UNIVERSITY OF LEEDS

Periplasmic mechano-transduction networks

Samuel John Hickman

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Astbury Centre for Structural Molecular Biology

September 2016

Acknowledgements

David Brockwell has been an inspiring mentor who was always enthusiastic and came up with brilliant ideas that have resulted in this PhD project being so successful. I am indebted to his guidance, support and freedom to explore different experimental techniques throughout my PhD. I would also like to equally thank my co-supervisor, **Emanuele Paci** who has been a great support over the course of my PhD. He introduced me to molecular dynamics and computational biology, helping me flourish in a technique far from my comfort zone.

Many people were involved in the project directly or indirectly who shared their expertise and provided their time, which I am very thankful for. **Oliver Farrance** was my mentor for all things AFM related during the beginning of my PhD. **Nicholas Housden** provided cells, plasmids and protocols that have been essential for the work of this thesis. **Theo Karamanos** performed all of the NMR experiments; I am enormously grateful for all the time he gave up for me. **Sam Lenton** provided excellent knowledge of SAXS, helping me with the data collection and processing. The BtuB simulation would not have been possible without the initial help of **Luca Bellucci**. I would also like to thank **Rachael Cooper** for making lots of BtuB mutants, **Matthias Gantner** for help with the initial purification and characterisation of BtuB and **Nasir Khan** for sorting all of my purchases and helping out when equipment was playing up.

I would also like to thank **Sheena Radford** and everyone in the Radford/Brockwell lab between 2013-2016 who shared their wisdom and made my time at Leeds memorable.

Out of the lab, my biggest support has been **Theana Greenfield**, who has always been there for me through all the tough times. I am forever in your debt for all the patience, encouragement and love you seem to have an endless supply of.

Last but not least, thanks to my family for always supporting me.

.

Table of Contents

List of Figures

List of Tables

List of Equations

List of Abbreviations

AFM	Atomic force microscopy
APTES	(3-Aminopropyl)triethoxysilane
ATP	Adenosine triphosphate
APS	Ammonium persulphate
BAM	Barrel assembly complex
β-OG	Octyl β-D-glucopyranoside
CTD	C-terminal domain
CCCP	Carbonyl cyanide-m-chlorophenylhydrazone
CD	Circular dichroism
CSP	Chemical shift perturbation
DMSO	Dimethyl sulphoxide
DIPEA	N,N-Diisopropylethylamine
DTT	1,2-dithiothreitol
DDM	n-dodecyl-β-D-maltopyranoside
DEAE	Diethylaminoethyl
dNTP	deoxyribonucleic triphosphate
ΔL_c	Difference in contour length
DLS	Dynamic light scattering
EDTA	Ethylenediaminetetraacetic acid
EPR	Electron paramagnetic resonance

EOM	Ensemble optimisation method
F_U	Force of unbinding/unfolding
GFP	Green fluorescent protein
HSQC	Heteronuclear Single Quantum Coherence
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IM	Inner-membrane
LPS	Lipopolysaccharide
LIS	Lithium 3,5-diiodosalicylate
LB	Luria broth
LC-ESI-MS	Liquid chromatography electrospray ionisation mass spectrometry
L_c	Contour length
KLD	Kinase, ligase and DpnI
MWCO	Molecular weight cut off
MALLS	Multi-angle laser light scattering
MS PEG	methyl succinimide polyethylene glycol
MST	Microscale thermophoresis
NMR	Nuclear magnetic resonance
NAMD	Nanoscale Molecular Dynamics
Ni-NTA	Nickel nitrilotriacetic acid
MD	Molecular dynamics
OM	Outer-membrane

OMP	Outer-membrane protein
OD_{600}	Optical density at 600 nm
PCR	Polymerase Chain Reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PG	Peptidoglycan
PMSF	Phenylmethylsulphonyl fluoride
PPII	Polyproline type II
pmf	Proton motif force
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPE	1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine
PSD	Plug subdomain
PAGE	Polyacrylamide gel electrophoresis
rpm	Rotations per minute
RMSD	Root mean square deviation
SDS	Sodium dodecyl sulphate
SMFS	Single-molecule force spectroscopy
SMD	Steered molecular dynamics
SAXS	Small-angle X-ray scattering
SEC	Size exclusion chromatography
SM PEG	Succinimide maleimide polyethylene glycol
TAE	Tris-acetate-EDTA

TEMED	Tetramethylethylenediamine
TBDT	TonB-dependent transporter
TM	Transmembrane
TMD	Transmembrane domain
TBS	Tris-buffered saline
VMD	Visual molecular dynamics
WLC	Worm-like chain
XL	Cross-link

List of Amino Acid Abbreviations

Alanine	Ala	A
Cysteine	Cys	C
Aspartate	Asp	D
Glutamate	Glu	E
Phenylalanine	Phe	F
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Lysine	Lys	K
Leucine	Leu	L
Methionine	Met	M
Asparagine	Asn	N
Proline	Pro	P
Glutamine	Gln	Q
Arginine	Arg	R
Serine	Ser	S
Threonine	Thr	T
Valine	Val	V
Tryptophan	Trp	W
Tyrosine	Tyr	Y

Abstract

Unlike the inner-membrane, the outer-membrane of Gram-negative bacteria cannot be energised by a proton gradient and the periplasmic space is absent of ATP. This poses a significant problem for active processes such as transport against concentration gradients. TonB-dependent transporters (TBDT) are a class of outer-membrane proteins responsible for the scavenging and import of scarce metallo-organic complexes from the environment. They are structurally characterized by a 22-stranded β-barrel with the lumen occluded by an N-terminal globular 'plug' domain which contains a conserved binding motif known as the Ton box. Upon substrate binding, the Ton box becomes exposed to the periplasmic space where it forms a non-covalent complex with the C-terminus of a cytoplasmic membrane protein, TonB. The transport process requires energy from a proton motive force coupled with an inner membrane complex TonB-ExbB-ExbD. Regardless of a wealth of structural information, current models of the TonB-dependent transport mechanism are speculative. Conversely, despite no current experimental evidence, it is generally accepted that the plug domain must undergo a large conformational change facilitated by mechanical force exerted onto the Ton box tether by TonB.

In this thesis, force spectroscopy, protein engineering, molecular dynamics and bacterial growth assays are used to investigate the effects of force on TonB:TBDT complexes from *E. coli*. These experiments demonstrate that the channel of the vitamin B_{12} transporter (BtuB), reconstituted into synthetic liposome, can be opened by the application of force onto the plug domain via the non-covalent binding partner, TonB. Using wild-type BtuB and several of its mutants together with a related receptor (FhuA), the extent of plug remodelling is found to be highly controlled and determined by the cargo the receptor has evolved to transport. For both receptors, the plug domain can be regarded as comprising a mechanically weak channel forming sub-domain, and a mechanically strong sub-domain used both for allosteric signalling and to limit the size of the channel to allow passage of molecules no larger than its cargo.

Alongside these findings, structural and biophysical analysis of the periplasmic spanning protein TonB reveals conformation within the proline-rich linker domain, which allows speculation of the origin of the pulling force used for plug remodelling.

1 Introduction

Using a homemade single-lensed microscope, the first single cell organisms were discovered in rainwater by Antonie van Leeuwenhoek in 1675. The organisms swam swiftly in random directions and were estimated to be around 3 µM in size. Leeuwenhoek named the organisms 'animalcules' (from Latin *animalculum* = "tiny animal")[1]. 340 years after this landmark discovery, our understanding of animalcules (bacteria) has progressed tremendously. Over this time the advent of advanced microscopy, genetics, molecular biology, biochemistry, structural and biophysical analysis have helped reveal, in molecular detail, their central importance to human health and disease[2].

1.1 The outer-membrane of Gram-negative bacteria

During the 19th century, the Danish bacteriologist Christian Gram developed a revolutionary staining procedure that allowed the classification of nearly all bacteria into two large groups: bacteria that retain the crystal violet-iodine complex stain, designated Gram-positive, and those that did not, designated Gram-negative. The stain retention was due to differences in the structure of the bacterial cell envelope. Gram-positive bacteria retain the crystal violet due to an external peptidoglycan cell wall, whilst in Gram-negative bacteria, the peptidoglycan is shielded by an additional envelope, the outer-membrane (OM). The presence of the Gram-negative OM was not fully recognised until the advancement of electron microscopy in the 1950's.

The first clear demonstration that the plasma membrane and 'cell wall' of *E. coli* were composed of two separate entities were shown by Kellenberger and Ryter in 1958[3] using electron microscopy on specially fixed and stained phage infected cells. The separated 'cell wall' was found to exhibit a three layered structure, consisting of two dense outer layers separated by a less dense layer, with a total thickness of 60 to 90 Å (Figure 1.1).

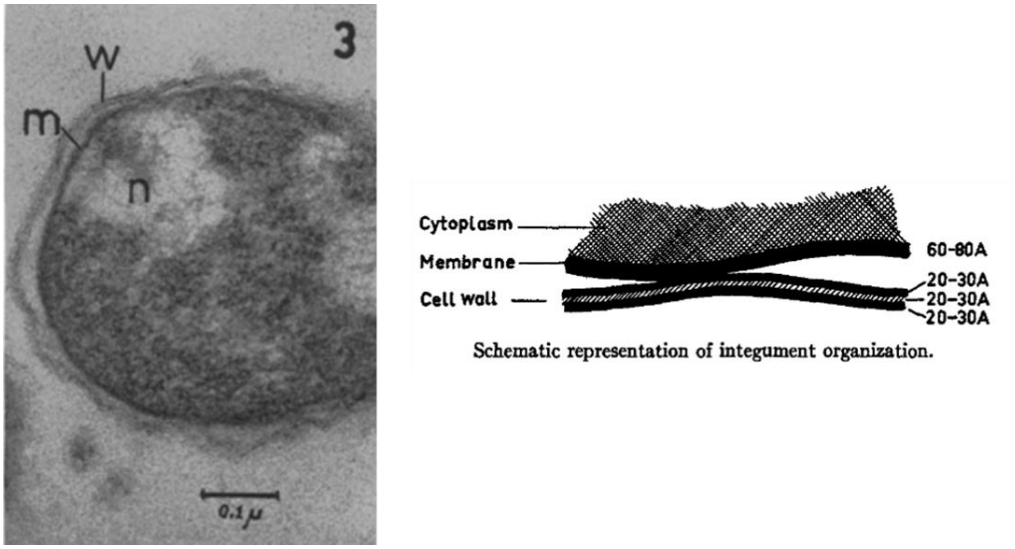


Figure 1.1 - The first observation by electron microscopy of the composition of the cell wall in *E. coli*. Left) x100,000 magnification of a T2 phage infected *E. coli*, no phage are visible, however the vegetative phage DNA is shown (n). The cytoplasmic membrane (m) and 'double structured' cell wall (w) are visible. Right) cartoon schematic of the organisation of the observable double envelope with measurements shown. Images taken from reference [3].

It is now accepted that the three principal layers observed by Kellenberger and Ryter in Gram-negative bacteria are; the outer membrane (OM), the peptidoglycan cell wall, and the cytoplasmic or inner membrane (IM). The compartment between the IM and the OM was coined the periplasmic space by Peter Mitchell in 1961 which was later found in 1967 to contain a distinct set of hydrolytic enzymes including nucleases and phosphatases[4]. The periplasmic space occupies around ~10 % of the total cell volume, and has been described as a highly viscous environment due to the high content of soluble proteins, with little free space available[5]. The periplasm is also an oxidising environment and contains protein disulfide isomerases[6], therefore protein folding occurs within this space, and interestingly the periplasm is completely void of ATP. Importantly the periplasmic space also hosts the peptidoglycan cell wall, which contributes to cell shape and protects the cell from lysis. The peptidoglycan is covalently linked to the OM by Braun's lipoprotein[7] which will be discussed in more detail later.

The two membranes in Gram-negative bacteria differ in composition completely. The lipid component of the IM is exclusively composed of phospholipids (both inner and outer leaflet), namely phosphatidylethanolamine (~70–80%), phosphatidylglycerol and cardiolipin. The OM on the other hand is asymmetric with respect to the inner and outer

leaflets, while the inner leaflet of the OM has the same lipid composition as the IM, the outer leaflet consists of glycolipids, principally lipopolysaccharides (LPS)[8].

1.1.1 Lipopolysaccharide

Bacterial smooth LPS (produced by many Gram-negative bacteria including the families Enterobacteriaceae and Pseudomonadaceae) typically consists of a hydrophobic domain known as lipid A, a "core" oligosaccharide domain, and a distal polysaccharide (named the O-antigen)(Figure 1.2)[9].

Lipid A is a glucosamine-based phospholipid that contains six or seven 14/12 carbon acyl chains that anchors the LPS to the outer leaflet of the OM. The architecture of lipid A is highly conserved across Gram-negative bacteria and is detected at picomolar concentrations by Toll-like receptors of the innate immune system[10].

The core oligosaccharide is divided into two regions, the inner core (directly attached to lipid A) and the outer core (attachment site for the O-antigen). The inner core typically contains Kdo (keto-deoxyoctulosonate) (1-3 units) and heptose (L-glycero-D-mannoheptose) (2-3 units). Kdo is found in all known cores and is linked to lipid A by a ketosidic bond. The minimal LPS required for the growth of *E. coli* consists of lipid A and Kdo (deep-rough phenotype). The inner core can also be phosphorylated and modified with pyrophosphorylethanolamine or phosphorylcholine. The outer core is more structurally diverse than the inner core, for example *E. coli* has been found to display five core types, the preference of which is dependent on the infectious strain. The outer core is composed of branched hexose unts (including: D-glucose, D-mannose, D-galactose) that are attached to the preceding heptose of the inner core. LPS that includes lipid A and a complete core oligosaccharide (inner and outer) is known as rough LPS. The O-antigen is ligated directly to a hexose unit within the outer core.

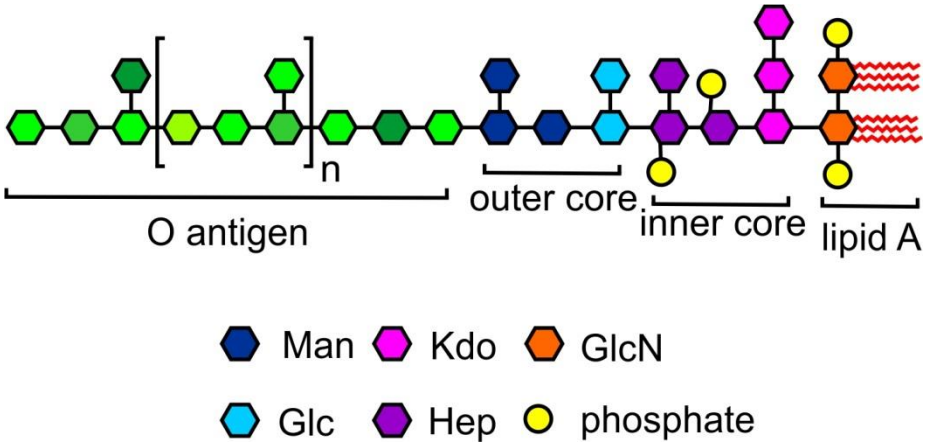


Figure 1.2 – Simplified cartoon structure of the organisation of lipopolysaccharide with the individual components (lipid A, the inner and outer core and the O-antigen) highlighted in different colours with labels.

The structural diversity of the O-antigen is extraordinary; with more than 60 different monosaccharides and 30 noncarbohydrate components recognised. The repeating glycan units (homo- or heteropolyers), the position and stereochemistry of the *O*-glycosidic linkages, the presence or absence of noncarbohydrate substituents, the various number of units, branching and nonstoichiometric modifications (such as *O*-acetylation or glycosylation) make the possible O-antigens conformations almost infinite. If an LPS extract is run on an SDS-PAGE from a smooth LPS containing strain, laddering is observed due to the extensive heterogeneity of the O-antigen. The number of unique O-antigens within a species can vary considerably, for example. *E. coli* produces around 170 O serotypes. As the O-antigen is fully exposed to the environment, its major role is the protection of the cell. In pathogenic Gram-negative species, the O-antigen and its chain length play a role in the evasion of the host immune system, in particular the complement cascade membrane attack complex[11] and neutrophil bactericidal/permeability-increasing protein[12]. The presence of the O-antigen also provides the bacterium with a protective barrier from hydrophobic molecules including antibiotics (macrolides, erythromycin and rokitamycin)[13]. To serve as an efficient barrier, the negatively charged LPS molecules bind avidly to each other laterally, especially if neutralising cations are present. The acyl chains in lipid A are mostly saturated and are tightly packed as a result. The combination of these LPS properties decreases the fluidity of the OM (described as 'gel-like'[14]) creating a very effective barrier for hydrophobic

molecules. Alongside protection, the O-antigen also has roles in the colonisation, invasion, inter- and intracellular spreading of infectious species[9].

1.1.2 Biogenesis of the OM

The synthesis of LPS occurs at the cytoplasmic side of the IM; therefore it is a remarkable feat how this amphipathic molecule with large hydrophobic (lipid A) and hydrophilic domains (O-antigen) is transported through the IM, across the periplasmic space and through the OM, irrespective of the type of LPS. For a long time it remained a mystery how the bacteria achieved the asymmetric OM, however in 2004 the discovery of the first component of the Lpt (LPS transport) system in *N. meningitides (*ΔlptD resulted in a reduction of LPS at the cell surface[15]) paved the way to a better understanding into the molecular mechanism. LptD exists as a complex in the OM with a lipoprotein LptE which is inserted into the lumen of the LptD 26 anti-parallel stranded β-barrel (Figure 1.3B)[16,17]. Other components of the Lpt system include the periplasmic protein LptA, integral IM proteins LptC (which contains a large soluble periplasmic domain), and LptFG and a cytoplasmic component LptB (sequence suggests a nucleotide binding domain), which is anchored to the IM complex in a stoichiometry of 2:1:1[18] (Figure 1.3A).

LptA and LptC have been shown to bind LPS, with LptA exclusively binding to lipid A[19]. When the crystal structure of LptA was determined (16 stranded antiparallel β-strand, with a slightly twisted β-jelly roll topology (typically eight β-strands arranged in two four-stranded antiparallel β-sheets) (Figure 1.3B)[20]), it was initially surprising that there was no hydrophobic cavity to accommodate lipid A, however, LPS cross linking experiments located a LPS binding groove inside of the jelly roll. Interestingly, LptD has a periplasmic N-terminal extension which shares structural homology to LptA. The current model (Figure 1.3A) predicts that this N-terminal periplasmic extension is in continuum with a polymer of LptA (forming a trans-envelope machine) in which LPS is shuttled from the IM to the OM[15]. LptA has been shown to form oligomers which are long enough to span the length of the periplasmic space and interacts with the periplasmic domain of LptC and N-terminal periplasmic domain of LptD (all sharing the same fold). Therefore LPS in the outer leaflet of the IM (flipped by MsbA) is thought to be extracted from the IM by LptFG using the ATPase activity of LptB, transferred to LptC where it is passed onto a

chain of LptA which crosses the periplasmic space and enters LptD via the N-terminal extension. The LPS is thought to exit the LptD barrel laterally through the first and last strands, as cross-linking these strands yields a lethal phenotype[16].

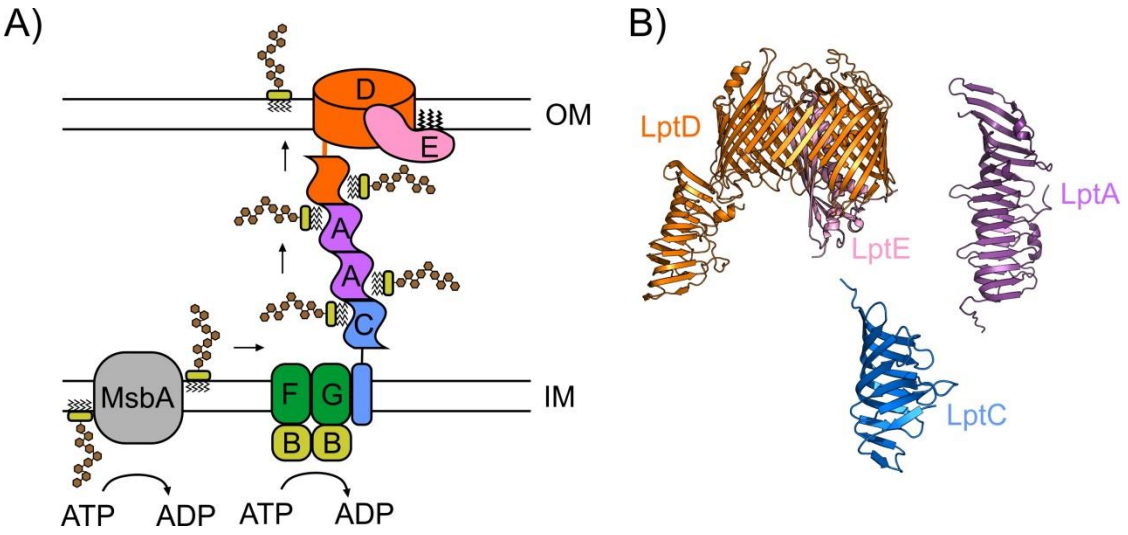


Figure 1.3 – LPS transport, the Lpt system A) Cartoon schematic for LPS transport with individual components labelled, figure adapted from reference [15]. LPS (gold and brown) is initially flipped across the IM by the ATP-binding cassette transporter MsbA, LptBFG then extracts the LPS from the IM and it is transported across the periplasm via a bridge of LptC, LptA and the periplasmic domain of LptD. It is then inserted into the outer leaflet of the OM by the integral membrane domain of LptD and lipoprotein LptE. B) The crystal structures of the Lpt components (labelled), the structure of LptA contains two subunits and the LptC structure is of the periplasmic domain only. (PDB: 5IV9, 2RI9 and 3MY2 for LptDE, LptA and LptC respectively).

Despite the huge advance over the last two decades in our understanding in molecular detail of how Gram-negative bacteria insert LPS into the outer leaflet of the OM, it is surprising that almost nothing is known about the transport of glycerophospholipids to the inner leaflet of the OM.

1.1.3 Lipoproteins of the OM

More than 100 lipoproteins are predicted to be present in *E. coli.* Although most lipoproteins have no known functions, they are expected to play various important activities in the periplasm that include signalling, chaperones, OM integrity and biogenesis[21]. Lipoproteins are attached to either the IM or OM by an N-terminal N-acyl-diacylglyceryl cysteine, which is transferred onto the protein in the IM by the enzyme Lgt (phosphatidylglycerol diacylglyceryl transferase). The signal sequence is cleaved by signal peptidase II and the amino terminus of the cysteine is N-acylated by Lnt

(phospholipid transacylase) generating the mature lipoprotein (Figure 1.4), this process is highly conserved in all Gram-negative bacteria[22].

Lipoproteins destined for the OM are identified by a lack of an amino acid retention signal (Asp at +2, (+1 being the lipidated Cys) and Asp, Glu or Gln at +3) flanking the lipidated cysteine. Lipoproteins lacking this signal are recognised and transported to the inner leaflet of the OM by the localisation of lipoproteins (Lol) system. In *E. coli* Lol is composed of 5 components (LolA-E), LolCDE is an ATP binding cassette complex found in the IM which extracts the lipoprotein from the IM to the periplasmic chaperone LolA. LolE recognises and binds the OM destined lipoprotein, LolC recruits LolA and LolD is the cytoplasmic ATPase which is tethered to the IM by LolC and E (Figure 1.4A). LolA contains a large hydrophobic cavity to accommodate the acyl chain of the lipoprotein substrate (Figure 1.4B), and delivers the lipoprotein to the final component of the system, LolB. LolB is an OM lipoprotein and also contains a large hydrophobic cavity with a higher affinity to the acyl chain than LolA. The mechanism of how LolB incorporates lipoproteins into the OM is currently unknown[22].

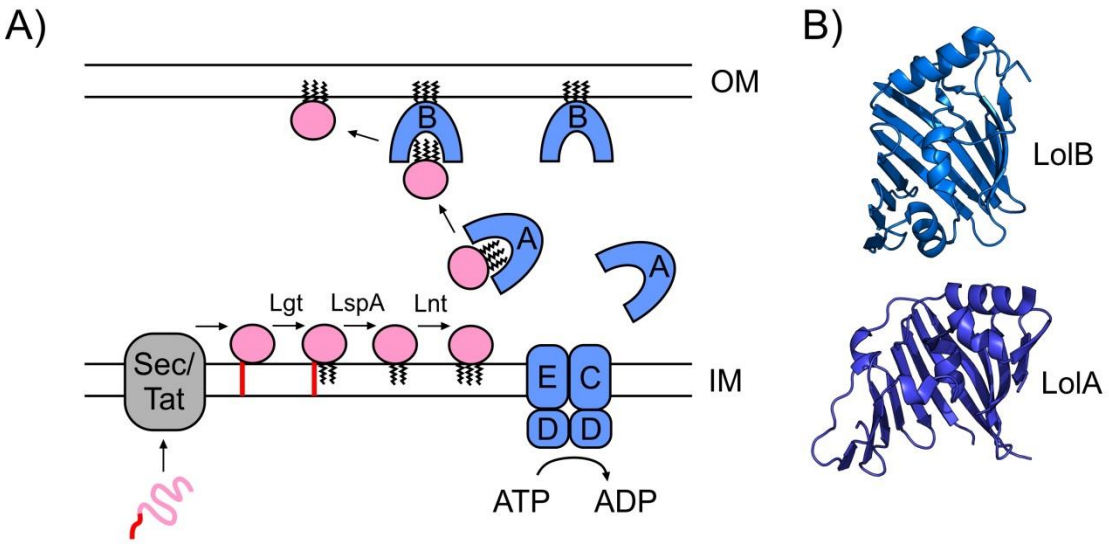


Figure 1.4 – Lipoprotein transport, The Lol system A) Simplified cartoon schematic of the lipoprotein maturation and export pathway by the Lol system (components in blue), figure adapted from reference [22]. The lipoprotein (pink) is synthesised in the cytoplasm with the N-terminal signal sequence (SS) (red) which targets the lipoprotein for translocation by the Sec/Tat translocon (grey). The lipoprotein initially remains anchored to the IM by the SS and Lgt adds a diacylglyceryl to the cysteine residue. LspA then cleaves the SS from the lipoprotein and Lnt adds another acyl chain to the new N-terminus. The LolCDE complex then powers extraction of the lipoprotein from the IM by ATP hydrolysis and forms a complex with the chaperone LolA which delivers the lipoprotein to the OM acceptor protein LolB. Finally LolB inserts the lipoprotein into the inner leaflet of the OM. B) Crystal structures of LolB and LolA (PBD: 1IWM and IUA8 respectively).

Two of the best studied lipoproteins in *E. coli* are Lpp (also known as Braun's lipoprotein/murein lipoprotein) and Pal (Peptidoglycan-associated lipoprotein), both function (as the names imply) in connecting the OM to the peptidoglycan (PG) cell wall. Lpp is the most abundant lipoprotein in *E. coli*, with an estimated 500,000 copies per cell. It is a small 6 kDa protein and is found either covalently cross-linked to the peptidoglycan (by the C-terminal Lys ϵ-amino group to the diaminopimelate in the PG[23]) or in a 'free form'. Pal, in contrast, interacts with the PG by a non-covalent interaction. Lpp and Pal have been shown to contribute to the OM stability as knockouts lead to OM vesiculation, release of periplasmic proteins and increased sensitivity to detergents[22]. Pal is a member of the Tol system, where TolAQR are IM proteins with TolA having a large periplasmic spanning domain and TolB is found in the periplasmic space[24]. TolA has been shown *in vivo* to use the *pmf* (using TolQR) to interact with Pal via TolB and stabilise the OM[25,24]. Not all lipoproteins in the OM have connections to the peptidoglycan, LptE

(part of the LPS transport system) is vital in the folding and assembly of the integral β-barrel LptD. Similarly, BamC (part of the barrel assembly machine complex (see Section 1.2.2)) is thought to help stabilise other components of the complex.

1.2 Outer-membrane proteins (OMPs)

The OM, like all biological membranes contains a bilayer of lipids which has little permeability for hydrophilic solutes including most nutrients required for survival. To overcome this barrier, the OM contains integral channel-forming proteins for the purpose of nutrient uptake. The proteins of the OM can be divided into two classes, lipoproteins and β-barrel proteins[23]. Strictly speaking, the term outer-membrane protein (OMP) refers to the integral β-barrel (various numbers of β strands wrapped into cylinders that traverse the OM) proteins only. The OMPs either facilitate passive diffusion of small molecules (also called porins), specific diffusion or active transport of large scarce ligands (known as the TonB-dependent transporters). There are also enzymatic OMPs such as PagP, OmpT and PldA which function as a lipid A palmitoyltransferase, aspartyl protease and phospholipase respectively[26-28]. OMPs also have roles in adhesion, assembly machinary and act as virulence factors for infectious Gram-negative species[29]. The details of the different types of OMPs involved in substrate transport and their insertion into the OM will be covered in the following sections.

1.2.1 Structural features of OMPs

The first non-specific activity of the OMP 'porin' channels were shown in 1976 by reconstituting OMPs into lipid vesicles, which were found to be permeable to a variety small compounds and sugars with a molecular weight less than 700 Da[30]. The 'porins' were soon found in all species of Gram-negative bacteria and even the mitochondrion[31]. The most significant advance in the study of OMPs was the elucidation of the structures of PhoE and OmpF by X-ray crystallography in 1992[32], which revealed the antiparallel β-barrel architecture and the trimeric nature of the porins. Before the structural determination of β-barrels, the protein sequences of OMPs baffled researchers due to predominant polar residues and a lack of hydrophobic rich regions, which are usually found in integral membrane proteins. It is now postulated that the OM is populated with β-barrels because proteins rich in hydrophobic residues would become stuck in the

inner-membrane or aggregate in the periplasmic space[33]. The wealth of OMP crystal structures now available have revealed that the main structural differences within the β-barrel is the number of β-strands (ranging from 8 to 26 (Figure 1.5)) and the shear number, which is a measure of the stagger between strands. Structural analysis found that these two parameters determine the geometry of the barrel, such as the tilt angle (angle of the β-strands against the barrel axis, ranging from 30-60 ˚) and the radius (ranging from ~7-20 Å (excluding FimD and LptD)) (Table 1.1)[34]. As the β-barrels have a right-handed β-twist, the shear is slightly larger than the strand number; also the shear value is always even and positive due to the hydrogen bonding between strands (Table 1.1). The shape of the lateral cross-section of the β-barrels also shows considerable variation that is dependent on the interior of the barrel (domains, peptide segments or inward facing side chains).

Name	β-strands	Shear	Radius (Å)	Strand tilt (˚)	Trimer
OmpX	8	8	7.2	37	no
OmpA	8	10	7.9	43	no
OmpT	10	12	9.5	42	no
OmpF	16	20	15.5	43	yes
PhoE	16	20	15.5	43	yes
LamB	18	20	17.1	40	yes
ScrY	18	20	17.1	40	yes
FhuA	22	24	19.9	39	no
FepA	22	24	19.9	39	no

Table 1.1 – Analysis of β-barrel structural information such as strand number, shear, radius, tilt and oligomerisation[34].

The type of cross-sections observed have been: circular (OmpA), ellipsoid (OmpT) or bean-shaped (FimD) (Figure 1.5). Another interesting feature of OMPs is the aromatic girdles which are found at both periplasmic and extracellular sides of the protein. The aromatic girdle is defined as outward facing aromatic residues found near the polar head groups of the lipid bilayer. The girdle was originally proposed to stabilise and anchor the β-barrel in a vertical position with respect to the membrane[35], which was later confirmed by mutagenesis[36,37].

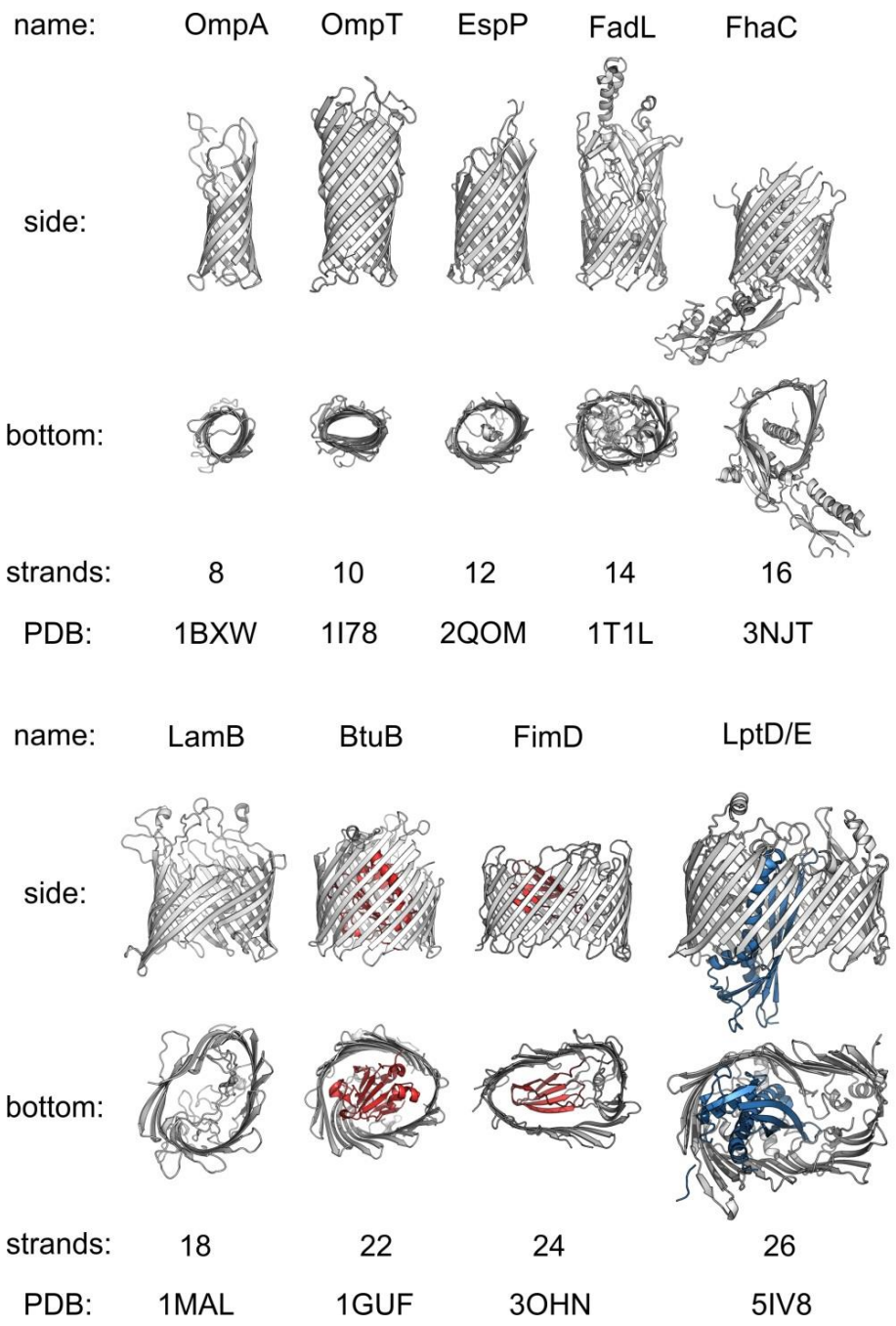


Figure 1.5 – Cartoon representations of the X-ray crystal structures of integral OMPs with 8-26 antiparallel β-barrels from the OM of Gram-negative bacteria. Globular plug domains are found within barrels greater than 18 strands (shown in red). For the enormous 26 stranded barrel of LptD, a separate lipoprotein (LptE) occludes the lumen, which is shown in blue. The monomeric unit of the LamB trimer is shown for simplicity. The proteins shown function as: a non-specific diffusion pore (OmpA), proteases (OmpT and EspP), fatty acid transporter (FadL), haemagglutinin transporter (FhaC), specific diffusion pore (LamB), vitamin B_{12} transporter (BtuB), pilus biogenesis usher (FimD) and LPS transporter (LptDE) (PDB codes shown).

The exterior of the β-barrel in contact with the non-polar groups of the lipid bilayer contains few charged residues (2.8 %) whilst aromatic and non-polar residues make up

29.3 and 49.2 % respectively, as expected[33]. The β-barrel interiors are much more polar. The β-strands are connected by short turns at the periplasmic side (1-12 residues), however on the extracellular side, long (up to ~46 residues), highly mobile and hydrophilic loops are often found, which have exceptionally high sequence variability between OMPs (due to environmental exposure)[38]. In fact, the entire sequence variability of OMPs is high in comparison to soluble proteins. The external loops are involved in the biological function of the OMPs, such as substrate binding (especially in the case of TonB-dependent receptors). In OmpF loop 3 folds inside of the β-barrel which narrows the lumen of the barrel enough to allow sugar diffusion[32]. Not all OMPs exist as trimers (Table 1.1), the trimeric OMPs (examples from *E. coli*: OmpF, OmpC, PhoE, and LamB) form a non-polar core at the three fold-axis. The hydrophobic trimer interface adds to the robustness of trimeric OMPs, as removal of specific monomer interactions results in reduction of thermodynamic stability[39].

The membrane embedded β-barrel structure of OMPs makes them extremely thermodynamically stable (some OMPs resist denaturation in the presence of 5 M guanidium hydrochloride or 2 % SDS at 70 °C[33]), which is intuitive as they are exposed to often harsh environmental conditions. When an OMP is inserted into the OM, their removal requires an energy that does not exist in the periplasmic space, therefore the cell cannot remove OMPs after insertion. Bacteria recycle OMPs by migrating them to the poles during cell division and passing old OMPs to daughter cells[40]. As folded OMPs are so stable, a large energy barrier must be overcome to insert them into the OM. The only essential OMPs found in *E. coli* are LptD (a huge 26 stranded β-barrel with a crucial role of inserting LPS into the OM[41]) and BamA (16 stranded β-barrel, part of a large complex with a crucial role in binding, folding and inserting β-barrel proteins into the OM). The latter will be discussed in more detail in the following section.

1.2.2 Folding OMPs into the OM

Like all proteins in Gram-negative bacteria, OMPs are initially synthesised in the cytoplasm by ribosomes. Therefore proteins destined to the OM must traverse the hydrophobic IM, navigate through the aqueous and crowded periplasmic space and then be correctly folded into a functional state at the OM. This must be achieved without

aggregation and maintaining the protein in a suitable unfolded state for OM insertion. To make matters more complicated, the bacteria also have to accomplish this extraordinary feat in the absence of chemical energy in the periplasm and OM. As mentioned in the previous section, two OMPs are essential for Gram-negative bacterial physiology, LptD (covered in Section 1.1.2) and BamA. BamA is a highly conserved 16-stranded β-barrel and is a component of the β-barrel assembly machinery (BAM) complex, which as the name suggests, is involved in the insertion of OMPs into the OM. The BAM complex contains five components, BamA, the integral membrane protein component which also contains a N-terminal periplasmic region called the POTRA (polypeptide transport-associated) domain, which contains five individual folded domains. The other components (BamB, BamC, BamD, BamE) are all lipoproteins anchored to the inner leaflet of the OM (Figure 1.6). BamA and BamD form the core of the complex (both components are essential[42]) with BamD and BamB interacting directly with BamA, whilst BamC and BamE directly interact with BamD[43]. The presence of all five components is required for the most efficient OMP insertion, despite only BamAD being essential. The structure of each component of the BAM complex has been solved by X-ray crystallography, and also the structure of the entire complex (BamABCDE)(Figure 1.7) with and without BamB[44]. Despite a wealth of structural information, the exact mechanism of how the BAM complex folds OMPs independent of ATP is not clear. Structural insight from the BamA component and cross-linking studies has led to speculations that client OMPs are released into the OM by a lateral gating mechanism. Cross-linking the first and last β-strands (which are only weakly associated) of the BamA barrel yields a lethal phenotype, and MD simulation demonstrated free opening of these strands[45]. The crystal structures of both BamABCDE and BamACDE complexes show a laterally open BamA barrel[44]. Other models of BAM insertion include thinning and destabilisation of the membrane near the lateral opening site, which would lower the kinetic energy barrier of insertion[46]. The signal that the BAM complex is thought to recognise is the C-terminal strand of OMPs, which all contain a terminal phenylalanine[47].

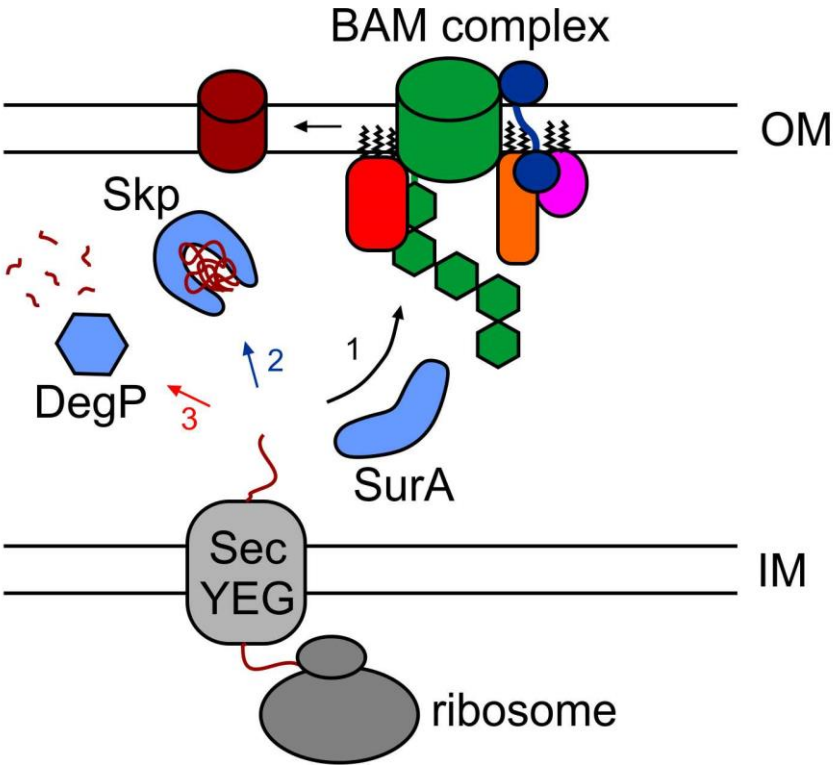


Figure 1.6 – Simplified cartoon showing OMP biogenesis, regulation and insertion into the OM of Gram-negative bacteria (image adapted from reference [43]). Prefolded OMPs (maroon line) are initially synthesised in the bacterial cytoplasm and cross the IM by the SecYEG translocon. In the periplasmic space, the unfolded OMP can then be delivered by the primary pathway (1) by SurA or a secondary pathway (2) by Skp to the BAM complex (BamA (green) with POTRA domains (hexagons), BamB (red), BamC (blue), BamD (orange) and BamE (pink)) for OM insertion. If the OMP becomes misfolded (3), it is targeted by DegP for degradation.

The BAM complex serves the function of OMP insertion into the OM, where periplasmic chaperones aid the passage of the unfolded OMP through the periplasmic space, which unfortunately is less well understood than OM insertion. The most thoroughly studied and essential chaperones for OMP biogenesis present in *E. coli* are Survival protein A (SurA), Seventeen kilodalton protein (Skp) and DegP (Figure 1.6)[43]. SurA is a peptidyl-proline isomerase which facilitates the correct folding of OMPs, and was first identified for its crucial role in cell survival during stationary growth phase. Variants of SurA have been found to reduce the amount of OMPs in the OM and also increase the bacterial sensitivity to detergent and antibiotics (likely due to the phenotypical impact of reduced OMP activity, such as LPS export) which is a classic sign of OMP misfolding[48]. Phage display demonstrated that SurA preferentially recognises the sequence repeat motif Ar-X-Ar (Ar = aromatic, X = any)[49], which is commonly found in the C-terminus of OMPs and

is also thought to target the OMPs to BAM. The structure of SurA (Figure 1.7) is composed of 4 domains, an NT and CT domain (N and C terminal) which have been shown to bind to OMPs specifically and two prolyl isomerase (PPIase) domains which are non-essential for OMP biogenesis[50]. Interestingly *in vitro* experiments have demonstrated that SurA does not directly fold OMPs into lipid bilayers and an interaction with BamA has been identified[51]. The current hypothesis is that SurA delivers the OMP to the BAM complex in an 'insertion-competent' state[43].

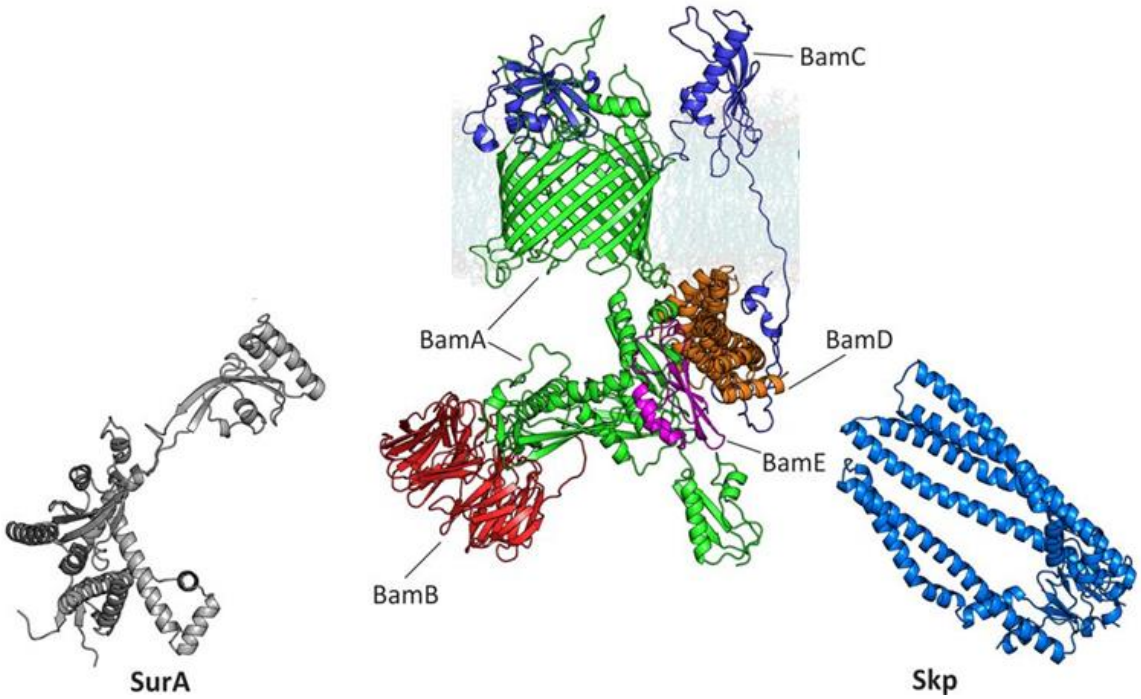


Figure 1.7 – Cartoon representations of the crystal structures of the components of the OMP biogenesis pathway. The entire BamABCDE OM complex (PDB: 5D0O) with individual components coloured and labelled. Chaperones SurA (PDB: 1M5Y missing regions built using MODELLER) and Skp (PDB: 1U2M, missing regions built using PyMol) are shown in grey and blue respectively. Image taken and adapted from reference [52].

Skp is a jellyfish-like trimer (Figure 1.7) that has a large central cavity between the three helical 'legs' and is believed to be an antifolding chaperone, which does not facilitate OMP folding directly. In the unfolded state, OMPs preferentially bind to the hydrophobic cavity of Skp where they are shielded from the aqueous environment of the periplasmic space. Skp is specific for the unfolded β-barrel domain of OMPs only[53] and NMR has shown the unfolded OMP to be highly dynamic within the Skp cage[54]. Skp has an electrostatic dipole with the 'head' of the jelly-like structure being highly positively

charged. As the POTRA domains of BamA contain patches of negative charge, there is a possibility that the POTRAs attract Skp to deliver the unfolded OMP to the BAM complex[43]. It has been recently demonstrated that Skp acts as a multivalent chaperone, where the size of the unfolded OMP substrate is dependent on the copies of Skp required to prevent folding. For example a Skp/Omp ratio of 4:1 is required for 16 stranded OMPs, whilst a 2:1 ratio was sufficient for 8 stranded barrels[227].

The final periplasmic component to be discussed is the serine protease DegP, which is responsible for the clearance of misfolded, aggregated OMPs by cleaving between pairs of hydrophobic residues. DegP exists as homo-multimers (6, 12 and 24-mers) where misfolded substrates trigger oligomerisation into cage-like structures, activating the DegP protease cleavage site (by removal of a regulatory loop). DegP has homologs in eukaryotic chloroplasts and mitochondria, therefore the presence of DegP in organisms and organelles with an intermembrane space suggest an important role in the regulation of OMP biogenesis[55]. Where *surA* knockout mutants cause an impairment of OMP biogenesis, knockouts of *degP* and *skp* have a lesser effect. Even double knockouts of *skp degP* display cellular viability[56]; therefore Skp and DegP have been suggested to not be directly involved in OMP biogenesis like SurA, but act in a secondary pathway, such as quality control (Figure 1.6).

1.3 OMP transporters

1.3.1 Non-specific diffusion, the porins

Porins (by strict definition, a nonspecific diffusion channel[14]) are highly abundant in the bacterial OM with up to 10^4-10^6 copies per cell depending on the bacterial species and the environmental conditions[33]. The first to be discovered in *E. coli* were OmpC and OmpF, which are 16-stranded β-barrel trimers (Figure 1.8) and are termed the “classical porins”. They consist of water-filled channels and serve as general pathways for the influx of small hydrophilic molecules with a molecular weight cut off of ~700 Da, excluding the size of many antibiotics. The classical porins show no particular substrate specificity, however they display some selectivity for either cations or anions[33], for example OmpC is slightly more cation-selective than OmpF despite sharing 60 % sequence homology. This observation was originally predicted to be due to a smaller

pore size in OmpC[14], however when the *E. coli* OmpC structure was solved, the structure was found to be identical to OmpF (RMSD = 0.78 Å)[57] with the electrostatic potential mediated by residues lining the pore determining the physiological differences. A key feature in OmpC, OmpF and PhoE (a classical porin with selectivity for anions) is the presence of a conserved extracellular loop (L3), that folds into the channel creating a constricted pore (known as the eyelet) approximately half way through the channel. A 'carboxyl cluster' positions L3 into the barrel interior by electrostatic interactions with luminal facing Arg/Lys residues from the barrel. Acidic residues in L3 and basic residues in the opposing β-strand create an electrostatic field that regulates the diffusion of molecules through the constricted region. The eyelet of the classical porins therefore contributes significantly to the permeability properties, such as the exclusion limit and ion selectivity of the pore. Another selective filter which gives individual porins their unique substrate preference is the overall charge of the channel mouth[58].

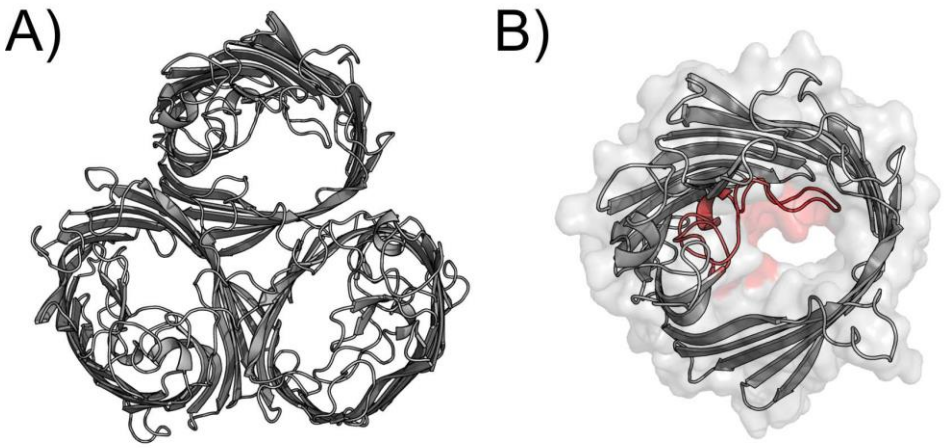


Figure 1.8 – Cartoon representation of the X-ray crystal structure of the nonspecific OmpF porin (PDB 1MPF). A) The trimeric state of OmpF shown from above the plane of the membrane. B) The L3 'eyelet' of OmpF is shown in red which restricts the size of the channel's lumen (surface representation).

It is not surprising that environmental conditions have a strong influence on which classical porins are upregulated. For example OmpC has been shown to be preferentially expressed at high osmolarity, whereas OmpF is expressed in low osmolarity[59]. As OmpF allows permeation of slightly larger solutes than OmpC, it would be detrimental to have high OmpF expression at high osmolarity. Low pH is another condition that increases

OmpC levels while decreasing OmpF levels and phosphate deprivation leads to the upregulation of PhoE which has a slight preference for phosphate uptake. The classical porins serve as the main route into the cell through the OM for small hydrophilic antibiotics including β-lactams, tetracycline, chloramphenicol and fluoroquinolones[14] and antibiotic resistance due to reduction in porin expression or a change in the porin upregulated has been reported for a range of Gram-negative bacteria[60].

1.3.2 Substrate-specific diffusion

Size and charge is not the only solute filter the OMPs passively facilitate. Channels have evolved for the specific diffusion of substrate into the periplasmic space and the best studied example is the maltose transporter LamB from *E. coli* (often referred as 'maltoporin', but named LamB due to the prior observation of it being the receptor for λ –bacteriophage). Maltose is less than 600 Da in molecular weight and therefore can diffuse through the classical porins, however as diffusion rates rely on a concentration gradient across the membrane, when the environmental maltose is below the millimolar range, uptake by OmpF/C is greatly diminished. As maltose is the major breakdown product from starch in the intestine, the efficient uptake of this carbon source is vital for bacterial survival[14]. The X-ray crystal structure of LamB was elucidated in 1995 revealing a trimer of 18 stranded β-barrels, which like the classical porins contains an extracellular loop (L3) that enters the interior of the barrel creating a narrow central constriction[61]. The constricted region is narrower than the classical porins and the pore is further occluded by inward folding of loops 1 and 6. Soaking LamB crystals in maltose, maltotriose, and maltohexaose revealed the substrate binding site as a row of aromatic residues arranged in a helical manner that are opposed by polar residues from L3 (named the greasy slide and ionic track respectively (Figure 1.9)). It is proposed that the hydrophobic face of the sugar residues are in van der Waals' contact with the greasy slide, while hydrogen bonds are formed between the sugar hydroxyls and the side chains of the ionic track.

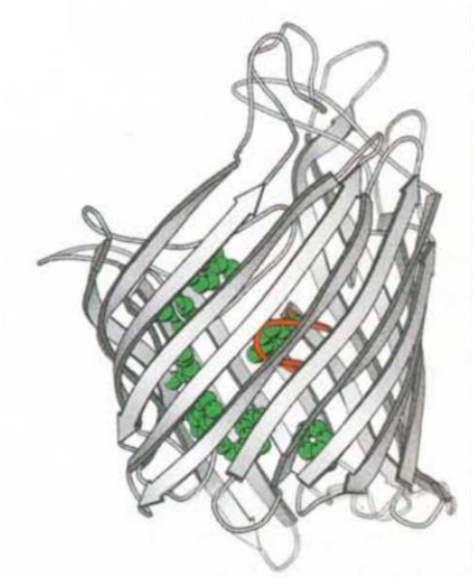


Figure 1.9 – Cartoon representation of a single unit of the LamB trimer with the residues of the greasy slide (W74*, Y41, Y6, W420, W358, F227 and Y118) shown as green spheres and the ionic track from L3 (which restricts the channel) shown in red. Image taken from reference [62].

Therefore the sugars 'slide' along the aromatic residues while continuous disruption and formation of hydrogen bonds are made with the ionic track. It should be noted that LamB does not solely transport maltose; however sugars containing fructose (such as sucrose) are prevented. A transporter called ScrY (from *Salmonella typhimurium*) has been found to be specific for sucrose and differs from LamB at the restriction site (LamB is more restricted than ScrY)[33].

Not all OMP passively diffuse their substrates, active transport is required for larger substrates such as iron-siderophore complexes, heme or vitamin B_{12}. The receptors involved in active transport belong to the TonB-dependent transporter family and are the main focus of this thesis.

1.4 TonB-dependent transporters

1.4.1 Discovery

The Ton system was initially discovered by seminal work in 1943 by Luria and Delbrück when a virus-resistant strain of *E. coli* B was isolated that was resistant to bacteriophage T1[63]. The resistant strain contained mutations in genes which were subsequently named *tonA* and *tonB* (Ton being named after the **T on**e phage). This was the first demonstration that inheritance (in this case viral resistance) in bacteria follows

Darwinian principles and genetic mutations can occur randomly in the absence of a selection pressure.

It was not until 1973 that developments in biochemical techniques allowed Braun and colleagues to purify the *tonA* gene product from the OM envelope of *E. coli*[64], the second protein ever to be isolated from the OM (the first being Braun's lipoprotein). The purified protein was able to bind T5 phage and colicin M (another ligand for TonA) and was shown by electrophoresis to be a polypeptide chain of around 85 kDa. During the same year as the isolation of TonA, two more OMPs were purified from the OM, the lambda phage receptor LamB[65] (which was later shown to transport maltose[66]) and the receptor of colicin E3[67] (later found to be transporter for vitamin B_{12}[68]), BtuB. TonA was renamed to FhuA (for *f*erric *h*ydroxamate *u*ptake) after the discovery that TonA also facilitates the transport of ferrichrome (an iron-complexing hydroxamate siderophore)[69].

The initial finding that mutation within *tonA* (FhuA) and *tonB* prevented T1 phage uptake suggested a functional relationship between the two gene products. In 1951, T1 infection was shown to occur in two distinct steps, an initial reversible step (requiring tonA) followed by an irreversible step (requiring tonB) upon which the phage could not be eluted from the cell by agitation or dilution[70]. The irreversible step was proposed to require an energised cell, which was confirmed by Hancock and Braun, who found that an energised inner-membrane (provided by the electron transport chain/ATP hydrolysis) was required for the irreversible step[71].

FhuA is a highly diverse receptor for the phages T5, T1, and φ80, colicin M and ferrichrome. It was found that T5 phage can infect cells independent of TonB function, however ferrichrome uptake, colicin M killing and infection with T1 and φ80 phages requires both functional FhuA and TonB[72]. By the late 1970s, tonB had been shown by numerous studies to be essential for the lethal actions of colicins B, I, and M, the uptake of iron chelates[73] and vitamin B_{12}[74]. The requirement for both tonB and an energised inner-membrane led to the hypothesis that they are coupled. Hantke and Braun tested this and found that TonB is required for the transport function for a class of OM receptors, as *tonB* deficient cells were resistant to T5 phage killing when ligand (ferrichrome) was present, therefore blocking the T5 binding site of the receptor with

bound untransported ligand[73]. Thus the term 'TonB-dependent transporters' (TBDTs) was coined for receptors that require TonB to facilitate transport.

1.4.2 Interaction with TonB: discovery of the Ton box binding motif

During the 1980's when cloning and DNA sequencing was becoming more achievable, an explosion of DNA sequences became available, including the DNA sequences of the TonB-dependent receptors BtuB[75], FhuA[76], IutA[77] and FepA[78] which had all been mapped, cloned and sequenced by 1986. After sequencing the FepA gene, Lundrigan and Kadner performed a multiple sequence alignment to uncover regions of homology in the receptors[78]. It was discovered that the sequences shared high homology near their N-termini. A BtuB variant had been studied by Heller and Kadner that was able to bind but not transport vitamin B_{12}[75]. Therefore the authors suspected that this mutant was TonB binding deficient. The sequence of this BtuB mutant (named btuB451) contained a leucine to proline substitution at position 8 (L8P), located in the highly homologous N-terminus. It was thus proposed that this region interacts with TonB and was named the 'TonB box or Ton box'. Interestingly it had also been shown that the sequences of colicins B and M also contained the Ton box motif[38]. In 1989 the Kadner group performed extensive mutagenesis (32 substitutions) within the Ton box region (residues 6-13) of BtuB to find the residues necessary to transport vitamin B_{12}[79]. They found that only mutations L8P and V10G/P were completely transport defective, whilst other mutants (T7N, L8R, V9G, V9P, T11N and N13T) exhibited significantly reduced transport efficiency. A similar finding was observed by the Braun group who performed a mutagenesis screen on the Ton box of FhuA (I9P analogous to BtuB L8P had a reduced transport phenotype but was still T5 sensitive)[80]. The conclusion from both studies was that TonB likely interacts with the receptor by a secondary structure facilitated manner (prolines being the most disruptive) rather than a sequence specific manner and that TonB interacts with other residues other than just the residues within the Ton box.

During the 1990s to mid-2000s, multiple biochemical and biophysical studies were carried out to find whether TonB does indeed interact with the Ton box motif. Tuckman and Osburne demonstrated a synthetic Ton box pentapeptide (Glu-Thr-Val-Ile-Val) could inhibit the growth of *Escherichia coli* in low iron medium, prevent killing by colicin B and

Ia and stop Bacteriophage phi-80 infection, presumably by blocking the TonB binding site[81]. The Postle group used a combination of *in vivo* formaldehyde cross-linking and Western-blotting to show that TonB interacts with TonB-dependent receptors by forming a high molecular weight species of 195 kDa, which was absent when fepA was knocked out[82]. They also created C-terminal truncations in TonB and found that the final 48 residues were essential for an interaction with FepA[83]. Similar formaldehyde cross-linking experiments showed that the TonB interaction with FecA was enhanced by the presence of the FecA ligand (ferric citrate)[84]. In 1999, Cadieux and Kadner demonstrated by engineering cysteines at specific positions within TonB (Q160C, Q162C and Y163C) and each residue within the Ton box of BtuB (^{6}DTLVVTA^{12}) that specific residues preferentially reacted with each other. This demonstrated that the Ton box of BtuB assumes a β-strand structure with alternating residue side chains in the same direction relative to TonB (ie. a parallel β-strand augmentation). The disulfide bridge formation was also enhanced by the presence of vitamin B_{12}, suggesting that the Ton box is exposed upon substrate binding[85]. In 2005, NMR chemical shift perturbations were used to analyse the structural changes when synthetic Ton box peptides derived from FhuA, FepA, and BtuB interact with TonB[86]. Large chemical shifts all occurred within the same region of TonB (along the β3-strand, specifically residues Gln160, Ala167, Gly174, Ile232, and Asn233), indicating a common binding site. Isothermal calorimetry between the structured C-terminus of TonB and the FhuA Ton box peptide (FepA and BtuB Ton boxes were insoluble at high concentrations) revealed a relatively weak binding (K_D) of 36 ±7 μM[86]. The weak binding was unexpected as the NMR spectrum for the protein-peptide interaction was in the slow exchange regime, a feature of tight binding. However the authors suspected this was due to a slow association and dissociation rate and that TonB may undergo structural rearrangement of the previously hidden β4-strand revealed in the solution structure (discussed in the next section) before binding the Ton box. In 2006, X-ray crystal structures of the C-terminus of TonB in complex with FhuA[87] and BtuB[88] were solved at 3.3 and 2.1 Å resolution respectively (Figure 1.18C), confirming that the Ton box interacts with TonB by a parallel β-strand addition augmentation as predicted by Cadieux and Kadner. The complex revealed that hydrogen bonds are formed between residues 6-12 of BtuB and residues 226-232 of TonB (Figure 1.10D). A salt-bridge between conserved TonB residues Arg166 and Arg158 with FhuA-Glu56 and

BtuB-Asp7 respectively was also observed. This salt-bridge was proposed to be important for the initiation of the intermolecular β-strand between TonB and receptor.

As the Ton box is essential for transport, it is important to understand how the substrate-bound TBDT signals to TonB to facilitate the binding. Crystallographic studies have demonstrated that the conformation of the Ton box does not correlate well with the presence of substrate[89]. However, electron paramagnetic resonance spectroscopy (EPR) studies from the Cafiso group were able to demonstrate that substrate binding to BtuB resulted in an disordered Ton box, whereas in the apo state the Ton box displays a refolded conformation, forming part of the plug domain[90]. It was also shown that reagents used in protein crystallisation can prevent this conformational switching[91].

1.4.3 TonB, a multi-domain periplasmic spanning membrane protein

TonB is present (with equivalent forms) in all Gram-negative bacteria and has been postulated as an energy transducer between membranes since the mid-1970s. Little was known about the gene product until 1983 when Postle and Good cloned the *tonB* gene and acquired the DNA sequence[92] revealing a protein of approximately 26 kDa (239 amino acids) with a 17 % proline content. The high proline content gave an anomalous mass by SDS-PAGE of 36-40 kDa. The proline-rich region (residues 66-100) was found to contain an unusual amino acid sequence $(EP)_4$-IPEPPKEAPVVIE-$(KP)_6$, a synthetic peptide of this sequence was subsequently studied by proton NMR revealing an extended conformation, which was proposed to have a periplasmic spanning function[93]. In 1988, direct evidence emerged that TonB remains associated with the IM (found in the spheroplast fraction)[94], in agreement with the DNA sequence which showed a hydrophobic N-terminus[95]. Fusions of periplasmic protein β-lactamase to the TonB C-terminus were still active, suggesting that the C-terminus of TonB remains in the periplasmic space[96], rather than being associated with both the inner and outer-membrane. IM attachment of TonB has also been shown to be necessary for TonB-dependent function, as deletion of the N-terminal transmembrane (TM) domain of TonB (TonBΔ1-32), engineering a peptidase cleavage site near the TM domain or by substituting the TM domain with a tetA or penicillin-binding protein 3 TM domain[97] abolishes TonB activity. The residues within the TM anchor of TonB associate with an

additional IM protein named ExbB (see next section) and conserved residues Ser16 and His20 were found to be essential for TonB activity and TonB:ExbB heterodimer formation[98]. The spacing residues between (^{16}SXXXH^{20}) can be substituted but not deleted, implying a spacing function only[98].

Residues 66-100 are described as the proline-rich region of TonB, functioning to span the periplasmic space as a 10 nm rigid rod[93], despite this, deletion of residues within the proline-rich domain (TonBΔ(66–100)) is fully capable of irreversible ϕ80 adsorption[99]. However in high saline growth conditions (which results in an expanded periplasm), the irreversible step was inhibited in TonBΔ(66–100) but not in wild-type TonB. A more recent spin-label EPR double electron–electron resonance (DEER) study was used to make long range distance measurements within the proline-rich domain of TonB[100]. Six double cysteine mutants were engineered between residues 56-126 of TonB (59/69, 59/76, 69/76, 69/84, 88/106, and 106/120) for spin label attachment. This resulted in an average spin-label distance of 0.26 nm per residue and a full length of 14.4 nm. With a periplasmic space size of ~20 nm (under physiological conditions)[101], this length would be sufficient for TonB to reach the OM when the structured C-terminal domain is included. Circular dichroism of TonB residues 56–126 shows a strong signal for a polyproline type II (PPII) helix, which were also in agreement with the EPR measurements. The EPR data also indicated some flexibility of the proline-rich domain, rather than a completely rigid rod.

The C-terminal domain of TonB (residues 150-239) that interacts with the outer-membrane receptors[85] has been studied by X-ray crystallography and NMR. The first structure of TonB (residues (165-239) solved in 2001[102] revealed a strand-exchanged dimer (Figure 1.10A), however this was later found to be a consequence of construct truncation within the structured region of the domain[103]. In 2005 a TonB construct (residues 150-239) was solved by X-ray crystallography[104] and solution NMR[86], which were almost identical (RMSD 0.9 Å)(Figure 1.10B and Figure 1.10C) and they both contained a 3-stranded antiparallel β-sheet packed against an α-helix. The difference was that the NMR structure was monomeric and contained an addition β-strand (β4, residues 235–239 (Figure 1.10C)) that folds back onto the Ton box binding site (β3). The same region in the crystal structure (residues 220–236) extends outwards and hydrogen

bonds to an adjacent β3 from another molecule of TonB, forming a homodimer (Figure 1.10B). Interestingly the crystal structure of TonB (residues 153-233 are visible) in complex with BtuB lacks electron density for the final C-terminal 6 residues that are visible in the NMR structure. This suggests that the β4-strand from the NMR structure has been replaced by the Ton box of BtuB (purple strand, Figure 1.10D), implying a conformational rearrangement before Ton box binding, explaining the low binding affinity (36 ± 7 μM) found by isothermal calorimetry[86].

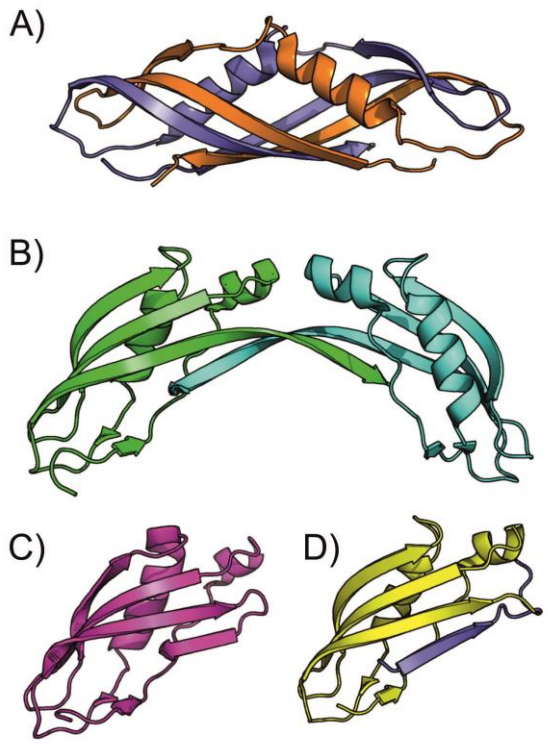


Figure 1.10 – Cartoon representations of the various X-ray crystal and solution NMR structures of the C-terminal domain (CTD) of TonB. A) The first crystal structure of a truncated $TonB_{CTD}$ (residues 165-239), the individual subunits of the intertwined dimer are coloured in orange and purple (PDB: 1IHR). B) Crystal structure of $TonB_{CTD}$ (residues 150-239) which formed dimers in the crystal environment (PDB: 1U07). C) NMR solution structure of monomeric $TonB_{CTD}$ (residues 150-239), the additional β4 is resolved (PDB: 1XX3). D) Crystal structure of monomeric $TonB_{CTD}$ (residues 150-239, yellow) bound to BtuB (Ton box shown in purple) (PDB: 2GSK), β4 in the NMR structure has been replaced by the Ton box.

TonB dimerisation has been proposed to have physiological relevance *in vivo*[105] despite a monomeric solution structure[86], monomeric forms in the crystal structures bound to BtuB and FhuA[72,73], and the authors of the crystal dimer structure (residues 150-239 (Figure 1.10B)) show that the dimer was only present in the crystal environment, not in

solution[104]. Double electron-electron resonance (DEER) has demonstrated that a significant population (~50 %) of dimeric TonB (residues 33-239, missing only the TM domain) exists in solution, however when associated with the TBDT the dimer is converted to the monomeric form[106]. This monomer-dimer inter-conversion has also been proposed to be important in the transport cycle shown by *in vivo* formaldehyde cross-linking[107]. *In vitro* studies investigating a full-length TonB construct (including the TM domain) are yet to be reported, therefore the biological relevance of the TonB homodimer still remains questionable.

1.4.4 The accessory proteins ExbB and ExbD

While TonB has clearly been shown as the energy transducer, two additional IM proteins, ExbB and ExbD have been shown to be important for energy transduction in TonB-dependent processes. Exb functions of *E. coli* are related to TonB as the presence of Exb enhances TonB-dependent functions. Exb was initially identified in *E. coli* resistant to colicin B and reduced sensitivity to group B colicins[108]. The Exb locus of *E. coli* was cloned and sequenced in 1989, revealing two open reading frames (named ExbB and ExbD), which encoded proteins of 244 (26 kDa) and 141 (17.8 kDa) amino acids respectively[109]. The proteins were found to contain homology to the TolQ and TolR gene products[110] which are involved in the uptake of colicins A, E1-3 and K. The sequence of ExbB/D revealed extensive hydrophobic regions implying that both are integral IM proteins which was confirmed by cellular fractionation[109]. The transmembrane topologies of both proteins have been determined with ExbD possessing a single transmembrane (TM) domain near the N-terminal (similar to TonB)[111] whilst ExbB has three TM domains, oriented so the majority of the protein located in the cytoplasm (Figure 1.11)[112].

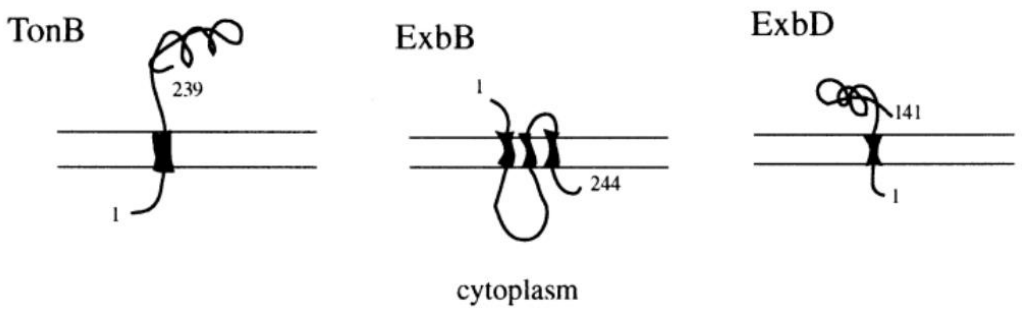


Figure 1.11 – Simple schematic of the transmembrane topologies of the TonB energy transduction complex in the IM, the N- and C-terminal domains are shown by the residue number (image taken from reference [113]).

Early evidence indicated that ExbB/D and TonB are found as a complex[114,115], with ExbB appearing to stabilise both TonB and ExbD. The existence of a membrane complex was further supported by *in vivo* formaldehyde cross-linking experiments which found TonB interacts with ExbB by its transmembrane domain[116] and both the periplasmic domains of ExbD and TonB interact[113]. Point mutations within ExbD revealed that when the single charged residue in the TM domain (D25) was mutated to D25N, Ton related activity was lost[115]. The D25 residue is believed to be responsive to the *pmf*.

The only structural information currently available on ExbB/D is the periplasmic domain of ExbD, where the NMR solution structure revealed that residues 64 to 133 form a well-defined folded globular domain (two alpha-helices that are located on one side of a single beta-sheet) where the surrounding residues (43 to 63 and 134 to 141) are unstructured[117]. The structure was solved at pH 3 due to a high propensity to aggregate, therefore the physiological relevance is questionable. Using the ExbD construct studied by NMR, peptides representing the TonB PE (residues 70–83) and PK (residues 84–102) repeats, along with the TonB C-terminal domain (residues 150-239) were tested for binding by chemical shift perturbations. Only weak binding was found in the PE peptide to the structured domain of ExbD. In contrary to this, *in vivo* disulfide cross-linking experiments have shown the periplasmic domains of ExbD and TonB to interact[118] and the absence of TonB favours ExbD homodimer formation[119].

Despite no structural information available, ExbB has been shown to have a functional role in the stabilisation of TonB and ExbD[120], acting as a scaffold for the TonB:ExbD/B

complex formation[121], and does not utilise the *pmf*[122]. ExbB TM domain (TMD) 1 interacts with TonB TM domain, whilst ExbB TMD2 and TMD3 (highly conserved) are believed to be involved in signal transduction from the cytoplasm to the periplasmic space[122]. Quantification of the stoichiometry of the TonB:ExbB:ExbD complex has been performed by several research groups yielding different ratios. The first quantification used both iron-rich and iron-depleted growth conditions where the complex was found to exhibit a stoichiometry of $TonB_1$-$ExbD_2$-$ExbB_7$[123]. More recently, the full TonB:ExbB:ExbD complex was co-purified and analysed in DDM detergent by size exclusion chromatography with multi-angle laser light scattering (SEC-MALLS) revealing a stoichiometry of $TonB_1$-$ExbB_4$-$ExbD_1$. This reported stoichiometry was further complimented by a low resolution negative stain EM topology structure[124].

A study by mass spectrometry found the formation of subcomplexes of ExbBD without TonB, with major species being $ExbB_6ExbD_1$ and $ExbB_6$ (with a minor amount of $ExbB_5ExbD_1$ also observed). Importantly, ExbD dimers were never observed[125]. Despite these findings, *in vivo* formaldehyde cross-linking has shown the presence of ExbD homodimers, and ExbD homodimerisation has even been suggested to be of functional importance to the transport cycle[107]. Therefore the exact stoichiometry remains uncertain, although the presence of multiple copies of ExbB and a single copy of TonB in the complex is consistently observed.

1.4.5 Structural features of TonB-dependent transporters

The first crystal structures were solved in the late 1990s of two *E. coli* TBDTs, the ferrichrome transporter (FhuA[126,127]) and ferric enterobactin transporter (FepA[128]), which revealed that TBDTs contain a 22-stranded β-barrel with a unexpected globular plug domain folded into the interior of the barrel occluding the large (≈35–40 Å diameter) pore of the barrel (Figure 1.12).

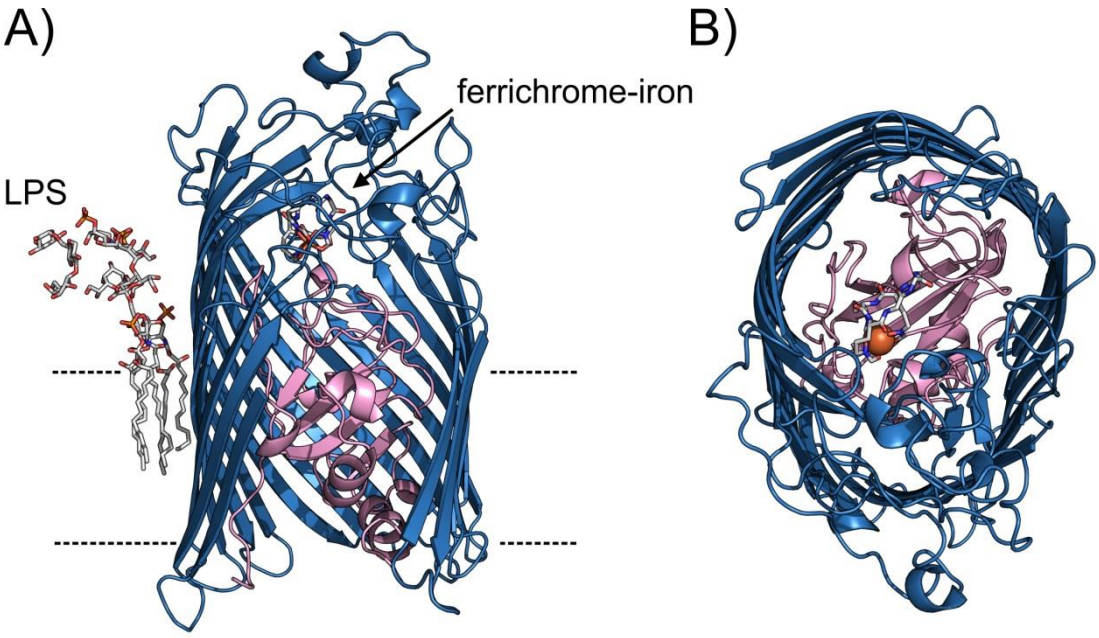


Figure 1.12 - Revealing the novel structure of the TonB-dependent transporters: the first X-ray crystal structure of FhuA (blue and pink cartoon) in complex with ferrichrome-iron and a molecule of LPS (stick representations, with labels) (PDB: 2FCP). A) Strand residues 621 to 723 are hidden to allow visualisation of the interior globular plug domain (pink), the extracellular loops involved in substrate binding are shown at the top of the barrel, encaging the ferrichrome-iron ligand. The dashed lines indicate the positions of the upper and lower aromatic girdles which sit on either side of the lipid bilayer. B) Top view of FhuA[126].

In 2005, a comprehensive structural analysis of four TBDT structures from *E. coli* (FhuA (apo/liganded), FepA, BtuB (apo/liganded) and FecA (apo/liganded)) was performed[129]. The four TBDTs contained the same domain architecture with ligand binding sites in the long extracellular loops of the β-barrel and extracellular facing residues of the plug domain. The Ton box was located N-terminally and is found at the periplasmic side of the plug domain either folded within the barrel or disordered and not visible in the structures (ligand bound FecA and both apo and liganded FhuA). The ordering of the Ton box did not correlate well with either substrate bound or apo forms from crystallographic studies[89]. A structure-based sequence alignment has revealed that most of the conserved residues were found at the plug barrel interface and this interface is also highly solvated resembling a transient protein complex (Figure 1.13B). Barrel strands β1-8 show no sequence conservation; whilst β9-18 contain some strongly conserved residues, most of which face the lumen of the barrel. An interesting feature of the barrel was revealed, loop 7 (of strands β13 and 14, named the β-cantilever (Figure 1.13A, blue strands)) folds towards the barrel lumen breaking the regular inter-strand

hydrogen bonding of the barrel and forms extensive interactions with the 'latch region' of the plug domain (Figure 1.13A and Figure 1.14). The latch region tightly interacts with conserved residues within strands β9-13.

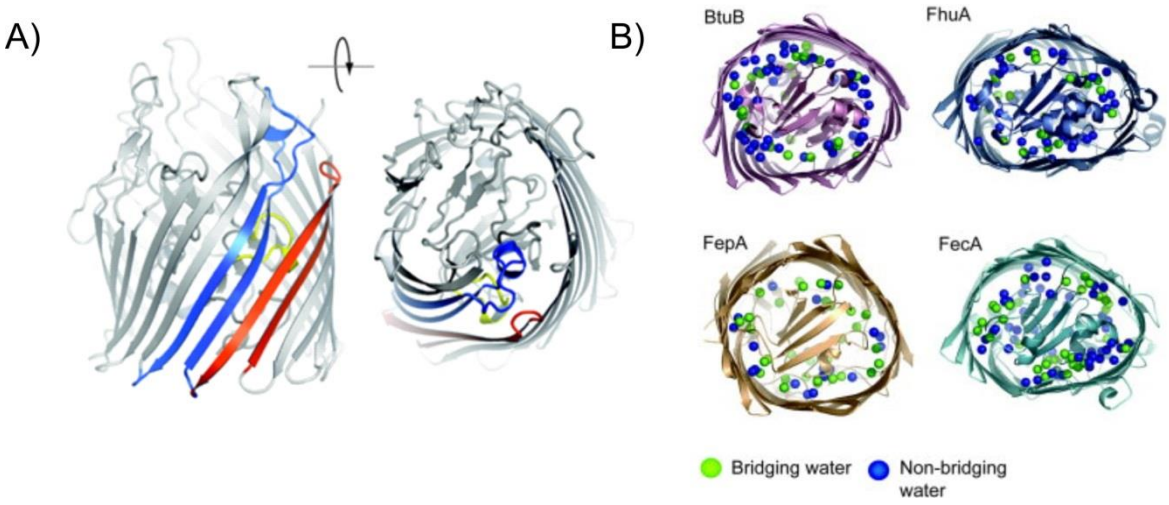
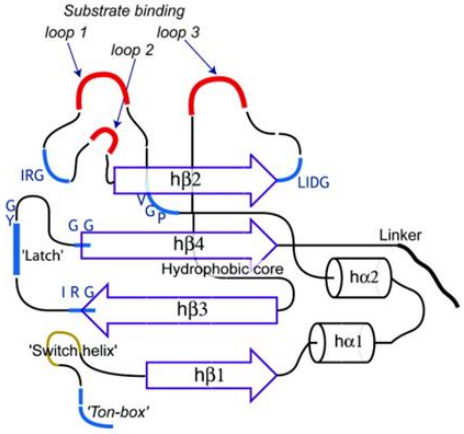


Figure 1.13 - Comparative structural analysis of TonB-dependent transporters (images taken directly from reference [129]). A) Membrane and extracellular view of the the vitamin B_{12} receptor BtuB. The β-cantilever (strands 13 and 14, blue) tilting from strands 11 and 12 (red) and the latch motif within the plug domain (yellow) are shown. B) Intracellular view of the structures of BtuB, FepA, FhuA, FecA (labelled) with bridging (green, protein-water-protein hydrogen bonds) and non-bridging (blue) waters at the plug–barrel interface.

Some highly conserved motifs other than the Ton box were discovered within the plug domain which are also noted in a more recent structure-based sequence alignment of 12 unique TBDTs[89]. These include the TEE, PGV, IRG box, LIDG box, RP box, latch and Hβ4 (Figure 1.14). The PGV motif forms three backbone hydrogen bonds to the IRG box (PGV and IRG are strictly conserved in TBDTs), forcing a substrate-binding loop 1 (SB1) of the plug domain to extend toward the extracellular side of the transporter. Similarly the IRG motif forms a sharp turn producing the second substrate-binding loop (SB2). Substrate binding loop 3 (SB3) is formed by the LIDG-box. Hβ4 (Hatch (alternative name for the plug domain) β-strand 4) forms the hydrophobic core of the plug domain.



Structure motif	Sequence	BtuB	FepA	FhuA	FecA
Ton box	DTLxxTAN	6-12	12-18	7-13	80-86
TEE	TxEE	30-33	32-35	54-57	108-111
PGV box	PGV	50-52	53-55	74-76	128-130
SB1 Gly	G	58	64	82	139
IRG box	IRG	68-70	74-76	92-94	149-151
LIDG box	LIDG	79-82	85-88	105-108	163-166
RP box	RxEϕϕRGP	106-113	121-128	128-135	191-198
latch	xYG	117-119	132-134	139-141	202-204
Hβ4	GGVVNxxT	124-131	139-146	146-153	209-216

x = non-conserved, ϕ = hydrophobic

Figure 1.14 - Comparative structural analysis of TonB-dependent transporter plug domains (plug domain schematic taken directly from reference [129]). The table shows structurally conserved motifs within the plug domain and a 2D secondary structure schematic (β-strands as purple arrows, α-helices as cylinders and substrate binding loops in red) of the plug domain of TBDTs with conserved motifs highlighted.

There are currently (May 2016) 58 X-ray crystal structures of TonB-dependent transporters, representing 15 unique transporters, substrate-bound or apo forms and in complex with TonB[87,88] or colicin E2[130] and E3[131]. All share the features highlighted above, however despite a wealth of structural information, the exact transport mechanism still remains uncertain.

1.4.6 Proposed mechanisms of transport

Despite structural and biochemical evidence of the TonB interaction with the TonB-dependent transporters, it is still unclear how this interaction results in the transport of the substrate. The first hypothesis (Figure 1.15) was proposed in 1980 by Bradbeer when both TonB and the *pmf* were known to be required for the uptake of vitamin B_{12} through BtuB in *E. coli* and TonB was also known to be an IM protein. TonB was proposed to use the *pmf* to either generate a secondary messenger, or a *pmf*-dependent efflux of a secondary messenger into the periplasmic space. The hypothetical secondary messenger molecule then antiports exchanges with vitamin B_{12} or energises BtuB for substrate transport[132].

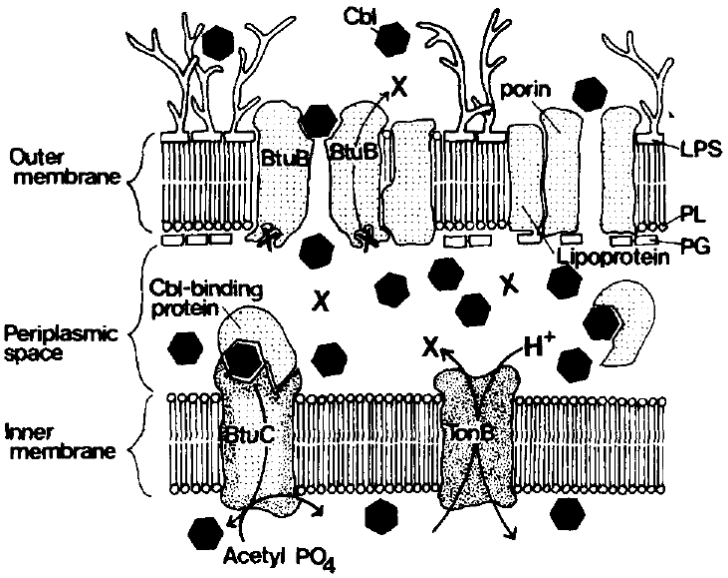


Figure 1.15 – Diagram of the first suggested model of TonB-dependent transport of cobalamin (Cbl, black hexagons) before structural or genetic information of the gene products were available (image taken from reference [132]) X represents a hypothetical messenger molecule generated in the periplasm by TonB using the *pmf*, the messenger was hypothesised to bind to the OM receptor and facilitate transport by either an allosteric effect or antiport with the substrate.

In 1982, TonB was postulated to be a periplasmic permease, either by acting as a positive regulator of a permease associated with the TBDTs, or TonB itself is a permease and modifies the TBDTs[133]. The shuttle hypothesis emerged in 1997 when sucrose density gradient fractionation of cell lysates revealed that TonB was distributed between both the IM and the OM in a 60:40 ratio respectively. This led to the idea that TonB shuttles

between the IM and OMs to transduce energy for substrate transport[134]. The propeller model was proposed in 2001 when the accessory proteins ExbBD were shown to be homologous to the flagella motor proteins MotAB[135] along with the crystal structure of the intertwined dimer of TonB (residues 164-239) (Figure 1.10A)[102]. The authors of the dimeric crystal structure speculated that ExbBD provides torsional motion of TonB and two proline-rich regions would provide a stiffer structure to directly transduce torsional force to the plug domain of the TBDT. The application of mechanical force by TonB to remodel the plug domain of the transporter ('pulling hypothesis' model) originates from single-molecule force spectroscopy experiments of β-stranded proteins[136,137], where it was found that parallel β-strands serve as a force resistant clamp in an orientation dependent manner (perpendicular directed forces being significantly weaker than parallel (see Section 1.5.4)). Chimento *et al* performed comparative structural analysis on the available crystal structures of four TBDTs[129], the plug domains were noted to contain a core four-stranded β-sheet. If a force by TonB was applied perpendicularly (ie. away from the receptor) to the strands then only a very modest mechanical force would be sufficient to cause substantial conformational change or unfolding of the plug domain (Figure 1.18A and B).

The three main hypotheses (shuttle, propeller and pulling) will be each discussed in depth.

1.4.6.1 The shuttle model (1997-2011)

As mentioned above, the shuttle model was initially postulated from the results of a cellular fractionation where TonB was found in both IM and OM fractions[134]. In the shuttling model (Figure 1.16), TonB is initially in an unenergised conformation in complex with ExbBD in the IM. The *pmf* is used by ExbBD to convert TonB to an 'energised conformation', which then dissociates from the IM carrying the stored energy and associates with the OM transporters to transfer the energy to drive transport. TonB has been shown to cross-link *in vivo* to OmpA and Lpp[138], therefore could temporarily dock at these proteins waiting for TBDTs to signal ligand occupancy (Ton box exposure).

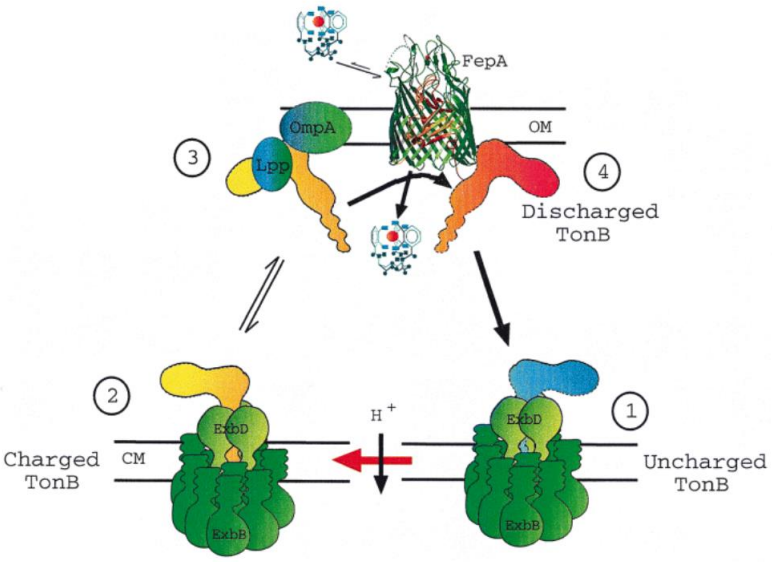


Figure 1.16 – The shuttle model of TonB-dependent transport, cartoon representative of the crystal structure TonB-dependent receptor FepA is shown in the OM, whilst the other components (with no structure at the time of publication) are shown as coloured blobs. (1) A 'uncharged TonB' (blue) is converted to charged TonB (2) by the *pmf* using the ExbB/D complex (green). The charged TonB (yellow) leaves the cytoplasmic membrane (CM) and shuttles to the OM where it may dock with non-transporter proteins (such as OmpA or Lpp) until FepA binds its substrate (enterochelin) (3). Substrate binding results in a conformational change in FepA causing an interaction with charged TonB which releases the stored potential energy to drive substrate transport (4). Image taken from reference [113].

The 'energised' state of TonB was supported by the formation of a proteinase K-resistant wild-type TonB from spheroplasts when carbonylcyanide *m*-chlorophenylhydrazone (CCCP) was used to inhibit the *pmf*[139]. Additionally, *in vivo* labelling experiments demonstrated that a fluorescent label (Oregon Green(R) 488 maleimide) that is permeable to the OM but not the IM could label an engineered cysteine (L3C) on full length TonB. This implies that the extreme N-terminal domain of wild-type TonB was exposed at some point to the periplasm and supports the notion of TonB leaving the IM[140]. The shuttle model received its first major criticism when GFP was fused to the N-terminus of TonB to prevent it from leaving the IM, full TonB-dependent activity was observed and the authors declared TonB does not shuttle[141]. The shuttle model was then dismissed in 2011 by the original author in the aptly named publication 'Death of the TonB Shuttle Hypothesis' where they showed that fusing the N-terminus of TonB to the cytoplasmic domain of ToxR was capable of 100 % specific activity (ferrichrome uptake) as wild-type TonB (ie. does not leave the IM). It was also shown that Oregon Green(R)

488 maleimide (the strongest evidence of the shuttle model) was able to cross the IM and label cytoplasmic located TonB ΔTMD L3C. The 40% of OM located TonB was concluded to be tightly associated with a TBDT and was artificially pulled out of the IM during cellular fractionation[142].

1.4.6.2 The propeller/ROSET model (2001-present)

In the propeller model, TonB undergoes a rotary motion that is initiated by ExbB, ExbD and the *pmf* (Figure 1.17). When TonB associates with the ligand bound TBDT, the rotary motion initiated at the IM causes conformational change to the plug domain and causes the release of the bound ligand into the periplasmic space.

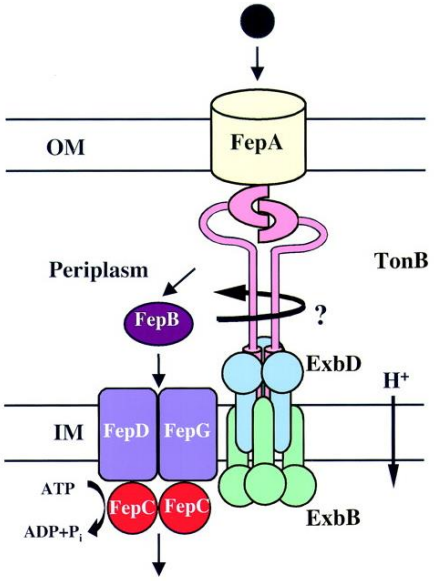


Figure 1.17 – The first schematic depicting the hypothesis of TonB rotation driving conformational change to the plug domain of a TBDT (the propeller model). After the intertwined dimeric crystal structure of TonB(165-239) (Figure 1.10A) was solved, TonB (pink) was proposed to exist as a dimer *in vivo* where it associates with the plug domain as a dimer and torque is applied by the ExbBD (blue and green) complex using the *pmf*. The substrate is then delivered to the cytoplasm by FepBCDCG components (labelled). Image taken from reference [102].

The idea of TonB rotation originated from the discovery of homology between ExbBD and MotAB[135], however supporting evidence towards the model was not investigated until 2013 when a GFP-TonB fusion construct was created for the analysis of GFP anisotropy in living bacteria[143]. Fluorescence polarisation measurements demonstrated motion of TonB *in vivo*, the motion (predicted to be rotational) of GFP-TonB was slower

than that of free GFP in the cytoplasm. However, dissipation of the *pmf* by exposure to CCCP led to further reductions in motion of GFP-TonB. Deletion of *exbBD* rendered TonB motion insensitive to the same metabolic inhibitors, suggesting a dependency of TonB motion on ExbBD and the *pmf*. The motion was not directly shown to be rotational in this study, however due to rapid (2.5 ns) reorientation of the fluorescence transition dipole of GFP, other types of membrane protein motions such as lateral diffusion and rigid body motion are excluded[144]. In a recent review, the 'propeller model' (name originating from the aberrant dimer crystal structure[102]) was rebranded 'rotational surveillance and energy transfer (ROSET) model'[144]. TonB is proposed to interact weakly only as a dimer with peptidoglycan (PG) (due to structural homology between the aberrant dimer structure of TonB to the PG binding motif LysM[141]). Moving laterally as a complex with ExbBD in the IM, TonB scans the OM until a exposed Ton box is encountered, the dimer then dissociates to a monomeric form (no longer able to associate with the PG) forming a complex with the Ton box. Using the constant rotational motion generated by ExbBD, kinetic energy is then transferred to the plug domain of the TBDT creating a conformational change, and transporting the substrate into the periplasmic space. The notion of a pulling force to facilitate transport is not dismissed in this revised model.

1.4.6.3 The pulling model (2005-present)

The pulling model proposes that TonB reaches across the periplasm using the extended proline-rich domain and binds to the Ton box of a ligand bound TBDT. The TonB:ExbBD complex then applies a force on the plug domain by extension away from the receptor by an unknown mechanism, which partially unfolds the β-sheets of the plug domain and allows for the translocation of the substrate. The high resolution crystal structure of the C-terminus of TonB (residues 153-233) and vitamin B_{12} non-covalently bound to BtuB confirmed that the Ton box:TonB is complexed by a parallel β-strand augmentation rotated ~90° with respect to the β-strands of the plug domain (Figure 1.18C)[88]. The authors speculated that only partial unfolding by mechanical force of the plug domain by TonB would be sufficient for substrate transport and the degree of unfolding may be a function of substrate size.

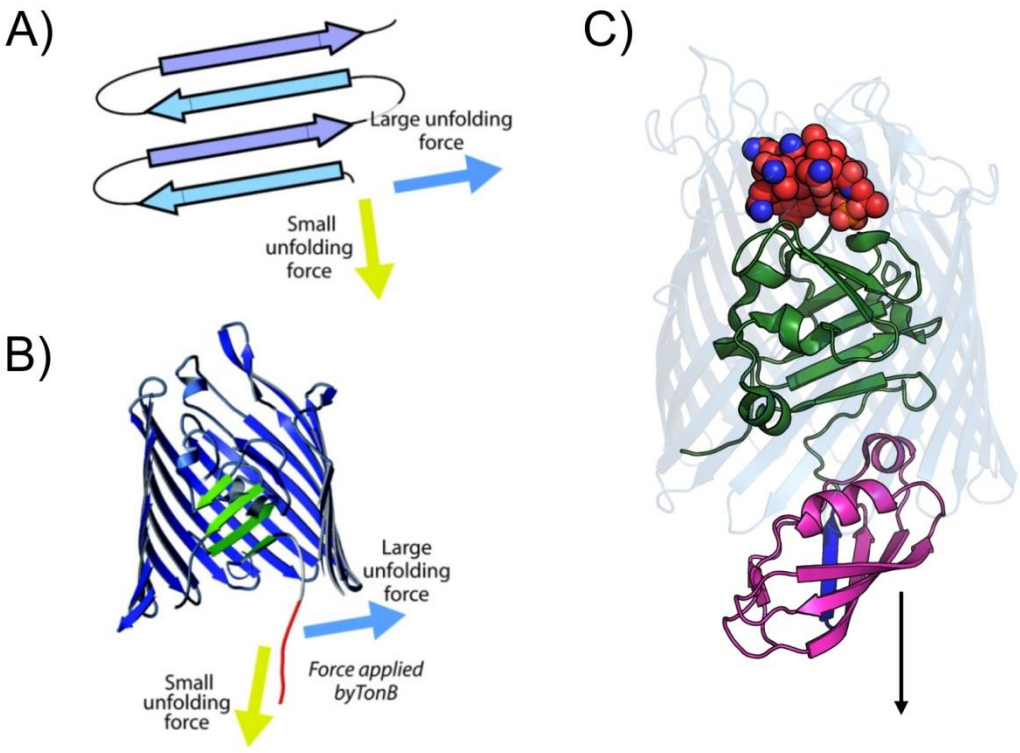


Figure 1.18 – The mechanical unfolding of the plug domain by TonB (the pulling model). A) Single-molecule force spectroscopy experiments demonstrated that the direction of force applied onto the β-sheets correlates with the amount of force required to unfold them. B) The structures of TBDT plug domains indicate that if force was applied by TonB perpendicular to the β-sheets (yellow arrow) of the plug domain (green) by the Ton box tether (red) then remodelling would occur under a small amount of force. C) Cartoon representative of the X-ray crystal structure of TonB(residues 150-239) (pink) bound to the Ton box (blue) of BtuB (PDB: 2GSK). The plug domain is shown in green, vitamin B_{12} in red and blue spheres and the barrel domain is transparent for clarity. The Ton box is perpendicular to strands of the plug domain, indicating that the unbinding of the Ton box will require more force (if extended in the direction of the black arrow) than the remodelling of the strands of the plug domain. Image A and B taken from reference [129].

In silico steered molecular dynamics (SMD) was performed on the crystal structure of BtuB bound to TonB (vitamin B_{12} was excluded due to lack of parameters at the time). The BtuB:TonB complex was inserted into a POPE lipid bilayer and TonB was extended away from the receptor by applying constant velocity (2.5 Å/ns) to the N-terminal Cα atom (Pro153) of TonB. TonB and the Ton box remained tightly bound (TonB-Arg158 and BtuB-Asp6 salt bridge was maintained during extension and oriented TonB so that the Ton box was parallel to the plug domain strands) as TonB was extended by 9 nm. Longer extension could not be performed as a larger system would require prohibitive computational effort[145]. The first β-strand and α-helix downstream of the Ton box

unfolded during the 9 nm extension (Figure 1.19A), whilst the TonB:Ton box interaction was maintained. After this time, the simulation was extended by cycles of deleting the unfolded N-terminal residues and extending with the new N-terminus of BtuB. This was performed until a sufficient amount of plug residues were unfolded to accommodate substrate passage, which was found after 20 nm of extension (Figure 1.19B). SMD was also applied to the centre of mass of the plug domain testing the feasibility of intact plug removal from the barrel; however biologically inaccessible amounts of force (~4500 pN) were required for complete dislodging.

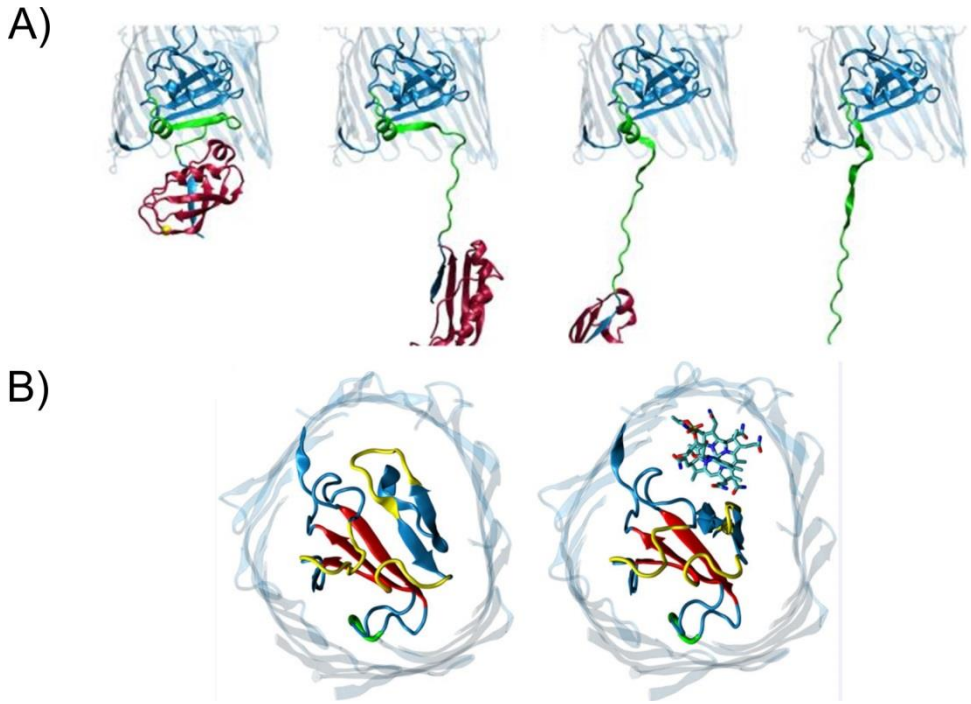


Figure 1.19 – Steered molecular dynamics (SMD) study of applying force to the TonB:BtuB complex. A) Partial unfolding of the plug domain by TonB, TonB (red) was extended away from BtuB (blue, barrel shown as light blue and transparent with the front strands removed for visualisation of the plug domain) by its N-terminal residue (yellow sphere). The part of the plug domain remodelled during the simulation is shown in green (residues 12-38). States of unfolding are shown at (left to right) 0 Å, 55 Å, 70 Å, and 90 Å of extension. B) Continued SMD unfolding of the plug domain of BtuB (extracellular view, loops removed for clarity). The core β-sheet is shown in red, the substrate binding loops in yellow and the latch domain (See Figure 1.14) in green. The states shown are after approximately (*left*) 150 Å and (*right*) 215 Å of pulling. After 215 Å of unfolding cyanocobalamin (blue sticks) is placed to scale in the open space created in the transporter. Images taken from reference [145].

The possibility of TonB generating a pulling motion was suggested in a study on the proline-rich region of TonB[75], were the rotatory motion provided by ExbBD causes the

predominant poly-proline type II (PPII) helix to form an unstable PPI conformation. This conformational change is accompanied by a contraction of up to 40% of the length of the proline-rich segment, generating an inwards force. Other than the study mentioned above, there is currently no *in vitro* or *in vivo* evidence supporting the pulling model (with the ability of TonB to mechanically remodel the plug domain or the origin of the pulling force).

1.4.7 The Btu system: transport of vitamin B_{12} in *E. coli*

E. coli cannot fully synthesise vitamin B_{12} and it is not an essential nutrient, however it is advantageous for *E. coli* to acquire vitamin B_{12} because the synthesis of methionine is 100 times more efficient if vitamin B_{12} is utilised by the methyltransferase MetH[146].

As vitamin B_{12} is large (> 700 Da) and scarce, it fulfils the specifications for transport in a TonB-dependent manner. The Btu (B-twelve uptake) (Figure 1.20) system consists of the OMP BtuB, periplasmic cobalamin-binding protein BtuF and an IM ATP-binding cassette complex $BtuBC_2D_2$ (which also forms complex with BtuF). $BtuC_2$ functions to span the IM (Figure 1.20, green) by forming 20 transmembrane helices grouped around a translocation pathway, whilst BtuD serves as a nucleotide binding domain to hydrolyse ATP (at the dimer interface) driving the conformational change required for transport. Initially, vitamin B_{12} binds the extracellular side of BtuB, where TonB-dependent activity translocates it across the OM. Once in the periplasm, vitamin B_{12} binds within a large cleft in the periplasmic protein BtuF which shuttles it to the $BtuC_2D_2$ complex. BtuF binding to the periplasmic side of $BtuC_2D_2$ initiates hydrolysis of ATP by BtuD, which opens the dimer, pulling the TM helices apart. The passage of vitamin B_{12} through $BtuC_2D_2$ has been described as peristaltic, whereby the helices force the substrate into the cytoplasm[147].

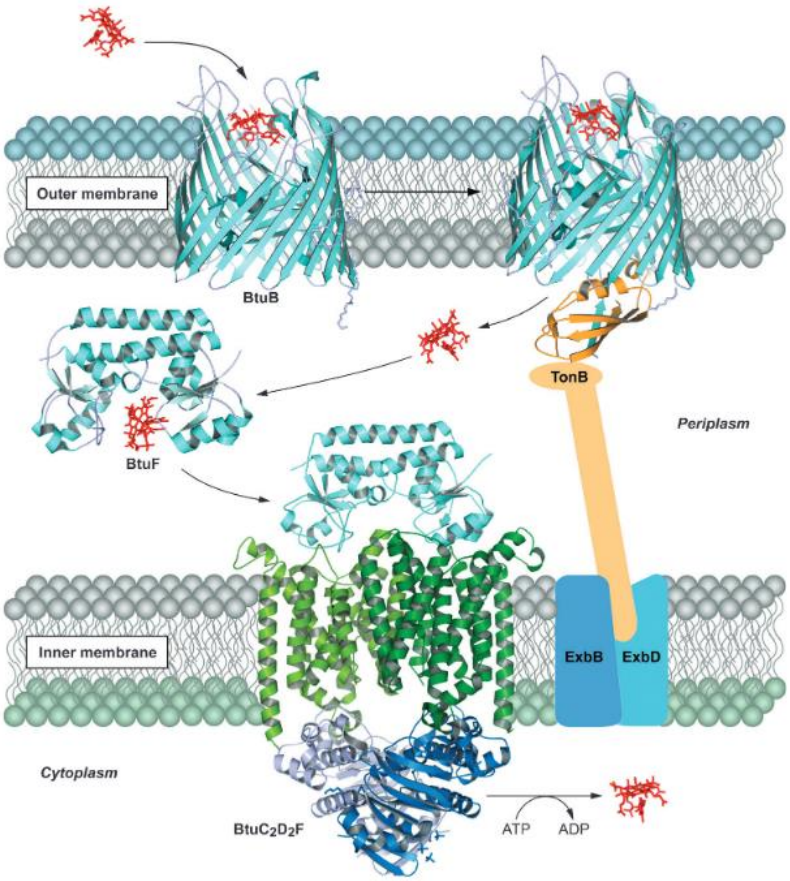


Figure 1.20 – Schematic of the Btu system of *E. coli*. Vitamin B_{12} is shown as red sticks, the structures of BtuB and BtuF (labelled) are shown in cyan cartoon representation, BtuC in green (different shades of green show the individual monomer subunits within the dimer), BtuD in blue (again, different shades show the monomeric units) and $TonB_{CTD}$ in complex with BtuB shown in yellow. Image taken from reference [148].

1.5 Study of force-induced remodelling of proteins and their complexes

One of the primary goals of this thesis is to test whether mechanical force applied by TonB is able to remodel the plug domain of a TonB-dependent transporter (in accordance with the pulling model), therefore in this section the application of force to study biological systems will be covered.

The first mechanical unfolding experiments were performed over 20 years ago on the giant muscle protein titin using atomic force microscopy (AFM)[149], since this study our understanding of protein and protein complex mechanical stability has improved by multiple methods developed to investigate biological forces. As mechanical remodelling is now known to play a crucial role in many biological processes such as enzymatic activation, protein degradation, ion channel opening, translocation, cellular adhesion,

signal transduction and so on, it is important to gain insight into the mechanisms at the molecular level *in vitro* if we want to understand the function *in vivo*. Three techniques that have been extensively used to explore the effects of force on proteins are AFM, optical tweezers and patch clamping, the former being extensively used in this thesis. Each technique has different force ranges and resolution which are summarised in Table 1.2

	AFM	**Optical Tweezers**	**Patch clamp**
Temporal resolution (ms)	1	0.1	0.001
Spatial resolution (nm)	0.1 (vertical)	0.1	50
Force range (pN)	1-10,000	0.1-100	-

Table 1.2 – The force range and resolution of techniques used to explore the role of mechanical force on biological molecules. Values taken from reference [150].

1.5.1 Atomic force microscopy (single-molecule force spectroscopy)

Due to the dynamic range of forces (1-10,000 pN), high positional accuracy (0.1 nm), high sensitivity (~10^{-2} pN) and operational compatibility under physiological conditions, AFM has been an invaluable and extensively employed technique to measure intra-molecular unfolding forces of individual proteins[151] and the unbinding of non-covalent molecular interactions[152]. A typical AFM instrument comprises of two major components, an optical head and a XY stage scanner. A thin (10 µm) flexible cantilever composed of gold coated silicon nitride with a sharp stylus is used as the force-sensitive probe which is attached to the AFM optical head. A diode laser is focussed onto the end of a flexible cantilever probe, and movement of the lever leads to the position of the reflected spot changing on a position sensitive photo-diode, and hence the deflection of the lever can be measured with picometer sensitivity (Figure 1.21). The known spring constant of the cantilever allows force to be measured according to Hooke's law (Equation 2.13). As the stiffness of the spring affects sensitivity, spring constants of 10-40 pN/nm are ideal for single-molecule force spectroscopy (SMFS) experiments of proteins. The AFM probe is positioned on the Z-axis by piezo expansion, which

approaches and retracts the cantilever to and from the surface at a constant velocity. The cantilever is pushed against the hard surface (Figure 1.21Bii) until a defined load is reached; the withdrawal velocity defines the amount of force which is loaded onto the molecule or complex of interest.

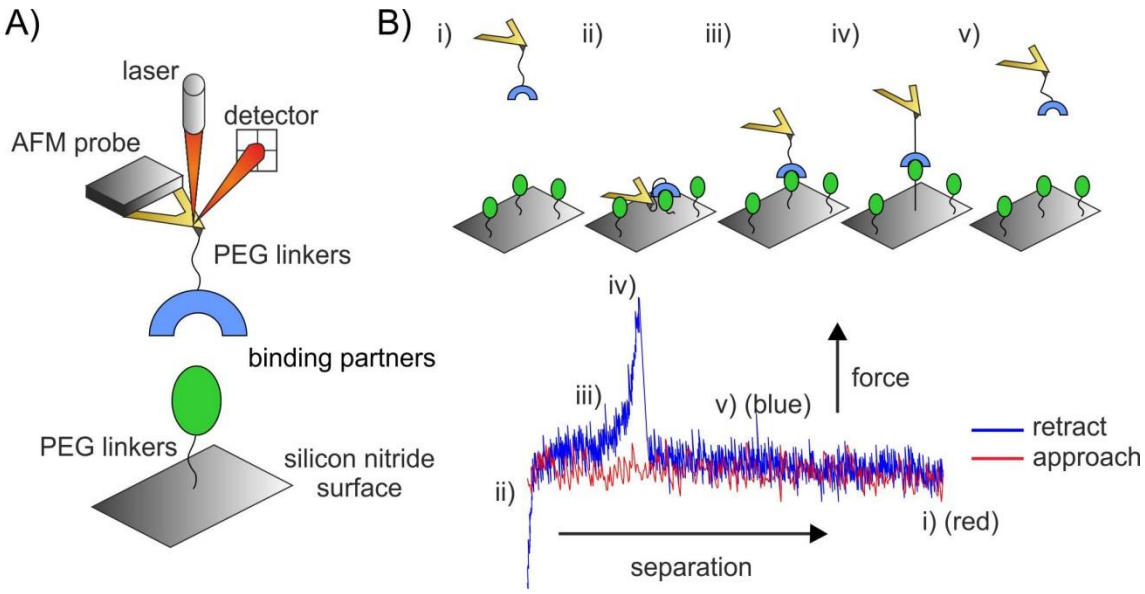


Figure 1.21 – Studying protein-protein interactions under force by single-molecule force spectroscopy. A) Schematic of the AFM set up for single-molecule force spectroscopy (SMFS) experiments to identify the strength of interacting partners such as a protein:protein interaction. The silicon nitride AFM probe contains a triangle flexible cantilever which is coated with a layer of reflective gold, linkers are used to covalently attach biomolecules to the cantilever. A laser is positioned onto the back of the cantilever which detects the deformation caused by the resisting force. B) Step by step of an ideal single-molecule unbinding event in a SMFS experiment between two binding partners and a force-extension trace. i) The cantilever is initially away from the surface, the piezo causes the cantilever to press against the surface (ii) which brings a single complex together, the complex is then extended by retracting the cantilever (iii) until a force (which bends the cantilever (iv)) that breaks the interacting partners (the rupture force) is reached. The cantilever returns back to its original position (v). This is called an approach-retract cycle.

For SMFS measurements of the mechanical strength of a protein:protein or protein:ligand interaction, the substrates are covalently attached to the AFM probe and surface by the use of polyethylene glycol (PEG) linkers derived with a maleimide at one end. The use of PEG linkers is also advantageous as it prevents non-specific protein sticking, enables substrate flexibility and has a signature resisting force when stretched[153]. The SMFS procedure involves bringing the AFM probe into contact with the surface to form a complex between the substrates (Figure 1.21B). The AFM probe is

retracted from the surface and if the interaction has formed, an entropic restoring force will be initially observed, which increases in a parabolic fashion until the interaction breaks and the force sharply drops to zero (force extension trace Figure 1.21B). The force required for dissociation (rupture force) is the force at the apex of the parabolic curve (Figure 1.21B). SMFS is a powerful technique to study protein:protein/ligand complex as it allows estimation of the unbinding forces (Table 1.3), dissociation rate constants (see Section 1.5.3), energy landscapes, the nature of force resistant bonds and the length of the complex before dissociation (by using the Worm-like chain (WLC) model (see Section 1.5.2)).

Molecular partners	**Pulling velocity (μm/s)**	**Unbinding force (pN)**
Avidin/biotin	5	173 ± 19
Streptavidin/biotin	5	326 ± 19
Human serum albumin(HSA)/anti HSA antibody (Ab)	0.2	244 ± 22
Intercellular adhesion molecule-1 (ICAM-1)/anti-ICAM-1 Ab	4.7	100 ± 50
Actin/Myosin	0.03	25 ± 1.4
Nitrilotriacetate(NTA)/hexa-histidine (His_6)	0.09-0.27	150-194
Lactose/lactose-binding immunoglobulin G (IgG)	1.2	36 ± 4
Recombinant P-selectin/P-selectin	2.8	~165

Table 1.3 – Examples of the strength of molecular complexes studied by SMFS, values taken from reference [152].

1.5.2 Worm-like chain model

The force extension data collected from SMFS experiments in this thesis are analysed using a model for polymer elasticity called the worm-like chain (WLC) model. The WLC model (Equation 1.1) was derived empirically from the stretching of DNA[154,155] and predicts that entropic restoring force (f) is generated upon extension (x) of a polymer:

$$f(x) = \frac{k_B T}{p}\left(\frac{1}{4}\left(1 - \frac{x}{L_c}\right)^{-2} - \frac{1}{4} + \frac{x}{L_c}\right) \quad (1.1)$$

where L_c is the contour length (the length of the biomolecule under study when fully extended), T is temperature, k_B is Boltzmann's constant and p is the persistence length (describes the decay in correlations of the tangent vectors of the WLC model (or the typical length scale over which a polymer bends)). The persistence length for a protein is the length between two α-carbons of two amino acids (0.4 nm)[156].

The WLC model has been extensively used in the study of protein domain unfolding[84,149,157], and also used for protein-protein interactions[158,159] as the force-extension profiles are well defined by the model. Farrance *et al* demonstrated that the WLC model gave accurate measurements of the contour length of various protein:protein/peptide complexes when extended by AFM[159]. The contour length is important for identifying whether an interaction between two binding partners is specific, as the expected L_c can be estimated from structural information. Contour lengths greater than the expected value give good indication of protein remodelling under force. It is also possible to calculate the amount of protein unfolded by subtracting observed L_c from the expected L_c and by using disulfide cross-links to prevent unfolding. Other than finding the L_c, a differential WLC model (Equation 2.14) can be used to calculate the instantaneous loading rate (the rate of force applied to the complex) which is strongly affected by the L_c of the complex under study. The differential WLC calculates the slope (WLC_{slope}) which is the change in force per change in distance (ie. the stiffness), which when multiplied by the retraction velocity, allows the loading rate (r_f) for a specific single event to be calculated.

1.5.3 Dynamic force spectroscopy

For most interactions, force acts to decrease the stability of the folded or bound state of a protein or complex relative to the unfolded or dissociated state. The unfolding or unbinding of a protein or complex can be considered as a two-state process, a low-energy conformation which represents the folded/bound state and a high-energy unfolded/unbound state. Between the states is an energy barrier that has to be

overcome to switch between states. A simplified example of a 'free-energy landscape' is shown by the black trace in Figure 1.22. Based on Bell's model for off-rates[160], it has been demonstrated by Evans and Ritchie[161,162] that an externally applied force lowers the energy barrier between the low- and high-energy states (Figure 1.22). Additionally the force required to unfold a protein or break a protein complex depends on the loading rate, and in general the higher the velocity, more force is required to unfold/break the complex, therefore a linear dependence is usually observed when the most probable force (F*) is plot as a function of ln(r_f) (See inset Figure 1.22 and Equation 1.2)[163]:

$$F* = \frac{k_B T}{x_u} \ln \frac{r_f x_u}{k_B T \, k_0} \tag{1.2}$$

where k_B is Boltzmann's constant, T is temperature, x_u is the distance from the low energy state to the transition state and k_0 is the spontaneous unfolding when no force is applied.

Pulling a protein or a protein complex at various velocities (dynamic force spectroscopy), allows the quantification of parameters that are used to describe the free-energy landscape (energy barrier height and width), such as x_u and the free-energy of activation (ΔG, the height of the transition state) which determines the rate at which the protein spontaneously unfolds (k_0) or complex dissociates (k_{off}^{0F}) under no applied force ($k_0 = 1/t_0 \exp(-\Delta G/k_B T)$). t_0 is the diffuse relaxation time.

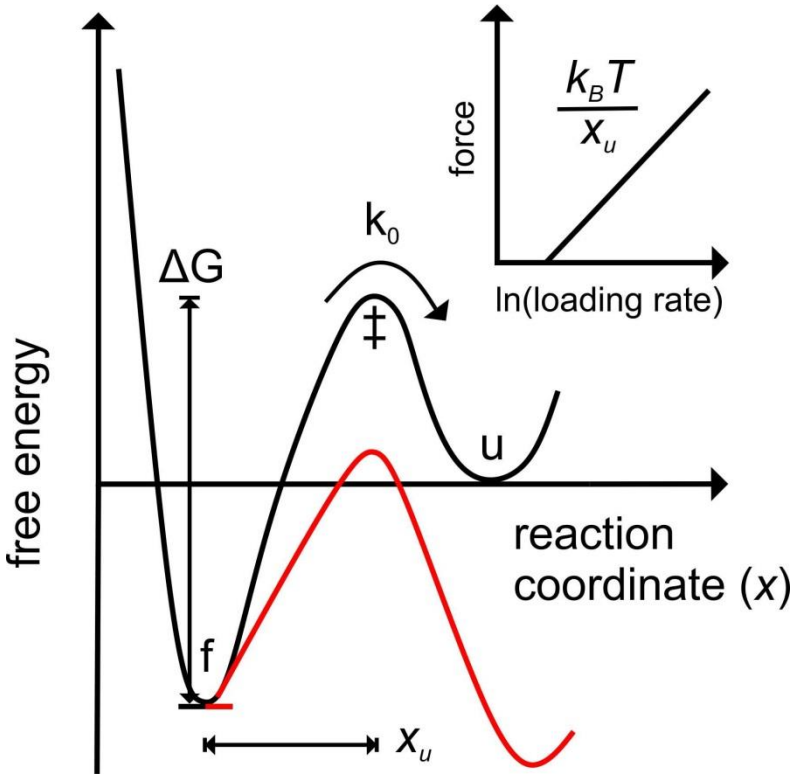


Figure 1.22 – Simple free-energy landscape of a folded protein under externally applied force. In this two-state unfolding model, a single barrier separates the folded (f) and unfolded (u) states of a protein (black trace). The activation free-energy of unfolding is given by ΔG, while x_u represents the distance between the folded and the transition state (‡) along the reaction coordinate *x*. The energy barrier is spontaneously crossed at a transition rate k_0 (or in the case of a protein ligand complex unbinding (k_{off}^{0F})). Application of an external force tilts the energy landscape (red trace), lowering the energy barrier. The inset depicts a theoretical dependence of the most probable rupture force on the loading rate: the dynamic force spectrum, which is governed by a single linear regime, with a slope proportional to $1/x_u$. The y-intercept allows the extrapolation of the k_0 or k_{off}^{0F}.

The application of force to a complex tilts the energy landscape reducing the free-energy barrier to unbinding of the complex (or indeed unfolding of the protein). So the probability of the complex breaking apart or protein unfolding increases with the application of force. In the classic Bell–Evans model, it is assumed that the energy barrier is so deep that its position does not change during tilting, where only the height of the energy barrier is lowered by force, the shape of the landscape does not change (red trace, Figure 1.22)[164].

The Bell model describes interactions that show an exponential decrease in lifetime with increased force (slip bonds), however not all biological interactions display this

behaviour, in some cases more than one energy barrier can exist[164] and some interactions are strengthened by force (catch bonds). A slip bond is an interaction that dissociates at a higher rate when subjected to force, which is commonly observed for protein:protein interactions. In catch bonds, the interaction lifetime increases with force (mechanical activation adhesion), which was initially observed in fluid flow experiments where cellular adhesion was enhanced by shear flow[164]. The first direct demonstration of catch bond behaviour was by SMFS on P-selectin, the bond lifetime was found to be longer under low forces, but acted as a slip bond at high forces[165]. Catch bonds are predominantly found in cellular adhesion proteins. Trip bonds are found in tight protein interactions which have very large dissociation lifetimes under zero force (> days), however upon the application of a small force, the binding partners rapidly dissociate, exemplified by the colicin immunity protein complex, Im9:E9[158].

1.5.4 Protein structure and mechanical stability

As all proteins have evolved to provide a specific function to a cell, it was therefore likely that the fold (secondary structure) of a protein with a mechanical role will influence its ability to either withstand force or be susceptible to force. This question was addressed soon after the first application of SMFS to study titin unfolding in 1997[149]. The conclusion being that α-helical proteins are force sensitive[166], β-sheet proteins force resistant and mixed topology proteins show various responses to force. The array of non-covalent interactions that form β-sheeted proteins provide much more stability under force compared to the hydrophobic contacts of α-helices which unfold in a step-wise manner.

The mechanical strength of a protein is not only dependent on the type of secondary structure. The direction of the force application relative to the topology of the secondary structure has also been found to show a significant effect. In a study by Brockwell *et al,* a β-stranded protein E2lip3 was shown to resist mechanical deformation when force was applied parallel to the β-strands of the protein (187 ± 10 pN at 700 nms^{-1}) however if the geometry of the force was applied perpendicular to the strands, the unfolding of the protein became undetectable by AFM (< 15 pN)[136]. Similar observations have been found for ubiquitin[137] and GFP[167]. The reason for this effect can be rationalised by an

analogy to a zip, sequential 'unzipping' of the hydrogen bonds requires little force whereas breaking all the bonds simultaneously requires much more force.

Proteins with parallel β-strands that are directly bonded are more mechanically robust than antiparallel β-strands. Brockwell *et al* provided evidence of this mechanical durability on a small β-stranded protein with simple topology and parallel terminal β-strands[168]. It was shown that although this protein has no known mechanical function, the parallel strands where the force was directly propagated rendered the protein to be mechanically strong when extended by AFM (152 pN, 700 nms^{-1}). Other structural features of proteins which confer mechanical strength include hydrophobic packing, solvent accessibility of the hydrogen bonds, bond patterns and sequence motifs[169].

1.6 Thesis aims

The principal aim of the work described in this thesis is to provide molecular insight into the mechanism of TonB dependent transporters and their force-induced remodelling both *in vitro* and *in vivo*.

As discussed in this chapter (Section 1.4.6.3), force induced remodelling of the plug domain of TBDTs by the Ton box tether is an attractive model for substrate transport. Whilst this model is well received, there is no direct experimental evidence *in vitro* or *in vivo* supporting the pulling tether model and more generally, for other membrane proteins predicted to function by a tether gating mechanism[170,171].

Therefore the work in this thesis aims to provide evidence for:

1. Membrane protein gating by proteinaceous tethers
2. Periplasmic inside-out energy transduction mechanisms
3. Whether the extent of mechanical plug domain unfolding is dependent on the substrate size

Alongside this, biophysical methods will also be used to address structural features of the periplasmic spanning protein TonB, especially the poorly charactered proline-rich linker domain.

The first results chapter provides an extensive analysis of TonB, including measurement of the mechanical strength and bond properties of the TonB:Ton box complex. Structural

and biophysical features of the proline-rich domain will also be explored, such as the effects on binding, dimerisation and stability of the protein.

After demonstrating the mechanical durabily of the TonB:Ton box interaction, the second results chapter uses SMFS to validate that TonB is capable of partially unfolding the plug domain of two TBDTs, BtuB and FhuA. These findings are supported by cross-linking experiments and *in silico* methods.

Finally, the physiological consequences of partial plug domain unfolding are explored *in vivo*. A variant of BtuB is engineered to allow complete plug unfolding and using antibiotic sensitivity assays it is revealed that TonB is capable of unfolding the entire plug in this variant.

2 Materials and method

2.1 Materials

2.1.1 Centrifuges

Bench top centrifuge: GenFuge 24D (Progen Scientific, London, UK).

Centrifuge: Avanti J25 (Beckman Coulter, CA, USA).

Ultracentrifuge: Beckman Optima XL-90 (Beckman Coulter, CA, USA).

2.1.2 Incubators

Stuart SI500 Shaking Incubator (Bibby Scientific, Staffordshire, UK).

New Brunswick™ Innova 44 (Eppendorf, Hamburg, Germany).

2.1.3 Protein purification equipment

ÄKTAprime plus (GE healthcare, Buckinghamshire, UK).

Resource™ S 6 mL cation exchange column (GE healthcare, Buckinghamshire, UK).

Superdex™ 75 Hiload 26/60 gel filtration column (GE healthcare, Buckinghamshire, UK).

Superdex™ 75 HR 10/30 gel filtration column (GE healthcare, Buckinghamshire, UK).

Superdex™ peptide HR 10/30 gel filtration column (GE healthcare, Buckinghamshire, UK).

HiTrap® DEAE FF 5 mL anion exchange column (GE healthcare, Buckinghamshire, UK).

Superdex™ 200 increase 3.2/300 gel filtration column (GE healthcare, Buckinghamshire, UK).

2.1.4 Spectrophotometers

NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Scientific, MA, USA).

Ultrospec 2100 pro UV/Visible spectrophotometer (GE Healthcare, Buckinghamshire, UK).

2.1.5 PCR thermocycler

T100 thermal cycler (BioRad, CA, USA).

2.1.6 AFM

MFP-3D™ Stand Alone AFM (Asylum Research, Buckinghamshire, UK).

2.1.7 AFM probes

MLCT silicon nitride with reflective gold AFM probe (Bruker, CA, USA).

2.1.8 Chemicals

A	**Supplier**
Acetic acid, glacial	Fisher Scientific, Loughborough, UK
Acrylamide 30 % (v/v)	Severn Biotech, Kidderminster, UK
Agar	Fisher Scientific, Loughborough, UK
Agarose	Melford Laboratories, Suffolk, UK
Ampicillin sodium	Formedium, Norfolk, UK
L-(+)-Arabinose	Sigma Life Sciences, MO, USA
Arginine	
(3-Aminopropyl)triethoxysilane (APTES)	Sigma Life Sciences, MO, USA
Ammonium persulphate (APS)	Sigma Life Sciences, MO, USA
Alexa Fluor® 488 C_5 maleimide	Fisher Scientific, Loughborough, UK
B	
Benzamidine dihydrocholride	Sigma Life Sciences, MO, USA
Bacitracin zinc	Bio Basic Inc, Ontario, Canada
C	
Cyanocobalamin (vitamin B_{12})	Sigma Life Sciences, MO, USA
Chloroform	Fisher Scientific, Loughborough, UK
Calcium chloride ($CaCl_2$)	Melford Laboratories, Suffolk, UK

D	
Disodium phosphate	Fisher Scientific, Loughborough, UK
Dithiothreitol (DTT)	Formedium, Norfolk, UK
Dimethyl sulfoxide (DMSO)	GE Healthcare, Buckinghamshire, UK
E	
Ethidium bromide	Sigma Life Sciences, MO, USA
Ethanol	Sigma Life Sciences, MO, USA
Ethylenediaminetetraacetic acid (EDTA)	Fisher Scientific, Loughborough, UK
F	
Ferrichrome iron	EMC Microcollections, Tübingen, Germany
G	
Guanidine hydrochloride	Sigma Life Sciences, MO, USA
Glycerol	Fisher Scientific, Loughborough, UK
H	
Hydrogen peroxide (H_2O_2)	Sigma Life Sciences, MO, USA
Hydrochloric acid (HCl)	Fisher Scientific, Loughborough, UK
I	
Imidazole	Sigma Life Sciences, MO, USA
Isopropanol	Fisher Scientific, Loughborough, UK
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Melford Laboratories, Suffolk, UK
L	
Lithium 3,5-diiodosalicylate (LIS)	Sigma Life Sciences, MO, USA
Lithium chloride (LiCl)	Sigma Life Sciences, MO, USA
M	
Monosodium phosphate	Fisher Scientific, Loughborough, UK
Methanol	Fisher Scientific, Loughborough, UK
N	
N-Hydroxysuccinimide -PEG_{24}-maleimide (SM PEG)	Fisher Scientific, Loughborough, UK

N-Hydroxysuccinimide -PEG_{24}-methyl (MS PEG)	Fisher Scientific, Loughborough, UK
N,N-diisopropylethylamine (DIPEA)	Fisher Scientific, Loughborough, UK
N-octyl-β-D-glucoside (β-OG)	Avanti Polar Lipids Inc, Delfzijl, NL
Nickel nitrilotriacetic acid (Ni-NTA)	GE Healthcare, Buckinghamshire, UK
P	
Phenylmethanesufonyl fluoride (PMSF)	Sigma Life Sciences, MO, USA
Potassium chloride (KCl)	Fisher Scientific, Loughborough, UK
S	
Sodium azide (NaN_3)	Sigma Life Sciences, MO, USA
Sodium chloride (NaCl)	Fisher Scientific, Loughborough, UK
Sulphuric acid	Fisher Scientific, Loughborough, UK
Sodium hydroxide (NaOH)	Fisher Scientific, Loughborough, UK
Sodium dodecyl sulphate (SDS)	Severn Biotech, Kidderminster, UK
Streptomycin sulfate	Sigma Life Sciences, MO, USA
T	
Triton X-100	Sigma Life Sciences, MO, USA
Tris	Fisher Scientific, Loughborough, UK
Tetramethylethylenediamine (TEMED)	Sigma Life Sciences, MO, USA
U	
Urea	MP biomedicals, Loughborough , UK

2.1.9 Gel ladders and dyes

1 kb, 100 bp DNA ladder (Promega Hampshire, UK).

Precision plus protein dual color standards protein ladder (BioRad, CA, USA).

Blue/orange 6x loading dye (Promega Hampshire, UK).

2.1.10 Kits

Wizard® Plus SV Minipreps DNA Purification System (Promega Hampshire, UK).

Q5® High-Fidelity DNA Polymerase kit (NEB, Hertfordshire, UK).

QIAquick Gel Extraction Kit (Qiagen, Manchester, UK).

pGEM®-T Easy Vector System (Promega, Hampshire, UK).

Q5® Site-Directed Mutagenesis Kit (NEB, Hertfordshire, UK).

2.1.11 Buffers

2.1.11.1 Imidazole wash buffer

20 mM Tris-HCl pH 8.0, 300 mM NaCl, 20 mM imidazole, 2 mM DTT, 0.025 % (w/v) sodium azide, 1 mM PMSF, 2 mM benzamidine

2.1.11.2 Imidazole elution buffer

20 mM Tris-HCl pH 8.0, 300 mM NaCl, 250 mM imidazole, 2 mM DTT, 0.025 % (w/v) sodium azide, 1 mM PMSF, 2 mM benzamidine.

2.1.11.3 SDS-PAGE resolving gel buffer

1 M Tris-HCl pH 8.45, 0.1 % (w/v) SDS, 15 % (v/v) acrylamide, 13 % (v/v) glycerol, 0.7 % (w/v) APS, 0.07 % (v/v) TEMED.

2.1.11.4 SDS-PAGE stacking gel buffer

750 mM Tris-HCl pH 8.45, 0.07 % (w/v) SDS, 4 % (v/v) acrylamide, 0.32 % (w/v) APS, 0.16 % (v/v) TEMED.

2.1.11.5 SDS-PAGE loading buffer (x2 concentrated stock)

2 % (w/v) SDS, 10 % (v/v) glycerol, 0.1 % bromophenol blue, 100 mM DTT.

2.1.11.6 SDS-PAGE cathode buffer

1 M Tris, 1 M Tricine, 1 % (w/v) SDS.

2.1.11.7 SDS-PAGE anode buffer

2 M Tris-HCl pH 8.9.

2.1.11.8 TAE buffer (x25 concentrated stock)

1 M Tris, 25 mM EDTA, 2.8 % (v/v) glacial acetic acid.

2.1.11.9 1.5 % agarose gel

40 mM Tris, 1 mM EDTA, 0.1 % (v/v) glacial acetic acid, 1.5 % (w/v) agarose, 2.5 µM ethidium bromide.

2.1.11.10 BtuB purification buffers

Buffer A - 10 mM Tris-HCl pH 8.0, 0.25 % (w/v) lithium diiodosalicylic acid (LIS)

Buffer B - 10 mM Tris-HCl pH 8.0, 0.25 % (w/v) LIS, 2 % (v/v) Triton X-100

Buffer C - 10 mM Tris-HCl pH 8.0

Buffer D - 10 mM Tris-HCl pH 8.0, 1 % (w/v) n-octyl-β-D-glucopyranoside (β-OG), 5 mM EDTA

Buffer E - 50 mM Tris-HCl pH 7.5, 5 mM EDTA, 0.54 % (w/v) β-OG

Buffer F - 50 mM Tris-HCl pH 7.5, 5 mM EDTA, 1 M LiCl, 0.54 % (w/v) β-OG

2.1.11.11 Liposome detergent free dialysis buffer

10 mM Tris-HCl pH 7.5, 300 mM KCl, 400 µM $CaCl_2$, 0.01% (w/v) NaN_3

2.1.12 Enzymes

Restiction enzymes (*XhoI*, *NdeI* and *NcoI*), quick ligase, alkaline phosphatase, Q5® and Taq polymerases were purchased from NEB, Hertfordshire, UK. Lysozyme and DNAse I were purchased from Sigma Life Sciences, MO, USA.

2.1.13 Lipids

E. coli polar lipid extract (Avanti Polar Lipids Inc, Delfzijl, Netherlands).

2.1.14 Peptides/DNA oligios

All DNA primers were purchased from Eurofins MWG Operon, Ebersberg, Germany. Peptide oligos were purchased from Genscript, NJ, USA.

2.1.15 *E. coli* strains

E. coli XL1-Blue Competent cells (efficiency: >1 x 10^8 cfu/µg) (Agilent technologies, Berkshire, UK).

E. coli BL21(DE3) Competent cells (efficiency: >1 x 10^6 cfu/µg) (Agilent technologies, Berkshire, UK).

E. coli DH5-α Competent cells (NEB, Hertfordshire, UK).

E. coli JM109 cells (Stratagene, Cambridge, UK).

E. coli RK5016 (MC4100 *btub metE70 argH recA*) cells were kindly provided by Nicholas Housden, University of Oxford, UK.

E. coli TNE012 (K12 $tsx^{-}ompA^{-}ompB^{-}$) were kindly provided by Nicholas Housden, University of Oxford, UK.

2.1.16 Plasmids

pAG1 (wild-type BtuB from *E. coli* under a native promotor, pUC8 vector) was kindly provided by Nicholas Housden, University of Oxford, UK.

pNGH15 (wild-type BtuB from *E. coli* in a pBAD expression vector) was kindly provided by Nicholas Housden, University of Oxford, UK.

pSH11 (wild-type FhuA from E. coli in a pBAD expression vector).

His_6-$TonB_{\Delta TMD}$ pET-23a(+) (*E. coli* His_6-tagged TonB without the transmembrane domain (residues 1-33) in an expression vector).

His_6-$TonB_{CTD}$ pET-23a(+) (*E. coli* His_6-tagged TonB C-terminal domain (residues 144-239) in an expression vector).

His_6-$TonB_{\Delta TMD}$ (V32C) pET-23a(+) (*E. coli* His_6-tagged TonB without the transmembrane domain (residues 1-33) with N-terminally located cysteine in an expression vector).

His_6-$TonB_{CTD}$ (K144C) pET-23a(+) (*E. coli* His_6-tagged TonB C-terminal domain (residues 144-239) with N-terminally located cysteine in an expression vector).

2.2 Methods

2.2.1 Molecular biology

2.2.1.1 Preparation of competent *E. coli* for transformation

10 mL of LB medium was inoculated with a desired strain of *E. coli*; this was grown at 37 ˚C, 200 rpm for 16 hours. 5 mL of the culture was used to inoculate 100 mL of LB medium and was grown to an optical density at 600 nm (OD_{600}) of 0.45 before the cells were harvested by centrifugation at 4,000 rpm for 10 minutes at 4 ˚C (Eppendorf 5804R Refrigerated Benchtop Centrifuge). The supernatant was discarded and the pellet gently resuspended in 10 mL of sterile pre-chilled (4 ˚C) 100 mM $CaCl_2$. The resuspension was incubated on ice for 10 minutes before centrifugation to re-pellet the cells. The supernatant was discarded and the pellet was gently resuspended in 2 mL of pre-chilled (4 ˚C) 100 mM $CaCl_2$, 30 % (w/v) glycerol and 100 µL aliquots were prepared in sterile Eppendorf tubes sat on dry ice. Once frozen, the competent cells were stored at -80 ˚C.

2.2.1.2 Extraction of *E. coli* genomic DNA

The genome of *E. coli* was extracted by inoculating 5 mL of LB broth with a subculture of *E. coli* cells; the culture was then incubated at 200 rpm, 37 ˚C until an OD_{600} of 0.25 was recorded. 20 µL of the culture was transferred to 980 µL of sterile purified H_2O and boiled at 98 ˚C for 5 minutes.

2.2.1.3 Q5® High-Fidelity PCR

For Q5® PCR the following reaction was set up in 0.2 mL PCR tubes:

	Volume (µL)
Q5® 5X master mix	10.0
dNTP (10 mM)	1.0
Forward primer (10 µM)	2.5
Reverse primer (10 µM)	2.5
Template DNA	10.0 (~1 ng-1 µg)
Nuclease-free water	23.5
Q5® polymerase	0.5

This was then subjected to the following thermal cycle:

Temperature (°C)	Duration (s)	
98	30	
98	5-10	
Primer T_m	30	X 35
72	30	
72	120	
4	∞	

2.2.1.4 Vent PCR

For Vent PCR the following reaction was set up in 0.2 mL PCR tubes:

	Volume (µL)
Template DNA (50 ng/µL)	2
Forward primer (10 µM)	2
Reverse primer (10 µM)	2
Nuclease-free water	To final 100
$MgSO_4$ (100 mM)	0, 2 or 4
DMSO	1
Vent polymerase	0.5
4 X PCR master mix*	25

* ThermoPol Reaction Buffer (4X), 10 mM dNTP

This was then subjected to the following thermal cycle:

Temperature (°C)	Duration (s)	
95	30	
95	30	
Primer T_m	30	X 20
72	120/kb	
72	600	
4	∞	

The primer melting temperatures (T_m) were calculated using the NEB online T_m calculator (www.neb.com/tools-and-resources/interactive-tools/tm-calculator).

2.2.1.5 Restriction digest

	Volume (µL)
DNA (~300 ng/µL)	6
CutSmart™ buffer (10X)	2
NdeI	1
XhoI	1
Nuclease-free water	10

The reaction mixture was incubated at 37 °C for 2 hours and enzymes were removed by 1.5 % agarose gel electrophoresis of the products.

2.2.1.6 Site-directed mutagenesis

Site-directed mutagenesis was performed using the reagents and protocol from the Q5® Site-Directed Mutagenesis Kit (Section 2.1.10). Primers were designed using the NEBasechanger software (http://nebasechanger.neb.com/) and purchased from Eurofins MWG Operon. The PCR set up was identical to the Q5® PCR (Section 2.2.1.3) however only 25 cycles were performed rather than 35. 1 µL of the PCR product was added to 5 µL of 2 x Kinase, Ligase & DpnI (KLD) buffer, 1 µL of 10 x KLD enzyme mix and 3 µL of purified H_2O and incubated for 5-10 minutes at 21 °C. 5 µL of the reaction product was transformed into DH5α cells by heat shock (42 °C for 30 seconds). This was then plated onto antibiotic selection agar plates after an hour incubation at 37 °C, 200 rpm in SOC medium provided with the kit.

The mutagenesis primers for BtuB are listed in Table 2.1:

Primer name:	**Forward sequence:**	**Reverse sequence:**
BtuB L8P	5’-CCC GGA TAC TCC CGT CGT TAC TG-3’	5’-CTG GTA TCC TGT GCC CAA-3’
BtuB L23C	5’-CAG CAC TGT GTG TGC ACC AAC CAC C-3’	5’-CGC GGC TGT TCA AAA CGG-3’
BtuB S374C	5’-CTG GCA AAC CTG CGC CGG TTG-3’	5’-GTT CCA TGA CGA CCA AAC TGT G-3’
BtuB V29C	5’-AAC CAC CGT TTG CAC CCG TCA GGA TAT CGA C-3’	5’-GGT GCA AGC ACA GTG CTG-3’
BtuB V45C	5’-GGT CAA TGA TTG CCT GCG CCG TCT TCC GGG-3’	5’-GAG GTC GAC TGC CAG CGG-3’
BtuB T27C	5’-TGC ACC AAC CTG CGT TGT GAC CCG TCA GG-3’	5’-AGC ACA GTG CTG CGC GGC-3’
BtuB Y109C	5’-CGT GTT GAA TGT ATC CGT GGG-3’	5’-CTG GAC AAG CGC AAT AGG-3’
BtuB I80A/L85A	5'-GTA CGC GCG AAT CTG GCG GGG GTG AGT-3'	5'-GCC ATC AGC TAA CAC CAA CAC ATG ACT GGC-3'
BtuB L96A	5'-TTC TGC CGA CGC TAG CCA GTT CCC TAT TGC G-3'	5'-CCA CTC ACC CCC GCC AGA-3'

Table 2.1 – Primers for BtuB mutagenesis

2.2.1.7 Preparation of plasmid DNA for sequencing

Single colonies were picked from antibiotic selection agar plates and used to inoculate 10 mL of LB medium with the same selection antibiotic as the agar plate (100 µg/mL ampicillin). The inoculated medium was incubated for 16 hours at 37 °C, 200 rpm and the cells were subsequently harvested by centrifugation at 4,000 rpm for 10 minutes, 4 ˚C (Eppendorf 5804R Refrigerated Benchtop Centrifuge). Using a Wizard® Plus SV Minipreps DNA Purification System kit (Section 2.1.10), the plasmids were purified from the bacteria and the concentration in water was calculated using a nanodrop 1000 spectrophotometer using the optical density at 260 nm (A_{260}) (concentration (µg/mL) = 50 µg/mL x A_{260}). 15 µL of the plasmid DNA (50 ng/µL) was then prepared and sent for sequencing (Beckman Coulter Genomics).

2.2.1.8 Agarose gel electrophoresis

1 L of 1x TAE buffer (Section 2.1.11.8) was prepared, 150 mL was transferred to a conical flask with 2.25 g agarose (1.5 % gel) and heated in a microwave until the agarose was fully dissolved into the TAE buffer. This was allowed to cool for 10 minutes before 15 µL of 10 mg/mL ethidium bromide was added, mixed and poured into a sealed gel mould with a comb inserted. Once the gel had set, the seals were removed and the gel mould containing the gel was submerged into an electrophoresis bath filled with the remaining 1x TAE buffer. The comb was removed and 20 µL DNA samples were loaded with 4 µL of Blue/orange 6x loading dye. 10 µL of 1 kb or 100 bp DNA ladder was applied to a single well for size (bp) determination. The gel was electrophoresed at 100 mV for ~45 minutes and DNA stained with ethidium bromide and visualised with a UV transilluminator.

2.2.1.9 Extraction of DNA from agarose gel

Agarose gel electrophoresis was used to resolve DNA species after PCR and restriction digests. In order to purify the target DNA and other DNA fragments from the gel itself, DNA bands were extracted using a scalpel and a QIAquick Gel Extraction Kit (Section 2.1.10) was used to purify the DNA. The DNA concentration was then determined using a nanodrop 1000 spectrophotometer (Section 2.2.1.7).

2.2.2 Protein expression and purification

2.2.2.1 Starter culture

200 mL of sterile LB medium containing 0.1 % (w/v) ampicillin was inoculated with a single colony from an ampicillin selection plate of freshly transformed expression cells (BL21 (DE3) for TonB constructs and TNE012 (K12 *tsx*$^{-}$*ompA*$^{-}$*ompB*$^{-}$) for BtuB and FhuA). The inoculated medium was incubated at 37 °C, 200 rpm for 16 hours.

2.2.2.2 Protein expression procedure for TonB constructs

1 L (x10) of sterile LB medium containing 0.1 % (w/v) ampicillin was inoculated with 10 mL of starter culture and grown at 37 °C, 200 rpm to an OD_{600} of 0.7. Protein expression was then induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) to give a final concentration of 1 mM. After 16 hours incubation, cells were harvested via centrifugation at 15,000 rpm (4 °C) using a contifuge (Contifuge® stratos) or 4,000 rpm (4 °C) for 20 minutes using a Beckman Avanti J-26 XP centrifuge with a JLA-8.1 fixed-angle rotor. The cell pellet was either frozen for storage (-80 °C) in a plastic zip bag or directly resuspended (10 g/50 mL) into pre-cooled (4 °C) imidazole wash buffer (Section 2.1.11.1).

2.2.2.3 Protein expression procedure for TonB-dependent receptors

1 L (x10) of LB medium containing 0.1 % (w/v) ampicillin was inoculated with 10 mL of starter culture of TNE012 cells transformed with pNGH15 plasmid (Section 2.1.16), or derivatives of the plasmid expressing either mutants of BtuB or FhuA. This was grown at 37 °C, 200 rpm to an OD_{600} of 0.7 and protein expression was induced with 0.15 % (w/v) L-(+)-arabinose. After inducing protein expression, the temperature was decreased to 15 °C and the culture was incubated for a further 16 hours with shaking at 200 rpm. The cells were harvested by centrifugation at 15,000 rpm 4 °C using a contifuge or 4,000 rpm in a Beckman Avanti J-26 XP centrifuge with a JLA-8.1 fixed-angle rotor. The cell pellet was either frozen for storage (-80 °C) in a plastic zip bag or directly resuspended BtuB purification buffer A (Section 2.1.11.10) at 4 °C.

2.2.2.4 Cell lysis after protein expression

Cells were disrupted by sonication (Sonics Vibra-cell, 6 mm Microtip, Amp 70 %) on ice for 10 sec on / 20 sec off for a total process time of 4 minutes. After cell disruption, the cell debris was removed via centrifugation for 20 minutes at 10,000 rpm, 4 °C (Beckman Avanti J-26 XP centrifuge with a JA-25.50 Fixed-Angle Rotor). The supernatant was collected for protein purification, whilst the pellet was subject to subsequent rounds of sonication which increased the yield of protein significantly.

2.2.2.5 Purification of His_6-TonB constructs by nickel affinity

The following steps were all carried out at 4 ˚C and in the presence of 1 mM PMSF and 2 mM benzamidine. The supernatant from cell lysis was mixed with 6 mL Ni-NTA resin (GE Healthcare) in a 50 mL Falcon tube pre-equilibrated with imidazole wash buffer (Section 2.1.11.1). The His_6-tagged protein was bound to the resin by inverting to suspend the resin in the cleared cell lysate and incubated on rollers for 20 minutes at 4 °C. The protein-bound resin was pelleted by centrifugation at 4,000 rpm, 4 °C for 5 minutes (Beckman Avanti J-26 XP centrifuge with a JS-5.3 swinging bucket rotor), the supernatant was discarded and 40 mL of imidazole wash buffer was added, mixed by inverting and centrifuged at 4,000 rpm, 4 °C for 5 minutes. This wash process was repeated three times and was followed by two elution steps using imidazole elution buffer (Section 2.1.11.2) giving a total elution volume of ~30 mL. The imidazole was removed by dialysis into pre-cooled 25 mM Tris-HCl (pH 8.0), 128 mM NaCl, 1 mM PMSF, 2 mM benzamidine using a 3.5K MWCO SnakeSkin dialysis tubing (Thermo), followed by centrifugation (4,000 rpm, 4 ˚C) and sterile filtration (0.22 µm filter) to remove any insoluble aggregate or resin. Further purification was performed with gel filtration using a Superdex 75 26/60 gel filtration column (GE Healthcare)($TonB_{CTD}$ constructs) equilibrated with 25 mM Tris-HCl pH 7.5, 128 mM NaCl, 1 mM PMSF, 2 mM benzamidine. For $TonB_{\Delta TMD}$ constructs, anion exchange chromatography with 6 mL Resource™ S column (GE Healthcare) equilibrated with 25 mM Tris-HCl pH 7.5, 128 mM NaCl, 1 mM PMSF, 2 mM benzamidine was used and the protein was eluted with a linear 0-0.5 M NaCl gradient over 16 column volumes. The fractions were pooled and concentrated to 1 mg/mL (after dialysis into 25 mM Tris-HCl pH 7.5, 128 mM NaCl) with a vivaspin column

with an MWCO of 5 kDa ($TonB_{CTD}$ constructs) and 10 kDa ($TonB_{\Delta TMD}$ constructs). The purified proteins were snap frozen in liquid nitrogen and stored at −20 °C.

2.2.2.6 Purification of FhuA, BtuB and BtuB mutants

The supernatant from cell lysis (2.2.2.4) was centrifuged at 45,000 rpm for 1 hour, 4 °C using a Beckman L-70 ultracentrifuge with a Beckman 45 Ti Fixed-Angle rotor. The pellet (total membrane fraction) was retained and the supernatant discarded. The membrane pellet was resuspended in 56 mL BtuB purification buffer B (Section 2.1.11.10), homogenised by hand then ultracentrifuged at 45,000 rpm for 1 hour, 4 °C with the same rotor. The supernatant (containing extracted inner membrane proteins) was discarded and this step was repeated. The pellet was then resuspended in 56 mL BtuB purification buffer buffer C (Section 2.1.11.10), homogenised by hand, ultracentrifuged at 45,000 rpm for 1 hour, 4 °C and then the supernatant discarded. (This step decreases the amount of Triton X-100 carried over to the next stages of purification). The pellet was then resuspended in 56 mL BtuB purification buffer D (Section 2.1.11.10), homogenised and ultracentrifuged at 45,000 rpm for 1 hour at 4 °C. The supernatant from this step contained the extracted outer membrane proteins.

A disposable 5 mL DEAE-Sepharose column was equilibrated in 90 % BtuB purification buffer E and 10 % buffer F (Section 2.1.11.10) at room temperature using an ÄKTA prime plus. Variation of the detergent by 0.04 % at this stage had adverse effects on the purification; therefore was crucial. The supernatant containing the outer-membrane proteins was loaded onto the column and washed with 40 mL 90 % buffer E, 10 % buffer F. A gradient of 90% buffer E /10% buffer F to 50% buffer E / 50 % buffer F was then run over the next 65 mL. A further 30 mL of 100 % buffer F was run through the column which causes the target protein to elute from the column. The chromatogram from the DEAE-Sepharose column gave two major absorbance peaks. The first peak eluted from the column as soon as the gradient started and contained the majority of the impurities. The second peak eluted upon switching to 100 % BtuB purification buffer F. This peak contained high purity BtuB that precipitated immediumtely after elution. To enhance precipitation, the eluted fractions were stored at 4 °C for 16 hours. The precipitate was then pelleted by centrifugation for 20 minutes at 4,000 rpm, 4 °C.

For FhuA purification, the protein eluted with impurities at the DEAE-sepharose stage. Consequently, the FhuA precipitate was pelleted by centrifugation, dissolved into 1 % (w/v) β-OG, and further purified using a Superdex™ 75 HR 10/30 column pre-equilibrated with 25 mM Tris-HCl pH 8.0, 128 mM NaCl, 1 % (w/v) β-OG. The major peak contained pure FhuA and was used for the SMFS experiments.

2.2.2.7 Growth and purification of ^{15}N-labelled TonB for NMR

A 100 mL overnight culture in LB medium with selection antibiotic (100 µg/mL ampicillin) was set up from a single colony of freshly transformed BL21 (DE3) with the pET-23a(+) expression plasmid containing the TonB construct. 7.5 mL of the overnight culture was used to inoculate 2 L (x8) conical flasks containing pre-warmed (37 ˚C) 500 mL (x 8) of M9 minimal medium:

Component	g/L for 5 x stock
Na_2HPO_4 (anhydrous)	34
KH_2PO_4	15
NaCl	2.5
$^{15}NH_4Cl$	5

A 5 X stock of the M9 minimal medium salts was initially prepared and used to make 500 mL of 1 X medium in 2 L conical flasks which was autoclaved before use.

Before the medium was inoculated with the overnight culture, it was supplemented with sterile filtered 0.5 mL 100 mg/mL ampicillin, 1 mL 1 M $MgSO_4$, 10 mL 20 % (w/v) D-glucose and 50 µL 1 M $CaCl_2$.

The cells were grown to mid-log phase (OD_{600} = 0.7) before 0.5 mL of 1 M IPTG was added to each flask to induce protein expression. The induced cells were then incubated for a further 16 hours at 37 ˚C, 200 rpm before the cells were harvested by centrifugation at 4000 rpm, 4 ˚C and stored at -80 ˚C. The ^{15}N-labelled TonB constructs were purified as described in Section 2.2.2.5, the protease inhibitor cocktail (1 mM pmsf, 2 mM bezamidine) was present throughout the entire purification and the buffers used

were all pre-cooled to 4 ˚C. The total yield from 4 L was ~50 mg/30mg for $TonB_{\Delta TMD}$ and $TonB_{CTD}$ respectively.

2.2.2.8 Quantification of protein concentration

To calculate protein concentrations after purification, the absorption at 280 nm was measured (using a NanoDrop 2000 UV-Vis Spectrophotometer) and the estimated extinction coefficient (M^{-1} cm^{-1}) from the protein sequence (ExPASy online software Protparam) was used. Readings were typically taken in duplicate and the average used.

2.2.3 Biochemistry techniques

2.2.3.1 SDS-PAGE

SDS resolving and stacking gel solutions (Section 2.1.11.3 and 2.1.11.4) were freshly prepared and APS and TEMED were added directly before pouring the gel mixture into a sealed clean casting chamber. Once the resolving gel had set, the stacking gel was added and a comb inserted into the top of the casting chamber. Excess stacking gel was removed using a Pasteur pipette. The rubber seals were removed once the gel had set and the casting chamber containing the gel was inserted into the cathode chamber of an electrophoresis cell, which was then filled with 1x cathode buffer (Section 2.1.11.6). The anode chamber was then filled with 1x anode buffer (Section 2.1.11.7). 10 µL of the protein sample was mixed in a 1:1 ratio with 2x SDS loading buffer (Section 2.1.11.5) and boiled for 5 minutes. 20 µL of this solution was loaded into the wells of the stacking gel. 5 µL of Dual Xtra Recombinant Protein Standards protein ladder was also loaded into a single well for molecular weight determination. The electrodes were then connected to a power supply and 35 mV per gel was applied until the stained sample rested at the surface of the resolving gel. The voltage was then increased to 65 mV per gel and was electrophoresed until the stain left the bottom of the resolving gel. The gel was removed from the casting chamber and submerged into Instantblue coomassie stain (Expedeon) and incubated on a rocking table for ~2 hours before samples became visible for analysis.

2.2.4 Surface functionalisation (for single-molecule force spectroscopy)

2.2.4.1 Oxidisation of silicon nitride AFM probe and surface

To oxidise the surfaces of the silicon nitride AFM probe and surfaces (1 cm^2 cut from a silicon nitride disc (Rockwood electronic material)), piranha solution (3:1 0.5 M (>95 %) H_2SO_4 to 30 % (v/v) H_2O_2) was prepared and surfaces were submerged for 5 minutes whilst the solution was still hot (and most active). The piranha solution was removed from the surfaces by washing with purified H_2O, and dried with N_2. The AFM probes were submerged in the freshly prepared piranha solution for 30 seconds, washed with purified H_2O then dried with N_2. The surfaces and AFM probes were then placed on a microscope slide inside a petri dish with a hole in the lid and placed under a UV lamp (UVIlite, UVItec) set to 254 nm for 30 minutes (ozone cleaning).

2.2.4.2 Aminosilanisation of silicon nitride

Oxidised AFM probes and surfaces were placed into a desiccator along with 80 µL of (3-aminopropyl)triethoxysilane (APTES) and 20 µL of N,-N-diisopropylethylamine (DIPEA) held in separate 1.5 mL Eppendorf tube lids. The desiccator was evacuated using a vacuum pump for 1 minute and left to incubate at room temperature for 2 hours. After the incubation, the APTES and DIPEA solutions were removed and the desiccator was flooded with N_2 and left to cure for 48 hours.

2.2.4.3 Attachment of NHS-PEG_{24}-maleimide linkers

Amino-silanised AFM probes and surfaces were immersed in ~1 mL chloroform containing 20 µL of 250 mM N-hydroxysuccinimide-PEG_{24}-maleimide (SM PEG) linkers (in DMSO) and left to incubate at room temperature for 1 hour. Both AFM probe and surface were then washed with chloroform and dried with N_2. In the proteoliposome experiments, a 1:9 ratio of SM PEG: N-hydroxysuccinimide-PEG_{24}-methyl (MS PEG) was used to reduce the frequency of multiple binding events, by reducing the amount of available maleimide groups on the surface for the protein to react with.

2.2.4.4 Protein attachment

Protein (0.5-1.0 mg/mL) or oligopeptide ligand (1.0 mg/mL) both containing an engineered cysteine residue were deposited over the surfaces or AFM probe and left to incubate in a covered container for 30 minutes at room temperature. Unreacted protein/oligopeptide was then washed from the surface and AFM probe with the reaction buffer (25 mM Tris-HCl pH 8.0, 128 mM NaCl).

2.2.5 Microbiology techniques

2.2.5.1 Davis minimal agar plates for vitamin B_{12} uptake assay

A sterile concentrated stock of Davis and Mingioli minimal medium salts was added to a volumetric flask, and made to volume with purified H_2O (the total volume of the supplements (antibiotics, glucose, amino acids and vitamins) was subtracted from the total volume). Agar was then added and autoclaved together with the minimal salts. The autoclaved minimal salts and agar were cooled to 50 ˚C in a water bath for 10 minutes, and then concentrated stocks of the supplements were added to give a final concentration shown in Table 2.2. Once all the supplements were applied, the medium was poured into sterile petri dishes next to a flame and allowed time to set.

Component	Concentration (µg/mL)
Thiamine	1
Arginine	5
Glucose	2500
Ampicillin	100
Streptomycin	25
Vitamin B_{12}	0.1 – 100 nM
Methionine (Control plates only)	5

Table 2.2 - Davis and Mingioli minimal medium plate supplements for the methionine assay

2.2.5.2 Phenotype growth assay

To assess the effect mutation of BtuB upon substrate transport, the plasmid pAG1 (btuB (or mutant) under a native promotor) was transformed into competent RK5016

(MC4100 *btub metE70 argH recA*) *E. coli* cells by heat shock at 42 ˚C for 30 seconds, and plated onto LB selection plates supplemented with 100 µg/mL ampicillin, 25 µg/mL streptomycin and 5 µg/mL arginine. After 42 hours of incubation at 37 ˚C, a single colony was selected and grown for 16 hours in 5 mL LB (containing the same concentration of ampicillin, streptomycin and arginine), then 0.5 mL was mixed with 0.5 mL of 50 % sterile glycerol and snap frozen in liquid nitrogen. Using a sterile pipette tip, a bead was taken from the glycerol stock and plated onto LB selection plates using a sterile loop. After 42 hours of incubation at 37 ˚C, large colonies (~2 mm diameter) were picked with a sterile loop and spread onto Davis minimal medium agar supplemented with 100 µg/mL ampicillin, 25 µg/mL streptomycin, 5 µg/mL arginine, 1 µg/mL thiamine, various concentrations of vitamin B_{12} (0.1-100 nM) or 5 µg/mL methionine as a control. The Davis minimal plates were incubated at 37 ˚C for 42 hours before colony size within the final streaked area was analysed. Colonies that were >2 mm in diameter were considered to have robust growth and smaller colonies were considered to have partial or reduced growth.

2.2.5.3 Bacitracin sensitivity growth assay (plate reader method)

Single colonies of RK5016 cells transformed with pAG1 containing either wild-type *btuB* or its variants were used to inoculate 96 well plate wells containing 200 µL LB supplemented with ampicillin (100 µg/mL), streptomycin (25 µg/mL), 50, 100, 150 or 200 µg/mL bacitracin and for a control where TonB activity is prevented, 200 µg/mL Ton box pentapeptide (ETVIV) was added to the medium. The plates (Corning Costar® assay plate, 96 well, black with clear flat bottom, non-treated, 1 cm path length) were incubated at 37 °C with 200 rpm shaking for 24 hours with an OD_{600} reading every 4 minutes (CLARIOstar® high performance monochromator multimode microplate reader (BMG LABTECH)).

2.2.6 Biophysical techniques

2.2.6.1 Circular dichroism (CD)

To gain secondary structural information of the purified proteins, Far UV (190-260 nm) circular dichroism (CD) was performed in a 1 mm path length curvette (Hellma) using a Jasco J715 CD spectropolarimeter.

200 µL of 0.05 mg/mL protein solution was placed into the cuvette and incubated at 25 ˚C before a far-UV CD spectrum was acquired using a 1 nm bandwidth, 1 nm timestep and an average of 2 scans (190-260 nm) was taken per sample.

BtuB, BtuB mutants and FhuA were resuspended in 15 mM NaH_2PO_4 pH 7.5, 1 % (w/v) β-OG and the protein concentration adjusted to 0.05 mg/mL.

2.2.6.2 Fluorescence spectroscopy (equilibrium denaturation)

For the TonB constructs, a 9 M urea, 25 mM Tris-HCl pH 7.5, 128 mM NaCl stock solution was prepared along with a 1 mg/mL solution of the protein being analysed in 25 mM Tris-HCl pH 7.5, 128 mM NaCl (± 0.5 M additional NaCl). Five stocks of 100 µg/mL protein containing 0, 2, 4, 6 and 8 M urea were prepared and used to create 40 1 mL solutions increasing in urea concentration by 0.2 M (between 0-8 M). The samples were then equilibrated at 15 °C for 16 hours in a circulating water bath (NesLab) before analysis. A Photon Technology International (PTI) fluorimeter was used to monitor the intrinsic tryptophan emission fluorescence when excited at 280 nm. In order to identify the emission wavelength that gave the largest change in fluorescence intensity between the 0 and 8 M urea protein samples, an emission scan was carried out by exciting at 280 nm and scanning the emission between 290-400 nm with the slit width set to 4 nm. The protein samples were placed in a 1 mL quartz cuvette with a 1 cm path length and were sampled at 15 °C. The largest emission change was used for a time-drive scan. Each of the 40 samples were excited at 280 nm and the average emission intensity was recorded for 60 seconds with 1 second intervals. The average signal was calculated for each denaturant concentration and plot as a function of urea concentration. A two state transition chemical denaturant model (Equation 2.1) was fit to the data using Igor Pro 6.32A (wavemetrics) software

$$S_{obs} = \frac{\left(\frac{(a[D] + b)\exp\left(\Delta G_{UN} - M_{UN}[D]\right)}{RT + (c[D] + d)}\right)}{1 + \exp\left(\Delta G_{UN} - \frac{M_{UN}[D]}{RT}\right)} \quad (2.1)$$

where S_{obs} is the observed signal *a* and *c* are the native and unfolded baseline gradients respectively, *b* and *d* are the native and unfolded y-intercept respectively, *[D]* is the denaturant concentration (molarity), *R* is the ideal gas constant (0.0083 kJ K^{-1} mol^{-1}), T is the temperature in Kelvin (288.15 K), ΔG_{UN} is the urea-induced unfolding free-energy (kJ mol^{-1}) and M_{UN} is the denaturant dependence of ΔG_{UN}(correlates with the amount of protein surface exposed to solvent upon unfolding) (kJ mol^{-1} M^{-1}). The observed signal and fit was normalised to fraction of natively folded protein (f_N) using Equation 2.2:

$$f_N = \frac{S_{obs} - (c[D] + d)}{(a[D] + b) - (c[D] + d)} \quad (2.2)$$

To assess the tertiary fold of OMPs (BtuB, its variants and FhuA), the OMP concentration was adjusted to 0.005 mg/mL in 1 % (w/v) β-OG, 25 mM Tris-HCl pH 8.0, 128 mM NaCl. A separate protein solution was created where the buffer contained 8 M urea. Each sample was allowed an hour at room temperature to equilibrate. An emission scan was carried out by exciting at 280 nm and scanning the emission between 290-400 nm with a slit width of 4 nm. The protein samples were placed in a 1 mL quartz cuvette with a 1 cm path length and were sampled at 15 °C.

2.2.6.3 Dynamic light scattering (DLS) of proteoliposomes

The proteoliposomes from the dialysis (Section 2.2.8) were suspended by pipetting and a 2 µL sample was added to 198 µL of filtered and degassed 10 mM Tris-HCl pH 7.5, 300 mM KCl, 400 µM $CaCl_2$, to give a final lipid concentration of approximately 40 µM. Using a Minidawn™ Treos® (Wyatt) detector in batch mode, 1 mL of 10 mM Tris-HCl pH 7.5, 300 mM KCl, 400 µM $CaCl_2$ buffer was injected into the flow cell (prewashed with 0.1 M nitric acid followed by purified H_2O) through a 0.22 µm syringe filter. A baseline in the DLS with fluctuations within four decimal places was recorded for 5 minutes before the

sample of proteoliposomes was injected into the flow cell. Light scattering of the proteoliposomes was recorded for 5 minutes before the buffer was injected back into the flow cell for a post-sample baseline reading. The data was analysed using the Astra 6.1 software (Wyatt Technology).

The correlation function for a monodisperse sample is given by:

$$g^{(2)}(\tau) = B + \beta \exp(-2\Gamma\tau) \tag{2.3}$$

where *B* is the baseline of the correlation function at infinite delay, β is the correlation function amplitude at zero delay, and Γ is the decay rate. Γ can be converted to the diffusion constant *D* by the relation:

$$D = \frac{\Gamma}{q^2} \tag{2.4}$$

where *q* is the magnitude of the scattering vector. The diffusion coefficient can be used to calculate the hydrodynamic radius (R_H) according to the Stokes-Einstein equation:

$$D = \frac{k_B T}{6\pi\eta R_H} \tag{2.5}$$

where *T* is the temperature of the system k_B is Boltzmann's constant, η is the viscosity of the medium and D the diffusion constant.

2.2.6.4 Size exclusion chromatography – multi-angle laser light scattering (SEC-MALLS)

Size exclusion chromatography in combination with multi-angle laser light scattering (SEC-MALLS) was performed at Diamond light source beamline B21. A 2.4 mL Superdex™ 200 increase 3.2/300 column connected to an Agilent 1200 HPLC system was equilibrated with 25 mM Tris-HCl pH 8.0, 128 mM NaCl. 45 µL of 464 µM $TonB_{\Delta TMD}$ was

injected onto the column at flow rate of 0.075 ml/min. The MALLS instrument used was an 18 angle DAWN HELEOS II with an OPTILab rEX refractive index monitor (Wyatt Technology). The MALLS data was processed using the ASTRA 6.1 software (Wyatt Technology).

The molecular weight (M_w) of a protein from the light scattering can be found by the Zimm expression[172]:

$$\frac{K^*c}{R(\theta,c)} = \frac{1}{M_w P(\theta)} + 2A_2c \quad (2.6)$$

where *K** is an experimental constant dependent on the square of the solvent refractive index increment and the inverse fourth power of the incident wavelength, c is the solute concentration, R is the Rayleigh excess ratio of the solution (function of scattering angle (θ) and solute concentration), M_w is the average solute molecular weight, *P*(θ) describes the angular dependence of the scattered light and A_2 is the second virial coefficient in the virial expansion of osmotic pressure.

2.2.6.5 Microscale thermophoresis (MST)

MST was carried out using a Monolith NT.115 series instrument (NanoTemper) using Alexa Fluor® 488 as the fluorescent probe and a blue LED filter (excitation 460–480 nm and emission 515–530 nm). The fluorescently labelled Ton Box_{BtuB} peptide (Section 2.2.9) was adjusted to a final concentration of 500 nM (optimal fluorescence at 250 nM). TonB constructs were concentrated to 1 mM using a 3 kDa MWCO Vivaspin column and centrifugation at 15,000 g. The concentration was determined by ultraviolet absorbance spectroscopy at 280 nm and an extinction coefficient of 5,500 M^{-1} cm^{-1} (calculated using the ExPASy ProtParam tool). A serial dilution over 16 concentrations was set up for TonB in 25 mM Tris-HCl pH 7.5, 128 mM NaCl. This dilution series was mixed with equal parts of the 500 nM Ton Box_{BtuB}-Alexa fluor® 488 and incubated at room temperature for 5 minutes. The samples were then loaded into capillaries (K002 MonolithTM NT.115 standard treated capillaries) and analysed with the LED power set

to 95 % and MST power (IR- laser) set to 40 %. The normalised fluorescence from T-jump and thermophoresis (F_{norm}):

$$F_{norm} = 1 + \left(\frac{\delta F}{\delta T} - S_T\right) \Delta T \tag{2.7}$$

(where δF/δT is the fluorescence change due to temperature increase, ΔT is the change in temperature and S_T is the Soret coefficient) was then plot against the $\log_{10}$ of the protein concentration and the data was fit with the Nanotemper analysis software using Equation 2.8 to yield the K_D [173]:

$$\begin{aligned} & S_{obs} \\ & = S_U + (S_B - S_U) \\ & \times \frac{([A_0] + [B_0] + K_D) \pm \sqrt{([A_0] + [B_0] + K_D)^2 - 4[A_0][B_0]}}{2[A_0]} \end{aligned} \tag{2.8}$$

where S_{obs} is the observed signal, S_U is the signal from unbound state, S_B is the signal from the bound state, $[A_0]$ and $[B_0]$ are the total concentrations of *A* and *B* respectively.

2.2.6.6 Small angle X-ray scattering (SAXS)

Size exclusion chromatography in combination with small-angle X-ray scattering (SEC-SAXS) experiments were performed at Diamond light source beamline B21. A 2.4 mL Superdex™ 200 increase 3.2/300 column connected to an Agilent 1200 HPLC system was equilibrated with 25 mM Tris-HCl pH 8.0, 128 mM NaCl (± 0.5 M NaCl). 55 µL of TonB construct (10 mg/mL) was loaded into a well of a 96 well plate; the amount loaded onto the column was 45 µL. A fixed wavelength of 1.0 Å (12.4 keV) was used with the X-ray detector (PILATUS 2M) placed 4 meters from the sample (suitable for particles with an R_g < 200 Å), the flow rate through the detector was 0.075 mL/min.

The buffer contributions were subtracted using the program SCATTER (version 3.0a) and a Guinier approximation (Equation 2.9) of the scattering vector ((ln[I(q)] vs q^2) where I(q) is the background corrected intensity and q is the momentum transfer ($q = 4\pi \sin(\theta)/$

λ, where 2θ is the angle between the incident and scattered beam and λ is the beam wavelength)) was used to estimate the radius of gyration (R_g) by limiting q x R_g < 1.3.

$$\ln I(q) = \ln I(0) - \frac{{R_g}^2}{3}.q^2 \quad (2.9)$$

The pair-distance distribution function (P(r)) (Equation 2.10) is an indirect fourier transform (due to a limited range of I(q)) of the scattering intensity). The distance distribution is a real space representation of the scattering data and allows the generation of a scattering envelope of the solute.

$$p(r) = \frac{r}{2\pi^2}\int_0^\infty I(q)q\sin(qr)\,dq \quad (2.10)$$

where: r is the distance vector. The maximum dimension (D_{max}) is the value of r at P(r) = 0[174].

2.2.6.7 Ensemble optimisation method (EOM)

To gain structural information of the intrinsically disordered domain of $TonB_{\Delta TMD}$ from the SAXS experimental data, the software EOM (version 2.0)[175] was used. EOM fits averaged theoretical scattering intensity from an ensemble of conformations into experimental SAXS data. A pool of 10,000 independent models based upon the sequence of $TonB_{\Delta TMD}$ (see Appendix) with structural information (PDB: 1XX3) was initially generated. As the proline-rich linker is predicted to be intrinsically disordered to some extent, completely random configurations of the alpha-carbons (residues 33-150, including the His_6-tag and N-terminal methionine) were created based upon the sequence provided. A genetic algorithm compares the averaged theoretical scattering intensity from the independent ensembles of conformations against the experimental SAXS data. The conformations that best describe the experimental data are selected.

2.2.6.8 Nuclear magnetic resonance (NMR) spectroscopy

All NMR experiments and data processing were performed by Dr Theo Karamanos (University of Leeds).

HSQC spectra of $TonB_{\Delta TMD}$ or its truncated variant ($TonB_{CTD}$) (both 10 mg/mL) were recorded on an AVANCE III Bruker spectrometer (600 MHz) equipped with a cryogenic probe in the presence of 25 mM Tris-HCl pH 8.0, 128 mM NaCl (± 0.5 M NaCl). 2048 and 256 complex points were detected in the direct and indirect dimension respectively. Spectra were processed in NMRPipe and analysed in CCPN analysis.

Chemical shift perturbations (CSP) for assigned residues were calculated by the formula[176]:

$$CSP = \sqrt{(5\Delta H)^2 + \Delta N^2} \tag{2.11}$$

where ΔH and ΔN are the chemical shift variations between the low and high salt conditions in the proton and nitrogen dimensions, respectively.

2.2.7 Atomic force microscopy (AFM)

2.2.7.1 Cantilever calibration

The AFM probe (MLCT with reflective gold, Bruker) with attached oligopeptide ligand or TonB construct was inserted into a cantilever holder and secured. The silicon surfaces with attached protein were attached to a microscope slide with Loctite superglue and secured to the XY scanner with magnetic bars. For the proteoliposome experiments, a 1 cm mica disc was attached to a metal disc.

A droplet of reaction buffer was applied to the silicon or mica surface and was held by surface tension. The AFM probe in the holder was mounted to the MFP-3D head (Asylum) and approached towards the surface until the probe was fully submerged in the buffer droplet. Using the built in optics, the laser was positioned to the tip of cantilever D (manufacturers spring constant: 30 pN/nm) and the deflection was set to zero using the PD disc. The cantilever was engaged using the Asylum Research software (MFP version 11), engagement causes Z-piezo voltage maxima (+150), an indication of

full Z-piezo extension and zero surface contact. Using the thumb wheel on the MFP-3D head (Asylum), the cantilever was approached to the surface. The z-voltage was adjusted to 70, assuring that the piezo was in the middle of its z range (7.5 µm). Spring constant calibration was carried out as a two-step procedure: determining the slope of contact from a force curve to find the sensitivity of the cantilever (nm/V) and then performing the non-destructive thermal tune to determine resonant frequency of the cantilever[177]. A single force-extension plot was recorded with the trigger (amount of deflection the cantilever undergoes before retraction) set to 20-40 nm, which gives a quantifiable deflection slope. The slope of the contact region (inverse optical lever sensitivity (InvOLS)) was measured by a linear fit. The cantilever was withdrawn from the surface and the deflection was set to 0. A thermal tune was carried out to detect the natural thermal fluctuation of the cantilever by performing ~50 frequency sweeps (0-1 MHz). The natural frequency (first major resonance peak) was selected and a Lorentzian function was fit. The area of the thermal fluctuations (P) is used to find the spring constant (*k*) using Equation 2.12:

$$k = \frac{K_B T}{P} \tag{2.12}$$

The spring constant was always within error of the manufacturer's guidelines before data collection. Hooke's Law (Equation 2.13) allows the calculation of the force applied (*F*) to the cantilever with a known spring constant by the deflection (or extension (*x*)).

$$F = kx \tag{2.13}$$

2.2.7.2 Data collection for Ton box peptide unbinding experiments

The feedback loop trigger point was set to 10 nm, the retraction distance was set to 600 nm, approach velocity was set to 2 μms^{-1} during data collection and kept constant whilst using various retraction velocities and sample rates (10 kHz per μms^{-1} velocity). The buffer on the sample surface was replenished frequently to prevent evaporation, which

would vary the NaCl concentration and dry the sample. Force maps of 20 μm^2 with 600 (100x6 array) approach-retract cycles were taken to maximise surface coverage, and between force maps the AFM probe was repositioned manually using the XY scanner. The retraction velocities used in this study were 200, 500, 1000, 3000 and 5000 nms^{-1}.

2.2.7.3 Data collection for proteoliposome experiments

The feedback loop trigger point was set to 3-4 nm with a 0.5-1 s surface dwell, the retraction distance was set to 0.5-0.6 µm, approach velocity was set to 2 μms^{-1} during data collection and kept constant whilst using various retraction velocities (200, 500, 1000 and 5000 nms^{-1}, (wild-type BtuB only)) and a sample rate of 10 kHz per μms^{-1} velocity. For the cross-linked BtuB variants, destabilised plug subdomain and FhuA the retraction velocity used was 1000 nms^{-1}. Force maps of 20 μm^2 with 500-1000 force-extensions were taken to maximise surface coverage.

2.2.7.4 Data processing

All force spectroscopy data was analysed using IGOR pro 6.32A with an Asylum Research extension (MFP3DXop v30).

The hard contact (0 nm) and baseline (0 pN) of all the force-extension retraction traces were manually set, by taking a section of the hard contact or the retraction with no events and setting the zero to the centre of the data (force baseline has a thermal noise of ± 10 pN). The WLC model (with a fixed persistence length of 0.4 nm) was fit manually to all unbinding events that are parabolic by inserting locks at the apex and the base of the curve and recording the contour length (L_c) and extension (x) (Figure 3.10).

Force-extension profiles were binned for analysis if: 1) the data fitted to the WLC model (single molecule events should display WLC like behaviour where the force-distance profile is not linear) and 2) the contour length was greater than the length of the PEG_{24} linkers used. For both the $TonB_{\Delta TMD}$:Ton Box_{BtuB} and $TonB_{CTD}$:Ton box_{BtuB} experiments, fewer than 0.1% of force curves showed more than one event. This indicates that there are no detectable unfolding events for either protein domain prior to dissociation of the bound complex. Single Gaussian distributions were fit to unbinding force and contour length histograms in order to determine the most probable force and contour length at

rupture for each retraction velocity investigated. For each pulling velocity used, data were collected in triplicate (using a freshly prepared cantilever for each repeat) to calculate the modal value and its average error.

Loading rates were calculated by fitting a WLC model to the rising edge of each unbinding profile when plotted as force versus tip-sample separation. The instantaneous gradient of this fit at rupture (WLC_{slope}) was calculated by inserting the derived contour length and extension at rupture into a differentiated form of the same equation (equation 2.14). The loading rate at rupture was then obtained by multiplying this value by the retraction velocity.

$$WLC_{slope} = \frac{k_B T}{p}\left(\frac{1}{2L_c\left(1-\frac{x}{L_c}\right)^3}\right)+\frac{1}{L_c} \tag{2.14}$$

where; *p* is the persistence length, L_c is the contour length, *x* is the extension, *kB* is the Boltzmann constant and T is the temperature.

The natural logarithm of the mean loading rate (Ns^{-1}) at each velocity was plot against the mean rupture force (N) which gives a linear relationship (dynamic force spectrum). The Bell-Evans model[162] (Equation 1.2) was rearranged to use the gradient of the linear fit to calculate the distance from the transition state (x_u) (Equation 2.15) and the y-intercept for the off rate at zero force (k_{off}^{0F}) (Equation 2.16).

$$x_u = \frac{K_B T}{m} \tag{2.15}$$

$$\frac{1}{k_{off}^{0F}} = \frac{K_B T}{x_u}\exp\left[\frac{c\,x_u}{K_B T}\right] \tag{2.16}$$

To calculate the error of k_{off}^{0F} and the x_u, manual bootstrapping was carried out, where a single data point was withdrawn from the triplicate results and a linear regression was

used to extrapolate the k_{off}^{0F} and x_u. This was performed for all of the data set, the average was taken for the values and the standard deviation was used for the error.

2.2.8 Formation of proteoliposomes and mica immobilisation

TonB-dependent receptors were inserted into a lipid bilayer at a high density by solubilising the precipitated receptor (0.1 mg) in 200 µL 25 mM Tris-HCl pH 7.5, 128 mM NaCl, 1 % (w/v) β-OG as previously described[178]. The lipids (0.2 mg *E. coli* polar lipid extract (Avanti)) were dried from a chloroform stock (10 mg/mL) using a nitrogen flow and solubilised with the protein and detergent by pepitting. The ligand for the receptor (cyanocobalamin (synthetic vitamin B_{12}) for BtuB and ferrichrome for FhuA) was added to the buffer to be in 100 molar excess of the receptor. The components were then placed into a 200 µL 12-14 kDa MWCO dialysis bag (D-tube dialyser mini, (Merck Millipore)) and dialysed against 300 mL of detergent-free adsorption buffer (Section 2.1.11.11) for 7 days at room temperature with daily buffer changes. To attach the proteoliposomes to mica, the concentration of lipid was diluted to 0.1 mg/mL in the detergent-free adsorption buffer and 150 µL was pooled over a 1 cm disc of freshly cleaved (using scotch tape) mica and incubated for 30 minutes at room temperature. After the incubation, the unbound proteoliposomes were thoroughly washed from the surface by pipetting 25 mM Tris-HCl pH 8.0, 128 mM NaCl over the mica.

2.2.9 Covalent labelling of proteins

The Ton box peptide from BtuB (sequence: **PDTLVVTANR**GSWSC, Ton box residues in bold) was fluorescently labelled with Alexa Fluor® 488 C_5 maleimide by creating a 1 mg/mL solution of the peptide in 25 mM Tris-HCl pH 7.5, 128 mM NaCl and titrating 100 µL of 10 mg/mL dye dissolved in DMSO. This was then stirred for 16 hours at room temperature in the dark. The reaction mixture was then loaded onto a pre-equilibrated Superdex™ peptide HR 10/30 gel filtration column equilibrated with 25 mM Tris-HCl pH 7.5, 128 mM NaCl. The peaks from the chromatogram were then analysed by liquid chromatography electrospray ionisation mass spectrometry (LC-ESI-MS).

2.2.10 Molecular dynamic simulations

All molecular dynamics (MD) simulations in this study were performed using NAMD (v2.10)[179] and the CHARMM27 force field[180]. The system was prepared using VMD (v1.9.2)[181]. 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) was used to model the lipid bilayer, the parameters for vitamin B_{12} were obtained from Gumbart *et al*[182], whereas the TIP3P force field was used for water[183]. The X-ray crystal structure of BtuB in complex with TonB and vitamin B_{12} (PDB 2GSK), including calcium ions and crystal waters, was inserted into the homogenous POPC lipid bilayer placed in the XY plane. The dimension of the simulation box was 100x100x300 Å. A total of 100211 molecules of TIP3P water, 149 Na^{2+} and 141 Ca^{2+} were then used to fill the simulation box. The total system comprised 342,843 atoms. The simulations were conducted using periodic boundary conditions. The bonds between hydrogen and heavy atoms were constrained with SHAKE[184]. The r-RESPA multiple time step method [185] was employed with 2 fs for bonded potentials, 2 fs for short-range non-bonded potentials and 4 fs for long-range non-bonded potentials[179]. Long-range electrostatic interactions were treated with the particle-mesh-Ewald (PME) method[186]. The distance cut off for non-bonded interactions was set to 10 Å, and a switching function was applied to smooth interactions between 9 and 10 Å. All simulations were conducted in the NPT ensemble. The temperature was set to 300 K and regulated via a Langevin thermostat[187]; the pressure was set to 1 atm and regulated via an isotropic Langevin piston manostat[188].

A multistage equilibration was used to relax the system. Initially all of the atoms except water and ions were fixed. 2000 steps of conjugate gradient geometry optimization were initially conducted. After energy minimization, the system was simulated for 1 ns to equilibrate the solvent around the protein and membrane. Next, the membrane was relaxed around the protein by removing the restraints on the fatty acid tails for 1 ns, and keeping the protein and vitamin B_{12} atoms restrained to their initial positions by applying an harmonic potential on the heavy atoms using a force constant of 1 kcal/(mol·Å^2). The protein and ligand was then gradually released over the next 3 ns by decreasing the force constant to 0.5 kcal/(mol· Å). Finally, all restraints were removed and the system was equilibrated for 10 ns. The final structure from the unrestrained MD trajectory was used as the starting point in the pulling simulations. The pulling was performed by using

the steered molecular simulation (SMD) module as implemented in NAMD[179]. In a SMD simulation, a selection of atoms are connected to an ideal spring with an elastic constant and pulled apart at a constant velocity. The SMD simulation was performed by restraining the phosphate head groups of the upper bilayer leaflet to its initial coordinates (to prevent the membrane being dragged) and extending a harmonic spring attached to either the Ton box of BtuB (Pro5 Cα) or the centre of mass of TonB (all Cα residues of TonB). The spring constant was set to 0.5 kcal/(mol·Å^2), while the pulling velocity was set to 2.5 Å/ns along the z-direction. The trajectories of the simulations were analysed using the software: wordom, VMD, MEMBplugin (v 1.1) and MDAnalysis.

The Root Mean Square Deviation (RMSD) is defined as:

$$RMSD = \sqrt{\sum_i^N \frac{\left(\left(x_i - x_{ref}\right)^2 + \left(y_i - y_{ref}\right)^2 + \left(z_i - z_{ref}\right)^2\right)}{N}} \qquad (2.17)$$

where *i* refers to the atom, x, y and z are the atom coordinates and N is the total number of atoms.

The Membrane Thickness tool of the MEMBplugin software measures the distance (Å) between two density peaks, formalised as the first and second central moment of the mass density profile of phosphorus atoms along the membrane[189]. The position of the membrane protein within the membrane (centre of mass (COM) difference) was calculated using a script kindly provided by Bob Schiffrin using the MDanalysis software. The centre of mass of the protein (BtuB) was calculated and the average centre of the membrane. The average difference between the two centres was then calculated over the trajectory.

3 Biophysical analysis of TonB and the influence of the proline-rich linker domain

3.1 Objectives

The primary objective of this thesis is to test the feasibly of the pulling model of TonB-dependent transport. The known biophysical properties of the periplasmic domain of TonB (residues 33-239) are limited, therefore in this chapter a thorough structural and biophysical characterisation is performed on TonB constructs with and without the proline-rich linker domain. Alongside this characterisation, the strength of the TonB:Ton box interaction is tested *in vitro* using single-molecule force spectroscopy, providing evidence that this non-covalent interaction mediated by parallel β-strand augmentation can withstand forces that can remodel folded proteins.

3.2 Results

3.2.1 Cloning and purification of TonB constructs

3.2.1.1 Cloning recombinant TonB from *E. coli*

PCR was used to amplify $TonB_{\Delta TMD}$ (TonB missing the transmembrane domain (residues 1-33)) and $TonB_{CTD}$ (the structured C-terminal domain of TonB (residues 147-239)) from the *tonB* gene of *E. coli* BL21(DE3) (Figure 3.1).

10	20	30	40	50
MTLDLPRRFP	WPTLLSVCIH	GAVVAGLLYT	SV**H**QVIELPA	PAQPISVTMV
60	70	80	90	100
TPADLEPPQA	VQPPPEPVVE	PEPEPEPIPE	PPKEAPVVIE	KPKPKPKPKP
110	120	130	140	150
KPVKKVQEQP	KRDVKPVESR	PASPFENTAP	ARLTSSTATA	ATSK**P**VTSVA
160	170	180	190	200
SGPRALSRNQ	PQYPARAQAL	RIEGQVKVKF	DVTPDGRVDN	VQILSAKPAN
210	220	230		
MFEREVKNAM	RRWRYEPGKP	GSGIVVNILF	KINGTTEIQ	

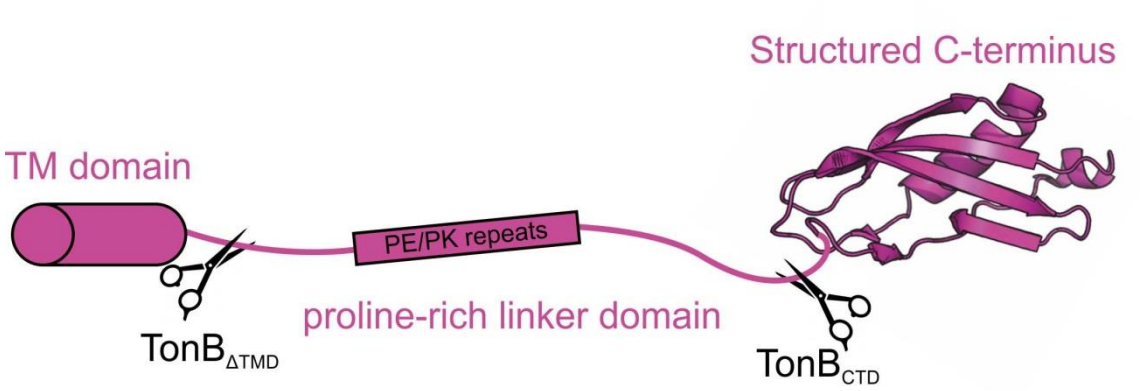


Figure 3.1 – Sequence and structural schematic of TonB from *E. coli*. The blue text indicates the TM domain, black text the proline-rich linker domain and pink text the structured C-terminal domain. The bold red text denotes the start position of each construct (TonB$_{\Delta TMD}$ H33 and TonB$_{CTD}$ P145). The schematic below shows the location of the truncation sites, TM domain (pink cylinder), the PE/PK repeats (pink box) and the NMR solution structure (PDB: 1XX3) of the structured C-terminal domain is shown by cartoon representation.

Primers were designed to include an N-terminally located cysteine for surface attachment for force spectroscopy experiments. Constructs containing cysteine are designated TonB$_{\Delta TMD}$(V32C) and TonB$_{CTD}$(K144C). The primers also included *NdeI* and *XhoI* restriction sites at 5' and 3' respectively and a His$_6$-tag at the N-terminus for purification. The primer sequences are shown in Table 3.1 and the protein sequences are shown in the Appendix (Section 7.1).

$TonB_{CTD}$(K144C)	Forward	$^{5'}$AATTTACAACATATGCATCACCATCACCATCACGGATGTCCGGT TACCAGTGTGGCTTCA$^{3'}$
	Reverse	$^{5'}$GGAATTTGACTCGAG**TTA***CTGAATTTCGGTGGTGCCGTTA*$^{3'}$
$TonB_{\Delta TMD}$(V32C)	Forward	$^{5'}$AATAATTAACATATGCATCACCATCACCATCACGGCTGTCATCA GGTTATTGAACTACCTGC$^{3'}$
	Reverse	$^{5'}$GGAATTTGACTCGAG**TTA***CTGAATTTCGGTGGTGCCGTTA*$^{3'}$

Table 3.1 - DNA Primers used for PCR gene amplification of TonB constructs for force-spectroscopy experiments. Underlined text indicates the restriction sites, green text is the start codon, blue text is the His_6-tag sequence, red text is the engineered cysteine and italic text is the sequence of *tonB*.

Q5® PCR was used for $TonB_{CTD}$(K144C) (Figure 3.2A) and Vent PCR was used for $TonB_{\Delta TMD}$(V32C) at three different $MgSO_4$ concentrations (Figure 3.2B). PCR products were analysed by a 1.5 % agarose gel stained with ethidium bromide to confirm successful amplification. $TonB_{\Delta TMD}$(V32C) was initially ligated into a pGEM-T Easy vector (Figure 3.2C), before insertion into a pET-23a(+) vector (Figure 3.2D). $TonB_{CTD}$(K144C), was directly ligated into a pET-23a(+) vector from the PCR amplification. The constructs were confirmed by DNA sequencing using T7P primer.

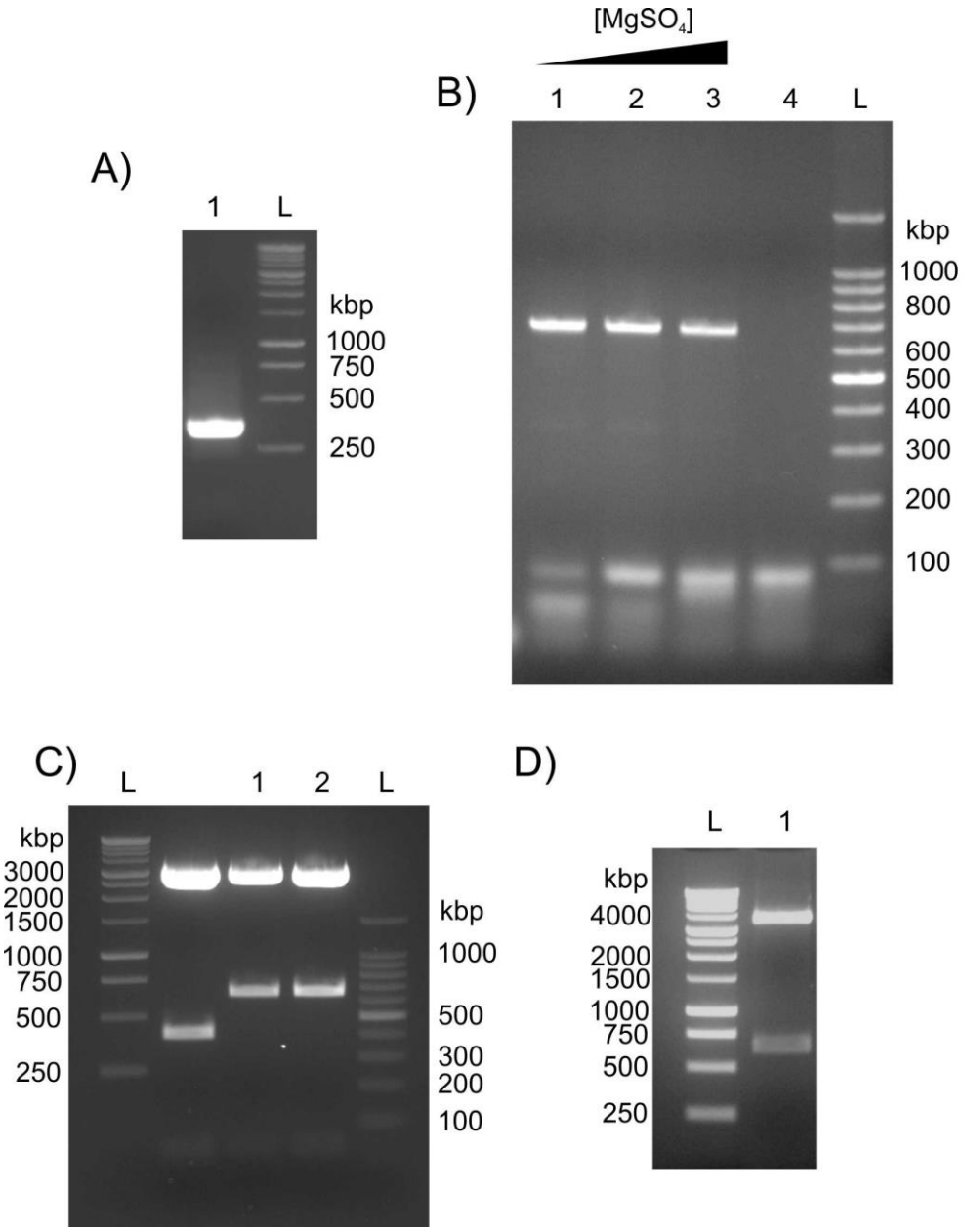


Figure 3.2 - Agarose gels (1.5 %) stained with ethidium bromide showing the stages of cloning of the *tonB* gene product from *E. coli* and insertion into pET-23a(+) expression vector. A) The amplification of $TonB_{CTD}$ (312 bp) by Q5® PCR, L is the ladder. B) The amplification of $TonB_{\Delta TMD}$(V32C) (678 bp) by Vent PCR under increasing concentrations of $MgSO_4$, lane 4 contains no $MgSO_4$. C) The cleavage of pGEM-T Easy vectors containing protein inserts. Lanes 1 and 2 contain pGEM-T Easy (3015 bp) with $TonB_{\Delta TMD}$(V32C) insert digested with NdeI/XhoI restriction enzymes. The unlabelled lane contains a construct not used in this thesis. D) The cleavage of pET-23a(+) vectors (3666 bp) with $TonB_{\Delta TMD}$(V32C) insert (678 bp). Lane 1 contains pET-23a(+) with $TonB_{\Delta TMD}$(V32C) digested by NdeI/XhoI restriction enzymes.

To prevent the engineered cysteines in the TonB constructs for AFM attachment forming covalent homodimers during biophysical characterisation, Q5® site-directed mutagenesis was subsequently used to delete the cysteine in both $TonB_{\Delta TMD}$(V32C) and

$TonB_{CTD}$(K144C) constructs. Table 3.2 shows the DNA primers designed for mutagenesis. The constructs without cysteine are designated $TonB_{\Delta TMD}$ and $TonB_{CTD}$. Site-directed mutagenesis was performed using the Q5® Site-Directed Mutagenesis Kit (Section 2.1.10) and the method described in Section 2.2.1.6.

$TonB_{\Delta TMD}$ Δcys	Forward	CATCAGGTTATTGAACTACCTG
	Reverse	GCCGTGATGGTGATGGTG
$TonB_{CTD}$ Δcys	Forward	CCGGTTACCAGTGTGGCT
	Reverse	TCCGTGATGGTGATGGTG

Table 3.2 – DNA mutagenesis primers for the removal of the engineered cysteine in the TonB constructs for AFM attachment.

3.2.1.2 Expression and purification of TonB constructs

TonB constructs (both with and without the N-terminally located cysteine) in pET-23a(+) were transformed into *E. coli* BL21 (DE3) cells and protein expression was carried out as described in Section 2.2.2.2. The proteins were purified using Ni-NTA batch purification as described in Section 2.2.2.5. At each step of the purification, a sample was taken for SDS-PAGE analysis (Figure 3.3A and C), which confirmed the presence of the TonB constructs. After removal of the imidazole by dialysis, centrifugation and filtration to remove any insoluble material, the $TonB_{\Delta TMD}$ construct (± cysteine) was further purified after Ni-NTA purification using a Resource™ S 6 mL cation exchange column (Figure 3.3B). The purity of the eluted protein was confirmed by SDS-PAGE (Figure 3.3A, lane 7). For the $TonB_{CTD}$ constructs (± cysteine), a Superdex™ 75 Hiload 26/60 gel filtration column was used for further purification (Figure 3.3D). The purity was assessed by SDS-PAGE (Figure 3.3E), confirming the successful isolation of $TonB_{CTD}$.

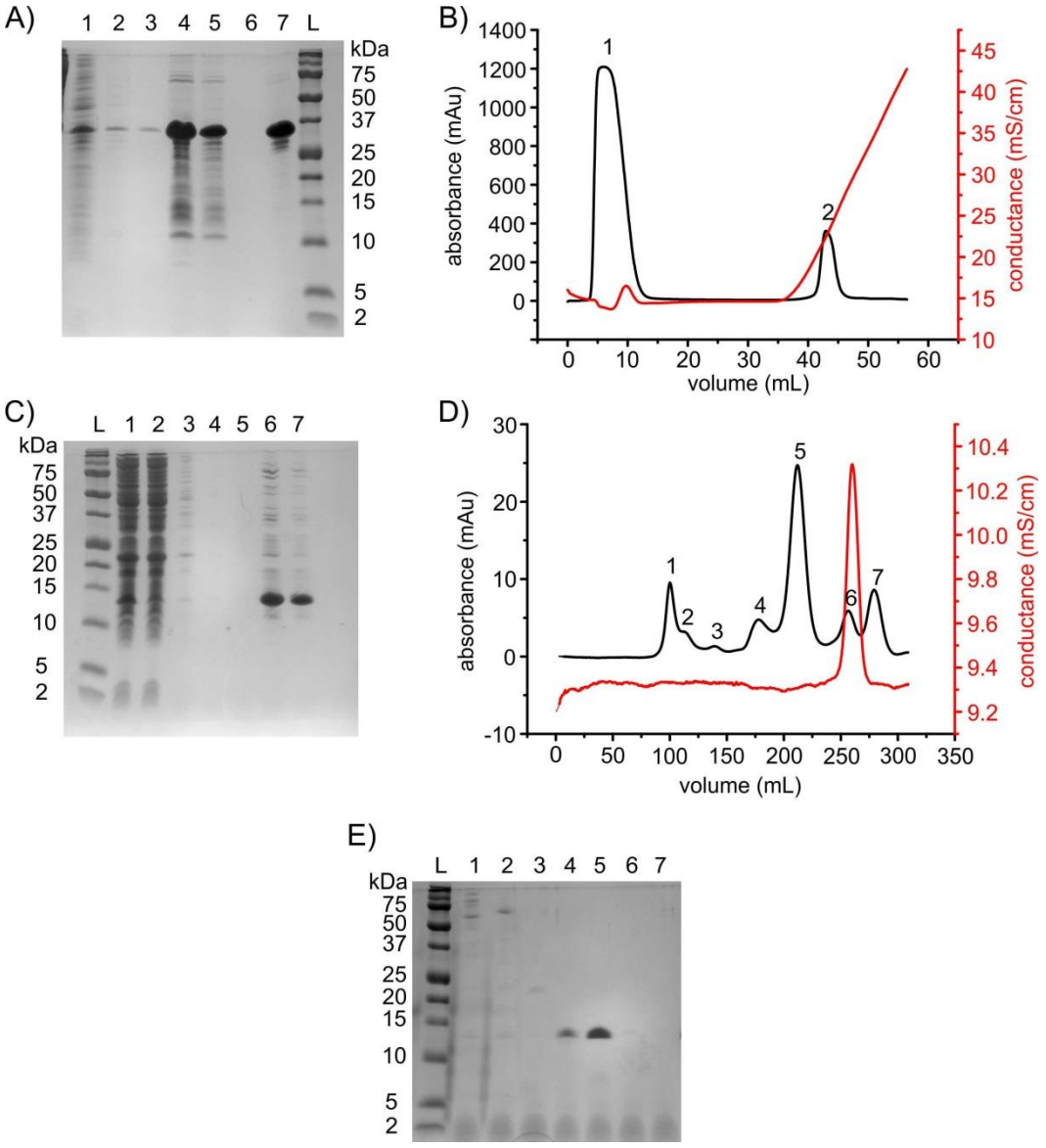


Figure 3.3 – Analysis of TonB purification from *E. coli* BL21 (DE3). A) SDS-PAGE of $TonB_{\Delta TMD}$ purification steps: Lanes 1-3 are the three wash steps in imidazole wash buffer (Section 2.1.11.1) when $TonB_{\Delta TMD}$ is bound to the Ni-NTA resin, lanes 4 and 5 are two elution steps using imidazole elution buffer (Section 2.1.11.2) and lane 6 and 7 are peaks 1 and 2 respectively from cation exchange chromatography (B). $TonB_{\Delta TMD}$ is 23.6 kDa in size, however the presence of the proline-rich linker domain (residues: 70-102) gives an apparent mass of 36 kDa[92]. B) Cation exchange chromatography of $TonB_{\Delta TMD}$ using a Resource™ 6 mL S cation exchange column and eluting with a 1 M NaCl gradient. Peak 2 from the NaCl gradient contains pure $TonB_{\Delta TMD}$. C) SDS-PAGE of the purification steps of $TonB_{CTD}$: Lanes 1 and 2 contain the cell lysate (1/10 dilution), lanes 3-5 contain wash steps and lanes 6 and 7 contain the imidazole elution steps. $TonB_{CTD}$ is 11.7 kDa in size. D) Size exclusion chromatography and E) SDS-PAGE analysis of the peaks from the final stage of $TonB_{CTD}$ purification. $TonB_{CTD}$ elutes in peak 4) and peak 5, which is confirmed by SDS-PAGE (numbers over the gel correspond to the peak number in the gel filtration chromatogram in (D)).

3.2.1.3 Spectroscopic analysis of protein fold

To ensure that the TonB constructs were correctly folded (only constructs without the engineered cysteine were analysed) after purification and were in agreement with previous findings, secondary structure analysis was carried out using Far-UV CD spectroscopy (Figure 3.4A).

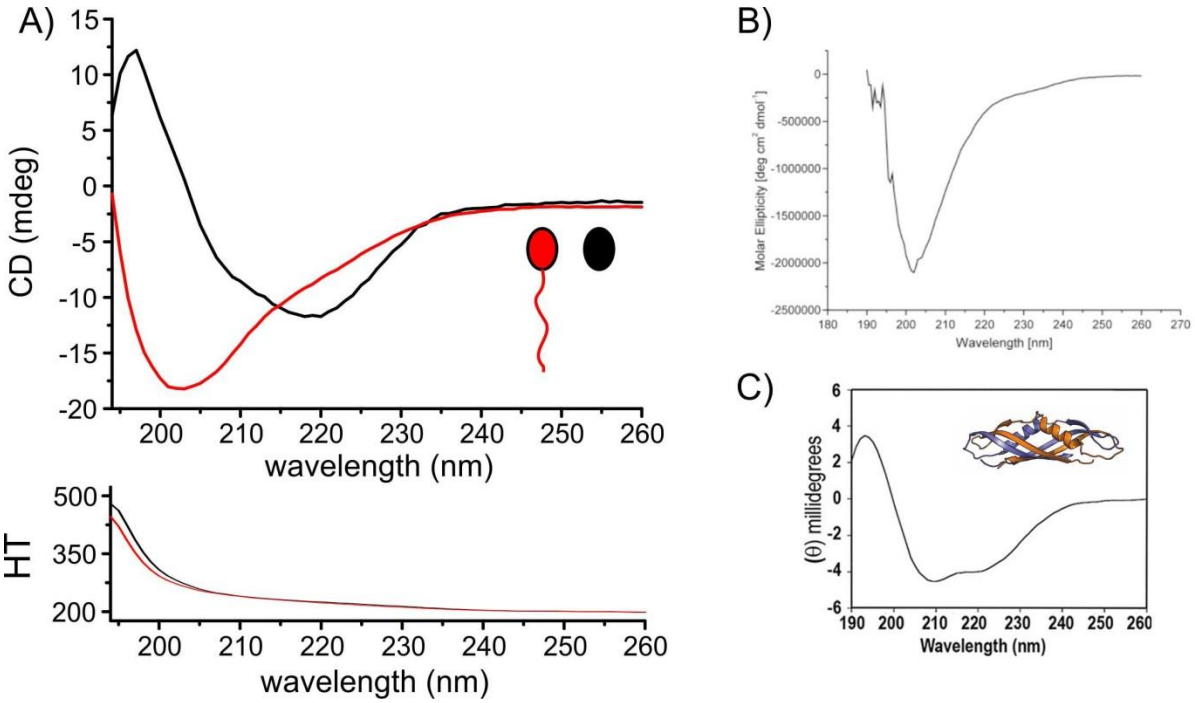


Figure 3.4 – Circular dichroism of TonB constructs. A) Far-UV CD spectra of purified TonB constructs ($TonB_{\Delta TMD}$ red, $TonB_{CTD}$ black) with HT below. $TonB_{\Delta TMD}$ contains a distinctive minima at 205 nm, which is representative of a polyproline helix type II (PPII) from the periplasmic spanning domain. The spectra for $TonB_{CTD}$ is distinct from $TonB_{\Delta TMD}$ and is consistent with the mixed α/β topology of the NMR solution structure. B) Far-UV CD spectra of TonB (residues 56–126 (proline-rich linker domain)), taken from reference [100]. C) Far-UV CD of a truncated TonB C-terminal domain (residues 170-239), which is five residues shorter than the construct that gave the intertwined homodimer X-ray crystal structure (see inset), spectra taken from reference[141].

The results of the far-UV CD secondary structure analysis show that the $TonB_{\Delta TMD}$ spectra is dominated by the signal for a PPII helix, represented by the signal at 205 nm (Figure 3.4A) in agreement with the CD spectra of the periplasmic linker domain alone (residues 56–126) (Figure 3.4B)[100]. The spectra of the C-terminal domain ($TonB_{CTD}$: residues 147-293) is different to the truncated C-terminal domain construct (residues 170-239)(Figure 3.4C). The signal at 209 nm is not visible for the $TonB_{CTD}$ construct used

in this study, which has a lower α-helical content than the truncated intertwined dimer (Figure 1.10).

To investigate whether the TonB constructs maintain tertiary structure after purification, intrinsic tryptophan fluorescence was monitored using a Photon Technology Instrument (PTI) fluorometer. TonB contains a single tryptophan (W213) within the hydrophobic core of the periplasmic C-terminal domain. Using the method described in Section 2.2.6.2, an emission spectrum (290-400 nm) from a 280 nm excitation was carried out for both TonB constructs ± 8 M urea (Figure 3.5). The decrease in fluorescence in 8 M urea and redshift (λmax 322 to 348) was indicative that W213 was no longer packed into the hydrophobic core of the TonB globular C-terminal domain in both constructs. The spectroscopic data provides evidence that both $TonB_{CTD}$ and $TonB_{\Delta TMD}$ constructs retain secondary and tertiary structure after purification.

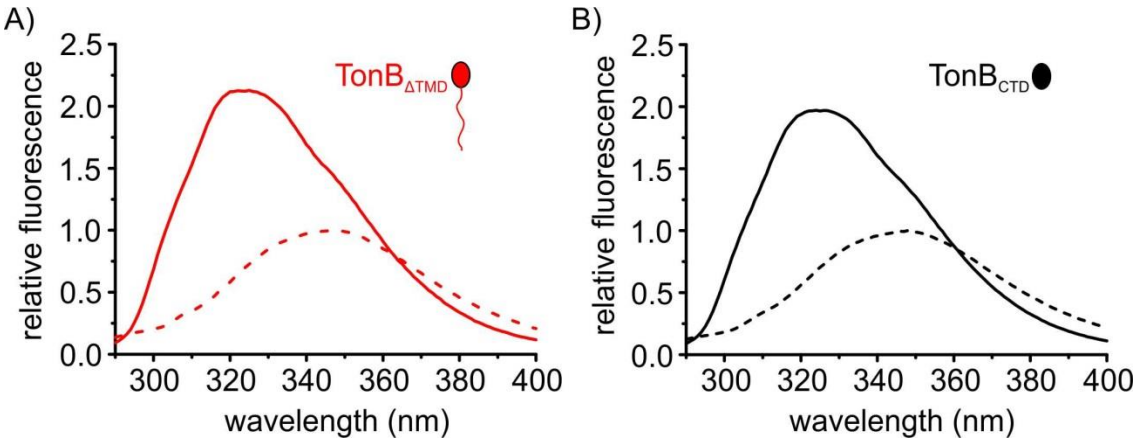


Figure 3.5 – Intrinsic tryptophan emission spectra to assess the folding state of the structured C-terminal domain in both TonB constructs after purification. A) TonBΔTMD and B) TonBCTD in the absence (line) and presence of 8 M urea (dashes).

3.2.2 Binding affinity of TonB constructs to Ton box from BtuB

Before carrying out single-molecule force spectroscopy experiments on the interaction between TonB and the Ton box peptide from BtuB ((Ton box_{BtuB}) sequence: **PDTLVVTANR**GSWSC, residues in bold are the Ton box_{BtuB}), microscale thermophoresis (MST) was performed to ensure the purified TonB constructs were capable of binding the synthetic peptide.

The Ton box$_{BtuB}$ peptide was fluorescently labelled with Alexa Fluor® 488 C_5 maleimide using the method described in Section 2.2.9. The chromatogram from the gel filtration chromatography step of peptide labelling had two major peaks (Figure 3.6A), peak 1 was analysed by liquid chromatography electrospray ionisation mass spectrometry (LC-ESI-MS). The mass spectrum (Figure 3.6B) contained two major species: a 2303.86 Da species, which was the correct mass of the Ton box$_{BtuB}$ peptide (1604.76 Da) attached to Alexa Fluor® 488 C_5 maleimide (698.66 Da). The other species contained an additional 53 Da in mass, and accounts for 25 % of the sample. This additional weight likely arises from metal ions (possibly Fe) chelating to the peptide-Alexa Fluor® 488 C_5 maleimide conjugate, as Alexa Fluor® is known to bind metal ions[190].

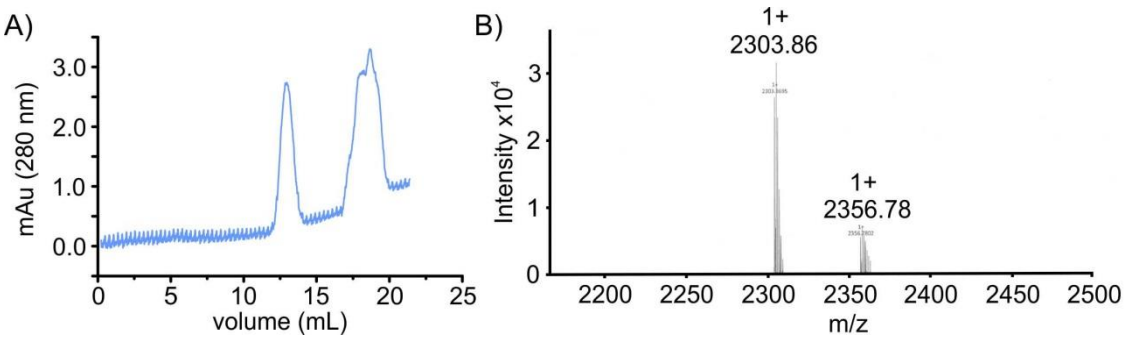


Figure 3.6 – Purification of the Alexa Fluor® 488 labelled Ton box$_{BtuB}$ peptide. A) Size exclusion chromatography using a Superdex™ peptide HR 10/30 gel filtration column (Section 2.2.9) column of the Ton box$_{BtuB}$ peptide labelled with Alexa Fluor® 488 C_5 maleimide. The labelled peptide eluted from the column at 13 mL and is represented by the first absorbance peak in the chromatogram. B) LC-ESI-MS of the first absorbance peak in the chromatogram (A), the major species corresponds to the correct size of the labelled peptide (2303.7 Da), the second peak is likely an adduct on the Alexa Fluor®.

The binding affinity of the fluorescently labelled Ton box$_{BtuB}$ peptide to both TonB constructs was assessed using MST (Section 2.2.6.5). The results are shown in Figures 3.7 and 3.8 with the K_D values in Table 3.3. The results demonstrate that the TonB$_{\Delta TMD}$ construct has a similar binding constant to the TonB$_{CTD}$ construct, indicating that the proline-rich linker does not affect the binding properties of the structured C-terminal domain. The K_D values reported are within the same order of magnitude to that previously found with isothermal calorimetry with Ton Box$_{FhuA}$ and the same TonB$_{CTD}$ construct (36 μM)[86].

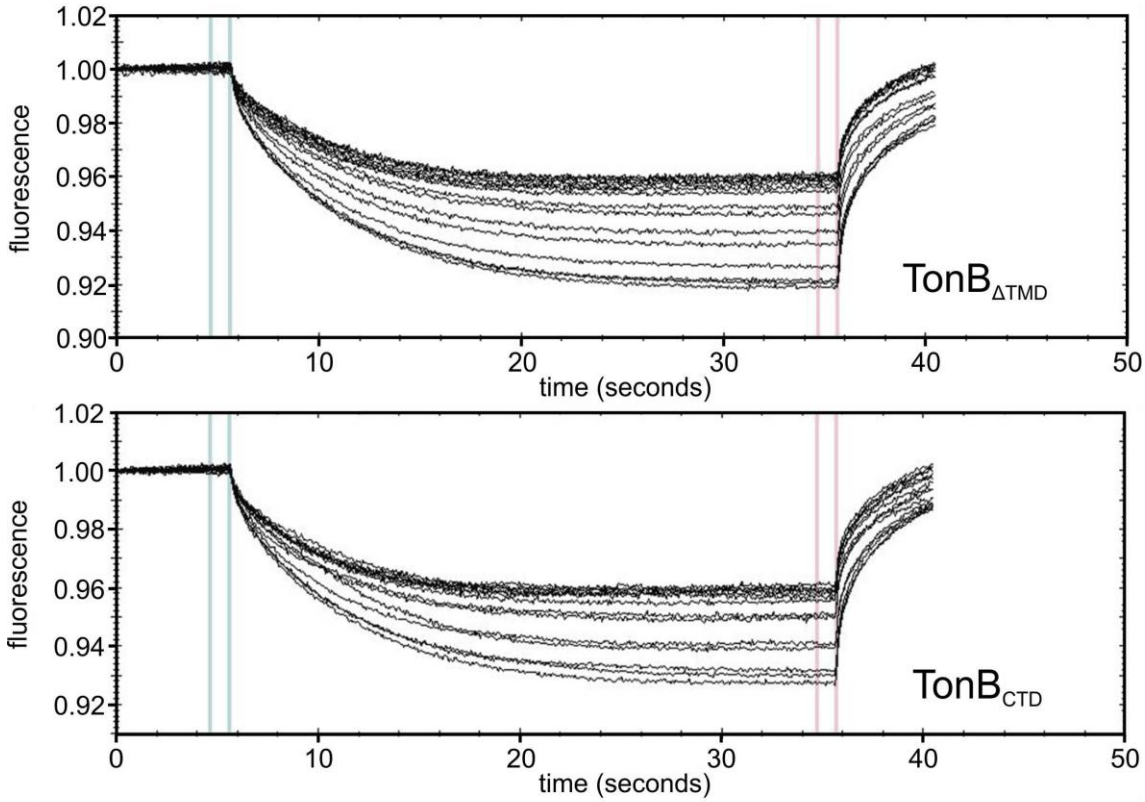


Figure 3.7 - Normalised MST fluorescent signal (LED: blue, 95 % power) for different concentrations of the TonB constructs (labelled) and 250 nM of Ton box-Alexa Fluor® 488. The IR-laser (40 % power) is activated at t =5 s, and is active for 30 seconds before inactivation at t = 35 s. The diffusion (Δfluorescence) is greater when the complex is formed.

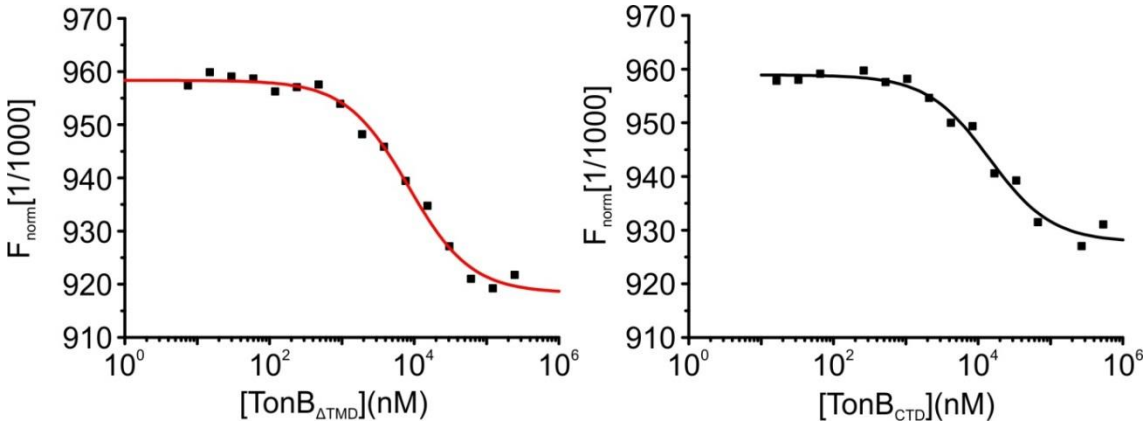


Figure 3.8 – Measuring the binding constant of the Ton box$_{BtuB}$ to TonB constructs. Plot of the F_{norm}[1/1000] (see Section 2.2.6.5) over the log of the TonB construct concentration (nM). The solid lines show the fit using the law of mass action (Equation 2.8) to yield the K_D.

	K_D (μM)
$TonB_{\Delta TMD}$	8.4 ± 0.3
$TonB_{CTD}$	14.1 ± 0.8

Table 3.3 - Equilibrium binding constant values from MST.

3.2.3 Assessing the mechanical strength of the TonB:Ton box$_{BtuB}$ interaction

The TonB:Ton box interaction is mediated via a parallel β-strand augmentation that is predicted to be force-resistant. To test this hypothesis a single-molecule force spectroscopy experiment (SMFS) was designed (Figure 3.9).

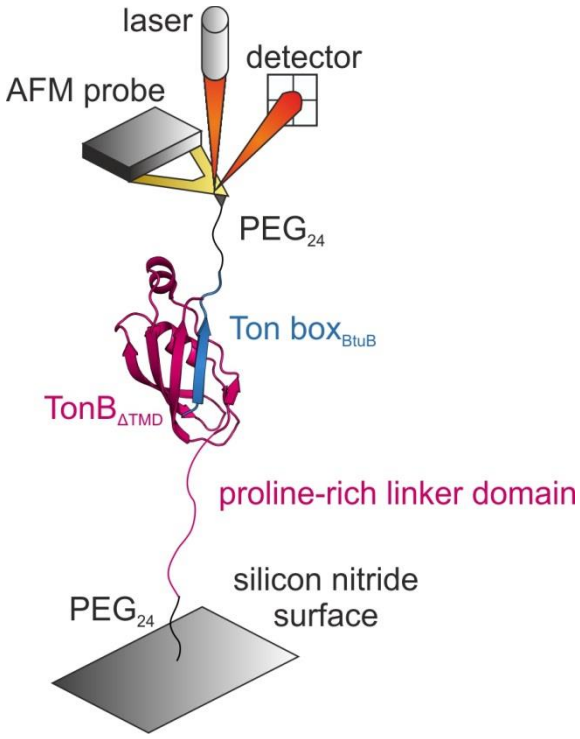


Figure 3.9 – Schematic of the SMFS experimental set up to assess the strength of the TonB:Ton box$_{BtuB}$ interaction. The complex will be brought together by approaching the AFM probe towards the surface which is covalently labelled with $TonB_{\Delta TMD}$ (pink) by heterobifunctional PEG_{24} linkers (black). The Ton box$_{BtuB}$ (blue) is attached to the cantilever using the same chemistry. Withdrawing the AFM probe will propagate force onto the interaction, and the amount of force required for dissociation will be investigated. The cartoon structure representation of the TonB:Ton box$_{BtuB}$ complex is from PDB: 2GSK.

The Ton box$_{BtuB}$ peptide was immobilised onto a silicon nitride AFM probe using flexible heterobifunctional PEG linkers (SM PEG) (see Section 2.2.4). The TonB constructs

(containing cysteine) were attached to a derivatised silicon nitride surface using the same chemistry. The binding partners were brought together by approaching the AFM probe in buffered solution (see Section 2.2.7.2) towards the labelled surface (trigger 300 pN, dwell 0 s), and then the interaction was mechanically pulled apart by withdrawing the probe from the surface at various velocities (200-5000 nms^{-1}). To ensure sufficient surface coverage, 20 μm^2 force maps with 600 force cycles (approach and retract) were taken, and the cantilever was repositioned to a new area between force maps (to ensure adequate surface coverage and maximise the number of unbinding events in the data set). A typical data set at a single retraction velocity contains 2-4 force maps, depending on the hit-rate. Approximately 1 in 10 approach-retract cycles resulted in a true unbinding event (Figure 3.10A). All detected events were manually fit to the worm-like chain (WLC) model (Equation 1.1), where the persistence length was fixed (0.4 nm), the hard contact was used to identify the zero distance and the retraction baseline was used to zero the force (Figure 3.10A).

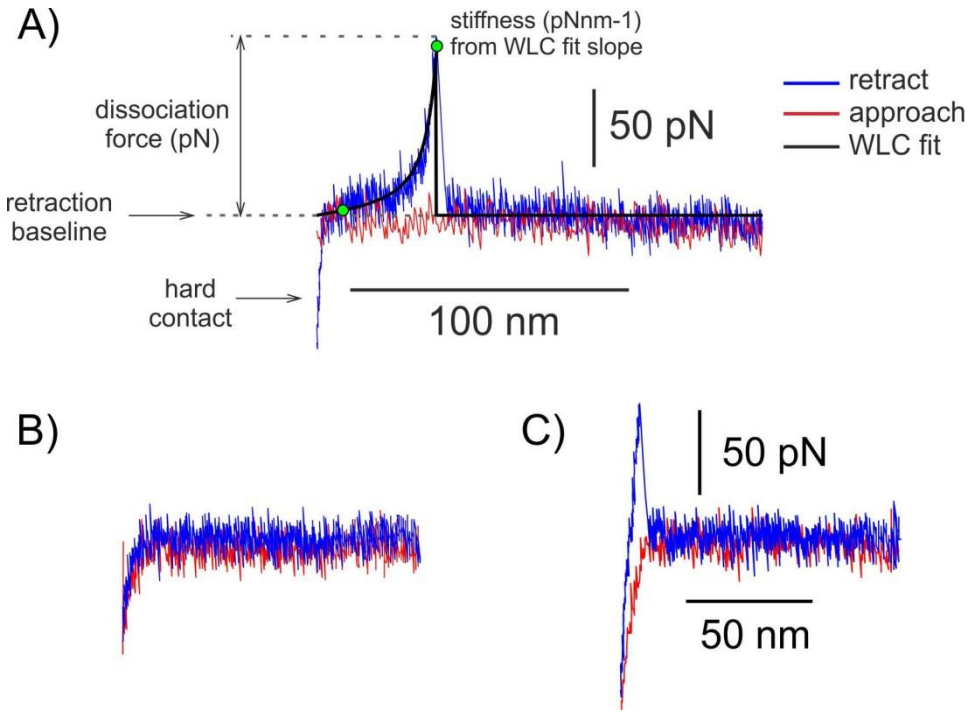


Figure 3.10 – Example force-extension traces from SMFS experiments. A) An annotated force-extension trace with the WLC fit from the SMFS of $TonB_{\Delta TMD}$ with Ton $box_{BtuB.}$ The hard contact (probe pushed against the silicon surface) was used as the zero for the extension scale (nm) and the zero force level was set to the retraction baseline after rupture. The WLC model with a fixed persistence length (0.4 nm) is fit to the curvature of the rupture peak by applying locks to the apex and the base of the peak (green circles). From the fit, a contour length (L_c) and stiffness (WLC slope, see Section 2.2.7.4) are extrapolated. The amount of force required to break to complex is shown. The WLC fit was performed manually using Igor Pro 6.32A (wavemetrics) with an MFP3A AFM software add on (Asylum Research). B) An example force-extension trace with no events C) An example of non-specific tip-sample interactions with a linear force-distance profile.

Detected events for tip-sample separation values of less than 10 nm were in general found to be consistent with non-specific tip-sample interactions and displayed a linear force-distance profile (Figure 3.10C). Force-extension profiles were analysed if the force-extension profile was parabolic and fit well to the WLC model with a fixed persistence length of 0.4 nm. The $TonB_{\Delta TMD}$:Ton box experiment displayed fewer than 0.1% of force curves with more than a single rupture peak. This indicates that there are no detectable unfolding events in $TonB_{\Delta TMD}$ prior to dissociation of the Ton box_{BtuB}.

Analysis of ~400 of the force-extension profiles (per velocity) allows quantification of the mechanical strength and the end-to-end length of the complex (Figures 3.11, 3.12 and 3.13 and Table 3.4).

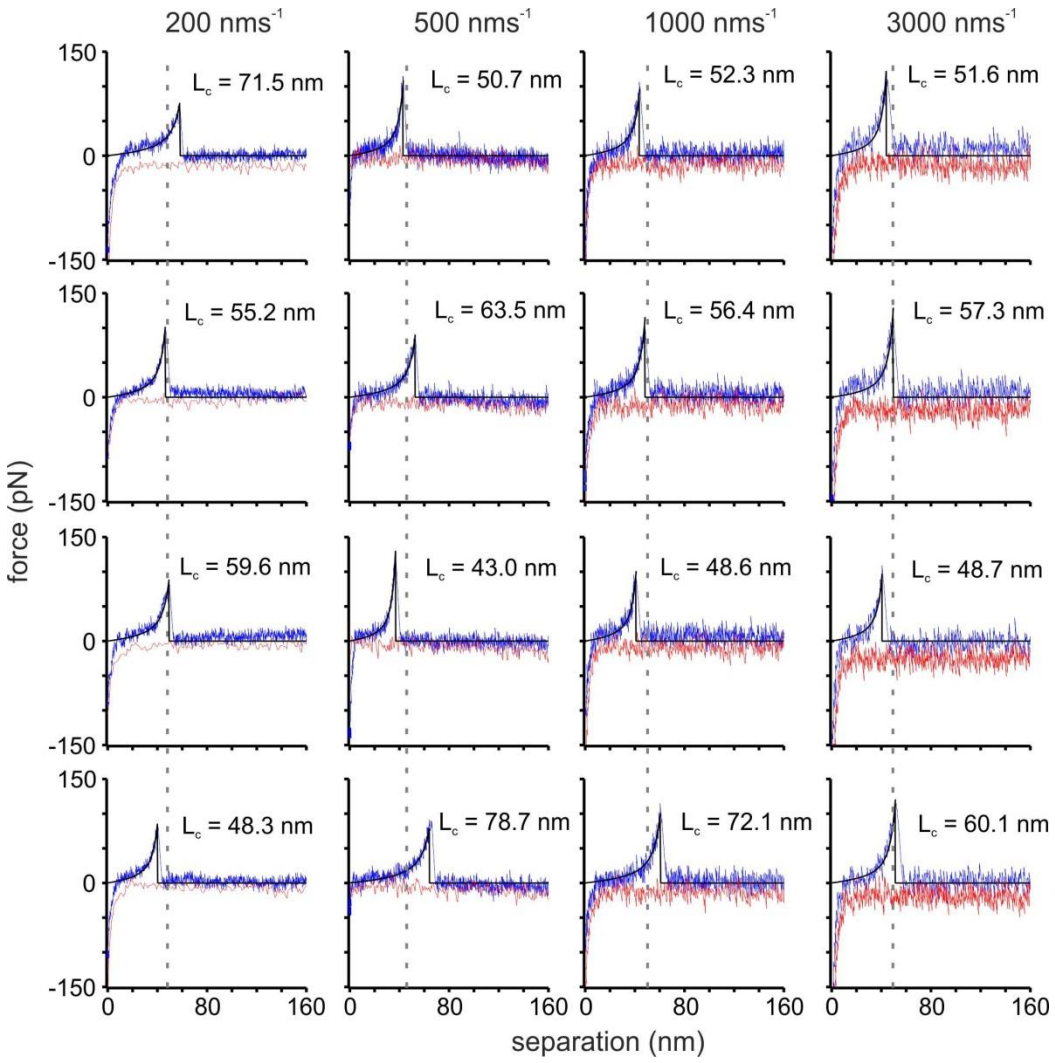


Figure 3.11 – Example dissociation events for four different retraction velocities (200, 500, 1,000, and 3,000 nms^{-1}) from the SMFS of $TonB_{\Delta TMD}$ with Ton box_{BtuB}. Black lines show the WLC fit, red traces are from the approach and blue traces are from the retraction. The extrapolated contour length for each fit is shown. The dashed grey vertical lines represent the average contour length for the whole dataset.

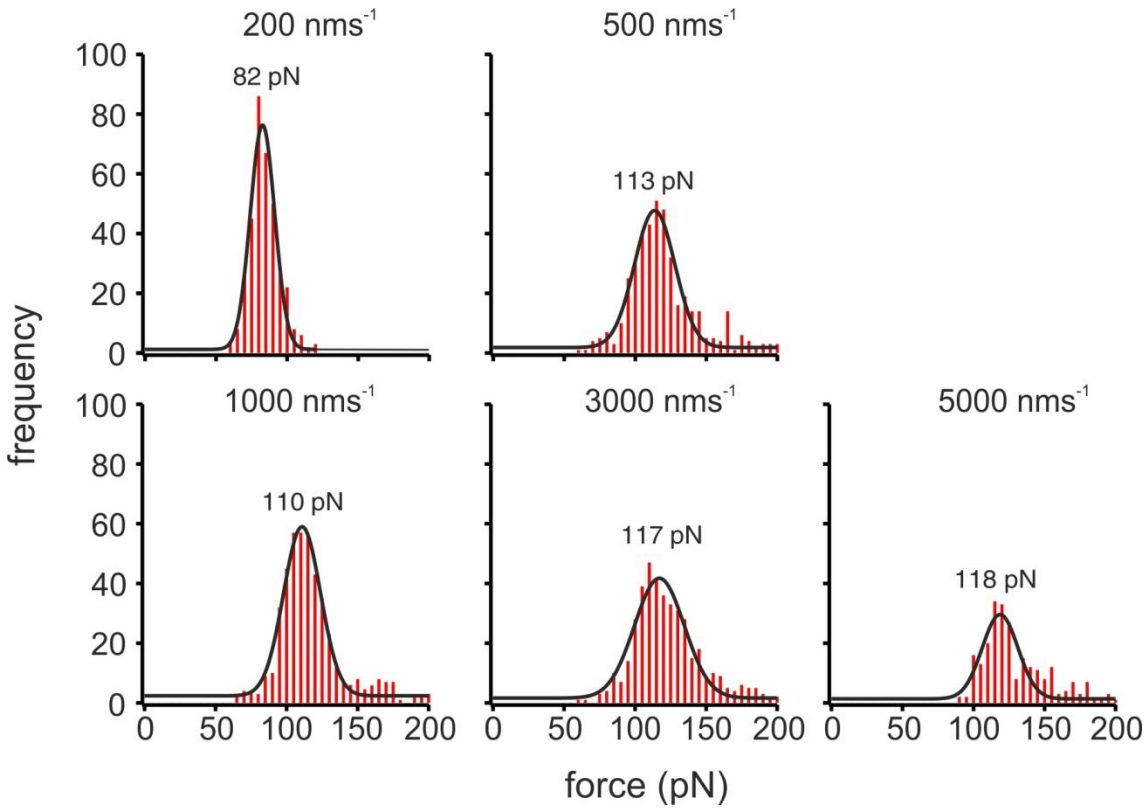


Figure 3.12 - Force-frequency distributions for the $TonB_{\Delta TMD}$:Ton box_{BtuB} dissociation events with fits to a single Gaussian distribution for the five pulling velocities (200, 500, 1000, 3000, and 5000 nms^{-1}) used in this study. Histograms contain the full triplicate data collected.

Speed (nms^{-1})	#	force (pN)	average (pN)
200	1	80	
	2	87	84 ± 7
	3	84	
500	1	109	
	2	110	105 ± 15
	3	95	
1000	1	119	
	2	115	113 ± 13
	3	106	
3000	1	121	
	2	123	119 ± 9
	3	114	
5000	1	151	
	2	121	133 ± 30
	3	126	

Table 3.4 – AFM data of the unbinding of $TonB_{\Delta TMD}$ from the Ton box of BtuB at various retraction velocities. Force values were calculated by plotting the distribution of events and fitting with a Gaussian to obtain the mean value. The error on the average is the range of the triplicate averages.

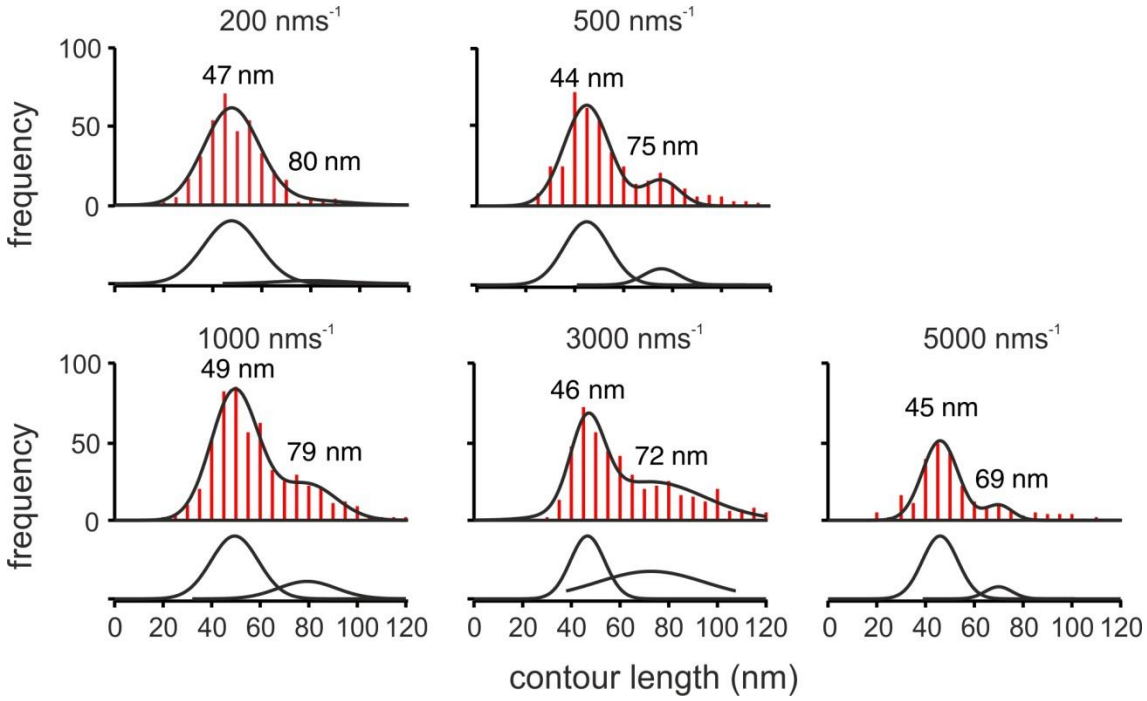


Figure 3.13 - Contour-frequency distributions for the $TonB_{\Delta TMD}$:Ton box_{BtuB} dissociation events fit to a double Gaussian distribution for the five pulling velocities (200, 500, 1000, 3000, and 5000 nms^{-1}) used in this study. Individual Gaussians are shown below. Histograms contain the full triplicate data collected.

Combining these data in a scatterplot allows the specificity of the interaction to be assessed by the presence (specific) or absence (non-specific) of a 'hot-spot' indicative of correlated forces and distances (Figure 3.14). A binding-deficient variant of Ton box_{BtuB} (L8P)[79] was used as a control to assess the specificity of these measurements. In experiments with Ton box_{BtuB}(L8P), random events at a lower hit rate were observed with no hot-spot observable in the scatterplot (Figure 3.14 bottom panels).

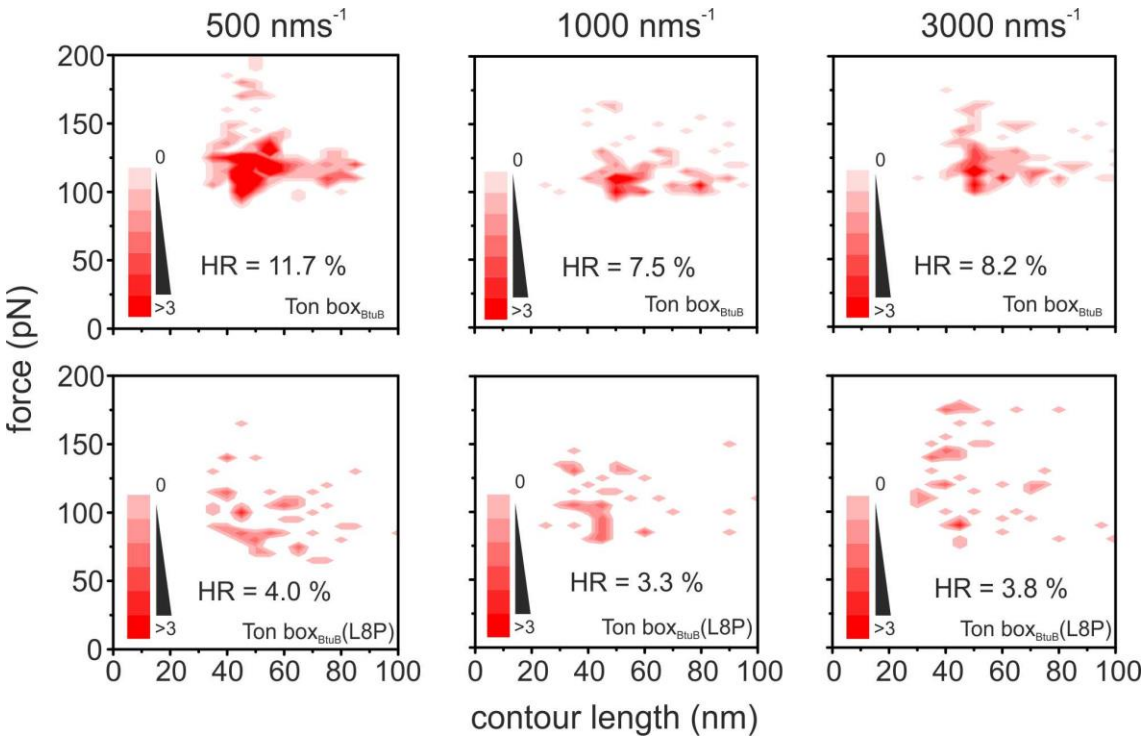


Figure 3.14 - Force spectroscopy data (rupture force and contour length) plot as scatterplots at three different retraction velocities of Ton box$_{BtuB}$ or the Ton box$_{BtuB}$ L8P peptide unbinding from TonB$_{\Delta TMD}$. Each contour map contains all filtered data reflecting single-molecule dissociation events after 1200 surface-approaches over a large surface area. The 'hot spot' of specific unbinding events are lost in the L8P mutant which confirms that the wild-type data is specific to the TonB$_{\Delta TMD}$:Ton box$_{BtuB}$ interaction. The hit rate (HR) is shown in the inset of each graph and denotes the percentage of approach/retract cycles that result in force-extension profiles that meet the criteria discussed in Section 2.2.7.4. The points were binned by a contour length of 5 nm and force of 10 pN, the number of points that occupy a quadrant are indicated by the colour of the contours.

Loading rates were calculated by fitting the WLC model to the unbinding event of a force-extension plot (Figure 3.10A). The instantaneous gradient of this fit at rupture (WLC$_{slope}$) was calculated by inserting the derived contour length and extension at rupture into a differentiated form of the same equation (Equation 2.14). The loading rate at rupture was then obtained by multiplying this by the retraction velocity. For each retraction velocity used, force and loading rate frequency histograms were used to obtain modal values, and the average of triplicate values was taken for each velocity. The natural logarithm of the loading rate (Ns^{-1}) plot against the rupture force (N) gave a linear dependence of force against loading rate (Figure 3.15). A linear fit was used to extrapolate the off rate at zero force (k_{off}^{0F}) using Equation 2.16 and x_u with Equation 2.15, the errors were calculated using the bootstrapping method (Section 2.2.7.4).

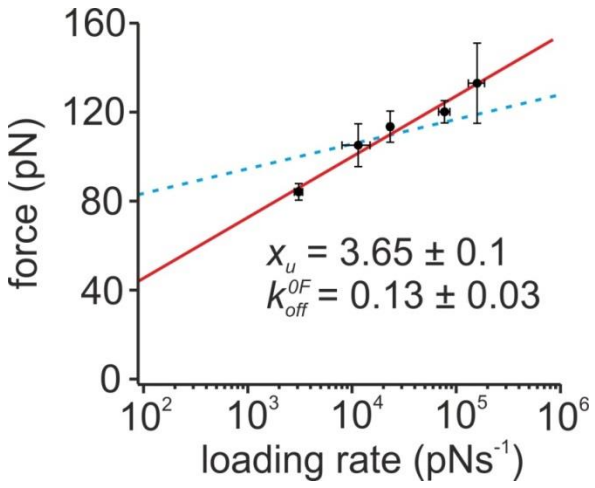


Figure 3.15 - The dynamic force spectrum (a plot of rupture force versus the natural logarithm of the force loading rate (pNs^{-1}) of the TonB$_{\Delta TMD}$:Ton box$_{BtuB}$ interaction. The error bars show the range of measurements from the triplicate mean. Data is fit to a linear regression (red line). The dynamic force spectrum obtained for the highly avid E9:Im9 (E9 cross-linked between residues 20-66) is shown for comparison (dashed blue line)[158]. The units for x_u and k_{off}^{0F} are Å and s^{-1} respectively.

The resultant dynamic force spectrum reveals that rupture occurs at 84-132 pN (Figure 3.15) at the relatively high loading rates applied using AFM. This is a remarkably strong interaction for a complex with a relatively low affinity. For example, at similar loading rates, the dissociation of E9$_{20\text{-}66}$:Im9 (K_D = 10^{-14} M)[158] and an antibody and its epitope (K_D = 10^{-9} M)[191] occurs at forces of ~100 pN and ~160 pN respectively. It has been previously suggested that due to the additional β$_4$-strand in the NMR structure of TonB$_{CTD}$ (Figure 1.10C), the association rate could be slow (as conformational rearrangement would be required for Ton box binding), which is why the K_D for the interaction is in the micromolar range (slow association, slow dissociation)[86]. Using the AFM derived dissociation rate and the MST derived binding constant for TonB$_{\Delta TMD}$, the calculated association rate of the Ton box$_{BtuB}$ is 10^4 M^{-1} s^{-1} which is 4 orders of magnitude slower than the theoretical diffusion-limited biomolecular collision rate[192].

Interestingly, the force-extension profiles, distance histograms and scatterplots of these data all reveal that while dissociation occurs at a single force, two distributions are evident for the end-to-end length of the complex at rupture (see Figures 3.13 and 3.14) with Gaussian fit values of 47 ± 2 and 76 ± 3 nm. These measured lengths suggest the

presence of both a compact form and extended form of the proline-rich linker domain of TonB within the structured ensemble.

As shown in Figure 3.16, the lengths of the PEG linkers[159] and the structured TonB C-terminal domain in complex with the Ton box of BtuB (PDB: 2GSK) are known, therefore by subtracting 24.5 nm from the most probable contour lengths observed gives the length of proline-rich linker domain under extension. The data shows that the linker domain predominantly exhibits a ~22 nm length under force; however the less frequently observed longer linker length is ~51 nm. The linker domain contains 118 amino acids, if this assumed an unstructured chain; the length would be ~47 nm, in good agreement with the observed contour length distribution.

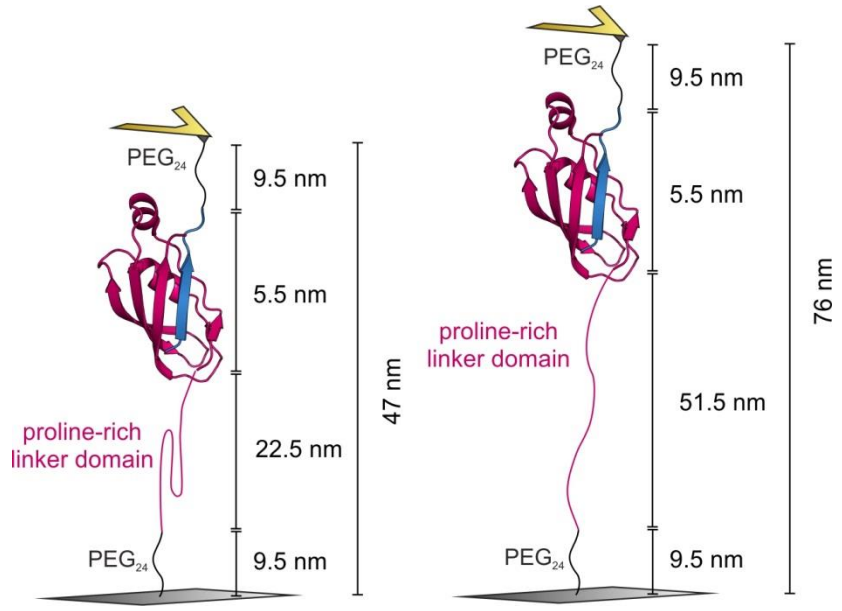


Figure 3.16 – Individual lengths of each component of the system. The known length of the PEG_{24} linkers (9.5 nm), $TonB_{CTD}$:Ton box complex (5.5 nm) are shown. Additional length observed will be from the proline-rich linker domain, the two lengths observed are shown which are consistent with a compact and extended form.

To confirm that this bimodality did arise from the proline-rich linker domain, rather than the structured C-terminal domain being unfolded at a force below the sensitivity of the AFM experiment (~20 pN), the $TonB_{CTD}$(K144C) construct missing the linker domain (Figure 3.1) was used for the SMFS experiments using the same method as for $TonB_{\Delta TMD}$(V32C). The results clearly show that the bimodality is lost when the linker

domain is removed; however the rupture force is the same, indicating that the proline-rich linker domain does not affect the strength of the interaction (Figure 3.17).

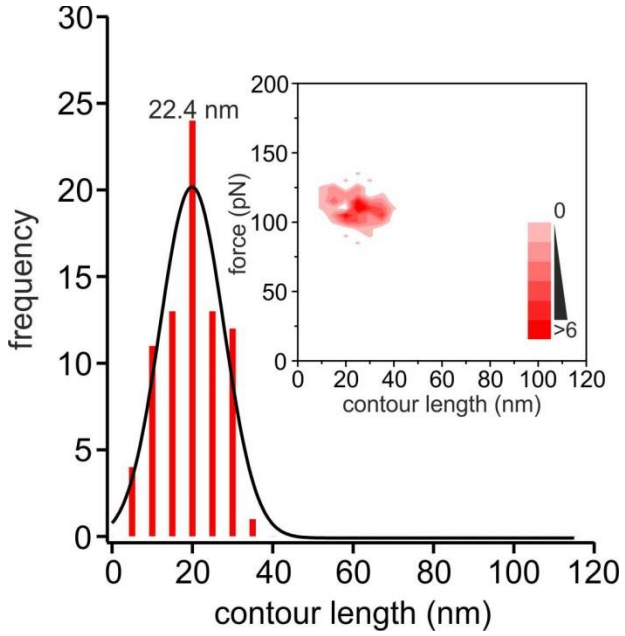


Figure 3.17 - Contour-frequency distribution for the $TonB_{CTD}$:Ton Box_{BtuB} dissociation events at a pulling velocity of 500 nms^{-1} fit to a single Gaussian. The inset shows a scatterplot of contour length vs rupture force for the same data set. Points were binned by a contour length of 5 nm and force of 10 pN and the number of points that occupy a quadrant are indicated by the colour of the contours.

As the proline-rich linker contains PE and PK repeats (see Figure 3.1), an electrostatic interaction within the linker was suspected to be stabilising a force-resistant compact conformation. The effect of disrupting these potential salt bridges using high concentrations of NaCl was investigated using SMFS.

3.2.4 The influence of salt on the contour length of extended $TonB_{\Delta TMD}$ at dissociation

As the contour length distribution can give an accurate measure of biomolecules extended under force[159], SMFS was used to investigate the effect of high concentrations of NaCl on the length of the proline-rich linker domain of TonB. SMFS was carried out using the same protocol as discussed in section 2.2.7.2; however after data collection in Tris-buffered saline (25 mM Tris-HCl pH 8.0, 128 mM NaCl) at a retraction velocity of 1000 nms^{-1} (Figure 3.18B and C); the buffer was changed 25 mM Tris-HCl pH 8.0, 628 mM NaCl. Remarkably the contour length distribution changed so that the major species

was the longer contour length (Figure 3.18D). It should be also noted that the different lengths and presence of additional NaCl did not change the amount of force to break the interaction (Figure 3.19), nor the binding constant (Section 3.2.6) which was expected for an interaction mediated by H-bonds. Also no additional characteristics (i.e. unfolding events) were found in the force-extension traces (Figure 3.18B).

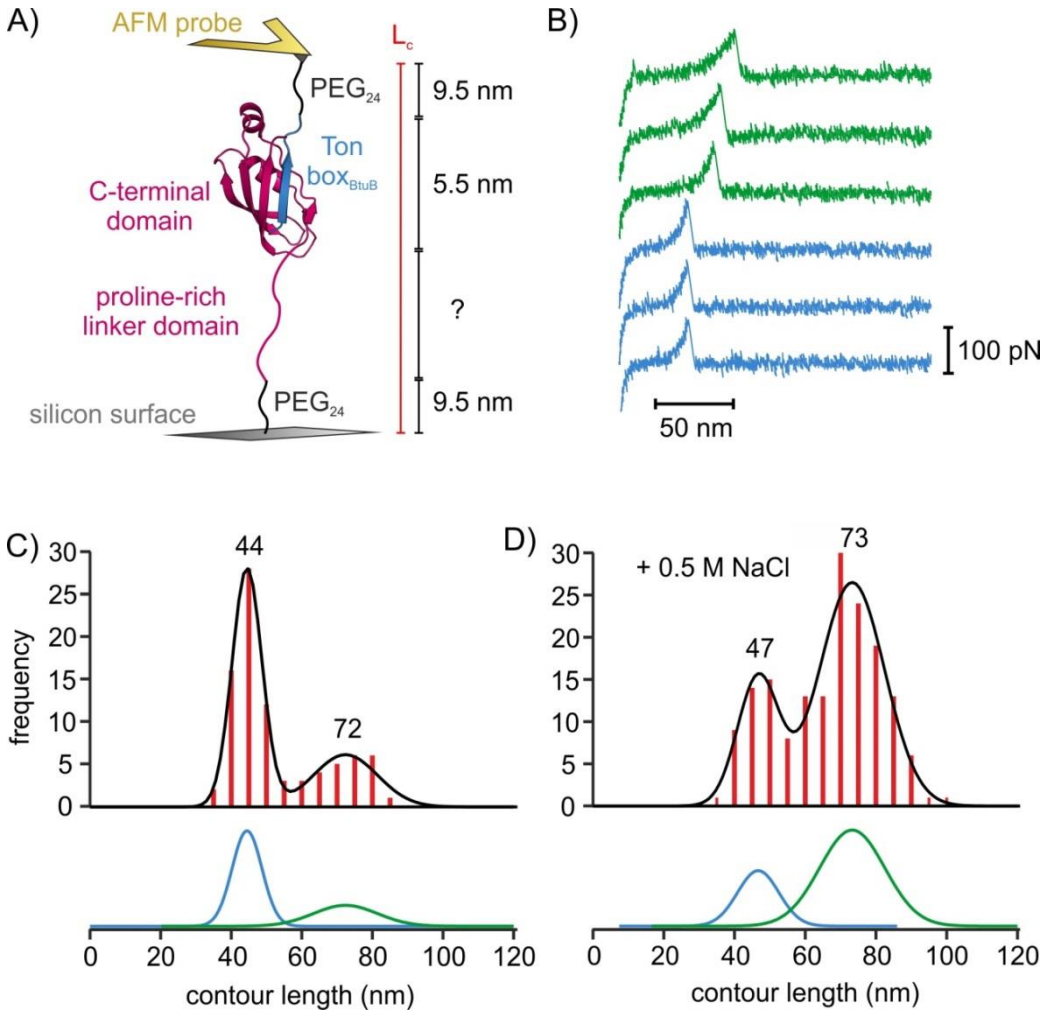


Figure 3.18 – The influence of NaCl on the contour length of $TonB_{\Delta TMD}$. A) Schematic showing the length of individual components of the $TonB_{\Delta TMD}$ (pink): Ton box_{BtuB} (blue) complex with PEG_{24} linkers (labelled) used in the SMFS experiments. The proline-rich linker domain contains 118 residues. B) Force-extension curves collected at 500 nms^{-1} showing single unbinding events of $TonB_{\Delta TMD}$ with Ton box_{BtuB}. The green traces are events that show a longer contour length of ~73 nm, and the blue traces are the shorter events that have a contour length of ~45 nm C) Contour length distribution with Gaussian fits below of the $TonB_{\Delta TMD}$:Ton box_{BtuB} interaction pulled apart by SMFS at 500 nms^{-1} in 25 mM Tris-HCl, pH 8.0, 128 mM NaCl (Tris-buffered saline (TBS)) and D) TBS + 0.5 M NaCl. Modal values from Gaussian fits are shown.

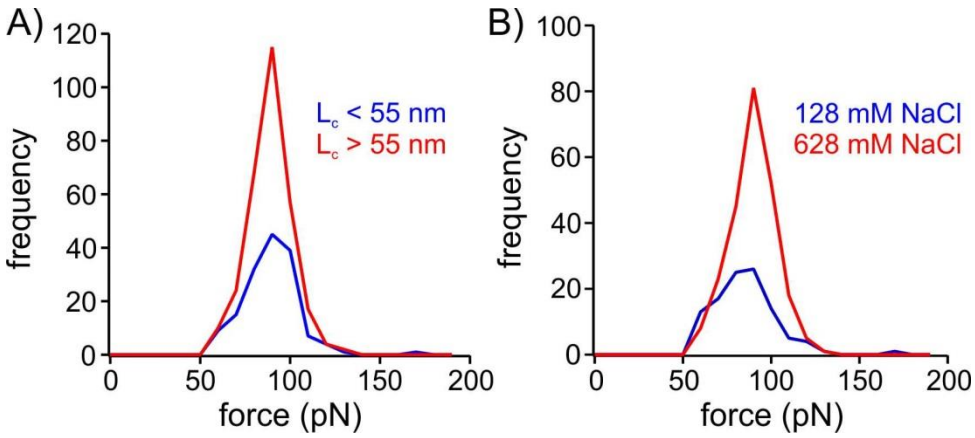


Figure 3.19 – Force distributions of the $TonB_{\Delta TMD}$:Ton box_{BtuB} interaction by SMFS at 1000 nms^{-1} A) The difference in contour length (L_c) or the presence of additional 0.5 M NaCl (B) does not affect the rupture force of the complex.

The results of this experiment imply that the proline-rich linker domain contains a mechanically strong structure that is lost in more saline conditions. To further understand the structural and biophysical properties of the linker domain and its influence on the C-terminal domain, denaturation equilibrium, NMR and SAXS experiments were carried out in the presence and absence of additional 0.5 M NaCl.

3.2.5 Effect of salt on the stability of the C-terminal domain of TonB

To understand the effect of high concentrations of NaCl on the structured C-terminal domain of TonB and to investigate whether the proline-rich linker domain affects its thermodynamic stability, a denaturant equilibrium study using urea was carried out using the method described in Section 2.2.6.2. The stability of the TonB constructs ± the proline-rich linker domain and ± 0.5 M additional NaCl were monitored using the fluorescence of the single buried tryptophan (W213) within the structured domain. The resulting fit of a two state transition chemical denaturation model (Figure 3.20) shows that the proline-rich linker domain was slightly destabilising to the structured domain, as the ΔG_{UN} (kJ mol^{-1}) was lower for $TonB_{\Delta TMD}$ (13.4 ± 0.4) than the $TonB_{CTD}$ construct (15.4 ± 0.3) (Table 3.5), where the m_{UN} values do not vary more than 10 %, indicating that the linker domain does not significantly interact with the globular domain. However upon the addition of 0.5 M NaCl, the C-terminal domain of TonB containing the proline-rich linker domain ($TonB_{\Delta TMD}$) was significantly stabilised (ΔG_{UN} = 16.0 ± 0.4) implying that charged residues within the linker domain contact and possibly destabilise the

structured domain or that dimerisation via the linker domain destabilises the C-terminal domain (see Section 3.2.9).

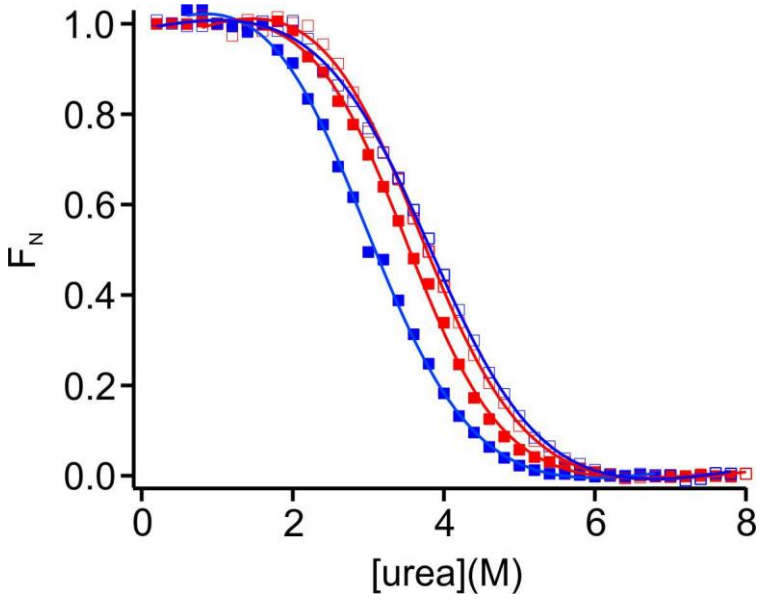


Figure 3.20 - Urea equilibrium denaturation of the TonB constructs ($TonB_{\Delta TMD}$ in blue and $TonB_{CTD}$ in red) in the absence (closed squares) and presence (open squares) of 0.5 M NaCl. Data points are fit to a two state model (Section 2.2.6.2). The data was normalised using Equation 2.2.

	construct	ΔG_{UN} (kJ mol^{-1})	m_{UN} (kJ mol^{-1} M^{-1})
	$TonB_{\Delta TMD}$	8.9 ± 1.0	3.1 ± 0.3
	$TonB_{CTD}$	12.4 ± 0.6	3.6 ± 0.2
+ 0.5 M NaCl	$TonB_{\Delta TMD}$	11.4 ± 0.8	3.2 ± 0.2
	$TonB_{CTD}$	11.3 ± 0.6	3.0 ± 0.2

Table 3.5 - Thermodymanic parameters of TonB constructs with and without an additional 0.5 M NaCl.

3.2.6 Effect of salt on the binding properties of TonB with the Ton box$_{BtuB}$

As demonstrated in the previous section, high concentrations of NaCl alter the stability of the structured C-terminal domain of TonB, and the presence of the proline-rich linker domain in normal buffer conditions was destabilising. As the C-terminal domain functionally binds to the Ton box of ligand bound TBDTs, the effect of high salt

concentrations on the binding of the Ton box was investigated using an Alexa Fluor® 488 labelled Ton box$_{BtuB}$ with MST (Figure 3.21) (For method, see Section 2.2.6.5).

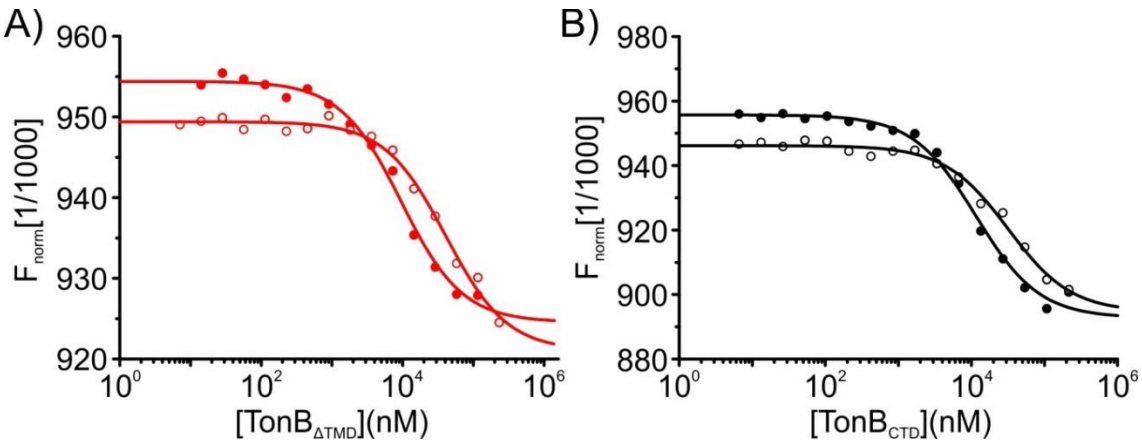


Figure 3.21 – The effect of 0.5 M NaCl on the binding of TonB constructs to the Ton box$_{BtuB}$ peptide. A plot of the F_{norm}[1/1000] (see Section 2.2.6.5) over the log of the TonB construct concentration (nM) is shown. The solid lines show the fit using the law of mass action to yield the K_D. Open circles show data from buffer conditions with additional 0.5 M NaCl.

	construct	K_D (μM)
	TonB$_{\Delta TMD}$	9.5 ± 0.3
	TonB$_{CTD}$	11.5 ± 0.4
+ 0.5 M NaCl	TonB$_{\Delta TMD}$	40.8 ± 1.3
	TonB$_{CTD}$	31.6 ± 1.1

Table 3.6 - Equilibrium binding constants in the presence and absence of 0.5 M NaCl.

The addition of 0.5 M NaCl to the buffer slightly alters the binding constant of TonB to the Ton box$_{BtuB}$. The TonB$_{\Delta TMD}$ construct was 4 fold weaker at binding the Ton box in high salt, whilst TonB$_{CTD}$ was 3 fold weaker (Table 3.6). An explanation for this observation is that a conserved salt bridge between TonB and the Ton box of BtuB (Arg158 of TonB and Asp6 of BtuB) is believed to be a nucleation site for the initiation of the β-strand augmentation[145]; therefore in high salt conditions this electrostatic interaction was equally dampened for both constructs. However in comparison to other avid protein-protein complexes (such as Im9:E9[193], ribonuclease inhibitor:angiogenin[194] and Barnase:Barstar[195]), the effect of additional NaCl on the TonB:Ton box$_{BtuB}$ complex is minimal, highlighting that the interaction is mediated by H-bonds.

3.2.7 NMR chemical shift perturbation demonstrates the proline-rich linker domain has an effect on the C-terminal domain of TonB in high salt conditions

The NMR chemical shifts of the C-terminal domain of TonB have previously been assigned[86], therefore ^{1}H-^{15}N HSQC was used to find any chemical shift perturbations within the structured C-terminus of TonB influenced by the proline-rich linker domain in the presence of absence of 0.5 M NaCl. The $TonB_{\Delta TMD}$ and $TonB_{CTD}$ constructs were ^{15}N-labelled using the procedure described in Section 2.2.2.7 and concentrated to 400 µM. 10 % D_2O was applied to the sample before data collection (Section 2.2.6.8). Figure 3.22 shows the HSQC of both $TonB_{\Delta TMD}$ (black) and $TonB_{CTD}$ (pink, with assignments) overlaid, where the additional chemical shifts in the $TonB_{\Delta TMD}$ spectrum correspond to the proline-rich linker domain and are clustered around 8.2-8.5 (^{1}H) and 121-127 (^{15}N) ppm. These chemical shifts show a larger than expected 1H dispersion for an unfolded protein, consistent with the formation of a helical structure, in agreement with the PPII conformation indicated by CD spectrometry (Figure 3.4A). The presence of the proline-rich linker domain did not alter the structure of the C-terminal domain, as no significant chemical shifts were observed (Figure 3.22 and Figure 3.23A).

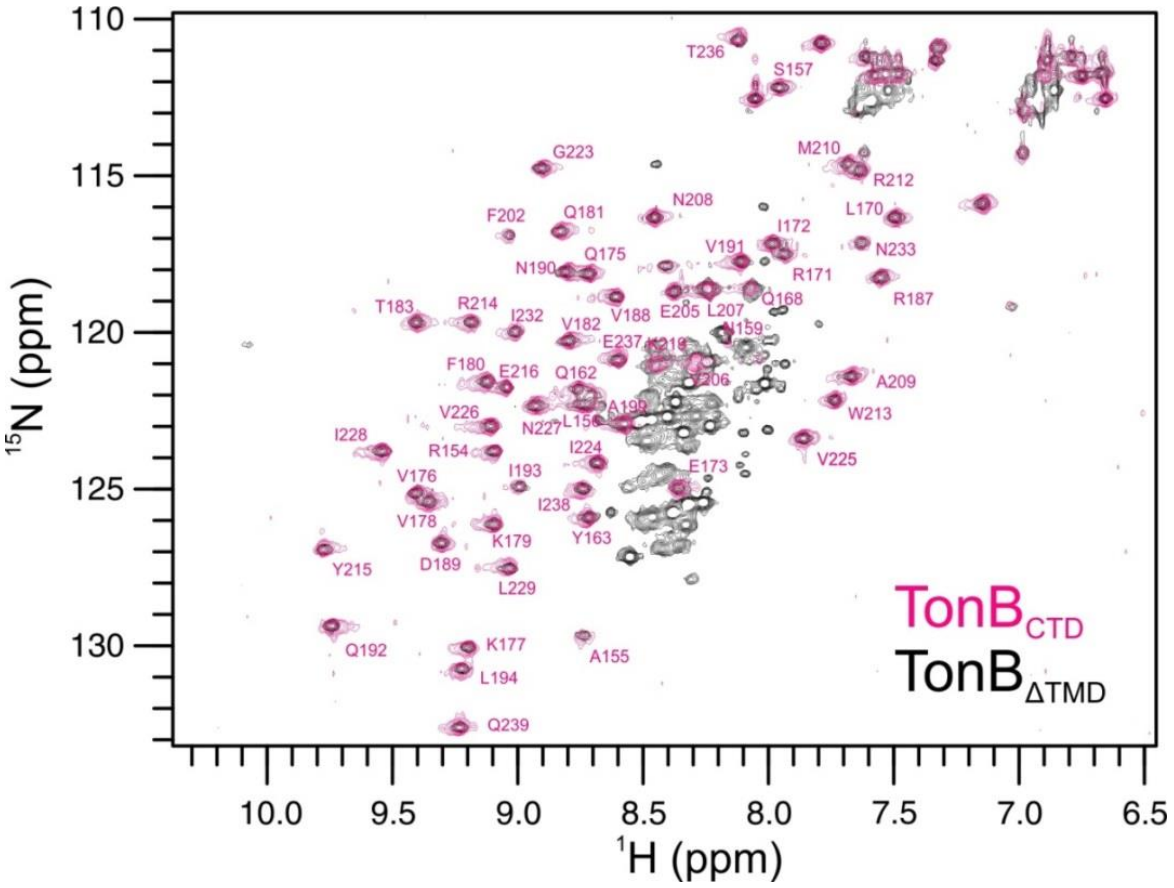


Figure 3.22 – HSQC spectra of TonB$_{CTD}$ and TonB$_{\Delta TMD}$ (labelled pink assigned peaks and black peaks respectively). No pertubations were found in the structured C-terminal domain when the proline-rich linker domain is present (also see Figure 3.23). Additional peaks (no pink overlay) correspond to the residues from the proline-rich linker domain.

The presence of 0.5 M NaCl caused large chemical shift perturbations (CSP) in the C-terminal domain of both constructs (Figure 3.23B and C). Additional perturbations were observed in the TonB$_{\Delta TMD}$ construct that were absent in the TonB$_{CTD}$ construct (Figure 3.23B). This indicates that the proline-rich linker domain interacts with the C-terminal domain in more saline conditions, suggesting a collapse around the structured domain.

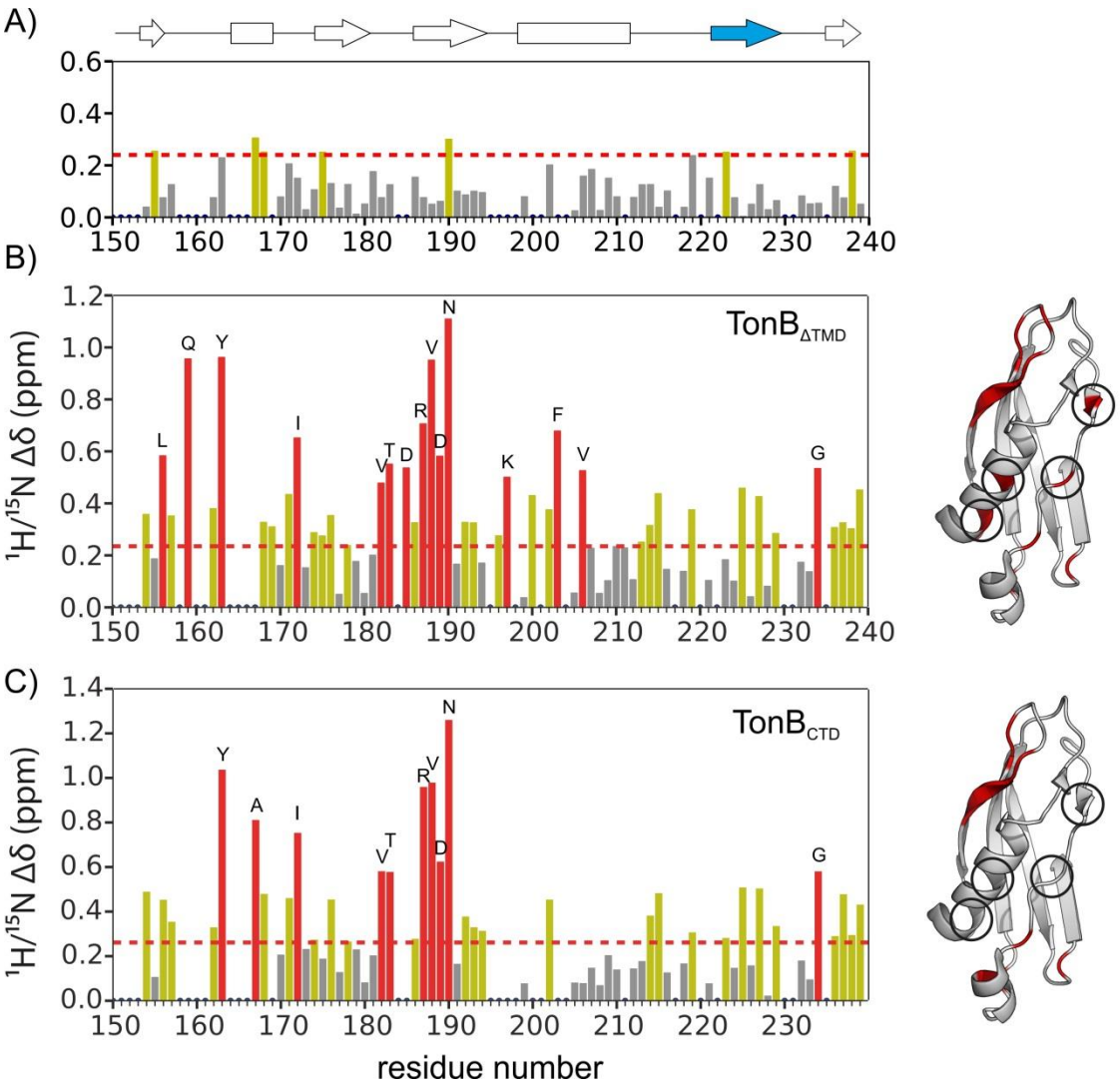


Figure 3.23 – Chemical shift pertubations within the C-terminal domain of TonB caused by the presence of the linker domain (A) or additional 0.5 M NaCl in constructs with the proline-rich linker domain (B, TonB$_{\Delta TMD}$) and without (C, TonB$_{CTD}$). The dashed red line shows the average global pertubation (from B), red bars show perturbations that are two standard deviations greater than the average (residue names shown) and yellow bars are shifts equal to the average and are less than two standard deviations from the average. The secondary structure schematic above the histograms shows the Ton box binding site in blue, which is unperturbed. Cartoon structures of the C-terminal domain are shown (right) with the location of significant pertubations coloured in red. The circles highlight the shifts observed in TonB$_{\Delta TMD}$ that are absent in TonB$_{CTD}$.

Global perturbations were also observed from the proline-rich linker domain (Figure 3.24) which indicates a conformational change in more saline conditions, in agreement with the loss of structure in the SMFS experiments.

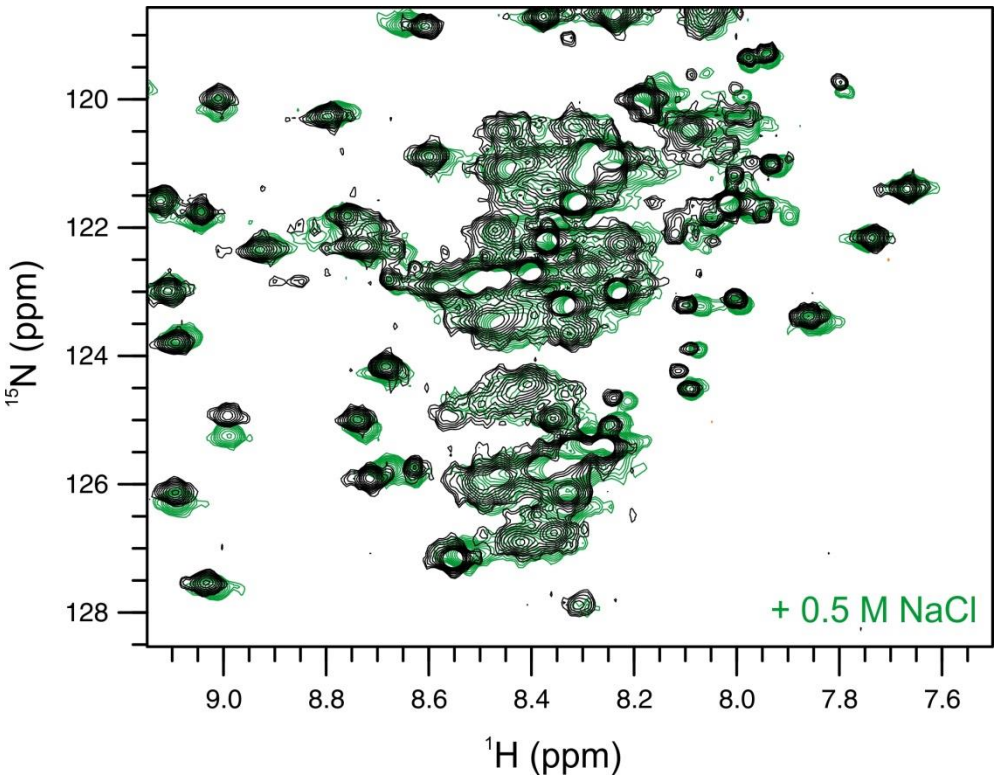


Figure 3.24 – Chemical shifts from the proline-rich linker domain of $TonB_{\Delta TMD}$ in the absence (black) and presence (green) of additional 0.5 M NaCl. The global shifts suggest the presence of a salt-dependent conformational change within this region when 0.5 M NaCl is present.

3.2.8 SAXS reveals salt sensitive change of the proline-rich linker

For more insight of the proline-rich linker domain of TonB, size-exclusion chromatography with small angle X-ray scattering (SEC SAXS) was performed at Diamond beamline 21 using the method described in Section 2.2.6.6. $TonB_{\Delta TMD}$ is known to form dimers[106]; therefore SEC was required to separate the monomers prior to SAXS evaluation. $TonB_{CTD}$ was purely monomeric and eluted from the column as a single peak, whereas $TonB_{\Delta TMD}$ eluted as two peaks (Figure 3.26A) corresponding to dimer and monomer (later confirmed by SEC-MALLS, see Section 3.2.9). The SAXS for the monomeric species of both TonB constructs are shown in Figure 3.25A. $TonB_{CTD}$ shows an elongated globular structure in full agreement with the monomeric NMR solution structure (Figure 3.25B inset), whereas the Kratky plot for $TonB_{\Delta TMD}$ indicates a globular domain in combination with a disordered region (Figure 3.25A). By performing a Guinier approximation, the radius of gyration (R_g) was found to be 19 and 44 Å for $TonB_{CTD}$ and $TonB_{\Delta TMD}$ respectively. The large R_g value for $TonB_{\Delta TMD}$ agrees with the notion of the proline-rich linker domain forming an extended PPII conformation[100].

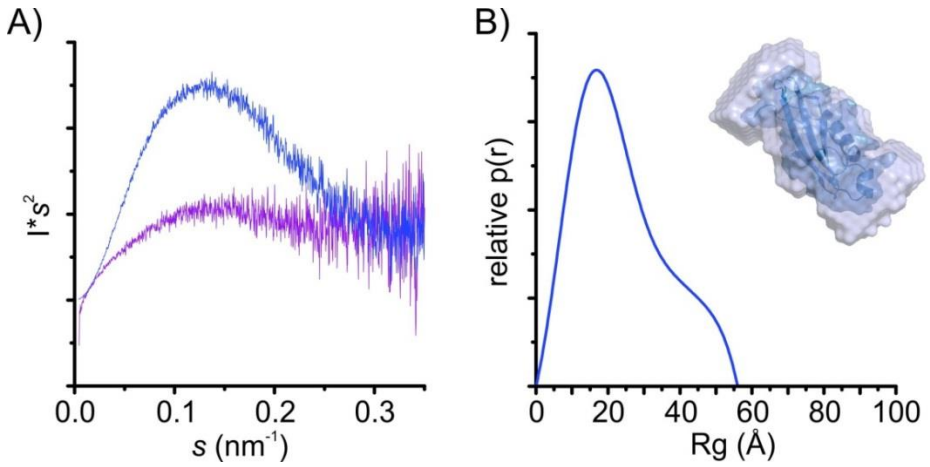


Figure 3.25 – Small angle X-ray scattering (SAXS) to determine structural information of the proline-rich linker domain of TonB. A) Kratky plot for both TonB constructs ($TonB_{\Delta TMD}$ (purple) and $TonB_{CTD}$ (blue)). The features of a globular protein are shown by the $TonB_{CTD}$ SAXS whilst $TonB_{\Delta TMD}$ shows a combination of globular and disordered domains as the plot displays a clear maximum, corresponding to the folded domain with a continuous rise produced by the presence of disordered regions. B) The pair-distance distribution (p(r)) of $TonB_{CTD}$, the additional shoulder of the distribution describes the elongated structure of the globular domain. The inset shows the scattering envelope (transparent light blue) over the NMR solution structure of $TonB_{CTD}$ (transparent blue surface and blue cartoon representation) which are in full agreement. The real space R_g from the p(r) (18.6 Å) matches the reciprocal space R_g (19 Å) from the Guinier approximation.

In order to test whether the proline-rich linker domain collapses in more saline conditions, SEC SAXS was performed on $TonB_{\Delta TMD}$ in the presence and absence of 0.5 M NaCl. By performing a Guinier approximation of the monomeric species, R_g values of 44 and 36 Å were found in the absence and presence of 0.5 M NaCl respectively. This suggests that under physiological saline conditions (128 mM NaCl) the proline rich linker domain is more extended, which collapses in increased saline conditions.

To gain more structural information about the conformation of the proline-rich linker domain from the SAXS, an ensemble optimisation method (EOM) was used (Section 2.2.6.7). EOM generates a large pool of random structures using sequence and structural information. A genetic algorithm then compares the averaged theoretical scattering intensity from *n* independent ensembles of conformations against the SAXS data and the conformations that best describe the experimental SAXS data (Figure 3.26B) are selected. The results show that under physiological conditions, $TonB_{\Delta TMD}$ populates a more extended conformation (indicative of the higher R_g distribution (Figure 3.26C)) where as under increased saline conditions, a higher frequency of low R_g dominates.

Figure 3.26D shows example conformations in physiological and high saline conditions that best describe the experimental SAXS data.

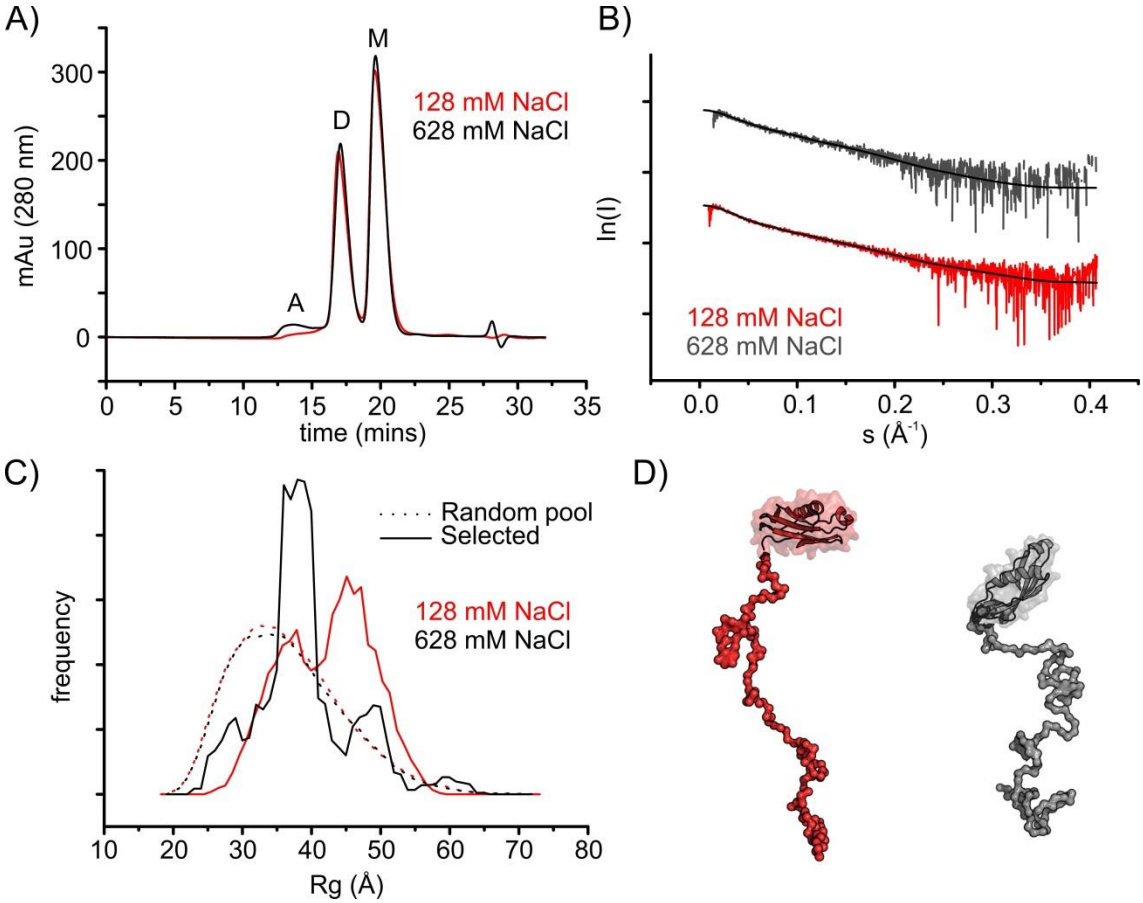


Figure 3.26 – SEC SAXS and EOM analysis of $TonB_{\Delta TMD}$ in the presence and absence of additional 0.5 M NaCl. A) Chromatogram of the gel filtration of $TonB_{\Delta TMD}$ in the absence (red) and presence (black) of additional 0.5 M NaCl. The two major peaks show dimeric (D) and monomeric (M) $TonB_{\Delta TMD}$ species with a small proportion of aggregates (A) (confirmed by SEC-MALLS, Figure 3.28A). B) The experimental SAXS with the fit from EOM analysis (black) obtained for an ensemble of 5 models (low salt, red) and 4 models (high salt, grey). C) R_g distributions from the EOM-selected ensemble (line) and that corresponding to the random pool (dashs) in the presence (black) and absence (red) of additional 0.5 M NaCl. D) Example conformations of the proline-rich linker domains from EOM analysis in low (red) and high (grey) salt concentrations.

The combined results from SMFS and SAXS suggest that under physiological saline conditions, the proline-rich linker domain is more extended and appears to contain structure when extended under force. In high saline conditions, the proline-rich linker domain is collapsed and structure is lost under force extension (summarised in Figure 3.27).

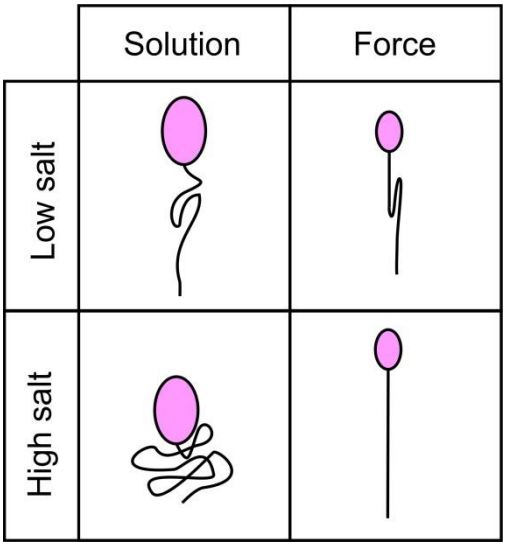


Figure 3.27 – Possible conformations of the proline-rich linker domain under different conditions. In low salt conditions, the proline-rich linker domain (black) is more extended, however when extended under force, the linker contains some structure. In high salt conditions, the proline-rich linker domain is collapsed and makes contact with the C-terminal domain (pink), however under force, the linker domain contains no structure and is a fully extended polypeptide chain.

3.2.9 Dimerisation of the TonB$_{\Delta TMD}$ construct

The dimerisation of TonB depending on the construct size has been well documented, for example truncating the protein within the structured C-terminal domain (residues 150-239) results in an intertwined dimer artefact (Figure 1.10A). TonB constructs are monomeric in solution until residues containing the proline-rich domain (PE/PK repeats) are included[196], which leads to the presence of a significant population of TonB dimer[106]. This indicates that TonB dimerisation is driven by the PE/PK region, to confirm that the two peaks in the chromatogram in Figure 3.26A where indeed monomer and dimer, SEC-MALLS was performed on TonB$_{\Delta TMD}$ which confirmed the presence of a monomer-dimer population in the concentrated (464 µM) sample (Figure 3.28A). The molecular mass of TonB$_{\Delta TMD}$ predicted from the static angle scattering was inaccurate by ± 10 kDa which was evident from the poor fit (R^2 = 0.84) to the data with the standard Zimm model (due to the elongated nature of TonB$_{\Delta TMD}$), however the doubling of mass in the dimer peak gave confidence that a TonB homodimer was present (Figure 3.28A). Interestingly, dynamic light scattering (DLS) of the monomer and dimer peak gave the same radius of hydration (R_H) (Figure 3.28B), suggesting that although a dimer is present, the globular C-terminal domain (visible by light scattering) do not contact each other.

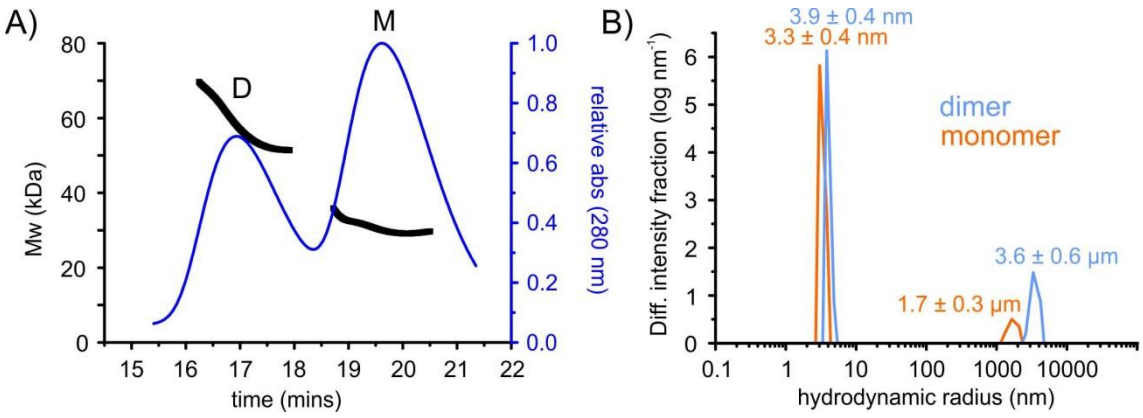


Figure 3.28 – SEC-MALLS of $TonB_{\Delta TMD}$. A) Chromatogram of TonB revealing dimer (D) and monomer (M). The calculated M_w from the static angle light scattering is shown in black. The true molecular weight of the $TonB_{\Delta TMD}$ construct is 23580 Da. B) Regularisation plot of the $TonB_{\Delta TMD}$ dimer (blue) and monomer (orange) with radius of hydration (R_H) shown. Each peak of the chromatogram contained a small quantity of µm sized particles (likely aggregates or microbubbles).

To find if $TonB_{\Delta TMD}$ dimerisation is concentration dependent, a concentration titration was performed on $TonB_{\Delta TMD}$ using gel filtration. $TonB_{\Delta TMD}$ elutes as three distinguishable peaks, the first being aggregates (confirmed by SEC-MALLS), followed by dimer and monomeric TonB (Figure 3.29A). Surprisingly, at concentrations between 1.5 mM to 25 µM (incubated at room temperature for 16 hours), the amount of dimer in the sample (calculated by the area under a Gaussian fit (Figure 3.29B)) remained constant (16.8 ± 3.9 % dimer) at all concentrations tested (Figure 3.29C), less than that previously predicted by EPR (40-50 %)[106]. Also the monomer-dimer conversion is dynamic as gel filtration of the isolated monomer or dimer fraction resulted in the reappearance of a double peak.

As $TonB_{\Delta TMD}$ dimer formation appears to be electrostatically driven by the PE/PK repeats within the proline-rich linker domain, gel filtration was performed at various concentrations of NaCl with 16 hour incubations at room temperature. Although a population of dimer was always present, the total percentage of dimer decreased with increasing NaCl concentration (Figure 3.29D). This was not observed in Figure 3.26A as the protein was buffer exchanged into higher NaCl during SEC with no incubation period, which implies the monomer-dimer equilibrium is slow.

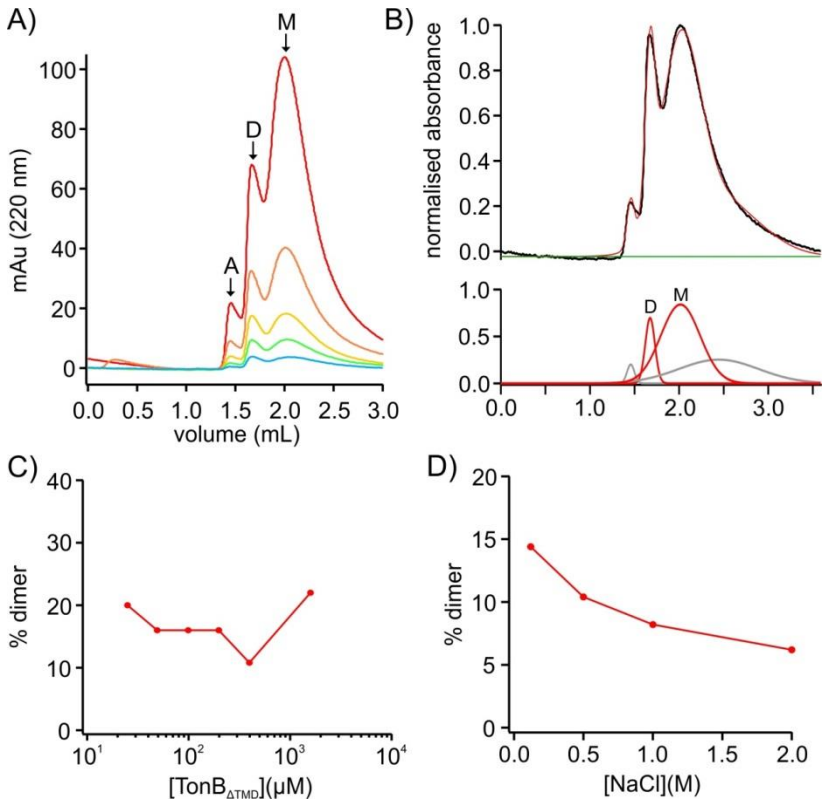


Figure 3.29 – Finding the effect of dimerisation of TonB$_{\Delta TMD}$. A) Concentration titration of TonB$_{\Delta TMD}$ (red = 397, orange = 198, yellow = 99, green = 49 and blue = 25 μM) on a Superdex™ 200 increase 3.2/300 gel filtration column. Aggregates (A), dimer (D) and monomer (M) are shown. B) Calculating the amount of dimer by fitting multiple Gaussians to the chromatogram (red). The sum of 4 Gaussians is shown in black. The individual Gaussians are shown below, the dimer (D) and monomer (M) Gaussians are labelled. C) The % dimer at various concentrations of TonB$_{\Delta TMD}$ calculated from the area under the Gaussians. D) The % dimer at various NaCl concentrations.

3.3 Discussion

3.3.1 Mechanical strength of the TonB:Ton box interaction

The results presented in this chapter have demonstrated several important findings about TonB and the interaction with the Ton box of a TonB-dependent receptor. Arguably, the most significant finding is that the interaction between TonB and the Ton box from BtuB displays huge mechanical resilience when extended by AFM. When the TonB$_{\Delta TMD}$:Ton Box$_{BtuB}$ interaction was extended at retraction velocities of 200 – 5000 nms^{-1} (corresponding to force loading rates of 10^3-10^5 pNs^{-1}) the resultant dynamic force spectrum reveals that rupture occurs at 84-132 pN (Figure 3.15).

These findings indicate that the TonB:Ton box complex is mechanically robust as predicted. The β-strand augmentation mechanism, which underlies the complex formation, is well-known for being mechanically "strong". For example, this motif is present in the pili subunit FimG where the extraordinary stable β-strand augmentation with neighbouring subunit FimF is necessary for immobilising uropathogenic *E. coli* to host endothelium under hydrodynamic flow.

The key difference between the β-strand augmentation of FimG:FimF and TonB:BtuB is in the latter the interaction is facilitated by the docking of a strand (the Ton box in this instance) to the leading edge of the pre-existing β-sheet of the protein (TonB). In the Fim complex, the β-strand ligand complements its binding partner by completing the core of the protein fold (known as donor strand complementation). These difference could be key in determining the strength of the interaction, where it is beneficial that the Fim complex never unbinds, whilst for TonB:Ton box, eventual unbinding is necessary to prevent the unfolding of the entire receptor.

These results provide the first evidence *in vitro* towards the pulling model of TonB-dependent transport, which will be developed further in the following chapter.

3.3.2 Conformation with the linker domain

An unexpected but interesting finding was that TonB populates two distinct conformations when extended by SMFS. Since the discovery of the proline-rich linker domain of TonB, little has been known about its structural organisation, other than it exists as a 15 nm extended PPII rod to allow the C-terminal domain of TonB to cross the periplasmic space[86]. Here we find that the proline-rich linker domain must contain a force-resistant structure as the observed contour length for the entire complex (PEG_{24}-Ton box:$TonB_{\Delta TMD}$-PEG_{24} =~44 nm)(Figure 3.18C) was much shorter than the expected contour length (~74 nm) (Figure 3.16). This suggests that the proline-rich linker domain does not exist as a purely unstructured domain or a continuous PPII helix. Similar force-resistant collapses within an unstructured polypeptide chain have also been observed in PolyQ chains (associated with Huntington's disease)[197]. The collapsed random globular states in PolyQ were found to be so mechanically resilient that no unfolding was observed when a force-clamp of 180 pN was used. MD simulation revealed that these

force-resistant globules were extensively H-bonded. The force-resistant structure within the linker domain of TonB could have implications in its function *in vivo* (see discussion below). The SAXS data in this chapter provides evidence of an extended proline-rich linker domain as previously shown by EPR spectroscopy[100]. It is interesting that the presence of an extra 0.5 M NaCl causes the collapse of this extended rod (indicative by the smaller R_g and EOM distribution), which implies that the elongated conformation is also electrostatically driven.

A hypothetical possibility is that the PE and PK repeats form a force-resistant electrostatic hinge, whilst the rest of the linker domain exists as a PPII rod (Figure 3.30). This force-resistant conformation gives an estimated linker length close to that observed in the SMFS experiments (~22 nm), and together with the C-terminal domain is almost sufficient to span the length of the periplasmic space of the mid cell (estimates of 30 ± 2 nm in rich growth medium[198]).

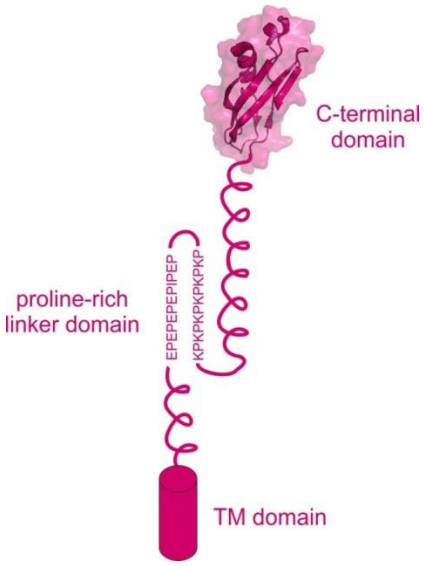


Figure 3.30 – Predicted model of the proline-rich linker domain of TonB. The helix shows a continuous PPII conformation.

The fact that the linker domain conformation is lost (or changes) in the presence of high NaCl concentrations (indicative by the more abundant longer contour length and chemical shift perturbations) supports the predicted model. The function of the hinge remains purely speculative, however it could represent an off- or on-state of TonB. For example, in the off-state, the linker domain contains no structure allowing TonB to

search for exposed Ton boxes. Upon binding to the Ton box, the *pmf* possibly drives the hinge conformation (possibly by interactions with ExbD) causing a 28 nm retraction of the C-terminal domain, driving a pulling force. The main evidence against this hypothesis is that TonB(Δ66-100) which lacks the PE/PK repeats is still functional *in vivo* under physiological growth conditions[99]. There is also the possibility that the PE/PK driven conformation acts as a length adaptor, where the salt-dependent hinge opens allow TonB to reach the OM when the bacteria are exposed to more saline growth conditions (resulting in an expanded periplasm).

3.3.3 TonB dimer conformation

The construct used in this study ($TonB_{\Delta TMD}$: residues 33-239) was found to form dimers where $TonB_{CTD}$ (residues 144-239) did not, in accordance with other studies[86,106,196]. This strongly suggests that TonB dimers are formed by the proline-rich linker domain (specifically residues 33-112 which contain the PE/PK repeats). Although it is accepted that TonB is monomeric when in contact with the receptor[144,196], there is sufficient evidence by formaldehyde cross-linking *in vivo* that TonB forms homodimers at some stage during transport[103,105,199]. The combination of SEC-MALLS, SAXS and analytical gel filtration in this thesis has revealed a possible conformation of the TonB homodimer. The R_H values calculated by MALLS for dimeric TonB were the same as monomeric TonB, suggesting that the structured C-terminal domains in TonB homodimers do not contact, as the R_H would double. The R_g approximation from the SAXS on the dimeric TonB (50 Å) was only slightly larger than the monomer (44 Å), indicating that a head-on-head dimer is not present, as doubling of the R_g would be expected. Therefore it is likely that the dimer is aligned so that the proline-rich linker domains are parallel. Analytical gel filtration demonstrated that dimer formation is electrostatically driven, as the amount of dimer decreases with increasing NaCl conditions; it is therefore likely that the PE/PK repeats form the TonB homodimer. A predicted model of the TonB homodimer from these observations is shown in Figure 3.31.

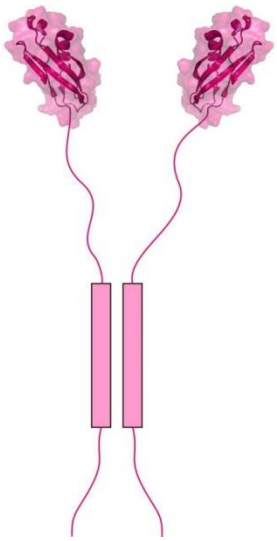


Figure 3.31 – Predicted structure of the TonB homodimer from MALLS and SAXS. The structure of the monomeric C-terminal domain is shown, whilst the proline-rich region (PE/PK repeats) are shown by the pink boxes.

$TonB_{CTD}$ is purely monomeric, and has the same binding constant to the Ton box_{BtuB} peptide as $TonB_{\Delta TMD}$ (which contains a population of dimers) (Figure 3.8), this indicates that the homodimer formation does not affect the binding function of the C-terminal domain to the Ton box.

4 Testing the pulling model mechanism of TonB-dependent transport

4.1 Objectives

In the previous chapter the unbinding of the Ton box from BtuB (Ton box_{BtuB}) was found to require a magnitude of force previously shown to unfold many globular proteins[200]. In this chapter the natively folded TonB-dependent receptors BtuB and FhuA will be isolated from the OM of *E. coli* and reconstituted into *E. coli* lipid extract liposomes. SMFS will then be used to test whether TonB can remodel the plug domain by the Ton box tether under mechanical force, therefore directly testing the pulling model of TonB-dependent transport *in vitro*.

4.2 Results

4.2.1 Cloning FhuA from *E. coli*

An arabinose inducible expression plasmid (pBAD) containing the wild-type *btuB* gene from *E. coli* (pNGH15) was kindly provided by Dr Nicholas Housden, University of Oxford. The ferric hydroxamate uptake receptor (*fhuA*) gene was cloned from the genome of *E. coli* (strain: JM109) using Q5® PCR (Section 2.2.1.3) and the DNA primers shown in Table 4.1.

FhuA	Forward	5′CATTAATCACCATGGCGCGTTCCAAAACTGC 3′
	Reverse	5′TTGAATATTCTCGAGTTAGAAACGGAAGGTTGCGGTTG 3′

Table 4.1– DNA primers for the cloning of the fhuA gene from *E. coli.* The underlined text shows the restriction sites (*XhoI* and *NcoI*), green text is the start codon, red text is the stop codon and grey text is a random sequence preceding and following the start and stop codons respectively.

The amplified gene product (2270 bp) was confirmed by agarose gel electrophoresis (Figure 4.1A) and extracted using the method described in Section 2.2.1.9. The BtuB expression plasmid and amplified *fhuA* gene were cleaved with *XhoI* and *NcoI* restriction enzymes (Method described in Section 2.2.1.5)(Figure 4.1B) and the digested plasmid was purified using agarose gel electrophoresis (Figure 4.1B). The cleaved *fhuA* gene was then ligated into the pBAD vector, transformed into DH5α cells and plated onto ampicillin selection plates. After incubation for 16 hours at 37 ˚C, several colonies were

randomly selected, a restriction digest was performed on the purified plasmids and *fhuA* insertion was confirmed by agarose gel electrophoresis (Figure 4.1C). The constructs were further confirmed by DNA sequencing.

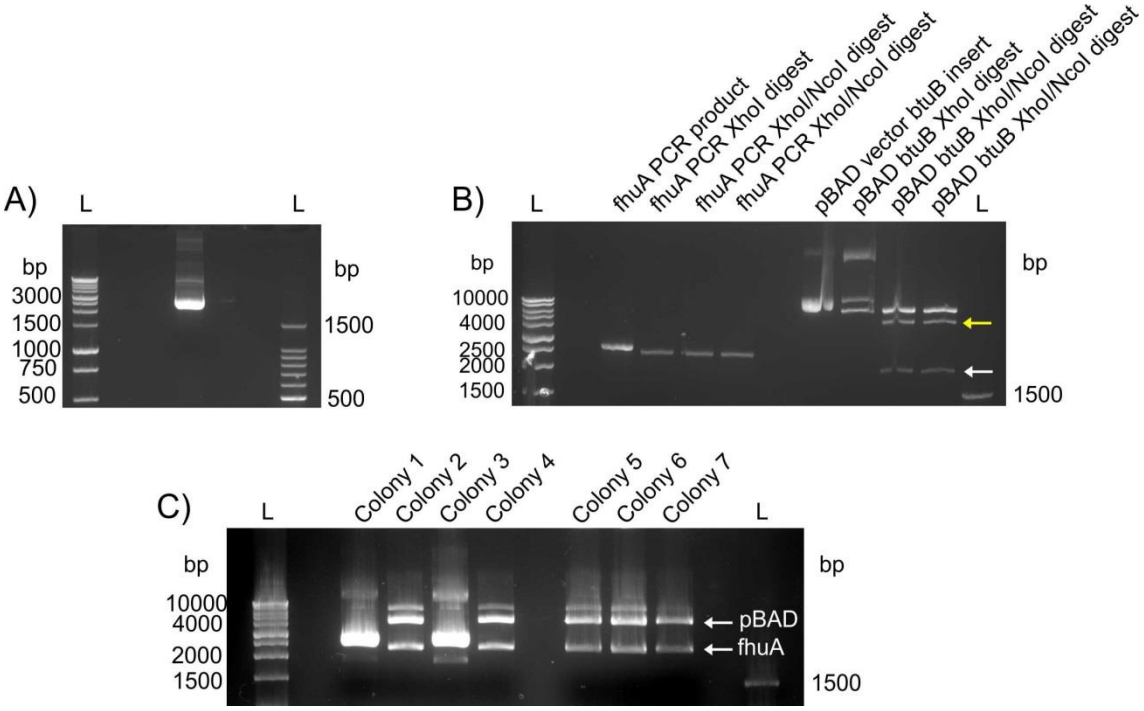


Figure 4.1 – Agarose gels stained with ethidium bromide showing the stages of cloning the *fhuA* gene product from *E. coli* into an arabinose inducible expression plasmid (pBAD). A) The amplification of the *fhuA* gene (2270 bp) by PCR (L is the DNA ladder). B) Restriction digest of the amplified *fhuA* gene and pNGH15 (pBAD vector with *btuB* gene insert). The yellow arrow indicates the position of the digested pBAD vector without insert, and the white arrow indicates the position of the *btuB* gene. C) After ligation and transformation into competent cells, single colonies were selected and the plasmids isolated and digested by *XhoI*/*NcoI* restriction enzymes before agarose gel electrophoresis was used to confirm the insertion. The labelled arrows show the correct band for the pBAD vector and *fhuA* gene.

4.2.2 Over-expression and purification of TonB-dependent receptors

4.2.2.1 BtuB purification

BtuB was overexpressed from the pBAD vector (pNGH15, kindly provided by Nicholas Housden, University of Oxford) using a freshly transformed porin-deficient strain of *E. coli* (TNE012 (K12 *tsx*$^{-}$ *ompA*$^{-}$ *ompB*$^{-}$), also provided by Dr Nicholas Housden) using the method described in Section 2.2.2.3. After protein induction by 0.15 % (w/v) L-(+)-arabinose and incubation for 16 hours at 15 ˚C, the cells were pelleted and lysed using a sonicator (Section 2.2.2.4). The supernatant of the lysed cells was ultracentrifuged at

45,000 rpm to isolate the membrane fraction, inner membrane contaminates were removed with Triton-X 100 (Figure 4.2A) and the outer membrane pellet was solubilised into a solution containing n-octyl β-D-glucopyranoside (β-OG)(See Section 2.2.2.6). The solubilised OMPs were loaded onto a HiTrap DEAE FF 5 mL anion exchange column (Section 2.2.2.6) and impurities were removed with a 1 M LiCl gradient (Figure 4.2B). BtuB elutes from the column when 100 % of the 1 M LiCl buffer is applied (Figure 4.2B, peak 3 and Figure 4.2C) and forms a white precipitate containing pure BtuB. BtuB was allowed to precipitate for 16 hours at 4 ˚C before centrifugation at 4,000 rpm to pellet the precipitated protein. This was then stored at 4 ˚C until further use. Typically from a 10 L growth, ~5 mg of BtuB was isolated.

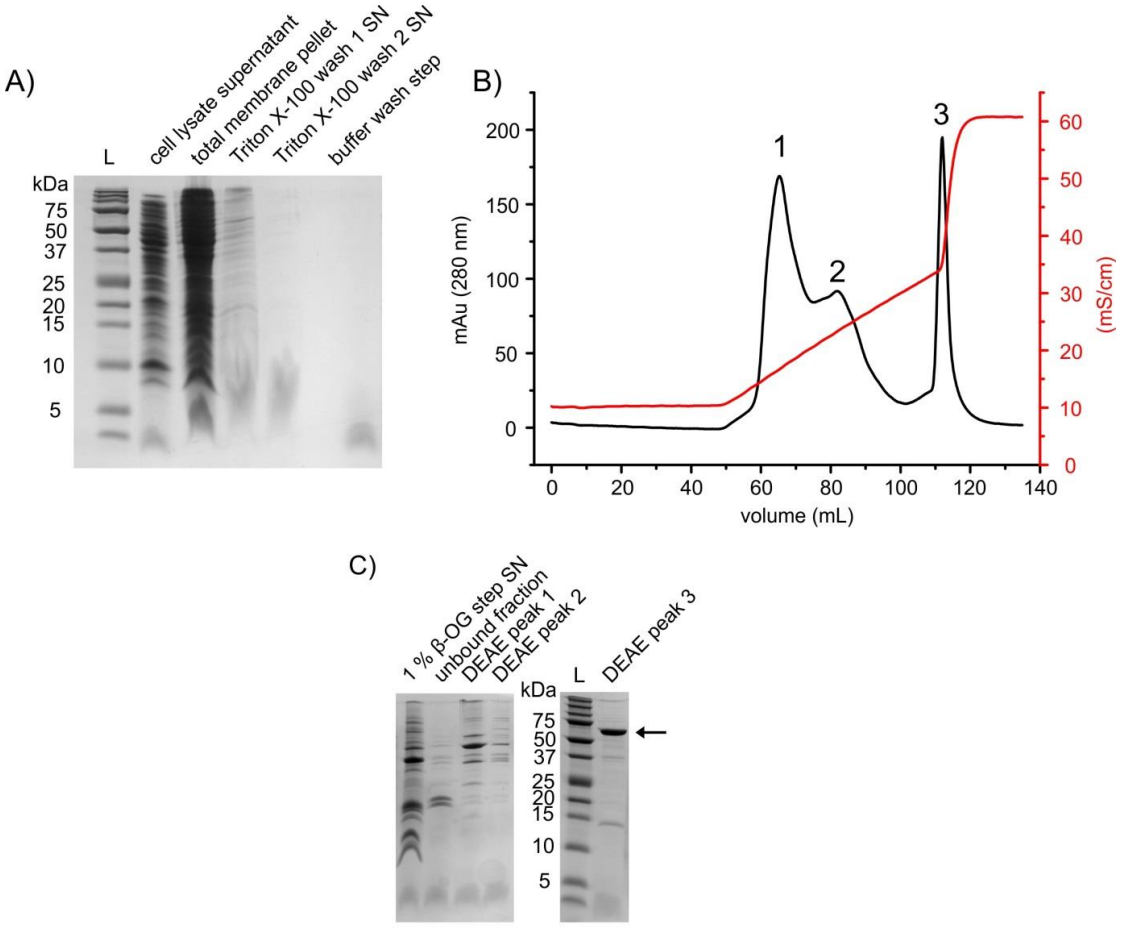


Figure 4.2 – Purification of BtuB. A) SDS-PAGE of the initial centrifugation purification steps (SN = supernatant), L is the protein ladder B) Anion exchange chromatography using a 5 mL HiTrap DEAE FF column of the solubilised OM fraction, the impurities eluted in peaks labelled 1 and 2, and pure BtuB elutes at 100 % 1 M LiCl in peak 3. C) SDS-PAGE of the peaks during anion exchange, the expected mass (66 kDa) of BtuB is shown by the black arrow.

4.2.2.2 FhuA purification

The purification of FhuA was carried out using the same protocol as for BtuB (using the same expression vector and porin-deficient strain of *E. coli*). After the OM pellet was solubilised in 1 % (w/v) β-OG, SDS-PAGE confirmed the presence of FhuA (Figure 4.3A) in the supernatant. During anion exchange chromatography using a 5 mL HiTrap DEAE FF column, FhuA eluted from the column with the impurities during the 1 M LiCl gradient (peak 1 Figure 4.3B and Figure 4.3C). Fortunately, FhuA began to precipitate in a similar fashion to BtuB (due to the β-OG concentration being slightly below the CMC). The precipitant was isolated by centrifugation, solubilised in 1 % (w/v) β-OG and a Superdex™ 75 HR 10/30 gel filtration column equilibrated in 25 mM Tris-HCl pH 8.0, 128 mM NaCl, 1 % (w/v) β-OG was used to further purify FhuA (Figure 4.3D and E). The total yield of protein was ~ 2 mg.

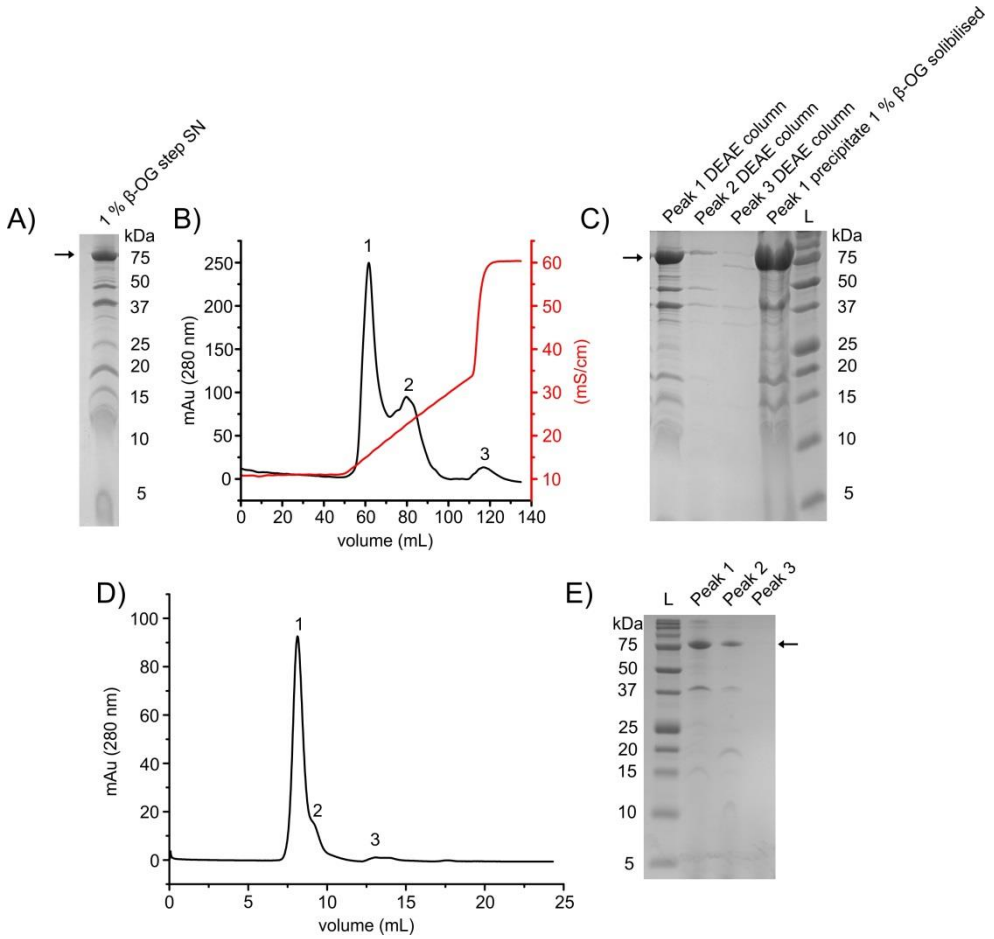


Figure 4.3 – Purification of FhuA. A) SDS-PAGE of the solubilised OM fraction (in 1 % (v/v) β-OG) with the expected mass (79 kDa) of FhuA shown by the black arrow. B) Anion exchange chromatography of the solubilised OM fraction using a 5 mL HiTrap DEAE FF column, FhuA eluted with impurities in peak 1. C) SDS-PAGE of the labelled peaks from anion exchange (B) and the precipitate formed. The black arrow shows the expected position of FhuA. D) Gel filtration of the solubilised FhuA precipitate using a Superdex™ 75 HR 10/30 gel filtration column, FhuA eluted in the major peak (peak 1). E) SDS-PAGE of the three absorbance peaks from gel filtration (D), the black arrow shows the expected position of FhuA.

4.2.2.3 Spectroscopic analysis of protein fold

To show that the purified TBDTs were correctly folded when solubilised in 1 % (w/v) β-OG, secondary structure analysis was carried out using Far-UV CD (Section 2.2.6.1) and the tertiary structure was investigated using intrinsic tryptophan fluorescence (Section 2.2.6.2). The CD spectra demonstrated that both BtuB (Figure 4.4A) and FhuA (Figure 4.4C) have a predominant β-sheet secondary structure, which is in agreement with the X-ray crystal structures of BtuB (Figure 1.18C), and FhuA (Figure 1.12). Tryptophan fluorescence emission (excitation 280 nm, emission 290-400 nm) was used to monitor

the tertiary fold in the presence and absence of 8 M urea for both BtuB and FhuA. Upon the addition of 8 M urea, 15 minutes of equilibration was required before complete BtuB unfolding was observed (Figure 4.4B), demonstrating the high kinetic stability of β-barrel proteins. The decrease in fluorescence and red shift in 8 M urea in both BtuB and FhuA (Figure 4.4B and D) is indicative that the proteins had unfolded, therefore in the absence of denaturant; BtuB and FhuA contain tertiary structure.

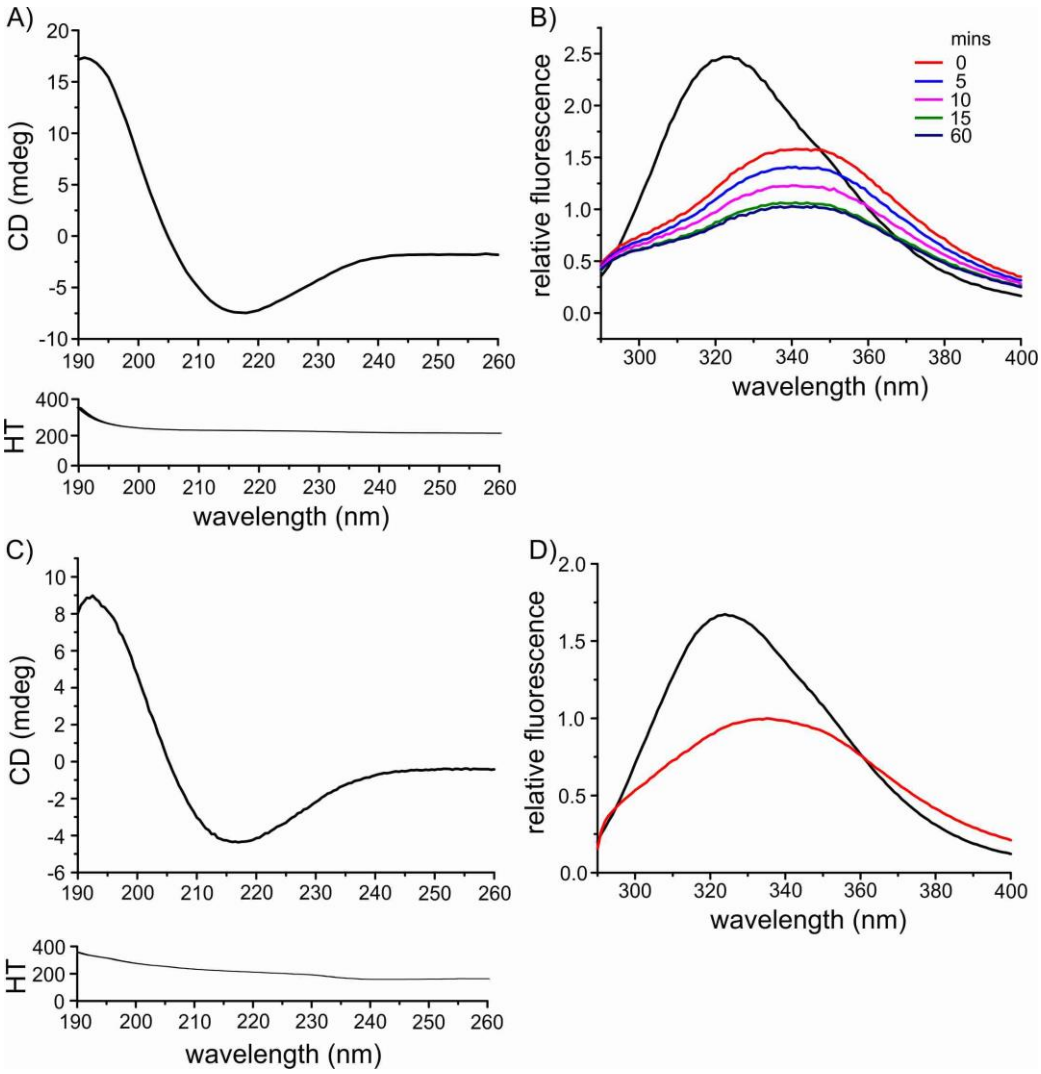


Figure 4.4 – Monitoring the secondary and tertiary fold of BtuB and FhuA after purification and solubilisation into detergent. A) Far-UV CD spectrum of BtuB (below shows a plot of the high tension voltage (HT), values greater than 500 indicate detector saturation). B) Tryptophan emission spectra (from an excitation at 280 nm) of BtuB (black trace). The protein was then incubated in 8 M urea and the various time points are shown (inset, coloured traces). The protein was fully unfolded after 15 minutes. C) Far-UV CD spectrum of FhuA D) Tryptophan emission spectrum (from an excitation at 280 nm) of FhuA (black) and after 1 hour incubation in 8 M urea (red).

4.2.3 Insertion of TonB-dependent transporters into proteoliposomes

To perform the SMFS experiments, the TBDTs were densely packed into proteoliposomes composed of *E. coli* polar lipid extract using a dialysis method previously described for the insertion of FhuA into liposomes for an AFM unfolding study[178] (Section 2.2.8). After dialysis, a gel-like sediment (large proteoliposomes) formed at the bottom of the dialysis bag; this sediment was gently suspended by pipetting and stored at 4 ˚C. To assess whether the proteoliposomes had formed and contained the protein (all analysis was performed on BtuB proteoliposomes) SDS-PAGE, DLS and AFM imaging were carried out. To assess whether the protein was indeed inserted into liposomes, a sample from the dialysis was boiled in SDS for 5 minutes whilst another sample was left at room temperature. These samples were then loaded onto a SDS-PAGE gel and electrophoresed. In the sample that was not boiled, the protein could not enter the stacking gel due to being trapped inside the large stable proteoliposomes, whilst boiling allowed the protein to be released from the proteoliposomes and enter the gel with the correct electrophoretic mass (66 kDa) (Figure 4.5).

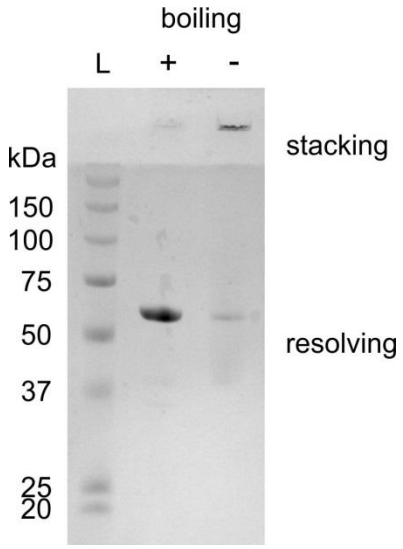


Figure 4.5 – SDS-PAGE of BtuB inserted into proteoliposomes composed of *E. coli* polar lipid extract using the dialysis method. The protein enters the resolving gel after boiling in SDS, indicating proteoliposomes had been disrupted and BtuB unfolds and has electrophoretic mobility as expected for its mass (66 kDa). In the unboiled sample, the large stable liposomes prevent the protein entering the resolving gel where it becomes trapped in the stacking gel.

Dynamic light scattering (DLS) measures the scattered intensity over a range of different angles, which can be used to reveal particle hydrodynamic radius distribution between

0.001 to several microns in solution. Therefore DLS was used in order to measure the size of the proteoliposomes formed from the dialysis method. A sample of the suspended proteoliposomes was diluted to give a final lipid concentration of approximately 40 µM. Using the method described in Section 2.2.6.3, DLS was performed on the diluted sample. The correlation curves accumulated over 3 minutes were used to obtain a distribution of particle sizes (Figure 4.6, black trace). The distribution of the hydrodynamic radii for the proteoliposomes indicates that there is a large variation in size in the population (100-2000 nm), which was expected as the proteoliposomes were not extruded through a filter. For a negative control, the proteoliposomes were suspended in 1 % (w/v) SDS and boiled for 5 minutes to disrupt the proteoliposomes before DLS was performed. Upon boiling in detergent, the large structures are lost (Figure 4.6, red trace), which provides evidence of the presence of large proteoliposomes in the sample.

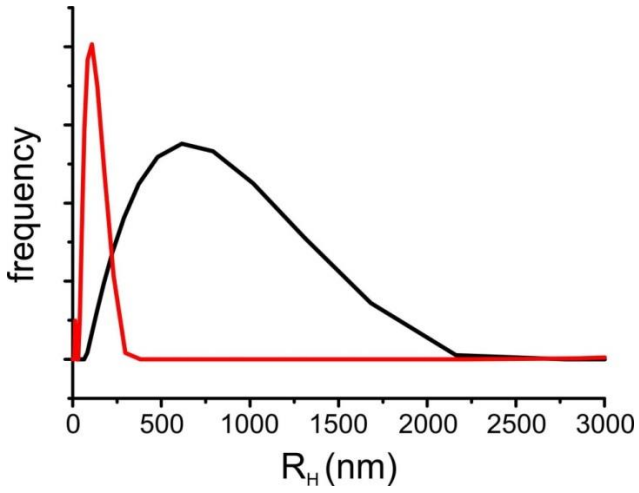


Figure 4.6 - Distribution of the hydrodynamic radius of vesicles of *E. coli* polar lipid extract containing BtuB obtained by dynamic light scattering (black). Addition of SDS and boiling caused breakdown of liposomes (red).

Finally, AFM imaging was performed on the proteoliposomes rolled onto mica to further confirm that the receptor was densely packed into a lipid bilayer and was surface immobilised when placed on mica. The proteoliposomes were resuspended by pipetting and diluted to a final lipid concentration of 0.1 mg/mL. 200 µL was pooled over a 1 cm diameter disc of freshly cleaved mica and incubated at room temperature for 30 minutes. The surface was then thoroughly rinsed with 25 mM Tris-HCl pH 8.0, 128 mM NaCl. AFM imaging with a FastScan AFM (Bruker) equipped with a Fast Scan (D probe)

cantilever (Bruker) was conducted at a scan frequency of 80 kHz which confirmed the presence of membrane packed with BtuB immobilised to the surface (Figure 4.7).

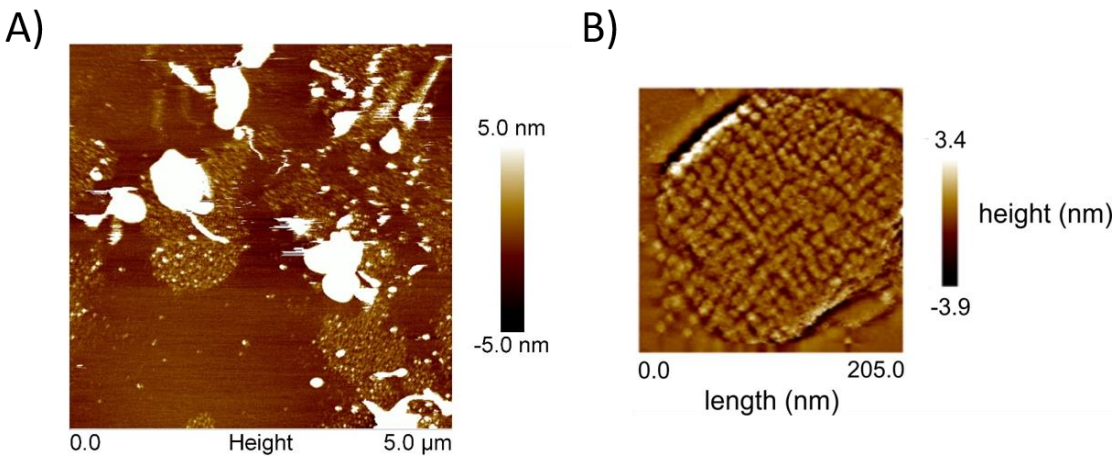


Figure 4.7 – AFM imaging of BtuB proteoliposomes rolled onto mica. A) Low magnification of BtuB proteoliposomes on mica, the large white areas are likely unwashed vesciles. B) Higher magnification of BtuB proteoliposomes on mica (image taken by Jonny Burns, University of Sheffield).

4.2.4 Single-molecule force spectroscopy of the $TonB_{\Delta TMD}$:BtuB interaction

As shown in Chapter 3, the $TonB_{\Delta TMD}$:Ton box_{BtuB} complex is able to resist forces that are usually associated with the unfolding of moderately mechanically robust proteins at similar loading rates. Therefore it is likely that the plug domain, which is contiguous with the Ton box, may fully or partially unfold. To directly test this hypothesis, the BtuB proteoliposomes were surface immobilised onto a freshly cleaved mica (Section 2.2.8) and $TonB_{\Delta TMD}$(V32C) was immobilised to the AFM probe (Section 2.2.4). Figure 4.8 depicts a schematic of the system.

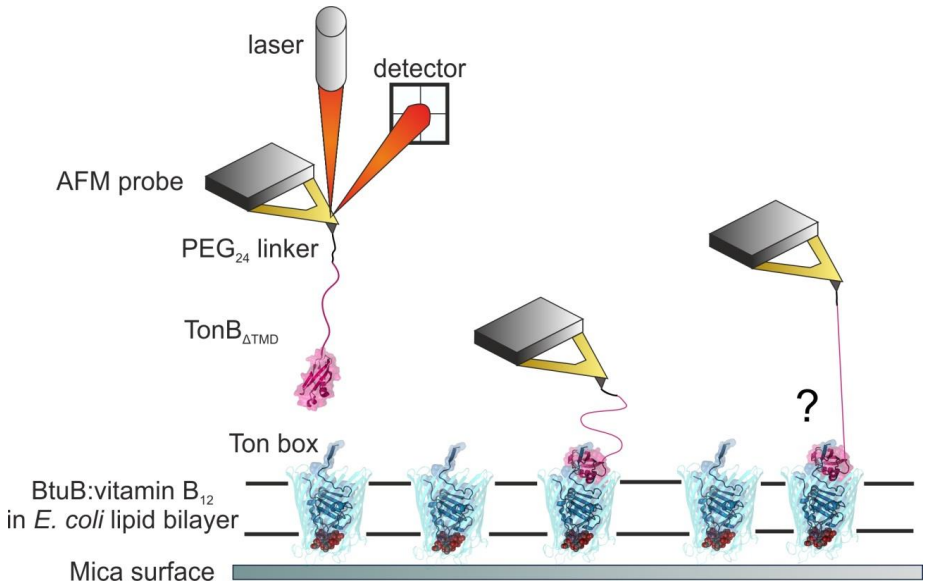


Figure 4.8 - Schematic to describe the single-molecule force spectroscopy (SMFS) approach used in this study. $TonB_{\Delta TMD}$ (pink) is attached to the AFM probe by heterobifunctional PEG_{24} linkers and BtuB (blue cartoon structure) bound to vitamin B_{12} (red) is inserted into *E. coli* polar lipid extract liposomes and immobilised to a mica surface. The forces between the interacting partners and any remodelling will be measured using an AFM.

Pressing the $TonB_{\Delta TMD}$-labelled AFM probe against the surface with 60-90 pN of force, dwelling for 1 second and withdrawing at a velocity of 1000 nms^{-1} (force loading rate ~10 nNs^{-1}) led to distinctive force-extension profiles displaying two rupture events (Figure 4.9A) at forces of 61 ± 4 and 91 ± 23 pN and contour lengths of 58 ± 3 and 77 ± 7 nm (Table 4.2). These events indicate that the complex undergoes partial unfolding before dissociation. As $TonB_{\Delta TMD}$ does not unfold at these forces, force would presumably be propagated into the plug domain via the contiguous Ton box tether. These data thus suggest that the strength of intra-polypeptide interactions in the plug domain are slightly weaker than those of the inter-polypeptide interactions between $TonB_{\Delta TMD}$ and BtuB, allowing unfolding before unbinding to occur in the majority of cases. As the mechanical strength of proteins and their complexes is kinetically controlled, the loading rate may affect unfolding and unbinding differently. Interestingly, analysis of the speed dependence of the putative unfolding and unbinding events demonstrates that the relative strength of intra- and inter-protein interactions at the same loading rate remains constant across the dynamic range of experiment (Figure 4.9B). For a negative control, SMFS of a BtuB L8P variant was performed which produced no detectable events. Interestingly, SMFS of wild-type BtuB in the absence of

vitamin B_{12} resulted in events at a similar rate to that of the substrate bound receptor, which implies that the Ton box of BtuB was always accessible for TonB binding, independent of the substrate.

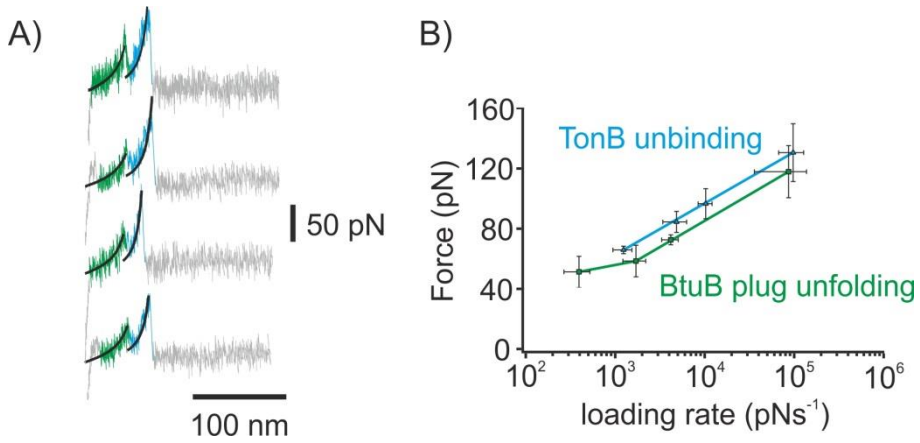


Figure 4.9 – Pulling on the $TonB_{\Delta TMD}$:BtuB complex in the presence of vitamin B_{12}. A) Representative force-extension profiles show two rupture events (the leading edge of each event is coloured green and blue, respectively) fitted to the WLC model (black). B) Dynamic force spectrum of unfolding (green) and unbinding (blue) events at 200, 500, 1,000 and 5,000 nms^{-1}, error bars show the range of the mode of triplicate data sets.

To quantify the extent of unfolding, force-extension profiles showing two events were analysed further by fitting the worm-like chain model (Section 1.5.2) to the unfolding (L_c1) and unbinding (L_c2) events, as the difference in these values (ΔL_c) reveals how much of the protein complex becomes unfolded at the first rupture event (Figure 4.10A). These data in the form of a histogram are shown in Figure 4.10B and the distribution for ΔL_c (red histogram) reveals a modal value of 20 ± 1 nm that is independent of the retraction velocity used (Table 4.2), or importantly, the presence or absence of the TonB linker domain (Figure 4.11 and Table 4.3). As the length of an extended amino-acid is known (0.4 nm) this distance can be directly translated to the unfolding of 50 amino-acids, approximately half of the residues within the plug domain (Figure 4.10B inset and Figure 4.10D).

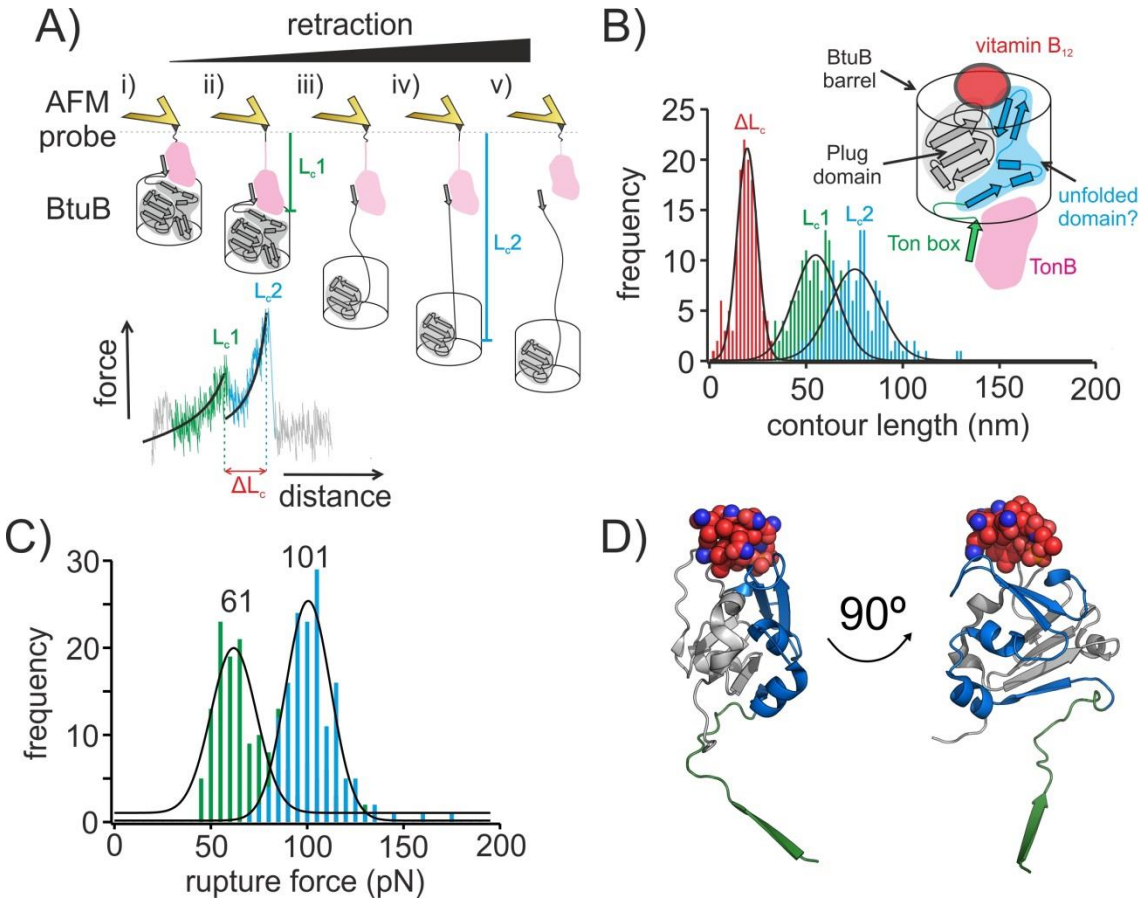


Figure 4.10 – Measuring the amount of plug domain unfolding. A) Schematic describing the origin of each rupture event (an example trace with WLC fits is shown below) and their relationship to L_c1, L_c2 and ΔL_c. The barrel of BtuB is shown as an open cylinder with the secondary structure of the plug domain (grey shaded region) and TonB, which is attached to the AFM probe, shown in pink. i) The AFM probe is pressed against the surface bringing the TonB:BtuB complex together; ii) TonB bound to BtuB is fully extended causing the force to increase; iii) part of the plug domain is unfolded and the force sharply decreases; iv) the unfolded region is extended causing an increase in the entropic restoring force and v) Ton Box_{BtuB} dissoicates from TonB and the force sharply falls to zero. The difference in fitted contour lengths (ΔL_c) between the two peaks can be used to calculate the extent of unfolding. B) Contour length frequency distributions from WLC fits to double-peaked force extension profiles from a triplicate datasets (n = 154) accumulated at a retraction velocity 1,000 nms^{-1}. Frequency distributions obtained for the remodelling/unfolding event (L_c1, green), unbinding (L_c2, blue) and ΔL_c (red) are fitted to Gaussian functions (black lines). Inset: the presence of an unfolding event prior to dissociation reveals that the plug domain comprises a mechanically weak sub-domain (the 50 amino acids downstream of the Ton box (residues 23-73, shaded blue)) and a mechanically recalcitrant sub-domain (shaded grey). Residues that extend prior to the first event are coloured green. Ton Box_{BtuB} is shown as a green arrow bound to TonB (shaded in pink). C) Force histograms for the plug unfolding (green) and TonB unbinding (blue) from the Ton box_{BtuB} with the modal values shown. D) The crystal structure of the BtuB plug domain with unfolded residues shown in blue, Ton box in green and vitamin B_{12} as coloured spheres.

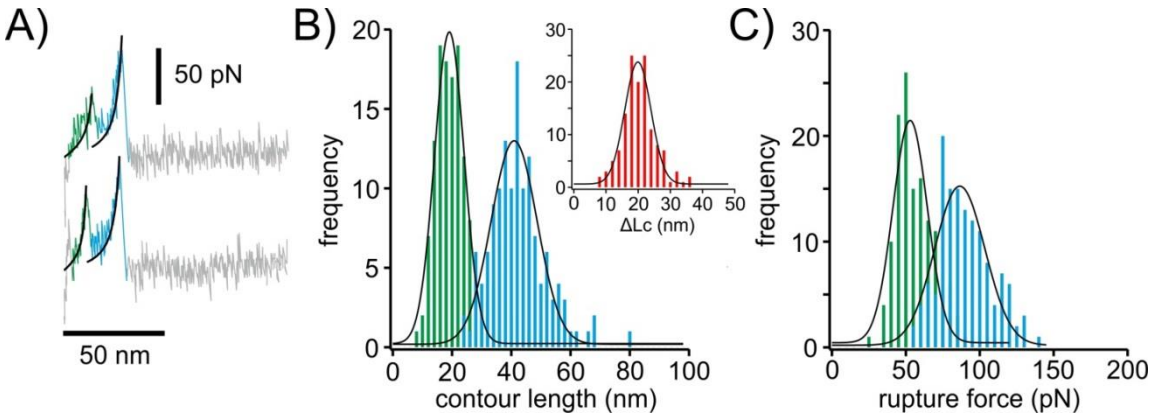


Figure 4.11 - SMFS results of wild-type BtuB bound to vitamin B_{12} unfolding using $TonB_{CTD}$(K144C) construct at 1000 nms^{-1}. A) Example force-extension traces of $TonB_{CTD}$:BtuB:vitamin B_{12} fitted to the worm-like chain model (p= 0.4 nm). B) Contour length distributions of unfolding (green) and unbinding (blue) and ΔLc (red inset), modes = 17, 40 and 21 nm respectively. C) Force distributions of unfolding (green) and unbinding (blue), modes = 53 and 88 pN respectively.

					Plug unfolding				TonB unbinding					
BtuB wt:	#	N	n	Hits	Force	Average	L_c	Average	Force	Average	L_c	Average	ΔL_c	Average
$TonB_{\Delta TMD}$		attempts	events*	(%)	(pN)		(nm)		(pN)		(nm)		(nm)	
200 nms^{-1}	1	2400	33	1.4	62		54		68		80		21	
	2	4300	42	1.0	41	51 ± 10	55	53 ± 4	63	66 ± 3	79	76 ± 7	19	19 ± 2
	3	2237	19	0.8	51		49		66		69		18	
500 nms^{-1}	1	1200	30	2.5	70		51		77		80		26	
	2	1050	27	2.6	57	59 ± 10	48	53 ± 6	90	85 ± 8	67	76 ± 9	20	22 ± 4
	3	1167	24	2.0	49		59		87		81		20	
1,000	1	2400	54	2.2	63		58		105		70		20	
nms^{-1}	2	1200	18	1.5	62	61 ± 4	55	58 ± 3	68	91 ± 23	80	77 ± 7	21	20 ± 2
	3	1948	81	4.1	57		61		100		82		18	
5,000	1	1000	17	1.7	130		41		125		75		24	
nms^{-1}	2	1200	11	0.9	98	118 ± 20	43	45 ± 5	115	131 ± 21	61	71 ± 10	25	24 ± 2
	3	1200	44	3.7	125		50		152		76		22	

Table 4.2 - AFM data of $TonB_{\Delta TMD}$:wild-type BtuB unbinding at various retraction velocities. Force, contour length (L_c) and ΔL_c values were calculated by plotting the distribution of events and fitting with a Gaussian to obtain the mean value. Hits denotes the percentage of approach/retract cycles that result in force-extension profiles that meet the criteria discussed in Section 2.2.7.4. The error on the average is the range of the triplicate averages. *number of double events.

					Plug unfolding				TonB unbinding					
Variant	#	N attempts	n events	Hits (%)	Force (pN)	Average	L_c (nm)	Average	Force (pN)	Average	L_c (nm)	Average	ΔL_c (nm)	Average
BtuB:$TonB_{CTD}$	1	9815	69	0.7	59		17		94		40		21	
(1000 nms^{-1})	2	4617	42	0.9	50	52 ± 7	21	19 ± 2	82	86 ± 8	43	41 ± 2	21	20 ± 2
	3	8550	26	0.3	48		20		82		40		18	

Table 4.3 - AFM data of $TonB_{CTD}$:wild-type BtuB at 1000 nms^{-1}. Force, contour length (L_c) and ΔL_c values were calculated by plotting the distribution of events and fitting with a Gaussian to obtain the mean value. Hits denotes the percentage of approach/retract cycles that result in force-extension profiles that meet the criteria discussed in Section 2.2.7.4. The error on the average is the range of the triplicate averages.

4.2.5 Using disulfide bridges to pin point the unfolding region within the BtuB plug domain

While these data reveal the extent of domain unfolding it does not reveal its location. To locate this region, it was reasoned that the 50 residues directly downstream of the Ton box motif (residues 23-73) would most likely unfold first (see Figure 4.10D) as these residues are initially exposed to mechanical deformation as force is propagated through the plug domain.

To test this assumption, disulfide bridges were created within the plug domain of BtuB. A variant designated XL_{barrel} (Figure 4.12 schematic) was designed to prevent any unfolding by linking the N-terminus of the plug domain to the barrel (BtuB L23C/S374C). A V29C/V45C mutant (XL_{loop}) was designed to create a covalently linked 15 amino acid loop within the region predicted to unfold which if this hypothesis is correct, would shorten the ΔL_c by ~6 nm. A final variant (T27C/Y109C) was created (designated XL plug subdomain (XL_{PSB}), Figure 4.12) linking the first β-strand in the predicted mechanically weak subdomain to the mechanically strong subdomain, diverting the force propagation pathway.

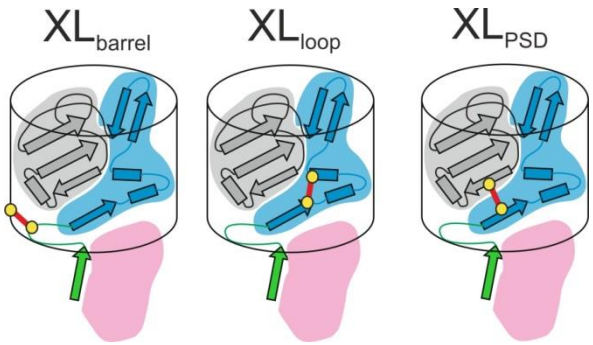


Figure 4.12 - Secondary structure topology of the BtuB plug domain showing location of the disulfide bridges (bold red line linked by yellow circles).The mechanically weak and strong sub-domains are shaded blue and grey, respectively, the barrel domain as an empty cylinder and Ton Box_{BtuB} is shown as a green arrow bound to TonB (shaded in pink).

The BtuB variants were all expressed and purified using the same method as for the wild-type receptor (see Section 2.2.2.6) and, after purification, adoption of the correct fold was analysed using far-UV CD and intrinsic tryptophan fluorescence. Both techniques

demonstrated that the cross-link variants displayed negligible differences in the λ-max from the wild-type receptor, however the slight change in intensity could reflect a change in the local tryptophan environment around the point mutations (Figure 4.13).

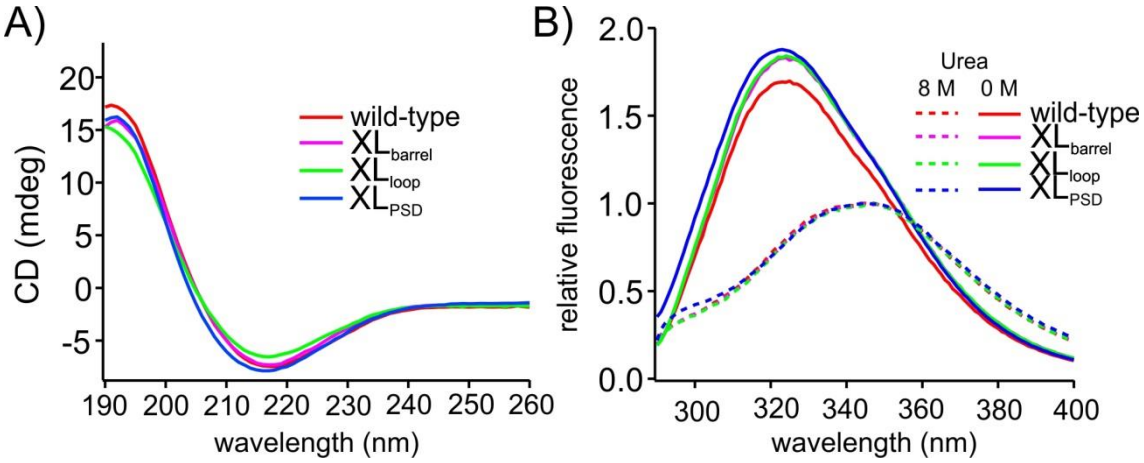


Figure 4.13 – Confirming the correct secondary and tertiary structure of the disulfide cross-link mutants. A) Far-UV CD spectra of wild-type BtuB and variants. B) Intrinsic tryptophan emission spectra (excitation 280 nm) of BtuB and variants in the absence (solid line) and presence of 8M urea (dashes).

SMFS was performed using the same experimental procedures as for the wild-type BtuB receptor with $TonB_{\Delta TMD}$(V32C) (Section 2.2.7.3). For the XL_{barrel} variant, greater than 74 % of the force-extension profiles analysed contained single rupture events with an F_U of 100 ± 6 pN (Figure 4.14B) at a L_c of 47 ± 2 nm (Figure 4.14A). The L_c value is in accord with the distance expected for the rupture of the complex without remodelling (based on the length of the first rupture event in the wild-type experiments) while the F_U value is close to that observed for the dissociation of the minimal Ton Box_{BtuB}:TonB complex observed previously at the same retraction velocity (1 μms^{-1}, 113 ± 13 pN). Addition of reductant (2 mM DTT) to the buffer solution during the AFM experiments resulted in the reappearance of the double peaked events as observed for the wild-type receptor (Figure 4.14C, green and blue histograms and inset) with a modal ΔL_c value of 21 ± 4 nm (Figure 4.14C grey histogram) and F_U = 64 ± 1 and 99 ± 7 pN for the unfolding and unbinding forces, respectively) (Figure 4.14D and Table 4.4).

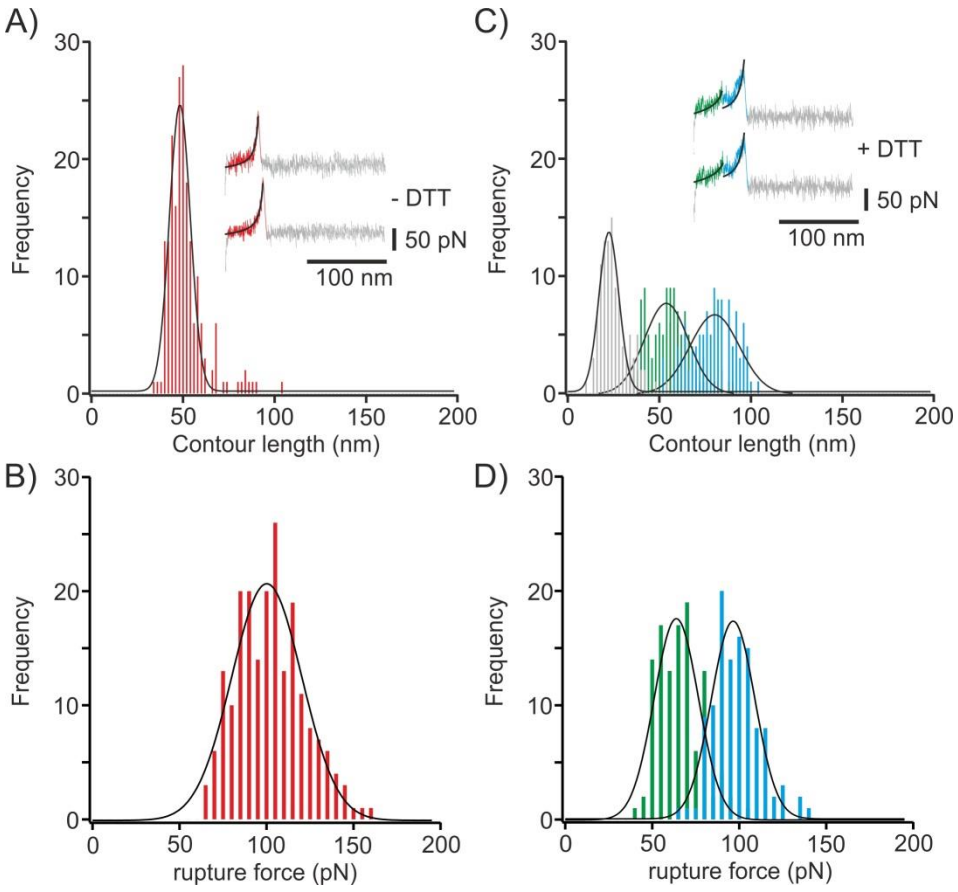


Figure 4.14 – SMFS of BtuB XL_{barrel} bound to vitamin B_{12} A) L_c distributions of single rupture events observed for BtuB XL_{barrel}:$TonB_{\Delta TMD}$ in the absence of reductant (inset shows force-extension traces) B) The force distribution of the single unbinding events of BtuB XL_{barrel}:$TonB_{\Delta TMD}$. C) After the addition of 2 mM DTT, double rupture events resulting from partial unfolding (L_c1, green histogram and green force-extension profiles, inset) and unbinding (L_c2, blue histogram and blue force-extension profiles, inset) of XL_{barrel} plug domain from Ton Box_{BtuB}. When fitted to a Gaussian distribution, the ΔL_c-frequency histogram (grey bars) for these double events had a modal value of 21 ± 4 nm. D) The force distribution of the double unbinding events of BtuB XL_{barrel}:$TonB_{\Delta TMD}$ in the presence of a reducing agent. Plug unfolding is shown in green and Ton box_{BtuB} unbinding in blue.

Similar to the wild-type data, force-extension profiles for BtuB XL_{loop}:$TonB_{\Delta TMD}$ dissociation contained a high proportion (94 %) of double peaked events (Figure 4.15A, orange and red traces inset). Most importantly, the ΔL_c-frequency histogram (Figure 4.15A, grey histogram) was found to give a significantly shorter modal value relative to that for wild-type (14 ± 3 nm and 20 ± 1 nm respectively), confirming unfolding in the N-terminal portion of the plug domain. Once again, the force extension profiles for BtuB XL_{loop}:$TonB_{\Delta TMD}$ dissociation reverted to wild-type profiles upon the addition of 2 mM DTT (ΔL_c = 20 ± 1 nm) (Figure 4.15B, grey histogram).

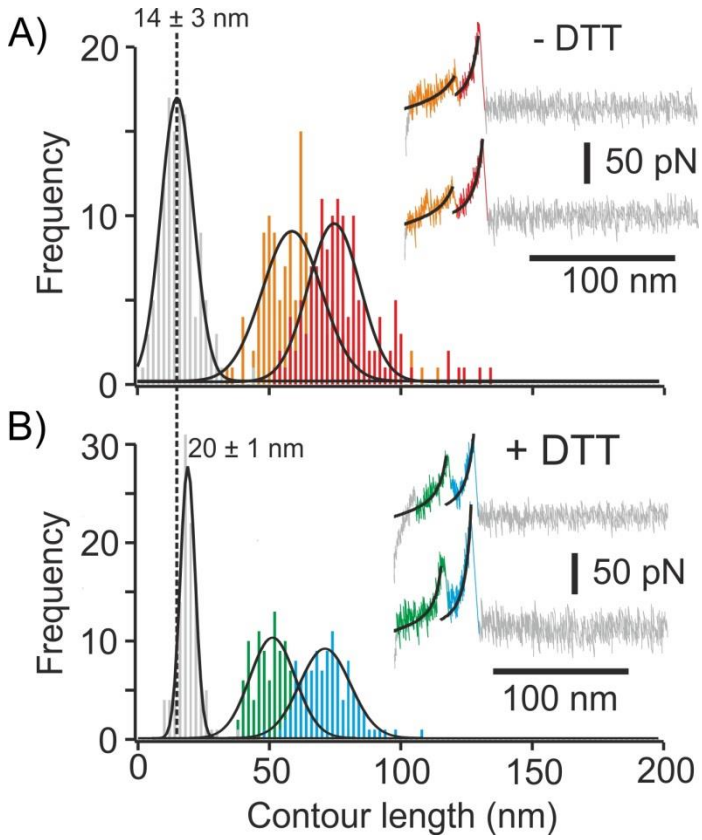


Figure 4.15 – SMFS of BtuB XL_{loop} bound to vitamin B_{12} A) L_c distributions and force-extension profiles (inset) for BtuB XL_{loop}:$TonB_{\Delta TMD}$ in the absence of DTT. Partial unfolding (L_c1) and unbinding (L_c2) are shown in orange and red, respectively, for both histograms and force-extension profiles. The ΔL_c (grey histogram) modal value is shown by the dashed line. B) L_c distributions and force-extension profiles (inset) for BtuB XL_{loop}:$TonB_{\Delta TMD}$ in the presence of 2mM DTT. Partial unfolding (L_c1) and unbinding (L_c2) are shown in green and blue, respectively, for both histograms and force-extension profiles. In the presence of DTT, the ΔL_c (grey histogram) modal value (20 ± 1) reverts to that of wild-type BtuB (20 ± 2). All data in Figure 4.15 was collected at 1,000 nms^{-1}.

The XL_{PSD} variant displayed the same single peaked force-extension AFM profiles (Figure 4.16A and Table 4.5) as XL_{barrel} (Figure 4.14A inset) in the absence of reducing agent, which suggests that diverting the force propagation network into the stronger plug subdomain prevents partial unfolding. Again, the addition of 2 mM DTT restored the presence of double rupture events and the ΔL_c (20 ± 1 nm) was the same as the wild-type data.

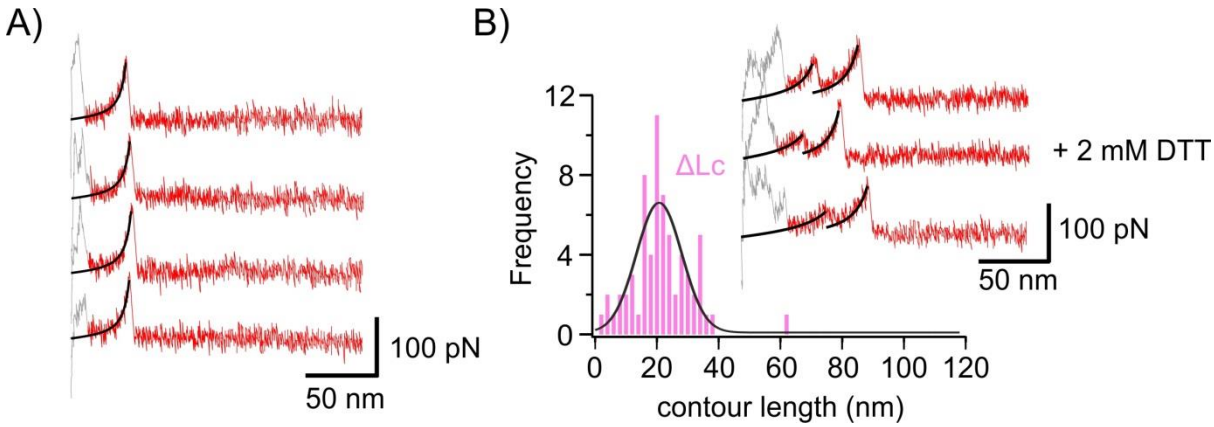


Figure 4.16 - SMFS data of the XL_{PSD} mutant of BtuB with $TonB_{\Delta TMD.}$ A) Example force-extension traces for the XL_{PSD} mutant. The non-specific surface adhesion is coloured in light grey. The bold black trace show the WLC fit to the data. B) The ΔL_c distribution (pink) of double peak events (see inset) upon the addition of reducing agent (2 mM DTT) to the sample.

Together these data suggest that approximately 50 residues of the plug domain directly downstream of the Ton box are denatured before the non-covalent TonB:BtuB complex dissociates. The structural consequence of partially unfolding the plug domain, was visualised by removing the fifty residues directly downstream of Ton Box_{BtuB} from the crystal structure of BtuB bound to vitamin B_{12} and $TonB_{CTD}$ (PDB: 2GSK) which resulted in a 12.5 × 30.2 Å-wide channel through the receptor (Figure 4.17).

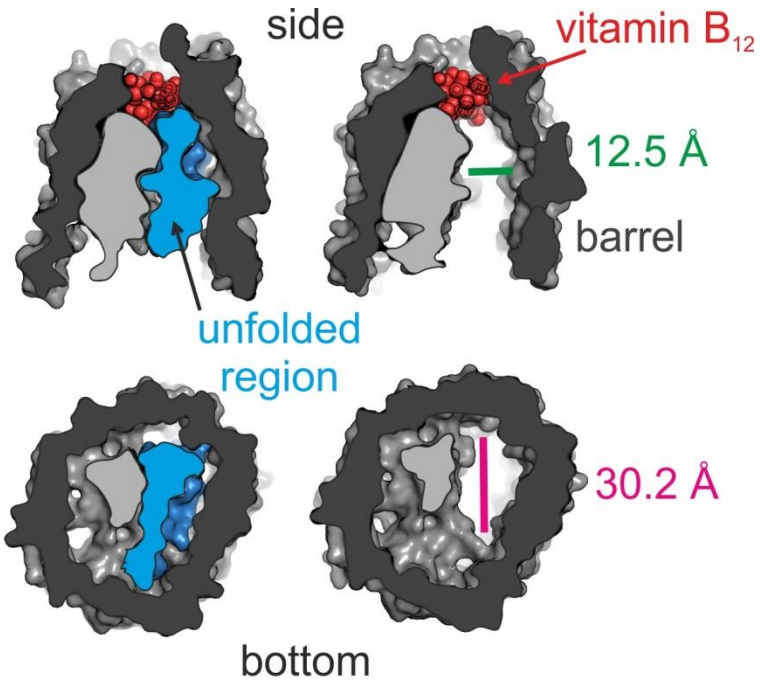


Figure 4.17 - Vertical and horizontal surface representation cross-sections of BtuB showing the effect of removing 50 amino acids (blue) downstream of the Ton box. The dimensions of the resultant channel are shown (PDB: 2GSK). Measurements were made using PyMOL.

					Plug unfolding				TonB unbinding					
Variant	#	N attempts	n events[a]	Hits (%)	Force (pN)	Average	L_c (nm)	Average	Force (pN)	Average	L_c (nm)	Average	ΔL_c (nm)	Average
BtuB wt	1	2400	54	2.2	63		58		105		70		20	
	2	1200	18	1.5	62	60 ± 3	54	57 ± 4	68	91 ± 23	80	77± 7	21	20 ± 2
	3	1948	81	4.1	57		61		100		82		18	
XL_{loop}	1	6386	88	1.4	56		58		103		76		17	
(V29C/V45)	2	7365	44	0.6	51	55 ± 4	55	53 ± 7	90	98 ± 8	71	69 ± 8	11	14 ± 3
No DTT	3	2000	22	1.1	58		46		101		61		15	
XL_{barrel}	1	6725	85[b]	1.3[b]	-		-		102		48		-	
(L23C/S374C)	2	2492	38[b]	1.5[b]	-	-	-	-	104	100 ± 6	45	47 ± 2	-	-
No DTT	3	3304	73[b]	2.2[b]	-		-		94		47		-	
XL_{loop}	1	8324	34	0.4	60		49		99		70		21	
(V29C/V45)	2	6420	71	1.1	65	61 ± 4	52	49 ± 3	102	102 ± 3	72	69 ± 4	19	20 ± 1
2 mM DTT	3	3579	17	0.5	57		47		105		65		19	
XL_{barrel}	1	3419	18	0.5	63		52		92		82		25	
(L23C/S374C)	2	3669	69	1.9	65	64 ± 1	46	51 ± 5	105	99 ± 7	69	76 ± 7	21	21 ± 4
2 mM DTT	3	3753	22	0.6	64		55		100		77		19	

Table 4.4 - AFM data of $TonB_{\Delta TMD}$:BtuB variants (± 2 mM DTT) at 1000 nms^{-1}. Force, contour length (L_c) and ΔL_c values were calculated by plotting the distribution of events and fitting with a Gaussian to obtain the mean value. Hits denotes the percentage of approach/retract cycles that result in force-extension profiles that meet the criteria discussed in Section 2.2.7.4. The error on the average is the range of the triplicate averages. [a]number of double events, [b]number/% of single events.

					Plug unfolding				TonB unbinding					
Variant	#	N attempts	n events	Hits (%)	Force (pN)	Average	L_c (nm)	Average	Force (pN)	Average	L_c (nm)	Average	ΔL_c (nm)	Average
BtuB	1	4656	150	3.2	-		-		95		59		-	
XL_{PSD}	2	5003	95	1.9	-	-	-	-	120	108 ±	52	52 ± 7	-	-
No DTT	3	5000	118	2.4	-		-		108	13	46		-	
BtuB	1	3500	16	0.4	72		45		114		62		19	
XL_{PSD}	2	3870	37	1.0	67	62 ± 14	42	52 ± 18	113	106 ±	69	76 ± 20	19	20 ± 1
2 mM DTT	3	3540	29	0.8	48		70		90	16	96		21	

Table 4.5 - AFM data of $TonB_{\Delta TMD}$:BtuB XL_{PSD} (± 2 mM DTT) at 1000 nms^{-1}. Force, contour length (L_c) and ΔL_c values were calculated by plotting the distribution of events and fitting with a Gaussian to obtain the mean value. Hits denotes the percentage of approach/retract cycles that result in force-extension profiles that meet the criteria discussed in Section 2.2.7.4. The error on the average is the range of the triplicate averages

4.2.6 MD simulations of the TonB: BtuB interaction

The SMFS results suggest that TonB can remodel approximately half of the plug domain, thus creating a channel through the BtuB receptor of sufficient dimensions to allow passage of the vitamin B_{12} substrate (Section 2.2.10). To test the consequence of unfolding half of the plug domain and visualise the process, MD simulations were performed. The structure of BtuB in complex with $TonB_{CTD}$ (residues 153-233) and vitamin B_{12} (PDB: 2GSK) was inserted into a POPC lipid bilayer and surrounded by a water box sufficiently large for the system to be entirely solvated when the Ton box is extended up to 20 nm. Using this system, it is possible to determine how much of the plug domain can be remodelled by the non-covalent β-strand augmentation between the C-terminus of TonB and the Ton box of BtuB, albeit at a much higher loading rate (x250,000) than the SMFS experiments. The system was relaxed and equilibrated using a step by step method described in Section 2.2.10 for a total time of 15.2 ns, the equilibrated system is shown in Figure 4.18A. The positional root mean square deviation (RMSD) from the crystal structure plateaus at about 1.2 Å and shows that the system is stable and well equilibrated (Figure 4.18B). The positioning of the integral membrane protein within the lipid bilayer and the membrane thickness also reached a plateau within the initial 15 ns, showing that the system was sufficiently equilibrated (Figure 4.18C and D). Longer equilibration did not alter the structure further (Figure 4.20C).

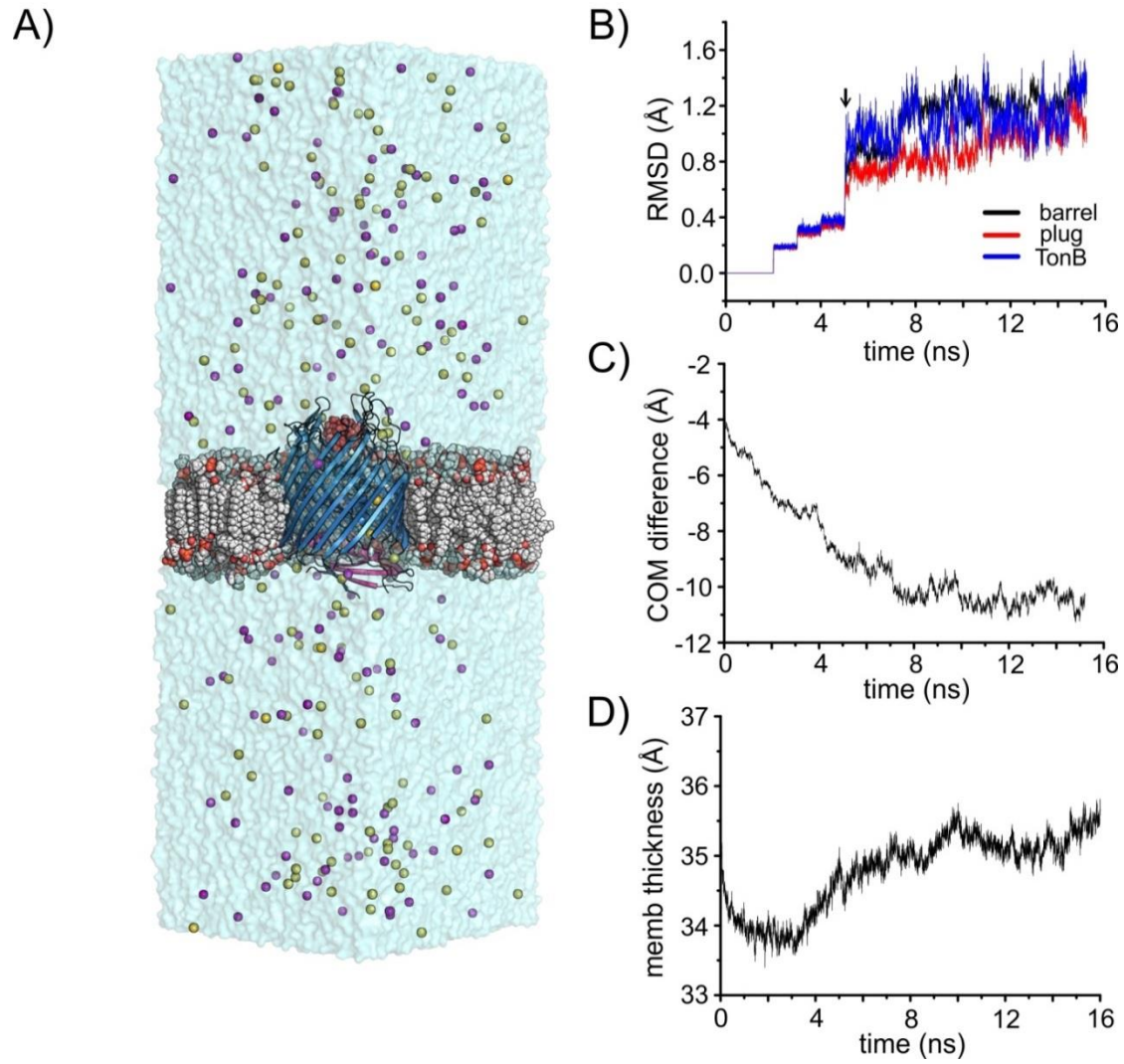


Figure 4.18 - Molecular dynamics simulation of the TonB:BtuB:vitamin B_{12} complex. A) The equilibrated system of BtuB (blue cartoon) in complex with TonB (pink cartoon) and vitamin B_{12} (red spheres) (PDB: 2GSK) in a POPC lipid bilayer (red and white spheres) and a large water box (blue surface representation). The Na^{2+} (purple) and Ca^{2+} ions (yellow) are shown as coloured spheres. B) RMSD of the individual components (labelled) during the equilibration of the system. The black arrow indicates when all restraints were removed. C) The average centre of mass (COM) difference between BtuB and the POPC lipid bilayer during the equilibration of the system (calculated using MD analysis[201] and a script provided by Robert Schiffrin, University of Leeds). The data indicates that the protein position within the membrane remains constant after 8 ns equilibration. D) The average membrane thickness during the equilibration of the system (calculated using membrane plugin v1.1[189])

4.2.6.1 Pulling the Ton box of BtuB

Initially a force ramp was applied to the Ton box (Cα proline 5, BtuB) by retracting a harmonic spring at a constant velocity (also called steered MD (SMD)) for 200 Å (Figure 4.19A and B). The plug domain unfolded sequentially (residues 5-74) and a channel was

formed through the receptor (Figure 4.19F). As the weaker subdomain of the plug is in contact with vitamin B_{12}, its unfolding could potentially alter the substrate position within the binding pocket. However, only a small displacement of 3 Å occurred during unfolding of the residues in contact with the substrate before returning to its original position (Figure 4.19D). A 20 ns equilibration of the open state also did not change the position of the substrate. Remarkably, the mechanically strong plug subdomain and the barrel domain did not undergo any conformational change during (Figure 4.19C) and after unfolding (Figure 4.19E) of the weak subdomain, despite exposure of hydrophobic residues usually buried by the folded weak subdomain.

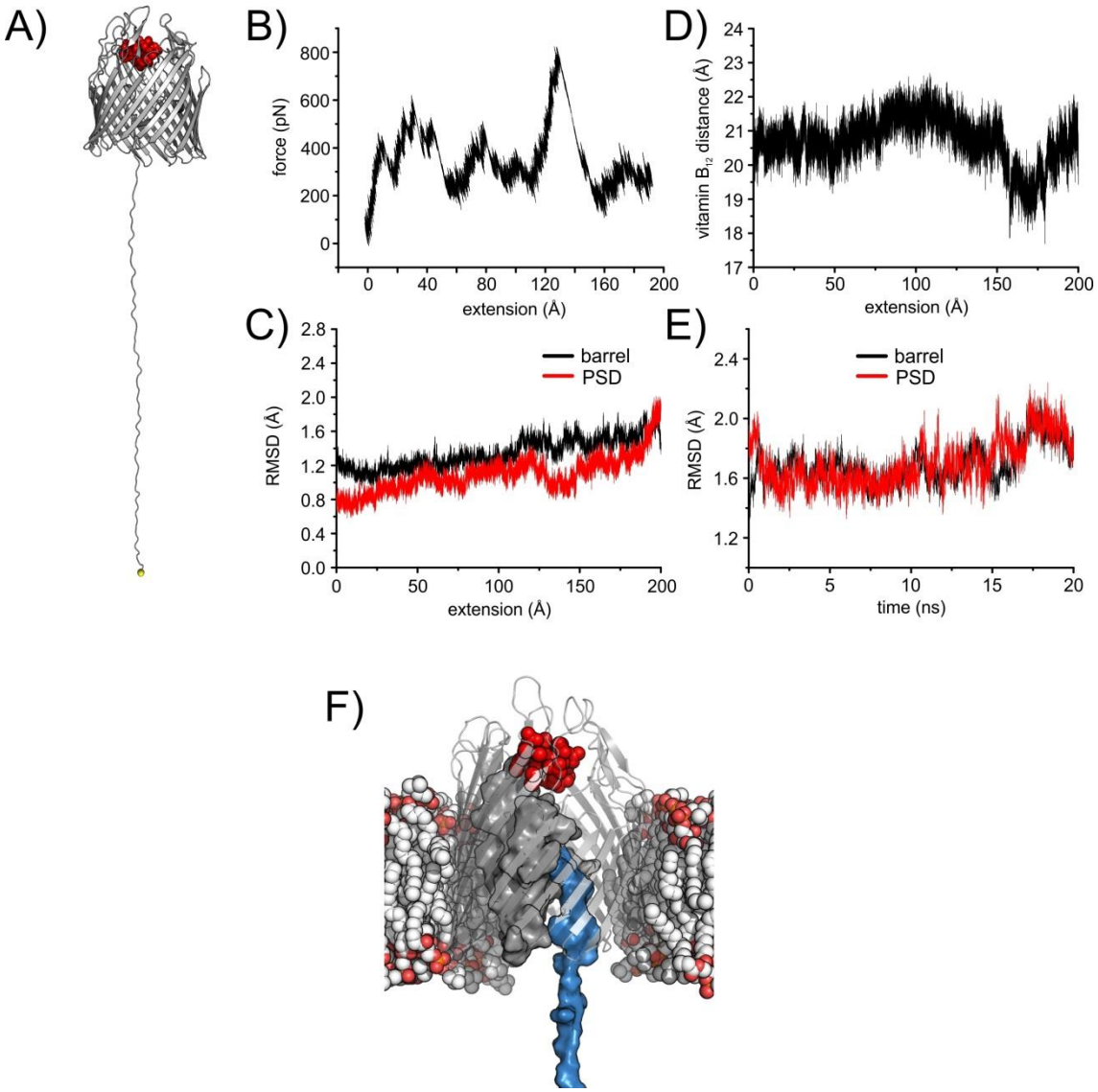


Figure 4.19 – Unfolding half of the plug domain *in silico*. A) Snapshot after atom P5 Cα (yellow sphere) had been extended 200 Å from its initial position using SMD, the membrane and water are hidden for clarity. BtuB is shown as a grey cartoon with vitamin B_{12} shown as red spheres. B) Force-extension plot of the Ton box during SMD. C) RMSD of the mechanically strong plug subdomain (PSD) and barrel domain of BtuB during the unfolding of the weak plug subdomain. D) Distance from the centre of the plug domain (L77) from the cobalt core of vitamin B_{12} during unfolding of the weak plug subdomain. E) Stability of the open state after 200 Å extension of the Ton box, showing the RMSD of the strong plug subdomain (PSD) and the barrel domain of BtuB. F) Snapshot of BtuB equilibrated in a POPC bilayer (red and white spheres) after 200 Å extension of the Ton box by SMD. The weak plug subdomain is shown in blue, the strong sub-domain in grey surface representation and vitamin B_{12} as red spheres.

4.2.6.2 Pulling TonB to unfold the plug domain

To find the extent of plug remodelling by the non-covalent Ton box $TonB_{CTD}$ complex, SMD was applied to the centre of mass of TonB (all Cα). The interaction between TonB and the Ton box demonstrated outstanding mechanical strength even at the high loading rate of the simulation (with compatibly of the current limits on such a large system), remaining tightly bound and unfolding 46 residues of the BtuB plug domain (up to residue G51, ~130 Å of extension) before dissociation (Figure 4.20A and D). The same extent of unfolding as the SMFS experiments was not observed. A longer equilibration of the system before SMD was performed to ensure the system was equilibrated, however the same amount of unfolding was again observed (Figure 4.20B). To find if the vitamin B_{12} was stabilising residues 47-74 (which it is in contact with) a system was set up without the vitamin B_{12} in the BtuB binding pocket. Again, 46 residues were unfolded before the complex dissociated (Figures 4.20B). Comparing the force-extension profiles of the SMD from the Ton box and TonB (Figure 4.19B and 4.20B), it is clear that the residues left unfolded by the non-covalent complex (residues 47-74) require ~800 pN of force to unfold (Figure 4.19B), which the non-covalent interaction can only withstand ~500-600 pN at the loading rate used (Figure 4.20B). However, despite this, it can be concluded that the non-covalent β-strand augmentation does display huge mechanical resilience, even at high loading rates.

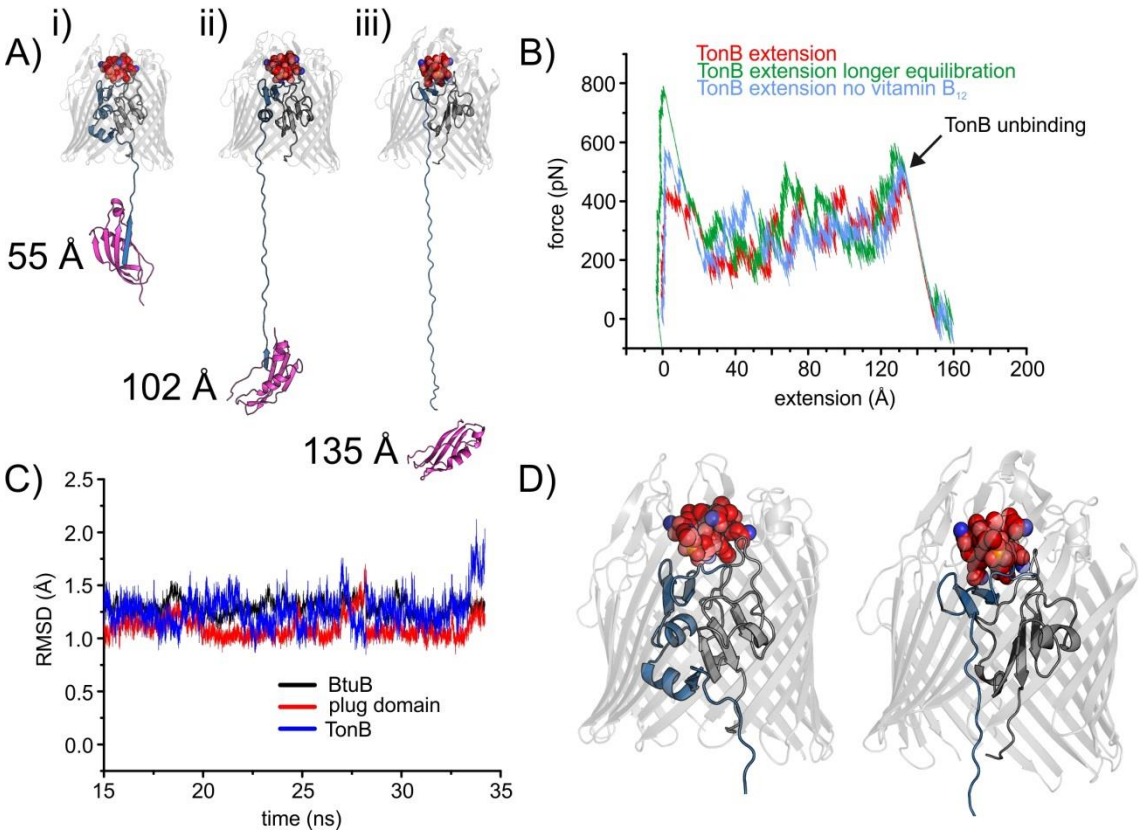


Figure 4.20 – Finding whether $TonB_{CTD}$ can unfold the same extent of plug domain as the SMFS experiments. A) SMD snap shots of TonB (pink) unfolding the plug domain of BtuB (barrel domain shown as grey transparent cartoon, plug weak subdomain shown in blue, strong plug subdomain in grey and vitamin B_{12} as spheres. The membrane and water is hidden for clarity. The extensions are shown. In iii) 46 residues (residues 5-51) have been unfolded by the non-covalent TonB:Ton box complex. B) Force-extension plots of SMD of the interaction between BtuB:vitamin B_{12} with TonB (red) TonB with a longer equilibration (green) and TonB with apo BtuB (blue). When the complex is extended via TonB, the complex breaks apart at 135 Å of extension (arrow). C) RMSD of the protein components during a longer equilibration before SMD. D) Detail of the BtuB plug domain before (left, zoom of Ai) and after (right, zoom of Aiii) SMD was used to extend the non-covalent complex by TonB. Residues 5-74 (the weak plug subdomain) are shown in blue.

4.2.7 Modelling vitamin B_{12} passage

The crystal structure of BtuB (PDB: 2GSK) with residues 5-74 removed (residues 5-22 Ton box, 23-74 the unfolding domain) shows a clear channel throughout the receptor, and the dimensions of the channel appear to be sufficient to accommodate the passage of vitamin B_{12} (Figure 4.17). To find whether vitamin B_{12} could pass through the channel, an equilibrated open state after 200 Å Ton box extension was used (i.e. with residues 5-74 of the plug domain unfolded).

The open state was initially equilibrated for 20 ns before SMD was applied to the centre of mass of vitamin B_{12}. The vitamin B_{12} was tightly bound to the binding pocket, requiring 400-500 pN of force for displacement (Figure 4.21B). After binding site displacement, vitamin B_{12} began to move rapidly through the channel until roughly midpoint, where it briefly stalled at the narrowest point of the channel before continuing to pass through the receptor with little resistance. As vitamin B_{12} moved through BtuB, the RMSD of both the barrel domain and the mechanically strong subdomain displayed no significant changes (apart from one repeat, where the PSD distorted (Figure 4.21C, red trace)) (Figures 4.21C and D). This indicated that by unfolding residues 5-74, the passage of vitamin B_{12} is accommodated without further conformational change of the receptor. Mapping the contacts vitamin B_{12} made with the receptor during translocation, revealed that surprisingly few contacts with the strong plug domain were made. Other than the initial substrate binding pocket, the major contacts were with the barrel domain at the narrowest point of the channel (Figure 4.21E).

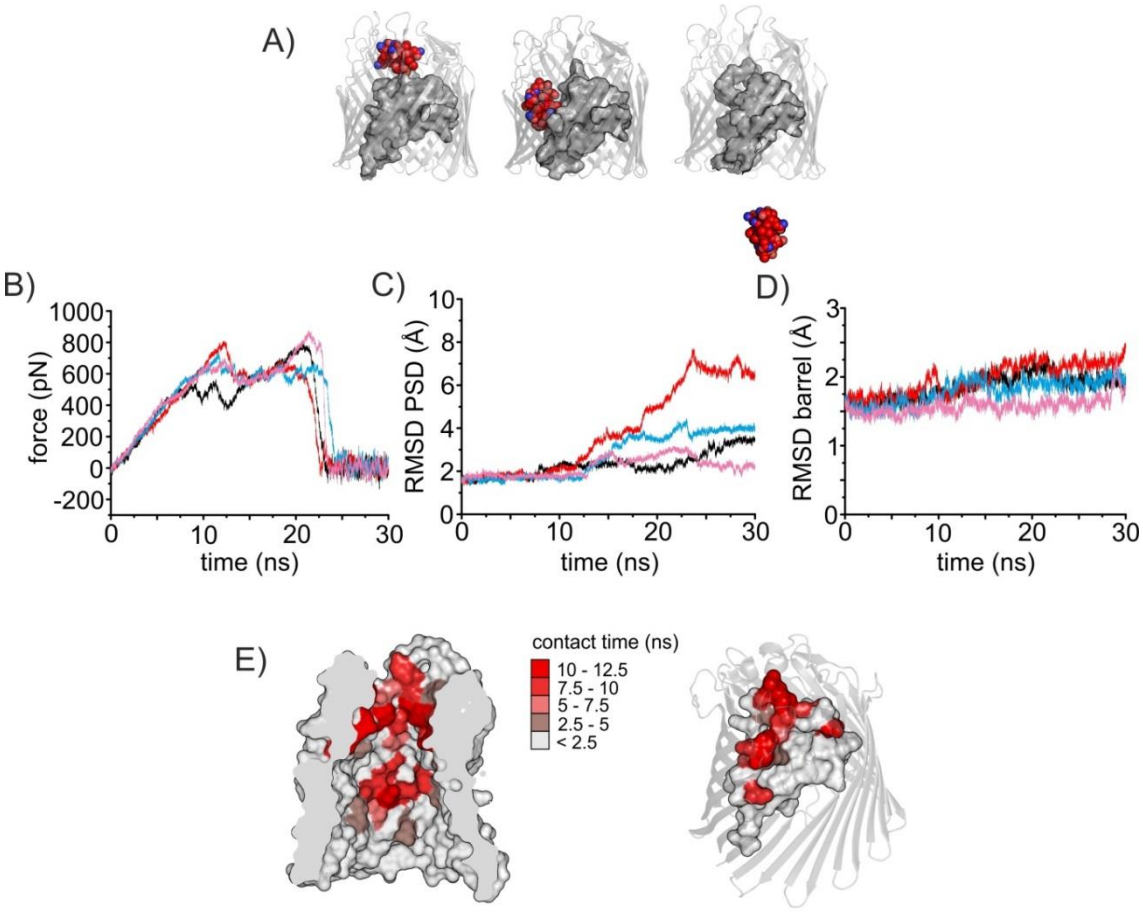


Figure 4.21 – Simulation of vitamin B_{12} passage. A) Snapshots as vitamin B_{12} moves through the channel created by 200 Å of extension. The strong plug subdomain is shown as grey surface representation; the barrel is shown as transparent cartoon and vitamin B_{12} as coloured spheres. B) Force applied to displace vitamin B_{12} through the BtuB channel. C) RMSD of the BtuB strong plug subdomain (PSD). D) RMSD of the BtuB barrel domain. In B-D, four separate repeats are shown. E) The contacts vitamin B_{12} makes as it passes through the receptor channel (calculated by residues within 4 Å of vitamin B_{12}).

4.2.8 Single-molecule force spectroscopy of the $TonB_{\Delta TMD}$:FhuA interaction

TBDTs transport substrates of various size, with vitamin B_{12} (1.35 kDa) being one of the largest substrates. In order to determine whether plug remodelling is a BtuB-specific process, and whether the extent of remodelling is dependent on the substrate size, the FhuA receptor for ferricrocin/ferrichrome (740.52/773.5 Da respectively) (Figure 4.22) was studied by SMFS.

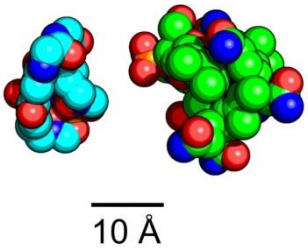


Figure 4.22 – Comparison of TBDT substrate size. The structure of ferricrocin-iron is shown as blue spheres (with N in navy and O in red) and vitamin B_{12} is shown as green spheres with N and O navy and red respectively. The structures are from the TBDT bound states of FhuA and BtuB (PBD: 2GRX and 2GSK respectively).

Using the experimental method for BtuB and its variants, FhuA (cloned, expressed and purified as described in Section 4.2.1 and Section 4.2.2.2) proteoliposomes (formed using the method described in Section 2.2.8) were immobilised onto a mica surface and an AFM probe covalently attached to $TonB_{\Delta TMD}$(V32C) was used. As observed for BtuB, two rupture events were detected (74 ± 11 and 105 ± 24 pN at a distance of 62 ± 4 and 85 ± 9 nm), Figure 4.23A) but were separated by a larger change in contour length (ΔL_c = 25 ± 1 nm) (Table 4.6 and Figure 4.23B), corresponding to the removal of 63 residues (residues 20-83)(Figure 4.23B inset). The plug domain (residues 18-160) of FhuA is larger than BtuB (residues 17-136), despite the cargo being smaller. In this case, while remodelling results in the formation of a channel throughout the receptor, its dimensions are slightly smaller (14 X 22 Å) reflecting the smaller size of its cargo (Figure 4.24).

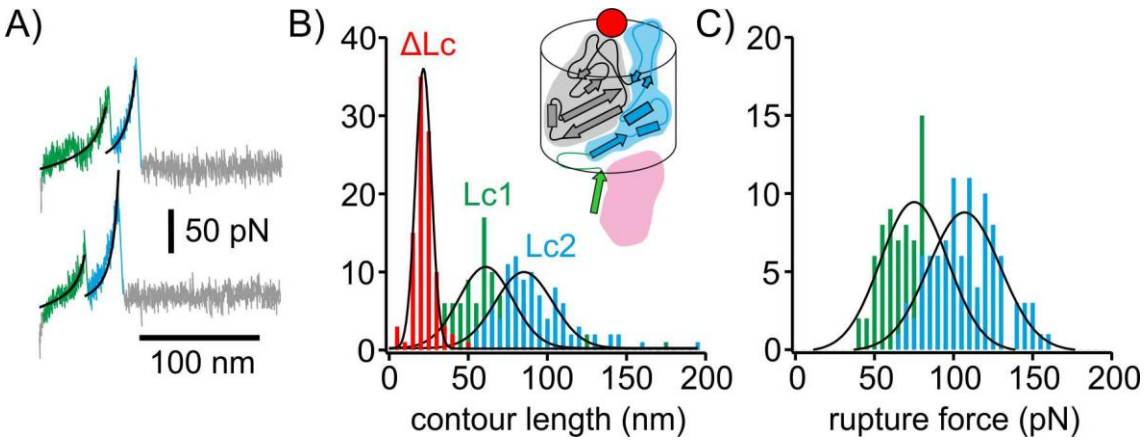


Figure 4.23 - SMFS of the FhuA:TonB$_{\Delta TMD}$ interaction in the presence of ferrichrome. A) Examples of the double rupture events with the WLC fit (black) B) L_c1 (green), L_c2 (blue) and ΔL_c (red) distributions of the double rupture peaks with Gaussian fits; the inset shows a secondary structure schemetic of the FhuA plug domain with the residues unfolded (inferred from the ΔL_c value) shown in blue, TonB in pink, the Ton box in green and the ferricrocin substrate in red. C) Force-distribution of the double rupture events, with the plug unfolding in green and TonB unbinding from the Ton Box$_{FhuA}$ in blue.

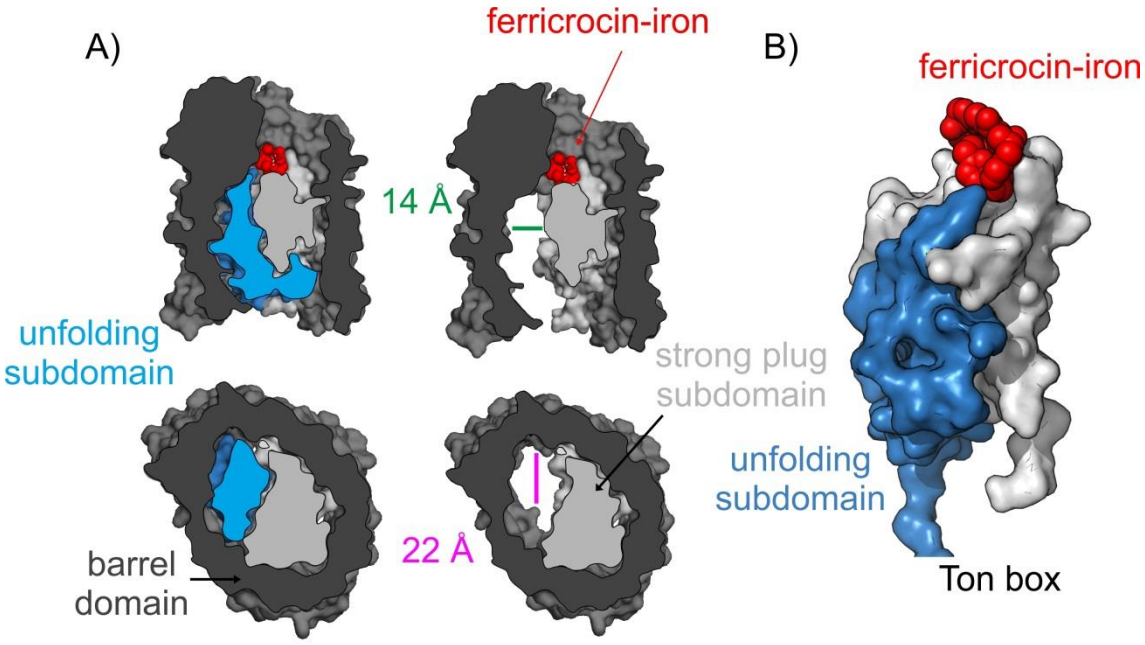


Figure 4.24 – Removal of the unfolding region from the crystal structure of FhuA. A) Surface representation of vertical and horizontal cross-sections of FhuA (PDB: 2GRX) showing the effect of removing 62 amino acids (blue) downstream of the Ton box. The dimensions of the resultant channel are shown. B) Surface representation of the plug domain of FhuA, coloured blue are the residues that are unfolded (inferred from the ΔL_c value).

					Plug unfolding				TonB unbinding					
Protein	#	N attempts	n events	Hits (%)	Force (pN)	Average	L_c (nm)	Average	Force (pN)	Average	L_c (nm)	Average	ΔL_c (nm)	Average
FhuA	1	5207	34	0.7	89		66		124		94		26	
	2	6139	41	0.7	70	74 ± 11	62	62 ± 4	100	105 ± 24	82	85 ± 9	24	25 ± 1
	3	8150	24	0.3	63		59		91		79		24	

Table 4.6 - AFM data of $TonB_{\Delta TMD}$:FhuA at 1000 nms^{-1}. Force, contour length (L_c) and ΔL_c values were calculated by plotting the distribution of events and fitting with a Gaussian to obtain the mean value. Hits denotes the percentage of approach/retract cycles that result in force-extension profiles that meet the criteria discussed in Section 2.2.7.4. The error on the average is the range of the triplicate averages.

4.3 Discussion

When the first crystal structure of a TBDT revealed that the lumen of the barrel was completely occluded by a globular plug domain[126], it became obvious that in order for the substrate to be transported across the receptor, the plug domain would have to undergo conformational change to allow substrate passage. The SMFS data presented in Chapter 3 and Chapter 4 has demonstrated that the intermolecular interaction between TonB and the Ton box of BtuB is stronger than intra-molecular interactions within the plug domain. This allows mechanical unfolding of the globular plug domain of both BtuB and FhuA before its dissociation under mechanical force *in vitro*. The extent of plug remodelling appears to be highly controlled and is determined by the cargo the receptor has evolved to transport. For both receptors, the plug domain can be regarded as comprising a mechanically weak channel forming subdomain, and a mechanically strong subdomain used both for allosteric signalling[202] and to limit the size of the channel to allow passage of molecules no larger than its cargo. The strong plug sub-domain identified in this study is likely structurally supported (to prevent dislodging) by the conserved latch and β-cantilever domains identified[129] (Figure 1.13 and Figure 1.14). It was surprising that the open conformation of BtuB created by steered molecular dynamics was stable and did not undergo further conformational change despite exposing the core of the folded domain. This strongly suggests that the plug domain has evolved to perform this function, which is also evident from the highly solvated plug-barrel interface in the crystal structures of TBDTs (Figure 1.13).

Interestingly, the mechanically weak subdomains of BtuB (Figure 4.17) and FhuA (Figure 4.24) also both make direct contact to their ligands. As these subdomains are presumably pulled into the periplasmic space during ligand transport, it is plausible that *in vivo* the substrate is dragged through the receptor. Ligand-plug contacts are also present in the substrate-bound crystal structures of FecA[203], HasR[204], FpvA[205] and FptA[205] , therefore this could represent a universal feature of TBDTs. The fact that such a mechanism was not observed in the MD simulations of BtuB remodelling (Figure 4.19D) could be due to the high pulling velocity used.

The direct identification of two distinct subdomains in the luminal plugs of FhuA and BtuB rationalise a wealth of previous observations from mutational and cross-linking

studies. For example, the first 51 residues of the FepA plug domain were found to be more susceptible to labelling relative to residues 91–142[206], while the introduction of cross-links between the N- or C-terminal portions of the FhuA plug to the barrel were found to prevent or support ferrichrome transport respectively[207,208]. A lack of plug domain co-operativity has also been observed *in vitro*: EPR and trypsin cleavage have shown that the N-terminal regions of the BtuB and FhuA plug domains (respectively) unfold in the presence of 4 M urea[209,210] and that for BtuB, the unfolded sub-domain rapidly refolds upon removal of the denaturant[209]. Finally, previous AFM studies which investigated the mechanical unfolding pathway of the barrel of FhuA (by adsorbing FhuA to the AFM tip in a non-specific manner in the absence of TonB) showed that the plug domain unfolded in two steps (residues 1 to 91, F_U= ~50 pN and residues 92-161, F_U = 163 ± 50 pN at a retraction velocity of 2200 nms^{-1})[178].

In the next chapter, the work in the current chapter will be supported by an *in vivo* uptake assay to find the effects of preventing the weak plug subdomain from unfolding using disulfide cross-linking. The effect of destabilising the stronger plug subdomain by removal of the hydrophobic core (hence entire plug unfolding) will also be explored both *in vitro* and *in vivo*.

5 Testing the physiological role of partial plug domain unfolding in BtuB

5.1 Objectives

The SMFS experiments have shown that TonB can partially remodel half of the plug domain of both BtuB and FhuA under mechanical extension. To test whether preventing unfolding in BtuB prevents vitamin B_{12} uptake *in vivo*, an assay using *E. coli* cells which require vitamin B_{12} to proliferate in minimal medium is used. Conversely, to examine the effect of entire plug unfolding, a construct of BtuB was engineered in which the plug domain is significantly destabilised. This construct was then used in an antibiotic sensitivity assay *in vivo,* and the consequences of entire plug unfolding explored.

5.2 Results

5.2.1 Vitamin B_{12} uptake assay

Several of the residues in the mechanically weak plug sub-domains of both BtuB and FhuA directly contact the cargo at the extra-cellular surface suggesting that the resulting channel is the likely route for cell entry. To assess the consequence of preventing channel formation *in vivo*, or of diverting the force propagation pathway to the strong plug subdomain, the growth phenotypes of the *E. coli* strain RK5016 expressing wild-type BtuB or its cross-linked variants in the presence of different concentrations of vitamin B_{12} was examined (Section 2.2.5.2).

In *E. coli* there are two methionine synthases, MetE (vitamin B_{12}-independent methionine synthase) and MetH (vitamin B_{12}-dependent methionine synthase). It is advantageous for *E. coli* to use MetH if vitamin B_{12} is present as the turnover rate is 100 times greater than MetE[146]. When *metE* strains of *E. coli* are grown in the absence of methionine, vitamin B_{12} becomes a nutritional requirement for growth. Therefore *metE* strains grown in Davis minimal medium (Section 2.2.5.1) with stringent concentrations of vitamin B_{12} (0.1-1 nM) require a functional BtuB for growth, demonstrated in Figure 5.1.

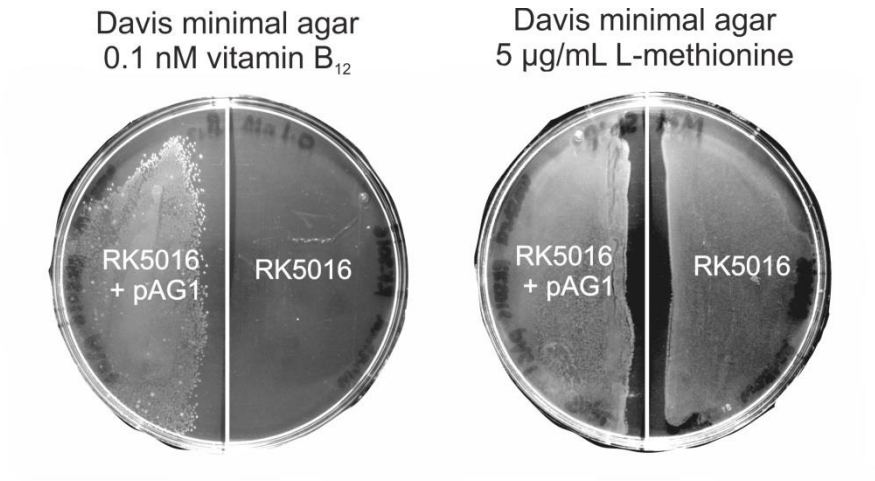
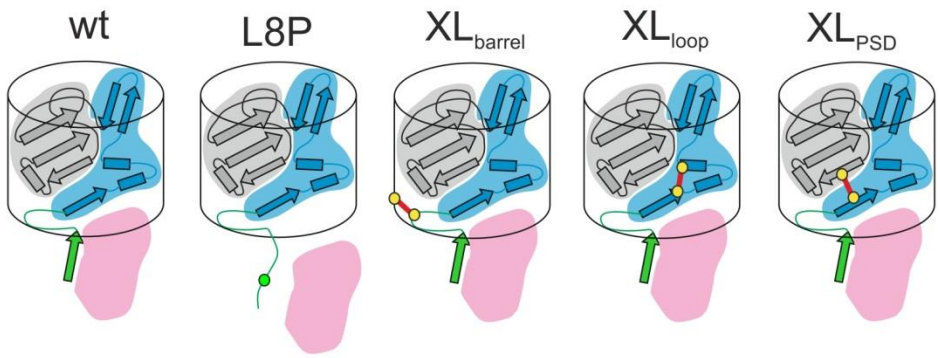


Figure 5.1 – *metE* strains of *E. coli* (RK5016) require BtuB (pAG1) for growth in the absence of methionine.

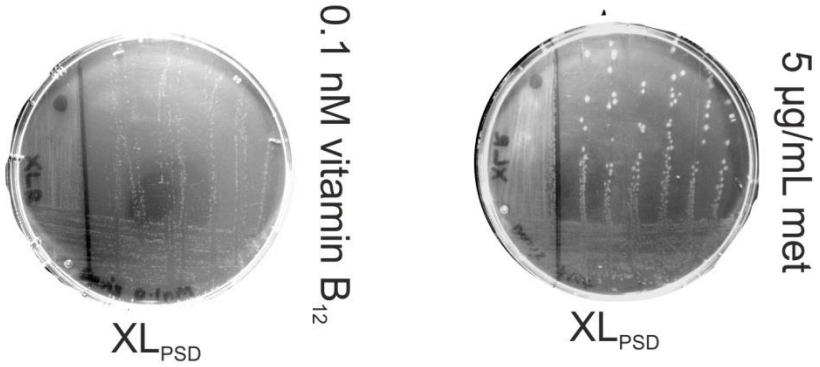
E. coli RK5016 transformed with wild-type BtuB, the cross-link variants or L8P (Figure 5.2A) under the control of a native promoter (pAG1)[79] were grown on Davis minimal medium agar plates supplemented with various concentrations of vitamin B_{12} (0.1-100 nM) or methionine (5 µg/mL) (Section 2.2.5.2) and the growth phenotype was assessed by comparing colony size after 48 hour incubation at 37 ˚C.

Full growth on vitamin B_{12} supplemented plates (thus fully functional BtuB) resulted in colonies similar in size to those in the presence of methionine (2 mm diameter) (Figure 5.2B). Under stringent conditions of vitamin B_{12} (0.1-1 nm), full growth was observed for wild-type and XL_{loop} BtuB (Figure 5.2C), whilst reduced growth was observed at all vitamin B_{12} concentrations for BtuB XL_{barrel} and XL_{PSD}. The phenotypes of these variants concur that application of force (via TonB) unravels the mechanically weak plug sub-domain from the N-terminus to create a channel for substrate transport. Diverting the propagated force to the mechanically strong plug subdomain or the barrel domain prevents transport of vitamin B_{12}, in agreement with the findings of Chapter 4.

A)



B)



C)

Strain	Growth on:				
		[CN-Cbl (vitamin B12)] (nM)			
RK5016	Met	0.1	1	10	100
pAG1 (wt BtuB)	++	++	++	++	++
L8P	++	-	-	+	+
XL_{barrel}	++	+	+	+	+
XL_{loop}	++	++	++	++	++
XL_{PSD}	++	+	+	+	+

Figure 5.2 – Vitamin B_{12} uptake assay. A) Cartoon schematics of the BtuB mutants, disulfide bridges are shown by yellow circles with a red bridge, the L8P mutant is shown by the green circle. The mechanically weak plug subdomain is shown in blue, TonB in pink and the Ton box in green. B) Photographs of *E. coli* RK5016 BtuB XL_{PSD} colonies on agar plates supplemented with stringent vitamin B_{12} or methionine (labelled). The reduced growth (+) phenotype after 48 hours incubation is exemplified by the vitamin B_{12} plate, and full growth (++) on the methionine plate. C) Growth of *E. coli* RK5016 transformed with BtuB variants after 48 hours incubation on Davis minimal agar supplemented with 0.1 – 100 nM vitamin B_{12} or 5 µg/mL methionine. – indicates no observable colonies, + indicates partial growth (small colonies) and ++ indicates full growth (2 mm colonies).

5.2.2 Destabilising the mechanically strong plug subdomain of BtuB

To find the consequence of entire plug unfolding *in vivo*, the hydrophobic core of the strong plug subdomain (residues 73-136) of BtuB was located using the high resolution crystal structure of BtuB in complex with TonB (Figure 5.3). Site-directed mutagenesis was used to engineer a variant with a weakened strong subdomain (Section 2.2.1.6 and Table 2.1).

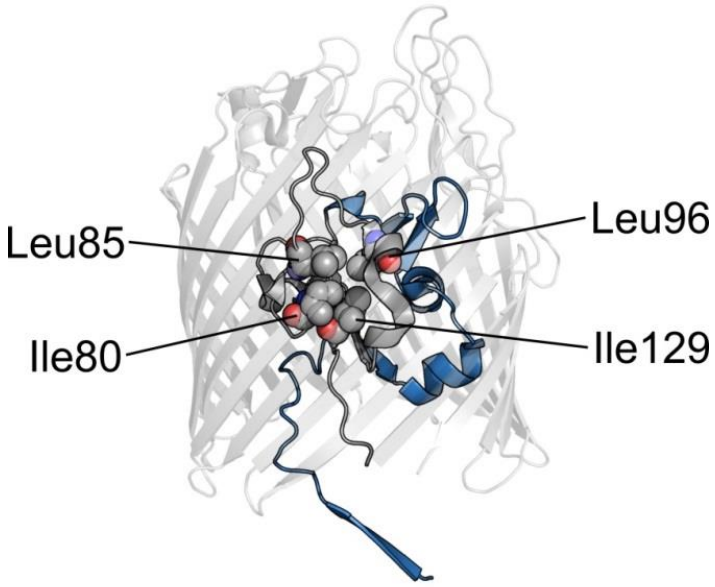


Figure 5.3 – The hydrophobic core of the BtuB strong plug subdomain (grey) with the residues that form the core shown as spheres (labelled).The weak plug subdomain is shown in blue, whilst the barrel domain is transparent. TonB and vitamin B_{12} are hidden for clarity. (PDB: 2GSK).

Two mutants were created: BtuB(I80A/L85A/L96A) (designated $BtuB_{3A}$) and BtuB(I80A/L85A/L96A/I129A) (designated $BtuB_{4A}$) into the pNGH15 plasmid (Section 2.1.16). Once the mutants were confirmed by DNA sequencing, the BtuB variants were expressed and purified using the same method as for the wild-type receptor (Section 2.2.2.6). The yield of protein from the $BtuB_{4A}$ mutant was less than $BtuB_{3A}$ (Figure 5.4A), implying that this mutant was preventing correct folding into the OM to some extent.

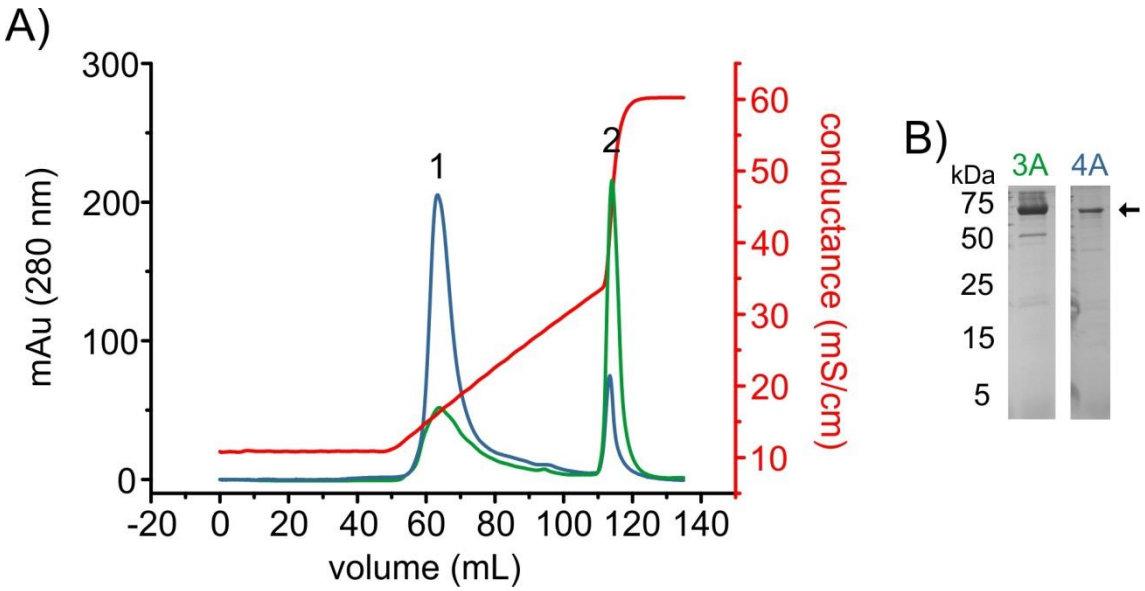


Figure 5.4 – Purification of the weakened plug mutants of BtuB. A) Anion exchange chromatography of $BtuB_{3A}$ (green) and $BtuB_{4A}$(blue) using a HiTrap DEAE FF 5 mL column, the peak at 115 mL (peak 2) is the elution of the BtuB variants from the column. B) SDS-PAGE of the BtuB elution peaks from anion exchange. The black arrow shows the expected band from the M_w of BtuB.

After purification, far-UV CD and intrinsic tryptophan fluorescence spectroscopy (Sections 2.2.6.1 and 2.2.6.2 respectively) were performed to confirm that the purified proteins were correctly folded into detergent by comparison to the wild-type receptor (Figure 5.5).

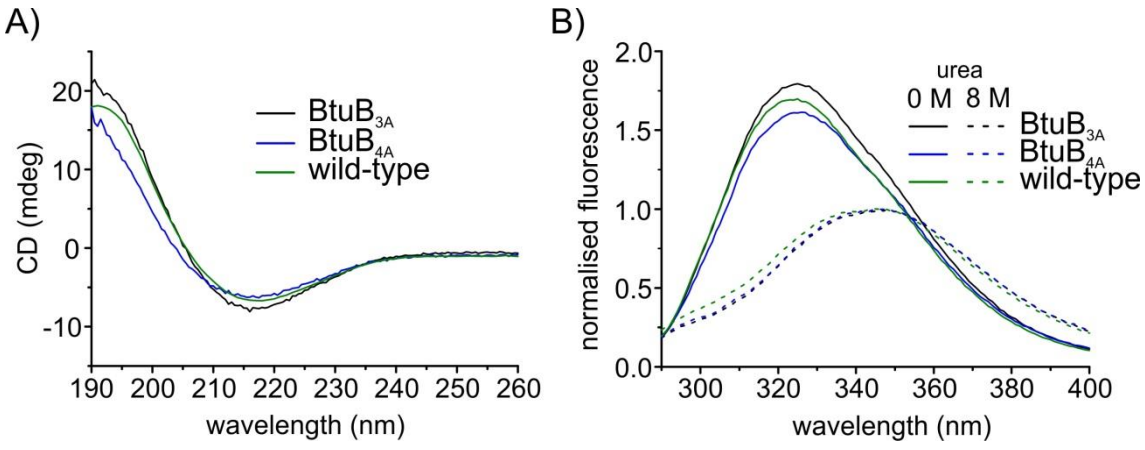


Figure 5.5 – Spectroscopic analysis of the protein fold of the weakened plug mutants. A) Far-UV CD spectra of BtuB mutants and wild-type. B) Intrinsic tryptophan fluorescence in the absence (lines) and presence of 8 M urea (dashes) of BtuB mutants and wild-type.

The variants were then inserted into proteoliposomes using the method described in Section 2.2.8. However, the $BtuB_{4A}$ variant formed a precipitate during dialysis; therefore only the $BtuB_{3A}$ variant was taken forward for analysis by SMFS.

5.2.3 SMFS of the destabilised plug

Once $BtuB_{3A}$ had been inserted into proteoliposomes, it was immobilised onto a mica surface (Section 2.2.8) and SMFS was carried out using an AFM probe covalently attached to $TonB_{\Delta TMD}$ and the same experimental method for wild-type BtuB and disulfide mutants (Section 2.2.7.3). By contrast to wild-type FhuA, BtuB and the other variants studied, the force-extension profiles for $BtuB_{3A}$ displayed three rupture events (Figure 5.6A and Table 5.1). The increase in contour length between the first and second rupture event (ΔL_c1 = 19 ± 4 nm) was similar to that found for wild-type BtuB: $TonB_{\Delta TMD}$ (ΔL_c = 20 ± 1 nm) identifying this event as the unfolding of the mechanically weak subdomain. If the additional rupture event (F_U = 94 ± 9 pN) observed for $BtuB_{3A}$ was due to the unfolding of the remaining 64 plug residues of the strong plug subdomain, the second ΔL_c would be expected to be ~25 nm, which was in good agreement with the measured value (ΔL_{c2} = 26 ± 3 nm) (Figure 5.6B). The difference in contour length between the first and last event (total ΔL_c = 41 ± 7 nm) is equivalent to the unfolding of 103 ± 18 residues, also in good agreement with the number of plug residues in BtuB (114 (residues 22-136), Figure 5.6B inset).

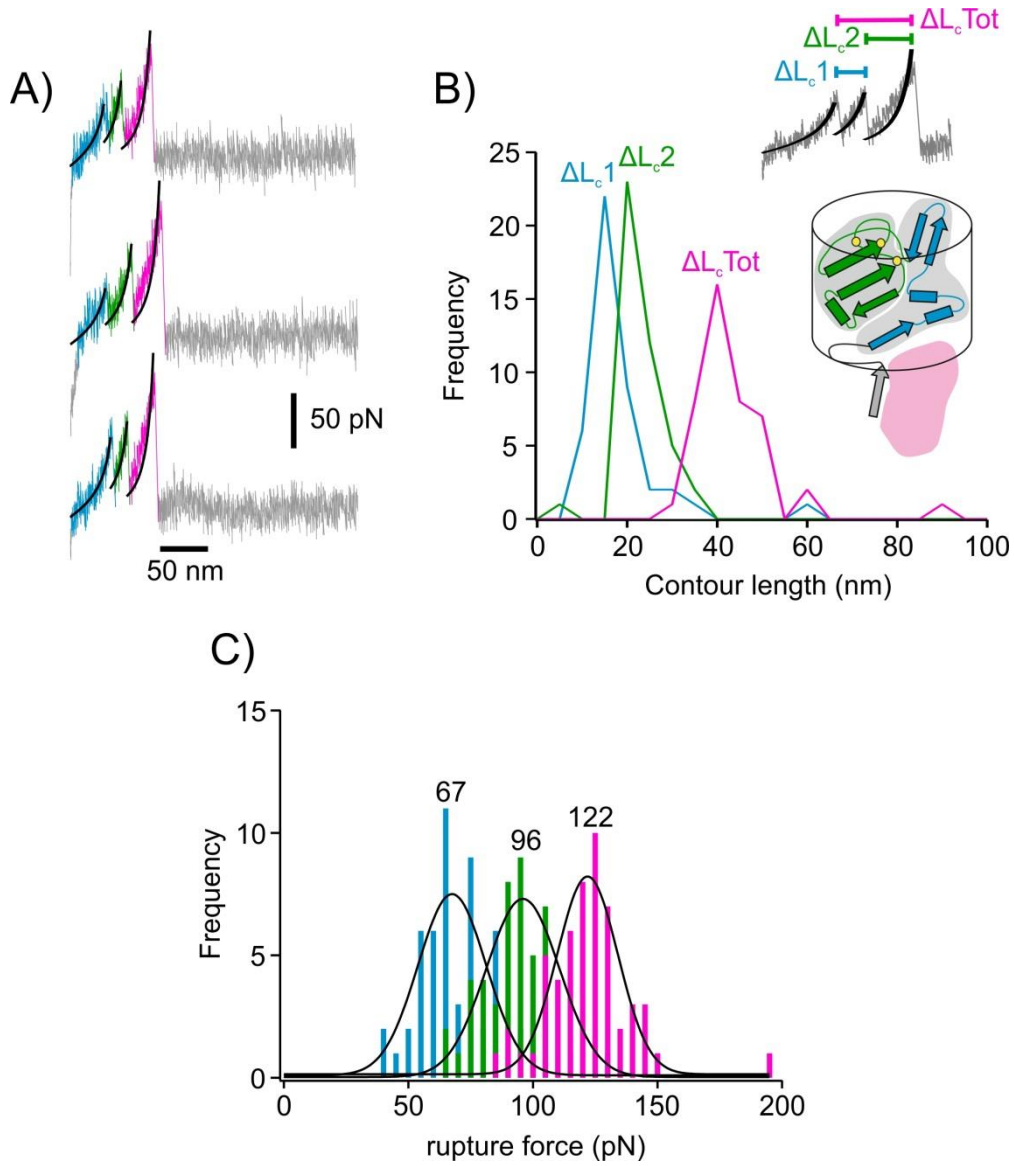


Figure 5.6 - Removal of the hydrophobic core from the mechanically strong plug subdomain leads to complete unfolding of the plug domain *in vitro*. A) Force extension profiles for $BtuB_{3A}$:vitamin B_{12}:$TonB_{\Delta TMD}$ display triple rupture events. Each event is fit to the WLC model (black). Data were obtained at a retraction velocity of 1,000 nms^{-1}. B) ΔL_c frequency histograms for the difference in L_c between the first and second event (ΔL_c1, blue histogram), the second and third event (ΔL_{c2}, green histogram) and the first and third event (ΔL_cTot, pink histogram) (see top inset). As the length of an amino-acid is 0.4 nm, these data can be used to calculate the number of amino-acids unfolded at each event, allowing the unfolding events to be mapped onto the secondary structure schematic (bottom inset: residues unfolded at the first and second rupture events coloured blue and green, respectively. Yellow dots represent the location of each alanine mutation). C) Force frequency histogram for each of the three rupture peaks. Modal values (pN) are shown.

		Peak 1		Peak 2		Peak 3 (TonB unbinding)				
protein	#	Force (pN)	L_c (nm)	Force (pN)	L_c (nm)	Force (pN)	L_c (nm)	ΔL_c 1 (nm)	ΔL_c 2 (nm)	ΔL_cTot (nm)
$BtuB_{3A}$	1	68	50	85	75	117	98	23	28	44
	2	62	51	96	72	120	94	17	23	37
	3	62	49	99	62	125	90	18	27	41
	av	64 ± 4	50 ± 1	94 ± 9	70 ± 5	121 ± 4	94 ± 4	19 ± 4	26 ± 3	41 ± 7

Table 5.1 - AFM data of $TonB_{\Delta TMD}$:$BtuB_{3A}$ at 1000 nms^{-1}. Force, contour length (L_c) and ΔL_c values were calculated by plotting the distribution of events and fitting with a Gaussian to obtain the mean value. The error on the average is the range of the triplicate averages.

5.2.4 Bacitracin sensitivity assay

To find whether TonB can completely unfold the destabilised plug domain *in vivo*, an antibiotic sensitivity screen was performed (See Section 2.2.5.3) using bacitracin, a large (1.4 kDa) inhibitor of cell wall synthesis[211]. As the OM provides a natural barrier to this large antibiotic, *E. coli* are usually unaffected by its presence unless used at high concentrations (200 µg/mL)[211]. Sensitivity to bacitracin has been shown to increase if the integrity of the OM is compromised by addition of polymyxin B[212] or by the expression of a TBDT with a deleted plug domain (FhuAΔ5-160)[213]. Accordingly, the growth of *E. coli* RK5016 expressing wild-type BtuB in LB medium containing various concentrations of bacitracin was unaffected until concentration of 200 µg/mL bacitracin was used (Figure 5.7A). The same strain expressing $BtuB_{3A}$ exhibited reduced growth at concentrations equal or greater than 100 µg/mL bacitracin, showing an increased accessibility of bacitracin to the periplasm (Figure 5.7B), suggesting that complete plug unfolding occurs in the $BtuB_{3A}$ variant.

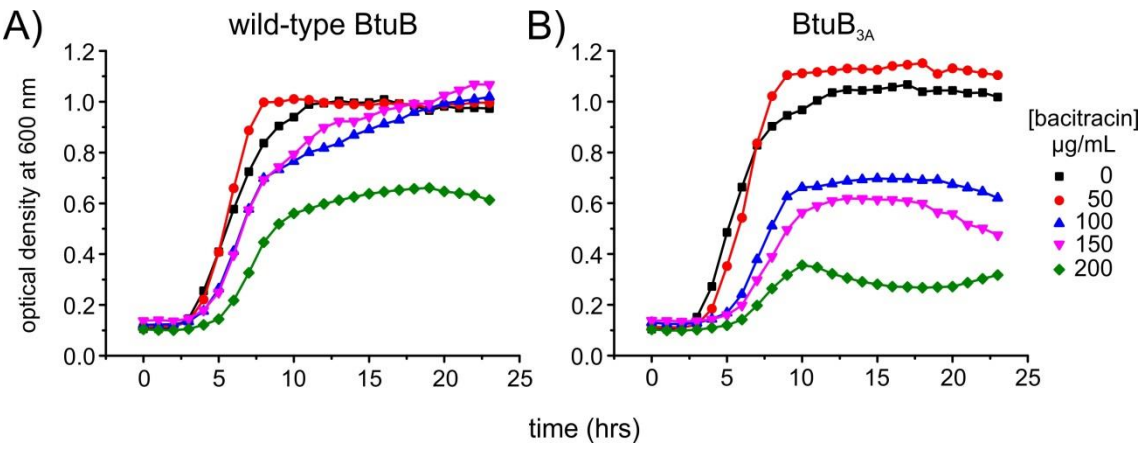


Figure 5.7 - Bacitracin sensitivity assay. RK5016 cells transformed with wild-type BtuB (A) or $BtuB_{3A}$ (B) under a native promoter (pAG1) were used to inoculate LB medium supplemented with various concentrations of bacitracin. The OD_{600} was monitored over 24 hours, the data shows the average of triplicate growth curves.

The increase in bacitracin sensitivity provides evidence that destabilising the mechanically strong subdomain results in an increase BtuB channel size but cannot preclude the possibility that this phenotype is independent of TonB activity. To obviate the possibility that the increase in bacitracin sensitivity was due to the inability of $BtuB_{3A}$ to form a natively folded plug domain during OM insertion, 200 µg/mL of Ton box

pentapeptide (ETVIV) was added to the growth medium to prevent TonB-dependent activity[81]. By preventing TonB-dependent activity, the $BtuB_{3A}$ variant growth phenotype was restored to a similar level as the wild-type receptor under the same conditions (Figure 5.8).

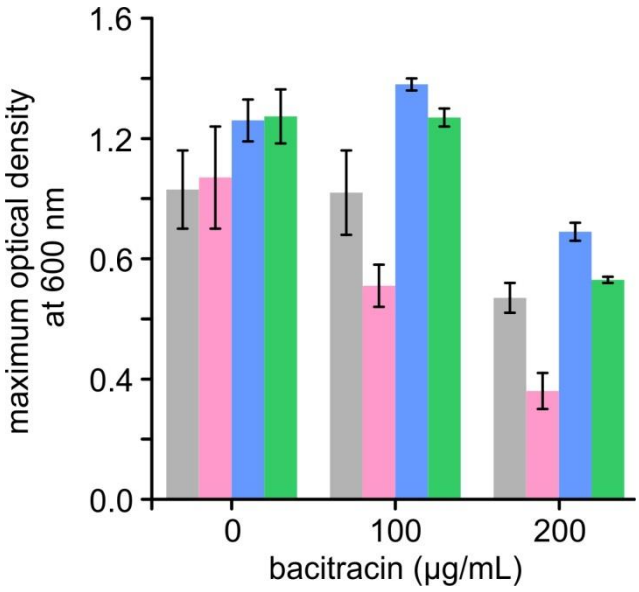


Figure 5.8 - Bacitracin sensitivity assay of RK5016 *E. coli* expressing wild-type BtuB (grey), $BtuB_{3A}$ (pink), and wild-type BtuB and $BtuB_{3A}$ in the presence of 200 µg/mL Ton box pentapeptide (blue and green respectively). The maximum optical density at 600 nm over a 24 hour growth is shown; the error bars show the standard deviation of triplicate growths.

These results suggest that $BtuB_{3A}$ is able to form a native-like plug domain *in vivo* as preventing TonB activity reduces the sensitivity to the antibiotic. It is worth noting that the presence of the Ton box pentapeptide slightly increases the growth of cells expressing the wild-type BtuB receptor, therefore whether the peptide is elicting a stress response is inconclusive. The differential sensitivity to bacitracin of wild-type and $BtuB_{3A}$ in the presence and absence of Ton box pentapeptide accords with the *in vitro* AFM findings above, suggesting that *in vivo*, TonB drives partial (wild-type BtuB) or complete unfolding of the plug ($BtuB_{3A}$) before complex dissociation.

5.3 Discussion

Here, the *in vitro* SMFS findings of Chapter 4 are fully supported by an *in vivo* vitamin B_{12} uptake assay, where the mechanically weak plug subdomain identified in Chapter 4 is cross-linked to the strong plug subdomain or barrel domain to prevent unfolding. By preventing unfolding of this weak subdomain, vitamin B_{12} uptake is reduced/prevented,

which strongly suggests that this region of the plug domain requires unfolding for substrate transport. Removal of the hydrophobic core of the strong plug subdomain allows complete unfolding in a TonB-dependent manner, allowing the free diffusion of large antibiotics such as bacitracin into the periplasmic space. This observation, taken with others[145,206,209,210], suggests that TBDTs have evolved to minimise the deleterious effect of compromising its outermost barrier by allowing the remodelling of the fewest residues required for transit of their cargo.

6 Conclusions and future directions

Membrane proteins are the gatekeepers of cellular compartments and are therefore vital for the transduction of messages across the membrane. By necessity, membrane proteins that transport cargo into or out of compartments introduce “holes” into the membrane barrier. As these pores can be deleterious to the organism (by allowing the entry / exit of toxins/ions/metabolites) their formation and deformation are closely regulated (gated) by a variety of mechanisms. The potential of mechanical deformation to gate membrane proteins is currently highly topical due to its attractiveness as a conceptually simple method to transduce signals across membranes. Several mechano-sensitive networks in Eukayotic cells have been identified with functions as diverse as adhesive GPCR activation[214], Notch and E-cadherin signalling[215] and activation of transmembrane channel–like 1 and 2 by the tip link of stereocilia[216]. However, the molecular mechanism of how force is used to gate proteins is only now beginning to be understood.

Membrane protein transporters can be mechanically gated by either the application of a remodelling force by the lipid bilayer in which the protein resides (force-from-lipid principle), or by application of a remodelling force pulling on tethers on the membrane protein by another protein in the transduction cascade. While equally plausible, the evidence for each mechanism contrasts remarkably. The force from lipid principle was originally demonstrated using MscL/S from *E. coli* 25 years ago and has recently been shown for the Eukaryotic TRAAK and TREK1 K+ potassium channels[217]. By comparison, while potential tethers for membrane transporters have been identified and their importance to mechano-transduction confirmed[218], there is no direct evidence for the ability of one protein to gate a membrane protein. This lack of evidence has been highlighted in a review of mechano-transduction in membrane systems[219]. The strongest evidence for tethered gating was recently reported by Zhang *et al*. who used patches excised from cells and deletion studies *in vivo* to show mechano-gating in the *Drosophila* cation TRP channel NompC[218]. NompC contains a cytoplasmic ankyrin repeat tether domain which was shown to interact with the microtubule cytoskeleton and was absolutely essential for NompC activation. While this provides indirect evidence for

tethered gating, it does not describe a molecular mechanism between two protein partners.

In this thesis, the channel of a purified membrane protein (BtuB), reconstituted into a synthetic liposome has been shown to be gated by the application of force onto the plug domain that blocks the channel. Importantly, in these *in vitro* experiments, the force is directly applied onto the plug domain via its known binding partner *in vivo*, TonB. The use of an antibiotic sensitivity assay in combination with a BtuB variant engineered to have a weakened plug domain has shown that this interaction also drives remodelling *in vivo*. The *in vitro* and *in vivo* data shows for the first time that a non-covalent interaction can mediate the gating of a membrane protein, giving experimental credence to a long held theory.

The potential of using the *pmf* across the inner membrane to drive processes at the outer membrane (inside-out energy transduction) is conceptually attractive. However despite models to explain colicin import[220], TonB-dependent import of nutrients[87], and the Tol-Pal complex for OM integrity and invagination during cell division[221], there has been no direct evidence that a non-covalent protein complex could withstand the effects of force long enough to drive the remodelling required. Therefore, alongside describing the molecular mechanism for TonB-dependent transport, the work of this thesis also provides strong evidence supporting the feasibility of periplasmic inside-out energy transduction for other related systems.

6.1 Open questions and future directions

The major miles stones of understanding TonB-dependent transport over the last half-century are summarised in Figure 6.1 (reviewed in detail in Chapter 1), which demonstrates advances in biochemistry, microbiology, DNA sequencing, X-ray crystallography and spectroscopy are continually opening new exciting discoveries and theories in the field. The entire story is still far from complete; discovery only opens new questions and long standing questions still remain unanswered, some of which will be discussed.

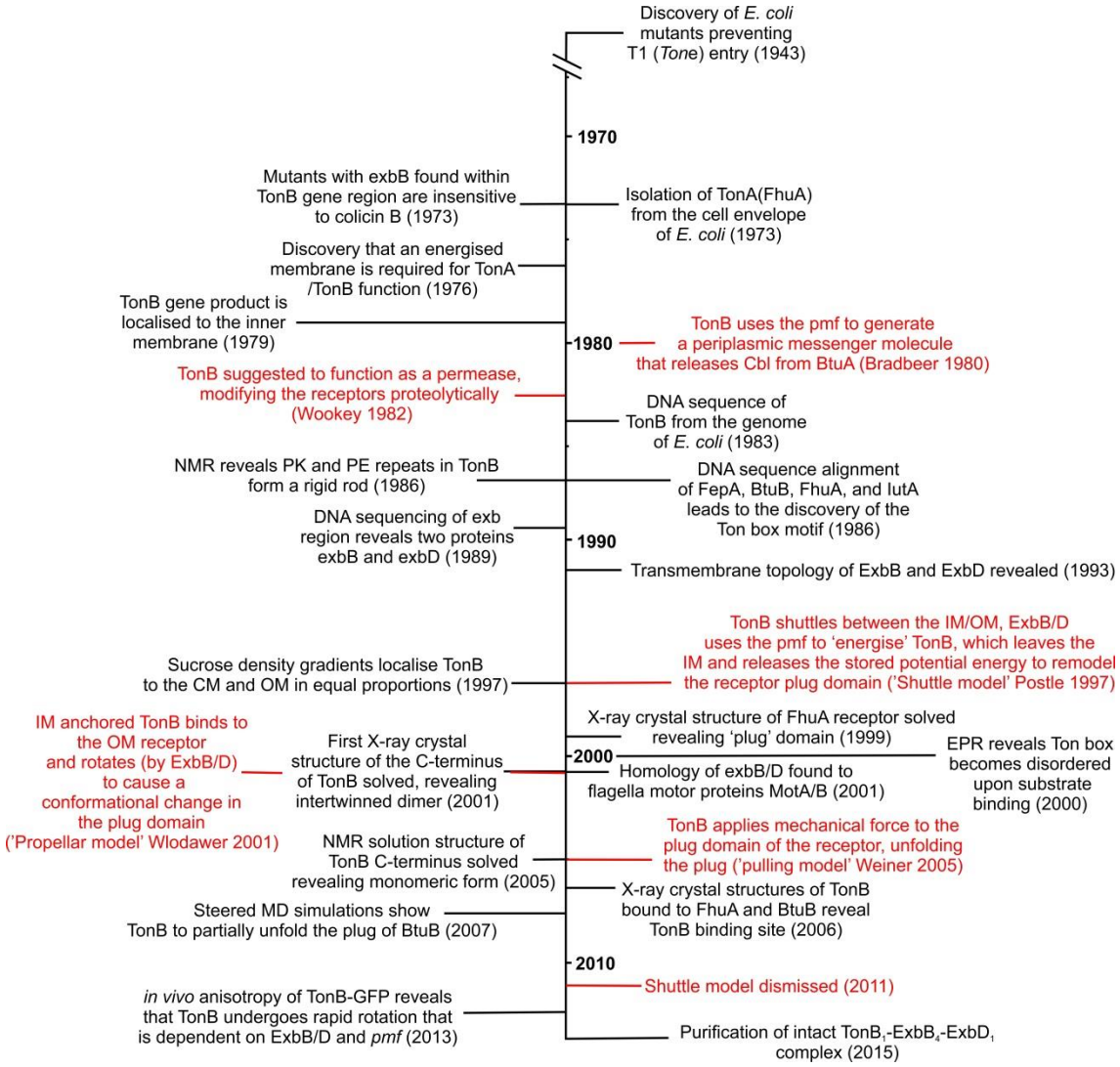


Figure 6.1 – Timeline of the major advances (black) and theories (red) for research into the TonB-dependent transport system.

1) How does TonB apply a pulling force in vivo?

For TonB-dependent transport to function *in vivo*, it requires i) the *pmf* ii) an intact TonB C-terminal domain and transmembrane domain (conserved residues Ser16 and His20 spaced apart by three amino acids) and iii) the inner membrane proteins ExbB and ExbD. There are passive ways in which TonB could assert a force onto the plug domain of the receptor, such as variations in the width of the periplasmic space[198] and /or the differential diffusion of the inner and outer membrane proteins[222]. In these cases, TonB would act as an anchor to the exposed Ton box tether, leaving the functional roles of ExbBD and the *pmf* unclear. Active models which include ExbBD and the *pmf* in the functional process could include a conformational change within the TonB proline-rich

linker domain in an ExbBD/*pmf* dependent manner as discussed in Chapter 3 (Section 3.3.2). TonB has recently been shown to rapidly rotate which is ExbBD/*pmf* dependent[143] and could apply torque to the TonB:TBDT complex that could possibly induce changes in the proline-rich linker domain. It has previously been suggested that torque could destabilise the PPII helix into a PPI helix, driving a 40 % collapse of the domain[100] which would facilitate a pulling force. The formation of a proteinase K-resistant TonB conformation has been shown to only occur when ExbBD and the *pmf* is available and the TM domain of TonB is intact[139]. This distinct energised conformation could represent a pulling conformation or an on-state, such as the electrostatic hinge or the collapsed PPI helix. The role of TonB dimerisation (addressed in question 4) could also play a role in the regulation of the transport cycle.

2) What are the roles of ExbB and ExbD in the transport cycle?

Our current understanding of ExbB and ExbD is limited. With sequence homology to the Mot flagella stator proteins, the notion of ExbBD acting as a motor to rotate TonB has long been a concept, with recent supporting evidence[143]. Structural evidence has shown that the stator protein TolR (homologous to ExbD) exists in two conformations (closed and open) that are likely driven by the *pmf*[223]. In the closed state, TolR closely associates with the IM, however upon activation by the *pmf*, a large conformational change disorders ~25 amino acids that allow the C-terminus of TolR to cross the periplasm and associate with the peptidoglycan (PG). A similar role for ExbD could be present, in that the *pmf* activates ExbD's periplasmic domain where it can then interact with TonB[107] and possibly facilitate a conformational change within the linker domain of TonB that applies the pulling force to the Ton box tether (Figure 6.2). It is also possible that upon Ton box binding, an allosteric change within TonB activates ExbD, where the periplasmic domain is released into the periplasm, opening the channel for proton influx (as proposed for TolR[223]).

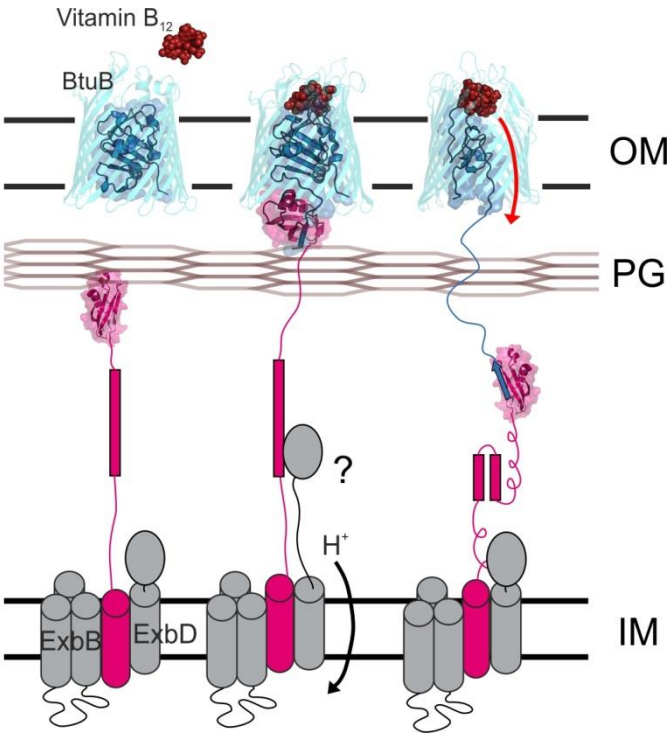


Figure 6.2 – Potential role of the periplasmic domain of ExbD. Upon TonB association with the Ton box, ExbD is activated by the *pmf*, which causes release of its periplasmic domain. This domain can then associate with the proline-rich region of TonB (pink box). This interaction could drive a conformational change which applies a pulling force to the plug domain of the receptor.

As TonB requires to be firmly anchored to the IM to prevent dislodging of the single TM domain under force, it is quite likely that ExbB (the scaffold element) prevents this (TM interactions between ExbB and TonB have been shown). The bulk of ExbB is predicted to be located in the cytoplasm, which has no known function. It would be interesting to investigate whether this region interacts with the bacterial cytoskeleton (MreB) as a related system in *M. xanthus* uses IM proteins (AglR/AglS) and a cytosolic track provided by MreB to generate the force for gliding motility[224]. A similar function could exist for ExbB/D (Figure 6.3), because OMPs have restricted mobility (due to residing in OMP islands[40]) the TonB-dependent receptor will be laterally immobilised. Therefore the lateral movement of the TonB:ExbBD complex energetically driven by ExbD (utilising the *pmf*) moving ExbB along the track provided by MreB and would apply force onto the plug domain. To test this theory directly, a MreB polymerisation inhibitor could be used in combination with a colicin/phage sensitivity assay to detect for changes in TonB-dependent activity.

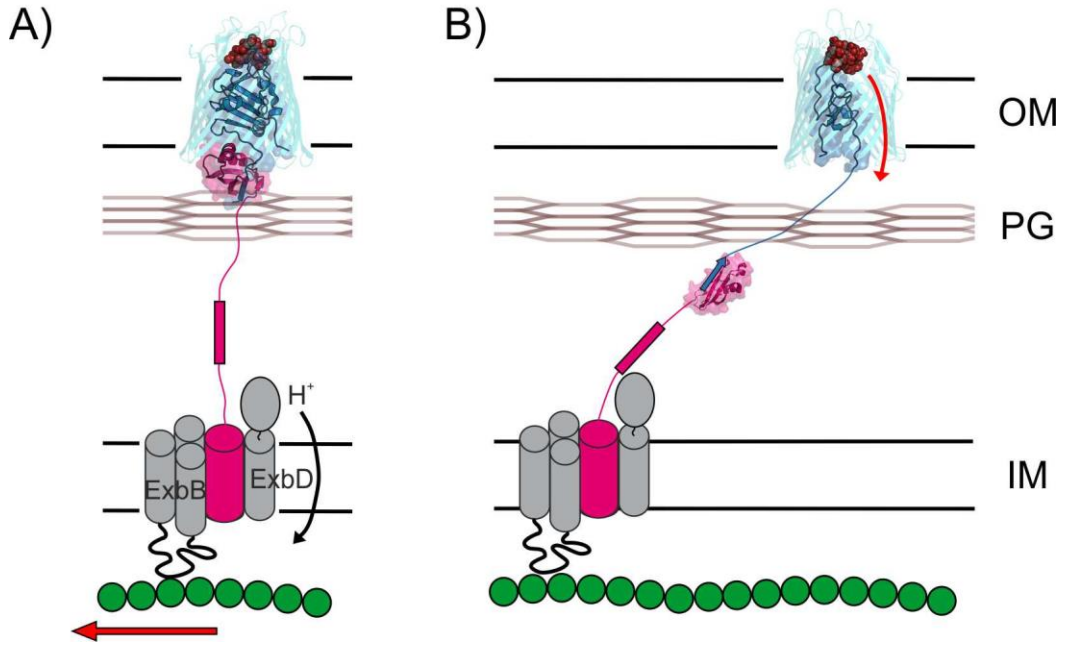


Figure 6.3 – Theoretical role for ExbBD based on the related AlgRS system for bacterial gliding motility in *M. xanthus*. The bulk of ExbB is located in the cytoplasm, where it could interact with the bacterial cytoskeleton (MreB (green circles)). Upon TonB association with the Ton box, ExbD uses the *pmf* to power the lateral movement of ExbB along the MreB track, driving a pulling force to the plug domain of a TBDT (BtuB shown). The TBDT cannot move laterally due to its restriction within an OMP island.

3) Do the TonB PE/PK repeats have a functional role?

The PE/PK rich region of the linker domain has had little known function since its initial discovery. The findings of this thesis could hint towards a possible hinge conformation which could potentially drive a pulling force by withdrawing the C-terminal domain of TonB by up to 28 nm. What is really interesting is that this withdrawal distance is within the range of protein unfolded in both BtuB (20 nm) and FhuA (25 nm) to create a substrate channel through the receptor. Mutagenesis and cross-linking within the PE/PK region would provide valuable insight to whether this conformation is significant or not to TonB function. How the hinge recycles between open and closed states remains uncertain, however this could be regulated by either the ExbD periplasmic domain (shown to weakly interact with the PE repeats by NMR[117]), or the formation of TonB homodimers by the hinge region (Figure 6.4), or even the torque applied to TonB by the ExbBD complex. The fact that TonB-dependent activity was observed in TonB(Δ66-100) *in vivo*[99], reduces the functional relevance of the hinge, however, the notion of being a

length adaptor to compensate for changes in the periplasmic width is equally plausible (see Section 3.3.2).

4) Possible functional role for TonB dimerisation?

The current evidence from previous work and this thesis suggests that TonB homodimers are likely formed by the PE/PK regions (residues 60-110), and could represent an extended off-state during the transport cycle (Figure 6.4). One could speculate that in the dimeric form, TonB does not form the electrostatic hinge and is fully extended allowing TonB to reach the OM. Upon binding of the Ton box, TonB then monomerises (perhaps in an ExbD-dependent fashion), where the electrostatic hinge is formed and the linker domain collapses into a PPII helix causing the retraction of the C-terminal domain and subsequent unfolding of the plug domain (Figure 6.4). However it is worth noting that analysis of the stoichiometry of the TonB:ExbBD complex has never shown more than one copy of TonB per complex (Section 1.4.4), therefore whether multiple complexes form *in vivo* will require further assessment.

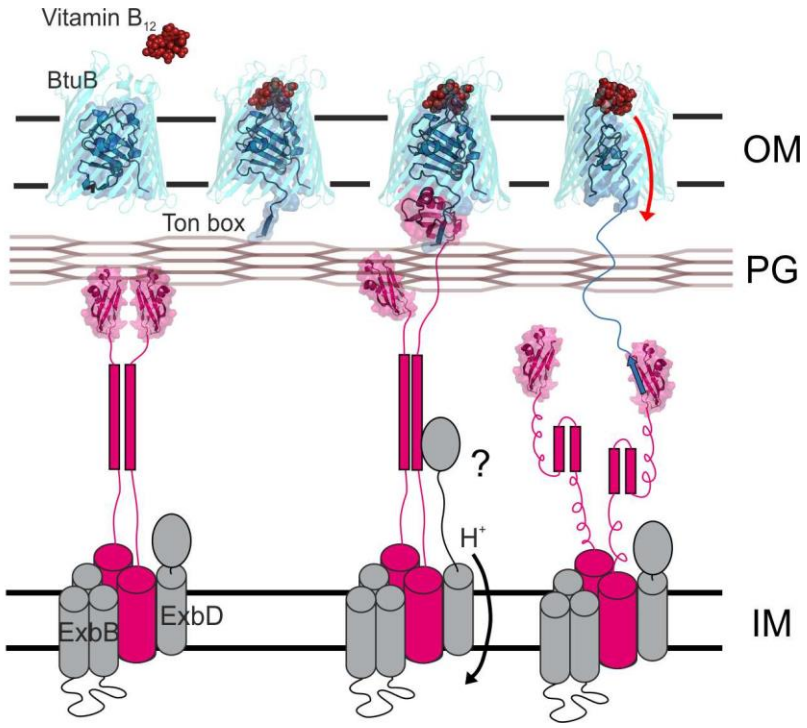


Figure 6.4 – Speculative model of the TonB homodimer function *in vivo*. In the off-state, TonB exist as a homodimer driven by the PE/PK repeat motifs (pink box), the separate C-terminal domains search for an exposed Ton box. Upon binding of the Ton box, the homodimer breaks apart (possibly facilitated by the periplasmic domain of ExbD). The monomeric TonB then forms the electrostatic hinge and a continuous PPII helix (on-state), retracting whilst still in contact with the Ton box, which drives the unfolding of the plug domain.

5) What role does peptidoglycan play?

Often overlooked, the PG sacculus is thought to play a key role in other periplasmic machinery such as the Tol and Mot systems involved in outer-membrane integrity and flagella rotation respectively. Truncated dimeric $TonB_{CTD}$ was shown to weakly interact with PG; similarly TolR (homologous to ExbD) has also been shown to only interact with PG in a dimeric form. However it is not entirely clear whether these interactions are an artefact of construct size and exist *in vivo*. The sequence of TonB was found to contain the lysin (LysM) PG-binding motif[141], but only when $TonB_{CTD}$ existed as an intertwined dimer. The authors suggested that as this motif is away from the Ton box binding site it could function to allow TonB to weakly associate with the PG when scanning for an exposed Ton box. Further studies will be required to elucidate whether the PG interaction is present in longer constructs of TonB, and whether dimeric $TonB_{\Delta TMD}$ is preferential over monomeric forms. Interactions of the periplasmic domain of ExbD with

PG are also likely due to its homology to TolR. Whether or not ExbD anchors the TonB-ExbBD complex to PG during the transport cycle (acting as a stator for TonB rotation) is still unknown.

6) Is plug domain remodelling different for colicins?

An interesting question to address is whether the plug remodelling for colicins is different in comparison to the natural substrate. A good example of this is demonstrated by the pectocin M2 toxin (active against the phytopathogens *Pectobacterium*), which unlike other colicins does not contain an intrinsically unstructured translocation domain (IUTD). The lack of an IUTD implies that the pectocins cannot make direct contact with the Tol or Ton complexes in the periplasm and as a result they are unable to utilise the *pmf* for translocation. Therefore in order to gain access to the cell interior, the pectocin exploits the ferredoxin domain it is coupled to. However the folded ferredoxin (a 30 kDa globular protein, with dimensions 30 x 30 Å) would require dislodgement of the entire plug domain of a TonB-dependent transporter in order to transport this large substrate[225]. It would be interesting to find whether the ferredoxin receptors of *Pectobacterium* have complete plug domain unfolding in comparison to the receptors from *E. coli* studied in this thesis.

7) What happens to the unfolded domain during periplasmic exposure?

Another interesting question to address would be whether periplasmic chaperones (such as Skp and SurA) are involved in the TonB-dependent transport process, as ≥ 50 amino acids from the TBDT are exposed upon substrate transport. Skp has been shown to envelope unfolded OMPs in the periplasmic space (Section 1.2.2), therefore Skp sequestering the unfolded plug domain of a TBDT during transport is conceivable. Other periplasmic chaperones may also aid the refolding of the weak plug subdomain. The plug domain could also contain binding motifs for the recruitment of the periplasmic substrate binding proteins such as BtuF (Figure 1.20).

6.2 Final remarks

The work of this thesis rationalises a large volume of previous observations (Figure 6.1) into a detailed molecular mechanism. The previous section highlights the vast effort still

required to fully understand this complex and interesting system. The next desirable advance will be a high-resolution structure of the entire TonB-ExbBD complex, as with all molecular systems, atomic structures always lead to rapid progress in understanding of function. The isolation of the intact complex ($TonB_1$-$ExbB_4$-$ExbD_1$) in detergent has recently been demonstrated[226] and with constant advances in structural determination techniques, and in particular cubic lipid phase crystallography and cryo-electron microscopy, it is only a matter of time before this major accomplishment will be achieved.

With the world on the verge of a post-antibiotic era, reversing decades of medical progress, the time to resolve the mechanisms of vital molecular machinery governing bacterial survival and virulence could not be more urgent. Over the past 40 years, technology to study life in atomic detail has progressed tremendously. Undoubtedly, over forthcoming years our knowledge of cellular machinery will be ever more detailed with the ultimate goal of developing novel antibiotic therapy to literally put a spanner in the works.

7 Appendix

7.1 Protein and peptide sequences

TonB$_{\Delta TMD}$ (V32C)

MHHHHHHGCHQVIELPAPAQPISVTMVTPADLEPPQAVQPPPEPVVEPEPEPEPIPEPPKEAPVV
IEKPKPKPKPKPKPVKKVQEQPKRDVKPVESRPASPFENTAPARPTSSTATAATSKPVTSVASGPRA
LSRNQPQYPARAQALRIEGQVKVKFDVTPDGRVDNVQILSAKPANMFEREVKNAMRRWRYEPG
KPGSGIVVNILFKINGTTEIQ

TonB$_{CTD}$ (K144C)

MHHHHHHGCPVTSVASGPRALSRNQPQYPARAQALRIEGQVKVKFDVTPDGRVDNVQILSAKP
ANMFEREVKNAMRRWRYEPGKPGSGIVVNILFKINGTTEIQ

TonB$_{\Delta TMD}$

MHHHHHHGHQVIELPAPAQPISVTMVTPADLEPPQAVQPPPEPVVEPEPEPEPIPEPPKEAPVVIE
KPKPKPKPKPKPVKKVQEQPKRDVKPVESRPASPFENTAPARPTSSTATAATSKPVTSVASGPRALS
RNQPQYPARAQALRIEGQVKVKFDVTPDGRVDNVQILSAKPANMFEREVKNAMRRWRYEPGKP
GSGIVVNILFKINGTTEIQ

TonB$_{CTD}$

MHHHHHHGPVTSVASGPRALSRNQPQYPARAQALRIEGQVKVKFDVTPDGRVDNVQILSAKPA
NMFEREVKNAMRRWRYEPGKPGSGIVVNILFKINGTTEIQ

BtuB

MIKKASLLTACSVTAFSAWAQDTSPDTLVVTANRFEQPRSTVLAPTTVVTRQDIDRWQSTSVNDV
LRRLPGVDITQNGGSGQLSSIFIRGTNASHVLVLIDGVRLNLAGVSGSADLSQFPIALVQRVEYIRGP
RSAVYGSDAIGGVVNIITTRDEPGTEISAGWGSNSYQNYDVSTQQQLGDKTRVTLLGDYAHTHGY
DVVAYGNTGTQAQTDNDGFLSKTLYGALEHNFTDAWSGFVRGYGYDNRTNYDAYYSPGSPLLDT
RKLYSQSWDAGLRYNGELIKSQLITSYSHSKDYNYDPHYGRYDSSATLDEMKQYTVQWANNVIVG
HGSIGAGVDWQKQTTTPGTGYVEDGYDQRNTGIYLTGLQQVGDFTFEGAARSDDNSQFGRHGT
WQTSAGWEFIEGYRFIASYGTSYKAPNLGQLYGFYGNPNLDPEKSKQWEGAFEGLTAGVNWRIS

GYRNDVSDLIDYDDHTLKYYNEGKARIKGVEATANFDTGPLTHTVSYDYVDARNAITDTPLLRRAK
QQVKYQLDWQLYDFDWGITYQYLGTRYDKDYSSYPYQTVKMGGVSLWDLAVAYPVTSHLTVRG
KIANLFDKDYETVYGYQTAGREYTLSGSYTF

FhuA

MARSKTAQPKHSLRKIAVVVATAVSGMSVYAQAAVEPKEDTITVTAAPAPQESAWGPAATIAAR
QSATGTKTDTPIQKVPQSISVVTAEEMALHQPKSVKEALSYTPGVSVGTRGASNTYDHLIIRGFAAE
GQSQNNYLNGLKLQGNFYNDAVIDPYMLERAEIMRGPVSVLYGKSSPGGLLNMVSKRPTTEPLKE
VQFKAGTDSLFQTGFDFSDSLDDDGVYSYRLTGLARSANAQQKGSEEQRYAIAPAFTWRPDDKTN
FTFLSYFQNEPETGYYGWLPKEGTVEPLPNGKRLPTDFNEGAKNNTYSRNEKMVGYSFDHEFNDT
FTVRQNLRFAENKTSQNSVYGYGVCSDPANAYSKQCAALAPADKGHYLARKYVVDDEKLQNFSV
DTQLQSKFATGDIDHTLLTGVDFMRMRNDINAWFGYDDSVPLLNLYNPVNTDFDFNAKDPANS
GPYRILNKQKQTGVYVQDQAQWDKVLVTLGGRYDWADQESLNRVAGTTDKRDDKQFTWRGG
VNYLFDNGVTPYFSYSESFEPSSQVGKDGNIFAPSKGKQYEVGVKYVPEDRPIVVTGAVYNLTKTN
NLMADPEGSFFSVEGGEIRARGVEIEAKAALSASVNVVGSYTYTDAEYTTDTTYKGNTPAQVPKH
MASLWADYTFFDGPLSGLTLGTGGRYTGSSYGDPANSFKVGSYTVVDALVRYDLARVGMAGSNV
ALHVNNLFDREYVASCFNTYGCFWGAERQVVATATFRF

Ton box$_{BtuB}$

PDTLVVTANRGSWSC

Ton box$_{BtuB}$ L8P

PDTPVVTANRGSWSC

Ton box pentapeptide

ETVIV

References

1. Lane, N. The unseen world: reflections on Leeuwenhoek (1677) 'Concerning little animals'. *Philos Trans R Soc Lond B Biol Sci* **370**(2015).
2. Kleanthous, C. & Armitage, J.P. The bacterial cell envelope. *Philos Trans R Soc Lond B Biol Sci* **370**(2015).
3. Kellenberger, E. & Ryter, A. Cell wall and cytoplasmic membrane of *Escherichia coli*. *J Biophys Biochem Cy* **4**, 323-& (1958).
4. Heppel, L.A. Selective release of enzymes from bacteria. *Science* **156**, 1451-& (1967).
5. Vanwielink, J.E. & Duine, J.A. How big is the periplasmic space. *Trends Biochem Sci* **15**, 136-137 (1990).
6. Nakamoto, H. & Bardwell, J.C. Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochim Biophys Acta* **1694**, 111-9 (2004).
7. Braun, V. Covalent lipoprotein from the outer membrane of *Escherichia coli*. *Biochim Biophys Acta* **415**, 335-77 (1975).
8. Kamio, Y. & Nikaido, H. Outer membrane of Salmonella typhimurium: accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium. *Biochemistry* **15**, 2561-70 (1976).
9. Raetz, C.R. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**, 635-700 (2002).
10. Bryant, C.E., Spring, D.R., Gangloff, M. & Gay, N.J. The molecular basis of the host response to lipopolysaccharide. *Nat Rev Microbiol* **8**, 8-14 (2010).
11. Burns, S.M. & Hull, S.I. Comparison of loss of serum resistance by defined lipopolysaccharide mutants and an acapsular mutant of uropathogenic *Escherichia coli* O75:K5. *Infect Immun* **66**, 4244-53 (1998).
12. Weiss, J., Hutzler, M. & Kao, L. Environmental modulation of lipopolysaccharide chain length alters the sensitivity of *Escherichia coli* to the neutrophil bactericidal/permeability-increasing protein. *Infect Immun* **51**, 594-9 (1986).
13. Tsujimoto, H., Gotoh, N. & Nishino, T. Diffusion of macrolide antibiotics through the outer membrane of Moraxella catarrhalis. *J Infect Chemother* **5**, 196-200 (1999).
14. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol R* **67**, 593-656 (2003).
15. Putker, F., Bos, M.P. & Tommassen, J. Transport of lipopolysaccharide to the Gram-negative bacterial cell surface. *Fems Microbiol Rev* **39**, 985-1002 (2015).
16. Dong, H.H., Xiang, Q.J., Gu, Y.H., Wang, Z.S., Paterson, N.G., Stansfeld, P.J., He, C., Zhang, Y.Z., Wang, W.J. & Dong, C.J. Structural basis for outer membrane lipopolysaccharide insertion. *Nature* **511**, 52-U569 (2014).
17. Qiao, S., Luo, Q.S., Zhao, Y., Zhang, X.J.C. & Huang, Y.H. Structural basis for lipopolysaccharide insertion in the bacterial outer membrane. *Nature* **511**, 108-U523 (2014).
18. Narita, S. & Tokuda, H. Biochemical characterization of an ABC transporter LptBFGC complex required for the outer membrane sorting of lipopolysaccharides. *Febs Lett* **583**, 2160-2164 (2009).

19. Tran, A.X., Trent, M.S. & Whitfield, C. The LptA protein of *Escherichia coli* is a periplasmic lipid a-binding protein involved in the lipopolysaccharide export pathway. *J Biol Chem* **283**, 20342-20349 (2008).
20. Suits, M.D.L., Sperandeo, P., Deho, G., Polissi, A. & Jia, Z.C. Novel structure of the conserved Gram-negative lipopolysaccharide transport protein A and mutagenesis analysis. *J Mol Biol* **380**, 476-488 (2008).
21. Miyadai, H., Tanaka-Masuda, K., Matsuayama, S. & Tokuda, H. Effects of lipoprotein overproduction on the induction of DegP (HtrA) involved in quality control in the *Escherichia coli* periplasm. *J Biol Chem* **279**, 39807-39813 (2004).
22. Konovalova, A. & Silhavy, T.J. Outer membrane lipoprotein biogenesis: Lol is not the end. *Philos Trans R Soc Lond B Biol Sci* **370**(2015).
23. Silhavy, T.J., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harbor perspectives in biology* **2**, a000414 (2010).
24. Kleanthous, C. Swimming against the tide: progress and challenges in our understanding of colicin translocation. *Nat Rev Microbiol* **8**, 843-848 (2010).
25. Cascales, E., Gavioli, M., Sturgis, J.N. & Lloubes, R. Proton motive force drives the interaction of the inner membrane TolA and outer membrane Pal proteins in *Escherichia coli*. *Mol Microbiol* **38**, 904-915 (2000).
26. Hwang, P.M., Choy, W.Y., Lo, E.I., Chen, L., Forman-Kay, J.D., Raetz, C.R., Prive, G.G., Bishop, R.E. & Kay, L.E. Solution structure and dynamics of the outer membrane enzyme PagP by NMR. *P Natl Acad Sci USA* **99**, 13560-5 (2002).
27. Vandeputte-Rutten, L., Kramer, R.A., Kroon, J., Dekker, N., Egmond, M.R. & Gros, P. Crystal structure of the outer membrane protease OmpT from *Escherichia coli* suggests a novel catalytic site. *Embo J* **20**, 5033-9 (2001).
28. Snijder, H.J., Ubarretxena-Belandia, I., Blaauw, M., Kalk, K.H., Verheij, H.M., Egmond, M.R., Dekker, N. & Dijkstra, B.W. Structural evidence for dimerization-regulated activation of an integral membrane phospholipase. *Nature* **401**, 717-21 (1999).
29. Garcia, E.C., Brumbaugh, A.R. & Mobley, H.L. Redundancy and specificity of *Escherichia coli* iron acquisition systems during urinary tract infection. *Infect Immun* **79**, 1225-35 (2011).
30. Nakae, T. Outer membrane of Salmonella - Isolation of protein complex that produces transmembrane channels. *J Biol Chem* **251**, 2176-2178 (1976).
31. Schein, S.J., Colombini, M. & Finkelstein, A. Reconstitution in planar lipid bilayers of a voltage-dependent anion-selective channel obtained from paramecium mitochondria. *J Membrane Biol* **30**, 99-120 (1976).
32. Cowan, S.W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R.A., Jansonius, J.N. & Rosenbusch, J.P. Crystal structures explain functional properties of two E. coli porins. *Nature* **358**, 727-33 (1992).
33. Koebnik, R., Locher, K.P. & Van Gelder, P. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol Microbiol* **37**, 239-253 (2000).
34. Schulz, G.E. The structure of bacterial outer membrane proteins. *Bba-Biomembranes* **1565**, 308-317 (2002).
35. Schulz, G.E. Structure-function-relationships in the membrane channel porin as based on a 1.8 angstrom resolution crystal-structure. *Jerus Sym Q* **25**, 403-412 (1992).

36. Wimley, W.C. & White, S.H. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Biol* **3**, 842-848 (1996).
37. Yau, W.M., Wimley, W.C., Gawrisch, K. & White, S.H. The preference of tryptophan for membrane interfaces. *Biochemistry* **37**, 14713-14718 (1998).
38. Kock, J., Olschlager, T., Kamp, R.M. & Braun, V. Primary structure of colicin-M, an inhibitor of murein biosynthesis. *J Bacteriol* **169**, 3358-3361 (1987).
39. Van Gelder, P. & Tommassen, J. Demonstration of a folded monomeric form of porin PhoE of *Escherichia coli* in vivo. *J Bacteriol* **178**, 5320-2 (1996).
40. Rassam, P., Copeland, N.A., Birkholz, O., Toth, C., Chavent, M., Duncan, A.L., Cross, S.J., Housden, N.G., Kaminska, R., Seger, U., Quinn, D.M., Garrod, T.J., Sansom, M.S.P., Piehler, J., Baumann, C.G. & Kleanthous, C. Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria. *Nature* **523**, 333-6 (2015).
41. Dong, H., Xiang, Q., Gu, Y., Wang, Z., Paterson, N.G., Stansfeld, P.J., He, C., Zhang, Y., Wang, W. & Dong, C. Structural basis for outer membrane lipopolysaccharide insertion. *Nature* **511**, 52-6 (2014).
42. Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T.J. & Kahne, D. Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* **121**, 235-245 (2005).
43. Rollauer, S.E., Sooreshjani, M.A., Noinaj, N. & Buchanan, S.K. Outer membrane protein biogenesis in Gram-negative bacteria. *Philos Trans R Soc Lond B Biol Sci* **370**(2015).
44. Gu, Y.H., Li, H.Y., Dong, H.H., Zeng, Y., Zhang, Z.Y., Paterson, N.G., Stansfeld, P.J., Wang, Z.S., Zhang, Y.Z., Wang, W.J. & Dong, C.J. Structural basis of outer membrane protein insertion by the BAM complex. *Nature* **531**, 64-9 (2016).
45. Noinaj, N., Kuszak, A.J., Balusek, C., Gumbart, J.C. & Buchanan, S.K. Lateral opening and exit pore formation are required for bama function. *Structure* **22**, 1055-1062 (2014).
46. Noinaj, N., Kuszak, A.J., Gumbart, J.C., Lukacik, P., Chang, H.S., Easley, N.C., Lithgow, T. & Buchanan, S.K. Structural insight into the biogenesis of beta-barrel membrane proteins. *Nature* **501**, 385-90 (2013).
47. Struyve, M., Moons, M. & Tommassen, J. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer-membrane protein. *J Mol Biol* **218**, 141-148 (1991).
48. Rouviere, P.E. & Gross, C.A. SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Gene Dev* **10**, 3170-3182 (1996).
49. Bitto, E. & McKay, D.B. The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *J Biol Chem* **278**, 49316-49322 (2003).
50. Bitto, E. & McKay, D.B. Crystallographic structure of SurA, a molecular chaperone that facilitates folding of outer membrane porins. *Structure* **10**, 1489-1498 (2002).
51. Sklar, J.G., Wu, T., Kahne, D. & Silhavy, T.J. Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Gene Dev* **21**, 2473-2484 (2007).

52. Horne, J.E. & Radford, S.E. A growing toolbox of techniques for studying beta-barrel outer membrane protein folding and biogenesis. *Biochem Soc T* **44**, 802-9 (2016).
53. Walton, T.A., Sandoval, C.M., Fowler, C.A., Pardi, A. & Sousa, M.C. The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *P Natl Acad Sci USA* **106**, 1772-1777 (2009).
54. Burmann, B.M., Wang, C.W. & Hiller, S. Conformation and dynamics of the periplasmic membrane-protein- chaperone complexes OmpX-Skp and tOmpA-Skp. *Nat Struct Mol Biol* **20**, 1265-U186 (2013).
55. Pallen, M.J. & Wren, B.W. The HtrA family of serine proteases. *Mol Microbiol* **26**, 209-221 (1997).
56. Rizzitello, A.E., Harper, J.R. & Silhavy, T.J. Genetic evidence for parallel pathways of chaperone activity in the periplasm of *Escherichia coli*. *J Bacteriol* **183**, 6794-6800 (2001).
57. Basle, A., Rummel, G., Storici, P., Rosenbusch, J.P. & Schirmer, T. Crystal structure of osmoporin OmpC from *E. coli* at 2.0 angstrom. *J Mol Biol* **362**, 933-942 (2006).
58. Van Gelder, P., Saint, N., Phale, P., Eppens, E.F., Prilipov, A., van Boxtel, R., Rosenbusch, J.P. & Tommassen, J. Voltage sensing in the PhoE and OmpF outer membrane porins of *Escherichia coli*: role of charged residues. *J Mol Biol* **269**, 468-72 (1997).
59. Alphen, W.V. & Lugtenberg, B. Influence of osmolarity of the growth medium on the outer membrane protein pattern of *Escherichia coli*. *J Bacteriol* **131**, 623-30 (1977).
60. Delcour, A.H. Outer membrane permeability and antibiotic resistance. *Bba-Proteins Proteom* **1794**, 808-816 (2009).
61. Schirmer, T., Keller, T.A., Wang, Y.F. & Rosenbusch, J.P. Structural basis for sugar translocation through maltoporin channels at 3.1-angstrom resolution. *Science* **267**, 512-514 (1995).
62. Dutzler, R., Wang, Y.F., Rizkallah, P.J., Rosenbusch, J.P. & Schirmer, T. Crystal structures of various maltooligosaccharides bound to maltoporin reveal a specific sugar translocation pathway. *Structure* **4**, 127-134 (1996).
63. Luria, S.E. & Delbruck, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**, 491-511 (1943).
64. Braun, V., Schaller, K. & Wolff, H. Common receptor protein for phage T5 and colicin M in outer membrane of *Escherichia coli*-B. *Biochim Biophys Acta* **323**, 87-97 (1973).
65. Randall-Hazelbauer, L. & Schwartz, M. Isolation of the bacteriophage lambda receptor from *Escherichia coli*. *J Bacteriol* **116**, 1436-46 (1973).
66. Szmelcman, S. & Hofnung, M. Maltose transport in *Escherichia coli* K-12: involvement of the bacteriophage lambda receptor. *J Bacteriol* **124**, 112-8 (1975).
67. Sabet, S.F. & Schnaitman, C.A. Purification and properties of the colicin E3 receptor of *Escherichia coli*. *J Biol Chem* **248**, 1797-806 (1973).
68. Di Masi, D.R., White, J.C., Schnaitman, C.A. & Bradbeer, C. Transport of vitamin B12 in *Escherichia coli*: common receptor sites for vitamin B_{12} and the E colicins on the outer membrane of the cell envelope. *J Bacteriol* **115**, 506-13 (1973).

69. Kadner, R.J., Heller, K., Coulton, J.W. & Braun, V. Genetic control of hydroxamate-mediated iron uptake in *Escherichia coli*. *J Bacteriol* **143**, 256-64 (1980).
70. Garen, A. & Puck, T.T. The first two steps of the invasion of host cells by bacterial viruses. II. *J Exp Med* **94**, 177-89 (1951).
71. Hancock, R.W. & Braun, V. Nature of the energy requirement for the irreversible adsorption of bacteriophages T1 and phi80 to *Escherichia coli*. *J Bacteriol* **125**, 409-15 (1976).
72. Hantke, K. & Braun, V. Membrane receptor dependent iron transport in *Escherichia coli*. *Febs Lett* **49**, 301-5 (1975).
73. Hantke, K. & Braun, V. Functional interaction of the tonA/tonB receptor system in *Escherichia coli*. *J Bacteriol* **135**, 190-7 (1978).
74. Bassford, P.J., Bradbeer, C., Kadner, R.J. & Schnaitman, C.A. Transport of vitamin B_{12} in tonb mutants of *Escherichia coli*. *J Bacteriol* **128**, 242-247 (1976).
75. Heller, K. & Kadner, R.J. Nucleotide-sequence of the gene for the vitamin B_{12} receptor protein in the outer-membrane of *Escherichia coli*. *J Bacteriol* **161**, 904-908 (1985).
76. Coulton, J.W., Mason, P., Cameron, D.R., Carmel, G., Jean, R. & Rode, H.N. Protein fusions of beta-galactosidase to the ferrichrome-iron receptor of *Escherichia coli* K-12. *J Bacteriol* **165**, 181-192 (1986).
77. Krone, W.J.A., Stegehuis, F., Koningstein, G., Vandoorn, C., Roosendaal, B., Degraaf, F.K. & Oudega, B. Characterization of the Pco1v-K30 encoded cloacin Df13/Aerobactin outer-membrane receptor protein of *Escherichia coli* - Isolation and purification of the protein and analysis of its nucleotide-sequence and primary structure. *Fems Microbiol Lett* **26**, 153-161 (1985).
78. Lundrigan, M.D. & Kadner, R.J. Nucleotide-sequence of the gene for the ferrienterochelin receptor Fepa in *Escherichia coli* - Homology among outer-membrane receptors that interact with TonB. *J Biol Chem* **261**, 797-801 (1986).
79. Gudmundsdottir, A., Bell, P.E., Lundrigan, M.D., Bradbeer, C. & Kadner, R.J. Point mutations in a conserved region (Tonb Box) of *Escherichia coli* outer-membrane protein BtuB affect vitamin B_{12} transport. *J Bacteriol* **171**, 6526-6533 (1989).
80. Schoffler, H. & Braun, V. Transport across the outer-membrane of *Escherichia coli*-K12 via the Fhua receptor is regulated by the TonB protein of the cytoplasmic membrane. *Mol Gen Genet* **217**, 378-383 (1989).
81. Tuckman, M. & Osburne, M.S. *In vivo* inhibition of TonB-dependent processes by a TonB box consensus pentapeptide. *J Bacteriol* **174**, 320-323 (1992).
82. Skare, J.T., Ahmer, B.M.M., Seachord, C.L., Darveau, R.P. & Postle, K. Energy transduction between membranes - TonB, a cytoplasmic membrane-protein can be chemically cross-linked in vivo to the outer-membrane receptor FepA. *J Biol Chem* **268**, 16302-16308 (1993).
83. Larsen, R.A., FosterHartnett, D., McIntosh, M.A. & Postle, K. Regions of *Escherichia coli* TonB and FepA proteins essential for in vivo physical interactions. *J Bacteriol* **179**, 3213-3221 (1997).
84. Ogierman, M. & Braun, V. Interactions between the outer membrane ferric citrate transporter FecA and TonB: Studies of the FecA TonB box. *J Bacteriol* **185**, 1870-1885 (2003).

85. Cadieux, N. & Kadner, R.J. Site-directed disulfide bonding reveals an interaction site between energy-coupling protein TonB and BtuB, the outer membrane cobalamin transporter. *P Natl Acad Sci USA* **96**, 10673-10678 (1999).
86. Peacock, R.S., Weljie, A.M., Howard, S.P., Price, F.D. & Vogel, H.J. The solution structure of the C-terminal domain of TonB and interaction studies with TonB box peptides. *J Mol Biol* **345**, 1185-1197 (2005).
87. Pawelek, P.D., Croteau, N., Ng-Thow-Hing, C., Khursigara, C.M., Moiseeva, N., Allaire, M. & Coulton, J.W. Structure of TonB in complex with FhuA, *E. coli* outer membrane receptor. *Science* **312**, 1399-1402 (2006).
88. Shultis, D.D., Purdy, M.D., Banchs, C.N. & Wiener, M.C. Outer membrane active transport: Structure of the BtuB : TonB complex. *Science* **312**, 1396-1399 (2006).
89. Noinaj, N., Guillier, M., Barnard, T.J. & Buchanan, S.K. TonB-dependent transporters: regulation, structure, and function. *Annu Rev Microbiol* **64**, 43-60 (2010).
90. Fanucci, G.E., Cadieux, N., Kadner, R.J. & Cafiso, D.S. Competing ligands stabilize alternate conformations of the energy coupling motif of a TonB-dependent outer membrane transporter. *P Natl Acad Sci USA* **100**, 11382-11387 (2003).
91. Fanucci, G.E., Lee, J.Y. & Cafiso, D.S. Spectroscopic evidence that osmolytes used in crystallization buffers inhibit a conformation change in a membrane protein. *Biochemistry* **42**, 13106-13112 (2003).
92. Postle, K. & Good, R.F. DNA sequence of the *Escherichia coli* tonB gene. *P Natl Acad Sci USA* **80**, 5235-9 (1983).
93. Evans, J.S., Levine, B.A., Trayer, I.P., Dorman, C.J. & Higgins, C.F. Sequence-imposed structural constraints in the TonB protein of *Escherichia coli*. *Febs Lett* **208**, 211-216 (1986).
94. Postle, K. & Skare, J.T. *Escherichia coli* TonB protein is exported from the cytoplasm without proteolytic cleavage of its amino terminus. *J Biol Chem* **263**, 11000-11007 (1988).
95. Plastow, G.S. & Holland, I.B. Identification of an *Escherichia coli* inner membrane polypeptide specified by a lambda-TonB transducing bacteriophage. *Biochem Bioph Res Co* **90**, 1007-1014 (1979).
96. Hannavy, K., Barr, G.C., Dorman, C.J., Adamson, J., Mazengera, L.R., Gallagher, M.P., Evans, J.S., Levine, B.A., Trayer, I.P. & Higgins, C.F. Tonb protein of Salmonella-Typhimurium - a model for signal transduction between membranes. *J Mol Biol* **216**, 897-910 (1990).
97. Karlsson, M., Hannavy, K. & Higgins, C.F. A sequence-specific function for the N-terminal signal-like sequence of the TonB protein. *Mol Microbiol* **8**, 379-388 (1993).
98. Larsen, R.A. & Postle, K. Conserved residues Ser(16) and His(20) and their relative positioning are essential for TonB activity, cross-linking of TonB with ExbB, and the ability of TonB to respond to proton motive force. *J Biol Chem* **276**, 8111-8117 (2001).
99. Larsen, R.A., Wood, G.E. & Postle, K. The conserved proline-rich motif is not essential for energy transduction by *Escherichia coli* Tonb protein. *Mol Microbiol* **10**, 943-953 (1993).
100. Kohler, S.D., Weber, A., Howard, S.P., Welte, W. & Drescher, M. The proline-rich domain of TonB possesses an extended polyproline II-like conformation of

sufficient length to span the periplasm of Gram-negative bacteria. *Protein Sci* **19**, 625-630 (2010).
101. Graham, L.L., Harris, R., Villiger, W. & Beveridge, T.J. Freeze-substitution of Gram-Negative Eubacteria - general cell morphology and envelope profiles. *J Bacteriol* **173**, 1623-1633 (1991).
102. Chang, C.S., Mooser, A., Pluckthun, A. & Wlodawer, A. Crystal structure of the dimeric C-terminal domain of TonB reveals a novel fold. *J Biol Chem* **276**, 27535-27540 (2001).
103. Postle, K., Kastead, K.A., Gresock, M.G., Ghosh, J. & Swayne, C.D. The TonB dimeric crystal structures do not exist In Vivo. *Mbio* **1**(2010).
104. Kodding, J., Killig, F., Polzer, P., Howard, S.P., Diederichs, K. & Welte, W. Crystal structure of a 92-residue C-terminal fragment of TonB from *Escherichia coli* reveals significant conformational changes compared to structures of smaller TonB fragments. *J Biol Chem* **280**, 3022-3028 (2005).
105. Sauter, A., Howard, S.P. & Braun, V. In vivo evidence for TonB dimerization. *J Bacteriol* **185**, 5747-5754 (2003).
106. Freed, D.M., Lukasik, S.M., Sikora, A., Mokdad, A. & Cafiso, D.S. Monomeric TonB and the Ton Box are required for the formation of a high-affinity transporter-TonB complex. *Biochemistry* **52**, 2638-2648 (2013).
107. Gresock, M.G., Kastead, K.A. & Postle, K. From homodimer to heterodimer and back: elucidating the TonB energy transduction cycle. *J Bacteriol* **197**, 3433-3445 (2015).
108. Guterman, S.K. Colicin-B - Mode of action and inhibition by enterochelin. *J Bacteriol* **114**, 1217-1224 (1973).
109. Eickhelmerich, K. & Braun, V. Import of bio-polymers into *Escherichia coli* - nucleotide-sequences of the ExbB and ExbD genes are homologous to those of the TolQ and TolR genes, respectively. *J Bacteriol* **171**, 5117-5126 (1989).
110. Sun, T.P. & Webster, R.E. Nucleotide sequence of a gene cluster involved in entry of E colicins and single-stranded DNA of infecting filamentous bacteriophages into *Escherichia coli*. *J Bacteriol* **169**, 2667-74 (1987).
111. Kampfenkel, K. & Braun, V. Membrane topology of the *Escherichia coli* ExbD protein. *J Bacteriol* **174**, 5485-5487 (1992).
112. Kampfenkel, K. & Braun, V. Topology of the ExbB protein in the cytoplasmic membrane of *Escherichia coli*. *J Biol Chem* **268**, 6050-6057 (1993).
113. Postle, K. & Kadner, R.J. Touch and go: tying TonB to transport. *Mol Microbiol* **49**, 869-882 (2003).
114. Skare, J.T. & Postle, K. Evidence for a TonB-dependent energy transduction complex in *Escherichia coli*. *Mol Microbiol* **5**, 2883-2890 (1991).
115. Braun, V., Gaisser, S., Herrmann, C., Kampfenkel, K., Killmann, H. & Traub, I. Energy-coupled transport across the outer membrane of *Escherichia coli*: ExbB binds ExbD and TonB in vitro, and leucine 132 in the periplasmic region and aspartate 25 in the transmembrane region are important for ExbD activity. *J Bacteriol* **178**, 2836-2845 (1996).
116. Larsen, R.A., Thomas, M.G., Wood, G.E. & Postle, K. Partial suppression of an *Escherichia coli* TonB transmembrane domain mutation (delta V17) by a missense mutation in ExbB. *Mol Microbiol* **13**, 627-40 (1994).

117. Garcia-Herrero, A., Peacock, R.S., Howard, S.P. & Vogel, H.J. The solution structure of the periplasmic domain of the TonB system ExbD protein reveals an unexpected structural homology with siderophore-binding proteins. *Mol Microbiol* **66**, 872-889 (2007).
118. Ollis, A.A. & Post, K. Identification of Functionally important TonB-ExbD periplasmic domain interactions In vivo. *J Bacteriol* **194**, 3078-3087 (2012).
119. Ollis, A.A. & Postle, K. The same periplasmic ExbD residues mediate In vivo interactions between ExbD homodimers and ExbD-TonB heterodimers. *J Bacteriol* **193**, 6852-6863 (2011).
120. Fischer, E., Gunter, K. & Braun, V. Involvement of ExbB and TonB in transport across the outer-membrane of *Escherichia coli* - phenotypic complementation of exb mutants by overexpressed TonB and physical stabilization of TonB by ExbB. *J Bacteriol* **171**, 5127-5134 (1989).
121. Jana, B., Manning, M. & Postle, K. Mutations in the ExbB cytoplasmic carboxy terminus prevent energy-dependent interaction between the TonB and ExbD periplasmic domains. *J Bacteriol* **193**, 5649-5657 (2011).
122. Baker, K.R. & Postle, K. Mutations in *Escherichia coli* ExbB transmembrane domains identify scaffolding and signal transduction functions and exclude participation in a proton pathway. *J Bacteriol* **195**, 2898-2911 (2013).
123. Higgs, P.I., Larsen, R.A. & Postle, K. Quantification of known components of the *Escherichia coli* TonB energy transduction system: TonB, ExbB, ExbD and FepA. *Mol Microbiol* **44**, 271-281 (2002).
124. Sverzhinsky, A., Chung, J.W., Deme, J.C., Fabre, L., Levey, K.T., Plesa, M., Carter, D.M., Lypaczewski, P. & Coulton, J.W. Membrane protein complex ExbB(4)-ExbD(1)-TonB(1) from *Escherichia coli* demonstrates conformational plasticity. *J Bacteriol* **197**, 1873-1885 (2015).
125. Pramanik, A., Hauf, W., Hoffmann, J., Cernescu, M., Brutschy, B. & Braun, V. Oligomeric structure of ExbB and ExbB-ExbD isolated from *Escherichia coli* as revealed by LILBID mass spectrometry. *Biochemistry* **50**, 8950-8956 (2011).
126. Ferguson, A.D., Hofmann, E., Coulton, J.W., Diederichs, K. & Welte, W. Siderophore-mediated iron transport: Crystal structure of FhuA with bound lipopolysaccharide. *Science* **282**, 2215-2220 (1998).
127. Locher, K.P., Rees, B., Koebnik, R., Mitschler, A., Moulinier, L., Rosenbusch, J.P. & Moras, D. Transmembrane signaling across the ligand-gated FhuA receptor: Crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell* **95**, 771-778 (1998).
128. Buchanan, S.K., Smith, B.S., Venkatramani, L., Xia, D., Esser, L., Palnitkar, M., Chakraborty, R., van der Helm, D. & Deisenhofer, J. Crystal structure of the outer membrane active transporter FepA from *Escherichia coli*. *Nat Struct Biol* **6**, 56-63 (1999).
129. Chimento, D.P., Kadner, R.J. & Wiener, M.C. Comparative structural analysis of TonB-dependent outer membrane transporters: Implications for the transport cycle. *Proteins* **59**, 240-251 (2005).
130. Sharma, O., Yamashita, E., Zhalnina, M.V., Zakharov, S.D., Datsenko, K.A., Wanner, B.L. & Cramer, W.A. Structure of the complex of the colicin E2 R-domain and its BtuB receptor - The outer membrane colicin translocon. *J Biol Chem* **282**, 23163-23170 (2007).

131. Kurisu, G., Zakharov, S.D., Zhalnina, M.V., Bano, S., Eroukova, V.Y., Rokitskaya, T.I., Antonenko, Y.N., Wiener, M.C. & Cramer, W.A. The structure of BtuB with bound colicin E3 R-domain implies a translocon. *Nat Struct Biol* **10**, 948-954 (2003).
132. Reynolds, P.R., Mottur, G.P. & Bradbeer, C. Transport of vitamin B_{12} in *Escherichia coli* - some observations on the roles of the gene-products of BtuC and TonB. *J Biol Chem* **255**, 4313-4319 (1980).
133. Wookey, P. The TonB gene-product in *Escherichia coli* - energy-coupling or molecular processing of permeases. *Febs Lett* **139**, 145-154 (1982).
134. Letain, T.E. & Postle, K. TonB protein appears to transduce energy by shuttling between the cytoplasmic membrane and the outer membrane in *Escherichia coli*. *Mol Microbiol* **24**, 271-283 (1997).
135. Cascales, E., Lloubes, R. & Sturgis, J.N. The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB. *Mol Microbiol* **42**, 795-807 (2001).
136. Brockwell, D.J., Paci, E., Zinober, R.C., Beddard, G.S., Olmsted, P.D., Smith, D.A., Perham, R.N. & Radford, S.E. Pulling geometry defines the mechanical resistance of a beta-sheet protein (vol 10, pg 731, 2003). *Nat Struct Biol* **10**, 872-872 (2003).
137. Carrion-Vazquez, M., Li, H.B., Lu, H., Marszalek, P.E., Oberhauser, A.F. & Fernandez, J.M. The mechanical stability of ubiquitin is linkage dependent. *Nat Struct Biol* **10**, 738-743 (2003).
138. Higgs, P.I., Letain, T.E., Merriam, K.K., Burke, N.S., Park, H., Kang, G.C. & Postle, K. TonB interacts with nonreceptor proteins in the outer membrane of *Escherichia coli*. *J Bacteriol* **184**, 1640-1648 (2002).
139. Larsen, R.A., Thomas, M.G. & Postle, K. Protonmotive force, ExbB and ligand-bound FepA drive conformational changes in TonB. *Mol Microbiol* **31**, 1809-1824 (1999).
140. Larsen, R.A., Letain, T.E. & Postle, K. In vivo evidence of TonB shuttling between the cytoplasmic and outer membrane in *Escherichia coli*. *Mol Microbiol* **49**, 211-218 (2003).
141. Kaserer, W.A., Jiang, X.X., Xiao, Q.B., Scott, D.C., Bauler, M., Copeland, D., Newton, S.M.C. & Klebba, P.E. Insight from TonB hybrid proteins into the mechanism of iron transport through the outer membrane. *J Bacteriol* **190**, 4001-4016 (2008).
142. Gresock, M.G., Savenkova, M.I., Larsen, R.A., Ollis, A.A. & Postle, K. Death of the TonB shuttle hypothesis. *Front Microbiol* **2**(2011).
143. Jordan, L.D., Zhou, Y.Y., Smallwood, C.R., Lill, Y., Ritchie, K., Yip, W.T., Newton, S.M. & Klebba, P.E. Energy-dependent motion of TonB in the Gram-negative bacterial inner membrane. *P Natl Acad Sci USA* **110**, 11553-11558 (2013).
144. Klebba, P.E. ROSET model of TonB action in Gram-Negative bacterial iron acquisition. *J Bacteriol* **198**, 1013-21 (2016).
145. Gumbart, J., Wiener, M.C. & Tajkhorshid, E. Mechanics of force propagation in TonB-dependent outer membrane transport. *Biophys J* **93**, 496-504 (2007).
146. Whitfield, C.D., Steers, E.J., Jr. & Weissbach, H. Purification and properties of 5-methyltetrahydropteroyltriglutamate-homocysteine transmethylase. *J Biol Chem* **245**, 390-401 (1970).

147. Locher, K.P. Mechanistic diversity in ATP-binding cassette (ABC) transporters. *Nat Struct Mol Biol* **23**, 487-493 (2016).
148. Gruber, K., Puffer, B. & Krautler, B. Vitamin B_{12}-derivatives-enzyme cofactors and ligands of proteins and nucleic acids. *Chem Soc Rev* **40**, 4346-63 (2011).
149. Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J.M. & Gaub, H.E. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276**, 1109-12 (1997).
150. Chen, Y., Radford, S.E. & Brockwell, D.J. Force-induced remodelling of proteins and their complexes. *Curr Opin Struc Biol* **30**, 89-99 (2015).
151. Crampton, N. & Brockwell, D.J. Unravelling the design principles for single protein mechanical strength. *Curr Opin Struc Biol* **20**, 508-517 (2010).
152. Lee, C.K., Wang, Y.M., Huang, L.S. & Lin, S.M. Atomic force microscopy: Determination of unbinding force, off rate and energy barrier for protein-ligand interaction. *Micron* **38**, 446-461 (2007).
153. Riener, C.K., Stroh, C.M., Ebner, A., Klampfl, C., Gall, A.A., Romanin, C., Lyubchenko, Y.L., Hinterdorfer, P. & Gruber, H.J. Simple test system for single molecule recognition force microscopy. *Anal Chim Acta* **479**, 59-75 (2003).
154. Marko, J.F. & Siggia, E.D. Stretching DNA. *Macromolecules* **28**, 8759-8770 (1995).
155. Bustamante, C., Marko, J.F., Siggia, E.D. & Smith, S. Entropic elasticity of lambda-phage DNA. *Science* **265**, 1599-1600 (1994).
156. Carrion-Vazquez, M., Oberhauser, A.F., Fowler, S.B., Marszalek, P.E., Broedel, S.E., Clarke, J. & Fernandez, J.M. Mechanical and chemical unfolding of a single protein: A comparison. *P Natl Acad Sci USA* **96**, 3694-3699 (1999).
157. Marszalek, P.E., Lu, H., Li, H.B., Carrion-Vazquez, M., Oberhauser, A.F., Schulten, K. & Fernandez, J.M. Mechanical unfolding intermediates in titin modules. *Nature* **402**, 100-103 (1999).
158. Farrance, O.E., Hann, E., Kaminska, R., Housden, N.G., Derrington, S.R., Kleanthous, C., Radford, S.E. & Brockwell, D.J. A force-activated trip switch triggers rapid dissociation of a colicin from its immunity protein. *Plos Biol* **11**(2013).
159. Farrance, O.E., Paci, E., Radford, S.E. & Brockwell, D.J. Extraction of accurate biomolecular parameters from single-molecule force spectroscopy experiments. *Acs Nano* **9**, 1315-1324 (2015).
160. Bell, G.I. Theoretical-models for the specific adhesion of cells to cells or to surfaces. *Adv Appl Probab* **12**, 566-567 (1980).
161. Evans, E. Probing the relation between force - lifetime - and chemistry in single molecular bonds. *Annu Rev Bioph Biom* **30**, 105-128 (2001).
162. Evans, E. & Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys J* **72**, 1541-1555 (1997).
163. Bippes, C.A. & Muller, D.J. High-resolution atomic force microscopy and spectroscopy of native membrane proteins. *Rep Prog Phys* **74**(2011).
164. Rakshit, S. & Sivasankar, S. Biomechanics of cell adhesion: how force regulates the lifetime of adhesive bonds at the single molecule level. *Phys Chem Chem Phys* **16**, 2211-2223 (2014).
165. Marshall, B.T., Long, M., Piper, J.W., Yago, T., McEver, R.P. & Zhu, C. Direct observation of catch bonds involving cell-adhesion molecules. *Nature* **423**, 190-193 (2003).

166. Rief, M., Pascual, J., Saraste, M. & Gaub, H.E. Single molecule force spectroscopy of spectrin repeats: Low unfolding forces in helix bundles. *J Mol Biol* **286**, 553-561 (1999).
167. Dietz, H. & Rief, M. Exploring the energy landscape of GFP by single-molecule mechanical experiments. *P Natl Acad Sci USA* **101**, 16192-16197 (2004).
168. Brockwell, D.J., Beddard, G.S., Paci, E., West, D.K., Olmsted, P.D., Smith, D.A. & Radford, S.E. Mechanically unfolding the small, topologically simple protein L. *Biophys J* **89**, 506-519 (2005).
169. Hoffmann, T., Tych, K.M., Hughes, M.L., Brockwell, D.J. & Dougan, L. Towards design principles for determining the mechanical stability of proteins. *Phys Chem Chem Phys* **15**, 15767-15780 (2013).
170. Anishkin, A., Loukin, S.H., Teng, J.F. & Kung, C. Feeling the hidden mechanical forces in lipid bilayer is an original sense. *P Natl Acad Sci USA* **111**, 7898-7905 (2014).
171. Knecht, Z.A., Gaudet, R. & Garrity, P.A. The touching tail of a mechanotransduction channel. *Cell* **162**, 1214-1216 (2015).
172. Wyatt, P.J. Light-scattering and the absolute characterization of macromolecules. *Anal Chim Acta* **272**, 1-40 (1993).
173. Seidel, S.A., Dijkman, P.M., Lea, W.A., van den Bogaart, G., Jerabek-Willemsen, M., Lazic, A., Joseph, J.S., Srinivasan, P., Baaske, P., Simeonov, A., Katritch, I., Melo, F.A., Ladbury, J.E., Schreiber, G., Watts, A., Braun, D. & Duhr, S. Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. *Methods* **59**, 301-15 (2013).
174. Putnam, C.D., Hammel, M., Hura, G.L. & Tainer, J.A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* **40**, 191-285 (2007).
175. Tria, G., Mertens, H.D.T., Kachala, M. & Svergun, D.I. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *Iucrj* **2**, 207-217 (2015).
176. Kalverda, A.P., Thompson, G.S., Vogel, A., Schroder, M., Bowie, A.G., Khan, A.R. & Homans, S.W. Poxvirus K7 protein adopts a Bcl-2 fold: biochemical mapping of its interactions with human DEAD box RNA helicase DDX3. *J Mol Biol* **385**, 843-853 (2009).
177. Hutter, J.L. & Bechhoefer, J. Calibration of atomic-force microscope tips. *Rev Sci Instrum* **64**, 1868-1873 (1993).
178. Thoma, J., Bosshart, P., Pfreundschuh, M. & Muller, D.J. Out but not in: the large transmembrane beta-barrel protein FhuA unfolds but cannot refold via beta-hairpins. *Structure* **20**, 2185-2190 (2012).
179. Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L. & Schulten, K. Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781-802 (2005).
180. MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D. & Karplus, M. All-atom empirical potential for

molecular modeling and dynamics studies of proteins. *J Phys Chem B* **102**, 3586-616 (1998).
181. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-8, 27-8 (1996).
182. Gumbart, J., Wiener, M.C. & Tajkhorshid, E. Coupling of calcium and substrate binding through loop alignment in the outer-membrane transporter BtuB. *J Mol Biol* **393**, 1129-1142 (2009).
183. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926-935 (1983).
184. Ryckaert, J.P., Ciccotti, G. & Berendsen, H.J.C. Numerical-integration of cartesian equations of motion of a system with constraints - Molecular-dynamics of N-alkanes. *J Comput Phys* **23**, 327-341 (1977).
185. Tuckerman, M., Berne, B.J. & Martyna, G.J. Reversible multiple time scale molecular-dynamics. *J Chem Phys* **97**, 1990-2001 (1992).
186. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H. & Pedersen, L.G. A smooth particle mesh Ewald method. *J Chem Phys* **103**, 8577-8593 (1995).
187. Grest, G.S. & Kremer, K. Molecular-dynamics simulation for polymers in the presence of a heat bath. *Phys Rev A* **33**, 3628-3631 (1986).
188. Feller, S.E., Zhang, Y.H., Pastor, R.W. & Brooks, B.R. Constant-pressure molecular-dynamics simulation - the langevin piston method. *J Chem Phys* **103**, 4613-4621 (1995).
189. Guixa-Gonzalez, R., Rodriguez-Espigares, I., Ramirez-Anguita, J.M., Carrio-Gaspar, P., Martinez-Seara, H., Giorgino, T. & Selent, J. MEMBPLUGIN: studying membrane complexity in VMD. *Bioinformatics* **30**, 1478-1480 (2014).
190. Sumner, J.P. & Kopelman, R. Alexa Fluor 488 as an iron sensing molecule and its application in PEBBLE nanosensors. *Analyst* **130**, 528-533 (2005).
191. Schwesinger, F., Ros, R., Strunz, T., Anselmetti, D., Guntherodt, H.J., Honegger, A., Jermutus, L., Tiefenauer, L. & Pluckthun, A. Unbinding forces of single antibody-antigen complexes correlate with their thermal dissociation rates. *P Natl Acad Sci USA* **97**, 9972-9977 (2000).
192. Wallis, R., Moore, G.R., James, R. & Kleanthous, C. Protein-protein interactions in colicin E9 DNase-immunity protein complexes. 1. Diffusion-controlled association and femtomolar binding for the cognate complex. *Biochemistry* **34**, 13743-50 (1995).
193. Li, W., Keeble, A.H., Giffard, C., James, R., Moore, G.R. & Kleanthous, C. Highly discriminating protein-protein interaction specificities in the context of a conserved binding energy hotspot. *J Mol Biol* **337**, 743-59 (2004).
194. Lee, F.S., Shapiro, R. & Vallee, B.L. Tight-binding inhibition of angiogenin and ribonuclease A by placental ribonuclease inhibitor. *Biochemistry* **28**, 225-30 (1989).
195. Schreiber, G. & Fersht, A.R. Interaction of barnase with its polypeptide inhibitor barstar studied by protein engineering. *Biochemistry* **32**, 5145-50 (1993).
196. Koedding, J., Howard, P., Kaufmann, L., Polzer, P., Lustig, A. & Welte, W. Dimerization of TonB is not essential for its binding to the outer membrane siderophore receptor FhuA of *Escherichia coli*. *J Biol Chem* **279**, 9978-9986 (2004).

197. Dougan, L., Li, J.Y., Badilla, C.L., Berne, B.J. & Fernandez, J.M. Single homopolypeptide chains collapse into mechanically rigid conformations. *P Natl Acad Sci USA* **106**, 12605-12610 (2009).
198. Sochacki, K.A., Shkel, I.A., Record, M.T. & Weisshaar, J.C. Protein diffusion in the periplasm of *E. coli* under osmotic stress. *Biophys J* **100**, 22-31 (2011).
199. Ghosh, J. & Postle, K. Disulphide trapping of an in vivo energy-dependent conformation of *Escherichia coli* TonB protein. *Mol Microbiol* **55**, 276-288 (2005).
200. Hoffmann, T. & Dougan, L. Single molecule force spectroscopy using polyproteins. *Chem Soc Rev* **41**, 4781-4796 (2012).
201. Michaud-Agrawal, N., Denning, E.J., Woolf, T.B. & Beckstein, O. Software news and updates MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* **32**, 2319-2327 (2011).
202. Ferguson, A.D., Amezcua, C.A., Halabi, N.M., Chelliah, Y., Rosen, M.K., Ranganathan, R. & Deisenhofer, J. Signal transduction pathway of TonB-dependent transporters. *P Natl Acad Sci USA* **104**, 513-518 (2007).
203. Ferguson, A.D. & Deisenhofer, J. TonB-dependent receptors - structural perspectives. *Bba-Biomembranes* **1565**, 318-332 (2002).
204. Krieg, S., Huche, F., Diederichs, K., Izadi-Pruneyre, N., Lecroisey, A., Wandersman, C., Delepelaire, P. & Welte, W. Heme uptake across the outer membrane as revealed by crystal structures of the receptor-hemophore complex. *P Natl Acad Sci USA* **106**, 1045-1050 (2009).
205. Cobessi, D., Celia, H., Folschweiller, N., Schalk, I.J., Abdallah, M.A. & Pattus, F. The crystal structure of the pyoverdine outer membrane receptor FpvA from *Pseudomonas aeruginosa* at 3.6 angstrom resolution. *J Mol Biol* **347**, 121-134 (2005).
206. Devanathan, S. & Postle, K. Studies on colicin B translocation: FepA is gated by TonB. *Mol Microbiol* **65**, 441-453 (2007).
207. Endriss, F., Braun, M., Killmann, H. & Braun, V. Mutant analysis of the *Escherichia coli* FhuA protein reveals sites of FhuA activity. *J Bacteriol* **185**, 4683-4692 (2003).
208. Eisenhauer, H.A., Shames, S., Pawelek, P.D. & Coulton, J.W. Siderophore transport through *Escherichia coli* outer membrane receptor FhuA with disulfide-tethered cork and barrel domains. *J Biol Chem* **280**, 30574-30580 (2005).
209. Jimenez, R.H.F. & Cafiso, D.S. The N-terminal domain of a TonB-dependent transporter undergoes a reversible stepwise denaturation. *Biochemistry* **51**, 3642-3650 (2012).
210. Udho, E., Jakes, K.S. & Finkelstein, A. TonB-dependent transporter FhuA in planar lipid bilayers: partial exit of its plug from the barrel. *Biochemistry* **51**, 6753-6759 (2012).
211. Toscano, W.A., Jr. & Storm, D.R. Bacitracin. *Pharmacol Ther* **16**, 199-210 (1982).
212. Cornelissen, F. & Van den Bossche, H. Synergism of the antimicrobial agents miconazole, bacitracin and polymyxin B. *Chemotherapy* **29**, 419-27 (1983).
213. Braun, M., Killmann, H. & Braun, V. The beta-barrel domain of FhuA Delta 5-160 is sufficient for TonB-dependent FhuA activities of *Escherichia coli*. *Mol Microbiol* **33**, 1037-1049 (1999).
214. Langenhan, T., Piao, X. & Monk, K.R. Adhesion G protein-coupled receptors in nervous system development and disease. *Nat Rev Neurosci* (2016).

215. Seo, D., Southard, K.M., Kim, J.W., Lee, H.J., Farlow, J., Lee, J.U., Litt, D.B., Haas, T., Alivisatos, A.P., Cheon, J., Gartner, Z.J. & Jun, Y.W. A mechanogenetic toolkit for interrogating cell signaling in space and time. *Cell* **165**, 1507-18 (2016).
216. Lin, X. Perception of sound and gravity by TMC1 and TMC2. *J Clin Invest* **121**, 4633-4636 (2011).
217. Brohawn, S.G., Su, Z. & MacKinnon, R. Mechanosensitivity is mediated directly by the lipid membrane in TRAAK and TREK1 K+ channels. *P Natl Acad Sci USA* **111**, 3614-9 (2014).
218. Zhang, W., Cheng, L.E., Kittelmann, M., Li, J., Petkovic, M., Cheng, T., Jin, P., Guo, Z., Gopfert, M.C., Jan, L.Y. & Jan, Y.N. Ankyrin repeats convey force to gate the NOMPC mechanotransduction channel. *Cell* **162**, 1391-403 (2015).
219. Anishkin, A., Loukin, S.H., Teng, J. & Kung, C. Feeling the hidden mechanical forces in lipid bilayer is an original sense. *P Natl Acad Sci USA* **111**, 7898-905 (2014).
220. Housden, N.G., Hopper, J.T., Lukoyanova, N., Rodriguez-Larrea, D., Wojdyla, J.A., Klein, A., Kaminska, R., Bayley, H., Saibil, H.R., Robinson, C.V. & Kleanthous, C. Intrinsically disordered protein threads through the bacterial outer-membrane porin OmpF. *Science* **340**, 1570-4 (2013).
221. Gerding, M.A., Ogata, Y., Pecora, N.D., Niki, H. & de Boer, P.A.J. 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**, 1008-1025 (2007).
222. Mullineaux, C.W., Nenninger, A., Ray, N. & Robinson, C. Diffusion of green fluorescent protein in three cell environments in *Escherichia coli*. *J Bacteriol* **188**, 3442-3448 (2006).
223. Wojdyla, J.A., Cutts, E., Kaminska, R., Papadakos, G., Hopper, J.T.S., Stansfeld, P.J., Staunton, D., Robinson, C.V. & Kleanthous, C. Structure and function of the *Escherichia coli* Tol-Pal stator protein TolR. *J Biol Chem* **290**, 26675-26687 (2015).
224. Kaimer, C., Berleman, J.E. & Zusman, D.R. Chemosensory signaling controls motility and subcellular polarity in *Myxococcus xanthus*. *Curr Opin Microbiol* **15**, 751-757 (2012).
225. Grinter, R., Josts, I., Zeth, K., Roszak, A.W., McCaughey, L.C., Cogdell, R.J., Milner, J.J., Kelly, S.M., Byron, O. & Walker, D. Structure of the atypical bacteriocin pectocin M2 implies a novel mechanism of protein uptake. *Mol Microbiol* **93**, 234-46 (2014).
226. Sverzhinsky, A., Chung, J.W., Deme, J.C., Fabre, L., Levey, K.T., Plesa, M., Carter, D.M., Lypaczewski, P. & Coulton, J.W. Membrane protein complex ExbB4-ExbD1-TonB1 from *Escherichia coli* demonstrates conformational plasticity. *J Bacteriol* **197**, 1873-85 (2015).
227. Schiffrin, B., Calabrese, A. N., Devine, P. W., Harris, S. A., Ashcroft, A. E., Brockwell, D. J., & Radford, S. E. Skp is a multivalent chaperone of outer-membrane proteins. *Nat Struct Mol Biol* **23**, 786-793 (2016).