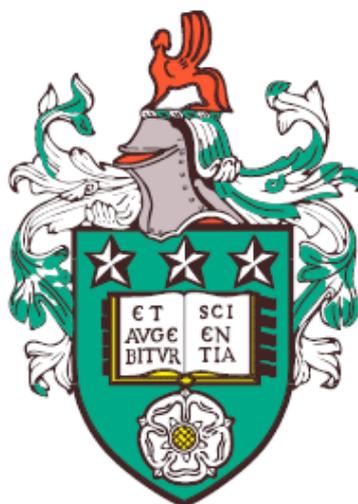


The roles of periplasmic chaperones and the β -barrel assembly machinery complex in outer membrane protein biogenesis



James Edward Andrew Horne

School of Molecular and Cellular Biology
Astbury Centre for Structural Molecular Biology
University of Leeds

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

December 2019

Declaration

The candidate confirms that the work submitted is his/her/their own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others. This copy has been supplied on the understanding that it is copyright material and that no quotation from this thesis may be published without proper acknowledgement.

Throughout this thesis the work directly attributable to the candidate is as follows: (i) Literature research and compilation of the manuscript stated above. (ii) The candidate performed all the experimental work and data analysis unless otherwise stated.

Chapter 3 includes experiments and results from OmpA-Skp crosslinking experiments and the development of the tag-transfer crosslinking methodology which was reproduced in the following publication: Horne, J.E., Walko, M., Calabrese, A.N., Levenstein, M.A., Brockwell, D.J., Kapur, N., Wilson, A.J. and Radford, S.E., (2018). Rapid Mapping of Protein Interactions Using Tag-Transfer Photocrosslinkers. *Angewandte Chemie International Edition*, 57(51), pp.16688-16692. All work shown in this chapter was performed by the candidate except for the synthesis of tag-transfer crosslinkers which was performed by M.W. The design and building of the UV LED lamp and acrylic chip was performed by J.E.H., M.A.L., and N.K. Mass spectrometry was performed by the University of Leeds Mass Spectrometry Facility and A.N.C., all data analysis was performed by J.E.H.

Chapter 5 includes contributions from Dr. Paul White, University of Leeds and Benji Bateman, Central Laser Facility, Research Complex at Harwell. P.W. and J.E.H. cloned constructs for *in vivo* experiments, prepared samples for, and performed cryoSTORM imaging. P.W. performed *in vivo* assays on BamA-fusion protein function. J.E.H. performed data analysis. B.B. performed cryoSTORM image acquisition and prepared custom built cryoSTORM microscope.

James Edward Andrew Horne
December 2019

Acknowledgements

Firstly, I would like to thank my supervisors Professor Sheena Radford and Dr David Brockwell for their support, encouragement and enthusiasm for the project and for keeping me going despite a couple of years of failed experiments and a general lack of data at the start of the project. Not to mention their good ideas to escape from the odd scientific quagmires that appeared along the way, and their willingness to read drafts of this thesis despite the late hour at which I presented them!

Thank you too all the members of the OMPire, old and new. Dr Bob Schiffrin who guided me through my first forays into the land of the β -barrel and was always happy to discuss bits of literature and nagging questions I had (even when his ears were plugged). To Dr Anna Higgins who was my long time office desk neighbour and who had the misfortune to be the closest person to speak to when a random thought or interesting tidbit came to me. Dr Julia Humes for all her help when I started and her general enthusiasm for the OMPs. Dr Anton Calabrese for his bottomless knowledge of mass spectrometry and his capacity to find time to help out newbies with experiments, despite always seeming to be very busy. Dr Tom Watkinson for introducing me to climbing in my early days in the lab, as well as my soul mate, Bakery 164, and being up for the odd jam session with Owen. Dr Paul White who integrated seamlessly and quickly into the Radfords and who has been an integral part of pushing the cryoSTORM project forward (Chapter 5), and who never once made me feel shame at scoffing down £6 worth of snacks every evening when we were working down in Harwell. To Dr Matt Watson for all his assistance in getting the ALEX box up and running being a generally friendly chap. And to the newest members of the OMPire, Sam and James who I'm sure will go on to do great PhDs.

Of course, the rest of the Radford/Brockwell lab have been brilliant and I feel lucky to have found a kind and sociable group of lab mates. Mike Davies for being an enabler of my addiction to Bakery 164 sandwiches and Miro coffee, as well as a steadfast climbing buddy. Dr Matt Jackson for his potent blend of nihilism and kind-heartedness. My sole yeargroup buddy Emma Cawood who I could always rely on to work even crazier hours

than I did (and who never pulled out of footie at the last minute). All the members of climbing crew past and present, those who join me for football, and those who enjoy a drink (think that pretty much covers everyone!).

An extra special shout has to go to Nasir Khan for having the unenviable task of trying to keep the lab running despite all the challenges that have been thrown at him (particularly in the last year or so!). Despite this, he always happily plied me with food, even regular homemade curries, and I never once saw him fail to make time to help others.

Outside of the Radford lab, thanks to my first year friends providing plenty of joy and opportunities to socialise, and particularly to the 135 Ash Road boys (Ray Li, Dan O'Brien, Dave Kealy and Christian Nathan), the 9 De Lacy mount Krew (Dave Kealy, Christian Nathan, and Sophie Hesketh), Jack Bravo for letting me squash him at squash, and Alex Borodavka for making sure I stayed caffeinated.

I had the opportunity to spend three months on a PIPS placement at the Research Campus in Harwell in the Central Laser Facility and I'm grateful to the team there for their support, and their ongoing collaboration which led to Chapter 5, particularly Dr Laura Zanetti-Domingues who looked after me for those three months in Oxford and has also been a constant source of encouragement and enthusiasm for the project.

Thanks as always has to go to my Mum, Dad, and brother, for their unwavering and unconditional support. Knowing they were always there for me when if I needed help made the whole process of doing a PhD a lot easier. Thanks to my friends near and far for sticking with me despite the gaps in geography and lifestyle and my terrible ability to keep in contact. Looking forward to catching up with all of them some more now this thesis is complete! I'm very grateful to have met Rom here at Leeds, who has kept me sane, fed, and watered throughout this last year, and I look forward to returning the favour when it's your turn to write up!

Finally, I am grateful to the BBSRC White Rose DTP whose sponsorship has allowed me to do what I love for a living.

Abstract

The outer membrane of bacteria is a complex and important structure representing the first (and most 'fortified') line of defense against insults from the extracellular medium as well as a platform for adhesion, recognition, and nutrient acquisition. As such, the outer membrane is a prime target for rational design of new drugs - both in terms of designing new classes of bacteriocidal agents, and also to sensitize bacteria to improve the efficacy of existing antibiotics. Our current knowledge suggests that all essential routes to the assembly of the outer membrane itself (both the lipids and proteins that form its structure) pass through the β -barrel assembly machinery complex (BAM complex), either directly or indirectly. This nanoscale machine is conserved across all bacteria containing an outer membrane and although the exact constituent parts vary, there is a common architecture scaffolded around one completely conserved protein, BamA. The BAM complex is responsible for the ATP- and protonmotive-independent assembly of integral transmembrane β -barrel proteins commonly referred to as outer membrane proteins (OMPs). Despite its essential role, the availability of high resolution structures, and over 15 years of biochemical studies, many questions about its mechanism of action remain unanswered.

In this thesis new methods for studying OMP biogenesis through the use of cross-linking mass spectrometry and cryogenic super-resolution microscopy have been developed and applied to study the interaction with a model OMP, OmpA, and its chaperones Skp and SurA, as well as mapping its interaction with the BAM complex during folding, giving new insights into the mechanism of SurA chaperoning and suggesting a possible mechanism and route for the transit of an OMP from SurA and through BAM. Cryogenic super-resolution microscopy is used to provide preliminary insights into the nanoscale organisation of the BAM complex and OmpA, as well as two-colour co-localisation of these proteins. Kinetic assays are used alongside fluorescent probes of lipid order and single-molecule FRET to study the role of the lipid environment on BAM-catalysed, BamA-catalysed, and uncatalysed folding of tOmpA, both with and without SurA.

In Chapter 3, a new crosslinking method was developed and validated on the Skp-OmpA chaperone-substrate pair. This was then applied to the other major OMP chaperone, SurA, where it could be shown that the binding activity of SurA resides almost exclusively in the core N- and C-terminal domains. Finally, this approach was used to try and capture a folding intermediate of OmpA as it was passed from SurA through the BAM complex and then analyse the interactions from OmpA to these partners during folding. A position at the bottom of the first (N-terminal) β -strand of OmpA makes crosslinks with the POTRA1, 4, and 5 domains of BamA as well as with the N-terminus and P2 of SurA, suggesting a greater recruitment of SurA P2 during OMP folding. This pattern of crosslinks from OmpA also implies a possible route from POTRA1, via BamD near the interface with BamA-POTRA5, onto β 1 of BamA, with the crosslinks at POTRA4 formed last as the final β -strand is appended to the nascent barrel.

In Chapter 4, the hypothesis that BAM functions by disordering lipids in the membrane was tested by using a number of techniques. DMPC was used as a model bilayer to be able to control the phase of lipids by conducting experiments at, below, or above, the transition temperature of 24 °C. The kinetics of tOmpA folding into DMPC liposomes showed that the full BAM complex is a much better catalyst for OMP folding (as measured by t_{50}) than BamA alone when below or at the transition temperature of DMPC, and slightly better when above. While the BAM complex could accelerate the formation of folded tOmpA almost 16X over uncatalyzed folding at the transition temperature (24 °C), folding via BamA was only marginally faster (at 1.5 fold the uncatalyzed t_{50}) which prompted further studies into the ability of these proteins to affect lipid order. The packing of lipids was assessed directly using the lipid order probe, laurdan, and the dynamics and conformational ensemble of the BAM complex was measured at a single-molecule level using FRET (smFRET). Laurdan experiments found that the presence of the BAM complex causes a broadening of the phase transition region as well as a 2 °C fall in the transition temperature implying a stabilisation of the liquid phase by BAM. smFRET studies showed that two populations of the BAM complex exist in solution, corresponding to the predicted FRET efficiencies of the open and closed states and these do not appear to interconvert on a 100s of μ s to 100s of ms timescale.

In Chapter 5, the organisation of the BAM complex is probed by using a novel method in super-resolution microscopy, cryoSTORM. By plunge-freezing samples of *E. coli* expressing BamA and/or OmpA fused with fluorescent proteins or the self-labelling HaloTag protein, these proteins could be visualised in their assembled state on the surface of a bacteria trapped in a frozen-hydrated state with sub 5 nm localisation precision. This showed the arrangement of molecules of BAM into discrete ‘islands’ spotted throughout

the cell surface as well as smaller islands formed by OmpA showing that OMPs are prone to cluster together in small islands. Initial two-colour studies of OmpA and BamA suggest a relatively low degree of co-localisation for these proteins.

Table of contents

List of figures	xxi
List of tables	xxvii
Nomenclature	xxix
1 Introduction	1
1.1 The importance of understanding the outer membrane and outer membrane protein biogenesis	1
1.1.1 Outer membranes are of huge importance to human health and understanding the biosphere	1
1.1.2 The discovery of BamA and its central role in OM biogenesis	3
1.1.3 BamA becomes the BAM complex	6
1.1.4 BamA remains, the subunits change	7
1.1.5 The BAM complex as a major hub of OM biogenesis	9
1.2 The diversity of OMPs	9
1.2.1 The OMP-ome of <i>E. coli</i>	10
1.2.2 The scope of BAM-dependent substrates	15
1.3 Outer membrane proteins <i>en route</i> to the outer membrane	19

1.3.1	From translation to the inner membrane	19
1.3.2	Emergence into the periplasm: more ATP-independent chaperones come into play	21
1.4	Recognition of OMPs and their delivery to the BAM complex	26
1.4.1	SurA and a possible role in bridging the IM and OM	26
1.4.2	BamB plays an important role in OMP biogenesis	28
1.4.3	BamD may be a site for initial engagement and prime BamA for folding	29
1.4.4	BamE modulates the conformations of BamA and installs a stress sensor	32
1.4.5	Recognition of OMPs by BamA	34
1.5	The final step: mechanisms of membrane insertion	36
1.5.1	Thermodynamic barriers to folding	36
1.5.2	Direct role of the BAM complex in OMP assembly: lowering the energy barriers	37
1.5.2.1	Effects of BAM at the lipid interface and the role of POTRAs	38
1.5.2.2	Templated folding of OMPs by the β 1- β 16 seam	39
1.5.3	Summary of possible BAM mechanisms	43
1.6	Conclusion and open questions in the field	46
1.7	Aims of this thesis	47
2	Materials and Methods	49
2.1	Materials	49
2.2	Bacterial strains	52

2.3	Plasmids and primers	53
2.4	Molecular biology (preparatory) methods	58
2.4.1	Culture medium	58
2.4.2	Preparation of competent cells	59
2.4.3	Transformation of <i>E. coli</i> strains	59
2.4.4	Site-directed mutagenesis	60
2.4.5	Plasmid DNA purification	60
2.4.6	<i>In vivo</i> growth and expression of BamA and OmpA variants in the JCM166 cell line	61
2.4.7	Agarose gel electrophoresis	62
2.4.8	Sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE)	62
2.4.9	Low SDS-PAGE gel	63
2.4.10	Production of large unilamellar liposomes (LUVs)	63
2.4.11	Protein expression and purification from inclusion bodies (BamA, OmpA, and tOmpA)	64
2.4.12	Protein expression and purification of periplasmic chaperones (Skp and SurA)	65
2.4.13	Protein expression and purification of the BAM complex and cysteine variants	66
2.4.14	Determining BAM concentration (Pierce BCA Protein Assay)	67
2.4.15	Determining molecular mass by denaturing MS	68
2.4.16	Refolding and enrichment of folded BamA	69
2.4.17	Production of BamA- and BAM-containing proteoliposomes	69
2.4.18	Labelling BAM with fluorescent dyes	70

2.4.19	Reconstitution of BAM into DMPC for smFRET studies	71
2.5	Biochemical (experimental) methods	72
2.5.1	Band-shift assay (end-point)	72
2.5.2	Band-shift assay (Kinetics of Tertiary Structure formation by Electrophoresis, KTSE)	72
2.6	Biophysical methods	73
2.6.1	Fluorescence spectroscopy (intrinsic): steady state emission scans	73
2.6.2	Fluorescence spectroscopy (intrinsic): kinetics	73
2.6.3	Dynamic light scattering (DLS)	74
2.6.4	Laurdan doped liposomes: measuring lipid phase and T_m	74
2.6.5	Single-molecule FRET experiments	75
2.6.6	cryoSTORM super-resolution microscopy	77
2.6.7	Data fitting	78
2.7	Bioinformatic methods	79
2.7.1	Multiple sequence alignment	79
2.7.2	<i>E. coli</i> K-12 OMP-ome identification	79
2.7.3	Cys and Lys abundances by subcellular location	80
2.7.4	BAM folding rate <i>in vivo</i> / metabolic cost of OMPs	80
2.8	Cross-linking experiments	80
2.8.1	Labelling with maleimido-benzophenone	80
2.8.2	Labelling with MTS-daizirine or MTS-TFMD	81
2.8.3	Crosslinking sample preparation	81
2.8.4	Photo-crosslinking	81

2.8.5	Separation and enrichment of cross-linked products	82
2.8.5.1	Method 1	82
2.8.5.2	In-gel digestion	83
2.8.5.3	Method 2	83
2.8.5.4	Method 3	84
2.8.5.5	Mass spectrometry	84
2.9	UV LED system manufacture and testing	85
2.9.1	Acrylic chip fabrication	85
2.9.2	UV LED lamp, optics, and sample holder construction	86
2.9.3	Thermocouple measurements of UV LED with tube holder	86
2.9.4	Infrared thermography measurements of sample heating	86
2.10	Other software tools	87
3	The development and application of tag-transfer crosslinkers	89
3.1	Introduction	89
3.2	Results	91
3.2.1	Limitations of benzophenone crosslinkers for capturing dynamic and transient interactions	91
3.2.2	Validating the non-perturbing nature of a new crosslinker design on OmpA folding	97
3.2.3	Accelerating crosslinking rates to capture transient interactions with a new UV LED lamp design	103
3.2.3.1	Design of custom UV LED-based lamps	103
3.2.3.2	Lamp-associated sample heating across designs	105

3.2.3.3	UV LED lamp dramatically improves the rate of crosslinking	108
3.2.4	Development of the tag-transfer workflow	108
3.2.4.1	Strategies for transfer of an MS-stable tag to crosslinked target proteins	111
3.2.4.2	Improving signal quality and number of PSMs with novel enrichment strategies	113
3.2.5	Interrogating the interaction between OmpA and Skp	117
3.2.5.1	OmpA tumbles inside the cavity of Skp in a dynamic, disordered state	117
3.2.5.2	Positions on OmpA have no strong motif or structural preference for locations within the cavity of Skp	120
3.2.6	OmpA binds to conserved domains of SurA in a manner consistent with ‘beads on a string’ model of chaperoning	124
3.2.6.1	Chaperoning activity of SurA primarily involves its core domain	124
3.2.7	Diazirine crosslinkers are small, unbiased reactive probes	125
3.2.7.1	The MTS-diazirine tag-transfer reagent is one of the smallest crosslinking molecules	125
3.2.7.2	Diazirine photocrosslinkers are globally reactive and unbiased	128
3.2.8	Towards capturing the reaction coordinate of OmpA on BAM	129
3.2.8.1	Tag-transfer crosslinking after initiating folding on BAM	129
3.2.8.2	Putative interactions of an OMP folding through BAM	130
3.2.8.3	Improving crosslink enrichment for crosslinking to BAM	132
3.3	Discussion	135

3.3.1	A new tool for integrative structural biology	136
3.3.2	Prospects for further refinements of the tag-transfer method and UV LED lamp	136
3.3.3	Probing dynamic interactions of OMPs reveals mechanisms of chaperoning	138
3.3.4	Capturing the folding of OmpA through BAM	139
3.3.5	Summary	142
4	Lipid order and the role of the β-barrel assembly machinery	145
4.1	Introduction	145
4.2	Results	150
4.2.1	Designing an assay to assess the impact of lipid phase and packing on catalysed and uncatalysed OMP folding	150
4.2.2	BAM provides considerable catalytic enhancement at the transition temperature while BamA has only a minor effect	157
4.2.3	SurA enhances BAM activity and slows uncatalysed folding rate at 30 °C liquid phase, has minor or negative effect at other phases and with BamA	163
4.2.4	Effects of BamA and the BAM complex on lipid order and phase transitions	167
4.2.5	The conformational ensemble of the BAM complex is not signifi- cantly altered between lipid phases	175
4.3	Discussion	181
4.3.1	Folding into gel phase lipids	181
4.3.2	Energetic costs of poor folding yields	182
4.3.3	Comparison between the biophysics of DMPC phases amd the outer membrane of <i>E. coli</i>	184

4.3.4	Effect of SurA on OMP folding	186
4.3.5	First evidence for physiologically compatible folding rates	188
4.3.6	Effect of the BAM complex on membrane physical properties	189
4.3.7	Single-molecule FRET showing conformational states of BAM	190
4.4	Conclusion	191
5	Towards visualising the nanoscale organization of OMP biogenesis <i>in vivo</i>	193
5.1	Introduction	193
5.2	Results	200
5.2.1	Approaches for labelling of OMPs	200
5.2.2	Imaging BamA-sfGFP	207
5.2.3	Imaging OmpA-sfGFP	211
5.2.4	Two-colour imaging of BAM and OmpA	213
5.3	Discussion	214
5.3.1	Limitations of the methodology	214
5.3.1.1	Limitations of a plasmid-expressed FP-fusion of OmpA	214
5.3.1.2	Expression levels of proteins	216
5.3.2	Localisation of OMPs	217
5.3.3	Future directions	218
5.3.3.1	Localisation limits of the technique	218
5.3.3.2	Genomic manipulations	219
5.3.3.3	Organisation of the whole biogenesis machinery	219
5.4	Conclusion	220

6 General Discussion	221
6.1 Overall conclusion of results	221
6.2 Where do we stand with models of the mechanism of BAM's catalytic activity?	226
6.3 Future studies	230
6.4 Final thoughts	232
References	233
Appendix A Modified peptides from Skp	285
Appendix B Modified peptides from SurA	289
Appendix C Full gels of each kinetic run	293
Appendix D Single and double exponential fits to folding kinetics by gel	303
Appendix E Parameters from double exponential fits to gel kinetics	323
Appendix F Parameters from single exponential fits to gel kinetics	325

List of figures

1.1	Bacterial cell envelope organisation	3
1.2	Structure of lipopolysaccharide	4
1.3	Structures of lipid A from various species	5
1.4	Subunits of BAM in <i>E. coli</i>	7
1.5	TolB-Pal Interaction	9
1.6	Gallery of all <i>E. coli</i> K-12 OMPs	16
1.7	OMP biogenesis pathway	20
1.8	Structures of periplasmic OMP chaperones	23
1.9	Structures of BAM and Omp85 homologues	30
1.10	Evidence for disordering of lipid around BamA	40
1.11	The lateral gate and exit pore of the BAM complex	41
1.12	Mechanisms of catalysis by the BAM complex	45
3.1	Diazirine photodecomposition and reaction scheme	90
3.2	Subcellular abundance of cysteine	92
3.3	Structure of OmpA and mutant positions	93
3.4	Maleimido-benzophenone OmpA(N _{cys}) x BAM	95

3.5	Structure of Skp trimer in apo and tOmpA-bound states	96
3.6	Maleimido-benzophenone OmpA(N _{cys}) x Skp	98
3.7	Illustration of products from a benzophenone XL reaction between OmpA(N _{cys}) and Skp	99
3.8	Tag-transfer reagents and workflow	100
3.9	Conjugation efficiencies of MTS-diazirine and MTS-TFMD to OmpA(W7C) or OmpA(T144C)	101
3.10	Conjugation of tag-transfer crosslinkers doesn't perturb OmpA folding	102
3.11	Design of new UV LED crosslinking lamps	104
3.12	IR thermography measurements of sample and system during UV crosslinking	106
3.13	Thermocouple measurements of tube sample during UV crosslinking	107
3.14	Sample heating at irradiation times required to reach specific yields	108
3.15	Relative rates of crosslinking measured by SDS-PAGE	109
3.16	Enrichment and analysis strategies for MTS-diazirine and MTS-TFMD	110
3.17	Reaction scheme for crosslinking and alkylation by MTS-diazirine	111
3.18	Masses and structures of crosslinkers, derivatives, and tags used throughout this methodology	112
3.19	Growth of search space for BS3 versus tag-transfer crosslinkers	114
3.20	Output of each stage in tag-transfer processing and enrichment	116
3.21	Number of unique modifications detected for each tag-transfer method	118
3.22	Method 3 representative mass spectra	119
3.23	The results of tag-transfer MS-XL mapped onto Skp	121
3.24	Abundance of crosslinks at each position on Skp	122

3.25	Comparison of binding hotspots with structural and sequence properties of Skp	123
3.26	All crosslinks observed from OmpA to SurA	126
3.27	Comparison of sizes of tag-transfer reagents with common crosslinkers	127
3.28	Reactivity profile of tag-transfer reagents to accessible residues of Skp	129
3.29	Crosslinking OmpA to BAM	131
3.30	Crosslinks from OmpA to BAM and SurA under folding conditions	133
3.31	Example of outcome of residue-specific crosslinking experiment versus tag-transfer	135
3.32	Illustrations of potential modes of the BAM-elongation model	141
3.33	BamA-BamD electrostatic network and crosslinks	143
4.1	Electrostatic profile of the BAM complex	147
4.2	BamA conformations	148
4.3	Thermotropic behaviour of DMPC	149
4.4	DLS of liposomes	153
4.5	Effect of extrusion on BAM-catalysed rates	154
4.6	Curvature sensed by tOmpA	155
4.7	tOmpA occupancy on SurA as a function of concentration	155
4.8	Illustration of folding method	156
4.9	Comparison of single and double exponential fits	157
4.10	Example gels for each condition	158
4.11	t_{50} rates and catalytic fold change for tOmpA folding	160
4.12	t_{50} fold change vs. 24 °C	162

4.13	Fraction folded, and change in yield for tOmpA folding	164
4.14	Change in t_{50} due to SurA	166
4.15	Change in yield due to SurA	168
4.16	Laurdan incubation with liposomes	171
4.17	Lipid order measured by laurdan	172
4.18	T_m by laurdan for multiple BAM batches	173
4.19	Change in GP during BamA folding	175
4.20	Labelling of BAM for smFRET	177
4.21	Distinguishing the open and closed state of a labelled BAM(R127C/N520C) complex	177
4.22	smFRET studies of BAM(R127C/N520C) in DMPC at different lipid phases	178
5.1	Modelled BAM cluster size	196
5.2	Labelling strategies and localisation uncertainty	197
5.3	BamA tolerates N-terminal fusion proteins.	201
5.4	Examples of failed methodologies	206
5.5	Single-molecule localisation workflow.	208
5.6	Example sfGFP-BamA cryoSTORM reconstructions.	210
5.7	Example OmpA-sfGFP cryoSTORM reconstructions.	212
5.8	Example two-colour images of BamA and OmpA	215
1	tOmpA folding via BAM (with and without SurA), 20 °C	294
2	tOmpA folding via BAM (with and without SurA), 24 °C	295
3	tOmpA folding via BAM (with and without SurA), 30 °C	296

4	tOmpA folding via BamA (with and without SurA), 20 °C	297
5	tOmpA folding via BamA (with and without SurA), 24 °C	298
6	tOmpA folding via BamA (with and without SurA), 30 °C	299
7	tOmpA folding into empty liposomes (with and without SurA), 20 °C . .	300
8	tOmpA folding into empty liposomes (with and without SurA), 24 °C . .	301
9	tOmpA folding into empty liposomes (with and without SurA), 30 °C . .	302
10	Exponential fits to tOmpA folding via BAM without SurA, 20 °C	304
11	Exponential fits to tOmpA folding via BAM with SurA, 20 °C	305
12	Exponential fits to tOmpA folding via BAM without SurA, 24 °C	306
13	Exponential fits to tOmpA folding via BAM with SurA, 24 °C	307
14	Exponential fits to tOmpA folding via BAM without SurA, 30 °C	308
15	Exponential fits to tOmpA folding via BAM with SurA, 30 °C	309
16	Exponential fits to tOmpA folding via BamA without SurA, 20 °C	310
17	Exponential fits to tOmpA folding via BamA with SurA, 20 °C	311
18	Exponential fits to tOmpA folding via BamA without SurA, 24 °C	312
19	Exponential fits to tOmpA folding via BamA with SurA, 24 °C	313
20	Exponential fits to tOmpA folding via BamA without SurA, 30 °C	314
21	Exponential fits to tOmpA folding via BamA with SurA, 30 °C	315
22	Exponential fits to tOmpA folding into empty liposomes without SurA, 20 °C	316
23	Exponential fits to tOmpA folding into empty liposomes with SurA, 20 °C	317
24	Exponential fits to tOmpA folding into empty liposomes without SurA, 24 °C	318
25	Exponential fits to tOmpA folding into empty liposomes with SurA, 24 °C	319

- 26 Exponential fits to tOmpA folding into empty liposomes without SurA,
30 °C 320
- 27 Exponential fits to tOmpA folding into empty liposomes with SurA, 30 °C 321

List of tables

1.1	List of all <i>E. coli</i> K-12 OMPs	11
1.2	Abundance of <i>E. coli</i> K-12 OMPs	13
1.3	Validated OMP substrates for Skp and SurA	27
2.1	Materials	49
2.2	List of plasmids	53
2.3	Protein sequences	54
2.4	List of primers	57
2.5	SDS-PAGE recipe	63
2.6	Protein properties	65
2.7	Dye properties	71
3.1	UV LED lamp components	104
3.2	UV LED and commercial UV lamp cost comparison	104
3.3	Sample heating for each lamp design	105
3.4	Rate of crosslinking for each lamp design	109
3.5	Identified crosslinks from OmpA to BAM and SurA	130
4.1	Lipids and transition temperatures	151

4.2	Average t_{50} values	159
4.3	t_{50} fold change vs. 24 °C	161
4.5	Change in yield vs. 24 °C	163
4.4	Average folding yield and effect of BamA, BAM, or SurA	163
4.6	Change in t_{50} due to SurA	166
4.7	Change in yield due to SurA (Δ % Folded)	167
4.8	Change in yield due to SurA (Δ % Fold Change)	167
5.1	Periplasmic fluorescent protein reports	203
5.2	Cryogenic fluorescent protein reports	205
1	All modified Skp peptides (OmpA-MTS-diazirine XL)	285
2	All modified Skp peptides (OmpA-MTS-TFMD XL)	287
3	All modified SurA peptides (OmpA-MTS-diazirine XL)	289
4	All modified SurA peptides (OmpA-MTS-TFMD XL)	291
5	Fitting parameters from double exponential fits	324
6	Fitting parameters from single exponential fits	326

Nomenclature

Greek Symbols

β ME β -mercaptoethanol

μ M Micromolar

Other Symbols

K_d Dissociation constant

T_m Transition temperature

y_0 y-intercept

Acronyms / Abbreviations

Ab Antibody

ACF Autocorrelation function

ADP Adenosine diphosphate

AFM Atomic force microscopy

ALEX Alternating laser excitation

APD Avalanche photodiode

ASIL Asymmetric solid immersion lenses

ATP Adenosine triphosphate

BAM β -barrel assembly machinery

BCA Bicinchoninic acid

BP	Maleimido benzophenone
BS3	Bis(sulfosuccinimidyl)suberate
CG-MD	Coarse-grained molecular dynamics
CMC	Critical micelle concentration
DDM	n-dodecyl β -D-maltoside
DLPC	1,2-dilauroyl-sn-glycero-3-phosphocholine
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DMSO	Dimethyl sulfoxide
DOPC	1,2-dioleoyl-sn-glycero-3-phosphocholine
DOPE	1,2-dioleoyl-sn-glycero-3-phosphatidylethanolamine
DOPG	1,2-dioleoyl-sn-glycero-3-phosphatidylglycerol
DPH	1,3-diphenyl-1,3,5-hexatriene
DPPC	1,2-dipalmitoyl-sn-glycero-3-phosphocholine
DTPC	1,2-ditridecanoyl-sn-glycero-3-phosphatidylcholine
DTT	Dithiothreitol
DUPC	1,2-diundecanoyl-sn-glycero-3-phosphatidylcholine
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
ESI-IM-MS	Electrospray ionisation ion-mobility mass spectrometry
FHA	Filamentous hemagglutinin
FP	Fluorescent protein
FRET	Förster resonance energy transfer
GP	Generalized polarisation
Hg-Xe	Mercury-xenon

-
- HSQC Heteronuclear single quantum coherence
- IAA Iodoacetamide
- IM Inner membrane
- IMP Inner membrane protein
- IPTD Isopropyl β -D-1-thiogalactopyranoside
- I-TASSER Iterative threading and assembly refinement
- KTSE Kinetics of tertiary structure formation by electrophoresis
- LED Light emitting diode
- LPR Lipid-to-protein ratio
- LPS Lipopolysaccharide
- LUV Large unilamellar vesicle
- mAb Monoclonal antibody
- MBP Maltose binding protein
- MD Molecular dynamics
- MS/MS Mass spectrometry / mass spectrometry (Tandem mass spectrometry)
- MS Mass spectrometry
- MTS Methanethiosulfonate
- MWCO Molecular weight cutoff
- MWM Molecular weight marker
- NA Numerical aperture
- NMR Nuclear magnetic resonance
- OBPC 1-oleoyl-2-(9,10-dibromostearoyl)-phosphocholine
- OD600 Optical density at 600 nm
- OEP80 Outer envelope protein 80

OM	Outer membrane
OMP	Outer membrane protein
OMV	Outer membrane vesicle
P1	Peptidyl prolyl isomerase domain 1
P2	Peptidyl prolyl isomerase domain 2
PALM	Photo-activated localization microscopy
pBpa	p-Benzoyl-L-phenylalanine
PBS	Phosphate buffered saline
PC	Phosphatidylcholine
PDB	Protein data bank
PEG	Polyethylene glycol
PE	Phosphatidylethanolamine
PFAM	Protein families
PG	Phosphatidylglycerol
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPE	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine
POTRA	Polypeptide transport-associate
PPIase	Peptidyl prolyl isomerase
PSM	Peptide spectrum matches
RASP	Recurrence analysis of single particles
SAM	Sorting and assembly machinery
SANS	Small angle neutron scattering
SAXS	Small angle X-ray scattering
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

-
- SEC Size exclusion chromatography
- sfGFP Superfolder GFP
- smFRET Single-molecule FRET
- STORM Stochastic optical reconstruction microscopy
- SUV Small unilamellar vesicle
- T2SS Type 2 secretion system
- T3SS Type 3 secretion system
- T4SS Type 4 secretion system
- TAM Translocation and assembly module
- TBDR TonB-dependent receptor
- TCEP Tris(2-carboxyethyl)phosphine
- TFMD Trifluoro-methyl-phenyldiazirine
- TOB Topogenesis of outer membrane β -barrel proteins
- TOC Translocon at the outer envelope membrane of chloroplasts
- TPR Tetratricopeptide repeat
- uOMP Unfolded outer membrane protein
- UV Ultra-violet
- WT Wild-type
- XL-MS Crosslinking mass-spectrometry

Chapter 1

Introduction

1.1 The importance of understanding the outer membrane and outer membrane protein biogenesis

1.1.1 Outer membranes are of huge importance to human health and understanding the biosphere

Prokaryotes make up ~20 % of all biomass on the Earth and are the most abundant living cells on the planet by far, making up >99 % of cells (Whitman et al., 1998; Kallmeyer et al., 2012; Landenmark et al., 2015). Prokaryotes themselves can be divided into bacteria and archaea. Although archaea form a natural part of human and animal microbiomes (Lurie-Weinberger and Gophna, 2015), archaea have never been directly implicated in pathogenesis (Cavicchioli et al., 2003; Gill and Brinkman, 2011) and so understanding how bacteria work not only teaches us about a major part of the biosphere, but is also of utmost relevance to human health and disease.

Bacteria can be divided into two groups dependent on the structure of their cell wall, those which contain an outer membrane (OM) and those which do not (Figure 1.1). The Gram stain is commonly used to distinguish between these two types where those without an OM often have a thick layer of peptidoglycan which can retain crystal violet dye (giving them the name Gram-positive) and those with an OM generally have a much thinner peptidoglycan layer allowing the dye to be washed out (hence, Gram-negative) (Beveridge, 2001). However, exceptions to these rules exist:

- the phylum Firmicutes are phylogenetically ‘Gram-positive’ but contain at least two classes - the Negativicutes and Halanaerobiales - which have an OM and stain Gram-negative (Antunes et al., 2016; Poppleton et al., 2017)
- the Tenericutes phylum stain Gram-negative but in fact lack both peptidoglycan and an OM (Skenner et al., 2016)
- the phylum Actinobacteria contain the families Mycobacteria and Corynebacteria which are phylogenetically Gram-positive but have an ambiguous Gram-stain and an unusual and unique OM bilayer structure composed of mycolic acid lipids with chain lengths of 30 to 90 carbons that fold back on themselves (Hoffmann et al., 2008; Zuber et al., 2008)
- the Deinococcus-Thermus phylum contains examples of species which stain Gram-positive (due to a thick peptidoglycan layer) but are in fact classical diderms (species containing two membranes) as well as those which stain Gram-negative as expected (Castán et al., 2002; Yu and Lu, 2019)
- the Chloroflexi phylum stain Gram-negative but only contains one membrane (Sutcliffe, 2011)
- the newly discovered Caldiserica phylum (with a single member, *Caldisericum exile*) stains Gram-negative and electron microscopy appears to show an outer membrane but lacks any genes or proteins commonly associated with such (Mori et al., 2009; Heinz and Lithgow, 2014; Antunes et al., 2016)
- ...and many more unusual architectures likely remain to be discovered

In this thesis diderm bacteria (those containing an inner membrane and an OM with integral transmembrane β -barrel proteins) will be referred to as ‘Gram-negative’ in accordance with the literature tradition and all others as ‘Gram-positive’. Although a huge amount of microbial diversity has been discovered through environmental metagenomic surveys (suggesting 92 phyla of bacteria) (Hug et al., 2016; Parks et al., 2017), under the International Code of Nomenclature of Bacteria, organisms must be cultured to allow their phenotypic description before they can be named. Of the 34 phyla of bacteria into which all of these described species are contained (as documented by the List of Prokaryotic Names with Standing in Nomenclature [<http://www.bacterio.net/>]) (Parte, 2018), only four (Actinobacteria, Chloroflexi, Firmicutes, Tenericutes) have had Gram-positive examples described (with Firmicutes appearing to have lost its OM evolutionarily recently) (Heinz and Lithgow, 2014; Antunes et al., 2016). The Gram-negative envelope thus appears

to be the most common architecture for bacteria, although it should be cautioned that a large grouping of uncultured and undescribed bacteria (the super-phylum, Candidate Phyla Radiation) may not contain an OM (Méheust et al., 2019). Nonetheless, likely owing to this abundance (at least in environments relevant to human habitation), but also to the protective qualities of the OM (which may have evolved as an antibiotic barrier, Gupta (2011)), 9 out of 12 bacteria on a list published by the World Health Organisation which were deemed as priorities for the research and development of new antibiotics were Gram-negative bacteria (Tacconelli et al., 2018).

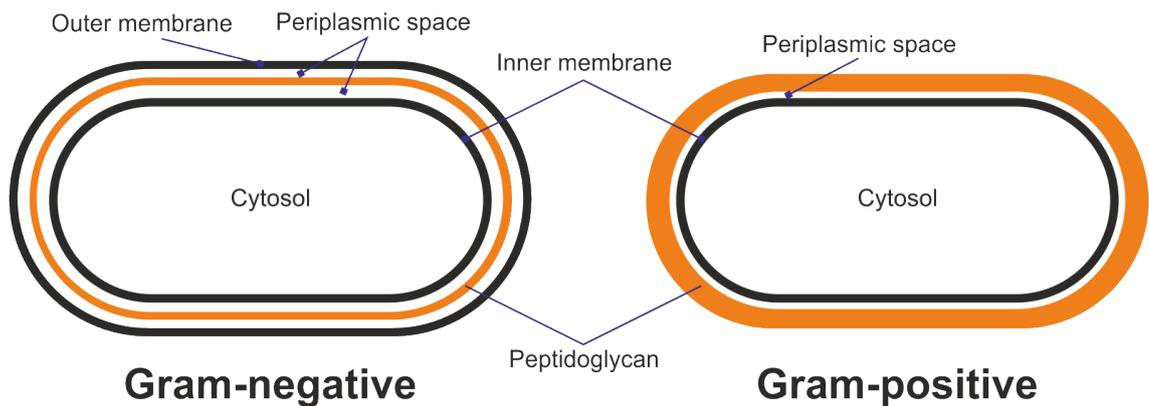


Fig. 1.1 The organisation of the bacterial cell wall divides bacteria into two major groups: Gram-positive or Gram-negative, or alternatively, monoderms (one membrane) and diderms (two membranes). Gram-positive bacteria have a thicker layer of peptidoglycan that cover a relatively small and crowded periplasmic space (Zuber et al., 2006). Gram-negative bacteria have an additional membrane, the outer membrane, and a thinner layer of peptidoglycan. The presence of a double membrane allows more control over the periplasmic space and allows more space for cellular processes to occur in this region, relatively isolated from the cytoplasm.

The defining characteristics of the OM are usually taken to be the presence of lipopolysaccharide (LPS) (a lipid with multiple acyl chains embedded in the outer leaflet of the OM and a large polysaccharide headgroup - see Figure 1.2 and Figure 1.3) and transmembrane β -barrel proteins, also known as outer membrane proteins, or OMPs. The presence of LPS is not an absolute requirement, with some bacteria lacking major LPS synthesis and transport genes but containing alternative glycolipids (such as the phyla Spirochaetes, Deinococcus-Thermus, and Thermotogae) (Schultz et al., 1998; Vinogradov et al., 2004; Sutcliffe, 2010; Suda et al., 2012). Although the exact lipid composition of the OM varies, all described bacterial species which contain OMPs also contain the machinery for assembling them: the β -barrel assembly machinery complex (BAM complex) (Sutcliffe, 2010). Although the exact architecture of this protein machine varies, homologues of the central component, BamA, are always present (Webb et al., 2012a). Its essentiality, conservation throughout bacterial lineages, and its partial surface exposure, make it an obvious target for new antibiotics (Lehman and Grabowicz, 2019) and new therapeutics are (or were) being developed or investigated by a number of pharmaceutical companies

(e.g. Genentech, GlaxoSmithKline, Regeneron Pharmaceuticals) (Wedegge et al., 2013; Machutta et al., 2017; Vij et al., 2018; Storek et al., 2018a; Psonis et al., 2019).

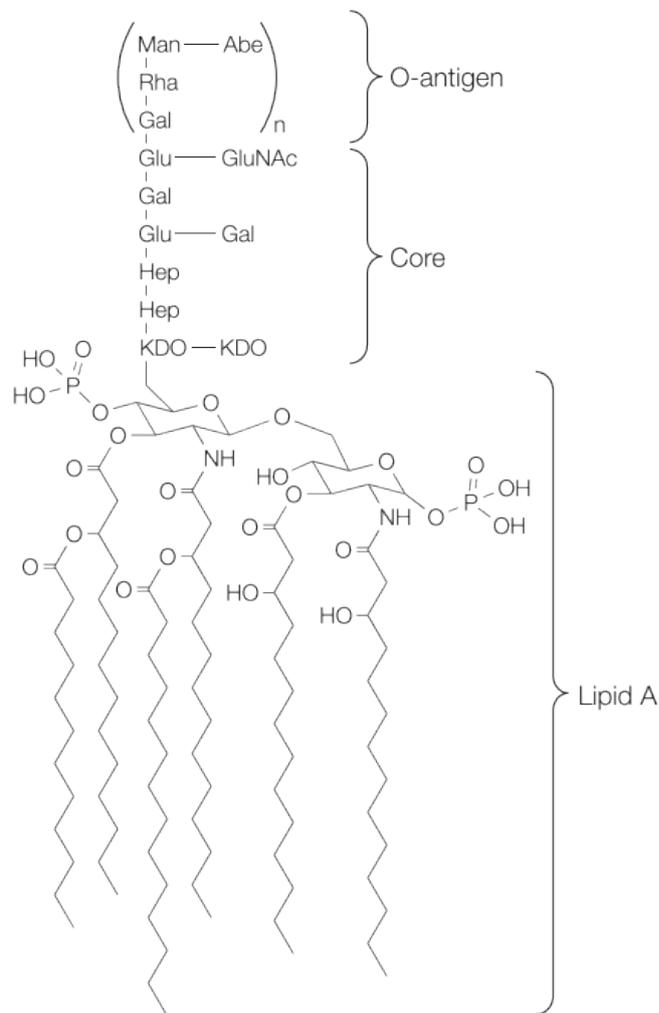


Fig. 1.2 Generic structure of lipopolysaccharide (LPS). LPS comprises three regions: lipid A is made up of 4–8 acyl chains and a disaccharide diphosphate headgroup (Kim et al., 2016a), a core region that is split into a highly conserved inner core with unusual sugars (such as 3-deoxy-D-manno-octulosonic acid, Kdo; and L-glycero-D-manno heptose, Hep) and an outer core with more common sugars (glucose, Glu; galactose, Gal; N-acetylglucosamine, GluNAc), and the O-antigen region (rhamnose, Rha; mannose, Man; abesquiose, Abe) which is highly variable - even between strains of the same species (Erridge et al., 2002). Figure modified from Miller et al. (2005).

1.1.2 The discovery of BamA and its central role in OM biogenesis

The role of BamA (originally YaeT) in the assembly of OMPs in bacteria was first discovered in *Neisseria meningitidis* where it was called Omp85 (establishing the protein family, Omp85) (Voulhoux et al., 2003; Genevrois et al., 2003). Around the same time an Omp85 homologue was found to play an essential role in assembling transmembrane β -barrel proteins in mitochondrial outer membranes of *Saccharomyces cerevisiae* (where it is called

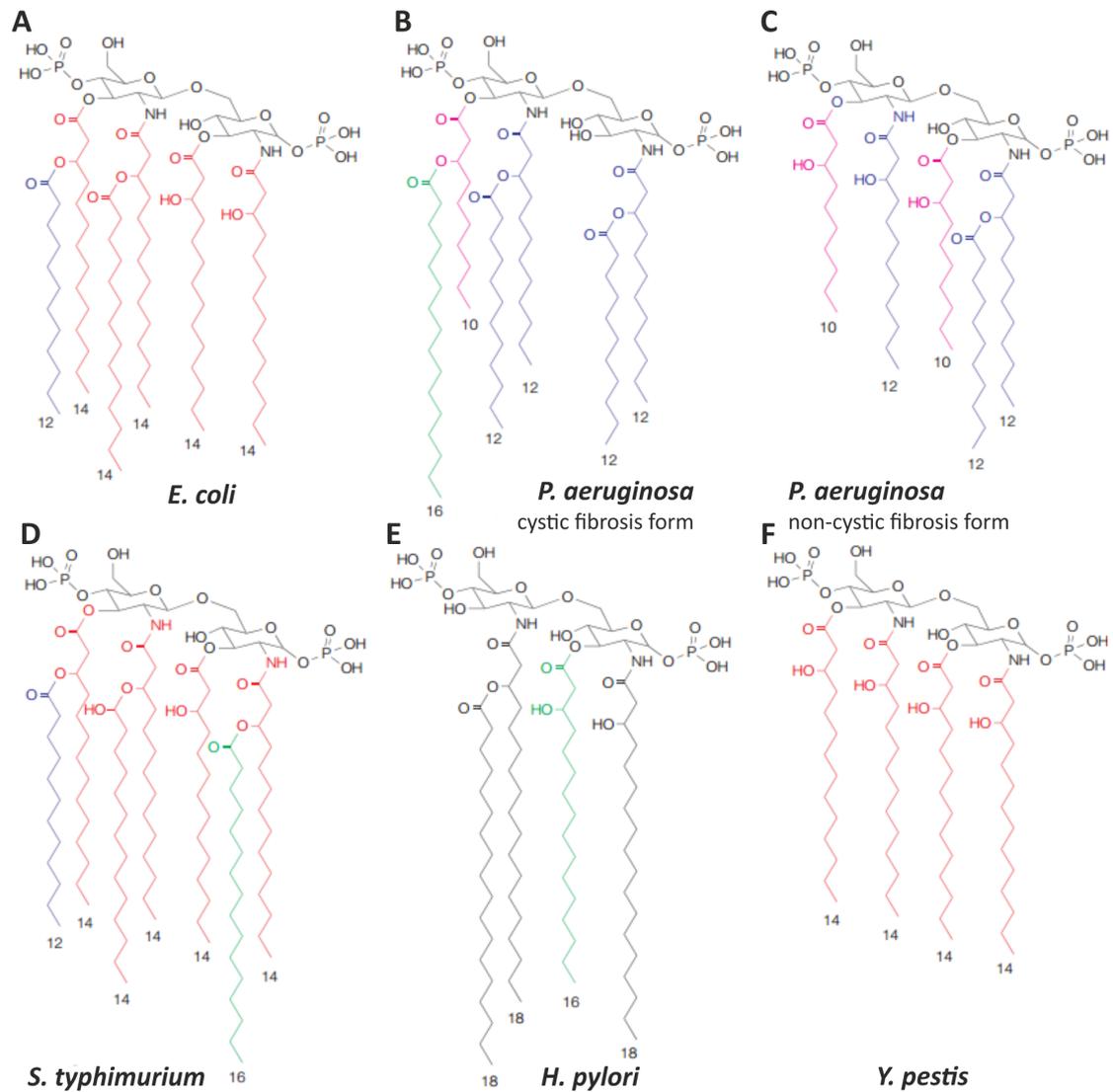


Fig. 1.3 Examples of some alternative forms of lipid A in Gram-negative bacterial outer membranes. This can vary in the number of acyl chains (4-7), the degree of branching, and the chain length. Each of these characteristic would be expected to affect the degree of packing, hydrophobic mismatch, and membrane elastic tension / stress. Image redrawn from Miller et al. (2005).

Sam50 [aka Tob55], as part of the Sorting and Assembly Machinery (SAM) complex [aka Topogenesis of Outer membrane β -barrel proteins (TOB) complex] (Wiedemann et al., 2003; Paschen et al., 2003). Although homologues of the protein responsible for the biogenesis of β -barrel proteins in chloroplasts were reported soon after (Gentle et al., 2004), it took a number of years before the protein responsible for the biogenesis of β -barrel proteins in chloroplasts was identified. This was likely due to the presence of a number of Omp85 homologues in the genomes of plant model organisms (e.g. *Arabidopsis thaliana* encodes at least four; Toc75-I, III, IV, and V – Translocon at the Outer Envelope Membrane of Chloroplasts, TOC) and the role of the major protein Toc75-III in chloroplast import as part of the TOC complex (Patel et al., 2008). However, it soon became clear that Toc75-V played a unique role unrelated to import and was thus renamed OEP80 (Outer Envelope Protein 80) (Inoue and Potter, 2004). Surprisingly, although OEP80 is the most likely candidate, direct experimental evidence for a role of OEP80 in transmembrane β -barrel biogenesis is still lacking and this function could reside in other homologues (Huang et al., 2011; Day et al., 2019).

1.1.3 BamA becomes the BAM complex

In the following years, it became clear that BamA did not operate in isolation and instead formed a complex with other proteins to aid the assembly of OMPs. In *E. coli*, the BAM complex has been most fully characterised and was found to comprise four additional lipoproteins: BamB (originally YfgL), BamC (originally NlpB), BamD (originally YfiO), and BamE (originally SmpA) (Ruiz et al., 2005; Wu et al., 2005; Sklar et al., 2007a) (see Figure 1.4), with BamD being the only essential lipoprotein under normal growth conditions (Malinverni et al., 2006). These form a 1:1:1:1 hetero-oligomeric protein complex which has been confirmed by X-ray crystallography and cryo-electron microscopy (cryoEM) structures (Bakelar et al., 2016; Gu et al., 2016; Han et al., 2016; Iadanza et al., 2016). However, complementary evidence suggests that these structures may have some artefactual inconsistencies due to the purification and preparation process. Nuclear magnetic resonance (NMR) analysis and crystal structures of BamC have shown it has an N-terminal ~75-residue disordered region, followed by two helix-grip domains with low sequence identity (~12 %), but significant structural homology to each other (Knowles et al., 2009; Warner et al., 2011; Albrecht and Zeth, 2011; Kim et al., 2011b). The structure of an isolated BamCD complex showed the disordered N-terminal tail to interact with BamD, but not the helix grip domain. This is recapitulated in all structures of the full BAM complex, where density corresponding to the disordered N-terminal region of BamC is always present bound to BamD, but the helix grip domains are generally much more

poorly resolved (inspection of the structure of the one exception in 5D0Q (Gu et al., 2016) where both domains are present suggests this is the result of crystal packing artefacts). Immunofluorescence against BamC in non-permeabilised *E. coli* cells showed robust labelling localised to clusters at the outer membrane, and protease shaving experiments indicated that the two helix-grip domains of this protein are surface exposed *in vivo* (Webb et al., 2012b; Rassam et al., 2015; Gunasinghe et al., 2018). Early structural studies on BamE by crystallography, NMR, and analytical size-exclusion chromatography (SEC) suggested that this protein was dimeric, or could exist in both monomeric and dimeric forms, although it was unclear whether this represented a domain-swapped artefact due to the expression and purification procedures, or the true physiological state of the protein (Albrecht and Zeth, 2011; Kim et al., 2011c; Knowles et al., 2011). Again, the solved crystal and cryo-EM structures of the BAM complex all suggested a monomeric form of BamE, but native mass spectrometry experiments have suggested that at least some population of the BAM complex exists in a BamABCD(E)₂ conformation, both when prepared by detergent extraction from the OM and when ejected directly from native *E. coli* membranes (Iadanza et al., 2016; Chorev et al., 2018).

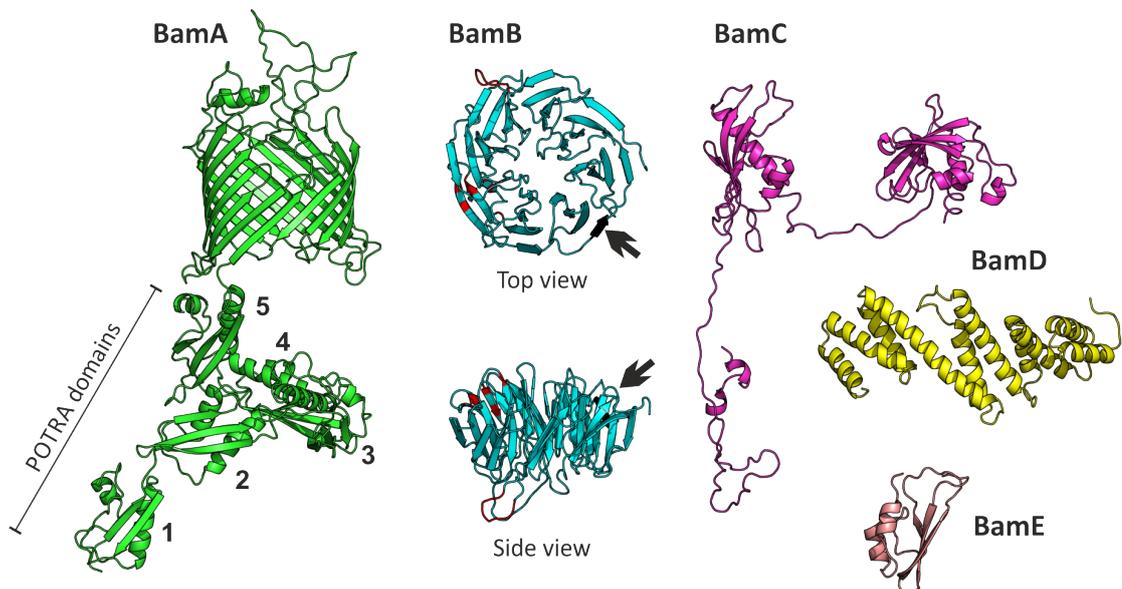


Fig. 1.4 Subunits of the BAM complex. Location of β -augmentation site observed in BamB crystal structure shown in black with black arrow (β -strand D of blade 6) (Heuck et al., 2011; O'Neil et al., 2015). Site of proposed BamB hydrophobic interactions with unfolded OMPs shown in red (Heuck et al., 2011). In the BamB top view, the side binding to the BamA POTRA domains is facing forward; the side view of BamB shows the POTRA binding region facing down (cf. Figure 1.7). Structures of individual subunits taken from the O'Neil et al. (2015) homology model of the full BAM complex with BamC modelled with its helix-grip domains in a surface exposed orientation.

1.1.4 BamA remains, the subunits change

This canonical organisation of the BAM complex and its four subunits, BamBCDE, can vary between species, with some subunits being omitted, but generally the members of the BAM complex in other species comprise homologues of these proteins (Webb et al., 2012a). Nonetheless, a number of unique BAM complex members have been reported. In *Caulobacter crescentus*, the BAM complex comprising BamABDE also associates with the additional subunits Pal (which is an essential protein like BamA and BamD and comprises an N-terminal disordered domain, and a C-terminal OmpA-like peptidoglycan binding domain) (Anwari et al., 2010) and BamF (which has a similar conserved BamD-binding sequence as BamC) (Anwari et al., 2012). Interestingly, in *E. coli* two homologues of Pal are known: Pal and YiaD. Pal is anchored to the outer membrane and interacts with OmpA, as well as forming a transmembrane complex with TolAQR complex in the inner membrane and TolB in the periplasm as part of the Tol-Pal complex (Godlewska et al., 2009). TolB has a β -propeller architecture similar to BamB (PDB: 2HQS) (Bonsor et al., 2007) and YiaD has been shown to be a multicopy suppressor for a temperature-sensitive mutant of BamD, with its peptidoglycan binding capacity being key for this suppression activity (Tachikawa and Kato, 2011). This provides both evidence for interaction with the BAM complex and a structural mechanism (via BamB) for achieving that, suggesting that homologous roles could be played in *E. coli* which are yet to be discovered. The BAM complex of *Neisseria meningitidis* comprises BamACDE and also has a stably associated homologue of Pal, RmpM, as well as two homologues of *E. coli* BamE (BamE and Mlp) (Volokhina et al., 2009). Reports have shown that BamE may be surface-displayed, and in $\Delta bamE$ strains, BamD can become surface exposed (Sikora et al., 2018). In *Borrelia burgdorferi*, immunoblot analysis of outer membrane vesicles (OMVs) show BamA associated with high molecular weight bands (with major bands at 200 kDa and 1 MDa) indicating that it formed large complexes. Analysis of the *B. burgdorferi* genome indicated the presence of only one homologue of the canonical subunits, BamD (BB0324), and BamA co-immunoprecipitation yielded a novel subunit, a 349-residue lipoprotein, BB0028 (Lenhart et al., 2012). Surprisingly, a $\Delta bamD$ mutant in this species showed no growth defects (Dunn et al., 2015). Although the role of the BB0028 subunit is unclear, structural modelling suggested a β -propeller architecture similar to BamB (Dunn et al., 2015), and varied reports suggested it may be surface exposed or face the periplasm (Lenhart et al., 2012; Dowdell et al., 2017). In *E. coli*, another Omp85 homologue called TamA forms part of the TAM complex along with the inner membrane protein, TamB, and is thought to be a ‘helper’ complex for the assembly of autotransporter proteins. *B. burgdorferi* encodes no homologue of TamA, but TamB was shown to associate directly with the BAM complex (Iqbal et al., 2016). The

authors also found that TamB is essential in *B. burgdorferi*, whereas deletion of this gene in *E. coli* shows no effect on growth (Selkrig et al., 2012). This validated the hypothesis that TamA arose from a gene duplication of BamA in Proteobacteria, while TamB was an ancestral protein involved in OMP biogenesis (Heinz et al., 2015) (as the Spirochaetes phylum which *B. burgdorferi* are a member of, diverged before Proteobacteria which *E. coli* are a member of) (Hug et al., 2016).

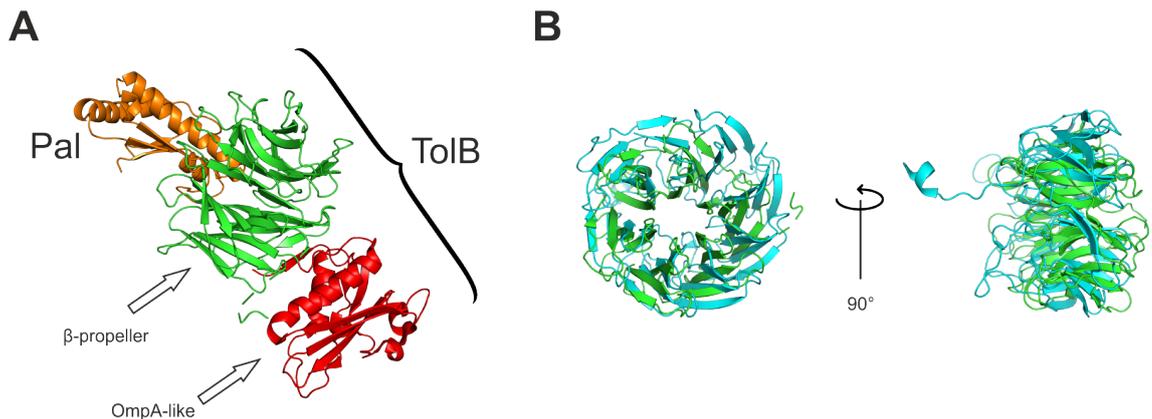


Fig. 1.5 Interaction between the periplasmic proteins TolB and Pal. (A) Interaction between TolB and Pal (2HQS) (Bonsor et al., 2007). (B) β -propeller domain of TolB (green) with BamB of the BAM complex (cyan) (5LJO).

1.1.5 The BAM complex as a major hub of OM biogenesis

In *E. coli*, all essential routes to the assembly of the lipids and proteins that comprise the OM pass through, or interact with, the BAM complex, highlighting its position as a major organisational hub. Knockout and transposon insertion studies on *E. coli* have identified the genes which are essential for the growth of *E. coli* and only 7 of these (*bamA*, *bamD*, *lolA*, *lolB*, *lptA*, *lptD*, and *lptE*) are predicted to reside in the periplasm or the OM (Baba et al., 2006; Yamazaki et al., 2008; Goodall et al., 2018; Loos et al., 2019). BamD is a lipoprotein which forms part of the BAM complex, LolA and LolB are part of the Lol pathway which shuttles and installs lipoproteins into the OM (including BamBCDE, although recent evidence suggests parallel pathways do exist) (Konovalova and Silhavy, 2015; Grabowicz and Silhavy, 2017; Grabowicz, 2018), and LptADE are part of the larger LPS transport pathway (comprising LptB-FG-C-A-DE, from cytoplasm to OM) (Okuda et al., 2016) where the BAM complex is responsible for assembling LptD upon which LptE and LptA are scaffolded (Lee et al., 2016).

1.2 The diversity of OMPs

OMPs carry out a diverse array of different functions in bacteria from maintaining the structure of the OM, cell-cell and cell-surface adhesion, host invasion, host immune evasion, passive and active nutrient uptake, regulation of ion flux, secretion of soluble extracellular factors, export of small molecules, installation of LPS, modification of phospholipids, cleavage of antimicrobial peptides and host factors, anchoring of pili and fimbria, sheathing or anchoring flagella and conjugation machinery, amongst many others. To carry out these varied roles, bacteria have evolved a large number of different architectures of OMPs built around a central transmembrane β -barrel. Most OMPs are single polypeptide chains with strand numbers varying from 8–36 (Lauber et al., 2018) while split oligomeric barrels can range from 12–60 strands. Remarkably, it appears that almost all OMPs come from just two common ancestors: single polypeptide OMPs arising from within-gene duplications of 2-stranded β -hairpin units and split oligomeric barrels arising from higher order oligomers of 4-stranded repeats (Remmert et al., 2010; Franklin et al., 2018b,a). Lysins, which are secreted extracellular toxins forming β -barrel pores in unrelated organisms in competitive environments, form a third convergent group of transmembrane β -barrel proteins, but these are not assembled into the host OM and can assemble spontaneously (Remmert et al., 2010; Franklin et al., 2018a). The OMP TolC is itself involved in the secretion of lysin monomers (α -hemolysin and cytolysin A) in *E. coli* and genes for cytolysin A are maintained in the laboratory K-12 strain (Kerényi et al., 2005; Thomas et al., 2014).

1.2.1 The OMP-ome of *E. coli*

The genome of *E. coli* encodes at least 33 families of proteins (comprising a total of 84 genes) which form transmembrane structures in the OM (see Table 1.1). Excluding those which form α -helical transmembrane-spanning regions in the OM (e.g. Wza and GfcE of the Outer Membrane Auxiliary (oma) Protein family (Dong et al., 2006), and TraF which forms the membrane spanning region of the type IV secretion system, Chandran et al. (2009), or those for which the structure is unknown (i.e. the membrane spanning region of the flagellum, FlgH, Fujii et al. (2017)), and those located on the F-plasmid which has been lost in many strains of K-12 (OmpP, YuaO, YuaQ, TraF) (Blattner et al., 1997) leaves 30 different protein families (77 genes) that comprise the OM. OMPdb, a database of β -barrel outer membrane proteins gathered by searching through the Uniprot database with hidden Markov model profiles of candidate families generated from manual curation of solved 3D structures, the PFAM database, and literature research, reports 105 known

families of OMPs in all organisms (Tsirigos et al., 2011). Remarkably, this suggests that within *E. coli*, almost a third of all known architectures of OMPs are represented. This data also shows that ~1.6 % of the *E. coli* K-12 genome encodes for transmembrane β -barrel OMP (the EcoCyc database records 4686 genes identified in strain MG1655 as of Sep. 2019), but when pseudogenes (see Table 1.2) are also excluded this falls to ~1.4 % (63 ‘true’ OMP genes, 4540 ‘true’ genes identified in strain MG1655 as of Sep. 2019). This is slightly lower than previous estimates of 2–3 % (Wimley, 2003) but it is possible that in ‘wild’ and pathogenic strains where there is a greater selective pressure to maintain and diversify OMPs involved in biofilm formation, nutrient acquisition, and secretion of extracellular factors (due to increased cellular competition) that this fraction would be greater. Nonetheless, these figures indicate that a significant amount of the genome is dedicated to these proteins.

Although absolute copy numbers are not known for all OMPs in *E. coli*, Li et al. (2014), used ribosome profiling to calculate the synthesis rates of all genes (number of translation initiation events per gene per unit time) in *E. coli* K-12 MG1655 and this could then be converted to ‘molecules per generation’ with knowledge of the doubling time of this organism. This may differ from the absolute copy number due to aggregation in the cytoplasm or periplasm before folding (and subsequent degradation), or due to inherited OMPs present at the poles in daughter cells (there appears to be no active mechanism to degrade folded OMPs apart from dilution into new cells upon division, and the formation of OMVs) (Rassam et al., 2015). Nonetheless, it provides a good genome-wide approximation of both the metabolic cost of OMP synthesis and the number of each substrate that requires folding into the OM. These data are summarised in Table 1.2.

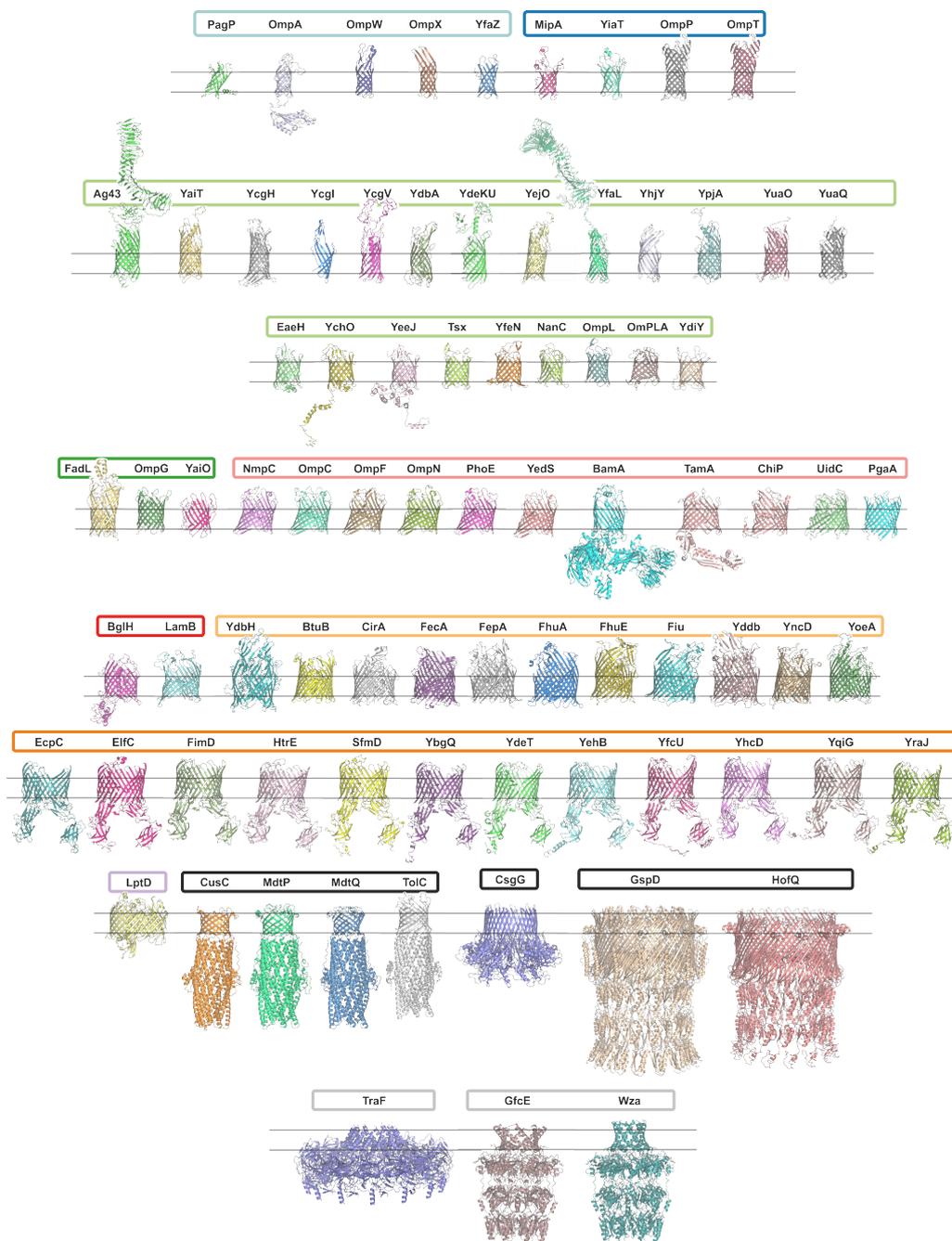


Fig. 1.6 Structures of all transmembrane proteins in the outer membrane of *E. coli* K-12 MG1655. A list of all transmembrane proteins in the outer membrane of *E. coli* K-12 strain MG1655 was manually curated, creating the “OMP-ome” (see Materials & Methods 2.7.2 for details on how this list was generated). The Protein Data Bank (<https://www.rcsb.org/>) was searched for solved structures of these proteins or close homologues. Where no solved 3D structures were available, homology models were generated using the I-TASSER server (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) (Yang and Zhang, 2015). For three proteins, BcsC, NfrA, and FlgH, no homology models could be generated. See Table 1.1 for details of which structures were used. OMPs are grouped here by strand number and then by protein family. The colour of the box surrounding the protein names represents the number of strands. Light blue = 8, dark blue = 10, light green = 12, dark green = 14, light red = 16, red = 18, light orange = 22, orange = 24, pink = 26, black = oligomeric split β -barrel, grey = α -helical transmembrane region.

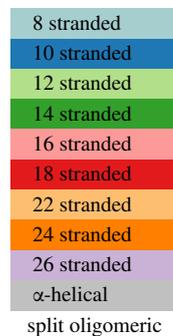
Table 1.1 Curated list of all OMPs encoded for by the *E. coli* K-12 MG1655 genome. Pseudogenes and fragments have been reconstructed to give an indication of the ancestral OMP-ome that would use the biogenesis machinery of this *E. coli* strain. For details of how this list was curated, see Materials & Methods 2.7.2. Protein name uses the gene name listed on Uniprot (<https://www.uniprot.org/>) (UniProt Consortium, 2019) or EcoCyc (<https://ecocyc.org/>) (Keseler et al., 2017), where they differ the name most commonly used in the literature is chosen. Strands shows the number of transmembrane β -strands the protein contains (or the assembled complex for proteins which contain a split oligomeric barrel). *, the number of strands has been inferred from the structure of homologs in the same family with solved structures; for OMPs with no solved 3D structures in its family the number of strands in a homology model built by the I-TASSER server (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) (Yang and Zhang, 2015) is displayed. **, proteins for which models could not be built in I-TASSER and for which there are no homologues with solved structures are given the number of strands predicted in literature searches for these proteins, based on homology models in the literature, or the results of β -barrel prediction algorithms. X, this protein is predicted to contain transmembrane α -helices rather than β -strands. ?, no structural data could be found either by homologous proteins, homology modelling by I-TASSER, or through literature searches, but localisation studies suggest it has transmembrane regions in the OM. For split oligomeric barrels they are described as [# of β -strands per subunit \times # of subunits]. Structure refers to where the structure in Figure 1.6 and the number of β -strands was derived from. For solved structures, this is given as the PDB code and highlighted in green. Structures with homology models generated using the I-TASSER server are described as 'ITASSER'. Those for which no models could be generated, and no structures of homologues exist, are described as 'None' and highlighted in blue. The Uniprot ID for each protein in *E. coli* K-12 is given in the column 'Uniprot'. YdeK and YdeU have two entries but are actually fragments immediately downstream of each other separated by a stop codon (which is not present in homologues from other *E. coli* strains). YcgH is a pseudogene with an in-frame stop codon splitting it into two fragments as well as a single nucleotide deletion causing a frameshift in fragment 2. It has no entry in Uniprot but has the accession ID of G6609 in the EcoCyc database. The homologue from *E. coli* strain O157:H7 was used in analyses (Uniprot: Q8X6C1).

Protein	Strands	Family	Structure	Uniprot
PagP	8	The Antimicrobial peptide resistance and lipid A acylation protein (PagP) Family	3GP6	P37001
OmpA	8	The OmpA Family	2GE4	P0A910
OmpW	8	The OmpW Family	2F1T	P0A915
OmpX	8	The Outer Membrane Protein X (OmpX) Family	1QJ8	P0A917
YfaZ	8*	The YfaZ Outer Membrane Protein Family	ITASSER	P76471
MipA	10*	The MipA/OmpV Family	ITASSER	P0A908
YiaT	10*	The MipA/OmpV Family	ITASSER	P37681
OmpP	10	The OmpT (OmpT) Family	ITASSER	P34210
OmpT	10	The OmpT (OmpT) Family	I178	P09169
Ag43	12*	The Autotransporter (AT) Family	ITASSER	P39180
YaiT	12*	The Autotransporter (AT) Family	ITASSER	P77199
YcgH	12*	The Autotransporter (AT) Family	ITASSER	-
YcgI	12*	The Autotransporter (AT) Family	ITASSER	P76000
YcgV	12*	The Autotransporter (AT) Family	ITASSER	P76017
YdbA	12*	The Autotransporter (AT) Family	ITASSER	P33666
YdeK / YdeU	12*	The Autotransporter (AT) Family	ITASSER	P32051 / P77286
YejO	12*	The Autotransporter (AT) Family	ITASSER	P33924
YfaL	12*	The Autotransporter (AT) Family	ITASSER	P45508
YhjY	12*	The Autotransporter (AT) Family	ITASSER	P37663
YpjA	12*	The Autotransporter (AT) Family	ITASSER	P52143
YuaO	12*	The Autotransporter (AT) Family	ITASSER	Q9JMS5
YuaQ	12*	The Autotransporter (AT) Family	ITASSER	Q9JMS3
EaeH	12	The Intimin/Invasin Family	4E1S	P36943
YchO	12*	The Intimin/Invasin Family	ITASSER	P39165

Protein	Strands	Family	Structure	Uniprot
YeeJ	12*	The Intimin/Invasin Family	ITASSER	P76347
NfrA	12**	The N4 bacteriophage Receptor (nfrA) Family	None	P31600
Tsx	12	The Nucleoside-specific Channel-forming Outer Membrane Porin (Tsx) Family	1TLY	P0A927
YfeN	12*	The Nucleoside-specific Channel-forming Outer Membrane Porin (Tsx) Family	ITASSER	P45564
NanC	12	The Oligogalacturonate-specific Porin (KdgM) Family	2WJQ	P69856
OmpL	12*	The Oligogalacturonate-specific Porin (KdgM) Family	ITASSER	P76773
OmPLA	12	The Outer Membrane Phospholipase (OMPLA) Family	1QD5	P0A921
YdiY	12*	The Salt-stress induced outer membrane protein (SspA) Family	ITASSER	P76206
BcsC	14**	The Bacterial Cellulose Synthase Operon Protein C (BcsC) Family	None	P37650
FadL	14	The FadL Outer Membrane Protein (FadL) Family	1T16	P10384
OmpG	14	The OmpG Porin (OmpG) Family	2IWW	P76045
YaiO	14*	The yaiO Outer Membrane Protein Family	ITASSER	Q47534
NmpC	16*	The General Bacterial Porin (GBP-1) Family 1	ITASSER	P21420
OmpC	16	The General Bacterial Porin (GBP-1) Family 1	2J1N	P06996
OmpF	16	The General Bacterial Porin (GBP-1) Family 1	2OMF	P02931
OmpN	16*	The General Bacterial Porin (GBP-1) Family 1	ITASSER	P77747
PhoE	16	The General Bacterial Porin (GBP-1) Family 1	1PHO	P02932
YedS	16*	The General Bacterial Porin (GBP-1) Family 1	ITASSER	P76335
BamA	16	The Outer Membrane Protein Insertion Porin (OmpIP/Omp85) Family	5LJO	P0A940
TamA	16	The Outer Membrane Protein Insertion Porin (OmpIP/Omp85) Family	4C00	P0ADE4
ChiP	16	The Outer Membrane Porin (OprD) Family	5MDQ	P75733
UidC	16*	The Outer Membrane Porin (OprD) Family	ITASSER	Q47706
PgaA	16	The Poly Acetyl Glucosamine Porin (PgaA) Family	4Y25	P69434
BglH	18*	The Sugar Porin (SP) Family	ITASSER	P26218
LamB	18	The Sugar Porin (SP) Family	1MPM	P02943
YdbH	22*	The Dicarboxylate Transport Family	ITASSER	P52645
BtuB	22	The Outer Membrane Receptor (OMR-TonB Dependent Receptor) Family	2GUF	P06129
CirA	22	The Outer Membrane Receptor (OMR-TonB Dependent Receptor) Family	5HDI	P17315
FecA	22	The Outer Membrane Receptor (OMR-TonB Dependent Receptor) Family	1KMO	P13036
FepA	22	The Outer Membrane Receptor (OMR-TonB Dependent Receptor) Family	1FEP	P05825
FhuA	22	The Outer Membrane Receptor (OMR-TonB Dependent Receptor) Family	2FCP	P06971
FhuE	22	The Outer Membrane Receptor (OMR-TonB Dependent Receptor) Family	6E4V	P16869
Fiu	22	The Outer Membrane Receptor (OMR-TonB Dependent Receptor) Family	6BPN	P75780
Yddb	22*	The Outer Membrane Receptor (OMR-TonB Dependent Receptor) Family	ITASSER	P31827
YncD	22*	The Outer Membrane Receptor (OMR-TonB Dependent Receptor) Family	ITASSER	P76115
YoeA	22*	The Outer Membrane Receptor (OMR-TonB Dependent Receptor) Family	ITASSER	P76356
EcpC	24*	The Outer Membrane Fimbrial Usher Porin (FUP) Family	ITASSER	P77802
ElfC	24*	The Outer Membrane Fimbrial Usher Porin (FUP) Family	ITASSER	P75857
FimD	24	The Outer Membrane Fimbrial Usher Porin (FUP) Family	3RFZ	P30130
HtrE	24*	The Outer Membrane Fimbrial Usher Porin (FUP) Family	ITASSER	P33129
SfmD	24*	The Outer Membrane Fimbrial Usher Porin (FUP) Family	ITASSER	P77468
YbgQ	24*	The Outer Membrane Fimbrial Usher Porin (FUP) Family	ITASSER	P75750
YdeT	24*	The Outer Membrane Fimbrial Usher Porin (FUP) Family	ITASSER	P76137
YehB	24*	The Outer Membrane Fimbrial Usher Porin (FUP) Family	ITASSER	P33341
YfcU	24*	The Outer Membrane Fimbrial Usher Porin (FUP) Family	ITASSER	P77196
YhcD	24*	The Outer Membrane Fimbrial Usher Porin (FUP) Family	ITASSER	P45420
YqiG	24*	The Outer Membrane Fimbrial Usher Porin (FUP) Family	ITASSER	P76655
YraJ	24*	The Outer Membrane Fimbrial Usher Porin (FUP) Family	ITASSER	P42915
LptD	26	The Imp/OstA Family	4RHB	P31554
CusC	12 (4x3)	The Outer Membrane Factor (OMF) Family	4K7R	P77211
MdtP	12 (4x3)*	The Outer Membrane Factor (OMF) Family	ITASSER	P32714
MdtQ	12 (4x3)*	The Outer Membrane Factor (OMF) Family	ITASSER	P33369

Protein	Strands	Family	Structure	Uniprot
TolC	12 (4x3)	The Outer Membrane Factor (OMF) Family	1EK9	P02930
CsgG	36 (4x9)	The Curli Fiber Subunit Porin, CgsA, CsgG (CsgG) Family	4UV3	P0AEA2
GspD	60 (4x15)	The Secretin Family	5WQ7	P45758
HofQ	60 (4x15)*	The Secretin Family	ITASSER	P34749
TraF	X	The F plasmid transfer operon (TraF) Family	3JQO	P14497
GfcE	X*	The Outer Membrane Auxiliary (oma) Protein Family	ITASSER	P0A932
Wza	X	The Outer Membrane Auxiliary (oma) Protein Family	2J58	P0A930
FlgH	? (?x26)	Flagellar L-ring Protein (FlgH) Family	None	P0A6S0

Table 1.2 Abundance of OMPs in *E. coli* K-12 strain MG1655 as measured by absolute synthesis rates. The copy number data in this table was taken from ribosome profiling experiments conducted by Li et al. (2014). In these experiments, cells were grown in three different media but here are shown the data for cells grown in a MOPS-buffered rich defined medium. The gene name in Uniprot is used (except for *pldA* which is commonly referred to by its protein name, OmpLA). Note that for split oligomeric OMPs, each 'copy' is a single subunit. Gene names are coloured according to the number of β -strands in their barrel domain (following the colour scheme of Figure 1.6, except for oligomeric split barrels which are not coloured in this table). The table is split at each order of magnitude to ease comparisons. See Table 1.1 for details of the nomenclature used for β -strands. Note that expression of genes will be highly dependent on the sources and abundances of nutrients, as well as other growth factors, so these data serve only as a general guide of expression under laboratory conditions. pseudo = pseudogene according to EcoCyc (<https://ecocyc.org/>) (Keseler et al., 2017) (not expressed under any conditions). F plasmid = gene is encoded on the F mating plasmid which was not assessed in Li et al. (2014).



Gene	Copy Number	β -strands
<i>ompA</i>	207618	8
<i>ompC</i>	163538	16
<i>ompX</i>	125295	8
<i>ompF</i>	88988	16
<i>ompT</i>	40237	10
<i>mipA</i>	20925	10
<i>tsx</i>	14911	12
<i>tolC</i>	8768	12 (4x3)
<i>fadL</i>	6912	14
<i>ydiY</i>	5888	12
<i>bamA</i>	3904	16
<i>yfaZ</i>	3657	8
<i>btuB</i>	2295	22
<i>lptD</i>	2267	26
<i>fhuA</i>	1208	22

Gene	Copy Number	β -strands
<i>ompla</i>	1175	12
<i>yncD</i>	353	22
<i>ompW</i>	331	8
<i>tamA</i>	195	16
<i>fiu</i>	136	22
<i>ag43</i>	122	12
<i>lamB</i>	100	18
<i>ydbH</i>	90	22
<i>cirA</i>	79	22
<i>gfcE</i>	74	X
<i>fepA</i>	69	22
<i>phoE</i>	62	16
<i>fecA</i>	51	22
<i>nfrA</i>	50	12
<i>bcsC</i>	39	14
<i>yiaT</i>	32	10
<i>ychO</i>	26	12
<i>yhjY</i>	25	12
<i>fhuE</i>	20	22
<i>yddB</i>	19	22
<i>fimD</i>	17	24
<i>pagP</i>	16	8
<i>flgH</i>	13	X
<i>csgG</i>	11	36 (4x9)
<i>ompN</i>	4	16
<i>chiP</i>	3	16
<i>gspD</i>	3	60 (4x15)
<i>yaiO</i>	3	14
<i>yeeJ</i>	3	12
<i>yfaL</i>	3	12
<i>bglH</i>	2	18
<i>hofQ</i>	2	60 (4x15)
<i>ycgV</i>	2	12
<i>yfeN</i>	2	12
<i>yhcD</i>	2	24
<i>ypjA</i>	2	12
<i>ompL</i>	1	12
<i>pgaA</i>	1	16
<i>cusC</i>	0	12 (4x3)
<i>ecpC</i>	0	24
<i>elfC</i>	0	24
<i>htrE</i>	0	24
<i>mdtP</i>	0	12 (4x3)
<i>nanC</i>	0	12
<i>ompG</i>	0	14
<i>sfmD</i>	0	24
<i>uidC</i>	0	16
<i>wza</i>	0	X
<i>ybgQ</i>	0	24
<i>yehB</i>	0	24
<i>yraJ</i>	0	24
<i>eaeH</i>	pseudo	12
<i>mdtQ</i>	pseudo	12 (4x3)

Gene	Copy Number	β -strands
<i>nmpC</i>	pseudo	16
<i>yaiT</i>	pseudo	12
<i>ycgH</i>	pseudo	12
<i>ycgI</i>	pseudo	12
<i>ydbA</i>	pseudo	12
<i>ydeKU</i>	pseudo	12
<i>ydeT</i>	pseudo	24
<i>yedS</i>	pseudo	16
<i>yejO</i>	pseudo	12
<i>yfcU</i>	pseudo	24
<i>yoeA</i>	pseudo	22
<i>yqiG</i>	pseudo	24
<i>ompP</i>	F plasmid	10
<i>traF</i>	F plasmid	X
<i>yuaO</i>	F plasmid	12
<i>yuaQ</i>	F plasmid	12

1.2.2 The scope of BAM-dependent substrates

The diversity of OMPs present, even in the relatively reduced genome of the laboratory strain, *E. coli* K-12, raises the question: how many of these OMPs require BAM for their assembly into the outer membrane? *In vitro* studies suggest that BAM is required for the assembly of many different OMPs into *E. coli* polar lipids (including even the smallest OMPs with a single polypeptide chain, no additional subunits, no large folded domains - e.g. OmpA, OmpX, OmpT, OmpLA, OmpG) (Burgess et al., 2008; Hagan et al., 2010, 2013; Patel and Kleinschmidt, 2013; Gessmann et al., 2014; Iadanza et al., 2016; Hussain and Bernstein, 2018). Other studies have shown that BAM plays an essential role in the biogenesis of probably all classes of autotransporter (confirmed for Type Va, Vb, Vc, and Ve autotransporters, type Vd not confirmed but this group appears to have arisen from a fusion between ancestral Va and Vb, Leo et al. (2012)), so likely to also require BAM) (Ieva and Bernstein, 2009; Sauri et al., 2009; Bodelón et al., 2009; Norell et al., 2014; Roman-Hernandez et al., 2014). Members of the Outer Membrane Porin (OprD) family in *Pseudomonas aeruginosa* were shown to be significantly reduced in copy number at the OM upon depletion of SurA, which is the major chaperone responsible for delivery of OMPs to BAM, suggesting indirectly that BAM catalyses the folding of these proteins (Klein et al., 2019). Levels of LamB of the Sugar Porin family were reduced in Δ bamB mutants or under conditions of BamA depletion in *E. coli* (Ruiz et al., 2005; Wu et al., 2005). TonB-dependent receptors (TBDRs) were shown to cluster with the BAM complex during their biogenesis (Rassam et al., 2015) and deletion of BamE severely reduces levels of at least three TBDRs in *Caulobacter crescentus* (Ryan et al., 2010). TolC

levels were shown to be significantly reduced in coordination with depletion of BamA, suggesting that split oligomeric barrels of the Outer Membrane Factor family also require BAM (Werner and Misra, 2005). Levels of the Fimbrial Usher Porin family protein, FimD, was shown to be reduced under conditions of BamA depletion or BamB deletion (BamC or BamE deletion individually only marginally affected the levels of FimD, but their double deletion showed a similar phenotype to $\Delta bamB$) (Palomino et al., 2011). The essential outer membrane protein responsible for inserting LPS into the OM, LptD, requires BAM for its assembly (Ruiz et al., 2005; Chimalakonda et al., 2011; Lee et al., 2016, 2018). In *Neisseria meningitidis*, the assembly of the split oligomeric barrel protein, PilQ, was shown to be dependent on the BAM complex with depletion of BamA, BamD (which may or may not be essential in *N. meningitidis* - Fussenegger et al. (1996); Volokhina et al. (2009)) or BamE, but not BamC, reducing levels of the assembled complex (Voulhoux et al., 2003; Volokhina et al., 2009). This protein was also shown to be weakly dependent on BAM in *Pseudomonas aeruginosa*, but more significantly dependent on LolB, suggesting a gradient of dependence for this substrate (Hoang et al., 2011). PilQ is a lipoprotein and part of the secretin superfamily (sometimes referred to as the GspD-PilQ family) which forms large oligomeric β -barrel pores with numbers of subunits that vary between species and machinery (e.g. type II secretion system [T2SS], type III secretion system [T3SS], type IV pili secretion), but which likely all donate four β -strands per monomer (Burkhardt et al., 2011; D'Imprima et al., 2017). Another secretin in *P. aeruginosa*, PscC, forms part of the T3SS and was also shown to weakly depend on BAM, but more significantly to depend on LolB (Hoang et al., 2011). The biogenesis of the T3SS and flagella apparatus were also shown to be affected by deletion of BamB or BamD in *Salmonella enterica*, although it was not clear whether this was due to a transcriptional feedback loop reducing expression levels or a direct role of BAM leading to misassembly (Fardini et al., 2009). Deletion of BamE in *Caulobacter crescentus* led to reduced levels of assembled CpaC, a secretin for type IV pili. Even the BAM complex itself, via the lipoprotein subunits (BamB and BamD), is thought to be at least partially responsible for the assembly of BamA in *E. coli* (perhaps suggesting that the Lol machinery is ancestral to OMPs) (Hagan et al., 2013).

Interestingly, a number of proteins embedded in the outer membrane have been shown capable of assembling in a BAM-independent manner. PulD, a secretin homologous to GspD, forms part of the T2SS in *Klebsiella oxytoca* and does not require BAM for assembly (Collin et al., 2007; Huysmans et al., 2015). XcpQ, which forms the secretin for the T2SS in *P. aeruginosa*, does not require BAM or the Lol machinery for assembly (Hoang et al., 2011). Levels of the GspD secretin in *E. coli* are not affected by BamA depletion (Dunstan et al., 2015). The filamentous phage secretin pIV can assemble without accessory proteins *in vivo* and spontaneously *in vitro* (Kazmierczak et al., 1994; Nickerson

et al., 2012). The split oligomeric barrel of the curli fibre biogenesis system (sometimes called the Type VIII secretion system, T8SS), CsgG, can be assembled independently of BamA in *E. coli* (Dunstan et al., 2015). Finally, and perhaps not surprisingly, α -helical transmembrane proteins in the OM do not require BAM for their assembly (Dunstan et al., 2015).

It is tempting to speculate that the distribution of different OMP families in the OMP-ome of *E. coli* might have driven the evolution of the BAM complex, particularly whether more complex protein families which are well represented (see Table 1.1) might make special use of additional subunits of BAM, BamBCDE (e.g. the 12-stranded Autotransporter Family which have to secrete huge extracellular domains on the order of hundreds of kilodaltons to megadaltons, the 22-stranded TBDR Family which have to fold and insert a plug domain and engage with energised machinery in the inner membrane, and the 24-stranded Fimbrial Usher Porin Family which have multiple folded domains at their N- and C-termini as well as inserted into periplasmic loops). The BAM complex, therefore, has to fold a diverse palette of OMP families, some of which may have unique or special requirements to be correctly folded and inserted into the OM. The conservation of BamA throughout bacteria suggests that a basic common mechanism for the activity of BAM exists for every OMP substrate but, as mentioned above, some of these families seem to rely on some subunits more than others. Other species may also tailor their biogenesis machinery to cater for their particular OMP-ome. *Caulobacter crescentus* is a bacterium found in soil, freshwater, and saltwater, where it inhabits environmental niches where the nutrient concentration is very low. From analysis of its genome, it has a very different OMP-ome bias to *E. coli*, with 67 TBDRs (for active nutrient import), and no general porins (which are used for passive diffusion of small molecules and nutrients) (Nierman et al., 2001; Neugebauer et al., 2005). As mentioned in Section 1.1.4, the BAM complex in this species lacks BamC, but contains the additional subunits Pal and BamF. Furthermore, unlike the mild growth defect of $\Delta bamE$ in *E. coli* and *S. enterica*, or the slightly greater defect in *N. meningitidis*, $\Delta bamE$ strains of *C. crescentus* show much more severe growth defects (Sklar et al., 2007a; Lewis et al., 2008; Volokhina et al., 2009; Ryan et al., 2010; Sikora et al., 2018), suggestive of a more significant role of BamE in OM biogenesis in these bacteria.

1.3 Outer membrane proteins *en route* to the outer membrane

1.3.1 From translation to the inner membrane

As a nascent OMP emerges from the ribosome it is bound preferentially by an ATP-independent cytoplasmic chaperone, trigger factor (TF) (Oh et al., 2011) (see Figure 1.7 for an overview of this pathway). This preference for TF over the signal recognition particle (SRP) is thought to be driven by the lower hydrophobicity of N-terminal signal sequences of secreted proteins and some recognition of the physicochemical properties of the polypeptide chain such as the hydrophobicity of β -strands in OMPs (Eisner et al., 2006). The next step, either co- or post-translationally, involves delivery of the OMP to the SecYEG translocon (Figure 1.7) for export into the periplasm. It has been shown *in vitro* that the post-translational SecA/SecB pathway is predominant for OMP secretion over a parallel pathway utilizing the signal recognition peptide (SRP) that is mainly used by inner-membrane proteins (Behrmann et al., 1998; Koch et al., 1999; Bornemann et al., 2014). These pathways are thought to diverge as: i) OMPs may have hydrophobic surfaces which require specific kinds of chaperones to complement their folded state and/or prevent misfolding; ii) the SecYEG translocon needs to distinguish between proteins for secretion, and those to be integrated into the inner membrane (IM). *In vitro* studies with native *E. coli* IM lipids showed that the OMP PhoE (Figure 1.6) interacts with another ATP-independent chaperone in the cytoplasm, SecB. SecB may act as an aid for targeting proteins to the membrane-bound receptor SecA for delivery to SecYEG (Kusters et al., 1989; de Cock and Tommassen, 1992). SecB docks with the dimeric SecA motor domain and the polypeptide is fed into the SecY channel. The mechanism by which the Sec translocon moves OMPs through this channel is disputed but there are two main competing models. In the first, secreted polypeptides move in a ratchet-like motion through the SecY channel powered by hydrolysis of ATP in SecA (Lycklama A Nijeholt and Driessen, 2012). In a second model, Sec works by a ratchet mechanism but is primarily powered by biased Brownian motion. The translocating polypeptide can move freely forward and backward through the SecY channel but will stop when the chain reaches a 'block' in the channel (due to bulky residues or partially formed secondary structure in the substrate), this is sensed by the channel and causes nucleotide exchange (ADP out, ATP in) which leads to an opening in the channel permitting free diffusion of this bulky region of the polypeptide. ATP hydrolysis at SecA subsequently recloses the channel and traps the bulky region on the periplasmic or cytosolic side, binding of chaperones and other secretion or chaperoning factors in the

periplasm will further bias this Brownian motion in a periplasmic direction (Allen et al., 2016; Fessl et al., 2018). Whichever model is correct, the result is the emergence of the OMP into the periplasm.

1.3.2 Emergence into the periplasm: more ATP-independent chaperones come into play

During translocation or at the earlier stages after this is completed, OMPs may be held near the inner membrane on the periplasmic side awaiting their recognition and recruitment by periplasmic folding factors. PpiD is a periplasmic (inactive) PPIase embedded in the inner membrane with a single-span α -helix and has been shown to interact directly with the inner membrane translocon SecYEG (Sachelaru et al., 2014) as well as crosslinking to a translocation intermediate of OmpA and mediating its release into the periplasm (Antonoaea et al., 2008; Fürst et al., 2018). It was initially discovered as a chaperone important for OMP biogenesis where it was shown to be a multicopy suppressor of a deletion mutant of a key periplasmic OMP chaperone (Δ surA), with a Δ ppiD mutant leading to reduced OMP levels and a Δ ppiD Δ surA mutant being lethal (Dartigalongue and Raina, 1998). In contrast to that study, others have found PpiD cannot compensate for a lack of SurA and so is unlikely to play a *direct* role in OMP biogenesis, but that its overexpression can rescue the lethality of a double mutant of periplasmic OMP chaperones (Δ skp Δ surA), in a manner that requires it to be anchored to the inner membrane (Matern et al., 2010). Deletion of PpiD also does not affect the rate of folding of OmpA *in vivo* into the OM as measured by radiography in a pulse-chase experiment, and was suggested to act by modulating the dwell time of nascent secreted proteins as they emerge from SecYEG (Fürst et al., 2018). A role for both SurA and PpiD in OMP biogenesis was also found in *Campylobacter jejuni* (which contains at least two proteins containing homology to SurA) (Taylor et al., 2017). As the nascent OMP emerges from the SecYEG machinery into the periplasm its N-terminal signal peptide is cleaved by signal peptidase (Paetzel, 2014) whereupon the OMP is bound by periplasmic chaperones (although it is unclear which event occurs first or if they occur simultaneously, and whether cleavage of the signal peptide occurs before or after recognition by PpiD). In *E. coli* there are five known periplasmic folding factors with direct chaperoning activity for OMPs: SurA, Skp, DegP, FkpA and Spy (Figure 1.8) (Mogensen and Otzen, 2005; Quan et al., 2011). Skp was identified initially in a search for chaperones which specifically interact with periplasmic proteins and OMPs (Chen and Henning, 1996; Missiakas et al., 1996) and Skp may be involved primarily in the earliest stages of OMP secretion into the periplasm. *In vivo* cross-linking studies with spheroplasts showed that the OMP PhoE interacts with Skp while PhoE was still in transit across

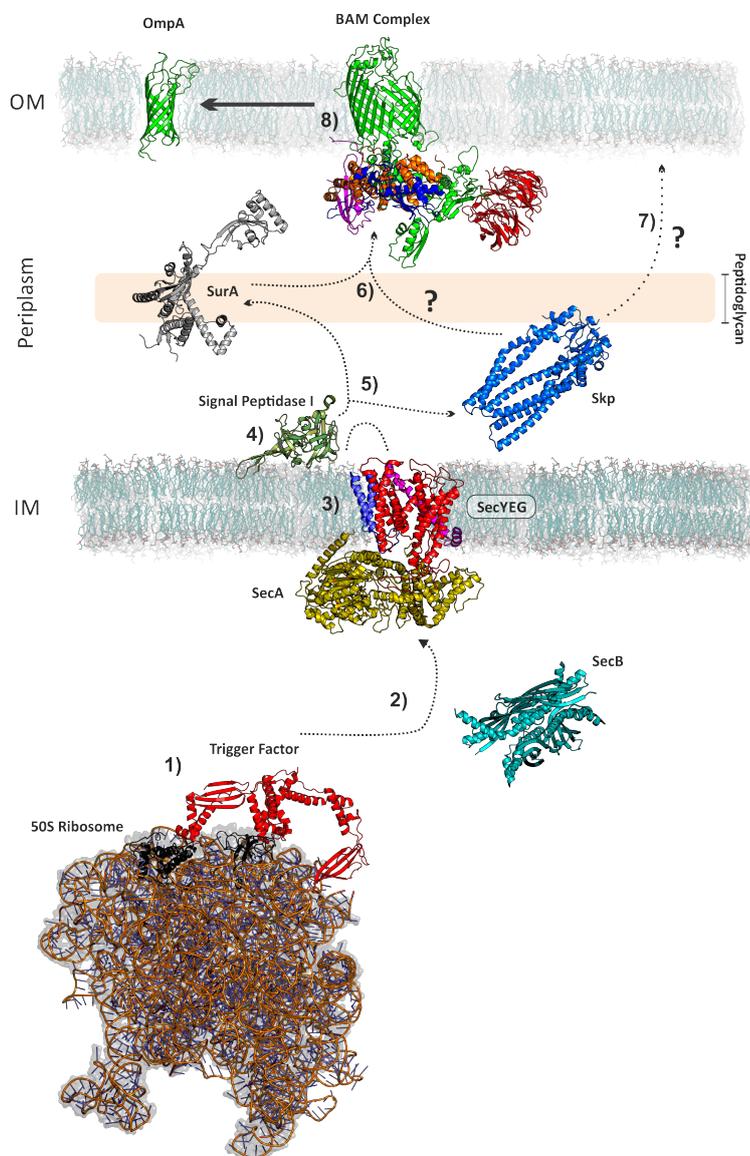


Fig. 1.7 A canonical folding pathway of an OMP including key members of the pathway. A nascent OMP emerges from the ribosome and is bound by TF (1) before being passed to SecA directly or via SecB (2). The unfolded OMP (uOMP) passes through the SecY channel (3) and the uOMP's N-terminal signal sequence is inserted into the inner membrane (IM) (4). This sequence is cleaved by SP1, signal peptidase I, and the uOMP is bound by the chaperones Skp and/or SurA (5). The uOMP can then be delivered to the BAM complex (6) or directly to the outer membrane (OM) (7). The BAM complex then catalyses the OMP's folding into the OM (8). SecYEG complex components: SecY – red, SecE – blue, SecG – magenta, SecA – yellow. All proteins are shown to scale. The length of the periplasmic space from leaflet to leaflet has been reported to be between 150 to >500 Å but around the mid-cell is likely to be on the lower end of these estimates. The distance between headgroups in the IM and OM is scaled to 18 nm in this figure. The thickness of the peptidoglycan (PG) has been scaled to 3 nm to match the thickness of a belt of density found in cryo-electron tomography reconstructions of the T3SS of *S. enterica* (a closely related species to *E. coli*) (Hu et al., 2017). The distance from the OM to the peptidoglycan layer is also scaled to be 7 nm as found by molecular dynamics simulations of PG binding by OmpA and Braun's lipoprotein (the most abundant PG binding proteins in *E. coli*) in native membranes (Samsudin et al., 2017). PDB ID of structures: OmpA (1G90) (Arora et al., 2001); BamACDE (5EKQ) (Bakelar et al., 2016); BamB (4XGA) (Chen et al., 2016); SurA (1M5Y, missing regions rebuilt using MODELLER) (Bitto and McKay, 2002); Skp (1U2M, missing regions rebuilt using MODELLER) (Walton and Sousa, 2004); SP1, signal peptidase I (1KN9) (Paetzl et al., 2002); SecYEG+SecA (3DIN) (Zimmer et al., 2008); SecB (1OZB) (Zhou and Xu, 2003); TF, Trigger Factor (3GU0) (Martinez-Hackert and Hendrickson, 2009); 50S ribosome (2D3O) (Schlünzen et al., 2005).

the IM (Harms et al., 2001). Similar *in vitro* studies with inside-out plasma membrane vesicles showed the same interactions with OmpA (Figure 1.6) (Schäfer et al., 1999). Sequential deletion of N-terminal residues from PhoE reduced Skp binding efficiency in co-immunoprecipitation experiments. By varying the location and length of N-terminal deletions the authors proposed two independent binding sites for Skp, one located in the N-terminal 100 residues and the second within the next 100 residues (Harms et al., 2001). Skp has also been co-purified with inner-membrane fractions of *E. coli* and shows partial resistance to proteolysis, possibly indicating shielding of cleavage sites through association with the IM (Schäfer et al., 1999). *In vitro* and *in silico* studies have shown that a selection of OMPs (OmpC, transmembrane OmpA, and OmpF) bind to Skp N-terminus first and enter into the central cavity of the chaperone through the bottom ‘legs’ of the Skp trimer (see Figure 1.8), possibly through a ‘climbing’ mechanism involving the making and breaking of sequential salt-bridges between the unfolded OMP and Skp (Lyu et al., 2012). Recent small-angle neutron scattering (SANS) studies with OmpA and OmpW (Figure 1.6) have shown that Skp binds in a clamp-like manner with the unfolded OMPs entering the central cavity within Skp’s α -helical tentacles before these ‘legs’ close in around the OMP (Zaccai et al., 2016). The crystal structure of Skp (see Figure 1.8) showed a cavity whose size was incompatible with the full sequestration of OMPs larger than ~8-strands. *In vitro* studies using electrospray ionisation ion-mobility mass spectrometry (ESI-IM-MS), molecular dynamics (MD), and folding kinetics on 8-, 10-, and 16-stranded OMPs indicated that the cavity of Skp expands to accommodate larger substrates and when that is not sufficient, or in parallel with this, multiple copies of the Skp trimer can bind to a single OMP substrate (Schiffrin et al., 2016). This cavity expansion mechanism was investigated further through microsecond timescale MD, NMR, and small angle X-ray scattering (SAXS) to show that the tentacle like ‘legs’ of Skp can flex backwards by a conserved ‘hinge’ at their midpoint to allow larger substrates to bind (Holdbrook et al., 2017). This increases the largest substrate size (imagined as an idealised sphere) from a radius of ~1.5 nm in the smallest structures observed in MD to ~3 nm in the largest structures observed (an ~8X increase in volume). Furthermore, these states exist as part of a conformational ensemble that can interconvert on a sub-millisecond timescale. *In vitro* single-molecule fluorescence experiments took aggregated OmpC and added Skp or SurA to monitor their effects and found that Skp was able to rescue the aggregated state of OmpC while SurA could not suggesting that although the roles of Skp and SurA were shown to be overlapping genetically (as each can be deleted individually but not together), they each have unique functions (Li et al., 2018). Inside the cavity, NMR studies on tOmpA and OmpX in complex with Skp showed that OMPs tumble dynamically in an unfolded conformational ensemble that interconverts on a sub-millisecond timescale (Burmamann et al., 2013). This OMP-Skp complex has a lifetime of hours in the absence of any

other factors (Burmam et al., 2013), with the substrate showing non-specific interactions throughout the cavity region but excluded from the trimersation crown (shown in pink in Figure 1.8) and the tips of the legs (Qu et al., 2009; Walton et al., 2009; Burmam et al., 2013; Callon et al., 2014).

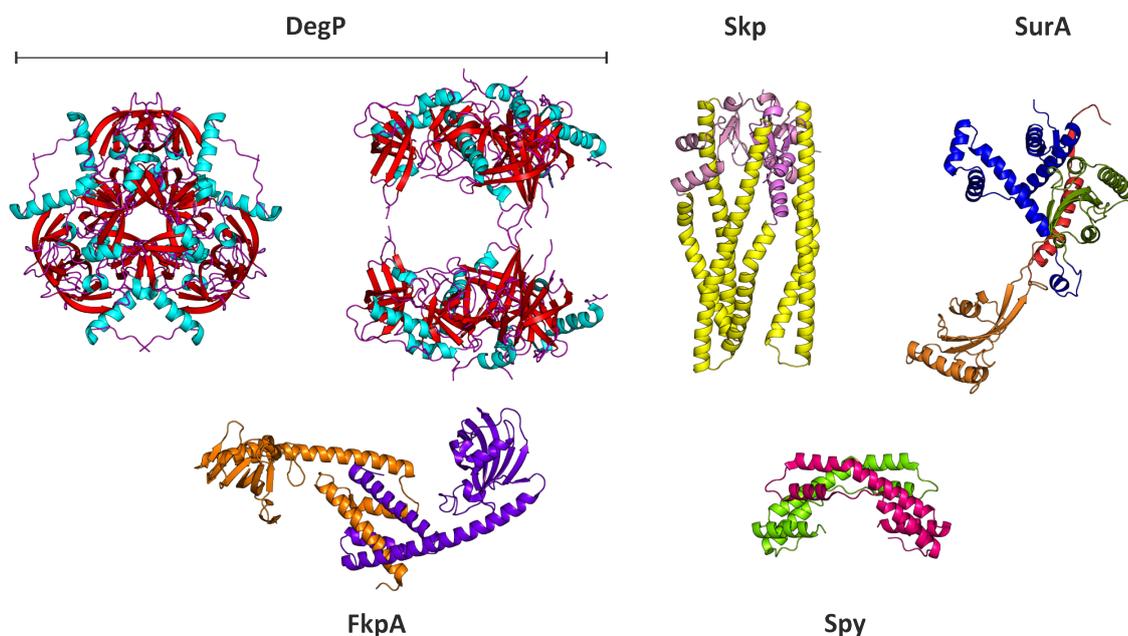


Fig. 1.8 Structures of periplasmic proteins with known chaperoning activity. DegP is shown from two angles: left – top view, right – side view. DegP has been shown to adopt a number of oligomeric states including a 6-mer resting state and 12- and 24-mer activated states (Ortega et al., 2009). Shown here is the 6-mer coloured according to secondary structure (helix – cyan, β -sheet – red, loop – magenta). Skp monomers assemble into a trimer as shown here. The ‘legs’ of Skp are shown in yellow with the trimerization ‘head’ shown in pink. SurA is coloured according to its folding domains (N-domain – blue, peptidyl-prolyl isomerase [PPI] domain 1 (P1) – green, PPI2 domain (P2) – orange, C-domain – red). FkpA and Spy form functional dimers and each monomeric unit is shaded differently for clarity. Shown to scale. PDB ID of structures: DegP (3MH5) (Krojer et al., 2010); Skp (1U2M, missing regions rebuilt using MODELLER) (Walton and Sousa, 2004); SurA (1M5Y, missing regions rebuilt using MODELLER) (Bitto and McKay, 2002); FkpA (1Q6H) (Saul et al., 2004); Spy (3O39) (Quan et al., 2011).

Some authors have suggested that SurA is also capable of interacting with OMPs as they emerge from the IM. *In vivo* pulse-chase kinetic analysis of LamB (Figure 1.6) maturation from translation to OM insertion found the rate of conversion from its post-SecYEG secreted, unfolded form to monomeric folded form was reduced from 0.007 s^{-1} in WT strains to 0.00033 s^{-1} in $\Delta surA$ strains (Ureta et al., 2007). Cell viability and OM density is also maintained in $\Delta skp \Delta degP$ double mutants suggesting that SurA is capable of carrying out the role of Skp in its absence (Sklar et al., 2007b). Nonetheless, evidence for SurA interaction with OMPs at the IM remains indirect and sparse. Despite Skp showing interactions with translocating OMPs it may be that Skp is not necessarily required for the release of unfolded OMPs from the translocon / IM if they are sufficiently soluble or ‘targeted’ by another chaperone. This effect was shown by the dependence of Skp for release of OmpA, but not PhoE, from the IM (Harms et al., 2001). SurA may be the first chaperone to ‘capture’ the unfolded OMPs as they are released into the periplasm.

Furthermore, SurA was shown to fractionate with OM fractions and soluble proteins, but not IM fractions, in gradient centrifugation experiments (Hennecke et al., 2005).

Skp may undergo conformational changes to detach from the exit-channel of SecYEG at the inner-membrane in order to switch its function from a binding/receiving mode, to a ‘holdase shuttle’ to transport the unfolded OMPs across the periplasmic space. *In vivo* proteolysis studies have shown a small population of Skp to be trypsin resistant (~20 %) when whole cells were partially permeabilised and incubated with protease (De Cock et al., 1999). These results were recapitulated *in vitro* where the degree of protease resistance of Skp was found to be modulated by the presence of phospholipids, Mg^{2+} , and LPS (De Cock et al., 1999). Skp has since been shown to exist as a trimer (Schlapschy et al., 2004; Walton and Sousa, 2004) and it remains unclear whether these results represent alternative conformations of the trimer, or an artefact relating to Skp assembly / disassembly from or to its monomeric state (Sandlin et al., 2015). Fluorescence studies of a Skp-OmpA complex with site-specific labels showed that in the presence of the negatively charged LPS, the C-terminal loop regions of OmpA are released from their Skp-bound state (Qu et al., 2009). Other studies showed that Skp is able to release bound OmpA and allow folding in the presence of negatively-charged membranes (containing a fraction of dioleoylphosphatidylglycerol [DOPG, *diC*_{18:1}PG]) or LPS, but to retard folding in the presence of neutral charge / zwitterionic dioleoylphosphatidylcholine (DOPC, *diC*_{18:1}PC) or DOPC/dioleoylphosphatidylethanolamine (DOPE, *diC*_{18:1}PE) membranes (Patel et al., 2009). However, subsequent bioinformatic and NMR studies on Skp-LPS interactions found that the previously proposed LPS binding site on Skp (Walton and Sousa, 2004) was in fact not conserved when a larger multiple sequence alignment was performed, furthermore, LPS was found to bind non-specifically to the trimerization ‘crown’ of Skp causing the trimer to dissociate and Skp to partially denature (Burmam et al., 2015). LPS is likely to be present at only very low free concentrations in the periplasm as it is shuttled to the OM via a specific bridge (Okuda et al., 2016) suggesting that an LPS-based release mechanism is unlikely, but negative charges are present on the inner leaflet of the outer membrane (and the outer leaflet of the inner membrane) due to the presence of PG headgroups (at around 15—20 % of the total headgroup content) (Morein et al., 1996). Taken together, these results suggest that Skp is capable of reacting to its environment to facilitate ‘correct’ OMP folding (McMorran et al., 2013).

As well as a holdase role for nascently secreted OMPs, Skp plays a more general chaperoning role in the periplasm. For example, Skp has been shown to crosslink to the autotransporter EspP (Figure 1.6) while EspP was undergoing a transport reaction across the OM and EspP itself was cross-linked to BamA (Ieva and Bernstein, 2009). This

suggests that Skp protects unfolded OMP chains during transit across the periplasm, as well as delivering them to the OM or to the vicinity of assembly machinery such as the BAM complex. Another study investigating the interactions of periplasmic chaperones with the BAM complex *in vivo* found a direct interaction of BAM with SurA, but not Skp or DegP (Sklar et al., 2007b). Nonetheless, it is interesting to note that most OMPs cluster together in discrete BAM-complex enriched 'islands' in the outer-membrane of *E. coli* (Rassam et al., 2015), and that one of the components of the BAM complex, BamE, is reported to bind and recruit phosphatidylglycerol (PG) lipids (Endo et al., 2011; Knowles et al., 2011) which, as mentioned above, was shown to promote OMP release from Skp in synthetic lipid vesicles (Patel et al., 2009). Further studies on EspP suggested a mechanism where Skp binds first and then passes substrates to SurA and components of the BAM complex, thereby acting as a periplasmic 'shuttle' (Ieva et al., 2011; Pavlova et al., 2013). However, SurA has a dissociation constant, K_d , for OMP-derived peptides between 1.6–10.9 μM (Bitto and McKay, 2003; Hennecke et al., 2005), while Skp binds OMPs with a K_d between 0.3–83 nM (Qu et al., 2007; Patel et al., 2009) making this interaction unlikely in the absence of additional modifying factors. *In vivo* mutagenesis studies found that folding of the transmembrane component of the LPS assembly machinery, LptD (see Figure 1.6), and not other OMPs studied, was impaired in the absence of either Skp or SurA, and that overexpression of either SurA or Skp alone, could not abrogate these effects (Schwalm et al., 2013). This suggests that, at least for some OMPs, there is a specific need for multiple chaperones to act in concert for the proper assembly of a single OMP.

SurA, an ATP-independent chaperone, is thought to recognise and bind to carboxy-terminal sequences of its OMP substrates via their aromatic-random-aromatic (Ar-X-Ar) sequence which is enriched within OMPs (28 % of OMPs versus <1 % for proteins localized in any other compartment, Bitto and McKay (2003)). More generally, SurA has a preference for substrates with aromatic-rich sequences with specific side-chain orientations involving at least two aromatic residues on the same 'face' of the β -strand (Bitto and McKay, 2003; Hennecke et al., 2005; Xu et al., 2007). SurA has been proposed as the major periplasmic chaperone for OMPs on the basis of general OM and OMP folding defects in ΔsurA strains, as well as decreased OM density and reduced levels of LamB and OmpA in ΔsurA , but not Δskp ΔdegP , *E. coli* strains (Lazar and Kolter, 1996; Rouvière and Gross, 1996; Sklar et al., 2007b). However, these results may be confounded by the selectivity of SurA for the chosen OMPs and the relative abundance of SurA's targets compared with the total *E. coli* OMP proteome. Differential proteomics studies comparing the levels of 41 OM lipoproteins and 23 β -barrel OMPs in wild-type versus ΔsurA strains showed the abundance of only 8 β -barrel proteins to be affected 2-fold or more by the deletion of SurA (Table 1.3) (Vertommen et al., 2009). However, these included OmpA

and OmpF which are present at copy numbers of ca. 10^5 and 10^4 copies per cell and which, along with OmpC (ca. 2×10^4 per cell), make up almost the entire mass of integral OMPs in *E. coli* (Henning et al., 1973; Rosenbusch, 1974; Lugtenberg and Van Alphen, 1983; Li et al., 2014). In comparison to SurA, proteomics studies involving a *Strep*-tagged Skp fusion protein identified direct interactions of Skp with at least 19 β -barrel OMPs (Jarchow et al., 2008). Furthermore, one of the few OMPs which essentially require SurA for proper assembly is LptD, whose function in LPS assembly is related to cell-wall integrity and whose misfolding may disproportionately destabilise the OM (Schwalm et al., 2013). Conversely, differential proteomics of *skp* null mutants showed none of the 23 reproducibly detectable β -barrel proteins was decreased in abundance more than 2-fold, implying that no OMPs *absolutely* require Skp for their biogenesis. Another analysis found the levels of 14 of the 20 reproducibly detectable OMPs to fall more than 2-fold for Δ *skp*, SurA depletion mutants (Denoncin et al., 2012). These results suggest that in regards to chaperoning of the most abundant OMPs, other periplasmic chaperones can rescue Skp function almost completely, partial redundancy of chaperoning exists for OMPs on the SurA pathway, and little redundancy exists for chaperoning OMPs outside of Skp and SurA pathways.

1.4 Recognition of OMPs and their delivery to the BAM complex

1.4.1 SurA and a possible role in bridging the IM and OM

Regardless of the relative importance of SurA and Skp at the IM and in the periplasm, SurA is thought to be the major chaperone interacting directly with the BAM complex. *In vivo* cross-linking with His-tagged BamA found SurA to be the only chaperone to co-purify with the complex (Sklar et al., 2007b). SurA with the photocrosslinker p-Benzoyl-L-phenylalanine (pBpa) introduced as an unnatural amino acid throughout its chain could be crosslinked to BamA *in vivo* via its N-domain and second peptidyl-prolyl isomerase (PPIase) domain (P2) (shown in blue and orange, respectively, in Figure 1.8), while crosslinks to OmpF and LamB were found exclusively in the N-domain (Wang et al., 2016). This study also showed crosslinks from the N-domain of SurA to the inner-membrane protein, PpiD, and an enrichment of BamA and SurA in fractions with SecYEG and PpiD in sucrose density gradient centrifugation experiments. This was remarkable as it suggested the presence of a supercomplex spanning the inner and outer membranes and involving the first proteins involved in this secretory process (PpiD and SecYEG), and

Table 1.3 Validated substrates for Skp and SurA. Crosses show that an interaction has been proposed between the named OMP and SurA (column 1), Skp (column 2), or both (column 3). Mass-spectrometry based differential proteomics coupled with genetic knockouts or depletions or coupled with a Strep-based pulldown of the chaperone with its bound substrates were used in these studies. Double crosses indicate interactions which have unambiguously been shown to be due to chaperone loss directly, and not downregulation of the gene as part of an envelope stress response. a) Data from Vertommen et al. (2009). b) Data from Jarchow et al. (2008). c) Data from Denoncin et al. (2012).

OMP (Integral)	$\Delta surA^a$	Skp _{pulldown} ^b	$\Delta skp:SurA$ depletion ^c
BamA		X	
BtuB		X	X
CirA		X	
FadL	X	X	X
FecA	X	X	X
FepA		X	X
FhuA	X X	X	X
Fiu		X	X
LamB	X	X	X
LptD	X X	X	X
OmpA	X	X	X
OmpC		X	X
OmpF	X	X	X
OmpG		X	
OmpLA		X	
OmpT			X
OmpX	X	X	X
TolC		X	X
Tsx		X	X
YncD		X	

the last protein involved in its folding (BamA). Further evidence for this supercomplex has come from a preprint showing immunoprecipitation experiments to pull down SecG, His-tag pull downs on BamA, and sucrose density gradient centrifugation of from extracted native membranes. This data shows an interaction between the BAM complex, SurA, and the Sec holotranslocon (which would comprise the supercomplex: [SecYEG-YajC-YidC-SecDF]-SurA-BamABCDE, with the Sec holotranslocon shown in square brackets) where the interaction is suggested to be mediated through SecDF (Alvira et al., 2019). This supercomplex interaction (albeit without direct detection of the involvement of SurA) has also been confirmed in a study on interacting partners of the *E. coli* cell envelope proteome using a membrane-mimetic system called peptidiscs to help solubilise membrane proteins and maintain their integrity after extraction and preparation for proteomics (Carlson et al., 2019). *In vitro* studies in the presence of the full BAM complex found SurA to increase the rate of folding of OmpT (Figure 1.6) in a concentration-dependent manner, as qualified by cleavage of a fluorogenic peptide by natively folded OmpT (Hagan et al., 2010). The same study also showed that the activity of SurA-delivered OmpT (and thus by proxy, the rate of β -barrel assembly by BAM) was significantly reduced in the absence of BamB (with approximately 3x less reporter peptide cleaved after 30 mins) indicating a coordinated interaction between the chaperone and this subunit (Hagan et al., 2010). This result recapitulated earlier *in vivo* studies showing that BamB and SurA null mutants had identical phenotypes (Ureta et al., 2007).

1.4.2 BamB plays an important role in OMP biogenesis

The crystal structure of BamB shows a β -propeller fold (Figure 1.4); these folds are often involved in scaffolding protein-protein interactions. On this basis it was suggested BamB may stabilise an interaction between SurA and BamA through modulating the flexibility of the POTRA domains of BamA (see Figure 1.7 and Figure 1.4), allowing more efficient handover of OMPs without being involved in substrate binding itself (Noinaj et al., 2011). Analysis of the crystal contacts in BamB structures have led other groups to suggest that BamB may bind substrate OMPs directly by β -augmentation through unsatisfied edges of its β -strands or through interactions of aromatic residues from unfolded OMPs with a hydrophobic pocket on the 'top' of the propeller (see Figure 1.4) (Heuck et al., 2011). Supporting this, photocrosslinking studies identified a weak interaction between the autotransporter OMP EspP and BamB during biogenesis (Ieva et al., 2011). As attractive as this seems, a few lines of evidence disfavour this model as a general mechanism for recognition of all OMPs by BamB. Isothermal calorimetry experiments failed to show an interaction between BamB and peptides derived from LamB and BtuB that represented

either β -strand regions or the C-terminal OMP consensus sequence (β -signal) (Jansen et al., 2012). Furthermore, most β -propeller motifs have inbuilt ‘design features’ to prevent non-specific pairing of unsatisfied hydrogen-bonds in β -strands at the edges (Richardson and Richardson, 2002). BamB follows this principle with most β -strands containing charged residues in the centre of potentially exposed regions, as well as β -bulges to introduce local twists which disfavour hydrogen-bonding on its convex side (Jansen et al., 2012). It has been suggested that the β -augmentation seen is actually an artefact of the artificial constraints imposed by crystal lattice packing in β -sheet proteins (O’Neil et al., 2015). Homologues of BamB are not present in all Alphaproteobacteria (a sub-phylum in the same phylum as *E. coli*), while SurA homologues are very well represented (Humes et al., 2019). Two examples are *Brucella melitensis* and *Neisseria* spp. which lack BamB but retain SurA homologues (Gatsos et al., 2008). This again suggests that BamB is not the sole interaction partner for SurA and unfolded OMPs. However, it was found that certain OMPs may prefer pathways which have different dependencies for SurA and BamB. LptD and FhuA were assembled normally in $\Delta bamB$ strains and also relied on SurA but could not be rescued by overexpression of SurA in a $\Delta skp \Delta fkpA$ background showing these chaperones are also required. However, simpler and more abundant substrates such as LamB and OmpA show the reverse phenotype with a lower dependence on SurA but a strong dependence on BamB (Schwalm et al., 2013). Initially it was thought that BamB may be only transiently associated with the BAM complex (which would allow it to interact with other proteins) on the basis of crystal structures solved independently by three groups which showed an open state of BAM that lacked BamB (5EKQ and 5D0Q - see Figure 1.9) versus a closed state where it was present (5D0O and 5AYW) (Bakelar et al., 2016; Gu et al., 2016; Han et al., 2016). A solution cryo-EM structure of BAM subsequently showed that an open BAM structure was not incompatible with BamB binding, suggesting that the lack of BamB was an artefact of weaker binding leading to dissociation during crystallisation or crystal packing artefacts, and that in fact BamB may play an important role in mediating the 3D arrangement of the POTRA domains (see 5LJO in Figure 1.9) (Iadanza et al., 2016). Finally, BamB was shown to be involved in mediating the formation of BAM islands *in vivo* through connections between two BamB molecules of neighbouring BAM complexes (Gunasinghe et al., 2018). This suggests it would not have a large accessible surface area for interactions with other proteins.

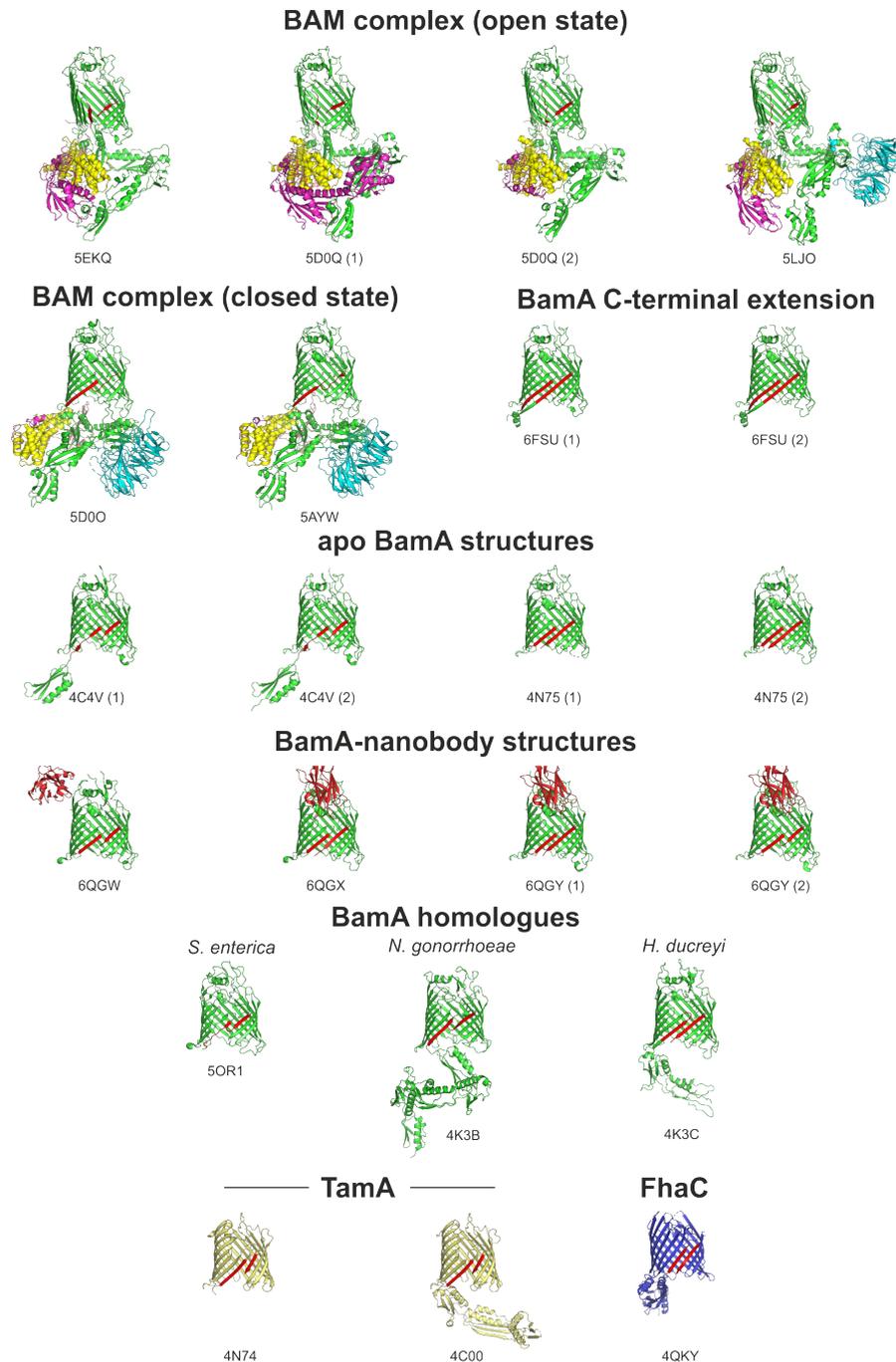


Fig. 1.9 Solved crystal and cryo-EM structures of BAM and Omp85 homologues. $\beta 1$ and $\beta 16$ strands are highlighted in red. Structures of the BAM complex in the open state (left to right): 5EKQ (Bakelar et al., 2016), 5D0Q (asymmetric unit 1) (Gu et al., 2016), 5D0Q (asymmetric unit 2), 5LJO (Iadanza et al., 2016). BAM complex in the closed state (left to right): 5D0O (Gu et al., 2016), 5AYW (Han et al., 2016). tBamA C-terminal extension: 6FSU (asymmetric unit 1) (Hartmann et al., 2018), 6FSU (asymmetric unit 2). Apo BamA(WT) structures (left to right): 4C4V, BamA($\Delta P1$ -P4) (asymmetric unit 1) (May and Grabowicz, 2018a), 4C4V (asymmetric unit 2), 4N75, tBamA (asymmetric unit 1) (Ni et al., 2014), 4N75 (asymmetric unit 2). BamA-nanobody structures, nanobody is coloured in red (left to right): 6QGW (Kaur et al., 2019), 6QGX, 6QGY (asymmetric unit 1), 6QGY (asymmetric unit 2). BamA homologues in other species (left to right): 5OR1, tBamA (*Salmonella enterica*) (Gu et al., 2017), 4K3B, BamA (*Neisseria gonorrhoeae*) (Noinaj et al., 2013), 4K3C, BamA($\Delta P1$ -P3) (*Haemophilus ducreyi*) (Noinaj et al., 2013). Omp85 homologues (left to right): 4N74, tTamA (*E. coli*) (Gruss et al., 2013), 4C00, TamA (*E. coli*) (Gruss et al., 2013), 4QKY, FhaC (*Bordetella pertussis*) (Maier et al., 2015).

1.4.3 BamD may be a site for initial engagement and prime BamA for folding

Other authors have hypothesized that BamD may be involved in binding substrates at the OM, and BamB and BamD alone are sufficient to assemble BamA into the OM (Hagan et al., 2013). A number of cross-linking studies have found interactions with BamD and unfolded OMPs. Pull down experiments with His-tagged BamD found BamD was bound to BamA and OmpA (Hagan et al., 2013), peptides derived from β -strands of the *Haemophilus influenzae* autotransporter, Hia, were able to cross-link with BamD *in vitro* (Albrecht and Zeth, 2011), and a peptide derived from the C-terminal β -signal of BamA bound to BamD and inhibited the assembly of OMPs by the BAM complex *in vivo* and *in vitro* (Hagan et al., 2015). Crystal structures of BamD have shown that it contains an elongated fold with five tetratricopeptide (TPR) repeats, an architecture which is known to be involved in recognition of targeting signals, as in Hop-Hsp90 and PEX5 which bind to the extended C-terminal tails of their respective cytosolic substrate proteins (Albrecht and Zeth, 2011; Sandoval et al., 2011). Sequence analysis of the TPR1-3 folds in BamD show three conserved patches which were suggested to have different roles in substrate and chaperone recognition (Albrecht and Zeth, 2011). The crystal structure of BamCD showed BamC to bind in one of these conserved patches which could indicate a mechanism of regulation for the binding of substrates in the periplasm (Kim et al., 2011a). This initial recognition of OMP substrates may trigger a conformational change in BamD which is then transmitted to BamA. A mutant of LptD with a mutation in a putative β -signal motif, Y721D, showed an early folding defect in the periplasm. A peptide comprising this β -signal region was found to bind to BamD with micromolar affinity in microscale thermophoresis (MST) experiments (Lee et al., 2018). A late-stage folding defective mutant, LptD4213, that fails to properly oxidise essential pairs of cysteines, could be crosslinked to BamD from positions L717 and Y726 when these were substituted with an unnatural amino acid photocrosslinker, pBpa (Lee et al., 2018). The LptD_{Y721D} mutant stalls on BamD and suppressor mutants to the OM defects caused by this stall were mapped to BamA_{F494L} and BamA_{G669A} in loop 6. These data suggest that correct binding to BamD is essential for communication between BamA and BamD and activation of the BAM complex (Lee et al., 2018). A temperature sensitive lethal mutant of BamA, BamA_{E373K}, containing a negative to positive charge substitution in POTRA5 near the interface with BamD and BamE was shown to disrupt the interaction between BamA and BamD, splitting the BAM complex into BamAB and BamCDE sub-complexes (Ricci et al., 2012). Suppressor mutants were mapped to BamD, BamD_{R197X}, where X was Leu, Ser, or His, which restores function but which doesn't regenerate a stable interaction between the two parts of the

full BAM complex (note that this doesn't rule out a specific but weak or transient BamAD interaction). Furthermore, in a mutant strain where levels of BamA are diminished and growth defects are normally observed (named *bamA101*, containing a transposon insertion and duplication upstream of the AUG start codon that lowers expression levels)(Aoki et al., 2008)), BamD_{R197L} could in fact rescue this growth defect implying that this is an 'activator' mutation of BamD compared to wild-type. Studies into the conformational plasticity of BamA in *bamD_{R197L}* strains showed a greater sensitivity to externally added proteinase K and increased labelling of native cysteines in loop 6 by PEG-maleimide (Rigel et al., 2013). The labelling of WT BamA loop 6 at lower levels (as opposed to no labelling) implies that this conformation is still accessible but less populated under normal conditions indicating a conformational equilibrium between the two states. This same study also isolated intra-gene suppressors of detergent sensitivity in BamA_{E373A} (which has a milder phenotype than E373K) that mapped to Q693P in loop 6 implying an allosteric interaction between POTRA5 and loop 6, one which may be controlled by BamD (Rigel et al., 2013). Investigating the structure of the BAM complex showed that there is a large, highly conserved electrostatic interaction network between BamD and POTRA5 which involves D363, R366, and E373 in BamA and Y176, E177, and R197 in BamD (Bakelar et al., 2016; McCabe et al., 2017). Charge-repelling mutations at this interface are lethal (i.e. positive-positive, negative-negative) but an intragenic suppressor of BamA_{E373K}, BamA_{K351E}, is able to restore function without restoring binding to BamD, showing that these residues are not key for function *per se* but instead alter the conformational ensemble of BamA needed for coordinated action with BamD (McCabe et al., 2017). This idea is supported by solution and solid state NMR experiments which show the local conformation of POTRA5 as part of an isolated P4-P5 construct or BamA(Δ P1-P3) to be dynamic, dependent on its internal electrostatic network, and shifting within a conformational ensemble (Sinnige et al., 2015). Taken together, the literature suggests that BamD may help regulate a conformational cycle of BamA, one where initial recognition of a substrate on BamD switches BamA from a 'resting' state to a 'substrate ready' state. Once a substrate has passed fully from BamD to BamA and folding is completed, this allosteric switch may be turned off, returning BamA to its 'resting' state. These two states may be the open and closed states seen in solved structures of BAM (Figure 1.9) but it is not clear which is the substrate ready, and which is the resting state, if indeed these represent all major conformations or whether BamA is in a dynamic conformational ensemble in a resting state with a particular conformation selected and 'locked in' upon substrate engagement.

1.4.4 BamE modulates the conformations of BamA and installs a stress sensor

Although the biochemistry of BamE has been less thoroughly studied, there is no current evidence suggesting a direct involvement in OMP substrate engagement and deletion causes only minor OMP assembly defects in *E. coli* (Sklar et al., 2007a). The phenotype of $\Delta bamC$ and $\Delta bamE$ are very similar but double deletion mutants of $\Delta bamC$ or $\Delta bamE$ with other mutations (that lower the levels of BamA, cause misfolding of LptD, delete *surA*, or delete both *skp degP*) caused much more severe effects when paired with $\Delta bamE$ than $\Delta bamC$ (Rigel et al., 2012). The double deletion of $\Delta bamC \Delta bamE$ also shows the least severe phenotype of all double deletions of OMP biogenesis components. Furthermore, BamA as part of a BamABCD complex in $\Delta bamE$ strains is extremely sensitized to the addition of proteinase K to the extracellular medium, causing it be completely degraded (an effect not observed upon deletion of *bamB* or *bamC* and implying that BamE may regulate the conformation of BamA or the permeability of the membrane around it (Rigel et al., 2012). Further studies on this protease sensitivity phenotype found that *bamE* deactivation is copied by the presence of a BamD_{R197L} activating mutant that prevents a stable complex between BamA and BamD in a WT background, and the presence of both mutations is lethal implying the two subunits have opposite roles in regulating BamA dynamics (Rigel et al., 2013). As discussed above for BamD_{R197L}, the conformational change in BamA caused by BamE deletion results in greater surface exposure of loop 6 as measured by labelling the native cysteines C690 and C700 by PEG-maleimide (Rigel et al., 2013). A role in modulating the local lipid environment around BamA also seems plausible as BamE has been shown to bind PG headgroups (in DHPG micelles) *in vitro* in NMR experiments (n.b. no binding was observed to PE or cardiolipin) (Knowles et al., 2011). Furthermore, intact MS of BAM ejected directly from the *E. coli* OM was stably bound to up to three cardiolipin molecules which the authors hypothesized was mediated by BamE (Chorev et al., 2018). RcsF is a lipoprotein sensor of outer leaflet defects in the outer membrane which is threaded through the centre of OMP substrates (primarily OmpA, OmpC, and OmpF) during their assembly on the BAM complex (Konovalova et al., 2014; Cho et al., 2014; Konovalova et al., 2016). Deletion of BamE causes major defects in the assembly of RcsF/OmpA complexes suggesting that BamE functions to help assemble this stress sensor network (Konovalova et al., 2016). A double deletion mutant of $\Delta bamB \Delta bamE$ causes major growth and OM defects and this phenotype was shown to be caused by lethal jamming of the BAM complex by RcsF which was relieved in a triple deletion mutant including $\Delta rcsF$ (Tata and Konovalova, 2019; Hart et al., 2019). A gain-of-function suppressor of BamA, BamA_{F494L}, had been shown to restore cell viability

in $\Delta bamB \Delta bamE$ strains, reduce the dependence of cells on BamD for viability, and improve engagement of a substrate with a mutation near the β -signal motif suggesting that it reduces the dependence of BamA for activation by BamD (Tellez and Misra, 2012; Misra, 2015; Lee et al., 2018). The suppression of the $\Delta bamB \Delta bamE$ phenotype by BamA_{F494L} therefore implies that the jamming by RcsF was due to a lack of coordination between BamA and BamD and this in turn suggests BamE modulates the engagement of BamA by BamD to allow proper folding of an OMP substrate when RcsF is also bound to BamA (as BamB does not directly contact BamD - see Figure 1.9) (Hart et al., 2019; Tata and Konovalova, 2019). As RcsF is unusual for a lipoprotein in having its lipid anchor embedded in the outer leaflet of the OM, BamE may also be involved in aiding this surface translocation via BamA. RcsF was also shown to play a key role in the lethality of mutant strains lacking PG and CDL (Shiba et al., 2004), pointing to an additional link between the function of BamE in binding lipids with RcsF. Recently, the BAM complex has been shown to be capable of assembling cell surface proteins independently of any β -barrel partner (González-Rivera et al., 2019) and it will be interesting to see if BamE forms part of a general pathway for surface-exposure of proteins via BAM. *In silico* co-evolution studies of the whole proteome of *E. coli* on a residue-by-residue basis indicated a co-evolved interface between SurA and BamE, with residues particularly concentrated in the P2 domain of SurA (as well as the core N- and C-domains) (Cong et al., 2019) implying a hitherto undiscovered role of BamE in chaperone docking.

1.4.5 Recognition of OMPs by BamA

Although in *E. coli*, BamD and BamA both have key roles, the position of BamA as the only fully conserved subunit of the BAM complex in all species with two membranes suggests it would be the primary recognition site for OMPs. Commensurate with the above discussion, co-immunoprecipitation experiments with wild-type BamB and mutant BamB sequences containing weakened BamA binding regions (L173S, L175S, R176A), or BamA alone as bait, found that as BamB variants become more defective in BamA binding, the amount of SurA precipitated was reduced (Vuong et al., 2008). This suggests that SurA bypasses BamB when delivering unfolded OMPs to BAM by direct interaction with BamA or another BAM lipoprotein. The N-terminal periplasmic region of BamA contains a string of five polypeptide-transport associated (POTRA) domains (Figure 1.4). The crystal structure of POTRA1-4 (POTRA1-5 constructs could not originally be crystallized), and *in vivo* mutagenesis studies of BamA POTRA deletions of varying length, gave the first indications that unfolded OMPs may interact with these domains via β -augmentation (Kim et al., 2007; Gatzeva-Topalova et al., 2008). While removal of any one POTRA domain

causes some degree of OMP assembly defect *in vivo*, deletion of POTRA 3-5 show the most marked effect. Interestingly, similar POTRA deletion studies in the BamA homologue of *N. meningitidis* (Figure 1.9) showed little or no OMP assembly defects upon deletion of POTRA 1-4, but found POTRA5 to be essential (Bos et al., 2007). The POTRA domains are structural hallmarks of the Omp85 superfamily of protein secretion and OMP assembly machines and have been implicated in the binding of unfolded protein substrates in other macromolecular systems such as FhaC (Figure 1.6) (Clantin et al., 2007). In the chloroplast outer membrane, an Omp85 homologue, Toc75 contains 3 POTRA domains and has the unusual feature of importing proteins through the ‘top’ of the barrel (i.e. via the loop end) where they are hypothesised to interact with POTRA domains upon emergence into the inter-membrane space. Crystal structures of the POTRA domains showed a remarkable structural homology to *E. coli* POTRA domains despite the evolutionary divergence of the host, and these were shown to interact directly with a chloroplast pre-protein, preSSU, by nickel-affinity pull down on His-tagged POTRA in the presence of preSSU (O’Neil et al., 2017).

Detailed characterisations of the POTRA domains have given more insight into their potential interaction with OMPs. ^1H - ^{15}N -resolved NMR studies of POTRA 1-2 screened against β -strand and β -hairpin peptides derived from the porin PhoE showed chemical shift perturbations in the β -strands of POTRA 1 and 2 indicative of binding (see Figure 1.4 for orientation of POTRA domains) (Knowles et al., 2008). However, at 1.5 mM the peptide failed to saturate binding as observed by chemical shift perturbations in a ^1H - ^{15}N HSQC spectra, indicating that this binding interaction is weak. Despite this, the increased number of individual POTRA domains (i.e. POTRA 1, 2, 3, 4, 5) in a single BamA molecule, or a cluster of BAM molecules present in OM OMP islands *in vivo* (Rassam et al., 2015), may provide avidity to the interaction with substrate OMPs. However, these experiments were obtained in the presence of a 50x molar excess of peptide and the authors were unable to examine peptides from C-terminal β -signal sequences due to their propensity to aggregate. These β -signal motifs are almost always found at the C-terminus of an OMP and have been proposed to be involved in recognition of OMPs by the chaperone and BAM machinery in the periplasm and OM. Studies on mutant strains of *E. coli* deficient in OMP assembly found that a single in-frame deletion of R64 in POTRA 1 of BamA was sufficient to cause OM permeability and OMP assembly defects (Bennion et al., 2010). BamA-mutant mediated pull down assays showed the region of POTRA 1 from R36-K89 to be important for interactions with SurA, and indicated an interaction site with helix 2 of POTRA 1 (Bennion et al., 2010). Chimeric BamA proteins featuring mix-and-match POTRA from different species showed that there are no species-specific interactions with unfolded OMPs or chaperones at the putative POTRA binding interfaces within

POTRA 1-3, suggesting that the interactions in these regions are driven by gross structural features (such as β -augmentation) rather than conserved sequence recognition (Browning et al., 2015). Similarly, hybrid studies in mitochondria using Sam50 (which natively has one POTRA) with *E. coli* POTRA domains, or the *E. coli* barrel and Sam50 POTRA domains showed that Sam50 could tolerate the replacement of its only POTRA with POTRA5 from *E. coli*, and to a lesser degree with a fusion of the native Sam50 with POTRA1-4 (Pfitzner et al., 2016). Replacement of the barrel domain, however, was lethal in all variations. This agreed with the bacterial mix-and-match experiments that the POTRA domains may recognise a structural feature (or play a generic modulatory role) as the sequences of mitochondrial OMPs would be highly divergent due to the large evolutionary time since mitochondria branched off from free-living bacteria. Interestingly, bioinformatic analyses showed the C-terminal β -signal varied at certain positions in a manner specific to the OMP class (defined by its number of β -strands) and taxonomic group / species (Paramasivam et al., 2012). BamA reconstituted in lecithin planar lipid bilayers could not fold either PhoE lacking its C-terminal phenylalanine or PorA from *N. meningitidis*, and C-terminal 11 or 12-mer peptides derived from both failed to ‘activate’ BamA (defined as observing a channel opening event) (Robert et al., 2006). This evidence points to a *structural* recognition occurring within the POTRA domains followed by the C-terminal β -signal being recognised at the β -barrel interface of BamA (see regions shaded red in Figure 1.9). A few bacteria are known to have variant forms of LPS (Miller et al., 2005) (Figure 1.3) or unusual OM characteristics (cholesterol glycosides are found in *Helicobacter* OMs (Hirai et al., 1995) and LPS is non-essential in *Neisseria* (Bos and Tommassen, 2005)). This may lead to species-specific ‘design’ of certain regions of OMPs to accommodate the structural differences between OMs and could explain the observation that *Helicobacter* and *Neisseria* OMPs are difficult to express in *E. coli* (Gotschlich et al., 1987; Carbonetti and Sparling, 1987; Humphries et al., 2002). Alternatively, BamA may have co-evolved with the OMs as the β -barrel is in direct contact with the bilayer and as this is putatively the recognition site of the C-terminal β -signal, this motif itself may have co-evolved with BamA. *Helicobacter* C-terminal signals have an unusually strong preference for tyrosine at position +5, and *Neisseria* have preference for positively charged residues at position +2 (Paramasivam et al., 2012). Could these differences be related to the membrane structure, rather than a species-specific ID tag? Recognition at the BamA β -barrel may be driven primarily by the thermodynamics or sterics of β -strand insertion or templating at the putative BamA β 1- β 16 ‘gate’ (see Figure 1.9). Evolution does not tend to design, but rather works with what is already in existence, and other systems such as periplasmic chaperones and proteases may have evolved to recognise the consistency of the C-terminal sequence as a consensus sequence, removed from its original thermodynamic considerations. Supporting this idea, an *in vitro* PhoE expression system

coupled to folding in detergent in the absence of any folding catalysts, $\Delta\text{Phe}_{\text{C-terminal}}$ mutants showed a reduction in both stability and folding rates (de Cock et al., 1997). *In vitro* folding experiments with transmembrane OmpA showed that a tOmpA-Skp complex could be released into synthetic DUPC membranes (which tOmpA alone is capable of folding into) in the presence of either BamA or just the transmembrane domain, tBamA (without POTRAs), but not in the presence of SurA or soluble POTRAs alone (Schiffrin et al., 2017b). This suggested that some kind of recognition of BamA by Skp or by tOmpA, is possible even in the absence of soluble POTRA domains.

1.5 The final step: mechanisms of membrane insertion

1.5.1 Thermodynamic barriers to folding

Before addressing the energetic role played by the BAM complex in OMP assembly, it is worth reviewing which factors augment OMP folding and assembly or inhibit folding *in vivo* and *in vitro*.

Measured in 1-palmitoyl-2-oleoyl-phosphatidylcholine (POPC, $\text{C}_{16:0}:\text{C}_{18:1}$) liposomes, the insertion of each peptide bond into the bilayer (and thus the dehydration of the protein backbone) has an energetic cost of $\sim 5.0 \text{ kJ}\cdot\text{mol}^{-1}$ (Wimley and White, 1996). This energy cost must be overcome by the interactions formed through secondary and tertiary structure when a protein folds into lipids. Formation of secondary structure within the bilayer by both helical proteins and OMPs contributes favourable backbone hydrogen bonds. Few studies have investigated the energetics of hydrogen bond formation in β -barrel membrane proteins, but studies on a hexapeptide which forms β -sheet aggregates in POPC and 1-oleoyl-2-(9,10-dibromostearoyl)-phosphocholine (OBPC, $\text{C}_{18:1}:\text{diBr-C}_{18:0}$) showed free energy reductions of around 2.1 to 2.5 $\text{kJ}\cdot\text{mol}^{-1}$ per residue for formation of β -sheets (Wimley et al., 1998). This agrees well with a more robust analysis of the energetics of formation of eight hydrogen-bonds throughout the α -helical bacteriorhodopsin which supported an average contribution of $-2.5 \text{ kJ}\cdot\text{mol}^{-1}$ per hydrogen-bond (Joh et al., 2008). Using OmpA as an example, folding of the β -barrel region (excluding loops and turns) buries around 62 peptide bonds ($+311.5 \text{ kJ}\cdot\text{mol}^{-1}$) and forms around 72 hydrogen-bonds ($-180.9 \text{ kJ}\cdot\text{mol}^{-1}$) which still leaves over $130 \text{ kJ}\cdot\text{mol}^{-1}$ unaccounted for if the fold is to be energetically favourable. This difference can be attributed to a number of factors: the entropic benefit of shielding hydrophobic regions/residues from water, enthalpic benefits of hydrophobic interactions with acyl chains (Ferguson et al., 2000; Pautsch and Schulz, 2000), formation

of charged and polar interactions with LPS and other phospholipids (Ferguson et al., 2000), formation of salt bridges within the barrel interior (Hong et al., 2006), the degree of barrel tilt or shear (Chou et al., 1990), β -strand twist and relief (Dou et al., 2018; Kikuchi et al., 2018), formation of dimers or trimers (Ma et al., 2018), membrane-water interfacial aromatic residues (Hong et al., 2007; Chaturvedi and Mahalakshmi, 2018b), hydrophobic mismatch (Xu et al., 2008; Marsh, 2008; Muhammad et al., 2011; Yin and Kindt, 2012; Srivastava et al., 2018), bilayer curvature stress (Hong and Tamm, 2004; Pocanschi et al., 2006b; Marsh et al., 2006), π -interactions (π -stacking, cation- π , anion- π) (Gallivan and Dougherty, 1999; Thompson and Smithrud, 2002; Hong et al., 2006, 2007; Dou et al., 2018), aromatic-glycine (so-called ‘mortise-tenon’) interactions (Merkel and Regan, 1998; Jackups and Liang, 2005; Leyton et al., 2014; Michalik et al., 2017), and likely many more yet to be discovered. Unlike α -helical membrane proteins which are highly hydrophobic and insoluble in both their folded and unfolded states outside a membrane environment, β -barrel membrane proteins are relatively hydrophilic (Kleinschmidt, 2003) and hence are soluble in chaotropes such as urea and guanidinium chloride, and some at low concentrations in water itself (at least for a physiologically relevant length of time – 5–20 min and *E. coli* generation time is ~20 min) (Danoff and Fleming, 2015a; Humes et al., 2019). One might consider, therefore, that the driving force for burying an OMP into a lipid bilayer would be low compared with the energetic barrier to insertion which would be relatively high. Despite this, the OM of Gram-negative bacteria is densely packed with β -barrel structures (see Figure 1.6 and Table 1.1) with copy numbers of the most abundant OMPs in *E. coli* between 10,000 – 200,000 per cell (Henning et al., 1973; Rosenbusch, 1974; Lugtenberg and Van Alphen, 1983; Li et al., 2014). Furthermore, for the few OMPs whose thermodynamic free energy profiles have been well characterised (AIDA-I (Mogensen et al., 2005), Ail (Chaturvedi and Mahalakshmi, 2018a), OmpA (Hong and Tamm, 2004; Hong et al., 2006, 2007; Sanchez et al., 2008; Andersen et al., 2012; Pocanschi et al., 2013), OmpLA (Moon and Fleming, 2011; Moon et al., 2011; McDonald and Fleming, 2016), OmpX (Chaturvedi and Mahalakshmi, 2013, 2018a,b), OmpW (Moon et al., 2013), PagP (Huysmans et al., 2010; Moon et al., 2013; Iyer and Mahalakshmi, 2016; Iyer et al., 2017; Marx and Fleming, 2017; Iyer et al., 2018; Iyer and Mahalakshmi, 2019), and PulD (Guilvout et al., 2017)) the folded state under *in vitro* conditions is remarkably stable with $\Delta G^0_{\text{Folding}}$ values between ~ -10 to ~-140 kJ.mol⁻¹.

1.5.2 Direct role of the BAM complex in OMP assembly: lowering the energy barriers

As discussed in Section 1.1, BamA is the only completely conserved component of the BAM complex. This suggests that while other lipoproteins may augment OMP folding, assembly and insertion, BamA likely has the greatest effect on lowering the activation energy barrier posed by the lipid bilayer.

1.5.2.1 Effects of BAM at the lipid interface and the role of POTRAS

Studies using small unilamellar vesicles (SUVs) formed from native *E. coli* polar lipid extract or synthetic large unilamellar vesicles (LUVs) containing 20:80 didecanoyl (*diC*_{10:0}) PE:PC showed OMP folding to be severely retarded compared with folding into *diC*_{10:0}PC liposomes alone (Gessmann et al., 2014) (n.b. the inner leaflet of the OM contains around 80 % PE within *E. coli* (Morein et al., 1996)). However, the addition of pre-folded BamA into the SUVs/LUVs increased the folding rates for OmpA, OmpLA, and OmpX (Gessmann et al., 2014). This suggests that the lipid head groups of the native membrane in *E. coli* do not allow spontaneous insertion for most β -barrel proteins at a physiologically relevant rate. Hence, the BAM complex may be required to accelerate this step *in vivo*. *In vitro* experiments on a number of OMPs have shown that lower hydrophobic thickness and increased curvature promote the folding of OMPs (Kleinschmidt and Tamm, 2002; Hong and Tamm, 2004; Pocanschi et al., 2006b; Burgess et al., 2008). Crystal structures of BamA show a shortening on one side of the transmembrane barrel around strands β 1 and β 16 suggesting the bilayer may thin in this region to account for the hydrophobic mismatch between the protein and the lipid bilayer (see Figure 1.9). MD simulations support this, showing thinning and that the lipid tails are also less ordered in this region (Figure 1.10) (Noinaj et al., 2013, 2015; Fleming et al., 2016; Lundquist et al., 2018). Long time-scale simulations of BamA from a number of homologues have shown the importance of a ‘glycine kink’ at G807 (in *E. coli* gene numbering) which is conserved across BamA homologues (Lundquist et al., 2018). Located in the terminal (β 16) strand (see for example 1.9 5D0Q & 5D0O versus 4N75), this glycine causes the last 3 residues to ‘kink’ inward to the barrel lumen, preventing hydrogen bonds that would otherwise be formed with β 1. *In vivo* assays showed that substitutions of this residue are poorly tolerated, causing severe OM permeability defects (Lundquist et al., 2018). These kinked and non-kinked forms were able to be conformationally selected experimentally through the use of nanobodies which complex with the loops of BamA (structures shown in Figure 1.9 [BamA-nanobody]) (Kaur et al., 2019). The binding site for these nanobodies are in the

loops of BamA (loops 3, 4, and 6) indicating that conformational changes to the gate region of BamA can be controlled by the extracellular loops (and presumably *vice versa*). A bactericidal monoclonal antibody (mAb), MAB1, against BamA also showed binding on loop 4, while a high affinity (but neutral) antibody, MAB2, bound to loop 6 (Storek et al., 2018a). The structures of the *nanobody*-bound BamA show that loop 4 binding tended to stabilise the kinked state, while binding to loop 3 + 6 stabilised the closed form. If the effect of loop 4 binding is consistent between mAbs and nanobodies, then a permanently kinked form of BamA may be lethal, underlining the importance of being able to switch conformational states. Furthermore, a genetic interaction between MAB1 and the fluidity of the outer membrane was discovered in sequencing of spontaneous MAB1-resistant strains which revealed mutations in the *lpxM* gene that prevented the transfer of an additional acyl chain to the lipid A moiety of LPS (see Figure 1.2 and Figure 1.3). These mutations were shown to cause a decrease in membrane fluidity and that BAM is sensitive, perhaps paradoxically, to high membrane fluidity (Storek et al., 2018a,b). Early simulations also revealed a transient opening of the BamA β -barrel between strands $\beta 1$ and $\beta 16$ associated with motions of the POTRA5 domain (Noinaj et al., 2013; Lundquist et al., 2018). The recent crystal structures of the full BAM complex supports this observation and shows that the POTRA domains occlude OMP entry via the bottom of the barrel lumen. These structures suggest that binding of the subunits BamBCDE can cause a twist in the barrel (cf. open structures of the BAM complex in Figure 1.9 versus closed BAM complex and other BamA alone structures) breaking the interaction between strands $\beta 1$ and $\beta 16$, leading to the formation of an exit pore in the extracellular side of the barrel - see Figure 1.11) (Bakelar et al., 2016; Gu et al., 2016; Iadanza et al., 2016). It is not clear how these subunits are responsible for modulating the conformational equilibria of BamA, but it appears to be driven by the POTRA domains whose conformation is correlated with the open-close state of the barrel. Although BamB binds at a flexible hinge region of the POTRA domains (between POTRA2 and POTRA3), it does not appear essential for driving this motion as it is also seen in structures of BAM which contain or lack BamB (Bakelar et al., 2016; Iadanza et al., 2016). Nonetheless, BamB may constrain the conformational ensemble of the POTRA domains which could be key for function. Solution NMR studies of POTRA1-5 in isolation have shown that, in approximate terms, the POTRA domains behave as two rigid bodies with a flexible hinge between POTRA1-2 and POTRA3-5 (Warner et al., 2017) - the same location where BamB binds. This study also used disulfide crosslinking to tether POTRA2 and POTRA3 and restrict their dynamics in a BamA depletion strain complemented with the tethered BamA mutant and found that growth of these strains is impaired and is lethal after a few generations. AFM studies on BamA in native OMVs immobilised on a mica surface and unfolded by pulling on the N-terminus found that the force required to unfold BamA β -strands 1-4 was lowered by ~ 30 – 50 pN in the absence

of the POTRA domains indicating a mechanical link between POTRA domains and this region of BamA which characterises the open and closed state of BAM (Thoma et al., 2018).

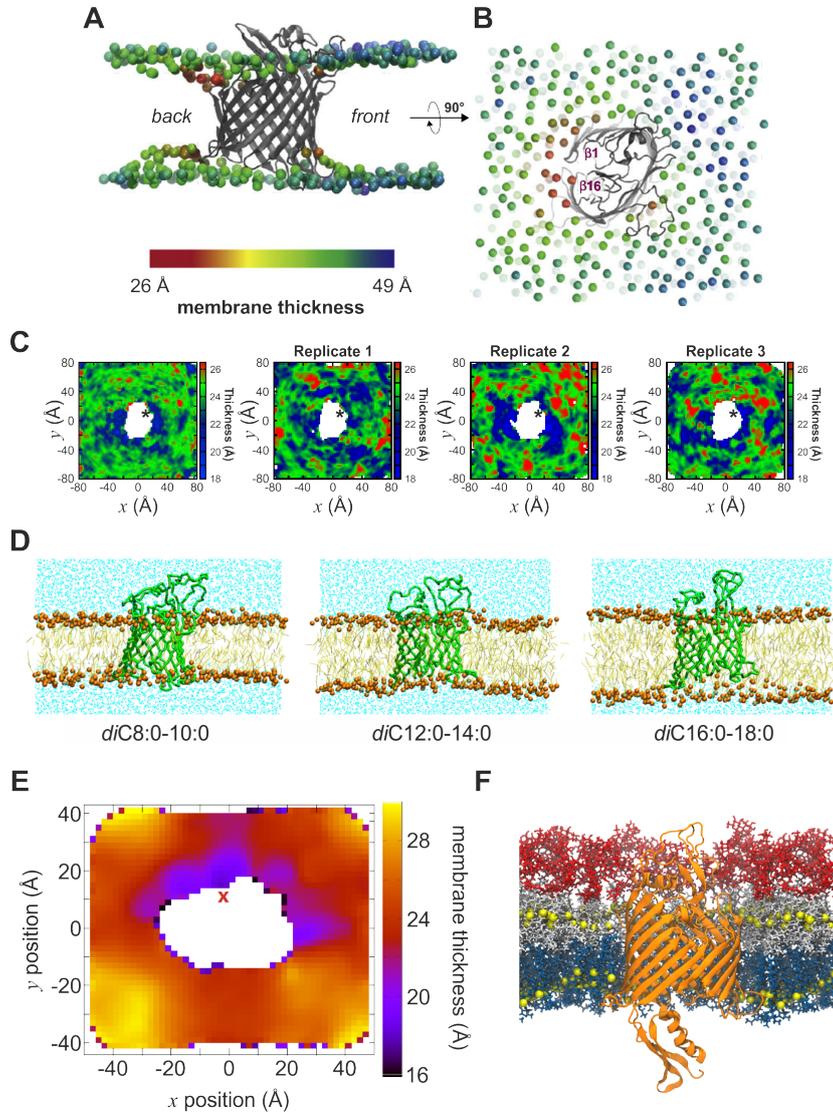


Fig. 1.10 BamA can create local disorder and/or membrane thinning of lipid bilayers. (A-B) Atomistic MD simulations of the BamA barrel from *N. gonorrhoeae* in a DMPE bilayer showing localized thinning of the membrane around the region of $\beta 1$ - $\beta 16$ by as much as 16 Å. (A) View of the BamA barrel from the side with phosphate headgroups represented as spheres (acyl tails not shown). (B) Same as (A) but shown from the top looking down into the barrel. Figure from Noinaj et al. (2015), data from Noinaj et al. (2013). (C) The thickness of asymmetric lipid bilayers surrounding BamA from atomistic MD simulations of a 'native' lipid bilayer with full length BamA embedded. Thickness defined as the average distance between the C2 and C4 atoms of lipid A of LPS and the C2 positions of phospholipids. White space is the area taken up by BamA. The starting structure was a homology model of *E. coli* described in Noinaj et al. (2014). The bilayer contained LPS (lipid A and R1 core) in the outer leaflet and a mixture of lipid with PE, PG, and cardiolipin headgroups. Approximate location of the $\beta 1$ - $\beta 16$ seam is indicated with a black asterisk. Figure from Fleming et al. (2016). (D) Coarse-grained (CG) MD simulations of the transmembrane region of BamA in PC bilayers of varying hydrophobic thickness (CG-MD represents multiple atoms as single 'beads' so the named lipid type is only an approximation). Figure from Schiffrin et al. (2017b). (E-F) Atomistic MD simulations of *E. coli* BamA homology model from Noinaj et al. (2014) embedded in an asymmetric lipid bilayer model of the *E. coli* OM. The outer leaflet contained LPS (lipid A and core oligosaccharide) and the inner leaflet contained POPE ($C_{16:0}$: $C_{18:1}$ PE). (E) Average membrane thickness over 4.2 μ s of simulation. (F) Structure of *E. coli* BamA homology model embedded in asymmetric membrane (snapshot from MD). Figure from Lundquist et al. (2018).

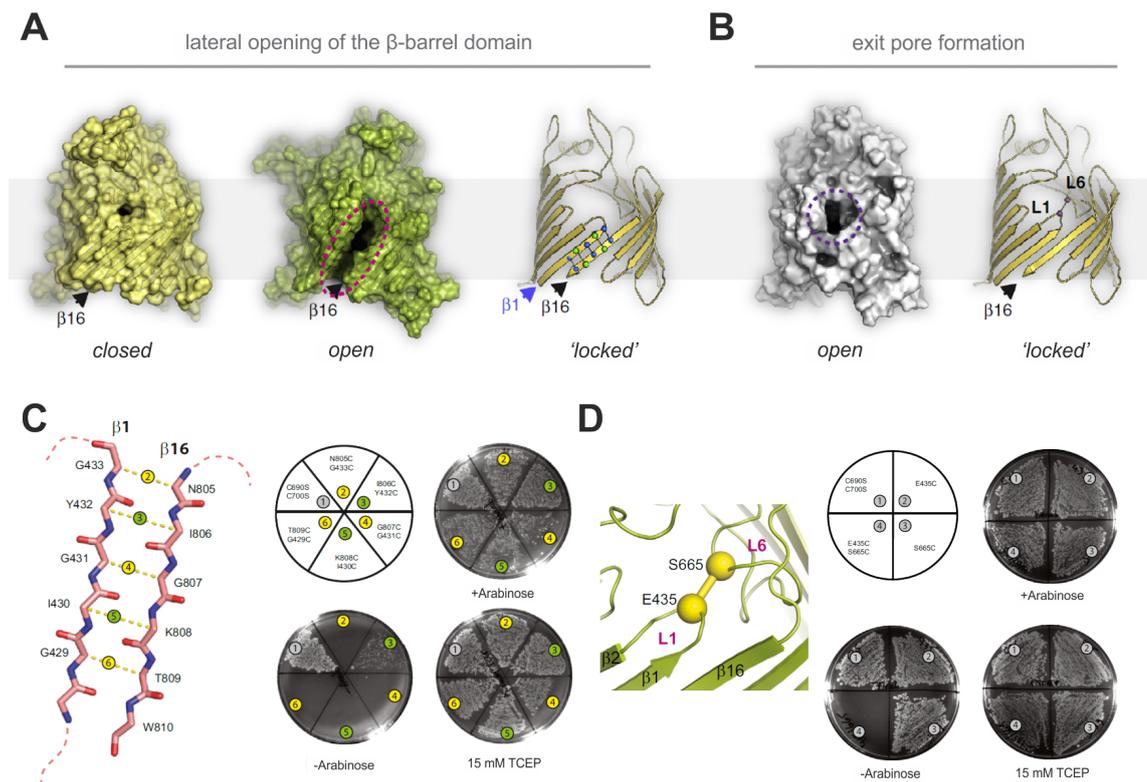


Fig. 1.11 BamA can open between $\beta 1$ and $\beta 16$ forming a 'lateral gate' as well as forming an 'exit pore' on its extracellular side which is coordinated by loop 6. (A-B) 'Closed' structure shows homology model of *E. coli* BamA built from solved by X-ray crystal structure of *H. ducreyi* and *N. gonorrhoeae* BamA, 'open' structure is a static snapshot taken from an MD simulation where the same protein was used as a starting structure. 'Locked' shows the locations of double cysteine mutants of *E. coli* BamA which have been used *in vivo* and *in vitro* to assess the functional role and importance of both the lateral gate and the exit pore. (C) Left: Zoomed in view of the $\beta 1$ - $\beta 16$ lateral gate of BamA showing the residue numbers and locations of double cysteine mutants used to assess viability and function. Right: *In vivo* plate assays of BamA disulfide locked mutants in a BamA(WT) depletion strain. Depletion of WT BamA only occurs in the absence of arabinose allowing both toxicity (expression of WT and mutant, +arabinose) and functional lethality (expression of mutant only, -arabinose) to be assessed. (D) Left: Zoomed in view of the BamA 'exit pore' showing the residue number and location of a 'lid-locked' mutant. Right: As in (C), assessed for the lid-lock mutant. Figure A+B from Noinaj et al. (2015). Figure C+D from Noinaj et al. (2014).

1.5.2.2 Templated folding of OMPs by the β 1- β 16 seam

The observation of opening of BamA at the β 1- β 16 seam prompted researchers to suggest the possibility of a model of BAM-assisted OMP folding where the β -strands of the nascent OMP (initially guided by the β -signal) would interact in a hairpin-by-hairpin manner with β 1 of BamA (and possibly β 16) to gradually form a super-barrel (see Figure 1.12A) (Noinaj et al., 2017). The first direct experimental evidence for this model came from studies on the Sam50-mediated folding of Tom40, Por1, and VDAC1 in mitochondria of *Saccharomyces cerevisiae*. Using truncated variants of the 19-stranded mitochondrial OMP, Por1, the authors created a large number of cysteine mutants throughout the first and terminal strands of Por1 and throughout Sam50 (the mitochondrial homologue of BamA) (Höhr et al., 2018). They coupled this with a number of oxidising agents and Cys-to-Cys crosslinkers to increase the available crosslinking distances (which normally require direct stable contact for disulfide formation if using Cys only). The β -signal strand from Por1 and Tom40 could be crosslinked to Sam50 β 1 but not β 16, furthermore, these residues from Sam50 β 1 were facing into the barrel lumen implying lumenal insertion rather than directly through the membrane. The N-terminal strand of Por1 was able to interact with both lumen and lipid facing residues of Sam50 β 16 showing that this interface is much more weakly defined. This lumen-mediated insertion model is not without precedent in Omp85 proteins. FhaC from *Bordetella pertussis* has been shown to translocate its substrate filamentous hemagglutinin (FHA) through its lumen to the outside of the cell (Baud et al., 2014), the contact dependent inhibition protein CdiA may use *E. coli* BamA as a conduit for translocation from the outside of the cell (Aoki et al., 2008), and Toc75 of the chloroplast outer envelope imports preproteins through its pore (O'Neil et al., 2017). Furthermore, *E. coli* BamA has a relatively hollow core with a large number of conserved Gly residues (particularly in β 1) and it also lacks a plug domain which is often found in other larger OMPs. In isolated mitochondria the authors found that deleting Sam50's only POTRA domain didn't prevent crosslinking to β 1 or β 16 by Por1 showing that recognition by Sam50 can be independent of the POTRA domains. Finally, an interaction with loop 6 of Sam50 was observed from loops of Por1 and a mutation of the conserved VRGF motif reduced the ability of Por1 to pull down Sam50. Mitochondria have a much more limited repertoire of OMPs (the human genome is only known to encode 7 transmembrane β -barrels) and it was not clear whether these results were generally applicable to bacterial BamA which has to fold a huge variety of substrates (as discussed in Section 1.2.1 and shown in Table 1.1). Furthermore, the use of a truncated substrate for crosslinking could have caused some artefacts in its folding pathway through Sam50. Hartmann et al. (2018) created a variant of the *E. coli* BamA barrel with a 31-residue C-terminal extension corresponding to an OmpX hairpin in order to try and stabilise the conformational heterogeneity of

BamA and allow assignment of its NMR spectra (Hartmann et al., 2018). Trimming this extension down they found that the minimum extension length to inhibit this heterogeneity as measured by NMR was 9-residues and solved the crystal structure of this construct (see Figure 1.9 [BamA C-terminal extension]). This showed that the extension was capable of forming hydrogen bonds with $\beta 1$ and ‘zipping’ up the seam, in a manner analogous to a templating intermediate of the β -signal from a substrate OMP. Evidence is also now available for a templated mechanism of BamA-mediated OMP folding in *E. coli* using a stalled intermediate of the EspP autotransporter (from the pathogenic *E. coli* strain O157:H7) where the native extracellular secreted domain was replaced with the fast folding maltose binding protein (MBP) (forming MBP-EspP) to trap an EspP intermediate (Doyle and Bernstein, 2019). Cysteines were then introduced into MBP-EspP and BamA at various positions and crosslinking could be observed from the C-terminal strand of MBP-EspP (the β -signal) to $\beta 1$ of BamA. In contrast to disulphide bonding studies within $\beta 1$ - $\beta 16$ of BamA which showed these strands can slide out of register by as much as 14 Å (an offset of 4 residues) (Doerner and Sousa, 2017), the MBP-EspP $\beta 12$ / BamA $\beta 1$ interface showed dramatically reduced crosslinking when out of register by just 1 residue indicating a specific and rigid interaction. An interaction was also observed between the N-terminal residues of MBP-EspP ($\beta 1$) and BamA $\beta 16$ although this was much more diffuse and showed lower crosslinking efficiencies than the β -signal $\beta 1$ interaction. However, using a Cys-to-Cys crosslinker, much greater crosslinking yields could be observed between MBP-EspP $\beta 1$ and both BamA $\beta 16$ and $\beta 15$. Overall this study suggests a templating mechanism as shown in Sam50, but unlike that protein, BamA doesn’t appear to thread β -strands through the BamA lumen and instead OMPs may be folded in the periplasm (either partially before, or after engagement with BamA $\beta 1$) consistent with studies on EspP (Ieva et al., 2008; Ieva and Bernstein, 2009; Pavlova et al., 2013), LptD (Lee et al., 2016; Wzorek et al., 2017; Lee et al., 2018), and the trimeric autotransporter UpaG (Sikdar et al., 2017). With the assembled OMP still attached to BamA, the authors propose a model where $\beta 1$ of BamA swings up into the membrane to insert a pre-folded barrel.

1.5.3 Summary of possible BAM mechanisms

Based on the evidence presented in this introduction, Figure 1.12 suggests three possible mechanisms by which BAM aids in OMP folding and insertion: 1) in the ‘BamA-assisted’ model unfolded OMPs are recognised by BamA POTRA domains, BamD, or are delivered directly to the membrane, where a destabilised region of the bilayer has been generated by BamA thereby lowering the activation energy barrier of insertion. OMPs then thread into the membrane strand-by-strand or insert via a concerted mechanism as has been

shown for uncatalyzed folding (Kleinschmidt et al., 1999; Kleinschmidt and Tamm, 1999; Kang et al., 2012). OMPs do not make any direct contact with the BamA barrel domain in this model (Figure 1.12A); 2) for the ‘BamA-budding’ model unfolded OMPs enter through the barrel lumen via their C-terminal strand and are fed into the destabilised membrane by templating β -stands against $\beta 1$ and $\beta 16$ of BamA. This causes a ‘blebbing’ or ‘budding’ of a nascent barrel from BamA which eventually splits away when sufficient structure is formed (Figure 1.12B); 3) The BamA $\beta 1$ -elongation model is a hybrid of models 1) and 2) where strands are templated against $\beta 1$ of BamA but a significant amount of structure is pre-formed in the periplasm, possibly in the cradle of the periplasmic ring of POTRA domains. This could explain the widespread conservation of five POTRA domains in bacteria, with some species in the Acidobacteria and Thermus-Deinococcus phylum, and the δ -Proteobacteria sub-phylum, having six or seven, while the more evolutionarily ancient phylum Fusobacteria having four and the earliest bacterial ancestors of the Omp85 family (the Cyanobacteria) appearing have the minimum of three (Arnold et al., 2010; Heinz and Lithgow, 2014; Hug et al., 2016). This periplasmic folding could occur before or after recognition of an OMP β -signal, and folding could be completed by either a more passive insertion by the intrinsic affinity of the OMP for the hydrophobic bilayer (Schiffrin et al., 2017a) or by the active ‘swinging’ of the folded barrel up and into the membrane by a movement of BamA $\beta 1$ (Figure 1.12C). These mechanisms are not mutually exclusive and it may be that BAM utilises different routes for different substrates, a process which could be modulated by subunits BamB–E. The differences in numbers of POTRA domains between the mitochondrial BamA homologues (Sam50, which has a single POTRA domain) and bacterial BamA homologues (which rapidly increase from three to five or more upon diverging from Cyanobacteria) suggests that the mechanism may have also diverged from the last common ancestor of mitochondria and bacteria. This would explain the disparities between the recent crosslinking studies on Sam50 and BamA (Höhr et al., 2018; Doyle and Bernstein, 2019) where the essential element of recognition of the β -signal at Omp85 $\beta 1$ is retained, but the subsequent steps seemed to differ. Sam50, which is located in the protective environment of the eukaryotic cell and has subunits (Sam35 and Sam37) bound on or near its cytosol-facing loops, may be sterically more free to open and dilate in size than BamA homologues which tend to have more POTRA domains and more bound periplasmic lipid-anchored subunits. Furthermore, for bacteria, which are generally free-living, maintaining the impermeability of the OM (to molecules larger than ~600 Da) is essential for survival against external threats and the formation of a large super-barrel / hybrid-barrel threatens to form unprotected pores in the OM, particularly for larger substrates such as TBDRs (22-strands), fimbrial ushers (24-strands), and LptD (26-strands), which are abundant in the OMP-ome (Table 1.2). For mitochondrial outer membranes, these are freely permeable to uncharged molecules up to 5 kDa (Vander Heiden

et al., 2000), the cytosol is isotonic to the intermembrane space, and it seems unlikely that many proteins larger than 5 kDa could readily access any temporary hybrid-barrel pore. For these reasons, a hybrid barrel model (which appears to fit better with the data from Höhr et al. (2018)) could predominate in mitochondria, while a β 1-templated model (which fits better with the data from Doyle and Bernstein (2019)) may better explain the mechanism in bacteria. Both of these models could also incorporate lipid disordering and both mitochondria and bacteria may take different approaches for different substrates.

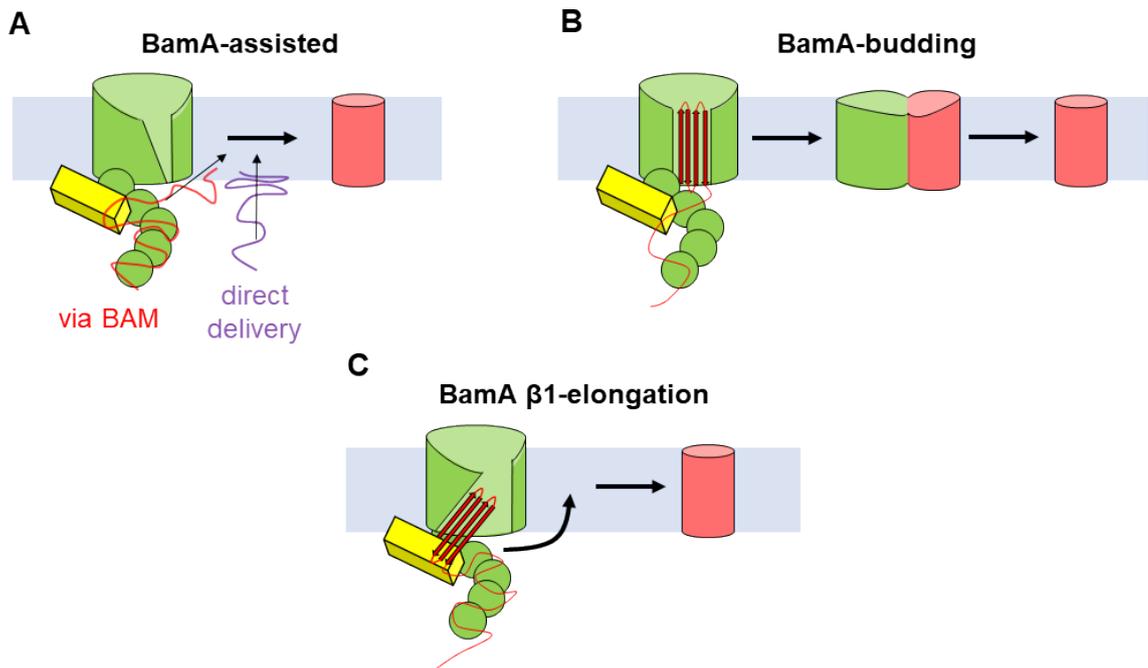


Fig. 1.12 Possible mechanisms of catalysis by BamA / the BAM complex. BamA is shown in green, BamD in yellow (other subunits omitted for clarity), OMP substrate in red. (A) In the BamA assisted model, the BamA barrel plays no direct role in OMP folding (e.g. through recognition or binding) but instead creates a region of disordered lipid in its vicinity which lowers the activation energy barrier to OMP folding imposed by acyl chain packing and lipid head groups (Noinaj et al., 2015). In theory an OMP can take two routes to this mechanism: via BAM (through recognition at the POTRA domains or BamD), or direct delivery to the disordered membrane by chaperones. (B) The BamA-budding model proposes that BamA binds to OMPs via β 1 and β 16 of BamA, with recognition of the OMP β -signal driving the initial β -strand to β -strand binding event. This doesn't preclude an initial recognition step mediated by POTRA domains or BamD preceding this. From here, an OMP is folded by inserting β -hairpins directly into a membrane-embedded barrel which slowly grows laterally ('blebs') from a nucleation point at the β 1- β 16 seam on BamA forming a 'super-barrel'. Once all strands are inserted the substrate β 1 and β -signal strand would have greater affinity for each other than BamA, allowing the OMP to dissociate and diffuse away (Noinaj et al., 2015). (C) The BamA β 1-elongation model represents a hybrid of (A) and (B) where recognition and binding of the OMP β -signal occurs on BamA β 1 but most of the folding occurs in the periplasm. Again, this is still compatible with initial recognition at POTRA domains or BamD. This model is based on observations that the 'open' state of BamA displays its β 1-strand at an angle off the axis of the membrane normal (i.e. pointed slightly towards the periplasm). OMP substrate β -strands 'template' from BamA β 1 in the open state and elongate outwards toward the periplasm (rather than into the membrane as in (B)). Variants on this model depend on whether the final step of insertion is passive due to the disrupted lipid interface around the BamA β 1- β 16 seam (Schiffirin et al., 2017a) or active and caused by β 1 of BamA 'swinging' upwards (i.e. becoming more upright, see (B)), thereby pulling the 'pre-folded' barrel into the membrane with it (Doyle and Bernstein, 2019). Figure adapted from Schiffirin et al., (2017a)

1.6 Conclusion and open questions in the field

Rapid progress has been made in the last 15 years in understanding the biogenesis of the bacterial outer membrane and particularly how outer membrane proteins are chaperoned, delivered, and folded into the OM. The slow trickle of structures of members of this pathway from the ribosome to the OM has been rapidly accelerated in the last 5 years with the solution of each individual component as well as numerous structures of the full complex, mutant, and stabilised conformations of BamA, and other Omp85 homologues. These structures (Figure 1.9), along with the design of genetic and molecular biology tools for assaying activity *in vitro* (plasmids for expressing the full BAM complex)(Roman-Hernandez et al., 2014) and *in vivo* (BamA depletion strains with complementing BamA plasmids)(Wu et al., 2005; Kim et al., 2007), have allowed researchers to design and test mechanistic hypothesis about how BAM functions to recognise substrates and accelerate OMP folding. Nonetheless, many questions remain. One interesting future line of study would be the structural characterisation of the many phenotypes described for mutants of BamA, BamD, and deletion of BamE. For example, mutations in POTRA5 (E373K and E373A) have been reported to alter the conformation of loop 6, as has deletion of BamE, and the mutations BamD_{R197L/S/H}. A genetic interaction between POTRA5 and loop 6 was found where BamA_{Q693P} rescued permeability defects in the presence of BamA_{E373A}. BamA_{F494L} has been isolated in a number of independent studies as a suppressor for lethal phenotypes caused by alterations of the conformation of BamA. A complementary biochemical and kinetic analysis of the folding of different classes of substrate which vary in their chaperone and BAM subunit dependence (e.g. LptD or FhuA, and LamB or other ‘simple’ OMPs with sizes <18-strands) would also tease apart whether these mutants cause lethal effects through a general kinetic effect (which would be recapitulated in an *in vitro* system) or an effect specific to the *in vivo* environment. Similar to the differences in severity for lid-locked and lateral-locked cysteine mutants of BamA assays *in vivo* and *in vitro*. The destabilisation of the lipid bilayer by BamA has been observed *in silico* in MD simulations but no direct experimental evidence for BAM’s role in destabilising lipids has been described and proving this would be important for putting the mechanisms of BAM function in context. A high resolution map of the pathway of an OMP through the BAM machinery, either through structural methods that capture stalled intermediates, or a higher resolution crosslinking approach than described in Section 1.5.2.2 would also resolve questions about how OMPs are delivered to and manipulated by BAM. Understanding the roles of larger OMP substrates in these folding processes and the differential roles of BAM subunits (BamBCDE) in this process, picking from the list of OMPs found in *E. coli*, will allow us to dissect all possible routes OMPs can take in this organism. Finally, with

this wealth of genetic and biochemical knowledge about BAM *in vitro* and *in vivo*, and the complexity of the network of cell envelope chaperones, stress regulator and monitors, and biogenesis machinery, it is time for more biophysical and structural research to be conducted in a fully *in vivo* environment.

1.7 Aims of this thesis

In this thesis, Chapter 3 discusses the development of new methods to probe the interaction between two proteins using crosslinking with mass spectrometry which allow rich datasets to map protein-protein interactions even when they are dynamic or transient. This is then applied to studying the interaction between OmpA with its cognate chaperones, Skp and SurA, as well as capturing its interactions with SurA and BAM during folding. In Chapter 4, the role of the lipid phase and lipid order in the folding of the model OMP, tOmpA, into DMPC liposomes is explored in order to dissect the contributions of SurA, BamA, and the lipoproteins of the BAM complex. Kinetic folding experiments are combined with direct measurements of lipid phase transitions for BamA- and BAM-embedded liposomes, as well as single-molecule FRET to probe the conformational ensemble of BAM in different lipid phases. Finally, Chapter 5 presents preliminary experiments towards super-resolution imaging of the organisation of BamA and its co-localisation with OmpA at cryogenic temperatures, working towards an understanding of how different members of the OMP biogenesis machinery interact *in vivo*.

Chapter 2

Materials and Methods

2.1 Materials

Deionised 18 M Ω water used in all methods.

Table 2.1 Materials

A	Supplier	Catalogue Number
Acetic acid, glacial	Fisher Scientific, Loughborough, UK	A/0400/PB17
Acrylamide 30 % (w/v):bis-acrylamide 0.8 % (w/v)	Severn Biotech, Kidderminster, UK	20-2100-10
Agar	Melford Laboratories, Suffolk, UK Fisher Scientific, Loughborough, UK	A20250-500.0 BP-1423-500
Agarose	Melford Laboratories, Suffolk, UK	MB1200
Alexa Fluor 488 C5 Maleimide	Fisher Scientific, Loughborough, UK	A10254
Alexa Fluor 594 C5 Maleimide	Fisher Scientific, Loughborough, UK	A10256
L-(+)-Arabinose	Sigma Life Sciences, MO, USA	A3256
Ammonium persulfate (APS)	Sigma Life Sciences, MO, USA	A7460
Ampicillin sodium salt	Formedium, Norfolk, UK	AMP25
B		
D-Tube dialyzer maxi, 12-14 kDa MWCO	EMD Millipore, MA, USA	71510-3
D-Tube dialyzer mini, 12-14 kDa MWCO	EMD Millipore, MA, USA	71505-3
BCA protein assay kit (Pierce)	Fisher Scientific, Loughborough, UK	23227
Benzamidine hydrochloride	Sigma Life Sciences, MO, USA	B6506

β -mercaptoethanol	Acros Organics, Geel, Belgium	125472500
	Sigma Life Sciences, MO, USA	M-7154
Bio-Beads SM-2 Adsorbents Resin	Bio-Rad Laboratories, CA, USA	1523920
Bromophenol blue	Sigma Life Sciences, MO, USA	B0126
C		
Carbenicillin disodium	Formedium, Norfolk, UK	CAR0025
Chloramphenicol	Sigma Life Sciences, MO, USA	C0378
Chloroform	Sigma Life Sciences, MO, USA	366927
Calcium chloride (CaCl ₂)	Acros Organics, Geel, Belgium	219171000
D		
Dithiothreitol (DTT)	Formedium, Norfolk, UK	DTT025
Dimethyl sulfoxide (DMSO)	Sigma Life Sciences, MO, USA	P841
	Fisher Scientific (Invitrogen), Loughborough, UK	D12345
Dimyristoyl-phosphatidylcholine (DMPC)	Avanti Polar Lipids, AL, USA	850345P
DNA ladders	New England Biolabs, MA, USA	N0552G
	Promega, WI, USA	G5711
n-dodecyl β -D-maltoside (DDM)	Anatrace, OH, USA	D310
DyLight 594 Maleimide	Fisher Scientific, Loughborough, UK	46608
E		
<i>E. coli</i> polar lipid extract	Avanti Polar Lipids, AL, USA	100600P
Ethidium bromide (EtBr)	Sigma Life Sciences, MO, USA	E-8751
Ethanol	Sigma Life Sciences, MO, USA	E/0650DF/17
Ethylenediaminetetraacetic acid (EDTA)	Acros Organics, Geel, Belgium	409930010
Mini, EDTA-free protease inhibitor cocktail tablets	Roche Applied Science	11836170001
G		
Glycerol	Fisher Scientific, Loughborough, UK	G/0650/17
Glycine	Fisher Scientific, Loughborough, UK	G/0800/60
Guanidine hydrochloride	Sigma Life Sciences, MO, USA	50950
H		
Hydrochloric acid (HCl)	Fisher Scientific, Loughborough, UK	H/1100/PB17
I		
Imidazole	Sigma Life Sciences, MO, USA	I202
Instant Blue Coomassie Blue Stain	Expedeon, CA, USA	ISB1LUK
Isopropanol	Honeywell Research Chemicals, Seelze, Germany	190764
Isopropyl β -D-1-thiogalactopyranoside (IPTG)	Formedium, Norfolk, UK	IPTG100
L		
Laurdan	Cayman Chemical, Ann Arbor, MI, USA	19706

LB Broth	Fisher Scientific, Loughborough, UK Merck KGaA, Darmstadt, Germany	1289-1650 1.10285.0500
M		
Methanol	Fisher Scientific, Loughborough, UK Sigma Life Sciences, MO, USA	M/4000/17 34860-1L-R
Molecular weight marker (Precision Plus Dual Xtra Standards)	Bio-Rad Laboratories, CA, USA	161-0377
N		
Nickel sepharose	GE Healthcare, Buckinghamshire, UK	
Nickel(II) sulfate heptahydrate	Fluorochem, Hadfield, UK	510236
Nickel nitrilotriacetic acid (Ni-NTA)	GE Healthcare, Buckinghamshire, UK	
P		
Phenylmethanesulfonyl fluoride (PMSF)	Sigma Life Sciences, MO, USA	P7762
0.1 µm polycarbonate membrane	Whatman Inc., NJ, USA	800309
Potassium chloride (KCl)	Fisher Scientific, Loughborough, UK	P/4200/60
Potassium hydroxide (KOH)	Fisher Scientific, Loughborough, UK	P/5600/53
Q		
Q5 Site-directed mutagenesis kit	New England Biolabs, MA, USA	E0554
S		
SnakeSkin dialysis tubing, 3.5K MWCO	Fisher Scientific, Loughborough, UK	68035
Sodium azide (NaN ₃)	Sigma Life Sciences, MO, USA	S-8032
Sodium chloride (NaCl)	Fisher Scientific, Loughborough, UK	S/3160/60
Sodium dodecyl sulphate (SDS)	Fisher Scientific, Loughborough, UK Severn Biotech, Kidderminster, UK	S/P530/53 20-4000-01
Sodium hydroxide (NaOH)	Sigma Life Sciences, MO, USA Fisher Scientific, Loughborough, UK	L4509 S/4920/60
Sodium succinate dibasic hexahydrate	Sigma Life Sciences, MO, USA	S2378
Syringe filter (nylon) (0.22 µm)	Camlab Ltd., Cambridge, UK	1181466
Syringe filter (PES) (0.22 µm & 0.45 µm)	Jet Biofil, Guangzhou, China	FPE-204-025, FPE-404-025
Sucrose	Fisher Scientific, Loughborough, UK	S/8600/53
Super Optimal Catabolite (SOC)	New England Biolabs, MA, USA	B90205
T		
Triton X-100	Sigma Life Sciences, MO, USA	X100-500

	Calbiochem, CA, USA	648463-50
Tris	Fisher Scientific, Loughborough, UK	BP152-1
Tetramethylethylenediamine (TEMED)	Sigma Life Sciences, MO, USA	T9281
Thiopropyl Sepharose 6B	GE Healthcare, Buckinghamshire, UK	17042001
Tris(2-carboxyethyl)phosphine (TCEP)	Sigma Life Sciences, MO, USA	C4706
Tris-tricine SDS running buffer 10X, cathode buffer, pH 8.3	Alfa Aesar, Heysham, UK	J60992
Tryptone	Fisher Scientific, Loughborough, UK	1285-1660
U		
Urea	MP biomedical, Loughborough, UK	04821527
	Fisher Scientific, Loughborough, UK	29700
V		
Vivaspin 20 centrifugal concentrators (3K, 5K, 10K, 30K, 50K, 100K MWCO)	Sartorius, Göttingen, Germany	VS2092, VS2012, VS2002, VS2022, VS2032, VS2042
Vivaspin 500 centrifugal concentrators (3K MWCO)	Sartorius, Göttingen, Germany	VS0192
W		
Wizard Plus SV Minipreps DNA purification systems	Promega, WI, USA	A1460
X		
Xylene cyanol FF	Sigma Life Sciences, MO, USA	X-4126
Y		
Yeast Extract	Melford Laboratories, Suffolk, UK	Y20025-2000.0
Z		
ZebaSpin Desalting Column 7K MWCO (0.5 ml)	Fisher Scientific, Loughborough, UK	89883
ZipTip Pipette Tips (with 0.6 µL C ₁₈)	Merck KGaA, Darmstadt, Germany	ZTC18S096

2.2 Bacterial strains

E. coli DH5 α derivative strain, NEB 5-alpha (F⁻ *fhuA2* Δ (*argF-lacZ*)U169 *phoA glnV44* Φ 80 Δ (*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*) purchased from NEB (Cat#: C2987H). *E. coli* BL21(DE3) strains (F⁻ *dcm ompT hsdS(rB- mB-)* *gal* λ (DE3)) originally purchased from Agilent (Cat#: 200132) and made competent in lab. *E. coli* JCM166

(BamA depletion strain) derived from MC4100 ($F^- [araD139]_{B/r} \Delta(argF-lac)169^* \lambda^- e14-flhD5301 \Delta(fruK-yeiR)725 (fruA25) \ddagger relA1 rpsL150(strR) rbsR22 \Delta(fimB-fimE)632(::IS1) deoC1 ara^{r/-} \Delta yaeT \Delta(\lambda att-lom)::bla P_{BAD} yaeT araC Ap^r$).

2.3 Plasmids and primers

E. coli transmembrane OmpA₁₋₁₇₁ (tOmpA), full-length *E. coli* OmpA₁₋₃₂₅ (FL-OmpA), and WT-BamA in pET11a vectors containing a T7 promoter under control of a lac operon were obtained from Karen Fleming, John Hopkins University. All subsequent mutant OmpA plasmids were derived from these (see Table 2.2). Skp with an N-terminal 6X His-tag in a pET28b vector was supplied by Sebastian Hiller, University of Basel. SurA with an N-terminal 6X His-tag in a pET28b vector was provided by Daniel Kahne, Harvard University. BAM complex (BamABCDE) with an 8X His-tag on BamE in a pTRC99a vector was obtained from Harris Bernstein, NIH Bethesda. BamA with an intact signal sequence and an 6X His-tag at the N-termini of the mature protein in a pZS21 plasmid was obtained from Tom Silhavy, Princeton University.

Table 2.2 List of plasmids used in this thesis and description of gene insert and mutations.

Plasmid name and vector	Description of sequence
OmpA-pET11a	FL-OmpA(WT)(22-346). Mature full length OmpA gene (1-325) with an initiating methionine N-terminal to position 1. Ampicillin resistance marker.
OmpA ₁₋₁₇₁ -pET11a	tOmpA(WT)(22-192). As above but with a deletion of the C-terminal periplasmic domain leaving only the 8-stranded N-terminal β -barrel domain (1-171 with initiating methionine). Ampicillin resistance marker.
OmpA(S290C+S302C)-pET11a	FL-OmpA(no cys). OmpA-pET11a with a substitution of cysteine-290 to serine and a substitution of cysteine-302 to serine.
OmpA(M0_A1insC+S290C+S302C)-pET11a	FL-OmpA(N _{cys}). OmpA(S290C+S302C)-pET11a with a cysteine insertion between the initiating methionine.
OmpA(W7C+S290C+S302C)-pET11a	FL-OmpA(W7C). OmpA(S290C+S302C)-pET11a with a substitution of tryptophan-7 to cysteine.
OmpA(T144C+S290C+S302C)-pET11a	FL-OmpA(W7C). OmpA(S290C+S302C)-pET11a with a substitution of threonine-144 to cysteine.

Plasmid name and vector	Description of sequence
BamA-pET11a	Full-length mature BamA(22-810) without its signal sequence and the first alanine of the mature sequence, and an initiating methionine N-terminal to the first position. The presence of A21 was found to inhibit expression (B. Schiffrin, unpublished results). Ampicillin resistance marker.
HT-Skp-pET28b	Full-length mature Skp(21-161) with an N-terminal 6X His-tag and a thrombin cleavage site-containing linker region. Without signal sequence, for cytoplasmic expression. Kanamycin resistance marker.
HT-SurA-pET28b	Full length mature SurA(21-428) with an N-terminal 6X His-tag. Without signal sequence, for cytoplasmic expression. Kanamycin resistance marker.
BAM complex-pTRC99a (pJH114)	Each component of the BAM complex including signal sequences in tandem on a single plasmid. For overexpression into the outer membrane. Ampicillin resistance marker.
BAM(BamA[C690S+C700S])-pJH114	Full BAM complex with cysteines in loop 6 of BamA mutated to serines. BAM <i>no cys</i> background.
BAM(BamA[R127C+N520C+C690S+C700S])-pJH114	pJH114 plasmid with R127C and N520C cys mutations added on BamA into a <i>no cys</i> background (BamA[C690S/C700S]). R127 is in POTRA2, N520 is on turn 3.
HT-BamA-pZS21	N-terminally His-tagged mature BamA with intact signal sequence for <i>in vivo</i> expression. 6X His-tag (HHHHHHAA) was inserted between A21 and E22. Can complement BamA knockdown in the JCM166 cell line. Kanamycin resistance marker.

Table 2.3 Master protein sequences encoded for by each plasmid (not including point mutants or fusions, which are variants on these).

Protein	Sequence
tOmpA	MAPKDNTWYTGAKLGWSQYHDTGFINNNGPTHENQLGAGAFGGYQVNPYV GFEMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDIYTRLGG MVWRADTKSNVYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGD AHTIGTRPDNGMLSLGVSYRFG
OmpA	MAPKDNTWYTGAKLGWSQYHDTGFINNNGPTHENQLGAGAFGGYQVNPYV GFEMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDDLDIYTRLGG MVWRADTKSNVYGKNHDTGVSPVFAGGVEYAITPEIATRLEYQWTNNIGD AHTIGTRPDNGMLSLGVSYRFGQGEAAPVVAPAPAPAPEVQTKHFTLKSD VLFNFKATLKPEGQAALDQLYSQLSNLDPKDGSVVVLGYTDRIIGSDAYN QGLSERRAQSVDYLLISKGIPADKISARGMGESNPVTGNTCDNVKQRAAL IDCLAPDRRVEIEVKGIKDVVTQPQA

Protein	Sequence
HT-Skp	MGSSHHHHHHSSGLVPRGSHMADKIAIVNMGSLFQQVAQKTGVSNTLENE FKGRASELQRMETDLQAKMKKLQSMKAGSDRTKLEKDVMAQRQTFAQKAQ AFEQDRARRSNEERGKLVTRIQTAVKSVANSQDIDLVV DANAVAYNSSDV KDITADVLKQVK
HT-SurA	MGSSHHHHHHSSGLVPRGSHMAPQVVDKVA AVVNNGVVLESDVDGLMQSV KLNAAQARQQLPDDATLRHQIMERLIMDQIILQMGQKMGVKISDEQLDQA IANIAKQNNMTLDQMRSLAYDGLNYNTYRNQIRKEMIISEVRNNEVRRR ITILPQEVESLAQQVGNQNDASTELNLSHILIPLENPTSDQVNEAESQA RAIVDQARNGADFGKLAIAHSADQQALNGGQMGWGRIQELPGIFAQALST AKKGDIVGPIRSGVGFHILKVNDLRGESKNI SVTEVHARHILLKPSIMT DEQARVKLEQIAADIKSGKTTFAAAAKEFSQDPGSANQGGDLGWATPDIF DPAFRDALTRLNKGQMSAPVHSSFGWHLIELLDTRNVDKTDAAQKDRAYR MLMNRKFSEEASWMQEQRASAYVKILSN
BamA (pET11a)	MAEGFVVKDIHFEGLRVAVGAALLSMPVRTGDTVNDEDISNTIRALFAT GNFEDVRVLRDGD TLLVQVKERPTIASITFSGNKSVKDDMLKQNLASGV RVGESLDRTTIADIEKGLEDFYYSVGKYSASVKAVVTPLPRNRVDLKLVF QEGVSAEIQQINIVGNHAFTTDELISHFQLRDEVPWWNVVGD RKYQKQKL AGDLETLSYLLDRGYARFNIDSTQVSLTPDKKGIYVTVNITEGDQYKLS GVEVSGNLAGHSAEIEQLTKIEPGELYNGTKVTKMEDDIKLLGRYGYAY PRVQSMPEINDADKTVKLRVNV DAGRNFYVRKIRFEGNDTSKDAVLRREM RQMEGAWLGSDLVDQKERLNRLGFFETVD TDTQRVPGSPDQVDVVYKVK ERNTGSFNFGIGYGTESGVSFQAGVQQDNWLGTGYAVGINGTKNDYQTYA ELSVTNPYFTVDGVSLGGRLFYNDFQADDADLSDYTNKSYGTDVTLGFP I NEYNSLRAGLGYVHNSLSNMQPQVAMWRYLYSMGEHPSTSDQDNSFKTDD FTFNYGWTYNKLD RGYFPTDGSRVNLTGKVTIPGSDNEYKVTLDTATYV PIDDDHKWVVLGRTRWGYGDGLGGKEMPFYENFYAGGSSTVRGFQSNTIG PKAVYFPHQASNYDPDYDYECATQDGAKDLCKSDDAVGGNAMAVASLEFI TPTPFI SDKYANSVRTSFFWDMGTVWDTNWDSSQYSGYPDYSDPSNIRMS AGIALQWMSPLGPLVFSYAQPFFKKYDGDKAEQFQFNIGKTW

Protein	Sequence
HT-BamA (pZS21)	MAMKKLLIASLLFSSATVYGAHHHHHAAEGFVVKDIHFEGLQRVAVGAA LLSMPVRTGDTVNDEDISNTRALFATGNFEDVRVLRDGDITLLVQVKERP TIASITFSGNKSVKDDMLKQNLASGVRVGEISLDRTTIADIEKGLEDFYY SVGKYSASVKAVVTP LPRNRVDLKLVFQEGVSAEIQQINIVGNHAF TTDE LISHFQLRDEVPWWNVVGD RKYQKQKLAGDLET LRSYYLDRGYARFNIDS TQVSLTPDKKGIYVTVNI TEGDQYKLSGVEVSGNLAGHSAEIEQLTKIEP GELYNGTKVTKMEDDIKKLLGRYGYAYPRVQSMPEINDADKTVKLRVNVD AGNRFYVRKIRFEGNDTSKDAVLRREMRQMEGAWLGSDLVDQGKERLNRL GFFETVDTDTQRVPGSPDQVDVYKVKERN TG SFNFGIGYGTESGVSFQA GVQQDNWLGTYAVGINGTKNDYQTYAELSVTNPYFTVDGVS LGGRLFYN DFQADDADLSDYTNKSYGTDVTLGFP INEYNSLRAGLGYVHNSLSNMQPQ VAMWRYLYSMGEHPSTSDQDNSFKTDDFTFNYGWTYNKLD RGYFP TDGSR VNL TGKVTIPGSDNEYKVTLDTATYVP IDDDHKWVVLGRTRWGYGDGLG GKEMPFYENFYAGGSSTVRGFQSNTIGPKAVYFPHQASNYDPDYDECAT QDGAKDLCKSDDAVGGNAMAVASLEF ITP TPFISDKYANSVRTSFFWDMG TVWDTNWDSSQYSGYPDYSDPSNIRMSAGIALQWMSPLGPLVFSYAQPFFK KYDGDKAEQFQFNIGKTW
BAM complex	BamA: MAMKKLLIASLLFSSATVYGAEGFVVKDIHFEGLQRVAVGAALLS MPVRTGDTVNDEDISNTRALFATGNFEDVRVLRDGDITLLVQVKERP TIA SITFSGNKSVKDDMLKQNLASGVRVGEISLDRTTIADIEKGLEDFYYSVG KYSASVKAVVTP LPRNRVDLKLVFQEGVSAEIQQINIVGNHAF TTDELIS HFQLRDEVPWWNVVGD RKYQKQKLAGDLET LRSYYLDRGYARFNIDSTQV SLTPDKKGIYVTVNI TEGDQYKLSGVEVSGNLAGHSAEIEQLTKIEP GEL YNGTKVTKMEDDIKKLLGRYGYAYPRVQSMPEINDADKTVKLRVNVDAGN RFYVRKIRFEGNDTSKDAVLRREMRQMEGAWLGSDLVDQGKERLNRLGFF ETVDTDTQRVPGSPDQVDVYKVKERN TG SFNFGIGYGTESGVSFQAGVQ QDNWLGTYAVGINGTKNDYQTYAELSVTNPYFTVDGVS LGGRLFYNDFQ ADDADLSDYTNKSYGTDVTLGFP INEYNSLRAGLGYVHNSLSNMQPQVAM WRYLYSMGEHPSTSDQDNSFKTDDFTFNYGWTYNKLD RGYFP TDGSRVNL TGKVTIPGSDNEYKVTLDTATYVP IDDDHKWVVLGRTRWGYGDGLGGKE MPFYENFYAGGSSTVRGFQSNTIGPKAVYFPHQASNYDPDYDECATQDG AKDLCKSDDAVGGNAMAVASLEF ITP TPFISDKYANSVRTSFFWDMGTW DTNWDSSQYSGYPDYSDPSNIRMSAGIALQWMSPLGPLVFSYAQPFFKYD GDKAEQFQFNIGKTW BamB: MQLRKLLLPGLLSVTL LSGCSLFNSEEDVVKMSPLPTVENQFTPT TAWSTSVGSGIGNFYSNLHPALADNVVYAADRAGLVKALNADDGKEIWSV SLAEKDGWF SKEPALLSGGVTVSGGHVYIGSEKAQVYALNTSDGTVAWQT KVAGEALS R PVSDGLVLIHTSNGQLQALNEADGAVKWTVNLDMP SLSLR GESAPTTAFGAAVVGDN GRVSAVLMEQGQMIWQQRISQATGSTEIDRLS DVD TTPVVVNGVVFALAYNGNLTALDL RSGQIMWKRELGSVNDF IVDG NR IYLVDQND RVMALTIDGGVTLWTQSDLLHRLLTSPVLYNGNLVVG DSEGY LHWINVEDGRFVAQQKVDSSGFQTEPVAADGKLLIQAKDGTVYSITR

Protein	Sequence
BamC:	MAYSVQKSRLAKVAGVSLVLLLAACSSDSRYKRQVSGDEAYLEAA PLAELHAPAGMILPVTSGDYAIPVTNGSGAVGKALDIRPPAQPLALVSGA RTQFTGDTASLLVENGRGNTLWPQVSVLQAKNYTITQRDDAGQTLTTDW VQWNRLDEDEQYRGRYQISVKPQGYQQAVTVKLLNLEQAGKPVADAASMQ RYSTEMMNVISAGLDKSATDAANAAQNRASSTMDVQSAADDTGLPMLVVR GPFNVVWQRLPAALEKVGKMTDSTRSQGNMAVTKPLSDSDWQELGASD PGLASGDYKLVQVGLDNRSSLQFIDPKGHTLTQSQNDALVAVFQAASF BamD: MTRMKYLVA AATLSLFLAGCSGSKEEVPDNPNEIYATAQQKLQD GNWRQAITQLEALDNRYPFPGPYSQQVQLDLIYAYYKNADLPLAQAAIDRF IRLNPTHPNIDYVMYMRGLTNMALDDSSALQGGFGVDRSDRDPQHARAAFS DFSKLVRGYPNSQYTTDATKRLVFLKDRDLAKYEYSVAEYYTERGAWVAVV NRVEGMLRDYPTQATRDALPLMENAYRQMOMNAQAEKVAKIIAANSNT BamE: MRCKTLTAAAVLLMLTAGCSTLERVVYRPDINQGNILTANDVSK IRVGMTQQQVAYALGTPLMSDPFGTNTWFYVFRQQPGHEGVTQQTLLTLTF NSSGVLTNIDNKPALSGNGGHHHHHHHH

Table 2.4 List of primers used for mutagenesis and cloning. All primers were synthesized by Eurofins Genomics as unmodified DNA oligonucleotides and purified by high purity salt-free purification (HPSF).

Primer name	Sequence	Details
OmpA-M0_A linsC_F	TGCGCTCCGAAAGATAACACC	N _{cys} cysteine insertion (forward)
OmpA-M0_A linsC_R	CATATGTATATCTCCTTCTTAA AG	N _{cys} cysteine insertion (reverse)
OmpA-W7C_F	ATAACACCTGcTACACTGGTGC	W7C cysteine substitution (forward)
OmpA-W7C_R	CTTTCGGAGCCATATGTATATC	W7C cysteine substitution (reverse)
OmpA-T144C_F	ATACCAGTGGTGCAACAACATC GGTG	T144C cysteine substitution (forward)
OmpA-T144C_R	TCCAGACGGGTAGCGATT	T144C cysteine substitution (reverse)
OmpA-C290S+C302S_F	TGCTGCACTGATCGACTCCCTG GCTCCGGATCGTGC	C290S and C302S cysteine substitutions (forward)
OmpA-C290S+C302S_R	CGCTGTTTCACGTTGTCAGAGG TGTTGCCAGTAACCGG	C290S and C302S cysteine substitutions (reverse)
BAM(BamA- C690S+C700S)_F	CGCGAAAGACCTGAGCAAATCG GATGATGCTGTAG	C690S and C700S cysteine substitutions to make Cys-free BAM (forward)
BAM(BamA- C690S+C700S)_R	CCGTCCTGAGTCGCGCTTTTCGT AATCATAGTCCGG	C690S and C700S cysteine substitutions to make Cys-free BAM (reverse)
BAM(BamA-R127C)_F	ATCCCTCGATTGCACCACCATT GC	R127C cysteine substitution in BamA (forward)

Primer name	Sequence	Details
BAM(BamA-R127C)_R	TCGCCACACGCACACCA	R127C cysteine substitution in BamA (reverse)
BAM(BamA-N520C)_F	CTTCCCGATTGCGAATATAAC TCGCTG	N520C cysteine substitution in BamA (forward)
BAM(BamA-N520C)_R	CCCAACGTACGTCTGTA	N520C cysteine substitution in BamA (reverse)

2.4 Molecular biology (preparatory) methods

2.4.1 Culture medium

LB medium was prepared at a concentration of 25 g.L⁻¹ (10 g tryptone, 5 g yeast extract, 10 g NaCl) for growth of all proteins except the BAM complex, and for *in vivo* experiments with the JCM166 strain. BAM complex was grown in 2xTY medium (16 g bacto-tryptone, 10 g yeast extract, 5 g NaCl, made up to 1 L). All media for protein production was autoclaved before use and used within two days. JCM166 strains (both untransformed, and transformed with pZ series vectors) were grown in M9 minimal medium with added supplements. All media for *in vivo* growth experiments were handled exclusively under flame or in microbiological safety cabinets after sterilisation. 1L of 10X M9 salts stock (67.8 g Na₂HPO₄, 30 g KH₂PO₄, 5 g NaCl, 10 g NH₄Cl) was made with deionised water and autoclaved. 1X solutions were made by 1:10 dilution with sterile deionised water to give a final salts concentration of 33.7 mM Na₂HPO₄, 22.0 mM KH₂PO₄, 8.55 mM NaCl, 9.35 mM NH₄Cl. This was supplemented by 0.1 mM CaCl₂ (from 0.1 M autoclaved stock), 2 mM MgSO₄ (from 1 M autoclaved stock), 0.002 % (w/v) thiamine (from 0.2 % 0.22 µm filter-sterilised stock), 0.1 mM FeSO₄ (from 0.1 M 0.22 µm filter-sterilised stock), 0.05 % (w/v) casamino acids (from 10 % autoclaved stock). For untransformed JCM166 growth 0.2 % (w/v) L-arabinose (from 10 % 0.22 µm filter-sterilised stock) and 100 µg.mL⁻¹ carbenicillin (from 1000X 0.22 µm filter-sterilised stock) were also added. For JCM166 transformed with BamA or OmpA variants in pZ-series vectors, 0.4 % (w/v) D-glucose (from 20 % 0.22 µm filter-sterilised stock) and either 50 µg.mL⁻¹ kanamycin, 50 µg.mL⁻¹ chloramphenicol, or both were added, dependent on the resistance marker. FeSO₄ stocks were made fresh daily to reduce the load of oxidised Fe(III) oxide and Fe(III) sulfate formed by reaction with oxygen in the air and dissolved in the aqueous stock. Thiamine, arabinose, and casamino stocks were made fresh weekly. Salt stocks, casamino acids and

glucose were stored at room temperature, iron sulfate and thiamine were stored at 4 °C, antibiotics and arabinose were stored at -20 °C.

Agar plates were prepared with 15 g.L⁻¹ agar and 25 g.L⁻¹ LB medium. LB-agar solutions were sterilised by autoclaving at 120 °C for 20 minutes and allowed to cool below 60 °C before addition of the relevant antibiotics (final concentrations: carbenicillin, 100 µg.mL⁻¹; kanamycin, 50 µg.mL⁻¹; and chloramphenicol, 25 µg.mL⁻¹) from 0.22 µm filter-sterilised 1000x stocks. For growth of JCM166 cells, arabinose or glucose was also added as described above. Around 25 mL of LB-agar solution was poured into petri dishes under sterile conditions and allowed to cool to prevent condensation build-up before immediate use or storage at 4 °C.

2.4.2 Preparation of competent cells

An *E. coli* host strain (e.g. DH5α, BL21(DE3)pLysS, JCM166) was streaked out on an antibiotic-free LB agar plate and was grown overnight at 37 °C. A single colony was picked and grown overnight in LB medium at 37 °C, 200 rpm in a 50 mL Falcon tube. Alternatively, a 50 µL aliquot from stocks of a non-competent strain in 25 % (v/v) glycerol were used to inoculate 10 mL of LB and grown as above. 5 mL of this culture was added to 100 mL of LB medium and grown until the optical density at 600 nm (OD₆₀₀) was 0.45 at 37 °C, 200 rpm. The cells were then harvested by centrifugation for 10 minutes at 4,000 × g in a JLA-16.250 rotor (Beckman Coulter) pre-chilled to 4 °C. The supernatant was then discarded and the pellet was resuspended in 10 mL pre-chilled, 0.22 µm filter-sterilised 100 mM CaCl₂ before incubating on ice for 15 minutes. Cells were then pelleted by centrifugation for 10 minutes at 3,220 × g in an Eppendorf 5810R centrifuge with an A-4-62 bucket rotor pre-chilled to 4 °C. The supernatant was discarded and the pellet was resuspended in 2 mL of pre-chilled, 0.22 µm filter-sterilised 100 mM CaCl₂, 30 % (v/v) glycerol. 100 µL aliquots were added to 2 mL transformation tubes pre-chilled on dry ice and stored at -80 °C.

2.4.3 Transformation of *E. coli* strains

E. coli NEB 5-alpha (C2987H) competent cells were thawed on ice and 5 µL of 1–100 ng plasmid DNA or the product of the KLD mix from Q5 site-directed mutagenesis (Section 2.4.4) was added and mixed by flicking 5 times. This mixture was placed on ice for 30 minutes, heat shocked at 42 °C for 30 seconds and then returned to ice for 5 minutes.

950 μL of room temperature super optimal broth with catabolite repression (SOC) was added [SOC: 2 % (w/v) vegetable peptone, 0.5 % (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl_2 , 10 mM MgSO_4 , 20 mM glucose]. This was incubated at 37 °C, 250 rpm, for 60 minutes. BL21(DE3) strains were defrosted on ice and 50 μL was transferred into a transformation tube. 2 μL of 25 $\text{ng}\cdot\mu\text{L}^{-1}$ plasmid DNA added and flicked gently 4 times. Cells were placed on ice for 30 minutes and subsequently heat-shocked at 42 °C for 45 seconds before returning to ice for 5 minutes. The process then continued as for NEB 5-alpha cells. JCM166 strains were prepared as for NEB 5-alpha except 700 μL of room temperature LB supplemented with arabinose was added to JCM166 cells after the 5 minute ice incubation. For NEB 5-alpha and BL21(DE3) cells 50 μL was then plated out on LB agar plates containing the appropriate antibiotic selection marker, for JCM166 the cells were pelleted and then resuspended in 50 μL before plating out. Plates were then grown at 37 °C overnight.

2.4.4 Site-directed mutagenesis

Point mutations were introduced into the OmpA or BAM plasmid by exponential amplification of double stranded plasmid DNA through polymerase chain-reaction (PCR) using non-overlapping primers designed through the NEBaseChanger tool provided by New England Biolabs (<http://nebasechanger.neb.com/>) (see also Table 2.4). A Q5 Site-Directed Mutagenesis kit was used according to the manufacturer's guidelines with 25 $\text{ng}\cdot\mu\text{L}^{-1}$ of OmpA or BAM plasmid added. The product of the PCR mutagenesis reaction was run on an agarose gel to check for successful amplification (indicated by the presence or absence of a band around 6 or 7 kb for OmpA, 10 kb for BAM). The PCR product was then incubated with a kinase, ligase, DpnI (KLD) mix according to manufacturer's guidelines. 5 μL of this mixture was then used to transform NEB 5-alpha competent *E. coli* cells. The correct mutation was confirmed through sequencing of the purified plasmid DNA (full OmpA insert, or full BamA gene in the BAM complex).

2.4.5 Plasmid DNA purification

Plasmid DNA was purified from a 10 mL overnight culture in LB medium using a Wizard Plus SV Minipreps DNA Purification system according to manufacturer's guidelines. DNA concentration was then quantified on a NanoDrop 2000 UV-Vis spectrophotometer (Thermo Fisher Scientific) by measuring its absorbance at 260 nm and the DNA then stored at -20 °C.

2.4.6 *In vivo* growth and expression of BamA and OmpA variants in the JCM166 cell line

The JCM166 cell line (Wu et al., 2005) contains a deletion of the chromosomal BamA using recombineering methods (Ellis et al., 2001) and an insertion of a cassette containing an ampicillin resistance marker and BamA under control of an arabinose-inducible promoter, P_{BAD}. This cassette was installed at the λ attachment site using a lambda-phage based specialized vector (Boyd et al., 2000). This strain has a genetic background that makes it unable to catabolize arabinose allowing it to be grown in/on arabinose-containing media to stably induce expression. N-terminally His-tagged BamA cloned into a pZS21 vector can complement BamA knockdown in this strain. The pZS21 vector has a pSC101 origin of replication giving ~5 plasmids per cell (Thompson et al., 2018), a kanamycin resistance marker allowing stable complementary selection of both the JCM166 strain and the plasmid, and the gene insert is under control of the P_{LtetO-1} promoter which allows constitutive expression of BamA. Untransformed JCM166 cells were streaked onto LB agar supplemented with carbenicillin and arabinose to induce expression of chromosomal BamA, and grown overnight at 37 °C. A colony was picked, grown overnight at 37 °C with 220 rpm shaking in a starter culture of LB supplemented with carbenicillin and arabinose, and subsequently made chemically competent as described in Section 2.4.2. JCM166 was transformed with pZ series vectors containing HT-BamA or OmpA transmembrane domain, or variants where a fluorescent protein or HaloTag had been fused to the N-terminus (for BamA) or C-terminus (for OmpA). These were grown on medium supplemented with just the vector's resistance marker (no carbenicillin) and arabinose. Before each imaging experiment (Section 2.6.6), colonies were picked from plates containing the transformants and inoculated into 5–10 mL LB + glucose + kanamycin (and/or chloramphenicol) starter cultures. These were grown in 50 ml Falcon tubes overnight at 37 °C, 200 rpm. The following day 50 μ L of this culture was inoculated 1:100 into M9 medium with supplements and grown at 37 °C, 200 rpm until reaching an OD₆₀₀ of 0.6–0.8 (approx. 2–3 hours). 1 mL of this was then spun down at 3,000 \times g, 4 °C, 3 min, the supernatant discarded, and resuspended in 1 mL sterile Dulbecco's phosphate buffered saline (DPBS) [137 mM NaCl, 2.7 mM KCl, 10 mM phosphate pH 7.4]. This was repeated 3X with the final resuspension in 100 μ L to give a high cell density (and account for minor losses from cells which remained in the supernatant). These samples were then kept on ice and immediately used for preparation of imaging samples for cryoSTORM super-resolution microscopy. For experiments with HaloTag-BamA the protocol was slightly modified. After an overnight grow, the cells were inoculated into new growth medium as described above but at an OD₆₀₀ of 0.3 2 \times 1 ml was spun down at 3,000 \times g, 4 °C, 3 min, in 1.5 ml Eppendorf

tubes and then resuspended in 300 μL of PBS, 1 % DMSO, 2 μM chloroalkane-derivatized (HaloTag ligand, HTL) *Janelia Fluor 549* (HTL-JF549) or JF646 (PBS labelling buffer made from 200 μM stock of HTL-JF dye in DMSO). This was then incubated at 37 °C, 200 rpm, for 30 min to allow labelling. Cells were washed 4X with 500 μL of PBS (pelleting between each wash at 3,000 \times g, 4 °C, 3 min, and discarding the supernatant) before a final resuspension in 500 μL M9 medium without FeSO_4 or antibiotics. The cells were grown for a further 15–30 min to allow recovery from labelling. These cells were spun at 3,000 \times g, 4 °C, 3 min, the supernatant discarded, and resuspended in 50–75 μL of PBS ready for preparation of imaging samples.

2.4.7 Agarose gel electrophoresis

1 % or 1.5 % (w/v) agarose was added to a final volume of 150 mL 1X TAE buffer from a 50X stock (2 M Tris base, 1 M acetic acid, 50 mM EDTA pH 8.0) and heated in a conical flask until the agarose had completely dissolved. Once the solution had cooled to below 60 °C, 15 μL of 10 $\text{mg}\cdot\text{mL}^{-1}$ ethidium bromide was added to give a final concentration of 1 $\mu\text{g}\cdot\text{mL}^{-1}$. This was mixed and poured into a 12 \times 15 cm gel cast with lane combs in. 5 or 10 μL of DNA sample was added to 1 or 2 μL , respectively, of 6X DNA loading buffer (0.25 % (w/v) bromophenol blue, 0.25 % (w/v) xylene cyanol, 40 % (w/v) sucrose) and mixed on a piece of parafilm. 5 or 10 μL of each DNA sample and 7 μL of 1 kb DNA ladder was loaded onto the gels. The gels were ran in 1X TAE buffer at 100 V until the bands were well resolved and subsequently imaged in an INGenius UV transilluminator (Syngene).

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

Tris-tricine buffered SDS-PAGE gels were set in 8 \times 10 cm casts according to Table 2.5 below (for two gels). For low-SDS gels and time-resolved gels, see the appropriate subsection. For all other gels, samples were then diluted 3:4 into 4X loading buffer (100 mM Tris-Cl pH 6.8, 200 mM DTT, 4 % (w/v) SDS, 0.2 % (w/v) bromophenol blue, 20 % (w/v) glycerol) and boiled for 10 minutes, spun in a benchtop centrifuge, and then mixed thoroughly before loading 15 μL of sample into each lane. A protein standards marker was loaded in the first lane of the gel. The inner reservoir was filled with 1X cathode buffer (100 mM Tris-Cl, 100 mM tricine, 0.1 % (w/v) SDS, pH 8.25) from a 10X

Table 2.5 Recipe for regular Tris-tricine SDS-PAGE gels used in this thesis.

Solution component	Resolving gel	Stacking gel
	Volume added (ml)	Volume added (ml)
30 % w/v acrylamide:0.8 % w/v <i>bis</i> -acrylamide	7.5	0.83
3 M Tris-Cl, 0.3 % (w/v) SDS pH 8.45	5	1.55
H ₂ O	0.44	3.72
Glycerol	2	-
10 % (w/v) ammonium persulfate	0.4	0.2
TEMED	0.01	0.005

stock, and the outer reservoir with 1X anode buffer (200 mM Tris-Cl, pH 8.9) from a 10X stock. Gels were run at a constant current of 35 mA until the dye front passed the stacking gel. The current was then adjusted to 75 mA until the dye front reached the bottom of the gel. Gels were then stained for 15 - 60 minutes using InstantBlue, washed, and imaged under a white light transilluminator.

2.4.9 Low SDS-PAGE gel

As above for standard SDS-PAGE gel but with the following modifications: 0.1 % (w/v) SDS final in loading buffer (vs. 1 % (w/v) SDS), no SDS in the gel (3 M Tris-Cl pH 8.45 only), run at 4 °C at a low current (10–30 mA) overnight or across 4–8 hours, respectively.

2.4.10 Production of large unilamellar liposomes (LUVs)

1,2-diundecanoyl-sn-glycero-3-phosphatidylcholine (DUPC) and 1,2-dimyristoyl-sn-glycero-3-phosphatidylcholine (DMPC) were used to make large unilamellar vesicles. Lipids were dissolved from powdered form in 80:20 (v/v) HPLC grade chloroform:methanol in Pyrex glass test tubes. 25 mg.mL⁻¹ stocks were stored in sealed glass bottles at -20 °C until use. To make up liposomes, an appropriate volume of lipid stock was measured out using a 1 mL gas-tight glass syringe (Hamilton). Solvent was then removed by drying under a gentle stream of N₂ while agitating vigorously in a 42 °C water bath. This was followed by further drying in a desiccator under high vacuum for > 3 h. The resulting thin lipid film was resuspended above its transition temperature in buffer (50 mM glycine pH 9.5 at room temperature for DUPC; 150 mM NaCl, 20 mM Tris-Cl pH 8.0 at 37 °C for DMPC) to give a 40 mM lipid suspension, vortexed briefly and left to stand at room temperature (for DUPC) or 37 °C (for DMPC) for 30 min to hydrate, before vortexing again. The

large multilamellar vesicles that formed were disrupted by five freeze-thaw cycles. 100 nm Large Unilamellar Vesicles (LUVs) were prepared by extruding the lipid mixtures >13 times through a 0.1 μm polycarbonate membrane using a mini-extruder (Avanti, Alabaster, AL, USA). For DMPC, extrusion was performed at 37 °C using a pre-heated extruder in a metal block. All LUVs were stored at 4 °C immediately following extrusion and used within a week.

2.4.11 Protein expression and purification from inclusion bodies (BamA, OmpA, and tOmpA)

5 mL of overnight culture of *E. coli* BL21(DE3) or BL21(DE3)pLysS cells transformed with a pET11a plasmid vector containing a lactose-inducible copy of the gene of interest under control of the T7 promoter were added to 500 mL of pre-warmed LB medium in 2 L baffled flasks and incubated at 200 rpm, 37 °C until the optical density at 600 nm (OD₆₀₀) reached 0.6. Protein expression was induced by addition of IPTG to a final concentration of 1 mM from a 1000X stock in H₂O and cultured for a further 3.5–4 hours. The cells were then pelleted by centrifugation at 5,739 \times g, 4 °C in a JLA-8.1000 rotor and the pellet was transferred to a 50 mL Falcon tube and stored at -20 °C overnight. The pellet was thawed at room temperature and thoroughly resuspended in 20 mL of lysis reagent 1 (50 mM Tris-Cl pH 8.0, 1 mM PMSF, 2 mM benzamidine hydrochloride, 5 mM EDTA). This solution was then sonicated at 44 % amplitude on a Vibra-Cell VCX-130PB (Sonics) with a 6 mm diameter tip (Cat#: 630-0435) on a cycle of 6x 1 minute on, 1 minute wait, and kept on ice at all times. The insoluble fraction was pelleted at 25,000 \times g, 4 °C on a JA-25.50 rotor (Beckman Coulter) for 30 minutes and the supernatant was discarded. The pellet was thoroughly resuspended in 20 mL lysis reagent 2 (50 mM Tris-Cl pH 8.0, 2 % (w/v) Triton X-100) and incubated for 1 hour at room temperature with rocking. The insoluble fraction was pelleted as described above, the supernatant was discarded and the pellet was thoroughly resuspended with 20 mL 50 mM Tris-Cl pH 8.0 and mixed on a rocker at room temperature for 1 hour. This was repeated once more to ensure the removal of all detergent and the inclusion bodies were then pelleted as above, the supernatant was discarded and the pellet was stored at -20 °C.

The inclusion body pellet was solubilised and thoroughly resuspended in 10 mL of solubilisation buffer (double-filtered 6 M guanidine hydrochloride [0.45 μm then 0.22 μm , as the raw powder contains many particulate contaminants which may interfere with downstream spectroscopic measurements], 25 mM Tris-Cl, \pm 1 mM TCEP pH 8.0 – dependent on if protein contains Cys residues). The insoluble fraction was pelleted at 25,000 \times g,

Table 2.6 Properties proteins used in *in vitro* studies in this thesis. Isoelectric point calculated using EMBOSS Pepstats (https://www.ebi.ac.uk/Tools/seqstats/emboss_pepstats/), value in brackets calculated using ExPASy ProtParam (<https://web.expasy.org/protparam/>). * = BAM complex molecular weight not including the mass of lipid anchors on the lipoproteins BamBCDE.

Protein	Length (aa)	Molecular weight (Da)	Reduced extinction coefficient, ϵ ($M^{-1} cm^{-1}$)	Isoelectric point, pI
tOmpA	172	18,874.9	46,870	6.08 (5.74)
FL-OmpA	326	35,303.4	52,830	5.69 (5.59)
FL-OmpA(C290S/C302S)	326	35,271.4	52,830	5.69 (5.59)
FL-OmpA(no cys) (N_{Cys})	328	35,477.6	52,830	5.69 (5.59)
FL-OmpA(no cys)(W7C)	327	35,390.6	52,830	5.69 (5.59)
FL-OmpA(no cys)(T144C)	326	35,190.3	47,330	5.69 (5.59)
BamA	790	88,486.2	140,040	4.65 (4.87)
BAM-HT	1813	199,963.8*	291,650	4.80 (5.01)
BAM-HT(BamA(no cys)(R127C-N520C))	1813	199,867.7*	291,650	4.79 (5.00)
HT-SurA	429	47,372.5	29,450	7.10 (6.58)
HT-Skp	162	17,968.3	1,490	10.40 (9.65)

4 °C for 20 minutes (JA-25.50 rotor). The supernatant was then decanted into a fresh 50 mL Falcon tube and syringe-filtered (0.22 μ m). The sample was then loaded onto a 5 mL loop and ran at 2 mL.min⁻¹ onto a Superdex 75 HiLoad 26/60 (for tOmpA and Omp) or a Sephacryl S-200 HR HiPrep 26/60 (for BamA) pre-equilibrated with solubilisation buffer using an ÄKTA Prime Plus (GE Healthcare). 3 mL fractions were collected in regions on the chromatogram showing absorbance peaks at 280 nm (A280) and were buffer exchanged into 8 M urea using 0.5 mL ZebaSpin 7K MWCO desalting columns and analysed on an SDS-PAGE gel. Fractions containing the protein of interest were pooled and concentrated down to 1–2 mL using a VivaSpin 20 centrifugal concentrator with a MWCO 0.25X the molecular weight of the purified protein. The concentration was measured by UV spectroscopy from the calculated extinction coefficients at 280 nm (Table 2.6) on a NanoDrop 2000 and 50 μ L aliquots were snap frozen in liquid nitrogen before storage at -80 °C.

2.4.12 Protein expression and purification of periplasmic chaperones (Skp and SurA)

5 mL of an overnight culture of BL21(DE3) cells transformed with a pET28b vector containing Skp or SurA (with a hexa-histidine tag) was inoculated into 500 mL of LB medium containing $30 \mu\text{g}\cdot\text{mL}^{-1}$ kanamycin and grown at 37°C at 200 rpm in 2 L baffled flasks until reaching an OD₆₀₀ of 0.6 whereupon the temperature was lowered to 20°C and expression was induced with 0.4 mM IPTG. Cells were harvested after overnight expression and resuspended in 20 mM imidazole, 150 mM NaCl, 25 mM Tris-Cl pH 7.2 with a cocktail of EDTA-free protease inhibitor before being lysed in a cell disrupter (Constant Cell Disruption Systems, UK). After cell disruption, lysate was pelleted (40 mins, 4°C , $48,000 \times g$) to remove cell debris. For Skp-containing lysate (Skp pI \sim 9.7), this was applied to 3×5 mL HiTrap SP FF cation exchange columns (GE Healthcare) which were washed with start buffer (50 mM NaPO₄, 5 mM EDTA, pH 8.0) and Skp eluted over 5 column volumes of elution buffer (1 M NaCl, 50 mM NaPO₃, 5 mM EDTA, pH 8.0). Skp-containing fractions as determined by SDS-PAGE were pooled, diluted 1:2, and dialysed overnight (SnakeSkin 3.5K MWCO) into 25 mM Tris-Cl, 150 mM NaCl, 20 mM imidazole, pH 7.6 at 4°C . The dialysed Skp sample, or the SurA lysate, was applied to 3×5 mL HisTrap columns (GE Healthcare) which were washed with 25 mM Tris-HCl, 150 mM NaCl, 20 mM imidazole, pH 7.6 and denatured on-column with 25 mM Tris-Cl, 150 mM NaCl, 6 M guanidine-HCl, pH 7.6. Skp or SurA were eluted with a 0–500 mM imidazole gradient over 5 column volumes in 25 mM Tris-Cl, 150 mM NaCl, 6 M guanidine-HCl, pH 7.6 and Skp/SurA-containing fractions pooled. SurA was then loaded onto a Superdex 75 HiLoad 26/60 column and purified further by size-exclusion chromatography in unfolding running buffer (6 M guanidine-HCl, 25 mM Tris-Cl pH 7.20, 1 mM EDTA). The purest fractions (as determined by SDS-PAGE) were then pooled. Unfolded Skp- or SurA-samples were diluted 1:5 (to 1.2 M guanidine-HCl) and dialysed (SnakeSkin 3.5K MWCO) into 50 mM Tris-Cl pH 7.2 overnight at 4°C . The dialysed samples were then applied to 3×5 mL Q HP anion exchange columns (GE Healthcare) for purification in the case of SurA, and to bind impurities in the case of Skp (which doesn't bind to the column due to its high pI). For SurA, this was eluted over 5 column volumes of elution buffer (1 M NaCl, 20 mM Tris-Cl pH 8.0, 5 mM EDTA). SurA was then dialysed against 50 mM Tris-Cl pH 7.2 overnight at 4°C . The unbound fraction of Skp, and the eluted fractions of SurA were concentrated in a 5K MWCO Vivaspin 20 (Sartorius, UK) centrifugal concentrator, aliquoted, snap-frozen in liquid nitrogen and stored at -80°C .

2.4.13 Protein expression and purification of the BAM complex and cysteine variants

The complete BamABCDE complex was expressed and purified using a protocol adapted from the lab of Harris Bernstein (Roman-Hernandez et al., 2014). *E. coli* BL21(DE3) cells were transformed with plasmid pJH114 containing all five Bam genes (BamABCDE-HT) and grown overnight (37 °C, 200 rpm) in LB containing 100 µg.mL⁻¹ carbenicillin. Cells were diluted 1:100 into fresh 2xTY broth with the same antibiotic selection and grown (37 °C, 200 rpm) to an OD600 of 0.6 before addition of 0.4 mM IPTG to induce expression. Following 1.5 h expression, cells were harvested with a Beckman JLA-8.1000 rotor (4000 rpm, 15 min, 4 °C). The cell pellet was resuspended and homogenised in 10 mL per litre of growth medium of 20 mM Tris-Cl pH 8.0, lysed with a cell disruptor (Constant Cell Disruption Systems, UK), then centrifuged (6000 × g, 10 min, 4 °C). The supernatant was ultracentrifuged with a 50.2Ti rotor (244,280 × g, 45,000 rpm, 30 min, 4 °C) to pellet membranes. Pelleted membranes were incubated with 10 mL/L cold 50 mM Tris-Cl pH 8.0, 150 mM NaCl, 1 % (w/v) DDM at 4 °C for 2 h and the ultracentrifugation repeated to remove insoluble material. Supernatants were then incubated overnight at 4 °C with 2 mL/L Ni-NTA agarose on a tube roller. Ni-NTA beads were washed with one column volume of 20 mM Tris-Cl pH 8.0, 150 mM NaCl, 0.05 % (w/v) DDM, 50 mM imidazole and BamABCDE was eluted using two column volumes of 20 mM Tris-Cl pH 8.0, 0.05 % (w/v) DDM, 500 mM imidazole. The protein was concentrated to ~10 mg/mL, using a Vivaspin 100K MWCO concentrator and further purified by gel filtration chromatography using an Analytical Superdex 200 10/300 GL column, equilibrated with 150 mM NaCl, 20 mM Tris-Cl, 0.05 % (w/v) DDM running at 0.5 mL.min⁻¹. 0.5 mL fractions were collected and those containing complete BamABCDE complexes were identified by SDS-PAGE, pooled and concentrated. Protein concentration was determined using the Pierce BCA Protein Assay according to the manufacturer's instructions. The purified BamABCDE complex was concentrated to 10–15 mg.mL⁻¹, flash frozen in liquid N₂, and stored at -80 °C.

2.4.14 Determining BAM concentration (Pierce BCA Protein Assay)

The protein concentrations of purified BAM complex and proteoliposomes created by dialysis were measured using the Pierce BCA (bicinchoninic acid) Protein Assay. Briefly, using the 2 mg.mL⁻¹ BSA provided, a series of 10 standards from 0 – 2000 µg.mL⁻¹ BSA were prepared in the buffer of choice. This included TBS (50 mM Tris, 150 mM NaCl pH 8.0), TBS + 0.05 % DDM, 50 mM glycine-NaOH pH 9.5, or TBS + 10 % (w/v)

glycerol to match the sample to be measured. The sample was diluted in the same buffer to be within the 25 – 2000 $\mu\text{g}\cdot\text{mL}^{-1}$ range. Using a clear-bottom 96-well plate, 25 μL of each standard (in duplicate) and sample (in triplicate) was added to 200 μL of working reagent. The plate was covered, incubated at 37 °C for 10 minutes and absorbance at 562 nm was measured in a Clariostar Microplate Reader (BMG Labtech) equilibrated at 37 °C. The average blank-corrected measurement for the standards provides a standard curve when plotted against protein concentration. In situations where the R^2 of this fit was poor, readings were repeated at longer 37 °C incubation times. This was then used to determine BAM concentration by using the standards curve where the BSA trend line was most linear - i.e. the highest R^2 (at low time periods, low concentrations of BSA do not show above the background noise, at long time periods the high BSA can become overdeveloped). PC phospholipids have been shown to interfere with the BCA assay (Kessler and Fanestil, 1986) but this effect can be ameliorated by the inclusion of SDS (Morton and Evans, 1992) in the working reagent. For BAM and BamA DMPC proteoliposomes, the above protocol was modified to include the presence of 1 % (w/v) SDS in the working reagent (diluted out of a 20 % (w/v) stock solution).

2.4.15 Determining molecular mass by denaturing MS

Denaturing mass spectrometry was carried out on OmpA variants and labelled BAM(BamA [R127C-N520C]). Protein was separated from buffer by chloroform–methanol precipitation. For this a sample of protein (50 μL , 10 μM) was taken, and methanol (150 μL) and chloroform (50 μL) were added. The solution was mixed by vortexing, water (100 μL) was then added and the solution was mixed again before centrifuging (10,000 \times g, 2 min). The upper aqueous phase was removed (leaving the white protein pellet and the lower organic phase) and methanol (150 μL) was then added. The solution was mixed by vortexing, centrifuged (10,000 \times g, 2 min) and the supernatant removed. The precipitated protein was air dried in a laminar flow hood. The dried protein pellet was resuspended in formic acid (4 μL) and 18 M Ω H₂O was then added (46 μL) for subsequent MS analyses. Proteins were analysed intact using online desalting liquid chromatography–MS on a nanoAcquity LC system interfaced to a Xevo G2-S mass spectrometer (Waters Ltd., Wilmslow, Manchester, UK). The sample (2 μL) was loaded onto a MassPREP protein desalting column (Waters Ltd, Wilmslow, Manchester, UK), which was washed with 2 % (v/v) solvent B in solvent A (solvent A was 0.1 % (v/v) formic acid in water, solvent B was 0.1 % (v/v) formic acid in acetonitrile) for 5 minutes at 40 $\mu\text{L}\cdot\text{min}^{-1}$. After valve switching, the bound proteins were eluted using a fast gradient of 2–40 % (v/v) solvent B in A over 1 minute at 0.5 $\mu\text{L}\cdot\text{min}^{-1}$. The column was subsequently washed with 95 % (v/v) solvent B in A for 6 minutes

and re-equilibrated with 5 % (v/v) solvent B in solvent A for the next injection. The column eluent was infused into a Xevo G2-S mass spectrometer (Waters Ltd, Wilmslow, Manchester, UK). Data were processed using MassLynx v4.1, (Waters Corporation, UK) and UniDec (Marty et al., 2015).

2.4.16 Refolding and enrichment of folded BamA

BamA was purified from inclusion bodies and solubilised in 6 M guanidine-HCl, 5 mM DTT, 20 mM Tris-Cl pH 8.0 (see Section 2.4.11). This was then refolded into 0.5 % (w/v) N,N-dimethyldodecylamine N-oxide (LDAO), 500 mM arginine, 10 mM DTT, 300 mM NaCl, 50 mM Tris-Cl pH 8.0 which has been shown to lead to BamA giving well dispersed NMR spectra (Hartmann et al., 2018). BamA first was diluted down to 60 μ M in 6 M guanidine-HCl buffer and pre-equilibrated to 4 °C. 8 mL of refolding buffer was also pre-equilibrated to 4 °C and stirred vigorously with a magnetic flea. 2 mL of 60 μ M BamA solution was added 25 μ L at a time to give a final concentration of \sim 1 mg.mL⁻¹. This was then stirred for a further 24 hours at 4 °C. This was then dialyzed against 1 L of 0.1 % (w/v) LDAO, 20 mM Tris-Cl pH 8.0 for 24 hours with one change of buffer. Well folded BamA was then separated from unfolded and degraded BamA by anion exchange chromatography. A 5 mL HiTrap Q HP column (GE Healthcare) was equilibrated with buffer A (0.1 % (w/v) LDAO, 20 mM Tris-Cl pH 8.0) before BamA was loaded overnight by circulating through a peristaltic pump. BamA was then eluted with buffer B (buffer A + 500 mM NaCl) (12 column volumes (CV) 0 % B, 3 CV 0–40 % B, 6 CV 40 % B, 12 CV 40–65 % B, 3 CV 65–100 % B, 3 CV 100 % B). A low SDS gel was run on peak fractions and well-folded protein was found to elute \sim 45–55 % B. Fractions were pooled and concentrated in a 50K MWCO VivaSpin 20 concentrator at 4 °C before aliquoting, snap freezing in liquid N₂, and storing at -80 °C.

2.4.17 Production of BamA- and BAM-containing proteoliposomes

BAM complex was reconstituted into proteoliposomes using a procedure established for the outer-membrane protein FhuA (Thoma et al., 2012). For *E. coli* polar lipid proteoliposomes, 20 μ L DDM-solubilized BamABCDE (0.3 mg) was mixed with 200 μ L *E. coli* polar lipid films resuspended in 150 mM NaCl, 20 mM Tris-Cl pH 8.0 (TBS) using a 2:1 (w/w) final ratio of lipid to protein. This was placed into a 200 μ L 12–14 kDa MWCO D-tube dialyser mini and dialysed against detergent-free buffer (0.01 % (w/v) sodium azide, 150 mM KCl, 20 mM Tris-Cl pH 8.0) at room temperature for 3–4 days. Proteoliposomes were

harvested by centrifuging at $100,000 \times g$ (42,900 rpm), 4°C , 30 minutes in a TLA-110 rotor (Beckman Coulter) on a Optima MAX-XP Ultracentrifuge, supernatant discarded, and resuspended in TBS. This was repeated a second time and a sample of the pellet was boiled in 4X SDS-loading buffer for 10 minutes, with another sample run unboiled. Samples were run on regular SDS-PAGE gels to verify the presence of all 5 BAM components intact and show that BamA is folded (according to band-shift vs. the boiled protein). Proteoliposomes created using this procedure resulted in the vast majority of protein incorporated into the liposomes. Empty liposomes were made with $200\ \mu\text{L}$ *E. coli* polar lipids resuspended as above, with $20\ \mu\text{L}$ of 0.05 % (w/v) DDM in 150 mM NaCl, 20 mM Tris-Cl pH 8.0 added. For BAM DMPC proteoliposomes, DDM-solubilized BamABCDE was mixed with DMPC prepared as above and resuspended in TBS to give a final lipid-to-protein ratio (LPR) of 1600:1 (mol/mol) and final DDM concentration of 0.00442 % (w/v) ($694.2\ \mu\text{L}$ DMPC suspension, $176.8\ \mu\text{L}$ BamABCDE in 0.05 % (w/v) DDM TBS, $1129\ \mu\text{L}$ TBS). This mixture was transferred to a 2 mL 12–14 kDa MWCO D-tube dialyser maxi. Dialysis and preparation as above with the dialysis buffer temperature checked periodically to ensure it remained above 24°C . For BamA DMPC proteoliposomes, LDAO-solubilized BamA was mixed with DMPC as for BAM to give 1600:1 (mol/mol) LPR and a final LDAO concentration of 0.016 % (w/v) LDAO ($200\ \mu\text{L}$ DMPC suspension, $326.8\ \mu\text{L}$ BamA in 0.1 % (w/v) LDAO TBS, $1473.2\ \mu\text{L}$ TBS). Dialysis was carried out as for BAM proteoliposomes. Empty DMPC liposomes were prepared as for BAM-proteoliposomes but the BamABCDE volume replaced with 0.05 % (w/v) DDM, TBS. For both BAM- and BamA-proteoliposomes, and empty DMPC liposomes, after harvesting the dialysed liposomes twice by ultracentrifugation they were extruded through $0.1\ \mu\text{m}$ polycarbonate membranes $>21\ \times$ as part of a Mini-Extruder kit (see Section 2.4.10) with the kit and liposome sample pre-incubated at 37°C . This was then spun at $100,000 \times g$ for a final time, supernatant discarded, resuspended in TBS, aliquoted and snap-frozen before storing at -80°C . Protein concentration was determined by BCA protein assay and size checked by dynamic light scattering (DLS) (Section 2.6.3).

2.4.18 Labelling BAM with fluorescent dyes

Dye containing solutions were protected from light at all times to reduce photobleaching. 1 mg of lyophilized Alexa Fluor 488 and DyLight 594 dyes were reconstituted in DMSO to a stock concentration of 10 mM. $200\ \mu\text{L}$ of $\sim 95\ \mu\text{M}$ BAM(R127C/N520C) was buffer-exchanged using a 0.5 mL ZebaSpin 7K MWCO desalting column into reducing labelling buffer (10 mM DTT, 0.05 % (w/v) DDM, 1 mM EDTA, 150 mM NaCl, 50 mM Tris-Cl pH 7.6) and allowed to reduce for 45 min. This was then buffer exchanged into nitrogen-

Table 2.7 Properties of fluorescent dyes used for labelling. A280 correction factors were used to subtract the absorbance at 280 nm due to the presence of the dyes, in order to calculate the protein concentration from its extinction coefficient at 280 nm.

Fluorescent label	Absorbance maximum, λ_{\max} (nm)	Emission maximum, λ_{fl} (nm)	A280 correction factor (fraction of abs. at λ_{\max})
Alexa Fluor 488	495	519	0.11
Alexa Fluor 594	590	617	0.56
DyLight 594	592	616	0.585

sparged labelling buffer (as above, but without any reducing agent). The molar excess of each dye over BAM was varied until good 50:50 labelling of each dye was achieved, the best conditions were found to be a ~1.75X excess of Alexa Fluor 488, and a ~8.8X excess of DyLight 594. Dyes were pre-mixed and added to BAM(R127C/N520C) 10 μL at a time while mixing. This was then put on a roller and allowed to react at room temperature for 1 hour. This reaction was quenched with 100X molar excess of DTT over the dyes. The volume was then adjusted with labelling buffer to 200 μL before loading on a 0.5 mL capillary loop as part of an ÄKTA protein purification system. Separation of labelled protein from free dye was performed on a Superdex 200 10/30 GL pre-equilibrated with running buffer (0.05 % (w/v) DDM, 150 mM NaCl, 20 mM Tris-Cl pH 8.0). Separation was performed at 0.5 mL.min⁻¹ with constant monitoring at A280, A495, and A590. 0.5 mL fractions were collected and pooled in the region corresponding to the double-labelled protein. Labelled fractions were pooled and concentrated to around 500 μL in a 50 kDa MWCO VivaSpin 20 centrifugal concentrator. Protein concentration and labelling stoichiometry was determined by UV spectroscopy (NanoDrop 2000) with A280 corrections to subtract the additional absorbance at 280 nm from the dyes themselves (see Table 2.7). This gave a stoichiometry of 0.48 for Alexa Fluor 488, 0.52 for DyLight 594. 3 μL aliquots were snap frozen in liquid nitrogen and stored at -80 °C.

2.4.19 Reconstitution of BAM into DMPC for smFRET studies

Thin lipid films of DMPC were formed and dessicated as described in Section 2.4.10. This was then resuspended in 150 mM NaCl, 20 mM Tris-Cl pH 8.0 (TBS) to 10 mg.mL⁻¹. A 3 μL aliquot of BAM(R127C/N520C)[Alexa488/DyLight594] was rapidly diluted by mixing with 241 μL 10 mg.mL⁻¹ DMPC, 756 μL TBS in a 1.5 mL microfuge tube to give a final LPR of ~82,000:1 (mol/mol) (corresponding to 1 BAM per 100 nm LUV). This was incubated at room temperature for ~2 mins before ~50 Bio-Beads were added (to adsorb

detergent) and incubated in a heatblock at 50 °C for 10 minutes to allow integration of BAM into DMPC liposomes by removal of the detergent (dilution method of proteoliposome reconstitution). The sample was then loaded into a Mini Extruder (Avanti) but the size of the Bio-Beads prevented them from being taken up by the extruder syringe. >13 extrusions were performed and this sample was then diluted to ~100–500 pM in TBS (concentration adjusted to give well-separated fluorescent bursts during data collection).

2.5 Biochemical (experimental) methods

2.5.1 Band-shift assay (end-point)

OmpA samples were diluted from an unfolded state in high urea (8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0) into a folding buffer (final concentrations: 1600:1 LPR (mol/mol) DUPC LUVs, 0.24 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0) with a final volume of 100 µL and final protein concentrations of 2–6 µM to ensure bands are clearly visible on a gel after Coomassie or colloidal Coomassie stain. Two aliquots were taken for each folding reaction and mixed with 4X loading buffer with one being boiled for 10 minutes and the other not. These samples were then run side-by-side on a standard SDS-PAGE under standard conditions.

2.5.2 Band-shift assay (Kinetics of Tertiary Structure formation by Electrophoresis, KTSE)

All samples were pre-incubated in a Grant-bio PCH1 block heater (± 0.1 °C resolution, ± 0.1 °C accuracy) and the temperature further confirmed before each experiment with an Electronic Temperature Instruments Precision Plus PT100 thermometer and liquid probe (± 0.01 °C resolution, ± 0.08 °C accuracy) immersed in an Eppendorf tube containing 1 mL H₂O. tOmpA(WT) was diluted from an unfolded state in high urea (8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0) into either a folding buffer containing empty 100 nm DMPC LUVs, BamA-containing DMPC LUVs, or BAM complex-containing DMPC LUVs in 150 mM NaCl, 20 mM Tris-Cl pH 8.0 (1:10 tOmpA dilution from 10 µM tOmpA stock) giving final concentrations of 2 µM BamA or BAM, 1 µM tOmpA(WT), 0.8 M urea, 1600:1 LPR (mol/mol) DMPC, 150 mM NaCl, 20 mM Tris-Cl pH 8.0. For SurA-containing samples, tOmpA in high urea was initially diluted 1:15 from 150 µM to 10 µM into SurA-containing buffer (final conditions: 10 µM tOmpA, 100 µM SurA, 1 M

urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0), and then 26 μL of this solution diluted 1:10 into empty, BamA-containing, or BAM-containing DMPC liposomes to give the same final concentrations as for the no SurA sample. A timer was started immediately after dilution and mixing continued for 15 s. Samples were removed at set time-points and the folding reaction quenched by mixing 20 μL of sample with 4 μL 6X SDS loading buffer (150 mM Tris-Cl pH 6.8, 150 mM DTT, 6 % (w/v) SDS, 0.3 % (w/v) bromophenol blue, 30 % (v/v) glycerol). These were then ran on a regular SDS-PAGE gel, stained with InstantBlue, and imaged on a Uvitec Cambridge Q9 Alliance gel dock. ImageJ (Schneider et al., 2012) was used to perform densitometry on the gel lanes to calculate the relative intensities of the folded and unfolded monomer bands, and the fraction folded was calculated from the ratio of these band intensities according to:

$$\textit{Fraction folded} = \frac{\textit{intensity folded band}}{(\textit{intensity folded band} + \textit{intensity unfolded band})}$$

2.6 Biophysical methods

2.6.1 Fluorescence spectroscopy (intrinsic): steady state emission scans

OmpA samples were buffer exchanged into unfolding buffer (8 M urea, 50 mM glycine pH 9.5) using a 0.5 mL ZebaSpin 7K MWCO desalting column and the concentration was adjusted to 40 μM . Proteins were diluted 1:20 in folding buffer (final: 3200:1 LPR (mol/mol) of DUPC LUVs, 50 mM glycine pH 9.5, 0.24 M urea) or unfolding buffer supplemented with 3200:1 LPR (mol/mol) of DUPC LUVs, to give a final protein concentration of 2 μM . For refolding into detergent all conditions were held the same but phospholipids were replaced with 0.05 % (w/v) DDM (detergent-to-protein ratio, DPR = 1583 to 2438). These samples were allowed to fold overnight at 25 °C before diluting 1:5 into 500 μL to give a final concentration of 0.4 μM per scan. Scans were carried out in triplicate on a PTI QuantaMaster spectrofluorometer (Photon Technology International) in QS quartz cuvettes with excitation and emission slit widths set at 2 nm. Step size was 1 nm and integration time 1 s. For tryptophan emission, samples were excited at 280 nm and scanned from 300 to 400 nm. Final traces were an average of three scans.

2.6.2 Fluorescence spectroscopy (intrinsic): kinetics

Samples and buffers were prepared as described in Section 2.6.1. From a stock concentration of 40 μM , proteins were rapidly diluted out of 8 M urea-containing buffer into folding buffer to a final concentration of 3200:1 LPR (mol/mol) (DUPC LUVs) or 1600:1 LPR (mol/mol) (DMPC LUVs), 0.4 μM protein, 150 mM NaCl, 20 mM Tris-Cl pH 8.0, 0.24 M urea (1:100 dilution). Folding reactions were initiated by adding the protein sample on top of the folding buffer in the cuvette pre-equilibrated to 25 °C (DUPC LUVs) or 30 °C (DMPC LUVs), adding the cuvette lid, and then mixing by rapid agitation by hand, before placing into the fluorimeter and immediately starting the scan. Data was accumulated on a PTI QuantaMaster spectrofluorometer with excitation and emission slit widths set at 0.4 nm and 5 nm, respectively. Intrinsic protein fluorescence was followed by excitation at 280 nm and emission measured at 335 nm. For time-based single-wavelength monitoring, the integration time was 1 s. Scans were continued until a sufficient end-point baseline had been acquired. Protein, buffers, cuvettes and cuvette holder were pre-warmed to 25 °C for DUPC LUVs, 30 °C for DMPC LUVs. For BamA-containing samples, BamA was present in the DMPC LUVs at a final concentration of 0.8 μM (giving a 2:1 BamA:tOmpA stoichiometry).

2.6.3 Dynamic light scattering (DLS)

Empty DMPC liposomes, BAM-, and BamA-proteoliposomes formed by the dialysis method were diluted 1:1200 (to ~ 1 nM protein or ~ 2 μM DMPC) in 150 mM NaCl, 20 mM Tris-Cl pH 8.0 (TBS). All buffers were freshly filtered at 0.22 μm . Samples had been pre-incubated at 25 °C for 30 minutes before loading. After baselines of pure TBS were collected for 5 minutes, 250 μL of sample was injected onto a miniDawn Treos system (Wyatt Technology) in a batch-mode DLS setup and data collected for 10 minutes before being flushed out with fresh TBS and a second baseline collected. Data analysis was performed using Astra 6 software (Wyatt Technology).

2.6.4 Laurdan doped liposomes: measuring lipid phase and T_m

Laurdan was dissolved in DMSO to make a 1 $\text{mg}\cdot\text{mL}^{-1}$ (2.83 mM) stock and syringe filtered to 0.22 μm through a nylon filter. Liposomes were prepared to a final concentration of 0.8 μM BamA, BAM, or the equivalent concentration of pure DMPC empty liposomes. 1.5 μL of laurdan stock per 1 mL of liposomes were mixed together by first adding

laurdan to the bottom of an empty 1.5 mL Eppendorf tube then diluting by adding the liposome mixture on top to give a final lipid-to-probe ratio of ~300:1 (mol/mol), probe-to-BAM/BamA ratio of ~5:1, and a final DMSO concentration of 0.1508 % (w/v). This rapid mixing by adding the laurdan first prevents locally high concentrations of laurdan which may lead to a subpopulation of highly labelled liposomes. This concentration of DMSO should not significantly affect the structure of DMPC liposomes (Chang and Dea, 2001; Bonora et al., 2005; Mazeres et al., 2014; Cheng et al., 2015). Liposomes were then incubated at 25 °C overnight (16 hours) to allow partitioning of the probe into the liposomes.

2.6.5 Single-molecule FRET experiments

BAM(R127C/N520C) stochastically labelled with Alexa Fluor 488 and DyLight 594 was reconstituted into DMPC liposomes to approximately 1 BAM per liposome (see Section 2.4.19). 100 µL of samples at 100–500 pM were produced by diluting with TBS buffer incubated at the indicated temperatures and spotting on to a glass coverslip mounted on an objective set within a piezo-controlled stage connected to a water bath which was set at the indicated temperature. The smFRET microscope had been allowed to equilibrate to the water bath temperature for 3 hours before each experiment. This was covered with a plastic lid to minimise evaporation while the sample equilibrated with the temperature of the stage (5 minutes).

Experiments were performed on a custom-built Alternating Laser Excitation (ALEX) (Kapanidis et al., 2004) setup as described previously (Sharma et al., 2014). Briefly, the design is that of a custom inverted confocal microscope coupled with a diode-pumped 488 nm laser and a He-Ne 594 nm laser used for excitation both run in continuous wave mode (for donor [Alexa Fluor 488] and acceptor [DyLight 594], respectively). Laser switching was controlled by an acousto-optical modulator. Laser light was passed through a polarizer and shaped into a Gaussian beam profile using two 50 mm lenses and a 15 µm pinhole. This was then guided with a series of mirrors into the objective, illuminating the sample in the confocal volume. Fluorescence from the sample was then passed through a pinhole and filters and split with a dichroic mirror into two avalanche photodiodes (APDs). Laser powers were checked before each experiment and adjusted to 98 µW at 488 nm, 110 µW at 594 nm. Acousto-optic modulators attenuated the 488 nm laser to ~90 nW, and 594 nm laser to 15 nW, with an alternation period of 40 µs. A mirror was placed in the emission beamline to redirect scattered light to a camera which was used to monitor the distance of the focal plane relative to the coverslip and a piezo-controller (Piezo System

Jena) was used to adjust the focal height to 20 μm above the surface of the coverslip. 3×10 minute collections were performed before a fresh 100 μL sample was prepared, thermally equilibrated, and another 3×10 minute period collected. Control of the setup and data collection was managed using software developed in the LabView graphical programming environment (LabView 7.1, National Instruments, Austin TX, USA). Photon streams were converted and stored in the photon-hdf5 open file format (Ingargiola et al., 2016a). Fluorescence bursts were analysed as described previously with custom Python 2.7 scripts (Fessl et al., 2018) which are built around the FRETbursts Python toolkit (Ingargiola et al., 2016b). The functionality in the FRETbursts package allows the background signal to be estimated as a function of time, identification and removal of artefacts due to photophysical effects, and the optimisation of the signal-to-noise ratio. Efficiency values (E) were calculated after the application of correction factors: the γ -factor which accounts for differences in the efficiency of excitation of each dye, donor leakage into the acceptor channel, and direct acceptor excitation by the donor excitation laser. The data from each set of acquisitions was merged before further analysis. The ALEX setup allows the stoichiometry of dye labelling (S) to be approximated and a cutoff can be applied to only select double-labelled molecules with a 50:50 labelling ratio. ALEX-2CDE (Tomov et al., 2012) was used to filter the data and remove bursts arising from donor-, or acceptor-only labelled protein (either singly or doubly). This is approximately those molecules which fall within an S of ~ 0.25 – 0.75 . 1D histograms were produced from the filtered bursts and a 2D Gaussian kernel density estimation (Scott, 1992) was used to approximate probability density functions in plots of all data from each condition. Recurrence analysis of single particles (RASP) (Hoffmann et al., 2011) allows the dynamics of molecules on a 100 μs to ~ 100 ms timescale to be interrogated. In this method, very dilute samples have the property that there is a greater probability of the same molecule returning to the confocal volume than a new molecule entering. The approach used here is identical to that of Fessl et al., 2018. Building a RASP plot (i.e. performing a RASP analysis) involves selecting all molecules with a starting efficiency within a small range (e.g. 0.10 to 0.15, or 0.50 to 0.55), E_1 , and then comparing this to the E -value detected for all molecules detected within a certain window of time after E_1 (e.g. 10–12 ms after a burst with E_1 was detected), which are defined as E_2 . This window of time is called the recurrence interval. For molecules which exhibit no dynamics on this timescale, E_1 and E_2 should be identical (as they come from the same molecule, whose conformation has not changed). If dynamics occurs on this timescale, E_1 and E_2 will differ. The recurrence interval is then systematically varied to detect interconversion between conformational subpopulations. The longest usable recurrence time was calculated by estimating the recurrence probability using an autocorrelation function. This approach assumes that bursts from different molecules should be uncorrelated, but bursts from the same molecule will have a much greater

correlation. Only those bursts with a probability > 0.95 of being from the same molecule were chosen, according to the equation:

$$P_{same}(\tau) = 1 - \frac{1}{g(\tau)}$$

where $P_{same}(\tau)$ is the probability that the burst pair is from the same molecule, $g(\tau)$ is the burst time autocorrelation function of all detected bursts. 2D RASP histograms were created by plotting E1 versus E2 for those burst pairs with a $P_{same} > 0.95$, within the indicated recurrence interval. 2D Gaussian kernel density estimation analysis was used to create contour plots from this data.

2.6.6 cryoSTORM super-resolution microscopy

The cryoSTORM methodology was applied as described previously using a custom-built setup (Wang et al., 2019). Briefly, biological samples are applied directly to the surface of a solid immersion lens (SIL) which is made of a high-refractive index (RI) material (cubic zirconium). This is coupled to a 0.55 NA dry objective, but the optical properties of the SIL dramatically increases the effective NA of the objective up to the RI of the material used for the SIL. For cubic zirconium, the RI at 512 nm is 2.17. This gives a dramatic magnification boost to the imaging setup and greatly increases the efficiency of light collection. In the cryoSTORM method, the SIL-imaging concept is used under cryogenic conditions where the SIL assembly is held at 77 K in liquid nitrogen vapour. Under cryogenic conditions, the photophysical properties of dyes can be dramatically altered and can often result in very large increases in the photon budget of fluorophores due to reduced photobleaching. The switching properties of fluorophores can also be altered with molecules / fluorescent proteins which normally show little to no blinking behaviour undergoing cycles of on-off switching. These two properties can be combined to perform super-resolution single-molecule localisation methods (such as STORM and PALM) on traditionally ‘poor’ super-resolution probes (such as GFP derivatives), and the enhanced photon-budget can reduce the localisation uncertainty, increasing the effective resolution of the images.

E. coli samples for cryoSTORM microscopy were grown and prepared according to Section 2.4.6. Asymmetric solid-immersion lenses (SILs) were cleaned by application of piranha solution (13.8 M sulfuric acid, 7.5 % (w/v) hydrogen peroxide) for 15 minutes, thoroughly washed, then glow-discharged for 120 s on a Q150T ES Plus (Quorum Tech-

nologies) with air as the process gas to apply negative charge to the surface of the SILs to reduce the hydrophobicity of the surface. SILs were then loaded onto a FEI Vitrobot MKIV with a chamber equilibrated at room temperature, 95 % relative humidity. 2.5 μL of sample was applied and blotted manually at the edge of the SIL before plunge-freezing into liquid ethane (88 K). Where required, fiducials were applied *after* blotting the bacterial sample, blotted manually, then another 2.5 μL of sample was applied, and blotted manually again before plunge freezing. These were then transferred to a grid holder stored in liquid nitrogen (77 K) until use. Sample-loaded SILs were loaded onto a custom-built grid holder mounted in a cryo-stage cooled by liquid nitrogen (Linkam, CMS-196), as part of a custom-built microscope. A 470 nm LED was used to illuminate the sample to find a good field-of-view before imaging was performed with a 405/488/561/642 nm continuous wave laser for excitation coupled to an EMCCD camera (Andor, iXon+ DU-897) to collect emitted fluorescence. 405 nm laser used at 5 mW laser power, 488 nm laser was used at 50 mW laser power. 10,000–20,000 frames were collected per sample with 50 ms exposure time per frame. For two-colour images, imaging was sequential. Features detected in both channels were used to register two-colour super-resolution images during image post-processing (this was either fiducial markers or fluorescent features of the SIL surface).

2.6.7 Data fitting

Tryptophan fluorescence kinetics, crosslinking time-courses, KTSE, and laurdan doping were fitted using IgorPro 7 (WaveMetrics) to a single exponential equation of the form:

$$y = y_0 + A \exp^{-kt}$$

where y_0 , y-intercept; A, amplitude, e, exponential; k , the observed rate constant; t, time (in seconds). y_0 , A, and k were free variables. KTSE data was also fitted to a double exponential equation of the form:

$$y = y_0 + A_1 \exp^{-k_1 t} + A_2 \exp^{-k_2 t}$$

where A_1/k_1 and A_2/k_2 represent the two phases of the graph.

t_{50} values from single exponential fits were derived by rearranging the exponential equation and solving for t when y is $y_0/2$ with parameters derived from the fits to the

experimental data:

$$t_{50} = \frac{\ln\left(\frac{\frac{y_0}{2} - y_0}{A}\right)}{-k}$$

t_{50} values for double exponential fits were calculated numerically by solving the double exponential equation for y with all parameters except t derived from the fits to the experimental data. Solutions for all values of t from ~ 0 –86400 s in 0.1 s steps were solved and the value of t closest to the 50 % of maximum fluorescence yield value (aka $y_0/2$) was taken as the calculated t_{50} . There exists no general algebraic / arithmetic solution to equations in the form of sums of exponentials where the exponent varies.

Reported values are the average of at least three repeats and the error bars on the fit are the standard error of the mean according to:

$$\sigma_M = \frac{\sigma}{\sqrt{N}}$$

where σ_M is the standard error of the mean; σ , the standard deviation; and N , the sample size.

HT-SurA-tOmpA binding was simulated with a custom Python 3.X script implemented in the Anaconda Distribution. The K_d for the HT-SurA-tOmpA interaction (1.8 μM (Humes et al., 2019)) was used in a quadratic binding equation with the concentration of receptor (SurA) or ligand (tOmpA) varied.

2.7 Bioinformatic methods

2.7.1 Multiple sequence alignment

The MultAlin server (<http://multalin.toulouse.inra.fr/multalin/>) (Corpet, 1988) was used for sequence alignments. The postscript output file from this server was then fed into the ESPript 3.0 server (<http://esprict.ibcp.fr/ESPript/ESPript/>) which automatically annotates the sequences (Robert and Gouet, 2014).

2.7.2 *E. coli* K-12 OMP-ome identification

The identity of the *E. coli* K-12 OMP-ome was manually curated through literature searches, transmembrane beta-barrel prediction algorithms, the StepDB 2.0 (beta) database (Loos et al., 2019), and OMPdb database (ompdb.org) (Tsirigos et al., 2011). Candidate OMPs were further confirmed through literature searches, searches for structural homologues in the PDB, assessing ontology terms in Uniprot (uniprot.org) (UniProt Consortium, 2019), PFAM domain identification (pfam.xfam.org) (El-Gebali et al., 2019), assessing curated *E. coli* gene information on EcoCyc (ecocyc.org) (Keseler et al., 2017), visual assessment of the local genomic region surrounding the gene, and structural modelling using I-TASSER (zhanglab.ccmb.med.umich.edu/I-TASSER/) (Yang and Zhang, 2015). For OMP fragments (where part of the gene has undergone deletion or an insertion has occurred within the gene inactivating it - usually due to the integration of insertion-sequence elements from phages), the gene was reconstructed from homologues as it is very likely this OMP would still be a true substrate of the *E. coli* K-12 OMP biogenesis machinery, had it not been inactivated by phages.

2.7.3 Cys and Lys abundances by subcellular location

Amino acid sequences and Uniprot IDs for the *E. coli* K-12 proteome were downloaded from the Uniprot database. These were then cross-referenced with the StepDB 2.0 (beta) database (Loos et al., 2019) to attach subcellular localisations for the *E. coli* K-12 core proteome and merged into a single CSV file. A custom Python 3.X script was then used to calculate the abundances of a chosen amino acid in its subcellular compartment. Periplasm+ was defined as those proteins with a localisation assigned to either peripherally associated with the plasma membrane facing the periplasm, inner membrane lipoprotein, periplasmic, outer membrane lipoprotein, peripherally associated with the outer membrane facing the periplasm, outer membrane protein beta-barrel protein, or peripherally associated with the outer membrane facing the extra-cellular space.

2.7.4 BAM folding rate *in vivo* / metabolic cost of OMPs

The absolute number of OMPs of each type was acquired using the data of absolute synthesis rates from ribosome profiling experiments (Li et al., 2014). To get an upper bound on the folding rate required by BAM to support physiologically relevant *E. coli* growth, the doubling time of *E. coli* was taken as 20 minutes and divided by the sum total

of OMPs. To calculate the metabolic cost of OMPs, a custom Python 3.X script was used to calculate the per molecule energy cost to an *E. coli* cell for producing each type of OMP based on its sequence and the per-residue energy cost (Akashi and Gojobori, 2002). This was then multiplied by the absolute number of each OMP to give the *per cell* sum total energy cost to produce the OMP-ome.

2.8 Cross-linking experiments

2.8.1 Labelling with maleimido-benzophenone

Aliquots of OmpA cysteine variants in 6 M guanidine, 50 mM Tris-Cl pH 8.0 were buffer exchanged into 8 M urea, 2 or 5 mM TCEP, 50 mM Tris-Cl pH 5.0 using ZebaSpin 0.5 ml 7K MWCO columns. 4-(*N*-maleimido)benzophenone was dissolved in DMSO to create a 200 mM stock. DMSO and an OmpA cysteine variant was mixed to give the following concentrations: 200 μ M OmpA, 6 % (v/v) DMSO, 2–5 mM TCEP, 8 M urea, 50 mM Tris-Cl pH 7.0. This was then incubated at room temperature for 30 min on a roller and shielded from direct light before being moved to 4 °C and continued rolling overnight. The following day the reaction was quenched by the addition of 120 mM DTT final concentration (10 X molar excess over crosslinker) and buffer exchanged 3 X using ZebaSpin 0.5 ml 7K MWCO spin columns into 50 mM glycine pH 9.5.

2.8.2 Labelling with MTS-daizirine or MTS-TFMD

Aliquots of OmpA cysteine mutants were buffer exchanged into 6 M guanidine-HCl, 50 mM Tris-Cl, 10 mM DTT, 1 mM EDTA, pH 8.0 which had been sparged with nitrogen gas and left for 15 min at room temperature to ensure all cysteines are reduced. The presence of EDTA to chelate heavy metal contaminants and the displacement of dissolved O₂ by nitrogen reduce the propensity for the re-oxidation of cysteine or disulfide bond formation before addition of the MTS-based crosslinking reagent. A second buffer exchange was performed into the same buffer without DTT. Stock solutions of MTS-diazirine or MTS-TFMD in DMSO were added in 20-times molar excess over OmpA. Final concentrations were: 200–280 μ M OmpA, 4 mM MTS-diazirine or MTS-TFMD, 20 % (v/v) DMSO, 4.8 M guanidine-HCl, 40 mM Tris-Cl, and 0.8 mM EDTA. This solution was incubated with rolling at room temperature for 1 hr. The solution was then buffer exchanged using 0.5 mL Zeba Spin Desalting Columns, 7K MWCO into 6 M guanidine-HCl, 50 mM

Tris-Cl, 1 mM EDTA, pH 8.0 and stored at 4 °C or aliquoted, snap frozen in liquid nitrogen and stored at -80 °C.

2.8.3 Crosslinking sample preparation

For all irradiation methods, samples were prepared in the following way. Aliquots of Skp trimers, SurA, or BAM were buffer exchanged into 50 mM Tris-Cl, 150 mM NaCl (TBS) pH 7.6, and crosslinker-conjugated OmpA into 8 M urea, TBS pH 7.6, using 0.5 mL Zeba Spin Desalting Columns, 7K MWCO (Thermo).

2.8.4 Photo-crosslinking

To form the chaperone-substrate complex, urea-denatured OmpA was diluted into a Skp containing buffer from high (8 M) to low (0.8 M) urea at a final stoichiometry of 2:1 (mol/mol) Skp trimer to OmpA monomer (8 μ M Skp trimer/4 μ M OmpA). For Hg-Xe lamp experiments, 30–130 μ L of sample was added to a thin-walled PCR tube and placed on its side on a stack of petri dishes. This was then placed under a covered LF-206.LS 6 W 365 nm lamp (Uvitec) at a distance of 12 mm from the source for different times up to 30 min. For UV LED chip experiments, acrylic chip chambers were washed with TBS pH 7.6 buffer and 33 μ L of sample was added. These were then clamped into the chip holder and irradiated with the 365 nm LED for up to 30 s. A pipette was used to remove the sample from the chip via the lower filling hole. For UV LED tube experiments, 30–130 μ L of sample was added to a thin-walled PCR tube and placed inside a 0.5 mL microfuge tube which itself was in a 1.5 mL microfuge tube that had been wrapped in aluminium foil (to reflect the UV irradiation within the containers). Samples were irradiated with the 365 nm LED for 30 s.

For SurA crosslinking experiments, urea-denatured OmpA in a 120 μ M stock was diluted into a SurA containing buffer from high (8 M) to low (1 M) urea at a final stoichiometry of 10:1 (mol/mol) SurA to OmpA (100 μ M SurA/10 μ M OmpA). Slightly higher urea was used in these experiments due to the higher concentration of OmpA being more prone to aggregation, while urea helps to keep it soluble. The crosslinking reaction and enrichment then proceeded as for Skp described above.

For BAM crosslinking experiments, OmpA was rapidly diluted from 8 M urea into buffer containing BAM proteoliposomes of *E. coli* polar lipids and HT-SurA giving a final

volume of 30 μL and a final concentration of 10 μM BAM, 50 μM SurA, 10 μM OmpA in 1 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0. This mixture was then allowed to fold for 5 min at room temperature in a 200 μL PCR tube before being transferred to an acrylic chip and irradiated for 1 min at 365 nm. After crosslinking samples were kept on ice until ready for enrichment. 3X 30 μL samples were pooled together to increase the quantity of crosslinked material.

2.8.5 Separation and enrichment of cross-linked products

To separate and/or enrich crosslinked samples three methods were used as described in Chapter 3.

2.8.5.1 Method 1

Crosslinked samples were mixed with non-reducing SDS-PAGE loading buffer (final concentrations: 50 mM Tris-Cl pH 6.8, 2 % (w/v) SDS, 0.1 % (w/v) bromophenol blue, 10 % (w/v) glycerol) and run directly on a 15 % (w/v) non-reducing Tris-tricine SDS-PAGE gel to separate cross-linked product from non-cross-linked (or intramolecular cross-linked) material. Gels were stained using InstantBlue protein stain (Expedeon). The cross-linked band was then excised for in-gel digestion (see below).

2.8.5.2 In-gel digestion

Gel bands were cut into approximately 1 mm³ pieces, and destained by incubating in 30 % (v/v) ethanol at 60 °C for 30 min. Reduction was performed by incubating with 50 μL of 10 mM DTT (in 25 mM ammonium bicarbonate) at 56 °C for 1 hr, followed by alkylation with 50 μL 55 mM iodoacetamide (in 25 mM ammonium bicarbonate pH 8) at room temperature for 45 min in the dark. The gel pieces were dehydrated with 100 % acetonitrile, and dried in a laminar flow hood for 60 min. The gel pieces were rehydrated with 20 μL 0.02 $\mu\text{g}\cdot\mu\text{L}^{-1}$ trypsin solution (Promega) in 25 mM ammonium bicarbonate pH 8, and incubated at 37 °C for 18 hr with shaking. Peptides were recovered by incubating gel pieces with 50 μL 60 % (v/v) acetonitrile/5 % (v/v) formic acid (x3) for 10 min. The peptides were then evaporated to dryness and resuspended in 20 μL with 5 % (v/v) acetonitrile/0.1 % (v/v) formic acid prior to MS analysis.

2.8.5.3 Method 2

30–130 μL of cross-linked sample was buffer exchanged into 8 M urea, 10 mM DTT, TBS pH 7.6 using 0.5 mL Zeba™ Spin Desalting Columns, 7K MWCO (Thermo Scientific) and rotated for 30 min at room temperature. This reduced sample was buffer exchanged again into 8 M urea, TBS pH 7.6 to remove excess DTT and then added to 200 μL of a thiopropyl Sepharose 6B (GE Healthcare) slurry (100 μL of settled medium), mixed by pipetting, and incubated at room temperature for 60 min with rotation. The sample-bound thiopropyl Sepharose beads were pelleted by centrifugation at $14,000 \times g$ for 1 min on a bench-top centrifuge and the supernatant removed. The beads were washed with at least double the volume of the medium plus sample using 8 M urea, TBS pH 7.6, centrifuged, and the supernatant removed. This step was repeated until no protein was detected in the supernatant by Coomassie stain after SDS-PAGE analysis (around 6×1 mL washes for OmpA-Skp in the concentrations used here). Bound proteins were then eluted by the addition of 100 μL of 8 M urea, 100 mM β -mercaptoethanol, TBS pH 7.6. This was incubated for 30 min at room temperature with rotation before the beads were pelleted by centrifugation and the supernatant retained. The eluted material was separated by SDS-PAGE (see Method 1) and the gels were stained using InstantBlue protein stain (Expedeon). The band corresponding to the target protein (modified with the XL reagent) was then excised for in-gel digestion (see above).

2.8.5.4 Method 3

Method 3 proceeded in the same way as Method 2 (above), but diverged before the final elution step. Instead of eluting with β -mercaptoethanol, TBS pH 7.6 was added to the beads to dilute the urea concentration to approximately 1 M. Sequencing grade trypsin (Promega) was added to a final 1:50 (w/w) ratio of trypsin:protein and the mixture was incubated at 37 °C, 600 RPM for between 4 hr to overnight. The beads were isolated by centrifugation and the supernatant was discarded. The beads were washed with 2×1 mL of 8 M urea, TBS pH 7.6, followed by 5×0.5 mL TBS pH 7.6. Bound peptides were eluted by addition of 100 μL of 8 M urea, 40 mM β -mercaptoethanol, TBS pH 7.6 and the mixture was incubated with rotation at room temperature for 30 min. The beads were pelleted by centrifugation and the supernatant removed and retained. Free thiols were alkylated by addition of iodoacetamide to a final concentration of 133 mM (50 μL of a 0.4 M stock) and incubated with rotation for 1 hr at room temperature. This sample was then desalted using 10 μL C18 ZipTip pipette tips with 0.6 μL bed volume (Merck) according to manufacturer's instructions. The residual solvent was then evaporated to

dryness in a vacuum evaporator (Thermo Scientific). Samples were frozen at -80 °C for later analysis or resuspended in 20 μL 5 % (v/v) acetonitrile/0.1 % (v/v) formic acid for injection onto an LC/MS system.

2.8.5.5 Mass spectrometry

Peptides (5 μL) were injected onto a reverse-phase Acquity M-Class C18, 75 μm x 150 mm column (Waters) and separated by gradient elution of 1–50 % solvent B (0.1 % (v/v) formic acid in acetonitrile) in solvent A (0.1 % (v/v) formic acid in water) over 60 min at 300 $\text{nL}\cdot\text{min}^{-1}$. The eluate was either infused either into a Xevo G2-XS (Waters) or Orbitrap Q Exactive (Thermo) mass spectrometer operating in positive ion mode. In all cases, data processing and modification localization was performed using PEAKS Studio 7 (Bioinformatics Solutions).

For the Xevo G2-XS, mass calibration was performed by infusion of aqueous sodium iodide at a concentration of 2 $\mu\text{g}\cdot\mu\text{L}^{-1}$. [Glu1]-Fibrinopeptide B (GluFib) was used as a lock mass calibrant with a 0.5 second lock spray scan taken every 30 seconds during acquisition. The lock mass correction factor was determined by averaging ten scans. Data acquisition was performed in DDA mode with a one second MS scan over an m/z range of 350-2000. The four most intense ions in the MS spectrum were selected for MS/MS by CID, each with a 0.5 second scan over an m/z range of 50-2000. The collision energy applied was dependent upon the charge and mass of the selected ion. Dynamic exclusion of 60 seconds was used.

Orbitrap calibration was performed using Ultramark solution (Thermo). Data acquisition was performed in DDA mode and fragmentation was performed using HCD. Each high-resolution full scan (m/z range 500-2000, $R=120,000$) was followed by high-resolution product ion scans ($R=15,000$), with a normalised collision energy of 30 %. The 15 most intense ions in the MS spectrum were selected for MS/MS. Dynamic exclusion of 60 seconds was used.

2.9 UV LED system manufacture and testing

2.9.1 Acrylic chip fabrication

Sample chips for UV irradiation were fabricated from laser-cut poly(methyl methacrylate) (PMMA, Perspex) pieces using a solvent-assisted bonding process (Liga et al., 2016). Each chip comprises 3 layers: a top UV-transparent layer with inlets/outlets for insertion of a 100 μL pipette tip, a middle layer with a 33 μL well, and a bottom UV-transparent layer which seals the well. Firstly, chip components were designed and cut from PMMA sheets into $50 \times 25 \text{ mm}^2$ pieces with a VLS6.60 CO_2 laser (Universal Laser Systems). Middle pieces were made from 1 mm thick sheets and the top/bottom pieces from 1.8 mm thick sheets. After cutting, these components were rinsed with ethanol and air dried. Next, the pieces were bonded with an Instron 3300 Dual Column Universal Testing System. The system was modified into a heated press using bespoke top and bottom stainless steel plates, each embedded with two cartridge heaters and a thermocouple. Both plates were maintained at 70 $^\circ\text{C}$ using a panel mounted proportional-integral-differential (PID) temperature controller (RS Components Ltd). The bottom and middle layers were bonded by pipetting 60 μL of ethanol between them and placing them in the press, which was operated using Bluehill testing software (Instron). Bonding was controlled by a predefined program which compresses the PMMA pieces until a force of 1700 N (1.36 MPa) is reached and held for 2 min. This process is then repeated to bond the top and middle pieces.

2.9.2 UV LED lamp, optics, and sample holder construction

The heart of the UV lamp was a 4 UV LED array mounted on a standard star circuit board (Led Engin part LZ4-44UV00-0000, peak wavelength 365 nm, radiant flux 4.1 W, viewing angle 110°). This was attached to a heat sink (DragonStar part ILA-HSINK-STAR-50X60MM-BLK-K) to dissipate heat from the LED array. Two additional holes were drilled into the heatsink to allow passage of the electrical wires beneath the mounting plate. A lens was attached to the front of the lamp through the use of a mounting plate (Thorlabs CP02T/M) attached to the heatsink with two threaded bars (M3 thread), to which a lens tube (Thorlabs SM1L20) was attached. A condenser lens (Thorlabs ACL2520U-A, $\varnothing 25 \text{ mm}$, $f=20.1 \text{ mm}$, $\text{NA}=0.60$), located within this lens tube, was used to focus the light to a 6 mm spot. As a guide, the centre of the lens was positioned $\sim 20 \text{ mm}$ from the LED and $\sim 25 \text{ mm}$ from the sample. Power was provided by a dedicated constant current LED

controller (eldoLED PWR180D1) regulated to 1000 mA output current. In turn this was powered from a standard desktop power supply with an output of 19 V. The duration of illumination was controlled by a timer-box which controlled a relay that made or broke the circuit between the controller and LED. This approach bypassed the gentle start-up offered by the controller and minimised variation in light intensity. For the chip based experiments, the lamp was mounted conveniently using an optical rail (Thorlabs), with a holder used to reproducibly position the chip. For the Eppendorf unit, the lamp was mounted above a block that held the tube.

2.9.3 Thermocouple measurements of UV LED with tube holder

Heating of the solutions due to UV irradiation was investigated using a Pico TC-08 Thermocouple Data Logger. Briefly, a K-type thermocouple was inserted into a thin-walled PCR tube containing 200 μ L of deionised water. Subsequently, a range of UV exposure times between 5 s and 5 min were conducted during constant temperature monitoring.

2.9.4 Infrared thermography measurements of sample heating

Heating of the wells in the PMMA chips due to the UV LED was monitored by IR Thermography using a FLIR 640 SC camera and ResearchIR software. Images taken directly after exposure times between 5 s and 5 min were utilised to determine the temperature increase, where pixel values corresponding to the well were averaged and compared to their average value before UV exposure. An unheated piece of black electrical tape (emissivity ~ 1) was used as a reference in each image. The same process was applied to monitor heating of a thin-walled PCR tube containing 200 μ L of deionized water after exposure times between 15 s to 20 min from a 6 W 365 nm lamp Hg-Xe lamp.

2.10 Other software tools

All graphs and category plots were made in Igor Pro 7. Chemical structures made in ChemDraw Prime 17.1. Protein structures all rendered in PyMol 2.X except where otherwise indicated. Single-molecule FRET RASP and E vs S plots made in Python 3.X using Matplotlib.

Chapter 3

The development and application of tag-transfer crosslinkers

3.1 Introduction

Chemical crosslinking is a powerful technique for interrogating the structure and dynamics of protein complexes both *in vivo* and *in vitro*. When coupled with mass spectrometry (crosslinking-mass spectrometry, XL-MS) it can allow high-resolution information about interaction surfaces to be acquired in a relatively high-throughput manner (Rappsilber, 2011; Calabrese and Pukala, 2013; Sinz, 2014; Schneider et al., 2018). Photoactivatable groups allow more precise control over the crosslinking reaction by masking a reactive intermediate behind a photolabile moiety. These molecules are relatively inert in the dark, allowing their installation onto proteins *in vitro* or *in vivo* until the experimenter is ready to utilize them by irradiation with specific wavelengths of UV or visible light. Requiring an input of energy (in the form of photons) to trigger unmasking allows access to higher energy, and thus more reactive, molecules.

Diazirines are one such class of photolabile groups and consist of a carbon bound to two nitrogen atoms, which are double bonded to each other forming a three-membered ring (Figure 3.1A). The diazirine group has a characteristic absorbance band around 350 nm which, upon irradiation with UV light, can generate an excited $n-\pi^*$ state that decomposes via two major pathways to form a carbene-containing product (Figure 3.1B) (Yamamoto et al., 1994). This highly reactive carbene is capable of inserting into any X–H bond (C–H, O–H, N–H, S–H) with hardly any discrimination between sites (Geurink et al., 2012),

a nanosecond lifetime (Toscano et al., 1995; Admasu et al., 1998), and reaction rates approaching the diffusion limit (Griller et al., 1984).

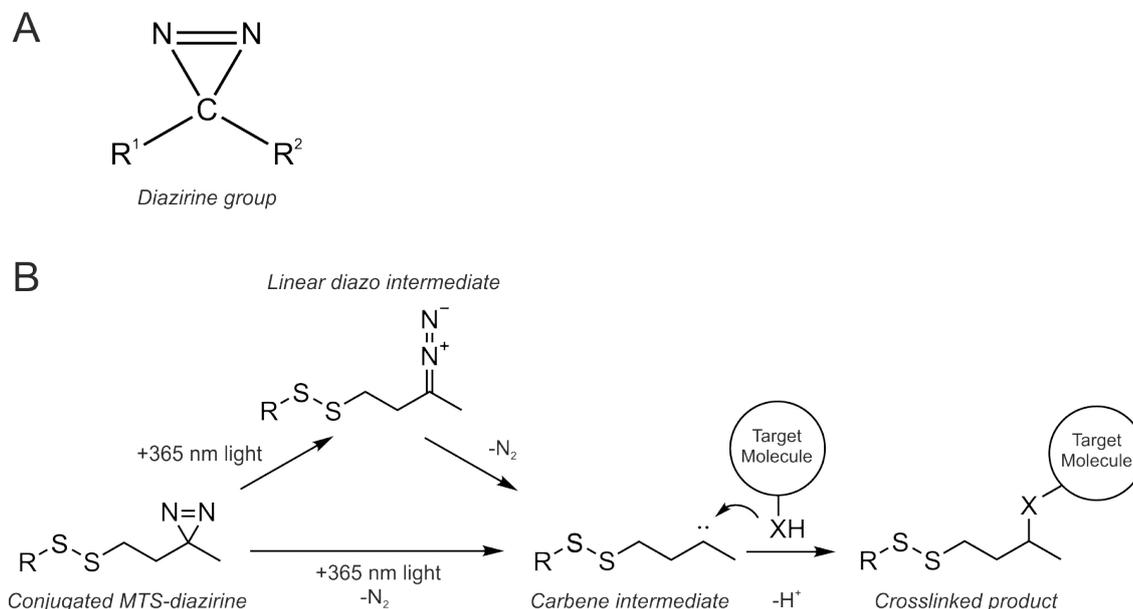


Fig. 3.1 Diazirine photodecomposition and reaction scheme. (A) Structure of a diazirine reactive group. (B) Mechanism of activation and bond-insertion by diazirines via 365 nm illumination. Irradiation of a diazirine moiety with 365 nm light directly unmasks a highly reactive carbene intermediate which will insert non-preferentially into any X-H bond. Alternatively, the reaction can proceed via the formation of a linear diazo intermediate which preferentially reacts with nucleophilic residues.

The first reports of photocrosslinking of biological molecules appeared in the 1960s (Singh et al., 1962). Diazirine functional groups were first introduced into more complex molecules around this time (Church et al., 1965) and developed for use as a biological photocrosslinking reagent through the 1970s and early 1980s (Smith and Knowles, 1973, 1975; Bayley and Knowles, 1978; Brunner et al., 1980). Nonetheless, it lagged in popularity compared with other photoactivatable groups such as benzophenones and aryl azides. The relative neglect of diazirine chemistry likely stemmed from the easier synthesis routes and yields of benzophenones and azides (and therefore, commercial availability) (Bayley, 1983), and the problems with internal rearrangement of smaller alkyl diazirines (Erni and Khorana, 1980). The use of diazirines is now catching up with other photoactivatable groups due to more widespread availability and the discovery of improved synthesis strategies of the alkyl diazirine form (MacKinnon et al., 2007; Bond et al., 2009).

The last 20 years has seen a huge increase in our structural understanding of the proteins which comprise the outer membrane protein (OMP) biogenesis pathway with high-resolution structures available for almost every protein known to be involved in the pathway (see Introduction Figure 1.7). This wealth of structural knowledge betrays our relative lack

of understanding about how these proteins interact with each other, and particularly how they work mechanistically (Plummer and Fleming, 2016; Bakelar et al., 2017; Konovalova et al., 2017; Noinaj et al., 2017; Schiffrin et al., 2017a; Mas and Hiller, 2018; Ranava et al., 2018; Ricci and Silhavy, 2019; Wu et al., 2019). Understanding OMP biogenesis means being able to follow the pathway from translation on the ribosome, delivery to and translocation through the Sec apparatus, chaperoning through the periplasmic space, delivery to the BAM complex, and finally folding of the OMP into its final tertiary structure. Throughout this whole process an OMP remains in an unfolded and dynamic state, making it recalcitrant to traditional structural techniques. XL-MS provides a method that can handle the conformational heterogeneity inherent in interactions involving one or more partners in a disordered and structurally dynamic state (Leitner et al., 2016; Chen and Rappsilber, 2018; Sinz, 2018).

In this chapter a new method for performing XL-MS using photoactivatable and enrichable tag-transfer crosslinkers based on diazirine chemistry is shown. The development of this method and the creation of a new low cost UV LED-based lamp to allow the time-resolved study of transient, non-equilibrium reactions is also discussed. The power of this new tag-transfer crosslinking methodology is used to study the chaperoning of OmpA by Skp and SurA to confirm the hydrophobic cavity capture mechanism of Skp and show for the first time how SurA captures client OMPs almost entirely via its N- and C-terminal (core) domains with little to no interactions observed between OmpA and the two peptidyl-prolyl-cis/trans-isomerase (PPIase) domains of SurA, P1 and P2. Finally, work towards capturing the pathway of OmpA during SurA-mediated, BAM-catalysed folding is presented showing the presence of crosslinks from OmpA to BamA and the possible involvement of the SurA P2 domain in this process.

3.2 Results

3.2.1 Limitations of benzophenone crosslinkers for capturing dynamic and transient interactions

The very low abundance of cysteine residues in periplasmic proteins and OMPs (Figure 3.2.1), combined with their flexible chemistry, makes cysteine-based site-directed mutagenesis a powerful tool for installing novel reactivity to probe the local interactions and conformations of the OMP biogenesis machinery. Heterobifunctional benzophenone crosslinkers which utilise maleimide chemistry to allow site-specific incorporation of this

crosslinker and such reagents into proteins are readily available commercially. Three single-cysteine mutants of OmpA were created, expressed, and purified (see Materials & Methods 2.3, 2.4.4, & 2.4.11 for details), with a cysteine introduced at the N-terminus (N_{cys}), the bottom of $\beta 1$ (W7C) and in loop 3 near the membrane (T144C), respectively (Figure 3.3).

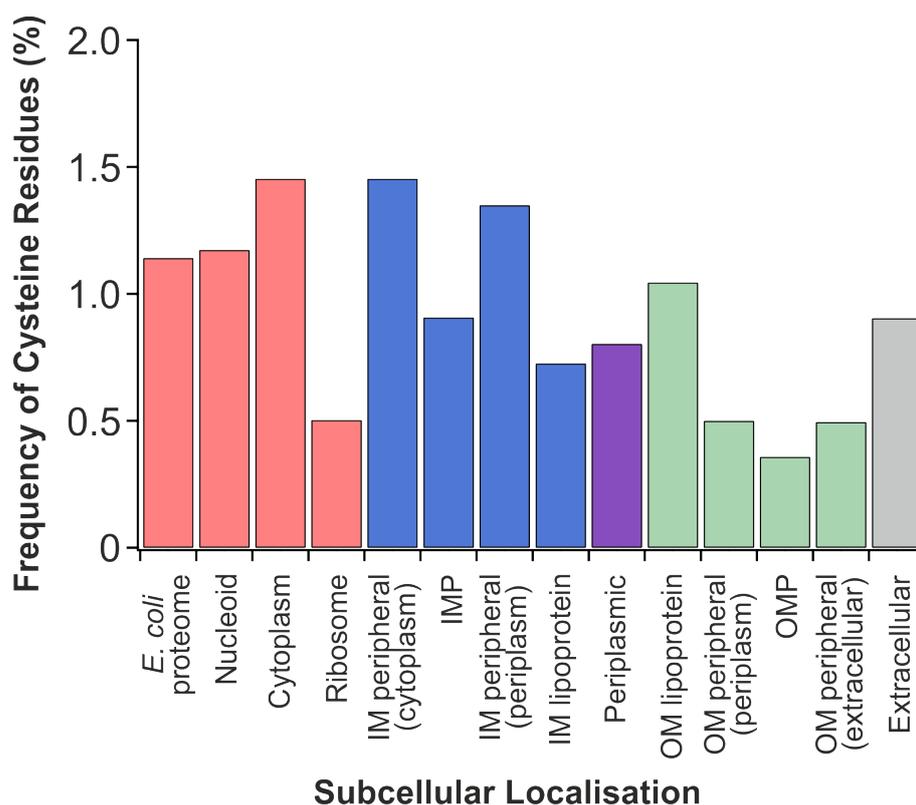


Fig. 3.2 Relative frequency of cysteine residues in proteins from *E. coli* K-12 MG1655. Subcellular localisation of proteins was acquired from STEPdb 2.0 beta (<http://stepdb.eu/>) (Loos et al., 2019) and the corresponding sequences taken from UniProt (<https://www.uniprot.org/>) (UniProt Consortium, 2019). IM = inner membrane, OM = outer membrane, IMP = inner membrane protein, OMP = outer membrane protein.

OmpA(N_{cys}) in 5 mM Tris(2-carboxyethyl)phosphine (TCEP), 8 M urea, 50 mM Tris-Cl pH 7.0 was mixed with a 50 X molar excess of maleimido-benzophenone from a DMSO stock, giving a residual 6 % (v/v) DMSO concentration. This was left at room temperature for 30 min before incubating overnight at 4 °C. Excess maleimide was quenched by the addition of 120 mM dithiothreitol (DTT) and was removed by three buffer exchange steps on a ZebaSpin 0.5 ml 7K MWCO spin column into 8 M urea, 50 mM glycine pH 9.5 (see Materials & Methods 2.8.1 for details). This construct now contained maleimido-benzophenone site-specifically introduced on its cysteine (henceforth, OmpA(N_{cys})-BP). For crosslinking reactions, OmpA(N_{cys})-BP was diluted from high (8 M) to low (<0.4 M) urea in the presence of the full BAM complex in *E. coli* polar lipids. This sample was then irradiated with a 6 W Hg-Xe benchtop UV light source (Uvitec) at 254 nm

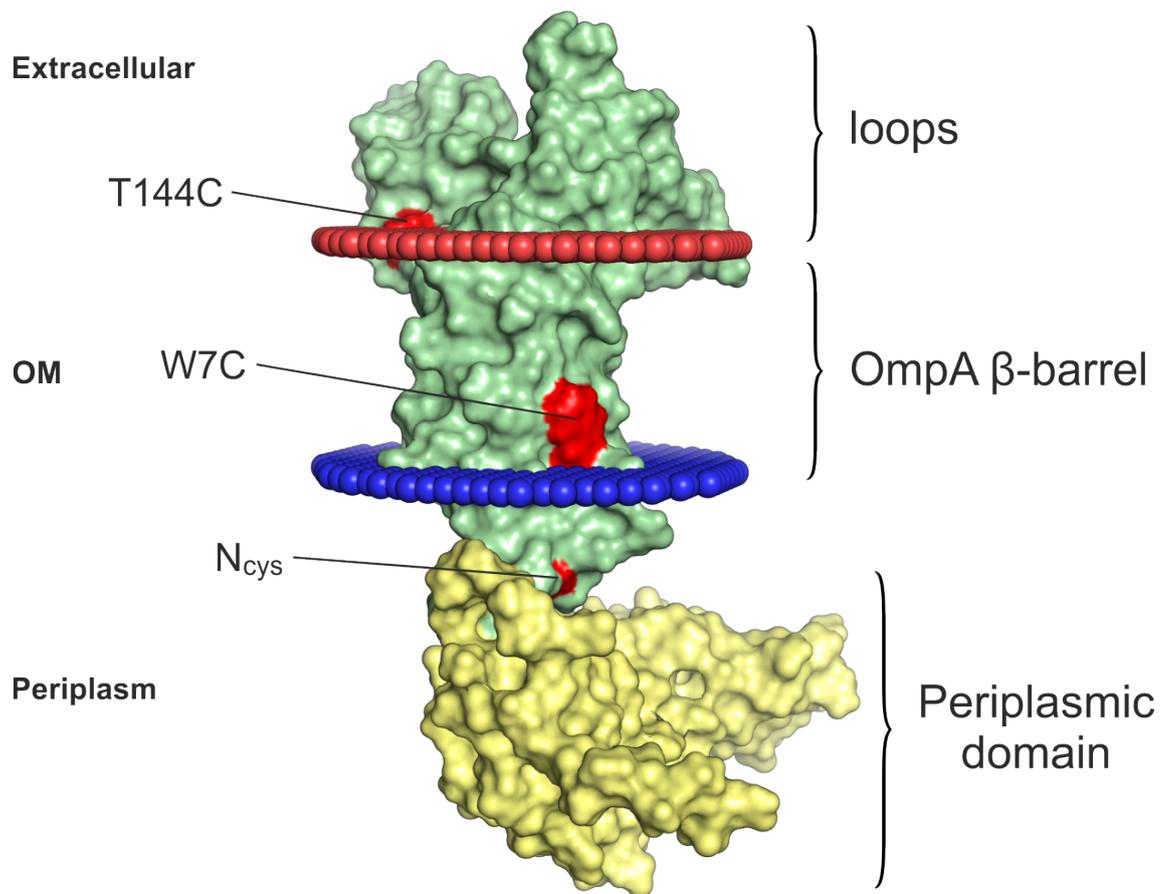


Fig. 3.3 Structure and orientation of OmpA. Surface representation of OmpA with the approximate locations of the three cysteine variants marked as red surfaces (N_{cys} , W7C and T144C). The approximate position of the membrane is shown by red (outer leaflet) and blue (inner leaflet) spheres (aligned from the Orientations of Proteins in Membranes database: <https://opm.phar.umich.edu/>). Transmembrane domain of OmpA shown in light green (PDB: 2GE4) (Cierpicki et al., 2006). Periplasmic domain of OmpA shown in light yellow (PDB: 2MQE) (Ishida et al., 2014).

for 30 min and then run on an SDS-PAGE gel to assess for crosslinked products. A BamA (88,326 Da) x OmpA(N_{cys}) (35,407 Da) crosslinked product would be expected to migrate at around 125 kDa but there is no strong band appearing with this migration (Figure 3.4A). When the contrast is enhanced very faint extra bands can be observed above BamA just under the 150 kDa band and at ~175 kDa, however, these are also present in the BAM and OmpA only lanes (Figure 3.4B) and may represent independent OmpA and/or BamA oligomers. Rationalising the lack of distinct crosslinks being observed three possibilities can be considered: i) a lack of interaction between OmpA and BAM at the N-terminus of OmpA, ii) the crosslinking reaction being too slow to capture folding intermediates, iii) a poor crosslinker yield or problems with the crosslinking chemistry. To address the third point a 'test' system which should definitely interact was needed. Skp is a trimeric periplasmic chaperone of OMPs which binds to OmpA with low nanomolar affinity (Qu et al., 2007). Skp binds substrates in a cavity formed by its jellyfish-like legs and holds them in an unfolded, dynamically tumbling state but with a lifetime for the substrate-chaperone complex on the order of hours (Burmam et al., 2013) (Figure 3.5). These properties make the OmpA-Skp substrate-chaperone pair an ideal testbed for assessing crosslinking strategies for dynamic and transient interactions.

OmpA(N_{cys})-BP was diluted from high (8 M) to low (0.27 M) urea in the presence of increasing concentrations of the periplasmic chaperone, Skp. This mixture was then irradiated with a 6 W Hg-Xe UV light source (Uvitec) for 30 minutes and run on an SDS-PAGE gel. A 55 kDa band appeared in the OmpA-Skp containing lanes corresponding to the mass of a crosslinked product of OmpA(N_{cys}) and a Skp monomer (17,986 Da) (Figure 3.6). The band corresponding to the cross-linked product was excised, digested, and the resultant peptides analysed by tandem mass-spectrometry (MS/MS) (for details see Materials & Methods 2.8.5.5). Despite multiple repeats, no crosslinks were able to be confidently assigned to positions in Skp at either the residue or peptide level. Intact denaturing MS of OmpA(N_{cys}) and the other cys variants of OmpA revealed variable undesired proteolytic cleavage of the N-terminal Met and Ala residues (Figure 3.9). This phenomenon has been observed by other groups who found that it is only present from inclusion-body purified OmpA and is absent when overexpressed proteins are targeted to the OM with a native signal peptide (A. Tsirigotaki, personal communication, March, 2018). While this would have no significant downstream affect for most variants, the OmpA(N_{cys}) variant, where a cysteine was introduced between Met and Ala (N_{term}-MCAP...), showed considerable heterogeneity and poor labelling efficiency due the variability of the size of this fragment in crosslinked peptides and its non-specifically proteolysed population (see Figure 3.7 for an illustration of the products of these in-gel trypsin cleavage reactions). For

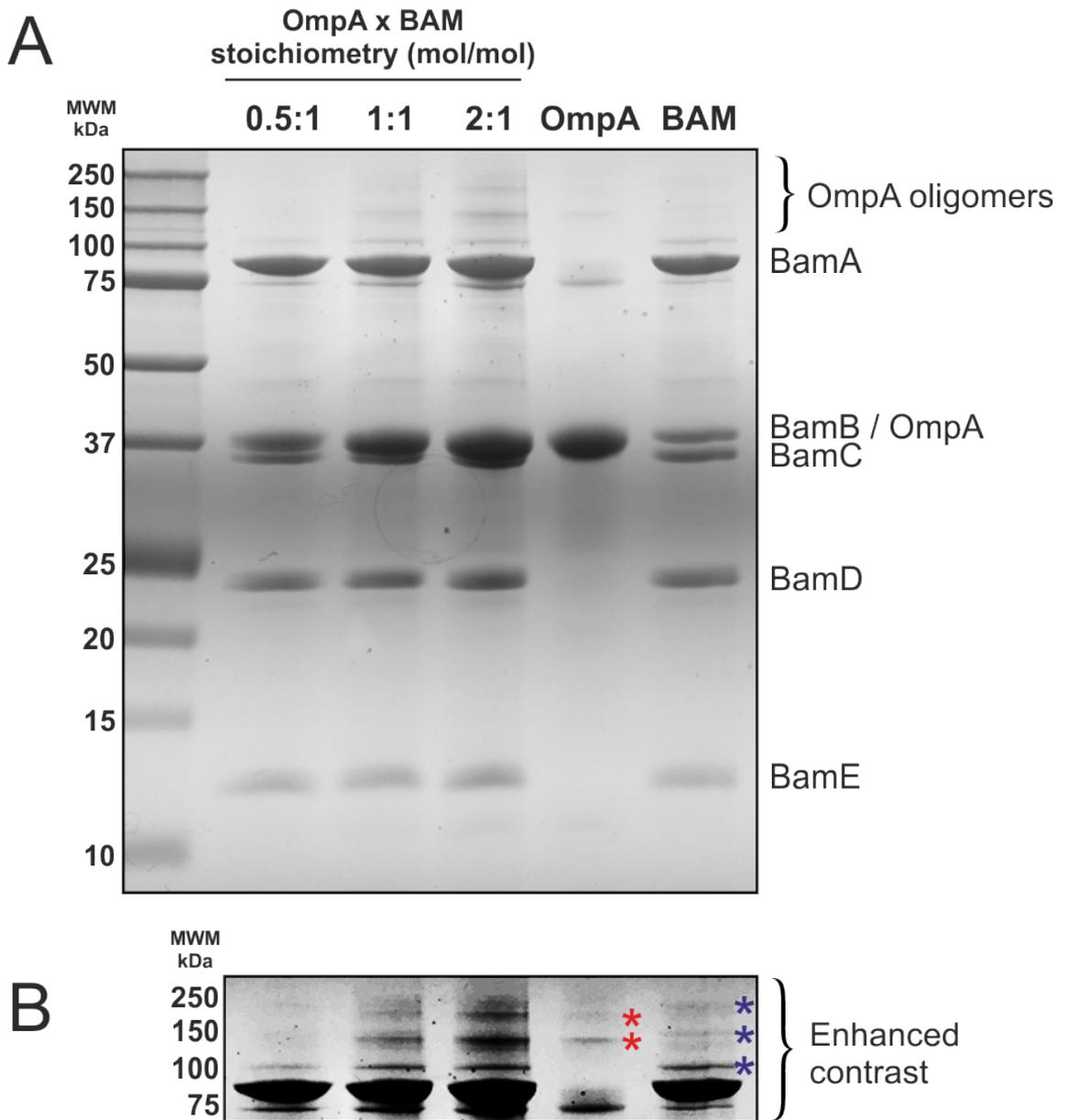


Fig. 3.4 Attempt to crosslink OmpA(N_{cys}) to the BAM complex in *E. coli* polar lipids. Final conditions: 50 mM glycine pH 9.5, 0.1–0.4 M urea (from 0.5:1 to 2:1), 3–12 μM OmpA (from 0.5:1 to 2:1), 3 μM BAM. OmpA only, 12 μM OmpA. BAM only, 3 μM BAM. (A) Full gel. (B) Enhanced contrast in region from 75–150 kDa to more clearly show the presence of low intensity bands at high molecular weights. High order bands in the OmpA-only, and BAM-only lanes are marked with red and dark blue asterixes, respectively. MWM, molecular weight marker.

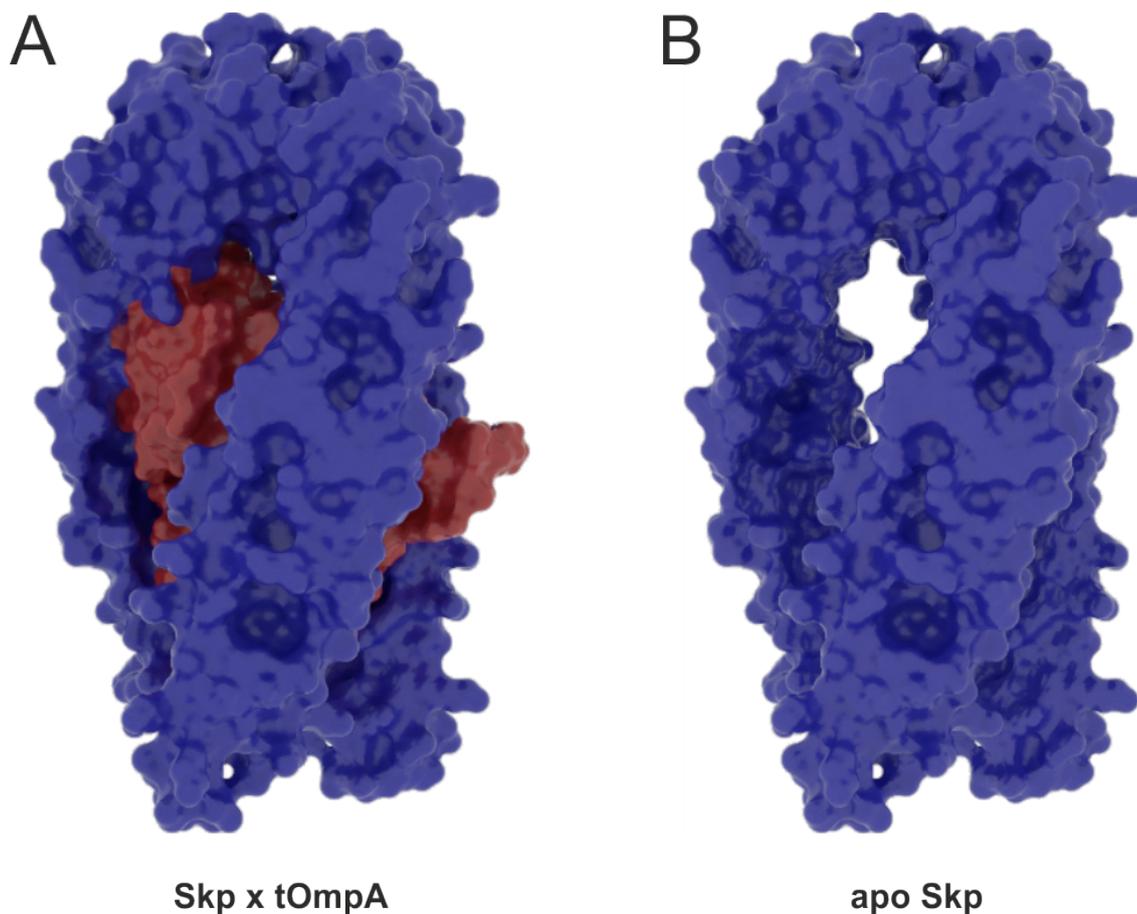


Fig. 3.5 Structure of Skp trimer in tOmpA-bound and apo states. Skp binds to OMPs during their transit across the periplasm and keeps them in an unfolded state within the cavity formed by the 'legs' of each monomer. Shown here is a model of a collapsed tOmpA chain sitting in this cavity. Missing loops of Skp (PDB: 1U2M) (Korndörfer et al., 2004; Walton and Sousa, 2004) were built in using Modeller (<https://salilab.org/modeller/>) (courtesy of Dr. Bob Schiffrin, University of Leeds), and tOmpA collapsed state was generated from a 3 ns *in vacuo* MD simulation starting from a model of the fully extended chain. Objects generated in PyMol and rendered in LightTracer (<https://www.lighttracer.org/>).

highly heterogeneous systems (such as those involving dynamic and transient complexes), it was apparent that this strategy would not be appropriate. For dynamic complexes, the large number of low-abundance crosslinks, the large search space of potential cross-linked sites created from a non-specific photoactivatable cross-linker, and the complexity of fragmentation patterns of cross-linked peptides combine to significantly degrade the ability to confidently identify true-positive matches using MS/MS. In transient complexes (such as an OMP folding through BAM), this is further compounded by the slow speed of crosslinking using standard mercury-xenon-based UV lamps which typically require reaction times of 30 min+ to reach completion (preventing time-resolved crosslinking, and poorly sampling transient interactions) and the biased reactivity of benzophenones towards certain amino acids (strongly toward Met but also Pro, Thr, Arg, Leu, Ser in water) (Wittelsberger et al., 2006; Dormán et al., 2016) prevent ‘true’ snapshots of the system at a moment in time. To overcome these issues, new strategies and new tools were required.

3.2.2 Validating the non-perturbing nature of a new crosslinker design on OmpA folding

To overcome these challenges, a new set of ‘tag-transfer’ crosslinkers were designed and synthesized. These tag-transfer crosslinkers use diazirine photochemistry for the reactive photoactivatable group, and methanethiosulfonate (MTS) chemistry for the site-specific conjugation to cysteine residues. Maleimide functional groups are widespread and allow the formation of an essentially irreversible covalent bond with thiol groups through Michael addition. However, these reactions can also occur on Lys residues and protein N-termini – particularly at higher pH values > 8, or where the protein local environment lowers the pKa of specific lysine residues (Sharpless and Flavin, 1966; Brewer and Riehm, 1967; Matos et al., 2018). MTS conjugation, on the other hand, is specific for thiol groups and results in the formation of a covalent disulfide bond that can be selectively broken in downstream applications by the addition of reducing agents such as β -mercaptoethanol (β ME), DTT, or TCEP.

Two kinds of diazirine-crosslinkers were synthesized (by Dr. Martin Walko, University of Leeds) (Figure 3.8A): i) the aliphatic MTS-diazirine which is smaller and so in terms of steric hindrance should be relatively non-perturbing, and ii) MTS-trifluoro-methyl-phenyl-diazirine (MTS-TFMD), which is larger, but has better photochemical properties (Preston and Wilson, 2013; Preston et al., 2014). These tags were conjugated to the OmpA(W7C) and OmpA(T144C) variants. OmpA(N_{cys}) was not used further due to the variable cleavage

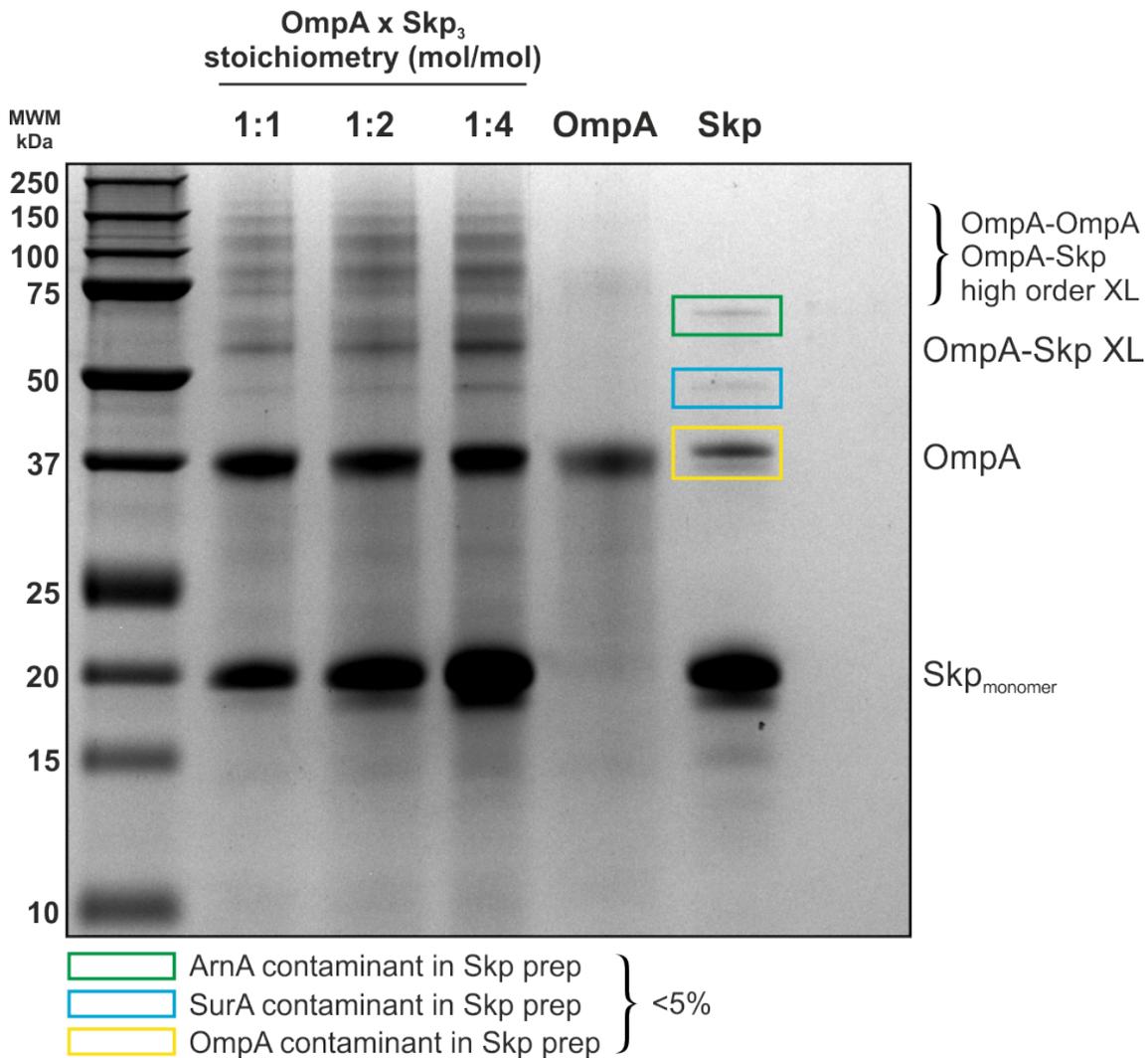
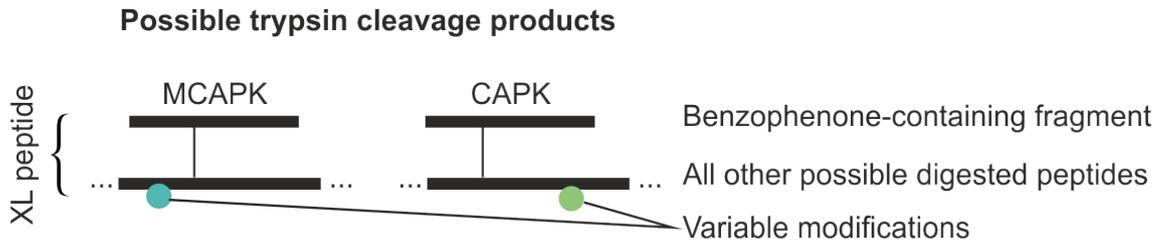


Fig. 3.6 OmpA can be crosslinked to Skp using site-specifically incorporated benzophenone. OmpA(N_{cys})-BP was rapidly diluted from high to low urea in the presence of Skp and irradiated with 254 nm UV light for 30 minutes. Final conditions: 50 mM glycine pH 9.5, 0.27 M urea, 4 μM OmpA, 4–16 μM Skp. Due to the high concentration of Skp loaded in the Skp-only lane, the presence of low-abundance contaminants becomes visible (identity by mass spectrometry analysis of the bands). These make up < 5 % of the total molar concentration of this sample (when considering band density with respect to molecular weight).



Skp = 38 peptides (1 missed cleavage, no mod)

OmpA(N_{cys}) = 53 peptides (1 missed cleavage, no mod)

Total = 91 peptides * 2 = 182 possible di-linked peptide precursors

Fig. 3.7 Illustration of products from a maleimido-benzophenone XL reaction between OmpA(N_{cys}) and Skp. The N-terminal degradation of OmpA(N_{cys}) meant that the peptide containing the site-specifically attached benzophenone varied in size, immediately doubling the search space. The product of a crosslinking reaction that results in a covalent bond formed between two side chains is a covalently linked di-peptide. During a data-dependent acquisition in a MS/MS experiment a fixed number of ions are selected and analysed, usually with a selection bias toward those with the strongest signal (i.e. the most abundant). These precursor ions are then fragmented (to form product ions) and subjected to a second round of mass analysis. As photocrosslinkers such as benzophenones can crosslink to one of many different residues in each peptide, this fragmentation pattern can vary even within the same precursor ion. The presence of variable modifications such as deamidation of Asn and Gln residues, oxidation of Met, or carbamylation of Lys and Cys, can further complicate the search space for identifying crosslinked peptides.

of the cysteine residue. High conjugation efficiencies were observed (>95 %) (Figure 3.9). To determine whether these small tag-transfer reagents alter OmpA folding, conjugated OmpA was folded into DDM detergent micelles or *di*C_{11:0}PC (DUPC) lipid membranes and its folding kinetics (Figure 3.10C), steady state fluorescence profile (Figure 3.10B), and the presence of a heat-modifiable SDS-PAGE band-shift (Figure 3.10A) was assessed. These experiments showed that the conjugation of a tag-transfer reagent to cysteine variants of OmpA does not prevent folding (as judged by SDS-PAGE and fluorescence spectroscopy) or alter kinetics (derived from time-resolved fluorescence spectroscopy). As these constructs are folding competent, the results derived from tag-transfer crosslinking reactions are now able to provide snapshots of its folding intermediates.

3.2.3 Accelerating crosslinking rates to capture transient interactions with a new UV LED lamp design

The rate of folding of OMPs via the BAM complex *in vivo* is not known, but must be a rapid process to be compatible with a doubling time of 20 minutes for *E. coli* and to prevent aggregation of nascent OMPs in the periplasm. *In vitro* BAM-mediated OMP folding into proteoliposomes of *E. coli* polar lipids occurs on a timescale of minutes to tens of

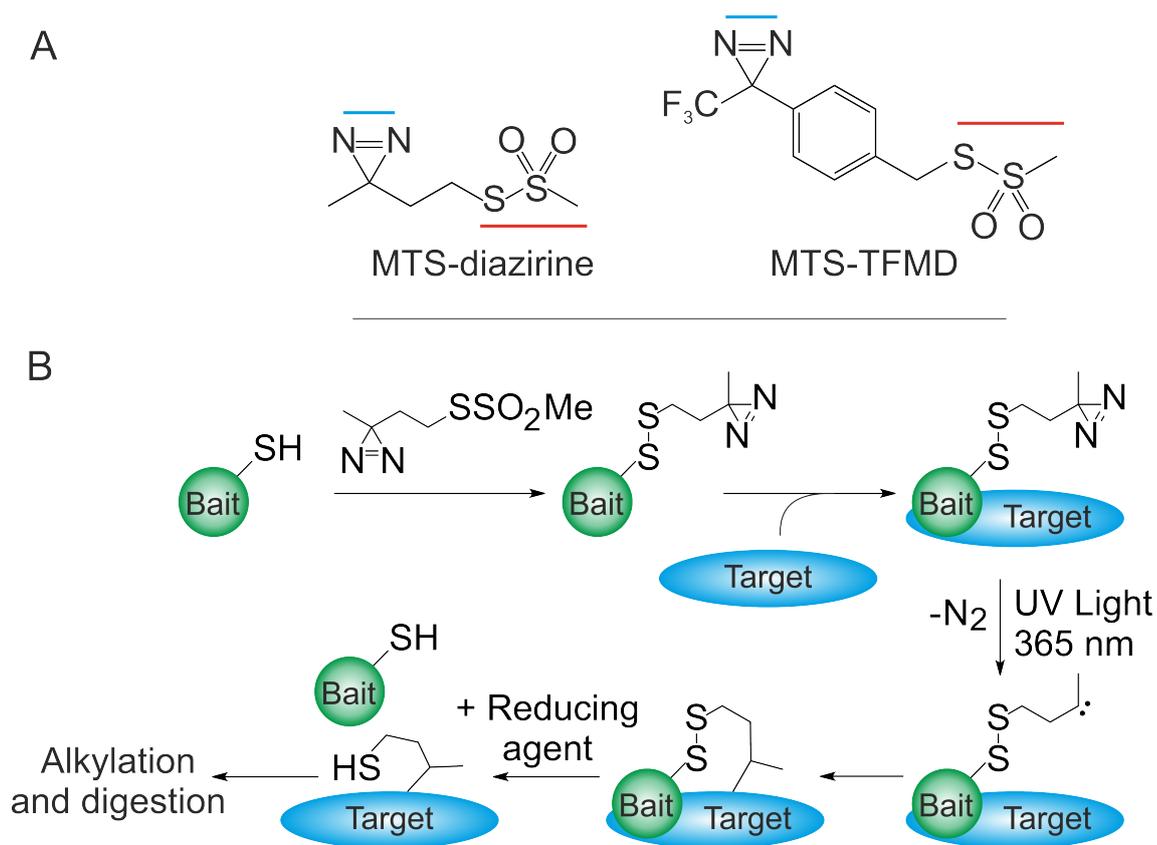


Fig. 3.8 Tag-transfer reagents and workflow. (A) Structures of MTS-diazirine and MTS-TFMD: red lines indicate the MTS active groups and blue lines indicate the diazirine photoactivatable group. (B) Crosslinking workflow schematic: A Cys-containing bait protein is conjugated with the reagent (here MTS-diazirine). After adding the target protein, the sample is irradiated with 365 nm UV light, revealing a carbene that reacts with the target. Reductant is added, leaving a sulfhydryl tag on the target at the interaction site.

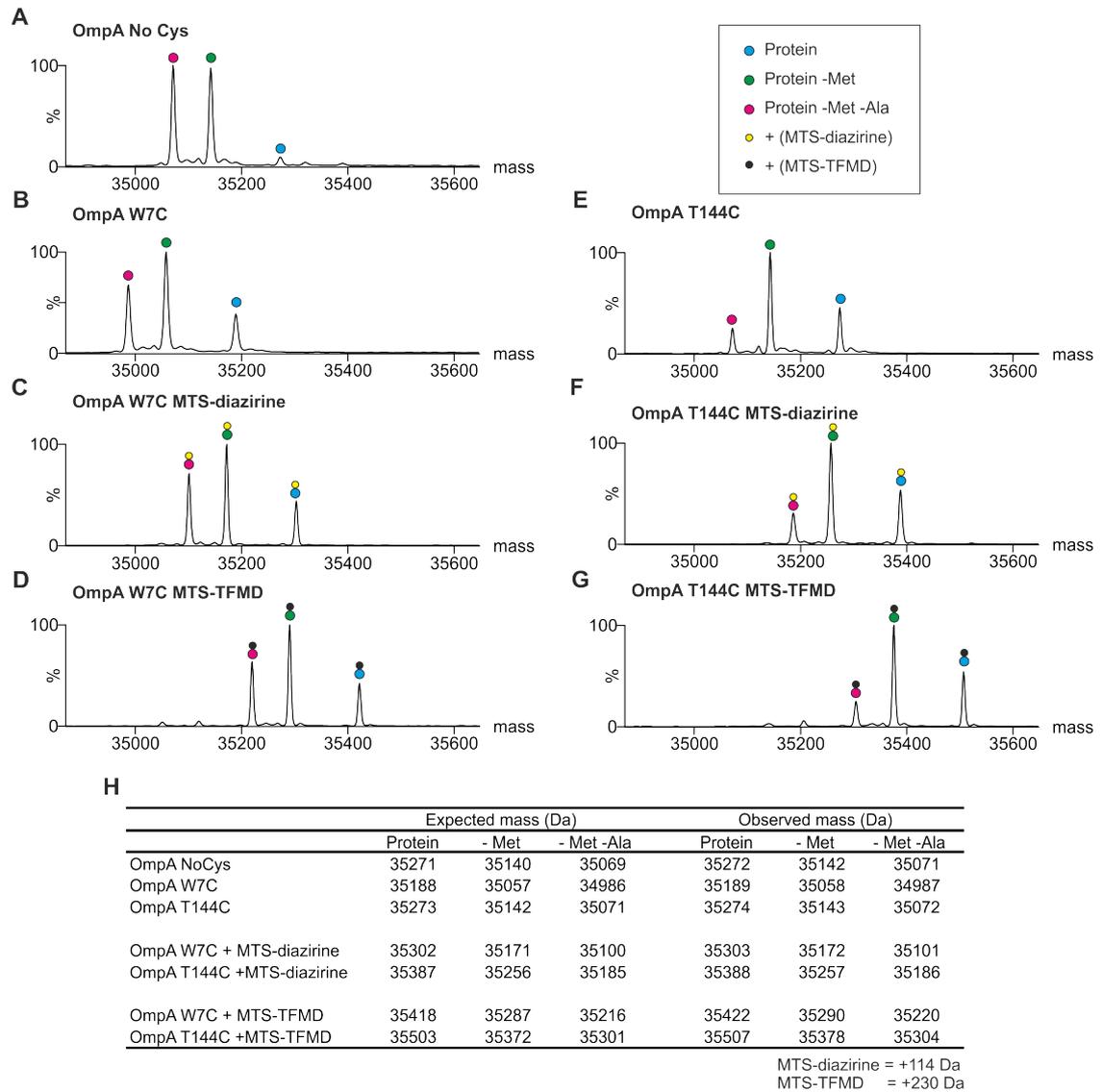


Fig. 3.9 Conjugation efficiencies of MTS-diazirine and MTS-TFMD to OmpA(W7C) or OmpA(T144C). Cysteine variants of OmpA were labelled with either MTS-diazirine or MTS-TFMD. Labelling reactions were performed for 1 h at room temperature, and intact MS was performed to determine the presence of labelled protein. (A) The deconvoluted mass distribution of OmpA purified from inclusion bodies shows variable loss of the N-terminal two residues (Met and Ala). (B-D) Deconvoluted mass distributions of (B) OmpA W7C, and OmpA W7C labelled with (C) MTS-diazirine or (D) MTS-TFMD. (E-G) Deconvoluted mass distribution spectra of (E) OmpA T144C, and OmpA T144C labelled with (F) MTS-diazirine or (G) MTS-TFMD. The distribution of peak intensities shows that 100 % labelling of both Cys variants of OmpA was achieved. (H) Table of expected and observed masses for the unlabelled and labelled OmpA proteins.

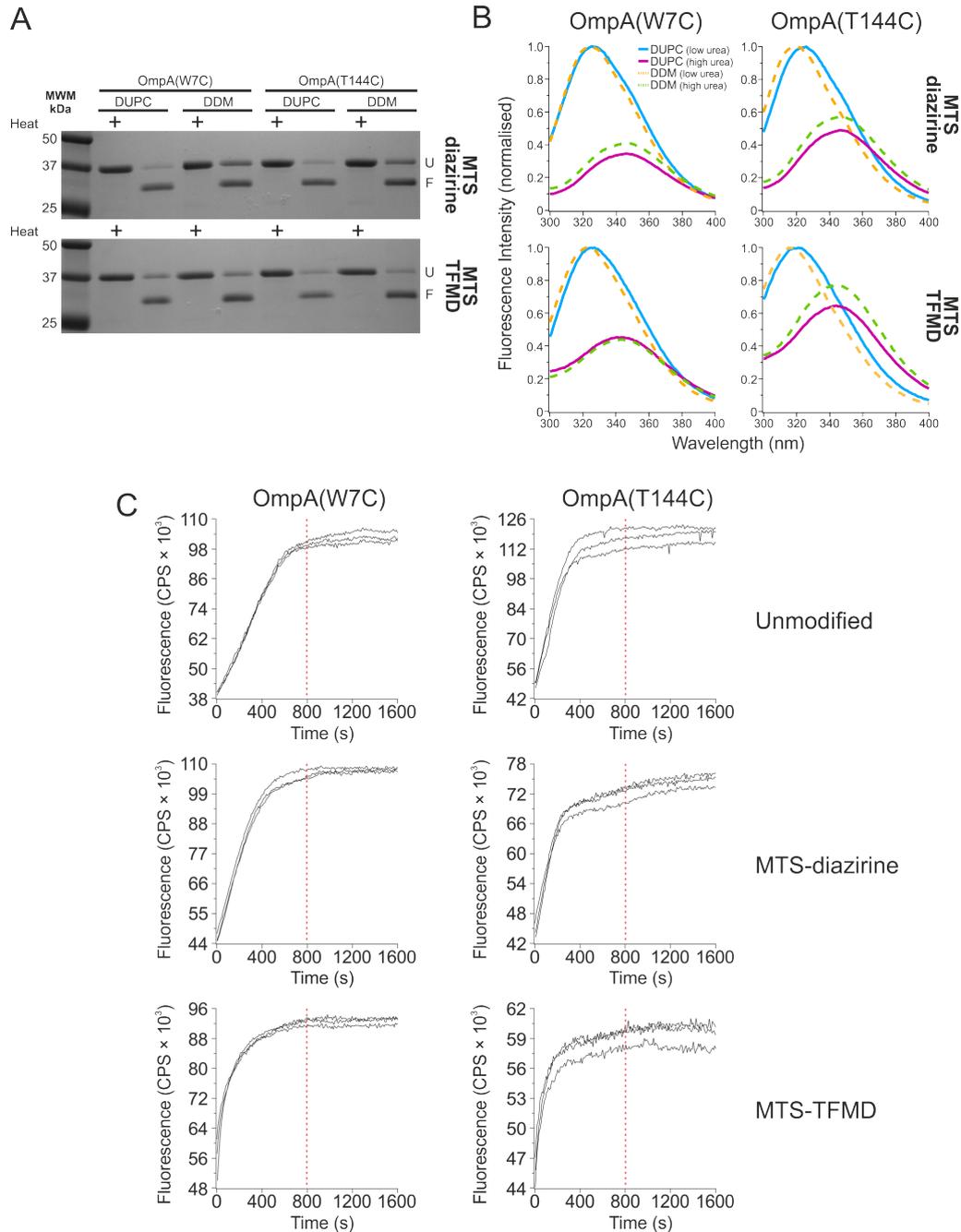


Fig. 3.10 Conjugation of tag-transfer crosslinkers does not perturb OmpA folding. OmpA(W7C) and OmpA(T144C) conjugated with either MTS-diazirine or MTS-TFMD are able to fold into 100 nm large unilamellar vesicles (LUVs) of *diC*_{11:0}PC (DUPC) or 0.05 % (w/v) DDM micelles. (A) Band-shift assay after overnight folding at 25 °C of 2 μM protein showing the presence of a heat-modifiable band characteristic of a folded OMP in all samples conjugated with the tag-transfer reagent. (B) Fluorescence emission spectra in low (0.8 M) and high (8 M) urea indicate the presence of a folded, membrane-embedded state. 0.4 μM samples were excited with 280 nm UV light and emission collected between 300–400 nm. 0.9 nm excitation slitwidths, 1 nm emission slitwidth, 1 s integration, 25 °C. (C) Kinetics of folding into DUPC as measured by fluorescence emission at 335 nm after excitation at 280 nm. Other conditions as in (B). Red dotted lines are to guide the eye to the same timepoint (800 s) after initiation of folding to aid comparison.

minutes for BamA, OmpT, and EspP (Hagan et al., 2010; Hagan and Kahne, 2011; Hagan et al., 2013; Roman-Hernandez et al., 2014; Hussain and Bernstein, 2018). While the microscopic rate constants for OMP folding are not known, when BAM is in stoichiometric excess over the OmpA substrate, folding is completed within tens of minutes in *E. coli* polar lipids. However, at a single molecule level it is entirely possible that an OMP will proceed rapidly through a BAM-associated intermediate to reach a slower intrinsic folding and insertion step. This means that very rapid crosslinking will be needed to catch the earliest stages of BAM-assisted OMP folding: recognition and insertion. The diazirine photoactivatable group provides the means to create a ‘snapshot’ of a system, as opposed to benzophenone photochemistry for which the photoactivatable group can activate and then diffuse before reacting with another amino acid, or deactivate to become reactivated again later (continuing the photograph analogy, this would be a ‘long exposure’ image) (Dormán et al., 2016). However, there is a second element controlling this snapshot – the speed at which the photocrosslinkers can be activated (more specifically, the proportion of photocrosslinkers activated per unit time). This should be directly proportional to the flux of light at the activation wavelength (around 364 nm) into the system. Therefore, to improve the ability to capture transient interactions a higher-powered lamp at 364 nm was needed.

3.2.3.1 Design of custom UV LED-based lamps

A custom-designed UV LED lamp was built by colleagues in the Department of Engineering at the University of Leeds (Dr. Mark Levenstein, Prof. Nik Kapur) according to these requirements. During the prototyping process, two different designs were built and tested. Both lamps comprise a UV LED with an emission profile centred at ~360 nm and a focusing lens to reduce the diffusivity of the light source and increase the power density at the sample plane to ~15 W cm⁻¹. The power level and ‘on’ time can be controlled electronically (Figure 3.11 and Table 3.1). The designs diverge in the layout and function of the sample holder compartment: design i) holds custom built acrylic chips designed for low volumes (30 µL) at high surface area-to-volume ratios; design ii) contains an Eppendorf holder that can be used in a ‘Russian doll’ configuration suitable for 1.5 ml, 0.5 ml, or 0.2 ml thin-walled PCR tubes. Both design i) and ii) are identical in their illumination profile but design (i) may be more appropriate when working with very low volumes and allows faster prototyping of acrylic chips for microfluidic mixing to initiate time-resolved reactions (such as protein folding) and reduce dead time. Another advantage of this custom design is that it can be easily assembled by someone with a passing

Table 3.1 Components and properties of the UV LED lamp

Part	Manufacturer	Model Number	Key Parameters
LED	LED ENGIN	LZ4-04UV00	Peak 365 nm wavelength, 1 A, 4.10 W
Lens	ThorLabs	ACL2520U-A	\varnothing 25 mm, f=20.1 mm, NA=0.60 ARC: 350-700 nm
LED Driver	eldoLED	PWR180D1	Input voltage: 12-32 V Output current: 1 A
Heat Sink	ILS	ILA-HSINK-STAR-50X60MM-RED-K	-
LED Cage	ThorLabs	CP02T/M	-
Lens Tube	ThorLabs	SM1L10	-

Table 3.2 Cost comparison between custom UV LED lamp and commercially available lamps

Manufacturer	Product	Initial Cost (£)	Source	Type
UVP, LLC	CL-1000	1590.00	VWR International	Hg-Xe
Boekel Scientific	UV Crosslinker AH (234100)	1556.93	Boekel Scientific	Hg-Xe
UVItec	UV Crosslinker CL-508	1575.00	Cleaver Scientific	Hg-Xe
ThorLabs	CS2010 UV Curing LED System	1580.49	ThorLabs	UV LED
This thesis	Minimal UV LED	276.08	Various	UV LED

familiarity with electronics from ‘off-the-shelf’ parts, and is considerably cheaper than commercially available lamps (Table 3.2).

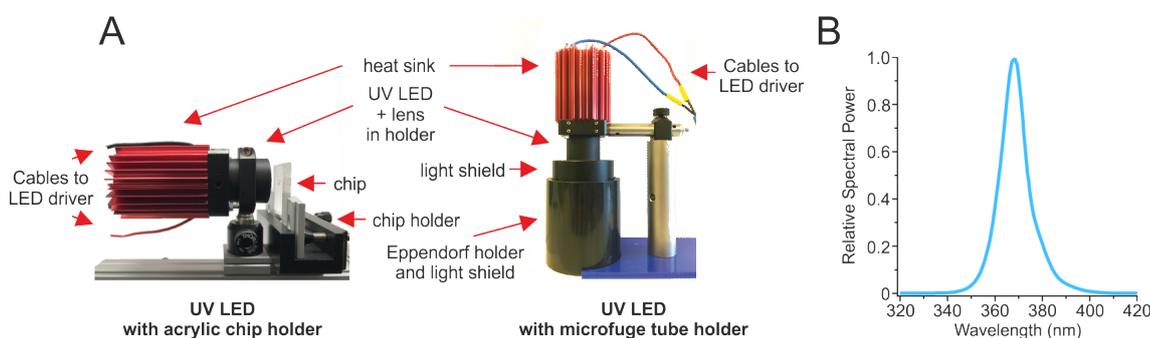


Fig. 3.11 Design of new UV LED crosslinking lamps. (A) Left: design (i) holds custom-milled UV-transparent acrylic chips with a 30 μ L sample well aligned to sit completely within the beam of the UV LED. Right: design (ii) can hold various sizes of microtubes inside the sample compartment and is lined with reflective foil to maximize illumination throughout the sample volume. (B) Emission profile of the UV LED showing an emission maxima aligned closely to the absorbance maxima of the photoactive group of diazirines (\sim 364 nm).

3.2.3.2 Lamp-associated sample heating across designs

Although crosslinking in 2–5 minutes has been reported (Weber and Beck-Sickinger, 1997; Zhang et al., 2011; Liang et al., 2017), these experiments were performed on ice to reduce heating from the very high-intensity mercury-xenon based lamps used to gain an appreciable amount of cross-linking. Furthermore, for crosslinking experiments where the goal is to assess the presence/absence of an interaction, it is not necessary to maximize the

Table 3.3 Degree of sample heating at fixed time points for different lamp designs

Time (s)	Temperature Change (°C)			
	UV LED (Tube) Thermocouple	UV LED (Tube) IR Camera	UV LED (Chip) IR Camera	Hg-Xe (Tube) IR Camera
5	1.0	0.4	0.9	-
15	1.8	0.9	1.8	0.6
30	2.8	2.1	2.6	-
300	16.9	15.9	8.5	5.7
1200	-	-	-	10.3

absolute yield of crosslinked materials as long as sufficient XL are formed to identify the proteins through proteomics using MS/MS. Where residue-level identification is required, maximising yield becomes more important to identify low-abundance interactions. To assess the degree of sample heating through a standard mercury-xenon lamp versus the UV LED lamp, the temperature of samples at different time points was measured using a fixed volume of water in a thin-walled PCR tube or in the custom acrylic chip. Infrared thermography was used to assess temperature changes of the whole system at fixed time points for each of the three experimental setups (6 W mercury-xenon strip lamp, UV LED lamp design (i), and UV LED lamp design (ii)) (see Figure 3.12) and a thermocouple thermometer was used to measure the temperature directly in the sample for the UV LED design (ii) (see Figure 3.13). These are summarised in Table 3.3. Both UV LED designs have similar sample heating profiles with ~1 °C of heating after 5 s, ~2 °C after 15 s, and ~2–3 °C after 30 s. However, this increases significantly by 5 min (300 s) where design (i) (Chip) has heated by 8.5 °C while design (ii) (Tube) has increased by ~16–17 °C. In contrast, the Hg-Xe strip lamp has only risen by 0.6 °C after 15 s (around 1/3 of the UV LED) and only reaches 10.3 °C after 20 min (1200 s). Pure temperature changes only tell a part of the story, however, as the high intensity of the UV LED lamps means that the crosslinking reaction has reached completion at a much earlier timepoint, and so shorter irradiation times can be used. Figure 3.14 illustrates that at the same yield of crosslinking, far less sample heating is incurred when using the UV LED lamp design compared with a Hg-Xe UV lamp. At the maximum observed yield (100 %), the UV LED designs show around ~2 °C heating, while over 10 °C heating occurs for the Hg-Xe UV lamp.

3.2.3.3 UV LED lamp dramatically improves the rate of crosslinking

The high reactivity of diazirines allows them to insert into water or buffer components – this is not observed in benzophenones and only weakly for aryl azides (Weber and Beck-Sickinger, 1997; Preston et al., 2014). This means that use of diazirines can capture

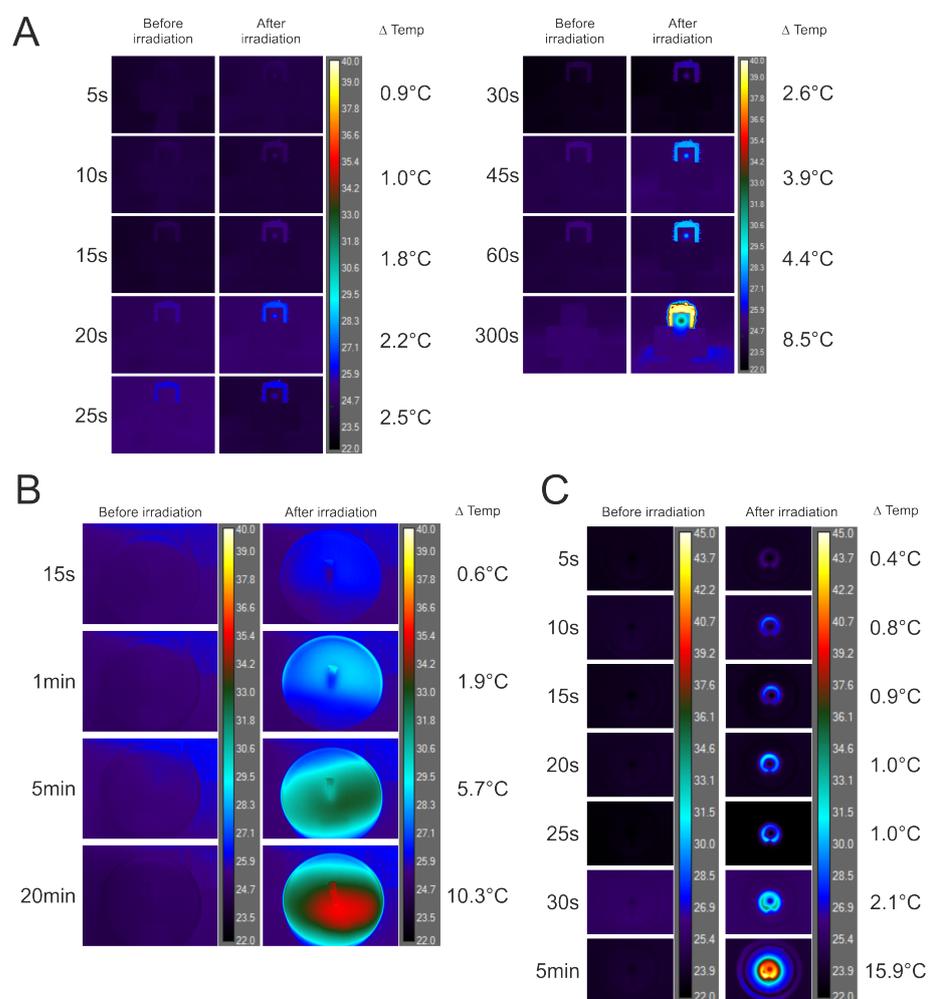


Fig. 3.12 Temperature response of sample and system during UV crosslinking. Infrared (IR) thermal imaging show that sample heating by the UV LED setup is small in the time window required for full crosslinking experiments as compared to that for a Hg-Xe lamp. (A) IR thermograms of the acrylic chip after irradiation by the UV LED. Each image is colour-normalised relative to the scale in the 5 minute image. The analysis was performed from the chip well, rather than the warm 'halo' in the images which is caused by the heat sink. (B) IR thermograms of before and after irradiation of a thin-walled PCR tube containing 200 μL of deionised water. The tube was placed on a transparent support and placed under a 365 nm Hg-Xe lamp for the indicated times. The change in temperature data was normalised against a strip of known emissivity placed on the support and the before irradiation images to correct for changes in ambient temperature and specular reflections in the room. (C) IR thermograms of the tube holder UV LED setup looking down onto the tube and holder before and after irradiation. The IR camera primarily measures infrared radiation from surfaces that are in its plane of focus. We found it challenging to accurately measure temperatures for the tube holder setup using the IR camera due to the narrow diameter of the tube and the depth of the liquid below the main surface of the tube holder. It was for this reason that a thermocouple was used for measuring the temperature in this setup.

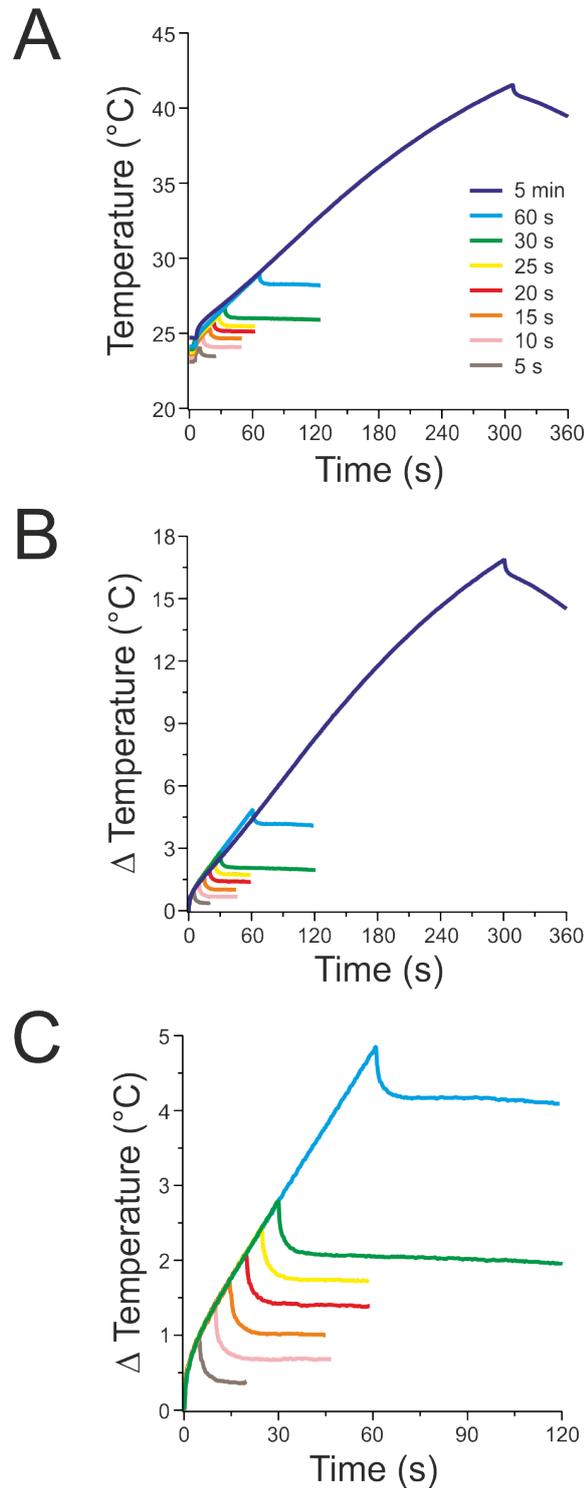


Fig. 3.13 Temperature response of tube sample during UV crosslinking. Thermocouple measurements show that sample heating by the UV LED setup is small in the time window required for full crosslinking experiments. This figure shows full thermocouple data for heating of a 200 μ L sample of deionised water in a thin-walled PCR tube. (A) raw, uncorrected data for 5–60 s timepoints as well as an extended 5 minute timepoint. The deviation of the 5 minute timepoint compared to 5–60 s may be due to heating of the UV LED source itself from extended use. Overheating can cause changes in the light intensity. (B) Same data as (A) but corrected for initial temperature and with the irradiation start points synchronised. The fall in intensity seen after the indicated times is due to the light source being turned off. There is an initial steep drop in temperature of approximately 0.5 $^{\circ}$ C followed by a gradual re-equilibration to room temperature. (C): Same data as (B) but zoomed in on the 5–60 s period.

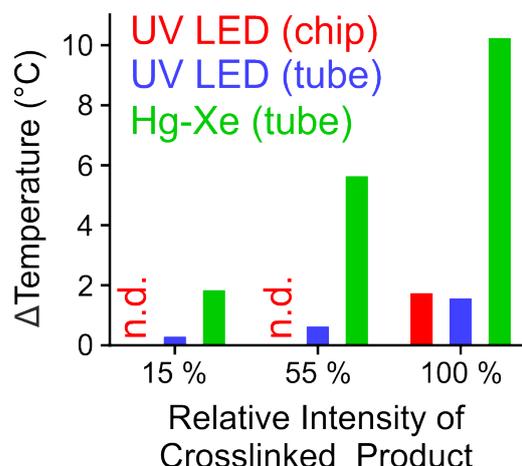


Fig. 3.14 Sample heating at irradiation times required to reach 15, 55, and 100 % maximal crosslinked product for the Hg-Xe and UV LED lamps. n.d. = not determined.

Table 3.4 Rate of crosslinking for each lamp design derived from crosslinking between OmpA(W7C)[MTS-diazirine] and Skp. The data shown in Figure 3.15 was fitted to a single exponential equation and the results of the fitting parameters are displayed here. y_0 , the y-intercept; A, amplitude; k_{obs} , the observed rate constant; t_{95} , the time taken to reach 95 % of the maximum crosslinking yield.

Parameter	UV LED (Tube)		UV LED (Chip)		Hg-Xe (Tube)	
		% error		% error		% error
y_0	0.9868 ± 0.0217	2.2	1.0155 ± 0.0231	2.3	1.0341 ± 0.0275	2.7
A	-1.0153 ± 0.0527	5.2	-1.0661 ± 0.0565	5.3	-1.0533 ± 0.0318	3
k_{obs} (s^{-1})	0.3875 ± 0.0472	12.2	0.3966 ± 0.0492	12.4	0.0027 ± 0.0003	9.5
Average Lifetime, $1/k_{\text{obs}}$ (s)	2.6	-	2.5	-	366.6	-
t_{95} (s)	7.8 ± 1.2	15.6	7.7 ± 1.2	16.0	1104.9 ± 200.5	18.1

a ‘snapshot’ of the local environment of the diazirine but it also means that absolute crosslinking yields can be low for dynamic systems. To measure crosslinking rates, OmpA(W7C) conjugated with MTS-diazirine was incubated with Skp, irradiated using either the UV LED lamp or mercury-xenon lamp, and illumination terminated at different time points. Samples were then run on an SDS-PAGE gel and the crosslinked band intensity measured by densitometry. These values were normalized against the crosslinked band intensity of the longest time point (Figure 3.15). The rate of crosslinking was well-described by a single-exponential, as would be expected for a XL reaction depending only on the concentration of crosslinker-conjugated protein (Table 3.4). The UV LED lamp was able to reach 95 % of the maximum yield in $7.7\text{--}7.8 \pm 1.2$ s (k_{obs} : $0.3875\text{--}0.3966 \pm 0.0472$ s^{-1}) compared with the Hg-Xe lamp which took 1105 ± 200 s to reach 95 % yield (k_{obs} : 0.0027 ± 0.0003 s^{-1}). This corresponds to a ~145-fold rate improvement over the Hg-Xe lamp and allows crosslinking to be completed on timescales consistent with OMP folding *in vitro* ($k_{\text{obs}} \sim 0.001$ s^{-1} for tOmpA folding into BAM *E. coli* proteoliposomes – Dr B Schiffrin, University of Leeds, data not shown).

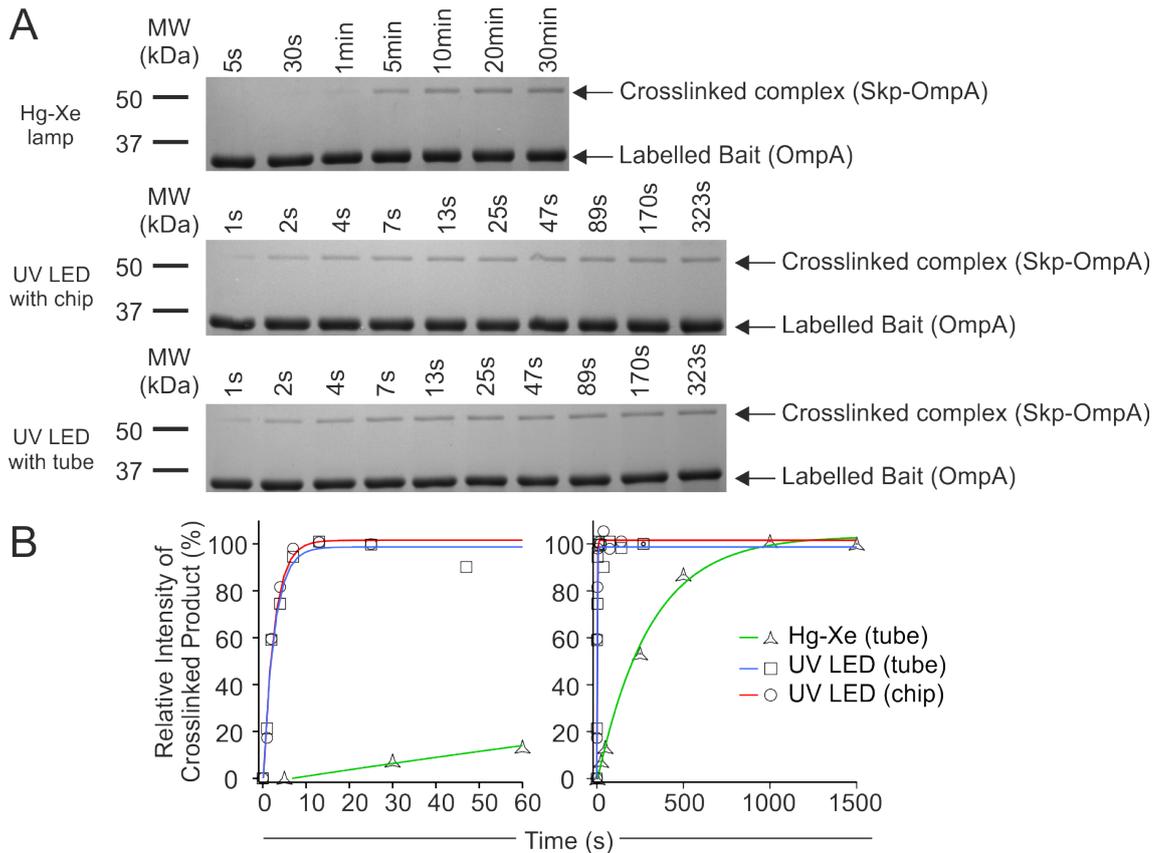


Fig. 3.15 Relative rates of crosslinking measured by SDS-PAGE. OmpA(W7C)[MTS-Diaz] was incubated with Skp and irradiated with UV light for the indicated time periods. Final conditions: 4 μ M OmpA, 24 μ M Skp, 0.8 M urea, 150 mM NaCl, 50 mM Tris-Cl pH 7.6. Irradiation was stopped and samples were run on an SDS-PAGE gel (A). Crosslinked band intensity was measured by densitometry and normalized to the longest time point. (B) This was then fitted to an exponential rate equation to derive rate constants and parameters to calculate the time to 95 % folded (t_{95}) (see Table 3.4). The graph on the left and the right are the same experiment, plotted on different scales to show the full length of the Hg-Xe timecourse.

3.2.4 Development of the tag-transfer workflow

The tag-transfer crosslinker designs take a fundamentally different approach to localizing crosslinks in XL-MS experiments. After irradiation with 365 nm UV light a crosslink is formed between a bait (here OmpA) and a target (here Skp) protein if there is an interaction in the local vicinity of the site-specifically installed tag-transfer reagent (Figure 3.8B). Importantly, this covalent crosslink contains a stable but reducible disulfide bridge as part of its structure. Depending on the aims of the XL experiments this reducible disulfide bond can be exploited by three main methodologies of increasing resolution (but also increasing number of preparatory steps). These are shown in Figure 3.16 and the development of each methodology is described below.

3.2.4.1 Strategies for transfer of an MS-stable tag to crosslinked target proteins

In Method 1, the simplest implementation of the tag-transfer workflow (Figure 3.16 - yellow box), the products of a XL experiment can be assessed by running them on an SDS-PAGE gel and looking for the appearance of a band not present in the isolated components and which disappears when run in the presence of a reducing agent (adding a further level of validation). In a heterogeneous mixture, the identity of the crosslinked product can be assessed by excising the gel band, digesting it with trypsin, analyzing the resulting peptides by MS/MS and then performing a database search to identify the proteins contained in the band. In a mixture of known components, the same approach can be taken to site-specifically assess whether interaction occurs from bait to target protein at that location on the bait – the unbiased reactivity of diazirines and the high stability of the crosslink formed makes this method more robust than typical disulfide crosslinking (where a cysteine is introduced into a position on both bait and target) or crosslinking by other functional groups (such as aryl azides or benzophenones). The disulfide bond can be exploited to perform higher resolution analysis of the exact location of the interaction on the target protein – to the exact residue or within a region spanning few residues. In a standard MS/MS workflow, after digestion proteins are reduced with DTT and alkylated with iodoacetamide (IAA) to cap any free thiol groups and prevent them from recombining in the mass spectrometer. As the tag-transfer reagents contain a disulfide bond, reduction of the bond regenerates the cysteine mutant of the bait protein, but leaves behind a small thiol-bearing tag on the target protein (MTS-diazirine: +87.02 Da after reduction, +145.06 Da after alkylation with IAA; MTS-TFMD: +204.02 Da after reduction, +261.04 Da after alkylation with IAA) which is also capped and stabilized by this alkylation step (Figure 3.17 and Figure 3.18).

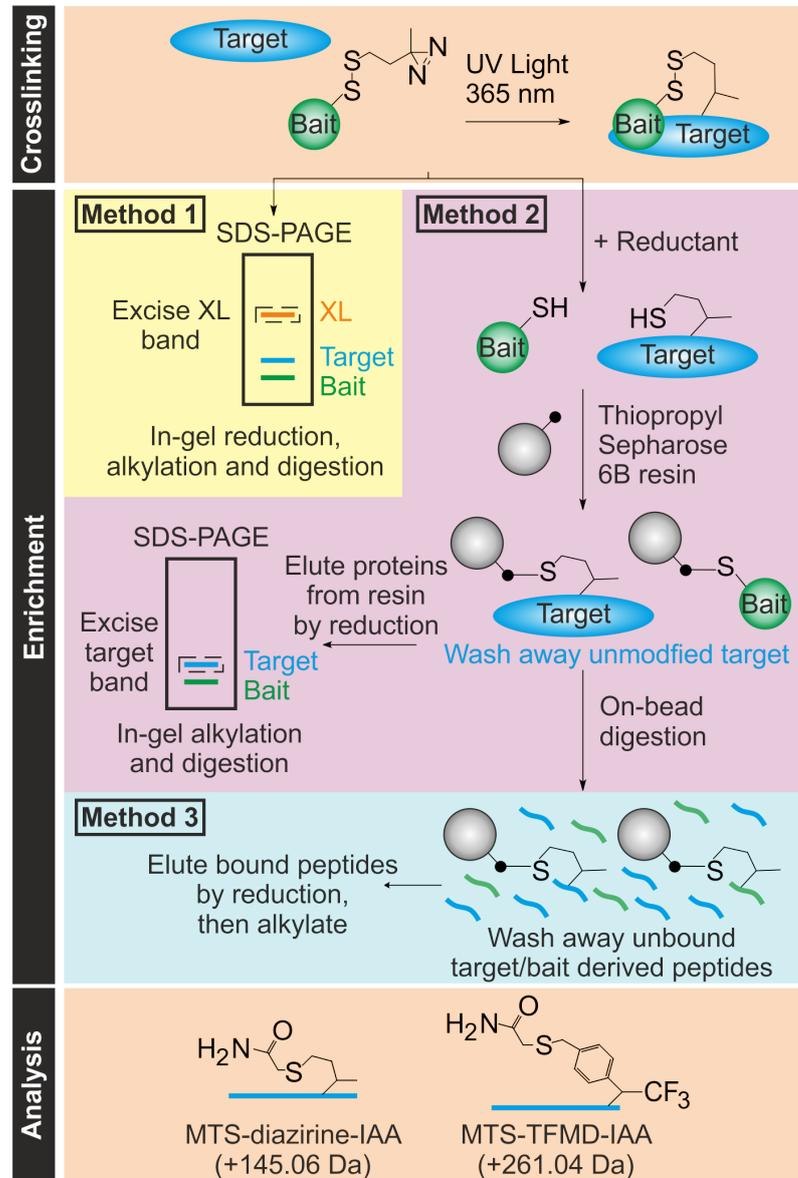


Fig. 3.16 Enrichment and analysis strategies for MTS-diazirine and MTS-TFMD. Irradiation at 365 nm results in crosslinking of bait and target proteins (n.b. low reaction efficiencies mean that uncrosslinked material will remain). Enrichment can be performed using one of three methods. MS analysis of the peptides is then performed, and the data searched to identify the peptides/residues modified with the crosslinking reagent (with the free thiol capped by reaction with IAA).

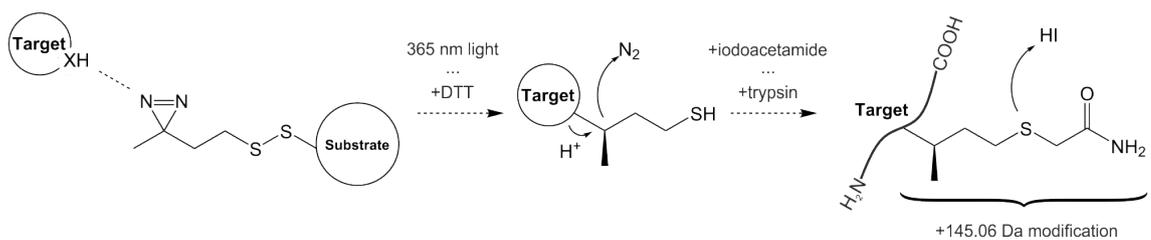


Fig. 3.17 Reaction scheme for crosslinking and alkylation by MTS-diazirine.

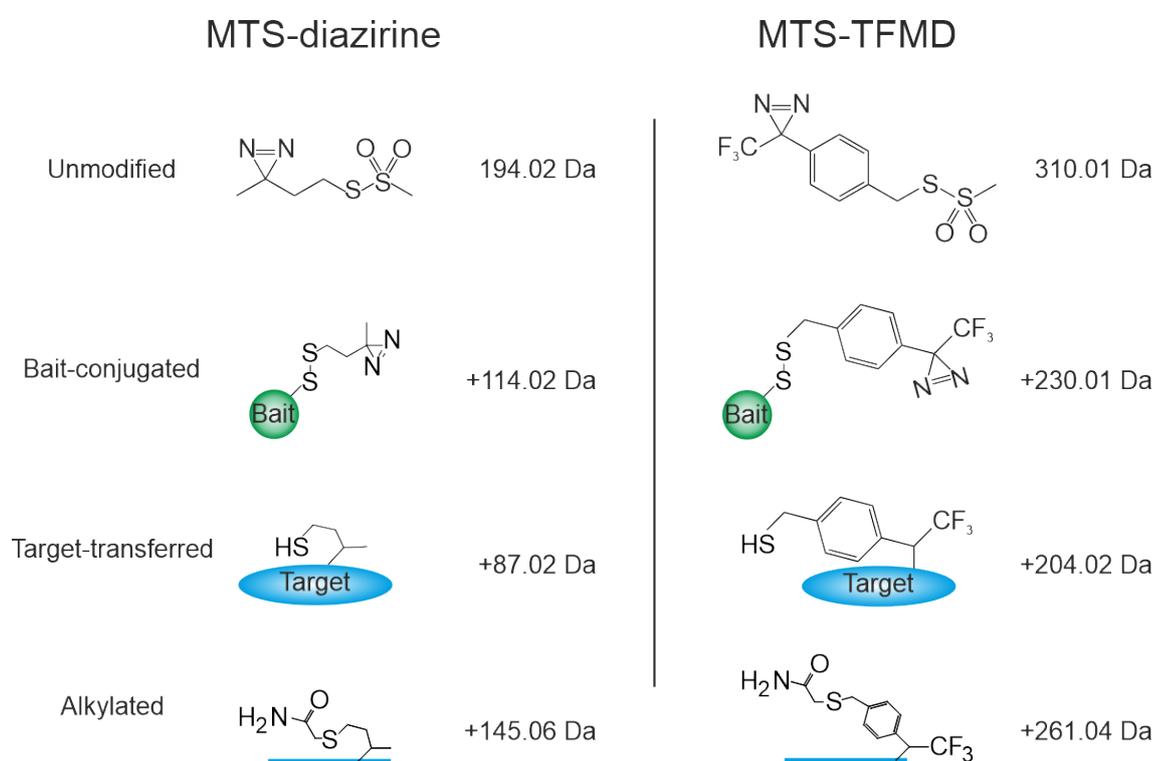


Fig. 3.18 Masses and structures of crosslinkers, derivatives, and tags used throughout this methodology. The raw MTS- and diazirine- containing crosslinker masses and structures are shown as well as the resulting structures and mass adducts after conjugation, tag-transfer, and alkylation steps.

The complicated fragmentation patterns of cross-linked peptides – the product of common photocrosslinkers – requires specialist software tools and algorithms to identify these peptides by MS/MS and calculate the positions where crosslinking occurred (Yilmaz et al., 2018). In these cases, because the site of interaction is unknown *a priori* for both peptides, the algorithmic search space scales as n_{tot}^2 (where n_{tot} is the number of crosslinkable residues in the bait and target combined). The fixed-mass adduct of the alkylated tag-transfer crosslinkers simplifies this search in three ways. Firstly, knowledge of the location of the crosslink on one of the partners (that is, the location of the Cys at which the tag-transfer reagent was installed) reduces the search to n_{target} – the number of crosslinkable residues in the target, which for diazirines is all residues. For small proteins or peptides, the difference between n_{tot} and n_{tot}^2 is small but it rapidly increases with increasing protein size and/or number of components. For example, consider an experiment with the chaperone Skp as target, and its substrate OmpA as bait. Bis(sulfosuccinimidyl)suberate (BS3) is a popular crosslinker which uses the well-established reactivity of *N*-hydroxysulfosuccinimide (NHS) esters towards primary amines (such as Lys and protein N-termini) on either end of an ~ 11.4 Å linker arm to bridge lysine residues intra- or inter-molecularly. Skp (with 17 Lys) and OmpA (with 15 Lys) have a search space of 1024 possible cross-linked peptides (32^2) (Figure 3.19). However, with tag-transfer methods this is reduced to 161 (the length of Skp) as we can ignore the contribution from OmpA (although we would still search OmpA separately as a control – this would only increase the search space to 486 [+325 OmpA residues]). Secondly, as a corollary to the first point, the addition of a mass adduct to residue side-chains or backbones (as opposed to a covalent bond to a second peptide) leads to much more efficient fragmentation in the second stage of MS/MS and simpler spectra making automated assignment of peptide identities much more rapid and robust. Finally, the tag-transfer adduct can be treated like a variable post-translational modification to peptides – a search task for which there are many well-established (and well-maintained) commercial algorithms and software available. Performing these tag-transfer crosslinking experiments using Method 1, and analysing the crosslinked product directly excised from an SDS-PAGE gel band, it is possible to identify 6 unique peptides corresponding to crosslink locations clustered in two discrete regions near the middle of the 'legs' of Skp (see Figure 3.21 Top, for the location in the 1D sequence, and Figure 3.23B, for these locations mapped onto the crystal structure of Skp). A list of all modified peptides for all Skp experiments in this chapter can be found in Appendix A. The resolution of the tag-transfer approach allowed 4 of these peptides to have the location of crosslinking narrowed down to an individual residue (Figure 3.21 Bottom).

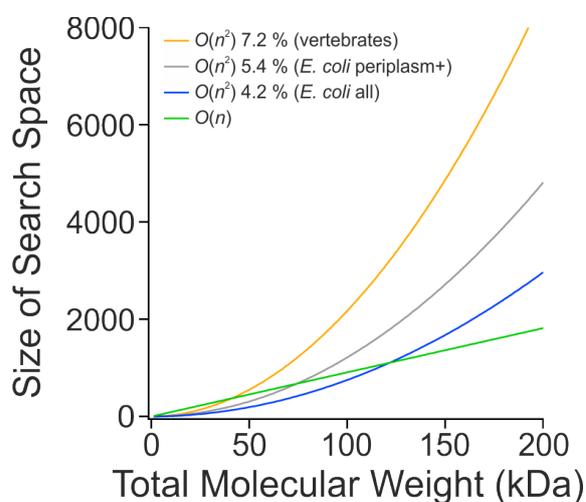


Fig. 3.19 Growth of search space for BS3 versus tag-transfer crosslinkers. BS3 is a homobifunctional crosslinker that crosslinks two primary amines (such as Lys) in the same or different proteins. This mechanism of action means that the search grows as $((n^2 - n)/2)$ where n is the total number of crosslinkable residues (here, number of Lys) across all proteins in the XL-MS experiment. As this is protein-dependent, the average number of Lys in vertebrates, *E. coli* K-12 secretome (periplasm+), and *E. coli* K-12 total, have been used to demonstrate some representative proportions of Lys. This has been scaled to 'total molecular weight' (sum of all protein molecular weights in an experiment) using the approximation of 110 Da per amino acid. For the tag-transfer crosslinkers, the exact position of one site is known, and the diazirines can react with all amino acids sidechains, so the search space grows as n where n is the total number of residues. Big O notation used to represent the growth of algorithmic complexity of each method is indicated on the graph.

3.2.4.2 Improving signal quality and number of PSMs with novel enrichment strategies

3.2.4.2.1 Sepharose bead-based enrichment with Method 2 For systems where the target and/or bait protein is large (e.g. forming a complex >50–70 kDa), unmodified peptides from the crosslinked band will contribute considerable background to the experiment, degrading the resolution – particularly for low-abundance crosslinks. This is where Method 2 can improve on current approaches (Figure 3.16 - pink box). Consider a situation where tag-transfer labelled OmpA is crosslinked to Skp and ran on a gel (as shown in Figure 3.7 for maleimido-benzophenone crosslinkers): for each molecule in the crosslinked band 91 peptides are generated but, as each OmpA bait protein is labelled with just a single tag-transfer reagent, only 1 peptide per complex will be modified ($1/91 \approx 1\%$). For a larger complex, such as between OmpA and BamA, this is even more significant with only $\approx 0.5\%$ of peptides being modified (trypsin digest with 1 missed cleavage: OmpA 53 peptides + BamA 126 peptides = 179 peptides) and assuming 100% cross-linking efficiency. Should the complex be dynamic, the population of each modified peptide in a mixture would be even lower in abundance. To overcome this the free thiol transferred by the tag-transfer reagents can be exploited as an affinity tag. Thiopropyl Sepharose[®] 6B (GE Healthcare) is a Sepharose-based resin bearing a β -thiopyridyl disulfide group that can undergo a disulfide exchange reaction with free thiols, attaching thiol-bearing proteins, peptides, or small molecules, and liberating β -thiopyridone. To make use of this reagent, the product of the crosslinking experiment is buffer exchanged into buffer containing a high concentration of reducing agent (100 mM β -mercaptoethanol or 50 mM DTT) and incubated to allow complete reduction of the disulfide bond in the linker arm (see Materials & Methods Section 2.8.5.3 for a detailed protocol). A second buffer exchange step removes the reducing agent before the sample is incubated for 30–60 minutes with resin. The resin is then thoroughly washed before the bound proteins are eluted by the addition of a high concentration of DTT (>50 mM) to saturate all binding sites on the resin. Running this eluate on an SDS-PAGE gel, only target protein which had a tag-transfer thiol appended to it during crosslinking appears in the corresponding band. The bait protein is not removed by this process as it contains the cysteine used for initial installation of the tag-transfer reagent. This method assumes that the molecular weight of the bait and target are different enough to be separated on an SDS-PAGE gel, and that the target protein contains no native cysteine residues (or that these have been mutated out). The output of each step in this workflow is summarized on an SDS-PAGE gel in Figure 3.20. The band (lower band highlighted in red in Figure 3.20) no longer contains the contribution of peptides derived from the bait protein, improving the signal-to-background during MS analysis. This results in the identification of 14 unique peptides spread across the sequence

of Skp, including 7 where the location of the crosslink could be confidently determined to a single residue (see Figure 3.21). Compared to the results described for Method 1, the interaction between OmpA and Skp now appears to comprise 4 clusters spread across the 'legs' of Skp (Figure 3.21 Top for the location in the 1D sequence, and Figure 3.23B for these locations mapped onto the crystal structure of Skp).

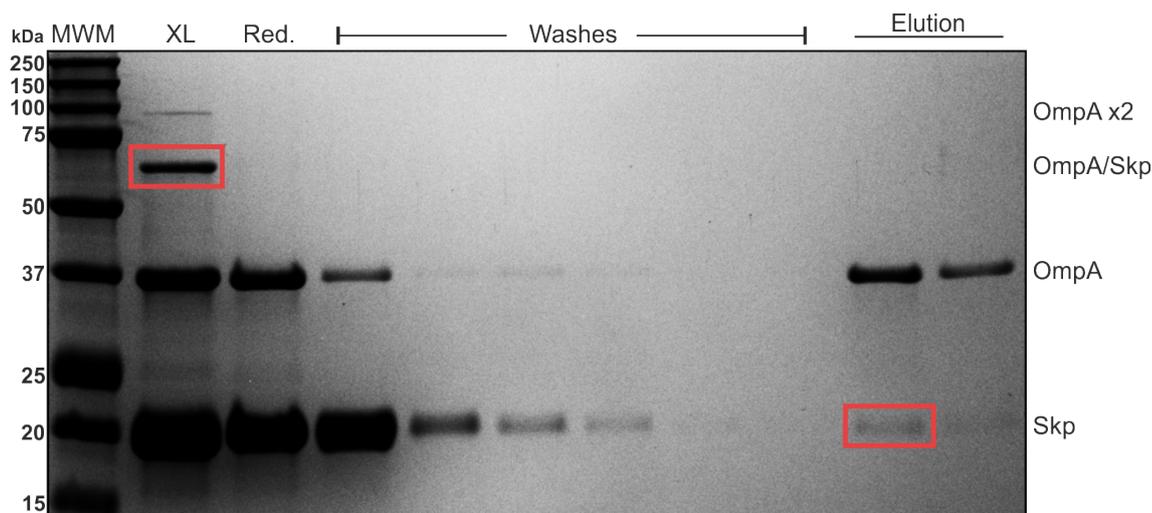


Fig. 3.20 Output of each stage in tag-transfer processing and enrichment. An aliquot from each step of Method 2 (and periodic samples from the wash steps) were run on a non-reducing SDS-PAGE gel. The bands to be analysed in Method 1 or 2 are highlighted in red. MWM, molecular weight marker. XL, crosslinked samples before addition of reducing agent. Red., same as XL but with reducing agent added. Washes, supernatant collected from wash steps taken at increasing numbers of washes as you move from left to right. Elution, supernatant collected after addition of reducing agent to elute samples from the Thiopropyl Sepharose beads. First and second elution lanes represent the initial elution and a second elution step to remove any protein still retained after the first.

3.2.4.2.2 On-bead digestion with Method 3 The third and final method, Method 3 (Figure 3.16 - blue box), is an extension of Method 2 in which the Thiopropyl Sepharose[®] 6B-conjugated proteins are incubated with trypsin on resin after the wash stage. This on-bead digestion creates peptides which can be washed away, leaving behind only peptides which were initially crosslinked (and thus bear a transferred thiol tag that is disulfide-bonded to the resin) and those which contain a cysteine residue (such as the region of the bait protein that the tag-transfer reagent was conjugated to). This means that irrespective of the size of the proteins in the experiment, only 'informative' peptides and the single peptide containing the cysteine residue of the bait, will be enriched for downstream MS analysis. This reduces the likelihood of a high background signal from unmodified peptides masking the presence of a low-abundance modified peptide, especially important if both elute at similar times from the initial LC separation step before injection into the mass spectrometer. After this digestion and wash step, specifically bound peptides

are eluted with a high concentration of DTT or β ME (see Materials & Methods 2.8.5.4), desalted using C18 media (ZipTip[®] pipette tips), and eluted in a volatile medium for drying before resuspension in a mass-spectrometry compatible buffer or storage at -80 °C. This method results in the identification of 35 unique peptides with 21 of those giving confident assignments of the interaction on Skp to the residue level (see Figure 3.21 Bottom). These are spread across the whole sequence of Skp (Figure 3.21 Top), distributed throughout the 'legs' as seen in Figure 3.23B. Example mass spectra of peptides derived from this method for both tag-transfer reagents are shown in Figure 3.22.

3.2.4.2.3 Comparison of analysis methods The OmpA/Skp substrate-chaperone complex is an excellent test-bed to assess the performance of each method at detecting the interactions between a dynamic interface. Figure 3.21 shows the number of unique modifications detected in XL-MS experiments on OmpA(T144C) with MTS-diazirine conjugated (see Figure 3.3 for the position of T144C in one of the loops of OmpA) being crosslinked to Skp. For a complicated system of this kind, using Method 1 cross-linked peptides were identified at the residue level (shown in red in Figure 3.21) or to the sub-peptide level (blue in Figure 3.21). MS-MS sequencing of peptides can result in incomplete fragmentation or poor detection of some fragments which means that the presence of a tag-transfer modification cannot always be confirmed down to an exact residue. However, by combining the known mass of the precursor ion from the first stage of MS and the size of the detected fragments, the identity of the peptide can be determined and the location of the modification narrowed down to two or more residues (its presence being known due to an increased mass of +145.06 Da or +261.04 Da of the precursor - see Figure 3.18). This is referred to here as a 'sub-peptide'. By enriching the target protein using Method 2 the number of modified peptides detected increases >2-fold (from 6 to 14) whilst Method 3 improves this a further 2.5-fold (from 14 to 35) to give a ~6-fold increase in detection over Method 1. However, a perhaps more important aspect of these differences is the distribution of the detected modifications. Using Method 1 only, an experimenter may interpret the OmpA-Skp interaction to be localized and specific to a few regions of the chaperone but as the resolution is increased and low abundance interactions are able to be detected, it becomes clear that this single position in OmpA (that is T144C in loop 3 of OmpA) makes contacts all over Skp - a model which is incompatible with a localized interaction. Example mass spectra of peptides derived from this Method 3 for both tag-transfer reagents are shown in Figure 3.22.

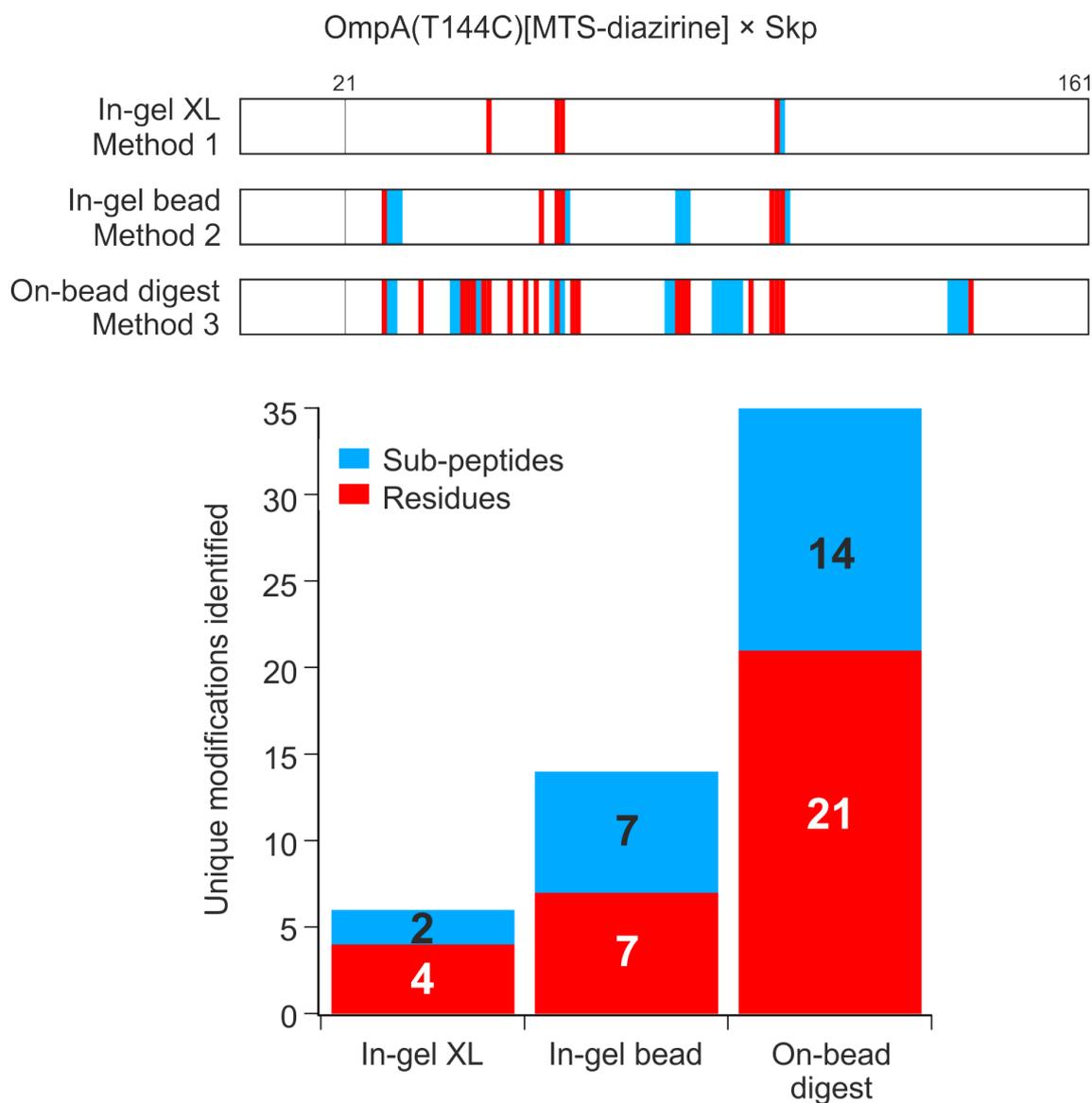


Fig. 3.21 Number of unique cross-linked residues detected on Skp for each tag-transfer method shown in Figure 3.16 with OmpA(T144C)[MTS-diazirine] as the bait. Residue-level confident identifications are shown in red, sub-peptide level identifications are shown in blue. Top: the distribution of unique crosslinks identified for each method mapped onto the primary structure of Skp (shown as a box - residues 1–21 comprise the His-tag and linker region). Bottom: the total number of unique sites identified by each enrichment strategy.

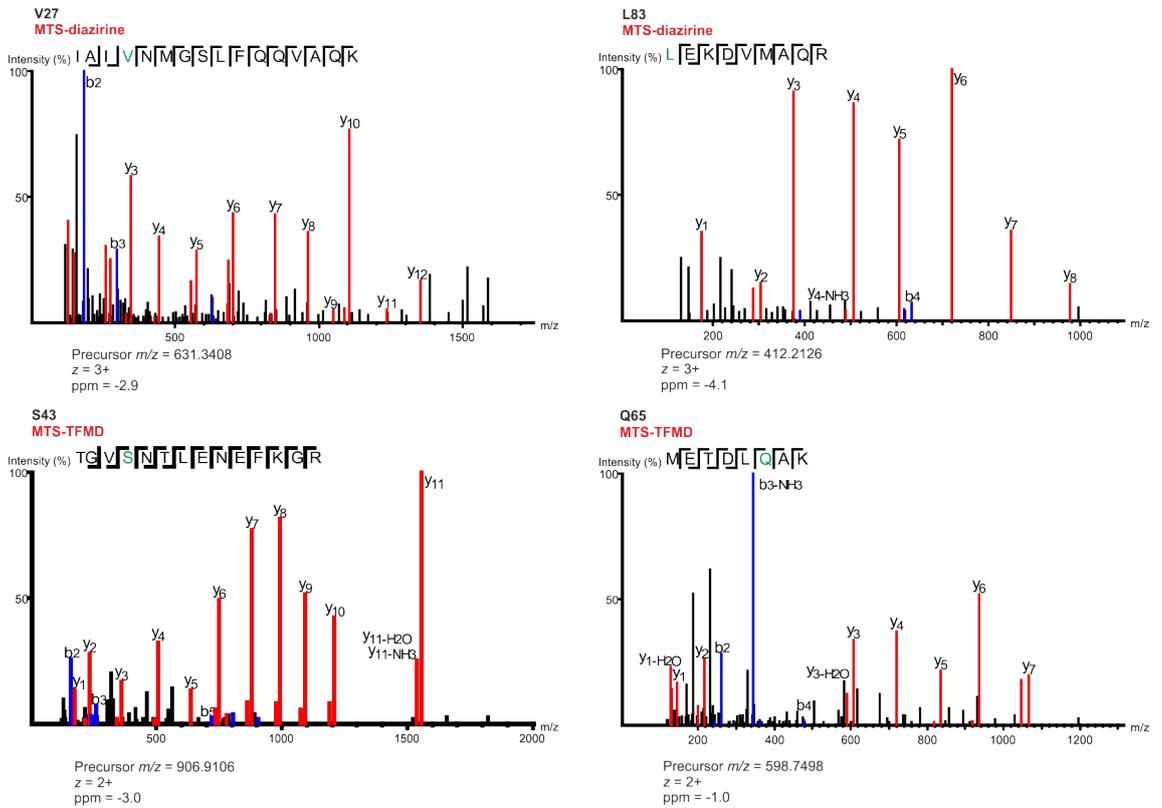


Fig. 3.22 Representative tandem MS mass spectra of peptides derived from Skp and enriched using Method 3. These peptides contain the modification associated with tag-transfer from MTS-diazirine or MTS-TFMD labelled OmpA. The site of modification is indicated above the spectrum and in green in the peptide sequence. The precursor m/z , charge state and mass error (ppm) are shown. y-ions are shown in red, b-ions in blue, unassigned peaks in black.

3.2.5 Interrogating the interaction between OmpA and Skp

3.2.5.1 OmpA tumbles inside the cavity of Skp in a dynamic, disordered state

Using Method 3, the interaction between OmpA and Skp can be probed at high resolution - down to the residue level. To validate the tag-transfer method for both tag-transfer reagent designs (MTS-daizirine and MTS-TFMD – see Figure 3.8A) and to determine any further insights into the well-studied mechanism of Skp chaperoning of OMPs, tag-transfer XL with enrichment (Method 3) was performed using OmpA(W7C) and OmpA(T144C) bound to Skp, using both tag-transfer reagents (for a total of 4 datasets - 2 mutants x 2 tag-transfer reagents). The results of these experiments are shown mapped onto the Skp trimer for a pooled dataset of OmpA(T144C)[MTS-diazirine] in Figure 3.23A. The pattern of crosslinks is consistent with the model of OMP substrates tumbling dynamically within the cavity formed by the legs of the Skp trimer (Walton et al., 2009; Burmann et al., 2013). There are no obvious specific interaction regions in the cavity and legs of Skp, but OmpA appears to be excluded from the trimerisation ‘crown’ and charged ‘tips’ of Skp, consistent with crystal structures of apo-Skp (PDB: 1U2M, 1SG2) and literature which shows that the tips are only involved in initial recognition and OmpA is held near the centre of the cavity (Walton and Sousa, 2004; Korndörfer et al., 2004; Walton et al., 2009; Lyu et al., 2012; Burmann et al., 2013; Schiffrin et al., 2016; Zaccai et al., 2016). This pattern is recapitulated using MTS-TFMD, and also with the crosslinker located in β 1 on the OmpA(W7C) variant with both tag-transfer reagents (Figure 3.23B).

3.2.5.2 Positions on OmpA have no strong motif or structural preference for locations within the cavity of Skp

A more detailed analysis of the crosslinks can be done by assessing the number of detected peptide-spectrum matches (PSMs) from the enriched Method 3 datasets as a function of location in the primary sequence to determine whether if there are any ‘hotspots’. A caveat to this analysis is that the number of detected PSMs is not purely a function of how often that site was crosslinked, but is also a measure of how well those peptides were passed through each enrichment stage, how well the LC step separated the peptide from other peptides, and how efficiently the peptides were able to be ionized and detected in the mass spectrometer. Figure 3.24 shows that there is a common ‘hotspot’ between residues ~40–60 for both locations (β 1 [W7C] and loop 3 [T144C]), and similar hotspots of ~85–105 for loop 3, ~100–113 for β 1. These hotspots map to the ‘joint’ regions of the Skp legs (Figure 3.25A) where the common regions for both locations are shown in blue,

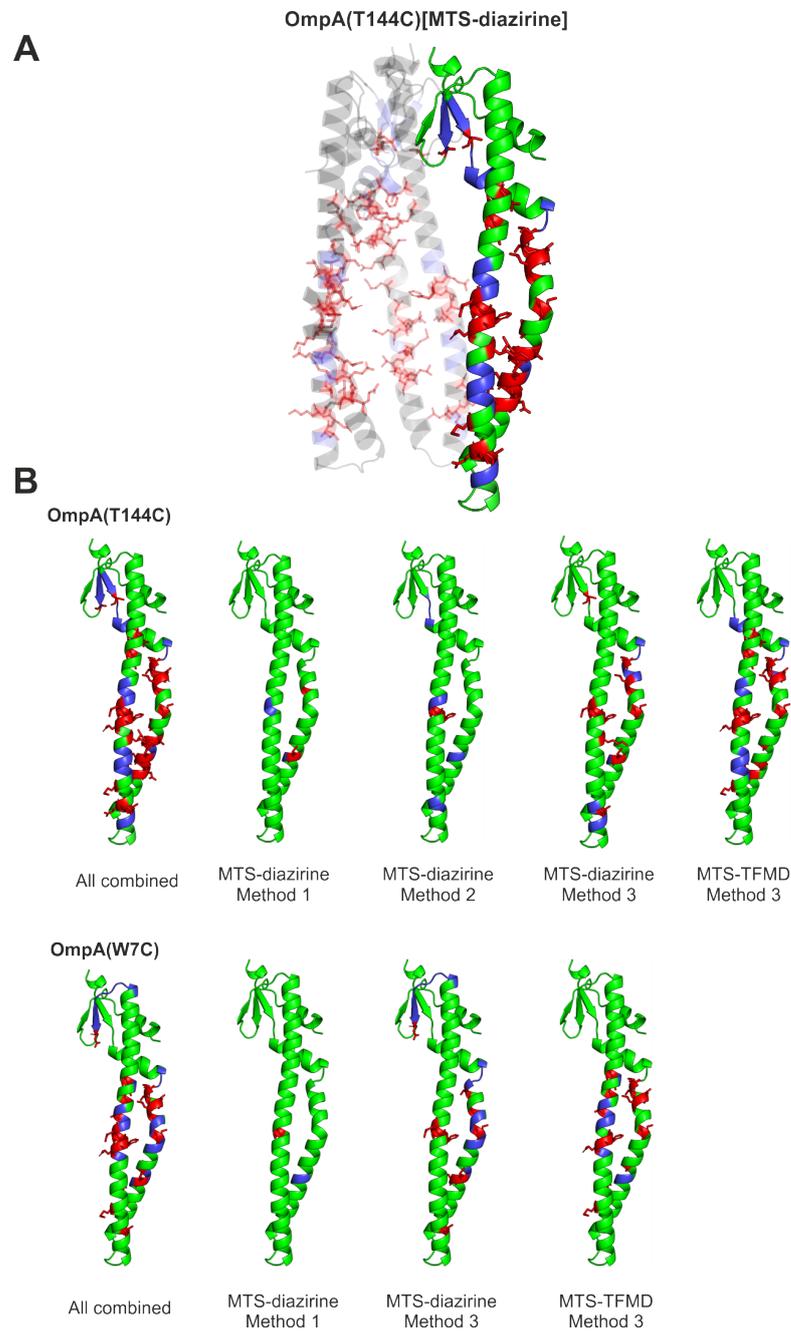


Fig. 3.23 The results of tag-transfer MS-XL mapped onto the structure of Skp. Residue-level identities of crosslinked positions are indicated as red sticks while subpeptide-level identities are shown in blue. (A) Pooled dataset of Method 1–3 results for OmpA(T144C)[MTS-diazirine] mapped onto the Skp trimer indicate that crosslinks are consistent with the OmpA substrate tumbling dynamically inside the cavity of Skp and making transient contacts throughout the length of the cavity. (B) Breakdown of the results for both positions from OmpA (β 1 - W7C; and loop 3 - T144C) and with both tag-transfer reagents using Methods 1 through 3 for enrichment.

the loop 3 only region in orange, and the $\beta 1$ only region in red. Comparing these locations to their hydrophobicity, electrostatic potential, or their conservation level (Figure 3.25B-D) showed no clear correlation with hydrophobicity or electrostatics, except that the more hydrophobic $\beta 1$ region tended to crosslink deeper into the cavity - where the Skp surface is also more hydrophobic. However, these hotspots correlate with the more conserved regions of the Skp legs (Figure 3.25D) as shown by the higher conservation score between homologues compared to the 'tips' and 'crown'.

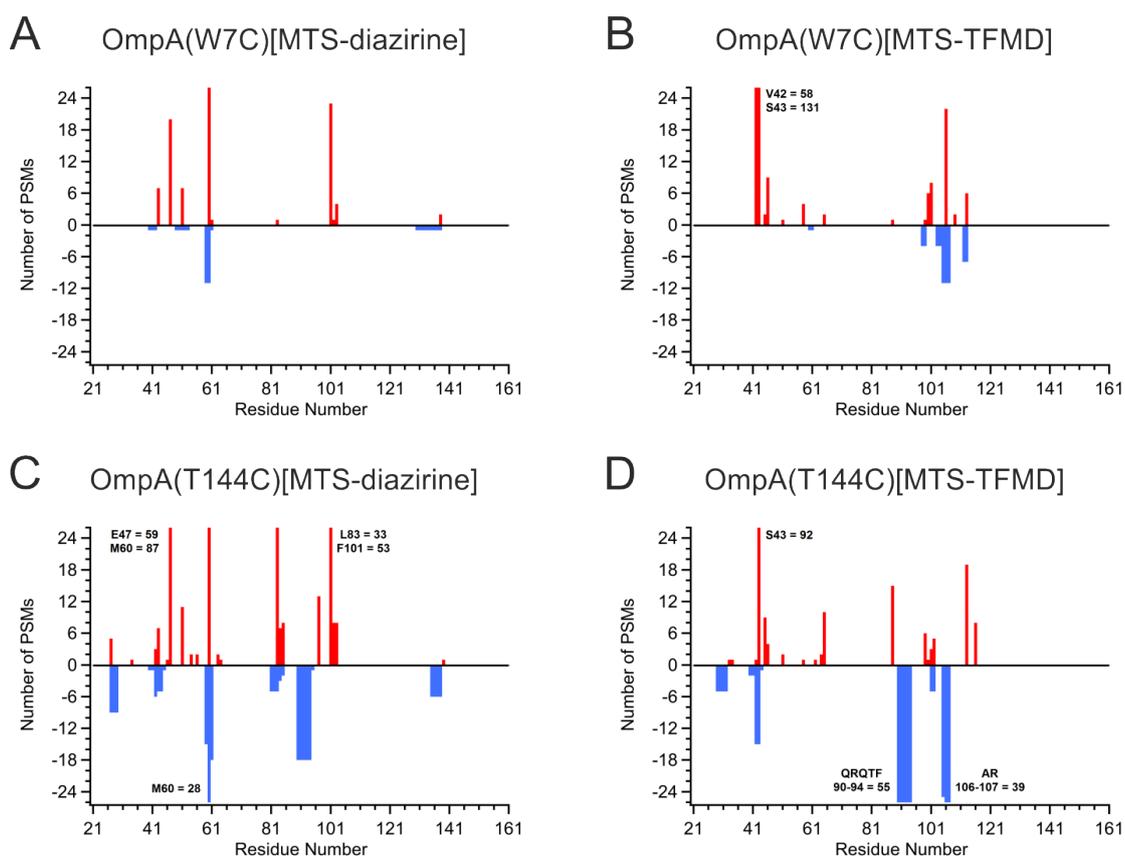


Fig. 3.24 Frequency of crosslinks at each location on Skp as measured by number of peptide spectrum matches. Residue numbering reflects the gene sequence (where residues 21–161 comprise the mature protein). Residue level-identifications are shown in the top-half of the plots as red bars. Sub-peptide-level identifications are shown in the bottom-half of the plots as blue bars. For sub-peptides, although in actuality only one modification exists in the region of uncertainty comprising the sub-peptide, a score of '1' PSM is given to each residue – i.e. the scoring has not been normalized. For sub-peptides with regions overlapping other sub-peptides, residue scores in these overlap regions are combined. To ease visualisation of abundances in the whole dataset, which spans 3 orders of magnitude, the scale has been capped at 26. For residues with more than 26 PSMs, the exact number has been annotated next to the bar on the chart along with the residue identity and number.

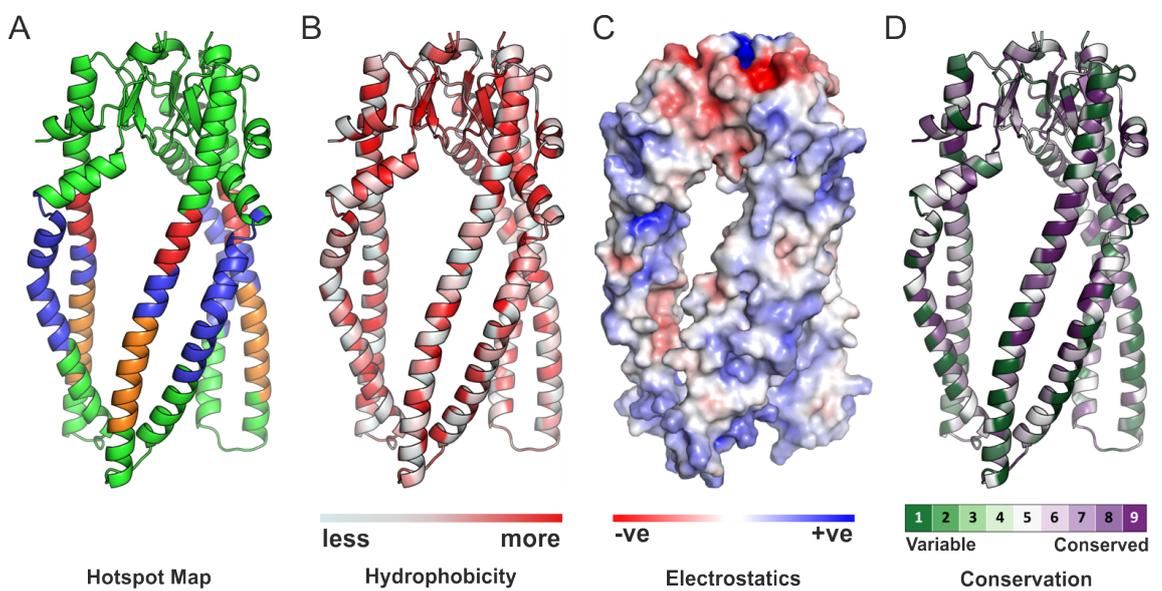


Fig. 3.25 Comparison of OmpA binding hotspots with structural and sequence properties of Skp. (A) Common binding hotspots for both OmpA(W7C) and OmpA(T144C) are shown in blue, OmpA(W7C) unique hotspots shown in red, and OmpA(T144C) unique hotspots shown in orange. (B) Structure-mapped hydrophobicity of Skp using the Normalized Eisenberg scale of hydrophobicity (Eisenberg et al., 1984), darker reds indicate more hydrophobic residues. (C) Electrostatic profile of Skp generated using the APBS plugin in PyMol2.0. The scale runs from -5.0 to $+5.0$ $k_B T e^{-1}$ units of electrostatic potential (where k_B is Boltzmann's constant, T is temperature, and e is the charge of an electron). (D) Is the per-residue conservation score of Skp generated using the ConSurf server (<https://consurf.tau.ac.il/>).

3.2.6 OmpA binds to conserved domains of SurA in a manner consistent with ‘beads on a string’ model of chaperoning

SurA is another major periplasmic chaperone of OMPs (Rizzitello et al., 2001; Sklar et al., 2007b) but less is known about its mechanism of chaperoning and the site of interaction for substrates than is known for Skp. In *E. coli* SurA comprises four domains: a N-terminal and C-terminal domain which make up the ‘core’ and two Parvulin-like PPIase domains, P1 and P2 (Rouvière and Gross, 1996; Bitto and McKay, 2002). Whilst the core domains are present in all Gram-negative bacteria with a *surA* homologue, the P1 and P2 domains are variably conserved (Humes et al., 2019). Although these domains are PPIases by homology, P1 shows no activity and P2 shows only low levels of activity (~1 % of Parvulin) suggesting that their function has diverged from proline isomerization (Rouvière and Gross, 1996). The crystal structure of SurA (PDB:1M5Y) (Bitto and McKay, 2002) shows an extended conformation with P1 packed against the core domain and P2 held away on a flexible linker. NMR experiments have shown that SurA holds the 22-stranded OMP, FhuA, in a dynamic unfolded state (Thoma et al., 2015). To gain more understanding of the role of these different domains in chaperoning a substrate, and to test the tag-transfer methodology on a dynamic (Thoma et al., 2015), μM affinity (Bitto and McKay, 2004; Humes et al., 2019), solvent exposed protein-protein interaction surface, crosslinking experiments were carried out between SurA and both OmpA(W7C) and OmpA(T144C) with MTS-diazirine and MTS-TFMD (to give four datasets, 2 OmpA variants x 2 tag-transfer reagents). For these experiments OmpA was incubated with SurA at a 10X molar excess (final conditions: 10 μM OmpA, 100 μM SurA, 1 M urea, 150 mM NaCl, 50 mM Tris-Cl pH 7.6) due to the lower affinity for the complex between SurA and OmpA compared to that for Skp and OmpA (see Materials & Methods Section 2.8.4 for experimental details, and Appendix B for a list of modified peptides and the UV irradiation and enrichment methods used).

3.2.6.1 Chaperoning activity of SurA primarily involves its core domain

Figure 3.26A shows the results of these crosslinking experiments pooled across both OmpA positions and both tag-transfer reagents. Crosslinks were observed throughout the N- and C-domains which form the core, with one sub-peptide level crosslink identified for each of the P1 and P2 domains. Figure 3.26B shows these crosslinks mapped onto the structure of SurA. A list of all modified peptides from SurA crosslinking experiments can be found in Appendix B. This shows that the OMP binding activity of SurA is performed almost entirely in the N- and C-domains of the protein. The crosslink identified in the P1 domain (IQE 236-238) is at the side of the P1 domain which faces the core and near a region

where activator mutations had been identified previously. These activator mutations were identified in a screen for suppressor mutants which rescue a partial loss-of-function *bamA* allele termed *bamA616* that showed OMP assembly defects and a less stable interaction between SurA and BamA *in vivo* (Ricci et al., 2013). One suppressor mutation is a deletion of L215 and A216 in SurA P1 (although this was found to also activate the Cpx stress response which effects expression of many envelope-related proteins Raivio et al. (2013)), and the other is a point mutant in the SurA P1 domain (S220A). This was found to destabilize the interface between SurA P1 and the core N- and C- domains without restoring BAM binding activity, implying a functional interaction occurs at that interface (Ricci et al., 2013; Soltes et al., 2016). Although the crosslink identified in the P2 domain appears to be very distal to the site of most abundant crosslinks in the ‘core’ domain this could be an artefact of the constraints imposed by the crystal structure of SurA. The flexible linkers connecting P2 to P1 and the C-domain could allow flexing at this hinge point and bring this domain into proximity with the chaperoning ‘hub’ located in the N-domain.

3.2.7 Diazirine crosslinkers are small, unbiased reactive probes

3.2.7.1 The MTS-diazirine tag-transfer reagent is one of the smallest crosslinking molecules

One of the goals of the new tag-transfer design was to create a reagent and workflow that minimizes perturbation of a protein’s native structure while maximising reactivity, residue accessibility, and analytical detectability. Perturbation of the structure and function of a protein by a crosslinker comes in part from the size of the reagent in comparison to the protein being studied. To put the size of the tag-transfer reagents in context, a range of both chemical and photoactivatable crosslinkers, as well as bulky amino acids residues, were built onto a common peptide backbone fragment using ChemDraw® (PerkinElmer, Waltham, MA) and 3D models built and energy minimised using Chem3D® (PerkinElmer, Waltham, MA). From these models, straight-line-distances were measured from the C α to the farthest non-hydrogen atom and the active atom for photo- or chemical-crosslinking. The results are summarized in Figure 3.27. The installation of MTS-diazirine on a Cys residue results in an adduct that is similar in length to a large amino acid such as arginine (8.4 Å for Cys \times MTS-diazirine, 7.3 Å for arginine), and of a similar length but less bulky than unnatural amino acid photocrosslinkers installed by amber suppression methods (8.4 Å for MTS-diazirine compared to 7.9 Å for *pAzpa*, 8.9 Å for *TfmdPhe*, and 10.0 Å for *pBpa*). Although MTS-TFMD is larger and bulkier, its conjugated form is within a couple

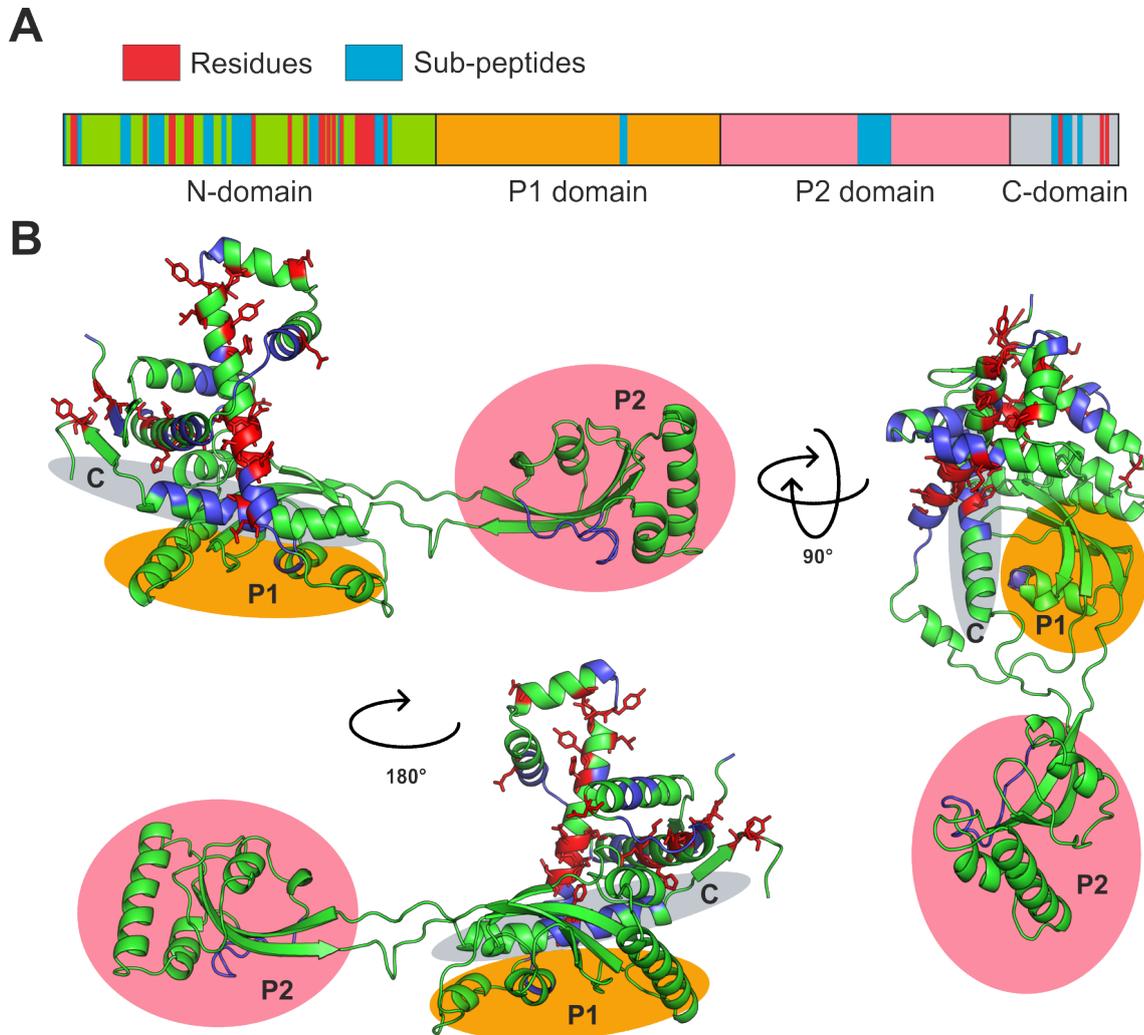


Fig. 3.26 All crosslinks observed from OmpA(W7C) and OmpA(T144C) to SurA with both tag-transfer reagents pooled and mapped onto SurA. (A) Sequence of SurA represented in 1D as a correctly scaled box with domain boundaries marked and coloured. Crosslinks identified to the residue level are indicated in red, while those identified to the 'sub-peptide' level are marked in blue. (B) Crosslinks mapped onto the crystal structure of SurA (PDB: 1M5Y) (Bitto and McKay, 2002) with the loops built in by Modeller (<https://salilab.org/modeller/>) (courtesy of Dr. Bob Schiffrin, University of Leeds). Crosslinks are coloured as in (A) with residue-level identifications are shown in red with the side chains displayed as sticks over the cartoon backbone, 'sub-peptide' level identifications are shown in blue on the cartoon representation. Three views of SurA are shown. P1, P2, and C-terminal domains of SurA are approximately highlighted with background ellipses with the same colour scheme as in (A). The N-terminal domain is not highlighted.

of angstroms in length of reaction intermediates of popular “zero-length” crosslinkers (12.7 Å for MTS-TFMD, 11.5 Å for EDC, 10.3 Å for DCC, and 9.3 Å for CDI).

3.2.7.2 Diazirine photocrosslinkers are globally reactive and unbiased

The rich dataset afforded by the tag-transfer crosslinking and enrichment methodology allows examination of the reactivity of the diazirine group in these crosslinkers. Other chemical and photoactivatable crosslinking groups have specific reactivities or strong biases in their reactivity toward certain amino acids (Wittelsberger et al., 2006; Dormán et al., 2016), but the carbene unmasked upon irradiating diazirines is extremely reactive and should show little to no residue selectivity (Geurink et al., 2012; Preston and Wilson, 2013). This was assessed empirically by first defining a ‘substrate accessible’ region of Skp that excludes the crown region involved in trimerization but not substrate binding (shown in red in Figure 3.28A). Next, the fraction of unique positions of each residue type that were crosslinked was quantified – pooling data across all experiments. If certain residue types are significantly avoided or selected by the diazirine moiety, then only a small fraction (or none) of that residue type will be modified as any preferential residues with faster reaction kinetics will ‘mop up’ the carbene first. Conversely, if there is a bias toward certain residues this should be apparent as a highly dominant residue. Figure 3.28B shows that although there are some differences between the two tag-transfer reagents (the data suggest MTS-diazirine has a greater preference toward carboxylic acids – Asp and Glu), the reactivity is broadly similar. Considering the global reactivity when both tag-transfer reagents are considered (Figure 3.28C) it is clear that the tag-transfer reagents are capable of capturing interactions across 13 of 15 residue types present (Cys, His, Pro, Tyr, and Trp residues are not present in this region of the protein; no crosslinks were observed on Ile and Gly residues). Leucine has the highest fraction of modified residues and this observation is not easily explained chemically (as the methyl groups of Leu would not be expected to show increased reactivity), but instead this could be a function of the higher conservation of Leu residues within the cavity of Skp suggesting an important role in holding substrates. There is only one Ile in the accessible region of Skp (I120) and no crosslinks to this residue were detected in these experiments. However, I120 resides high in the hydrophobic cavity in a pocket which is partially occluded by Met29, Phe33, Val117 and Gln121 in the crystal structure of Skp (PDB: 1U2M) (Korndörfer et al., 2004; Walton and Sousa, 2004) and so may be inaccessible in solution. No crosslinks are observed to glycine, which may be a combination of the small size of glycine offering much less accessible surface area for crosslinking and the fact that only two of the five glycines (Gly30 and Gly114) face toward the cavity, the others are located on the outside edge of legs or the tips. Only ~10 % of

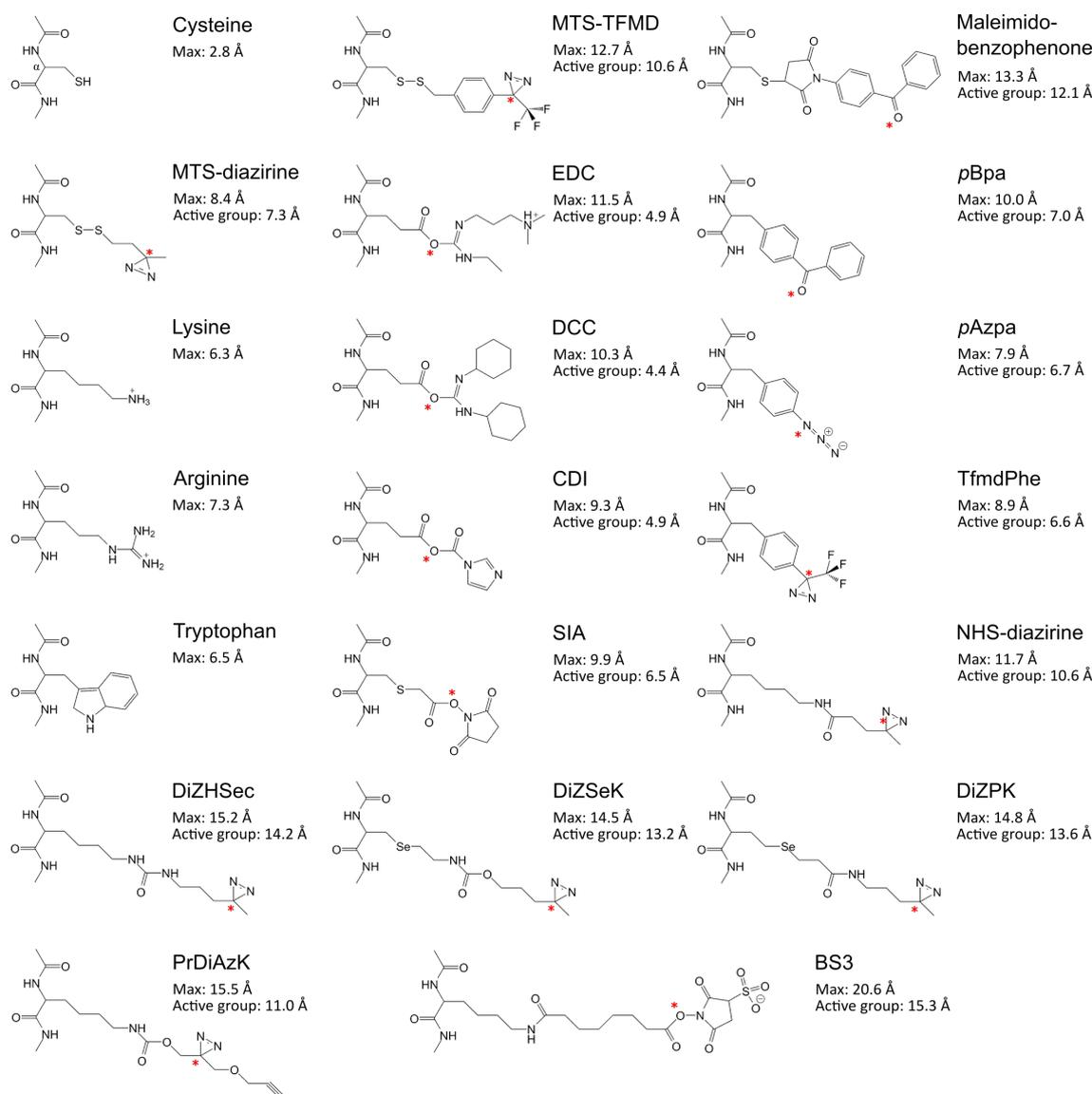


Fig. 3.27 Comparison of the sidechain length of MTS-diazirine and MTS-TFMD conjugated to Cys with other bulky amino acid sidechains and crosslinkers including ‘zero-length’ crosslinkers and unnatural amino acids used for crosslinking experiments. Crosslinkers are shown in their conjugated, or intermediate forms (i.e. after reaction of one half of a bifunctional crosslinker) to give an indication of the ‘maximum bulk’ conferred by the reagent. This is particularly pertinent for so-called ‘zero-length’ crosslinkers which actually proceed through a large, bulky, intermediate stage. Distances were calculated as straight-line-distances from energy minimised 3D models of the above structures using the MMFF94 force field in Chem3D (CambridgeSoft). “Max” distances were calculated from the centre of the C α carbon (marked on cysteine for reference) to the most distant non-proton atom of the residue side-chain. “Active group” distances were measured from the C α carbon to the reactive atom involved in the final crosslinking reaction to a target. These are marked with a red asterisk on the image. EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride; pBpa, *p*-benzoyl-L-phenylalanine; DCC, *N,N'*-dicyclohexylcarbodiimide; pAzpa, *p*-azido-L-phenylalanine; CDI, *N,N'*-carbodiimide; SIA, succinimidyl iodoacetate; NHS-diazirine, *N*-hydroxysuccinimide-diazirine; BS3, bis(sulfosuccinimidyl)suberate.

arginine residues were crosslinked, this may be partially explained by the fact that 6 of the 9 solvent accessible Arg in Skp face away from the binding site and so may be inaccessible to OmpA held in the cavity. One of the 3 remaining Arg was crosslinked (Arg113) but it is not clear why crosslinking was not observed to the other two.

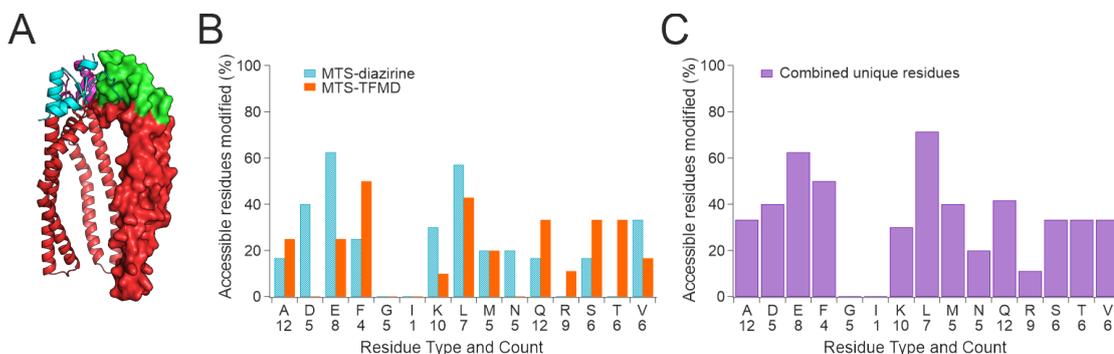


Fig. 3.28 Reactivity profile of MTS-diazirine and MTS-TFMD-conjugated OmpA to accessible residues of Skp. (A) The ‘crown’ region of Skp (green spacefill) is inaccessible to substrate but any part of the cavity or ‘tentacles’ may come in contact with OmpA (Walton and Sousa, 2004; Burmann et al., 2013; Schiffrin et al., 2016), and therefore might be crosslinked. We defined these as ‘accessible residues’ (comprising the cavity or tentacle region of Skp, corresponding to residues V27-Q121 and D138-A143, as indicated in red, with one ‘tentacle’ shown in spacefill). (B) MTS-diazirine and MTS-TFMD react with different residues in Skp. The percent of each ‘accessible residue’ type modified was calculated from the number of unique residue-level crosslink assignments. The absolute number of each residue type present in the Skp ‘accessible residue’ region are shown below each bar. Cysteine, histidine, proline, tyrosine and tryptophan residues are not present in this region of Skp and have been omitted. (C) The same data as (B), but not split by crosslinker type. Note that no Ile and Gly crosslinks were observed, but the former is poorly represented in Skp.

3.2.8 Towards capturing the reaction coordinate of OmpA on BAM

3.2.8.1 Tag-transfer crosslinking after initiating folding on BAM

With the tag-transfer methodology established and validated, these reagents were next used to map the folding of OmpA through the BAM complex. OmpA(W7C)[MTS-diazirine] was rapidly diluted out of 8 M urea into a solution containing SurA and BAM in *E. coli* polar lipid proteoliposomes. After a 5 minute delay to allow OmpA to engage with BAM and begin folding, samples were loaded into acrylic chips and irradiated for 60 s by the UV LED. For experimental details see Figure 3.29 and Materials & Methods Section 2.8.4. A captured intermediate of OmpA folding should be retained in the BAM proteoliposomes, while SurA-bound crosslinked OmpA will remain in the supernatant. The sample was thus centrifuged at $14,000 \times g$ to pellet the proteoliposomes, and this pellet resuspended in buffer containing 8 M urea and 0.001 % (w/v) DDM (below its CMC) to unfold the BAM

Table 3.5 Identified crosslinks from OmpA(W7C)[MTS-diazirine] to BamA and SurA. Residue numbering is according to the gene sequence. Sequence represents the identity of the residue or 'sub-peptide' region. Enrichment method identifies whether which tag-transfer crosslinking methodology was used to identify the crosslinks (see Section 3.2.4.1). Location shows domain or region of the protein the crosslink was identified in. Conservation score was calculated using the ConSurf server (<https://consurf.tau.ac.il/>) which compares residue conservation across homologues generated in a multiple sequence alignment, where 1 represents little to no conservation of the residue and 9 represents a very highly conserved residue.

Protein	Residue	Sequence	Enrichment Method	Location	Conservation score (out of 9)
BamA	44–52	LSMPVRTGD	3	Potra 1	6.4
	291	E	2	Potra 4	2
	304–305	ME	3	Potra 4	6.5
	323–325	QSM	3	Potra 4	4.7
	353	R	2	Potra 5	5
	354	F	2	Potra 5	7
	795	D	2	Loop 7	1
	SurA	44–46	GLM	3	N-terminal
107–109		NNM	3	N-terminal	5.7
120		Y	3	N-terminal	1
135–137		EMI	3	N-terminal	6
328–341		FSQDPGSANQGGDL	2	P2	6.6
359–366		RLNKGQMS	3	P2	4.4

complex and help solubilise the liposome. Samples were split and tag-transfer Method 2 and 3 were used to enrich the crosslinked products. Figure 3.29A shows that using Method 2 a number of bands were observed on an SDS-PAGE gel. These were excised, digested, and analysed by tandem MS according to the scheme shown on the right in Figure 3.29A. At the same time, on-bead digest samples (Method 3) were also analysed. Good spectra of putatively crosslinked residues were only acquired for BamA in band 3, but putative sub-peptide level identifications were also identified in the P2 domain of SurA in band 2 and 3. The on-bead digest of Method 3 did not provide much improvement, but the data identified crosslinks to the sub-peptide level for BamA and sub-peptide and residue-level identifications for SurA (Table 3.5).

3.2.8.2 Putative interactions of an OMP folding through BamA

Figure 3.30 shows the location of crosslinks from OmpA(W7C) found by Method 2 and 3 to BamA and SurA that were identified from crosslinking after initiation of OmpA folding through BamA via SurA. The location of crosslinks map almost entirely to the periplasmic domain of BamA in P5, P4 and P1 with one crosslink also observed at the top of loop 7 (D795). Two of these crosslinks are in locations which are hard to reconcile with

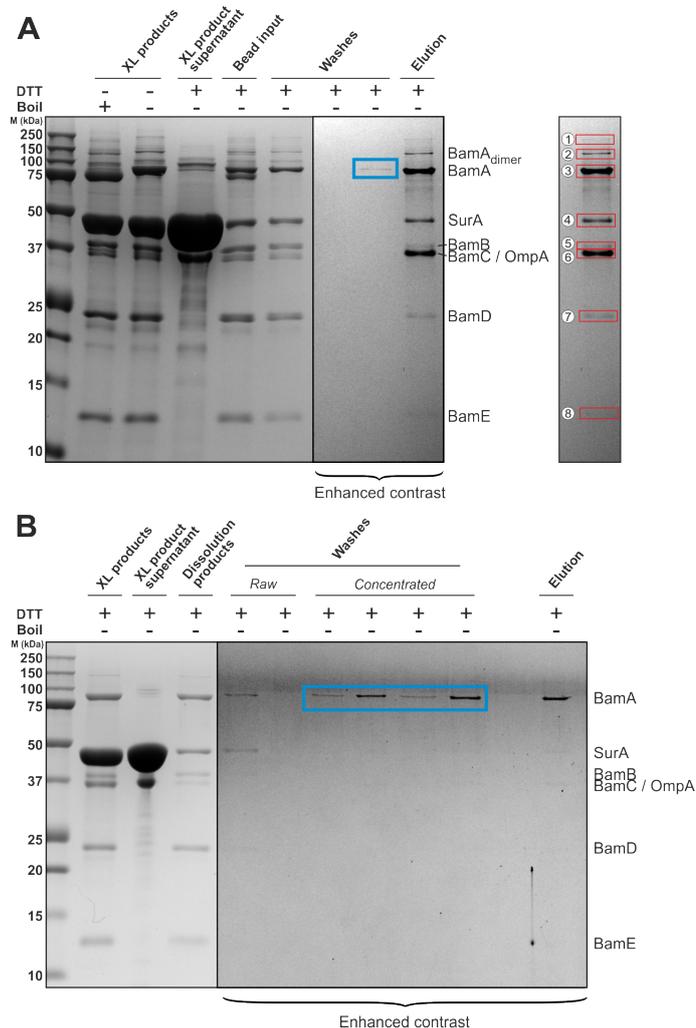


Fig. 3.29 Using tag-transfer crosslinking to capture the transit of OmpA as it folds via BAM. (A) 10 μ M BAM proteoliposomes of *E. coli* polar lipids, 10 μ M OmpA(W7C)[MTS-diazirine], and 50 μ M SurA were mixed together and OmpA allowed to fold for 5 minutes before being irradiated in the chip-based UV LED for 60 s. Final conditions: 1 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0. Proteoliposomes were pelleted, supernatant removed, resuspended in wash buffer (0.001 % (w/v) DDM, 10 mM DTT, 8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0), and allowed to incubate for 1 hour before continuing with the tag-transfer Method 2. The products were then enriched according to Method 2. Beads were washed by 6 rounds of 1 ml washes with wash buffer (without DTT added). (B) 5 μ M BAM proteoliposomes of *E. coli* polar lipids, 5 μ M OmpA(T144C)[MTS-diazirine], and 50 μ M SurA were mixed together and OmpA allowed to fold for 90 s before being irradiated with tube-based UV LED for 30 s. Final conditions: 0.5 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0. Proteoliposomes prepared as in A but resuspended in 'dissolution buffer' (1 % (w/v) DDM, 10 mM DTT, 8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0) instead of wash buffer. Beads were more thoroughly washed by 12 rounds of 1 ml washes with 6X wash buffer A (0.001 % (w/v) DDM, 8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0) followed by 6X wash buffer B (wash buffer A without DDM). Wash fractions show aliquots taken as these washes progressed. Raw indicates an aliquot from the 1 ml wash supernatant was loaded directly, 'concentrated' lanes were spun in a 3.5K MWCO centrifugal concentrator to ~10-30X concentrated (variable) to assess for low-abundance carry-over between washes. Contrast has been increased globally in the indicated regions on A and B to more clearly visualize low intensity bands.

what is known about BAM, particularly D795 in loop 7 and E291 in P4. D795 faces the interior of the proteoliposomes and is far from the expected final location of W7C which would sit at the bottom of β 1 near the interfacial region of the outside leaflet of the proteoliposomes in its final folded state. Although P4 forms part of the periplasmic “cradle” formed by the POTRA domains, E291 faces away from the cradle enclosed by the POTRA domains and sits on the outside of P4. Both of these residues are also poorly conserved between homologues when BamA is analysed through the ConSurf server (see Table 3.5). Nonetheless, E291 could have been crosslinked by folded OmpA which remained in close proximity to the BAM complex after folding. Other crosslinks were identified in P1, P4 and P5 which are in locations consistent with OmpA being passed from P1 to the ‘entrance’ of the BamA barrel at the bottom of P5/P4 (possibly via BamD). These residues also showed higher levels of conservation between homologues (see Table 3.5) suggesting they may be important for function. Crosslinks were also seen from OmpA to SurA with most mapping in the N-terminal domain of the core N- and C-terminal regions, consistent with the results shown in Section 3.2.6 where this region is primarily responsible for binding OmpA. However, in contrast to the results described in Section 3.2.6, a larger proportion of crosslinks are seen to the P2 domain of SurA (2 out of 6, 33 %). With much fewer crosslinks overall than the SurA-OmpA dataset it is difficult to tell whether this is a significant effect indicating greater recruitment of SurA P2 during OMP handover to BAM or merely a chance enrichment due to some other factor (greater solubility of the peptides, better ionisation efficiency etc.). The presence of SurA peptides in the final elution was surprising as it was expected that SurA would remain in the supernatant after pelleting and be removed before the resuspended pellet was incubated with the Thiopropyl Sepahrose® 6B beads. The persistence of SurA suggests that either it was stably associated with BAM during the pelleting step (and had been crosslinked to OmpA during chaperoning), or that it was crosslinked to folded / partially inserted OmpA due to the fact that W7C, where the tag-transfer reagent is conjugated, is at the bottom of β 1 and so crosslinks to a bulky target (i.e. SurA) may still allow the formation of a stable β -barrel.

3.2.8.3 Improving crosslink enrichment for crosslinking to BAM

Although these were promising initial results, there were a few indications that the protocol used for the enrichment process was not ideal: far fewer crosslinks identified than for the OmpA-chaperone complexes, in the final wash step a faint band corresponding to BamA was seen indicating not all ‘free’ / unbound protein had been removed (Figure 3.29A blue box), SurA was present in the elution despite the fact it should have been removed as supernatant after pelleting, an apparent stable dimer of BamA was present in the

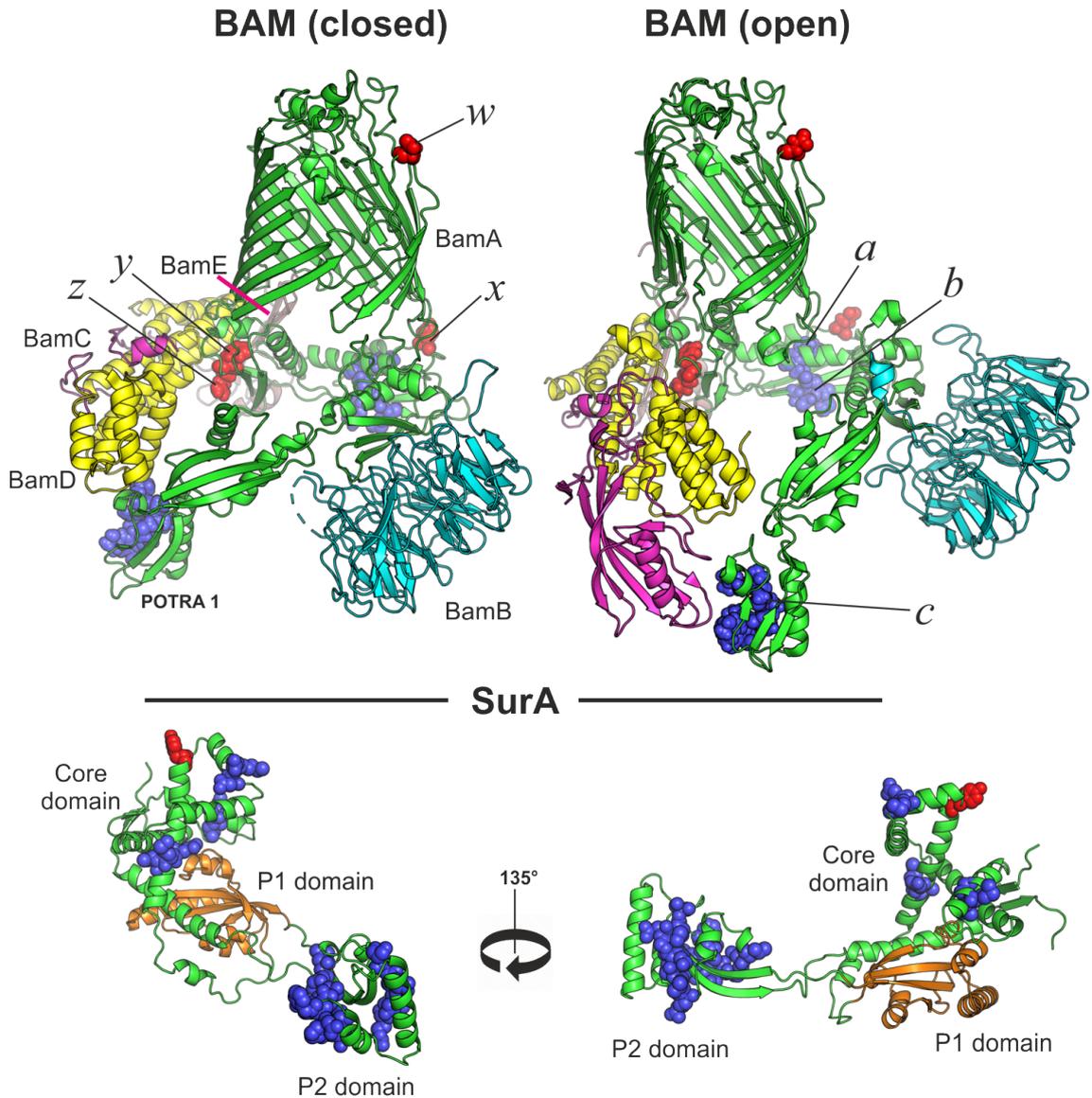


Fig. 3.30 Crosslinks observed from OmpA to BAM and SurA under folding conditions. Crosslinks which are confident to the residue level are shown as red spheres, while those confident to a sub-peptide level are shown as blue spheres. (A) The same view of the BAM complex is shown for the closed (PDB: 5D0O) (Gu et al., 2016) and open (PDB: 5LJO) (Iadanza et al., 2016) structures of BAM after alignment on the back of the BamA barrel. Residue-level identifications are lettered as follows: w = D795, x = E291, y = F354, z = R353. Sub-peptide level identifications are lettered as: a = 304-305, b = 323-325, c = 44-52. BAM subunits are coloured individually, BamA = green, BamB = cyan, BamC = magenta, BamD = yellow, BamE = salmon. (B) crosslinks plotted onto the structure of SurA (PDB: 1M5Y) (Bitto and McKay, 2002). Only the P1 domain of SurA is coloured separately to distinguish it from the core N- and C-domains.

elution, and there were relatively strong elution bands seen for subunits where little to no crosslinking would be expected (BamB/BamE). A logical explanation for these results is that there is a lot of ‘carry-over’ from earlier steps even when no binding to the Sepharose® beads has taken place. The most likely cause of this is that BAM is not being unfolded and released from proteoliposomes upon the addition of low concentration of detergent and high concentrations of urea. Folded OMPs are very thermodynamically stable and therefore even 8 M urea may be insufficient to unfold BamA within the BAM complex. BAM subunits (BamBCDE) may unfold but their N-terminal acylation could keep them stably associated with the proteoliposomes. If the liposomes themselves remain intact in these conditions, they are large enough that most if not all of them would be pelleted along with the Sepharose® beads. To improve the protocol for use with proteoliposomes, two modifications were implemented. First, the pellet produced after centrifuging the crosslinked material was resuspended in a ‘dissolution buffer’ containing 1000X more detergent in combination with the high concentration of urea (1 % (w/v) DDM, 8 M urea) to ensure complete liposome disruption and hopefully unfold BamA. A much longer and thorough wash was also employed (12 × 1 ml washes) with the higher detergent concentration also used (8 M urea, 1 % (w/v) DDM). The results of Method 2 enrichment using this protocol are shown in Figure 3.29B where the supernatants from the final wash steps were concentrated 10–30X to better visualise any low abundance carry-over. This method seems to have successfully removed carry-over from BAM subunits and SurA, but as can be seen in the final wash lane (Figure 3.29B blue box), BamA continues to be carried across experiments suggesting it is poorly removed by washing and suggesting that the elution band merely corresponds to this carry-over. It was clear that using the method of pelleting and resuspension of the Thiopropyl Sepharose® beads for washing was not appropriate when the sample contains material that can be co-sedimented. To overcome this limitation, empty spin columns (Bio-Rad Micro Bio-Spin®) were employed for the wash steps under the assumption that material bound to the beads (which are on the order of 45–165 µm) will be retained at the top of the column while sub-micron material (such as the 100 nm diameter *E. coli* polar lipid large unilamellar vesicles) will pass through the 30 µm polyethylene bed and can be discarded. Method 3 was used to enrich the products of crosslinking experiments from four samples comprising of OmpA(W7C) and OmpA(T144C) conjugated with each tag-transfer reagent, according to the methods described for Figure 3.29B. After binding to the beads, all further wash and elution steps were performed using spin columns. Unfortunately, mass spectrometry analysis showed the presence of large amounts of an unidentified polymer masking the signal from the peptides in all samples (data not shown). Due to the presence of this contaminant in all samples it is likely the source of the contamination is shedding of polyethylene from the bed of the spin columns which is likely enriched by the C18 media at the final cleanup

step. Thoroughly washing the spin columns before use (with water or buffer) removed any trace of polymer from the flow through. Unfortunately, the results of BAM crosslinking using this method were unable to be acquired before preparation of this thesis but these methodological alterations should allow future studies to study BAM-OmpA interactions using the tag-transfer reagents and workflow described in this chapter.

3.3 Discussion

In this chapter the development and application of a new method for mapping protein-protein interactions using XL-MS has been presented. This tag-transfer crosslinking method allows site-directed “one-to-many” interrogation of protein-protein interactions complementing existing chemical and photo-crosslinking strategies which can be used to study a broader spread of residues across the length of a bait (and target) in a single experiment, but at much lower resolution. These existing strategies could be used, for example, to prospect for sites of interest before a tag-transfer approach is used to discover all interacting residues in that region – something which is particularly powerful for dynamic and transient interactions as shown here for OmpA on Skp, SurA, and BAM. Alternatively, a series of Cys mutants spread along the primary sequence of a protein could be produced and used to obtain detailed information on whether even low-abundance interactions occur between a bait and target at each position (Figure 3.31).

3.3.1 A new tool for integrative structural biology

The potential for tag-transfer crosslinking extends further than used here. Structural proteomics, using XL-MS to discover surface-exposed regions of proteins and derive distance restraints for modelling, is a rapidly growing field bolstered by improvements in modelling software like Xplor-NIH (Schwieters et al., 2018), Integrative Modeling Platform - IMP (Webb et al., 2018), and Multiscale Modeling of Macromolecules - MMM (Jeschke, 2018) among others. Integrative structural biology increasingly relies on restraints from crosslinking studies to help fit and dock protein complexes together (Ferber et al., 2016; Politis and Schmidt, 2018; Braitbard et al., 2019; Chavez et al., 2019). Recent high profile examples include the RNA polymerase II pre-initiation complex (Murakami et al., 2013), transcription factor II H complex (Luo et al., 2015), mitochondrial respiratory complexes I and III (Schweppe et al., 2017), and perhaps most impressively the nuclear pore complex (Kim et al., 2018). The crosslinkers used in all these examples are Lys-

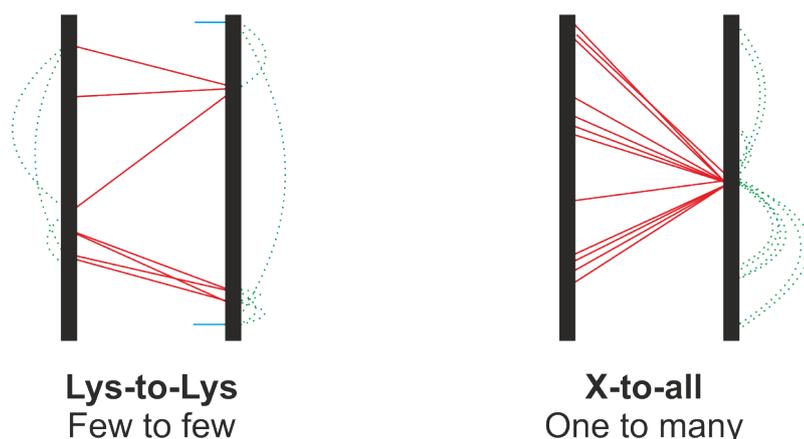


Fig. 3.31 Example of outcome of residue-specific crosslinking experiment versus tag-transfer. Residue-specific crosslinkers such as BS3 which crosslinks lysines to lysines can sample to and from multiple positions on both interacting proteins. However, the limited number of any one residue type and the potential for regions lacking any reactive residue can have two drawbacks. First, in this example assuming each bold black line represents the primary sequence of a different 100-residue long protein, on average in *E. coli* only 5 Lys will be present in each - thus "few-to-few". Second, some regions of the chain will be 'invisible' to the crosslink due to the lack of relevant amino acids. Tag-transfer crosslinkers only sample from one point for one protein, but this point can be selected by the experimenter, and it can crosslink to potentially any other position - thus "one-to-many". Red lines indicate crosslinks, blue ticks indicate 'dead-end' crosslinks, green dashed lines indicate intra-protein crosslinks.

to-Lys and very long (with linker arms over a nm in length). These long linker arms preclude the ability to generate high resolution maps of protein complexes in the absence of complementary data (such as EM or crystal structures of some subunits, or a low resolution structure of the whole complex in to which to fit a model). However, the inclusion of some short-distance constraints alongside longer constraints can help reduce the computation time and improve the overall model by reducing the conformational space that docking and modelling algorithms need to work within (Brodie et al., 2017).

3.3.2 Prospects for further refinements of the tag-transfer method and UV LED lamp

There are a number of ways in which the tag-transfer method could be further improved in the future. One issue with all diazirines is that after absorbing a photon of light, two chemical pathways are available for the molecule: (1) direct evolution of a molecule of N_2 and generation of a carbene, and (2) production of a linear diazo intermediate en-route to the production of (1) (see Figure 3.1B). The generation of this linear diazo form can be problematic as unlike the highly unbiased carbene, linear diazo groups have a preference for nucleophilic side chains (Preston and Wilson, 2013) (the relative order of reactivity

of unprotonated functional groups in amino acid residues is $R-S^- > R-NH_2 > R-COO^- = R-O^-$ (Bischoff and Schlüter, 2012)) – although this biased reactivity can also be exploited as a secondary reporter for carboxylic acids groups in XL-MS experiments (Iacobucci et al., 2018). This undesired isomerization of diazirines was one of the driving forces for the creation of the TFMD version of the diazirine group – the electron withdrawing trifluoromethyl and phenyl groups help to stabilise the diazirine ring system and increase the ratio of the singlet carbene over linear diazo isomer (Brunner et al., 1980). If the linear diazo form does not react, it can undergo spontaneous decomposition to the desired carbene. However, after irradiation at 360 nm there is an initial increase in the absorption band between 260–315 nm (Nassal, 1983) and irradiation near this band at ~302 nm has been shown to efficiently catalyse conversion from the linear diazo to the carbene form (Hosoya et al., 2004; Hashimoto and Hatanaka, 2006; Chee et al., 2010). The efficiency of diazirine crosslinkers could therefore be increased, and their residual bias reduced, by co-illuminating the samples with a 302 nm UV LED at medium intensity to drive any intermediate species into the singlet carbene form. For the chip design this could take the form of a second LED sitting behind the sample, or for the tube design, having the second LED underneath the tube.

The MTS-diazirine tag is one of the smallest crosslinking tags reported in the literature and when installed by Cys mutation and tag conjugation from a wild-type Lys, Trp, or Arg residue, it is a very conservative change in terms of pure steric size (see Figure 3.27). However, if mutating from smaller residues the max length of Cys-conjugated MTS-diazirine (8.4 Å) is still longer than any natural amino acid side chain. Recently, a synthetic route to an even smaller diazirine has been reported (Chang et al., 2018) that shortens the aliphatic linker by an extra carbon, reducing it by a further ~1.5 Å. This could be incorporated into the MTS-diazirine tag-transfer reagent to bring it within the lengths of natural side chains, a property which might be particularly important when the desired target site is in a small binding pocket or the protein-protein interaction is thought to be very sterically restrained. As mentioned above, this could also be used to derive even smaller restraints in docking and modelling of protein complexes (or *de novo* structure prediction from XL-MS).

In this chapter, time-resolved crosslinking (or at least, time-controlled crosslinking) was achieved by improving photon flux and efficiency through the creation of a new UV LED lamp. This allowed crosslinking to be initiated at set timepoints with a very small ~9 s uncertainty (the time it takes to reach maximum yield - Figure 3.14). The issue of temporal resolution in diazirine crosslinking has been overcome by other groups through the use of low (~10–100 $\mu\text{J}/\text{pulse} = \sim 100\text{--}1000\text{ W}$) to high energy (~600 $\text{mJ}/\text{pulse} = \sim 6\text{ MW}$)

pulsed lasers to achieve full conversion in 10–120 s or less than 10 ns, respectively (Jumper and Schriemer, 2011; Manzi et al., 2016). Compared to the UV LED lamp described in this chapter, the cost (£10k+) and safety considerations of buying and using pulsed lasers precludes their use outside of specialist groups. However, another approach is to exploit the fact that, unlike chemically triggered crosslinkers which need to diffuse to and react with a specific target residue, photocrosslinkers can be ‘dosed’ and activated by light even in the solid phase. This means that reactions can be quenched by snap freezing and then irradiated in the frozen state (which is transparent to ~350 nm UV light) to capture intermediates and gather time-resolved information. Although this would be a more technically challenging setup (handling liquid nitrogen, plunge freezing samples, then holding them at 77 K while they are irradiated) it should not have a significant financial cost and may come with a number of other advantages. Aliphatic-diazirines (such as MTS-diazirine) have been reported to show a preference towards glutamate and aspartate residues, likely due to a kinetic effect (the nucleophilic carboxylic acid moiety reacts faster and can ‘mop up’ carbene located near proximal residues). It was observed for surface labelling (footprinting) of calmodulin that this effect could be ameliorated by crosslinking in the frozen state where the exact temperature altered how well the XL-MS data recapitulated the theoretical surface-exposure of residues based on the crystal structure of calmodulin in the order 77 K > 195 K > 273 K (-196 °C > -78 °C > 0 °C) (Jumper et al., 2012). This was also shown for lysozyme and the deubiquitinating enzyme, USP5, confirming that this is a general property of diazirines (Manzi et al., 2016). Furthermore, the photolabeling yield for aliphatic diazirines increased 6X when labeling in the solid phase at 77 K versus in the aqueous phase at 293 K (20 °C) (Ziemianowicz et al., 2017).

3.3.3 Probing dynamic interactions of OMPs reveals mechanisms of chaperoning

The interaction between Skp and its substrate OMPs has been examined thoroughly for 8-stranded β -barrels, particularly OmpA and OmpX (Bulieris et al., 2003; Qu et al., 2007, 2009; Walton et al., 2009; Lyu et al., 2012; Burmann et al., 2013; Schiffrin et al., 2016; Zaccai et al., 2016; Holdbrook et al., 2017; Li et al., 2018). The data presented in this chapter agree with the model that has emerged from this extensive characterisation where OMPs (at least ones smaller than ~25 kDa) form a 1:1 complex with Skp where they are held in an unfolded, but dynamically tumbling state, within a cavity formed by the legs of the Skp trimer. In this model, the ‘crown’ region of Skp is not involved in chaperoning but plays a structural role in forming the trimer, and the charged tips of the Skp legs act only during the initial recognition stage to funnel the chain into the cavity (Korndörfer

et al., 2004; Walton and Sousa, 2004; Lyu et al., 2012). As shown in Figure 3.23B, in a crosslinking experiment between a dynamic, undefined protein-protein interface where either the yield or detection efficiency of crosslinks is poor (c.f. Method 1), incorrect conclusions may be drawn about the specificity of the interaction. The ability to enrich and detect low-abundance crosslinks (Method 3) allows us to recapitulate the literature conclusions drawn from high-resolution methods such as NMR (Burmam et al., 2013), validating the power of the tag-transfer technique to study these kinds of interactions.

Interactions between SurA and OMPs are much more poorly understood than the interaction with Skp. Capturing unfolded OMPs on SurA is also more challenging due to the lower affinity for clients (μM for OmpA) (Humes et al., 2019) versus Skp (nM for OmpA) (Qu et al., 2007). The data described in this chapter provide some of the highest resolution evidence to date that SurA captures OmpA via its most conserved N- and C-terminal domains, and that the P1 and P2 domain play only an indirect auxiliary role in OmpA binding. These N- and C-domains are the most evolutionarily ancient (Humes et al., 2019), suggesting that additional domains (P1 and P2) were added by evolution to either modulate the function of the core or to provide scaffolds for additional functionality or binding partners. Crosslinks between OmpA and SurA are found throughout the N- and C-domains, but particular hotspots are seen around an α -helical feature in the C-domain that extends from the bottom to the top of SurA (between residues E94–E145) suggesting that OmpA may ‘wrap-around’ the N-domain. A crosslink is also observed in P1 near the packing interface between P1 and the core domain, and it is interesting to note that a gain-of-function mutation (S220A) in SurA that rescues OM permeability and OMP assembly defects in a BAM-defective strain (*bamA616* containing the mutations R91H, R162L, Y317H in BamA) is located near this crosslink (Ricci et al., 2013; Soltes et al., 2016). A crosslink is also found to the P2 domain, but its distance is far from any other identified crosslink suggests that this domain may be able to ‘hinge’ on the flexible linkers to the P1 and C-domains, bringing into close proximity to the regions of high crosslink density on the C-domain.

3.3.4 Capturing the folding of OmpA through BAM

Tag-transfer crosslinking was used to capture the initial interactions between $\beta 1$ of OmpA as it was folding through BAM via SurA. In this chapter, it was shown that the binding surface of SurA for OmpA was contained almost entirely in its core domains, but this does not preclude the involvement of the P1 and P2 domains of SurA in handover and delivery of OMPs to the BAM complex. The presence of additional crosslinks to the SurA

P2 domain, particularly in the absence of many crosslinks overall (4 of 6 SurA XL were to SurA P2), and only 2 to the core, suggests a greater recruitment of SurA P2 to the OMP client in the presence of the full BAM complex. The crosslinks from OmpA β 1 to BamA show an intriguing pattern. Their presence in POTRA1, POTRA4 and POTRA5, suggests a route of OmpA from ‘arrival’ at the POTRA1 domain, snaking up through the periplasmic cradle of the POTRA domains to the conserved interface between POTRA5 and BamD. Recent evidence suggests that during OMP biogenesis, the C-terminal β -strand of OMPs (containing the β -signal) templates onto β 1 of Omp85 (BamA) homologues, at least for Por1 of *Saccharomyces cerevisiae* and EspP of *E. coli* (Höhr et al., 2018; Doyle and Bernstein, 2019). During this process, β 1 or the most recently appended β -hairpin of the OMP substrate forms weak or transient interactions with β 16 of BamA while β 1 of BamA and the C-terminal β -strand of OMPs form a tight interaction (see Introduction Figure 1.12). The hybrid barrel model suggests this takes place primarily in the membrane, while the elongation model suggests that the β -hairpins elongate in a C- to N-terminal manner in the periplasm (Schiffrin et al., 2017a). The original elongation model suggests that the open state of BAM is the acceptor while an alternative version of this model suggests that the closed state of BAM is the acceptor and a ‘swinging’ motion of β 1 of BamA from the closed to open state pushes the nascent barrel up and into the destabilized membrane generated by BAM concurrent with closing of β 1 and the terminal β -strand of the substrate (Doyle and Bernstein, 2019) (see Introduction Figure 1.12 and Figure 3.32). Alternatively, the ‘budding’ model of biogenesis suggests that β 16 of BamA plays a larger role in coordinating the growth of the nascent OMP substrate barrel until all the secondary structure is formed and then the substrate OMP barrel can close, dissociate from BamA, and diffuse away (Höhr et al., 2018). To assess the feasibility of these models in explaining the observed crosslinks, a model of BamA β 1-templated elongation was built using an 8-strand stretch of LptD which shows little to no curvature (i.e. forming a straight and flat set of 8 β -strands), and an 8-strand stretch of NanC which shows an intermediate level of curvature (compared to the final curvature of an 8-stranded β -barrel) (Figure 3.32). This shows that a linear elongation model is incompatible with any crosslinks observed (and also sterically clashes with the POTRA domains and BAM subunits), but a curved elongation model where OmpA curves into the lumen of BamA and the periplasmic cradle could be compatible with the crosslinks seen in POTRA5 and POTRA4 and eliminate steric clashes of the β -strands of OmpA. Interestingly, in the closed state the length of BamA β 1 up to the location of the glycine kink in β 16 matches well with the length of the C-terminal strand of a substrate. Together these data provide support for the model of periplasmic elongation where closed BAM is the acceptor state. Alternatively, these data do not rule out a model where SurA plays a more significant role in delivery of OMPs. SurA could form a lid on the bottom of the periplasmic ring formed by the POTRA

domains and insert P2 into this cradle to exclude volume in a manner similar to GroEL/ES – lowering the entropic cost of folding by limiting the conformational freedom of the substrate (Hayer-Hartl et al., 2016). This would allow SurA to improve folding in two ways: the chaperoning of free OMP by sequestering it in its core domain could improve folding *yields* by reducing irreversible aggregation (but not affect folding *rate*) and the ‘trapping’/‘spooling’ into the BAM periplasmic cradle could accelerate folding *rates* by reducing the entropic stabilization of intermediates (i.e. intermediates which are kinetically trapped due to their conformational heterogeneity versus the folded state).

BamD has been implicated in recognition of OMP substrates from both indirect evidence (presence of TPR-domains in its fold which are used for C-terminal motif recognition in non-homologous proteins (Albrecht and Zeth, 2011; Sandoval et al., 2011), *bamD* localized suppressor mutations of OMP folding defects (Ricci et al., 2012)), and direct evidence (pull down of BamA and OmpA by His-tagged BamD *in vitro* (Hagan et al., 2013), pull down of BamA(765-779) by His-tagged BamD, titratable inhibition of folding of BamA and OmpA when this BamA(765-779) fragment was added *in trans*, and *in vivo* photocrosslinking (Hagan et al., 2015)). The crosslinks found in POTRA5 at R353 and F354 are near a conserved electrostatic interface between BamA-P5 and BamD (Figure 3.33A) which has been shown to be important for coordination of BamA and BamD OMP folding activities (Sinnige et al., 2015; McCabe et al., 2017). Previous studies have shown that a single point mutant in this network (*bamAE373K*) can break the BAM complex into BamAB and BamCDE subcomplexes and cause a conditionally lethal phenotype, but surprisingly function can be restored by a compensatory mutation in BamD (*bamDR197L/S/H*), or in BamA (*bamAK351E*) despite the complex remaining split (Ricci et al., 2012; McCabe et al., 2017). These studies suggested that substrate binding triggers a conformational change that is transmitted between BamA and BamD, but that formation of the full BamABCDE complex is not essential for this. It was previously hypothesized that binding of a substrate near an “exposed nexus of charged residues” could trigger an allosteric event between BamA and BamD (McCabe et al., 2017). Investigating the interface around these residues and β 1 of BamA (Figure 3.33B-C) shows that in the closed state the crosslinked residues line up well with β 1 of BamA, with access to this strand being somewhat occluded in the open state by turn 7 between BamA β 14 and β 15, and the rotation of the periplasmic region of BAM relative to the barrel. In the closed state a groove formed by POTRA5, POTRA2, BamD and BamC leads to this the exposed edge strand of BamA β 1 and this cavity expands with the switch to the open state (Figure 3.33C - R353 and F354 sit just at the bottom edge of this groove, just out of view in this rendering). It is tempting to speculate that this may represent a channel for shuttling OMPs from an initial ‘weak’ recognition site around POTRA1, possibly mediated by SurA, to higher specificity

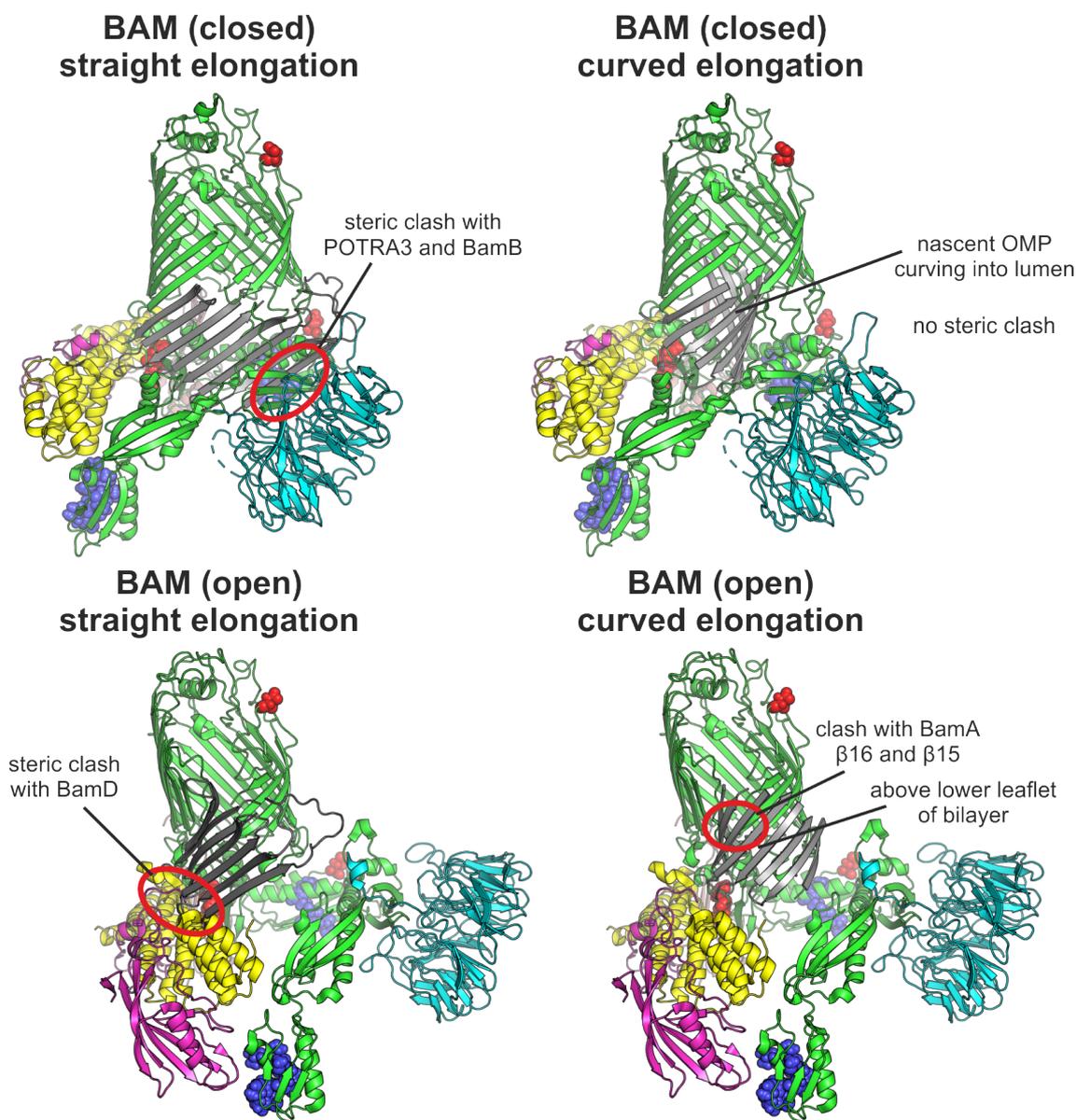


Fig. 3.32 Illustrations of potential modes of the BAM-elongation model. The elongation model of BAM-catalysed OMP folding proposes that the final strand of an OMP substrate templates onto $\beta 1$ of BamA followed by hairpin-by-hairpin polymerisation of strands C- to N-terminally (substrate shown here as grey strands). The compatibility of this model with the observed crosslinks from OmpA $\beta 1$ is assessed by building a model of 8-strands templating and extending linearly or with curvature, from either the open or closed states of the BAM complex. Model built from an 8-strand stretch of the 26-stranded LptD (straight elongation) (PDB: 4RHB) or the 12-stranded NanC (curved elongation) (PDB: 2WJQ).

recognition sites on BamD which could trigger an allosteric change to allow access to $\beta 1$ of BamA and initiate templated folding of OMPs. This possibility could be easily interrogated and validated (or refuted) by using the enhanced protocol for tag-transfer crosslinking from OMPs to BAM discussed in Section 3.2.8.3 and a larger palette of OmpA variants spread through the length of the chain (including the C-terminal ' β '-signal).

3.3.5 Summary

In summary, tag-transfer crosslinkers are a powerful new tool for tackling questions about dynamic and transient protein-protein interactions using XL-MS. The cavity-based chaperoning mechanism of Skp was confirmed and used to validate the tag-transfer system developed using OmpA as a substrate. This system was then used to dissect the binding mechanism of an OMP to SurA and to show, for the first time, that client OMPs bind almost exclusively to the SurA N- and C-terminal domains which make up the 'core'. Finally, with the improvements in crosslinker and lamp technology described in this chapter, and the data from OmpA-SurA binding, a very challenging dynamic transient tripartite system comprising SurA-BAM-OmpA was crosslinked, providing preliminary data on possible mechanisms of SurA-mediated OmpA delivery and folding through BAM.

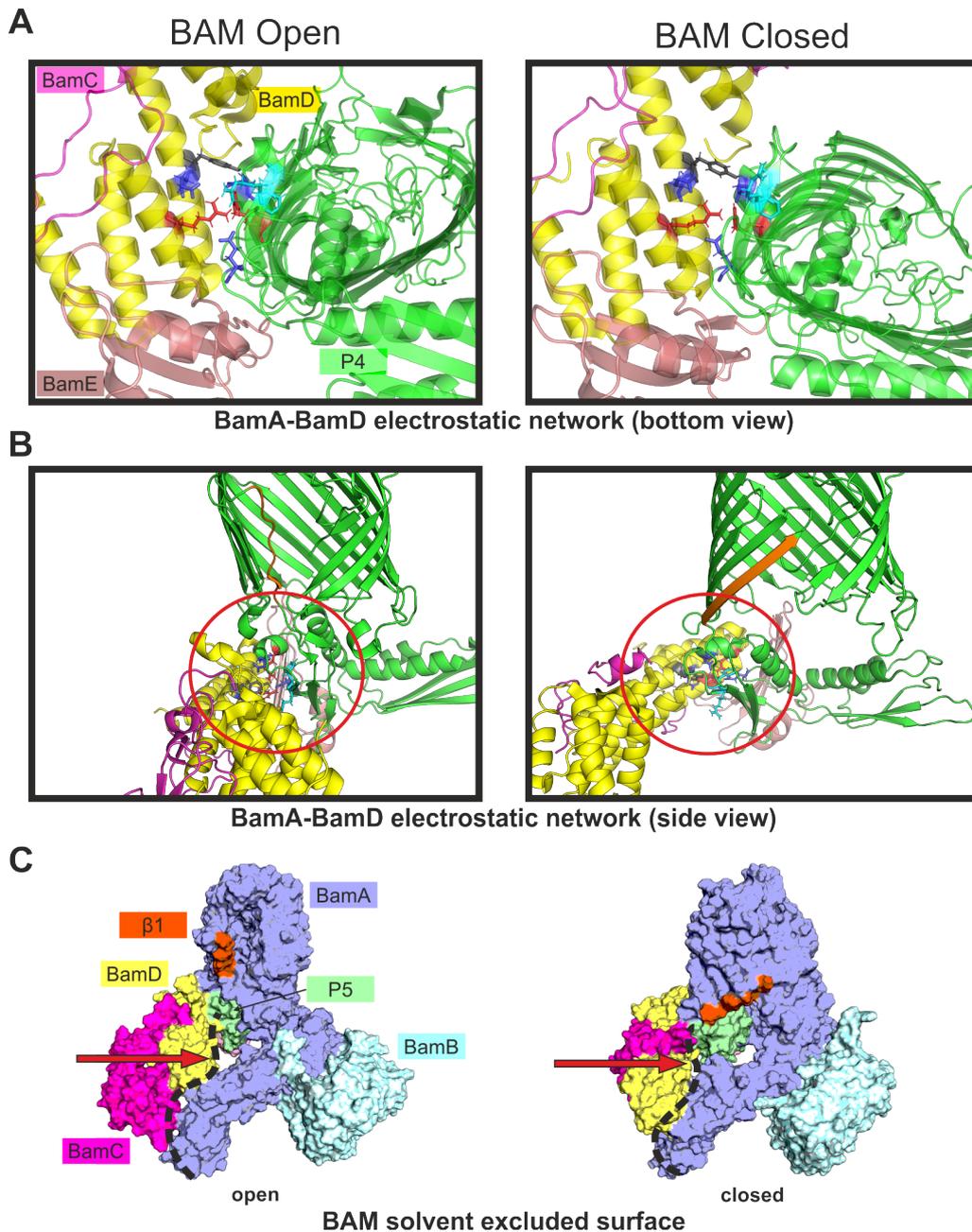


Fig. 3.33 BamA-BamD conserved electrostatic network and crosslinks in this region may couple with access to BamA $\beta 1$. Left and right views of the same location in BAM but for the open (left) and closed (right) states. (A) View of the conserved electrostatic network and crosslinks from OmpA(W7C) with positively charged residues shown in red (BamA-R366/BamD-R197), negatively charged residues shown in blue (BamA-D362/BamA-E373/BamD-E177), aromatics in grey (BamA-Y176), and crosslinked residues in cyan (BamA-R353/BamA-F354). BamA POTRA3-1 and BamB have been hidden. (B) As in (A) but viewed from the side and $\beta 1$ of BamA coloured in orange to give local context as to this electrostatic network (circled in red) (C) Solvent excluded surface of the BAM complex showing the location of a groove leading from periplasm to the exposed edge of BamA $\beta 1$ (groove indicated by a red arrow). The proposed route of a nascent OMP based on crosslinking data in this chapter is shown as a black dashed line.

Chapter 4

Lipid order and the role of the β -barrel assembly machinery

4.1 Introduction

One of the proposed mechanisms by which the BAM complex accelerates folding of OMPs is by lowering the activation energy barrier imposed by insertion of the polypeptide chain into the lipid bilayer. Bilayer insertion has two major associated energy costs: (1) dehydration of soluble groups on the polypeptide chain, and (2) disruption of lipid-packing (breaking electrostatic and hydrogen bonding interactions between lipid headgroups, and van der Waals interactions between acyl chains). With regards to (1), the instantaneous cost of dehydration may be lowered or removed by coordinating formation of hydrogen-bonding and electrostatic networks with membrane insertion (thereby offsetting the unfavourable enthalpy of breaking hydrogen bonds with formation of new bonds) and current *in vitro* evidence suggests that folding of OMPs proceeds from a membrane-associated state with concurrent formation of secondary and tertiary structure (Kleinschmidt, 2003; Rath et al., 2019). While BAM may aid in this process, the rapid and spontaneous folding of OMPs into detergent and thin or less-packed lipid bilayers in the absence of the BAM complex indicates that (1) is not a considerable kinetic barrier to folding and nature may have encoded the solution within the sequence and folding pathway of these proteins. Instead, the challenge posed by (2) may represent the biggest contribution to the activation energy barrier for OMPs folding *in vivo* and lowering it could represent the most ancestral or 'essential' function of the BAM machinery. BamA is the only protein that is conserved across all bacteria containing an outer membrane, and the only OMP biogenesis-related

bacterial protein with homologues involved in the biogenesis of transmembrane β -barrels in eukaryotes (e.g. Sam50 in *H. sapiens* mitochondria, Toc75 in *Arabidopsis thaliana* chloroplasts) (Paschen et al., 2005).

Attempts to fold varied OMPs into small unilamellar vesicles (SUVs) composed of *E. coli* polar lipids results in moderate folding yields for OmpA, OmpT and BamA, and poor to no folding for OmpX, PagP, OmpW, OmpLA, and FadL (Burgess et al., 2008; Gessmann et al., 2014). Folding into synthetic *diC*_{10:0}PC (DDPC) large unilamellar vesicles (LUVs) could occur spontaneously but was suppressed when the LUVs were doped with lipids containing charged lipid headgroups such as the positively charged phosphatidylethanolamine (PE) or the negatively charged phosphatidylglycerol (PG), and this suppression was partially rescued by the presence of BamA or a truncated version missing the last N-terminal 4 POTRA domains (BamA Δ _{P1-P4}) (Gessmann et al., 2014). These results suggested that spontaneous folding of OMPs *in vivo*, into a membrane which does not contain any zwitterionic PC headgroups, is kinetically repressed unless BamA is locally present. This kinetic affect could be due to increased lateral bilayer pressure and tighter packing of the PE/PG containing membranes (Murzyn et al., 2005), but it is possible that it is a charge-dependent effect. The electrostatic profile of the BAM complex, particularly BamA, does not suggest a mechanism for charge screening or neutralization of charges at the periplasmic face of the bilayer which might overcome this, making a charge-based kinetic relief mechanism unlikely (Figure 4.1). Instead, given BamA's position as an integral membrane protein, and the ability of the barrel domain only of BamA to relieve kinetic inhibition by PE phospholipids, a lipid-disordering mechanism is a more feasible explanation. A number of molecular dynamics studies on BamA and the full BAM complex have shown that BamA can switch from a closed to an open conformation, and also observed thinning and disordering of the membrane in the vicinity of the β 1- β 16 seam of BamA (Noinaj et al., 2013, 2014; Gu et al., 2016; Fleming et al., 2016; Schiffrin et al., 2017b; Lundquist et al., 2018). There are also a number of crystal and solution structures of BamA in isolation or as part of the BAM complex which show that the barrel is thinner in the β 1- β 16 region and that the BamA barrel can explore at least three distinct and potentially membrane-influencing conformations: a closed fully zipped barrel, a closed partially zipped barrel, and an open barrel (Noinaj et al., 2013; Ni et al., 2014; May and Grabowicz, 2018a; Bakelar et al., 2016; Gu et al., 2016; Han et al., 2016; Iadanza et al., 2016; Gu et al., 2017; Hartmann et al., 2018; Kaur et al., 2019) (Figure 4.2). BamA has also been shown to have a greater catalytic effect on tOmpA folding (greater degree of catalytic fold rate enhancement) as the hydrophobic thickness of the bilayer is increased from \sim 19.5 Å in *diC*_{12:0}PC (DLPC) to \sim 23.0 Å in *diC*_{14:0}PC (DMPC) LUVs suggesting that hydrophobic mismatch and/or lipid disordering plays an important role in the mechanism

of BamA-assisted folding (Schiffirin et al., 2017b). Despite the relatively large number of studies on this topic, there currently only exists indirect structural, molecular biology, and *in silico* evidence for the proposed membrane remodelling capacity of BamA.

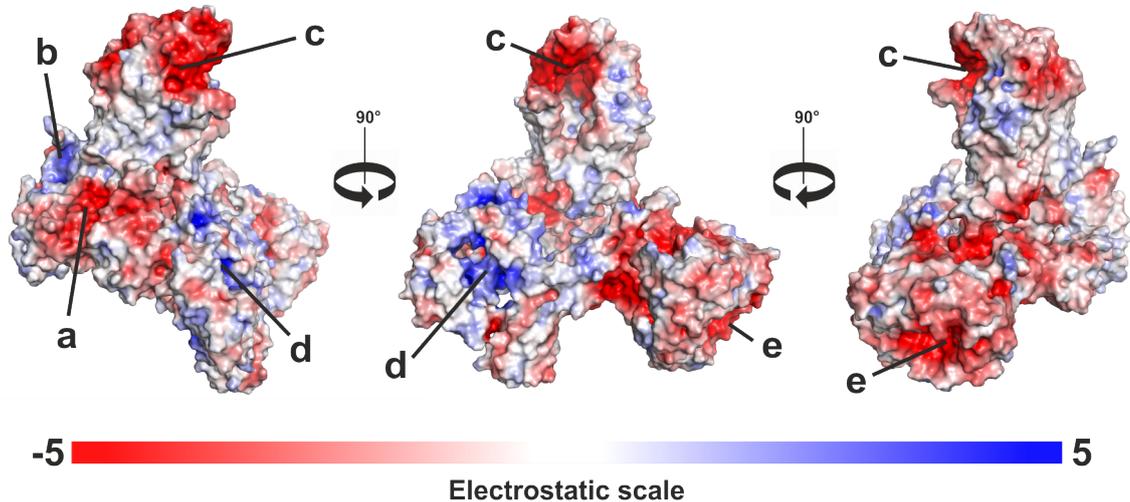


Fig. 4.1 The electrostatic profile of the BAM complex in its open state as viewed from three angles around the lateral gate. There are some charged hotspots located in the BAM complex such as the negative charges around the back of BamD (a), the face of BamB (e), and in the 'lid' of the BamA loops (c), and positively charged patches on BamD near POTRA1-2 (d). However, none of these patches come into direct contact with the membrane apart from a positively charged patch in BamE (b) suggesting that electrostatic interactions with charged headgroups are unlikely to play a major role in the mechanism of BAM-mediated OMP folding catalysis. Structure of the open state of the full BAM complex from PDB: 5LJO (Iadanza et al., 2016).

Pure lipid mixtures, such as DMPC, often show two clear transitions between gel ($L_{\beta'}$), ripple ($P_{\beta'}$), and liquid crystalline (L_{α}) phases as the temperature is increased (Riske et al., 2009; Akabori and Nagle, 2015). The gel to ripple transition is relatively small and broad and usually called the pre-transition (T_p), while the ripple to liquid crystalline phase change shows a sharp intense change termed the main transition (T_m) (Figure 4.3). These phases have two defining properties: their reversibility, and the characteristic temperature at which they occur. For DMPC, the pre-transition has a very low enthalpy ($\sim 3 \text{ kJ.mol}^{-1}$) and is centred at $14 \text{ }^{\circ}\text{C}$, however, it is not always observed experimentally (Prenner et al., 1999). The main transition has a large enthalpy change associated with it ($\sim 25\text{--}26 \text{ kJ.mol}^{-1}$) and occurs around $24 \text{ }^{\circ}\text{C}$ (Bonora et al., 2005; Sanderson, 2005). These transition temperatures can be altered by solutes which bind to lipid headgroups, or partition deeper into the membrane, as well as by peptides and proteins. Membrane-interacting peptides (such as gramicidins) and membrane proteins which have a propensity to cluster (such as bacteriorhodopsin) can reduce the enthalpy of transitions, lower the transition temperature by as much as $3 \text{ }^{\circ}\text{C}$, induce lipid domain formation, and broaden the transition region (Sankaram et al., 1994; Schram and Thompson, 1997; Píknová et al., 1997). During the transition between L_{α} (liquid) and $P_{\beta'}$ (ripple) phases the packing of

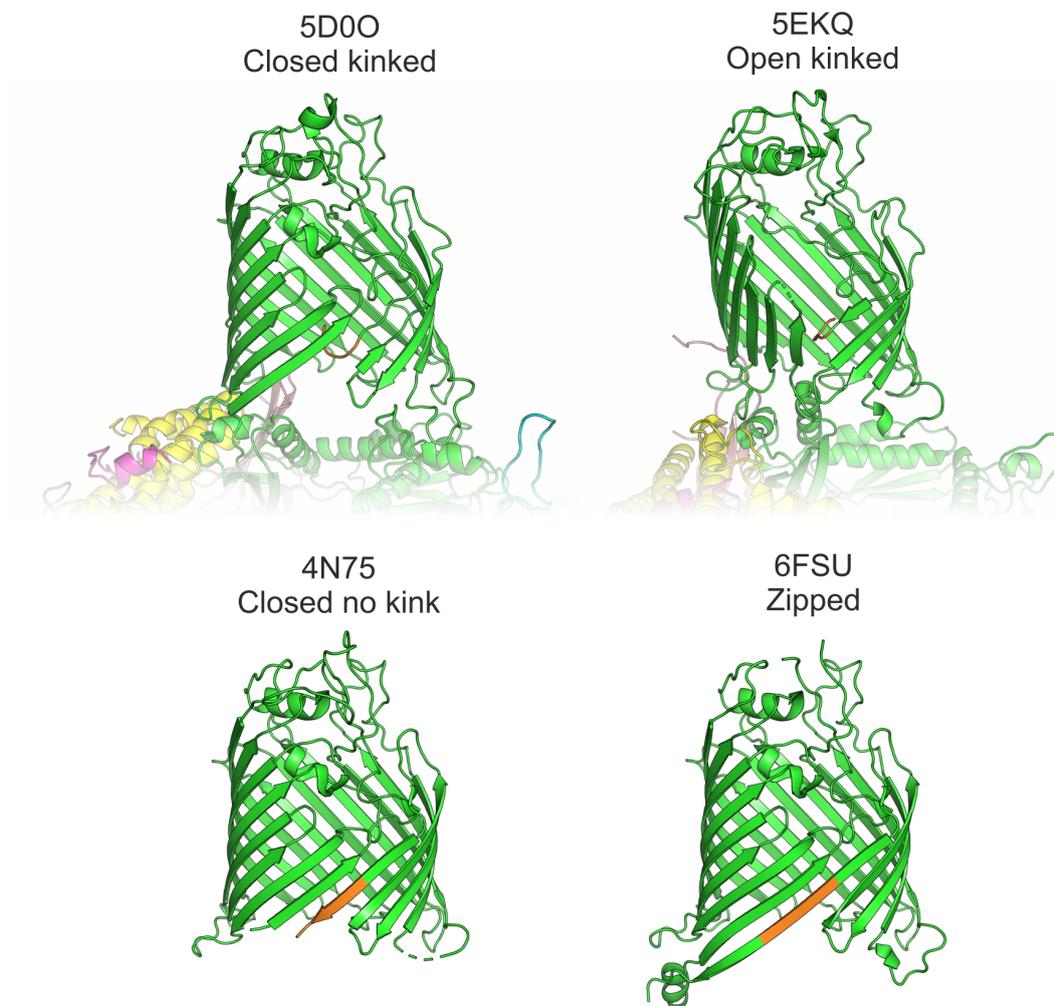


Fig. 4.2 Example structures of *E. coli* BamA adopting different conformations around the location of the β 1- β 16 seam. In all structures of the BAM complex to date β 16 of BamA adopts a kinked conformation at a highly conserved glycine (G807) in both the open (PDB: 5EKQ) (Bakelar et al., 2016) and closed (PDB: 5D0O) (Gu et al., 2016) states of the gate. Residues comprising the kink region are indicated in orange (I806–W810). This kink is also observed in structures of BamA from *Salmonella enterica* (PDB: 5OR1) (Gu et al., 2017) and *Neisseria gonorrhoeae* (PDB: 4K3B) (Noinaj et al., 2013) (not shown) and in the BamA homologue, TamA, which also plays a role in OMP assembly (not shown). BamA with a closed gate and no kink has been observed in isolation (4N75) (Ni et al., 2014), and in a hybrid BamA containing a C-terminal 9-residue extension comprised of part of turn 3 and β 7 from OmpX which may represent a mimic of an OMP-BamA folding intermediate.

the lipids undergoes a dramatic change which has been reported to cause local packing defects due to co-existing phases (Nielsen et al., 2000a,b; Enders et al., 2004). This occurs due to nanoscale fluctuations interacting and building into macroscopic scales at a critical temperature. The size of domains is highly heterogeneous and although they can spread from nucleation points, appear concurrently across the whole bilayer. This kind of behaviour suggests that there is not a characteristic length scale and that the transition is coupled across the whole bilayer. The phase transition of the lipid bilayer may also play an important role in modulating the function of proteins associated with it, such as protein binding and enzyme activity, either through direct lipid-protein interactions, or through the formation of nanoscopic lipid domains (Hønger et al., 1996; Gil et al., 1998; Mouritsen et al., 2006). However, at the transition temperature the permeability of a membrane is maximal (Heimburg, 2007) – likely due to defects at domain boundaries – and so it is likely that *in vivo* the formation of bilayer regions with characteristics of the transition temperature would be tightly controlled, or localised to specific regions where it can be carefully monitored in order to maintain the barrier function of the membrane.

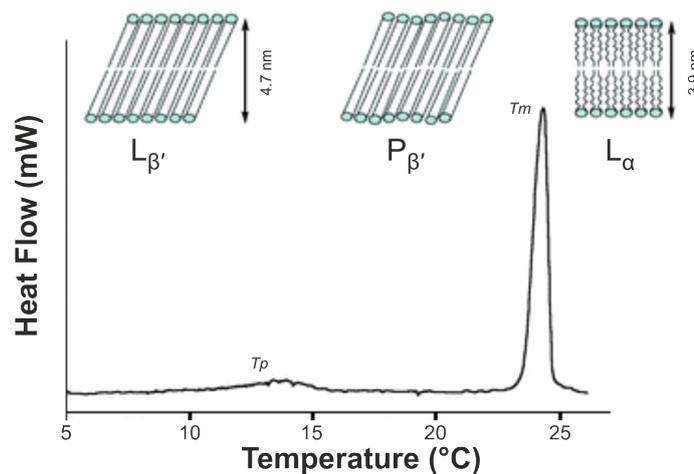


Fig. 4.3 DMPC undergoes a major and minor phase transition in response to temperature. Top: change in acyl chain packing and order between different phases. Bottom: DSC thermogram showing broadness of transition and degree of enthalpy change. Figure adapted from Sanderson, 2005.

In vitro folding studies of the transmembrane domain of OmpA (OmpA₁₋₁₇₁, henceforth, tOmpA) into LUVs of DMPC at temperatures below (20 °C), at (24 °C), and slightly above (25 and 26 °C) the transition temperature showed that the folding rate is maximal at or near the transition temperature, but rapidly fell when the bilayer was in the gel or fluid phase (Danoff and Fleming, 2015b). Danoff & Fleming speculated that OMPs may utilize packing defects during folding, and that the folding rate is maximal at the transition temperature where the incidence of defects is maximal. It has also been shown that the catalytic enhancement of tOmpA folding by BamA at 30 °C is greater in thicker lipids in

the order DMPC > DTPC > DLPC (Schiffrin et al., 2017b). In this chapter, the possibility that BamA functions by affecting the packing of lipids and experimentally simulating these sorts of packing defects is explored. The role of the additional subunits in the BAM complex of *E. coli* (BamBCDE) and the major BAM-interacting periplasmic chaperone, SurA, in this process is also assessed.

4.2 Results

4.2.1 Designing an assay to assess the impact of lipid phase and packing on catalysed and uncatalysed OMP folding

The yield of folded protein for most OMPs can be assessed by exploiting the resistance of the β -barrel domain to unfolding by SDS and running the protein on an SDS-PAGE gel without first boiling the samples. The folded and unfolded states will run with different electrophoretic mobilities with the folded band disappearing and the intensity of the unfolded band increasing upon boiling (Nakamura and Mizushima, 1976). The relative proportion of folded (SDS-resistant) and unfolded (SDS-labile) protein can be assessed by measuring the intensity of the bands through densitometry (Surrey and Jähnig, 1992). The progress of folding along the reaction coordinate can be assessed by quenching a sample of an *in vitro* folding reaction in SDS-containing PAGE loading buffer at set timepoints. SDS can bind to the unfolded polypeptide chain and prevent any further folding while the already folded barrel remains stable due to the high kinetic barrier to unfolding in SDS (estimated unfolding rate for OmpA at 30 °C: $\sim 10^{-8} \text{ s}^{-1}$, average lifetime $[\tau] > 1 \text{ year}$) (Ohnishi and Kameyama, 2001). This method, termed Kinetics of Tertiary Structure formation by Electrophoresis (KTSE), is well-established in extracting kinetic information about OMP folding (Kleinschmidt and Tamm, 1996; Schüßler et al., 2019).

tOmpA was chosen as the substrate for these experiments due to the wealth of knowledge and biophysical characterisation of its *in vitro* spontaneous folding into various lipids and detergents accumulated over the last 30+ years, its position as the most abundant OMP in *E. coli* (Henning et al., 1973; Rosenbusch, 1974; Li et al., 2014), the experimentally tractable rate of folding into DMPC LUVs, and the separation of its folded and unfolded state electrophoretic mobilities from SurA and the components of the BAM complex (see below). To reduce the complexity of the system, and to pare down variables in order to create a biophysical model of the environment of BAM and OMPs, a single saturated lipid system using DMPC was chosen as the lipid bilayer. The outer membrane of *E. coli* is

approximately the same hydrophobic thickness as a DMPC bilayer (~23–25 Å for the outer membrane, ~23–26 Å for DMPC in the liquid phase, ~29–30 Å for DMPC in the gel phase) (Lewis and Engelman, 1983; Balgavý et al., 2001; Tristram-Nagle et al., 2002; Kucerka et al., 2005; Pencer et al., 2005; Lomize et al., 2006; Wu et al., 2014; Kim et al., 2016b), the PC headgroup is zwitterionic and bilayer-forming, and its transition temperature sits conveniently around room temperature making it amenable to experimental manipulation (Table 4.1). Furthermore, this synthetic lipid system allows a comparison between BAM-catalysed, BamA-catalysed, and uncatalysed folding as tOmpA can fold into DMPC under all these conditions on a tractable timescale. Most *in vitro* studies on the BAM complex to date have been conducted in bilayers formed from *E. coli* polar lipid extract (Hagan et al., 2010, 2015). While this is a better mimic of the headgroup distribution encountered by OMPs *in vivo* (PE, PG and cardiolipin – PC headgroups are not native to *E. coli*), the drawback is that OMPs cannot fold spontaneously into this lipid type on measurable timescales (data not shown) and BamA alone does not appear to be functional for folding of OmpT (i.e. it cannot catalyse OmpT's folding), and very weakly active for folding of BamA (as a substrate) when reconstituted into *E. coli* polar lipid proteoliposomes (Hagan et al., 2013; Iadanza et al., 2016). *E. coli* is also tolerant of modifications to lipid headgroups and lipid types with synthetic viable strains having been developed where PE synthesis is eliminated (DeChavigny et al., 1991), PG synthesis is eliminated (Kikuchi et al., 2000; Matsumoto, 2001), cardiolipin synthesis is eliminated (Tan et al., 2012), PC synthesis is induced synthetically (Chen et al., 2009), gluco- or galacto-lipids utilised (Wikström et al., 2009), or even archaeal lipids incorporated into the membrane (Caforio et al., 2018). Although in many of these strains growth is affected and some membrane defects are observed, it highlights the fact that there is no absolute requirement for particular lipid types in OM assembly in *E. coli*. It should be noted also that *E. coli* polar lipid extract itself forms only an approximate mimic of the outer membrane as it does not contain lipopolysaccharide (LPS) which is present in the outer leaflet of the *E. coli* outer membrane, it does not produce an asymmetric leaflet structure, and has a greater hydrophobic thickness.

Folding rates of OMPs into lipid bilayers have been shown to be affected by a number of parameters including the degree of bilayer curvature (Pocanschi et al., 2006b), curvature elastic stress (Huysmans et al., 2012), lipid phase (Danoff and Fleming, 2015b), hydrophobic thickness (Kleinschmidt and Tamm, 2002), lipid-to-protein ratio (LPR) (Kleinschmidt and Tamm, 2002; Huysmans et al., 2012), OMP concentration (Kleinschmidt and Tamm, 2002), temperature (Pocanschi et al., 2006a), pH (Moon et al., 2011), denaturant concentration (Andersen et al., 2012), and lipid headgroup (Huysmans et al., 2012; Gessmann et al., 2014). All experiments described here were carried out in 20 mM Tris-Cl, 150 mM

Table 4.1 Comparison of transition temperatures for saturated PC phospholipids, common unsaturated PC lipids, and *E. coli* polar lipid extract (Koynova and Caffrey, 1998; White et al., 2000).

Phospholipid		Transition Temperature, T_m (°C)
Shorthand	Abbreviation	Full name
Saturated		
<i>diC</i> _{18:0}	DSPC	1,2-distearoyl-sn-glycero-3-phosphocholine
<i>diC</i> _{17:0}	DHDPC	1,2-diheptadecanoyl-sn-glycero-3-phosphocholine
<i>diC</i> _{16:0}	DPPC	1,2-dipalmitoyl-sn-glycero-3-phosphocholine
<i>diC</i> _{15:0}	DPDPC	1,2-dipentadecanoyl-sn-glycero-3-phosphocholine
<i>diC</i> _{14:0}	DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
<i>diC</i> _{13:0}	DTPC	1,2-ditridecanoyl-sn-glycero-3-phosphocholine
<i>diC</i> _{12:0}	DLPC	1,2-dilauroyl-sn-glycero-3-phosphocholine
<i>diC</i> _{11:0}	DUPC	1,2-diundecanoyl-sn-glycero-3-phosphocholine
<i>diC</i> _{10:0}	DDPC	1,2-didecanoyl-sn-glycero-3-phosphocholine
Unsaturated		
<i>C</i> _{16:0} / <i>C</i> _{18:1}	POPC	1-palmitoyl-2-oleoyl-glycero-3-phosphocholine
<i>diC</i> _{18:1}	DOPC	1,2-dioleoyl-sn-glycero-3-phosphocholine
Other		
-	<i>EcPL</i>	Avanti <i>E. coli</i> polar lipid extract

NaCl, 0.8 M urea, pH 8.0, at a lipid-to-protein ratio (LPR) of 1600:1 (mol/mol), with 1 μ M tOmpA, in DMPC lipid bilayers. This controls the effect of hydrophobic thickness, LPR, OMP concentration, pH, denaturant concentration and lipid headgroup. Empty liposomes, and BamA- or BAM-containing proteoliposomes were prepared in the same way (see Materials & Methods Section 2.4.10 and Section 2.4.17). Folded BamA or BAM solubilised in detergent was diluted to around the critical micelle concentration of the detergent by addition to a suspension of DMPC lipids. This mixture was then dialyzed for 3-4 days with regular changes of the dialysis buffer to remove residual detergent and form proteoliposomes. The process was the same for empty liposomes with the same mass of DMPC added, except an equivalent volume of dodecylmaltoside (DDM) detergent replaced the solubilised BamA/BAM protein. After dialysis, the liposomes or proteoliposomes were extruded through 100 nm pores at 37 °C (well above the T_m of DMPC) to form highly monodisperse, ~100 nm diameter LUVs (Figure 4.4). The process of extrusion did not have any significant effect on the rate of folding into BAM proteoliposomes at 24 °C (Figure 4.5). The assembly and insertion of outer membrane proteins *in vivo* has been observed to occur speckled across the outer membrane as seen for LamB in *E. coli* (Ursell et al., 2012) and OmpF in *Salmonella typhimurium* (Smit and Nikaido, 1978), localised at the old pole seen for IcsA in *Shigella flexneri* (Steinhauer et al., 1999; Charles et al., 2001), or occurring primarily at the mid-cell (i.e. new pole or constriction point of dividing bacteria) seen for Cir and BtuB in *E. coli* (Rassam et al., 2015) and Omp25 in *Brucella abortus* (Vassen et al., 2019). In each case the curvature is small or negligible and while

100 nm LUVs display a curved surface, on the length-scale of tOmpA this curvature is also very small (Figure 4.6). The remaining uncontrolled variable is temperature as the contribution of temperature to folding rates is inherently convolved with the phase changes of DMPC that are being induced. However, over the 10 °C range in this experiment it is not expected that the thermal energy imparted to the unfolded state of tOmpA will contribute significantly to the intrinsic ability of tOmpA to overcome the activation energy barrier, as the folding rate of OmpA into A8-35 amphipols shows only a weak temperature dependence (Pocanschi et al., 2013). The stoichiometry of BamA or BAM to tOmpA was always 2:1 (mol/mol) to ensure that the catalyst (BamA/BAM) could never be saturated by substrate. Each condition was also run with and without SurA to assess the effect of this chaperone on the rate of tOmpA folding. In SurA-containing experiments, tOmpA was initially diluted into SurA-containing buffer to a concentration of 10 μM tOmpA, 100 μM SurA, with 1 M urea, and this was then immediately added to the final mix containing either empty DMPC liposomes, or BamA/BAM-containing proteoliposomes to give a final concentration for the folding reaction of 1 μM tOmpA, 10 μM SurA. The K_d of His-tagged (HT)-SurA for tOmpA is 1.8 μM (Humes et al., 2019) and so tOmpA can be expected to be ~98 % bound to SurA in the pre-incubation step and ~84 % bound at the initiation of folding (assuming equilibrium is reached quickly) (Figure 4.7). This assay and its parameters are summarised in Figure 4.8.

At the end of each folding experiment, samples were run directly on a 15 % Tris-tricine SDS-PAGE gel, stained using a Coomassie-based stain, and imaged (see Materials & Methods Section 2.4.8). Example gels for each folding condition at 24 °C are shown in Figure 4.10 and the full set of gels can be found in Appendix C. The fraction folded was calculated as the ratio of the intensity of the folded bands to the sum of the folded and unfolded monomer bands, as measured by densitometry, and calculated according to Eq. 1. The shape of the kinetic trace was variable and did not always fit well to a single exponential (Figure 4.9A), as might be expected for a rugged energy landscape where an OMP might populate a number of folding intermediates, fold via parallel pathways, or diverge to off-pathway misfolded or aggregated states – all of which have been reported in the literature (Kleinschmidt and Tamm, 1996, 1999; Pocanschi et al., 2006a; Kang et al., 2012; Andersen et al., 2012; Huysmans et al., 2012; Danoff and Fleming, 2017). The limited number of data points available from KTSE experiments, as compared to spectroscopic methods such as using intrinsic fluorescence, means that fitting to more complicated kinetic schemes would be under-restrained and likely result in overfitting of the data. However, most experiments fitted well to a double-exponential (Figure 4.9B) (with the exception of 30 °C empty liposomes which fitted well to single exponential equations – see Appendix D) allowing the rate of folding to be compared quantitatively using

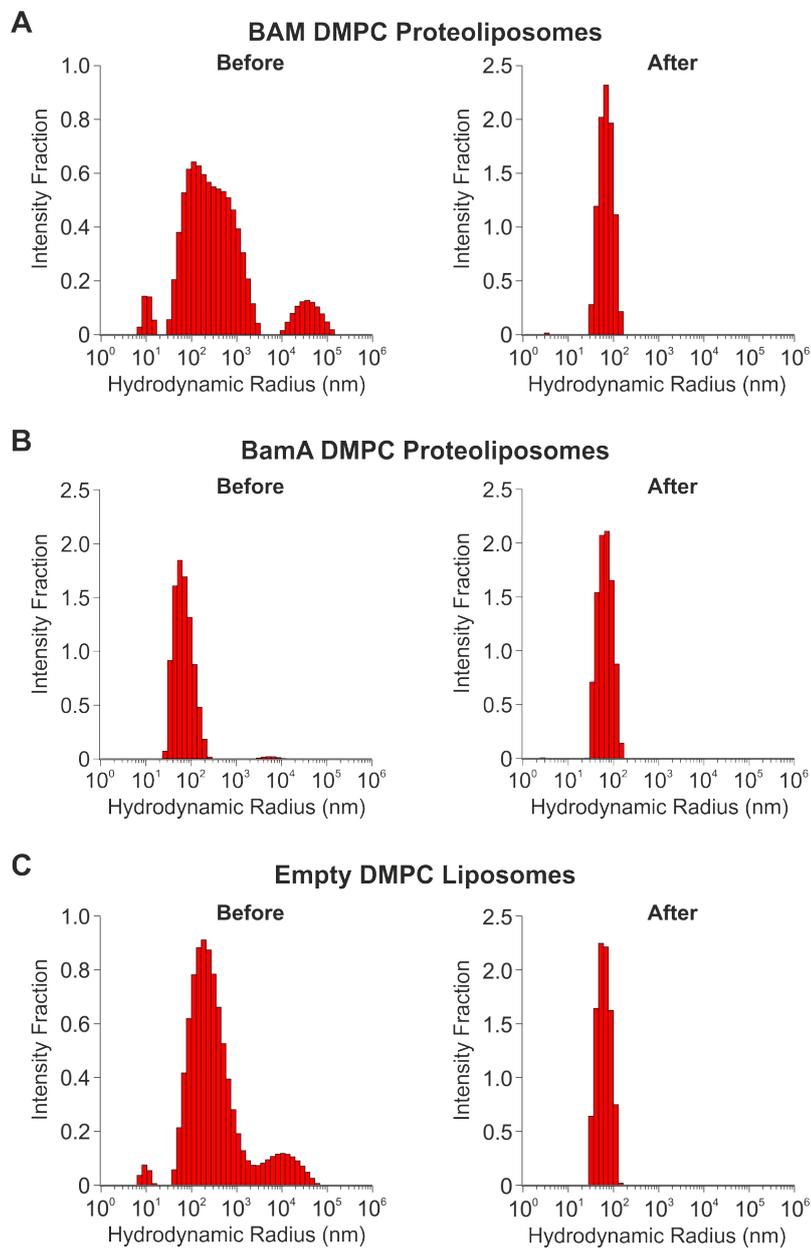


Fig. 4.4 Dynamic Light Scattering (DLS) intensity weight plot of hydrodynamic radius versus intensity of scattering for liposomes formed by dialysis. After dialysis and before extruding, samples show a variable and broad distribution of liposome sizes from small (10–30 nm diameter) to giant (>1000 nm diameter). After extrusion all samples show narrow distributions with hydrodynamic radii centred around 50–70 nm (100–140 nm diameter). All samples in 150 mM NaCl, 20 mM Tris-Cl pH 8.0.

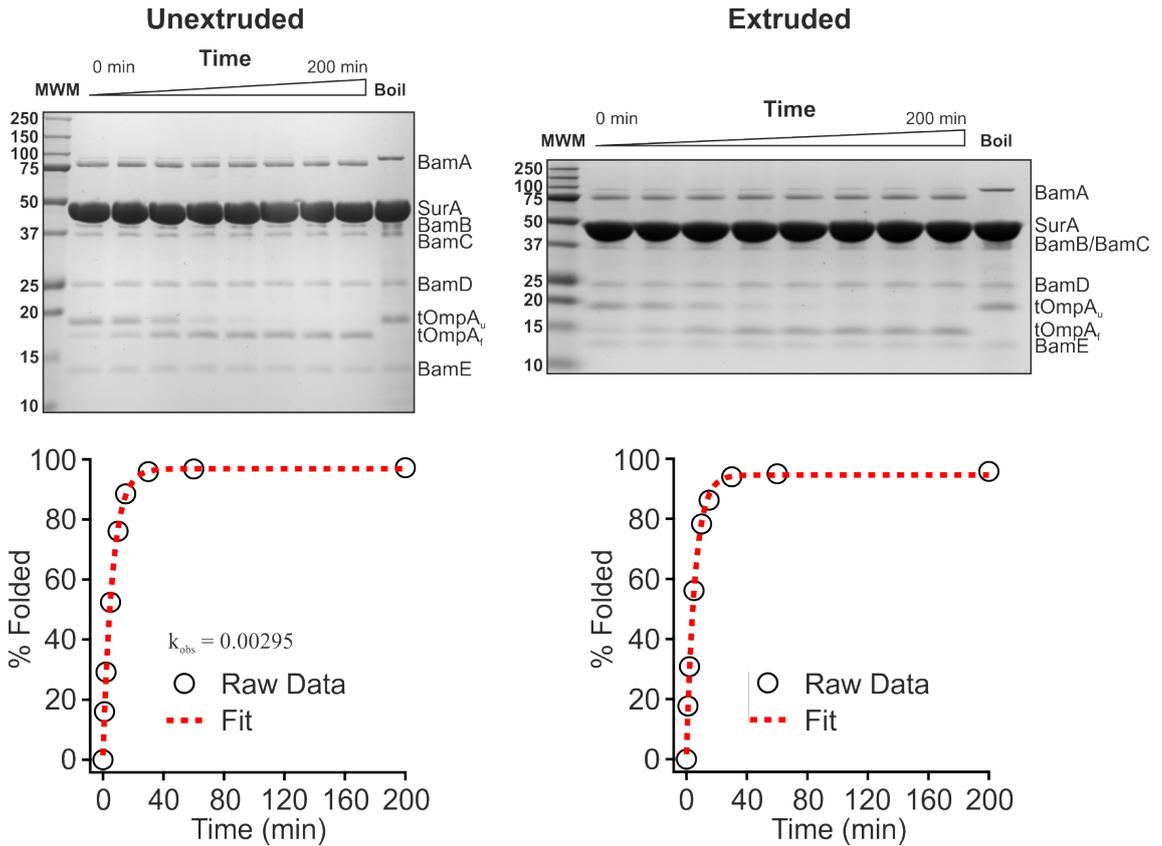


Fig. 4.5 Extrusion has no effect on the integrity of the BAM complex or the rate of folding. tOmpA was folded into 1600:1 LPR (mol/mol) BAM DMPC proteoliposomes at 24 °C under the following conditions: 1 μ M tOmpA, 0.5 μ M BAM, 10 μ M SurA, 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0. The kinetic data fitted well to a single exponential curve of the form $y = y_0 + A \exp^{-k_{obs}t}$ where y_0 is the y-intercept (and represents the yield), A is the amplitude, k_{obs} is the observed rate constant, and t is the time elapsed. Both sets of proteoliposomes gave similar observed rate constants and yields. Unextruded: $k_{obs} = 0.0026 \pm 0.0008 \text{ s}^{-1}$, $t_{50} = 258.9 \text{ s}$, $y_0 = 96.9 \pm 0.7 \%$. Extruded: $k_{obs} = 0.0030 \pm 0.0001 \text{ s}^{-1}$, $t_{50} = 229.1 \text{ s}$, $y_0 = 94.6 \pm 0.8 \%$. This shows that extrusion does not have a significant effect on the activity of BAM proteoliposomes.

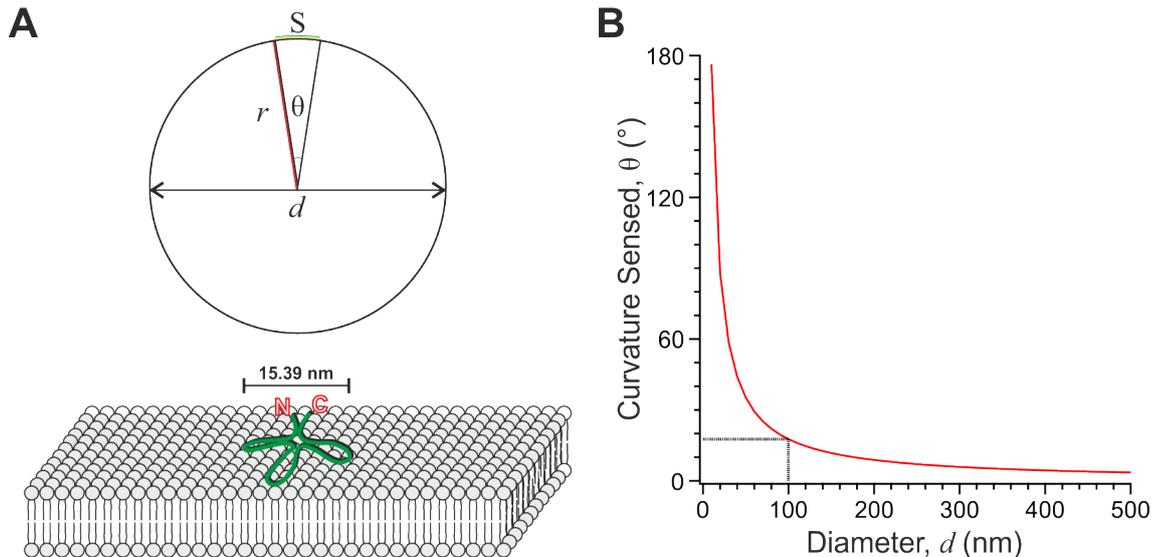


Fig. 4.6 Degree of curvature 'sensed' by tOmpA upon contact with a lipid bilayer. For spontaneous folding of tOmpA into lipid bilayers, a 'rosette'-like initial organisation (see A, bottom) of the polypeptide chain upon binding to a bilayer surface has been proposed based on data from site-directed fluorescence quenching, time-resolved circular dichroism, and kinetic modelling (Kleinschmidt et al., 2011; Danoff and Fleming, 2017). In this model, an unstructured membrane-associated state quickly rearranges to a state with loosely organised β -hairpins but little secondary structure. The periplasmic-facing turns are at the 'edges' of the rosette while the residues which will traverse through the bilayer are located nearer the centre. (A, bottom) The width of this membrane-associated 'rosette' can be approximated by assuming the formation of equal-length hairpins (which will also include loops and turns). tOmpA is a 171-residue 8-stranded OMP. Each residue of an extended polypeptide chain contributes ~ 0.36 nm to its length, so for tOmpA, the distance from the edge of one side of the 'rosette' to another is: $[(0.36 \text{ nm} \times 171 \text{ aa}) / 8 \text{ strands}] \times 2 = 15.39$ nm. (A, top) On the surface of a sphere (our LUVs) this width corresponds to an arc length, S , and knowing the diameter, d , and therefore radius, r , of the liposomes we can calculate the arc radius, θ , which is the curvature 'sensed' by a membrane-associated OMP. The green line marked at S represents the true relative size of tOmpA on a 100 nm diameter liposome. (B) The change in relative curvature, θ , sensed by tOmpA as the diameter the liposome increases. 100 nm is marked by black dotted lines.

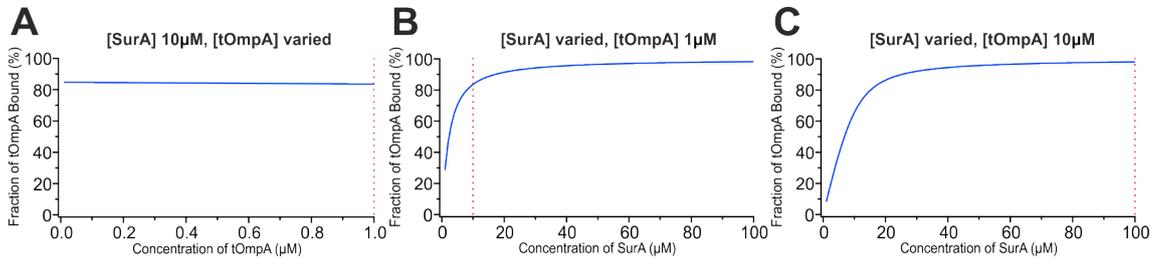


Fig. 4.7 Comparison of tOmpA occupancy on SurA as concentration of one species is varied. Fraction of tOmpA bound to SurA was calculated using a quadratic binding equation, assuming a 1:1 stoichiometry of tOmpA with SurA. Actual concentrations used are indicated by vertical dashed red lines. (A) The final reaction mixture contains 10 μM SurA and 1 μM tOmpA, however, as more tOmpA folds it removes itself from the pool of free substrate for SurA to bind. At these concentrations, this barely changes the occupancy of free tOmpA on SurA. (B) With a final concentration of 1 μM tOmpA, 10 μM SurA strikes a balance between saturating the SDS-PAGE gel with SurA (at higher SurA concentrations) and having very little tOmpA bound to SurA (at lower SurA concentrations). (C) In the pre-mixing of tOmpA and SurA, tOmpA should be almost completely bound and therefore protected from aggregating before folding is initiated.

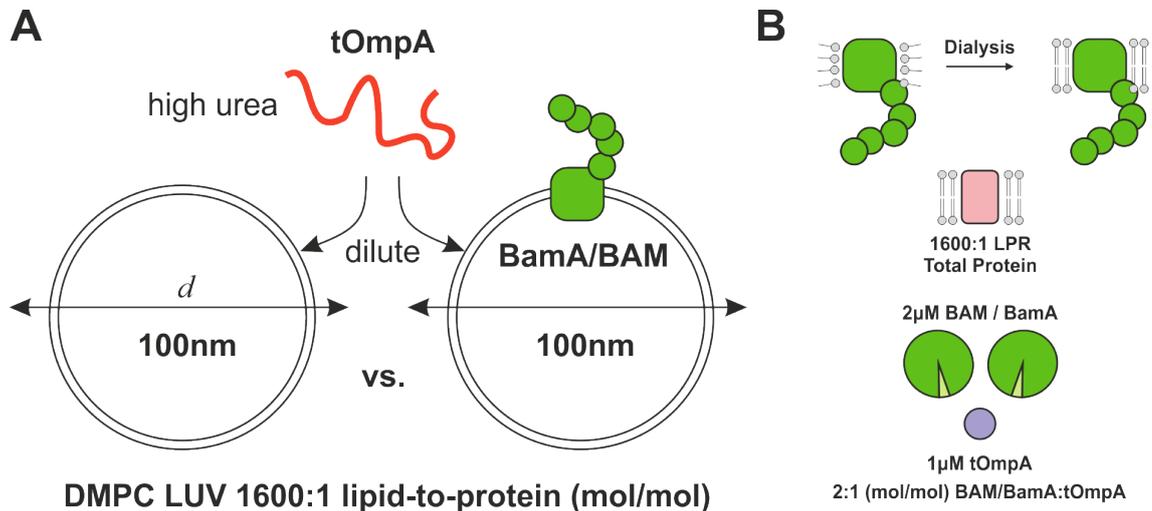


Fig. 4.8 Schematic illustration of kinetic folding experiments. tOmpA is rapidly diluted from high urea (8 M) in the presence of either empty liposomes, or liposomes containing BamA or the BAM complex, each of which have a diameter of ~ 100 nm. Liposomes are all formed by dialysis, the LPR of total membrane embedded protein is 1600:1 (mol/mol), and the ratio of BAM or BamA to tOmpA is always 2:1 (mol/mol).

the phenomenological measure of t_{50} , the time taken to reach 50 % of the maximum folding yield in that experiment. This was extracted from the data by finding a numerical solution to Eq. 2 using the parameters derived from the double exponential fit. Fits to single and double exponential equations can be found for all folding experiments in Appendix D.

Equation 1:

$$\text{Fraction folded} = \frac{\text{intensity folded band}}{(\text{intensity folded band} + \text{intensity unfolded band})}$$

Equation 2:

$$y = y_0 + A_1 \exp^{-k_1 t} + A_2 \exp^{-k_2 t}$$

where y is equal to $\frac{y_0}{2}$ at t_{50} ; y_0 is the y-intercept; A is the amplitude; k is the observed rate constant; t is the time elapsed

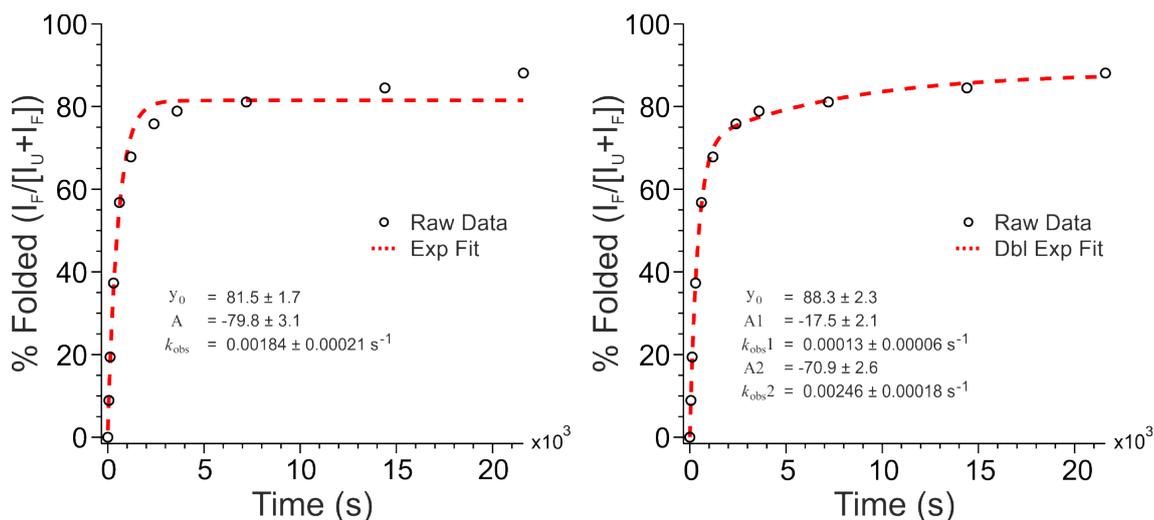


Fig. 4.9 Example fits from a gel phase (30 °C) folding experiment of tOmpA into BAM-containing DMPC proteoliposomes. Final conditions: 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0, 1 μ M tOmpA, 2 μ M BAM, 1600:1 (mol/mol) LPR DMPC proteoliposomes. Shown are single or double-exponential fits (red dashed line) to the raw data from densitometry performed on SDS-PAGE gels of tOmpA folding kinetics. Single exponentials were fit of the form $y = y_0 + A \exp^{-k_{\text{obs}} t}$ where y_0 = y-intercept, A = amplitude, t = time (x-axis), and k_{obs} = the observed rate constant. Double exponential fits were of the form $y = y_0 + A_1 \exp^{-k_{\text{obs}1} t} + A_2 \exp^{-k_{\text{obs}2} t}$. Fitted parameters are indicated in the inset of the graphs. Visual inspection of the graphs shows that the single exponential underestimates the final folded yield due to an additional slow phase which the double exponential is able to capture.

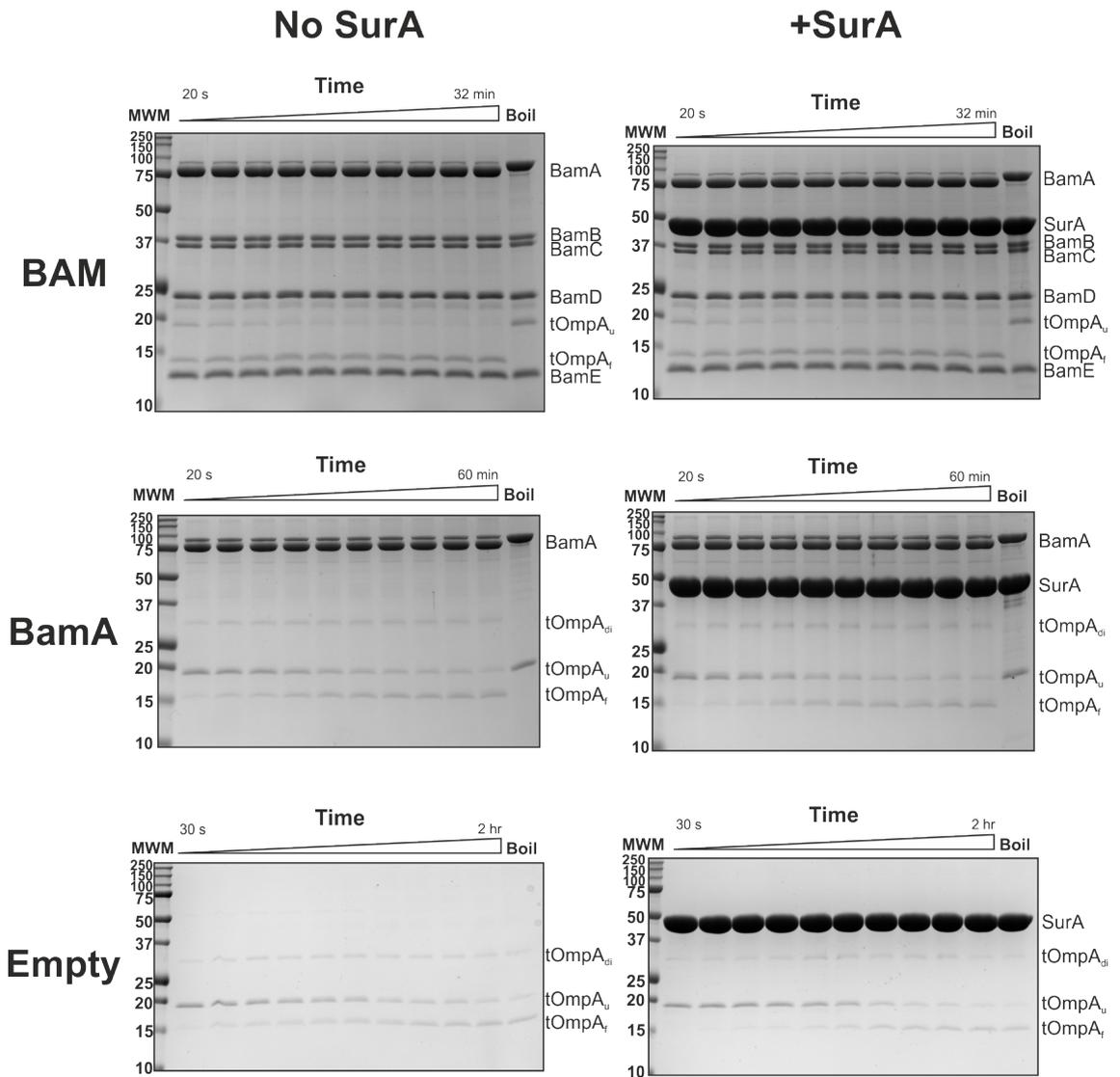


Fig. 4.10 Example SDS-PAGE gels for each folding condition at the transition temperature (24 °C). tOmpA_u = unfolded tOmpA, tOmpA_f = folded tOmpA, tOmpA_{di} = tOmpA dimer. All gels were run with 0.1 % (w/v) SDS, and folding was quenched by mixing with 6X SDS loading buffer giving 1 % (w/v) final SDS concentration. Under these conditions, BamA, unlike tOmpA, does not show complete retention of the folded state in the absence of boiling.

4.2.2 BAM provides considerable catalytic enhancement at the transition temperature while BamA has only a minor effect

Table 4.2 shows a summary of the results of these folding experiments. Full details of the fitting parameters and individual t_{50} values can be found in Appendix D and Appendix E. In this section, only the data from the no SurA conditions will be considered with the effect of SurA addition considered later. For completeness, both data sets are presented. The average t_{50} values in the no SurA conditions for each temperature are compared in Figure 4.11. In agreement with the literature (Danoff and Fleming, 2015b), the folding rate of tOmpA is maximal at the transition temperature but considerably slower in the gel (20 °C) or liquid (30 °C) phases. The catalytic power of the BAM complex is immediately apparent, both over the uncatalysed rate and over the rate with BamA alone too. In the gel phase (20 °C), the full BAM complex shows a ~130X rate enhancement over the uncatalysed rate and a ~40X rate enhancement over uncatalysed folding in the liquid phase (30 °C) (Figure 4.11B). BamA was also able to enhance folding rates significantly, showing a ~3X enhancement in the gel phase, and 16X enhancement in the liquid phase (Figure 4.11C). Comparing the enhancement of folding rate at the transition temperature is particularly interesting, the full BAM complex provides a catalytic enhancement of 13X, but BamA provides only a modest enhancement of ~1.5X. This suggests that the catalytic power of BamA to accelerate the folding of tOmpA comes primarily from local lipid disordering, as when lipid disorder and lipid packing defects are maximised in the lipid background (at the transition temperature) the presence of BamA only marginally accelerates folding rate further. In contrast, the full BAM complex is able to provide an ~13X rate enhancement over uncatalysed folding — suggesting that the additional subunits of the BAM complex (BamBCDE) are either independently contributing to accelerating folding through a non-lipid-order related mechanism, enhancing the efficiency of the BamA ‘disorderase’ activity, or priming BamA to perform other mechanistic roles in assisting OMP folding (e.g. templating β -strands of an incoming substrate).

Considering the degree of catalytic enhancement of folding between BamA and the full BAM complex it is clear that the context of the lipid bilayer is important in defining the catalytic power of BAM/BamA needed to facilitate folding (Table 4.2 [$\Delta\Delta$ BAM/BamA] and Figure 4.11D). In all conditions, the full BAM complex is a much better catalyst than BamA (from 2.4X in liquid phase lipids, ~10X at the transition temperature, and 48X in the gel phase). This suggests that the lipid conformation of gel phase lipids somehow constrains the catalytic ability of BamA, but the additional lipoproteins of the BAM complex help to overcome this energetic barrier to folding.

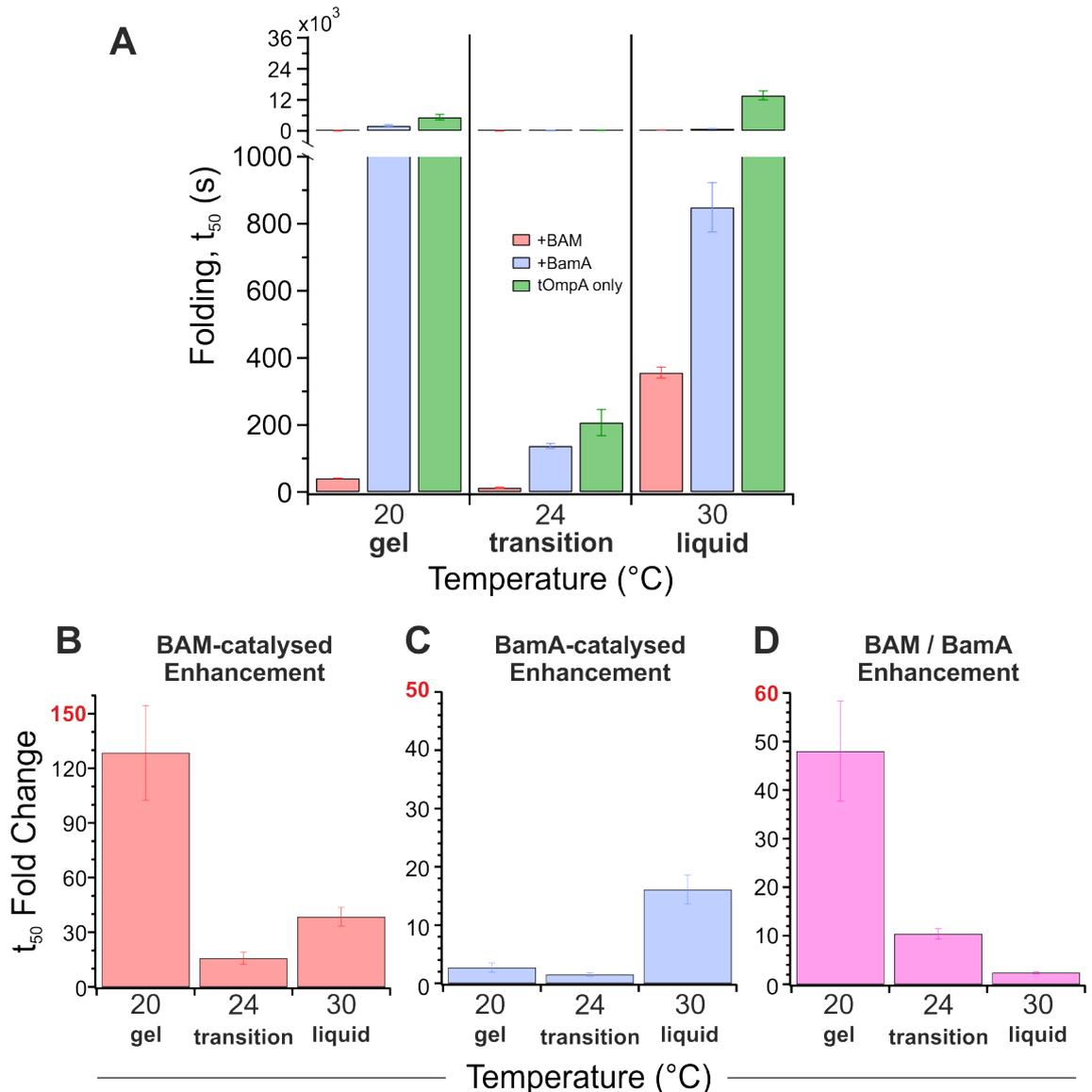


Fig. 4.11 Folding rate in the absence of SurA in the gel phase (20 °C), at the transition temperature (24 °C), and in the liquid phase (30 °C) and the catalytic affect of BamA and the full BAM complex. (A) Average t_{50} values for each condition compared. Error bars represent standard error of the mean from at least $n = 3$ repeats. (B)—(C) Fold enhancement of tOmpA folding rate caused by the presence of the full BAM complex (B) or BamA (C) over uncatalysed folding as measured by the average t_{50} . (D) Fold enhancement of tOmpA folding by the full BAM complex compared to BamA alone (average t_{50} of BAM-catalysed folding / average t_{50} of BamA-catalysed folding). Errors in (B-D) were propagated from (A). Final conditions: 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0, 1 μ M tOmpA, 2 μ M BAM or BamA or nothing (tOmpA only), 1600:1 (mol/mol) LPR DMPC proteoliposomes.

Table 4.2 Average t_{50} values for the folding of tOmpA into DMPC LUVs at different temperatures (different lipid phases), catalysed by the BAM complex, BamA, or uncatalysed. Average values are derived from at least $n = 3$ repeats for all conditions. Fold change values represent the acceleration of folding caused by the presence of the full BAM complex, or BamA, respectively. $\Delta\Delta$ BAM/BamA shows a measure of the relative increase in catalytic activity of the BAM complex compared to BamA alone.

Average t_{50} (s)						
No SurA	+BAM	+BamA	-BAM/BamA	Fold Change BAM	Fold Change BamA	$\Delta\Delta$ BAM/BamA
20 °C	40.9	1961.3	5258.9	128.6	2.7	48.0
24 °C	13.2	137.5	207.0	15.7	1.5	10.4
30 °C	356.1	849.5	13699.1	38.5	16.1	2.4
with SurA	+BAM	+BamA	-BAM/BamA	Fold Change BAM	Fold Change BamA	$\Delta\Delta$ BAM/BamA
20 °C	36.9	1694.9	4525.0	122.6	2.7	45.9
24 °C	12.3	135.3	346.4	28.2	2.6	11.0
30 °C	81.8	1268.3	36788.0	449.7	29.0	15.5

Comparing the folding rate between temperatures within the same conditions (e.g. only BAM-containing experiments, or only uncatalysed experiments etc.) it is easier to see trends caused by the lipid phase by setting the 24 °C “max” rate as the reference state (Table 4.3 and Figure 4.12). Here we see that BAM is able to maintain a high catalytic enhancement even at 20 °C in the gel phase with an average t_{50} only 3.1X slower than at 24 °C, while folding is slowed to a much greater degree (27X slower) in the liquid phase at 30 °C. For BamA, this trend is reversed with the folding rate at 20 °C being ~14X slower in the gel phase, but only ~6X slower in the liquid phase. Uncatalysed folding is considerably slower in either the gel or liquid phases as compared to the transition state, but surprisingly, it is faster at 20 °C in the gel phase than at 30 °C in the liquid phase.

Table 4.3 Fold change in average t_{50} values *within* each experimental condition relative to the 24 °C transition state. Positive values indicate slower folding compared to 24 °C data.

Average t_{50} fold Change versus 24 °C State			
No SurA	+BAM	+BamA	-BAM/BamA
20 °C	-3.1	-14.3	-25.4
24 °C	1.0	1.0	1.0
30 °C	-27.0	-6.2	-66.2
with SurA	+BAM	+BamA	-BAM/BamA
20 °C	-3.0	-12.5	-13.1
24 °C	1.0	1.0	1.0
30 °C	-6.7	-9.4	-106.2

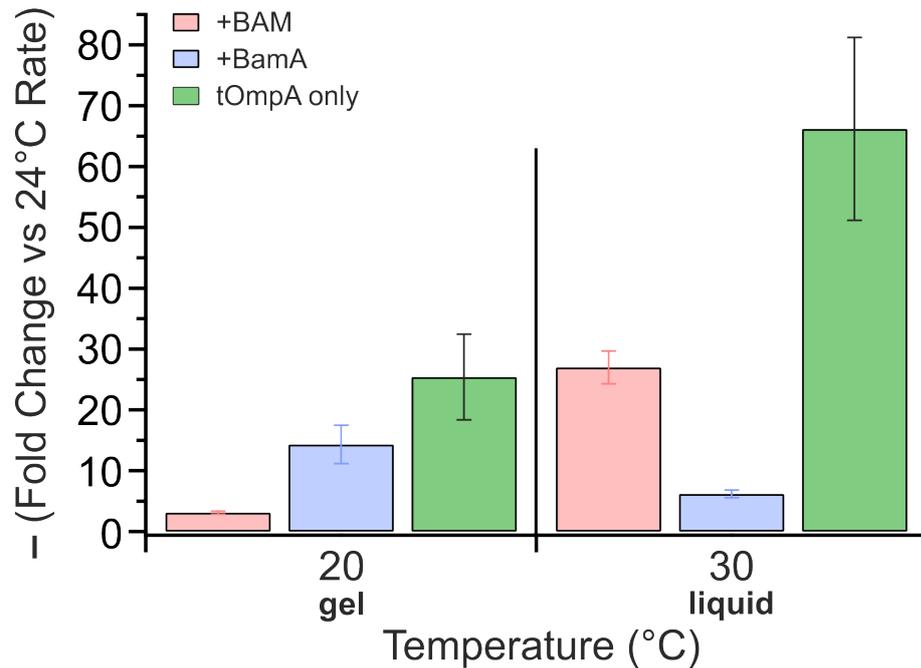


Fig. 4.12 Fold change in folding rate measured by average t50 in different lipids phases, relative to the transition temperature, for BAM-catalysed, BamA-catalysed, and uncatalysed folding. Assessing the relative difficulty for tOmpA to fold into different lipid phases under the same catalytic conditions.

In vivo, not only is the rate of folding critically important for cell survival (as rates must be compatible with the doubling time of a species, e.g. 20 min for *E. coli*), but there is also a large energetic cost to producing OMPs (which make up a very large fraction of the total mass of protein in *E. coli*) and so maximising the yield of correctly folded protein is likely to also be an important factor in survival. Approximate folding yields were extracted from the y_0 intercept value from double exponential fits to the data (Table 4.4 and Figure 4.13). The full BAM complex was able to improve absolute folding yields by around ~5 % at the transition temperature, ~7 % in the gel phase (20 °C), and ~8 % in the liquid phase (30 °C) (Figure 4.13B). BamA enhanced folding yields in the liquid phase (30 °C) by ~7 % but reduced folding yields at the transition temperature (a fall of ~12 %) and in the gel phase (a fall of ~21 %) (Figure 4.13C). Comparing the relative folding yield improvement for BamA over BAM (Table 4.4, Table 4.5, and Figure 4.13D), we see that the yields are roughly equal for BAM, regardless of the state of the lipids or temperature (~87—89 %), suggesting that for folding via BAM the yield is independent of the state of the lipid bilayer. Similarly, for uncatalysed folding, yields were between ~80—82 % (Figure 4.13A) suggesting that the yield of tOmpA folding is not directly correlated to the lipid phase or the rate of folding (as the folding rate of uncatalysed tOmpA folding spans 3 orders of magnitude in different lipid phases — Table 4.2). However, for BamA the yield is ~28 % lower than BAM in the gel phase (20 °C), ~17 % lower at the transition

Table 4.5 Change in average fraction folded (yield) values *within* each experimental condition relative to the 24 °C transition state. Positive values indicate greater yields compared to 24 °C data, negative values indicates falls in the yield.

Δ Yield (Absolute % Folded) versus 24 °C State (%)			
No SurA	+BAM	+BamA	-BAM/BamA
20 °C	1.1	-14.7	-1.0
24 °C	-	-	-
30 °C	0.5	23.9	-2.3
with SurA	+BAM	+BamA	-BAM/BamA
20 °C	6.1	0.9	0.9
24 °C	-	-	-
30 °C	6.2	29.0	-68.9

temperature, and similar in the gel phase (30 °C). It is also lower than the uncatalysed yield at the transition temperature and in the gel phase despite the *rate* being faster, suggesting that the mechanism of BamA-accelerated folding comes at a cost of folding yield.

Table 4.4 Average yield values (% folded) derived from y_0 of fits to data for the folding of tOmpA into DMPC LUVs at different temperatures (different lipid phases), catalysed by the BAM complex, BamA, or uncatalysed. Average values are derived from at least $n = 3$ repeats for all conditions. Yield change values represent the increase in fraction folded caused by the presence of the full BAM complex, or BamA, respectively. $\Delta\Delta$ BAM/BamA shows a measure of the relative increase in tOmpA folded yield in the presence of the BAM complex compared to BamA alone. Green boxes show improvements in yield, blue boxes show falls in yield.

Average Yield (% Folded)						
No SurA	+BAM	+BamA	-BAM/BamA	Yield Change BAM (%)	Yield Change BamA (%)	$\Delta\Delta$ BAM/BamA (%)
20 °C	88.4	60.2	81.3	7.1	-21.1	28.2
24 °C	87.4	70.5	82.1	5.3	-11.5	16.9
30 °C	87.9	87.4	80.2	7.6	7.2	0.4
with SurA	+BAM	+BamA	-BAM/BamA	Yield Change BAM (%)	Yield Change BamA (%)	$\Delta\Delta$ BAM/BamA (%)
20 °C	93.6	63.7	82.9	10.7	-19.1	29.8
24 °C	88.2	71.4	87.6	0.6	-16.2	16.8
30 °C	93.6	92.1	27.2	66.4	64.8	1.6

4.2.3 SurA enhances BAM activity and slows uncatalysed folding rate at 30 °C liquid phase, has minor or negative effect at other phases and with BamA

For each experimental condition, for tOmpA folding in BAM- and BamA-proteoliposomes, or empty liposomes, a separate experiment under the same conditions, but with the addition

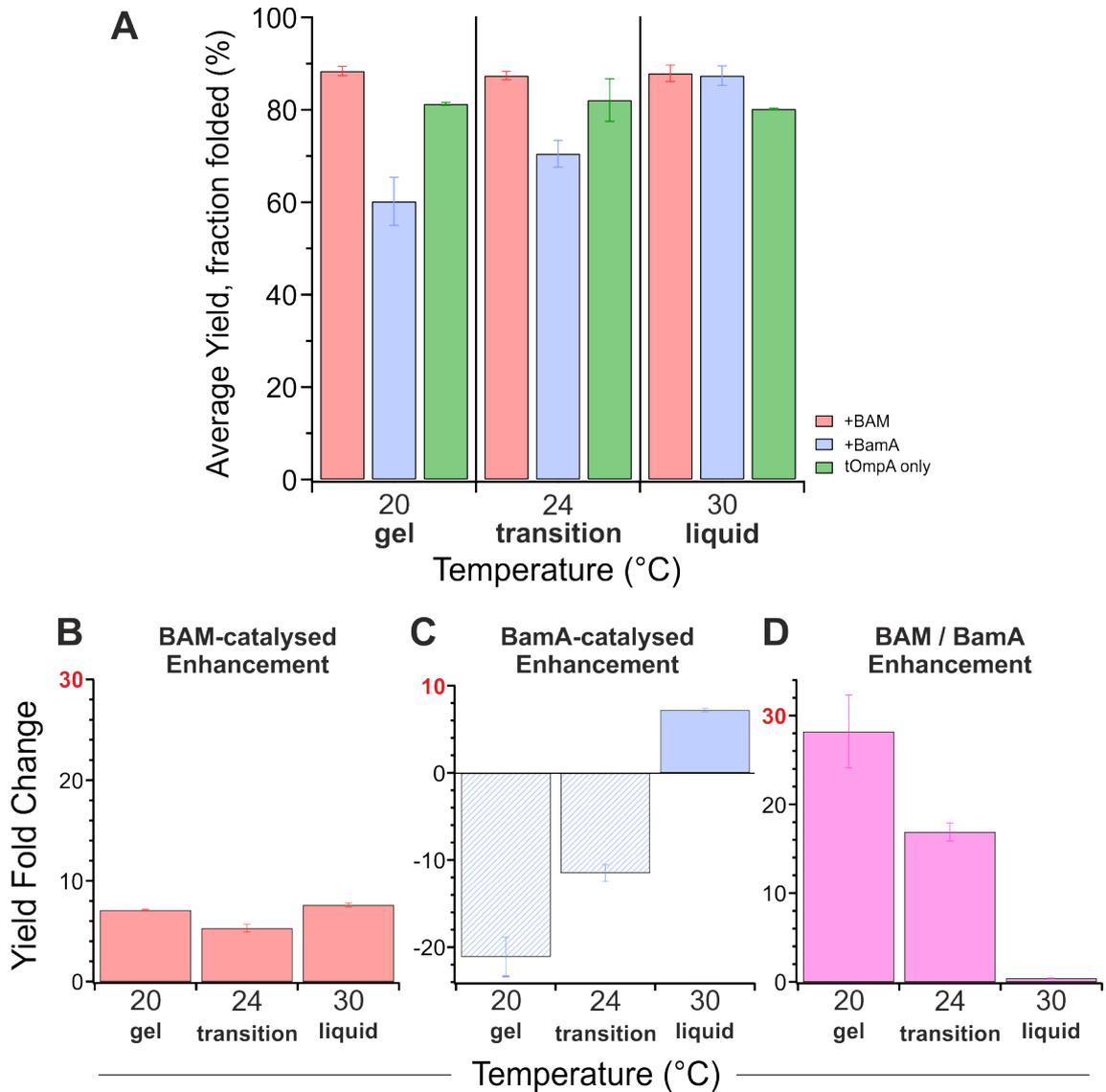


Fig. 4.13 Folding yield in the absence of SurA in the gel phase (20 °C), at the transition temperature (24 °C), and in the liquid phase (30 °C) and the effect of BamA and the full BAM complex in enhancing the yield of folding tOmpA as measured by band-shift. Yields were derived from the y_0 intercept of a double-exponential fit to the data. (A) Average folding yield for each condition compared. Error bars represent standard error of the mean from at least $n = 3$ repeats. (B)—(C) Fold change in folding yield of tOmpA caused by the presence of the full BAM complex (B) or BamA (C) versus uncatalysed folding as measured by the average yield. Negative values represent a poorer folding yield in the presence of BamA. (D) Fold yield enhancement of tOmpA in the presence full BAM complex compared to BamA alone (average yield of BAM-catalysed folding / average yield of BamA-catalysed folding). Final conditions: 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0, 1 μ M tOmpA, 2 μ M BAM or BamA or nothing (tOmpA only), 1600:1 (mol/mol) LPR DMPC proteoliposomes.

of SurA was conducted in parallel. This allows the differential effect on t_{50} and folding yield caused by SurA to be dissected. As shown in Table 4.6 and Figure 4.14, under most conditions SurA had a negligible effect on the folding rate as measured by t_{50} . Nonetheless, notable exceptions to this trend are observed. In the presence of BAM in the liquid phase (30 °C) there is 4.4X enhancement in average t_{50} , from an average t_{50} of $\sim 360 \pm 16$ s to $\sim 82 \pm$ s. It is unlikely that this occurs by SurA improving solubility of tOmpA (which might direct more protein down a fast-folding kinetic pathway) as comparing the t_{50} of no SurA BAM in the liquid phase (30 °C) against no SurA uncatalysed t_{50} at the transition temperature (24 °C) shows a similar t_{50} (207 ± 39 s uncatalysed vs 356 ± 16 s with BAM) (Table 4.2), but in this case (uncatalysed folding at the transition temperature, 24 °C) the t_{50} actually falls by 1.7X in the presence of SurA. The large positive change in t_{50} under liquid phase conditions (30 °C) occurs only with the full BAM complex and could suggest two possibilities: (1) that SurA can ‘activate’ BAM only when it is in a certain conformation and that this conformation is suppressed at the transition temperature (24 °C) or in the gel phase (20 °C) or (2), that the rate-accelerating mechanism of SurA is only manifested when OMPs fold into a liquid phase membrane. The other situations where SurA has a more dramatic effect on folding yield (causing a decrease of 1.5—2.7X — shown in blue in Table 4.6) may be due to the balance between SurA’s overlapping roles as a ‘holdase’ and fold-promoting chaperone tipping towards ‘holding’ on to tOmpA and preventing it from accessing a faster folding pathway. In other words, in one case, SurA would slow folding by holding on to the polypeptide chain and preventing it from being accessible for membrane recognition and insertion. In the other case, SurA would accelerate folding by keeping OMPs in a folding competent state and preventing off-pathway intermediates or aggregates being formed. The interplay between these forces may be complicated as evidenced by the varied response in t_{50} to the presence of SurA shown here.

Table 4.6 Fold change in average t_{50} values caused by the presence of SurA in each experimental condition. Positive values (green) indicate conditions where SurA causes an acceleration of folding rate while negative values (blue) indicate SurA slows folding.

Average t_{50} Fold Change			
	BAM \pm SurA	BamA \pm SurA	-BAM/BamA \pm SurA
20 °C	1.1	1.2	1.2
24 °C	1.1	1.0	-1.7
30 °C	4.4	-1.5	-2.7

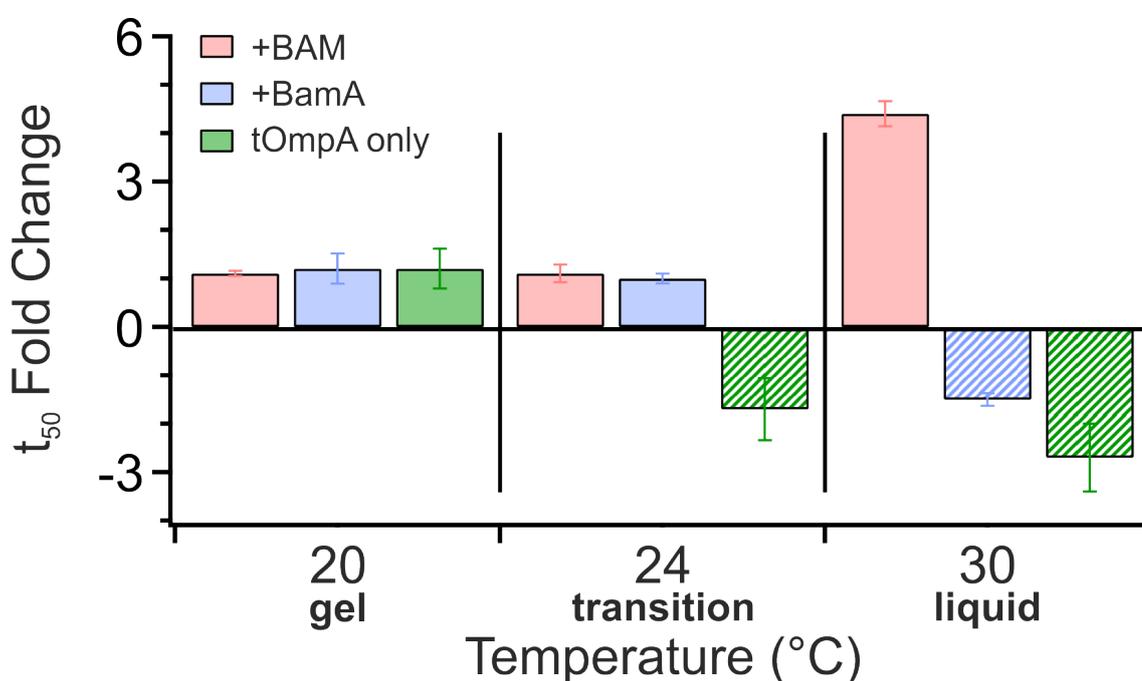


Fig. 4.14 Change in the folding rate of tOmpA as measured by t_{50} caused by the presence of SurA. Negative values represent a decrease in the folding rate in the presence of SurA. Final conditions: 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0, 1 μ M tOmpA, 2 μ M BAM or BamA or nothing (tOmpA only), 10 μ M SurA (when present), 1600:1 (mol/mol) LPR DMPC proteoliposomes.

Although the additional effect on folding rate in the presence of SurA is fairly minor in comparison to the enhancement by the BAM complex or BamA alone, the role of SurA may be to keep OMPs soluble and thus improve folding yields. To assess this idea, the change in folding yields caused by the presence of SurA was assessed (see Table 4.7, Table 4.8, and Figure 4.15 for absolute changes in fraction folded, and relative (%) changes in fraction folded, respectively). Similar to the t_{50} values, relatively small changes in folding yield (~1-6 %) were observed except for uncatalysed folding in the liquid phase (30 °C) where there was a very large decrease in folding yield (-53 %). Combined with the increase in t_{50} (from 13699 s / 228 mins to 36788 s / 613 mins), this suggests that

under certain conditions the holdase activity of SurA can be detrimental to both folding yield and rate. That this decrease in folding yield is not recapitulated in the presence of BAM or BamA suggests that these proteins may help to enable efficient folding, and SurA may hold OMPs in a state which is *not* generically ‘folding competent’, but which is more easily passed to BamA or the full BAM complex for folding.

Table 4.7 Change in final folding yield caused by the presence of SurA for each experimental condition. Yields were derived from the y_0 intercept value from double exponential fits to the data. Positive values (green) indicate conditions where SurA increased the folding yield, negative values (blue) where SurA decreased folding yield.

Δ Yield (Absolute % Folded)			
	BAM \pm SurA	BamA \pm SurA	-BAM/BamA \pm SurA
20 °C	5.2	3.5	1.6
24 °C	0.8	0.9	5.5
30 °C	5.8	4.7	-53.0

Table 4.8 Change in final folding yield caused by the presence of SurA as a percent of the original value for each experimental condition. Note that this is not the absolute change in folding yield (as % folded) but rather a normalized measure of yield change. Yields were derived from the y_0 intercept value from double exponential fits to the data. Positive values (green) indicate conditions where SurA increased the folding yield, negative values (blue) where SurA decreased folding yield.

Average Yield Percent Change (%)			
	BAM \pm SurA	BamA \pm SurA	-BAM/BamA \pm SurA
20 °C	5.9	5.9	2.0
24 °C	0.9	1.2	6.7
30 °C	6.6	5.3	-66.0

4.2.4 Effects of BamA and the BAM complex on lipid order and phase transitions

The dramatic catalytic effects of the BAM complex on folding rate of tOmpA into DMPC, particularly in the gel phase (20 °C) lipids, prompted the question as to whether the BAM complex has a measurable effect on lipid order or the phase of lipid bilayers. Secondly, is the effect comparable to BamA (suggesting the catalytic enhancement of the BAM complex is due to factors other than lipid disordering as the same level of disordering would be insufficient to explain the rate enhancement observed previously) or is it distinct

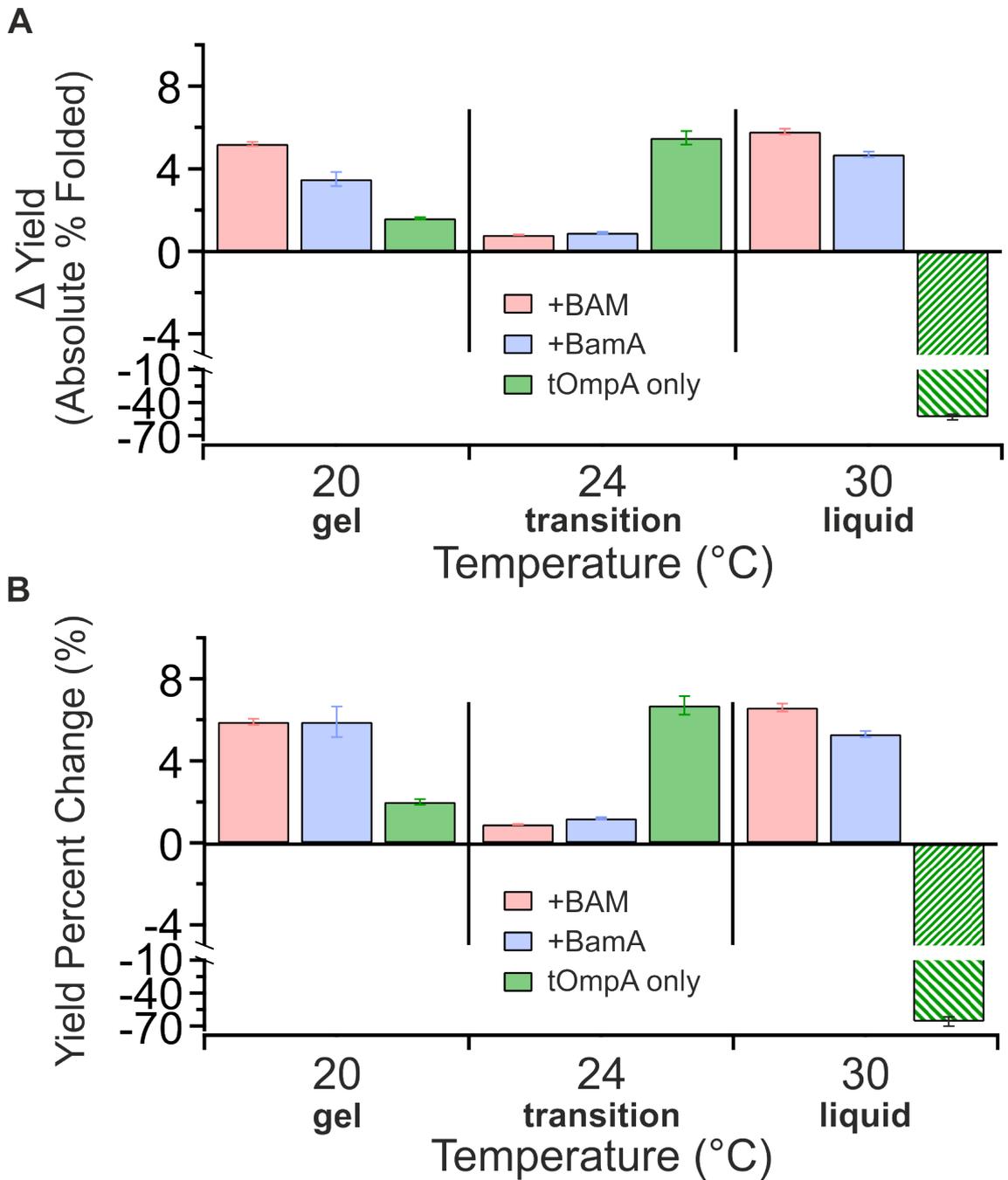


Fig. 4.15 Change in folding yields of tOmpA caused by the presence of SurA. Negative values represent a decrease in the folding yield in the presence of SurA. The y-axis is split between -5 to -10 %, and rescaled between -10 and -70 % to allow the large decreases in the folding yield of uncatalysed tOmpA folding in the presence of SurA to be shown while still allowing the relationship between other conditions to be clear. Final conditions: 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0, 1 μ M tOmpA, 2 μ M BAM or BamA or nothing (tOmpA only), 10 μ M SurA (when present), 1600:1 (mol/mol) LPR DMPC proteoliposomes.

from BamA (implying the lipoproteins of the BAM complex enhance the ‘disorderase’ activity)?

Laurdan is a fluorescent probe comprising a 12-carbon acyl ‘tail’ and a naphthalene-based fluorescent headgroup (see Figure 4.17A). Through its hydrophobic tail, it can partition into lipid bilayers, with its headgroup sitting in the interfacial region (Parasassi et al., 1991; Jurkiewicz et al., 2012). The fluorescence of laurdan is sensitive to its local environment, specifically, to the local dielectric character (Vequi-Suplicy et al., 2014, 2015). This location of laurdan in the interfacial region between the headgroups and acyl tails of lipid bilayers, and the sensitivity of its emission profile to the dielectric environment, makes laurdan a good probe of the level of hydration of a lipid bilayer due to penetration of water molecules from the bulk solvent (Parasassi et al., 1991). It has been observed that water penetrates deeper into the bilayer when lipids are in the liquid phase versus the gel phase (Tu et al., 1996; Stepniewski et al., 2010). This effect is intimately linked to the packing of lipid acyl chains (and by proxy their level of order) in the bilayer, where more densely packed lipids do not expose any ‘openings’ between headgroups in which a molecule of water could enter, but disordered lipid tails can create transient voids in the bilayer due to packing defects caused by their conformational heterogeneity (Disalvo et al., 2015; Pasenkiewicz-Gierula et al., 2016). In order to quantitate the changes in emission spectra and thus in the environment of the laurdan probe, a measure called the generalized polarization (GP) was derived (Parasassi et al., 1991). This is essentially a ratio of two regions of the fluorescence spectrum of laurdan, one of which is most sensitive to the gel phase and the other to the liquid phase and its form is given in Eq. 3. Measurements of GP values have been used to infer the phase of lipid bilayers and the presence of lipid rafts both *in vivo* and *in vitro*, and have been used to calculate the transition temperature of pure and complex lipid mixtures (Harris et al., 2002; Vanounou et al., 2002; Velázquez and Fernández, 2006; Sanchez et al., 2012). Here, laurdan was used to assess changes in the lipid bilayer packing caused by the presence of BamA or the BAM complex and provided a simplified method for measuring transition temperatures using the first derivative of the change in GP with temperature.

Equation 3:

$$GP = \frac{I_{440} - I_{490}}{I_{440} + I_{490}}$$

Laurdan was dissolved in DMSO to 1 mg ml⁻¹, and then doped into BAM- or BamA-proteoliposomes, or empty DMPC liposomes to a final concentration of 10 μ M giving a lipid-to-probe ratio of 300:1 (mol/mol), a BAM-to-probe ratio of 5:1 (mol/mol), and a residual DMSO concentration of 0.01 % (v/v). This sample was then incubated overnight

for ~16 hours at 25 °C to allow full partitioning of the probe into the lipid bilayer (see Figure 4.16). This DMSO loading method has been shown to give identical results to LUVs made with laurdan present in the initial chloroform/methanol mixture used to form the liposomes and is the method used for live-cell labelling (Mazeres et al., 2014). Samples were then incubated in a fluorescence spectrometer at 6 °C for 1 hour before full emission spectra from 400—600 nm were recorded after excitation at 340 nm. The temperature was then ramped in 1—2 °C intervals up to up to 56 °C and back down again with a 5 minute equilibration time at each temperature. Example emission spectra for DMPC only liposomes are given in Figure 4.17B. From these spectra, GP values could be calculated according to Eq. 3 and the results are shown in Figure 4.17C. The data show that the transitions are fully reversible and exhibit no hysteresis. The laurdan GP value can be used to suggest the presence of the gel or liquid phase in a single sample at a fixed temperature, with values over 0.5 or under 0 suggesting the presence of pure gel or liquid phases, respectively (Fidorra et al., 2006; Sanchez et al., 2007). By calculating GP as the temperature is altered, a more detailed analysis of the lipid bilayer properties can be undertaken. Taking the first derivative of these data revealed the transition temperature of the lipids in the sample, similar to previous approaches to measure known or unknown T_m values (Vanounou et al., 2002; Velázquez and Fernández, 2006).

Comparing the data in each of these three liposome systems shows that BamA proteoliposomes and empty DMPC liposomes have an identical transition temperature (24 °C) with a very similar degree of co-operativity and change in enthalpy (taking the magnitude of the Δ Slope and the width of the transition as a measure of these parameters, similar to a thermogram measured by DSC). Remarkably, for BAM-proteoliposomes a 2 °C decrease in the transition temperature was observed, as well as a broadening of the transition region and a decrease in the magnitude of change of the GP slope, indicating a less cooperative transition in the presence of the full BAM complex. To obtain a more quantitative measure of the width or broadness of the transition, the second derivative of the GP data was taken (Figure 4.17E). Due to the coarse-grained nature of the data (taken at 1 °C intervals) the onset and termination of the transition region was calculated as the point at which the second derivative came within 1 standard deviation (calculated from the absolute gel and liquid regions of the spectra) of the zero-line. For BamA and DMPC this was from 20—28 °C (8 °C width), and for BAM this was from 16—26 °C (10 °C width). These results were recapitulated in different liposome batches and made on different days (Figure 4.18).

BamA is capable of being refolded from a urea-denatured state into detergent micelles and DMPC liposomes and this fact was exploited to measure changes in GP values of DMPC bilayers caused by the gradual insertion of BamA to assess whether BamA caused

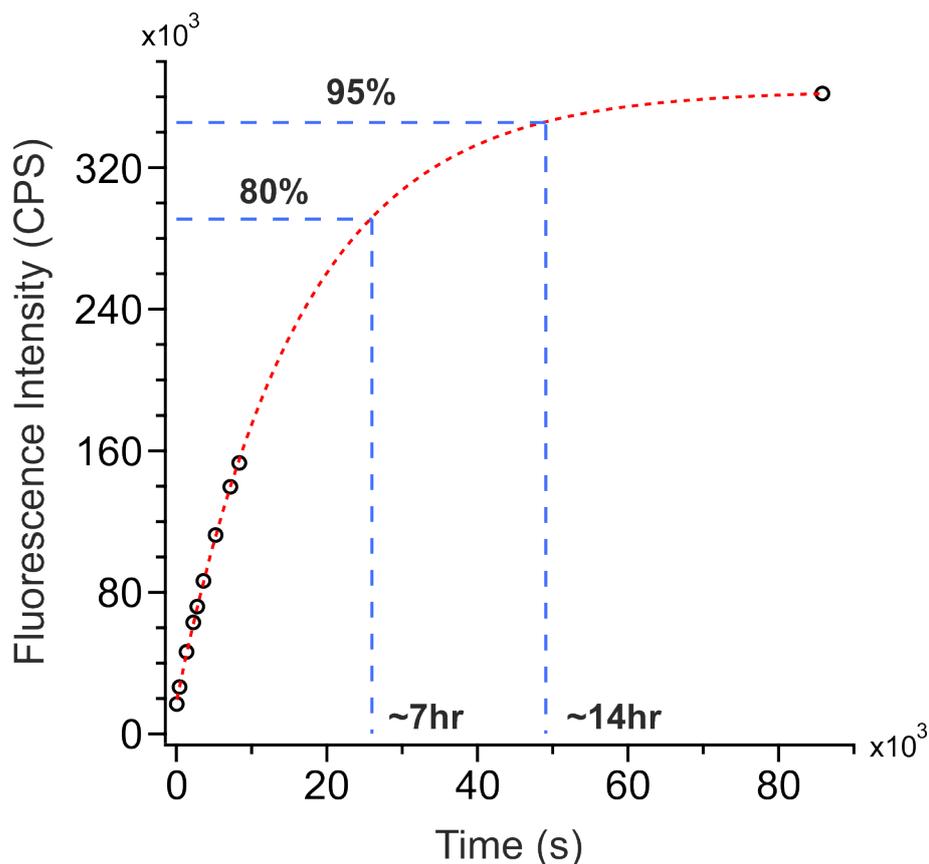


Fig. 4.16 Change in laurdan fluorescence intensity at 440nm after doping into BAM DMPC proteoliposomes incubated at 25 °C. Laurdan was diluted out of a $1 \text{ mg}\cdot\text{ml}^{-1}$ DMSO stock into the presence of empty, BamA-, or BAM-containing DMPC liposomes (shown here, BAM proteoliposomes). Excitation at 340 nm, emission collected between 420–520 nm. Laurdan is poorly soluble in water and the observed kinetics likely result from a slow partitioning of laurdan from DMSO-associated aggregates into the lipid bilayer (Mazeret et al., 2014). Full spectral scans were taken periodically over a 24-hour time period and the data fitted to a single exponential equation (red dashed line). The time to reach 80 % and 95 % incorporation were calculated from this fit (blue dashed lines) and the respective times are indicated in hours above the x-axis. Final conditions: 150 mM NaCl, 20 mM Tris-Cl pH 8.0.

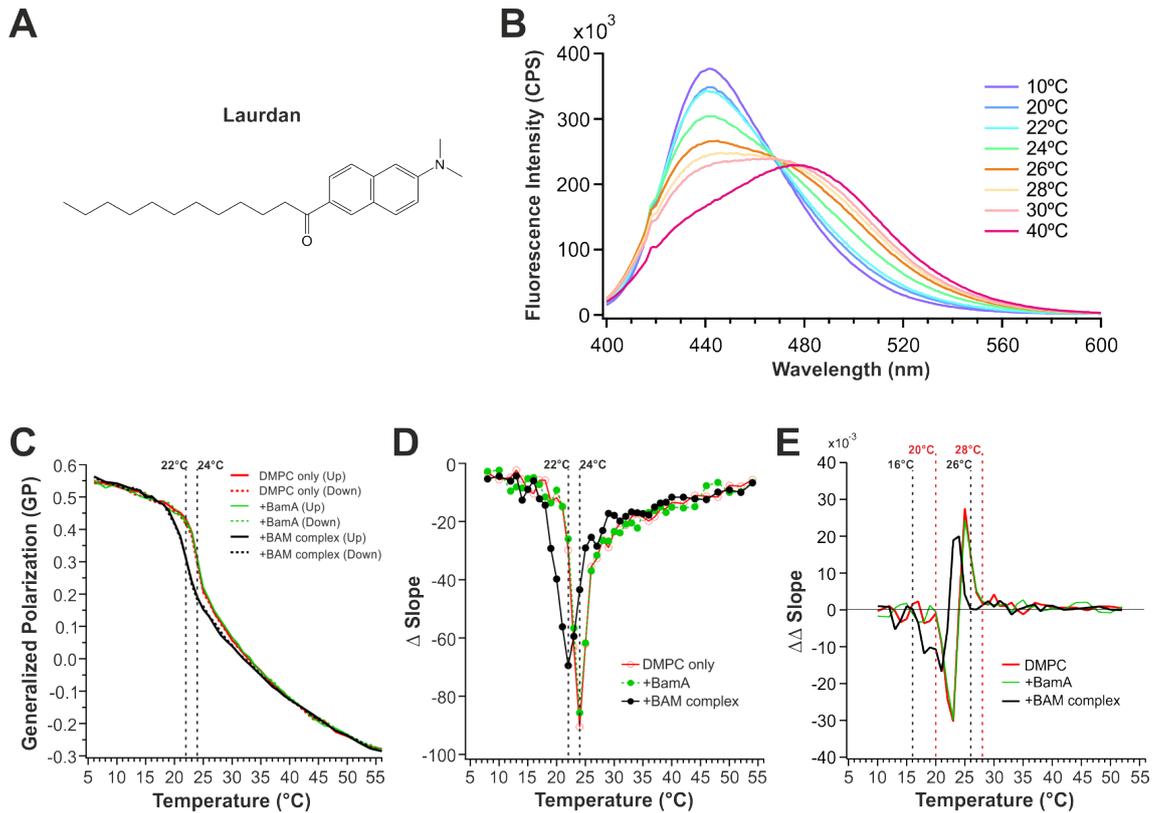


Fig. 4.17 Laurdan can probe changes in lipid order and lipid packing in pure DMPC LUVs, as well as BamA- and BAM-proteoliposomes. Laurdan was doped into liposomes or proteoliposomes to a final lipid-to-probe concentration of 300:1 (mol/mol) and changes in bilayer packing assessed by measuring the GP value. (A) Structure of a laurdan molecule. (B) Example emission spectra after exciting at 340 nm showing the change in emission properties as a pure DMPC bilayer transitions from the gel to liquid phase. (C) GP values for empty, BamA-, and BAM-containing liposomes as the temperature was increased from 6–56 °C (solid lines) and relaxed back down again (56–6 °C) (dashed lines, n.b. due to the close overlay these are mostly occluded by the up ramp). (D) First derivative of the data shown in (C) showing the change in lipid order and transition temperature in BamA and empty liposomes versus BAM proteoliposomes. The transition region broadens and the peak shifts downwards to 22 °C in the presence of the full BAM complex. (E) Second derivative of the data shown in (C) more clearly delineating the start and end of the transition region observed in (B).

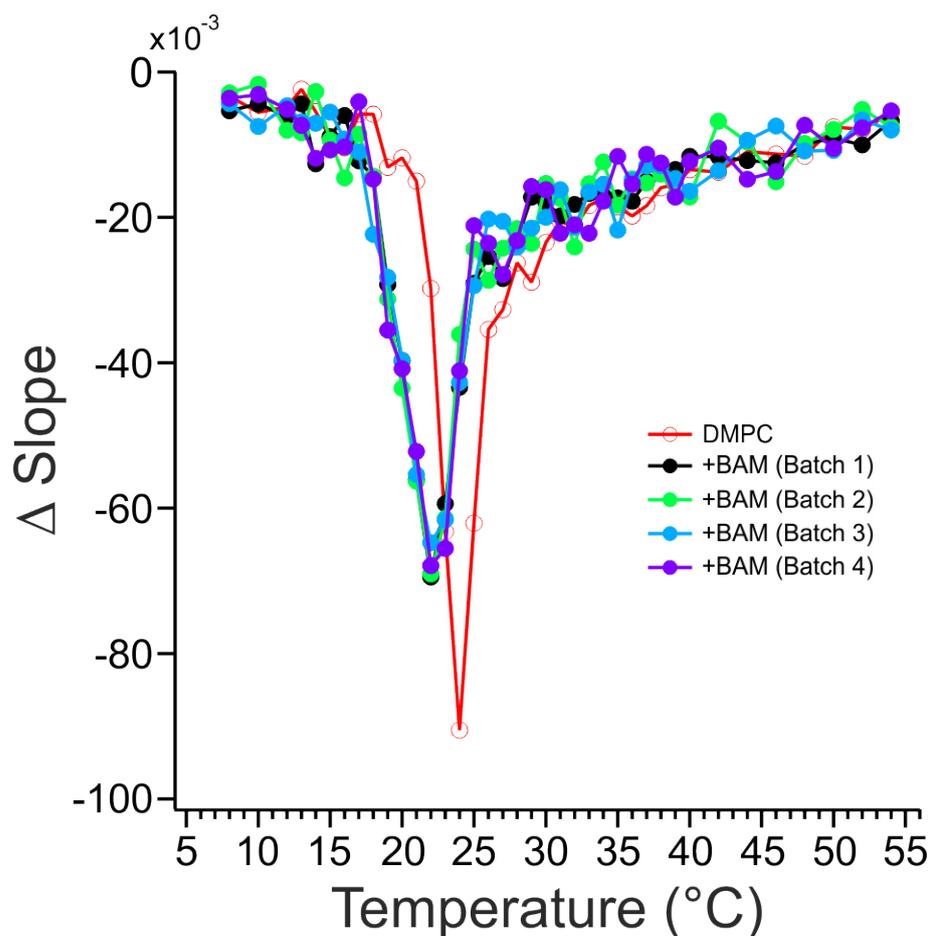


Fig. 4.18 First derivative of the change of laurdan GP with respect to temperature for four different batches of BAM DMPC 1600:1 (mol/mol) LPR LUV proteoliposomes. The same experiments as shown in Figure 4.17 were repeated for different batches of BAM DMPC proteoliposomes used for the kinetics experiments reported in Section 4.2.2. DMPC and BAM (Batch 1) are the same as shown in Figure 4.17D.

even minor changes in the lipid order of DMPC bilayers. The ratio of the fluorescence intensity at 440nm and 490nm was measured directly at high temporal resolution immediately after diluting BamA out of a urea-denatured state (8 M urea, 50 mM glycine pH 9.5) in the presence of laurdan-doped DMPC liposomes at 25 °C. A higher pH (9.5) was used in these experiments as this has been shown to eliminate aggregation of OMPs over the time course of these experiments (Schiffirin et al., 2017b). Figure 4.19 shows the result of this experiment. Black and grey traces show the static values of GP measured before addition of BamA (~0.203), and a few hours after completion of folding (~0.214). Surprisingly, instead of decreasing the GP value, as would be expected for a transformation to a more fluid membrane, the GP value actually increased from its original value on completion of the folding reaction although the magnitude of this change was very small ($\Delta 0.011$). Immediately after initiation of folding, the GP values show an initial decrease from ~0.203 to ~0.192. There are two ways to interpret this data: that the GP values reflect true changes in bilayer fluidity, or that they reflect changes in hydration of laurdan independent of changes in lipid order due to the presence of protein at or near the interfacial region. The first explanation agrees with the generic properties of OMPs which are believed to rigidify the outer membrane due to their low deformability and implying that when lipid order is measured globally (as is the case for the laurdan probe) BamA may influence the membrane like any other OMP, but this does not rule out a local disordering effect (e.g. imagine a case where the lipids radiating outwards from the $\beta 1$ - $\beta 16$ seam are much more disordered, but those around the sides and back of the barrel rigidified). Alternatively, if laurdan clustered near BamA, the protein may exclude more water than the interfacial region of the bilayer which would be a probe-dependent, rather than lipid-dependent, effect. The initial decrease in GP could be due to the early insertion steps of BamA- (and generic OMP-) folding: ‘opening up’ the membrane and allowing greater penetration of water molecules. An initial membrane-association step might be expected to cause an *increase* in rigidity due to the carpeting of the polypeptide chain restricting the diffusion of lipid molecules as is observed with antimicrobial peptides (Smith-Dupont et al., 2010) and also excluding water from the surface. The dead-time of this experiment is ~10 seconds so this step would not be captured if it rapidly proceeded to the insertion step. It is interesting to note that the data fitted well to a double exponential equation with the amplitude of the first phase corresponding approximately to the change from the start value to the *pre-folding* GP value, and the slower second phase to the transition between the pre-folding value to the final value. This suggests that the physical origin of the double exponential kinetics is a fast initial binding and insertion phase occurring near the interfacial region of the membrane, and a second equilibration phase occurring deeper in the membrane.

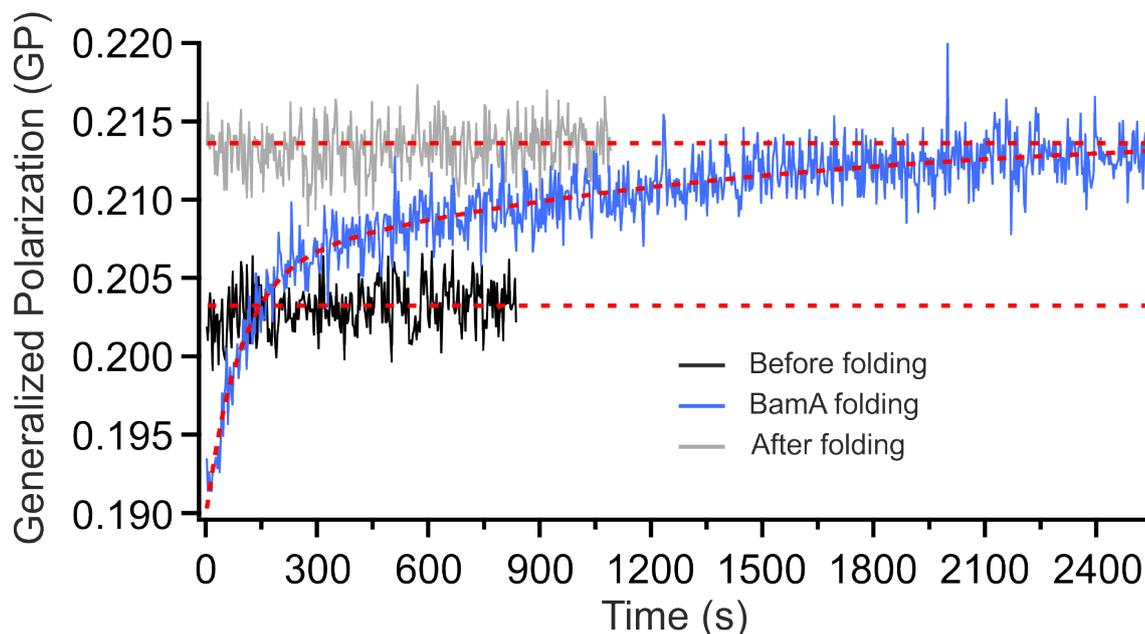


Fig. 4.19 The folding of BamA into DMPC LUVs causes changes in the laurdan GP value. BamA was diluted from high (8 M) to low (0.32 M) urea in the presence of 100 nm DMPC LUVs pre-loaded with laurdan at 300:1 (mol/mol) lipid-to-probe ratio. Final conditions: 0.8 μ M BamA, 0.32 M urea, 50 mM glycine pH 9.5, 25 $^{\circ}$ C. The GP of laurdan was measured before (black trace), after (grey trace), and immediately after adding BamA (blue trace). Red dashed lines are added to guide the eye. For BamA folding, the kinetic transient is fitted to a double exponential equation (red dashed line).

4.2.5 The conformational ensemble of the BAM complex is not significantly altered between lipid phases

The kinetic results and data from the laurdan lipid order experiments suggested that BAM is able to remodel lipid membranes in a manner that is distinct from BamA. Structural studies of the full BAM complex have observed two distinct states — open kinked and closed kinked (see Figure 4.2) — with *in silico* studies suggesting that a continuum between these states may be accessible (though not yet captured structurally) (Noinaj et al., 2013, 2014; Fleming et al., 2016; Gu et al., 2016). These two states (open and closed) are correlated with large movements of the POTRA domains relative to the barrel due to a movement of POTRA5 that is propagated down POTRA4-1. To exploit this large change in the orientation of the POTRA domains relative to the barrel domains, single-molecule FRET (smFRET) experiments were carried out. smFRET allows distance changes between two dyes to be monitored at a per molecule level and with ms temporal resolution, allowing conformational changes to be measured through both space and time. To determine whether BAM alters its conformational equilibrium in response to changes in the lipid environment a FRET pair that is sensitive to changes between the open and closed state was designed.

R127 in POTRA2 and N520 in turn 3 of BamA were chosen as locations to install dyes due to their C_{α} - C_{α} distances of 71.1 Å in the open, and 61.7 Å in the closed state (an ~1 nm shift), location away from putative OMP assembly locations, poor conservation across proteobacteria, and lack of literature evidence suggesting they are in key locations for function. A double cysteine variant of the BAM complex (R127C/N520C) was expressed, purified, and stochastically labelled with Alexa Fluor 488 C5 maleimide (donor dye), and DyLight 594 maleimide (acceptor dye) (Figure 4.20). DyLight 594 was chosen over Alexa Fluor 594 (a more common single-molecule FRET pair) due to its higher hydrophilicity, thereby reducing the possibility of membrane partitioning by the polyaromatic xanthene when conjugated at the N520C position. Absorbance spectra suggested an approximately 50/50 labelling yield of BAM(R127C/N520C) with each dye. An accurate dye-to-dye distance was calculated using the FRET Positioning and Screening tool (FPS) developed by the group of Claus Seidel (Kalinin et al., 2012) (Figure 4.21A). This method uses the concept of available volumes and performs a simulation of each dye (modelled as an ellipsoid with a flexible linker attached to the sulfur of the cysteine side-chain) around its attachment point (treating the protein as a rigid body), generating a ‘cloud’ of accessible volume for each dye. Although the structure of DyLight 594 is proprietary, it is a highly sulfonated xanthene derivative with an excitation and emission spectrum similar to Alexa Fluor 594 (also a sulfonated xanthene) and so Alexa Fluor 594 was used as the structural model. Accessible volumes were calculated on BAM(R127C/N520C) in the open (PDB: 5LJO) (Iadanza et al., 2016) and closed (PDB: 5D0O) (Gu et al., 2016) states embedded in a DMPC bilayer. Mean dye-to-dye distances were similar for both possible donor/acceptor labelling locations giving 72.5 Å in the closed state and 83.1 Å in the open state. The E, or efficiency value, is a measure of how efficient the process of FRET is with a value of 1.0 (100 %) meaning perfect FRET (all donor energy is transferred to the acceptor dye) and 0 (0 %) meaning no energy is transferred from donor to acceptor. Efficiency values were calculated assuming an R_0 (the distance at which the FRET efficiency is 50 %) of 60 Å based on the near identical absorption spectra of Alexa Fluor 594 and DyLight 594 and the known R_0 of the Alexa Fluor 488 / Alexa Fluor 594 pair (60 Å). This gives an E value of 0.28 in the closed state, and 0.15 in the open state. Derivation of this E value is based on the assumption of free rotation of each dye, however the dye at position N520C may be restrained and so the true E value may vary from this calculated figure.

For smFRET experiments, DDM-solubilised labelled BAM was diluted by 1:333 into DMPC-LUVs to a final LPR of ~82,000:1 (mol/mol), incubated at ~50 °C for 30 minutes in the presence of Bio-Beads, and extruded through a 100 nm pore. This dilution method for preparation of BAM into LUVs was used in place of the dialysis method discussed earlier as it is rapid (whereas dialysis takes 3–5 days) and allows samples to be prepared

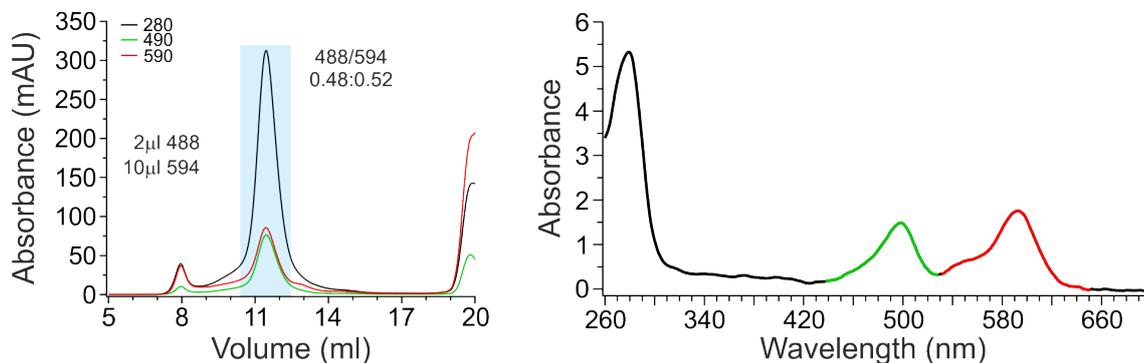
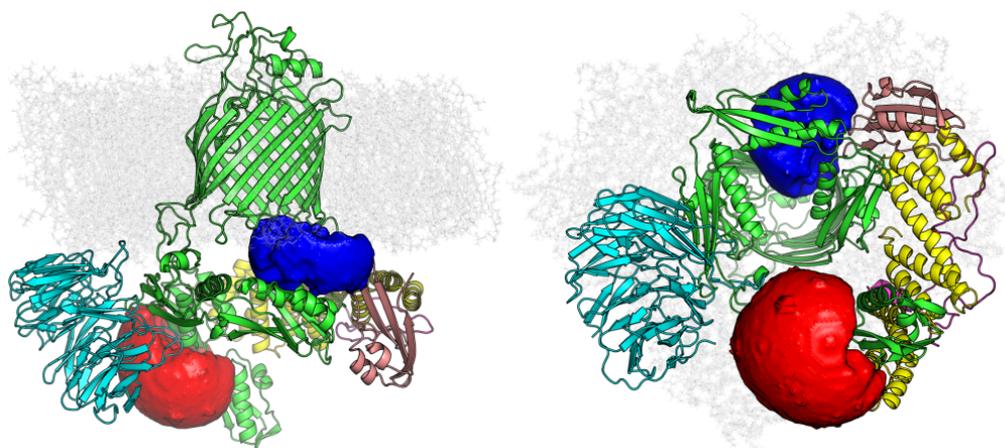


Fig. 4.20 Labelling of BAM(R127C/N520C) with Alexa Fluor 488 and DyLight 594. (Left) ÄKTA purification trace of double labelled protein. Highlighted in blue are the regions of the trace taken forward for smFRET studies. Indicated on the graph are the labelling stoichiometries as determined by the absorbance of the dyes, and the amount of 10 mM dye stock used for labelling. (Right) UV-Vis trace of the pooled and concentrated fractions taken on a NanoDrop 2000 and used to calculate the corrected A280 values.

A BAM Complex (Closed State)



Side View

Top View

Closed(5D0O): $E = 0.28$ $\langle R_{DA} \rangle = 72.5 \text{ \AA}$

Open(5LJO): $E = 0.15$ $\langle R_{DA} \rangle = 83.1 \text{ \AA}$

Fig. 4.21 Single-molecule FRET of the BAM complex shows small changes in the conformational landscape of BAM in response to changes in lipid phase. (A) Image of BAM(R127C/N520C) modelled in the closed state with Alexa Fluor 594 attached to position R127C and Alexa Fluor 488 attached to N520C. Atomistic MD of BamA in a DMPC bilayer was performed for 50 ns to allow the BamA barrel and lipids to relax (courtesy of Dr B. Schiffrin, University of Leeds). BamA was then exchanged with the full BAM complex by alignment on the back of the barrel. Accessible volumes of each dye was then simulated using the FPS software and the average inter-dye distance and average FRET efficiency calculated (Kalinin et al., 2012). This was repeated with the dyes attached to the opposite positions, and with the BAM complex open state.

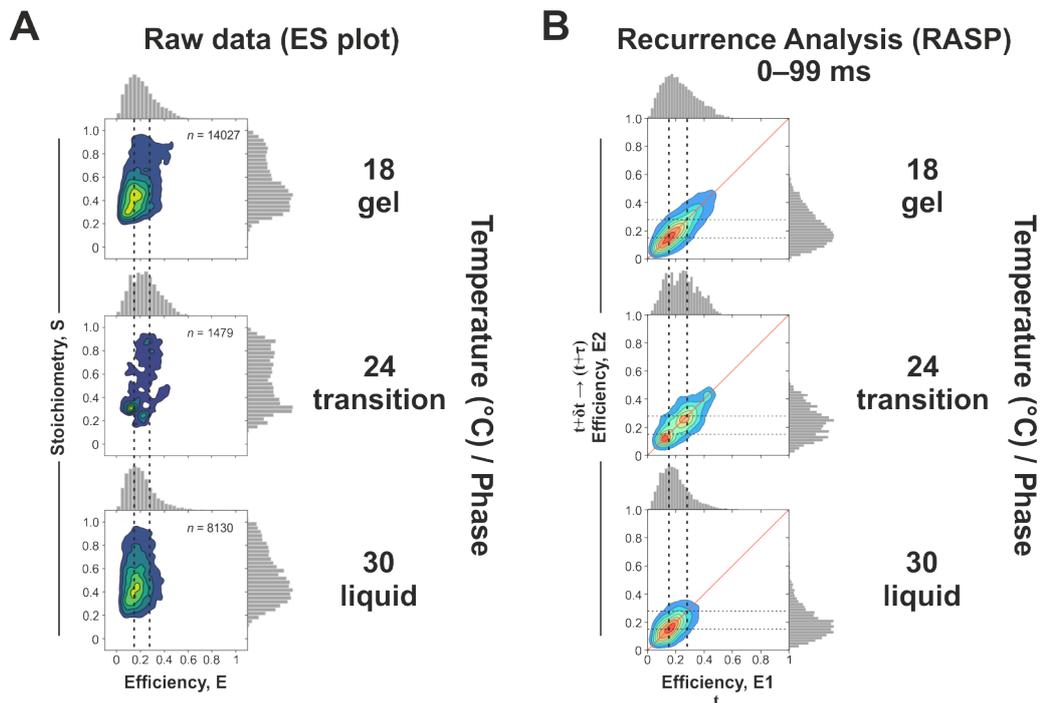


Fig. 4.22 Analysis of smFRET data of labelled BAM(R127C/N520C) in DMPC at different lipid phases. (A) Raw efficiency vs. stoichiometry (ES) plots from ALEX smFRET experiments on labelled BAM(R127C/N520C) embedded in DMPC bilayers at different temperatures. This represents the pooled sum of all bursts detected throughout the experiment. Dashed lines represent the predicted values for the open (left line) and closed (right vertical line) states of BAM. 1D histograms at the top and left show number of events versus efficiency or stoichiometry, respectively. Bin sizes are 0.05. The central plots show a kernel density estimator (KDE) function applied to a 2D plot of the histograms. (B) RASP analysis performed on the data in (A) (see main text and Material and Methods). Dashed lines represent the predicted values of the open (left vertical line, bottom horizontal line) and closed (right vertical line, top horizontal line) states. The red diagonal line represents the identity line between E1 and E2. Shown is analysis performed between fluorescent bursts detected at time t (E1) and all bursts detected after t until time $t+\tau$ where τ is 99 ms. 1D histograms at the top and left show number of events versus E1 or E2. Bin sizes are 0.05. The central plots show a kernel density estimator (KDE) function applied to a 2D plot of the histograms.

immediately before smFRET measurements were taken, reduces the risk of hydrolysis or a retro-Michael reaction causing the dye to dissociate (Lyon et al., 2014), and the risk of opening of the dye ring by hydrolysis leading to deactivation of the fluorophore. The downsides (BAM more prone to aggregation, less densely populated vesicles, poorer constitution efficiency) are not important under the single-molecule conditions employed here as the very low concentration reduces the risk of aggregation, very low density vesicles are favourable (1 per liposome), and the massive excess of LUVs over BAM improves constitution efficiency. This LPR corresponds to approximately one molecule of BAM per liposome according to Eq. 4. 100 μ L Samples of BAM at ~100—500 pM were incubated on a heated confocal stage at the stated temperatures for ~5 minutes before smFRET data was collected using an Alternating Laser Excitation (ALEX) setup (Kapanidis et al., 2005). The sample was replenished every 15 minutes until sufficient data had been collected. The ALEX technique utilises two lasers (488 nm and 594 nm) to selectively and alternately excite the donor and acceptor dyes, and a dichroic mirror to split the emitted fluorescence to two avalanche photodiode (APD) detectors for the donor and acceptor emission (setup described in more detail in Material and Methods). This allows the separation of signal into: $D_{ex}D_{em}$, $D_{ex}A_{em}$, $A_{ex}A_{em}$, and $A_{ex}D_{em}$ (where D_{ex} is excitation by the donor dye line, A_{ex} is excitation by the acceptor dye line, D_{em} is emission collected in the donor channel, and A_{em} is emission collected in the acceptor channel). From this, correction factors and stoichiometries of dye labelling can be derived and thus just the signal arising from acceptor emission of a donor excited molecule in a correctly double-labelled molecule – the true FRET signal – can be analysed.

Equation 4:

$$N_{total} = \frac{\left[4\pi \left(\frac{d}{2}\right)^2 + 4\pi \left[\frac{d}{2} - h\right]^2\right]}{a}$$

where N_{total} is the total number of lipid molecules in the liposomes, d is the diameter of the liposome (100 nm), h is the bilayer thickness (3.6–5.0 nm for DMPC), a is the lipid head group area (~0.71 nm² for PC lipids)

Figure 4.22A shows the FRET efficiency (E) vs dye stoichiometry (S) plots derived from the ALEX experiment for each temperature. In the gel phase, at the transition temperature, and in the liquid phase, the 2D plots are similar suggesting that the conformational ensemble that the BAM complex populates is also similar, irrespective of the phase of the lipid bilayer (and thus the packing of the lipids). The FRET E values match well with the predicted E values from the modelling of the dye accessible volumes suggesting that most molecules are in the open state (with some tending toward the closed state). The location of the dyes precludes a direct assessment of the state of the gate of the BamA

barrel, only the relative distance between turn 3 at the back of the barrel (which shows very little change between all known structures) and POTRA2. However, all open structures of the BAM complex show this coupling with movements in the POTRA domains – most likely due to the connection of the POTRA domains to the bottom of $\beta 1$ of BamA which swings outwards in the open state. Therefore, it is likely this FRET pair is reporting on the presence of at least two conformations of the BAM complex being populated in solution.

To try to identify whether this represented static or dynamic heterogeneity (that is, whether these states were interconverting on slow (>1 s) or fast (μs – ms) timescales), the data were analysed using Recurrence Analysis of Single Particles (RASP) (Hoffmann et al., 2011) (Figure 4.22B). RASP exploits the very low concentration of protein present in diffusion-based smFRET experiments and the randomness of diffusion caused by Brownian motion to extract information about protein dynamics on a μs to ms timescale. Briefly, each burst that makes up panel A in Figure 4.22 has a timestamped arrival time association with it. An autocorrelation function (ACF) can be applied to this dataset to estimate the probability that two bursts (an initial burst with efficiency, E_1 , and a second burst with efficiency, E_2) are from the same molecule diffusing back into the confocal volume, or a new molecule from the bulk solution, after a certain amount of time has elapsed between detections. In other words, whether any burst detected at within a time window, τ , after the first burst is a *recurrence* of that molecule re-entering the confocal volume. Only bursts which have a 95 % probability of being from the same molecule (according to the ACF) are analysed further. For the 100 nm proteoliposomes used here, each liposome should contain on average one molecule of BAM and its diffusion coefficient will be very slow (in comparison to free protein molecules, or detergent-solubilised proteins). This slower diffusion coefficient means that most molecules of BAM will take 10s to 100s of milliseconds to completely diffuse away from the confocal volume and be replaced by a new molecule. If dynamics occur on a timescale faster than diffusion then the measured E value will change between detection 1 (E_1) and detection 2 (E_2). For the plots shown in Figure 4.22B dynamics would be represented as off-diagonal density. Looking at this data, there does not appear to be considerable off-diagonal density, but more of a continuous bridging of density between states suggesting conformational exchange on a timescale slower than the diffusion time of the proteoliposomes (>100 ms).

Although the differences in the FRET states populated by BAM in each lipid condition (gel, transition, liquid) are small, the cumulative RASP data shown in Figure 4.22B, which is essentially an average of the states occupied by a population of BAM molecules across a 100 ms time window, make these trends easier to see. By representing the efficiency values in two dimensions (and collected within a fixed time window) it is

easier to see the ‘spread’ of density representing the conformational ensemble occupied by BAM. Overall, the conformational ensemble occupied by BAM is very similar but appears broadened in the gel phase with more BAM molecules occupying higher FRET states. On the basis of the accessible volume simulations of dyes, these higher FRET states could represent a greater occupation of the closed state of BAM in the gel phase. At the transition temperature, the ensemble appears to become more bimodal between the open and closed states. However, further analysis with FRET pairs in other locations is needed to substantiate these conclusions.

4.3 Discussion

4.3.1 Folding into gel phase lipids

The observation of uncatalysed tOmpA folding into gel phase DMPC lipids was surprising, as previous literature stated that OMPs require lipid bilayers to be in the liquid disordered phase in order to support productive folding, and that they are unable to fold into gel-phase lipids (Surrey and Jähnig, 1992; Rodionova et al., 1995; Kleinschmidt and Tamm, 1996; Dewald et al., 2011; Kleinschmidt, 2015). In fact, for uncatalysed and BAM-catalysed folding, faster folding kinetics were observed in gel phase compared with the liquid phase DMPC LUVs. Despite the prevailing wisdom that gel-phase lipids cannot support folding, evidence that this assertion may be incorrect, at least for some OMPs, is present in the literature. Danoff & Fleming observed tOmpA folding into DMPC 100 nm LUVs at 20 °C (gel phase) and 26 °C (liquid phase) on a timescale of hours (10,000s of seconds) (Danoff and Fleming, 2015b). While no quantitation was performed on that data, the results showed that although the folding yields were dramatically different, the rate of folding of the 26 °C dataset was less than an order of magnitude faster than at 20 °C. It was also shown that OmpA can fold into DMPC small unilamellar vesicles (SUVs) at 20 °C, but not 4 °C (Rodionova et al., 1995), suggesting that bilayers continue to become highly packed as you move further below the transition temperature. At temperatures further above the transition temperature (i.e. 30 °C used in this chapter), the folding may be even slower. It has been shown that a mutant version of the 12-stranded autotransporter EspP(Δ 5) which contains a truncated passenger domain can fold into BAM-containing *di*C_{16:0} (DPPC) proteoliposomes at 30 °C which is below the transition temperature of DPPC (41 °C) (Hussain and Bernstein, 2018). In that study, EspP folded with a similar t_{50} into DLPC, DMPC, and DPPC BAM proteoliposomes at 30 °C (respective T_m of -2, 24, 41 °C) of $132\text{--}162 \pm 12\text{--}18$ s. A more varied result was observed for OmpA into

the same BAM proteoliposomes with t_{50} values of 120 s, 198 s, and 264 s for DLPC, DMPC, and DPPC, respectively. In contrast to the data shown here for tOmpA, folding of EspP($\Delta 5$) into BAM DMPC proteoliposomes was slower at 25 °C ($t_{50} = 432$ s) than at 30 °C or 37 °C (t_{50} : 150—156 s) – despite the proximity to the transition temperature.

In comparisons between the catalytic rate enhancement of BamA alone versus the full BAM complex it was observed that the greatest difference in rate enhancement is observed in gel phase lipids. In the gel phase, lipids are tightly packed which can alter the mechanical properties of the bilayer. This tighter packing might help to ‘lock’ BamA in a gate closed conformation (Figure 4.2), or otherwise constrain any mechanistically important dynamics, reducing its ability to disorder lipids or perform other functions to accelerate tOmpA folding. This could also explain the inability of BamA alone to catalyse folding of tOmpA into *E. coli* polar lipids (Iadanza et al., 2016), in that case the tighter packing of lipids would not come from the gel phase (as liposomes formed from *E. coli* polar lipids have a transition temperature ~ 3 °C — see Table 4.1) but due to the longer acyl chain lengths (C16—C18) increasing the van der Waals packing forces. The data from the laurdan lipid order probe experiments presented here suggest that the full BAM complex is able to affect the packing of lipids, resulting in a lowering of the transition temperature of the DMPC proteoliposome system. This broadens the transition region and allows the BAM complex to create a larger temperature window in which lipids are in a ‘more disordered’ state. However, as seen in the first and second order derivatives of the laurdan data, at even lower temperatures there is not enough energy in the system to overcome the forces driving tight packing of lipids and the system reverts to being fully gel-phase. In summary, it appears that the additional lipoproteins of the full BAM complex enhance the efficiency of lipid disordering which, in this DMPC model system, has the affect of lowering and broadening the transition temperature. This may mean that the 20 °C samples in BAM proteoliposomes still have partial ‘transition’ character rather than being fully gel-phase.

4.3.2 Energetic costs of poor folding yields

For a OM-containing bacteria such as *E. coli*, robust, fast, and efficient folding of OMPs is important for a number of reasons: to maintain the physiological function of the OM in protecting against antimicrobial agents and acquiring nutrients (which depends on OMPs being correctly folded into it) (Noinaj et al., 2010; May and Grabowicz, 2018b), to prevent toxicity from aggregated and misfolded OMPs (Cho et al., 2014; Mitchell and Silhavy, 2019), and to reduce the metabolic waste of producing OMPs which will misfold and

require degradation as well as the cost of upregulating stress response machineries (Guo and Gross, 2014). The first two points have often been focussed on as the key problems that the BAM complex helps to overcome, but the final point, that folding yields are important for metabolic efficiency, has not been given as much weight. In the laboratory, most bacteria are grown as monocultures in nutrient rich environments which have an abundance of energy sources and very little to no ecological competition (Good et al., 2017). However, *in vivo* in a host or in the environment, energy sources can be a limiting factor to growth and parsimonious utilization of internal energy supplies can be key to microbial competition and the ability to inhabit, infiltrate, and thrive within an ecological niche (Hibbing et al., 2010). It has been estimated that the synthesis of a single *E. coli* cell costs between 20 to 60 billion equivalents of high energy phosphate bonds (as found in ATP) (Stouthamer, 1973; Akashi and Gojobori, 2002). Using known absolute synthesis rates of proteins in *E. coli* (Li et al., 2014), metabolic costs of per amino acid biosynthesis in *E. coli* (Akashi and Gojobori, 2002), and the curated list of OMPs found in *E. coli* K-12 (see Introduction Table 1.1), it is possible to calculate an approximate metabolic burden of OMP synthesis in *E. coli*. Integrating these datasets gives a value of ~4.7 billion equivalent high energy phosphate bonds consumed for OMP biogenesis per generation of *E. coli* cells. Or around 8—23 % of the entire energy expenditure of the cell. From the data presented in Section 4.2.2 and Section 4.2.3 (Figure 4.13 and Figure 4.15) the presence of the BAM complex gives a ~1.5—30 % yield improvement of tOmpA folding yield over BamA alone, dependent on the lipid phase, and SurA gives a ~5—6 % yield improvement to BAM-catalysed folding of tOmpA in the gel or liquid phase. Even a 5 % improvement in yield could save 0.5 % of the energy cost of the whole cell. Although the folding data described here are only for tOmpA, this is also calculated to have the largest metabolic cost in an *E. coli* cell due to its position as the most abundant OMP (100,000—200,000 copies per cell) (Henning et al., 1973; Rosenbusch, 1974; Li et al., 2014). It should be noted that the *per protein* cost of larger OMPs is greater than OmpA (cf. OmpA = 7756 ~P, LptD = 18790 ~P – where ~P represents an equivalent high energy phosphate bond) and if these proteins are more aggregation prone, or their intrinsic folding propensity is lower, then the boost in folding yields provided by SurA and BAM may be higher and thus the energy savings higher also.

Although the intrinsic folding propensity of different OMPs and their degree of dependence on the full BAM complex versus BamA alone has not been investigated in detail, we can draw some conclusions about this from the literature. *In vitro* folding studies on a broad selection of OMPs showed that some OMPs have a high folding propensity (as judged by their final folded yield) when folding into LUVs composed of PC lipids, these included OmpA (8-strands), OmpX (8-strands), PagP (8-strands), and BamA (16-strands) (Burgess

et al., 2008), while others were comparatively worse, including OmpW (8-strands), OmpT (10-strands), OmpLA (12-strands), FadL (14-strands), and OmpF (16-strands). It is not clear if this is due to intrinsic properties of the different sequences of these proteins or their folding landscapes, or whether these proteins are simply less soluble, or a mix of both. There does however, appear to be a trend that larger OMPs have poorer folding yields. In the presence of the BAM complex, and/or SurA, at least one of these proteins, OmpT, has been shown to fold more efficiently and its aggregation propensity reduced (Hagan et al., 2010; Iadanza et al., 2016; Humes et al., 2019). Differential proteomics experiments showed that deletion of SurA reduces the levels of LptD (26-strands) and FhuA (22-strands) protein detected, but not their expression levels, suggesting that the folding yields of these proteins are dependent on the presence of SurA (Vertommen et al., 2009). Deletion of SurA also caused a 30-fold reduction in the levels of the usher protein FimD (24-strands) and this effect was titratable in a *surA* depletion strain (Justice et al., 2006; Palomino et al., 2011). Subunits of the BAM complex have also been linked to improving folding of certain OMPs. *In vitro* OmpT folding was less efficient without BamB (Hagan and Kahne, 2011); deletion or depletion of BamB also caused a 5-fold reduction in levels of FimD (Palomino et al., 2011); levels of OmpF and LamB (18-strands), and to a lesser extent, OmpA, in the OM decrease upon deletion of BamB, while TolC (3x4-strands) and OmpC (16-strands) levels were unchanged (Charlson et al., 2006). In *Salmonella enterica*, which has homologues of all *E. coli* BAM subunits, BamD was found not to be essential for cell viability, but it was the only component of the BAM complex (besides BamA) that was required for full expression of the type 3 secretion system and the bacterial flagellar proteins (Fardini et al., 2009). In *Caulobacter crescentus*, a BamE deletion strain showed misassembly of the outer-membrane secretin of the type IV secretion system and reduced amounts of TonB-dependent receptors (Ryan et al., 2010). No reports of a BamC-dependent effect on specific OMP levels could be found in the literature. Together, these studies suggest that some OMPs, particularly large and specialized proteins, may require specific subunits of the BAM complex and/or the chaperone SurA to ensure efficient assembly into the outer membrane.

4.3.3 Comparison between the biophysics of DMPC phases and the outer membrane of *E. coli*

In biophysical experiments a complicated and highly multivariate ‘real’ biological system is often constrained *in vitro* into a much smaller parameter space where the influence of just one or a small number variables on the property of interest are investigated. The test of a ‘good’ biophysical experiment might therefore be how well the lessons learned in

this simplified system can be related to the ‘true’ environment in the cell. In this chapter, DMPC LUVs were used as a mimic of the native lipid bilayer of BAM and OMPs and so it is valuable to consider what is known about *E. coli* membranes and particularly, the OM.

E. coli is known to alter the lipid content, particularly the length and degree of saturation of acyl chains, in response to changes in growth temperature. This process, termed ‘homeoviscous adaptation’ (Sinensky, 1974), suggested that bacteria actively maintain their membranes at a particular level of ‘fluidity’ or in a particular phase. Total lipid extract from *E. coli* K-12 W3110 grown at 30, 37, 42, or 45 °C showed approximately the same headgroup content with a minor monotonic increase and decrease in cardiolipin and phosphatidylglycerol, respectively (Velázquez and Fernández, 2006). More significant was the increase in the ratio of saturated over unsaturated acyl chains, indicating a change to a more ‘rigid’ mixture of phospholipids at higher growth temperatures. Using laurdan GP values a T_m of these extracts of <14 °C was calculated for the bacteria grown at 30 or 37 °C, ~20—22 °C when grown at 42 °C, and ~27 °C when grown at 45 °C (Velázquez and Fernández, 2006). These results suggest that lipids in *E. coli* are natively in the liquid phase, and as much as ~20 °C above the transition temperature.

Few *in vivo* studies on bacterial membrane order are available in the literature, but they suggest that the true picture is more complicated than that derived from lipid extracts described above. Using laurdan and 1,3-diphenyl-1,3,5-hexatriene (DPH) labelled fixed *E. coli* suggested the presence of membranes that are predominantly in the liquid state, but a state which is heterogeneous, with at least two distinct phases detected – one more liquid, and one less liquid (Vanounou et al., 2002). However, the authors’ interpretation of their data may be incorrect. The paper assumes that these dyes are localised primarily to the inner membrane based on previous literature (Fishov and Woldringh, 1999) on hydrophobic DPH and the amphiphilic FM 4-64, this assumption may not be correct as more recent studies have shown FM 4-64 to partition specifically into the OM (Rojas et al., 2018) and biochemically one would assume that hydrophobic molecules would eventually equilibrate equally between both hydrophobic environments, or be trapped in the first hydrophobic environment they encounter. Furthermore, the reported GP values of these two phases at 37 °C were 0.22 in one environment and 0.11 in the other. Although they report both to be indicative of the liquid phase, the value of 0.22 is in fact very close to the values measured for the transition state in this chapter (second derivative suggests transition region spread between GP values of 0.11—0.46 for pure DMPC). The GP dependence on temperature was measured for each of these two distinct phases using the ability to photoselect specific molecules at different excitation wavelengths. Using the same method used in this chapter, the first derivative of the more liquid phase (showing 0.11 GP at

37 °C) showed a phase change with a peak ~24 °C but the less liquid phase (with 0.22 GP at 37 °C) showed no clear phase change between 10—40 °C suggesting that the phase transition occurs above the growth temperature, shows very weak cooperativity, or both. Anisotropy of laurdan, which is more sensitive to local order of the lipids as well as the phase, suggested a pre-transition ~32 °C for the less liquid phase, again supporting the idea of a phase transition slightly above the growth temperature. Pre-transitions, where they have been observed in synthetic lipid mixtures, often occur ~8—10 °C below the main transition (see Figure 4.3 for a comparison to DMPC). Finally, the anisotropy values at 37 °C in the less liquid phase (0.22 GP) were paradoxically higher than the more liquid phase (0.11 GP), indicating greater lipid dynamics despite being closer to the gel phase. This paradox could be explained, however, if the probe was measuring an environment near a phase transition.

On the basis of the physicochemical properties of LPS extracts, and the outer membrane, other authors have argued that the OM would be more likely to exist in the gel phase at physiological temperatures (Nikaido, 2003). Realistic *in vitro* models of the outer membrane which include an asymmetric bilayer with LPS in the outer leaflet were studied by neutron reflectometry and attenuated total reflection fourier transform infrared spectroscopy (ATR-FTIR) and showed an unusual mixed characteristic with elements of liquid phases and gel phases (Paracini et al., 2018). In this study two transition midpoints were observed below (~36.2 °C) and above (~39.3 °C) 37 °C, for the outer (LPS), and inner (DPPC) leaflets, respectively. The exact *in vivo* composition of the inner leaflet lipids and LPS varies from this study so the absolute melting temperatures may vary, but it suggests that the LPS component of the asymmetric OM may confer greater rigidity. How this would affect the folding of OMPs is unclear, but as membrane-spanning proteins they would need to insert through both leaflets of the outer membrane. A higher transition temperature boundary for control of the fluidity of the outer membrane by *E. coli* is also suggested by the recent evidence of an mRNA thermostat for the regulation of the *lpxT* gene. This inner membrane protein can covalently modify LPS before it reaches the outer membrane and the protein was shown to be stable between 28 °C to 42 °C, however, its mRNA levels fall dramatically across this same range (with intermediate levels at 37 °C) (Sciandrone et al., 2019).

E. coli polar lipid extract is often used as the most ‘correct’ *in vitro* mimic of native lipid membranes due to its headgroup composition and mixture of acyl chains, but as discussed in Section 4.2.1 this still misses a lot of the diversity of native outer membranes (wrong hydrophobic thickness, no LPS, no asymmetric leaflet structure, no integral or interacting proteins, much lower transition temperature). Therefore, it could be argued that

DMPC bilayers are no better or worse mimics, dependent on the question under study by the experimenter.

4.3.4 Effect of SurA on OMP folding

The effect of SurA on OMP-folding and on BAM-catalysed folding has been investigated previously. It was shown that OmpA folded slower into BAM-containing POPC proteoliposomes in the presence of SurA (t_{50} : 192 ± 48 s versus 306 ± 30 s) and tOmpA + SurA also showed the same rate as OmpA + SurA (although tOmpA without SurA was not quantified) (Hussain and Bernstein, 2018). In this chapter, a more detailed analysis of the affect of SurA on tOmpA folding was undertaken. Overall, SurA appeared to have only a marginal effect on folding rate under most conditions, with the most dramatic effects occurring for folding into liquid phase lipids (30 °C). Here, the presence of SurA accelerated BAM-mediated folding by ~4-fold, but slowed uncatalysed folding by almost 2-fold. However, for folding into gel or transition temperature lipids via BAM, the increase in yield was nominal. SurA has been reported to interact directly with the BAM-complex *in vitro* (Bennion et al., 2010), *in vivo* (Sklar et al., 2007b; Wang et al., 2016; Alvira et al., 2019; Carlson et al., 2019), and by *in silico* predictions (Cong et al., 2019). It is possible that under certain conditions, SurA is able to cause an allosteric change in the BAM complex or to select certain conformations which are more favourable to OMP delivery, thereby increasing the observed rate of folding. The results of smFRET experiments presented here indeed suggested that the conformational landscape of BAM is slightly altered dependent on the lipid conditions (Figure 4.22). For the much lower uncatalysed folding rate at 30 °C caused by SurA, although this shows that SurA is capable of *frustrating* OMP folding, it may be an effect which is magnified by the very slow timescales of spontaneous folding into thicker membranes and not necessarily relevant in a cellular context where a vast periplasmic quality control system exist to degrade OMPs which fall off their folding pathway (Soltes et al., 2017).

Perhaps more surprising is the relative lack of change in folded yield of tOmpA caused by the presence of SurA (with the exception of the aforementioned 30 °C uncatalysed rate). This suggests that, at least under these conditions, SurA is not critical for tOmpA folding. However, its function *in vivo* may be different. There is growing evidence for the presence of a inner-to-outer membrane “supercomplex” between the BAM complex and SecYEG (Wang et al., 2016; Alvira et al., 2019; Carlson et al., 2019). In these supercomplex models, SurA acts as a bridge between the two machineries, suggesting a role where it accepts incoming substrates through the SecYEG machinery and provides

them a ‘safe path’ guiding them to the BAM complex, rather than capturing OMPs and diffusing across the periplasm like a shuttle between BAM and SecYEG (Costello et al., 2016). In the experiments performed here, SurA is utilised more like a shuttle and so it may not be recapitulating the situation seen *in vivo*. Alternatively, the headgroup content of the lipid bilayer may have an important role in modulating the activity of BAM. In the inner leaflet of the outer membrane, and in *E. coli* polar lipids, there is a high proportion of PE-containing lipid types (as well as PG and cardiolipin) (Morein et al., 1996). PE lipids are typically non-bilayer forming and their incorporation into lipid bilayers increases the stored curvature stress. Either the charge of these headgroups, or their ability to increase lateral pressure in a membrane (Booth and Curnow, 2009), could change the conformational landscape accessible by BAM in a manner which is partially relieved by SurA. Early studies on *in vitro* reconstitution of the BAM complex showed that SurA increased the rate of folding of OmpT, as measured by the cleavage of a fluorogenic peptide by the correctly assembled protein (Hagan et al., 2010). Nonetheless, the viability of SurA deletion strains (Sklar et al., 2007b) suggests that this protein is not essential for OMP biogenesis *in vivo*.

4.3.5 First evidence for physiologically compatible folding rates

One of the unresolved questions in the field of OMP biogenesis is how to reconcile the slow OMP folding rates observed so far in *in vitro* in purified systems with values that are needed to support the rates of doubling observed in *E. coli*. Even in the presence of the BAM complex in *in vitro* reconstituted systems, the rates of folding are far below that required to support growth. Using a similar approach as described in Section 4.3.2 to calculate the metabolic cost of OMP biogenesis, we can estimate the rate the BAM complex needs to achieve in order to fold a full complement of OMPs. Taking the absolute synthesis rates of all OMPs (Li et al., 2014), and the curated list of *E. coli* K-12 OMPs, we know the number of molecules of BamA, the total number of OMPs, and the doubling time of *E. coli* (~20 mins). From this a rate of 0.139 s^{-1} is calculated. The average lifetime would therefore be $(1/0.139) \sim 7.2 \text{ s}$, very close to the t_{50} value measured for BAM-mediated folding at the transition temperature (24 °C) of 13.2 s. Considering for losses of protein which aggregate or misfold before reaching BAM, this is the first time a physiologically relevant rate of folding in lipid bilayers has been observed. Furthermore, this rate was measured into a bilayer with a hydrophobic thickness matching the outer membrane. These results suggest that, with the right modulation of the bilayer properties – by controlling the global phase (perhaps by homeoviscous adaptation *in vivo*), and the local environment

(through BAM) – the BAM complex alone is sufficient to allow the rapid incorporation of OMPs required for bacterial growth.

Although the proteins, and the calculations, made here are relevant for lab-grown *E. coli*, this is one of the fastest growing bacteria known to science. The fastest outer-membrane containing bacteria so far reported in the literature is *Vibrio natriegens* which is a marine bacterium with a doubling time of 10 minutes. The small difference in generation time between that bacteria and *E. coli* could easily be accommodated by a greater copy number of the BAM complex, or a smaller burden of OMPs. The vast majority of bacteria have much slower growth rates in their native environment, and this may be one explanation for the large number of components (BamABCDE) in many well-studied bacteria (which have often been selected for their experimentally tractable growth rates) (Anwari et al., 2012; Webb et al., 2012a).

4.3.6 Effect of the BAM complex on membrane physical properties

Recent studies using bactericidal monoclonal antibodies which target the loops of BamA in the BAM complex in LPS-truncated strains of *E. coli* have shown that *E. coli* which develop spontaneous resistance to this BAM-mediated killing have mutations in the *lpxM* gene which transfers a C14 (myristoyl) chain to penta-acylated LPS, creating hexa-acylated LPS (Storek et al., 2018a). Antibody sensitivity was restored (i.e. bactericidal effects of the antibody was restored) when *lpxM* was expressed from a plasmid. Assays of membrane fluidity using a pyrene-based probe showed that membrane fluidity decreased in the resistant strains, and this effect was recapitulated in other conditions that decrease membrane fluidity (high salt, longer LPS sugar region, lower temperature). The levels of OMPs were not reduced in $\Delta lpxM$ strains in the absence of the antibody. This suggests that there is mechanistic link between BamA / the BAM complex and membrane fluidity, as BAM is most sensitive to inhibition when the membrane is more fluid. The authors in that study hypothesised that BamA activity (as part of the BAM complex) may be lower when the OM is excessively fluid. Although the rate enhancement observed by BAM is greater in the liquid (more fluid) phase than the gel phase (more rigid) (see Figure 4.11 and Table 4.2), the uncatalysed folding of OMPs is considerably slower. This means that high levels of catalysis may be essential to obtain physiologically-relevant folding rates in membranes in more fluid phases. Furthermore, the absolute rates of folding are actually faster for BAM-catalysed folding in gel (more rigid) versus liquid (more fluid) phases agreeing with the authors hypothesis of lower absolute activities in more fluid membranes.

The results of laurdan experiments in Section 4.2.4 (Figure 4.17) show that the first derivative of the slope of laurdan fluorescence can recapitulate the thermotropic response of pure DMPC bilayers. This was used to interrogate the effect of BamA and BAM integration into the lipid bilayer, on the transition temperature - a metric which is closely tied to the physical properties of the bilayer, and particularly the packing of acyl chains in its core. BamA-containing DMPC LUVs showed no difference compared with pure DMPC LUVs, but incorporation of the full BAM complex resulted in a 2 °C decrease in the transition temperature, a shallower slope, and a broader onset and termination of the transition. Changes in the transition temperature of pure lipid bilayer systems have been observed before, with DMPC and DPPC embedded in Nanodiscs, showing an *increase* in T_m of 5 °C, and 2 °C, respectively, with good agreement between measurements performed by DSC and laurdan fluorescence (Denisov et al., 2005). The shallower slope (lowered peak height in first derivative) suggests a decrease in enthalpy of the phase transition, an effect which is due to reduced cooperativity in the transition. In the above Nanodisc study in pure lipids, a decrease in enthalpy of the transition was also observed which was attributed to a boundary layer of lipids in contact with the Nanodisc. These boundary lipids display a greater level of disorder versus the bulk lipid (Pourmousa and Pastor, 2018) and this may be analogous to lipids around the BAM complex. The increase in T_m for DMPC in Nanodiscs has been shown to be due to stabilisation of the gel phase due to the restricted geometry of the Nanodisc system increasing the lateral pressure (in the liquid phase, the bilayer area expands) (Denisov et al., 2005). Conversely, BAM can be said to stabilise the liquid phase according to the laurdan data shown in this chapter. Hydrophobic mismatch has been suggested as a reason for changes in the transition temperature of lipid systems with embedded proteins (Sperotto and Mouritsen, 1988; Marsh, 2008) and the decrease in T_m observed with BAM could point to a greater degree of hydrophobic mismatch versus BamA alone - perhaps driven by the opening of the $\beta 1$ - $\beta 16$ gate which has only been observed for the full BAM complex (see Figure 4.2 for examples of currently known possible states of BamA). Theoretical calculations suggest that larger hydrophobic mismatch increases the range at which a protein exerts its influence of lipid order (Katira et al., 2016). Furthermore, the same study suggested that the largest effect will occur in disordered membranes close to a transition between liquid-ordered and liquid-disordered phases (n.b. liquid-disordered corresponds to the “liquid” state in this chapter, while liquid-ordered phases are induced by the presence of certain lipids and proteins, such as cholesterol). The differences between BamA and the full BAM complex suggest that the additional subunits of BAM (i.e. BamBCDE) may increase the efficiency of lipid disordering, either by direct interaction with the membrane (Iadanza et al., 2016), allosterically through BamA, or by clustering multiple BAM molecules together via BamB (Gunasinghe et al., 2018) to grow local disordered patches into larger

regions of disorder or perhaps lipid rafts. This latter idea would be analogous to AFM observations on phase transitions in DMPC which showed that phase transitions grew out of nucleation points that may correspond to random packing defects in the bilayer (Xie et al., 2002).

4.3.7 Single-molecule FRET showing conformational states of BAM

Structures of BamA and the full BAM complex have shown the presence of two major conformational states of BamA, an open gate and a closed gate (Figure 4.2). In all structures solved to date, a movement of the POTRA domains appears to be coupled to this gate motion. A solution structure of the BAM complex in DDM solved by cryoEM (Iadanza et al., 2016) showed a single state (the open state) with no indication of subclasses which may represent closed states. This opens the question as to whether the resting state of the BAM complex is open, closed, or a mixture of both. Structures of BamA alone have always been solved in the closed state, but molecular dynamics simulations suggest that this can open even in the absence of the other subunits (Noinaj et al., 2013). Nonetheless, it remained unclear whether the closed structures of the full BAM complex were well-populated in lipid bilayers or whether they were simply stabilised in the crystal lattice but sparingly present in solution (or perhaps, only present as part of a reaction cycle). If they were stabilised by the crystal lattice, then tighter packing of lipids (and a reduced bending modulus leading to increased lateral pressure) should drive the BAM complex into a closed state. To answer this question, and explore whether the catalytic power of the BAM complex in different gel phases could be related to changes in its conformational ensemble, smFRET experiments were performed on BAM labelled with a FRET pair between a periplasmic turn on BamA and POTRA2.

Figure 4.22 shows that two major populations are present for BAM in a DMPC bilayer and these agree well with simulated FRET efficiencies for the closed and open state, based on crystal and cryoEM structures. However, little difference is observed in FRET distribution between the different lipid phases. Attempts to quantify the dynamics and timescale of conversion between these states using RASP was not successful but this in itself, and the continuous density between the states (Figure 4.22B), suggests dynamic heterogeneity on a much slower timescale than the measurements (seconds and greater). The RASP dataset can be used as an approximate ‘fingerprint’ of the conformational ensemble of the BAM complex and, surprisingly, appears to suggest a greater degree of conformational heterogeneity in the gel phase than the liquid phase. However, minor

changes in FRET efficiency may be due to the change in the bilayer thickness between the gel and liquid states and so a definitive conclusion cannot be drawn at this time.

4.4 Conclusion

The results presented here support the view that the full BAM complex is a more catalytically active OMP-folding enzyme than BamA alone (Hagan et al., 2010; Iadanza et al., 2016). Depending on the phase of the lipid, BAM is capable of accelerating the folding of tOmpA by ~130X compared with uncatalysed folding (in the gel phase, 20 °C), or as much as ~450X when paired with SurA (in the liquid phase, 30 °C). At the transition temperature, where lipid packing is maximally disordered, BamA alone shows only very modest enhancement (1.5X) compared with the rate of uncatalysed tOmpA folding under the same conditions and this may point to a primary (or ancestral) role of BamA as a pure lipid-disorderase. In contrast, the full BAM complex is capable of providing additional catalytic enhancement. Remarkably, under these conditions BAM is able to fold tOmpA into membranes with a hydrophobic thickness mimicking that of the outer membrane at a rate which is physiologically compatible with the doubling time of *E. coli*. This is the first time *in vitro* folding studies using lipid bilayers, and especially ones with a representative thickness and degree of curvature of the OM, have been able to recapitulate such a high rate of OMP folding meaning in theory, no extra components or conditions need to be met (apart from the presence of the BAM complex) to achieve observed rate of OMP-flux *in vivo*. Mechanistically, BAM may achieve this by the presence of the additional subunits of the BAM complex which can drive even greater disordering of lipids, and may combine this effect with templating of β -strands of nascent OMPs (Schiffrin et al., 2017a; Höhr et al., 2018; Doyle and Bernstein, 2019).

The BAM complex is a complicated machine, and this chapter provides the first direct evidence that it may exert a strong influence directly on lipid bilayer. Combined with what we know about β -strand templating by the BAM complex (Höhr et al., 2018; Doyle and Bernstein, 2019), a picture is emerging of a multi-faceted mechanism by which BAM accelerates the insertion and folding of β -barrels like an enzymatic cooper with multiple tools to complete the job. It remains to be seen whether BAM uses these different mechanisms of action in a concerted manner (multiple mechanisms working at the same time), in parallel (proteins can take either route if the flux is high enough to saturate one), in tandem (one proceeds the other), or on a case-by-case basis (where some OMPs require minimal intervention, and others may require the full gamut of BAM mechanisms). More

work, using approaches such as described here, will be able to answer these questions for different OMPs.

Chapter 5

Towards visualising the nanoscale organization of OMP biogenesis *in vivo*

5.1 Introduction

The BAM complex is one of the two major machineries that are essential for building the outer membrane (OM) of *E. coli* (the other being the Lol machinery, made up of LolABCDE) (Konovalova and Silhavy, 2015). From the periplasmic space to the OM, only 7 proteins have been identified as essential for bacterial growth under all conditions (Baba et al., 2006; Yamazaki et al., 2008; Goodall et al., 2018). These are *bamA*, *bamD*, *lolA*, *lolB*, *lptA*, *lptD*, and *lptE*. All of these proteins are involved in the biogenesis of the OM, but all apart from *lolA* and *lolB* are themselves dependent on the BAM complex (LptA and LptE require LptD for assembly (Chimalakonda et al., 2011; Sherman et al., 2018), which is itself assembled by the BAM complex (Bos et al., 2004; Wu et al., 2006); BamD is a component of the BAM complex and has never been identified in the absence of BamA (Anwari et al., 2012; Webb et al., 2012a), although apparent BamA only species do exist (e.g. *Thermus thermophilus* & *Thermosynechococcus elongates*) (Webb et al., 2012a)). This stresses the importance of the BAM complex as a major hub in the organisation and biogenesis of the OM. Despite the key role that BAM plays in maintaining the OM, relatively little is known about its organisation in the OM *in vivo*, and the organisation of its substrates and interacting partners (such as chaperones, proteases, and envelope stress sensors) (Ricci et al., 2013; Narita et al., 2013; Cho et al., 2014; Konovalova et al., 2016; Soltes et al., 2017; Konovalova et al., 2018; Hart et al., 2019; Tata and Konovalova, 2019).

Early studies in *Salmonella typhimurium* monitored the appearance of ferritin-antibody-bound OmpF in the OM at time points after induction (caused by shifting from high-to-no NaCl media) using electron microscopy (Smit and Nikaido, 1978). Discrete clusters ~50 nm in diameter spread throughout the whole length of the cell were observed 15 minutes after induction, and the size of the clusters grew to ~100 nm in diameter after constitutive expression of OmpF during outgrowth as well as their number on the cell surface increasing. Intriguingly, experiments where induction continued for 15–30 minutes before the cells were plasmolyzed (causing shrinkage of the IM away from the OM), fixed, and imaged, showed that many clusters appeared at sites of adhesion between the inner and OM. Fluorescence microscopy studies using LamB with an Sfp transferase labelling tag in a loop were able to monitor the appearance of LamB as its loops emerged and became accessible for dye labelling (Ursell et al., 2012). LamB was seen to appear as discrete puncta spread throughout the cell length, but were about twice as likely to appear near the mid-cell than the poles. As cells grew (without further labelling) fluorescence which was initially distributed across the whole cell periphery moved towards the poles, puncta intensity reduced and size of puncta split. This suggested the insertion of new unlabelled material ‘diluting’ existing LamB molecules, although it was not clear whether this was caused by slowly diffusing outer membrane proteins (OMPs) joining and mixing with existing puncta, or active insertion into puncta (which would indicate retention of active BAM in or near these puncta). *In vivo* live cell confocal TIRF imaging studies using fluorescently-labelled colicins (toxins which bind particular OMP receptors) showed that the TonB-dependent receptors (TBDRs) CirA and BtuB clustered in ~500 nm diameter ‘islands’, which first appeared near the constriction site of a dividing bacteria, and moved to the poles during successive divisions (Rassam et al., 2015). BamA with an HA-epitope inserted in loop 7 was also labelled using anti-HA Alexa Fluor 488 (AF488)-conjugated antibody (Ab), as well as an anti-BamC Ab targeted with an AF488 secondary Ab. These were also found to co-localise and migrate with OMP islands, with ~20–30 % colocalization by fluorescence overlap. This study also found restricted diffusion of OMPs, agreeing with *in vitro* and *in silico* studies (Jarosławski et al., 2009; Casuso et al., 2012; Goose and Sansom, 2013). It is likely that these promiscuous ‘sticky’ interactions are of two types: the first are sequence-mediated due to interactions of aromatic residues between OMPs (Jarosławski et al., 2009; Rassam et al., 2015), and the second due to hydrophobic mismatch between the OMPs and the OM, where hydrophobic surfaces that match the thickness of an OMP are enriched around it (either lipids with matching acyl lengths, or other OMPs) (Yin and Kindt, 2012; Katira et al., 2016). This second effect is likely due to the energetic cost of distorting the acyl chains or exposing hydrophobic regions (of lipid or protein) to polar moieties (versus forming clusters, rafts, or islands) (Marsh, 2008). This may be particularly pronounced for TonB-dependent receptors like BtuB which are

predicted to have a large and asymmetric hydrophobic mismatch with the OM (Ellena et al., 2011). Super-resolution single-molecule localisation microscopy studies of BAM on fixed *E. coli* cells showed BAM distributed throughout the cell but tending to cluster in islands of around 150 nm diameter at a density of $\sim 11 \mu\text{m}^{-2}$ spread relatively uniformly along the cell length (Gunasinghe et al., 2018). The authors also found that these islands formed smaller domains in the absence of active protein synthesis triggered through the addition of rifampicin (to ~ 126 nm diameter, $\sim 15 \mu\text{m}^{-2}$). Further crosslinking and super-resolution microscopy studies showed that BamB is important for BAM-BAM clustering. ΔbamB strains had smaller islands (~ 100 nm diameter) with more BAM seen as not part of a distinct cluster ($\sim 8 \mu\text{m}^{-2}$). In contrast to the clustering observed for BAM, these authors observed that the autotransporter Ag43 was spread diffusely throughout the OM. Other studies of OMPs which have not explicitly focussed on OMP organisation and biogenesis have observed that clustering of OMPs is a common trend, but is not universal. Fluorescence microscopy of fixed *E. coli* cells chimerically expressing the *Klebsiella oxytoca* type II secretion system secretin, PulD, fused to the fluorescent protein (FP) mCherry showed the formation of distinct foci in the OM (Buddelmeijer et al., 2009). The TBDR, FepA, was also observed in clusters in the OM in live-cell imaging experiments and this clustering did not change when ExbBD (which form the TonB-ExbBD inner membrane complex that binds TBDRs) was deleted (Lill et al., 2016). Studies in *Brucella abortus* found that newly inserted Omp25 and Omp2b (an OmpF homologue) appeared primarily at the new pole and constriction site of bacteria and tended to colocalize with new peptidoglycan and a specific isoform of LPS (rough LPS) (Vassen et al., 2019). The autotransporter IcsA of *Shigella flexneri* localizes specifically at the poles of the cell and its polar organisation is thought to be pre-organized in the cytoplasm before secretion by the cytoskeletal filament, MreB (Steinhauer et al., 1999; Charles et al., 2001; Krokowski et al., 2019). This effect is maintained across diverse autotransporters in Proteobacteria (IcsA and SepA, *Shigella flexneri*; AIDA-I, *E. coli*; BrkA, *Bordetella pertussis*), is dependent on the presence of full-length LPS, and is even observed during heterologous expression of NalP from *Neisseria meningitidis* in *E. coli* or *S. flexneri* IcsA in *E. coli*, *Salmonella typhimurium*, *Yersinia pseudotuberculosis* or *Vibrio cholerae* (Charles et al., 2001; Jain et al., 2006). The eukaryotic TOM complex imports preproteins from the cytosol into mitochondria and the channel is a β -barrel transmembrane protein, Tom40, which is located in the mitochondrial OM. Preproteins labelled with quantum dots were stalled during transit and showed that this import machinery is localised to discrete clusters in the OM, and interestingly it formed a supercomplex with the inner mitochondrial membrane (Gold et al., 2014).

Although these studies shed light on the underlying organisation of OMP biogenesis, a number of questions remain unanswered: is the punctate organisation (and insertion)

of OMPs spread exclusively around BAM complexes or are they randomly distributed relative to BAM (indicating either BAM-independent insertion events, or some mechanism by which BAM can clear newly inserted OMPs from its local neighbourhood)? What is the organisational (oligomeric) structure of a BAM island? How are major periplasmic OMP chaperones such as Skp and SurA organised relative to the OMP biogenesis machinery? How is the inner membrane Sec secretion machinery organised relative to the BAM complex, does it co-localize laterally (suggesting the formation of transient super-complexes as suggested by Wang et al. (2016); Alvira et al. (2019) and Carlson et al. (2019))? Answering these questions requires high resolution information about the localisation of molecules in the envelope and membranes of *E. coli*, on the order of individual BAM complexes (see Figure 5.1).

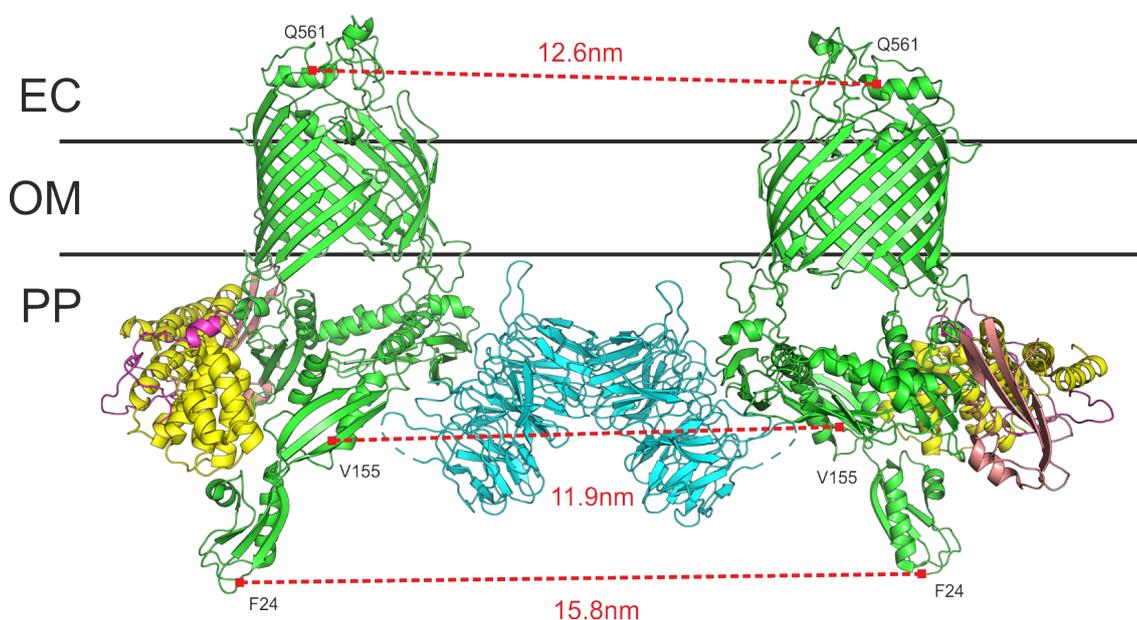


Fig. 5.1 Possible minimum BAM-BAM distance within a cluster. Black lines indicate the approximate position of the OM. For a clustering mechanism of BAM mediated by BamB that tries to satisfy some of the crosslinks found in Gunasinghe et al. (2018), two complexes of BAM were aligned and distances between assessed. Distances were measured in PyMol between the locations marked by small squares (F24-F24, POTRA1; V155-V155, POTRA2; Q561-Q561, loop 4). EC = extracellular, OM = OM, PP = periplasm. BAM (PDB: 5D0O).

With the exception of one paper (Gunasinghe et al., 2018), all imaging studies on OMP biogenesis have been performed using diffraction-limited imaging techniques, which limits our ability to understand the nanoscale organisation of this process. Gunasinghe et al. chemically fixed *E. coli* BW25113 cells before labelling with primary antibodies against BamA POTRA domains (in detergent solubilised bacteria), BamC (which is located on the extracellular surface of the OM), or loops of the autotransporter Ag43. These were then labelled with an Alexa Fluor 647-labelled secondary antibody. Although this approach produced excellent data, provided new insights into the organisation of the BAM

complex, and allowed the labelling of wild-type proteins with a very bright and far-red dye (eliminating issues of cellular autofluorescence) there were also some drawbacks. The use of chemical fixants has been shown to effect the distribution of LamB in the OM of *E. coli*, possibly by causing extraction of proteins (Gibbs et al., 2004), and this may effect other OMPs. The use of secondary Abs risks disrupting protein-protein interactions or causing artificial clustering due to the multivalency of antibody binding, and the large distance of the fluorophore (~27 nm) from the protein of interest, will degrade image resolution (Figure 5.2). Finally, for performing accurate counting of molecules the problem of multivalent binding arises (exacerbated by the secondary Ab) as well as an unknown affinity of the Ab for its target *in vivo*, and a variable degree-of-labelling of the secondary Ab.

Since the initial development of super-resolution microscopy through stimulated emission depletion (STED) (Willig et al., 2006), photo-activated localization microscopy (PALM) (Betzig et al., 2006; Hess et al., 2006), and stochastic optical reconstruction microscopy (STORM) (Rust et al., 2006), there has been huge interest in developing new techniques and pushing the resolution limits on existing methods. One approach that has seen increasing growth in the last decade is cryogenic single-molecule fluorescence microscopy. The first reports of single-molecule detection at cryogenic temperatures (1–2 K) came about in the 1990s and it was discovered that the photon yield was increased substantially (by 2-3 orders of magnitude) at low temperatures due to reduced photobleaching (Moerner and Orrit, 1999). At that time, before the advent of widespread usage of super-resolution methodologies, three issues held back further development of cryogenic imaging: technical challenges of imaging at low temperatures, the limited palette of fluorophores, and background fluorescence which could be considerable due to the concomitant increase in fluorescence intensity from minor impurities in the sample, optical parts, and other materials. Furthermore, preparation of the sample was key to avoid the generation of amorphous ice which can scatter laser light into emission channels. Over the last decade many of the technical hurdles have been overcome as the potential benefits come to light. For example, it was found that a wide range of dyes and FPs across the visible spectrum show an improvement in photon yield of one to three orders of magnitude (Weisenburger et al., 2013, 2014; Li et al., 2015; Hulleman et al., 2018; Dahlberg et al., 2018). This huge photon budget can theoretically allow a localisation precision beyond the size of individual atoms, but has been practically demonstrated in 3D to the Ångstrom level (2–5 Å) by localising multiple fluorescent labels in isolation or attached to a single protein (or complex) (Weisenburger et al., 2017; Furubayashi et al., 2017).

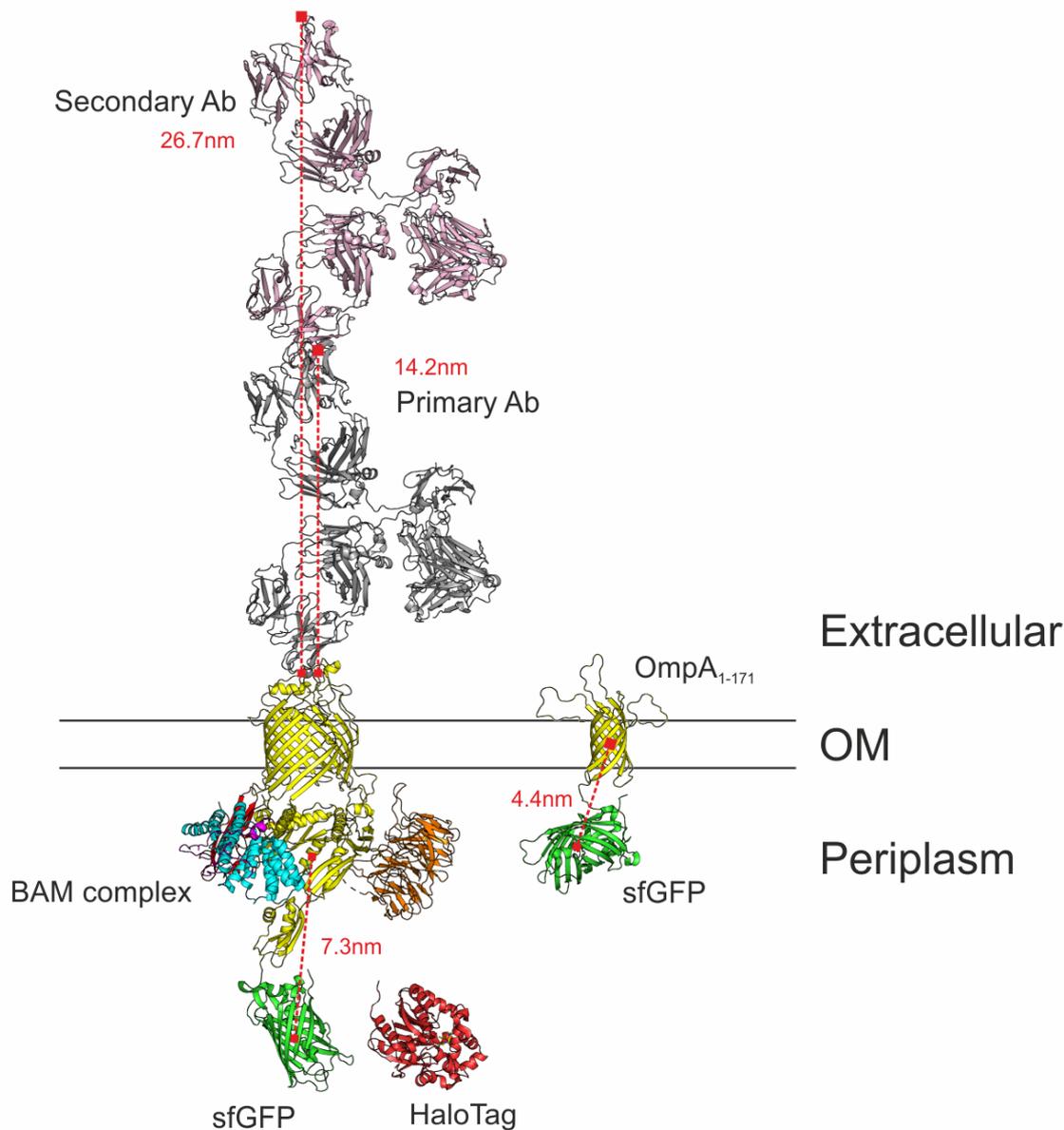


Fig. 5.2 Illustration of differences between location of fluorescence emitter and location of labelled target for different labelling strategies. Distances were measured in PyMol between the locations marked by small squares. For most commercially available labelled antibodies, the degree of labelling (how many fluorophores per Ab) and the conjugation points of the labels, is usually either not homogeneous or not published. Here, the distance to the furthest loop has been used to approximate a “worst case” scenario. Linkers are all approximate (not actual linker used). Residues measured: OmpA_{L139}, sfGFP_{chromophore residue 66}, BamA-POTRA2_{K111}, BamA-loop_{6F675}, IgG1-primary_{chain D, Y56}, IgG1-secondary_{chain D, R71}. BAM complex (PDB: 5D0O) (Iadanza et al., 2016), sfGFP (PDB: 4LQT) (Xia et al., 2013), HaloTag (PDB: 5Y2Y) (Kang et al., 2017), OmpA (PDB: 2GE4) (Cierpicki et al., 2006), IgG1 (PDB: 1IGY) (Harris et al., 1998).

These advances have allowed some specialist groups to begin developing the technique into specific methodologies and demonstrating its use on biological systems *in vivo*. These include the use of correlated cryo-PALM cryo-EM to image the type VI secretion system (T6SS) with a component fused to photoactivatable (PA)-GFP in *Myxococcus xanthus* with <160 nm localisation uncertainty, allowing the T6SS to be picked and refined from a large number of candidate structures in cryoET (Chang et al., 2014); cryogenic super-resolution imaging using the photoactivatable red FP PAmKate to look at the polar organisational scaffold protein, PopZ, in *Caulobacter crescentus* yielding an average localisation uncertainty of 8 nm (Dahlberg et al., 2018); and cryoSTORM to visualize an inner membrane ABC transporter McjD fused to EGFP in live *E. coli* with <8 nm localisation uncertainty and ~12 nm resolution (Wang et al., 2019).

In this chapter the cryoSTORM method is used to perform cryogenic single-molecule localisation microscopy. This method allows samples to be fixed in a frozen (vitreous) hydrated state, eliminating the fixation artefacts caused by other methods (Bleck et al., 2010). It also takes advantage of the unusual photophysics at cryogenic temperatures which causes some FPs to blink in an analogous way to the organic dyes used in room temperature STORM (hence the name), broadening the palette of fluorescent fusion proteins which are compatible with single-molecule localisation. The method comprises two essential elements: a solid-immersion lens, and the vitrification and imaging of the sample at 77 K. The numerical aperture (NA) of an optical system is directly related to its light collection efficiency and magnification. In most microscopy applications, an objective lens (the collection optics) is coupled to a sample on a glass coverslip through air (in the case of dry objective / lenses) or through a liquid (water or oil, for immersion lenses). For dry objectives, the NA is typically limited to a maximum of 1 and for immersion lenses this ranges from 1.3—1.75. Solid-immersion lenses (SILs) in the shape of a truncated sphere (known as a Weierstrass or superSIL geometry) can increase the numerical aperture of a system, and therefore the magnification and light collection efficiency, by as much as n^2 where n is the refractive index of the SIL (Barnes et al., 2002). Furthermore, with a dry objective, the maximum theoretical collection efficiency is 50 % due to the spherical / isotropic emission profile of a single emitter (i.e. half of the emission is away from the collection lens). When a fluorophore is coupled to a dielectric medium, it preferentially emits photons in the direction of the medium with the higher refractive index, with higher refractive indices resulting in stronger coupling and thus more anisotropy in fluorescence emission (Yoshita et al., 2002; Wang et al., 2019). Wang et al., use a cubic zirconium SIL to improve the NA of the cryoSTORM system to ~2.17 and calculated a theoretical collection efficiency of 91 %, 3.3-fold higher than a 0.9 NA dry objective. Combining these improved photon collection properties of the SIL with the enhanced photon yield

at cryogenic conditions was responsible for the improved resolution of the cryoSTORM technique.

In this chapter preliminary data on super-resolution imaging of the OMP biogenesis machinery including transmembrane-only OmpA (tOmpA) and the BAM complex using the cryoSTORM method is presented. These data show that BAM organises in discrete clusters distributed throughout the cell, with tOmpA also organising in clusters but which are smaller and more disperse. Two-colour super-resolution imaging appears to show only a weak co-localisation between the two systems. The work in this chapter was undertaken in collaboration with Dr Paul White (University of Leeds) [creation of OMP-FP constructs, *in vivo* assays, sample preparation for imaging], Benji Bateman (Central Laser Facility [CLF], Research Complex at Harwell [RCaH]) [microscope development, imaging], Dr Laura Zanetti-Domingues (Central Laser Facility, Research Complex at Harwell) [technique development, ASIL preparation]. The CLF at RCaH kindly provided training and access to cryoSTORM apparatus.

5.2 Results

5.2.1 Approaches for labelling of OMPs

In approaching the labelling of BamA for super-resolution imaging, we aimed to satisfy a number of criteria with the choice and method of labelling: (1) the tag must not perturb the function of BamA, (2) the tag should report on a location very close to the BAM complex (so that resolution is not significantly reduced by the length of the linker), (3) the fluorophore should be functional in the periplasm, and (4) the fluorophore should have favourable photophysical properties under cryogenic conditions (bright and blinking). These criteria were assessed for a number of fusions to BamA of FPs and self-labelling protein tags as discussed below.

BamA is an essential protein in OM-containing bacteria (Wu et al., 2005), and its complete depletion or loss of function is lethal to cells. To test the function of BamA and the effect of mutations and deletions, a BamA depletion strain of *E. coli*, JCM166 has been developed (Wu et al., 2005), in which the chromosomal BamA gene is replaced with a gene cassette encoding BamA under the control of an arabinose-inducible promoter and an ampicillin resistance marker. To complement the depletion of BamA in these strains, cells are transformed with BamA containing an N-terminal His-tag in a pZS21

vector maintained with a kanamycin resistance marker (Kim et al., 2007). In this vector BamA is under control of a $P_{LTetO-1}$ promoter which is constitutively expressed when the cognate repressor, *tetR*, is not expressed - as is the case in the JCM166 strain which contains no copy of *tetR*. This allows the construction of BamA with N-terminal fusion proteins, and a method for testing the effect of these on BamA function. When grown in the absence of arabinose and the presence of glucose (to further suppress 'leaky' chromosomal expression of BamA) (Guzman et al., 1995), practically all BamA in the cell is expressed from the pZS21 vector and so BamA fusions which prevent its essential function will be incompatible with growth. A simple test for (1), therefore, involves plating out JCM166 cells transformed with a BamA-fusion-protein on solid media containing glucose and no arabinose and comparing this to a positive (wild-type BamA) control. The results of these plate assays are shown in Figure 5.3 (performed by Dr. Paul White, University of Leeds) and indicate that N-terminal fusions of BamA are well-tolerated by *E. coli* cells under growth in rich medium. Double-transformants with a second plasmid encoding OmpA FP fusions are also shown.

As shown in Figure 5.2, an approximate lateral (xy) and axial (z) uncertainty for a FP (modelled as sfGFP) or a labelled HaloTag is $\sim 2\text{-}3$ nm in xy , and $\sim 5\text{-}7$ nm in z (dependent on exact linker length). This is in comparison to a primary or secondary antibody which can be located as far as 14 nm (primary Ab) - 27 nm (secondary Ab) away from their target. Although alternative methods exist which in principle would allow even closer labelling (such as direct labelling with dyes either by unnatural amino acid incorporation (Lang and Chin, 2014), or by use of a tag-binding motif such as the tetracysteine motif recognised by FAsH and ReAsH dyes (Griffin et al., 1998)), the use of fluorescent protein fusions can still fulfil criteria (2).

Finding a protein that satisfied both (3) and (4) was challenging as both areas are relatively understudied, and both properties are unrelated (folding in the oxidising environment of the periplasm, and having favourable photophysics at 77 K), requiring testing of both these conditions in parallel. Promising reports of cryogenic-active FPs must be checked for their effect on BamA activity in the periplasm, and vice versa.

The palette of FPs available to researchers is huge with over 700 different variants described on the (non-exhaustive) fluorescent protein database, FPbase (Lambert, 2019). However, many of these FPs have only been characterised in cytosolic environments. Periplasmic expression of fluorescent proteins has been much more difficult, owing to problems with secretion in an unfolded state through the Sec machinery and incorrect folding in the periplasm (Feilmeier et al., 2000), and the oxidising environment of the periplasm that can cause problems in disulfide bond formation or oxidation (Kadokura and Beckwith,

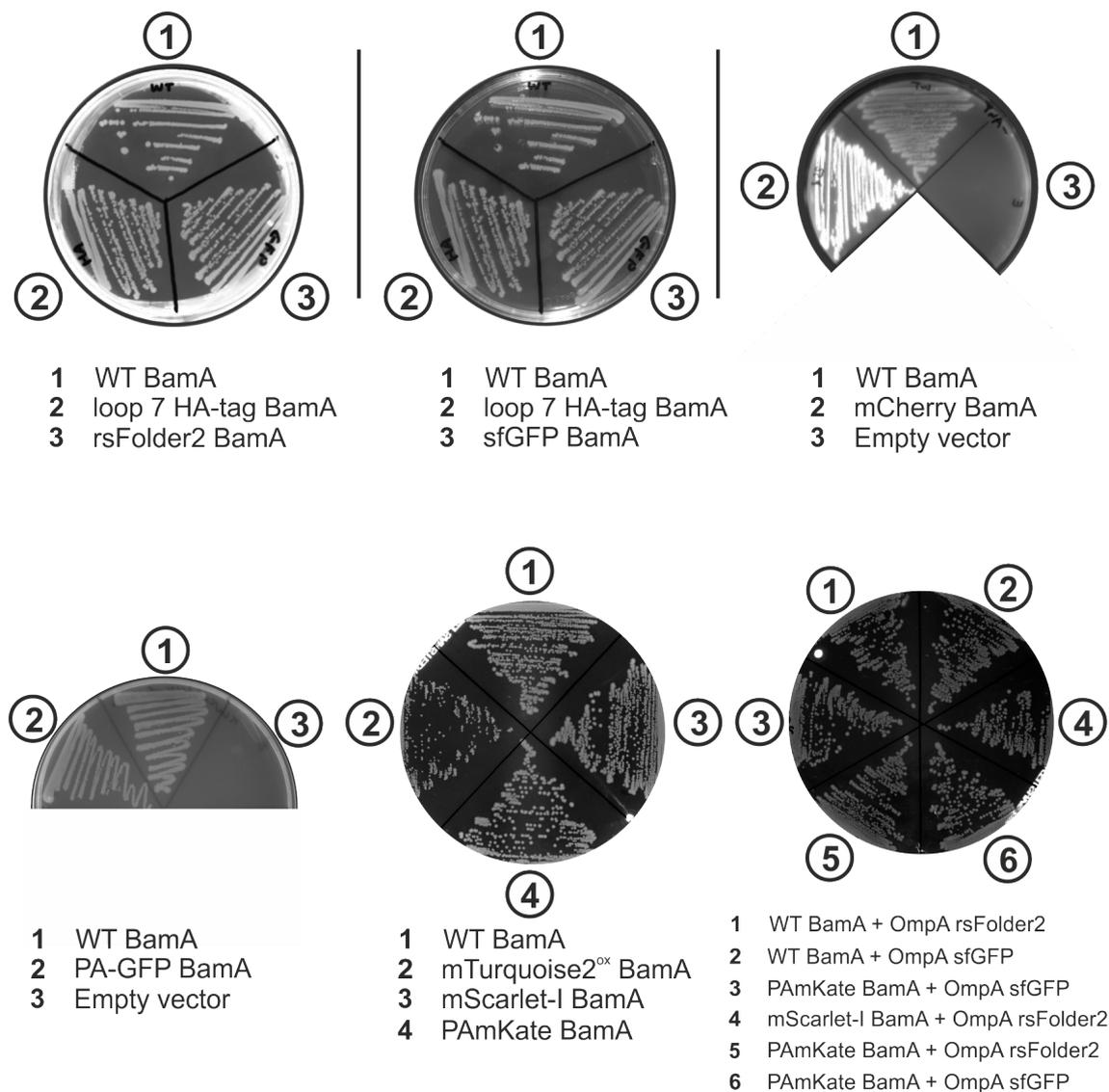


Fig. 5.3 The results of growth assays of JCM166 cells transformed with pZS21 plasmids containing BamA and double-transformants containing BamA and OmpA. JCM166 cells were transformed with plasmids containing the gene fusions indicated in the figure. These were then allowed to outgrow overnight for 1 hour at 37 °C in LB media supplemented with arabinose before being streaked onto plates containing antibiotic and glucose. (Left) Wild-type (WT) or BamA-FP fusions show good growth on solid media (LB agar) containing glucose and kanamycin to suppress expression of chromosomal BamA. This indicates that BamA can tolerate N-terminal fusions. (Right) Double-transformants containing BamA expressed from a pZS21 vector and OmpA from a pZS31 vector containing a chloramphenicol resistance marker. Plasmids are maintained by double selection with kanamycin and chloramphenicol and have been streaked onto LB agar solid media containing glucose to suppress chromosomal BamA expression. Details of this assay can be found in Materials & Methods 2.4.6.

2010; Denoncin and Collet, 2013) and maturation of the chromophore (Bartkiewicz et al., 2018). Although successful secretion of FPs folded in the cytoplasm and secreted in a folded state through the twin-arginine translocation (Tat) pathway has been demonstrated for EGFP, sfGFP, mRFP1, and mCherry (Santini et al., 2001; Chen et al., 2005; Cava et al., 2008; Bageshwar et al., 2016), export was not always efficient and this pathway is inappropriate for OMPs. However, some reports of FPs retaining fluorescence after Sec-mediated export have been reported and are summarised in Table 5.1.

Table 5.1 Literature reports of FPs targeted to the periplasm via the Sec pathway. Positive indicates reports where fluorescence was observed. Negative indicates the opposite. The table is coloured relative to the emission properties of the protein according to the traditional colour names for each wavelength range (n.b. these are not always the same as the true colour, for example YFP actually appears more green to human eyes). Protein names marked with ** are cysteine-less derivatives of the named FP. Fusion partner indicates the protein it was fused to (if any). Signal sequence indicates any special targeting signal used and ‘fusion’ indicates the native signal sequence of the fusion protein was retained. MBP**1 is a signal sequence derived from MBP and optimized for co-translational secretion by increasing its hydrophobicity (Lee and Bernstein, 2001). Lipobox* is the signal peptide from the *Klebsiella oxytoca* lipoprotein, PulA, with the lipobox motif altered for IM or OM display. Termini represents which terminus of the fusion partner the FP was fused to. Pathway indicates which secretory pathway (if known), the signal sequence pushes the protein towards. Membrane = embedded membrane protein (via Sec), post1 = SecB-dependent post-translational secretion (via Sec), post2 = SecB-independent post-translational secretion (via Sec), co = co-translational secretion (via Sec). Species shows the species the experiments were conducted in. *S. flexneri* = *Shigella flexneri*, *S. enterica* = *Salmonella enterica*, *Y. pseudotuberculosis* = *Yersinia pseudotuberculosis*, *E. carotovora* = *Erwinia carotovora*.

Fluorescent protein	λ_{exc}	Fusion partner	Signal sequence	Termini	Pathway	Species	References
Positive							
mTurquoise2 ^{Ox}	434	tOmpA, PBP5	fusion, DsbA	C, N	post1, co	<i>E. coli</i>	(Meiresonne et al., 2019)
rsFolder2	478	-	PeIB	-	post	<i>E. coli</i>	(El Khatib et al., 2016)
sfGFP	485	-	MBP**1, DsbA	-	co, co	<i>E. coli</i>	(Aronson et al., 2011; Dinh and Bernhardt, 2011)
sfGFP	485	MalE, EnvC, Pal	fusion	C	post1 / co?	<i>E. coli</i>	(Dinh and Bernhardt, 2011)
mNeonGreen	506	-	DsbA	-	co	<i>E. coli</i>	(Meiresonne et al., 2017, 2019)
mNeonGreen	506	-	OmpA, PBP5	C	post1, membrane	<i>E. coli</i>	(Meiresonne et al., 2017)
sfYFP	513	-	PeIB	-	post	<i>E. coli</i>	(Dammeyer et al., 2013)
mOrange**	548	PBP5	DsbA	N	co	<i>E. coli</i>	(Meiresonne et al., 2019)
mOrange2**	549	PBP5	DsbA	N	co	<i>E. coli</i>	(Meiresonne et al., 2019)
mScarlet-I	569	PBP5	DsbA	N	co	<i>E. coli</i>	(Meiresonne et al., 2019)
mStrawberry**	574	PBP5	DsbA	N	co	<i>E. coli</i>	(Meiresonne et al., 2019)
mRFPI	584	RseP, DsbA, MBP	fusion	C	membrane, post1, co	<i>E. coli</i>	(Chen et al., 2005)
mRFPI	584	-	lipobox*	-	co?	<i>E. coli</i> , <i>S. flexneri</i> , <i>S. enterica</i> , <i>Y. pseudotuberculosis</i> , <i>E. carotovora</i>	(Lewenza et al., 2006)
mCherry	587	-	MBP**1, DsbA, MalE	-	co, co, post1	<i>E. coli</i>	(Aronson et al., 2011; Dinh and Bernhardt, 2011)
mCherry	587	OmpA, tOmpA, Pal, PulD	fusion	C	post1, post1, co?, co?	<i>E. coli</i>	(Gerding et al., 2007; Buddelmeijer et al., 2009; Verhoeven et al., 2013)
Negative							
EGFP	488	MBP	fusion	C	co	<i>E. coli</i>	(Feilmeier et al., 2000)
mGFP	488	-	MBP**1	-	co	<i>E. coli</i>	(Aronson et al., 2011)
sfGFP	485	-	DsbA, MBP, PhoA	-	co, post1, post2	<i>E. coli</i>	(Fisher and DeLisa, 2008)
sfGFP	485	-	MalE	-	post1	<i>E. coli</i>	(Dinh and Bernhardt, 2011)
EYFP	513	RseP	fusion	C	membrane	<i>E. coli</i>	(Chen et al., 2005)

Reports of FPs used at cryogenic temperatures are similarly sparse but are summarised in Table 5.2. On this basis, a number of initial candidate FPs were chosen for testing. These included the photoswitchable / photoactivatable GFP derivatives, rsFolder2 and PA-GFP, a fast-folding highly stable 'superfolder' variant of GFP, sfGFP, and the red FP, mCherry, which were all well-tolerated as N-terminal fusions on BamA (see Figure 5.3).

rsFolder2, sfGFP, and mCherry were taken forward for testing under cryogenic conditions (see Section, 5.2.2, for details of imaging workflows), but only sfGFP-BamA showed bright fluorescence with well-separated blinks at 77 K. rsFolder2 and mCherry showed little to no fluorescence above the background autofluorescence of the cell as measured at cryogenic conditions (see Figure 5.4 for example mCherry data). For mCherry-BamA, imaging was also undertaken after incubating cells with β -mercaptoethanol to induce photoactivation and switching behaviour of mCherry as had been reported previously at room-temperature to convert mCherry to a super-resolution compatible label (Cloin et al., 2017). Unfortunately, this also failed to convert mCherry into a cryo-compatible probe (data not shown). Earlier reports of BamA-labelling had used a construct with a haemagglutinin (HA) epitope installed into loop 7 which could be targeted by an Alexa Fluor 488-labelled anti-HA primary antibody (Rassam et al., 2015). This was also tested, but similar to the reports of the original authors, we observed large amounts of unbound label (see Figure 5.4 for example), even after extensive washing, which swamped the acquisition and so this approach was not taken further.

Although sfGFP had been shown to be a cryo-compatible probe, for two-colour imaging a second colour would be needed. Based on literature reports (see Table 5.2) and discussions with groups developing this method (Dr L. Zanetti-Domingues, CLF, RCaH, personal communication), mTurquoise2^{Ox}, mScarlet-I, PA-mKate, and the self-labelling protein tag, HaloTag, were cloned onto the N-terminus of BamA. These fusions did not significantly affect growth of JCM166 cells (see Figure 5.3) (cloning and *in vivo* assays were performed (primarily) by Dr. Paul White, University of Leeds). mTurquoise2^{Ox} was not used further due its blue-shifted excitation wavelength which would likely lead to high levels of autofluorescence, which are very significant at cryogenic temperatures (Carter et al., 2018). Cells grown with expression of mScarlet-I and PA-mKate BamA fusions showed low levels of fluorescence at 77 K (data not shown, see Figure 5.4 for similar result obtained with mCherry-BamA). Attempts to overexpress mScarlet-I BamA from the P_{LtetO-1} promoter through the addition of anhydrotetracycline did not result in any increase in fluorescence intensity (data not shown). We reasoned that the low fluorescence stemmed from the longer maturation times of these proteins compared to sfGFP which is both highly stable and matures in ~13 minutes (Balleza et al., 2018) (cf. mScarlet-I, 36 minutes, Balleza

Table 5.2 Literature reports of FPs used under cryogenic conditions. Positive indicates reports where the FP was used successfully, Negative showing the opposite. The table is coloured relative to the emission properties of the protein according to the traditional colour names for each wavelength range (n.b. these are not always the same as the true colour, for example YFP actually appears more green to human eyes). Temperature as reported in the reference or 77 K where only “liquid nitrogen” was specified.

Fluorescent protein	Temperature	References
<i>Positive</i>		
Dronpa	77 K / 113 K	(Liu et al., 2015; Tuijtel et al., 2019)
EGFP	77 K / 113 K	(Liu et al., 2015; Tuijtel et al., 2019; Wang et al., 2019)
mEGFP	77 K	(Kaufmann et al., 2014)
mGeos-M	113 K	(Liu et al., 2015)
PA-GFP	77 K / 128 K	(Chang et al., 2014; Nahmani et al., 2017; Dahlberg et al., 2018; Tuijtel et al., 2019)
Padron	77 K	(Tuijtel et al., 2019)
rs-EGFP2	77 K	(Tuijtel et al., 2019)
rs-FastLime	77 K	(Tuijtel et al., 2019)
mVenus	77 K	(Kaufmann et al., 2014)
mEos3.2	113 K	(Liu et al., 2015)
mIrisFP	77 K	(Tuijtel et al., 2019)
PA-mCherry	113 K	(Liu et al., 2015)
PA-TagRFP	113 K	(Liu et al., 2015)
PS-mOrange	128 K	(Nahmani et al., 2017)
PA-mKate	77 K	(Dahlberg et al., 2018)
<i>Negative</i>		
Dronpa	77 K	(Chang et al., 2014)
PS-CFP2	77 K	(Chang et al., 2014)
PA-mCherry	77 K	(Chang et al., 2014; Dahlberg et al., 2018)
PA-mRFP1	77 K	(Chang et al., 2014)
mEosFP2	77 K	(Chang et al., 2014)
Dendra2	77 K	(Chang et al., 2014)

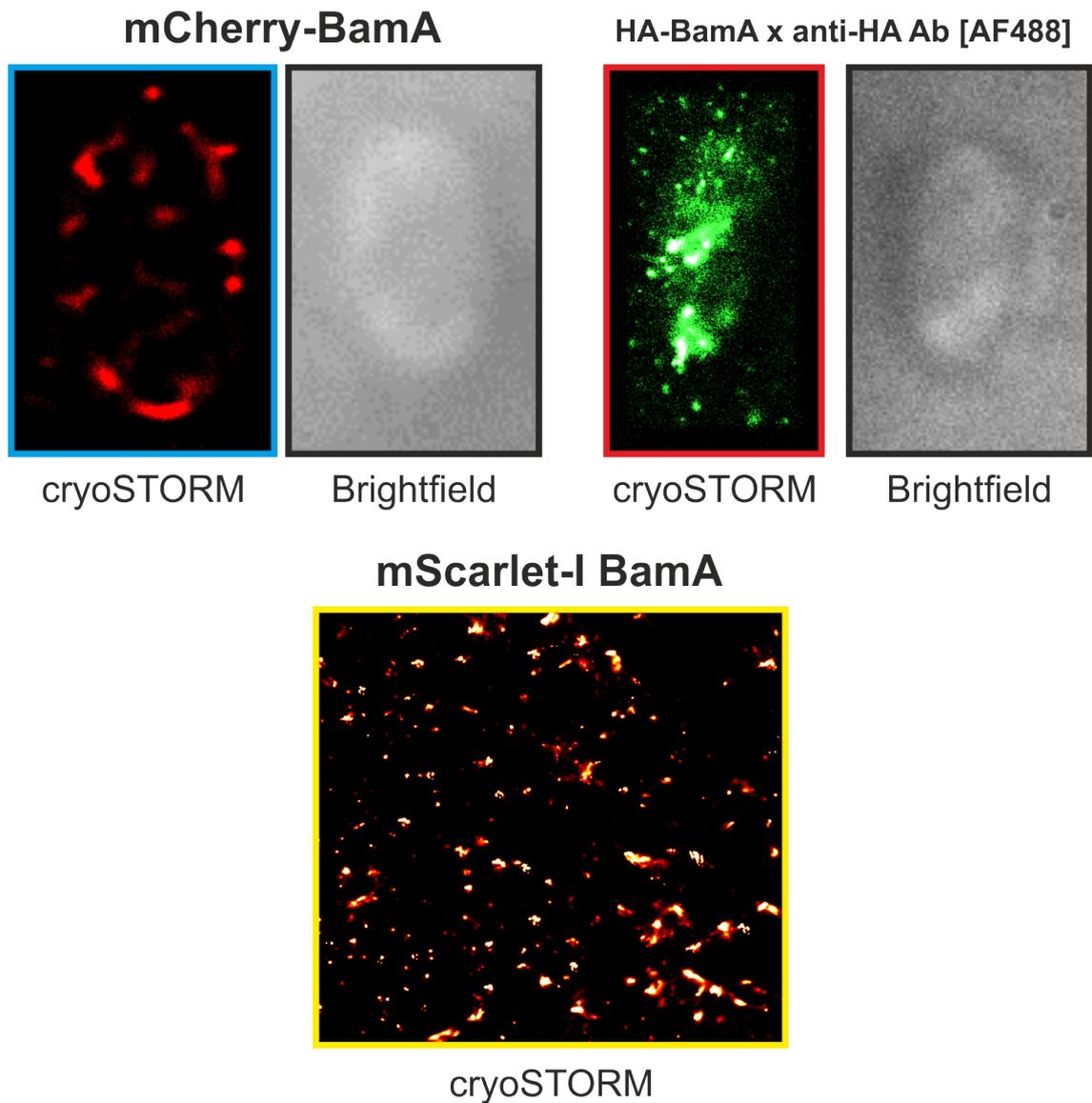


Fig. 5.4 Examples of failed experiments using BamA-FP fusions or antibody labelling. cryoSTORM reconstructions using ThunderSTORM and brightfield images are shown for mCherry and Alexa Fluor 488-conjugated antibody (AF488). mScarlet-I BamA fusion imaged after an outgrow at 20 °C to allow chromophore maturation.

et al. (2018); PAmKate, 19 minutes, Gunewardene et al. (2011)) and reports that indicate slower maturation in the oxidative environment of the periplasm compared to reported cytoplasmic rates (Meiresonne et al., 2017, 2019). To attempt to ameliorate this, overnight cultures of mScarlet-I BamA cells were shifted to lower temperatures (20 °C) during the outgrow before imaging to attempt to lower the doubling time and allow the chromophore to mature. Although this resulted in much brighter fluorescence of the protein, large levels of fluorescence outside of the bacteria were observed indicating membrane defects or cell lysis (see Figure 5.4, mScarlet-I BamA), suggesting that these low growth conditions combined with the BamA-fusion may put too large a stress on the OM. Labelling HaloTag-BamA with externally added, membrane-permeable organic fluorophores (Janelia Fluor 549) was much more successful and is discussed in Section 5.2.4.

5.2.2 Imaging BamA-sfGFP

Glycerol stocks of JCM166 strains of *E. coli* transformed with BamA-sfGFP in pZS21 vectors were freshly streaked onto LB agar plates containing 0.4 % (w/v) glucose (to suppress chromosomal wild-type BamA expression) and 50 µg mL kanamycin (to maintain the plasmid) at the start of each week. Before each imaging session, colonies were freshly picked the day before and grown overnight at 37 °C in LB containing 0.4 % (w/v) glucose and 50 µg mL kanamycin. The following day these cultures were diluted 1:100 into M9 minimal medium containing glucose and kanamycin (concentrations as above) and grown at 37 °C until reaching an optical density at 600 nm (OD₆₀₀) of ~0.6. 1 ml of each culture was then pelleted under gentle conditions (3,000 × *g*, 3 min) at 4 °C, washed and resuspended in phosphate buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 10 mM phosphate pH 7.4), and this was repeated twice more before a final resuspension in 100 µL of PBS. 2.5 µL of this high OD resuspension was applied to the flat side of custom-built hemispherical cubic zirconium asymmetric solid immersion lenses (ASILs), manually blotted to remove excess liquid, and then plunge-frozen in liquid ethane (90 K). This was then stored in liquid nitrogen (77 K) until it was ready to be imaged.

Samples were imaged on a custom-built microscope with a cryostage and holder for the ASILs (see Materials & Methods Section 2.6.6 and Wang et al., 2018, for more details). A stream of liquid nitrogen vapour maintained the sample at 77 K during imaging. The focus was manually adjusted in brightfield illumination to find the flat surface of the ASIL before areas on the surface containing *E. coli* were found by illuminating with a 470 nm LED. *E. coli*-containing fields of view were then imaged for 10,000–20,000 frames with 488 nm continuous wave excitation (200 mW, 9.11 kW.cm⁻¹) and 405 nm

activation (2 mW , 0.01 kW.cm^{-1}) lasers. Emitted fluorescence was collected on an EMCCD camera. Raw image files were processed using ThunderSTORM (Ovesný et al., 2014) in ImageJ (Schneider et al., 2012) to calculate the lateral position of fluorescent bursts in xy along with their intensity and the uncertainty in their localisation. This method is summarised for a field of view containing BamA-sfGFP transformed JCM166 cells in Figure 5.5.

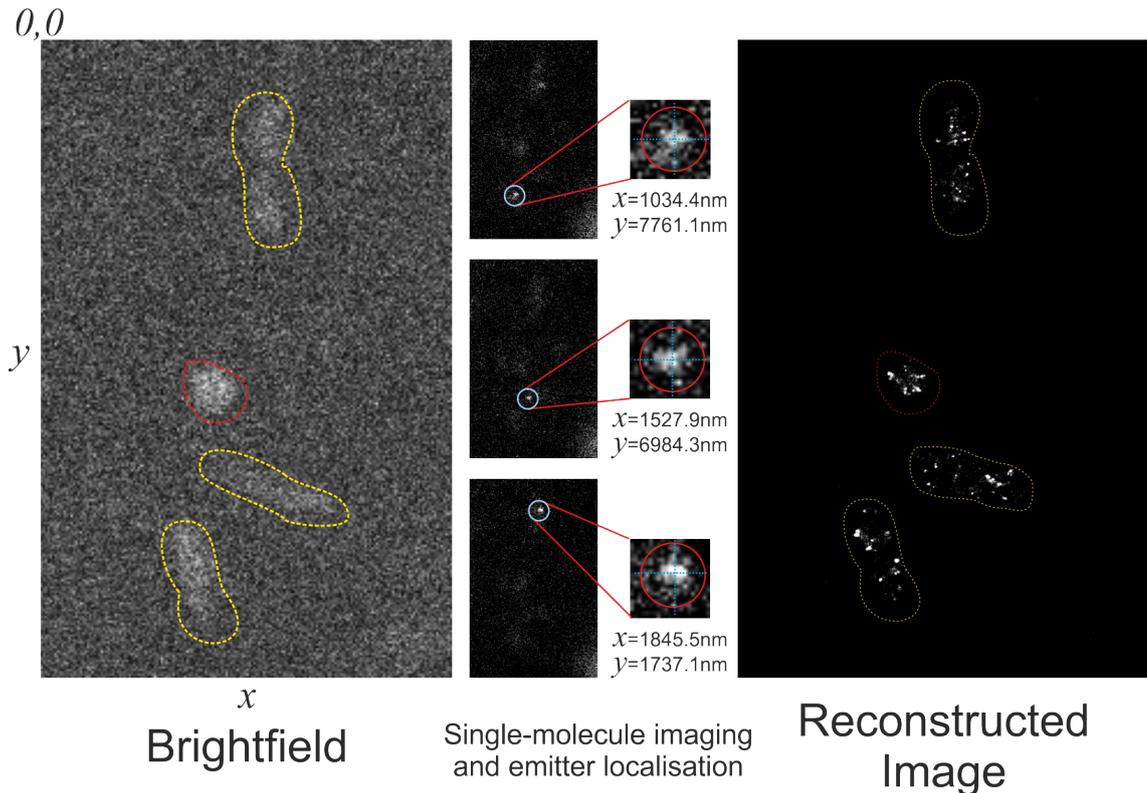


Fig. 5.5 Summary of single-molecule imaging and localisation process for samples in this chapter. (Brightfield) Bacteria are visible on the surface of the ASIL under brightfield imaging and their approximate shape is circled in yellow. Some bacteria were frozen in an upright position (circled in red) and these were not included in any downstream analyses. During image acquisition, for each frame the intensity of emitted (or reflected) light is encoded on an EMCCD camera chip as a grid of pixels (with the $0,0$ position indicated in the top left) with an associated intensity value and typically output as a stack of TIFF images. (Single-molecule imaging and emitter localisation) Image stacks are imported into ImageJ and the ThunderSTORM plugin uses a series of algorithms to detect fluorescent bursts from single emitters and localise them with sub-pixel accuracy. Shown are three example bursts from each bacterium illustrating the principles of burst detection and localisation. With knowledge of the camera pixel size (here 34 nm) these values can be converted into actual co-ordinates (in nm). To convert the relative intensity values into absolute intensity values (in photons) the photoelectrons per analogue to digital count (A/D unit), the base (background) noise level of the camera (electronic bias offset), and the EM gain (12.6 , 323 , and 1000 , respectively, in this acquisition session) are required. (Reconstructed image) The initial output of the ThunderSTORM analysis is a table of localisations with associated intensity (in photons), uncertainty (in nm), and σ -values (the standard deviation of the 2D Gaussian function used for sub-pixel localisation).

Reconstructions of BamA-sfGFP show the presence of distinct clusters of localisations throughout the cell, rather than a diffuse haze of localisations. A representative example from three *E. coli* cells is shown in Figure 5.6A. This indicates that BamA as part of the BAM complex is clustered in discrete areas in the OM, much smaller than suggested by Rassam et al. (2015) in diffraction-limited confocal studies but agreeing with the STORM analysis performed by Gunasinghe et al. (2018) on chemically fixed cells at room temperature. Compared with the study by Gunasinghe et al. (2018) the clusters appear much more sparsely populated across the surface of the bacteria, with far few clusters overall. To try and obtain a more quantitative understanding of this clustering more detailed analysis was undertaken. During the collection of these data, a super-resolution study of the BAM complex and its clustering was published (Gunasinghe et al., 2018). This is discussed in the introduction to this chapter alongside the limitations of the methodology those authors performed and potential advantages to be gained from use of the cryoSTORM method. Code to automate the clustering analysis of BAM super-resolution STORM data was made freely available by the authors and was utilised to form the basis of the analyses shown in Figure 5.6B-D. The degree of clustering was calculated using spatial analysis based on Ripley's K function. This function takes a dataset (consisting of N points) and calculates how much the distribution of neighbouring points (radiating outwards from each point, n) deviates from a spatially homogeneous (i.e. random) distribution. This is often normalised, so that the expected value for a random distribution scales linearly with the area of the radius giving the L-function, $L(r)$. As a random distribution would appear as a diagonal identity line when $L(r)$ is plotted against the radius of the circle, r , this can be hard to interpret. Commonly, $L(r)$ is normalised again so that the expected value for a random distribution is equal to zero (it has identity with the x-axis) by subtracting r , i.e. $L(r)-r$. This is also sometimes referred to as the H-function, $H(r)$ (Kiskowski et al., 2009).

The analysis of these data was challenging due to the blinks displaying a distribution of fluorescence intensities, not all of which were significantly above the background noise or background autofluorescence. The low intensity of a large proportion of the bursts means that filtering of the initial raw images by applying stricter thresholds for burst identification also discards a lot of fluorescence coming from within the bacteria, which may be 'true' bursts (that is, bursts occurring from a sfGFP molecule) of low intensity. The second challenge is the high enhancement of the cellular autofluorescence which results in large bursts even in wild-type cells containing no fluorescent protein and has been observed before in attempts to perform cryogenic super-resolution microscopy on cells (Carter et al., 2018). Finally, the ASIL itself can contribute large bursts and blinks due to defects in the crystal structure of the cubic zirconium which are optically active (Rabouw et al., 2016; Feng et al., 2018; Wang et al., 2019). These three components can degrade the

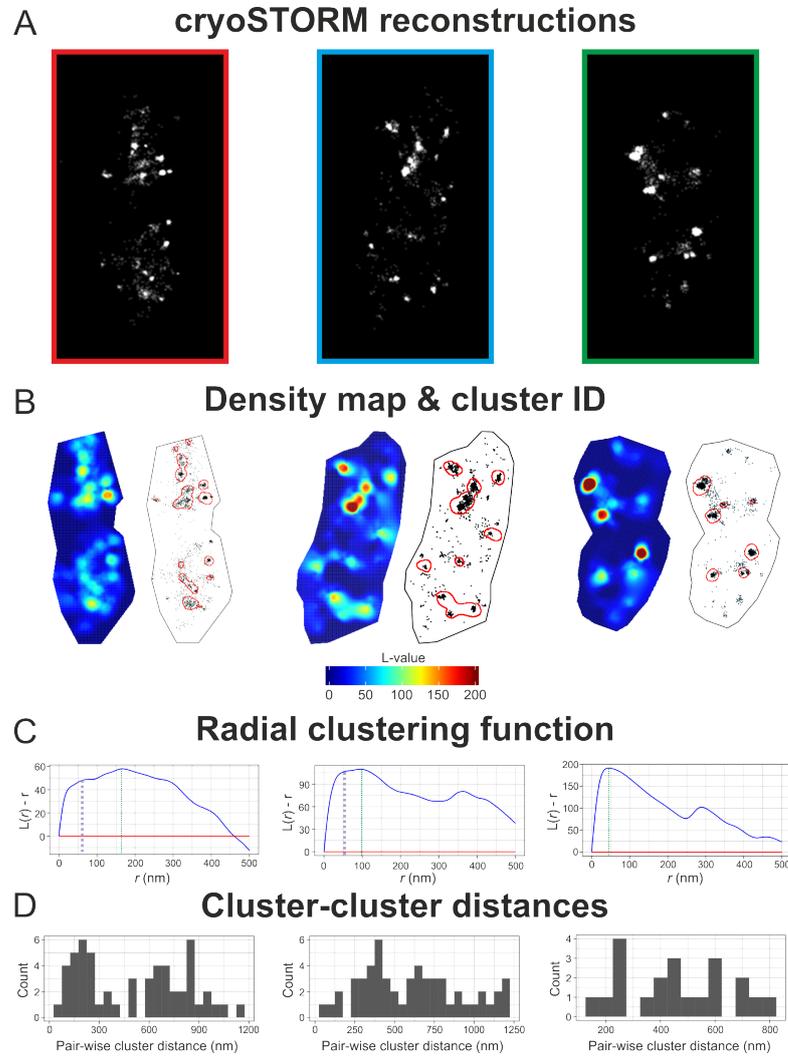


Fig. 5.6 Single-molecule localisation microscopy reconstructions and spatial analysis of BamA-sfGFP in *E. coli* imaged using cryoSTORM. (A) Average shifted histograms of single-molecule localisation performed in ThunderSTORM of BamA-sfGFP showing the location of BAM complexes in punctate clusters around *E. coli*. These are all taken from the same field of view as shown in Figure 5.5. (B-D) Spatial analysis of the single-molecule localisation data performed by the method described by Gunasinghe et al., 2018, using R and Matlab scripts described therein and accessed from <https://github.com/monashmicroimaging/autoclustr>. (B) The pattern of localisations was analysed using the normalised L-function for clustering. This outputs an L-value which is a measure of the degree of clustering where zero on the y-axis (red line) is the expected value for a completely homogeneous distribution of points and higher values equal points which are more clustered than would be expected by a random Poisson distribution (which is itself based on the number of points in the dataset, and the size of the bounding box – the ROI). Negative values indicate areas which are more disperse than a random distribution. These heat maps can be thought of as visualizing the ‘degree of clustering’ in a more quantitative manner. Alongside each heatmap is a visualization of the location of areas algorithmically predicted to form clusters. (C) Radial clustering function showing degree of clustering for a particular radial length scale (degree of clustering within a circle of radius, r) for increasing values of r . $L(r)-r$ is the normalised version of the L-function. Marked with purple dashed lines are the first local maximum identified (where different from the global maxima), and green dashed lines indicate the global maximum. Peaks on this plot indicate the size of circle (or rather radius) that maximises the clustering score. In other words, a measure of the ‘average’ cluster size. (D) Sum all of all pairwise distances between clusters identified in (B) with 50 nm bin size.

signal-to-noise, or more specifically, the ability to readily distinguish ‘true’ bursts coming from BamA from other sources of fluorescence when the emitter coupled to BamA (in this case sfGFP) is not intense enough, with the end result being a reduction in the confidence of any quantitative analysis.

Nonetheless, a few observations can be made. Figure 5.6B shows a heatmap indicating areas where localisations of BamA are particularly clustered. For the far left and far right example bacteria which appear to be dividing, the central region is devoid of any clusters or large numbers of localisations. This is in marked contrast to previous studies which pointed to BamA and other OMPs being inserted at the new pole / point of constriction (Ursell et al., 2012; Rassam et al., 2015). It is likely that the effect observed here is due to the maturation time of sfGFP. Older molecules of BamA have been reported to move towards the old pole of *E. coli* and some molecules can be retained in daughter cells for many generations, giving time for the chromophore to mature. The sparse numbers of localisations after filtering for background fluorescence made it challenging for the script to algorithmically select clusters. The cluster picking algorithm was also set so that objects smaller than 1000 nm² were not selected. Those regions identified as discrete clusters are indicated as red boxes in the bacteria to the right of the heatmaps in Figure 5.6B. The radial clustering function in Figure 5.6C indicates the size of a circle of radius, r , which exhibits ‘peak’ clustering. The broad distributions for the left and middle bacteria indicate that there is not a characteristic scale for the size of a cluster, but indicates a minimum cluster size of ~100–120 nm in diameter (according to the leftmost local maxima) with peak clustering (global maxima) at larger sizes. The rightmost bacteria shows a more homogeneous cluster size (with a peak cluster ~100 nm in diameter) but it should be noted there are also far fewer localisations in this region of interest. Figure 5.6D shows all the pairwise distances from clusters identified in Figure 5.6B. The left and middle bacteria show a bimodal distribution, which likely reflects the visual partitioning of the cell into two caused by the insertion of new molecules at the mid-cell / site of constriction as discussed above.

5.2.3 Imaging OmpA-sfGFP

To determine whether the clustering effect and pattern of BAM organisation was recapitulated by its substrates, similar experiments were undertaken with the most abundant OMP substrate for the BAM complex, OmpA. OmpA comprises a N-terminal β -barrel domain and periplasmic C-terminal peptidoglycan binding domain which is dispensable for cell viability (as is OmpA) *in vivo* and OmpA folding *in vitro* (Baba et al., 2006; Danoff and Fleming, 2011; Verhoeven et al., 2013). Applying a strategy similar to that for

BamA-sfGFP, a pZS31 plasmid with a chloramphenicol selection marker, harbouring an OmpA-sfGFP fusion where the C-terminal peptidoglycan binding domain was replaced with sfGFP, was created and showed no toxicity when expressed in JCM166 cells (Figure 5.3). cryoSTORM data were acquired and analysed as described for sfGFP-BamA in Section 5.2.2. Representative images from two bacteria are shown in Figure 5.7.

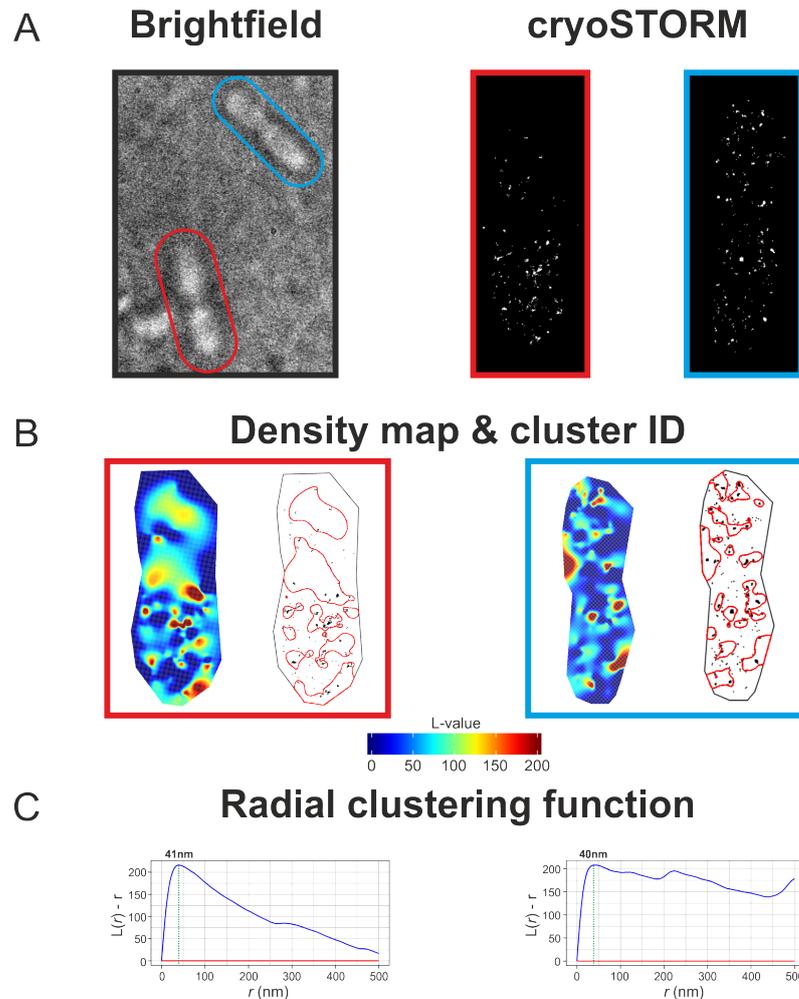


Fig. 5.7 Example images from a single field of view showing JCM166 *E. coli* expressing OmpA-sfGFP. (A, left) Brightfield image of the field of view from which these bacteria were collected. (A, right) Reconstruction of the cryoSTORM image set performed in ThunderSTORM. (B-C) Spatial analysis of the data in (A, right). Data were analysed as described for BamA in Section 5.2.2, and in Figure 5.6. (B) Heatmap generated from the set of localisations and the result of attempts to cluster the data into separate puncta. (C) Radial clustering function where the green dashed line represents the radius of a circle that corresponds to the global maximum of the cluster function.

Qualitatively analysing the pattern of localisations from the cryoSTORM analysis suggests that OmpA-sfGFP also forms discrete clusters but that these are fewer and smaller than observed for sfGFP-BamA (Figure 5.7A). As discussed for BamA, the small number of localisations after filtering makes quantitative analysis challenging and this is particularly the case for the OmpA-sfGFP data. Many of the bacteria collected showed low levels of

fluorescence and few blinks from OmpA-sfGFP and this can be observed strikingly in the upper half of the dividing bacteria in highlighted in red in Figure 5.7. The disparity between the two halves of the dividing bacteria is not understood but it may be that the lower half represents a repository cell as suggested by Rassam et al. (2015), where the load of “old” OMPs are asymmetrically divided between two cells with one receiving the old pole and older OMPs, and the daughter receiving the new pole and newer OMPs. If the maturation time of the chromophore in OmpA-sfGFP is slowed, this could also exhibit itself as a dark region. This sparse data prevents accurate assignment of clusters (Figure 5.7B) and also leads to poor results from radial clustering analysis (Figure 5.7C) where the data on the right show no strong spatial scale for clustering, despite a visual analysis of the data suggesting a trend towards smaller clusters. Regardless, both example bacteria show peak clustering at a diameter of ~80 nm, which agrees with the visual analysis that the average cluster size is smaller (compared to the 200–300 nm for BamA - see Figure 5.6C).

5.2.4 Two-colour imaging of BAM and OmpA

To understand how BamA and an OMP substrate, OmpA, might colocalise, a methodology for two-colour imaging needed to be developed. As discussed in Section 5.2.1, the use of red fluorescent proteins was not successful and so a different approach involving the self-labelling protein, HaloTag, was used. HaloTags are self-labelling tags which can form a covalent bond between themselves and chloroalkane ligands which can be derivatized with fluorescent dyes (Los et al., 2008). These have been shown to be functional in the oxidizing environment of the periplasm (Ke et al., 2016) and can be labelled with very bright, photostable, Janelia Fluor 549 dyes with a labelling efficiency of ~80 % (Lepore et al., 2019). JCM166 cells were co-transformed with pZS21 and pZS31 plasmids encoding an N-terminal fusion of BamA to HaloTag and OmpA-sfGFP, respectively. These were maintained by double selection for kanamycin and chloramphenicol and were well tolerated by the *E. coli* (see Figure 5.3). Cells were grown as described for single-colour imaging, but at an OD600 of 0.3, cells were inoculated with an aliquot of chloroalkane-derivatized Janelia Fluor 549 (HTL-JF549) from DMSO stocks at a final dye concentration of 2 μ M (see Materials & Methods 2.4.6). These bacteria were grown for a further 30 minutes at 37 °C to allow labelling to occur. Cells were then pelleted and washed thoroughly with phosphate buffered saline (PBS) and outgrown for another 15 minutes at 37 °C as a recovery step. The cells were then prepared as described for single-colour imaging until the preparation of the ASILs for plunge freezing. To accurately register two images to nanometer precision requires the use of fiducial markers. A number of possible fiducial markers were tried including 15 nm and 30 nm colloidal gold nanoparticles (MP Biomedicals, Cat#: 154012 &

154014) which reflect light in both green and red channels, as well as 20 nm FluoSpheres (Invitrogen, Cat#: F8787) which are latex particles coated with a variety of dyes across a range of excitation and emission wavelengths. ASILs were first coated with bacterial samples before being manually blotted, and fiducial markers applied, these too were manually blotted and a second round of application and blotting of bacterial samples was applied. Two-colour images were collected sequentially with 10,000 frames being collected for each colour.

Unfortunately, during testing and image analysis it was found that the gold beads were too dim for accurate registration of images and the FluoSpheres were too bright for the low dynamic range of the EMCCD camera and swamped the signal from the labelled bacteria. Shown in Figure 5.8 are the reconstructions from these two-colour cryoSTORM experiments. The inability to register the images automatically using fiducial markers means only an approximate low-resolution alignment can be performed by manually aligning on surface features of the ASIL and the fluorescence around the periphery of the bacteria. The results suggest poor correlation between features from OmpA and BamA, suggesting a low degree of colocalisation. Due to the low resolution of this method of aligning images, no quantitative analysis was performed.

5.3 Discussion

5.3.1 Limitations of the methodology

5.3.1.1 Limitations of a plasmid-expressed FP-fusion of OmpA

OmpA was chosen as a model substrate due to its position as the most abundant client of the BAM complex in *E. coli* (see Introduction Table 1.2). Native OmpA contains a ~170-residue N-terminal β -barrel domain and a C-terminal ~150-residue soluble periplasmic peptidoglycan-binding domain. In the construct used in this chapter, the C-terminal domain was replaced by a FP (sfGFP). How might this substitution change the biogenesis of OmpA? *In vitro* studies on the chaperoning of OmpA by Skp and SurA have shown that the periplasmic domain of OmpA is able to fold independently of the barrel domain (Danoff and Fleming, 2011; Zaccai et al., 2016), work on OmpA aggregation have shown that the soluble domain has only a minor self-chaperoning activity on OmpA (Danoff and Fleming, 2011) (although this may simply be a solubility effect, similar to the use of MBP as a solubility tag for protein expression), and *in vitro* folding experiments show a

BamA-HaloTag[JF549] / OmpA-sfGFP

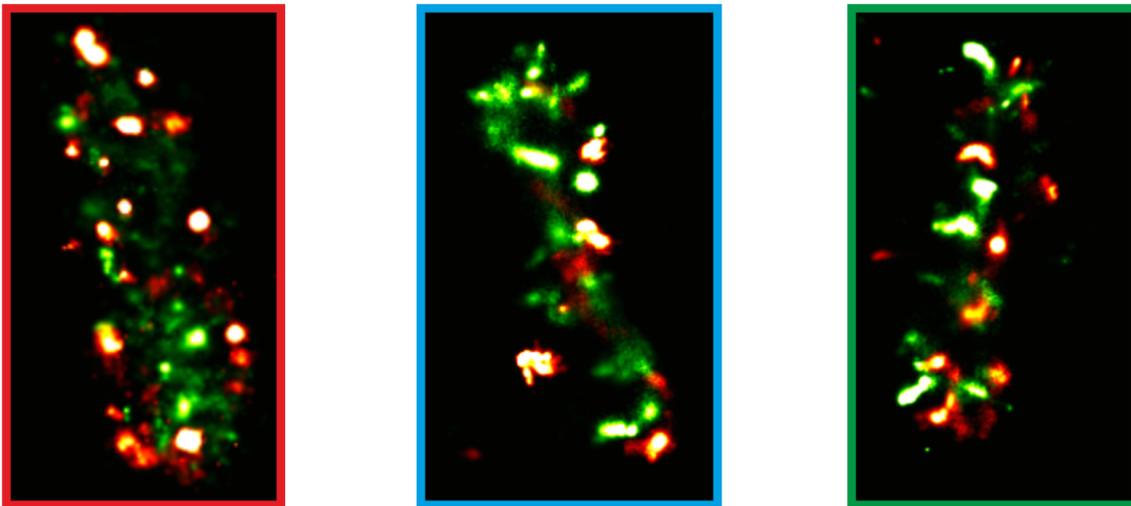


Fig. 5.8 Example two-colour images of HaloTag[JF549]-BamA and OmpA-sfGFP. JCM166 cells co-transformed with pZS21 plasmid containing BamA N-terminally fused to HaloTag, and pZS31 containing OmpA-sfGFP were labelled with Janelia Fluor 549 and imaged using cryoSTORM. Each colour was sequentially imaged for 10,000 frames before reconstructions were performed in ThunderSTORM. Images were manually aligned based on features on the surface of the ASIL and the shape of peripheral fluorescence in the bacteria. Reconstructions were plotted using a Gaussian reconstruction which essentially blends together localisations which are within a certain distance of each other (according to their measured uncertainty), 'smearing' the density across the width of a Gaussian distribution. This was done to help visualize areas where colocalisation might occur, despite the low resolution alignment.

negligible decrease in folding rate for full-length OmpA compared to transmembrane-only OmpA (Danoff and Fleming, 2011). Together, these results suggest that the intrinsic folding, chaperoning, and delivery of OmpA to BAM should be independent of its C-terminal domain. In this respect, the localisation of the OmpA fusion in the OM gives an indication as to the fate of a generic OMP after folding via BAM, as well as being able to assess the co-localisation of an OMP with the BAM complex after folding. However, these data do not necessarily give a physiologically relevant picture of the organisation of native OmpA (with its peptidoglycan binding domain intact).

5.3.1.2 Expression levels of proteins

In this study, OmpA and BamA were expressed from low copy number plasmids (~5 per cell) where the fusion protein was under the control of the $P_{\text{LtetO-1}}$ promoter. This promoter results in constitutive expression of BamA and OmpA due to the absence of the *tetR* gene in the JCM166 strain for repression. To manipulate essential genes often requires plasmid-based complementation as if the mutation (or in this case, protein fusion) introduced disrupts function, the bacteria will not grow and it can be difficult to determine whether the cause was a non-functional gene product, a toxic gene product, or simply a failed cloning reaction. The downside of their usage is that the protein cannot take advantage of any downstream or upstream code in the genome which modulates its expression in response to physiological cues. Furthermore, for some proteins, their exact location in the genome can influence the protein's localisation and expression. This was shown in *Caulobacter crescentus* for the OmpA homologue, OmpA2, where its localisation pattern was controlled by its position in the genome as translation took place at particular locations near the inner membrane, controlling the distance the protein diffused through the periplasm before integration into the OM (Ginez et al., 2014, 2019). Furthermore, for complicated protein complexes like BAM a higher level of coordination between the genome and translation machinery may be required to allow efficient secretion of the ~80 kDa protein and allow its proper assembly alongside the other subunits of BAM (BamBCDE) (Schwarz and Beck, 2019). For physiological studies of bacterial behaviour, these aspects should not be overlooked as there are many highly intertwined feedback mechanisms which govern the proper assembly of the OM and it is not always clear which aspects are important and which can be bypassed (Guo and Gross, 2014; Mitchell and Silhavy, 2019).

5.3.2 Localisation of OMPs

In this chapter, preliminary steps into developing a system to study the *in vivo* organisation of OMPs has been shown. The results agree well with the super-resolution studies on BamA performed by Gunasinghe et al. (2018) with similar clustering behaviour observed. The results shown in this chapter differ in the number of localisations and number of clusters seen. This could be a result of the use of plasmid-based expression systems to label BamA, as opposed to the use of antibodies to natively label *E. coli*. The observation of dark patches at the midcell and old poles which are attributed to new BamA/OmpA whose sfGFP fusion partner has not yet matured, also corroborates the findings of Rassam et al. (2015) who saw that new OMPs are inserted at these locations with old OMPs moving towards the old poles as the cells elongate and age. Unfortunately, the current data do not allow more detailed questions about the organisation of BamA within islands to be answered. This would require more intense fluorophores, reduced background autofluorescence, redder dyes, or all of the above. It is useful to note that Gunasinghe et al. (2018) used the far-red, and very photostable dye Alexa Fluor 647 for their imaging studies. Autofluorescence arises from NAD(P)H, flavins, and flavoproteins in the cell and have emission maxima around 440–460 nm, and 530 nm, respectively. This is also highly enhanced at cryogenic temperatures (König et al., 2014; Carter et al., 2018). The use of far-red dyes is well established for removing the vast majority of signal from background autofluorescence, and minimizing fluorescence due to optical components and minor contaminants (Zhang et al., 2018). Another option is to use spectral demixing to separate blinks which are due to autofluorescence (and generally have very broad emission profiles) versus those arising from fluorescent dyes or proteins (which have much narrower profiles) (Carter et al., 2018). Improvements in resolution would also be likely with better correction for the small amounts of drift that occur across the ~20 minute acquisition time. This would be possible with robust fiducial markers and could be achieved with the 20 nm fluospheres discussed in Section 5.2.4 coupled with a scientific CMOS (sCMOS) camera which would allow collection of fields of view containing higher dynamic ranges.

For two colour imaging, the lack of clear localisation between OmpA and BamA may reflect differences in their migration in the OM with OmpA - the most abundant OMP in *E. coli* (see Introduction Table 1.2) - possibly showing weaker clustering behaviour or less likelihood to form clusters with other proteins. This might make sense for a structural protein which is thought to give rigidity to the OM, to be better able to disperse throughout the cell than BAM. Furthermore, it would seem unfavourable for BAM to retain OMPs nearby as this has the potential to create a queue for biogenesis whereby newly inserted

OMPs crowd the local environment around BamA, preventing or slowing the insertion of new proteins which can be toxic for bacteria (Mitchell and Silhavy, 2019).

5.3.3 Future directions

5.3.3.1 Localisation limits of the technique

Although a localisation uncertainty as low as 3–5 nm could be achieved in these studies, the linker of the FP fusions adds an additional ~3–5 nm of uncertainty to the location of BamA or OmpA (see Figure 5.2), and low levels of uncorrected drift could also add 1–2 nm uncertainty. As shown in Figure 5.1, the minimum possible BAM-BAM distance is 12–16 nm which is very close to the range of 7–12 nm achieved from summing up possible sources of uncertainty. Dependent on the exact organisation of BAM islands, this may be sufficient to get information on their lateral organisation (i.e. how molecules of BAM are organized and packed relative to each other) in future studies. However, a number of alternative approaches exist which could reduce linker-length uncertainty and remove problems arising from FP maturation. Here, the limitations benefits of these possible alternative approaches will be discussed. FIAsh and ReAsH are fluorogenic dyes that are quenched in their unbound state by arsenic-groups but can bind specifically to tetra-cysteine (CCXXCC) motifs via these arsenic groups, resulting in a massive increase in fluorescence intensity (20–100-fold dependent on the exact motif and cellular environment) (Griffin et al., 1998). These dyes have been shown to be compatible with single-molecule localisation microscopy methods (Lelek et al., 2012). However, they have some drawbacks in that there can be a large diffuse background fluorescence (Stroffekova et al., 2001; Fessenden and Mahalingam, 2013), cytotoxicity (Langhorst et al., 2006), and that their application in oxidising environments (such as the periplasm) requires the use of harsh (and highly toxic) cell-permeable reductants such as British anti-Lewisite (BAL) and ethanediol (EDT) (Hoffmann et al., 2010) (which we were unable to use in the microscopy facility at RCaH). An alternative approach may be the use of bis-boronic acid-based dyes which behave similarly to FIAsh/ReAsH, but instead target tetra-serine (SSPGSS) motifs (Halo et al., 2009). This approach reduces the need for reducing agents but may have some issues with the stability of the boronate ester formed upon binding (António et al., 2019). The Sfp system uses a modified version of the Sfp phosphopantetheinyl transferase enzyme from *Bacillus subtilis* to transfer a fluorescent dye-coenzyme a (CoA) conjugate onto a serine residue in the small 11-residue motif (DSLEFIASKLA) known as a ybbR tag within 15 minutes (Yin et al., 2006). A longer 20-residue version of this tag was installed in a loop of LamB and the system used to follow the appearance of this OMP in the OM

of *E. coli* (Ursell et al., 2012). This system allows relatively non-perturbing installation of a labelling tag due to the degeneracy of OMP loops, low background fluorescence due to the tags accessibility on the extracellular side of the OM, and high labelling efficiency (>80 %) (Yin et al., 2005). Unfortunately, this system is no longer available commercially meaning the dye-conjugated CoA and the Sfp transferase must be synthesized and produced in house. However, the chemical reagents and plasmids for this are readily available, and their synthesis and expression relatively trivial (Worthington and Burkart, 2006; Yin et al., 2006). Unnatural amino acid (UAA) incorporation is another option for labelling, and has been used previously to install a benzophenone-based crosslinker into BamA (Hagan et al., 2015; Wzorek et al., 2017). However, incorporation of unnatural amino acids using amber stop codon suppression can cause premature termination of translation (and therefore truncation of the protein) at the site of incorporation due to competition between the UAA tRNA and amber-recognizing release factor protein 1 (RF1) (Wals and Ova, 2014). In the aforementioned studies on BamA, wild-type BamA was still expressed from the chromosome which would help ameliorate toxicity from truncated BamA in the OM. In a depletion strain, care would need to be taken to ensure that truncation products were not interfering with the normal function of the BAM complex. However, this could be achieved by labelling near the N-terminus of BamA which would reduce any truncation product to the signal sequence and (after cleavage by signal peptidase I) a small peptide.

5.3.3.2 Genomic manipulations

As mentioned when discussing the limitation of the methodology, there are a number of drawbacks to plasmid-mediated expression for physiological studies of bacteria. A way to overcome this would be use the JCM166 depletion strain for initially testing and validating the viability of BamA-fusions (which has been shown in this chapter), and then creating a new strain where the native BamA locus is modified to add the fusion protein to the gene product. Although this is slower and more challenging than plasmid-based manipulation, it would allow confidence that the organisation of BamA observed would more closely reflect the true physiological state of the bacteria. Furthermore, a wealth of tools now exist for manipulation of *E. coli* genomes (Jiang et al., 2015; Wiles et al., 2018).

5.3.3.3 Organisation of the whole biogenesis machinery

A large number of studies have been conducted on the organisation of OMPs and this chapter represents the third that focuses on BamA (Rassam et al., 2015; Gunasinghe

et al., 2018). However, no studies have yet interrogated the relative organisation of OMP chaperones and other proteins associated with the BAM machinery such as SurA, Skp, FkpA, DegP, RcsF, and BepA. Direct interactions of BamA between SurA, RcsF, and BepA have been reported (Wang et al., 2016; Konovalova et al., 2014; Cho et al., 2014; Daimon et al., 2017) but many questions remain unanswered. Are these proteins enriched at BAM clusters, or at specific sites of insertion, or do they distribute randomly throughout the periplasm? Super-resolution studies of these systems could also be combined with single-molecule tracking in live cells to discriminate populations of freely diffusing and bound proteins. Furthermore, there is growing evidence for a super-complex between BAM in the OM and SecYEG in the inner membrane (Wang et al., 2016; Alvira et al., 2019; Carlson et al., 2019). Super-resolution microscopy could help resolve its existence by looking for colocalisation of these complexes, particularly the absence and presence of active OMP synthesis.

5.4 Conclusion

This chapter presents preliminary data which builds on existing fluorescence microscopy studies of OMP organisation to try and develop a new strategy and workflow that can theoretically allow the resolution of individual molecules of BAM and other larger OMPs to be resolved inside a cluster. While much work remains to be done, the potential for pushing the barrier in our understanding of OMP biogenesis is huge. The theoretical resolution afforded by cryoSTORM is to the level of individual molecules, and the frozen-hydrated nature of the sample preparation inherently lends itself to applications in correlated cryo-EM and cryo-STORM. This would allow not only the lateral organisation of all molecules in OMP biogenesis to be discerned, but potentially the structure of BAM and its supercomplexes to be revealed by cryo-electron tomography. Many hurdles remain, most importantly overcoming background autofluorescence (or accurately filtering it), finding fluorescent labels or strategies that are both very bright with well-separated blinks at 77 K but also non-perturbing to BAM function, incorporating chromosomal fusions of BamA to ensure native assembly, and finding good two-colour combinations for labelling accessory proteins alongside BamA. Despite these challenges the door to super-resolution cryo-fluorescence has been opened and its possibilities are irresistible.

Chapter 6

General Discussion

6.1 Overall conclusion of results

Outer membrane protein biogenesis is challenging to study *in vitro* and *in vivo* due to the transient and dynamic nature of the protein-protein interactions made between unfolded OMPs and chaperones, membranes, and the BAM machinery; the difficulty in measuring localised changes in the membrane around the BAM complex; and the nanoscopic size of these components in the crowded environment of the outer membrane *in vivo*. This thesis has aimed to discover new insights into this process through the development and application of new methodologies (Chapter 3, tag-transfer crosslinking; Chapter 5, cryoSTORM super-resolution microscopy) and the application of existing methods in new ways (Chapter 4, the role of lipid order in BAM catalysis). Together, this expands our understanding of OMP binding and chaperoning by SurA and Skp, the relative contribution of the subunits of BAM to its catalytic activity and ability to disorder lipids, and builds towards acquiring a nanoscopic picture of the relative organisation of OMP substrates with BAM and the shape of BAM / OMP 'islands'.

The data presented in Chapter 3 showed the development of a new method for cross-linking mass-spectrometry using tag-transfer crosslinkers. This method involves the novel combination of a diazirine photoactivatable moiety with a methanethiosulfonate (MTS) conjugation group for site-specific attachment onto a cysteine residue installed into a bait protein using site-directed mutagenesis. A new crosslinking UV LED lamp design allows the completion of crosslinking reactions in ~7 s (Figure 3.15) and the unbiased reactivity of the diazirine group can crosslink to any residue side chain (for example, Figure 3.28) or even the protein backbone, allowing a true snapshot of the local environment to be taken.

This method was validated on the OMP chaperone-substrate pair, OmpA and Skp. The use of a transferred tag at a specific site of interaction on the target protein (in this case, from OmpA to Skp) has two benefits: firstly, it can be used as an enrichment tag using Thiopropyl Sepharose resin that will bind specifically to thiol group (of which the tag bears one) and secondly, it can be chemically ‘capped’ by alkylation creating a stable mass adduct at that position which can be mapped to residue-level accuracy using MS/MS coupled with a search for a +145.06 Da or +261.04 Da (dependent on crosslinker used) modification using standard mass spectrometry proteomics software – aiding the simplicity and automation of the assignment process. The small size of both tag-transfer reagents (MTS-diazirine and MTS-TFMD) did not significantly affect the folding of OmpA after installation on cys mutants meaning that experiments performed with this tag installed represent OmpA in a folding-competent state (Figure 3.10). The results of the crosslinking experiments showed a massive improvement on previous approaches using benzophenones with 35 unique modified positions in Skp identified in MS/MS experiments using tag-transfer with resin-based enrichment of peptides (Figure 3.21), compared to none which could be confidently assigned using benzophenone-based crosslinking and in-gel digestion. This is remarkable due to the dynamic nature of the complex where OmpA is tumbling dynamically inside the cavity of Skp (Burmann et al., 2013), with only a very short residence time at any one residue and shows the power of this approach for studying these kinds of dynamic interactions. Furthermore, using standard mercury-xenon based UV lamps took over 20 min for the reaction to reach completion while the UV LED lamp took <10 s, a more than 100X rate improvement (Figure 3.15). In experiments carried out by Dr Martin Walko, University of Leeds (not shown in this thesis – see Horne et al. (2018)), this method was also validated for a tight and stable protein-protein interaction where side chains closely interdigitate using the potential anti-cancer target of the regulatory pro/anti-apoptotic pair Mc11 and Bid (Lessene et al., 2008; Liu et al., 2010). Maintaining such a tight interaction would be incompatible with large and bulky crosslinkers such as benzophenones, but the small MTS-diazirine tag-transfer reagent showed minimal perturbation to binding with either little change in the nanomolar affinity of BID for Mc11 or retained low μ M affinity, dependent on the exact location of installation (see Figure 3.27 for comparison with sizes of other commercially available crosslinking reagents). Unlike the dynamic Skp-OmpA interaction where absolute crosslinking yields were relatively low (due to crosslinking to accessible solvent) and the size of the complex (~70 kDa) meant that in-gel analysis (Method 1) suffered from poor signal-to-noise., the small complex size (~18 kDa) and high yields of this tight Mc11-Bid interaction allowed robust quantitative proteomics to be performed on each crosslinked site that could be mapped back on to the known NMR structure of the complex and showed a distance-dependent intensity drop-off based on the site of tag-transfer reagent installation. This shows that as well as providing a new method

for studying dynamic complexes, this method could also be used to acquire restraints for protein-complex modelling and so is also applicable to structural mass spectrometry on ‘more static’ protein-protein interactions. The development of this methodology led to its publication and the commercialisation of the tag-transfer crosslinkers (Horne et al., 2018).

Once this method had been developed and validated in these well-studied protein-protein interactions, it was applied to more challenging systems. SurA the other major periplasmic chaperone for OMPs and perhaps has a more significant role in their delivery in a folding competent state to the BAM complex (Sklar et al., 2007b; Wang et al., 2016; Soltes et al., 2016). Applying the tag-transfer method from OmpA to SurA it could be shown that the binding activity of SurA resides almost exclusively in the core N- and C-terminal domains, raising the question as to what the roles of the less conserved (Humes et al., 2019) P1 and P2 domains are (Figure 3.26). Finally, this approach was used to try and capture a folding intermediate of OmpA as it was passed from SurA through the BAM complex and the interactions from OmpA to these partners during folding. This interaction is both transient and dynamic meaning that a method was needed that maximises reactive residues so as not to bias the observed transit path (diazirine groups react with all 20 amino acids), one that can enrich crosslinked peptides due to the low abundance of each crosslinked residue (assuming no long-lived intermediate states exist, any one position on OmpA will likely pass by many residues of SurA and BAM during its path to a membrane-embedded folded state), and an ability to complete the crosslinking reaction on a timescale much faster than OMP folding (minutes). The tag-transfer crosslinking method with the UV LED lamp fulfils those requirements and could show that position W7 at the bottom of the first (N-terminal) β -strand of OmpA makes crosslinks with the POTRA1, 4, and 5 domains of BamA as well as with the N-terminus and P2 of SurA (Table 3.5). Although it can’t be stated with confidence whether these crosslinks were formed as part of a ternary complex (SurA – OmpA – BAM) or a binary complex (SurA – OmpA) or (OmpA – BAM), when compared to the SurA-OmpA data collected in isolation, they suggested a greater recruitment of SurA P2 during OMP folding. This pattern of crosslinks from OmpA suggest a possible route from POTRA1, via BamD near the interface with BamA-POTRA5, onto β 1 of BamA, with the crosslinks at POTRA4 formed last as the final β -strand is appended to the nascent barrel (i.e. furthest from the β -signal containing C-terminal strand) (see Chapter 3 Figure 3.32 and Figure 3.33).

In Chapter 4 the hypothesis that BAM functions by disordering lipids in the membrane was tested by using a number of techniques, as well as the relative importance of BAM subunits, SurA, and the phase of lipid bilayer in folding the transmembrane domain of OmpA, tOmpA. To measure the kinetics of tOmpA folding into DMPC liposomes a gel-

based assay was used. This assay exploits the fact that the folded state of most OMPs is SDS-stable and exhibits anomalous migration on an SDS-PAGE gel when ran without prior heating (while their unfolded state migrates according to their molecular weight) and this folded state can be unfolded by boiling in the presence of SDS before loading (the so-called ‘heat modifiability’ of OMPs). It could be shown that the full BAM complex is a much better catalyst for OMP folding (as measured by t_{50}) than BamA alone when below or at the transition temperature of DMPC, and slightly better when above (see Figure 4.11 and Table 4.2). While the BAM complex could accelerate the formation of folded tOmpA almost 16X over uncatalyzed folding at the transition temperature (24 °C), folding via BamA as only marginally faster (at 1.5 fold the uncatalyzed t_{50}) (see Figure 4.12 and Table 4.3). In the transition state lipids are at a high level of ‘disorder’ with more packing defects present (that are proposed to nucleate spontaneous OMP folding) (Nielsen et al., 2000a,b; Enders et al., 2004; Danoff and Fleming, 2015a) and so a protein folding enzyme which works by creating this kind of disorder would show less apparent activity. This result implies two possibilities: 1) that isolated BamA functions primarily by disordering lipids while the BAM complex has some additional capacity (unrelated to lipid order) to enhance folding rate, or 2) that the subunits of BAM (BamBCDE) further enhance the ability of BamA to create membrane defects that support accelerated OMP folding (or perhaps both). SurA was shown to have a relatively minor effect under most conditions, increasing the yield by <6 % in all but one instance, and changing the rate of folding by less than 1.5X in all but two conditions (change in rate: Table 4.6, change in yield: Table 4.7). The exceptions appeared to be for BAM-catalysed folding in the liquid disordered phase (30 °C) where the presence of SurA accelerated folding by ~4X and uncatalyzed folding in the same phase and at the transition temperature (24 °C) where SurA caused the rate to fall by 2.7X and 1.7X, respectively. While the yield was less affected under most conditions, one remarkable exception was uncatalyzed folding in the liquid disordered phase at 30 °C where the yield fell by over 50 %. These results for SurA, in light of the known interaction of BAM and SurA *in vivo* (Sklar et al., 2007b; Wang et al., 2016), suggest that SurA may only bind to or select certain conformations of the BAM complex as shown by the minor or negligible change in rate and yield caused by SurA in the presence of apo-BamA or the BAM complex below or at the transition temperature. Furthermore, this data suggested that the exact packing of the lipids themselves could bias the conformations of BAM. To test these hypotheses, the packing of lipids was assessed directly as the dynamics and conformational ensemble of the BAM complex was measured at a single-molecule level using FRET (smFRET). The lipid order probe laurdan was used to calculate the transition temperature of DMPC liposomes in an empty state, or reconstituted with BamA or the BAM complex, and found that the presence of the BAM complex causes a broadening of the phase transition region as well as a 2 °C fall in the transition temperature implying a

stabilisation of the liquid phase by BAM (Figure 4.17 and Figure 4.18). smFRET studies showed that two populations of BAM complex exist in solution, corresponding to the predicted FRET efficiencies of the open and closed states and that these do not appear to interconvert on a 100s of μ s to 100s of ms timescale (Figure 4.21 and Figure 4.22). The data suggested a slow interconversion (>seconds to minutes on average), but doesn't rule out the possibility that this could be modulated by the binding of substrates or SurA. Furthermore, although the change is small and would need to be confirmed through further studies, it appears that the conformational ensemble of BAM, as measured through RASP analysis of smFRET data, is more restricted in the liquid crystalline phase (30 C) as compared to in the gel phase (20 °C), which could offer a reason for the ability of SurA to more significantly affect the rate of folding at this temperature (if it does indeed select particular conformations of BAM).

Chapter 5 outlined the current progress in studying the nanoscale organisation of the BAM complex by using a novel method in super-resolution microscopy, cryoSTORM. By plunge-freezing samples of *E. coli* expressing BamA and/or OmpA fused with fluorescent proteins or the self-labelling HaloTag protein, these proteins could be visualised in their assembled state on the surface a bacteria trapped in a frozen-hydrated state with sub 5 nm localisation precision. This not only showed the proof-of-principle for future imaging studies on these proteins but also showed the arrangement of molecules of BAM into discrete 'islands' spotted throughout the cell surface (Figure 5.6) as well as smaller islands formed by OmpA (Figure 5.7) showing that OMPs are prone to cluster together in islands much smaller than suggested by diffraction-limited microscopy studies (Rassam et al., 2015) and recapitulating data which was published during this study (Gunasinghe et al., 2018). Although a number of hurdles remain for improving the quality of the signal over the background autofluorescence (which is enhanced at cryogenic temperatures), this establishes a pipeline for imaging and data analysis to investigate the organisation of individual OMPs and members of the OMP biogenesis machinery (e.g. SurA, Skp, BepA, DegP, SecYEG). The appearance of dark patches at the mid-cell (i.e. the new pole) of dividing bacteria implicated it as an insertion site for new OMPs (due to the slow maturation of the fluorescent protein chromophore) in agreement with previous diffraction-limited studies (Rassam et al., 2015). Initial two-colour studies of OmpA and BamA could not be accurately registered due to challenges of dynamic range for fiducials at cryogenic temperatures (gold nanoparticles are too dim behind the bright blinks of labelled proteins, while fluorescent beads are too bright) but manual alignment suggested a relatively low degree of co-localisation for these proteins (Figure 5.8). Further studies with better alignment will be needed to confirm this result, but if correct, implicates a

mechanism for rapid clearance (that is lateral movement rather than degradation) of newly formed OMPs from the vicinity of BAM islands.

6.2 Where do we stand with models of the mechanism of BAM's catalytic activity?

In the introduction, three possible models of BAM's mechanism were presented:

- The BamA-assisted model (Figure 1.12A) where BamA/BAM plays a passive role in the insertion process and only acts to create a local destabilised region of lipid which lowers the activation energy of OMP folding. The POTRA domains and BamD may still be involved in the initial recognition of an OMP β -signal in order to recruit the OMP substrate to the region of disorder, but there would be no protein-protein interactions with the β -barrel domain of BamA.
- The BamA-budding model (Figure 1.12B) would involve initial recognition at the POTRA domains and at BamD but also specific binding of the β -signal of an OMP to $\beta 1$ of BamA. The barrel domain of the substrate OMP would then be folded by expanding outwards laterally into the membrane, with β -strands being added pairwise, N-terminally to the nascent barrel, after threading through the lumen of BamA. This would temporarily create a 'super-barrel' / hybrid-barrel in the membrane that would eventually bud-off once the final strands were added, as the β -signal would have a greater affinity for its own $\beta 1$ -strand than $\beta 1$ of BamA (although this remains an untested assumption).
- The BamA $\beta 1$ -elongation model (Figure 1.12C) is a hybrid of the above two. Again, this is compatible with initial recognition in the POTRA domains or BamD, but the POTRA domains here play a greater role in providing a 'safe space' for OMP 'pre-folding'. $\beta 1$ of BamA recognises the β -signal of an incoming OMP substrate but instead of threading β -hairpins through the BamA lumen and into the membrane, β -hairpins are appended to the nascent barrel *in the periplasmic space*. Either upon barrel closure, or upon recognition of a second β -signal or other cue, the formed barrel (or partially formed barrel) is inserted into the membrane through the destabilised lipids around the $\beta 1$ - $\beta 16$ seam. This is achieved either passively (by affinity of the OMP for the lipids) or actively (by a 'swinging' of $\beta 1$ upwards / inwards).

The data presented in this thesis helps to understand the relative importance of each of these possible mechanisms for the BAM complex of *E. coli*. Evidence for direct binding of the β -signal / C-terminal substrate β -strand to β 1 of BamA has now been shown *in vivo* for Por1 on Sam50 in mitochondria (*S. cerevisiae*) and for EspP on BamA in bacteria (*E. coli*) so it seems likely that folding via BAM involves this direct pairing with the BamA barrel. SurA has also been shown to bind directly to BamA (Sklar et al., 2007b; Wang et al., 2016) suggesting delivery occurs directly to the POTRA domains, and although deletion of POTRA1 or 2 causes only relatively minor defects in OMP assembly (perhaps by effecting a SurA binding site?), defects are much more severe for deletion of POTRA3 and 4 and lethal for deletion of POTRA5 (Kim et al., 2007; Gatzeva-Topalova et al., 2008). The essentiality of BamD in *E. coli* and the direct interaction observed *in vitro* with OmpA (Hagan et al., 2013), peptides derived from the *Haemophilus influenzae* autotransporter Hia (Albrecht and Zeth, 2011), and peptides derived from BamA (as a substrate) (Hagan et al., 2015), as well as *in vivo* crosslinking from BamD to LptD (Lee et al., 2018) provide strong evidence that BamD is also involved in substrate recognition. Together, this would appear to rule out a pure BamA-assisted model for OMP folding, at least in a general sense. The data presented in Chapter 4 suggest that in isolation BamA is unable to significantly accelerate folding when large numbers of membrane defects are already present (at the transition temperature), suggesting that for BamA alone, or for the ancestral BAM complex (which would have simply been BamA and possibly a single POTRA domain), lipid destabilisation causing lowering of the activation energy barrier to insertion may have been the primary mechanism for catalysis. Furthermore, recent genetic evidence pointed to a direct link between OMP biogenesis and membrane fluidity in *E. coli* (Storek et al., 2018a). The full BAM complex, however, may take a two-pronged approach to improve folding: enhanced lipid destabilisation and greater co-ordination in β -1 templating.

Although two conformational states for BamA alone cannot be ruled out, all structures of isolated BamA or homologues of BamA show only the closed state (16 individual crystal structures, see Figure 1.9). The smFRET data in Chapter 4 show that the BAM complex is in conformational equilibrium between two distinct states that match the open and closed structures. These states may exhibit a conformational cycling that promotes the enhanced lipid destabilisation observed by laurdan fluorescence as well as an additional benefit unrelated (or indirectly related) to lipid packing but which could be conformationally selected for by binding of an OMP or SurA. A conformational selection model is feasible as BamD, which can bind to OMP substrates, appears to play a role in altering the conformation of BamA by transmitting an allosteric signal via BamA POTRA5 (Ricci et al., 2012; Rigel et al., 2013; McCabe et al., 2017). Crosslinking experiments to from OmpA

to BamA presented in Chapter 3 support this possibility by showing an interaction from OmpA β 1 (the final strand to be assembled in these models, not the β -signal) to POTRA1 and POTRA5. This suggests that substrates pass from an initial site near the bottom of the BamA POTRA domains, via BamD, to the vicinity of this site (the small number of crosslinks observed and challenges with enrichment of crosslinked protein means that *direct* binding to BamD cannot be ruled out). Furthering the idea of conformational selection, smFRET studies on the BAM complex in Chapter 4 in DMPC bilayers at different lipid phases suggested that the conformational ensemble is reactive to the local lipid environment with the ensemble broadening as the temperature is lowered from liquid disordered (30 °C), to transition (24 °C), and the gel phase (20 °C). It was initially surprising that SurA caused such a minor effect on both the rate and yield of OMP folding, considering the proven direct interaction mentioned earlier. However, in the liquid disordered phase (30 °C) the presence of SurA conferred a 4X enhancement in folding speed despite having little to no effect under other conditions of BamA or BAM-catalysed folding. This would make sense, however, in a model where the conformational ensemble of BAM was better primed to accept SurA (conformational selection). The crosslinking in Chapter 4 was performed on a BAM-SurA-OmpA mixture with BAM embedded in *E. coli* polar lipids, but nonetheless, the crosslinking was performed at room temperature where *E. coli* polar lipids would also be well above their transition temperature. In these experiments, a greater degree of crosslinking to SurA P2 was observed compared to experiments with OmpA-SurA alone. This could reflect a state of SurA bound to BAM where the P2 domain is not actively involved in OMP binding (this is reserved for the N- and C-domains) but gets crosslinked as a ‘bystander’ – for example if the BAM-bound form of SurA inserts the P2 domain into the ‘hole’ formed by the ring of periplasmic POTRA domains. This seems feasible as examples of closed cages to exclude volume and reduce the entropic penalty for folding are found in nature such as the GroEL/ES cage (Hayer-Hartl et al., 2016) and SurA plugging the bottom may allow a ‘safe space’ for the OMP to ‘pre-fold’ as suggested by the BamA β 1-elongation model (see Introduction Figure 1.12C, Chapter 3 Figure 3.32) (Schiffirin et al., 2017a; Doyle and Bernstein, 2019). Although in the example of GroEL/ES binding and release of substrate is driven by ATP hydrolysis, in the ATP-free environment of the periplasm this substrate release may be driven by the very high free energy change of OMP folding (\sim -10 to \sim -140 kJ.mol⁻¹) acting as a thermodynamic “sink” driving OMP biogenesis along the biogenesis pathway (Hong and Tamm, 2004; Huysmans et al., 2010; Moon et al., 2013; Fleming, 2015). An enhanced co-ordination of β 1-templating could also explain the incidental observation that minor bands of OmpA dimers are present in the uncatalysed folding and BamA-catalysed folding reactions, but none are observed in the presence of the full BAM complex (see Appendix B). Dimers and higher order oligomers of tOmpA have been observed before (Danoff and Fleming, 2017) and likely arise from

strand-swapping of OMPs during their folding or rearrangement on the membrane. Initial recognition of tOmpA on $\beta 1$ of BamA is favourable in this regard as it helps to ensure the N- and C-terminal β -strands of OMPs will only pair with themselves.

Finally, the crosslinking experiments in Chapter 3 also show crosslinks observed to POTRA4. The crosslink to E294 (labelled x in Figure 3.30) doesn't seem to satisfy any model of OMP folding, but it may represent a crosslink formed shortly after folding. W7 on OmpA is located at the bottom of the N-terminal β -strand of OmpA and although the aromatic ring of the tryptophan 'snorkels' upwards into the membrane in the crystal structure of OmpA (1BXW), it sits at the bottom of lipid interface and MTS-diazirine-conjugated cysteine would be likely to slightly disfavour a membrane embedded state due to the electron rich disulfide bond of formed between MTS-diazirine and W7C, as well as the electron rich diazirine group itself, and so would be expected to point towards the aqueous environment. E294 is located around the back of the BamA barrel so OmpA would need to diffuse around it to explain this crosslink. However, the periplasmic domain of OmpA and the presence of BamB immediately around the opening of $\beta 1$ - $\beta 16$ seam of BamA may occlude immediate free diffusion in that direction but if we consider a BAM dimer model where two BAMs connect via their BamB subunit (see Chapter 5 Figure 5.1) we see that a newly released OmpA (with its periplasmic domain) could be shuttled toward this exact location. Two more crosslinks were observed to POTRA4 (labelled a and b in Figure 3.30) which are not compatible with labelling from above (i.e. by a folded OmpA diffusing) but appear to be crosslinked from within the cradle formed by the POTRA domains. If these were crosslinked at a late stage in BamA $\beta 1$ -templating they are only compatible with a model where the substrate OMP is bound to the closed state of BAM and the growing OMP curves into the lumen of BamA (see Figure 3.32 BAM (closed) curved elongation). Alternatively, the POTRA chain may sit 'in wait' in the POTRA cradle after being spooled through from SurA via POTRA1, BamD, and POTRA5 and at this point it could have crosslinked to these residues. This begs the question, however, of what would drive the movement of a polypeptide chain along this route and into the POTRA cradle. The thermodynamic sink model would mean that the driving force is the formation of new hydrogen bonds 'zipping' along exposed β -strands and pulling in the chain in a C- to N-terminal fashion. If this is correct, there shouldn't be any 'slack' that would sit in the POTRA cradle as the unfolded chain would only be pulled through when a section of chain was ready to template a new β -strand. If, on the other hand, folding was aborted due to misfolding of the OMP, or mis-timing of a conformational cycle, then it is possible that a chain could be dropped back down to re-attempt a productive cycle of folding. The fact that stalled folding intermediates on the BAM complex (e.g. LptD_{Y721D}, LptD₄₂₁₃, LptE₆, Hbp_{110C/348C}) (Sauri et al., 2009; Chimalakonda et al., 2011; Lee et al., 2016, 2018) are not

lethal conditions and can be recovered from implies that mechanisms exist for correcting misfolding. A number of proteases have been identified which degrade misfolded OMPs at different stages in this assembly process: the periplasmic protein DegP degrades substrates which misfold before reaching BAM, YcaL is an OM lipoprotein that degrades OMPs stalled after binding to BAM but possible before initiating folding, and BepA degrades substrates which have misfolded at a late stage of barrel formation (Soltes et al., 2017). YcaL remains poorly studied but the structure of BepA is known and crosslinking studies have shown that it interacts with BamA, BamC, and BamD, placing it at the bottom of the periplasmic cradle and midway along the route proposed by the crosslinking data shown in Chapter 3 and discussed above (Daimon et al., 2017). As BepA has been shown to degrade a late-stage folding intermediate of LptD (Soltes et al., 2017), this supports a model where mostly folded intermediates remain relatively exposed to the periplasmic space (albeit in the protected environment of the BAM complex). Single-molecule studies on Skp and SurA have shown that Skp (but not SurA) is able to rescue aggregated forms of OmpC and bring them back onto the folding pathway (Li et al., 2018). While there is no evidence to suggest that Skp plays a role at such late stages of assembly, the possibility remains that other BAM subunits, or known (or as yet undiscovered) BAM-associated proteins could also rescue late-stage misfolding intermediates rather than simply degrade them.

Chapter 5 showed that BAM forms discrete puncta - so called 'islands' - in the outer membrane of *E. coli*, agreeing with previous studies (Rassam et al., 2015; Gunasinghe et al., 2018). While it remains unclear what the physiological benefit of this clustering of BAM would be, some authors have suggested that it may form a protected environment for the formation of OMPs which rely on oligomerisation (such as OmpC and OmpF) (Gunasinghe et al., 2018). However, as shown in Figure 5.1, the dimerization across BamB subunits results in an orientation where the $\beta 1$ - $\beta 16$ seam of each BamA points in opposite directions making it unclear how assembly could be coordinated. The lack of co-localisation of OmpA and BamA observed in Chapter 5 would need to be further verified, but as these represent a population of slightly older OMP molecules (on the basis that the mid-cell / new pole was dark due to new OMP-FP fusion whose chromophores hadn't yet matured) it could be possible that clustering of BAM molecules would help to 'clear' new OMPs. This could happen due to unfavourable hydrophobic mismatch in the narrowed thickness surrounding the BamA seam compared to the hydrophobic thickness of most OMP β -barrels and the bulk OM. Although thinner membranes would present lower activation energies for initial insertion, the fully assembled barrel would move away to an area of greater hydrophobic thickness that was more energetically favourable (Katira et al., 2016).

6.3 Future studies

With so much now understood about OMP biogenesis *in vitro* and *in vivo*, there seems to be a clear path toward finally fully understanding the key aspects of OMP biogenesis through BAM. Future studies should aim to look at:

- What is the affinity of β -signal strands of OMPs for their own β 1 strands versus the β 1 strand of BamA? If the affinity of an OMPs β -signal strand for its own β 1 strand is greater than for β 1 of BamA, this suggests a mechanism by which nascent OMPs can dissociate from BAM to complete folding.
- What is the conformational state of BamA in proteoliposomes in solution? Single-molecule FRET studies on BamA using a similar strategy as shown here for BAM in Chapter 4 would show if the open state is accessible to BamA at all and whether it represents a minor population, or a significant population as shown for BAM. SurA and an OMP substrate could also be added to try and alter the conformational state. If two states are seen, this could suggest that the lack of solved structures of BamA in the open state is related to the crystallisation process rather than it being the only accessible conformation. If only one state is observed (the closed state) this suggests that the role of additional subunits is to stabilise open conformations to allow for BamA-budding or BamA β 1-elongation and that BamA alone cannot undergo conformational cycling.
- What is the structure of BAM with ‘activator’ mutations, or mutations that are suggested to alter the conformation? Deletion of BamE, mutations in BamD (R197L/S/H), and in BamA (POTRA5 E373K/A, loop 3 F494L, loop 6 Q693P and G669A,) (Ricci et al., 2012; Rigel et al., 2013; Lee et al., 2018) have been shown to change the conformation of BamA *in vivo*. Solving the structures of these mutants may tell us how changes in distant regions are propagated allosterically from different subunits and POTRA domains, to the BamA barrel.
- Using the improved protocol for tag-transfer crosslinking with the full BAM complex, map the folding pathway of OmpA from SurA and to BamA/BamD and any other complex members involved. The methodology presented in Chapter 3 is ripe to be exploited and a series of cysteine mutants along the chain of OmpA, including the β -signal could yield much higher resolution information than that shown in this thesis, in theory as rich as was shown for Skp and SurA. A long standing goal in this field has been the solution of a stalled intermediate, but that only represents a static snapshot of a more complex process. In this situation, spatial resolution should

be sacrificed for better temporal resolution in order to understand the biogenesis pathway.

- Building on the preliminary data shown in Chapter 5, use cryogenic super-resolution microscopy to study the colocalisation of BamA with SecYEG and understand how prevalent ‘super-complexes’ across the IM and OM are. Furthermore, other groups have recently shown the possibility of performing correlative super-resolution microscopy and electron microscopy (Hoffman et al., 2019) and this could eventually be expanded to study the structure of super-complexes by cryo-electron tomography.

6.4 Final thoughts

With over 15 years since the discovery of BamA’s involvement in OMP biogenesis (Voulhoux et al., 2003; Genevrois et al., 2003), and almost 30 years since the first refolding studies of OMPs (Dornmair et al., 1990), it truly feels as if we are very close to understanding all the key details of OMP folding *in vitro* and *in vivo*. It seems likely that there will be many variations on the mechanism of OMP folding, whether you are a small and simple 8-stranded OMP like OmpX, a large and complicated multi-domain protein like the 22-stranded fimbrial usher FimD, or a split-oligomeric barrel like TolC. Nonetheless, solved structures of homologous family members of both OMPs and BAM subunits from diverse species show a remarkable level of structural conservation and Omp85 proteins have a lineage that connects even the most disparate species (including eukaryotes) together, and so despite the differences in subunits or number of BamA POTRA domains between species it is almost certain that the basic mechanisms will also be conserved. While the details will no doubt be ironed out over the coming decades, it seems as though only a few major discoveries remain and with improvements in the tools and techniques of molecular biology, structural biology, bioinformatics, and biophysics that have appeared in the last few years, these discoveries are sure to come in fast and finally blow the lid off the complex mechanisms of OMP biogenesis.

References

- Admasu, A., Gudmundsdóttir, A. D., Platz, M. S., Watt, D. S., Kwiatkowski, S., and Crocker, P. J. (1998). A laser flash photolysis study of p-tolyl(trifluoromethyl)carbene. *Journal of the Chemical Society, Perkin Transactions 2*, (5):1093–1100.
- Akabori, K. and Nagle, J. F. (2015). Structure of the DMPC lipid bilayer ripple phase. *Soft matter*, 11(5):918–26.
- Akashi, H. and Gojobori, T. (2002). Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America*, 99(6):3695–700.
- Albrecht, R. and Zeth, K. (2011). Structural basis of outer membrane protein biogenesis in bacteria. *The Journal of biological chemistry*, 286(31):27792–803.
- Allen, W. J., Corey, R. A., Oatley, P., Sessions, R. B., Baldwin, S. A., Radford, S. E., Tuma, R., and Collinson, I. (2016). Two-way communication between SecY and SecA suggests a Brownian ratchet mechanism for protein translocation. *eLife*, 5(MAY2016):1–23.
- Alvira, S., Watkins, D. W., Troman, L., Lorrیمان, J., Daum, B., Gold, V. A. M., and Collinson, I. (2019). Trans-membrane association of the Sec and BAM complexes for bacterial outer-membrane biogenesis. *bioRxiv*, pages 1–35.
- Andersen, K. K., Wang, H., and Otzen, D. E. (2012). A kinetic analysis of the folding and unfolding of OmpA in urea and guanidinium chloride: single and parallel pathways. *Biochemistry*, 51(42):8371–83.
- António, J. P. M., Russo, R., Carvalho, C. P., Cal, P. M. S. D., and Gois, P. M. P. (2019). Boronic acids as building blocks for the construction of therapeutically useful bioconjugates. *Chemical Society reviews*, 48(13):3513–3536.
- Antonoaea, R., Fürst, M., Nishiyama, K.-I., and Müller, M. (2008). The periplasmic chaperone PpiD interacts with secretory proteins exiting from the SecYEG translocon. *Biochemistry*, 47(20):5649–56.
- Antunes, L. C., Poppleton, D., Klingl, A., Criscuolo, A., Dupuy, B., Brochier-Armanet, C., Beloin, C., and Gribaldo, S. (2016). Phylogenomic analysis supports the ancestral presence of LPS-outer membranes in the Firmicutes. *eLife*, 5(AUGUST):1–21.

- Anwari, K., Poggio, S., Perry, A., Gatsos, X., Ramarathinam, S. H., Williamson, N. A., Noinaj, N., Buchanan, S., Gabriel, K., Purcell, A. W., Jacobs-Wagner, C., and Lithgow, T. (2010). A modular BAM complex in the outer membrane of the alpha-proteobacterium *Caulobacter crescentus*. *PloS one*, 5(1):e8619.
- Anwari, K., Webb, C. T., Poggio, S., Perry, A. J., Belousoff, M., Celik, N., Ramm, G., Lovering, A., Sockett, R. E., Smit, J., Jacobs-Wagner, C., and Lithgow, T. (2012). The evolution of new lipoprotein subunits of the bacterial outer membrane BAM complex. *Molecular microbiology*, 84(5):832–44.
- Aoki, S. K., Malinverni, J. C., Jacoby, K., Thomas, B., Pamma, R., Trinh, B. N., Remers, S., Webb, J., Braaten, B. A., Silhavy, T. J., and Low, D. A. (2008). Contact-dependent growth inhibition requires the essential outer membrane protein BamA (YaeT) as the receptor and the inner membrane transport protein AcrB. *Molecular microbiology*, 70(2):323–40.
- Arnold, T., Zeth, K., and Linke, D. (2010). Omp85 from the thermophilic cyanobacterium *Thermosynechococcus elongatus* differs from proteobacterial Omp85 in structure and domain composition. *The Journal of biological chemistry*, 285(23):18003–15.
- Aronson, D. E., Costantini, L. M., and Snapp, E. L. (2011). Superfolder GFP is fluorescent in oxidizing environments when targeted via the Sec translocon. *Traffic (Copenhagen, Denmark)*, 12(5):543–8.
- Arora, A., Abildgaard, F., Bushweller, J. H., and Tamm, L. K. (2001). Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. *Nature structural biology*, 8(4):334–8.
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, 2:2006.0008.
- Bageshwar, U. K., VerPlank, L., Baker, D., Dong, W., Hamsanathan, S., Whitaker, N., Sacchettini, J. C., and Musser, S. M. (2016). High Throughput Screen for *Escherichia coli* Twin Arginine Translocation (Tat) Inhibitors. *PloS one*, 11(2):e0149659.
- Bakelar, J., Buchanan, S. K., and Noinaj, N. (2016). The structure of the β -barrel assembly machinery complex. *Science (New York, N.Y.)*, 351(6269):180–6.
- Bakelar, J., Buchanan, S. K., and Noinaj, N. (2017). Structural snapshots of the β -barrel assembly machinery. *The FEBS journal*, 284(12):1778–1786.
- Balgavý, P., Dubnicková, M., Kucerka, N., Kiselev, M. A., Yaradaikin, S. P., and Uhríková, D. (2001). Bilayer thickness and lipid interface area in unilamellar extruded 1,2-diacylphosphatidylcholine liposomes: a small-angle neutron scattering study. *Biochimica et biophysica acta*, 1512(1):40–52.
- Balleza, E., Kim, J. M., and Cluzel, P. (2018). Systematic characterization of maturation time of fluorescent proteins in living cells. *Nature methods*, 15(1):47–51.

- Barnes, W., Björk, G., Gérard, J., Jonsson, P., Wasey, J., Worthing, P., and Zwiller, V. (2002). Solid-state single photon sources: light collection strategies. *The European Physical Journal D - Atomic, Molecular and Optical Physics*, 18(2):197–210.
- Bartkiewicz, M., Kazazić, S., Krasowska, J., Clark, P. L., Wielgus-Kutrowska, B., and Bzowska, A. (2018). Non-fluorescent mutant of green fluorescent protein sheds light on the mechanism of chromophore formation. *FEBS letters*, 592(9):1516–1523.
- Baud, C., Guérin, J., Petit, E., Lesne, E., Dupré, E., Loch, C., and Jacob-Dubuisson, F. (2014). Translocation path of a substrate protein through its Omp85 transporter. *Nature communications*, 5:5271.
- Bayley, H. (1983). Reagents for photoaffinity labeling. In Bayley, H., editor, *Laboratory Techniques in Biochemistry and Molecular Biology*, volume 12, chapter 3, pages 25–65.
- Bayley, H. and Knowles, J. R. (1978). Photogenerated reagents for membrane labeling. 2. Phenylcarbene and adamantylidene formed within the lipid bilayer. *Biochemistry*, 17(12):2420–3.
- Behrmann, M., Koch, H. G., Hengelage, T., Wieseler, B., Hoffschulte, H. K., and Müller, M. (1998). Requirements for the translocation of elongation-arrested, ribosome-associated OmpA across the plasma membrane of *Escherichia coli*. *The Journal of biological chemistry*, 273(22):13898–904.
- Bennion, D., Charlson, E. S., Coon, E., and Misra, R. (2010). Dissection of β -barrel outer membrane protein assembly pathways through characterizing BamA POTRA 1 mutants of *Escherichia coli*. *Molecular microbiology*, 77(5):1153–71.
- Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science (New York, N.Y.)*, 313(5793):1642–5.
- Beveridge, T. J. (2001). Use of the gram stain in microbiology. *Biotechnic & histochemistry : official publication of the Biological Stain Commission*, 76(3):111–8.
- Bischoff, R. and Schlüter, H. (2012). Amino acids: chemistry, functionality and selected non-enzymatic post-translational modifications. *Journal of proteomics*, 75(8):2275–96.
- Bitto, E. and McKay, D. B. (2002). Crystallographic structure of SurA, a molecular chaperone that facilitates folding of outer membrane porins. *Structure (London, England : 1993)*, 10(11):1489–98.
- Bitto, E. and McKay, D. B. (2003). The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *The Journal of biological chemistry*, 278(49):49316–22.
- Bitto, E. and McKay, D. B. (2004). Binding of phage-display-selected peptides to the periplasmic chaperone protein SurA mimics binding of unfolded outer membrane proteins. *FEBS letters*, 568(1-3):94–8.

- Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997). The complete genome sequence of *Escherichia coli* K-12. *Science (New York, N.Y.)*, 277(5331):1453–62.
- Bleck, C. K. E., Merz, A., Gutierrez, M. G., Walther, P., Dubochet, J., Zuber, B., and Griffiths, G. (2010). Comparison of different methods for thin section EM analysis of *Mycobacterium smegmatis*. *Journal of microscopy*, 237(1):23–38.
- Bodelón, G., Marín, E., and Fernández, L. A. (2009). Role of periplasmic chaperones and BamA (YaeT/Omp85) in folding and secretion of intimin from enteropathogenic *Escherichia coli* strains. *Journal of bacteriology*, 191(16):5169–79.
- Bond, M. R., Zhang, H., Vu, P. D., and Kohler, J. J. (2009). Photocrosslinking of glycoconjugates using metabolically incorporated diazirine-containing sugars. *Nature protocols*, 4(7):1044–63.
- Bonora, S., Markarian, S., Trincherio, A., and Grigorian, K. (2005). DSC study on the effect of dimethylsulfoxide (DMSO) and diethylsulfoxide (DESO) on phospholipid liposomes. *Thermochimica Acta*, 433(1-2):19–26.
- Bonsor, D. A., Grishkovskaya, I., Dodson, E. J., and Kleanthous, C. (2007). Molecular mimicry enables competitive recruitment by a natively disordered protein. *Journal of the American Chemical Society*, 129(15):4800–7.
- Booth, P. J. and Curnow, P. (2009). Folding scene investigation: membrane proteins. *Current opinion in structural biology*, 19(1):8–13.
- Bornemann, T., Holtkamp, W., and Wintermeyer, W. (2014). Interplay between trigger factor and other protein biogenesis factors on the ribosome. *Nature communications*, 5(May):4180.
- Bos, M. P., Robert, V., and Tommassen, J. (2007). Functioning of outer membrane protein assembly factor Omp85 requires a single POTRA domain. *EMBO reports*, 8(12):1149–54.
- Bos, M. P., Tefsen, B., Geurtsen, J., and Tommassen, J. (2004). Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface. *Proceedings of the National Academy of Sciences of the United States of America*, 101(25):9417–22.
- Bos, M. P. and Tommassen, J. (2005). Viability of a capsule- and lipopolysaccharide-deficient mutant of *Neisseria meningitidis*. *Infection and immunity*, 73(9):6194–7.
- Boyd, D., Weiss, D. S., Chen, J. C., and Beckwith, J. (2000). Towards single-copy gene expression systems making gene cloning physiologically relevant: lambda InCh, a simple *Escherichia coli* plasmid-chromosome shuttle system. *Journal of bacteriology*, 182(3):842–7.
- Braitbard, M., Schneidman-Duhovny, D., and Kalisman, N. (2019). Integrative Structure Modeling: Overview and Assessment. *Annual review of biochemistry*, 88(1):113–135.

- Brewer, C. F. and Riehm, J. P. (1967). Evidence for possible nonspecific reactions between N-ethylmaleimide and proteins. *Analytical Biochemistry*, 18(2):248–255.
- Brodie, N. I., Popov, K. I., Petrotchenko, E. V., Dokholyan, N. V., and Borchers, C. H. (2017). Solving protein structures using short-distance cross-linking constraints as a guide for discrete molecular dynamics simulations. *Science advances*, 3(7):e1700479.
- Browning, D. F., Bavro, V. N., Mason, J. L., Sevastyanovich, Y. R., Rossiter, A. E., Jeeves, M., Wells, T. J., Knowles, T. J., Cunningham, A. F., Donald, J. W., Palmer, T., Overduin, M., and Henderson, I. R. (2015). Cross-species chimeras reveal BamA POTRA and β -barrel domains must be fine-tuned for efficient OMP insertion. *Molecular microbiology*, 97(4):646–59.
- Brunner, J., Senn, H., and Richards, F. M. (1980). 3-Trifluoromethyl-3-phenyldiazirine. A new carbene generating group for photolabeling reagents. *The Journal of biological chemistry*, 255(8):3313–8.
- Buddelmeijer, N., Krehenbrink, M., Pecorari, F., and Pugsley, A. P. (2009). Type II secretion system secretin PulD localizes in clusters in the Escherichia coli outer membrane. *Journal of bacteriology*, 191(1):161–8.
- Bulieris, P. V., Behrens, S., Holst, O., and Kleinschmidt, J. H. (2003). Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone Skp and by lipopolysaccharide. *The Journal of biological chemistry*, 278(11):9092–9.
- Burgess, N. K., Dao, T. P., Stanley, A. M., and Fleming, K. G. (2008). Beta-barrel proteins that reside in the Escherichia coli outer membrane in vivo demonstrate varied folding behavior in vitro. *The Journal of biological chemistry*, 283(39):26748–58.
- Burkhardt, J., Vonck, J., and Averhoff, B. (2011). Structure and function of PilQ, a secretin of the DNA transporter from the thermophilic bacterium Thermus thermophilus HB27. *The Journal of biological chemistry*, 286(12):9977–84.
- Burmann, B. M., Holdbrook, D. A., Callon, M., Bond, P. J., and Hiller, S. (2015). Revisiting the interaction between the chaperone Skp and lipopolysaccharide. *Biophysical journal*, 108(6):1516–1526.
- Burmann, B. M., Wang, C., and Hiller, S. (2013). Conformation and dynamics of the periplasmic membrane-protein-chaperone complexes OmpX-Skp and tOmpA-Skp. *Nature structural & molecular biology*, 20(11):1265–72.
- Caforio, A., Siliakus, M. F., Exterkate, M., Jain, S., Jumde, V. R., Andringa, R. L. H., Kengen, S. W. M., Minnaard, A. J., Driessen, A. J. M., and van der Oost, J. (2018). Converting Escherichia coli into an archaeobacterium with a hybrid heterochiral membrane. *Proceedings of the National Academy of Sciences of the United States of America*, 115(14):3704–3709.
- Calabrese, A. N. and Pukala, T. L. (2013). Chemical Cross-linking and Mass Spectrometry for the Structural Analysis of Protein Assemblies. *Australian Journal of Chemistry*, 66(7):749.

- Callon, M., Burmann, B. M., and Hiller, S. (2014). Structural mapping of a chaperone-substrate interaction surface. *Angewandte Chemie (International ed. in English)*, 53(20):5069–72.
- Carbonetti, N. H. and Sparling, P. F. (1987). Molecular cloning and characterization of the structural gene for protein I, the major outer membrane protein of *Neisseria gonorrhoeae*. *Proceedings of the National Academy of Sciences of the United States of America*, 84(24):9084–8.
- Carlson, M. L., Stacey, R. G., Young, J. W., Wason, I. S., Zhao, Z., Rattray, D. G., Scott, N., Kerr, C. H., Babu, M., Foster, L. J., and Duong Van Hoa, F. (2019). Profiling the *Escherichia coli* membrane protein interactome captured in Peptidisc libraries. *eLife*, 8:1–26.
- Carter, S. D., Mageswaran, S. K., Farino, Z. J., Mamede, J. I., Oikonomou, C. M., Hope, T. J., Freyberg, Z., and Jensen, G. J. (2018). Distinguishing signal from autofluorescence in cryogenic correlated light and electron microscopy of mammalian cells. *Journal of structural biology*, 201(1):15–25.
- Castán, P., Zafra, O., Moreno, R., de Pedro, M. A., Vallés, C., Cava, F., Caro, E., Schwarz, H., and Berenguer, J. (2002). The periplasmic space in *Thermus thermophilus*: evidence from a regulation-defective S-layer mutant overexpressing an alkaline phosphatase. *Extremophiles : life under extreme conditions*, 6(3):225–32.
- Casuso, I., Khao, J., Chami, M., Paul-Gilloteaux, P., Husain, M., Duneau, J.-P., Stahlberg, H., Sturgis, J. N., and Scheuring, S. (2012). Characterization of the motion of membrane proteins using high-speed atomic force microscopy. *Nature nanotechnology*, 7(8):525–9.
- Cava, F., de Pedro, M. A., Blas-Galindo, E., Waldo, G. S., Westblade, L. F., and Berenguer, J. (2008). Expression and use of superfolder green fluorescent protein at high temperatures in vivo: a tool to study extreme thermophile biology. *Environmental microbiology*, 10(3):605–13.
- Cavicchioli, R., Curmi, P. M. G., Saunders, N., and Thomas, T. (2003). Pathogenic archaea: do they exist? *BioEssays : news and reviews in molecular, cellular and developmental biology*, 25(11):1119–28.
- Chandran, V., Fronzes, R., Duquerroy, S., Cronin, N., Navaza, J., and Waksman, G. (2009). Structure of the outer membrane complex of a type IV secretion system. *Nature*, 462(7276):1011–5.
- Chang, C.-F., Mfuh, A., Gao, J., Wu, H.-Y., and Woo, C. M. (2018). Synthesis of an electronically-tuned minimally interfering alkynyl photo-affinity label to measure small molecule–protein interactions. *Tetrahedron*, 74(26):3273–3277.
- Chang, H. H. and Dea, P. K. (2001). Insights into the dynamics of DMSO in phosphatidylcholine bilayers. *Biophysical chemistry*, 94(1-2):33–40.
- Chang, Y.-W., Chen, S., Tocheva, E. I., Treuner-Lange, A., Löbach, S., Søggaard-Andersen, L., and Jensen, G. J. (2014). Correlated cryogenic photoactivated localization microscopy and cryo-electron tomography. *Nature methods*, 11(7):737–9.

- Charles, M., Pérez, M., Kobil, J. H., and Goldberg, M. B. (2001). Polar targeting of Shigella virulence factor IcsA in Enterobacteriaceae and Vibrio. *Proceedings of the National Academy of Sciences of the United States of America*, 98(17):9871–6.
- Charlson, E. S., Werner, J. N., and Misra, R. (2006). Differential effects of yfgL mutation on Escherichia coli outer membrane proteins and lipopolysaccharide. *Journal of bacteriology*, 188(20):7186–94.
- Chaturvedi, D. and Mahalakshmi, R. (2013). Methionine mutations of outer membrane protein X influence structural stability and beta-barrel unfolding. *PloS one*, 8(11):e79351.
- Chaturvedi, D. and Mahalakshmi, R. (2018a). Folding Determinants of Transmembrane β -Barrels Using Engineered OMP Chimeras. *Biochemistry*, 57(13):1987–1996.
- Chaturvedi, D. and Mahalakshmi, R. (2018b). Position-Specific contribution of interface tryptophans on membrane protein energetics. *Biochimica et biophysica acta. Biomembranes*, 1860(2):451–457.
- Chavez, J. D., Mohr, J. P., Mathay, M., Zhong, X., Keller, A., and Bruce, J. E. (2019). Systems structural biology measurements by in vivo cross-linking with mass spectrometry. *Nature protocols*, page 1.
- Chee, G.-L., Yalowich, J. C., Bodner, A., Wu, X., and Hasinoff, B. B. (2010). A diazirine-based photoaffinity etoposide probe for labeling topoisomerase II. *Bioorganic & medicinal chemistry*, 18(2):830–8.
- Chen, F., Zhao, Q., Cai, X., Lv, L., Lin, W., Yu, X., Li, C., Li, Y., Xiong, M., and Wang, X.-G. (2009). Phosphatidylcholine in membrane of Escherichia coli changes bacterial antigenicity. *Canadian journal of microbiology*, 55(11):1328–34.
- Chen, J. C., Viollier, P. H., and Shapiro, L. (2005). A membrane metalloprotease participates in the sequential degradation of a Caulobacter polarity determinant. *Molecular microbiology*, 55(4):1085–103.
- Chen, R. and Henning, U. (1996). A periplasmic protein (Skp) of Escherichia coli selectively binds a class of outer membrane proteins. *Molecular microbiology*, 19(6):1287–94.
- Chen, Z., Zhan, L. H., Hou, H. F., Gao, Z. Q., Xu, J. H., Dong, C., and Dong, Y. H. (2016). Structural basis for the interaction of BamB with the POTRA3-4 domains of BamA. *Acta crystallographica. Section D, Structural biology*, 72(Pt 2):236–44.
- Chen, Z. A. and Rappsilber, J. (2018). Protein Dynamics in Solution by Quantitative Crosslinking/Mass Spectrometry. *Trends in biochemical sciences*, 43(11):908–920.
- Cheng, C.-Y., Song, J., Pas, J., Meijer, L. H. H., and Han, S. (2015). DMSO induces dehydration near lipid membrane surfaces. *Biophysical journal*, 109(2):330–9.

- Chimalakonda, G., Ruiz, N., Chng, S.-S., Garner, R. A., Kahne, D., and Silhavy, T. J. (2011). Lipoprotein LptE is required for the assembly of LptD by the beta-barrel assembly machine in the outer membrane of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 108(6):2492–7.
- Cho, S.-H., Szewczyk, J., Pesavento, C., Zietek, M., Banzhaf, M., Roszczenko, P., Asmar, A., Laloux, G., Hov, A.-K., Leverrier, P., Van der Henst, C., Vertommen, D., Typas, A., and Collet, J.-F. (2014). Detecting envelope stress by monitoring β -barrel assembly. *Cell*, 159(7):1652–64.
- Chorev, D. S., Baker, L. A., Wu, D., Beilsten-Edmands, V., Rouse, S. L., Zeev-Ben-Mordehai, T., Jiko, C., Samsudin, F., Gerle, C., Khalid, S., Stewart, A. G., Matthews, S. J., Grünewald, K., and Robinson, C. V. (2018). Protein assemblies ejected directly from native membranes yield complexes for mass spectrometry. *Science (New York, N.Y.)*, 362(6416):829–834.
- Chou, K. C., Heckel, A., Némethy, G., Rumsey, S., Carlacci, L., and Scheraga, H. A. (1990). Energetics of the structure and chain tilting of antiparallel beta-barrels in proteins. *Proteins*, 8(1):14–22.
- Church, R. F. R., Kende, A. S., and Weiss, M. J. (1965). Diazirines. I. Some Observations on the Scope of the Ammonia-Hydroxylamine-O-sulfonic Acid Diaziridine Synthesis. The Preparation of Certain Steroid Diaziridines and Diazirines. *Journal of the American Chemical Society*, 87(12):2665–2671.
- Cierpicki, T., Liang, B., Tamm, L. K., and Bushweller, J. H. (2006). Increasing the accuracy of solution NMR structures of membrane proteins by application of residual dipolar couplings. High-resolution structure of outer membrane protein A. *Journal of the American Chemical Society*, 128(21):6947–51.
- Clantin, B., Delattre, A.-S., Rucktooa, P., Saint, N., Méli, A. C., Locht, C., Jacob-Dubuisson, F., and Villeret, V. (2007). Structure of the membrane protein FhaC: a member of the Omp85-TpsB transporter superfamily. *Science (New York, N.Y.)*, 317(5840):957–61.
- Cloin, B. M. C., De Zitter, E., Salas, D., Gielen, V., Folkers, G. E., Mikhaylova, M., Bergeler, M., Krajnik, B., Harvey, J., Hoogenraad, C. C., Van Meervelt, L., Dedeker, P., and Kapitein, L. C. (2017). Efficient switching of mCherry fluorescence using chemical caging. *Proceedings of the National Academy of Sciences of the United States of America*, 114(27):7013–7018.
- Collin, S., Guilvout, I., Chami, M., and Pugsley, A. P. (2007). YaeT-independent multimerization and outer membrane association of secretin PulD. *Molecular microbiology*, 64(5):1350–7.
- Cong, Q., Anishchenko, I., Ovchinnikov, S., and Baker, D. (2019). Protein interaction networks revealed by proteome coevolution. *Science (New York, N.Y.)*, 365(6449):185–189.
- Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic acids research*, 16(22):10881–90.

- Costello, S. M., Plummer, A. M., Fleming, P. J., and Fleming, K. G. (2016). Dynamic periplasmic chaperone reservoir facilitates biogenesis of outer membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 113(33):E4794–800.
- Dahlberg, P. D., Sartor, A. M., Wang, J., Saurabh, S., Shapiro, L., and Moerner, W. E. (2018). Identification of PAmKate as a Red Photoactivatable Fluorescent Protein for Cryogenic Super-Resolution Imaging. *Journal of the American Chemical Society*, 140(39):12310–12313.
- Daimon, Y., Iwama-Masui, C., Tanaka, Y., Shiota, T., Suzuki, T., Miyazaki, R., Sakurada, H., Lithgow, T., Dohmae, N., Mori, H., Tsukazaki, T., Narita, S.-i., and Akiyama, Y. (2017). The TPR domain of BepA is required for productive interaction with substrate proteins and the β -barrel assembly machinery complex. *Molecular microbiology*, 106(5):760–776.
- Dammeyer, T., Timmis, K. N., and Tinnefeld, P. (2013). Broad host range vectors for expression of proteins with (Twin-) Strep-tag, His-tag and engineered, export optimized yellow fluorescent protein. *Microbial cell factories*, 12(1):49.
- Danoff, E. J. and Fleming, K. G. (2011). The soluble, periplasmic domain of OmpA folds as an independent unit and displays chaperone activity by reducing the self-association propensity of the unfolded OmpA transmembrane β -barrel. *Biophysical chemistry*, 159(1):194–204.
- Danoff, E. J. and Fleming, K. G. (2015a). Aqueous, Unfolded OmpA Forms Amyloid-Like Fibrils upon Self-Association. *PloS one*, 10(7):e0132301.
- Danoff, E. J. and Fleming, K. G. (2015b). Membrane defects accelerate outer membrane β -barrel protein folding. *Biochemistry*, 54(2):97–9.
- Danoff, E. J. and Fleming, K. G. (2017). Novel Kinetic Intermediates Populated along the Folding Pathway of the Transmembrane β -Barrel OmpA. *Biochemistry*, 56(1):47–60.
- Dartigalongue, C. and Raina, S. (1998). A new heat-shock gene, *ppiD*, encodes a peptidyl-prolyl isomerase required for folding of outer membrane proteins in *Escherichia coli*. *The EMBO journal*, 17(14):3968–80.
- Day, P. M., Inoue, K., and Theg, S. M. (2019). Chloroplast Outer Membrane β -Barrel Proteins Use Components of the General Import Apparatus. *The Plant cell*, page tpc.00001.2019.
- De Cock, H., Schäfer, U., Potgeter, M., Demel, R., Müller, M., and Tommassen, J. (1999). Affinity of the periplasmic chaperone Skp of *Escherichia coli* for phospholipids, lipopolysaccharides and non-native outer membrane proteins. Role of Skp in the biogenesis of outer membrane protein. *European journal of biochemistry*, 259(1-2):96–103.
- de Cock, H., Struyvé, M., Kleerebezem, M., van der Krift, T., and Tommassen, J. (1997). Role of the carboxy-terminal phenylalanine in the biogenesis of outer membrane protein PhoE of *Escherichia coli* K-12. *Journal of molecular biology*, 269(4):473–8.

- de Cock, H. and Tommassen, J. (1992). SecB-binding does not maintain the translocation-competent state of prePhoE. *Molecular microbiology*, 6(5):599–604.
- DeChavigny, A., Heacock, P. N., and Dowhan, W. (1991). Sequence and inactivation of the pss gene of Escherichia coli. Phosphatidylethanolamine may not be essential for cell viability. *The Journal of biological chemistry*, 266(8):5323–32.
- Denisov, I. G., McLean, M. A., Shaw, A. W., Grinkova, Y. V., and Sligar, S. G. (2005). Thermotropic phase transition in soluble nanoscale lipid bilayers. *The journal of physical chemistry. B*, 109(32):15580–8.
- Denoncin, K. and Collet, J.-F. (2013). Disulfide bond formation in the bacterial periplasm: major achievements and challenges ahead. *Antioxidants & redox signaling*, 19(1):63–71.
- Denoncin, K., Schwalm, J., Vertommen, D., Silhavy, T. J., and Collet, J.-F. (2012). Dissecting the Escherichia coli periplasmic chaperone network using differential proteomics. *Proteomics*, 12(9):1391–401.
- Dewald, A. H., Hodges, J. C., and Columbus, L. (2011). Physical determinants of β -barrel membrane protein folding in lipid vesicles. *Biophysical journal*, 100(9):2131–40.
- D’Imprima, E., Salzer, R., Bhaskara, R. M., Sánchez, R., Rose, I., Kirchner, L., Hummer, G., Kühlbrandt, W., Vonck, J., and Averhoff, B. (2017). Cryo-EM structure of the bifunctional secretin complex of Thermus thermophilus. *eLife*, 6:1–23.
- Dinh, T. and Bernhardt, T. G. (2011). Using superfolder green fluorescent protein for periplasmic protein localization studies. *Journal of bacteriology*, 193(18):4984–7.
- Disalvo, E. A., Pinto, O. A., Martini, M. F., Bouchet, A. M., Hollmann, A., and Frías, M. A. (2015). Functional role of water in membranes updated: A tribute to Träuble. *Biochimica et biophysica acta*, 1848(7):1552–62.
- Doerner, P. A. and Sousa, M. C. (2017). Extreme Dynamics in the BamA β -Barrel Seam. *Biochemistry*, 56(24):3142–3149.
- Dong, C., Beis, K., Nesper, J., Brunkan-Lamontagne, A. L., Clarke, B. R., Whitfield, C., and Naismith, J. H. (2006). Wza the translocon for E. coli capsular polysaccharides defines a new class of membrane protein. *Nature*, 444(7116):226–9.
- Dormán, G., Nakamura, H., Pulsipher, A., and Prestwich, G. D. (2016). The Life of Pi Star: Exploring the Exciting and Forbidden Worlds of the Benzophenone Photophore. *Chemical reviews*, 116(24):15284–15398.
- Dornmair, K., Kiefer, H., and Jähnig, F. (1990). Refolding of an integral membrane protein. OmpA of Escherichia coli. *The Journal of biological chemistry*, 265(31):18907–11.
- Dou, J., Vorobieva, A. A., Sheffler, W., Doyle, L. A., Park, H., Bick, M. J., Mao, B., Foight, G. W., Lee, M. Y., Gagnon, L. A., Carter, L., Sankaran, B., Ovchinnikov, S., Marcos, E., Huang, P.-S., Vaughan, J. C., Stoddard, B. L., and Baker, D. (2018). De novo design of a fluorescence-activating β -barrel. *Nature*, 561(7724):485–491.

- Dowdell, A. S., Murphy, M. D., Azodi, C., Swanson, S. K., Florens, L., Chen, S., and Zückert, W. R. (2017). Comprehensive Spatial Analysis of the *Borrelia burgdorferi* Lipoproteome Reveals a Compartmentalization Bias toward the Bacterial Surface. *Journal of bacteriology*, 199(6):1–20.
- Doyle, M. T. and Bernstein, H. D. (2019). Bacterial outer membrane proteins assemble via asymmetric interactions with the BamA β -barrel. *Nature communications*, 10(1):3358.
- Dunn, J. P., Kenedy, M. R., Iqbal, H., and Akins, D. R. (2015). Characterization of the β -barrel assembly machine accessory lipoproteins from *Borrelia burgdorferi*. *BMC microbiology*, 15(1):70.
- Dunstan, R. A., Hay, I. D., Wilksch, J. J., Schittenhelm, R. B., Purcell, A. W., Clark, J., Costin, A., Ramm, G., Strugnell, R. A., and Lithgow, T. (2015). Assembly of the secretion pores GspD, Wza and CsgG into bacterial outer membranes does not require the Omp85 proteins BamA or TamA. *Molecular microbiology*, 97(4):616–29.
- Eisenberg, D., Schwarz, E., Komaromy, M., and Wall, R. (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *Journal of molecular biology*, 179(1):125–42.
- Eisner, G., Moser, M., Schäfer, U., Beck, K., and Müller, M. (2006). Alternate recruitment of signal recognition particle and trigger factor to the signal sequence of a growing nascent polypeptide. *The Journal of biological chemistry*, 281(11):7172–9.
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., Qureshi, M., Richardson, L. J., Salazar, G. A., Smart, A., Sonnhammer, E. L. L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S. C. E., and Finn, R. D. (2019). The Pfam protein families database in 2019. *Nucleic acids research*, 47(D1):D427–D432.
- El Khatib, M., Martins, A., Bourgeois, D., Colletier, J.-P., and Adam, V. (2016). Rational design of ultrastable and reversibly photoswitchable fluorescent proteins for super-resolution imaging of the bacterial periplasm. *Scientific reports*, 6(November 2015):18459.
- Ellena, J. F., Lackowicz, P., Montgomery, H., and Cafiso, D. S. (2011). Membrane thickness varies around the circumference of the transmembrane protein BtuB. *Biophysical journal*, 100(5):1280–7.
- Ellis, H. M., Yu, D., DiTizio, T., and Court, D. L. (2001). High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proceedings of the National Academy of Sciences of the United States of America*, 98(12):6742–6.
- Enders, O., Ngezahayo, A., Wiechmann, M., Leisten, F., and Kolb, H.-A. (2004). Structural calorimetry of main transition of supported DMPC bilayers by temperature-controlled AFM. *Biophysical journal*, 87(4):2522–31.
- Endo, T., Kawano, S., and Yamano, K. (2011). BamE structure: the assembly of β -barrel proteins in the outer membranes of bacteria and mitochondria. *EMBO reports*, 12(2):94–5.

- Erni, B. and Khorana, H. G. (1980). Fatty acids containing photoactivable carbene precursors. Synthesis and photochemical properties of 3,3-bis(1,1-difluorohexyl)diazirine and 3-(1,1-difluorooctyl)-3H-diazirine. *Journal of the American Chemical Society*, 102(11):3888–3896.
- Erridge, C., Bennett-Guerrero, E., and Poxton, I. R. (2002). Structure and function of lipopolysaccharides. *Microbes and infection*, 4(8):837–51.
- Fardini, Y., Trotereau, J., Bottreau, E., Souchard, C., Velge, P., and Virlogeux-Payant, I. (2009). Investigation of the role of the BAM complex and SurA chaperone in outer-membrane protein biogenesis and type III secretion system expression in *Salmonella*. *Microbiology (Reading, England)*, 155(Pt 5):1613–22.
- Feilmeier, B. J., Iseminger, G., Schroeder, D., Webber, H., and Phillips, G. J. (2000). Green fluorescent protein functions as a reporter for protein localization in *Escherichia coli*. *Journal of bacteriology*, 182(14):4068–76.
- Feng, J., Deschout, H., Caneva, S., Hofmann, S., Lončarić, I., Lazić, P., and Radenovic, A. (2018). Imaging of Optically Active Defects with Nanometer Resolution. *Nano letters*, 18(3):1739–1744.
- Ferber, M., Kosinski, J., Ori, A., Rashid, U. J., Moreno-Morcillo, M., Simon, B., Bouvier, G., Batista, P. R., Müller, C. W., Beck, M., and Nilges, M. (2016). Automated structure modeling of large protein assemblies using crosslinks as distance restraints. *Nature methods*, 13(6):515–20.
- Ferguson, A. D., Welte, W., Hofmann, E., Lindner, B., Holst, O., Coulton, J. W., and Diederichs, K. (2000). A conserved structural motif for lipopolysaccharide recognition by prokaryotic and eukaryotic proteins. *Structure (London, England : 1993)*, 8(6):585–92.
- Fessenden, J. D. and Mahalingam, M. (2013). Site-specific labeling of the type 1 ryanodine receptor using biarsenical fluorophores targeted to engineered tetracysteine motifs. *PloS one*, 8(5):e64686.
- Fessl, T., Watkins, D., Oatley, P., Allen, W. J., Corey, R. A., Horne, J., Baldwin, S. A., Radford, S. E., Collinson, I., and Tuma, R. (2018). Dynamic action of the Sec machinery during initiation, protein translocation and termination. *eLife*, 7:e35112.
- Fidorra, M., Duelund, L., Leidy, C., Simonsen, A. C., and Bagatolli, L. A. (2006). Absence of fluid-ordered/fluid-disordered phase coexistence in ceramide/POPC mixtures containing cholesterol. *Biophysical journal*, 90(12):4437–51.
- Fisher, A. C. and DeLisa, M. P. (2008). Laboratory evolution of fast-folding green fluorescent protein using secretory pathway quality control. *PloS one*, 3(6):e2351.
- Fishov, I. and Woldringh, C. L. (1999). Visualization of membrane domains in *Escherichia coli*. *Molecular microbiology*, 32(6):1166–72.
- Fleming, K. G. (2015). A combined kinetic push and thermodynamic pull as driving forces for outer membrane protein sorting and folding in bacteria. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 370(1679).

- Fleming, P. J., Patel, D. S., Wu, E. L., Qi, Y., Yeom, M. S., Sousa, M. C., Fleming, K. G., and Im, W. (2016). BamA POTRA Domain Interacts with a Native Lipid Membrane Surface. *Biophysical journal*, 110(12):2698–2709.
- Franklin, M. W., Nepomnyachiy, S., Feehan, R., Ben-Tal, N., Kolodny, R., and Slusky, J. S. G. (2018a). Efflux Pumps Represent Possible Evolutionary Convergence onto the β -Barrel Fold. *Structure (London, England : 1993)*, 26(9):1266–1274.e2.
- Franklin, M. W., Nepomnyachiy, S., Feehan, R., Ben-Tal, N., Kolodny, R., and Slusky, J. S. (2018b). Evolutionary pathways of repeat protein topology in bacterial outer membrane proteins. *eLife*, 7:1–19.
- Fujii, T., Kato, T., Hiraoka, K. D., Miyata, T., Minamino, T., Chevance, F. F. V., Hughes, K. T., and Namba, K. (2017). Identical folds used for distinct mechanical functions of the bacterial flagellar rod and hook. *Nature communications*, 8(May 2016):14276.
- Fürst, M., Zhou, Y., Merfort, J., and Müller, M. (2018). Involvement of PpiD in Sec-dependent protein translocation. *Biochimica et biophysica acta. Molecular cell research*, 1865(2):273–280.
- Furubayashi, T., Motohashi, K., Wakao, K., Matsuda, T., Kii, I., Hosoya, T., Hayashi, N., Sadaie, M., Ishikawa, F., Matsushita, M., and Fujiyoshi, S. (2017). Three-Dimensional Localization of an Individual Fluorescent Molecule with Angstrom Precision. *Journal of the American Chemical Society*, 139(26):8990–8994.
- Fussenegger, M., Facius, D., Meier, J., and Meyer, T. F. (1996). A novel peptidoglycan-linked lipoprotein (ComL) that functions in natural transformation competence of *Neisseria gonorrhoeae*. *Molecular microbiology*, 19(5):1095–105.
- Gallivan, J. P. and Dougherty, D. A. (1999). Cation- π interactions in structural biology. *Proceedings of the National Academy of Sciences of the United States of America*, 96(17):9459–64.
- Gatsos, X., Perry, A. J., Anwari, K., Dolezal, P., Wolyneć, P. P., Likić, V. A., Purcell, A. W., Buchanan, S. K., and Lithgow, T. (2008). Protein secretion and outer membrane assembly in Alphaproteobacteria. *FEMS microbiology reviews*, 32(6):995–1009.
- Gatzeva-Topalova, P. Z., Walton, T. A., and Sousa, M. C. (2008). Crystal structure of YaeT: conformational flexibility and substrate recognition. *Structure (London, England : 1993)*, 16(12):1873–81.
- Genevrois, S., Steeghs, L., Roholl, P., Letesson, J.-J., and van der Ley, P. (2003). The Omp85 protein of *Neisseria meningitidis* is required for lipid export to the outer membrane. *The EMBO journal*, 22(8):1780–9.
- Gentle, I., Gabriel, K., Beech, P., Waller, R., and Lithgow, T. (2004). The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *The Journal of cell biology*, 164(1):19–24.

- Gerding, M. A., Ogata, Y., Pecora, N. D., Niki, H., and de Boer, P. A. J. (2007). The trans-envelope Tol-Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in *E. coli*. *Molecular microbiology*, 63(4):1008–25.
- Gessmann, D., Chung, Y. H., Danoff, E. J., Plummer, A. M., Sandlin, C. W., Zaccai, N. R., and Fleming, K. G. (2014). Outer membrane β -barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. *Proceedings of the National Academy of Sciences of the United States of America*, 111(16):5878–83.
- Geurink, P. P., Prely, L. M., van der Marel, G. A., Bischoff, R., and Overkleeft, H. S. (2012). Photoaffinity labeling in activity-based protein profiling. *Topics in current chemistry*, 324:85–113.
- Gibbs, K. A., Isaac, D. D., Xu, J., Hendrix, R. W., Silhavy, T. J., and Theriot, J. A. (2004). Complex spatial distribution and dynamics of an abundant *Escherichia coli* outer membrane protein, LamB. *Molecular microbiology*, 53(6):1771–83.
- Gil, T., Ipsen, J. H., Mouritsen, O. G., Sabra, M. C., Sperotto, M. M., and Zuckermann, M. J. (1998). Theoretical analysis of protein organization in lipid membranes. *Biochimica et biophysica acta*, 1376(3):245–66.
- Gill, E. E. and Brinkman, F. S. L. (2011). The proportional lack of archaeal pathogens: Do viruses/phages hold the key? *BioEssays : news and reviews in molecular, cellular and developmental biology*, 33(4):248–54.
- Ginez, L. D., Osorio, A., Camarena, L., and Poggio, S. (2019). Establishment of a Protein Concentration Gradient in the Outer Membrane Requires Two Diffusion-Limiting Mechanisms. *Journal of bacteriology*, 201(17).
- Ginez, L. D., Osorio, A., and Poggio, S. (2014). Localization of the outer membrane protein OmpA2 in *Caulobacter crescentus* depends on the position of the gene in the chromosome. *Journal of bacteriology*, 196(15):2889–900.
- Godlewska, R., Wiśniewska, K., Pietras, Z., and Jagusztyn-Krynicka, E. K. (2009). Peptidoglycan-associated lipoprotein (Pal) of Gram-negative bacteria: function, structure, role in pathogenesis and potential application in immunoprophylaxis. *FEMS microbiology letters*, 298(1):1–11.
- Gold, V. A. M., Ieva, R., Walter, A., Pfanner, N., van der Laan, M., and Kühlbrandt, W. (2014). Visualizing active membrane protein complexes by electron cryotomography. *Nature communications*, 5(May):4129.
- González-Rivera, C., Khara, P., Awad, D., Patel, R., Li, Y. G., Bogisch, M., and Christie, P. J. (2019). Two pKM101-encoded proteins, the pilus-tip protein TraC and Pep, assemble on the *Escherichia coli* cell surface as adhesins required for efficient conjugative DNA transfer. *Molecular microbiology*, 111(1):96–117.
- Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E., and Desai, M. M. (2017). The dynamics of molecular evolution over 60,000 generations. *Nature*, 551(7678):45–50.

- Goodall, E. C. A., Robinson, A., Johnston, I. G., Jabbari, S., Turner, K. A., Cunningham, A. F., Lund, P. A., Cole, J. A., and Henderson, I. R. (2018). The Essential Genome of *Escherichia coli* K-12. *mBio*, 9(1):1–18.
- Goose, J. E. and Sansom, M. S. P. (2013). Reduced lateral mobility of lipids and proteins in crowded membranes. *PLoS computational biology*, 9(4):e1003033.
- Gotschlich, E. C., Seiff, M. E., Blake, M. S., and Koomey, M. (1987). Porin protein of *Neisseria gonorrhoeae*: cloning and gene structure. *Proceedings of the National Academy of Sciences of the United States of America*, 84(22):8135–9.
- Grabowicz, M. (2018). Lipoprotein Transport: Greasing the Machines of Outer Membrane Biogenesis: Re-Examining Lipoprotein Transport Mechanisms Among Diverse Gram-Negative Bacteria While Exploring New Discoveries and Questions. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 40(4):e1700187.
- Grabowicz, M. and Silhavy, T. J. (2017). Redefining the essential trafficking pathway for outer membrane lipoproteins. *Proceedings of the National Academy of Sciences of the United States of America*, 114(18):4769–4774.
- Griffin, B. A., Adams, S. R., and Tsien, R. Y. (1998). Specific covalent labeling of recombinant protein molecules inside live cells. *Science (New York, N.Y.)*, 281(5374):269–72.
- Griller, D., Nazran, A. S., and Scaiano, J. C. (1984). Flash photolysis studies of carbenes and their reaction kinetics. *Accounts of Chemical Research*, 17(8):283–289.
- Gruss, F., Zähringer, F., Jakob, R. P., Burmann, B. M., Hiller, S., and Maier, T. (2013). The structural basis of autotransporter translocation by TamA. *Nature structural & molecular biology*, 20(11):1318–20.
- Gu, Y., Li, H., Dong, H., Zeng, Y., Zhang, Z., Paterson, N. G., Stansfeld, P. J., Wang, Z., Zhang, Y., Wang, W., and Dong, C. (2016). Structural basis of outer membrane protein insertion by the BAM complex. *Nature*, 531(7592):64–9.
- Gu, Y., Zeng, Y., Wang, Z., and Dong, C. (2017). BamA β 16C strand and periplasmic turns are critical for outer membrane protein insertion and assembly. *The Biochemical journal*, 474(23):3951–3961.
- Guilvout, I., Brier, S., Chami, M., Hourdel, V., Francetic, O., Pugsley, A. P., Chamot-Rooke, J., and Huysmans, G. H. M. (2017). Prepore Stability Controls Productive Folding of the BAM-independent Multimeric Outer Membrane Secretin PulD. *The Journal of biological chemistry*, 292(1):328–338.
- Gunasinghe, S. D., Shiota, T., Stubenrauch, C. J., Schulze, K. E., Webb, C. T., Fulcher, A. J., Dunstan, R. A., Hay, I. D., Naderer, T., Whelan, D. R., Bell, T. D. M., Elgass, K. D., Strugnell, R. A., and Lithgow, T. (2018). The WD40 Protein BamB Mediates Coupling of BAM Complexes into Assembly Precincts in the Bacterial Outer Membrane. *Cell reports*, 23(9):2782–2794.

- Gunewardene, M. S., Subach, F. V., Gould, T. J., Penoncello, G. P., Gudheti, M. V., Verkhusha, V. V., and Hess, S. T. (2011). Superresolution imaging of multiple fluorescent proteins with highly overlapping emission spectra in living cells. *Biophysical journal*, 101(6):1522–8.
- Guo, M. S. and Gross, C. A. (2014). Stress-induced remodeling of the bacterial proteome. *Current biology : CB*, 24(10):R424–34.
- Gupta, R. S. (2011). Origin of diderm (Gram-negative) bacteria: antibiotic selection pressure rather than endosymbiosis likely led to the evolution of bacterial cells with two membranes. *Antonie van Leeuwenhoek*, 100(2):171–82.
- Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of bacteriology*, 177(14):4121–30.
- Hagan, C. L. and Kahne, D. (2011). The reconstituted Escherichia coli Bam complex catalyzes multiple rounds of β -barrel assembly. *Biochemistry*, 50(35):7444–6.
- Hagan, C. L., Kim, S., and Kahne, D. (2010). Reconstitution of outer membrane protein assembly from purified components. *Science (New York, N.Y.)*, 328(5980):890–2.
- Hagan, C. L., Westwood, D. B., and Kahne, D. (2013). Bam Lipoproteins Assemble BamA in vitro. *Biochemistry*, 52(35):6108–13.
- Hagan, C. L., Wzorek, J. S., and Kahne, D. (2015). Inhibition of the β -barrel assembly machine by a peptide that binds BamD. *Proceedings of the National Academy of Sciences of the United States of America*, 112(7):2011–6.
- Halo, T. L., Appelbaum, J., Hobert, E. M., Balkin, D. M., and Schepartz, A. (2009). Selective recognition of protein tetraserine motifs with a cell-permeable, pro-fluorescent bis-boronic acid. *Journal of the American Chemical Society*, 131(2):438–9.
- Han, L., Zheng, J., Wang, Y., Yang, X., Liu, Y., Sun, C., Cao, B., Zhou, H., Ni, D., Lou, J., Zhao, Y., and Huang, Y. (2016). Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nature structural & molecular biology*, 23(3):192–6.
- Harms, N., Koningstein, G., Dontje, W., Muller, M., Oudega, B., Luirink, J., and de Cock, H. (2001). The early interaction of the outer membrane protein phoE with the periplasmic chaperone Skp occurs at the cytoplasmic membrane. *The Journal of biological chemistry*, 276(22):18804–11.
- Harris, F. M., Best, K. B., and Bell, J. D. (2002). Use of laurdan fluorescence intensity and polarization to distinguish between changes in membrane fluidity and phospholipid order. *Biochimica et biophysica acta*, 1565(1):123–8.
- Harris, L. J., Skaletsky, E., and McPherson, A. (1998). Crystallographic structure of an intact IgG1 monoclonal antibody. *Journal of molecular biology*, 275(5):861–72.

- Hart, E. M., Gupta, M., Wühr, M., and Silhavy, T. J. (2019). The Synthetic Phenotype of Δ bamB Δ bamE Double Mutants Results from a Lethal Jamming of the Bam Complex by the Lipoprotein RcsF. *mBio*, 10(3):1–12.
- Hartmann, J.-B., Zahn, M., Burmann, I. M., Bibow, S., and Hiller, S. (2018). Sequence-Specific Solution NMR Assignments of the β -Barrel Insertase BamA to Monitor Its Conformational Ensemble at the Atomic Level. *Journal of the American Chemical Society*, 140(36):11252–11260.
- Hashimoto, M. and Hatanaka, Y. (2006). Practical conditions for photoaffinity labeling with 3-trifluoromethyl-3-phenyldiazirine photophore. *Analytical biochemistry*, 348(1):154–6.
- Hayer-Hartl, M., Bracher, A., and Hartl, F. U. (2016). The GroEL-GroES Chaperonin Machine: A Nano-Cage for Protein Folding. *Trends in biochemical sciences*, 41(1):62–76.
- Heimburg, T. (2007). Permeability. In *Thermal Biophysics of Membranes*, chapter 17, pages 289–300. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.
- Heinz, E. and Lithgow, T. (2014). A comprehensive analysis of the Omp85/TpsB protein superfamily structural diversity, taxonomic occurrence, and evolution. *Frontiers in microbiology*, 5(JULY):370.
- Heinz, E., Selkrig, J., Belousoff, M. J., and Lithgow, T. (2015). Evolution of the Translocation and Assembly Module (TAM). *Genome biology and evolution*, 7(6):1628–43.
- Hennecke, G., Nolte, J., Volkmer-Engert, R., Schneider-Mergener, J., and Behrens, S. (2005). The periplasmic chaperone SurA exploits two features characteristic of integral outer membrane proteins for selective substrate recognition. *The Journal of biological chemistry*, 280(25):23540–8.
- Henning, U., Höhn, B., and Sonntag, I. (1973). Cell envelope and shape of *Escherichia coli* K12. The ghost membrane. *European journal of biochemistry*, 39(1):27–36.
- Hess, S. T., Girirajan, T. P. K., and Mason, M. D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical journal*, 91(11):4258–72.
- Heuck, A., Schleiffer, A., and Clausen, T. (2011). Augmenting β -augmentation: structural basis of how BamB binds BamA and may support folding of outer membrane proteins. *Journal of molecular biology*, 406(5):659–66.
- Hibbing, M. E., Fuqua, C., Parsek, M. R., and Peterson, S. B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nature reviews. Microbiology*, 8(1):15–25.
- Hirai, Y., Haque, M., Yoshida, T., Yokota, K., Yasuda, T., and Oguma, K. (1995). Unique cholesteryl glucosides in *Helicobacter pylori*: composition and structural analysis. *Journal of bacteriology*, 177(18):5327–33.

- Hoang, H. H., Nickerson, N. N., Lee, V. T., Kazimirova, A., Chami, M., Pugsley, A. P., and Lory, S. (2011). Outer membrane targeting of *Pseudomonas aeruginosa* proteins shows variable dependence on the components of Bam and Lol machineries. *mBio*, 2(6):1–8.
- Hoffman, D. P., Shtengel, G., Xu, C. S., Campbell, K. R., Freeman, M., Wang, L., Milkie, D. E., Pasolli, H. A., Iyer, N., Bogovic, J. A., Stabley, D. R., Shirinifard, A., Pang, S., Peale, D., Schaefer, K., Pomp, W., Chang, C.-L., Lippincott-Schwartz, J., Kirchhausen, T., Solecki, D. J., Betzig, E., and Hess, H. F. (2019). Correlative three-dimensional super-resolution and block face electron microscopy of whole vitreously frozen cells. *bioRxiv*.
- Hoffmann, A., Nettels, D., Clark, J., Borgia, A., Radford, S. E., Clarke, J., and Schuler, B. (2011). Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: recurrence analysis of single particles (RASP). *Physical chemistry chemical physics : PCCP*, 13(5):1857–71.
- Hoffmann, C., Gaietta, G., Zürn, A., Adams, S. R., Terrillon, S., Ellisman, M. H., Tsien, R. Y., and Lohse, M. J. (2010). Fluorescent labeling of tetracysteine-tagged proteins in intact cells. *Nature protocols*, 5(10):1666–77.
- Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J. M., and Engelhardt, H. (2008). Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proceedings of the National Academy of Sciences of the United States of America*, 105(10):3963–7.
- Höhr, A. I. C., Lindau, C., Wirth, C., Qiu, J., Stroud, D. A., Kutik, S., Guiard, B., Hunte, C., Becker, T., Pfanner, N., and Wiedemann, N. (2018). Membrane protein insertion through a mitochondrial β -barrel gate. *Science (New York, N.Y.)*, 359(6373):eaah6834.
- Holdbrook, D. A., Burmann, B. M., Huber, R. G., Petoukhov, M. V., Svergun, D. I., Hiller, S., and Bond, P. J. (2017). A Spring-Loaded Mechanism Governs the Clamp-like Dynamics of the Skp Chaperone. *Structure (London, England : 1993)*, 25(7):1079–1088.e3.
- Hong, H., Park, S., Jiménez, R. H. F., Rinehart, D., and Tamm, L. K. (2007). Role of aromatic side chains in the folding and thermodynamic stability of integral membrane proteins. *Journal of the American Chemical Society*, 129(26):8320–7.
- Hong, H., Szabo, G., and Tamm, L. K. (2006). Electrostatic couplings in OmpA ion-channel gating suggest a mechanism for pore opening. *Nature chemical biology*, 2(11):627–35.
- Hong, H. and Tamm, L. K. (2004). Elastic coupling of integral membrane protein stability to lipid bilayer forces. *Proceedings of the National Academy of Sciences of the United States of America*, 101(12):4065–70.
- Hønger, T., Jørgensen, K., Biltonen, R. L., and Mouritsen, O. G. (1996). Systematic relationship between phospholipase A2 activity and dynamic lipid bilayer microheterogeneity. *Biochemistry*, 35(28):9003–6.

- Horne, J. E., Walko, M., Calabrese, A. N., Levenstein, M. A., Brockwell, D. J., Kapur, N., Wilson, A. J., and Radford, S. E. (2018). Rapid Mapping of Protein Interactions Using Tag-Transfer Photocrosslinkers. *Angewandte Chemie (International ed. in English)*, 57(51):16688–16692.
- Hosoya, T., Hiramatsu, T., Ikemoto, T., Nakanishi, M., Aoyama, H., Hosoya, A., Iwata, T., Maruyama, K., Endo, M., and Suzuki, M. (2004). Novel bifunctional probe for radioisotope-free photoaffinity labeling: compact structure comprised of photospecific ligand ligation and detectable tag anchoring units. *Organic & biomolecular chemistry*, 2(5):637–41.
- Hu, B., Lara-Tejero, M., Kong, Q., Galán, J. E., and Liu, J. (2017). In Situ Molecular Architecture of the Salmonella Type III Secretion Machine. *Cell*, 168(6):1065–1074.e10.
- Huang, W., Ling, Q., Bédard, J., Lilley, K., and Jarvis, P. (2011). In vivo analyses of the roles of essential Omp85-related proteins in the chloroplast outer envelope membrane. *Plant physiology*, 157(1):147–59.
- Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., Butterfield, C. N., Hermsdorf, A. W., Amano, Y., Ise, K., Suzuki, Y., Dudek, N., Relman, D. A., Finstad, K. M., Amundson, R., Thomas, B. C., and Banfield, J. F. (2016). A new view of the tree of life. *Nature microbiology*, 1(5):16048.
- Hulleman, C. N., Li, W., Gregor, I., Rieger, B., and Enderlein, J. (2018). Photon Yield Enhancement of Red Fluorophores at Cryogenic Temperatures. *Chemphyschem : a European journal of chemical physics and physical chemistry*, 19(14):1774–1780.
- Humes, J. R., Schiffrin, B., Calabrese, A. N., Higgins, A. J., Westhead, D. R., Brockwell, D. J., and Radford, S. E. (2019). The Role of SurA PPIase Domains in Preventing Aggregation of the Outer-Membrane Proteins tOmpA and OmpT. *Journal of molecular biology*, 431(6):1267–1283.
- Humphries, H. E., Christodoulides, M., and Heckels, J. E. (2002). Expression of the class 1 outer-membrane protein of *Neisseria meningitidis* in *Escherichia coli* and purification using a self-cleavable affinity tag. *Protein expression and purification*, 26(2):243–8.
- Hussain, S. and Bernstein, H. D. (2018). The Bam complex catalyzes efficient insertion of bacterial outer membrane proteins into membrane vesicles of variable lipid composition. *The Journal of biological chemistry*, 293(8):2959–2973.
- Huysmans, G. H., Guilvout, I., Chami, M., Nickerson, N. N., and Pugsley, A. P. (2015). Lipids assist the membrane insertion of a BAM-independent outer membrane protein. *Scientific Reports*, 5:1–15.
- Huysmans, G. H. M., Baldwin, S. A., Brockwell, D. J., and Radford, S. E. (2010). The transition state for folding of an outer membrane protein. *Proceedings of the National Academy of Sciences of the United States of America*, 107(9):4099–104.

- Huysmans, G. H. M., Radford, S. E., Baldwin, S. A., and Brockwell, D. J. (2012). Malleability of the folding mechanism of the outer membrane protein PagP: parallel pathways and the effect of membrane elasticity. *Journal of molecular biology*, 416(3):453–64.
- Iacobucci, C., Götze, M., Piotrowski, C., Arlt, C., Rehkamp, A., Ihling, C., Hage, C., and Sinz, A. (2018). Carboxyl-Photo-Reactive MS-Cleavable Cross-Linkers: Unveiling a Hidden Aspect of Diazirine-Based Reagents. *Analytical chemistry*, 90(4):2805–2809.
- Iadanza, M. G., Higgins, A. J., Schiffrin, B., Calabrese, A. N., Brockwell, D. J., Ashcroft, A. E., Radford, S. E., and Ranson, N. A. (2016). Lateral opening in the intact β -barrel assembly machinery captured by cryo-EM. *Nature communications*, 7(1):12865.
- Ieva, R. and Bernstein, H. D. (2009). Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane. *Proceedings of the National Academy of Sciences of the United States of America*, 106(45):19120–5.
- Ieva, R., Skillman, K. M., and Bernstein, H. D. (2008). Incorporation of a polypeptide segment into the β -domain pore during the assembly of a bacterial autotransporter. *Molecular Microbiology*, 67(1):188–201.
- Ieva, R., Tian, P., Peterson, J. H., and Bernstein, H. D. (2011). Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain. *Proceedings of the National Academy of Sciences of the United States of America*, 108(31):E383–91.
- Ingargiola, A., Laurence, T., Boutelle, R., Weiss, S., and Michalet, X. (2016a). Photon-HDF5: An Open File Format for Timestamp-Based Single-Molecule Fluorescence Experiments. *Biophysical journal*, 110(1):26–33.
- Ingargiola, A., Lerner, E., Chung, S., Weiss, S., and Michalet, X. (2016b). FRET-Bursts: An Open Source Toolkit for Analysis of Freely-Diffusing Single-Molecule FRET. *PLoS one*, 11(8):e0160716.
- Inoue, K. and Potter, D. (2004). The chloroplastic protein translocation channel Toc75 and its paralog OEP80 represent two distinct protein families and are targeted to the chloroplastic outer envelope by different mechanisms. *The Plant journal : for cell and molecular biology*, 39(3):354–65.
- Iqbal, H., Kenedy, M. R., Lybecker, M., and Akins, D. R. (2016). The TamB ortholog of *Borrelia burgdorferi* interacts with the β -barrel assembly machine (BAM) complex protein BamA. *Molecular microbiology*, 102(5):757–774.
- Ishida, H., Garcia-Herrero, A., and Vogel, H. J. (2014). The periplasmic domain of *Escherichia coli* outer membrane protein A can undergo a localized temperature dependent structural transition. *Biochimica et biophysica acta*, 1838(12):3014–24.
- Iyer, B. R. and Mahalakshmi, R. (2016). Distinct Structural Elements Govern the Folding, Stability, and Catalysis in the Outer Membrane Enzyme PagP. *Biochemistry*, 55(35):4960–70.

- Iyer, B. R. and Mahalakshmi, R. (2019). Hydrophobic Characteristic Is Energetically Preferred for Cysteine in a Model Membrane Protein. *Biophysical journal*, pages 1–11.
- Iyer, B. R., Vetal, P. V., Noordeen, H., Zadafiya, P., and Mahalakshmi, R. (2018). Salvaging the Thermodynamic Destabilization of Interface Histidine in Transmembrane β -Barrels. *Biochemistry*, 57(48):6669–6678.
- Iyer, B. R., Zadafiya, P., Vetal, P. V., and Mahalakshmi, R. (2017). Energetics of side-chain partitioning of β -signal residues in unassisted folding of a transmembrane β -barrel protein. *The Journal of biological chemistry*, 292(29):12351–12365.
- Jackups, R. and Liang, J. (2005). Interstrand pairing patterns in beta-barrel membrane proteins: the positive-outside rule, aromatic rescue, and strand registration prediction. *Journal of molecular biology*, 354(4):979–93.
- Jain, S., van Ulsen, P., Benz, I., Schmidt, M. A., Fernandez, R., Tommassen, J., and Goldberg, M. B. (2006). Polar localization of the autotransporter family of large bacterial virulence proteins. *Journal of bacteriology*, 188(13):4841–50.
- Jansen, K. B., Baker, S. L., and Sousa, M. C. (2012). Crystal structure of BamB from *Pseudomonas aeruginosa* and functional evaluation of its conserved structural features. *PloS one*, 7(11):e49749.
- Jarchow, S., Lück, C., Görg, A., and Skerra, A. (2008). Identification of potential substrate proteins for the periplasmic *Escherichia coli* chaperone Skp. *Proteomics*, 8(23-24):4987–94.
- Jarosławski, S., Duquesne, K., Sturgis, J. N., and Scheuring, S. (2009). High-resolution architecture of the outer membrane of the Gram-negative bacteria *Roseobacter denitrificans*. *Molecular microbiology*, 74(5):1211–22.
- Jeschke, G. (2018). MMM: A toolbox for integrative structure modeling. *Protein science : a publication of the Protein Society*, 27(1):76–85.
- Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., and Yang, S. (2015). Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Applied and environmental microbiology*, 81(7):2506–14.
- Joh, N. H., Min, A., Faham, S., Whitelegge, J. P., Yang, D., Woods, V. L., and Bowie, J. U. (2008). Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature*, 453(7199):1266–70.
- Jumper, C. C., Bomgarden, R., Rogers, J., Etienne, C., and Schriemer, D. C. (2012). High-resolution mapping of carbene-based protein footprints. *Analytical chemistry*, 84(10):4411–8.
- Jumper, C. C. and Schriemer, D. C. (2011). Mass spectrometry of laser-initiated carbene reactions for protein topographic analysis. *Analytical chemistry*, 83(8):2913–20.
- Jurkiewicz, P., Cwiklik, L., Jungwirth, P., and Hof, M. (2012). Lipid hydration and mobility: an interplay between fluorescence solvent relaxation experiments and molecular dynamics simulations. *Biochimie*, 94(1):26–32.

- Justice, S. S., Lauer, S. R., Hultgren, S. J., and Hunstad, D. A. (2006). Maturation of intracellular *Escherichia coli* communities requires SurA. *Infection and immunity*, 74(8):4793–800.
- Kadokura, H. and Beckwith, J. (2010). Mechanisms of oxidative protein folding in the bacterial cell envelope. *Antioxidants & redox signaling*, 13(8):1231–46.
- Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P. J., Berger, S., Restle, T., Goody, R. S., Gohlke, H., and Seidel, C. A. M. (2012). A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nature methods*, 9(12):1218–25.
- Kallmeyer, J., Pockalny, R., Adhikari, R. R., Smith, D. C., and D'Hondt, S. (2012). Global distribution of microbial abundance and biomass in subseafloor sediment. *Proceedings of the National Academy of Sciences of the United States of America*, 109(40):16213–6.
- Kang, G., López-Peña, I., Oklejas, V., Gary, C. S., Cao, W., and Kim, J. E. (2012). Förster resonance energy transfer as a probe of membrane protein folding. *Biochimica et biophysica acta*, 1818(2):154–61.
- Kang, M.-G., Lee, H., Kim, B. H., Dunbayev, Y., Seo, J. K., Lee, C., and Rhee, H.-W. (2017). Structure-guided synthesis of a protein-based fluorescent sensor for alkyl halides. *Chemical communications (Cambridge, England)*, 53(66):9226–9229.
- Kapanidis, A. N., Laurence, T. A., Lee, N. K., Margeat, E., Kong, X., and Weiss, S. (2005). Alternating-laser excitation of single molecules. *Accounts of chemical research*, 38(7):523–33.
- Kapanidis, A. N., Lee, N. K., Laurence, T. A., Doose, S., Margeat, E., and Weiss, S. (2004). Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 101(24):8936–41.
- Katira, S., Mandadapu, K. K., Vaikuntanathan, S., Smit, B., and Chandler, D. (2016). Pre-transition effects mediate forces of assembly between transmembrane proteins. *eLife*, 5(FEBRUARY2016):e13150.
- Kaufmann, R., Schellenberger, P., Seiradake, E., Dobbie, I. M., Jones, E. Y., Davis, I., Hagen, C., and Grünwald, K. (2014). Super-resolution microscopy using standard fluorescent proteins in intact cells under cryo-conditions. *Nano letters*, 14(7):4171–5.
- Kaur, H., Hartmann, J.-B., Jakob, R. P., Zahn, M., Zimmermann, I., Maier, T., Seeger, M. A., and Hiller, S. (2019). Identification of conformation-selective nanobodies against the membrane protein insertase BamA by an integrated structural biology approach. *Journal of biomolecular NMR*, (0123456789).
- Kazmierczak, B. I., Mielke, D. L., Russel, M., and Model, P. (1994). pIV, a filamentous phage protein that mediates phage export across the bacterial cell envelope, forms a multimer. *Journal of molecular biology*, 238(2):187–98.

- Ke, N., Landgraf, D., Paulsson, J., and Berkmen, M. (2016). Visualization of Periplasmic and Cytoplasmic Proteins with a Self-Labeling Protein Tag. *Journal of bacteriology*, 198(7):1035–43.
- Kerényi, M., Allison, H. E., Bártai, I., Sonnevend, A., Emödy, L., Plaveczyk, N., and Pál, T. (2005). Occurrence of hlyA and sheA genes in extraintestinal Escherichia coli strains. *Journal of clinical microbiology*, 43(6):2965–8.
- Keseler, I. M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S., Kothari, A., Krummenacker, M., Latendresse, M., Muñoz-Rascado, L., Ong, Q., Paley, S., Peralta-Gil, M., Subhraveti, P., Velázquez-Ramírez, D. A., Weaver, D., Collado-Vides, J., Paulsen, I., and Karp, P. D. (2017). The EcoCyc database: reflecting new knowledge about Escherichia coli K-12. *Nucleic acids research*, 45(D1):D543–D550.
- Kessler, R. J. and Fanestil, D. D. (1986). Interference by lipids in the determination of protein using bicinchoninic acid. *Analytical biochemistry*, 159(1):138–42.
- Kikuchi, N., Fujiwara, K., and Ikeguchi, M. (2018). β -Strand twisting/bending in soluble and transmembrane β -barrel structures. *Proteins*, 86(12):1231–1241.
- Kikuchi, S., Shibuya, I., and Matsumoto, K. (2000). Viability of an Escherichia coli pgsA null mutant lacking detectable phosphatidylglycerol and cardiolipin. *Journal of bacteriology*, 182(2):371–6.
- Kim, K. H., Aulakh, S., and Paetzel, M. (2011a). Crystal structure of β -barrel assembly machinery BamCD protein complex. *The Journal of biological chemistry*, 286(45):39116–21.
- Kim, K. H., Aulakh, S., Tan, W., and Paetzel, M. (2011b). Crystallographic analysis of the C-terminal domain of the Escherichia coli lipoprotein BamC. *Acta crystallographica. Section F, Structural biology and crystallization communications*, 67(Pt 11):1350–8.
- Kim, K. H., Kang, H.-S., Okon, M., Escobar-Cabrera, E., McIntosh, L. P., and Paetzel, M. (2011c). Structural characterization of Escherichia coli BamE, a lipoprotein component of the β -barrel assembly machinery complex. *Biochemistry*, 50(6):1081–90.
- Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C., and Kahne, D. (2007). Structure and function of an essential component of the outer membrane protein assembly machine. *Science (New York, N.Y.)*, 317(5840):961–4.
- Kim, S., Patel, D. S., Park, S., Slusky, J., Klauda, J. B., Widmalm, G., and Im, W. (2016a). Bilayer Properties of Lipid A from Various Gram-Negative Bacteria. *Biophysical journal*, 111(8):1750–1760.
- Kim, S. J., Fernandez-Martinez, J., Nudelman, I., Shi, Y., Zhang, W., Raveh, B., Herricks, T., Slaughter, B. D., Hogan, J. A., Upla, P., Chemmama, I. E., Pellarin, R., Echeverria, I., Shivaraju, M., Chaudhury, A. S., Wang, J., Williams, R., Unruh, J. R., Greenberg, C. H., Jacobs, E. Y., Yu, Z., de la Cruz, M. J., Mironska, R., Stokes, D. L., Aitchison, J. D., Jarrold, M. F., Gerton, J. L., Ludtke, S. J., Akey, C. W., Chait, B. T., Sali, A., and Rout, M. P. (2018). Integrative structure and functional anatomy of a nuclear pore complex. *Nature*, 555(7697):475–482.

- Kim, Y., Kim, W., and Park, J. W. (2016b). Principles and Applications of Force Spectroscopy Using Atomic Force Microscopy. *Bulletin of the Korean Chemical Society*, 37(12):1895–1907.
- Kiskowski, M. A., Hancock, J. F., and Kenworthy, A. K. (2009). On the use of Ripley's K-function and its derivatives to analyze domain size. *Biophysical journal*, 97(4):1095–103.
- Klein, K., Sonnabend, M. S., Frank, L., Leibiger, K., Franz-Wachtel, M., Macek, B., Trunk, T., Leo, J. C., Autenrieth, I. B., Schütz, M., and Bohn, E. (2019). Deprivation of the Periplasmic Chaperone SurA Reduces Virulence and Restores Antibiotic Susceptibility of Multidrug-Resistant *Pseudomonas aeruginosa*. *Frontiers in microbiology*, 10(FEB):100.
- Kleinschmidt, J. H. (2003). Membrane protein folding on the example of outer membrane protein A of *Escherichia coli*. *Cellular and molecular life sciences : CMLS*, 60(8):1547–58.
- Kleinschmidt, J. H. (2015). Folding of β -barrel membrane proteins in lipid bilayers - Unassisted and assisted folding and insertion. *Biochimica et biophysica acta*, 1848(9):1927–43.
- Kleinschmidt, J. H., Bulieris, P. V., Qu, J., Dogterom, M., and den Blaauwen, T. (2011). Association of neighboring β -strands of outer membrane protein A in lipid bilayers revealed by site-directed fluorescence quenching. *Journal of molecular biology*, 407(2):316–32.
- Kleinschmidt, J. H., den Blaauwen, T., Driessen, A. J., and Tamm, L. K. (1999). Outer membrane protein A of *Escherichia coli* inserts and folds into lipid bilayers by a concerted mechanism. *Biochemistry*, 38(16):5006–16.
- Kleinschmidt, J. H. and Tamm, L. K. (1996). Folding intermediates of a beta-barrel membrane protein. Kinetic evidence for a multi-step membrane insertion mechanism. *Biochemistry*, 35(40):12993–3000.
- Kleinschmidt, J. H. and Tamm, L. K. (1999). Time-resolved distance determination by tryptophan fluorescence quenching: probing intermediates in membrane protein folding. *Biochemistry*, 38(16):4996–5005.
- Kleinschmidt, J. H. and Tamm, L. K. (2002). Secondary and tertiary structure formation of the beta-barrel membrane protein OmpA is synchronized and depends on membrane thickness. *Journal of molecular biology*, 324(2):319–30.
- Knowles, T. J., Browning, D. F., Jeeves, M., Maderbocus, R., Rajesh, S., Sridhar, P., Manoli, E., Emery, D., Sommer, U., Spencer, A., Leyton, D. L., Squire, D., Chaudhuri, R. R., Viant, M. R., Cunningham, A. F., Henderson, I. R., and Overduin, M. (2011). Structure and function of BamE within the outer membrane and the β -barrel assembly machine. *EMBO reports*, 12(2):123–8.
- Knowles, T. J., Jeeves, M., Bobat, S., Dancea, F., McClelland, D., Palmer, T., Overduin, M., and Henderson, I. R. (2008). Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Molecular microbiology*, 68(5):1216–27.

- Knowles, T. J., McClelland, D. M., Rajesh, S., Henderson, I. R., and Overduin, M. (2009). Secondary structure and (1)H, (13)C and (15)N backbone resonance assignments of BamC, a component of the outer membrane protein assembly machinery in *Escherichia coli*. *Biomolecular NMR assignments*, 3(2):203–6.
- Koch, H. G., Hengelage, T., Neumann-Haefelin, C., MacFarlane, J., Hoffschulte, H. K., Schimz, K. L., Mechler, B., and Müller, M. (1999). In vitro studies with purified components reveal signal recognition particle (SRP) and SecA/SecB as constituents of two independent protein-targeting pathways of *Escherichia coli*. *Molecular biology of the cell*, 10(7):2163–73.
- König, K., Uchugonova, A., and Breunig, H. G. (2014). High-resolution multiphoton cryomicroscopy. *Methods (San Diego, Calif.)*, 66(2):230–6.
- Konovalova, A., Grabowicz, M., Balibar, C. J., Malinverni, J. C., Painter, R. E., Riley, D., Mann, P. A., Wang, H., Garlisi, C. G., Sherborne, B., Rigel, N. W., Ricci, D. P., Black, T. A., Roemer, T., Silhavy, T. J., and Walker, S. S. (2018). Inhibitor of intramembrane protease RseP blocks the σ E response causing lethal accumulation of unfolded outer membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 115(28):E6614–E6621.
- Konovalova, A., Kahne, D. E., and Silhavy, T. J. (2017). Outer Membrane Biogenesis. *Annual review of microbiology*, 71(1):539–556.
- Konovalova, A., Mitchell, A. M., and Silhavy, T. J. (2016). A lipoprotein/ β -barrel complex monitors lipopolysaccharide integrity transducing information across the outer membrane. *eLife*, 5(JUNE2016):1–17.
- Konovalova, A., Perlman, D. H., Cowles, C. E., and Silhavy, T. J. (2014). Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of β -barrel proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 111(41):E4350–8.
- Konovalova, A. and Silhavy, T. J. (2015). Outer membrane lipoprotein biogenesis: Lol is not the end. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 370(1679).
- Korndörfer, I. P., Dommel, M. K., and Skerra, A. (2004). Structure of the periplasmic chaperone Skp suggests functional similarity with cytosolic chaperones despite differing architecture. *Nature structural & molecular biology*, 11(10):1015–20.
- Koynova, R. and Caffrey, M. (1998). Phases and phase transitions of the phosphatidylcholines. *Biochimica et biophysica acta*, 1376(1):91–145.
- Krojer, T., Sawa, J., Huber, R., and Clausen, T. (2010). HtrA proteases have a conserved activation mechanism that can be triggered by distinct molecular cues. *Nature structural & molecular biology*, 17(7):844–52.
- Krokowski, S., Atwal, S., Lobato-Márquez, D., Chastanet, A., Carballido-López, R., Salje, J., and Mostowy, S. (2019). *Shigella* MreB promotes polar IcsA positioning for actin tail formation. *Journal of cell science*, 132(9):1–9.

- Kucerka, N., Liu, Y., Chu, N., Petrache, H. I., Tristram-Nagle, S., and Nagle, J. F. (2005). Structure of fully hydrated fluid phase DMPC and DLPC lipid bilayers using X-ray scattering from oriented multilamellar arrays and from unilamellar vesicles. *Biophysical journal*, 88(4):2626–37.
- Kusters, R., de Vrije, T., Breukink, E., and de Kruijff, B. (1989). SecB protein stabilizes a translocation-competent state of purified prePhoE protein. *The Journal of biological chemistry*, 264(35):20827–30.
- Lambert, T. J. (2019). FPbase: a community-editable fluorescent protein database. *Nature methods*, 16(4):277–278.
- Landenmark, H. K. E., Forgan, D. H., and Cockell, C. S. (2015). An Estimate of the Total DNA in the Biosphere. *PLoS biology*, 13(6):e1002168.
- Lang, K. and Chin, J. W. (2014). Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chemical reviews*, 114(9):4764–806.
- Langhorst, M. F., Genisyuerek, S., and Stuermer, C. A. O. (2006). Accumulation of FlAsH/Lumio Green in active mitochondria can be reversed by beta-mercaptoethanol for specific staining of tetracysteine-tagged proteins. *Histochemistry and cell biology*, 125(6):743–7.
- Lauber, F., Deme, J. C., Lea, S. M., and Berks, B. C. (2018). Type 9 secretion system structures reveal a new protein transport mechanism. *Nature*, 564(7734):77–82.
- Lazar, S. W. and Kolter, R. (1996). SurA assists the folding of Escherichia coli outer membrane proteins. *Journal of bacteriology*, 178(6):1770–3.
- Lee, H. C. and Bernstein, H. D. (2001). The targeting pathway of Escherichia coli presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proceedings of the National Academy of Sciences of the United States of America*, 98(6):3471–6.
- Lee, J., Sutterlin, H. A., Wzorek, J. S., Mandler, M. D., Hagan, C. L., Grabowicz, M., Tomasek, D., May, M. D., Hart, E. M., Silhavy, T. J., and Kahne, D. (2018). Substrate binding to BamD triggers a conformational change in BamA to control membrane insertion. *Proceedings of the National Academy of Sciences of the United States of America*, 115(10):2359–2364.
- Lee, J., Xue, M., Wzorek, J. S., Wu, T., Grabowicz, M., Gronenberg, L. S., Sutterlin, H. A., Davis, R. M., Ruiz, N., Silhavy, T. J., and Kahne, D. E. (2016). Characterization of a stalled complex on the β -barrel assembly machine. *Proceedings of the National Academy of Sciences of the United States of America*, 113(31):8717–22.
- Lehman, K. M. and Grabowicz, M. (2019). Countering Gram-Negative Antibiotic Resistance: Recent Progress in Disrupting the Outer Membrane with Novel Therapeutics. *Antibiotics*, 8(4):163.
- Leitner, A., Faini, M., Stengel, F., and Aebersold, R. (2016). Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends in biochemical sciences*, 41(1):20–32.

- Lelek, M., Di Nunzio, F., Henriques, R., Charneau, P., Arhel, N., and Zimmer, C. (2012). Superresolution imaging of HIV in infected cells with FIAsh-PALM. *Proceedings of the National Academy of Sciences of the United States of America*, 109(22):8564–9.
- Lenhart, T. R., Kenedy, M. R., Yang, X., Pal, U., and Akins, D. R. (2012). BB0324 and BB0028 are constituents of the *Borrelia burgdorferi* β -barrel assembly machine (BAM) complex. *BMC microbiology*, 12:60.
- Leo, J. C., Grin, I., and Linke, D. (2012). Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 367(1592):1088–101.
- Lepore, A., Taylor, H., Landgraf, D., Okumus, B., Jaramillo-Riveri, S., McLaren, L., Bakshi, S., Paulsson, J., and Karoui, M. E. (2019). Quantification of very low-abundant proteins in bacteria using the HaloTag and epi-fluorescence microscopy. *Scientific reports*, 9(1):7902.
- Lessene, G., Czabotar, P. E., and Colman, P. M. (2008). BCL-2 family antagonists for cancer therapy. *Nature reviews. Drug discovery*, 7(12):989–1000.
- Lewenza, S., Vidal-Ingigliardi, D., and Pugsley, A. P. (2006). Direct visualization of red fluorescent lipoproteins indicates conservation of the membrane sorting rules in the family Enterobacteriaceae. *Journal of Bacteriology*, 188(10):3516–3524.
- Lewis, B. A. and Engelman, D. M. (1983). Lipid bilayer thickness varies linearly with acyl chain length in fluid phosphatidylcholine vesicles. *Journal of molecular biology*, 166(2):211–7.
- Lewis, C., Skovierova, H., Rowley, G., Rezuchova, B., Homerova, D., Stevenson, A., Sherry, A., Kormanec, J., and Roberts, M. (2008). Small outer-membrane lipoprotein, SmpA, is regulated by sigmaE and has a role in cell envelope integrity and virulence of *Salmonella enterica* serovar Typhimurium. *Microbiology (Reading, England)*, 154(Pt 3):979–88.
- Leyton, D. L., Johnson, M. D., Thapa, R., Huysmans, G. H. M., Dunstan, R. A., Celik, N., Shen, H.-H., Loo, D., Belousoff, M. J., Purcell, A. W., Henderson, I. R., Beddoe, T., Rossjohn, J., Martin, L. L., Strugnell, R. A., and Lithgow, T. (2014). A mortise-tenon joint in the transmembrane domain modulates autotransporter assembly into bacterial outer membranes. *Nature communications*, 5(1):4239.
- Li, G., He, C., Bu, P., Bi, H., Pan, S., Sun, R., and Zhao, X. S. (2018). Single-Molecule Detection Reveals Different Roles of Skp and SurA as Chaperones. *ACS chemical biology*, 13(4):1082–1089.
- Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell*, 157(3):624–35.
- Li, W., Stein, S. C., Gregor, I., and Enderlein, J. (2015). Ultra-stable and versatile widefield cryo-fluorescence microscope for single-molecule localization with sub-nanometer accuracy. *Optics express*, 23(3):3770–83.

- Liang, J., Zhang, L., Tan, X.-L., Qi, Y.-K., Feng, S., Deng, H., Yan, Y., Zheng, J.-S., Liu, L., and Tian, C.-L. (2017). Chemical Synthesis of Diubiquitin-Based Photoaffinity Probes for Selectively Profiling Ubiquitin-Binding Proteins. *Angewandte Chemie (International ed. in English)*, 56(10):2744–2748.
- Liga, A., Morton, J. A. S., and Kersaudy-Kerhoas, M. (2016). Safe and cost-effective rapid-prototyping of multilayer PMMA microfluidic devices. *Microfluidics and Nanofluidics*, 20(12):164.
- Lill, Y., Jordan, L. D., Smallwood, C. R., Newton, S. M., Lill, M. A., Klebba, P. E., and Ritchie, K. (2016). Confined Mobility of TonB and FepA in Escherichia coli Membranes. *PloS one*, 11(12):e0160862.
- Liu, B., Xue, Y., Zhao, W., Chen, Y., Fan, C., Gu, L., Zhang, Y., Zhang, X., Sun, L., Huang, X., Ding, W., Sun, F., Ji, W., and Xu, T. (2015). Three-dimensional super-resolution protein localization correlated with vitrified cellular context. *Scientific reports*, 5(October):13017.
- Liu, Q., Moldoveanu, T., Sprules, T., Matta-Camacho, E., Mansur-Azzam, N., and Gehring, K. (2010). Apoptotic regulation by MCL-1 through heterodimerization. *The Journal of biological chemistry*, 285(25):19615–24.
- Lomize, A. L., Pogozheva, I. D., Lomize, M. A., and Mosberg, H. I. (2006). Positioning of proteins in membranes: a computational approach. *Protein science : a publication of the Protein Society*, 15(6):1318–33.
- Loos, M. S., Ramakrishnan, R., Vranken, W., Tsirigotaki, A., Tsare, E.-P., Zorzini, V., Geyter, J. D., Yuan, B., Tsamardinos, I., Klappa, M., Schymkowitz, J., Rousseau, F., Karamanou, S., and Economou, A. (2019). Structural Basis of the Subcellular Topology Landscape of Escherichia coli. *Frontiers in Microbiology*, 10(July):1–22.
- Los, G. V., Encell, L. P., McDougall, M. G., Hartzell, D. D., Karassina, N., Zimprich, C., Wood, M. G., Learish, R., Ohana, R. F., Urh, M., Simpson, D., Mendez, J., Zimmerman, K., Otto, P., Vidugiris, G., Zhu, J., Darzins, A., Klaubert, D. H., Bulleit, R. F., and Wood, K. V. (2008). HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS chemical biology*, 3(6):373–82.
- Lugtenberg, B. and Van Alphen, L. (1983). Molecular architecture and functioning of the outer membrane of Escherichia coli and other gram-negative bacteria. *Biochimica et biophysica acta*, 737(1):51–115.
- Lundquist, K., Bakelar, J., Noinaj, N., and Gumbart, J. C. (2018). C-terminal kink formation is required for lateral gating in BamA. *Proceedings of the National Academy of Sciences of the United States of America*, 115(34):E7942–E7949.
- Luo, J., Cimermancic, P., Viswanath, S., Ebmeier, C. C., Kim, B., Dehecq, M., Raman, V., Greenberg, C. H., Pellarin, R., Sali, A., Taatjes, D. J., Hahn, S., and Ranish, J. (2015). Architecture of the Human and Yeast General Transcription and DNA Repair Factor TFIIH. *Molecular cell*, 59(5):794–806.
- Lurie-Weinberger, M. N. and Gophna, U. (2015). Archaea in and on the Human Body: Health Implications and Future Directions. *PLoS pathogens*, 11(6):e1004833.

- Lycklama A Nijeholt, J. A. and Driessen, A. J. M. (2012). The bacterial Sec-translocase: structure and mechanism. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 367(1592):1016–28.
- Lyon, R. P., Setter, J. R., Bovee, T. D., Doronina, S. O., Hunter, J. H., Anderson, M. E., Balasubramanian, C. L., Duniho, S. M., Leiske, C. I., Li, F., and Senter, P. D. (2014). Self-hydrolyzing maleimides improve the stability and pharmacological properties of antibody-drug conjugates. *Nature biotechnology*, 32(10):1059–62.
- Lyu, Z.-X., Shao, Q., Gao, Y. Q., and Zhao, X. S. (2012). Direct observation of the uptake of outer membrane proteins by the periplasmic chaperone Skp. *PloS one*, 7(9):e46068.
- Ma, H., Khan, A., and Nangia, S. (2018). Dynamics of OmpF Trimer Formation in the Bacterial Outer Membrane of Escherichia coli. *Langmuir : the ACS journal of surfaces and colloids*, 34(19):5623–5634.
- Machutta, C. A., Kollmann, C. S., Lind, K. E., Bai, X., Chan, P. F., Huang, J., Ballell, L., Belyanskaya, S., Besra, G. S., Barros-Aguirre, D., Bates, R. H., Centrella, P. A., Chang, S. S., Chai, J., Choudhry, A. E., Coffin, A., Davie, C. P., Deng, H., Deng, J., Ding, Y., Dodson, J. W., Fosbenner, D. T., Gao, E. N., Graham, T. L., Graybill, T. L., Ingraham, K., Johnson, W. P., King, B. W., Kwiatkowski, C. R., Lelièvre, J., Li, Y., Liu, X., Lu, Q., Lehr, R., Mendoza-Losana, A., Martin, J., McCloskey, L., McCormick, P., O’Keefe, H. P., O’Keefe, T., Pao, C., Phelps, C. B., Qi, H., Rafferty, K., Scavello, G. S., Steinginga, M. S., Sundersingh, F. S., Sweitzer, S. M., Szewczuk, L. M., Taylor, A., Toh, M. F., Wang, J., Wang, M., Wilkins, D. J., Xia, B., Yao, G., Zhang, J., Zhou, J., Donahue, C. P., Messer, J. A., Holmes, D., Arico-Muendel, C. C., Pope, A. J., Gross, J. W., and Evindar, G. (2017). Prioritizing multiple therapeutic targets in parallel using automated DNA-encoded library screening. *Nature communications*, 8(May 2017):16081.
- MacKinnon, A. L., Garrison, J. L., Hegde, R. S., and Taunton, J. (2007). Photo-leucine incorporation reveals the target of a cyclodepsipeptide inhibitor of cotranslational translocation. *Journal of the American Chemical Society*, 129(47):14560–1.
- Maier, T., Clantin, B., Gruss, F., Dewitte, F., Delattre, A.-S., Jacob-Dubuisson, F., Hiller, S., and Villeret, V. (2015). Conserved Omp85 lid-lock structure and substrate recognition in FhaC. *Nature communications*, 6(May):7452.
- Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R., and Silhavy, T. J. (2006). YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in Escherichia coli. *Molecular microbiology*, 61(1):151–64.
- Manzi, L., Barrow, A. S., Scott, D., Layfield, R., Wright, T. G., Moses, J. E., and Oldham, N. J. (2016). Carbene footprinting accurately maps binding sites in protein-ligand and protein-protein interactions. *Nature communications*, 7(1):13288.
- Marsh, D. (2008). Energetics of hydrophobic matching in lipid-protein interactions. *Biophysical journal*, 94(10):3996–4013.
- Marsh, D., Shanmugavadivu, B., and Kleinschmidt, J. H. (2006). Membrane elastic fluctuations and the insertion and tilt of beta-barrel proteins. *Biophysical journal*, 91(1):227–32.

- Martinez-Hackert, E. and Hendrickson, W. A. (2009). Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone. *Cell*, 138(5):923–34.
- Marty, M. T., Baldwin, A. J., Marklund, E. G., Hochberg, G. K. A., Benesch, J. L. P., and Robinson, C. V. (2015). Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Analytical chemistry*, 87(8):4370–6.
- Marx, D. C. and Fleming, K. G. (2017). Influence of Protein Scaffold on Side-Chain Transfer Free Energies. *Biophysical journal*, 113(3):597–604.
- Mas, G. and Hiller, S. (2018). Conformational plasticity of molecular chaperones involved in periplasmic and outer membrane protein folding. *FEMS microbiology letters*, 365(13):1–9.
- Matern, Y., Barion, B., and Behrens-Kneip, S. (2010). PpiD is a player in the network of periplasmic chaperones in Escherichia coli. *BMC microbiology*, 10:251.
- Matos, M. J., Oliveira, B. L., Martínez-Sáez, N., Guerreiro, A., Cal, P. M. S. D., Bertoldo, J., Maneiro, M., Perkins, E., Howard, J., Deery, M. J., Chalker, J. M., Corzana, F., Jiménez-Osés, G., and Bernardes, G. J. L. (2018). Chemo- and Regioselective Lysine Modification on Native Proteins. *Journal of the American Chemical Society*, 140(11):4004–4017.
- Matsumoto, K. (2001). Dispensable nature of phosphatidylglycerol in Escherichia coli: dual roles of anionic phospholipids. *Molecular microbiology*, 39(6):1427–33.
- May, K. L. and Grabowicz, M. (2018a). The bacterial outer membrane is an evolving antibiotic barrier. *Proceedings of the National Academy of Sciences of the United States of America*, 115(36):8852–8854.
- May, K. L. and Grabowicz, M. (2018b). The bacterial outer membrane is an evolving antibiotic barrier. *Proceedings of the National Academy of Sciences of the United States of America*, 115(36):8852–8854.
- Mazeres, S., Joly, E., Lopez, A., and Tardin, C. (2014). Characterization of M-laurdan, a versatile probe to explore order in lipid membranes. *F1000Research*, 3:172.
- McCabe, A. L., Ricci, D., Adetunji, M., and Silhavy, T. J. (2017). Conformational Changes That Coordinate the Activity of BamA and BamD Allowing β -Barrel Assembly. *Journal of bacteriology*, 199(20):1–11.
- McDonald, S. K. and Fleming, K. G. (2016). Aromatic Side Chain Water-to-Lipid Transfer Free Energies Show a Depth Dependence across the Membrane Normal. *Journal of the American Chemical Society*, 138(25):7946–50.
- McMorran, L. M., Bartlett, A. I., Huysmans, G. H. M., Radford, S. E., and Brockwell, D. J. (2013). Dissecting the effects of periplasmic chaperones on the in vitro folding of the outer membrane protein PagP. *Journal of molecular biology*, 425(17):3178–91.

- Méheust, R., Burstein, D., Castelle, C. J., and Banfield, J. F. (2019). The distinction of CPR bacteria from other bacteria based on protein family content. *Nature communications*, 10(1):4173.
- Meiresonne, N. Y., Consoli, E., Mertens, L. M. Y., Chertkova, A. O., Goedhart, J., and den Blaauwen, T. (2019). Superfolder mTurquoise2ox optimized for the bacterial periplasm allows high efficiency in vivo FRET of cell division antibiotic targets. *Molecular microbiology*.
- Meiresonne, N. Y., van der Ploeg, R., Hink, M. A., and den Blaauwen, T. (2017). Activity-Related Conformational Changes in d,d-Carboxypeptidases Revealed by In Vivo Periplasmic Förster Resonance Energy Transfer Assay in *Escherichia coli*. *mBio*, 8(5):1–18.
- Merkel, J. S. and Regan, L. (1998). Aromatic rescue of glycine in beta sheets. *Folding & design*, 3(6):449–55.
- Michalik, M., Orwick-Rydmark, M., Habeck, M., Alva, V., Arnold, T., and Linke, D. (2017). An evolutionarily conserved glycine-tyrosine motif forms a folding core in outer membrane proteins. *PloS one*, 12(8):e0182016.
- Miller, S. I., Ernst, R. K., and Bader, M. W. (2005). LPS, TLR4 and infectious disease diversity. *Nature reviews. Microbiology*, 3(1):36–46.
- Misra, R. (2015). Entry and exit of bacterial outer membrane proteins. *Trends in microbiology*, 23(8):452–4.
- Missiakas, D., Betton, J. M., and Raina, S. (1996). New components of protein folding in extracytoplasmic compartments of *Escherichia coli* SurA, FkpA and Skp/OmpH. *Molecular microbiology*, 21(4):871–84.
- Mitchell, A. M. and Silhavy, T. J. (2019). Envelope stress responses: balancing damage repair and toxicity. *Nature reviews. Microbiology*, 17(7):417–428.
- Moerner, W. E. and Orrit, M. (1999). Illuminating single molecules in condensed matter. *Science (New York, N.Y.)*, 283(5408):1670–6.
- Mogensen, J. E., Kleinschmidt, J. H., Schmidt, M. A., and Otzen, D. E. (2005). Misfolding of a bacterial autotransporter. *Protein science : a publication of the Protein Society*, 14(11):2814–27.
- Mogensen, J. E. and Otzen, D. E. (2005). Interactions between folding factors and bacterial outer membrane proteins. *Molecular microbiology*, 57(2):326–46.
- Moon, C. P. and Fleming, K. G. (2011). Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proceedings of the National Academy of Sciences of the United States of America*, 108(25):10174–7.
- Moon, C. P., Kwon, S., and Fleming, K. G. (2011). Overcoming hysteresis to attain reversible equilibrium folding for outer membrane phospholipase A in phospholipid bilayers. *Journal of molecular biology*, 413(2):484–94.

- Moon, C. P., Zaccai, N. R., Fleming, P. J., Gessmann, D., and Fleming, K. G. (2013). Membrane protein thermodynamic stability may serve as the energy sink for sorting in the periplasm. *Proceedings of the National Academy of Sciences of the United States of America*, 110(11):4285–90.
- Morein, S., Andersson, A., Rilfors, L., and Lindblom, G. (1996). Wild-type *Escherichia coli* cells regulate the membrane lipid composition in a "window" between gel and non-lamellar structures. *The Journal of biological chemistry*, 271(12):6801–9.
- Mori, K., Yamaguchi, K., Sakiyama, Y., Urabe, T., and Suzuki, K.-i. (2009). *Caldisericum exile* gen. nov., sp. nov., an anaerobic, thermophilic, filamentous bacterium of a novel bacterial phylum, *Caldiserica* phyl. nov., originally called the candidate phylum OP5, and description of *Caldiseriaceae* fam. nov., *Caldisericales* ord. no. *International journal of systematic and evolutionary microbiology*, 59(Pt 11):2894–8.
- Morton, R. E. and Evans, T. A. (1992). Modification of the bicinchoninic acid protein assay to eliminate lipid interference in determining lipoprotein protein content. *Analytical biochemistry*, 204(2):332–4.
- Mouritsen, O. G., Andresen, T. L., Halperin, A., Hansen, P. L., Jakobsen, A. F., Jensen, U. B., Jensen, M. O., Jørgensen, K., Kaasgaard, T., Leidy, C., Simonsen, A. C., Peters, G. H., and Weiss, M. (2006). Activation of interfacial enzymes at membrane surfaces. *Journal of physics. Condensed matter : an Institute of Physics journal*, 18(28):S1293–304.
- Muhammad, N., Dworeck, T., Fioroni, M., and Schwaneberg, U. (2011). Engineering of the *E. coli* outer membrane protein FhuA to overcome the hydrophobic mismatch in thick polymeric membranes. *Journal of nanobiotechnology*, 9:8.
- Murakami, K., Elmlund, H., Kalisman, N., Bushnell, D. A., Adams, C. M., Azubel, M., Elmlund, D., Levi-Kalisman, Y., Liu, X., Gibbons, B. J., Levitt, M., and Kornberg, R. D. (2013). Architecture of an RNA polymerase II transcription pre-initiation complex. *Science (New York, N.Y.)*, 342(6159):1238724.
- Murzyn, K., Róg, T., and Pasenkiewicz-Gierula, M. (2005). Phosphatidylethanolamine-phosphatidylglycerol bilayer as a model of the inner bacterial membrane. *Biophysical journal*, 88(2):1091–103.
- Nahmani, M., Lanahan, C., DeRosier, D., and Turrigiano, G. G. (2017). High-numerical-aperture cryogenic light microscopy for increased precision of super-resolution reconstructions. *Proceedings of the National Academy of Sciences of the United States of America*, 114(15):3832–3836.
- Nakamura, K. and Mizushima, S. (1976). Effects of heating in dodecyl sulfate solution on the conformation and electrophoretic mobility of isolated major outer membrane proteins from *Escherichia coli* K-12. *Journal of biochemistry*, 80(6):1411–22.
- Narita, S.-i., Masui, C., Suzuki, T., Dohmae, N., and Akiyama, Y. (2013). Protease homolog BepA (YfgC) promotes assembly and degradation of β -barrel membrane proteins in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38):E3612–21.

- Nassal, M. (1983). 4-(1-Azi-2,2,2-trifluoroethyl)benzoic Acid, a Highly Photolabile Carbene Generating Label Readily Fixable to Biochemical Agents. *Liebigs Annalen der Chemie*, 1983(9):1510–1523.
- Neugebauer, H., Herrmann, C., Kammer, W., Schwarz, G., Nordheim, A., and Braun, V. (2005). ExbBD-dependent transport of maltodextrins through the novel MalA protein across the outer membrane of *Caulobacter crescentus*. *Journal of bacteriology*, 187(24):8300–11.
- Ni, D., Wang, Y., Yang, X., Zhou, H., Hou, X., Cao, B., Lu, Z., Zhao, X., Yang, K., and Huang, Y. (2014). Structural and functional analysis of the β -barrel domain of BamA from *Escherichia coli*. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 28(6):2677–85.
- Nickerson, N. N., Abby, S. S., Rocha, E. P. C., Chami, M., and Pugsley, A. P. (2012). A Single Amino Acid Substitution Changes the Self-Assembly Status of a Type IV Piliation Secretin. *Journal of bacteriology*, 194(18):4951–8.
- Nielsen, L. K., Bjørnholm, T., and Mouritsen, O. G. (2000a). Fluctuations caught in the act. *Nature*, 404(6776):352.
- Nielsen, L. K., Vishnyakov, A., Jørgensen, K., Bjørnholm, T., and Mouritsen, O. G. (2000b). Nanometre-scale structure of fluid lipid membranes. *Journal of Physics: Condensed Matter*, 12(8A):A309–A314.
- Nierman, W. C., Feldblyum, T. V., Laub, M. T., Paulsen, I. T., Nelson, K. E., Eisen, J. A., Heidelberg, J. F., Alley, M. R., Ohta, N., Maddock, J. R., Potocka, I., Nelson, W. C., Newton, A., Stephens, C., Phadke, N. D., Ely, B., DeBoy, R. T., Dodson, R. J., Durkin, A. S., Gwinn, M. L., Haft, D. H., Kolonay, J. F., Smit, J., Craven, M. B., Khouri, H., Shetty, J., Berry, K., Utterback, T., Tran, K., Wolf, A., Vamathevan, J., Ermolaeva, M., White, O., Salzberg, S. L., Venter, J. C., Shapiro, L., Fraser, C. M., and Eisen, J. (2001). Complete genome sequence of *Caulobacter crescentus*. *Proceedings of the National Academy of Sciences of the United States of America*, 98(7):4136–41.
- Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and molecular biology reviews : MMBR*, 67(4):593–656.
- Noinaj, N., Fairman, J. W., and Buchanan, S. K. (2011). The crystal structure of BamB suggests interactions with BamA and its role within the BAM complex. *Journal of molecular biology*, 407(2):248–60.
- Noinaj, N., Guillier, M., Barnard, T. J., and Buchanan, S. K. (2010). TonB-dependent transporters: regulation, structure, and function. *Annual review of microbiology*, 64(1):43–60.
- Noinaj, N., Gumbart, J. C., and Buchanan, S. K. (2017). The β -barrel assembly machinery in motion. *Nature reviews. Microbiology*, 15(4):197–204.
- Noinaj, N., Kuszak, A. J., Balusek, C., Gumbart, J. C., and Buchanan, S. K. (2014). Lateral opening and exit pore formation are required for BamA function. *Structure (London, England : 1993)*, 22(7):1055–62.

- Noinaj, N., Kuszak, A. J., Gumbart, J. C., Lukacik, P., Chang, H., Easley, N. C., Lithgow, T., and Buchanan, S. K. (2013). Structural insight into the biogenesis of β -barrel membrane proteins. *Nature*, 501(7467):385–90.
- Noinaj, N., Rollauer, S. E., and Buchanan, S. K. (2015). The β -barrel membrane protein insertase machinery from Gram-negative bacteria. *Current opinion in structural biology*, 31:35–42.
- Norell, D., Heuck, A., Tran-Thi, T.-A., Götzke, H., Jacob-Dubuisson, F., Clausen, T., Daley, D. O., Braun, V., Müller, M., and Fan, E. (2014). Versatile in vitro system to study translocation and functional integration of bacterial outer membrane proteins. *Nature communications*, 5:5396.
- Oh, E., Becker, A. H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., Nichols, R. J., Typas, A., Gross, C. A., Kramer, G., Weissman, J. S., and Bukau, B. (2011). Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell*, 147(6):1295–308.
- Ohnishi, S. and Kameyama, K. (2001). Escherichia coli OmpA retains a folded structure in the presence of sodium dodecyl sulfate due to a high kinetic barrier to unfolding. *Biochimica et biophysica acta*, 1515(2):159–66.
- Okuda, S., Sherman, D. J., Silhavy, T. J., Ruiz, N., and Kahne, D. (2016). Lipopolysaccharide transport and assembly at the outer membrane: the PEZ model. *Nature reviews. Microbiology*, 14(6):337–45.
- O’Neil, P. K., Richardson, L. G. L., Paila, Y. D., Piszczek, G., Chakravarthy, S., Noinaj, N., and Schnell, D. (2017). The POTRA domains of Toc75 exhibit chaperone-like function to facilitate import into chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America*, 114(24):E4868–E4876.
- O’Neil, P. K., Rollauer, S. E., Noinaj, N., and Buchanan, S. K. (2015). Fitting the Pieces of the β -Barrel Assembly Machinery Complex. *Biochemistry*, 54(41):6303–11.
- Ortega, J., Iwanczyk, J., and Jomaa, A. (2009). Escherichia coli DegP: a structure-driven functional model. *Journal of bacteriology*, 191(15):4705–13.
- Ovesný, M., Křížek, P., Borkovec, J., Svindrych, Z., and Hagen, G. M. (2014). ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics (Oxford, England)*, 30(16):2389–90.
- Paetzel, M. (2014). Structure and mechanism of Escherichia coli type I signal peptidase. *Biochimica et biophysica acta*, 1843(8):1497–508.
- Paetzel, M., Dalbey, R. E., and Strynadka, N. C. J. (2002). Crystal structure of a bacterial signal peptidase apoenzyme: implications for signal peptide binding and the Ser-Lys dyad mechanism. *The Journal of biological chemistry*, 277(11):9512–9.
- Palomino, C., Marín, E., and Fernández, L. Á. (2011). The fimbrial usher FimD follows the SurA-BamB pathway for its assembly in the outer membrane of Escherichia coli. *Journal of bacteriology*, 193(19):5222–30.

- Paracini, N., Clifton, L. A., Skoda, M. W. A., and Lakey, J. H. (2018). Liquid crystalline bacterial outer membranes are critical for antibiotic susceptibility. *Proceedings of the National Academy of Sciences of the United States of America*, 115(32):E7587–E7594.
- Paramasivam, N., Habeck, M., and Linke, D. (2012). Is the C-terminal insertional signal in Gram-negative bacterial outer membrane proteins species-specific or not? *BMC genomics*, 13(1):510.
- Parasassi, T., De Stasio, G., Ravagnan, G., Rusch, R. M., and Gratton, E. (1991). Quantitation of lipid phases in phospholipid vesicles by the generalized polarization of Laurdan fluorescence. *Biophysical journal*, 60(1):179–89.
- Parks, D. H., Rinke, C., Chuvochina, M., Chaumeil, P.-A., Woodcroft, B. J., Evans, P. N., Hugenholtz, P., and Tyson, G. W. (2017). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nature microbiology*, 2(11):1533–1542.
- Parte, A. C. (2018). LPSN - List of Prokaryotic names with Standing in Nomenclature (bacterio.net), 20 years on. *International journal of systematic and evolutionary microbiology*, 68(6):1825–1829.
- Paschen, S. A., Neupert, W., and Rapaport, D. (2005). Biogenesis of beta-barrel membrane proteins of mitochondria. *Trends in biochemical sciences*, 30(10):575–82.
- Paschen, S. A., Waizenegger, T., Stan, T., Preuss, M., Cyrklaff, M., Hell, K., Rapaport, D., and Neupert, W. (2003). Evolutionary conservation of biogenesis of beta-barrel membrane proteins. *Nature*, 426(6968):862–6.
- Pasenkiewicz-Gierula, M., Baczynski, K., Markiewicz, M., and Murzyn, K. (2016). Computer modelling studies of the bilayer/water interface. *Biochimica et biophysica acta*, 1858(10):2305–2321.
- Patel, G. J., Behrens-Kneip, S., Holst, O., and Kleinschmidt, J. H. (2009). The periplasmic chaperone Skp facilitates targeting, insertion, and folding of OmpA into lipid membranes with a negative membrane surface potential. *Biochemistry*, 48(43):10235–45.
- Patel, G. J. and Kleinschmidt, J. H. (2013). The lipid bilayer-inserted membrane protein BamA of *Escherichia coli* facilitates insertion and folding of outer membrane protein A from its complex with Skp. *Biochemistry*, 52(23):3974–86.
- Patel, R., Hsu, S.-C., Bédard, J., Inoue, K., and Jarvis, P. (2008). The Omp85-related chloroplast outer envelope protein OEP80 is essential for viability in *Arabidopsis*. *Plant physiology*, 148(1):235–45.
- Pautsch, A. and Schulz, G. E. (2000). High-resolution structure of the OmpA membrane domain. *Journal of molecular biology*, 298(2):273–82.
- Pavlova, O., Peterson, J. H., Ieva, R., and Bernstein, H. D. (2013). Mechanistic link between β barrel assembly and the initiation of autotransporter secretion. *Proceedings of the National Academy of Sciences of the United States of America*, 110(10):E938–47.

- Pencer, J., Nieh, M.-P., Harroun, T. A., Krueger, S., Adams, C., and Katsaras, J. (2005). Bilayer thickness and thermal response of dimyristoylphosphatidylcholine unilamellar vesicles containing cholesterol, ergosterol and lanosterol: a small-angle neutron scattering study. *Biochimica et biophysica acta*, 1720(1-2):84–91.
- Pfützner, A.-K., Steblau, N., Ulrich, T., Oberhettinger, P., Autenrieth, I. B., Schütz, M., and Rapaport, D. (2016). Mitochondrial-bacterial hybrids of BamA/Tob55 suggest variable requirements for the membrane integration of β -barrel proteins. *Scientific reports*, 6(November):39053.
- Piknová, B., Marsh, D., and Thompson, T. E. (1997). Fluorescence quenching and electron spin resonance study of percolation in a two-phase lipid bilayer containing bacteriorhodopsin. *Biophysical journal*, 72(6):2660–8.
- Plummer, A. M. and Fleming, K. G. (2016). From Chaperones to the Membrane with a BAM! *Trends in biochemical sciences*, 41(10):872–882.
- Pocanschi, C. L., Apell, H.-J., Puntervoll, P., Høgh, B., Jensen, H. B., Welte, W., and Kleinschmidt, J. H. (2006a). The major outer membrane protein of *Fusobacterium nucleatum* (FomA) folds and inserts into lipid bilayers via parallel folding pathways. *Journal of molecular biology*, 355(3):548–61.
- Pocanschi, C. L., Patel, G. J., Marsh, D., and Kleinschmidt, J. H. (2006b). Curvature elasticity and refolding of OmpA in large unilamellar vesicles. *Biophysical journal*, 91(8):L75–7.
- Pocanschi, C. L., Popot, J.-L., and Kleinschmidt, J. H. (2013). Folding and stability of outer membrane protein A (OmpA) from *Escherichia coli* in an amphipathic polymer, amphipol A8-35. *European biophysics journal : EBJ*, 42(2-3):103–18.
- Politis, A. and Schmidt, C. (2018). Structural characterisation of medically relevant protein assemblies by integrating mass spectrometry with computational modelling. *Journal of proteomics*, 175:34–41.
- Poppleton, D. I., Duchateau, M., Hourdel, V., Matondo, M., Flechsler, J., Klingl, A., Beloin, C., and Gribaldo, S. (2017). Outer Membrane Proteome of *Veillonella parvula*: A Diderm Firmicute of the Human Microbiome. *Frontiers in microbiology*, 8(June):1215.
- Pourmousa, M. and Pastor, R. W. (2018). Molecular dynamics simulations of lipid nanodiscs. *Biochimica et biophysica acta. Biomembranes*, 1860(10):2094–2107.
- Prenner, E. J., Lewis, R. N., Kondejewski, L. H., Hodges, R. S., and McElhaney, R. N. (1999). Differential scanning calorimetric study of the effect of the antimicrobial peptide gramicidin S on the thermotropic phase behavior of phosphatidylcholine, phosphatidylethanolamine and phosphatidylglycerol lipid bilayer membranes. *Biochimica et biophysica acta*, 1417(2):211–23.
- Preston, G. W., Radford, S. E., Ashcroft, A. E., and Wilson, A. J. (2014). Analysis of amyloid nanostructures using photo-cross-linking: in situ comparison of three widely used photo-cross-linkers. *ACS chemical biology*, 9(3):761–8.
- Preston, G. W. and Wilson, A. J. (2013). Photo-induced covalent cross-linking for the analysis of biomolecular interactions. *Chemical Society reviews*, 42(8):3289–301.

- Psonis, J. J., Chahales, P., Henderson, N. S., Rigel, N. W., Hoffman, P. S., and Thanassi, D. G. (2019). The small molecule nitazoxanide selectively disrupts BAM-mediated folding of the outer membrane usher protein. *The Journal of biological chemistry*.
- Qu, J., Behrens-Kneip, S., Holst, O., and Kleinschmidt, J. H. (2009). Binding regions of outer membrane protein A in complexes with the periplasmic chaperone Skp. A site-directed fluorescence study. *Biochemistry*, 48(22):4926–36.
- Qu, J., Mayer, C., Behrens, S., Holst, O., and Kleinschmidt, J. H. (2007). The trimeric periplasmic chaperone Skp of *Escherichia coli* forms 1:1 complexes with outer membrane proteins via hydrophobic and electrostatic interactions. *Journal of molecular biology*, 374(1):91–105.
- Quan, S., Koldewey, P., Tapley, T., Kirsch, N., Ruane, K. M., Pfizenmaier, J., Shi, R., Hofmann, S., Foit, L., Ren, G., Jakob, U., Xu, Z., Cygler, M., and Bardwell, J. C. A. (2011). Genetic selection designed to stabilize proteins uncovers a chaperone called Spy. *Nature structural & molecular biology*, 18(3):262–9.
- Rabouw, F. T., Cogan, N. M. B., Berends, A. C., Stam, W. V. D., Vanmaekelbergh, D., Koenderink, A. F., Krauss, T. D., and Donega, C. D. M. (2016). Non-blinking single-photon emitters in silica. *Scientific reports*, 6(February):21187.
- Raivio, T. L., Leblanc, S. K. D., and Price, N. L. (2013). The *Escherichia coli* Cpx envelope stress response regulates genes of diverse function that impact antibiotic resistance and membrane integrity. *Journal of bacteriology*, 195(12):2755–67.
- Ranava, D., Caumont-Sarcos, A., Albenne, C., and Ieva, R. (2018). Bacterial machineries for the assembly of membrane-embedded β -barrel proteins. *FEMS microbiology letters*, 365(10):1–12.
- Rappsilber, J. (2011). The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *Journal of structural biology*, 173(3):530–40.
- Rassam, P., Copeland, N. A., Birkholz, O., Tóth, C., Chavent, M., Duncan, A. L., Cross, S. J., Housden, N. G., Kaminska, R., Seger, U., Quinn, D. M., Garrod, T. J., Sansom, M. S. P., Piehler, J., Baumann, C. G., and Kleanthous, C. (2015). Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria. *Nature*, 523(7560):333–6.
- Rath, P., Sharpe, T., Kohl, B., and Hiller, S. (2019). Two-State Folding of the Outer Membrane Protein X into a Lipid Bilayer Membrane. *Angewandte Chemie (International ed. in English)*, 58(9):2665–2669.
- Remmert, M., Biegert, A., Linke, D., Lupas, A. N., and Söding, J. (2010). Evolution of outer membrane beta-barrels from an ancestral beta beta hairpin. *Molecular biology and evolution*, 27(6):1348–58.
- Ricci, D. P., Hagan, C. L., Kahne, D., and Silhavy, T. J. (2012). Activation of the *Escherichia coli* β -barrel assembly machine (Bam) is required for essential components to interact properly with substrate. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9):3487–91.

- Ricci, D. P., Schwalm, J., Gonzales-Cope, M., and Silhavy, T. J. (2013). The activity and specificity of the outer membrane protein chaperone SurA are modulated by a proline isomerase domain. *mBio*, 4(4):1–9.
- Ricci, D. P. and Silhavy, T. J. (2019). Outer Membrane Protein Insertion by the β -barrel Assembly Machine. *EcoSal Plus*, 8(2):1–9.
- Richardson, J. S. and Richardson, D. C. (2002). Natural beta-sheet proteins use negative design to avoid edge-to-edge aggregation. *Proceedings of the National Academy of Sciences of the United States of America*, 99(5):2754–9.
- Rigel, N. W., Ricci, D. P., and Silhavy, T. J. (2013). Conformation-specific labeling of BamA and suppressor analysis suggest a cyclic mechanism for β -barrel assembly in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 110(13):5151–6.
- Rigel, N. W., Schwalm, J., Ricci, D. P., and Silhavy, T. J. (2012). BamE modulates the *Escherichia coli* beta-barrel assembly machine component BamA. *Journal of bacteriology*, 194(5):1002–8.
- Riske, K. A., Barroso, R. P., Vequi-Suplicy, C. C., Germano, R., Henriques, V. B., and Lamy, M. T. (2009). Lipid bilayer pre-transition as the beginning of the melting process. *Biochimica et biophysica acta*, 1788(5):954–63.
- Rizzitello, A. E., Harper, J. R., and Silhavy, T. J. (2001). Genetic evidence for parallel pathways of chaperone activity in the periplasm of *Escherichia coli*. *Journal of bacteriology*, 183(23):6794–800.
- Robert, V., Volokhina, E. B., Senf, F., Bos, M. P., Van Gelder, P., and Tommassen, J. (2006). Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS biology*, 4(11):e377.
- Robert, X. and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic acids research*, 42(Web Server issue):W320–4.
- Rodionova, N. A., Tatulian, S. A., Surrey, T., Jähnig, F., and Tamm, L. K. (1995). Characterization of two membrane-bound forms of OmpA. *Biochemistry*, 34(6):1921–9.
- Rojas, E. R., Billings, G., Odermatt, P. D., Auer, G. K., Zhu, L., Miguel, A., Chang, F., Weibel, D. B., Theriot, J. A., and Huang, K. C. (2018). The outer membrane is an essential load-bearing element in Gram-negative bacteria. *Nature*, 559(7715):617–621.
- Roman-Hernandez, G., Peterson, J. H., and Bernstein, H. D. (2014). Reconstitution of bacterial autotransporter assembly using purified components. *eLife*, 3:e04234.
- Rosenbusch, J. P. (1974). Characterization of the major envelope protein from *Escherichia coli*. Regular arrangement on the peptidoglycan and unusual dodecyl sulfate binding. *The Journal of biological chemistry*, 249(24):8019–29.
- Rouvière, P. E. and Gross, C. A. (1996). SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Genes & development*, 10(24):3170–82.

- Ruiz, N., Falcone, B., Kahne, D., and Silhavy, T. J. (2005). Chemical conditionality: a genetic strategy to probe organelle assembly. *Cell*, 121(2):307–17.
- Rust, M. J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature methods*, 3(10):793–5.
- Ryan, K. R., Taylor, J. A., and Bowers, L. M. (2010). The BAM complex subunit BamE (SmpA) is required for membrane integrity, stalk growth and normal levels of outer membrane {beta}-barrel proteins in *Caulobacter crescentus*. *Microbiology (Reading, England)*, 156(Pt 3):742–56.
- Sachelaru, I., Petriman, N.-A., Kudva, R., and Koch, H.-G. (2014). Dynamic interaction of the sec translocon with the chaperone PpiD. *The Journal of biological chemistry*, 289(31):21706–15.
- Samsudin, F., Boags, A., Piggot, T. J., and Khalid, S. (2017). Braun's Lipoprotein Facilitates OmpA Interaction with the Escherichia coli Cell Wall. *Biophysical journal*, 113(7):1496–1504.
- Sanchez, K. M., Gable, J. E., Schlamadinger, D. E., and Kim, J. E. (2008). Effects of tryptophan microenvironment, soluble domain, and vesicle size on the thermodynamics of membrane protein folding: lessons from the transmembrane protein OmpA. *Biochemistry*, 47(48):12844–52.
- Sanchez, S. A., Tricerri, M. A., and Gratton, E. (2012). Laurdan generalized polarization fluctuations measures membrane packing micro-heterogeneity in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19):7314–9.
- Sanchez, S. A., Tricerri, M. A., Gunther, G., and Gratton, E. (2007). Laurdan Generalized Polarization: from cuvette to microscope. *Modern Research and Educational Topics in Microscopy*, 48:1689–1700.
- Sanderson, J. M. (2005). Peptide-lipid interactions: insights and perspectives. *Organic & biomolecular chemistry*, 3(2):201–12.
- Sandlin, C. W., Zaccari, N. R., and Fleming, K. G. (2015). Skp Trimer Formation Is Insensitive to Salts in the Physiological Range. *Biochemistry*, 54(48):7059–62.
- Sandoval, C. M., Baker, S. L., Jansen, K., Metzner, S. I., and Sousa, M. C. (2011). Crystal structure of BamD: an essential component of the β -Barrel assembly machinery of gram-negative bacteria. *Journal of molecular biology*, 409(3):348–57.
- Sankaram, M. B., Marsh, D., Gierasch, L. M., and Thompson, T. E. (1994). Reorganization of lipid domain structure in membranes by a transmembrane peptide: an ESR spin label study on the effect of the Escherichia coli outer membrane protein A signal peptide on the fluid lipid domain connectivity in binary mixtures of dimyristoyl. *Biophysical journal*, 66(6):1959–68.

- Santini, C. L., Bernadac, A., Zhang, M., Chanal, A., Ize, B., Blanco, C., and Wu, L. F. (2001). Translocation of jellyfish green fluorescent protein via the Tat system of *Escherichia coli* and change of its periplasmic localization in response to osmotic up-shock. *The Journal of biological chemistry*, 276(11):8159–64.
- Saul, F. A., Arié, J.-P., Vulliez-le Normand, B., Kahn, R., Betton, J.-M., and Bentley, G. A. (2004). Structural and functional studies of FkpA from *Escherichia coli*, a cis/trans peptidyl-prolyl isomerase with chaperone activity. *Journal of molecular biology*, 335(2):595–608.
- Sauri, A., Soprova, Z., Wickström, D., de Gier, J.-W., Van der Schors, R. C., Smit, A. B., Jong, W. S. P., and Luirink, J. (2009). The Bam (Omp85) complex is involved in secretion of the autotransporter haemoglobin protease. *Microbiology (Reading, England)*, 155(Pt 12):3982–91.
- Schäfer, U., Beck, K., and Müller, M. (1999). Skp, a molecular chaperone of gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins. *The Journal of biological chemistry*, 274(35):24567–74.
- Schiffrin, B., Brockwell, D. J., and Radford, S. E. (2017a). Outer membrane protein folding from an energy landscape perspective. *BMC biology*, 15(1):123.
- Schiffrin, B., Calabrese, A. N., Devine, P. W. A., Harris, S. A., Ashcroft, A. E., Brockwell, D. J., and Radford, S. E. (2016). Skp is a multivalent chaperone of outer-membrane proteins. *Nature structural & molecular biology*, 23(9):786–793.
- Schiffrin, B., Calabrese, A. N., Higgins, A. J., Humes, J. R., Ashcroft, A. E., Kalli, A. C., Brockwell, D. J., and Radford, S. E. (2017b). Effects of Periplasmic Chaperones and Membrane Thickness on BamA-Catalyzed Outer-Membrane Protein Folding. *Journal of molecular biology*, 429(23):3776–3792.
- Schlapschy, M., Dommel, M. K., Hadian, K., Fogarasi, M., Korndörfer, I. P., and Skerra, A. (2004). The periplasmic *E. coli* chaperone Skp is a trimer in solution: biophysical and preliminary crystallographic characterization. *Biological chemistry*, 385(2):137–43.
- Schlünzen, F., Wilson, D. N., Tian, P., Harms, J. M., McInnes, S. J., Hansen, H. A. S., Albrecht, R., Buerger, J., Wilbanks, S. M., and Fucini, P. (2005). The binding mode of the trigger factor on the ribosome: implications for protein folding and SRP interaction. *Structure (London, England : 1993)*, 13(11):1685–94.
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature methods*, 9(7):671–5.
- Schneider, M., Belsom, A., and Rappsilber, J. (2018). Protein Tertiary Structure by Crosslinking/Mass Spectrometry. *Trends in biochemical sciences*, 43(3):157–169.
- Schram, V. and Thompson, T. E. (1997). Influence of the intrinsic membrane protein bacteriorhodopsin on gel-phase domain topology in two-component phase-separated bilayers. *Biophysical journal*, 72(5):2217–25.

- Schultz, C. P., Wolf, V., Lange, R., Mertens, E., Wecke, J., Naumann, D., and Zähringer, U. (1998). Evidence for a new type of outer membrane lipid in oral spirochete *Treponema denticola*. Functioning permeation barrier without lipopolysaccharides. *The Journal of biological chemistry*, 273(25):15661–6.
- Schüßler, A., Herwig, S., and Kleinschmidt, J. H. (2019). Kinetics of Insertion and Folding of Outer Membrane Proteins by Gel Electrophoresis. *Methods in molecular biology (Clifton, N.J.)*, 2003:145–162.
- Schwalm, J., Mahoney, T. F., Soltes, G. R., and Silhavy, T. J. (2013). Role for Skp in LptD assembly in *Escherichia coli*. *Journal of bacteriology*, 195(16):3734–42.
- Schwarz, A. and Beck, M. (2019). The Benefits of Cotranslational Assembly: A Structural Perspective. *Trends in Cell Biology*, 29(10):791–803.
- Schwepe, D. K., Chavez, J. D., Lee, C. F., Caudal, A., Kruse, S. E., Stuppard, R., Marcinek, D. J., Shadel, G. S., Tian, R., and Bruce, J. E. (2017). Mitochondrial protein interactome elucidated by chemical cross-linking mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America*, 114(7):1732–1737.
- Schwieters, C. D., Bermejo, G. A., and Clore, G. M. (2018). Xplor-NIH for molecular structure determination from NMR and other data sources. *Protein science : a publication of the Protein Society*, 27(1):26–40.
- Sciandrone, B., Forti, F., Perego, S., Falchi, F., and Briani, F. (2019). Temperature-dependent regulation of the *Escherichia coli* lpxT gene. *Biochimica et biophysica acta. Gene regulatory mechanisms*, 1862(8):786–795.
- Scott, D. W. (1992). Kernel Density Estimators. In *Multivariate Density Estimation: Theory, Practice, and Visualization*, chapter 6, pages 125–193. John Wiley & Sons, Inc., 1 edition.
- Selkrig, J., Mosbahi, K., Webb, C. T., Belousoff, M. J., Perry, A. J., Wells, T. J., Morris, F., Leyton, D. L., Totsika, M., Phan, M.-D., Celik, N., Kelly, M., Oates, C., Hartland, E. L., Robins-Browne, R. M., Ramarathinam, S. H., Purcell, A. W., Schembri, M. A., Strugnell, R. A., Henderson, I. R., Walker, D., and Lithgow, T. (2012). Discovery of an archetypal protein transport system in bacterial outer membranes. *Nature structural & molecular biology*, 19(5):506–10, S1.
- Sharma, A., Leach, R. N., Gell, C., Zhang, N., Burrows, P. C., Shepherd, D. A., Wigneshweraraj, S., Smith, D. A., Zhang, X., Buck, M., Stockley, P. G., and Tuma, R. (2014). Domain movements of the enhancer-dependent sigma factor drive DNA delivery into the RNA polymerase active site: insights from single molecule studies. *Nucleic acids research*, 42(8):5177–90.
- Sharpless, N. E. and Flavin, M. (1966). The reactions of amines and amino acids with maleimides. Structure of the reaction products deduced from infrared and nuclear magnetic resonance spectroscopy. *Biochemistry*, 5(9):2963–71.
- Sherman, D. J., Xie, R., Taylor, R. J., George, A. H., Okuda, S., Foster, P. J., Needleman, D. J., and Kahne, D. (2018). Lipopolysaccharide is transported to the cell surface by a membrane-to-membrane protein bridge. *Science (New York, N.Y.)*, 359(6377):798–801.

- Shiba, Y., Yokoyama, Y., Aono, Y., Kiuchi, T., Kusaka, J., Matsumoto, K., and Hara, H. (2004). Activation of the Rcs signal transduction system is responsible for the thermosensitive growth defect of an *Escherichia coli* mutant lacking phosphatidylglycerol and cardiolipin. *Journal of bacteriology*, 186(19):6526–35.
- Sikdar, R., Peterson, J. H., Anderson, D. E., and Bernstein, H. D. (2017). Folding of a bacterial integral outer membrane protein is initiated in the periplasm. *Nature communications*, 8(1):1309.
- Sikora, A. E., Wierzbicki, I. H., Zielke, R. A., Ryner, R. F., Korotkov, K. V., Buchanan, S. K., and Noinaj, N. (2018). Structural and functional insights into the role of BamD and BamE within the β -barrel assembly machinery in *Neisseria gonorrhoeae*. *The Journal of biological chemistry*, 293(4):1106–1119.
- Sinensky, M. (1974). Homeoviscous adaptation—a homeostatic process that regulates the viscosity of membrane lipids in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 71(2):522–5.
- Singh, A., Thornton, E. R., and Westheimer, F. H. (1962). The photolysis of diazoacetylchymotrypsin. *The Journal of biological chemistry*, 237:3006–8.
- Sinnige, T., Weingarh, M., Daniëls, M., Boelens, R., Bonvin, A. M. J. J., Houben, K., and Baldus, M. (2015). Conformational Plasticity of the POTRA 5 Domain in the Outer Membrane Protein Assembly Factor BamA. *Structure (London, England : 1993)*, 23(7):1317–24.
- Sinz, A. (2014). The advancement of chemical cross-linking and mass spectrometry for structural proteomics: from single proteins to protein interaction networks. *Expert review of proteomics*, 11(6):733–43.
- Sinz, A. (2018). Cross-Linking/Mass Spectrometry for Studying Protein Structures and Protein-Protein Interactions: Where Are We Now and Where Should We Go from Here? *Angewandte Chemie (International ed. in English)*, 57(22):6390–6396.
- Skennerton, C. T., Haroon, M. F., Briegel, A., Shi, J., Jensen, G. J., Tyson, G. W., and Orphan, V. J. (2016). Phylogenomic analysis of Candidatus 'Izimaplasma' species: free-living representatives from a Tenericutes clade found in methane seeps. *The ISME journal*, 10(11):2679–2692.
- Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D., and Silhavy, T. J. (2007a). Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 104(15):6400–5.
- Sklar, J. G., Wu, T., Kahne, D., and Silhavy, T. J. (2007b). Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes & development*, 21(19):2473–84.
- Smit, J. and Nikaido, H. (1978). Outer membrane of gram-negative bacteria. XVIII. Electron microscopic studies on porin insertion sites and growth of cell surface of *Salmonella typhimurium*. *Journal of bacteriology*, 135(2):687–702.

- Smith, R. A. and Knowles, J. R. (1973). Aryldiazirines. Potential reagents for photolabeling of biological receptor sites. *Journal of the American Chemical Society*, 95(15):5072–3.
- Smith, R. A. G. and Knowles, J. R. (1975). The preparation and photolysis of 3-aryl-3H-diazirines. *Journal of the Chemical Society, Perkin Transactions 2*, (7):686.
- Smith-Dupont, K. B., Guo, L., and Gai, F. (2010). Diffusion as a probe of the heterogeneity of antimicrobial peptide-membrane interactions. *Biochemistry*, 49(22):4672–8.
- Soltes, G. R., Martin, N. R., Park, E., Sutterlin, H. A., and Silhavy, T. J. (2017). Distinctive Roles for Periplasmic Proteases in the Maintenance of Essential Outer Membrane Protein Assembly. *Journal of bacteriology*, 199(20):1–10.
- Soltes, G. R., Schwalm, J., Ricci, D. P., and Silhavy, T. J. (2016). The Activity of Escherichia coli Chaperone SurA Is Regulated by Conformational Changes Involving a Parvulin Domain. *Journal of bacteriology*, 198(6):921–9.
- Sperotto, M. and Mouritsen, O. (1988). Dependence of lipid membrane phase transition temperature on the mismatch of protein and lipid hydrophobic thickness. *European Biophysics Journal*, 16(1):1–10.
- Srivastava, S. R., Zadafiya, P., and Mahalakshmi, R. (2018). Hydrophobic Mismatch Modulates Stability and Plasticity of Human Mitochondrial VDAC2. *Biophysical journal*, 115(12):2386–2394.
- Steinhauer, J., Agha, R., Pham, T., Varga, A. W., and Goldberg, M. B. (1999). The unipolar Shigella surface protein IcsA is targeted directly to the bacterial old pole: IcsP cleavage of IcsA occurs over the entire bacterial surface. *Molecular microbiology*, 32(2):367–77.
- Stepniewski, M., Bunker, A., Pasenkiewicz-Gierula, M., Karttunen, M., and Róg, T. (2010). Effects of the lipid bilayer phase state on the water membrane interface. *The journal of physical chemistry. B*, 114(36):11784–92.
- Storek, K. M., Auerbach, M. R., Shi, H., Garcia, N. K., Sun, D., Nickerson, N. N., Vij, R., Lin, Z., Chiang, N., Schneider, K., Weckslar, A. T., Skippington, E., Nakamura, G., Seshasayee, D., Koerber, J. T., Payandeh, J., Smith, P. A., and Rutherford, S. T. (2018a). Monoclonal antibody targeting the β -barrel assembly machine of Escherichia coli is bactericidal. *Proceedings of the National Academy of Sciences of the United States of America*, 115(14):3692–3697.
- Storek, K. M., Vij, R., Sun, D., Smith, P. A., Koerber, J. T., and Rutherford, S. T. (2018b). The Escherichia coli β -barrel assembly machinery is sensitized to perturbations under high membrane fluidity. *Journal of bacteriology*, 201(1):1–15.
- Stouthamer, A. H. (1973). A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van Leeuwenhoek*, 39(3):545–65.

- Stroffekova, K., Proenza, C., and Beam, K. G. (2001). The protein-labeling reagent FLASH-EDT2 binds not only to CCXXCC motifs but also non-specifically to endogenous cysteine-rich proteins. *Pflugers Archiv : European journal of physiology*, 442(6):859–66.
- Suda, Y., Okazaki, F., Hasegawa, Y., Adachi, S., Fukase, K., Kokubo, S., Kuramitsu, S., and Kusumoto, S. (2012). Structural characterization of neutral and acidic glycolipids from *Thermus thermophilus* HB8. *PloS one*, 7(7):e35067.
- Surrey, T. and Jähnig, F. (1992). Refolding and oriented insertion of a membrane protein into a lipid bilayer. *Proceedings of the National Academy of Sciences of the United States of America*, 89(16):7457–61.
- Sutcliffe, I. C. (2010). A phylum level perspective on bacterial cell envelope architecture. *Trends in microbiology*, 18(10):464–70.
- Sutcliffe, I. C. (2011). Cell envelope architecture in the Chloroflexi: a shifting frontline in a phylogenetic turf war. *Environmental microbiology*, 13(2):279–82.
- Tacconelli, E., Carrara, E., Savoldi, A., Harbarth, S., Mendelson, M., Monnet, D. L., Pulcini, C., Kahlmeter, G., Kluytmans, J., Carmeli, Y., Ouellette, M., Outterson, K., Patel, J., Cavaleri, M., Cox, E. M., Houchens, C. R., Grayson, M. L., Hansen, P., Singh, N., Theuretzbacher, U., Magrini, N., and WHO Pathogens Priority List Working Group (2018). Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *The Lancet. Infectious diseases*, 18(3):318–327.
- Tachikawa, T. and Kato, J.-i. (2011). Suppression of the temperature-sensitive mutation of the *bamD* gene required for the assembly of outer membrane proteins by multicopy of the *yiaD* gene in *Escherichia coli*. *Bioscience, biotechnology, and biochemistry*, 75(1):162–4.
- Tan, B. K., Bogdanov, M., Zhao, J., Dowhan, W., Raetz, C. R. H., and Guan, Z. (2012). Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. *Proceedings of the National Academy of Sciences of the United States of America*, 109(41):16504–9.
- Tata, M. and Konovalova, A. (2019). Improper Coordination of BamA and BamD Results in Bam Complex Jamming by a Lipoprotein Substrate. *mBio*, 10(3):1–10.
- Taylor, A. J., Zakai, S. A. I., and Kelly, D. J. (2017). The Periplasmic Chaperone Network of *Campylobacter jejuni*: Evidence that SalC (Cj1289) and PpiD (Cj0694) Are Involved in Maintaining Outer Membrane Integrity. *Frontiers in microbiology*, 8(MAR):531.
- Tellez, R. and Misra, R. (2012). Substitutions in the BamA β -barrel domain overcome the conditional lethal phenotype of a Δ bamB Δ bamE strain of *Escherichia coli*. *Journal of bacteriology*, 194(2):317–24.
- Thoma, J., Bosshart, P., Pfreundschuh, M., and Müller, D. J. (2012). Out but not in: the large transmembrane β -barrel protein FhuA unfolds but cannot refold via β -hairpins. *Structure (London, England : 1993)*, 20(12):2185–90.

- Thoma, J., Burmann, B. M., Hiller, S., and Müller, D. J. (2015). Impact of holdase chaperones Skp and SurA on the folding of β -barrel outer-membrane proteins. *Nature structural & molecular biology*, 22(10):795–802.
- Thoma, J., Sun, Y., Ritzmann, N., and Müller, D. J. (2018). POTRA Domains, Extracellular Lid, and Membrane Composition Modulate the Conformational Stability of the β Barrel Assembly Factor BamA. *Structure (London, England : 1993)*, 26(7):987–996.e3.
- Thomas, S., Holland, I. B., and Schmitt, L. (2014). The Type 1 secretion pathway - the hemolysin system and beyond. *Biochimica et biophysica acta*, 1843(8):1629–41.
- Thompson, M. G., Sedaghatian, N., Barajas, J. F., Wehrs, M., Bailey, C. B., Kaplan, N., Hillson, N. J., Mukhopadhyay, A., and Keasling, J. D. (2018). Isolation and characterization of novel mutations in the pSC101 origin that increase copy number. *Scientific reports*, 8(1):1590.
- Thompson, S. E. and Smithrud, D. B. (2002). Carboxylates stacked over aromatic rings promote salt bridge formation in water. *Journal of the American Chemical Society*, 124(3):442–9.
- Tomov, T. E., Tsukanov, R., Masoud, R., Liber, M., Plavner, N., and Nir, E. (2012). Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis. *Biophysical journal*, 102(5):1163–73.
- Toscano, J. P., Platz, M. S., and Nikolaev, V. (1995). Lifetimes of Simple Ketocarbenes. *Journal of the American Chemical Society*, 117(16):4712–4713.
- Tristram-Nagle, S., Liu, Y., Legleiter, J., and Nagle, J. F. (2002). Structure of gel phase DMPC determined by X-ray diffraction. *Biophysical journal*, 83(6):3324–35.
- Tsirigos, K. D., Bagos, P. G., and Hamodrakas, S. J. (2011). OMPdb: a database of {beta}-barrel outer membrane proteins from Gram-negative bacteria. *Nucleic acids research*, 39(Database issue):D324–31.
- Tu, K., Tobias, D. J., Blasie, J. K., and Klein, M. L. (1996). Molecular dynamics investigation of the structure of a fully hydrated gel-phase dipalmitoylphosphatidylcholine bilayer. *Biophysical journal*, 70(2):595–608.
- Tuijtel, M. W., Koster, A. J., Jakobs, S., Faas, F. G. A., and Sharp, T. H. (2019). Correlative cryo super-resolution light and electron microscopy on mammalian cells using fluorescent proteins. *Scientific reports*, 9(1):1369.
- UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic acids research*, 47(D1):D506–D515.
- Ureta, A. R., Endres, R. G., Wingreen, N. S., and Silhavy, T. J. (2007). Kinetic analysis of the assembly of the outer membrane protein LamB in *Escherichia coli* mutants each lacking a secretion or targeting factor in a different cellular compartment. *Journal of bacteriology*, 189(2):446–54.

- Ursell, T. S., Trepagnier, E. H., Huang, K. C., and Theriot, J. A. (2012). Analysis of surface protein expression reveals the growth pattern of the gram-negative outer membrane. *PLoS computational biology*, 8(9):e1002680.
- Vander Heiden, M. G., Chandel, N. S., Li, X. X., Schumacker, P. T., Colombini, M., and Thompson, C. B. (2000). Outer mitochondrial membrane permeability can regulate coupled respiration and cell survival. *Proceedings of the National Academy of Sciences of the United States of America*, 97(9):4666–71.
- Vanounou, S., Pines, D., Pines, E., Parola, A. H., and Fishov, I. (2002). Coexistence of domains with distinct order and polarity in fluid bacterial membranes. *Photochemistry and photobiology*, 76(1):1–11.
- Vassen, V., Valotteau, C., Feuillie, C., Formosa-Dague, C., Dufrêne, Y. F., and De Bolle, X. (2019). Localized incorporation of outer membrane components in the pathogen *Brucella abortus*. *The EMBO journal*, 38(5):e100323.
- Velázquez, J. B. and Fernández, M. S. (2006). GPS, the slope of Laurdan generalized polarization spectra, in the study of phospholipid lateral organization and *Escherichia coli* lipid phases. *Archives of biochemistry and biophysics*, 455(2):163–74.
- Vequi-Suplicy, C. C., Coutinho, K., and Lamy, M. T. (2014). Electric dipole moments of the fluorescent probes Prodan and Laurdan: experimental and theoretical evaluations. *Biophysical reviews*, 6(1):63–74.
- Vequi-Suplicy, C. C., Coutinho, K., and Lamy, M. T. (2015). New insights on the fluorescent emission spectra of Prodan and Laurdan. *Journal of fluorescence*, 25(3):621–9.
- Verhoeven, G. S., Dogterom, M., and den Blaauwen, T. (2013). Absence of long-range diffusion of OmpA in *E. coli* is not caused by its peptidoglycan binding domain. *BMC microbiology*, 13(1):66.
- Vertommen, D., Ruiz, N., Leverrier, P., Silhavy, T. J., and Collet, J.-F. (2009). Characterization of the role of the *Escherichia coli* periplasmic chaperone SurA using differential proteomics. *Proteomics*, 9(9):2432–43.
- Vij, R., Lin, Z., Chiang, N., Vernes, J.-M., Storek, K. M., Park, S., Chan, J., Meng, Y. G., Comps-Agrar, L., Luan, P., Lee, S., Schneider, K., Bevers, J., Zilberleyb, I., Tam, C., Koth, C. M., Xu, M., Gill, A., Auerbach, M. R., Smith, P. A., Rutherford, S. T., Nakamura, G., Seshasayee, D., Payandeh, J., and Koerber, J. T. (2018). A targeted boost-and-sort immunization strategy using *Escherichia coli* BamA identifies rare growth inhibitory antibodies. *Scientific reports*, 8(1):7136.
- Vinogradov, E., Paul, C. J., Li, J., Zhou, Y., Lyle, E. A., Tapping, R. I., Kropinski, A. M., and Perry, M. B. (2004). The structure and biological characteristics of the *Spirochaeta aurantia* outer membrane glycolipid LGLB. *European journal of biochemistry*, 271(23-24):4685–95.
- Volokhina, E. B., Beckers, F., Tommassen, J., and Bos, M. P. (2009). The beta-barrel outer membrane protein assembly complex of *Neisseria meningitidis*. *Journal of bacteriology*, 191(22):7074–85.

- Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M., and Tommassen, J. (2003). Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science (New York, N.Y.)*, 299(5604):262–5.
- Vuong, P., Bennion, D., Mantei, J., Frost, D., and Misra, R. (2008). Analysis of YfgL and YaeT interactions through bioinformatics, mutagenesis, and biochemistry. *Journal of bacteriology*, 190(5):1507–17.
- Wals, K. and Ovaa, H. (2014). Unnatural amino acid incorporation in E. coli: current and future applications in the design of therapeutic proteins. *Frontiers in chemistry*, 2(APR):15.
- Walton, T. A., Sandoval, C. M., Fowler, C. A., Pardi, A., and Sousa, M. C. (2009). The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proceedings of the National Academy of Sciences of the United States of America*, 106(6):1772–7.
- Walton, T. A. and Sousa, M. C. (2004). Crystal structure of Skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation. *Molecular cell*, 15(3):367–74.
- Wang, L., Bateman, B., Zanetti-Domingues, L. C., Moores, A. N., Astbury, S., Spindloe, C., Darrow, M. C., Romano, M., Needham, S. R., Beis, K., Rolfe, D. J., Clarke, D. T., and Martin-Fernandez, M. L. (2019). Solid immersion microscopy images cells under cryogenic conditions with 12 nm resolution. *Communications biology*, 2(1):74.
- Wang, Y., Wang, R., Jin, F., Liu, Y., Yu, J., Fu, X., and Chang, Z. (2016). A Supercomplex Spanning the Inner and Outer Membranes Mediates the Biogenesis of β -Barrel Outer Membrane Proteins in Bacteria. *The Journal of biological chemistry*, 291(32):16720–9.
- Warner, L. R., Gatzeva-Topalova, P. Z., Doerner, P. A., Pardi, A., and Sousa, M. C. (2017). Flexibility in the Periplasmic Domain of BamA Is Important for Function. *Structure (London, England : 1993)*, 25(1):94–106.
- Warner, L. R., Varga, K., Lange, O. F., Baker, S. L., Baker, D., Sousa, M. C., and Pardi, A. (2011). Structure of the BamC two-domain protein obtained by Rosetta with a limited NMR data set. *Journal of molecular biology*, 411(1):83–95.
- Webb, B., Viswanath, S., Bonomi, M., Pellarin, R., Greenberg, C. H., Saltzberg, D., and Sali, A. (2018). Integrative structure modeling with the Integrative Modeling Platform. *Protein science : a publication of the Protein Society*, 27(1):245–258.
- Webb, C. T., Heinz, E., and Lithgow, T. (2012a). Evolution of the β -barrel assembly machinery. *Trends in microbiology*, 20(12):612–20.
- Webb, C. T., Selkrig, J., Perry, A. J., Noinaj, N., Buchanan, S. K., and Lithgow, T. (2012b). Dynamic association of BAM complex modules includes surface exposure of the lipoprotein BamC. *Journal of molecular biology*, 422(4):545–55.
- Weber, P. J. and Beck-Sickinger, A. G. (1997). Comparison of the photochemical behavior of four different photoactivatable probes. *The journal of peptide research : official journal of the American Peptide Society*, 49(5):375–83.

- Wedge, E., Lie, K., Bolstad, K., Weynants, V. E., Halstensen, A., Herstad, T. K., Kreutzberger, J., Nome, L., Naess, L. M., and Aase, A. (2013). Meningococcal omp85 in detergent-extracted outer membrane vesicle vaccines induces high levels of non-functional antibodies in mice. *Scandinavian journal of immunology*, 77(6):452–9.
- Weisenburger, S., Boening, D., Schomburg, B., Giller, K., Becker, S., Griesinger, C., and Sandoghdar, V. (2017). Cryogenic optical localization provides 3D protein structure data with Angstrom resolution. *Nature methods*, 14(2):141–144.
- Weisenburger, S., Jing, B., Hänni, D., Reymond, L., Schuler, B., Renn, A., and Sandoghdar, V. (2014). Cryogenic colocalization microscopy for nanometer-distance measurements. *Chemphyschem : a European journal of chemical physics and physical chemistry*, 15(4):763–70.
- Weisenburger, S., Jing, B., Renn, A., and Sandoghdar, V. (2013). Cryogenic localization of single molecules with angstrom precision. page 88150D.
- Werner, J. and Misra, R. (2005). YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of Escherichia coli. *Molecular microbiology*, 57(5):1450–9.
- White, G. F., Racher, K. I., Lipski, A., Hallett, F. R., and Wood, J. M. (2000). Physical properties of liposomes and proteoliposomes prepared from Escherichia coli polar lipids. *Biochimica et biophysica acta*, 1468(1-2):175–86.
- Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences of the United States of America*, 95(12):6578–83.
- Wiedemann, N., Kozjak, V., Chacinska, A., Schönfisch, B., Rospert, S., Ryan, M. T., Pfanner, N., and Meisinger, C. (2003). Machinery for protein sorting and assembly in the mitochondrial outer membrane. *Nature*, 424(6948):565–71.
- Wikström, M., Kelly, A. A., Georgiev, A., Eriksson, H. M., Klement, M. R., Bogdanov, M., Dowhan, W., and Wieslander, A. (2009). Lipid-engineered Escherichia coli membranes reveal critical lipid headgroup size for protein function. *The Journal of biological chemistry*, 284(2):954–65.
- Wiles, T. J., Wall, E. S., Schlomann, B. H., Hay, E. A., Parthasarathy, R., and Guillemin, K. (2018). Modernized Tools for Streamlined Genetic Manipulation and Comparative Study of Wild and Diverse Proteobacterial Lineages. *mBio*, 9(5):1–19.
- Willig, K. I., Rizzoli, S. O., Westphal, V., Jahn, R., and Hell, S. W. (2006). STED microscopy reveals that synaptotagmin remains clustered after synaptic vesicle exocytosis. *Nature*, 440(7086):935–9.
- Wimley, W. C. (2003). The versatile beta-barrel membrane protein. *Current opinion in structural biology*, 13(4):404–11.
- Wimley, W. C., Hristova, K., Ladokhin, A. S., Silvestro, L., Axelsen, P. H., and White, S. H. (1998). Folding of beta-sheet membrane proteins: a hydrophobic hexapeptide model. *Journal of molecular biology*, 277(5):1091–110.

- Wimley, W. C. and White, S. H. (1996). Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature structural biology*, 3(10):842–8.
- Wittelsberger, A., Thomas, B. E., Mierke, D. F., and Rosenblatt, M. (2006). Methionine acts as a "magnet" in photoaffinity crosslinking experiments. *FEBS letters*, 580(7):1872–6.
- Worthington, A. S. and Burkart, M. D. (2006). One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Organic & biomolecular chemistry*, 4(1):44–6.
- Wu, E. L., Fleming, P. J., Yeom, M. S., Widmalm, G., Klauda, J. B., Fleming, K. G., and Im, W. (2014). E. coli outer membrane and interactions with OmpLA. *Biophysical journal*, 106(11):2493–502.
- Wu, R., Stephenson, R., Gichaba, A., and Noinaj, N. (2019). The big BAM theory: An open and closed case? *Biochimica et Biophysica Acta (BBA) - Biomembranes*, page 183062.
- Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J., and Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. *Cell*, 121(2):235–45.
- Wu, T., McCandlish, A. C., Gronenberg, L. S., Chng, S.-S., Silhavy, T. J., and Kahne, D. (2006). Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 103(31):11754–9.
- Wzorek, J. S., Lee, J., Tomasek, D., Hagan, C. L., and Kahne, D. E. (2017). Membrane integration of an essential β -barrel protein prerequires burial of an extracellular loop. *Proceedings of the National Academy of Sciences of the United States of America*, 114(10):2598–2603.
- Xia, Y., DiPrimio, N., Keppel, T. R., Vo, B., Fraser, K., Battaile, K. P., Egan, C., Bystroff, C., Lovell, S., Weis, D. D., Anderson, J. C., and Karanicolas, J. (2013). The designability of protein switches by chemical rescue of structure: mechanisms of inactivation and reactivation. *Journal of the American Chemical Society*, 135(50):18840–9.
- Xie, A. F., Yamada, R., Gewirth, A. A., and Granick, S. (2002). Materials science of the gel to fluid phase transition in a supported phospholipid bilayer. *Physical review letters*, 89(24):246103.
- Xu, Q., Kim, M., Ho, K. W., Lachowicz, P., Fanucci, G. E., and Cafiso, D. S. (2008). Membrane hydrocarbon thickness modulates the dynamics of a membrane transport protein. *Biophysical Journal*, 95(6):2849–2858.
- Xu, X., Wang, S., Hu, Y.-X., and McKay, D. B. (2007). The periplasmic bacterial molecular chaperone SurA adapts its structure to bind peptides in different conformations to assert a sequence preference for aromatic residues. *Journal of molecular biology*, 373(2):367–81.

- Yamamoto, N., Bernardi, F., Bottoni, A., Olivucci, M., Robb, M. A., and Wilsey, S. (1994). Mechanism of Carbene Formation from the Excited States of Diazirine and Diazomethane: An MC-SCF Study. *Journal of the American Chemical Society*, 116(5):2064–2074.
- Yamazaki, Y., Niki, H., and Kato, J.-i. (2008). Profiling of Escherichia coli Chromosome database. *Methods in molecular biology (Clifton, N.J.)*, 416(6):385–9.
- Yang, J. and Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. *Nucleic acids research*, 43(W1):W174–81.
- Yin, F. and Kindt, J. T. (2012). Hydrophobic mismatch and lipid sorting near OmpA in mixed bilayers: atomistic and coarse-grained simulations. *Biophysical journal*, 102(10):2279–87.
- Yin, J., Lin, A. J., Golan, D. E., and Walsh, C. T. (2006). Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nature protocols*, 1(1):280–5.
- Yin, J., Straight, P. D., McLoughlin, S. M., Zhou, Z., Lin, A. J., Golan, D. E., Kelleher, N. L., Kolter, R., and Walsh, C. T. (2005). Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proceedings of the National Academy of Sciences of the United States of America*, 102(44):15815–20.
- Yilmaz, u., Shiferaw, G. A., Rayo, J., Economou, A., Martens, L., and Vandermarliere, E. (2018). Cross-linked peptide identification: A computational forest of algorithms. *Mass spectrometry reviews*, 37(6):738–749.
- Yoshita, M., Koyama, K., Hayamizu, Y., Baba, M., and Akiyama, H. (2002). Improved High Collection Efficiency in Fluorescence Microscopy with a Weierstrass-Sphere Solid Immersion Lens. *Japanese Journal of Applied Physics*, 41(Part 2, No. 7B):L858–L860.
- Yu, J. and Lu, L. (2019). BamA is a pivotal protein in cell envelope synthesis and cell division in Deinococcus radiodurans. *Biochimica et biophysica acta. Biomembranes*, 1861(7):1365–1374.
- Zaccai, N. R., Sandlin, C. W., Hoopes, J. T., Curtis, J. E., Fleming, P. J., Fleming, K. G., and Krueger, S. (2016). Deuterium Labeling Together with Contrast Variation Small-Angle Neutron Scattering Suggests How Skp Captures and Releases Unfolded Outer Membrane Proteins. *Methods in enzymology*, 566:159–210.
- Zhang, M., Lin, S., Song, X., Liu, J., Fu, Y., Ge, X., Fu, X., Chang, Z., and Chen, P. R. (2011). A genetically incorporated crosslinker reveals chaperone cooperation in acid resistance. *Nature chemical biology*, 7(10):671–7.
- Zhang, Y., Song, K.-H., Tang, S., Ravelo, L., Cusido, J., Sun, C., Zhang, H. F., and Raymo, F. M. (2018). Far-Red Photoactivatable BODIPYs for the Super-Resolution Imaging of Live Cells. *Journal of the American Chemical Society*, 140(40):12741–12745.
- Zhou, J. and Xu, Z. (2003). Structural determinants of SecB recognition by SecA in bacterial protein translocation. *Nature structural biology*, 10(11):942–7.

- Ziemianowicz, D. S., Bomgarden, R., Etienne, C., and Schriemer, D. C. (2017). Amino Acid Insertion Frequencies Arising from Photoproducts Generated Using Aliphatic Diazirines. *Journal of the American Society for Mass Spectrometry*, 28(10):2011–2021.
- Zimmer, J., Nam, Y., and Rapoport, T. A. (2008). Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature*, 455(7215):936–43.
- Zuber, B., Chami, M., Houssin, C., Dubochet, J., Griffiths, G., and Daffé, M. (2008). Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *Journal of Bacteriology*, 190(16):5672–5680.
- Zuber, B., Haenni, M., Ribeiro, T., Minnig, K., Lopes, F., Moreillon, P., and Dubochet, J. (2006). Granular layer in the periplasmic space of gram-positive bacteria and fine structures of *Enterococcus gallinarum* and *Streptococcus gordonii* septa revealed by cryo-electron microscopy of vitreous sections. *Journal of bacteriology*, 188(18):6652–60.

Appendix A

List of modified peptides identified by XL-MS/MS from Skp x OmpA crosslinking experiments

OmpA(W7C)[MTS-diazirine] and OmpA(T144C)[MTS-diazirine]

Table 1 Modified peptides of Skp identified by MS/MS from crosslinking experiments using OmpA(W7C) and OmpA(T144C) conjugated to MTS-diazirine crosslinked to Skp. See main text for experimental details and discussion. A number of residues had modifications arising from the treatment with the presence of urea (carbamylation) and protein ageing / treatments used for preparation (deamidation and oxidation). These are marked to the right of residue. ‡ = deamidation, † = carbamylation, * = oxidation.

Sequence	Start Residue	End Residue	Modified Residue	# of Spectra	Enrichment	Lamp
T144C						
AIVNMGSLFQQVAQK	25	39	25-31	1	Method 1	Hg-Xe
TGVSNTLENEFK	40	51	47	1	Method 1	Hg-Xe
METDLQAK	60	67	60	1	Method 1	Hg-Xe
M*ETDLQAK	60	67	60-61	1	Method 1	Hg-Xe
METDLQAK	60	67	61	1	Method 1	Hg-Xe
AQAFEQDR	98	105	102-103	1	Method 1	Hg-Xe
AQAFEQDR	98	105	102	1	Method 1	Hg-Xe
IAIVNMGSLFQQVAQK	24	39	27	3	Method 2	Hg-Xe
AIVNMGSLFQQVAQK	25	39	27-29	10	Method 2	Hg-Xe
AIVNMGSLFQ‡QVAQK	25	39	27-30	1	Method 2	Hg-Xe
AIVNMGSLFQQVAQK	25	39	27-29	1	Method 2	LED
ASELQ‡R	54	59	57	1	Method 2	Hg-Xe
M*ETDLQAK	60	67	60-61	1	Method 2	LED
M*ETDLQAK	60	67	60	25	Method 2	Hg-Xe
METDLQAK	60	67	60	8	Method 2	Hg-Xe
METDLQAK	60	67	60-61	1	Method 2	Hg-Xe
M*ETDLQAK	60	67	60-61	3	Method 2	Hg-Xe
METDLQ‡AK	60	67	60-61	1	Method 2	Hg-Xe
M*ETDLQ‡AK	60	67	60-62	1	Method 2	Hg-Xe
E†TDLQAK	61	67	61	3	Method 2	Hg-Xe
TK†LEK†DVM*AQR	81	91	84-85	1	Method 2	LED
L†EKDVM*AQR	83	91	83-84	1	Method 2	LED
QTFAQK†AQAFEQDR	92	105	102	2	Method 2	LED
AQAFEQDR	98	105	101	1	Method 2	LED

Sequence	Start Residue	End Residue	Modified Residue	# of Spectra	Enrichment	Lamp
AQAFEQDR	98	105	103-104	1	Method 2	LED
AQAFEQDR	98	105	101	22	Method 2	Hg-Xe
AQAFEQ \ddagger DR	98	105	102	9	Method 2	Hg-Xe
AQAFEQDR	98	105	102	1	Method 2	Hg-Xe
AQAFEQDR	98	105	103	1	Method 2	Hg-Xe
GSHMADKIAIVNMGSLFQQVAQK	17	39	27-29	1	Method 3	LED
IAIVNMGSLFQQVAQK	24	39	27	5	Method 3	LED
IAIVNMGSLFQQVAQK	24	39	27-29	8	Method 3	LED
IAIVNMGSLFQQVAQK	24	39	34	1	Method 3	LED
TGVSNTLENEFKGR	40	53	40-42	1	Method 3	LED
TGVSNTLENEFKGR	40	53	42	3	Method 3	LED
TGVSNTLENEFKGR	40	53	42-44	4	Method 3	LED
TGVSNTLENEFKGR	40	53	42-45	1	Method 3	LED
TGVSNTLENEFKGR	40	53	46	1	Method 3	LED
TGVSNTLENEFKGR	40	53	47	17	Method 3	LED
TGVSNTLENEFK	40	51	47	24	Method 3	LED
T \ddagger GVSNTLENEFK	40	51	47	9	Method 3	LED
TGVSNTLENEFKG	40	52	47	9	Method 3	LED
TGVSNTLENEFKGR	40	53	51	11	Method 3	LED
TGVSNTLENEFKGR	40	53	43	7	Method 3	LED
TGVSNTLENEFKGR	40	53	44	1	Method 3	Hg-Xe
TGVSNTLENEFK	40	51	47	24	Method 3	Hg-Xe
TGVSNTLENEFKGR	40	53	47	18	Method 3	Hg-Xe
TGVSNTLENEFKGR	40	53	51	12	Method 3	Hg-Xe
ASELQR	54	59	54	2	Method 3	Hg-Xe
ASELQR	54	59	56	2	Method 3	Hg-Xe
ASELQRMETDLQAK	54	67	60	8	Method 3	Hg-Xe
ASELQRMETDLQAK	54	67	59-61	5	Method 3	Hg-Xe
RM*ETDLQAK	59	67	59-60	10	Method 3	Hg-Xe
M \ddagger ETDLQAK	60	67	60	23	Method 3	LED
M*ETDLQAK	60	67	60-61	8	Method 3	LED
METDLQAK	60	67	60-61	4	Method 3	LED
M*ETDLQAK	60	67	60	40	Method 3	Hg-Xe
METDLQAK	60	67	60	16	Method 3	Hg-Xe
METDLQAK	60	67	60-61	1	Method 3	Hg-Xe
METDLQ \ddagger AK	60	67	63	2	Method 3	Hg-Xe
METDLQ \ddagger AK	60	67	64	1	Method 3	Hg-Xe
AGSDRTKLEKDVMAQR	76	91	83	3	Method 3	LED
AGSDRTKLEKDVMAQR	76	91	83-85	1	Method 3	LED
AGSDRTKLEKDVMAQR	76	91	83	3	Method 3	Hg-Xe
AGSDRTKLEKDVMAQR	76	91	84	3	Method 3	Hg-Xe
T \ddagger KLEKDVMAQR	81	91	81-83	2	Method 3	LED
T \ddagger KLEKDVMAQR	81	91	81-84	1	Method 3	LED
TKLEK \ddagger DVMAQR	81	91	81-85	1	Method 3	LED
TKLEKDVMAQR	81	91	83	2	Method 3	Hg-Xe
TKLEKDVMAQR	81	91	84	4	Method 3	Hg-Xe
TKLEKDVMAQR	81	91	85	7	Method 3	Hg-Xe
TKLEKDVMAQRQTFAQK	81	97	90-94	13	Method 3	Hg-Xe
LEKDVMAQR	83	91	83	18	Method 3	Hg-Xe
LEKDVMAQ \ddagger R	83	91	83	7	Method 3	Hg-Xe
LEKDVMAQR	83	91	85	1	Method 3	Hg-Xe
DVMAQRQTFAQK	86	97	90-95	1	Method 3	Hg-Xe
DVMAQRQTFAQK	86	97	90-94	4	Method 3	Hg-Xe
QTFAQKAQAFEQDR	92	105	97	4	Method 3	LED
QTFAQKAQAFEQDR	92	105	102	5	Method 3	LED

Sequence	Start Residue	End Residue	Modified Residue	# of Spectra	Enrichment	Lamp
QTFAQKAQAFEQDRAR	92	107	97	2	Method 3	Hg-Xe
QTFAQKAQAFEQDR	92	105	97	7	Method 3	Hg-Xe
AQAFEQDR	98	105	101	27	Method 3	LED
A†QAFEQDR	98	105	102	8	Method 3	LED
AQAFEQDR	98	105	103	3	Method 3	LED
AQAFEQDR	98	105	101	26	Method 3	Hg-Xe
AQAFEQDR	98	105	103	5	Method 3	Hg-Xe
SVANSQDIDLVDANAVAYNSSDVKDITADV LK	126	158	135-138	6	Method 3	Hg-Xe
SVANSQDIDLVDANAVAYNSSDVKDITADV LK	126	158	139	1	Method 3	Hg-Xe
W7C						
M*ETDLQAK	60	67	60-61	1	Method 1	LED
M*ETDLQAK	60	67	60	1	Method 1	LED
METDLQAK	60	67	61	1	Method 1	LED
M*ETDLQAK	60	67	60-61	1	Method 1	Hg-Xe
METDLQAK	60	67	61	1	Method 1	Hg-Xe
AQAFEQDR	98	105	102	2	Method 1	LED
AQAFEQDR	98	105	101	1	Method 1	LED
AQAFEQDR	98	105	102	1	Method 1	Hg-Xe
TGVSNTLENEFKGR	40	53	40-42	1	Method 3	LED
TGVSNTLENEFKGR	40	53	43	7	Method 3	LED
TGVSNTLENEFKGR	40	53	47	7	Method 3	LED
TGVSNTLENEFK	40	51	47	13	Method 3	LED
TGVSNTLENEFKGR	40	53	49-53	1	Method 3	LED
TGVSNTLENEFKGR	40	53	51	7	Method 3	LED
RM*ETDLQAK	59	67	59-60	10	Method 3	LED
RM*ETDLQAK	59	67	59-61	1	Method 3	LED
M*ETDLQAK	60	67	60	18	Method 3	LED
METDLQAK	60	67	60	8	Method 3	LED
M*ETDLQAK	60	67	61	1	Method 3	LED
LEKDVMAQR	83	91	83	1	Method 3	LED
AQAFEQDR	98	105	101	23	Method 3	LED
AQAFEQDR	98	105	102	1	Method 3	LED
AQAFEQDR	98	105	103	4	Method 3	LED
SVANSQ‡DIDLVDANAVAYNSSDVKDITADV LK	126	158	130-138	1	Method 3	LED
SVANSQDIDLVDANAVAYNSSDVKDITADV LK	126	158	138	2	Method 3	LED

OmpA(W7C)[MTS-TFMD] and OmpA(T144C)[MTS-TFMD]

Table 2 Modified peptides of Skp identified by MS/MS from crosslinking experiments using OmpA(W7C) and OmpA(T144C) conjugated to MTS-TFMD crosslinked to Skp. See main text for experimental details and discussion. A number of residues had modifications arising from the treatment with the presence of urea (carbamylation) and protein ageing / treatments used for preparation (deamidation and oxidation). These are marked to the right of residue. ‡ = deamidation, † = carbamylation, * = oxidation.

Sequence	Start Residue	End Residue	Modified Residue	# of Spectra	Enrichment	Lamp
T144C						
GSHMADKIAIVNMGSLFQQVAQK	17	39	29-32	5	Method 3	LED
GSHMADKIAIVNMGSLFQQVAQK	17	39	33	1	Method 3	LED
GSHMADKIAIVNMGSLFQQVAQK	17	39	34	1	Method 3	LED

Sequence	Start Residue	End Residue	Modified Residue	# of Spectra	Enrichment	Lamp
TGVSNTLENEFKGR	40	53	40-41	2	Method 3	LED
TGVSNTLENEFKGR	40	53	42-43	14	Method 3	LED
TGVSNTLENEFK	40	51	42	1	Method 3	LED
TGVSNTLENEFKGR	40	53	43	45	Method 3	LED
TGVSNTLENEFK	40	51	43	46	Method 3	LED
T†GVSNTLENEFKGR	40	53	43	1	Method 3	LED
TGVSNTLENEFKGR	40	53	42-44	1	Method 3	LED
TGVSNTLENEFKGR	40	53	45	9	Method 3	LED
TGVSNTLENEFKGR	40	53	46	4	Method 3	LED
TGVSNTLENEFKGR	40	53	51	2	Method 3	LED
ASELQR	54	59	58	1	Method 3	LED
METDLQAK	60	67	62	1	Method 3	LED
METDLQAK	60	67	64	2	Method 3	LED
METDLQAK	60	67	65	10	Method 3	LED
TKLEKDVMAQR	81	91	88	15	Method 3	LED
TKLEKDVMAQRQTFAQK	81	97	90-94	38	Method 3	LED
LEKDVMAQRQTFAQK	83	97	90-94	5	Method 3	LED
DVMAQRQTFAQK	86	97	90-94	12	Method 3	LED
AQAFEQDR	98	105	99	6	Method 3	LED
AQAFEQDR	98	105	100	1	Method 3	LED
AQAFEQDRAR	98	107	101-102	5	Method 3	LED
AQAFEQDR	98	105	101	3	Method 3	LED
AQAFEQDRAR	98	107	102	2	Method 3	LED
AQAFEQDR	98	105	102	3	Method 3	LED
AQAFEQDRAR	98	107	105-107	25	Method 3	LED
AQAFEQDRAR	98	107	106-107	14	Method 3	LED
RSNEERGKLVTR	108	119	113	12	Method 3	LED
SNEERGKLVTR	109	119	113	7	Method 3	LED
GKLVTR	114	119	116	8	Method 3	LED
W7C						
TGVSNTLENEFKGR	40	53	42	27	Method 3	LED
TGVSNTLENEFK	40	51	42	31	Method 3	LED
TGVSNTLENEFKGR	40	53	43	62	Method 3	LED
TGVSNTLENEFK	40	51	43	65	Method 3	LED
T†GVSNTLENEFKGR	40	53	43	4	Method 3	LED
TGVSNTLENEFKGR	40	53	45	2	Method 3	LED
TGVSNTLENEFKGR	40	53	46	9	Method 3	LED
TGVSNTLENEFKGR	40	53	51	1	Method 3	LED
ASELQR	54	59	58	4	Method 3	LED
METDLQAK	60	67	60-61	1	Method 3	LED
METDLQAK	60	67	65	2	Method 3	LED
TKLEKDVMAQR	81	91	88	1	Method 3	LED
AQAFEQDR	98	105	98-99	4	Method 3	LED
AQAFEQDR	98	105	99	1	Method 3	LED
AQAFEQDR	98	105	100	6	Method 3	LED
AQAFEQDR	98	105	101	8	Method 3	LED
AQAFEQDRAR	98	108	103-107	4	Method 3	LED
AQAFEQDRAR	98	108	105-107	7	Method 3	LED
AQAFEQDRAR	98	108	106	22	Method 3	LED
RSNEER	108	113	109	2	Method 3	LED
RSNEERGKLVTR	108	119	113	4	Method 3	LED
SNEERGKLVTR	109	119	112-113	7	Method 3	LED
SNEERGKLVTR	109	119	113	2	Method 3	LED

Appendix B

List of modified peptides identified by XL-MS/MS from SurA x OmpA crosslinking experiments

OmpA(W7C)[MTS-diazirine] and OmpA(T144C)[MTS-diazirine]

Table 3 Modified peptides of SurA identified by MS/MS from crosslinking experiments using OmpA(W7C) and OmpA(T144C) conjugated to MTS-diazirine crosslinked to SurA. See main text for experimental details and discussion. A number of residues had modifications arising from the treatment with the presence of urea (carbamylation) and protein ageing / treatments used for preparation (deamidation and oxidation). These are marked to the right of residue. ‡ = deamidation, † = carbamylation, * = oxidation.

Sequence	Start Residue	End Residue	Modified Residue	# of Spectra	Enrichment	Lamp
T144C						
IMDQILQM*GQK	55	66	55-58	2	Method 2	Hg-Xe
IMDQILQM*GQK	55	66	55-58	1	Method 2	Hg-Xe
IMDQILQM*GQK	55	66	62-63	4	Method 2	Hg-Xe
LAYDGLNYNTYR	98	109	100	1	Method 2	Hg-Xe
K†EM*IISEVR	114	122	114	2	Method 2	Hg-Xe
K†EMIISEVR	114	122	114-115	3	Method 2	Hg-Xe
K†EM*IISEVR	114	122	114-115	2	Method 2	Hg-Xe
KEMIISEVR	114	122	115-116	5	Method 2	Hg-Xe
KEMIISEVR	114	122	116	6	Method 2	Hg-Xe
EMIISEVR	115	122	115	3	Method 2	Hg-Xe
EMIISEVR	115	122	118	1	Method 2	Hg-Xe
EM*IISEVR	115	122	120-121	2	Method 2	Hg-Xe
EMIISEVR	115	122	120	17	Method 2	Hg-Xe
IQELPGIFAQALSTAK	216	231	216-218	3	Method 2	Hg-Xe
FSQ‡DPGSANQGGDLGWATPDIFDPAFR	308	334	308-320	10	Method 2	Hg-Xe
GSHMAPQVVVDKVAAVVNNGVVLESVDVDGLMQSVK	-4	30	-4-1	1	Method 3	LED
GSHMAPQVVVDK	-4	7	4	1	Method 3	LED
GSHMAPQVVVDK	-4	7	5	1	Method 3	LED
GSHMAPQVVVDK	-4	7	6-7	4	Method 3	LED
AAVVNNGVVLESVDVDGLM*QSVK	9	30	23-26	8	Method 3	LED
AAVVNNGVVLESVDVDGLM*QSVK	9	30	23-26	6	Method 3	Hg-Xe
QQLPDDATLR	38	47	42	1	Method 3	Hg-Xe
QQLPDDATLR	38	47	43	2	Method 3	Hg-Xe
QQLPDDATLRHQIMER	38	53	48	15	Method 3	LED

Sequence	Start Residue	End Residue	Modified Residue	# of Spectra	Enrichment	Lamp
LIMDQIILQMGQKMGVK	54	70	66-70	1	Method 3	LED
ISDEQLDQAIANIAK	71	85	71-74	16	Method 3	Hg-Xe
ISDEQLDQAIANIAK	71	85	71-75	1	Method 3	Hg-Xe
QNNMTLDQMR	86	95	88	1	Method 3	LED
SRLAYDGLNYNTYR	96	109	100	10	Method 3	LED
LAYDGLNYNTYRNQIRK	98	114	98-100	1	Method 3	LED
LAYDGLNYNTYR	98	109	100	8	Method 3	Hg-Xe
L†AYDGLNYNTYR	98	109	100	1	Method 3	Hg-Xe
LAYDGLNYNTYR	98	109	108	5	Method 3	Hg-Xe
LAYDGLNYNTYR	98	109	100	13	Method 3	LED
LAYDGLNYNTYR	98	109	105	1	Method 3	LED
LAYDGLNYNTYR	98	109	108	4	Method 3	LED
KEMIISEVR	114	122	115-116	9	Method 3	Hg-Xe
KEM*IISEVR	114	122	115-116	1	Method 3	Hg-Xe
KEMIISEVRNNEVRR	114	128	114	1	Method 3	LED
KEMIISEVRNNEVR	114	127	114-115	7	Method 3	LED
KEMIISEVRNNEVR	114	127	114-116	2	Method 3	LED
KEM*IISEVR	114	122	115-116	2	Method 3	LED
KEMIISEVR	114	122	115-116	14	Method 3	LED
KEMIISEVR	114	122	119	4	Method 3	LED
E†MIISEVR	115	122	115	20	Method 3	Hg-Xe
EMIISEVR	115	122	115	15	Method 3	Hg-Xe
EM*IISEVR	115	122	115-116	5	Method 3	Hg-Xe
EM*IISEVR	115	122	117	1	Method 3	Hg-Xe
EM*IISEVR	115	122	118	2	Method 3	Hg-Xe
E†MIISEVR	115	122	118	1	Method 3	Hg-Xe
E†MIISEVR	115	122	120	22	Method 3	Hg-Xe
EMIISEVR	115	122	120	19	Method 3	Hg-Xe
E†MIISEVR	115	122	119-120	2	Method 3	Hg-Xe
EM*IISEVR	115	122	119-120	1	Method 3	Hg-Xe
EMIISEVR	115	122	119-120	1	Method 3	Hg-Xe
E†MIISEVR	115	122	119-121	4	Method 3	Hg-Xe
EMIISEVRNNEVRR	115	128	117	4	Method 3	LED
EMIISEVRNNEVR	115	127	115	13	Method 3	LED
EMIISEVRNNEVR	115	127	117	1	Method 3	LED
EMIISEVRNNEVR	115	127	119-120	1	Method 3	LED
EMIISEVRNNEVR	115	127	121-122	1	Method 3	LED
EMIISEVRNNEVR	115	127	123-125	1	Method 3	LED
EMIISEVRNNEVR	115	127	125	3	Method 3	LED
EMIISEVR	115	122	115	14	Method 3	LED
EMIISEVR	115	122	120	19	Method 3	LED
KFSEEAASWMQEQR	385	398	386-389	4	Method 3	LED
KFSEEAASWMQEQR	385	398	393-394	2	Method 3	LED
FSEEAASWMQEQR	386	398	388-390	1	Method 3	LED
ASAYVK	399	404	402	10	Method 3	Hg-Xe
ASAYVKILSN	399	408	404	8	Method 3	Hg-Xe
ASAYVK	399	404	402	8	Method 3	LED
ASAYVKILSN	399	408	404	5	Method 3	LED
W7C						
GSHMAPQVVDK	-4	7	5	1	Method 3	LED
AAVVNNGVVLSDVDGLM*QSVK	9	30	23-26	22	Method 3	LED
ISDEQLDQAIANIAK	71	85	71-75	2	Method 3	LED
ISDEQLDQAIANIAK	71	85	73-74	3	Method 3	LED
ISDEQLDQAIANIAK	71	85	74	9	Method 3	LED

Sequence	Start Residue	End Residue	Modified Residue	# of Spectra	Enrichment	Lamp
SRLAYDGLNYNTYR	96	109	100	9	Method 3	LED
SRLAYDGLNYNTYR	96	109	101	1	Method 3	LED
LAYDGLNYNTYR	98	109	100	27	Method 3	LED
LAYDGLNYNTYR	98	109	105	1	Method 3	LED
LAYDGLNYNTYR	98	109	108	6	Method 3	LED
KEMIISEVR	114	122	115	2	Method 3	LED
KEMIISEVR	114	122	115-116	2	Method 3	LED
EMIISEVR	115	122	115	14	Method 3	LED
EMIISEVRNNEVR	115	127	117	2	Method 3	LED
EMIISEVR	115	122	118	1	Method 3	LED
EMIISEVR	115	122	120	17	Method 3	LED
KFSEEAASWMQEQR	385	398	385-390	1	Method 3	LED
FSEEAASWMQEQR	386	398	386	1	Method 3	LED
FSEEAASWMQEQR	386	398	386-388	1	Method 3	LED

OmpA(W7C)[MTS-TFMD] and OmpA(T144C)[MTS-TFMD]

Table 4 Modified peptides of SurA identified by MS/MS from crosslinking experiments using OmpA(W7C) and OmpA(T144C) conjugated to MTS-TFMD crosslinked to SurA. See main text for experimental details and discussion. A number of residues had modifications arising from the treatment with the presence of urea (carbamylation) and protein ageing / treatments used for preparation (deamidation and oxidation). These are marked to the right of residue. ‡ = deamidation, † = carbamylation, * = oxidation.

Sequence	Start Residue	End Residue	Modified Residue	# of Spectra	Enrichment	Lamp
T144C						
LNAAQAR	31	37	32	1	Method 3	LED
QQLPDDATLRHQIMER	38	53	50	1	Method 3	LED
QNNMTLDQMR	86	95	94	12	Method 3	LED
SRLAYDGLNYNTYR	96	109	96-97	10	Method 3	LED
R†LAYDGLNYNTYR	97	109	97-101	1	Method 3	LED
LAYDGLNYNTYR	98	109	103	1	Method 3	LED
LAYDGLNYNTYR	98	109	105	5	Method 3	LED
LAYDGLNYNTYR	98	109	107-108	3	Method 3	LED
LAYDGLNYNTYR	98	109	108	32	Method 3	LED
EMIISEVR	115	122	115-116	5	Method 3	LED
EMIISEVR	115	122	116	7	Method 3	LED
EMIISEVR	115	122	117	2	Method 3	LED
EMIISEVR	115	122	119	6	Method 3	LED
EMIISEVRNNEVR	115	127	121-122	8	Method 3	LED
MLMNRK	380	385	383-384	1	Method 3	LED
MLMNRK	380	385	384-385	2	Method 3	LED
ASAYVK	399	404	402	2	Method 3	LED
W7C						
LNAAQARQQLPDDATLR	31	47	34-39	1	Method 3	LED
QQLPDDATLRHQIMER	38	53	49	4	Method 3	LED
SRLAYDGLNYNTYR	96	109	96-97	10	Method 3	LED
SRLAYDGLNYNTYR	96	109	107-108	1	Method 3	LED
LAYDGLNYNTYR	98	109	100	1	Method 3	LED

Sequence	Start Residue	End Residue	Modified Residue	# of Spectra	Enrichment	Lamp
LAYDGLNYNTYR	98	109	108	19	Method 3	LED
KEMIISEVRNNEVR	114	127	118-122	1	Method 3	LED
KEMIISEVRNNEVR	114	127	121-122	1	Method 3	LED
EMIISEVRNNEVRR	115	128	119-122	1	Method 3	LED
EMIISEVRNNEVRR	115	128	119-127	1	Method 3	LED
EMIISEVRNNEVRR	115	128	119-125	1	Method 3	LED
EMIISEVRNNEVR	115	127	121-122	13	Method 3	LED

Appendix C

Full gels from catalysed and uncatalysed folding kinetics

BAM 20°C

No SurA

+SurA

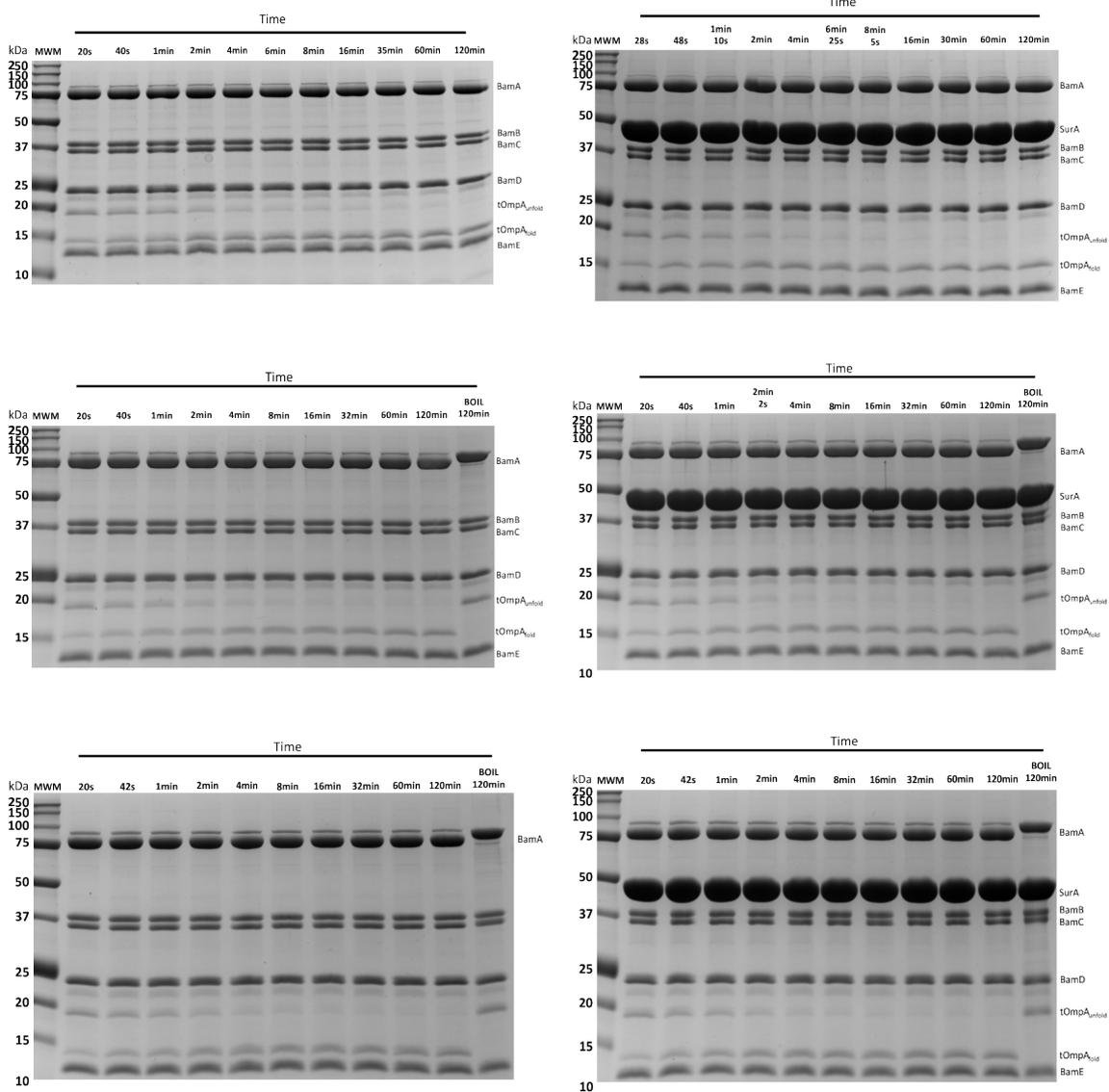


Fig. 1 Full gels of tOmpA kinetics folding into BAM-containing DMPC proteoliposomes in the presence or absence of SurA at 20 °C. Final conditions: (10 μ M SurA), 2 μ M BAM, 1 μ M tOmpA, 1600:1 LPR DMPC, 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0.

BAM 30°C

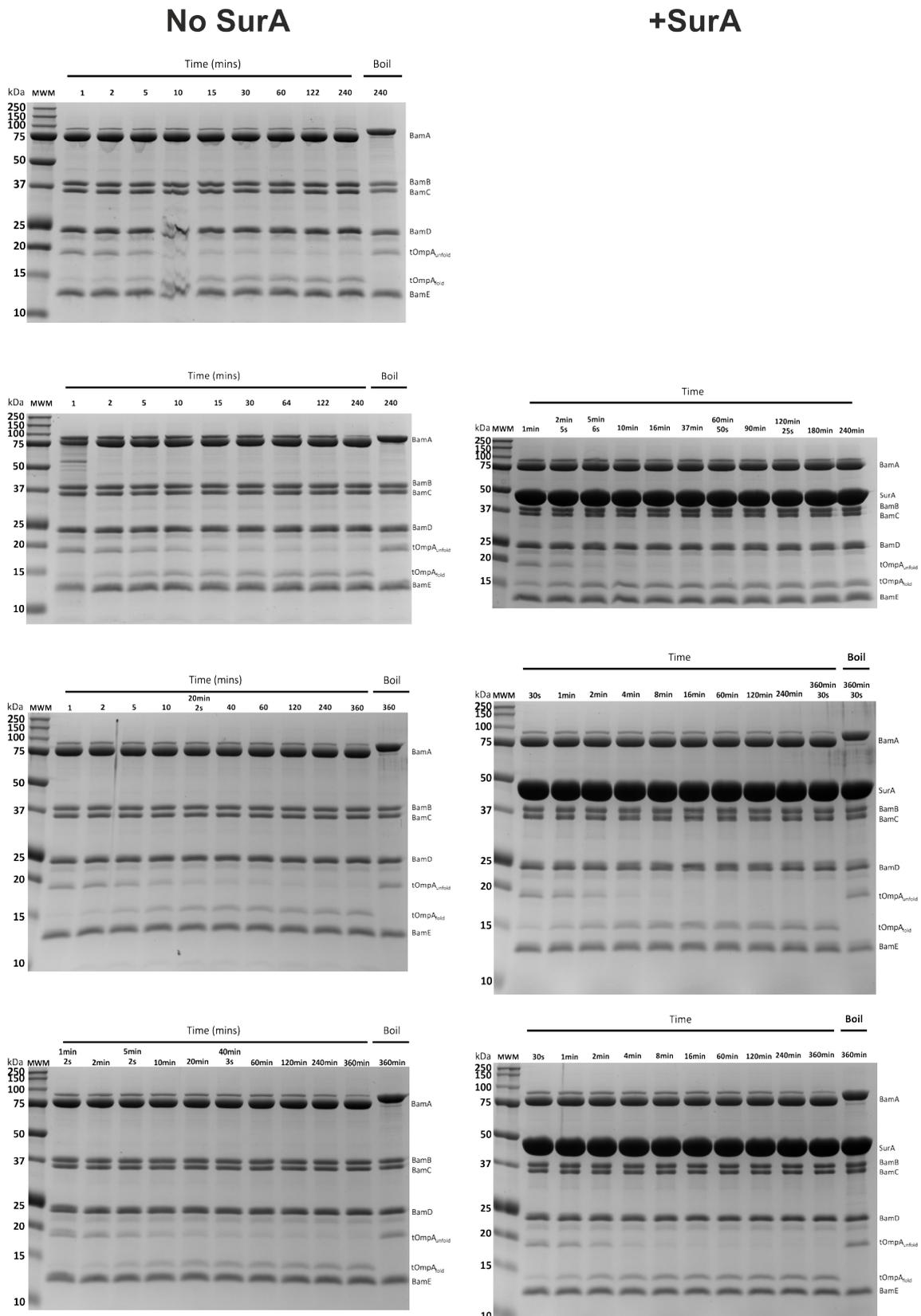


Fig. 3 Full gels of tOmpA kinetics folding into BAM-containing DMPC proteoliposomes in the presence or absence of SurA at 30 °C. Final conditions: (10 μ M SurA), 2 μ M BAM, 1 μ M tOmpA, 1600:1 LPR DMPC, 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0.

BamA 20°C

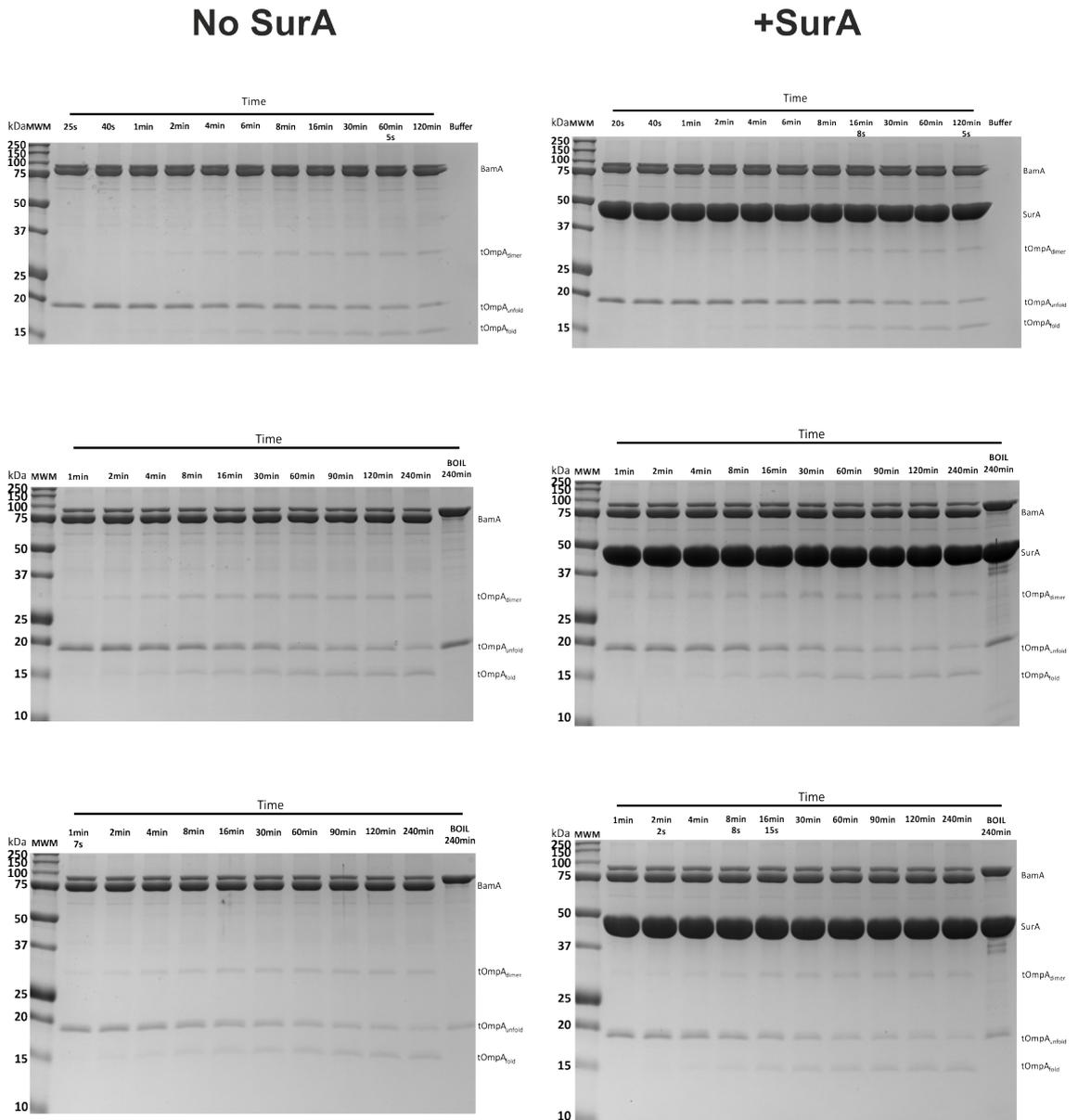


Fig. 4 Full gels of tOmpA kinetics folding into BamA-containing DMPC proteoliposomes in the presence or absence of SurA at 20 °C. Final conditions: (10 μ M SurA), 2 μ M BamA, 1 μ M tOmpA, 1600:1 LPR DMPC, 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0.

BamA 24°C

No SurA

+SurA

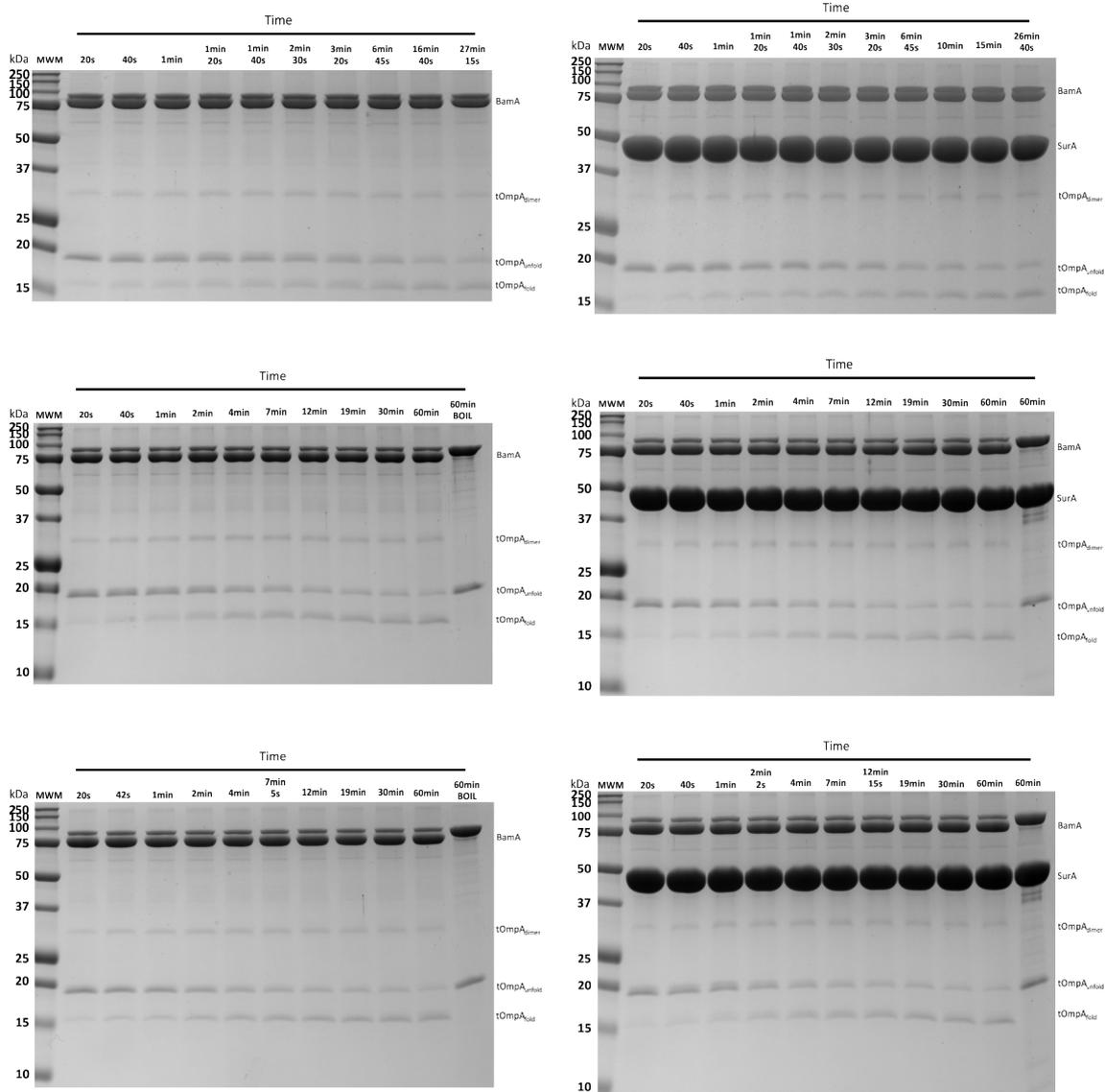


Fig. 5 Full gels of tOmpA kinetics folding into BamA-containing DMPC proteoliposomes in the presence or absence of SurA at 24 °C. Final conditions: (10 μ M SurA), 2 μ M BamA, 1 μ M tOmpA, 1600:1 LPR DMPC, 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0.

BamA 30°C

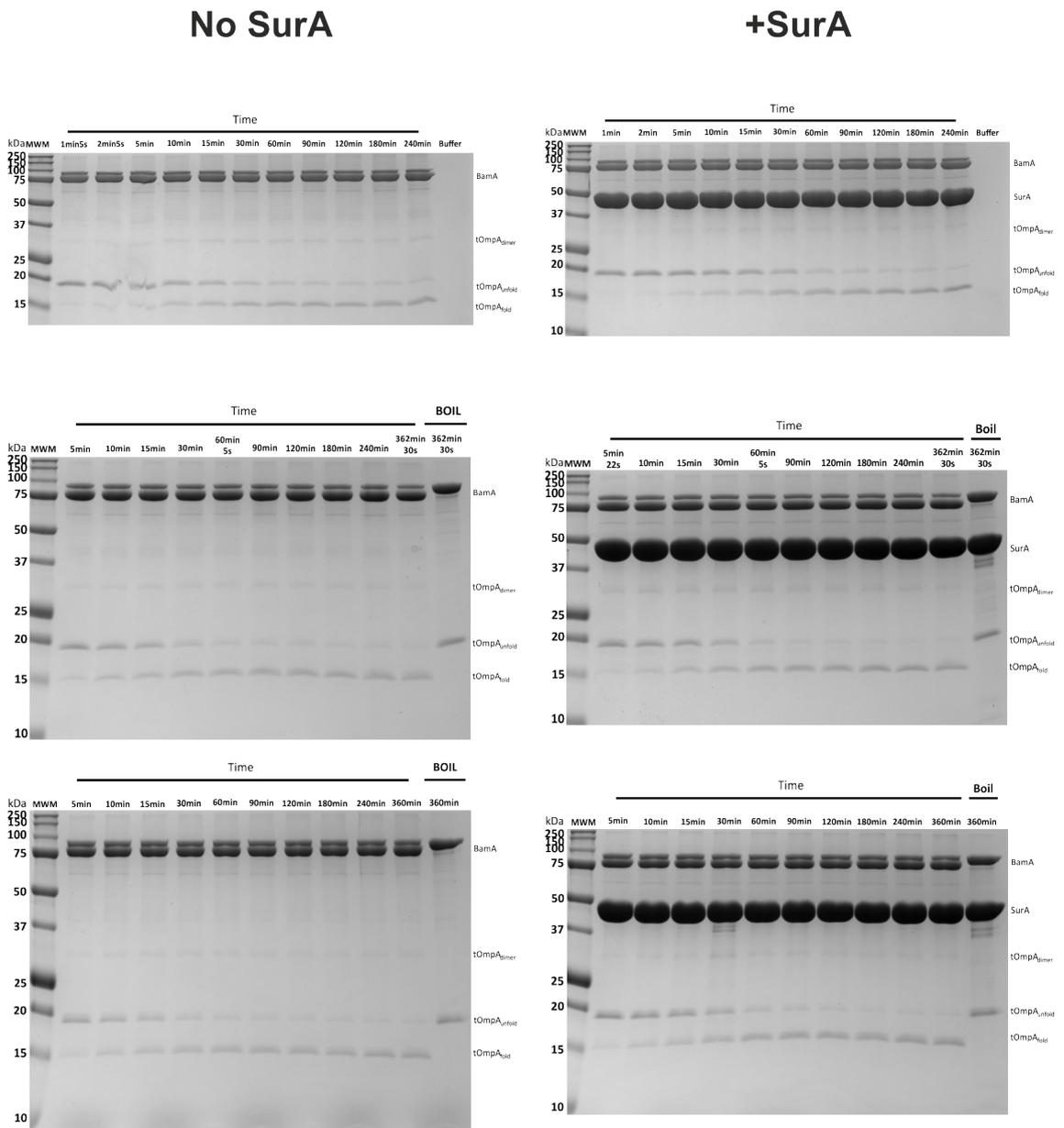


Fig. 6 Full gels of tOmpA kinetics folding into BamA-containing DMPC proteoliposomes in the presence or absence of SurA at 30 °C. Final conditions: (10 μ M SurA), 2 μ M BamA, 1 μ M tOmpA, 1600:1 LPR DMPC, 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0.

Empty 20°C

No SurA

+SurA

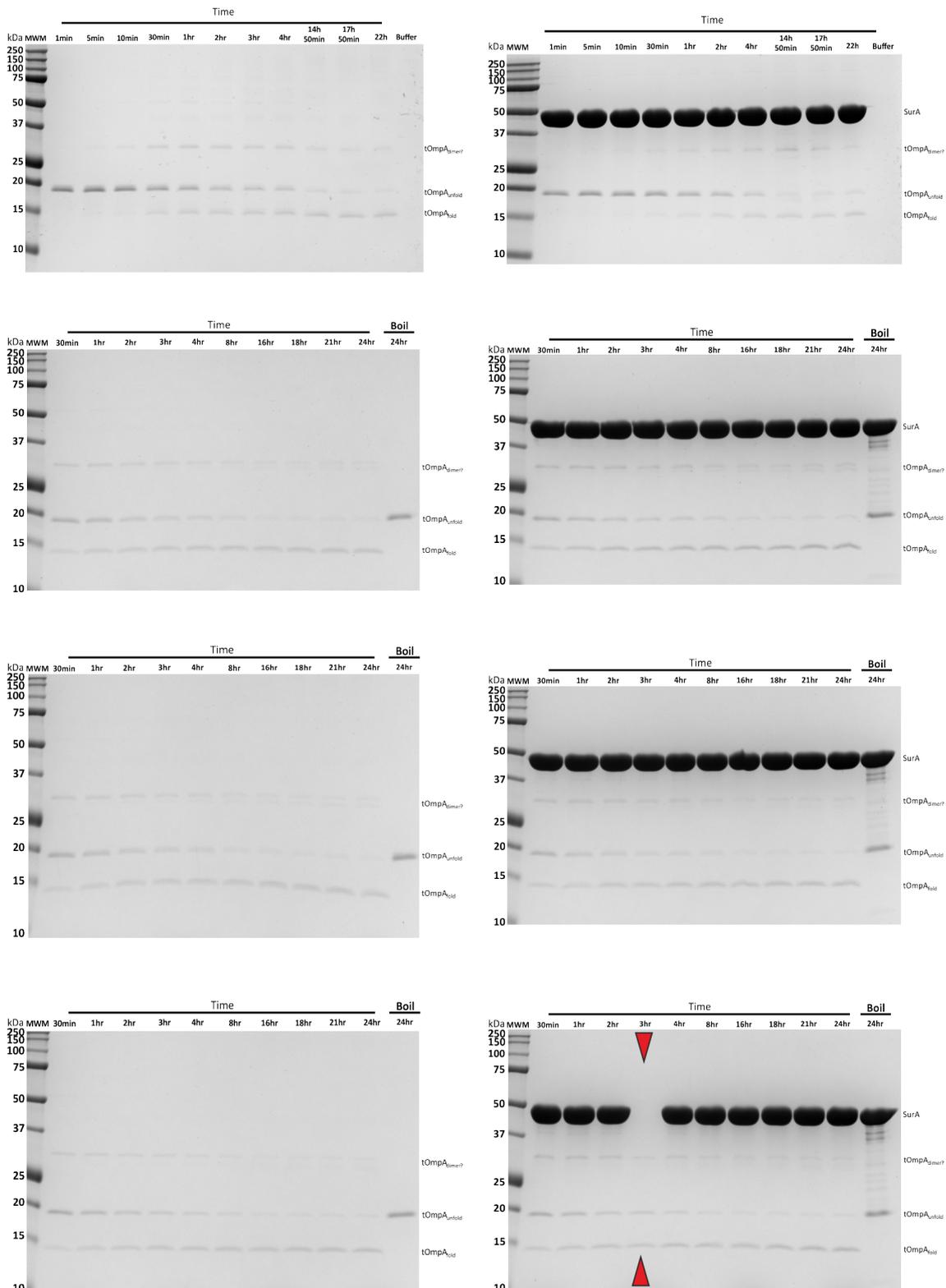


Fig. 7 Full gels of tOmpA kinetics folding into empty DMPC liposomes in the presence or absence of SurA at 20 °C. Final conditions: (10 μ M SurA), 1 μ M tOmpA, 1600:1 LPR DMPC, 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0. Red arrows indicate lanes which were not used in the analysis.

Empty 24°C

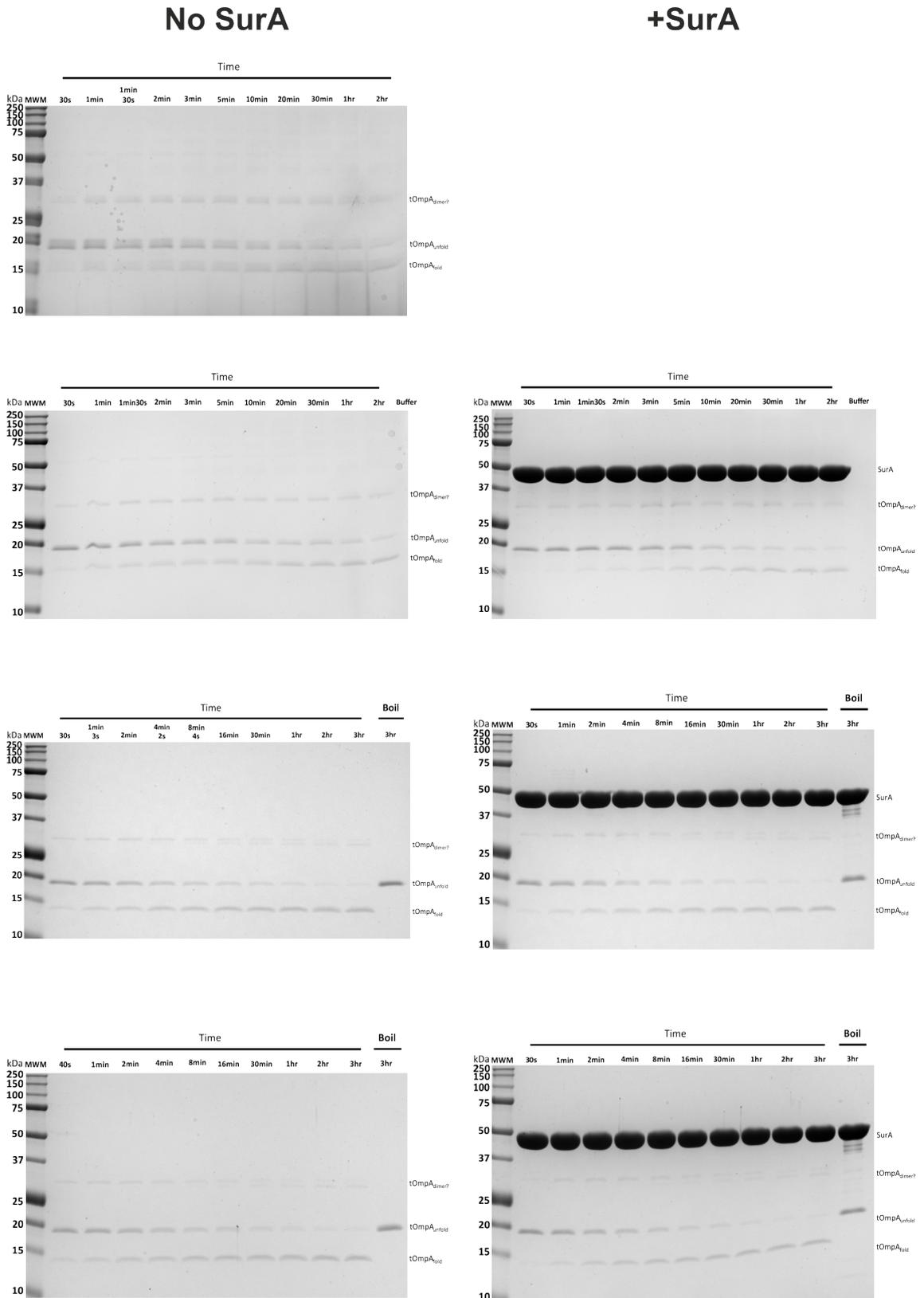


Fig. 8 Full gels of tOmpA kinetics folding into empty DMPC liposomes in the presence or absence of SurA at 24 °C. Final conditions: (10 μM SurA), 1 μM tOmpA, 1600:1 LPR DMPC, 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0.

Empty 30°C

No SurA

+SurA

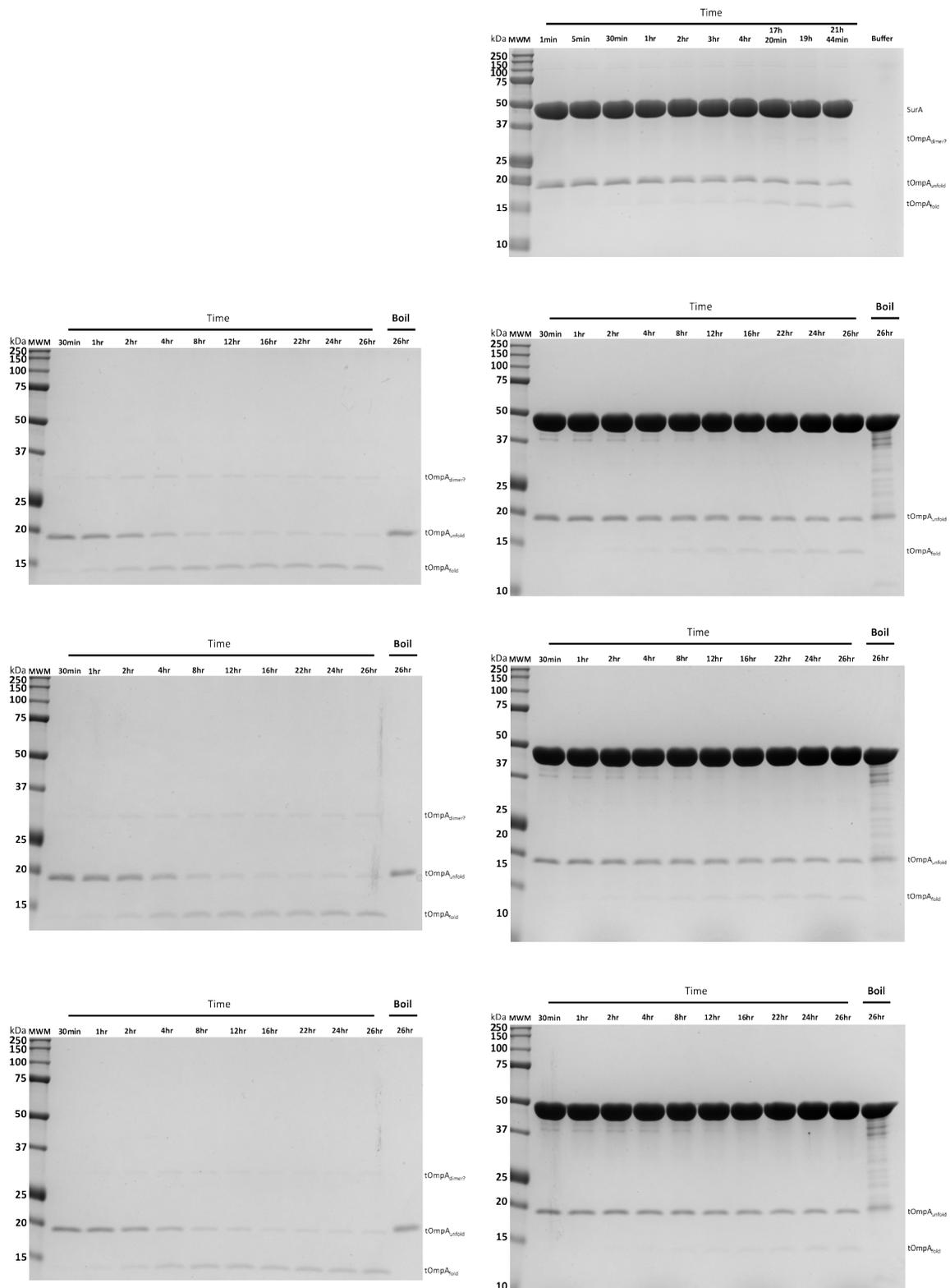


Fig. 9 Full gels of tOmpA kinetics folding into empty DMPC liposomes in the presence or absence of SurA at 30 °C. Final conditions: (10 μ M SurA), 1 μ M tOmpA, 1600:1 LPR DMPC, 0.8 M urea, 150 mM NaCl, 20 mM Tris-Cl pH 8.0.

Appendix D

Single and double exponential fits to folding kinetics by gel

BAM 20°C

No SurA

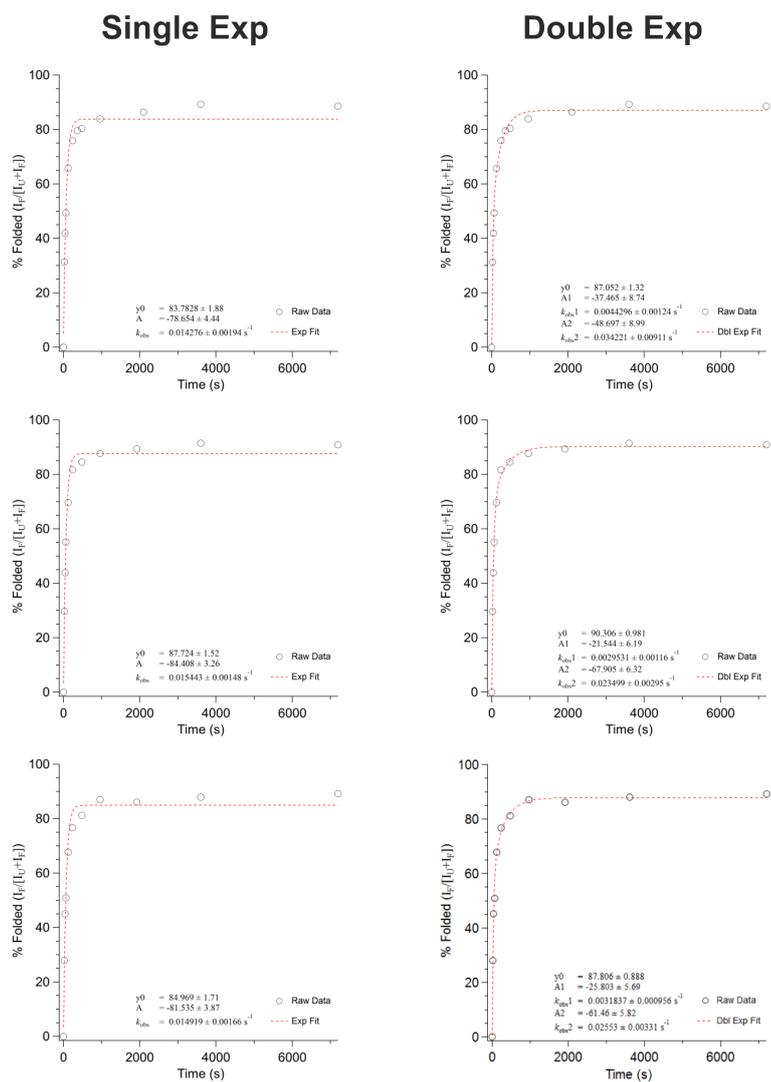


Fig. 10 Exponential fits to tOmpA folding via BAM without SurA at 20 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

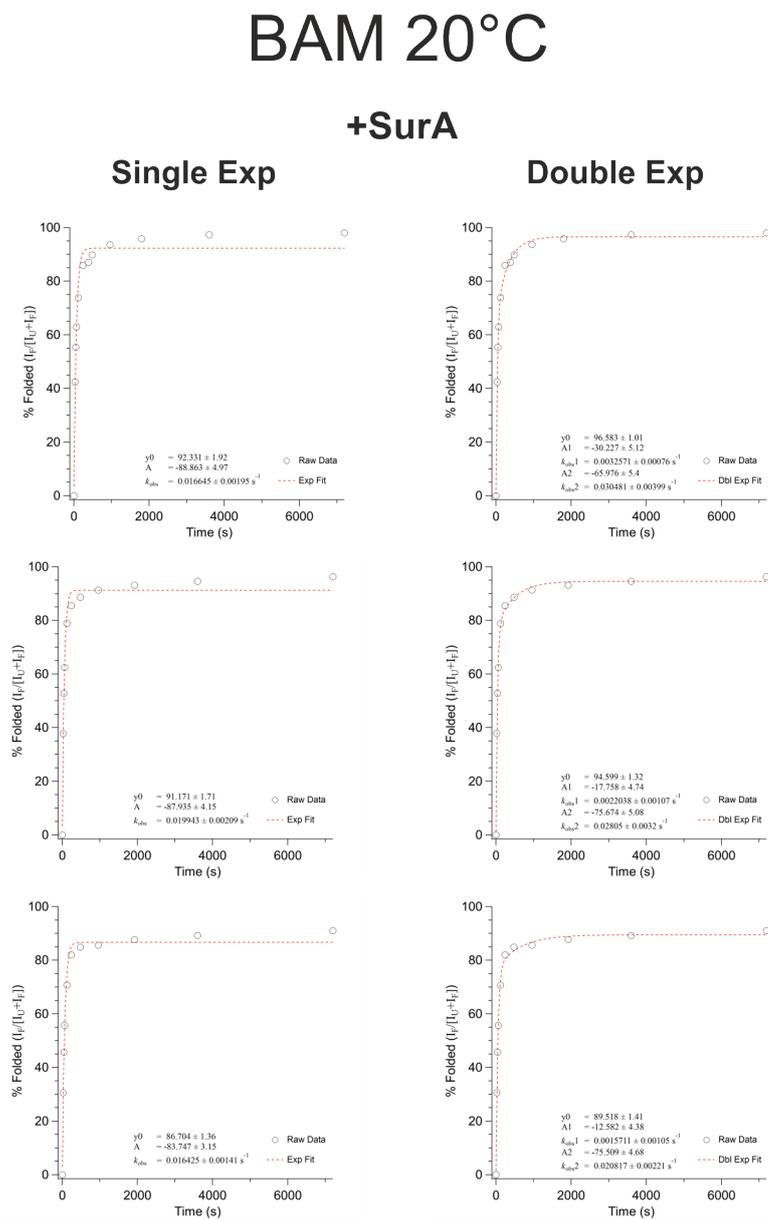


Fig. 11 Exponential fits to tOmpA folding via BAM with SurA at 20 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

BAM 24°C

No SurA

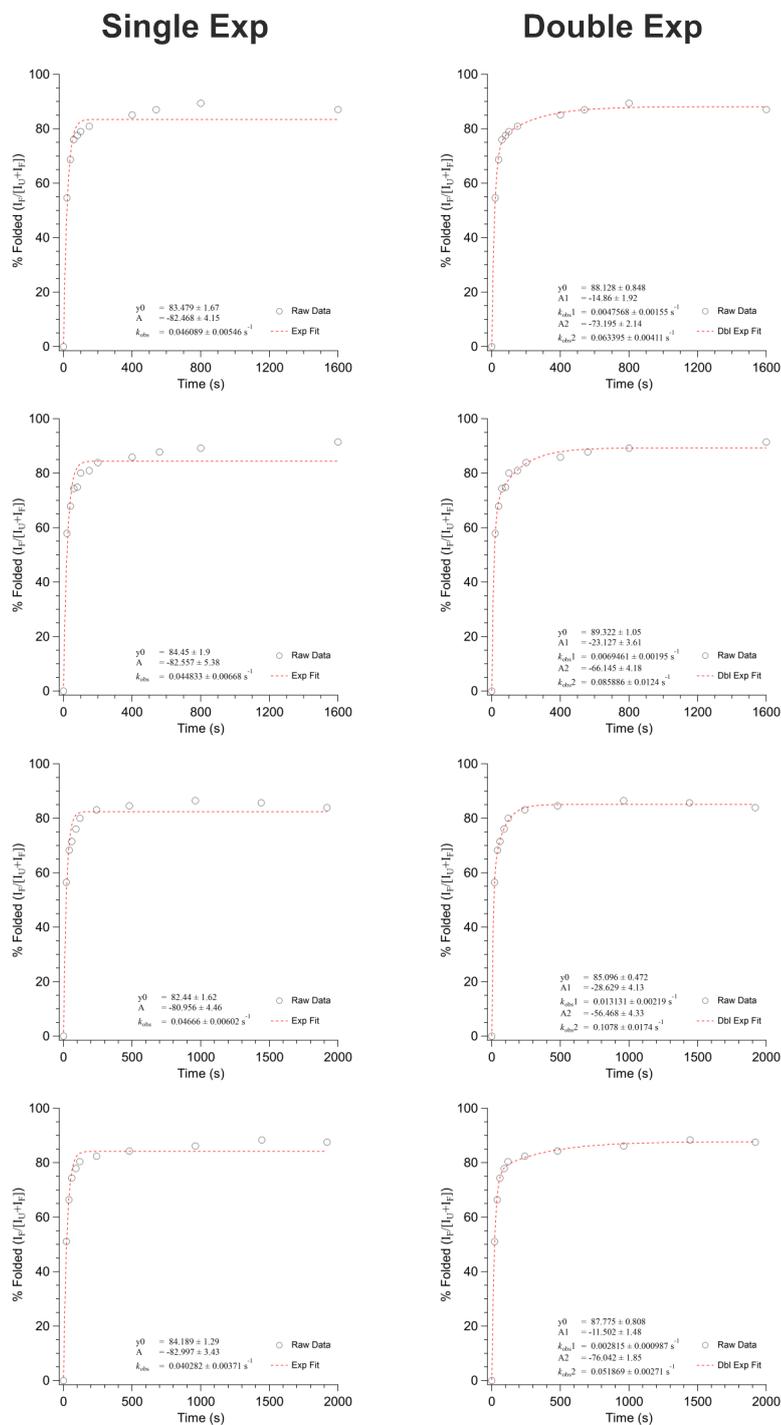


Fig. 12 Exponential fits to tOmpA folding via BAM without SurA at 24 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

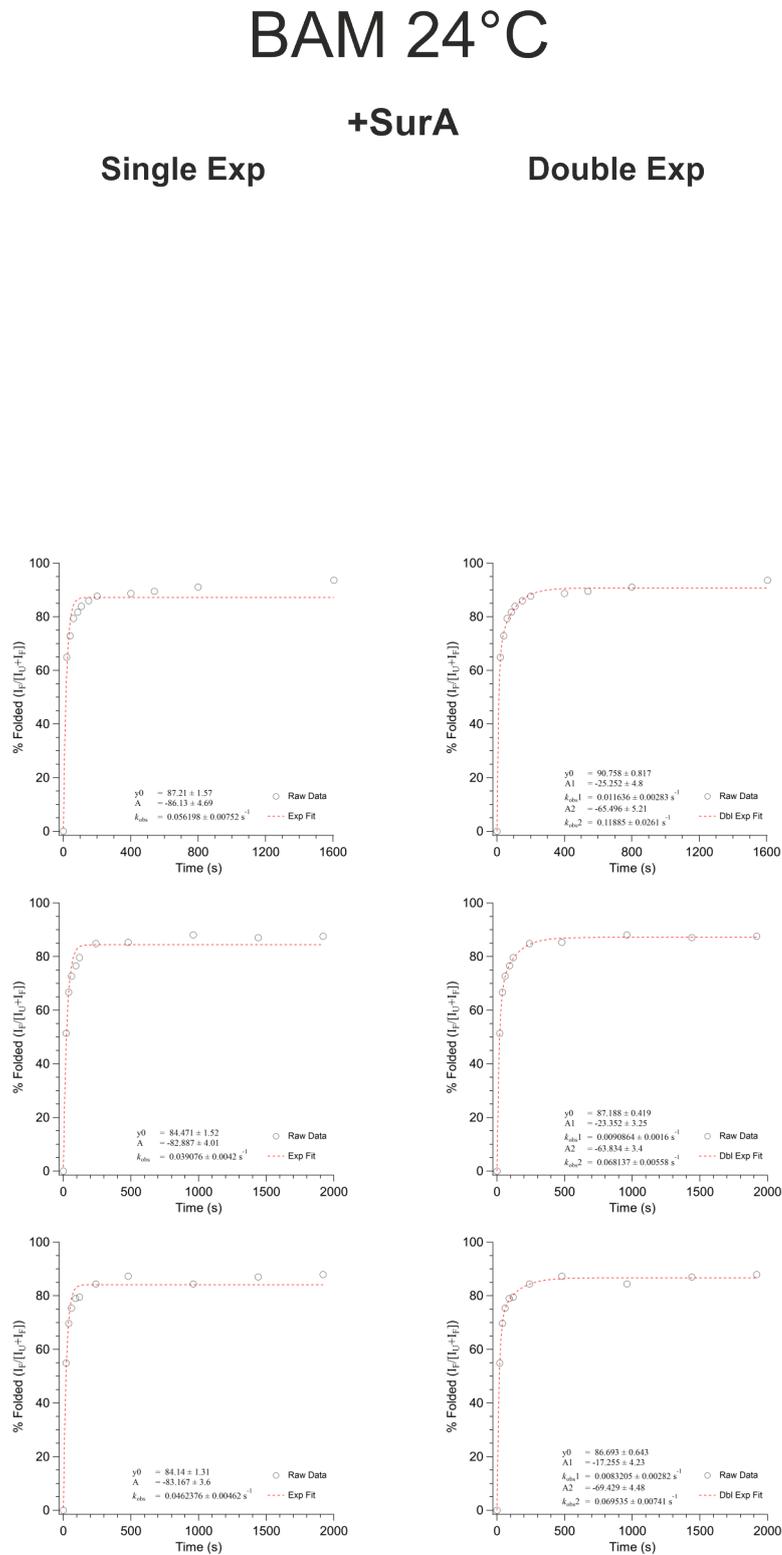


Fig. 13 Exponential fits to tOmpA folding via BAM with SurA at 24 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

BAM 30°C

No SurA

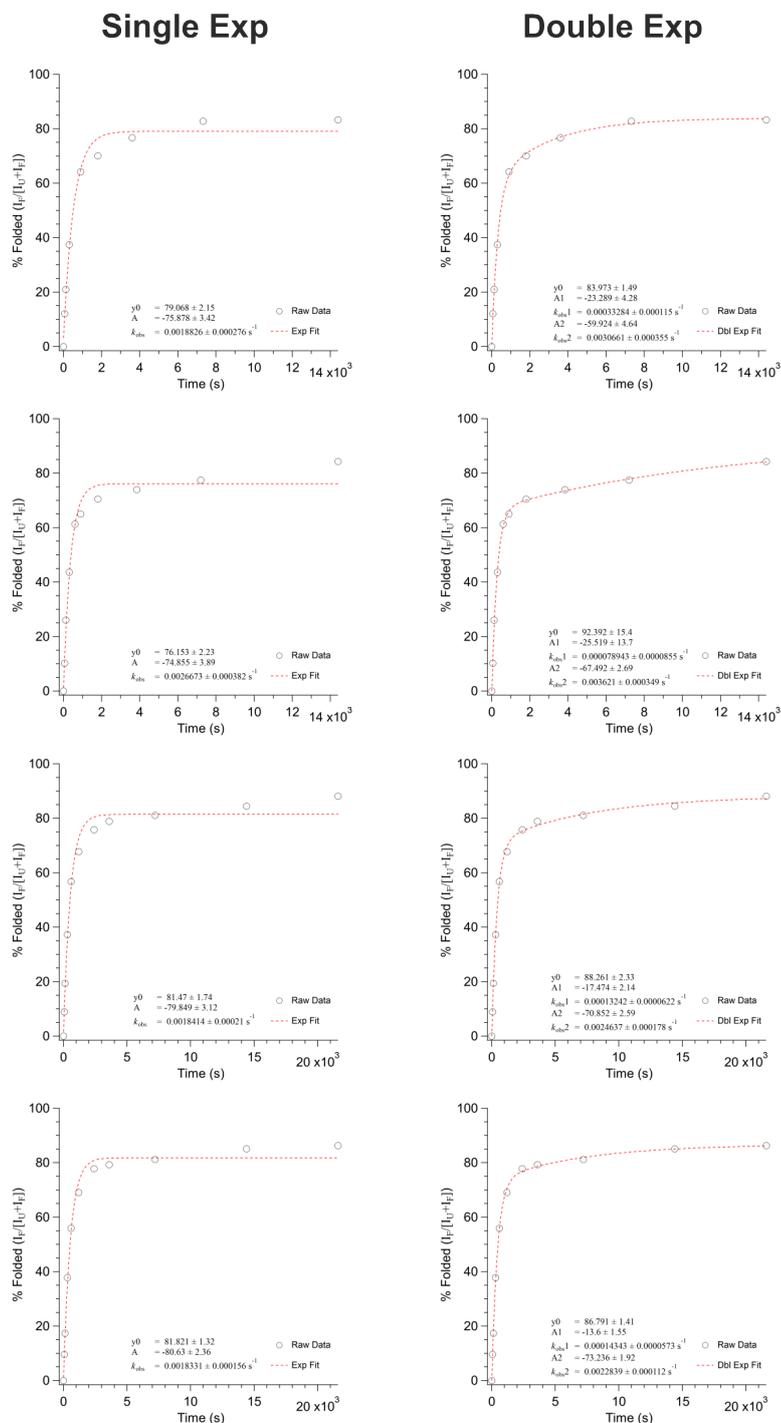


Fig. 14 Exponential fits to tOmpA folding via BAM without SurA at 30 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

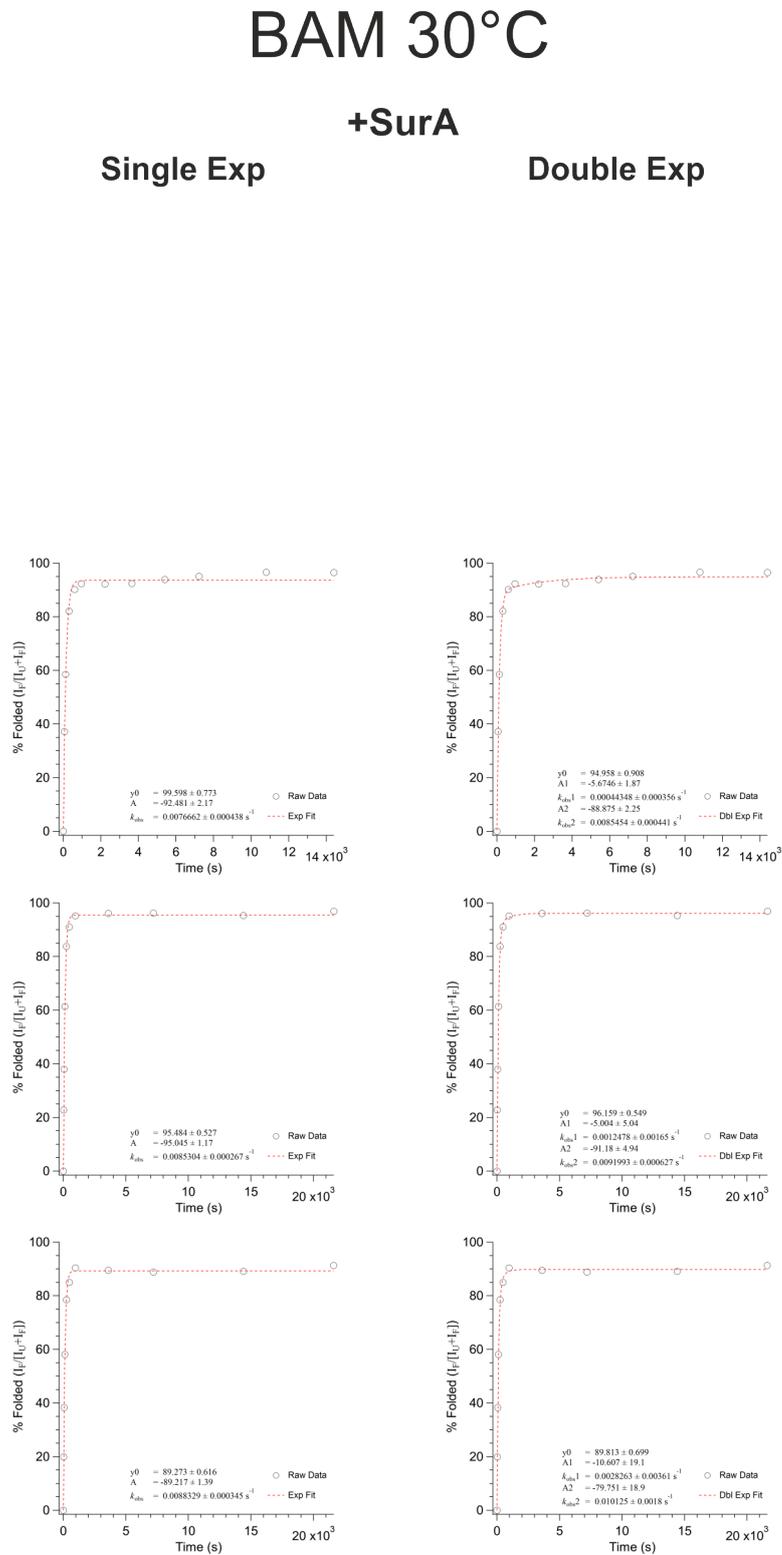


Fig. 15 Exponential fits to tOmpA folding via with SurA at 30 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

BamA 20°C

No SurA

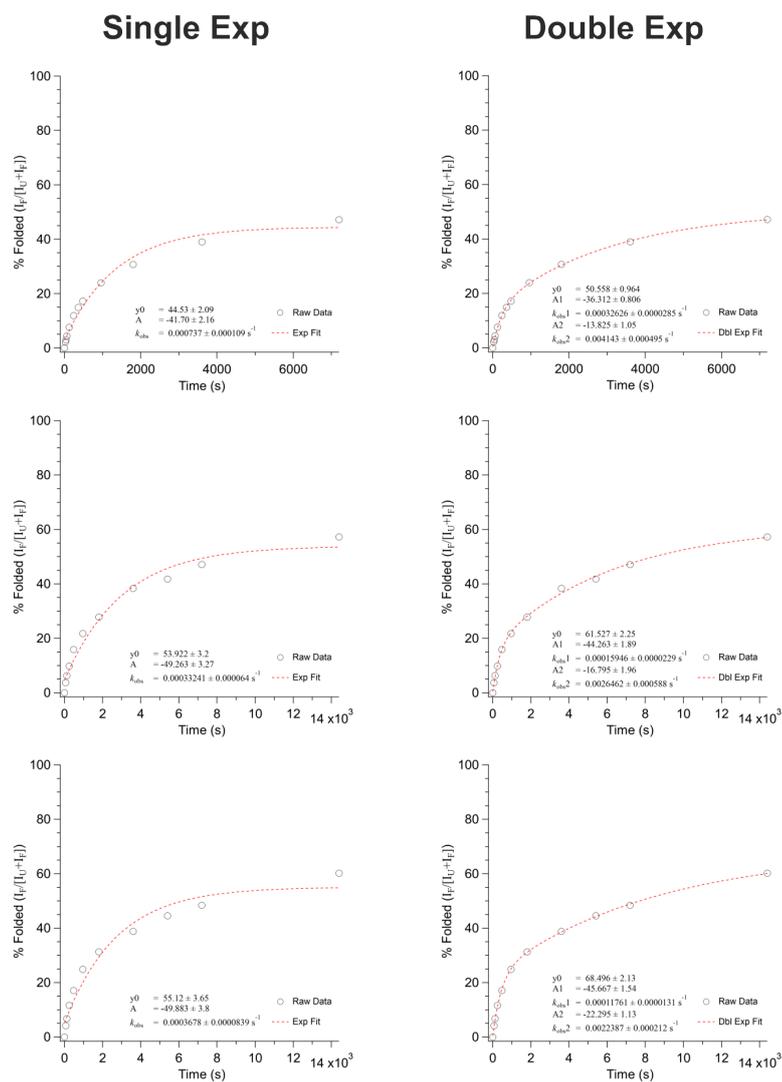


Fig. 16 Exponential fits to tOmpA folding via BamA without SurA at 20 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

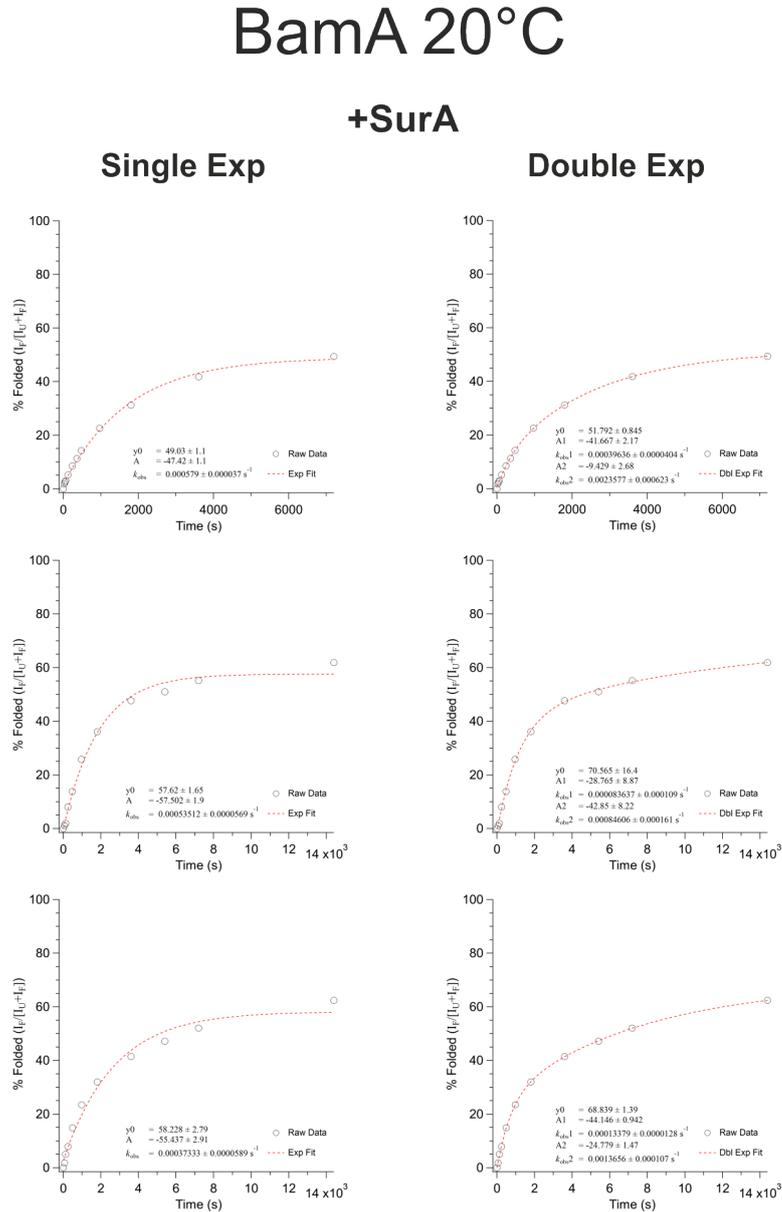


Fig. 17 Exponential fits to tOmpA folding via BamA with SurA at 20 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

BamA 24°C

No SurA

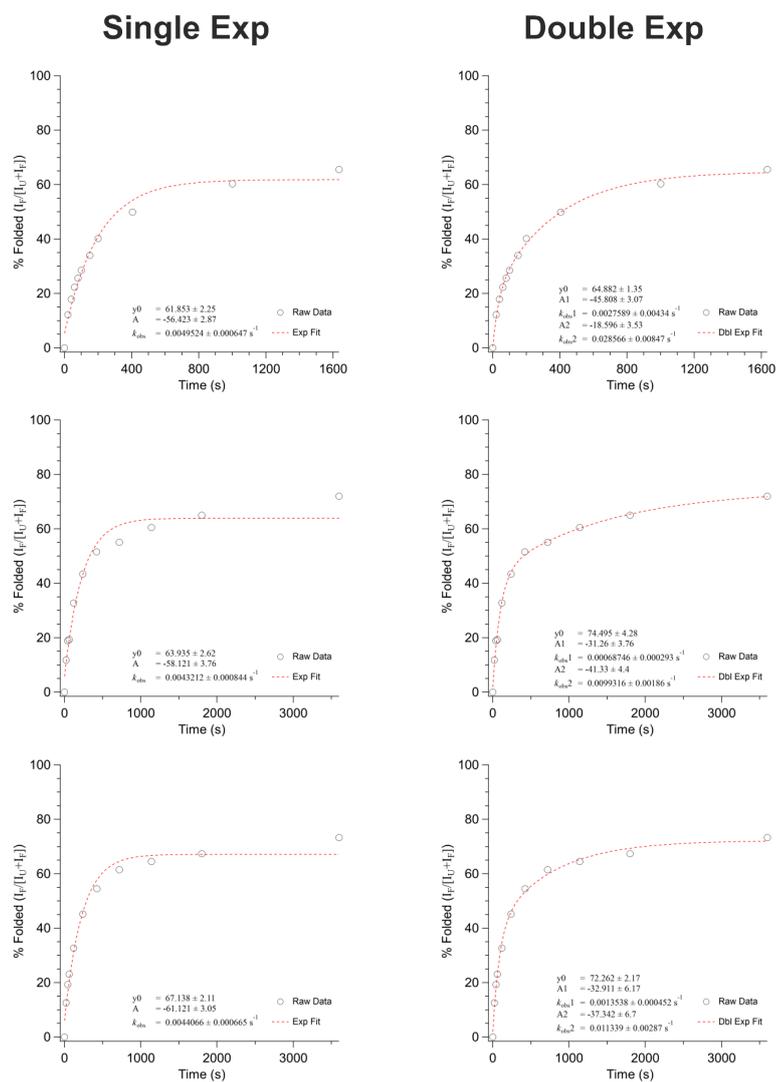


Fig. 18 Exponential fits to tOmpA folding via BamA without SurA at 24 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

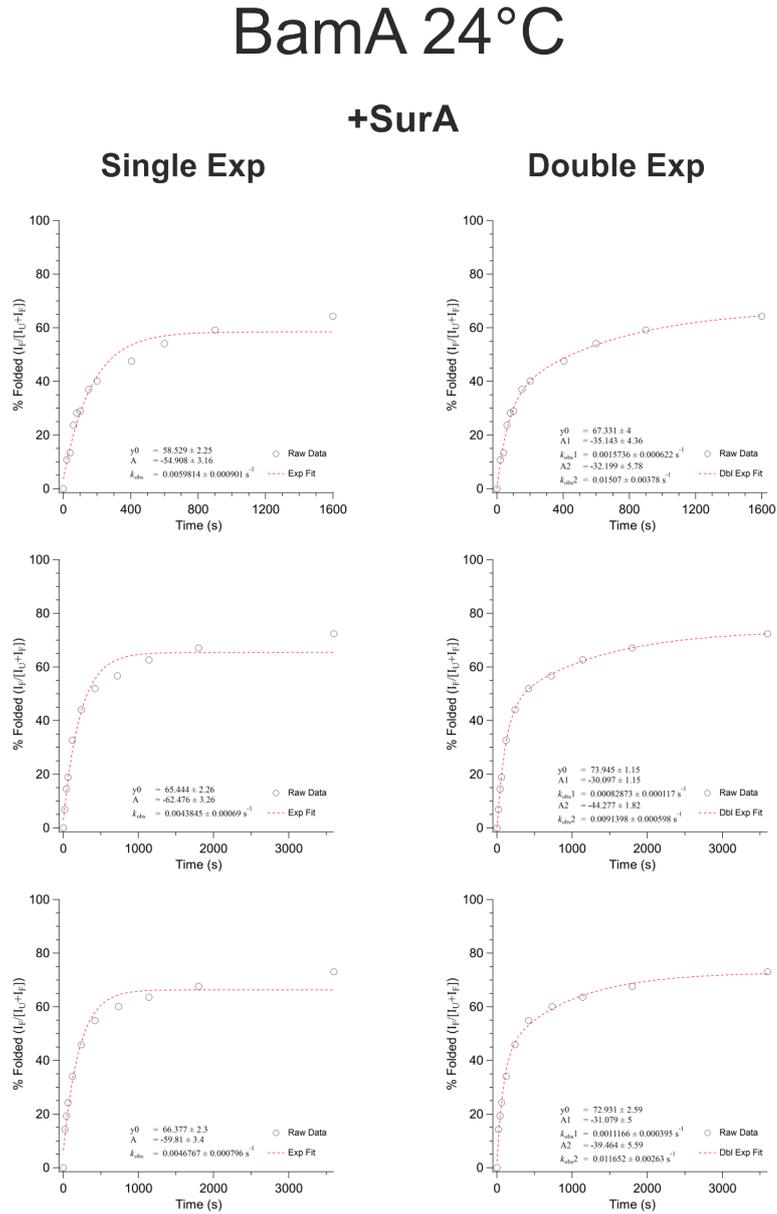


Fig. 19 Exponential fits to tOmpA folding via BamA with SurA at 24 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

BamA 30°C

No SurA

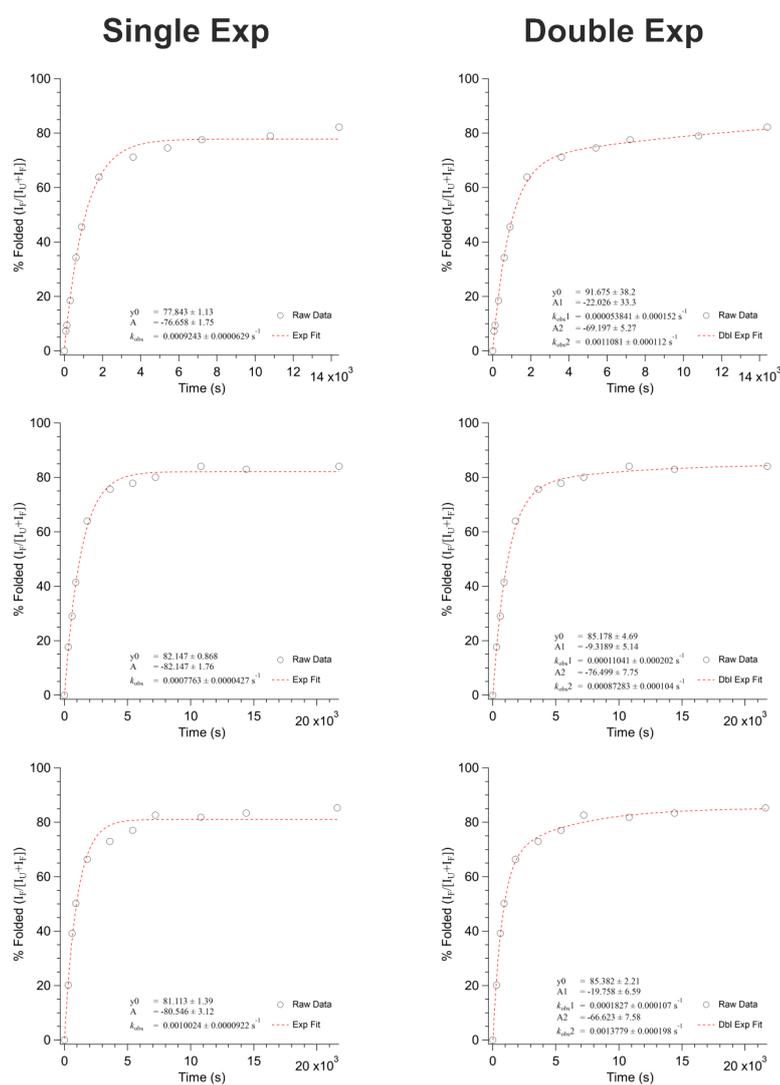


Fig. 20 Exponential fits to tOmpA folding via BamA without SurA at 30 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

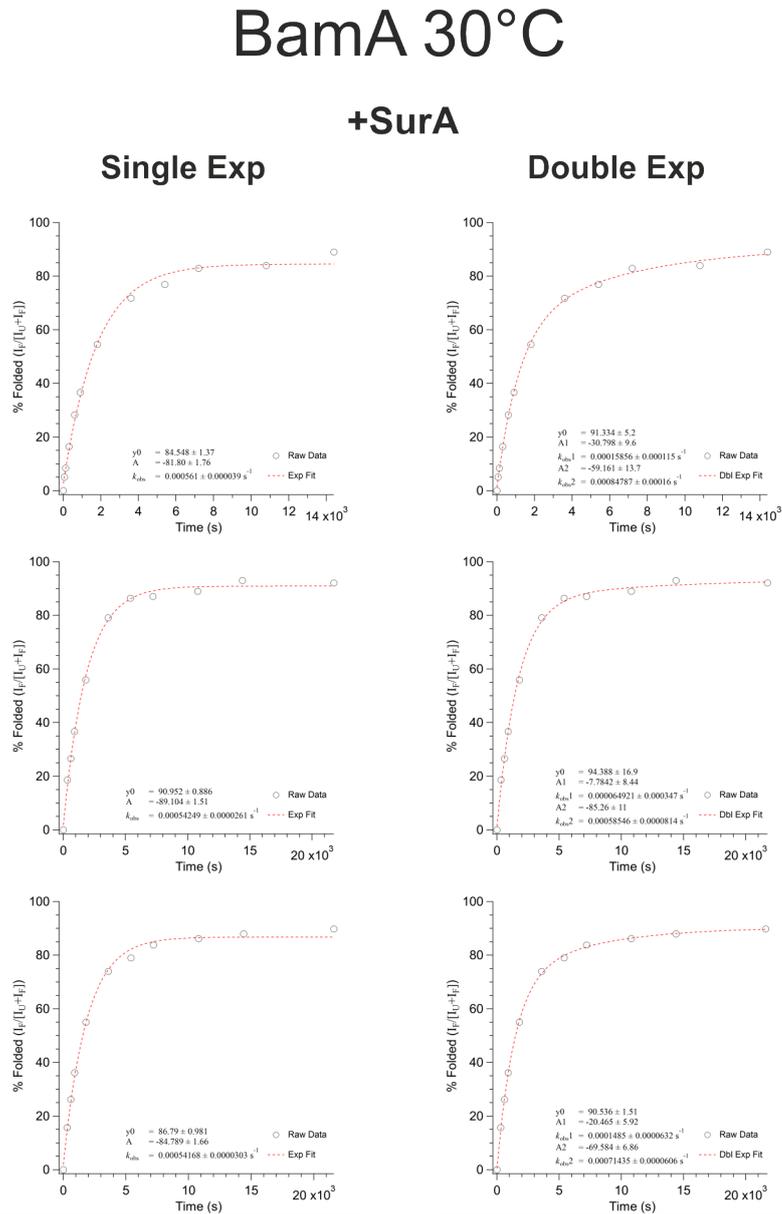


Fig. 21 Exponential fits to tOmpA folding via BamA with SurA at 30 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

Empty 20°C

No SurA

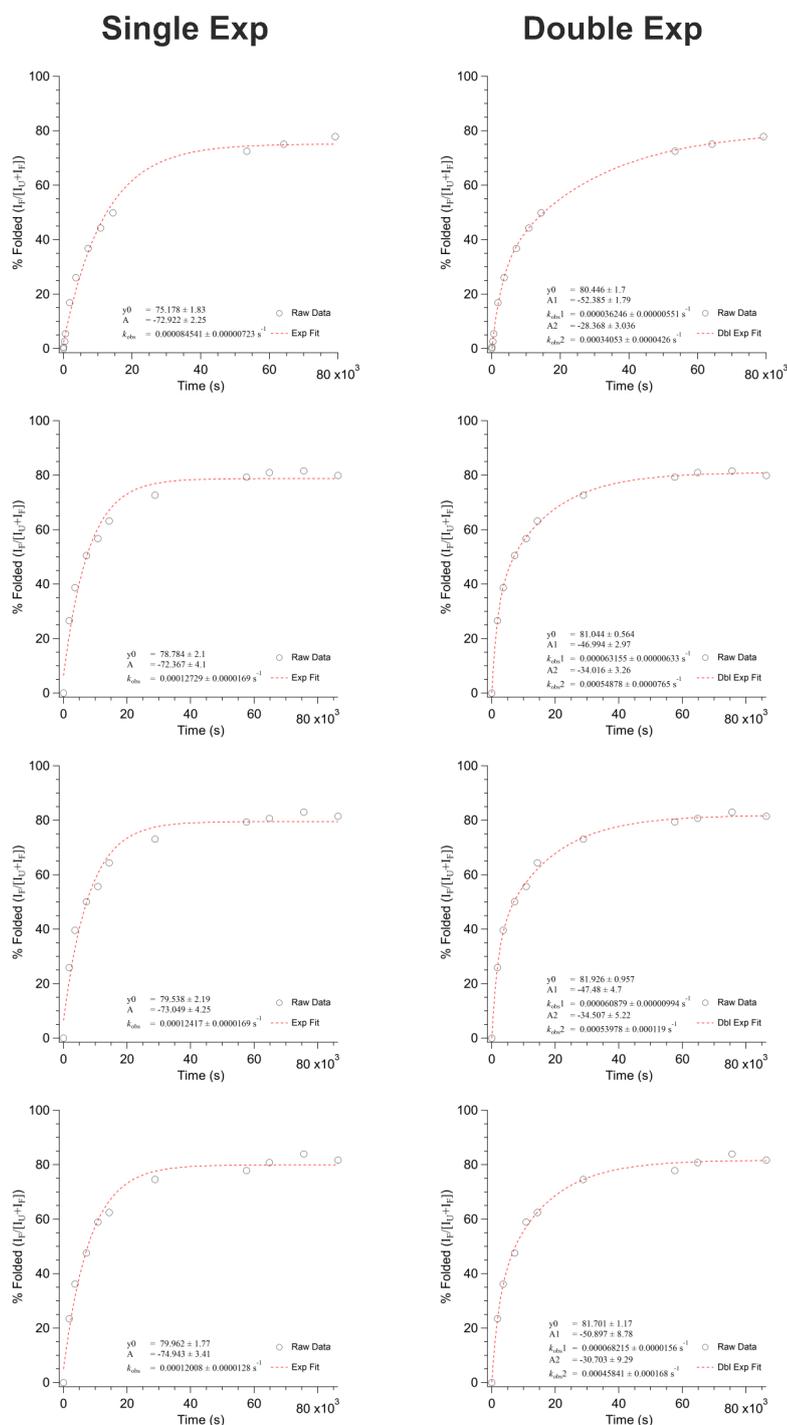


Fig. 22 Exponential fits to tOmpA folding into empty liposomes without SurA at 20 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

Empty 20°C

+SurA

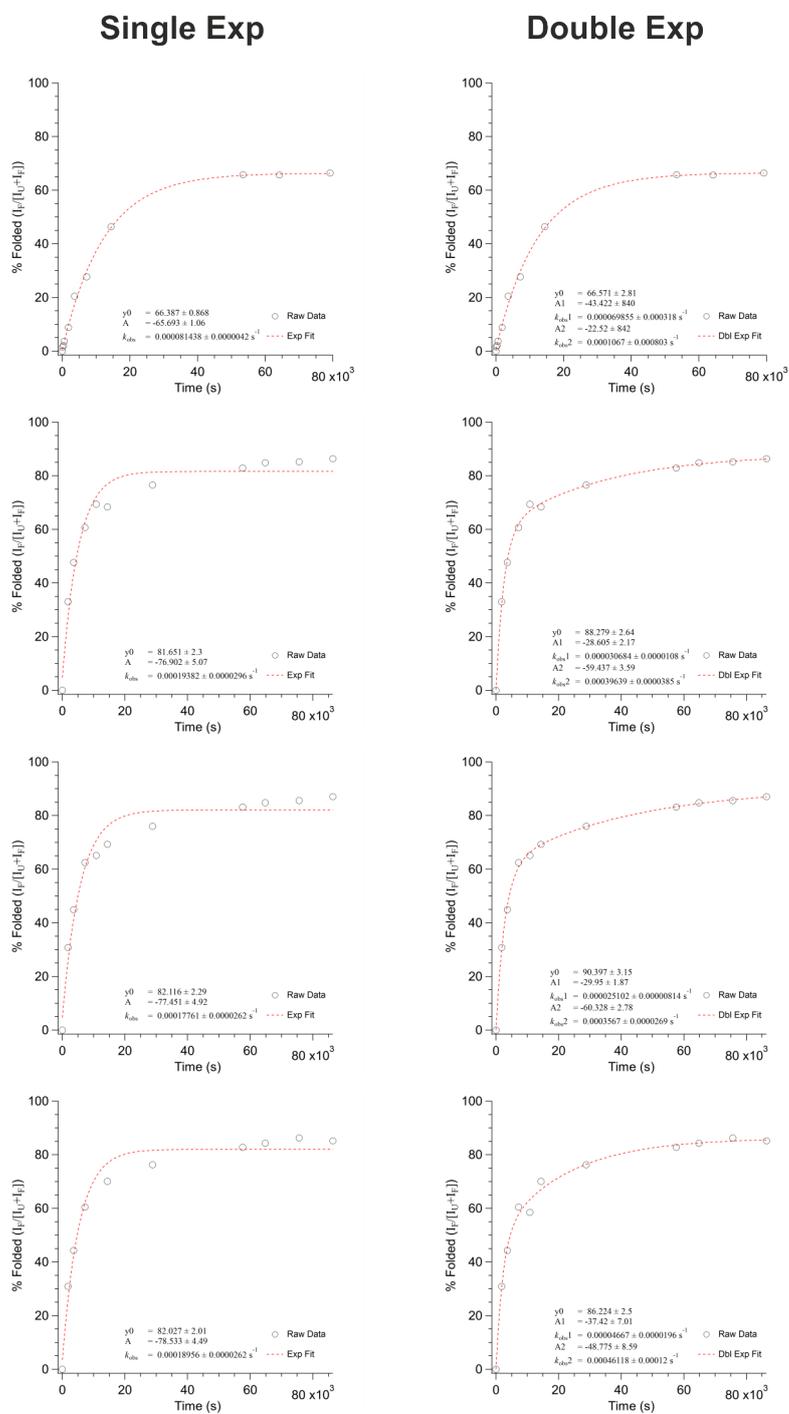


Fig. 23 Exponential fits to tOmpA folding into empty liposomes with SurA) at 20 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

Empty 24°C

No SurA

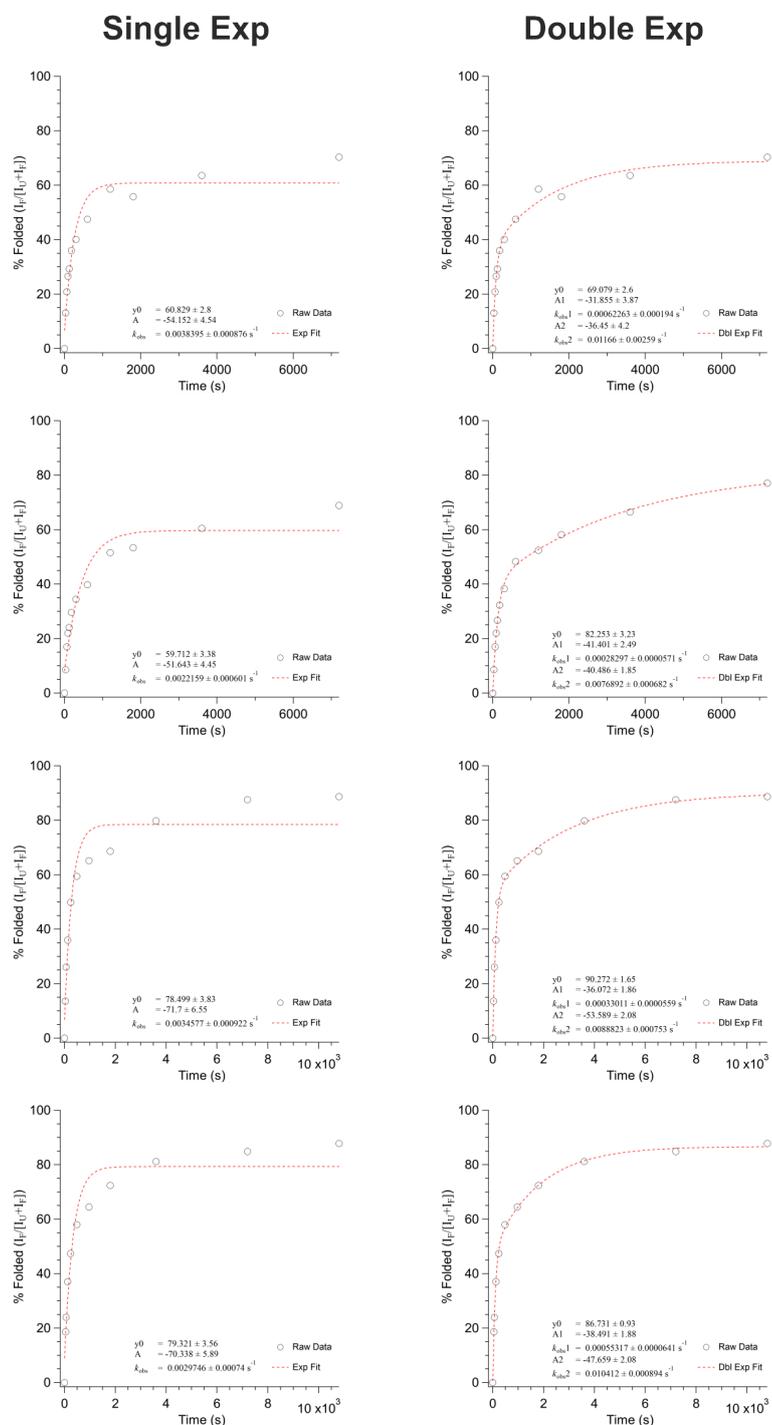


Fig. 24 Exponential fits to tOmpA folding into empty liposomes without SurA at 24 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

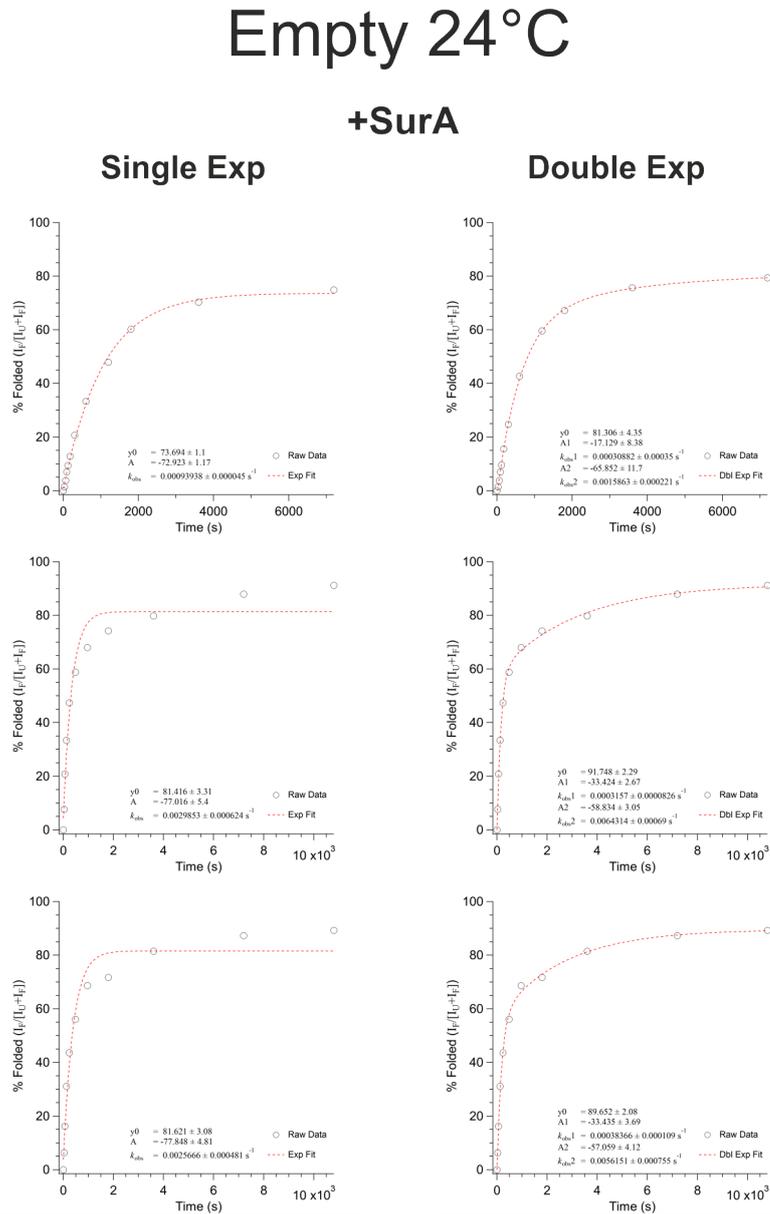


Fig. 25 Exponential fits to tOmpA folding into empty liposomes with SurA at 24 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

Empty 30°C

No SurA

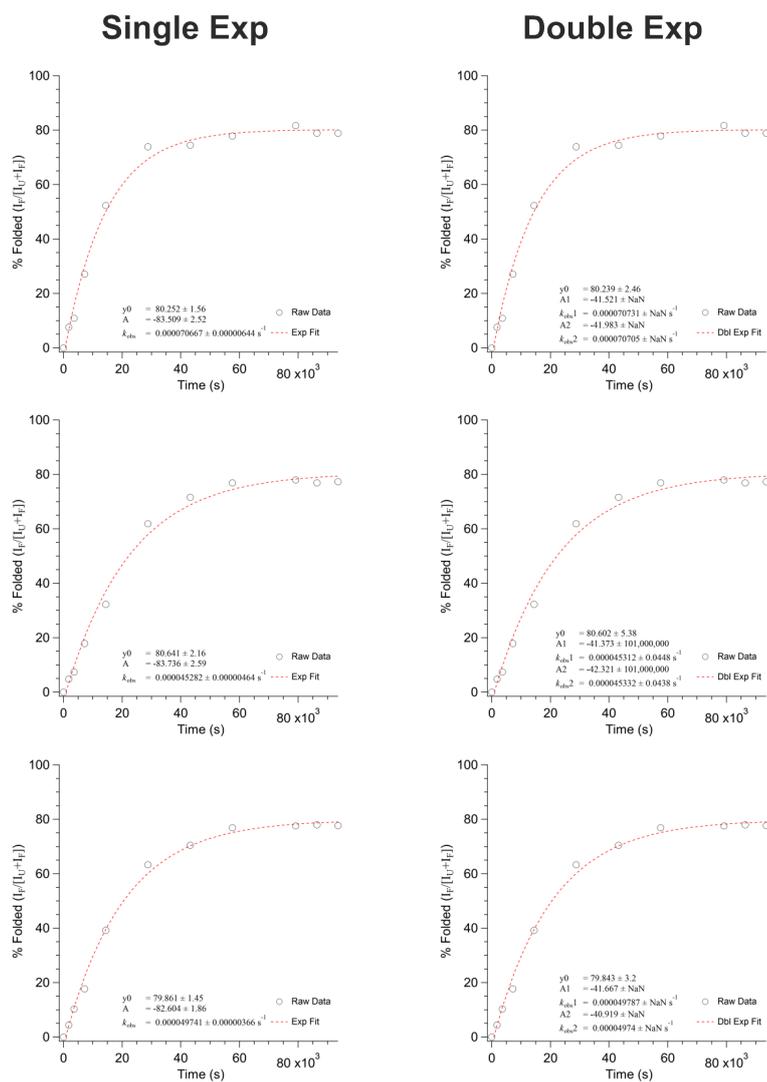


Fig. 26 Exponential fits to tOmpA folding into empty liposomes without SurA at 30 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

Empty 30°C

+SurA

Single Exp

Double Exp

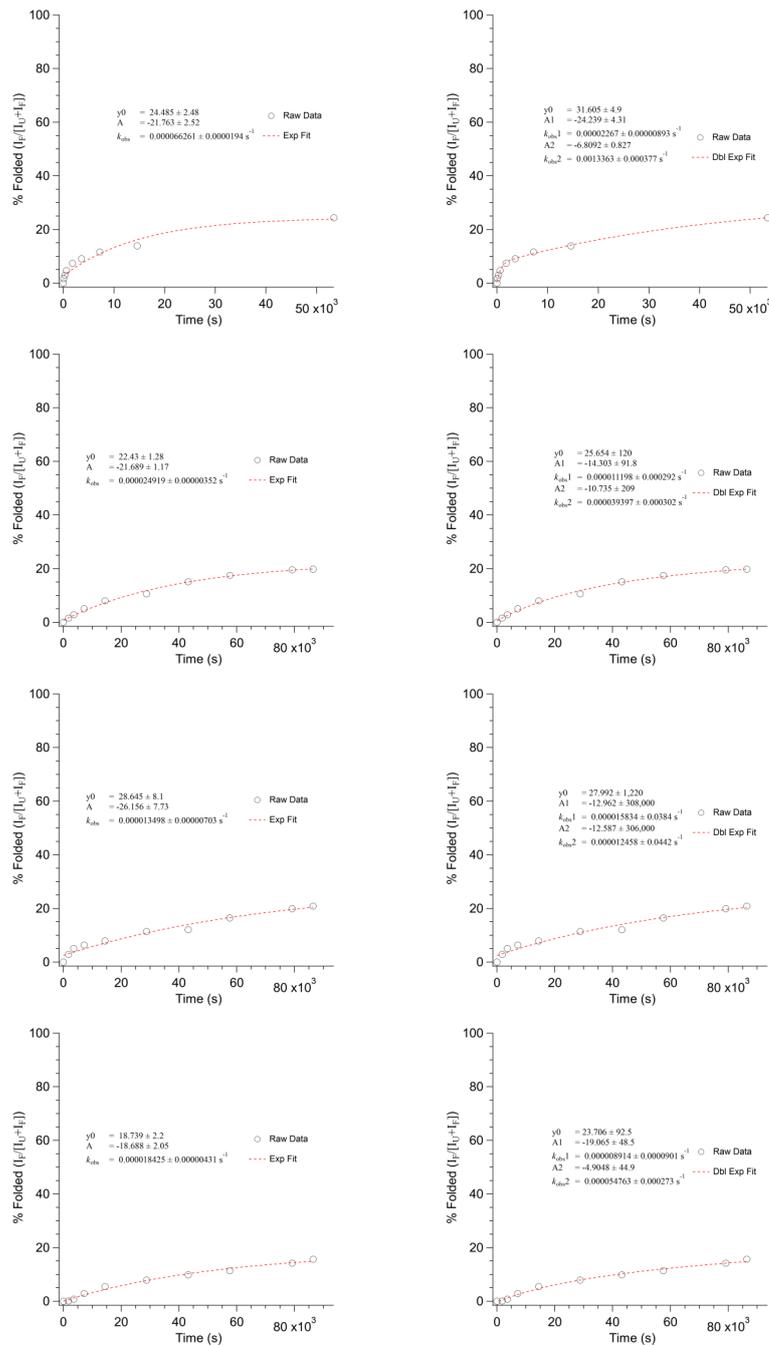


Fig. 27 Exponential fits to tOmpA folding into empty liposomes with SurA at 30 °C with the calculated parameters are shown here. Densitometry was performed on each gel shown in Appendix C to get the amount of folded and unfolded monomer. From this the percent folded was calculated as the intensity of the folded band over the sum of intensities for folded and unfolded. This was plotted against time and fitted to a single or double exponential equation.

Appendix E

Double exponential fits

Table 5 Fitting parameters from double exponential fits. See Materials & Methods for details of fitting methods and derivation of t_{50} values. t_{50} = time taken to reach 50 % folded. k_{obs} = observed rate constant. A = amplitude. σ_M = standard error of the mean.

Sample	BAM Catalysed																	
	t_{50} (s)	$k_{obs,1}$ (s ⁻¹)	$k_{obs,2}$ (s ⁻¹)	$k_{obs,2}$ error (s ⁻¹)	A1	A1 error	A2	A2 error	Y_0	Y_0 error	A1 (%)	A2 (%)	Average t_{50}	Average $k_{obs,1}$	Average $k_{obs,2}$	Average Y_0	Y_0 σ_M	
20°C no SurA Repeat 1	40.4	0.00443	0.00124	0.03422	0.00911	-37.5	8.7	-48.7	9.0	87.1	1.3	43.5	56.5					
20°C no SurA Repeat 2	41.5	0.00295	0.00116	0.02350	0.00295	-21.5	6.2	-67.9	6.3	90.3	1.0	24.1	75.9					
20°C no SurA Repeat 3	40.7	0.00318	0.00096	0.02553	0.00331	-25.8	5.7	-61.5	5.8	87.8	0.9	29.6	70.4		0.00352	0.02775	88.4	1.0
24°C no SurA Repeat 1	13.9	0.00476	0.00155	0.06340	0.00411	-14.9	1.9	-73.2	2.1	88.1	0.8	16.9	83.1					
24°C no SurA Repeat 2	12.0	0.00695	0.00195	0.08589	0.01240	-28.1	3.6	-66.1	4.2	89.3	1.1	25.9	74.1					
24°C no SurA Repeat 3	10.7	0.01313	0.00219	0.10780	0.01740	-58.6	4.1	-85.1	4.3	85.1	0.5	33.6	66.4					
24°C plus SurA Repeat 1	861	0.00232	0.00069	0.03187	0.00191	-11.3	1.3	-60.0	1.9	87.8	1.8	33.1	66.9					
24°C plus SurA Repeat 2	330	0.00612	0.00142	0.06259	0.00422	-21.9	2.9	-59.9	2.7	87.9	0.9	27.0	72.6					
24°C plus SurA Repeat 3	316.3	0.00008	0.00009	0.00362	0.00035	-24.5	13.7	-67.5	2.9	82.4	15.4	27.0	72.6					
30°C no SurA Repeat 1	383.8	0.00013	0.00006	0.00246	0.00018	-17.5	2.1	-70.9	2.6	88.3	2.3	19.8	80.2					
30°C no SurA Repeat 2	383.2	0.00014	0.00006	0.00228	0.00011	-13.6	1.6	-73.2	1.9	86.8	1.4	15.7	84.3		0.00017	0.00286	87.9	1.8
20°C plus SurA Repeat 1	36.8	0.00326	0.00076	0.03048	0.00399	-30.2	5.1	-66.0	5.4	96.6	1.0	31.4	68.6					
20°C plus SurA Repeat 2	34.0	0.00220	0.00107	0.02805	0.00320	-17.8	4.7	-75.7	5.1	94.6	1.3	19.0	81.0					
20°C plus SurA Repeat 3	39.8	0.00157	0.00055	0.02082	0.00221	-12.6	4.4	-75.5	4.7	89.5	1.4	14.3	85.7		0.00234	0.02645	93.6	2.1
24°C plus SurA Repeat 1	8.9	0.01164	0.00283	0.11885	0.02610	-23.3	4.8	-65.5	5.2	90.8	0.8	27.8	72.2					
24°C plus SurA Repeat 2	14.8	0.00909	0.00160	0.06814	0.00558	-23.4	3.3	-63.8	3.4	87.2	0.4	26.8	73.2					
24°C plus SurA Repeat 3	13.1	0.00832	0.00282	0.06954	0.00741	-17.3	4.2	-69.4	4.5	86.7	0.9	19.9	80.1		0.00968	0.08551	88.2	1.3
30°C plus SurA Repeat 1	87.6	0.00044	0.00036	0.00855	0.00044	-5.7	1.9	-88.9	2.3	95.0	0.9	6.0	94.0					
30°C plus SurA Repeat 2	80.3	0.00125	0.00165	0.00920	0.00063	-5.0	5.0	-91.2	4.9	96.2	0.5	5.2	94.8					
30°C plus SurA Repeat 3	77.5	0.00285	0.00061	0.01013	0.00180	-10.6	19.1	-79.8	18.9	89.8	0.7	11.7	88.3		0.00151	0.00929	93.6	1.9
Sample	BAMx Catalysed																	
	t_{50} (s)	$k_{obs,1}$ (s ⁻¹)	$k_{obs,2}$ (s ⁻¹)	$k_{obs,2}$ error (s ⁻¹)	A1	A1 error	A2	A2 error	Y_0	Y_0 error	A1 (%)	A2 (%)	Average t_{50}	Average $k_{obs,1}$	Average $k_{obs,2}$	Average Y_0	Y_0 σ_M	
20°C no SurA Repeat 1	1125.9	0.00033	0.00003	0.00414	0.00050	-36.3	0.8	-13.8	1.1	50.6	1.0	72.4	27.6					
20°C no SurA Repeat 2	2289.5	0.00016	0.00002	0.00265	0.00059	-44.3	1.9	-16.8	2.0	61.5	2.3	72.5	27.5					
20°C no SurA Repeat 3	2468.6	0.00012	0.00001	0.00224	0.00021	-45.7	1.5	-22.3	1.1	68.5	2.1	67.2	32.8		0.00020	0.00301	60.2	5.2
24°C no SurA Repeat 1	130.1	0.00276	0.00434	0.02857	0.00847	-45.8	3.1	-18.6	3.5	64.9	1.4	71.1	28.9					
24°C no SurA Repeat 2	152.4	0.00029	0.00093	0.00186	-31.3	6.8	-41.3	4.4	74.5	4.3	43.1	56.9						
24°C no SurA Repeat 3	130.1	0.00135	0.00045	0.01134	0.00287	-32.9	6.7	-37.3	6.7	72.3	2.2	46.8	53.2					
30°C no SurA Repeat 1	923.1	0.00005	0.00011	0.00011	0.00011	-22.0	33.3	-69.2	5.3	91.7	38.2	24.1	75.9		0.00160	0.01661	70.5	2.9
30°C no SurA Repeat 2	923.2	0.00011	0.00020	0.00087	0.00010	-9.3	5.1	-76.5	7.8	85.2	4.7	10.9	89.1					
30°C no SurA Repeat 3	702.3	0.00018	0.00011	0.00138	0.00020	-19.8	6.6	-66.6	7.6	85.4	2.2	22.9	77.1		0.00012	0.00112	87.4	2.1
20°C plus SurA Repeat 1	1246.7	0.00040	0.00026	0.00236	0.00062	-41.7	2.2	-9.4	2.7	51.8	0.8	81.5	18.5					
20°C plus SurA Repeat 2	1685.0	0.00008	0.00001	0.00008	0.00010	-29.8	8.9	-43.9	8.7	70.1	16.6	40.2	59.8					
20°C plus SurA Repeat 3	2151.0	0.00013	0.00001	0.00137	0.00014	-44.1	0.9	-24.8	1.5	68.8	1.4	62.0	36.0		0.00020	0.00152	63.7	6.0
24°C plus SurA Repeat 1	125.8	0.00157	0.00062	0.01507	0.00378	-35.1	4.4	-32.2	5.8	67.3	4.0	57.2	47.8					
24°C plus SurA Repeat 2	156.9	0.00083	0.00012	0.00914	0.00060	-30.1	1.2	-44.3	1.8	73.9	1.2	40.5	59.5					
24°C plus SurA Repeat 3	123.2	0.00112	0.00040	0.01165	0.00263	-31.1	5.0	-39.5	5.6	72.9	2.6	44.1	55.9		0.00117	0.01195	71.4	2.1
30°C plus SurA Repeat 1	1254.3	0.00016	0.00012	0.00085	0.00016	-30.8	9.6	-59.2	13.7	91.3	5.2	34.2	65.8					
30°C plus SurA Repeat 2	1291.1	0.00006	0.00035	0.00059	0.00008	-7.8	8.4	-85.3	11.0	94.4	16.9	8.4	91.6					
30°C plus SurA Repeat 3	1259.6	0.00015	0.00006	0.00071	0.00006	-20.5	5.9	-69.6	6.9	90.5	1.5	22.7	77.3		0.00012	0.00072	92.1	1.2
Sample	Uncatalysed																	
	t_{50} (s)	$k_{obs,1}$ (s ⁻¹)	$k_{obs,2}$ (s ⁻¹)	$k_{obs,2}$ error (s ⁻¹)	A1	A1 error	A2	A2 error	Y_0	Y_0 error	A1 (%)	A2 (%)	Average t_{50}	Average $k_{obs,1}$	Average $k_{obs,2}$	Average Y_0	Y_0 σ_M	
20°C no SurA Repeat 1	8418.2	0.00036	0.00006	0.00034	0.00043	-52.4	1.8	-28.4	3.0	80.4	1.7	64.9	35.1					
20°C no SurA Repeat 2	3947.6	0.00063	0.00006	0.00059	0.00077	-47.0	3.0	-34.0	3.3	81.0	0.6	58.0	42.0					
20°C no SurA Repeat 3	4054.3	0.00061	0.00010	0.00650	0.00119	-47.5	4.7	-34.5	5.2	81.9	1.0	27.9	42.1					
24°C no SurA Repeat 1	4813.9	0.00063	0.00007	0.00683	0.00148	-49.9	2.8	-34.5	1.9	82.3	3.2	50.6	49.4					
24°C no SurA Repeat 2	1592	0.00063	0.00094	0.01660	0.00259	-31.9	3.9	-36.5	4.2	69.1	2.6	46.6	53.4					
24°C no SurA Repeat 3	1771	0.00030	0.00056	0.00882	0.00075	-36.1	1.9	-53.6	2.1	90.3	1.7	40.2	59.8					
24°C plus SurA Repeat 1	167.9	0.00053	0.00064	0.01042	0.00071	-41.5	1.9	-47.7	2.1	86.7	0.9	44.7	55.3		0.00045	0.00966	82.1	4.6
30°C no SurA Repeat 1	10365.5	0.00071	Na/N	0.00071	Na/N	-41.5	Na/N	-42.0	Na/N	80.2	2.5	49.7	50.3					
30°C no SurA Repeat 2	16124.3	0.00045	0.04480	0.00045	0.04380	-41.4	101000000.0	-42.3	101000000.0	80.6	5.4	49.4	50.6					
30°C no SurA Repeat 3	14607.5	0.00050	Na/N	0.00050	Na/N	-41.7	Na/N	-40.9	Na/N	79.8	3.2	50.5	49.5		0.00016	0.00006	80.2	0.2
20°C plus SurA Repeat 1	8419.5	0.00070	0.00031	0.00107	0.00083	-43.4	840.0	-22.5	842.0	66.6	2.8	65.8	34.2					
20°C plus SurA Repeat 2	3005.8	0.00031	0.00011	0.000396	0.00039	-28.6	2.2	-59.4	3.6	88.3	2.6	32.5	67.5					
20°C plus SurA Repeat 3	3434.1	0.00025	0.00008	0.000357	0.00027	-30.0	1.9	-60.3	2.8	90.4	3.2	33.2	66.8					
24°C plus SurA Repeat 1	3240.4	0.00047	0.00020	0.00461	0.00020	-37.4	7.0	-48.8	8.6	86.2	2.5	43.4	56.6		0.00004	0.00033	82.9	5.5
24°C plus SurA Repeat 2	3780.9	0.00039	0.00007	0.000386	0.00007	-31.1	8.4	-63.9	11.7	81.5	1.4	40.6	63.4					
24°C plus SurA Repeat 3	4054.3	0.00061	0.00010	0.00650	0.00119	-47.5	4.7	-34.5	5.2	81.9	1.0	27.9	42.1					
24°C plus SurA Repeat 4	2452	0.000384	0.000109	0.0005615	0.000253	-33.4	3.7	-57.1	4.1	89.7	2.1	36.9	63.1		0.00034	0.00454	87.6	3.2
30°C plus SurA Repeat 1	16022.2	0.00027	0.00009	0.001336	0.000302	-24.2	4.3	-6.8	0.8	31.6	4.9	78.1	21.9					
30°C plus SurA Repeat 2	32922.6	0.00011	0.00022	0.000039	0.00037	-14.3	91.8	-10.7	209.0	25.7	120.0	57.1	42.9					
30°C plus SurA Repeat 3	42565.3	0.00016	0.03840	0.00012	0.03400	-13.0	308000.0	-12.6	306000.0	28.0	1220.0	50.7	49.3					
30°C plus SurA Repeat 4	55552.1	0.00009	0.00090	0.00055	0.00023	-19.1	48.5	-4.9	44.9	23.7	92.5	79.5	20.5		0.00002	0.00036	27.2	1.7

Appendix F

Single exponential fits

Table 6 Fitting parameters from single exponential fits. See Materials & Methods for details of fitting methods and derivation of t_{50} values. t_{50} = time taken to reach 50 % folded. k_{obs} = observed rate constant. A = amplitude.

BAM Catalysed									
Sample	t_{50} (s)	k_{obs} (s^{-1})	k_{obs} error (s^{-1})	A	A error	y_0	y_0 error	Average t_{50}	Average k_{obs}
20°C no SurA Repeat 1	44.1	0.01428	0.00194	-78.7	4.4	83.8	1.9		
20°C no SurA Repeat 2	42.4	0.01544	0.00148	-84.4	3.3	87.7	1.5		
20°C no SurA Repeat 3	43.7	0.01492	0.00166	-81.5	3.9	85.0	1.7	43.4	0.01488
24°C no SurA Repeat 1	14.8	0.04609	0.00546	-82.5	4.2	83.5	1.7		
24°C no SurA Repeat 2	15.0	0.04483	0.00668	-82.6	5.4	84.5	1.9		
24°C no SurA Repeat 3	14.5	0.04666	0.00602	-81.0	4.5	82.4	1.6		
24°C no SurA Repeat 4	16.9	0.04028	0.00371	-83.0	3.4	84.2	1.3	15.3	0.04447
30°C no SurA Repeat 1	346.3	0.00188	0.00028	-75.9	3.4	79.1	2.2		
30°C no SurA Repeat 2	253.4	0.00267	0.00038	-74.9	3.4	76.2	2.2		
30°C no SurA Repeat 3	365.5	0.00184	0.00021	-79.8	3.1	81.5	1.7		
30°C no SurA Repeat 4	370.1	0.00183	0.00016	-80.6	2.6	81.8	1.3	333.8	0.00206
20°C plus SurA Repeat 1	39.3	0.016645	0.00195	-88.9	5.0	92.3	1.9		
20°C plus SurA Repeat 2	32.9	0.019943	0.00209	-87.9	4.2	91.2	1.7		
20°C plus SurA Repeat 3	40.1	0.016425	0.00141	-83.7	3.2	86.7	1.4	37.5	0.01767
24°C plus SurA Repeat 1	12.1	0.056198	0.00752	-86.1	4.7	87.2	1.6		
24°C plus SurA Repeat 2	17.3	0.039076	0.00420	-82.9	4.0	84.5	1.5		
24°C plus SurA Repeat 3	14.7	0.046238	0.00462	-83.2	3.6	84.1	1.3	14.7	0.04717
30°C plus SurA Repeat 1	88.8	0.007666	0.00044	-92.5	2.2	93.6	0.8		
30°C plus SurA Repeat 2	80.7	0.008530	0.00027	-95.0	1.2	95.5	0.5		
30°C plus SurA Repeat 3	78.4	0.008833	0.00035	-89.2	1.4	89.3	0.6	82.6	0.00834
BamA Catalysed									
Sample	t_{50} (s)	k_{obs} (s^{-1})	k_{obs} error (s^{-1})	A	A error	y_0	y_0 error	Average t_{50}	Average k_{obs}
20°C no SurA Repeat 1	851.8	0.00074	0.00011	-41.7	2.2	77.5	2.1		
20°C no SurA Repeat 2	1813.4	0.00033	0.00006	-49.3	3.3	53.9	3.2		
20°C no SurA Repeat 3	1613.2	0.00037	0.00008	-49.9	3.8	55.1	3.7	1426.1	0.00048
24°C no SurA Repeat 1	121.4	0.00495	0.00065	-56.4	2.9	61.9	2.3		
24°C no SurA Repeat 2	138.3	0.00432	0.00084	-58.1	3.8	63.9	2.6		
24°C no SurA Repeat 3	136.0	0.00441	0.00067	-61.1	3.1	67.1	2.1	131.9	0.00456
30°C no SurA Repeat 1	733.3	0.00092	0.00006	-76.7	1.8	77.8	1.1		
30°C no SurA Repeat 2	892.9	0.00078	0.00004	-82.1	1.8	82.1	0.9		
30°C no SurA Repeat 3	684.5	0.00100	0.00009	-80.5	3.1	81.1	1.4	770.2	0.00090
20°C plus SurA Repeat 1	1139.5	0.000579	0.000037	-47.4	1.1	49.0	1.1		
20°C plus SurA Repeat 2	1291.5	0.000535	0.000057	-57.5	1.9	57.6	1.7		
20°C plus SurA Repeat 3	1725.1	0.000373	0.000059	-55.4	2.9	58.2	2.8	1385.4	0.00050
24°C plus SurA Repeat 1	105.2	0.005981	0.000901	-54.9	3.2	58.5	2.3		
24°C plus SurA Repeat 2	147.5	0.004385	0.000690	-62.5	3.3	65.4	2.3		
24°C plus SurA Repeat 3	125.9	0.004677	0.000796	-59.8	3.4	66.4	2.3	126.2	0.00501
30°C plus SurA Repeat 1	1176.7	0.000561	0.000039	-81.8	1.8	84.5	1.4		
30°C plus SurA Repeat 2	1239.9	0.000542	0.000026	-89.1	1.5	91.0	0.9		
30°C plus SurA Repeat 3	1236.6	0.000542	0.000030	-84.8	1.7	86.8	1.0	1217.7	0.00055
Uncatalysed									
Sample	t_{50} (s)	k_{obs} (s^{-1})	k_{obs} error (s^{-1})	A	A error	y_0	y_0 error	Average t_{50}	Average k_{obs}
20°C no SurA Repeat 1	7838.6	0.000085	0.000007	-72.9	2.3	75.2	1.8		
20°C no SurA Repeat 2	4778.0	0.000127	0.000017	-72.4	4.1	78.8	2.1		
20°C no SurA Repeat 3	4896.9	0.000124	0.000017	-73.0	4.3	79.5	2.2		
20°C no SurA Repeat 4	5232.5	0.000120	0.000013	-74.9	3.4	80.0	1.8	5686.5	0.00011
24°C no SurA Repeat 1	247.3	0.002216	0.000601	-51.6	4.5	59.7	3.4		
24°C no SurA Repeat 2	150.2	0.003840	0.000876	-54.2	4.5	60.8	2.8		
24°C no SurA Repeat 3	174.3	0.003458	0.000922	-71.7	6.6	78.5	3.8		
24°C no SurA Repeat 4	192.6	0.002975	0.000740	-70.3	5.9	79.3	3.6	191.1	0.00312
30°C no SurA Repeat 1	10371.6	0.000071	0.000006	-83.5	2.5	80.3	1.6		
30°C no SurA Repeat 2	16139.1	0.000045	0.000005	-83.7	2.6	80.6	2.2		
30°C no SurA Repeat 3	14614.1	0.000050	0.000004	-82.6	1.9	79.9	1.5	13708.2	0.00006
20°C plus SurA Repeat 1	8382.3	0.000081	0.000004	-65.7	1.1	66.4	0.9		
20°C plus SurA Repeat 2	3267.1	0.000194	0.000030	-76.9	5.1	81.7	2.3		
20°C plus SurA Repeat 3	3573.3	0.000178	0.000026	-77.5	4.9	82.1	2.3		
20°C plus SurA Repeat 4	3427.0	0.000190	0.000026	-78.5	4.5	82.0	2.0	4662.4	0.00016
24°C plus SurA Repeat 1	726.7	0.000939	0.000045	-72.9	1.2	73.7	1.1		
24°C plus SurA Repeat 2	213.6	0.002985	0.000624	-77.0	5.4	81.4	3.3		
24°C plus SurA Repeat 3	251.6	0.002567	0.000481	-77.8	4.8	81.6	3.1	397.3	0.00216
30°C plus SurA Repeat 1	8682.3	0.000066	0.000019	-21.8	2.5	24.5	2.5		
30°C plus SurA Repeat 2	26467.9	0.000025	0.000004	-21.7	1.2	22.4	1.3		
30°C plus SurA Repeat 3	44617.5	0.000013	0.000007	-26.2	7.7	28.6	8.1		
30°C plus SurA Repeat 4	37472.0	0.000018	0.000043	-18.7	2.1	18.7	2.2	29309.9	0.00003