This electronic thesis or dissertation has been downloaded from the King's Research Portal at https://kclpure.kcl.ac.uk/portal/

Examining the role of structural dynamics in the assembly and function of the multidrug efflux pump AcrAB-TolC

Russell Lewis, Ben

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT

Unless another licence is stated on the immediately following page this work is licensed

under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

licence. https://creativecommons.org/licenses/by-nc-nd/4.0/

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 09. May. 2024

Examining the role of structural dynamics in the assembly and function of the multidrug efflux pump AcrAB-TolC

Benjamin Russell Lewis

Department of Chemistry, King's College London

This thesis is submitted for the degree of Doctor of Philosophy

2023

Abstract

Membrane efflux pumps are a leading cause of increasing multidrug resistant bacterial infections, which pose a significant threat to global human health. Understanding the mechanisms that underpin their function is critical for the development of inhibitors targeting these systems, with the aim being to 'revive' the activities of pre-existing antibiotics known to suffer bacterial resistance. AcrAB-TolC is a membrane spanning, tripartite multidrug efflux pump native to *Escherichia coli* and prototypical of homologous systems across other ESKAPE bacteria. The work in this thesis investigates the role of structural dynamics in the function, assembly, and inhibition of AcrAB-TolC, with a focus on the membrane fusion protein (MFP) AcrA, to reveal critical information on how these efflux systems work, which could be essential for developing novel methods of inhibition to combat multidrug resistance. Throughout this work, structural mass spectrometry (MS) techniques such as hydrogen deuterium exchange mass spectrometry (HDX-MS) and native MS were used alongside a range of complementary biophysical/biochemical techniques to investigate AcrAB-TolC.

This work reveals that AcrA lipidation promotes the propensity of AcrA to form oligomers, whereas a non-lipidated, water soluble AcrA construct ($AcrA^{S}$) is still monomeric. Moreover, HDX-MS showed $AcrA^{S}$ exhibits increased backbone structural dynamics at pH 6.0 compared to pH 7.4, yet this was largely tempered by the presence of magnesium. In the periplasm, the pH can often be ~1.7 pH units lower than in the cytosol, and there is a significantly higher concentration of magnesium ions (7.56 times). This suggests a regulatory role of magnesium to help AcrA function within the periplasmic environment. To expand the investigations on $AcrA^{S}$ further, a soluble pseudo-dimer construct ($AcrA^{SD}$) was used to infer biological information on the AcrA functional dimer. It was found the pseudo-dimer has unique structural dynamics compared to AcrA, with extensive protection in the α-helices and in regions of the αβ-barrel and MP domains. Furthermore, whilst $AcrA^{S}$ and $AcrA^{SD}$ appeared to bind peptidoglycan similarly, $AcrA^{SD}$ had a higher propensity to form higher order complexes with AcrB. This suggests dimerization may help prime the AcrA protomers for interactions with its binding partners.

Traditionally, efflux pumps inhibitors (EPIs) have been targeted against AcrB, but none have made it past clinical trials, often due to toxicity issues. This has led to a switch in focus for the next generation of EPIs, with AcrA becoming a promising target. In this work, HDX-MS and native MS were used in combination with molecular dynamics (MD) simulations, to investigate the effect of a recently identified EPI, NSC 60339, on the structural dynamics of $AcrA^{S}$. The data showed NSC 60339 likely

binds to AcrA in a cleft bridging the lipoyl and αβ-barrel domains, stabilising these areas as well as the MP domain which usually exhibits intrinsic disorder; NSC 60339 inhibition of $AcrA^{SD}$ presented the same. This work proposes the first mechanism of action regarding an AcrA inhibitor and reveals a promising new way to target the AcrAB-TolC complex.

Due to the hydrophobic nature of membrane proteins, a suitable membrane mimetic is required for *in vitro* investigations. As the AcrAB-TolC multidrug efflux pump spans the entire Gram-negative cell envelope and are therefore membrane proteins, studying them in lipid environments rather than detergents is essential as they provide a more representative environment. In this work, HDX-MS was used to show MBX-3756 stabilises the hydrophobic trap of AcrB in membrane scaffold protein (MSP) nanodiscs. Furthermore, a novel SMALP-liposome-SMALP assay was utilised to show that previously designed AcrB antimicrobial peptides did not make the AcrB trimer, purified in styrene maleic acid lipid particles (SMALPs), dissociate into monomers. Lastly, assembly of the AcrAB-TolC complex was probed using two different pull-down assays and native polyacrylamide gel electrophoresis (PAGE), however the heterogeneity and hydrophobicity of SMALPs complicated these investigations, combined with the slow energetics of this assembly *in vitro*.

Acknowledgements

There are many people to thank that have helped me throughout this work and to create this thesis. Firstly, I would like to thank my supervisor Dr Eamonn Reading, who gave me the opportunity to complete a PhD within his laboratory, and to become the first student within his group. He has been a fantastic mentor, both personally and professionally, and has taught me invaluable lessons that will stay with me throughout my career. I have thoroughly enjoyed working together throughout my PhD.

Secondly, I would like to thank the entire Reading group, past and present for the help and support over the last four years. In particular, I would like to thank Dr Dietmar Hammerschmid, who trained me on the mass spectrometry instruments throughout the early years of my PhD, with exceptional patience and dedication. I thank him for all the time he put in to help me, and it has been key to my scientific development. I would also like to thank Ryan Lawrence for the laughs over the last two years and having to suffer my fear of flying!

I would like to thank my second supervisor Professor Paula Booth, who allowed me to join the Booth laboratory for the last few months of my PhD. Furthermore, I would like to thank the entire Booth group past and present, who I shared an office with, for the support during my PhD. A particular thanks goes to Dr Heather Findlay, who was always available with any questions I ever had.

I would like to thank the external collaborators I have worked with throughout my PhD: Dr Muhammad Uddin, Dr Mohammad Moniruzzaman, Professor J.C Gumbart, Katie Kuo, Professor Frank Sobott, Dr Anna Higgins, Dr Anna Olerinyova and Dr Weston Struwe. A special thank you goes to Professor Helen Zgurskaya who has provided me with invaluable knowledge and ideas on AcrAB-TolC that has sparked fascinating investigations.

I would also like to thank my family and friends for their continued support throughout my PhD. A particular thank you goes to my parents who always support me in whatever I do, and I am eternally grateful. Finally, I would like to thank my partner Emma Deans, for her unwavering support throughout this journey.

Publications

- Conformational restriction shapes the inhibition of a multidrug efflux adaptor protein. B Russell Lewis, M R Uddin, M Moniruzzaman, K M Kuo, A J Higgins, L M N Shah, F Sobott, J M Parks, D Hammerschmid, J C Gumbart, H I Zgurskaya, E Reading. *Nature Communications*, (2023) 14, 3900. DOI: 10.1038/s41467-023-39615-x

  *Evening Standard article about this work: https://www.standard.co.uk/news/health/kings-college-scientists-tool-antibiotics-superbugs-b1095412.html#comments-area

- Structural mass spectrometry approaches to understand multidrug efflux systems. B Russell Lewis, R Lawrence, D Hammerschmid, E Reading. *Essays In Biochemistry*, (2023) 67, 255–267. DOI: 10.1042/EBC20220190

- Dietmar Hammerschmid, Valeria Calvaresi, Chloe Bailey, Benjamin Russell Lewis et al. Chromatographic Phospholipid Trapping for Automated H/D Exchange Mass Spectrometry of Membrane Protein–Lipid Assemblies. *Analytical Chemistry* **2023** *95* (5), 3002-3011 DOI: 10.1021/acs.analchem.2c04876

Contents

List of Figures

List of Tables

List of equations

List of abbreviations

ABC	ATP binding cassette
A. baumanni	*Acinetobacter baumanni*
AcrA^{L}	Lipidated AcrA
AcrA^{S}	Soluble AcrA
AcrA^{SD}	Soluble AcrA pseudo-dimer
ADH	Alcohol dehydrogenase
AMR	Antimicrobial resistance
ASM	Auxiliary solvent manager
BAM	Barrel assembly machinery
BSA	Bovine serum albumin
BSM	Binary solvent manager
CD	Circular dichroism
CI	Confidence intervals
CID	Collison induced dissociation
CL	Cardiolipin
CMC	Critical micelle concentrations
CRM	Charge residual model
Cryo-EM	Cryogenic electron microscopy
Cryo-ET	Cryogenic electron tomography
CSD	Charge state distributions
CV	Column volumes
DC	Direct current
DDM	*n*-Dodecyl-β-D-maltopyranoside
DIBMA	Di-isobutylene-alt-maleic acid
DLS	Dynamic light scattering
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol

E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EPI	Efflux pump inhibitor
EPR	Electron paramagnetic resonance
ESI	Electrospray ionisation
ESKAPE	Multidrug resistant Gram-negative bacteria
fD	Fraction denatured
Fos-choline	n-Dodecyl-phosphocholine
FT	Fourier transform
FTICR	Fourier transform ion cyclotron resonance
FWHM	Full width half maximum
GlcNAc	*N*-acetylglucosamine
GuHCl	Guanidine hydrochloride
HAE-1	Hydrophobe/amphiphile efflux-1
HDX-MS	Hydrogen deuterium exchange mass spectrometry
HME	Heavy metal efflux
HTH	Helix-turn-helix
IEM	Ion evaporation model
IM	Ion mobility
IPTG	Isopropyl ß- d-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
LB	Luria-Bertani broth
LC	Liquid chromatography
LILBID	Laser-induced liquid bead ion desorption
LPS	Lipopolysaccharide
MALDI	Matrix-assisted laser desorption ionisation
MATE	Multidrug and toxin extrusion
MaxD	Maximally deuterated control
MBX	Pyranopyridine inhibitors
MCP	Microchannel plate

MD	Molecular dynamics
MFP	Membrane fusion protein
MFS	Major facilitator superfamily
MP	Membrane proximal
MRSA	Methicillin resistant *Staphylococcus aureus*
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSP	Membrane scaffold protein
MTS-R6G	MTS-rhodamine 6G
MurNAc	*N*-acetylmuramic acid
nESI	nano electrospray ionisation
NMR	Nuclear magnetic resonance
NSC 60339	2-chloro-4',4"-bis(2-imidazolin-2-yl)terephthalanilide
OMF	Outer membrane factor
OMP	Outer membrane protein
PAGE	Polyacrylamide gel electrophoresis
PAβN	Phenyl-arginine-β-naphthylamide
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PE	Phosphatidyl ethanolamine
PG	Phosphatidyl glycerol
pI	Isoelectric point
PLE	Polar lipid extract (*Escherichia coli*)
PLGS	ProteinLynx Global Sever
PMSF	Phenylmethylsulfonyl fluoride
POPC	1-palmitoyl-2-oleoyl-glycero-3-phosphocholine
PS	Phosphatidyl serine
P. aeruginosa	*Pseudomonas aeruginosa*
Q-ToF	Quadrupole-Time of Flight
RF	Radio frequency

RFU	Relative fractional uptake
RMSD	Root mean-square deviations
RMSF	Root mean-square fluctuations
RND	Resistance nodulation cell division
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
RU	Response units
S. aureus	*Staphylococcus aureus*
SASA	Solvent accessible surface area
SD	Standard deviation
SDS	Sodium dodecyl-sulphate
SEC	Size exclusion chromatography
SEC MALS	Multi-angle laser light scattering size exclusion chromatography
SMALP	Styrene maleic acid lipid particle
SMA-PAGE	Styrene maleic acid - polyacrylamide gel electrophoresis
SMILP	Polystyrene-co-maleimide
SMR	Small multidrug resistance
SPR	Surface plasmon resonance
SRP	Signal recognition pathway
TCEP	Tris(2-carboxyethyl)phosphine
TM	Transmembrane
Tm	Melting temperature
TMP-SMX	Trimethoprim sulfamethoxazole
ToF	Time of Flight
UHMR	Ultra-high mass resolution
UPLC	Ultra-high pressure liquid chromatography
β-ME	β-mercaptoethanol
ΔHDX	Differential hydrogen deuterium exchange mass spectrometry

Thesis layout

Chapter 1 is split into two parts, providing a detailed introduction into the structural mass spectrometry (MS) techniques used in this thesis, and the background biology regarding the AcrAB-TolC multidrug efflux pump. The first part of **chapter 1** details the core principles of MS that underpins its application to biological systems. Then, detailed backgrounds are provided on hydrogen deuterium exchange mass spectrometry (HDX-MS) and native MS, which are used extensively throughout this thesis. This covers the theory, practical considerations, suitability to the study of membrane proteins and data analysis. Finally, part 1 provides a short introduction to mass photometry.

Part 2 of **chapter 1** begins by introducing biological membranes and the structure of the Gram-negative bacterial cell envelope. There is a short introduction into the biogenesis of inner and outer membrane proteins. Part 2 then introduces bacterial drug resistance, and the several mechanisms from which resistance can arise. Multidrug efflux pumps are introduced, and their role in multidrug resistance, and more specifically the varying roles of the resistance nodulation and cell division (RND) superfamily of efflux pumps. There is a detailed introduction into the AcrAB-TolC multidrug efflux pump and the reasons for studying this complex are highlighted, providing a rationale to this work. The structural features of the inner membrane protein AcrB, membrane fusion protein (MFP) AcrA, and outer membrane factor (OMF) protein TolC are described, as well as the structure of the assembled complex. The rotational mechanism of efflux and the assembly of the complex are detailed, whilst current gaps in the literature are acknowledged throughout.

Each experimental chapter contains its own experimental methods section describing the methods used in that particular chapter.

Chapter 2 presents, discusses, and concludes the first set of experimental results. In this chapter, AcrA is expressed and purified from *Escherichia coli* in both a lipidated ($AcrA^{L}$) and soluble ($AcrA^{S}$) form. The effect of pH and the lipidation on the oligomeric state of AcrA is shown using native MS. Furthermore, native MS and HDX-MS are used to help characterise $AcrA^{S}$ as a folded protein with defined secondary structure with areas of intrinsic disorder. This chapter aims to investigate the role of the periplasmic environment on AcrA, which often exhibits more acidic pH's than the cytosol and contains over 7 times higher magnesium concentrations.[1–3] HDX-MS investigations show AcrA exhibits increased dynamics at pH 6.0 compared to pH 7.4, but that Mg^{2+} appears to temper the increased dynamics at pH 6.0 whilst having little observed effect at pH 7.4. This suggests a regulatory role of Mg^{2+} in the function of AcrA.

Chapter 3 presents, discusses, and concludes the second set of experimental results. In this chapter, a soluble pseudo-dimer construct of AcrA ($AcrA^{SD}$) is utilised to gain biological insights into the AcrA functional dimer. The work reveals pseudo-dimerization leads to extensive stabilisation across the α-helical domains, as well as stabilisation in portions of the αβ-barrel and MP domains. Furthermore, native polyacrylamide gel electrophoresis (PAGE) and mass photometry are used to characterise the binding of $AcrA^{SD}$/$AcrA^{S}$ to AcrB in styrene maleic acid lipid particle (SMALP) nanodiscs to reveal differences in the stoichiometries of the complexes, with $AcrA^{SD}$ promoting higher order binding to AcrB. However, when comparing the affinity to peptidoglycan using a peptidoglycan pull-down assay, it reveals there is no difference between the two constructs suggesting dimerization is not necessary for this interaction.

Chapter 4 presents, discusses, and concludes the third set of experimental results. This chapter utilises structural MS techniques combined with molecular dynamics (MD) simulations and cellular accumulation assays to propose the first mechanism of inhibition study for an AcrA inhibitor. This work is performed in collaboration with the Zgurskaya group (University of Oklahoma) and the Gumbart group (Georgia Institute of Technology). This chapter reveals the previously identified NSC 60339 efflux pump inhibitor (EPI) binds as a molecular wedge in a cleft on AcrA between the lipoyl and αβ-barrel domains, to reduce the structural dynamics of AcrA across all four domains, impacting its function in the assembled complex.

Chapter 5 presents, discusses, and concludes the final set of experimental results. This chapter investigates the AcrAB-TolC multidrug efflux pump purified in different membrane mimetic environments, with different levels of complexity compared to native membranes, to gain biological information. This chapter shows that a novel online delipidation HDX-MS workflow can be used to study an EPI (MBX-3756) binding AcrB in membrane scaffold protein (MSP) nanodiscs, without the knowledge of any binding parameters.[4] Furthermore, the work in this chapter uses a novel SMALP-liposome-SMALP assay to monitor the effect of antimicrobial peptides on the oligomeric state of AcrB in lipid environments. Lastly, work in this chapter details efforts to assemble the AcrAB-TolC complex *in vitro*, using SMA-PAGE and pull-down assays.

Chapter 6 contains the concluding remarks and future outlooks.

Chapter 1: Introduction

Part 1: An introduction to structural mass spectrometry techniques and mass photometry

1.1 Introduction

Mass spectrometry (MS) has been developed throughout the last 100 years to become a key tool in structural biology.[5] The first biomolecules were analysed at the end of the 1950's, yet it wasn't until softer ionisation techniques such as electrospray ionisation (ESI), developed by John Fenn, that the field was truly born.[6,7] The ability to analyse intact non-covalent interactions in the gas phase, and being able to maintain protein complexes in a mass spectrometer, led to the publication of literature that shed light on various protein systems previously beyond investigation.[8–10] Now, 30 years after the initial reports, MS can provide structural information on protein complexes in the megadalton mass range such as viruses and ribosomal complexes.[9,11–13] MS has had success in a range of biological areas, including elucidating subunit stoichiometry, monitoring molecular dynamics interactions and substrate binding, defining crystallographic interactions and assigning unknown density to lipids.[10,14] Today, MS is a key technique for investigating biological systems, especially when complemented with other biophysical methods such as x-ray crystallography, nuclear magnetic resonance spectroscopy (NMR) and single particle cryo-electron microscopy (cryo-EM).[15–17]

MS can investigate increasingly complex species, which reflects the increase in MS applications regarding membrane proteins. Membrane proteins pose a greater problem to study than water-soluble proteins due to their hydrophobicity, lower yields of functional and stable protein, and the need for a membrane mimetic environment.[18,19] Characterising membrane proteins is critical in structural biology as they make up 30% of the open reading frames in the genome, and count for ~60% of therapeutic drug targets.[20,21] Investigating membrane proteins by MS techniques is challenging due to their insolubility in aqueous buffers, the fact their subunits readily dissociate and the use of anionic surfactants can often lead to signal suppression.[18,22–26] It is important to note that membrane proteins do not just pose an issue for MS techniques, but the structural biology field in general. These challenges are exemplified by the fact that there are only <100 solved structures of membrane proteins from human cells, yet there are ~8000 known; this is very poor compared to the ratio for soluble proteins (as of 2022).[27,28] However, it is important to note that recent advances in cryo-EM and the emergence of AlphaFold are helping to rectify some of the issues associated with obtaining structural information of membrane proteins.[29–31]

Regardless of the problems posed by membrane proteins, MS techniques such as hydrogen-deuterium exchange MS (HDX-MS) and native MS are increasingly used to study these systems. The work presented in this thesis focuses on the AcrAB-TolC multidrug efflux pump (see **chapter 1, part 2**), and MS techniques have been used extensively to study various multidrug efflux pumps to elucidate critical information. Please read our review on this topic for a more in-depth discussion on the MS

approaches to studying multidrug efflux systems.[32] A few select examples are mentioned in this section.

MS techniques are deployed to reveal critical biological information regarding the function, assembly, and inhibition of multidrug efflux pumps (**Figure 1.1**). HDX-MS is an increasingly popular technique to monitor protein dynamics over time.[33] It can be applied to efflux pumps to help decipher their function. Javad *et al.* (2022) utilised HDX-MS to probe conformational states of the bacterial ATP binding cassette (ABC) transporter BmrA during the ATPase cycle.[34] They revealed the key steps of the catalytic cycle and highlighted the importance of an ADP bound inward-facing conformation of BmrA during the post-hydrolytic step. Reading *et al.* (2022) used both HDX-MS and native MS to understand AcrB inhibition by phenyl-arginine-β-naphthylamide (PAβN). The HDX-MS data supported an inhibitory mode of action whereby PAβN restricts AcrB structural dynamics in the drug binding pockets and the switch loop.[35]

Aside from HDX-MS, native MS can reveal useful information on protein-ligand interactions. Lyu *et al.* (2022) used native MS to determine the binding affinities of a range of lipids to the ABC transporter, MsbA[36]. They revealed MsbA copurifies with copper and has an enhanced affinity to lipids when MsbA was fully loaded with copper. Bolla *et al*. (2020) used native MS to study the proteobacterial antimicrobial compound efflux (PACE) protein AceI, and its binding to nucleic acids, lipids and drugs under a range of conditions.[37] Crosslinking MS is another technique that can be used to investigate efflux pumps. As mentioned in **section 1.10.8,** Shi *et al.* (2019) used *in vivo* crosslinking of AcrA and TolC to peptidoglycan, and analysed the interactions by liquid chromatography (LC)-MS/MS.[38]

Overall, the application of MS based structural biology techniques is widely applied to the study of efflux proteins, and yields a host of critical information, essential for the understanding and ultimately inhibition of these systems. This chapter will introduce the core principles of MS and detail the fundamentals of HDX-MS and native MS, and their applications for studying membrane proteins.

.

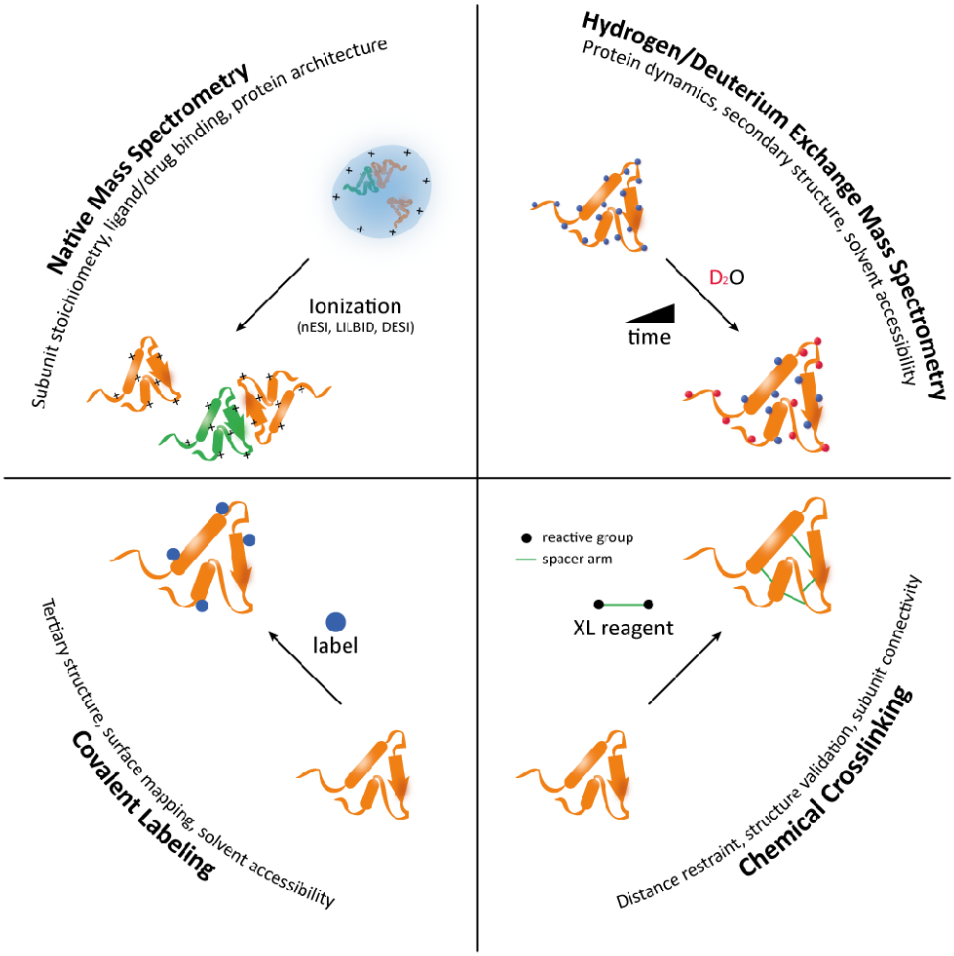


Figure 1.1. Structural MS toolbox. Abbreviations: ESI, electrospray ionisation; nESI, nano-electrospray ionisation; LILBID, laser-induced liquid bead ion desorption; DESI, desorption electrospray ionisation; MS, mass spectrometry. From Russell Lewis et al. (2023).[32]

1.2 General principles of MS

1.2.1 Ion sources

1.2.1.1 Electrospray ionisation

ESI is a soft ionisation technique, and together with matrix-assisted laser desorption ionisation (MALDI) paved the way for MS to study biomolecules.[39] Before these techniques were developed, by John Fenn (ESI), Michael Karas and Franz Hillenkamp (MALDI), MS was mostly limited to organic compounds.[7,40] The introduction of soft ionisation techniques revolutionised the field, and led to John Fenn and Koichi Tanaka sharing the 2002 Nobel Prize in Chemistry for their contributions to MS. This section will mainly discuss ESI, but alternatives will also be introduced.[41]

In ESI a sample in solution is injected through a hypodermic needle or a stainless steel capillary (with an inner diameter of ~0.1 mm) at a low flow rate (1-40 µL/min), and a very high voltage (2.5-3.0 kV) is applied to the tip of the needle relative to the orifice of the mass spectrometer which is typically 1-3 cm away.[42–44] ESI can produce both positive and negative ions depending on the sign of the applied electrical field, and usually for proteins the positive ion mode is utilised.[7,42,45] The applied voltage provides the electric field gradient required for the charge separation at the liquid surface. These accumulated like charges are destabilised and the meniscus is deformed into a structure called the Taylor cone, which is formed at the capillary outlet.[46] The Taylor cone ejects a fine jet of liquid from its apex at high field strength and a sheath gas (e.g. N_2) flow around the capillary helps atomize the solution and direct it towards the heated capillary.[47,48] The charges on the droplets are equidistantly spaced to minimise potential energy.[49] Solvent evaporation occurs as the droplets travel towards the heated capillary, and hence the droplet size decreases until it reaches the Rayleigh Limit (**Equation 1.1**), where the surface tension can no longer sustain the charge repulsion and 'Coulomb fission' occurs.[42,50,51] A drying gas sweeps away the solvent vapour, along with any uncharged material.[52] This results in the disintegration of the parent droplet into smaller offspring droplets that carry around 2% of the mass but 15% of the charge, giving them a much higher charge to mass ratio.[53] The process of solvent evaporation and Coulomb fission repeats to form smaller and smaller progeny droplets, and from these smaller droplets the gas phase ions are formed.

$$Z_R.e = 8\pi(\varepsilon_\theta \gamma R^3)^{\frac{1}{2}}$$

Equation 1.1. The Rayleigh equation. *Z* is the charge number, *e* is the elementary charge, ε_θ is the permittivity of the surrounding medium, γ is the surface tension of the droplet and *R* is the radius of the droplet.

The charged, desorbed analyte ions released from the droplet, pass through a sampling cone or the orifice of a heated capillary and are transported into the first vacuum stage of the mass spectrometer through two different mechanisms.[54] A schematic of both models is shown in **Figure 1.2**. The first is the charge residual model (CRM) and it can be applied to large globular proteins.[54–56] The CRM states that after serial events of ion evaporation and Coulomb fission, a nanodroplet (~1 nm) is formed which contains only one analyte molecule. Desolvation of this droplet causes its surface charge to be transferred to the analyte molecule. This model was originally developed by Dole *et al*. (1968) and it was expanded upon by Schmelzeisen-Recker *et al*. (1989).[55,56] Assuming the nanodroplet only contains one neutral analyte molecule and is only marginally bigger than the protein itself, and that the density of the globular protein is the same as for water, **Equation 1.2** can be created by combining this with the Rayleigh equation.[42,57] **Equation 1.2** shows the observed agreement between the charges on the proteins and the water droplets of approximately identical size, at the Rayleigh limit, is a consequence of multiply charged proteins being formed by the CRM.

$$\left(\frac{4}{3}\pi R^3\right) p N_A = M$$

$$Z_R = 4\left(\frac{\pi\gamma\varepsilon_\theta}{pe^2 N_A}\right)^{\frac{1}{2}} \times M^{\frac{1}{2}}$$

Equation 1.2. The charge residual model. *Z* is the charge number, *e* is the elementary charge, ε_θ is the permittivity of the surrounding medium, γ is the surface tension of the droplet, *R* is the radius of the droplet, N_A is Avogadro's constant, *p* is the density of water and *M* is the molecular mass of the protein.

The second model is the ion evaporation model (IEM) and this applies to low molecular weight species.[54,58] The IEM posits that after repeated ion evaporation and Coulomb fission, the size of the droplet decreases to a point where the electric field is strong enough to directly emit solvated ions; when the droplet reaches ~10 nm, ion emission dominates over Rayleigh fission. Thus, in the IEM a nanodroplet containing a single molecule does not have to be produced. **Equation 1.3** shows the rate constant for ion emission.[42,58] Activation is influenced by three factors: the attraction between

escaping ion and the solvents composing the droplet, Coulombic repulsion of the escaping ion and the remaining surface charges on the droplet, and the ion desolvation energy.

$$K_1 = \frac{K_b T}{h} \exp\left(-\frac{\Delta G^+}{RT}\right)$$

Equation 1.3. The ion evaporation model. $K_b$is the Boltzmann constant, h is Planck's constant, T is the temperature of the droplet and ΔG^+ is the activation free energy.

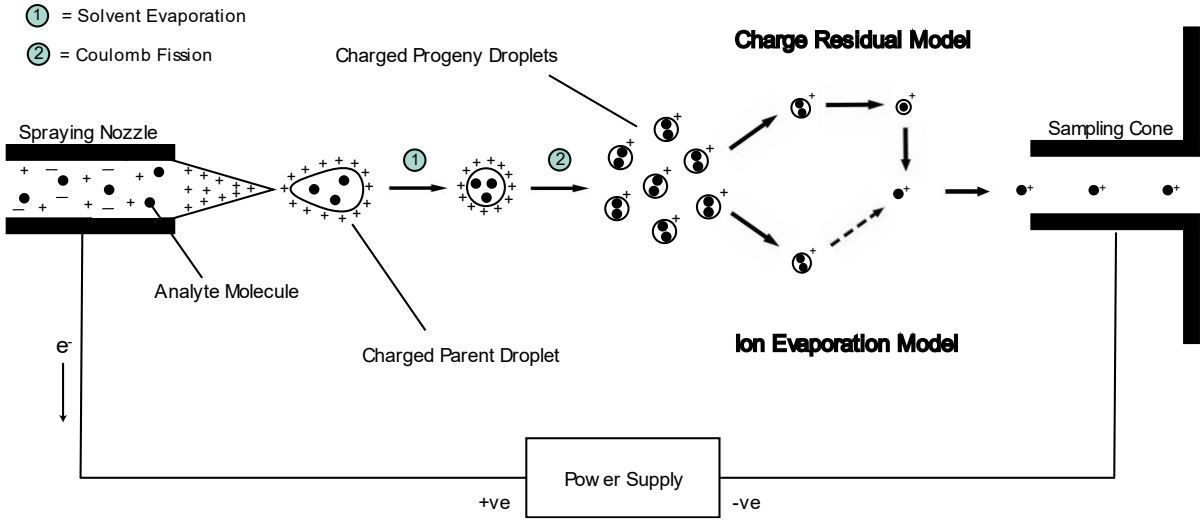


Figure 1.2. Simple schematic of electrospray ionisation. The voltage applied to the spraying nozzle causes charge separation at the liquid surface, causing the destabilization of like charges and the formation of the Taylor cone at the end. This ejects a fine jet of liquid made up of parent droplets. As these travel towards the sample cone, solvent evaporation decreases the size of the droplets until the Rayleigh limit is reached and Coulomb fission occurs. The parent droplet disintegrates into much smaller progeny droplets. The top route shows the Charge Residual Model (CRM), where a nanodroplet forms with one analyte molecule in, and desolvation causes the transfer of charge to the analyte. The lower route represents the Ion Evaporation Model (IEM), which does not require a single analyte molecule in a nanodroplet. When the electric field is strong enough the droplets directly emit analyte ions.

1.2.1.2 Alternatives to ESI

An alternative ionisation technique is nanoESI (nESI), the miniaturized version of ESI, and it is often used for native MS (see **section 1.3**). It was developed by Wilm and Mann, and it most notably differs from ESI in the size of the initial droplets formed, at ~180 nm compared to the µm range for ESI.[47,59] This is achieved by using a smaller needle orifice with only a 1 µm inner diameter, and no external pumping is required. nESI has two major advantages over conventional ESI. Firstly, higher signal intensities are observed in nESI, which serves a particular advantage when analysing glycoproteins as oligosaccharides give a strong intensity.[60] Secondly nESI exhibits a higher tolerance to salt contaminations in buffers, which is particularly advantageous when studying biological samples.[59] More information on the advantages of nESI for native MS can be found in **section 1.3.3**.

Another alternative soft ionisation technique is laser-induced liquid bead ion desorption (LILBID), developed by Morgner *et al*. (2006).[61] Here, biological samples in buffered solution are laser desorbed from microdroplets into vacuum.[62] For desorption to occur the infrared laser is tuned to the absorption maximum of water (~3 µm) and at a threshold intensity of around 100 MW/cm^2 a very fast phase transition is induced. Accompanying this is subsequent disruption of the droplet and the emission of ions from liquid into the gas phase. Typically, 100-200 droplets are needed to get a mass spectrum, but this can vary with sample concentration. Depending on the laser intensity applied, LILBID has shown potential in different applications.[41] At low laser intensity LILBID is a gentle technique enabling the characterisation of intact protein complexes. Laser intensity can be increased, and this is useful to break hydrophobic interactions between subunits to analyse them as individuals, or to release membrane proteins from their mimetic environments, as demonstrated on AcrB in styrene maleic acid lipid particle (SMALP) nanodiscs.[63]

1.2.2 Ion transmission, filtering and focusing

After the production of gaseous ions, to analyse their *m/z* ratios, they must be focused and guided along their intended flight path and separated in a mass analyser, before detection.[64] Effective ion transmission is essential in order to detect ions.[65] Ions are guided through the instrument by a voltage gradient, and ion transmission is optimised using a combination of voltages and pressures; this is dependent on ion mass.[64,66]

ESI occurs at atmospheric pressure, and ions enter the first vacuum of the system through an aperture whilst being carried by the bath gas.[52] A high proportion of ions deviate from their ideal trajectory during the transition to vacuum, due to the free jet expansion of the gas stream. Ions of all sizes, including large protein assemblies, will be moving at a high velocity. Collisional focusing in a radio

frequency (RF) ion guide is traditionally used to focus the ions moving along wild trajectories at high kinetic energies.[64] Collisions between comparatively large ions and neutral gas molecules dissipate ion kinetic energy into the surrounds to allow for more efficient ion focusing. Douglas and French expanded upon this to give the RF quadrupole ion guide.[66] Increased pressure in the RF quadrupole ion guide leads to increased collisions with the bath gas, so the axial and radial velocity of the ions exhibit 'collisional dampening', allowing ions to be captured more efficiently in the dynamic potential well at the centre axis of the ion guide.[64,66,67]

Some mass spectrometers used to measure large protein complexes, and the majority of instruments used throughout this work, use an orthogonal time of flight (ToF) mass analyser, where the ion beam is injected orthogonal to the axis of the ToF analyser.[64,68] Ions in the orthogonal ToF mass analyser are separated by their axial velocity as they enter this region, and if it is in excess they can overshoot the detector.[69] Thus the pressure has to be at the correct level; if too low large ions may not be slowed down enough but too high and they may not have enough kinetic energy to reach the detector at all. Mass spectrometers have been modified for their required needs, so ion guides, ion optics, and pressure stages are all tailored for the ion transmission of both small and large ions. **Figure 1.3** displays the general layout of the instrumentation used in this thesis.

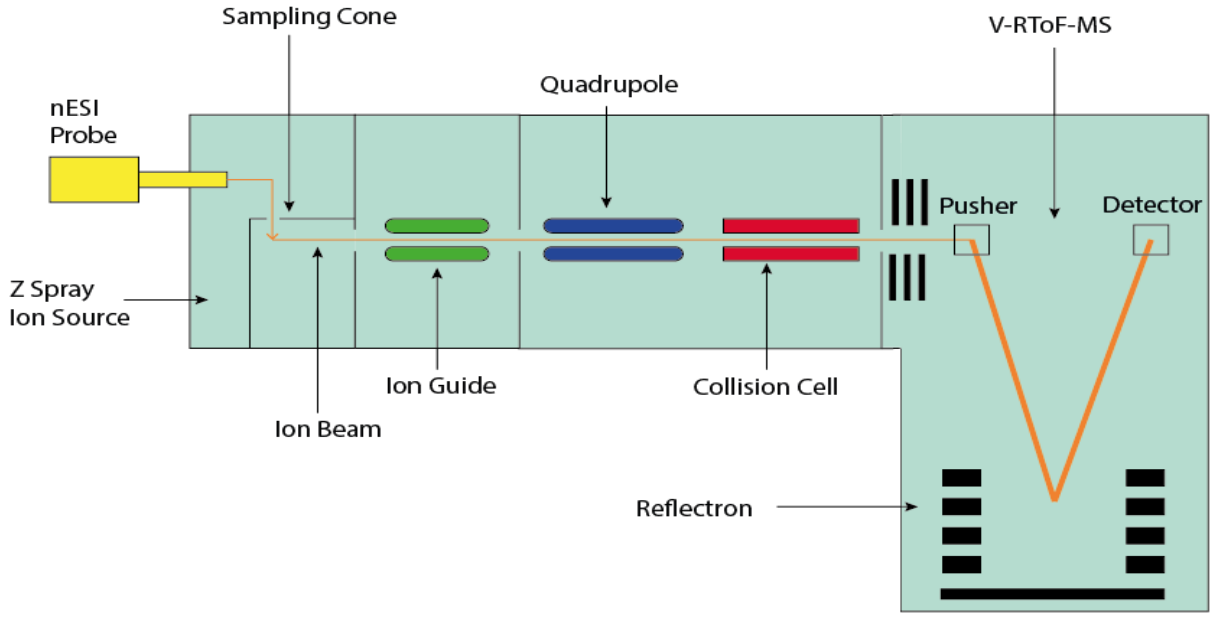


Figure 1.3. A general schematic of a nESI-Q-ToF instrument. The orange line represents the ion beam moving from left to right throughout the instrument. Analyte ions are produced in the ion source and drawn into the first vacuum stage of the instrument. The ion beam traverses the ion guide (green), RF quadrupole (blue), and the collision cell (red) designed for ion focusing, ion selection, and fragmentation respectively, before reaching the ToF analyser and hitting the detector.

1.2.3 Ion activation

Tandem MS (MS/MS) is a general method involving at least two stages of mass analysis.[70] MS/MS usually requires the fragmentation of precursor ions selected by the first mass analyser in order to allow the second mass analyser to analyse the product ions. Activation of ions in MS occurs in the collision cell, and is crucial for structure elucidation, with collision induced dissociation (CID) the most common method.[71] Precursor ions are injected into the collision cell pressurized with inert gas molecules (e.g Argon, Helium or Nitrogen).[42,72,73] As a result of the collisions, internal energy is built up gradually and redistributed among different vibrational degrees of freedom.[74] **Equation 1.4** shows the fraction of translational energy that is converted to internal energy.[70] Consequently, an unstable excited precursor ion decomposes into product ions. The fragmentation pathway depends on the amount of energy deposited, not the activation method used.[70]

$$E_{cm} = E_{lab} \frac{M_t}{M_i + M_t}$$

Equation 1.4. Energy fraction of translational energy to internal energy. M_i is the ion mass, M_t is the target gas mass, E_{lab} is the ion kinetic energy in a lab frame of reference and E_{cm} is the maximum energy fraction converted to internal energy.

Ion activation provides useful information on protein sequence and subunit composition.[75] Gas phase activation of protein complexes involves multiple steps: cleaning, restructuring, unfolding, dissociating and fragmenting. During cleaning, gentle activation of ions leads to the removal of residual solvent and/or buffer salt present.[74,76] As internal energy increases protein complexes begin to restructure leading to a compromised complex conformation. Following this, the protein starts to unfold before a subunit dissociates. Unfolding increases the accessible area of the complex so charge migration to this area occurs, and most complexes will eject a single subunit with disproportionate charge.[77] Finally when internal energy is high, fragmentation of protein chain occurs as well as dissociation. See **section 1.4.3.3.1** for a MS/MS spectrum displaying peptide fragmentation data. CID has also proven useful in the characterisation of membrane protein complexes. Membrane proteins need to have their hydrophobic environment (e.g. detergent micelle, SMALP) removed in order to study them using MS.[78,79] This is achieved by stripping the hydrophobic environment in a collision cell under high voltage, releasing the membrane protein.

1.2.4 Mass analysers

1.2.4.1 Quadrupole mass analyser

The principle of quadrupoles was first described by Wolfgang Paul (Nobel prize in 1989) in 1953.[80] A quadrupole mass analyser is composed of four parallel rods with a circular cross section (**Figure 1.4**).[81] A direct current (DC, termed *U*) is applied to two of the rods, and the other two are linked to an alternating RF potential (potential termed *v*, frequency termed *ω*). Ions are pulsed towards a quadrupole by an electric field, and positively charged ions will head towards negatively charged rods. However, once the polarity is changed, the ions movement path will change before it hits the rod. They undergo complex oscillation and with the correct values of *U*, *V* and *ω*, only ions within a narrow range of *m/z* will survive the path towards the detector, as these have 'stable' trajectories. The remaining ions possess unstable trajectories and will eventually collide with one of the electrodes. The advantages are the low cost, small size, and robustness of the quadrupole, yet it is restricted by its limited mass detection range and low resolving power.

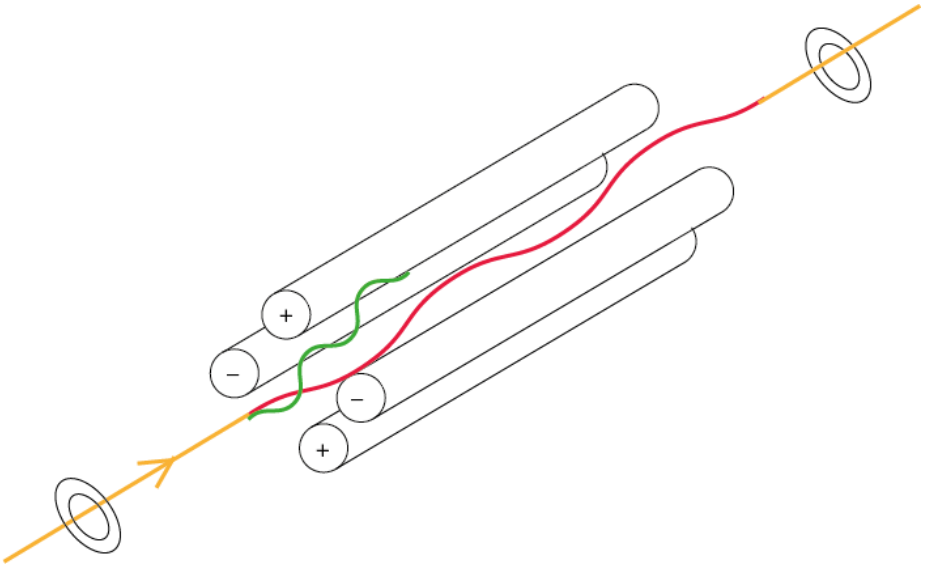


Figure 1.4. A quadrupole mass analyser. Four parallel rods with varying DC or RF potentials. The two positively charged and two negatively charged rods are marked. The ion beam from the ion source is shown as the yellow line. Ions with stable trajectories follow the red path to the detector, whereas ions with unstable trajectories follow the green path and hit the rod.

This simplified idea of the quadrupole mechanism is based on the Mathieu equation (**Equation 1.6**).[82] Ions travelling along the Z axis experience a total electric field, consisting of a quadrupolar alternative field superimposed on a constant field, resulting from the alternating DC and RF potentials applied to the rods (**Equation 1.5**).[70]

$$\phi_0 = +(U - VCos\omega t)$$

$$-\phi_0 = -(U - VCos\omega t)$$

Equation 1.5. Alternating DC and RF potentials. ϕ_0 represents the potential applied to the rods, ω represents the angular frequency (in radians per second = $2\pi v$ where v is the frequency of the RF field), U is the direct potential and V is the amplitude of the RF voltage.

The Z axial accelerated ions enter the space between the quadrupole rods, maintaining their velocity. They are subjected to accelerations along the X and Y axes (which determine the position of the ion from the centre of the rods) from the forces induced by the electric fields. The trajectory of the ion will be stable if the values of X or Y do not reach r_0 (the inner diameter distance between the rods) as the ions will never hit the rods.[83] Mathieu established an equation describing wave propagation in membranes in 1868.[70] This equation can be changed and rearranged to reflect a second order differential known as the Mathieu equations (**Equation 1.6**).

$$\frac{d^2u}{d\xi^2} + (a_u - zq_u cos2\xi)u = 0$$

$$a_u = a_x = -a_y = \frac{8zeU}{m\omega^2 {r_0}^2}$$

$$q_u = q_x = -q_y = \frac{4zeV}{m\omega^2 {r_0}^2}$$

Equation 1.6. The Mathieu equations. The first equation is the 1866 equation describing wave propagation in membranes. The second and third equations show the changed and rearranged equations that establish a relationship between the coordinates of an ion and time. u represents either X or Y, r_0 is the inner diameter distance between the rods, ξ is time length ($\frac{\omega t}{2}$).

Equations 1.6 establish a relationship between the coordinates of an ion and time. If X and Y both remain less than r_0 the ion can pass through the quadrupole without colliding with the rods, so it is not discharged, and it reaches the detector.

Quadrupole mass analysers in the RF scanning mode are used to select and filter analyte ions, and it is achieved by keeping the RF frequency constant whilst altering the direct potential and RF amplitude.[84] This enables ions of a particular *m/z* to be selected, with a stability diagram (**Figure 1.5**) showing the ratio of *U* and *V* where that particular ion is stable.[85] Where ions fall within the triangle in the stability plots, they will be detected. **Figure 1.5** shows a trajectory diagram where the upper and lower surfaces represent the rod surfaces. As velocity in the Z axis is constant, the horizontal axis represents the length of rods and time taken for the ion to traverse the filter. These stability plots are mass dependant, so a constant *U/V* ratio scan can allow ions of different masses but similar *m/z* ratios to be detected, expanding the mass limit.

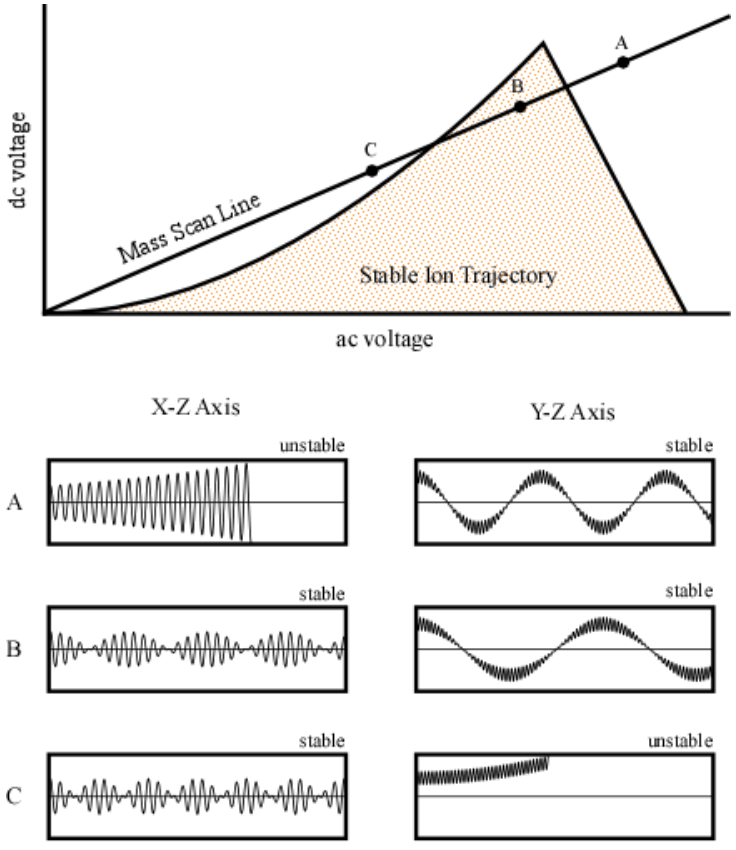


Figure 1.5. UV stability diagram and trajectory plots. The ion at points A and C have unstable trajectories in the X or Y dimension, respectively. Figure from Henchman and Steel (1998).[85]

A quadrupole can also be in the RF only mode to act as an ion guide.[70,84,86,87] When *U*=0, all of the ions with a mass higher than a given limit, selected by adjusting the value of the RF voltage (*V*), have a stable trajectory. Ions within a transmission mass range are systematically brought back to the centre of the rods, even in when they are deflected by collisions or close to colliding with one of the rods due to their initial trajectories. Ions below this mass range follow an unstable trajectory, whilst ions above the mass range are also lost due to poor focusing. The highest *m/z* that can be transmitted is given by **Equation 1.7**.[86]

$$M_{Max} = \frac{7 \times 10^6 V_m}{f^2 r_0^2}$$

Equation 1.7. Determining the highest *m/z* in an RF only quadrupole ion guide. V_m is the RF voltage, f is the frequency and r_0 is the inner diameter distance between the rods.

1.2.4.2 Time of Flight mass analysers

ToF mass analysers are used to detect an ions *m/z* ratio, by measuring an ions velocity through a field free drift region (*D*), called a flight tube.[70,88] The velocity at which the ions enter the drift region is determined by **Equation 1.8**.

$$V = (\frac{2zeEs}{m})^{\frac{1}{2}}$$

Equation 1.8. **Velocity of ions entering the drift region.** *Es* is the accelerating potential, *e* is the charge on an electron, *z* is the number of charges and *m* is the mass.

The time (*t*) taken to traverse the drift region is dependent on the mass, highlighted in **Equation 1.9**.

$$t = (\frac{m}{2zeEs})^{\frac{1}{2}} D$$

Equation 1.9. Time taken to traverse the drift region. *Es* is the accelerating potential, *e* is the charge on an electron, *z* is the number of charges, *m* is the mass, *t* is time and *D* is the drift region.

Equation 1.10 shows the conversion of the time spectrum to the mass spectrum.

$$\frac{M}{Z} = 2eEs(\frac{t}{D})^2$$

Equation 1.10. Calculating the mass spectrum. *Es* is the accelerating potential, *e* is the charge on an electron, *z* is the number of charges, *m* is the mass, *t* is time, *D* is the drift region and *m/z* is the mass to charge ratio.

To obtain the timing information it is pertinent to know the time of ion formation or extraction. Thus, most ToF analysers use pulsed ionisation, such as in MALDI. There are a variety of different ToF instrumentation that differ in terms of data recording. Linear ToF mass spectrometers are the simplest, as these use pulsed ionisation techniques and ions travel in a straight line through the flight tube towards the detector. However, ESI techniques are not pulses, as instead they form a continuous beam.[64] Therefore ESI-ToF mass spectrometers often use orthogonal acceleration ToF analysis, developed initially by O'Halloran *et al.* (1964).[70,89] In orthogonal ToF the continuous beam is deflected and injected orthogonally to the ToF mass analyser.[64,70] The orthogonal accelerator is filled with new ions from the ion source while the ions in the flight tube are being analysed simultaneously. New ions cannot be injected into the mass analyser until the current ions have hit the detector; the ion beam fills the orthogonal accelerator faster than it takes the ions in the flight tube to hit the detector, so part of the ion beam is not pushed into the ToF mass analyser and is lost. Typically, this instrumentation exhibits poor duty cycle, especially with large masses. This can be combatted with the use of ion traps preceding the orthogonal extraction of the primary beam, modulating the continuous beam into a source of ions pulsed at a frequency timed to match the extracting timing of the ToF analyser.[64,90]

One way to improve mass resolution is to use reflectron ToF instrumentation. Two ions formed in the same location and with the same mass may have different kinetic energies, and hence different velocities. Reflectrons can be used to compensate for the differences in kinetic energy.[88] They create a retarding field that acts like a mirror by deflecting the ions and sending them back through the flight tube. V-type reflectrons (**Figure 1.3**) are often used, and they have a linear potential applied to them. Reflectrons consist of a series of equally spaced grid electrodes connected through a resistive network of equal valve resistors. They correct the kinetic energy dispersion of the ions of equal *m/z*, due to ions with greater kinetic energy and hence velocity penetrating the reflectron more deeply. Thus, faster ions spend more time in the reflectron and reach the detector at the same time as the slower ions.

1.2.4.3 Orbitrap mass analysers

An orbitrap is a Fourier-Transform (FT) based mass analyser, which functions differently to quadrupoles.[91] Whilst quadrupoles measure *m/z* as a function of ion stability in an electromagnetic field, an FT based analyser measures *m/z* as a function of ion frequency during oscillatory motion. Orbitraps consist of an outer electrode that is shaped as a barrel and is cut into two equal parts with a small interval, and a central spindle shaped electrode (**Figure 1.6**).[70,92] An electrostatic voltage of several kilovolts is applied to the central electrode, whilst the outer electrode is at ground potential. Ions oscillate in the axial dimension around the central electrode due to electromagnetic forces experienced in the orbitrap due to the quadro-logarithmic potential distribution obtained by the DC voltage and geometry of the trap.[70,91] Ions are injected with an electric potential ramp that causes the ions to squeeze close to the central axial electrode; ions squeeze depending on their *m/z*, with lower *m/z* populations tighter to the central axis. This allows for populations of different *m/z* to oscillate around the central axis simultaneously without bumping into each other.

As **Figure 1.6** shows, the orbitrap is cut cross sectionally so that there is effectively a left and right electrode. As ions oscillate in the axial dimension, they move from one electrode to another, yet because they are charged, the electrodes sense the moving charge; this is the signal collected to determine ion oscillatory frequency, which depends on the *m/z*. During a scan, ions within a *m/z* range are injected into the orbitrap and the oscillatory motion for each population is measured simultaneously. This data is converted to the frequency domain using FT, which is then converted to *m/z* spectral data to produce a mass spectrum.

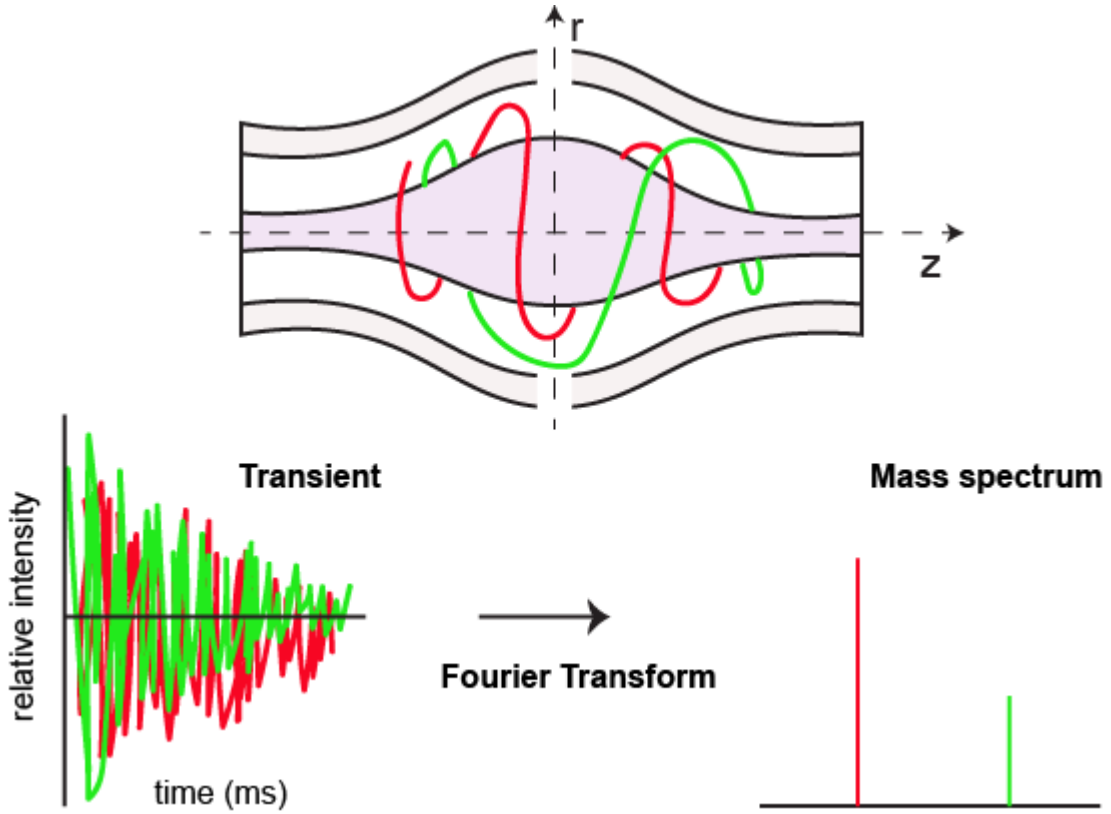


Figure 1.6. Orbitrap mass analyser. Ions in the orbitrap oscillate around the central electrode in the axial (Z) dimension and move back and forth between the left and right electrodes. This yields a transient (frequency of ion oscillation in the axial dimension), which is transformed to a mass spectrum via Fourier Transformation.

1.2.5 Detectors

There are three primary methods for detecting ions in MS, yet they all function by the generation of a current due to incident ions, which is measured, representing ion abundance.[93] The first method is called direct charge detection, an example being the Faraday cup. This type of detector has been important historically in MS, but mainly used in magnetic sector instruments today due to its relatively low sensitivity. The second method is called image charge detection, an example being the inductive detector. This technique is less sensitive than direct charge detection, but it is the only non-destructive detection modality. This is important for FT instruments such as Fourier transform ion cyclotron resonance (FTICR) and Orbitrap mass analysers. The final method is called secondary electron generation, including electron multiplier and microchannel plates (MCP). ToF instruments require large areas, rapid response times for good timing resolutions and corresponding accurate *m/z* determinations, and high sensitivity. Electron multipliers and MCP are best suited to this, as they can produce secondary electrons.[94] In this chapter MCP will be introduced as this was the detector used in the MS instruments throughout this thesis.

1.2.5.1 Microchannel plates

In an MCP, a conversion dynode is held at high potential, oppositely charged to the analyte ions.[70] A positive or negative ion striking the conversion dynode causes the emission of several secondary particles. When positive ions hit the negative conversion dynode the secondary particles of interest are negative ions and electrons. When negative ions hit the positive conversion dynode the secondary particles of interest are positive ions, which are converted to electrons at the first dynode. These secondary electrons are amplified in a cascade effect within the electron multipliers to form a detectable current.[95]

The MCP detector is an array of miniature electron multipliers (channels) orientated parallel to each other (**Figure 1.7**).[96] The channel diameter can range from 4-25 µm and the centre-centre distance ranging from 6-32 µm. There is a potential across the channel, with the input side being kept at a negative potential of ~1 kV compared to the output side. The channel matrix is made from lead glass, to optimize secondary emission characteristics of each individual channel, and to make the walls of the channels semi-conductive, allowing for the charge replenishment from an external voltage. Parallel electrical contact between the channels is provided by the deposition of a metallic coating such as nichrome, ensuring electron multiplication by giving off secondary electrons. The cascade effect within a channel can multiply the number of electrons by 10^5. A plate allows amplification between 10^2-10^4, but with several plates this can reach 10^8. A metal anode at the output side gathers

the stream of secondary electrons at every channel exit, so ions with different *m/z* ratios that reach different parts of the detector can be counted simultaneously. This increases sensitivity and efficiency, leading to lower detection limits.

The amplifying power is the product of the conversion factor (the number of secondary particles emitted per ion collision) and the multiplying factor of the continuous dynode electron multiplier. The conversion factor is reliant on the impact velocity of incoming ions, and their characteristics (mass, charge, structure). Therefore, the channels are characterised by the mass discrimination effect, as the slower velocity of larger ions (i.e. protein complexes) produce fewer secondary electrons, decreasing efficiency. The conversion dynodes at high voltages can reduce this effect by accelerating ions to high velocity to enhance efficiency, and therefore enhance signal intensity and sensitivity.

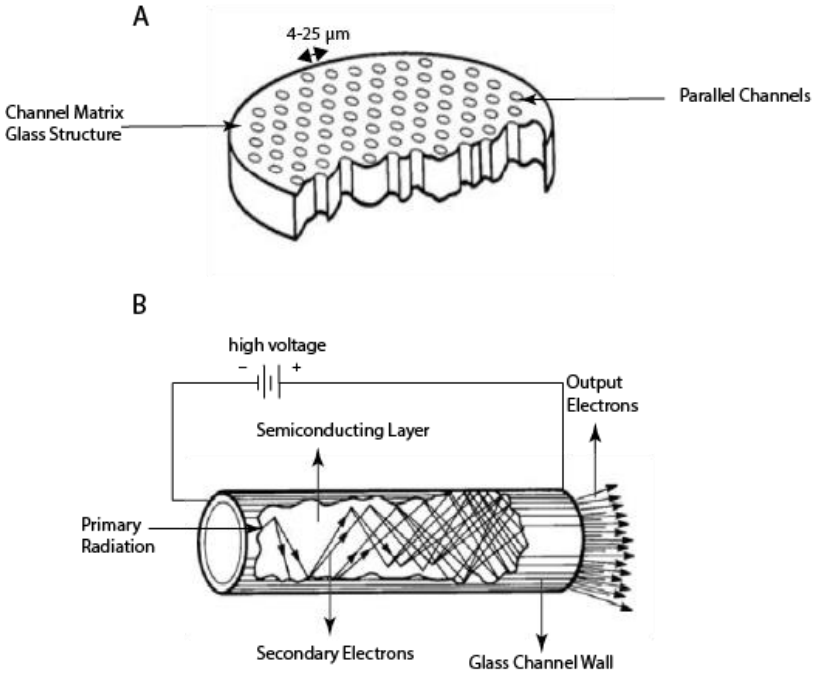


Figure 1.7. Schematic of a microchannel plate detector. A. Cross sectional view of an MCP plate. **B.** Representation of an MCP channel. The ion hits the dynode, causing the emission of secondary electrons, which are amplified in the cascade effect. This is repeated all over the MCP to generate a current, which is measured. Figure taken and adapted from Wiza (1979).[96]

1.3 Native Mass Spectrometry

1.3.1 Introduction

All cellular processes are determined by a range of proteins and their interactions with each other or other biomolecules. The development of soft ionisation techniques such as ESI allowed the visualisation of such interactions via MS.[7] This allows for the preservation of non-covalent interactions in the gas phase and thus protein bio-complexes are maintained.[8–10] The term 'native' in native MS corresponds to the status of the biological sample prior to ionisation, as the sample is in a non-denaturing buffer.[97] Since native MS is a gas phase method, this cannot be claimed to be a native state. Therefore, it is imperative that parameters such as pH and ionic strength are carefully controlled to best maintain the folded state of the protein in solution. Thus, by applying certain protocols and optimised workflows, it is possible to draw conclusions about the physiological state of the protein sample through native MS.

Native MS can be applied to elucidate key biological information. At its simplest, native MS can be used to calculate an accurate molecular weight of a protein sample and/or separate a mixture of proteins based on size to a much greater resolving power than other techniques such as native polyacrylamide gel electrophoresis (PAGE) and multi-angle laser light scattering size exclusion chromatography (SEC MALS).[97] Furthermore, native MS can be used to probe the biochemistry of protein complexes; it can be used to measure the mass of intact assemblies, probe the stoichiometry of assemblies, determine direct interactions between subunits, observe their relative position (core or periphery) and measure the relative strength of interactions between subunits.[10,97,98] Furthermore, it can be used to investigate protein-protein interactions and protein-ligand interactions, and it can be used to provide accurate determinations of binding parameters.[99] Native MS is also amenable to the study of membrane proteins, and can provide structural information on membrane protein complexes not easily achieved by other techniques.[28] Moreover, it can be used to study the binding of lipids to membrane proteins, which are essential to their function.[100] Native MS has been used in several studies to reveal critical information regarding multidrug efflux pumps – please refer to **section 1.1** for some examples.

1.3.2 Sample preparation

A high-quality protein sample is key to the success of native MS experiments. Usually this is achieved by overexpressing the protein(s) in expression systems such as bacteria, with the appropriate affinity tags, however samples can be obtained from more natural sources such as organ tissue.[98] It is important to ensure the protein sample is free of aggregation and purified to a concentration of 1-20

µM. However, during purification, substances which can cause signal suppression or extensive adduct formation during native MS measurements are often used.[101,102] This includes metal cations, inorganic ions, guanidinium salts, and common biological buffers such as HEPES, phosphate buffer saline (PBS), MES, MOPS and Tris. Furthermore, membrane protein samples can require the presence of detergents to ensure their solubility in solution, which can have detrimental effects to the observed spectra. Therefore, protein samples are buffer exchanged to volatile native MS buffers prior to experimentation. The most common is ammonium acetate at a pH between 6-8, however ammonium bicarbonate or ammonium formate can both be used as alternatives.[98] Ammonia and acetic acid are both volatile and readily evaporate in the ESI process.[103] Usually 100-200 mM ammonium acetate is used, and the high concentration helps in reducing the effect of non-volatile buffering components.[98,102] Commonly, samples are buffer exchanged using miniaturised gel filtration columns, molecular weight cut-off spin concentrators or dialysis devices. However it is important to note that ammonium acetate buffers are not a good reflection of physiological conditions.[104]

1.3.2.1 Optimisation of protein samples

Optimisation of protein samples for native MS is essential. Firstly, some experiments or protein complexes may require the presence of other buffer components, such as metal ions, cofactors or reducing agents.[102] The concentration of additional buffering components should be kept below 1 mM but be as low as possible. Higher concentrations of ammonium acetate can be used to offset the presence of additional components. Therefore, careful optimisation of the buffering composition is required. Secondly, certain proteins samples may aggregate throughout the buffer exchange process. This can be due to several reasons; the protein sample may be unhappy in ammonium acetate, the current pH may be unsuitable or it may be 'sticking' to the membrane of molecular weight cut-off concentrators.[98,105] Therefore, the type of volatile buffer and method of buffer exchange needs to be optimised per protein sample.

1.3.2.2 Optimisation of membrane protein samples

Native MS of membrane proteins pose an additional complication. Purified membrane proteins require the presence of a membrane mimetic, most commonly detergent micelles.[22] Therefore, protein-detergent samples must be buffer exchanged into a volatile buffer that contains 2x the critical micelle concentration (CMC) of the particular detergent used.[106] In native MS experiments, the membrane protein-micelle complex is ionised, which can cause heterogeneous mass spectra due to micelle adducts. Therefore, micelles need to be removed by ion activation inside the mass spectrometer (see **section 1.2.3**). It is essential to strike a balance between adding enough energy to

remove the micelle, but not too much energy as to compromise the integrity of the membrane protein. This can sometimes be inevitable if the intermolecular forces stabilizing the protein's secondary structure is weaker than the electrostatic forces stabilizing the micelle, or if there are large portions of the membrane protein residing outside the micelle, with no protection from collisional activation.[107,108] Therefore, selecting a suitable detergent for native MS with membrane proteins is imperative. They need to preserve the native fold of the membrane protein whilst also being optimised for producing high quality mass spectra.[106] Non-ionic detergents, characterised by uncharged hydrophilic head groups, are often best suited to native MS experiments as they are mild and less denaturing, yet cause little ion suppression and are generally easier to remove in the gas phase.[106,109] Early native MS studies utilised saccharide detergents such as *n*-Dodecyl-β-D-maltopyranoside (DDM) and *n*-octyl-β-D-glucopyranoside (OG).[106,109] The use of second-wave detergents was pioneered by Robinson *et al*. such as tetraethylene glycol monooctyl ether (C_8E_4) and lauryldimethylamine-*N*-oxide (LDAO); these detergents exhibit charge-reducing properties in the ionisation process and are therefore often removed from membrane protein samples at a lower collisional activation energy compared to non-charge-reducing detergents.[107,110] The structures of the aforementioned detergents are shown in **Figure 1.8a**, and these detergents continue to be commonly used in native MS experiments. In this work, DDM was the detergent used in the native MS spectra of membrane proteins presented in this thesis.

The use of detergents to study membrane proteins has disadvantages. Detergents strip away the native lipids found in the cell membrane that usually interact with the protein, which can have detrimental effects to its stability and function.[110,111] Therefore other mimetic environments have been developed to provide a more 'native' lipid environment. One option are membrane scaffold protein (MSP) nanodiscs, which are self-assembled proteolipid particles, containing MSP proteins encapsulating a well-defined mixture of phospholipids.[32,112,113] Another is the use of SMALP nanodiscs, which contain the membrane protein surrounded by its native lipids.[114] Whilst native MS protocols are readily amenable to proteins in MSP nanodiscs, traditional native MS methods are not suitable to SMALP nanodiscs.[106] Due to the heterogeneity of the lipids and polymer in the SMALP, the membrane protein needs to be ejected from the nanodisc. However, this is difficult with CID due to the overall stability of the nanodisc. Therefore, native MS protocols using LILBID ionisation has been applied to membrane proteins in SMALPs (see **section 1.2.1.2**).[63] For more information on native MS of membrane proteins, see this excellent review by Keener *et al.* (2021).[106]

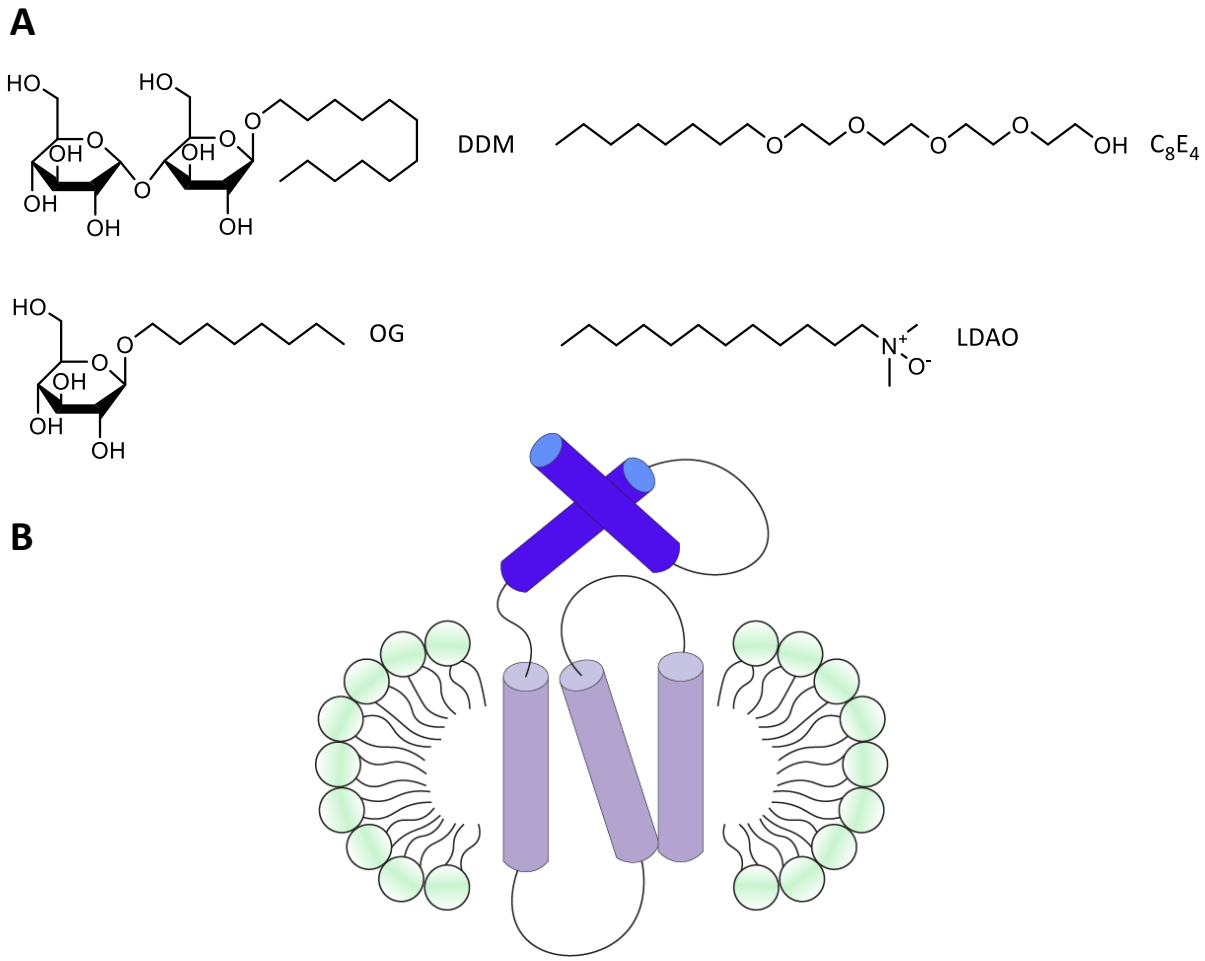


Figure 1.8. The detergent micelle. A. The chemical structures of four commonly used detergents in native MS (DDM, OG, C_8E_4, LDAO). **B.** Cartoon representation of a membrane protein in a detergent micelle. Purple α-helices are hydrophobic and reside in the micelle, whereas blue α-helices are hydrophilic and reside outside of the micelle.

1.3.3 Advantages of nanoelectrospray ionisation for native MS

The most commonly used ionisation method for native MS is nESI (see **section 1.2.1.2**).[59] nESI produces ions following the same steps as ESI, however nESI uses small borosilicate glass or quartz capillaries that have been pulled to produce a fine tip of ~1 µm, and given a metallic coating (usually gold) to hold an electric potential.[64] This differs from ESI, which have much wider tips (~0.1 mm) and use metallic capillaries.[42–44] 1-3 µL of sample is loaded into the capillary, and the fine tip leads to slow flow rates of 1-50 nL/min, compared to ESI which uses µL/min.

This confers several advantages for native MS experiments. Firstly, the lower flow rate reduces sample consumption, which is ideal for studying biological samples as their production can often be time consuming and have low yields.[64] This is particularly advantageous for the study of membrane protein samples, as they are often inherently unstable in solution and traditionally have lower yields than soluble proteins.[115] Secondly, nESI is more tolerable to the presence of non-volatile salts, which can be critical for certain protein complexes.[102] This increased tolerance can be attributed to the nuanced differences between nESI and traditional ESI. In both processes, the first charged droplets undergo solvent evaporation until the Rayleigh limit is reached and Coulomb fission occurs (see **section 1.2**).[42,50,51] This increases the concentration of the analyte, as well as the concentration of any non-volatile salts as they do not undergo solvent evaporation.[64] Therefore, the more fission events required to form the charged residue (**Figure 1.1**) the higher the non-volatile salt concentration in the final droplet with the single charged protein molecule. Thus, the size of the original droplet formed at the Taylor cone at the capillary tip directly effects the salt concentration in the final droplet. nESI capillary tips have a much smaller diameter and consequently produces primary droplets generally one order of magnitude smaller than those produced in ESI (150-200 nm compared to µm range of ESI).[47,59] This difference in primary droplet size removes roughly one round of Rayleigh fission and thus a salt concentration step.[64] Lastly, the desolvation process of nESI is a softer method of introducing protein complexes into the gas phase. The smaller droplets produced reduces the number of fission events and reduces the energy of collision used during desolvation. nESI is usually initiated by applying a potential of 0.5-1.5 kV to the capillary, however it is common for an auxiliary backing gas to be applied to the sample to help initiate and maintain a steady stream.[64]

It is possible to use submicron borosilicate glass capillaries for nESI experiments to confer several further advantages such as desalting large macromolecules from biochemical buffers with high concentrations of non-volatile salts, preventing aggregation inside electrospray droplets to identify complexes that exist in solution, and to improve the measurement of dissociation constants between ligands bound to biopolymers.[116] However, the use of submicron nESI emitters can be problematic

due to issues with sample clogging and difficulty obtaining reproducible tip diameters and ion currents. Therefore, despite the possible advantages, the use of submicron nESI emitters is not commonplace in the field.

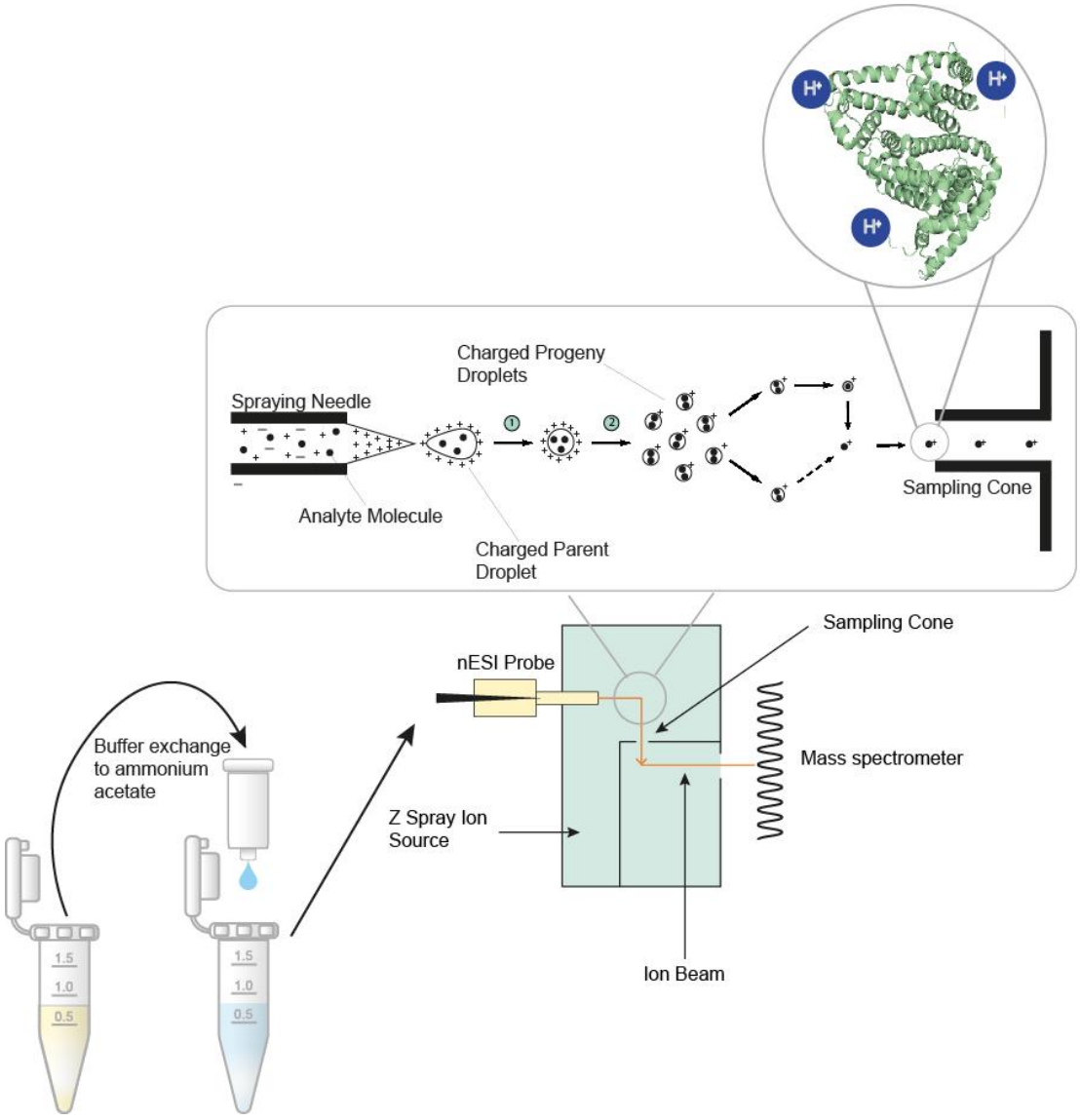


Figure 1.9. Schematic of a native MS experiments. The sample is buffer exchanged to a volatile native MS buffer such as ammonium acetate. 1-3 µL of sample is added to a metallic coated capillary adapted for nESI. It is then connected to the MS instrument, where an electric potential is applied across the capillary and ionisation occurs.

1.3.4 Interpreting electrospray mass spectra

nESI can produce multiply charged ions that can be analysed by MS and appear as multiple peaks of different *m/z* in a Gaussian distribution, with each peak resulting from a different charge state.[98] The Gaussian distributions are known as charge state distributions (CSD), with the number of acquired charges often correlating to the surface area of the protein sample. Primarily, the charges are localised at basic residues in positive ion mode and acidic residues in negative ion mode.[117] The molecular mass and charge states of a protein sample can be derived from **Equations 1.11-1.13**. These simultaneous equations are based on adjacent peaks within a CSD, assuming the peaks within a CSD differ by one charge (n, n+1).[60]

$$(m/z)_1 = \frac{m^{protein} + (z+1)m^{proton}}{(z+1)}$$

$$(m/z)_2 = \frac{m^{protein} + zm^{proton}}{z}$$

Equation 1.11. **Simultaneous equations of adjacent peaks in a CSD.** *z* is the number of charges and *m* is the mass of the protein.

If $(m/z)_2 > (m/z)_1$ and $m^{proton} = 1$, the above simultaneous equations can be combined to eliminate *m*.

$$z(m/z) = \frac{(m/z)_1 - 1}{(m/z)_2 - (m/z)_1}$$

Equation 1.12. Combined equation from 1.11. *z* is the number of charges and *m* is the mass of the protein.

Then the value of *z* can be added back to the equation to give the mass of the protein.

$$m^{protein} = z((m/z) - 1)$$

Equation 1.13. Calculating the protein mass from a CSD. *z* is the number of charges and *m* is the mass of the protein.

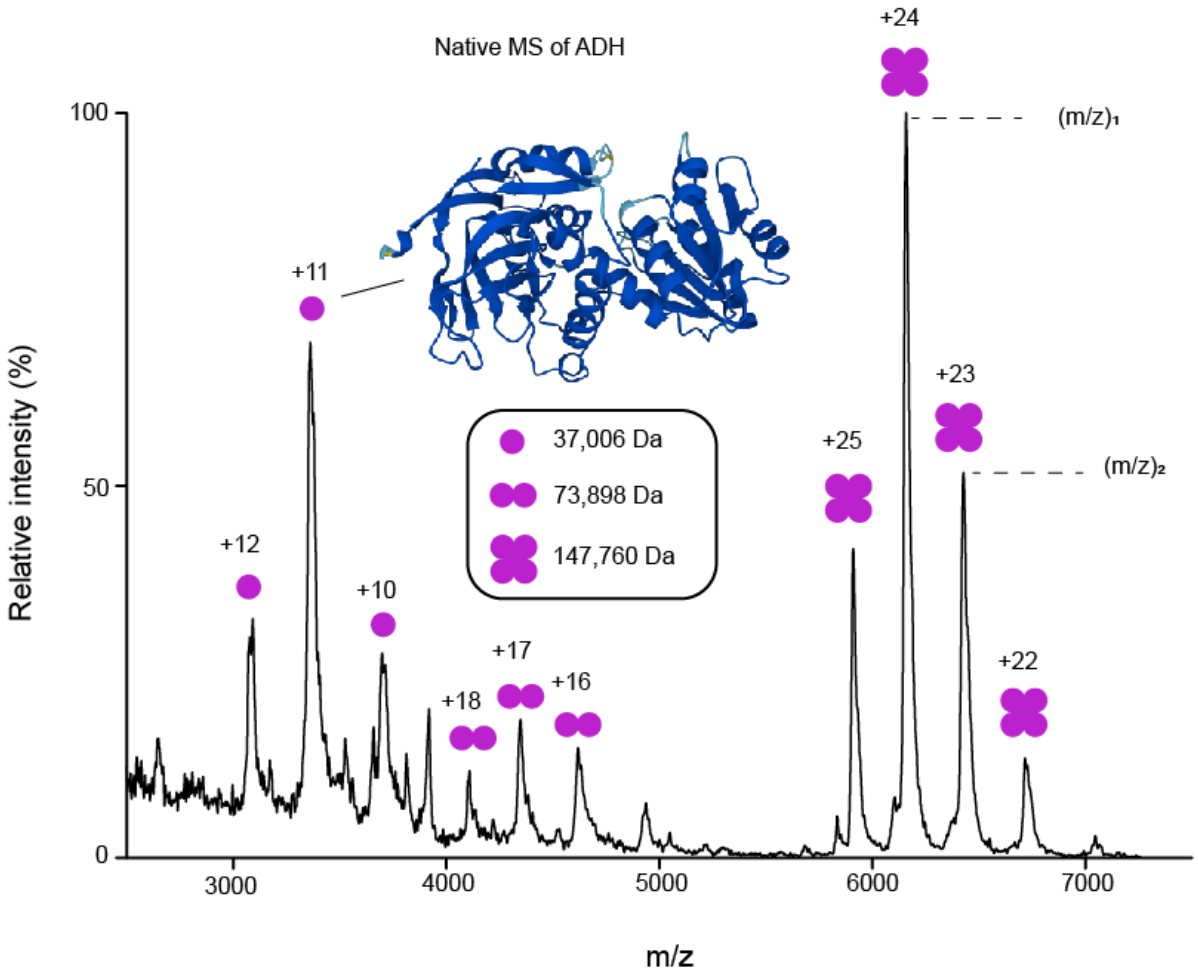


Figure 1.10. Native MS of ADH. Typical native MS spectrum of a soluble protein. The sample is ADH from *Saccharomyces cerevisiae* (Sigma-Aldrich). ADH presents as a monomer, dimer, and tetramer in the spectrum. $(m/z)_1$ and $(m/z)_2$ from **Equation 1.11** is labelled on the tetramer CSD. ADH protein structure predicted from AlphaFold (accession code: A0A3G3NDH9).[30,31]

To assist with the analysis of native MS data and assigning masses to protein peaks, several deconvolution software is available. In this thesis, UniDec was used to facilitate the analysis of native MS, which uses a Bayesian deconvolution algorithm.[118] UniDec requires minimal input from the user and allows for fast and quantitative analysis of mass spectra.[119] It is important to note that measured masses of are often larger than theoretical masses.[98] This is due to incomplete desolvation, and therefore buffer/salt adducts are observed in the spectra, adding to the mass. By calculating the difference between the theoretical and measured masses, it is possible to estimate what the adduct is. Furthermore, the CSD of the protein reflects its folded state in the gas phase. A folded protein typically has narrower CSDs with a lower average charge, whereas unfolded proteins have broader CSDs with higher charge.[98] This is because unfolded proteins have a larger surface area that is available

to take up charge during the ionisation process. It is for this reason that CSDs can also indicate proteins containing intrinsic disorder (see **chapter 2**).[120–122]

1.3.5 Instrumentation for native MS

Throughout this thesis, two different instruments were used to acquire native MS of protein samples. The first is the Synapt G2-Si mass spectrometer (Waters); this is a quadrupole-time of flight (Q-ToF) mass spectrometer (**Figure 1.3**). The Q-ToF is modified for the transmission of high-mass protein complexes. Firstly, the pressure of all pumping stages is increased throughout the system, and the RF of the quadrupole is reduced to 300 kHz to allow transmission of ions up to 32,000 *m/z*.[86,98] The Synapt G2-Si mass spectrometer contains a Triwave collision cell, consisting of two collision cells (trap and transfer) flanking a travelling wave ion mobility (IM) cell. IM separates gas phase ions based on their charge and shape, analogous to electrophoresis in the condensed phase.[123] Ions are separated by their interactions with a buffer gas in the IM cell (e.g. Helium or Nitrogen), adding an additional dimension that is particularly useful for separating complex ions.[124–127] Due to the flanking collision cells, this also allows for IM of membrane proteins released from its mimetic environment. No IM data is presented in this thesis, so it is not detailed in the introduction. For more information on IM please see this review by Lanucara *et al*. (2014).[123]

The second instrument used is the Q-Exactive Plus ultra-high mass resolution (UHMR) mass spectrometer (ThermoFischer). The Q-Exactive plus is a quadrupole-orbitrap mass spectrometer, with high resolving power of 140,000 full-width half-maximum (FWHM). This resolving power is due to the orbitrap mass analyser and 'in-source' trapping, which allows for greater desolvation of protein complexes and the dissociation of oligomers in the front end of the instrument. It has been used to analyse large macromolecular complexes up to 50,000 *m/z*.[128] **Figure 1.11** shows the features of this instrument in more detail.[129] The high resolution makes this instrument ideal for observing protein-ligand complexes, as shown in **chapter 4**.

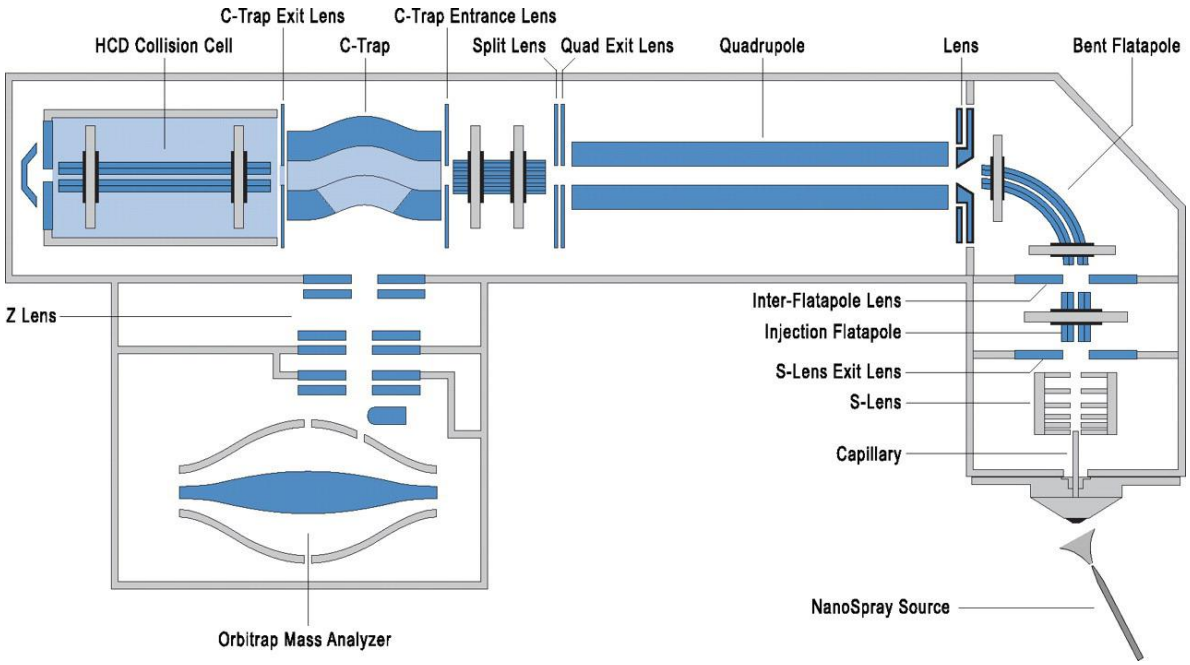


Figure 1.11. Schematic of the Q-Exactive Plus mass spectrometer. Ions are injected into the source by nESI. The injection flatapole is pulsed down to a negative voltage to improve desolvation, and the inter-flatapole lens has a high positive potential to prevent ions eluting out. Trapping occurs through restoring voltage levels which allows low energy elution of ions into the bent flatapole. This focuses ions using an axial DC field and a focusing RF field. The ions enter the quadrupole optimised for the transmission of high mass ions. Ions can then enter the C-trap where ion packages are sent to the Orbitrap for mass analysis or HCD collision cell for fragmentation before mass analysis. Taken from Michalski *et al.* (2011).[129]

1.4 Hydrogen Deuterium Exchange Mass Spectrometry

1.4.1 Introduction

Proteins are not static molecules in solution, and their dynamics are essential for their function. Most biophysical techniques lack the resolution to link protein motions with protein function, making it hard to fully understand protein mechanisms.[130] HDX-MS is a powerful analytical tool for probing protein conformational dynamics.[33] HDX-MS monitors isotopic exchange of hydrogens, in the protein backbone amides, to heavier deuterium. The rate of exchange is dependent on the folded state of the protein, and both the dynamics and intrinsic chemical properties of the underlying amino acid sequence.[131] Thus HDX is a sensitive method for investigating protein dynamics along the backbone. Pioneering work was done in the 1950's by Linderstrøm and Lang, and the coupling of HDX to MS was introduced by Zhang and Smith in the 1990's (previously NMR was used).[33,130,132–134] HDX-MS is a versatile technique that has many advantages; an unlimited mass range, low sample concentrations and virtually any solution-phase condition or buffer system can be used.[135]

HDX-MS can be widely applied across structural biology. It can investigate protein folding pathways, protein dynamics over time or in relation to stability, protein-ligand interactions (including epitope mapping) and conformational characterisation studies.[136–138] Advances in the field has led to HDX-MS being amenable to study membrane proteins. For example, Merkle *et al.* (2018) used HDX to investigate the translocation mechanism in the neurotransmitter sodium symporter protein LeuT.[139] They uncovered LeuT segments involved in substrate binding, and the dynamics associated between an outward and inward facing configuration. Applying HDX to membrane proteins involved in multidrug resistance is becoming increasingly important. Eisinger *et al*. (2018) studied NorM_PS, a membrane transporter from the multidrug and toxic compound extrusion (MATE) family.[140] HDX revealed an occlusion in the proposed binding site, and the closure of a cytoplasmic cavity coupled to the formation of a periplasmic cavity. In this chapter, the fundamentals and workflow of HDX-MS will be detailed, as well as its experimental application to membrane proteins.

1.4.2 Fundamentals of HDX-MS

HDX focuses on backbone amide hydrogens because at physiological pH they exchange with deuterons within 1-10s when incubated with D_2O, whereas hydrogen bonded to main chain carbons exchange too slowly and side chain hydrogens exchange too quickly to be measured.[33] Amide HDX can be acid, base or water catalysed, but base catalysed HDX is the dominant mechanism at physiological pH.[130] The mechanisms of acid and base catalysed HDX is shown in **Figure 1.12**.

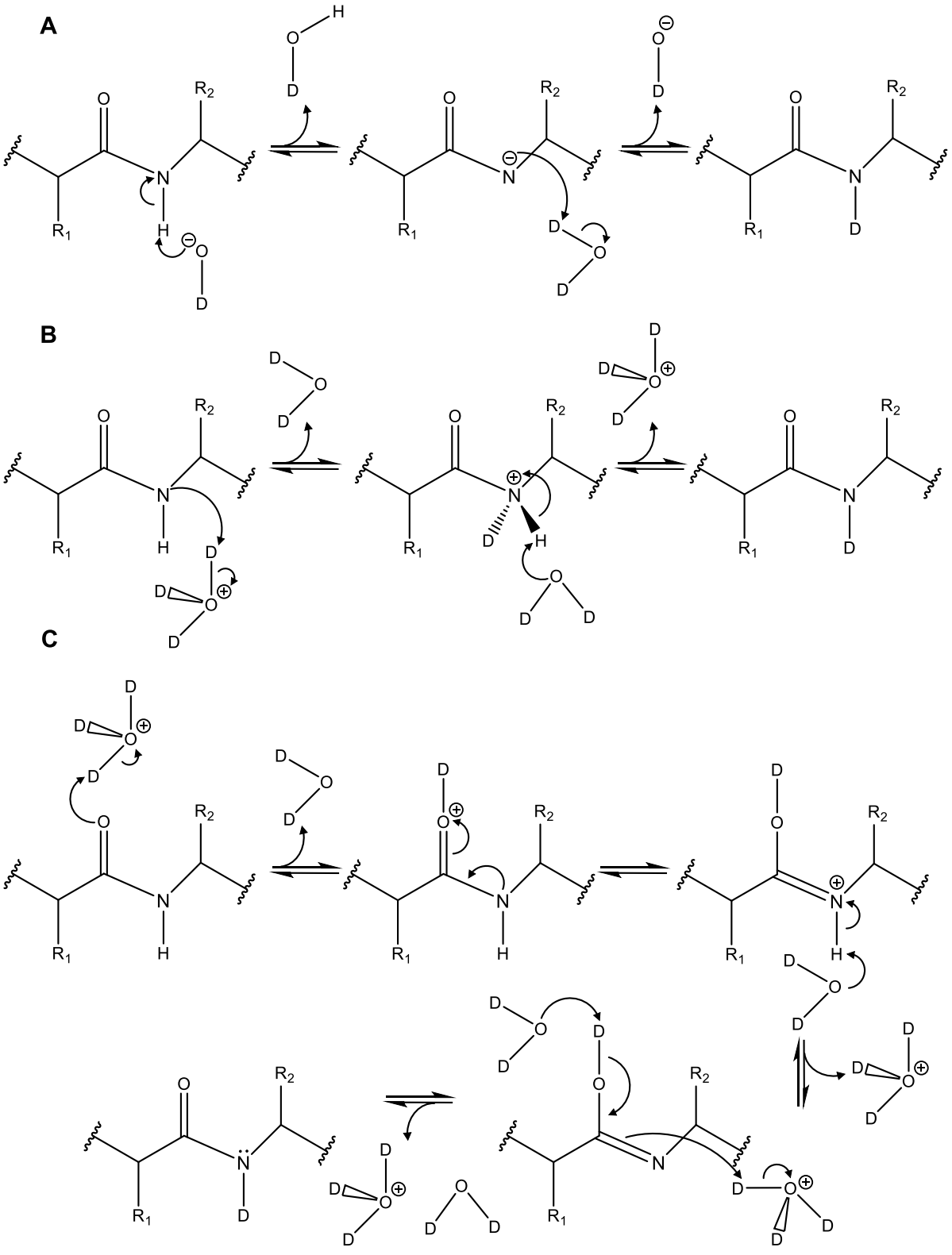


Figure 1.12. HDX mechanisms. A. Base catalysed HDX. **B.** Acid catalysed *N* protonation. **C.** Acid catalysed *O* protonation.

As the pH can affect the mechanism by which HDX can occur, HDX is evidently dependent on pH.[33,141] **Figure 1.13a** shows HDX is lowest at pH ~2.5, and this allows the reaction to be sufficiently quenched by lowering the pH, in order to measure deuterium incorporation via MS. Since the rate of HDX is affected by pH, this means labelling experiments conducted at pH values other than physiological pH will either exchange slower or faster within the same labelling time point. Therefore, a time window expansion method can be applied to modify the time points for an accurate reflection of HDX under physiological conditions. The effect of pH on protein conformational dynamics can be described by **Equation 1.14**.[142,143]

$$k_{ch} \sim k_{OH}[OH^-] = A \exp\left(\frac{-E_a}{RT}\right)[OH^-]$$

Equation 1.14. Effect of pH on protein dynamics. *A* is the frequency factor, E_a is the activation of the dominant base catalysed HDX exchange in the range of pH 5-10.

With temperature remaining constant, **Equation 1.15** can be derived from **Equation 1.14** to calculate the ratio of HDX rate constants that are applicable to two pH conditions.

$$\frac{k_{ch1}}{k_{ch2}} = \frac{[OH^-]_1}{[OH^-]_2} = \frac{k_w / [H^+]_1}{k_w / [H^+]_2} = \frac{10^{-pH_2}}{10^{-pH_1}} = 10^{pH_1 - pH_2}$$

Equation 1.15. HDX time window expansion. When temperature is constant the above equation can be derived from the Arrhenius equation to calculate the ratio of amide hydrogen exchange rate constants that are applicable to the two pH conditions. For example, this shows the intrinsic HDX rate decreases $10^{7.4\text{-}6}$ fold when the pH decreases from 7.4 to 6.0.

Therefore, **Equation 1.15** states that the intrinsic HDX rate decreases $10^{7.5\text{-}x}$ fold when decreasing the pH from 7.5 to x. For example, to have the same amount of HDX exchange between pH 7.4 and pH 6.0, one must deconvolve this effect by converting the labelling times at pH 7.4 to new, equivalent, labelling times at pH 6.0. Thus, 60 seconds at pH 7.4 becomes 1507 seconds at pH 6.0 (60 x $10^{7.4\text{-}6.0}$).

HDX is also affected by temperature (**Figure 1.13b**) through an increase in the water ionisation constant and the rate of diffusion.[33,144] **Equation 1.16** shows the exchange rate for an unstructured peptide.[145]

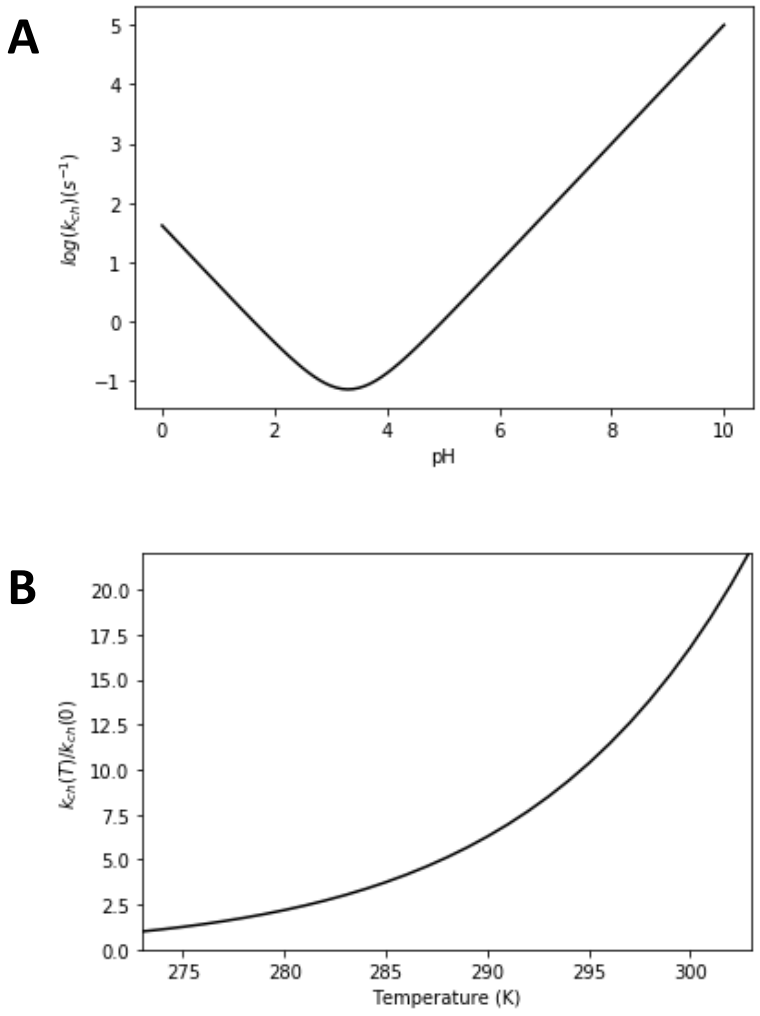


Figure 1.13. pH and temperature effects on HDX. **A.** pH profile of the chemical hydrogen exchange rate *(k_{ch})*. **B.** The chemical hydrogen exchange rate (k_{ch}) as a function of temperature. Created using Matplotlib.

$$k_{ch} = k_{int,H}[H^+] + k_{int,OH}[OH^-] + k_{int,H_2O}[H_2O]$$

Equation 1.16. Exchange rate of amide hydrogens in an unstructured peptide. k_{ch} is the exchange rate, $k_{int,H/OH/H2O}$ is the intrinsic rate constants for acid/base/water catalysed reactions.

For each site in an unstructured polypeptide, it is possible to calculate k_{ch} using pH, temperature, primary sequence and ionic strength of solution data.[145] The primary sequence affects the HDX rate by steric blocking of neighbouring side chains and charge delocalisation. The influence of the ionic strength of solution on the rate depends on local electrostatic fields within the peptide.

The situation changes considerably when dealing with folded proteins. Whilst experimentally controlling pH and temperature, the rate of HDX is dependent on solvent accessibility to exchange sites and intramolecular hydrogen bonding.[130] In folded proteins, some amide hydrogens exchange quickly (unstable areas) whilst others only exchange after months (more stable). **Equation 1.17** shows the rate constant for HDX at each individual amide linkage in a folded protein.[146]

$$k_{ex} = k_f + k_u = (\beta + k_{unf})K_2$$

Equation 1.17. Exchange rate of amide hydrogens in a folded protein. k_f is the contributions of exchange from folded form, k_u is the contributions of exchange from unfolded form.

k_{ex} is a sum of the contributions of exchange for the folded (k_f) and unfolded (k_u) forms of the protein. For amides near the surface and not participating in intramolecular hydrogen bonding (may still exhibit hydrogen bonding to H_2O) exchange from the folded form dominates (**Equation 1.18**).

$$k_f = \beta k_2$$

Equation 1.18. Rate constant for exchange in a folded state. β is the probability factor for exchange from folded forms, k_2 is the rate constant for HDX at each amide linkage in an unstructured peptide.

Exchange in unfolded forms requires substantial movement of the backbone to expose backbone amides to deuterium; this can be localised or global. The rate constant for exchange from unfolded forms (**Equation 1.19**) depends on the rate constant for exchange from an unfolded peptide (k_2) and the unfolding dynamics (k_1, k_{-1}).

$$F_H \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} U_H \xrightarrow[D_2O]{k_2} U_D \underset{k_1}{\overset{k_{-1}}{\rightleftharpoons}} F_D$$

Equation 1.19. Linderstrøm-Lang model for exchange in unfolded proteins. F is folded, U is unfolded, H is hydrogenated, D is deuterated. k_1/k_{-1} are folding/unfolding kinetics, k_2 is the rate constant for exchange from unfolded protein.

Under physiological conditions, it is more common to see EX2 kinetics where $k_{-1} >> k_2$. **Equation 1.20** gives the rate constant for exchange from unfolded proteins in EX2 kinetics. In this system the protein exhibits many random local and rapid folding/unfolding events to a state capable of exchange, yet the probability of exchange for a single event is small.

$$k_u = \frac{k_1}{k_{-1}} k_2 = k_{unf} k_2$$

Equation 1.20. EX2 kinetics. k_u is the rate constant for exchange from unfolded protein, k_1/k_{-1} are the folding/unfolding dynamics, k_2 is the rate constant for HDX at each amide linkage in an unstructured peptide, k_{unf} is the equilibrium constant describing the unfolding process.

A few proteins can naturally exhibit EX1 kinetics, where $k_2 >> k_{-1}$ and the rate constant for exchange from unfolded forms is given by the unfolding rate constant k_1 (**Equation 1.21**). Here, a protein undergoes a cooperative unfolding event involving several residues that all exchange before refolding. EX1 kinetics can be induced using denaturant or by increasing the pH.[147,148]

$$k_u = k_1$$

Equation 1.21. EX1 kinetics. k_u is the rate constant for exchange from unfolded protein, k_1 is the unfolding rate constant.

It is possible to differentiate between EX1 and EX2 kinetics in protein regions through characteristic isotope patterns in the mass spectra (**Figure 1.14**).[33,130,149] Regions undergoing EX2 kinetics show binomial isotope patterns, where the central *m/z* value of the isotope envelope increases gradually due to multiple folding/unfolding events. EX1 kinetics show a bimodal isotope pattern with distinct isotope envelopes separated on the *m/z* scale. It is possible to have regions display both EX1 and EX2 kinetics simultaneously.[130]

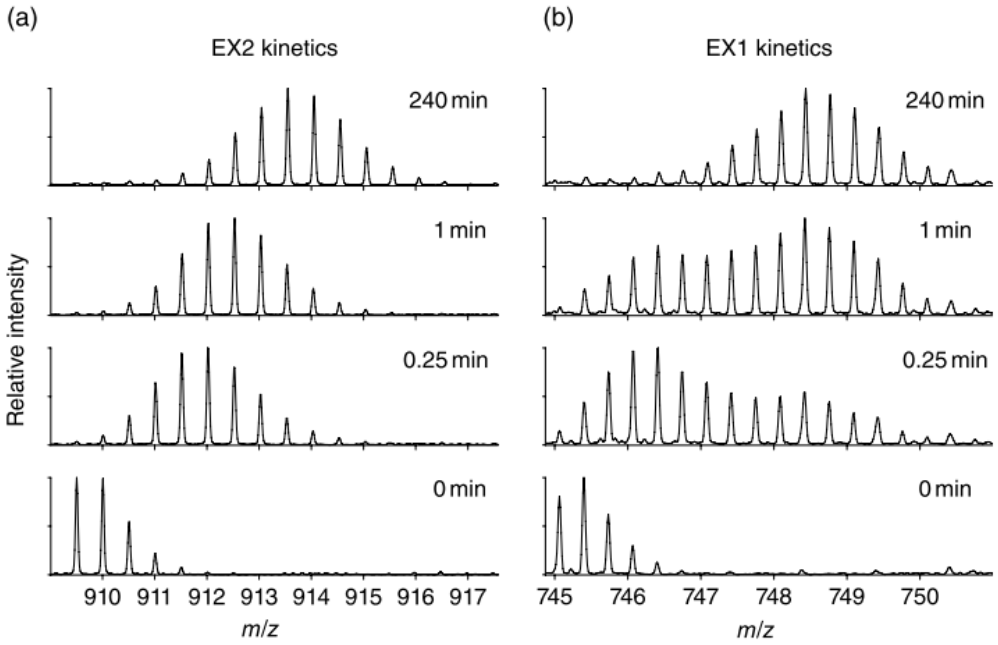


Figure 1.14. MS characteristics and data analysis of EX1/EX2 kinetics. Figure from Weis *et al.* (2016).[33]

1.4.3 Workflow

1.4.3.1 Deuterium Labelling/ Exchange Reaction

Correct preparation of protein samples and buffers is essential to the HDX experiment. Proteins must be overexpressed in a buffer that maintains it in an active and folded state.[33] Overexpression itself must be in a suitable expression system, and a size exclusion chromatography (SEC) step is advised to ensure the protein sample is homogenous and soluble.[150] HDX requires a lower protein concentration than other biophysical techniques like NMR and x-ray crystallography, but stock concentrations need to be higher to account for the 10-50x dilution factor upon initialization with the exchange reaction. Therefore stock concentrations need to be at a level where 10-100 picomoles of protein is present per injection, but not so high they contribute to heterogeneity, through aggregation, misfolding or precipitation.[33] The D_2O labelling buffer should be chemically identical to the native buffer, just with a D_2O content of 99%, ensuring a final D_2O concentration of >90%.[130] The pD of the labelling buffer must be corrected to be equivalent to the aqueous buffer (**Equation 1.22**).[151]

$$pD_{corrected} = pH_{read} + 0.4$$

Equation 1.22. pD correction.

The first step in the workflow (**Figure 1.15**) is the labelling reaction, where proteins are diluted in D_2O under desired conditions to start the exchange reaction.[141] The number of experiments in the HDX workflow for a given protein depends on the number of time points of HDX monitored. A typical amount would be 5 time points, each with 3 repeats: 15 experiments. Hydrogen exchange is usually automated by robotics today, thus increasing the efficiency and reproducibility of experiments. There are two types of labelling experiment; continuous and pulsed labelling.[33,130,152] In continuous labelling a protein is exposed to deuterated buffer while under conformational flux; multiple rounds of dynamic reactions between multiple states. It averages the HDX across all the dynamic states to provide an average dynamic snapshot for a specific time point. It is very useful at measuring most unfolding events in proteins, as these happen at relatively slow rates. Pulsed labelling is effective when timescales between a proteins conformational state is longer than the HDX, so short durations of HDX are essential.[33] Typically the sample is perturbed to equilibrium by a denaturant/substrate as a starting point. It is used to monitor short-lived folding intermediates as reaction conditions change, thus probes folding mechanisms/kinetic intermediates.[149,152–154] At each time point an aliquot of the reaction is exposed to a quick pulse (millisecond) of deuterated buffer, rapidly quenched and characterised by MS.

At the end of the labelling period, HDX must be quenched (step 2, **Figure 1.15**) to retain the extent of deuteration incorporation reproducibly by MS.[33] This is achieved by lowering the pH to 2.3-2.5 and the temperature to ~0°C (**Figure 1.13**) via a chilled acidic quench solution.[144] The quench solution needs to bring the aqueous buffers pH down to ~2.5 so the same occurs in the labelling buffer. The composition of the quench buffer needs to be empirically determined for each protein/buffer system to give reproducible results. Formic acid and trifluoracetic acid are typically used, and additional reagents including denaturants, reducing agents and detergents can be included to aid the digestion step (see **section 1.3.3.2**).

1.4.3.1.1 Back-exchange consideration

Even though HDX is greatly reduced at quench conditions, it still continues. The loss of deuterium is known as back-exchange, and this mostly occurs when in contact with the aqueous mobile phase during digestion and separation. Back-exchange is affected by pH, temperature, time and the physiochemical properties of the peptide.[33,155–157] It is key that digestion and separation steps are carried out reproducibly for the same duration (rapidly) and at quench conditions.[158] The level of back-exchange for a particular HDX-workflow must be characterised in detail through analysis of model proteins (e.g. Phosphorylase B) that are incubated in D_2O until 100% deuterated.[131] **Equation 1.23** shows how the back-exchange for peptides in the model system can be estimated.[130,133,159] In well

controlled systems, the overall back-exchange ranges between 20-30%, and only very few peptides should have back-exchange levels over 50%.[133,136]

$$Back\ Exchange = \left(1 - \frac{m_{100\%} - m_{0\%}}{N \times D_{frac}}\right) \times 100$$

Equation 1.23. Calculating back-exchange. $m_{0\%}$ is the non-deuterated peptide centroid mass, $m_{100\%}$ is the maximally labelled peptide centroid mass, N is the theoretical number of backbone amides in the peptide, D_{frac} is the fraction of D/H in the labelling buffer used.

Equation 1.24 shows how the number of deuterons exchanged in each peptide can be adjusted for back-exchange.[133]

$$D_{corr} = \frac{m_{expt} - m_{0\%}}{m_{100\%} - m_{0\%}} \times N$$

Equation 1.24. Adjusting for back-exchange. D_{corr} is the corrected deuterium exchanged in the peptide, m_{expt} is the experimentally determined mass of peptide, $m_{0\%}$ is the non-deuterated peptide mass, $m_{100\%}$ is the fully deuterated peptide mass, N is the theoretical number of backbone amides in the peptide. N excludes the N terminus and prolines, as N terminal amides exchange too rapidly for experimental analysis and prolines do not have a backbone amide.

1.4.3.1.2 Methods for back-exchange correction

Given that the sequence of a polypeptide governs the extent of back-exchange under quench conditions, each peptide will exhibit a different amount of back-exchange.[145] The extent of back-exchange is determined by k_{ch} of each backbone amide hydrogen of each residue composing that peptide.[160] Therefore, to accurately measure the absolute deuterium incorporation of every peptide, a fully deuterated protein sample should be measured and then each individual peptide should be back-exchange corrected using the corresponding fully deuterated peptide (MaxD). It is important to note the MaxD for a peptide excludes the N terminus and any prolines.[133] Conversely, in experiments measuring the relative difference between the same proteins in two or more different states, called differential HDX (ΔHDX), it is not necessary to correct for back-exchange as the generated peptides will exhibit the same back-exchange in both states, as long as the experimental conditions remain constant. Therefore, the observed states would exhibit the same difference regardless of being corrected for back-exchange.

Achieving MaxD values is challenging, as it requires all backbone amide positions to be in an exchange competent state, often meaning breaking hydrogen bonds affecting any backbone amide.[160] Additionally, it is not possible to know if complete deuteration is actually achieved. However, there are several methods to obtain MaxD values for proteins to provide the most accurate results, each with their own pros and cons. One of the earliest methods by Zhang and Smith in 1993 involved incubating a protein sample in 100% deuterated buffer at an elevated temperature.[161] As long as protein aggregation is avoided, this can be successful for some proteins. Also it has been found that heating a protein to a temperature 5 ˚C below the melting point (Tm) and labelling for 10 mins provided successful results.[162] The shortened deuteration time is ideal, and reduces the possibility of the histidine C2 to be deuterated.[163] However, it requires knowledge of the proteins melting point, proteins may aggregate and sometimes this will not lead to full deuteration. Another method is to expose a protein to 100% deuterated buffer for an extended period of time (>24 hours) and allow natural fluctuations to lead to full deuteration.[164] Nonetheless, some proteins may have exceptionally stable regions that do not become fully deuterated even after several weeks. Sowole and Konermann (2014) detailed an acid denaturation protocol to achieve MaxD values, but this can require long labelling times and lead to protein aggregation or modification.[165] The use of chemical chaotropes such as guanidinium and urea is a possible way to denature the proteins and does not require as long incubation times as acid denaturation, but that can cause chemical modifications and species in solution may interact with the protein, interfering with the HDX reaction.[166] One alternative to denaturing the folded protein is to pre-digest the protein, extract the peptides and lyophilise them, then deuterate and analyse them using LC-MS.[167] However, incorporating this method into HDX workflows whilst maintaining the same experimental conditions, without losing peptides due to weak signals, is fairly challenging.[160]

Often choosing the right method to collect MaxD values depends on the protein system at hand and may involve some trial and error. Often some of the most effective ways to achieve MaxD values is to combine some of the mentioned methods. Peterle *et al.* (2022) detailed a protocol that involves chemical denaturation, heating and short incubation times, which showed successful results.[160]

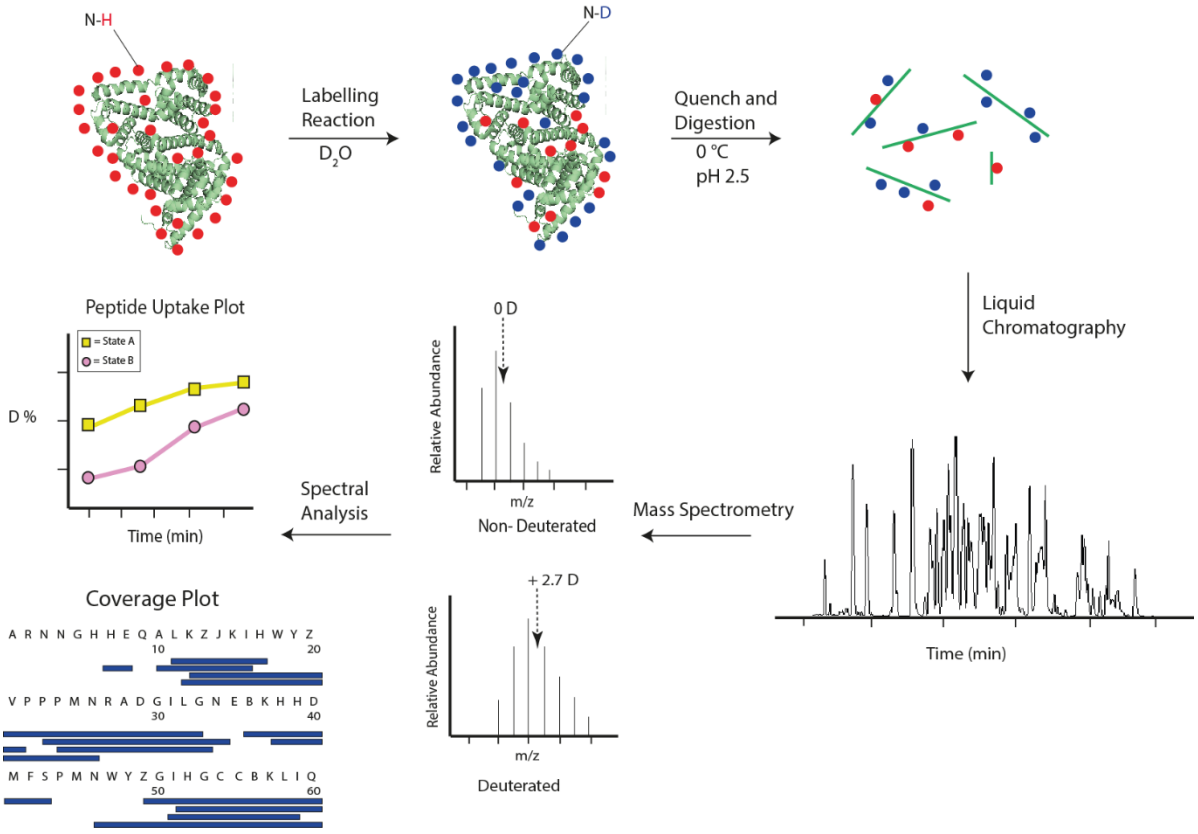


Figure 1.15. Workflow of a bottom-up HDX-MS experiment. Protein samples equilibrated in their aqueous buffer are diluted with the equivalent deuterated buffer in the labelling reaction. The reaction is quenched at various time points (pH 2.5, 0°C) and digested with an acid protease (usually pepsin). These deuterated peptides undergo liquid chromatography separation (UPLC) before they undergo MS analysis. This measures deuterium incorporation (blue dots). Further analysis of the data can be global (coverage plot) or local (peptide uptake plots).

1.4.3.2 Proteolytic Digestion

The protein is digested to increase the resolution and localize information in an HDX experiment (step 2, **Figure 1.15**).[33,133] Proteolytic digestion must be done at quench conditions, so acid proteases are used. The most common is pepsin, a well characterised, non-specific protease with activity at pH 2-4.[155,168] The sites of cleavage cannot be predicted from the amino acid sequence, but under the same conditions it will cut in the same place so reproducibility is enabled.[130] One issue with pepsin digestion is the production of undesirably long peptides (>15 amino acids) with a large protein sample. It is challenging to interpret small deuterium changes in a long peptide; for example, a two-deuterium incorporation in a 20 amino acid peptide cannot be attributed to changes in the dynamics of the whole peptide. Other acid proteases can be used (e.g. aspergillipepsin), and the use of multiple acid

proteases with different specificities in combination can yield greater resolution.[169,170] As HDX resolution is limited to the resolution achieved by the protease fragments obtained, single amino acid resolution of protein dynamics is possible through information constructed from overlapping peptides from multiple protease digestion.

Proteolytic digestion can be offline or online. Online involves a column packed with immobilized pepsin that is connected to the liquid chromatography (LC) system, whilst a cooling system limits back-exchange.[158] The protein injected at one end of the column is rapidly digested and the resulting peptides come out the other end, entering the LC. Immobilized pepsin is commercially available as a pre-packed column, or slurry, otherwise it can be made in the lab.[171,172] The high quantity of pepsin stays in the column and is used for other samples; this requires routine cleaning and monitoring to prevent peptide carryover from other experiments.[136,173] Online digestion usually takes 1-3 minutes.[136] Offline digestion involves the addition of protease, either in solution or immobilized on solid supports, to the sample (kept at 0°C) for a predetermined time before injection into the LC system.[33] Immobilized protease is removed prior to injection into the LC, to prevent blockages. Offline digestion offers the advantage of eliminating the possibility of peptide carryover.[173]

1.4.3.3 Liquid Chromatography and Mass Spectrometry

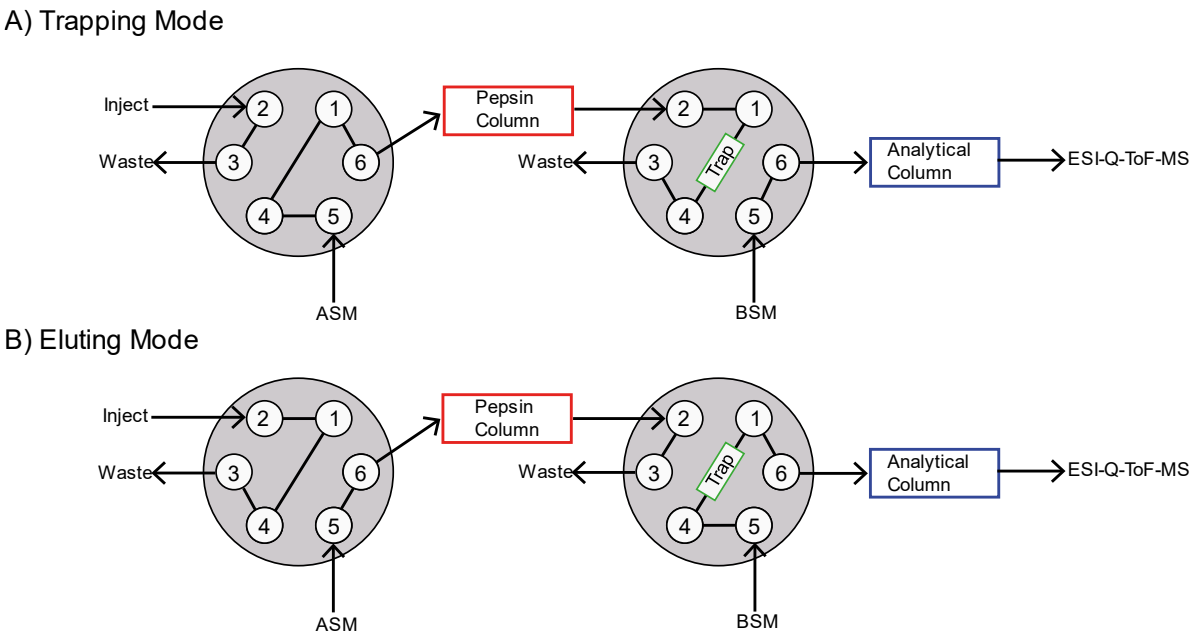


Figure 1.16. Layout of the two valve LC system. The ASM (Auxiliary Solvent Manager) pumps through aqueous phase buffer. The BSM (Binary Solvent Manager) pumps through organic phase buffer or aqueous phase buffer. **A.** Shows the layout in trapping mode. **B.** Shows the layout in eluting mode.

Figure 1.16 shows the layout of the two valve LC system. Separation is needed for better MS performance, and to allow D_2O in any side chain positions to be washed away, as this happens very quickly. After digestion the peptides are trapped on the C18 guard column (trap column) where they are washed with the mobile phase buffer (0.1% FA, pH 2.5, 0°C), simultaneously removing salts and concentrating the peptides. Peptides are eluted from the trap via a gradient of organic phase buffer (e.g. acetonitrile). The gradient required is usually quite sharp to minimise back-exchange (e.g. 2-60%, 10 mins). The interaction of peptides with the aqueous mobile phase buffer leads to increased back-exchange so quench conditions must be maintained, and the LC run time must be kept short and reproducible. Ultra-high pressure liquid chromatography (UPLC) is compatible and highly effective for the reverse phase LC separation of peptides over a fast run time of 8-10 minutes.[158,174] Peptides eluting from the UPLC are directly injected into the ESI source for MS analysis.

Peptides are analysed by MS/MS. MS^E is selected, which uses high and low energy CID to measure precursor ions (peptides) and fragments simultaneously (see **sections 1.2.3** and **1.4.3.3.1**), so the peptides can be identified from paired fragments with the help of computer software and relevant algorithms.[133] The peptides are identified from an undeuterated sample, to act as a reference, and this can be used to identify labelled peptides as regardless of deuterium incorporation each unique peptide has an unchanged unique retention time (time taken for LC elution). Therefore, at each timepoint, the output is the reproducibly digested peptides, each with multiple charge states, and each peptide's specific retention time. The instrument is calibrated before every experiment with NaI, and a Leucine-Enkephalin solution is infused constantly throughout along with the mobile phase flow to verify the exact mass; both these steps ensure a high mass accuracy is achieved.

1.4.3.3.1 Interpreting peptide fragmentation data

In the instrumentation used in this thesis, MS^E activation is utilised. This involves switching the CID conditions between a low and high energy regime, so alternate spectra can be generated that contains intact precursor ions (low energy) and their product ions (high energy).[175,176] With time alignment product ions are correlated with their intact precursor all in one scan event. **Figure 1.17a** details the different ways a peptide can fragment, creating different types of ions that can be seen in a spectrum.[177–179] In the main chain peptide, cleavage can occur between Cα-C, C-N and N-Cα, yielding six fragments; three amino terminal fragments (a_m, b_m, c_m) and three carboxy terminal fragments (x_n, y_n, z_n). The subscript m/n refers to the number of R groups on the fragment. Lowest energy bonds fragment first, so in a low energy collision cell you mostly get b- and y- ions, which then lose small molecules such as H_2O from its side chains.[70] **Figure 1.17c** shows an example of peptide fragmentation by CID, and the production of b-, y- and a- ions. At high energy all the fragment types seen in **Figure**

1.17a can be produced theoretically but not all will be observed, as the types of fragments seen depends on amino acid composition, peptide sequence, the amount of internal energy transferred and ion activation methods etc.[70]

Two other fragments may also be observed. One is an internal fragment which have lost the initial N- and C-terminal sides.[180] It is represented by a series of letters corresponding to the fragment species. Internal fragments are not widely abundant and appear in the low masses.[70] The second is immonium ions, which are seen in the low masses and labelled by a letter representing the parent amino acid. These are rarely observed for all amino acids. Additionally, the fragments defined in **Figure 1.17a** can also fragment if they have a mobile H^+ (e.g. b- ions can fragment to a- ions).

Through analysing the series of b-/y- ions, a particular fragment can be identified.[177] In an ideal spectrum, the b-/y- ions will form two ladders. **Figure 1.17c** shows an example of b- and y- ladders for the example peptide: FGVEQNVDMVFASFIR.[181] For b- ions, the ladder refers to the peaks corresponding to the prefix ions observed sequentially in the spectrum, with each prefix offset from the previous by the mass of an amino acid.[182] The y- ion ladder represents the suffix ions sequentially, and by concatenating the amino acids deduced by the sequential mass differences either of the ladders, the sequence of the fragment can be constructed.

The MS/MS spectrum can be complex; some expected fragments will not be observed, the presence of amino acid isomers (e.g. leucine and isoleucine), amino acid isobars (e.g. glutamine and lysine) and some amino acid combinations can equal that of a single amino acid. The dissociation of multiply charged ions further complicates the spectra. Protons on various protonation sites induces varying methods of fragmentation, and the different charge states on ions affects their isotopic distributions (**Figure 1.18**).[70,177] This can be solved by turning a peptide sequencing problem into a database searching problem, with various algorithms and computer software available.[183,184]

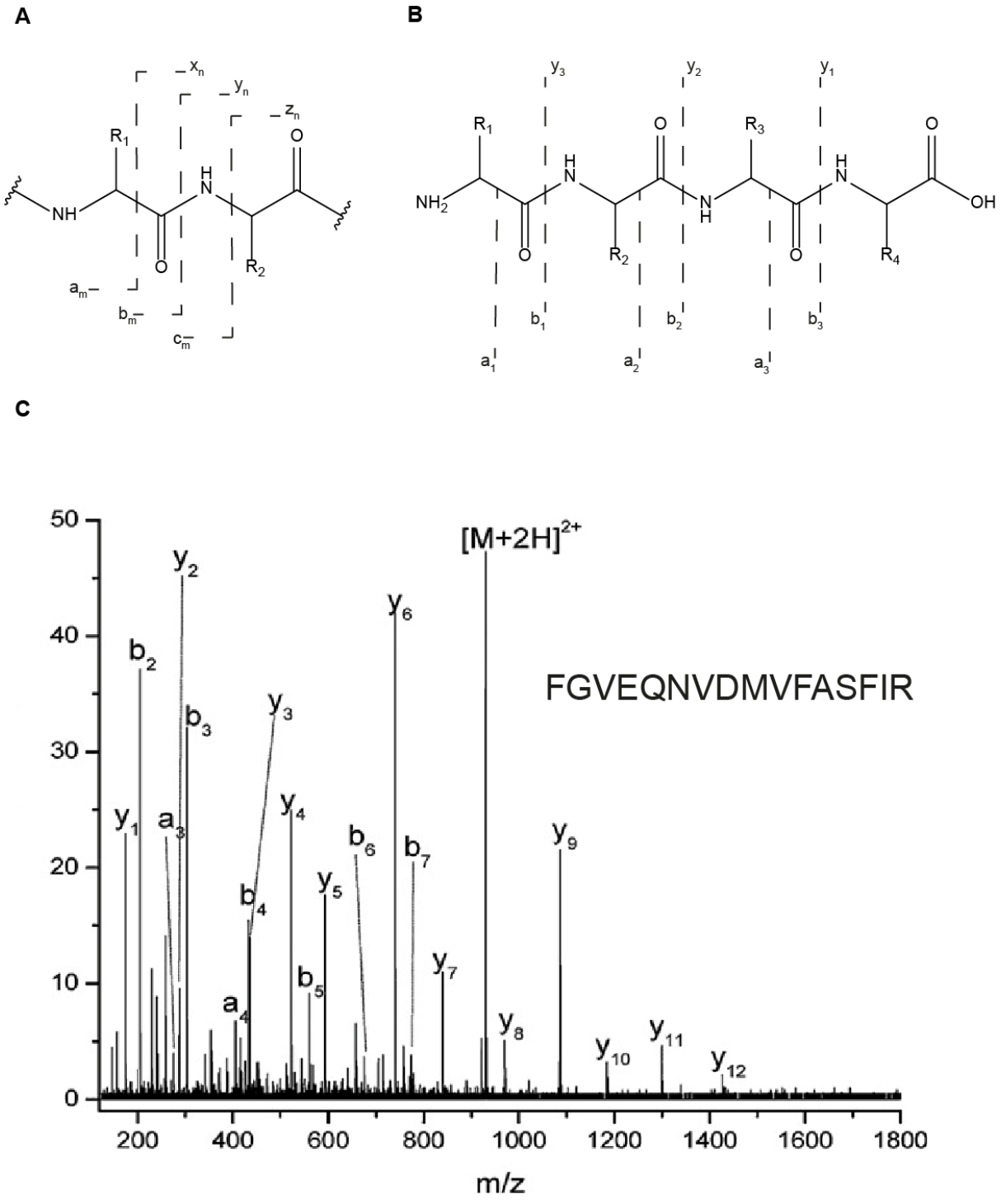


Figure 1.17. Fragmentation data. A. The theoretical six fragments made from main chain cleavage. Three amino terminal fragments a_m, b_m, c_m and three carboxy terminal fragments x_n, y_n, z_n. **B.** Classic fragmentation observed in a low energy CID collision cell. **C.** Ion spectrum from sample peptide: FGVEQNVDMVFASFIR. **A** and **B** created on ChemDraw, **C** adapted from Lau et al (2009).[181]

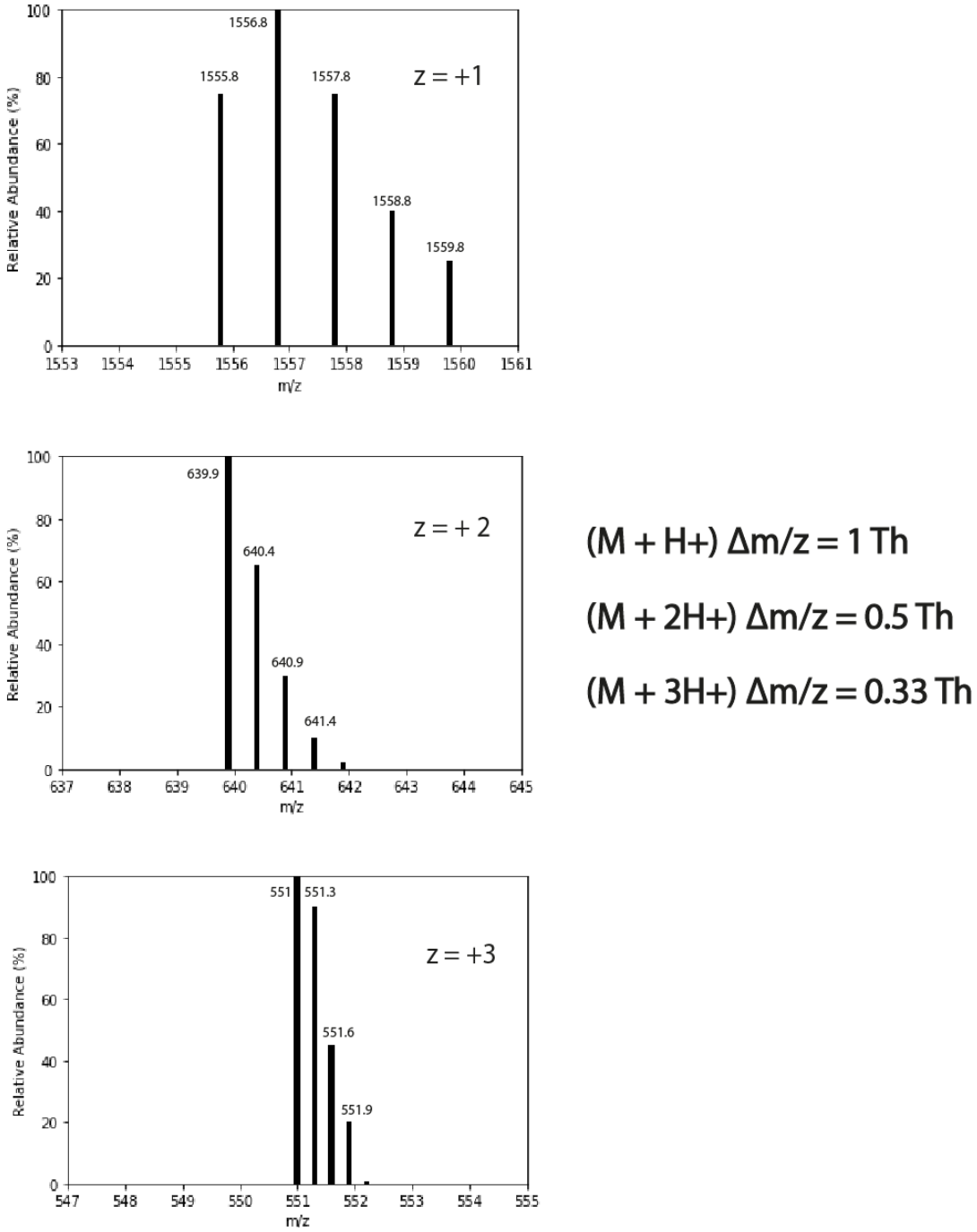


Figure 1.18. Isotopic distributions of peptides. Peptide clusters at charges +1, +2, +3. +1 = 1 Th difference between peaks in the cluster, +2 = 0.5 Th difference between peaks in the cluster, +3 = 0.33 Th difference between peaks in the cluster. Th is a Thompson unit, equivalent to 1 *u/e* where *u* is the atomic mass units and *e* is the elementary charge unit.[70]

1.4.4 HDX and Membrane Proteins

The HDX-MS practicalities discussed so far apply to soluble protein samples. A recent surge in structural studies of membrane proteins using cryo-EM or x-ray crystallography, coupled with an increase in sensitivity and the protein size/buffer tolerance has allowed HDX-MS to become applicable to membrane proteins.[135,185] Due to the intrinsic hydrophobicity of membrane proteins, they are manipulated with detergents to replace their natural environment.[18] Detergents present a problem for MS techniques as they give high signals, thus they usually require removal from the proteins/peptides prior to analysis.[23,33] Traditionally detergent removal methods (e.g. adsorption to polystyrene beads, trapping onto cyclodextrin) have not been suited to HDX-MS, as detergent levels remained too high and removal had to be quick to preserve labelling data. However, developments in this area (e.g. C18 guard column flush with dichloromethane) have improved HDX-MS applications to membrane proteins.[186,187] Additionally, more MS-friendly detergents have become available, and with correct sample/LC optimisation the detrimental effects of detergents on MS detection can be minimized.[186,188]

A newer development in studying membrane proteins is the use of native nanodiscs. Detergents significantly differ from the natural membrane, and this can destabilise some membrane proteins.[33] SMA based polymers can act as a 'cookie cutter' to directly solubilize the membrane, thus the sample is surrounded by a 'native' lipid environment.[114,185,189] This adds complications to the HDX-MS workflow; a clean-up step is needed to remove the nanodisc before LC separation which needs to be quick enough to reduce back-exchange. For SMALP removal, zirconia (ZrO_2) beads are added for lipid desorption, and a small amount of detergent cholate in the quench buffer added to dissemble the nanodisc before online digestion.[190]

A problem with HDX-MS experiments with membrane proteins is the sequence coverage can be quite low. Efficient digestion requires unhindered access of the protease to the sample, but this can be hampered by detergent micelles around transmembrane (TM) domains.[115] TM peptides may precipitate when in contact with aqueous solutions, and they can aggregate on LC columns. Pepsin also preferably cuts at hydrophobic amino acids, which are abundant in TM domains, thus creating very small peptides that may not be retained by the reverse-phase pre column. Improving sequence coverage is possible by taking extra steps. Optimising the quench conditions, specifically with suitable denaturants can significantly improve sequence coverage. Furthermore, the choice of chromatographic stationary phase and method for peptide separation can tuned for optimal signal/noise ratios. Digestion can be optimised by testing different proteases individually or as combinations and increasing the pressure to ~7000 psi. IM coupled with MS^E further improves peptide

detection. For SMALP samples, automation of the zirconia beads increases the efficiency of contaminant removal, which increases peptide resolution and coverage.[4,191,192]

1.4.5 Data Analysis

MS is an ensemble averaging technique where data from multiple individual peptide molecules in the gas phase are averaged to give a final mass spectral envelope (**Figure 1.19**).[33] The envelope reflects the natural abundance of heavier isotopes (e.g. ^{13}C, ^{15}N). In HDX-MS experiments deuterium incorporation increases the mass of a peptide, but this increase is not uniform across all samples of the specific peptide. This results in a mixture of peptides with varying levels of deuterium incorporation, giving the characteristic binomial distribution shape of the spectral envelope.[193] **Figure 1.19** shows that with deuterium incorporation the spectral envelope of the deuterated sample shifts to a higher *m/z* ratio; the size of the shift is dependent on the amount of deuterium incorporation. For data analysis and interpretation, the average mass for a peptide is determined through the measurement of the centroid. The centroid of a binomial distribution can be defined by the weighted mean of the distribution. **Equation 1.25** shows how to calculate the centroid of the spectral peak. The magnitude of each peak in the spectral envelope can be defined by the area under the curve.[194]

$$Centroid = \frac{\sum m_i I_i}{\sum I_i}$$

Equation 1.25. Calculation of the centroid of the spectral envelope. m_i is the *m/z* ratio of a particular spectral peak (*i*) in the spectral envelope and I_i is the magnitude (intensity) of that particular spectral peak.

Summing the *m/z* ratios and magnitudes of each peak gives us the 'area under the curve', and thus dividing this by the total magnitude of all the peaks gives the centroid (mean). Comparison of the centroids between deuterated and nondeuterated samples, after correcting for charge, allows analysis of changes within that sample (**Figure 1.19**). After centroid calculations, the relative deuterium levels can be calculated from the MS raw data (**Equation 1.26**).[136]

$$D_a = [z.(m_n - m_0)]$$

Equation 1.26. Relative deuterium level calculation. D_a is the relative deuterium level, *z* is the charge state of a particular peptide, m_n is the *m/z* ratio of the centroid of the isotopic envelope at a certain time point (*n*), m_0 is the *m/z* ratio of the centroid of the isotopic envelope for the undeuterated reference sample.

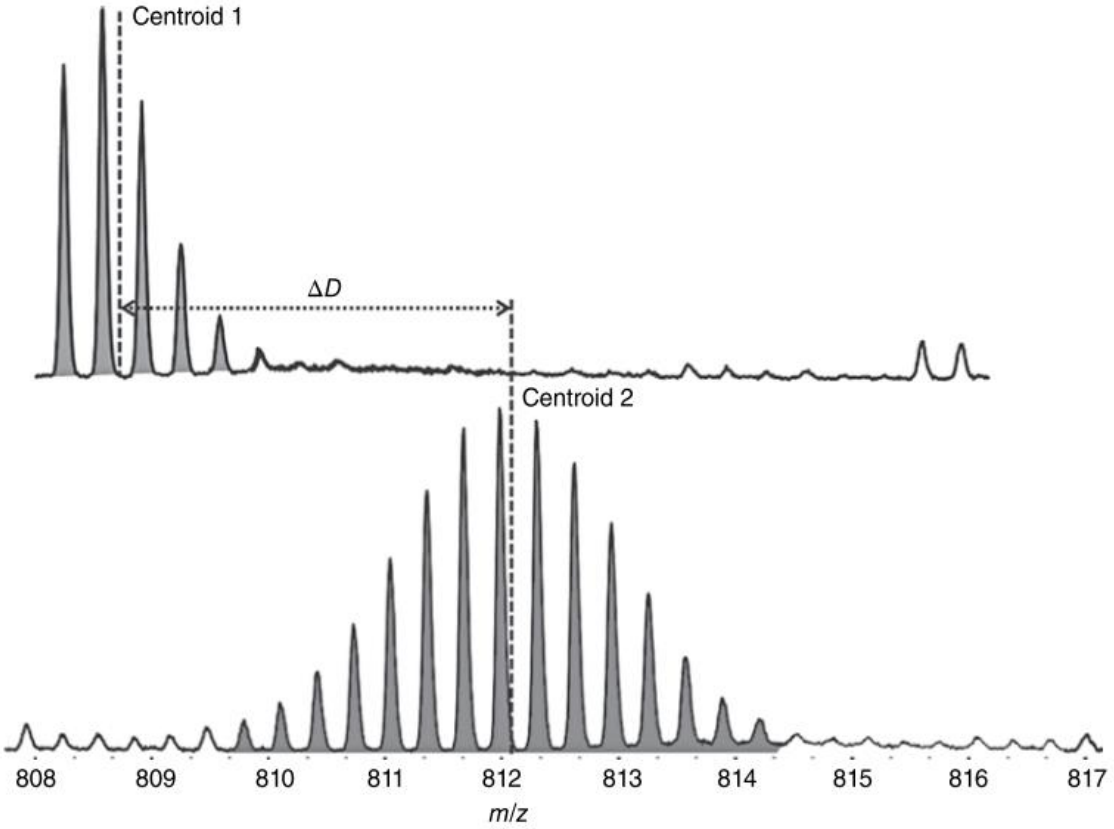


Figure 1.19. Comparing centroids for HDX analysis. Calculation of the centroid is given by **Equation 1.19**, and it is used to quantify HDX-MS data. The centroids (dotted lines) are defined by the mean of distribution; the area under the curve for each peak (grey) must be determined. When multiplied by charge, the difference between the centroids of an undeuterated sample and a deuterated sample is the average deuterium uptake for that peptide (ΔD). Taken from Weis (2016).[33]

HDX-MS data visualisation is an important part of deciphering information on conformational dynamics for a protein sample. After adjusting for back-exchange (**Equation 1.24**) and calculating relative deuterium levels (**Equation 1.26**), one of the most effective methods is peptide uptake plots, where relative deuterium level is plotted as a function of labelling time.[133,136,141] These can be of varying complexity; either showing all the peptides in one figure or only peptides of interest. For a narrow span of time scales monitoring fast exchanging amides, a linear time axis is used; for longer time scales a log scale is needed. The shape of these plots provides information on the conformational state of the peptide plotted, and hence information on the region of the protein it is native to (**Figure 1.20**).[195] Peptides with a consistently high D_a overtime represents backbone amide hydrogens that are very solvent accessible and/or in highly unstructured regions of a protein. Conversely peptides exhibiting consistently low D_a overtime often reside in areas of stable secondary structures and/or areas protected from exchange. A slow and consistent increase in D_a indicates a dynamic region that frequently visits an exchange competent state.

Uptake plots provide the opportunity for statistical tests for detect differences between conditions.[33] False positives are significant in HDX, so conservative methods for detecting changes are incorporated where possible. The high degree of amino acid redundancy resulting from most digestion protocols provides a useful approach for testing significance; in many cases groupings of peptides in the same areas of the protein yield the same trend in deuteration, so statistical outliers can be spotted.[196] Other visualisation techniques are also often used; heat maps show ΔHDX data represented on the primary amino acid sequence, which are colour coded to reflect the extent of exchange, and difference maps are utilised to highlight the effect of perturbants on dynamics.

There are many software packages available for data visualisation, data analysis and statistical analysis. DynamX is the only commercially available processing tool accessible through a mass spectrometer vendor.[197] DynamX allows the user to view peptide uptake plots through standard kinetic plots or through a spectrum view for each time point, whilst also supporting difference charts, and butterfly plots for visualisation. DynamX supports the extraction of deuteration data from LC-IM-MS analysis (IM can be utilised to resolve peak overlaps with mixed charge states). Another software called HDExaminer, a third-party software from Sierra Analytics, also supports IM data, but provides statistical analysis unlike DynamX through a data quality scoring algorithm. There are several free software available such as HDX Workbench and Mass Spec Studio.[198,199]

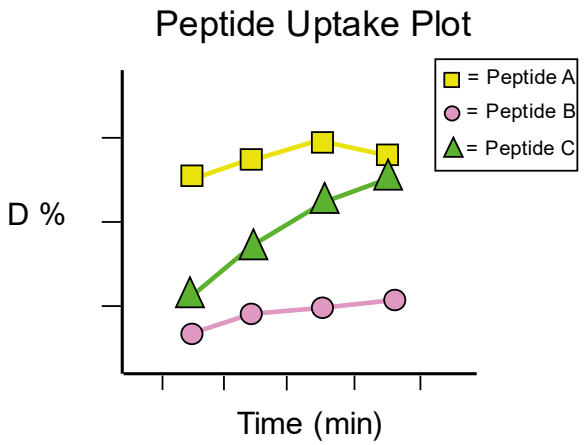


Figure 1.20. Peptide uptake plots. The Y-axis is the relative deuterium incorporation; the X-axis is time. Peptide A (yellow squares) has consistently high D% over time, so backbone amide hydrogens are likely very solvent accessible and/or reside in unstructured regions of the protein. Peptide B (pink circles) has consistently low D% over time, so backbone amide hydrogens are likely in stable secondary structures and/or are solvent inaccessible. Peptide C (green triangles) has a slow and consistent increase in D% over time, so backbone amide hydrogens are likely in a region that is dynamic and reaches an exchange competent state frequently.

1.5 Mass photometry

1.5.1 Introduction

The field of biophysics revolves around monitoring proteins and their bindings partners, which is key to understanding the cellular processes that govern life. Several existing biophysical techniques monitor proteins in different ways. Some can be classed as sized based approaches that perform size based separation or quantification by actual protein size or a diffusion coefficient (e.g. SEC or PAGE).[200] Other biophysical techniques monitor physical interactions with a functionalised surface (e.g. surface plasmon resonance (SPR)), direct mass measurements (e.g. MS), monitor enthalpy changes (e.g. isothermal titration calorimetry (ITC)) or light scattering (e.g. dynamic light scattering (DLS)).[200] Whilst these methods can reveal key information regarding protein structure and function, and many have been used throughout this thesis, they all exhibit practical shortcomings; these include the use of non-native environments, artefacts caused by protein immobilization/labelling, an inability to use low protein concentrations and low resolution.[201] A new technique, called mass photometry has been developed to measure the mass of proteins and their complexes.[201] This is a single molecule method, thus uses little amounts of protein (nanomolar), it is label free and simple, and it is possible to calculate binding affinities. Whilst mass photometry does not form a core part of this thesis, it is used in several chapters. This section aims to provide a short background into this new technology.

1.5.2 Fundamentals

Mass photometry is built on the principles of interference reflection microscopy and interferometric scattering microscopy.[202,203] It measures the mass of biomolecules by their light scattering as they bind non-specifically to a microscope cover glass surface.[201] At the glass/water interface, the binding of a biomolecule changes the refractive index, which alters the local reflectivity. By taking advantage of the interference between scattered and reflected light, the reflectivity change caused by a biomolecule binding is proportional to the molecular mass, which in turn can be calculated by using relevant calibrants (**Figure 1.21**).

1.5.3 Practical considerations

Mass photometry can tolerate most biological buffers; therefore buffer exchanging protein samples is usually not required (as it is with native MS).[204] Buffers with salts less than 10 mM should be avoided and glycerol concentrations should be kept below 5% as it may affect focusing. Furthermore, mass photometry is amenable to the study of membrane proteins. Olerinyova *et al*. (2021) have shown that membrane proteins in several mimetic environments such as detergents, amphipols, as well as MSP

nanodiscs and native nanodiscs such as SMALPs can be studied using mass photometry.[205] This is particularly useful for characterising a samples mass, purity and heterogeneity for downstream structural analysis. The low protein concentrations required for mass photometry also suit the study of membrane proteins.

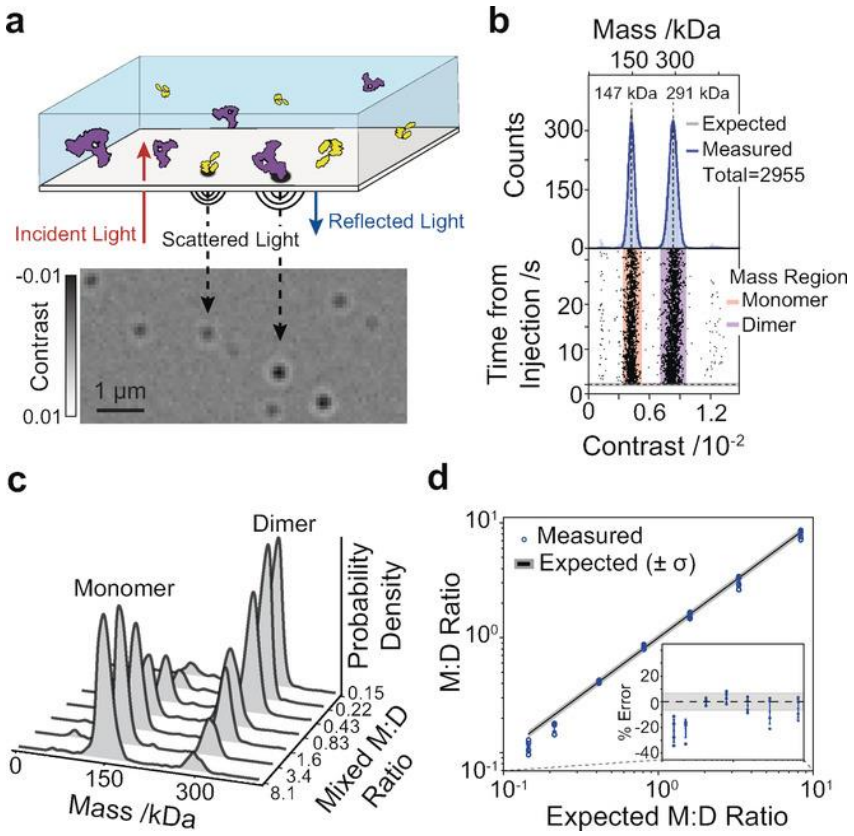


Figure 1.21. Principles of mass photometry. A. Imaging the interference of scattered and reflected light caused by the non-specific binding of a biomolecule at the glass/water interface. **B.** Scatter plot of single molecule contrasts and mass distribution of an example protein (PDB: 2G12) consisting of a 1:1 monomer:dimer mixture. **C.** Mass distributions of varying ratios of monomer and dimer. **D.** Comparisons of monomer:dimer distributions measured by mass photometry compared to the expected ratios calculated by UV-Vis spectroscopy. Taken from Soltermann *et al*. (2020).[201]

Chapter 1: Introduction

Part 2: An introduction into membrane proteins and the AcrAB-TolC multidrug efflux pump

1.6 Introduction

The rapid rise in antimicrobial resistance (AMR) is a major threat to human health, and if the situation continues at the current trajectory, current standards of living will be significantly impacted.[206,207] An increased burden to health systems, an inability to perform basic hospital procedures safely and thus a reversion to older less effective techniques, and impacts across other sectors such as agriculture are just some of the consequences of increasing AMR. The seriousness of the situation is exemplified by the Centres for Disease Control in the United States stating humanity is heading for a 'post-antibiotic' era.[207,208] In 2019, AMR caused by bacterial resistance caused more deaths than HIV and malaria combined.[209] A particular concern is the rise of a group of multidrug resistant Gram-negative and Gram-positive bacteria that are resistant to all current classes of antibiotics.[210] To focus and guide research relating to new antibiotic treatments, this group of bacteria – *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumanni*, *Pseudomonas aeruginosa*, and *Enterobacter species* – have been termed "ESKAPE" bacteria, and are referred to as this hereafter.[211,212]

Bacteria deploy several resistance mechanisms, and the acquisition of multiple resistance factors can lead to the emergence of multidrug resistant bacteria.[213] One key mechanism behind bacterial multidrug resistance is the array of multidrug efflux pumps expressed, which can effuse a wide variety of chemically diverse compounds, including antibiotics. Efflux pumps are membrane proteins found in all three domains of life, including both Gram-negative and Gram-positive bacteria.[214,215] In Gram-negative bacteria, efflux pumps can either be tripartite assemblies that span the entire cell envelope or inner membrane proteins. Efflux pumps can cooperate as part of a system that first move efflux substrates into the periplasm, then out the cell envelope through tripartite assemblies.[216] There are six superfamilies of efflux pumps in bacteria (**Figure 1.22**); the ATP binding cassette (ABC), multidrug and toxin extrusion (MATE), major facilitator superfamily (MFS), proteobacterial antimicrobial compound efflux (PACE), resistance-nodulation-cell division (RND) and small multidrug resistance (SMR).[217–222]

This thesis focuses on the AcrAB-TolC multidrug efflux pump of *Escherichia coli* (*E. coli*), a member of the RND superfamily.[223,224] AcrAB-TolC has been well characterised and it is prototypical of homologs across other ESKAPE bacteria, thus this research confers effectively to other systems.[211] This chapter first aims to introduce membrane proteins and their environment, to provide a background on their biology and biosynthesis. Then a detailed introduction on AcrAB-TolC will be provided, discussing the structure of the three proteins, their similarities and differences to other RND homologs from other bacterial species, the energetics of efflux, pump assembly and potential inhibition. This chapter will

draw attention to gaps in the current understanding of AcrAB-TolC and how this research contributes to this gap.

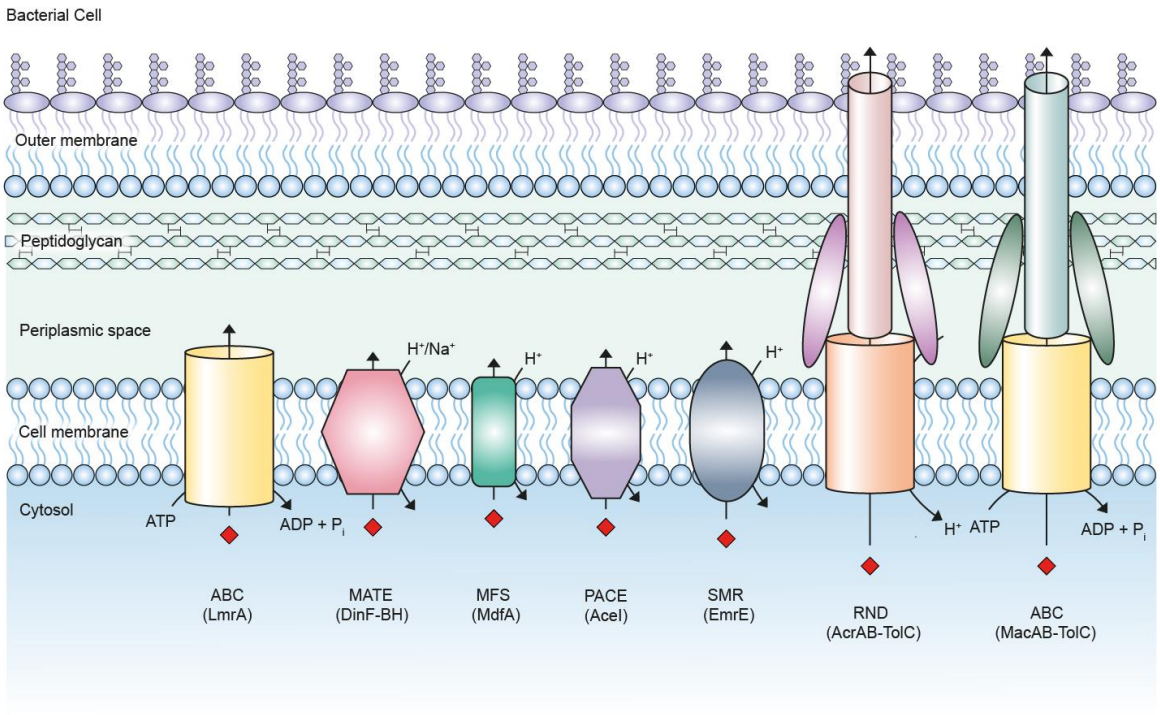


Figure 1.22. Classes of efflux pumps in bacteria. Examples of each pump are labelled in brackets. Abbreviations: ABC, ATP-binding cassette; MATE, multidrug and toxin extrusion; MFS, major facilitator superfamily; PACE, proteobacterial antimicrobial compound efflux; RND, resistance-nodulation-cell division; SMR, small multidrug resistance. Taken from Russell Lewis *et al*. (2023).[32]

1.7 Biological membranes

1.7.1 Biological membranes and their lipids and proteins

Biological membranes are dynamic structures consisting of lipids, proteins, and carbohydrates.[225] The basic feature of the membrane is a bilayer formed by phospholipids or sphingolipids. Phospholipids contain two fatty acid chains and a phosphate group, linked together by a glycerol or an alcohol. Phospholipids containing glycerol are called glycerophospholipids or they can also be sphingolipids if they contain sphingosine. Glycolipids contain either glycerol or sphingosine but have a sugar in place of the phosphate head. Lipids are amphipathic in nature, meaning they have a polar hydrophilic head group and a non-polar hydrophobic tail; this means they form fluid lamellar bilayers with the hydrophilic head facing out towards the aqueous environment and the hydrophobic tails packing inward.[226] This is thermodynamically favourable as the buried, packed hydrophobic tails are hidden from the water molecules, thus the water molecules have a higher degree of movement and the

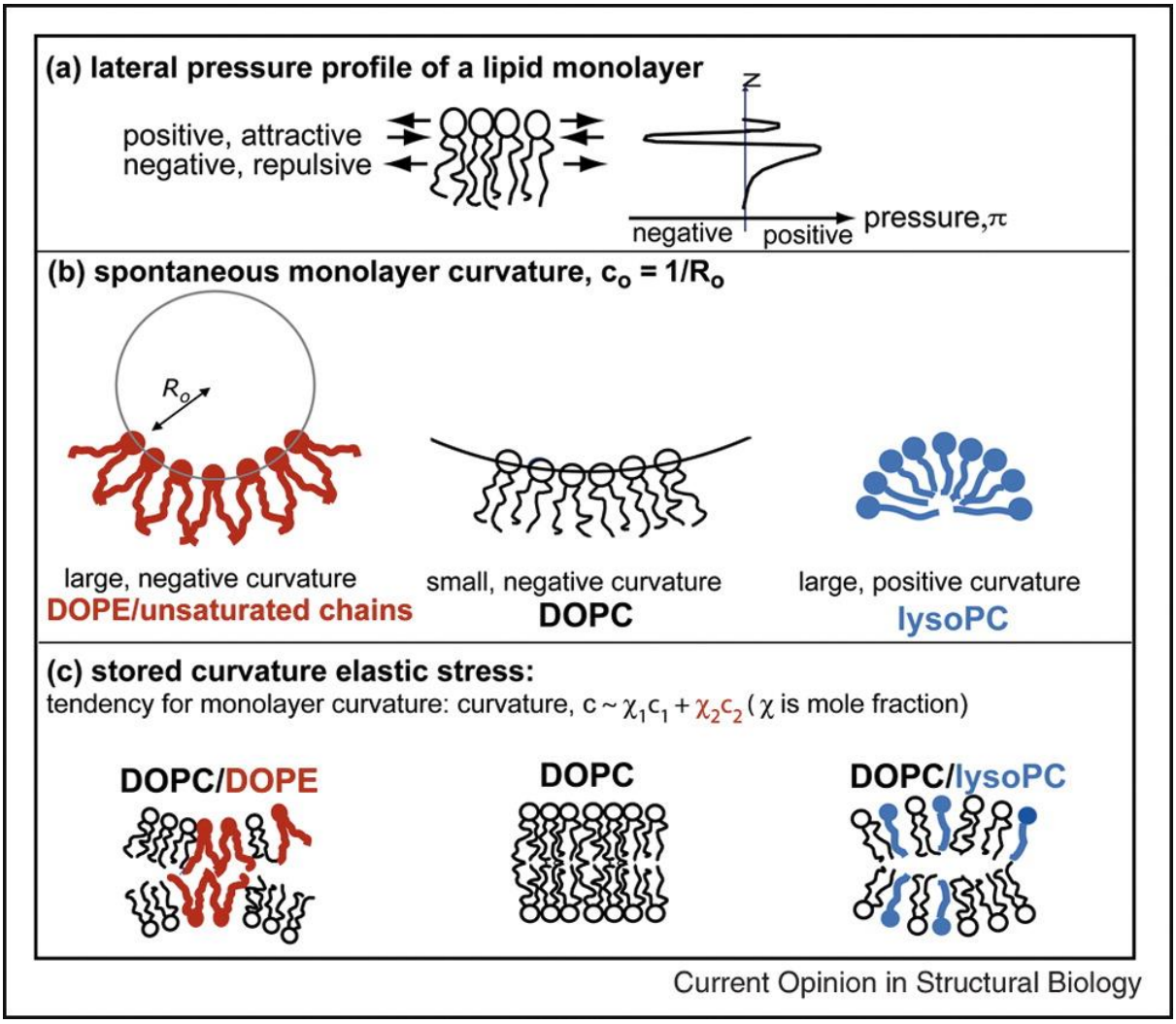


Figure 1.23. Schematic diagram of relevant lipid bilayer properties. A. Lateral pressure profile of a lipid monolayer. In the polar to non-polar interface there is a positive, attractive pressure due to the hydrophobic effect. Negative lateral pressures act in the oppose direction on the headgroup or chain regions. **B.** An imbalance of pressures causes the monolayer to curve away from or to the water. This is measured by the spontaneous monolayer curvature, C_0. **C.** This curvature cannot be satisfied in a lipid bilayer, leading to stored curvature elastic stress. Different lipids have different C_0; phosphatidylcholine (PC) lipids have negligible C_0 and form fluid, lamellar phases. Phosphatidylethanolamine (PE) lipids have a larger negative C_0 and form non-bilayer phases. Therefore, addition of PE to a PC bilayer increases the monolayer curvature and the stored elastic stress. Taken from Booth *et al.* (2009).[227]

system has increased entropy. The membrane itself provides a barrier between the organelles of a cell and the external environment.[228] However, the diversity of lipids found in biological membranes far exceeds the requirement for only barrier functions. Lipids have an array of functions within the cell; they act as membrane structural components, energy sources, signalling molecules, scaffolds for protein anchoring and protein substrates.[100,229–232] While this partly explains the diversity of lipids, it does not explain the effect this has on the membrane itself. The composition of the lipids making up

a membrane will affect its physiochemical properties.[228] The size of the lipid head group and the length of the fatty acid tail affects the curvature of the membrane. Lipids with long, saturated fatty acids increase the thickness of the membrane whilst reducing the fluidity, due to the tight packing of the hydrophobic chains. Unsaturated lipids have the opposite effect, due to the presence of a *cis* (*z*-configuration) double bond in the acyl chain forming a kink in the hydrocarbon, which prevents tight packing. Lipid-lipid interactions may assist the stabilisation of nanodomains within a bilayer (heterogeneity within a membrane).[233] **Figure 1.23** shows a schematic of relevant lipid bilayer properties.[227]

The other components of biological membranes are proteins. Half of all genes encode for membrane proteins; they make up roughly 33% of the dry weight of a cell and they are over 60% of all drug targets, due to the large variety of biological functions and importance for cellular viability.[21,234] Membrane proteins can be divided into two classes.[235] Peripheral membrane proteins are attached to the membrane; this can be through weak van der Waals interactions with lipid head groups, protein-protein interactions or through a covalent anchor (e.g. protein lipidation). Integral membrane proteins are embedded within the membrane.[225] They can be α-helical, β-barrel, or a mixture of both (**Figure 1.24**).[236] α-helical proteins have transmembrane (TM) helices which are hydrophobic structures lasting between 17-25 residues.[225,237] These proteins can range from having a single TM helix to over 20 which form a serpentine structure in the membrane.[238,239] β-barrel proteins consist of β-sheets to form a barrel-like structure. The residues alternate from outward (lipid) facing to inward (inside the barrel) facing, resulting in the residue sequence switching between hydrophobic and polar. Therefore, the β-barrel forms a polar channel through-which water soluble molecules can pass through.[235] These proteins are abundant in the outer membrane of Gram-negative bacteria (see **section 1.7.2.1**), as well as the outer membrane of mitochondria and chloroplasts.[240] Membrane proteins cover a vast array of functions within the cell. They can transport electrons or protons, ions, small molecules/metabolites, and proteins or RNA. They can be involved in chemical signalling, cell-cell adhesion and can generate electrical impulses. Membrane proteins can also have structural roles, such as controlling membrane lipid composition and maintaining the shapes of organelles or cells.[225]

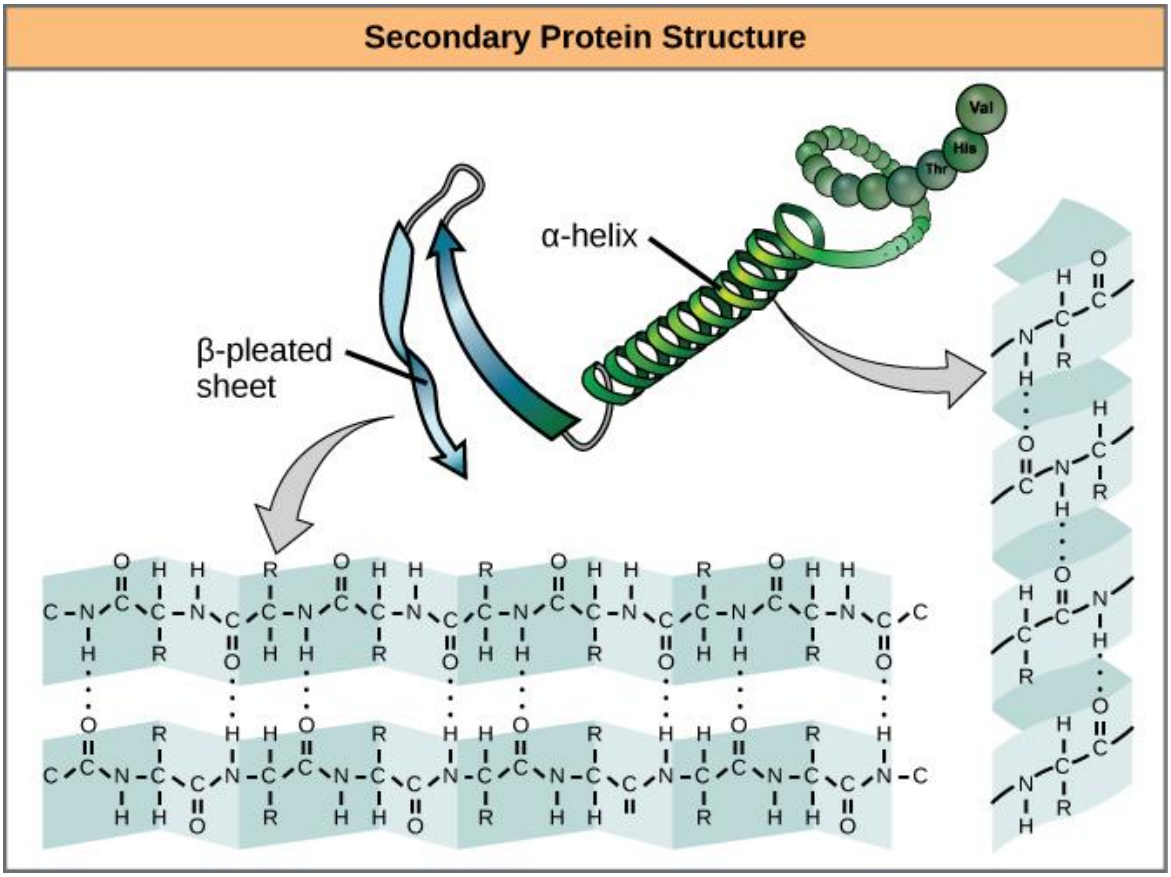


Figure 1.24. Secondary structural features of proteins. A right hand α-helix and a β-pleated sheet are common secondary structural units of proteins. The α-helix can be hydrophobic or hydrophilic and the β-sheet can twist to form a β-barrel. Taken from web page (https://teaching.ncl.ac.uk/bms/wiki/index.php/Protein).

1.7.2 Gram-negative bacterial membrane

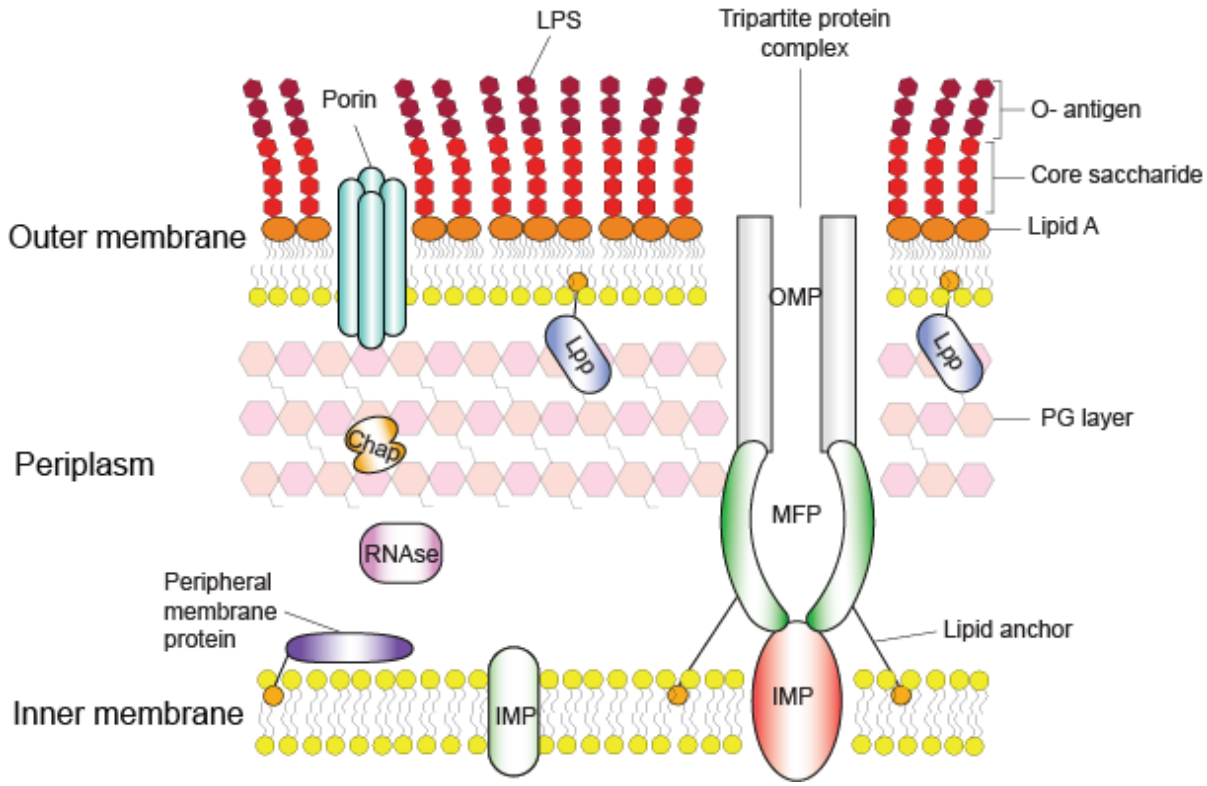


Figure 1.25. The Gram-negative cell envelope. The cell envelope of Gram-negative bacteria, showing the outer membrane, periplasm, and inner membrane. Abbreviations: LPS, lipopolysaccharide; OMP, outer membrane protein; Lpp, Brauns lipoprotein; Chap, chaperone; MFP, membrane fusion protein; PG, peptidoglycan; IMP, inner membrane protein.

Bacteria can be classed into two groups depending on the result of the Christian Gram stain; Gram-positive or Gram-negative.[241] This work predominantly uses *E. coli* which is a Gram-negative species, so only Gram-negative membrane biology will be discussed (**Figure 1.25**). The main difference between the two types of bacteria is Gram-negative have an outer membrane. The outer membrane is a lipid bilayer, but the outer leaflet contains glycolipids rather than phospholipids. The terms Gram-negative/positive originate from the Gram stain technique developed by Hans Christian Gram in 1882.[242] The test uses a crystal violet dye and a subsequent solvent, such as ethanol or acetone, to remove the dye. Initially all bacteria take up the dye, but with the use of solvent the outer membrane is dissolved and hence the stain is lost – this produces a Gram-negative result. However, Gram-positive results arise from the dehydration of the cell wall and thick peptidoglycan layer, resulting in the closure of pores and retention of the dye.

1.7.2.1 The outer membrane

The outer membrane is a glycolipid bilayer that functions as an extra permeability barrier; the glycolipid found is predominantly lipopolysaccharide (LPS).[243] LPS is critical to the function of the outer membrane; it is a heavily acylated glucosamine disaccharide with a polysaccharide core and an extended polysaccharide chain named the *O*-antigen.[244] LPS is a notorious antigen that elicits a response from the mammalian innate immune system, with the pathogenicity of *E. coli* being classified by the properties of their *O*-antigen (as well as the flagella protein, flagellin – not mentioned).[245,246] Furthermore, LPS molecules are able to bind each other in the presence of Mg^{2+}, which neutralize the negative charge from the phosphate groups. Tight packing is ensured by the acyl chains being saturated. This creates an effective barrier that prevents the passive diffusion of hydrophobic molecules such as antibiotics and detergents.[247] Proteins are also found in the outer membrane; they tend to be either lipoproteins attached to the inner portion of the outer membrane by a lipid moiety or β-barrels.[248] These integral membrane proteins are known as outer membrane proteins (OMPs). Porins are an important type of OMP (OmpF, OmpC) as they also contribute to the selective nature of the membrane. They allow passive diffusion of small hydrophilic molecules into the cell but only up to ~600 Daltons.[249] Other OMPs include enzymes, structural proteins and larger gated channels.[245,250]

1.7.2.2 The periplasm and peptidoglycan

The space between the outer membrane and the inner membrane is the periplasm. The periplasm functions as a multipurpose environment separate from the cytosol and with its own distinct environment.[251] The periplasm is often more acidic than the cytosol (by ~1.7 pH units) due to the proton gradient across the inner membrane, and contains a higher concentration of divalent cations such as Mg^{2+} (7.56 times higher).[1,3,252] The periplasm is a distinct oxidising environment compared to the cytosol that allows for efficient protein oxidation, folding and quality control.[251] The major feature of the periplasm is the peptidoglycan layer. Its main function is to maintain cellular integrity by withstanding turgor.[253] Furthermore its rigidity determines cell shape and it serves as a scaffold for other periplasmic components such as proteins.[254] The main structural features of peptidoglycan are linear glycan strands of alternating *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) sugars, connected by β-1,4 glycosidic bonds.[253,255] The lactoyl group of each MurNAc residue is covalently attached to a peptide stem – this stem is usually L-Ala-γ-D-Glu-*meso*-A_2pm (or L-Lys)-D-Ala-D-Ala. The presence of the dibasic amino acid A_2pm is essential for it to be able to form cross-links to other peptide chains. This occurs through Ala^4 to A_2pm^3 or two A_2pm^3 residues of neighbouring chains from adjacent glycans. The cross-links between glycan strands form a mesh-like structure of peptidoglycan. The outer membrane is stapled to the peptidoglycan layer through a lipoprotein called

Braun's lipoprotein (Lpp).[256] This helps to constrain the position of the peptidoglycan layer. Lpp is the most abundant protein in *E. coli* and is anchored to the outer membrane via an triacylated N-terminus.[251,257]

Aside from maintaining cellular integrity, the periplasm is densely packed with proteins for an array of functions. Due to the compartmentalization of the periplasm, bacteria can harbour potentially harmful enzymes there, such as numerous proteases, RNases and alkaline phosphatase.[245,258] Other proteins that exist in this space include membrane fusion proteins (MFPs) that form a part of large protein complexes to enable the transport of diverse substrates across the entire cell envelope, and chaperone-like proteins involved in envelope biogenesis.[259,260]

1.7.2.3 The inner membrane

The inner membrane of Gram-negative bacteria exists as a phospholipid bilayer. In *E. coli,* the principal phospholipids are phosphatidyl ethanolamine (PE) and phosphatidyl glycerol (PG), but there are smaller amounts of other lipids such as cardiolipin (CL) and phosphatidyl serine (PS).[261] The lipid composition of biological membranes is not constant, but instead it is dynamic and ever-changing in response to environmental stresses. This is crucial for bacterial survival and adaption. Furthermore, the lipid distribution within biological membranes is often asymmetrical; CL is often found at the poles of the inner membrane of the Gram-negative species *Shigella flexneri*.[262,263] The role of bacterial phospholipids determines cell integrity and function. As bacteria lack organelles, all of the membrane associated functions carried out at eukaryotic organelles occur at the inner membrane. It has the selective nutrient transport of the eukaryotic plasma membrane, protein translocation and lipid biosynthesis of the endoplasmic reticulum and oxidative phosphorylation capabilities of the mitochondrial inner membrane.[264] In contrast to the outer membrane, the integral inner membrane proteins tend to be α-helical.

1.7.3 Membrane proteins and antimicrobial resistance

Membrane proteins play essential roles in AMR and bacterial virulence. A major type of membrane protein are efflux pumps, which export antibiotics (as well as a range of other substrates) from the bacterial cell before they can enact its effect.[223] Efflux pumps are able to export antibiotics at fast rates due to bacteria's ability to dynamically increase their expression in response to an environmental stressor. Furthermore, mutations in local repressor genes, regulatory genes, promoter regions, or in insertion elements located upstream of the efflux pump can be selected for, increasing the efflux response. Many efflux pumps can also export different classes of chemically diverse antibiotics; this is

of clinical relevance as they can contribute to bacterial infections becoming untreatable with current available antibiotics.

Aside from efflux pumps, membrane proteins can be involved in biofilm formation.[265] A biofilm is an assemblage of microbial cells that are strongly associated with a surface, whilst enclosed in a matrix of self-produced, predominantly polysaccharide material.[266] Biofilms are distinct from free-living bacterial cells and have emergent properties. Biofilms can exhibit increased tolerance to antimicrobials, due to the biofilm matrix leading to increased entrapment or inactivation of the drugs.[267] Moreover, the high cell density in biofilms combined with increased genetic competence and accumulation of mobile genetic elements provides a perfect environment for horizontal gene transfer, and the spread of resistance genes.[268] An example is CmpX, which is a cytoplasmic membrane protein in *Pseudomonas aeruginosa* (*P. aeruginosa*) and a regulator of PA1611, which is part of the PA1611-RetS-GacS/A-RsmA/Y/Z pathway that controls virulence factors such as biofilm formation.[269]

OMPs also have an important role in AMR. The outer membrane acts as an additional barrier blocking the transport of many toxic compounds such as antibiotics and detergents; generally, chemicals with a molecular weight of >600 Da cannot penetrate the outer membrane.[247,270] Therefore many antibiotics with a larger molecular weight, such as vancomycin and daptomycin, cannot simply diffuse through the outer membrane. However, it does contain an abundance of porin proteins, transmembrane pore-forming β-barrels, which form a water filled open channel that allows the passage of hydrophilic molecules into the cell.[249] Since porins allow for the passive diffusion of antibiotics into the cell, they are closely associated with AMR. For example, β-lactams and fluoroquinolone antibiotics often penetrate the outer membrane through the OmpF porin, and consequently, *ompF* mutants designed to disrupt antibiotic binding in the pore, lead to β-lactam resistance in several Gram-negative pathogens due to their inability to cross the outer membrane.[249,271]

1.8 Bacterial membrane protein biogenesis

1.8.1 The SRP pathway

Membrane proteins must be produced so they can insert into the membrane and be properly folded. For inner membrane proteins, the nascent chain coming off the ribosome during translation are targeted co-translationally to the inner membrane.[272–274] This happens via the signal recognition pathway (SRP), comprising SRP and its receptor, FtsY. SRP is a ubiquitous ribonucleoprotein, consisting of one small RNA (4.5S RNA) and the Ffh protein.[275] SRP binds to the ribosome at the large subunit proteins L23 and L29, near the nascent chain exit site.[276] This overlaps with the binding sites of the

chaperone trigger factor and SecA, which have multiple functions in protein folding and translocation.[277] Ffh and part of the 4.5S RNA form a largely hydrophobic groove so SRP can interact with a hydrophobic segment of the membrane protein emerging from the ribosome.[278] The recruitment of SRP to its ribosomal binding site may be promoted by conformational changes at the nascent chain exit channel in response to the presence of any short nascent peptide.[274] The SRP protein is suggested to undergo a 'sorting' mechanism that takes place near the membrane. When nascent chain peptides emerge from the exit site, if the peptide is a signal anchor sequence, the SRP-ribosome nascent chain (RNC) complex is maintained, yet ribosomes revealing other peptide sequences are released. For example, SRP targets the first TM helix of AcrB as it exits the ribosome N-terminally; AcrB is the inner membrane protein of the AcrAB-TolC efflux pump and the focus of this work (see **section 1.10**).[279]

The SRP-RNC complex binds to the FtsY receptor on the cell membrane, to mediate the transfer of the nascent chain to the Sec-translocon. Ffh is primed for binding FtsY after binding to an RNC complex, increasing its affinity to GTP, and FtsY is primed in a GTP-bound state by interactions with lipids and Sec-translocon proteins.[277] Both Ffh and FtsY are GTPases and release of the nascent chain and dissociation of Ffh-FtsY requires GTP hydrolysis.[280] The membrane protein RNC complex is transferred to the SecYEG channel, facilitated by the direct interaction of FtsY and SecY.[281]

1.8.2 SRP independent recognition and targeting

Translocation to the inner membrane can be independent of the SRP pathway.[272,277] mRNA encoding integral membrane proteins localize to the cytoplasmic membrane; mRNA encoding integral membrane proteins have a higher uracil content than cytoplasmic proteins, suggesting a correlation between uracil content and membrane association.[282] Furthermore, traditional chaperones such as GroEL may play a role in the post-translational insertion of lactose permease (LacY) and may deliver bacteriorhodopsin to the membrane.[283,284] Secretory proteins must pass through the inner membrane, and are post-translationally targeted by the SecB chaperone to the Sec-translocon.[285] SecB keeps the pre-protein in an unfolded conformation and directs it to the motor protein SecA. SecA binds SecYEG and ATP binds SecA, to facilitate the translocation of the pre-protein across the inner membrane.

1.8.3 The Sec-translocon

Typically, bacterial membrane insertion into the inner membrane occurs through the Sec translocon. It is composed of the membraned embedded SecYEG and SecDFYajC complexes and SecA. SecYEG is a heterotrimer of evolutionary conserved integral membrane proteins, and is a key player in protein translocation, functioning as the protein-conducting channel.[277,285,286] The first high resolution crystal

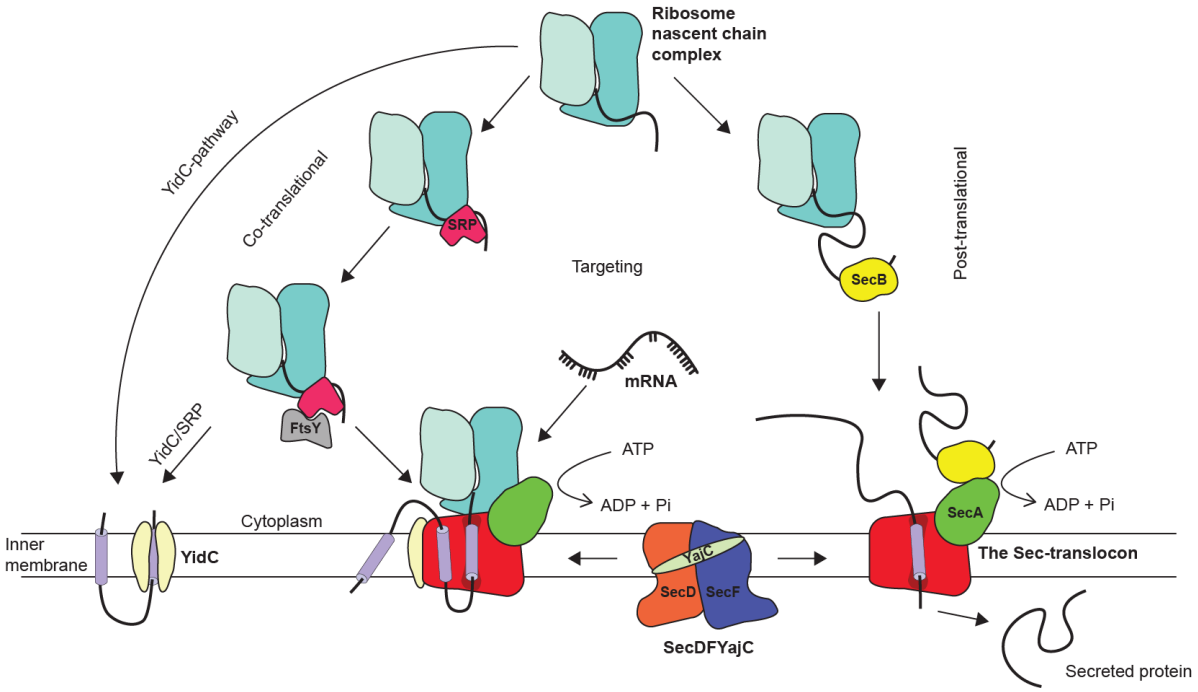


Figure 1.26. Biogenesis of inner membrane proteins in *E. coli*. For full description see text. Ribosome nascent chain complexes (RNCs) are targeted co-translationally to the inner membrane by the signal recognition pathway (SRP) (SRP protein and FtsY receptor). RNCs dock at the Sec-translocon to facilitate the translocation of hydrophilic polypeptide chains and the insertion of transmembrane helices (TMs) into the lipid bilayer. Translocation of sizeable periplasmic loops requires the ATPase SecA. YidC insertase can assist the Sec-translocon in protein translocation and insertion. The SecDFYajC complex can help the Sec-translocon with membrane proteins and secreted proteins. Some membrane proteins are targeted directly to YidC via the SRP pathway. mRNAs encoding membrane proteins can localize to the inner membrane. Secretory proteins are targeted to the inner membrane post-translationally via SecB; the translocation of secreted proteins is SecA dependant.

structure of the trimeric Sec-translocon complex from the archaeon *Methanococcus jannaschii* was essential in understanding Sec mediated translocation/insertion.[287] SecY has TMs which form the actual channel, and a small α-helical domain which plugs the channel. TM helices 1-5 and TM helices 6-10 are pseudo-symmetrically aligned to form a 'clamshell' conformation. This conformation causes an hourglass shape to the central channel, with six hydrophobic residues in the middle of the membrane likely forming a seal to prevent leakage of water or ions.[288,289] TMs 2 and 7 are thought to form a flexible gate to allow the release of TM helices during co-translational insertion of integral membrane proteins. *E.coli* SecE contains three TMs, which enwrap SecY in a V-shaped manner. Only the C-terminal TM helix of SecE is required for a functional Sec translocon.[290] SecE is thought to stabilise the SecY complex.[291] SecG has two TMs, with its N- and C-termini in the periplasm. Its located more peripherally than SecE, and whilst it is not essential for translocation or cell viability, it facilitates Sec-translocon function.[285,292–294]

The additional SEC components such as SecA, SecD, SecF and YajC are not discussed further in this thesis.[292] Furthermore, YidC, an insertase that can work dependently or independently of the Sec-translocon, is not discussed either.[295] For more information regarding the importance and functions of additional Sec components or the YidC insertase, please see the following review by Luirink *et al.* (2012).[277]

1.8.4 Outer membrane protein biogenesis

Like all bacterial proteins, OMPs are synthesized by the ribosome in the cytosol. They are delivered to the Sec translocase by one of the mechanisms in **sections 1.8.1-1.8.2**.[296] OMPs are targeted to the Sec translocase by an N-terminal signal peptide, which is then subsequently cleaved by periplasmic proteases.[297] They traverse the inner membrane in an unfolded state through the Sec-translocon, and as they emerge from the periplasmic exit they interact with periplasmic chaperones to avoid misfolding and aggregation in the aqueous environment.[298] Three classes of periplasmic chaperones have been identified; first are disulphide bond catalysts which are essential for forming disulphide bonds in the oxidising periplasmic environment.[299] Second are peptidyl-prolyl *cis-trans* isomerases which catalyse the *cis-trans* isomerisation of proline peptide bonds. Lastly are general chaperones, such as SurA, Skp and DegP.[298] SurA has been shown to have general chaperone activity, and it is the primary chaperone escorting the bulk of OMPs across the periplasm.[300] Skp is another general chaperone which can trimerize, and bind OMPs in their internal cavity, shielding them from the aqueous periplasmic environment, preventing their aggregation.[301] However, DegP has both protease and chaperone activity, but primarily functions as a protease to degrade misfolded or aggregated OMPs.[302] It is important to note that the shuttling and quality control pathways in the periplasm are less well understood compared to the pathways in the cytosol.[298]

The OMPs are delivered by the soluble periplasmic chaperones to the barrel assembly machinery (BAM) complex, which folds and inserts the OMPs into the outer membrane.[303] In *E. coli* the 200 kDa BAM complex is composed of the single β-barrel protein BamA and up to four lipoproteins (BamB-E) that bind to the polypeptide transport-associated domains of BamA.[296,303,304] β-barrels are renowned for their structural stability, but BamA has the unusual ability to open laterally.[305] It is thought the β-barrel of BamA opens laterally to allow for the OMPs to be inserted into the outer membrane.[306] However, although the structure of the BAM complex has been solved, the mechanism and energetics of insertion by the BAM complex is still unknown.[296,303]

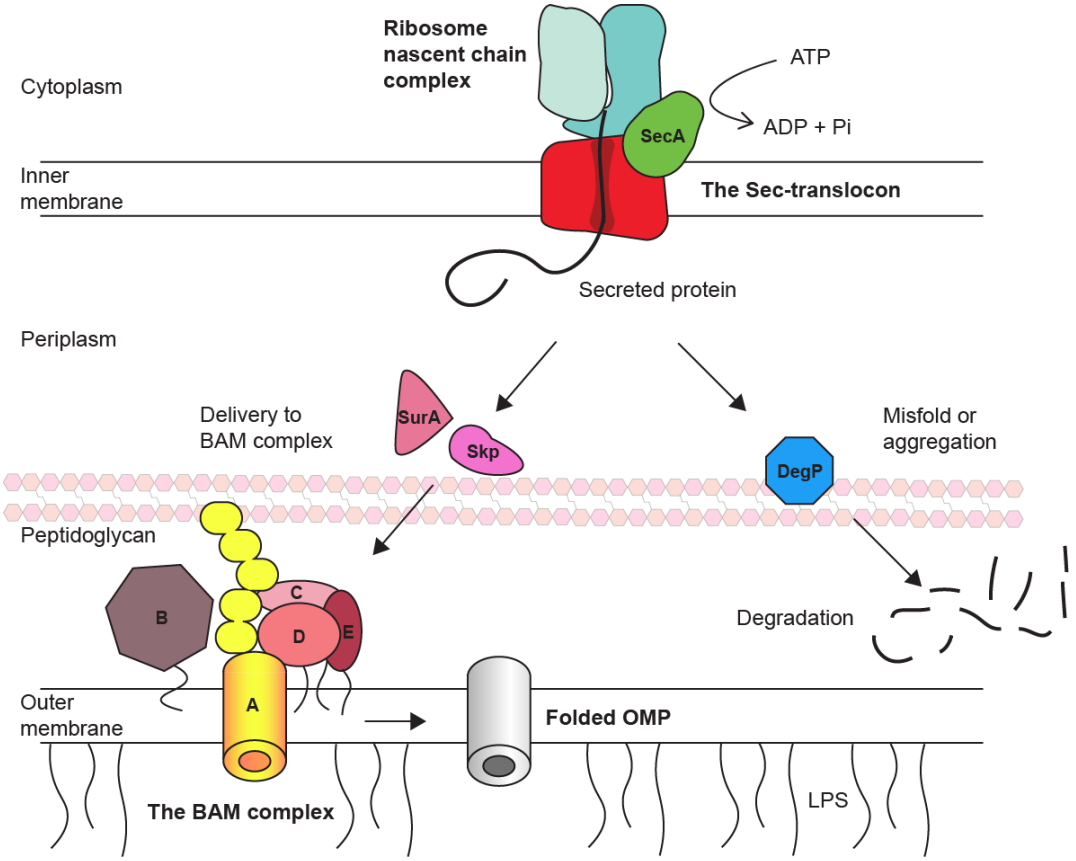


Figure 1.27. Biogenesis of outer membrane proteins in *E. coli*. For a full description see the main text. Outer membrane proteins (OMPs) are synthesized in the cytosol by the ribosome and pass through the inner membrane via the Sec-translocon. Once in the periplasm, OMPs are delivered to the barrel assembly machinery (BAM) complex by periplasmic chaperones such as SurA and Skp. Misfolded or aggregated OMPs are degraded by DegP. The BAM complex ensures correct OMP folding and insertion into the outer membrane.

1.9 Antimicrobial resistance in bacteria

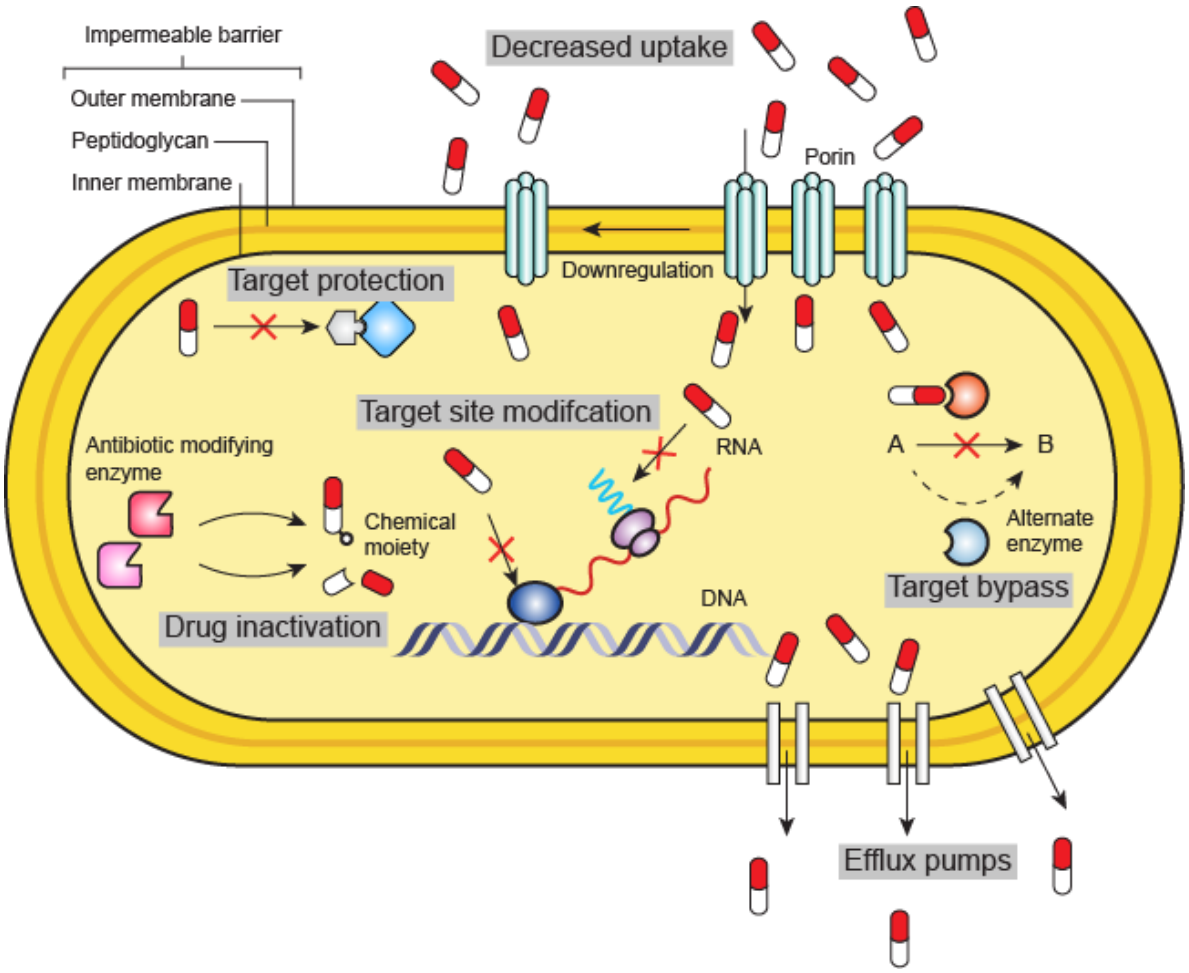


Figure 1.28. Bacterial mechanisms for drug resistance. Drug inactivation involves enzymes degrading the antibiotic or modifying antibiotics by adding a chemical moiety. Target site modification occurs by mutations to the antibiotic target protein, which effects the antibiotics affinity for the target. Efflux pumps are overexpressed when exposed to antibiotic and remove the antibiotic from the cellular environment. Target bypass involves the production of an alternate enzyme that undertakes similar biochemical activity but isn't inhibited by the antibiotic. Target protection involves the binding of a target protection protein preventing inhibition by the antibiotic. Decreased uptake often occurs due to changes in the membrane structure, such as downregulation of porins.

Antibiotics are molecules which are either bactericidal or bacteriostatic and can therefore be used to treat infection by specifically killing bacteria or inhibiting their growth. Antibiotics are divided into different classes depending on their mechanism of action.[307] The main five groups inhibit cell wall synthesis (e.g. β lactams), depolarize the cell membrane (e.g. daptomycin), inhibit protein synthesis (e.g. aminoglycosides and macrolides), inhibit nucleic acid synthesis (e.g. quinolones) or inhibit metabolic pathways (e.g. platensimycin). Antibiotics are commonly produced by microorganisms, to compete with other species in complex environments.[308] However, it is only natural that microorganisms have some intrinsic resistance to these antibiotic molecules. Intrinsic resistance is a trait shared universally within a bacterial species and is independent of previous antibiotic exposure and horizontal gene transfer.[307] Problematically, resistance can also be acquired through horizontal gene transfer (transformation, transposition or conjugation) or via random mutations to its own chromosomal DNA.[307] Therefore the use of antibiotics as therapeutics comes with a catch; the use of these drugs can lead to the selection of high-level resistance in successive bacterial generations, the selection of hypermutable strains, and may increase the ability for bacteria to acquire new resistance to different antimicrobial agents. For example, after the widespread use of penicillin in the 1940's, it only took several years for resistance to become a substantial clinical crisis.[309] Furthermore, human action has to led to the significant increase in resistance resulting in the growing problem seen today. Whilst even the proper use of antibiotics increases the prevalence of bacterial drug resistance, this is exacerbated by the overuse of antimicrobials by physicians and the improper use of antibiotics by patients leading to an excessive use of antibiotics in humans.[310] Another major reason is the widespread use of antimicrobials in the food industry. Animal feed often contains antibiotics (ranging from below therapeutic levels to full therapeutic levels) to treat and prevent infection whilst raising food animals. There is evidence suggesting this leads to an increase in antibiotic resistant microorganisms that can be transferred to humans during consumption.[311,312]

Bacteria can exhibit resistance through a multitude of mechanisms (**Figure 1.28**). The key mechanisms are discussed below.

1.9.1 Limiting uptake

Most antibiotics need to cross the bacterial cell envelope to exert their activity.[313] This provides Gram-negative bacteria a natural advantage, due to the extra barrier provided by the LPS layer (see **section 1.7.2.1**). For example, the glycopeptide vancomycin is inactive against Gram-negative bacteria due to its inability to cross the outer membrane.[314] Hydrophilic molecules such as β-lactam antibiotics rely on porins to diffuse across the outer membrane.[315] Therefore, alterations to porins affect the permeability of the outer membrane; this may affect the penetration of some antibiotic compounds.

Bacteria achieve this by three general mechanisms: a shift in the type of porins expressed, changing porin expression level and impairment of porin function.

Gram-positive bacteria lack the outer membrane and are therefore more permeable to many antibiotics. Some bacteria lack a cell wall, such as mycoplasma, consequently making them intrinsically resistant to β-lactams and glycopeptides.[316] Gram-positive bacteria can change the composition of the cytosolic membrane to affect membrane fluidity, which has been shown to affect its permeability to certain antibiotics.[317] Furthermore, some species of bacteria such as mycobacteria, produce a unique outer cell envelope containing a lipid barrier of mycolic acids and a capsule-like coat of polysaccharides and proteins, therefore can be considered to lie somewhere between Gram-positive and Gram-negative.[318,319] Another widely seen phenomenon is the formation of biofilms by a bacterial community.[307] The biofilm matrix contains polysaccharides, proteins and DNA from resident bacteria to create a thick, sticky environment, which makes it very difficult for antibiotics to reach their target.

1.9.2 Antibiotic modification

One strategy bacteria have is modification of the antibiotic molecule. The drugs can be inactivated in two ways, either by degradation of the drug, or transfer of a chemical group to the molecule. One of the most widely used groups of antibiotics are β-lactams. They contain a four-sided β-lactam ring, and they interrupt bacterial cell wall formation by covalently binding to penicillin-binding proteins involved in peptidoglycan cross-linking.[320] The main mechanism of resistance against these antibiotics is their destruction by β-lactamases.[307,313,314] They hydrolyse a specific site in the β-lactam ring, causing it to open and affecting their ability to bind penicillin-binding proteins. It is the most important resistance mechanism against penicillin and cephalosporin drugs.[321] Another example is the inactivation of tetracyclines by tetracycline destructases.[322] Moreover, drug modifying enzymes have been identified for a range of antibiotics such as aminoglycosides, macrolides, rifamycins, streptogramins, lincosamides and phenicols. For example, aminoglycosides can be modified by *N*-acetyl transferases, *O*-phosphotransferases, and *O*-adenyltransferases that acetylate, phosphorylate or adenylate respectively.[323] The modification of the hydroxyl or amino groups of the drug, reduces its affinity to its target.[324]

1.9.3 Target modification/bypass

Essential to the selective function of most antibiotics is their high specificity for bacterial cellular targets. Therefore, modification to a drug target or protection by chemical moieties can interfere with the drugs affinity to the target, conferring resistance. One example of this is fluoroquinolone resistance. These inhibit the DNA gyrase topoisomerase enzymes involved in supercoiling DNA.

However, point mutations S83 and D87 of DNA Gyrase weaken fluoroquinolones affinity for the enzyme, whilst retaining function.[325] Mutations in the genes encoding for penicillin-binding proteins can also confer resistance to β-lactam antibiotics. Moreover, rather than modification of the drug targets, the addition of moieties can prevent antibiotic function. This is well documented in macrolide resistance. The 16S rRNA target can be methylated by ribosomal methyltransferases, which prevents the binding of a macrolide molecule.[326] Some resistance mechanisms do not require any target modification, instead they are able to ensure target bypass by evolving new targets that carry out similar biochemical functions but are not inhibited by the antibiotic.[314] The best-known example of this is in methicillin resistant *Staphylococcus aureus* (MRSA). *S. aureus* can acquire the *mecA* gene, which encodes an exogenous penicillin binding protein that is homologous to the original but with a lower affinity for β-lactam antibiotics.[327] Target bypass can also be achieved by increasing the expression of the antimicrobial target with the aim of overloading the drug with too many targets. One example of this method is resistance to trimethoprimsulfamethoxazole (TMP-SMX). TMP-SMX alters the production of folate, which effects the biosynthesis of purines and some amino acids. Therefore, bacteria can overproduce two enzymes critical for folate synthesis to bypass the effect of the antibiotic.[328]

1.9.4 Efflux pumps

Efflux of antibiotics is a common mechanism of antibiotic resistance and the main focus of this thesis. Efflux pumps are found in all domains of life, and they are responsible for bacterial resistance against a range of structurally diverse compounds.[215,329] Some are expressed constitutively whilst others are overexpressed in response to antibiotic exposure; some high-level resistant strains carry mutations that modify the transporter.[307] Even though these pumps are found in both Gram-positive and Gram-negative bacteria, they are particularly important in resistance in the latter class.[313] Efflux pumps combined with the impermeable double membrane makes these species intrinsically resistant to multiple classes of antibiotics, raising a significant public health problem. A recent study at a hospital in Greece saw that patients with multidrug resistant Gram-negative bacteraemia exhibited a 50.3% mortality rate.[330]

There are six superfamilies of multidrug efflux pumps, classed by their composition, number of TM regions, energy sources and substrates (**Figure 1.22**).[329] These are the ABC, MATE, MFS, PACE, SMR and the RND superfamilies.[217–222] It is members of the RND superfamily that confer the most clinically relevant levels of resistance in Gram-negative bacteria.[331] The AcrAB-TolC multidrug efflux pump is the RND efflux pump native to *E. coli,* and the focus of the work presented in this thesis.[223,224] Below,

the three proteins that make up this tripartite system are introduced, and the relevant literature is discussed.

1.10 The AcrAB-TolC multidrug efflux pump

1.10.1 The RND superfamily

RND efflux pumps were first described by Ma *et al*. (1993) in *E. coli* and Poole *et al*. (1993) in *P. aeruginosa*.[332,333] These complexes were AcrAB-TolC and MexAB-OprM respectively. In Gram-negative bacteria, RND efflux pumps from the hydrophobe/amphiphile efflux-1 (HAE-1) family contribute to the intrinsic resistance towards antibiotics and has the largest number of known RND transporters.[215,334,335] Alongside the heavy metal efflux (HME) family, they are only active as tripartite assemblies consisting of an RND inner membrane protein, a periplasmic MFP and an outer membrane factor (OMF) protein.[336,337] The three proteins form a continuous, sealed channel across the Gram-negative cell envelope. The RND protein has been suggested to transport substrates from the outer leaflet of the inner membrane or the periplasm.[334] This process is powered by the proton motive force. All RND proteins have a conserved TM region (most RND proteins have 12 TMs), as well as differing periplasmic and luminal domains that comprise 60% of the protein.[338] The RND proteins are large in comparison to other bacterial proteins and tend to be homotrimers.[216]

Originally, it was assumed RND efflux pumps evolved because of selection pressures due to the presence of antibiotics, however it has been shown they are part of an ancient family of proteins with homologs across all three domains of life.[215] Thus it has been hypothesised RND pumps have important physiological roles alongside drug resistance. This is supported by the redundancy in antibiotic specificity across multiple pumps within the same species and the ability to efflux non-antibiotic substrates such as dyes, detergents, disinfectants and fatty acids.[339–341] It is thought efflux pumps may impact virulence, through roles in colonialization, phagocytosis and biofilms, as well as the general efflux of toxic bacterial metabolites.[336]

Section 1.10 will introduce the roles of RND pumps in virulence, before specifically introducing the AcrAB-TolC efflux pump from *E. coli*. Then, the regulation of AcrAB-TolC expression will be introduced, followed by an overview of the structure of the assembled complex. Then there will be a deeper dive into the structure and chemistry of the individual components, how they compare to other members of the family, and finally a more detailed insight into the interactions and assembly of the complex.

1.10.1.1 RND pump's role in colonization

There have been several studies that suggest RND pumps have a role in host colonization, possibly due to their ability to effuse host antimicrobials and bacterial toxins.[336] Studies concerning RND pumps of *P. aeruginosa* provided early evidence of their role in colonization; firstly, the MexAB-OprM, MexCD-OprJ and MexEF-OprN pumps were rapidly isolated in a rat pneumonia infection model, even without the presence of antibiotics.[342] Furthermore, the MexAB-OprM RND pump was shown to be essential for the invasion of Madine-Darby canine kidney cells lines, and the MuxABC-OpmB RND pump has been shown to increase the twitching motility of bacteria.[343,344] Interestingly, over expression of MexCD-OprJ or MexEF-OprN in *P. aeruginosa* can lead to downregulation of type III secretion proteins, likely reducing its virulence; clinical isolates have also shown that these RND pumps can reduce *exsA* expression, which encodes a regulator protein which activates the expression of various genes in the type III secretion system.[345] This shows that different RND pumps may have a different impact on virulence in *P. aeruginosa*.[336]

Furthermore, four RND pumps have been shown to influence colonization in *Vibrio cholerae* (VexAB, VexCD, VexIJK, and VexGH), and the same group of RND pumps have been shown to be required for the optimal production of the cholera toxin and the toxin regulated pili, possibly due to the efflux of effector molecules that usually suppress the transcription of the toxin genes.[346,347] There has been further evidence of RND roles in colonization across other species of bacteria such as *Salmonella* Typhimurium and the plant pathogen *Pseudomonas syringae*.[348–350]

1.10.1.2 RND pumps and phagocytosis

When a bacterial pathogen infects a host cell, it encounters the innate immune response; one aspect of this is phagocytosis. Reactive oxygen species (ROS) and reactive nitrogen species (RNS) are important molecules created by the host phagocytes in the defence response.[336] Several studies have suggested that ROS and RNS compounds can cause the overexpression of RND pumps, implying they play an important role in bacterial defence against these host species. For example, SoxRS is a two-component regulatory system that responds to oxidative stress as part of the pathogenic defence mechanism against the host cell. SoxR is the sensor protein that responds to superoxide and nitric oxide ions by activating the expression of *soxS*. SoxS then activates genes that protect against oxidative damage.[351] Several studies have shown SoxS can result in the upregulation of the AcrAB pump across multiple species of bacteria, suggesting oxidative stress and the upregulation of AcrAB is widespread.[336]

1.10.1.3 RND pump's role in biofilms

As mentioned in **section 1.7.3**, a biofilm is an assemblage of microbial cells that are strongly associated with a surface, whilst enclosed in a matrix of self-produced predominantly polysaccharide material.[266] Biofilms are distinct from free-living bacterial cells, with their own unique properties and they are clinically very important due to their persistence and resistance to antibiotics. Efflux pumps appear to have several roles in biofilm formation and fitness. The earliest studies monitored gene expression during biofilm expression in *E. coli*. It was found that the efflux gene *mdtF* was expressed at significantly higher levels during biofilm formation compared to usual growth or stationary phase.[352] Furthermore, MdtEF-TolC (which is an homolog of the AcrAB-TolC efflux pump) had a greater expression under anaerobic conditions, and it has a role in exporting nitrosyl indole derivatives.[353] In the core of biofilms, anaerobic conditions are common as the cells on the outer regions respire the available oxygen; this may force cells in the core to switch to anaerobic respiration, and the overexpression of MdtEF-TolC may protect cells from damage.[354] The role of efflux pumps in biofilm formation has been further investigated in range of other bacterial species and efflux pumps, with the proposed roles shown in **Figure 1.29**. For more information on efflux pumps and biofilms, please see this great review by Alav *et al*. (2018).[354]

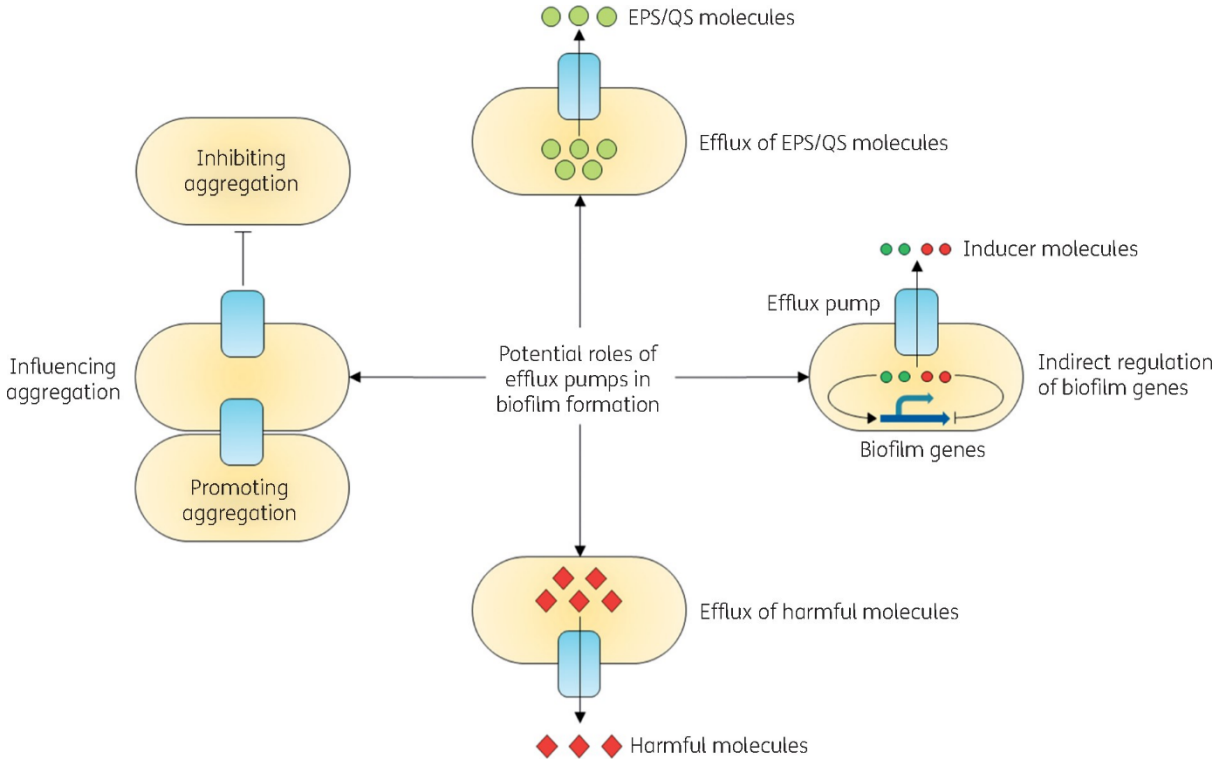


Figure 1.29. Possible roles of efflux pumps in biofilm formation. Taken from Alav *et al*.(2018).[354]

1.10.2 The overexpression of AcrAB-TolC

RND pump expression is tightly mediated by local and global regulators. As mentioned in **section 1.9.4**, a big cause of multidrug resistance in bacteria is the overexpression of efflux pumps in response to antibiotic exposure. Overexpression can occur via multiple mechanisms. At the lowest level, local repression occurs by AcrR in *E. coli*, *Salmonella Spp* and *Klebsiella pneumoniae,* which acts as a modulator to prevent overexpression of *acrAB*.[355,356] AcrR is a member of the TetR family and thus contains a predicted DNA binding helix-turn-helix (HTH) motif. It likely represses *acrAB* transcription by directly binding to the 24-base pair palindrome sequence in the *acrAB* promoter region.[357] Other local regulators of AcrAB-TolC include AcrS/EnvR, the histone-like nucleoid structuring protein (H-NS) and the suppressor of division inhibition SdiA.[358] These regulators are only thought to play a minor role in the regulation of *acrAB* and *tolC*.

In *E. coli* the *acrAB-tolC* genes are regulated by the multiple antibiotic resistance operon, Mar. MarR is a protein that blocks transcription of *marRAB* in the absence of any environmental signal by binding to two palindromic sequences within the operator DNA sequence *marO* that contains its promoter.[359] However, derepression of the Mar operon leads to the expression of *marA*, encoding MarA of the AraC/XyIS family of proteins.[355] Depression can be due to the presence of certain ligands such as phenolic compounds that bind MarR, antibiotics, oxidative stress, or mutations of *marR*, *marO,* or MarR binding site.[360] MarA contains a 100 amino acid domain with two HTH motifs that bind DNA; it binds to a sequence of DNA called the marbox, which is highly degenerate and asymmetric, with over 10,000 copies in the *E. coli* chromosome.[355] MarA is able to bind to two major DNA grooves using its two HTH motifs and bends the DNA by 35˚.[361] This results in the activation of many genes including *acrAB* and *tolC* that contain upstream marboxes. [362]

A protein called MarB also acts to increase MarA expression via an unknown mechanism. Furthermore, MarA homologues SoxS and Rob (which also belong to the AraC family) can also activate transcription of *marRAB* and *acrAB-tolC.*[363,364] The expression of these three transcription factors is influenced by a range of specific environmental stimuli to ensure appropriate AcrAB-TolC regulation from a variety of stress signals. Overexpression of these transcription factors have been found in many antibiotic resistant isolates, which often harbour multiple mutations.[355] In *Salmonella enterica acrAB-tolC* is primarily regulated by the MarA homolog, RamA of the AraC/XyIS family.[365] With high RamA expression, overexpression of the AcrAB-TolC proteins can contribute to multidrug resistance.

1.10.3 Structure of AcrAB-TolC

The focus of this thesis is the AcrAB-TolC multidrug efflux pump from *E. coli*. AcrAB-TolC is well characterised, and is prototypical of homologous RND pumps across other ESKAPE bacteria, making it an excellent system to study and findings translate well across the field.[216,223,366,367] For example, homologs include the MexAB-OprM, MexCD-OprJ and MexXY-OprM pumps in *Pseudomonas ssp* and the AdeABC pump in *Acinetobacter baumanni* (*A. baumanni*).[355] The RND pumps are tripartite systems, with an inner membrane RND protein (AcrB), which undergoes a rotational efflux mechanism powered by the proton motive force, a periplasmic MFP (AcrA) and an OMF protein (TolC).[366] AcrAB-TolC confers resistance to a range of structurally diverse antimicrobial compounds, such as tetracycline, chloramphenicol, β-lactams, novobiocin, fusidic acid, nalidixic acid and fluoroquinolones.[368] Furthermore, it can efflux sodium dodecylsulfate (SDS), Triton X-100, detergent-like bile salts, cationic dyes, disinfectants and non-polar solvents.[339–341] Over a number of years, there has been debate over the stoichiometry of the assembled complex. Originally, a 3:3:3 arrangement was suggested, in the adapter wrapping model; this hypothesised a tip to tip interaction between AcrB and TolC, with three AcrA protomers wrapping the outer portions of the AcrB:TolC complex.[369,370] However, the cryo-electron microscopy (cryo-EM) structure of the AcrAB-TolC pump by Du *et al.* (2014) coupled with the negative stain EM structure reported by Kim *et al.* (2015) provided evidence for a 3:6:3 model, with an AcrA trimer of dimers.[366,371–373] Therefore, the literature supports an adapter bridging model where there is no direct interaction between AcrB and TolC in the assembled complex. Instead, they are bridged in the periplasm entirely by AcrA (**Figure 1.30**).

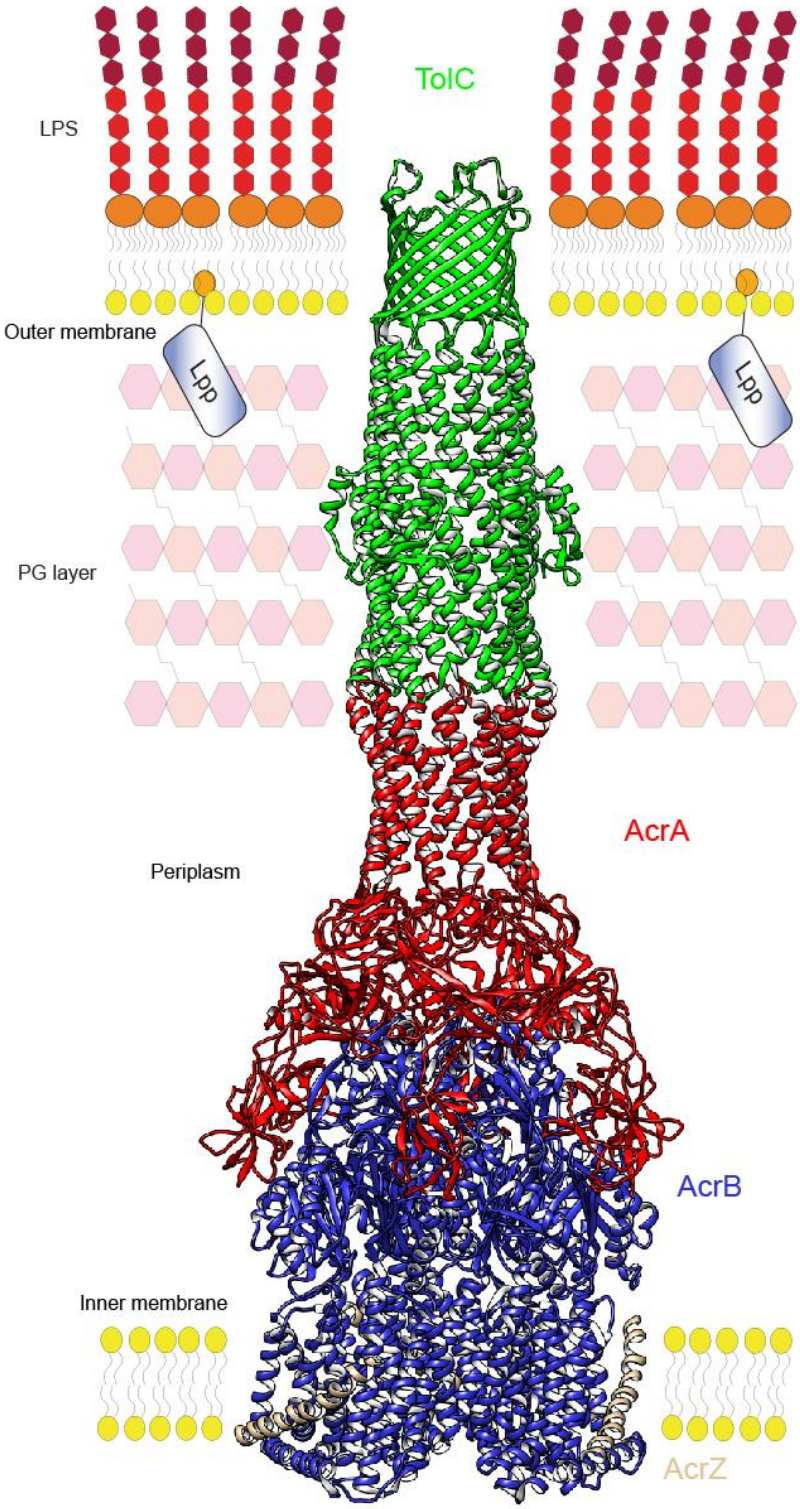


Figure 1.30. Structure of AcrAB-TolC. AcrB in blue, AcrA in red, TolC in green, AcrZ in gold. PDB 5O66.

1.10.4 AcrB

AcrB is a 1049 amino acid inner membrane RND protein that exists as a homotrimer (**Figure 1.31**).[374] It has 12 TM-spanning helices (TM1-12), which contains two large, characteristic periplasmic portions between TM1-2 and TM7-8. These are a typical feature of RND proteins, consisting of over 300 amino acids each. These periplasmic sections contain the funnel and porter domains. The funnel domain is made up of two subdomains, DN and DC, whereas the porter domain is located nearer to the inner membrane and is made up of four subdomains, PN1, PN2, PC1 and PC2.[371] PN1/2 are located in the N-terminal half of AcrB between TM helices 1-2 and PC1/2 are located in the C-terminal half between TM helices 7-8. PN1-PC2 and PN2-PC1 both pack together to form β-sheet structures.[375] The PC1/2 subdomains form a cleft at the bottom of the periplasmic section to form the proximal (access) binding pocket. The distal (deep) binding pocket is located in the PN2-PC1 unit. PN1/2 form the drug exit pathway that leads to the central cavity.

AcrB contains multiple substrate entry channels leading to the binding pockets.[334,376] Channel 1 (CH1) is located in the TM/PC2 interface above the TM8/TM9 groove, and it is open to the outer leaflet of the inner membrane.[334,377] It is open in the L and T state monomers (see **section 1.10.4.2**), but is closed in the O state due to the conformational movements of a loop between TM8 and the PC2 subdomain. CH2 is located between the PC1 and PC2 subdomains and is open to the periplasm in the L and T state. Both CH1 and CH2 are connected to the proximal binding pocket. CH3 is open from the central cavity of the AcrB trimer and leads directly to the distal binding pocket, without passing the proximal binding pocket or the switch-loop.[376] A recently proposed CH4 is located in the groove between TM1/TM2 and also leads to the distal binding pocket.[378,379] The different channels contribute to the broad substrate specificity of AcrB, as drugs with certain physiochemical properties favour different channels. Low molecular mass drugs ($M < 500$ g mol^{-1}) such as chloramphenicol and minocycline have been proposed to prefer CH1 whereas high molecular mass drugs ($M > 500$ g mol^{-1}) such as novobiocin and erythromycin are suggested to utilise CH2.[379] On the other hand, planar aromatic compounds such as ethidium bromide and rhodamine 6G are preferentially taken up by CH3.

Aside from the substrate entry channels, AcrB has several other structural features that aid its function. AcrB's binding pockets are lined with aromatic, polar residues, adept at interacting with large substrates, termed the hydrophobic trap. The binding pockets are separated by an 11 amino acid switch loop containing four glycine residues that provides the flexibility necessary to adapt to different substrates in the proximal and distal binding pockets.[380] It is likely the structure of these pockets contribute to AcrB's broad substrate polyspecificity, as the proximal binding pocket has been shown to favour interactions with larger molecules, whereas the distal binding pocket appears to be a general

binding pocket for AcrB substrates.[216,380–382] There is a flexible hinge termed the hoisting loop between the TM and porter domains, which links TM8 and the PN2 subdomain.[377] It changes conformation from a random coil to a α-helix during the export cycle, and thus makes a hoisting movement which enables large motion of the PN2 subdomain passively, whilst also closing the CH1 substrate channel.[383] The centre of the trimer contains a cavity where drugs are transported to in the efflux mechanism. The central cavity is plugged by an ordered lipid bilayer, and contains specific hydrophobic protein-lipid interactions.[384] Furthermore, there is also a distinct lipid belt around the TM region and it is suggested that lipids support and harmonize AcrB's peristaltic motions as it undergoes its efflux mechanism (see **section 1.10.4.2**).

1.10.4.1 AcrZ

Hobbs *et al.* (2012) discovered a small 49 amino acid inner membrane protein, termed AcrZ, that interacts with AcrB.[385] The *acrZ* gene is highly conserved, and it is coregulated with *acrAB* through the transcription regulators MarA, Rob and SoxS. AcrZ folds into a long hydrophobic α-helix that fits into a wide groove in the TM region of AcrB.[366] AcrZ is inclined 45˚ compared to the membrane, and whilst it is unusual for a TM α-helix to have this inclination, it optimizes the interfacial complementarity. The interfacial residues are highly conserved across homologs, suggesting similar interactions occur across other RND proteins. Work has to be done to show strains of *E. coli* lacking *acrZ* are susceptible to some antibiotics usually exported by the AcrAB-TolC complex. It is not clear how AcrZ modulates the activity of AcrB for certain antibiotics, but potential mechanisms could involve modifications to the drug pockets or restricting their access.[371] It is thought small proteins such as AcrZ help create RND proteins polyspecificity to antibiotics.

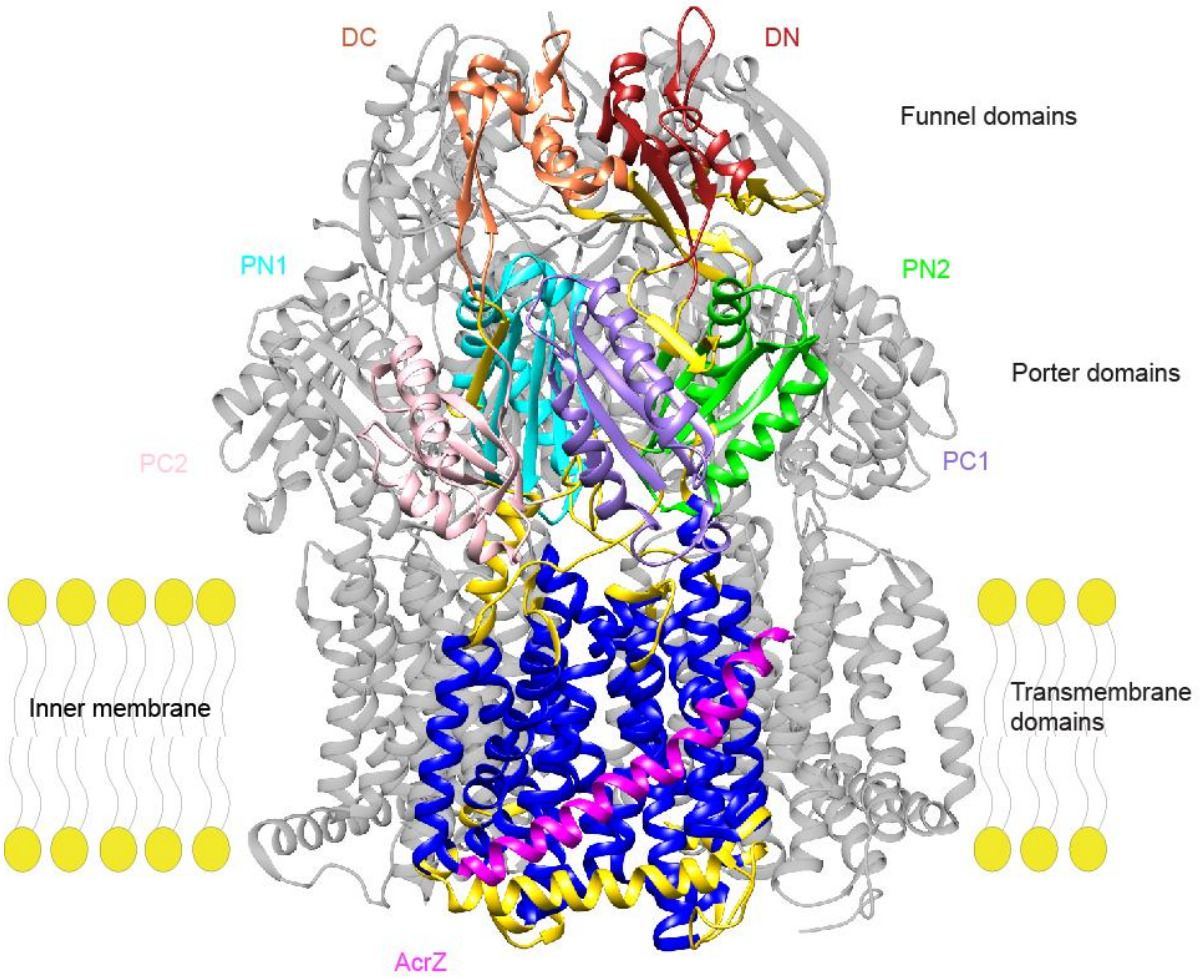


Figure 1.31. Structure of AcrB. AcrB trimer, with only one protomer coloured. Transmembrane domains coloured blue. Porter domain containing PN1 (cyan), PN2 (green), PC1 (purple), PC2 (pink) subdomains. Funnel domain containing DN (red) and DC (coral) subdomains. AcrZ for the coloured protomer is shown in magenta. PDB 5O66.

1.10.4.2 Rotational efflux mechanism

It has been suggested that AcrB sequesters substrates from the outer leaflet of the inner membrane or from the periplasm.[334] Hydrophobic compounds readily penetrate the outer leaflet of the inner membrane and thus can be effused by AcrB. However, substrates that reach the cytosol are transported to the periplasm by single-component multidrug efflux pumps of the SMR, MFS, MATE or ABC superfamilies, before being transported across the outer membrane by AcrAB-TolC.

AcrB loses its molecular symmetry upon drug binding.[371] Each protomer cycles through three distinct conformations: Loose (L), Tight (T) and Open (O), corresponding to drug access, binding and extrusion respectively (**Figure 1.32**).[386,387] PN2 and PC1 undergo cyclic changes due to their role in substrate binding, whilst PN1 and PC2 remain rigid. However, all PN and PC domains undergo large conformational changes when cycling between two states. Throughout the transport cycle, the TM domains undergo relative motion; from L to T, they exhibit a rocking motion, but from T to O they exhibit a shearing motion. The central cavity created by the DN and DC subdomains remains consistent throughout the efflux mechanism. The cycle is powered by the proton motive force.

Focusing on a single protomer, the efflux mechanism functions as follows. When a substrate binds to the PN2 and PC1 subdomains in the T state, it triggers the movement of TM2 and the subsequent formation of water molecules to residues D407 and D408. Once the aspartic acid residues have been protonated, the protomer cycles to the O state. Protonation causes the shearing motion of the TM helices which disrupts the water chain and prevents them from returning to the periplasm. This also induces the movement of TM2 and TM8, which causes PN1 and PC2 to undergo a significant conformational change that closes the PC1 and PC2 cleft. This seals off the path to the periplasm and creates a new drug exit tunnel towards the funnel domain (the exit duct). Simultaneously, the movement of TM2 causes the flexible PN2 and PC1 to close the distal binding pocket, which squeezes the drug out in a peristaltic manner as shown in **Figure 1.32.**[386,388] This mechanism occurs in each protomer asymmetrically.

1.10.4.3 Comparisons to other RND proteins

AcrB is the principal RND efflux system in *E. coli*, and it is also present in *Salmonella* and *Klebsiella pneumoniae*.[389,390] Overall, *E. coli* has five other RND pumps with varying percentage identity to AcrB: AcrD (66%), AcrF (78%), MdtF (71%), MdtBC (28/28%) and CusA (22%) and many of the RND pumps in *E. coli* are phylogenetically closely related to those of *Salmonella*.[391–394] Aside from *E.coli*, AcrB has homologs across other bacterial species, including ESKAPE bacteria. Examples include MexB in *Pseudomonas spp.*, CmeB in *Campylobacter spp.,* and AdeB in *A. baumannii*.[395]

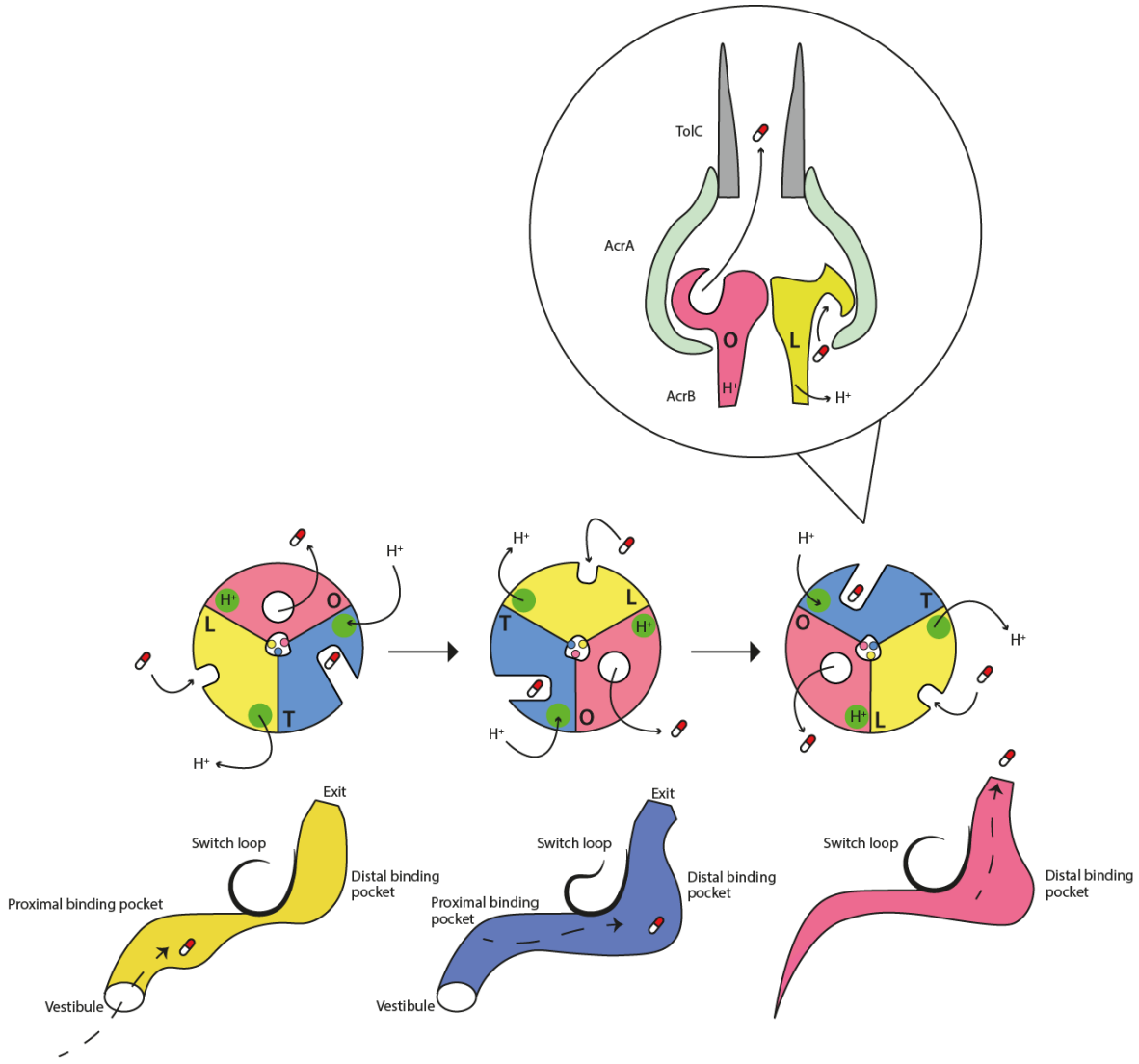


Figure 1.32. AcrB rotational efflux mechanism. The transport cycle is shown. The AcrB trimer cycles through three states, access (L (loose)), binding (T (tight)) and extrusion (O (open)). L is yellow, T is blue, and O is red. Protonation is required for the transport cycle. The bottom of the diagram shows the peristaltic mechanism of drug efflux through the AcrB protomer. They are colour-coded in the same manner as the transport cycle above it. A cross section is shown at the top corresponding to the third arrangement of the AcrB protomer. AcrA and TolC are also present, and it shows the drug leave an AcrB protomer in the O state and enter the TolC channel.

In general, the RND transporters of the HAE-1 and HME families have a very similar appearance.[335] As with AcrB they form hetero- or homo-trimers, with each protomer containing TM, porter and funnel domains. The TM domains are the most conserved portion of the RND proteins. They usually have 12 TMs and are arranged in two pseudo-symmetric 6 TM bundles. In particular, three residues forming salt bridges (D407, D408 and K940 – AcrB *E. coli* numbering) were 100% conserved in 135 RND proteins, highlighting the crucial role of proton translocation during the efflux cycle.[395] On the other hand, the porter domain was variable across the RND proteins, except for D568 at the interface-lock and R637 in the PC1 subdomain.

Among homologs, substrate recognition can differ to varying degrees. Similarities and differences in the binding pockets can help to understand this. The hydrophobic trap in the distal binding pocket almost always contains hydrophobic residues which can partially explain similarities in substrate recognition and binding structure between homologs – for example AcrB in *E. coli* and AdeB in *A. baumannii* have quite different amino acids in the hydrophobic trap yet show drugs bound at the same location within the distal binding pocket, with different amino acids interacting with the drug molecules.[395] AcrDs are the big outlier for the hydrophobic trap conservation; their trap contains Asn, Ser, Pro Tyr and two Phes, making it far more hydrophilic than the traps of the other RND transporters. This may explain why AcrD in *E. coli* can export aminoglycosides whilst other AcrB substrates such as ciprofloxacin and tetracycline are poorly exported. Furthermore, outside of the hydrophobic trap, the distal binding pocket can vary significantly, allowing for differences in substrate recognition between transporters.

1.10.5 AcrA

AcrA is a 397 amino acid MFP with four linearly arrange domains that are connected by flexible linkers and has an elongated shape essential for its function as a flexible adaptor protein (**Figure 1.33**).[396–398] It has been proposed to form a trimer of dimers in the assembled complex to maintain a sealed channel, and accommodate the conformational movement of AcrB as it cycles through its rotational efflux mechanism.[399] Residues 1-24 constitutes a classic signal peptide allowing its transport to the inner membrane, where it is then cleaved off, exposing an N-terminal cysteine residue. This residue is lipidated, to anchor AcrA to the inner membrane; however previous work has shown AcrA can still retain function in the absence of the lipid moeity.[396,397,400,401]

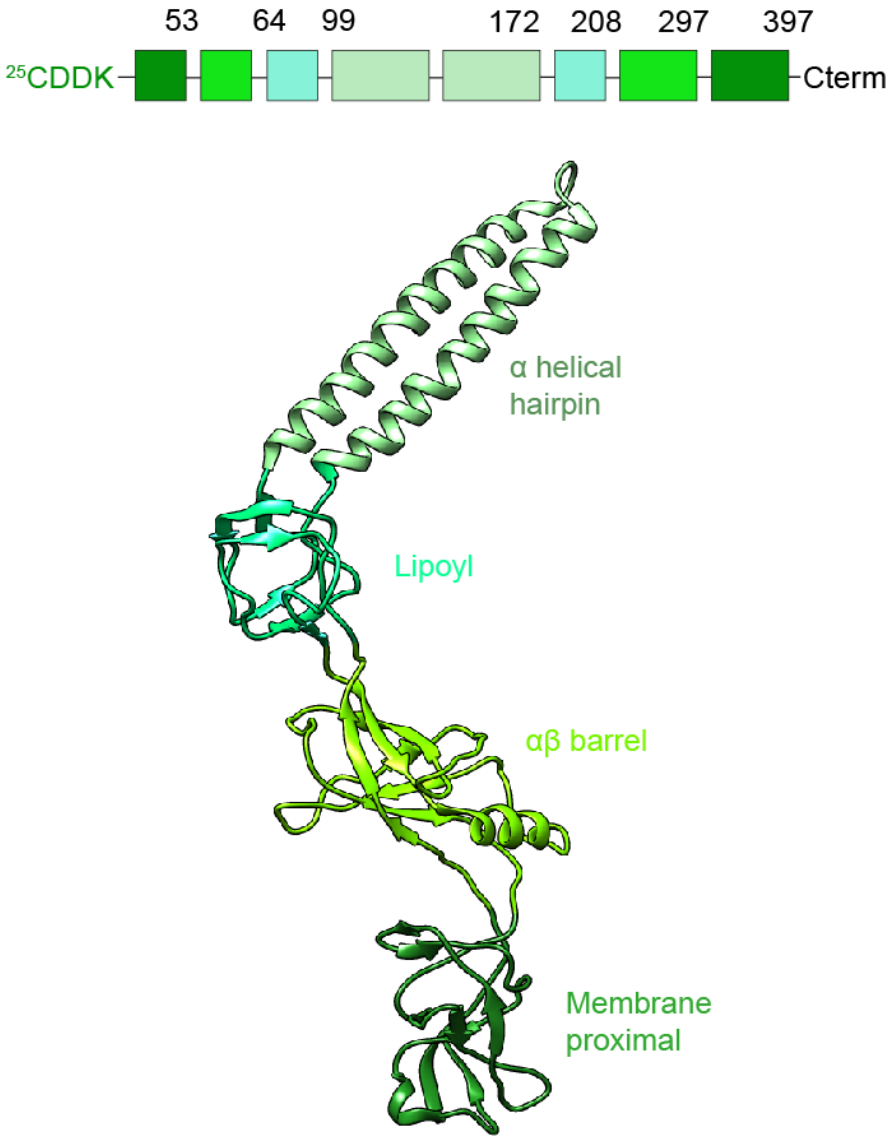


Figure 1.33. Structure of AcrA. Domain map shown at the top. All four domains coloured in different shades of green. PDB 5O66.

The AcrA α-helical domain is the docking site for TolC in the assembled complex. It is 58 Å in length with five heptad repeats per helix. It can assume multiple conformations due to the flexibility of the hinge region linking it to the lipoyl domain. The lipoyl domain should be thought of as two lipoyl half-motifs separated by an α-helical hairpin formed by helices α1 and α2. The two half-motifs are identical; they consist of four β-strands that come together to form a β-sandwich.[402] The exact role of the lipoyl domain is unknown, but it is hypothesised to stabilise the assembled complex.[403] The αβ-barrel domain consists of six anti-parallel β-strands and a small α-helix (α3). Usually, in related proteins, this domain is involved in ligand binding, but it is unclear whether AcrA or other MFPs interact with specific

ligands.[404] The membrane proximal (MP) domain is presumed to be unstructured, and are often unresolved in crystal structures of AcrA alone, as well as being proteolytically labile. The αβ-barrel and MP domains are involved in defined interactions with AcrB. The binding of protein partners, ligands, lipid modifications and interprotomer interactions can stabilise the MP domain into a β-roll structure.[403,404] The positioning of the αβ-barrel and MP domains can change the shape of AcrA from a crescent chape to a more extended conformation.[398]

1.10.5.1 Comparisons to other MFPs

The majority of RND transporters and some MFS and ABC transporters in Gram-negative bacteria require the presence of MFPs to function.[405] In *E. coli* there are five RND associated MFPs, including AcrA. AcrA has varying percentage identity to the other four MFPs: AcrE (69%), MdtE (55%), MdtA (29%) and CusB (24%). AcrA's closest homolog is AcrE, and previous work in *Salmonella* has demonstrated the interchangeability between the two MFPs when forming a functional efflux pump with AcrD.[406] AcrA has homologs among other RND efflux pumps across ESKAPE bacteria, such as MexA of the MexAB-OprM efflux pump from *P. aeruginosa* with 59% sequence similarity.[259] AcrA and MexA are the two best characterised proteins in the MFP family, and they are two members of a phylogenetic cluster of MFPs that function with RND efflux pumps that confer clinically significant levels of resistance.[405,407] This means research on AcrA translates well between different systems in other bacterial species.

MFPs share the general architecture exhibited by AcrA; they all have an elongated structure with a "hairpin" arrangement, whereby the polypeptide chain comes together in the tertiary fold, so each domain is made from two sections of the amino acid sequence: an N-terminal side and a C-terminal side (**Figure 1.33**). The majority of MFPs have an α-helical hairpin that forms a coiled-coil arrangement. This domain is widely variable between MFPs, with the hairpin shortened or extended by the deletion or insertion of heptad repeats; for example the AcrA α-helical hairpin is seven residues longer than MexA's.[397,408] The AcrA lipoyl and αβ-barrel domain are typical of MFPs, and the ~100 C-terminal amino acids are the most conserved.[405] This lies in AcrA's MP domain, which is present in many MFPs but not all. It's formed of the N- and C-terminal ends of the protein and is highly dynamic due to the flexible connection to the αβ-barrel. Both AcrA and MexA MP domains form a β-roll structure when stabilised.[369,403,404,408] MFPs can be anchored to the inner membrane through the cleavage of an N-terminal signal peptide and subsequent cysteine lipidation, such as with AcrA/MexA type MFPs, or through an N-terminal transmembrane helix.[408] MFPs that associate with HME-RND transporters can also present additional N- and C-terminal domains.[408]

AcrA has flexible linkers between its linearly arranged domains, which is another key feature among MFPs.[408] Due to the hairpin nature of the MFP polypeptide chains, they consist of two-antiparallel strands. They have a distinctive structures with a degree of inter-strand hydrogen bonding. A common thread between solved structures of different MFPs, such as AcrA, MexA, CusB and ZneB, is that these linkers can accommodate a range of different angular and rotational flexibility between the domains.[394,397,398,404,409] This feature is critical to the function of the MFP in tripartite assemblies, as they have to interact with an outer membrane and inner membrane component simultaneously, whilst accommodating conformational changes during the efflux cycle.

The functional unit of an MFP is a dimer, with each protomer engaging in a specific interface with a RND protein and OMF.[410] In tripartite assemblies, the MFP dimer trimerizes to form an inverted-funnel structure creating a path for the substrate through the periplasm. *In vitro*, MFPs lacking the N-TM or lipidation tend to be monomeric, as has been shown for AcrA, MF-type MFP EmrA, and ABC-type MacA.[405] However, the whole-length MFPs have a propensity for oligomerisation.[410,411] In AcrA, the functional dimer unit consists of two AcrA protomers; this is the same for homologs such as AcrE and MexA. However, the RND efflux pump TriABC-OpmH from *P. aeruginosa* requires two different MFPs (TriA and TriB) that play non-equivalent roles in the pumps function.[412]

1.10.6 TolC

TolC is a 471 amino acid outer membrane β-barrel that acts as the OMF protein and exit duct for the AcrAB-TolC complex (**Figure 1.34**).[413] It is a 12 stranded β-barrel, and quite unique for an outer membrane protein. As a trimer, each monomer comes together to form one barrel. Each monomer contributes four anti-parallel β-strands and four anti-parallel α-helical strands to create TolC's channel and tunnel domains respectively.[414] TolC is a tapered hollow cylinder, 140 Å in length. The tunnel domain is 40 Å in length, and the channel domain is 100 Å in length, residing entirely in the periplasm. It also has a third domain containing an αβ structure, which forms a belt around the periplasmic tip of the tunnel. This is another example of TolC's unique structure, as most outer membrane proteins have a restriction domain at the equator of the tunnel.[415] The internal diameter of TolC is 35 Å for most of its length, yet reduces to 4 Å at the restricted domain.

The periplasmic portion of TolC holds distinctive features. Each of the 12 helices is packed in a 'knobs-into-holes' structure; each helix is packed laterally with two helices as neighbours, and the intermeshing of constituent side chains stabilises the arrangement. They follow a left-handed superhelical twist in the α-barrel, but the helices are underwound in the β-barrel proximal half. This enables them to lie on a cylindrical surface. Bulkier side chains reside on the outside of the barrel to

further facilitate this. Tunnel assembly is stabilised by hydrogen bonding and possible salt bridges between monomers. At the periplasmic tip, one coil from each protomer folds inwards to help reduce the diameter in the restricted domain; this ensures TolC is in the closed state in isolation, and only opened with defined interactions in the assembled complexes.

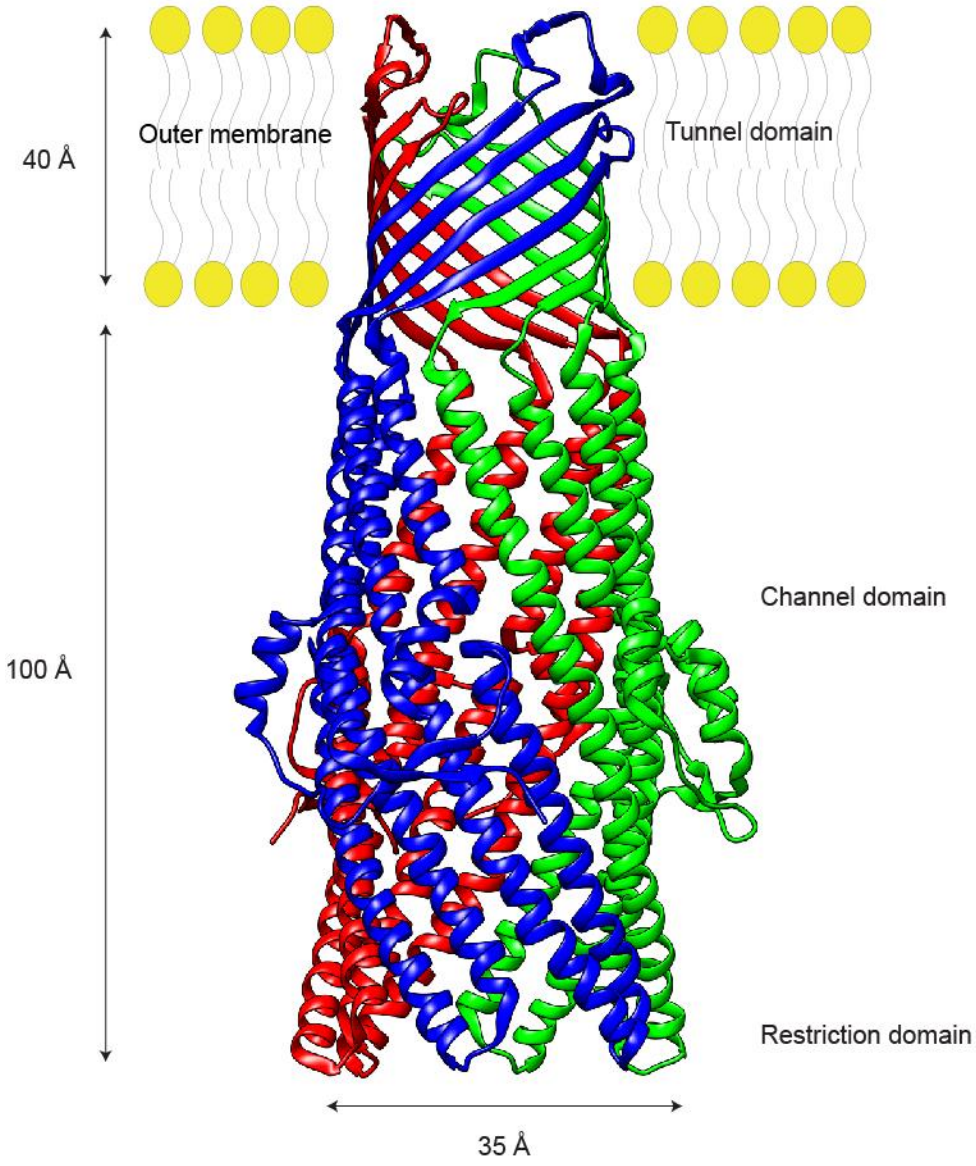


Figure 1.34. Structure of TolC. Each monomer shown as a different colour. Structure from PDB 5O66, therefore TolC is shown in the 'open' position.

1.10.6.1 Comparisons to other OMFs

Sequence analysis with other OMFs show they are highly divergent, with only two subtle common motifs in the coiled-coil regions.[402] OMFs can be divided into three clades based on their efflux functions: a) multidrug efflux, such as OprM from *P. aeruginosa*, b) cation efflux, such as CusC from *E. coli*, c) protein export (type I secretion system), such as TolC from *E. coli*.[416] In *E. coli* there are four OMF proteins including TolC (MdtP, MdtQ and CusC), yet TolC inactivation increases susceptibility to multiple antibiotic agents, suggesting it is a major conduit in *E. coli*.[417] In fact, TolC interacts with multiple different classes of multidrug efflux pumps such as those from the RND, ABC and MFS superfamilies and bacteria often have multiple homologs of TolC that interact with a number of different pumps to confer broad, overlapping substrate specificities.[418] This is in contrast to *P. aeruginosa*, which contain 18 OMFs divided into two subfamilies; the OprM subfamily with 11 homologs involved with multidrug efflux pumps, and the divergent AprF protein export family (type I secretion system), with AprF and OpmH being the most closely related to TolC.[416,419]

1.10.7 AcrAB-TolC interactions in the assembled complex

1.10.7.1 AcrA – AcrB interaction

Many of the AcrAB-TolC interactions have been well characterised. AcrA acts as a trimer of dimers in the assembled complex, with each protomer in an AcrA dimer binding AcrB differently. The model by Du *et al.* (2014) shows one AcrA protomer bridges the upper regions of the PC1, PC2 and DN subdomains of AcrB.[366] However, in the second protomer, the MP domain shifts towards the AcrB PN2 subdomain, so only the upper regions of PN2 and DN contact AcrA. This model was supported by Kim *et al.* (2015).[373] McNeil *et al.* (2019) further analysed available cryo-EM data of assembled AcrAB complexes.[389] They identified four binding sites on AcrA that interact with AcrB (**Figure 1.35**). For the first AcrA protomer, site 1 is the αβ-barrel domain to the DN subdomain of AcrB (Nα4 helix and β-hairpin1). Site 2 is αβ-barrel domain of the DC subdomain (Cα4 helix and β-hairpin2). Site 3 is the MP domain to the base of the DC subdomain, and Nβ8-Nβ9 of the neighbouring AcrB protomer. Site 4 is the MP domain to the AcrB pore domain (PC1). For the second protomer, site 1 is the αβ-barrel domain to the DC α-hairpin of the following AcrB subunit. Site 2 is the αβ-barrel domain to the AcrB α-hairpin and the β-hairpin of the DN subdomain. Site 3 is the MP domain to the base of the funnel domain, and site 4 is the MP domain to the AcrB pore domain (PN2). From these four binding sites, they identified nine discrete linear sequence "binding boxes" in AcrA (**Figure 1.35**). This analysis was performed on AcrA from Salmonella *enterica* and aligned to the sequence of AcrA from *E. coli*, which confers >90% sequence homology.

1.10.7.2 AcrA – TolC interaction

An essential function of AcrA is to bind to TolC and alter it to the 'open' state so it is a continuous channel in the assembled complex.[366] The Du *et al*. (2014) model suggests as follows: the AcrA α-helical hairpins pack into a cylindrical α-barrel that interacts with the TolC periplasmic α-barrel. The TolC α-barrel is partially inserted into the slightly larger AcrA α-barrel, and they suggest the dilation of TolC is driven by chelate cooperativity (due to AcrA hexamerization) and allosteric cooperativity (due to the breaking of interprotomer interactions in TolC). Kim *et al*. (2015) proposes a different mechanism of how TolC switches from the closed to open state; they suggest an intermeshing cogwheel interaction, otherwise known as tip-to-tip model.[373,420] They argue this model fits available EM density more accurately and supports previous biochemical and genetic data reporting three conserved residues (a Leucine, Arginine and Serine) in the adaptor proteins α-helical domain, and 24 conserved amino acids in the periplasmic tip of TolC, all essential for binding.[421,422] Furthermore, recent work on MFPs binding to the OMF protein suggest they bind two separate binding sites on the

OMF, each with a separate role; one MFP subunits 'grasps' the OMF whilst the other opens the channel.[410,412,423] Interestingly, it has been shown that pH is important for the AcrA:TolC interaction, with the subcomplex becoming more unstable at pH ranges above 6.[411] This is likely an adaption to the often more acidic periplasmic environment.[1,2] Overall, further investigations are needed to fully decipher the how TolC switches to the open state.

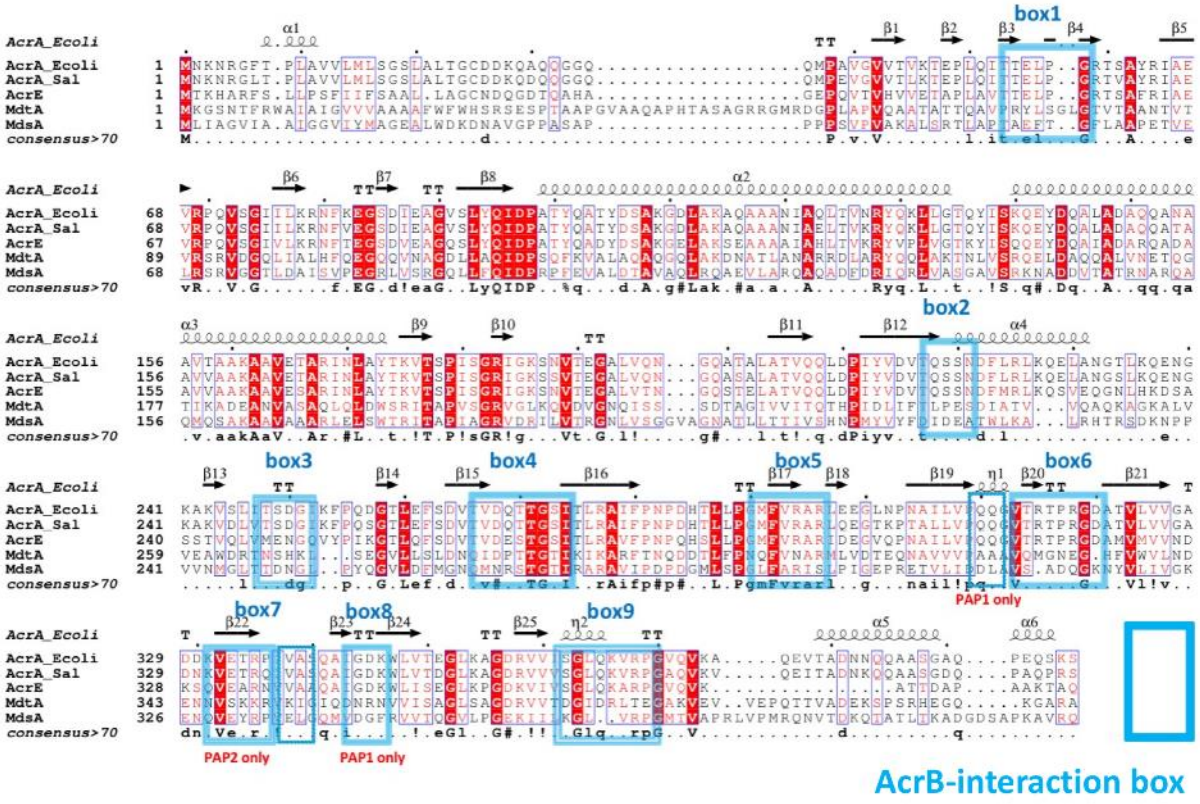


Figure 1.35. The discrete "binding boxes" of periplasmic adaptor proteins. A multiple sequence alignment of 4 *Salmonella* MFPs combined with the structural alignment of the *E. coli* AcrA structure (PDB: 5O66) reveals clear domain boundaries and correspondingly high likelihood of secondary structure conservation. Identical residues in red. Figure taken from McNeil *et al.* (2019).[389]

1.10.8 AcrAB-TolC assembly

The mechanism of pump assembly is continually debated in the literature. The model by Du *et al.* (2014) finds no interaction between AcrB and TolC in the assembly. This was supported by *in vivo* crosslinking studies and pull-down assays which found AcrA interacts independently with AcrB and TolC, and no interaction between AcrB:TolC.[424] This model also accounted for thermodynamic measurements, which suggests the AcrB:TolC interaction doesn't happen spontaneously like the other interactions. Daury *et al*. (2016) investigated the assembly of RND pumps in nanodiscs and did not observe any contact between AcrB and TolC through EM analysis.[372] They did conclude the observed set up may have been an intermediate formation, and under *in vivo* conditions there may be another conformational step where AcrB and TolC are in direct contact. Interestingly, work by Zgurskaya *et al.* (2011) contradicts these findings; using surface plasmon resonance (SPR) they observed TolC binding to AcrB, independently of AcrA, with nanomolar affinity, and proposed a mechanism where the tripartite complex is assembled between AcrA and TolC bound to AcrB.[367] A recent study by Shi *et al*. (2019) suggests a sequential mechanism of assembly, beginning with the formation of the AcrAB subcomplex, and highlights the importance of peptidoglycan in pump function and assembly.[38] From cryo-EM density and crosslinking MS, they conclude the α-helical domain of AcrA is in contact with peptidoglycan in the periplasm, and that this interaction is essential to the AcrAB subcomplex and tripartite assembly. This was further supported by molecular dynamics (MD) simulations by Gumbart *et al.* (2021), which suggested Lpp positions peptidoglycan at the AcrA-TolC interface in pump assembly.[257] Furthermore, combined with the N-terminus being anchored to the inner membrane by its lipidation, AcrA can communicate the conformational changes of TolC and keep them coupled to the substrate binding of AcrB.

1.11 Thesis aims

This thesis presents an investigation into the role of structural dynamics in the function, assembly, and inhibition of the AcrAB-TolC multidrug efflux pump. By using a combination of structural MS techniques supplemented with relevant biochemical/biophysical studies, the aim was to gain a unique insight into the structural dynamics of the constituent proteins, with a significant focus on the MFP AcrA. MFPs are known to be flexible, dynamic proteins, that link an inner membrane protein and an outer membrane protein, to form a sealed channel across the Gram-negative cell envelope, however there is a lack of structural biology information available on AcrA's functional dynamics, making it an ideal candidate for investigation by structural MS methods.[366,369]

The results chapters are **chapters 2-5**, with each chapter having its own methods section. In **chapter 2**, the first aim was to characterise AcrA using both HDX-MS and native MS. Using native MS, the aim was to investigate the effect of the N-terminal lipidation (by using two constructs either containing or missing the lipidation) and pH on the oligomeric state of AcrA. Then HDX-MS aimed to provide a first look at the backbone structural dynamics of AcrA, and explored whether weakly acidic pH or Mg^{2+}, reflecting the periplasmic environment, had any effect on the dynamics of AcrA.[1–3]

Chapter 3 builds on the work shown in **chapter 2.** The functional unit of an MFP is a dimer, and the two protomers can function slightly differently; this is exemplified by the differences each AcrA protomer in the dimer unit binds AcrB.[389,410] **Chapter 3** aimed to gain insight into the effects of AcrA dimerization, by characterising a novel pseudo-dimeric AcrA construct. The aim was to use circular dichroism (CD) to investigate whether pseudo-dimerization increased the thermal stability of AcrA and to use HDX-MS to see how the structural dynamics were affected. Furthermore, the effect of pseudo-dimerization on AcrA binding to AcrB and peptidoglycan was explored.

Chapter 4 aimed to determine the first mechanism of action of the recently discovered AcrA inhibitor, NSC 60339 by using a combination of HDX-MS, native MS, MD simulations, biophysical techniques, and cellular inhibition assays.[403,425] This was performed in collaboration with the Zgurskaya laboratory (University of Oklahoma) and the Gumbart laboratory (Georgia Institute of Technology).

Chapter 5 shifted away from AcrA and aimed to probe AcrAB-TolC in increasingly complex lipid environments. The aim was to purify the constituent proteins in different membrane mimetic environments for downstream experiments, and to explore whether styrene maleic acid (SMA) co-polymer could be used to purify the OMF TolC. Recently, Hammerschmid *et al*. (2023) detailed an online deplidation workflow for HDX-MS experiments; as a validation of this protocol, the aim was to monitor the effect of an AcrB inhibitor MBX-3756 on the structural dynamics of AcrB in a lipid

environment (membrane scaffold protein (MSP) nanodiscs).[4] Furthermore, this chapter aimed to develop a novel SMALP-liposome-SMALP assay, that could be applied to answer a pre-existing question regarding the effect of antimicrobial peptides on AcrB's oligomeric state in lipid environments.[426] The last aim was to develop a quick biochemical assay to probe AcrAB-TolC assembly *in vitro*, to be able to monitor assembly in the presence of different ligands and conditions.

Chapter 2: Investigating the structural dynamics of the periplasmic adaptor protein AcrA

Sections of this chapter are currently being written up for publication:

'Mg2+ and pH coregulate the functional dynamics of a multidrug efflux adaptor protein'

2.1 Introduction

AcrA forms a trimer of dimers in the assembled AcrAB-TolC complex, and functions as a flexible membrane fusion protein (MFP).[366,403] Its primary role is to connect AcrB and TolC, whilst keeping a sealed channel for the efflux of toxins from the bacterial cytosol or periplasm. Furthermore, recent work by Chen *et al. (*2022) combined their cryo-electron tomography (cryo-ET) work on AcrAB-TolC with all currently available structural data to propose a model for the function of AcrAB-TolC.[366,386,387,400,403,427,428] They suggest that drug binding to AcrB and the subsequent conformational change in its porter domain, is communicated to TolC through the movements of AcrA. AcrA has to accommodate AcrB's rotational movement, whilst keeping a sealed channel; the AcrA protomers end up twisting which in turn opens the TolC channel, so the complex is in an efflux competent state. Therefore, for efflux to happen, AcrA must be highly dynamic to carry out its functions.

In the assembled AcrAB-TolC complex, AcrA resides entirely in the periplasm.[366] The periplasmic environment differs from the bacterial cytosol significantly (see **section 1.7.2.2**).[251] Its environment is dominated by the presence of the peptidoglycan layer, which has been suggested to play a role in the assembly of the AcrAB-TolC complex, and has been shown to bind AcrA at the AcrA:TolC interface.[38,257] In *Escherichia coli* (*E. coli*), half of the proton motive force across the cytoplasmic membrane comes from a proton gradient.[2,252] Therefore, there is a higher concentration of H^+ ions in the periplasm; pH in the cytosol is often ~1.7 pH units higher than the periplasm. It is possible that mediation of pH in the periplasm may be key to regulating the dynamics and conformation of AcrA in pump assembly and structure; as Wang *et al*. (2012) suggest.[2] Furthermore, the periplasm contains a far higher concentration of divalent cations than the cytosol such as Mg^{2+} (7.56 times more), and its role in the structure-function of AcrA is understudied.[1,3,252]

Various other biophysical studies and modelling have reported AcrA to be highly dynamic.[2,398] Wang *et al*. (2012) performed modelling and, combined with previous electron paramagnetic resonance (EPR) data, showed AcrA as a highly dynamic protein that exhibited a pH-induced conformational change.[1,2] This was attested to the protonation/deprotonation of His285, which was suggested to act as a conformational switch in AcrA. In one model, protonated His285 gained another hydrogen bond to Asp86 in the lipoyl domain, strengthening the local hydrogen bonding network and reducing dynamics across AcrA's structure, whereas in a second model this hydrogen bonding network was weakened, and an overall increase in backbone dynamics was observed. They observed these pH-induced conformational changes to be reversible, and suggested changes in pH could trigger AcrA conformational arrangement to begin complex assembly. Furthermore, molecular dynamics (MD) simulations by Hazel *et al.* (2019) revealed free AcrA exhibited a range of orientations but had two

main conformational basins.[398] One was a *cis*-like formation where the membrane proximal (MP) and α-helical domains point in the same direction, and a *trans*-like conformation where they point in opposite directions. Cryo-electron microscopy (cryo-EM) structures of AcrAB-TolC show AcrA to be in the *trans* conformation.[366,373] Moreover, they observed locking AcrA in the *cis* conformation comprised the assembly of the complex. Regardless of these studies, there is no experimental structural biology information regarding the specific dynamics of AcrA.

AcrA contains a lipidation which is attached to Cys25 on the matured N-terminus. In this thesis, the AcrA construct containing its lipidation is denoted $AcrA^{L}$. The lipidation has been characterised and proposed to be *N*-acyl-*S*-diacylglycerol containing two palmitoyl residues and one oleoyl residue, based on the measured masses reported by native mass spectrometry (MS).[400] The lipidation has been suggested to promote AcrA oligomerisation by stabilising AcrA's MP domain.[367] However, an AcrA variant lacking the lipidation (denoted $AcrA^{S}$ throughout this thesis) has been shown to remain active and has been used for several previous biochemical studies.[396,397,400,401] The exact oligomeric state of $AcrA^{L}$ and $AcrA^{S}$ *in vivo* and *in vitro* remains uncertain, with early work by Zgurskaya *et al*. (2000) suggesting AcrA formed trimers *in vivo*.[429] Furthermore, $AcrA^{S}$ has been suggested to oligomerise at acidic pH and work by Tikhonova *et al.* (2009) showed membrane fusion proteins (MFPs) are prone to oligomerisation, but $AcrA^{S}$ had a lower propensity to oligomerise than other similar proteins.[1,411] A more comprehensive study on the two constructs *in vitro* by Zgurskaya *et al*. (2011) used size exclusion chromatography (SEC) to investigate oligomers of $AcrA^{L}$ and $AcrA^{S}$ at both pH 7.4 and pH 6.0.[367] They observed $AcrA^{S}$ monomers at both pHs, but $AcrA^{L}$ saw a range of higher order oligomers at pH 7.4, and dimers at pH 6.0.

Structural characterisation of proteins is traditionally carried out by crystallographic, EM or nuclear magnetic resonance (NMR) studies.[15–17] These techniques are essential for elucidating structure-function relationships, but only provide a snapshot of the protein in question. Proteins are not static in solution, and many biophysical techniques lack the resolution to link protein dynamics with protein function. Therefore, the use of structural MS techniques to gain further insights on protein function is becoming increasingly popular. Hydrogen deuterium exchange MS (HDX-MS) is a powerful tool for probing protein structural dynamics (see **chapter 1.4**).[33] In particular, differential HDX (ΔHDX) is adept for analysing the differences in HDX between two states (such as wildtype vs mutant, apo vs holo), allowing for the characterisation and localisation of the effect of a condition on the dynamics of a protein. Native MS is effective at monitoring protein oligomers and binding partners, which is utilised to fully characterise the two AcrA constructs. MS is always most effective with complementary biophysical and biochemical techniques.

The aims of this chapter were to firstly express and purify the two AcrA constructs ($AcrA^{L/S}$) and investigate the effect of the lipidation and pH on AcrA oligomerisation using native MS. Then, focusing on $AcrA^{S}$, HDX-MS was deployed to reveal the first look at the backbone structural dynamics of $AcrA^{S}$. After characterising $AcrA^{S}$, the last aim was to investigate the effects of the periplasmic environment (pH and Mg^{2+}) on AcrA's structural dynamics, primarily using structural MS techniques.

2.2 Results

2.2.1 Overexpression and purification of the AcrA constructs

2.2.1.1 Purification of AcrA^{L} in DDM

The first attempts of AcrA^{L} purification did not yield any pure, intact protein for several reasons (gel not shown). Whilst AcrA^{L} was purified from detergent solubilised membrane fractions, detergent was not present in the buffers throughout the purification. This led to sample aggregation, due to the lipid moiety present in the sample. Furthermore, AcrA^{L} exhibited large amounts of degradation, with only a band at ~28 kDa visible on the gel. This has been characterised previously as the proteolytic stable core of AcrA.[400] This was likely due to the high amount of periplasmic proteases in the periplasmic environment degrading AcrA through the purification process.[258] Therefore, the protocol was modified to that described in **methods 2.4.3.1** to include *n*-Dodecyl-β-D-maltopyranoside (DDM) in all AcrA^{L} buffers, and the protease inhibitor phenylmethylsulfonyl fluoride (PMSF) was increased from 100 µM to 1 mM, in align with previously described protocols.[396,400,411]

Figure 2.1 shows the results of a typical optimised AcrA^{L} purification in DDM. AcrA^{L} was solubilised from the cell membrane and purified by Ni^{2+} affinity chromatography and SEC. AcrA^{L} was purified in these two stages to ensure any aggregates from the purification are successfully removed. A SEC chromatogram often contains a void peak containing aggregates, or shouldered peaks which contain degraded or truncated protein. Using SEC, they can be visualised and discarded. **Figure 2.1a** shows a SEC chromatogram for AcrA^{L} and **Figure 2.1b** shows the corresponding sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel containing the collected fractions. The SEC chromatogram shows a void peak just after 40 mL (fractions 4-7), and then a broad trace which is likely caused by a range of oligomers formed by AcrA^{L} at pH 7.4. The SEC chromatogram in **Figure 2.1a** resembles the SEC seen for AcrA^{L} at pH 7.4 by Tikhonova *et al.* (2011), which also found AcrA^{L} forms a range of oligomers at this pH.[367]

The SDS-PAGE gel (**Figure 2.1b**) shows a clean sample with no contaminants present. Monomeric AcrA^{L} has a theoretical molecular weight of 41,624 Da, and the protein can be visualised on the gel at ~40 kDa. In fact, AcrA^{L} presents as a double band, suggesting there were two populations in the sample. Possibilities could be degradation, conformational differences, or variety within the lipidation of AcrA. Further investigations were required to explain this double banding (see **section 2.2.2.3**). Fraction 28 presented as nothing on the gel, and was likely degraded products or contaminants, as a 90 mL elution volume corresponded to a mass far lower than that of AcrA^{L}. Furthermore, a band can be seen at ~90 kDa and was present throughout the SEC fractions. To determine whether this was AcrA^{L} or a co-

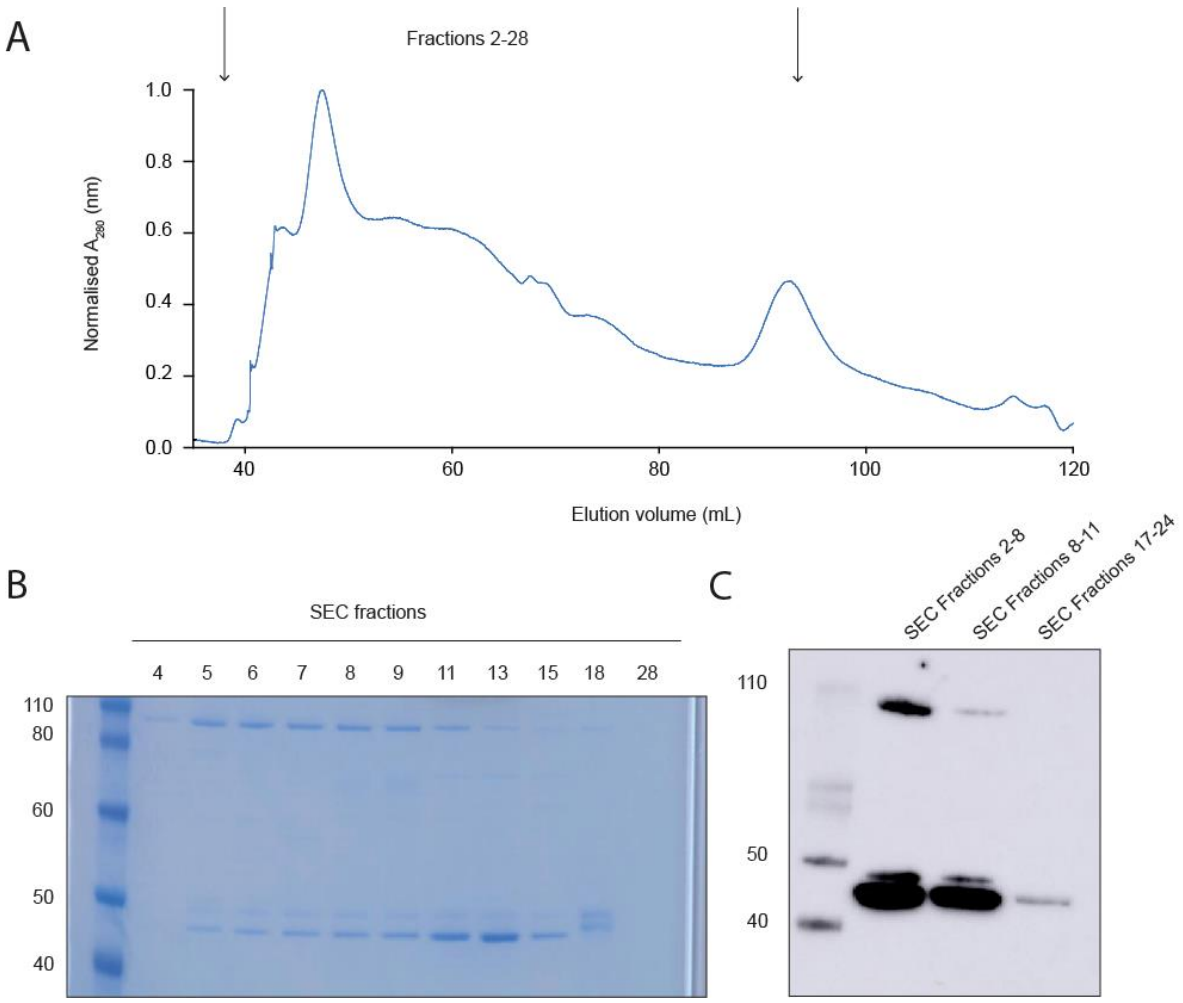


Figure 2.1. Characterisation of AcrA^L purification. A. Characteristic size exclusion chromatogram for AcrA^L. Fractions were loaded onto the gel and/or western blot labelled 2-28. Absorbance normalised. Void peak shown at ~40 mL and broad shoulder across the trace suggests various oligomers of AcrA^L. **B.** Characteristic SDS-PAGE running the size exclusion chromatography fractions shown in part A. Gel shows a double band at around ~40 kDa and a higher single band at ~90 kDa for AcrA^L. **C.** Western blot of pooled fractions from the size exclusion chromatogram. Western blot ran with Anti-PolyHistidine-HRP Antibody. It shows all bands seen on SDS-PAGE are the AcrA^L construct.

puritant, a western blot using Anti-PolyHistidine-HRP Antibody was performed (**Figure 2.1c**). It confirmed that both bands in the double band at ~40 kDa and the higher band at ~90 kDa was our $AcrA^L$ construct. Further detail on this higher band can be found in **section 2.2.2.1**. Furthermore, both the SDS-PAGE and western blot show the bottom band in the ~40 kDa double band is more intense and dominant in the sample. This purification protocol yielded 0.1-0.2 mg/mL of $AcrA^L$, equivalent to between 2-5 µM.

2.2.1.2 Purification of $AcrA^S$

$AcrA^S$ lacks signal peptide 1-24, so it is not transported to the membrane and is hence purified from the cytosol like other soluble cytoplasmic proteins. Furthermore, its Cys25 residue is mutated to a Met, so it lacks the lipidation, and does not require the presence of detergent in the sample buffers. A characteristic purification of $AcrA^S$ is shown in **Figure 2.2**. The SEC trace (**Figure 2.2a**) shows a significant difference to the lipidated sample. The void peak was observed at ~45 ml, then a single peak was measured at ~75 mL constituting the entire sample. A peak at ~75 mL corresponds to a mass of ~40 kDa, so this suggests $AcrA^S$ is monomeric. This is in agreement with Tikhonova *et al.* (2011) who also observed $AcrA^S$ as a monomeric sample.[367] Only the middle three fractions of the main peak were pooled and stored for further experimentation.

The SDS-PAGE gel (**Figure 2.2b**) shows a very intense band at ~40 kDa representing $AcrA^S$. The molecular weight of $AcrA^S$ is slightly smaller than $AcrA^L$ due to the lack of lipid moiety; it has a weight of 40,816 Da. Fractions were loaded onto the gel as they were eluted from the SEC column, to provide an immediate idea of concentration before measuring specifically on a nanodrop (Implen NanoPhotometer, Geneflow). The band at ~40 kDa presents as a singlet rather than a doublet; this suggests the doublet is to do with the lipidation on AcrA. There is a small band at ~90 kDa as seen previously, that is likely AcrA. Further experiments were needed to determine whether this is due to dimers in the sample, or if this is an artefact of AcrA as a non-globular MFP. The western blot (**Figure 2.2c**) confirmed the presence of our $AcrA^S$ construct. This purification protocol yielded a far greater concentration of protein, at ~6 mg/mL, equivalent to ~150 µM.

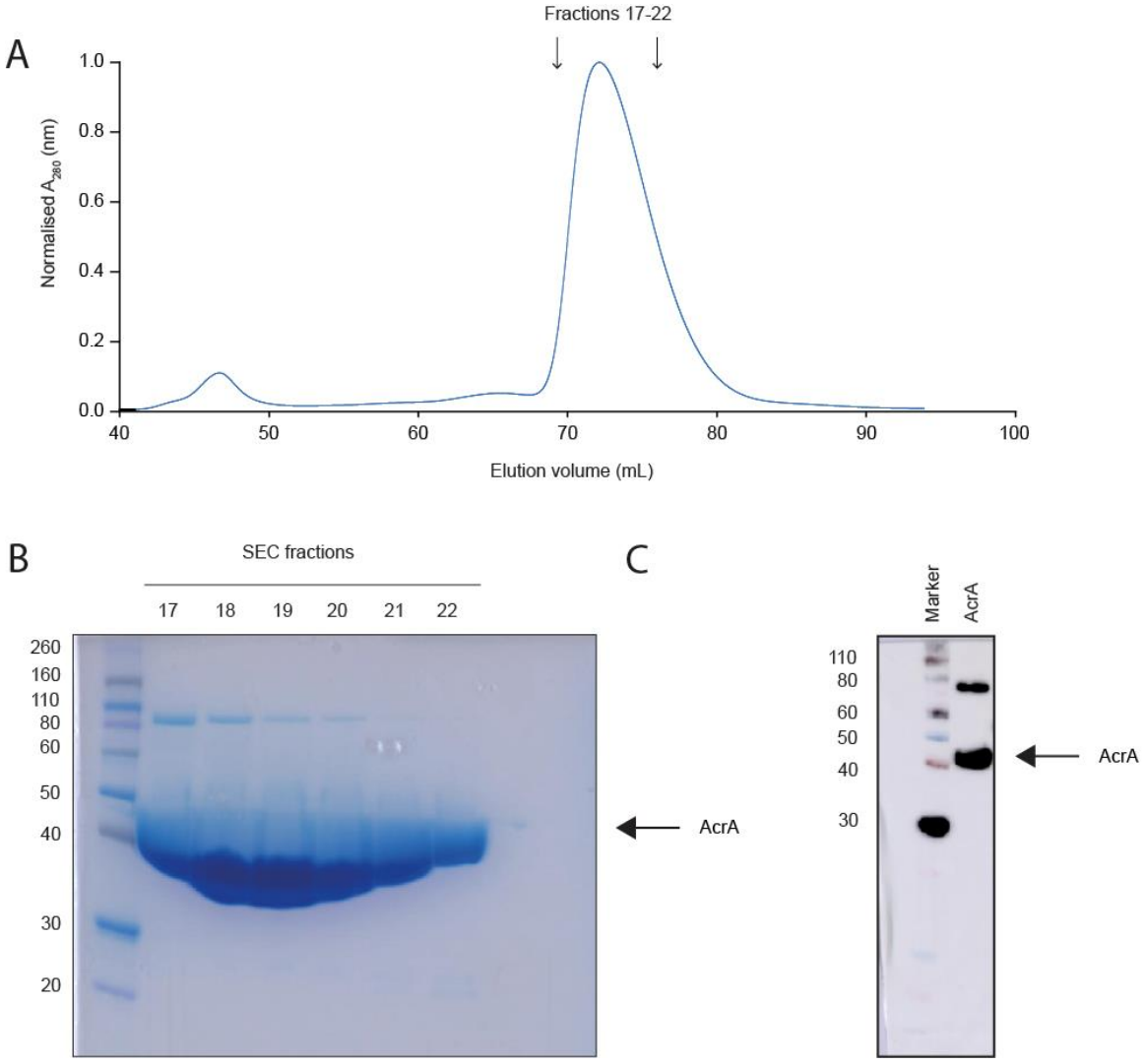


Figure 2.2. Characterisation of AcrA^S purification. A. Characteristic size exclusion chromatogram for AcrA^S. Fractions of the main peak were loaded onto the gel and/or western blot labelled 17-22. Absorbance normalised. Void peak shown at ~40 mL and main peak at ~75 mL suggests monomeric AcrA^S. **B.** Characteristic SDS-PAGE running the size exclusion chromatography fractions shown in part A. Gel shows a single band at around ~40 kDa and a fainter higher single band at ~90 kDa for AcrA^S. **C.** Western blot of pooled fractions from the size exclusion chromatogram. Western blot ran with Anti-PolyHistidine-HRP Antibody. It shows all bands seen on SDS-PAGE are the AcrA^S construct.

2.2.2 Native MS of the AcrA constructs

Native MS was carried out as described in **methods 2.4.5.1**. Native MS is a simple and effective technique to determine protein mass, stoichiometries, homogeneity and post-translational modifications. nano-electrospray ionisation (nESI)-MS only uses 2 µL of sample so it is well adapted to minimise the amount of sample needed, important for protein purifications with relatively low yield. To release AcrA^{L} from the detergent micelle, higher energy was required, and therefore the trap collision energy used ranged from 50-200 V and the sampling cone was set to 120 V. For these experiments, AcrA^{L} was buffer exchanged into 100 mM ammonium acetate with 0.03% DDM, at either pH 7.4 or 6.0.

2.2.2.1 Investigating AcrA oligomerisation

Figure 2.3a shows the native mass spectrum of AcrA^{L} at pH 7.4, and **Table 2.1** shows the theoretical and measured masses. The native mass spectrum of AcrA^{L} at pH 7.4 shows a range of detectable oligomers, from monomer to pentamer. UniDec analysis provided a breakdown of sample intensities for each oligomer.[118] The monomer and dimer species had relative intensities of 100% and 98.23%, so were nearly at a 1:1 ratio. The abundance of trimer was 80% less with an intensity of 20.42%. Then the intensities of tetramers and pentamers was 2.08% and 1.16% respectively. This supports the SEC trace seen in **Figure 2.1a**, that AcrA^{L} is heterogenous in nature and has a propensity to form oligomers at pH 7.4. The theoretical mass of an AcrA^{L} monomer is 41,624 Da, and agrees well with the measured mass which was 41,632 ± 7 Da. The mass of the theoretical dimer and the measured dimer differed by 26 Da, which is likely due to a sodium ion adduct. The mass of the theoretical trimer and tetramer differed to the measured sample by 7 and -11 Da respectively. The pentamer differed by 1903 Da, which was likely due to associated DDM micelles. Overall, the detected masses matched the theoretical masses well.

The native mass spectrum at pH 6.0 (**Figure 2.3b**) reveals a different distribution of oligomers. Only AcrA^{L} monomers and dimers were observed, with no high order oligomers like at pH 7.4. UniDec analysis revealed the ratio of monomers and dimer was 5:1 unlike the 1:1 ratio seen at pH 7.4.[118] The measured mass of the monomer differed to the theoretical mass by 3 Da, whereas the measured dimer differed by -27. This is in contrast to the findings by Tikhonova *et al.* (2011), which suggested AcrA^{L} exists as dimers at pH 6.0. Instead, AcrA^{L} is likely an equilibrium of monomers and dimers at pH 6.0.

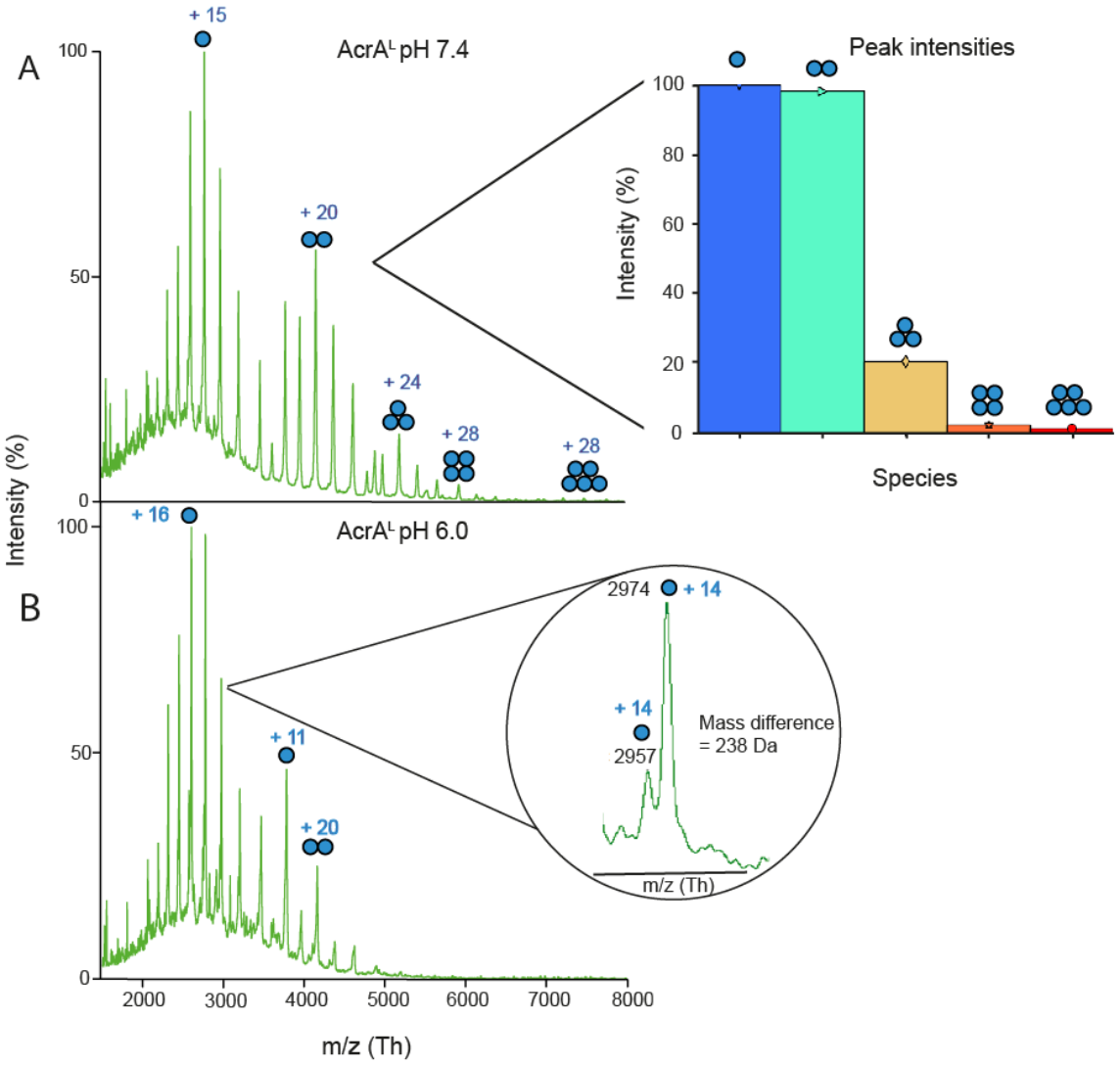


Figure 2.3. Native MS of AcrA^{L}. A. Native MS characterisation of AcrA^{L} at pH 7.4. Proteins were buffer exchanged in 100 mM ammonium acetate buffer prior to MS. AcrA^{L} required the presence of 2 x critical micelle concentration (CMC) of DDM at 0.03%. At pH 7.4 AcrA^{L} presents as a range of oligomers up to pentamers. **B.** Native MS characterisation of the AcrA constructs at pH 6.0. AcrA^{L} presents as a mix of monomers and dimers. Insert shown detailing the mass difference in AcrA^{L} satellite peaks (238 ± 12 Da) representing differences in the lipidation. Masses found in **Table 2.1**.

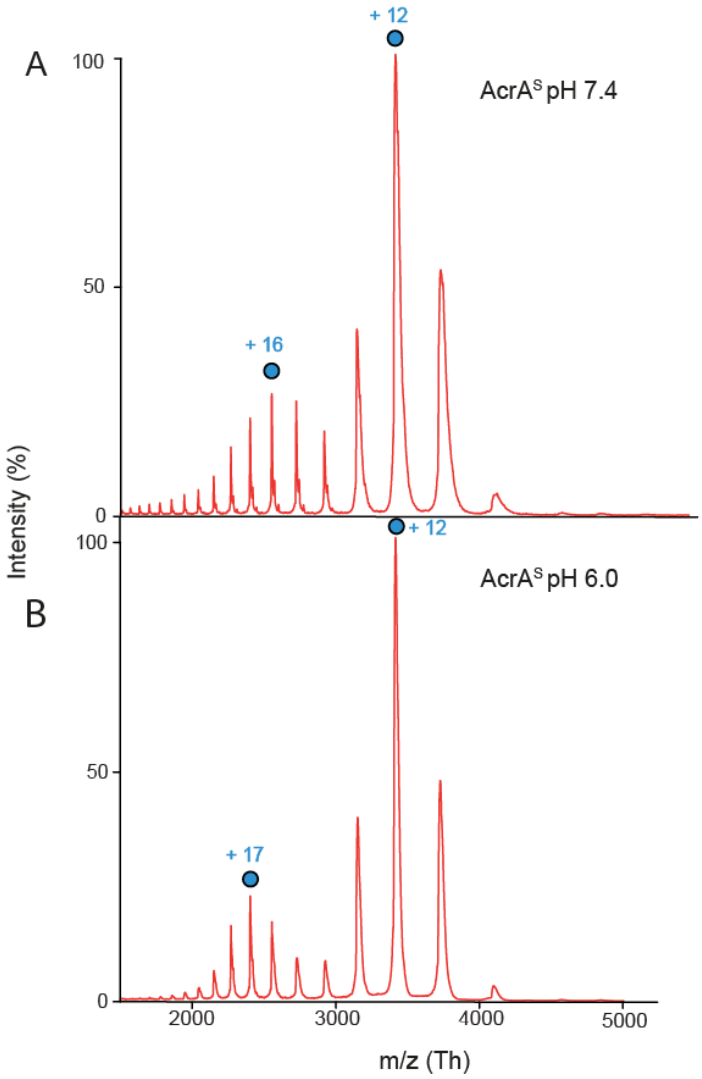


Figure 2.4. Native MS of AcrA^S. A. Native MS characterisation of AcrA^S at pH 7.4. Proteins were buffer exchanged in 100 mM ammonium acetate buffer prior to MS. **B.** Native MS characterisation of the AcrA constructs at pH 6.0. AcrA^S presented as monomers at both pHs. Masses found in **Table 2.1**.

Table 2.1. Native MS masses for AcrA$^{L/S}$. Reported is the standard error of the mean within a single spectrum. Positive mass differences can be attributed to salt and/or detergent adducts.

	Measured mass (Da)	Standard Error (± Da)	Theoretical mass†,* (Da)	Mass difference (Da)
AcrA^{L} pH 6.0 Monomer	41,627	8	41,624	3
AcrA^{L} pH 6.0 Dimer	83,221	2	83,248	-27
AcrA^{L} pH 7.4 Monomer	41,632	7	41,624	8
AcrA^{L} pH 7.4 Dimer	83,274	8	83,248	26
AcrA^{L} pH 7.4 Trimer	124,879	10	124,872	7
AcrA^{L} pH 7.4 Tetramer	166,485	25	166,496	-11
AcrA^{L} pH 7.4 Pentamer	210,023	33	208,120	1903
AcrA^{S} pH 6.0 Monomer	40,849	2	40,817	32
AcrA^{S} pH 7.4 Monomer	40,846	2	40,817	29
AcrA^{S} pH 6.0 (UHMR) Monomer	40,841	2	40,817	24

†*The theoretical masses for AcrA^{L} were calculated for AcrA modified with N-acyl-S-diacylglycerol containing two palmitoyl residues and one oleoyl residue.*

If AcrA^{L} oligomerisation occurred non-specifically, and the affinity of binding between different combinations of oligomers was the same, then the distribution of oligomers would expect to follow a Poisson distribution.[430] A Poisson distribution is a discrete probability function, and can only be used if individual events (such as oligomerisation) happened independently and randomly. Instead, at pH 7.4, a high and equal amount of monomer and dimer was observed, but the amount of trimer, tetramer and pentamer was far lower than what would be observed in a Poisson distribution. This suggests the affinity for the monomer + monomer interaction to form dimer is much stronger than the other interactions exhibited by the other oligomers, hence the very high proportion of dimer in the sample. Interestingly, at pH 6.0, the ratio of dimers is much lower (**Figure 2.3b**). This may have a significant biological effect; in the assembled complex, AcrA forms a trimer of dimers., with the functional unit of AcrA being a dimer.[366,410] Therefore, periplasmic pH may effect AcrA conformationally, and decrease the propensity of AcrA to dimerize. The regulation of AcrA dimerization may help prevent AcrA forming higher order oligomers, which may be detrimental to

complex formation. To draw any definite conclusions, further work is needed, as these experiments did not directly measure binding constants between AcrA oligomers.

Figure 2.4 shows the native mass spectra of $AcrA^S$ at both pH 6.0 and pH 7.4. In agreement with findings by Tikhanova *et al.*(2011) and the SEC traces (**Figure 2.2a**), $AcrA^S$ was observed to be entirely monomeric at both pHs. Therefore, it can be concluded that the lipid moiety increases AcrA's propensity to oligomerise. **Table 2.1** shows the detected masses compared to the theoretical values. Our values differ by 29 and 32 Da for pH 7.4 and 6.0 respectively. This is likely due to salt adducts bound to the protein in the gas phase. Common adducts include Na^+ (23 Da), H_2O (18 Da) and Cl^- (35.5 Da). The SEC and the native mass spectrum combined suggests the $AcrA^S$ sample is entirely monomeric. Therefore, the higher order band seen on the SDS-PAGE gels (**Figures 2.1a and 2.2a**) is unlikely to be due to dimers/oligomers in the buffered, purified sample. This may be an artefact of SDS-PAGE, which has been seen in various studies.[431–435] One possibility for this could be an effect of SDS in the sample buffer, causing a small amount of dimerization to the sample, which has been seen for other proteins.[434,435] This highlights the importance of using several techniques when characterising a protein sample, to ensure artefacts caused by certain techniques can be accurately identified.

2.2.2.2 Charge state distributions suggest intrinsic disorder in AcrA

Aside from gaining information on AcrA stoichiometry, native MS can reveal information on protein structure. nESI transfers proteins to the gas phase with an array of protons, and MS separates these due to their different charges.[121] Desolvation gives rise to charge state distributions (CSDs) that resemble a Gaussian distribution (see **section 1.3.4**). The CSDs observed are strongly influenced by global protein compactness at the moment of transfer to the gas phase.[436] Proteins with a lower degree of compactness have wider CSDs as they have more solvent accessible sites available with high proton affinity, and more compact, folded proteins present as more compact CSDs.[121,437] This analysis offers a valuable insight into protein conformational properties.

Figures 2.3 and 2.4 show AcrA monomers to have two different CSDs for both the lipidated/non-lipidated samples at both pH's. This feature is harder to see and not labelled for the $AcrA^L$ at pH 7.4 due to overlapping dimer peaks. Therefore, for clarity, only the $AcrA^S$ spectra will be discussed here. $AcrA^S$ has a lower charged, narrow CSD at ~3500 *m/z* centred around a +12 ion charge at pH 6.0 and 7.4. It also has a higher charged, wider CSD at ~2500 *m/z* centred around a +16 and +17 ion charge for pH 6.0 and 7.4 respectively. This type of spectra is characteristic of proteins that contain intrinsic disorder. They exhibit wide CSDs with multiple Gaussian distributions which each represent different conformers of the protein in the gas phase.[120] Furthermore, the lower charge CSD ion peaks are

significantly more intense than the higher charge CSD ion peaks which suggests AcrA^{S} has both structured and disordered regions. Characterisation by HDX-MS was needed to further investigate AcrA structural dynamics (see **section 2.2.3.2**).

To ensure the CSDs observed were not artefacts of MS detection or protein purification, a fresh purification of AcrA^{S} was measured by native MS using an orbitrap system (see **section 1.2.4.3**) (**Figure 2.5**).[92] Orbitrap coupled with ultra-high mass resolution (UHMR) provides a powerful native MS system, closing the gap in upper mass limit to time of flight (ToF) MS whilst providing excellent desolvation and ion cooling capabilities.[119] The same CSDs were observed for AcrA^{S} at pH 6.0 when ran on The Q-Exactive Plus UHMR (ThermoFisher Scientific) as observed on the Synapt G2-Si mass spectrometer (Waters). **Figure 2.5** has been annotated to highlight the two different CSDs, and the distribution resulting from the unstructured conformers has been denoted with an Asterix (*). This provided confidence these CSDs were caused by AcrA's conformations in the gas phase.

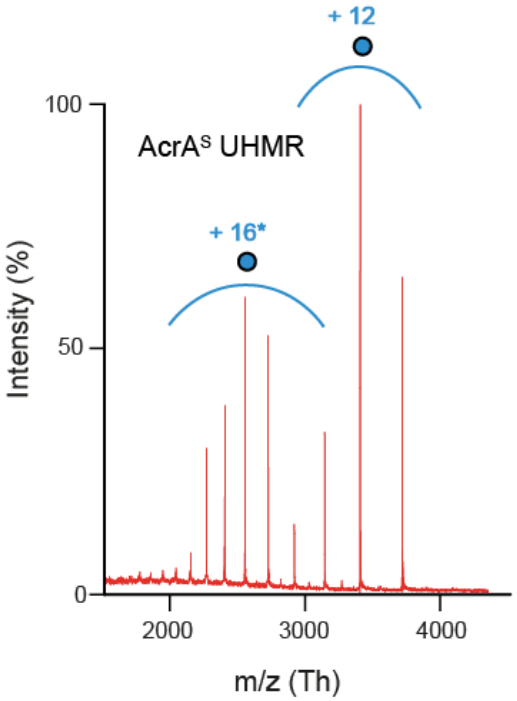


Figure 2.5. The Q-Exactive Plus UHMR native MS of AcrA^{S}. Native MS characterisation of AcrA^{S} at pH 6.0. Proteins were buffer exchanged in 100 mM ammonium acetate buffer prior to MS. AcrA^{S} presented as entirely monomeric with two charge state distributions (CSDs). Envelope marked with an * represent the higher charge state indicative of intrinsic disorder. Masses found in **Table 2.1**.

Interestingly, only AcrA monomers present with multiple CSDs. Focusing on AcrA^{L} (**Figure 2.3**), the oligomers only present as a single narrow CSDs. This suggests that oligomerisation may be stabilising the regions of intrinsic disorder within AcrA. This may be a partial function of dimerization, and by

default the lipidation, to help stabilise AcrA's disordered regions. This may have an impact on its affinity for its binding partners AcrB and TolC.

2.2.2.3 Measuring the AcrA lipidation

AcrA^{L} shows as a double band ~40 kDa on SDS-PAGE gels (**Figure 2.1b**). The native MS results suggest this could be due to differences in the lipid moiety. **Figure 2.3b** contains an insert that zooms in on one of the peaks. It shows the main protein peak has a smaller satellite peak to the left. This was representative of all the peaks for the Gaussian distribution of the protein species, and it was also present in the sample at pH 7.4; this was not shown due to overlapping peaks in this spectrum, so it was less clear to see. The difference in mass between the main peaks in the CSD depicted in **Figure 2.3b** and the satellite peaks was 238 ± 12 Da. This represents the mass of a palmitoyl group (239 Da), a component of the lipidation attached to AcrA.[438] Therefore, the double bands are likely differences in the lipidation, and thus differences in the migration through the gel.

This has been seen in previous work with palmitoylated proteins.[439] Wang *et al.*(2018) observed two proteins with similar molecular weight, but only one containing a palmitoylation, being separated on an SDS-PAGE gel.[439] The palmitoylated protein ran lower than the non-palmitoylated protein, likely due to their interactions with the gel. This matches the relative intensities seen in **Figure 2.1b/c** for the AcrA^{L}. The bottom band is far more intense than the upper band. Furthermore, this is supported by the native MS findings; at least 80 % of the sample was fully lipidated according to the measured masses (**Table 2.1**). This provided enough evidence to conclude the double banding at ~40 kDa for AcrA^{L} is due to differences in the lipidation of AcrA. This may be due to the plasmid construct itself. The plasmid used for these purifications of AcrA^{L} contained a T7 promoter. The T7 promoter is adept at expressing high amounts of the target protein.[440] One explanation for this may be the high amounts of target protein produced saturating the *E. coli* palmitoylation machinery, resulting in partial lipidation for some of the sample. If this is the case, it highlights a limitation of using T7 expression systems. One other possibility is due to the nature of palmitoylations themselves; they are a reversible post-translational modification that can hydrolyse over time, so this may explain the two populations at ~40 kDa.[441]

To test the T7 promoter hypothesis, an AcrA^{L} pUC18 construct with the native *acrAB* promoter (Zgurskaya Group, Oklahoma) was cloned into C43(DE3) Δ*acrAB E. coli* cells, and AcrA^{L} expressed and purified using the same protocol as described in **section 2.4.3.1**. Expression through the native promoter is expected to be lower than the T7 promoter; the lower level of AcrA expression may help sample homogeneity as the palmitoylation machinery of the cell is less likely to be overloaded. If this is the case, it can be expected AcrA^{L} purification through the pUC18 construct would produce a single

band at around ~40 kDa. SDS-PAGE of this sample showed a single band at ~40 kDa (**Appendix 1**), suggesting the T7 promoter hypothesis was correct.

2.2.3 Investigating AcrA^S structural dynamics through HDX-MS

2.2.3.1 Optimisation of HDX-MS experiments

The practise of HDX-MS experiments can be challenging, despite its increasing popularity.[442] HDX-MS optimisation is critical to the success and reliability of downstream experiments. Sample preparation is an important first feasibility stage for designing HDX-MS experiments, to extract reliable conclusions *in vitro*. Samples must be homogenous and at a high enough concentration to inject 10-100 pmols into the mass spectrometer, after dilution in deuterated labelling buffer (with >90% final deuterium concentration) and quench buffer.[33] Sample homogeneity is an important consideration because HDX-MS is an averaging technique, so therefore it would be impossible to differentiate the impact of different oligomers on the structural dynamics of AcrA under different conditions. Therefore, the AcrA^S soluble construct was taken forward for HDX-MS investigations. AcrA^S yields a very high protein concentration, is a homogenous monomeric sample at both pH's, and most importantly has been shown to still be functional.[396,397,400,401]

For HDX-MS experiments to provide as much information as possible, protein coverage of digested peptides should be as high as possible. Optimisation the quench conditions is a crucial stage, as quench conditions have a great impact on digestion results and hence downstream operations.[442] Quench buffers usually contain denaturants (such as urea or guanidine hydrochloride (GuHCl)), reductants and detergents, to help unfold the protein prior to digestion, all at acidic pH. AcrA^S does not have disulphide bonds, therefore a reducing agent was not required in the quench buffer. **Table 2.2** shows the original five quench conditions trialled in the first tests. Quench tests involved diluting protein with quench buffer to lower the pH to between 2.3-2.5, injecting the sample into the mass spectrometer and observing the chromatogram of peptides and analysing the number of high-quality peptides as determined by ProteinLynx Global Server 2.5.1 (PLGS) and DynamX (following the parameters detailed in **section 2.4.5.2**). The number of peptides yielded for all five conditions was low (<100 high-quality peptides with a score of 6.62 or higher), which inevitably led to low redundancy scores for the amino acids (between 1-3) This was largely due to problems with the protein sample used (the first optimisation tests were before the purification protocol was optimised). However, preliminary results suggested quench buffers containing GuHCl and *n*-Dodecyl-phosphocholine (fos-choline) yielded the most peptides (**Table 2.2**).

Table 2.2. Original quench test trials. Reported are the five quench conditions trialled for AcrA digestion.

Quench buffer	Number of peptides	Coverage (%)	Redundancy
1.2% formic acid	70	89.8	2.37
1.2% formic acid, 1.6 M GuHCl	68	90	2.37
1.2% formic acid, 3.2 M GuHCl	23	52.5	1.35
1% formic acid, 0.1% fos-choline	83	91.6	2.92
50 mM $NaHPO_4$, 0.1% DDM (pH 2.5)	31	69	1.41

Quench buffers containing GuHCl, and fos-choline were taken forward for further quench buffer optimisation tests. Furthermore, these tests were completed on protein sample that was purified using the optimised protocol. **Table 2.3** displays the results of these quench conditions and reports the number of high-quality peptides seen under each condition, the coverage percentage, and the redundancy. All of the results provided higher peptide numbers, higher coverage and higher redundancy compared to the first tests. Quench buffer 2 gave the least number of peptides in these experiments and the lowest coverage at 96.6%, and thus buffers 1 and 3 were combined to provide the final quench buffer composition used for AcrA experiments: 1.6 M GuHCl, 0.1% fos-choline, formic acid, pH 1.9. This was mixed at a 1:1 ratio with the labelled protein to get a final pH of 2.4.

Table 2.3. Second quench test optimisation. Reported are the three final quench conditions trialled for AcrA digestion. Observed peptides and protein coverage shown as reported by PLGS.

Quench buffer	Number of peptides	Coverage (%)	Redundancy
Formic acid, 1.6 M GuHCl	189	99.5	6.33
Formic acid, 3.2 M GuHCl	187	96.6	6.54
Formic acid, 0.1% fos-choline	191	99.0	6.66

Another technical challenge during HDX-MS data acquisition is the minimisation of peptide carryover.[173,443] This phenomenon is observed when some so-called 'sticky' peptides remain in the MS system, often due to their chemical properties such as hydrophobicity. These peptides can remain in various places throughout the MS system, such as the injection syringe barrel, dead volumes within the system, the rotor seal, tubing:tubing connections, and the analytical or trap columns.[443] Unsurprisingly, material sticking to the trap column can be more prone to carryover in the next run; this is due to strong hydrophobic interactions between the peptides and the stationary phase (in the case of this thesis, C18 octadecyl alkyl chains) that can remain in some cases where the organic solvent gradient is not strong enough to elute some peptides. Peptides exhibiting carryover can disrupt HDX-MS analysis, as they exhibit extensive back-exchange, and are eluted in a later run. Therefore, these peptides manifest in the mass spectral data as isotopic doublets consisting of a deuterated peptide feature, and a undeuterated feature caused by the peptide that remained in the system. Often, this signature can be falsely identified as EX1 kinetics (see **section 1.4.2**).[443] To avoid any confusion in the HDX-MS data analysis stage, minimisation of carryover is essential. When undergoing AcrA investigations, some protein samples and conditions led to an increase in carryover in some experiments. **Figure 2.6** displays an example peptide in an AcrA experiment. The spectrum shows an example AcrA peptide from a dataset that exhibited significant carryover. Circled in red, peptide carryover can be seen, and this was representative of >50% of filtered peptides. Therefore, multiple methods were implemented to reduce observed carryover within HDX-MS datasets. Firstly, AcrA samples were buffered exchanged using SEC; buffer exchanging to remove glycerol from storage buffer was always performed using a centrifugal exchange device (Micro Bio-Spin 6, Bio-Rad). However, use of SEC allowed for an additional 'clean-up' stage to remove any sample aggregates. Secondly, was the addition of 0.1% fos-choline to the pepsin wash. Fos-choline reduces carryover by solubilising hydrophobic peptides prone to carryover, and it does not interfere with the analysis of peptic fragments as it elutes out much later than most peptides.[442] It was ensured that the cleanblank had no fos-choline present, to help remove excess detergent from the MS system between runs. The final implementation was the increase in pepsin washes between runs, from two to three. The bottom spectrum in **Figure 2.6** shows the same representative peptide after the optimisation steps had been added to the HDX-MS protocol and highlights the significant reduction in observed carryover.

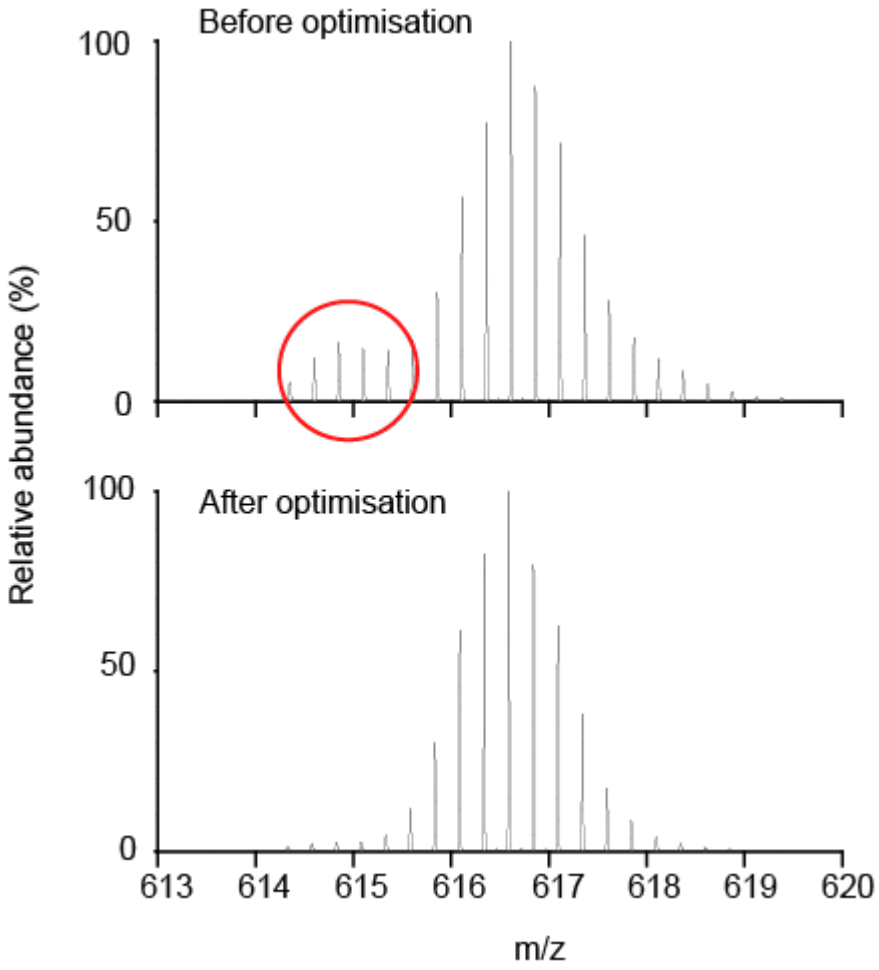


Figure 2.6. Reducing peptide carryover in HDX-MS experiments. Representative peptide chosen from two datasets with the same conditions. The first spectrum was before optimisation steps, the second was after. Representative peptide had a retention time of 5.9 mins. Red circle highlights the envelope caused by peptide carryover. This was significantly reduced after optimisation steps were implemented.

2.2.3.2 Structural dynamics of AcrA at pH 6.0

As discussed in the chapter introduction (**section 2.1**), the periplasm is often more acidic than the cytosol, and it has been hypothesised that pH could be critical to modulating AcrA.[2,252] Since AcrA exists in the periplasm, to investigate it in a more physiological environment, HDX-MS experiments were performed at pH 6.0. The rate of HDX is strongly influenced by pH (see **section 1.4.2**). Therefore, labelling times had to be modified using **Equation 1.15** (see **sections 2.4.5.2.2, or 1.4.2**). This elongated the labelling times compared to the equivalent HDX at pH 7.4. The labelling times adjusted between pH 7.4 and pH 6.0 are shown in **Table 2.4**. These three labelling timepoints were used to investigate the structural dynamics over time, and represented 10s, 1 minute, 10 minutes at pH 7.4.

Longer time points at pH 6.0 would lead to the protein being labelled at room temperature for many hours/days, leading to protein aggregation and a decrease in the accuracy of results. AcrA^S was buffer exchanged into equilibration buffer (50 mM $NaHPO_4$, 150 mM NaCl, pH 6.0) and was deuterated in labelling buffer (50 mM $NaHPO_4$, 150 mM NaCl, pD 5.6) for the three mentioned time points in quadruplet (n=4). The sample was then quenched 1:1 with quench buffer (1.6 M GuHCl, 0.1 % fos-choline, formic acid, pH 1.9) to lower the final pH to 2.4 and injected into the MS for liquid chromatography (LC)/MS analysis.

Table 2.4. Labelling time adjustments between pH 6.0 and pH 7.4. Reported are the three labelling time points at both pH 6.0 and pH 7.4. Temperature was constant at 20 ˚C.

HDX reaction conditions	
pH 6.0 (seconds)	**pH 7.4 (seconds)**
251	10
1507	60
15071	600

Figure 2.7 shows a coverage map for the 125 peptides identified for AcrA^S at pH 6.0, post-analysis. AcrA^S coverage was 95.3% with a redundancy of 4.27. Peptide redundancy refers to the number of detected peptides that describe the deuterium uptake of an amino acid residue, averaged over the whole protein.[33] Most of the protein has multiple peptides per region, which suggest efficient protein digestion. The few gaps in coverage appear between residues 217-222 (αβ-barrel domain), 342-347, 361-362 and 380-385 (MP domain).

The aim of this HDX-MS experiment was to observe the first look at AcrA^S structural dynamics. Since this HDX-MS experiment was only looking at AcrA in a single state, to most accurately characterise AcrA^S structural dynamics, a back-exchange correction was applied. Therefore, maximally labelled (MaxD) AcrA^S samples were obtained as described in **methods 2.4.5.2.3**, and a back-exchange correction (see **section 1.4.3.1**) was applied to the AcrA^S deuterium uptake data at pH 6.0. Only common peptides between the AcrA^S pH 6.0 dataset and the MaxD dataset were kept. **Figure 2.8** displays the RFU heatmap of AcrA^S for all timepoints.

Figure 2.7. Coverage map of AcrA^{S} at pH 6.0. 125 peptides post analysis, resulting in 95.3% coverage and a 4.27 redundancy.

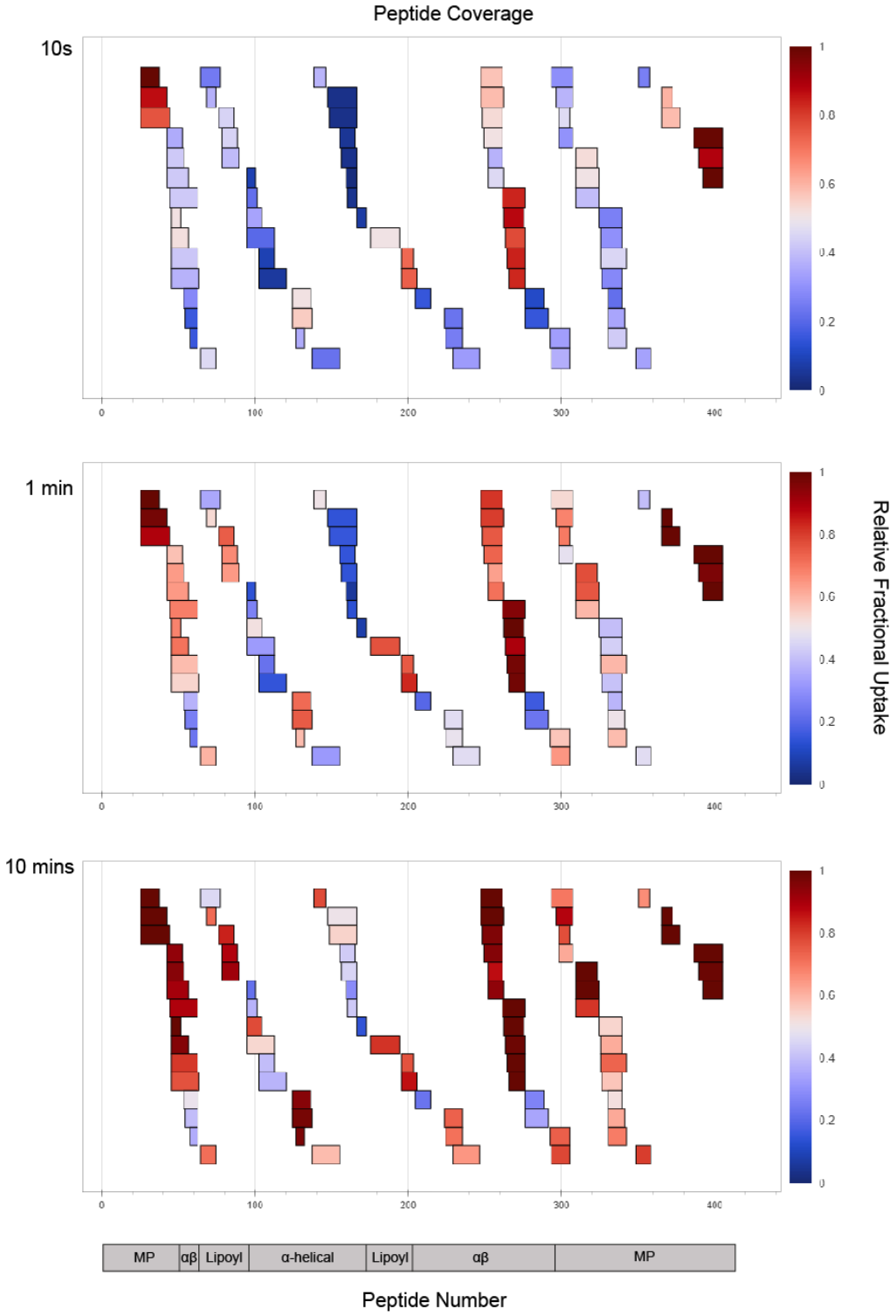


Figure 2.8. Relative fractional uptake of AcrA^S at pH 6.0. Relative fractional deuterium uptake (RFU) analysis of AcrA^S at pH 6.0 for three time points (pH 7.4 corrected). RFU was normalized using a MaxD control for AcrA^S (see **section 2.4.5.2.3**). Areas which take up near-maximal deuteration at the earliest timepoints (10s) are indicative of protein regions which have no measurable secondary structure and are likely intrinsically disordered. Plotted using PyHDX.[444]

A relative fractional uptake (RFU) analysis of AcrA^S reveals areas with time-dependant exchange, characteristic of a folded protein with differences in secondary structure and dynamics. The α-helices show a strong level of protection, and hence a low RFU throughout the entire HDX time course, suggesting this is the most structurally stable area of AcrA^S. Portions of the αβ-barrel domain also show strong protection across all time points, indicating further stable regions within the protein. Interestingly, areas of the MP domain at the N- and C-termini show close to maximum RFU at the earliest timepoint. This is indicative of unstructured regions, as a lack of stable secondary structure means HDX can occur rapidly.[445] This suggests it is the MP domain that gives rise to the characteristic CSDs seen in the native MS spectra of the AcrA constructs (**Figures 2.3, 2.4, 2.5**), and this domain is capped by unstructured regions on either side.

To evaluate the structure of AcrA^S further, **Figure 2.9** shows the AlphaFold2 predicted structure, to compare this to the HDX heat map.[30,31] AlphaFold2 provides a per-residue confidence score (pLDDT) between 0-100 for each residue and regions with a score of <50 may be unstructured. Regions 1-36 and 379-397 are both in the MP domain and contain many residues with a pLDDT score <50. Furthermore, the predicted structure shows these regions as an exposed strand with no secondary structure. This agrees with the experimental MS results that the MP contains unstructured regions. Therefore, AcrA can be classified as a folded protein with areas of intrinsic disorder. This likely benefits AcrA functionally as an MFP. One of the unstructured regions is at the N-terminus of AcrA, where *in vivo* it is attached to the inner membrane via a lipidation. The unstructured nature of this strand may allow AcrA to be flexible and stretch in response to changes in the periplasm, which is a dynamic environment that can change size under different conditions. AcrA's ability to be dynamic and keep a sealed channel under changing conditions is critical for the function of the AcrAB-TolC pump.[251]

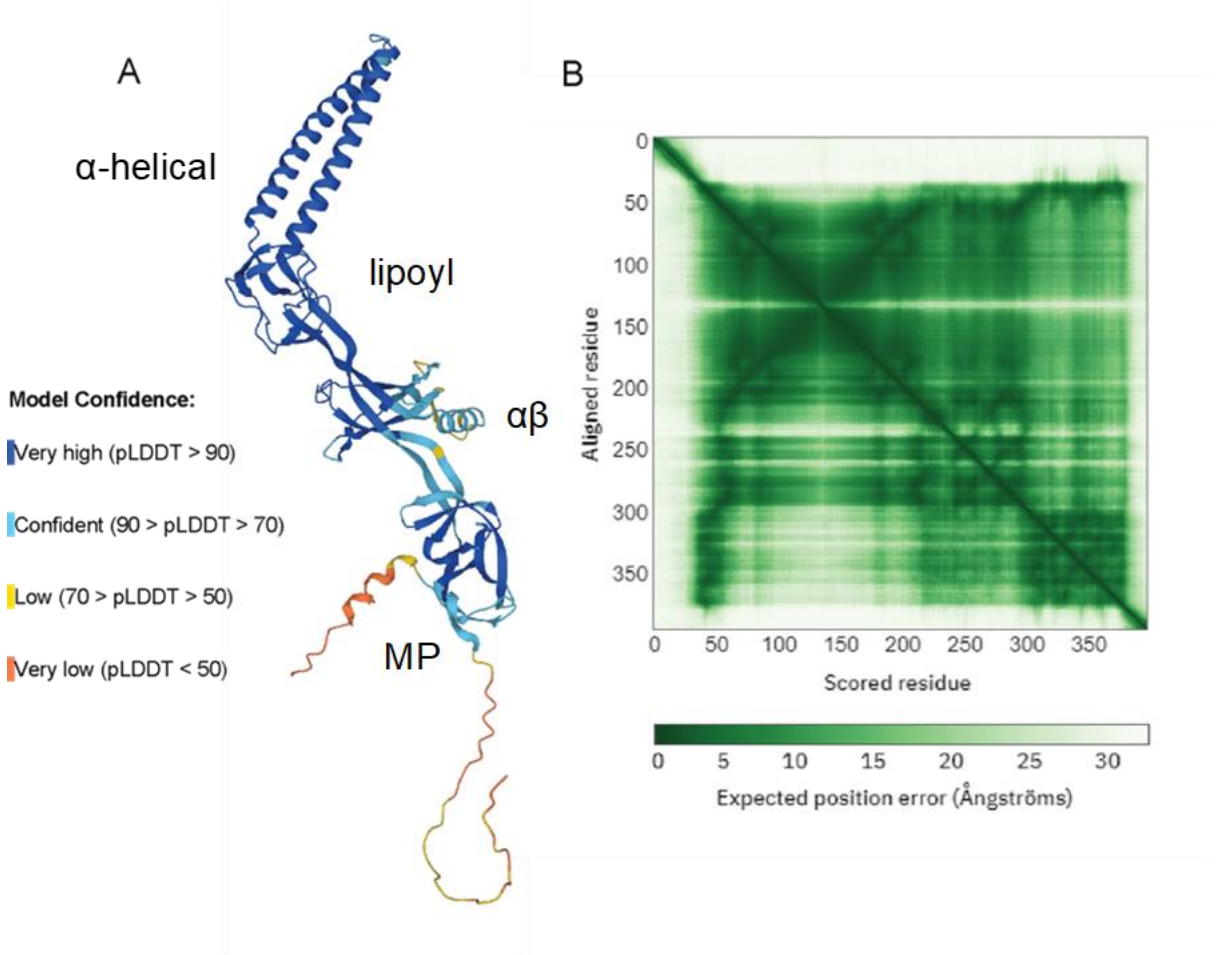


Figure 2.9. AlphaFold2 prediction of AcrA structure. A. Predicted structure of AcrA. Colour coded based on the per-residue confidence score (pLDDT). A very low confidence score (<50) can indicate unstructured regions. **B.** Predicted aligned error plot. The colour position at (x, y) indicates Alphafold's expected position error at residue X, when the predicted and true structures are aligned on residue Y. The colour bar represents how confident AlphaFold2's prediction is – dark green is high confidence; light green is low confidence.[30,31]

2.2.3.3 Comparing AcrA^{S} structural dynamics at pH 6.0 and pH 7.4

The periplasm can exhibit a dynamic range of pH values, so it is important to investigate the effect of cytosolic and acidic pH on AcrA. Previous work has hypothesised that an acidification of AcrA's environment causes conformational changes.[1,2] To investigate whether AcrA's structural dynamics are different at pH 6.0 compared to cytosolic pH (~7.4), ΔHDX was performed. When performing ΔHDX, statistics are essential to ensure any differences seen have biological significance and are not due to chance. Since the experiments in this work do not contain biological replicates, to ensure confidence in drawing conclusions from the data, a conservative ΔHDX significance cut-off was set at 0.5 Da, which is larger than the confidence interval (CI) calculated according to the **methods 2.4.5.2.1**, which is based off the standard deviation (SD) of deuterated peptides.[446] Thus, only peptides exhibiting a 0.5

Da difference between two states passed this cut-off. Therefore, peptides that passed this cut-off and a Welch's *t*-test were deemed significant. All conditions were kept the same except the pH of the HDX buffers, and the pH of the quench buffer was modified to ensure the pH of the quenched sample at pH 7.4 was reduced to same final pH (2.4) when added 1:1 (vol/vol). To compare across states accurately, the labelling times must match according to **Table 2.4** (**methods 2.4.5.2.2**). Therefore, to match the pH 6.0 data, HDX data of $AcrA^S$ at pH 7.4 was collected at three time points (10s, 1min, 10mins). ΔHDX data was analysed as described in the methods.

Figure 2.10 displays the differences in $AcrA^S$ between the two pH's. **Figure 2.10a** is a chiclet plot which shows the ΔHDX for $AcrA^S$ pH 6.0 - $AcrA^S$ pH 7.4, for the three time points investigated. Acidification increases backbone HDX throughout $AcrA^S$ at the latest time point, suggesting a protein wide increase in structural dynamics. The earlier time points show less difference which may suggest that the overall protein fold is not significantly influenced, as secondary structure loss would likely lead to increased exposure of backbone amides to rapid and extensive HDX. There was a local area of reduced dynamics observed (peptides ranging from residues 223-246) within the αβ-barrel domain which shows decreased HDX at the first or second time points. HDX protection seen in this portion of the αβ-barrel domain may be due to secondary structure stabilisation and/or reduced solvent accessibility due to increased local hydrogen bonding arising from conformational changes of $AcrA^S$ at more acidic pH. Peptide 127-136 in the α-helical domain shows interesting changes across the HDX time course. It is protected at the earliest time point, but then becomes deprotected at the latest time point (**Figure 2.10**).

Interestingly, the previous MD simulations performed by Wang *et al.*(2012) were fitted to two different models based on the different orientations of the MP domain, akin to the two conformations identified by Hazel *et al*. (2019) (*cis* and *trans* conformations).[2,398] Model 1 resembled the *cis* conformation, with the MP domain in line with the αβ-barrel and lipoyl domains, whereas in model 2 (resembling the *trans* conformation) the MP domain freely rotates 85˚. Interestingly, both models showed contrasting simulation results at acidic pH. In model 1, protonated His285 gained another hydrogen bond to Asp86 in the lipoyl domain, strengthening the local hydrogen bond network and preventing interdomain movements, which caused increased and decreased Cα root mean squared fluctuations (RMSF) within specifics parts of AcrA's structure. However, model 2 observed that this hydrogen bonding network was weaker, and interdomain movement and backbone dynamics increased in most regions of the protein, matching the HDX results in **Figure 2.10**. The HDX-MS results supports the building evidence that pH has a significant role in regulating the backbone dynamics of AcrA. It is possible that AcrA favours a more *trans*-like conformation where the MP domain undergoes large movements away from the axis of the lipoyl and αβ-barrel domains at pH 6.0, and this may

perturb backbone hydrogen bonds between, and within, domains. However, a deeper investigation is needed to elucidate the exact mechanism of pH activation of AcrA, but it can be concluded that acidification has a significant effect on AcrA dynamics.

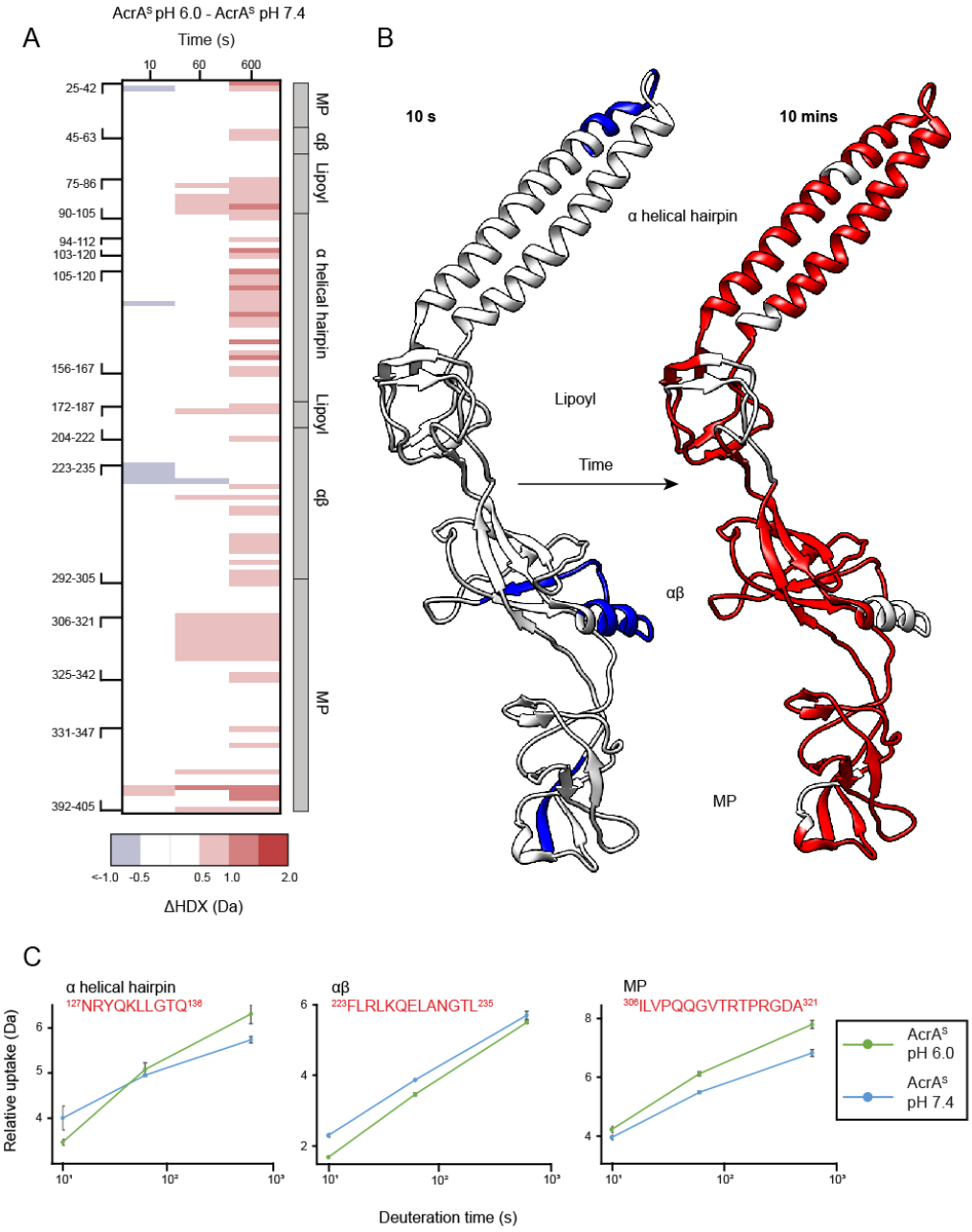


Figure 2.10. Structural dynamics of AcrA at pH 6.0 vs pH 7.4. A. Chiclet plot displaying the differential HDX (ΔHDX) plots for AcrA^S pH 6.0 - AcrA^S pH 7.4 for all time points collected. Blue signifies areas with decreased HDX between states and red signifies areas with increased HDX. Significance was defined to be ≥ 0.5 Da change with a *P*-value ≤ 0.01 in a Welch's *t*-test (n=4). White areas represent regions with insignificant ΔHDX. **B.** ΔHDX for the earliest and latest time point is painted onto the AcrA structure (PDB:5O66) using HDeXplosion and Chimera.[447,448] **C.** Uptake plots for three peptides in different domains of AcrA. Uptake plots are the average deuterium uptake and error bars indicate the standard deviation.

2.2.4 Investigating the effect of magnesium on AcrA^S

Another core aim of this chapter was to investigate if Mg^{2+} has any effect on the structural dynamics of AcrA^S. The functions of the AcrA and other MFPs are well documented, and there is enough evidence to show they are essential components to the efflux function of resistance nodulation and cell division (RND) complexes.[366,386,387,400,403,427,428] Therefore, there has been several hypotheses on the mechanism of action of these proteins, such as pH activation as discussed in **section 2.2.3.3**. Some homologous MFP's from the RND superfamily have been shown to interact with divalent cations, to induce conformational changes essential for function.[404,449] However, it is uncertain whether other MFPs such as AcrA binds divalent cations.[403] Due to the abundance of Mg^{2+} in the *E. coli* periplasm, and the uncertainty around AcrA binding ligands, this section aims to investigate if Mg^{2+} affects AcrA.[1,3,252]

2.2.4.1 Magnesium is predicted to bind AcrA^S

The binding of Mg^{2+} to proteins has been well characterised.[450] At the physiological pH range, Mg^{2+} is coordinated to six H_2O molecules with a large hydration energy. As a hard metal, Mg^{2+} is prone to bind oxygens from the side chain carboxylate groups in Glu and Asp residues (these residues in AcrA are highlighted in **Figure 2.11a**). There are several classes of recognised Mg^{2+} binding sites that can be found across classes of enzymes that use Mg^{2+} as a co-factor which all commonly share an abundance of Glu and Asp residues.[450] To investigate whether AcrA was possibly metal-binding competent, *Me*Bi*Pred* was first used to predict whether AcrA could bind a variety of mono/di/trivalent metal cations, including Ca, Co, Cu, Fe, K, Mg, Mn, Na, Ni, and Zn; only Mg was identified as a likely binder (**Table 2.5**).[451] Interestingly, other MFPs were investigated, and several were shown to bind Mg^{2+}, suggesting this may be more general across RND associated MFPs and not just AcrA specific. Based on this information, the aim was to search for possible Mg^{2+} binding sites within AcrA, and the *Metal Site Hunter* revealed a possible Mg^{2+} binding site was revealed within the αβ-barrel domain (**Figure 2.11**).[452]

Table 2.5. MeBiPred predictions of cation binding in MFPs

			Metal										
Protein	**Organism**	**Residues**	*Ca*	*Co*	*Cu*	*Fe*	*K*	*Mg*	*Mn*	*Na*	*Ni*	*Zn*	
AcrA	*E. coli*	25-397	0.25	0.31	0.08	0.07	0.29	**0.58**	0.21	0.46	0.22	0.23	**Lipidated**
MdtE	*E. coli*	21-385	0.16	0.23	0.00	0.00	0.36	**0.58**	0.14	0.41	0.23	0.17	
AcrE	*E. coli*	24-385	0.14	0.30	0.18	0.28	0.46	**0.58**	0.20	**0.52**	0.27	0.40	
MexA	*P. aeruginosa*	24-383	**0.66**	0.41	0.06	0.00	0.38	**0.54**	0.45	0.49	0.17	0.31	
CusB	*E. coli*	29-407	**0.74**	0.23	**0.68**	0.18	0.06	0.47	0.45	0.42	0.00	0.11	
MdtA	*E. coli*	22-415	0.29	0.37	0.13	0.36	**0.53**	0.20	0.39	0.47	0.48	0.21	
MdsA	*S. enterica*[#]	26-408	0.27	0.31	0.27	0.42	0.38	0.49	0.22	0.40	**0.50**	0.46	
ZneB	*C. metallidurans* †	37-385	**0.63**	0.34	0.26	0.40	0.43	0.33	0.13	0.46	0.43	**0.49 ***	

(**Green box**) = above 0.5 threshold as recommended in Aptekmann *et al.* (2022).[451]
[*]Confirmed Zn^{2+} binder and close to threshold
[#]*Salmonella enterica* serovar Typhimurium
†*Cupriavidus metallidurans*

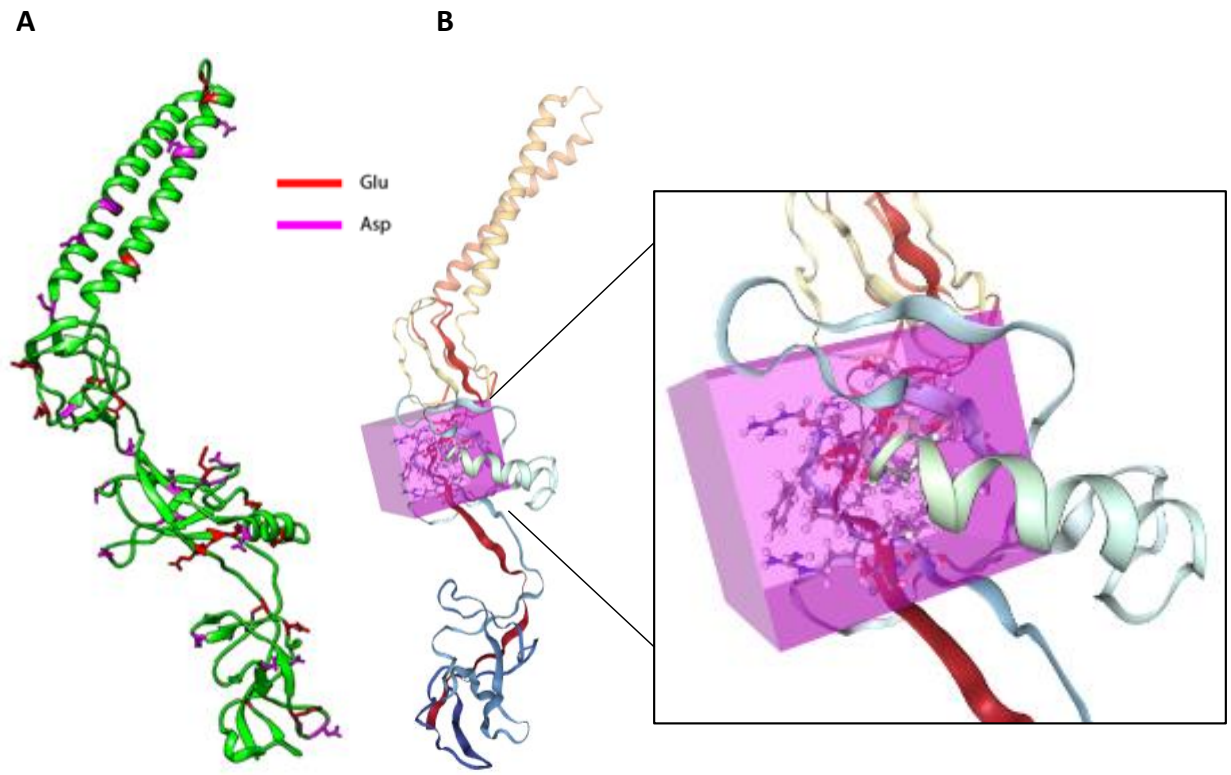


Figure 2.11. Putative Mg^{2+} binding site in AcrA. A. Mapping the Asp and Glu residues on AcrA. Asp residues labelled pink, Glu residues labelled red. **B.** Possible Mg^{2+} binding site revealed using *Metal Site Hunter*.[452] Pink box highlights site location within the αβ-barrel domain.

2.2.4.2 Circular dichroism investigations of $AcrA^S$ and magnesium

To investigate whether Mg^{2+} had any effect on $AcrA^S$ secondary structure, circular dichroism (CD) spectroscopy scans of $AcrA^S$ ± Mg^{2+} were compared. CD is great for rapidly evaluating the secondary structure, folding and binding properties of proteins.[453] This is because when the chromophores of the backbone amides align in arrays, their optical transitions are split or shifted due to 'excitation' interactions. This means each secondary structure element has a characteristic CD spectrum. Therefore, $AcrA^S$ was buffer exchanged into buffer (50 mM $NaHPO_4$, 150 mM NaCl, ± 1 mM $MgCl_2$, pH 6.0) and glycerol was removed as it interfered with the measurement. Protein samples were diluted to 0.32 mg/mL (7.8 µM) and a 0.05 cm pathlength coverslip was used. Full length scans (280-185 nm) were completed in triplicate. The data was provided in measured ellipticity (mdeg) and needed to be converted into mean residue ellipticity for analysis of secondary structure (see **methods 2.4.6**). The data were analysed using the BeStSel online program, with the analysed scans given in delta epsilon (M^{-1} cm^{-1}).[454] **Figure 2.12** shows the scans of $AcrA^S$ ± Mg^{2+}.

The CD scans of $AcrA^S$ show slight differences with and without Mg^{2+}. In the $AcrA^S$ CD spectrum with Mg^{2+} the portion between 230-250 nm appears tilted further to the left than the other spectrum without Mg^{2+}. This suggests a higher amount of β-sheet in the presence of Mg^{2+}. The BeStSel online program also provided a breakdown of secondary structure elements; this agreed with the observation from **Figure 2.12**. The BeStSel program said there was 12.7% antiparallel and 4.7% parallel β-sheet for $AcrA^S$, but with Mg^{2+} the antiparallel β-sheet increased to 25.2%. The BeStSel program provided root mean-square deviations (RMSD) number for each analysis to describe spectra discrepancies, between the raw and fitted data; these were 0.0561 for $AcrA^S$ and 0.0449 for $AcrA^S$ + Mg^{2+} which is within the accepted range from previous experiments.[454] An increase in β-sheet content was also seen when another MFP ZneB bound Zn^{2+}. Furthermore, in the presence of Mg^{2+} the spectrum is more compact; this suggested a lower α-helical content, due to the fact α-helical contributions dominate the spectrum. Again, this agreed with BeStSel analysis; $AcrA^S$ had an α-helical content of 30.2%, yet in the presence of Mg^{2+} $AcrA^S$ only had 17.4%. Previous work combined with the results in **Figure 2.11** suggests Mg^{2+} could bind in the αβ-barrel domains, making it more likely this result is due to perturbations in the small α3 helix in the αβ-barrel domain rather than the α-hairpins.[403,404,452] BeStSel analysis reports similar contents of turn and other structures for both states.

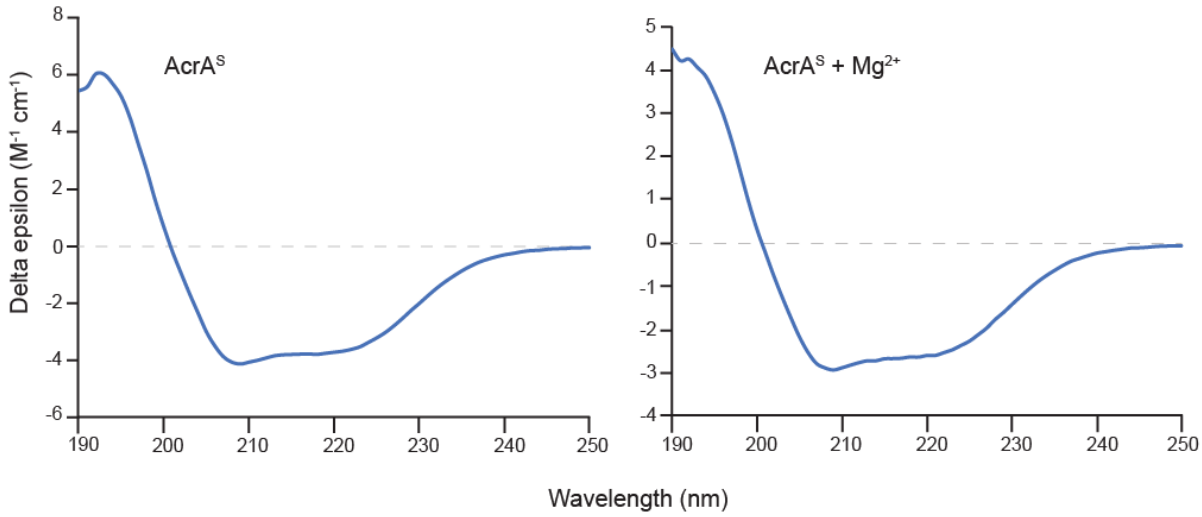


Figure 2.12. Circular dichroism scans of AcrA ± Mg^{2+}. Full length scans of AcrA ± Mg^{2+} at pH 6.0. Only 190-250 nm analysed due to noisy data at the ends. Proteins loaded at 0.32 mg/mL and Mg^{2+} at 1 mM. Each scan was repeated in triplicate, and the averaged scan was analysed. Delta epsilon conversion done by BeStSel online program.[454]

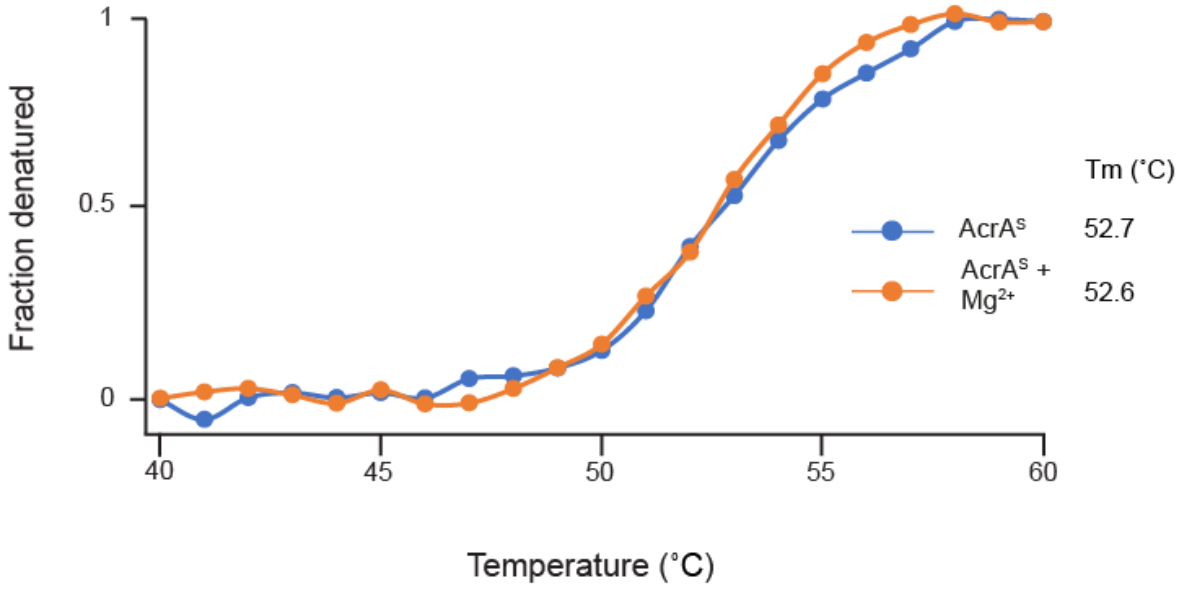


Figure 2.13. Circular dichroism thermal melts of $AcrA^S$ ± Mg^{2+}. Scans at 222 nm taken between 40-60 ˚C with 1 ˚C increments of AcrA ± Mg^{2+} at pH 6.0. Proteins loaded at 0.0075 mg/mL and Mg^{2+} at 1 mM. Plot shows fraction denatured vs temperature. Tms reported are 52.7 ˚C for $AcrA^S$ and 53.6 for $AcrA^S$ + Mg^{2+}.

To see if the presence of Mg^{2+} increased the thermal stability of $AcrA^S$, CD thermal melts (± Mg^{2+}) were performed. CD scans were first completed at 15 temperatures ranging from 30-95 ˚C, to identify where the transition from folded to denatured occurred. Then a second scan was completed, ranging between 40-60 ˚C with 1 ˚C increments, to accurately determine the melting temperatures (Tms) of $AcrA^S$ under both conditions (**Figure 2.13**). $AcrA^S$ was buffer exchanged into the same buffer used for previous CD experiments (**Figure 2.12**) and diluted to 0.0075 mg/mL. The CD value at 222 nm at each temperature point was used to calculate thermodynamic parameters, as described in **section 2.4.6**.[455] **Figure 2.13** shows the calculated fraction denatured (fD) at each temperature, and the calculated Tms. The Tms are the same regardless of Mg^{2+} suggesting Mg^{2+} is not affecting the thermal stability of AcrA. However, the data does suggest Mg^{2+} has an effect on secondary structure at pH 6.0.

2.2.4.3 Investigating the effect of magnesium on $AcrA^S$ structural dynamics

The next task was to investigate whether Mg^{2+} had an effect on the structural dynamics of AcrA. HDX-MS experiments were performed on $AcrA^S$ at pH 6.0 ± Mg^{2+}. The HDX buffers were the same as in **section 2.2.3.3** (and **methods 2.4.5.2**) except for the presence of $MgCl_2$ for the sample with Mg^{2+}. First attempts at this data used 5 mM Mg^{2+} in an attempt to saturate $AcrA^S$ with the divalent cation. However, upon analysis there was far fewer peptides identified through PLGS than usually observed. This could have been caused by several possibilities; one reason for this may have been because of a proposed phenomenon called ion suppression, which can have deleterious effects on ESI.[101] Effectively, the less volatile solutes (i.e. Mg^{2+}) change the efficiency of droplet formation or desolvation, thus a lower number of charged gaseous ions reach the detector, and a lower signal is observed. Another reason may have been due to effects on the protease during digestion caused by the high amount of $MgCl_2$. The most likely reason is that the high salt concentration affected the desalting ability on the trap column, leading to [M + Mg^{2+}] peptides than cannot be accurately identified using PLGS due to the increased mass and would therefore reduce the ion population of the corresponding [M + H^+] peptide. Therefore, an optimised concentration of 1 mM Mg^{2+} was used, which provided high amounts of divalent cation in the sample and a strong peptide signal.

To investigate whether Mg^{2+} influenced the structural dynamics of AcrA, ΔHDX experiments were performed on $AcrA^S$ at pH 6.0 comparing HDX profiles ± $MgCl_2$. **Figure 2.14** displays the ΔHDX of $AcrA^S$ ± 1 mM $MgCl_2$ at pH 6.0. The chiclet plot in **Figure 2.14a** shows Mg^{2+} is having a significant stabilisation effect on $AcrA^S$ throughout the HDX time course, as there is a decrease in deuterium across all four domains. There is stabilisation across all time points in the core of the MP domain, from residues 306-342. **Figure 2.14c** shows the uptake plot for an example peptide in this region (^{309}PQQGVTRTPRGDATVL^{324}), showing ~1 Da difference in HDX across all time points for $AcrA^S$ + Mg^{2+}.

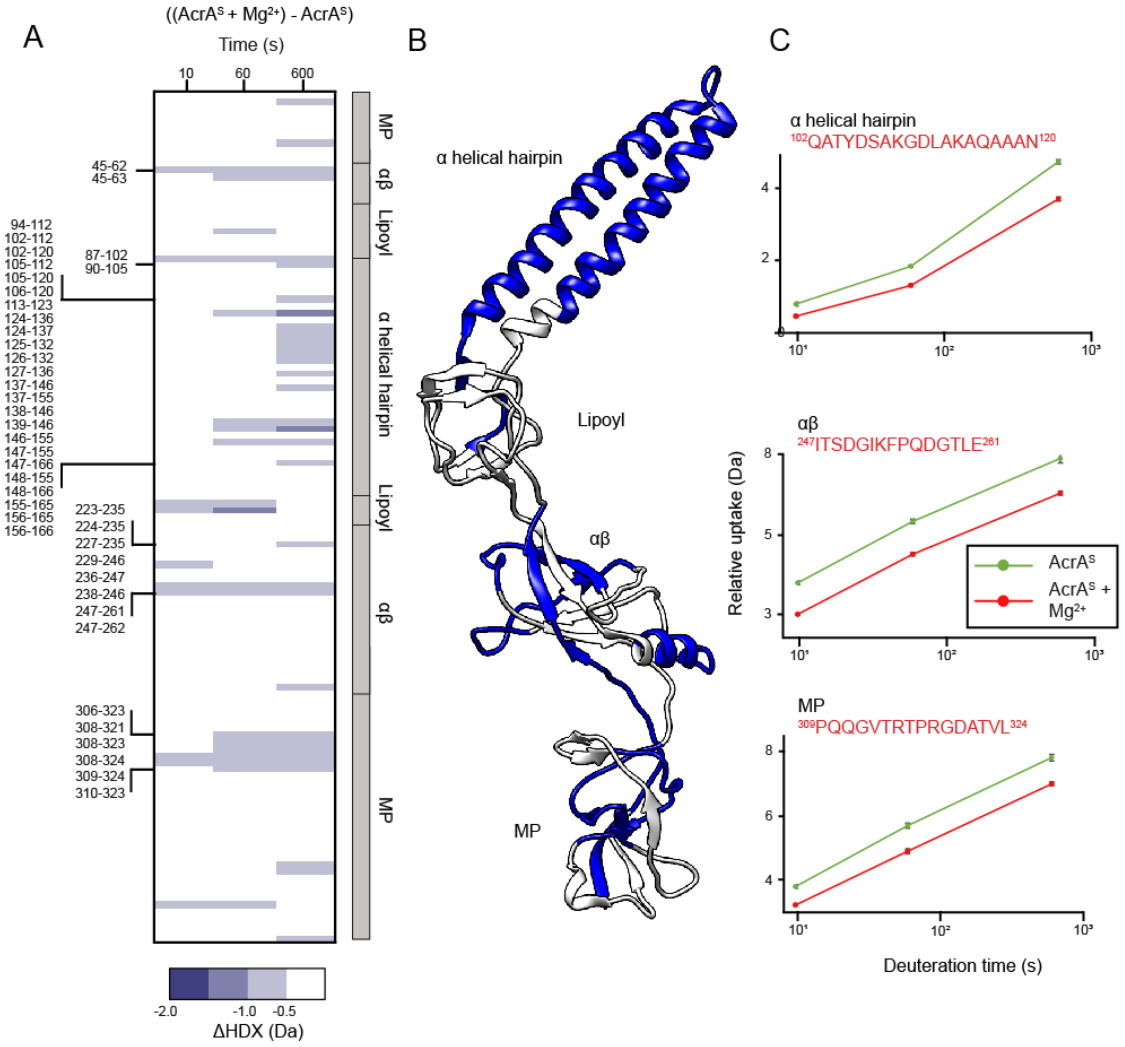


Figure 2.14. Structural dynamics of AcrA and magnesium at pH 6.0. A. Chiclet plot displaying the differential HDX (ΔHDX) plots for ((AcrA^S + Mg^{2+}) - AcrA^S), at pH 6,0 for all time points collected. Blue signifies areas with decreased HDX between states and white signifies areas with no significant change in HDX. Significance was defined to be ≥ 0.5 Da change with a *P*-value ≤ 0.01 in a Welch's *t*-test (n=4). **B.** ΔHDX for the latest time point is painted onto the AcrA structure (PDB:5O66) using HDeXplosion and Chimera.[447,448] **C.** Uptake plots for three peptides in different domains of AcrA. Uptake plots are the average deuterium uptake and error bars indicate the standard deviation.

Furthermore, there is protection in both sides of the αβ-barrel domain for all three time points (residues 42-63 and 223-262); peptide ^{247}ITSDGIKFPQDGTLE^{261} representing this effect (**Figure 2.14c**). Interestingly, there is protection on several flexible linkers between AcrA's domains; regions between the αβ-barrel and MP domains and lipoyl and α--helical domains exhibit reduced HDX. The flexible linkers are key to AcrA's ability to freely position its four domains and change conformations.[366,397] It is likely that the restriction of the flexible linkers is what causes the expansive stabilisation of the α-helical hairpins, as the ability of AcrA to freely position them is hindered (see **Figure 2.14c** peptide: ^{102}QATYDSAKGDLAKAQAAAN^{120}). The α-helices show a protection across most of the domain, whereas the other three domains have large portions with insignificant ΔHDX, suggesting their backbone dynamics haven't been affected by Mg^{2+}. The lipoyl domain has the largest portion with insignificant ΔHDX.

A possible site has been identified within the αβ-barrel domain of AcrA which could be responsible for binding Mg^{2+} (**Figure 2.11**), however it is difficult to directly localize these effects using peptide level HDX-MS. Furthermore, as Mg^{2+} likely binds side chains which are not measured by HDX-MS studies, they will not appear in the data unless it impacts backbone dynamics as well.

Our HDX results agree with a previous investigation with AcrA and Mg^{2+}. Zgurskaya *et al.* (1999) found that adding Mg^{2+} lead to a small but statistically significant difference in the conformation of $AcrA^S$.[396] Dynamic light scattering (DLS), and velocity centrifugation studies saw a 16% shrink in length when Mg^{2+} was added. Furthermore, without Mg^{2+} AcrA had a more variable structure, judged by the heterogeneity of the sedimentary coefficient and polydispersity parameter from DLS, compared to one predominant conformation when Mg^{2+} was added. Therefore, they concluded some regions in the tertiary structure of $AcrA^S$ are more rigid in the presence of Mg^{2+}, complementing the stabilisation effect seen with the HDX results (**Figure 2.14**). However, it is important to note the DLS studies were performed at pH 8.0, whereas the HDX-MS experiments were performed at pH 6.0. Furthermore, Mg^{2+} has been shown to stabilise the dynamics of another periplasmic protein, alkaline phosphatase, through the use of hydrogen-tritium exchange, which supports that it could have a general role in ensuring periplasmic protein robustness.[456] Zgurskaya *et al.* (1999) also observed no effect on AcrA oligomeric state by Mg^{2+}.[396] Native MS analysis supports this showing $AcrA^S$ to be monomeric in the presence of Mg^{2+} (**Figure 2.15**). Increased peaks widths observed suggest Mg^{2+} ions remain bound to the protein in the gas phase; however, stoichiometry could not be resolved with the quadrupole-ToF (Q-Tof) mass spectrometer used.

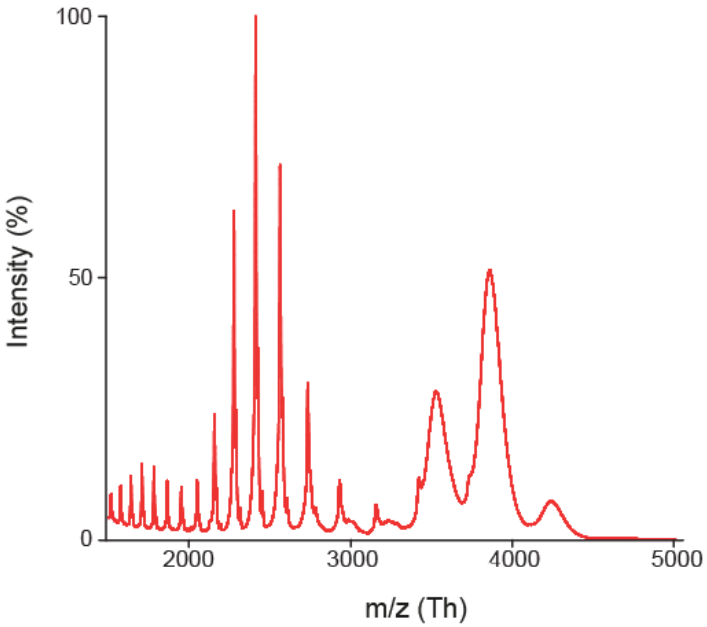


Figure 2.15. Native MS of AcrA^{S} and Mg^{2+}. Native MS characterisation of AcrA^{S} at pH 6.0 with 100 µM $MgCl_2$. Proteins were buffer exchanged in 100 mM ammonium acetate buffer prior to MS. Two monomeric CSD's were observed for AcrA^{S} + Mg^{2+}.

Analysing the coordination of a Zn^{2+} molecule to ZneB, another MFP from the heavy metal efflux (HME) subclass of RND efflux pumps, may assist in interpreting Mg^{2+} effect on AcrA^{S}.[404] In ZneB, Zn^{2+} is tetrahedrally coordinated between a water molecule, H220 (αβ-barrel domain), E328 (MP domain) and H284 (linker between αβ-barrel and MP domains). These regions show significant stabilisation in AcrA^{S}, suggesting Mg^{2+} could be interacting in a similar way within a similar site in the αβ-barrel domain identified in **Figure 2.11**. This would also explain why a global stabilisation is observed; a metal ion bound between three different regions is likely to restrict conformational movement across a protein with four linearly arranged domains. However, it must be noted the residues in the ZneB are His and Glu, as Zn^{2+} is a soft metal and prefers to bind a nitrogen atom from the side chain of His. However, Asp and Glu residues positioned in a similar arrangement may be responsible for coordinating Mg^{2+}. Furthermore, crystal structures of ZneB showed Zn^{2+} only bound in the crescent (*trans*-like) conformation and not in the extended (*cis*) conformation. This suggested Zn^{2+} caused a conformational change in ZneB to a more compact state; similar results have been seen for CusB as well.[449] These conformations are similar to those observed in MexA, which is an AcrA homolog and a MFP from *Pseudomonas aeruginosa* (*P. aeruginosa*), which have rotated/unrotated conformers akin to the *cis*/*trans* conformations of AcrA.[409] Thus it is possible metal binding aids the function of the MFP as it cycles through conformational changes during the pumps rotational mechanism. Furthermore, if

the metal binding site occurs in residues across several domains, conformational changes in the MFP could possibly disrupt the protein-metal interaction.[404]

2.2.4.3.1 The effect of magnesium at pH 7.4

To test whether this Mg^{2+} effect is pH dependant, HDX-MS experiments were repeated ± Mg^{2+} at pH 7.4 (**Figure 2.16**). The AcrA structure shown represents the ΔHDX (($AcrA^S$ + Mg^{2+}) - $AcrA^S$) at the 10 minute time point, equivalent to the labelling time shown for $AcrA^S$ at pH 6.0 (**Figure 2.14**). Unlike at pH 6.0, Mg^{2+} has very little significant effect on AcrA dynamics. At pH 7.4, the stabilising effect of Mg^{2+} at pH 6.0 is not observed. The three peptides shown are the same peptides that exhibited a significant difference in HDX at pH 6.0, and all show no change in the presence of Mg^{2+} at pH 7.4. Due to the experiment being at pH 7.4, a longer 4 hour time point was sampled, to see if Mg^{2+} had an effect on dynamics over a longer period of time; again, there were no observed peptides that saw a change in uptake at this time point (**Appendix 2**). One possibility for this is that AcrA does not interact with Mg^{2+} with the same affinity as seen at pH 6.0. Furthermore, looking at the protonation states of amino acids, the only difference at pH 6.0 is the protonation of His side chains, which is not proposed to commonly bind Mg^{2+} ions; therefore, it may be due to different conformers sampled at the different pH's. If the cation binding site is similar to that seen in other MFPs, where the three residues are from different regions of the protein, then the conformations exhibited at pH 7.4 may not see the residues required for binding come together in the proper arrangement.

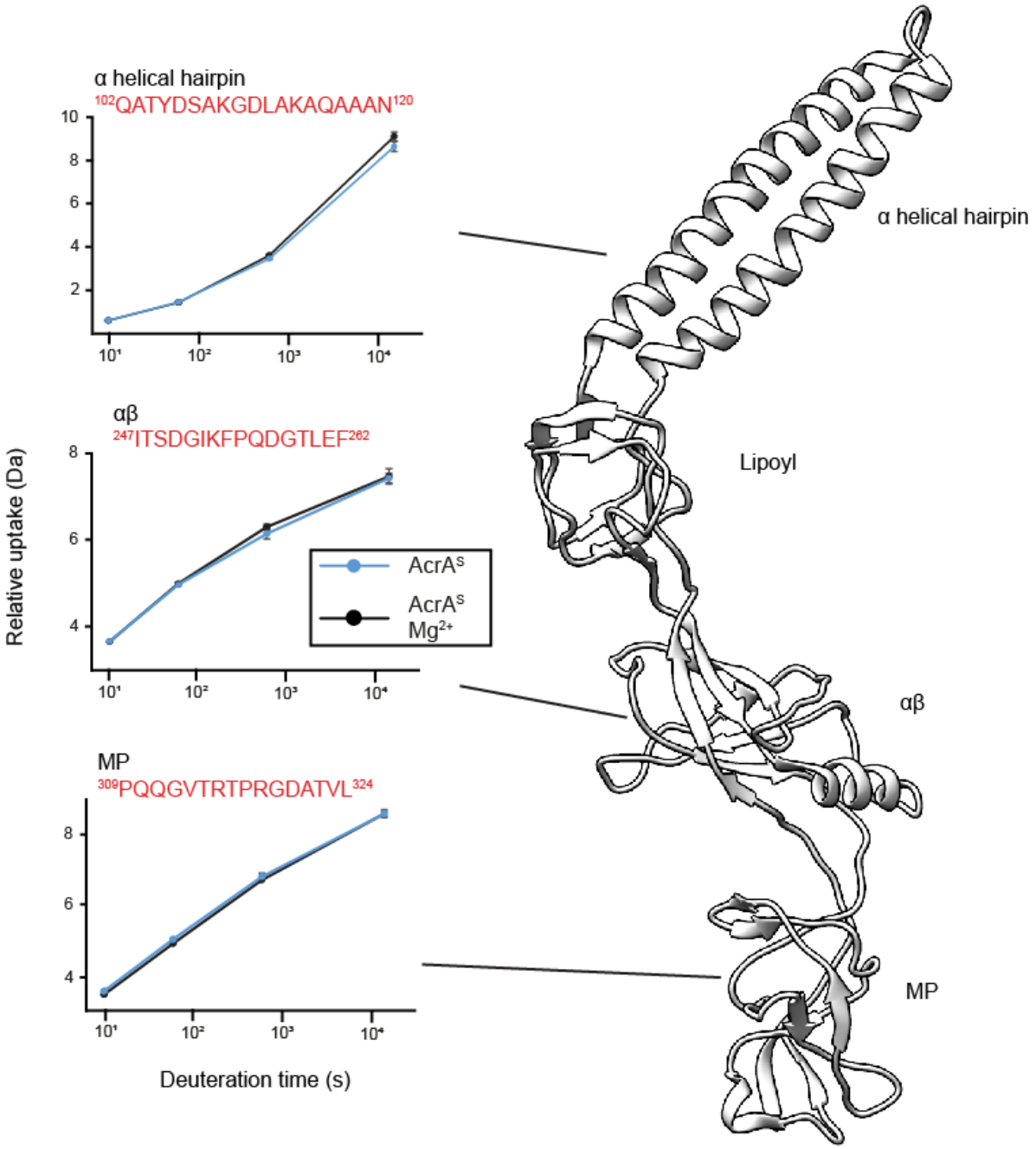


Figure 2.16. Structural dynamics of AcrA and magnesium at pH 7.4. The differential HDX (ΔHDX) plots for (($AcrA^S$ + Mg^{2+}) - $AcrA^S$), at pH 7.4 for the 10 min time point painted onto the structure of AcrA (PDB: 5066) using HDeXplosion and Chimera.[447,448] White signifies areas with no significant change in HDX. Significance was defined to be ≥ 0.5 Da change with a *P*-value ≤ 0.01 in a Welch's *t*-test (n=4)**.** Uptake plots for three peptides in different domains of AcrA. Uptake plots are the average deuterium uptake and error bars indicate the standard deviation.

2.2.4.3.2 Magnesium tempers increased dynamics of $AcrA^S$ at pH 6.0

To investigate the relationship between pH and Mg^{2+} binding further, ΔHDX was performed on (($AcrA^S$ + Mg^{2+} at pH 6.0) - ($AcrA^S$ + Mg^{2+} at pH 7.4)) (**Figure 2.17**), as Mg^{2+} is constantly present in the periplasm throughout the different pH ranges it can experience. It was previously observed that $AcrA^S$ at pH 6.0 had a significant increase in dynamics at the longest time point compared to pH 7.4 (**Figure 2.10**). This analysis reveals that Mg^{2+} largely compensates for the increased dynamics exhibited by $AcrA^S$ at pH 6.0, resulting in insignificant HDX changes across the protein between the different pH regimes. There are far fewer peptides possessing increased HDX in the latest timepoint, and the degree of ΔHDX is much smaller (**Figure 2.17a**). Two peptides shown in **Figure 2.17c** demonstrate this effect (^{137}YISKQEYDQA^{146}, ^{308}VPQQGVTRTPRGDA^{321}). Therefore, the data shows the increased dynamics detected at pH 6.0 is primarily rectified by Mg^{2+} binding. This balance may be key to ensuring that the efflux pump system works and remains stable even when the periplasm is challenged by more acidic environments which Gram-negative bacteria occupy. Therefore, Mg^{2+} may play a role in modulating the structural dynamics of AcrA, enabling it to function across a broad pH range.

Within the ΔHDX analysis between the two pH states in the context of Mg^{2+} ($AcrA^S$ + Mg^{2+} at pH 6.0) - ($AcrA^S$ + Mg^{2+} at pH 7.4) a clear and focused reduction in HDX was observed (**Figure 2.17c**). The largest area is within the αβ-barrel domain (e.g. peptide ^{223}FLRLKQELANGTL^{235}) which is protected across the entire HDX time course). One possibility for this increased stabilisation is that it could be the site of Mg^{2+} coordination. This would agree with the predictions made in **Figure 2.11**, and the stabilisation of this region without Mg^{2+} present may suggest that this site is better primed for Mg^{2+} binding at pH 6.0. The αβ-barrel domain has been previously hypothesised to bind ligands in homologous MFPs, and peptide 223-235 does contain a Glu residue. To draw any certain biological conclusions, further investigations would be needed to decipher the Mg^{2+} binding site, such as targeted mutagenesis and x-ray crystallography.

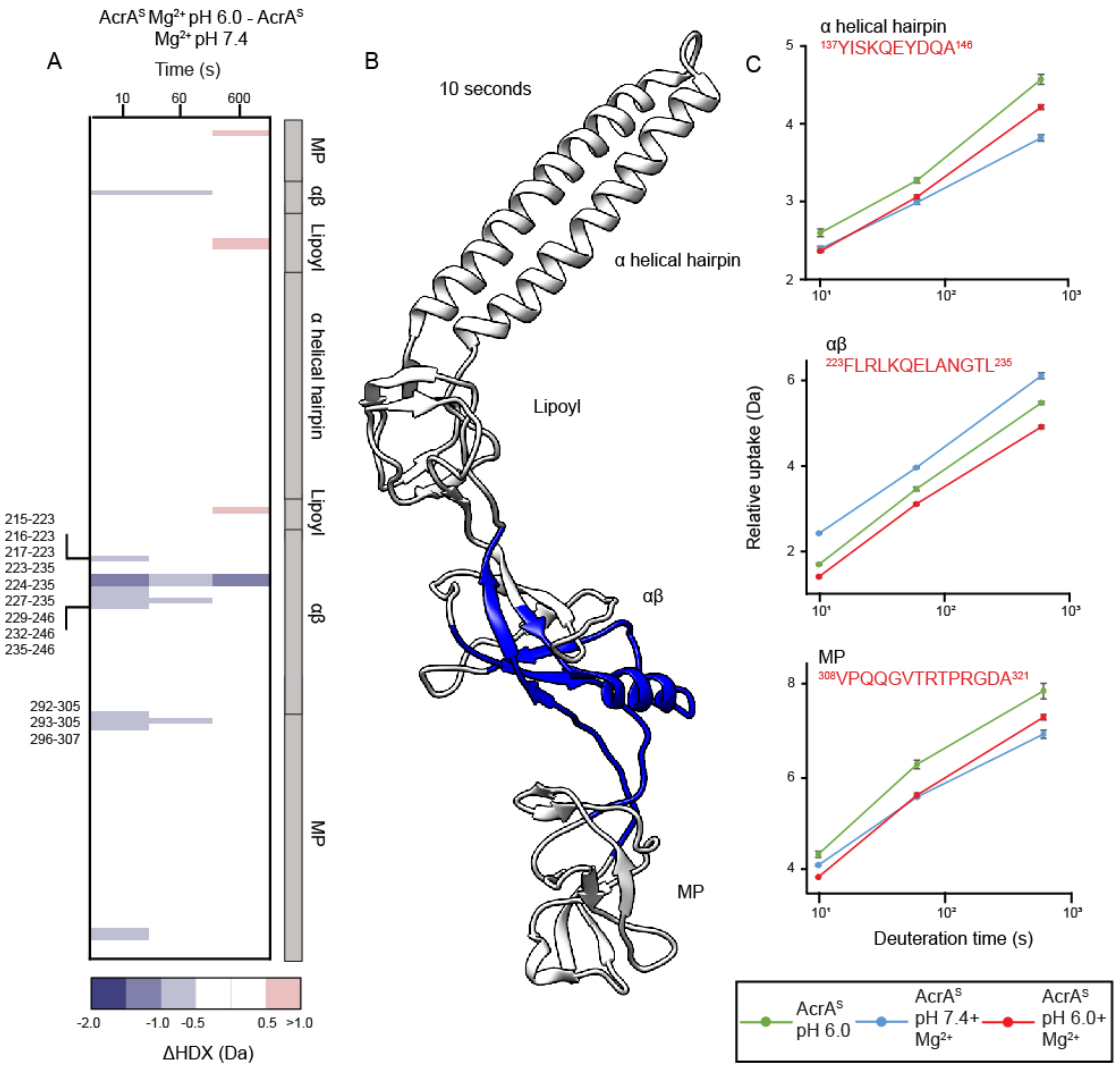


Figure 2.17. Structural dynamics of AcrA + magnesium at pH 6.0 and 7.4. A. Chiclet plot displaying the differential HDX (ΔHDX) plots for ((AcrA^S + Mg^{2+} pH 6.0) - AcrA^S + Mg^{2+} pH 7.4), for all time points collected. Blue signifies areas with decreased HDX between states, red signifies areas with increased HDX between states and white signifies areas with no significant change in HDX. Significance was defined to be ≥ 0.5 Da change with a *P*-value ≤ 0.01 in a Welch's *t*-test (n=4).**B.** ΔHDX for the earliest time point is painted onto the AcrA structure (PDB:5O66) using HDeXplosion and Chimera.[447,448] **C.** Uptake plots for three peptides in different domains of AcrA. Uptake plots are the average deuterium uptake and error bars indicate the standard deviation.

2.3 Conclusions

AcrA is an MFP in the AcrAB-TolC multidrug efflux pump, that forms a trimer of dimers in the assembled complex to maintain a sealed channel from the periplasm, and it binds to both AcrB and TolC.[366,367,410,457–460] Therefore, AcrA has four linearly arranged domains connected by flexible linkers, and must function as a highly dynamic protein to ensure stability and function of the assembled complex.[397] In this chapter, AcrA was extensively characterised using structural MS techniques, to reveal how its lipidation promotes oligomerisation *in vitro*, but this is somewhat regulated by acidic pH, and how AcrA is a folded protein with defined secondary structure and areas of intrinsic disorder. Furthermore, the impact of periplasmic conditions on AcrA backbone dynamics was investigated, revealing that weakly acidic conditions increase backbone dynamics across the whole protein, yet this is largely modulated by the presence of Mg^{2+}.

The results in **chapter 2** suggest that the dynamics of AcrA are regulated by cross play between pH and Mg^{2+}. Although the periplasm is on average ~1.7 pH units more acidic than the cytosol, due to its proximity to the external medium it can undergo rapid changes in pH as bacteria experience changes in the pH of their environment. Therefore, periplasmic proteins are expected to be able to withstand dramatic changes in pH. The HDX-MS data in this chapter reveal AcrA as a pH conformer and identifies Mg^{2+} as a putative co-factor. Furthermore, previous work has identified His285 as a key conformational regulator in AcrA, as the His side chain is protonated at pH 6.0.[2] The conformational dynamics and conformation of AcrA is critical to the efflux activity of the AcrAB-TolC multidrug efflux pump, thus it is likely pH, Mg^{2+} and His285 work together to modulate the conformational states and dynamics of AcrA in the assembled complex under varying periplasmic conditions.

The findings also suggest a wider role for Mg^{2+} in the function of MFPs, and thus multidrug efflux systems, to work across a variety of conditions the cell may exhibit. As discussed in **section 1.10.5.1**, AcrA has many homologs across Gram-negative bacteria, with some more closely related than others. For example, MexA from *P. aeruginosa*, is closely related to AcrA from *E. coli* and also shares the general architecture of four linearly arranged domains connected by flexible linkers.[369,409] Furthermore, MD simulations have also shown that MexA is highly flexible and can adopt different conformations with different orientations of the α-helical and MP domains, like AcrA, with its flexibility critical for the MexAB-OprM assembly and dynamics.[458] Therefore, due to the shared characteristics between MFPs, cation binding may be more general across the family. Two MFPs from the HME-RND family have already been shown to bind cations (ZneB and Zn^{2+}, CusB and Cu^{+}/Ag^{+}).[404,449]

As described in **section 2.2.4.1**, *Me*Bi*Pred* predicted several closely related homologs of AcrA to bind Mg^{2+} (**Table 2.5**); MdtE and AcrE from *E. coli* and MexA from *P. aeruginosa* were also suggested to

bind Mg^{2+}.[451] However, less closely related homologs of AcrA, as adjudged by McNeil *et al*. (2019) were not suggested to bind Mg^{2+}, such as MdtA from *E. coli* or MdsA from *Salmonella enterica*.[389] Furthermore, closely related MFPs such as AcrA and AcrE have been shown to exhibit interchangeability between cognate systems, whereas less closely related MFPs cannot. Thus, it may be possible that Mg^{2+} regulates the functional dynamics of AcrA-like MFPs in a similar fashion across RND efflux pump systems. The HDX-MS data has shown that Mg^{2+} binding to AcrA broadly rectifies increased backbone dynamics exhibited under acidic conditions, whilst specifically stabilising the αβ-barrel domain portion. It is possible the localised stabilisation of the αβ-barrel domain by Mg^{2+} specifically in mildly acidic conditions may offer a route to specialised conformations for robust efflux within these regimes.

These findings raise many questions about the regulation of the AcrAB-TolC multidrug efflux pump, and if bacteria regulate the conditions of the periplasmic environment to modify the stability or activity of the efflux pump under certain stress conditions. Is Mg^{2+} required to help AcrA bind to AcrB or TolC with higher affinity by promoting certain conformations, or does it have functions in the efflux mechanism or stability of the pump under varying periplasmic conditions. Future work is discussed in section **6.2.1**. Overall, this chapter provides a useful basis for future studies on the role of pH and Mg^{2+} in the function of MFPs and efflux pumps in general.

2.4 Materials and methods

2.4.1 Reagents

All reagents purchased from ThermoFischer Scientific or Sigma Aldrich/Merck unless otherwise stated. The highest quality reagents were always prioritised.

2.4.2 Molecular biology

2.4.2.1 AcrA^{L} constructs

A pET28a plasmid was engineered to contain full length AcrA with an LE linker and His_6-tag. AcrA was isolated and amplified from *E. coli* genomic DNA (AcrA^{L} cloning plan shown in **Appendix 3**). The polymerase chain reaction (PCR) products were purified by agarose gel electrophoresis and extracted using the NEB monarch DNA extraction kit. The pET28a vector was cut with NcoI and XhoI restriction enzymes, and again extracted using the same DNA extraction kit. Vector and insert were joined using the EcoDry Infusion HD enzyme premix. The construct was verified by DNA sequencing (Eurofins).

Another AcrA^{L} construct was provided from the Zgurskaya laboratory, which contained full length AcrA^{L} in a pUC18 plasmid which contained the native *acrAB* promoter and the lac promoter, with ampicillin resistance.[367] Construct contained an LE linker and His_6-tag. All cloning and sequencing for this construct was performed by the Zgurskaya laboratory.

2.4.2.2 AcrA^{S} construct

The AcrA^{S} construct was provided from the Zgurskaya laboratory, which contained AcrA lacking signal peptide 1-24 and Cys25Met mutation to ensure no lipidation, in a pEt28a plasmid.[367] The construct contained an LE linker and His_6-tag. The AcrA^{S} sequence is shown in **Appendix 4.** All cloning and sequencing for this construct was performed by the Zgurskaya laboratory.

2.4.2.3 Bacterial strains

All cloning was performed using NEB 5-alpha Competent *E. coli* cells. Any protein expression was performed in C43(DE3)Δ*acrAB E. coli* cells. C43(DE3)Δ*acrAB E. coli* was a kind gift from Klaas Martinus Pos (Goethe Universität, Frankfurt, Germany).

2.4.3 Protein expression and purification

2.4.3.1 AcrA^{L}

AcrA^{L} was purified according to previous protocols.[396,400,411] Briefly, pET28a or pUC18 containing AcrA^{L} was transformed into C43(DE3)Δ*acrAB E. coli* cells. 7 mL of an overnight Luria-Bertani (LB) culture was added to 1 L of pre-warmed LB broth containing 30 µg/mL kanamycin or 100 µg/mL ampicillin respectively. Cells were grown at 37 ˚C until an OD_{600} of 0.5-0.6 was reached, then 1 mM isopropylthio-β-galactoside (IPTG) was added to induce protein expression. Cells were harvested by centrifugation at 4200 x g for 30 mins at 4 ˚C after 3 hours of growth and washed with ice-cold phosphate buffer saline (PBS). The cell pellet was frozen at -20 ˚C overnight.

The cell pellet was thawed and resuspended in 40 mL Buffer A (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 5 mM imidazole). The cell suspension was supplemented with protease inhibitor mini tablets (1 tablet per 10 mL – as per Roche recommendations), 1 mM PMSF, 2 µL Benzonase, 1 mM ethylenediaminetetraacetic acid (EDTA) and 20 mg/mL lysozyme. The cell suspension was sonicated (Fisherbrand™ Model 120 Sonic Dismembrator, ThermoFisher) at 40 amps, 6 times for 15 seconds, with a 90 second break between each sonication. Insoluble cell debris was removed by centrifugation at 20,000 x g for 20 mins at 4 ˚C.

The membrane fraction was isolated by high-speed centrifugation at 200,000 x g for 1 hour at 4 ˚C. The membrane pellet was resuspended in Buffer A using a Potter-Elvehjem Teflon pestle and glass tube. This suspension was supplemented with protease inhibitor mini tablets (1 tablet per 10 mL – as per Roche recommendations) and 1 mM PMSF. The membrane fraction was solubilised in 1% DDM for 2 hours at 4 ˚C. Insoluble material was removed by centrifugation at 100,000 x g for 30 mins at 4 ˚C.

The solubilised membrane fraction was put through an AKTA purification system for Ni^{2+} chelation chromatography and then SEC. The sample was loaded onto a 1 mL HiTrap Ni^{2+} column equilibrated in Buffer B (50 mM $NaHPO_4$, pH 7.4, 300 mM NaCl, 20 mM imidazole, 10 % (v/v) glycerol, 0.03% (w/v) DDM). The column was washed with 10 column volumes (CVs) of Buffer B, then 20 CVs of Buffer B with 50 mM imidazole, then AcrA^{L} was eluted with Buffer C (50 mM $NaHPO_4$, pH 7.4, 300 mM NaCl, 500 mM imidazole, 10 % (v/v) glycerol, 0.03% (w/v) DDM). The samples were buffer exchanged directly through injection onto a Superdex 16/600 GL SEC column (GE Healthcare) equilibrated in Buffer D (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 10% v/v glycerol, 0.03% (w/v) DDM). Peak fractions containing pure AcrA^{L} were pooled. AcrA stored at -80 ˚C for long term storage.

2.4.3.2 AcrA^S

AcrA^S was purified according to previous protocols.[396,400,411] pET28a containing AcrA^S was transformed into C43(DE3)Δ*acrAB E. coli* cells. 7 mL of an overnight LB culture was added to 1 L of pre-warmed LB broth containing 30 µg/mL kanamycin. Cells were grown at 37 ˚C until an OD_{600} of 0.5-0.6 was reached, then 1 mM IPTG was added to induce protein expression. Cells were harvested by centrifugation at 4200 x g for 30 mins at 4 ˚C after 3 hours of growth and washed with ice-cold PBS. The cell pellet was frozen at -20 ˚C overnight.

The cell pellet was thawed and resuspended in 40 mL Buffer A (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 5 mM imidazole). The cell suspension was supplemented with protease inhibitor mini tablets (1 tablet per 10 mL – as per Roche recommendations), 1 mM PMSF, 2 µL Benzonase, 1 mM EDTA and 20 mg/mL lysozyme. The cell suspension was sonicated (Fisherbrand™ Model 120 Sonic Dismembrator, ThermoFisher) at 40 amps, 6 times for 15 seconds, with a 90 second break between each sonication. Insoluble cell debris was removed by centrifugation at 20,000 x g for 20 mins at 4 ˚C.

The membrane fraction was isolated by high-speed centrifugation at 200,000 x g for 1 hour at 4 ˚C. The supernatant was put through an AKTA purification system for Ni^{2+} chelation chromatography and then SEC. The supernatant was loaded onto a 1 mL HiTrap Nickel column in Buffer B (50 mM $NaHPO_4$, pH 7.4, 300 mM NaCl, 20 mM imidazole, 10% (v/v) glycerol). The column was washed with 10 CVs of Buffer B, then 20 CVs of Buffer B with 50 mM imidazole, then AcrA^S eluted with Buffer C (50 mM $NaHPO_4$, pH 7.4, 300 mM NaCl, 500 mM imidazole, 10% (v/v) glycerol). The samples were buffer exchanged directly injected onto a Superdex 16/600 GL SEC column (GE Healthcare) equilibrated in Buffer D (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 10% v/v glycerol). Peak fractions containing pure AcrA^S were pooled. Stored at 4 ˚C for up to a month, or at -80 ˚C for longer term.

2.4.4 Polyacrylamide gel electrophoresis (PAGE) and Western blots

2.4.4.1 SDS-PAGE

Proteins were diluted in 5x Laemmli sample buffer (312.5 mM Tris, pH 6.8, 50 % glycerol, 100 mg/mL SDS, 80 mg/mL dithiothreitol (DTT), 0.1% bromophenol blue). Samples were run on either a 10 or 12% pre-cast NuPAGE Bis-Tris or Novex Tris-Glycine gel. The Bis-Tris gels were run with a 20x NuPAGE MOPS SDS running buffer (50 mM MOPS, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.7). The Tris-Glycine gels were run with a 10x Tris-Glycine running buffer (25 mM Tris Base, 192 mM glycine, pH 8.3). The protein ladder used was the Novex Sharp Pre-Stained Protein Standard. 25 µL of proteins loaded per well, and samples run for 50 mins at 200 V. Bands were visualised by Brilliant Blue Coomassie stain.

2.4.4.2 Western blot

SDS-PAGE gels were transferred onto a nitrocellulose (0.2 µm pore size) membrane. Membranes were equilibrated in transfer buffer (25 mM Tris, 192 mM Glycine, 0.77% SDS (w/v), 10% methanol (v/v)) and proteins transferred from the cell to the membrane using a Cytiva Amersham™ TE 77 PWR Semi-dry Transfer Unit for 1 hour at 45 A per gel. Membranes were then blocked in 5% milk powder (w/v) in PBS-Tween (0.1% Tween (v/v) for 1 hour at room temperature or 4 ˚C overnight. Membranes were further incubated with Anti-PolyHistidine-HRP Antibody (1:10000) dilutions for 1 hour at room temperature. Membranes were then washed every 15 minutes in PBS-Tween, four times. Membranes developed with 1 mL Amersham™ ECL Select™ and imaged using an A1600 Imager (GE Healthcare).

2.4.5 Mass spectrometry

2.4.5.1 Native mass spectrometry

Purified $AcrA^{L}$ was exchanged into a volatile solution (100 mM ammonium acetate, pH 6.0/7.4, 0.03% DDM) using a centrifugal exchange device (Micro Bio-Spin 6, Bio-Rad) according to manufacturer's instructions, or via a SEC Superdex 10/200 Increase column at 0.4 ml/min flow rate. Native MS experiments were performed on a Synapt G2-Si mass spectrometer (Waters). Each sample was loaded into homemade gold-coated borosilicate glass capillaries and mounted onto the mass spectrometer where native MS experiments were performed. nESI was performed, and generated protein ions were drawn into the vacuum of the mass spectrometer. The following instrument parameters were carefully optimised to avoid ion activation and protein unfolding - capillary voltage: 1.6 kV, sampling cone: 120 V, trap DC bias; 15 V, trap collision energy: 50-200 V, transfer collision energy: 2 V. Pressures were set to 5.91 $x10^{-2}$ mbar in the source region (backing) and to 1.58 x 10^{-2} in both trap and collision cells. The collision gas was He.

For $AcrA^{S}$, the procedure was the same, except volatile buffer lacked DDM and native MS experiments were performed on either a Synapt G2-Si mass spectrometer (Waters) or The Q-Exactive Plus UHMR (ThermoFisher Scientific). The parameters for the native MS experiments on the Synapt G2-Si mass spectrometer (Waters) were the same as for $AcrA^{L}$, except the sampling cone was set to 30 V and the trap collision energy 2 V. This is because less energy was needed as no hydrophobic environment was present for $AcrA^{S}$. The UHMR settings used were: 1.5 kV spray voltage, capillary temperature 60 ˚C, ion source temperature off, extended trapping 60.

Data were processed and analysed using MassLynx v4.1 (Waters) and UniDec.[118]

2.4.5.2 Hydrogen deuterium exchange mass spectrometry

HDX-MS experiments were performed on a nanoAcquity ultra-performance liquid chromatography (UPLC) Xevo G2-XS Q-Tof mass spectrometer system (Waters). Optimised peptide identification and peptide coverage for $AcrA^S$ was performed from undeuterated controls. The optimal sample workflow for HDX-MS of $AcrA^S$ was as follows: 5 µL of $AcrA^S$ (20 µM) was diluted into 95 µL of either equilibration buffer (50 mM sodium phosphate, 150 mM NaCl, ± 1mM $MgCl_2$, pH 6.0/pH 7.4) or labelling buffer (deuterated equilibration buffer) at 20 ˚C. After fixed times of deuterium labelling, the samples were mixed with 100 µL of quench buffer (formic acid, 1.6 M GuHCl, 0.1% fos-choline, pH 1.9) to provide a quenched sample at pH 2.4. 70 µL of quenched sample was then loaded onto a 50 µL sample loop before being injected onto an online Enzymate™ pepsin digestion column (Waters) in 0.1% formic acid in water (200 µL/min flow rate) at 20 ˚C. The peptic fragments were trapped onto an Acquity BEH c18 1.7 µM VANGUARD pre-column (Waters) for 3 min. The peptic fragments were then eluted using an 8-35% gradient of 0.1% formic acid in acetonitrile (40 µL/min flow rate) into a chilled Acquity UPLC BEH C18 1.7 µM 1.0 x 100mm column (Waters). The trap and UPLC were both maintained at 0 ˚C. The eluted peptides were ionised by electrospray ionisation (ESI) into the Xevo G2-XS Q-Tof mass spectrometer. MS^E data were acquired with a 20–30 V trap collision energy ramp for high-energy acquisition of product ions. Argon was used as the trap collision gas at a flow rate of 2 ml/min. Leucine enkephalin was used for lock mass accuracy correction and the mass spectrometer was calibrated with sodium iodide. The online Enzymate™ pepsin digestion column (Waters) was washed three times with pepsin wash (1.5 Gu-HCl, 4% MeOH, 0.8% formic acid, 0.1% fos-choline) between runs.

All deuterium time points and controls were performed in triplicate/quadruplicate. Sequence identification was performed from MS^E data of digested undeuterated samples of $AcrA^S$ using PLGS software (Waters, v. 2.5.1). The output peptides were then filtered using DynamX (v. 3.0) using these parameters: minimum intensity of 1481, minimum and maximum peptide sequence length of 5 and 20 respectively, minimum tandem MS (MS/MS) products of 1, minimum products per amino acid of 0.11, and a maximum MH^+ error threshold of 5 ppm.[461] All the spectra were visually examined and only those with a suitable signal to noise ratio were used for analysis. The amount of relative deuterium uptake for each peptide was determined using DynamX (v. 3.0) and are only corrected for back-exchange when specified. The RFU was calculated from the following equation, where *Y* is the deuterium uptake for peptide a at incubation time *t*, and *D* is the percentage of deuterium in the final labelling solution:

$$RFU_a = \frac{Y_{a,t}}{MaxUptake_a \; X \; D}$$

Equation 2.1. Relative fractional uptake. *Y* is the deuterium uptake for peptide a, *t* is the incubation time, *D* is the percentage deuterium in the final labelling solution.

2.4.5.2.1 Statistics and reproducibility

For ΔHDX experiments a significance threshold between two states was established based on earlier approaches.[462,463] A CI was calculated using the SD of deuterated peptides for time points performed in quadruplets. For each state, SDs were averaged using the root mean square (**Equation 2.2**).

$$\mathrm{SD_{state}} = \sqrt{\frac{\sum \mathrm{SD_i^2}}{\mathrm{N}}}$$

Equation 2.2. Root mean square. *N* is the number of peptides considered multiplied by the number of time points performed in quadruplet.

A pooled SD for the difference between the two states was calculated using **Equation 2.3**.

$$\mathrm{SD_{pool}} = \sqrt{\mathrm{SD_{stateA}}^{\,2} + \mathrm{SD_{stateB}}^{\,2}}$$

Equation 2.3. **Pooled SD difference between two HDX states**.

The pooled SD was used to calculate the CI at the 99% significance level, considering a two-tailed distribution with three degrees of freedom (n=3) by using **Equation 2.4**.

$$\mathrm{CI} = 5.841 \times \frac{\mathrm{SD_{pool}}}{\sqrt{\mathrm{n}}}$$

Equation 2.4. Confidence interval calculation.

ΔHDX experiments in this chapter did not contain biological repeats, therefore a more conservative significance level cut-off was set that was higher than the calculated CI value, at 0.5 Da.[446] Therefore, significance for ΔHDX measurements of any individual time point was deemed as passing the 0.5 Da cut-off and scoring a *P*-value ≤ 0.01 in a Welch's *t*-test using HDeXplosion software.[447] Only peptides that satisfied both criteria were deemed significant. All ΔHDX structure figures were generated from the data using HDeXplosion and Chimera.[447,448]

2.4.5.2.2 Adjusting labelling times for pH 6.0

At pH 6.0, the rate of deuterium exchange is slower (see **section 1.4.2**). Therefore, to compare proteins to pH 7.4, labelling times were adjusted using **Equation 1.15**.

$$\frac{K_{ch1}}{K_{ch2}} = \frac{[OH^-]_1}{[OH^-]_2} = \frac{K_w/[H^+]_1}{K_w/[H^+]_2} = \frac{10^{-pH_2}}{10^{-pH_1}} = 10^{pH_1 - pH_2}$$

Equation 1.15. HDX time window expansion. When temperature is constant the above equation can be derived from the Arrhenius equation to calculate the ratio of amide hydrogen exchange rate constants that are applicable to the two pH conditions. This shows the intrinsic HDX rate decreases $10^{7.4\text{-}6}$ fold when the pH decreases from 7.4 to 6.0.

2.4.5.2.3 Maximally labelled control (MaxD)

5 µL of protein was diluted in 95 µL labelling buffer at pH 6.0 and labelled for 251 minutes at 50 ˚C. This is following the previously determined protocol of labelling for 10 mins at pH 7.4 at ~5 ˚C below the Tm of the protein.[160] This is to ensure maximum deuterium incorporation of unfolded protein. The deuterium content in the reaction mixture is identical to the corresponding HDX experiment. After 251 minutes, the proteins were left at room temperature for 2 minutes, then on ice for 2 minutes before being flash frozen and stored at -80 ˚C until LC-MS analysis.

2.4.6 Circular Dichroism

The CD experiments performed in this chapter were completed on a Chirascan V2 instrument. For standard CD scans, $AcrA^S$ was buffer exchanged into protein buffer (50 mM $NaHPO_4$, 150 mM NaCl, ± 1 mM $MgCl_2$, pH 6.0). Proteins were analysed at a concentration of 0.32 mg/mL. A coverslip was used with a pathlength of 0.05 cm. Scans were repeated three times between wavelengths 185-280 nm. BeStSeL online algorithm analysed the secondary structure.[454]

$AcrA^S$ was buffer exchanged into the same buffer as before but diluted to 0.0075 mg/mL. Thermal melts were performed at 15 temperatures from 30-95 ˚C, with 5 ˚C increments to first identify where the transition from a folded state to an unfolded state occurs. Then a more accurate scan ranging between 40-60 ˚C with 1 ˚C increments was completed to find the Tm. From this scan, values at 222 nm were taken for each temperature and thermodynamic parameters calculated from the following equations.

$$f_D = \frac{y - y_F}{y_D - y_F}$$

Equation 2.5. Calculating the fraction of denatured protein. f_D is the fraction denatured, y_F is the gradient of folded protein and y_D is the gradient of the denatured protein.

$$K_D = \frac{f_D}{1 - f_D}$$

$$\Delta G_D = -RT\ln K_D$$

$$\Delta H_m = T_m \Delta S_m$$

Equation 2.6. Calculating thermodynamic parameters. ΔG_D is the Gibbs free energy of denaturation, R is the gas constant with a value of 8.314 J $K^{-1}mol^{-1}$, T is the temperature in Kelvin, ΔH_m is the enthalpy change, T_m is the melting temperature and ΔS_m is the entropy change.

Spectra analysed on SigmaPlot.

Chapter 3: Use of a pseudo-dimer for the investigation into AcrA structural dynamics

Sections of this chapter has been adapted and modified from publication with additional discussion and detail:

Russell Lewis, B., Uddin, M.R., Moniruzzaman, M. et al. Conformational restriction shapes the inhibition of a multidrug efflux adaptor protein. Nat Commun 14, 3900 (2023). https://doi.org/10.1038/s41467-023-39615-x

3.1 Introduction

Membrane fusion proteins (MFPs) are essential components of tripartite efflux pumps that span the entire cell envelope of Gram-negative bacteria. They all bind a corresponding inner membrane transporter and an outer membrane exit channel in the periplasm. To efficiently do this, MFPs share several characteristic traits. They all have an elongated asymmetric shape of a protomer, they contain 3-4 linearly arranged domains with flexible linkers to create dynamic structures, and they oligomerise.[397,409,410,464] *In vitro* studies of several MFPs from different superfamilies of multidrug efflux pumps have been shown to form oligomers *in vitro*; AcrA, MexA, MacA and EmrA are a few examples.[367,411,458] This work utilised surface plasmon resonance (SPR) and size exclusion chromatography (SEC) to monitor the presence of oligomers, however there is no structural biology information available on the effect of oligomerisation on the structural dynamics of the corresponding MFP.[367,411] Furthermore, it has been suggested that experimental conditions such as pH can affect MFP oligomerisation; this was confirmed for AcrA using native mass spectrometry (native MS) shown in **chapter 2**. Importantly, *in vivo* genetic studies and *in vitro* biochemical investigations have shown that the functional unit of MFPs is a dimer, and that trimerization of the dimer is what leads to the formation of funnel-like structures that bind the transporter and exit duct to form a sealed channel across the periplasm for drug efflux.[366,367,410,457–460]

Within the functional dimer of an MFP, the two protomers can have a specific role and function slightly differently. For AcrA, this is highlighted by the differences that each protomer bind AcrB.[366,373,389] **Figure 3.1** shows the differences in how each protomer binds AcrB, based on the work by McNeil *et al*. (2019).[389] Models of *Salmonella* AcrA were constructed based on the direct correspondence of the sequence between it and experimentally determined *Escherichia coli* (*E. coli*) structures, with 92% identity, meaning the findings translate well to *E. coli* AcrA. The first AcrA protomer's αβ-barrel domain binds to the Nα4 helix and β-hairpin1 of the DN subdomain in AcrB and also contacts the Cα4 helix and β-hairpin2 of the DC subdomain. Then, the AcrA membrane proximal (MP domain) binds to the base of the DC subdomain and Nβ8-Nβ9 of the neighbouring AcrB protomer, whilst also binding to the PC1 subdomain of the original AcrB protomer. The second AcrA protomer's αβ-barrel domain binds to the DC α-hairpin of the neighbouring AcrB protomer and also binds to the AcrB α-hairpin and β-hairpin of the DN subdomain. The MP domain then binds to the base of the funnel domain and the PN2 subdomain. Furthermore, each protomer in an MFP dimer have been shown to interact with two different specific sites on the OMF (outer membrane factor), with distinctive roles: evidence suggests one protomer grasps the OMF whereas the other is responsible for opening the channel.[410,412,423] Therefore, it is clear the AcrA protomers in the functional dimer unit can act differently to each other

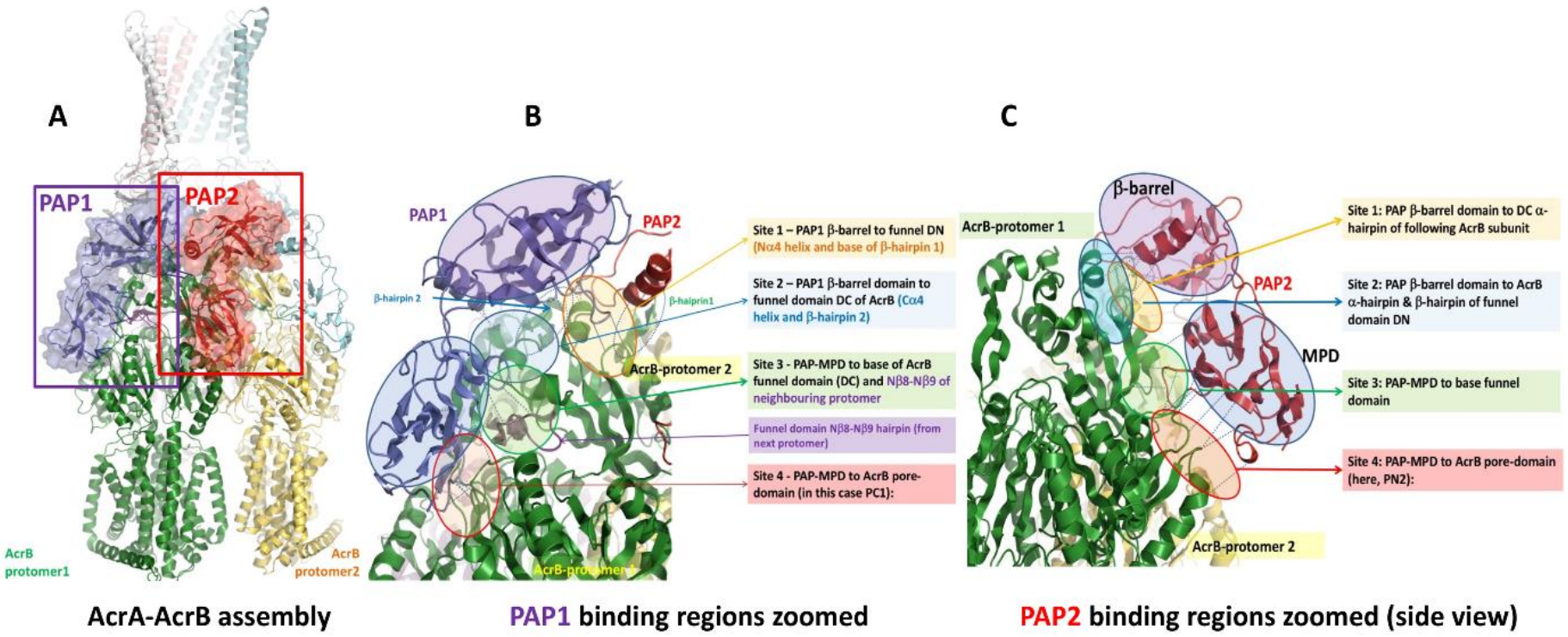


Figure 3.1. The AcrA-AcrB binding interaction. A. Positions of AcrA protomer 1 and AcrA protomer 2 relative to the AcrB protomer. Binding is relative to the green AcrB protomer. **B.** Binding regions of AcrA protomer 1 relative to the AcrB protomer. The binding region is annotated with the 4 principal binding sites and structural features of the surface. **C.** Binding regions of AcrA protomer 2 relative to the AcrB protomer. PDB 5O66. Taken from McNeil *et al.* (2019).[389]

in the assembly of the pump, so, understanding the biology of the AcrA dimer is critical for understanding the assembly and function of the AcrAB-TolC pump.

Studying the AcrA dimer is a challenging feat due to the heterogeneity of AcrA *in vitro*, making it difficult to isolate dimers for experimentation. Whilst the previously characterised AcrA^{G363C} mutant forms a covalently linked dimer under nonreducing conditions, *in vitro* it presents the same oligomerisation pattern as lipidated AcrA (AcrA^{L}).[367] Therefore, to further investigate the nature of the AcrA dimer, a pseudo-dimer construct was designed to gain structural biology information and biochemical insights on the AcrA functional dimer. The pseudo-dimer contained two soluble AcrA (AcrA^{S}) sequences connected by a 5 amino acid linker (TRRIT) to form one polypeptide chain. The idea was to use the soluble version of AcrA when designing the pseudo-dimer construct, which has been previously shown to be monomeric *in vitro* (confirmed by native MS studies, **chapter 2**), in order to make a homogenous pseudo-dimeric protein.[367] The pseudo-dimer construct was denoted AcrA^{SD} soluble dimer, as it contains two AcrA^{S} sequences. This homogenous construct allowed for more accurate interpretation of experimental results, and therefore better inference into the behaviour of the AcrA functional dimer unit.

As stated above, the two AcrA protomers in the AcrA functional dimer bind AcrB differently, and dimerization of AcrA has been shown to promote high affinity binding to AcrB.[367,389] Comparing differences in the way AcrA^{S} and AcrA^{SD} bind AcrB could highlight critical differences between the constructs. Two techniques are used to characterise the AcrA-AcrB interaction in this chapter, which are styrene maleic acid polyacrylamide gel electrophoresis (SMA-PAGE) and mass photometry.[201,465] Firstly, SMA-PAGE is an adapted native-PAGE protocol for proteins in styrene maleic acid lipid particles (SMALPs).[465] Native-PAGE separates charged proteins based on their mass, charge, and shape under non-disrupting conditions.[465,466] Electrophoretic migration through the gel occurs as most proteins carry a negative charge in the presence of alkaline running buffers, with higher charge density molecules moving faster. However, the gel matrix contains pores and acts like a sieve, which regulates the migration of proteins based on the size and shape of the native molecules. Usually, a protein ladder is run concurrently to provide a molecular weight standard. Therefore, native-PAGE can report on native masses, oligomeric states, stoichiometry of protein complexes and protein:protein interactions, as well as many more tasks.[466] Typically, native-PAGE has been somewhat unreliable for membrane proteins, often due to the presence of detergent which can interfere with non-covalent interactions between subunits, disrupting quaternary structures.[467] Furthermore, many membrane proteins bands visualised on the gels can have low clarity or resolution; bands can appear smeared or migrate differently than predicted by their mass, making interpretation difficult. This can be caused by an inconsistent amount of detergent bound among membrane proteins.[468] Furthermore the

mitigations for these issues often involve additives which make the protocol more complex without necessarily improving the data.[465,466] Therefore native-PAGE has been developed for use with membrane proteins in SMALPs, creating SMA-PAGE. Using the SMA-PAGE protocol, proteins migrate as intact SMALPs, with their migration being consistent with their size and whilst retaining lipids.

Secondly, mass photometry was used to probe the AcrA:AcrB interaction. For an in-depth introduction into mass photometry see **chapter 1.5**. Briefly, mass photometry is adept at determining molecular mass, oligomeric states and determining ligand binding, at a single molecule level.[201] Buffered samples are placed on a glass coverslip, which is illuminated with a laser. Single molecules bind non-specifically to the coverslip glass surface, which changes the refractive index at the glass water interface, altering the local reflectivity, leading to scattered light. This scattered light from biomolecules interferes with reflected light from the glass coverslip, and the degree of interference is directly proportional to mass. Mass is then calculated by comparing the interference value to that of relevant protein calibrants. This technique offers many benefits, such as the need for very little sample concentration or volume, and it is amenable to membrane proteins in SMALPs.[205]

Aside from AcrB, AcrA has been shown to have an affinity to peptidoglycan. Shi *et al.* (2019) used cryo-electron microscopy (cryo-EM) density and crosslinking MS to conclude the α-helical hairpins contact peptidoglycan in the periplasm and is essential to the complex assembly *in vivo*.[38] Furthermore, Xu *et al.* (2011) have also demonstrated the affinity of AcrA and TolC to *E. coli* peptidoglycans.[460] These findings have been supported by molecular dynamics (MD) simulations, which suggest peptidoglycan is positioned at the AcrA:TolC interface during pump assembly.[257] No studies have looked at the effect of dimerization on the affinity of AcrA to peptidoglycan, and investigating this may reveal further insights into the assembly and stability of AcrAB-TolC.

This chapter aimed to characterise the pseudo-dimerization of AcrA using the novel $AcrA^{SD}$ construct, to gain structural and biochemical insights into the effect of AcrA dimerization. Understanding the behaviour of the AcrA functional dimer could shed light on how the AcrAB-TolC multidrug efflux pump assembles *in vivo*, and possibly provide a new avenue to inhibit AcrA and corresponding MFPs. Firstly, $AcrA^{SD}$ was purified, and its oligomeric state characterised by native MS. Then the effect of AcrA pseudo-dimerization on AcrA thermal stability and structural dynamics was measured by circular dichroism (CD) spectroscopy and hydrogen deuterium exchange mass spectrometry (HDX-MS) respectively. Qualitative comparisons were made between $AcrA^{S}$ and $AcrA^{SD}$ binding to AcrB, using SMA-PAGE and mass photometry. Furthermore, a peptidoglycan pull-down assay was developed to monitor AcrA's affinity for peptidoglycan, so the effect of AcrA pseudo-dimerization on its ability to bind peptidoglycan could be investigated.

3.2 Results

3.2.1 Overexpression and purification of the AcrA pseudo-dimer construct

3.2.1.1 Purification of $AcrA^{SD}$

$AcrA^{SD}$ contains two $AcrA^{S}$ sequences connected by a five amino acid linker TRRIT. The first AcrA sequence lacks signal peptide 1-24, and the Cys25 is mutated to Met, with an additional Ala residue afterwards. The second AcrA sequence starts from Asp26 and ends with an LE linker and an His_6-tag. Therefore, $AcrA^{SD}$ is not transported to the membrane and does not have the post-translational lipidation, so it can be purified from the cytosol without the presence of a membrane mimetic such as detergents or nanodiscs. **Figure 3.2** represents a characteristic purification of $AcrA^{SD}$. The SEC trace (**Figure 3.2a**) presents differently to that of the lipidated $AcrA^{L}$ construct in **Figure 2.1**, as AcrA lacking the lipidation does not form oligomers *in vitro* according to previous work and results shown in **chapter 2** (**Figure 2.2a**).[367] A void peak is observed at ~45 mL and a broad, shouldered peak at ~55 ml, which may be due to the presence of $AcrA^{SD}$ aggregates or higher order oligomers respectively. Then, a main peak with the highest intensity was seen at ~65 ml, representing $AcrA^{SD}$, suggesting this is the most dominant species of the purification. For fair comparison to $AcrA^{S}$, only the three middle fractions of the main peak were pooled and stored for further experimentation.

The sodium dodecylsulfate (SDS)-PAGE gel (**Figure 3.2b**) shows the most intense band at ~80 kDa representing $AcrA^{SD}$, closely matching the theoretical molecular weight of $AcrA^{SD}$ (81,118 Da). As in **chapter 2**, the fractions were loaded onto the gel as they were eluted from the SEC column, to provide a quality control check of the protein sample, and then the sample concentration was measured on a nanodrop (Implen NanoPhotometer, Geneflow). Fractions 8-9 represent the SEC peak at ~55 mL and confirms this contains $AcrA^{SD}$ and not a protein contaminant. The main peak (fractions 12-16) shows an intense single band at ~80 kDa on the SDS-PAGE gel. $AcrA^{SD}$ presents as a single band at ~80 kDa as expected due to its molecular weight. Fractions 15-16 show slightly more degradation, represented by the faint laddering under the main band, so were discarded. Overall, the SDS-PAGE gel shows a very clean purification of $AcrA^{SD}$. A western blot analysis using Anti-PolyHistidine-HRP Antibody confirmed these bands were $AcrA^{SD}$ (**Figure 3.2c**). This purification protocol yielded less protein than $AcrA^{S}$, at ~4 mg/mL, equivalent to ~50 µM.

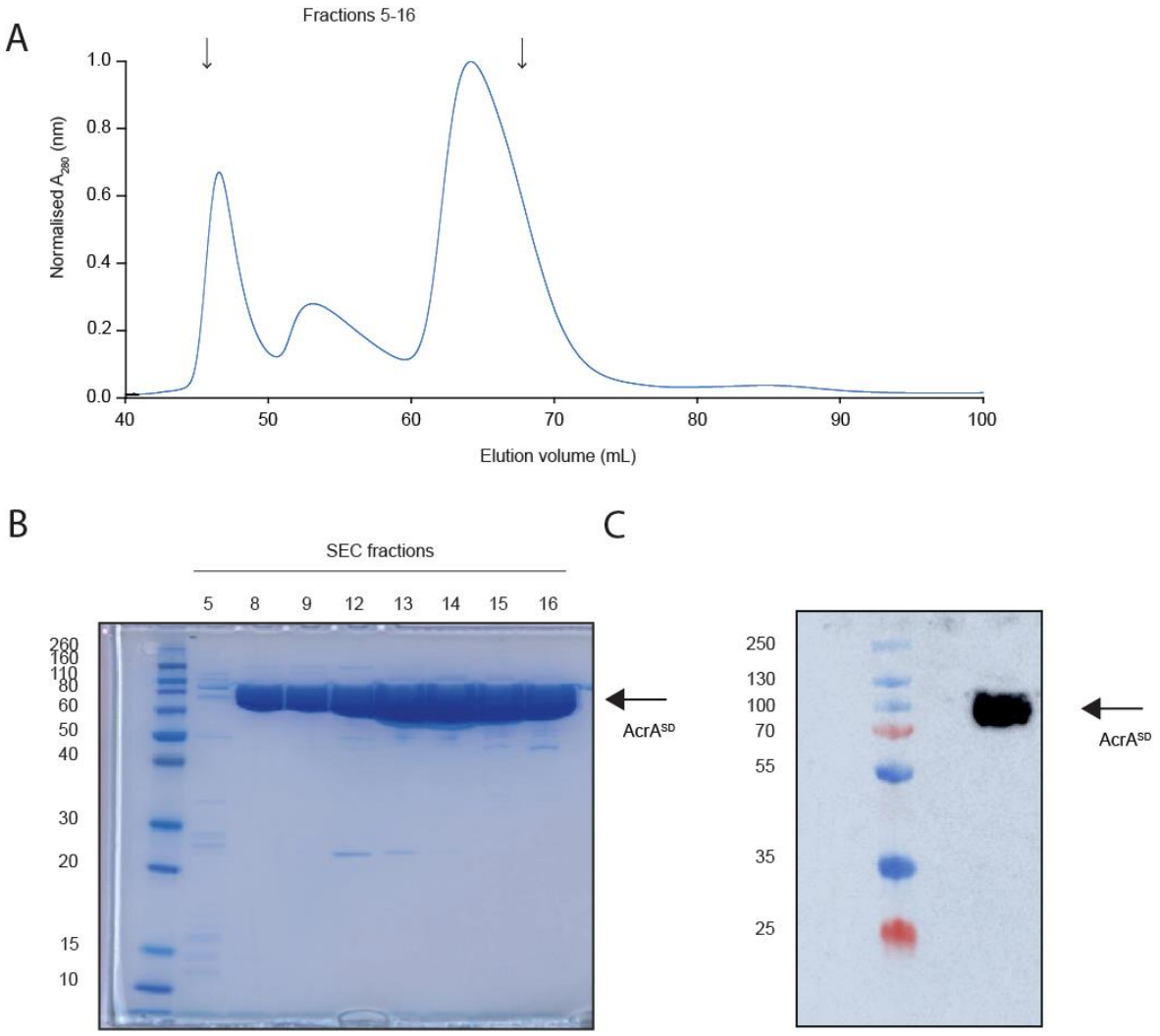


Figure 3.2. Characterisation of AcrA^{SD} purification. A. Characteristic size exclusion chromatogram for AcrA^{SD}. Fractions of the main peaks loaded onto the gel (fractions 5-16). Absorbance normalised. Void peak shown at ~45 ml, higher order AcrA^{SD} peak shown at ~55 ml, and main peak at ~75 ml. **B.** Characteristic SDS-PAGE running the size exclusion chromatography fractions shown in part A. Gel shows a single band at around ~80 kDa for AcrA^{SD}. **C.** Western blot using HRP anti-his antibodies. The western blot confirms the presence of the AcrA^{SD} construct.

3.2.2 Native MS of the AcrA pseudo-dimer construct

Native MS experiments were conducted as described in **methods 2.4.5.1** and **results 2.2.2.**

3.2.2.1 Investigating AcrA^{SD} oligomerisation

Native mass spectra of AcrA^{SD} were achieved using both a Synapt G2-SI (Waters) and a Q-Exactive Plus ultra-high mass resolution (UHMR) (ThermoFischer) instruments. **Figure 3.3** shows the native mass spectra for AcrA^{SD} on both instruments. AcrA^{SD} was observed to be entirely pseudo-dimeric at pH 6.0, as analysed by UniDec.[118] Please note, this is technically monomeric as AcrA^{SD} is a single polypeptide chain, but referred to as pseudo-dimeric. As stated in **section 3.2.1**, the theoretical mass of AcrA^{SD} is 81,118 Da. The measured masses for AcrA^{SD} were 81,005 ± 4 Da (Synapt, **Figure 3.3a**) and 81,000 Da ± 3 (UHMR, **Figure. 3.3b**) and are shown in **Table 3.1**. The measured masses differed by -113 and -118 Da respectively. This is most likely due to N-terminal formyl methionine (fMet) processing. In bacteria, as well as eukaryotic organelles such as mitochondria and chloroplasts, the N-terminal Met residue is N$^{\alpha}$-terminally formylated through a pre-translational mechanism.[469] Formyltransferase uses 10-formyltetrahydrofolate to formylate the Met α-amino group in the initiator tRNA$_i^{Met}$. The fMet becomes the first residue of the resulting nascent chain that emerges from the ribosome; this formyl moiety is co-translationally removed by peptide deformylase, which can bind to the ribosome by the exit tunnel. After the fMet is de-formylated, the resulting Met can be cleaved by Met-aminopeptidase. The removal of the de-formylated Met requires the residue at position 2 to be smaller than Val. This may explain why fMet processing is seen for the AcrA^{SD} construct but not AcrA^{S}. Val has a mass of 117 Da; the second residue in the AcrA^{SD} construct is Ala, which has a mass of 89 Da. Therefore, the starting Met residue can be cleaved. However, the second residue in the AcrA^{S} construct is Asp, which has a mass of 133 Da, and thus the starting Met cannot be cleaved. Therefore, the theoretical mass of AcrA^{SD} in **Table 3.1** has been adjusted for fMet processing, and the differences between the theoretical and measured masses become 18 and 13 Da respectively. These differences are likely due to a adducts that remain in the gas phase post-electrospray ionisation (ESI), e.g. + H_2O, + H or + NH_4.

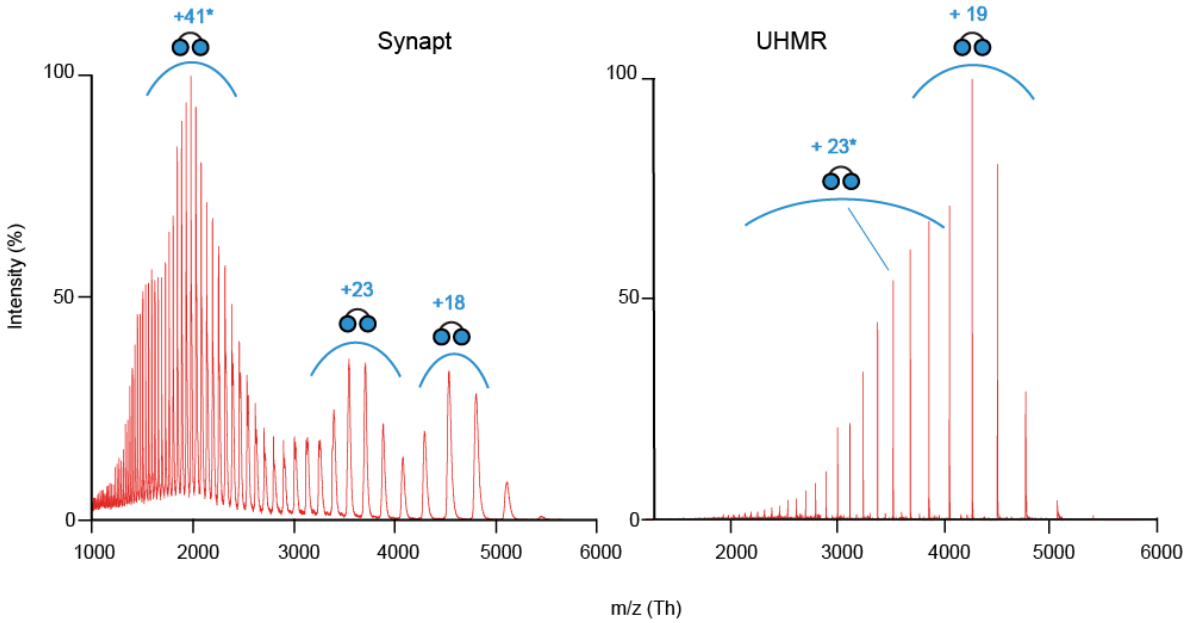


Figure 3.3. Native MS of AcrA^{SD} at pH 6.0. Native MS characterisation of AcrA^{S} at pH 6.0 on both the Synapt G2-SI (Waters) and Q-Exactive UHMR (ThermoFischer) instruments. Proteins were buffer exchanged in 100 mM ammonium acetate buffer prior to MS. AcrA^{SD} presented as pseudo-dimers in both instruments. Masses found in **Table 3.1**.

Table 3.1. Native MS masses for AcrA^{SD}

	Measured mass (Da)	**Standard Error (± Da)**	**Theoretical mass†,* (Da)**	**Mass difference (Da)**
***AcrA^{SD} pH 6.0 (Synapt)**	81,005	4	80,987	18
***AcrA^{SD} pH 6.0 (UHMR)**	81,000	3	80,987	13

**Theoretical masses for AcrA^{SD} construct were amended for fMet processing.*

3.2.2.2 Charge state distributions of AcrA^{SD}

Figure 3.3 shows that AcrA^{SD} produces multiple charge state distributions (CSDs), as previously seen with AcrA^{S} (**Figure 2.3, Figure 2.4**). The MS spectrum captured on the Synapt shows three different CSDs. There is a lower charge, narrow CSD at ~4500 *m/z* centred around a +18 ion charge, a middle charge CSD at ~3500 *m/z* centred around a +23 ion charge, and a higher charge, wider CSD at ~2000 *m/z* centred around a +41 ion charge. Firstly, the MS spectrum captured on the UHMR shows a greater resolution as expected due to the high resolving power of the instrument compared to the Synapt.[128] The low *m/z* CSD show a better peak separation in spectra acquired by the UHMR, and the peaks

exhibit a lower full width half maximum (FWHM). Like AcrA^{S}, AcrA^{SD} also presents multiple CSDs. It has a lower charge, narrow CSD at ~4500 *m/z* centred around a +19 ion charge, and a higher charge, much broader CSD at ~3500 *m/z*, centred around a +23 ion charge. This second, broader CSD may encompass both the +23 and +41 CSDs seen on the Synapt, just at much lower intensities. The reason for the different intensities of the CSDs measured across the two instruments is possibly due to differences in needle preparation for nano electrospray ionisation (nESI), or due to the Synapt's z spay source compared to the direct spray in the UHMR effecting the nESI mechanism of AcrA^{SD}.[119] The two different instruments present slightly different ionisation envelopes, but both still present multiple CSDs suggesting that intrinsic disorder is still present when AcrA is a pseudo-dimer. Therefore, it was important to characterise AcrA^{SD} using two different systems, to ensure the range of observed CSDs were not an instrumental artefact caused by nuanced differences between the instruments. It is important to remember that these CSDs are representative of AcrA^{SD} in the gas phase, and not necessarily solution populations, however it is possible to draw biological inferences. For more information regarding native MS and CSDs please see **section 2.2.2.2**.

3.2.3 The effect of dimerization on the thermal stability of AcrA

In some proteins, dimerization leads to an increase in thermodynamic stability.[470] To investigate whether AcrA pseudo-dimerization achieves this effect, CD thermal melts were completed on both AcrA soluble constructs. A different instrument was used in this chapter than the one used in **chapter 3**, so AcrA^{S} thermal melts were repeated. Proteins were buffer exchanged into Tris-HCl buffer at pH 6.0 and diluted to 0.4 mg/mL. Wavescans between 190-260 nm were performed at 14 temperatures ranging from 25-90 ˚C with 5 ˚C increments (**Figure 3.4a**) and the melting temperatures (Tms) were then calculated and analysed as described in **chapter 2** (**Figure 3.4b**). As **Figure 3.4** shows, the Tms of AcrA^{S} and AcrA^{SD} are the same, at 53.1 and 53.2 ˚C respectively, suggesting that pseudo-dimerization is not increasing the thermodynamic stability of AcrA. This does not necessarily suggest pseudo-dimerization has no increased thermodynamic stability, as this only tested thermostability, but has not looked at entropic or enthalpic factors. It is known dimerization can confer several other structural and functional advantages besides an increase in thermal stability, which does not always confer an advantage anyway.[471] For example, dimerization assists the assembly of large dynamic complexes from preformed subunits; this negates the need to expand the genome and avoids encountering problems with the folding of large proteins. Furthermore, dimerization helps regulate allostery by generating new binding faces at the dimer interface or induce structural changes in existing subunits, to increase or decrease specificity to binding partners. It is likely these are the advantages of dimerization for AcrA, as the functional dimer has to bind both AcrB and TolC, with each protomer in

an AcrA dimer binding AcrB and TolC differently. This is then followed by the formation of a trimer of dimers in the assembled pump to create a sealed channel through the periplasm.[366,373,389]

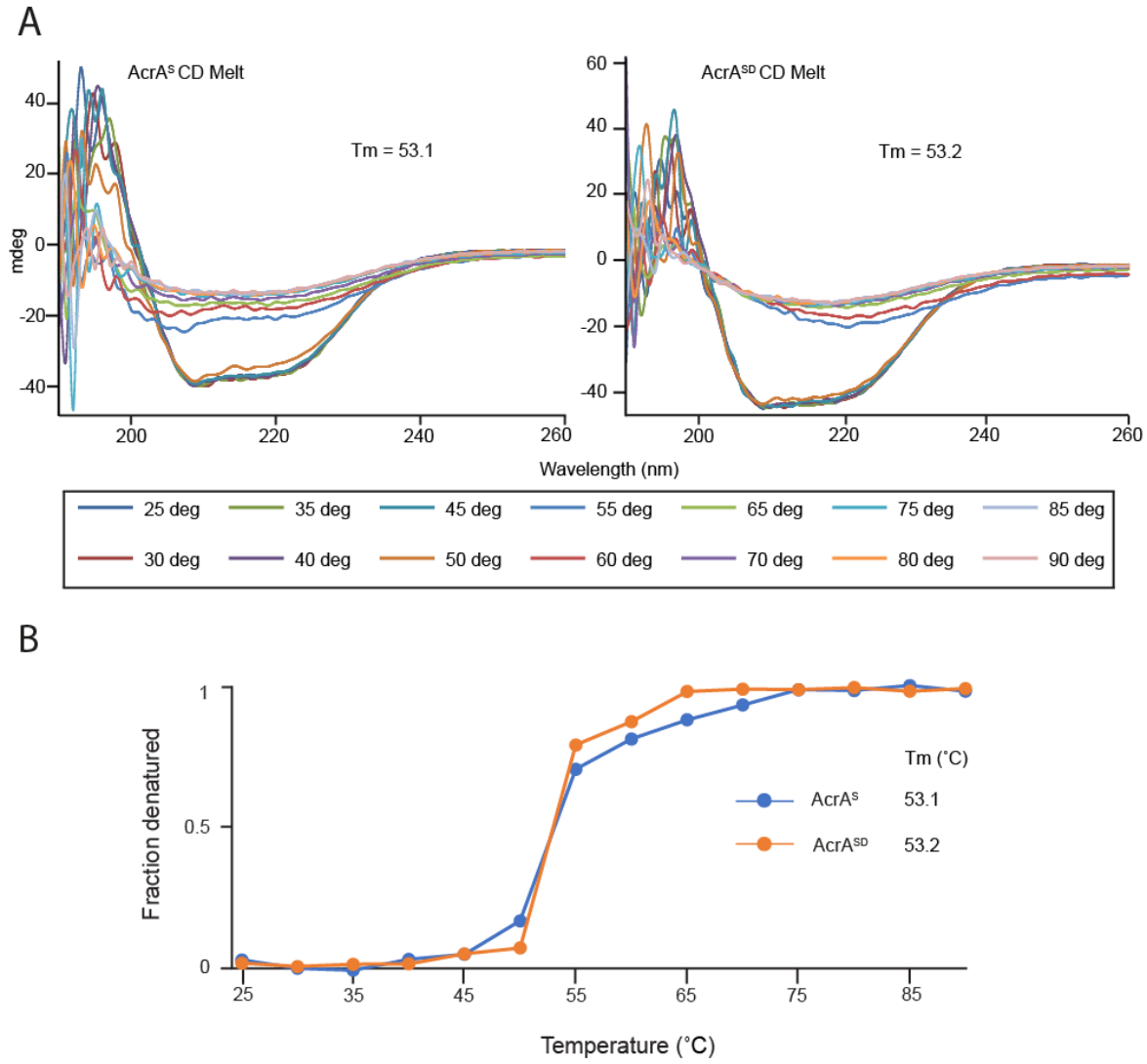


Figure 3.4. Circular dichroism thermal melts of AcrA^{S} and AcrA^{SD}. A. Circular dichroism thermal melts of AcrA^{S} and AcrA^{SD}. Proteins diluted to 0.4 mg/mL and CD measured from 190-260 nm in a 0.5 mm pathlength cell, from 25 ˚C to 90 ˚C, at 5 ˚C intervals. **B.** Scans at 222 nm taken for each temperature recorded. Plot shows fraction denatured vs temperature. Tms reported are 53.1 ˚C for AcrA^{S} and 53.2 ˚C for AcrA^{SD}.

3.2.4 Investigating the effect of dimerization on the structural dynamics of AcrA

Until now, there has been no experimental data available that shows the effect of dimerization on AcrA structural dynamics. To monitor this, differential HDX (ΔHDX) was performed for $AcrA^{SD}$ – $AcrA^{S}$ at pH 6.0. The experiments were carried out as previously described in **chapter 2**. The only difference was the presence of 5% dimethyl sulfoxide (DMSO) in the equilibration and labelling buffers for both proteins. This was because these experiments provided a no-drug reference state in **chapter 4** (more information found here). When performing ΔHDX, statistical tests are used to draw biological conclusions from any differences between two states. Similar to **chapter 2**, hybrid statistics were used that involved peptides passing a cut-off threshold for a statistically significant change between two states and a Welch's *t*-test. The confidence interval (CI) (**methods in section 2.4.5.2.1**) was calculated using the pooled standard deviation (SD) values of the datasets.[446] Therefore, only peptides that met a Δ0.42 Da cut-off and passed a Welch's *t*-test were deemed significant.

Figure 3.5 shows that $AcrA^{SD}$ has unique dynamics compared to the monomer even though they possess similar thermal stabilities (**Figure 3.4**). Predominantly, the dimer exhibits stabilisation across all four domains, but far more localised than previously seen between pH's and ± Mg^{2+} for $AcrA^{S}$. The most extensive protection is seen in the α-helical hairpin. **Figure 3.5b** reveals the majority of the helices are protected in the latest time point, with only the top left section of the α1 helix exhibiting insignificant ΔHDX. The early time points do not show much change, likely because they are a stable unit with defined secondary structure, as can be seen in the chiclet plot and uptake plot for peptide ^{147}LADAQQANAAVTAAKAAVET^{166} (**Figure 3.5a/c**). However, over the HDX time course, difference in ΔHDX increases. It has been well characterised that α-helices can pack together, which can affect their mutual orientations.[472] The α-helical surfaces should be complementary to each other, with the hydrophobic stripes of the helices fitting together in a 'jig-saw-like' manner, and polar side chains arranged in a complementary fashion. The packing of the hydrophobic residues is stabilised by van der Waals forces whereas the side chains of polar amino acids can be stabilised from hydrogen bonding.[473] This packing increases the stability of the hairpins and affects their flexible rotation. Therefore, dimerization may stabilize the α-helical domains flexibility to position them for interactions with TolC. The decreased HDX observed for the later time points is probably due a mixture of reduced solvent accessibility at the helix-helix interface and a decreased rotation of the helices due to dimerization.

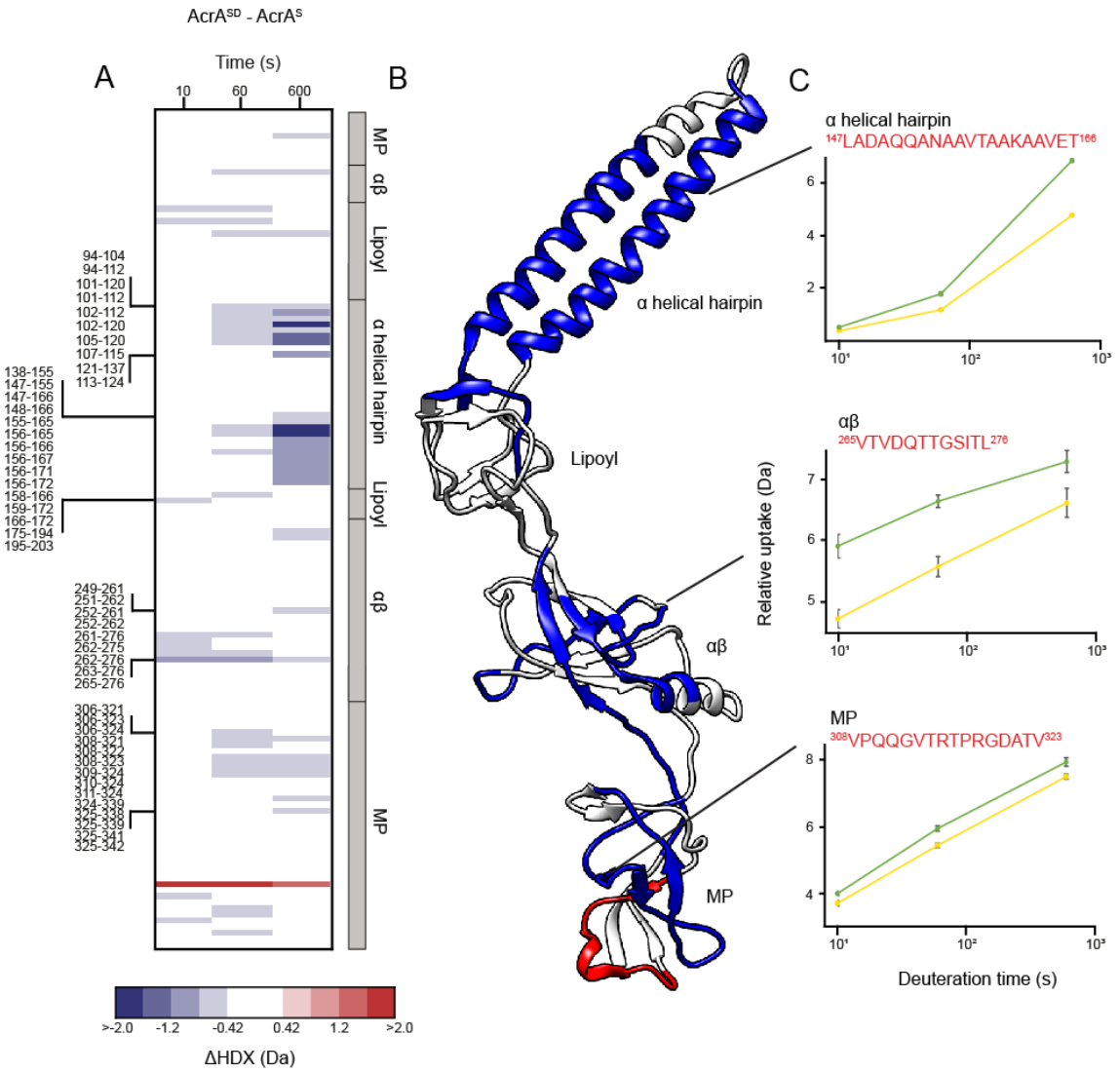


Figure 3.5. Effect of pseudo-dimerization on AcrA structural dynamics. A. Chiclet plot displaying the differential HDX (ΔHDX) plots for AcrA^{SD} - AcrA^{S} for all time points collected. Blue signifies areas with decreased HDX between states and red signifies areas with increased HDX between states. Significance was deemed to be ≥ 0.42 Da change with a *P*-value ≤ 0.01 in a Welch's *t*-test (n = 4 technical replicates). White areas represent regions with insignificant ΔHDX. **B.** ΔHDX for (AcrA^{SD} – AcrA^{S}) for the latest time point is painted onto the AcrA structure (PDB:5O66) using HDeXplosion and Chimera.[447,448] **C.** Uptake plots for three peptides in different domains of AcrA. Uptake plots are the average deuterium uptake and error bars indicate the standard deviation.

The αβ-barrel and MP domains also exhibit significant protection in AcrA^{SD}. Peptides in the αβ-barrel domain show more protection at earlier time points, as shown by the chiclet plot (**Figure 3.5a**). Peptide ^{265}VTVDQTTGSITL^{276} also shows this, with the ΔHDX larger at the first two time points, and smaller in the last, although still with a significant difference (**Figure 3.5c**). Pseudo-dimerization may be adding more structural order to this region, or there may be a contribution from decreased solvent accessibility at a dimer interface and therefore over time dynamics become closer. The MP domain

predominantly shows protection at the later time points, similar to the α-helical domain, but to a lesser degree. This stabilisation occurs between residues 306-342, which is the core of the MP domain. Peptide ^{308}VPQQGVTRTPRGDATV^{323} is representative of peptides found in this region (**Figure 3.5c**). Peptide 358-372 showed a marked increase in ΔHDX, as can be seen in the chiclet plot. However, no overlapping peptides in the same region saw this increase, so no biological conclusions can be drawn. Interestingly, the lipoyl domain shows very little change between AcrA^{SD} and AcrA^{S}, suggesting backbone dynamics in this region is not altered extensively by pseudo-dimerization. Overall, the HDX results show AcrA^{SD} has different structural dynamics than AcrA^{S}.

To summarise, HDX-MS has enabled the first look at the effect of pseudo-dimerization on the structural dynamics of AcrA. From this, it is possible to gain biological insights into the importance of the AcrA functional dimer. The most extensive protection is within the α-helical hairpins, which are involved in defined interactions with TolC. The increased stability within these regions may aid in the TolC opening mechanism. Furthermore, increased stability is seen across sections of the αβ-barrel domain and MP domain of AcrA, which both interact with AcrB. As AcrA is a very dynamic protein, shown by HDX-MS investigations in **section 2.2.3.2**, and it may be that the increased dynamic stability provided by dimerization allows the AcrA protomers to be primed for interactions with their binding partners and stabilises particular conformations for binding. One limitation of this investigation was the inability to differentiate between the two protomers in the AcrA^{SD} construct. As HDX-MS is an averaging technique, and the two protomers share the same amino acid sequence, analysed peptides will be an average of the deuterium uptake between them. MD simulations could be a suitable next step to complement these HDX-MS results, to model the conformations exhibited by the pseudo-dimer and its interactions with AcrB and TolC.

3.2.5 Differences in binding to AcrB

3.2.5.1 SMA-PAGE

It has been previously reported that AcrA binds AcrB as a dimer *in vivo*, with each protomer interacting with AcrB differently, and that an AcrA dimer promotes high affinity binding to AcrB.[337,366,367,389] In this section, AcrB is purified in SMALPs to provide a qualitative look at the AcrA-AcrB interaction; a more in depth introduction on membrane protein mimetic environments can be found in **chapter 5.1**. Therefore, to assess whether dimerization effects the binding to AcrB *in vitro*, SMA-PAGE was utilised. Previous work has characterised AcrA^{S} binding to AcrB; Tikhonova *et al.* (2011) using SPR to show AcrA^{S} binds AcrB at pH 6.0 with a K_D of 1.2 μM, and the data fit well to a 1:1 binding model.[367] There is no experimental data available that measures how an AcrA dimer can bind AcrB *in vitro* due to the

difficulty of isolating an AcrA dimer in solution. Moreover, previous work used AcrB in n-Dodecyl-β-D-maltopyranoside (DDM) detergent micelles, yet in this chapter AcrB was purified in SMALP native nanodiscs. This meant AcrB was in a nanodisc containing its own lipids, and thus is in a more native-like environment, which is likely to have an impact on its stability and function.[112,232]

Figure 3.6a shows the SMA-PAGE of AcrB in SMALPs, AcrA^{S}, and AcrA^{SD}. The AcrB trimer band appears at ~480 kDa, which is in agreement with previous mass determinations of AcrB in SMALPs by laser-induced liquid bead ion desorption (LILBID)-MS.[63] The theoretical molecular weight of the AcrB trimer is 342 kDa, with the extra mass accounting for the lipids within the SMALP nanodisc. Furthermore, a higher band can be seen at ~700 kDa, agreeing with previous work that AcrB can exist as a trimer-dimers in SMALPs due to SMA nanodisc interactions.[465,474] The two AcrA constructs appear to not effectively migrate through the gel in isolation, as they do not leave the wells they were loaded into. This is not due to the isoelectric points (pI) of AcrA^{S} and AcrA^{SD} which are 6.40 and 6.76 respectively, meaning they should carry negative charge in the sample buffer used (pH 8.2). The reason for the migration issue of AcrA$^{S/SD}$ constructs in SMA-PAGE is unknown.

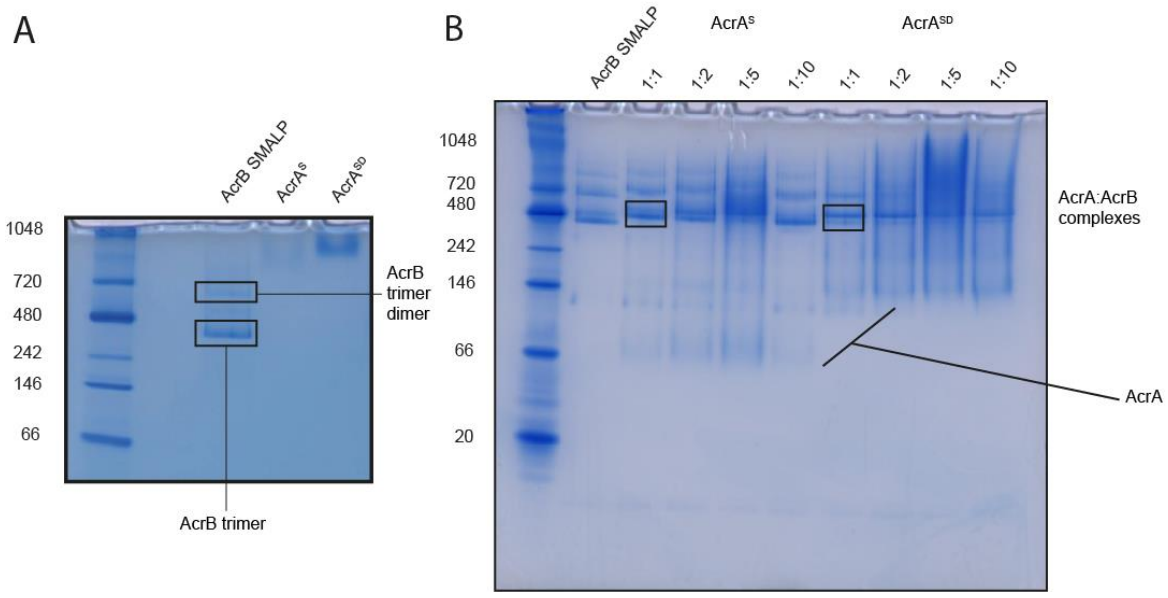


Figure 3.6. SMA-PAGE of AcrA constructs and AcrB SMALPs. A. SMA-PAGE of AcrB SMALPs, AcrA^{S} and AcrA^{SD}. Proteins loaded at 2 µM. AcrB runs as a trimer at ~400 kDa, with a trimer-dimer band also observed. AcrA constructs exhibit migration problems through the gel. **B.** SMA-PAGE of AcrA$^{S/SD}$ and AcrB, with increasing concentrations of AcrA. AcrB kept at 1 µM, and AcrA constructs range from 1-10 µM. Subcomplexes between AcrA and AcrB labelled on the gel.

Figure 3.6b examines the binding of AcrA$^{S/SD}$ to AcrB with increasing concentrations of AcrA relative to AcrB. Samples containing AcrB and AcrA$^{S/SD}$ show bands representing AcrA (highlighted in **Figure 3.6b**). This is likely due to the presence of free SMA or SMALPs, non-specifically associating with AcrA and providing extra charge for it to migrate successfully. For wells containing AcrA^{S}, a clear shift can be seen on the AcrB trimer band, representing a 1:1 binding model, as previously reported. AcrA^{S} concentration at 1:1 and 1:2 (AcrB:AcrA) look the same, but at 1:5 there is a larger band slightly higher than the 1:1 subcomplex band, suggesting more AcrA molecules may be able to bind in excess. However, at 1:10 excess of AcrA^{S}, this band disappears. For samples containing AcrA^{S}, the AcrB trimer band completely shifts, suggesting the entire AcrB sample binds AcrA in some regard, further highlighting the affinity of the binding interaction. Interestingly, the trimer-dimer bands of AcrB also shifts in the presence of AcrA^{S}, providing further confidence for this interaction.

The wells containing AcrA^{SD} also see a shift of the AcrA trimer band, slightly higher than for AcrA^{S} accounting for the increased size of the pseudo-dimer. Furthermore, at a ratio of 1:5, even higher intense bands can be seen suggesting higher order stoichiometries of binding to AcrB than seen for AcrA^{S} at the same ratio. This also suggests 1:5 is an optimum ratio of AcrB:AcrA for higher order binding. Overall, wells containing AcrA^{SD} contain more higher-order mass streaking than AcrA^{S}, suggesting a range of stoichiometries of binding, highlighting potential differences in how the constructs bind AcrB. However, there are limitations of SMA-PAGE; the presence of multiple bands, such as the trimer-dimers of AcrB, complicates interpretation. Furthermore, the streaking in certain wells does not allow for resolution of specific stoichiometries, which makes comparing between AcrA^{S} and AcrA^{SD} less reliable. However, it is clear from the results in **Figure 3.6** that AcrA^{SD} has a different propensity for binding AcrB than AcrA^{S}. Nonetheless, to further investigate these differences, mass photometry was utilised to study the interaction of AcrA and AcrB.

3.2.5.2 Mass photometry

To elucidate more information on the binding between the AcrA constructs and AcrB, mass photometry was used. Mass photometry uses the scattering of light caused by single molecules to measure mass.[201] One limitation of SMA-PAGE is the requirement of a loading buffer for analysis that differs from the experimental conditions; mass photometry allows proteins to be measured directly in their sample buffer, as long as glycerol is removed.[111] Therefore, proteins were diluted in phosphate buffer saline (PBS) at pH 7.0 to stocks of 200 nM, and measured on the mass photometer at 50-10 nM by droplet dilution. The mass photometer was calibrated using the unstained NativeMark protein ladder (ThermoFischer) according to manufacturer's recommendations.

Firstly, the AcrB SMALP sample was characterised using mass photometry (**Figure 3.7a**). A dominant peak with a high normalised count at 502 ± 60 kDa was observed, corresponding to AcrB SMALP; this was in good agreement with our SMA-PAGE value and previous determinations.[63] A peak at 960 ± 51 kDa could also be visualised, representing the trimer-dimer seen in SMA-PAGE results. Another peak at 51 ± 10 kDa was most likely free-SMALPs in the sample. To confirm this was unoccupied SMALPs, polar lipid extract (PLE) from *E. coli* (Avanti) was solubilised in 2.5% SMA and analysed on the mass photometer (see **Appendix 5**). Lastly, another peak was seen at 145 ± 66 kDa, which is likely a co-puritant solubilised by the SMA (see **chapter 5.2.1**). Overall the observed mass photometry results presented polydisperse spectra with several peaks assigned to different species, which has been previously reported for proteins in SMALPs.[205] Furthermore, the SD of the peaks corresponding to SMALPs are fairly high, likely arising from the varying number of lipids and SMA polymers per disc, increasing the heterogeneity of the samples and affecting the masses.

Figure 3.7b shows the three mass photometry traces representing AcrB SMALP (blue), $AcrA^{S}$:AcrB (green) and $AcrA^{SD}$:AcrB (pink). The $AcrA^{S}$:AcrB trace shows a decrease of the AcrB SMALP peak and a shifted new peak, representing the 1:1 $AcrA^{S}$:AcrB. It would be expected that the mass would shift ~40 kDa, representing the size of an AcrA monomer, but the mass for this is 512 ± 43 kDa. However, the SD is 43 kDa, so the mass shift is within error. AcrA is a small molecule for mass photometry, which only has a mass range from 30-40 kDa upwards, and combined with the SMALP disc heterogeneity due to the lipid and polymer content, this can lead to differences between expected and calculated mass.[201] However, visually the peak has shifted and has a shoulder not seen for AcrB alone, suggesting an $AcrA^{S}$:AcrB subcomplex. Furthermore, lots of $AcrA^{S}$ monomer can be seen at the 58 ± 10 kDa peak. Again, the observed mass is 18 kDa higher than the theoretical mass of $AcrA^{S}$, which is 40.8 kDa, due to the fact AcrA is near the mass resolution of the instrument. Interestingly, the trimer-dimer of AcrB appears to decrease, suggesting AcrA is able to disrupt the trimer-dimer association of AcrB in SMALPs.

Mass photometry of $AcrA^{SD}$:AcrB trace presents slightly differently. A broad peak spanning 500-750 kDa can be seen, with an average mass determination of 611 ± 74 kDa, representing the $AcrA^{SD}$:AcrB subcomplex, with the expected mass shift. However, the broadness of the peak and larger SD compared to AcrB alone (± 60 kDa) suggests heterogeneity in the $AcrA^{SD}$:AcrB subcomplexes formed, that is not just caused by the heterogeneity of SMALPs but by the binding of $AcrA^{SD}$, agreeing with the SMA-PAGE results. This also results in the lower number of counts observed for the $AcrA^{SD}$:AcrB, as the complexes are spread over a much larger mass area due to the heterogeneity. The spread of the peak and SD value suggest the two $AcrA^{SD}$ molecules may be able to bind AcrB, creating a pseudo-tetramer bound to AcrB. Furthermore, there is no $AcrA^{SD}$ alone visible in the trace, suggesting it is all

bound to AcrB. In addition to this, there is no AcrB trimer-dimer in this trace, implying AcrB has a much greater affinity to bind $AcrA^{SD}$.

Overall, the SMA-PAGE and mass photometry results suggest AcrB has a greater affinity to binding $AcrA^{SD}$, and different stoichiometries of binding can occur, not seen for $AcrA^{S}$. This supports the idea that dimerization is important for binding to AcrB, and can promote an increased affinity for the binding interaction.[367,410] The propensity for $AcrA^{SD}$ to from higher order complexes with AcrB may be a defining feature of how AcrA forms a trimer of dimers in the assembled AcrAB-TolC complex. One limitation of SMA-PAGE is the requirement of sample and running buffers, often at alkaline pH's, limiting the conditions that can be tested to monitor AcrA complexes with AcrB. Mass photometry bypasses this problem, as samples can be measured directly in their sample buffer, so theoretically, this means binding can be measured at different pH's; in this instance that was not possible due to the pH limitations of SMA. SMA is only soluble at neutral-basic pH's, because if the maleic acid group becomes protonated, SMA precipitates out of solution.[111] Therefore it was not possible to monitor the binding of AcrA to AcrB at pH 6.0; future experiments could use AcrB in polystyrene-co-maleimide lipid particles (SMILP) nanodiscs to bypass the pH dependence of SMA (see **chapter 5.1**).

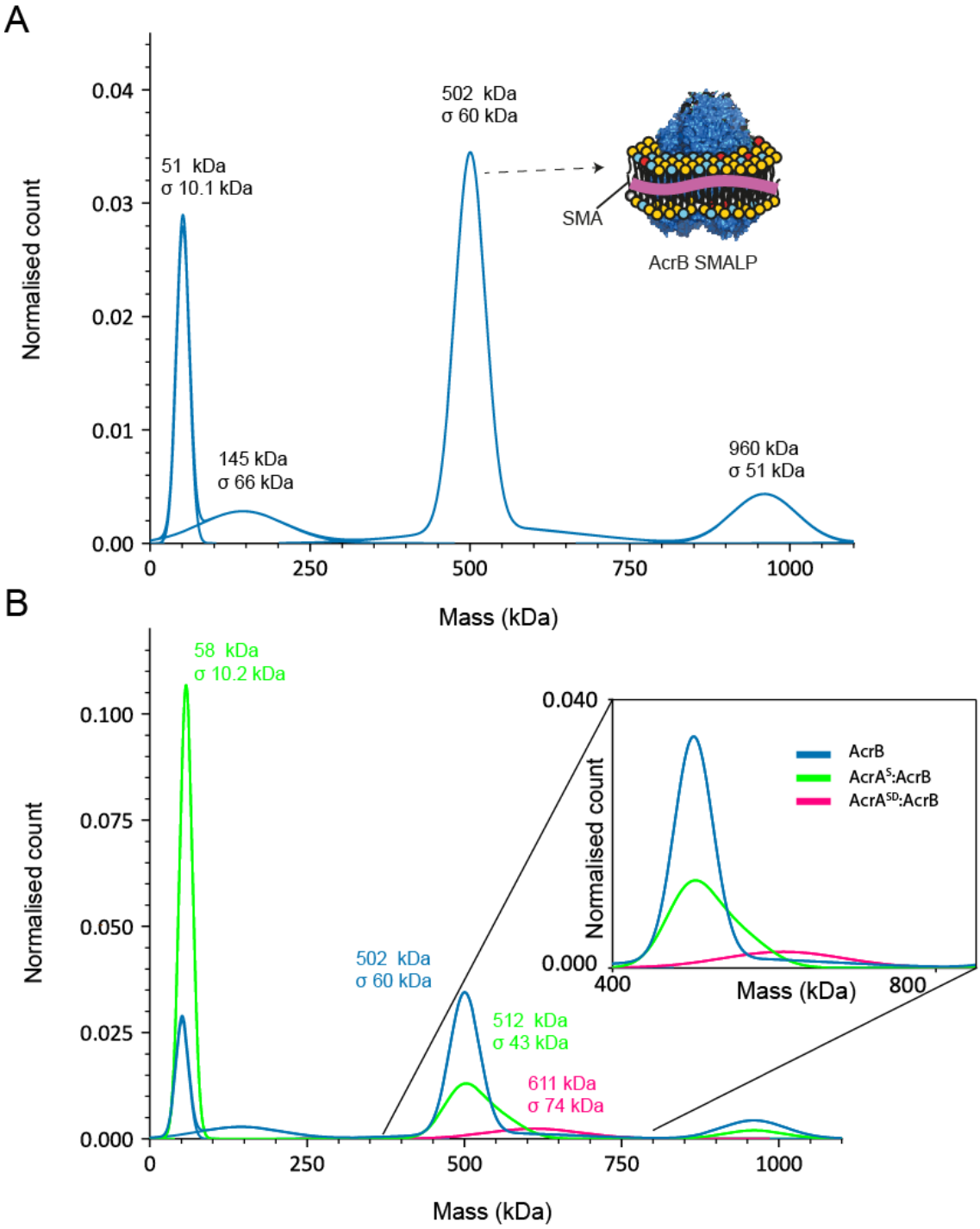


Figure 3.7. Mass photometry of AcrA$^{S/SD}$:AcrB binding. A. Mass photometry of AcrB SMALPs. **B.** Mass photometry of AcrB SMALPs (blue), AcrA^{S}:AcrB (green) and AcrA^{SD}:AcrB (pink). Insert of the AcrA:AcrB subcomplexes provided to better observe the differences in binding. Proteins added at a molar ratio of 1:1. Masses of identified peaks labelled on the plot.

3.2.6 Differences in binding to peptidoglycan

The affinity for AcrA to peptidoglycan has been reported previously.[38,460] Therefore, to monitor whether pseudo-dimerization effects the binding interaction to peptidoglycan, a pull-down assay was developed. Traditionally, pull-down assays are an *in vitro* technique used to detect protein-protein interactions (see **section 5.2.7.2**), but in this instance it can be used to identify interacting partners (AcrA and peptidoglycan).[475] **Figure 3.8** displays the workflow behind the peptidoglycan pull-down assay. Briefly, insoluble fragments of peptidoglycan from *E. coli* (Invivogen) were incubated with protein for 30 mins at room temperature, with constant shaking to prevent the peptidoglycan settling. The peptidoglycan was then pelleted by centrifugation at 17,000 x g for 5 mins. The supernatant represents unbound sample, so an aliquot was taken. The remaining supernatant was discarded, and the peptidoglycan washed three times in 1.5 mL PBS, to remove all unbound proteins. The washed pellet was resuspended in Laemmli buffer and boiled at 95 ˚C for 5 mins, and samples were analysed via SDS-PAGE.

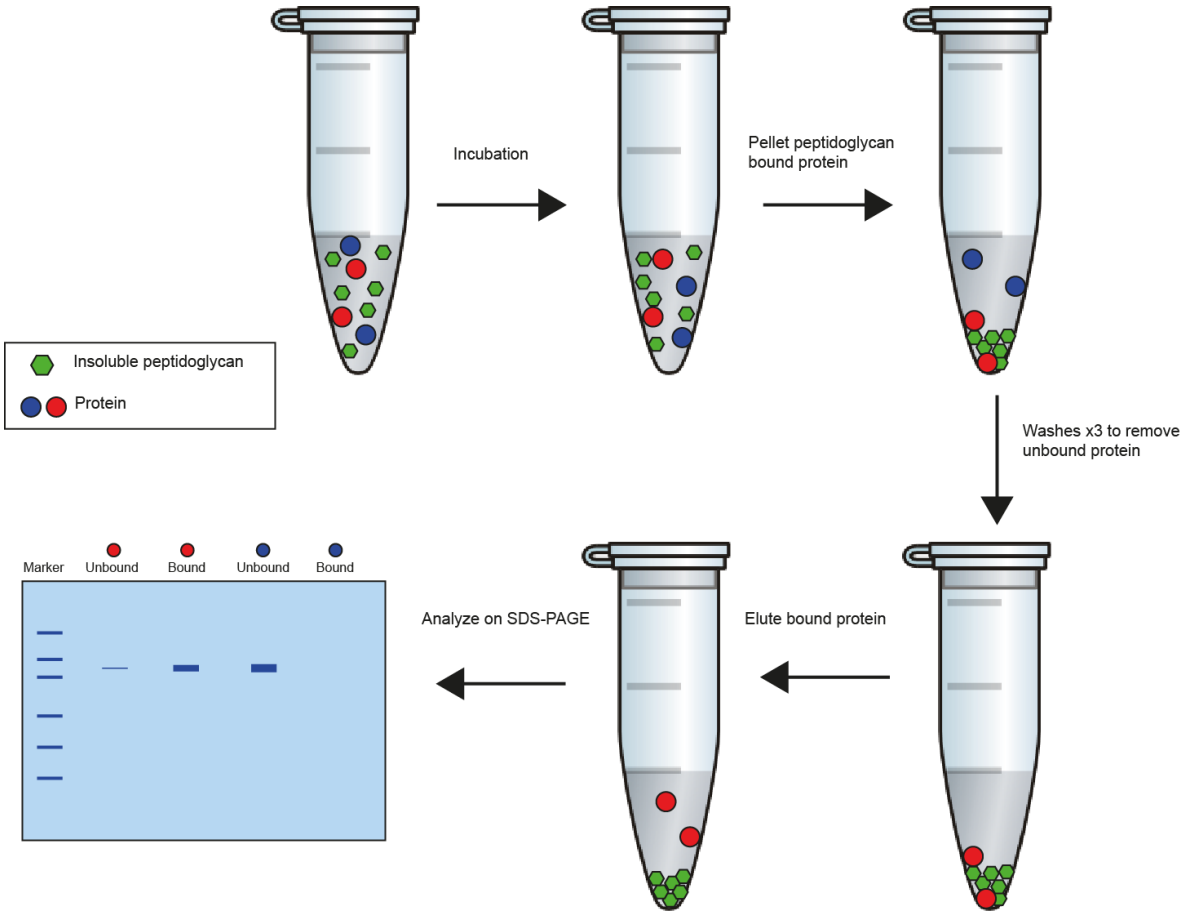


Figure 3.8. Peptidoglycan pull-down assay. Schematic detailing the pull-down assay.

The assay was optimised, as displayed in **Figure 3.9.** Firstly, differing amounts of AcrA was added to 200 µg peptidoglycan to see the optimal amount, based on previous work (**Figure 3.9a**).[476] 25, 50, 75 µg was added to the peptidoglycan, in a reaction volume of 150 µL. The input lane of AcrA^S shows it mainly appears at ~40 kDa, with a small fraction running at ~90 kDa as previously observed in **section 2.2.1.2**. **Figure 3.9** confirms AcrA's affinity for peptidoglycan as it can be observed in the bound fraction. The portion of AcrA^S in the bound fraction was small but looked most intense at 50 µg so this was the amount of protein selected for downstream experiments. To ensure the observed interaction was specific, a control experiment was designed. Firstly, insoluble peptidoglycan from *Staphylococcus aureus* (*S. aureus*) was trialled, but AcrA showed crossover ability to bind to this peptidoglycan (**Figure 5.18a**, **section 5.2.7.2**). Therefore, a control protein was chosen that should not bind peptidoglycan and had a similar mass and isoelectric point (pI) to AcrA. Alcohol dehydrogenase (ADH) from *Saccharomyces cerevisiae* (Sigma) has a molecular weight of 36.8 kDa and a pI of 6.21 and was therefore selected as a control as this closely matches AcrA which has a molecular weight of 40.8 kDa and a pI of 6.40. The input lane of ADH shows it runs slightly below 40 kDa as expected, but also presents some SDS resistant dimers. More information on the optimisation of this assay can be seen in **section 5.2.7.2**. **Figure 3.9b** shows a repeat of the experiment shown in **Figure 3.9a** using ADH instead of AcrA. As expected, no ADH was observed in any of the bound fractions.

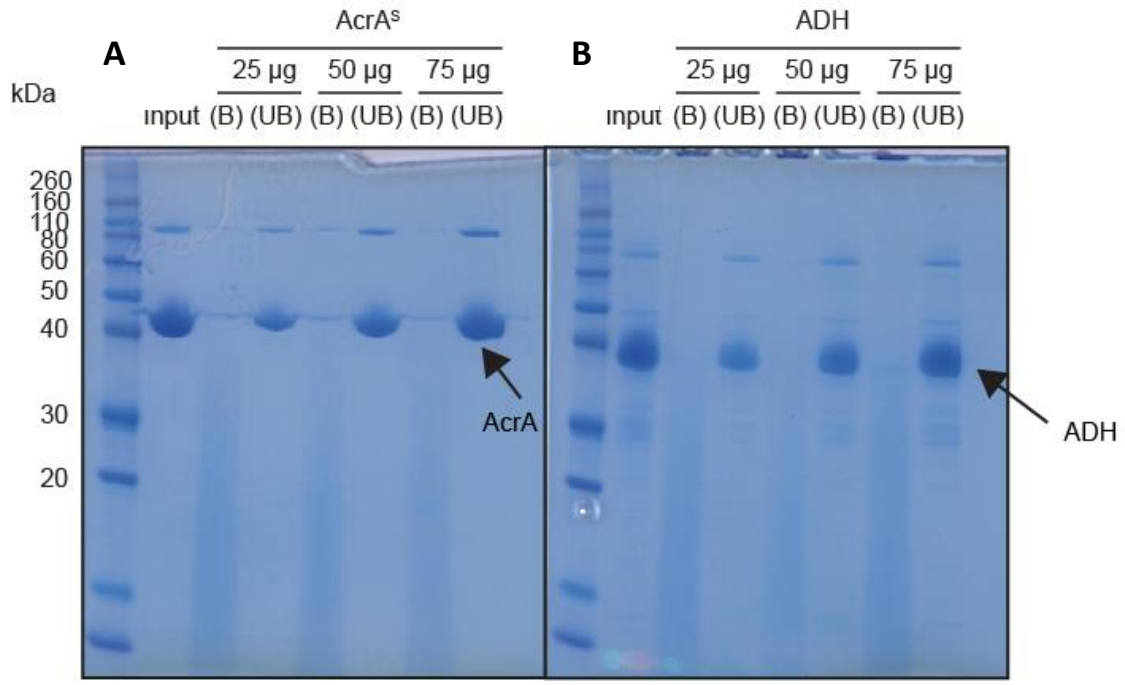


Figure 3.9. Peptidoglycan pull-down assay optimisations. A. 25, 50, 75 µg of AcrA^S incubated with 200 µg peptidoglycan from *E. coli*. AcrA^S is visible in the bound fraction at all three amounts. **B.** Experiment repeated with ADH. ADH not present in any of the bound fractions.

Using this assay, a qualitative comparison was made between $AcrA^{S}$ and $AcrA^{SD}$ binding peptidoglycan at pH 6.0 (**Figure 3.10**). 50 µg of AcrA proteins were added to 200 µg peptidoglycan. As can be seen in **Figure 3.10**, $AcrA^{SD}$ also has an affinity to peptidoglycan, and it is seen in the bound fractions. The relative intensities of the AcrA constructs in the bound fraction are similar, suggesting pseudo-dimerization does not increase the affinity of AcrA to peptidoglycan. This suggests the α-helical domain of one AcrA protomer is sufficient to bind peptidoglycan, and the packing of several hairpins does not enhance binding to peptidoglycan. MD simulations by Gumbart *et al*. (2021) suggests Lys131 and Lys140 form hydrogen bonds with peptidoglycan, further supporting a dimer interface is not required.[257] Peptidoglycan may provide a platform for AcrA to bind and open TolC; previous work has shown the affinity for both TolC and AcrA to peptidoglycan is enhanced when both proteins are present.[476] Furthermore, Mg^{2+} was added to AcrA constructs to see if it increased the amount of AcrA found in the bound fractions. As an *in vitro* experiment, the lack of the native periplasmic environment may affect the affinity of AcrA to peptidoglycan. However, it appeared Mg^{2+} had no effect on the binding to peptidoglycan.

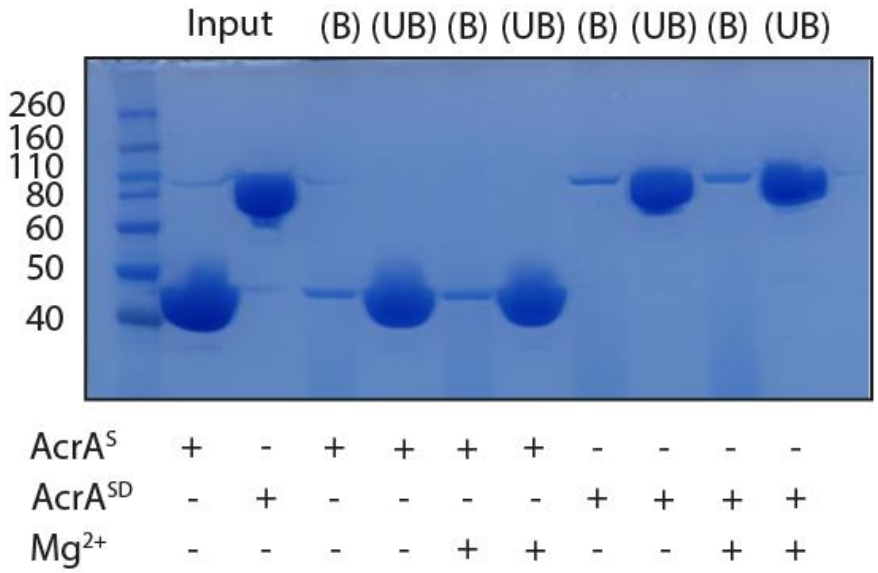


Figure 3.10. The effect of pseudo-dimerization on the binding of AcrA to peptidoglycan. 50 µg of $AcrA^{S/SD}$ incubated with 200 µg peptidoglycan from *E. coli* in the presence of ± Mg^{2+}. $AcrA^{S/SD}$ are visible in all bound fractions. B represents bound, UB represents unbound.

3.3 Conclusions

Previous work has established MFPs, such as AcrA, act as a functional dimer, that trimerize in assembled complexes to form a sealed channel in the periplasm.[366,367,410,457–460] Each protomer in the AcrA functional dimer binds differently to AcrB and TolC, further suggesting the importance of studying AcrA as a dimer.[389,410,412,423] However, due to the propensity of lipidated AcrA to oligomerise in solution, and $AcrA^{S}$ presenting as a monomer, it is somewhat challenging to isolate a stable AcrA dimer. Therefore, in this chapter a pseudo-dimer construct was designed, expressed, and purified to infer structural biology information on the AcrA functional dimer. It was found that $AcrA^{SD}$ was a homogenous pseudo-dimer (technically monomeric as $AcrA^{SD}$ consists of a single polypeptide chain) with unique backbone dynamics compared to $AcrA^{S}$. Furthermore, it was observed that $AcrA^{SD}$ has an increased propensity to form higher order complexes with AcrB, but there was no observed difference in the affinity to peptidoglycan between $AcrA^{SD}$ and $AcrA^{S}$.

MFPs bridge efflux transporters with the constituent OMF to assemble a tightly sealed channel; to ensure effective efflux, the MFP must bind and open the OMF.[410] Recent cryo-EM data has supported a tip-to-tip model where the α-helices of AcrA and TolC show inter-digitation to give a cogwheel-like structure.[373,420,427] MD simulations of the MexA-OprM interface, from *Pseudomonas aeruginosa* (*P. aeruginosa*), supported the same mechanism for this homologous system.[420] The HDX-MS data show reduced dynamics in the $AcrA^{SD}$ α-helical domain, suggesting pseudo-dimerization is somewhat restricting the movement of the α-helices. This may be important so AcrA can correctly position its α-helices to efficiently open TolC. Still, drug binding to AcrB triggers conformational movement that is propagated to TolC by AcrA, leading to the twisting of AcrA protomers and the subsequent opening of TolC during efflux.[428] Therefore, there still has to be a degree of flexibility in the α-helices in order for AcrA to function effectively in the efflux mechanism. MD simulations of MexA have also shown that dimers retain an elongated shape and flexibility in the α-helical domain.[458] Thus dimerization is likely to have to provide a balance, allowing the α-helices to be correctly positioned, whilst preserving conformational flexibility.

The use of a pseudo-dimer construct to gain insights into the functional dimer unit of AcrA presents advantages and disadvantages. Advantageously, $AcrA^{SD}$ provided a clean protein sample, that was easy to express and purify, and that was a homogenous soluble pseudo-dimer. This allowed for the inference of biological information regarding an AcrA dimer unit that was directly attributable to the pseudo-dimer itself and no other species of oligomers as per other samples of AcrA. Furthermore, its homogeneity combined with the fact it could be produced at high concentrations made it perfect for HDX-MS investigations, which in turn provided the first look at the differences in structural dynamics

between the different AcrA constructs. However, there are some limitations; as this construct is a rationally designed mutant that does not exist *in situ*, it is unknown how reflective results are of the native AcrA functional dimer unit. This pseudo-dimer is used to provide insights on how the functional dimer may behave, but it does not directly report on its behaviour. Furthermore, as this is a new construct, there is no structural information regarding the conformations that $AcrA^{SD}$ exhibit *in vitro* or *in vivo*. From the biophysical and HDX data reported in this chapter, it can be confirmed that $AcrA^{SD}$ is folded, but it is unknown how reflective this is of wild type AcrA dimers in the AcrAB-TolC complex. In conclusion, the use of the pseudo-dimer provides a novel way of gaining the first biological insights into how dimerization is affecting AcrA *in vitro*, which traditionally has been difficult to do.

3.4 Material and methods

For methods described in previous chapters, the reader will be redirected.

3.4.1 Reagents

All reagents purchased from ThermoFischer Scientific or Sigma Aldrich/Merck unless otherwise stated. The highest quality reagents were always prioritised.

3.4.2 Molecular biology

3.4.2.1 AcrA pseudo-dimer construct

The AcrA plasmid containing the soluble AcrA pseudo-dimer construct was received from the Zgurskaya laboratory. It was a pET21d+ plasmid that was composed of one AcrA molecule lacking the signal peptide, a Cys25 mutated to a Met residue and an additional Ala residue afterwards. This was connected to a second AcrA molecule lacking the signal peptide and Cys25 through a TRRIT linker. The construct had an LE linker and an His_6-tag on the second AcrA molecule. **Appendix 6** shows the sequence of this construct. All cloning and sequencing for this construct performed by the Zgurskaya laboratory.

3.4.3 Protein expression and purification

3.4.3.1 $AcrA^{SD}$

pET21d+ containing $AcrA^{S}$ was transformed into C43(DE3)Δ*acrAB E. coli* cells. 7 mL of an overnight Luria-Bertani (LB) culture was added to 1 L of pre-warmed LB broth containing 100 µg/mL ampicillin. Cells were grown at 37 ˚C until an OD_{600} of 0.5-0.6 was reached, then 1 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG) was added to induce protein expression. Cells were harvested by centrifugation at 4200 x g for 30 mins at 4 ˚C and washed with ice-cold PBS. The cell pellet was frozen at -20 ˚C overnight.

The cell pellet was thawed and resuspended in 40 mL Buffer A (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 5 mM imidazole). The cell suspension was supplemented with protease inhibitor mini tablets (1 tablet per 10 mL – as per Roche recommendations), 1 mM phenylmethylsulfonyl fluoride (PMSF), 2 µL Benzonase, 1 mM ethylenediaminetetraacetic acid (EDTA) and 20 mg/mL lysozyme. The cell suspension was sonicated (Fisherbrand™ Model 120 Sonic Dismembrator, ThermoFisher) at 40 amps, 6 times for 15 seconds, with a 90 second break between each sonication. Insoluble cell debris was removed by centrifugation at 20,000 x g for 20 mins at 4 ˚C.

The membrane fraction was isolated by high-speed centrifugation at 200,000 x g for 1 hour at 4 ˚C. The supernatant was put through an AKTA purification system for Ni^{2+} chelation chromatography and then SEC. The supernatant was loaded onto a 1 mL HiTrap Nickel column in Buffer B (50 mM $NaHPO_4$, pH 7.4, 300 mM NaCl, 20 mM imidazole, 10 % (v/v) glycerol). The column was washed with 10 column volumes (CV) of Buffer B, then 20 CVs of Buffer B with 50 mM imidazole, then $AcrA^{SD}$ was eluted with Buffer C (50 mM $NaHPO_4$, pH 7.4, 300 mM NaCl, 500 mM imidazole, 10 % (v/v) glycerol). The samples were buffer exchanged directly through injection onto a Superdex 16/600 GL SEC column (GE Healthcare) equilibrated in Buffer D (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 10% v/v glycerol). Peak fractions containing pure $AcrA^{SD}$ were pooled. Stored at 4 ˚C for up to a month, or at -80 ˚C for longer term.

3.4.4 Polyacrylamide gel electrophoresis (PAGE) and Western blots

SDS-PAGE and Western Blots were completed as previously described.

3.4.4.1 SMA-PAGE

SMA-PAGE performed as previously described.[465] Samples were run on either a precast Novex Value 4-20% Tris-Glycine gel or nativePAGE 4-20 % Bis-Tris gel. Proteins diluted in 2x Novex Tris-Glycine Native Sample Buffer (Tris HCl (100 mM), glycerol (10%), bromophenol blue (0.00025%), pH 8.6) or 4x NativePAGE Sample Buffer (Bis-Tris (50 mM), 6 N HCl, NaCl (50 mM), Glycerol (10%), Ponceau S (0.001%), pH 7.2). The Tris-Glycine gels were run with 25 mM Tris, 192 mM Glycine, pH 8.8 Running Buffer. The Tris-Bis gels were run with 20x NativePAGE Running Buffer (see Thermofischer for composition). The protein ladder used was the NativeMark Unstained Protein Standard. 30-40 µL of protein was loaded per well. All samples were run for 90-120 mins at 150 V and 4 ˚C. Bands visualised by Brilliant Blue Coomassie stain or silver stain.

3.4.5 Mass spectrometry

Native MS and HDX-MS experiments completed as previously described.

3.4.6 Mass photometry

The mass photometry experiments in this chapter were performed on a Samux MP or a Two MP instrument (Refeyn) with glass coverslips for 60-90s. Protein samples were diluted to 100 nm stocks in PBS buffer. A droplet of 10 µL PBS was added to a sample well and the machine focused. Then 10 µL of sample was added directly to the droplet and mixed via pipetting. AcrA:AcrB added at a 1:1 Molar ratio. Recorded videos analysed through DiscoverMP software (Refeyn, version 2.2).

3.4.7 Circular Dichroism

The CD experiments performed in this chapter were completed on an Aviv Biochemical Model 410 Circular Dichroism Spectrometer (Lakewood, NJ, USA). For the thermal melts, proteins were diluted to 0.4 mg/mL and buffer exchanged into Tris-HCl pH 6.0. Wavescans between 190-260 nm were performed at 14 temperatures ranging from 25-90 ˚C with 5 ˚C increments. Tms were then calculated and analysed as described previously.

3.4.8 Peptidoglycan pull down assay

200 µg of PGN-ECndi ultrapure peptidoglycan (Invivogen) were divided up into separate Eppendorf tubes. 50 µg of $AcrA^S$ (in Protein Buffer, see **section 2.4.3.2**) was added to the peptidoglycan. Protein Buffer was added until the reaction volume was 150 µL. Samples were incubated at room temperature on a ThermoShaker (ThermoMixer C, Eppendorf) at 400 revolutions per minute (RPM) for 30 mins. Suspended peptidoglycan was pelleted by centrifugation at 17,000 x g for 5 mins. A 50 µL aliquot of the supernatant was taken to represent the unbound fraction. The remaining supernatant was discarded, and the peptidoglycan washed three times with 1.5 mL PBS. The washed pellet was resuspended with 150 µL Laemmli Buffer and boiled at 95 ˚C for 5 mins. After the samples were allowed to cool, 15 µL of each fraction were loaded onto an SDS-PAGE gel and ran as described in **section 2.4.4.1**.

Chapter 4: Conformational restriction shapes the inhibition of AcrA

This chapter has been adapted and modified from publication with additional discussion and detail:

Russell Lewis, B., Uddin, M.R., Moniruzzaman, M. et al. Conformational restriction shapes the inhibition of a multidrug efflux adaptor protein. Nat Commun 14, 3900 (2023). https://doi.org/10.1038/s41467-023-39615-x

4.1 Introduction

Multidrug resistant infections arise from bacterial pathogens with the ability to survive lethal doses from an array of structurally diverse compounds.[329] Antimicrobial resistance (AMR) is rising at an alarming rate; in 2019 bacterial multidrug resistance directly caused 1.27 million deaths, overshadowing fatalities caused by human immunodeficiency virus (HIV) and malaria combined.[209] If AMR continues to increase at the current trajectory, current standards of living will be significantly impacted.[206,207] A major mechanism of multidrug resistance is the activity of efflux pumps.[210,477] They are overexpressed in the presence of antibiotic exposure, and subsequently export a wide range of chemically diverse compounds, to lower intracellular antibiotic concentration and confer resistance.[329] As previously described in this work, the AcrAB-TolC multidrug efflux pump is a member of the resistance nodulation and cell division (RND) superfamily and native to *Escherichia coli* (*E. coli*).[223,224] It is prototypical of homologs across other ESKAPE bacteria, so research on AcrAB-TolC will have implications to other systems.[211] Finding ways to inhibit these multi-protein complexes is an essential battle in the war against rising AMR.

Efflux pump inhibitors (EPIs) have the potential to be a useful adjunctive therapy to 'revive' the activities of antibiotics during multidrug resistant infections.[403,478] However, to this date no EPIs have entered clinical trials, often due to toxicity issues and the promiscuous nature of AcrB to transport its inhibitors.[479–481] AcrB is the inner membrane RND protein of the AcrAB-TolC multidrug efflux pump that undergoes the rotational mechanism of drug efflux, and therefore it has been the main target of EPI design.[366] The first reported EPI was Phe-Arg-β-napthylamide (PAβN) described by Lomovskaya *et al.* (2001), and has been shown to inhibit AcrB and other clinically relevant efflux pumps such as MexB, MexC, MexE, and MexX from *Pseudomonas aeruginosa* (*P. aeruginosa*), as well as other efflux pumps from other Enterobacteriaceae species.[480,482] PAβN has been shown to potentiate antibiotic activity by restraining drug-binding pocket dynamics and thereby preventing the binding of other substrates to this site.[35,480] However PAβN has not been used in clinical trials due to its cytotoxicity.[483] Another class of EPIs that inhibit AcrB are the pyranopyridines (MBX) compounds, first described by MicroBiotix in 2014.[484] One derivative MBX-2319 was shown to bind to the periplasmic portion of AcrB in the T state, in the hydrophobic trap, and prevent substrate binding via steric hinderance in a mode of competitive inhibition.[479,485] The pyranopyridines have also yet to advance to clinical trials due to unfavourable pharmacological properties, but they remain a promising preclinical candidate.[483] Other classes of EPIs exist, including several that come from natural products.[480,483]

Due to the lack of clinical success targeting AcrB, there has been a need to explore other avenues in the quest to generate successful EPIs, with AcrA emerging as a potential target for inhibition.[379,403]

Membrane fusion proteins (MFPs) such as AcrA are critical for the efflux of antimicrobials, as they assemble into a trimer of dimers to ensure a sealed channel across the periplasm for the substrate to reach TolC.[366,367] This assembly is highly dynamic and serves structural and functional roles in drug efflux. Moreover, recent studies have shown AcrA to have more diverse functions than previously thought. AcrA has been identified as a bacterial 'necrosignal' within *E. coli* swarms.[486] Bacteria can employ flagella-driven motility to colonize favourable niches in a collective 'swarm'. Swarms are clinically relevant due to their non-genetic resistance to lethal levels of antibiotics; resistance is caused by high cell densities within swarms and is phenomenologically similar to the adaptive resistance exhibited by bacterial biofilms. When a subpopulation of the swarm dies, dead cells release AcrA which binds TolC on the outside of other live cells, stimulating efflux within the affected area and the upregulation of various efflux pumps.[487] Therefore, AcrA's diverse functions and critical roles within antibiotic resistance mechanisms make it an exciting new druggable target.

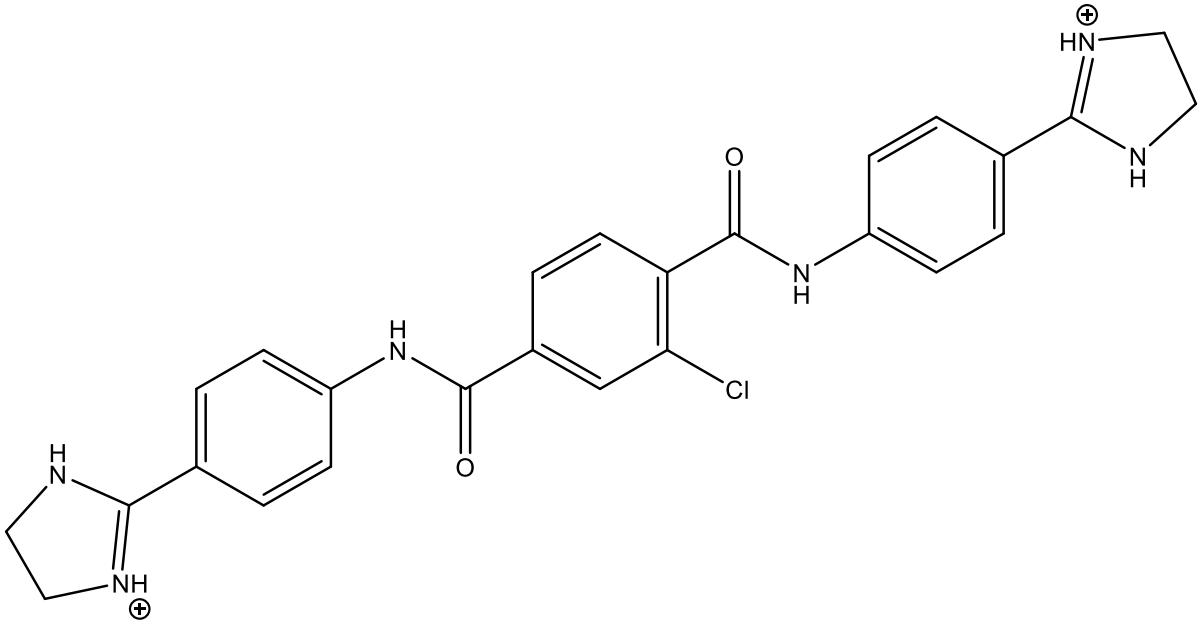


Figure 4.1. Structure of NSC 60339 EPI.

Recently, NSC 60339 was identified as an AcrA inhibitor through a joint experimental-computational screen.[403] NSC 60339 is a substituted phthalanilide (2-chloro-4',4"-bis(2-imidazolin-2-yl)terephthalanilide) and also an anti-cancer agent (**Figure 4.1**).[488] It was shown to potentiate both novobiocin and erythromycin in wild type-pore *E. coli* cells. The original work by Abdali *et al.* (2017) used *in vivo* proteolysis of AcrAB-TolC in the presence of NSC 60339 to show the structure of AcrA was different in the presence of the EPI, suggesting the mechanism of inhibition may involve structural changes to AcrA.[403] Computational methods by Darzynkiewicz *et al.* (2019) revealed six possible

binding sites on AcrA for NSC 60339 (**Figure 4.2**).[489] They were able to conclude NSC 60339 likely bound at or near site IV, located between the lipoyl and αβ-barrel domains, for several reasons. Firstly, NSC 60339 still bound to truncated AcrA lacking the membrane proximal (MP) domain, suggesting it wasn't the binding site. Secondly, mutations in site IV have the greatest impact on efflux; F81W impairs efflux significantly. Thirdly, fluorescence changes to AcrA mutants F81W and F254W (site IV) upon NSC 60339 binding most resembled corresponding changes in the wild type, whereas mutants at the other binding sites showed no change (except I343W, MP domain). This was further investigated by measuring Trp fluorescent quenching by potassium iodide with and without NSC 60339. NSC 60339 interfered with the quenching of F81W and F254W mutants of site IV only. Therefore, the evidence suggests that site IV is associated with NSC 60339 binding.

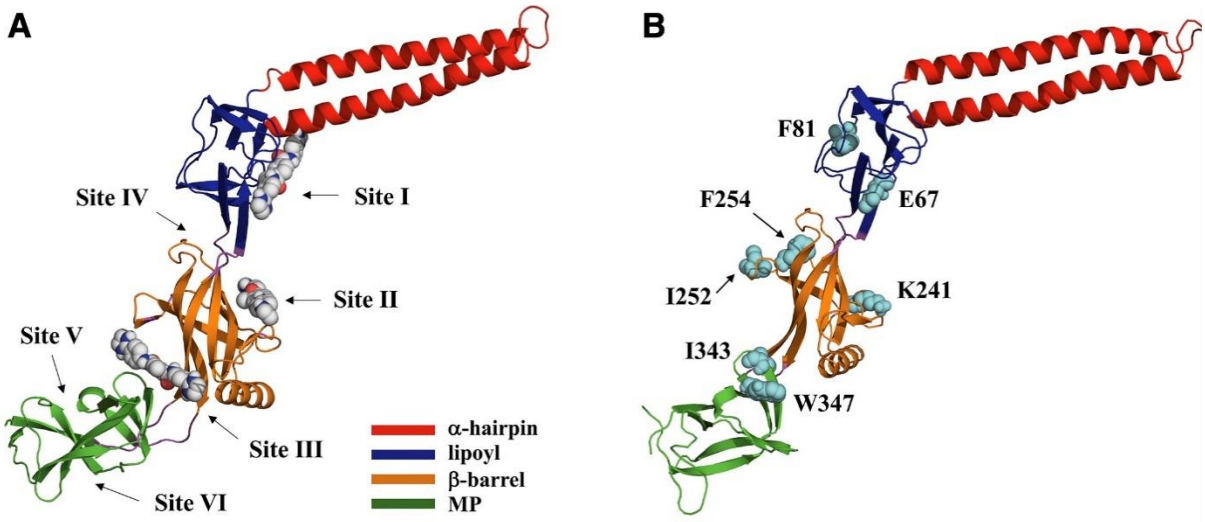


Figure 4.2. Predicted AcrA binding sites of EPIs. A. Inhibitor binding sites predicted from blind ensemble docking of NSC 60339, chlorobiocin and novobiocin (sites I-III) and FTMap (sites IV-VI). NSC 60339 is shown docked in some of the sites. **B.** Residues individually mutated to Trp in previous fluorescence quenching experiments shown as cyan spheres. Figure taken from Darzynkiewicz *et al*. (2019).[489]

Docking models of NSC 60339 in site IV presented by Darzynkiewicz *et al.* (2019) provide an idea of how NSC 60339 may be orientated in its binding site.[489] Due to the positive charge of the molecule, the side chain of R183 and the hydroxy group of T205 is pointed way from NSC 60339. It was suggested to form hydrogen bonding interactions with Q207 and the backbone carbonyl of H285. However, regardless of the biophysical evidence suggesting the binding site, and the subsequent docking

studies, the molecular mechanism of inhibition is not understood. Deciphering the mechanism of inhibition is important for the future design of EPIs.

Therefore, the aims of this chapter, was to determine the molecular mechanism of action for the AcrA inhibitor NSC 60339, using a combination of hydrogen deuterium exchange mass spectrometry (HDX-MS) and native mass spectrometry (native MS), molecular dynamics (MD) simulations, biophysical techniques, and cellular inhibition assays.

4.2 Results

4.2.1 Optimisation of HDX-MS studies of AcrA^S and NSC 60339

To investigate if NSC 60339 is having an effect on the conformation or backbone structural dynamics of soluble AcrA (AcrA^S), as previous results suggested, differential HDX (ΔHDX) between AcrA^S + NSC 60339 and AcrA^S alone was utilised.[403,489] However, before investigations could begin, a series of optimisation steps were required. Firstly, it was decided the investigations would be performed at pH 6.0. This is a closer reflection to the physiological environment of active AcrA, and previous binding kinetics had been completed at pH 6.0, showing NSC 60339 had a K_D of 78 µM to AcrA^S as determined by surface plasmon resonance (SPR) measurements.[403] Secondly, the solubility of NSC 60339 had to be considered for both AcrA^S and the HDX-MS experiments. NSC 60339 (MedChemExpress) required dimethyl sulfoxide (DMSO) to be soluble, and solubility was only possible up to 10 mM. Traditionally, ligand studies involve using a concentration of ligand 10 times higher than the Kd. Since the K_D is 78 µM, and the highest stock concentration of NSC 60339 is 10 mM, this would result in a high amount of DMSO in the sample, which would lower the deuterium content and possibly effect AcrA^S. As a quick control to see the effects of DMSO on AcrA^S, styrene maleic acid polyacrylamide gel electrophoresis (SMA-PAGE) of the AcrA^S:AcrB subcomplex was performed at different concentrations of DMSO (**Figure 4.3**).

As **Figure 4.3** shows, 5% DMSO showed no effect on the formation of the AcrA^S:AcrB subcomplex, suggesting AcrA is unaffected under this condition. From a stock solution of 10 mM NSC 60339 in 100% DMSO, this resulted in a 500 µM concentration of NSC 60339 (5% DMSO) in the labelling buffer. Furthermore, this gave an 87% protein:ligand complex and a 90.25% deuterium content in the final deuterated sample, which was deemed acceptable for HDX-MS experiments. It was also decided that AcrA^S be pre-incubated in 500 µM NSC 60339 for 30 minutes before HDX-MS experiments. Moreover, for samples without NSC 60339, 5% DMSO was kept consistent to ensure any observed changes are due to NSC 60339.

Lastly, before HDX-MS experiments, the effect of NSC 60339 and 5% DMSO on the quality of the HDX-MS data had to be checked. It was ensured that the pH of the quenched sample was still between pH 2.3-2.5. Furthermore, it was checked that 500 µM of NSC 60339 does not make 1 µM of AcrA^S aggregate. The number of peptides yielded in PLGS and DynamX analysis was the same as the number of peptides yielded without DMSO present. Lastly, drug solubility was tested at pH 6.0. The drug was soluble at pH 6.0 and 5% DMSO but would precipitate out of solution over the course of the HDX-MS experiments; this could be visually seen. Therefore, the 10 mM NSC 60339 stock was sonicated for 3

hours prior to HDX-MS experiments, which enabled the drug to stay in solution throughout the entire time course of the experiments.

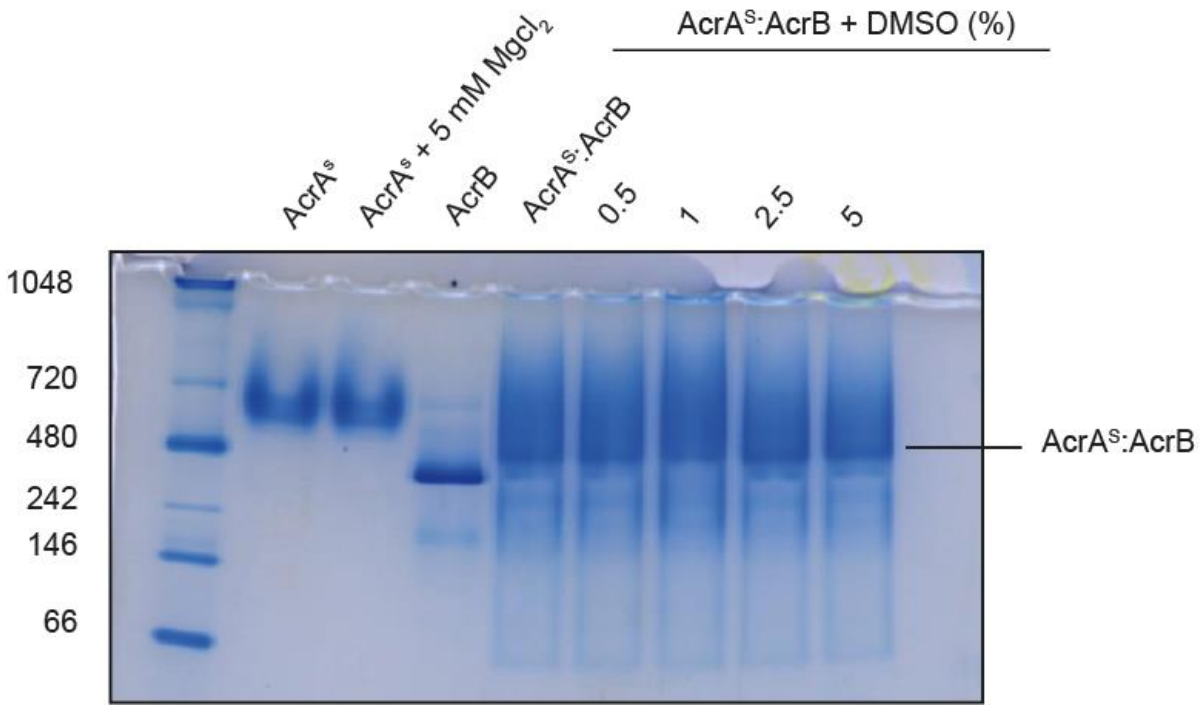


Figure 4.3. The effect of DMSO on the AcrA^S:AcrB subcomplex. AcrA^S and AcrB both added at 3.5 µM, at different amounts of DMSO ranging from 0.5-5%. The effect of $MgCl_2$ was also analysed on AcrA migration.

4.2.2 AcrA^S inhibition by NSC 60339

4.2.2.1 HDX-MS investigations

A ΔHDX experiment was performed on AcrA^S + NSC 60339 and AcrA^S alone at pH 6.0, for the same time points as the previous experiments in **chapters 2** and **3**. The buffers were the same as described in **section 2.2.3.2** except for the presence of 5% DMSO and ± 500 µM NSC 60339. The HDX-MS experiments AcrA^S + NSC 60339 and AcrA^S alone contained 2 biological repeats. The biological repeats required additional preparations of the protein sample and ensured the variability in exchange measurements due to different batches of protein expression/purification was quantified and averaged.[131] Therefore these HDX-MS experiments had 8 independent measurements ($n_{biological}$ = 2 and $n_{technical}$ = 4). Due to the presence of biological replicates, a ΔHDX cut-off was decided based on the actual confidence interval (CI) value calculated as previously described in **section 2.4.5.2.1**. Significance was deemed to be ≥ 0.33 Da change and with a P-value ≤ 0.01 in a Welch's *t*-test.

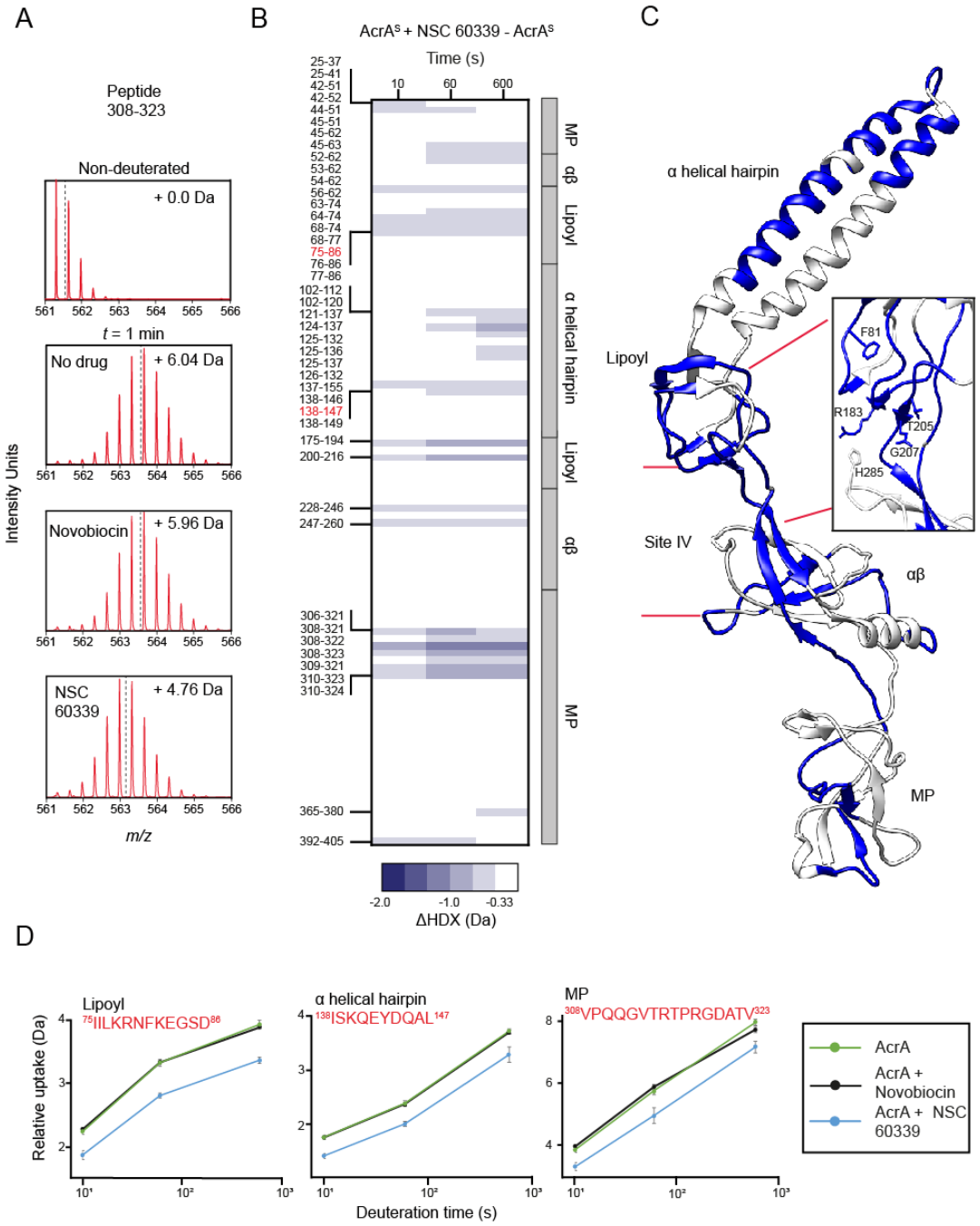


Figure 4.4. The effect of NSC 60339 on AcrA^S structural dynamics. A. The *m/z* spectrum for peptide 308-323 under non-deuterating conditions and deuterating conditions with DMSO, NSC 60339 and novobiocin. The centroid is represented by the dotted line, and the mass change of the deuterated samples is written in Daltons. **B.** Chiclet plot displaying the differential HDX (ΔHDX) plots for AcrA^S ± NSC 60339 for all time points collected. Blue signifies areas with decreased HDX between states. Significance was defined to be ≥ 0.33 Da change (see **section 2.4.5.2.1**) with a *P*-value ≤ 0.01 in a Welch's *t*-test (8 independent measurements: $n_{biological}$ = 2 and $n_{technical}$ = 4). White areas represent regions with insignificant ΔHDX. **C.** ΔHDX for the latest time point is painted onto the AcrA structure (PDB:5O66) using HDeXplosion and Chimera.[447,448] Zoomed in insert of site IV is shown, with the side chains of implicated residues highlighted.[489] **D.** Uptake plots for three peptides in different domains of AcrA. Uptake plots are the average deuterium uptake and error bars indicate the standard deviation.

In the presence of NSC 60339, AcrA^S exhibits stabilisation across all four domains, with the site IV region previously identified by Darzynkiewicz *et al.* (2019) being affected (**Figure 4.4**).[489] Site IV is shown on the AcrA structure in **Figure 4.4c**. Portions of site IV exhibited a significant decrease in deuterium uptake in the presence of NSC 60339 **(Figures 4.4b-d**), with peptides including residues F81 and F254 showing decreased uptake. For example, peptide ^{75}IILKRNFKEGSD^{86} exhibited significant protection in the presence of NSC 60339 across all three time points (**Figure 4.4d**). F81 (lipoyl domain) and F254 (αβ-barrel domain) have both been implicated in NSC 60339 binding, and F81 has been shown to be essential for efflux.[489] HDX reduction within these areas may be due to drug binding stabilising the local area through direct interactions with amide hydrogens, or a reduction in solvent accessibility within the implicated domains.

Aside from site IV, the HDX results reveal that NSC 60339 induced long-range stabilisation of AcrA backbone structural dynamics away from the proposed binding site. The α-helical hairpins show stabilisation on both helices and the flexible linker, primarily in the later time points. This is likely due to the stabilisation exhibited by the lipoyl domain due to NSC 60339, restricting the ability of AcrA to freely position the helices, possibly reducing its conformational plasticity. Additionally, the MP domain exhibits extensive protection in the presence of NSC 60339. The largest degree of protection in the MP domain occurs between residues 308-324, with multiple peptides within this region showing a sustained protection throughout the entire HDX time course. The raw data peptide envelope for 308-323 is shown in **Figure 4.4a**, and in the presence of NSC 60339 there is a far smaller shift in *m/z* than is seen for deuterated samples lacking NSC 60339, when compared to the non-deuterated reference sample.

Furthermore, in **chapter 2**, AcrA^S saw a reduction in structural dynamics in the presence of Mg^{2+}. Therefore, to check NSC 60339 still functioned the same way with Mg^{2+} present, as it would be in the periplasm, additional HDX-MS experiments were completed (**Appendix 7**). The same regions observed in **Figure 4.4** still show significant protection when NSC 60339 is added in the presence of Mg^{2+} and two uptake plots are shown in **Appendix 7**. Whilst this data is not part of this story, and requires further optimisation for further discussion, it provided a check to see if NSC 60339 still functions as expected in the presence of Mg^{2+}, which is a core part of the periplasmic environment.[3]

4.2.2.2 Molecular dynamics simulations

MD simulations were completed by Katie M. Kuo, Jerry M. Parks, and James C. Gumbart as part of a collaborative project.[425]

To further understand AcrA inhibition by NSC 60339, MD simulations were compared to the HDX-MS data. MD simulations are a powerful tool that provides an insight into protein structural dynamics by projecting a time dimension (μs – ms) to the static 3D coordinates of a protein structure solved by cryo-electron microscopy (cryo-EM) or crystallography.[490] However, a direct, quantitative comparison between HDX-MS and MD simulations can be difficult. Protein HDX is complex, with neighbouring residues having significant differences in their solvent interactions.[35] Furthermore, the different timescales, solution conditions, temperatures and complex kinetics involved complicates quantitative comparison between HDX and MD simulations.[490,491] Nevertheless, comparisons to MD calculated root mean square fluctuations (RMSF) and solvent accessible surface area (SASA) can provide a qualitative comparison to aid the interpretation of protein HDX.

MD simulations of the NSC 60339 bound monomer were initialised using the docked structure from Darzynkiewicz *et al.* (2019), and those with the apo used the same structure with the compound removed.[489] Specific details regarding the set-up of the simulations are described in **section 4.4.5**. Based on the docked structure (**Figure 4.5a**), F81 and F254 do not contact NSC 60339. However, in MD simulations F254 shows reduced RMSF (**Figure 4.5d**) and both F81 and F254 show reduced SASA (**Figure 4.5e**). This demonstrates how the binding of NSC 60339 has allosteric effects on both sides of site IV. The MD and HDX-MS results, together with the previous work, supports that site IV is the binding site of NSC 60339.

RMSF differences from the MD simulations of the bound and apo states show a high degree of stabilisation of residues 308-324 even compared to other residues (**Figure 4.5d**), supporting the HDX data and providing confidence that NSC 60339 can cause long range stabilisation of dynamics away from the proposed binding site. Peptide 308-323 shown in **Figure 4.4a** is highlighted as spheres in **Figure 4.5d** for clarity. The fact that NSC 60339 can bind to truncated AcrA lacking the MP domain, the MD simulations showing decreased MP domain RMSF in the docked structure with NSC 60339 in site IV and the HDX results showing decreased uptake in the MP domain supports that the effects seen in this domain are a downstream, allosteric effect of NSC 60339 rather than a secondary binding site.[489]

In contrast to the MP domain, the difference in RMSF for the lipoyl domain was modest (**Figure 4.5d**); however, portions of the lipoyl domain did see reduced SASA, possibly explaining the increased protection observed by HDX-MS (**Figure 4.4** and **Figure 4.5e).** Interestingly, some areas of the MP, lipoyl and αβ-barrel domains experienced increased SASA in the MD simulations but showed reduced RMSF or HDX (**Figure 4.4 and Figure 4.5**). This observation supports that, even with increased solvent accessibility, enhanced protection of the backbone amide to HDX in these regions is likely dominated by its restricted structural dynamics.

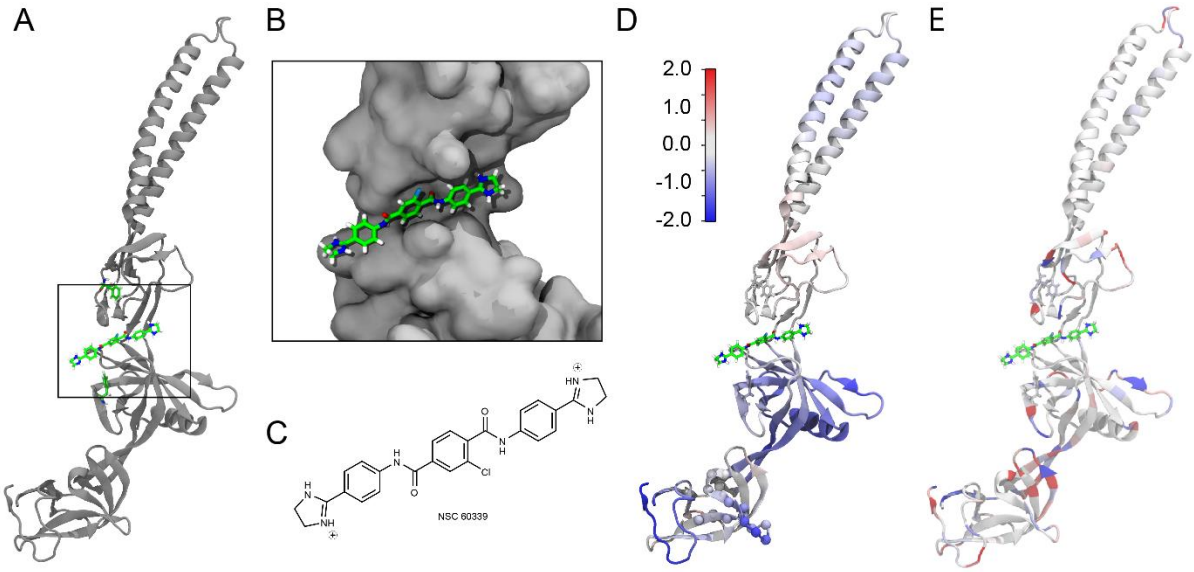


Figure 4.5. MD simulations of NSC 60339 bound to AcrA^S. A. Docked structure of NSC 60339 (centre) to AcrA from Darzynkiewicz *et al.* (2019).[489] F81 and F254 are shown above and below NSC 60339, respectively. **B.** Zoomed-in view of the NSC 60339 binding pocket cleft. AcrA is shown in a surface representation. **C.** NSC 60339 structure; calculated pKa values for the dihydroimidazoline groups are >9, suggesting that it is dicationic at pH 6.0.[492] **D.** AcrA coloured according to the difference in root-mean-square fluctuations (RMSF) between simulations of the bound and apo states, averaged over four replicas for each. Red indicates that the RMSF is greater in the bound state while blue indicates it is greater in the apo state (the colour range is from -2 Å, blue, to 2 Å, red as indicated by the colour bar). RMSF was calculated over the last 70 ns of each 100-ns simulation. **E.** Solvent accessible surface area (SASA) from MD simulations of an AcrA monomer. AcrA coloured according to the difference in SASA between the last 70 ns of 100-ns simulations of the bound and apo states, averaged over four replicas for each. Red indicates that the SASA is greater in the bound state while blue indicates it is greater in the apo state (the colour range is from -1 $Å^2$ to 1 $Å^2$). F81 and F254 are shown as sticks above and below the ligand.

4.2.3 Novobiocin control

To provide further confidence in the results describing AcrA^S inhibition by NSC 60339, a control was required. Novobiocin is an aminocoumarin antibiotic that is also a substrate that binds to AcrA (**Figure 4.6**). Previous work has shown that whilst novobiocin binds to AcrA, it is does not inhibit efflux activity and therefore is not classed as an EPI.[403] Furthermore, the addition of novobiocin had little effect on Trp fluorescent quenching by potassium iodide, suggesting this ligand does not significantly affect the structure of AcrA.[489] Therefore, as a control, HDX-MS experiments were repeated as in **section 4.2.2**, except with novobiocin. Novobiocin has a greater solubility than NSC 60339 and does not require DMSO. However, to provide consistency between the two experiments, DMSO was kept at 5%. Furthermore, the amount of protein:drug complex was kept consistent between experiments; previous SPR data confirmed novobiocin had a K_D of 4.3 µM to AcrA^S at pH 6.0.[403] Therefore, only 30 µM novobiocin was required in the labelling buffer and AcrA^S sample to achieve 87% fraction of ligand:protein complex. To match the previous experiments, significance was determined by the CI calculation previously described in **section 2.4.5.2.1**.[446] Therefore, significance was deemed to be ≥ 0.34 Da change and with a P-value ≤ 0.01 in a Welch's *t*-test.

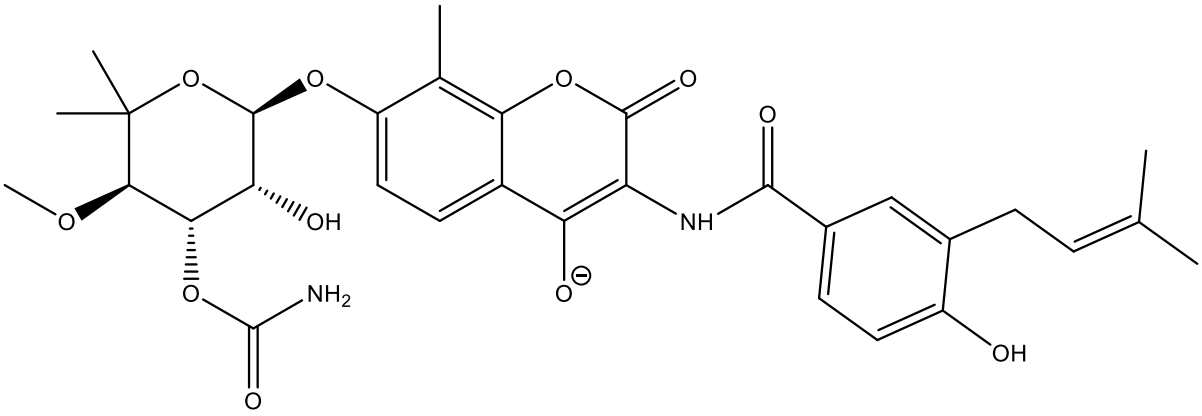


Figure 4.6. Structure of novobiocin. Novobiocin assigned a -1 charge as at pH 6.0 the molecule is deprotonated (pKa 4.3).[489]

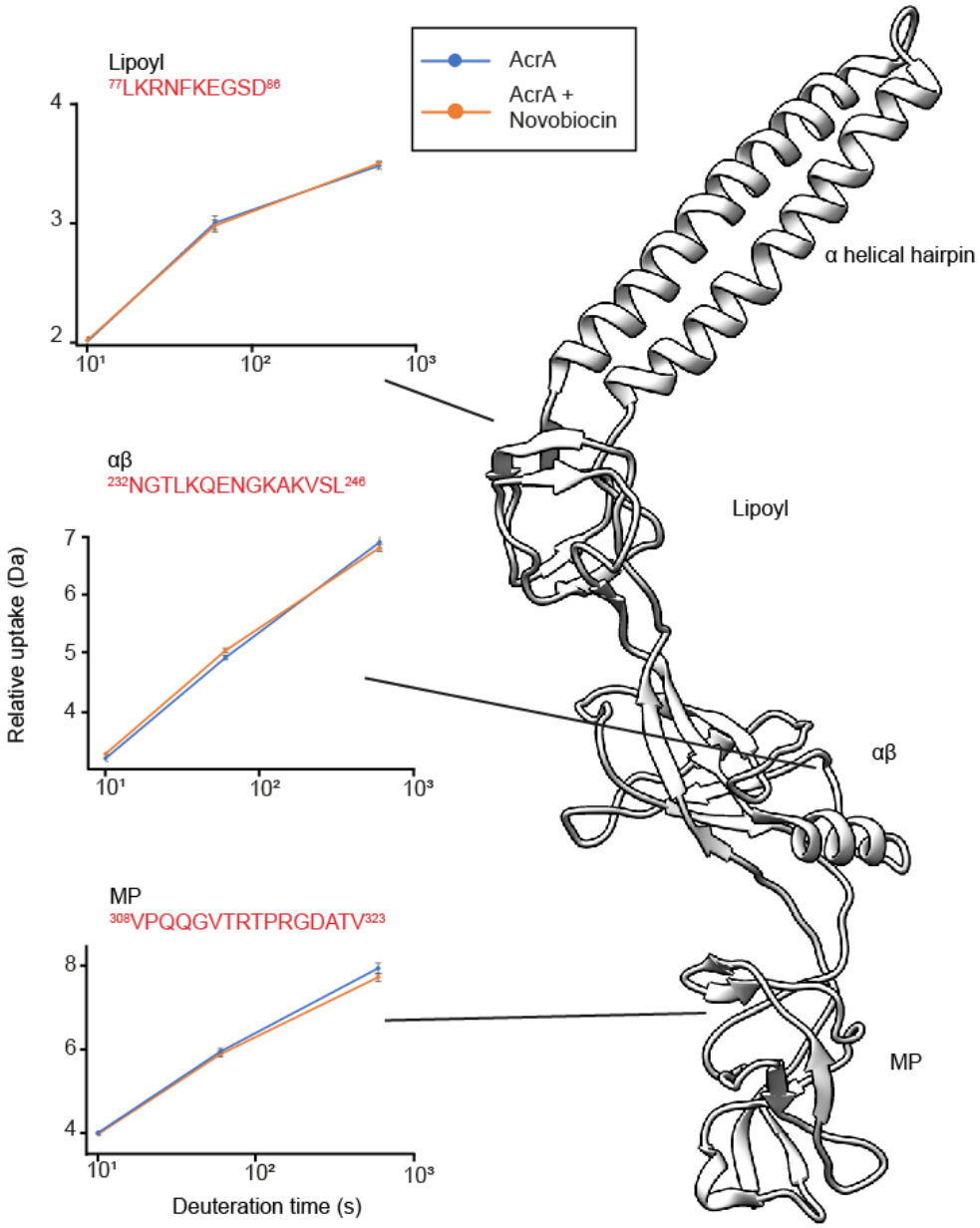


Figure 4.7. The effect of novobiocin on AcrA^S structural dynamics. ΔHDX for ((AcrA^S + novobiocin) – AcrA^S) for the latest time point is painted onto the AcrA structure (PDB:5O66) using HDeXplosion and Chimera.[447,448] Significance was defined to be ≥ 0.34 Da change (see **section 2.4.5.2.1**) with a *P*-value ≤ 0.01 in a Welch's *t*-test (n = 4 technical replicates). White areas represent regions with insignificant ΔHDX. Three peptide uptake plots can be shown, in areas that saw significant protection with NSC 60339. Uptake plots are the average deuterium uptake and error bars indicate the standard deviation.

Figure 4.7 shows the ΔHDX on AcrA^S + novobiocin and AcrA^S alone. The HDX data reveals that novobiocin has little observed effect on AcrA structural dynamics, within the examined time window, which is in stark contrast to NSC 60339. The AcrA structure in **Figure 4.7** is entirely white, revealing that no regions of AcrA exhibited significance ΔHDX in the presence of novobiocin. Furthermore, the uptake plots in **Figure 4.4d** show the deuterium uptake for AcrA^S + novobiocin (black) compared to AcrA^S alone (green) and AcrA^S + NSC 60339 (blue) in three domains that exhibit statistically significant protection when inhibited by NSC 60339. It shows the uptake is almost identical between AcrA^S + novobiocin and AcrA^S alone, and the uptake for AcrA^S + NSC 60339 is significantly reduced. The uptake plots in **Figure 4.7** provide examples of other peptides with no significant ΔHDX.

As a supporting experiment, native MS was deployed to monitor novobiocin binding to AcrA^S (**Figure 4.8, Table 4.1**) and the protein-ligand complex was observed. The mass of novobiocin is 613 Da, and the measured mass of AcrA^S was 40,841 Da. The lower charge state distribution (CSD) saw satellite peaks, reflecting measured mass increases of 617 and 1226 Da, reflecting the binding of one and two novobiocin molecules respectively (± 3 and 4 Da error). This provides further confidence to the HDX-MS experiments. Attempts were made to measure NSC 60339 binding to AcrA^S, but it was not possible, likely due to it having a larger K_D. Capturing protein-drug complexes with weaker affinity ligands can be difficult as the drug can often dissociate during ionisation.[493] Overall, it was observed that whilst it is a substrate of AcrA, novobiocin does not have global effects on structural dynamics like NSC 60339. This supports our premise that NSC 60339 is a successful inhibitor of AcrA due to its ability to stabilise AcrA across all four of its domains. AcrA relies on its dynamism and flexibility to function as an MFP, so restricting this conformational ability may be a promising target for inhibition in general.[398]

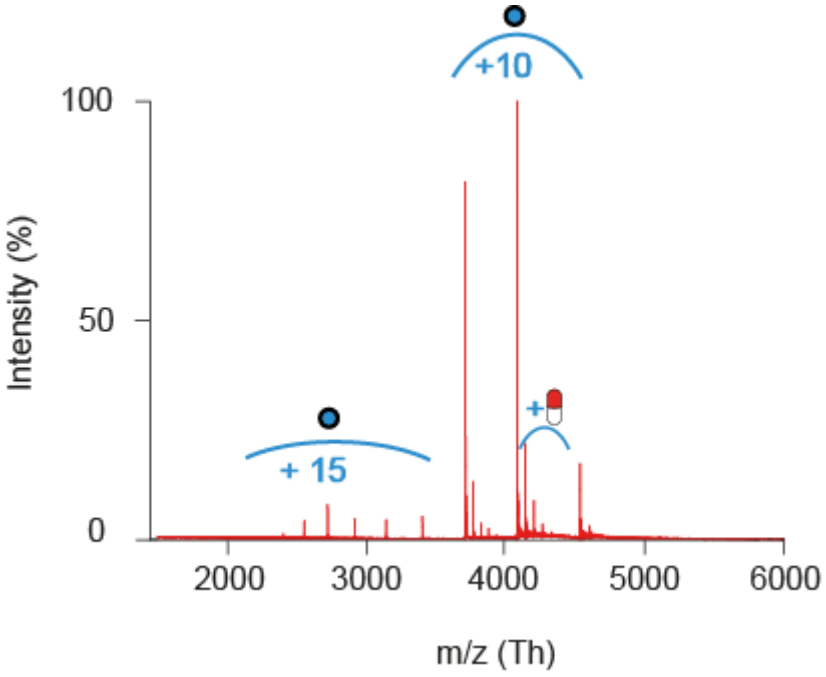


Figure 4.8. Native MS of AcrA^S and novobiocin at pH 6.0. Native-MS characterisation of AcrA^S construct with novobiocin at pH 6.0. Protein was buffer exchanged to 100 mM ammonium acetate prior to MS and diluted to 10 µM. Novobiocin added to a concentration of 30 µM, 5% DMSO final. Satellite peaks representing drug binding can be seen adjacent to peaks in the lower charge state distribution (CSD). See **Table 4.1** for masses.

Table 4.1. Native MS masses. Reported is the standard error of the mean within a single spectrum. Positive mass differences can be attributed to salt and/or detergent adducts. Novobiocin mass = 613 Da

	Measured mass (Da)	**Standard Error (± Da)**	**Mass difference (Da)**
AcrA^S pH 6.0	40,841	2	-
+ Novobiocin pH 6.0	41,458	3	617
+ Novobiocin pH 6.0 (x2)	42,067	4	1226

4.2.4 AcrA^{SD} inhibition by NSC 60339

4.2.4.1 HDX-MS investigations

To investigate whether AcrA pseudo-dimerization (AcrA^{SD}) effects inhibition by NSC 60339, the HDX-MS experiments were repeated with the AcrA^{SD} construct. Since no specific K_D has been calculated for AcrA^{SD} a qualitative SPR binding assay experiment was performed by the Zgurskaya group (**Figure 4.9a**). AcrA^{S} and AcrA^{SD} were immobilized onto a chip at similar densities of 4743 and 4054 response units (RU) respectively and increasing concentrations of NSC 60339 (6-200 µM). This was an effective way to qualitatively confirm similar binding of NSC 60339 to the two AcrA variants which suggests that pseudo-dimerization does not affect the mode of interaction of NSC 60339 with the AcrA binding site (**Figure 4.9a**). Therefore, the same K_D values between NSC 60339/novobiocin and AcrA^{S} were used for AcrA^{SD} experiments.

HDX experiments with AcrA^{SD} were repeated identically to those with AcrA^{S}. HDX-MS experiments AcrA^{SD} + NSC 60339 and AcrA^{S} alone also contained 2 biological repeats. Therefore, these HDX-MS experiments had 7 independent measurements ($n_{biological}$ = 2 and $n_{technical}$ = 3-4). As with the previous experiments in this chapter, the ΔHDX cut-off was decided based on the CI calculation described in **section 2.4.5.2.1.** Significance was deemed to be ≥ 0.35 Da change and with a P-value ≤ 0.01 in a Welch's *t*-test.

The results showed that NSC 60339 acts in a similar way even when presented with a dimer interface (**Figure 4.9**). There is statistically significant protection across all four domains as seen with AcrA^{S}, with pronounced HDX protection across the lipoyl domain, where site IV is located, suggesting the drug is acting in a similar area. Furthermore, AcrA^{SD} exhibits extensive protection in the MP domain, with the core of this stabilisation happening at the same region (residues 306-324). There is also a stabilisation across the α2 helix in the α-helical hairpin; however, fewer peptides and a smaller region compared to AcrA^{S}. One possibility for this may be due to the packing of the hairpins in the pseudo-dimer already making the region more stable compared to the monomer (**Figure 3.5**). Thus, the hairpins may be less effected by NSC 60339 inhibition in an AcrA dimer.

Interestingly, there is a reduced region of the αβ-barrel domain being stabilised, on the other side of site IV. When AcrA^{SD} is inhibited with NSC 60339, only residues 232-246 and 249-261 show significant protection, whereas AcrA^{S} shows protection between residues 228-260 and 52-62, which may suggest a difference in how the drug is interacting within the binding site. Overall, NSC 60339 appears to be inhibiting AcrA^{SD} in a similar way, by reducing AcrA's structural dynamics across the entire protein.

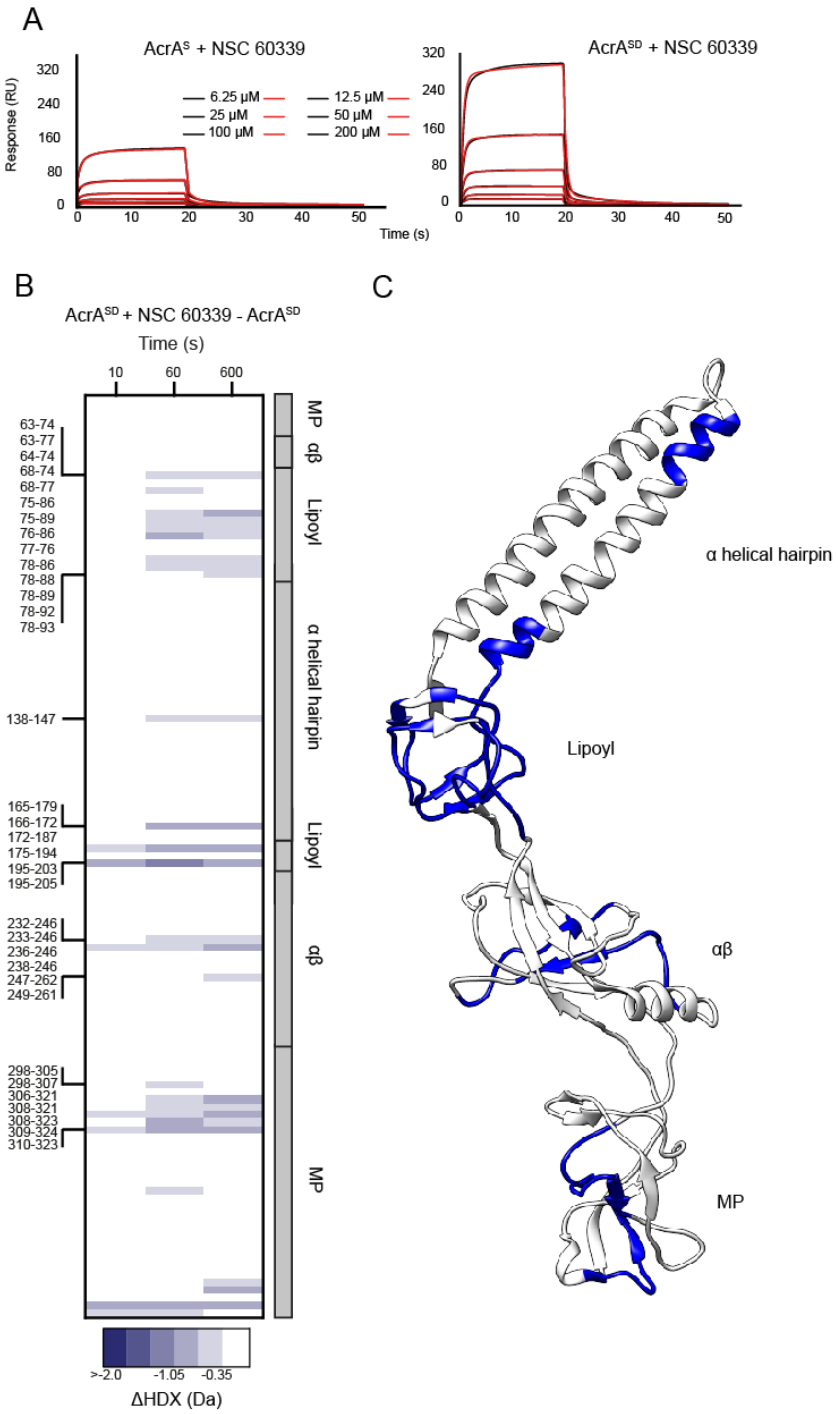


Figure 4.9. The effect of NSC 60339 on AcrA^{SD} structural dynamics. A. Surface Plasmon Resonance (SPR) for AcrA^{S} and AcrA^{SD} with NSC 60339. **B.** Chiclet plot displaying the differential HDX (ΔHDX) plots for ((AcrA^{SD} + NSC 60339) – (AcrA^{SD})) for all time points collected. Blue signifies areas with decreased HDX between states. Significance was defined as a ± ≥ 0.35 Da change (see **section 2.4.5.2.1.**), with a *P*-value ≤ 0.01 in a Welch's *t*-test (7 independent measurements: $n_{biological}$ = 2 and $n_{technical}$ = 3-4). White areas represent regions with insignificant ΔHDX. **C.** ΔHDX for ((AcrA^{SD} + NSC 60339) – AcrA^{SD}) for the 10 minute time point is painted onto the AcrA structure (PDB:5O66) using HDeXplosion and Chimera.[447,448]

4.2.4.2 Molecular dynamics simulations

MD simulations were completed by Katie M. Kuo, Jerry M. Parks, and James C. Gumbart as part of a collaborative project.[425]

As with $AcrA^{S}$, MD simulations were run of the bound and apo states of a modelled $AcrA^{SD}$ starting from an existing crystal structure (PDB 2F1M).[397] This conferred an advantage; $AcrA^{SD}$ consists of two AcrA molecules connected by a TRRIT linker. HDX-MS is an averaging technique, and as the two monomers produce the same peptides, the results are an average between the two protomers and differentiation between protomer 1 and 2 is not possible. However, in MD simulations, RMSF can be measured for protomer 1 and 2 individually (**Figure 4.10**). Each MD experiment was repeated four times, and the average RMSF for protomer 1 and 2 is shown in **Figure 4.10b**. Overall, it presents similarly to the MD seen for $AcrA^{S}$, with a reduction in RMSF for the bound state being observed over most of both copies of the protein compared to the apo state. Furthermore, no increase in RMSF is observed in the lipoyl domain, as was seen for $AcrA^{S}$, only a modest protection. Interestingly, protomer 1 shows a decrease in RMSF in the top section of the α2 helix, which is also highlighted in the HDX-MS experiments (**Figure 4.9c**). The αβ-barrel and MP domains exhibit extensive reduction in RMSF in both protomers. Overall, the MD simulations complement the HDX-MS data, and reveal NSC 60339 is functioning in a similar way in both $AcrA^{SD}$ and $AcrA^{S}$.

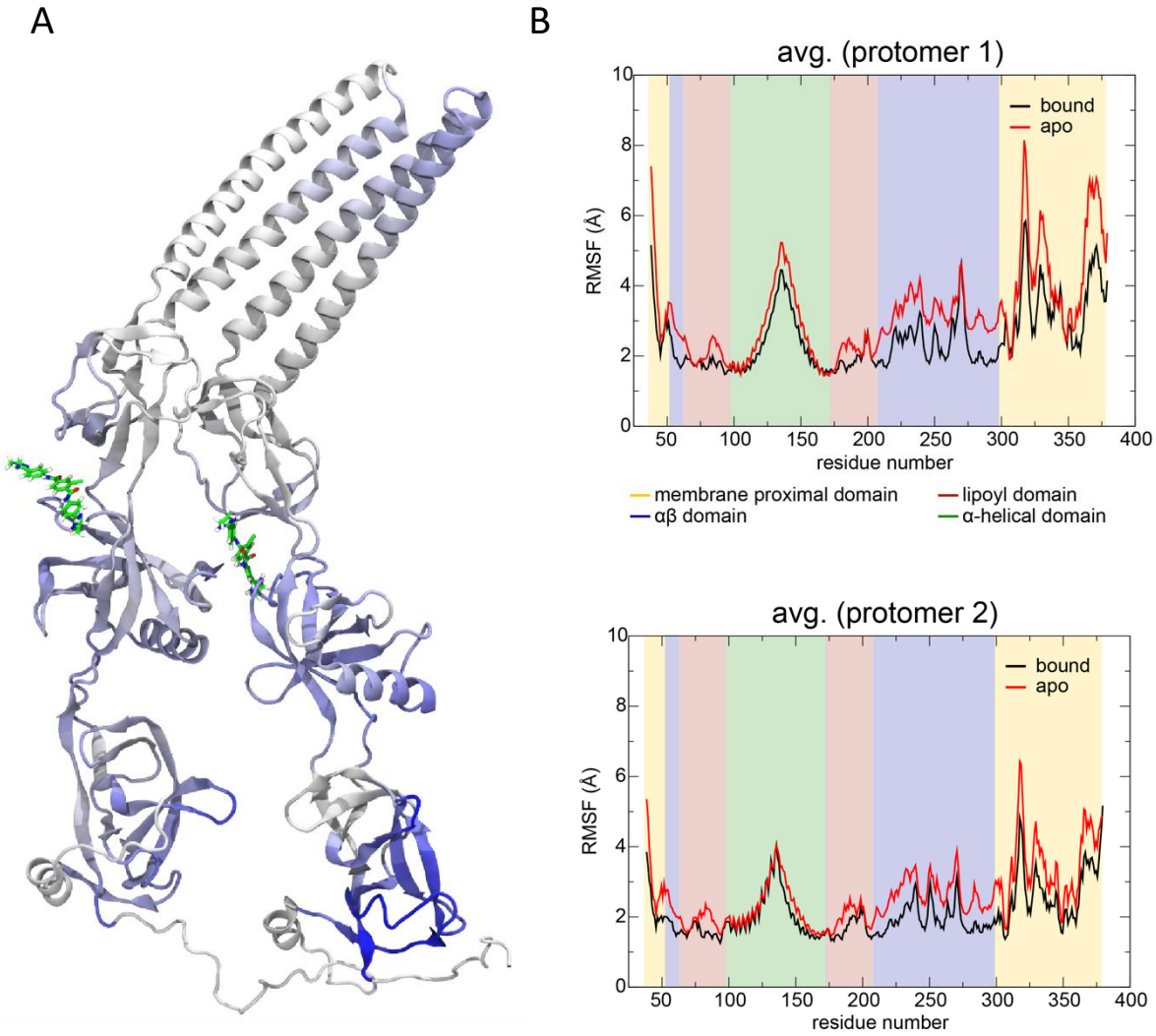


Figure 4.10. MD simulations of NSC 60339 bound to AcrA^{SD}. A. Simulated AcrA^{SD} coloured according to the difference in RMSF over the last 70 ns of 100-ns simulations of the bound and apo states, averaged over four replicas for each. Blue indicates that the RMSF is greater in the apo state than the bound state while white indicates they are similar in both states (the colour range is from -2 Å, blue, to 0 Å, white).**B.** RMSF of AcrA^{SD} from four independent simulations for the bound and apo states. RMSF was calculated over the last 70 ns of each 100-ns simulation. Average from the four replicas for the bound (black) and apo (red) states. The shading indicates the domains of AcrA as indicated. The first plot is for protomer 1, and the second plot is for protomer 2**.**

4.2.4.3 Novobiocin control

As with $AcrA^{S}$, HDX-MS was repeated with $AcrA^{SD}$ and novobiocin, ensuring the same fraction of bound protein and 5% DMSO, at pH 6.0. **Figure 4.11** shows the ΔHDX on $AcrA^{S}$ + novobiocin and $AcrA^{S}$ alone. As with the previous experiments in this chapter, the ΔHDX cut-off was decided based on the CI calculation described in **section 2.4.5.2.1.** Therefore, significance was deemed to be ≥ 0.38 Da change and with a P-value ≤ 0.01 in a Welch's *t*-test.

Again, as expected, novobiocin was observed to have no impact on $AcrA^{SD}$ structural dynamics, as previously seen for $AcrA^{S}$. Three peptides are shown in **Figure 4.11** that exhibited significant protection in the presence of NSC 60339 but see no change in the presence of novobiocin. This provides further confidence in the NSC 60339 results.

As with $AcrA^{S}$, native MS was deployed to monitor novobiocin binding to $AcrA^{SD}$ (**Figure 4.12, Table 4.2**) and the protein-ligand complex was once again observed. The measured mass of $AcrA^{SD}$ was 81,000 ± 3 Da and novobiocin has a mass of 613 Da. The lower ion charge CSD saw satellite peaks, reflecting measured mass increases of 613, 1228 and 1954 Da representing 1-3 novobiocin molecules bound respectively (± 10, 4, 2 Da error).

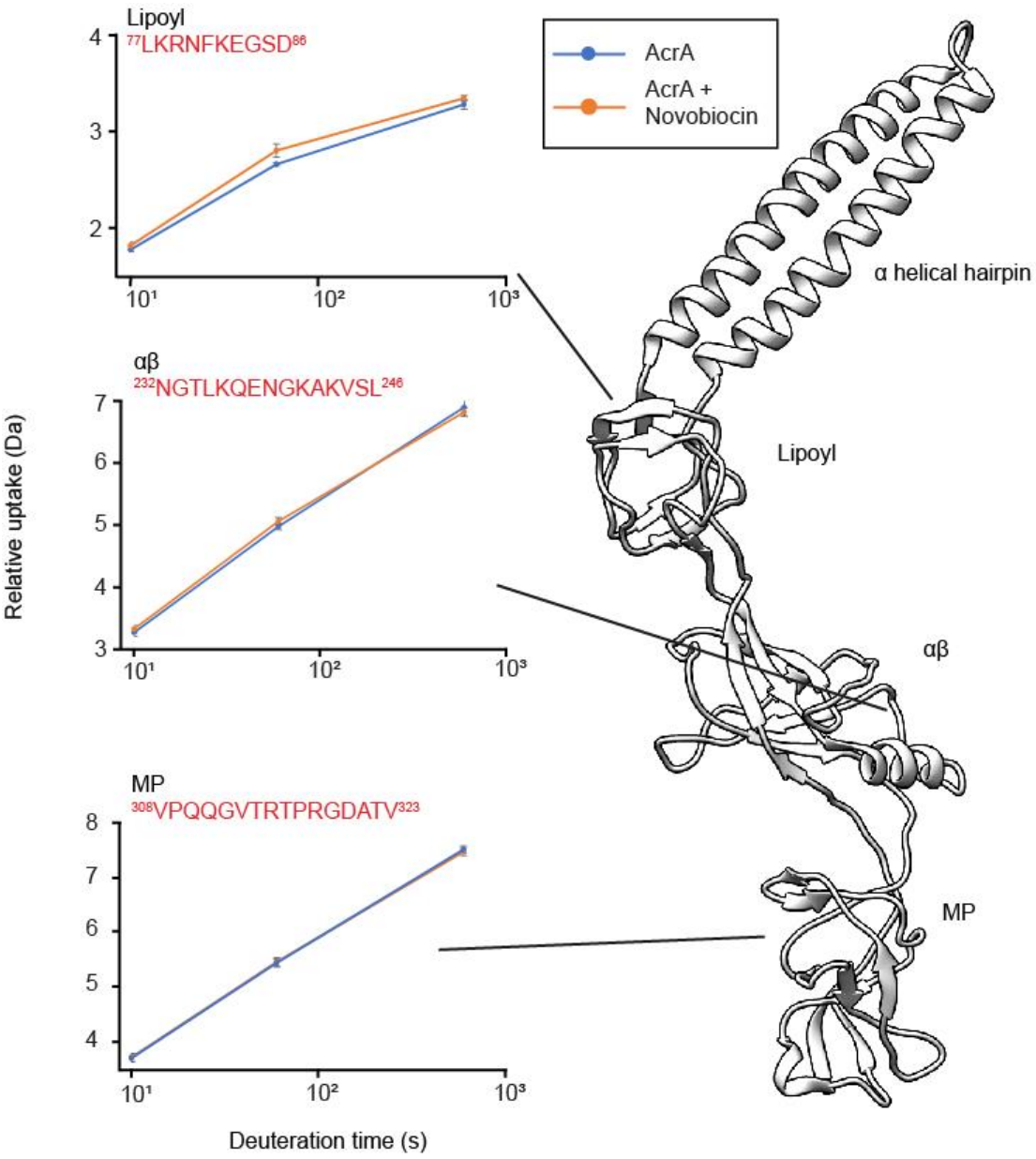


Figure 4.11. The effect of novobiocin on AcrA^{SD} structural dynamics. A Differential HDX (ΔHDX) plots for ((AcrA^{SD} + novobiocin) – AcrA^{SD}) for the latest time point is painted onto the AcrA structure (PDB:5O66) using HDeXplosion and Chimera.[447,448] Significance was defined to be ≥ 0.38 Da change (see **section 2.4.5.2.1**) with a *P*-value ≤ 0.01 in a Welch's *t*-test (n = 4 technical replicates). White areas represent regions with insignificant ΔHDX. Three peptide uptake plots can be shown, in areas that saw significant protection with NSC 60339. Uptake plots are the average deuterium uptake and error bars indicate the standard deviation.

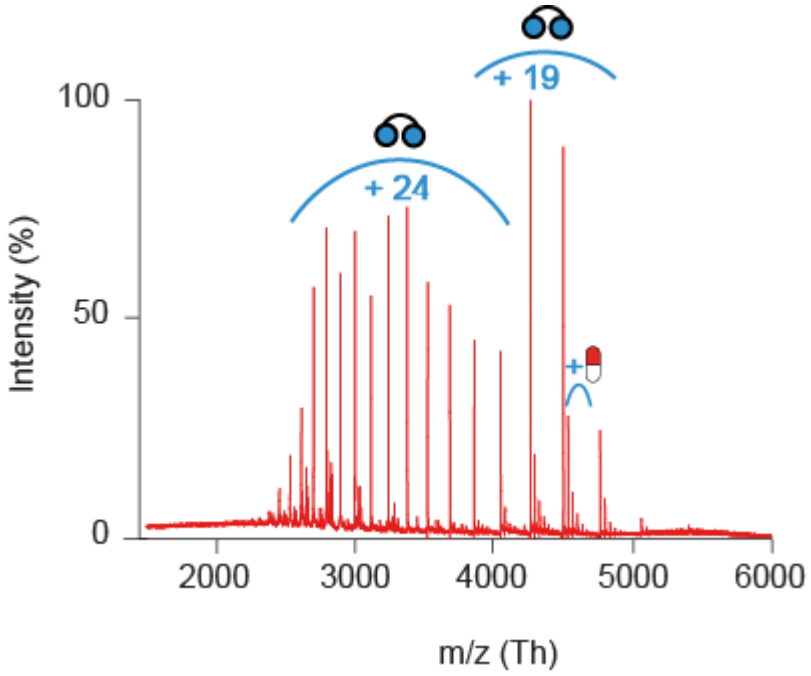


Figure 4.12. Native MS of AcrA^{SD} and novobiocin at pH 6.0. Native MS characterisation of AcrA^{SD} construct with novobiocin at pH 6.0. Protein was buffer exchanged to 100 mM ammonium acetate prior to MS and diluted to 10 µM. Novobiocin added to a concentration of 100 µM, 10% DMSO final. Satellite peaks representing drug binding can be seen adjacent to peaks in the lower ion charge, charge state distribution (CSD). See **Table 4.2** for masses.

Table 4.2. Native MS masses 2. Reported is the standard error of the mean within a single spectrum. Positive mass differences can be attributed to salt and/or detergent adducts. Novobiocin mass = 613 Da

	Measured mass (Da)	**Standard Error (± Da)**	**Mass difference (Da)**
***AcrA^{SD} pH 6.0**	81,000	3	-
+ Novobiocin pH 6.0	81,613	10	613
+ Novobiocin pH 6.0 (x2)	82,228	4	1228
+ Novobiocin pH 6.0 (x3)	82,954	2	1954

4.2.5 Targeting AcrA flexible linkers

The work in **section 4.2.5** was completed by Muhammad R. Uddin and Helen I. Zgurskaya Gumbart as part of a collaborative project.[425]

To determine whether targeting the flexible hinge regions of AcrA could lead to inhibition of antibiotic efflux by AcrAB-TolC *in vivo*, a covalent inhibition assay using live *E. coli* cells was developed. For this purpose, the plasmid borne *acrAB* was constructed with *acrA* variants containing unique Cys substitutions in the flexible linkers between the MP and αβ-barrel domains (Leu50Cys, Ile52Cys, Arg225Cys, Glu229Cys, Asn232Cys) and between the αβ-barrel and lipoyl domains (Arg183Cys, Thr205Cys, Asp284Cys) (**Figure 4.13c**). Cells producing the WT and AcrA(Cys)B variants were treated with a Cys-reactive probe MTS-rhodamine 6G (MTS-R6G). This probe is a substrate of AcrAB-TolC and its covalent binding to Cys residues of AcrA located in positions important for conformational flexibility and/or function of the protein is expected to inhibit efflux activity of AcrAB-TolC pump. In this assay, hyperporinated *E. coli* Δ9-Pore cells lacking all nine TolC-dependent transporters and carrying the plasmid borne AcrA(Cys)AcrB variants were used.[494] Hyperporination of the outer membrane has eliminated the permeability barrier of the outer membrane and facilitated MTS-R6G penetration into the periplasm to reach AcrA(Cys). Cells were pre-treated with MTS-R6G, the unreacted probe washed away and the kinetics of accumulation of a fluorescent substrate Hoechst 33342 (Hoechst) was analysed (**Figure 4.13a**). Kinetic data were fitted into a burst-single exponential decay function and the calculated initial rates of Hoechst accumulation (mM/s) are plotted as a function of the externally added concentration of Hoechst.[495] Therefore, less initial accumulation represents more efflux.

It was found that the MTS-R6G pre-treatment of cells producing wildtype AcrAB-TolC or carrying an empty vector did not affect the Hoechst accumulation, suggesting that the two intrinsic Cys residues of AcrB located in the transmembrane (TM) domain of the protein do not affect the outcome of the assay (**Figure 4.13b**). Then, the comparison of Hoechst accumulation levels in the cells producing different AcrA(Cys) variants and with and without pre-treatment with MTS-R6G was performed. All pumps assembled with AcrA(Cys) variants showed similar accumulation levels of Hoechst as the wildtype pump, suggesting they were as efficient in efflux, which is consistent with their ability to protect Δ9-Pore cells from novobiocin, erythromycin and SDS. The exception is AcrA(Leu50Cys) which appeared to be fully functional in MIC measurement but was only partially efficient in efflux of Hoechst (**Figure 4.13**).

The pre-treatment with MTS-R6G inhibited the activity of AcrA Leu50Cys, Thr205Cys and Asn232Cys variants only, whereas no significant inhibition was seen for other analysed AcrA(Cys) variants (**Figure 4.13b**). Hence, only these Cys residues of AcrA are vulnerable to MTS-R6G binding and inhibition,

whereas other Cys residues are either inaccessible in the AcrAB-TolC complex or binding of MTS-R6G in these sites does not affect the function of the complex. None of these residues are implicated in the characterised AcrA:AcrB binding sites, so the inhibitory effect on efflux is not likely due to perturbed interactions in the assembly of AcrAB-TolC.[389] One possibility for this inhibitory effect is that MTS-R6G binding causes the reduction in AcrA conformational flexibility, but definite conclusions on restricted dynamics can't be drawn here. However, the results show that these three sites within the flexible hinges between AcrA domains are critical for the efflux function of the AcrAB-TolC pump.

T205 is in site IV, the proposed binding site of NSC 60339, and further evidences this area as a druggable site for AcrA inhibition. Interestingly, two other residues were mutated in site IV, yet had no effect on efflux (R183 and D284). These residues reside on the periphery of site IV (**Figure 4.13c**) and are likely less critical for AcrA function than T205, which in the core of the cleft. Furthermore, T205 showed significant protection in the presence of NSC 60339 in both AcrA constructs, highlighting its importance for AcrA flexibility. The two other residues that had an inhibitory effect on efflux are located away from site IV, at the flexible linker between the αβ-barrel and MP domains. N232C and L50C are both located at the bottom of this region and orientated in the same direction (**Figure 4.13c**). The other residues that had no effect on efflux were positioned above N232 and L50 and orientated in the opposite direction. Interestingly, both positions were also found in significantly protected peptides in the presence of NSC 60339 across our HDX-MS experiments. This finding suggests NSC 60339 is stabilising this area downstream from its binding site. Therefore, this provides a novel druggable site on AcrA for the future design of EPIs.

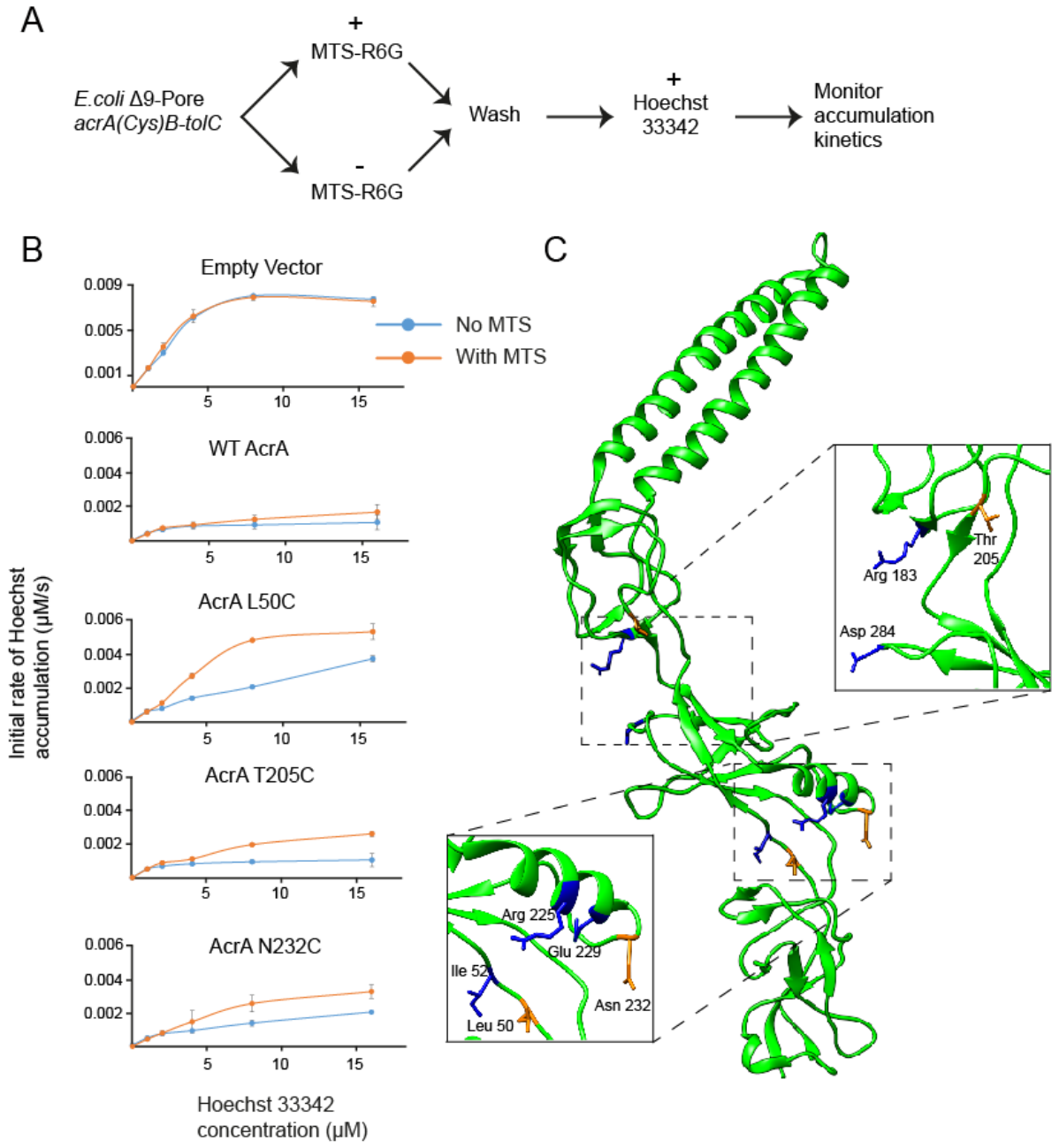


Figure 4.13. The effect of Cys-reactive MTS probe on the efficiency of AcrAB-TolC. A. Covalent inhibition assay experimental process. MTS-R6G is the Cys reactive probe and Hoechst 33342 is the fluorescent efflux substrate. **B.** *E. coli* D9-Pore cells producing AcrAB-TolC complex carrying the indicated AcrA variants were split into two aliquots and one of the aliquots was treated with a Cys-reactive probe MTS. After incubation for 15 min at 37 ˚C, cells were washed and the intracellular accumulation of Hoechst was analysed as described previously.[495] Kinetic data were fitted into a burst-single exponential decay function and the calculated initial rates of Hoechst accumulation (mM/s) are plotted as a function of the externally added concentration of Hoechst.[495] Error bars indicate standard deviation (n = 3). **C.** AcrA structure (PDB 5O66) with mutated residues highlighted. Blue residues have no effect on efflux and orange residues had an effect.

4.2.6 Proposed mechanism of NSC 60339 inhibition

The findings presented in this chapter support the previous data that NSC 60339 binds in a cleft between the lipoyl and αβ-barrel domains, termed site IV.[403,489] Furthermore, whilst previous work suggested NSC 60339 affects the conformation of AcrA, the HDX data reports NSC 60339 causes long-range restriction in backbone dynamics across all four of its domains.[403,489] This behaviour was reported for the monomer and pseudo-dimer constructs. Therefore, combining the previous results of NSC 60339 binding, the results presented in this chapter and the current literature detailing the function of AcrAB-TolC, it is possible to propose a mechanism of inhibition (**Figure 4.14**).

It can be proposed that NSC 60339 acts as a molecular wedge within the site IV cleft, significantly restricting the structural dynamics of AcrA, which could have implications for the conformational transitions required during the functional rotation of the AcrAB-TolC pump. A recent study of the assembled AcrAB-TolC multidrug efflux pump *in situ* confirms that one of the AcrA protomers in the AcrA dimer is anchored to the inner membrane and interacts with AcrB, whilst also stretching to interact with the peptidoglycan layer and TolC.[428] The N-terminal region of the second protomer is suggested to interact with the AcrB PC2 subdomain, which undergoes extensive conformational changes throughout the rotational mechanism.[427] This architecture could allow AcrA to communicate conformational changes in AcrB, to TolC, as it cycles through the rotational mechanism, resulting in TolC transitioning to the 'open' state. Thus, NSC 60339 inhibition may result in AcrA losing its ability to accommodate changes across the periplasm and communicate the conformational changes of AcrB to TolC, required for efflux. This could affect the pump in several ways; 1) The interaction of AcrA with AcrB and TolC may become disrupted as AcrA becomes more rigid, 2) AcrA may not be able to maintain a sealed channel during functional rotation, and 3) TolC may not be 'opened' as efficiently.

4.2.7 NSC 60339 and necrosignalling

AcrA has recently been identified as a bacterial necrosignal within *E. coli* swarms.[486] This functions by dead cells releasing AcrA which binds TolC on the outside of other live cells, stimulating efflux within the affected area and promoting the upregulation of various other efflux pumps.[487] Further work by Bhattacharyya *et al.* (2022) to characterise necrosignalling by AcrA, found NSC 60339 inhibited both efflux and necrosignalling.[487] Thus, in the case of *E. coli*, inhibition by NSC 60339 appears to be doubly as effective, inhibiting both efflux and necrosignalling. Whilst necrosignalling has been seen across a wide variety of both Gram-positive and Gram-negative bacteria, the nature of other necrosignals is unknown. Therefore, it is unknown if necrosignals mostly tend to be components of efflux systems.

Interestingly, novobiocin was observed to inhibit necrosignalling ability as well, but to a lesser degree than NSC 60339. This suggests one of two possibilities. Firstly, that restricting structural dynamics assists necrosignalling inhibition. The HDX-MS data in this chapter revealed novobiocin has no significant effect on AcrA structural dynamics. Therefore, the reason NSC 60339 may be more effective at inhibiting necrosignalling is due to the restriction of structural dynamics caused by NSC 60339 binding. It has been proposed that AcrA contacts TolC on the outside of live cells via the α-helical hairpins of AcrA contacting the TolC pore in the outer membrane. Therefore, a restriction in structural dynamics by NSC 60339, particularly in the lipoyl and α-helical domains as shown by the HDX data, may affect the ability of AcrA to freely position the α-helices for optimum binding to TolC. It may be for this reason that NSC 60339 inhibits necrosignalling more effectively than novobiocin.

The second possibility is that restricting structural dynamics is not necessary for the inhibition of necrosignalling. Since novobiocin inhibits necrosignalling to some degree and does not cause a change to AcrA structural dynamics as shown by HDX-MS, it is possible novobiocin is inhibiting necrosignalling in a different fashion, or that restricting structural dynamics is not the important feature of this mechanism. It is possible other effects such as steric hinderance, or the blocking of certain residues, play a more important role. The necrosignalling function of AcrA has only recently been discovered, and therefore further investigations into the mechanisms of necrosignalling are required to understand this ability, and how best to inhibit it to combat bacterial multidrug resistance that occurs via this process. Mapping the interaction between AcrA and TolC on the outside of the outer membrane, and investigations into whether other MFPs in homologous systems across different bacterial species can perform necrosignalling would help to shed further light on the mechanisms of necrosignalling.

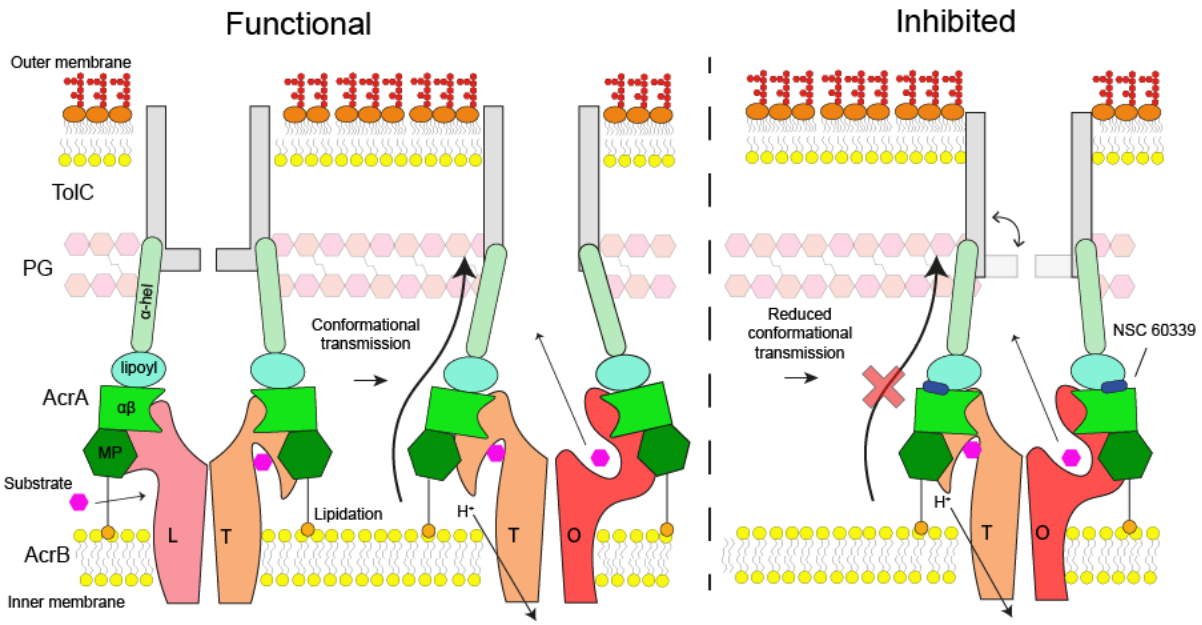


Figure 4.14. Schematic of AcrAB-TolC inhibition by NSC 60339. Under normal conditions, as AcrB cycles through its three states of its rotational mechanism (L = Loose, T = Tight, O = Open) the conformational transition information is transmitted through AcrA to TolC, ensuring it is in the 'open' state for efflux. AcrA becomes conformationally restricted once NSC 60339 'wedges' between its lipoyl and αβ-barrel domains, reducing its ability to transmit the conformational movements of AcrB and, subsequently, inhibiting functional rotation and efflux.

4.3 Conclusions

This chapter reports the first mechanism of action for an AcrA inhibitor. Through a combination of HDX-MS investigations, MD simulations and previous work, it was found that NSC 60339 acts as a molecular wedge at a cleft between the lipoyl and αβ-barrel domains of AcrA, causing a restriction in conformational dynamics.[403,489] This is suggested to have implications for the conformational transitions required by AcrA during the functional rotation of the AcrAB-TolC multidrug efflux pump, which could result in AcrA losing its ability to accommodate changes across the periplasm and communicating AcrB's conformational changes to TolC, and thus affecting the efficiency of efflux.[428] Furthermore, cellular accumulation assays revealed a second potential druggable site on AcrA, on the flexible linkers between the αβ-barrel and MP domains.

MFPs share a several common characteristics; they have an elongated structure with 3-4 linearly arranged domains that are connected by flexible linkers.[410] It is well documented that MFPs are highly flexible and must be able to adopt different conformations, in order to effectively interact with an outer membrane and inner membrane component, whilst also accommodating conformational changes throughout the efflux cycle.[398,410,428] This work provides a novel model in targeting the AcrAB-TolC multidrug efflux pump, providing a platform to understand and develop the next generation of EPIs. Targeting the flexible linkers between the domains of MFPs may be an effective strategy to restrict the conformational dynamics of these proteins and enable the design of inhibitors against homologous RND efflux pumps across other ESAKPE bacteria. More generally, this work highlights how HDX-MS can be used in combination with MD simulations, high-resolution structural information, and microbiology assays to understand these dynamics and to determine the molecular mechanisms of inhibition, as previously seen with AcrB.[35]

4.4 Material and methods

For methods described in previous chapters, the reader will be redirected.

4.4.1 Reagents

All reagents purchased from ThermoFischer Scientific or Sigma Aldrich/Merck unless otherwise stated. The highest quality reagents were always prioritised.

4.4.2 Protein expression and purification

AcrA^{S} and AcrA^{SD} were expressed and purified as described in **chapter 2** and **chapter 3** respectively.

4.4.3 Polyacrylamide gel electrophoresis (PAGE)

Sodium dodecylsulfate (SDS)-PAGE and SMA-PAGE were completed as previously described.

4.4.4 Mass spectrometry

4.4.4.1 Native mass spectrometry

Native MS data in this chapter was collected in collaboration with Anna J. Higgins and Frank Sobott as part of a collaborative project.[425]

AcrA$^{S/SD}$ were measured using The Q-Exactive Plus ultra-high mass resolution (UHMR) (ThermoFisher Scientific) with novobiocin. The drug was added to the protein and left to incubate for 30 mins prior to MS analysis. Samples acquired with and without DMSO. The UHMR settings used were: 1.5 kV spray voltage, capillary temperature 60 ˚C, ion source temperature off, extended trapping 60.

4.4.4.2 Hydrogen deuterium exchange mass spectrometry

HDX-MS spectrometry experiments ran and analysed as previously described. The differences in sample preparation are described below.

4.4.4.2.1 Preparation of ligands for hydrogen deuterium exchange mass spectrometry

NSC 60339 and novobiocin was purchased from MedChemExpress and Caymen Chemical respectively. Stock solutions of NSC 60339 (10 mM) and novobiocin (600 µM) were made in 100% DMSO and sonicated for 2-3 hours to ensure solubility. Obtaining a maximal percentage of protein:drug complex during deuterium labelling is an important consideration for HDX-MS experiments. NSC 60339 and novobiocin both bind to AcrA^{S} with µM affinity (NSC 60339 with a K_D of 78 µM and novobiocin with a

K_D of 4.3 µM, as measured by previous SPR measurements at pH 6.0).[403] To ensure maximum protein:drug complex was present in our labelling conditions, the following equation was used:[496]

$$\text{fraction of bound protein} = \frac{(L_T + P_T + K_D) - \sqrt{(L_T + P_T + K_D)^2 - 4L_T P_T}}{2P_T}$$

Equation 4.1. Estimating the percentage of protein:ligand complex. L_T is the total ligand concentration and P_T is the protein concentration, K_D is the equilibrium dissociation constant.

Protein was incubated with 500 µM NSC 60339 or 30 µM of novobiocin before being diluted into labelling buffer containing the same concentration of drug. 5% DMSO was kept consistent throughout experimentation to ensure drug solubility.

4.4.4.2.2 Statistics and reproducibility

In this chapter, several of the experiments contained biological replicates so the cut-off significance threshold between two states was established based on the calculated CI value for all experiments.[462,463] This was calculated as described in **section 2.4.5.2.1.**

4.4.5 Molecular dynamics simulations

All MD simulations was completed by Katie M. Kuo, Jerry M. Parks, and James C. Gumbart as part of a collaborative project.[425]

MD simulations of the NSC 60339-bound monomer were initialised using the docked structure from Darzynkiewicz *et al.* (2019)[489] while those of the apo used the same structure with the compound removed. Each of the two systems was solvated in an ~(170-Å)^3 cubic water box in order to allow for tumbling of the protein without using orientational restraints and ionised with 150 mM NaCl, resulting in a system size of ~476000 atoms. The dimer system was constructed starting from the AcrA dimer structure in PDB 2F1M.[397] The MP domain was added based on the monomer structure, and residues in the lipoyl and αβ-barrel domains were adjusted to match their positions in the NSC 60339-bound structure, after which the compound was modelled in to both copies of AcrA. Apo and bound systems were solvated in an ~(210-Å)^3 cubic water box and ionised with 150 mM NaCl, resulting in a system size of ~906000 atoms.

Each system was equilibrated for 0.5 ns with all protein and ligand atoms restrained, followed by 4.5 ns with only the protein backbone and ligand atoms restrained. Then, each system was equilibrated for 100 ns in quadruplet. Simulations were run using either NAMD2.14[497] or NAMD3[498] depending on

the computational resource used, and the CHARMM36m force field.[499] Force-field parameters for NSC 60339 were generated using the CGenFF webserver[500]. A time step of 2 fs was used, with short-range non-bonded interactions (cut off of 12 Å with a switching function starting at 10 Å) updated every time step and long-range electrostatics updated every other time step using the particle mesh Ewald method.[501] A constant temperature of 310 K was maintained using Langevin dynamics, while a constant pressure of 1 atm was maintained using a Langevin piston. All results presented were averaged over four replicas.

All system preparation and analysis were carried out using VMD. RMSF and SASA were measured using VMD's "measure rmsf" and "measure sasa" functions. In the case of RMSF, the fluctuations are measured using the average position over the sampled frames as a reference.

4.4.6 Surface plasmon resonance

$AcrA^{S}$ and $AcrA^{SD}$ protein were immobilised using the amine coupling method. For this purpose, CM5 chip (BiaCore) surfaces were activated with 0.05 M *N*-hydroxysuccinimide and 0.2 M *N*-ethyl-*N*- (3 -diethylaminopropyl) carbodiimide. $AcrA^{S}$ and $AcrA^{SD}$ were injected over surfaces immediately after activation. After immobilisation, the excess of reactive groups was blocked by injecting 0.5 M ethanolamine HCl (pH 8.0). The immobilisation and subsequent binding experiments were conducted in running buffer containing 20 mM HEPES-KOH (pH 7.0), 150 mM NaCl, and 0.03% n-Dodecyl-β-D-maltopyranoside (DDM) supplemented with 5% DMSO. The CM5 chip contains four chambers, whereas the first (control surface) was activated and processed in the same way but the protein was omitted during the immobilisation step. The second and third chambers contained the immobilised $AcrA^{S}$ and $AcrA^{SD}$ (ligand). The immobilised densities of both proteins (ligand) were 4743 and 4054 RU, respectively. The sensor-grams were collected and analysed as described before.[411,502]

4.4.7 Covalent inhibition experiments

Covalent inhibition experiments were completed by Muhammad R. Uddin and Helen I. Zgurskaya Gumbart as part of a collaborative project.[425]

4.4.7.1 Site-directed mutagenesis

All amino acid substitutions in *acrA* were introduced by QuickChange II XL Site-Directed Mutagenesis Kit using p151*acrAB*His as the template.[503] Primer design and polymerase chain reaction (PCR) reaction for each substitution were performed by following manufacturer's protocol.

E. coli Δ9-Pore strain (*acrB* Δ*acrD* Δ*acrEF*::*spc* Δ*emrB* Δ*emrY* Δ*entS*::*cam* Δ*macB* Δ*mdtC* Δ*mdtF* *att*Tn*7*::mini-Tn*7*T Tp^r *araC* P_{araBAD} *fhuA*Δ*C/*Δ*4L*) was constructed previously.[494] This strain was used in antibiotic susceptibility and covalent inhibition studies.

4.4.7.2 Minimal inhibitory concentrations

Susceptibilities of the *E. coli* Δ9-Pore cells containing plasmid-borne AcrA(Cys)AcrB variants against SDS, Novobiocin, Erythromycin, and Vancomycin, were determined by two-fold broth microdilution. Briefly, overnight cultures were sub-cultured in Luria−Bertani (LB) broth (tryptone, 10 g/L; yeast extract, 5 g/L; NaCl, 5 g/ L), and cells were grown at 37 °C in a shaker at 225 revolutions per minute (RPM) until OD_{600} reached 0.2-0.3. For the proper expression of the pore, L-arabinose (final concentration of 0.1%) was added to each culture, and cells were further grown until OD_{600} reached 1.0. The minimum inhibitory concentration of each bacterial strain against different antibiotics was measured in 96-well plates. Exponentially growing cells were added to each well and incubated for 18 hrs. Plates were scanned to determine the final OD_{600} by Spark 10 M microplate reader (TECAN). Minimal inhibitory concentrations were taken as the lowest concentration of drug that prevented bacterial growth as judged by the OD_{600} measurement.

4.4.7.3 Covalent inhibition experiments

Overnight culture of *E. coli* Δ9-Pore cells carrying p151-AcrABHis plasmid with the indicated substitutions in AcrA were sub-cultured with 0.1% arabinose for 6 hours to achieve OD_{600} 1.0. For covalent inhibition of AcrA, cells were split into two aliquots and one aliquot was incubated with 20 μM MTS-rhodamine 6G (MTS-R6G) for 15 min at 37°C. The comparator cell aliquot was incubated with a blank solvent (DMSO) under the same conditions. After incubation, cells were washed twice with HMG buffer (50 mM HEPES-KOH buffer pH 7.0, 1 mM magnesium sulphate and 0.4 mM glucose) and resuspended in HMG buffer to an OD_{600} of ~1.0, at room temperature. Different concentrations of Hoechst 33342 were tested to measure the substrate efflux efficiency of each mutant cells by measuring kinetics of Hoechst accumulation at λex = 350 nm and λem = 450 nm. Hoechst is a water soluble membrane permeable compound that fluoresces when bound to the minor groove of DNA or in a hydrophobic environment such as a lipid membrane.[504] The initial burst rate of Hoechst accumulation was used to understand efflux by understanding the accumulation of dye at the periplasm and cytosol. This was done for different starting concentrations of Hoechst ± MTS for wild type AcrAB-TolC and empty vector originally before the AcrA(Cys) mutants (**Figure 4.13**). Less initial accumulation suggests more efflux, and this model has previously shown efflux is at least 420 times faster than diffusion into the cell.[495]

Data were normalized and kinetic parameters were calculated as described previously by Westfall *et al*. (2017) using a MATLAB program.[495] Briefly, the time courses of Hoechst uptake were fit to the burst-single exponential decay function $F = A_1+A_2\cdot(1-\exp(-kt))$, where A_1 and A_2 describe the magnitude of the fast and slow steps, respectively, and k is the rate of the slow step. The fast and slow steps were attributed to Hoechst binding to the lipids (i.e., in the periplasm) and chromosomal DNA (i.e., in the cytoplasm), respectively. The initial rates for Hoechst accumulation in the cytoplasm were calculated as $V_I = A_2\cdot k$ and plotted in **Figure 4.13**. Representative time courses of Hoechst accumulation in *E. coli* Δ9-Pore cells with wild type AcrA and empty vector is shown in **Appendix 8**.

Chapter 5: The study of AcrAB-TolC in increasingly complex lipid environments

5.1 Introduction

In situ, integral membrane proteins are embedded within a biological membrane, due to their hydrophobic α-helical and/or β-barrel regions.[225] In the middle of the bilayer is a hydrophobic core, which shields the hydrophobic protein regions from the aqueous phase (see **section 1.7.1**). This adds complexity to their study *in vitro*; for reliable structure-function elucidation, membrane proteins need to be purified in a mimetic hydrophobic environment to ensure their solubilisation.[505] Traditionally, the most common strategy for this was the use of detergents.[506] Detergents are able to extract a membrane protein from the native membrane, forming a spherical micelle containing the protein and detergent molecules. Whilst the use of detergents has been critical in advancing the understanding of membrane protein biology, their use has several drawbacks.[112] One key issue with detergents is they strip the protein of its native lipid environment, which can affect the structure/function of the membrane protein as lipids can be essential for a proteins function.[100,338] This can lead to proteins adopting non-physiological conformations in micelles or having compromised function. The other key issue is the difference in physiochemical properties of a detergent micelle and a phospholipid bilayer.[112] Micelles have a single hydrophilic surface that is curved, and a low order hydrophobic core. Furthermore, detergent molecules exist in monomer-micelle equilibrium, causing detergent monomers to exchange micellar and soluble conformations. This unfavourably increases the dynamics of the membrane proteins environment. Lateral pressures and water permeability are also vastly different in detergent micelles compared to native membranes.[507] These physiochemical differences lead to a general reduction of membrane protein stability in micelles, and they are prone to aggregation.

Alternatives to detergent systems exist to provide a more stabilising environment. Amphipols are a class of amphipathic polymer comprising a polyacrylate backbone with pendant hydrophobic and hydrophilic sidechains.[508] Membrane protein-amphipol complexes are more stable than detergent micelles, due to the low exchange rate between protein-bound amphipols and free monomers in solution.[112] Another alternative are the use of liposomes; these are self-closed phospholipid bilayers, and can be used to reconstitute membrane proteins into a unilamellar lipid vesicle.[18,509] They are relatively easy to construct via extrusion or ultrasonication methods. The most common method of reconstitution involves mixing a detergent solubilised membrane protein sample with the phospholipid vesicle of choice, then slowly removing the detergent through dialysis, gel filtration or Bio-Bead adsorption. When the detergent levels reach the critical micelle concentration (CMC), the membrane protein will simultaneously associate with the liposome to form proteoliposomes. If care is taken to control the size distribution of the liposomes and the stoichiometry of proteins-to-liposomes, this method can yield success for some systems.[510] However, it is hard to control the

orientation of the protein within the liposome, and the amount of reconstituted protein can be limited.[509] Furthermore liposomes can be prone to aggregation due to the high curvature of the membrane, which can lead to cracks in the hydrophilic outward facing surface during storage.[511] This can also be unfavourable to reconstituted integral membrane proteins.

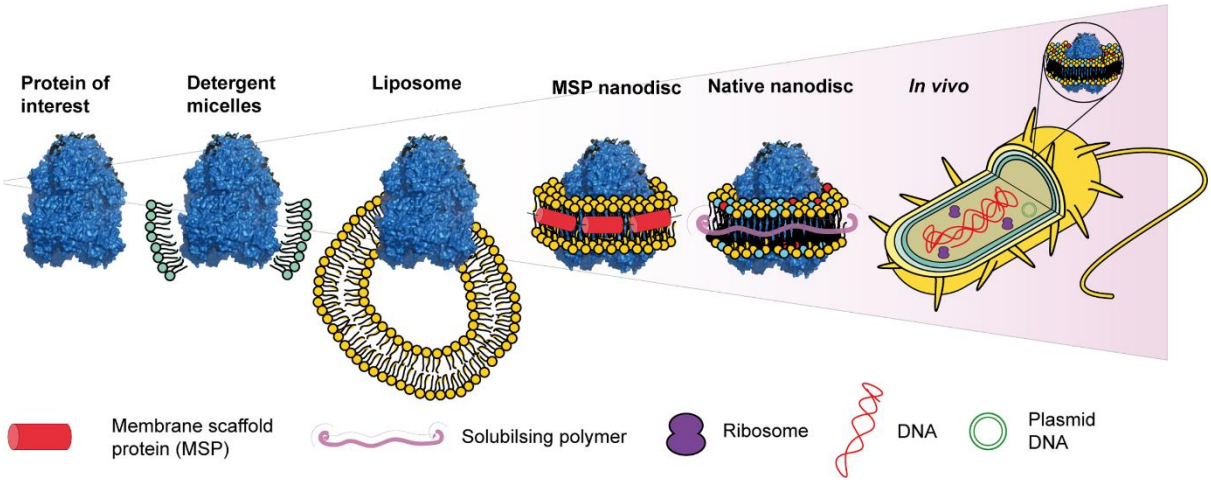


Figure 5.1. The increasing complexity of studying membrane proteins. Schematic showing the different environments used to study membrane proteins, with increasing complexity and representability. PDB 5O66. Taken from Russell Lewis *et al*. (2023).[32]

The term 'nanodisc' was invented by Silgar and co-workers in 2002, to represent a new method to incorporate membrane proteins in a phospholipid bilayer.[512] It involves transferring a detergent solubilised membrane protein into lipid nanodiscs, which in turn is bound by membrane scaffold proteins (MSP). MSP is an amphipathic helical protein engineered from human apolipoprotein A-1 that shield the hydrophobic core of lipids from the aqueous phase.[112] MSP nanodiscs have been found to be generally applicable to most types of membrane protein to form constructs with fairly high stability. MSP nanodiscs confer many advantages over detergents and liposomes; the ability to use different genetically engineered variants of MSP, and to modify the protein-lipid ratio in the final nanodisc, means the final size of the nanodisc can be somewhat controlled.[513,514] The size of an MSP nanodisc can range from ~6 nm to ~17 nm. Furthermore, the MSPs can be functionalised, or affinity tags/labels can be introduced. The lipid composition can be controlled as with other bilayer systems, and membrane proteins can even be introduced to an exclusively native lipid environment from detergent solubilised membranes.[515] However, whilst MSP nanodiscs provided a more stable environment for

studying membrane proteins than detergents or liposomes, they still require detergents to extract native membrane proteins from cellular membranes, which can be highly destabilising.

A novel approach to studying membrane proteins is the use of styrene maleic acid (SMA) copolymers.[114] SMA is able to insert into biological membranes to directly extract membrane proteins and form small discs of native bilayer surrounded by the SMA polymer – SMA lipid particles (SMALPS).[516] Therefore, SMALPs provide the encapsulated membrane protein an environment containing its native lipids, which is important for protein function.[100,338] It has been shown AcrB is significantly more active in SMALPs than DDM micelles.[474] SMA is amphipathic (**Figure 5.2**) due to its hydrophobic aromatic styrene group and its hydrophilic maleic acid groups. Usually, SMA polymers with either 2:1 and 3:1 styrene:maleic acid ratios are used to solubilise membrane proteins. This provides a huge advantage for studying membrane proteins, as they are always kept in their native lipid environment, even throughout the purification process as SMALPs are amenable to affinity chromatography.[516] The resulting SMALP is a small (~10 nm), stable, soluble molecule that can be used in many downstream techniques for biophysical analysis of membrane proteins.[63,474]

As with all approaches, SMALPs have several disadvantages. Firstly, their small disc size of 10 nm can be too small for large proteins and complexes with significant membrane spanning domains.[516] However, it is possible to tune the size of the SMALP by optimising the polymer:lipid ratio or the styrene:maleic acid ratio.[517,518] Secondly, SMALPs have a significant intolerance to divalent cations such as Mg^{2+} and Ca^{2+}.[516] It is thought the two carboxyl groups of the maleic acid (see blue colour, **Figure 5.2a**) chelate the divalent cations, to induce a conformational strain in the polymer, and if this occurs on too many of the maleic acid groups surrounding a single SMALP, it causes the SMA to precipitate. This results in the SMALP no longer being water soluble and therefore the membrane protein precipitates too. SMALPs with a 2:1 ratio of styrene:maleic acid precipitate with a Mg^{2+} concentration of 4 mM, yet SMALPs with a 3:1 ratio precipitate with a Mg^{2+} concentration of <1mM.[519] However, restricting Mg^{2+} below these concentration levels may not be effective as SMA will still bind Mg^{2+}, making it unavailable to the protein. This is a particular problem for proteins binding divalent cations as cofactors such as ATPases. Thirdly, the use of SMALPs is restricted by an intolerance to acidic pH.[520] At acidic pH, the maleic acid groups become protonated and the polymer becomes insoluble. Therefore, pH values above 7 have to be used for proteins in SMALPs. Lastly, the aromatic styrene group shows a strong UV absorbance at 260 nm, which overlaps with the absorption from many aromatic amino acid residues, complicating spectroscopic techniques to study proteins such as UV-Vis spectroscopy.[521]

As of today, there are many different polymers used to create native nanodiscs, each with their own advantages and disadvantages. Only two main alternatives were used in this chapter, and hence only they will be discussed. For more information on different polymer derivatives, please see this review by Orekhov *et al*. (2022).[522] Di-isobutylene-alt-maleic acid (DIBMA, **Figure 5.2b**) can be used to directly solubilise membrane proteins from their native environment to form DIBMALPs.[521] DIBMA deals with some of the limitations of SMA. Firstly, they lack the aromatic styrene group and thus do not absorb at 260 nm.[521] DIBMALPs form larger discs than SMALPs (10-50 nm), and thus allows for a higher degree of protein conformational movement and the ability to solubilise large proteins and complexes.[523] Solubilisation by DIBMA, which has a lower hydrophobicity than SMA, only leads to mild perturbation of the packing and phase temperatures of the lipids, and the lipid dynamics are less constrained than in SMALPs.[524] Furthermore, DIBMALPs have a lower rate of collisional lipid transfer than SMALPs.[525] DIBMA can also tolerate divalent cations far better than SMA; even though DIBMA has a higher percentage of maleic acid groups than SMA, it has been hypothesised that the larger disc alleviates the strain caused on the polymer by divalent cation chelation.[526] An alternate hypothesis is that the higher percentage of maleic acid groups in DIBMA reduces the hydrophobicity and therefore it can tolerate a higher degree of neutralisation by divalent cations.[527] On the other hand, DIBMA is not as effective at solubilising membrane proteins as SMA, and exhibited a 2.5 fold lower yield of purified protein from *Escherichia coli* (*E. coli*) membranes.[526] Additionally, DIBMALPs often contain more contaminants, which may be attributed to the larger disc size. Furthermore, proteins exhibit a slight decrease in stability over long term storage in DIBMALPs compared to SMALPs, and DIBMALPs still exhibit the same pH restrictions as they still contain the maleic acid groups.[526]

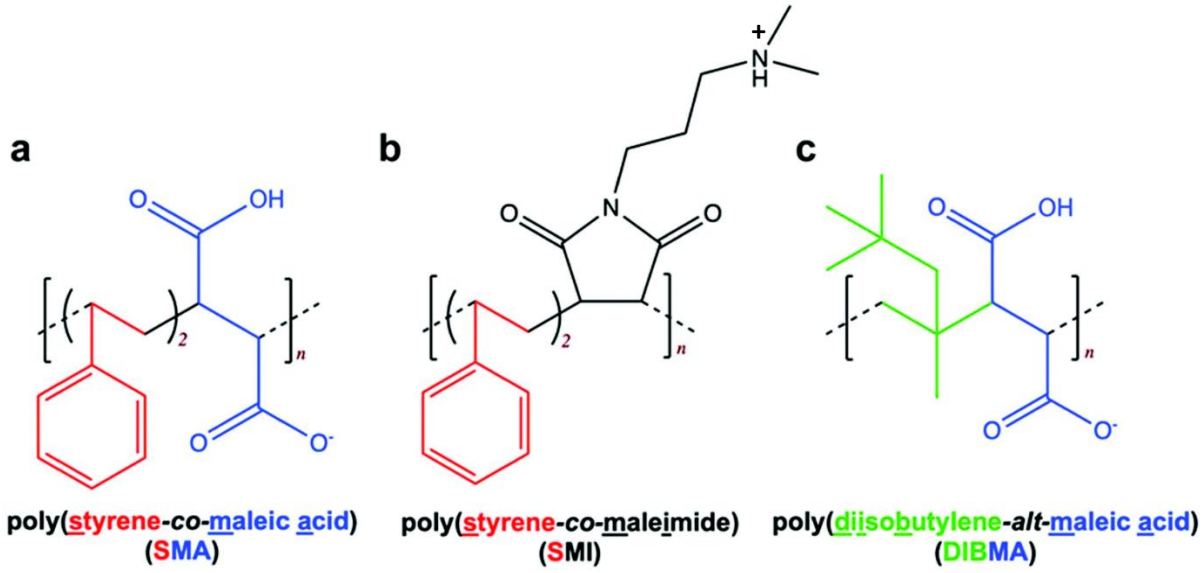


Figure 5.2. Structures of copolymers for membrane protein solubilisation. A. SMA. **B.** SMI. **C.** DIBMA. Taken from Grime *et al.* (2021).[528]

The second polymer alternative used in this chapter is polystyrene-*co*-maleimide (**Figure 5.2c**). It is an amphipathic copolymer of styrene and dimethylaminopropylamine maleimide (2:1 ratio) and self assembles into phospholipid nanodiscs (SMILPs) in the same way as SMA, except under acidic conditions.[529] This addresses the main limitation of SMA, which is the pH restriction, as SMILPs have the opposite pH dependence and are only soluble below pH 7.8; importantly SMILPs do work at physiological pH.[529] SMILPs also have a very high tolerance to divalent cations unlike SMALPs.[530] However, SMI still contains the styrene functional group so will disrupt UV-Vis spectroscopy. Furthermore, at pH 7, SMI is less efficient at solubilising proteins from the membrane as SMA and *n*-Dodecyl-β-D-maltopyranoside (DDM), which may be due to the smaller size of the SMILP nanodisc. Therefore, with the range of different polymers available for studying membrane proteins, it is important to choose the correct polymer depending on the which experiments are required and which protein system is being used.[528]

As the field of studying membrane proteins has advanced over the last 20 years to the use of lipid nanodiscs, techniques such as structural mass spectrometry (MS) have had to develop new methods to analyse these systems. Traditional native MS methods are amenable to membrane proteins in detergents or MSP nanodiscs, but polymer nanodiscs complicates experiments.[106] Due to the heterogeneity of the lipids and polymer in the nanodisc, the membrane protein needs to be ejected from the polymer nanodisc in the gas phase. However, this is hard to do using traditional collisional induced dissociation (CID) due to the stability of the nanodisc. Therefore, novel ionisation methods have been developed, such as laser-induced liquid bead ion desorption (LILBID), to perform native MS with membrane proteins in polymer nanodiscs (for more information, see **sections 1.2.1.2** and **1.3.2**).[63]

Hydrogen deuterium exchange mass spectrometry (HDX-MS) methods have also had to adapt to study proteins in lipid environments. Lipids can cause several problems in bottom-up HDX workflows.[131] The efficiency of protein digestion can be reduced due to interference with the protease, the LC system can be fouled, and possible spectral complication due to peptide-lipid co-elution and peptide ion suppression can occur. Hebling *et al*. (2010) and Reading *et al*. (2017) detailed HDX-MS workflows for membrane proteins in MSP nanodiscs and SMALPs respectively.[188,190] Both of these protocols used ZrO_2 coated beads, as an additional step post deuterium labelling to remove lipids from the protein sample. However, these steps are offline, so are not completed by the robotics systems commonly used in HDX-MS experiments. Therefore, as the beads need to be removed pre injection into the mass spectrometer, the workflows can be laborious and time consuming.[4] In turn, this can effect reproducibility between repeats and can increase back-exchange. Recently, Hammerschmid *et al*. (2023) developed a protocol to enable a completely online workflow, which uses a regenerable

phospholipid trapping ZrO_2 bead column.[4] This exploits the Lewis acid chemistry of ZrO_2 to offer an automated HDX system to analyse membrane proteins in lipid environments, eliminating the need for additional steps such as filtration or bead disposal.

This chapter aimed to investigate AcrAB-TolC in increasingly complex lipid environments. It is important to consider which type of mimetic environment should be used for a particular experiment. There should be a compromise between aiming for the most native membrane environment possible for the protein sample, and the amenability of the mimetic to the particular technique being used. Therefore, the constituent proteins have been purified in a range of different mimetic environments, such as detergents, liposomes, MSP nanodiscs and native nanodiscs, each with varying complexity in their lipid environments, and their own unique characteristics, for downstream experiments.

In this chapter, an HDX investigation into AcrB in MSP nanodiscs was performed with an MBX inhibitor to monitor the effects on AcrB structural dynamics in a lipid environment; this aimed to validate the novel online delipidation HDX-MS workflow detailed by Hammerschmid *et al.* (2023).[4] Furthermore, the optimised SMALP samples were utilised and a novel SMALP-liposome-SMALP assay was applied to a separate investigation with the aim of investigating the effect of anti-microbial peptides on AcrB in a lipid environment. Lastly, various pull-down assays and SMA polyacrylamide gel electrophoresis (SMA-PAGE) protocols were optimised to try and observe complex assembly *in vitro* in native nanodiscs or DDM micelles. Whilst AcrAB-TolC complex assembly has been monitored by other groups (see **section 1.10.8**), often by use of cryo-electron microscopy (cryo-EM), the aim was to develop a quick biochemical assay to probe assembly *in vitro*, and to monitor assembly in the presence of different ligands and conditions such as inhibitors (NSC 60339, MBX-3756), Mg^{2+} concentrations and pH.[38,367,372,428,531]

5.2 Results

5.2.1 Purifying AcrB in different polymer native nanodiscs

With the different polymers available to form native nanodiscs, the aim was to purify AcrB in three different native nanodiscs for downstream experimentation. This also provided an opportunity to see how the different types of native nanodiscs solubilised AcrB, to act as a guide for future work. As an integral membrane protein, AcrB had to be solubilised from *E. coli* inner membranes. The first AcrB purification was with SMA, to form AcrB SMALPs. **Figure 5.3a** shows a traditional AcrB purification of AcrB in SMALPs. Usually, this protocol yielded a protein concentration of ~1.2 mg/mL, equivalent to ~10 µM. AcrB is a 343 kDa trimer, making the monomeric unit 114 kDa; **Figure 5.3a** shows the sodium dodecylsulfate (SDS)-PAGE gel for a typical purification of AcrB in SMALPs. The gel shows the purification from the SMA solubilised membrane and throughout the Ni^{2+} affinity chromatography, and clearly shows an intense band ~110 kDa representing the AcrB monomer. This reveals that SMA is able to solubilise AcrB efficiently, and SMALPs are able to bind Ni^{2+} and not be lost in the resulting washes or buffer exchange. It is important to note, that proteins encapsulated in SMALPs bind Ni^{2+} less efficiently than detergent micelles, likely due to interactions between the Ni^{2+} resin and free SMA, which is why SMALP protocols involve binding to Ni^{2+} overnight.[505,516] The sample shows as an intense band, with a high purity. Smaller faint bands can be seen in the sample lane, which could be co-puritants within the SMALP or degradation. Proteins in SMALPs often contain significant amounts of free SMA in the buffer, which can interfere with UV detection methods to estimate protein concentration.[114] Therefore, a size exclusion chromatography (SEC) protocol was optimised for SMALPs as a further purification step (**Figure 5.3b**). SMALPs can exhibit non-specific binding to surfaces such as the stationary phase of SEC columns, however running SEC methods at room temperature rather than 4 ˚C seemed to significantly reduce this. As **Figure 5.3b** shows, the SEC trace reveals two main peaks, and absorbance was followed at 280 nm representing the protein, and 254 nm representing the polymer.[530] The first peak showed strong absorbance for both 280 and 254 nm at ~10 mL, although a higher mAU for the 280 nm peak, corresponding to the AcrB trimer in the SMALP. The second peak was only seen at 254 nm and showed the free SMA polymer being removed from the sample. Further characterisation by mass photometry and SMA-PAGE confirmed AcrB to be a stable trimer in SMALPs (**Figure 3.6/7**).

Next, AcrB was purified in DIBMA native nanodiscs (**Figure 5.3c**). As stated in **section 5.1**, DIBMA lacks the styrene moiety and forms larger nanodiscs than SMALPs.[526] The SDS-PAGE gel shows DIBMA effectively solubilised the AcrB from the inner membrane, and protein is retained throughout the Ni^{2+} affinity chromatography. However, the AcrB elution band is far less intense than what is observed for

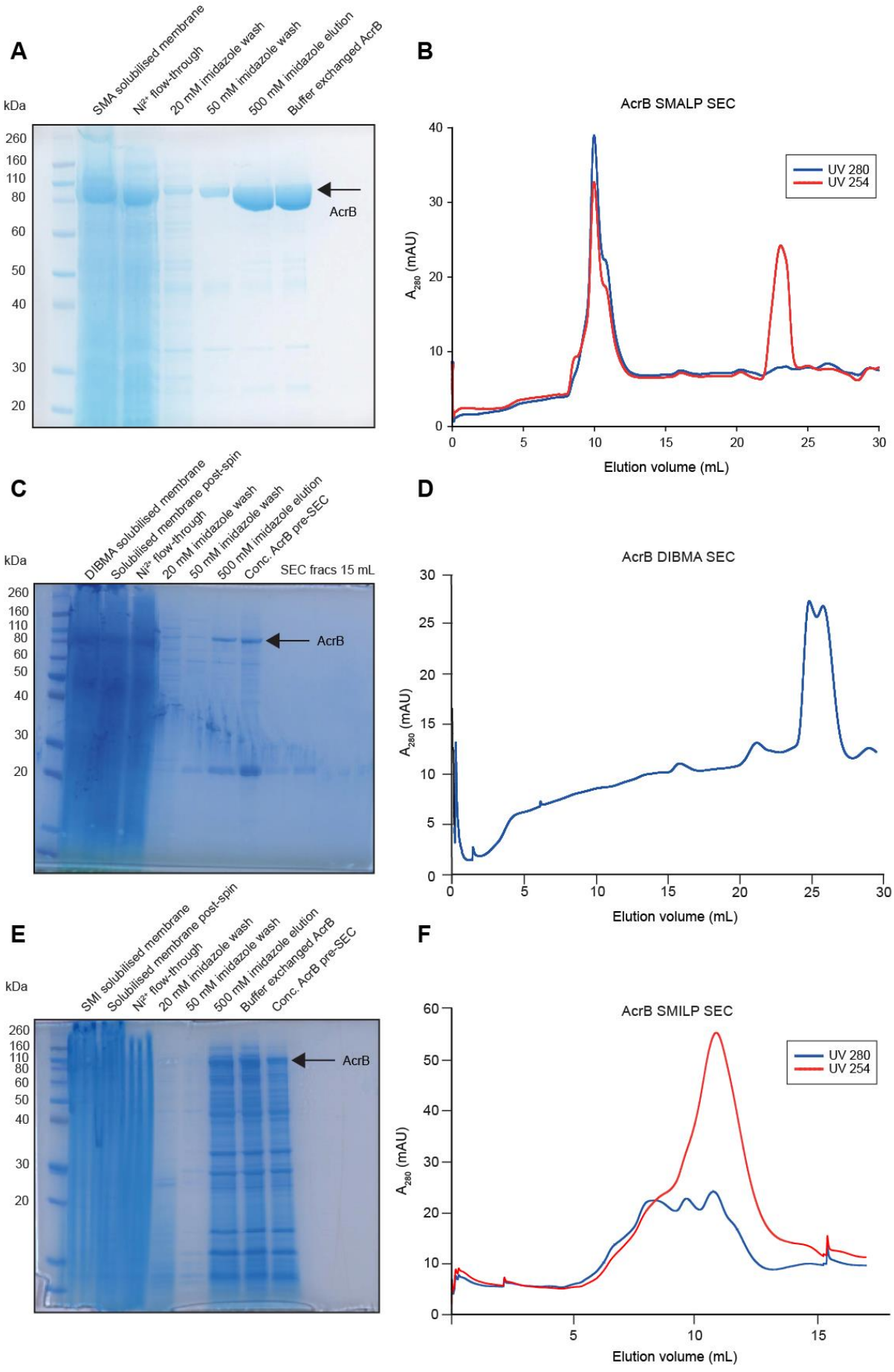

A
SMA solubilised membrane
Ni2+ flow-through
20 mM imidazole wash
50 mM imidazole wash
500 mM imidazole elution
Buffer exchanged AcrB
kDa
260
160
110
80
60
50
40
30
20
AcrB
B
AcrB SMALP SEC
UV 280
UV 254
A280 (mAU)
Elution volume (mL)
C
DIBMA solubilised membrane
Solubilised membrane post-spin
Ni2+ flow-through
20 mM imidazole wash
50 mM imidazole wash
500 mM imidazole elution
Conc. AcrB pre-SEC
SEC fracs 15 mL
kDa
AcrB
D
AcrB DIBMA SEC
A280 (mAU)
Elution volume (mL)
E
SMI solubilised membrane
Solubilised membrane post-spin
Ni2+ flow-through
20 mM imidazole wash
50 mM imidazole wash
500 mM imidazole elution
Buffer exchanged AcrB
Conc. AcrB pre-SEC
kDa
AcrB
F
AcrB SMILP SEC
UV 280
UV 254
A280 (mAU)
Elution volume (mL)

Figure 5.3. AcrB purifications in different polymer nanodiscs. A. SDS-PAGE showing AcrB solubilisation from the *E. coli* inner membrane with SMA, and subsequent Ni^{2+} affinity purification. **B.** Size exclusion chromatography of AcrB SMALPs. AcrB SMALPs (blue) elute at ~10 mL whereas free SMA polymer elutes at ~22.5 ml. **C.** SDS-PAGE showing AcrB solubilisation from the *E. coli* inner membrane with DIBMA, and subsequent Ni^{2+} affinity purification. **D.** Size exclusion chromatography of AcrB DIBMALPs. UV at 254 nm was not needed as DIBMA lacks the styrene moiety. Even in the presence of 0.2 M arginine, DIBMALPs associated with the SEC column and did not elute properly. **E.** SDS-PAGE showing AcrB solubilisation from the *E. coli* inner membrane with SMI, and subsequent Ni^{2+} affinity purification. **F.** Size exclusion chromatography of AcrB SMILPs. SMILPs (blue) elute at ~7-12 mL but individual AcrB SMILPs were hard resolve due to the number of contaminants. Free SMI polymer elutes at ~10-13 mL.

SMALPs. This was reflected in the concentration of AcrB obtained; this protocol yielded ~0.2 mg/mL of AcrB, equivalent to ~1.7 µM. This is a roughly 5-fold less yield than is obtained using the SMA polymer. This is likely due to DIBMA being more disruptive to the His-Ni^{2+} interaction during the affinity chromatography stage, as has been seen previously.[526,532] Furthermore, there is an intense band at 25 kDa, which is a contaminant. This protein appears in the SMALP purification too (**Figure 5.3a**) but is significantly more intense in the DIBMA purification. The more intense impurities is attest to the larger disc size of the DIBMALP, which allows for other proteins to also be encapsulated in the same disc as the target protein.[526] Therefore, as a further purification step, SEC was attempted to try and remove the impurity (**Figure 5.3d**). Previous work by Pellowe *et al.* (2020) have shown that DIBMA nanodiscs can interact with the stationary phase of SEC columns, preventing its elution from the column.[532] To counter this all mobile phase buffers required the presence of 0.2 M arginine to mitigate the interactions of the polymer.[533] However, even in the presence of arginine the DIBMA did not elute properly and stuck to the column, resulting in the trace seen in **Figure 5.3d**.

The last polymer nanodisc that AcrB was purified in was SMI. SMI has the opposite pH dependence to SMA, and SMILPs are only soluble below pH 7.8.[529] SMILPs also have a very high tolerance to divalent cations unlike SMALPs.[530] **Figure 5.3e** shows the SDS-PAGE for this protocol. Again, the gel shows SMI can solubilise AcrB from the inner membrane, but the purified protein is very dirty with many equally as intense proteins bands as the AcrB band. This is likely due to the different conditions of the Ni^{2+} affinity. Due to SMI only being soluble at acidic conditions, Ni^{2+} was completed with the same buffers as the SMA protocol but at pH 6.5. At pH 6.5, more His residues become protonated and are less efficient at binding Ni^{2+}. Therefore, the His_8-tag on AcrB was less efficient and non-specific binding was more prevalent. To rectify this the protocol can be modified; for purifications using SMI, the NaCl concentration should be increased to 300 mM to reduce non-specific binding due to electrostatic interactions with the resin. Furthermore, 20 mM imidazole should be added when binding to Ni^{2+} to outcompete other contaminants binding to the resin. As with the SMA and DIBMA purifications, SEC

was completed on AcrB SMILPs at pH 6.5 (**Figure 5.3f**). Unlike DIBMA, the SMILPs did not stick to the column and did elute. However due to the vast amounts of contaminants, several peaks came off the SEC in close proximity which affected the resolution.

Therefore, AcrB can be purified in using three different native nanodiscs, which is important for downstream experiments. Since the different polymers have different advantages, the experiment required can govern which polymer should be used for AcrB experiments. SMA produced the most protein and the cleanest purification, but DIBMA and SMI purifications can be further optimised to produce high quality samples too. This purification can be compared to the AcrB purification in DDM, which isn't shown in this chapter but in **Appendix 9**.

5.2.2 Studying the effect of MBX-3756 inhibitor on AcrB in a lipid environment

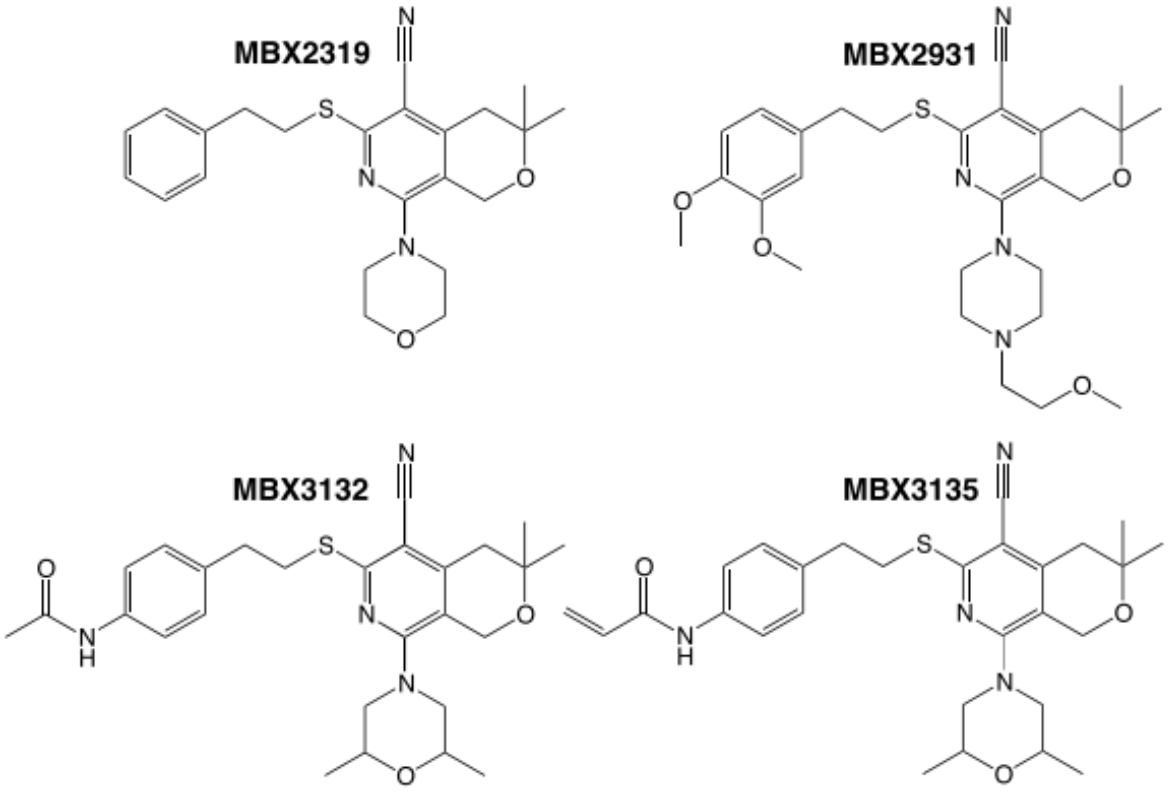


Figure 5.4. Pyranopyridine series of EPIs. From Wang *et al.* (2017).[427]

MBX-2319 (**Figure 5.4**) is an efflux pump inhibitor (EPI) that was discovered in 2014 and was shown to have potent activity against resistance nodulation and cell division (RND) efflux pumps of Enterobacteriaceae species.[483,484] MBX-2319 was able to fully potentiate the activity of levofloxacin and piperacillin at concentrations as low at 3 µM, and does not exhibit membrane-disrupting or antibacterial activity. Through a combination of crystal structures, molecular dynamics (MD)

simulations, docking studies and cellular assays, the mechanism of action has been proposed.[484,485,534] MBX-2319 binds tightly to the lower part of the distal binding pocket of AcrB, and can prevent the binding of substrates to the distal binding pocket through steric hinderance. Furthermore, MBX-2319 has defined interactions with residues in the hydrophobic trap of AcrB. This may prevent the AcrB protomer from making the Tight (T) – Open (O) conformational transition which would effectively halt the functional rotational mechanism of AcrB. Since MBX-2319, structural information of its binding to AcrB has led to the production of several MBX derivatives (**Figure 5.4**) to improve the drug-like properties and activity, thus creating the pyranopyridine series of EPIs.[427,485] MBX-3756 is a pyranopyridine, and a *trans*-isomer of MBX-3132 (*trans*-isomer of the dimethyl-morpholino group) that was send to the Reading group by a collaborator. However, there was no information on exact K_D measurements for MBX-3756, so exact amounts of protein:ligand occupancy could not be calculated, and experiments were essentially a 'shot in the dark'. Therefore, the aim was to utilise HDX-MS to observe any effect of MBX-3756 on AcrB structural dynamics. This would provide a first look at how the drug is interacting with AcrB and enable further experiments to be designed to investigate its mechanism of action.

5.2.2.1 Optimising the AcrB sample for HDX-MS experiments

Due to advances in HDX-MS practical methods, such as online lipid removal using a phospholipid trap column, having AcrB in a lipid environment was important to provide a more "native" look AcrB structural dynamics.[4] SMALPs were not suited to online HDX-MS, as under quench conditions (pH 2.5) SMALPs would precipitate out of solution and causes blockages throughout the MS tubing.[112] The use of SMILPs would be ideal due to their suitability to HDX-MS quench conditions, however the current purification protocol is not optimised to a high enough standard to produce a reliably clean sample.[529] Therefore, AcrB in MSP nanodiscs were used for HDX-MS investigations utilising the novel online delipidation workflow.

Since the lipid composition of MSP nanodiscs can be controlled, *E. coli* polar lipid extract (PLE) lipids were trialled first as they would best represent the native *E. coli* membranes (**Appendix 10**). However, the yield for this purification was too low for HDX-MS experiments under two states and in triplicate, and for coverage map/ delipidation tests. Therefore, 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) MSP nanodiscs were used instead, as this produced enough sample to be used for HDX-MS. **Figure 5.5** shows the characterisation for the MSP nanodiscs. **Figure 5.5a** shows the SEC trace, which showed one void peak and then a main AcrB MSP peak ~10 ml– this is the volume where AcrB SMALPs also elute from, suggesting they are similar in size. This has also been seen for MexB in POPC MSP nanodiscs and SMALPs.[372] The SEC fractions were run on an SDS-PAGE (**Figure 5.5b**), with fractions 8,

9 and 10 showing the most intense bands of both AcrB and MSP1E3D1; these fractions also represented the middle of the main peak. These were therefore collected for HDX-MS experimentation. **Figure 5.5c** shows the dynamic light scattering (DLS) data for confirming the size and heterogeneity of the AcrB MSP nanodiscs – the data showed a homogenous peak with a nanodisc size of 13.89 nm. The yield was calculated by estimating the overall protein concentration; determining the concentration of only AcrB was tricky due to the presence of MSP, which is a protein and also absorbs UV at 280 nm. Therefore, the molecular extinction coefficients were combined, and a total protein concentration calculated. This was ~0.33 mg/mL, equivalent to ~2 µM of sample.

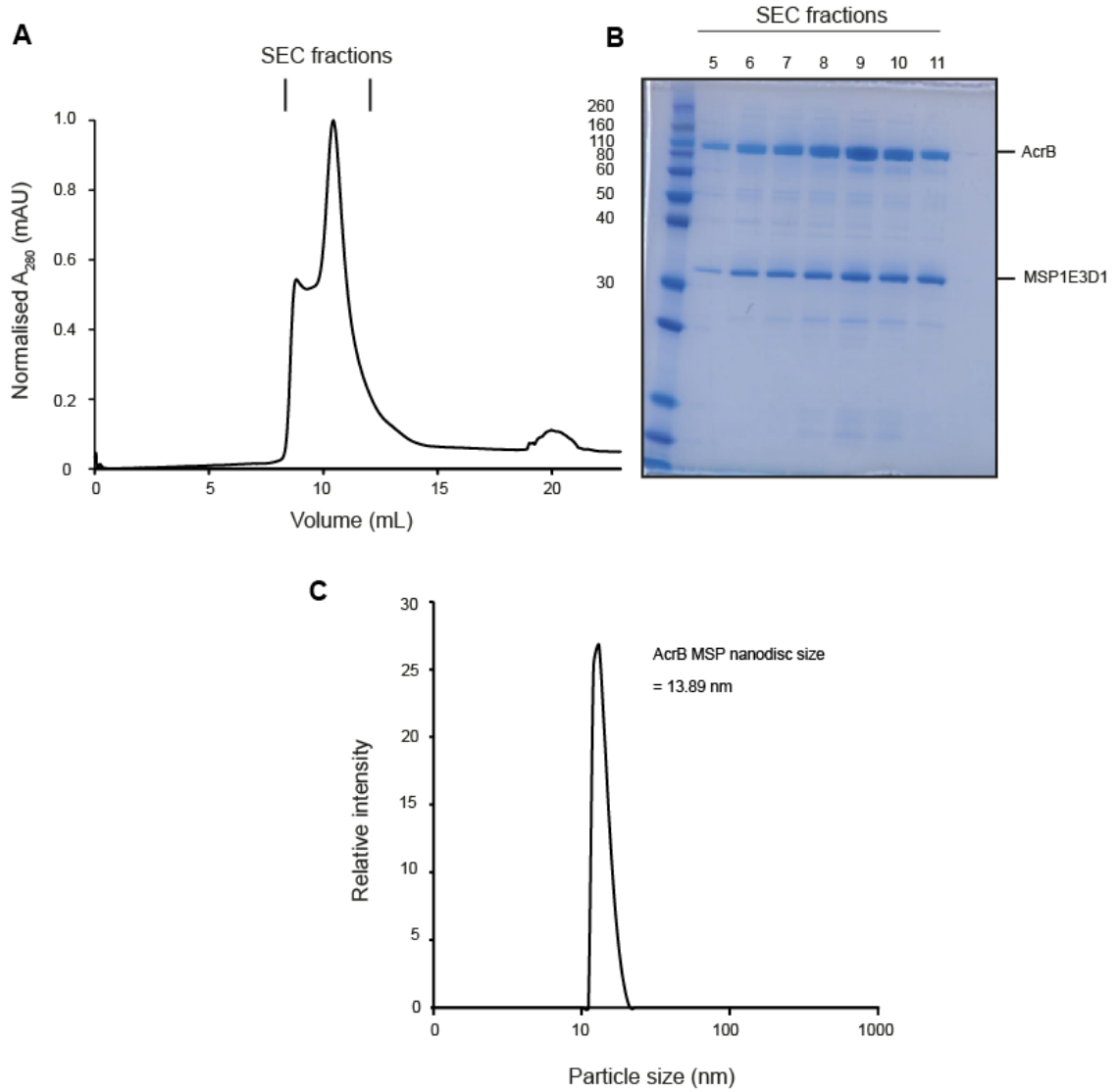


Figure 5.5. Characterising AcrB MSP nanodiscs. MSP nanodiscs were made using POPC lipids. **A.** Size exclusion chromatography for AcrB MSP nanodiscs. Void peak can be seen at ~9 mL and AcrB MSP nanodisc elutes ~10 mL. **B.** The fractions from the SEC run on an SDS-PAGE. AcrB and MSP1E3D1 can be observed, with the most intense bands in fractions 8-10, the core of the peak at 10 mL. **C.** Dynamic light scattering (DLS) of AcrB MSP nanodiscs. DLS shows sample homogeneity and a size of 13.89 nm.

5.2.2.2 HDX-MS of AcrB MSP nanodiscs

HDX of AcrB in MSP nanodiscs was collected for three labelling time points (10s, 100s, 1000s) at pH 7.4. The first test was to achieve a coverage map of AcrB MSP nanodiscs and optimise the experimental conditions before collecting further experiments. As any experiment containing MBX-3756 would have 5% dimethyl sulfoxide (DMSO), to ensure ligand saturation and solubility, 5% DMSO was kept consistent throughout all samples. **Figure 5.6** shows the coverage map achieved for AcrB, post analysis. A total of 177 peptides were identified to a high-quality standard, with a peptide coverage of 77.8%. This confirmed that the online deplidation protocol led to high membrane protein coverage.[4] In fact, previous HDX-MS experiments with AcrB in DDM resulted in a peptide coverage of 72%, therefore the novel workflow yielded better peptide coverage than traditional methods.

Figure 5.6. Coverage map of AcrB in MSP nanodiscs. 177 peptides post analysis, resulting in 77.8% coverage and a 2.90 redundancy.

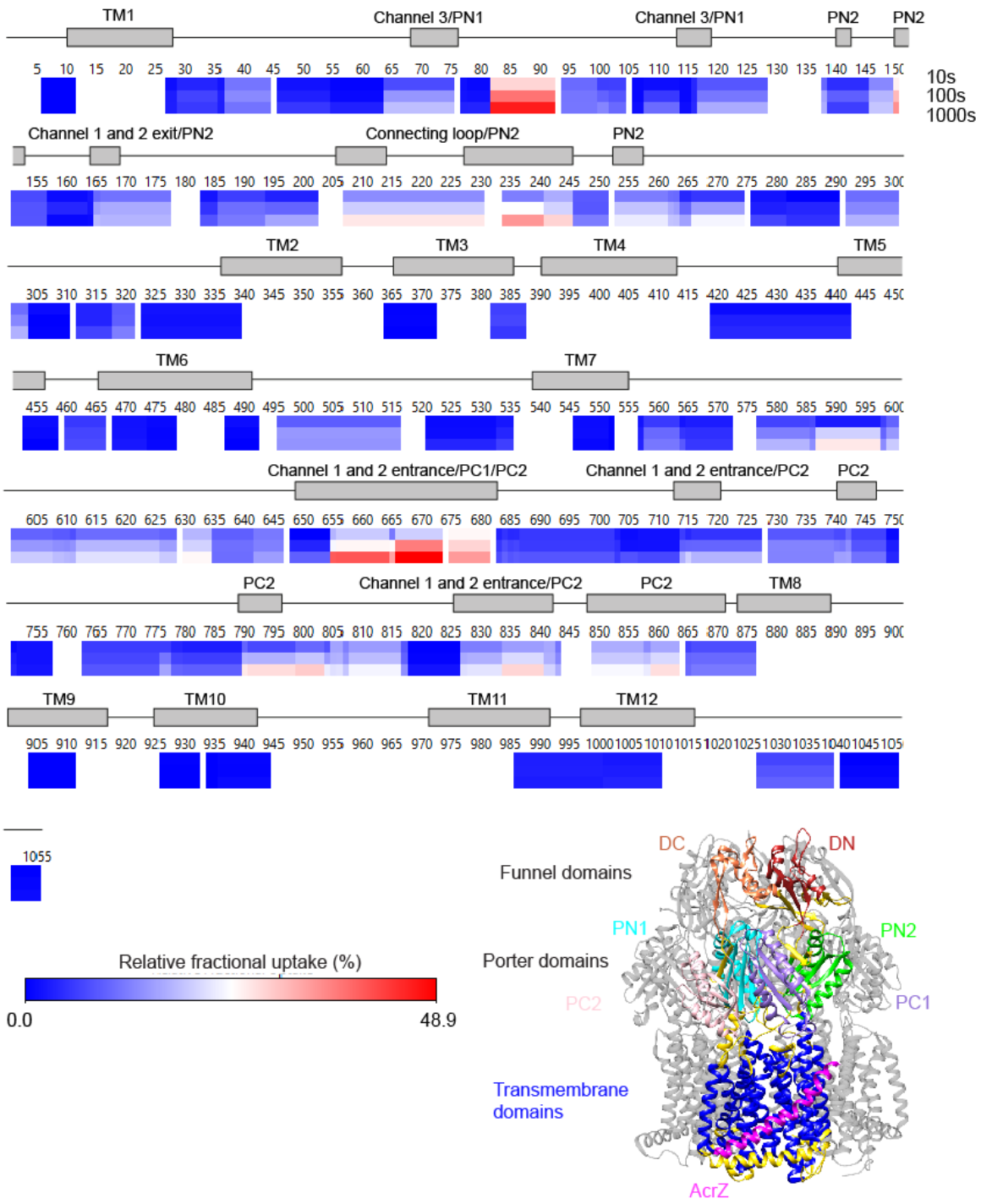


Figure 5.7. Relative fractional uptake of AcrB in MSP nanodiscs. Relative fractional deuterium uptake analysis of AcrB in MSP nanodiscs for all three timepoints (10s, 100s, 1000s).

Figure 5.7 shows the relative fractional uptake (RFU) of AcrB displayed as a heatmap. It shows that many regions of AcrB form part of stable structures, due to marginal amounts of deuterium incorporation, after all labelling timepoints, across significant portions of the protein. Of particular note, the transmembrane (TMs) helices of AcrB show little-to-no deuterium uptake across the entire HDX time course, which is likely due to their extensive protection in the hydrophobic lipid environment of the MSP nanodisc.[113] The most flexible regions of AcrB are the subdomains of the porter domain (PC1, PC2, PN1, PN2), which agrees with previous RFU analysis of AcrB.[35] Interestingly, this is the region where the proximal and distal binding pockets are formed. This increased dynamism in the porter domain compared to the rest of the protein, likely benefits its efflux mechanism, which involves substrates moving through the binding pockets in a peristaltic manner.[386]

5.2.2.3 The effect of MBX-3756 on AcrB structural dynamics

To measure the effect of MBX-3756 on AcrB, differential HDX (ΔHDX) was performed ((AcrB + MBX-3756) -AcrB). As there was no kinetic information available for MBX-3756 binding to AcrB, it was not possible to work out a concentration of ligand to saturate AcrB. Therefore 100 µM of MBX-3756 was added, to ensure as much ligand as possible was present and DMSO was kept at 5%. The buffers were the same as previous HDX-MS experiments (see **chapters 2-4**), except for the use of a 500 mM glycine-HCl at pH 2.3 quench buffer, which was essential for the online delipidation workflow and an extended LC gradient of 9.0 minutes from 8-40% solvent B at 40 µL/min, due to the large mass of AcrB. This experiment contained three technical replicates, as there was not a high enough sample concentration for more. Therefore, due to the limited technical replicates and lack of biological replicates, this HDX experiment acted as a check to validate the online delipidation workflow, and to see if MBX-3756 binding could be observed. Due to the fact the sample concentration only allowed for measurements to be made in triplicate, Deuteros 2.0 was used to calculate significant peptides.[535] Deuteros 2.0 uses (n+1) to calculate the degrees of freedom when calculating the confidence interval (CI), rather than (n-1) which is used in the previous chapters; this is more suitable for triplicates and provides less stringent CI values, with the aim of not missing statistically significant changes (see **section 2.4.5.2.1**).[463,535,536] Furthermore, previous work with AcrB that monitored the effect of another EPI Phe-Argβ-napthylamide (PAβN), used Deuteros 2.0 for calculating statistics.[35] Therefore, significance was deemed to be a ≥ 0.25 Da change and with a *P*-value ≤ 0.05 in a Welch's *t*-test.

In the presence of MBX-3756, AcrB exhibits protection across the entire HDX time course between residues 610-630 (**Figure 5.8**). **Figure 5.8c** shows an example peptide in this region (^{611}AVNGFGFAGRGQNTGIAF^{628}) showing a statistically significant protection in the presence of MBX-3756. This area of AcrB is located within the substrate binding pocket, and contains several residues

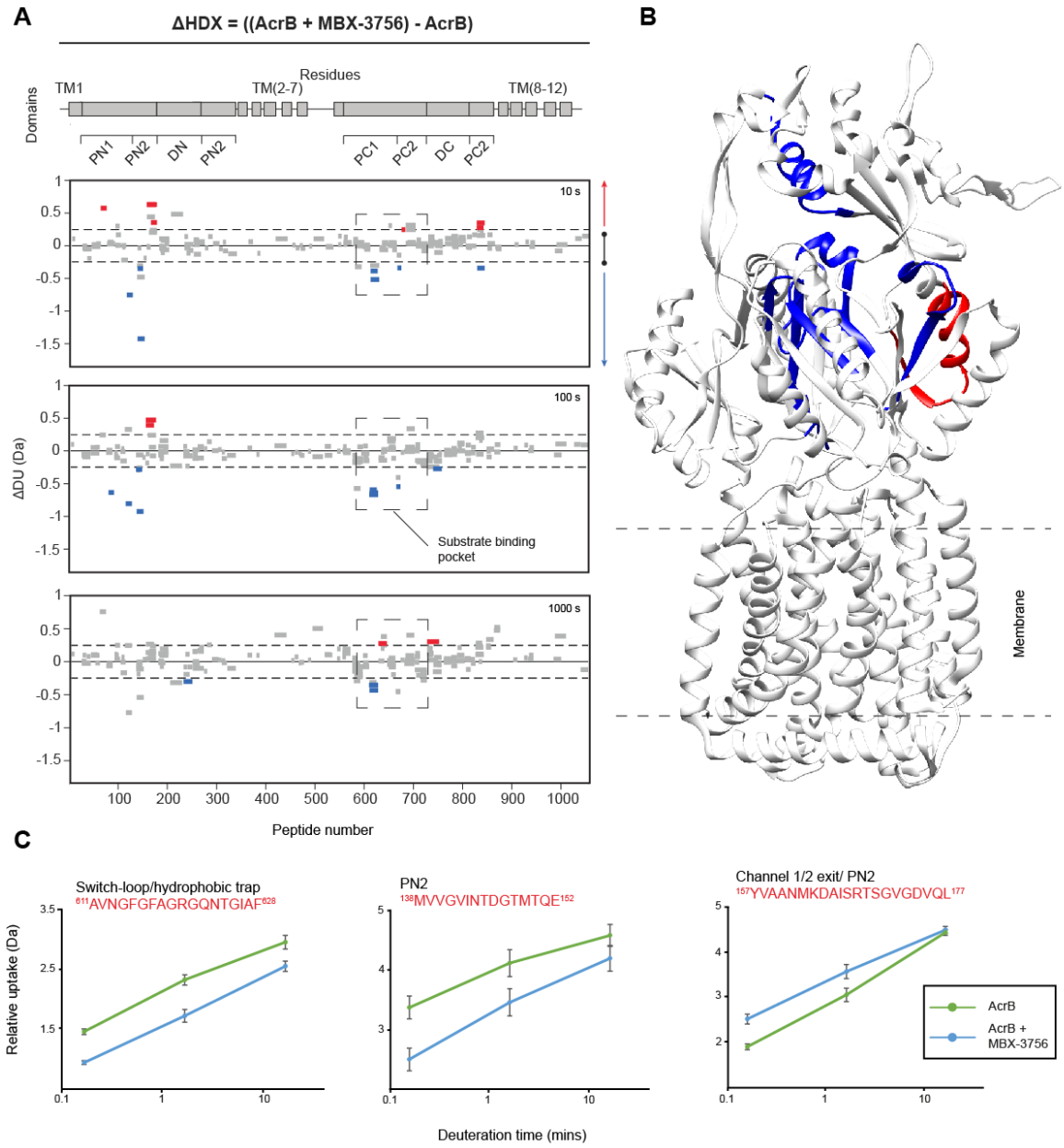


Figure 5.8. The effect of MBX-3756 on AcrB structural dynamics within a lipid environment. A. Woods plot displaying the differential HDX (ΔHDX) plots for ((AcrB + MBX-3756) – (AcrB)) for all time points collected. Blue signifies areas with decreased HDX between states, red signifies areas with increased HDX. Significance was defined to be ≥ 0.25 Da change with a *P*-value ≤ 0.05 in a Welch's *t*-test ($n_{technical}$ = 3). Gray areas represent regions with insignificant ΔHDX. **B.** ΔHDX for the middle time point is painted onto the AcrA structure (PDB:5O66) using Deuteros and Chimera.[448,535] **C.** Uptake plots for three peptides in areas of AcrB. Uptake plots are the average deuterium uptake and error bars indicate the standard deviation.

found in the hydrophobic trap (F610, V612 F615, I626, F628).[35] The HDX data also reveals regions 116-149 contains several protected peptides in the first two time points, which also contain several residues found in the distal binding pocket (S128, E130, S132, S134, F136, V139). This result is similar

to results found for other MBX derivatives. The cryo-EM structure of MBX-3132, of which MBX-3756 is a *trans*-isomer, observed its interactions with multiple residues in the hydrophobic trap, including V612 and F615 which are also in the significant peptides highlighted by HDX.[427] Furthermore crystal structures of MBX-2319 bound to a soluble construct of AcrB lacking the TM domains, showed it also bound to AcrB with many hydrophobic interactions with residues in the distal binding pocket and the hydrophobic trap.[485] The central aromatic ring of MBX-2319 was oriented parallel to the F628 side chain resulting in an extensive π–π stacking interaction. Furthermore, the phenyl and morpholinyl groups interacted with F178 and F615, and the F610 side chain was packed against the dimethylenesulfide moiety of MBX-2319. Our HDX data suggests MBX-3756 is still interacting with the hydrophobic trap, and thus may be inhibiting AcrB through steric blocking of other substrates binding. However, it is likely the orientation of the molecule is somewhat different than MBX-3132, as the *trans*-isomerisation creates a structurally unique molecule which likely binds slightly differently in the binding pocket, and further experimentation such as cryo-EM and MD simulations are needed to reveal how MBX-3756 is interacting within this region.

The HDX data reveals some similarities to another EPI, outside the pyranopyridine series. PAβN was discovered by Lomovskaya *et al.* (2001) and has been shown to inhibit RND pumps of several bacterial species, including AcrB from *E. coli.*[482] Previous work has suggested that PAβN binds to the bottom of the of the distal binding pocket, interacting with the AcrB switch loop (residues 615-620) and other nearby regions to disrupt substrate binding and possibly trap AcrB in a T-like conformation, restricting its rotational transport mechanism.[35,537,538] Reading *et al*. (2020) performed a similar experiment as described in this section, performing HDX to monitor the effect of PAβN on AcrB structural dynamics. As observed with MBX-3756, PAβN caused protection among residues in the binding pocket, including the switch loop. However, aside from stabilisation effects, HDX revealed that PAβN can cause destabilisation in areas of the PN2 and PN1 subdomains, suggesting an increase in dynamics. **Figure 5.8a** shows that MBX-3756 also causes some regions of AcrB's PN2 subdomain to become more dynamic in the first two time points, suggesting there is some similarities in the way the two classes of EPI are affecting AcrB. Peptide ^{157}YVAANMKDAISRTSGVGDVQL^{177} (**Figure 5.8c**) shows an example of a peptide in the PN2 subdomain exhibiting increased uptake in the presence of MBX-3756. Whilst PAβN has never made it to clinical trials due to toxicity issues, it shows that drugs binding to the distal binding pocket of AcrB may be common mechanism of AcrB EPIs.[483]

Overall, this study works as a validation of the novel online delipidation workflow. Here, it has been possible to observe HDX on AcrB within MSP nanodiscs, and observe drug binding effects on AcrB structural dynamics, all using an entirely automated workflow to remove lipids before analysis using MS. Furthermore, the data collected resulted in a higher peptide coverage than traditional HDX-MS

methods with AcrB in DDM micelles.[35] The HDX-MS data provides a useful insight into MBX-3756 binding, however further experiments are needed to determine its precise interactions in the binding pocket and to elucidate its mechanism of action.

5.2.3 Utilising a novel SMALP-liposome-SMALP assay to study the effect of antimicrobial peptides on AcrB within a lipid environment

In 2020, a study by Jesin *et al*. (2020) described a peptide based approach to inhibit AcrB.[426] They designed two antimicrobial peptides to be targeted to the inner membrane, and mimic TM1 and TM8 of AcrB. The aim being to promote binding of the peptide to its endogenous binding partner to disrupt the native packing of AcrB. Synthetic peptides have been used previously to disrupt membrane protein subunit interactions in the lipid bilayer, by impeding the formation and function of membrane-embedded oligomers, on the small multidrug resistance (SMR) class of efflux pumps.[539] Previous work on these peptides showed that they potentiated antimicrobials, showing their activity as an EPI, and no effect was observed in a Δ*acrB* strain, proving their specificity to AcrB. Nile red assays also showed the addition of the antimicrobial peptides decreased efflux by 20-27%.

However, no data was available to see if the peptides were able to cause AcrB to dissociate into respective monomers, by binding and outcompeting the native AcrB interactions, or if they disrupt the function of AcrB by a different mechanism. Therefore, utilising the optimised SMALP protocols from **section 5.3.1**, SMA-PAGE was first applied to see the effect of these peptides on AcrB within SMALP nanodiscs, and thus a native lipid environment rather than detergent micelles which can make membrane proteins more unstable anyway.[540] Two peptides were sent from the Deber group, Peptide-TM1 and Peptide-TM8, mimicking TM1 and TM8 of AcrB respectively. A further two peptides were also sent as controls, Peptide-TM1scr and Peptide-TM8scr, which had scrambled sequences so were expected to have no effect on AcrB.

Figure 5.9a shows the first attempts at this assay. Proteins and peptides were incubated at a 1:1 ratio at 6 µM and incubated at 37 ˚C for 1 hour. This was repeated in triplicate (**Appendix 11**) but only one gel is shown in **Figure 5.9a.** Since binding affinities were unknown, these were the conditions that best reflected the previous work carried out by the Deber group.[426] The peptides were stored in DMSO, but the final DMSO concentration was kept at 0.4%, which had no effect on the SMA-PAGE of AcrB (**Figure 5.9a**). As the gel shows, none of the peptides, including the controls, appear to have any effect on the oligomeric state of AcrB, as the trimer band could be seen under each condition, with and without peptides, appearing the same as the AcrB input. If AcrB fell apart into respective monomers/dimers, bands would appear under the trimer band, but no new species are observed. Therefore, the assay was repeated at a higher 10:1 ratio of peptide:protein, to see if this had any effect (**Figure 5.9b**). The

incubation time and temperature were kept the same, however a 30 minute incubation step (on ice) was added after the incubation at 37 ˚C, as this reduced smearing on the gel.

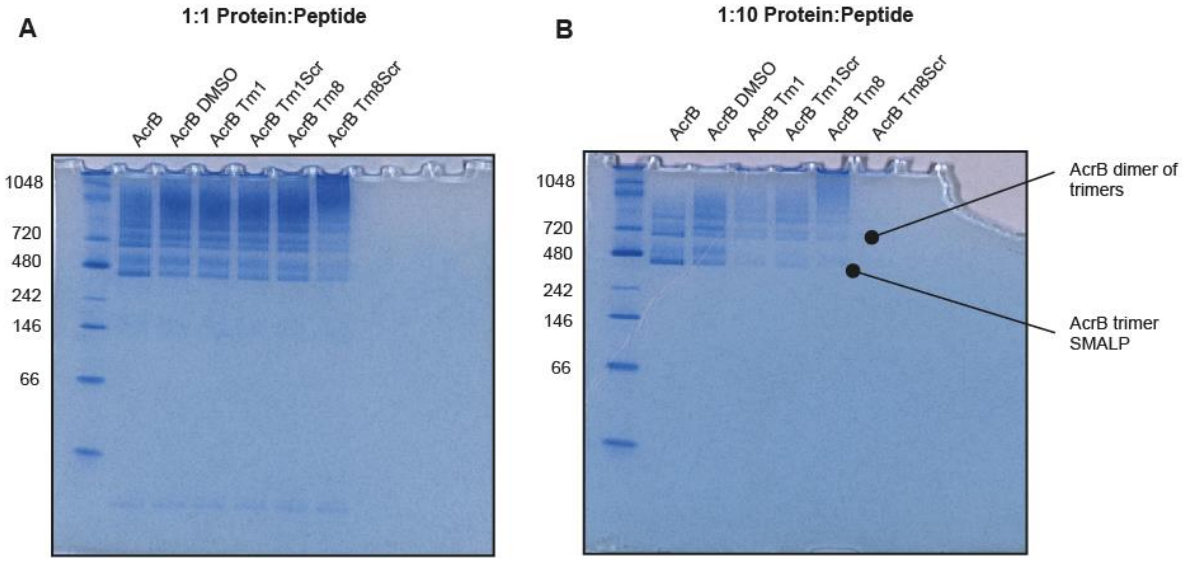


Figure 5.9. SMA-PAGE of AcrB and antimicrobial peptides. AcrB incubated with peptides at 37 ˚C for 1 hour at 400 revolutions per minute (RPM). Gel ran at 150 V for 90 mins at 4 ˚C. **A.** Shows the gel for a 1:1 peptide:protein ratio. **B.** Shows the gel for a 10:1 peptide:protein ratio.

Figure 5.9b shows that the AcrB trimer is severely reduced and even almost entirely drops out in the presence of the peptides, whereas it is fine in the two input lanes. However, due to the lack of lower molecular weight bands that would suggest monomers/dimers present and the fact the peptide controls exhibit the same effect, it appears as though the integrity of the nanodisc is being compromised with the excess of peptide present, rather than the peptides causing AcrB to dissociate. It may be the peptides are able to insert into the nanodisc and disrupt its structural integrity, and at high enough peptide concentrations significant amounts of sample are lost.

One possibility for the results seen in **Figure 5.9a** may be the peptides are struggling to access AcrB due to the SMALP itself. The bulky, negatively charged nanodisc may be preventing the peptides accessing AcrB and enacting its effect.[541] Thus, a novel assay was developed to bypass this possibility, and investigate whether the peptides are able to dissociate AcrB trimers (**Figure 5.10a**). Essentially, AcrB SMALPs were reconstituted into liposomes containing *E. coli* PLE lipids, before incubation with the peptides, to allow the peptides to reach AcrB easier, and then the whole reaction was re-solubilised in SMALPs and analysed on a SMA-PAGE. The experiment was performed on the same day as the AcrB reconstitution into liposomes, after a quality control check using DLS (**Figure 5.10c**). Again,

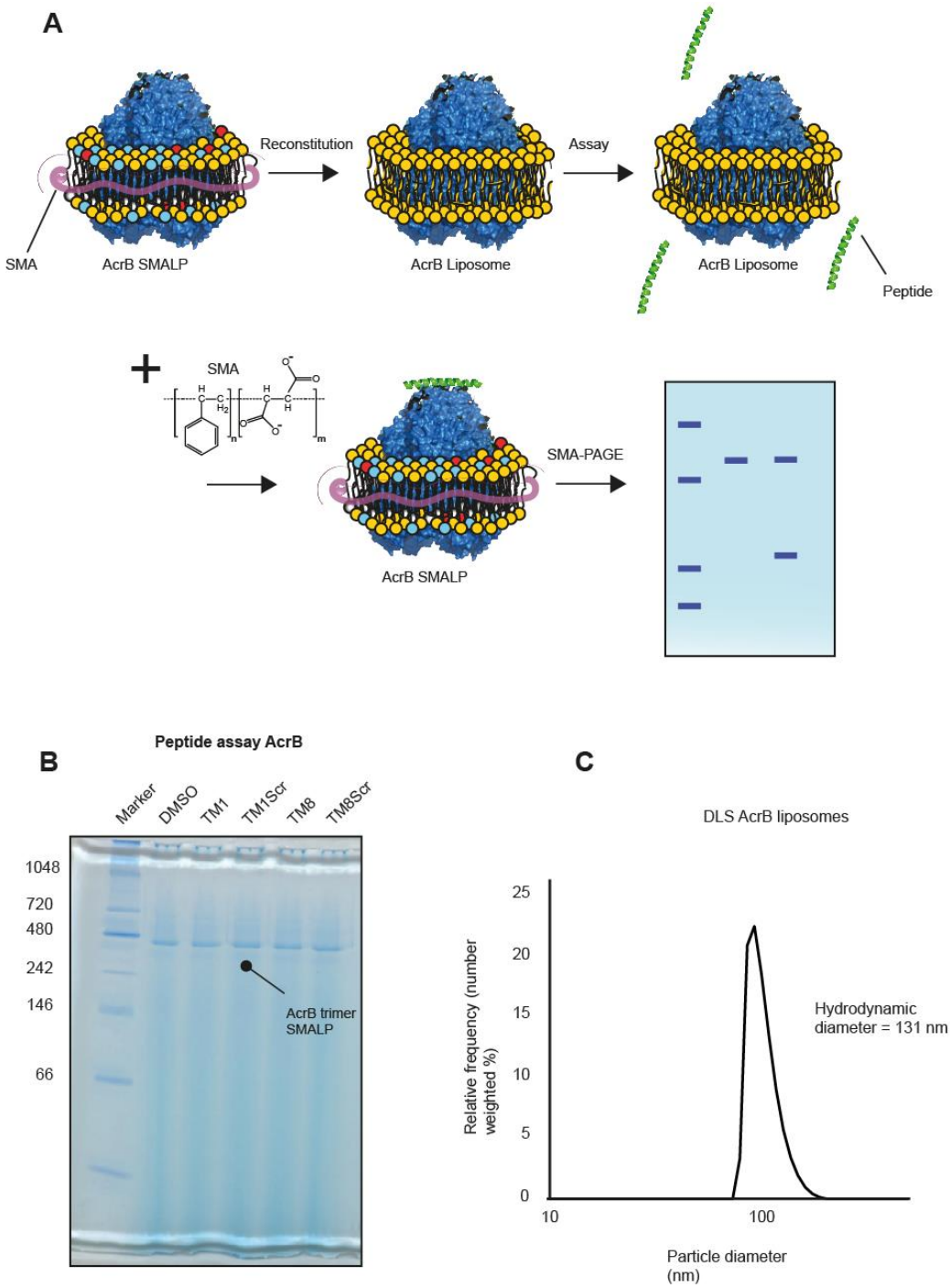


Figure 5.10. SMA-liposome-SMA assay of AcrB and antimicrobial peptides. **A.** Schematic detailing the workflow. AcrB SMALPs were reconstituted into liposomes containing *E. coli* polar lipid extract (Avanti). AcrB was incubated at room temperature with 2 µM peptides (0.4% DMSO) for 1 hour. Samples were then resolubilised in 2.5% SMA, before analysis via SMA-PAGE. **B.** SMA-PAGE of AcrB + peptides. AcrB trimer band can be seen at ~480 kDa for all conditions. Gel ran at 150 V for 2 hours at 4 ˚C. **C.** DLS of AcrB liposomes showing a hydrodynamic radius of 131 nm.

this was repeated in triplicate (**Appendix 11**) but **Figure 5.10c** shows one repeat as they all gave the same result. The peptides were added to AcrB in liposomes at a 1:1 concentration, and then incubated for 1 hour at room temperature, then re-solubilised in 2.5% SMA for 1 hour at room temperature,

followed by SMA-PAGE analysis. **Figure 5.10b** shows neither of the antimicrobial peptides or control peptides caused the AcrB trimer to dissociate into monomers, as the trimer band is the only band present across all protein lanes. Therefore, the results of the assay suggest that Peptide-TM1 and Peptide-TM8 are not causing the AcrB trimer to dissociate. Previous results clearly show these peptides have EPI activity against AcrB, however the mechanism of action is unknown.[426] It should be noted, this assay only examines the purified AcrB trimer *in vitro*; *in vivo* it is entirely possible the peptides affect the oligomerisation of AcrB during protein synthesis or membrane translocation. Interestingly, Lu *et al*. (2011) showed that AcrB trimerizes from folded monomers, so it is possible the antimicrobial peptides may affect the association of the TM helices *in vivo*.[542] However, this was not possible to investigate with this assay. It is also possible the peptides work by affecting the TM regions of AcrB – further experiments could use HDX to monitor the effect of the peptides on AcrB structural dynamics. The movement of TM2 and TM8 are essential in the rotational efflux mechanism of AcrB, as they cause the porter domain to undergo significant conformational movement to close access to the drug binding pocket and squeeze the substrate out into the exit duct.[386,388] Any changes to the dynamics of these regions may affect the ability of AcrB to efflux. Further work is needed to characterise the mechanism of action of these peptides.

5.2.4 Expressing and purifying the TolC construct in SMALP nanodiscs

The aim was to purify TolC in SMALPs for downstream experimentations to attempt assembly of AcrAB-TolC in SMALP nanodiscs (see **section 5.2.6**). TolC is an outer membrane factor protein (OMF) and thus is purified from the outer membrane of *E. coli*. Several methods exist to separate the inner and outer membrane during protein purification, such as sucrose gradients to separate the membranes based on density or selective solubilisation of one of the membranes.[543] The use of sucrose gradients resulted in lost protein due to the practically difficult procedure of collecting the outer membrane fraction. Therefore, the outer membrane was separated from the inner membrane by selectively solubilising the inner membrane using lauryl sarcosine, leaving an outer membrane pellet. A characteristic purification of TolC in SMALPs is shown in **Figure 5.11**. **Figure 5.11a** shows a SDS-PAGE gel following the whole purification of TolC in SMALPs. As can be seen, TolC SMALPs can be purified by solubilising the outer membrane and by Ni^{2+} chromatography. The mature TolC construct (with the cleaved signal sequence 2-22) has a monomeric mass of 52,722 Da, and the TolC monomer presents as a double band at ~53 kDa. The double band is characteristic of TolC and has been seen previously; this is due to the C-terminal end being cleaved at approximately R459 by *E. coli* proteases.[544] This is why only a single band presents on the western blot for TolC, as the cleaved band has the His_6-tag removed (**Figure 5.11b**), which was analysed by Anti-PolyHistidine-HRP Antibodies.

The western blot confirms the TolC monomer is present throughout the purification process and is our final sample. This purification protocol yielded ~0.35-0.40 mg/mL of protein, equivalent to ~7 µM.

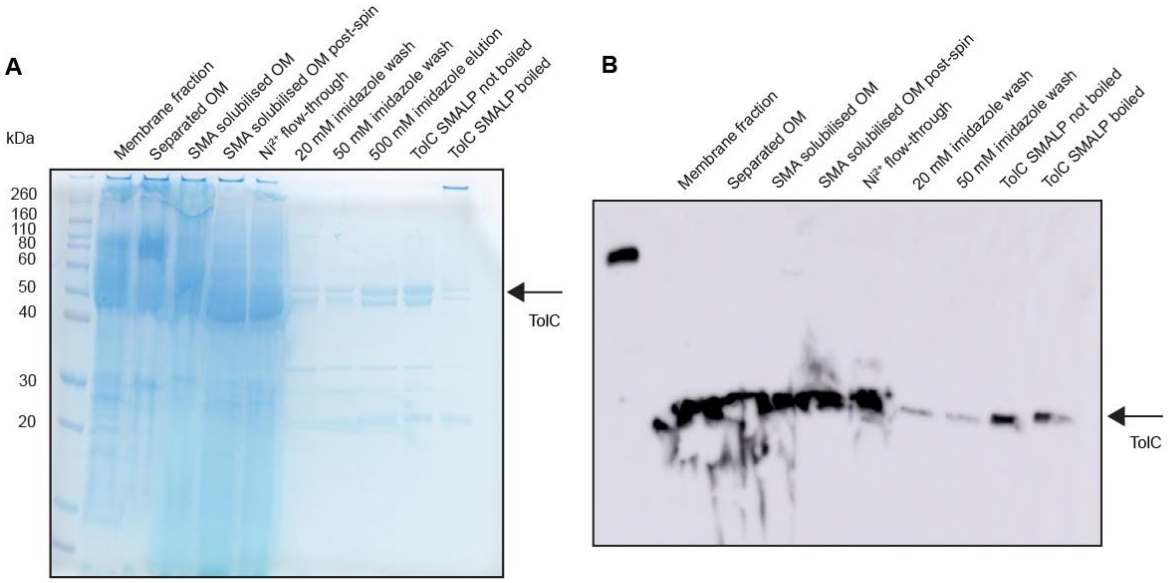


Figure 5.11. Characterisation of TolC SMALP purification. **A.** Characteristic SDS-PAGE of the TolC SMALP purification. The gel shows a double band at around ~53 kDa representing TolC. **B.** Western blot for TolC using Anti-PolyHistidine-HRP Antibodies, confirming the presence of the TolC-His construct.

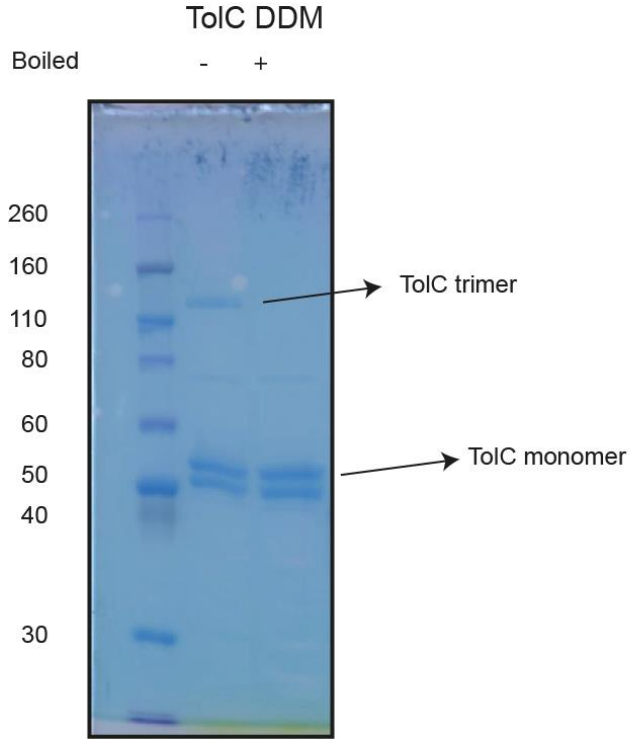


Figure 5.12. Boiled and unboiled samples of TolC DDM. A band at ~130 kDa can be seen for unboiled samples, showing a fraction of the SDS-resistant TolC trimer.

TolC is a unique β-barrel protein, as three monomer come together to form the TolC trimer, whereas most β-barrels form their structure from a single amino acid chain.[416] A common feature of β-barrel proteins is their unboiled samples run differently than samples that have been boiled, often due to extensive hydrogen bonding holding the β-strands together.[545] Previous work on TolC purified in DDM showed a fraction of TolC was stable in SDS without boiling, and ran at ~130 kDa. However, TolC SMALPs run entirely as a monomer on SDS-PAGE, regardless of boiling. Since TolC has not previously been purified in SMALPs, it is unknown whether this is an effect of SMA, or the construct itself, thus TolC was purified in DDM, and analysed on an SDS-PAGE with and without boiling. **Figure 5.12** shows that TolC in DDM without boiling has a fraction of the purification that runs at ~130 kDa, as previously seen, and the rest as a monomer. Therefore, it appears as if this increased susceptibility to SDS is a consequence of the SMALP.

5.2.5 Characterisation of TolC SMALPs

As TolC SMALPs present as a monomer on an SDS-PAGE, it was essential to examine whether the TolC trimer is present natively. For this mass photometry and SMA-PAGE were utilised to report on the size of the TolC SMALP sample (**Figure 5.13**). **Figure 5.13a** shows the SMA-PAGE gel for TolC in SMALPs, as well as for AcrB SMALPs and AcrA^{S}, except the latter two proteins are not discussed in this section. The SMA-PAGE reveals a TolC band ~242 kDa, which confirms the presence of the TolC trimer. The theoretical mass of the TolC trimer is 158 kDa, with mass difference owing to the size of the lipids in the nanodisc, as observed with AcrB (**Figure 3.6**, **section 3.2.5.1**). A higher band can be seen ~480 kDa, most likely resulting from TolC trimer-dimers, as is previously observed with AcrB in SMALPs.[465,474] **Figure 5.13b** shows mass photometry of TolC SMALPs, performed by Anna Olerinyova from the Struwe group (Oxford). The mass calculated by mass photometry for the TolC trimer was 255 kDa, which is excellent agreement with the SMA-PAGE result. Furthermore, mass photometry was also able to confirm the existence of TolC trimer-dimers, with a mass of 465 kDa. The mass photometry revealed a peak at ~80 kDa, which likely represented empty SMALPs as previously seen (**Appendix 5**). These two results confirmed TolC exists as a trimer within the SMALP nanodisc. Lastly, solubilisation of the outer membrane should result in high levels of lipopolysaccharide (LPS) encapsulated in the nanodisc. To confirm this, a dot blot was performed on protein buffer, AcrA, AcrB-SMALP, TolC-SMALP and bovine serum albumin (BSA), using anti *E. coli* LPS antibodies (**Figure 5.13c**). The dot blot revealed a strong presence of LPS within the TolC nanodisc, minimal presence in either AcrA or AcrB and no presence in either buffer or BSA. It can be confirmed LPS is successfully solubilised by SMA.

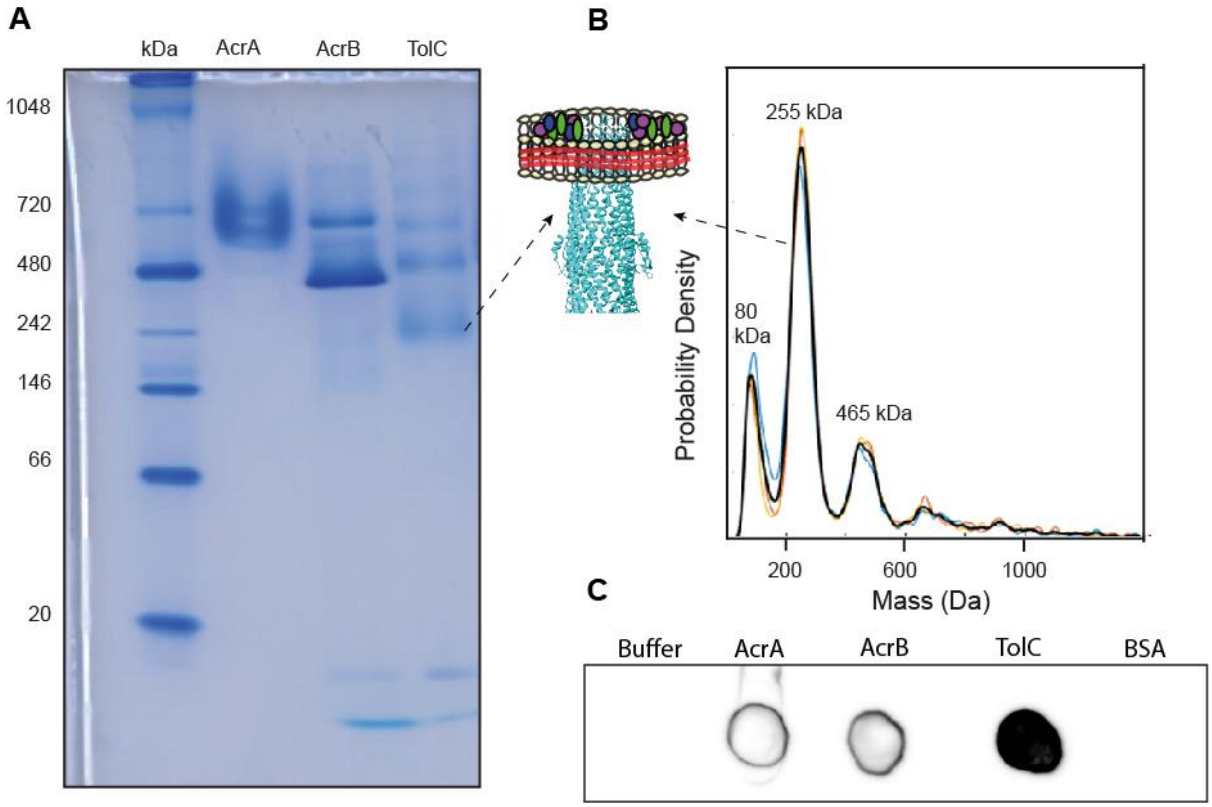


Figure 5.13. Further characterisation of TolC SMALPs. A. SMA-PAGE analysis of TolC. TolC trimer can be seen at ~242 kDa and a trimer-dimer band can be seen at ~480 kDa **B.** Mass photometry of TolC (90 nM) solubilised in SMALP nanodiscs. The TolC trimer has a mass of 255 kDa, and a trimer-dimer peak can be seen at 465 kDa. **C.** Dot blot analysis for LPS detection. Presence of LPS probed with anti *E. coli* LPS antibody, with anti-mouse secondary antibody. The buffer is 50 mM sodium phosphate, 150 mM NaCl, 10% glycerol, and AcrB and TolC are solubilised in SMALPs.

5.2.6 Attempting AcrAB-TolC complex formation using SMA-PAGE

With both AcrB and TolC purified in SMALPs, the aim was to try and assemble the complex *in vitro* and analyse it via SMA-PAGE. Daury *et al*. (2016) monitored the formation of *E. coli* AcrAB-TolC and the homologous MexAB-OprM multidrug efflux pump from *Pseudomonas aeruginosa* (*P. aeruginosa*) using native-PAGE, SEC and EM.[372] In their work, the assembled pumps were observed using the constituent proteins reconstituted into MSP nanodiscs containing POPC lipids. Therefore, the aim was to investigate whether complex formation could be observed with AcrB and TolC in native nanodiscs, which contained a native lipid environment.

Figure 5.14a shows the first attempt at complex assembly. The first trial used the soluble AcrA ($AcrA^S$) construct to avoid detergent in the reaction buffer which can compromise the integrity of SMALPs, as $AcrA^S$ still retains function.[367,396,397,487] The theoretical molecular weights of $AcrA^S$, AcrB and TolC are 40,816 Da, 114,396 Da and 52,722 Da. If the pump assembles at a 3:6:3 ratio of AcrB:AcrA:TolC, the

theoretical mass of the assembled complex would be 746,250 Da.[366,371–373] The proteins were added in a 1:1 molar ratio, except for in the AcrAB-TolC complex attempt where AcrA was in 10 molar excess. The gel shows the three input proteins on the gel; AcrA^S does not migrate through the gel as previously seen in **Figure 3.6** (**section 3.2.5.1**), AcrB appears as a trimer at ~480 kDa, and a trimer-dimer band at ~720 kDa, as previously seen in **Figure 3.6** and previous characterisation by SMA-PAGE and LILBID-MS.[63,465] The 1:1 AcrA^S:AcrB subcomplex can be observed on the gel, with the clear upward shift of the AcrB trimer band in the presence of AcrA^S (also shown and described in **section 3.2.5.1**). It is unknown why bands cannot be observed when AcrB and TolC are added together, so it is unknown if they form a complex together. In the presence of AcrA^S, it appears that the TolC trimer band is shifted upwards, also suggesting a complex with AcrA^S. The two lanes at the end represent all three constituent proteins incubated for 2 hours on ice or at room temperature. The characteristic AcrB trimer TolC trimer bands are shifted, and there is smearing present from ~800 kDa to ~600 kDa. It is possible some complex has formed, but this streaking suggests a high degree of heterogeneity, and therefore the resolution is not clear enough to observe a clear band to confirm the presence of the AcrAB-TolC complex.

Figure 5.14b shows a repeat of the experiment using a different type of native-PAGE gel. **Figure 5.14a** was performed on a NuPAGE™ 4 to 20%, Bis-Tris gel (ThermoFischer Scientific), whereas **Figure 5.14b** was performed on a Novex 4 to 20% Tris-Glycine gel (ThermoFischer Scientific). The aim was to investigate whether a change in gel type could increase the resolution of the bands observed in the SMA-PAGE experiments, to provide a higher chance of resolving any complex formation. The experiments performed by Daury *et al*. (2016) also used Tris-Glycine gels.[372] AcrA^S shows improved migration in the Tris-glycine gels, however it still only reaches ~720 kDa. The AcrB/TolC SMALPs appear as more distinct bands with a significant decrease in streaking. The AcrA^S:AcrB subcomplex is observed again with a clear shift upwards of the AcrB trimer band. When AcrA^S is added to TolC, the trimer band shifts upwards to a broad band at around ~400 kDa, suggesting there is possible subcomplexes forming, but likely with high heterogeneity. For this experiment, there was not enough sample or time (COVID-19 rules) to run AcrB with TolC in this experiment. For all samples containing SMALPs, SMA polymer can be observed as a band at the bottom of the gel (<20 kDa). Interestingly, whenever AcrA is run with another sample containing SMA, a new band appears at ~60 kDa, which is likely to represent the AcrA monomer. It is possible AcrA is able to migrate in the presence of the negatively charged SMA. The last two lanes containing all three proteins shows a similar result to **Figure 5.14a**, with large smearing at the higher molecular weight regions ~700 kDa. Again, this may contain some complex assembly, but the resolution of the gel is too low to draw any accurate conclusions.

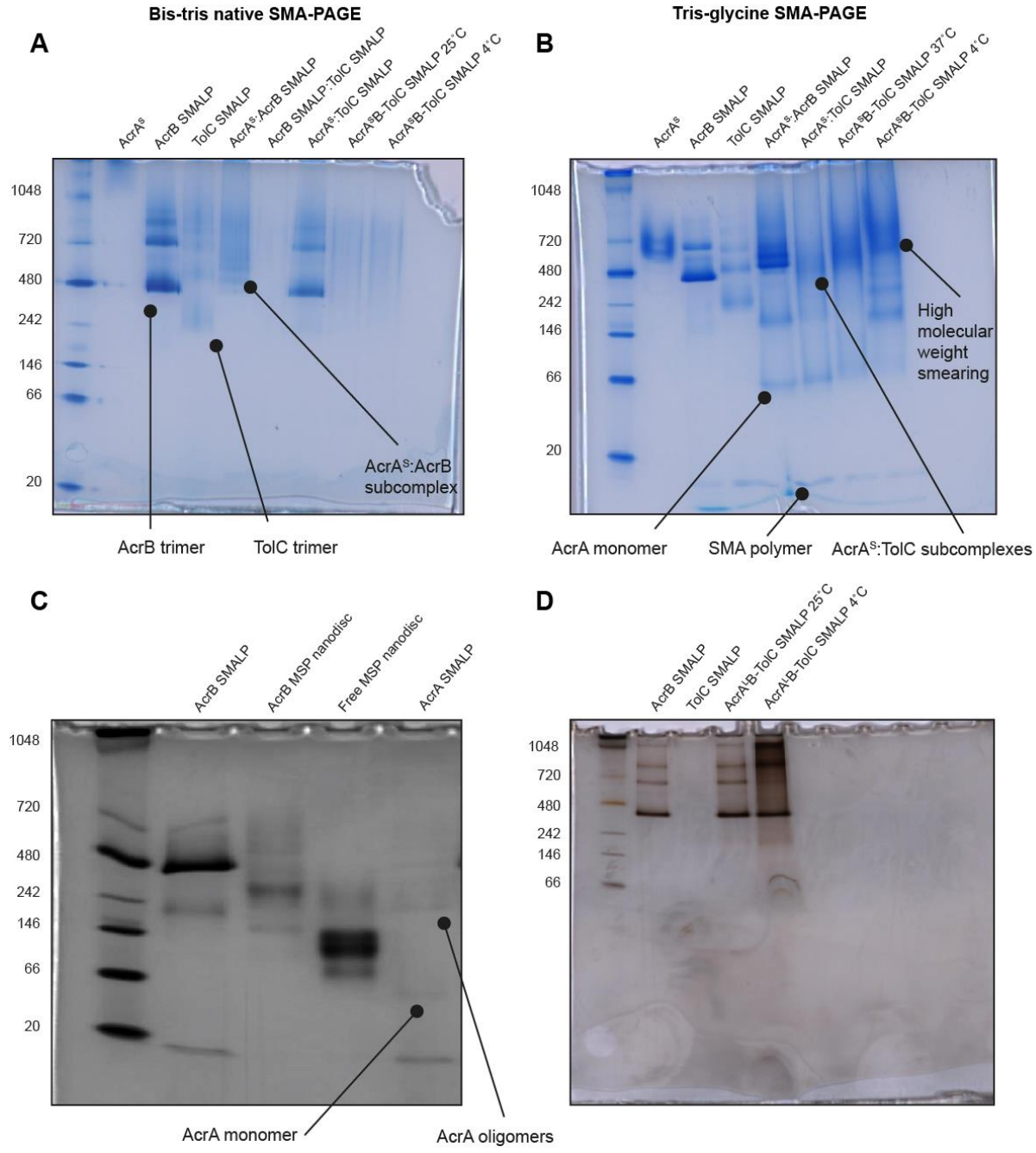


Figure 5.14. SMA-PAGE analysis of AcrAB-TolC complex assembly. A. SMA-PAGE analysis of individual and mixed components of AcrAB-TolC, using a NuPAGE™ 4-20% Bis-tris gel. AcrA^S, AcrB SMALP and TolC SMALP were used. Protein combinations mixed at a 1:1 molar ratio, except for when all three were added AcrA^S was at a 10-molar excess. AcrA^S does not migrate through the gel. AcrB SMALP presents as a trimer at ~480 kDa and TolC SMALP presents as a trimer ~242 kDa. Subcomplexes can be seen between AcrA^S and AcrB/TolC. AcrAB-TolC incubated for 2 hours at 25 ˚C or 4 ˚C but appears as a high molecular weight smear. **B.** The same experiment as **A** but repeated on a Novex 4-20% Tris-glycine gel. **C.** SMA-PAGE of AcrA^L in SMALPs. AcrA^L successfully migrates through the gel and appears as oligomers. **D.** SMA-PAGE of AcrB, TolC and AcrAB-TolC, with AcrA^L at a 10-molar excess. AcrAB-TolC incubated for 2 hours at 25 ˚C or 4 ˚C.

To see if lipidated AcrA ($AcrA^L$) purified in SMALPs (**Appendix 12**) could migrate more effectively on a SMA-PAGE than the $AcrA^S$ construct, another SMA-PAGE was run (**Figure 5.14c**). It showed that $AcrA^L$ SMALPs were able to migrate through the gel effectively. A monomer band can be observed at ~66 kDa, as well as several other oligomers up the lane. This coincides with previous results that $AcrA^L$ (**section 2.2.2.1**) forms a range of oligomers at pH 7.4. To test whether $AcrA^L$ SMALPs was more effective at complex assembly, it was combined with AcrB and TolC SMALPs at a 10-molar excess, and the resulting SMA-PAGE was analysed by silver staining (**Figure 5.14d**). Only the AcrB input was shown in this gel due to a limited amount of sample. However, **Figure 5.14d** did not show any complex formation, as the AcrB bands dominated the gel.

In conclusion, complex formation was not accurately observed by SMA-PAGE investigations. If complex formation was present in **Figure 6.14a/b** it was not possible to observe definite complex bands. This may be due to the heterogeneity of the polymer nanodisc, due to polymer and lipid content, compared to biologically derived MSP nanodiscs with monodisperse structures and uniform amphiphilic balances.[546,547] This heterogeneity leads to streaking within SMALP samples, which is made worse when multiple SMALP samples are added together. Daury *et al*. (2016) used MSP nanodiscs, which may have made resolving the native-PAGE gels easier, however at the expense of a non-native lipid environment.[372]

One other possibility is that no complex formed during these experiments. The formation of AcrAB-TolC *in vitro* is an energetically slow process; Daury *et al.* (2016) only observed 2.5% complex formation after 1 hour and 14% after 6 weeks.[372] Therefore, 2-hour incubation times may not have been long enough to study the formation of the complex. Furthermore, due to AcrB and TolC being in SMALP nanodiscs, the conditions of the reaction could not be as easily modified, due to the limitations of the nanodisc.[516] It has been well documented pH plays an important role in the binding affinities of AcrAB-TolC.[367] In fact, the TolC-AcrA interaction is dramatically weakened at pH > 6.0.[411] Lowering the pH to more acidic conditions may provide a more favourable environment for complex assembly, but due to the pH limitation of SMALPs this was not possible.[516] Furthermore, Mg^{2+} is prevalent in high concentrations within the periplasm but could not be added to the reaction mixture due to the intolerance of SMALPs to divalent cations. Future experiments would trial different reaction conditions and incubation lengths, utilising different membrane mimetic environments such as SMILPs or MSP nanodiscs.

5.2.7 Developing pull-down assays to monitor AcrAB-TolC assembly

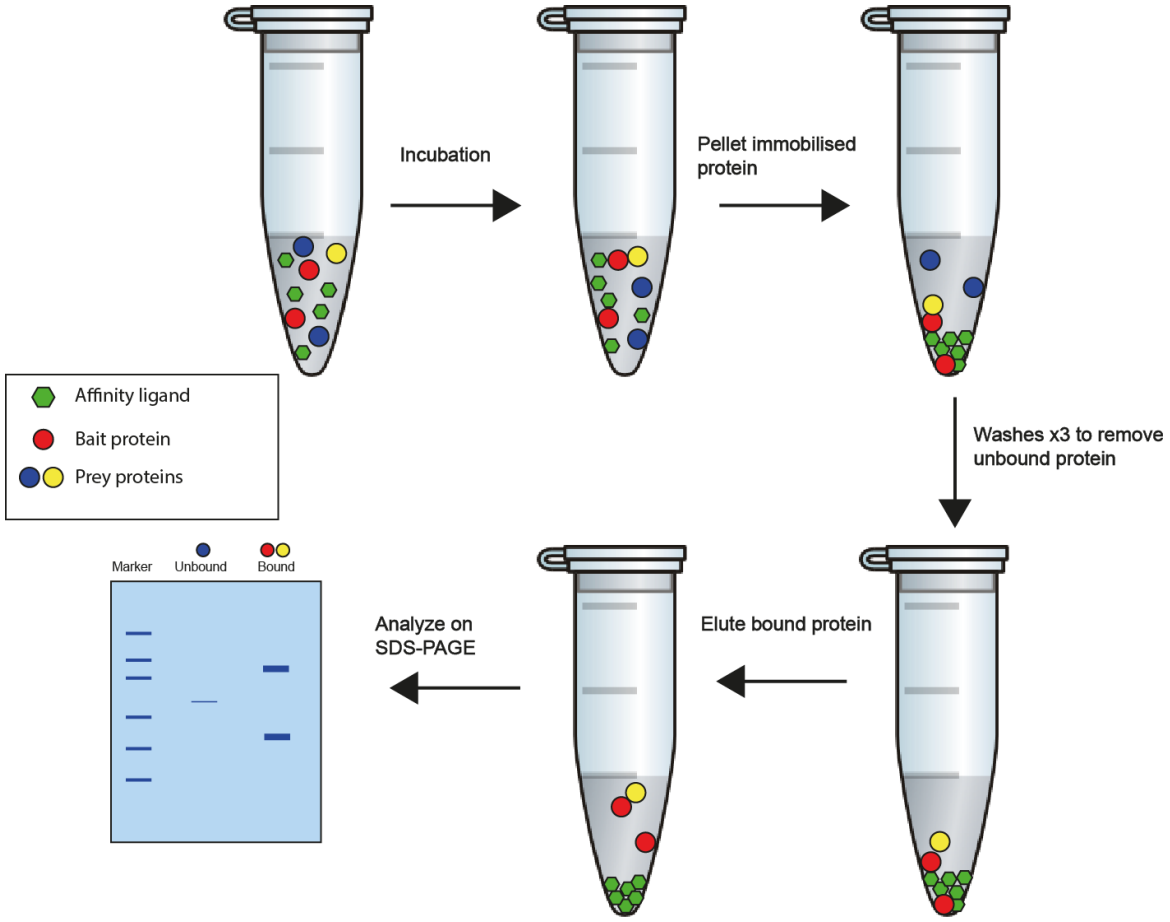


Figure 5.15. Schematic of a pull-down assay. Bait protein is immobilised on a ligand that it has specific affinity for and is then incubated with prey protein(s). The immobilised protein is isolated from the supernatant, and then washed multiple times to remove unbound prey proteins. The bait protein is eluted from its ligand and the samples analysed via SDS-PAGE, where bait and bound prey proteins can be observed.

Another *in vitro* method to monitor protein-protein interactions or complex assembly is pull-down assays. Pull-down assays tend to use a "bait" protein with an affinity tag or natural affinity to a ligand, which also binds to interacting proteins. The method involves immobilising the bait protein on an affinity ligand specific to the either the tag or the protein itself. This creates an affinity support to capture and purify other proteins ("prey") that interact with the bait. These prey proteins can be extracted from cell lysates or purified proteins. Once the prey proteins(s) have incubated with the immobilised bait protein, unbound proteins are washed away, and specifically bound proteins can be eluted with the bait using a specific elution buffer to the affinity ligand. Following the pull-down assay, samples are often analysed via SDS-PAGE to determine which proteins have bound the bait protein.

These types of assays require strict controls to demonstrate observed interactions are not artefacts caused by non-specific binding.

In this section two pull-down assays are developed to monitor interactions in the AcrAB-TolC efflux pump. The aim was to be able to monitor formation of the complex and subcomplexes, to probe assembly of the pump under differing conditions and addition of ligands (such as NSC 60339).

5.2.7.1 Biotin pull-down assay

The use of the avidin-biotin bond has been widely used for biotin pull-down assays.[531,548,549] Avidin is a tetrameric glycoprotein, originally derived from the eggs of aves, reptiles and amphibians, that binds to biotin with high specificity and affinity (K_D ~10^{-15} M) and it is the strongest known non-covalent interaction.[550] Therefore, AcrB purified with a biotinylated Avi-tag in SMALP nanodiscs can be used as a bait protein as it will bind with high affinity to avidin coated beads. This can be used to probe how $AcrA^L$ and TolC SMALPs bind to form the entire complex. The workflow of the assay was essentially a typical pulldown assay. Avidin beads were mixed with $AcrB^{Avi}$ (bait) and $AcrA^L$/TolC (prey), then separated from the supernatant using a magnetic rack or centrifugation, depending on the beads used. The beads were washed several times to remove unbound protein and the protein on the washed beads was eluted in Laemmli buffer at 95 ˚C and analysed on an SDS-PAGE.

The first decision to make was deciding which avidin beads should be used for the pull-down assay. The carbohydrate content and basic isoelectric point (pI) of avidin can result in a high amount of non-specific binding; there are alternative beads available that can be used. Streptavidin is a tetrameric binding protein from *Streptomyces avidinii* that also binds to biotin with high affinity.[550] Streptavidin has no glycosylation and a lower pI of 5, thus reducing the amount of non-specific binding it exhibits. However, it is more expensive than avidin and also contains a bacterial recognition sequence (RYD motif) which can bind cell surface receptors causing background signal in certain experiments.[551] Therefore, it was decided to use Neutravidin magnetic beads (Cytiva SpeedBeads™) for the pull-down assay, as it has a pI of 6.3, no glycosylation or RYD sequence, and it is cheaper to purchase. It was shown that $AcrB^{Avi}$ can bind to the neutravidin beads, and the optimum bead volume to use was 30 µL (**Appendix 13**).

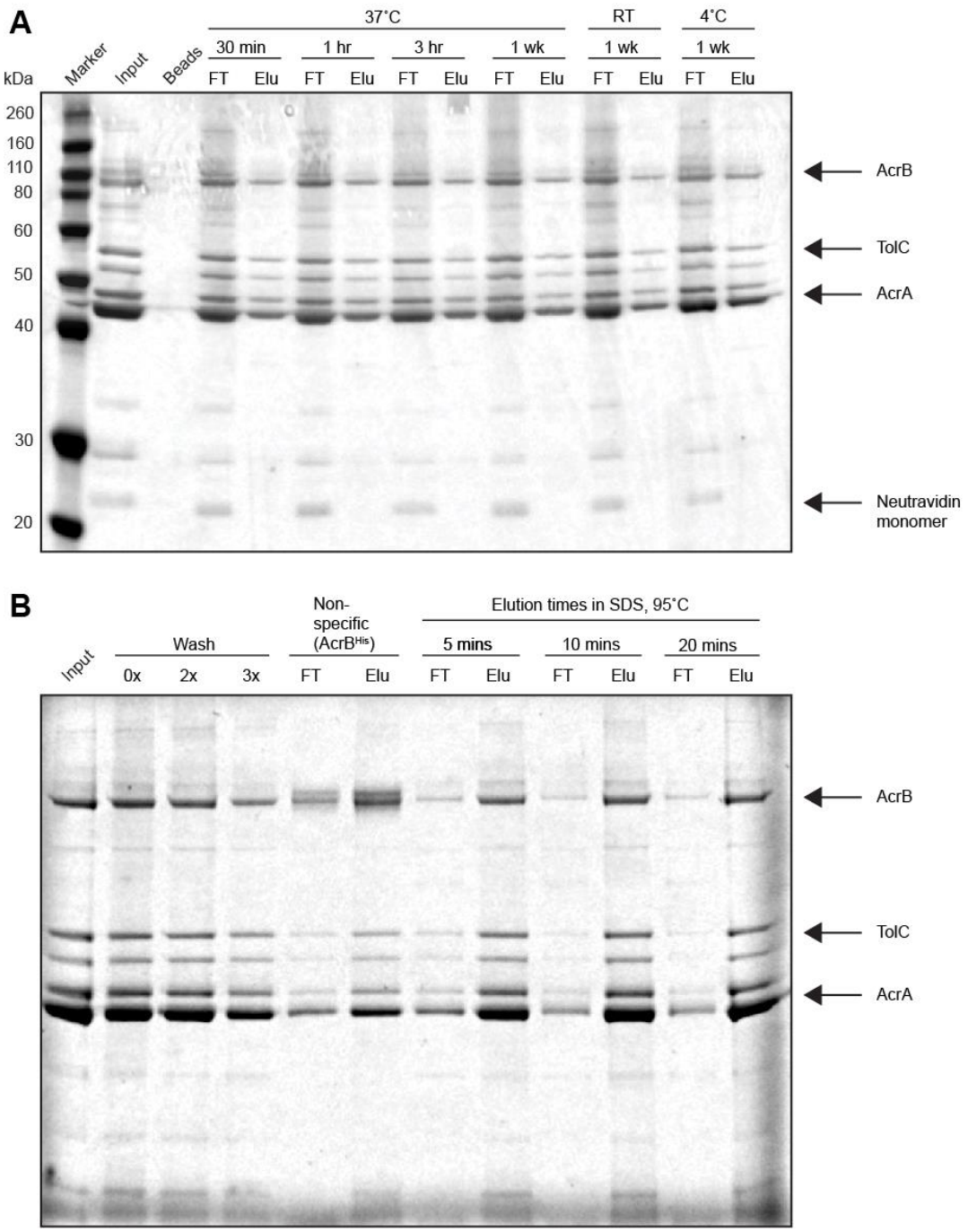


Figure 5.16. Biotin pull-down assays of AcrAB-TolC. A. SDS-PAGE showing the input proteins (AcrB^{Avi} SMALP, AcrA^{L} and TolC SMALP), the beads, and the flow-through (FT) and elution (Elu) of the pull-down assays under different incubation times and temperatures. Both AcrA^{L} and TolC are His-tagged. All three proteins are in the FT and Elu, suggesting AcrA and TolC have co-precipitated with AcrB^{Avi}. **B.** SDS-PAGE showing the effect of washes, non-specific binding with AcrB^{His} and elution times on the pull-down assay. This shows non-specific binding occurs as AcrB^{His} is present in the Elu.

Figure 5.16a shows the first pull-down assay with $AcrB^{Avi}$ and $AcrA^{L}$ and TolC. The aim was to monitor complex assembly over a variety of temperatures and timepoints. The proteins were incubated at 37 ˚C at equal concentrations, over a variety of time points ranging from 30 minutes to 1 week, and also 1 week at 4 ˚C and room temperature. The elution fractions of this pull-down assay show signals for all three proteins, at first suggesting that AcrA and TolC have co-precipitated with $AcrB^{Avi}$. The flow-through shows high signal for the three proteins as well, showing that binding of the proteins has not been saturated. Interestingly, the amount of AcrA/TolC do not appear to increase with the longer incubation times, which suggested this may be non-specific binding. Therefore, to test this, the pull-down assay was repeated with non-biotinylated AcrB ($AcrB^{His}$) which should have no affinity for the neutravidin beads, and therefore be removed in the wash; this means AcrA and TolC should not be present in the elution lane of this sample either. However, **Figure 5.16b** shows that $AcrB^{His}$ is present in the elution fraction and eluted from the neutravidin beads after three washes, along with AcrA and TolC. It cannot be concluded that any complex formation has occurred as AcrA and TolC are able to co-precipitate without the presence of a bait protein and may occur entirely through non-specific interactions. Non-specific interactions can occur due to several reasons, such as electrostatic interactions between proteins and beads or hydrophobic interactions between proteins and the resin containing immobilised neutravidin.[552]

In order to reduce non-specific binding, the beads can be blocked using a non-specific blocking agent; commonly BSA is used.[553] BSA is used to prevent non-specific binding by blocking leftover spaces over the solid resin containing the affinity ligand, in this case neutravidin. Furthermore, it is cheap and easy to use, making it an attractive option for optimising pull-down assays.[554] Therefore, a blocking step was added to the pull-down assay protocol, where the neutravidin beads were blocked in 5% BSA in phosphate buffer saline (PBS) overnight. **Figure 5.17a** shows another pull-down assay to test the binding of $AcrB^{Avi}$ SMALP, $AcrB^{His}$ SMALP and $AcrB^{His}$ DDM to BSA blocked neutravidin beads. As the gel shows, each AcrB sample was present in the elution, even after 4 washes. **Figure 5.17a** reveals AcrB exhibits non-specific binding to beads in both SMALPs and DDM micelles, albeit to a lesser extent in DDM judging by the less intense band in the elution than the flow-through. Furthermore, BSA can be seen in the elution fractions, showing it did effectively bind to the beads during the blocking stage.

As a final attempt to eliminate non-specific binding from the pull-down assays, different beads were used. Rather than magnetic neutravidin beads, monomeric avidin agarose beads (ThermoFischer Scientific) were used. Before incubation with AcrB and TolC, the beads were blocked with blocking buffer (2 mM biotin in PBS) to block irreversible biotin binding sites. The beads were then regenerated using a regeneration buffer (0.1 M glycine, pH 2.8), before being equilibrated in PBS. The aim of this procedure was to utilise a specific elution technique, the idea being to try bypassing any non-specific

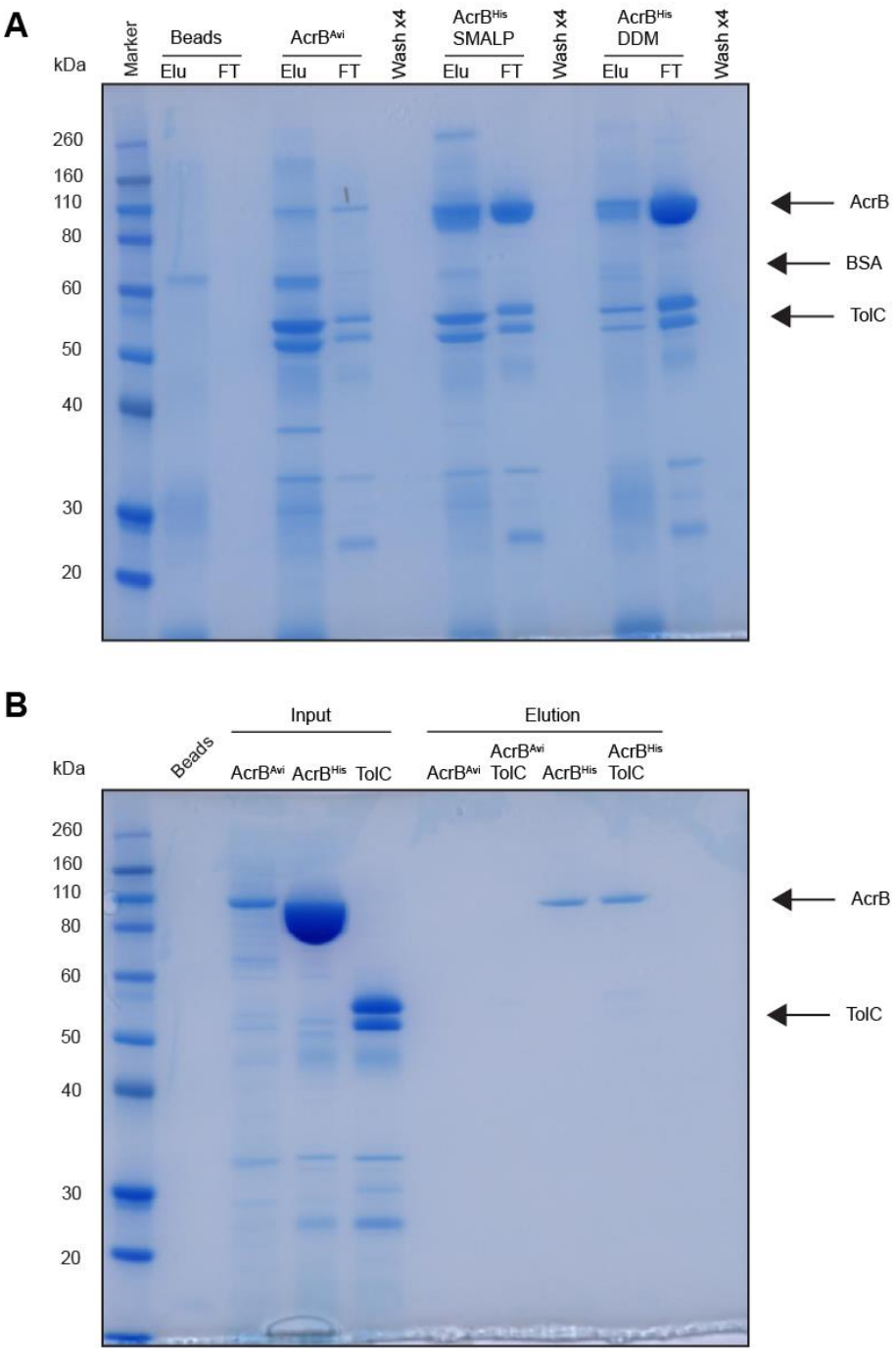


Figure 5.17. Characterising non-specific binding in biotin pull-down assays. A. SDS-PAGE showing a pull-down assay with $AcrB^{Avi}$ SMALP and $AcrB^{His}$ in SMALPs and DDM. Both $AcrB^{His}$ in SMALPs and DDM show non-specific binding to neutravidin magnetic beads blocked overnight in 5% bovine serum albumin. **B.** Pull-down assay using monomeric avidin agarose beads. Proteins eluted in 2 mM biotin. The $AcrB^{Avi}$ sample does not elute but the non-specifically bound $AcrB^{His}$ and $TolC^{His}$ does.

binding through only disrupting the biotinylated $AcrB^{Avi}$ avidin bond. Therefore, only proteins attached to $AcrB^{Avi}$ should be eluted and viewed in the elution lanes of the SDS-PAGE. The elution buffer was the same as the blocking buffer, and beads were incubated in the elution buffer for 5 minutes. The samples were then analysed on an SDS-PAGE (**Figure 5.17b**). The gel shows the presence of $AcrB^{His}$ and $TolC^{His}$ in the elution fractions, but not $AcrB^{Avi}$. This suggests the elution procedure was enough to elute non-specifically bound proteins but not biotinylated $AcrB^{Avi}$.

In conclusion, the biotin pull-down assay aimed to monitor complex assembly, and probe interactions under various different conditions. However, the excessive non-specific binding exhibited by the proteins did not allow for the characterisation of any binding interactions. **Figure 5.16** and **Figure 5.17** showed that both $AcrB^{His}$ and $TolC^{His}$ exhibited non-specific binding in SMALPs and DDM, although to a lesser degree with DDM. The assay was not developed further due to the workload and time required for optimisation. It is possible that pull-down assays using proteins in SMALP nanodiscs is troublesome, due to the hydrophobic nature of the nanodisc causing them to be more prone to non-specific binding.[474]

5.2.7.2 Peptidoglycan pull-down assay

The peptidoglycan pull-down assay has been introduced in **section 3.2.6**, to compare the differences between $AcrA^{S}$ and the AcrA pseudo-dimer ($AcrA^{SD}$) binding to peptidoglycan. The affinity of AcrA to peptidoglycan has been previously characterised; Xu *et al.* (2012) measured the affinity of AcrA and TolC to peptidoglycan through a pull-down assay.[476] Furthermore, cryo-electron tomography (cryo-ET) studies have shown that peptidoglycan interacts with AcrA and TolC at their binding interface in the periplasm.[428] Moreover, recent MD simulations have detailed interactions between peptidoglycan and AcrAB-TolC, and it may influence the assembly and stabilisation of the complex.[257] Regardless of these studies, the role of peptidoglycan on AcrA and the entire efflux pump is poorly understood. Therefore, optimisation of a pull-down assay using peptidoglycan could double as a binding assay for AcrA, whilst also probing AcrAB-TolC subunit interactions.

The procedure for the peptidoglycan assay was similar to the biotin pull-down assay and the same as stated in **section 3.2.6**. Briefly, insoluble fragments of peptidoglycan from *E. coli* (Invivogen) were incubated with protein for 30 mins at room temperature, with constant shaking to prevent the peptidoglycan settling. The peptidoglycan was then pelleted by centrifugation at 17,000 x g for 5 mins. The supernatant represents unbound sample, so an aliquot was taken. The remaining supernatant was discarded, and the peptidoglycan washed three times in 1.5 mL PBS, to remove all unbound proteins. The washed pellet was resuspended in Laemmli buffer and boiled at 95 ˚C for 5 mins, and samples were analysed via SDS-PAGE.

The first task was to develop a suitable control for the binding of AcrA^S to peptidoglycan. Two possible controls were tested in **Figure 5.18a**; insoluble peptidoglycan fragments isolated from *Staphylococcus aureus* (*S. aureus*) instead of *E. coli*, and Cytochrome C (Sigma). AcrA was incubated at different amounts ranging from 100-3 µg with a constant amount of peptidoglycan (200 µg) from either *E. coli* or *S. aureus*. The AcrA^S can sometimes present a dimer band in the SDS-PAGE, which has been discussed **in sections 2.2.1.2** and **2.2.2.1. Figure 5.18a** shows that the amount of AcrA^S binding to *E. coli* peptidoglycan increases with the amount of AcrA^S present, whereas for *S. aureus* peptidoglycan only a minimal amount of AcrA is present that does not change with concentration. This suggests AcrA^S only binds weakly to *S. aureus* peptidoglycan, whereas the binding to *E. coli* peptidoglycan appears more specific with a higher affinity. It is not surprising AcrA may have a weak affinity to peptidoglycans from other species, as they share many structural features such as the chemical composition of the glycans.[253] The variation in peptidoglycans between species mainly arise from the structure of the peptide stems and position of the interpeptide bridge, and these differences will be key for the affinities of proteins to their species-specific peptidoglycans. However, as the amount of protein captured by peptidoglycan fragments is low, peptidoglycan from alternate bacterial species is not the most effective control. Cytochrome C does not bind to the peptidoglycan, but as a small protein (12.3 kDa) it is hard to visualise on the gel, and therefore does not make for an effective control either.

Figure 5.18b shows a more effective control for AcrA^S peptidoglycan pull-down assays. Alcohol dehydrogenase (ADH) from *Saccharomyces cerevisiae* has a molecular weight of 36.8 kDa, which is close to the molecular weight of AcrA^S (40.8 kDa). Furthermore, they have similar pI values of 6.21 and 6.40 for ADH and AcrA^S respectively. Therefore, ADH better reflects AcrA^S; **Figure 5.18b** shows 25, 50, 75 µg of either AcrA^S or ADH incubated with *E. coli* peptidoglycan. AcrA^S can be seen in the bound fraction for each amount, however no ADH was observed in any of the bound fraction, suggesting it has no affinity for peptidoglycan, making it a suitable control for this reaction. Furthermore, the appearance of the gel in **Figure 5.18b** shows a reduced level of streaking compared to **Figure 5.18a**. The streaking was caused by the presence of peptidoglycan in the SDS-PAGE loaded samples, and efforts were made to remove them in future pull-down assays.

With a suitable control, the peptidoglycan pull-down assay provided a useful binding assay for AcrA^S. Aside from the data shown in **section 3.2.6** (**Figure 3.10**) which compared the affinities of AcrA$^{S/SD}$ to peptidoglycan, it can also be probed to monitor the effect of ligands on AcrA^S's interaction with peptidoglycan. Therefore, AcrA^S was incubated with peptidoglycan in the presence of 5% DMSO, 500 µM NSC 60339 (see **section 4.2.2**) and 1 mM Mg^{2+} (see **section 2.2.4**). The HDX-MS data shown in **section 4.2.2** shows AcrA^S exhibits a stabilisation across regions in four domains in the presence of NSC 60339, including peptides that contain Lys131 and 140, which have been suggested to form

hydrogen bonds with peptidoglycan by MD simualtions.[257] Furthermore, Mg^{2+} was observed to have an even wider stabilisation effect on HDX as shown by HDX-MS (**section 2.2.4.3**), and is abundant in the periplasm where the peptidoglycan:AcrA interface is.[1,3,252] A useful qualitative test was to observe if these ligands had an effect on $AcrA^S$ binding to peptidoglycan. Interestingly, **Figure 5.18c** shows that the affinity of $AcrA^S$ to peptidoglycan appears to be unchanged regardless of DMSO, NSC 60339 or Mg^{2+}, suggesting that restriction of the α-helices does not impact the binding to peptidoglycan. This experiment was repeated with ADH instead of $AcrA^S$ and no binding to peptidoglycan was observed.

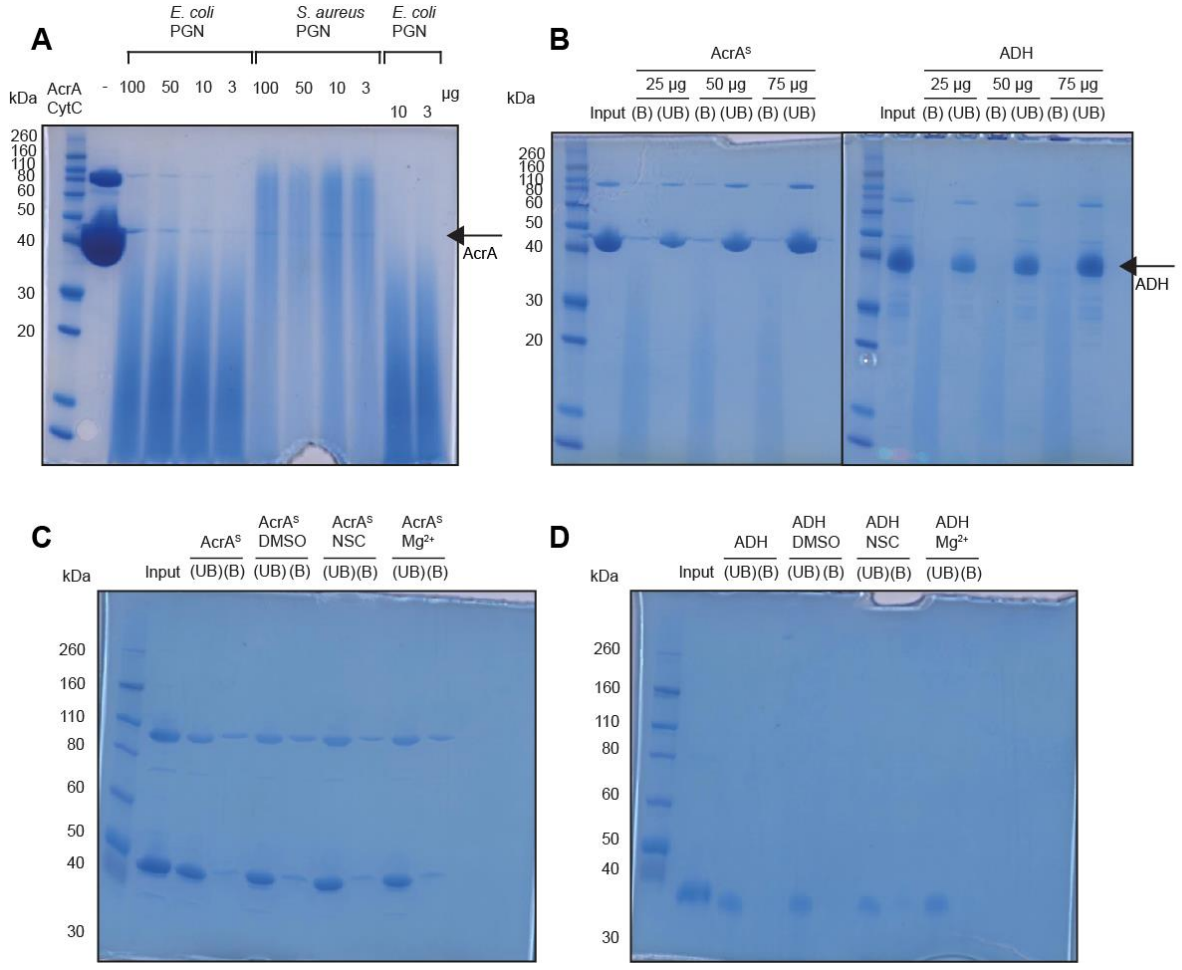


Figure 5.18. Optimisation of the peptidoglycan pull-down assay. A. Peptidoglycan pull-down assay with $AcrA^S$ at different amounts incubated with both *E. coli* peptidoglycan and *S. aureus* peptidoglycan. $AcrA^S$ appears in the bound fraction of both. Cytochrome C was trialled as a possible control but could not be visualised on the SDS-PAGE. **B.** Both $AcrA^S$ and ADH incubated in *E. coli* peptidoglycan, and the bound (B) and unbound (UB) fractions are shown on the SDS-PAGE. $AcrA^S$ is found in the bound fraction at all three amounts, but ADH is not present in the bound fraction. **C.** Peptidoglycan pull-down assay with $AcrA^S$ and peptidoglycan in the presence of 5% DMSO, 500 µM NSC 60339 and 1 mM Mg^{2+}. $AcrA^S$ affinity to peptidoglycan appears unchanged. **D.** C repeated with ADH. No ADH observed in the bound fractions.

To see if the pull-down assay could be used to monitor complex/subcomplex assembly, $AcrA^S$ was incubated with peptidoglycan and AcrB, to investigate whether the well-defined $AcrA^S$:AcrB 1:1 subcomplex could be co-precipitated (see **section 3.2.5**). It was chosen to use AcrB in DDM micelles, to bypass the non-specific binding commonly exhibited by proteins in SMALP nanodiscs. Due to the necessity of samples in detergent needing a concentration of the detergent at a concentration of 2x CMC in the sample buffer, in order to retain stability, all samples contained 0.03% DDM, ensuring the experimental conditions were kept constant.[109] The result is shown in **Figure 5.19a**; the gel shows that in the presence of DDM $AcrA^S$ does not bind to peptidoglycan. Perturbation of proteins binding to peptidoglycan due to DDM has been seen previously with MexA and OprM.[459] The maltoside head group can mimic the *N*-acetylglucosamine- *N*-acetylmuramic acid (GlcNAc-MurNAc) disaccharide structure of peptidoglycan, and can therefore disrupt the binding of proteins. It is possible DDM molecules in the buffer may outcompete $AcrA^S$ for binding to peptidoglycan. **Figure 5.19a** also shows AcrB present in the bound fraction, but AcrB has no reported affinity for peptidoglycan. This may not be a specific interaction, especially if the maltose head group of DDM molecules can associate with peptidoglycan and AcrB is in a DDM micelle. ADH exhibited no binding to peptidoglycan as expected.

This experiment was repeated with TolC in DDM instead of AcrB, to confirm the disruptive effect of DDM. So far, all pull-down assays were completed at pH 7.4 but the assay in **Figure 5.19b** was completed at pH 6.0, as the AcrA-TolC interaction is strongest at this pH.[411] The pull-down assay was completed with $AcrA^S$ in the presence of DDM in the buffer and without. As **Figure 5.19b** shows, $AcrA^S$ without DDM is able to bind to peptidoglycan but in the presence of DDM it is not, confirming the disruptive effect of DDM. TolC does not show in the bound fraction of peptidoglycan; previous work has shown TolC or homologous OMFs have an increased affinity for peptidoglycan in the presence of AcrA or its respective MFP.[459,476] However, no TolC or $AcrA^S$ is present in the bound fraction in the presence of DDM. When $AcrA^S$ and ADH were added together without DDM, $AcrA^S$ was present in the bound fraction, as was ADH. This could be due to ADH weakly binding to AcrA, or due to the ADH sample not being freshly made for this experiment, and a small fraction pelleting with the peptidoglycan.

To conclude, the pull-down assays shown in **Figure 5.19** show that DDM has a disruptive effect on $AcrA^S$ binding to peptidoglycan. Therefore, the next steps to optimise this pull-down assay to look at complex formation of the AcrAB-TolC efflux pump would be to trial different detergents that are structurally different from peptidoglycan, such as $C_{12}E_8$.[459] Nanodiscs could be trialled as well, except SMALPs does not allow for testing the binding to peptidoglycan over different pH's and SMILPs often contain many contaminants. MSP nanodiscs may be amenable to this assay, but the reconstitution of three proteins into this environment is time consuming and laborious. Nonetheless, the peptidoglycan

assay has been optimised for an AcrA^S binding assay, and different conditions and ligands can be tested to see how it affects AcrA^S affinity for peptidoglycan.

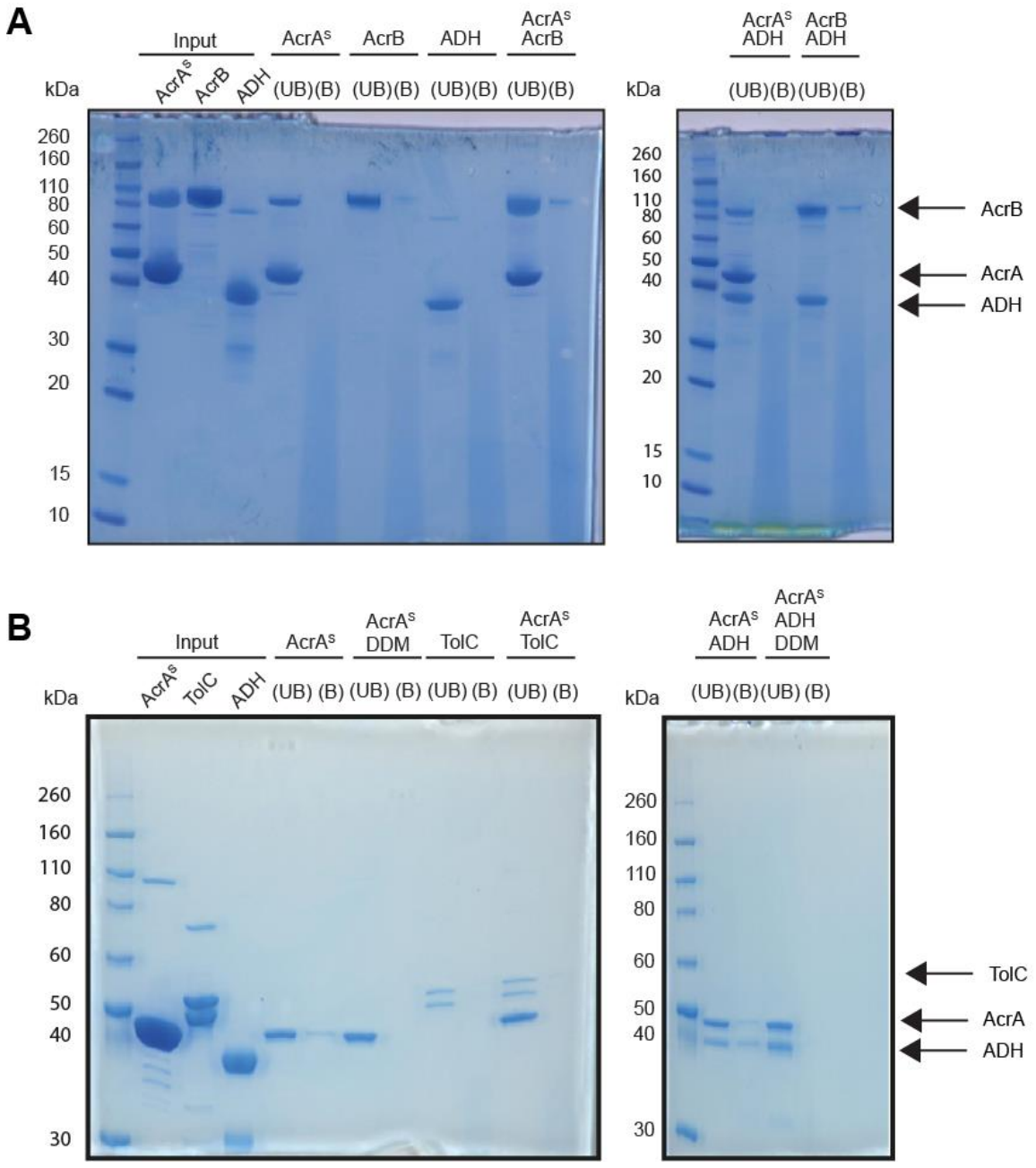


Figure 5.19. Peptidoglycan pull-down assays with AcrAB-TolC. A. Peptidoglycan pull-down assay at pH 7.4 with AcrA^S and AcrB in DDM. All buffers contained 0.03% DDM. ADH used as a control. Bound (B) and (UB) fractions shown on the SDS-PAGE. **B.** Peptidoglycan pull-down assay at pH 6.0 with AcrA^S and TolC in DDM. AcrA^S tested with and without 0.03% DDM to confirm its disruptive effects. ADH used as a control.

5.3 Conclusions

This chapter aimed to study the AcrAB-TolC multidrug efflux pump in more complex lipid environments. Membrane proteins require a mimetic hydrophobic environment to ensure solubility *in vitro*; traditionally detergent micelles have been used, but they do not provide a native lipid environment for the protein, which can affect stability, structure and function.[100,338] In this chapter, the components of the AcrAB-TolC multidrug efflux pump have been purified in different mimetic environments, including liposomes, MSP nanodiscs and polymer native nanodiscs, for downstream experiments in lipid environments.

A novel workflow by Hammerschmid *et al.* (2023) detailed an online delipidation workflow for HDX-MS investigations of membrane proteins. Therefore, as a validation of the protocol, AcrB was reconstituted to MSP nanodiscs for HDX-MS investigations with an inhibitor (MBX-3756). The workflow yielded high coverage of identified peptides and showed MBX-3756 stabilised residues in the hydrophobic trap of AcrB, akin to other MBX inhibitors. Furthermore, a novel SMALP-liposome-SMALP assay was applied to monitor the oligomeric state of AcrB in the presence of previously identified antimicrobial peptides, revealing they did not make AcrB trimers fall apart into monomers.

Another aim of this chapter was to develop pull-down assays and utilise SMA-PAGE to monitor AcrAB-TolC complex assembly and probe the effects of different ligands or conditions. Whilst the SMALP'ed proteins and some subcomplexes could be visualized using SMA-PAGE, complex formation could not be definitively observed. One reason for this may be that *in vitro* complex formation is too energetically slow for the time scales of the reactions performed in this chapter. Previous work has shown that complex formation was only 14% after 6 weeks, and for biochemical studies this is too long a time point for proteins or ligands to remain stable.[372] Another possibility is that the heterogeneity of SMALP samples makes resolution of any higher order complexes somewhat difficult using SMA-PAGE.[546,547] Different membrane protein environments such as more uniform MSP nanodiscs may be more amenable to observing complexes using native-PAGE.

Two pull-down assays were developed with the aim of monitoring complex assembly with proteins in SMALP native nanodiscs. The first was a biotin pull-down assay, utilising a biotinylated $AcrB^{Avi}$ construct and $AcrA^{His}$/$TolC^{His}$. However, this assay exhibited large amounts of non-specific binding, in the presence of different types of avidin beads and BSA blocked beads, attributed to the hydrophobic nature of the SMALP nanodisc.[474] Therefore, pull-down assays with proteins in SMALPs presents a challenge to overcome non-specific binding of bait or prey proteins. The second pull-down assay was a peptidoglycan assay, utilising the intrinsic affinity of AcrA to peptidoglycan. The assay was developed and used as an optimised binding assay for $AcrA^{S}$, and the effect of different conditions and ligands on

AcrA[5] binding to peptidoglycan could be probed. However, when trialling subcomplex assembly using this pull-down assay, DDM was found to have a disruptive effect on the binding to peptidoglycan due to its structural similarity to GlcNAc-MurNAc sugars, as was observed previously for MexA and OprM.[459] However, future work could trial other structurally different detergents and possibly polymer nanodiscs to optimise this pull-down assay further.

This chapter highlights various ways components of the AcrAB-TolC multidrug efflux pump can be studied using different *in vitro* mimetic environments that are currently available, and the importance of choosing the most suitable membrane mimetic systems for each experiment. Whilst the aim is to provide the most native-like environment each time, to ensure a more reflective environment of the protein, sometimes it is not possible; for example, SMALPs would not be suitable for the online delipidation workflow as it would precipitate at the low pH quench conditions and block the MS system. However, having multiple options in the biological toolkit allows membrane proteins, and thus efflux proteins, to be studied under various conditions.

However, the biggest challenge is the aim to study multidrug efflux systems *in situ* (**Figure 5.1**).[32] Proteins in any mimetic environment and not their native membrane is not necessarily representative of their state *in situ*.[555] HDX-MS methods have made significant progress towards an "in cell" approach. Donnarumma *et al*. (2018) were able to characterise OmpF in outer membrane vesicles naturally released by *E. coli* and reported areas of the protein that were buried or exposed based on the RFU.[556] Furthermore, Lin *et al*. (2022) have developed an *in vivo* HDX-MS protocol by overexpressing cells with the desired protein and diluting into deuterated Luria-Bertani (LB) buffer to achieve desired deuterium labelling.[557] However, further work is needed to create reliable, reproducible methods to study efflux pumps in their native membranes (see **section 6.2.4**).

5.4 Material and methods

For methods described in previous chapters, the reader will be redirected.

5.4.1 Reagents

All reagents purchased from ThermoFischer Scientific or Sigma Aldrich/Merck unless otherwise stated. The highest quality reagents were always prioritised.

5.4.2 Molecular biology

$AcrA^{S}$, $AcrA^{L}$ and $AcrA^{SD}$ are the same constructs as described in **chapters 2** and **3**.

5.4.2.1 AcrB constructs

The AcrB-His construct was already available in the Reading group. It is a pET15b plasmid containing full length AcrB with a C-terminal His_8-tag. The AcrB-Avi construct was also already available in the Reading group. It is a pET15b plasmid containing full length AcrB with a C-terminal Avi-tag. For this purification, pBirA plasmid was also used to express BirA, which biotinylated the Avi-tag for purification. All cloning and sequencing on these constructs were performed previously.

5.4.2.2 TolC construct

A pET28a plasmid was engineered to contain full length TolC with an LE linker and His_6-tag. TolC was isolated and amplified from *E. coli* genomic DNA (TolC cloning plan shown in **Appendix 14**). The polymerase chain reaction (PCR) products were purified by agarose gel electrophoresis and extracted using the NEB monarch DNA extraction kit. The pET28a vector was cut with NcoI and XhoI restriction enzymes, and again extracted using the same DNA extraction kit. Vector and insert were joined using the EcoDry Infusion HD enzyme premix. The construct was verified by DNA sequencing (Eurofins).

5.4.3 Protein expression and purification

$AcrA^{S}$, $AcrA^{L}$ and $AcrA^{SD}$ were expressed and purified as described in **chapters 2** and **3** respectively.

5.4.3.1 $AcrA^{L}$ in SMALPs

$AcrA^{L}$ construct, cell growth, harvest and membrane purification were the same as described in **section 2.4.3.1**. The membrane fraction was then solubilised in 2.5% SMA 2000 copolymer for 2 hours at room temperature. Insoluble material was removed by centrifugation at 100,000 x g for 30 minutes at 4 ˚C. 1 mL super nickel affinity resin equilibrated in Buffer B (50 mM $NaHPO_4$, pH 7.4, 300 mM NaCl,

20 mM imidazole, 10 % (v/v) glycerol) was added to the membrane and left overnight at 4 ˚C with gentle agitation. The next morning it was transferred to a gravity flow column and washed with 20 column volumes (CV) Buffer B, 10 CV Buffer B with 50 mM imidazole and eluted with 5 CV Buffer B with 500 mM imidazole. The eluted protein sample was buffer exchanged using a PD-10 desalting column (Cytiva) in Protein Buffer (50 mM $NaHPO_4$, 150 mM NaCl, 10% Glycerol). $AcrA^L$ was stored at -80 ˚C for long term storage.

5.4.3.2 AcrB

5.4.3.2.1 AcrB-Avi-tag SMALPs

pET15b AcrB-Avi and pBirA plasmids were transformed into C43(DE3)Δ*acrAB E. coli* cells. 7 mL of an overnight LB culture was added to 1 L of pre-warmed LB broth containing 100 µg/mL ampicillin, 10 µg/mL chloramphenicol and 50 µM biotin. Cells were grown at 37 ˚C for one hour, then grown at 25 ˚C until an OD_{600} of 0.6-0.8 was reached, then 1 mM isopropylthio-β-galactoside (IPTG) was added to induce protein expression. The temperature was reduced to 18 ˚C and the cells were grown for 16-18 hours. Cells were harvested by centrifugation at 4200 x g for 30 mins at 4 ˚C and washed with ice-cold PBS.

Cell pellets were immediately resuspended in Buffer A (50 mM $NaHPO_4$, pH 7.4, 300 mM NaCl) and supplemented with protease inhibitor mini tablets (1 tablet per 10 mL – as per Roche recommendations), 100 µM phenylmethylsulfonyl fluoride (PMSF), 2 µL Benzonase and 5 mM β-mercaptoethanol (β-ME). The cell suspension was then passed twice through a microfluidizer processor (Microfluidics) at 25,000 psi and 4 °C. Any insoluble material was removed by centrifugation at 20,000 x g for 30 mins at 4 °C. The membranes were then pelleted by centrifugation at 200,000 x g for 1 hour at 4 °C. Membrane pellets were then resuspended to 40 mg/mL in ice-cold Buffer B (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 10% (v/v) glycerol) supplemented with protease inhibitor mini tablets (1 tablet per 10 mL – as per Roche recommendations) and 100 µM PMSF, and homogenised with a Potter-Elvehjem Teflon pestle and glass tube.

The membrane fraction was then solubilised in 2.5% SMA 2000 copolymer for 2 hours at room temperature. Insoluble material was removed by centrifugation at 100,000 x g for 30 minutes at 4 ˚C. Monomeric avidin beads were added to a gravity flow column and washed with 10 CV Wash Buffer (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 10% glycerol) supplemented with 100 µM PMSF. In order to block non-reversible biotin binding sites, beads were blocked with 10 CV Blocking Buffer (50 mM sodium phosphate pH 7.4, 150 mM NaCl, 2 mM biotin, 10% glycerol) supplemented with 100 µM PMSF. The reversibly bound biotin was removed by washing with 30 CV Acidic Regeneration Buffer

(0.1 M glycine pH 2.8). The column was equilibrated by washing with at least 20 CV Wash Buffer. The protein sample was added to the column and to maximise binding of the biotinylated protein sample to monomeric avidin, the column was incubated on a rotating mixer at 4°C for 30 min. The column was washed with 30 CV Wash Buffer, then the biotinylated protein was eluted from the column by adding 10 CV Elution Buffer (2 mM biotin). The eluted protein sample was buffer exchanged using a PD-10 desalting column (Cytiva) in Protein Buffer (50 mM $NaHPO_4$, 150 mM NaCl, 10% Glycerol). $AcrB^{Avi}$ was stored at -80 ˚C for long term storage.

5.4.3.2.2 AcrB His-tag

The rest of the AcrB purifications use the AcrB His_8-tag construct. AcrB was purified in a host of different environments, as described below. The construct, cell growth, harvest and membrane purification were the same between all purifications. Firstly, pEt15b containing wildtype AcrB was transformed into C43(DE3)Δ*acrAB E. coli* cells. 7 mL of an overnight LB culture was added to 1 L of pre-warmed LB broth containing 100 µg/mL ampicillin. Cells were grown at 37 ˚C until an OD_{600} of 0.6-0.8 was reached, then 1 mM IPTG was added to induce protein expression. The temperature was reduced to 18 ˚C and the cells were grown for 16-18 hours. Cells were harvested by centrifugation at 4200 x g for 30 mins at 4 ˚C and washed with ice-cold PBS.

Cell pellets were immediately resuspended in Buffer A (50 mM $NaHPO_4$, pH 7.4, 300 mM NaCl) and supplemented with protease inhibitor mini tablets (1 tablet per 10 mL – as per Roche recommendations), 100 µM PMSF, 2 µL Benzonase and 5 mM β-ME. The cell suspension was then passed twice through a microfluidizer processor (Microfluidics) at 25,000 psi and 4 °C. Any insoluble material was removed by centrifugation at 20,000 x g for 30 min at 4 °C. The membranes were then pelleted by centrifugation at 200,000 x g for 1 hour at 4 °C. Membrane pellets were then resuspended to 40 mg/mL in ice-cold Buffer B (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 10% (v/v) glycerol) supplemented with protease inhibitor mini tablets (1 tablet per 10 mL – as per Roche recommendations) and 100 µM PMSF, and homogenised with a Potter-Elvehjem Teflon pestle and glass tube.

5.4.3.2.3 AcrB in DDM

Homogenised membranes were solubilised in 1% (w/v) DDM (Anatrace) for 2 hours at 4 ˚C with gentle agitation. Insoluble material was removed by centrifugation at 100,000 x g for 30 minutes at 4 ˚C. The sample was then filtered through a 0.22 µm filter (Fisher Scientific) and loaded onto a 1 mL HiTrap Nickel column in Buffer C (50 mM $NaHPO_4$, pH 7.4, 300 mM sodium chloride, 20 mM imidazole, 10 % (v/v) glycerol, 0.03% (w/v) DDM). The column was washed with 10 CVs of Buffer C, then 20 CVs of

Buffer C with 50 mM imidazole, then AcrB was eluted with Buffer C with 500 mM imidazole. The samples were buffer exchanged directly through injection onto a Superdex 16/600 GL SEC column (GE Healthcare) equilibrated in Buffer D (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 10% v/v glycerol, 0.03% (w/v) DDM). Peak fractions containing pure AcrB were pooled. AcrB was stored at -80 ˚C for long term storage.

5.4.3.2.4 AcrB in SMALPs

The membrane fraction was then solubilised in 2.5% SMA 2000 copolymer for 2 hours at room temperature. Insoluble material was removed by centrifugation at 100,000 x g for 30 minutes at 4 ˚C. 1 mL super affinity Ni^{2+} resin equilibrated in Buffer B (50 mM $NaHPO_4$, pH 7.4, 300 mM NaCl, 20 mM imidazole, 10 % (v/v) glycerol) was added to the membrane and left overnight at 4 ˚C with gentle agitation. The next morning it was transferred to a gravity flow column and washed with 20 CV Buffer B, 10 CV Buffer B with 50 mM imidazole and eluted with 5 CV Buffer B with 500 mM imidazole. The eluted protein sample was buffer exchanged using a PD-10 desalting column (Cytiva) in Protein Buffer (50 mM $NaHPO_4$, 150 mM NaCl, 10% Glycerol). AcrB was stored at -80 ˚C for long term storage.

5.4.3.2.5 AcrB in SMILPs

To solubilize the membranes, SMI co-polymer powder was added to the suspension at a final concentration of 2.5 % (w/v): this was performed using a SMI 5% (w/v) stock in buffer (Tris at pH 6.8), then adding this 1:1 to the membrane sample for a final concentration of 2.5%. The sample was incubated with gentle agitation at room temperature for 2 hours. Insoluble material was pelleted by centrifugation at 100,000 x g for 30 minutes at 4 °C.

1 mL super affinity Ni^{2+} resin (Generon) was equilibrated in Buffer B (50 mM $NaHPO_4$, pH 6.5, 150 mM NaCl, 10% glycerol) and was added to the solubilised membrane suspension and left overnight at 4 °C with gentle agitation. The beads were then transferred to a gravity-flow column and washed with 20 CVs of Buffer B with 20 mM imidazole and a further 10 CVs of Buffer B with 50 mM imidazole. The protein was then eluted with 5 CVs of Buffer B with 500 mM imidazole. This was performed at pH 6.5.

As a further clean-up stage and to buffer exchange, the eluent was injected into a Superdex 200 10/300 Increase SEC column (GE Healthcare) equilibrated in Buffer B at pH 6.5. The SEC column was brought to room temperature before running the program. Purified AcrB fractions were concentrated and aliquoted before being frozen at -80 °C.

5.4.3.2.6 AcrB in DIBMALPs

Homogenised membranes in Buffer B (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 10% Glycerol) were made sure to not contain any imidazole, as this prevented DIBMA binding to Ni^{2+} downstream. 5% (w/v) DIBMA solution was added in a 1:1 ratio to ensure a final concentration of 2.5% (w/v) DIBMA. The membranes were left to solubilise for 4 hours at 30 ˚C with gentle agitation, or overnight at 4 ˚C. Insoluble material was pelleted by centrifugation at 100,000 x g for 30 minutes at 4 °C. 1 mL super affinity Ni^{2+} resin (Generon) was equilibrated in Buffer B and was added to the solubilised membrane suspension and left overnight at 4 °C with gentle agitation. The beads were then transferred to a gravity-flow column and washed with 20 CVs of Buffer B with 20 mM imidazole and a further 10 CVs of Buffer B with 50 mM imidazole. The protein was then eluted with 5 CVs of Buffer B with 500 mM imidazole. As a further clean-up stage and to buffer exchange, the eluent was injected into a Superdex 200 10/300 Increase SEC column (GE Healthcare) equilibrated in Buffer B with 0.2 M arginine.[532] The SEC column was brought to room temperature before running the program. Purified AcrB fractions were concentrated, buffer exchanged via a PD-10 desalting column (Cytiva) to remove arginine and aliquoted before being frozen at -80.

5.4.3.2.7 AcrB in liposomes

AcrB was reconstituted into liposomes using AcrB in SMALPs. PLE lipids (Avanti) from *E. coli* were used to make liposomes. The lipids were first prepared; 100 mg of PLE was dissolved in 2 mL cyclohexane and divided into 200 µL aliquots. Each lipid was dried under N_2 to remove excess organic solvent, frozen in liquid N_2 and freeze dried for at least 4 hours. Lipids were frozen at -20 °C until further use.

Lipids were resuspended in 200 µL Protein Buffer (50 mM $NaHPO_4$, pH 7.4, 150 mM NaCl, 10% v/v glycerol), and freeze thawed in liquid N_2 six times. An aliquot of AcrB in SMALPs was added to the lipids at 10% of the lipid volume. The mixture was immediately extruded at room temperature and passed through 31 times. Sample was then incubated for 20 minutes at room temperature, 1 hour at 4 °C then kept on ice. The sample was then centrifuged at 110,000 x g for 35 minutes at 4 °C to pellet out the liposomes from the free SMA polymer. They were stored at 4 °C for 3-4 days and checked using DLS.

5.4.3.2.8 AcrB in MSP nanodiscs

AcrB was reconstituted into MSP nanodiscs using AcrB in DDM. MSP nanodiscs were made with either *E. coli* PLE lipids or POPC lipids (Avanti).

AcrB was reconstituted according to previous protocols.[372,558] Dried lipids were resuspended in Nanodisc Buffer (20 mM Tris pH 7.4, 0.1 M NaCl, 0.5 mM ethylenediaminetetraacetic acid (EDTA)) containing 200 mM sodium cholate. The tube was vortexed, heated and sonicated until the solution was clear and no lipids were on the side of the walls. MSP1E3D1 (Merck) was added to the sodium cholate solubilised lipids, and the mixture was incubated at the transition temperature of the lipid (~4 ˚C for both PLE and POPC) for at least 15 minutes. Detergent solubilised AcrB was then added to the reconstitution mixture. For PLE lipids, the final molar ratios were 60:1:0.2 lipid:MSP:AcrB with a final sodium cholate concentration of 19.2 mM. For POPC lipids, the final molar ratios were 40:1:0.5 lipid:MSP:AcrB with a final sodium cholate concentration of 16 mM.

After the addition of detergent solubilised AcrB, the mixture was incubated at 4 ˚C for up to 2 hours. The self-assembly process was initiated upon removal of detergent. Detergent was removed by the addition of 0.8 g per mL of reconstitution mixture of damp SM2 Bio-beads (Bio-Rad). The suspension was either incubated for 4 hours at room temperature on an orbital shaker, and the beads changed every 40 minutes, or left overnight at 4 ˚C. AcrB MSP nanodiscs samples were then filtered (0.22 µm) and injected onto a Superdex 200 10/300 increase GL column in Protein Buffer (50 mM $NaHPO_4$, 150 mM NaCl, 10% Glycerol). Peak fractions containing pure AcrB MSP nanodiscs were pooled. AcrB was stored at -80 ˚C for long term storage.

5.4.3.3 TolC

TolC was purified in DDM and SMALPs. The protocol for the expression and purification protocols was the same as described for AcrB, except for a few subtle differences. Firstly, the pEt28a plasmid used for TolC was kanamycin resistant, and 1 mM PMSF was used when necessary. Furthermore, for some purifications, the outer membrane was separated from the inner membrane through specific solubilisation of the inner membrane using lauryl sarcosine. The pelleted membranes were incubated for 30 minutes in 1% lauryl sarcosine, and then centrifuged at 100,000 for 30 minutes 2-3 times until the supernatant was clear. The pellet was the outer membrane.

5.4.4 Polyacrylamide gel electrophoresis (PAGE) and Western blots

SDS-PAGE and SMA-PAGE were completed as previously described.

5.4.5 Lipopolysaccharide dot blot

20 µL of Protein Buffer, $AcrA^L$, AcrB SMALP, TolC SMALP and BSA was dotted on a piece of nitrocellulose membrane. It was blocked in 5% milk in PBS-Tween for 1 hour at room temperature, then incubated with anti-*E. coli* LPS antibody (Abcam) for 1 hour. The membrane was then washed

every 15 minutes with PBS-Tween 4 times. It was then incubated with a secondary anti-mouse antibody (Abcam) for 1 hour and the wash steps repeated. Membranes were developed with 1 mL Amersham™ ECL Select™ and imaged using an A1600 Imager (GE Healthcare).

5.4.6 Dynamic light scattering (DLS)

DLS experiments for liposomes and MSP nanodiscs were carried out using a Litesizer 500 DLS (Anton Paar) in particle size mode, with 10 µL protein sample diluted in 990 µL water.

5.4.7 Mass photometry

Mass photometry experiments completed as previously described.

5.4.8 Hydrogen deuterium exchange mass spectrometry

HDX-MS experiments and ligand preparation were mostly the same as previously described, except for an extended set-up which was required for online delipidation.[4] This involved an additional valve positioned between the injection and trapping valves (**Figure 5.20**). This was equipped with an in-house packed (according to Hammerschmid *et al*. (2023)) ZrO_2 phospholipid trap column which was kept on ice.[4] The delipidation column was cleaned with 3% NH_4OH in methanol and re-equilibrated in solvent A during the subsequent wash run. This protocol required a quench buffer consisting of 500 mM glycine-HCl at pH 2.35. Furthermore, for AcrB, an extended 9.0-minute linear gradient was applied from 8-40% solvent B at 40 µL/min. The analysis was the same as previous workflows.

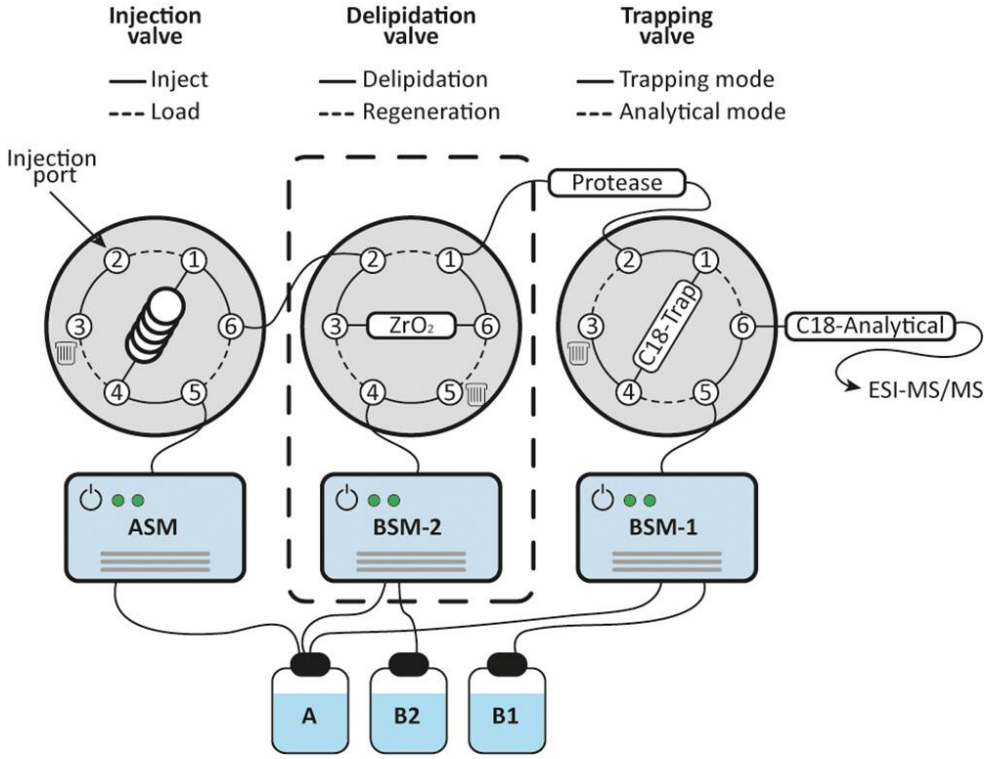


Figure 5.20. Schematic of the automated delipidation workflow. The dashed box represents the additional valve required for this workflow. It is equipped with a phospholipid trap column and operated by an additional binary solvent manager (BSM) to provide independence from standard HDX workflow solvents. In this workflow, the sample passes through the delipidation column where the lipids are retained, before passing over the protease column and following the standard bottom-up workflow. After delipidation, the phospholipid trap column can be cleaned and regenerated for the next sample by the BSM-2. Taken from Hammerschmid et al. (2023).[4]

5.4.9 SMALP-Liposome-SMALP assay

AcrB SMALPs were reconstituted into liposomes containing PLE lipids. AcrB was incubated at room temperature with 2 µM antimicrobial peptides (0.4% DMSO) for 1 hour. Samples were then resolubilised in 2.5% SMA, before analysis via SMA-PAGE.

5.4.10 Biotin pull down assay

Biotinylated AcrB, and constituent His-tagged proteins AcrA and TolC, were incubated under the desired experimental conditions (e.g. pH, temperature). 30 µL of monomeric avidin agarose (ThermoFischer Scientific) or magnetic neutravidin speed beads (Cytiva), referred to as resin, were equilibrated in PBS before blocking. For some permutations of the assay, the resin was blocked with 2 mM biotin to block non reversible binding sites. Biotin was removed from reversible binding sites

using 12 mL Regeneration Buffer (0.1 M glycine pH 2.8) per 2 mL of resin. Resin was then washed 3 times in PBS to remove any excess Regeneration Buffer. Other permutations of the assay saw the resin blocked in 5% BSA in PBS overnight and re-equilibrated in PBS. Samples were then added to the resin, and incubated for 1 hour, vortexing several times. The resin was washed with 1 mL PBS, resuspended by centrifugation for 2-3 minutes at 2500 x g or separated by a MagRack™ (Cytiva), and this was repeated 4 times. The sample was then eluted with 2 mM biotin or by boiling in SDS for 10 minutes, and the elute and flow through analysed via SDS-PAGE.

5.4.11 Peptidoglycan pull down assay

Peptidoglycan pull down assays completed as previously described.

Chapter 6: Conclusions and future directions

A note on the COVID-19 pandemic

The 2020 COVID-19 pandemic led to a university-wide shutdown of the laboratory for three months, in which no data was able to be collected. Furthermore, upon re-opening, occupancy limits within the department lasted most of 2020, also affecting the amount of time in the laboratory. Therefore, it was not possible to expand some of the avenues in this thesis further due to a lack of time.

6.1 Summary of results

Bacterial multidrug resistance is a threat to global healthcare, and it continues to spread at rapid rates. Multidrug resistance refers to the ability of bacterial pathogens to survive lethal doses from many structurally diverse classes of antibiotics.[329,425] A major driver of multidrug resistance is the activity of efflux pumps. The AcrAB-TolC multidrug efflux pump is native to *Escherichia coli* (*E. coli*) and a member of the resistance nodulation and cell division (RND) superfamily.[223,224] It consists of the RND inner membrane protein AcrB, membrane fusion protein (MFP) AcrA and outer membrane factor (OMF) protein TolC, and it is able to effuse a wide variety of chemically diverse compounds such as antibiotics, fatty acids, dyes and detergents, to confer drug resistance.[329] The AcrAB-TolC pump is well characterised and is prototypical of homologs across other ESKAPE bacteria, thus the research in this thesis confers effectively to other systems.[211] Studying the structural dynamics of the efflux pump to elucidate information on its function, stability, assembly and inhibition is key to finding ways to combat these molecular machines, in the fight against antibiotic resistance. Structural mass spectrometry (MS) techniques can be deployed to study the structural dynamics of proteins. In this thesis, hydrogen deuterium exchange mass spectrometry (HDX-MS) and native MS were utilised to answer a number of biological questions. HDX-MS reports on protein dynamics over time, whereas native MS can report on protein oligomeric state under different conditions and protein-ligand interactions.[33,97] Therefore, structural MS techniques supplemented with an array of different biophysical and biochemical techniques, were applied to study the structural dynamics of the AcrAB-TolC multidrug efflux pump.

AcrA has been extensively characterised using structural MS techniques (**chapter 2**). Native MS revealed the AcrA lipidation ($AcrA^L$) led to oligomerisation of $AcrA^L$, however this was somewhat tamed at pH 6.0 as only monomers and dimers were observed, whereas oligomers up to pentamers were observed at pH 7.4. This may be an important feature in pump assembly, as the functional unit of AcrA is a dimer, and the often acidic conditions of the periplasm may help the functional dimer form and not lead to higher order oligomers.[366,410] Then, the structural dynamics of the soluble AcrA construct ($AcrA^S$) was investigated by HDX-MS, which revealed AcrA was a folded protein with

secondary structure but had unstructured areas; this was also suggested by the multiple charge state distributions (CSDs) exhibited by AcrA in native MS. AcrA resides in the periplasm, yet the role of the periplasmic environment, which can differ significantly from the cytosol, on AcrA dynamics and function is understudied.[2,3,252] It was shown that Mg^{2+} is able to temper the increased dynamics exhibited by AcrA at pH 6.0, but has little effect at pH 7.4. This suggests a previously unknown role of Mg^{2+} ions in the regulation of AcrA, that could have implications for AcrAB-TolC assembly over a range of different pH conditions the periplasm can exhibit. Furthermore, a region in the αβ-barrel domain of AcrA was highlighted by HDX-MS and prediction software as a possible Mg^{2+} binding site, but further work is needed to draw specific conclusions (**section 6.2.1**).

Previous work has characterised a dimer as the functional unit of MFPs, however there is no structural dynamics information available on an AcrA dimer due to the difficulty of isolating them *in vitro* without crosslinking.[367,397,410,411] Therefore, a soluble pseudo-dimer AcrA construct ($AcrA^{SD}$) was designed to infer biological information on the AcrA functional dimer (**chapter 3**). Native MS confirmed $AcrA^{SD}$ was a homogenous pseudo-dimer, and still contained markers of intrinsic disorder in its native mass spectra. Differential HDX-MS (ΔHDX) between $AcrA^{SD}$ – $AcrA^{S}$ revealed $AcrA^{SD}$ had unique dynamics, with extensive protection observed in the α-helices, suggesting that dimerization may help stabilise the α-helices when binding TolC and during the 'opening' mechanism. The αβ-barrel and membrane proximal (MP) domains also saw areas of protection compared to $AcrA^{S}$. The study of pseudo-dimerization structural dynamics suggests dimerization may stabilize AcrA, which is a highly dynamic protein, possibly priming the protomers for interactions with AcrB and TolC. Furthermore, differences between the binding of the AcrA constructs to both AcrB and peptidoglycan was investigated. Whilst no differences were observed between $AcrA^{SD}$ and $AcrA^{S}$ when binding to peptidoglycan, suggesting that a dimer interface is not necessary for this interaction, the $AcrA^{SD}$ construct appeared to have a higher propensity to form higher order, heterogenous AcrA:AcrB complexes compared to the 1:1 complex seen for $AcrA^{S}$. This work highlights the importance of the AcrA dimer, and using the pseudo-dimeric construct it is possible to infer that AcrA dimerization plays an important role in stabilising AcrA dynamics and promotes higher order binding to AcrB. Understanding the role of the AcrA functional dimer is essential to shed light on how the AcrAB-TolC complex assembles in the periplasm; understanding these mechanisms could provide a new avenue to inhibit AcrA and corresponding MFPs.

Efflux pump inhibitors (EPIs) have the potential to be a useful adjunctive therapy in the case of multidrug resistant infections, to 'revive' the activities of available antibiotics.[403,478] Most EPIs to-date have focused on inhibiting AcrB, however none have made clinical trials due to toxicity issues and the promiscuous nature of AcrB to transport its inhibitors.[479–481] Therefore, new approaches are needed

to develop methods of inhibiting efflux pumps in the fight against antibiotic resistance. Therefore, the work in **chapter 4** revealed the first mechanism of action against an AcrA inhibitor, NSC 60339. Previous work and a combination of HDX-MS, molecular dynamics (MD) simulations and cellular inhibition assays suggests that NSC 60039 binds in a cleft between the lipoyl and αβ-barrel domains of AcrA.[403,489] HDX-MS revealed extensive protection across regions in all four domains of AcrA in the presence of NSC 60339, and this was supported by MD simulations of NSC 60339 docked in the proposed binding site. Furthermore, it was shown that NSC 60339 inhibits the $AcrA^{SD}$ construct in the same way. This work proposed NSC 60339 acts as a molecular wedge within the binding cleft, significantly restricting AcrA structural dynamics, which could have implications for the conformational transitions required during the functional rotation of the AcrAB-TolC efflux pump.[425] NSC 60339 may result in AcrA losing its ability to accommodate changes across the periplasm and communicate conformational signals between AcrB and TolC. This could disrupt the interactions between AcrA and AcrB/TolC, cause a leaky pump as AcrA may not be able to maintain a sealed channel during AcrB functional rotation, or TolC may not be opened efficiently. Moreover, cellular inhibition assays revealed a novel potential binding site between the αβ-barrel and MP domains, suggesting that targeting the flexible linkers between either the lipoyl and αβ-barrel domains or αβ-barrel and MP domains could be a promising approach to inhibit efflux. This work lays the foundation for the future production of inhibitors with optimised pharmacodynamics and pharmacokinetics targeted to flexible linkers of AcrA.

Studying the AcrAB-TolC multidrug efflux pump in lipid environments is key to understanding how they function within native membranes. Therefore, **chapter 5** highlighted the various ways components of the AcrAB-TolC multidrug efflux pump can be studied using different *in vitro* mimetic environments. A novel HDX-MS online delipidation workflow developed by Hammerschmid *et al*. (2023) was used to observe the effects of MBX-3756 on AcrB structural dynamics, with AcrB purified in membrane scaffold protein (MSP) nanodiscs.[4] Using this protocol, it was possible to observe the effects of MBX-3756 on AcrB, with residues in the hydrophobic trap of AcrB showing significant protection across all time points, likely due to drug interactions. Furthermore, it was shown that antimicrobial peptides targeting AcrB did not make the trimer fall apart into monomers, by developing and applying a novel styrene maleic acid lipid particle (SMALP)-liposome-SMALP assay. SMA-polyacrylamide gel electrophoresis (PAGE) analysis was utilised to try and observe AcrAB-TolC complex assembly, but complex formation could not be definitively determined, either due to the slow energetics of assembly *in vitro* or the heterogeneity of SMALPs, due to the polymer and lipid content, leading to a lack of resolution on the stained gels. Two pull-down assays were also developed to try to capture complex assembly of the AcrAB-TolC efflux pump. The first was a biotin pull-down assay using

a biotinylated $AcrB^{Avi}$ construct and aimed at forming the complex with the proteins in SMALPs. However, due to excessive non-specific binding attributed to the hydrophobicity of the SMALP nanodisc, this pull-down assay was abandoned. The second pull-down assay was a peptidoglycan pull-down assay, utilising AcrA's natural affinity to it. To bypass non-specific binding of SMALPs, proteins were trialled in DDM, but due to the presence of DDM disrupting protein binding to peptidoglycan, this pull-down assay was optimised as a binding assay for AcrA under a range of conditions instead. Overall, this work studied components of AcrAB-TolC in different membrane mimetic environments and highlighted possible complications of working with them. However, it was still possible to elucidate information on two types of inhibitor binding to AcrB.

6.2 Signifiance and future directions

6.2.1 A general role for pH and Mg^{2+} in the function of MFPs?

It is well established that the conformational flexibility of MFPs is critical to their function in their respective efflux pump systems.[410] As periplasmic proteins, they exhibit a vastly different environment to the cytosol; on average it is ~1.7 pH units more acidic and contains 7.56 times higher concentration of Mg^{2+} ions. However, they are also subject to rapid changes in pH conditions, due to the periplasm's proximity to the external medium, as bacteria move through different environments. For example, there will be varying acidity and Mg^{2+} concentrations through *E. coli's* enteric journey through the human gut, (e.g. stomach pH 1.4-4.0 and intestines pH 4.0-7.0) and within the microenvironment of macrophage phagosomes.[559] MFPs are therefore expected to be able to function in the dynamic periplasmic environment. **Chapter 2** revealed that the dynamics of AcrA are regulated by cross play between pH and Mg^{2+}; weakly acidic conditions increased the backbone dynamics of AcrA, yet Mg^{2+} was able rectify this increase in dynamics. It is likely that pH, Mg^{2+} and His285 regulate the hydrogen bonding network of AcrA and ensure functional conformations across the dynamic periplasmic environment.[1,2]

This work provides a detailed look into the role of the periplasmic environment on the structural dynamics of an MFP and may uncover a more general role for pH and cation binding in the function of MFPs. AcrA homologs such as AcrE and MdtE from *E. coli* and MexA from *Pseudomonas aeruginosa* (*P. aeruginosa*) were also suggested to bind Mg^{2+} according to the *MeBiPred* software (**Table 2.5**).[451] Interestingly, less closely related MFPs, such as MdtA (*E. coli*) and MdsA (*Salmonella enterica*) were not suggested to bind Mg^{2+} but were suggested to bind different cations. Furthermore, previous work has already shown the heavy metal efflux (HME)-RND MFP ZneB binds to Zn^{2+}, which leads to a conformational change to a more compact state, and CusB binds to Cu^{+}/Ag^{+}.[404,449] For these two

examples, crystal structures revealed the αβ-barrel and MP domains were involved in ligand binding; interestingly the folding of the β-barrel is similar between ZneB, CusB and MexA (except for the region facing the MP domain), and the αβ-barrel domain of AcrA exhibited protection in the presence of Mg^{2+} at pH 6.0 compared to pH 7.4 (**Figure 2.17**). Therefore, there is an increasing amount of evidence that pH and cations play a role in the function of MFPs during efflux. It may be the case that AcrA/MexA type MFPs are interacting with Mg^{2+}, yet more distantly related MFPs interact with different cations.

Nonetheless, more investigation is needed to elucidate the exact relationship between MFPs and pH or Mg^{2+}, and the role of this relationship in the function of multidrug efflux pumps. To see if other MFPs exhibit increased dynamics at weakly acidic pH's, HDX-MS characterisation of MexA could be performed at both pH 7.4 and 6.0. Whilst MD simulations on MexA have confirmed it is a flexible MFP analogous to previous results of AcrA, there is a lack of structural biology information available as there was with AcrA before this work.[398,458] Furthermore, to confirm the importance of His285 as a molecular switch in AcrA, HDX-MS experiments could be performed at both pH 6.0 and 7.4 with an AcrA H285A mutant, to measure if there are still differences in the dynamics between pH's. Furthermore, cellular accumulation assays could be performed to characterise the effect of the AcrA H285A mutant on the rate of efflux, to reveal further information on the importance and function of this residue. Lastly, regarding the role of Mg^{2+}, the most important experiment to do would be to obtain a crystal structure of AcrA with Mg^{2+} and reveal the binding site(s). The experiments performed in **chapter 2** could be repeated on AcrA like homologs such as AcrE and MexA to see if similar results are observed. This would have implications for MFPs in general and could confirm a wider role of divalent cations in the function and regulation of efflux pumps. Further work could also investigate whether different divalent cations such as Ca^{2+} can have the same effect on AcrA, or if it is Mg^{2+} specific.

6.2.2 Investigating the functional dimer unit of MFPs

Chapter 3 utilised a pseudo-dimer construct to infer biological information on the dimerization of AcrA, which exists as a functional dimer unit to form a trimer of dimers in the assembled complex.[366,367,410,457–460] Whilst this is not a perfect reflection on *in situ* dimerization as it is a rationally designed construct, useful information can be determined to provide insights on the behaviour of an AcrA dimer. Therefore, HDX-MS investigations in **chapter 3** revealed the AcrA pseudo-dimer has unique structural dynamics compared to the monomer. This provides structural biology evidence that compliments MD simulations to shed light on the role of the functional dimer unit; MD simulations of MexA dimers show they are stable in aqueous solution in the absence of protein partners, whilst retaining plasticity of the peripheral domains of the protein.[458] The HDX-MS data from this chapter

show a protection of the AcrA^{SD} α-helical domain, but there is still a degree of deuterium uptake suggesting there is still conformational plasticity. Therefore, the functional dimer unit likely provides a scaffold for the MFP to position its α-helices, whilst retaining conformational flexibility required for the efflux cycle.

It was also demonstrated in **chapter 3** that pseudo-dimerization had no observed effect on AcrA's affinity to peptidoglycan. Whilst a lot of work has been done to characterise the functional mechanisms of multidrug efflux pumps, only until recently has the role of peptidoglycan been considered. Recent work has demonstrated the affinity of MFPs AcrA and MexA to peptidoglycan, and revealed peptidoglycan interacts with AcrA and TolC at the AcrA-TolC interface.[38,257,428,459] Recent modelling by Gumbart *et al*. (2021) suggest peptidoglycan stabilises the α-helical hairpin of AcrA and preserves the AcrAB subcomplex in an assembly competent state before binding to TolC.[257] The results from **chapter 3** suggest the dimer unit does not enhance AcrA's affinity for peptidoglycan, providing further evidence this interaction is due to hydrogen bonding with Lys131 and Lys140.

The work from **chapter 3** can be built upon and expanded further. Additional characterisation of the AcrA^{SD} construct would provide more insights into AcrA; calculating binding affinities between AcrA^{SD} and AcrB/TolC using surface plasmon resonance (SPR), at a range of pH's, and comparing those to previous values calculated for AcrA^{S} would help quantify the differences in binding between the two constructs.[367] Moreover, a lipidated dimer could be investigated by native MS, to see if it has the same propensity to oligomerise as AcrA^{L}, and to see if it more readily forms a trimer of pseudo-dimers. Lastly, cryo-electron microscopy (cryo-EM) structures of the AcrA^{SD} construct would be useful to see if it folds akin to the function dimer unit of AcrA seen in previous cryo-EM structures.[366]

One limitation of HDX-MS experiments on the AcrA^{SD} construct is that differences between the two protomers cannot be observed. Since HDX-MS is an averaging technique, as the two protomers share the same sequence, potential differences in the dynamics between the two are averaged out. However, previous work has shown that the two protomers in the functional dimers of MFPs function differently, by the way they bind to the respective binding partners.[410] Therefore, it would be interesting to study TriAB using HDX-MS. TriABC-OpmH is an RND efflux pump from *P. aeruginosa*, yet requires two different MFPs, TriA and TriB, that play non-equivalent roles in the pumps function; TriA is responsible for the recruitment of OMF OpmH whilst TriB is responsible for the stimulation of the transporter TriC.[412] Therefore, HDX-MS investigations of TriA and TriB separately and a TriAB pseudo-dimer may provide unique insights into the differences in structural dynamics of an MFP functional dimer.

6.2.3 Developing the next generation of efflux pump inhibitors

Work presented in **chapter 4** proposed a mechanism of inhibition for an efflux EPI targeting AcrA. NSC 60339 appears to bind as a molecular wedge in a cleft between the lipoyl and αβ-barrel domains, causing a restriction of AcrA's dynamics across all four domains. Furthermore, cellular accumulation assays revealed a novel druggable site between the αβ-barrel and MP domains. This has laid the foundation for a new generation of EPIs targeted to AcrA, and hopefully bypassing the promiscuity of AcrB. Furthermore, this may provide a new avenue for the inhibition of homologous efflux pumps, by targeting the flexible linkers of MFPs. For the next generation of improved EPIs, future work could aim to achieve a cryo-electron microscopy (cryo-EM) structure of AcrA with NSC 60339 bound. This would allow for the mapping of the interactions between NSC 60339 functional groups and the amino acid residues in the cleft. This would therefore allow new drugs to be designed to strengthen certain interactions and add additional functional groups to the molecule to improve its efficacy as a drug; one key aspect of any EPI against Gram-negative bacteria would be the need to permeate the outer membrane, which NSC 60339 does not do efficiently for *E. coli*.[403] Furthermore, drug design targeting the flexible linkers of other MFPs such as MexA and MdtE should be tried to see if is effective across a range of different multidrug efflux pumps. Overall, the work in **chapter 4** shows that EPIs targeted to AcrA that restrict its structural dynamics, could be a promising avenue in the fight against bacterial multidrug resistance.

6.2.4 Studying multidrug efflux pumps within lipid environments

RND multidrug efflux pumps are tripartite complexes that span the entire cell envelope of Gram-negative bacteria, and therefore consist of inner and outer integral membrane proteins.[366] **Chapter 5** highlights the various ways membrane mimetic environments can be used to study AcrAB-TolC, with the aim to study these membrane protein systems in the presence of a lipid environment. As lipids can be essential for the function of membrane proteins, aiming for the most native membrane environment possible will provide a more reflective view of the protein *in situ*, and thus more accurate results.[100,338] The work in this chapter compliments work in the field and a general trend of analysing multidrug efflux pumps in more native membrane environments *in vitro*; Daury *et al*. (2016) monitored the assembly of AcrAB-TolC and MexAB-OprM in MSP nanodiscs, and Parmer *et al*. (2018) reported the structure of AcrB in SMALPs to a sub-nm resolution using cryo-EM.[372,560]

However, these mimetic environments still do not necessarily reflect the conditions exhibited in the natural membrane. Therefore, the development of methods to study AcrAB-TolC *in situ* would allow greater insights into how these systems are functioning in cell. HDX-MS has the potential to study

these systems *in vivo*, with recent work already demonstrating how cells can be diluted in deuterated Luria-Bertani (LB) buffer to achieve labelling of transporters.[557] With optimised protocols to reliably measure deuterium uptake 'in cell', analysing the dynamics of the AcrAB-TolC complex during the efflux of substrates would reveal critical information on how the three proteins coordinate conformational movements throughout the efflux mechanism. Being able to study such protein movements directly 'in cell' would revolutionise the way in which membrane protein systems are studied.

6.3 Final remarks

In summary, this thesis highlights how structural MS can be applied to multidrug efflux pumps to elucidate different information regarding its function and inhibition. Furthermore, it shows the power of structural MS when used in combination with supplementary biophysical and biochemical techniques. Specifically, this research makes a significant contribution to the understanding of AcrA's function and dynamics across the different conditions exhibited by the periplasm, and perhaps uncovers a wider role for pH and divalent cations in the function of MFPs. It has provided the first insight into the benefits of dimerization on AcrA's dynamics and its binding to AcrB, through the use of a novel pseudo-dimer construct. This thesis also reveals the first mechanism of inhibition for an AcrA inhibitor, showing that NSC 60339 functions as a molecular wedge between the lipoyl and αβ-barrel domain to reduce AcrA's structural dynamics across all four domains. Lastly, this work highlights how AcrAB-TolC can be studied in a range of mimetic protein environments such as (native) nanodiscs, and critical information can be obtained through the use of novel HDX-MS delipidation workflows and biochemical assays.[4]

References

1. Ip, H., Stratton, K., Zgurskaya, H. & Liu, J. pH-induced Conformational Changes of AcrA, the Membrane Fusion Protein of *Escherichia coli* Multidrug Efflux System. *J. Biol. Chem.* **278**, 50474–50482 (2003).

2. Wang, B., Weng, J., Fan, K. & Wang, W. Interdomain flexibility and pH-induced conformational changes of AcrA revealed by molecular dynamics simulations. *J. Phys. Chem. B.* **116**, 3411–3420 (2012).

3. Thanassi, D. G., Suh, G. S. B. & Nikaido, H. Role of outer membrane barrier in efflux-mediated tetracycline resistance of *Escherichia coli*. *J. Bacteriol.* **177**, 998–1007 (1995).

4. Hammerschmid, D. *et al.* Chromatographic Phospholipid Trapping for Automated H/D Exchange Mass Spectrometry of Membrane Protein–Lipid Assemblies. *Anal. Chem.* **95**, 3002–3011 (2023).

5. Yates, J. R. A century of mass spectrometry: From atoms to proteomes. *Nat. Methods.* **8**, 633–637 (2011).

6. Biemann, K. Four decades of structure determination by mass spectrometry: From alkaloids to heparin. *J. Am. Soc. Mass Spectrom.* **13**, 1254–1272 (2002).

7. Fenn, J. B., Mann, M., Meng, C. K. A. I., Wong, S. F. & Whitehouse, C. M. Electrospray Ionisation for Mass Spectrometry of Large Biomolecules. *Science.* **246**, 64–70 (1989).

8. Loo, J. A. Studying noncovalent protein complexes by electrospray ionisation mass spectrometry. *Mass Spectrom. Rev.* **16**, 1–23 (1997).

9. Hilton, G. R. & Benesch, J. L. P. Two decades of studying non-covalent biomolecular assemblies by means of electrospray ionisation mass spectrometry. *J. R. Soc. Interface.* **9**, 801–816 (2012).

10. Sharon, M. & Robinson, C. V. The Role of Mass Spectrometry in Structure Elucidation of Dynamic Protein Complexes. *Annu. Rev. Biochem.* **76**, 167–193 (2007).

11. Benesch, J. L. P. & Ruotolo, B. T. Mass spectrometry: Come of age for structural and dynamical biology. *Curr. Opin. Struct. Biol.* **21**, 641–649 (2011).

12. Uetrecht, C. *et al.* High-resolution mass spectrometry of viral assemblies: Molecular composition and stability of dimorphic hepatitis B virus capsids. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9216–9220 (2008).

13. Gordiyenko, Y., Deroo, S., Zhou, M., Videler, H. & Robinson, C. V. Acetylation of L12 Increases Interactions in the *Escherichia coli* Ribosomal Stalk Complex. *J. Mol. Biol.* **380**, 404–414 (2008).

14. Liko, I., Allison, T. M., Hopper, J. T. & Robinson, C. V. Mass spectrometry guided structural biology. *Curr. Opin. Struct. Biol.* **40**, 136–144 (2016).

15. Matthews, B. W. X-ray crystallographic studies of proteins. *Ann. Rev. Phys. Chem.* **27**, 493–523 (1976).

16. Meyer, B. & Peters, T. NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors. *Angew. Chemie - Int. Ed.* **42**, 864–890 (2003).

17. Nogales, E. The development of cryo-EM into a mainstream structural biology technique. *Nat. Methods.* **13**, 24–27 (2016).

18. Seddon, A. M., Curnow, P. & Booth, P. J. Membrane proteins, lipids and detergents: Not just a soap opera. *Biochim. Biophys. Acta - Biomembr.* **1666**, 105–117 (2004).

19. Drew, D., Fröderberg, L., Baars, L. & De Gier, J. W. L. Assembly and overexpression of membrane proteins in *Escherichia coli*. *Biochim. Biophys. Acta - Biomembr.* **1610**, 3–10 (2003).

20. Wallin, E. & Heijne, G. Von. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* **7**, 1029–1038 (2008).

21. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Discov.* **5**, 10 (2006).

22. Santoni, V., Molloy, M. & Rabilloud, T. Membrane proteins and proteomics: Un amour impossible? *Electrophoresis.* **21**, 1054–1070 (2000).

23. Loo, R. R. O., Dales, N. & Andrews, P. C. Surfactant effects on protein structure examined by electrospray ionisation mass spectrometry. *Protein Sci.* **3**, 1975–1983 (1994).

24. Esteban, O. *et al.* Stoichiometry and Localization of the Stator Subunits E and G in *Thermus thermophilus* H^+ -ATPase/Synthase. *J. Biol. Chem.* **283**, 2595–2603 (2010).

25. Sharon, M., Ilag, L. L. & Robinson, C. V. Evidence for micellar structure in the gas phase. *J. Am. Chem. Soc.* **129**, 8740–8746 (2007).

26. Rundlett, K. L. & Armstrong, D. W. Mechanism of Signal Suppression by Anionic Surfactants in Capillary Electrophoresis - Electrospray Ionisation Mass Spectrometry. *Anal. Chem.* **68**, 3493–3497 (1996).

27. Martin, J. & Sawyer, A. Elucidating the structure of membrane proteins. *Biotechniques.* **66**, 167–170 (2019).

28. Dafun, A. S. & Marcoux, J. Structural mass spectrometry of membrane proteins. *Biochim. Biophys. Acta - Proteins Proteomics.* **1870**, 140813 (2022).

29. Kuhlbrandt, W. Forty years in cryoEM of membrane proteins. *Mircroscopy* **71**, 30–50 (2022).

30. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature.* **596**, 583–589 (2021).

31. Varadi, M. *et al.* AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).

32. Russell Lewis, B., Lawrence, R., Hammerschmid, D. & Reading, E. Structural mass spectrometry approaches to understand multidrug efflux systems. *Essays Biochem.* **67**, 255–267 (2023).

33. Weis, D. D. *Hydrogen Exchange Mass Spectrometry of Protiens: Fundamentals, Methods and Applications*. (Wiley, 2016).

34. Javed, W. *et al.* Structural Insights into the Catalytic Cycle of a Bacterial Multidrug ABC Efflux Pump. *J. Mol. Biol.* **434**, (2022).

35. Reading, E. *et al.* Perturbed structural dynamics underlie inhibition and altered efflux of the multidrug resistance pump AcrB. *Nat. Commun.* **11**, 1–11 (2020).

36. Jixing Lyu, Chang Liu, Tianqi Zhang, Samantha Schrecke, Nicklaus P. Elam, Georg Hochberg, David Russell, M. Z. & Laganowsky, A. Structural basis for lipid and copper regulation of the ABC transporter MsbA. *bioRxiv* (2022).

37. Bolla, J. R., Howes, A. C., Fiorentino, F. & Robinson, C. V. Assembly and regulation of the chlorhexidine-specific efflux pump AceI. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 17011–17018 (2020).

38. Shi, X. *et al. In situ* structure and assembly of the multidrug efflux pump AcrAB-TolC. *Nat. Commun.* **10**, 4–9 (2019).

39. Wang, Y., Sun, J., Qiao, J., Ouyang, J. & Na, N. A 'soft' and 'hard' Ionisation Method for Comprehensive Studies of Molecules. *Anal. Chem.* **90**, 14095–14099 (2018).

40. Karas, M. & Hillenkamp, F. Laser Desorption Ionisation of Proteins with Molecular Masses Exceeding 10 000 Daltons. *Anal. Chem.* **60**, 2299–2301 (1988).

41. Peetz, O. *et al.* LILBID and nESI: Different Native Mass Spectrometry Techniques as Tools in Structural Biology. *J. Am. Soc. Mass Spectrom.* **30**, 181–191 (2019).

42. Banerjee, S. & Mazumdar, S. Electrospray Ionisation Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *Int. J. Anal. Chem.* **2012**, 1–40 (2012).

43. Karas, M., Bahr, U. & Dülcks, T. Nano-electrospray ionisation mass spectrometry: Addressing analytical problems beyond routine. *Fresenius. J. Anal. Chem.* **366**, 669–676 (2000).

44. Fenn, J. B., Mann, M., Meng, C. K. & Wong, S. F. Electrospray ionisation-principles and practice. *Mass Spectrom. Rev.* **9**, 37–70 (1990).

45. Cech, N. B. & Enke, C. G. Practical implications of some recent studies in electrospray ionisation fundamentals. *Mass Spectrom. Rev.* **20**, 362–387 (2001).

46. Taylor, G. Disintegration of water drops in an electric field. *Proc. R. Soc. Lond. A.* **280**, 383–397 (1964).

47. Wilm, M. S. & Mann, M. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *Int. J. Mass Spectrom. Ion Process.* **136**, 167–180 (1994).

48. Bruins, A. P. Mass spectrometry with ion sources operating at atmospheric pressure. *Mass Spectrom. Rev.* **10**, 53–77 (1991).

49. Fenn, J. B. Ion formation from charged droplets: Roles of geometry, energy, and time. *J. Am. Soc. Mass Spectrom.* **4**, 524–535 (1993).

50. Rayleigh, Lord. On the Equilibrium of Liquid Conducting Masses charged with Electricity. *Sci. Pap.* **14**, 184–186 (1882).

51. Konermann, L. A Simple Model for the Disintegration of Highly Charged Solvent Droplets during Electrospray Ionisation. *J. Am. Soc. Mass Spectrom.* **20**, 496–506 (2009).

52. Hiraoka, K. & Kudaka, I. Electrospray interface for liquid chromatography/mass spectrometry. *Rapid Commun. Mass Spectrom.* **4**, 519–526 (1990).

53. Keqi Tang and Richard D. Smith. Physical/Chemical Separations in the Break-up of Highly Charged Droplets from Electrosprays. *Am. Soc. Mass Spectrom.* **12**, 343–347 (2001).

54. Konermann, L., Ahadi, E., Rodriguez, A. D. & Vahidi, S. Unraveling the mechanism of electrospray ionisation. *Anal. Chem.* **85**, 2–9 (2013).

55. Dole, M. *et al.* Molecular beams of macroions. *J. Chem. Phys.* **49**, 2240–2249 (1968).

56. Schmelzeisen-Redeker, G., Bütfering, L. & Röllgen, F. W. Desolvation of ions and molecules in thermospray mass spectrometry. *Int. J. Mass Spectrom. Ion Process.* **90**, 139–150 (1989).

57. Fernandez De La Mora, J. Electrospray ionisation of large multiply charged species proceeds via Dole's charged residue mechanism. *Anal. Chim. Acta.* **406**, 93–104 (2000).

58. Iribarne, J. V. & Thomson, B. A. On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **64**, 2287–2294 (1976).

59. Wilm, M. & Mann, M. Analytical properties of the nanoelectrospray ion source. *Anal. Chem.* **68**, 1–8 (1996).

60. Juraschek, R., Dülcks, T. & Karas, M. Nanoelectrospray - More than just a minimized-flow electrospray ionisation source. *J. Am. Soc. Mass Spectrom.* **10**, 300–308 (1999).

61. Morgner, N., Barth, H. D. & Brutschy, B. A new way to detect noncovalently bonded complexes of biomolecules from liquid micro-droplets by laser mass spectrometry. *Aust. J. Chem.* **59**, 109–114 (2006).

62. Morgner, N., Kleinschroth, T., Barth, H. D., Ludwig, B. & Brutschy, B. A Novel Approach to Analyze Membrane Proteins by Laser Mass Spectrometry: From Protein Subunits to the Integral Complex. *J. Am. Soc. Mass Spectrom.* **18**, 1429–1438 (2007).

63. Hellwig, N. *et al.* Native mass spectrometry goes more native: Investigation of membrane protein complexes directly from SMALPs. *Chem. Commun.* **54**, 13702–13705 (2018).

64. Benesch, J. L. P., Ruotolo, B. T., Simmons, D. A. & Robinsons, C. V. Protein complexes in the gas phase: Technology for structural genomics and proteomics. *Chem. Rev.* **107**, 3544–3567 (2007).

65. Heck, A. J. R. & Van Den Heuvel, R. H. H. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.* **23**, 368–389 (2004).

66. Douglas, D. J. & French, J. B. Collisional focusing effects in radio frequency quadrupoles. *J. Am. Soc. Mass Spectrom.* **3**, 398–408 (1992).

67. Tolmachev, A. V., Udseth, H. R. & Smith, R. D. Modeling the ion density distribution in collisional cooling RF multipole ion guides. *Int. J. Mass Spectrom.* **222**, 155–174 (2003).

68. Rockwood, A. L., Kushnir, M. M. & Clarke, N. J. Mass Spectrometry. in *Principles and Applications of Clinical Mass Spectrometry* 33–65 (Elsevier Inc., 2018).

69. Chernushevich, I. V. & Thomson, B. A. Collisional Cooling of Large Ions in Electrospray Mass Spectrometry. *Anal. Chem.* **76**, 1754–1760 (2004).

70. Hoffman, E. & Stroobant, V. *Mass spectrometry*. (Wiley, 2007).

71. Mitchell Wells, J. & McLuckey, S. A. Collision-induced dissociation (CID) of peptides and proteins. *Methods Enzymol.* **402**, 148–185 (2005).

72. CWK Lam, C. H. *et al.* Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications. *Clin. Biochem. Rev.* **24**, 243–248 (2003).

73. Mcluckey, S. A. Principles of Collisional Activation in Analytical Mass Spectrometry. *Am. Soc. Mass Spectrom.* **3**, 599–614 (1992).

74. Benesch, J. L. P. Collisional Activation of Protein Complexes: Picking Up the Pieces. *J. Am. Soc.*

Mass Spectrom. **20**, 341–348 (2009).

75. Jennings, K. R. The changing impact of the collision-induced decomposition of ions on mass spectrometry. *Int. J. Mass Spectrom.* **200**, 479–493 (2000).

76. Smith, R. D., Loa, J. A., Barinaga, C. J., Edmonds, C. G. & Udseth, H. R. Collisional Activation and Collision-Activated Dissociation of Large Multiply Charged Polypeptides and Proteins Produced by Electrospray Ionisation. *Am. Soc. Mass Spectrom.* **1**, 53–65 (1990).

77. Pagel, K., Hyung, S., Ruotolo, B. T. & Robinson, C. V. Alternate Dissociation Pathways Identified in Charge-Reduced Protein Complex Ions. *Anal. Chem.* **82**, 5363–5372 (2010).

78. Barrera, N. P., Di Bartolo, N., Booth, P. J. & Robinson, C. V. Micelles protect membrane complexes from solution to vacuum. *Science.* **321**, 243–246 (2008).

79. Landreh, M., Liko, I., Uzdavinys, P., Coincon, M. & Jonathan, T. S. Controlling release, unfolding and dissociation of membrane protein complexes in the gas phase through collisional cooling. *Chem. Commun.* **51**, 15582–15584 (2016).

80. Wolfgang Paul, H. S. Ein neues Massenspektrometer ohne Magnetfeld. *Zeitschrift für Naturforsch. A* **8**, 448–450 (1953).

81. El-Aneed, A., Cohen, A. & Banoub, J. Mass spectrometry, review of the basics: Electrospray, MALDI, and commonly used mass analyzers. *Appl. Spectrosc. Rev.* **44**, 210–230 (2009).

82. March, R. E., Hughes, R. J. & Roberts, A. D. Quadrupole storage mass spectrometry. *Anal. Chim. Acta.* **243**, 323–324 (1991).

83. Dawson, P. H. Quadrupole mass analyzers: Performance, design and some recent applications. *Mass Spectrom. Rev.* **5**, 1–37 (1986).

84. Hofstadler, S. A., Bakhtiar, R. & Smith, R. D. Electrospray Ionisation Mass Spectrometry. *J. Chem. Educ.* **73**, 82–88 (1996).

85. Henchman, M. & Steel, C. Understanding the Quadrupole Mass Filter through Computer Simulation. *J. Chem. Educ.* **75**, 1049 (1998).

86. Sobott, F., Hernández, H., McCammon, M. G., Tito, M. A. & Robinson, C. V. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal. Chem.* **74**, 1402–1407 (2002).

87. Ruotolo, B. T. *et al.* Ion mobility-mass spectrometry reveals long-lived, unfolded intermediates in the dissociation of protein complexes. *Angew. Chemie - Int. Ed.* **46**, 8001–8004 (2007).

88. Cotter, R. J. Time-of-Flight Mass Spectrometry for the Structural Analysis of Biological Molecules. *Anal. Chem.* **64**, 1027–1039 (1992).

89. O'Halloran, G. J., Fluegge, R. A., Betts, J. F. & Everett, W. L. *Technical Documentary Report No. ASD-TDR-62-644*. (1964).

90. Giles, K. *et al.* Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun. Mass Spectrom.* **18**, 2401–2414 (2004).

91. Savaryn, J. P., Toby, T. K. & Kelleher, N. L. A researcher's guide to mass spectrometry-based proteomics. *Proteomics.* **16**, 2435–2443 (2016).

92. Zubarev, R. A. & Makarov, A. Orbitrap mass spectrometry. *Anal. Chem.* **85**, 5288–5296 (2013).

93. Liu, R., Li, Q. & Smith, L. M. Detection of large ions in time-of-flight mass spectrometry: effects of ion mass and acceleration voltage on microchannel plate detector response. *J. Am. Soc. Mass. Spectrom.* **25**, 1374–1383 (2014).

94. Geno, P. W. & Macfarlane, R. D. Secondary electron emission induced by impact of low-velocity molecular ions on a microchannel plate. *Int. J. Mass Spectrom. Ion Process.* **92**, 195–210 (1989).

95. Westmacott, G., Frank, M., Labov, S. E. & Benner, W. H. Using a superconducting tunnel junction detector to measure the secondary electron emission efficiency for a microchannel plate detector bombarded by large molecular ions. *Rapid Commun. Mass Spectrom.* **1861**, 1854–1861 (2000).

96. Wiza, J. L. Microchannel plate detectors. *Nucl. Instruments Methods.* **162**, 587–601 (1979).

97. Leney, A. C. & Heck, A. J. R. Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass Spectrom.* **28**, 5–13 (2017).

98. Barth, M. & Schmidt, C. Native mass spectrometry — A valuable tool in structural biology. *J. Mass Spectrom.* **55**, 1–17 (2020).

99. Pedro, L. & Quinn, R. J. Native Mass Spectrometry in Fragment-Based Drug Discovery. *Molecules.* **21**, 1–16 (2016).

100. Saliba, A. E., Vonkova, I. & Gavin, A.-C. The systematic analysis of protein-lipid interactions comes of age. *Nat. Rev. Mol. Cell Biol.* **16**, 753–761 (2015).

101. Annesley, T. M. Ion suppression in mass spectrometry. *Clin. Chem.* **49**, 1041–1044 (2003).

102. Hernandez, H. & Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat. Protoc.* **2**, 715–726 (2007).

103. Verkerk, U. H. & Kebarle, P. Ion–Ion and Ion–Molecule Reactions at the Surface of Proteins Produced by Nanospray. Information on the Number of Acidic Residues and Control of the Number of Ionized Acidic and Basic Residues. *Am. Soc. Mass Spectrom.* **16**, 1325–1341 (2005).

104. Konermann, L. Addressing a Common Misconception: Ammonium Acetate as Neutral pH 'Buffer' for Native Electrospray Mass. *Am. Soc. Mass Spectrom.* **28**, 1827–1835 (2017).

105. Heck, A. J. R. Native mass spectrometry: A bridge between interactomics and structural biology. *Nat. Methods.* **5**, 927–933 (2008).

106. Keener, J. E., Zhang, G. & Marty, M. T. Native Mass Spectrometry of Membrane Proteins. *Anal. Chem.* **93**, 583–597 (2021).

107. Reading, E. *et al.* The Role of the Detergent Micelle in Preserving the Structure of Membrane Proteins in the Gas Phase. *Angew. Chemie - Int. Ed.* **54**, 4577–4581 (2015).

108. Borysik, A. J., Hewitt, D. J. & Robinson, C. V. Detergent Release Prolongs the Lifetime of Native-like Membrane Protein Conformations in the Gas-Phase. *J. Am. Chem. Soc.* **135**, 6078–6083 (2013).

109. Laganowsky, A., Reading, E., Hopper, J. T. S. & Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nat. Protoc.* **8**, 639–651 (2013).

110. Laganowsky, A. *et al.* Membrane proteins bind lipids selectively to modulate their structure and function. *Nature.* **510**, 172–191 (2014).

111. Pollock, N. L., Lee, S. C., Patel, J. H., Gulamhussein, A. A. & Rothnie, A. J. Structure and function of membrane proteins encapsulated in a polymer-bound lipid bilayer. *Biochim. Biophys. Acta - Biomembr.* **1860**, 809–817 (2018).

112. Dörr, J. M. *et al.* The styrene–maleic acid copolymer: a versatile tool in membrane research. *Eur. Biophys. J.* **45**, 3–21 (2016).

113. Redhair, M., Clouser, A. F. & Atkins, W. M. Hydrogen-Deuterium Exchange Mass Spectrometry of Membrane Proteins in Lipid Nanodiscs. *Chem. Phys. Lipids.* **220**, 14–22 (2019).

114. Lee, S. C. *et al.* A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nat. Protoc.* **11**, 1149–1162 (2016).

115. Möller, I. R. *et al.* Improving the Sequence Coverage of Integral Membrane Proteins during Hydrogen/Deuterium Exchange Mass Spectrometry Experiments. *Anal. Chem.* **91**, 10970–10978 (2019).

116. Jordan, J. S., Xia, Z. & Williams, E. R. Tips on Making Tiny Tips: Secrets to Submicron Nanoelectrospray Emitters. *J. Am. Soc. Mass Spectrom.* **33**, 607–611 (2022).

117. Krusemark, C. J., Frey, B. L., Belshaw, P. J. & Smith, L. M. Modifying the Charge State Distribution of Proteins in Electrospray Ionisation Mass Spectrometry by Chemical Derivatization. *Am. Soc. Mass Spectrom.* **20**, 1617–1625 (2009).

118. Marty, M. T. *et al.* Bayesian deconvolution of mass and ion mobility spectra: From binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).

119. Tamara, S., Boer, M. A. Den & Heck, A. J. R. High-Resolution Native Mass Spectrometry. *Chem. Rev.* **122**, 7269–7326 (2022).

120. Mitra, G. Application of native mass spectrometry in studying intrinsically disordered proteins: A special focus on neurodegenerative diseases. *Biochim. Biophys. Acta - Proteins Proteomics.* **1867**, 140260 (2019).

121. Beveridge, R. *et al.* A Mass-Spectrometry-Based Framework To Define the Extent of Disorder in Proteins. *Anal. Chem.* **86**, 10979–10991 (2014).

122. Borysik, A. J., Kovacs, D., Guharoy, M. & Tompa, P. Ensemble Methods Enable a New Definition for the Solution to Gas-Phase Transfer of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **137**, 13807–13817 (2015).

123. Lanucara, F., Holman, S. W., Gray, C. J. & Eyers, C. E. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nat. Chem.* **6**, 281–294 (2014).

124. Kanu, A. B., Dwivedi, P., Tam, M., Matz, L. & Hill Jr., H. H. Ion mobility–mass spectrometry. *J. mass Spectrom.* **43**, 1–22 (2008).

125. McLean, J. A., Ruotolo, B. T., Gillig, K. J. & Russell, D. H. Ion mobility-mass spectrometry: A new paradigm for proteomics. *Int. J. Mass Spectrom.* **240**, 301–315 (2005).

126. Jin, L., Barran, P. E., Deakin, J. A., Lyon, M. & Uhrín, D. Conformation of glycosaminoglycans by ion mobility mass spectrometry and molecular modelling. *Phys. Chem. Chem. Phys.* **7**, 3464–3471 (2005).

127. Larson, M. E. *et al.* Metabolic Profiling of *Escherichia coli* by Ion Mobility-Mass Spectrometry with MALDI Ion Source. *J. Mass Spectrom.* **45**, 1383–1393 (2010).

128. Van De Waterbeemd, M. *et al.* High-fidelity mass analysis unveils heterogeneity in intact ribosomal particles. *Nat. Methods.* **14**, 283–286 (2017).

129. Michalski, A. *et al.* Mass spectrometry-based proteomics using Q exactive, a high-performance benchtop quadrupole orbitrap mass spectrometer. *Mol. Cell. Proteomics.* **10**, 1–11 (2011).

130. Wales, T. E. & Engen, J. R. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. *Mass Spectrom. Rev.* **25**, 158–170 (2006).

131. Masson, G. R. *et al.* Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods.* **16**, 595–602 (2019).

132. Aase Hvidt and K Linderstrom-Lang. Exchange of hydrogen atoms in insulin with deuteruim atoms in aqueous solutions. *Short Commun.* **14**, 574–575 (1954).

133. Zhang, Z. & Smith, D. Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. *Protein Sci.* **2**, 522–531 (1993).

134. Englander, J. J., Rogero, J. R. & Englander, S. W. Protein Hydrogen Exchange Studied by the Fragment Separation Method. *Anal Biochem.* **147**, 234–244 (2012).

135. Trabjerg, E., Nazari, Z. E. & Rand, K. D. Conformational Analysis of Complex Protein States by Hydrogen/Deuterium Exchange Mass Spectrometry (HDX-MS): Challenges and Emerging Solutions. *Trends Anal. Chem.* **106**, 125–138 (2018).

136. Engen, J. R., Wales, T. E. & Shi, X. Hydrogen Exchange Mass Spectrometry for Conformational Analysis of Proteins. in *Encyclopedia of Analytical Chemistry* (ed. Meyers, R. A.) 1–17 (New York: Wiley, 2011).

137. Deng, B., Lento, C. & Wilson, D. J. Hydrogen deuterium exchange mass spectrometry in biopharmaceutical discovery and development - A review. *Anal. Chim. Acta.* **940**, 8–20 (2016).

138. Pirrone, G. F., Iacob, R. E. & Engen, J. R. Applications of Hydrogen/Deuterium Exchange MS from 2012 to 2014. *Anal. Chem.* **87**, 99–118 (2014).

139. Merkle, P. S. *et al.* Substrate-modulated unwinding of transmembrane helices in the NSS transporter LeuT. *Sci. Adv.* **4**, 1–12 (2018).

140. Eisinger, M. L., Nie, L., Dörrbaum, A. R., Langer, J. D. & Michel, H. The xenobiotic extrusion mechanism of the MATE transporter NorM_PS from *Pseudomonas stutzeri*. *J. Mol. Biol.* **18**, 1–28 (2018).

141. Tsutsui, Y. & Wintrode, P. L. Hydrogen/Deuterium Exchange-Mass Spectrometry: A Powerful Tool for Probing Protein Structure, Dynamics and Interactions. *Curr. Med. Chem.* **14**, 2344–2358 (2007).

142. Li, J., Rodnin, M. V, Ladokhin, A. S. & Gross, M. L. Hydrogen−Deuterium Exchange and Mass Spectrometry Reveal the pH-Dependent Conformational Changes of Diphtheria Toxin T Domain. *Am. Chem. Soc.* **53**, 6849–6854 (2014).

143. Coales, S. J. *et al.* Expansion of time window for mass spectrometric measurement of amide hydrogen/deuterium exchange reactions. *Rapid Commun. Mass Spectrom.* **24**, 3585–3592 (2010).

144. Englander, S. W. & Kallenbach, N. R. Hydrogen exchange and structural dynamics of proteins

and nucleic acids. *Quart. Rev. Biophys.* **16**, 521–655 (1983).

145. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins.* **17**, 75–86 (1993).

146. Kim, K. S. & Woodward, C. Protein Internal Flexibility and Global Stability: Effect of Urea on Hydrogen Exchange Rates of Bovine Pancreatic Trypsin Inhibitor. *Biochemistry.* **32**, 9609–9613 (1993).

147. Deng, Y. & Smith, D. L. Identification of unfolding domains in large proteins by their unfolding rates. *Biochemistry.* **37**, 6256–6262 (1998).

148. Swint-Kruse, L. & Robertson, A. D. Temperature and pH dependences of hydrogen exchange and global stability for ovomucoid third domain. *Biochemistry.* **35**, 171–180 (1996).

149. Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T. & Dobson, C. M. Detection of transient protein folding populations by mass spectrometry. *Science.* **262**, 896–900 (1993).

150. Raynal, B., Lenormand, P., Baron, B., Hoos, S. & England, P. Quality assessment and optimisation of purified protein samples: why and how? *Microb. Cell Fact.* **13**, 1–10 (2014).

151. Glasoe, P. K. & Long, F. A. Use of glass electrodes to measure acidities in deuterium oxide. *J. Phys. Chem.* **64**, 188–190 (1960).

152. Deng, Y., Zhang, Z. & Smith, D. L. Comparison of Continuous and Pulsed Labeling Amide Hydrogen Exchange/Mass Spectrometry for Studies of Protein Dynamics. **10**, 675–684 (1999).

153. Wu, Y., Engen, J. R., December, R. & February, A. What mass spectrometry can reveal about protein function. *RSC Adv.* **129**, 290–296 (2004).

154. Hai, P. & David, S. Amide Hydrogen Exchange/Mass Spectrometry Applied to Cooperative Protein Folding: Equilibrium Unfolding of *Staphylococcus aureus* Aldolase. **380**, 285–308 (2004).

155. Hoofnagle, A. N., Resing, K. A. & Ahn, N. G. Protein analysis by hydrogen exchange mass spectrometry. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 1–25 (2003).

156. Molday, R. S., Englander, S. W. & Kallen, R. G. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Biochemistry.* **11**, 150–158 (1973).

157. Walters, B. T., Ricciuti, A., Mayne, L. & Englander, S. W. Minimizing Back Exchange in the Hydrogen Exchange-Mass Spectrometry Experiment. *J. Am. Soc. Mass. Spectrom.* **23**, 2132–2139 (2012).

158. Wales, T. E., Fadgen, K. E., Gerhardt, G. C. & Engen, J. R. High-Speed and High-Resolution UPLC Separation at Zero Degrees Celsius. *Anal. Chem.* **80**, 6815–6820 (2008).

159. Hoofnagle, A. N., Resing, K. A. & Ahn, N. G. Practical Methods for Deuterium Exchange/Mass Spectrometry. in *MAP Kinase Signalling Protocols.* (ed. Seger, R.) 283–298 (Humana Press, 2004).

160. Peterle, D., Wales, T. E. & Engen, J. R. Simple and Fast Maximally Deuterated Control (maxD) Preparation for Hydrogen-Deuterium Exchange Mass Spectrometry Experiments. *Anal. Chem.* **94**, 10142–10150 (2022).

161. Zhang, Z. & Smith, D. L. Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. *Protein Sci.* **2**, 522–531 (1993).

162. Mayne, L. Hydrogen Exchange Mass Spectrometry. in *Methods in Enzymology.* 335–356

(Elsevier Inc., 2016).

163. Miyagi, M. & Nakazawa, T. Determination of pKa values of individual histidine residues in proteins using mass spectrometry. *Anal. Chem.* **80**, 6481–6487 (2008).

164. Singh, H., Dai, Y., Outten, F. W. & Busenlehner, L. S. *Escherichia coli* SufE sulfur transfer protein modulates the SufS cysteine desulfurase through allosteric conformational dynamics. *J. Biol. Chem.* **288**, 36189–36200 (2013).

165. Sowole, M. A. & Konermann, L. Effects of protein-ligand interactions on hydrogen/deuterium exchange kinetics: Canonical and noncanonical scenarios. *Anal. Chem.* **86**, 6715–6722 (2014).

166. Yan, X. & Maiser, C. Hydrogen/Deuterium Exchange Mass Spectrometry. in *Mass Spectrometry of Proteins and Peptides.* 241–273 (Humana Press, 2009).

167. Hoofnagle, A. N., Resing, K. A. & Ahn, N. G. Protein analysis by hydrogen exchange mass spectrometry. *Annu. Rev. Biophys. Biomol. Struct.* **32**, 1–25 (2003).

168. Hamuro, Y., Coales, S. J., Molnar, K. S., Tuske, S. J. & Morrow, J. A. Specificity of immobilized porcine pepsin in H/D exchange compatible conditions. *Rapid Commun. Mass Spectrom.* **22**, 1041–1046 (2008).

169. Cravello, L., Lascoux, D. & Forest, E. Use of different proteases working in acidic conditions to improve sequence coverage and resolution in hydrogen/deuterium exchange of large proteins. *Rapid Commun. Mass Spectrom.* **17**, 2387–2393 (2003).

170. Liebert, M. A., Garcia, R. A., Pantazatos, D. & Villarreal, F. J. Investigating Protein-Ligand Interactions. *Assay Drug Dev. Technol.* **2**, 81–91 (2004).

171. Wang, L. & Smith, D. L. Downsizing improves sensitivity 100-fold for hydrogen exchange-mass spectrometry. *Anal. Biochem.* **314**, 46–53 (2003).

172. Busby, S. A., Chalmers, M. J. & Griffin, P. R. Improving digestion efficiency under H/D exchange conditions with activated pepsinogen coupled columns. *Int. J. Mass Spectrom.* **259**, 130–139 (2007).

173. Majumdar, R. *et al.* Minimizing carry-over in an online pepsin digestion system used for the H/D exchange mass spectrometric analysis of an IgG1 monoclonal antibody. *J. Am. Soc. Mass Spectrom.* **23**, 2140–2148 (2012).

174. Wu, Y., Engen, J. R. & Hobbins, W. B. Ultra performance liquid chromatography (UPLC) further improves hydrogen/deuterium exchange mass spectrometry. *J. Am. Soc. Mass Spectrom.* **17**, 163–167 (2006).

175. De Vijlder, T. *et al.* A tutorial in small molecule identification via electrospray ionisation-mass spectrometry: The practical art of structural elucidation. *Mass Spectrom. Rev.* **37**, 607–629 (2018).

176. Bateman, K. P. *et al.* MSE with mass defect filtering for *in vitro* and *in vivo* metabolite identification. *Rapid Commun. Mass Spectrom.* **21**, 1485–1496 (2007).

177. Steen, H. & Mann, M. The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* **5**, 699–711 (2004).

178. Roepstorff, P. & Fohlman, J. Letter to the editors. *Biol. Mass Spectrom.* **11**, 601–601 (1984).

179. Biemann, K. Contributions of mass spectrometry to peptide and protein structure. *Biol. Mass Spectrom.* **16**, 99–111 (1988).

180. Ziady, A. G. & Kinter, M. Protein sequencing with tandem mass spectrometry. *Methods Mol. Biol.* **544**, 325–341 (2009).

181. Lau, K. W. *et al.* Observations on the detection of b- and y-type ions in the collisionally activated decomposition spectra of protonated peptides. *Rapid Commun. Mass Spectrom.* **23**, 1508–1514 (2009).

182. Cleveland, J. P. & Rose, J. R. Identification of b-/y-ions in MS/MS spectra using a two stage neural network. *Proteome Sci.* **11**, S4 (2013).

183. Perkins, D. N., Pappin, D. J. C., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis.* **20**, 3551–3567 (1999).

184. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976–989 (1994).

185. Politis, A. & Martens, C. A glimpse into the molecular mechanism of integral membrane proteins through hydrogen – deuterium exchange mass spectrometry. *Protein Soc.* **29**, 1285–1301 (2020).

186. Rey, M. *et al.* Effective removal of nonionic detergents in protein mass spectrometry, hydrogen/deuterium exchange, and proteomics. *Anal. Chem.* **82**, 5107–5116 (2010).

187. Vissers, J. P. C., Hulst, W. P., Chervet, J. P., Snijders, H. M. J. & Cramers, C. A. Automated on-line ionic detergent removal from minute protein/peptide samples prior to liquid chromatography-electrospray mass spectrometry. *J. Chromatogr. B.* **686**, 119–128 (1996).

188. Hebling, C. M. *et al.* Conformational analysis of membrane proteins in phospholipid bilayer nanodiscs by hydrogen exchange mass spectrometry. *Anal. Chem.* **82**, 5415–5419 (2010).

189. Knowles, T. J. *et al.* Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J. Am. Chem. Soc.* **131**, 7484–7485 (2009).

190. Reading, E. *et al.* Interrogating Membrane Protein Conformational Dynamics within Native Lipid Compositions. *Angew. Chemie - Int. Ed.* **56**, 15654–15657 (2017).

191. Redhair, M., Clouser, A. F. & Atkins, W. M. Hydrogen-deuterium exchange mass spectrometry of membrane proteins in lipid nanodiscs. *Chem. Phys. Lipids.* **220**, 14–22 (2019).

192. Anderson, K. W., Gallagher, E. S. & Hudgens, J. W. Automated Removal of Phospholipids from Membrane Proteins for H/D Exchange Mass Spectrometry Workflows. *Anal. Chem.* **90**, 6409–6412 (2018).

193. Hotchko, M., Anand, G. S., Komives, E. A. & Ten Eyck, L. F. Automated extraction of backbone deuteration levels from amide H/^{2}H mass spectrometry experiments. *Protein Sci.* **15**, 583–601 (2006).

194. Igor A. Kaltashov and Stephen J. Eyles. Mass spectrometry-based approaches to study biomolecular dynamics: equilibrium intermediates. in *Mass Spectrometry in Biophysics: Conformation and Dynamics of Biomolecules.* 183–230 (Wiley, 2005).

195. Morgan, C. R. & Engen, J. R. Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry. *Curr. Protoc. Protein Sci.* **58**, 17.6.1-17.6.17 (2009).

196. Mayne, L. *et al.* Many overlapping peptides for protein hydrogen exchange experiments by the fragment separation-mass spectrometry method. *J. Am. Soc. Mass Spectrom.* **22**, 1898–

1905 (2011).

197. Waters. DynamX HDX Data Analysis Software 3.0. (2015).

198. Pascal, B. D. *et al.* HDX Workbench: Software for the Analysis of H/D Exchange MS Data. *J. Am. Soc. Mass Spectrom.* **23**, 1–16 (2012).

199. Rey, M. *et al.* Mass Spec Studio for Integrative Structural Biology. *Structure.* **22**, 1538–1548 (2015).

200. Vedadi, M., Arrowsmith, C. H., Allali-Hassani, A., Senisterra, G. & Wasney, G. A. Biophysical characterization of recombinant proteins: A key to higher structural genomics success. *J. Struct. Biol.* **172**, 107–119 (2010).

201. Soltermann, F. *et al.* Quantifying Protein – Protein Interactions by Molecular Counting with Mass Photometry. *Angew. Chemie - Int. Ed.* **132**, 10866–10871 (2020).

202. Verschueren, H. Interference reflection microscopy in cell biology: methodology and applications. *J. Cell Sci.* **75**, 279–301 (1985).

203. Ortega-Arroyo, J. & Kukura, P. Interferometric scattering microscopy (iSCAT): New frontiers in ultrafast and ultrasensitive optical microscopy. *Phys. Chem. Chem. Phys.* **14**, 15625–15636 (2012).

204. Wu, D. & Piszczek, G. Standard Protocol for Mass Photometry Experiments. *Eur. Biophys. J.* **50**, 403–409 (2021).

205. Olerinyova, A. *et al.* Mass Photometry of Membrane Proteins. *Chem* **7**, 1–13 (2021).

206. Michael, C. A., Dominey-Howes, D. & Labbate, M. The antimicrobial resistance crisis: Causes, consequences, and management. *Front. Public Heal.* **2**, 1–8 (2014).

207. Morrison, L. & Zembower, T. R. Antimicrobial Resistance: Global Report on Surveillance. *WHO.* **30**, 619–635 (2020).

208. Lahsoune, M., Boutayeb, H., Zerouali, K., Belabbes, H. & El Mdaghri, N. Antibiotic Resistance Threats In The United States. *CDC.* **37**, 828–831 (2013).

209. Murray, C. J. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet.* **399**, 629–655 (2022).

210. Nikaido, H. Multidrug Resistance in Bacteria. *Annu Rev Biochem.* **78**, 119–146 (2009).

211. Rice, L. B. Federal funding for the study of antimicrobial resistance in nosocomial pathogens: No ESKAPE. *J. Infect. Dis.* **197**, 1079–1081 (2008).

212. De Oliveira, D. M. P. *et al.* Antimicrobial resistance in ESKAPE pathogens. *Clin. Microbiol. Rev.* **33**, 1–19 (2020).

213. Lorusso, A. B., Carrara, J. A., Barroso, C. D. N., Tuon, F. F. & Faoro, H. Role of Efflux Pumps on Antimicrobial Resistance in *Pseudomonas aeruginosa*. *Int. J. Mol. Sci.* **23**, (2022).

214. Saier, M. H. *et al.* Evolutionary origins of multidrug and drug-specific efflux pumps in bacteria. *FASEB J.* **12**, 265–274 (1998).

215. Tseng, T. T. *et al.* The RND permease superfamily: An ancient, ubiquitous and diverse family that includes human disease and development proteins. *J. Mol. Microbiol. Biotechnol.* **1**, 107–125 (1999).

216. Du, D. *et al.* Multidrug efflux pumps: structure, function and regulation. *Nat. Rev. Microbiol.*

16, 523–539 (2018).

217. Lubelski, J., Konings, W. N. & Driessen, A. J. M. Distribution and Physiology of ABC-Type Transporters Contributing to Multidrug Resistance in Bacteria. *Microbiol. Mol. Biol. Rev.* **71**, 463–476 (2007).

218. Kuroda, T. & Tsuchiya, T. Multidrug efflux transporters in the MATE family. *Biochim. Biophys. Acta - Proteins Proteomics.* **1794**, 763–768 (2009).

219. Saier, M. H. *et al.* The major facilitator superfamily. *J. Mol. Microbiol. Biotechnol.* **1**, 257–279 (1998).

220. Hassan, K. A. *et al.* Pacing across the membrane: the novel PACE family of efflux pumps is widespread in Gram-negative pathogens. *Res. Microbiol.* **169**, 450–454 (2018).

221. Nikaido, H. & Takatsuka, Y. Mechanisms of RND Multidrug Efflux Pumps. *Biochim. Biophys. Acta.* **1794**, 769–781 (2009).

222. Jack, D. L., Yang, N. M. & Saier, M. H. The drug/metabolite transporter superfamily. *Eur. J. Biochem.* **268**, 3620–3639 (2001).

223. Piddock, L. J. V. Multidrug-resistance efflux pumps — not just for resistance. *Nat. Rev. Microbiol.* **4**, 629–636 (2006).

224. Nikaido, H. & Zgurskaya, H. I. AcrAB and related multidrug efflux pumps of *Escherichia coli*. *J. Mol. Microbiol. Biotechnol.* **3**, 215–218 (2001).

225. Luckey, M. *Membrane Structural Biology: With Biochemical and Biophysical Foundations*. (Cambridge University Press., 2008).

226. Watson, H. Biological membranes. *Essays Biochem.* **59**, 43–70 (2015).

227. Booth, P. J. & Curnow, P. Folding scene investigation: membrane proteins. *Curr. Opin. Struct. Biol.* **19**, 8–13 (2009).

228. Harayama, T. & Riezman, H. Understanding the diversity of membrane lipid composition. *Nat. Rev. Mol. Cell Biol.* **19**, 281–296 (2018).

229. Van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell. Biol.* **9**, 112–124 (2008).

230. Nakamura, M. T., Yudell, B. E. & Loor, J. J. Progress in Lipid Research Regulation of energy metabolism by long-chain fatty acids. *Prog. Lipid Res.* **53**, 124–144 (2014).

231. Shimizu, T. Lipid Mediators in Health and Disease: Enzymes and Receptors as Therapeutic Targets for the Regulation of Immunity and Inflammation. *Annu. Rev. Pharmacol. Toxicol.* **49**, 123–150 (2009).

232. Resh, M. D. Fatty Acylation of Proteins: The Long and the Short of it. *Prog. Lipid. Res.* **63**, 120–131 (2016).

233. Sezgin, E., Levental, I., Mayor, S. & Eggeling, C. The mystery of membrane organization: composition, regulation and roles of lipid rafts. *Nat. Publ. Gr.* **18**, 361–374 (2017).

234. Zimmerberg, J. & Gawrisch, K. The physical chemistry of biological membranes. *Nat. Chem. Biol.* **2**, 564–567 (2006).

235. Hedin, L. E. & Elofsson, A. An Introduction to Membrane Proteins. *J. Proteome Res.* **10**, 3324–3331 (2011).

236. Protein. *School of Biomedical Sciences Wiki* https://teaching.ncl.ac.uk/bms/wiki/index.php/Protein (2018).

237. Heijne, G. von. Membrane protein assembly: Rules of the game. *BioEssays.* **17**, 25–30 (1995).

238. Möller, S., Croning, M. D. R. & Apweiler, R. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics.* **17**, 646–653 (2001).

239. Taylor, P. D., Toseland, C. P., Attwood, T. K. & Flower, D. R. Alpha helical trans-membrane proteins: Enhanced prediction using a Bayesian approach. *Bioinformation.* **1**, 234–6 (2006).

240. Noinaj, N., Rollauer, S. E. & Buchanan, S. K. The β-barrel membrane protein insertase machinery from Gram-negative bacteria. *Curr. Opin. Struct. Biol.* **31**, 35–42 (2015).

241. HCJ, G. Uber die isolirte Farbung der Schizomyceten in Schnitt- and Trockenpraparaten. *Fortschr. Med.* **2**, 185–189 (1884).

242. Tripathi, N. & Sapra, A. *Gram Staining*. (StatPearls, 2023).

243. 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**, 2561–2570 (1976).

244. Raetz, C. R. H. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71**, 635–700 (2002).

245. Silhavy, T. J., Kahne, D. & Walker, S. The Bacterial Cell Envelope. *Cold Spring Harb Perspect Biol.* **2**, 1–16 (2010).

246. Matsuura, M. Structural modifications of bacterial lipopolysaccharide that facilitate gram-negative bacteria evasion of host innate immunity. *Front. Immunol.* **4**, 1–9 (2013).

247. Zhang, G., Meredith, T. C. & Kahne, D. On the Essentiality of Lipopolysaccharide to Gram-Negative Bacteria. *Curr. Opin. Microbiol.* **16**, 779–785 (2014).

248. Sankaran, K. & Wu, H. C. Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *J. Biol. Chem.* **269**, 19701–19706 (1994).

249. Choi, U. & Lee, C. R. Distinct Roles of Outer Membrane Porins in Antibiotic Resistance and Membrane Integrity in *Escherichia coli*. *Front. Microbiol.* **10**, 1–9 (2019).

250. 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).

251. Miller, S. I. & Salama, N. R. The gram-negative bacterial periplasm: Size matters. *PLoS Biol.* **16**, 1–7 (2018).

252. Wilks, J. C. & Slonczewski, J. L. pH of the cytoplasm and periplasm of *Escherichia coli*: Rapid measurement by green fluorescent protein fluorimetry. *J. Bacteriol.* **189**, 5601–5607 (2007).

253. Vollmer, W., Blanot, D. & De Pedro, M. A. Peptidoglycan structure and architecture. *FEMS Microbiol. Rev.* **32**, 149–167 (2008).

254. Dramsi, S., Magnet, S., Davison, S. & Arthur, M. Covalent attachment of proteins to peptidoglycan. *FEMS Microbiol. Rev.* **32**, 307–320 (2008).

255. Garde, S., Chodisetti, P. K. & Reddy, M. Peptidoglycan: Structure, Synthesis, and Regulation. *EcoSal Plus.* **9**, (2021).

256. Braun, V. Covalent lipoprotein from the outer membrane of *escherichia coli*. *BBA - Rev.*

Biomembr. **415**, 335–377 (1975).

257. Gumbart, J. C. *et al.* Lpp positions peptidoglycan at the AcrA-TolC interface in the AcrAB-TolC multidrug efflux pump. *Biophys. J.* **120**, 3973–3982 (2021).

258. Cook, R. A. Periplasmic proteases of *Escherichia coli*. *Crit. Rev. Biotechnol.* **8**, 159–175 (1988).

259. Krishnamoorthy, G., Tikhonova, E. B. & Zgurskaya, H. I. Fitting periplasmic membrane fusion proteins to inner membrane transporters: Mutations that enable *Escherichia coli* AcrA to function with *Pseudomonas aeruginosa* MexB. *J. Bacteriol.* **190**, 691–698 (2008).

260. Krojer, T. *et al.* Structural basis for the regulated protease and chaperone function of DegP. *Nature.* **453**, 885–890 (2008).

261. Rowlett, V. W. *et al.* The impact of membrane phospholipid alterations in *Escherichia coli* on cellular function. *J. Bacteriol.* **199**, 1–22 (2017).

262. Bogdanov, M. *et al.* Phospholipid distribution in the cytoplasmic membrane of Gram-negative bacteria is highly asymmetric, dynamic, and cell shape-dependent. *Sci. Adv.* **6**, 1–13 (2020).

263. Rossi, R., Lauren;, Y., Agaisse, H. & Payne, S. Cardiolipin Synthesis and Outer Membrane Localization Are Required for *Shigella flexneri* Virulence. *MBio.* **8**, 1–16 (2017).

264. Duong, F., Eichler, J., Price, A., Leonard, M. R. & Wickner, W. Biogenesis of the Gram-Negative Bacterial Envelope. *Cell.* **91**, 567–573 (1997).

265. Duan, K. & Kumar, A. Understanding Bacterial Antibiotic Resistance and Pathogenicity Through Investigating Bacterial Membrane Proteins. *J. Membr. Biol.* **251**, 1–3 (2018).

266. Donlan, R. Biofilms: Microbial Life on Surfaces. *Perspective* **8**, 881–890 (2002).

267. Flemming, H. C. *et al.* Biofilms: An emergent form of bacterial life. *Nat. Rev. Microbiol.* **14**, 563–575 (2016).

268. Fux, C. A., Costerton, J. W., Stewart, P. S. & Stoodley, P. Survival strategies of infectious biofilms. *Trends Microbiol.* **13**, 34–40 (2005).

269. Bhagirath, A. Y., Somayajula, D., Li, Y. & Duan, K. CmpX Affects Virulence in *Pseudomonas aeruginosa* Through the Gac/Rsm Signaling Pathway and by Modulating c-di-GMP Levels. *J. Membr. Biol.* **251**, 35–49 (2018).

270. O'Shea, R. & Moser, H. E. Physicochemical Properties of Antibacterial Compounds: Implications for Drug Discovery. *J. Med. Chem.* **51**, 2871–2878 (2008).

271. Ziervogel, B. K. & Roux, B. The Binding of Antibiotics in OmpF Porin. *Structure.* **21**, 76–87 (2013).

272. Schlegel, S., Hjelm, A., Baumgarten, T., Vikström, D. & De Gier, J. W. Bacterial-based membrane protein production. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843**, 1739–1749 (2014).

273. Dalbey, R. E., Wang, P. & Kuhn, A. Assembly of bacterial inner membrane proteins. *Annu. Rev. Biochem.* **80**, 161–187 (2011).

274. Bornemann, T., Jöckel, J., Rodnina, M. V. & Wintermeyer, W. Signal sequence-independent membrane targeting of ribosomes containing short nascent peptides within the exit tunnel. *Nat. Struct. Mol. Biol.* **15**, 494–499 (2008).

275. Cross, B. C. S., Sinning, I., Luirink, J. & High, S. Delivering proteins for export from the cytosol.

Nat. Rev. Mol. Cell Biol. **10**, 255–264 (2009).

276. Ullers, R. S. *et al.* Interplay of signal recognition particle and trigger factor at L23 near the nascent chain exit site on the *Escherichia coli* ribosome. *J. Cell Biol.* **161**, 679–684 (2003).

277. Luirink, J., Yu, Z., Wagner, S. & De Gier, J. W. Biogenesis of inner membrane proteins in *Escherichia coli*. *Biochim. Biophys. Acta - Bioenerg.* **1817**, 965–976 (2012).

278. Janda, C. Y. *et al.* Recognition of a signal peptide by the signal recognition particle. *Nature.* **465**, 507–510 (2010).

279. Schibich, D. *et al.* Global profiling of SRP interaction with nascent polypeptides. *Nat. Publ. Gr.* **536**, 219–223 (2016).

280. Saraogi, I., Akopian, D. & Shan, S. O. A tale of two GTPases in cotranslational protein targeting. *Protein Sci.* **20**, 1790–1795 (2011).

281. Angelini, S., Deitermann, S. & Koch, H. G. FtsY, the bacterial signal-recognition particle receptor, interacts functionally and physically with the SecYEG translocon. *EMBO Rep.* **6**, 476–481 (2005).

282. Nevo-Dinur, K., Nussbaum-Shochat, A., Ben-Yehuda, S. & Amster-Choder, O. Translation-independent localization of mRNA in *E. coli*. *Science.* **331**, 1081–1084 (2011).

283. Bochkareva, E., Seluanov, A., Bibi, E. & Girshovich, A. Chaperonin-promoted post-translational membrane insertion of a multispanning membrane protein lactose permease. *J. Biol. Chem.* **271**, 22256–22261 (1996).

284. Deaton, J. *et al.* Functional bacteriorhodopsin is efficiently solubilized and delivered to, membranes by the chaperonin GroEL. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 2281–2286 (2004).

285. Lycklama a Nijeholt, J. A. & Driessen, A. J. M. The bacterial Sec-translocase: Structure and mechanism. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 1016–1028 (2012).

286. Dalal, K. & Duong, F. The SecY complex: Conducting the orchestra of protein translocation. *Trends Cell Biol.* **21**, 506–514 (2011).

287. Berg, B. Van Den *et al.* X-ray structure of a protein-conducting channel. *Nature.* **427**, 36–44 (2004).

288. Gumbart, J. & Schulten, K. The Roles of Pore Ring and Plug in the SecY Protein-conducting Channel. *J. Gen. Physiol.* **136**, 709–719 (2008).

289. Saparov, S. M. *et al.* Determining the Conductance of the SecY Protein Translocation Channel for Small Molecules. *Mol. Cell.* **26**, 501–509 (2007).

290. Schatz, P. J., Biekerl, K. L., Ottemann, K. M., Silhavy, T. J. & Beckwith, J. One of three transmembrane stretches is sufficient the functioning of the SecE protein, a membrane component of the *E. coli* secretion machinery. *EMBO J.* **1**, 1749–1757 (1991).

291. Pohlschroder, M., Murphy, C. & Beckwith, J. *In Vivo* Analyses of Interactions between SecE and SecY, Core Components of the *Escherichia coli* Protein Translocation Machinery. *J. Biol. Chem.* **271**, 19908–19914 (1996).

292. Plessis, D. J. F., Nouwen, N. & Driessen, A. J. M. The Sec translocase. *Biochim. Biophys. Acta.* **1808**, 851–865 (2011).

293. Hanada, M., Nishiyama, K. I., Mizushima, S. & Tokuda, H. Reconstitution of an Efficient Protein Translocation Machinery Comprising SecA and the Three Membrane Proteins, SecY,

SecE, and SecG (p12). *J. Biol. Chem.* **269**, 23625–23631 (1994).

294. Nishiyama, K. & Mizushimal, S. A novel membrane protein involved in protein translocation across the cytoplasmic membrane of *Escherichia coli*. *EMBO J.* **12**, 3409–3415 (1993).

295. Samuelson, J. C. *et al.* YidC mediates membrane protein insertion in bacteria. *Nature.* **406**, 637–641 (2000).

296. Han, L. *et al.* Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nat. Struct. Mol. Biol.* **23**, 192–196 (2016).

297. Hegde, R. S. & Bernstein, H. D. The surprising complexity of signal sequences. *Trends Biochem. Sci.* **31**, 563–571 (2006).

298. Rollauer, S. E., Sooreshjani, M. A., Noinaj, N. & Buchanan, S. K. Outer membrane protein biogenesis in Gram-negative bacteria. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, (2015).

299. Nakamoto, H. & Bardwell, J. C. A. Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochim. Biophys. Acta - Mol. Cell Res.* **1694**, 111–119 (2004).

300. Vertommen, D., Ruiz, N., Leverrier, P., Silhavy, T. J. & Collet, J. F. Characterization of the role of the *escherichia coli* periplasmic chaperone SurA using differential proteomics. *Proteomics.* **9**, 2432–2443 (2009).

301. Walton, T. A. & Sousa, M. C. Crystal structure of Skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation. *Mol. Cell.* **15**, 367–374 (2004).

302. Ge, X. *et al.* DegP primarily functions as a protease for the biogenesis of β-barrel outer membrane proteins in the Gram-negative bacterium *Escherichia coli*. *FEBS J.* **281**, 1226–1240 (2014).

303. Hussain, S., Peterson, J. H. & Bernstein, H. D. Reconstitution of bam complex-mediated assembly of a trimeric porin into proteoliposomes. *MBio.* **12**, 1–16 (2021).

304. Wu, T. *et al.* Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell.* **121**, 235–245 (2005).

305. Noinaj, N. *et al.* Structural insight into the biogenesis of β-barrel membrane proteins. *Nature.* **501**, 385–390 (2013).

306. 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).

307. C Reygaert, W. An overview of the antimicrobial resistance mechanisms of bacteria. *AIMS Microbiol.* **4**, 482–501 (2018).

308. Clardy, J., Fischbach, M. A. & Currie, C. R. The natural history of antibiotics. *Curr. Biol.* **19**, 1–8 (2009).

309. Ventola, C. L. The Antibiotic Resistance Crisis Part 1: Causes and Threats. *Pharmacol. Ther.* **40**, 181–224 (2015).

310. Griffith, M., Postelnick, M. & Scheetz, M. Antimicrobial stewardship programs: methods of operation and suggested outcomes. *Expert Rev. Anti. Infect. Ther.* **10**, 63–73 (2012).

311. Landers, T. F., Cohen, B., Wittum, T. E. & Larson, E. L. A review of antibiotic use in food animals: Perspective, policy, and potential. *Public Health Rep.* **127**, 4–22 (2012).

312. Wegener, H. Antibiotic resistance - linking human and animal health. in *Improving Food*

Safety Through a One Health Approach 331–349 (Washington: National Academy of Sciences, 2012).

313. Darby, E. M. *et al.* Molecular mechanisms of antibiotic resistance revisited. *Nat. Rev. Microbiol.* **21**, 280–295 (2022).

314. Munita, J. M. & Arias, C. A. Mechanisms of Antibiotic Resistance. *Microbiol Spectr.* **4**, (2016).

315. Pagès, J., James, C. E. & Winterhalter, M. The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat. Rev. Microbiol.* **6**, 893–903 (2008).

316. Bébéar, C. M. & Pereyre, S. Mechanisms of drug resistance in *Mycoplasma pneumoniae*. *Curr. Drug Targets. Infect. Disord.* **5**, 263–271 (2005).

317. Mishra, N. N. *et al.* Daptomycin resistance in enterococci is associated with distinct alterations of cell membrane phospholipid content. *PLoS One.* **7**, 1–10 (2012).

318. Draper, P. The outer parts of the mycobacterial envelope as permeability barriers. *Front. Biosci.* **3**, 1253–1261 (1998).

319. Fu, L. M. & Fu-Liu, C. S. Is *Mycobacterium tuberculosis* a closer relative to Gram-positive or Gram – negative bacterial pathogens? *Tuberculosis* **82**, 85–90 (2002).

320. Bush, K. & Bradford, P. A. β-Lactams and β-Lactamase Inhibitors An Overview. *Cold Spring Harb. Perspect. Medicine.* **6**, 1–22 (2016).

321. Bush, K. & Jacoby, G. A. Updated functional classification of β-lactamases. *Antimicrob. Agents Chemother.* **54**, 969–976 (2010).

322. Markley, J. L. & Wencewicz, T. A. Tetracycline-inactivating enzymes. *Front. Microbiol.* **9**, 1–22 (2018).

323. Peterson, E. & Kaur, P. Antibiotic resistance mechanisms in bacteria: Relationships between resistance determinants of antibiotic producers, environmental bacteria, and clinical pathogens. *Front. Microbiol.* **9**, 1–21 (2018).

324. Ramirez, M. S. & Tolmasky, M. E. Aminoglycoside Modifying Enzymes. *Drug Resist. Updat.* **13**, 151–171 (2010).

325. Bush, N. G., Diez-Santos, I., Abbott, L. R. & Maxwell, A. Contributions to Antibiotic Resistance. *Molecules.* **25**, 5662 (2020).

326. Bhujbalrao, R. & Anand, R. Deciphering Determinants in Ribosomal Methyltransferases That Confer Antimicrobial Resistance. *J. Am. Chem. Soc.* **141**, 1425–1429 (2019).

327. Wielders, C. L. C., Fluit, A. C., Brisse, S., Verhoef, J. & Schmitz, F. J. *mecA* Gene Is Widely Disseminated in *Staphylococcus aureus* Population. *J. Clin. Microbiol.* **40**, 3970–3975 (2002).

328. Flensburg, J. & Skold, O. Massive overproduction of dihydrofolate reductase in bacteria as a response to the use of trimethoprim. *Eur. J. Biochem.* **162**, 473–476 (1987).

329. Sun, J., Deng, Z. & Yan, A. Bacterial multidrug efflux pumps: Mechanisms, physiology and pharmacological exploitations. *Biochem. Biophys. Res. Commun.* **453**, 254–267 (2014).

330. Tsachouridou, O. *et al.* Mortality due to Multidrug-Resistant Gram-Negative Bacteremia in an Endemic Region: No Better than a Toss of a Coin. *Microorganisms.* **11**, 1–16 (2023).

331. Alav, I. *et al.* Structure, Assembly and Function of Tripartite Effux and Type 1 Secretion Systems in Gram-Negative Bacteria. *Chem. Rev.* **121**, 5479–5596 (2021).

332. Ma, D. *et al.* Molecular cloning and characterization of *acrA* and *acrE* genes of *Escherichia coli*. *J. Bacteriol.* **175**, 6299–6313 (1993).

333. Poole, K., Krebes, K. & Neshat, S. Multiple Antibiotic Resistance in *Pseudomonas aeruginosa*: Evidence for Involvement of an Efflux Operon. *J. Bacteriol.* **175**, 7363–7372 (1993).

334. Kobylka, J., Kuth, M. S., Müller, R. T. & Geertsma, E. R. AcrB: a mean, keen, drug efflux machine. *Antimicrob. Ther. Rev.* **1459**, 38–68 (2020).

335. Nikaido, H. & Biology, C. RND transporters in the living world. *Res. Microbiol.* **169**, 363–371 (2019).

336. Fernando, D. M. & Kumar, A. Resistance-Nodulation-Division multidrug efflux pumps in Gram-negative bacteria: Role in virulence. *Antibiotics.* **2**, 163–181 (2013).

337. Zgurskaya, H. I., Weeks, J. W., Ntreh, A. T., Nickels, L. M. & Wolloscheck, D. Mechanism of coupling drug transport reactions located in two different membranes. *Front. Microbiol.* **6**, 1–13 (2015).

338. Klenotic, P. A., Moseng, M. A., Morgan, C. E. & Yu, E. W. Structural and functional diversity of RND transporters. *Chem. Rev.* **121**, 5378–5416 (2021).

339. Poole, K. Efflux-mediated multiresistance in Gram-negative bacteria. *Clin. Microbiol. Infect.* **10**, 12–26 (2004).

340. Tsukagoshi, N. & Aono, R. Entry into and release of solvents by *Escherichia coli* in an organic-aqueous two-liquid-phase system and substrate specificity of the AcrAB-TolC solvent-extruding pump. *J. Bacteriol.* **182**, 4803–4810 (2000).

341. White, D. G., Goldman, J. D., Demple, B. & Levy, S. B. Role of the *acrAB* locus in organic solvent tolerance mediated by expression of *marA*, *soxS*, or *robA* in *Escherichia coli*. *J. Bacteriol.* **179**, 6122–6126 (1997).

342. Join-Lambert, O. F. *et al.* Differential Selection of Multidrug Efflux Mutants by Trovafloxacin and Ciprofloxacin in an Experimental Model of *Pseudomonas aeruginosa* Acute Pneumonia in Rats. *Antimicrob. Agents Chemother.* **45**, 571–576 (2001).

343. Hirakata, Y., Kondo, A., Hoshino, K., Yano, H. & Arai, K. Efflux pump inhibitors reduce the invasiveness of *Pseudomonas aeruginosa*. *Int. J. Antimicrob. Agents.* **34**, 343–346 (2009).

344. Yang, L., Chen, L., Shen, L., Surette, M. & Duan, K. Inactivation of MuxABC-OmpB Transporter System in *Pseudomonas aeruginosa* Leads to Increased Ampicillin and Carbenicillin Resistance and Decreased Virulence. *J. Microbiol.* **49**, 107–114 (2011).

345. Linares, J. F. *et al.* Overexpression of the multidrug efflux pumps MexCD-OprJ and MexEF-OprN is associated with a reduction of type III secretion in *Pseudomonas aeruginosa*. *J. Bacteriol.* **187**, 1384–1391 (2005).

346. Bina, X. R., Provenzano, D., Nguyen, N. & Bina, J. E. *Vibrio cholerae* RND Family Efflux Systems Are Required for Antimicrobial Resistance, Optimal Virulence Factor Production, and Colonization of the Infant Mouse Small Intestine. *Infect. Immun.* **76**, 3595–3605 (2008).

347. Taylor, D. L., Bina, X. R. & Bina, J. E. *Vibrio cholerae vexH* encodes a multiple drug efflux pump that contributes to the production of cholera toxin and the toxin co-regulated pilus. *PLoS One.* **7**, 1–9 (2012).

348. Baucheron, S., Mouline, C., Praud, K., Chaslus-dancla, E. & Cloeckaert, A. TolC but not AcrB is essential for multidrug-resistant *Salmonella enterica* serotype Typhimurium colonization of

chicks. *J. Antimicrob. Chemother.* **55**, 707–712 (2005).

349. Lacroix, F. J. C. *et al. Salmonella typhimurium acrB*-like gene: identification and role in resistance to biliary salts and detergents and in murine infection. *FEMS Microbiol. Lett.* **135**, 161–167 (1996).

350. Stoitsova, S. O., Braun, Y., Ullrich, M. S. & Weingart, H. Characterization of the RND-Type Multidrug Efflux Pump MexAB-OprM of the Plant Pathogen *Pseudomonas syringae*. *Appl. Environ. Microbiol.* **74**, 3387–3393 (2008).

351. Koutsolioutsou, A., Pen, S. & Demple, B. Constitutive *soxR* Mutations Contribute to Multiple-Antibiotic Resistance in Clinical *Escherichia coli* Isolates. *Antimicrob. Agents Chemother.* **49**, 2746–2752 (2005).

352. Schembri, M. A., Kjærgaard, K. & Klemm, P. Global gene expression in *Escherichia coli* biofilms. *Mol. Microbiol.* **48**, 253–267 (2003).

353. Zhang, Y. *et al.* The Multidrug Efflux Pump MdtEF Protects against Nitrosative Damage during the Anaerobic Respiration in *Escherichia coli*. *J. Biol. Chem.* **286**, 26576–26584 (2011).

354. Alav, I., Sutton, J. M. & Rahman, K. M. Role of bacterial efflux pumps in biofilm formation. *J. Antimicrob. Chemother.* **73**, 2003–2020 (2018).

355. Weston, N., Sharma, P., Ricci, V. & Piddock, L. J. V. Regulation of the AcrAB-TolC efflux pump in Enterobacteriaceae. *Res. Microbiol.* **169**, 425–431 (2018).

356. Rosenberg, E. Y., Bertenthal, D., Nilles, M. L., Bertrand, K. P. & Nikaido, H. Bile salts and fatty acids induce the expression of *Escherichia coli* AcrAB multidrug efflux pump through their interaction with Rob regulatory protein. *Mol. Microbiol.* **48**, 1609–1619 (2003).

357. Li, M. *et al.* Crystal Structure of the Transcriptional Regulator AcrR from *Escherichia coli*. *J. Mol. Biol.* **374**, 591–603 (2007).

358. Li, X., Plésiat, P. & Nikaido, H. The Challenge of Efflux-Mediated Antibiotic Resistance in Gram-Negative Bacteria. *Clin. Microbiol. Rev.* **28**, 337–418 (2015).

359. Alekshun, M. N. & Levy, S. B. The *mar* regulon: multiple resistance to antibiotics and other toxic chemicals. *Trends Microbiol.* **7**, 410–413 (1999).

360. Cohen, S. P., Hachler, H. & Levy, Stuart, B. Genetic and Functional Analysis of the Multiple Antibiotic Resistance (*mar*) Locus in *Escherichia coli*. *J. Bacteriol.* **175**, 1484–1492 (1993).

361. Rhee, S., Martin, R. G., Rosner, J. L. & Davies, D. R. A novel DNA-binding motif in MarA: The first structure for an AraC family transcriptional activator. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 10413–10418 (2023).

362. Zhang, A., Rosner, J. L. & Martin, R. G. Transcriptional activation by MarA, SoxS and Rob of two *tolC* promoters using one binding site: a complex promoter configuration for *tolC* in *Escherichia coli*. *Mol. Microbiol.* **69**, 1450–1455 (2008).

363. Miller, P. F., Gambino, L. F., Sulavik, M. C. & Gracheck, S. J. Genetic Relationship between *soxRS* and *mar* Loci in Promoting Multiple Antibiotic Resistance in *Escherichia coli*. *Antimicrob. Agents Chemother.* **38**, 1773–1779 (1994).

364. Nakajima, H., Kobayashi, K. E. I., Kobayashi, M., Asako, H. & Aono, R. Overexpression of the *robA* Gene Increases Organic Solvent Tolerance and Multiple Antibiotic and Heavy Metal Ion Resistance in *Escherichia coli*. *Appl. Environ. Microbiol.* **61**, 2302–2307 (1995).

365. Grimsey, E. M., Weston, N., Ricci, V., Stone, J. W. & Piddock, L. J. V. Overexpression of RamA, Which Regulates Production of the Mutation Rate and Influences Drug Resistance Phenotype. *Antimicrob. Agents Chemother.* **64**, 1–8 (2020).

366. Du, D. *et al.* Structure of the AcrAB-TolC multidrug efflux pump. *Nature.* **509**, 512–515 (2014).

367. Tikhonova, E. B., Yamada, Y. & Zgurskaya, H. I. Sequential Mechanism of Assembly of Multidrug Efflux Pump AcrAB-TolC. *Chem. Biol.* **18**, 454–463 (2011).

368. Nishino, K., Nikaido, E. & Yamaguchi, A. Regulation and physiological function of multidrug efflux pumps in *Escherichia coli* and *Salmonella*. *Biochim. Biophys. Acta - Proteins Proteomics.* **1794**, 834–843 (2009).

369. Symmons, M. F., Bokma, E., Koronakis, E., Hughes, C. & Koronakis, V. The assembled structure of a complete tripartite bacterial multidrug efflux pump. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, 26–31 (2009).

370. Hinchliffe, P., Symmons, M. F., Hughes, C. & Koronakis, V. Structure and operation of bacterial tripartite pumps. *Annu. Rev. Microbiol.* **67**, 221–242 (2013).

371. Du, D., Veen, H. W. Van, Murakami, S., Pos, K. M. & Luisi, B. F. Structure, mechanism and cooperation of bacterial multidrug transporters. *Curr. Opin. Struct. Biol.* **33**, 76–91 (2015).

372. Daury, L. *et al.* Tripartite assembly of RND multidrug efflux pumps. *Nat. Commun.* **7**, 1–8 (2016).

373. Kim, J. S. *et al.* Structure of the tripartite multidrug efflux pump AcrAB-TolC suggests an alternative assembly mode. *Mol. Cells.* **38**, 180–186 (2015).

374. Elkins, C. A. & Nikaido, H. 3D structure of AcrB: the archetypal multidrug efflux transporter of *Escherichia coli* likely captures substrates from periplasm. *Drug Resist. Updat.* **6**, 9–13 (2003).

375. Rajapaksha, P., Pandeya, A. & Wei, Y. Probing the Dynamics of AcrB Through Disulfide Bond Formation. *ACS Omega.* **5**, 21844–21852 (2020).

376. Zwama, M. *et al.* Multiple entry pathways within the efflux transporter AcrB contribute to multidrug recognition. *Nat. Commun.* **9**, 1–9 (2018).

377. Zwama, M. & Yamaguchi, A. Molecular mechanisms of AcrB-mediated multidrug export. *Res. Microbiol.* **169**, 372–383 (2018).

378. Tam, H. *et al.* Binding and Transport of Carboxylated Drugs by the Multidrug Transporter AcrB. *J. Mol. Biol.* **432**, 861–877 (2020).

379. Alav, I., Bavro, V. N. & Blair, J. M. A. A role for the periplasmic adaptor protein AcrA in vetting substrate access to the RND efflux transporter AcrB. *Sci. Rep.* **12**, 1–13 (2022).

380. Müller, R. T. *et al.* Switch Loop Flexibility Affects Substrate Transport of the AcrB Efflux Pump. *J. Mol. Biol.* **429**, 3863–3874 (2017).

381. Takatsuka, Y., Chen, C. & Nikaido, H. Mechanism of recognition of compounds of diverse structures by the multidrug efflux pump AcrB of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6559–6565 (2010).

382. Cha, H. J., Müller, R. T. & Pos, K. M. Switch-loop flexibility affects transport of large drugs by the promiscuous AcrB multidrug efflux transporter. *Antimicrob. Agents Chemother.* **58**, 4767–4772 (2014).

383. Zwama, M. *et al.* Hoisting-loop in bacterial multidrug exporter AcrB is a highly flexible hinge

that enables the large motion of the subdomains. *Front. Microbiol.* **8**, 1–8 (2017).

384. Qiu, W. *et al.* Structure and activity of lipid bilayer within a membrane-protein transporter. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 12985–12990 (2018).

385. Hobbs, E. C., Yin, X., Paul, B. J., Astarita, J. L. & Storz, G. Conserved small protein associates with the multidrug efflux pump AcrB and differentially affects antibiotic resistance. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16696–16701 (2012).

386. Seeger, M. A. *et al.* Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science.* **313**, 1295–1298 (2006).

387. Murakami, S., Nakashima, R., Yamashita, E., Matsumoto, T. & Yamaguchi, A. Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature.* **443**, 173–179 (2006).

388. Vargiu, A. V. *et al.* Effect of the F610A mutation on substrate extrusion in the AcrB transporter: Explanation and rationale by molecular dynamics simulations. *J. Am. Chem. Soc.* **133**, 10704–10707 (2011).

389. McNeil, H. E. *et al.* Identification of binding residues between periplasmic adapter protein (PAP) and RND efflux pumps explains PAP-pump promiscuity and roles in antimicrobial resistance. *PLoS Pathog.* **15**, 1–28 (2019).

390. Li, Y., Cross, T. S. & Tobias, D. Analysis of AcrB in *Klebsiella pneumoniae* reveals natural variants promoting enhanced multidrug resistance. *Res. Microbiol.* **173**, 1–8 (2022).

391. Yamasaki, S., Zwama, M., Yoneda, T., Nishino, M. H.- & Nishino, K. Drug resistance and physiological roles of RND multidrug efflux pumps in *Salmonella enterica*, *Escherichia coli* and *Pseudomonas aeruginosa*. *Microbiology.* **169**, 1–20 (2023).

392. Pugh, H. L., Connor, C., Siasat, P., Mcnally, A. & Blair, J. M. A. *E.coli* ST11 (O157:H7) does not encode a functional AcrF efflux pump. *Microbiology.* **169**, 1–7 (2023).

393. Anes, J., McCusker, M. P., Fanning, S. & Martins, M. The ins and outs of RND efflux pumps in *Escherichia coli. Front. Microbiol.* **6**, 1–14 (2015).

394. Su, C. *et al.* Crystal structure of the membrane fusion protein CusB from *Escherichia coli*. *J. Mol. biol.* **393**, 342–355 (2010).

395. Zwama, M. & Nishino, K. Ever-Adapting RND Efflux Pumps in Gram-Negative Multidrug-Resistant Pathogens: A Race against Time. *Antibiotics.* **10**, 1–32 (2021).

396. Zgurskaya, H. I. & Nikaido, H. AcrA is a Highly Asymmetric Protein Capable of Spanning the Periplasm. *J. Mol. Biol.* **285**, 409–420 (1999).

397. Mikolosko, J., Bobyk, K., Zgurskaya, H. I. & Ghosh, P. Conformational flexibility in the multidrug efflux system protein AcrA. *Structure.* **14**, 577–587 (2006).

398. Hazel, A. J. *et al.* Conformational Dynamics of AcrA Govern Multidrug Efflux Pump Assembly. *ACS Infect. Dis.* **5**, 1926–1935 (2019).

399. Pos, K. M. Drug transport mechanism of the AcrB efflux pump. *Biochim. Biophys. Acta - Proteins Proteomics.* **1794**, 782–793 (2009).

400. Ge, Q., Yamadah, Y. & Zgurskaya, E. The C-terminal domain of AcrA is essential for the assembly and function of the multidrug efflux pump AcrAB-TolC. *J. Bacteriol.* **191**, 4365–4371 (2009).

401. Hayashi, K. *et al.* AcrB-AcrA fusion proteins that act as multidrug efflux transporters. *J. Bacteriol.* **198**, 332–342 (2016).

402. Johnson, J. M. & Church, G. M. Alignment and Structure Prediction of Divergent Protein Families: Periplasmic and Outer Membrane Proteins of Bacterial Efflux Pumps. *J. Mol. Biol.* **287**, 695–715 (1999).

403. Abdali, N. *et al.* Reviving antibiotics: Efflux pump inhibitors that interact with AcrA, a membrane fusion protein of the AcrAB-TolC multidrug efflux pump. *ACS Infect. Dis.* **3**, 89–98 (2017).

404. De Angelis, F. *et al.* Metal-induced conformational changes in ZneB suggest an active role of membrane fusion proteins in efflux resistance systems. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11038–11043 (2010).

405. Zgurskaya, H. I., Yamada, Y., Tikhonova, E. B., Ge, Q. & Krishnamoorthy, G. Structural and functional diversity of bacterial membrane fusion proteins. *Biochim. Biophys. Acta.* **1794**, 794–807 (2009).

406. Alav, I., Bavro, V. N. & Blair, J. M. A. Interchangeability of periplasmic adaptor proteins AcrA and AcrE in forming functional efflux pumps with AcrD in *Salmonella enterica* serovar Typhimurium. *J. Antimicrob. Chemother.* **76**, 2558–2564 (2021).

407. Ueda, O. *et al.* Sixteen Homologs of the Mex-Type Multidrug Resistance Efflux Pump in *Bacteroides fragilis*. *Antimicrob. Agents Chemother.* **49**, 2807–2815 (2005).

408. Symmons, M. F., Marshall, R. L., Bavro, V. N. & Shafer, W. W. Architecture and roles of periplasmic adaptor proteins in tripartite efflux assemblies. *Front. Microbiol.* **6**, 1–20 (2015).

409. Higgins, M. K., Bokma, E., Koronakis, E., Hughes, C. & Koronakis, V. Structure of the periplasmic component of a bacterial drug efflux pump. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9994–9999 (2004).

410. Zgurskaya, H. I., Rybenkov, V. V, Krishnamoorthy, G. & Leus, I. V. Trans-envelope multidrug efflux pumps of Gram-negative bacteria and their synergism with the outer membrane barrier. *Res. Microbiol.* **169**, 351–356 (2018).

411. Tikhonova, E. B., Dastidar, V., Rybenkov, V. V & Zgurskaya, H. I. Kinetic control of TolC recruitment by multidrug efflux complexes. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 16416–16421 (2009).

412. Ntreh, A. T., Weeks, J. W., Nickels, L. M. & Zgurskaya, H. I. Opening the Channel: the Two Functional Interfaces of *Pseudomonas aeruginosa* OpmH with the Triclosan Efflux Pump TriABC. *J. Bacteriol.* **189**, 3176–3185 (2016).

413. Luisi, B., Koronakis, V., Hughes, C., Sharff, A. & Koronakis, E. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature.* **405**, 914–919 (2000).

414. Koronakis, V. TolC - the bacterial exit duct for proteins and drugs. *FEBS Lett.* **555**, 66–71 (2003).

415. Gilardi, A. *et al.* Biophysical characterization of *E. coli* TolC interaction with the known blocker hexaamminecobalt. *Biochim. Biophys. Acta - Gen. Subj.* **1861**, 2702–2709 (2017).

416. Zgurskaya, H. I., Krishnamoorthy, G., Ntreh, A., Lu, S. & Bacteria, G. Mechanism and function of the outer membrane channel TolC in multidrug resistance and physiology of enterobacteria. *Front. Microbiol.* **2**, 1–13 (2011).

417. Sulavik, M. C. *et al.* Antibiotic Susceptibility Profiles of *Escherichia coli* Strains Lacking Multidrug Efflux Pump Genes. *Antimicrob. Agents Chemother.* **45**, 1126–1136 (2001).

418. Koronakis, V., Eswaran, J. & Hughes, C. Structure and Function of TolC: The Bacterial Exit Duct for Proteins and Drugs. *Annu. Rev. Biochem.* **73**, 467–489 (2004).

419. Jo, J. T. H., Brinkman, F. S. L. & Hancock, R. E. W. Aminoglycoside Efflux in *Pseudomonas aeruginosa*: Involvement of Novel Outer Membrane Proteins. *Antimicrob. Agents Chemother.* **47**, 1101–1111 (2003).

420. López, C. A., Travers, T., Pos, K. M., Zgurskaya, H. I. & Gnanakaran, S. Dynamics of Intact MexAB-OprM Efflux Pump: Focusing on the MexA-OprM Interface. *Sci. Rep.* **7**, 1–17 (2017).

421. Xu, Y. *et al.* The tip region of the MacA α-hairpin is important for the binding to TolC to the *Escherichia coli* MacAB–TolC pump. *Biochem. Biophys. Res. Commun.* **394**, 962–965 (2010).

422. Lee, M. *et al.* Membrane Fusion Proteins of Type I Secretion System and Tripartite Efflux Pumps Share a Binding Motif for TolC in Gram-Negative Bacteria. *PLoS One.* **7**, 1–7 (2012).

423. Bavro, V. N. *et al.* Assembly and Channel Opening in a Bacterial Drug Efflux Machine. *Mol. Cell.* **30**, 114–121 (2008).

424. Touzé, T. *et al.* Interactions underlying assembly of the *Escherichia coli* AcrAB-TolC multidrug efflux system. *Mol. Microbiol.* **53**, 697–706 (2004).

425. Lewis, B. R. *et al.* Conformational restriction shapes the inhibition of a multidrug efflux adaptor protein. *Nat. Commun.* **14**, 1–14 (2023).

426. Jesin, J. A., Stone, T. A., Mitchell, C. J., Reading, E. & Deber, C. M. Peptide-Based Approach to Inhibition of the Multidrug Resistance Efflux Pump AcrB. *Biochemsitry* **59**, 3973–3981 (2020).

427. Wang, Z. *et al.* An allosteric transport mechanism for the AcrAB-TolC multidrug efflux pump. *Elife.* **6**, 1–19 (2017).

428. Chen, M. *et al. In situ* structure of the AcrAB-TolC efflux pump at subnanometer resolution. *Structure.* **30**, 107–113 (2022).

429. Zgurskaya, H. I. & Nikaido, H. Cross-linked complex between oligomeric periplasmic lipoprotein AcrA and the inner-membrane-associated multidrug efflux pump AcrB from *Escherichia coli*. *J. Bacteriol.* **182**, 4264–4267 (2000).

430. Stockbridge, R. B. The application of Poisson distribution statistics in ion channel reconstitution to determine oligomeric architecture. *Methods Enzymol.* **652**, 321–340 (2022).

431. Atwood, C. S. *et al.* Copper Mediates Dityrosine Cross-Linking of Alzheimer's Amyloid-β. *Biochemistry.* **43**, 560–568 (2004).

432. Tulumello, D. V & Deber, C. M. SDS Micelles as a Membrane-Mimetic Environment for Transmembrane Segments. *Biochemistry.* **48**, 12096–12103 (2009).

433. Kumar, M. A. & Davidson, V. L. Methods to identify and avoid artifactual formation of interchain disulfide bonds when analyzing proteins by SDS-PAGE. *Biotechniques.* **12**, 198–202 (1992).

434. Watt, A. D. *et al.* Oligomers, fact or artefact? SDS-PAGE induces dimerization of β-amyloid in human brain samples. *Acta Neuropathol.* **125**, 549–564 (2013).

435. Pujol-pina, R., Vilaprinyó-pascual, S., Mazzucato, R. & Arcella, A. SDS-PAGE analysis of Aβ oligomers is disserving research into Alzheimer's disease: appealing for ESI-IM-MS. *Sci. Rep.* **5**,

1–13 (2015).

436. Natalello, A., Santambrogio, C. & Grandori, R. Are Charge-State Distributions a Reliable Tool Describing Molecular Ensembles of Intrinsically Disordered Proteins by Native MS? *J. Am. Soc. Mass Spectrom.* **28**, 21–28 (2017).

437. Beveridge, R., Chappuis, Q., Macphee, C. & Barran, P. Mass spectrometry methods for intrinsically disordered proteins. *Analyst.* **138**, 32–42 (2013).

438. Hoffman, M. D. & Kast, J. Mass spectrometric characterization of lipid-modified peptides for the analysis of acylated proteins. *J. mass Spectrom.* **41**, 229–241 (2006).

439. Wang, Z. & Schey, K. L. Proteomic analysis of s-palmitoylated proteins in ocular lens reveals palmitoylation of AQP5 and MP20. *Investig. Ophthalmol. Vis. Sci.* **59**, 5648–5658 (2018).

440. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* **5**, 1–17 (2014).

441. Drisdel, R. C. & Green, W. N. Labeling and quantifying sites of protein palmitoylation. *Biotechniques.* **36**, 276–285 (2004).

442. Hamuro, Y. & Coales, S. J. Optimisation of Feasibility Stage for Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **29**, 623–629 (2018).

443. Fang, J., Rand, K. D., Beunning, P. J. & Engen, J. R. False EX1 signatures caused by sample carryover during HX MS analyses. *Int. J. Mass Spectrom.* **302**, 19–25 (2011).

444. Smit, J. H. *et al.* Probing Universal Protein Dynamics Using Hydrogen-Deuterium Exchange Mass Spectrometry-Derived Residue-Level Gibbs Free Energy. *Anal. Chem.* **93**, 12840–12847 (2021).

445. Fowler, M. L. *et al.* Using hydrogen deuterium exchange mass spectrometry to engineer optimised constructs for crystallization of protein complexes: Case study of PI4KIIIβ with Rab11. *Protein Sci.* **25**, 826–839 (2016).

446. Hageman, T. S. & Weis, D. D. Reliable Identification of Significant Differences in Differential Hydrogen Exchange-Mass Spectrometry Measurements Using a Hybrid Significance Testing Approach. *Anal. Chem.* **91**, 8008–8016 (2019).

447. Zhang, N., Yu, X., Zhang, X. & Arcy, S. D. Structural bioinformatics HD-eXplosion: visualization of hydrogen – deuterium exchange data as chiclet and volcano plots with statistical filtering. *Bioinformatics.* **37**, 1926–1927 (2021).

448. Pettersen, E. F. *et al.* UCSF Chimera — A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

449. Bagai, I., Liu, W., Rensing, C., Blackburn, N. J. & McEvoy, M. M. Substrate-linked conformational change in the periplasmic component of a Cu(I)/Ag(I) efflux system. *J. Biol. Chem.* **282**, 35695–35702 (2007).

450. Tang, S. & Yang, J. J. Encyclopedia of Metalloproteins. in *Encyclopedia of Metalloproteins* 1243–1250 (Springer Science, 2013).

451. Aptekmann, A. A. *et al.* Sequence analysis mebipred: identifying metal-binding potential in protein sequence. *Bioinformatics.* **38**, 3532–3540 (2022).

452. Mohamadi, A., Jin, L. & Wang, J. An ensemble 3D deep-learning model to predict protein metal-binding site. *Cell Reports Phys. Sci.* **3**, 101046 (2022).

453. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–2890 (2007).

454. Micsonai, A. *et al.* BeStSel: A web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res.* **46**, 315–322 (2018).

455. Consalvi, V. *et al.* Thermal unfolding and conformational stability of the recombinant domain II of glutamate dehydrogenase from the hyperthermophile *Thermotoga maritima*. *Protein Eng.* **13**, 501–507 (2000).

456. Anderson, R. A., Bosron, W. F., Kennedy, F. S. & Vallee, B. L. Role of Magnesium in *Escherichia coli* Alkaline Phosphatase. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2989–2993 (1975).

457. Weeks, J. W., Nickels, L. M., Ntreh, A. T. & Zgurskaya, H. I. Non-equivalent roles of two periplasmic subunits in the function and assembly of triclosan pump TriABC from *Pseudomonas aeruginosa*. *Mol. Microbiol.* **98**, 343–356 (2015).

458. Cacciotto, P. *et al.* Molecular rationale for the impairment of the MexAB-OprM efflux pump by a single mutation in MexA. *Comput. Struct. Biotechnol. J.* **20**, 252–260 (2022).

459. Ma, M., Lustig, M., Mengin-lecreulx, D., Phan, G. & Broutin, I. MexAB-OprM Efflux Pump Interaction with the Peptidoglycan of *Escherichia coli* and *Pseudomonas aeruginosa*. *Int. J. Mol. Sci.* **22**, 1–12 (2021).

460. Xu, Y., Lee, M., Moeller, A. & Song, S. Funnel-like Hexameric Assembly of the Periplasmic Adapter Protein in the Tripartite Multidrug Efflux Pump in. *J. Biol. Chem.* **286**, 17910–17920 (2011).

461. Sørensen, L. & Salbo, R. Optimised Workflow for Selecting Peptides for HDX-MS Data Analyses. *J. Am. Soc. Mass Spectrom.* **29**, 2278–2281 (2018).

462. Calvaresi, V. *et al.* Conformational dynamics of free and membrane-bound human Hsp70 in model cytosolic and endo-lysosomal environments. *Commun. Biol.* **4**, 1–11 (2021).

463. Houde, D., Berkowitz, S. A. & Engen, J. R. The Utility of Hydrogen/Deuterium Exchange Mass Spectrometry in Biopharmaceutical Comparability Studies. *J. Pharm. Sci.* **100**, 2071–2086 (2012).

464. Hinchliffe, P. *et al.* Structure of the periplasmic adaptor protein from a major facilitator superfamily (MFS) multidrug efflux pump. *FEBS Lett.* **588**, 3147–3153 (2014).

465. Pollock, N. L. *et al.* SMA-PAGE: A new method to examine complexes of membrane proteins using SMALP nano-encapsulation and native gel electrophoresis. *BBA- Biomembr.* **1861**, 1437–1445 (2019).

466. Wittig, I., Braun, H. P. & Schägger, H. Blue native PAGE. *Nat. Protoc.* **1**, 418–428 (2006).

467. Crichton, P. G., Harding, M., Ruprecht, J. J., Lee, Y. & Kunji, E. R. S. Lipid, detergent, and coomassie blue G-250 affect the migration of small membrane proteins in blue native gels: Mitochondrial carriers migrate as monomers not dimers. *J. Biol. Chem.* **288**, 22163–22173 (2013).

468. Rath, A., Glibowicka, M., Nadeau, V. G., Chen, G. & Deber, C. M. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1760–1765 (2009).

469. Piatkov, K. I., Vu, T. T. M., Hwang, C. S. & Varshavsky, A. Formyl-methionine as a degradation

signal at the N-termini of bacterial proteins. *Microb. Cell.* **2**, 376–393 (2015).

470. Pollegioni, L. *et al.* Contribution of the dimeric state to the thermal stability of the flavoprotein D-amino acid oxidase. *Protein Sci.* **12**, 1018–1029 (2003).

471. Marianayagam, N. J., Sunde, M. & Matthews, J. M. The power of two: protein dimerization in biology. *Trends Biochem. Sci.* **29**, 618–625 (2004).

472. Efimov, A. V. Complementary packing of α-helices in proteins. *FEBS Lett.* **463**, 3–6 (1999).

473. Johnson, R. M., Heslop, C. L. & Deber, C. M. Hydrophobic Helical Hairpins: Design and Packing Interactions in Membrane. *Biochemistry.* **43**, 14361–14369 (2004).

474. Postis, V. *et al.* The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *Biochim. Biophys. Acta - Biomembr.* **1848**, 496–501 (2015).

475. Louche, A., Salcedo, S. P. & Bigot, S. Protein--Protein Interactions: Pull-Down Assays. in *Bacterial Protein Secretion Systems: Methods and Protocols* (eds. Journet, L. & Cascales, E.) 247–255 (Springer New York, 2017).

476. Xu, Y. *et al.* Assembly and channel opening of outer membrane protein in tripartite drug efflux pumps of gram-negative bacteria. *J. Biol. Chem.* **287**, 11740–11750 (2012).

477. Levy, S. B. & Bonnie, M. Antibacterial resistance worldwide: Causes, challenges and responses. *Nat. Med.* **10**, S122–S129 (2004).

478. Opperman, T. J. & Nguyen, S. T. Recent advances toward a molecular mechanism of efflux pump inhibition. *Front. Microbiol.* **6**, 1–16 (2015).

479. Compagne, N. *et al.* Update on the Discovery of Efflux Pump Inhibitors against Critical Priority Gram-Negative Bacteria. *Antibiotics.* **12**, (2023).

480. Lamut, A., Peterlin Mašič, L., Kikelj, D. & Tomašič, T. Efflux pump inhibitors of clinically relevant multidrug resistant bacteria. *Med. Res. Rev.* **39**, 2460–2504 (2019).

481. Lomovskaya, O. & Bostian, K. A. Practical applications and feasibility of efflux pump inhibitors in the clinic - A vision for applied use. *Biochem. Pharmacol.* **71**, 910–918 (2006).

482. Lomovskaya, O. *et al.* Identification and characterization of inhibitors of multidrug resistance efflux pumps in *Pseudomonas aeruginosa*: Novel agents for combination therapy. *Antimicrob. Agents Chemother.* **45**, 105–116 (2001).

483. Alenazy, R. Drug Efflux Pump Inhibitors: A Promising Approach to Counter Multidrug Resistance in Gram-Negative Pathogens by Targeting AcrB Protein from AcrAB-TolC Multidrug Efflux Pump from *Escherichia coli*. *Biology.* **11**, 1328 (2022).

484. Opperman, T. J. *et al.* Characterization of a novel pyranopyridine inhibitor of the AcrAB efflux pump of *Escherichia coli*. *Antimicrob. Agents Chemother.* **58**, 722–733 (2014).

485. Sjuts, H. *et al.* Molecular basis for inhibition of AcrB multidrug efflux pump by novel and powerful pyranopyridine derivatives. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3509–3514 (2016).

486. Bhattacharyya, S., Walker, D. M. & Harshey, R. M. Dead cells release a 'necrosignal' that activates antibiotic survival pathways in bacterial swarms. *Nat. Commun.* **11**, 1–12 (2020).

487. Bhattacharyya, S., Bhattacharyya, M., Pfannenstiel, D. M., Nandi, A. K. & Hwang, Y. Efflux-linked accelerated evolution of antibiotic resistance at a population edge. *Mol. Cell.* **82**, 4368–4385 (2022).

488. Yesair, D. W., Kohner, F. A., Rogers, W. I., Baronowsky, P. E. & Kensler, C. J. Relationship of phthalanilide-lipid complexes to uptake and retention of 2-chloro-4',4"-di(2-imidazolin-2-yl)terephthalanilide (NSC 60339) by sensitive and resistant P388 leukemia cells. *Cancer Res.* **26**, 202–207 (1966).

489. Darzynkiewicz, Z. M. *et al.* Identification of Binding Sites for Efflux Pump Inhibitors of the AcrAB-TolC Component AcrA. *Biophys. J.* **116**, 648–658 (2019).

490. Anderson, K. W. *et al.* HDX-MS and MD Simulations Provide Evidence for Stabilization of the IgG1—FcγRIa (CD64a) Immune Complex Through Intermolecular Glycoprotein Bonds. *J. Mol. Biol.* **434**, 167391 (2022).

491. Best, R. B. & Vendruscolo, M. Structural interpretation of hydrogen exchange protection factors in proteins: characterization of the native state fluctuations of CI2. *Structure.* **14**, 97–106 (2006).

492. Haynes, K. M. *et al.* Identification and Structure-Activity Relationships of Novel Compounds that Potentiate the Activities of Antibiotics in *Escherichia coli*. *J. Med. Chem.* **60**, 6205–6219 (2017).

493. Jaquillard, L., Saab, F., Schoentgen, F. & Cadene, M. Improved accuracy of low affinity protein-ligand equilibrium dissociation constants directly determined by electrospray ionisation mass spectrometry. *J. Am. Soc. Mass Spectrom.* **23**, 908–922 (2012).

494. Krishnamoorthy, G. *et al.* Breaking the permeability barrier of *Escherichia coli* by controlled Hyperporination of the outer membrane. *Antimicrob. Agents Chemother.* **60**, 7372–7381 (2016).

495. Westfall, D. A. *et al.* Bifurcation kinetics of drug uptake by Gram-negative bacteria. *PLoS One.* **12**, 1–18 (2017).

496. Yang, L., Broderick, D., Jiang, Y., Hsu, V. & Maier, C. S. Conformational dynamics of human FXR-LBD ligand interactions studied by hydrogen/deuterium exchange mass spectrometry: Insights into the antagonism of the hypolipidemic agent Z-guggulsterone. *Biochim. Biophys. Acta - Proteins Proteomics.* **1844**, 1684–1693 (2014).

497. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 60–76 (2005).

498. Phillips, J. C. *et al.* Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **153**, 1–33 (2020).

499. Huang, J. *et al.* CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods.* **176**, 139–148 (2017).

500. Vanommeslaeghe, K. *et al.* CHARMM General Force Field (CGenFF): A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).

501. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

502. Li, W. *et al.* Direct Inhibition of MmpL3 by Novel Antitubercular Compounds. *ACS Infect. Dis.* **5**, 1001–1012 (2019).

503. Tikhonova, E. B. & Zgurskaya, H. I. AcrA, AcrB, and TolC of *Escherichia coli* form a stable intermembrane multidrug efflux complex. *J. Biol. Chem.* **279**, 32116–32124 (2004).

504. Coldham, N. G., Webber, M., Woodward, M. J. & Piddock, L. J. V. A 96-well plate fluorescence assay for assessment of cellular permeability and active efflux in *Salmonella enterica* serovar Typhimurium and *Escherichia coli*. *J. Antimicrob. Chemother.* **65**, 1655–1663 (2010).

505. Gulamhussein, A. A. *et al.* Examining the stability of membrane proteins within SMALPs. *Eur. Polym. J.* **112**, 120–125 (2018).

506. Cockley, D. Detergents and their Uses in Membrane Protein Science. *Anatrace Man.* 1–17 (2007).

507. Van Den Brink-Van Der Laan, E., Chupin, V., Killian, J. A. & De Kruijff, B. Stability of KcsA Tetramer Depends on Membrane Lateral Pressure. *Biochemistry.* **43**, 4240–4250 (2004).

508. Tribet, C., Audebert, R. & Popot, J. L. Amphipols: Polymers that keep membrane proteins soluble in aqueous solutions. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15047–15050 (1996).

509. Shen, H., Lithgow, T. & Martin, L. L. Reconstitution of Membrane Proteins into Model Membranes: Seeking Better Ways to Retain Protein Activities. *Int. J. Mol. Sci.* **14**, 1589–1607 (2013).

510. Morgan, C. R. *et al.* Conformational transitions in the membrane scaffold protein of phospholipid bilayer nanodiscs. *Mol. Cell. Proteomics.* **10**, 1–11 (2011).

511. Borch, J. & Hamann, T. The nanodisc: a novel tool for membrane protein studies. *Biol. Chem.* **390**, 805–814 (2009).

512. Bayburt, T. H., Grinkova, Y. V & Sligar, S. G. Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. *Nano Lett.* **2**, 1–4 (2002).

513. Hagn, F., Etzkorn, M., Raschle, T. & Wagner, G. Optimised Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins. *J. Am. Chem. Soc.* **135**, 1919–1925 (2013).

514. Grinkova, Y. V, Denisov, I. G. & Sligar, S. G. Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers. *Protein Eng. Des. Sel.* **23**, 843–848 (2010).

515. Civjan, N. R., Bayburt, T. H., Schuler, M. A. & Sligar, S. G. Direct solubilization of heterologously expressed membrane proteins by incorporation into nanoscale lipid bilayers. *Biotechniques.* **35**, 556–563 (2003).

516. Pollock, N. L., Lee, S. C., Patel, J. H., Gulamhussein, A. A. & Rothnie, A. J. Structure and function of membrane proteins encapsulated in a polymer-bound lipid bilayer. *Biochim. Biophys. Acta - Biomembr.* **1860**, 809–817 (2018).

517. Arenas, R. C., Klingler, J., Vargas, C. & Keller, S. Nanoscale Influence of lipid bilayer properties on nanodisc formation mediated by styrene/maleic acid. *Nanoscale.* **8**, 15016–15026 (2016).

518. Craig, A. F. *et al.* Tuning the size of styrene-maleic acid copolymer-lipid nanoparticles (SMALPs) using RAFT polymerization for biophysical studies. *Biochim. Biophys. Acta - Biomembr.* **1858**, 2931–2939 (2016).

519. Morrison, K. A. *et al.* Membrane protein extraction and purification using styrene-maleic acid (SMA) copolymer: Effect of variations in polymer structure. *Biochem. J.* **473**, 4349–4360 (2016).

520. Scheidelaar, S. *et al.* Effect of Polymer Composition and pH on Membrane Solubilization by Styrene-Maleic Acid Copolymers. *Biophys. J.* **111**, 1974–1986 (2016).

521. Oluwole, A. O., Meister, A., Babalola, J. O., Vargas, C. & Keller, S. Solubilization of Membrane Proteins into Functional Lipid-Bilayer Nanodiscs Using a Diisobutylene/Maleic Acid Copolymer. *Angew. Chemie - Int. Ed.* **56**, 1919–1924 (2017).

522. Orekhov, P. S. *et al.* Mechanisms of Formation, Structure, and Dynamics of Lipoprotein Discs Stabilized by Amphiphilic Copolymers: A Comprehensive Review. *Nanomaterials.* **12**, 1–27 (2022).

523. Voskoboynikova, N. *et al.* Evaluation of DIBMA nanoparticles of variable size and anionic lipid content as tools for the structural and functional study of membrane proteins. *Biochim. Biophys. Acta - Biomembr.* **1863**, 183588 (2021).

524. Voskoboynikova, N. *et al.* Lipid Dynamics in Diisobutylene-Maleic Acid (DIBMA) Lipid Particles in Presence of Sensory Rhodopsin II. *Int. J. Mol. Sci.* **22**, 1–18 (2021).

525. Danielczak, B. & Keller, S. Collisional lipid exchange among DIBMA-encapsulated nanodiscs (DIBMALPs). *Eur. Polym. J.* **109**, 206–213 (2018).

526. Gulamhussein, A. A., Uddin, R., Tighe, B. J., Poyner, D. R. & Rothnie, A. J. A comparison of SMA (styrene maleic acid) and DIBMA (di-isobutylene maleic acid) for membrane protein purification. *Biochim. Biophys. Acta - Biomembr.* **1862**, 183281 (2020).

527. Danielczak, B., Meister, A. & Keller, S. Influence of Mg^{2+} and Ca^{2+} on nanodisc formation by diisobutylene/maleic acid (DIBMA) copolymer. *Chem. Phys. Lipids.* **221**, 30–38 (2019).

528. Grime, R. L. *et al.* Differences in SMA-like polymer architecture dictate the conformational changes exhibited by the membrane protein rhodopsin encapsulated in lipid nano-particles. *Nanoscale.* **13**, 13519–13528 (2021).

529. Hall, S. C. L. *et al.* An acid-compatible co-polymer for the solubilization of membranes and proteins into lipid bilayer-containing nanoparticles. *Nanoscale.* **10**, 10609–10619 (2018).

530. Hall, S. C. L. *et al.* Adsorption of a styrene maleic acid (SMA) copolymer-stabilized phospholipid nanodisc on a solid-supported planar lipid bilayer. *J. Colloid Interface Sci.* **574**, 272–284 (2020).

531. Enguéné, V. Y. N., Verchère, A., Phan, G., Broutin, I. & Picard, M. Catch me if you can: A biotinylated proteoliposome affinity assay for the investigation of assembly of the MexA-MexB-OprM efflux pump from *pseudomonas aeruginosa*. *Front. Microbiol.* **6**, 1–8 (2015).

532. Pellowe, G. A. *et al.* Capturing Membrane Protein Ribosome Nascent Chain Complexes in a Native-like Environment for Co-translational Studies. *Biochemistry.* **59**, 2764–2775 (2020).

533. Ejima, D., Yumioka, R., Arakawa, T. & Tsumoto, K. Arginine as an effective additive in gel permeation chromatography. *J. Chromatogr. A.* **1094**, 49–55 (2005).

534. Vargiu, A. V, Ruggerone, P., Opperman, T. J. & Nguyen, S. T. Molecular Mechanism of MBX2319 Inhibition of *Escherichia coli* AcrB Multidrug Efflux Pump and Comparison with Other Inhibitors. *Antimicrob. Agents Chemother.* **58**, 6224–6234 (2014).

535. Lau, A. M., Claesen, J., Hansen, K. & Politis, A. Deuteros 2.0: peptide-level significance testing of data from hydrogen deuterium exchange mass spectrometry. *Bioinformatics.* **37**, 270–272 (2021).

536. Weis, D. D. Comment on 'Houde, D.; Berkowitz, S. A.; Engen, J. R., The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies. J. Pharm. Sci. 2011, 100, 2071-2086.' *J. Pharm. Sci.* **108**, 807–810 (2019).

537. Yu, E. W., Aires, J. R., Mcdermott, G. & Nikaido, H. A Periplasmic Drug-Binding Site of the AcrB Multidrug Efflux Pump: a Crystallographic and Site-Directed Mutagenesis Study. *Am. Soc. Microbiol.* **187**, 6804–6815 (2005).

538. Kinana, A. D., Vargiu, A. V, May, T. & Nikaido, H. Aminoacyl β-naphthylamides as substrates and modulators of AcrB multidrug efflux pump. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 1405–1410 (2016).

539. Poulsen, B. E. & Deber, C. M. Drug Efflux by a Small Multidrug Resistance Protein Is Inhibited by a Transmembrane Peptide. *Antimicrob. Agents Chemother.* **56**, 2–7 (2012).

540. Yang, Z. *et al.* Membrane protein stability can be compromised by detergent interactions with the extramembranous soluble domains. *Protein Sci.* **23**, 769–789 (2014).

541. Ayub, H. *et al.* CD81 extracted in SMALP nanodiscs comprises two distinct protein populations within a lipid environment enriched with negatively charged headgroups. *Biochim. Biophys. Acta - Biomembr.* **1862**, 1–15 (2020).

542. Lu, W., Zhong, M. & Wei, Y. Folding of AcrB Subunit Precedes Trimerization. *J. Mol. Biol.* **411**, 264–274 (2011).

543. Thein, M., Sauer, G., Paramasivam, N., Grin, I. & Linke, D. Efficient subfractionation of gram-negative bacteria for proteomics studies. *J. Proteome Res.* **9**, 6135–6147 (2010).

544. Krishnamoorthy, G., Tikhonova, E. B., Dhamdhere, G. & Zgurskaya, H. I. On the role of TolC in multidrug efflux: the function and assembly of AcrAB-TolC tolerate significant depletion of intracellular TolC protein. *Mol. Microbiol.* **87**, 982–997 (2013).

545. Noinaj, N., Kuszak, A. J. & Buchanan, S. K. Heat Modifiability of Outer Membrane Proteins from Gram-Negative Bacteria. *Methods Mol. Biol.* **1329**, 51–56 (2015).

546. Hesketh, S. J. *et al.* Styrene maleic-acid lipid particles (SMALPs) into detergent or amphipols: An exchange protocol for membrane protein characterisation. *Biochim. Biophys. Acta - Biomembr.* **1862**, 1–8 (2020).

547. Kamilar, E., Bariwal, J., Zheng, W. & Liang, H. SMALPs Are Not Simply Nanodiscs: The Polymer-to-Lipid Ratios of Fractionated SMALPs Underline Their Heterogeneous Nature. *Biomolecules.* **24**, 1819–1838 (2023).

548. Xu, T.-H. *et al.* Streptavidin Bead Pulldown Assay to Determine Protein Homooligomerization. *Bio Protoc.* **7**, 1–8 (2017).

549. Guesdon, J.-L., Ternynck, T. & Avrameas, S. The Use of Avidin-Biotin Interaction in Immunoenzymatic Techniques. *J. Histochem. Cytochem.* **27**, 1131–1139 (1979).

550. Jain, A. & Cheng, K. The Principles and Applications of Avidin-Based Nanoparticles in Drug Delivery and Diagnosis. *J. Control. Release.* **254**, 27–40 (2017).

551. Alon, R., Bayer, E. A. & Wilchek, M. Streptavidin contains an RYD sequence which mimics the RGD receptor domain of fibronectin. *Biochem. Biophys. Res. Commun.* **170**, 1236–1241 (1990).

552. Adibnia, V., Mirbagheri, M., Salimi, S., Crescenzo, G. De & Banquy, X. Nonspecific interactions in biomedical applications. *Curr. Opin. Colloid Interface Sci.* **47**, 70–83 (2020).

553. Ahirwar, R., Bariar, S., Balakrishnan, A. & Nahar, P. BSA blocking in enzyme-linked immunosorbent assays is a non-mandatory step: a perspective study on mechanism of BSA blocking in common ELISA. *RSC Adv.* **5**, 100077–100083 (2015).

554. Sun, Y. & Zhu, X. Characterization of Bovine Serum Albumin Blocking Efficiency on Epoxy-Functionalized Substrates for Microarray Applications. *J. Lab. Autom.* **21**, 625–631 (2016).

555. Baksh, M. M., Kussrow, A. K., Mileni, M., Finn, M. G. & Bornhop, D. J. Label-free quantification of membrane-ligand interactions using backscattering interferometry. *Nat. Biotechnol.* **29**, 357–362 (2011).

556. Donnarumma, D. *et al.* Native State Organization of Outer Membrane Porins Unraveled by HDx-MS. *J. Proteome Res.* **17**, 1794–1800 (2018).

557. Lin, X., Zmyslowski, A. M., Gagnon, I. A., Nakamoto, R. K. & Sosnick, T. R. Development of *in vivo* HDX-MS with applications to a TonB-dependent transporter and other proteins. *Protein Sci.* **31**, 1–9 (2022).

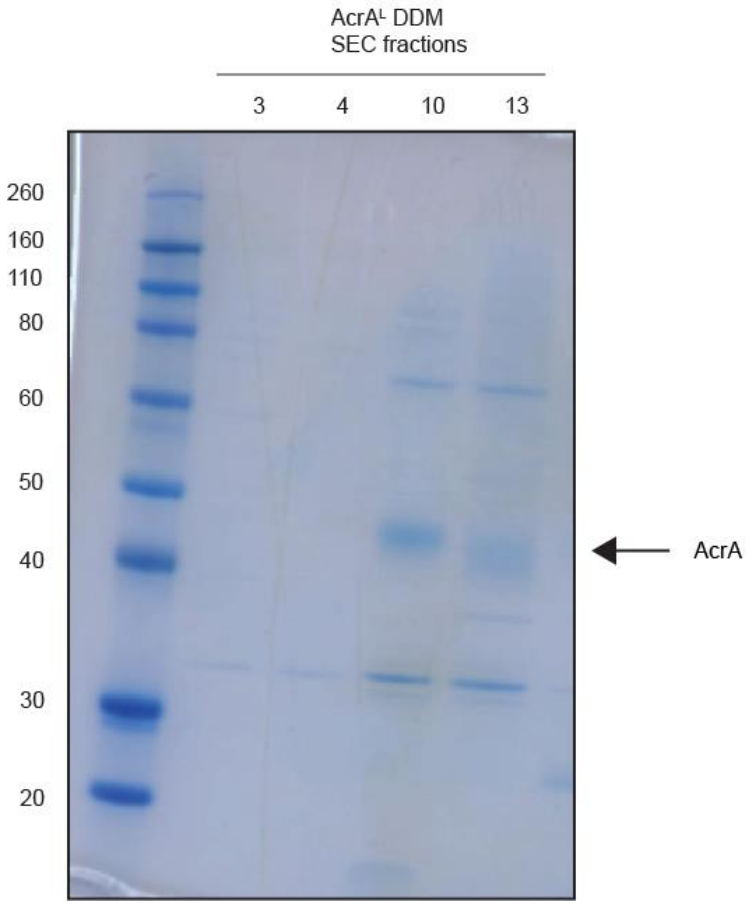
558. Denisov, I. G. *et al.* Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size. *J. Am. Chem. Soc.* **126**, 3477–3487 (2004).

559. Nagy, T. A., Crooks, A. L., Quintana, J. L. J. & Detweiler, C. S. Clofazimine Reduces the Survival of *Salmonella enterica* in Macrophages and Mice. *ACS Infect. Dis.* **6**, 1238–1249 (2020).

560. Parmar, M. *et al.* Using a SMALP platform to determine a sub-nm single particle cryo-EM membrane protein structure. *Biochim. Biophys. Acta - Biomembr.* **1860**, 378–383 (2018).

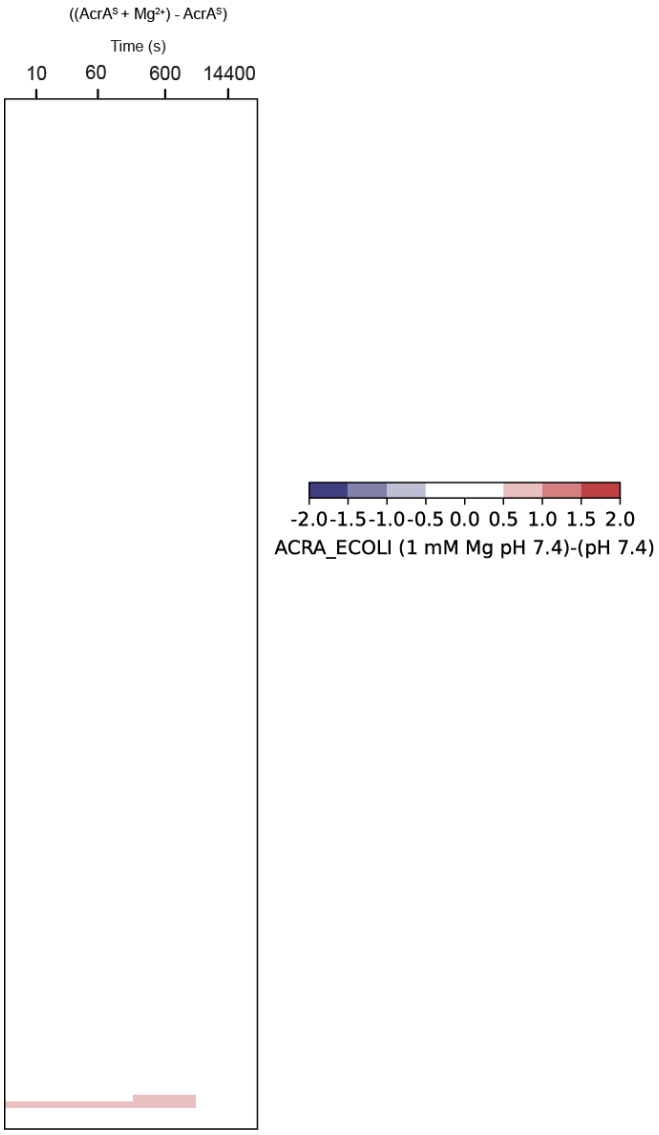
Appendix

Appendix 1. AcrA^{L} pUC18 purification



Appendix 1. SDS-PAGE of AcrA^{L} pUC18 construct. Size exclusion chromatography fractions shown on an SDS-PAGE. Fractions 3-4 are from the void peak and show no protein. Fractions 10-13 present a band at ~40 kDa representing AcrA. This presents as a single band, showing homogenous lipidation compared to the AcrA^{L} pET28a construct.

Appendix 2. ΔHDX of ((AcrA^S + Mg^{2+}) - AcrA^S) pH 7.4 with extended time point



Appendix 2. Chiclet plot of AcrA and Mg^{2+} at pH 7.4 with an extended time point. The differential HDX (ΔHDX) plots for ((AcrA^S + Mg^{2+}) - AcrA^S), at pH 7.4 for all time points (10s, 60s, 600s, 14400s) shown on a chiclet plot. Plotted using HDeXplosion.[447] White signifies areas with no significant change in HDX, whereas red signifies increased HDX. Significance was defined to be ≥ 0.50 Da change with a *P*-value ≤ 0.01 in a Welch's *t*-test (n=4)**.**

Appendix 3. AcrA^{L} cloning plan

AcrA DNA sequence:

>ENA|AAA67134|AAA67134.1 Escherichia coli 42 kDa protein

ATGAACAAAAACAGAGGGTTTACGCCTCTGGCGGTCGTTCTGATGCTCTCAGGCAGCTTAGCCCTAACAGGAT
GTGACGACAAACAGGCCCAACAAGGTGGCCAGCAGATGCCCGCCGTTGGCGTAGTAACAGTCAAAACTGAAC
CTCTGCAGATCACAACCGAGCTTCCGGGTCGCACCAGTGCCTACCGGATCGCAGAAGTTCGTCCTCAAGTTAG
CGGGATTATCCTGAAGCGTAATTTCAAAGAAGGTAGCGACATCGAAGCAGGTGTCTCTCTCTATCAGATTGAT
CCTGCGACCTATCAGGCGACATACGACAGTGCGAAAGGTGATCTGGCGAAAGCCCAGGCTGCAGCCAATATC
GCGCAATTGACGGTGAATCGTTATCAGAAACTGCTCGGTACTCAGTACATCAGTAAGCAAGAGTACGATCAG
GCTCTGGCTGATGCGCAACAGGCGAATGCTGCGGTAACTGCGGCGAAAGCTGCCGTTGAAACTGCGCGGATC
AATCTGGCTTACACCAAAGTCACCTCTCCGATTAGCGGTCGCATTGGTAAGTCGAACGTGACGGAAGGCGCAT
TGGTACAGAACGGTCAGGCGACTGCGCTGGCAACCGTGCAGCAACTTGATCCGATCTACGTTGATGTGACCC
AGTCCAGCAACGACTTCCTGCGCCTGAAACAGGAACTGGCGAATGGCACGCTGAAACAAGAGAACGGCAAA
GCCAAAGTGTCACTGATCACCAGTGACGGCATTAAGTTCCCGCAGGACGGTACGCTGGAATTCTCTGACGTTA
CCGTTGATCAGACCACTGGGTCTATCACCCTACGCGCTATCTTCCCGAACCCGGATCACACTCTGCTGCCGGGT
ATGTTCGTGCGCGCACGTCTGGAAGAAGGGCTTAATCCAAACGCTATTTTAGTCCCGCAACAGGGCGTAACCC
GTACGCCGCGTGGCGATGCCACCGTACTGGTAGTTGGCGCGGATGACAAAGTGGAAACCCGTCCGATCGTTG
CAAGCCAGGCTATTGGCGATAAGTGGCTGGTGACAGAAGGTCTGAAAGCAGGCGATCGCGTAGTAATAAGT
GGGCTGCAGAAAGTGCGTCCTGGTGTCCAGGTAAAAGCACAAGAAGTTACCGCTGATAATAACCAGCAAGCC
GCAAGCGGTGCTCAGCCTGAACAGTCCAAGTCTTAA

Protein sequence:

MNKNRGFTPLAVVLMLSGSLALTGCDDKQAQQGGQQMPAVGVVTVKTEPLQITTELPGRTSAYRIAEVRPQVSG
IILKRNFKEGSDIEAGVSLYQIDPATYQATYDSAKGDLAKAQAAANIAQLTVNRYQKLLGTQYISKQEYDQALADAQ
QANAAVTAAKAAVETARINLAYTKVTSPISGRIGKSNVTEGALVQNGQATALATVQQLDPIYVDVTQSSNDFLRLK
QELANGTLKQENGKAKVSLITSDGIKFPQDGTLEFSDVTVDQTTGSITLRAIFPNPDHTLLPGMFVRARLEEGLNPN
AILVPQQGVTRTPRGDATVLVVGADDKVETRPIVASQAIGDKWLVTEGLKAGDRVVISGLQKVRPGVQVKAQEVT
ADNNQQAASGAQPEQSKS-LEHHHHHH

Primer design:

F1 Primer: 5’ AGGAGATATACCATGAACAAAAACAGAGGGTTTACGC 3’

Overhang 15bp Primer 22 bp

Length: 40

GC%: 40

Tm (°C): 67.4

Gene Specific Tm (°C): 62.6

R1 Primer: 5’ GGTGGTGGTGCTCGAGAGACTTGGACTGTTCAGGC 3’

Length: 35

GC%: 60

Tm (°C): 76.1

Gene Specific Tm (°C): 59.4

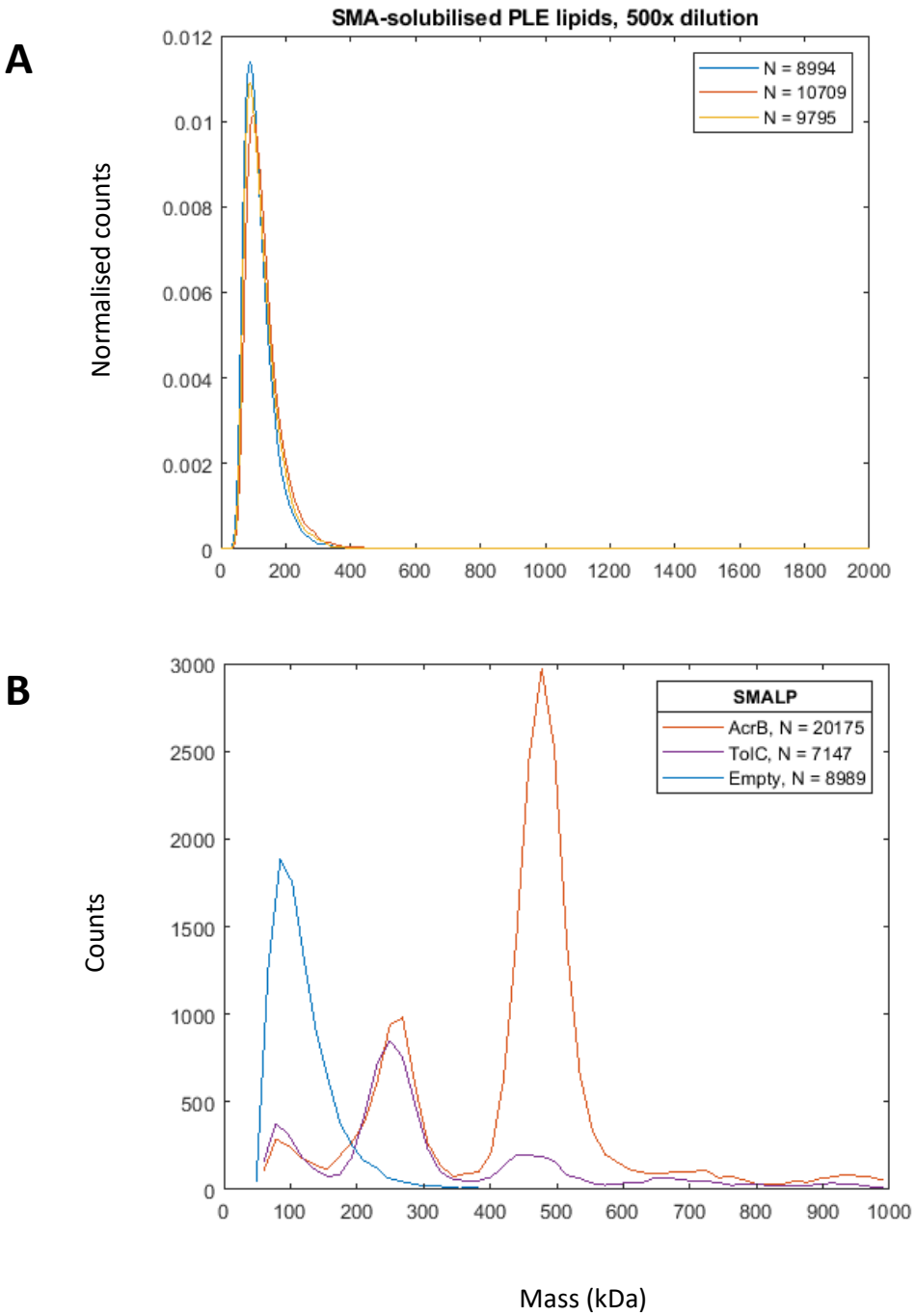
Appendix 4. AcrA^{S} protein sequence

Protein sequence:

MDDKQAQQGGQQMPAVGVVTVKTEPLQITTELPGRTSAYRIAEVRPQVSGIILKRNFKEGSDIEAGVSLYQIDPAT
YQATYDSAKGDLAKAQAAANIAQLTVNRYQKLLGTQYISKQEYDQALADAQQANAAVTAAKAAVETARINLAYTK
VTSPISGRIGKSNVTEGALVQNGQATALATVQQLDPIYVDVTQSSNDFLRLKQELANGTLKQENGKAKVSLITSDGI
KFPQDGTLEFSDVTVDQTTGSITLRAIFPNPDHTLLPGMFVRARLEEGLNPNAILVPQQGVTRTPRGDATVLVVGA
DDKVETRPIVASQAIGDKWLVTEGLKAGDRVVISGLQKVRPGVQVKAQEVTADNNQQAASGAQPEQSKSLEHHH
HHH

All cloning and sequencing performed by the Zgurskaya group (University of Oklahoma).

Appendix 5. Mass photometry SMA control



Appendix 5. Mass photometry of styrene maleic acid solubilised *E. coli* polar lipid extract. A. Mass photometry of styrene maleic acid (SMA) solubilised lipids. The main peak is centred at around 85kDa, which is consistent with the smallest peaks in the TolC and AcrB SMALP data. The peak is also slightly broader, with a right hand side shoulder, suggesting some variability of the discs produced. **B.** Comparison with AcrB and TolC SMALPs. From this we can see that the small molecular weight peaks of TolC and AcrB SMALPs match well with the empty SMALP, although the amplitude is different due to different dilutions and smaller amounts of empty SMALPs present in the AcrB and TolC sample.

Appendix 6. AcrA^{SD} sequence

DNA sequence:

ATGGCAGACGACAAACAGGCCCAACAAGGTGGCCAGCAGATGCCCGCCGTTGGCGTAGTAACAGTCAAAACT
GAACCTCTGCAGATCACAACCGAGCTTCCGGGTCGCACCAGTGCCTACCGGATCGCAGAAGTTCGTCCTCAAG
TTAGCGGGATTATCCTGAAGCGTAATTTCAAAGAAGGTAGCGACATCGAAGCAGGTGTCTCTCTCTATCAGAT
TGATCCTGCGACCTATCAGGCGACATACGACAGTGCGAAAGGTGATCTGGCGAAAGCCCAGGCTGCAGCCAA
TATCGCGCAATTGACGGTGAATCGTTATCAGAAACTGCTCGGTACTCAGTACATCAGTAAGCAAGAGTACGAT
CAGGCTCTGGCTGATGCGCAACAGGCGAATGCTGCGGTAACTGCGGCGAAAGCTGCCGTTGAAACTGCGCG
GATCAATCTGGCTTACACCAAAGTCACCTCTCCGATTAGCGGTCGCATTGGTAAGTCGAACGTGACGGAAGGC
GCATTGGTACAGAACGGTCAGGCGACTGCGCTGGCAACCGTGCAGCAACTTGATCCGATCTACGTTGATGTG
ACCCAGTCCAGCAACGACTTCCTGCGCCTGAAACAGGAACTGGCGAATGGCACGCTGAAACAAGAGAACGGC
AAAGCCAAAGTGTCACTGATCACCAGTGACGGCATTAAGTTCCCGCAGGACGGTACGCTGGAATTCTCTGACG
TTACCGTTGATCAGACCACTGGGTCTATCACCCTACGCGCTATCTTCCCGAACCCGGATCACACTCTGCTGCCG
GGTATGTTCGTGCGCGCACGTCTGGAAGAAGGGCTTAATCCAAACGCTATTTTAGTCCCGCAACAGGGCGTAA
CCCGTACGCCGCGTGGCGATGCCACCGTACTGGTAGTTGGCGCGGATGACAAAGTGGAAACCCGTCCGATCG
TTGCAAGCCAGGCTATTGGCGATAAGTGGCTGGTGACAGAAGGTCTGAAAGCAGGCGATCGCGTAGTAATAA
GTGGGCTGCAGAAAGTGCGTCCTGGTGTCCAGGTAAAAGCACAAGAAGTTACCGCTGATAATAACCAGCAAG
CCGCAAGCGGTGCTCAGCCTGAACAGTCCAAGTCT**ACCAGAAGAATTACC**GACGACAAACAGGCCCAACAAG
GTGGCCAGCAGATGCCCGCCGTTGGCGTAGTAACAGTCAAAACTGAACCTCTGCAGATCACAACCGAGCTTCC
GGGTCGCACCAGTGCCTACCGGATCGCAGAAGTTCGTCCTCAAGTTAGCGGGATTATCCTGAAGCGTAATTTC
AAAGAAGGTAGCGACATCGAAGCAGGTGTCTCTCTCTATCAGATTGATCCTGCGACCTATCAGGCGACATACG
ACAGTGCGAAAGGTGATCTGGCGAAAGCCCAGGCTGCAGCCAATATCGCGCAATTGACGGTGAATCGTTATC
AGAAACTGCTCGGTACTCAGTACATCAGTAAGCAAGAGTACGATCAGGCTCTGGCTGATGCGCAACAGGCGA
ATGCTGCGGTAACTGCGGCGAAAGCTGCCGTTGAAACTGCGCGGATCAATCTGGCTTACACCAAAGTCACCTC
TCCGATTAGCGGTCGCATTGGTAAGTCGAACGTGACGGAAGGCGCATTGGTACAGAACGGTCAGGCGACTGC
GCTGGCAACCGTGCAGCAACTTGATCCGATCTACGTTGATGTGACCCAGTCCAGCAACGACTTCCTGCGCCTG
AAACAGGAACTGGCGAATGGCACGCTGAAACAAGAGAACGGCAAAGCCAAAGTGTCACTGATCACCAGTGA
CGGCATTAAGTTCCCGCAGGACGGTACGCTGGAATTCTCTGACGTTACCGTTGATCAGACCACTGGGTCTATC
ACCCTACGCGCTATCTTCCCGAACCCGGATCACACTCTGCTGCCGGGTATGTTCGTGCGCGCACGTCTGGAAG
AAGGGCTTAATCCAAACGCTATTTTAGTCCCGCAACAGGGCGTAACCCGTACGCCGCGTGGCGATGCCACCGT
ACTGGTAGTTGGCGCGGATGACAAAGTGGAAACCCGTCCGATCGTTGCAAGCCAGGCTATTGGCGATAAGTG
GCTGGTGACAGAAGGTCTGAAAGCAGGCGATCGCGTAGTAATAAGTGGGCTGCAGAAAGTGCGTCCTGGTG
TCCAGGTAAAAGCACAAGAAGTTACCGCTGATAATAACCAGCAAGCCGCAAGCGGTGCTCAGCCTGAACAGT
CCAAGTCT

Yellow = AcrA(1)

Green = Linker

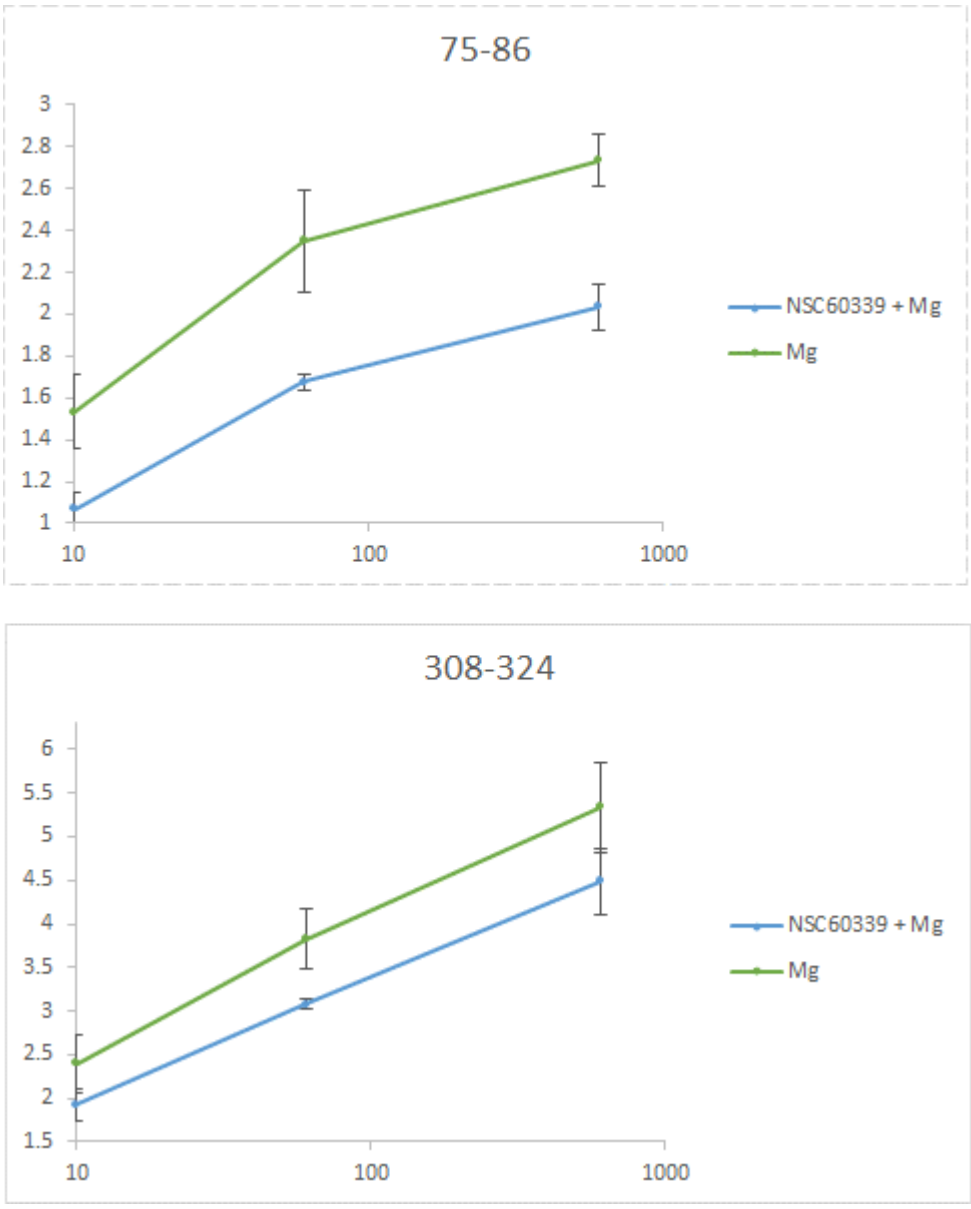
Gray = AcrA(2)

Protein sequence:

MADDKQAQQGGQQMPAVGVVTVKTEPLQITTELPGRTSAYRIAEVRPQVSGIILKRNFKEGSDIEAGVSLYQIDPA
TYQATYDSAKGDLAKAQAAANIAQLTVNRYQKLLGTQYISKQEYDQALADAQQANAAVTAAKAAVETARINLAYT
KVTSPISGRIGKSNVTEGALVQNGQATALATVQQLDPIYVDVTQSSNDFLRLKQELANGTLKQENGKAKVSLITSDG
IKFPQDGTLEFSDVTVDQTTGSITLRAIFPNPDHTLLPGMFVRARLEEGLNPNAILVPQQGVTRTPRGDATVLVVGA
DDKVETRPIVASQAIGDKWLVTEGLKAGDRVVISGLQKVRPGVQVKAQEVTADNNQQAASGAQPEQSKSTRRIT
DDKQAQQGGQQMPAVGVVTVKTEPLQITTELPGRTSAYRIAEVRPQVSGIILKRNFKEGSDIEAGVSLYQIDPATY
QATYDSAKGDLAKAQAAANIAQLTVNRYQKLLGTQYISKQEYDQALADAQQANAAVTAAKAAVETARINLAYTKV
TSPISGRIGKSNVTEGALVQNGQATALATVQQLDPIYVDVTQSSNDFLRLKQELANGTLKQENGKAKVSLITSDGIK
FPQDGTLEFSDVTVDQTTGSITLRAIFPNPDHTLLPGMFVRARLEEGLNPNAILVPQQGVTRTPRGDATVLVVGAD
DKVETRPIVASQAIGDKWLVTEGLKAGDRVVISGLQKVRPGVQVKAQEVTADNNQQAASGAQPEQSKSLEHHHH
HH

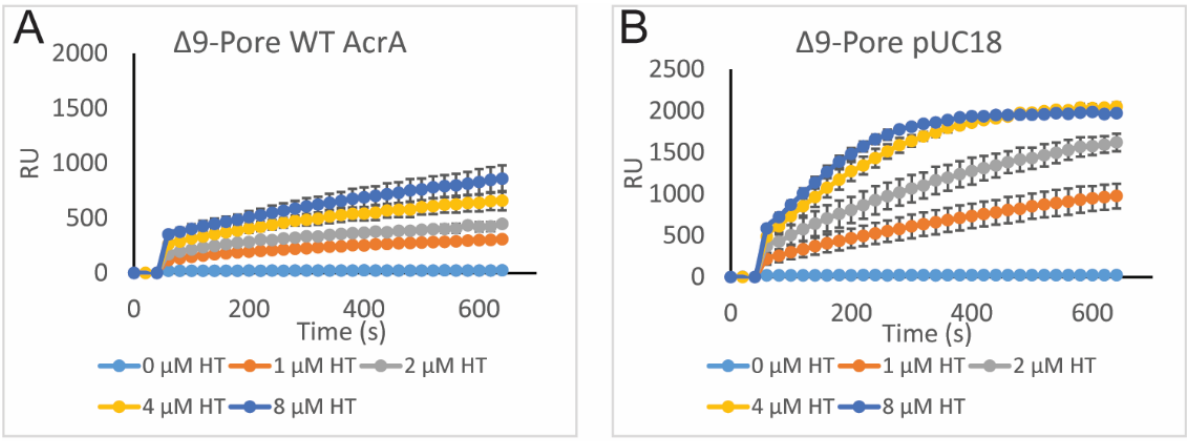
All cloning and sequencing performed by the Zgurskaya group (University of Oklahoma).

Appendix 7. NSC 60339 inhibition of AcrA with Mg^{2+}



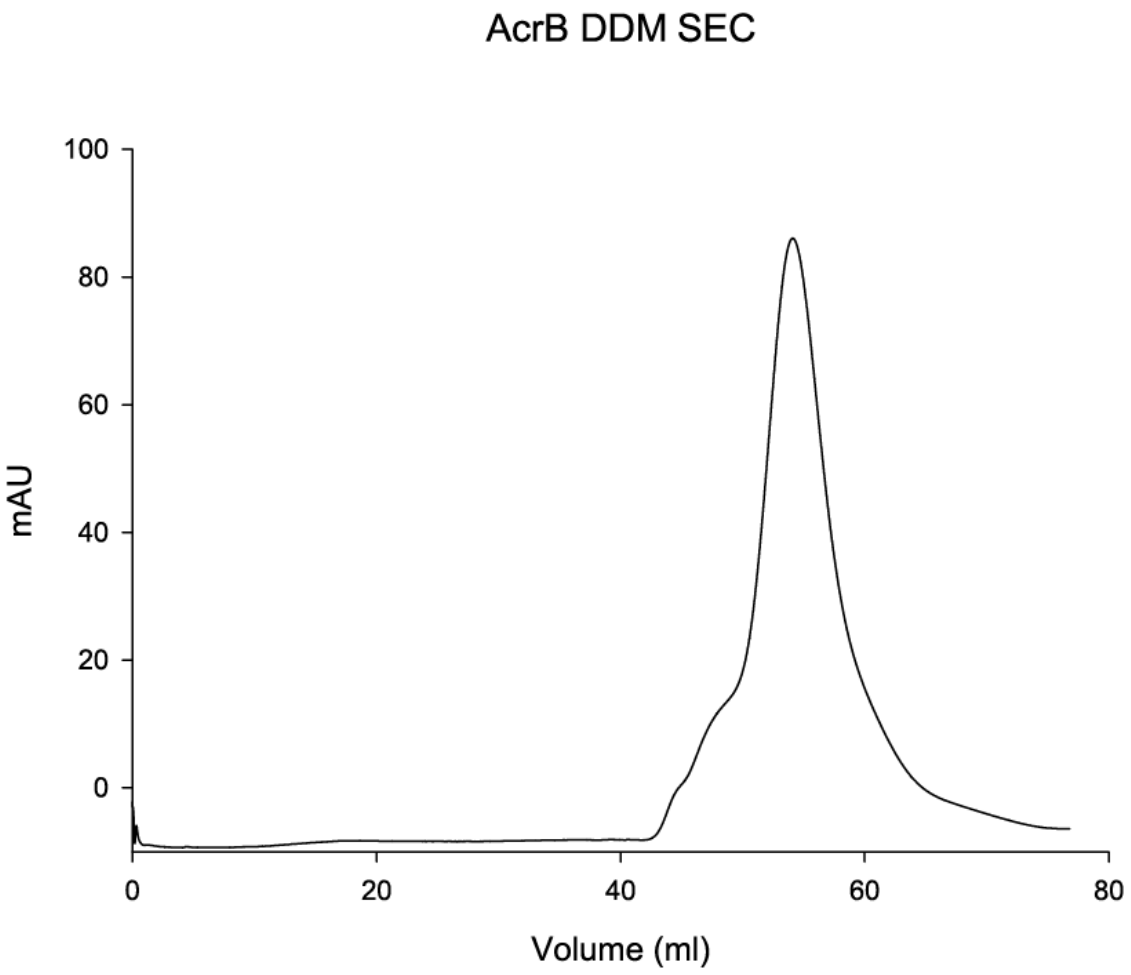
Appendix 7. Peptide uptake plots of NSC 60339 inhibition of AcrA with Mg^{2+}. Uptake plots for two peptides in areas with the largest change in deuterium exchange when AcrA is inhibited by NSC 60339. Uptake plots are the average deuterium uptake and error bars indicate the standard deviation. Green represents AcrA with Mg^{2+} and blue represents AcrA with Mg^{2+} and NSC 60339. The two peptides (^{75}IILKRNFKEGSD^{86}, ^{308}VPQQGVTRTPRGDATVL^{324}) show decreased uptake in the presence of NSC 60339 and Mg^{2+}, suggesting NSC 60339 still restricts AcrA dynamics in the presence of Mg^{2+}.

Appendix 8. Representative time courses of Hoechst accumulation in in Δ9-Pore cells



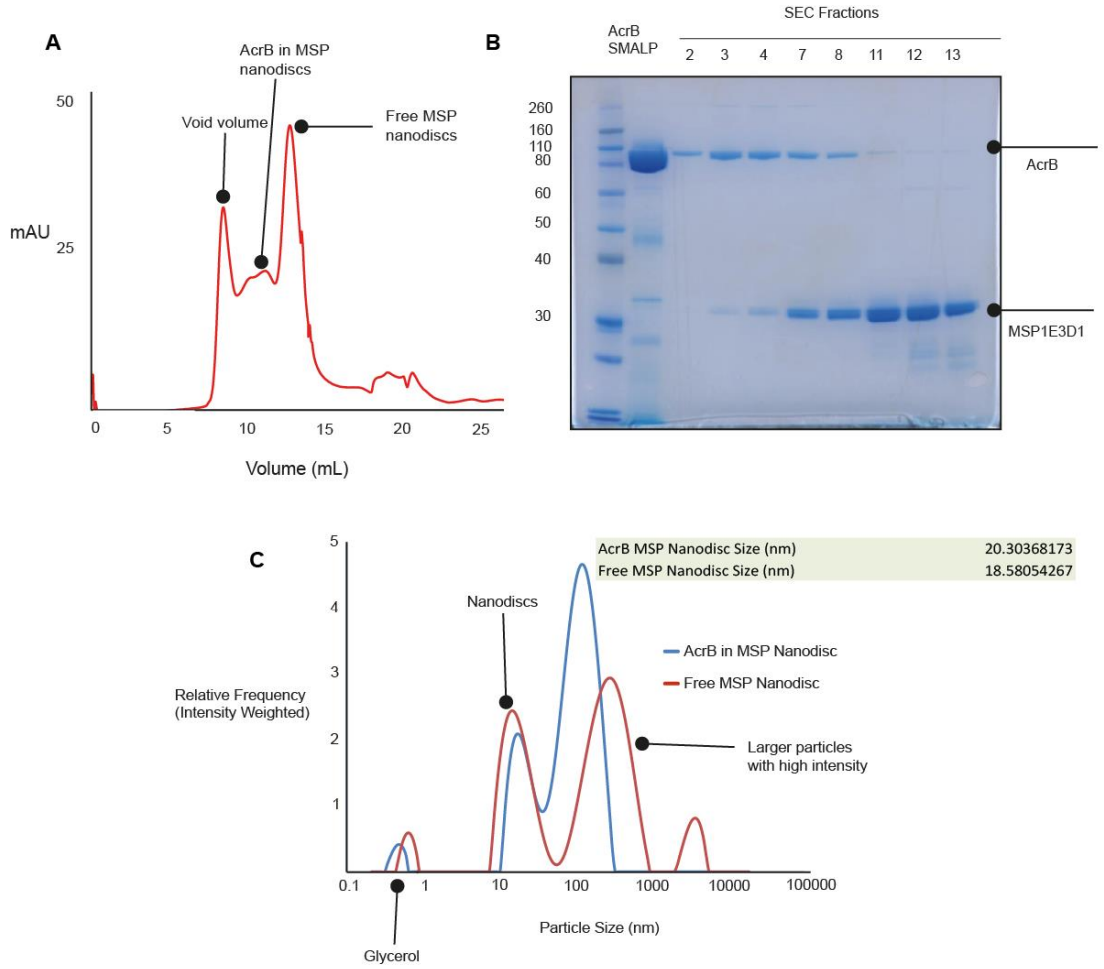
Appendix 8. Representative time courses of intracellular Hoechst accumulation in Δ9-Pore cells. A. Wild type AcrA **B.** pUC18 empty vector.

Appendix 9. AcrB DDM purification



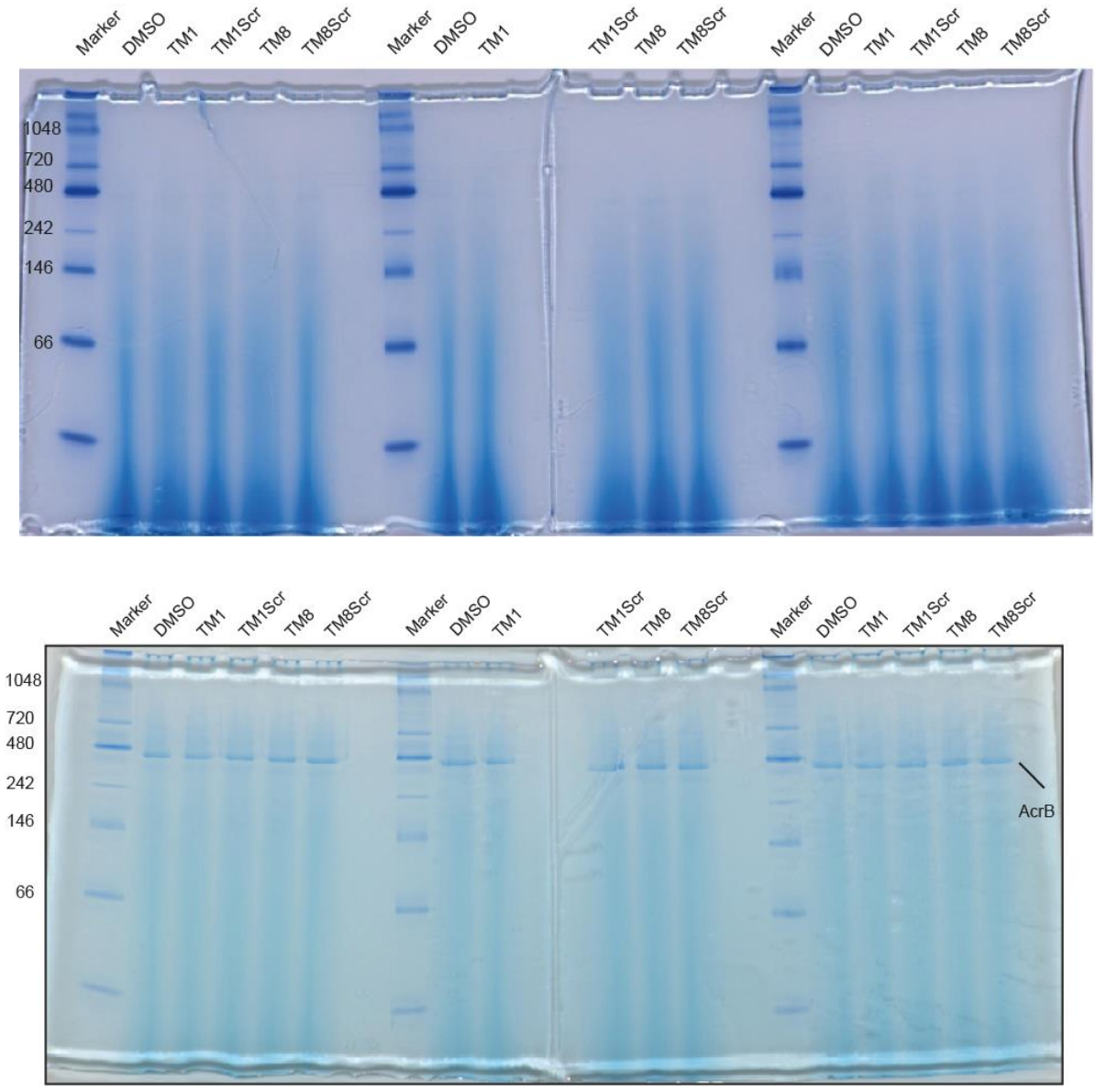
Appendix 9. Characteristic size exclusion chromatogram of AcrB DDM purification. Purified AcrB DDM can be seen on an SDS-PAGE in **Figure 5.19** in the input lane.

Appendix 10. AcrB *E. coli* polar lipid extract MSP nanodiscs characterisation



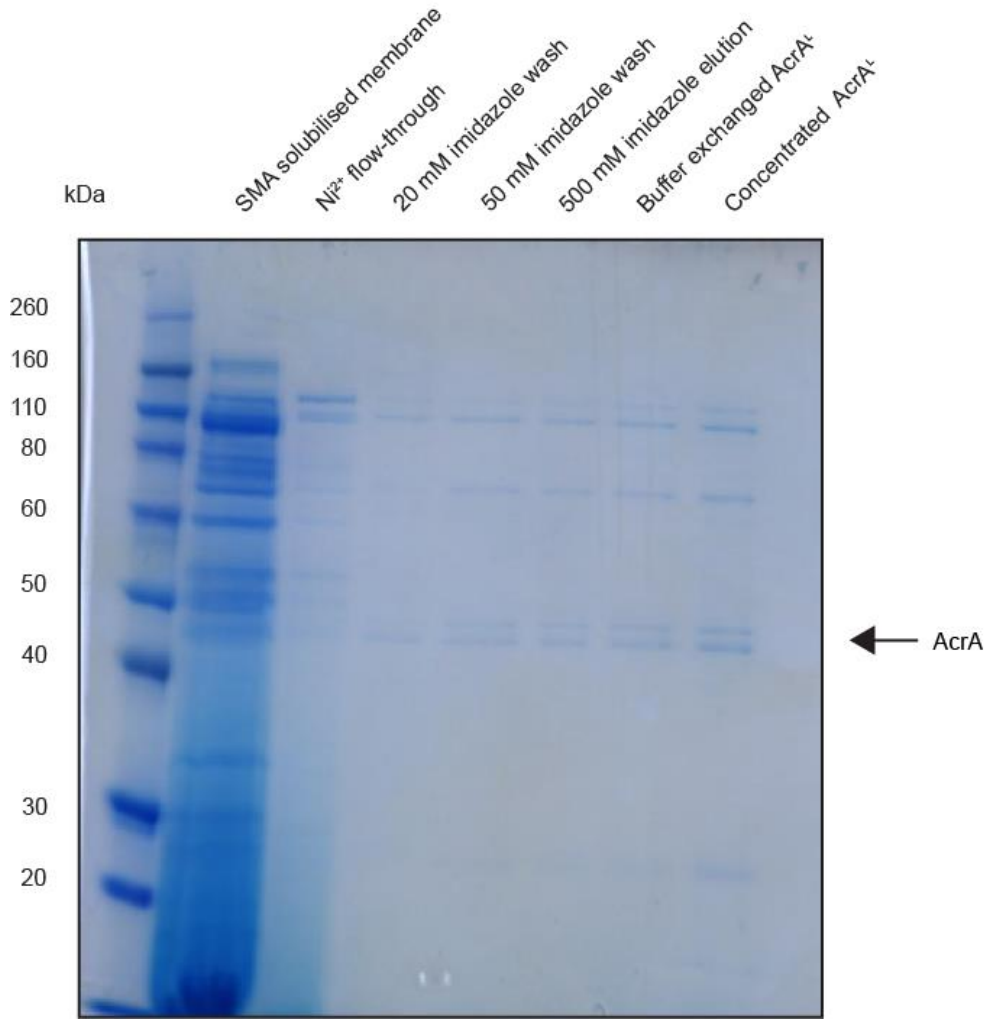
Appendix 10. Characterisation of AcrB membrane MSP nanodiscs in *Escherichia coli* PLE lipids. MSP nanodiscs were made using *E. coli* polar lipid extract (PLE) lipids. **A.** Size exclusion chromatography for AcrB MSP nanodiscs. Void peak can be seen at ~8mL and AcrB MSP nanodisc elutes ~10-12 mL. Free MSP discs elute at ~14 mL. **B.** The fractions from the SEC run on an SDS-PAGE. AcrB and MSP1E3D1 can be observed, with the most intense bands in fractions 7-8. **C.** Dynamic light scattering (DLS) of AcrB MSP nanodiscs. DLS shows a size of 20.3 nm for AcrB MSP nanodiscs and 18.5 nm for empty MSP nanodiscs.

Appendix 11. Triplicates of the SMALP-liposome-SMALP assay with AcrB and the antimicrobial peptides



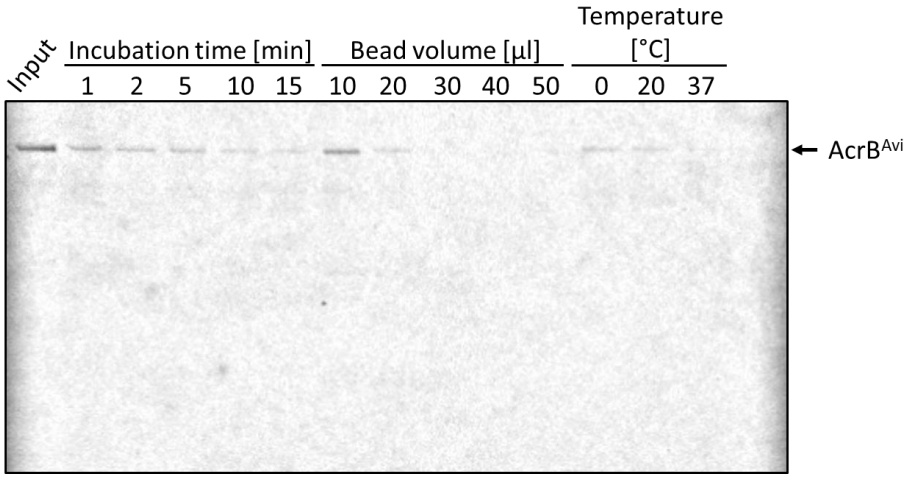
Appendix 11. SMA-liposome-SMA assay of AcrB and antimicrobial peptides in triplicate. Triplicates of the SMA-PAGE and SMA-liposome-SMA assays of AcrB and antimicrobial peptides shown in **Figures 5.9a and 5.10c**.

Appendix 12. AcrA^{L} SMALP purification



Appendix 12. AcrA^{L} purification in SMALPs. SDS-PAGE showing the purification of AcrA^{L} in SMALPS. SDS-PAGE shows AcrA^{L} is successfully solubilised from the cell membrane using SMA and can be purified using Ni^{2+} affinity chromatography.

Appendix 13. Optimisation of the biotin pull-down assay



Appendix 13. Optimising AcrB^{Avi} binding to neutravidin beads. AcrB incubated with different volumes of beads and the flow through analysed. No AcrB can be visualised in the flow through when using 30 or more µL of beads.

Appendix 14. TolC cloning plan

DNA sequence:

>ENA|AAC76071|AAC76071.2 Escherichia coli str. K-12 substr. MG1655 outer membrane channel TolC

ATGAAGAAATTGCTCCCCATTCTTATCGGCCTGAGCCTTTCTGGGTTCAGTTCGTTGAGCCAGGCCGAGAACCT
GATGCAAGTTTATCAGCAAGCACGCCTTAGTAACCCGGAATTGCGTAAGTCTGCCGCCGATCGTGATGCTGCC
TTTGAAAAAATTAATGAAGCGCGCAGTCCATTACTGCCACAGCTAGGTTTAGGTGCAGATTACACCTATAGCA
ACGGCTACCGCGACGCGAACGGCATCAACTCTAACGCGACCAGTGCGTCCTTGCAGTTAACTCAATCCATTTTT
GATATGTCGAAATGGCGTGCGTTAACGCTGCAGGAAAAAGCAGCAGGGATTCAGGACGTCACGTATCAGACC
GATCAGCAAACCTTGATCCTCAACACCGCGACCGCTTATTTCAACGTGTTGAATGCTATTGACGTTCTTTCCTAT
ACACAGGCACAAAAAGAAGCGATCTACCGTCAATTAGATCAAACCACCCAACGTTTTAACGTGGGCCTGGTAG
CGATCACCGACGTGCAGAACGCCCGCGCACAGTACGATACCGTGCTGGCGAACGAAGTGACCGCACGTAATA
ACCTTGATAACGCGGTAGAGCAGCTGCGCCAGATCACCGGTAACTACTATCCGGAACTGGCTGCGCTGAATGT
CGAAAACTTTAAAACCGACAAACCACAGCCGGTTAACGCGCTGCTGAAAGAAGCCGAAAAACGCAACCTGTC
GCTGTTACAGGCACGCTTGAGCCAGGACCTGGCGCGCGAGCAAATTCGCCAGGCGCAGGATGGTCACTTACC
GACTCTGGATTTAACGGCTTCTACCGGGATTTCTGACACCTCTTATAGCGGTTCGAAAACCCGTGGTGCCGCTG
GTACCCAGTATGACGATAGCAATATGGGCCAGAACAAAGTTGGCCTGAGCTTCTCGCTGCCGATTTATCAGGG
CGGAATGGTTAACTCGCAGGTGAAACAGGCACAGTACAACTTTGTCGGTGCCAGCGAGCAACTGGAAAGTGC
CCATCGTAGCGTCGTGCAGACCGTGCGTTCCTCCTTCAACAACATTAATGCATCTATCAGTAGCATTAACGCCT
ACAAACAAGCCGTAGTTTCCGCTCAAAGCTCATTAGACGCGATGGAAGCGGGCTACTCGGTCGGTACGCGTA
CCATTGTTGATGTGTTGGATGCGACCACCACGTTGTACAACGCCAAGCAAGAGCTGGCGAATGCGCGTTATAA
CTACCTGATTAATCAGCTGAATATTAAGTCAGCTCTGGGTACGTTGAACGAGCAGGATCTGCTGGCACTGAAC
AATGCGCTGAGCAAACCGGTTTCCACTAATCCGGAAAACGTTGCACCGCAAACGCCGGAACAGAATGCTATT
GCTGATGGTTATGCGCCTGATAGCCCGGCACCAGTCGTTCAGCAAACATCCGCACGCACTACCACCAGTAACG
GTCATAACCCTTTCCGTAACTGA

Protein sequence:

MKKLLPILIGLSLSGFSSLSQAENLMQVYQQARLSNPELRKSAADRDAAFEKINEARSPLLPQLGLGADYTYSNGYRD
ANGINSNATSASLQLTQSIFDMSKWRALTLQEKAAGIQDVTYQTDQQTLILNTATAYFNVLNAIDVLSYTQAQKEAI
YRQLDQTTQRFNVGLVAITDVQNARAQYDTVLANEVTARNNLDNAVEQLRQITGNYYPELAALNVENFKTDKPQ
PVNALLKEAEKRNLSLLQARLSQDLAREQIRQAQDGHLPTLDLTASTGISDTSYSGSKTRGAAGTQYDDSNMGQNK
VGLSFSLPIYQGGMVNSQVKQAQYNFVGASEQLESAHRSVVQTVRSSFNNINASISSINAYKQAVVSAQSSLDAME
AGYSVGTRTIVDVLDATTTLYNAKQELANARYNYLINQLNIKSALGTLNEQDLLALNNALSKPVSTNPENVAPQTPE
QNAIADGYAPDSPAPVVQQTSARTTTSNGHNPFRN-LEHHHHHH

Primer design:

F1 Primer: 5’ AGGAGATATACCATGAAGAAATTGCTCCCCATTC 3’

Overhang 15bp Primer 19 bp

Length: 33

GC%: 41.2

Tm (°C): 67.1

Gene Specific Tm (°C): 59.1

R1 Primer: 5’ GGTGGTGGTGCTCGAGGTTACGGAAAGGGTTATGACCG 3’

Length: 38

GC%: 57.9

Tm (°C): 75.9

Gene Specific Tm (°C): 60.3