH = 8.3) and HEPES running buffer (10 mM HEPES, 150 mM $NaCl_2$, 50 µM EDTA, 5 mM $CaCl_2$, pH = 7, 0.1% DM). For each binding run, the HEPES Nickel solution (0.5 M $NiCl_2$, 10 mM HEPES, 150 mM $NaCl_2$, 50 µM EDTA, pH = 7.5) flows across the measuring cell only. Subsequently, all proteins flow across both chambers. For experiments where OG and SDS were added to the P-domain to open up its structure even further than by 0.1 % DM addition only, P-domains were modified with DEPC first (as above) and then incubated with OG (0, 0.1 %, 0.25 %, 0.5 %, 0.731 % (=CMC), 1 %, 1.5 %.) or SDS (0.4 mM, 0.8 mM, 1.6 mM, 4 mM, 8 mM (=CMC), 16 mM) for 1 h at 37 °C. The procedure was adapted from Johnson *et al.* (2014).

Chapter 3. CNI overexpression optimisation

3.1 Introduction

Immunity proteins to pore-forming colicins are largely understudied, certainly in comparison to the immunity proteins of enzymatic colicins (Cascales *et al.*, 2007; Papadakos *et al.*, 2011; Farrance *et al.*, 2013; Vankemmelbeke *et al.*, 2013; Kim *et al.*, 2014). Most of the literature describes *in vivo* characterisation and although some purification has been carried out, the *in vitro* characterisation has not made significant progress (Chapter 1: Introduction). Functional immunity protein reconstitution into artificial membranes or micelles has not been successful and immunity protein activity or interaction with other proteins, most notably the cognate colicins, has not been demonstrated *ex vivo*.

The main difficulty lies in the function of the protein. The immunity protein guards the cell against the attack of colicins, i.e. making it “immune” to colicin attack, allowing it to grow and replicate. In an *in vitro* setting there is no cell and it is therefore difficult to show the immunity protein’s function. The activity of soluble proteins, like enzymes, can be shown through the creation of a product, while the inhibition of purified toxic proteins (Kleanthous *et al.*, 1999), like colicins, can be shown through killing assays. It is difficult to show the activity of a protein whose main task it is to prevent the action of another membrane protein, hence, maintaining the integrity of the membrane. When immunity proteins are present, we assume that colicins cannot form a pore and depolarise the membrane. So, one could argue that the main task of the immunity protein is to prevent this pore formation and depolarization. It is possible to show *in vitro* if a membrane is depolarised (Tokuda and Konisky, 1978a) and, also, if pores in a membrane are formed (Wilmsen *et al.*, 1990; Lakey and Slatin, 2001). Therefore, it should be possible to detect if this is prevented through the immunity protein. However, the difficulty lies in creating a suitable environment which is similar enough to the *E.coli* inner membrane to keep the protein active. One feature would require the application of an electrical membrane potential difference to open the channel which is formed by a colicin molecule.

3.1.1 Studying membrane proteins in vitro

The possibilities of how the complex systems within a biological membrane are simplified and mimicked include reconstitution of the membrane protein into artificial membranes, such as lipid vesicles, nanodiscs created with a styrene maleic acid polymer (Knowles *et al.*, 2009; Dorr *et al.*, 2014) or membrane scaffold proteins (Bayburt and Sligar, 2010), or membrane like environment such as detergents or amphipols (Popot, 2010; Arunmanee *et al.*, 2014; Della Pia *et al.*, 2014; Kleinschmidt and Popot, 2014; Calabrese *et al.*, 2015). A suitable detergent must outcompete the membrane lipids for interaction with the protein and stabilise the protein in solution through hydrophobic interaction with the hydrophobic part of the protein as well as through hydrophilic interaction with the aqueous environment (Lemmon and Engelman, 1994; Stangl and Schneider, 2015).

Less complex experiments, where detergent solubilisation is sufficient, might include protein – protein interaction assays with potential binding partners (Helenius and Simons, 1975), isothermal titration calorimetry (Evans *et al.*, 1996b), surface plasmon resonance (SPR)(Stora *et al.*, 1999), pull-down assays etc. These are routinely used for soluble proteins, but it remains to be investigated how suitable they are for pore-forming colicins and their immunity proteins. With these assays the binding specificity, binding affinity and/or possibly stoichiometry can be measured although binding does not necessarily translate directly into activity. Unspecific binding, for example, through hydrophobic interactions in an aqueous environment can lead to false positive results. This is potentially the case for pore-forming colicins and their immunity proteins based on the mutagenesis work carried out *in vivo,* which suggests that pore-forming colicins and their immunity protein interact through hydrophobic α-helix interactions (see Chapter 1: Introduction and Chapter 4, pull down section).

3.1.2 Difficulties with overexpressing and purifying membrane proteins

For soluble proteins, expression and purification is often the trivial part of the research project as these proteins often express well in the cytoplasm of cells and well established protocols are available to purify them using special affinity, e.g. heparin or biotin, or purification-tag chromatography, e.g. ion metal affinity chromatography (IMAC), in combination with separation by charge or size. Some particularly thermostable proteins can even be separated by denaturation of all other proteins with heat, such as recombinantly expressed proteins from thermophilic

species, like polymerases from archaea. For membrane proteins each step can be more complicated.

Firstly, natural expression can be poor, and overexpression toxic, because additional protein is deposited into the membrane, an already tightly packed and much smaller environment than the cytosol. Toxicity can arise from non-specific interactions with essential proteins or create difficulty in maintaining the cell turgor due to increased permeability of the bilayer. To counteract such toxicity the cells can target these proteins for degradation, thus reducing yields of overexpressed proteins.

Secondly, retrieving the protein from the membrane can be a fine balancing act between choosing a detergent which is able to solubilise the protein but does not denature the protein's tertiary structure or obstruct potential interaction sites. Although a huge assortment of ionic and non-ionic detergents and surfactants is commercially available and has been used extensively in studies of membrane proteins, it remains empirical guesswork which detergent is suitable to solubilise and stabilise which protein (Loll, 2014). Some membrane proteins, in particular, β-barrel proteins, such as OmpF or OmpA, are known to refold following the replacement of chemical denaturants, like urea, or denaturing detergents such as SDS with milder detergents (Watanabe, 2002; Visudtiphole *et al.*, 2004; Visudtiphole *et al.*, 2005; Watanabe and Inoko, 2009), while an example of an α-helical protein is bacteriorhodopsin (Booth, 1997; Curnow and Booth, 2007). Whether this is also possible with the α-helical immunity proteins to pore-forming colicins is unknown. Also, the most suitable detergent for extraction might not be suitable for storage or down-stream processing in terms of compatibility with purification columns or interference with UV absorbance measurements, e.g. Triton X-100. Membrane protein crystallography often uses a mixture of several detergents or detergent-lipid mixtures to stabilise and solubilise proteins and so the possibilities of what could constitute the perfect purification conditions are even greater than the plethora of detergents, surfactants and buffers.

Finally, purification steps which usually involve a kind of chromatography can be difficult with membrane proteins because chromatography resins, such as the highly cross-linked beaded agarose in Superose® columns, can interact with hydrophobic parts of the protein if the wrong detergent or the wrong amount of detergent is used. Membrane proteins are a lot less stable in a polar/aqueous environment and tend not

to stay monomeric in solution. Hydrophobic interactions between proteins minimise the number of water molecules in contact with apolar protein surfaces, leading to aggregation, denaturation and precipitation. For the protein to remain monomeric, detergent-protein interactions must overcome protein-protein interactions. A successful purification, therefore, depends on the detergent properties and concentration (Popot, 2010).

3.2 The sequence of the Colicin N immunity protein

While Pugsley 1988 showed expression of the immunity protein and attempted overexpression of CNI immunity protein, his success was limited and there was no follow-up for this work. The CNI sequence on the pCHAP vector, the vector from which ColN, its lysis gene and CNI are naturally expressed, has been re-sequenced by the Lakey group. Figure 12 - Figure 16 show the alignment of the nucleotide as well as the amino acid sequence. Minor mistakes in Pugsley's sequencing led to a frame shift and a change in amino acid sequence, consequently causing him to choose a wrong N-terminus for CNI (Figure 12). Based on his sequence, Pugsley chose the longest possible polypeptide as the sequence for CNI with 131 amino acids and 15 245.2 Da. We propose that the correct polypeptide starts much earlier but because Pugsley's sequence contains a frame shift of two bases the amino acid sequence deviates from the newly proposed sequence and leads to a truncated polypeptide (Figure 12 and Figure 15), resulting in a polypeptide which is 11 074.0 Da, with 96 amino acids.

An alignment of Pugsley's and the newly proposed amino acid sequence for CNI shows how the sequencing mistakes result in a truncated protein which is missing a large part of the N-terminus (Figure 15), including the first transmembrane helix. Pugsley proposed that CNI belongs to the E1-type group of colicin immunity proteins to pore formers because it has three transmembrane helices. Based on the revised sequence we propose that it belongs to the A-type group because it has 4 transmembrane helices. This is supported by the high sequence similarity with the other group A colicins such as Colicin A (ColA) and Colicin B (ColB), with which ColN-P has clear homology.

The newly proposed sequences (Figure 14) have been published in Stroukova and Lakey, 2015 (in press and attached in appendix) but not yet deposited in the NCBI

data base. The activity of the protein produced from this sequence was confirmed before this project was started by previous members of the Lakey group (Fridd *et al.* (2002) and unpublished personal communication), although, the sequencing mistake was only recently clarified. The preceding work benefited from the cloning of a larger fragment which included the full sequence of not only CNI but also the gene coding for Colicin N. Fridd *et al.* (2002) showed that the recombinant expression of Colicin A and N or their basic domains required the respective immunity proteins to be present because overexpression is otherwise toxic to the cells. However, the precise activity of protein produced from this newly proposed sequence has been extended and confirmed in this thesis.

```
Pugsley1988_NG_034343.1      ATGGATATAAAAGACAGAAATAAGATATCAAAAAAAAATATCATTCAGTCT 50
Lakey2011                    ATGGATATAAAAGACAGAAATAAGATATCAAAAAAAAATATCATTCAGTCT 50

Pugsley1988_NG_034343.1      TCTGCTCTTACTTTCCCCATTCGCATTAATATTTTTCAGTTATAATAATG 100
Lakey2011                    TCTGCTCTTACTTTCCCCATTCGCATTAATATTTTTCAGTTATAATAATG 100

Pugsley1988_NG_034343.1      CACAAT--ACACTCCTCGAAAAAATCATCGCATACCTATCCCTACCAGGA 148
Lakey2011                    CACCAATACCACTCCTCGAAAAAATCATCGCATACCTATCCCTACCAGGA 150
                                * ****
Pugsley1988_NG_034343.1      TTTCATTCATTAAACAACCCGCCCCTAAGCGAAGCATTCAATCTCTATGT 198
Lakey2011                    TTTCATTCATTAAACAACCCGCCCCTAAGCGAAGCATTCAATCTCTATGT 200

Pugsley1988_NG_034343.1      TCATACAGCCCCTTTAGCTGCAACCAGCTTATTCATATTCACACACAAAG 248
Lakey2011                    TCATACAGCCCCTTTAGCTGCAATCAGCTTATTCATATTCACACACAAAG 250
                                                *
Pugsley1988_NG_034343.1      AATTAGAGTTAAAACCAAAGTCGTCACCTCTGCGGGCACTAAAGATATTA 298
Lakey2011                    AATTAGAGTTAAAACCAAAGTCGTCACCTCTGCGGGCACTAAAGATATTA 300

Pugsley1988_NG_034343.1      ACTCCTTTCACTATTCTTTATATATCCATGATATACTGTTTCTTGCTAAC 348
Lakey2011                    ACTCCTTTCACTATTCTTTATATATCCATGATATACTGTTTCTTGCTAAC 350

Pugsley1988_NG_034343.1      TGACACAGAACTAACCTTGTCATCAAAAACATTTGTATTAATAGTCAAAA 398
Lakey2011                    TGACACAGAACTAACCTTGTCATCAAAAACATTTGTATTAAT-GTCAAAA 399
                                                                      **
Pugsley1988_NG_034343.1      AACGATCTGTTTTTGTCTTTTTTCTATATAACACTATATATTGGGATATA 448
Lakey2011                    AACGATCTGTTTTTGTCTTTTTTCTATATAACACTATATATTGGGATATA 449

Pugsley1988_NG_034343.1      TATATTCACATATTTGTACTTTTGGTTCCTTATAGGAACATATAAGCTAT 498
Lakey2011                    TATATTCACATATTTGTACTTTTGGTTCCTTATAGGAACATATAAGCTAT 499

Pugsley1988_NG_034343.1      TTACCAGGGGAGGAATACTCCCCTGACAC 527
Lakey2011                    TTACCAGGGGAGGAATACTCCGTCGACAC 528
                                           ***
```

Figure 12: The resequenced CNI gene from the natural pCHAP4 vector (NCBI NG_034343.1, Pugsley (1987a) and Pugsley (1988)) demonstrating the mistakes in the published sequence, resulting in two frame shifts and several wrong bases. (*) indicates changes.

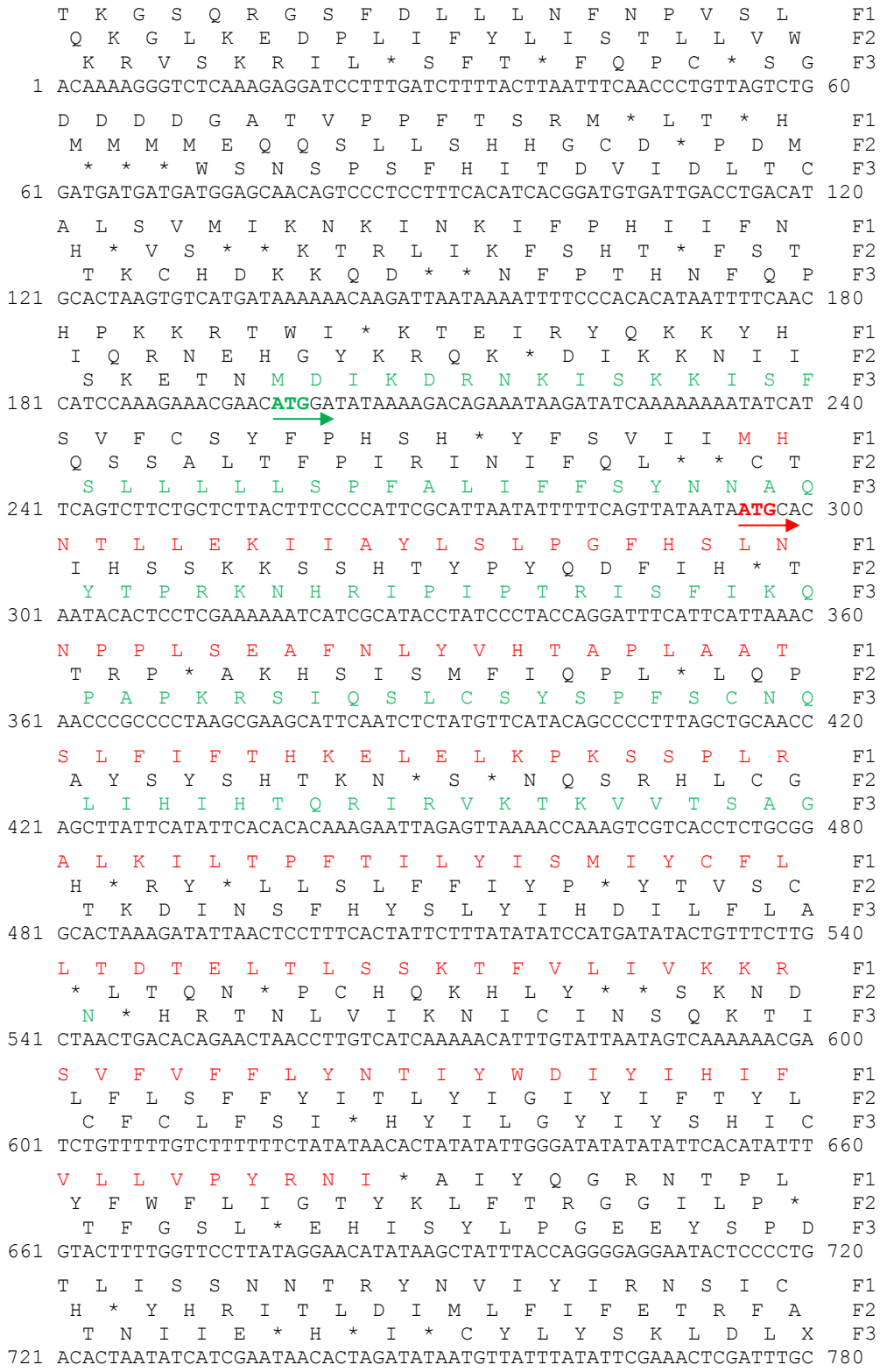


Figure 13: CNI sequence published by Pugsley (1988) translated into all 3 frames. Pugsley chose the longest possible polypeptide as the sequence for CNI with 131 amino acids and 15 245.2 Da (red). The correct polypeptide starts much earlier (green) but using Pugsley's sequence, is much shorter (green: 11 074.0 Da, 96 amino acids) due to the frame shift causing an incorrect premature stop codon.

```
   R  R  K  K  G  S  Q  E  D  P  L  I  F  Y  L  I  S  T  P  V     F1
    V  E  K  K  D  L  K  R  I  L  *  S  F  T  *  F  Q  P  L  L    F2
     *  K  K  R  I  S  R  G  S  F  D  L  L  L  N  F  N  P  C  *   F3
 1 CGTAGAAAAAAAGGATCTCAAGAGGATCCTTTGATCTTTTACTTAATTTCAACCCCTGTT 60

   S  L  D  D  D  D  G  A  T  V  P  P  F  T  S  R  M  *  L  T     F1
    V  W  M  M  M  M  E  Q  Q  S  L  L  S  H  H  G  C  D  *  P    F2
     S  G  *  *  *  W  S  N  S  P  S  F  H  I  T  D  V  I  D  L   F3
61 AGTCTGGATGATGATGATGGAGCAACAGTCCCTCCTTTCACATCACGGATGTGATTGACC 120
                                                        -35

   *  H  A  L  S  V  M  I  K  N  K  I  N  K  I  F  P  H  I  I     F1
    D  M  H  *  V  S  *  *  K  T  R  L  I  K  F  S  H  T  *  F    F2
     T  C  T  K  C  H  D  K  K  Q  D  *  *  N  F  P  T  H  N  F   F3
121 TGACATGCACTAAGTGTCATGATAAAAAAACAAGATTAATAAAATTTTCCCACACATAATT 180
                -10              Transcription start

   F  N  H  P  K  K  A  T  W  I  *  K  T  E  I  R  Y  Q  K  K     F1
    S  T  I  Q  R  K  Q  H  G  Y  K  R  Q  K  *  D  I  K  K  N    F2
     Q  P  S  K  E  S  N  M  D  I  K  D  R  N  K  I  S  K  K  I   F3
181 TTCAACCATCCAAAGAAAGCAACATGGATATAAAAGACAGAAATAAGATATCAAAAAAAA 240
              RBS       start of coding

   Y  H  S  V  F  C  S  Y  F  P  H  S  H  *  Y  F  S  V  I  I     F1
    I  I  Q  S  S  A  L  T  F  P  I  R  I  N  I  F  Q  L  *  *    F2
     S  F  S  L  L  L  L  L  S  P  F  A  L  I  F  F  S  Y  N  N   F3
241 TATCATTCAGTCTTCTGCTCTTACTTTCCCCATTCGCATTAATATTTTTCAGTTATAATA 300

   M  H  Q  Y  H  S  S  K  K  S  S  H  T  Y  P  Y  Q  D  F  I     F1
    C  T  N  T  T  P  R  K  N  H  R  I  P  I  P  T  R  I  S  F    F2
     A  P  I  P  L  L  E  K  I  I  A  Y  L  S  L  P  G  F  H  S   F3
301 ATGCACCAATACCACTCCTCGAAAAAATCATCGCATACCTATCCCTACCAGGATTTCATT 360

   H  *  T  T  R  P  *  A  K  H  S  I  S  M  F  I  Q  P  L  *     F1
    I  K  Q  P  A  P  K  R  S  I  Q  S  L  C  S  Y  S  P  F  S    F2
     L  N  N  P  P  L  S  E  A  F  N  L  Y  V  H  T  A  P  L  A   F3
361 CATTAAACAACCCGCCCCTAAGCGAAGCATTCAATCTCTATGTTCATACAGCCCCTTTAG 420

   L  Q  S  A  Y  S  Y  S  H  T  K  N  *  S  *  N  Q  S  R  H     F1
    C  N  Q  L  I  H  I  H  T  Q  R  I  R  V  K  T  K  V  V  T    F2
     A  I  S  L  F  I  F  T  H  K  E  L  E  L  K  P  K  S  S  P   F3
421 CTGCAATCAGCTTATTCATATTCACACACAAAGAATTAGAGTTAAAACCAAAGTCGTCAC 480

   L  C  G  H  *  R  Y  *  L  L  S  L  F  F  I  Y  P  *  Y  T     F1
    S  A  G  T  K  D  I  N  S  F  H  Y  S  L  Y  I  H  D  I  L    F2
     L  R  A  L  K  I  L  T  P  F  T  I  L  Y  I  S  M  I  Y  C   F3
481 CTCTGCGGGCACTAAAGATATTAACTCCTTTCACTATTCTTTATATATCCATGATATACT 540

   V  S  C  *  L  T  Q  N  *  P  C  H  Q  K  H  L  Y  *  C  Q     F1
    F  L  A  N  *  H  R  T  N  L  V  I  K  N  I  C  I  N  V  K    F2
     F  L  L  T  D  T  E  L  T  L  S  S  K  T  F  V  L  M  S  K   F3
541 GTTTCTTGCTAACTGACACAGAACTAACCTTGTCATCAAAAACATTTGTATTAATGTCAA 600

   K  T  I  C  F  C  L  F  S  I  *  H  Y  I  L  G  Y  I  Y  S     F1
    K  R  S  V  F  V  F  F  L  Y  N  T  I  Y  W  D  I  Y  I  H    F2
     N  D  L  F  L  S  F  F  Y  I  T  L  Y  I  G  I  Y  I  F  T   F3
601 AAAACGATCTGTTTTTGTCTTTTTTCTATATAACACTATATATTGGGATATATATATTCA 660

   H  I  C  T  F  G  S  L  *  E  H  I  S  Y  L  P  G  E  E  Y     F1
    I  F  V  L  L  V  P  Y  R  N  I  *  A  I  Y  Q  G  R  N  T    F2
     Y  L  Y  F  W  F  L  I  G  T  Y  K  L  F  T  R  G  G  I  L   F3
661 CATATTTGTACTTTTGGTTCCTTATAGGAACATATAAGCTATTTACCAGGGGAGGAATAC 720

   S  V  D  T                                                      F1
    P  S  T  X                                                     F2
     R  R  H                                                       F3
721 TCCGTCGACAC 731
```

Figure 14: Newly proposed CNI nucleotide sequence translated into all 3 possible frames. Frame 3 contains the correct polypeptide sequence (indicated green). Transcriptional elements (-35 site, -10 site and transcription start) as well as the ribosome binding site (RBS) and the first amino acid are indicated bold and underlined. The transcriptional elements have been determined using free bioinformatics software[3].

[3]http://www.fruitfly.org/seq_tools/promoter.html and
http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb

```
Green_Pugsley1988      MDIKDRNKISKKISFSLLLLLSPFALIFFSYNNA---------------QYTPRKN---- 41
correct_Lakey2011      MDIKDRNKISKKISFSLLLLLSPFALIFFSYNNAPIPLLEKIIAYLSLPGFHSLNNPPLS 60

Green_Pugsley1988      -------HRIPIPTRISFIK-----QPAPKRS-IQSLCSYSPFSCNQLIHIH-------T 81
correct_Lakey2011      EAFNLYVHTAPLAAISLFIFTHKELELKPKSSPLRALKILTPFTILYISMIYCFLLTDTE 120

Green_Pugsley1988      QRIRVKTKVVTSAGN---------------------------------------- 96
correct_Lakey2011      LTLSSKTFVLMSKNDLFLSFFYITLYIGIYIFTYLYFWFLIGTYKLFTRGGILRRH 176
```

Figure 15: Alignment of truncated polypeptide (green in Figure 12) and proposed amino acid sequence. The amino acid sequence deviates following the frame shift, which happens around 303 – 305 in the nucleotide sequence (Figure 12).

```
Pugsley1988      ----------------------------MHN----TLLEKIIAYLSLPGFHSLNNPPLS 27
Lakey2011        MDIKDRNKISKKISFSLLLLLSPFALIFFSYNNAPIPLLEKIIAYLSLPGFHSLNNPPLS 60
                                             *     **********************

Pugsley1988      EAFNLYVHTAPLAATSLFIFTHKELELKPKSSPLRALKILTPFTILYISMIYCFLLTDTE 87
Lakey2011        EAFNLYVHTAPLAAISLFIFTHKELELKPKSSPLRALKILTPFTILYISMIYCFLLTDTE 120
                 ************** ********************************************

Pugsley1988      LTLSSKTFVLIVKKRSVFVFFLYNTIYWDIYIHIFV----LLVPYRNI--------- 131
Lakey2011        LTLSSKTFVLMSKN-DLFLSFFYITLYIGIYIFTYLYFWFLIGTYKLFTRGGILRRH 176
                 **********  *    *  * * * *  ***         *   *
```

Figure 16: Alignment of Pugsley's and the newly proposed amino acid sequence for CNI. (*) indicates similarities. Underlined letters indicate predicted transmembrane α-helices. Pugsley's sequence is shorter and is missing the first transmembrane helix.

3.3 Overexpression of polyhistidine tagged CNI

The simplest way of purifying a protein is to attach a small purification tag, like a poly-histidine-tag, and purify the protein using affinity chromatography, like IMAC. In order to express CNI, the above CNI sequence was synthesised by GeneArt as a fusion with a 6His-tag, using pBAD322 (Cronan, 2006) as a backbone vector. As the insertion mechanism of CNI into the inner membrane is not known it was unclear whether an N-terminal or a C-terminal tag would be more beneficial or, indeed, make any difference for overexpression because CNI has no cleavable signal sequence, so both constructs were created. The pBAD322 vector is a low copy number vector which was designed as an expression system with an arabinose inducible promoter, where the amount of arabinose added regulated how much protein was expressed. This expression system was chosen because protein overexpression, and in particular membrane protein overexpression, can be a high burden to the cell metabolism and sometimes leads to toxic effects, resulting in poor cell growth or/and poor protein expression. BL21-AI *E. coli* cells are commercially available cells which are specifically designed for the overexpression of proteins from arabinose inducible expression systems and so these were chosen as hosts for an initial expression trial.

3.3.1 Polyhistidine-tagged CNI activity

The aim of the minimal inhibitory concentration (MIC) assay was to determine if the polyhistidine-tagged CNI is active and provides immunity against ColN to otherwise sensitive BL21-AI *E. coli* cells. The constructs pBAD322-*Para*-His-CNI (his-tag at N-terminal) and pBAD322-*Para*-CNI-His (his-tag at C-terminal) were transformed into BL21-AI cells and an empty pBAD322 vector was used as a negative control.

The MIC for CNI-6His and 6His-CNI was 100 nM ColN while it was 1000 times lower at 100 pM for the empty vector control. CNI-6His and 6His-CNI are both expressed and active. Both constructs were subsequently used for recombinant protein production and purification but yields sufficient for further *in vitro* analysis were not achieved. Here, purification attempts for 6His-CNI only are shown as examples.

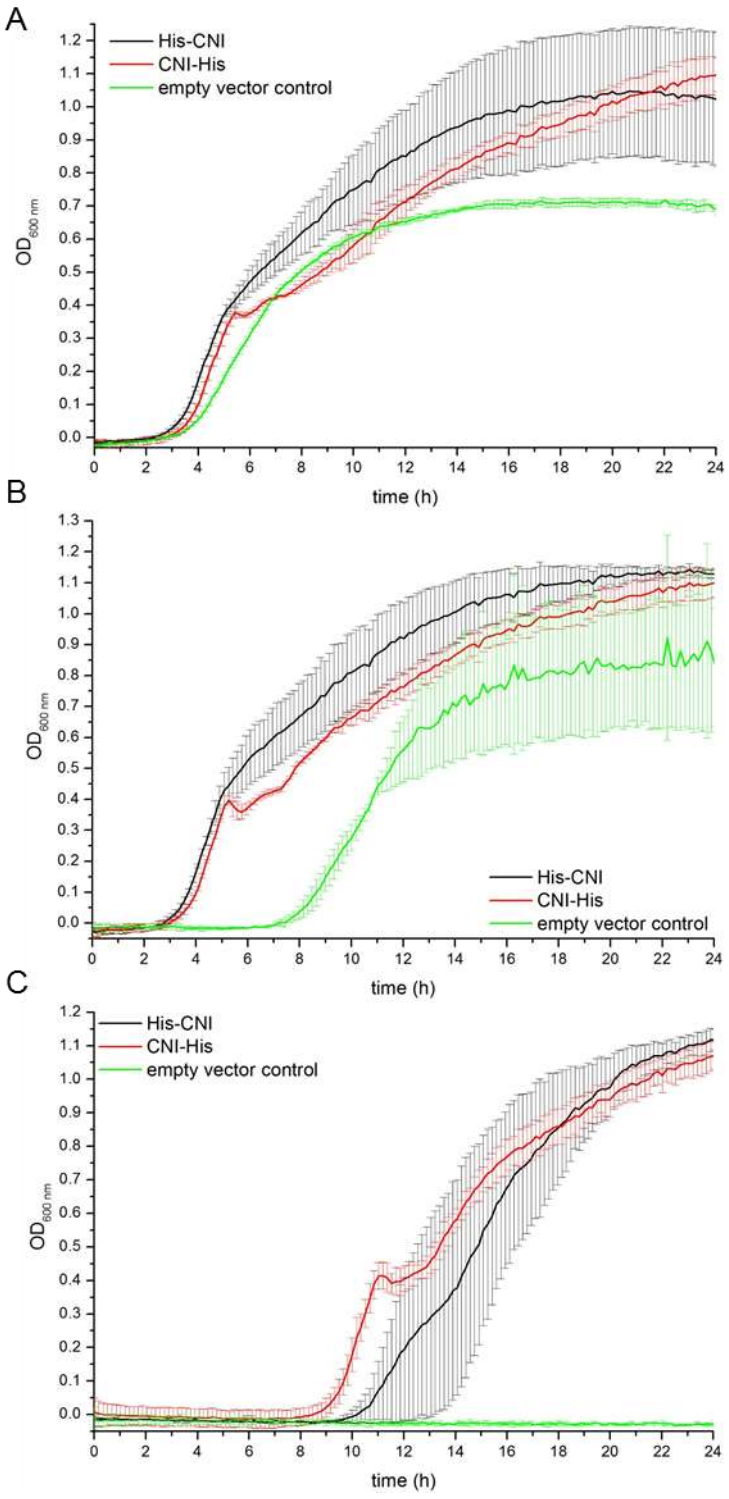


Figure 17: Polyhistidine-tagged CNI is active and protects otherwise sensitive BL21-AI cells against ColN. (A) No Colicin was added and culture growth is not significantly different. (B) When 10 pM of ColN were added, the culture viability of the empty vector control in comparison to the cultures expressing CNI is significantly reduced. (C) When 10 nM ColN were added, cultures expressing both versions of CNI are still viable while growth in the empty vector control is inhibited completely. The growth of BL21-AI cells expressing both versions of CNI is inhibited completely at 100 nM (not shown).

3.3.2 6HIS-CNI overexpression and protein aggregation

After showing that 6His-CNI was active and therefore expressed in the membrane, purification was attempted using immobilised metal affinity chromatography (IMAC) with n-dodecyl-β-D-maltoside (DDM) and a sodium phosphate buffer. Various attempts at purifying 6His-CNI from different cultures were carried out but the yield was very poor and, based on SDS-PAGE and western blot analysis probing for the polyhistidine-tag, the purified protein was mainly aggregated. Figure 18 shows an example of the purification fractions. The faint band at the top of the gel may also result from unspecific binding to large aggregates. It was not clear if these aggregates are artefacts of SDS-PAGE and whether they exist in solution. Due to the very low yield it was difficult to assess this further. Membrane proteins are likely to aggregate if an unsuitable detergent is used for solubilisation and so N-lauroylsarcosine, Genapol-X080 and Triton-X100 were used for further purification attempts in comparison to DDM.

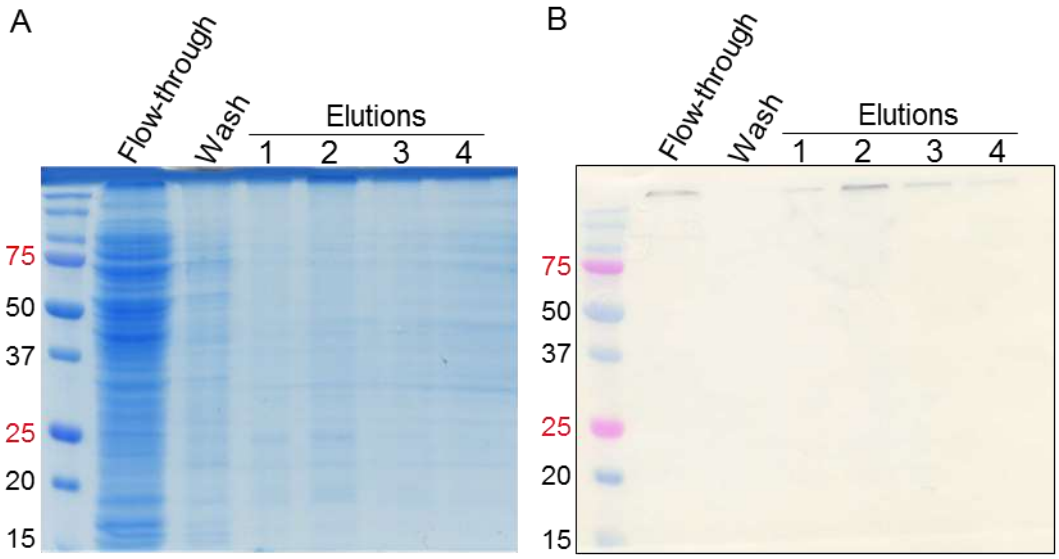


Figure 18: 6His-CNI purification. (A) Coomassie stained SDS-PAGE gel and (B) western blot of fractions samples from 6His-CNI purification using DDM. Protein transfer is greatly improved after protocol optimisation and the western blot probably shows the presence of 6His-CNI in the aggregates.

3.3.3 Detergent scouting

Four different detergents, n-dodecyl-β-D-maltoside, Genapol–X080, Triton-X100 and N-Lauroylsarcosine, were compared for their ability to solubilise 6His-CNI. The purification fractions were analysed using SDS-PAGE and a western blot (Figure 19). The Coomassie-stained SDS-PAGE gel (Figure 19A) shows differently sized bands in all four elutions and some bands are of approximately the expected size (21.62

kDa) for 6His-CNI but they could not be identified using the western blot and are therefore probably contaminants. The western blot only shows bands in the flow through lanes, which are likely caused by unspecific antibody binding (Figure 19B). Neither of the detergents tested seems to be particularly well suited to purify His-CNI.

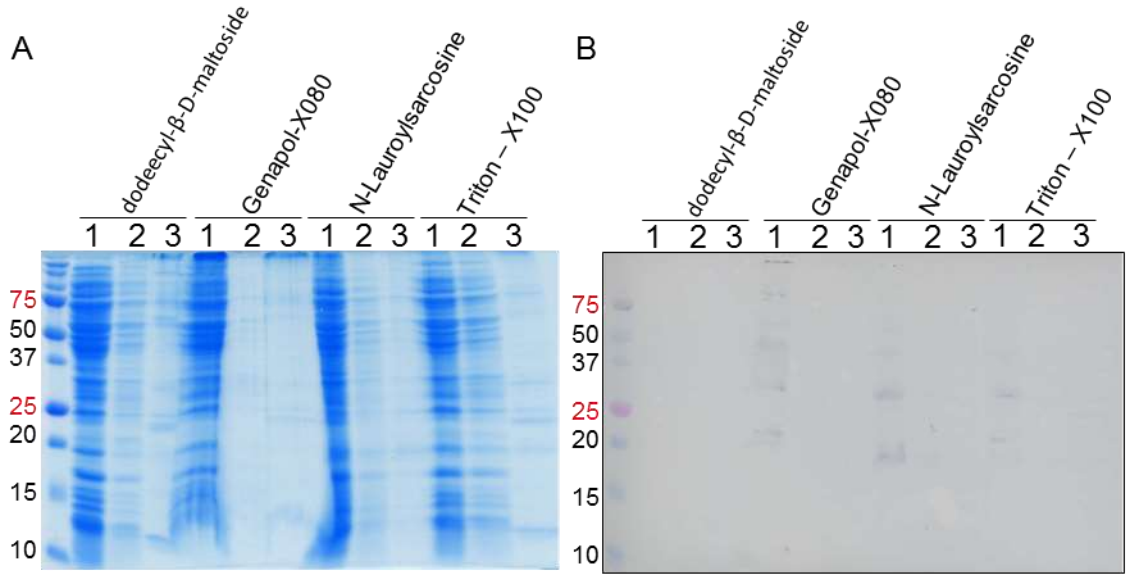


Figure 19: 6His-CNI purified with n-dodecyl-β-D-maltoside, Genapol–X080, Triton-X100 and N-Lauroylsarcosine. (A) Coomassie-stained SDS-PAGE gel and (B) western blot probing for the polyhistidine-tag show purification fractions. Lanes: 1) flow-through, 2) wash, 3) elution. The expected size of 6His-CNI is 21.62 kDa. Although the SDS-PAGE gel shows very faint bands of approximately the right size, the western blot only detects very faint bands in the flow-through lanes, which are probably due to unspecific anti-body binding.

In conclusion, 6His-CNI could not be purified in an amount suitable for further studies and purification attempts with other detergents than DDM did not lead to an improved yield. In fact, only *in vivo* activity data indicated that the protein was produced (3.3.1). Purification attempts with the C-terminally tagged CNI-His were also carried out but were equally unsuccessful (data not shown). A version of CNI fused to GFP was used for further purification attempts to make CNI easier to track during production and purification (Drew *et al.*, 2006; Carpenter *et al.*, 2008; Hammon *et al.*, 2009).

3.4 Overexpression of CNI-FLAG-TCS-sfGFP-6His and thrombin cleavage

3.4.1 Creating CNI-FLAG-TCS-sfGFP-6His

As well as creating a His-tagged CNI, a construct where CNI is C-terminally fused to superfolder GFP (sfGFP) and a polyhistidine –tag (6His) was also synthesised by GeneArt for *in vivo* localisation using fluorescent microscopy (Chapter 5). During my

Masters of Research Project at Newcastle University, it was established that this construct was active and could be overexpressed, however, purification still remained difficult. To continue this work, it was decided to tag CNI with a FLAG-tag and insert a thrombin cleavage site (TCS) between CNI-FLAG and sfGFP-6His (Figure 23) so that CNI could be cleaved from sfGFP-6His and characterised on its own. The thrombin cleavage site and the FLAG-tag were inserted using site-directed mutagenesis with partially annealing primers. Following successful mutagenesis, the overexpression was tested by assessing if it was toxic and if the produced protein was active. Membrane protein gene overexpression can be toxic due to the additional metabolic burden and the additional protein in an already crowded environment of the *E. coli* inner membrane. Culture viability is significantly reduced when CNI-FLAG-TCS-sfGFP-6His is expressed in comparison to the empty vector control cells which do not express the protein fusion (Figure 20). However, the protein expressing cultures still grow and overexpression may still be successful although culture volumes might have to be increased to compensate for reduced cell growth.

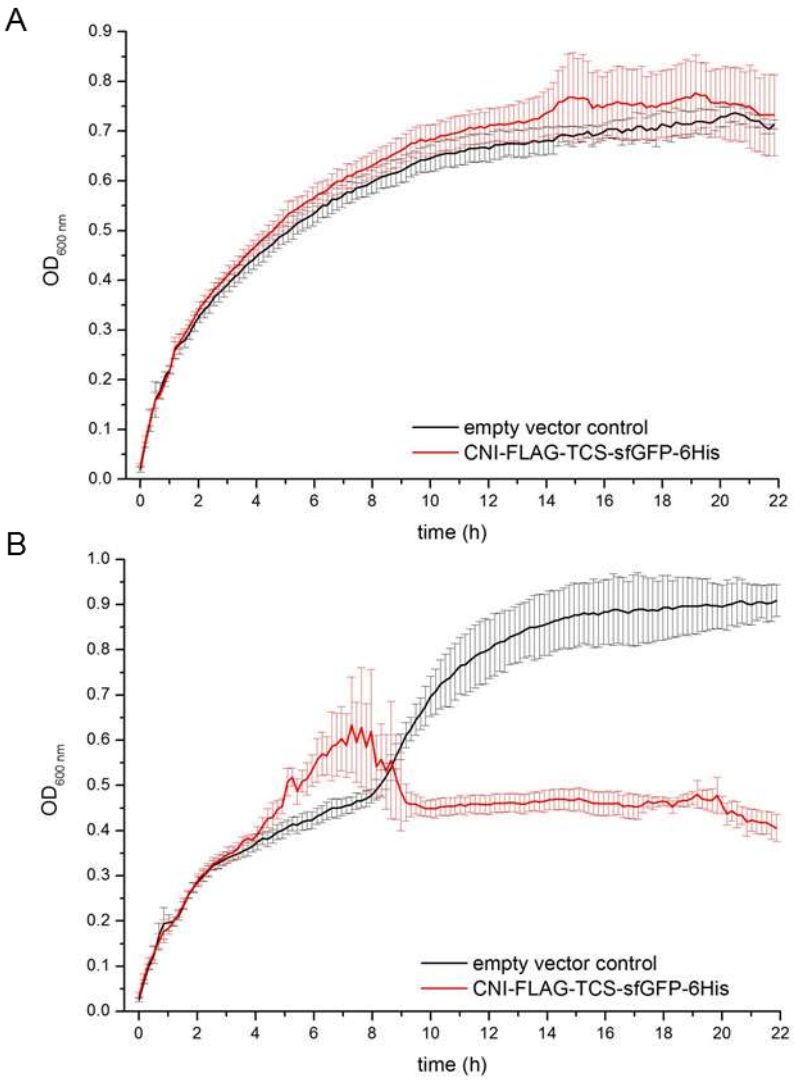


Figure 20: CNI-FLAG-TCS-sfGFP-6His over expression is toxic. (A) No induction with arabinose. CNI-FLAG-TCS-sfGFP-6His expressing BL21-AI cells and the empty vector control grow in the same way. (B) CNI-FLAG-TCS-sfGFP-6His overexpression induced with 0.2 % arabinose reduces culture viability through the additional metabolic burden. Cultures were grown in flasks at 37 °C, induced with arabinose at the start of culture growth in the plates and transferred to 96 well plates and a plate reader for recording over 22 h.

3.4.2 CNI-FLAG-TCS-sfGFP-6His activity

When overexpressing and modifying a protein with tags, especially, when, like here, the tags are larger than the protein itself, it is essential to determine if the fusion protein is still active. Protein activity proves that the protein is expressed, folded correctly and in this case, inserted correctly into its site of action, the *E. coli* inner membrane. The activity of CNI-FLAG-TCS-sfGFP-6His has been tested with a liquid culture killing assay (Figure 21). When no ColN is added, cells expressing CNI-FLAG-TCS-sfGFP-6His and the empty vector control show the same culture viability. Increasing ColN concentrations increased the lag times for both BL21-AI cells

expressing CNI-FLAG-TCS-sfGFP-6His and the empty vector control. All cultures initially contain the same amount of cells, estimated through optical density at $OD_{600\ nm}$ = 0.5. When colicin is added, the number of cells is greatly reduced, depending on antibiotic potency and concentration. Cultures are subsequently diluted 1:100, in order to dilute out the cell debris and the colicin to a hundredfold lower concentration. The optical density subsequently measured in the wells can almost entirely be attributed to living cells. The more cells are killed, the smaller is the inoculum for subsequent cultures during dilution and the longer it takes for visible growth, shown through well turbidity, to take off again. Therefore, the lethal impact of colicin addition is directly related to the culture's lag time. This is comparable to but much easier and accurate than the cell survival assay of Cavard and Lazdunski (1981). CNI-FLAG-TCS-sfGFP-6His activity was shown through a shorter lag time in comparison to the empty vector control at any ColN concentration higher than 1 pM. The greatest difference in lag time (ca. 8 h) between CNI-FLAG-TCS-sfGFP-6His expressing BL21-AI cells and the empty vector control is with 100 pM ColN. Given that the doubling time here is ca. 40 min, this corresponds to ca. 12 generations which corresponds to a factor of 2^{12} = 4096. After treatment with 100 pM ColN, there are 4000 times more cells in the immune culture than there are in the sensitive culture, which demonstrates a significant immunising effect of CNI-FLAG-TCS-sfGFP-6His. Cascales *et al.* (2007) reports that cells expressing immunity proteins protect against 10^4 -10^7 times the concentration of colicin needed to kill sensitive cells. Here, CNI-FLAG-TCS-sfGFP-6His protects up to 10^3 times the ColN concentration needed to kill sensitive cells. ColN concentrations higher than 1 µM have not been tested with this construct but have been tested with a different construct (Figure 35) and show that C41 cells expressing CNI-3C-Halo7-6His are killed by 5 µM. The lethal concentration of ColN for cells expressing CNI therefore lies between 1 µM and 5 µM and, while the lethal concentration of ColN to sensitive BL21-AI cells lies between 100 pM and 1 nM. The immunising factor is therefore in the range reported by Cascales *et al.* (2007).

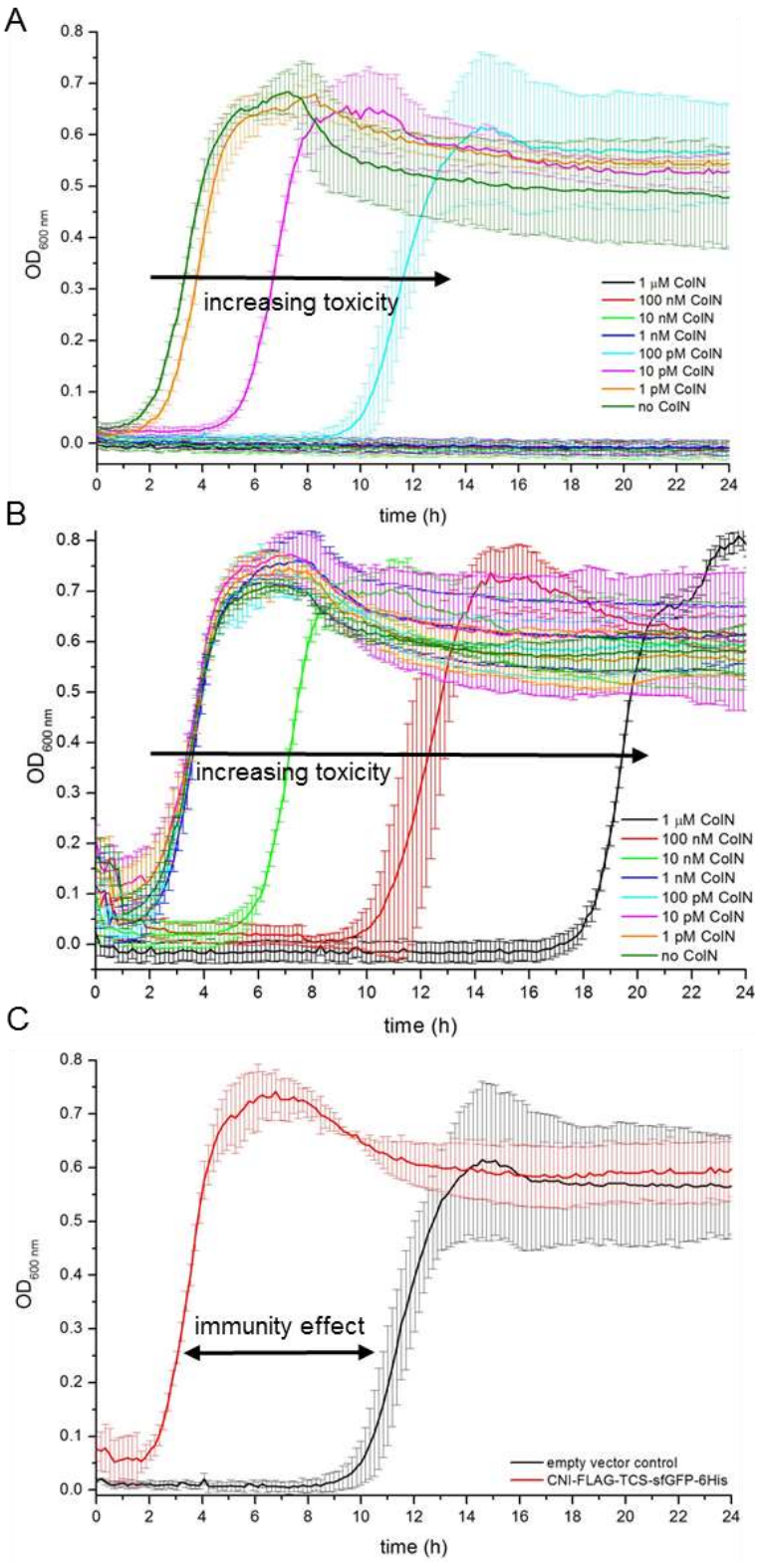


Figure 21: CNI-FLAG-TCS-sfGFP-6His protects otherwise sensitive cells against the lethal effect of ColN. (A) Empty vector control BL21-AI cells and (B) BL21-AI cells expressing CNI-FLAG-TCS-sfGFP-6His are affected by the increasing toxicity caused by increasing ColN concentrations. Given that the initial cell count is the same, the more ColN is added, the more cells are killed and the longer it takes for the cultures to reach a detectable exponential growth phase. Hence, increasing ColN concentrations lead to longer lag times. (C) CNI-FLAG-TCS-sfGFP-6His expressing BL21-AI cells and the empty vector control show the greatest difference in lag time (ca. 8 h) at 100 pM ColN.

3.4.3 CNI-FLAG-TCS-sfGFP-6His purification

After establishing that toxicity caused by overexpression was tolerable and the fusion protein was still active, the protein fusion was purified using a sodium phosphate buffer containing 2 % DM (n-decyl-β-D-maltoside). Figure 22 shows samples of fractions from the purification of CNI-FLAG-TCS-sfGFP-6His in a Coomassie stained gel, an anti-polyhistidine western blot and an in-gel fluorescence image. Solubilisation and purification of the protein fusion were successful although the purified protein fusion shows some minor aggregation in the eluates while the protein fusion still embedded in the membrane does not (whole cell and pellet fractions). This might indicate that the detergent DM has limited success in mimicking the membrane environment of CNI. It is possible that lipids or other proteins or physical factors, such as a certain pH or electrical potential, are important to keep the CNI soluble. A typical yield from a purification using these conditions was 2 mg from 5 l of culture. For the first time, purification of CNI in a fusion with GFP has been successful, although the individual parts of the process, such as the choice of the detergent, the buffer composition or the expression strain need further optimising.

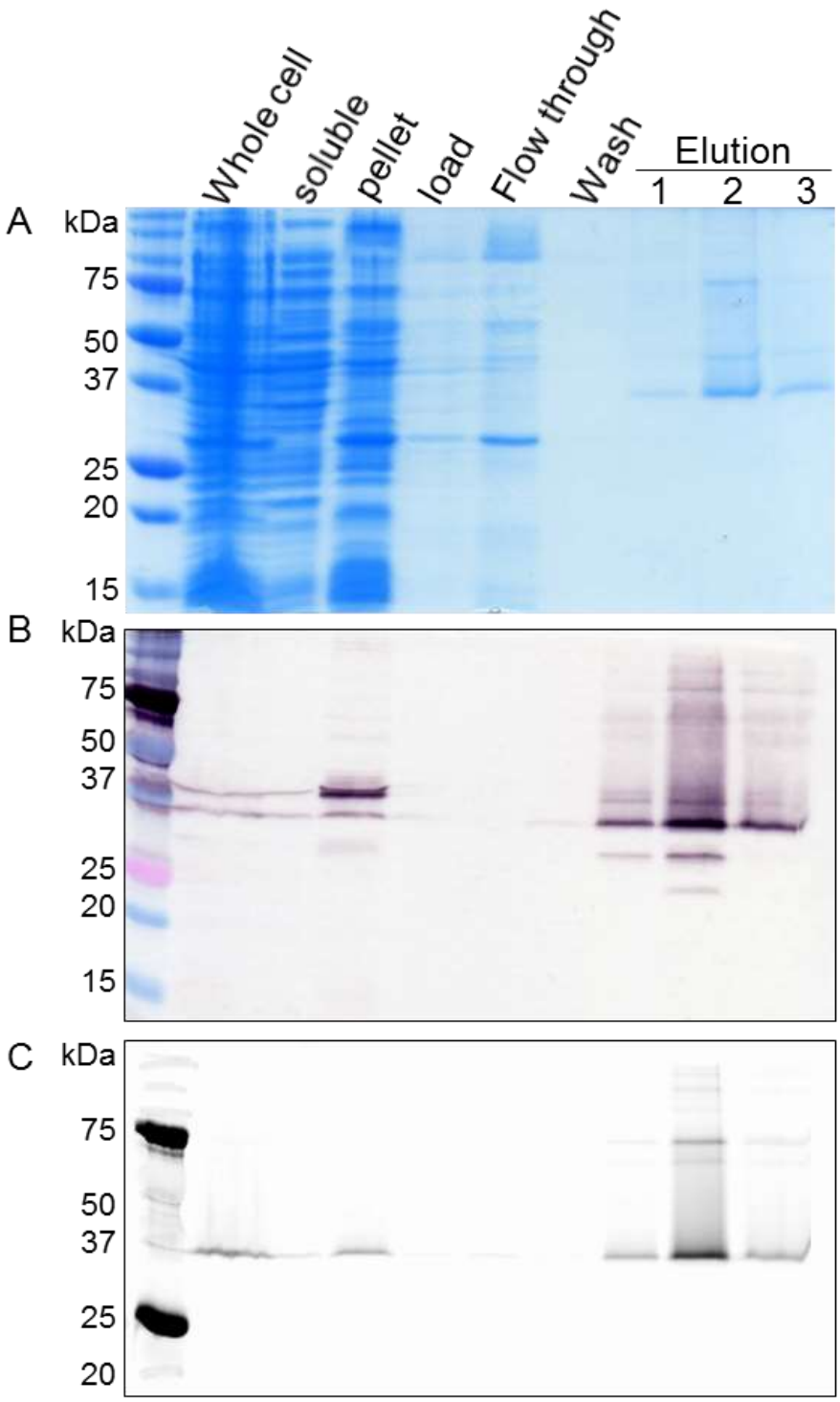


Figure 22: Purification of CNI-FLAG-TCS-sfGFP-6His. Purified from *E.coli* BL21-AI cells using 2 % DM, grown in LB at 37 °C. (A) Coomassie stained gel showing samples of fractions from the purification. (B) Western blot probing for the polyhistidine-tag at the C-terminal end of the protein fusion. (C) In-gel fluorescence (excitation at 485 nm for GFP) using a Typhoon imager. The purified protein fusion shows some minor aggregation (elutions) while the protein fusion still embedded in the membrane does not (whole cell and pellet).

3.4.4 Thrombin cleavage of CNI-FLAG-TCS-sfGFP-6His

Even though the purified protein showed signs of aggregation in the SDS-PAGE gel (Figure 22), it was unclear if this was a gel artefact because the protein solution looked clear to the eye and did not yield any visible precipitate upon spinning at 20 000 g in a bench top centrifuge. Therefore, protease cleavage along the previously introduced thrombin cleavage side (TCS, Figure 23) was carried out and anti-polyhistidine-tag as well as anti-FLAG-tag western blots were carried out to detect both cleavage products and establish cleavage efficiency. A cleavage control that was included in the commercial kit was used as a positive control to ensure thrombin was still active in the detergent containing buffer. The control proteins and thrombin have no tags. In the cleaved control lane of the Coomassie stained gel two cleavage products and thrombin are clearly visible, proving that thrombin is functional in the protein purification buffer (Figure 23B).

Cleavage of the protein fusion was only partially successful as some of the uncleaved fusion is still visible (white arrows, Figure 23). However, both cleavage products were also detected and more clearly visible in the blots than in the Coomassie-stained gel. Hence, thrombin cleavage was successful but still needed optimisation as some of the full length protein fusion is still present and there seems to be more of the sfGFP-6His-tag detected than of the cleaved CNI-FLAG protein. Apart from optimising the cleavage buffer composition and maybe the temperature during cleavage and the protease to protein fusion ratio, a more efficient protease may be needed.

The Coomassie-stained gel as well as the blots contain acid-denatured sample. Acetic acid has been used instead of heat denaturation, which is often used for soluble proteins, to unfold the protein structure. Heat denaturation is not commonly used to denature membrane proteins, especially α-helical proteins, because it induces aggregation. Acid treatment with acetic acid, a relatively safe alternative to trifluoroacetic acid (Sagne *et al.*, 1996), has been a useful method to unfold the structure of the protein fusion shortly before SDS-PAGE, allowing for a better detection with antibodies during the western blot and an assessment of the actual size of the protein. The protein bands shift up on the SDS-PAGE gel when acid is added and it appears as though there is, in total, more protein in the acid denatured lanes. In fact, the amount of protein in the first 4 lanes is the same but more protein is

detected in the acid treated samples because more epitopes are exposed. Acid treatment in the presence of SDS-PAGE loading buffer is therefore a good way of unfolding membrane proteins without heat. Unfortunately, the change in pH also leads to some aggregation but the greatly improved detection and the demonstration of actual size (Rath *et al.*, 2009) outweighs the disadvantages in this case. Only the protein fusion and the cleaved CNI-FLAG shift up in the acid denatured samples in comparison to the folded samples because they contain highly hydrophobic regions, while sfGFP-6His is unaffected by the acid addition. It is possible that the epitopes for the antibodies are much more occluded either by the protein fold or detergent molecules when no acid is added.

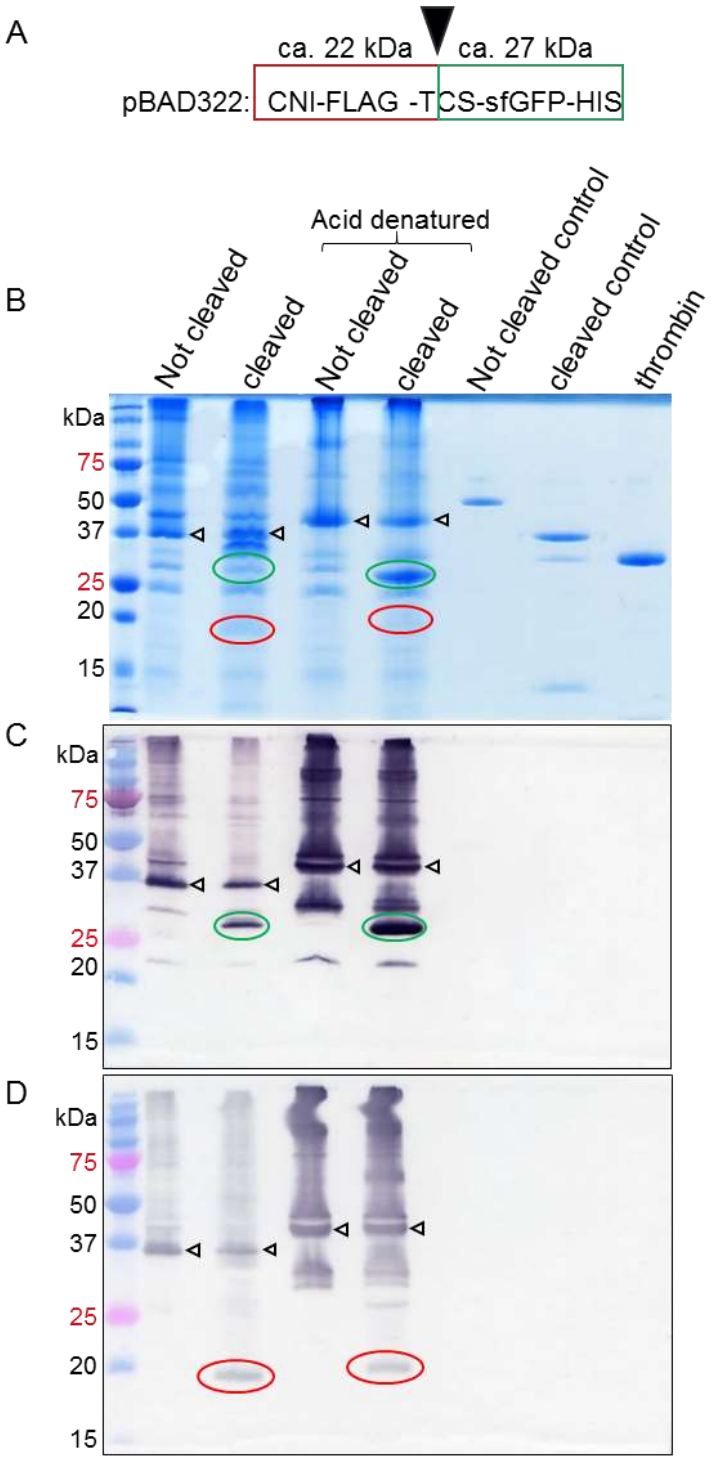


Figure 23: Thrombin cleavage of CNI-FLAG-TCS-GFP-6His. (A) CNI-FLAG-TCS-GFP-6His can be cleaved using thrombin at the thrombin cleavage site (TCS) between the FLAG-tag and sfGFP, giving products of ca. 22 kDa and ca. 27 kDa. (B) Coomassie stained gel showing cleaved and not cleaved CNI-FLAG-TCS-GFP-6His in a folded and acid denatured form. A cleavage control protein from the commercial kit was included to show thrombin is active in this buffer composition. A sample of thrombin protease was included to show it is pure and monomeric and for size comparison. (C) Anti-polyhistidine-tag western blot showing the protein fusion (white arrows) and cleaved sfGFP-6His (green circles). (D) Anti-FLAG-tag western blot showing the protein fusion (white arrows) and cleaved CNI-FLAG (red circles).

3.5 Protein solubilisation and stabilisation

3.5.1 Detergent screening using fluorescent size exclusion chromatography

Arguably, one of the most crucial parts in purifying membrane proteins is to choose the right detergent for solubilisation and stabilisation. Using a GFP-tagged version of CNI (CNI-FLAG-TCS-sfGFP-6His), 6 detergents were screened at the Membrane Proteins Laboratory (MPL, Harwell Research Complex, UK) using fluorescence size exclusion chromatography (FSEC, Hattori *et al.* (2012)), for their suitability to solubilise monomeric CNI. The MPL specialises in large scale membrane protein expression, purification and crystallization[4]. This GFP-tagged version of CNI is active *in vivo* (Figure 21). In preparation for FSEC, cell membranes from the same batch were homogenised in PBS at 4 °C. Aliquots of 900 µl were mixed with detergent (final concentration: 1 %), incubated and agitated for 1 h at 4 °C and ultra-centrifuged to pellet non-solubilised material. The supernatants containing the solubilised membrane protein were used in FSEC. Figure 24 shows relative fluorescence peaks obtained from FSEC; the higher the fluorescence the more GFP-tagged CNI has been solubilised. It is assumed that earlier peaks show higher order protein structures such as aggregates and polymers, while later peaks show monomeric proteins. In order of efficacy, N-Lauroylsarcosine, Cymal 6 and DM seem to be the best detergents, while Triton – X100, DDM and LDAO are poor at solubilising and stabilising the protein fusion. N-Lauroylsarcosine is not useful for most biophysical studies because it is likely to denature proteins. Triton – X100, DDM and N-Lauroylsarcosine were included in this screen because they have been used in the previous screen with 6 His-CNI (section 3.3.3)

[4] http://www.diamond.ac.uk/Beamlines/Mx/MPL.html

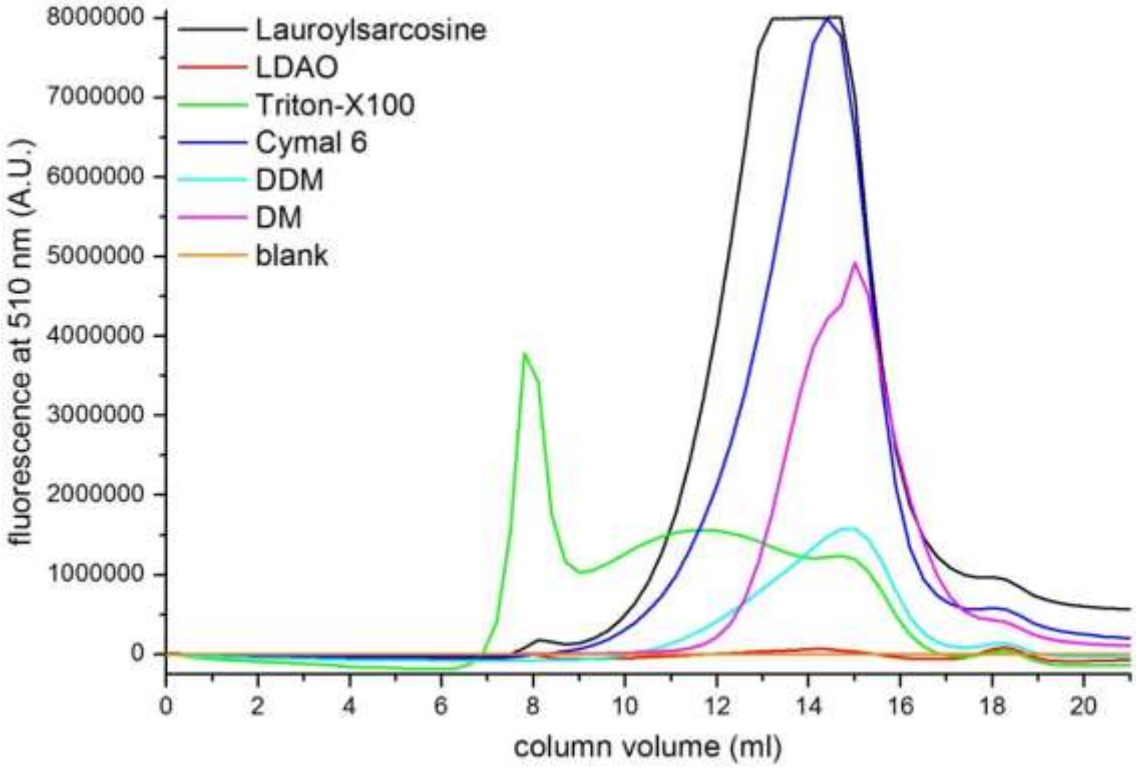


Figure 24: Solubilisation of CNI-FLAG-TCS-sfGFP using different detergents at 1 %. The sample solubilised by lauroylsarcosine has saturated the fluorescence detector, resulting in a cut off curve. In order of purification efficiency, lauroylsarcosine is the best detergent followed by Cymal 6, DM, DDM, Triton-X100 and LDAO. LDAO completely fails to solubilise CNI-Flag-TCS-sfGFP-6His. The blank is a buffer control.

3.5.2 Purification of CNI-FLAG-TCS-sfGFP-6His with Cymal 6 and thrombin cleavage

Cymal 6 is the non-denaturing detergent that solubilised the most fusion protein (Figure 24) and was therefore used in a large scale purification. The protein fusion was solubilised using a sodium phosphate buffer containing 1 % Cymal 6 and purified using IMAC. This high concentration of Cymal 6 was crucial to allow solubilisation of the protein but was subsequently reduced to 0.09%, which equals three times the CMC of Cymal 6, for IMAC and cleavage. To the elution (Figure 25, lane 9), 476 µM DTT, 2.4 mM $CaCl_2$ and 50 U thrombin were added in order to cleave the fusion protein while it was dialysed against Tris (pH = 7.2), 300 mM NaCl, 10 % glycerol, 0.09 % Cymal 6. Unfortunately, most of the protein aggregated and precipitated during dialysis. The aggregate was dissolved in 100 % acetic acid and loaded for comparison (Figure 25, lane 10). Either Cymal 6 is able to solubilise but not suitable to stabilise the protein or the buffer exchange or the cleavage itself resulted in aggregation. Another cause for aggregation may be the high concentration of imidazole at 500 mM in Figure 25, lane 8.

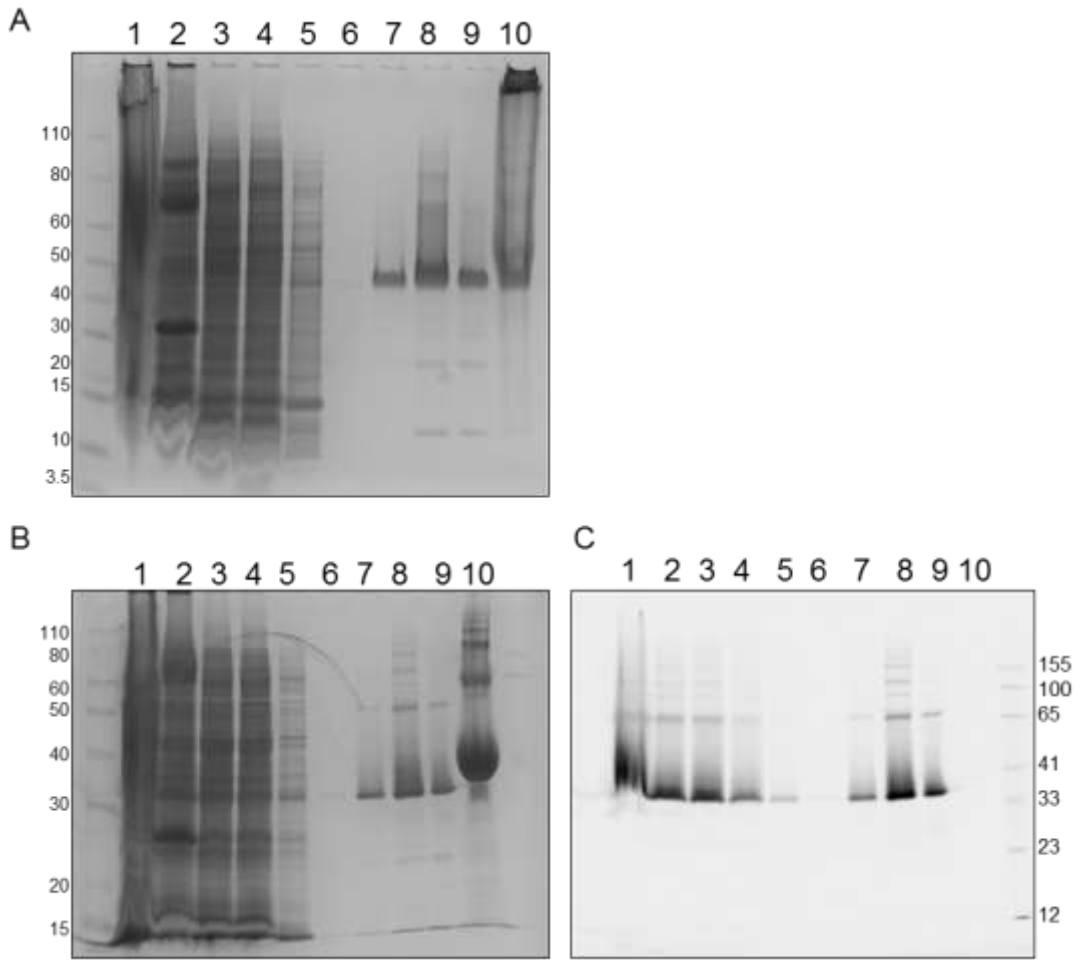


Figure 25: Large scale purification of CNI-FLAG-TCS-sfGFP-6His using Cymal 6, displayed in a Coomassie-stained 4-12 % Bis-tris SDS-PAGE gel (A), 12% Coomassie-stained SDS-PAGE Tris-gel (B) and in-gel fluorescence image (excitation at 485 nm) of the 12 % SDS-PAGE Tris-gel (C). Purification was carried out using IMAC on Ni^{2+}-NTA agarose. Lanes: 1: homogenised membrane in PBS, 2: solubilised membrane with 1 % Cymal 6, 3: load, 4: flow-through from IMAC, 5: Wash with a sodium phosphate buffer including 20 mM imidazole, 6: Wash with a sodium phosphate buffer including 50 mM imidazole, 7: Elution with a sodium phosphate buffer including 250 mM imidazole, 8: Elution with a sodium phosphate buffer including 500 mM imidazole, 9: supernatant after cleavage with thrombin and spin, 10: Aggregate in acetic acid. The aggregation ladder showing differently sized aggregates for the purified proteins (lanes 7 - 9) is visible in the Coomassie-stained gels (A and B) clearly originates from the CNI-FLAG-TCS-sfGFP-6His fusion as it is also seen when looking at in gel fluorescence (C).

The precipitated protein was spun down to determine what remained in solution (Figure 25, lane 9). With the remaining material now fully dialysed into the Tris thrombin cleavage buffer, cleavage was attempted again. The cleavage was not successful and more protein precipitated, maybe because Cymal 6, in contrast to the DM used in section 3.4.4 and Figure 23, is unsuitable to stabilise the protein fusion and potential cleavage products (Figure 26). However, even without the addition of

thrombin the purified protein seems to be unstable as both aggregation and degradation products are visible.

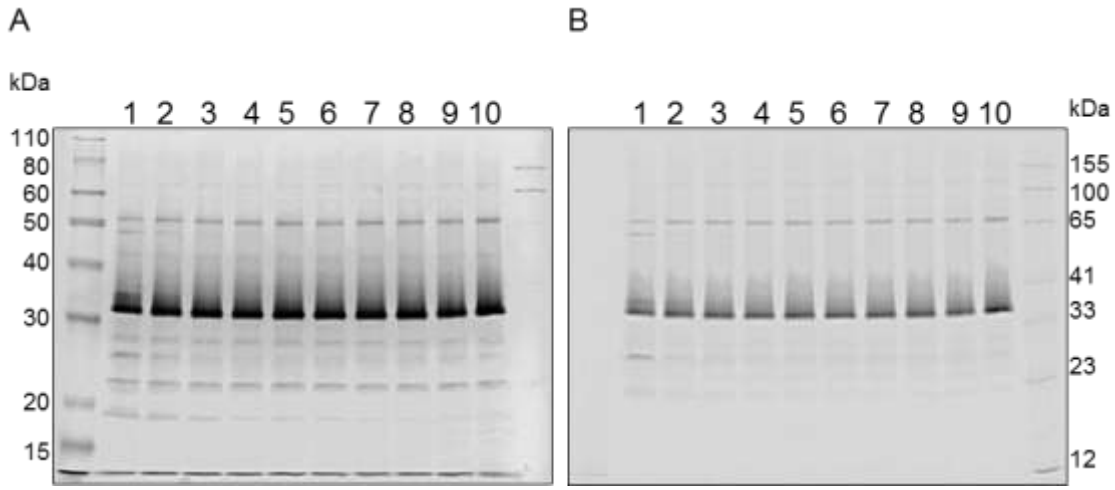


Figure 26: Thrombin cleavage of Cymal 6 purified CNI-Flag-TCS-GFP-His. (A) Coomassie stained SDS-PAGE and (B) in-gel fluorescence image. Lanes contain thrombin concentrations: 1: 100 U/ml, 2: 50 U/ml, 3: 25 U/ml, 4:12.5 U/ml, 5: 6.3 U/ml, 6: 3.1 U/ml, 7: 1.6 U/ml, 8: 0.8 U/ml, 9: 0.4 U/ml, 10: no thrombin. 1 µl of diluted thrombin was mixed with 9 µl of 41ug/ml and incubated for 16 h at 4 °C.

3.6 Overexpression of pWALDO-CNI-TEV-eGFP-6His

Following the unsuccessful cleavage with thrombin at MPL, a new construct was created where CNI was cloned into MPL's pWALDO vector using InFusion cloning creating a CNI-TEVCS-GFP-8His fusion protein. "TEV" indicates the location of the TEV cleavage site between CNI and GFP-8His. TEV is a protease from the tobacco etch virus and may be more efficient at cleaving the protein fusion than thrombin. It can also be made in large quantities in house (Figure 28). Switching the vector also has potential advantages for increasing the protein yield by increasing expression levels. While pBAD322 is a low copy number vector with an arabinose inducible promoter, pWALDO is a high copy number vector with an IPTG inducible T7 promoter. Also, this fusion includes 8 histidines in the C-terminal tag and may therefore bind better to the Ni^{+}-NTA resin because it is more likely to protrude outside of the detergent micelles.

The protein fusion expressed from this vector was active and apart from being used for protein production was also used to mutate the single cysteine residue of CNI and assess the activity of the C113S mutant. For a demonstration of activity through the protection of otherwise sensitive cells, refer to Chapter 4.

3.6.1 CNI-TEVCS-GFP-8His and CNI^{C113S}-TEVCS-GFP-8His purification

CNI-TEVCS-GFP-8His and CNI^{C113S}-TEVCS-GFP-8His were overexpressed and purified using the Cymal 6 protocol used to purify CNI-FLAG-TCS-sfGFP-6His at MPL (section 3.5.2) with 1 % Cymal 6. It was hypothesised that the serine mutant may be less dimerised than the WT because possible disulphide bonds are prevented. A Coomassie-stained SDS-PAGE gel shows that there is no difference between the WT and the mutated CNI in terms of purification yield and stability (Figure 27). Both constructs show signs of aggregation and degradation. Some of the protein bound to the Ni^{+} - NTA resin remained attached to the beads and did not elute from the beads so some protein was lost during the purification process, which indicates again that Cymal 6 is not a suitable detergent to keep the protein soluble. The aggregates in the eluted protein were spun down and the protein remaining in solution was used for protease cleavage with TEV. Due to the aggregation consistently observed with Cymal 6, it was decided that based on the screen performed at MPL (Figure 24), DM would be the next best suitable detergent for solubilising any CNI fusion proteins in further purification attempts as the fluorescent signal peak is not as high as for Cymal but was narrower and could indicate monomeric protein in detergent micelles.

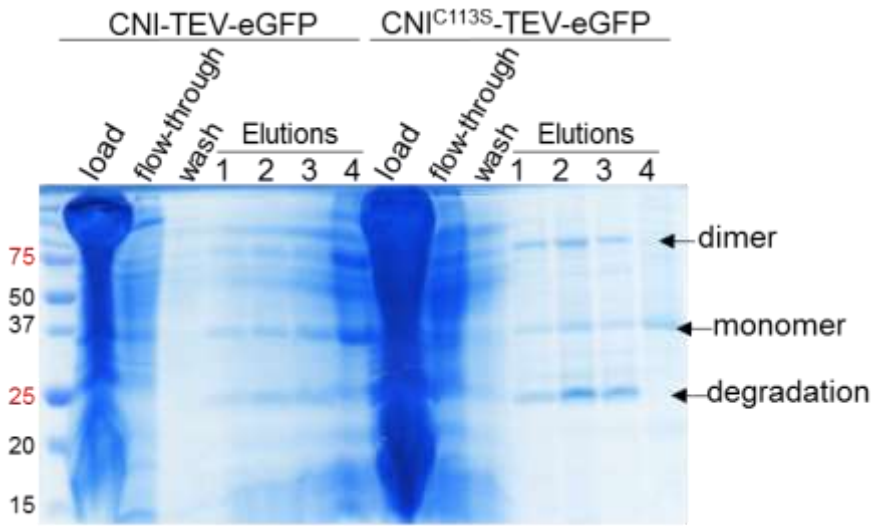


Figure 27: CNI-TEV-eGFP-8His and CNI^{C113S}-TEV-eGFP-8His mutant purification with 1 % Cymal 6. There is no difference between the WT and the mutated CNI in terms of purification yield and stability. Both constructs show signs of aggregation and degradation.

3.6.2 Protease cleavage with TEV protease

TEV protease cleavage was optimized using a protocol provided by MPL, which was adapted from Blommel and Fox (2007). A large amount of TEV-6His, in total 120 mg, was purified during the first IMAC from 5 l of culture and a second IMAC was performed to remove contaminants. The purified protease was used to cleave CNI-TEV-eGFP-8His and CNI^{C113S}-TEV-eGFP-8His.

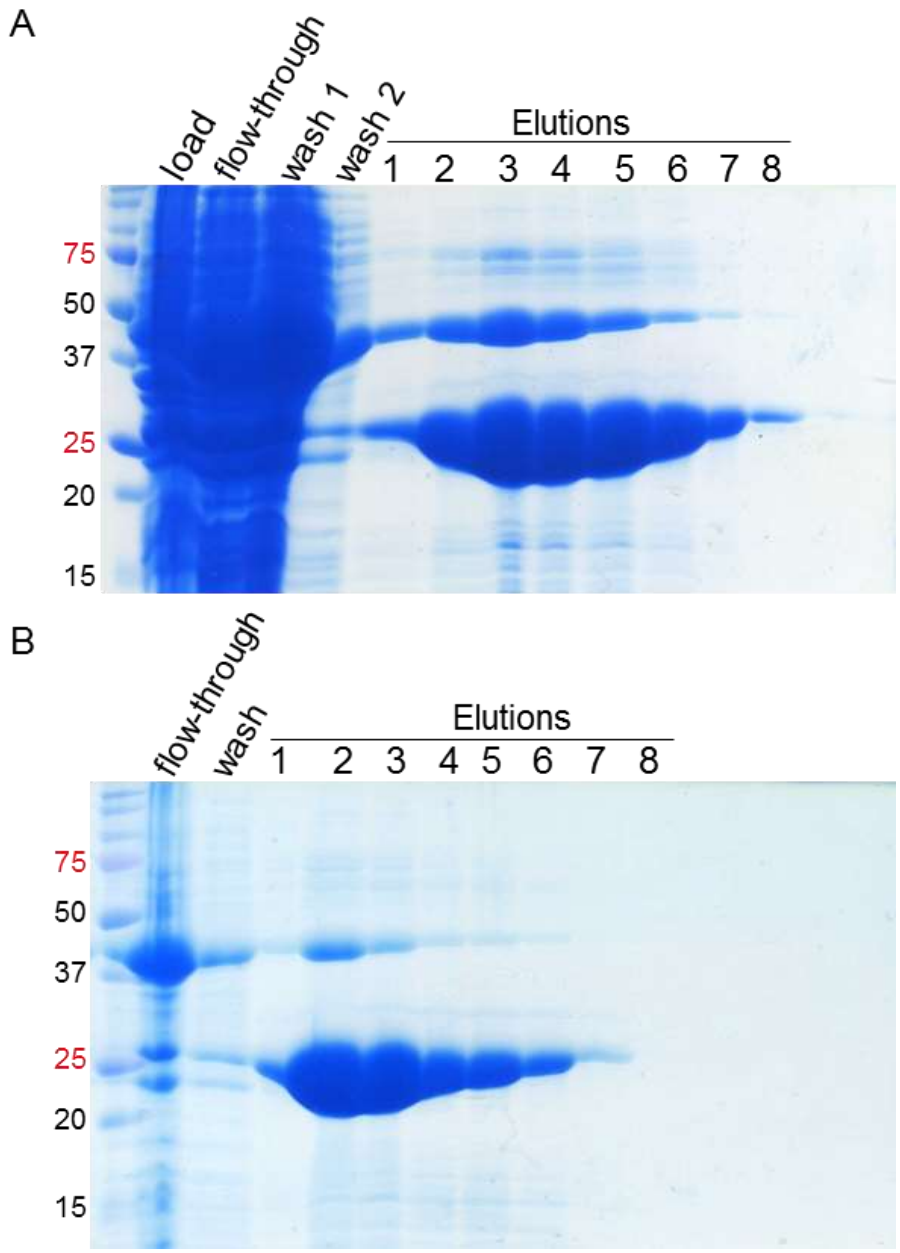


Figure 28: TEV-6His purification using IMAC was carried out twice. (A) First IMAC after cell disruption. (B) Second IMAC using the pooled elutions from the first IMAC as load. Total yield 120 mg from 5 l of culture.

CNI-TEV-eGFP-8His and CNI^{C113S}-TEV-eGFP-8His were cleaved with TEV protease at 4 °C for 16 h (Figure 29). Cleavage results were visualised using SDS-PAGE and

in-gel fluorescence. Cleavage seems to have been successful because less of the monomer is present the more TEV protease is added. However, it is difficult to determine what the cleavage products are due to the aggregation and possible degradation products, of which many are also visible with in-gel fluorescence and therefore must contain eGFP-8His. The bands that could represent CNI in the Coomassie stained gels are very faint and are certainly not adequate for further use. This is similar to the results achieved with thrombin and therefore indicates that this might be a general characteristic of CNI and unlikely to be caused by protease degradation of CNI itself. However, CNI may become unstable when cleaved from its tag. There is no difference in stability or aggregation between CNI-TEV-eGFP-8His and CNI^{C113S}-TEV-eGFP-8His.

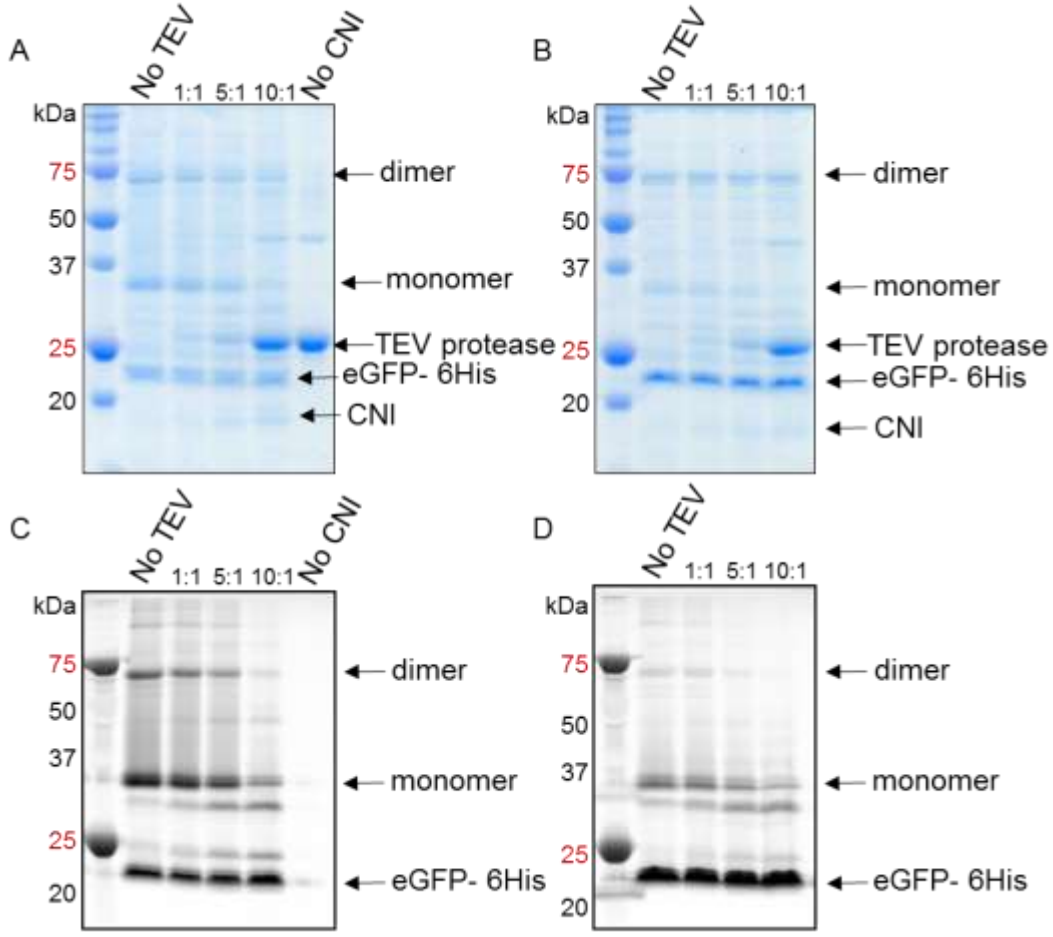


Figure 29: TEV cleavage of CNI-TEV-eGFP-8His and CNI^{C113S}-TEV-eGFP-8His with TEV protease at 4 °C for 16 h. TEV protease was mixed with CNI-TEV-eGFP-8His (10 µl of 0.74 mg/ml) and CNI^{C113S}-TEV-eGFP-8His (10 µl of 0.66 mg/ml) in molar ratios of 1:1, 5:1 and 10:1. Coomassie stained SDS-PAGE gel (A) and in-gel fluorescence image (C) of CNI-TEV-eGFP-8His after cleavage. Coomassie stained SDS-PAGE gel (B) and in-gel fluorescence image (D) of CNI^{C113S}-TEV-eGFP-8His after cleavage.

3.7 Optimising the expression system

3.7.1 Screening different tags, expression strains and media at OPPF, Harwell.

As well as attempting to find a suitable detergent for solubilisation and stabilisation of the protein, the expression system was also further optimised to improve overall yield. This included the use of protein fusions which are widely used for protein stabilisation and solubilisation, such as maltose binding protein (MBP), Glutathione S-transferase (GST) or Promega's Halo-tag (Los *et al.*, 2008). A screen of the expression levels of different tags was performed at the Oxford Protein Production Facility UK[5] (OPPF, Harwell Science Campus, UK) (Table 7). OPPF is a high-throughput cloning and protein production facility, focused on optimising expression of difficult proteins and scale-up of protein production. Except for the C-terminal His-tag, the 3C-GFP-tag and the Halo7-tag, all tags were N-terminal. All constructs were cloned into OPPF pOPIN vectors using InFusion cloning (Clontech) and transformed into *E. coli* strains C41, C43, Lemo21(DE3) or Rosetta™ 2. The vector abbreviations for CNI and CNI-TEV-GFP expression in Table 7 correspond to the OPPF nomenclature on their website.

Immunity proteins are known to interact with the pore-forming domains of colicins (Bishop *et al.*, 1985; Mankovich *et al.*, 1986; Benedetti *et al.*, 1991a), and so ColN P-domain was periplasmically co-expressed with CNI from a separate vector in case a CNI – ColN P-domain complex was more stable (Table 8). The vector abbreviations used for P-domain expression are shown in Table 8 and correspond to the OPPF nomenclature on their website. For co-expression of CNI fusions and P-domain from two different vectors in *E. coli* strains C41, C43 and Lemo21(DE3), the constructs where CNI was fused to the N- or C-terminal His-tag, N-His-GST or N-His-MBP were used.

EGFP-His and pWaldo-CNI-TEV-eGFP-6His, until then the most efficient construct, were used as controls. All screens were performed in two different media, Overnight Express™ Instant TB Medium (TBO[6]) and Power Broth™ (PB[7]), in 3 ml cultures. PB cultures were grown at 37 °C, induced with 1 mM IPTG at $OD_{600\ nm}$ = 0.5 and incubated over night at 20 °C. TBO is an auto-induction medium, so cultures were

[5] https://www.oppf.rc-harwell.ac.uk/OPPF/
[6] http://www.emdmillipore.com/life-science-research/overnight-express-instant-tb-medium/EMD_BIO-71491/p_iYSb.s1O34EAAAEjBRp9.zLX
[7] http://www.moleculardimensions.com/shopexd.asp?id=3455

grown at 37 °C during the day and 25 °C overnight. Cultures were harvested and the fusion proteins were purified using IMAC utilising the His-tag present on all constructs. Protein eluates were analysed qualitatively with SDS-PAGE. Ten gels in total were used for analysis; an example of this kind of gel is shown in Figure 30 (for other gels see appendix). All protein fusions which contained GFP and were examined separately using in-gel fluorescence at 510 nm. An example is shown in Figure 31 (for other gel see appendix). In-gel fluorescence identifies which constructs are overexpressed, the relative level of expression in comparison to other conditions and highlights protein degradation and some higher order structures. It also helps identifying proteins without a western blot. Constructs without additional solubilisation tags and fusion protein partners seem to break at the fusion point more easily than constructs with additional proteins. It is unclear why this is the case but possible that the bulk of protein is somehow sterically hindering proteases or, possibly, the three proteins are interacting and stabilising each other. ColN P-domain co-expression does not improve CNI-TEV-eGFP expression levels or stability.

Table 7: Protein and tag combination used during the overexpression screen, including the tag and protein fusion sizes. Vector names refer to nomenclature used by OPPF[8]. All were cloned into pOPIN vectors with spectinomycin (pOPINCD vectors) or ampicillin selection.

Vector name	Tag	Tag MW in Da	MW of fusions with CNI (20 431 Da) in Da	Fusions with CNI-TEV-GFP (49 882 Da)
pOPINCDE	C-His	969	21 400	x
pOPINCDF	N-His	2 158	22 589	x
pOPINS3C	N-His-SUMO3C	13 213	33 644	63 095
pOPINMSYB	N-His-MSYB	16 268	36 699	66 150
pOPINCDJ	N-His-GST	27 954	48 385	77 836
pOPINE-3C-eGFP	eGFP-6His-C	28 645	49 076	x
pOPINE-3C-HALO	HALO7-6His-C	35 343	55 774	x
pOPINCDM	N-His-MBP	42 711	63 142	92 593

Table 8: Periplasmic target sequences linked to ColN P-domain (in 23 Da) in pOPIN vectors with ampicillin selection for co-expression with CNI fusions in pOPINCD vectors. Vector names refer to nomenclature used by OPPF.

Vector	Periplasmic target sequence
pOPINO	Omp A SS (co-express with CD vectors)
pOPINP	PelB SS (co-express with CD vectors)
pOPINDsbA	Dsb A SS (co-express with CD vectors)
pOPINMalE	PelB SS (co-express with CD vectors)

[8] http://www.oppf.rc-harwell.ac.uk/OPPF/protocols/cloning.jsp

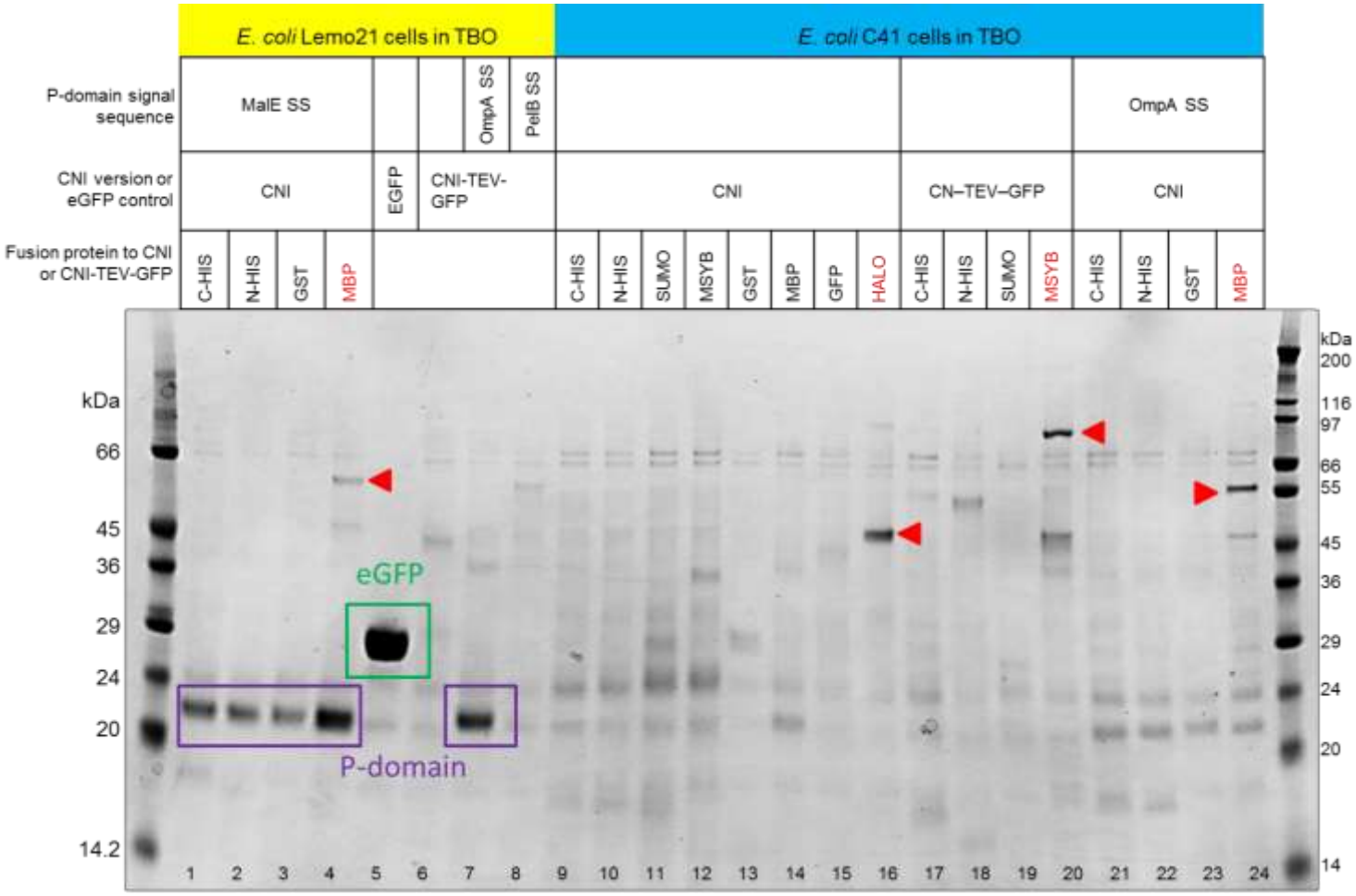


Figure 30: Example of a coomassie-stained gel showing CNI protein fusions purified by IMAC from *E. coli* Lemo21 and C41 cells, grown in TBO media. Constructs showing increased expression of correctly sized protein (Table 7) are highlighted red in the upper legend and indicated by red arrows. Where ColN P-domain (23kDa) is co-expressed, the signal sequence is given. Although ColN P-domain is coexpressed (purple square), it does not improve CNI expression. The eGFP control is indicated with a green square. See appendix for full set of expression screen gels. Numbers at the bottom indicate lane number.

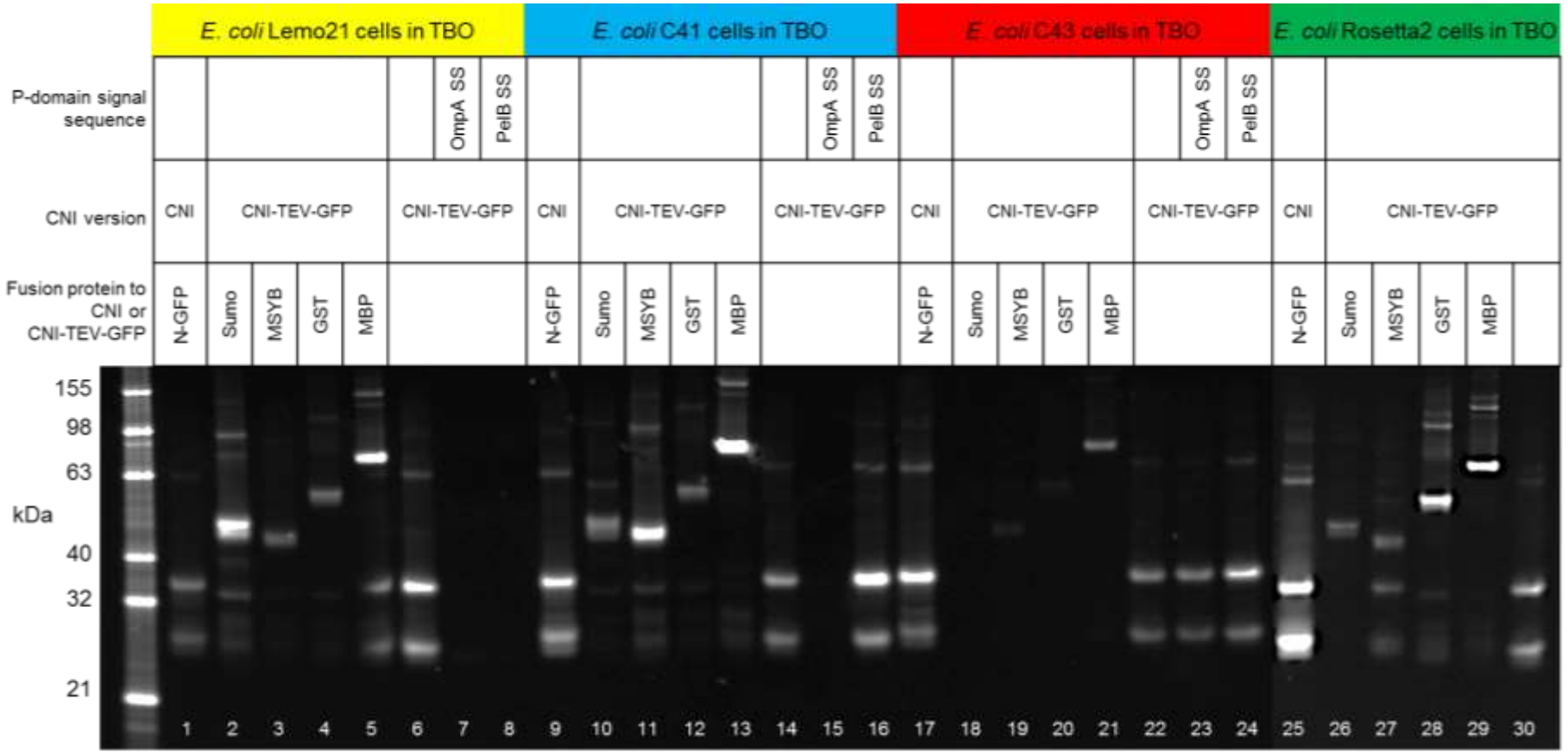


Figure 31: In-gel fluorescence of purified GFP-tagged CNI fusion proteins, expressed in TBO media. SDS-PAGE gel identifies which constructs are overexpressed, the relative level of expression in comparison to other conditions and highlights protein degradation and some higher order structures. CNI-TEV-GFP is the previously used reference construct. Constructs without additional solubilisation tags seem to break at the fusion point more easily than constructs with additional proteins. The broken off GFP-tag is around 28 kDa. Most purified proteins show dimerisation, which could indicate aggregation. Where ColN P-domain (23 kDa) is co-expressed, the signal sequence is given. ColN P-domain co-expression does not improve CNI-TEV-eGFP expression levels or stability. Numbers indicate lane number. See appendix for full set of expression screen gels.

The screen intended to find a combination which showed a considerable improvement of expression and construct stability as well as ease of subsequent fusion cleavage. Some constructs were certainly better than others, most notably Halo- and MBP-tagged fusions, while others failed to express at all. No single one condition stands out as considerably better than all the others. MBP and HALO7 linked constructs express consistently in both media and across all strains and so any of these could have been taken forward. Co-expression with P-domains does not seem to improve expression significantly. It may, however, improve CNI stability in solution but this was not further investigated. Based on the expression screen (Figure 30-Figure 31), I decided to take the pOPIN-CNI-3C-HALO7-6His construct further. It expressed consistently well across most conditions and did not show any detectable degradation products. The MBP-linked constructs showed some degradation products, which indicated that this fusion might not be as stable as the fusion with HALO7-6His. The HALO7-tag is a commercial and versatile tag, commonly used in protein immobilisation, purification and solubilisation (Los *et al.*, 2008) and in these constructs it can be cleaved by the 3C viral protease (Hedhammar *et al.*, 2006), also known under its commercial name PreScission Protease (GE Healthcare). A western blot against the polyhistidine-tag was carried out to confirm the identity of the purified protein-fusions and also to compare all strains and media used for CNI-HALO7-6His expression in the same gel (Figure 32). Figure 32A compared the expression of CNI-HALO7-6His (ca. 55 kDa) to CNI-TEV-eGFP-6His (ca. 48 kDa), the expression construct used prior to this screen. Arrows indicate the size where the purified protein fusion is expected. CNI-HALO7-6His is clearly expressed at a higher level than CNI-TEV-eGFP-His, when expression strains and media are the same, and so this screen has achieved signifcant improvement in expression level. The lower bands detected on the western blot indicate that some of the fusion protein has broken down, resulting in a band at 35 kDa, the size of the HALO7-tag alone. However, the Coomassie stained gel indicates that the breakdown is exaggerated by the western blot. The subsequent western blot compares the expression of CNI-HALO7-6His in different strains and media. A Coomassie-stained SDS-PAGE gel (Figure 32B) and the western blot (Figure 32C) show IMAC purified protein fusions. The western blot confirms the identity of the purified proteins but also highlights some degradation and aggregation and differences in expression levels between strains and media. Based on this western blot it was decided to select CNI-HALO7-6His expressed in the C41

strain and grown in TBO media for further scale-up in Newcastle and perform cleavage with 3C PreScission Protease.

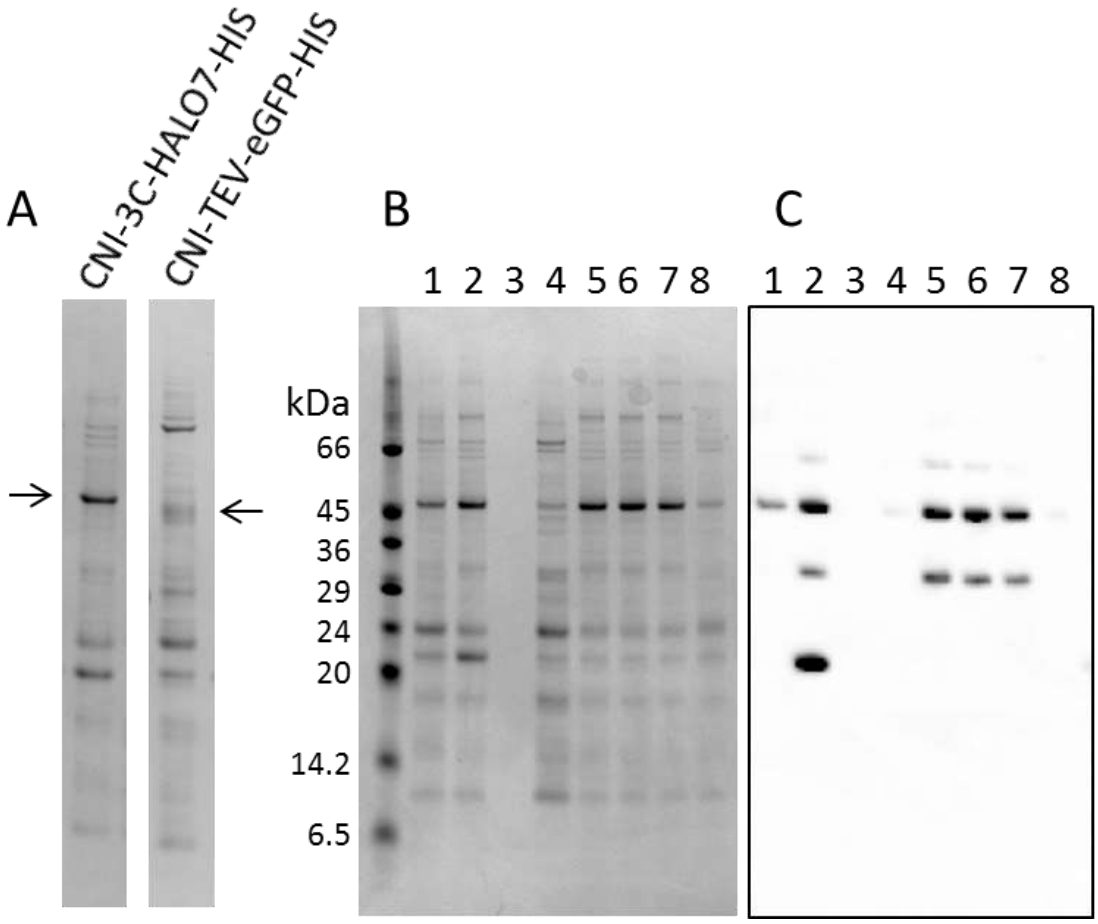


Figure 32: The protein fusion CNI-3C-HALO7-6His has a higher expression level than the previously used fusion CNI-TEV-eGFP-6His. (A) Expression comparison between CNI-3C-HALO7-6His and CNI-TEV-eGFP-6His in Coomassie-stained SDS-PAGE gel. (B, C) The CNI-3C-HALO7-6His was overexpressed in different strains and media. The Coomassie stained gel (B) and the western blot (C) show IMAC purified protein fusions. The protein fusion product is expected to be ca. 55 kDa, the HALO7-tag alone at ca. 35 kDa. Lanes: 1 - 4: PB, 5-8: TBO. 1 & 5: Lemo21, 2 & 6: C41, 3 & 7: C43, 4 & 8: Rosetta 2.

3.7.2 Expression scale up in Newcastle

Following successful small scale expression at OPPF, the expression was scaled up to 4x 500 ml in flasks in Newcastle. The fusion protein was solubilised in 1 % Tween and 1 % DM in a sodium phosphate buffer (50 mM NaP, 300 mM NaCl, pH = 7.4) and purified using IMAC. The western blot (Figure 33) shows the full length fusion at ca. 50 kDa and a broken off free HALO7-tag at ca. 35 kDa. However, it also shows some higher order structures. To examine these size exclusion chromatography was carried out with elution 2 of figure 3 (Figure 34). It showed that no very large aggregates such as those which do not enter the gel are present in solution, as there is no protein peak at the column void volume (ca. 6 ml). The largest protein peak (maximum around 11 ml) is however poorly resolved. A second smaller, peak is present, probably representing the broken off HALO7-6His-tag, eluting around 14.5 ml. Any higher order structures appearing in the SDS-PAGE gel are therefore likely to be gel artefacts, which are commonly caused in α-helical proteins by SDS (Sagne *et al.*, 1996; Kunji *et al.*, 2008). A Coomassie-stained SDS-PAGE gel of the fractions 6-15 from the size exclusion chromatography and the column load shows that only two very faint bands for fractions 8 and 9 are visible, which correspond to the broken off HALO7-6His-tag (size ca. 34 kDa).

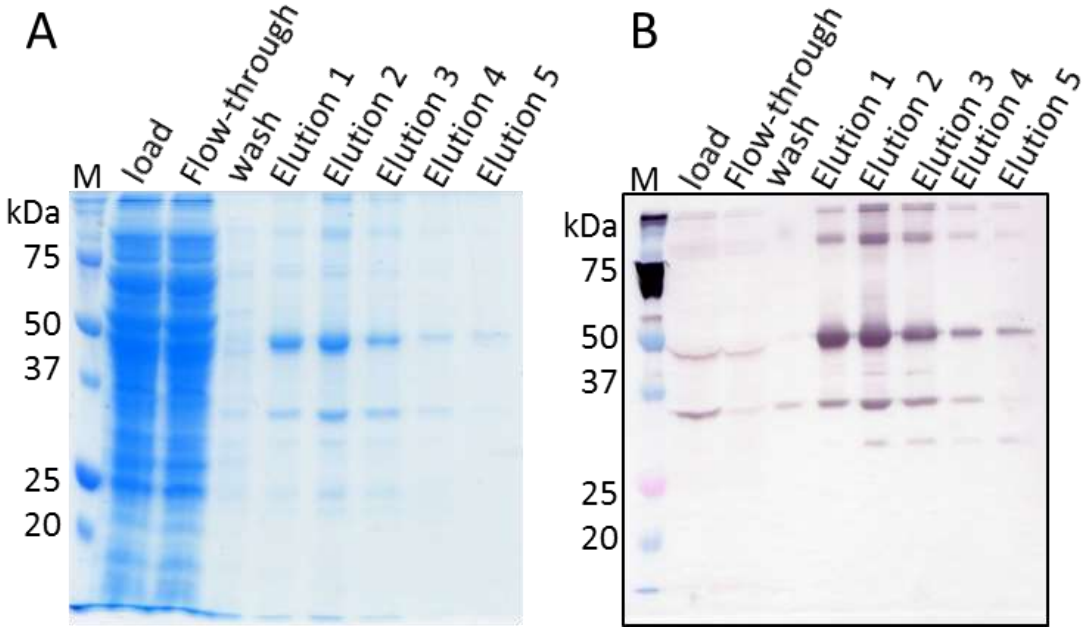


Figure 33: CNI-3C-HALO7-HIS was purified from C41 using 1 % DM and 1 % Tween. Cells were grown in TB and induced at $OD_{600\ nm}$ = 0.5 with IPTG overnight at 37 °C. The Coomassie stained gel (A) and western blot (B) show purification fractions.

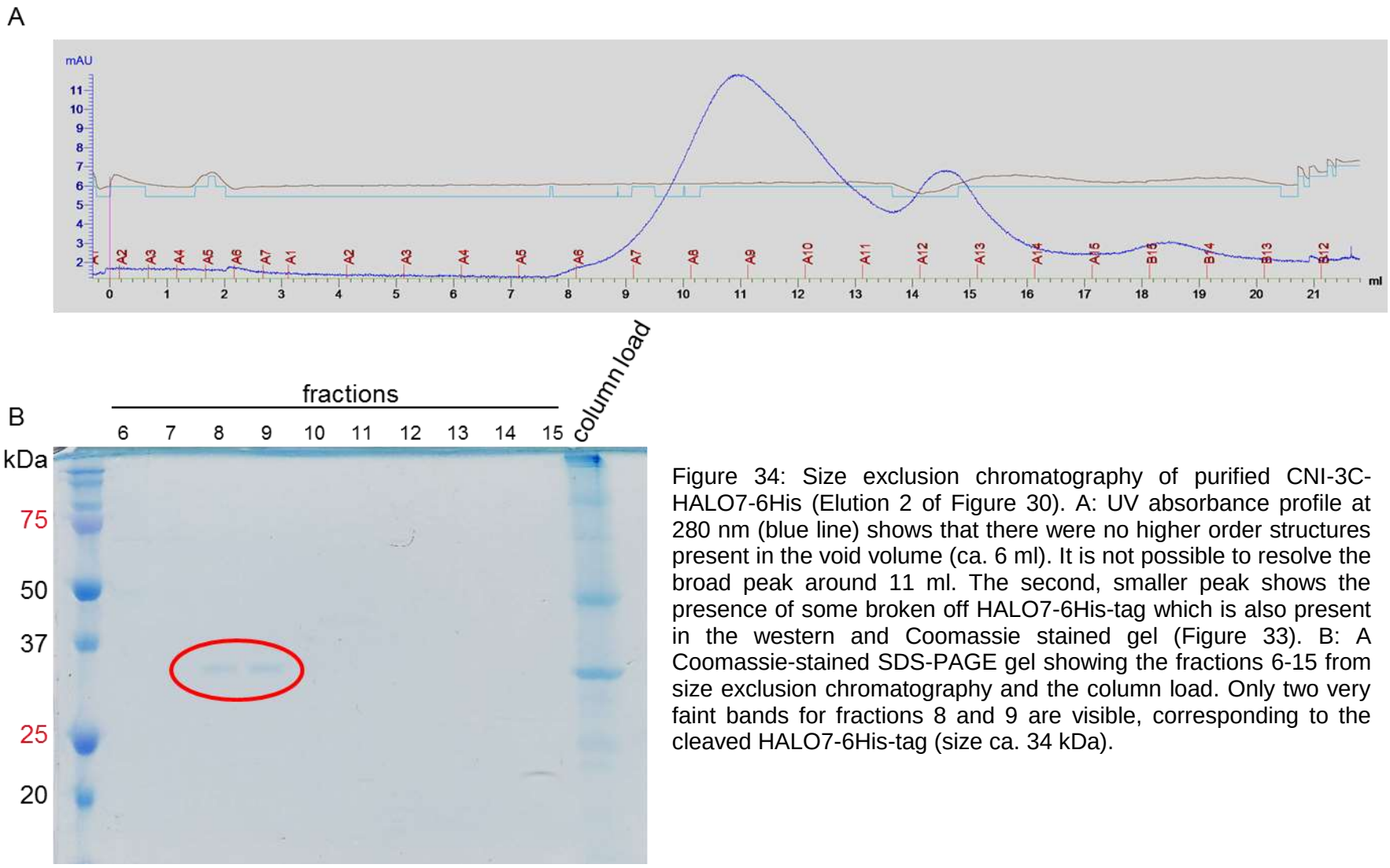


Figure 34: Size exclusion chromatography of purified CNI-3C-HALO7-6His (Elution 2 of Figure 30). A: UV absorbance profile at 280 nm (blue line) shows that there were no higher order structures present in the void volume (ca. 6 ml). It is not possible to resolve the broad peak around 11 ml. The second, smaller peak shows the presence of some broken off HALO7-6His-tag which is also present in the western and Coomassie stained gel (Figure 33). B: A Coomassie-stained SDS-PAGE gel showing the fractions 6-15 from size exclusion chromatography and the column load. Only two very faint bands for fractions 8 and 9 are visible, corresponding to the cleaved HALO7-6His-tag (size ca. 34 kDa).

3.7.3 CNI-3C-HALO7-HIS activity

In order to show that the successfully purified fusion protein CNI-3C-HALO7-6His was also active *in vivo*, a killing assay was carried out (Figure 35). C41 cells expressing CNI-3C-HALO7-6His where grown to the same optical density as sensitive cells (empty vector control), treated with different concentrations of ColN, diluted to reduce the ColN amount and grown overnight at 37 °C. CNI-3C-HALO7-6His is able to protect otherwise sensitive C41 cells, raising resistance levels from 50 pM ColN to 500 nM.

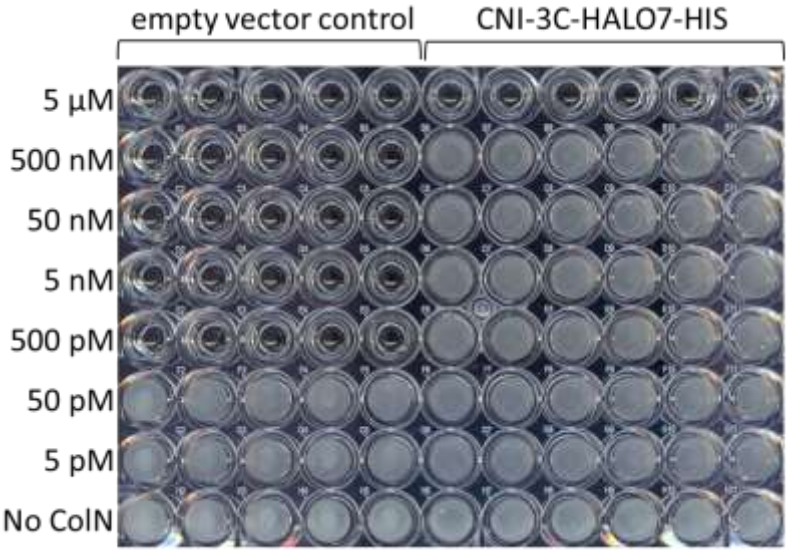


Figure 35: The killing assay in a plate shows that CNI-3C-HALO7-6His is able to protect otherwise sensitive C41 cells, raising resistance levels from 50 pM ColN to 500 nM. This picture was taken 20 h after incubation at 37 °C. Clear wells indicate no growth, cloudy wells show cell growth.

3.7.4 3C protease cleavage

In preparation for PreScission Protease 3C cleavage, CNI-HALO7-6His was purified using a purification buffer which matched the optimum cleavage buffer (Figure 36). It differs from the previous buffer by being a Tris buffer rather than a sodium phosphate buffer and containing 1 mM DTT and 1 mM EDTA and not containing Tween20 (Table 9). This buffer composition for CNI is the result of a long process of optimisation and achieves the cleanest purification result for CNI-HALO7-6His so far, resulting in minimal aggregation and degradation (Figure 36). Adjusting the buffer during purification also makes a dialysis step before cleavage redundant. The 350 mM imidazole is removed during subsequent steps, where the PreScission Protease, a 3C protease GST-fusion is removed using a GST-binding resin and the cleaved

Halo-tag is removed with a Halo-link resin. Alternatively, a His-tagged version of 3C protease could be purified in house and both the cleaved HALO7-HIS-tag as well as the His-tagged 3C protease removed using reverse IMAC, in which the desired protein is found in the flow through.

Table 9: Purification buffer components for the initial stages of purification are optimised for subsequent cleavage with PreScission protease. This buffer composition achieves the cleanest purification result for CNI-HALO7-6His with minimal aggregation and degradation.

Cell disruption buffer	Protein extraction buffer	Column Wash buffer	Protein elution buffer	Dialysis/Desalt buffer
50 mM Tris-HCl, pH = 7.5	50 mM Tris-HCl, pH = 7.5	50 mM Tris-HCl, pH = 7.5	50 mM Tris-HCl, pH = 7	50 mM Tris-HCl, pH = 7
1 M NaCl	1 M NaCl	300 mM NaCl	300 mM NaCl	300 mM NaCl
50 mM Imidazole	50 mM Imidazole	100 mM Imidazole	350 mM Imidazole	
Proteinase inhibitor, DNAse, RNAse, lysozyme				
			1 mM EDTA	1 mM EDTA
			1 mM DTT	1 mM DTT
	1 % DM	0.1% DM	0.1% DM	0.1% DM
	10% glycerol	10% glycerol	10% glycerol	10% glycerol

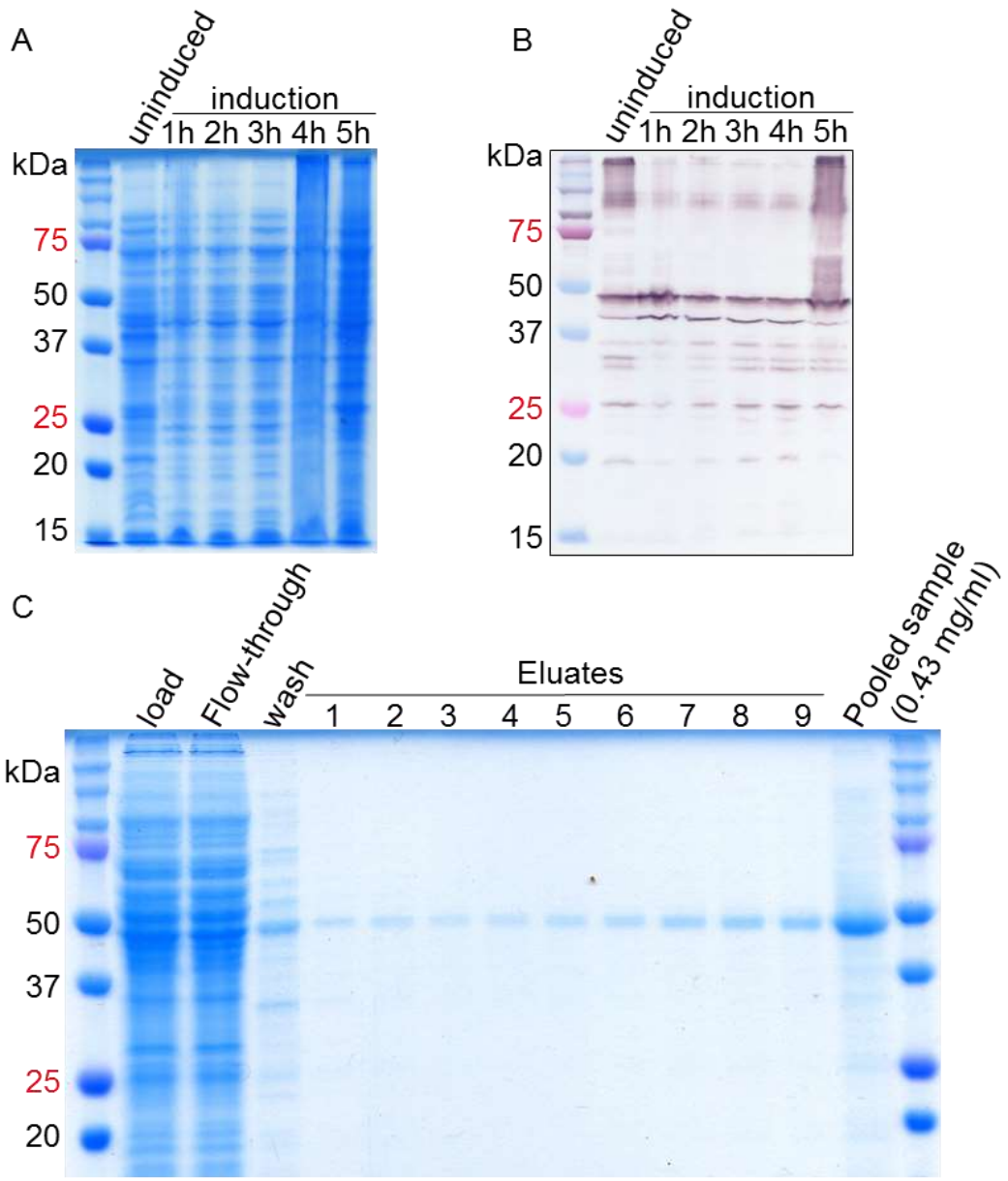


Figure 36: CNI-Halo-6His purification prior to PreScission Protease cleavage. (A) Coomassie-stained SDS-PAGE gel and (B) western blot of CNI-HALO7-6His expression in C41 cells. (C) Coomassie-stained SDS-PAGE gel showing samples from steps in CNI-HALO7-6His purification. The pooled sample was used for PreScission Protease cleavage.

According to the manufacturer one unit of PreScission Protease will cleave more than 90 % of 100 µg of a test GST-fusion protein in a buffer containing 50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, pH = 7.0 at 5 °C in 16 h. Here, the sodium chloride concentration was doubled to match the concentration previously used for solubilisation and DM and glycerol were added (Table 9). The increased sodium chloride concentration and the addition of DM and glycerol as well as the conformation of the protein fusion may affect protease efficiency. Two experiments were carried out to optimise the protease amount and the cleavage temperature

(Figure 37). One unit was sufficient to cleave most fusion protein, however, some uncleaved fusion protein still remains even with 5 units of PreScission protease which is a large excess. An increased temperature had no effect on the cleavage efficiency. There is no detectable aggregation which is a good improvement from the thrombin cleavage attempts (Figure 26). The cleaved HALO7-6His-tag is clearly visible, however, a clear band for cleaved CNI is missing. The double band highlighted with the red arrow could represent the cleaved CNI at around 20.5 kDa. Why there are two bands was not immediately clear.

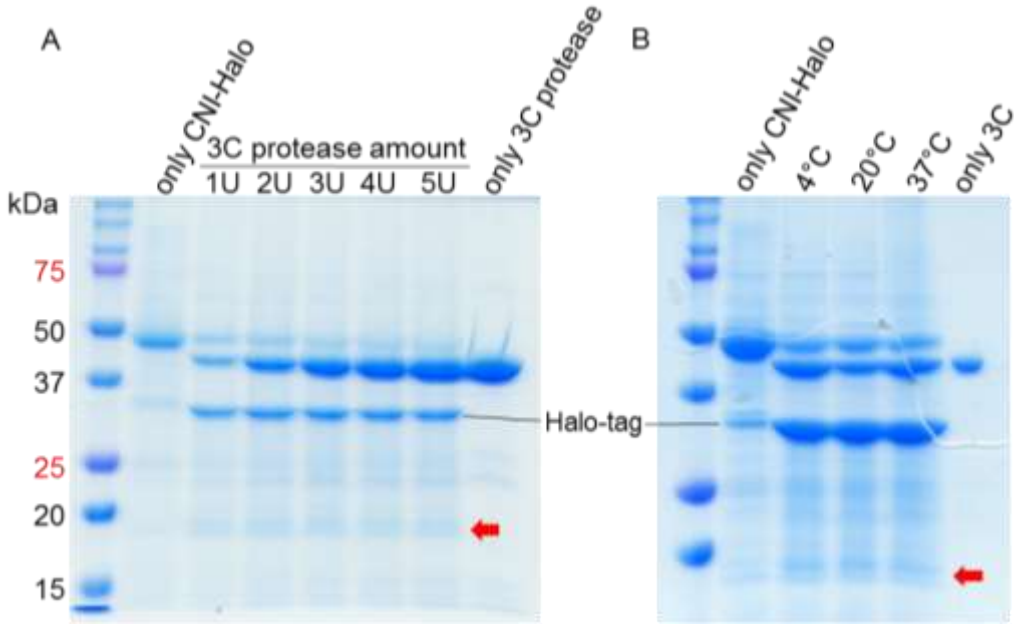


Figure 37: PreScission Protease 3C cleavage optimisation of the HALO7-6His-tag in Coomassie-stained 12 % SDS-PAGE gel. (A) Cleavage was performed for 16 h at 4 °C with 8.6 µg of CNI-HALO7-6His-tag in each condition and different amounts of protease, 1-5 units. (B) Cleavage was performed for 16 h at 4 °C, 20 °C, 37 °C, with 11.2 µg of CNI-HALO7-6His-tag and 1 unit of PreScission Protease in each condition. 1 PreScission unit will cut 100 µg of protein in optimum conditions. Bands that probably represent the cleaved immunity protein are highlighted with a red arrow.

3.7.5 LC-MS of CNI-HALO7-6His-tag and its cleavage products

Membrane proteins are not heat-denatured before they are loaded into SDS-PAGE gels to avoid aggregation and so some secondary structure and tertiary structure can remain intact leading to multiple bands or smears in SDS-PAGE gels if some of the protein partially unfolds. It is also not uncommon for proteins to degrade following purification or during protease cleavage.

To confirm that the protein fusion and the cleavage products are full length and the cleavage products have also been analysed with LC-MS (Figure 38-Figure 40). The

protein samples for LC-MS have been purified and cleaved using the procedure outlined in 3.7.4. LC-MS has been carried out as a commercial service by Dr Joseph Gray, Pinnacle lab, Newcastle University[9]. The predicted sizes are 55 541.6 Da for the CNI-HALO7-6His fusion protein, 34 270.2 Da for the cleaved HALO7-6HIS-tag and 21 289.4 Da for the cleaved CNI (20 431.4 Da for CNI and 858 Da for the PreScission Protease cleavage site). The purified protein fusion as well as the cleaved products are detected and shown to be full length, even though they may not run at their predicted sizes in the SDS-PAGE gels (Rath *et al.*, 2009). This is a common characteristic of highly hydrophobic proteins such as membrane proteins. For the full length protein-fusion and cleaved CNI, two additional protein species are detected, both 700 Da higher than the predicted sizes. The origin of these additional species is unclear. Random contaminations are possible but unlikely due to the exact difference in size for both the full length protein-fusion and the cleavage product. Post-translational modifications of CNI or membrane lipid attachment are more likely but have not been investigated further at this point. Looking back at some of the purifications carried out the full length fusion protein after purification appears as a double band in SDS-PAGE gels, western blots and in in-gel fluorescence images but its visibility seems to be concentration dependent (examples in: Figure 22: pellet lane, Figure 23: acid modified samples, Figure 25A: lanes 7-9, Figure 29: in-gel fluorescence image, Figure 36: lane 8). With this characteristic occurring across so many different expression and purification procedures over the years it is likely that this double band, which according to LC-MS is spaced 700 Da apart, is a genuine characteristic of CNI. It is possible that this is a post-translational modification or a non-annular lipid but this has not been investigated beyond this point.

3.8 Conclusions

- An overexpression system has been developed for CNI fused to a HALO7-6His which yields an amount of fusion protein that can be used for further study of the tagged version or CNI by itself after protease cleavage.
- From the detergents and buffers screened, DM in a Tris buffer was the most successful in solubilising and stabilising the protein fusion.

[9] http://www.ncl.ac.uk/camb/internal/pinnacle/contact/

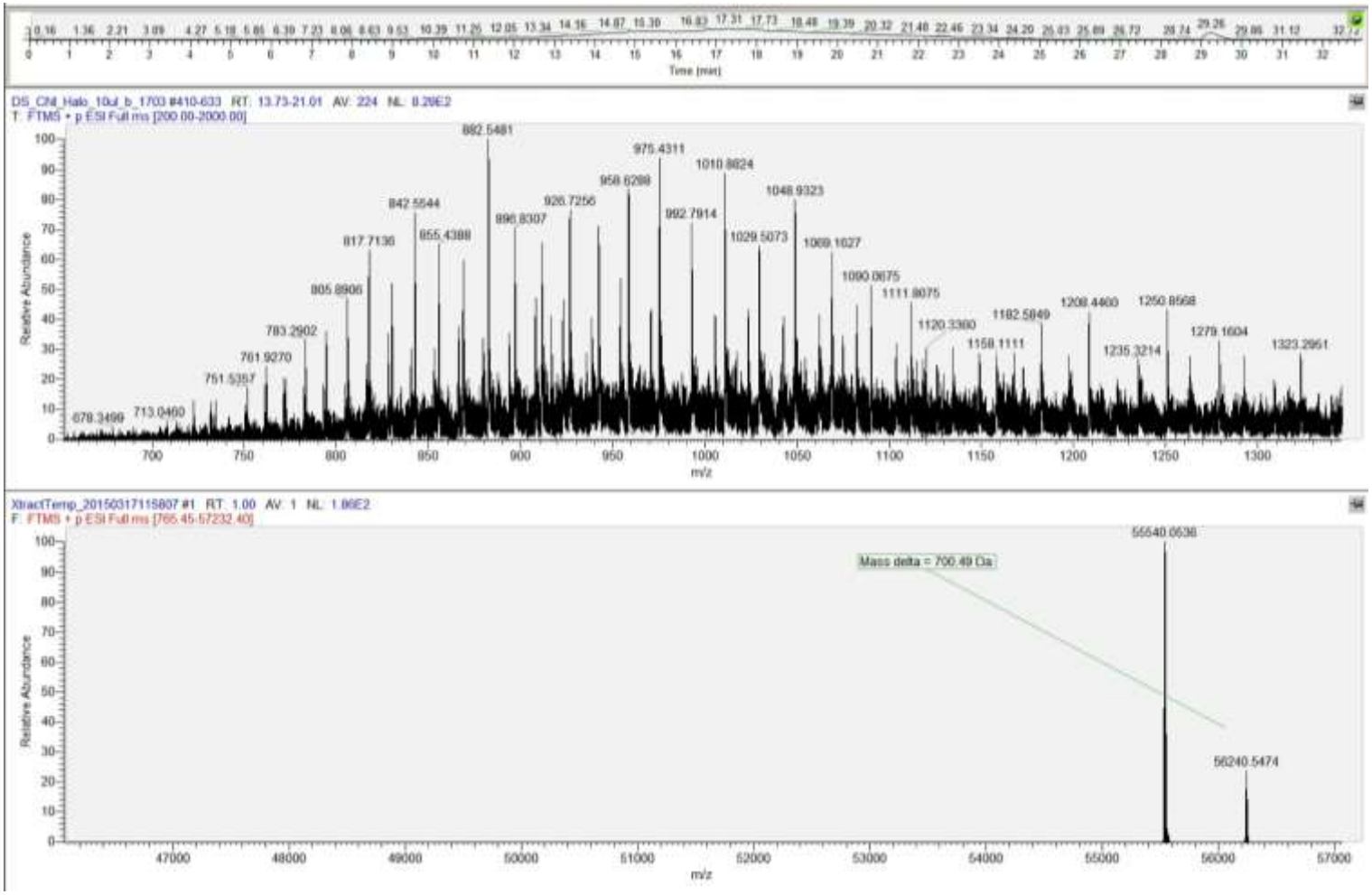


Figure 38: LC-MS analysis of Full length CNI-HALO7-His fusion protein. The mass spectrometry measurement matches the calculated size of 55 541.6 Da based on the amino acid sequence very closely, proving that the purified protein is the full length protein. A second protein species is detected at 56240. 55 Da, 700.49 Da higher than the full length protein. All measured values are accurate to ± 1 Da.

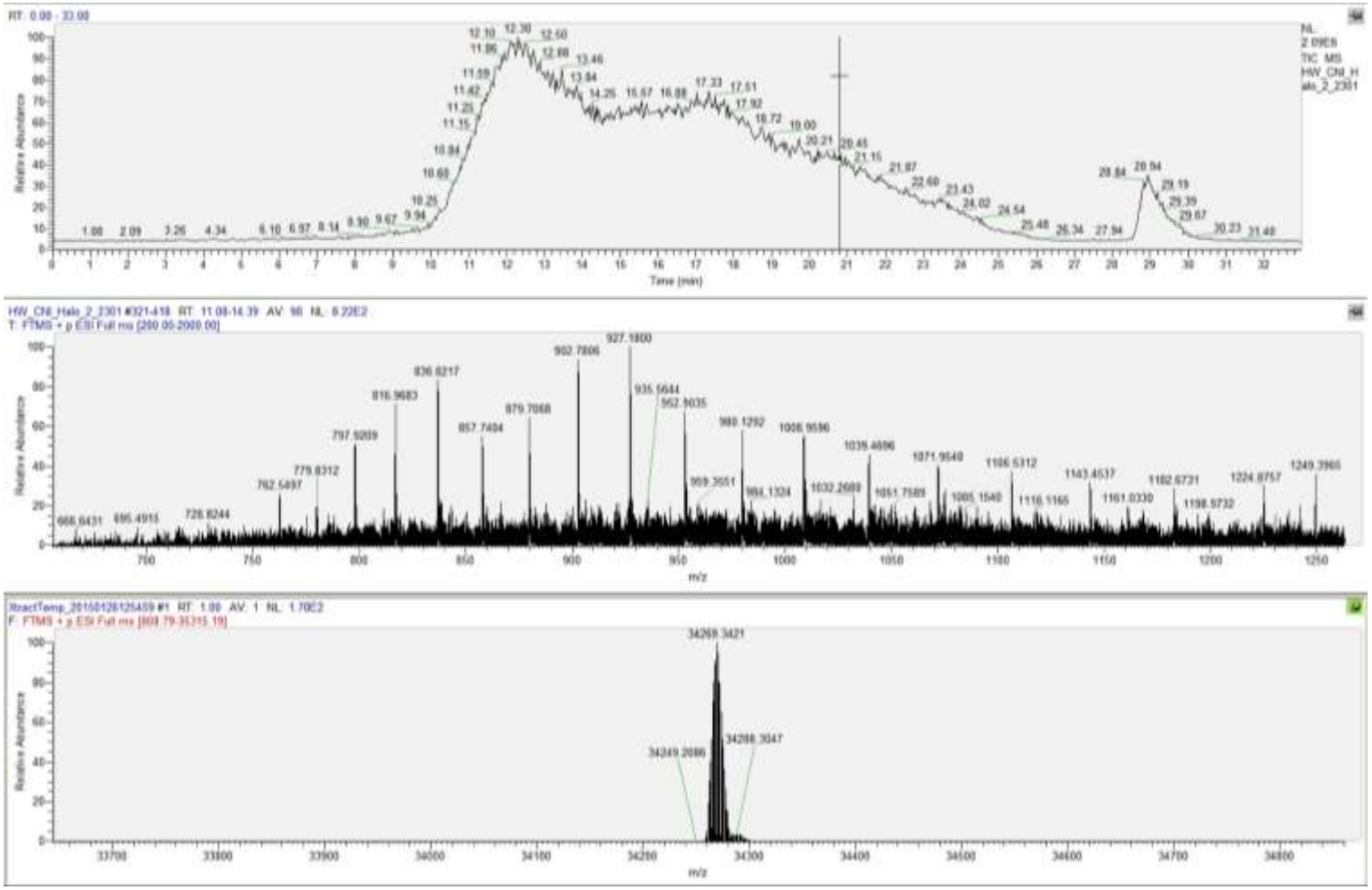


Figure 39: LC-MS analysis of CNI-HALO7-His cleavage products: The HALO7-6HIS-tag. One product is the Halo-6His-tag with a calculated size of 34270.2 Da which is matched closely by this LC-MS analysis confirming that cleavage was successful and was carried out at the correct site in the protein fusion. All measured values are accurate to ± 1 Da.

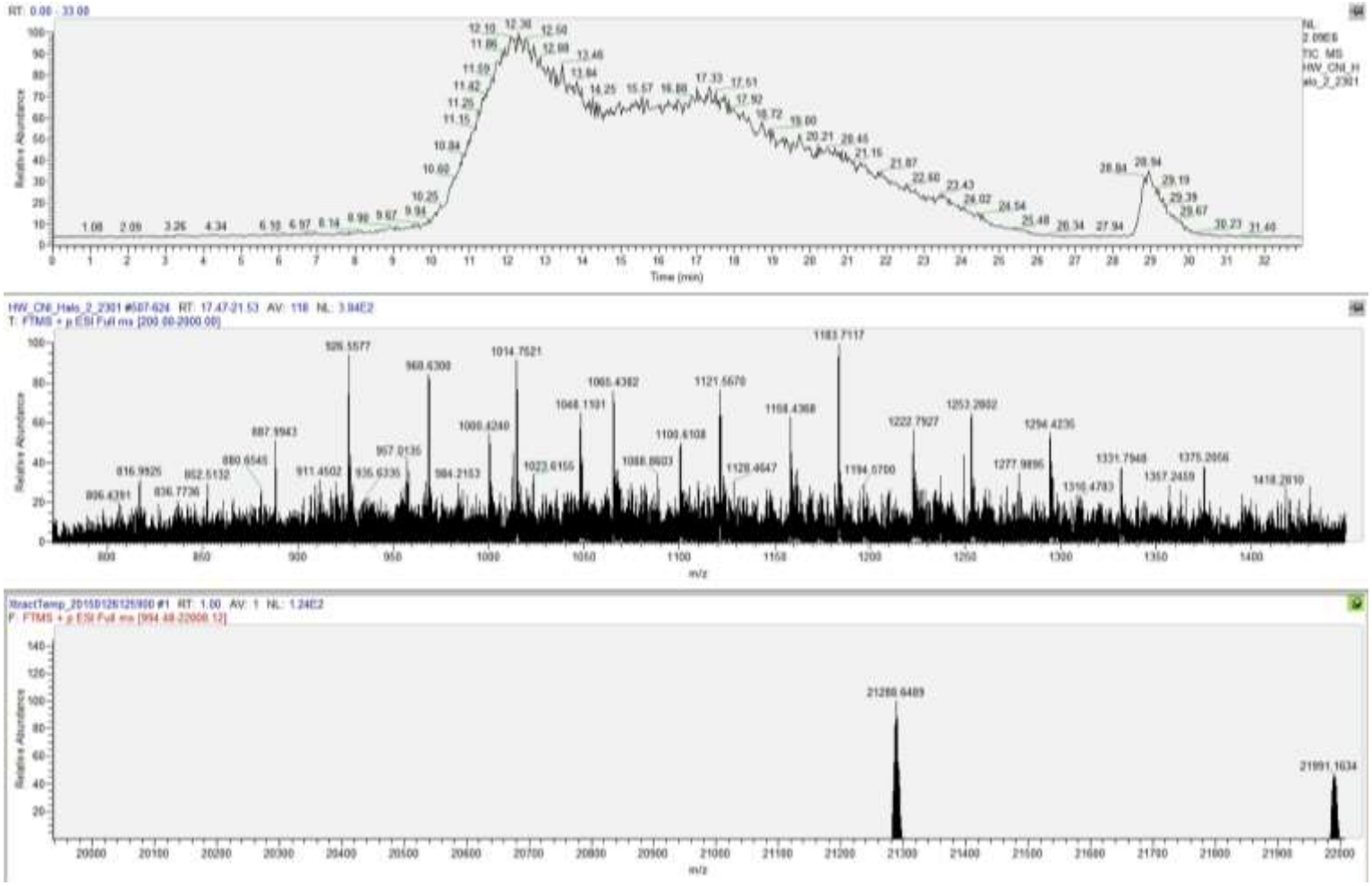


Figure 40: LC-MS analysis of CNI-HALO7-His cleavage products: CNI. The calculated size of CNI is 20431.4 Da, but including the amino acids left over after cleavage at the 3C cleavage site the calculated size is 21289.4 Da, which is matched closely by the lower peak detected by LC-MS. The higher peak is 702.5 Da larger than the full length cleavage product. Allowing for ±1 Da measuring error this matches the difference observed in Figure 38.

Chapter 4. *In vivo* and *in vitro* Colicin N immunity protein interaction with Colicin N

4.1 Introduction

Due to their hydrophobicity immunity proteins of pore-forming colicins are difficult to study *in vitro*. Hence, most published research (Chapter 1: Introduction) has focused on mutagenesis to define an interaction site *in vivo,* assuming a direct interaction between Colicin N and its immunity protein. While some immunity proteins are highly specific and only protect against one colicin, others provide immunity to several (Chapter 1: Introduction). The Colicin N immunity protein (CNI) is only known to provide immunity to Colicin N (ColN), while the Colicin A immunity protein (CAI) is only known to provide immunity to Colicin A (ColA). Inspired by mutagenesis studies of other Colicin immunity proteins (see Chapter 1: Introduction), this chapter focuses on the determination of an ColN-CNI interaction site using mutagenesis as well as attempts to recreate the interaction *in vitro* using pull-down assays and a surface plasmon resonance assay.

Presumably, if residues crucial for protein-protein interaction are mutated, CNI will no longer protect against ColN. However, mutagenesis always involves the risk of changing structurally important residues, therefore abolishing function through structural collapse rather than the change of functionally important sidechains. Therefore, usually, mutated proteins must be shown to be structurally intact, using biophysical techniques to assess stability such as CD, DSC, fluorescence spectroscopy and others. Since it is very difficult to purify stable CNI (chapter 3: CNI overexpression), here, structural integrity must be shown *in vivo*. Therefore, instead of just abolishing activity against ColN, it was attempted to create activity against ColA. Residues involved in specificity are not necessarily the same as residues involved in activity as seen in previous mutagenesis studies (chapter 1: introduction). A mutated immunity protein may still be able to perform its function of providing immunity against a colicin, even though it can change its specificity from Colicin N to Colicin A, for example. However, residues responsible for specificity and activity are likely to be located in the same region and given that structural analysis is difficult, analysing specificity rather than activity it is a good alternative. Hence, the following mutagenesis study aimed to establish the residues responsible for CNI specific

interaction with ColN, which may or may not be the same residues which are responsible for its activity.

Predominantly based on the hypothesis proposed by (Song and Cramer, 1991), this chapter tests if individual helices of CNI are involved in its interaction with Colicin N. It was planned to mutate whole helices first and then narrow down the interaction site to individual residues once the general area was known. The predicted helices of CNI and its closest homologue Colicin A immunity protein (CAI) have been swapped (Figure 41). CNI may cope well with mutagenesis of functional domains if the tertiary structure is flexible, but the secondary structure is fixed. The secondary structure is predicted to be predominantly α-helical. A diagram of the predicted topology is shown in Figure 42.

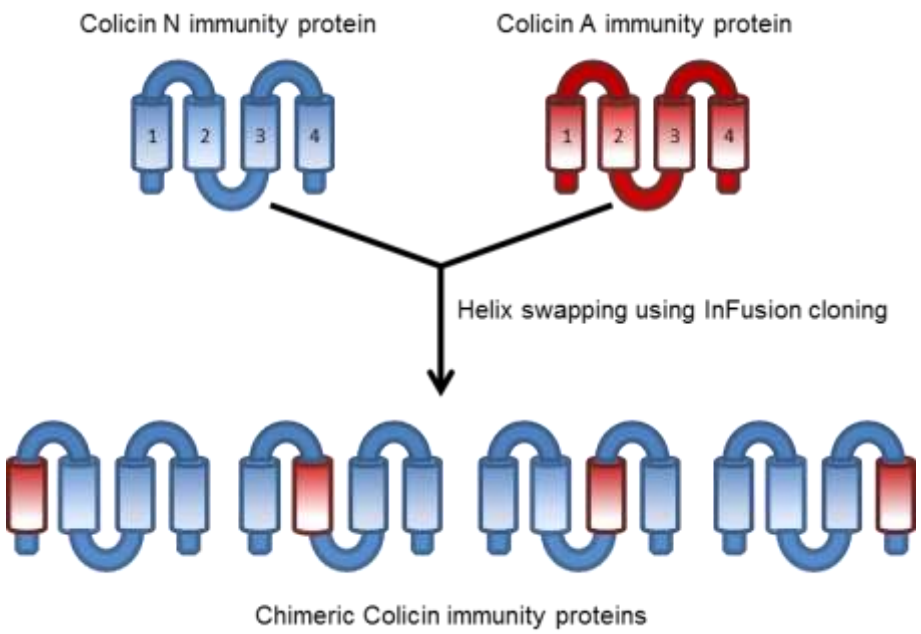
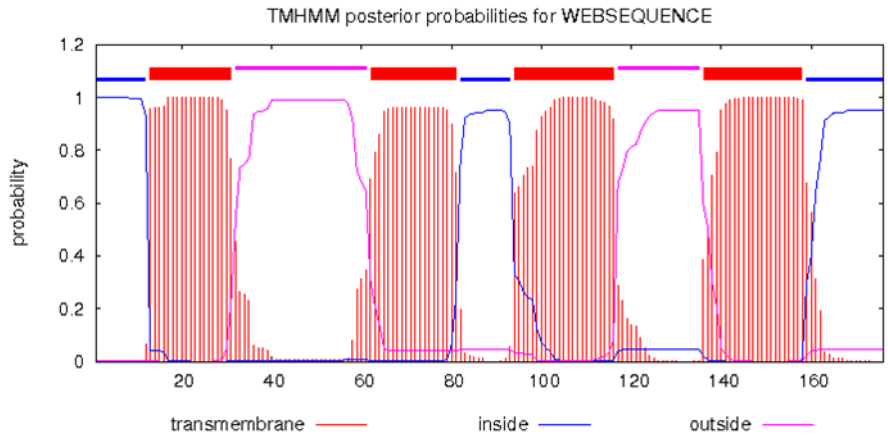


Figure 41: Helix swap between Colicin N and Colicin A immunity proteins. Helices are numbered 1-4 from N- to C-terminus. Each helix of CNI is replaced by the homologous helix of CAI. Ligation sites were chosen in regions of high sequence identity.

A



B

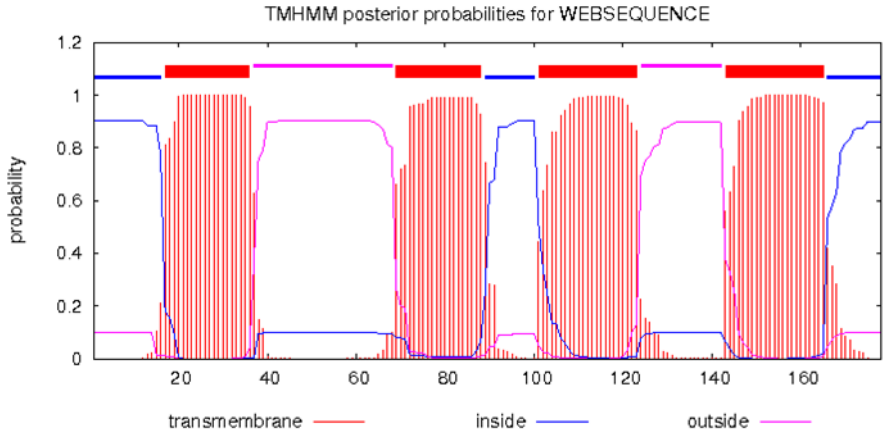


Figure 42: Predicted topology for CNI (A) and CAI (B). The TMHMM tool predicts 4 transmembrane helices spanning the *E. coli* inner membrane. N- and C- termini are located in the cytoplasm.

The sequence alignment of Colicin A immunity protein (NCBI Reference Sequence: WP_008323617.1) and Colicin N immunity protein (Stroukova and Lakey, 2015, in press, see appendix) are shown in Figure 43. Underlined amino acids indicate the predicted helical structure in both proteins (Krogh *et al.* (2001), Sonnhammer *et al.* (1998), online prediction tool[10]), while coloured regions mark swapped regions (yellow: helix 1, green: Helix 2, blue: helix 3, magenta: helix 4). The number indicates which helix of CNI was swapped for a helix from CAI. Overall, the proteins share a high nucleotide and amino acid sequence similarity. There is no region which stands out with special features, like a particularly high or low sequence identity, although, the region around helix 3 shows the least similarity.

Swapping a whole helix rather than individual residues not only increases the chance of swapping specificity but may be advantageous because a whole functional unit is

[10] http://www.cbs.dtu.dk/services/TMHMM/TMHMM2.0b.guide.php

swapped, increasing the likelihood of maintaining structural integrity. In homologues several residues must be changed in the same region to see a change in functionality and specificity and no single residue seems to be more important than the others (Lindeberg and Cramer, 2001; Smajs *et al.*, 2006; Smajs *et al.*, 2008).

The regions for swapping were chosen based on the predicted helix structure in both proteins (underlined in Figure 43). Maintaining the same residues at the ends of each region allowed a neat transition from one protein to the other. Furthermore, it was aimed to make the transition point in more flexible regions, like loops and turns, to allow tertiary structure folding and adjustment in case of steric hindrance. The predicted loops and turns connecting the helices were partially excluded from mutagenesis where possible based on previous research, which suggested that their role in protein – protein interaction is unlikely (Song and Cramer, 1991; Smajs *et al.*, 2006; Smajs *et al.*, 2008). The effect of mutagenesis was tested using two killing assays, a spot test assay on agar plates and a survival assay in liquid culture.

A

```
CAI             ATGATGAATGAACACTCAATAGATACGGACAACAGAAAGGCCAATAACGCATTGTATTTA 60
CNI             ---ATGGAT--------ATAAAAGACAGAAATAAGATAT-CAAAAAA---AATATCATTC 45
                   ***.**        *::*.* **.**.*:.***:*  *.**:**   *:*.*.:**.

CAI             TTTATAATAATCGGATTAATACCATTATTGTGCATTTTTGTTGTTTACTACAAAACGCCA 120
CNI             AGTCTTCTGCTCTTACTTTCCCCATTCGCATTAATATTTTTCAGTTATAATAATGCACCA 105
                : *.*:.*..**  * *:: .*****.  .* .**:*** * . *** :* **:.*.***

CAI             GACGCTTTACTTTTACGTAAAATTGCTACAAGCACTGAGAATCTCCCGTCAATAACATCC 180
CNI             ATACCACTCCT----CGAAAAAAT-CATCGCATACCTA-TCCCTACCAGGATTTCATTCA 159
                .:. *: *.**    **:****:* *::*... **  * :. **.**.  *:*:..:**.

CAI             TCCTACAACCCATTAATGACAAAGGTTATGGATATTTATTGTAAAACAGCGCCTTTCCTT 240
CNI             TTAAACAACCCGCCCCTAAGCGAAGCATTCAATCTCTATGTTCATACAGCCCCTTTAGCT 219
                * .:*******.  ..*.* ..*.* ::* .**.* ***  *.*:***** *****.  *

CAI             GCCTTAATACTATACATCCTAACCTTTAAAATCAGAAAATTAATCAACAACACCGACAGG 300
CNI             GCAATCAGCTTATTCATATTCACACACAAAGAATTAGAGTTAA-AACCAAAGTCGTCACC 278
                **.:*.* . ***:***. *.**. : ***.:.: *.*.**** .*.***.. **:**

CAI             AACACTG-TACTTAG-ATCTTGTTTATTAAGTCCATTGGTCTATGCAGCAATTGTTTATC 358
CNI             TCTGCGGGCACTAAAGATATTAACTCCTTTCACTATT-CTTTATATATCCATGATATACT 337
                :. .* *  ***:*. **.**.: *. *:: :* ***  * ***. * *.** .*:**

CAI             TATTCT-GCTTCCGAAATTTGAGTTAACAACAGCCGGAAGGCCTGTCAGATTAATGGCC 417
CNI             GTTTCTTGCTAAC-TGACACAGAACTAACCTTGTCATCAAAAACATTTGTATTAATGTCA 396
                 :**** ***:.* :.* : :**. ****.: . *.  **...*: * . ******* *.

CAI             ACCAATGACGCAACACTATTGTTATTTTATATTGGTCTGTACTCAATAATTTTCTTTACA 477
CNI             AAAAACGATCTGTTTTTGTCTTTTTTCTATATAACACTATATATTGGGATATATATATTC 456
                *..** **   .: : *.*  **:** *****:. :**.** : :. .**:*: :*:: .

CAI             ACCTATATCACGCTATTCACACCAGTCACTGCATTTAAATTATTAAAAAAAAGG------ 531
CNI             ACATATTTGTACTTTTGGTTCCTTATAGGAACATATAAGCTATTTACCAGGGGAGGAATA 516
                **.***:* :.  *:*  : .* :.*.. :.***:***. ****:*..*...*.

CAI             --CAGTAA---- 537
CNI             CTCCGTCGACAC 528
                  *.**..
```

B

```
CAI             MMNEHSIDTDNRKANNALYLFIIIGLIPLLCIFVVYYKTPDALLLRKIATSTENLPSITS 60
CNI             -MDIK----DRNKISKKISFSLLLLLSPFALIFFSYNNAPIPLLEKIIAY--LSLPGFHS 53
                 *: :    *..* .: : : ::: * *:  **. * ::* .** : **    .**.: *

CAI             SYNPLMTKVMDIYCKTAPFLALILYILTFKIRKLINNTDRNTVLRSCLLSPLVYAAIVYL 120
CNI             LNNPPLSEAFNLYVHTAPLAAISLFIFTHKELELKPKSSPLRALKILTPFTILYISMIYC 113
                  ** :::.:::* :***: *: *:*:*.*  :*  ::.   .*:     .::* :::*

CAI             FCFRNFELTTAGRPVRLMATNDATLLLFYIGLYSIIFFTTYITLFTPVTAFKLLKKRQ-- 178
CNI             FLLTDTELTLSSKTFVLMSKNDLFLSFFYITLYIGIYIFTYLYFWFLIGTYKLFTRGGIL 173
                * : : *** :.:.. **:.**  * :*** **  *:: **: ::  : ::**:.:

CAI             ---
CNI             RRH 176
```

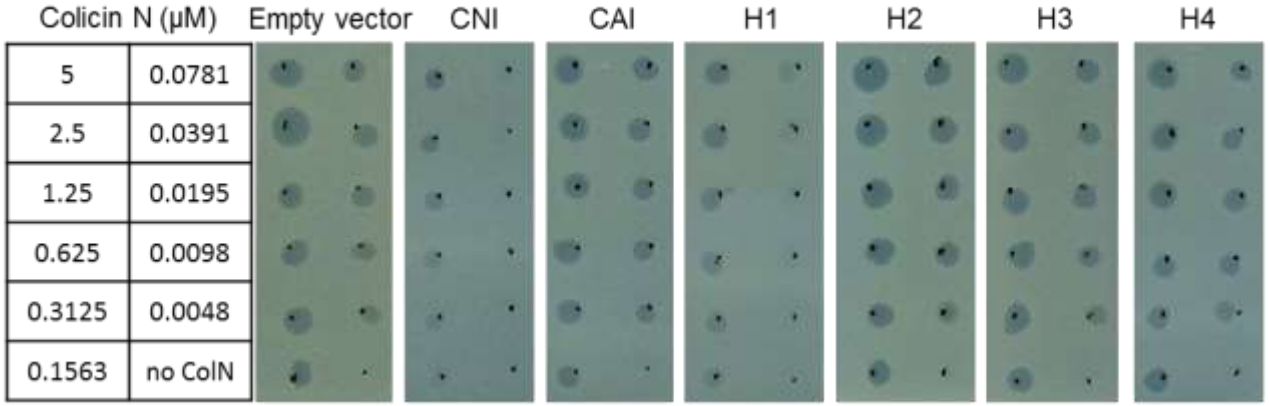
Figure 43: CAI and CNI alignment using ClustalW. (A) Nucleotide sequence. (B) amino acid sequence. Highlighted text marks swapped parts; helix 1: yellow, helix 2: green and red, helix 3: blue and red, helix 4: magenta. Nucleotides and amino acid highlighted red are the endpoint of Helix 2 in the Helix 2 mutant as well as the beginning of Helix 3 in the Helix 3 mutant. Underlined text indicates predicted α-helical structure based on TMHMM (Krogh, 2001).

4.2 Results and Discussion

4.2.1 Assessing the mutant activity using a spot test assay

Different concentrations of ColN and ColA were applied as 2 µl drops onto bacterial lawns of C41 *E. coli* cells expressing WT CNI-HALO7-6His, WT CAI-HALO7-6His, Helix 1, 2, 3 and 4 mutants of CNI-HALO7-6His and an empty vector control (Figure 44). Clearing zones demonstrate sensitivity. The empty vector control is equally sensitive to ColN (48 nM and above) and ColA (195 nM and above). WT CNI-HALO7-6His protects against ColN (up to 195 nM) but not ColA, while WT CAI protects against ColA (up to 1.563 µM) but not ColN. There is no cross-immunity to either colicin. Neither mutant is immune to ColA, so a change in specify from ColN to ColA was unsuccessful. If the mutants were immune to ColA, they would provide a protection level higher than the empty vector control, ideally similar to WT CAI-HALO7-6His. Helix mutants H2, H3 and H4 are also no longer resistant to ColN. The H1 mutant, however, still protects against ColN. This means, either the first 38 amino acids are not needed for specificity to ColN or their function can be replaced by the corresponding amino acids from CAI. It is possible that helix 1 of CAI provides enough structural support for the rest of the protein to still function. To investigate if those amino acids are important for structure and function, one would have to truncate the protein.

A



B

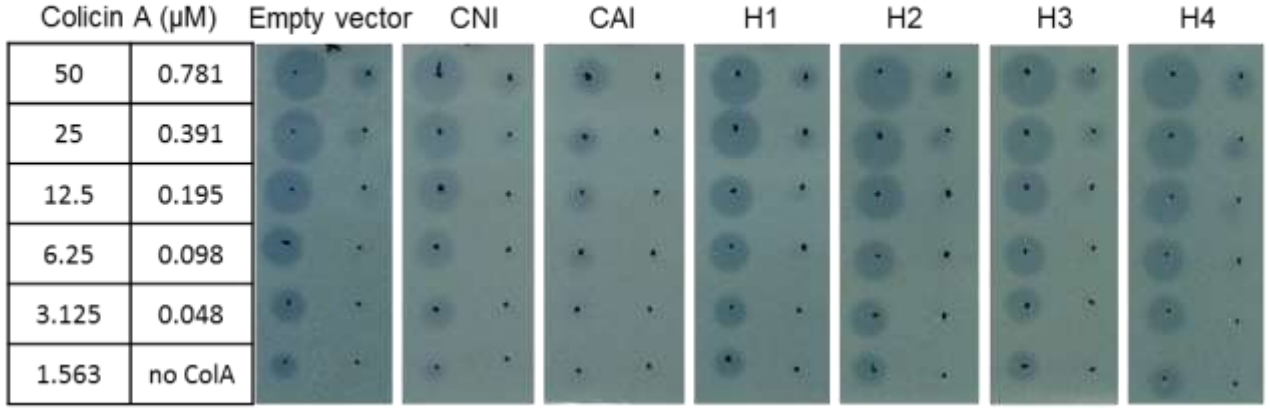


Figure 44: Immunity protein and mutant resistance to Colicin A and Colicin N. Colicin N and A were applied as 2 μl drops onto bacterial lawns of C41 cells. Helix mutants 2, 3, 4 and the empty vector control are not active against ColN and ColA. CNI and the helix 1 mutant are active against ColN, but not ColA. CAI is active against ColA, but not ColN. There is no change in specificity due to mutagenesis.

4.2.2 Measuring activity with the liquid culture cell survival assay

In comparison to the spot test assay, there is a more refined and quantitative way to establish culture viability through the length of the lag time. All cultures start out with approximately the same amount of cells at the beginning, estimated through optical density at $OD_{600\ nm}$ = 0.5. When colicin is added, the number of cells is greatly reduced, depending on antibiotic potency and concentration. Cultures are subsequently diluted 1:100, in order to dilute out the cell debris and the colicin to a hundredfold lower concentration. The optical density subsequently measured in the wells can almost entirely be attributed to living cells. The more cells are killed, the smaller is the inoculum for subsequent cultures during dilution and the longer it takes for visible growth, shown through well turbidity, to take off again. Therefore, the lethal impact of colicin addition is directly related to the culture lag time. The immunity protein activity is shown through a shorter lag time in comparison to the empty vector control (Figure 45).

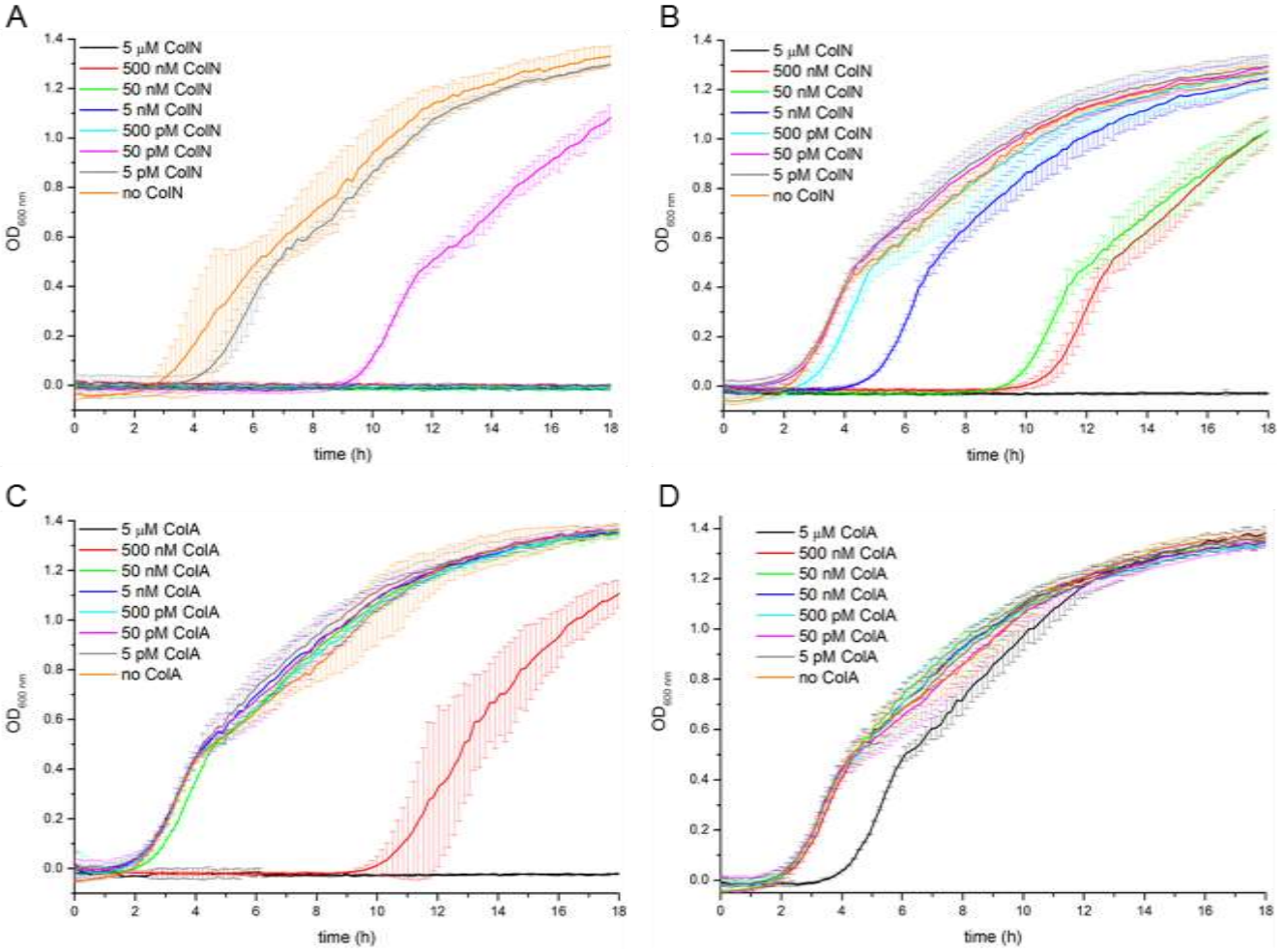


Figure 45: The effect of ColN and Col A concentration on the lag time of sensitive and immune C41 *E. coli* cultures. The higher the colicin concentration, the longer the lag time. A: Empty vector control with ColN. B: CNI-HALO7-6His with ColN. C: Empty vector control with ColA. D: CAI-HALO7-6His with ColA

.

Empty vector control C41 cells do not have a known protection mechanism but as with other toxic compounds, Colicin N and A require a certain minimal concentration or a certain molecule to cell ratio to be effective. For pore-forming colicins, theoretically, one molecule per cell is enough (Cascales *et al.*, 2007); practically, it should be around 600 molecules (Johnson *et al.*, 2013). Therefore, even empty vector control cells are resistant to Colicin N and A up to a certain concentration. The lethal impact is more clearly displayed through increased lag time than through well turbidity. Even the lowest concentration of 5 pM ColN is toxic to sensitive cells of the empty vector control and increases lag time by ca. 1h. For the immune cells a significantly higher concentration of 500 pM is needed to achieve a similar effect. For ColA higher concentrations are needed to impact on cell growth and the effect is less gradual. A concentration of 500 nM is needed to affect culture growth for the empty vector control, while 5 µM inhibits immune cultures expressing CAI-Halo7-6His.

4.2.3 Measuring the mutants' activity with the liquid culture cell survival assay

To confirm the results obtained in the spot test assay an activity assay was also carried out in liquid culture. For each mutant, Figure 46 (Helix 1 mutant), Figure 47 (Helix 2 mutant), Figure 48 (Helix 3 mutant) and Figure 49 (Helix 4 mutant) show the results of the liquid killing assay. All cultures were grown to the same optical density and exposed to different concentrations of ColN and ColA for 30 min, before being diluted 1:100 and grown in a plate reader. An exposure time of 30 min is more than adequate as ColA and ColN have been shown to act within minutes (Johnson *et al.*, 2013). Well turbidity and length of lag time are indicators of culture viability. Turbid wells mean culture growth, while clear wells mean all cells have been killed. Hence, simply looking at the plate gives a good qualitative indication about culture viability. Pictures were taken at the end of each run after 18 h.

All mutants were compared to a positive and a negative control. The empty vector control is a negative control as it is sensitive to both Colicin N and Colicin A. CNI-HALO7-6His is a positive control against ColN activity and CAI-HALO7-6His is a positive control against ColA activity. Including CNI-HALO7-6His in ColA treatments and vice versa shows that there is no cross immunity. CNI does not protect against ColA and CAI does not protect against ColN.

For the helix 1 mutant 500 nM ColN is toxic, just like for the positive control CNI-HALO7-6His (Figure 46). There is no difference in lag times between the mutant and

wild-type, which means that they are equally active. The empty vector control and CAI-HALO7-6His have no protective mechanism against ColN and so are sensitive to 500 pM and above. At 50 pM ColN, the lag times of CAI-HALO7-6His producing cells and the empty vector control are 6 h longer than those of CNI-HALO7-6His and the Helix 1 mutant, demonstrating a dramatic lethal effect of ColN on these cultures. The helix 1 mutant remains active against ColN which can be interpreted in two ways. Either, the N-terminal part of the protein is not essential for CNI-function or Helix 1 of CAI can replace Helix 1 of CNI without a change in activity and this part of the protein is not essential in the specific interaction between CNI and ColN. A truncation of the protein, so that the first 38 amino acids are missing, may be a way to answer this question.

The helix 1 mutant is as sensitive to ColA as the empty vector control and CNI and so does not protect against ColA. A transfer in specificity from ColN to ColA was unsuccessful. Helix mutants 2, 3, and 4 are completely inactive; they do not protect against ColN or ColA. They are sensitive to 500 pM ColN and 5 µM ColA, just like the negative controls. CAI is active against ColA and shows protection against 5 µM ColA. None of the mutants was protective against ColA, which means that immunity could not be changed from ColN to ColA. Probably, mutagenesis has detrimentally affected protein structure, rendering the protein inactive, but without a biophysical analysis of the purified protein it is difficult to prove that.

Lag times for the controls are consistent for all liquid cultures, demonstrating a high level of consistency and reproducibility for the method. Differences between growth curves are significant, showing either a qualitative result, dead or alive, or a huge difference in lag time of up to 6 h. Assuming a culture doubling time of 20 – 30 min, this is a highly significant result. Although growth curves for all wells were recorded, here only curves showing the biggest difference between the cultures are shown. Curves shown here are an average of 2-3 wells.

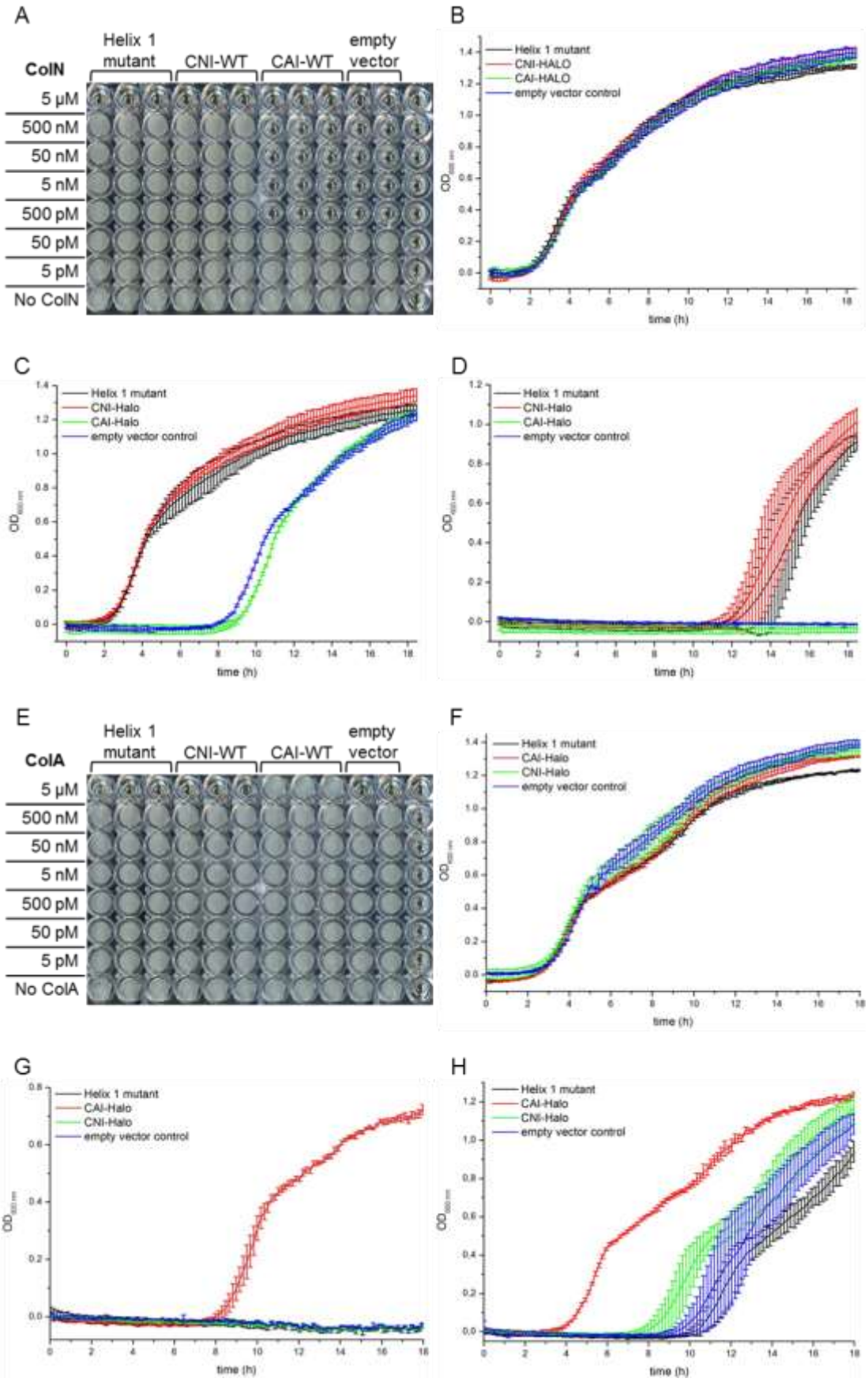


Figure 46: Helix 1 mutant is resistant to ColN but not ColA. Images of 96 well plates after 18 h with ColN (A) and ColA (E). Culture growth without ColN (B) and ColA (F). Culture growth with 500 nM ColN (C) and ColA (G). Culture growth with 5 µM ColN (D) and ColA (H).

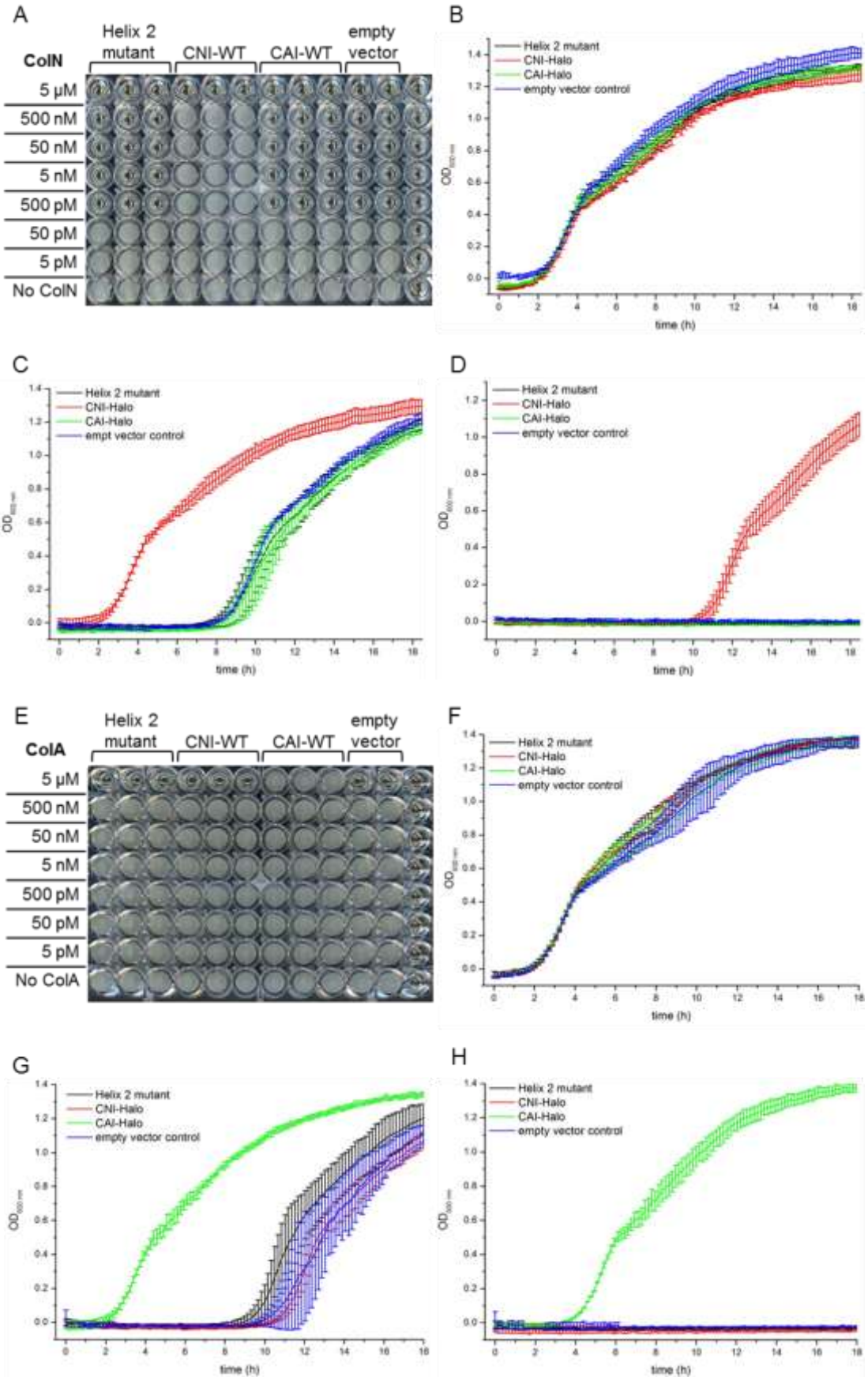


Figure 47: Helix 2 mutant is not resistant to ColN or ColA. Images of 96 well plates after 18 h with ColN (A) and ColA (E). Culture growth without ColN (B) and ColA (F). Culture growth with 500 nM ColN (C) and ColA (G). Culture growth with 5 µM ColN (D) and ColA (H).

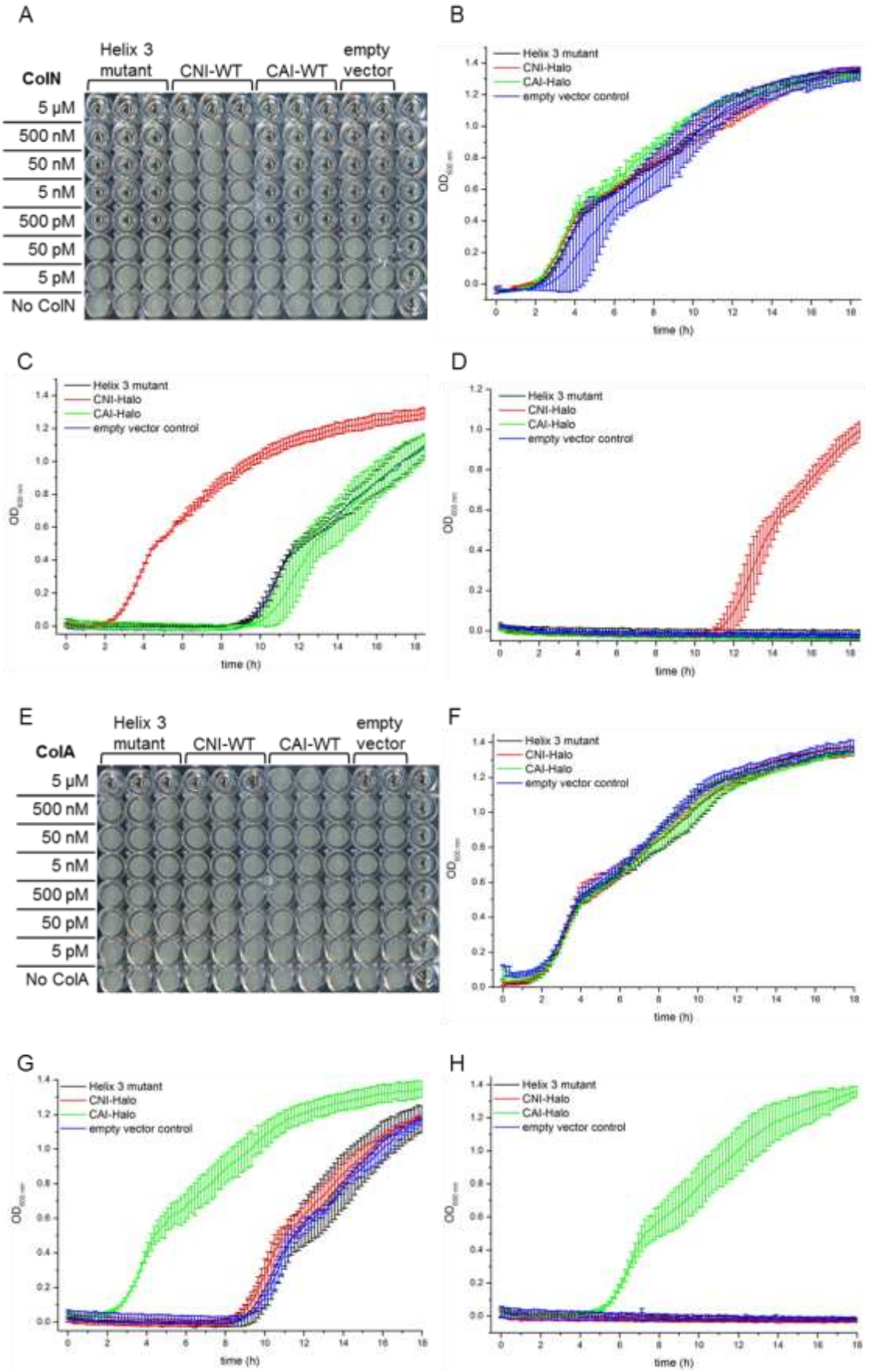


Figure 48: Helix 3 mutant is not resistant to ColN or ColA. Images of 96 well plates after 18 h with ColN (A) and ColA (E). Culture growth without ColN (B) and ColA (F). Culture growth with 500 nM ColN (C) and ColA (G). Culture growth with 5 µM ColN (D) and ColA (H).

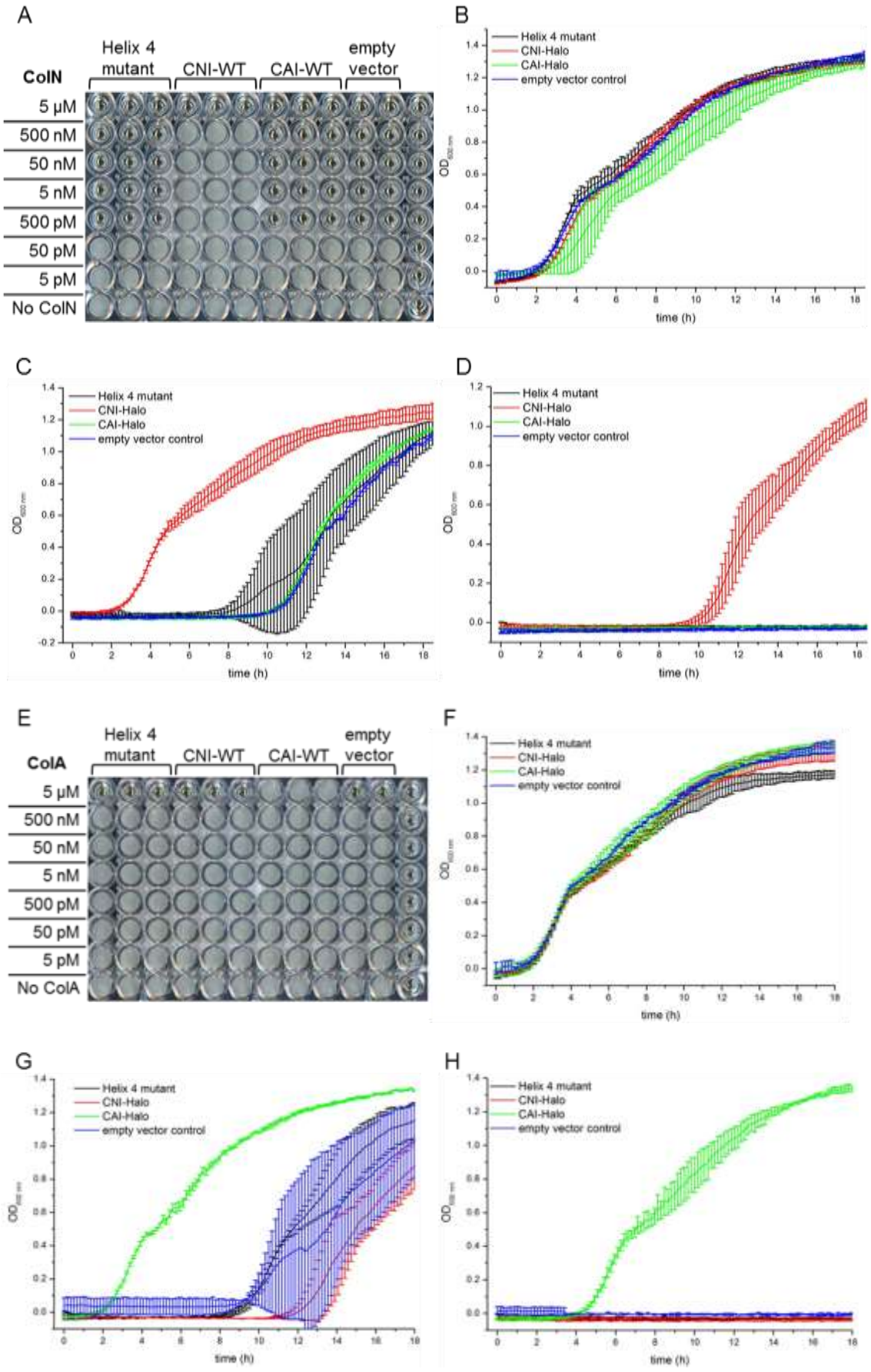


Figure 49: Helix 4 mutant is not resistant to ColN or ColA. Images of 96 well plates after 18 h with ColN (A) and ColA (E). Culture growth without ColN (B) and ColA (F). Culture growth with 500 nM ColN (C) and ColA (G). Culture growth with 5 µM ColN (D) and ColA (H).

4.2.4 Mutant overexpression detection by western blot

To show mutant proteins expression, a western blot with antibodies probing for the C-terminal polyhistidine-tag was carried out. It shows that CNI-HALO7-6His, CAI-HALO7-6His, Helix 1 mutant and Helix 4 mutant are expressed, while Helix 2 and 3 mutants are not. The promoters and genes of Helix 2 and 3 mutants were re-sequenced and confirmed as correct. The lack of expression must be caused by a problem in transcription or translation. To distinguish if genes are transcribed but not translated, RNA from these cells could be reverse transcribed and amplified with gene specific primers in a PCR using cDNA. If Helix 2 and 3 mutants are misfolded and unstable, they may be degraded quickly after translation. Degradation products would be expected in this case.

CAI-HALO7-6His and Helix mutants 1 and 4 are expressed but are clearly more degraded than CNI-HALO7-6His. It is unclear why they appear less stable than CNI-HALO7-6His and it is unclear whether this an artefact of SDS-PAGE or truly represents the state of the protein inside the cell. In Zhang *et al.* (2010), western blots of CAI show no degradation. Here, in agreement with Zhang *et al.* (2010), there is some dimerisation of CAI. The Helix 4 mutant seems to dimerise more than CNI-HALO7-6His and the Helix 1 mutant, where only minimal dimerisation can be observed. The mutation in Helix 4 does not add further cysteines for disulphide bonds to the protein and so this dimerisation may be based on hydrophobic helix-helix interaction. The visible degradation products can be attributed to the HALO7-6His-tag (ca. 34 kDa), because the polyhistidine is attached to its C-terminal. As CNI has no N-terminal tag it is not possible to visualise it with the anti-Histag antibody.

The degradation may not be a problem for cell immunity as CAI-HALO7-6His and the Helix 1 mutant are still active (Figure 46) and protect against ColA and ColN, respectively. While the lack of expression may explain the lack of activity for Helix mutants 2 and 3, it does not explain the lack of activity for the Helix 4 mutant. To continue this investigation, double and triple mutants and possibly a quadruple mutant could be made to test whether it is indeed the helices which interact with ColN and not the solvent exposed loops and termini.

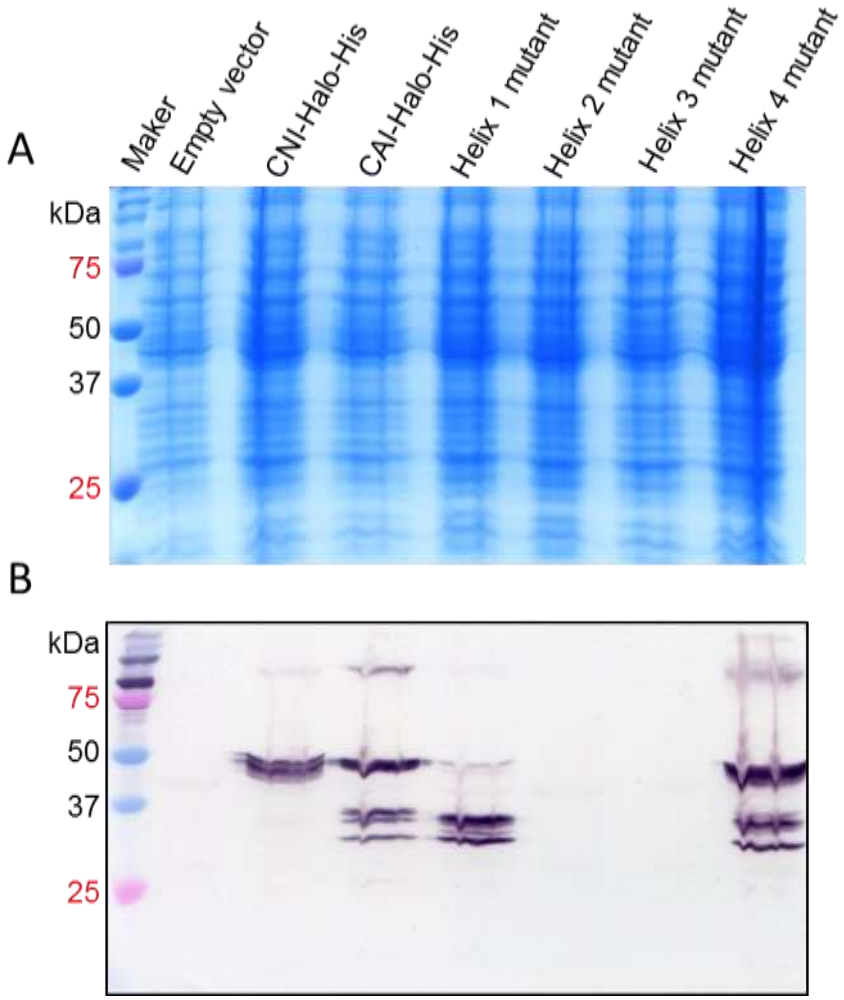


Figure 50: Mutant protein expression in C41 cells. (A) Coomassie stained SDS-PAGE of unboiled C41 cells. (B) Western blot probing for the polyhistidine-tag. CNI-HALO7-6His, CAI-HALO7-6His, Helix 1 mutant and Helix 4 mutant are expressed. Helix 2 and 3 mutants are not expressed. There is no protein expression in the negative empty vector control.

4.2.5 Cysteine residue mutagenesis, activity and purification

CNI contains only one cysteine residue at position 113 (Figure 43B). Cysteine residues often play an important role in protein stability or activity because they are able to form disulphide bonds. With the absence of further cysteines in the protein, a role in dimerisation is more likely than intramolecular interaction. However, Zhang *et al.* (2010) shows how cysteines in CAI are important for its structure and interaction with ColA. Therefore, C113 was mutated to a serine, a residue of similar size but without the ability to form disulphide bonds.

4.2.6 pWaldo-CNI^{C113S}-TEV-GFP-6His mutant activity

The ability of the C113S mutant to protect against ColN was assessed using the liquid culture cell survival assay in 96 well plates. The mutation was carried out in in the plasmid pWaldo-CNI-TEV-sfGFP-6His, which was also used for protein purification (Chapter 3). There is a 50-fold difference in minimal lethal concentration between CNI-sfGFP-6His and the empty vector control. The C113S mutant is just as active as the WT CNI-sfGFP-6His at 500 pM. At higher concentrations, however, such as 40 nM, the difference in lag time is significantly more pronounced. Here, a difference in lag time of 1.5 h was measured between WT CNI and the C113S mutant at 40 nM. There is no significant difference in growth between cultures expressing CNI-sfGFP-6His, the C113S mutant and the empty vector control when no ColN is added, demonstrating that protein expression is not toxic or a significant metabolic burden to the cells.

Similar to the effect described for CAI (Zhang *et al.*, 2010), the cysteine residue in CNI plays a role in protein activity, albeit the effect of the mutation is very small and only seen at higher ColN concentrations. It remains unclear whether the cysteine is involved in direct interaction with ColN or whether it has other, for example, structural purposes. ColN P-domain, which is the proposed site of interaction with CNI, does not contain any cysteine residues and so an interaction using a disulphide bond is not possible.

CNI does not form clear dimers on SDS-PAGE like CAI (Figure 50). Therefore, an effect of ColN on CNI dimerisation, like is described in Zhang *et al.* (2010) for CAI, was not assessed using a western blot. However, clustering upon ColN addition was investigated *in vivo* (Chapter 5). No cluster formation was observed. Cysteine

removal could affect the amount of protein in the membrane by altering expression levels or protein stability. However, upon purification (Figure 52), a difference in amount was not observed, although the mutant seemed to be less stable than the WT.

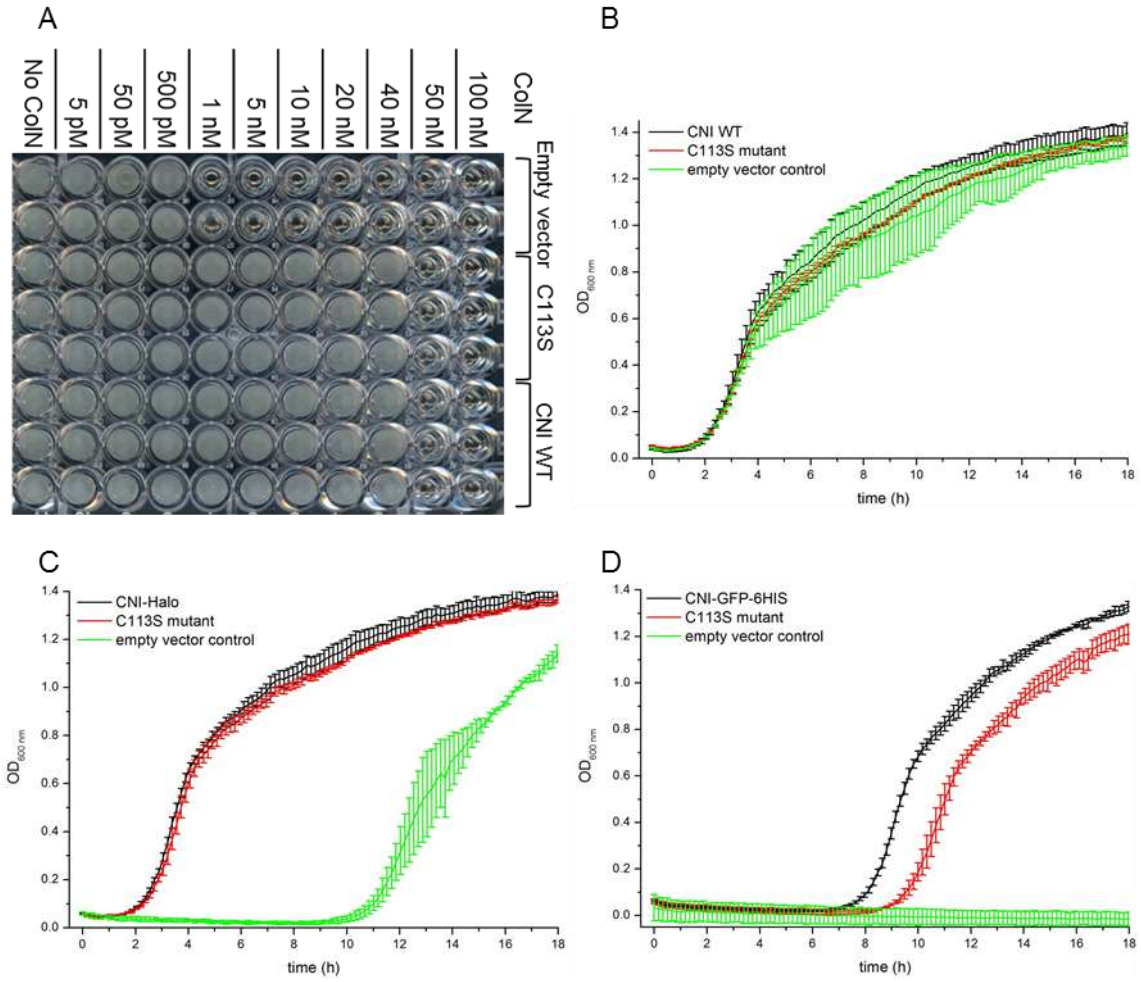


Figure 51: C113S mutant activity against ColN. (A) Image of 96 well plate after 18h with ColN. (B) Without ColN all cultures grow the same. (C) At lower concentrations, like 500 pM, both C113S mutant and WT CNI are similarly unaffected by the ColN, while the lag time of the empty vector control is significantly extended. (D) At higher concentrations, like 40 nM ColN, the mutant's lag time is significantly reduced, demonstrating a reduced culture viability.

A different reason to mutate the cysteine was to enable potentially easier purification of monomeric CNI if the potential disulphide bond leads to dimerisation. Figure 52 shows C113S mutant expression as well as C113S mutant and CNI purification. Protein expression was monitored over time using SDS-PAGE. The full size of the fusion protein is 48 kDa, but it migrates at 37kDa (Chapter 3) if not boiled because the protein is not fully unfolded during SDS-PAGE (Figure 52B). The size of GFP-6His is ca. 27 kDa, but it migrates faster, around 25 kDa, when not boiled and is

fluorescent inside the gel because SDS-PAGE does not unfold the fluorophore. The size of CNI is 20.4 kDa. As with other CNI fusion proteins, due to the low expression level, it is not clear from the Coomassie-stained gel if the protein is expressed (Figure 52A). Also, there is no evident difference in the protein band pattern when comparing the C113S mutant with WT CNI-sfGFP-6His and the empty vector control. However, following purification of both proteins (Figure 52C and D) using IMAC, a very small amount of protein can be detected in the eluates (ca. 2 mg in total from 5 l of culture).

No dimerisation of either protein is visible in the Coomassie stained SDS-PAGE gels showing purification fractions. There is also no difference in yield although there is more breakdown in the C113S sample. The in-gel fluorescence images taken with the Typhoon™ FLA 9500 biomolecular imager show fluorescent proteins excited 488 nm and emitting at 510 nm, suitable for sfGFP at around 37 kDa for the full length proteins and ca. 24k Da for the broken off GFP tag (Figure 52E and F). These images prove the presence of sfGFP–tagged CNI and C113S mutant but also show considerable amounts of break down products. It seems that the C113S mutant is already breaking down in the load and flow-through at the start of purification, while CNI-sfGFP-6HIS is largely stable until it is eluted. Similar to the helix mutants, mutating the cysteine residue of CNI seems to destabilise it and make it more difficult to purify as a single intact monomer. Notably, while the GFP of the protein is clearly visible at ca. 25 kDa in the Coomassie stained gel and identified as sfGFP in the in-gel fluorescence images, CNI (21 kDa) is not visible at all. It seems that it has been completely degraded.

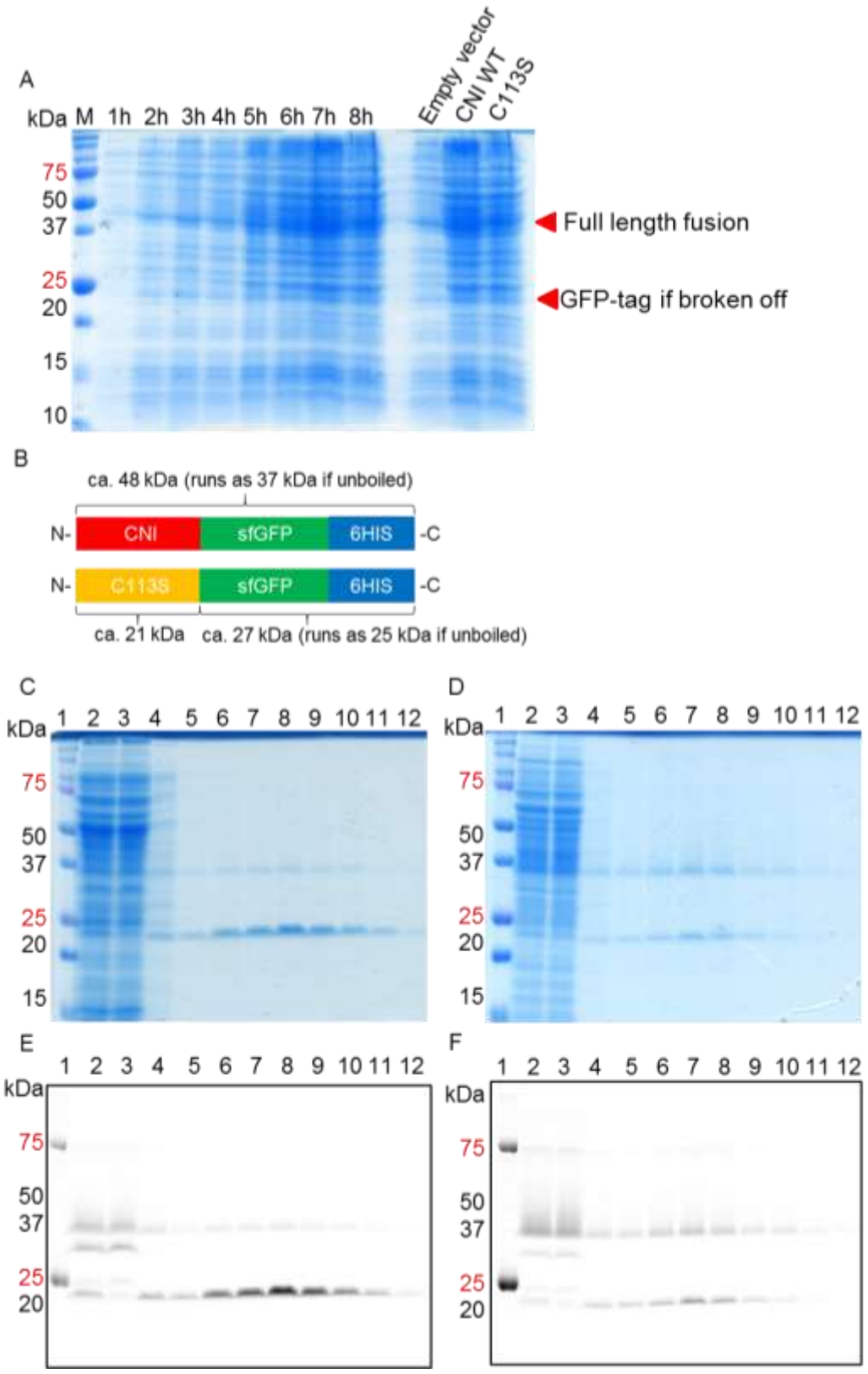


Figure 52: Cysteine mutant expression and purification. (A) Coomassie stained SDS-PAGE gel showing C113S expression over time and in comparison to WT CNI-sfGFP-6HIS and the empty vector control after 8 h. An increase in expression is not visible. Coomassie stained SDS-PAGE gel showing C113S (C) and CNI-WT (D) purification. In-gel fluorescence images of SDS-PAGE gels showing C113S (E) and CNI-WT (F) purification. There is no difference in dimerisation between C113S and CNI-WT, although the C113S mutant seems to degrade more easily. Lanes: 1) Marker, 2) Load, 3) Flow-through, 4) Wash, 5) - 12) Eluates.

4.2.7 Assessing in vitro protein-protein interaction with a pull down assay

A pull-down assay was used to assess protein-protein interaction between CNI and Colicin N. A protein pull down is a simple and popular method because it only requires minimal handling of proteins, which is particularly favourable when dealing with unstable membrane proteins. However, pull down assays are very prone to false positive and false negative results which can occur due to the change in protein conformation caused by the *in vitro* environment and the exposure of protein sites to each other which would not normally interact *in vivo.* Weak and transient interactions which would normally occur due to the confined space of a cell or even membrane are unlikely to be detected using pull-downs. Further caveats of this technique are discussed in Mackay *et al.* (2007).

Here it was attempted to recreate the natural environment for CNI-ColN interaction using detergents (Figure 53). CNI was bound via its HALO7-6His-tag to HaloLink resin and incubated with ColN, ColN P-domain, ColA and ColA P-domain in buffers containing different detergents. ColN, ColN P-domain and ColA were purified for this experiment Figure 54. A stock of ColA P-domain was obtained from Dr Yan Huang, formerly Lakey group. Figure 55 shows the purity level of each protein. After incubation the beads were washed to remove unbound protein. To establish what remained bound to the beads due to protein-protein interaction, they were resuspended in protein loading buffer and heat-denatured at 90 °C. Washes and "pulled-down" proteins were analysed using SDS-PAGE. The colicins and P-domains do not bind to the HaloLink resin if no CNI-HALO7-6His is present.

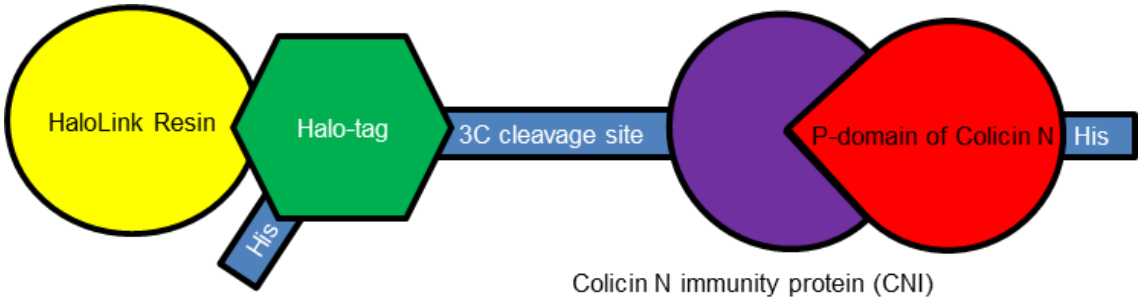


Figure 53: Desired arrangement of proteins during the pull-down assay. CNI was bound via its HALO7-tag to HaloLink resin and incubated with ColN, ColA and their P-domains in buffers containing different detergents. The colicins and P-domains should not bind to the HaloLink resin if no CNI-HALO7-6His is present.

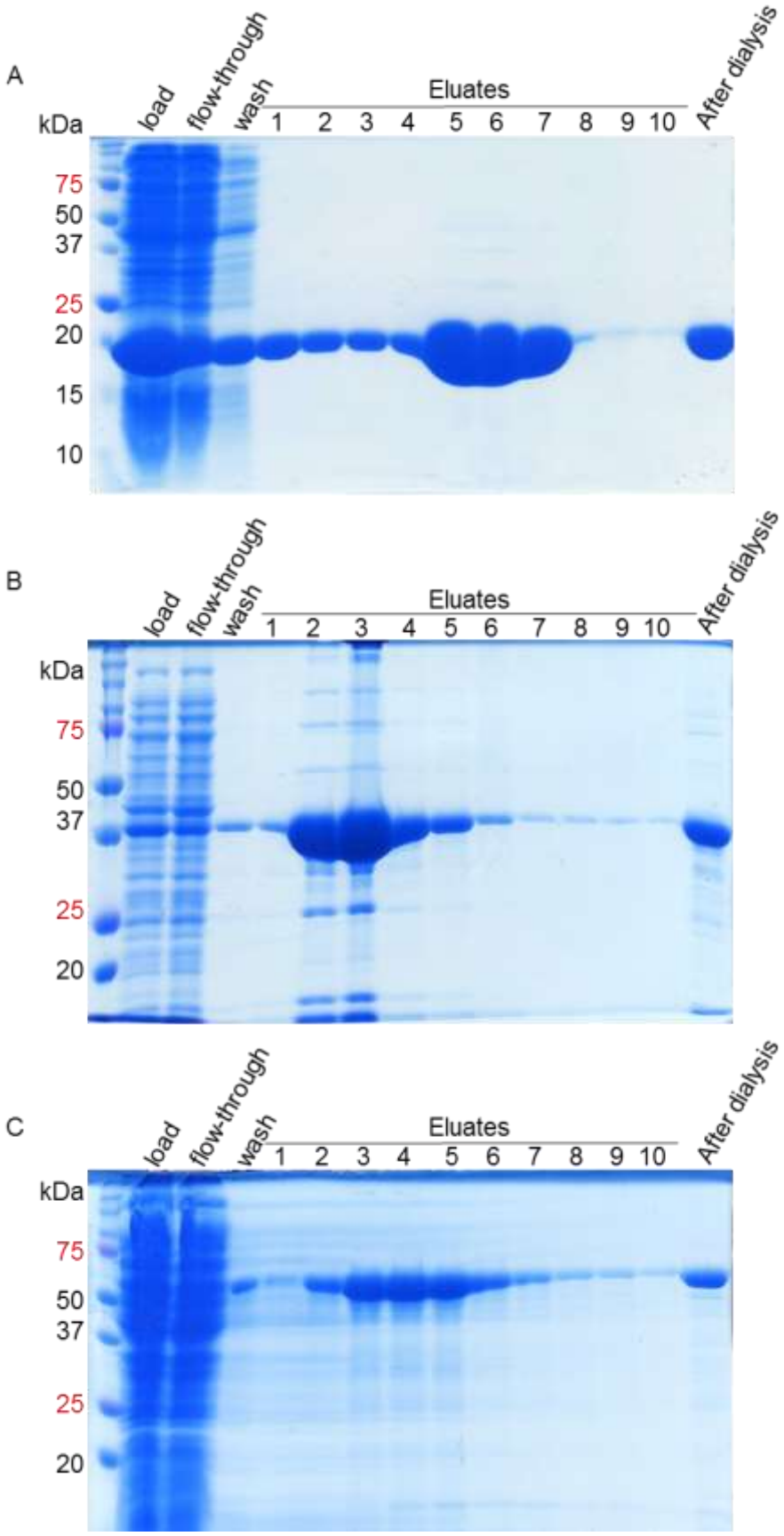


Figure 54: Protein purification steps of (A) ColN P-domain (23 kDa), (B) ColN, (ca. 42 kDa) and (C) ColA (ca. 63 kDa) in Coomassie-stained SDS-PAGE gels.

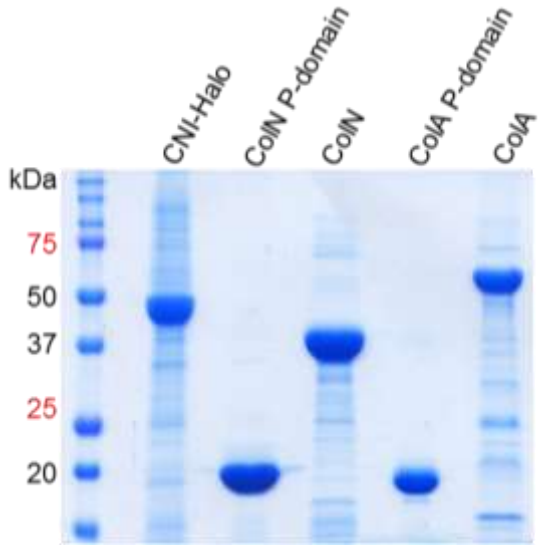


Figure 55: Coomassie stained SDS-PAGE of all five used proteins: CNI-Halo, ColN P-domain, ColN, ColA P-domain and ColA. The full length proteins show some degradation and CNI-Halo is also showing some aggregation.

A Circular Dichroism spectrum (Figure 56) shows how a detergent like DM can modify the tertiary structure of ColN P-domain. A less structured and more open conformation of P-domain is desirable for interaction with CNI because the P-domain changes conformation into a more open form during its insertion into the membrane. A P-domain which is locked shut through disulphide bonds, so that the helices cannot move apart for insertion, is not able to kill cells (Duche *et al.*, 1994). Colicin A and its P-domain are used as negative controls for ColN and its P-domain. Using the P-domain as well as the entire protein can exclude potential interactions of CNI with ColN T-domain or R-domain.

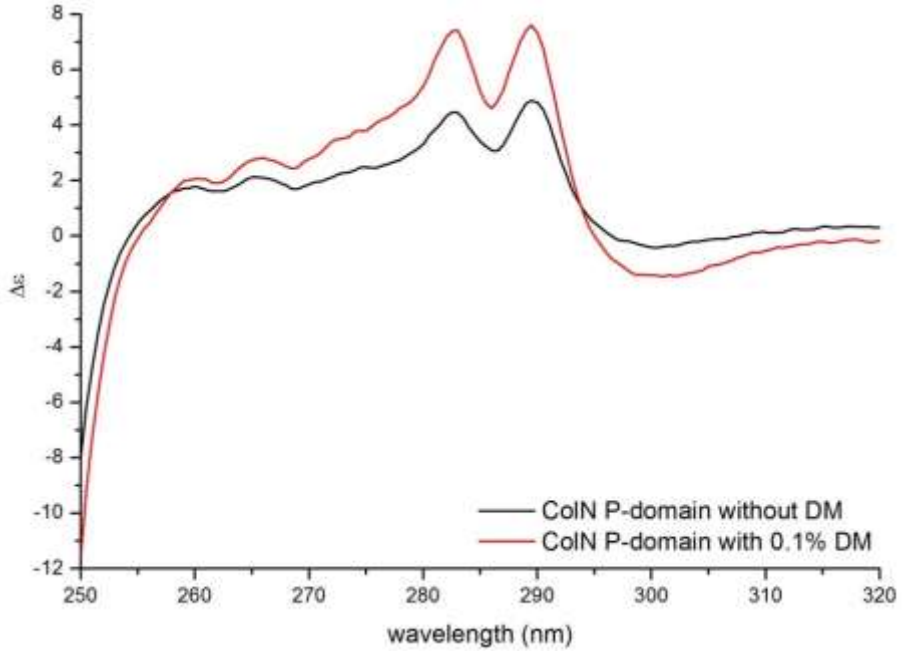


Figure 56: The effect of ColN P-domain modification with 0.1 % DM shown in a near-UV CD spectrum which is sensitive to tertiary structure destabilisation. After incubation of P-domain with 0.1 % DM for 24 h at 4 °C it loses some of its tertiary structure and is better suited for interaction with CNI.

Different detergents were added to a Tris buffer to establish the ideal conditions to mimic the specific interaction between ColN P-domain and CNI (Figure 57). ColN P-domain does not bind to the beads if CNI-HALO7-6His is not present. However, it binds to CNI in all three detergents equally well. The same amount of ColN P-domain and beads was used for each detergent. The washes remove all unbound protein before the beads are boiled and CNI-HALO7-6His and any ColN P-domain bound to it is released from the bead surface. Two different separation methods are possible to remove the complex from the bead surface – denaturing heat and 3C protease. 3C protease was not used at this initial stage as it is not active in SDS and it would also contaminate the sample with an additional band. DM is most suited to solubilise CNI-HALO7-6His. OG is also a non-ionic detergent but has a different structure to DM and may improve P-domain opening. SDS is an ionic detergent known to denature most proteins.

Following a successful demonstration of ColN P-domain binding to CNI-HALO7-6His (Figure 57), full length ColN, ColA and ColA P-domain were also used in a pull down assay (Figure 58). DM was chosen as the detergent because it is the detergent which best solubilises and stabilises CNI-HALO7-6His (chapter 3). CNI-HALO7-6His is bound to the HaloLink resin (Figure 58A) and does not dislodge until it is denatured with heat (boiled) (Figure 58 D). Full length ColN, ColA and ColA P-domain also bind to CNI-HALO7-6His and can only be removed when the beads are boiled. In this condition, there is no specific interaction between CNI-HALO7-6His with ColN P-domain and full length ColN as it also binds to ColA and ColA P-domain. SDS is an ionic detergent and therefore can mask potential unspecific charge interactions. The pull down with ColN, ColN P-domain, ColA and ColA P-domain was repeated with SDS replacing DM in the Tris buffer (Figure 59). It was also assessed if ColN, ColN P-domain, ColA and ColA P-domain bind to the beads if CNI-HALO7-6His is not present. There is some artefactual binding of ColN and ColN P-domain to the beads if CNI-HALO7-6His is absent, while there is no binding of ColA and ColA P-domain. When CNI-HALO7-6His is present all tested proteins (ColN, ColN P-domain, ColA and ColA P-domain) bind to CNI-HALO-6His. Adding SDS to the buffer does not remove unspecific binding of ColA and ColA P-domain to CNI and so this does not recreate the *in vivo* conditions, where CNI specifically binds ColN P-domain only and no other colicin.

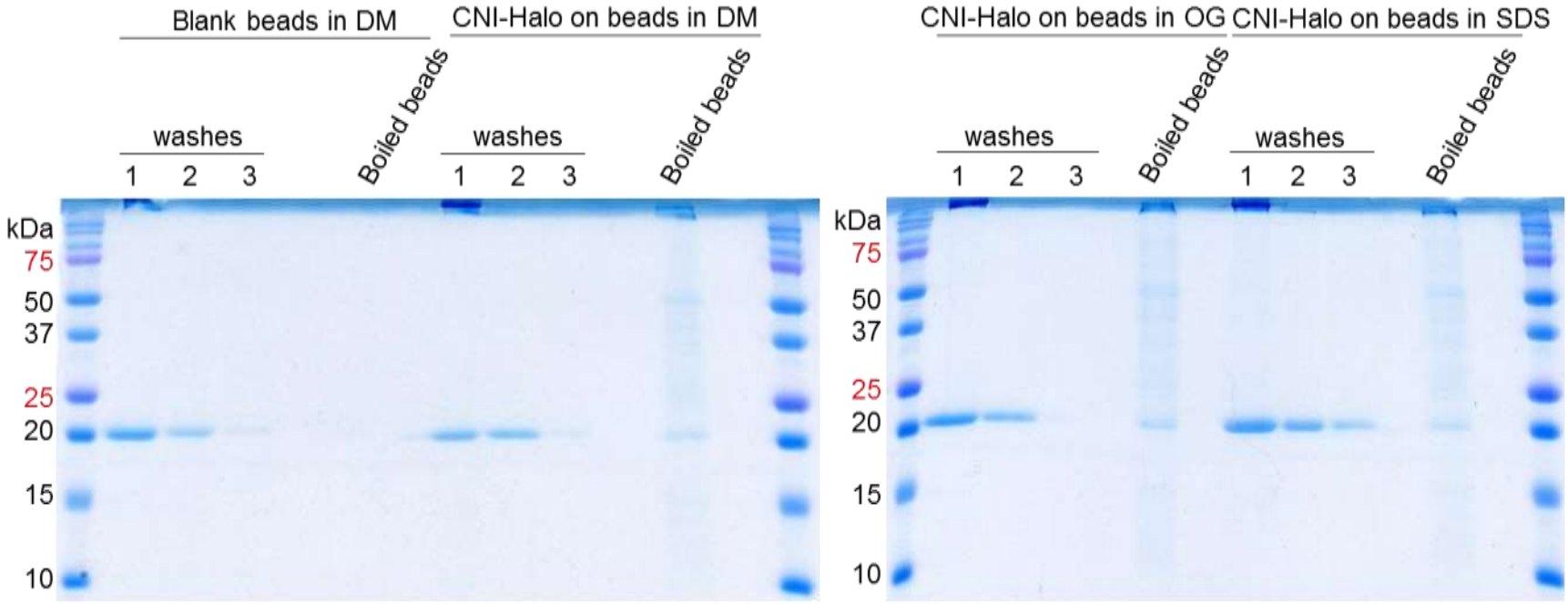


Figure 57: Screening detergents for an optimal CNI-HALO7-6His interaction with ColN P-domain (23kDa) using a pull-down assay and Coomassie stained SDS-PAGE. CNI-HALO7-6HIS binds to the HaloLink resin beads via the HALO7-6His-tag. ColN P-domain binds to CNI-HALO7-6His at 4 °C over night in Tris buffer containing 1.6 mM DM, 25 mM OG and 8 mM SDS, which corresponds to the detergent's critical micelle concentrations. ColN P-domain does not bind to blank beads without CNI-HALO7-6His.

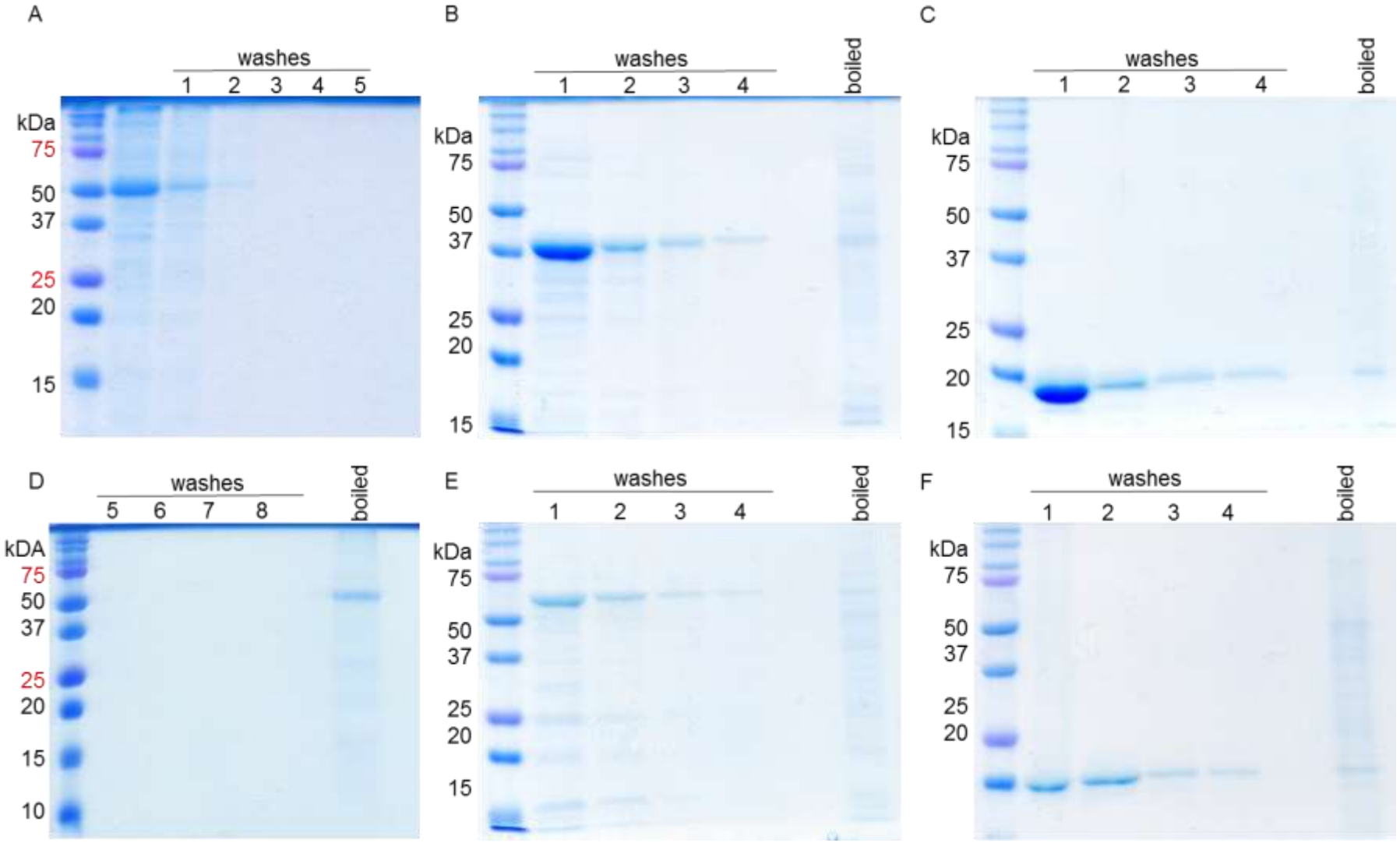


Figure 58: Pull down assay with various colicins in DM. Samples were incubated for 16 h at 4 °C with CNI-HALO7-6His bound to beads. (A) CNI-HALO7-6His binding to HaloLink resin beads and washing steps. Washing steps are continued after 24 h incubation in (D) and the boiled sample shows CNI-HALO7-6His eluting after the beads are boiled. (B) ColN bound to CNI-HALO7-6His, (C) ColN P-domain bound to CNI-HALO7-6His, (E) ColA bound to CNI-HALO7-6His (F) ColA P-domain bound to CNI-HALO7-6His.

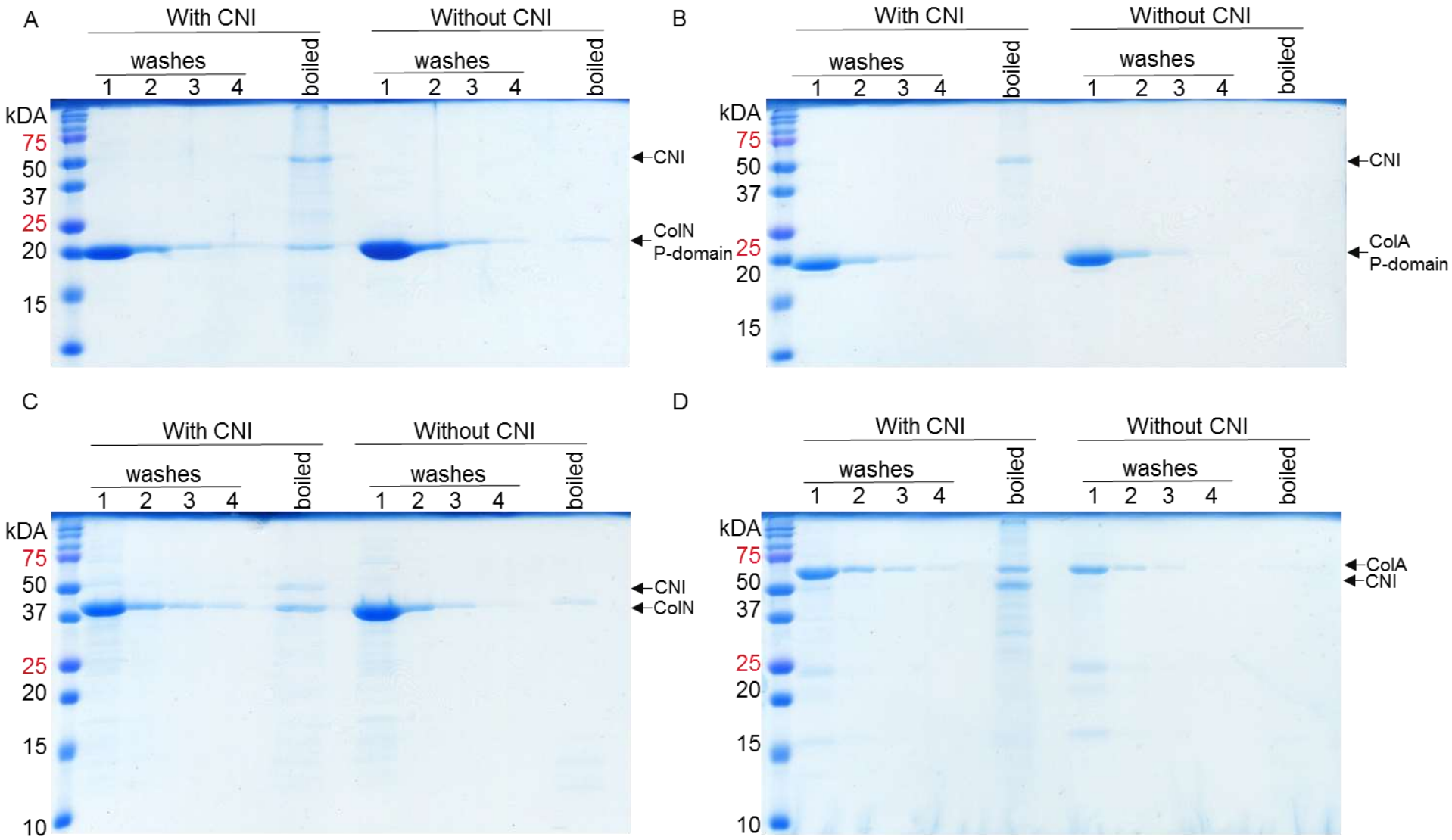


Figure 59: Pull down assay with SDS. Samples were incubated for 16 h incubation at 4 °C with or without CNI-HALO7-6His bound to beads. (A) ColN P-domain, (B) ColA P-domain, (C) ColN, (D) ColA. CNI-HALO7-6HIS binds to all proteins; there is no specific interaction with ColN or ColN P-domain. However, there is some interaction of ColN and ColN P-domain even without CNI-HALO7-6His.

4.2.8 In vitro interaction between CNI and ColN P-domain using Surface Plasmon Resonance

Unfortunately, no specific interaction between ColN and CNI could be established using the pull-down assay. It is possible that CNI-HALO7-6His can bind both ColN P-domain as well as ColA P-domain but cannot neutralise ColA. As an alternative approach surface plasmon resonance was then used to measure the binding affinity of purified CNI to ColN and ColA as a way of distinguishing between the two colicins. It was hypothesised that CNI would bind ColN tighter than ColA as it would be a more specific and energetically more favourable interaction.

Surface Plasmon Resonance is a well-established technique with many applications (Salamon *et al.*, 1997b; Salamon *et al.*, 1997a; Mariani and Minunni, 2014). Fundamentally, one molecule is bound to a surface and then the attachment of a second molecule to the first is measured through a change in the refractive index which is caused by increased surface density. Here, CNI-HALO7-6His was bound to a Ni^{+}-saturated NTA chip via its polyhistidine-tag and ColN, ColA and their P-domains were injected in order to measure their binding affinity to CNI (Figure 60).

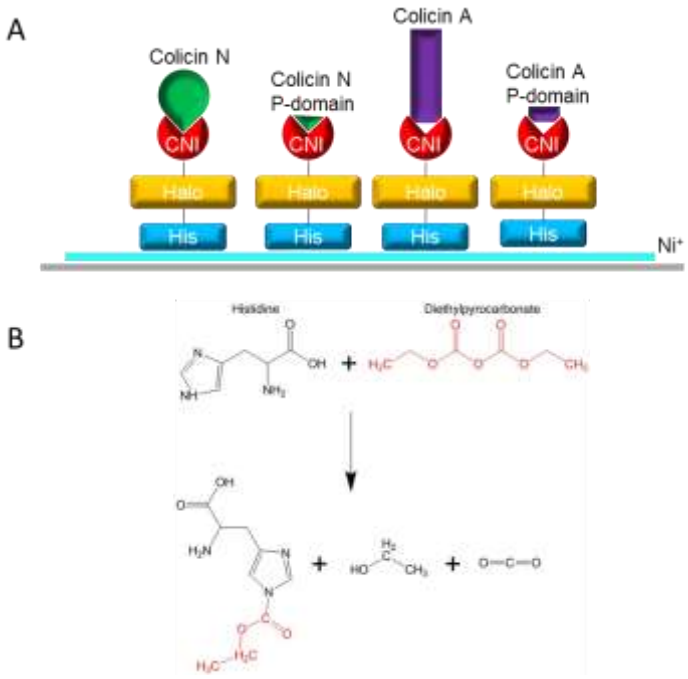


Figure 60: (A) Schematic representation of CNI-HALO7-6His bound to a Ni^{+}-NTA chip and ColN, ColN P-domain, ColA, ColA P-domain binding to CNI- HALO7-6His. (B) The polyhistidine tag was modified using diethylpyrocarbonate.

ColN, ColA and their P-domains have been purified using their His-tags and so, first, these His-tags had to be be removed so that ColN, ColA and their P-domains would no longer bind to the Ni^{+}-NTA surface but only to CNI-HALO7-6His (Figure 60, Figure 61). The histidines were modified using DEPC following a protocol adapted from Lundblad and Noyes (1984). ColN P-domain has one histidine in its native sequence and 6 additional ones at the C-terminal. When DEPC modification is successful the absorbance at 240 nm increases while the absorbance at 280 nm should stay the same. This indicates that the imidazole ring on the histidines is modified but the aromatic residues are not affected. Figure 61A shows ColN P-domain modifications with different DEPC amounts, based on the molar ratio between the histidines and the DEPC molecules. Figure 61A shows the end point modification over a spectrum of 230 nm to 320 nm, including at 240 nm and 280 nm after a 20 min reaction time. The curve changes depending on the concentration of DEPC, more DEPC leading to a bigger change in comparison to the unmodified molecules. The expected increase in absorbance at 240 nm is more apparent in Figure 61B, where the increase in absorbance at 240 nm only is measured over time. This graph shows that modification occurs rapidly at the beginning but then plateaus off. Again the more DEPC is added, the higher is the absorbance at 240 nm. As the exact molar ratio for histidine-tag modification was unknown, first a broad range (Figure 61 A) and then a more narrow range (Figure 61 B) of modifications was used. Whether the modification was successful was tested by applying the modified ColN P-domains onto the Ni^{+}-NTA chip, without adding CNI-HALO7-6His first. A 5:1 molar ratio, so adding 5 times more DEPC molecules than histidine molecules proved to be the optimum ratio. Saturation is reached during modification and these modified ColN P-domains also no longer bind to the Ni^{+}-NTA chip (Figure 61C and D). This ratio was therefore used for all subsequent modifications.

When trying to deposit CNI-HALO7-6His on the Ni^{2+}-NTA chip in preparation for the binding experiment it was discovered that CNI-HALO7-6His not only binds to the chip via the his-tag but also in a non-specific way when no Ni^{2+} ions are present. It also seems to bind to itself when it is repeatedly injected. A layer of CNI-HALO7-6His builds up further and further and does not dissociate. Possibly, the protein is aggregating on the chip. This is a major obstacle for carrying out this experiment as ideally a single layer of separate proteins with very few response units (R.U.) is needed to measure binding affinity accurately.

In subsequent runs, CNI-HALO7-6His was dialysed into and stored in OG and SDS but the result was the same. CNI-HALO7-6His built up on the surface and measuring binding was not possible. Figure 62 B shows the reference cell and C the measuring cell. CNI binds in both regardless of whether Ni^{+} ions are present or not. Figure 61 D and F show that a subtraction of the reference cell from the measuring cell with and without OG does not yield a net binding event and so carrying out further experiments with these conditions would not lead to reliable and meaningful results. Furthermore, ColN P-domain seems to bind to CNI-HALO7-6His in this assay. However, when run as a control by itself, without its binding partner CNI-HALO7-6His, ColN P-domain also seems to bind to the chip. Binding is increased with higher OG concentration and so it seems that the OG may be responsible for aggregation and unspecific binding. The binding was repeated with SDS but the effect was that, firstly, CNI-HALO7-6His again aggregated on the chip and secondly, when P-domain in SDS buffer was added to the chip, it would wash off the aggregated CNI-HALO7-6His from the chip.

Under the tested conditions measuring interaction between CNI-HALO7-6His and ColN P-domain is unfeasible, mainly, due to the instability of CNI-HALO7-6His and its aggregation on the chip. This makes any measurement of binding kinetics very difficult to interpret and measure.

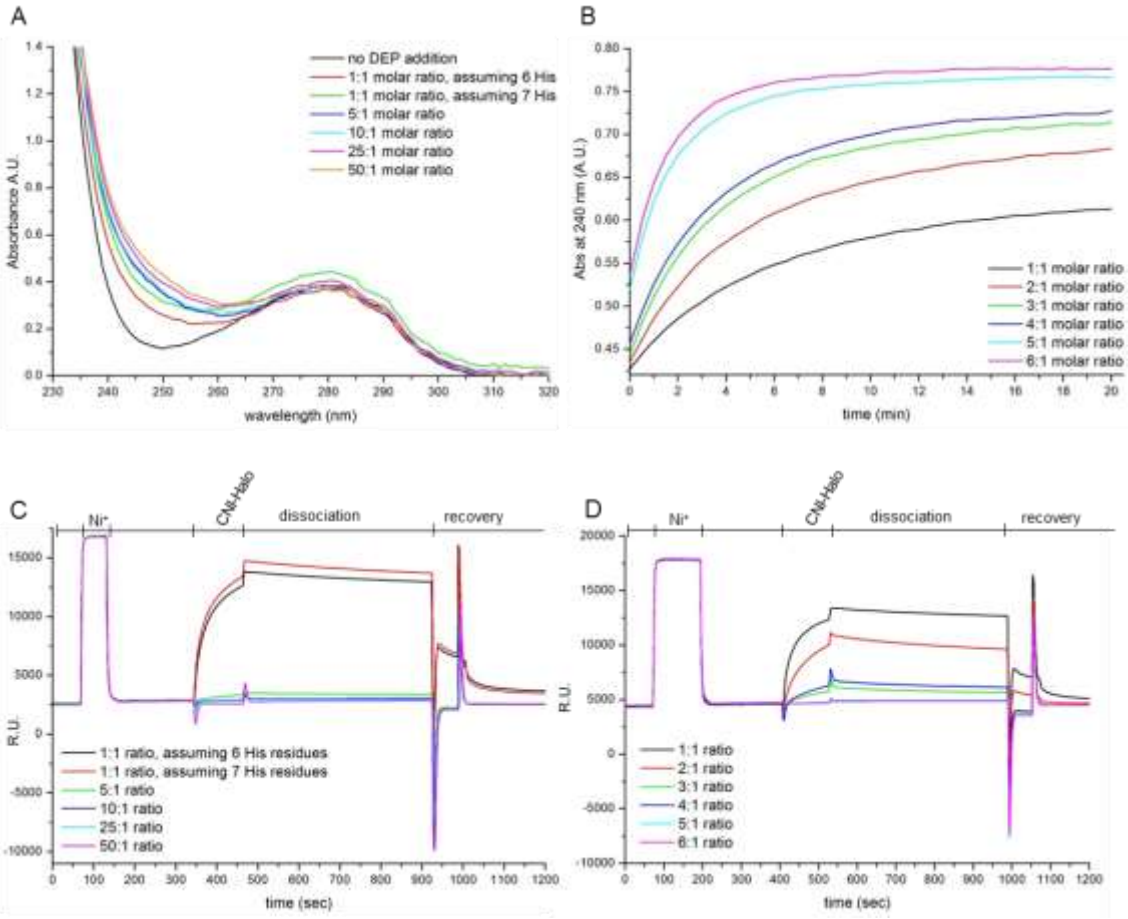


Figure 61: ColN P-domains modification with DEPC. (A) Histidine-tag modifications with DEPC; concentrations expressed as molar ratios between Histidine molecules to DEPC molecules. End point measurement of absorbance over a spectrum 230 nm – 320 nm. (B) Histidine-tag modifications with DEPC, concentration expressed as molar ratios between Histidine molecules to DEPC molecules. Continuous absorbance measurement at 240 nm over 20 min at room temperature. (C) and (D) Abolished binding of modified ColN-P-domain to the Ni^{+}-NTA chip. A 5:1 molar ratio, assuming 7 histidines in the ColN P-domain, during modification is enough to abolish binding to the Ni^{+}-NTA chip.

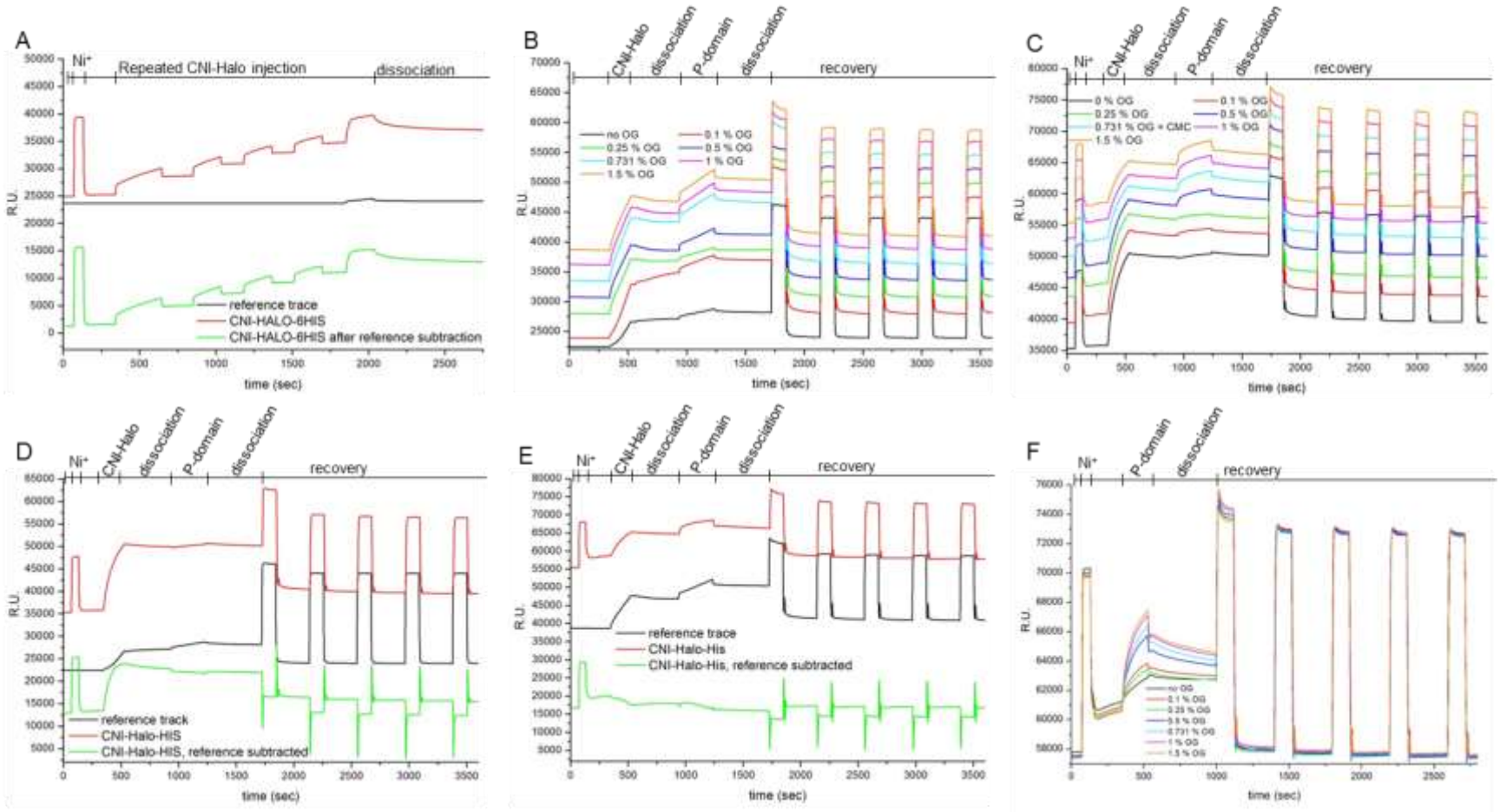


Figure 62: CNI-HALO7-6His aggregates on the Ni+-NTA chip and binding to ColN P-domain cannot be measured. (A) CNI-HALO7-6His builds with up with each 15 µM injection, indicating aggregation. (B) Reference cell of NTA-chip. CNI-HALO7-6His and ColN P-domain bind without Ni+ addition, demonstrating unspecific binding, not via the His-tag. (C) Measuring cell showing CNI-HALO7-6His and ColN P-domain binding to NTA-chip when Ni+ is added. (D) CNI-HALO7-6His and ColN P-domain binding without OG. Subtracting the reference cell from the measuring cell results in negative net binding. (E) CNI-HALO7-6His and ColN P-domain binding with 1.5 % G. Subtracting the reference cell from the measuring cell results in negative net binding. (F) Without CNI-HALO7-6His, increasing OG concentrations increase unspecific binding to the Ni+-NTA chip.

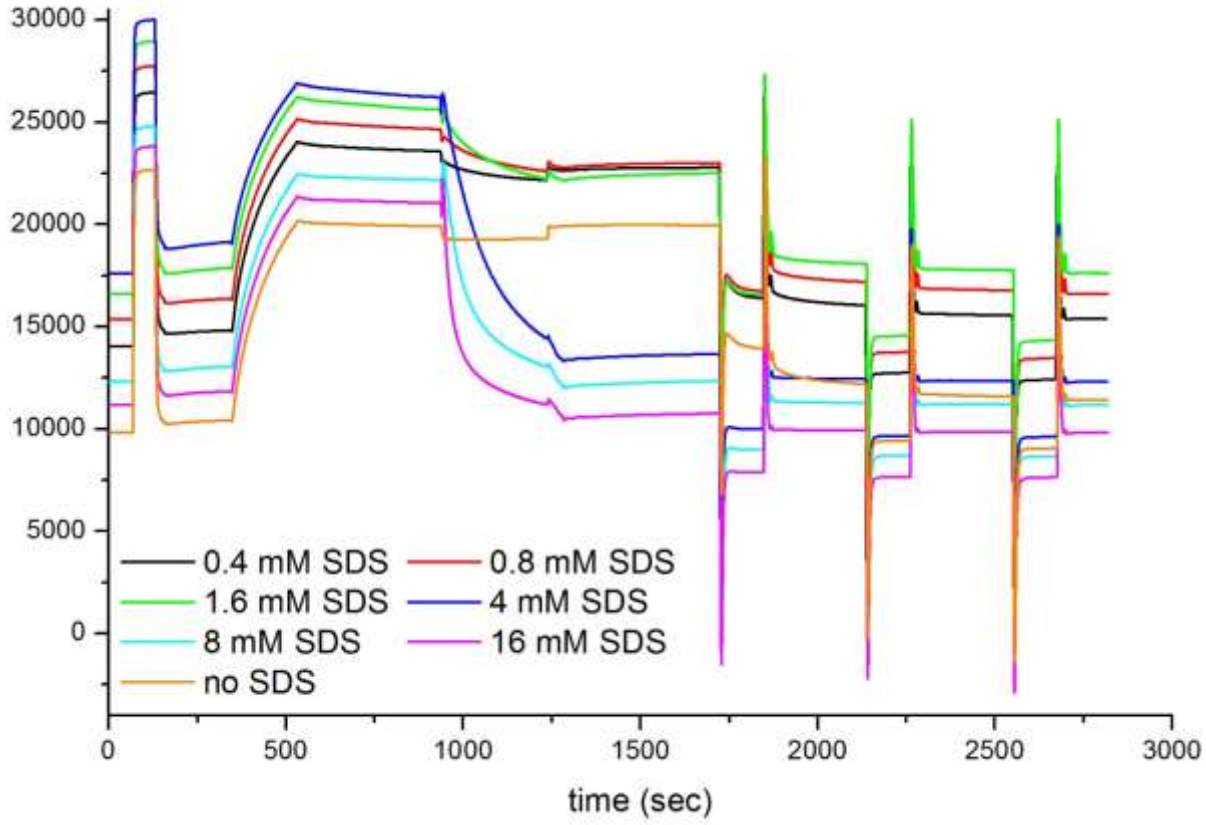


Figure 63: P-domain incubated in SDS washes CNI-HALO7-6His away. SDS is able to wash away aggregated CNI-HALO7-6His from the chip. A binding affinity cannot be measured using these conditions.

4.2.9 Conclusions

- Helix 1 of CNI is not needed for specific interaction with ColN. If Helix 1 is essential for CNI structure, Helix 1 of CAI is able to replace it.
- Helix 2, 3 and 4 are crucial for protein stability and function, indicating that the CNI helices have a function and the homologous helices of CAI are not able to replace their function.
- Mutagenesis of the only cysteine residue to a serine leads to a reduction in immunity but does not abolish protein function, indicating that the cysteine probably plays a structural role and that dimerisation through a disulphide bridge is not a crucial feature of CNI activity.
- Specific interaction between CNI and ColN could not be demonstrated *in vitro* using SPR and pull-down assays due to protein aggregation and unspecific binding.

Chapter 5. In vivo observation of GFP-tagged Colicin N immunity protein.

5.1 Introduction

In contrast to enzymatic colicins, there is a profound lack of understanding how immunity proteins to pore-forming colicins work *in vivo*. These immunity proteins are located in the *E. coli* inner membrane and protect cells from colicin induced death (Cascales *et al.*, 2007). However, there are no studies attempting to quantify the level of immunity and looking at the *in vivo* interaction of colicins and their immunity proteins. There have been attempts at producing immunity proteins fused to recognition tags (Zhang *et al.*, 2010), solubilisation tags and fluorescent proteins (Smajs *et al.*, 2008), but none of these studies have used a fluorescent tag to visualise the location of the immunity protein in the membrane and demonstrate protein-protein interaction *in vivo* using these tags. This chapter focuses on CNI localisation and how it is affected by the addition of its interaction partner Colicin N.

5.2 Results and Discussion

5.2.1 Expression system and fluorescent signal optimisation

CNI was tagged C-terminally with superfolder GFP (Pedelacq *et al.*, 2006) plus a 6His-tag and cloned into an arabinose inducible expression system which can also be used for protein purification (Chapter 3). As well as stabilising the protein and simplifying its purification, the tag also allows CNI visualisation *in vivo*. Following arabinose addition for 2 h prior to imaging, a fluorescent signal clearly localised to the cell periphery was observed (Figure 64). The protein fusion CNI-sfGFP-6His is evenly distributed in the membrane, not favouring any particular location. No obvious differences between dividing and non-dividing cells, or abnormal morphology that would indicate toxicity, which might be expected when overexpressing a membrane protein, were observed. As this is an overexpression system, induction with arabinose may lead to unspecific localisation caused by accumulation of protein in the membrane, which may mask any localisation patterns that would occur with natural amounts of protein.

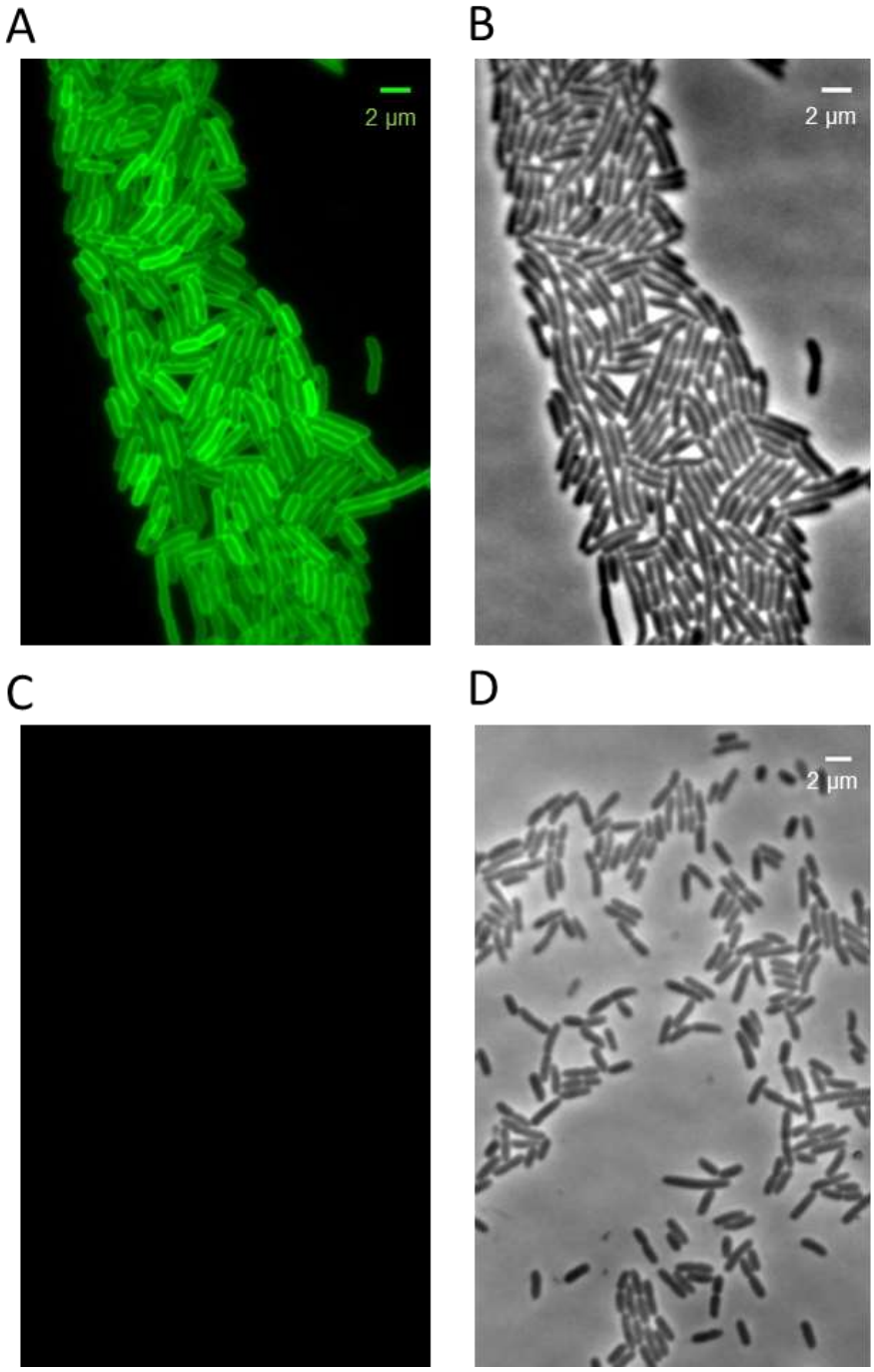


Figure 64: Overexpression of CNI-sfGFP-6His in MC1000 cells. Bright-field (B and D) and epifluorescence (A and C) images showing overexpression of CNI-sfGFP-6His in MC1000 cells (A and B) and the empty vector control MC1000 cells, which are not fluorescent (C and D). Overexpressing cells do not exhibit abnormal morphology. The fluorescent signal is localised to the edge of the cell, demonstrating a membrane localisation of the fusion protein.

To reduce the potential masking effect of overexpression and reveal potentially hidden localisation patterns, it was attempted to reduce the level of expression using the repressors glucose and fucose (Figure 65) to a more natural level, although the precise natural expression is unknown. Glucose is a known repressor of the arabinose inducible promoter (Guzman *et al.*, 1995) but is also a nutrient source

which could be broken down during cell growth and fail to repress expression at later stages of culture growth. Fucose is structurally very similar (Figure 65) and also represses expression from an arabinose inducible promoter but is not an easily degradable nutrient (Guzman *et al.*, 1995). The addition of repressors glucose and fucose at the same time as the inducer arabinose did not lead to a titration of the protein amount. Instead, protein expression was either induced or repressed (Figure 65), which is what is expected from a catabolite repression system (Lichenstein *et al.*, 1987; Schleif, 2000; Schleif, 2010). There was no difference in repression by glucose and fucose. Consequently, enough glucose (15 mM ≈ 0.2 %) was added to repress protein expression, even if some was used as a nutrient source.

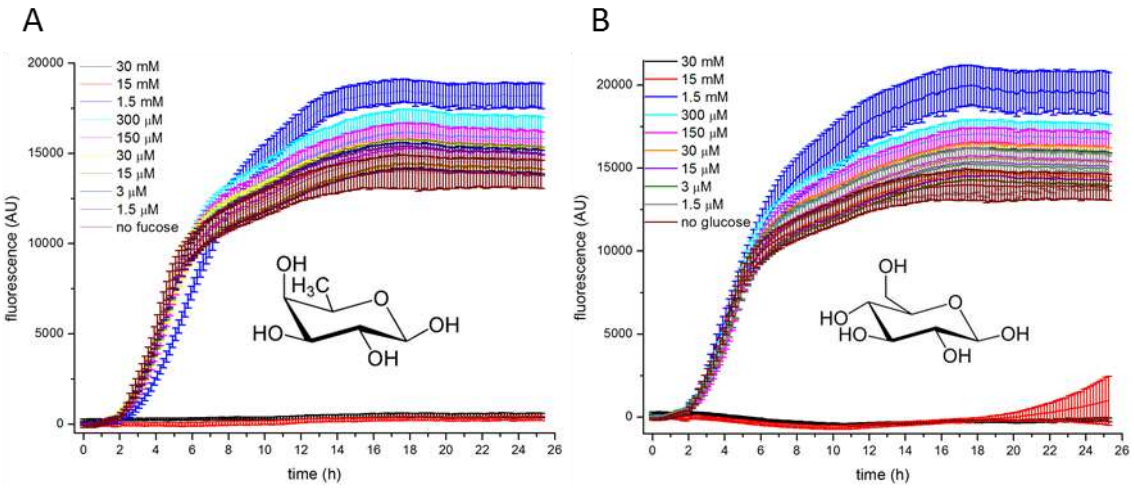


Figure 65: The effect of glucose (A) and fucose (B) addition on the expression of CNI-sfGFP-6His as measured by the plate reader

Another way of reducing protein expression is by delayed instead of simultaneous repressor addition. The repressor glucose was added at 0.2 % 10 sec, 20 sec and 30 sec after induction with 0.2 % arabinose to allow for a limited amount of transcription to occur before it is repressed (Figure 66). Gene expression is completely repressed when fucose or glucose are added, which allows these pulse induction for 10 sec, 20 sec or 30 sec. Protein production is visible even with a 10 sec induction, while no induction leads to no visible production of CNI-sfGFP-6His, meaning that the promoter is not significantly leaky.

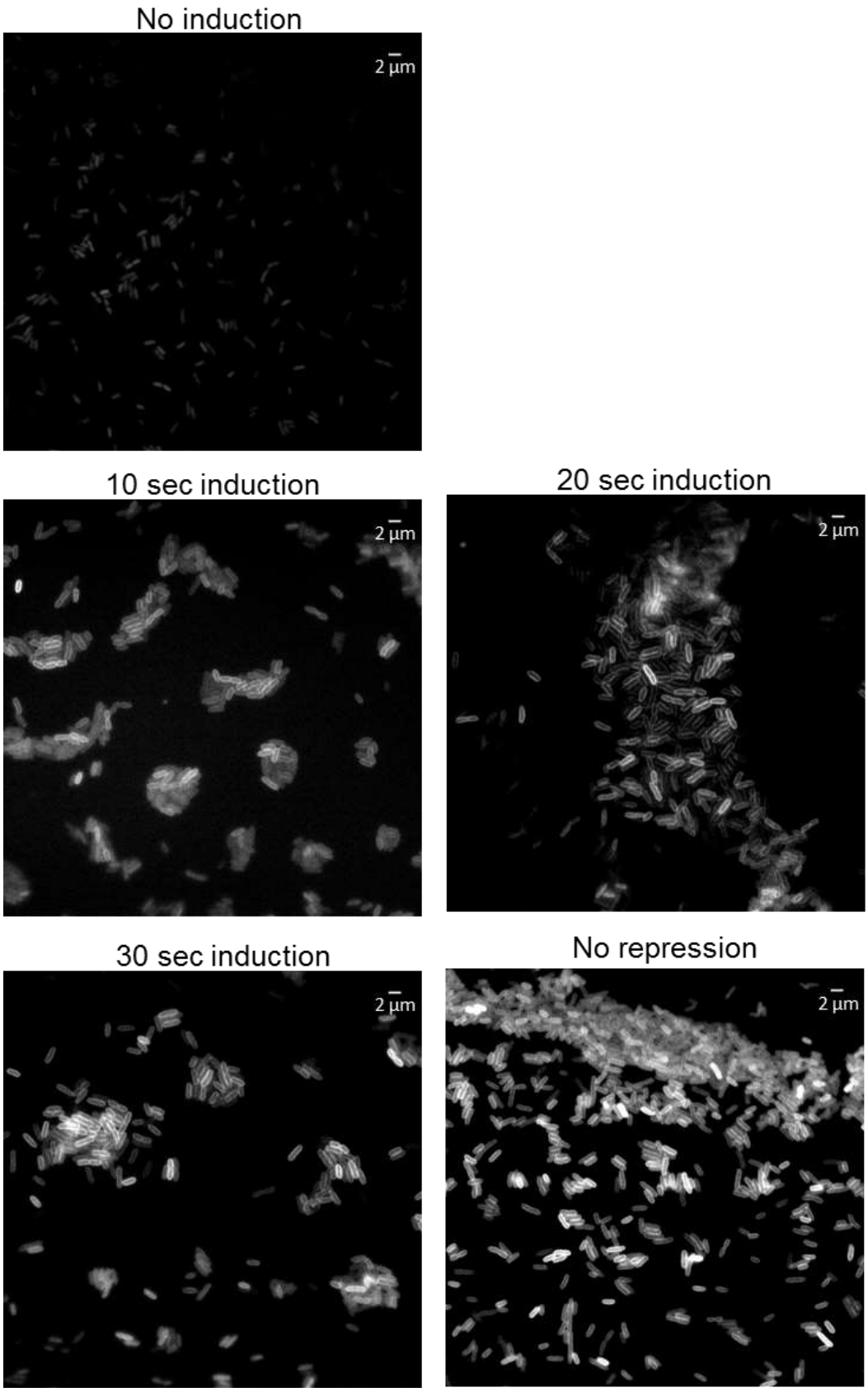


Figure 66: Optimisation of CNI-sfGFP-6HIS expression by delayed repression. CNI-sfGFP-6HIS expression in MC1000 cells is induced with 0.2% arabinose and repressed with 0.2% glucose after 10 sec, 20 sec or 30 sec.

Subsequently, 10 sec of arabinose induction were used prior to epifluorescent and Total Internal Reflection Fluorescence Microscopy (TIRF; (Axelrod, 2008; Jin *et al.*, 2008; Li *et al.*, 2008; Shawkat *et al.*, 2008; Hategan *et al.*, 2013)) in order to observe the distribution and movement of CNI-sfGFP-6His in the cell (Figure 67, Figure 68). At this low level of production the protein still appears evenly distributed in the cell membrane. There is no clear patterning or clustering of the protein. This expression level probably can no longer be considered overexpression but since the natural level of expression is unknown we cannot be certain.

The distribution of CNI-sfGFP-6His was compared to WALP23-eGFP-6His (eGFP-tagged WALP23 is a gift by Dr H. Strahl, Newcastle University), which is an artificial single helix transmembrane domain without any specific function but can act as a model for a generic membrane protein (Weiss *et al.*, 2003; Siegel *et al.*, 2006; Holt *et al.*, 2009; Kim and Im, 2010). Here it acts as a negative control, visualising how a randomly and evenly distributed, unclustered protein without an interaction partner would appear *in vivo*. WALP23-eGFP-6His is used here also to highlight any non-specific effects colicin addition might have on membrane proteins. CNI-sfGFP-6His appears smoother than the more coarse WALP23-eGFP-6His or the clearly localised TolA-eGFP-6His. TolA is essential for A-Type colicin translocation and toxicity and is known to localise strongly to the septum during cell division as well as being present throughout the membrane (Levengood *et al.*, 1991; Walburger *et al.*, 2002; Gerding *et al.*, 2007; Hecht *et al.*, 2010; Penfold *et al.*, 2012). This appearance could also be explained by an unusually high lateral diffusion of CNI-sfGFP-6His molecules leading to a blurred signal and therefore a smooth appearance.

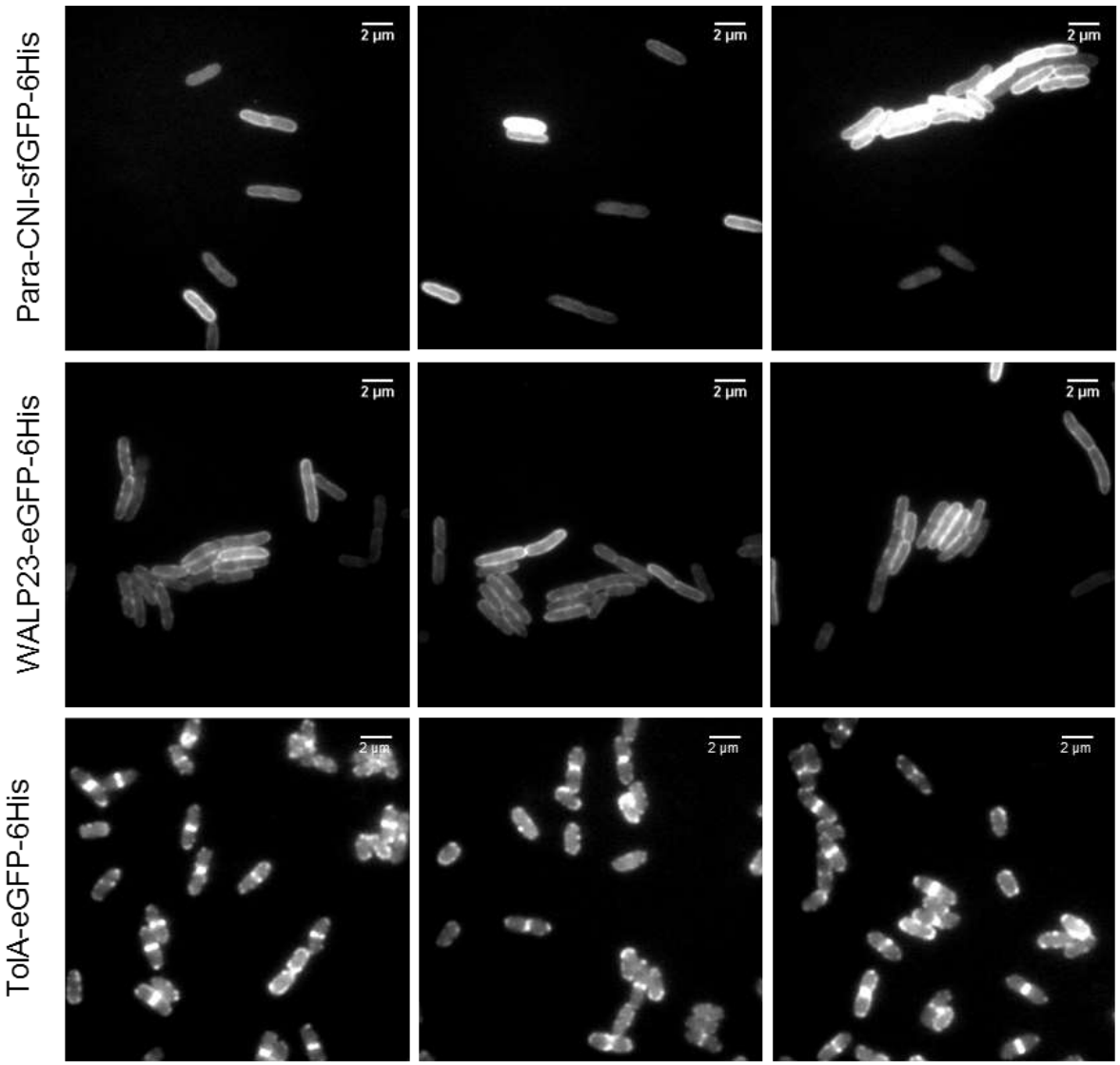


Figure 67: Epifluorescence images show 3 replicates of CNI-sfGFP-6HIS, WALP23-eGFP-6HIS and TolA-eGFP-6His. CNI-sfGFP-6HIS expression is heterogeneous between cells but looks very evenly distributed in the membrane. WALP23-eGFP-6HIS expression is also heterogeneous, but is also uneven in the membrane. TolA-eGFP-6His is present everywhere in the cell membrane but shows some clear localisation to the septum, which is consistent with the localisation pattern shown in Gerding *et al.* (2007).

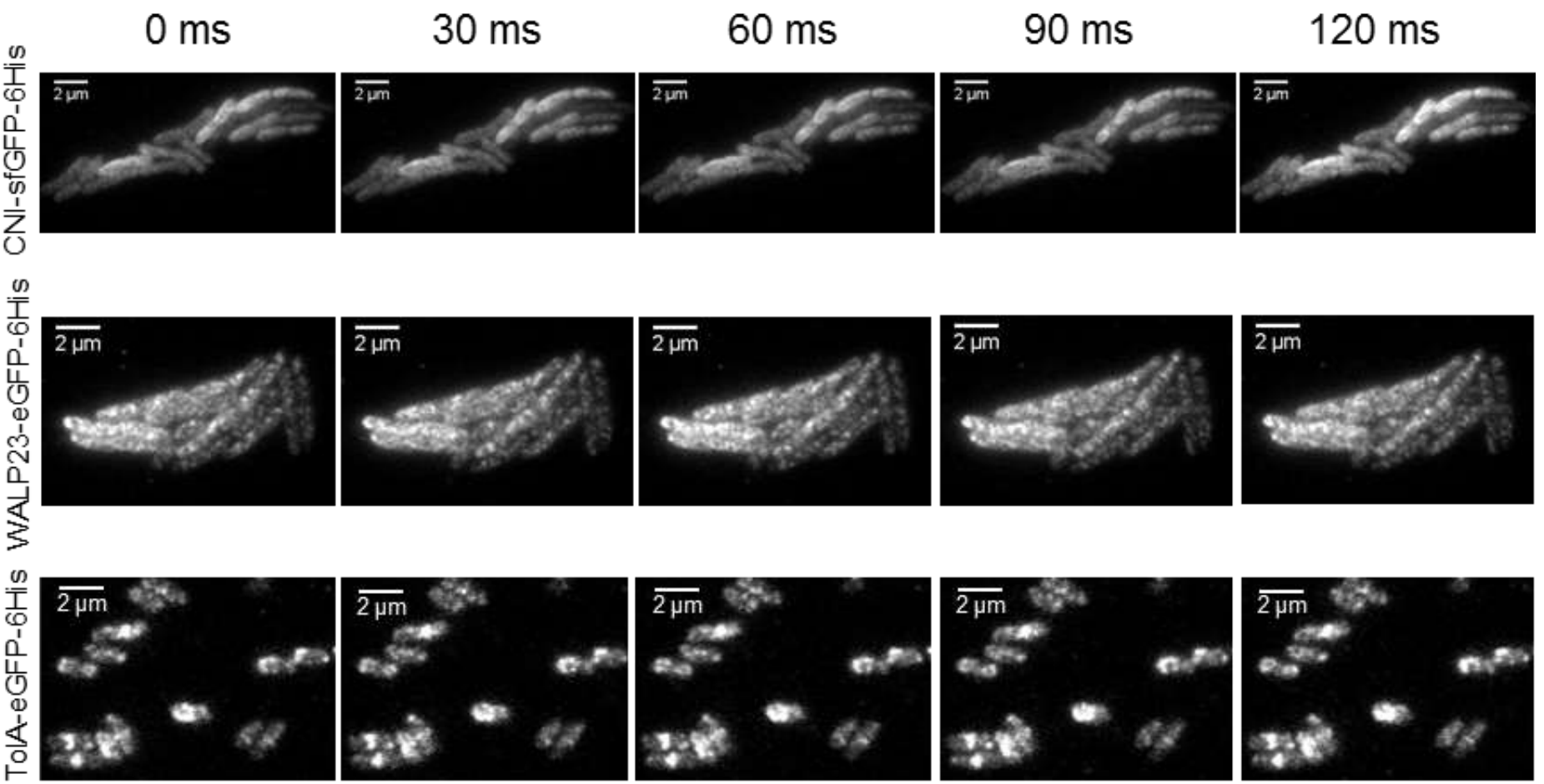


Figure 68: CNI-sfGFP-6His, WALP23-eGFP-6HIS, TolA-eGFP-6His distribution in the cell as seen in TIRF microscopy. Scale bar in upper right corner represents 2 µm.

5.2.2 Expression of CNI using its natural promoter

In an attempt to approximate natural expression levels the arabinose inducible promoter on pBAD322 was replaced by the natural promoter of CNI, P_{CNI}. InFusion cloning (Clontech Laboratories, Inc.) was used to delete the arabinose inducible promoter (location determined in Cronan, 2006) and insert the natural promoter (amplified from pCHAP4, the natural plasmid coding for ColN) in its place at the same time. Both pBAD322 and pCHAP4 are low copy number plasmids and so expression levels should be a good approximation to the natural state. The location of the natural promoter was determined by Pugsley (1988) somewhere between the BamHI cleavage site and the start of the gene. When this region is deleted, CNI was no longer expressed and could no longer protect sensitive cells against ColN. The promoter region was resquenced before cloning and some sequencing mistakes were discovered (Figure 69). The differences are probably due to the sequencing method used in 1988 which is more prone to errors than the methods used today, including Sanger ABI 3730xl sequencing used by GATC Biotech Ltd in 2014. The promoter was resequenced with 3 different primers in 3 different colonies with the same result. A spot test assay (Figure 70) and a liquid culture activity assay (Figure 71) were carried out to show promoter and protein activity. Similarly to CNI expressed from *pBAD322-Para-CNI-sfGFP-6His*, CNI regulated by its natural promoter (abbreviated as npCNI from here) was able to protect the cells from ColN.

A

```
                                                BamHI
gi|41117|emb|X06933.1|_NCBI_CN      ------ACAAAAGGGTCTCAAAGAGGATCCTTTGATCTTTTACTTAATTT 44
CNI_promoter_DS24052014             CGTAGAAAAAAAAGGATCTCAAG-AGGATCCTTTGATCTTTTACTTAATTT 1049
                                          *.******.******. ***************************

gi|41117|emb|X06933.1|_NCBI_CN      CAACCC-TGTTAGTCTGGATGATGATGATGGAGCAACAGTCCCTCCTTTC 93
CNI_promoter_DS24052014             CAACCCCTGTTAGTCTGGATGATGATGATGGAGCAACAGTCCCTCCTTTC 1099
                                    ****** ******************************************

gi|41117|emb|X06933.1|_NCBI_CN      ACATCACGGATGTGATTGACCTGACATGCACTAAGTGTCATGATAAAAAA 143
CNI_promoter_DS24052014             ACATCACGGATGTGATTGACCTGACATGCACTAAGTGTCATGATAAAAAA 1149
                                    **************************************************
                                                  -35                 -10       TS
gi|41117|emb|X06933.1|_NCBI_CN      CAAGATTAATAAAATTTTCCCACACATAATTTTCAACCATCCAAAGAAAC 193
CNI_promoter_DS24052014             CAAGATTAATAAAATTTTCCCACACATAATTTTCAACCATCCAAAGAAAG 1199
                                    **************************************************
                                                                          RBS
gi|41117|emb|X06933.1|_NCBI_CN      GAACATGGATATAAAAGACAGAAATAAGATATCAAAAAAAATATCATTCA 243
CNI_promoter_DS24052014             CAACATGGATATAAAAGACAGAAATAAGATATCAAAAAAAATATCATTCA 1249
                                     **************************************************

gi|41117|emb|X06933.1|_NCBI_CN      GTCTTCTGCTCTTACTTTCCCCATTCGCATTAATATTTTTCAGTTATAAT 293
CNI_promoter_DS24052014             GTCTTCTGCTCTTACTTTCCCCATTCGCATTAATATTTTTCAGTTATAAT 1299
                                    **************************************************

gi|41117|emb|X06933.1|_NCBI_CN      AATGCACAAT--ACACTCCTCGAAAAAAT------------ 320
CNI_promoter_DS24052014             AATGCACCAATACCACTGCCTGAAAAATCATGCTACTACAC 1340
                                    *******.*:  .**** *  ******:
```

B

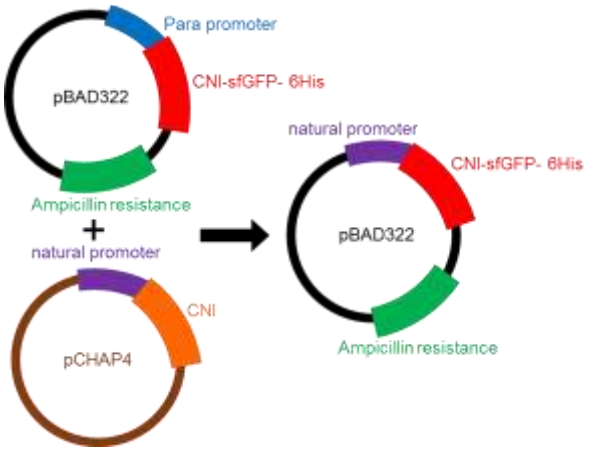


Figure 69: CNI natural promoter sequence and cloning diagram. (A) Promoter sequence alignment of Pugsley (1988), NCBI GenBank: X06933.1, and re-sequenced CNI Promoter data (2014). According to Pugsley (1988) the promoter lies between the BamHI site (green) and the start of the protein coding sequence. The new, here proposed, promoter sequence is underlined. Differences between sequences are highlighted in magenta. The proposed start of the protein coding sequence is highlighted with a yellow (Pugsley, 1988) and blue (DS, 2014) arrow and text highlight. Transcriptional elements (-35 site, -10 site and transcription start) as well as the ribosome binding site (RBS) and the start codon are indicated bold. The transcriptional elements have been determined using free bioinformatics software[11]. (B) The natural promoter of CNI was amplified from pCHAP4 and inserted in front of the CNI-sfGFP-6 gene into the pBAD322 plasmid using InFusion cloning technology. Spot test assay to show activity of the new construct npCNI-sfGFP-6His

[11]http://www.fruitfly.org/seq_tools/promoter.html and
http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb

A spot test assay was used to show that the newly created construct npCNI-sfGFP-6His was as effective in providing immunity against ColN as the previously used overexpression construct pBAD322-*Para*-CNI-sfGFP-6His and npCNI without a GFP –tag. ColA was used as a specificity control because npCNI-sfGFP should protect against ColN but not ColA. In the spot test assay, lethal ColN amounts caused clearing zones in the bacterial lawns (Figure 70). Generally, lethality depends on the concentration of colicin in the 2 µl drop, the colicin's ability to diffuse through the agar and the resistance of the cells. Therefore, clearing zones of different colicins used at the same concentration are not necessarily comparable. Cells expressing npCNI-sfGFP are protected against 600 nM Colicin N and Colicin A, while twice this concentration (1.25 µM) was lethal. Sensitive cells were also sensitive to 600 nM ColA but much lower concentrations of ColN (150 nM) were sufficient to kill sensitive cells. This 4 fold difference in potency between ColN and ColA has also been observed in spot test assays by Ridley and Lakey (2015) and is probably caused by differences in the translocation mechanism between ColA and ColN, as well as a down regulation of BtuB expression in rich media like LB. BtuB is the primary outer membrane receptor for ColA (Lundrigan *et al.*, 1987; Ravnum and Andersson, 1997; Lei *et al.*, 2011; Kim *et al.*, 2014) and its expression is regulated by vitamin B12 addition. The spot test results prove the activity of both the cloned natural promoter and the protein fusion. This assay also shows that npCNI-sfGFP activity is specific to ColN as it does not increase the immunity against ColA in comparison to sensitive cells. The activity levels of npCNI-sfGFP-6His are the same as npCNI without GFP, showing that GFP does not influence activity despite its size of ca. 28 kDa. The arabinose induced sample seems less resistant than npCNI-sfGFP-6His. This is probably due to the short induction time and because the agar does not contain arabinose or glucose so the level of protection essentially equals that of uninduced cells. This might not be an optimal assay to compare induced and naturally regulated CNI because the amount of resistant cells will depend on the timing of the pulse induction and for how long the culture is allowed to grow afterwards. WALP23-eGFP and the empty vector control show no resistance to ColN or ColA. As expected, all cultures are equally sensitive to ColA at ca. 1.25 µM, which indicates a specific interaction between CNI and ColN. CNI specifically protects against ColN to a certain extend and does not protect against a closely related protein.

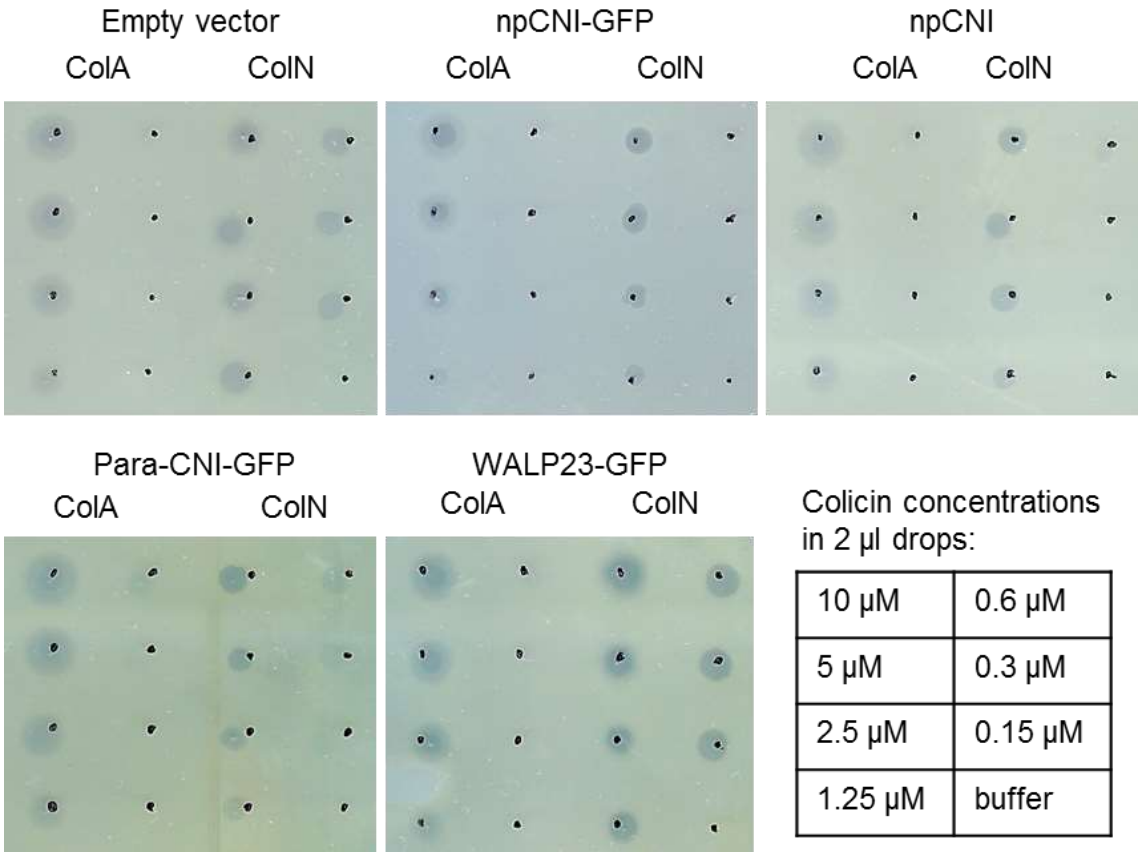


Figure 70: Naturally regulated (npCNI) and arabinose induced CNI (Para-CNI-sfGFP-6His) specifically protect MC1000 cells against ColN. Spot test assay to determine CNI specific activity against 10 µM, 5 µM , 2.5 µM, 1.25 µM, 0.6 µM, 0.3 µM and 0.15 µM ColN. Para-CNI-sfGFP-6His was induced for 10 sec with 0.2% arabinose, then repressed with 0.2% glucose. CNI does not protect against Colicin A. The GFP-tag has no impact on protein activity. WALP23-eGFP-6His does not provide protection against Colicin N.

5.2.3 Liquid culture cell survival assay demonstrates activity of the new construct npCNI-sfGFP-6His

To confirm the spot test assay results, the activity of npCNI-sfGFP-6His was also tested in liquid culture (Figure 71). When no ColN or ColA was added, all cultures (expressing CNI, WALP, TolA or not) grew similarly, meaning that membrane protein production is not significantly toxic. In liquid culture, npCNI-sfGFP-6His can protect cells up to 1 nM ColN, before culture viability is affected by an increase in lag time before log phase growth starts. This concentration is a lot smaller than what was measured on agar plates and highlights the differences in these two assays. Liquid cultures are constantly shaking, distributing the ColN evenly and allowing it to reach and kill more cells, while a drop placed on agar is more constrained. Therefore, the liquid culture assay is able to determine ColN impact on culture viability more accurately.

All MC1000 cells are sensitive to ColA above 100 pM (Figure 71D). Interestingly, even 1 µM ColN does not kill *all* CNI producing cells in liquid culture. The clearing zone created by 1.25 µM ColN in 2 µl on agar is not completely clear either although killing activity is clearly evident (Figure 70). Although the colicin makes a considerable impact in both assays, it fails to completely prevent growth. At 1 µM ColN, all empty vector control cells are killed (Figure 71C). CNI-sfGFP-6His expressed from an arabinose inducible promoter provides a higher level of protection than npCNI-sfGFP-6His. As in the spot test assay, GFP has no impact on immunity protein activity (Figure 71C).

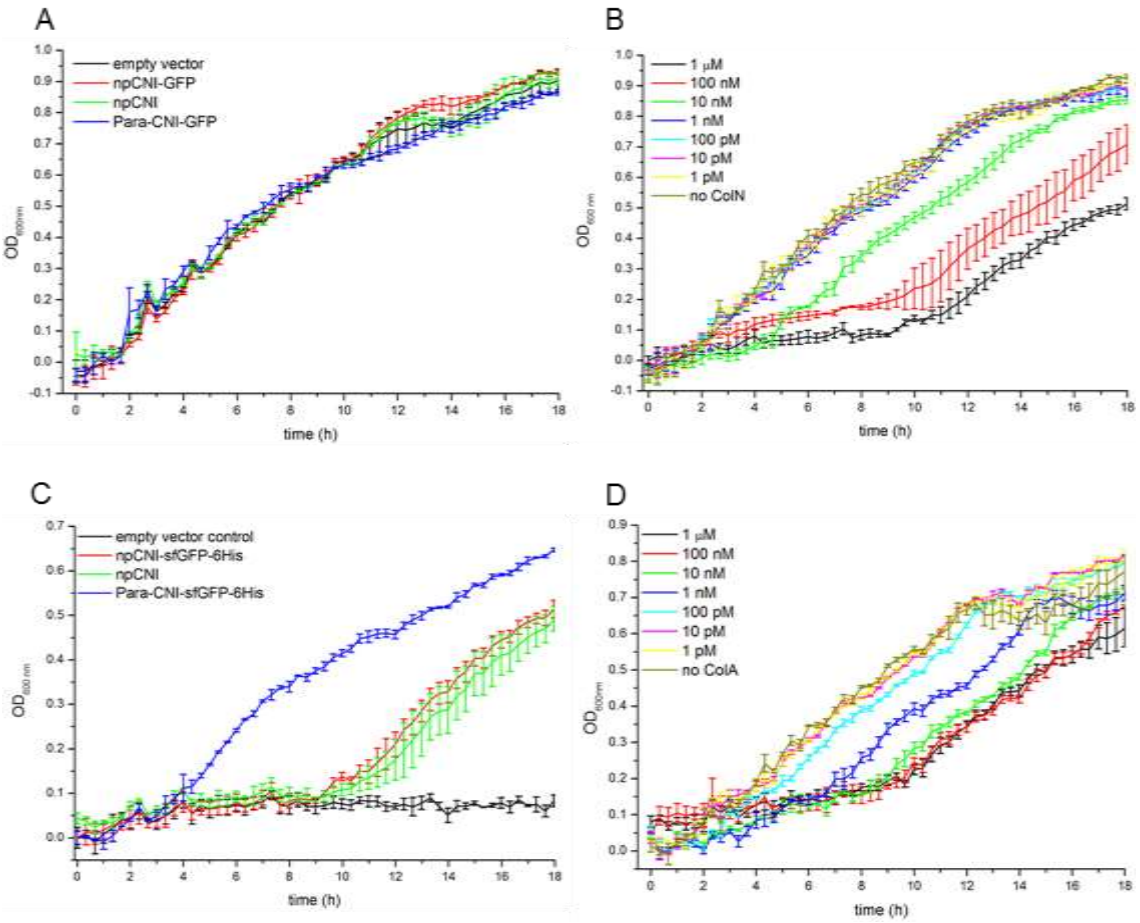


Figure 71: Activity of npCNI-sfGFP-6HIS in liquid culture. (A) When no ColN was added, all cultures grew similarly. (B) MC1000 cells expressing npCNI-sfGFP-6HIS with different Colicin N concentrations. npCNI-sfGFP-6His protects cells up to 1 nM ColN, before culture viability starts being affected. (C) At 1 µM ColN all empty vector control cells are killed. npCNI with and without GFP provides the same level of protection, while Para-CNI-sfGFP-6His provides a higher level of culture viability. (D) MC1000 cells are sensitive to ColA concentrations above 100 pM; cells expressing npCNI are shown as an example.

5.2.4 Monitoring protein production in vivo using the sfGFP fusion protein

As well as measuring resistance to ColN, protein production was also monitored using the fluorescence of sfGFP as a quantifiable reporter. Culture growth and protein production were monitored at the same time by measuring culture optical density at 600 nm and sfGFP fluorescence at 510 nm in the same culture over time. As cells have a certain level of auto-fluorescence around the 500 nm wavelength due to mildly fluorescent molecules like NADH and flavins (Bao *et al.*, 2008; Yang *et al.*, 2012), it is important to show that the increase in fluorescence at 510 nm is not only due to an increase in culture optical density. All three types of cultures have the same $OD_{600\ nm}$ (Figure 72 A), indicating that each of them contains approximately the same amount of cells. The expression of GFP-tagged npCNI, however, leads to a significantly higher fluorescence level at 510 nm than in the two controls, confirming once more the expression of the GFP-tagged CNI and the activity of the cloned natural promoter. The same method was used to show that CNI-sfGFP-6His expression and intracellular amount do not change when increasing concentrations of ColN are added (Figure 72 B), at least, not to an extent that can be measured by the plate reader. This suggests that CNI expression is not upregulated and the CNI turnover rate is not reduced as a consequence of ColN addition. It also agrees with the hypothesis of constitutive expression proposed for immunity proteins to pore-forming colicins and is a crucial prerequisite for the subsequent experiments investigating CNI diffusion rate *in vivo*.

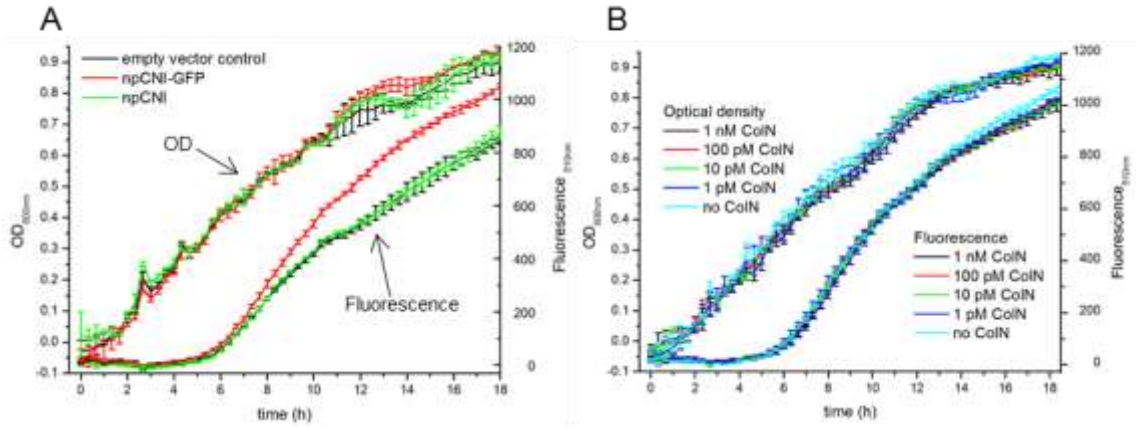


Figure 72: (A) MC1000 cells which express GFP-tagged CNI, are significantly more fluorescent than MC1000 cells that do not express CNI or the untagged version. The cloned natural promoter is therefore active. (B) The fluorescence of MC1000 cells expressing npCNI-sfGFP-6His does not increase with increased concentrations of ColN.

5.2.5 Colicin N and Colicin A have no effect on the distribution and diffusion rate of the Colicin N immunity protein

Following on from killing assays in Chapter 4 and this chapter, sections 0 and 5.2.3, where CNI activity was demonstrated through the increased viability of a culture, this section focuses on the demonstration of the interaction between CNI and ColN *in vivo* using fluorescent microscopy.

Figure 73 illustrates the difference between epifluorescent and Total Internal Reflection Fluorescence (TIRF) microscopy. In epifluorescent microscopy, light enters the cell and excites most GFP molecules within the whole cell. As a result, the cell membrane, which contains the CNI-sfGFP-6His fusion, appears brighter than the rest of the cell. It appears as if one is viewing a cross-section of the cell although actually one is looking completely through the cell. In TIRF microscopy, the laser angle is adjusted so that the light is totally reflected by the cover slide but creates an evanescent wave, a kind of energy wave, at the interface of the glass slide and the cell membrane or the agarose gel due to the change in refractive index. The angle is chosen for each slide individually so that the signal is a local maxima in intensity but the angle remains the same for repeats of the same condition which are on the same slide. The evanescent wave excites fluorophores close to the surface of one side of the cell, mostly within 200 nm from the cover slip, which allows viewing of the bacterial membrane in a quasi-plane with very little noise from the cytosol. The evanescent wave intensity decreases exponentially with increasing distance from the coverslip, so only the region of the bacterial membrane facing the laser is observed (Figure 73B). As the ends of rod shaped bacteria such as *Escherichia coli* are curved, they lie outside of the evanescent wave and cannot be seen in TIRF microscopy, but they are visible in epifluorescence microscopy. Therefore, bacterial cells in epifluorescence microscopy appear longer than in TIRF microscopy at any given magnification.

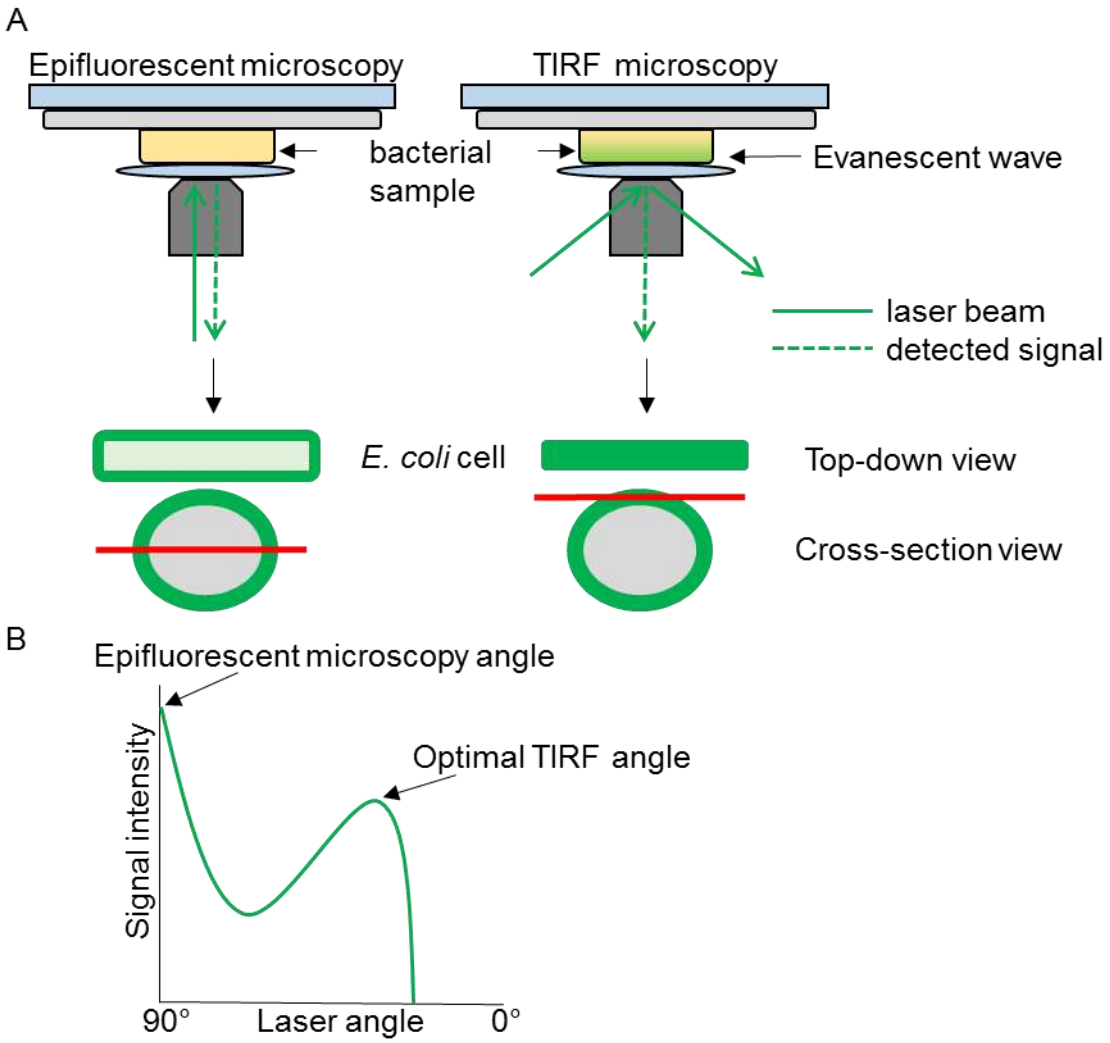


Figure 73: The difference between epifluorescent microscopy and TIRF microscopy. Laser angle adjustment (A) and signal strength (B) for epifluorescent microscopy and TIRF microscopy. In epifluorescent microscopy the fluorescent signal is stronger than in TIRF. The ridges of the cell appear brighter than the middle of the cell because it is a top-down view and the signal is mainly located in the cell membrane. In TIRF, a perfect angle is determined for each sample by adjusting until only the side of the cell nearest the slide, i.e. the membrane, is visible and signal strength reaches a local maximum (B). (C) The same field of view in epifluorescence and TIRF microscopy as an example for the difference in imaging between the methods.

Consistent with previous epifluorescence microscopy images CNI-sfGFP-6His appears to be uniformly distributed in the membrane when viewed using TIRF microscopy (Figure 74) and addition of ColN and ColA does not seem to obviously change distribution. While a drastic change in localisation cannot be observed by eye, it is still possible that the addition of ColN has an effect on CNI diffusion rate and clustering.

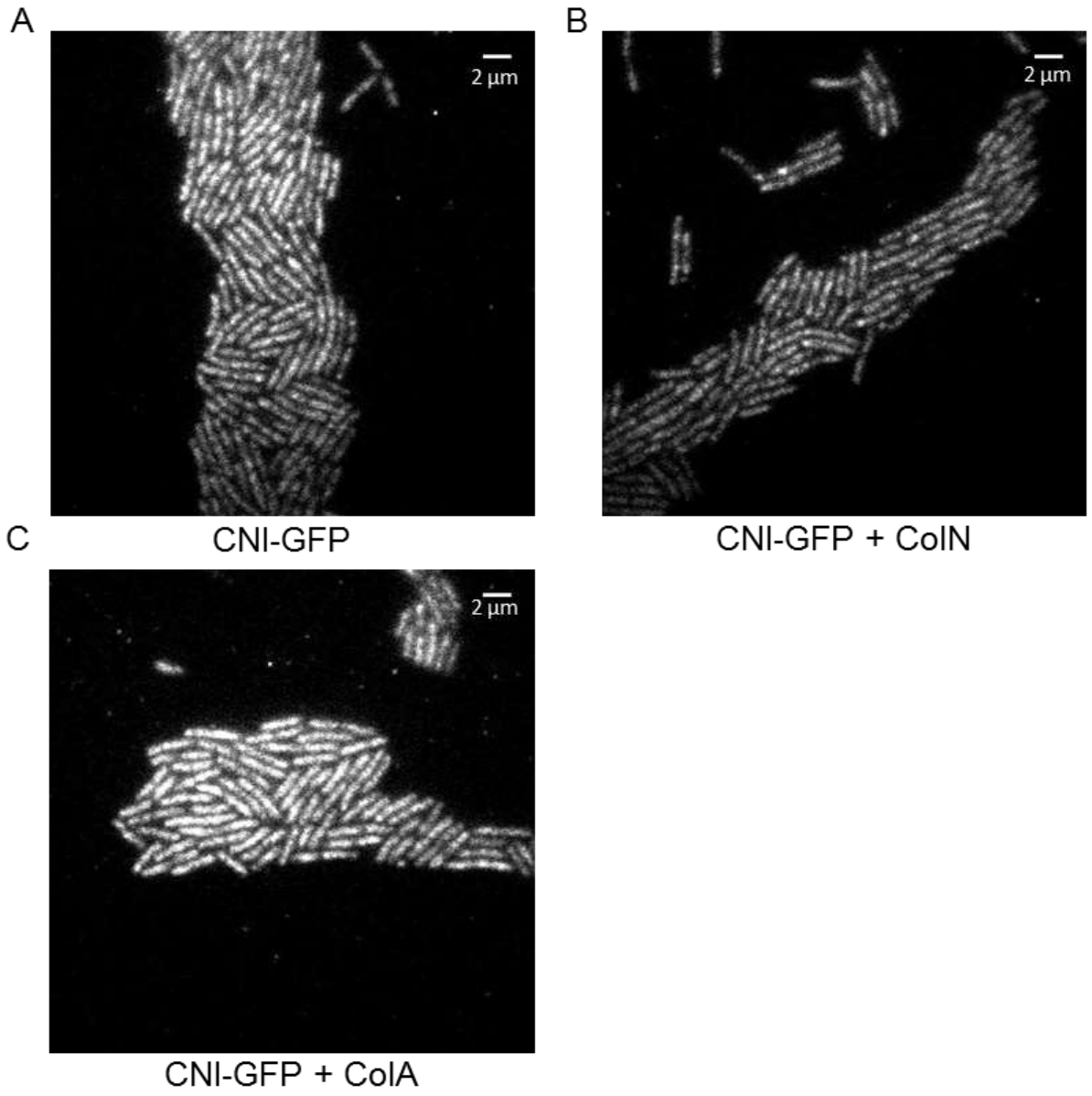


Figure 74: npCNI-sfGFP-6His distribution does no change upon addition of ColN or ColA in an obvious way. MC1000 cells expressing npCNI-sfGFP-6His.

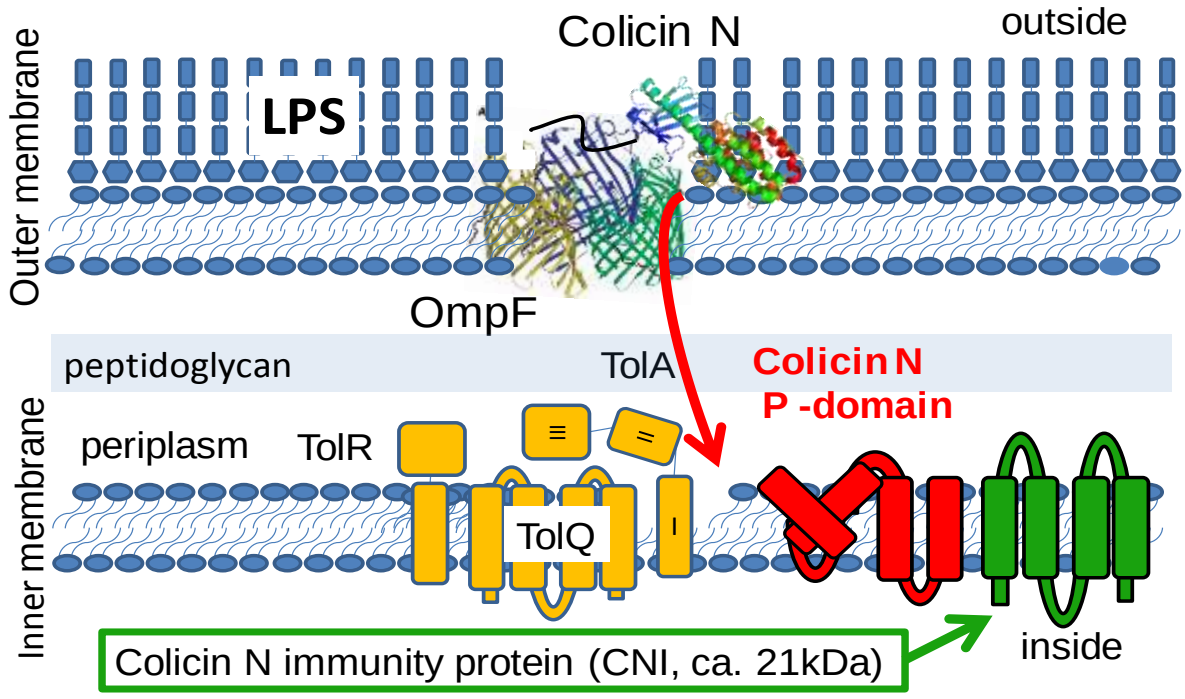


Figure 75: CNI may associate with the P-domain of ColN in the inner membrane. ColN in turn is bound to TolA through its TolA binding site (Ridley and Lakey, 2015). Figure is a reprint from Stroukova and Lakey, 2015 (in press, see appendix).

5.2.6 Investigating a change in diffusion rate using Total Internal Reflection Fluorescence Microscopy with Continuous Photobleaching (TIR-CP)

A change in diffusion rate can be tested by how fast a bleached site is replenished by unbleached molecules from other parts of the membrane. This is a principle used by fluorescence recovery after photo bleaching (FRAP) in eukaryotic cells (Reits and Neefjes, 2001; Rayan et al., 2010) but is difficult to apply to bacteria because of their small size, because a laser applied in epifluorescence mode would illuminate and bleach the entire bacterial cell rather than just a small portion. Therefore, bleaching kinetics were observed instead using Total Internal Reflection Fluorescence Microscopy with Continuous Photobleaching (TIR-CP), a method described by Slade *et al.* (2009) and used by Strahl et al. (2015) to investigate the diffusion rates of membrane proteins. The model to quantify the bleaching kinetics is defined in the appendix.

Here, TIR -CP was used to determine whether ColN addition would, firstly, slow down the overall diffusion rate of npCNI-sfGFP-6His through the membrane, and, secondly, induce CNI clustering around ColN entry sites. Colicin A and WALP23-eGFP were used as controls to highlight specific CNI-ColN interactions and to exclude any general effects the addition of a pore-forming colicin might have on the *E. coli* inner membrane.

Cells were grown in at 37 °C until $OD_{600\ nm}$ = 0.5 before aliquots were taken and treated with sub-lethal concentrations of Colicin N or Colicin A (both at 100 pM). WALP23-eGFP cultures were induced with 0.2 % arabinose 2 h prior to harvest and colicin treatment. npCNI-sfGFP-6His expression was not artificially induced. A drop of 0.5 µl of each culture was fixed on 1.2 % agarose on microscope slides and imaged over time at 100 ms per frame. Using the image processing software Image J, all images in the stack, which is composed of a series of single pictures taken 100 ms apart, were treated for background subtraction and then the overall signal intensity of each image was measured and plotted against time. The fluorescence intensity decays over time as more fluorophores are bleached than can be replenished. The fluorescence intensity decay was expressed in percentage, normalising the initial fluorescence intensity at the start of the measurement as 100 %. This allows the comparison of repeats of the same kind of condition, even if the stacks do not contain the same amount of cells and do not have the same absolute intensity. The curves in Figure 76 were fitted as a two phase exponential decay as per Slade *et al.* (2009) and Strahl *et al.* (2014), assuming that initial intensity = 100 %, the curves decay to 0 % and the initial fast decay rate is the same. Table 10 summarises how well the curves fit a two phase decay, giving the fast (K_{fast}, initial bleaching of resident molecules) and slow (K_{slow}, diffusion limited bleaching of molecules entering the TIRF field from the diminishing general pool), the standard error of the curve fit for both phases, the degrees of freedom and the goodness of fit (R^2). The goodness of fit values are very high, meaning that the curves are very likely to follow a two-phase decay. The model to quantify the bleaching kinetics is defined in the appendix.

Both npCNI-sfGFP-6His and WALP23-eGFP-6His bleaching curves can be fitted to a two-phased fluorescence decay when photo bleaching at 488 nm (Figure 76 and Table 10). The first phase is dominated by rapid bleaching of all GFP molecules already at one side of the bacteria and any background fluorescence. This step looks largely the same for both CNI-sfGFP-6His and WALP-eGFP-6His, although the GFP versions (sfGFP and eGFP) are different and therefore have different bleaching kinetics and fluorophore stability (Pedelacq *et al.*, 2006; Strahl *et al.*, 2014). When trying to fit the bleaching kinetics to a two-phase decay this was set to be the same/global (Table 10). During the second part, proteins which replenish the bleached area from other parts of the membrane are mainly bleached. Here, bleaching is limited by the diffusion efficiency of new GFP molecules entering the area visible in TIRF mode. The bleaching kinetics of this second part give the

best indication of how fast a protein diffuses because the speed of bleaching is inversely related to how fast this area is replenished. If molecules move fast into the bleached area, the area is quickly replenished and the overall bleaching is slower. If molecules move slower, the area is replenished slower and so the area is bleached faster. At the Colicin N and A concentration used here the diffusion rate of npCNI-sfGFP-6HIS and WALP23-eGFP-6His does not change significantly.

Furthermore, it appears as though CNI-sfGFP-6His bleaches slower than WALP23-eGFP-6His and therefore moves faster in the membrane. However, a comparison between the diffusion rates of WALP23-eGFP-6His and CNI-sfGFP-6His might not be meaningful because the speed of bleaching is not only determined by the diffusion rate of the proteins but also by their size and their abundance in the membrane, which is determined by protein expression levels and stability. The higher the protein density in the membrane, the faster a bleached area will be replenished. Expression levels are responsible for how well a membrane is replenished with proteins from the cytosol and determines to a certain extent how many molecules in total there are in the membrane and, therefore, how fast they move into the bleached spot. The bleaching happens much faster than the replenishing from the cytosol, so here the differences between the expression levels are irrelevant. However, if in a given space there are, in total, more protein molecules, the void is replenished faster than if there are fewer protein molecules, even if they have the same diffusion rate. Standardising to the initial fluorescence intensity to 100 % is a way to determine a starting and an end point. While a comparison between ColA, ColN or no colicin addition can be made, it cannot necessarily be made between CNI-sfGFP-6His and WALP23-eGFP-6His. Proteins are only comparable if one assumes similar expression levels, equally efficient transport to the membrane and equal stability in the membrane. These assumptions are very unlikely to be true for CNI and WALP23 because CNI-sfGFP-6His is a functioning protein regulated by its natural promoter and with potential interaction partners in the membrane, while WALP23-eGFP-6His is an artificial transmembrane helix which is transcribed from an arabinose inducible promoter and has no function. Therefore, WALP23-eGFP-6His could appear slower than CNI-sfGFP-6His if it is less abundant or less stable, even if it has the same or a faster diffusion rate. Also, the GFP versions used here are different. Superfolder GFP (sfGFP) is much more photostable than enhanced GFP (eGFP) and therefore has different bleaching kinetics (Pedelacq *et al.*, 2006; Strahl *et al.*, 2014).

Furthermore, in contrast to my previous assumption, WALP23-eGFP-6His might not be a suitable control for random movement as its distribution appears very coarse in Figure 68, as if it forms random clusters with itself. Why these supposed clusters are formed and whether the clusters are permanent and caused by WALP23-eGFP-6His interaction with lipids or something else has not been investigated further.

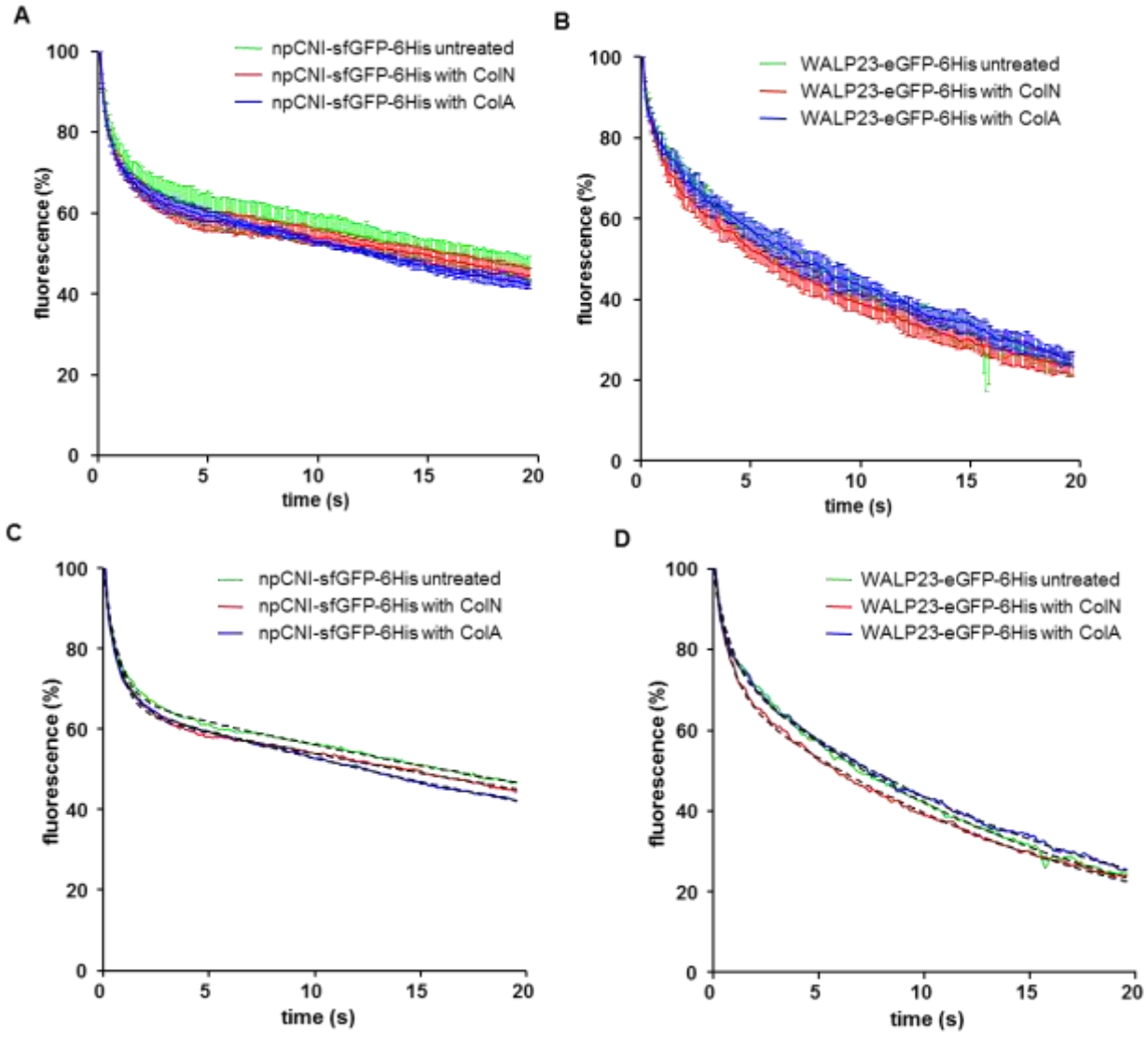


Figure 76: Fluorescence decay of CNI-sfGFP-6His is not significantly affected by the addition of ColN and ColA. MC1000 cells expressing npCNI-sfGFP-6His (A and C) and WALP23-eGFP-6His (B and D) treated with 1 nM ColN and 100 pM ColA. (A) and (B) show standard error, (C) and (D) show superimposed and fitted model of two phase exponential decay (dashed black line). Images taken every 100 ms, normalised to the same background and assessed for overall intensity over time, where fluorescence at t = 0 was equalised to 100 %. Fluorescence intensity corresponds to average intensity of all cells in one image over time (in sec). For each treatment 4 fields of view were averaged. All fields of view contain approximately the same number of cells. The model to quantify the bleaching kinetics is defined in the appendix.

	WALP untreated	WALP + ColN	WALP +ColA	Global (shared)	CNI untreated	CNI + ColN	CNI + ColA	Global (shared)
Best-fit values								
Kfast in 1% fluorescence/sec (bleaching)	1.337	1.337	1.337	1.337	1.452	1.452	1.452	1.452
Kslow in 1% fluorescence/sec (diffusion limited)	0.061	0.059	0.056		0.019	0.019	0.023	
Std. Error								
KFast	0.023	0.023	0.023	0.023	0.014	0.014	0.014	0.014
KSlow	0.00028	0.00031	0.00027		0.00015	0.00016	0.00016	
95% Confidence Intervals								
KFast	1.292 to 1.383	1.292 to 1.383	1.292 to 1.383	1.292 to 1.383	1.425 to 1.479	1.425 to 1.479	1.425 to 1.479	1.425 to 1.479
KSlow	0.060 to 0.062	0.058 to 0.060	0.055 to 0.056		0.019 to 0.020	0.018 to 0.019	0.023 to 0.024	
Goodness of Fit								
Degrees of Freedom				581				581
R^2	0.997	0.998	0.998	0.998	0.996	0.995	0.996	0.996

Table 10: Summary of two-phase decay fit values for WALP23-eGFP-6His and CNI-sfGFP-6His. Assuming the first bleaching phase (K_{fast}) is the same for WALP23-eGFP-6His and CNI- sfGFP-6His and fluorescence intensity bleaches from 100 % to 0 %, all curves fit a two-phase decay with high probability. K_{fast} indicates initial bleaching of GFP molecules already present in the TIRF area. K_{slow} indicates bleaching which is limited by the diffusion rate of the protein and therefore is an indirect estimate of molecule diffusion rate. K_{fast} and K_{slow} are given as 1 percent of fluorescence per second and are measured over the entire 20 sec. The model to quantify the bleaching kinetics is defined in the appendix.

5.2.7 CNI does not form clusters upon ColN addition

ColN requires the Tol-Pal system for translocation and so ColN molecules may attract CNI molecules to its site of entry, i.e. TolA, in the inner membrane (Figure 75). TolA is a slow moving protein itself and slows down even further with the addition of ColN and ColA (Figure 91, Figure 92, Table 11, Table 12). Therefore, if npCNI-sfGFP-6His becomes attracted by and attached to a TolA-bound ColN, this may lead to a reduction in movement and clustering of npCNI-sfGFP-6His molecules around the ColN entry site.

Cells were prepared as previously but imaged at 30 ms per frame. For each cell, a fluorescence intensity profile was created in ImageJ across the length of the cell (Figure 77B). The mean intensity and variance of each trace was calculated. The mean represents how much fluorescence is present in each cell and the variance indicates whether this fluorescence is evenly distributed along a line or fluctuates greatly around this mean which would indicate clustering (Figure 77A). As overall fluorescence intensity does not change upon ColN addition (Figure 72), the mean fluorescence in each treatment must be the same (Figure 77 B). If CNI became more clustered upon Colicin N addition a difference in variance would be apparent (Figure 77C). A cell with clusters would have a higher variance around the mean than a cell where npCNI-sfGFP-6His molecules are not clustered but evenly distributed. Figure 77C shows that all means and variances do not differ significantly in all three conditions. Hence, npCNI-sfGFP-6His does not form detectable clusters around potential entry sites or other locations upon addition of Colicins N and A, and, also, does not form randomly moving, unlocalised clusters (Figure 77).

In conclusion, npCNI-sfGFP-6His is evenly distributed in the cell membrane, while most of TolA is localised at the septum during cell division. Increasing Colicin N addition does not induce clustering of npCNI-sfGFP-6His in the membrane and does not increase the overall amount of molecules in a cell. npCNI-sfGFP-6His has a high diffusion rate in the cell membrane and is not slowed down significantly when Colicin N or Colicin A are added.

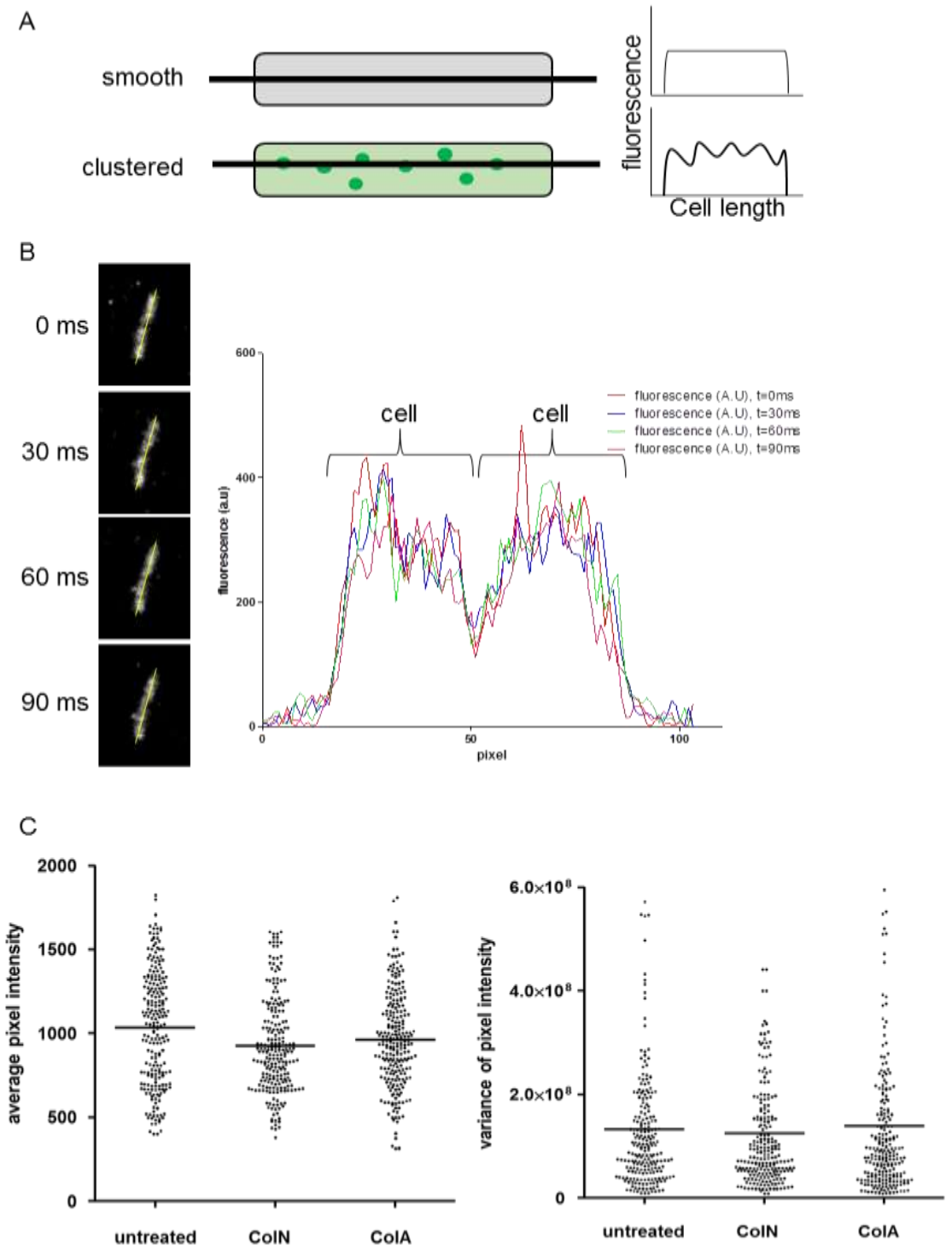


Figure 77: npCNI-sfGFP-6His does not cluster when ColN or ColA are added. (A) Schematic demonstration how protein clustering changes the detection of florescence intensity across the length of the cells. From each trace the average pixel intensity and the variance of pixel intensity are calculated. If clustering occurs, the variance will increase while the average pixel intensity remains the same. (B) The fluorescence intensity was measured along a three pixel line down the length of the cell, resulting in intensity plots. As npCNI-sfGFP-6His molecules move randomly in the cell, the signal is also fluctuating randomly. (C) The average pixel intensity of each intensity plot is represented as a dot. There is no significant difference between the means of average pixel intensities and there is no significant difference between the means of variances of pixel intensities. The number of cells analysed for each condition was 219 for untreated, 226 for ColN treated, 230 for ColA treated.

5.2.8 Single molecule tracking of Para-CNI-sfGFP-6His in the E. coli inner membrane

To directly quantify the diffusion rate of CNI and investigate the *in vivo* effect of ColN addition on single CNI-SFGFP-6HIS molecules, single molecule tracking was used to record the movement of Para-CNI-sfGFP-6His molecules in the inner membrane. Detecting and recording the movement of individual molecules is a direct measure of mobility in contrast to an indirect method like FRAP or the bleaching method described earlier (section 5.2.5, Chapter 5). Most importantly, it allows the comparison of different proteins with different expression levels and properties as well as the comparison of different GFP versions with different fluorophore stabilities (Pedelacq *et al.*, 2006), which is not possible with the previously described bleaching method. The Central Laser Facility at Research Complex Harwell, Oxford, provides the opportunity to image and track the single protein molecules in live cells. The principle of single molecule tracking is to take 400 - 1000 high resolution images of the same set of cells with a very high frame rate. Then complex image analysis software is used to identify single sources of fluorescence as single moving molecules and track their displacement from frame to frame. The software analyses where the source of light is in comparison to where it was in the previous image and connects the locations, creating a track. The length of each track is dependent on the durability of the signal, which in turn depends on the stability of the fluorophore, the laser strength and how long the molecules remains in the field of view. From the recording of a particular track in a given time, the mean square displacement can be calculated, from which the average diffusion rate of the molecule, given in μm^2/sec, can be calculated. The main challenge of this method is to achieve a high enough spatial and temporal resolution. For the software to track each single molecule accurately the sources of fluorescence need to be spaced far enough apart. Otherwise, the software will not be able to determine from frame to frame if a source of light is the continuation of a track or a separate molecule. If the equipment permits it, the frame rate should be as fast as possible to achieve temporal resolution, because this provides more data points within a track and makes the tracking more precise. This type of imaging has repeatedly been carried out with eukaryotic cells but bacteria, due to their small size (ca. 1 µm x 2 µm for *E. coli*), still remain a challenge. The software has been developed by the CLF team and a more in depth description of the software can be found in Rolfe *et al.* (2011). My part in this work was to prepare cells expressing GFP-tagged CNI, TolA and WALP23, label the

colicins with CF^{TM}640R fluorescent dye, treat the MC1000 *E. coli* cells with labelled ColN or ColA, prepare the samples for imaging and record the images. Image analysis and MSD calculation have been carried out by Dr Daniel J. Rolfe, CLF, Harwell.

5.2.9 Protein labelling

The imaging facility at CLF allows for the simultaneous imaging of up to 3 different fluorophores, so initially, it was intended to carry out a co-localisation experiment with ColN and CNI as well as measuring the diffusion rate of CNI for which only one channel would be needed. As CNI was already fused to GFP, a red dye (CF^{TM}640R) was chosen for ColN and ColA. This is a commercial dye which absorbs at 642 nm and emits at 662 nm and is sold with different linking groups and so can be used to label proteins through thioester bonds to cysteine residues or ester bonds to amines. ColN was labelled via a maleimide reaction in either one or two cysteine residues at positions 3 or 94 and 179 (Figure 80), while ColA was labelled with CF^{TM}640R succinimidyl ester via its primary amines, mostly lysine residues.

ColN does not naturally contain cysteine residues and so mutants were used, which were created using site-directed mutagenesis by Dr C. L. Johnson, a previous member of the Lakey group, for a different experiment. These existing plasmids were used to express the single mutant (S3C) and double mutant (V94C/Q179C) and then purified (Figure 79) and labelled (Figure 80 C). The dye absorbs at 280 nm as well as around 600 nm and 640 nm. 2 mg of S3C (concentration of 4.2 mg/ml = 91.91 mM) and 3 mg of V94C/Q179C (concentration of 7 mg/ml = 163.28 µM) were labelled.

>gi|116070|sp|P08083.1|CEAN_ECOLX RecName: Full=Colicin-N

MGSNGADNAHNNAFGGGKNPGIGNTSGAGSNGSASSNRGNSNGWSWSNKPHKNDGFH SDGSYHITFHGDNNSKPKPGGNSGNRGNNGDGASAKVGEITITPDNSKPGRYISSNPEYSL LAKLIDAESIKGTEVYTFHTRKGQYVKVTVPDSNIDKMRVDYVNWKGPKYNNKLVKRFVS QFLLFRKEEKEKNEKEALLKASELVSGMGDKLGEYLGVKYKNVAKEVANDIKNFHGRNIR SYNEAMASLNKVLANPKMKVNKSDKDAIVNAWKQVNAKDMANKIGNLGKAFKVADLAIK VEKIREKSIEGYNTGNWGPLLLEVESWIIGGVVAGVAISLFGAVLSFLPISGLAVTALGVIGI MTISYLSSFIDANRVSNINNIISSVIR

Figure 78: Cysteine mutations in ColN at positions 3, 94 and 179. Highlighted residues have been mutated to cysteines. S3C is single mutant in the ColN T-domain. V94C/Q179C is a double mutant in the ColN R-domain.

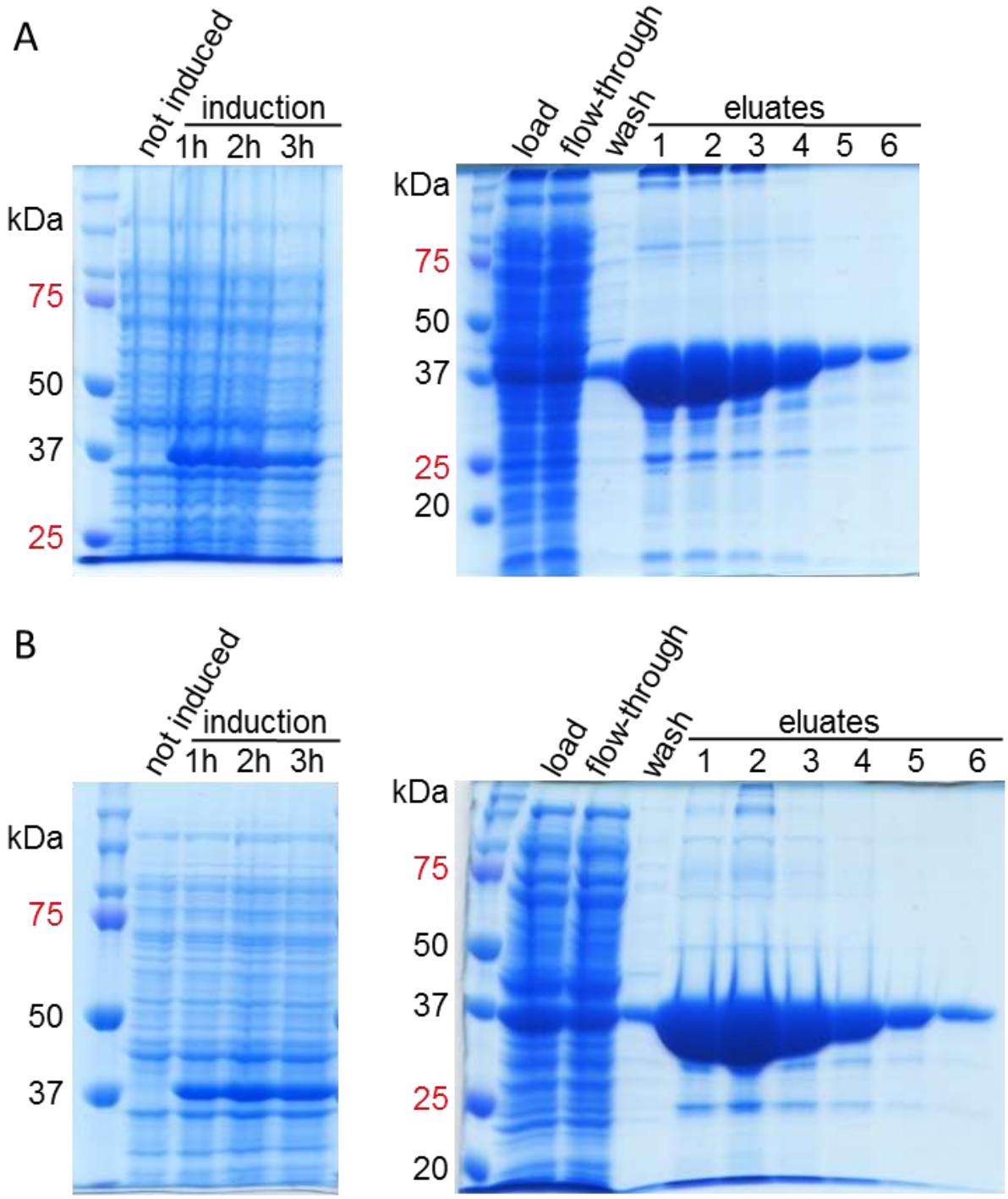


Figure 79: ColN mutant expression and purification profiles. (A) SC3; (B) V94C/Q179C. Coomassie stained SDS-PAGE gels show how both mutants are expressed by BL21-AI cells one hour after induction with 0.2 % arabinose. Purification fractions demonstrate successful purification.

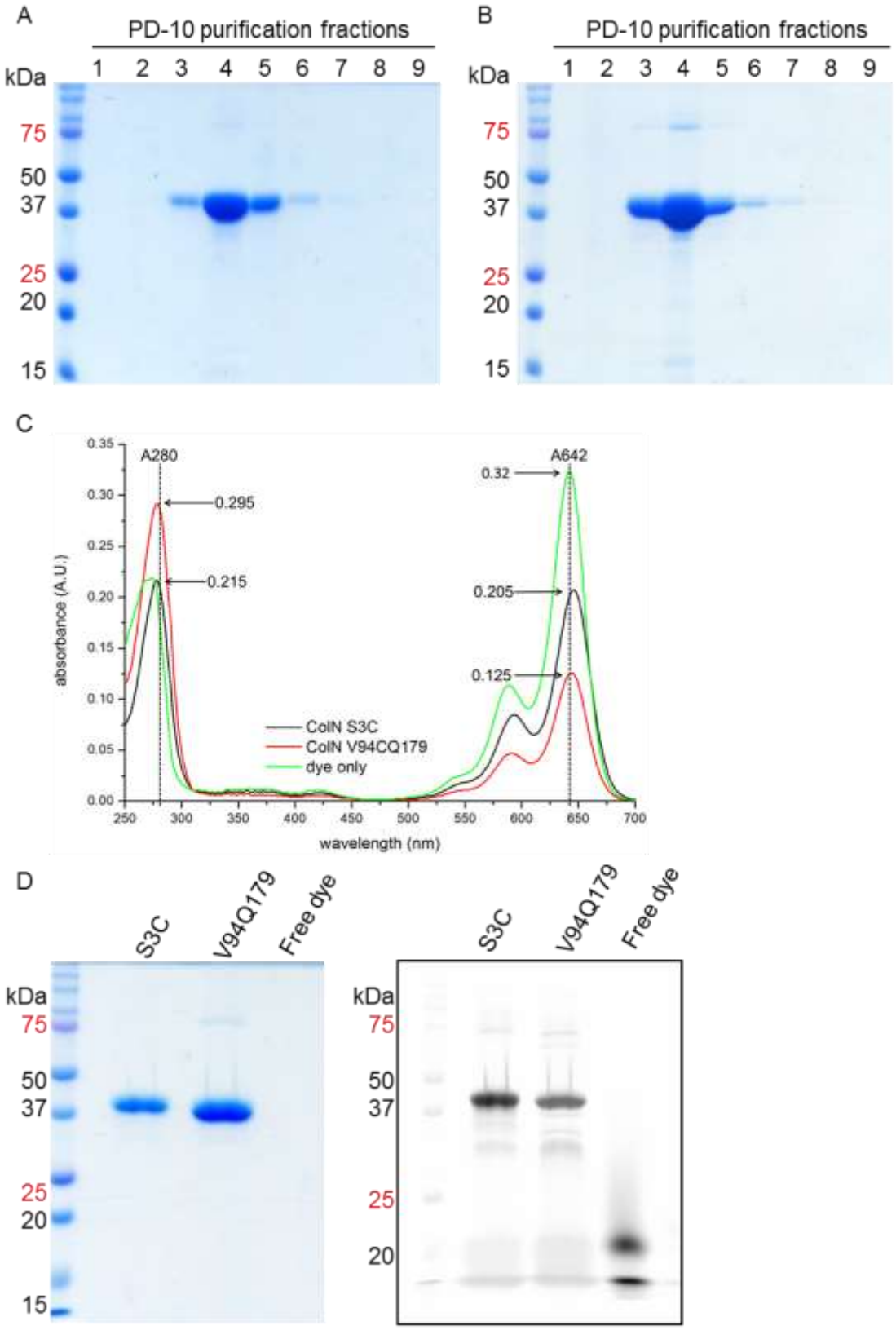


Figure 80: Colicin N mutant protein labelling. Both mutants were labelled with CF^{TM}640R maleimide via a thioester and separated from the free dye using a PD-10 column ((A) S3C, (B) V94C/Q179C). (C) An absorbance spectrum was measured between 250 nm – 700 nm to calculate labelling efficiency. (D) Typhoon imaging of fluorescent labelled protein and free dye at 640 nm (right). The free dye is invisible in the Coomassie-stained images (left).

ColA for which I do not have cysteine mutants (purified for Chapter 4, Figure 54), was labelled on exposed Lysine residues with CF^{TM}640R succinimidyl ester and excess dye was removed with a PD-10 column. The labelling success was evaluated with a Coomassie-stained SDS-PAGE gel and in-gel fluorescence imaging (Figure 81). In total, 4 mg of ColA (concentration of 2.1 mg/ml) were labelled. Even before the gel is stained with Coomassie dye, the labelling with CF^{TM}640R is visible within the gel (Figure 81). The loading buffer dye front is on the bottom while the labelled protein is at ca. 63 kDa, the expected size for ColA. Fluorescence is confirmed through in-gel fluorescence imaging. A free CF^{TM}640R dye front is at the bottom of the gel while fluorescent ColA is separated at its predicted size. For subsequent experiments purification fraction 2 was used as it has the least amount of free dye.

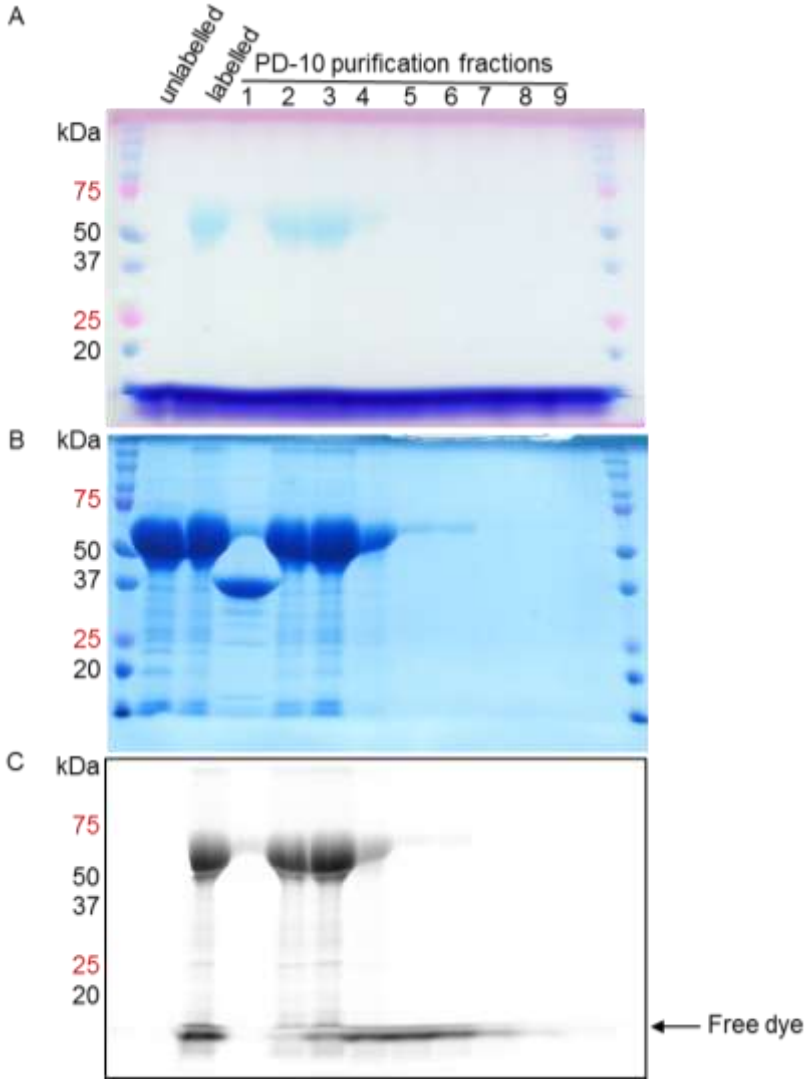


Figure 81: ColA (ca. 63 kDa) labelling with CF^{TM}640R Succinimidyl ester via its lysine residues. (A) SDS-PAGE gel before staining. Labelled protein is visible. (B) Coomassie-stained SDS-PAGE gel. All protein is visible. (C) In-gel fluorescence imaged with Typhoon imager. Labelled protein and free dye front are visible.

5.2.10 Protein activity

None of the ColN Cys mutations used for labelling are located in the P-domain, the part which causes cytotoxicity and is most likely to interact with CNI (see Chapter 1: Introduction), so it is assumed that this protein assumes the same pore conformation in the inner membrane as the wild-type protein. With the mutations located in the translocation and receptor binding domain the small dye molecule may interfere with binding to the receptor and translocator protein OmpF (Johnson *et al.*, 2013) or the LPS layer in the outer membrane (Johnson *et al.*, 2014) or the translocation partner TolA (Gokce *et al.*, 2000). Therefore, the activity of the labelled proteins was tested using a spot test assay (Figure 82).

Unlabelled ColN kills sensitive MC1000 cells and MC1000 cells expressing npCNI-sfGFP-6His. MC1000 cells which produce npCNI-sfGFP-6His are less sensitive to ColN than cells which do not. Both labelled ColN mutants kill MC1000 *E. coli* cells expressing TolA-eGFP-6His and npCNI-sfGFP-6His as well as ΔTolA *E. coli* cells (strain: JC207) complemented with TolA-eGFP-6His. Hence, the CF^{TM}640R label does not affect the activity of either labelled ColN mutants. It also does not affect ColN interaction with CNI, because CNI is still able to immunise cells against both mutants. When using the same concentration of ColN, the double mutant seems to be more active than the single mutant. This may be because the single mutation is in the ColN OmpF binding site (Johnson *et al.*, 2013) and, therefore, this may affect translocation. For subsequent single molecule tracking experiments, the double mutant was used as its activity corresponds more closely to the WT ColN. The activity of labelled ColA activity was also tested using a spot test assay (Figure 83). CF^{TM}640R labelled ColA is just as active as unlabelled ColA on MC1000 cells expressing TolA-eGFP-6His and npCNI-sfGFP-6His.

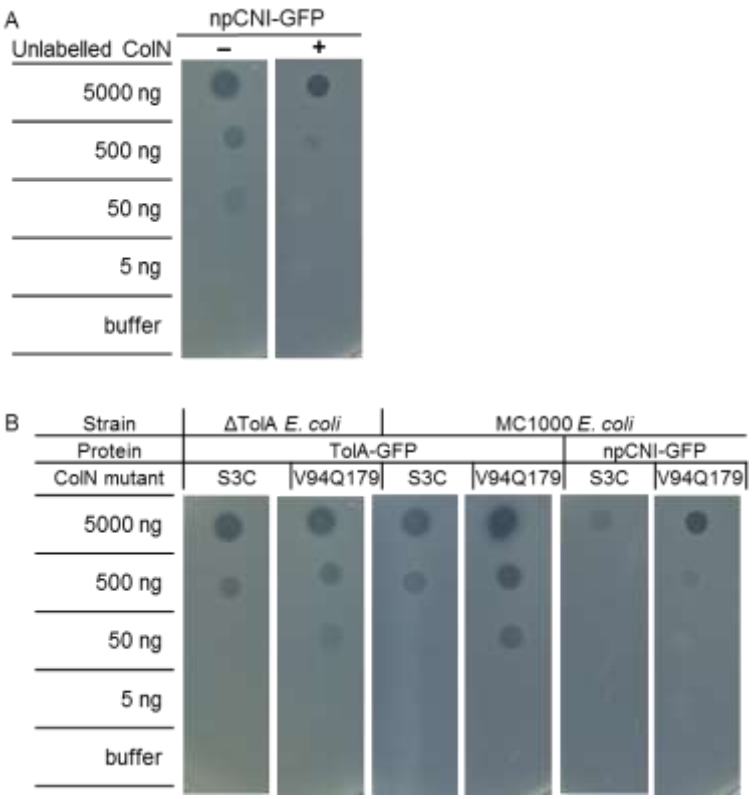


Figure 82: Labelled ColN mutants are as active as unlabelled ColN. Spot test assay testing the expression of TolA-eGFP-6His and npCNI-sfGFP-6His and the activity of labelled ColN mutants. (A) Unlabelled ColN kills sensitive MC1000 cells and MC1000 cells expressing npCNI-sfGFP-6His. (B) MC1000 cells which produce npCNI-sfGFP-6His are less sensitive to ColN than cells which do not. Both labelled ColN mutants kill MC1000 *E. coli* cells expressing TolA-eGFP-6His and npCNI-sfGFP-6His as well as ΔTolA *E. coli* cells (strain: JC207) complemented with TolA-eGFP-6His.

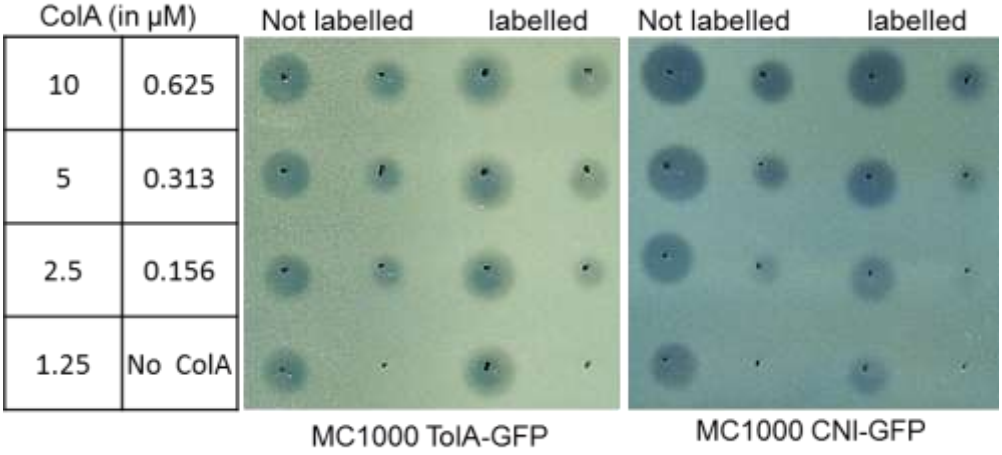


Figure 83: Labelled Colicin A activity assay. Labelled ColA is just as active as unlabelled ColA on MC1000 cells expressing TolA-eGFP-6His and npCNI-sfGFP-6His, demonstrating that ColA labelling with CF^{TM}640R succinimidyl ester via its lysine residues does not interfere with the colicin's lethal activity.

5.2.11 TIRF imaging optimization

In order to image the movement of Para-CNI-sfGFP-6His, WALP23-eGFP-6His and TolA-eGFP-6His, TIRF microscopy was used as it has advantages over epifluorescent microscopy for imaging membrane proteins (previously described in this chapter, section 5.2.5). Para-CNI-sfGFP-6His, WALP23-eGFP-6His and TolA-eGFP-6His are all expressed from an arabinose inducible promoter and were induced with 0.2 % arabinose. As previously, Para-CNI-sfGFP-6His was induced with 0.2 % arabinose only for 10 sec and then suppressed with 0.2 % glucose for 2 h to prevent overexpression. Subsequently, a series of 1000 pictures was taken with 60 ms exposure time for each image and a software designed by Dr Daniel J. Rolfe, Harwell Campus, Oxford, tracked the displacement (D) of each molecule from picture to picture. An example of how the software detects the single molecules in the image is shown in Figure 85.

Finally, knowing how many pixels each molecule has travelled within a given time, a mean square displacement curve (MSD) can be calculated. This MSD is the average squared displacement (position difference) as a function of time interval between all pairs of points within a molecule track. MSD curves were produced pooling measurements for each molecule separately, and also pooling all measurements from all tracks for a particular condition (Figure 87 - Figure 94). The gradient at the start of this curve is a measure of the diffusion coefficient on short timescales (instantaneous D). For clarification, the working process is depicted in a simplified form in Figure 84 and Figure 85.

The MSDs of Para-CNI-sfGFP-6His, TolA-eGFP-6His and WALP23-eGFP-6His were measured, with and without ColN and ColA addition. Similarly to previous experiments it was hypothesised that the addition of ColN would change the diffusion rate of CNI, i.e. increase the MSD, while ColA might not. TolA-eGFP-6His and WALP23-eGFP-6His were used as the positive and negative control, respectively. TolA-eGFP-6His should bind both ColN and ColA, while WALP23-eGFP-6His should bind neither but highlight any general and unspecific effects which colicin addition might have on inner membrane proteins, e.g. through the dissipation of membrane potential. ColA was the specificity control for ColN, as it should behave similarly towards WALP23-eGFP-6His, TolA-eGFP-6His and the cell membrane but not be neutralised by Para-CNI-sfGFP-6His.

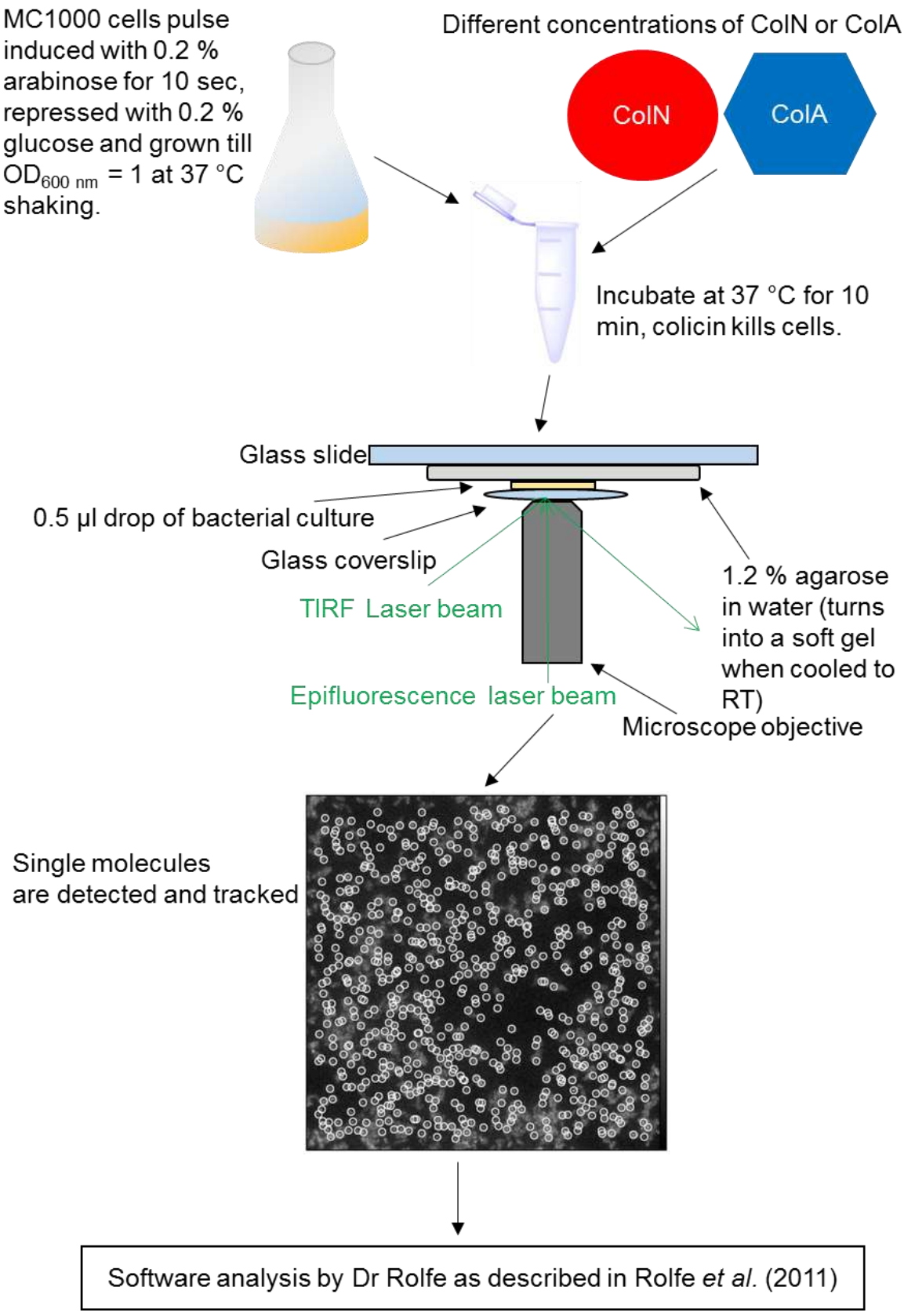


Figure 84: Simplified representation of steps carried out in preparation for and during single molecule tracking. Rolfe *et al.* (2011)

5.2.12 Co-localisation of Para-CNI-sfGFP-6His and ColN is unfeasible with the current methodology

While optimising imaging conditions, I discovered that it was possible to image ColN attachment to the cells in epifluorescence but not in TIRF mode (Figure 86). ColN not only sticks to the cells but also to the coverslip which was used to immobilise the cells. When viewed in TIRF ColN is evenly and randomly distributed on the slide and not necessarily attached to the cells and so any attempt to co-localise ColN with CNI would result in unreliable data. Co-localisation relies on the non-random distribution of both molecules, i.e. their dependence on each other. Two molecules can only be considered co-localized when they are found together more often than by chance. This condition is not achieved due to the random localisation of ColN on the cover slip as well as the cells and, therefore, co-localisation was not carried out. It is not feasible to carry out co-localisation in epifluorescent mode as the whole cell, not just one side of the membrane, is visible in this mode. Molecules which are separated in the 3rd dimension (z-axis) in the 3D space of a cell may falsely appear as co-localised in the 2D plane (x- and y – axes) of an image.

Furthermore, I discovered that extensive bleaching is needed to image overexpressed proteins, in particular Para-CNI-sfGFP-6His, because it is otherwise impossible to visually and technically separate one protein molecule from another and therefore track it. However, extensive bleaching may lead to false negative results. Although some Para-CNI-sfGFP-6His molecules may be bleached and invisible, they could still be in the membrane and may still interact with labelled ColN. Equally, visible Para-CNI-sfGFP-6His molecules may interact with invisible, unlabelled ColN molecules, because the labelling efficiency is not 100 %. Both cases would not lead to a successful detection of a co-localisation. It was therefore decided not to pursue co-localisation further but to focus on detecting a change in diffusion rate/MSD. Even if the bleaching could be avoided through a reduction in expression levels, which has already been attempted here, the inner membrane is a crowded space. Therefore, it is difficult to balance between having too many molecules in the membrane causing random collision and difficulties in separating single molecules and not having enough molecules to calculate a statistically significant co-localisation.

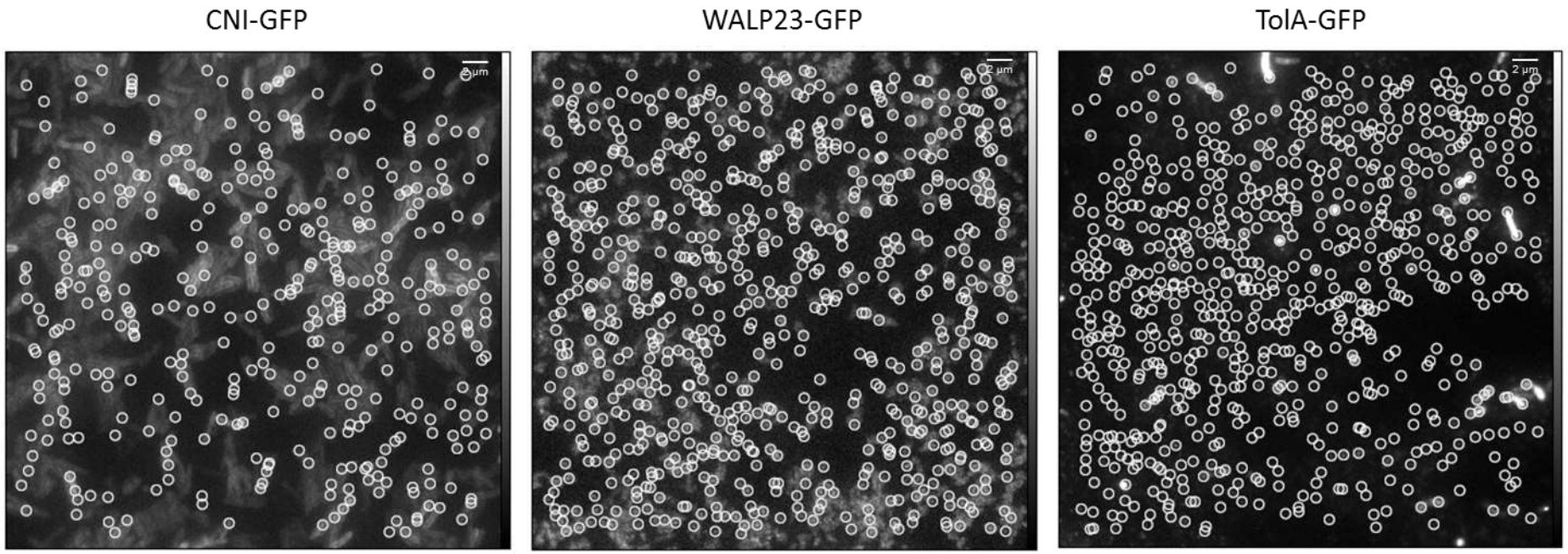


Figure 85: Examples of TIRF images of Para-CNI-sfGFP-6His, WALP23-eGFP-6His and TolA-eGFP-6His without colicin addition. Detected single molecules (circled) were tracked over time and used to calculate mean square displacement. Still images of Para-CNI-sfGFP-6His, WALP23-eGFP-6His and TolA-eGFP-6His expressed in MC1000 cells with ColN or ColA appear very similar and are not shown here as only tracking through time using a software reveals a difference. It is particularly difficult to detect single molecules for Para-CNI-sfGFP-6His and extensive bleaching is necessary to bleach background fluorescence.

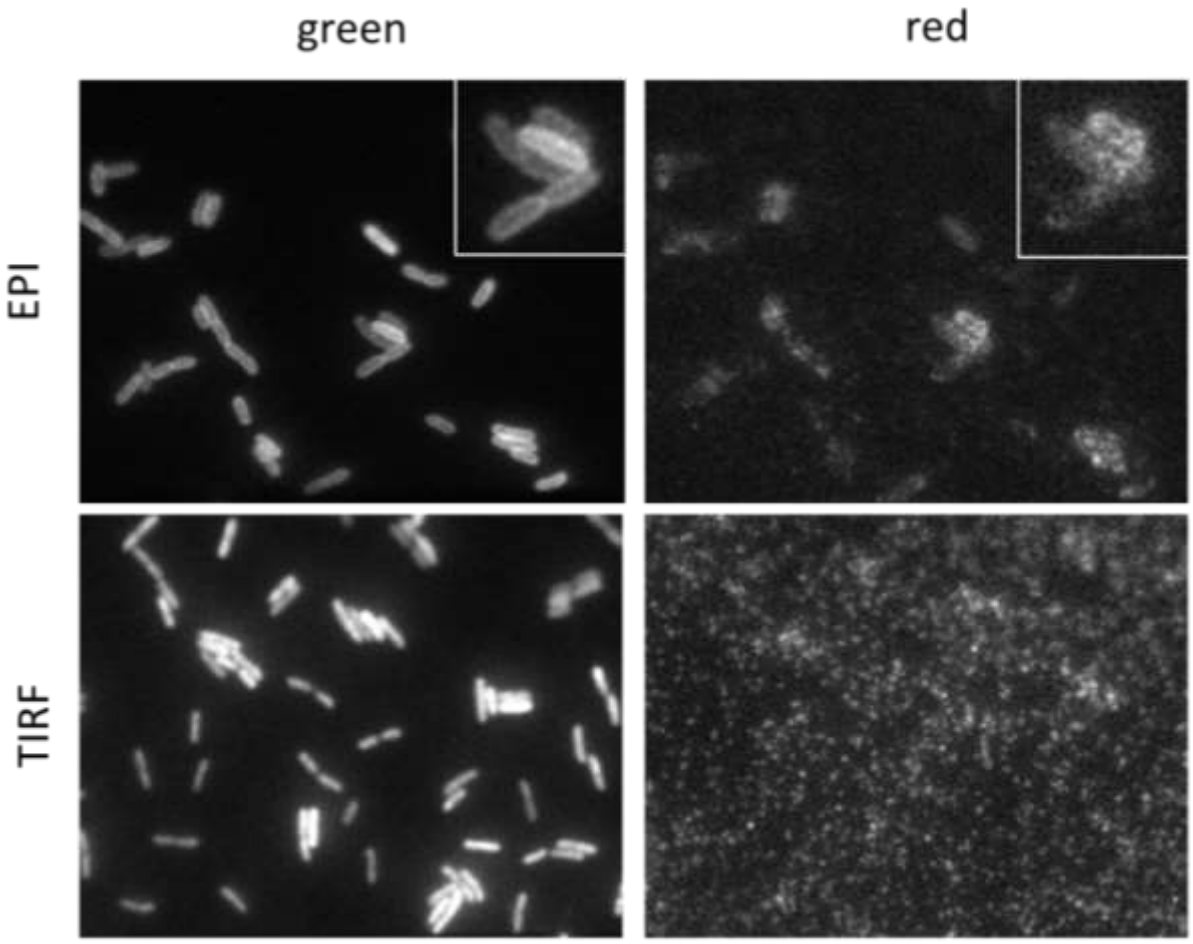


Figure 86: The difference between TIRF and epifluorescence for the red (ColN or ColA, 642 nm) and green (Para-CNI-sfGFP-6His, 510 nm) channels. In epifluorescence, the ColN or ColA molecules are attached to the cell membrane, while TIRF shows the glass cover slip indiscriminately covered in free colicin molecules. The green channel here is given for reference and shows continuity with Figure 73.

5.2.13 Mean square displacement - method assumptions and limitations

The single molecule tracking software (Rolfe *et al.*, 2011) detects all fluorescent molecules and tracks them whether they are moving or not. It is important to consider which molecules are pooled when calculating the MSD and values from it, since it will average the behaviours of the included molecules, and will thus be affected by the distribution of motions of those molecules, for example the proportion of moving and non-moving molecule included. Therefore, it is essential to consider the contribution of moving and non-moving sources of signal to the overall average. The total number of tracks and the number of tracks from moving molecules are indicated in Table 11, Table 12 and Figure 87 - Figure 94.

Here, we define a molecule as moving if its instantaneous D > sig(D) (in μm^2/sec). Sig(D) is one standard deviation of the diffusion rate "instantaneous D". The instantaneous D is estimated from the gradient of the first two points in the MSD curve. In Figure 87 - Figure 94 it is shown with a dashed line. The average value is more reliable the more tracks are measured and, ideally, a similar number of tracks for each treatment would make the data more comparable. This is not the case for Para-CNI-sfGFP-6His where 3 times more tracks were measured for 400 nM ColA and ColN and 4 µM ColA and ColN than for 40 nM ColA and ColN and many more tracks were measured for Para-CNI-sfGFP-6His without ColN. The track numbers clearly show how in some cases only a very small number of moving tracks actually contributes to the MSD. Whether one needs to consider all tracks or just the tracks from moving particles is a matter of definition and it is difficult to make a decision because there is still a lack of biological knowledge about this system. It is also unclear what the molecules that do not move are, i.e. whether they are immobile fluorescent proteins, artefacts from lysed cells or fluorescent non-protein artefacts on the slide. Comparing total track numbers is not meaningful as the track number is largely defined by how many cells are in a frame and how many data sets were taken. Also, cell sizes are different and expression is usually not homogeneous throughout a culture. Therefore, no conclusion can be made about whether addition of ColN or ColA changes the track number per cell.

Looking at the graphs where only moving particles are analysed the instantaneous D is much higher for all treatments. The general trend, however, remains the same. Overall, there is no significant difference between treatments and even at higher concentrations there is no significant difference between ColN and ColA. Overall, it is therefore more likely that the "slowing down" effect is a general effect on the membrane due to colicin addition, rather than one caused by specific protein-protein interaction. Here, it is crucial to consider the controls TolA-eGFP-6His and WALP23-eGFP-6His. TolA-eGFP-6His and WALP23-eGFP-6His were only observed with 40 nM of ColA and ColN due to time constraints and also because this concentration is already a lot higher than the minimal inhibitory concentration (MIC) of MC1000 cells (see chapter 4 for the MIC of sensitive cells). ColA and ColN will dissipate membrane potential and so may influence membrane fluidity and other proteins in the membrane. Because the membrane is such a crowded and tightly packed

environment, simply the presence of a colicin P-domain may lead to unspecific effects.

It is unknown if, upon binding to ColN, CNI slows down or stops entirely. Even if free diffusion is limited it is still likely to be moving around a little due to the membrane fluidity at 37 °C. It is unknown how many of the protein molecules are moving around or sitting still. It is also unknown what happens to ColN once it has entered the membrane, for example, whether it is neutralised by CNI and just stays there until it is removed in the general turnover of the inner and outer membrane. Since these mechanisms involve a lot of proteases it is conceivable that a neutralised ColN molecule could be degraded and removed during the general growth cycle. ColN is cytotoxic rather than cytostatic, so if the cell does not die, there should be no obstacle for it to stop growing, remodelling and dividing, as is demonstrated in the liquid culture killing assays in Chapter 4.

While it is easy to define mathematically what “stationary” means, it is difficult to do this biologically as the membrane is in constant movement and so in theory all proteins should be moving. Fundamentally, it seems to be a case of definition, whether one excludes “non-moving” sources of signal as false positives. Therefore, all results are given for all detected as well as all moving molecules. To exclude 90% of the signal is a rather large number and it is entirely possible and evident that not all molecules, even of the same kind, are in the same condition. However, it can be argued that, within this set up, I was trying to observe a change in diffusion rate and so must be observing something that is diffusing in the first place. When considering objects that are by definition not moving, it is not possible to observe a slowing down effect. And to observe the biggest possible effect of colicin binding I must focus on the moving molecules.

Finally, although I obtained quantitative numbers for the instantaneous D of Para-CNI-sfGFP-6His, TolA-eGFP-6His and WALP23-eGFP-6His, these proteins would probably have faster diffusion rate without the GFP fusion protein (ca. 27k Da). However, CNI and TolA can still carry out their biological function and so the measured diffusion rate should be a good approximation of the real movement.

5.2.14 Mean square displacement - the effect of colicin addition on membrane protein movement

Based on images like those shown in Figure 85, Dr Rolfe calculated the MSD for all treatments. The calculated values of mean square displacement and track numbers are summarised in Table 11, Table 12 and shown in Figure 87 - Figure 94.

When considering all detected molecules for Para-CNI-sfGFP-6His, both ColN and ColA addition slows down movement as the sample without colicin is the fastest with instantaneous D = 0.025 $\pm$ 0.002 µm ^{2}/sec (Table 12). ColN and ColA addition reduces diffusion rates to ca. 50% with 40 nM and 25% with 400 nM and 4 µM, so generally, the more colicin is added, the slower the protein. There is no significant difference between ColA and ColN at lower concentration of 40 nM and 400 nM. A significant difference is only apparent at the highest used concentration of 4 µM, where Para-CNI-sfGFP-6His treated with ColA is twice as fast as when treated with ColN. Indicating that CNI binds ColN and therefore slows down, while it does not bind ColA. However, when considering only the moving molecules there is no significant difference in instantaneous D upon colicin addition and there is no significant difference between adding ColN or ColA, even at 4 µM. This observation could lead to the conclusion that ColN addition causes a higher proportion of Para-CNI-sfGFP-6His molecules to become non-motile than ColA. Comparing the proportions of mobile molecules to all molecules between ColA (67/1247x100 = 5.37 % motile) and ColN (27/1046x100 = 2.58 % motile), this conclusion is only true for 4 µM. When no colicin is added the proportion of motile to all molecules for WALP23-eGFP-6His and TolA-eGFP-6His is ca. 10 % and 7 % for CNI-sfGFP-6His. When colicin is added this proportion remains at 10 % for WALP23-eGFP-6His and TolA-eGFP-6His, while it drops to ca. 3 % for CNI-sfGFP-6His. Whether this has any biological relevance is as yet unclear.

WALP23-eGFP-6His moves much faster than the proteins Para-CNI-sfGFP-6His and TolA-eGFP-6His, which is expected because it is not a protein but an artificial transmembrane helix (Figure 89 and Figure 90). Addition of ColA or ColN slows down WALP23-eGFP-6His to approximately 60 - 70 % percent of its original diffusion rate and so this is a considerable reduction. As expected there is no significant difference between ColA and ColN, as WALP23-eGFP-6His does not interact with either of the colicins in a specific way.

TolA-eGFP-6His is the positive control because it is an essential binding partner for both ColA and ColN during translocation through the periplasm and necessary in order to reach the inner membrane. TolA deletion impacts cells growth negatively but allows complete resistance to ColA and ColN. TolA is a large, elongated protein (ca. 43 kDa) with 3 domains. Domain I is anchored in the inner membrane while domains 2 and 3 are protruding into the periplasm and interacting with other proteins. In comparison to Para-CNI-sfGFP-6His and WALP23-eGFP-6His it is therefore a relatively slow moving protein which is reflected in the MSDs (Figure 91 and Figure 92). TolA-eGFP-6His follows a similar pattern as CNI-sfGFP-6His and WAL23-eGFP-6His in the way it reacts to ColN and ColA. When either colicin at 40 nM is added, the protein slows down. When considering all detected molecules the MSD of TolA-eGFP-6His is reduced by ColA from 0.0153± 0.0008 μm^2/sec to 0.0092 ± 0.0007 μm^2/sec, and to 0.0069 ± 0.0004 μm^2/sec by ColN, which corresponds to a reduction by ca. 40% and ca. 55 %, respectively. When considering only the moving molecules, the MSD of TolA-eGFP-6His is reduced by ColA from 0.033 ± 0.003 μm^2/sec to 0.023 ± 0.002 μm^2/sec, and to 0.016 ± 0.002 μm^2/sec by ColN, which corresponds to a reduction by ca. 30% and ca. 50 %, respectively.

It is certainly noteworthy that the reduction is not the same with both colicins and in fact bigger with ColN than with ColA, although ColA is the bigger protein. This suggests further that it may not be the physical addition of bulk which slows down TolA but the effect the colicins have on the membrane overall, and hence an unspecific effect. This is supported further by the observation that WALP23-eGFP-6His is also slowed down by colicin addition, although there is no known or anticipated interaction with either of the colicins.

In conclusion, using single molecule tracking for diffusion rate quantification had several advantages over TIR-CP. Single molecule tracking is a direct rather than an indirect measure of diffusion rate, which allowed the comparison of different protein fusions and demonstrated the variety of diffusion rates for the same protein in the same condition, leading to the conclusion that the diffusion rate measured with TIR-CP is actually an average of different diffusion rates. A significant drawback to this method is that nearly 90 % of the molecules which emit fluorescence do not move, which is a considerable issue when trying to measure the replenishment of an area, as was attempted with TIR-CP.

5.3 Conclusions

- Several methods were developed to observe GFP labelled CNI in the membrane of expressing cells by epifluorescent and TIRF microscopy.
- The mobility of CNI and WALP were compared by bleaching methods in the presence and absence of added ColN and ColA
- Fluorescently labelled ColN and ColA were prepared but adhere non-specifically to the slides preventing co-localisation studies.
- Single particle tracking was used to compare the diffusion rate (instantaneous D) of CNI, WALP and TolA in the membrane.
- The magnitude of instantaneous D was greater for WALP than CNI and TolA, and greater for CNI than TolA.
- The diffusion rate of all three proteins decreases on addition of ColN or ColA

Table 11: Instantaneous D (diffusion rate) values of Para-CNI-sfGFP-6His, WALP23-eGFP-6His and TolA-eGFP-6His (in μm^2/sec), when only moving molecules are considered, where D >1σ (D).

	no ColN	40 nM		400 nM		4 µM	
		ColN	ColA	ColN	ColA	ColN	ColA
Para-CNI-sfGFP-6His	0.043 ± 0.004 n = 431	0.060 ± 0.020 n = 8	0.030 ± 0.020 n = 10	0.031 ± 0.006 n = 51	0.034 ± 0.006 n = 138	0.040 ± 0.010 n = 27	0.045 ± 0.007 n = 67
WALP23-eGFP-6His	0.095 ± 0.006 n = 379	0.061 ± 0.006 n = 256	0.071 ± 0.003 n = 628				
TolA-eGFP-6His	0.033 ± 0.003 n = 1091	0.016 ± 0.002 n = 659	0.023 ± 0.002 n = 312				

Table 12: Instantaneous D (diffusion rate) values of Para-CNI-sfGFP-6His, WALP23-eGFP-6His and TolA-eGFP-6His (in μm^2/sec), when all molecules are considered.

	no ColN	40 nM		400 nM		4 µM	
		ColN	ColA	ColN	ColA	ColN	ColA
Para-CNI-sfGFP-6His	0.025 ± 0.002 n = 5961	0.013 ± 0.004 n = 257	0.013 ± 0.005 n = 301	0.006 ± 0.001 n = 1164	0.008 ± 0.002 n = 2197	0.007 ± 0.002 n = 1046	0.013 ± 0.002 n = 1247
WALP23-eGFP-6His	0.064 ± 0.003 n = 3721	0.033 ± 0.003 n = 2380	0.050 ± 0.003 n = 5313				
TolA-eGFP-6His	0.0153 ± 0.0008 n = 10103	0.0069 ± 0.0004 n = 5487	0.0092 ± 0.0007 n = 3718				

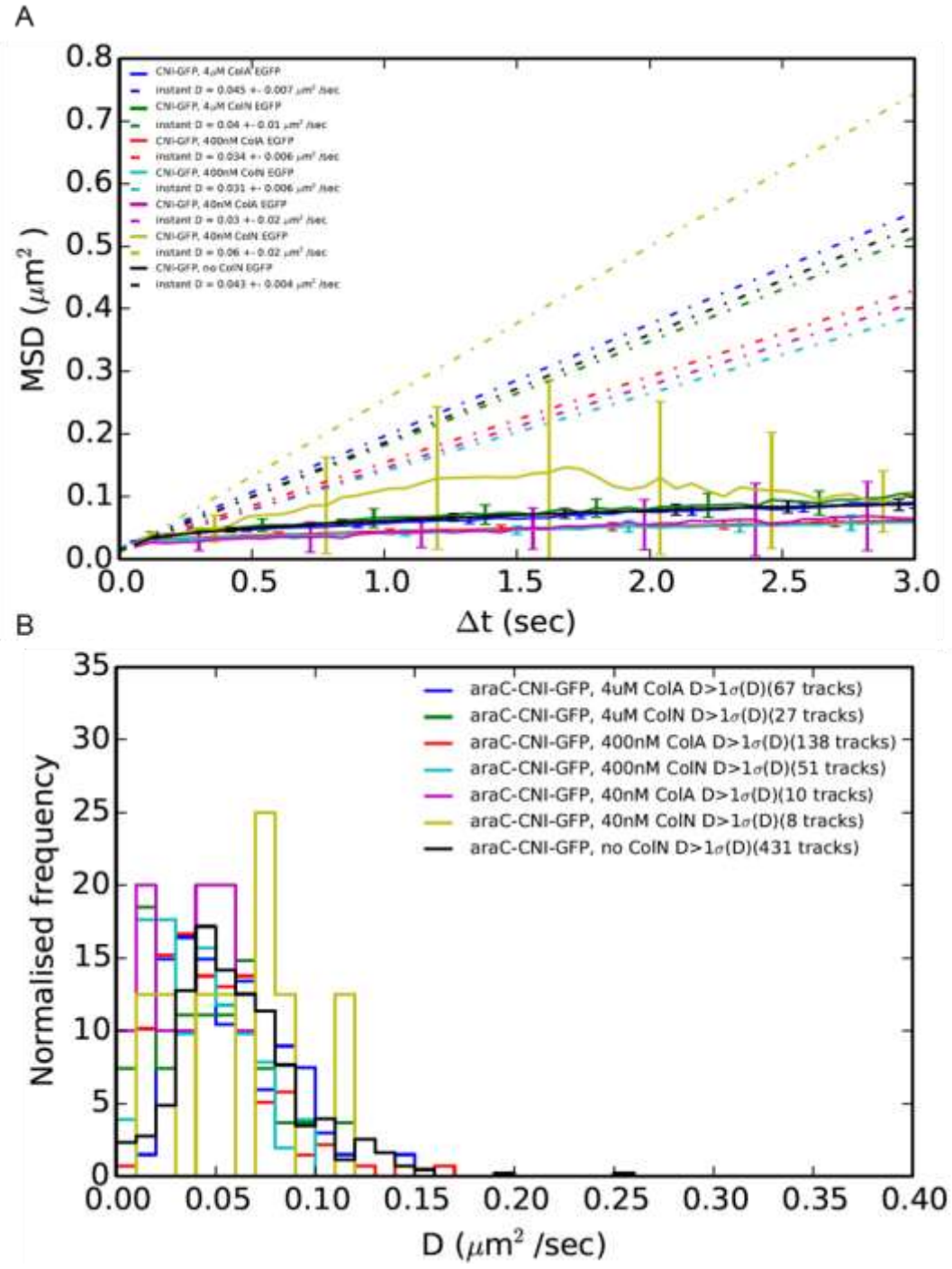


Figure 87: The effect of ColN and ColA on Para-CNI-sfGFP-6His MSD within the inner membrane. ColN and ColA were added at 4μM, 400 nM, and 40 nM. Colicin addition decreases diffusion rates but there is no significant difference between ColN and ColA. (A) MSD curves of moving particles, where displacement is higher than 1 standard deviation (D >1σ (D)). Instantaneous D is plotted with dashed line. (B) Normalised frequency of particles for each treatment based on their diffusion rate, when displacement is higher than 1 standard deviation (D >1σ (D)) and number of tracks for each treatment.

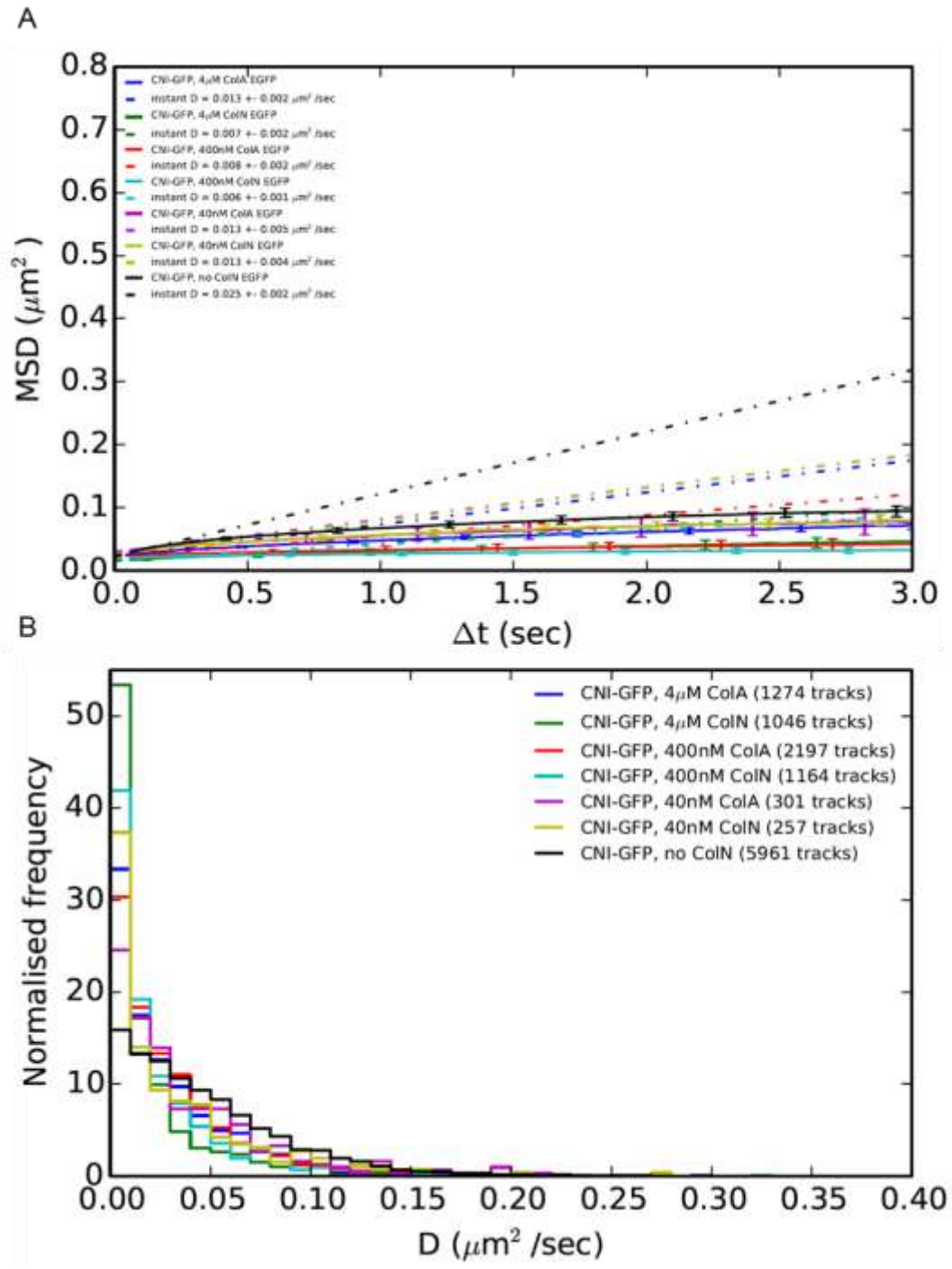


Figure 88: The effect of ColN and ColA on Para-CNI-sfGFP-6His MSD within the inner membrane. ColN and ColA were added at 4μM, 400 nM, and 40 nM. Colicin addition decreases diffusion rates but there is no significant difference between ColN and ColA. (A) MSD curves of all determined particle tracks. Instantaneous D is plotted with dashed line. (B) Normalised frequency of particles for each treatment based on their diffusion rate, and number of tracks for each treatment.

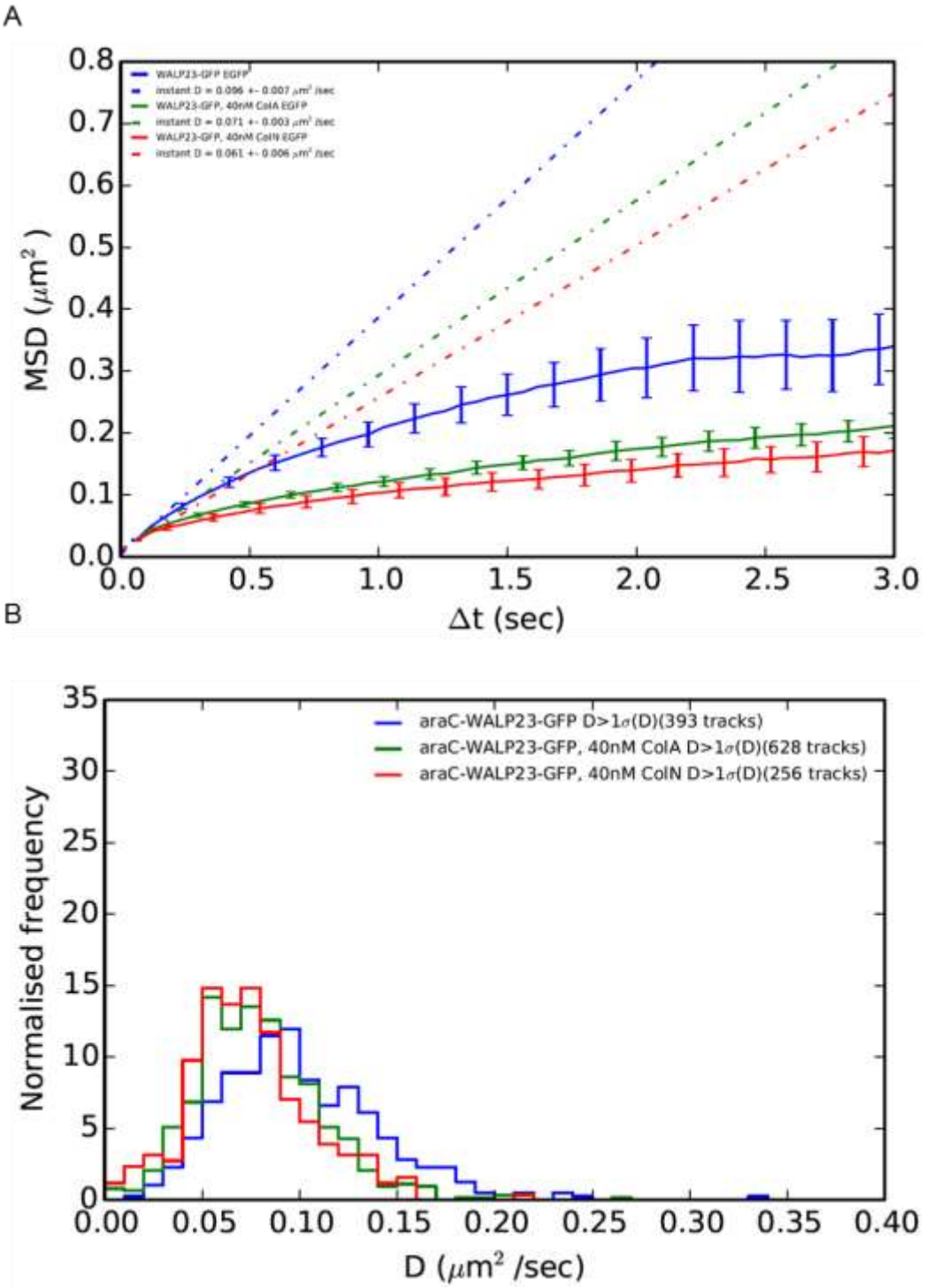


Figure 89: The effect of 40 nM ColN or ColA on WALP23-eGFP-6His MSD within the inner membrane. Colicin addition decreases diffusion rate but there is no significant difference between ColN and ColA. (A) MSD curves of moving particles, where displacement is higher than 1 standard deviation (D >1σ (D)). Instantaneous D is plotted with dashed line. (B) Normalised frequency of tracks for each treatment based on their diffusion rate, when displacement is higher than 1 standard deviation (D >1σ (D)) and number of tracks for each treatment.

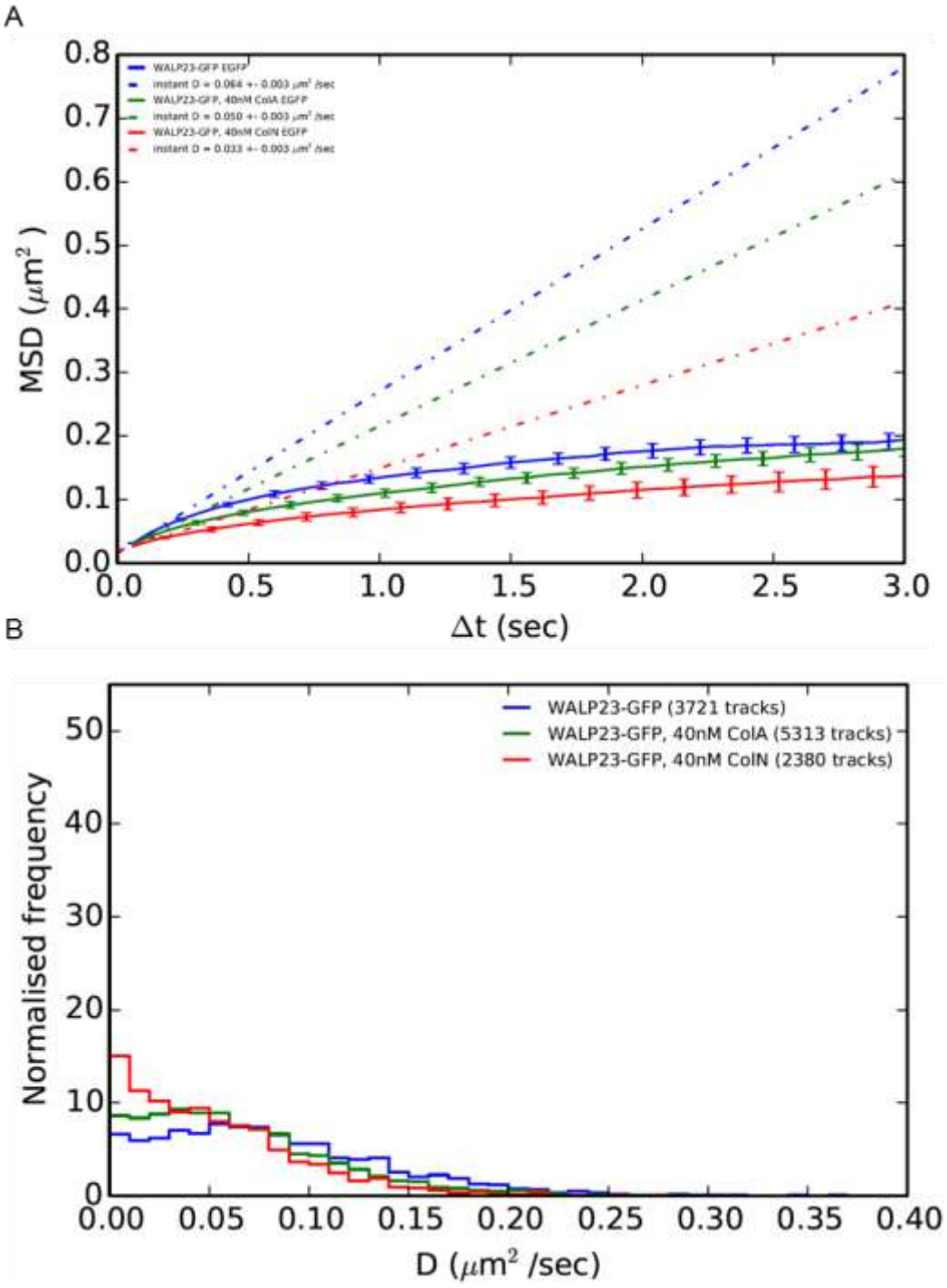


Figure 90: The effect of 40 nM ColN or ColA on WALP23-eGFP-6His MSD within the inner membrane. Colicin addition decreases diffusion rate but there is no significant difference between ColN and ColA. (A) MSD curves of all determined particle tracks. Instantaneous D is plotted with dashed line. (B) Normalised frequency of particles for each treatment based on their diffusion rate, and number of tracks for each treatment.

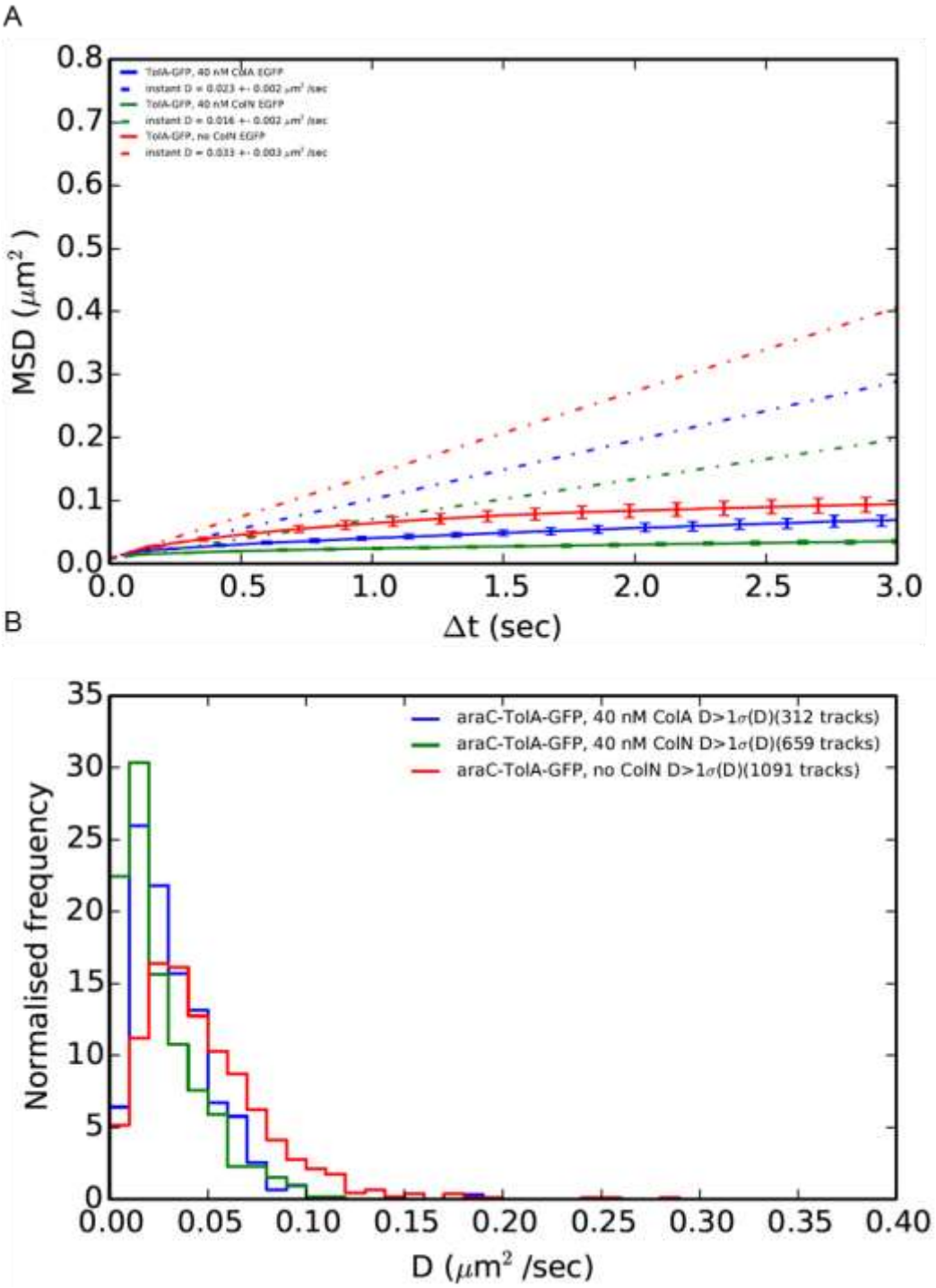


Figure 91: The effect of 40 nM ColN or ColA on TolA-eGFP-6His MSD within the inner membrane. Colicin addition decreased the diffusion rate but there is no significant difference between ColN and ColA. (A) MSD curves of moving particles, where displacement is higher than 1 standard deviation (D >1σ (D)). Instantaneous D is plotted with dashed line. (B) Normalised frequency of tracks for each treatment based on their diffusion rate, when displacement is higher than 1 standard deviation (D >1σ (D)) and number of tracks for each treatment.

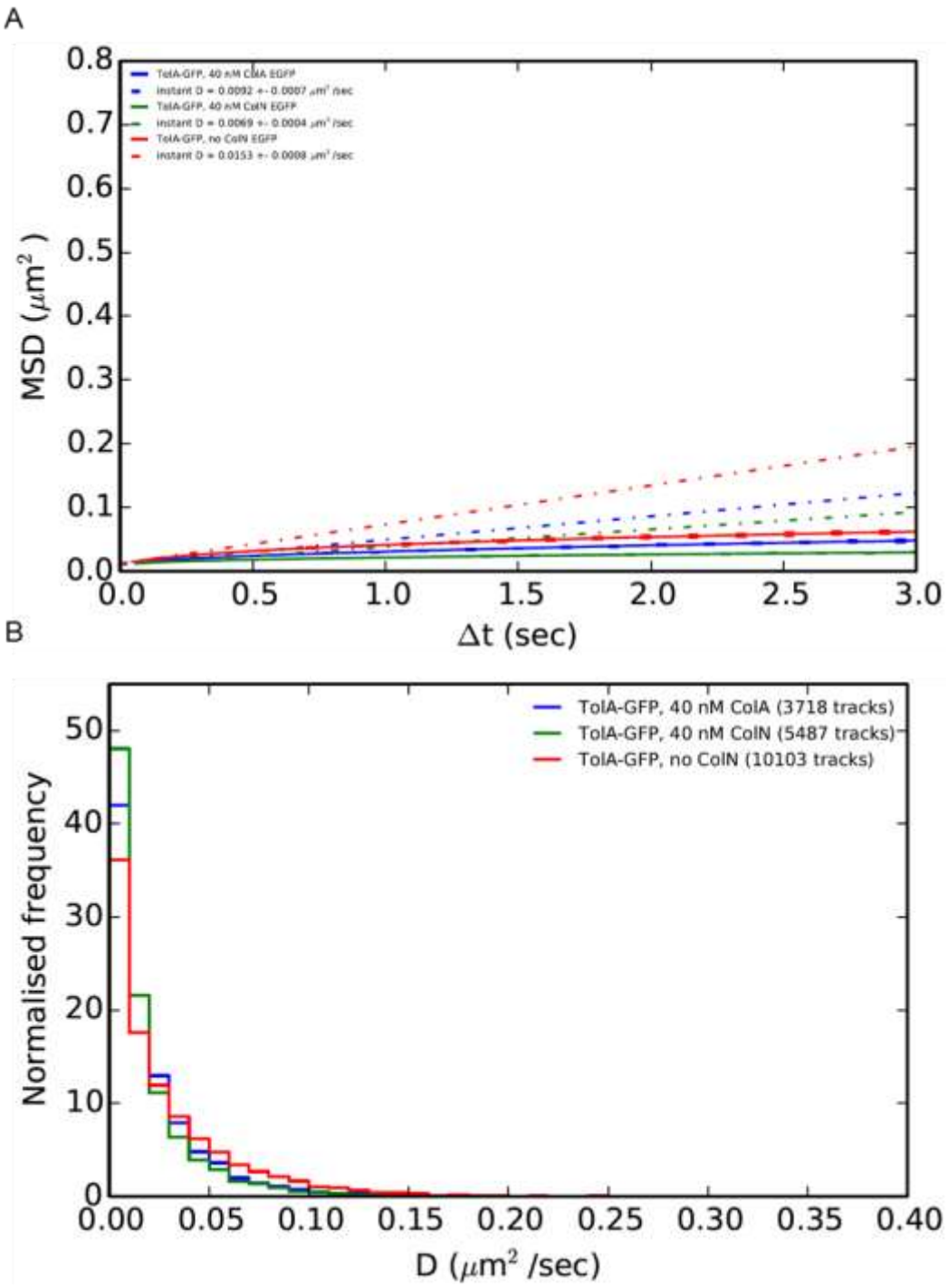


Figure 92: The effect of 40 nM ColN or ColA on TolA-eGFP-6His MSD within the inner membrane. Colicin addition decreased the diffusion rate but there is no significant difference between ColN and ColA. (A) MSD curves of all determined particle tracks. Instantaneous D is plotted with dashed line. (B) Normalised frequency of particles for each treatment based on their diffusion rate and number of tracks for each treatment.

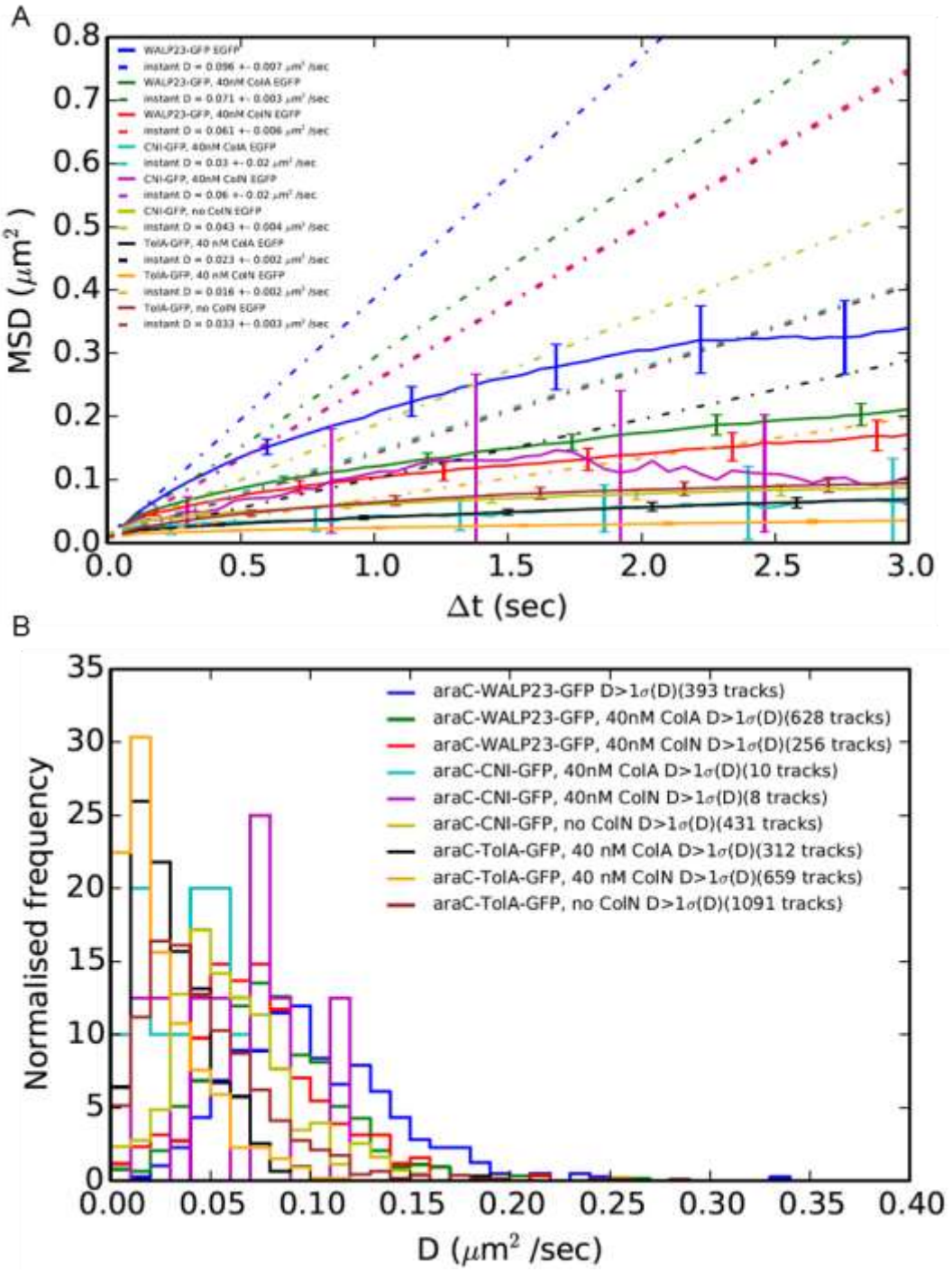


Figure 93: A comparison of MSD between Para-CNI-sfGFP-6His, WALP23-eGFP-6His and TolA-eGFP-6His at 40 nM ColN and ColA. For all treatments, the diffusion rate of Para-CNI-sfGFP-6His lies in between WALP23-eGFP-6His and TolA-eGFP-6His, as expected. (A) MSD curves of moving particles, where displacement is higher than 1 standard deviation (D >1σ (D)). Instantaneous D is plotted with dashed line. (B) Normalised frequency of tracks for each treatment based on their diffusion rate, when displacement is higher than 1 standard deviation (D >1σ (D)), and number of tracks for each treatment.

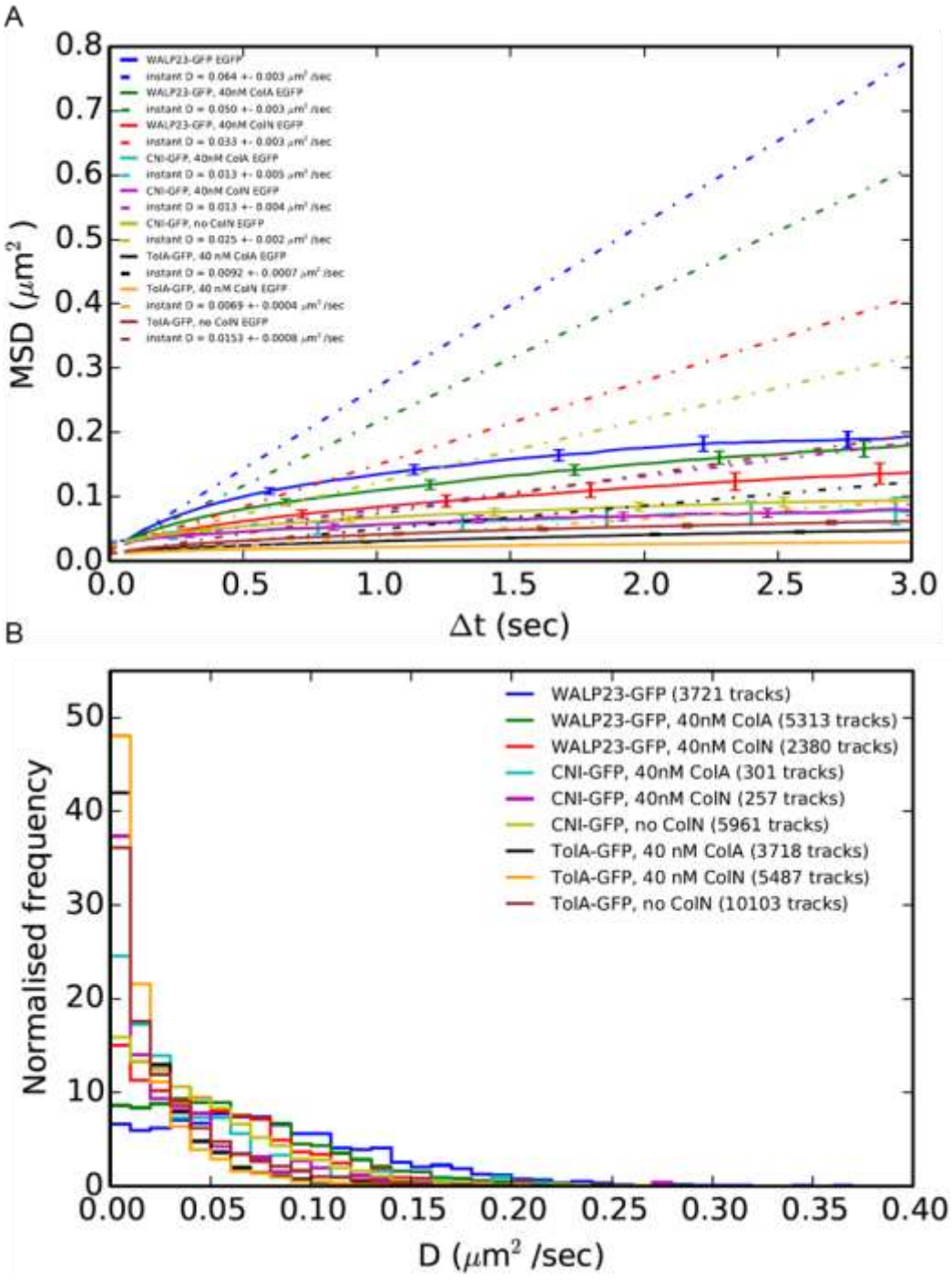


Figure 94: A comparison of MSD between Para-CNI-sfGFP-6His, WALP23-eGFP-6His and TolA-eGFP-6His at 40 nM ColN and ColA. For all treatments, the diffusion rate of Para-CNI-sfGFP-6His lies in between WALP23-eGFP-6His and TolA-eGFP-6His, as expected. (A) MSD curves of all determined particle tracks. Instantaneous D is plotted with dashed line. (B) Normalised frequency of particles for each treatment based on their diffusion rate, and number of tracks for each treatment.

Chapter 6. Discussion, Conclusions and Future work

6.1 Discussion

Colicins are an interesting and important field of study because they have diverse mechanisms of function, interacting with nucleic acids, peptidoglycan precursors, lipids and proteins of the inner and outer membrane (Cascales *et al.*, 2007; Kim *et al.*, 2014). Considering this breadth of function it is not surprising how common they are in natural microbial populations. Also, it is not unexpected that humans have started to consider their agricultural and medicinal applications as well as studying them academically as model systems for protein-nucleic acid interaction, protein-protein interaction and membrane translocation. However, the problem with all antimicrobial agents is the eventually emerging immunity and so more bacteria produce colicin immunity proteins than colicins (Riley, 1993; Gordon *et al.*, 1998), which is why colicin immunity proteins should also be studied.

When starting this work on the Colicin N immunity protein, only its sequence and general function were known, published knowledge on homologues was limited and the fundamental question of how colicinogenic bacteria protect themselves against their highly potent pore-forming toxin was and still remains unanswered. While for nuclease colicins, the mode of immunity protein binding and release as well as the force of affinity have been established, the interaction mechanism of pore-forming colicins and their immunity protein is unknown although most literature points towards a direct interaction. To test this hypothesis, *in vitro* interaction studies are crucial and so the first aim of this project was to express and purify CNI.

6.1.1 Fusion proteins improve CNI overexpression

The natural expression level of immunity proteins to pore-forming colicins is assumed to be constitutive and very low (Pugsley, 1988; Cascales *et al.*, 2007), and thus all previous work on immunity proteins to pore-forming colicins has used overexpression systems. Overexpressing membrane proteins requires a fine balance between a system that produces enough protein for purification but minimises the toxicity and metabolic burden. In nature, colicins as well as their lysis and immunity proteins are encoded on a plasmid, so expression from a well-characterised vector is an obvious

approach. In this work, a great variety of constructs with different promoters, plasmid copy numbers, fusion tags and fusion proteins was investigated. Ultimately, a CNI-HALO7-6His fusion expressed in C41 *E. coli* cells and purified using a DM containing Tris buffer was chosen as the optimal combination to produce the most stable monomeric protein. However, overexpression initially started with a polyhistidine-tagged version of CNI which should have given a product that did not require protease cleavage. Unfortunately, N- or C-terminally polyhistidine-tagged CNI could not be expressed in detectable quantities using either arabinose induction from a pBAD322 system or IPTG induction from a pOPIN vector in any of the tested conditions. Also, while CNI fusions with larger proteins, such as GFP, HALO7 or MBP and the cleaved tags were easily detected using Coomassie-stained gels, western blots and in the case of GFP-tags also in-gel fluorescence, the cleaved CNI protein was hardly detected. CNI cleaved from its fusion protein was only visible as very faint bands on Coomassie-stained SDS-PAGE gels and a western blot. LC-MS analysis of protease cleavage products confirmed that cleaved CNI was present and not degraded by the protease, but can give no indication of amounts. Cleaved CNI could be purified and concentrated by ion exchange chromatography but high salt concentrations can be a problem for membrane proteins because they require higher detergent concentrations for micelle formation. It is unclear why larger fusion protein tags were more beneficial for expression than smaller tags, but it may be related to a naturally high turnover rate of CNI, which is somehow inhibited by the large tag. Cell survival assays showed that cultures become more immune when the level of expression is increased by changing the promoter and plasmid copy number from arabinose induced expression using a Para promoter to a IPTG induced T7 promoter. The pCHAP4 and pBAD322 plasmids are low copy number while all the other plasmids used are high copy number plasmids. It may also be that the protein is not very stable and the tag provides a stabilising scaffold.

The difficulties of producing CNI by itself or with a small tag indicate that the larger fusion proteins are crucial in stabilising the protein outside of the membrane. CNI, therefore, seems to be amongst the many proteins which benefit from expression as a fusion. Fusions with larger proteins, especially GFP, are a popular approach to improve stability, solubility, detection, purification optimisation (Hammon *et al.*, 2009; Bird *et al.*, 2015) and even crystallography (Carpenter *et al.*, 2008; Suzuki *et al.*, 2010).

6.1.2 Alternative approaches to mimicking the inner membrane environment

These difficulties leave three routes for further characterisation. Some techniques, particularly biophysical ones, like cryo-electron microscopy (Zeev-Ben-Mordehai *et al.*, 2014) or small angle x-ray scattering (Gilbert *et al.*, 1999; Firer-Sherwood *et al.*, 2011; Breyton *et al.*, 2013) which determine the protein's shape, structure or rigidity, require as much pure protein as possible so further ways of purifying and stabilising CNI in a monomeric form are required. Past attempts all included overexpression of the immunity protein from a constructed system and purification with reducing agent, such as the DTT used in this work, and detergents like Triton X-100 (Mankovich *et al.*, 1986; Geli *et al.*, 1989b; Shirabe *et al.*, 1993) or solvents like butanol (Taylor *et al.*, 2000). However, none demonstrates specific *in vitro* interaction between the colicin and its immunity protein, indicating that the difficulty to achieve specific interaction is common to several immunity proteins.

After Weaver *et al.* (1981) determined the activity of the Colicin Ia and Ib immunity proteins (both E1-type) *in vivo* using vesicles produced from cells immune to Colicin Ia and Ib, Mankovich *et al.* (1986) determined the sequence of Colicin Ia and Ib immunity proteins and constructed an overexpression system because the natural expression levels were too low to attempt purification. Mankovich *et al.* (1986) extracted the Colicin Ia and Ib immunity proteins using a 10 mM Tris (pH = 8) buffer containing 1 % Triton – X 100, 20 mM $MgCl_2$, 100 mM 2-mercaptoethanol and purified them using an isoelectric focusing column. It is noteworthy that a reducing agent was used, just like DTT was used in this work, although it is not explained why it was added. The level of expression and purification yields were assessed using SDS-PAGE and autoradiography Following overexpression 31 µg of pure protein from 2.1 g (wet weight) of cells was produced. Mankovich *et al.* (1986) do not report any structural studies, *in vitro* interaction studies with the cognate colicins or attempts to reconstitute the immunity proteins. However, based on amino acid sequence analysis alone, they propose that immunity proteins interact directly with the hydrophobic region of the colicin molecules.

Geli *et al.* (1989b) tagged the Colicin A immunity protein (CAI) with an epitope from ColA for which a monoclonal antibody was available. They used a 75 mM Tris – acetate buffer (pH = 9.3) containing 1 % Triton X-100, 1.5 mM EDTA, 0.5 mM N-ethyl maleimide and 0.2 mM phenylmethyl sulfonyl fluride for purification, however, during the following chromatofocusing and immunoadsorbant chromatography the Triton X-

100 concentration was reduced to 0.1 %. This reduction in detergent concentration is the same approach that has been adopted in this work with DM and Cymal 6. The purified CAI protein has then been reconstituted into lipid vesicles by Geli *et al.* (1989b) but specific interactions with ColA and, therefore, activity of CAI could not be shown. Geli *et al.* (1989b) discuss various reasons why a specific interaction could not be demonstrated, including surface pressure of the vesicles, orientation of the reconstituted CAI protein and the lack of membrane potential. They also point out that there was unspecific binding of ColA to the lipid vesicles, which can be expected, because ColA would form pores in a membrane whether CAI is present or not. Perhaps this method could be improved by quantifying the binding affinity using more modern techniques such as SPR and determining a difference in binding affinity between ColA and vesicles and ColA and CAI in vesicles. Also, Geli *et al.* (1989b) uses very dramatic pH changes during their purification procedure, from alkaline to acidic and back to alkaline, which could cause the protein to denature or aggregate. This work (Chapter 3) shows how acetic acid can unfold properly folded and aggregated CNI prior to SDS-PAGE and western blotting to improve detection by the antibodies and it is possible that hydrochloric acid has the same effect on CAI. Geli *et al.* (1989b) does not demonstrate that the structure of CAI is intact after purification.

Shirabe *et al.* (1993) overexpressed Colicin E1 immunity protein using a specially constructed overexpression system and a *tac* promoter. The protein was solubilised using 0.5 % Brij® 58, a non-ionic detergent with varying chain length, and further purified with a hydroxyapatite gel and reverse phase chromatography. 360 µg of purified protein were obtained from 1 l culture. Other detergents were also tested for solubilisation (data not published) but only Triton X-100 and Brij® 58 were effective and Brij® 58 was more effective than Triton X-100. When the purified Colicin E1 immunity protein was added to solute-loaded liposomes, small but significant lytic activity was observed which demonstrated the hydrophobic nature of the protein. An interaction with Colicin E1 was not tested.

Taylor *et al.* (2000) overexpressed the immunity protein to Colicin E1 using a T7 promoter in BL21(DE3) cells, just like has been done in this work for CNI. Taylor *et al.* (2000) report that purification of the Colicin E1 immunity protein using several detergents was unsuccessful and that, therefore, purification with solvents was attempted instead. They report a lengthy and harsh purification procedure which involves gel filtration, ion-exchange chromatography, desiccation and a final refolding

step in chloroform/methanol/ H_2O (4:4:1). Using near-UV CD and NMR they confirmed that the protein was folded into a bundle of α-helices, as expected, and suggest that the protein's tertiary structure was dependent on helix-helix interactions. In this work, an increased degradation was observed in the chimeric proteins, where individual CNI helices have been replaced with homologous helices from CAI, which suggests that the chimeric CNI proteins have lost their WT structure and are therefore more prone to degradation. Similarly, to the Colicin E1 immunity protein it is therefore possible that the structure of CNI is determined by helix-helix interactions. Despite obtaining a purified and folded Colicin E1 immunity protein, an interaction with Colicin E1 is not reported by Taylor *et al.* (2000).

Potential different ways to stabilise CNI in a membrane-like environment could include styrene maleic acid lipid particle systems (Rodger *et al.*, 2002; Knowles *et al.*, 2009; Dorr *et al.*, 2014; Overduin, 2014; Paulin *et al.*, 2014; Postis *et al.*, 2015) and lipid-detergent mixtures, like those already used for lipid-cubic phase in membrane protein crystallography (Siegel *et al.*, 2006; Li and Caffrey, 2011; Caffrey *et al.*, 2012; Warren *et al.*, 2013; Li and Caffrey, 2014; Weierstall *et al.*, 2014).

6.1.3 Stabilising CNI through ColN P-domain

Another approach to improve stability could be to replace the fusion partner with a biological binding partner, such as the P-domain or full length ColN. This could be carried out as a continuation of the pull-down assays in this work. The fusion protein tag would remain bound to the beads, while CNI and P-domain could be cleaved using a PreScission protease, which in turn can be removed using GST-binding resin.

There is ample evidence suggesting that immunity proteins to pore-formers directly interact with the hydrophobic loop of the P-domain (Geli and Lazdunski, 1992; Zhang and Cramer, 1993). In this work, it was attempted to show specific interaction between CNI and ColN or ColN P-domain using SPR and pull-down assays, however, none of the conditions tested could imitate the inner membrane adequately. CNI is highly hydrophobic, which is explained by its amino acid sequence, its predicted secondary structure and the bioinformatics prediction that a large majority of the protein is embedded in the inner membrane. Size exclusion chromatography, SPR and, to a limited extent, pull-down assays were greatly hindered by CNI binding to the Superose 12 column, the SPR chip and ColA unspecifically and aggregating in

the HEPES buffer at room temperature during SPR. This suggests that the detergent was not able to mask the hydrophobic residues sufficiently, either because it is an unsuitable detergent or because 0.1% is not an adequate concentration. It was also observed that the CNI aggregates faster at room temperature than at 4 °C. A suitable detergent must therefore mask the hydrophobic areas and stop aggregation but still allow CNI and ColN to interact using those sites. This would be achieved when CNI has a greater affinity for ColN than for the detergent. *In vivo*, the interaction between the immunity proteins and colicins is greater than their interaction with lipids so a similar situation has to be achieved *in vitro*.

SPR can be carried out with lipid vesicles (Besenicar *et al.*, 2006), which could contain reconstituted immunity protein. However, reconstitution requires a large amount of protein since not all the proteins will be in the correct orientation. It could be possible to reconstitute the protein while it is tagged and cleave the tag afterwards. For CNI both N- and C- terminal are on the same side (the cytoplasmic side) and so if the protein inserts in the correct orientation, the tag will point inside the vesicle. If it is in the wrong orientation, with termini pointing outside of the vesicle, it will be cleavable using a protease and hence unspecific interactions between the tag and the binding partner, e.g. ColN P-domain, can be avoided. It has previously been hypothesised that lipids play a role in pore-formation (Sobko *et al.*, 2004a; Sobko *et al.*, 2004b; Sobko *et al.*, 2005; Sobko *et al.*, 2006a; Sobko *et al.*, 2006b; Sobko *et al.*, 2006c; Sobko *et al.*, 2010) because P-domain alone cannot form a pore as large as has been measured by electrophysiology. Therefore, it is possible that lipids are also important in the interaction of the colicin and its immunity protein.

6.1.4 Characterising tagged CNI using mutagenesis

In the absence of sufficient untagged CNI, a third approach of characterising CNI is by considering how it can be characterised while it is tagged, and overexpressed or regulated by its natural promoter. The caveat of using a tagged protein is always whether the tag somehow influences the protein's natural behaviour and possibly interferes with its function, especially if the tag is bigger than the protein. In this work, the activity of each new construct has been tested thoroughly and proven that tagged CNI maintains its function. There is no evidence which shows that the tags inhibit function or that the increased amount of protein in the membrane is significantly detrimental to the cells. Using CNI regulated by its natural promoter, a comparison

between tagged and untagged CNI was made using a spot test, which proved that tagged and untagged CNI provides the same level of protection against ColN. It is assumed that this is also the case for overexpressed CNI. In this work, pull-down assays and SPR have used the tags as anchors to solid surfaces but, as previously discussed, establishing specific interactions was not successful.

A tagged version of CNI has also been used to create mutants and assess their activity *in vivo.* Zhang and Cramer (1993) proposed pore-forming colicins and their immunity proteins interact through hydrophobic helix-helix interactions. Since then, this hypothesis has been supported by Sobko *et al.* (2009), Smajs *et al.* (2006), Smajs *et al.* (2008), Pilsl *et al.* (1998), Zhang *et al.* (2010) for homologous colicins. Based on this hypothesis, chimeric proteins were created where each predicted helix of CNI was swapped with the corresponding predicted helix in CAI, with the aim of changing the specific immunity from ColN to ColA. Although, a change in specify was not achieved, three other conclusions could be drawn. Firstly, CNI Helix 1 is not necessary for the specific interaction with ColN because the Helix 1 mutant still protects against ColN. Whether Helix 1 is entirely redundant for protein function or if CAI Helix 1 replaced its function could be determined in future through a truncation. Geli *et al.* (1988) replaced the first 12 amino acids of CAI with a part of ColA or β-galactosidase, which is similar to a truncation, and found that CAI was still functional, indicating that the N-terminal was not required for function. However, the first 12 amino acids are not part of the first helix of CAI as the helix starts with residue 17 and so the first helix of CAI could still replace the function of the first helix of CNI. Secondly, when Helix 2 and 3 of CNI were mutated in this thesis, no protein was produced. It was confirmed through sequencing that the promoter sequence and as well as the gene sequence do not contain unintended mutations, so whether the protein was transcribed but not translated is unclear and can only be resolved by analysing the cell RNA content.

Finally, following the hazy appearance of Coomassie stained bands from purified and cleaved CNI on SDS-PAGE gels, it may be that CNI is a very flexible protein, with a well-defined α-helical secondary structure connected through flexible loops or turns but no fixed tertiary structure. It was therefore hypothesised that helix swapping would have a minimally detrimental effect on protein structure. However, a western blot showed increased degradation in the mutants when compared with the WT CNI

and CAI proteins, indicating that Helix 1 and 4 are important in overall integrity of the protein and that therefore the protein must have a defined tertiary structure.

CNI contains only one cysteine residue, located in the predicted 3rd helix, and it was hypothesised to be involved in protein function or possibly lead to a dimerisation of the protein, like is observed with CAI (Zhang *et al.*, 2010). However, the CNI^{C113S} mutant fusion remained active, albeit was less effective than the WT CNI. There is no evidence to suggest that it is not a dimer *in vivo,* which could be tested by SDS-PAGE with a loading buffer that does not contain a reducing agent, but this potential dimerisation probably plays a minimal role in protein activity. The mutation also does not affect purification yields in a significant way. It is likely that the cysteine has a structural purpose, possibly in stabilising Helix 3, and, therefore, mutating it to serine only has a limited effect.

Possibly, the purified and cleaved CNI protein is unstable because it has been delipidated when extracted from the membrane using detergents. Lipid-protein interactions are crucial for membrane proteins because the lipids contribute significantly to the membrane characteristics, such as fluidity, lateral pressure, bilayer thickness and surface charge (Stangl and Schneider, 2015). Lipids can interact with membrane proteins in many ways, affecting their function and polymerisation state. Lipids and membrane proteins can interact through the polar lipid head groups and the polar part of the membrane loops or turns. For example, negatively charged lipid head groups bind to stretches of basic amino acids. In addition, the hydrophobic transmembrane domains of the protein, in case of CNI the helices, can bind to the hydrophobic fatty acid chains of the lipid.

There are three types of interactions between lipids and membrane proteins (Stangl and Schneider, 2015). Firstly, the bulk lipids that surround the protein loosely and make up the general environment. Secondly, the shell of annular or boundary lipids that associate loosely with the protein and the composition of which is determined by the local architecture of the protein (Yeagle, 2014). Thirdly, the non-annular lipids that bind closely and specifically to the membrane protein, usually in clefts, cavities and hydrophobic binding pockets and could be decisive for correct protein folding. Non-annular lipids often remain bound to membrane proteins even after purification with detergents and can sometimes even be co-crystallised with the protein.

The LC-MS and SDS-PAGE data in this work suggests that at least one such non-annular lipid (approximately 700 Da in size) is probably attached to CNI and co-purifies with approximately 50 % of the CNI molecules when non-ionic detergents are used. Both annular and non-annular lipids can play a crucial role in protein polymerisation, acting as a glue, or the prevention of polymerisation, acting as a shield or repellent (Stangl and Schneider, 2015). It is, therefore, possible that CNI appears smooth rather than clustered *in vivo* and shows no signs of dimerisation despite the cysteine residue because it is surrounded by lipids. Proteins may preferentially interact with specific lipids, possibly based on the length of the fatty acid chain. Therefore, a suitable detergent should have tails which match the lipids surrounding that membrane protein (Stangl *et al.*, 2012).

As well as playing a structural role within the molecule and aiding polymerisation, lipids can also aid in the interaction between different proteins and may be an alternative to disulphide bonds in a nonpolar environment. It has repeatedly been suggested that colicins and their immunity proteins interact through helix-helix interactions. Helix-helix interactions can be caused by van der Waals interactions, hydrogen bonds and intercalating aromatic residue interactions and so it is essential that interacting helices have complementary surfaces (Lemmon and Engelman, 1994; Cymer *et al.*, 2012).

In this work, a lot of non-specific binding was observed, although *in vivo* data shows that CNI interacts specifically with ColN. Possibly, the absence of lipids from the SPR and pull-down assays which used largely delipidated protein could be an explanation. CNI is highly hydrophobic and may be highly prone to bind unspecifically and ultimately aggregate but could be shielded from these non-specific interactions by lipids. DM used at 0.1 % may not be suitable to imitate this shielding effect and this is why there is so much unspecific interaction. Proteins may preferentially interact with specific lipids, possibly based on the length of the fatty acid chain. Therefore, a suitable detergent should have tails which match the lipids surrounding that membrane protein.

Considering that lipids are hypothesised to be involved in pore formation, CNI may be interacting with those same lipids to prevent pore-formation (Yeagle, 2014). As well as potentially being important in channel formation by ColN, lipids have been shown to be involved in the formation of the potassium channel Kir2.1 (Soom *et al.*,

2001). For the Kir 2.1 channel lipid binding induces a conformational change in a flexible linker region, which results in reorganisation of the entire channel structure and finally in channel activation (Hansen *et al.*, 2011). It is possible that ColN channel formation works according to a similar principle.

Proteins may preferentially interact with specific lipids, possibly based on the length of the fatty acid chain (Stangl *et al.*, 2012; Stangl and Schneider, 2015). Therefore, a suitable detergent should have tails which match the lipids surrounding that membrane protein. Possibly, none of the detergents tested here are perfect substitutes for non-annular lipids which would support the intramolecular structure of the protein as well as annular lipids which would prevent non-specific binding and aggregation.

Investigating the lipids associated with CNI would be interesting future work, which could be carried out with the creation of nanodiscs using SMALPs (Dorr *et al.*, 2014), followed lipid analysis using HPLC or thin layer chromatography. Also, an NMR structure may resolve the question if the additional 700 Da are a non-fuller lipid and what other lipids are associated with CNI.

When carrying out the spot and cell survival assay with ColN as well as ColA it was noticed that compared to ColN higher concentrations of ColA were needed to kill the cells. This could have a variety of reasons. ColA is prone to degradation and may have degraded or aggregated during handling and the samples may therefore contain fewer active molecules. However, Ridley and Lakey (2015) also reported a difference between ColA and ColN potency in spot test and cell survival assays, which are related to the differences in translocation through periplasm.

6.1.5 In vivo CNI observation using fluorescent microscopy

A GFP-tagged version of CNI proved particularly useful in characterising CNI. It allowed tracking the CNI-sfGFP-6His fusion during purification, was used to screen detergents for solubilisation and was also used in fluorescent microscopy and single molecule tracking. Although a GFP-tagged version of Colicin K immunity protein (Mulec *et al.*, 2003) has been created and observed using epifluorescent microscopy before, CNI is the first immunity protein to be clearly localised to the cell periphery. Furthermore, Mulec *et al.* (2003) does not demonstrate CKI-GFP activity, while CNI-sfGFP-His activity has been shown extensively in this work. It is also the first time an

immunity protein has been observed in TIRF and analysed for how its behaviour changes *in vivo* when ColN is added.

So far expression of immunity proteins to pore-forming colicins has always assumed to be constitutive and low (Cascales *et al.*, 2007), but a quantitative count of the molecules has not been carried out. When viewing the cells in TIRF for the single molecule tracking experiments, it was also attempted to image cells which express CNI regulated by its natural promoter $P_{CNI.}$ However, it was very difficult to see any fluorescent molecules at all (data not shown). A low signal to noise ratio is largely responsible for that but it also certainly indicates that the expression is very low. To quantify expression better, quantitative PCR could be used to establish transcription levels and establish if they are influenced by similar stresses as colicin expression. This work showed, using a liquid culture fluorescence assay (Chapter 5), that addition of sublethal ColN amounts did not increase the total amount of npCNI-sfGFP-6His in the cell, which means that expression was not upregulated and there was no increased stabilisation of the protein, like a protection from breakdown through ColN binding. This is further evidence suggesting that the CNI is constitutively expressed in nature and the few molecules that are present are extremely efficient at patrolling the inner membrane and protecting the cell. However, other factors such as culture density, temperature (growth temperature, heat or cold shock) and nutrient change or starvation, and other factors that can trigger Colicin N production naturally, such as DNA damage, for example, through mytomycin C, were not examined (Tsao and Goebel, 1969; Hausmann and Clowes, 1971; Nakazawa and Tamada, 1972; Pugsley, 1984a). This technique might not be accurate or sensitive enough to detect minor changes in expression or protein stability.

In addition to the *in vitro* experiments with SPR and the pull-down assay, single – molecule tracking and TIR-CP (total internal reflection fluorescent microscopy with continuous photo-bleaching) were used to show a direct and specific interaction of CNI with ColN *in vivo*. TIR-CP was used to assess ColN induced clustering, while single molecule tracking looked for a change in diffusion rate upon ColN addition.

Generally, the *E. coli* inner membrane is a dynamic structure, mainly due to constant cell growth, membrane turn over and random movement. Despite these dynamics, some membrane proteins can still maintain a certain degree of localisation within this environment, which is usually determined by their function. TolA, for example, is a

large, periplasmic space spanning protein which is anchored in the inner membrane but also interacts with Pal in the outer membrane (Ridley and Lakey, 2015). It localises to the cell septum late in cell division (Gerding *et al.*, 2007). The Tol system plays a prominent role in colicin translocation for several colicins, including ColA and ColN. Therefore, we hypothesised that CNI may co-localise with TolA and wait for ColN to reach the inner membrane before it inactivates it or somehow interfere with ColN binding to TolA. TIR-CP microscopy and single molecule tracking have not provided any evidence to support this hypothesis. CNI is, in comparison to TolA, more evenly distributed in the cell membrane and, overall, diffuses faster than TolA, therefore making a permanent interaction unlikely, although a transient interaction is still possible. It is remains unclear if CNI has any other interaction partners apart from ColN. Gerding *et al.* (2007) showed the distribution of other components of the Tol system but none resembles the distribution observed for CNI.

ColN P-domain goes through a dramatic transformation and helix rearrangement when inserting into the inner membrane, following the umbrella (Parker *et al.*, 1992; Padmavathi and Steinhoff, 2008; Bohme *et al.*, 2009) or penknife model (Lakey *et al.*, 1993; Massotte *et al.*, 1993; Duche *et al.*, 1996) or something in between, and so it was hypothesised that CNI could dramatically change distribution or diffusion in the membrane when ColN is added. A measurement of overall culture fluorescence showed that overall CNI abundance is not increased through a reduction in turnover or expression upregulation. Surprisingly, neither clustering, specific localisation nor a specific reduction in diffusion rate is observed. Interestingly, there was an overall, non-specific effect on protein diffusion rates in the membrane when ColN and ColA were added, which might be related to the dissipation of membrane potential. To test this hypothesis a membrane dissipating small molecule like Carbonyl cyanide m-chlorophenyl hydrazone (CCCP) could be used (Kasianowicz *et al.*, 1984).

Apart from the previously discussed difficulties in determining how to analyse the data and whether to include non-moving particles in the calculation of diffusion rates (Chapter 5), one fundamental discovery included that moving particles of the same type move at different diffusion rates. The frequency distribution plots in Chapter 5 show that the diffusion rate of a protein in a given condition can vary from 0 to 0.3 µm/sec. This discovery is important because it questions the purpose of citing a quantitative number for the diffusion coefficient, an average of speeds which are in terms of statistics not normally distributed. This work shows that the distribution is

dramatically skewed towards the slow and non-moving end of the spectrum. Also, the statistical distribution of speeds as well as the speed itself is unlikely to be the same for any two proteins because it is determined by the protein function and properties, such as size and shape, unless these proteins form a heterodimer. This discovery also highlights the well-known fact that the membrane is not a homogeneous environment, consisting of less and more rigid areas caused by different protein and lipid compositions, thereby influencing the molecules' movement within (Zerrouk *et al.*, 2008; Papanastasiou *et al.*, 2013; Bramkamp and Lopez, 2015). Interestingly, WALP23-eGFP, an artificial non-functional transmembrane helix appears to be more clustered than CNI, which makes it a poor negative control within the experiment but also poses the question how CNI is so evenly distributed in a non-homogeneous environment. This property may be essential for its ability to protect the entire inner membrane.

The underlying assumption for these *in vivo* measurements of clustering and speed is a direct interaction between CNI and ColN, but it is still unknown whether pore-forming colicins and their immunity proteins interact permanently and in a 1:1 stoichiometry or whether more molecules of an immunity protein are needed to neutralise a single colicin molecule. While a change in speed, if there is one, should be observable, it is doubtful if a clustering effect upon ColN addition can be seen if the ratio is indeed 1:1, as a cluster is defined as the accumulation of more than one molecule of CNI-sfGFP-6His in a given location. Whether a fluorescent signal is coming from one or several molecules can be determined by the way it bleaches under laser light. One molecule shows a sharp drop in signal, characteristic for an on/off switch, while several molecules would show a slower decrease in fluorescence, which looks like descending stairs, where one molecule is bleached at a time.

6.2 Conclusions

To the best of my knowledge, the only attempt to characterise the immunity protein to Colicin N is mentioned by Pugsley (1988) and has until now remained unstudied.

This works presents a revised sequence for the *cni* gene and its promoter P_{CNI} , which reclassifies CNI into an A-type immunity protein with 4, not 3, transmembrane helices. Recombinant expression and purification of CNI was optimised and its

activity characterised, concluding that expression as a fusion with a stabilising C-terminal HALO7-6His tag yielded the most protein when expressed in C41 *E. coli* cells at 37 °C overnight in Terrific Broth. The fusion protein increased resistance to Colicin N by a factor of 1000 and is therefore folded and localised correctly. A variety of detergents were screened to effectively solubilise and stabilise CNI and decyl β-D-maltopyranoside was determined to be the most effective detergent.

The creation of CNI/CAI chimeric proteins showed that all helices were crucial for protein stability and indicated that CNI has a fixed tertiary structure. The first helix is not essential for CNI specificity to ColN and the only cysteine residue of the protein has no significant role in protein function. An *in vitro* condition for specific CNI-ColN interaction in pull-down assays or SPR could not be created.

A fusion with sfGFP was used to determine CNI behaviour *in vivo* when ColN was added using TIR-CP and single molecule tracking. CNI is active when fused to sfGFP and localised to the cell periphery but does not compartmentalise or form clusters, in fact it is the most evenly distributed protein among those investigated here. CNI does not experience a change in diffusion rate specific to when ColN is added, although there seems to be a general effect on membrane protein mobility when a pore-forming colicin is added, possibly caused by the dissipation of membrane potential. This work has found no evidence of CNI interacting with any other membrane proteins, most notably TolA.

6.3 Future work

Future work could pursue any of the three approaches used here to study CNI. *In vitro* work should explore alternative methods of mimicking the cell membrane, such as membrane scaffold proteins, nanodiscs or lipid-detergents mixtures. The CNI capture after protease cleavage could also certainly be optimised.

Site-directed mutagenesis could be used to narrow down residues essential in CNI specificity to ColN. The hypothesis of hydrophobic helix-helix interaction between pore forming colicins and their immunity proteins could be tested further by creating a chimeric protein where all 4 helices are swapped.

In vivo work with sfGFP-tagged GFP could consider CNI localisation in mutant strains, such as ΔTolA strain. Continuing from the general effect observed on

membrane proteins when ColN is added, CCCP could be used to test if this change in diffusion rate is caused by membrane potential dissipation (Strahl and Hamoen, 2010).

Appendix

Supplementary figures for Chapter 3: CNI overexpression optimization

The below figures show the full overexpression optimization screen carried out at OPPF, Harwell, UK.

Below is a repeat of the protein fusion sizes for convenience.

Table 13: Protein and tag combination used during the overexpression screen, including the tag and protein fusion sizes. Vector names refer to nomenclature used by OPPF[12]. All were cloned into pOPIN vectors with spectinomycin (pOPINCD vectors) or ampicillin selection.

Vector name	Tag	Tag MW in Da	MW of fusions with CNI (20 431 Da) in Da	Fusions with CNI-TEV-GFP (49 882 Da)
pOPINCDE	C-His	969	21 400	x
pOPINCDF	N-His	2 158	22 589	x
pOPINS3C	N-His-SUMO3C	13 213	33 644	63 095
pOPINMSYB	N-His-MSYB	16 268	36 699	66 150
pOPINCDJ	N-His-GST	27 954	48 385	77 836
pOPINE-3C-eGFP	eGFP-6His-C	28 645	49 076	x
pOPINE-3C-HALO	HALO7-6His-C	35 343	55 774	x
pOPINCDM	N-His-MBP	42 711	63 142	92 593

Table 14: Periplasmic target sequences linked to ColN P-domain (in 23 Da) in pOPIN vectors with ampicillin selection for co-expression with CNI fusions in pOPINCD vectors. Vector names refer to nomenclature used by OPPF.

Vector	Periplasmic target sequence
pOPINO	Omp A SS (co-express with CD vectors)
pOPINP	PelB SS (co-express with CD vectors)
pOPINDsbA	Dsb A SS (co-express with CD vectors)
pOPINMalE	PelB SS (co-express with CD vectors)

[12] http://www.oppf.rc-harwell.ac.uk/OPPF/protocols/cloning.jsp

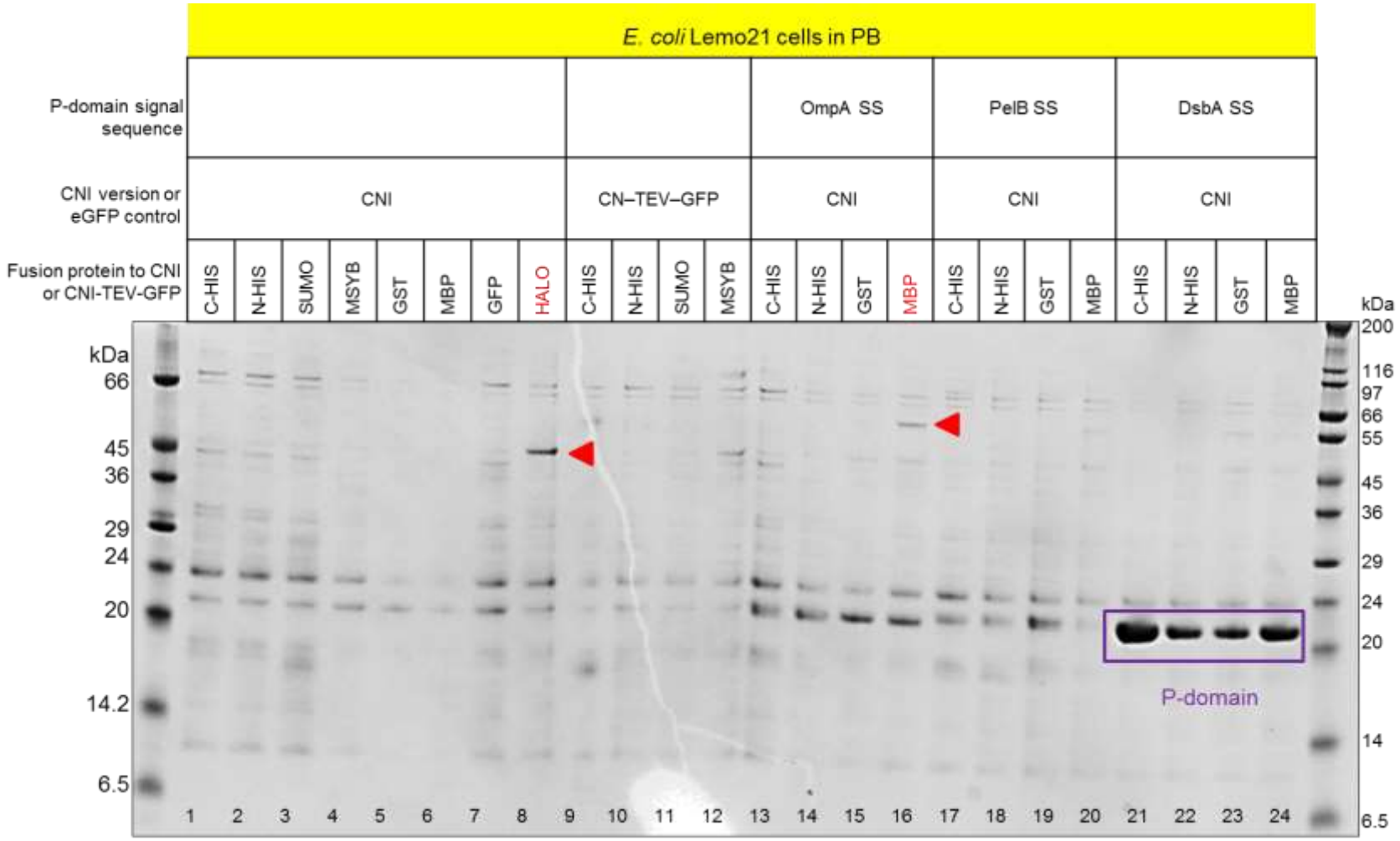


Figure 95: Coomassie-stained gel showing CNI protein fusions purified by IMAC from *E. coli* Lemo21 cells, grown in PB media. Constructs showing increased expression of correctly sized protein (Table 7) are highlighted red in the legend and indicated by red arrows. Where ColN P-domain (23kDa) is co-expressed, the signal sequence is given. Although ColN P-domain is coexpressed (purple square), it does not improve CNI expression.

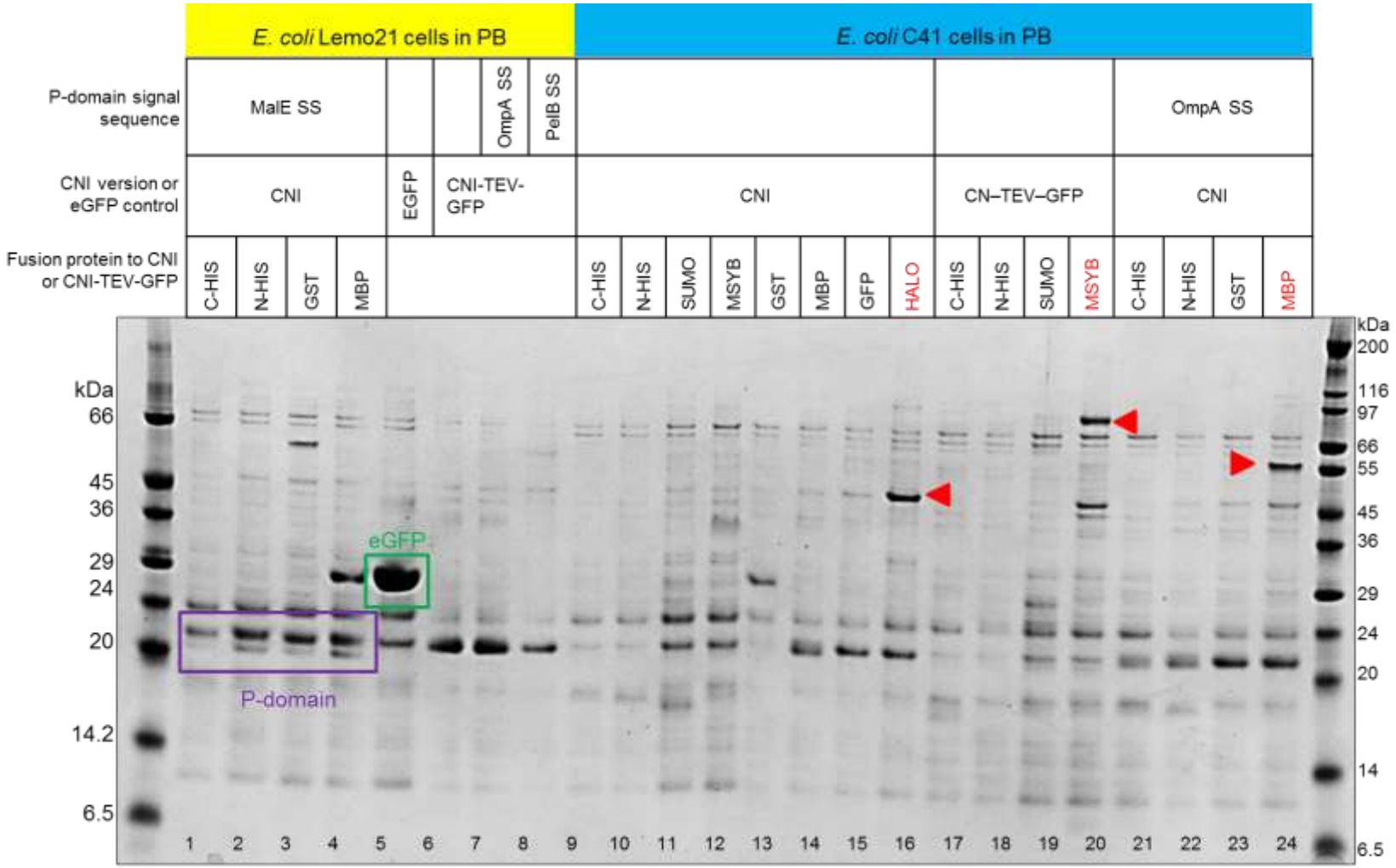


Figure 96: Coomassie-stained gel showing CNI protein fusions purified by IMAC from *E. coli* Lemo21 and C41 cells, grown in PB media. Constructs showing increased expression of correctly sized protein (Table 7) are highlighted red in the legend and indicated by red arrows. Where ColN P-domain (23kDa) is co-expressed, the signal sequence is given. Although ColN P-domain is coexpressed (purple square), it does not improve CNI expression. The eGFP control is indicated with a green square.

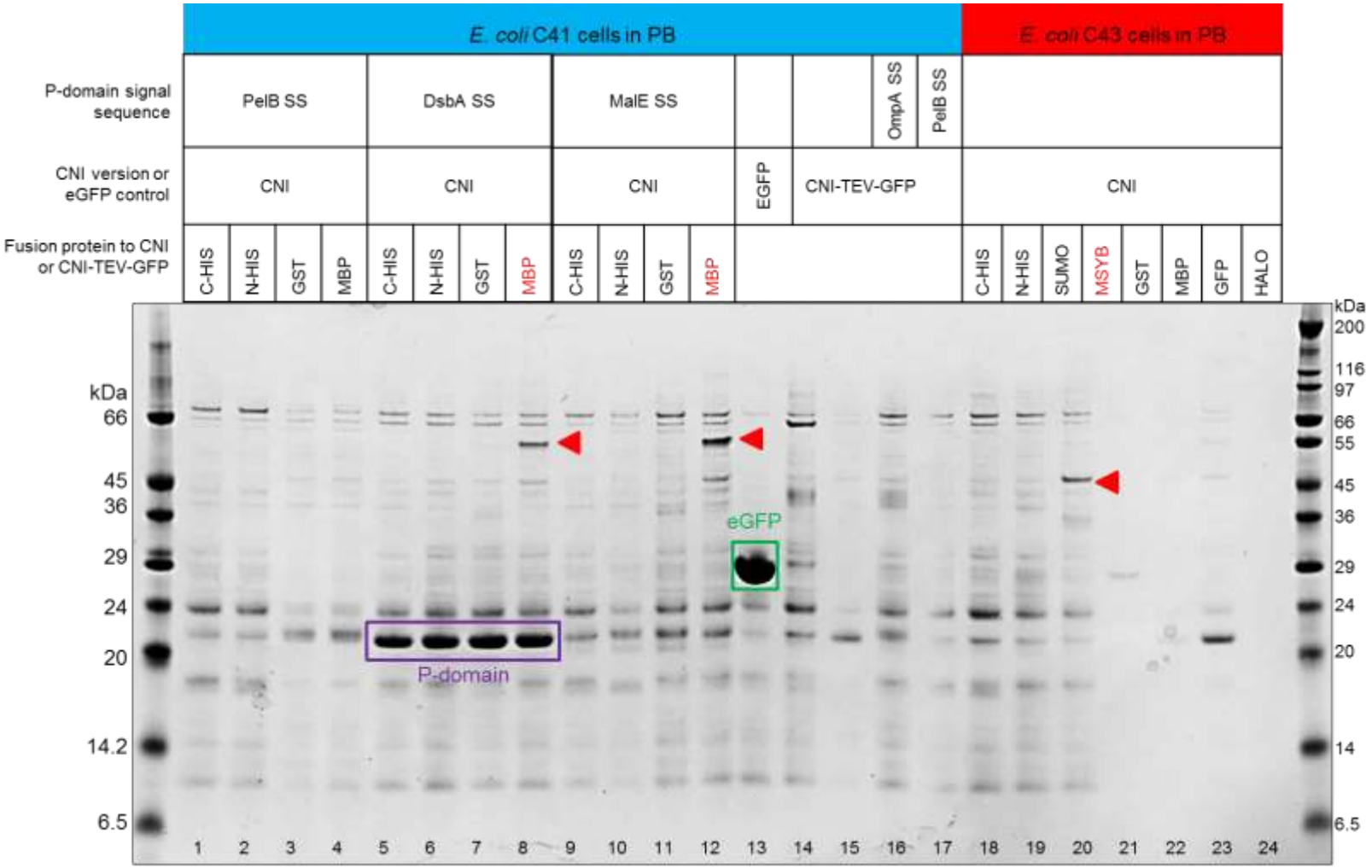


Figure 97: Coomassie-stained gel showing CNI protein fusions purified by IMAC from *E. coli* C41 and C43 cells, grown in PB media. Constructs showing increased expression of correctly sized protein (Table 7) are highlighted red in the legend and indicated by red arrows. Where ColN P-domain (23kDa) is co-expressed, the signal sequence is given. Although ColN P-domain is coexpressed (purple square), it does not improve CNI expression. The eGFP control is indicated with a green square.

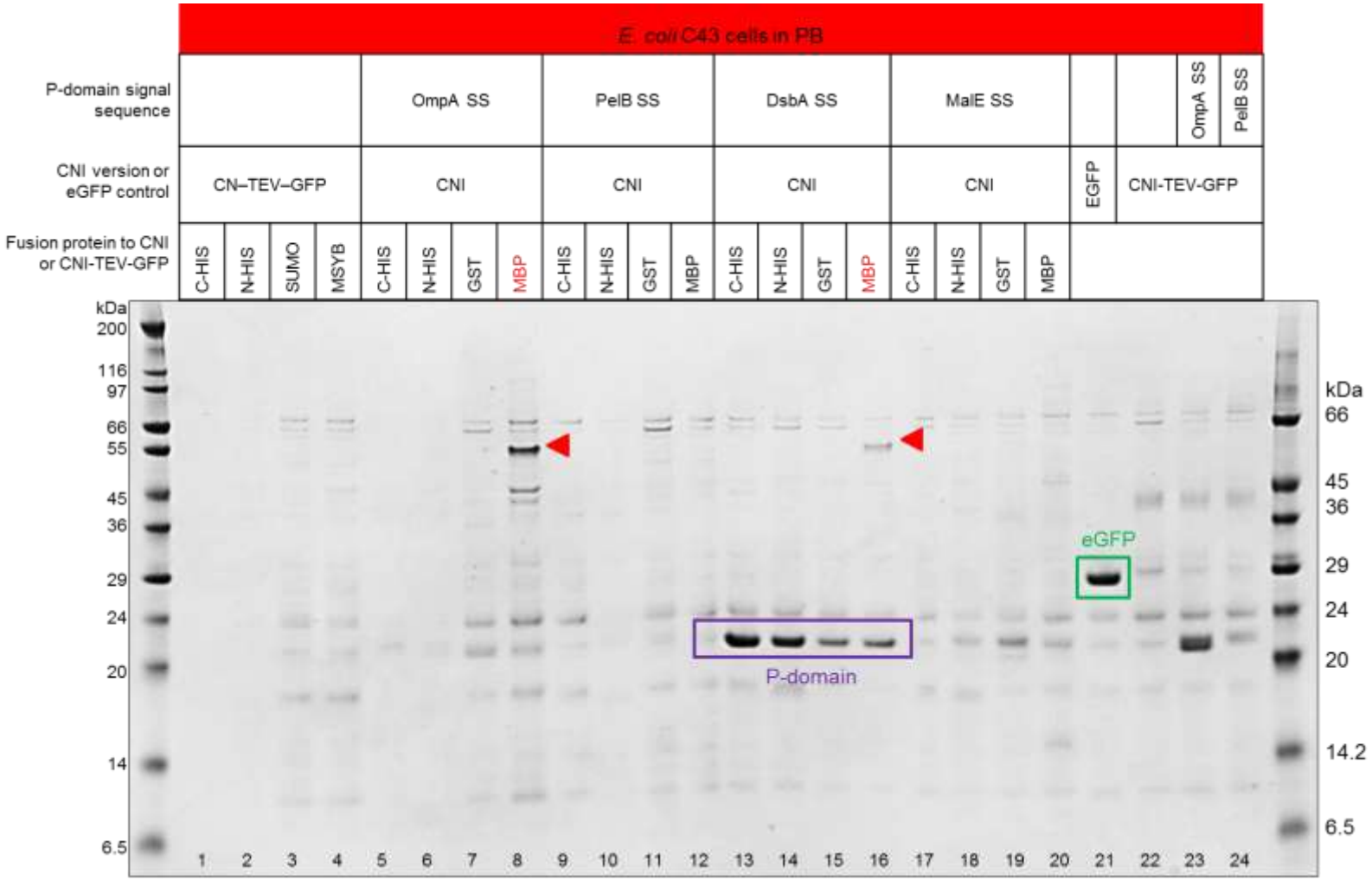


Figure 98: Coomassie-stained gel showing CNI protein fusions purified by IMAC from *E. coli* C43 cells, grown in PB media. Constructs showing increased expression of correctly sized protein (Table 7) are highlighted red in the legend and indicated by red arrows. Where ColN P-domain (23kDa) is co-expressed, the signal sequence is given. Although ColN P-domain is coexpressed (purple square), it does not improve CNI expression. The eGFP control is indicated with a green square.

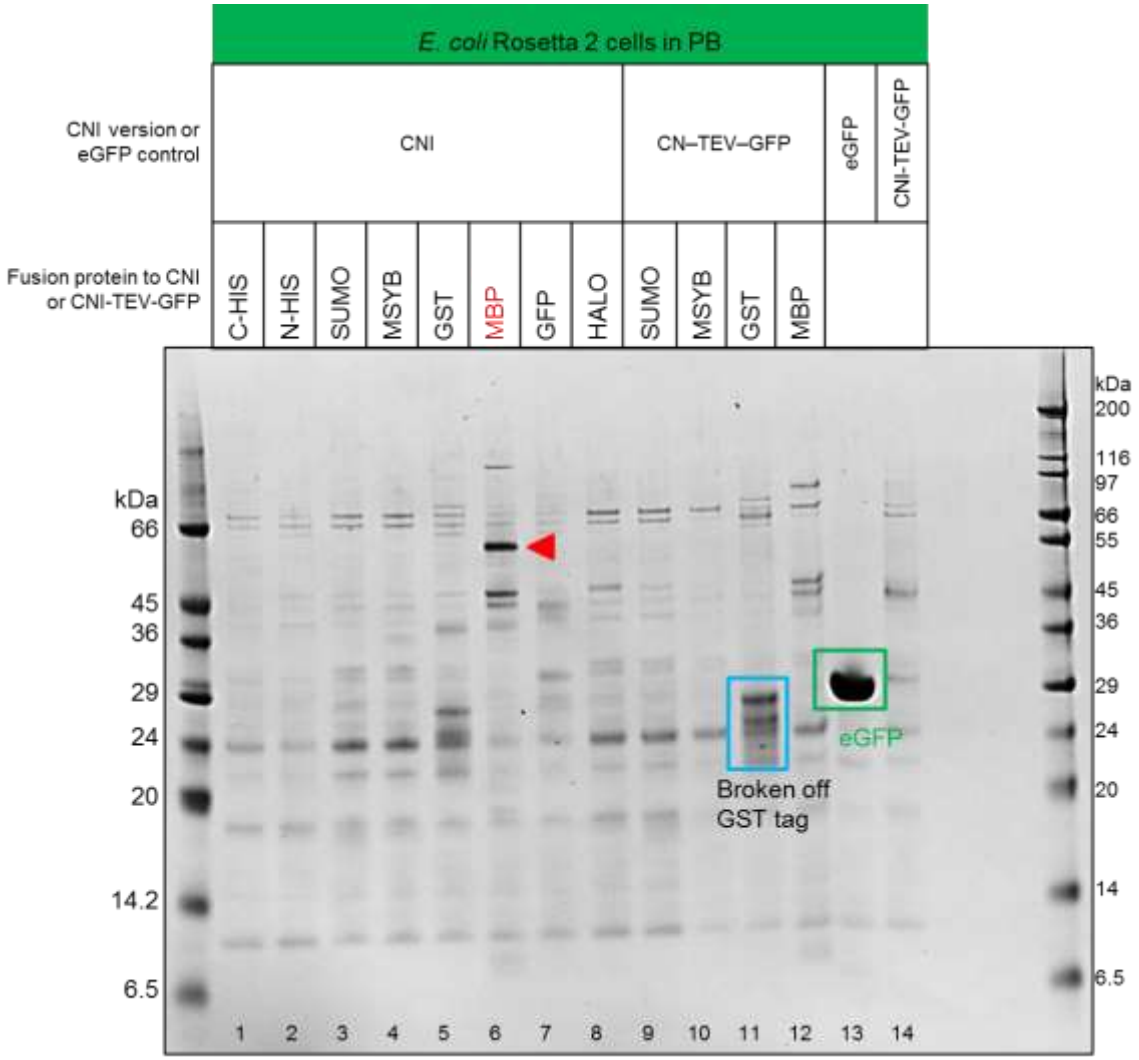


Figure 99: Coomassie-stained gel showing CNI protein fusions purified by IMAC from *E. coli* Rosetta 2 cells, grown in PB media. Constructs showing increased expression of correctly sized (Table 7) are highlighted red in the legend and indicated by red arrow. The eGFP control is indicated with a green square.

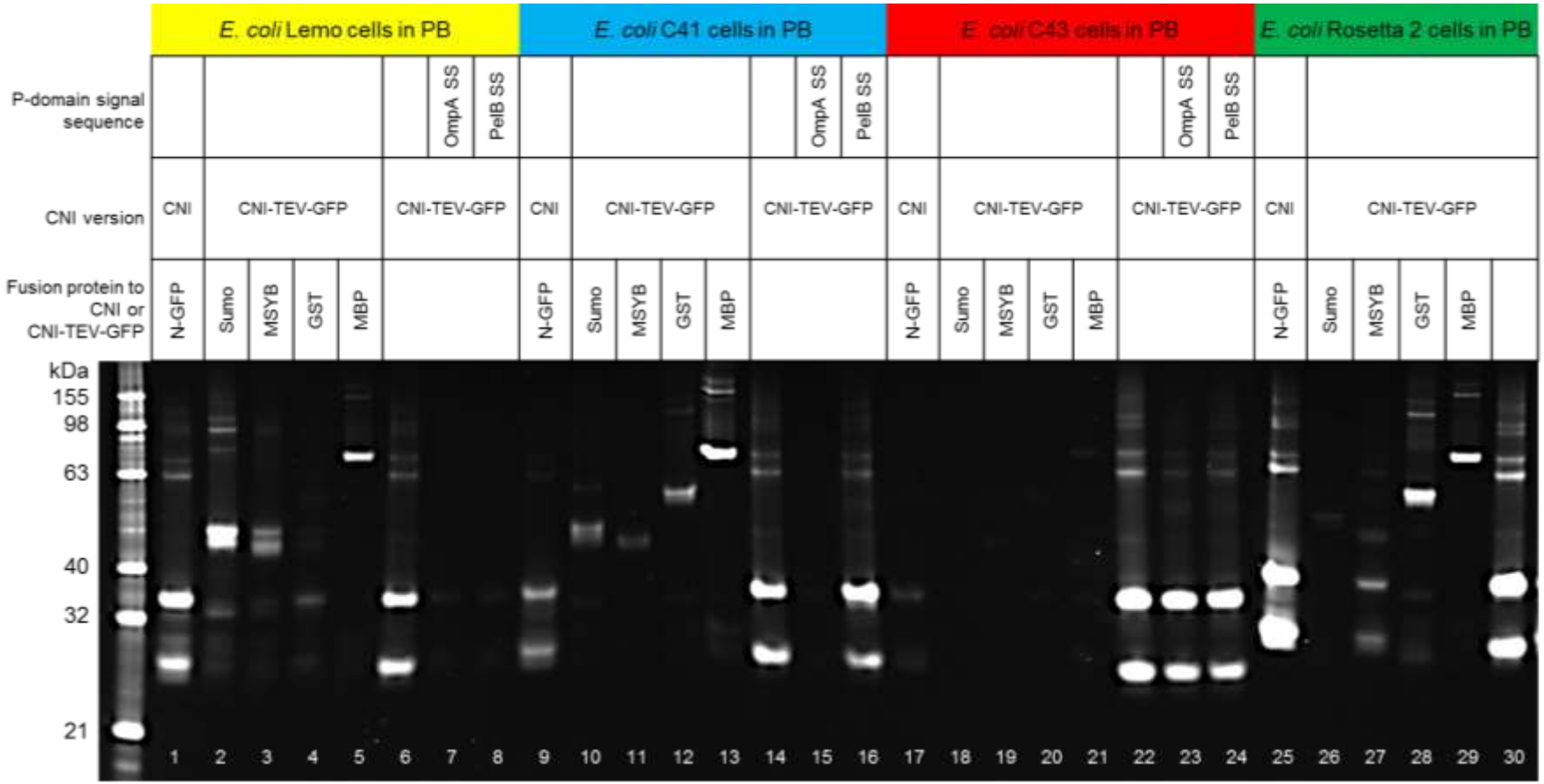


Figure 100: In-gel fluorescence of purified GFP-tagged CNI fusion proteins, expressed in PB media. SDS-PAGE gel identifies which constructs are overexpressed, the relative level of expression in comparison to other conditions and highlights protein degradation and some higher order structures. CNI-TEV-GFP is the previously used reference construct. Constructs without additional solubilisation tags seem to break at the fusion point more easily than constructs with additional proteins. The broken off GFP-tag is around 28 kDa. Most purified proteins show dimerisation, which could indicate aggregation. Where ColN P-domain (23 kDa) is co-expressed, the signal sequence is given. ColN P-domain co-expression does not improve CNI-TEV-eGFP expression levels or stability. Numbers indicate lane number.

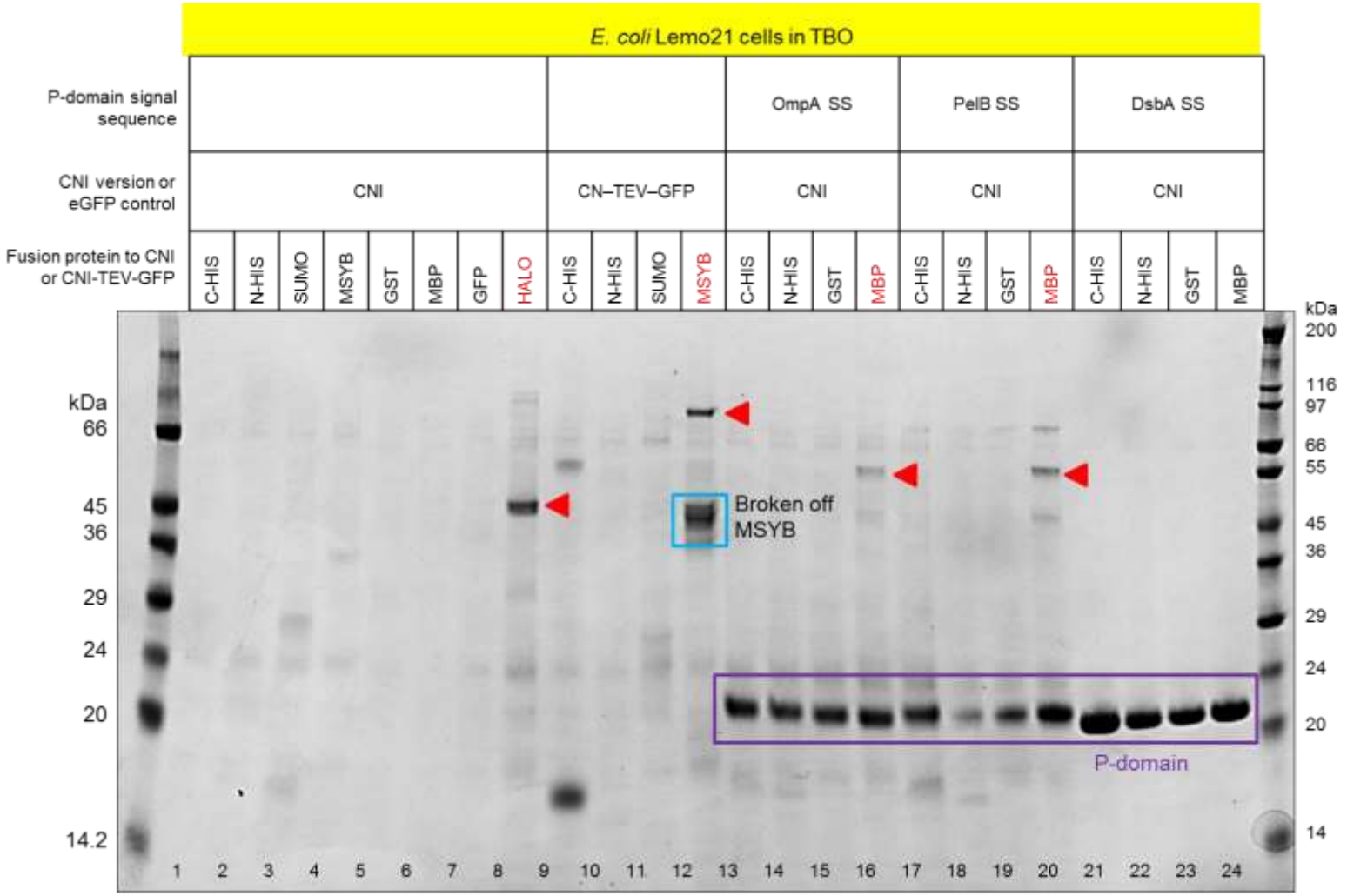


Figure 101: Coomassie-stained gel showing CNI protein fusions purified by IMAC from *E. coli* Lemo21 cells, grown in TBO media. Constructs showing increased expression of correctly sized protein (Table 7) are highlighted red in the legend and indicated by red arrows. Where ColN P-domain (23kDa) is co-expressed, the signal sequence is given. Although ColN P-domain is coexpressed (purple square), it does not improve CNI expression.

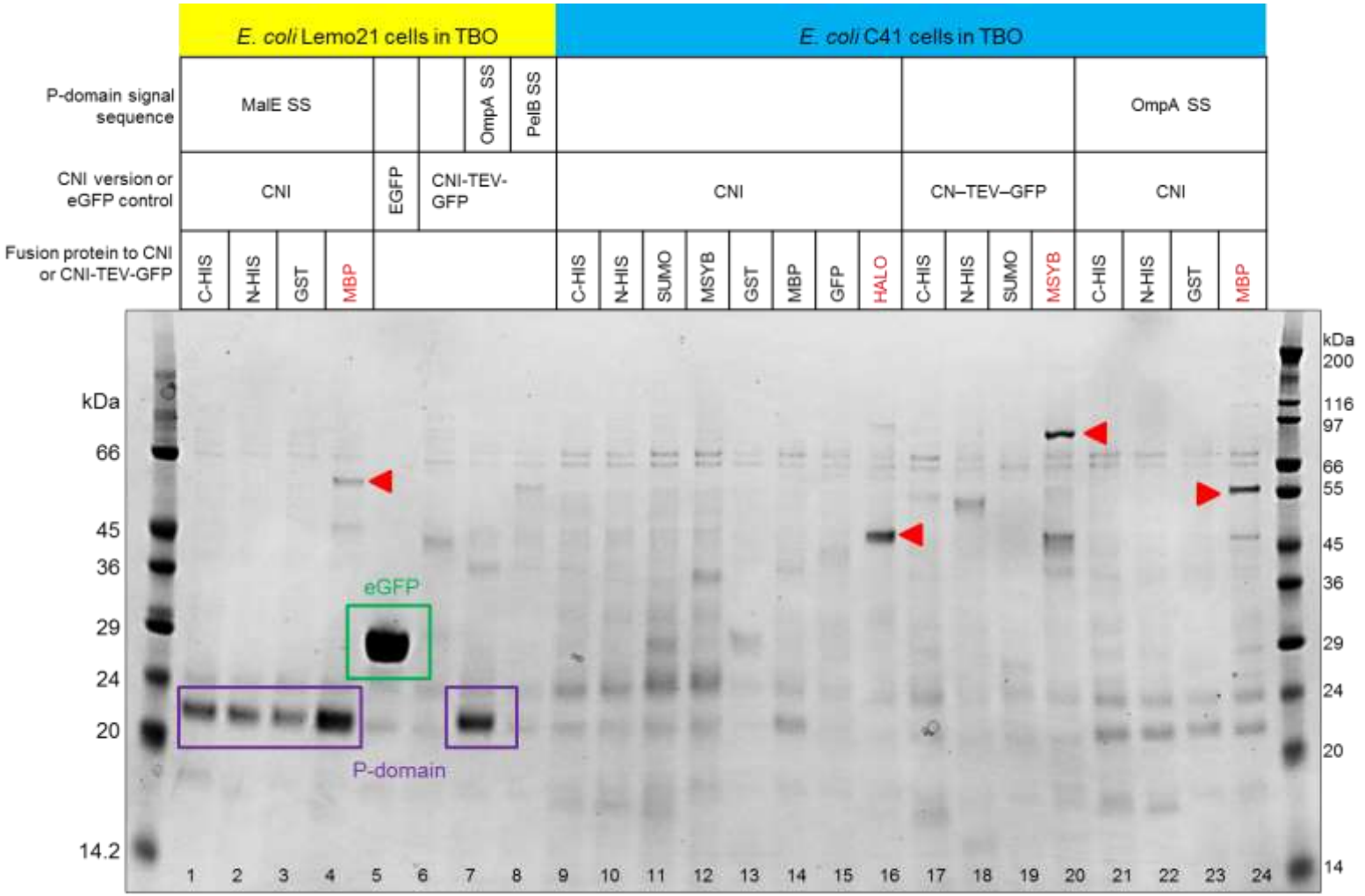


Figure 102: Coomassie-stained gel showing CNI protein fusions purified by IMAC from *E. coli* Lemo21 and C41 cells, grown in TBO media. Constructs showing increased expression of correctly sized protein (Table 7) are highlighted red in the legend. Where ColN P-domain (23kDa) is co-expressed, the signal sequence is given. Although ColN P-domain is coexpressed (purple square), it does not improve CNI expression. The eGFP control is indicated with a green square.

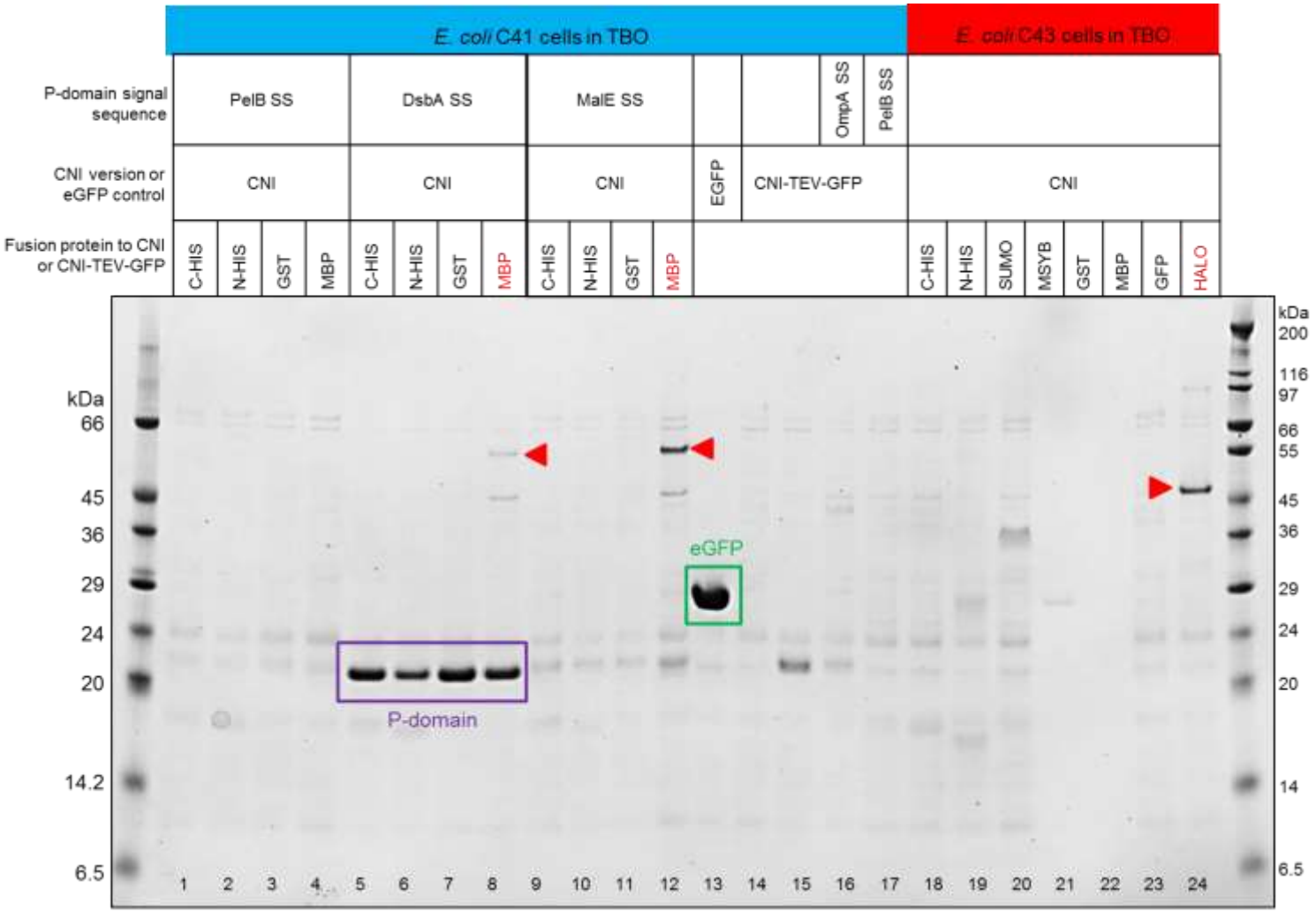


Figure 103: Coomassie-stained gel showing CNI protein fusions purified by IMAC from *E. coli* C41 and C43 cells, grown in TBO media. Constructs showing increased expression of correctly sized protein (Table 7) are highlighted red in the legend and indicated by red arrows. Where ColN P-domain (23kDa) is co-expressed, the signal sequence is given. Although ColN P-domain is coexpressed (purple square), it does not improve CNI expression. The eGFP control is indicated with a green square.

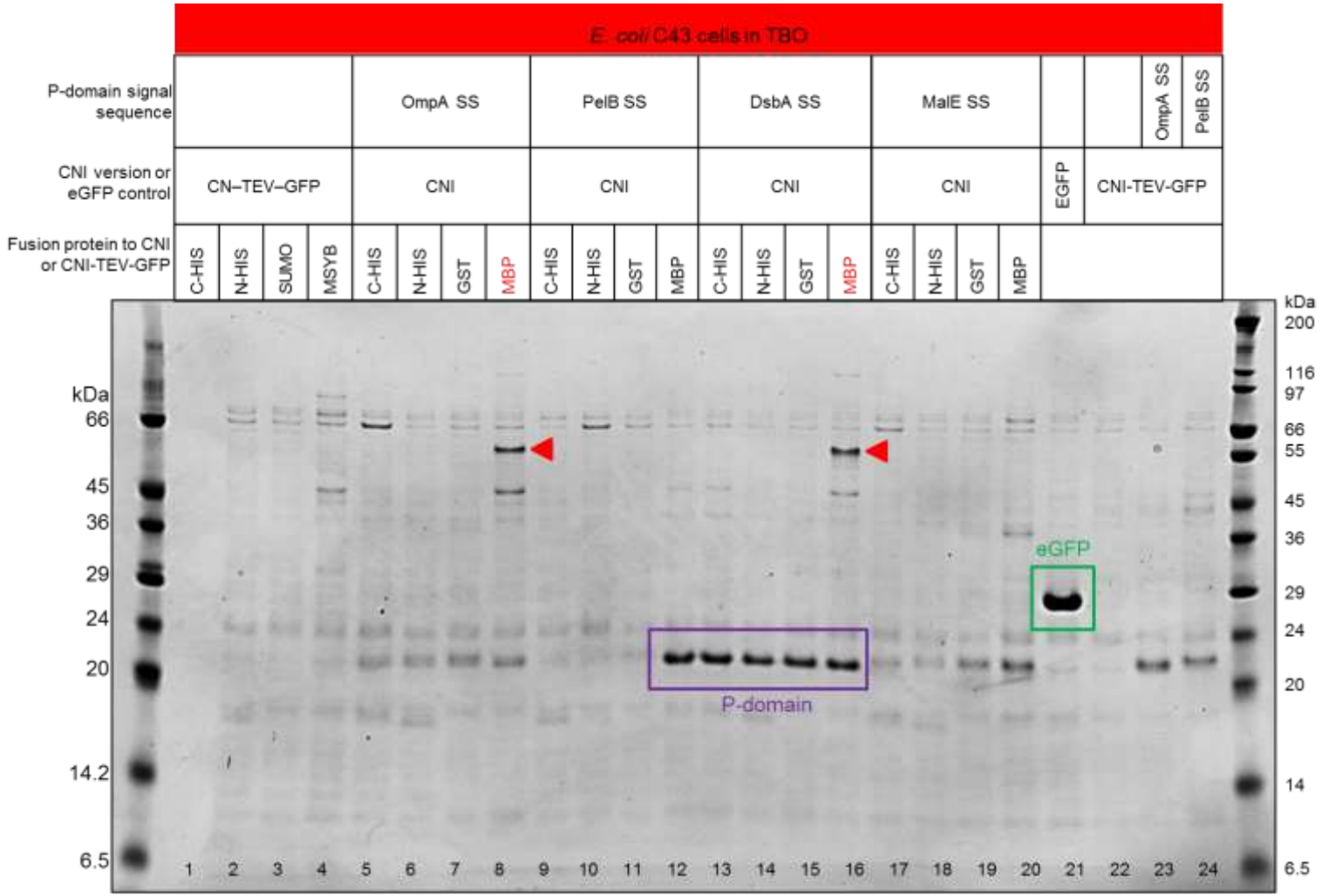


Figure 104: Coomassie-stained gel showing CNI protein fusions purified by IMAC from *E. coli* C43 cells, grown in TBO media. Constructs showing increased expression of correctly sized protein (Table 7) are highlighted red in the legend and indicated by red arrows. Where ColN P-domain (23kDa) is co-expressed, the signal sequence is given. Although ColN P-domain is coexpressed (purple square), it does not improve CNI expression. The eGFP control is indicated with a green square.

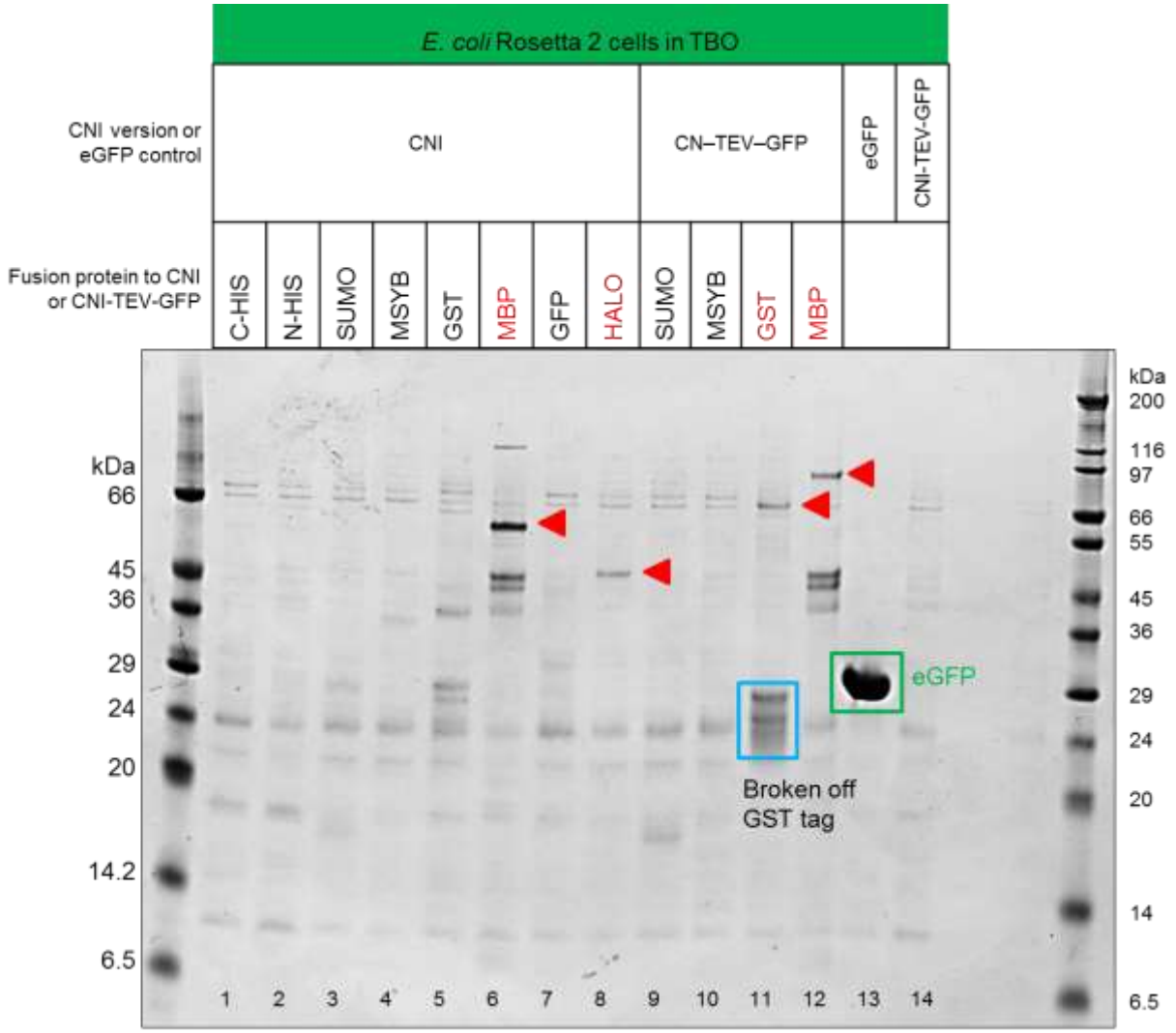


Figure 105: Coomassie-stained gel showing CNI protein fusions purified by IMAC from *E. coli* Rosetta 2 cells, grown in TBO media. Constructs showing increased expression of correctly sized protein (Table 7) are highlighted red in the legend and indicated by red arrows. The eGFP control is indicated with a green square.

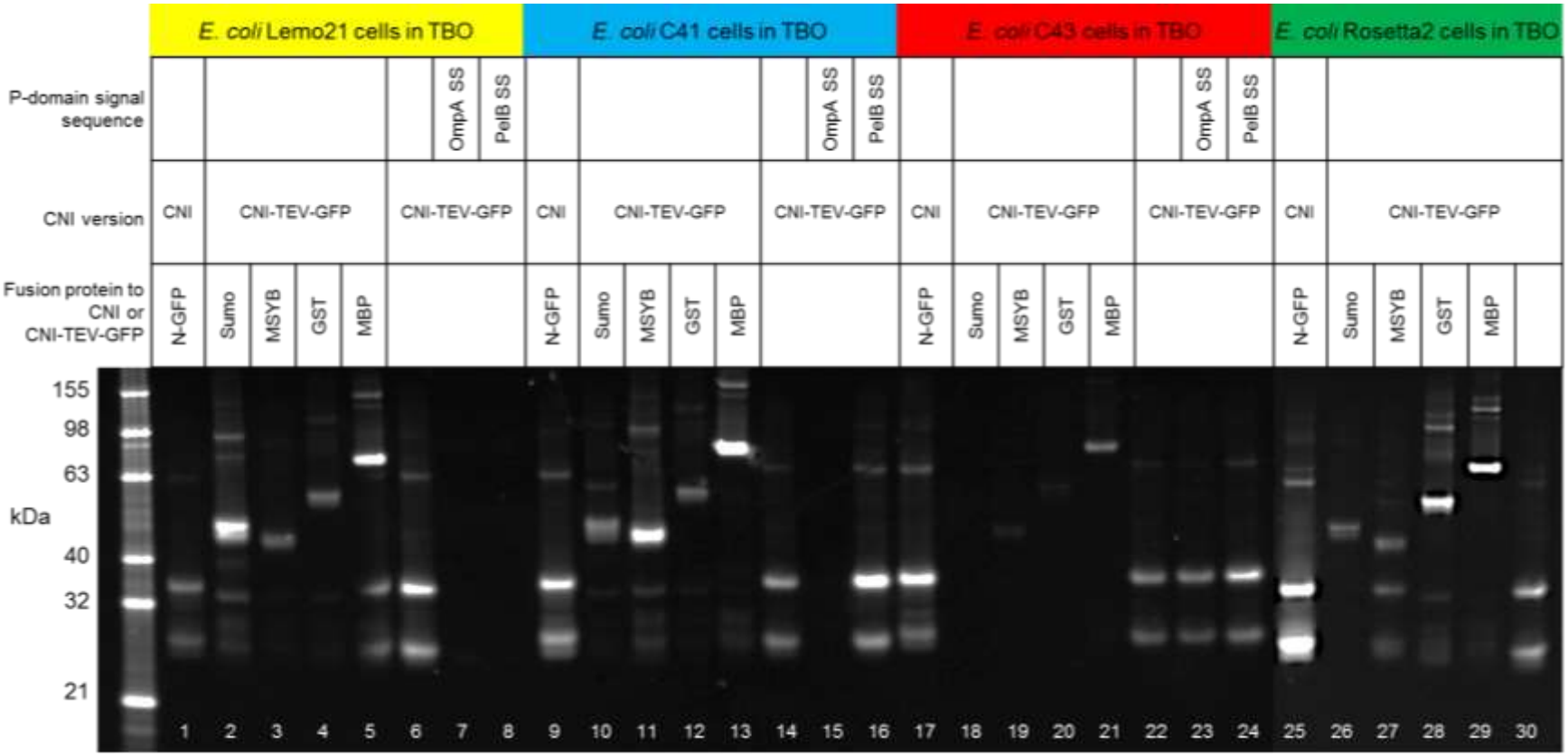


Figure 106: In-gel fluorescence of purified GFP-tagged CNI fusion proteins, expressed in TBO media. SDS-PAGE gel identifies which constructs are overexpressed, the relative level of expression in comparison to other conditions and highlights protein degradation and some higher order structures. CNI-TEV-GFP is the previously used reference construct. Constructs without additional solubilisation tags seem to break at the fusion point more easily than constructs with additional proteins. The broken off GFP-tag is around 28 kDa. Most purified proteins show dimerisation, which could indicate aggregation. Where ColN P-domain (23 kDa) is co-expressed, the signal sequence is given. ColN P-domain co-expression does not improve CNI-TEV-eGFP expression levels or stability. Numbers indicate lane number.

Supplementary formula for Chapter 5: *In vivo* observation of GFP-tagged Colicin N immunity protein

Model to quantify bleaching kinetics

To determine if there is a change in CNI mobility due to the addition of ColN or ColA, the bleaching kinetics of sfGFP fused to CNI have been calculated using the formula used in Strahl *et al.* (2014), which is based on the principles explained in detail in Slade *et al.* (2009). The normalised and background-subtracted fluorescence intensity decay of CNI-sfGFP is calculated using

$$y = y_0 \times e^{(-Kfast \times x)} + y_0 \times e^{(-Kslow \times x)},$$

where y is fluorescence intensity (starting at 100 % and decreasing towards 0 %), y_0 is fluorescence intensity at t = 0 (beginning of bleaching), K_{fast} is the constant associated with bleaching sfGFP molecules within the reach of the evanescent wave, and K_{slow} is the constant determined by diffusion limited bleaching of molecules entering the TIRF field from the diminishing general pool.

Statement of originality

I hereby declare that this dissertation is my own work and that, to the best of my knowledge, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of the university or any other institute of higher learning, except where due acknowledgment has been made in the text.

September 25th, 2015, Newcastle upon Tyne, UK

Bibliography

Aisenbrey, C., Cusan, M., Larnbotte, S., Jasperse, P., Georgescu, J., Harzer, U. and Bechinger, B. (2008) 'Specific isotope labeling of colicin E1 and B channel domains for membrane topological analysis by oriented solid-state NMR spectroscopy', *Chembiochem*, 9(6), pp. 944-951.

Aisenbrey, C., Sudheendra, U.S., Ridley, H., Bertani, P., Marquette, A., Nedelkina, S., Lakey, J.H. and Bechinger, B. (2007) 'Helix orientations in membrane-associated Bcl-X(L) determined by 15N-solid-state NMR spectroscopy', *Eur Biophys J*, 37(1), pp. 71-80.

Alonso, G., Vilchez, G. and Rodriguez Lemoine, V. (2000) 'How bacteria protect themselves against channel-forming colicins', *Int Microbiol*, 3(2), pp. 81-8.

Arnold, T., Zeth, K. and Linke, D. (2009) 'Structure and function of Colicin S4, a colicin with a duplicated receptor-binding domain', *J Biol Chem*, 284(10), pp. 6403-6413.

Arunmanee, W., Harris, J.R. and Lakey, J.H. (2014) 'Outer membrane protein F stabilised with minimal amphipol forms linear arrays and LPS-dependent 2D crystals', *J Membr Biol*, 247(9-10), pp. 949-56.

Audisio, M.C., Terzolo, H.R. and Apella, M.C. (2005) 'Bacteriocin from honeybee beebread Enterococcus avium, active against Listeria monocytogenes', *Appl Environ Microb*, 71(6), pp. 3373-3375.

Axelrod, D. (2008) 'Total internal reflection fluorescence microscopy', *Method Cell Biol, Vol 2: In Vivo Techniques*, 89, pp. 169-221.

Balciunas, E.M., Martinez, F.A.C., Todorov, S.D., Franco, B.D.G.D.M., Converti, A. and Oliveira, R.P.D. (2013) 'Novel biotechnological applications of bacteriocins: A review', *Food Control*, 32(1), pp. 134-142.

Bao, N., Jagadeesan, B., Bhunia, A.K., Yao, Y. and Lu, C. (2008) 'Quantification of bacterial cells based on autofluorescence on a microfluidic platform', *J Chromatogr A*, 1181(1-2), pp. 153-8.

Barneoud-Arnoulet, A., Barreteau, H., Touze, T., Mengin-Lecreulx, D., Lloubes, R. and Duche, D. (2010) 'Toxicity of the Colicin M catalytic domain exported to the periplasm Is FkpA Independent', *J Bacteriol*, 192(19), pp. 5212-5219.

Barreteau, H., Bouhss, A., Gerard, F., Duche, D., Boussaid, B., Blanot, D., Lloubes, R., Mengin-Lecreulx, D. and Touze, T. (2010) 'Deciphering the catalytic domain of Colicin M, a peptidoglycan lipid II-degrading enzyme', *J Biol Chem*, 285(16), pp. 12378-12389.

Barreteau, H., El Ghachi, M., Barneoud-Arnoulet, A., Sacco, E., Touze, T., Duche, D., Gerard, F., Brooks, M., Patin, D., Bouhss, A., Blanot, D., van Tilbeurgh, H., Arthur, M., Lloubes, R. and Mengin-Lecreulx, D. (2012) 'Characterization of Colicin M and its orthologs targeting bacterial cell wall peptidoglycan biosynthesis', *Microb Drug Resist*, 18(3), pp. 222-229.

Baty, D., Lakey, J., Pattus, F. and Lazdunski, C. (1990) 'A 136-amino-acid-residue cooh-terminal fragment of Colicin-a is endowed with ionophoric activity', *Eur J Biochem*, 189(2), pp. 409-413.

Bayburt, T.H. and Sligar, S.G. (2010) 'Membrane protein assembly into nanodiscs', *Febs Lett*, 584(9), pp. 1721-1727.

Benedetti, H., Frenette, M., Baty, D., Knibiehler, M., Pattus, F. and Lazdunski, C. (1991a) 'Individual domains of colicins confer specificity in colicin uptake, in pore-properties and in immunity requirement', *J Mol Biol*, 217(3), pp. 429-39.

Benedetti, H., Lazdunski, C. and Lloubes, R. (1991b) 'Protein import into *Escherichia coli* - Colicin-a and Colicin-E1 interact with a component of their translocation system', *Embo J*, 10(8), pp. 1989-1995.

Besenicar, M., Macek, P., Lakey, J.H. and Anderluh, G. (2006) 'Surface plasmon resonance in protein-membrane interactions', *Chem Phys Lipids*, 141(1-2), pp. 169-78.

Bird, L.E., Rada, H., Verma, A., Gasper, R., Birch, J., Jennions, M., Lwe, J., Moraes, I. and Owens, R.J. (2015) 'Green fluorescent protein-based expression screening of membrane proteins in *Escherichia coli*', *J Vis Exp*, (95), p. e52357.

Bishop, L.J., Bjes, E.S., Davidson, V.L. and Cramer, W.A. (1985) 'Localization of the immunity protein-reactive domain in unmodified and chemically modified cooh-terminal peptides of Colicin E1', *J Bacteriol*, 164(1), pp. 237-244.

Blommel, P.G. and Fox, B.G. (2007) 'A combined approach to improving large-scale production of tobacco etch virus protease', *Protein Expr Purif*, 55(1), pp. 53-68.

Bohme, S., Padmavathi, P.V.L., Holterhues, J., Ouchni, F., Klare, J.P. and Steinhoff, H.J. (2009) 'Topology of the amphipathic helices of the colicin A pore-forming domain in *E. coli* lipid membranes studied by pulse EPR', *Phys Chem Chem Phys*, 11(31), pp. 6770-6777.

Booth, P.J. (1997) 'Folding alpha-helical membrane proteins: kinetic studies on bacteriorhodopsin', *Fold Des*, 2(6), pp. R85-92.

Bramkamp, M. and Lopez, D. (2015) 'Exploring the existence of lipid rafts in bacteria', *Microbiol Mol Biol R*, 79(1), pp. 81-100.

Braun, V. (1974) 'Murein lipoprotein and receptor for T5 phage and Colicin-M - Defined structural and functional areas of outer membrane of *Escherichia coli*', *Zbl Bakt-Int J Med M*, 228(1-2), pp. 233-240.

Braun, V., Patzer, S.I. and Hantke, K. (2002) 'Ton-dependent colicins and microcins: modular design and evolution', *Biochimie*, 84(5-6), pp. 365-380.

Braun, V., Schaller, K. and Wabl, M.R. (1974) 'Isolation, characterization, and action of Colicin M', *Antimicrob Agents Ch*, 5(5), pp. 520-533.

Breukink, E. and de Kruijff, B. (1999) 'The lantibiotic nisin, a special case or not?', *BBA-Biomembranes*, 1462(1-2), pp. 223-234.

Breyton, C., Gabel, F., Lethier, M., Flayhan, A., Durand, G., Jault, J.M., Juillan-Binard, C., Imbert, L., Moulin, M., Ravaud, S., Hartlein, M. and Ebel, C. (2013) 'Small angle neutron scattering for the study of solubilised membrane proteins', *Eur Phys J E Soft Matter*, 36(7), p. 71.

Brown, C.L., Smith, K., McCaughey, L. and Walker, D. (2012) 'Colicin-like bacteriocins as novel therapeutic agents for the treatment of chronic biofilm-mediated infection', *Biochem Soc T*, 40, pp. 1549-1552.

Brown, C.L., Smith, K., Wall, D.M. and Walker, D. (2015) 'Activity of species-specific antibiotics against Crohn's disease-associated adherent-invasive *Escherichia coli*', *Inflamm Bowel Dis*, 21(10), pp. 2372-82.

Brown, S.P., Inglis, R.F. and Taddei, F. (2009) 'Evolutionary ecology of microbial wars: within-host competition and (incidental) virulence', *Evol Appl*, 2(1), pp. 32-39.

Caffrey, M., Li, D.F. and Dukkipati, A. (2012) 'Membrane protein structure determination using crystallography and lipidic mesophases: recent advances and successes', *Biochemistry-US*, 51(32), pp. 6266-6288.

Calabrese, A.N., Watkinson, T.G., Henderson, P.J.F., Radford, S.E. and Ashcroft, A.E. (2015) 'Amphipols outperform dodecylmaltoside micelles in stabilizing membrane protein structure in the gas phase', *Anal Chem*, 87(2), pp. 1118-1126.

Carpenter, E.P., Beis, K., Cameron, A.D. and Iwata, S. (2008) 'Overcoming the challenges of membrane protein crystallography', *Curr Opin Struct Biol*, 18(5), pp. 581-6.

Carr, S., Walker, D., James, R., Kleanthous, C. and Hemmings, A.M. (2000) 'Inhibition of a ribosome-inactivating ribonuclease: the crystal structure of the cytotoxic domain of colicin E3 in complex with its immunity protein', *Struct Fold Des*, 8(9), pp. 949-960.

Casadaban, M.J. and Cohen, S.N. (1980) 'Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*', *J Mol Biol*, 138(2), pp. 179-207.

Cascales, E., Buchanan, S.K., Duche, D., Kleanthous, C., Lloubes, R., Postle, K., Riley, M., Slatin, S. and Cavard, D. (2007) 'Colicin biology', *Microbiol Mol Biol Rev*, 71(1), pp. 158-229.

Cavard, D. (1994) 'Rescue by Vitamin-B-12 of *Escherichia coli* cells treated with Colicins A and E allows measurement of the kinetics of colicin binding on Btub', *Fems Microbiol Lett*, 116(1), pp. 37-42.

Cavard, D. and Lazdunski, C. (1981) 'Involvement of Btub and OmpF proteins in binding and uptake of Colicin A', *Fems Microbiol Lett*, 12(4), pp. 311-316.

Chao, L. and Levin, B.R. (1981) 'Structured habitats and the evolution of anticompetitor toxins in bacteria', *Proc Natl Acad Sci U S A*, 78(10), pp. 6324-8.

Cheng, Y.S., Hsia, K.C., Doudeva, L.G., Chak, K.F. and Yuan, H.S. (2002) 'The crystal structure of the nuclease domain of colicin E7 suggests a mechanism for binding to double-stranded DNA by the H-N-H endonucleases', J Mol Biol, 324(2), pp. 227-36.

Collarini, M., Amblard, G., Lazdunski, C. and Pattus, F. (1987) 'Gating processes of channels induced by Colicin A, its C-terminal fragment and Colicin E1 in planar lipid bilayers', *Eur Biophys J Biophy*, 14(3), pp. 147-153.

Corr, S.C., Li, Y., Riedel, C.U., O'Toole, P.W., Hill, C. and Gahan, C.G.M. (2007) 'Bacteriocin production as a mechanism for the antfinfective activity of *Lactobacillus salivarius* UCC118', *Proc Natl Acad Sci U S A*, 104(18), pp. 7617-7621.

Cotter, P.D., Hill, C. and Ross, R.P. (2005) 'Bacteriocins: Developing innate immunity for food', *Nat Rev Microbiol*, 3(10), pp. 777-788.

Cotter, P.D., Hill, C. and Ross, R.P. (2006) 'What's in a name? Class distinction for bacteriocins - Author reply', *Nat Rev Microbiol*, 4(2).

Cotter, P.D., Ross, R.P. and Hill, C. (2013) 'Bacteriocins - a viable alternative to antibiotics?', *Nat Rev Microbiol*, 11(2), pp. 95-105.

Cramer, W.A., Zhang, Y.L., Schendel, S., Merrill, A.R., Song, H.Y., Stauffacher, C.V. and Cohen, F.S. (1992) 'Dynamic properties of the Colicin E1 ion channel', *Fems Microbiol Immun*, 105(1-3), pp. 71-82.

Cronan, J.E. (2006) 'A family of arabinose-inducible *Escherichia coli* expression vectors having pBR322 copy control', *Plasmid*, 55(2), pp. 152-7.

Curnow, P. and Booth, P.J. (2007) 'Combined kinetic and thermodynamic analysis of alpha-helical membrane protein unfolding', *P Natl Acad Sci USA*, 104(48), pp. 18970-18975.

Cymer, F., Veerappan, A. and Schneider, D. (2012) 'Transmembrane helix-helix interactions are modulated by the sequence context and by lipid bilayer properties', *Biochim Biophys Acta*, 1818(4), pp. 963-73.

Czaran, T.L., Hoekstra, R.F. and Pagie, L. (2002) 'Chemical warfare between microbes promotes biodiversity', *P Natl Acad Sci USA*, 99(2), pp. 786-790.

de Jong, A., van Hijum, S.A.F.T., Bijlsma, J.J.E., Kok, J. and Kuipers, O.P. (2006) 'BAGEL: a web-based bacteriocin genome mining tool', *Nucleic Acids Res*, 34, pp. W273-W279.

de Zamaroczy, M. and Chauleau, M. (2011) 'Colicin killing: foiled cell defense and hijacked cell functions', *Prokaryotic Antimicrobial Peptides: From Genes to Applications*, pp. 255-287.

Della Pia, E.A., Hansen, R.W., Zoonens, M. and Martinez, K.L. (2014) 'Functionalized amphipols: a versatile toolbox suitable for applications of membrane proteins in synthetic biology', *J Membrane Biol*, 247(9-10), pp. 815-826.

Diep, D.B., Godager, L., Brede, D. and Nes, I.F. (2006) 'Data mining and characterization of a novel pediocin-like bacteriocin system from the genome of Pediococcus pentosaceus ATCC 25745', *Microbiology*, 152(Pt 6), pp. 1649-59.

Dirix, G., Monsieurs, P., Dombrecht, B., Daniels, R., Marchal, K., Vanderleyden, J. and Michiels, J. (2004) 'Peptide signal molecules and bacteriocins in Gram-negative bacteria: a genome-wide *in silico* screening for peptides containing a double-glycine leader sequence and their cognate transporters', *Peptides*, 25(9), pp. 1425-40.

Dorr, J.M., Koorengevel, M.C., Schafer, M., Prokofyev, A.V., Scheidelaar, S., van der Cruijsen, E.A., Dafforn, T.R., Baldus, M. and Killian, J.A. (2014) 'Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K+ channel: the power of native nanodiscs', *P Natl Acad Sci U S A*, 111(52), pp. 18607-12.

Dover, L.G., Evans, L.J.A., Fridd, S.L., Bainbridge, G., Raggett, E.M. and Lakey, J.H. (2000) 'Colicin pore-forming domains bind to *Escherichia coli* trimeric porins', *Biochemistry-US*, 39(29), pp. 8632-8637.

Drew, D., Lerch, M., Kunji, E., Slotboom, D.J. and de Gier, J.W. (2006) 'Optimization of membrane protein overexpression and purification using GFP fusions', *Nat Methods*, 3(4), pp. 303-313.

Duche, D., Izard, J., Gonzalez-Manas, J.M., Parker, M.W., Crest, M., Chartier, M. and Baty, D. (1996) 'Membrane topology of the Colicin A pore-forming domain analyzed by disulfide bond engineering', *J Biol Chem*, 271(26), pp. 15401-6.

Duche, D., Parker, M.W., Gonzalezmanas, J.M., Pattus, F. and Baty, D. (1994) 'Uncoupled steps of the Colicin A pore formation demonstrated by disulfide bond engineering', *J Biol Chem*, 269(9), pp. 6332-6339.

Dumon-Seignovert, L., Cariot, G. and Vuillard, L. (2004) 'The toxicity of recombinant proteins in *Escherichia coli*: A comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3)', *Protein Expres Purif*, 37(1), pp. 203-206.

El Ghachi, M., Bouhss, A., Barreteau, H., Touze, T., Auger, G., Blanot, D. and Mengin-Lecreulx, D. (2006) 'Colicin M exerts its bacteriolytic effect via enzymatic degradation of undecaprenyl phosphate-linked peptidoglycan precursors', *J Biol Chem*, 281(32), pp. 22761-22772.

Elkins, P., Bunker, A., Cramer, W.A. and Stauffacher, C.V. (1997) 'A mechanism for toxin insertion into membranes is suggested by the crystal structure of the channel-forming domain of colicin E1', *Structure*, 5(3), pp. 443-458.

Erukova, V., Sobko, A., Rokitskaya, T., Kotova, E., Antonenko, Y., Zakharov, S. and Cramer, W. (2004) 'Simultaneous electrical and optical measurements of colicin E1 ion channels', *Biophys J*, 86(1), pp. 205a-205a.

Espesset, D., Corda, Y., Cunningham, K., Benedetti, H., Lloubes, R., Lazdunski, C. and Geli, V. (1994) 'The Colicin A pore-forming domain fused to mitochondrial intermembrane space sorting signals can be functionally inserted into the *Escherichia coli* plasma membrane by a mechanism that bypasses the Tol proteins', *Mol Microbiol*, 13(6), pp. 1121-31.

Espesset, D., Duche, D., Baty, D. and Geli, V. (1996) 'The channel domain of colicin A is inhibited by its immunity protein through direct interaction in the *Escherichia coli* inner membrane', *Embo J*, 15(10), pp. 2356-2364.

Evans, L.J.A., Goble, M.L., Hales, K.A. and Lakey, J.H. (1996a) 'Different sensitivities to acid denaturation within a family of proteins: Implications for acid unfolding and membrane translocation', *Biochemistry-US*, 35(40), pp. 13180-13185.

Evans, L.J.A., Labeit, S., Cooper, A., Bond, L.H. and Lakey, J.H. (1996b) 'The central domain of colicin N possesses the receptor recognition site but not the binding affinity of the whole toxin', *Biochemistry-US*, 35(48), pp. 15143-15148.

Farrance, O.E., Hann, E., Kaminska, R., Housden, N.G., Derrington, S.R., Kleanthous, C., Radford, S.E. and Brockwell, D.J. (2013) 'A force-activated trip switch triggers rapid dissociation of a colicin from its immunity protein', *Plos Biol*, 11(2).

Firer-Sherwood, M.A., Ando, N., Drennan, C.L. and Elliott, S.J. (2011) 'Solution-based structural analysis of the decaheme cytochrome, MtrA, by small-angle X-ray scattering and analytical ultracentrifugation', *J Phys Chem B*, 115(38), pp. 11208-14.

Fridd, S.L., Gokce, I. and Lakey, J.H. (2002) 'High level expression of His-tagged colicin pore-forming domains and reflections on the sites for pore formation in the inner membrane', *Biochimie*, 84(5-6), pp. 477-483.

Geli, V., Baty, D. and Lazdunski, C. (1988) 'Use of a foreign epitope as a tag for the localization of minor proteins within a cell - the case of the immunity protein to Colicin A', *P Natl Acad Sci USA*, 85(3), pp. 689-693.

Geli, V., Baty, D., Pattus, F. and Lazdunski, C. (1989a) 'Topology and function of the integral membrane protein conferring immunity to Colicin A', *Mol Microbiol*, 3(5), pp. 679-87.

Geli, V., Knibiehler, M., Bernadac, A. and Lazdunski, C. (1989b) 'Purification and reconstitution into liposomes of an integral membrane-protein conferring immunity to Colicin A', *Fems Microbiol Lett*, 60(2), pp. 239-244.

Geli, V. and Lazdunski, C. (1992) 'An alpha-helical hydrophobic hairpin as a specific determinant in protein-protein interaction occurring in *Escherichia coli* Colicin A and B immunity systems', *J Bacteriol*, 174(20), pp. 6432-7.

Gerard, F., Brooks, M.A., Barreteau, H., Touze, T., Graille, M., Bouhss, A., Blanot, D., van Tilbeurgh, H. and Mengin-Lecreulx, D. (2011) 'X-ray structure and site-directed mutagenesis analysis of the *Escherichia coli* Colicin M Immunity Protein', *J Bacteriol*, 193(1), pp. 205-214.

Gerding, M.A., Ogata, Y., Pecora, N.D., Niki, H. and de Boer, P.A. (2007) 'The trans-envelope Tol-Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in *E. coli*', *Mol Microbiol*, 63(4), pp. 1008-25.

Giffard, C.J., Li, W., Moore, G.R., James, R. and Kleanthous, C. (1997) 'Protein-protein interactions in the colicin E2 DNase-immunity complex. Probing the importance of conserved residues.', *Faseb J*, 11(9), pp. A833-A833.

Gilbert, R.J., Heenan, R.K., Timmins, P.A., Gingles, N.A., Mitchell, T.J., Rowe, A.J., Rossjohn, J., Parker, M.W., Andrew, P.W. and Byron, O. (1999) 'Studies on the structure and mechanism of a bacterial protein toxin by analytical ultracentrifugation and small-angle neutron scattering', *J Mol Biol*, 293(5), pp. 1145-60.

Gokce, I. and Lakey, J.H. (2003) 'Production of an *E. coli* toxin protein; Colicin A in *E. coli* using an inducible system', *Turk J Chem*, 27(3), pp. 323-331.

Gokce, I., Raggett, E.M., Hong, Q., Virden, R., Cooper, A. and Lakey, J.H. (2000) 'The TolA-recognition site of Colicin N. ITC, SPR and stopped-flow fluorescence define a crucial 27-residue segment', *J Mol Biol*, 304(4), pp. 621-632.

Goldman, K., Suit, J.L. and Kayalar, C. (1985) 'Identification of the plasmid-encoded immunity protein for colicin E1 in the inner membrane of *Escherichia coli*', *FEBS Lett*, 190(2), pp. 319-23.

Gordon, D.M. and O'Brien, C.L. (2006) 'Bacteriocin diversity and the frequency of multiple bacteriocin production in *Escherichia coli*', *Microbiology*, 152(Pt 11), pp. 3239-44.

Gordon, D.M., Riley, M.A. and Pinou, T. (1998) 'Temporal changes in the frequency of colicinogeny in *Escherichia coli* from house mice', *Microbiology-Uk*, 144, pp. 2233-2240.

Graille, M., Mora, L., Buckingham, R.H., van Tilbeurgh, H. and de Zamaroczy, M. (2004) 'Structural inhibition of the Colicin D tRNase by the tRNA-mimicking immunity protein', *Embo J*, 23(7), pp. 1474-1482.

Griko, Y.V., Zakharov, S.D. and Cramer, W.A. (2001) 'Structural stability and domain organization of Colicin E1', *Biophys J*, 80(1), pp. 135a-135a.

Guihard, G., Benedetti, H., Besnard, M. and Letellier, L. (1993) 'Phosphate efflux through the channels formed by colicins and phage T5 in *Escherichia coli* cells is responsible for the fall in cytoplasmic ATP', *J Biol Chem*, 268(24), pp. 17775-80.

Guinane, C.M., Cotter, P.D., Hill, C. and Ross, R.P. (2005) 'Microbial solutions to microbial problems; lactococcal bacteriocins for the control of undesirable biota in food', *J Appl Microbiol*, 98(6), pp. 1316-1325.

Guzman, L.M., Belin, D., Carson, M.J. and Beckwith, J. (1995) 'Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter', *J Bacteriol*, 177(14), pp. 4121-30.

Hammon, J., Palanivelu, D.V., Chen, J., Patel, C. and Minor, D.L., Jr. (2009) 'A green fluorescent protein screen for identification of well-expressed membrane proteins from a cohort of extremophilic organisms', *Protein Sci*, 18(1), pp. 121-33.

Hansen, S.B., Tao, X. and MacKinnon, R. (2011) 'Structural basis of PIP2 activation of the classical inward rectifier K+ channel Kir2.2', *Nature*, 477(7365), pp. 495-8.

Hategan, A., Gersh, K.C., Safer, D. and Weisel, J.W. (2013) 'Visualization of the dynamics of fibrin clot growth 1 molecule at a time by total internal reflection fluorescence microscopy', *Blood*, 121(8), pp. 1455-1458.

Hattori, M., Hibbs, R.E. and Gouaux, E. (2012) 'A fluorescence-detection size-exclusion chromatography-based thermostability assay for membrane protein precrystallization screening', *Structure*, 20(8), pp. 1293-1299.

Hausmann, C. and Clowes, R.C. (1971) 'Mitomycin C and temperature induction of Colicin B in the absence of deoxyribonucleic acid synthesis', *J Bacteriol*, 107(3), pp. 633-5.

Hecht, O., Zhang, Y., Li, C., Penfold, C.N., James, R. and Moore, G.R. (2010) 'Characterisation of the interaction of Colicin A with its co-receptor TolA', *FEBS Lett*, 584(11), pp. 2249-2252.

Hedhammar, M., Jung, H.R. and Hober, S. (2006) 'Enzymatic cleavage of fusion proteins using immobilised protease 3C', *Protein Expres Purif*, 47(2), pp. 422-426.

Helenius, A. and Simons, K. (1975) 'Solubilization of membranes by detergents', *Biochim Biophys Acta*, 415(1), pp. 29-79.

Heng, N.C., Tagg, J.R. and Tompkins, G.R. (2007) 'Competence-dependent bacteriocin production by *Streptococcus gordonii* DL1 (Challis)', *J Bacteriol*, 189(4), pp. 1468-72.

Hilsenbeck, J.L., Park, H., Chen, G., Youn, B., Postle, K. and Kang, C.H. (2004) 'Crystal structure of the cytotoxic bacterial protein colicin B at 2.5 angstrom resolution', *Mol Microbiol*, 51(3), pp. 711-720.

Ho, D., Lugo, M.R., Lomize, A.L., Pogozheva, I.D., Singh, S.P., Schwan, A.L. and Merrill, A.R. (2011) 'Membrane topology of the Colicin E1 channel using genetically encoded fluorescence', *Biochemistry-US*, 50(22), pp. 4830-4842.

Ho, D., Lugo, M.R. and Merrill, A.R. (2013) 'Harmonic analysis of the fluorescence response of bimane adducts of Colicin E1 at helices 6, 7, and 10', *J Biol Chem*, 288(7), pp. 5136-5148.

Ho, D. and Merrill, A.R. (2009) 'Evidence for the amphipathic nature and tilted topology of helices 4 and 5 in the closed state of the Colicin E1 channel', *Biochemistry-US*, 48(6), pp. 1369-1380.

Ho, D. and Merrill, R. (2011) 'Characterization of the Colicin E1 channel using genetically Encoded Fluorescence', *Biophys J*, 100(3), pp. 208-208.

Holt, A., Koehorst, R.B., Rutters-Meijneke, T., Gelb, M.H., Rijkers, D.T., Hemminga, M.A. and Killian, J.A. (2009) 'Tilt and rotation angles of a transmembrane model peptide as studied by fluorescence spectroscopy', *Biophys J*, 97(8), pp. 2258-66.

Honigmann, A., Pulagam, L.P., Sippach, M., Bartsch, P., Steinhoff, H.J. and Wagner, R. (2012) 'A high resolution electro-optical approach for investigating transition of soluble proteins to integral membrane proteins probed by Colicin A', *Biochem Bioph Res Co*, 427(2), pp. 385-391.

Housden, N.G. and Kleanthous, C. (2012) 'Colicin translocation across the *Escherichia coli* outer membrane', *Biochem Soc T*, 40, pp. 1475-1479.

Jakes, K.S. (2014) 'Daring to be different:colicin N finds another way', *Mol Microbiol*, 92(3), pp. 435-439.

James, R., Schneider, J. and Cooper, P.C. (1987) 'Characterization of 3 group-a klebicin plasmids - localization of their E-Colicin immunity genes', *J Gen Microbiol*, 133, pp. 2253-2262.

Jin, L., Lee, R.K., Kong, S.K., Yuan, W., Ho, H.P. and Lin, C.L. (2008) 'Applications of total internal reflection fluorescence (TIRF) microscopy in cellular bio-imaging', *Aoe 2007: Asia Optical Fiber Commun & Optoelectronic Exposition & Conf, Conf Proc*, pp. 587-589.

Joerger, R.D. (2003) 'Alternatives to antibiotics: Bacteriocins, antimicrobial peptides and bacteriophages', *Poultry Sci*, 82(4), pp. 640-647.

Johnson, C.L., Ridley, H., Marchetti, R., Silipo, A., Griffin, D.C., Crawford, L., Bonev, B., Molinaro, A. and Lakey, J.H. (2014) 'The antibacterial toxin Colicin N binds to the inner core of lipopolysaccharide and close to its translocator protein', *Mol Microbiol*, 92(3), pp. 440-52.

Johnson, C.L., Ridley, H., Pengelly, R.J., Salleh, M.Z. and Lakey, J.H. (2013) 'The unstructured domain of Colicin N kills *Escherichia coli*', *Mol Microbiol*.

Kasianowicz, J., Benz, R. and Mclaughlin, S. (1984) 'The kinetic mechanism by which CCCP (Carbonyl Cyanide Meta-Chlorophenylhydrazone) transports protons across membranes', *J Membrane Biol*, 82(2), pp. 179-190.

Kemperman, R., Jonker, M., Nauta, A., Kuipers, O.P. and Kok, J. (2003) 'Functional analysis of the gene cluster involved in production of the bacteriocin Circularin A by *Clostridium beijerinckii* ATCC 25752', *Appl Environ Microb*, 69(10), pp. 5839-5848.

Kerr, B., Riley, M.A., Feldman, M.W. and Bohannan, B.J.M. (2002) 'Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors', *Nature*, 418(6894), pp. 171-174.

Kienker, P.K., Qiu, X.Q., Slatin, S.L., Finkelstein, A. and Jakes, K.S. (1997) 'Transmembrane insertion of the Colicin Ia hydrophobic hairpin', *J Membrane Biol*, 157(1), pp. 27-37.

Kim, T. and Im, W. (2010) 'Revisiting hydrophobic mismatch with free energy simulation studies of transmembrane helix tilt and rotation', *Biophys J*, 99(1), pp. 175-183.

Kim, Y.C., Tarr, A.W. and Penfold, C.N. (2014) 'Colicin import into *E. coli* cells: A model system for insights into the import mechanisms of bacteriocins', *BBA-Mol Cell Res*, 1843(8), pp. 1717-1731.

Kleanthous, C. (2010) 'Swimming against the tide: progress and challenges in our understanding of colicin translocation', *Nat Rev Microbiol*, 8(12), pp. 843-848.

Kleanthous, C. and Walker, D. (2001) 'Immunity proteins: enzyme inhibitors that avoid the active site', T Biochem Sci, 26(10), pp. 624-631.

Kleanthous, C., Kuhlmann, U.C., Pommer, A.J., Ferguson, N., Radford, S.E., Moore, G.R., James, R. and Hemmings, A.M. (1999) 'Structural and mechanistic basis of immunity toward endonuclease colicins', *Nat Struct Biol*, 6(3), pp. 243-252.

Kleinschmidt, J.H. and Popot, J.L. (2014) 'Folding and stability of integral membrane proteins in amphipols', *Arch Biochem Biophys*, 564, pp. 327-343.

Knowles, T.J., Finka, R., Smith, C., Lin, Y.P., Dafforn, T. and Overduin, M. (2009) 'Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer', *J Am Chem Soc*, 131(22), pp. 7484-5.

Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E.L. (2001) 'Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes', *J Mol Biol*, 305(3), pp. 567-80.

Kuhlmann, U.C., Pommer, A.J., Moore, G.R., James, R. and Kleanthous, C. (2000) 'Specificity in protein-protein interactions: The structural basis for dual recognition in endonuclease colicin-immunity protein complexes', *J Mol Biol*, 301(5), pp. 1163-1178.

Kunji, E.R., Harding, M., Butler, P.J. and Akamine, P. (2008) 'Determination of the molecular mass and dimensions of membrane proteins by size exclusion chromatography', *Methods*, 46(2), pp. 62-72.

Lakey, J.H., Baty, D. and Pattus, F. (1991) 'Fluorescence energy-transfer distance measurements using site-directed single cysteine mutants - the membrane insertion of Colicin A', *J Mol Biol*, 218(3), pp. 639-653.

Lakey, J.H., Duche, D., Gonzalezmanas, J.M., Baty, D. and Pattus, F. (1993) 'Fluorescence energy-transfer distance measurements - the hydrophobic helical hairpin of Colicin A in the membrane-bound state', *J Mol Biol*, 230(3), pp. 1055-1067.

Lakey, J.H. and Slatin, S.L. (2001) 'Pore-forming colicins and their relatives', *Curr Top Microbiol*, 257, pp. 131-161.

Lazzaroni, J.C., Fogninilefebvre, N. and Portalier, R.C. (1986) 'Effects of Lkyb-mutations on the expression of OmpF, OmpC and lamb porin structural genes in *Escherichia coli* K-12', *Fems Microbiol Lett*, 33(2-3), pp. 235-239.

Lei, G.S., Syu, W.J., Liang, P.H., Chak, K.F., Hu, W.S. and Hu, S.T. (2011) 'Repression of BtuB gene transcription in *Escherichia coli* by the GadX protein', *Bmc Microbiology*, 11.

Lemmon, M.A. and Engelman, D.M. (1994) 'Specificity and promiscuity in membrane helix interactions', *FEBS Lett*, 346(1), pp. 17-20.

Levengood, S.K., Beyer, W.F. and Webster, R.E. (1991) 'TolA - a membrane-protein involved in colicin uptake contains an extended helical region', *Proc Natl Acad Sci USA*, 88(14), pp. 5939-5943.

Li, D.F. and Caffrey, M. (2011) 'Lipid cubic phase as a membrane mimetic for integral membrane protein enzymes', *P Natl Acad Sci USA*, 108(21), pp. 8639-8644.

Li, D.F. and Caffrey, M. (2014) 'Renaturing membrane proteins in the lipid cubic phase, a nanoporous membrane mimetic', *Sci Rep*, 4.

Li, L., Tian, X.Z., Zou, G.Z., Shi, Z.K., Zhang, X.L. and Jin, W.R. (2008) 'Quantitative counting of single fluorescent molecules by combined electrochemical adsorption accumulation and total internal reflection fluorescence microscopy', *Anal Chem*, 80(11), pp. 3999-4006.

Lichenstein, H.S., Hamilton, E.P. and Lee, N. (1987) 'Repression and catabolite gene activation in the araBAD operon', *J Bacteriol*, 169(2), pp. 811-22.

Lindeberg, M. and Cramer, W.A. (2001) 'Identification of specific residues in Colicin E1 involved in immunity protein recognition', *J Bacteriol*, 183(6), pp. 2132-6.

Lindeberg, M., Zakharov, S.D. and Cramer, W.A. (2000) 'Unfolding pathway of the Colicin E1 channel protein on a membrane surface', *Biophys J*, 78(1), pp. 176a-176a.

Lloubes, R., Baty, D. and Lazdunski, C. (1986) 'The promoters of the genes for colicin production, release and immunity in the ColA plasmid: effects of convergent transcription and Lex A protein', *Nucleic Acids Res*, 14(6), pp. 2621-36.

Lloubes, R., Cascales, E., Walburger, A., Bouveret, E., Lazdunski, C., Bernadac, A. and Journet, L. (2001) 'The Tol-Pal proteins of the *Escherichia coli* cell envelope: an energized system required for outer membrane integrity?', *Res Microbiol*, 152(6), pp. 523-529.

Lloubes, R., Goemaere, E., Zhang, X., Cascales, E. and Duche, D. (2012) 'Energetics of colicin import revealed by genetic cross-complementation between the Tol and Ton systems', *Biochem Soc T*, 40, pp. 1480-1485.

Loll, P.J. (2014) 'Membrane proteins, detergents and crystals: what is the state of the art?', *Acta Crystallogr F Struct Biol Commun*, 70(Pt 12), pp. 1576-83.

Los, G.V., Encell, L.P., McDougall, M.G., Hartzell, D.D., Karassina, N., Zimprich, C., Wood, M.G., Learish, R., Ohana, R.F., Urh, M., Simpson, D., Mendez, J., Zimmerman, K., Otto, P., Vidugiris, G., Zhu, J., Darzins, A., Klaubert, D.H., Bulleit, R.F. and Wood, K.V. (2008) 'HaloTag: a novel protein labeling technology for cell imaging and protein analysis', *ACS Chem Biol*, 3(6), pp. 373-82.

Luna-Chavez, C., Lin, Y.L. and Huang, R.H. (2006) 'Molecular basis of inhibition of the ribonuclease activity in colicin E5 by its cognate immunity protein', J Mol Biol, 358(2), pp. 571-9.

Lundblad, R.L. and Noyes, C.M. (1984) 'Chemical reagents for protein modification', in Boca Raton, Florida, USA: CRC Press, Inc. , pp. 105 -125.

Lundrigan, M.D., De Veaux, L.C., Mann, B.J. and Kadner, R.J. (1987) 'Separate regulatory systems for the repression of MetE and BtuB by vitamin B12 in *Escherichia coli*', *Mol Gen Genet*, 206(3), pp. 401-7.

Mackay, J.P., Sunde, M., Lowry, J.A., Crossley, M. and Matthews, J.M. (2007) 'Protein interactions: is seeing believing?', *Trends Biochem Sci*, 32(12), pp. 530-531.

Mankovich, J.A., Hsu, C.H. and Konisky, J. (1986) 'DNA and amino acid sequence analysis of structural and immunity genes of Colicins Ia and Ib', *J Bacteriol*, 168(1), pp. 228-36.

Mariani, S. and Minunni, M. (2014) 'Surface plasmon resonance applications in clinical analysis', *Anal Bioanal Chem*, 406(9-10), pp. 2303-23.

Martirani, L., Varcamonti, M., Naclerio, G. and De Felice, M. (2002) 'Purification and partial characterization of bacillocin 490, a novel bacteriocin produced by a thermophilic strain of *Bacillus licheniformis*', *Microb Cell Fact*, 1.

Massotte, D., Yamamoto, M., Scianimanico, S., Sorokine, O., van Dorsselaer, A., Nakatani, Y., Ourisson, G. and Pattus, F. (1993) 'Structure of the membrane-bound form of the pore-forming domain of Colicin A: a partial proteolysis and mass spectrometry study', *Biochemistry-US*, 32(50), pp. 13787-94.

Mathur, H., Rea, M.C., Cotter, P.D., Hill, C. and Ross, R.P. (2015) 'The sactibiotic subclass of bacteriocins: an update', *Current Protein & Peptide Science*, 16(6), pp. 549-558.

Michel-Briand, Y. and Baysse, C. (2002) 'The pyocins of *Pseudomonas aeruginosa*', *Biochimie*, 84(5-6), pp. 499-510.

Miroux, B. and Walker, J.E. (1996) 'Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels', *J Mol Biol*, 260(3), pp. 289-98.

Muchmore, S.W., Sattler, M., Liang, H., Meadows, R.P., Harlan, J.E., Yoon, H.S., Nettesheim, D., Chang, B.S., Thompson, C.B., Wong, S.L., Ng, S.L. and Fesik, S.W. (1996) 'X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death', *Nature*, 381(6580), pp. 335-41.

Mulec, J., Podlesek, Z., Mrak, P., Kopitar, A., Ihan, A. and Zgur-Bertok, D. (2003) 'A cka-gfp transcriptional fusion reveals that the Colicin K activity gene is induced in only 3 percent of the population', *J Bacteriol*, 185(2), pp. 654-659.

Nakazawa, A. and Tamada, T. (1972) 'Stimulation of Colicin E1 synthesis by cyclic 3',5'-adenosine monophosphate in Mitomycin C-induced *Escherichia coli*', *Biochem Bioph Res Co*, 46(2), pp. 1004-&.

Nardi, A., Corda, Y., Baty, D. and Duche, D. (2001a) 'Colicin A immunity protein interacts with the hydrophobic helical hairpin of the Colicin A channel domain in the *Escherichia coli* inner membrane', *J Bacteriol*, 183(22), pp. 6721-5.

Nardi, A., Slatin, S.L., Baty, D. and Duche, D. (2001b) 'The C-terminal half of the Colicin A pore-forming domain is active *in vivo* and *in vitro*', *J Mol Biol*, 307(5), pp. 1293-1303.

Nedelkina, S., Gokce, I., Ridley, H., Weckerle, C., Magnin, T., Vallette, F., Pattus, F., Lakey, J.H. and Bechinger, B. (2008) 'High-yield expression and purification of soluble forms of the anti-apoptotic Bcl-x(L) and Bcl-2 as TolAIII-fusion proteins', *Protein Expr Purif*, 60(2), pp. 214-20.

Nes, I.F., Diep, D.B. and Holo, H. (2007) 'Bacteriocin diversity in Streptococcus and Enterococcus', *J Bacteriol*, 189(4), pp. 1189-98.

Nyholm, T.K.M., Ozdirekcan, S. and Killian, J.A. (2007) 'How protein transmembrane segments sense the lipid environment', *Biochemistry-US*, 46(6), pp. 1457-1465.

Olschlager, T. and Braun, V. (1987a) 'Sequence, expression, and localization of the immunity protein for Colicin M', *J Bacteriol*, 169(10), pp. 4765-4769.

Olschlager, T. and Braun, V. (1987b) 'Sequence, expression, and localization of the immunity protein for Colicin M', *J Bacteriol*, 169(10), pp. 4765-9.

Osborne, M.J., Breeze, A.L., Lian, L.Y., Reilly, A., James, R., Kleanthous, C. and Moore, G.R. (1996) 'Three-dimensional solution structure and C-13 nuclear magnetic resonance assignments of the Colicin E9 immunity protein Im9', *Biochemistry-US*, 35(29), pp. 9505-9512.

Overduin, M. (2014) 'MODA, HADDOCK, and SMALP: new developments for unravelling membrane protein structure and function: Honorary Lecture', *Biochem Cell Biol*, 92(6), pp. 603-604.

Padmavathi, P.V.L. and Steinhoff, H.J. (2008) 'Conformation of the closed channel state of Colicin A in proteoliposomes: An umbrella model', *J Mol Biol*, 378(1), pp. 204-214.

Papadakos, G., Wojdyla, J.A. and Kleanthous, C. (2011) 'Nuclease colicins and their immunity proteins', *Q Rev Biophys*, pp. 1-47.

Papanastasiou, M., Orfanoudaki, G., Koukaki, M., Kountourakis, N., Sardis, M.F., Aivaliotis, M., Karamanou, S. and Economou, A. (2013) 'The *Escherichia coli* peripheral inner membrane proteome', *Mol Cell Proteomics*, 12(3), pp. 599-610.

Parker, M.W., Pattus, F., Tucker, A.D. and Tsernoglou, D. (1989) 'Structure of the membrane-pore-forming fragment of Colicin A', *Nature*, 337(6202), pp. 93-96.

Parker, M.W., Postma, J.P.M., Pattus, F., Tucker, A.D. and Tsernoglou, D. (1992) 'Refined structure of the pore-forming domain of Colicin A at 2-Bullet-4 Angstrom resolution', *J Mol Biol*, 224(3), pp. 639-657.

Patterson, G.H., Knobel, S.M., Sharif, W.D., Kain, S.R. and Piston, D.W. (1997) 'Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy', *Biophys J*, 73(5), pp. 2782-2790.

Paulin, S., Jamshad, M., Dafforn, T.R., Garcia-Lara, J., Foster, S.J., Galley, N.F., Roper, D.I., Rosado, H. and Taylor, P.W. (2014) 'Surfactant-free purification of membrane protein complexes from bacteria: application to the staphylococcal penicillin-binding protein complex PBP2/PBP2a', *Nanotechnology*, 25(28).

Pedelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T.C. and Waldo, G.S. (2006) 'Engineering and characterization of a superfolder green fluorescent protein', *Nat Biotech*, 24(1), pp. 79-88.

Penfold, C.N., Li, C., Zhang, Y., Vankemmelbeke, M. and James, R. (2012) 'Colicin A binds to a novel binding site of TolA in the *Escherichia coli* periplasm', *Biochem Soc T*, 40, pp. 1469-1474.

Pilsl, H. and Braun, V. (1995) 'Evidence that the immunity protein inactivates Colicin 5 immediately prior to the formation of the transmembrane channel', *J Bacteriol*, 177(23), pp. 6966-6972.

Pilsl, H., Smajs, D. and Braun, V. (1998) 'The tip of the hydrophobic hairpin of Colicin U is dispensable for Colicin U activity but is important for interaction with the immunity protein', *J Bacteriol*, 180(16), pp. 4111-5.

Popot, J.L. (2010) 'Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions', *Annua Rev Biochem, Vol 79*, 79, pp. 737-775.

Postis, V., Rawson, S., Mitchell, J.K., Lee, S.C., Parslow, R.A., Dafforn, T.R., Baldwin, S.A. and Muench, S.P. (2015) 'The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy', *BBA-Biomembranes*, 1848(2), pp. 496-501.

Prieto, L. and Lazaridis, T. (2011) 'Computational studies of colicin insertion into membranes: the closed state', *Proteins-Structure Function and Bioinformatics*, 79(1), pp. 126-141.

Pugsley, A.P. (1984a) 'Genetic analysis of ColN plasmid determinants for colicin production, release, and immunity', *J Bacteriol*, 158(2), pp. 523-529.

Pugsley, A.P. (1984b) 'The ins and outs of colicins. Part I: Production, and translocation across membranes', *Microbiol Sci*, 1(7), pp. 168-75.

Pugsley, A.P. (1984c) 'The ins and outs of colicins. Part II. Lethal action, immunity and ecological implications', *Microbiol Sci*, 1(8), pp. 203-5.

Pugsley, A.P. (1987a) 'Nucleotide sequencing of the structural gene for Colicin N reveals homology between the catalytic, C-terminal domains of Colicin A and Colicin N', *Mol Microbiol*, 1(3), pp. 317-325.

Pugsley, A.P. (1987b) 'Taking toxins to pieces', *Microbiol Sci*, 4(10), p. 312.

Pugsley, A.P. (1988) 'The immunity and lysis genes of ColN plasmid pCHAP4', *Mol Gen Genet*, 211(2), pp. 335-41.

Qiu, X.Q., Jakes, K.S., Kienker, P.K., Finkelstein, A. and Slatin, S.L. (1996) 'Major transmembrane movement associated with Colicin Ia channel gating', *J Gen Physiol*, 107(3), pp. 313-328.

Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G. and Deber, C.M. (2009) 'Detergent binding explains anomalous SDS-PAGE migration of membrane proteins', *P Natl Acad Sci USA*, 106(6), pp. 1760-1765.

Ravnum, S. and Andersson, D.I. (1997) 'Vitamin B12 repression of the BtuB gene in *Salmonella typhimurium* is mediated via a translational control which requires leader and coding sequences', *Mol Microbiol*, 23(1), pp. 35-42.

Rebuffat, S. (2011) *Bacteriocins from Gram-negative bacteria: a classification?* New York, USA: Springer.

Ridley, H. and Lakey, J.H. (2015) 'Antibacterial toxin Colicin N and phage protein G3p compete with TolB for a binding site on TolA', *Microbiology*, 161(Pt 3), pp. 503-15.

Riley, M.A. (1993) 'Positive selection for colicin diversity in bacteria', *Mol Biol Evol*, 10(5), pp. 1048-1059.

Riley, M.A. (1998) 'Molecular mechanisms of bacteriocin evolution', *Annu Rev Genet*, 32, pp. 255-78.

Riley, M.A., Robinson, S.M., Roy, C.M., Dennis, M., Liu, V. and Dorit, R.L. (2012) 'Resistance is futile: the bacteriocin model for addressing the antibiotic resistance challenge', *Biochem Soc Trans*, 40(6), pp. 1438-42.

Riley, M.A. and Wertz, J.E. (2002a) 'Bacteriocin diversity: ecological and evolutionary perspectives', *Biochimie*, 84(5-6), pp. 357-64.

Riley, M.A. and Wertz, J.E. (2002b) 'Bacteriocins: evolution, ecology, and application', *Annul Rev Microbiol*, 56, pp. 117-137.

Rodger, A., Rajendra, J., Mortimer, R., Andrews, T., Hirst, J.D., Gilbert, A.T.B., Marrington, R., Dafforn, T.R., Halsall, D.J., Ardhammar, M., Norden, B., Woolhead, C.A., Robinson, C., Pinheiro, T.J.T., Kazlauskaite, J., Seymour, M., Perez, N. and Hannon, M.J. (2002) 'Flow oriented linear dichroism to probe protein orientation in membrane environments', *Roy Soc Ch*, (283), pp. 3-19.

Rolfe, D.J., McLachlan, C.I., Hirsch, M., Needham, S.R., Tynan, C.J., Webb, S.E., Martin-Fernandez, M.L. and Hobson, M.P. (2011) 'Automated multidimensional single

molecule fluorescence microscopy feature detection and tracking', *Eur Biophys J*, 40(10), pp. 1167-86.

Sagne, C., Isambert, M.F., Henry, J.P. and Gasnier, B. (1996) 'SDS-resistant aggregation of membrane proteins: application to the purification of the vesicular monoamine transporter', *Biochem J*, 316 (Pt 3), pp. 825-31.

Salamon, Z., Macleod, H.A. and Tollin, G. (1997a) 'Surface plasmon resonance spectroscopy as a tool for investigating the biochemical and biophysical properties of membrane protein systems. I: theoretical principles', *Biochim Biophys Acta*, 1331(2), pp. 117-29.

Salamon, Z., Macleod, H.A. and Tollin, G. (1997b) 'Surface plasmon resonance spectroscopy as a tool for investigating the biochemical and biophysical properties of membrane protein systems. II: applications to biological systems', *Biochim Biophys Acta*, 1331(2), pp. 131-52.

Schein, S.J., Kagan, B.L. and Finkelstein, A. (1978) 'Colicin-K acts by forming voltage-dependent channels in phospholipid bilayer membranes', *Nature*, 276(5684), pp. 159-163.

Schendel, S.L. and Cramer, W.A. (1994) 'On the nature of the unfolded intermediate in the *in vitro* transition of the Colicin E1 channel domain from the aqueous to the membrane phase', *Protein Sci*, 3(12), pp. 2272-2279.

Schleif, R. (2000) 'Regulation of the L-arabinose operon of *Escherichia coli*', *Trends Genet*, 16(12), pp. 559-65.

Schleif, R. (2010) 'AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action', *FEMS Microbiol Rev*, 34(5), pp. 779-96.

Schulz, S., Stephan, A., Hahn, S., Bortesi, L., Jarczowski, F., Bettmann, U., Paschke, A.K., Tuse, D., Stahl, C.H., Giritch, A. and Gleba, Y. (2015) 'Broad and efficient control of major foodborne pathogenic strains of *Escherichia coli* by mixtures of plant-produced colicins', *Proc Natl Acad Sci U S A*.

Shawkat, S., Karima, R., Tojo, T., Tadakuma, H., Saitoh, S.I., Akashi-Takamura, S., Miyake, K., Funatsu, T. and Matsushima, K. (2008) 'Visualization of the molecular

dynamics of lipopolysaccharide on the plasma membrane of murine macrophages by total internal reflection fluorescence microscopy', *J Biol Chem*, 283(34), pp. 22962-22971.

Shirabe, K., Yamada, M., Merrill, A.R., Cramer, W.A. and Nakazawa, A. (1993) 'Overproduction and purification of the Colicin E1 immunity protein', *Plasmid*, 29(3), pp. 236-240.

Siegel, D.P., Cherezov, V., Greathouse, D.V., Koeppe, R.E., Killian, J.A. and Caffrey, M. (2006) 'Transmembrane peptides stabilize inverted cubic phases in a biphasic length-dependent manner: Implications for protein-induced membrane fusion', *Biophys J*, 90(1), pp. 200-211.

Slade, K.M., Steele, B.L., Pielak, G.J. and Thompson, N.L. (2009) 'Quantifying green fluorescent protein diffusion in *Escherichia coli* by using continuous photobleaching with evanescent illumination', *J Phys Chem-US*, 113(14), pp. 4837-4845.

Slatin, S.L., Duche, D., Kienker, P.K. and Baty, D. (2004) 'Gating movements of Colicin A and colicin Ia are different', *J Membrane Biol*, 202(2), pp. 73-83.

Smajs, D., Dolezalova, M., Macek, P. and Zidek, L. (2008) 'Inactivation of Colicin Y by intramembrane helix-helix interaction with its immunity protein', *Febs Journal*, 275(21), pp. 5325-5331.

Smajs, D., Matejkova, P. and Weinstock, G.M. (2006) 'Recognition of pore-forming Colicin Y by its cognate immunity protein', *FEMS Microbiol Lett*, 258(1), pp. 108-13.

Smajs, D., Micenkova, L., Smarda, J., Vrba, M., Sevcikova, A., Valisova, Z. and Woznicova, V. (2010) 'Bacteriocin synthesis in uropathogenic and commensal *Escherichia coli*: Colicin E1 is a potential virulence factor', *Bmc Microbiology*, 10.

Sobko, A.A., Antonenko, Y.N., Zakharov, S.D. and Cramer, W.A. (2005) ''Giant' Colicin E1 channels', *Biophys J*, 88(1), pp. 249a-249a.

Sobko, A.A., Kotova, E.A., Antonenko, Y.N., Zakharov, S.D. and Cramer, W.A. (2004a) 'Effect of lipids with different spontaneous curvature on the channel activity of colicin E1: evidence in favor of a toroidal pore', *FEBS Lett*, 576(1-2), pp. 205-210.

Sobko, A.A., Kotova, E.A., Antonenko, Y.N., Zakharov, S.D. and Cramer, W.A. (2006a) 'Lipid dependence of the channel properties of a colicin E1-lipid toroidal pore', *J Biol Chem*, 281(20), pp. 14408-14416.

Sobko, A.A., Kotova, E.A., Mueller, P., Antonenko, Y.N., Zakharov, S.D. and Cramer, W.A. (2006b) 'Colicin E1 ionic channels and lipid flip-flop', *BBA-Bioenergetics*, pp. 393-393.

Sobko, A.A., Kotova, E.A., Zakharov, S.D., Cramer, W.A. and Antonenko, Y.N. (2006c) 'Lipid-mediated inactivation of colicin E1 channels by calcium ions', *Biochemistry-Moscow*, 71(1), pp. 99-103.

Sobko, A.A., Kovalchuk, S.I., Kotova, E.A. and Antonenko, Y.N. (2010) 'Induction of lipid flip-flop by Colicin E1 - a hallmark of proteolipidic pore formation in liposome membranes', *Biochemistry-Moscow*, 75(6), pp. 728-733.

Sobko, A.A., Rokitskaya, T.I. and Kotova, E.A. (2009) 'Histidine 440 controls the opening of Colicin E1 channels in a lipid-dependent manner', *BBA-Biomembranes*, 1788(9), pp. 1962-1966.

Sobko, A.A., Vigasina, M.A., Rokitskaya, T.I., Kotova, E.A., Zakharov, S.D., Cramer, W.A. and Antonenko, Y.N. (2004b) 'Chemical and photochemical modification of Colicin E1 and gramicidin a in bilayer lipid membranes', *J Membrane Biol*, 199(1), pp. 51-62.

Soelaiman, S., Jakes, K., Wu, N., Li, C. and Shoham, M. (2001) 'Crystal structure of colicin E3: implications for cell entry and ribosome inactivation', *Mol Cell*, 8(5), pp. 1053-62.

Song, H.Y., Cohen, F.S. and Cramer, W.A. (1991) 'Membrane topography of ColE1 gene-products - the hydrophobic anchor of the Colicin E1 channel is a helical hairpin', *J Bacteriol*, 173(9), pp. 2927-2934.

Song, H.Y. and Cramer, W.A. (1991) 'Membrane topography of ColE1 gene-products - the immunity protein', *J Bacteriol*, 173(9), pp. 2935-2943.

Sonnhammer, E.L., von Heijne, G. and Krogh, A. (1998) 'A hidden Markov model for predicting transmembrane helices in protein sequences', *Proc Int Conf Intell Syst Mol Biol*, 6, pp. 175-82.

Soom, M., Schonherr, R., Kubo, Y., Kirsch, C., Klinger, R. and Heinemann, S.H. (2001) 'Multiple PIP2 binding sites in Kir2.1 inwardly rectifying potassium channels', *FEBS Lett*, 490(1-2), pp. 49-53.

Stangl, M. and Schneider, D. (2015) 'Functional competition within a membrane: Lipid recognition vs. transmembrane helix oligomerization', *Biochim Biophys Acta*, 1848(9), pp. 1886-96.

Stangl, M., Veerappan, A., Kroeger, A., Vogel, P. and Schneider, D. (2012) 'Detergent properties influence the stability of the glycophorin A transmembrane helix dimer in lysophosphatidylcholine micelles', *Biophys J*, 103(12), pp. 2455-64.

Stora, T., Lakey, J.H. and Vogel, H. (1999) 'Ion-channel gating in transmembrane receptor proteins: Functional activity in tethered lipid membranes', *Angew Chem Int Edit*, 38(3), pp. 389-392.

Strahl, H., Burmann, F. and Hamoen, L.W. (2014) 'The actin homologue MreB organizes the bacterial cell membrane', *Nat Commun*, 5, p. 3442.

Strahl, H. and Hamoen, L.W. (2010) 'Membrane potential is important for bacterial cell division', *Proc Natl Acad Sci U S A*, 107(27), pp. 12281-12286.

Suzuki, N., Hiraki, M., Yamada, Y., Matsugaki, N., Igarashi, N., Kato, R., Dikic, I., Drew, D., Iwata, S., Wakatsuki, S. and Kawasaki, M. (2010) 'Crystallization of small proteins assisted by green fluorescent protein', *Acta Crystallogr D Biol Crystallogr*, 66(Pt 10), pp. 1059-66.

Taylor, R.M., Zakharov, S.D., Heymann, J.B., Girvin, M.E. and Cramer, W.A. (2000) 'Folded state of the integral membrane Colicin E1 immunity protein in solvents of mixed polarity', *Biochemistry-US*, 39(40), pp. 12131-12139.

Thumm, G., Olschlager, T. and Braun, V. (1988) 'Plasmid pColBM-Cl139 does not encode a colicin lysis protein but contains sequences highly homologous to the D protein (resolvase) and the oriV region of the miniF plasmid', *Plasmid*, 20(1), pp. 75-82.

Tokuda, H. and Konisky, J. (1978a) '*In vitro* depolarization of *Escherichia coli* membrane vesicles by Colicin Ia', *J Biol Chem*, 253(21), pp. 7731-7.

Tokuda, H. and Konisky, J. (1978b) '*In vitro* depolarization of *Escherichia coli* membrane-vesicles by Colicin Ia', *J Biol Chem*, 253(21), pp. 7731-7737.

Trautner, B.W., Hull, R.A. and Darouiche, R.O. (2005) 'Colicins prevent colonization of urinary catheters', *J Antimicrob Chemoth*, 56(2), pp. 413-415.

Tsao, S.S. and Goebel, W.F. (1969) 'Colicin K. 8. The immunological properties of mitomycin-induced colicin K', *J Exp Med*, 130(6), pp. 1313-35.

Vandergoot, F.G., Gonzalezmanas, J.M., Lakey, J.H. and Pattus, F. (1991) 'A molten-globule membrane-insertion intermediate of the pore-forming domain of Colicin A', *Nature*, 354(6352), pp. 408-410.

Vankemmelbeke, M., Housden, N.G., James, R., Kleanthous, C. and Penfold, C.N. (2013) 'Immunity protein release from a cell-bound nuclease colicin complex requires global conformational rearrangement', *Microbiologyopen*, 2(5), pp. 853-61.

Vankemmelbeke, M., Zhang, Y., Moore, G.R., Kleanthous, C., Penfold, C.N. and James, R. (2009) 'Energy-dependent immunity protein release during tol-dependent nuclease colicin translocation', *J Biol Chem*, 284(28), pp. 18932-18941.

Vetter, I.R., Parker, M.W., Tuckr, A.D., Lakey, J.H., Pattus, F. and Tsernoglou, D. (1998) 'Crystal structure of a Colicin N fragment suggests a model for toxicity', *Struct Fold Des*, 6(7), pp. 863-874.

Visudtiphole, V., Thomas, M.B., Chalton, D. and Lakey, J.H. (2004) '*In vitro* refolding of trimeric and dimeric OmpF', *Biophys J*, 86(1), pp. 499a-499a.

Visudtiphole, V., Thomas, M.B., Chalton, D.A. and Lakey, J.H. (2005) 'Refolding of *Escherichia coli* outer membrane protein F in detergent creates LPS-free trimers and asymmetric dimers', *Biochem J*, 392, pp. 375-381.

Walburger, A., Lazdunski, C. and Corda, Y. (2002) 'The Tol/Pal system function requires an interaction between the C-terminal domain of TolA and the N-terminal domain of TolB', *Mol Microbiol*, 44(3), pp. 695-708.

Wallis, R., Reilly, A., Rowe, A., Moore, G.R., James, R. and Kleanthous, C. (1992) '*In vivo* and *in vitro* characterization of overproduced Colicin E9 immunity protein', *Eur J Biochem*, 207(2), pp. 687-695.

Warren, A.J., Armour, W., Axford, D., Basham, M., Connolley, T., Hall, D.R., Horrell, S., McAuley, K.E., Mykhaylyk, V., Wagner, A. and Evans, G. (2013) 'Visualization of membrane protein crystals in lipid cubic phase using X-ray imaging', *Acta Crystallogr D*, 69, pp. 1252-1259.

Watanabe, Y. (2002) 'Effect of various mild surfactants on the reassembly of an oligomeric integral membrane protein OmpF porin', *Journal of Protein Chemistry*, 21(3), pp. 169-175.

Watanabe, Y. and Inoko, Y. (2009) 'Reassembly of an integral oligomeric membrane protein OmpF porin in n-octyl beta-D-Glucopyranoside-lipids mixtures', *Protein J*, 28(2), pp. 66-73.

Weaver, C.A., Redborg, A.H. and Konisky, J. (1981) 'Plasmid-determined immunity of *Escherichia coli* K-12 to Colicin Ia is mediated by a plasmid-encoded membrane protein', *J Bacteriol*, 148(3), pp. 817-828.

Weierstall, U., James, D., Wang, C., White, T.A., Wang, D.J., Liu, W., Spence, J.C.H., Doak, R.B., Nelson, G., Fromme, P., Fromme, R., Grotjohann, I., Kupitz, C., Zatsepin, N.A., Liu, H.G., Basu, S., Wacker, D., Han, G.W., Katritch, V., Boutet, S., Messerschmidt, M., Williams, G.J., Koglin, J.E., Seibert, M.M., Klinker, M., Gati, C., Shoeman, R.L., Barty, A., Chapman, H.N., Kirian, R.A., Beyerlein, K.R., Stevens, R.C., Li, D.F., Shah, S.T.A., Howe, N., Caffrey, M. and Cherezov, V. (2014) 'Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography', *Nat Commun*, 5.

Weiss, M.J. and Luria, S.E. (1978) 'Reduction of membrane-potential, an immediate effect of Colicin K', *P Natl Acad Sci USA*, 75(5), pp. 2483-2487.

Weiss, T.M., van der Wel, P.C.A., Killian, J.A., Koeppe, R.E. and Huang, H.W. (2003) 'Hydrophobic mismatch between helices and lipid bilayers', *Biophys J*, 84(1), pp. 379-385.

Wiener, M.C., Freymann, D.M., Williams, P., Ghosh, P. and Stroud, R.M. (1997) 'The crystal structure of colicin Ia.', *Biophys J*, 72(2), pp. Wamf1-Wamf1.

Wilmsen, H.U., Pugsley, A.P. and Pattus, F. (1990) 'Colicin N forms voltage- and pH-dependent channels in planar lipid bilayer membranes', *Eur Biophys J*, 18(3), pp. 149-58.

Woraprayote, W., Pumpuang, L., Tosukhowong, A., Roytrakul, S., Perez, R.H., Zendo, T., Sonomoto, K., Benjakul, S. and Visessanguan, W. (2015) 'Two putatively novel bacteriocins active against Gram-negative food borne pathogens produced by Weissella hellenica BCC 7293', *Food Control*, 55, pp. 176-184.

Yang, L., Zhou, Y., Zhu, S., Huang, T., Wu, L. and Yan, X. (2012) 'Detection and quantification of bacterial autofluorescence at the single-cell level by a laboratory-built high-sensitivity flow cytometer', *Anal Chem*, 84(3), pp. 1526-32.

Yeagle, P.L. (2014) 'Non-covalent binding of membrane lipids to membrane proteins', *Biochim Biophys Acta*, 1838(6), pp. 1548-59.

Zakharov, S.D. and Cramer, W.A. (1999) 'Transition from the soluble to membrane-bound state of Colicin E1 channel domain: Helix elongation precedes membrane insertion', *Biophys J*, 76(1), pp. A8-A8.

Zakharov, S.D. and Cramer, W.A. (2002) 'Insertion intermediates of pore-forming colicins in membrane two-dimensional space', *Biochimie*, 84(5-6), pp. 465-475.

Zakharov, S.D., Lindeberg, M. and Cramer, W.A. (1998a) 'Kinetic phases in membrane binding-insertion of the colicin E1 channel domain.', *Biophys J*, 74(2), pp. A228-A228.

Zakharov, S.D., Lindeberg, M., Griko, Y., Salamon, Z., Tollin, G. and Cramer, W.A. (1998b) 'Conformational changes of the Colicin E1 channel domain at the membrane surface.', *Biophys J*, 74(2), pp. A228-A228.

Zakharov, S.D., Lindeberg, M., Griko, Y., Salamon, Z., Tollin, G., Prendergast, F.G. and Cramer, W.A. (1998c) 'Membrane-bound state of the Colicin E1 channel domain as an extended two-dimensional helical array', *P Natl Acad Sci USA*, 95(8), pp. 4282-4287.

Zeev-Ben-Mordehai, T., Vasishtan, D., Siebert, C.A. and Grunewald, K. (2014) 'The full-length cell-cell fusogen EFF-1 is monomeric and upright on the membrane', *Nat Commun*, 5, p. 3912.

Zerrouk, Z., Alexandre, S., Lafontaine, C., Norris, V. and Valleton, J.M. (2008) 'Inner membrane lipids of *Escherichia coli* form domains', *Colloid Surface B*, 63(2), pp. 306-310.

Zeth, K., Romer, C., Patzer, S.I. and Braun, V. (2008) 'Crystal structure of Colicin M, a novel phosphatase specifically imported by *Escherichia coli*', *J Biol Chem*, 283(37), pp. 25324-25331.

Zhang, X.Y.Z., Lloubes, R. and Duche, D. (2010) 'Channel domain of Colicin A modifies the dimeric organization of its immunity protein', *J Biol Chem*, 285(49), pp. 38053-38061.

Zhang, Y.L. and Cramer, W.A. (1993) 'Intramembrane helix-helix interactions as the basis of inhibition of the Colicin E1 ion channel by its immunity protein', *J Biol Chem*, 268(14), pp. 10176-84.