



UNIVERSITY OF LEEDS

# **An *In Vivo* Platform for Identifying Protein Aggregation Inhibitors**

**Janet Catherine Saunders**

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

The University of Leeds

Astbury Centre for Structural Molecular Biology

**September 2014**

The candidate confirms that the submitted work is her 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 overleaf. The candidate confirms the 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.

© 2014 The University of Leeds and Janet Catherine Saunders



## Jointly Authored Publications

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.

### **Details of jointly authored publications and the contributions of other authors to these manuscripts:**

Chapter 4 contains work from the following manuscript submitted for publication:

Saunders, J.C.\*, Young, L.M.\*, Mahood, R.A., Revill, C.H., Foster, R.J., Tu, L.-H., Raleigh, D.P., Radford, S.E. and Ashcroft, A.E. Screening and classifying small molecule inhibitors of amyloid formation using ion mobility spectrometry-mass spectrometry. *Nature Chemistry*, In Press.

\*These authors contributed equally to this work.

In this work, I provided background information on the small molecules and their published interactions with human islet amyloid polypeptide (hIAPP) and amyloid  $\beta$  peptide residues 1-40 (A $\beta$ 40), and provided all the small molecules for *in vitro* analysis. I reviewed and selected the small molecules to be used for seeding the virtual screening library, and calculated LogP values of all 40 small molecules screened. I wrote the manuscript with L. M. Young, with help in manuscript preparation from the other authors. L. M. Young designed and carried out all ESI-IMS-MS experimental work, data interpretation and TEM analysis. R. A. Mahood purified the A $\beta$ 40 peptide and performed TEM of the sample. C. H. Revill and R. J. Foster performed the virtual screen. L. -H. Tu purified the hIAPP peptide. R. J. Foster, D. P. Raleigh, S. E. Radford and A. E. Ashcroft provided help with scientific discussions, data interpretation and manuscript preparation.



## Acknowledgements

It is with the greatest pleasure that I thank all the people who helped through my doctoral studies, be it through advice and support, or simply by distracting me with cake and gin.

My deepest gratitude goes to my supervisors, **Prof. Sheena Radford** and **Dr. David Brockwell** for their unwavering guidance, support and encouragement throughout my PhD. Their enthusiasm for science is infectious, which was indispensable during the particularly difficult periods (stupid bugs). This project would have fallen short many times without their ability to galvanise me into trying new things. I would also like to thank my industrial supervisors **Prof. Alastair Smith** and **Dr. Simon Webster**. The time and interest they have put into this project over the years has been invaluable.

Immense thanks are due to my incredible collaborator **Lydia Young** who carried out all of the mass spectrometry in this thesis. You enriched my work and help me complete countless tasks that I'd never have been able to finish alone. Collaborations have never been more fun.

I am very grateful I had the chance to work with all the wonderful members of the Radford and Brockwell laboratories; their knowledge and advice has been invaluable. In particular, I'd like to thank **Rachel Mahood** and **Dr. Sophia Goodchild**. It is a pleasure, not only to work with you beautiful ladies but to have you as such fun friends. I've loved every Costa, cocktail and sneaky lunch. A special thank you also goes to the past members of the group who helped me from the very beginning: **Dr. Claire Sarell** (literally the funniest person in the world), **Dr. Bethny Morrissey** (my liver destroyer), **Dr. Lindsay McMorran** (who I am saving the world with, one animal at a time) and **Dr. Alice Bartlett** (Super Scientist). **Dr. Clare Pashley**, you're still in the lab, but you are definitely in this list for the science (and not so science)-based advice. I set a little less stuff on fire in the lab thanks to you guys. You are all stars.

I would also like to express my gratitude to an unsung hero, our lab manager **Nasir Khan**. Thanks for keeping everything running so smoothly, and also the endless supply of doughnuts, biscuits and the spiciest vegetarian curries in the world.

My final, deepest gratitude goes to my parents **Gordon** and **Kath**, and my husband **Lewis**. I would not be the person I am today, or have ever reached this point in life without your endless, unwavering, patience, love and support. This is for you.



# Table of Contents

<b>JOINTLY AUTHORED PUBLICATIONS</b>	<b>1</b>
<b>ACKNOWLEDGEMENTS</b>	<b>3</b>
<b>TABLE OF CONTENTS</b>	<b>5</b>
<b>LIST OF FIGURES</b>	<b>13</b>
<b>LIST OF TABLES</b>	<b>17</b>
<b>LIST OF ABBREVIATIONS</b>	<b>19</b>
<b>LIST OF AMINO ACID ABBREVIATIONS</b>	<b>25</b>
<b>ABSTRACT</b>	<b>27</b>
<b>1 INTRODUCTION</b>	<b>29</b>
1.1 Principles of protein folding	29
1.2 Protein misfolding and aggregation	32
1.2.1 Aggregated species	33
1.2.2 Mechanisms of protein aggregation	36
1.3 Preventing protein aggregation	38
1.3.1 Intrinsic features that prevent aggregation	38
1.3.2 Cellular strategies for controlling aggregation	38
1.3.2.1 Aggregation prevention	39
1.3.2.1.1 The cytosol and the heat shock response	39
1.3.2.1.2 The ER and the unfolded protein response	41
1.3.2.2 Aggregation management	42
1.3.2.2.1 Sequestration of aggregates	42
1.3.2.2.2 Reversal of protein aggregation	44
1.3.2.2.3 Clearance of aggregates	44
1.4 Protein aggregation and disease	46

1.4.1	Loss of function-----	46
1.4.2	Gain of function -----	48
1.4.3	Functional amyloid -----	50
<b>1.5</b>	<b>Therapeutic approaches to protein aggregation -----</b>	<b>52</b>
1.5.1	Early diagnosis-----	52
1.5.2	Repairing cellular damage-----	53
1.5.3	Inhibiting production of monomeric form -----	53
1.5.4	Stabilising the native protein conformation-----	55
1.5.5	Increasing degradation of misfolded or aggregated protein-----	56
1.5.6	Accelerating fibril formation-----	57
1.5.7	Inhibiting or reversing protein oligomerisation-----	58
1.5.7.1	Small molecule inhibitors -----	59
<b>1.6</b>	<b>Screening for aggregation inhibitors -----</b>	<b>63</b>
1.6.1	<i>In vitro</i> screening techniques -----	63
1.6.2	<i>In vivo</i> screening techniques-----	65
1.6.2.1	Enzyme activity-based screening techniques -----	66
<b>1.7</b>	<b>The periplasmic system for identifying protein aggregation inhibitors-----</b>	<b>69</b>
1.7.1	Background to screen development-----	69
<b>1.8</b>	<b>Biopharmaceuticals -----</b>	<b>72</b>
1.8.1	History of biopharmaceuticals -----	72
1.8.2	Protein biopharmaceutical production -----	76
1.8.3	Biopharmaceutical stability and aggregation -----	77
1.8.4	Techniques employed to prevent biopharmaceutical aggregation -----	79
1.8.4.1	Protein engineering-----	79
1.8.4.2	Formulation screening -----	80
<b>1.9</b>	<b>An <i>in vivo</i> platform for assessing protein aggregation and identifying inhibitors -----</b>	<b>81</b>
<b>1.10</b>	<b>Aims of the current work -----</b>	<b>82</b>
<b>2</b>	<b>MATERIALS AND METHODS -----</b>	<b>83</b>
<b>2.1</b>	<b>Materials -----</b>	<b>83</b>
2.1.1	Technical equipment-----	83
2.1.2	Chemicals -----	85
2.1.3	Antibiotics, markers and dyes -----	88



2.1.4	Kits	88
2.1.5	Enzymes for molecular biology	89
2.1.6	Buffers	89
2.1.7	Media	90
<b>2.2</b>	<b>Molecular biology methods</b>	<b>91</b>
2.2.1	Bacterial strains	91
2.2.2	Polymerase chain reaction	91
2.2.3	Agarose gel electrophoresis	93
2.2.4	Restriction digest of plasmid DNA	94
2.2.5	Dephosphorylation of restriction endonuclease digests	94
2.2.6	Ligation of DNA	95
2.2.7	Transformation and cultivation of <i>E. coli</i> cells	95
2.2.8	Preparation of plasmids	95
2.2.9	Sequencing, stock maintenance and storage of DNA	96
2.2.10	Plasmids	96
2.2.11	Summary of plasmids created	97
<b>2.3</b>	<b>Protein expression and purification</b>	<b>98</b>
2.3.1	Purification of $\beta$ -lactamase constructs	98
2.3.1.1	Small-scale expression trials	98
2.3.1.2	Extraction of periplasmic fraction	98
2.3.1.3	Sodium dodecyl sulphate polyacrylamide gel electrophoresis	99
2.3.1.4	Large-scale expression of $\beta$ -lactamase constructs	100
2.3.1.5	Removal of sucrose from periplasmic preparation	100
2.3.1.6	Anion exchange chromatography	101
2.3.1.7	Size exclusion column chromatography	102
2.3.2	Purification of HEL4 and Dp47d antibody domains	103
2.3.2.1	Large-scale expression of antibody domains	103
2.3.2.2	Isolation of antibody domain inclusion bodies	103
2.3.2.3	Purification of HEL4 antibody domain	104
2.3.2.4	Purification of Dp47d antibody domain	105
2.3.2.5	Trichloroacetic acid precipitation	106
2.3.3	Purification of amyloid- $\beta_{40}$	106
2.3.4	Acquisition of IAPP and A $\beta$ 42	107
<b>2.4</b>	<b>Minimal inhibitory concentration (MIC) of antibiotic assays</b>	<b>108</b>
2.4.1	MIC assay in absence of small molecules	108
2.4.2	MIC assay in presence of increasing small molecule concentrations	108

2.4.3	MIC assay with bacterial lawns	109
2.4.4	MIC assay in presence of constant small molecule	110
2.4.5	MIC assay in presence of excipients	111
2.4.5.1	Sorbitol and glycerol	111
2.4.5.2	Arginine and glutamate	112
<b>2.5</b>	<b><i>In vitro</i> techniques</b>	<b>113</b>
2.5.1	Identification of fractions containing $\beta$ -lactamase during purification	113
2.5.2	Isoelectric point precipitation of WT $\beta$ -lactamase	113
2.5.3	Thermal denaturation of WT $\beta$ -lactamase	114
2.5.4	Freeze-thaw precipitation of WT $\beta$ -lactamase	114
2.5.5	Nephelometry	114
2.5.6	Transmission electron microscopy	115
2.5.7	Thioflavin T fluorometry	115
2.5.8	Mass spectrometry	115
2.5.8.1	Sample preparation	116
2.5.8.2	ESI-(IMS)-MS analysis	116
2.5.8.3	Analysis of ligand binding to monomeric hIAPP	116
2.5.9	Virtual screening for small molecules	117
<b>3</b>	<b>DEMONSTRATING ASSAY FEASIBILITY <i>IN VIVO</i></b>	<b>119</b>
<b>3.1</b>	<b>Objectives</b>	<b>119</b>
<b>3.2</b>	<b>The reporter protein <math>\beta</math>-lactamase</b>	<b>119</b>
<b>3.3</b>	<b>Wild-type <math>\beta</math>-lactamase is rendered inactive when aggregated</b>	<b>123</b>
3.3.1	Purification of wild-type $\beta$ -lactamase	123
3.3.1.1	Protein expression trials	123
3.3.1.2	WT $\beta$ -lactamase purification	125
3.3.2	Reduction in WT $\beta$ -lactamase enzyme activity correlates with aggregation	129
<b>3.4</b>	<b><math>\beta</math>-lactamase system allows identification of aggregation-prone proteins</b>	<b>131</b>
3.4.1	Selection of test proteins	131
3.4.1.1	Human and rat islet amyloid polypeptide	131
3.4.1.2	Amyloid $\beta$	134
3.4.2	Cloning test proteins into $\beta$ -lactamase linker	136
3.4.3	Split $\beta$ -lactamase system is amenable to a variety of culture conditions	136
3.4.4	Antibiotic resistance correlates with aggregation-propensity	139

3.5	Discussion	144
<b>4</b>	<b>DEVELOPING A PHENOTYPIC SMALL MOLECULE SCREEN</b>	<b>149</b>
4.1	Objectives	149
4.2	Assay development	149
4.2.1	Identification of positive and negative reference controls	149
4.2.2	Reducing the number of assay variables	155
4.2.3	Assay optimisation	159
4.2.3.1	Optimising choice of agar plate	159
4.2.3.2	Sample incubation with small molecule	160
4.2.3.3	Point of small molecule addition	161
4.2.3.4	Concentration of small molecule	162
4.2.3.5	Summary of optimised <i>in vivo</i> assay	163
4.2.4	Accounting for off-target effects	165
4.3	Assay validation	167
4.3.1	Selection of known small molecule inhibitors	167
4.3.2	<i>In vivo</i> screening of known inhibitors and non-inhibitors	170
4.3.3	Validating classification of hIAPP aggregation inhibitors	173
4.3.3.1	Transmission electron microscopy	173
4.3.3.2	Thioflavin T fluorescence	174
4.3.3.3	Mass spectrometry	175
4.3.4	<i>In vitro</i> analysis of $\beta$ la-hIAPP aggregation	183
4.3.4.1	Purification of $\beta$ la-linker and $\beta$ la-hIAPP	183
4.3.4.2	<i>In vitro</i> aggregation of $\beta$ la-linker and $\beta$ la-hIAPP	185
4.4	Investigating ambiguous inhibitors	188
4.5	Assay miniaturisation for high throughput screening	194
4.6	Investigating unknown inhibitors	196
4.7	Discussion	198
4.7.1	Assay development	198
4.7.2	The <i>in vivo</i> system as a screen for small molecules inhibitors of protein aggregation	202
4.7.3	Assay miniaturisation	203

<b>5</b>	<b>APPLYING THE <i>IN VIVO</i> ASSAY TO BIOPHARMACEUTICAL AGGREGATION</b>	<b>205</b>
<b>5.1</b>	<b>Objectives</b>	<b>205</b>
<b>5.2</b>	<b>Antibody biopharmaceuticals</b>	<b>205</b>
5.2.1	Monoclonal antibodies	206
5.2.2	Antibody fragments	207
5.2.3	Reducing the aggregation propensity of dAbs	210
<b>5.3</b>	<b>Using the <math>\beta</math>-lactamase tripartite fusion system to identify dAb aggregation propensity</b>	<b>212</b>
5.3.1	Cloning dAbs into $\beta$ -lactamase linker	213
5.3.2	Antibiotic resistance correlates with dAb aggregation-propensity	214
<b>5.4</b>	<b><i>In vivo</i> screening of excipients for anti-aggregation properties</b>	<b>216</b>
5.4.1	Dp47d purification in the absence of co-solvents	216
5.4.2	Selection of protein stabilisers	218
5.4.3	Selection of aggregation suppressors	220
5.4.4	Glycerol and sorbitol rescue bacterial growth in the presence of ampicillin	221
5.4.5	L-Arginine and L-glutamate elicit minimal rescue of bacterial growth	224
5.4.6	D-arginine is less toxic than L-arginine to bacterial growth	226
5.4.7	Addition of <i>in vivo</i> 'hit' excipients aid purification of Dp47d	228
<b>5.5</b>	<b>Discussion</b>	<b>229</b>
<b>6</b>	<b>CONCLUDING REMARKS AND FUTURE DIRECTIONS</b>	<b>235</b>
<b>7</b>	<b>APPENDICES</b>	<b>239</b>
<b>7.1</b>	<b>DNA and protein sequences of all peptides and constructs</b>	<b>239</b>
7.1.1	Wild-type TEM-1 $\beta$ -lactamase	239
7.1.2	$\beta$ la-linker <sub>SHORT</sub>	240
7.1.3	$\beta$ la-linker	241
7.1.4	$\beta$ la-hIAPP	242
7.1.5	$\beta$ la-rIAPP	243
7.1.6	$\beta$ la-A $\beta$ 40	244
7.1.7	$\beta$ la-A $\beta$ 42	245
7.1.8	$\beta$ la-HEL4	246
7.1.9	$\beta$ la-Dp47d	247

<b>7.2</b>	<b>Genes from which test proteins were cloned</b>	<b>248</b>
7.2.1	hIAPP	248
7.2.2	rIAPP	248
7.2.3	A $\beta$ 40 and A $\beta$ 42	249
<b>7.3</b>	<b>Plasmid maps of all plasmids obtained for this project</b>	<b>250</b>
7.3.1	pET28a- $\beta$ la-WT	250
7.3.2	pBM1- $\beta$ la-link-28	251
7.3.3	pBM1- $\beta$ la-link-64	252
7.3.4	pTXB1-hIAPP	253
7.3.5	pTXB1-rIAPP	254
7.3.6	pRSET-A $\beta$ 42	255
7.3.7	pEX-HEL4	256
7.3.8	pEX-Dp47d	257
<b>7.4</b>	<b>Small molecule compounds used in this study</b>	<b>258</b>
<b>REFERENCES</b>		<b>261</b>



## List of Figures

<b>FIGURE 1.1.</b> ENERGY LANDSCAPE SCHEME OF PROTEIN FOLDING AND AGGREGATION. ....	31
<b>FIGURE 1.2.</b> AN IDEALISED FUNNEL LANDSCAPE. ....	32
<b>FIGURE 1.3.</b> A SCHEMATIC REPRESENTATION OF THE PROTEIN AGGREGATION PATHWAY. ....	34
<b>FIGURE 1.4.</b> EXAMPLE OF AMYLOID FIBRIL MORPHOLOGY. ....	35
<b>FIGURE 1.5.</b> NUCLEATION-GROWTH MODEL FOR AMYLOID FORMATION. ....	36
<b>FIGURE 1.6.</b> CONTROLLING PROTEIN MISFOLDING IN THE CELL. ....	40
<b>FIGURE 1.7.</b> SPECIES-SPECIFIC PATHWAYS FOR AGGREGATE SEQUESTRATION. ....	43
<b>FIGURE 1.8.</b> ASYMMETRIC PARTITIONING OF PROTEIN AGGREGATES. ....	45
<b>FIGURE 1.9.</b> MECHANISMS OF PROTEIN AGGREGATION TOXICITY. ....	49
<b>FIGURE 1.10.</b> PRE-AGGREGATION THERAPEUTIC STRATEGIES AGAINST PROTEIN AGGREGATION. ....	54
<b>FIGURE 1.11.</b> POST-AGGREGATION THERAPEUTIC STRATEGIES AGAINST PROTEIN AGGREGATION. ....	57
<b>FIGURE 1.12.</b> $\pi$ -STACKING THEORY OF POLYPHENOL ANTI-AGGREGATION ACTIVITY. ....	61
<b>FIGURE 1.13.</b> CHEMICAL STRUCTURE OF THE POLYPHENOL (-)-EPIGALLOCATECHIN-3-GALLATE (EGCG). ....	62
<b>FIGURE 1.14.</b> CYTOPLASMIC-BASED SCREEN FOR INHIBITORS OF PROTEIN AGGREGATION. ....	66
<b>FIGURE 1.15.</b> CYTOPLASMIC ENZYME ACTIVITY SCREEN FOR INHIBITORS OF PROTEIN AGGREGATION. .	68
<b>FIGURE 1.16.</b> B-LACTAMASE TRIPARTITE FUSION SYSTEM TO MONITOR PROTEIN STABILITY. ....	70
<b>FIGURE 1.17.</b> PRINCIPLE OF B-LACTAMASE SCREEN FOR INHIBITORS OF PROTEIN AGGREGATION. ....	71
<b>FIGURE 1.18.</b> TOP NINE CATEGORIES OF BIOLOGIC DRUGS IN TERMS OF US SALES 2012 - 2013. ....	75
<b>FIGURE 1.19.</b> OVERVIEW OF THE PRODUCTION PROCESS FOR PROTEIN BIOPHARMACEUTICALS. ....	77
<b>FIGURE 3.1.</b> THE CORE STRUCTURES OF B-LACTAM ANTIBIOTICS. ....	119
<b>FIGURE 3.2.</b> BIOSYNTHESIS OF PEPTIDOGLYCAN AND ITS INHIBITION BY B-LACTAM ANTIBIOTICS. ....	120
<b>FIGURE 3.3.</b> RIBBON DIAGRAM OF TEM-1 B-LACTAMASE FROM <i>E. COLI</i> . ....	122
<b>FIGURE 3.4.</b> BLA-WT PROTEIN EXPRESSION UNDER PESSIMAL AND OPTIMAL CONDITIONS. ....	124
<b>FIGURE 3.5.</b> ANION EXCHANGE PURIFICATION OF WT B-LACTAMASE. ....	126
<b>FIGURE 3.6.</b> SCHEMATIC OF NITROCEFEN HYDROLYSIS BY B-LACTAMASE. ....	127
<b>FIGURE 3.7.</b> SIZE EXCLUSION CHROMATOGRAPHY OF WT B-LACTAMASE. ....	128
<b>FIGURE 3.8.</b> WT B-LACTAMASE ACTIVITY AFTER ISOELECTRIC POINT PRECIPITATION. ....	129
<b>FIGURE 3.9.</b> WT B-LACTAMASE ACTIVITY AFTER FREEZE-THAW-THERMAL DENATURATION CYCLING. .	130
<b>FIGURE 3.10.</b> SCHEMATIC OF THE <i>IN VIVO</i> SPLIT B-LACTAMASE ASSAY FOR PROTEIN AGGREGATION. .	132
<b>FIGURE 3.11.</b> THE AMINO ACID SEQUENCES OF HUMAN AND RAT IAPP. ....	133
<b>FIGURE 3.12.</b> THE NON-AMYLOIDOGENIC AND AMYLOIDOGENIC PATHWAYS OF APP PROCESSING. ....	134
<b>FIGURE 3.13.</b> THE AMINO ACID SEQUENCES OF AMYLOID B RESIDUES 1-40 AND 1-42. ....	135
<b>FIGURE 3.14.</b> COMPARISON OF MIC ASSAYS OVER TWO RANGES OF AMPICILLIN CONCENTRATION. ....	138
<b>FIGURE 3.15.</b> <i>IN VITRO</i> AGGREGATION CORRELATES WITH <i>IN VIVO</i> ANTIBIOTIC RESISTANCE. ....	141
<b>FIGURE 3.16.</b> <i>IN VITRO</i> AGGREGATION CORRELATES WITH <i>IN VIVO</i> ANTIBIOTIC RESISTANCE. ....	143
<b>FIGURE 4.1.</b> STRUCTURES OF (-)-EPIGALLOCATECHIN GALLATE (EGCG) (A) AND CURCUMIN (B). ....	150

<b>FIGURE 4.2. BACTERIAL GROWTH IN PRESENCE OF SMALL MOLECULE INHIBITORS OF AGGREGATION.</b>	151
<b>FIGURE 4.3. RELATIONSHIP BETWEEN PROTEIN EXPRESSION AND BACTERIAL GROWTH.</b>	153
<b>FIGURE 4.4. BACTERIAL GROWTH IN PRESENCE OF CURCUMIN WITH OPTIMISED CONDITIONS.</b>	154
<b>FIGURE 4.5. SCHEMATIC OF BACTERIAL LAWN-BASED ASSAY.</b>	155
<b>FIGURE 4.6. IDENTIFYING INHIBITORS OF AB42 AGGREGATION USING BACTERIAL LAWN ASSAY.</b>	157
<b>FIGURE 4.7. EXAMPLE OF SCORING DISCREPANCIES IN THE BACTERIAL LAWN ASSAY.</b>	158
<b>FIGURE 4.8. COMPARISON OF SCORING GROWTH WITHIN A 96-WELL AND A 48-WELL PLATE.</b>	160
<b>FIGURE 4.9. ASSAY SETUP FOR PRE-MIXING OF COMPOUND INTO THE AGAR.</b>	161
<b>FIGURE 4.10. SCHEMATIC OF OPTIMISED 48-WELL PLATE ASSAY.</b>	164
<b>FIGURE 4.11. CORRECTING FOR INTRINSIC EFFECTS OF SMALL MOLECULES ON BACTERIAL GROWTH.</b>	166
<b>FIGURE 4.12. SCHEMATIC OF ESI-MS SPECTRA FOR THE VARIOUS SMALL MOLECULE BINDING MODES.</b>	168
<b>FIGURE 4.13. ESI MASS SPECTRA OF SMALL MOLECULE/HIAPP INTERACTION.</b>	169
<b>FIGURE 4.14. GROWTH QUANTIFICATION OF BACTERIA IN THE PRESENCE OR ABSENCE OF A SMALL MOLECULE.</b>	170
<b>FIGURE 4.15. <i>IN VIVO</i> CLASSIFICATION OF SMALL MOLECULE INHIBITORS OF BLA-HIAPP AGGREGATION.</b>	171
<b>FIGURE 4.16. <i>IN VIVO</i> DOSE-RESPONSE RELATIONSHIP BETWEEN BACTERIAL GROWTH AND SMALL MOLECULE CONCENTRATION.</b>	172
<b>FIGURE 4.17. THT FLUORESCENCE INTENSITY OF HIAPP IN PRESENCE OR ABSENCE OF SMALL MOLECULES.</b>	174
<b>FIGURE 4.18. ESI-IMS-MS DEMONSTRATES THE MODE OF INHIBITION (SPECIFIC/COLLOIDAL/NON-SPECIFIC) OR LACK OF INHIBITION OF HIAPP AMYLOID FORMATION BY SMALL MOLECULES.</b>	177
<b>FIGURE 4.19. BLA-LINKER AND BLA-HIAPP PROTEIN EXPRESSION AND PURIFICATION.</b>	184
<b>FIGURE 4.20. STRUCTURES OF SMALL MOLECULES USED FOR <i>IN VITRO</i> ANALYSIS OF BLA-HIAPP AGGREGATION.</b>	185
<b>FIGURE 4.21. <i>IN VITRO</i> AGGREGATION OF BLA-HIAPP IN THE PRESENCE OF DIFFERENT SMALL MOLECULES.</b>	186
<b>FIGURE 4.22. <i>IN VITRO</i> AGGREGATION OF BLA-LINKER.</b>	187
<b>FIGURE 4.23. <i>IN VIVO</i> ANALYSIS OF SIX AMBIGUOUS INHIBITORS OF HIAPP AGGREGATION.</b>	190
<b>FIGURE 4.24. COMPARISON OF FALSE-POSITIVE AND FALSE-NEGATIVE RATES DURING ASSAY MINIATURISATION.</b>	195
<b>FIGURE 4.25. EXAMPLE OF DISPARITY IN COLONY SIZE PRODUCED BY DIFFERENT TRIPARTITE FUSION PROTEINS.</b>	199
<b>FIGURE 4.26. SUMMARY OF SMALL MOLECULE INHIBITOR OF PROTEIN AGGREGATION ASSAY OPTIMISATION. THE SELECTED CONDITIONS ARE HIGHLIGHTED IN GREEN.</b>	204
<b>FIGURE 5.1. STRUCTURE OF ANTIBODIES AND THEIR FRAGMENT-BASED THERAPEUTICS.</b>	207



<b>FIGURE 5.2.</b> STRUCTURE OF HEAVY CHAIN ANTIBODIES AND FRAGMENTS DERIVED FROM CAMELIDS.	208
<b>FIGURE 5.3.</b> SEQUENCE AND STRUCTURAL COMPARISONS OF THE V <sub>H</sub> DABS DP47D AND HEL4. ....	211
<b>FIGURE 5.4.</b> N- AND C-TERMINI DISTANCE BETWEEN MALTOSE-BINDING PROTEIN (MBP) AND HEL4. .	213
<b>FIGURE 5.5.</b> <i>IN VIVO</i> ANTIBIOTIC RESISTANCE CORRELATES WITH DAB AGGREGATION PROPENSITY....	215
<b>FIGURE 5.6.</b> PURIFICATION OF THE DABS HEL4 AND DP47D. ....	217
<b>FIGURE 5.7.</b> SCHEMATIC REPRESENTATION OF PREFERENTIAL HYDRATION OF PROTEIN VERSUS PREFERENTIAL BINDING OF CO-SOLUTE TO PROTEIN. ....	219
<b>FIGURE 5.8.</b> STRUCTURE OF THE POLYOLS (A) GLYCEROL AND (B) SORBITOL. ....	219
<b>FIGURE 5.9.</b> STRUCTURE OF (A) L-ARGININE AND (B) L-GLUTAMATE AT NEUTRAL PH. ....	220
<b>FIGURE 5.10.</b> BACTERIAL GROWTH IN PRESENCE OF SORBITOL.....	222
<b>FIGURE 5.11.</b> BACTERIAL GROWTH IN PRESENCE OF GLYCEROL. ....	223
<b>FIGURE 5.12.</b> BACTERIAL GROWTH IN PRESENCE OF EQUIMOLAR L-ARG AND L-GLU. ....	225
<b>FIGURE 5.13.</b> STRUCTURE OF D-ARGININE AT NEUTRAL PH. ....	226
<b>FIGURE 5.14.</b> COMPARISON OF BACTERIAL GROWTH IN PRESENCE OF L-ARG OR D-ARG.....	227
<b>FIGURE 5.15.</b> REFOLDING OF DP47D USING THE EXCIPIENT SORBITOL.....	228



## List of Tables

<b>TABLE 1.1.</b> THE MAIN HUMAN DISEASES ASSOCIATED WITH AMYLOID OR AMYLOID-LIKE DEPOSITS. ....	47
<b>TABLE 1.2.</b> EXAMPLES OF THE DIVERSITY OF FUNCTIONAL AMYLOID.....	51
<b>TABLE 1.3.</b> SELECT EXAMPLES OF SMALL MOLECULE INHIBITORS OF PROTEIN AGGREGATION.....	60
<b>TABLE 1.4.</b> FUNCTIONAL CLASSIFICATION OF PROTEIN THERAPEUTICS.....	74
<b>TABLE 1.5</b> COMPARISON OF PRODUCTION SYSTEMS FOR RECOMBINANT BIOPHARMACEUTICALS. ....	76
<b>TABLE 2.1.</b> ANTIBIOTICS USED IN THIS STUDY. ....	88
<b>TABLE 2.2.</b> DNA AND PROTEIN MARKERS USED IN THIS STUDY. ....	88
<b>TABLE 2.3.</b> DYES USED IN THIS STUDY. ....	88
<b>TABLE 2.4.</b> KITS USED IN THIS STUDY.....	88
<b>TABLE 2.5.</b> ENZYMES FOR MOLECULAR BIOLOGY USED IN THIS STUDY.....	89
<b>TABLE 2.6.</b> BUFFERS USED IN THIS STUDY. ....	89
<b>TABLE 2.7.</b> MEDIA USED IN THIS STUDY.....	90
<b>TABLE 2.8.</b> OLIGONUCLEOTIDE PRIMERS USED IN THIS THESIS.....	92
<b>TABLE 2.9.</b> TEMPERATURE CYCLE FOR A TYPICAL PCR REACTION. ....	93
<b>TABLE 2.10.</b> PLASMIDS OBTAINED FOR THIS THESIS. ....	97
<b>TABLE 2.11.</b> PLASMIDS CREATED IN THIS THESIS.....	97
<b>TABLE 2.12.</b> COMPONENTS OF A TRIS-TRICINE BUFFERED SDS-PAGE GEL.....	99
<b>TABLE 2.13.</b> ÄKTA PRIME PROGRAM FOR B-LACTAMASE CONSTRUCTS AND HEL4 PURIFICATION. ....	101
<b>TABLE 2.14.</b> ÄKTA PRIME PROGRAM FOR B-LACTAMASE CONSTRUCTS AND HEL4 PURIFICATION. ....	102
<b>TABLE 2.15.</b> ÄKTA PRIME PROGRAMME PARAMETERS FOR PURIFICATION OF DP47D. ....	105
<b>TABLE 2.16.</b> PROTOCOL FOR TITRATION OF AMPICILLIN INTO AGAR STOCK. ....	110
<b>TABLE 2.17.</b> PROTOCOL FOR PREPARATION OF SORBITOL AND GLYCEROL AGAR PLATES. ....	112
<b>TABLE 2.18.</b> PROTOCOL FOR PREPARATION OF MIXED L-ARGININE AND L-GLUTAMATE AGAR PLATES.....	112
<b>TABLE 3.1.</b> PARAMETERS OPTIMISED FOR MIC ASSAY. ....	137
<b>TABLE 3.2.</b> COMPARISON OF B-LACTAMASE CONSTRUCT LINKER LENGTHS.....	139
<b>TABLE 4.1.</b> PARAMETERS OPTIMISED FOR SMALL MOLECULE SCREEN.....	163
<b>TABLE 4.2.</b> KNOWN INHIBITORS OF HIAPP AGGREGATION. ....	167
<b>TABLE 4.3.</b> <i>IN VIVO</i> CLASSIFICATION COMPARED TO <i>IN VITRO</i> HIAPP AGGREGATION IN THE ABSENCE OF SMALL MOLECULE, OR IN THE PRESENCE OF CURCUMIN OR ACID FUCHSIN. ....	178
<b>TABLE 4.4</b> <i>IN VIVO</i> CLASSIFICATION COMPARED TO <i>IN VITRO</i> HIAPP AGGREGATION IN THE PRESENCE OF FAST GREEN FCF, EGCG OR CAFFEIC ACID.....	179
<b>TABLE 4.5.</b> <i>IN VIVO</i> CLASSIFICATION COMPARED TO <i>IN VITRO</i> HIAPP AGGREGATION IN THE PRESENCE OF HEMIN, 1 <i>H</i> -B-SA OR AZURE A.....	180
<b>TABLE 4.6.</b> <i>IN VIVO</i> CLASSIFICATION COMPARED TO <i>IN VITRO</i> HIAPP AGGREGATION IN THE PRESENCE OF SILIBININ, TRAMIPROSATE OR THIABENZADOLE. ....	181

<b>TABLE 4.7.</b> <i>IN VIVO</i> CLASSIFICATION COMPARED TO <i>IN VITRO</i> HIAPP AGGREGATION IN THE PRESENCE OF BENZIMIDAZOLE, ASPIRIN OR ORANGE G. ....	182
<b>TABLE 4.8.</b> PUBLISHED AMBIGUOUS INHIBITORS OF HIAPP AGGREGATION. ....	189
<b>TABLE 4.9.</b> <i>IN VIVO</i> CLASSIFICATION COMPARED TO <i>IN VITRO</i> HIAPP AGGREGATION IN THE PRESENCE OF ACRIDINE ORANGE, PHENOL RED OR MYRICETIN. ....	192
<b>TABLE 4.10.</b> <i>IN VIVO</i> CLASSIFICATION COMPARED TO <i>IN VITRO</i> HIAPP AGGREGATION IN THE PRESENCE OF RESVERATROL, MORIN HYDRATE OR CONGO RED. ....	193
<b>TABLE 4.11.</b> <i>IN VIVO</i> AND <i>IN VITRO</i> ANALYSIS OF FOUR UNKNOWN COMPOUNDS FOR THEIR ABILITY TO BIND TO HIAPP, AND/OR INHIBIT HIAPP AGGREGATION. ....	197
<b>TABLE 5.1.</b> ADVANTAGES OF SINGLE-DOMAIN ANTIBODY FRAGMENTS COMPARED TO CONVENTIONAL ANTIBODY FRAGMENTS. ....	209
<b>TABLE 7.1.</b> MOLECULAR WEIGHT, LOGP VALUE AND STRUCTURE OF SMALL MOLECULES USED IN THIS THESIS. ....	258
<b>TABLE 7.2.</b> MOLECULAR WEIGHT, LOGP VALUE AND STRUCTURE OF SMALL MOLECULES USED IN THIS THESIS CONTINUED. ....	259
<b>TABLE 7.3.</b> MOLECULAR WEIGHT, LOGP VALUE AND STRUCTURE OF SMALL MOLECULES USED IN THIS THESIS CONTINUED. ....	260

## List of Abbreviations

Amino acids are abbreviated according to their standard three-letter or single-letter codes. Other abbreviations are as follows:

1 <i>H</i> -B-SA	1 <i>H</i> -benzimidazole-2-sulfonic acid
Å	Ångström
$\alpha$ -1AT	Alpha 1-antitrypsin
A $\beta$	Amyloid beta
A $\beta$ 40	Amyloid beta residues 1-40
A $\beta$ 42	Amyloid beta residues 1-42
AD	Alzheimer's disease
AFM	Atomic force microscopy
ALS	Amyotrophic lateral sclerosis
amp	Ampicillin
APP	Amyloid precursor protein
BACE1	Aspartyl protease $\beta$ -site APP cleaving enzyme
BBB	Blood-brain barrier
<i>bla</i>	$\beta$ -lactamase encoding gene
bla	$\beta$ -lactamase
$\beta$ la-A $\beta$ 40	$\beta$ -lactamase-A $\beta$ 40
$\beta$ la-A $\beta$ 42	$\beta$ -lactamase-A $\beta$ 42
$\beta$ la-Dp47d	$\beta$ -lactamase-Dp47d
$\beta$ la-HEL4	$\beta$ -lactamase-HEL4
$\beta$ la-hIAPP	$\beta$ -lactamase-hIAPP
$\beta$ la-linker	$\beta$ -lactamase containing a 64-residue GS-rich linker
$\beta$ la-linker <sub>SHORT</sub>	$\beta$ -lactamase containing a 28-residue GS-rich linker
$\beta$ la-rIAPP	$\beta$ -lactamase-rIAPP
bp	Base pair

BSA	Bovine serum albumin
CDR	Complementary determining region
CFTR	Cystic fibrosis transmembrane conductance regulator
CJD	Creutzfeldt-Jakob disease
Da, kDa	Dalton, kiloDalton
dAb	Single-domain antibody
$\Delta G^{\circ}_{N-U}$	Free energy of folding
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DTT	1,2-dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
EGCG	(-)-epigallocatechin 3-gallate
ELISA	Enzyme-linked immunosorbent assay
EM	Electron microscopy
ER	Endoplasmic reticulum
ESI-IMS-MS	Electrospray ionisation-ion mobility spectrometry-mass spectrometry
EtOH	Ethanol
Fab	Antigen binding fragment
FAP	Familial amyloid polyneuropathy
Fc	Crystallisable fragment
Fv	Variable fragment
GAGs	Glycosaminoglycans
GFP	Green fluorescent protein
GlcNAc	<i>N</i> -acetyl glucosamine
GRAS	Generally regarded as safe

GS linker	Glycine-serine linker
GuHCl	Guanidine hydrochloride
h	Hour
HCAb	Heavy chain antibody
HD	Huntington's disease
HDPs	Host defence peptides
HEL	Hen egg lysozyme
hIAPP	Human islet amyloid polypeptide
HIV	Human immunodeficiency virus
HSA	Human serum albumin
rIAPP	Rat islet amyloid polypeptide
IAPP	Islet amyloid polypeptide, <i>or</i> Amylin
Ig	Immunoglobulin
IPOD	Insoluble protein deposit
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
JUNQ	Juxtannuclear quality-control compartment
kan	Kanamycin
kb	Kilo bases
kJ	Kilojoule
LB	Luria-Bertani
mA	Milliamp
mAb	Monoclonal antibody
MAC <sub>GROWTH</sub>	Maximum ampicillin concentration at which growth occurs
MBP	Maltose-binding protein
MCS	Multiple Cloning Site
MCD <sub>GROWTH</sub>	Maximal cell dilution allowing growth
MIC	Minimal Inhibitory Concentration
$\mu$ g	Microgram

mg	Milligram
mL	Millilitre
mol	mole
MWCO	Molecular weight cut off
MTOC	Microtubule organising centre
MurNAc	<i>N</i> -acetylmuramic acid
NEB	New England Biolabs
NMR	Nuclear magnetic resonance
OD <sub>600</sub>	Optical density at 600 nm
<i>P. anserina</i>	<i>Podospora anserina</i>
PBP	Penicillin binding protein
PCR	Polymerase Chain Reaction
PG	Peptidoglycan
PMSF	Phenylmethylsulphonyl fluoride
POI	Protein of interest
PrP	Prion protein
psi	Pounds per square inch
PTMs	Post-translational modifications
RAGE	Receptor for advanced glycation end products
RNA	Ribonucleic acid
RNAi	RNA interference
rpm	Rotations per minute
<i>S. coelicolor</i>	<i>Streptomyces coelicolor</i>
<i>S. enteric</i>	<i>Salmonella enterica</i>
s, ms	Second, millisecond
ScFV	Single chain variable fragment
SDS	Sodium dodecyl sulphate



SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SMMs	Small molecule microarrays
SOC	Super Optimal broth with catabolite repression
SOD1	Superoxide dismutase 1
TAE	Tris-acetate-EDTA
TB	Terrific Broth
TCA	Trichloroacetic acid
TEM	Transmission electron microscopy
TEMED	Tetramethylethylenediamine
tet	Tetracycline
ThT	Thioflavin T
TNF $\alpha$	Tumour necrosis factor alpha
Tris	Tris (hydroxymethyl)-aminomethane
TTR	Transthyretin
U	Units
UPR	Unfolded protein response
UV	Ultraviolet
V <sub>H</sub>	Variable heavy domain
V <sub>H</sub> H	Variable heavy domain from a camelid
V <sub>L</sub>	Variable light chain
v/v	volume:volume ratio
w/v	weight:volume ratio
X-Gal	Bromo-chloro-indolyl-galactopyranoside
$\lambda$	Wavelength



## List of Amino Acid Abbreviations

A	Ala	alanine
C	Cys	cysteine
D	Asp	aspartate
E	Glu	glutamate
F	Phe	phenylalanine
G	Gly	glycine
H	His	histidine
I	Ile	isoleucine
K	Lys	lysine
L	Leu	leucine
M	Met	methionine
N	Asn	asparagine
P	Pro	proline
Q	Gln	glutamine
R	Arg	arginine
S	Ser	serine
T	Thr	threonine
V	Val	valine
W	Trp	tryptophan
Y	Tyr	tyrosine



## Abstract

Protein aggregation is the basis of a vast array of diseases and one of the most expensive problems to overcome during production of biopharmaceuticals. The tendency of proteins to aggregate ensures that demanding purification techniques are a pre-requisite to *in vitro* analysis of aggregation mechanisms and screening for aggregation inhibitors.

In this thesis, a powerful new system was developed to identify aggregation-prone sequences *in vivo*, and to screen for inhibitors of aggregation. The screen is based on a  $\beta$ -lactamase-tripartite fusion system, where the minimal inhibitory concentration of antibiotic, conferred by the  $\beta$ -lactamase enzyme, is used to evaluate the level of test protein aggregation.

Using this *in vivo* system, the aggregation propensity of the two disease-related proteins human islet amyloid polypeptide (hIAPP) and amyloid beta peptide (A $\beta$ ) was found to be significantly higher than non-aggregating controls. Importantly, this system provides a new approach to assess aggregation-propensity without the need for purified protein.

The system was used to screen small molecules for their aggregation-inhibiting properties against hIAPP. It was found that many results correlated well with the published literature on these molecules, but notably, a number did not. *In vitro* analysis of hIAPP aggregation in the presence of these molecules validated the results from the *in vivo* assay, refuting a number of published studies and confirming the power of the tripartite system for identifying aggregation inhibitors.

Finally, the system was used to differentiate between an aggregating and non-aggregating human V<sub>H</sub> antibody domain, to demonstrate the application of the screen to biopharmaceuticals. The *in vivo* system successfully identified the aggregating test protein, wherein the addition of excipients prevented its aggregation *in vivo* in a titratable manner.

Overall the work presented herein describes a novel, and experimentally simple, *in vivo* system that provides rapid and accurate analysis of protein aggregation and its inhibition.



*“As our life expectancy increases, the chances of getting a degenerative disease also increase... caused by something conceptually quite simple... incorrect protein folding”*

- E. Reynaud, 2010<sup>1</sup>

## **1 Introduction**

### **1.1 Principles of protein folding**

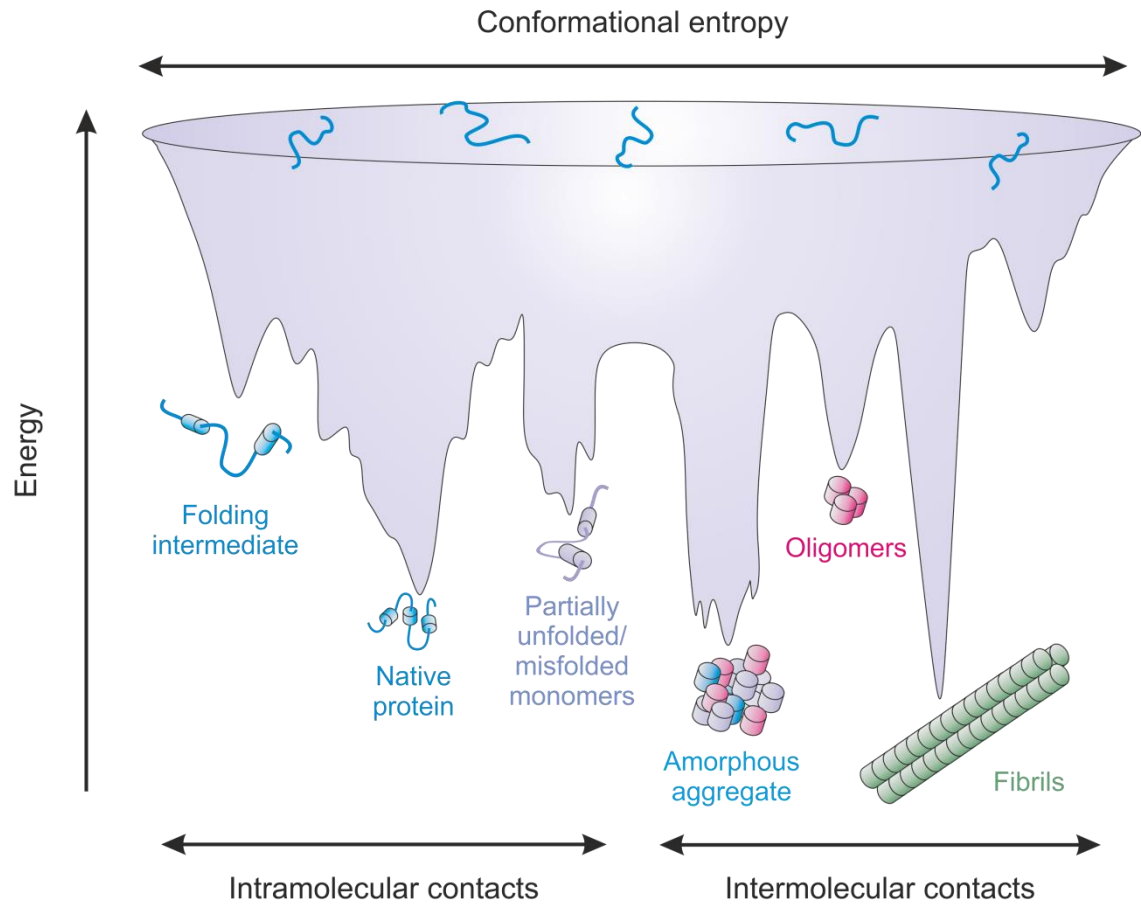
In order to function, most proteins must fold into a specific three-dimensional conformation termed the native fold. In the early 1960s, Anfinsen demonstrated that the folding of ribonuclease-A is reversible, and that its native structure is the conformation with the lowest free energy (and therefore highest stability)<sup>2</sup>. The conclusion from this work was that all the information required for a protein to reach its native three-dimensional conformation on a kinetically accessible timescale is contained in the polypeptide sequence. This notion seemed to give rise to a paradox: given the vast number of possible conformations available to the primary sequence of a protein, it is impossible for an exhaustive search for the global free energy minimum to occur on a biologically relevant timescale. Levinthal determined that this apparent paradox could be solved if protein folding is directed by specific pathways, guided by the rapid formation of local interactions, which determine the further folding of the polypeptide chain<sup>3</sup>.

The first proposed mechanism of protein folding was the nucleation-growth model<sup>4</sup>. This simplified the folding process by breaking it down into subprocesses that occur stepwise<sup>5</sup>. It postulated that the tertiary structure propagates rapidly from an initial nucleus of local secondary structure. However, this theory predicted the absence of folding intermediates and subsequently fell out of favour after the first proteins to be studied were shown to fold through kinetically observable intermediates<sup>6, 7</sup>. Two alternative models were suggested: the framework model<sup>8</sup>, and the related diffusion-collision model<sup>9, 10</sup>. These predicted that secondary structures would form first, and then diffuse until they collide and coalesce into the tertiary structure. Subsequently, the hydrophobic collapse model was proposed. This predicted that sequestration of hydrophobic side chains from the surrounding aqueous environment would produce a molten globule intermediate, constraining the volume in which the conformational search for the native state occurred<sup>11, 12</sup>. However, extensive studies on the folding of chymotrypsin inhibitor-2 (CI2) showed that this protein folds via simple two-state kinetics and that no intermediate accumulates<sup>13</sup>. Further investigation of CI2 by  $\phi$ -value analysis, which probes the presence of stabilising interactions in

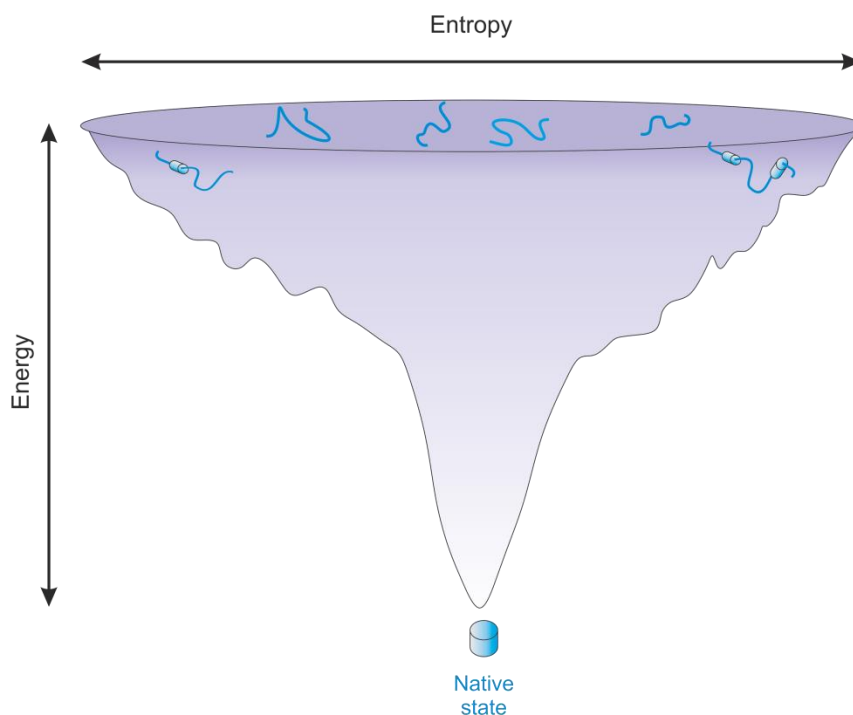
intermediate and transition states, showed that secondary and tertiary structures form in parallel<sup>14</sup>. These findings, coupled with the discovery of many proteins that fold with two-state kinetics<sup>15, 16</sup>, resulted in the resurgence of a nucleation-based mechanism of folding in the form of the nucleation-condensation model<sup>17</sup>. This model proposed that a small nucleus of weak secondary structure forms, stabilised by long range tertiary interactions. After this slow initiation step, the protein rapidly collapses around the nucleus to form the native state<sup>18</sup>. The boundaries that define all these models are somewhat artificial, and in 1993 the unified model of protein folding was proposed to combine the models in a continuum of mechanisms, depending on the relative strengths of secondary and tertiary interactions<sup>19</sup>. The framework and hydrophobic collapse models are the extremes of the nucleation-condensation model, with the relative stability of the secondary structure determining if the secondary and tertiary structural elements form in sequence or in parallel<sup>20</sup>.

Recently, energy landscapes and folding funnels have provided a new view of protein folding<sup>21, 22</sup>. They describe the structural ensemble acquired en route from a high energy unfolded state to the low energy native state. Energy landscapes can be plotted as the free energy of the polypeptide chain against the number of intra- and inter-molecular contacts (**Figure 1.1**). Folding funnels illustrate the relationship between the internal free energy of the protein and the conformational entropy. **Figure 1.2** shows an idealised folding funnel in which the internal free energy decreases concurrently with conformational entropy until the native state is reached at the energy minimum<sup>21</sup>. Realistic energy landscapes are likely to differ drastically from the idealised landscape. For proteins found to fold via a two-state transition, the energy landscape is relatively smooth. However the majority of proteins fold via the population of specific folding intermediates in order to reach their native conformation. These intermediate species give rise to a rough landscape, containing numerous high energy barriers and low energy kinetic traps. These 'on-pathway' species, however, are not the only conformations available to a folding polypeptide chain. Kinetically stable misfolded conformations can be populated, which require substantial reorganisation before the native state can be reached<sup>12</sup>.





**Figure 1.1.** Energy landscape scheme of protein folding and aggregation. The unfolded conformations funnel to the low energy native state via intramolecular contacts. Conversely, conformations can move towards amorphous aggregates, oligomers, or amyloid fibrils via intermolecular contacts. Figure redrawn and adapted from Hartl, 2001<sup>23</sup>.



**Figure 1.2.** An idealised funnel landscape. The vertical axis represents internal free energy. Conformational entropy is represented by the width of the funnel. Folding begins at the rim of the funnel, as the number of intramolecular contacts in the polypeptide increase, the internal free energy is lowered and conformational freedom reduced until the folded native state is reached. Figure redrawn and adapted from Bartlett, 2009<sup>24</sup>.

## 1.2 Protein misfolding and aggregation

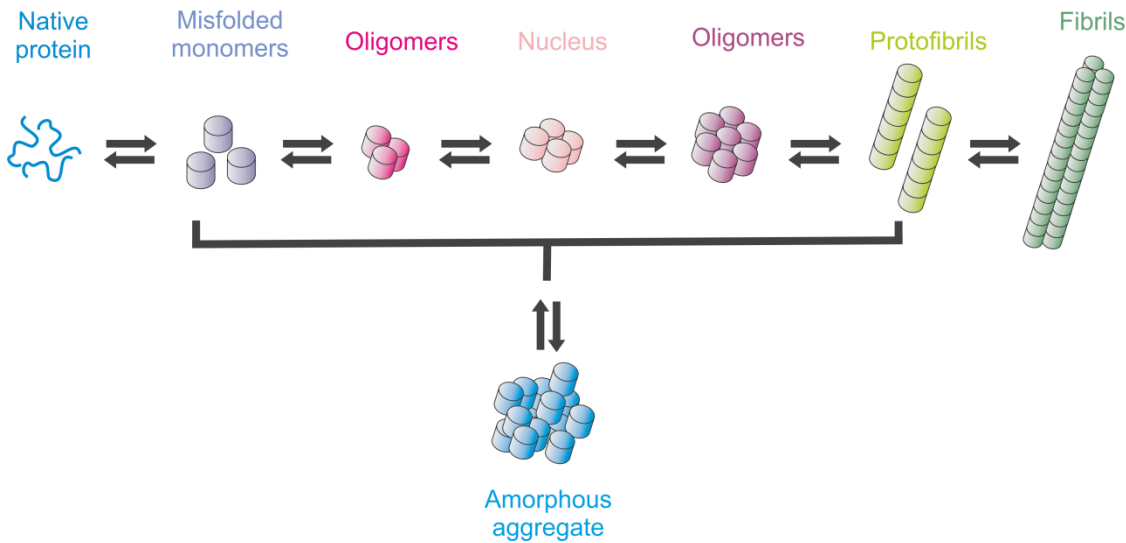
The native state of a protein is generally believed to correspond to the conformation that is most stable under physiological conditions<sup>25</sup>. However, native structures exhibit fluctuations around the minimal energy conformations, and may even undergo multiple localised unfolding reactions throughout the protein structure<sup>22</sup>. Indeed, many proteins adopt various other conformational states in addition to their native structure. These non-native, partially unfolded forms may represent functional or on-pathway species, but may also represent off-pathway states<sup>22</sup>. All species on the energy landscape are in constant flux, and rare or unfavourable species can also be produced as a consequence of changes in the thermodynamic stability, or interconversion kinetics, of the species. These changes can arise from destabilising or denaturing factors such as mutations, lack of ligands, improper proteolysis, and pH or temperature alterations<sup>26</sup>. Significantly, as the functional native state often only reflects a local free energy minimum at physiological concentrations and conditions, misfolding in many cases may actually lower the global free energy<sup>27-29</sup>. In

these unfolded, partially folded, or incorrectly folded species, the hydrophobic core and regions normally buried in the native state can become exposed to the solvent. These regions are prone to undesirable contacts with other molecules and, as folding must occur in the presence of 300-400 grams per litre of cytosolic protein and other macromolecules within the cell<sup>30</sup>, it is unsurprising that this can lead to aggregation and adverse interactions between non-native protein molecules.

### 1.2.1 Aggregated species

Due to the heterogeneous nature of protein aggregation, a diverse variety of species can form, ranging widely in both morphology and size. They can contain proteins bound together non-covalently; for example,  $\beta$ -lactoglobulin-A has been shown to form distinct non-covalent aggregates of tetramer, octamer and dodecamer<sup>31</sup>. Other forms of aggregates involve proteins covalently bound by intermolecular disulfide bonds. This occurs in aggregates of  $\beta$ -lactoglobulin-B and  $\kappa$ -casein-A during heat denaturation of milk<sup>32</sup>. The aggregation process is reversible in some cases, and irreversible in others. Reversible protein aggregation is commonly utilised in academic and industrial settings in the production of recombinant proteins in *E. coli* (see Section 1.8.2). Many proteins form inclusion bodies when over-expressed in bacteria, which can aid in purification as long as the protein can be resolubilised once extracted<sup>33</sup>. Protein aggregation in many neurodegenerative diseases, such as Alzheimer's and Parkinson's disease<sup>26, 34</sup>, is largely irreversible due to the very stable aggregate structures formed (see Section 1.4).

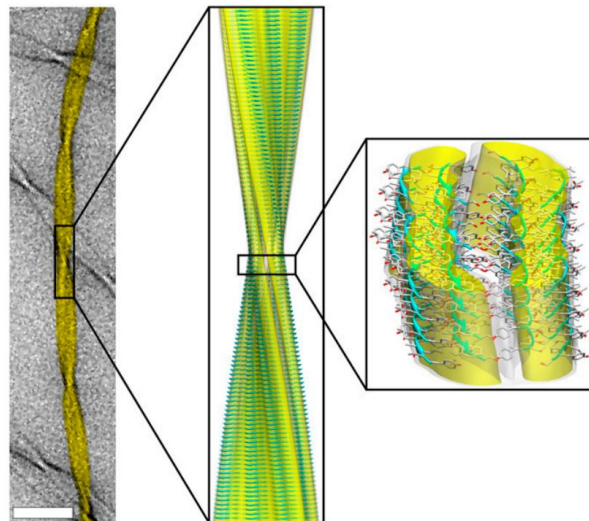
Protein aggregate composition and size varies greatly; from small soluble aggregates, to large oligomers, amorphous aggregates and amyloid fibrils (**Figure 1.3**)<sup>35</sup>. They also vary widely in their protein conformations. Some proteins, such as the human pancreatitis-associated protein, form fibrillar aggregates with a native-like conformation<sup>36</sup>, whereas other proteins undergo significant and sometimes dramatic alterations in conformation when they aggregate (islet amyloid polypeptide (IAPP) aggregation in type II diabetes mellitus, for example<sup>37</sup>). Protein aggregates and amyloid fibrils often contain a range of other components, including other proteins and carbohydrates; Glycosaminoglycans (GAGs), serum amyloid P component, proteoglycans, and apolipoprotein E are believed to help in the formation and stabilisation of fibrils<sup>38</sup>.



**Figure 1.3.** A schematic representation of the protein aggregation pathway. The starting reactant could be the native protein, or it could instead be a misfolded or partially unfolded monomer. Amorphous aggregates may be on or off the amyloid formation pathway.

It has been difficult to obtain high resolution structures of amyloid fibrils due to the insoluble and non-crystalline nature of these species. However, information from techniques such as chemical staining, hydrogen/deuterium exchange, electron microscopy (EM), atomic force microscopy, solid-state nuclear magnetic resonance (NMR), X-ray diffraction, circular dichroism, Fourier-transformed infrared spectroscopy, and electron paramagnetic resonance, has enabled the fibril structure to be deduced<sup>38</sup>. High resolution EM images show fibrils of different origin have similar morphology, consisting of 2-6 unbranched protofibrils associated laterally or twisted together to form fibrils with a 4-13 nm diameter (**Figure 1.4**)<sup>39</sup>. X-ray diffraction and solid-state NMR suggest a core cross- $\beta$ -sheet structure in which continuous  $\beta$ -sheets are formed, with the  $\beta$ -strands packed perpendicular to the fibril axis<sup>22</sup>. Furthermore, all amyloid fibrils stained with the diazo dye Congo red exhibit apple-green birefringence in cross-polarised light<sup>40</sup>. This characteristic, thought to occur through the linear arrangement of stacked dye molecules bound to the fibril<sup>40</sup>, is used as the diagnostic test for fibril presence. A shift in fluorescence after staining with thioflavin-T is also observed<sup>41</sup>.

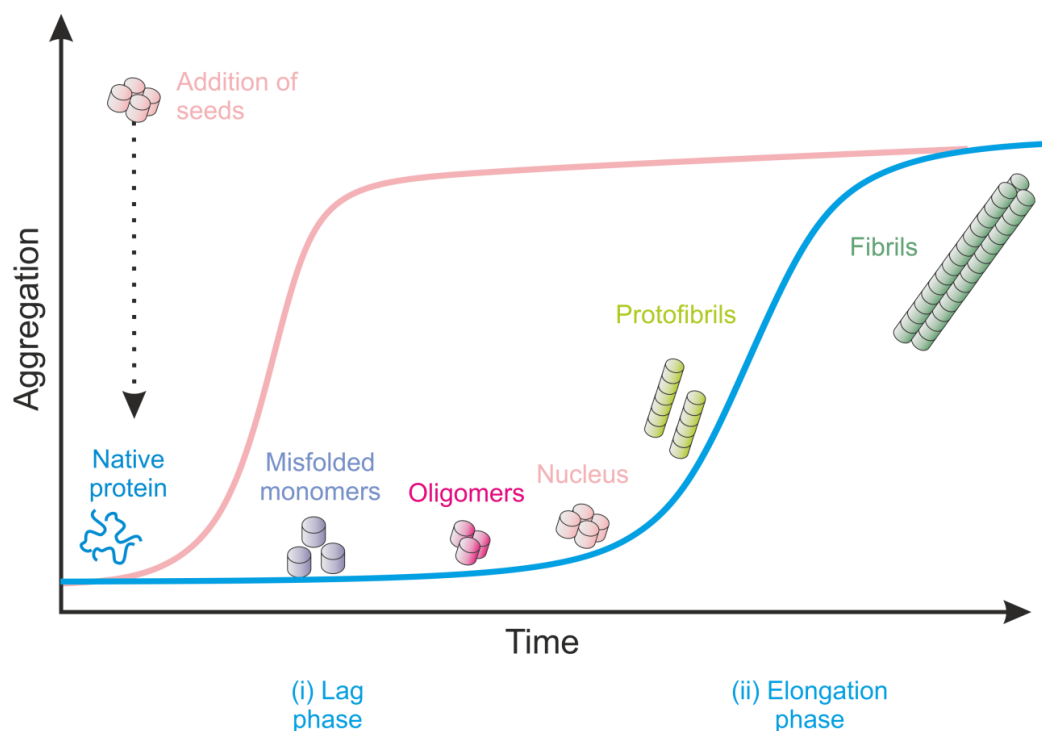
The native fold of a protein is established through the unique way in which its amino acid side chains interact with each other. Conversely, amyloid fibrils are formed via hydrogen bonding between the main chains of multiple peptides<sup>42</sup>. As the main chain of the polypeptide is preserved as a common feature of all natural proteins, it is possible that every protein possesses the ability to revert to this 'primordial' structure<sup>43</sup>. This hypothesis is supported by the fact that proteins able to form amyloid structures generally do not share any sequence similarities. Indeed, the secondary structural features of these proteins are quite distinct, ranging from all  $\alpha$ -helical through mixed  $\alpha$ -helical/ $\beta$ -sheet, to all  $\beta$ -sheet<sup>44</sup>. Furthermore, their native structures, sizes and biological functions vary widely, as do the conditions under which they form amyloid fibrils. To date, there are thirty proteins known to form amyloid fibrils in humans under physiological conditions, with a further six known to form either amyloid-containing, or non-amyloid inclusion bodies<sup>45</sup>. Moreover, it was shown that the src homology domain of bovine phosphatidylinositol 3-kinase (a protein with no link to protein-deposition diseases) has the ability to aggregate into fibrils that are structurally indistinguishable from those associated with human disease, when subjected to an acid pH<sup>46</sup>. There have since been many similar reports of the formation of amyloid-like fibrils<sup>25, 27, 47</sup>, leading to the conclusion that the amyloid fold is the universal global free-energy minimum of all polypeptide chains<sup>48</sup>.



**Figure 1.4.** Example of amyloid fibril morphology. A twisted fibril as visualised by transmission electron microscopy is shown on the left (scale bar = 50 nm). A cryo-EM reconstruction of an amyloid fibril, consisting of six 'protofilaments' winding around a hollow core, is shown in the centre. The constituent  $\beta$ -sheets shown in a ribbon representation, with the fibril density, indicated in yellow, is on the right. From Fitzpatrick *et al.*<sup>49</sup>.

## 1.2.2 Mechanisms of protein aggregation

Protein aggregation is generally proposed to occur via a nucleation-growth mechanism. Protein monomers are converted into fibrillar structures via a transiently populated aggregation nucleus, around which further deposition of monomers occurs (**Figure 1.5**)<sup>22</sup>. Nucleus formation is kinetically disfavoured, and hence the rate-limiting phase of the aggregation process. *In vitro* seeding experiments support this model by showing that addition of pre-formed fibril fragments greatly accelerates fibril formation (**Figure 1.5**)<sup>43</sup>. The addition of monomers to the existing aggregate is strongly favoured by the thermodynamic stability of the aggregate, and therefore the subsequent elongation of the fibril proceeds rapidly<sup>50</sup>. Amorphous aggregates and other prefibrillar species typically form faster than fibrils, via a nucleation-independent process, as there is no special conformational prerequisite; however, to date, there is no consensus as to whether these species are ‘on’ or ‘off’ the fibril formation pathway<sup>51</sup>.



**Figure 1.5.** Nucleation-growth model for amyloid formation. Amyloid formation proceeds via two phases: (i) the lag (or ‘nucleation’) phase and (ii) the elongation phase. During the lag phase, the native protein misfolds into an aggregation-competent conformation. This can act as a nucleus or self-associate to form an oligomeric nucleus. The elongation phase involves the rapid growth of the nuclei, via the addition of more monomers, into large fibrillar structures. The addition of pre-formed nuclei to the beginning of the reaction can overcome the rate-limiting nucleation phase, resulting in a significant reduction in the lag-time (pink line).

Much research has been directed towards determining the propensity of proteins to aggregate. Folding of a protein into a stable globular structure sequesters the hydrophobic, aggregation-prone residues, thereby dramatically decreasing the protein's overall propensity to aggregate<sup>52</sup>. Interestingly, natively unfolded and intrinsically unstructured proteins tend to have a lower propensity to aggregate than folded proteins (see Section 1.3.1)<sup>48</sup>. The reason for this appears to be due to the differences in mean net charge and mean hydrophobicity. Natively unstructured or unfolded proteins tend to have a higher net charge and lower hydrophobicity than folded proteins, thereby making aggregation thermodynamically less favourable<sup>53</sup>. Moreover, partially folded intermediates are more likely to aggregate than the unfolded species of the same protein<sup>25</sup>. This is because partial folding can bring together hydrophobic residues, creating an aggregation-prone hydrophobic surface. It has also been shown that helices with a high propensity for forming extended strand conformations are more likely to aggregate<sup>54</sup>.

Many other factors that destabilise native proteins (thereby increasing the likelihood of aggregation) have been found. Mutations can kinetically favour the formation of aggregates by increasing the mean hydrophobicity, increasing the propensity to form  $\beta$ -structure, or even by reducing the net charge of the polypeptide chain. For example, point mutations in the gene for lysozyme are associated with the aggregation of the protein in hereditary systemic amyloidoses<sup>55</sup>. Similarly, variants of the protein transthyretin which are significantly less stable than the native structure are involved in familial amyloidogenic neuropathy<sup>56</sup>. In other cases, full-length wild-type proteins can be encouraged to aggregate through changes in the environment or protein concentration. An example of this is the accumulation of the protein  $\beta_2$  microglobulin in patients undergoing long-term haemodialysis<sup>57</sup>. In this disorder, an up to 60-fold increase in monomer concentration leads to an increase in the concentration of a partially folded intermediate, with a high propensity for aggregation. This results in the deposition of the protein in amyloid plaques<sup>58</sup>. Furthermore, a high population of partially folded intermediates also risks exhausting the availability of cellular chaperones for nascent polypeptide chains, increasing aggregation propensity further<sup>59</sup>. In fact, impairment of any part of the protein processing machinery, from biogenesis to degradation, could have severely detrimental effects on the proteostasis network, and thereby increase the risk of aggregation (see Section 1.4).

## **1.3 Preventing protein aggregation**

### **1.3.1 Intrinsic features that prevent aggregation**

For life to have evolved to the level of complexity witnessed today, evolution must have eliminated protein sequences with a high intrinsic propensity to aggregate. For example, sequences of alternating polar and non-polar residues, as well as sequences of consecutive hydrophobic residues within the polypeptide chain, are strongly disfavoured<sup>48</sup>. However, polypeptide chains are far from optimised in their ability to resist aggregation. Proteins require flexibility to carry out their functions. Indeed, for some proteins to be functional they must be intrinsically disordered, or natively unfolded<sup>39</sup>. Evolution has developed a balance between a protein's propensity to aggregate and the dynamic requirements for biological function. All this must be encoded in the sequence, along with the ability to fold to the native state. Evolution will tend to produce sequences that improve an organism's function during its reproductive lifespan, as these confer a competitive advantage. However, due to the rapid advancements in science, technology and medicine, the changes in our society are fast exceeding the rate of evolution. A poignant example of this is the increase in sporadic Creutzfeldt–Jakob disease (CJD), believed to be due to the ingestion of aggregated tissue from cows. This has been linked to the recent outbreak of bovine spongiform encephalopathy in the UK that resulted from the highly unnatural process of feeding young cows with the infected remains of others<sup>42</sup>. Similarly 'unnatural' modern medical practices have also been linked to the amyloid diseases (e.g.  $\beta_2$ -microglobulin aggregation during long term haemodialysis)<sup>57</sup>.

### **1.3.2 Cellular strategies for controlling aggregation**

Within every kingdom, life has evolved intricate systems to prevent or control the aberrant misfolding and accumulation of cellular proteins<sup>30, 60-63</sup>. This indispensable and unceasing task is principally carried out by a class of proteins known as chaperones. Chaperones are ancient and universally conserved machines that assist in all stages of a protein's cellular life cycle; they aid in initial protein folding, assembly of multimeric complexes, inhibition of misfolding and aggregation, and final degradation of the protein into its constituent amino acids for recycling<sup>64, 65</sup>. The primary defence against protein aggregation is its prevention. However, with aggregation being inevitable in some cases, cells have evolved numerous 'management' systems to deal with the aggregate.



### 1.3.2.1 Aggregation prevention

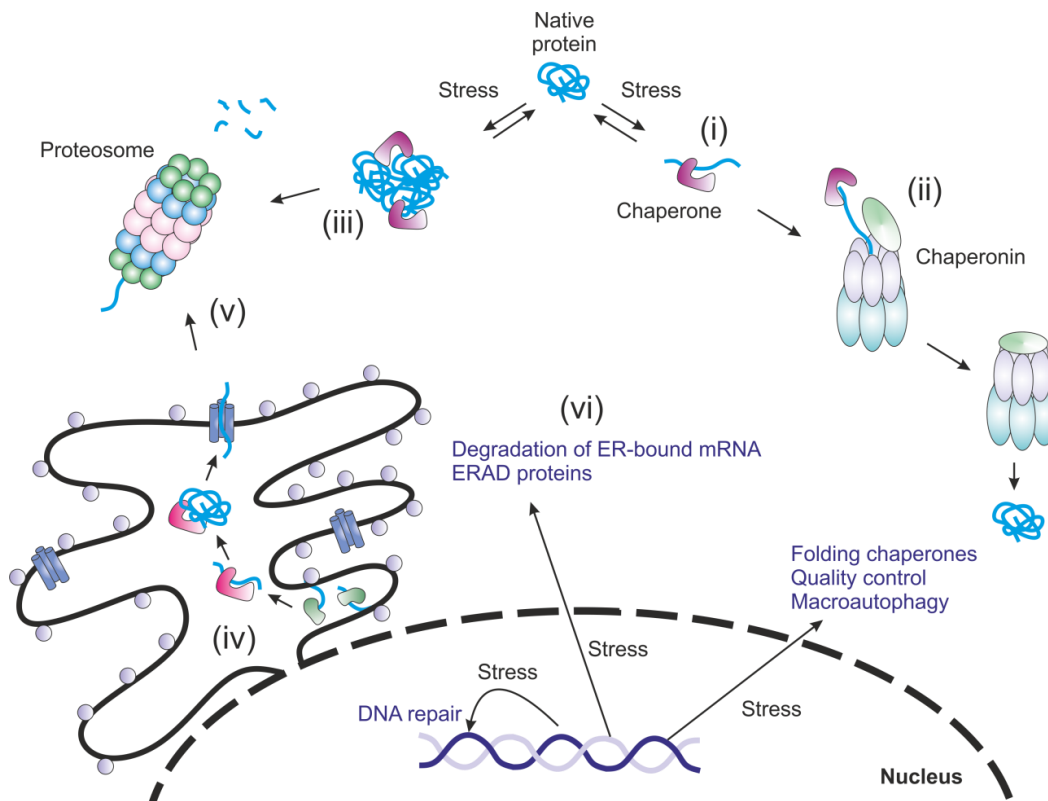
One of the most effective strategies used by the cell to control intracellular protein aggregation is to prevent the accumulation of misfolded conformers in the first place. The presence of misfolded proteins triggers a complex biological response; in the cytosol this is referred to as the 'heat shock response', and in the endoplasmic reticulum it is known as the 'unfolded protein response'<sup>47</sup>.

#### 1.3.2.1.1 The cytosol and the heat shock response

The heat shock response (HSR) is an ancient and universal mechanism to counter stress-induced aberrant interactions within the cell. This remarkably sensitive system can be triggered by a temperature increase of only a few degrees<sup>66</sup>. This is essential because cellular proteins require flexibility for their function; an increase in temperature, even a small one, can prompt unfolding<sup>66</sup>. Cells do not in fact recognise the temperature increase; it is believed that the trigger for the HSR is the appearance of the unfolding proteins. Despite what its name suggests, it is not just heat stress that triggers the HSR. Stresses including oxidative stress, heavy metals, or other toxic substances can activate this response mechanism<sup>67-69</sup>. From archaea to humans, studies have found that in order to tackle the cell-wide damage caused during times of stress, between 50 and 200 genes can be upregulated at any one time<sup>66</sup>. Expression of a wide variety of proteins is induced, however one class of proteins in particular is predominantly expressed: the heat shock proteins (Hsps). There are five major families of Hsps: Hsp100s, Hsp90s, Hsp70s, Hsp60s and small heat shock proteins (sHsps). Interestingly, many of these chaperones are also expressed under physiological conditions to aid the *de novo* folding of nascent polypeptide chains. As all cellular proteins are at risk of misfolding and aggregating, all Hsps have evolved to interact with a broad range of unfolded proteins. They recognise the hydrophobic regions, specific peptide sequences or structural elements that are exposed when a protein begins to unfold<sup>70-72</sup>. Through the energy-driven binding and release of the misfolding protein, they prevent unwanted interactions from occurring (**Figure 1.6 i**)<sup>66</sup>.

If the Hsp molecular chaperones alone are insufficient to promote correct folding of cytosolic proteins, the polypeptides can be transferred (via numerous cochaperone proteins) to a large, double ring complex known as a chaperonin (**Figure 1.6 ii**). Two groups of chaperonins exist: group I includes the seven-membered Hsp60s found in mitochondria, chloroplasts and bacteria (known as GroEL). Non-native protein chains (of

up to 60 kDa in GroEL<sup>73</sup>) can be bound in the central cavity of the chamber, closed in by the binding of another chaperone. This enables the chamber to close around the peptide chain and, in isolation from the rest of the cellular components, the protein can fold into its native state<sup>70, 74, 75</sup>. Interestingly, no group I chaperonin is found in the eukaryotic or archaeal cytosol. Instead, distantly related machines, denoted the group II chaperonins, carry out the same function<sup>76</sup>. These eight-membered rings, known as thermosomes in archaea, and TRiC or CCT in eukaryotes, have a similar architecture to the group I chaperonins, however the lid function is built in<sup>77</sup>.



**Figure 1.6.** Controlling protein misfolding in the cell. Protein unfolding or misfolding instigates the heat shock response and unfolded protein response. Chaperones bind the unfolding protein to prevent further misfolding or aggregation, and aid in the refolding to the native conformation (i). If protein refolding is initially unsuccessful, it is transferred to the chaperonin, wherein protein folding can occur in the protected interior environment (ii). If the protein is unable to refold, or aggregation occurs, the protein is transferred to the proteasome for degradation (iii). Within the ER, protein misfolding is prevented from occurring by ER chaperones (iv). They bind to the nascent polypeptide chain as it cotranslates into the ER lumen, where folding can occur. If folding is unsuccessful, the polypeptide chain is retrotranslocated into the cytosol where it is degraded by the proteasome (v). If the level of misfolded protein exceeds the capacity of the ER chaperones, the unfolded protein response (UPR) is initiated (vi).

Countless other chaperones and chaperone-associated proteins exist to aid in the folding, refolding and stabilisation of cellular proteins. Indeed, the quality control mechanism is so robust that as much as 30 % of newly synthesised proteins are degraded due to ineffective folding (**Figure 1.6 iii**)<sup>78</sup>.

### **1.3.2.1.2 The ER and the unfolded protein response**

Protein folding in the cell occurs either in the cytoplasm or within the secretory pathway<sup>79</sup>. Unlike in prokaryotes, where secretory proteins are secreted directly across the plasma membrane<sup>80</sup>, in eukaryotes transmembrane and secreted proteins are folded and assembled in the endoplasmic reticulum (ER). The ER provides a cellular 'checkpoint' for secretion, preventing export of proteins that are misfolded or modified incorrectly. The sheltered environment of the ER helps to spatially limit conformational freedom of the polypeptide chain, thereby reducing the risk of misfolding. Polypeptide chains are translocated into the lumen of the ER cotranslationally, where folding into the native structure (and assembly into multisubunit complexes, if required) is aided by ER chaperones (**Figure 1.6 iv**). The chaperones surround the translocon, thereby preventing the nascent polypeptide chain from interacting with neighbouring identical chains. This is vitally important in preventing protein aggregation, as the chance of aggregation is highest between partially folded proteins with identical sequences, often due to a process known as domain or strand swapping<sup>80-82</sup>. Proteins that fail to attain their native structure within the ER are retrotranslocated back into the cytoplasm for degradation by the proteasome (**Figure 1.6 v**). When the flux of unfolded proteins entering the ER surpasses the capacity of the folding machinery, a network of intracellular signalling pathways is activated, collectively called the unfolded protein response (UPR). The UPR transduces information to the cytosol and nucleus to upregulate proteins involved in folding, quality control, trafficking, ER-associated degradation and macroautophagy (**Figure 1.6 vi**). In addition, a widespread transcriptional program decreases protein translation and initiates the degradation of select ER-bound mRNA, thereby reducing the load of unfolded proteins at the ER<sup>83</sup>.

### 1.3.2.2 Aggregation management

The cellular quality control system adapts to the presence of misfolded proteins and works to clear them rapidly and efficiently from the cellular environment. However, when the generation of misfolded proteins exceeds the refolding and degradative capacity of the cell, protein aggregates can accumulate. It was initially believed that aggregation was the result of an uncontrolled, dead-end pathway; however mounting evidence suggests that it is, in fact, part of an organised response to imbalanced protein homeostasis<sup>84</sup>. It has been found, across all species, that protein aggregates are sequestered and deposited at specific cellular sites. It is believed that this protects the cellular environment from potentially dangerous species, such as the toxic soluble oligomers formed during aggregation (discussed in detail in Section 1.4.2). Indeed, it has been suggested that the formation of amyloid fibrils has a protective function through the sequestration of these toxic oligomers<sup>85-89</sup>. This intracellular organisation of aggregates may also facilitate their removal from the cell, either through degradation<sup>90, 91</sup>, or asymmetric distribution into a single daughter cell during cell division<sup>92-96</sup>.

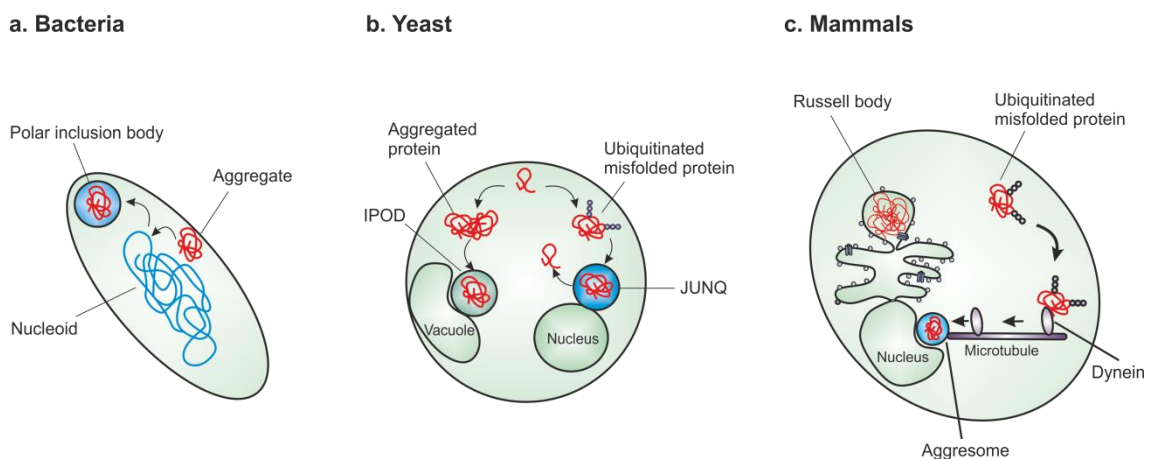
#### 1.3.2.2.1 Sequestration of aggregates

The specific localisation site of sequestered protein aggregates differs between organisms, and depends on the particular aggregating protein, the cellular compartment and even the particular stress conditions that have caused the aggregation. In bacteria, for example, misfolded proteins accumulate in inclusion bodies, particularly during heat or oxidative stress<sup>97, 98</sup>. Typically, one or two inclusion bodies form at the cell poles, and, for reasons that are still unclear, they usually localise to the old cell pole (**Figure 1.7a**)<sup>92, 94</sup>. The underlying mechanisms of how this localisation occurs remains controversial, with some studies suggesting it is an energy-driven process<sup>99</sup>, whereas others indicate that a passive mechanism of simple nucleoid occlusion is sufficient (**Figure 1.7a**)<sup>100</sup>.

Unlike bacteria, eukaryotic cells possess distinct protein quality-control compartments. For example, in yeast cells there are two recently identified specialised quality-control compartments: the juxtannuclear quality-control compartment (JUNQ) and the insoluble protein deposit (IPOD)<sup>91</sup>. The JUNQ is adjacent to the nuclear membrane, and transiently accumulates ubiquitinated misfolded proteins (**Figure 1.7b**). Substrates at the JUNQ are still mobile, and can diffuse back into the cytoplasm if they have successfully disaggregated and refolded to a native conformation. Conversely, the IPOD, located

adjacent to the vacuole, sequesters and harbours terminally misfolded and aggregated proteins, such as amyloidogenic proteins and yeast prion proteins (**Figure 1.7b**)<sup>91</sup>. The fate of the aggregated proteins in the IPOD compartment, however, remains unclear<sup>84</sup>.

Within mammalian cells, specialised inclusion bodies form in the ER and cytosol, known respectively as Russell bodies and aggresomes. Aggresome formation is initiated by the accumulation of smaller aggregates within the cell. These move in a dynein-based manner along the microtubule cytoskeleton to an indentation in the nuclear envelope called the microtubule organising centre (MTOC) (**Figure 1.7c**)<sup>101</sup>. Evidence suggests this mechanism prompts increased chaperone expression to counter the accumulation of aggregated species<sup>102</sup>. Similarly, Russell bodies are believed to form in order to sequester potentially toxic aggregating proteins in the ER, which might otherwise interfere with the secretory pathway (**Figure 1.7c**)<sup>103</sup>.



**Figure 1.7.** Species-specific pathways for aggregate sequestration. **(a)** In bacteria, aggregates accumulate into inclusion bodies, and, by either an energy-driven process or nucleoid occlusion, they are localised to the cell pole. **(b)** Yeast cells possess distinct compartments for the sequestration of misfolded or aggregated proteins. The juxtannuclear quality-control compartment (JUNQ) accumulates ubiquitinated misfolded proteins, which can diffuse back into the cytosol if refolding is successful. Aggregated proteins are sequestered in the insoluble protein deposit (IPOD), however their fate remains unclear. **(c)** In mammal cells, specialised inclusions of aggregated protein form in the cytosol and ER, known as aggresomes and Russell bodies, respectively. Aggresome formation requires the ubiquitinated protein aggregates to be transported along the microtubule to an indentation in the nuclear envelope called the microtubule organising centre. Figure redrawn from Tyedmers, 2010<sup>84</sup>.

### 1.3.2.2.2 Reversal of protein aggregation

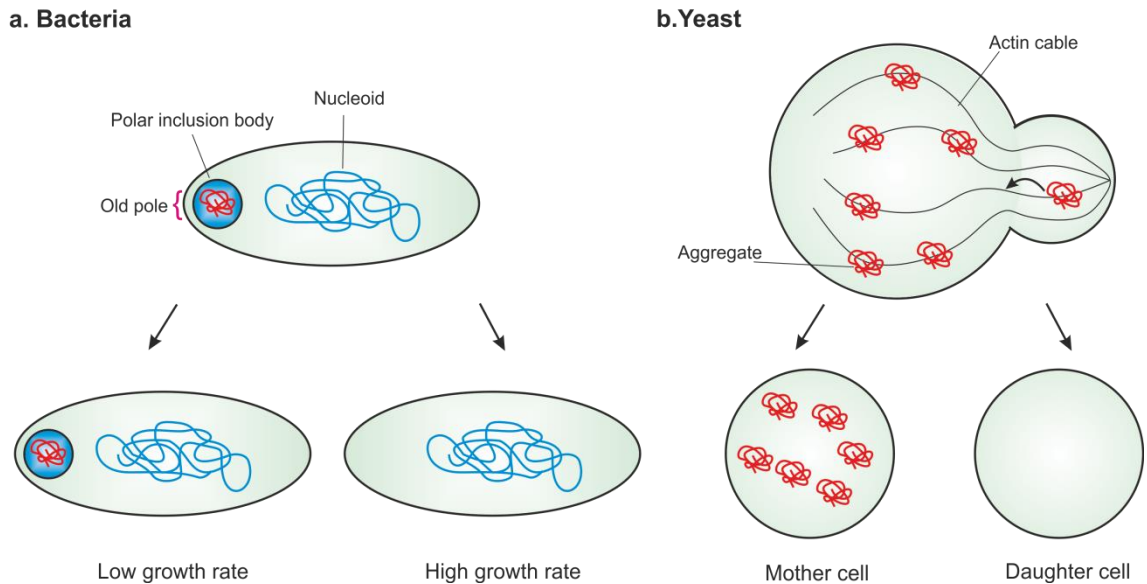
Disaggregation and refolding of protein aggregates has been observed in a diverse range of cells. This process is carried out by a bi-chaperone system, and has been demonstrated in the cytosol of bacteria, plants and several unicellular eukaryotes, and in mitochondria and chloroplasts. The aggregate is bound by Hsp70 (or its corresponding orthologue), which restricts the access of proteases to the aggregate. This enables the aggregate to be transferred to the second chaperone in this bichaperone system, the hexameric Hsp104 (or ClpB)<sup>84, 104, 105</sup>. The precise mechanisms by which this occurs are not fully understood, however it is believed that the protein aggregate is in some way remodelled to expose an area that is recognised by Hsp104 or ClpB<sup>84</sup>. One-by-one, the misfolded proteins are extracted from the aggregate and threaded through the pore of the chaperone, enabling refolding of the protein<sup>106, 107</sup>. Hsp104 and ClpB homologues have only been found in the mitochondria and chloroplasts of higher eukaryotes. Nevertheless, protein aggregates can be disassembled in the cytosol of these species<sup>108, 109</sup>, although the characteristics of this activity are not yet clear.

### 1.3.2.2.3 Clearance of aggregates

The sequestration, and when possible the disaggregation, of protein aggregates has a clear protective role within the cell. Despite these actions, the presence of protein aggregates is inevitably damaging. Consequently, degradation and removal are alternative routes for the elimination of the protein aggregate. In eukaryotes, aggregates are degraded via macroautophagy or the ubiquitin proteasome system<sup>83</sup>. The 26S proteasome, the central degradation machine in the eukaryote cytosol, degrades aggregates that have been marked for degradation by specific ubiquitin ligases. Conversely, in macroautophagy, entire protein aggregates can be engulfed in a double membrane vesicle termed the autophagosome. They fuse to the lysosome, which leads to the degradation of the cargo via lysosomal enzymes<sup>83</sup>.

A drastic approach to isolate a protein aggregate is the complete asymmetric sequestration of the aggregate into a single cell. This enables the generation of aggregate-free daughter cells upon cell division (**Figure 1.8**). This ancient mechanism appears to be used by all organisms. Within bacterial cells, the inclusion body localises to the old cell pole, and consequently stays in the older cell during cell division<sup>84</sup> (**Figure 1.8a**). Similarly, in yeast cells oxidatively damaged and aggregated proteins are partitioned to

the mother cell during cell division through tethering to the actin skeleton (**Figure 1.8b**)<sup>84, 96</sup>. Yeast cells can only undergo a finite number of divisions before they die, and it is thought that the limiting factor in this process is the accumulation of damaged proteins in the mother cell<sup>96</sup>.



**Figure 1.8.** Asymmetric partitioning of protein aggregates. **(a)** In bacterial cells, inclusion bodies are localised to the old cell pole, and remain, therefore, in the older cell during division. The older cell displays a reduction in growth rate. **(b)** In yeast cells, aggregates remain in the mother cell during cell division, and are even transported from the budding daughter cell back into the mother cell along the actin cable. Figure redrawn and adapted from Tyedmers, 2010<sup>84</sup>.

## 1.4 Protein aggregation and disease

A wide variety of diseases result from the incorrect folding of proteins. Many of the diseases are associated with the formation of extracellular amyloid fibrils (amyloidoses), or intramolecular inclusions with amyloid-like characteristics. These diseases range from neurodegenerative disorders such as Alzheimer's and Huntington's disease, to non-neuropathic localised amyloidoses and systemic amyloidoses. In many cases the disease manifests sporadically in a population, however there are also numerous associated hereditary causes. Furthermore, cases of transmissible aggregation-based diseases, and medical intervention-associated amyloidoses, are known (see **Table 1.1** for examples). The disease can arise from a loss of function of the protein, or a gain of toxic function. Moreover, additional complications that arise from the presence of misfolded and/or aggregated proteins can exacerbate the disease state further, as discussed below.

### 1.4.1 Loss of function

By definition, a misfolded protein is not in a functionally correct conformation, and is therefore incapable of performing its normal biological activity. Several diseases are caused by a loss of protein function because of misfolding, aggregation, or a general reduction in cellular levels of the protein. For example, cystic fibrosis (CF) is caused by a defect in the Cl<sup>-</sup> channel protein CFTR (cystic fibrosis transmembrane conductance regulator). More than a thousand mutants are known to lead to CF, but the most common (more than 70 %) is the deletion of the F508 residue<sup>110</sup>. This mutation results in the chaperone-mediated retention of the protein in the ER and subsequent degradation by the proteasome. What is greatly unfortunate is that the  $\Delta$ F508 mutant could function correctly at the plasma membrane, if it were able to reach it<sup>110</sup>.

Loss of function has also been hypothesised to account for the neurodegenerative disorders Huntington's disease (HD) and amyotrophic lateral sclerosis (ALS) (**Table 1.1**). In ALS, the misfolded protein superoxide dismutase 1 (SOD1) is unable to catalyse the conversion of toxic superoxide anions into hydrogen peroxide, thereby resulting in a build-up of superoxide radicals<sup>111</sup>. Similarly, the huntingtin protein (the protein that aggregates in Huntington's disease) is known to play a protective role in neurons against apoptosis<sup>112</sup>. It must be noted, however, that evidence against the loss of function hypothesis has also been found. Although mice homozygous for huntingtin die early in embryonic development<sup>112</sup>, patients that are heterozygous for huntingtin mutations have



similar clinical features to patients that are homozygous<sup>111</sup>. As for ALS, SOD1 knockout studies have shown no increase in neuronal degeneration<sup>113</sup>. The prominent theory in the field now proposes that protein misfolding diseases are a complex combination of loss of protein function, amplified through toxic gain of function and general cellular damage.

Disease class	Disease	Protein involved	Location of aggregates	Clinical features	Ref.
Neuro-degenerative	Transmissible spongiform encephalopathies	Prion protein	Extracellular amyloid fibrils in CNS	Dementia, ataxia, psychiatric problems	114
	Alzheimer's	Amyloid- $\beta$ and tau	Extracellular amyloid plaques, cytoplasmic neurofibrillar tangles	Progressive dementia	115
	Parkinson's	$\alpha$ -synuclein	Cytoplasmic Lewy bodies	Movement disorder	116
	Amyotrophic lateral sclerosis	Superoxide dismutase	Cytoplasmic and axon inclusions	Movement disorder	117
	Huntington's	Huntingtin	Intranuclear neuron inclusions	Dementia, motor and psychiatric problems	118
Non-neuropathic localised amyloidoses	Type II diabetes mellitus	hIAPP (amylin)	Pancreatic amyloid plaques	Hyperglycaemia, insulin deficiency and resistance	119
	Injection-localised amyloidosis	Insulin	Injection localised plaques	Sub-therapeutic levels of insulin	120
	Apolipoprotein A1 amyloidosis	Apo A-1 fragments	Renal, hepatic, cardiac, laryngeal, or cutaneous fibrillar deposits	Progressive renal and/or hepatic disease	121
Non-neuropathic systemic amyloidosis	Amyloid light chain amyloidosis	Immuno-globulin light chain fragments	Fibrillar deposits in various organs, most commonly kidneys	Kidney/renal failure	122
	Amyloid A amyloidosis	Serum amyloid A1 protein fragments	Kidney, liver and spleen	Renal failure	123
	Haemodialysis-related amyloidosis	$\beta_2$ -microglobulin	Amyloid plaques in joints	Renal failure and paraplegia	124
	Lysozyme amyloidosis	Lysozyme	Fibrillar deposits in leukocytes and macrophages	Renal failure	125

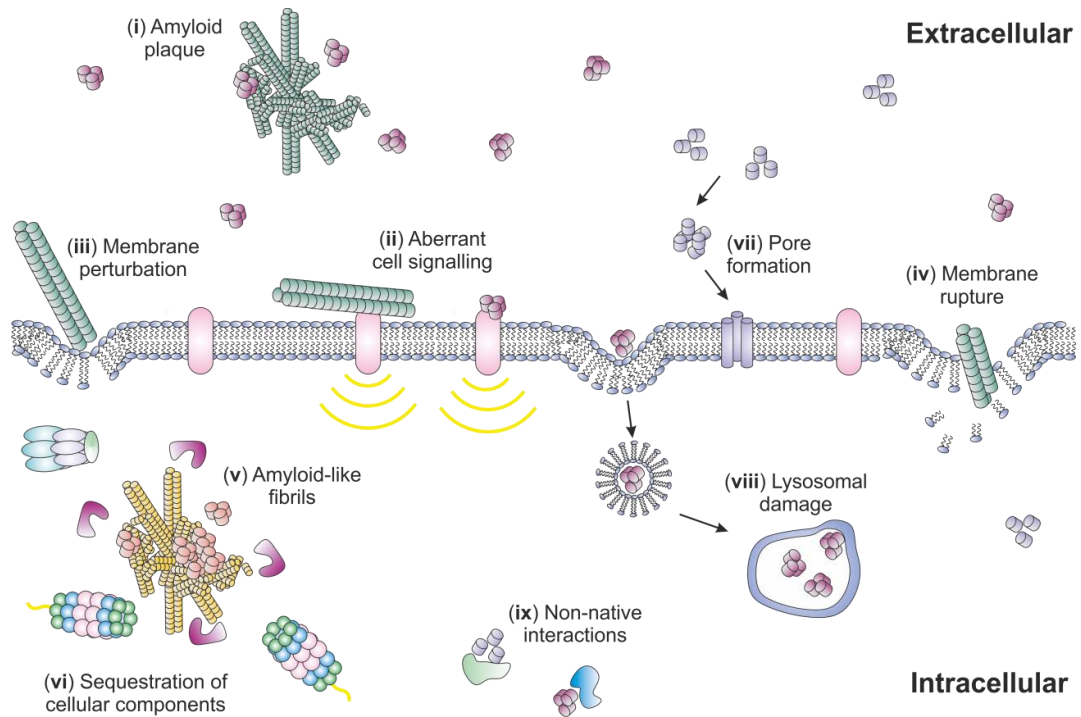
**Table 1.1.** The main human diseases associated with amyloid or amyloid-like deposits.

### 1.4.2 Gain of function

Misfolded and aggregated proteins exhibiting partial or total loss of native structure can impair the health of an individual through a gain of toxic function<sup>26</sup>. Indeed, there is significant evidence for this hypothesis. Animals transgenic for human misfolded proteins subsequently acquire neurodegeneration<sup>111</sup>, and, even more markedly, oligomerisation of non-disease related proteins is inherently cytotoxic<sup>126</sup>.

The toxic activity of protein aggregates depends on both the location of the aggregate, and its morphology. Extracellular aggregation disorders are associated with the deposition of elongated, unbranched, fibrillar protein assemblies (**Figure 1.9 i**). These diseases can be acquired or hereditary, and affect a multitude of systems and organs including the brain, nervous system, liver or heart (**Table 1.1**). These extracellular aggregates can activate a signal transduction pathway that leads to apoptosis by interacting with specific cell surface receptors (**Figure 1.9 ii**). An example of this is RAGE (receptor for advanced glycation end products), a cell surface receptor that has been demonstrated to interact with amyloid fibrils made from amyloid  $\beta$  ( $A\beta$ ), the prion protein (PrP), and human islet amyloid polypeptide (hIAPP). This interaction leads to the activation of cellular stress responses and the aberrant expression of NF- $\kappa$ B (a key regulator of the immune system)<sup>127</sup>. There is also significant evidence that these large fibrillar structures can disrupt membrane structure (**Figure 1.9 iii**)<sup>128-131</sup>. This perturbation can hinder normal cellular activity, and in some cases the cell will fully rupture (**Figure 1.9 iv**).

Intracellular fibrils possess many of the same physical characteristics as extracellular amyloid fibrils and are therefore often referred to as amyloid-like fibrils (**Figure 1.9 v**)<sup>41</sup>. These aggregates can be localised in the cytoplasm, known as aggresomes, or as Lewy or Russell bodies in the nucleus and endoplasmic reticulum. Intracellular aggregates can damage cells via the recruitment of essential factors for cell viability. For example, components of the proteasome and numerous chaperones are recruited to the forming aggregates, thereby preventing them from carrying out their essential functions in other parts of the cell (**Figure 1.9 vi**)<sup>132, 133</sup>. As with extracellular fibrillar structures, intracellular amyloid-like fibrils can also damage cell component membranes<sup>131, 134, 135</sup>. An example of this is thalassaemia, a disease that is characterised by an excess of  $\alpha$ - and  $\beta$ -globin subunits that do not associate into functional tetramers<sup>60</sup>. The excess subunits precipitate and distort erythrocyte shape, leading to their destruction by the spleen. This loss of red blood cells eventually leads to anaemia.



**Figure 1.9.** Mechanisms of protein aggregation toxicity. Extracellular fibrillar aggregates are known as amyloid plaques (i). These can cause cell toxicity through aberrant cellular signalling (ii), and membrane perturbation (iii) and rupturing (iv). Intracellular fibrils, known as amyloid-like fibrils (v), can sequester cellular components away from their required functions (vi). Aggregates on pathway to amyloid fibrils exert significant cytotoxic effects via pore formation (vii), cellular component damage (viii) and even non-native interactions with other cellular components (ix).

Until relatively recently, it was generally assumed that the most toxic form of protein aggregate was the mature amyloid fibril<sup>136-139</sup>. However, although aggregates are a typical feature of many neurodegenerative and systemic disease, the determination of whether or not they are directly involved in the pathogenesis of the disease has been difficult. Indeed, post-mortem studies have shown a poor correlation between plaque load and disease severity in Alzheimer's disease<sup>111</sup>. Recently, much interest has focused on the characteristics and formation of the pre-fibrillar species (also known as oligomers, amorphous aggregates, protein micelles or protofibrils), as mounting evidence suggests these, rather than the mature amyloid fibrils, are the toxic species<sup>47, 140-143</sup>. This hypothesis helps to explain the lack of a direct correlation between fibrillar plaque load and disease severity. Furthermore, it has been found that the pre-fibrillar aggregates formed *in vitro*

by non-disease related proteins are severely toxic to cells, in contrast to the relatively harmless effects of the mature fibrils<sup>126</sup>.

Part of the toxicity of these pre-fibrillar species appears to result from their ability to form 'pore-like' structures in cellular membranes (**Figure 1.9 vii**). This can lead to an alteration in ion-homeostasis and dysregulation of cell signal transduction, eventually leading to apoptotic or necrotic cell death<sup>47, 144, 145</sup>. The toxicity of the early species on the aggregation pathway also appears to result from their aberrant interactions with cellular processes. The endocytosis of A $\beta$  oligomers, for example, leads to their accumulation in lysosomes, which eventually leads to lysosomal damage (**Figure 1.9 viii**)<sup>146</sup>. Some protein aggregates have even been found to induce oxidative stress within cells, leading to lipid oxidation, elevation of intracellular calcium, and mitochondrial dysfunction<sup>111, 147, 148</sup>.

The greater toxicity of the smaller oligomeric species relative to the larger fibrillar aggregates is not unexpected, since the smaller the constituents, the larger the surface area-to-volume ratio. As these misfolded oligomers display groups on their surface that would, under normal conditions, be inaccessible, the potential for aberrant interactions is greatly increased. These non-native surfaces have been found to interact inappropriately with a multitude of cellular components, ranging from other proteins, to nucleic acids and lipid membranes (**Figure 1.9 ix**)<sup>27, 130, 149</sup>. These observations have further fuelled the theory that the mature fibrils act as a reservoir of toxic species by reducing their surface-to-volume ratio, and thereby extent of hydrophobic region exposure<sup>150</sup>. These pre-fibrillar species are also of key interest in the study of amyloid disease transmissibility. It is believed that transmission of the prion group of diseases results from the ingestion of prion proteins that have already begun to aggregate, thereby 'seeding' the formation of amyloid fibrils. Such a seeding mechanism could also explain the rapid progression of sporadic diseases such as Alzheimer's disease<sup>42</sup>.

### 1.4.3 Functional amyloid

In the interest of completeness, it should be noted that functional aggregates do exist. Indeed, examples have been found in species ranging from bacteria through to insects and fish (**Table 1.2**). Fibril formation is now believed to be an evolutionarily conserved biological pathway used to generate natural, stable structures with novel biological functions<sup>151</sup>. The same qualities that make fibrils a useful biological material in these applications, make them a challenge in disease-states: they are robust and durable. The

production of functional amyloid is under strict cellular control, specifically by key chaperones that are expressed under the same promoter as the amyloid proteins<sup>152</sup>. Importantly, as the operons from different species are not related, and the chaperones are not homologous, it suggests that the controlled production of functional amyloid has evolved numerous times independently. Understanding the different mechanisms nature uses to control such a potentially dangerous cellular component could provide valuable insights into novel therapeutic avenues.

Kingdom	Class	Protein	Role	Mechanism
Prokaryota	Gram negative bacteria (e.g. <i>E. coli</i> and <i>S. enterica</i> )	CurlI <sup>153</sup>	Protection, host invasion	Environmental conditions control the production of extracellular fibrils which play a role in cell surface adhesion, biofilm formation and host invasion.
	Gram positive bacteria (e.g. <i>S. coelicolor</i> )	Chaplin <sup>154</sup>	Sporing	Form an insoluble mat of fibres at an air-water interface to lower the surface tension, enabling filaments to grow out of water (sporing).
Fungi	Yeast ( <i>S. cerevisiae</i> )	Sup35; URE2p <sup>155</sup>	Non-Mendelian phenotypic inheritance; metabolic regulation	Sup35 is a translation termination factor. Under times of stress it aggregates, which prevents translation termination. This generates phenotypic diversity. URE2p regulates nitrate catabolism; when nitrogen levels are low, URE2p aggregates, which leads to an upregulation in other nutrient catabolism.
	Sordariomycetes (e.g. <i>P. anserine</i> )	HET-s <sup>156</sup>	Prevention of unviable cell fusions	Cellular fusion between genetically incompatible colonies is prevented if a cell expressing the prion HET-s attempts to fuse with a cell expressing the soluble HET-S. When coexpressed, cell death is induced, thereby preventing any further colony fusion.
Animalia	Arachnida (e.g. all spiders)	Spidroin <sup>157</sup>	Structural; protection	Multiple forms of silk formed for various uses, including web building (structural) and drag lines (the strongest type of silk).
	Reptilia, amphibian, aves, insect	Chorion <sup>158</sup>	Structural protection	Fibres in egg shell protect the egg from hazards such as proteases, microorganisms and physical stress.
	Mammalia	Pmel7 (M $\alpha$ ) <sup>159</sup>	Sequestering toxic species	Fibrils formed in melanosomes accelerate the polymerisation of the toxic monomers indolequinone into melanin by orientating the molecules along the fibre.

**Table 1.2.** Examples of the diversity of functional amyloid.

## 1.5 Therapeutic approaches to protein aggregation

Due to the complex mechanisms involved in aggregation-based diseases, there are currently alarmingly few therapies available. Indeed, as the toxic species in many of these disorders remains elusive, the commonly available therapies focus only on ameliorating the symptoms<sup>160</sup>. There are, however, a number of promising therapeutic avenues being intensively investigated, whose diversity demonstrates the multifactorial nature of these diseases.

### 1.5.1 Early diagnosis

The search for diagnostics that can identify sub-clinical manifestations of amyloid disorders is an active area of research as patients can be asymptomatic for a long period of time, so the damage caused is both extensive and currently irreversible. Early diagnosis, therefore, is vital for enabling intervention before extensive cellular damage has occurred. Furthermore, the identification of high-risk groups and individuals would enable targeted treatment. For example, roughly 500 families worldwide live with an inherited form of Alzheimer's disease that strikes much earlier than the sporadic form of the disease<sup>161</sup>. As testing preventative treatment in the general population is neither feasible nor ethical, these families are the focus of a worldwide effort to understand, diagnose and prevent the disease at the earliest possible stage. Because the risk of Alzheimer's disease is relatively low in the general population, and onset currently impossible to predict, such a trial would have to enrol thousands of people and would subject many who would never have developed the disease to the unknown long-term risks of taking novel therapeutics<sup>161</sup>. By studying these families, it is much easier to tell if the therapeutic is working early on. Furthermore, for a clinical trial to be balanced and unbiased, a placebo group must be included. For patients in the late stages of the disease, when fatality is inevitable, there is much debate on the placement of such individuals on placebo treatment. Through studying these families, it has been established that the concentration of amyloid  $\beta$  ( $A\beta$ ) in cerebrospinal fluid decreases up to 25 years before the onset of Alzheimer's disease<sup>162, 163</sup>. This is believed to indicate that the  $A\beta$  is being taken out of circulation as it begins to build up in the brain. This discovery will enable clinicians to identify patients at the earliest stage of the disease for prophylactic treatment.

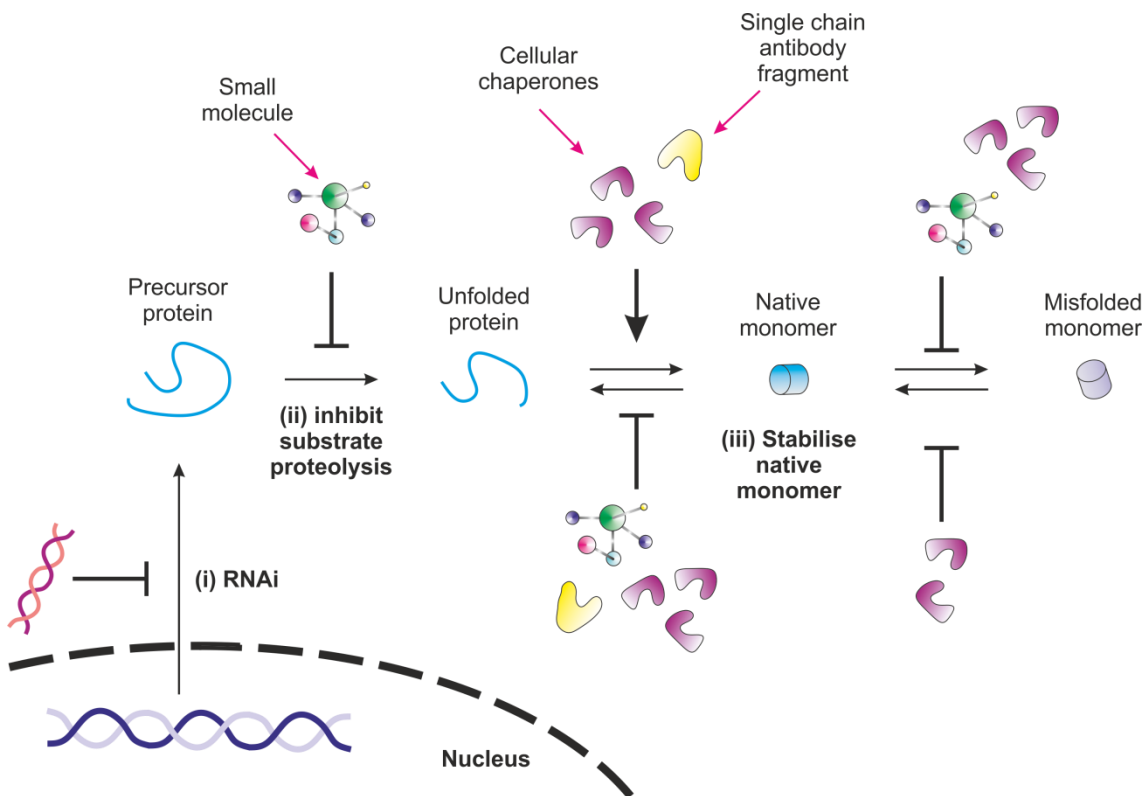
### 1.5.2 Repairing cellular damage

As there is a prominent lack of successful therapeutics against the amyloidosis group of disorders, a key avenue of research is the reversal of cellular damage caused during the disease lifetime. For example, embryonic stem cells<sup>164</sup> and neuronal precursor cells<sup>165, 166</sup> can migrate within the brain to sites of injury (including neuronal degeneration<sup>167</sup>) and differentiate into various neuronal cell types. Unfortunately, however, the level of neuronal loss in neurodegenerative diseases is already severe by the time the symptoms have arisen (it can be ten years before the symptoms of Alzheimer's disease show). Despite this, reversing the accumulated damage caused by these diseases will most likely play a complementary role in the treatment of patients, once successful prophylactics have been discovered.

Similarly, it is becoming apparent that toxicity in these diseases is not only directly caused by protein aggregation. Oxidative damage, inflammation and mitochondrial impairments have all been observed in Alzheimer's disease<sup>168-170</sup>, and as such, there has been significant research into slowing disease progression with antioxidants<sup>171</sup>, anti-inflammatory drugs<sup>172</sup> or mitochondrial protectors<sup>173</sup>. It is widely suspected that these will play an important complementary role in any future prophylaxes for protein aggregation diseases.

### 1.5.3 Inhibiting production of monomeric form

One obvious therapeutic avenue to prevent protein aggregation is the direct reduction in the level of abnormal protein within the cell. RNA interference (RNAi), the use of small double-stranded RNA molecules to silence specific gene expression, is a powerful biotechnological tool. Interfering RNAs target specific messenger RNAs for degradation, thereby silencing their expression (**Figure 1.10 i**). Incorporation of these RNAs into replication-deficient viral vectors, and the integration of the virus into neuronal DNA, results in stable long-term expression of these RNA molecules<sup>174-177</sup>. Indeed, RNAi has shown promise against polyglutamine diseases<sup>178, 179</sup>, Alzheimer's disease<sup>180</sup> and prion diseases<sup>181</sup>. However, the delivery of such a therapeutic would have to overcome the formidable challenge of crossing the blood-brain barrier (BBB) to access the neurons affected by neurodegenerative disorders. This may be feasible in diseases such as Parkinson's disease, which only effect a small region of the brain, as the viral vectors could be directly injected<sup>182</sup>.



**Figure 1.10.** Pre-aggregation therapeutic strategies against protein aggregation. Preventing production of the aggregation-prone peptide can be achieved by RNA interference (RNAi) (i) or through inhibition of precursor protein proteolysis (ii). The native monomer can also be stabilised to prevent unfolding or misfolding (iii). This has been achieved via upregulation of cellular chaperones and treatment with small molecule stabilisers and single chain antibody fragments.

In about half of the human amyloidogenic diseases, the amyloidogenic component is produced by cleavage of a larger precursor protein<sup>183</sup>. A viable therapeutic strategy, therefore, is to inhibit the proteases that generate these fragments (**Figure 1.10 ii**). For example, there are inhibitors of both  $\gamma$ - and  $\beta$ -secretase to prevent the processing of the amyloid precursor protein (APP) into the A $\beta$  peptide that is associated with Alzheimer's disease<sup>184, 185</sup>. A challenge with this therapeutic avenue, however, is to develop a drug that specifically inhibits the cleavage of the amyloidogenic peptide without affecting the cleavage of alternative substrates. Indeed, a phase III clinical trial of a  $\gamma$ -secretase was recently stopped due to side effects including worsening of cognitive impairments. It is suspected that this was due to the prevention of alternative cleavage product formation<sup>170</sup>.



Consequently, another avenue of drug development focuses on modulating these proteases, rather than inhibiting them. In this vein, a compound that can decrease the level of the most amyloidogenic cleavage product in favour of less toxic species is currently being sought<sup>186</sup>. However, it has still not been excluded that these other peptides also contribute to the toxicity of Alzheimer's disease. Furthermore, this strategy is inapplicable to many amyloid disorders, such as Creutzfeldt-Jakob disease, Parkinson's disease and Huntington's disease, in which the full length protein aggregates.

Significantly, this form of therapeutic requires blocking or down-regulating the expression of a native protein that has evolved to carry out a particular biological function. As the function of many of the proteins and peptides involved in these diseases is still not fully understood, it may not be possible to identify inhibitors that are safe to use *in vivo*<sup>187</sup>.

#### **1.5.4 Stabilising the native protein conformation**

The use of chemical chaperones to bind to a protein and stabilise the folded state is one of the most promising therapeutic approaches to ameliorating aggregation-based diseases (**Figure 1.10 iii**). Indeed, the only regulatory-approved drug that can slow the progression of a human amyloid disease (familial amyloid polyneuropathy, FAP) is a chemical chaperone. In FAP, tetramer dissociation of the protein transthyretin (TTR) into a misfolded monomer leads to the formation of non-fibrillar and amyloid aggregates in the plasma<sup>79</sup>. Tetramer dissociation is the rate-limiting step, and all TTR mutations associated with the familial form of the disease decrease tetramer stability<sup>188</sup>. Using structure-based drug design, the Kelly lab discovered a compound (Tafamidis) that binds to the correctly folded tetramer and kinetically stabilises it (see Section 1.6.1)<sup>189-191</sup>.

Single-domain fragments of a camelid antibody have also been utilised to increase protein stability of an amyloidogenic lysozyme variant. The antibody functions by restoring the cooperativity between two structural domains within the native lysozyme fold, thereby reducing the probability of partially unfolded state formation<sup>192</sup>. Chemical chaperones that prevent misfolding of the prion protein PrP<sup>193</sup> and A $\beta$ <sup>194</sup> have also been found.

A less direct approach to stabilise the native protein is to target cellular components such as chaperones. Alpha 1-antitrypsin ( $\alpha$ 1-AT) is a member of the serpin family of protease inhibitors and is secreted by liver cells to control the proteolytic activity of circulating enzymes<sup>79</sup>. In  $\alpha$ 1-AT deficiency, a misfolded but otherwise functional mutant protein is

retained in the ER. Osmolytes such as 4-phenylbutyric acid and glycerol have been shown to increase the secretion of  $\alpha$ 1-AT mutant variants. They do this without influencing the secretion efficiency of other proteins and, as they are known to not bind to  $\alpha$ 1-AT, it is speculated that they upregulate the chaperone system to bring about their effects<sup>195</sup>.

### 1.5.5 Increasing degradation of misfolded or aggregated protein

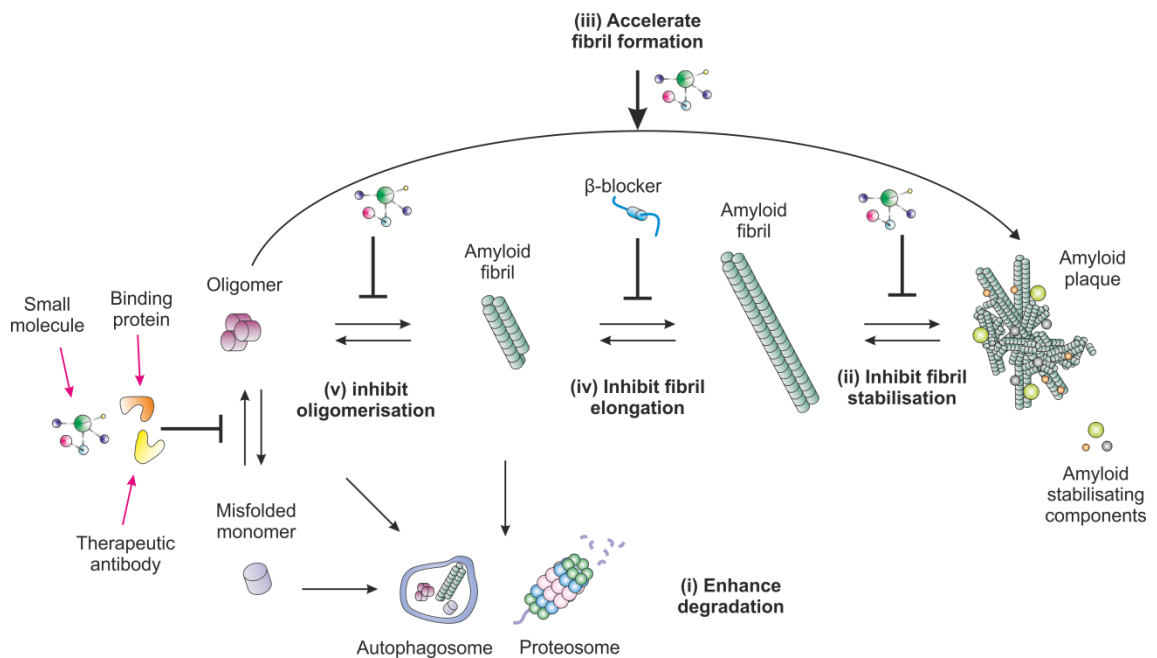
Efforts are being directed towards enhancing cellular defence mechanisms through the use of proteostasis regulators (**Figure 1.11 i**). Small molecule compounds have been developed that bind to intracellular targets in the endosomal-lysosomal pathway, and facilitate degradation of specific misfolded toxic proteins. Other compounds are currently under development to enhance cellular chaperone levels and to aid clearance of toxic protein aggregates. Geldanamycin, for example, can modulate and enhance chaperone levels in a defective heat shock response by stabilising functional heat shock factor trimers<sup>196</sup>. However, this compound, in its current form, exhibits substantial toxicity and cannot cross the blood-brain barrier<sup>34</sup>.

There is also significant interest in increasing autophagy in cells containing protein aggregates. This relatively non-selective process for degrading aggregates (via portions of cytoplasm being engulfed and delivered to the lysosome) has the benefit of not requiring the protein aggregate to be disentangled prior to degradation<sup>197</sup> (**Figure 1.11 i**). Finally, proteasome stimulation is also being considered as a potential therapeutic strategy. The phenothiazine methylene blue has been reported to decrease soluble A $\beta$  concentration and reduce memory deficits in transgenic mice through augmentation of proteasome function<sup>198</sup>. However this could still be potentially dangerous as it may alter the turnover of other molecules normally regulated by proteasome degradation<sup>34</sup>.

Another mechanism to increase clearance of amyloid fibrils *in vivo* is inhibiting amyloid-stabilising proteins (**Figure 1.11 ii**). Accessory proteins, such as serum amyloid-P, apolipoprotein-E and glycosaminoglycans, bind to and stabilise amyloid fibrils<sup>22, 183</sup>. Sulphonated aromatics are known to interfere with this interaction, thereby accelerating the disassembly and clearance of the fibrils<sup>199</sup>. However, caution must be taken with this therapeutic approach as evidence strongly suggests that fibrils act as reservoirs for toxic species (see Section 1.3.2.2).

### 1.5.6 Accelerating fibril formation

Controversially, another therapeutic strategy currently being investigated is the use of compounds to accelerate the aggregation of oligomers into the less toxic mature fibrils (**Figure 1.11 iii**)<sup>200</sup>. Indeed, peptides with recognition elements homologous to A $\beta$  have been found that can accelerate A $\beta$  aggregation kinetics, whilst simultaneously reducing toxicity *in vivo*<sup>201</sup>. However, the pathological consequences of amyloid fibril accumulation cannot be excluded; there is extensive evidence that fibrils can damage membranes (see Section 1.4.2), induce inflammatory reactions<sup>202</sup>, and even catalyse the formation of toxic oligomeric species<sup>203</sup>.



**Figure 1.11.** Post-aggregation therapeutic strategies against protein aggregation. Protein aggregation can be prevented via enhancement of proteostasis regulators (**i**). If fibrils have formed, further production can be destabilised, by targeting amyloid stabilising components (**ii**), or, controversially, enhanced using small molecules that accelerate fibrillation (**iii**).  $\beta$ -blocker peptides can disrupt the fibrillation process to prevent elongation (**iv**). Preventing multiple stages of oligomerisation can be achieved using small molecule inhibitors, binding proteins and therapeutic antibodies (**v**).

### 1.5.7 Inhibiting or reversing protein oligomerisation

An attractive strategy for preventing protein aggregation disease is to target the aggregation pathway, in order to inhibit, or reverse, aggregation. Initially, therapeutic strategies focused on inhibiting fibril formation, or initiating fibril disaggregation. Several groups have focused their attention on the inhibition of amyloidosis by the incorporation of synthetic peptide derivatives. N-terminal modifications, for example, have been demonstrated to be very effective in preventing fibril extension. By incorporating a peptide with a bulky group, such as a steroid, at its terminus, further A $\beta$  association is hindered<sup>204</sup> (**Figure 1.11 iv**). N-methyl amino acid incorporation is another successful strategy. The peptides correspond to the amyloidogenic region of the protein; one side of the peptide presents a hydrogen-bonding complementary face to the amyloidogenic protein, whilst the other side has N-methyl groups in place of backbone NH groups. This 'blocking' face can prevent fibril formation<sup>205</sup> and even break down preformed fibrils<sup>206</sup>. In a similar manner, the substitution of key residues with prolines has also been shown to reduce  $\beta$ -sheet propensity of numerous amyloidogenic proteins, including A $\beta$ <sup>136</sup> and PrP<sup>187</sup>. These ' $\beta$ -sheet breaker' peptides, when produced as all-D amino acids, have the added benefit of increased protease resistance, allowing them sufficient stability *in vivo* to cross the blood-brain barrier<sup>207</sup>. Similar peptide-based inhibitors have also been found for PrP,  $\alpha$ -synuclein, human islet amyloid polypeptide (hIAPP) and poly-Q using similar strategies<sup>208-212</sup>.

Therapeutic antibodies and other binding proteins have also been developed for treating amyloidoses (**Figure 1.11 v**). A recent study engineered a small binding protein (known as an Affibody) that could bind to A $\beta$  with nanomolar affinity, and subsequently prevent the initial association of A $\beta$  monomers into oligomers by occluding the aggregating region<sup>213</sup>. Similarly, antibodies can specifically inhibit aggregation by traditional antigen-antibody interactions<sup>214, 215</sup>. The antigen epitope can consist of a partial sequence of the amyloidogenic protein, or even amyloid-specific secondary or tertiary structures. Interestingly, an oligomer-specific antibody that can inhibit A $\beta$  aggregation<sup>216</sup> and A $\beta$  induced neurotoxicity<sup>217</sup> can also inhibit the toxicity attributed to other oligomers of amyloidogenic proteins, including  $\alpha$ -synuclein, hIAPP, poly-Q and lysozyme<sup>214</sup>. This suggests that many oligomers contain a common conformational binding epitope. One of the major drawbacks with this therapeutic approach is that antibodies cannot cross the blood-brain barrier. However, as immunisation of transgenic A $\beta$  mice can prevent the onset of Alzheimer's disease<sup>218, 219</sup>, it is believed that peripheral antibodies can bind to A $\beta$

peptides in the blood, hence lowering the available A $\beta$  in the brain for aggregation<sup>183</sup>. These findings demonstrate the potential for this therapeutic approach.

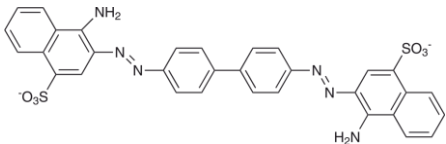
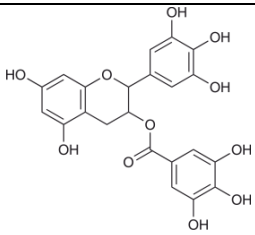

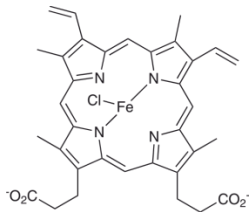
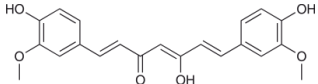
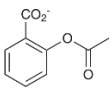
As evidence accumulated for the toxicity of the pre-fibrillar species, focus shifted away from inhibiting amyloid formation and towards inhibiting accumulation of the smaller oligomeric species. As the toxic species of many of these disorders are still unknown, by inhibiting the earliest possible stage in the aggregation pathway *all* potential toxic species will be prevented from forming (**Figure 1.11 v**). Recent clinical trials seem to support this hypothesis, as numerous drugs that inhibit A $\beta$  fibril formation by sequestering monomeric A $\beta$  peptides, such as cyclohexanehexol (Phase III)<sup>220</sup> and curcumin (Phase II)<sup>221, 222</sup>, are at various stages of development<sup>222, 223</sup>. Although tramiprosate, one of the most promising drugs against Alzheimer's disease, failed in phase III clinical trials (it was not significantly better than a placebo in improving cognitive function), the study proved that small molecule inhibitors of protofibril formation are capable of reaching late stage development and may still be of therapeutic value<sup>223</sup>.

A related approach is to trap the aggregation-prone protein in an on-pathway intermediate state. For example, the polyphenol exifone can inhibit  $\alpha$ -synuclein fibril formation by driving the aggregation pathway towards non-toxic oligomeric species, thereby preventing cytotoxicity<sup>224</sup>. Chaperones have also been targeted to developing protein aggregates in order to arrest any further aggregation. The molecule SLF-CR, for example, has been developed to recruit the FK506 binding protein family of chaperones to a developing A $\beta$  aggregate and inhibit subsequent aggregation<sup>225</sup>. It was created by linking Congo red, a dye known to bind A $\beta$  amyloid, and a synthetic ligand of FK binding protein 12 (SLF).

### 1.5.7.1 Small molecule inhibitors

Extensive work has been carried out using small molecules (low molecular weight organic compounds) to prevent or interfere with the aggregation pathway. Small molecule inhibitors of protein aggregation have some advantages over peptide inhibitors: they can more easily cross the blood-brain barrier, they can avoid an immunological response, and are more stable in biological fluids and tissues<sup>226</sup>. The small molecule inhibitor approach was initially based on the early observations that aromatic compounds such as thioflavin T and Congo red could bind to amyloid fibrils and, at high enough concentrations, inhibit their formation<sup>227-230</sup>. Since then, a vast array of compounds has been demonstrated to

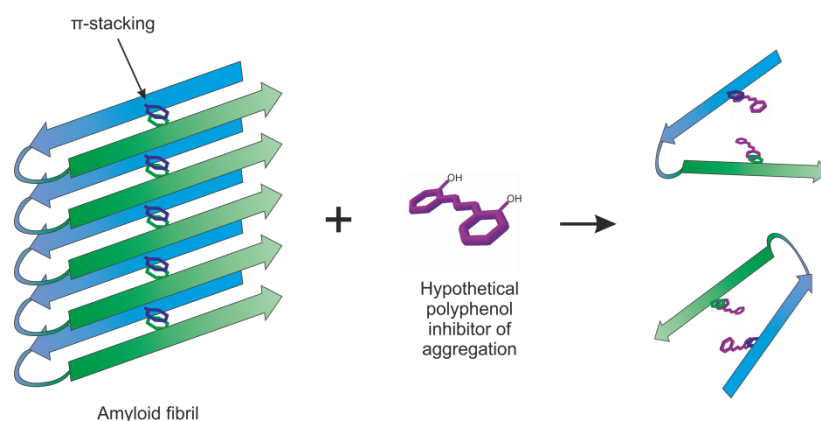
interfere with single, or multiple, points in the aggregation pathways of many proteins. **Table 1.3** contains select examples to highlight the diversity in structural and chemical properties, which suggests that they may be binding to different sites within protein aggregates or amyloid fibrils.

Small Molecule	Structure	Protein Target
Congo red		A $\beta$ <sup>228, 231</sup> , insulin <sup>232</sup> , Ig domain <sup>233</sup>
EGCG		A $\beta$ <sup>234</sup> , hIAPP <sup>235-237</sup> , TTR <sup>238</sup> , $\kappa$ -casein <sup>239</sup> , $\alpha$ -synuclein <sup>240</sup>
Tramiprosate		A $\beta$ <sup>222</sup>
Hemin		A $\beta$ <sup>241, 242</sup>
Curcumin		A $\beta$ <sup>243, 244</sup> , hIAPP <sup>245, 246</sup> , $\alpha$ -synuclein <sup>247</sup> , lysozyme <sup>248</sup>
Aspirin		A $\beta$ <sup>249</sup>

**Table 1.3.** Select examples of small molecule inhibitors of protein aggregation.

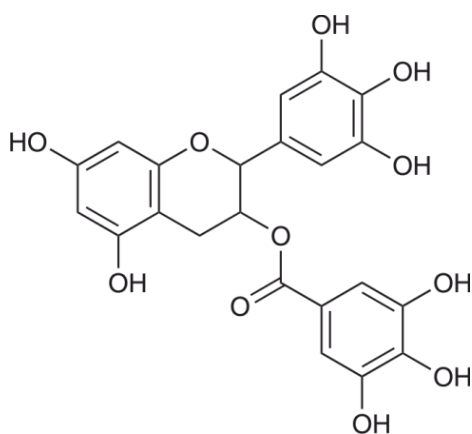
The most common small-molecule inhibitors of aggregation belong to the family of polyphenol compounds and contain two functional moieties: aromatic rings, and polar groups on the rings<sup>250</sup>. These natural compounds are found in high concentrations in wine, tea, berries, cocoa and many other plants. Their physiological role is largely protective, from disease to exposure to ultraviolet radiation. Consequently, beneficial effects of polyphenol consumption have been linked to their possible antibacterial, antiviral, anticancer and neuroprotective properties<sup>214</sup>. Interest in this group of compounds has led to the discovery and isolation of more than 8000 polyphenolic compounds<sup>227</sup>.

The anti-aggregation and anti-amyloid activities of polyphenols are believed to be due to competitive interactions with key residues in the amyloidogenic proteins. Stacking of aromatic residues (' $\pi$ -stacking') is believed to play an important role in the formation of many amyloid fibrils. The aromatic rings of inhibitors can therefore competitively interact with aromatic residues in the proteins. This prevents  $\pi$ - $\pi$  interactions, and subsequently blocks any further self-assembly (**Figure 1.12**)<sup>214, 251-253</sup>. Indeed, structural analysis has shown the polyphenol Congo red interacts with the aromatic moieties of insulin to prevent fibril formation<sup>232</sup>. It must be noted, however, that recent evidence suggests that aromatic stacking is not the only feature that directs aggregation; replacing the aromatic side chains of the A $\beta$  peptide with leucine or isoleucine actually leads to enhanced amyloid formation<sup>254</sup>. This suggests that there may be other therapeutic targets within amyloidogenic proteins rather than just the amyloidogenic aromatic residues.



**Figure 1.12.**  $\pi$ -stacking theory of polyphenol anti-aggregation activity.  $\pi$ -stacking of aromatic groups is believed to play an important role in protein aggregation and fibril formation. Polyphenolic small molecules can competitively interact with the aromatic residues in amyloidogenic proteins, thereby preventing the  $\pi$ - $\pi$  interaction and blocking self-assembly.

The phenolic hydroxyls are also believed to play a pivotal role in aggregation inhibition, potentially via hydrogen bonding to the peptide backbone of the amyloidogenic protein<sup>234, 236</sup>. One of the most intensively studied small molecules, the tea-derived flavanol (-)-epigallocatechin-3-gallate (EGCG), has the ability to prevent oligomerisation and fibrillation of a number of amyloidogenic peptides, and also to disaggregate preformed fibrils *in vivo* and *in vitro*<sup>227, 234, 235, 237, 239, 240, 255-257</sup>. The presence of numerous phenolic hydroxyl groups (**Figure 1.13**) is believed to be key to its anti-amyloidogenic properties; when these hydroxyl groups are removed sequentially, the anti-aggregation properties of the molecule correspondingly decrease<sup>236</sup>.



**Figure 1.13.** Chemical structure of the polyphenol (-)-epigallocatechin-3-gallate (EGCG). The high number of phenolic hydroxyl groups is pivotal to its broad anti-amyloidogenic properties<sup>234, 236</sup>.

One disadvantage of polyphenol compounds as aggregation inhibitors is their lack of specificity. The preserved features of amyloid fibrils mean that polyphenol compounds are able to bind to a very wide variety of amyloid fibrils. This could be problematic for functional amyloid, such as that found in mammalian melanocytes and related cell types<sup>258</sup>. Furthermore, the lack of understanding of the nature of the toxic species makes inhibitors of fibril formation potentially very dangerous. As every species on the aggregation pathway is inherently non-native, each one could potentially contribute towards the toxicity. The search for small molecule inhibitors has therefore shifted in favour of those that prevent the earliest stages of aggregation.



## 1.6 Screening for aggregation inhibitors

Over the past decade, significant progress has been made in the development of small molecules that can inhibit the interaction between two proteins<sup>259</sup>. However, the difficulties encountered have been significant; the disparities in size, and geometry, of protein-protein interfaces and small molecules are immense. These interfaces are often relatively featureless (covering an area of 750 – 1500 Å<sup>2</sup>)<sup>259</sup> and devoid of pockets in which a small molecule can bind. The major breakthrough in the development of small molecule inhibitors of protein-protein interactions was the discovery of ‘hotspots’ on the interface surface. Alanine scanning determined that there were particular areas that contributed disproportionately to the binding energy<sup>260, 261</sup>. Furthermore, these hotspots are often enriched in aromatic residues<sup>223, 260</sup>, thus providing key targets for the aromatic-enriched polyphenol inhibitors of protein aggregation. Prediction software has since been developed to identify protein-protein interaction hotspots, based on data from crystallographic methods or alanine scanning<sup>262, 263</sup>.

### 1.6.1 *In vitro* screening techniques

For folded proteins, structure-based drug design has played a pivotal role in the search for inhibitors of protein aggregation<sup>187</sup>. Indeed, the only clinically-approved drug for an amyloid disease - familial amyloid polyneuropathy - was discovered using a combination of inhibitor screens and structure-based drug design. Several hundred small molecules, structurally similar to the ligand of the amyloidogenic protein TTR, were screened for their anti-aggregation abilities. Crystal structures of TTR-inhibitor complexes were then utilised to identify the mode of interaction. These crystal structures further indicated where substitutions could be made, that increase the small molecule affinity for TTR by increasing the surface area of van der Waals interactions<sup>264, 265</sup>. This combination of approaches led to the production of Tafamidis, a drug for the treatment of familial amyloid polyneuropathy.

For aggregation-prone proteins that lack defined structure, however, structure-based drug design is unsuitable. Discovery of small molecule inhibitors of aggregation is limited, therefore, to screening using biophysical techniques or dye binding studies<sup>252</sup>. Traditionally, the extrinsic fluorescence of the benzothiazole dye thioflavin T (ThT) was used to quantify the extent of fibrillation *in vitro*<sup>266-268</sup>. ThT binds to  $\beta$ -sheet-rich structures, such as the cross- $\beta$ -sheet found in amyloid fibrils, which causes it to fluoresce

strongly. Free ThT, conversely, shows only weak fluorescence. Although the mechanistic interaction between ThT and amyloid fibrils remains poorly understood, it is believed that it most probably involves intercalation of ThT molecules within grooves of the fibril<sup>269-271</sup>. This binding sterically restricts the ThT molecule, which prevents the formation of a less-radiative conformation<sup>272</sup>. For many years, the ThT assay was used as the sole measure of the *in vitro* anti-aggregation activity of a variety of small molecules<sup>273-276</sup>. However, as many of the compounds are chromophoric and even intrinsically fluorescent<sup>277-279</sup>, false-positive result reporting was inevitable<sup>280-283</sup>.

Dye binding assays have been used to identify numerous small molecule inhibitors of fibril formation<sup>266, 284-286</sup>, however this is also a significant drawback of these assays: they rely on a reporter that only detects amyloid fibrils rather than then early, potentially toxic species. Consequently other biophysical techniques have been utilised, such as transmission electron microscopy (TEM)<sup>235, 240, 287</sup> and atomic force microscopy (AFM)<sup>288-291</sup>, to assess the effects of small molecules on aggregation as a whole. However, these techniques have, so far, not been converted into a high throughput methodology. Conversely, *in vitro* assays such as dynamic light scattering can be set up in 96- or 384-well format and are therefore often utilised for semi- or high-throughput analysis of aggregation<sup>292, 293</sup>. Light scattering also provides a more quantitative analysis of the extent of aggregation in a sample, when compared to visual techniques such as TEM and AFM<sup>294</sup>. High-throughput identification of compounds that bind to specific proteins has also been accomplished using small molecule microarrays (SMMs). SMMs are glass slides on which libraries of small molecules are covalently immobilised in an array of microspots<sup>295</sup>. The slides are probed with a fluorophore- or epitope-tagged protein, and compounds that bind are detected by automated fluorescence read-out<sup>296</sup>.

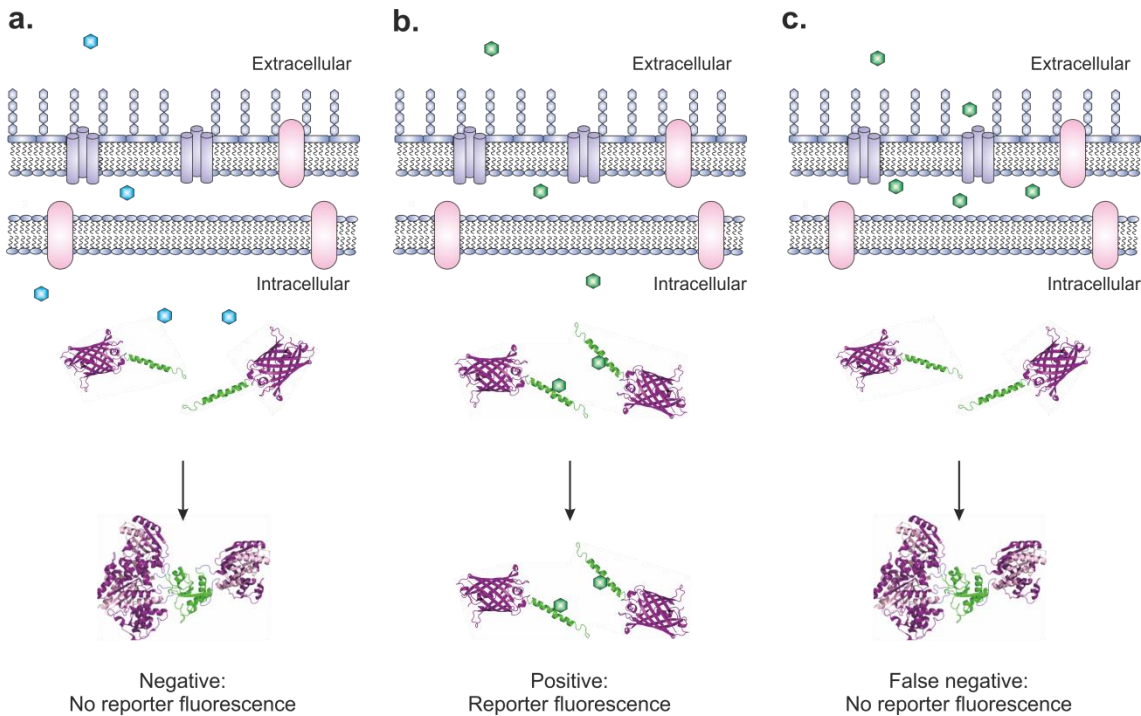
Although simple to perform, these, and other *in vitro* techniques, are significantly hampered by the need for substantial quantities of purified protein. This requires the *in vivo* expression, purification or expensive chemical synthesis of these challenging proteins, whose hydrophobic nature and tendency to form secondary structures, in particular  $\beta$ -sheet aggregates, hinders their isolation in significant yields<sup>297</sup>. These difficulties result in labour intensive and expensive procedures to obtain sufficient quantities of pure protein, in the required aggregation state, for *in vitro* screening of potential aggregation-inhibiting small molecule therapeutics. Furthermore, the synthetic peptide samples are difficult to produce in a form that is free of oligomeric 'seeds' that can nucleate further aggregation<sup>297</sup>. The presence of these pre-existing oligomers during

inhibitor screening could lead to some of the most important inhibitors, those that prevent low molecular weight oligomer formation, being overlooked.

### 1.6.2 *In vivo* screening techniques

In contrast to *in vitro* biochemical assays, an *in vivo* assay allows the investigator to probe the effects of small molecules on the protein of interest (POI), without the prerequisite for purified protein. Furthermore, bacterial systems provide a more physiological environment in which to assess aggregation<sup>298</sup>. Arrays of screens and selections have been developed for detecting protein aggregation in the cytoplasm of living cells. These approaches commonly involve a genetic fusion between a protein of interest and a reporter protein. In such systems, if the POI misfolds or aggregates, the reporter protein to which it is fused is rendered inactive (**Figure 1.14a**). If the POI folds into a soluble conformation, the reporter protein is functional (**Figure 1.14b**). A significant benefit to these approaches is that they can be used even when structural or functional information about the target is lacking.

Previous studies have shown the potential for utilising *in vivo* systems to study protein aggregation and its inhibition, however these have been solely confined to experiments within the cytosol<sup>299-303</sup>. The critical limitation of working in the cytosolic milieu is the restrictive nature of the inner membrane. This significantly limits the size and type of small molecule that can cross, potentially leading to the reporting of false negative results (**Figure 1.14c**). Furthermore, these studies involved the fusion of the aggregation-prone peptide to a reporter protein such as green fluorescent protein (GFP)<sup>299, 302, 303</sup>. A disadvantage of this approach is that the folding of the reporter protein can be dependent on the solubility of the upstream sequence. A false-positive result could arise from a small molecule with the ability to stabilise soluble, low molecular weight oligomers as the reporter could, in principle, still fold next to this low molecular weight aggregate. Moreover, these assays require pre-screening at the reporter protein fluorescence wavelength, and the subsequent elimination of any molecule that fluoresces at this wavelength<sup>299</sup>.



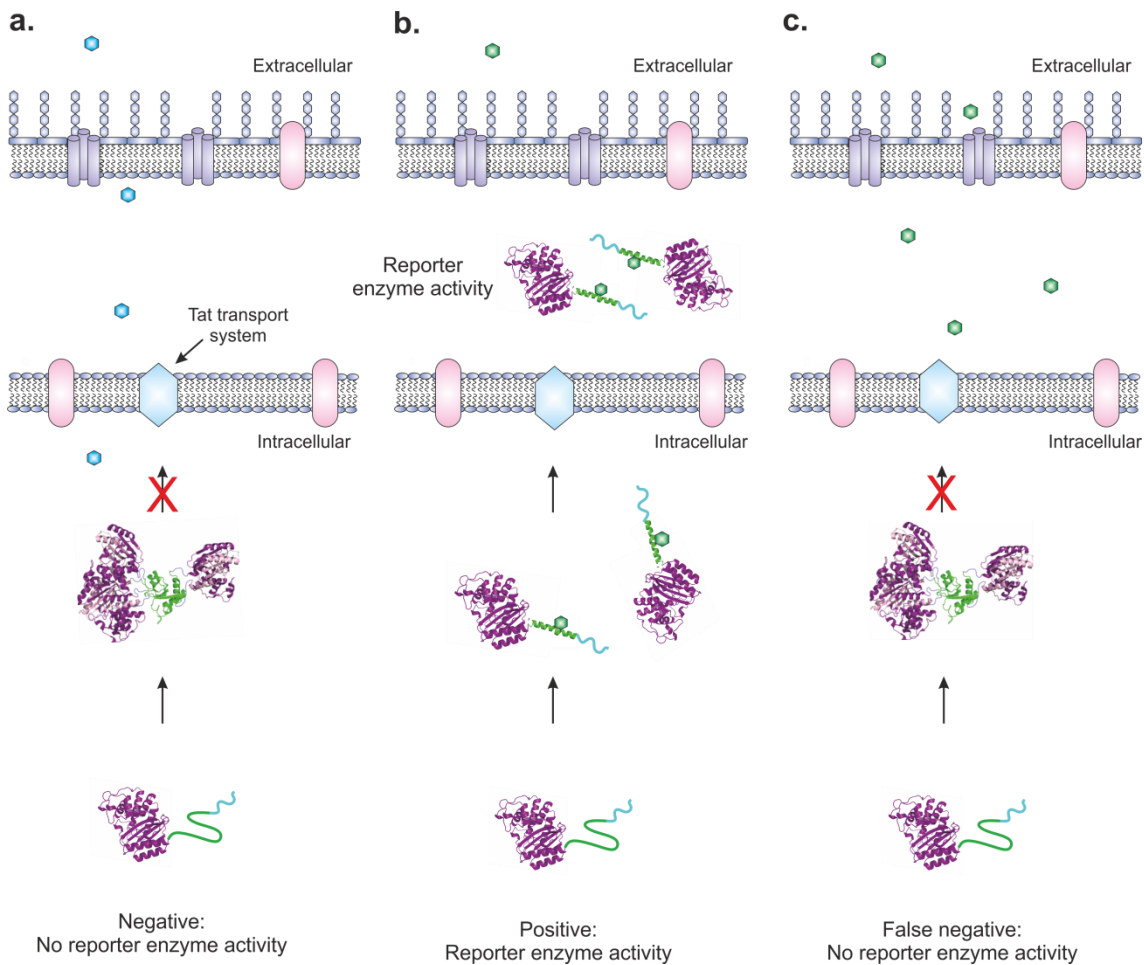
**Figure 1.14.** Cytoplasmic-based screen for inhibitors of protein aggregation. The protein of interest (POI) (green) is fused to a fluorescent reporter protein (purple) and expressed in the cytoplasm of bacteria. When grown in the presence of small molecules, if the molecule cannot prevent the POI aggregation (blue molecule), the reporter protein will not be able to fluoresce (**a**). If however, the molecule can bind and prevent POI aggregation (green molecule), reporter protein fluorescence is rescued (**b**). The molecules must, however, be able to traverse the inner membrane of the bacteria. If an inhibitor cannot cross this barrier a false-negative result will be observed (**c**).

### 1.6.2.1 Enzyme activity-based screening techniques

To overcome the disadvantages of using fluorescence-based screening assays, enzyme activity assays have more recently been utilised, particularly exploiting  $\beta$ -lactamase enzymes.  $\beta$ -lactamase has been used as a reporter system for many years due to its relatively small size (29 kDa) and monomeric nature. In addition, it has been expressed successfully in both prokaryotic and eukaryotic cells<sup>304</sup>. Initially,  $\beta$ -lactamases were used in assays based on the complementation of enzyme fragments. One simple example of this is the *in vitro*  $\beta$ -lactamase protein fragment complementation assay, in which protein-protein interactions are probed<sup>305</sup>. The two domains of  $\beta$ -lactamase are genetically separated and each is attached to the gene of a POI. If these proteins interact with each other, the two domains of  $\beta$ -lactamase are brought together and can form an active complex able to cleave a fluorescent substrate<sup>305</sup>. A recent enzymatic-based approach for

screening inhibitors of protein aggregation is the *in vivo* colocalisation of  $\beta$ -lactamase and the POI to the periplasm. For example, a POI is sandwiched between  $\beta$ -lactamase and a periplasmic export signal protein, such as the twin arginine transport (Tat)-dependent export signal ssTorA<sup>301</sup>. The Tat signal peptide directs the complex to the twin arginine transport system which transports fully folded proteins into the periplasm<sup>306</sup>. If the POI does not fold correctly or aggregates in the cytoplasm, it cannot be transported into the periplasm where the  $\beta$ -lactamase enzyme is required to function (**Figure 1.15a**). Conversely, if the POI is prevented from aggregating by small molecule inhibitors, the tripartite fusion complex can be translocated into the periplasm, where  $\beta$ -lactamase confers antibiotic resistance (**Figure 1.15b**). This assay has been used to screen for aggregation inhibitors of the Alzheimer's disease-related A $\beta$  peptide<sup>301</sup>, however like the cytoplasmic-based screens, the inhibitors must still be permeable to the inner membrane of bacteria to be identified using this approach (**Figure 1.15c**).

Although all of the approaches described above have been successfully used to select protein aggregation inhibitors, a generally applicable method that can additionally screen the multitude of small molecules that cannot enter the cytoplasmic space has yet to be established.



**Figure 1.15.** Cytoplasmic enzyme activity screen for inhibitors of protein aggregation. The protein of interest (POI) (green) is sandwiched between a periplasmic enzyme (purple) and a tat signal peptide (blue) and expressed in the cytoplasm of bacteria. When the bacteria are grown in the presence of small molecules unable to prevent POI aggregation (blue molecules), the tripartite construct cannot translocate into the periplasm and enzyme activity is not reported (**a**). Conversely, if the molecules bind and prevent POI aggregation (green molecule), the tripartite construct folds correctly and translocates into the periplasm where the reporter enzyme is active (**b**). The molecules must, however, be able to traverse the inner membrane of the bacteria. If an inhibitor cannot cross this barrier a false-negative result will be observed (**c**).

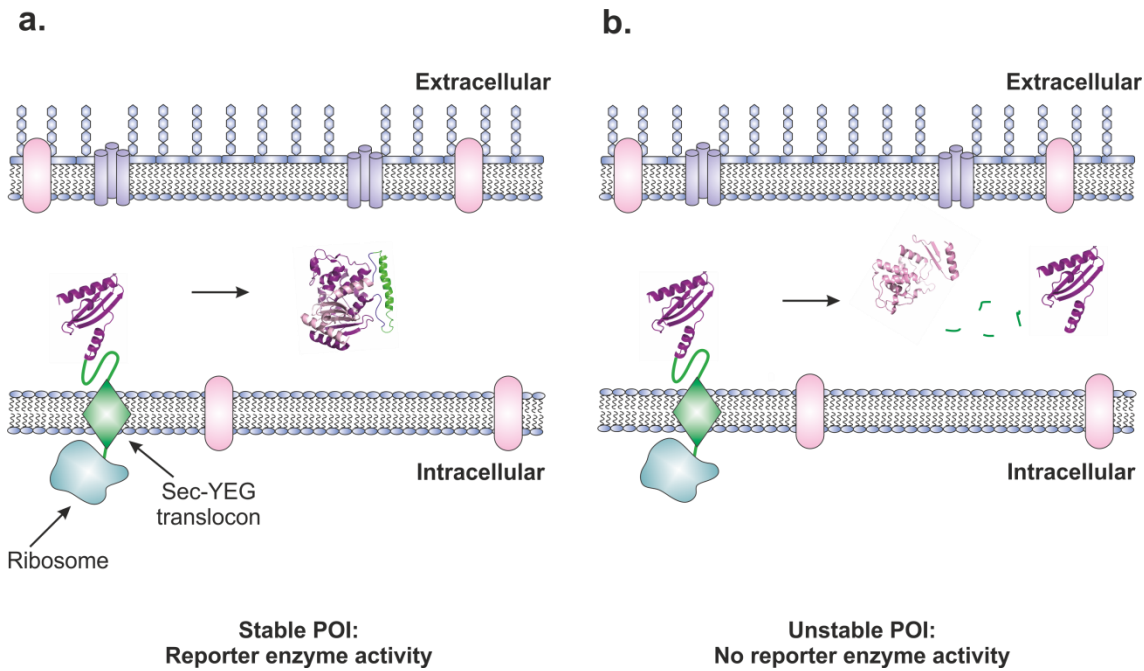
## 1.7 The periplasmic system for identifying protein aggregation inhibitors

In this work, a powerful and diversely applicable screening technique was developed for protein aggregation inhibitors. The assay is performed in the periplasm of *E. coli*, thereby avoiding the risk of false negative results produced through the limited access of small molecules to the target of interest in cytoplasmic screens.

### 1.7.1 Background to screen development

Due to the presence of outer membrane porins, the periplasm of *E. coli* (and other Gram negative bacteria) is permeable to molecules smaller than  $\sim 600$  Da<sup>307</sup>. This provides a significant advantage over cytoplasmic aggregation screens within eukaryotes<sup>300</sup> or the cytosol of prokaryotes<sup>299, 301-303</sup> as it includes the multitude of small molecules that can only enter the periplasmic space. Consequently, to overcome the disadvantage of screening for protein aggregation inhibitors within the cytoplasm, assaying for inhibitors of protein aggregation in the periplasmic space offers significant advantages.

An assay developed by Foit *et al.*<sup>308</sup> to analyse protein stability *in vivo* is different to previous  $\beta$ -lactamase reporter assays, as it takes advantage of the N-terminal 23 amino acid signal peptide of  $\beta$ -lactamase. This signal allows the unfolded  $\beta$ -lactamase polypeptide chain to be transported into the periplasm through the general secretory pathway. Unlike in the Tat system for protein translocation, in the general secretory pathway as soon as the N-terminus of the nascent polypeptide chain exits the ribosome, it associates with the cytosolic chaperones SecA and SecB<sup>306</sup>. These maintain  $\beta$ -lactamase in a partially unfolded, pre-protein state and target it to a multi-subunit Sec YEG translocon within the inner membrane<sup>306</sup>. The unfolded polypeptide chain is translocated in an ATP-dependent manner into the periplasm where the signal peptide is cleaved. The mature polypeptide is then able to fold (**Figure 1.16**).

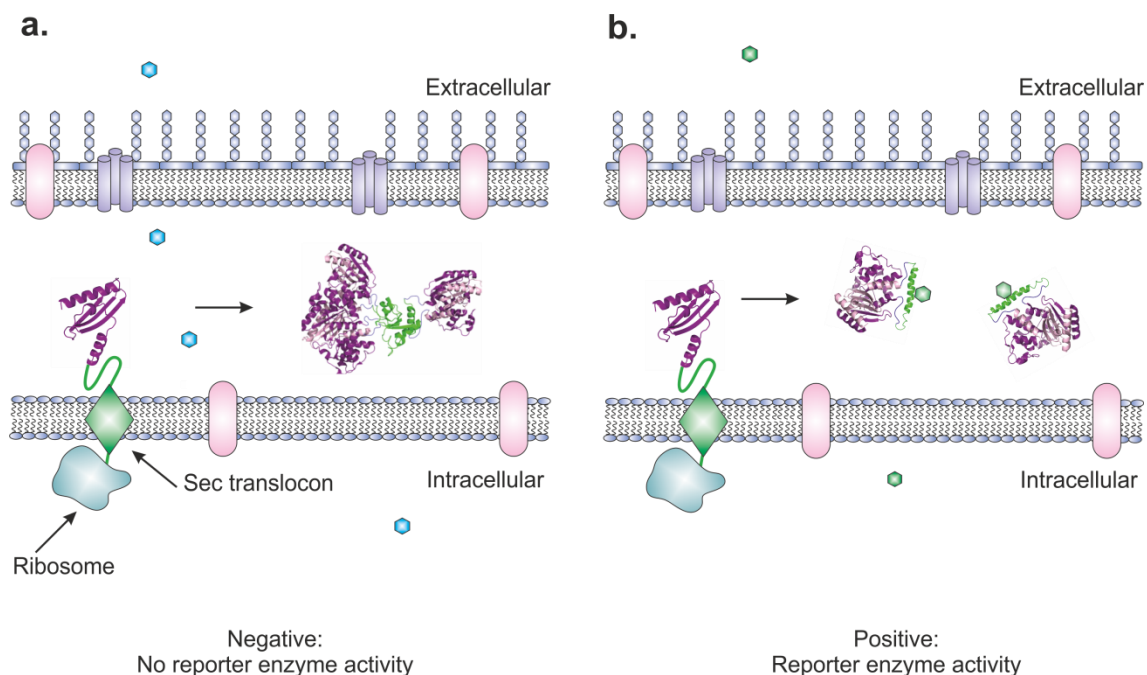


**Figure 1.16.**  $\beta$ -lactamase tripartite fusion system to monitor protein stability. The protein of interest (POI, green) is cloned between the two domains of the periplasmic enzyme  $\beta$ -lactamase (purple and pink). The unfolded polypeptide chain is translocated into the periplasm via the sec-YEG-translocon. **(a)** If the POI is stable, the two domains of  $\beta$ -lactamase associate and enzyme activity is achieved. **(b)** if the test protein does not fold correctly, or is unstable, it will be degraded by periplasmic proteases and enzyme activity is lost.

Foit *et al.*<sup>308</sup> exploited the known tolerance of the antibiotic resistance enzyme TEM-1  $\beta$ -lactamase to host the insertion of a test protein in a loop on its surface<sup>309-311</sup> in order to analyse protein stability *in vivo*. Upon correct folding of the test protein, the two halves of  $\beta$ -lactamase are brought close enough together to associate and the bacteria will be resistant to  $\beta$ -lactam antibiotics (**Figure 1.16a**). If, however, protein stability is compromised (for example, by mutation), the two halves of  $\beta$ -lactamase separate, either through proteolytic cleavage or sequestration by periplasmic chaperones, resulting in reduced resistance of the bacterium to  $\beta$ -lactam antibiotics (**Figure 1.16b**)<sup>308</sup>. One advantage of this system is that it does not require any knowledge of the protein's structure or function, and can in principle be performed with the primary sequence of the protein of interest as the only available information<sup>312</sup>. Furthermore, the oxidising milieu of the periplasm allows the formation of disulfide bonds, thereby providing another advantage over previous cytoplasmic assays.



This  $\beta$ -lactamase tripartite fusion system has been utilised to develop proteins with sequences that increase stability and expression yield<sup>308</sup>, and to identify chemical chaperones that aid protein folding in the periplasm<sup>313</sup>. The work described in this thesis aimed to determine whether this method could be further developed into a platform for screening for small molecule inhibitors of protein aggregation. **Figure 1.17** illustrates how this screen would function.



**Figure 1.17.** Principle of  $\beta$ -lactamase screen for inhibitors of protein aggregation. The protein of interest (POI) is cloned between the two domains of the periplasmic enzyme  $\beta$ -lactamase (purple and pink). The unfolded polypeptide chain is translocated into the periplasm via the sec translocon. (a) If the POI aggregates,  $\beta$ -lactamase enzyme activity is lost. (b) If, however, the POI is prevented from aggregating by small molecule inhibitors, the two domains of  $\beta$ -lactamase associate and enzyme activity is rescued.

## 1.8 Biopharmaceuticals

The subject of protein aggregation inhibition is not just limited to biological disorders. With the advent of commercial recombinant DNA technology in the mid-1970s it has become possible to produce vast quantities of proteins, from bacterial fermentation to mammalian cell cultures, to be used as therapeutic agents<sup>314, 315</sup>. Indeed, therapeutic proteins such as antibodies now constitute the most rapidly growing class of pharmaceuticals and are used as treatments in numerous clinical settings, from cancer treatment to kidney transplantation<sup>316</sup>. Unfortunately, the manufacturing process subjects the biologics to various environments and stresses that may modify their physical and chemical properties. These alterations increase the risk of protein misfolding, aggregation and degradation, which may compromise the quality, stability, and even safety of the drug product. The high concentrations required for many biologic drugs (such as monoclonal antibodies) also increases these risks, and as a consequence, control of protein aggregation during product development, storage, transportation, and patient administration is imperative.

### 1.8.1 History of biopharmaceuticals

Biopharmaceuticals are traditionally defined as pharmaceuticals with active agents biological in nature and manufactured using biotechnology<sup>317</sup>. The first part of this definition distinguishes biopharmaceuticals from small molecule drugs which are distinctively chemical in nature and manufactured using chemical methods<sup>317</sup>. Interpretation of the latter part of the definition is more complicated and depends on the country of origin or regulatory agency. Most regulatory agencies (including the United States of America Food and Drug Administration) have a 'broad view' of what constitutes a biopharmaceutical and classify them as anything manufactured by what is known as new technologies (e.g. monoclonal antibodies, and recombinant proteins) or old technologies (e.g. proteins and vaccines derived from non-engineered organisms as well as plasma-derived products)<sup>317, 318</sup>. Conversely, European Union regulations subscribe to the 'new' biotechnology view and define biopharmaceuticals as "a protein or nucleic acid-based pharmaceutical substance used for therapeutic or *in vivo* diagnostic purposes, which is produced by means other than direct extraction from a native (non-engineered) biological source"<sup>318</sup>. Europe consequently uses the broader term *biotechnology medicines* to denote "all pharmaceutical products produced in part or in full by biotechnological means, either

traditional or modern"<sup>319</sup>. Despite the disagreement over what constitutes a biopharmaceutical, it has become largely synonymous with recombinant or engineered versions of eukaryote proteins as almost all the biopharmaceuticals available today are proteins or peptides<sup>317</sup>. These therapeutic proteins have had a major impact on health care and, as such, over the past ten years the market for biopharmaceuticals such as monoclonal antibodies has grown exponentially<sup>320</sup>. These biopharmaceuticals are subject to the same physical laws as natural proteins, and hence protein aggregation has become one of the greatest obstacles in protein therapeutic research and development.

Based on the 'broad' view of what constitutes a biopharmaceutical, their use dates back over 200 years to 1798, when Edward Jenner developed the first vaccine for smallpox using live vaccinia virus from cows. By using a naturally weak form of the disease (cowpox), the severity and duration of illness was greatly reduced. This success was followed by Louis Pasteur's breakthroughs in the principles of vaccination less than 100 years later. Pasteur discovered that artificially weakening bacteria before infection lead to immunity. This discovery revolutionised the treatment of infections as a naturally weak form of the disease organism did not need to be found. Pasteur created rabies, anthrax and chicken cholera vaccinations, and gave them the generic name "vaccines" in honour of Edward Jenner. The dawn of bacteriology followed, in which large quantities of biopharmaceuticals could be produced, and by 1930 antitoxins and vaccines against tuberculosis, cholera, plague, typhoid, diphtheria and tetanus had been developed and were in mass production<sup>321</sup>.

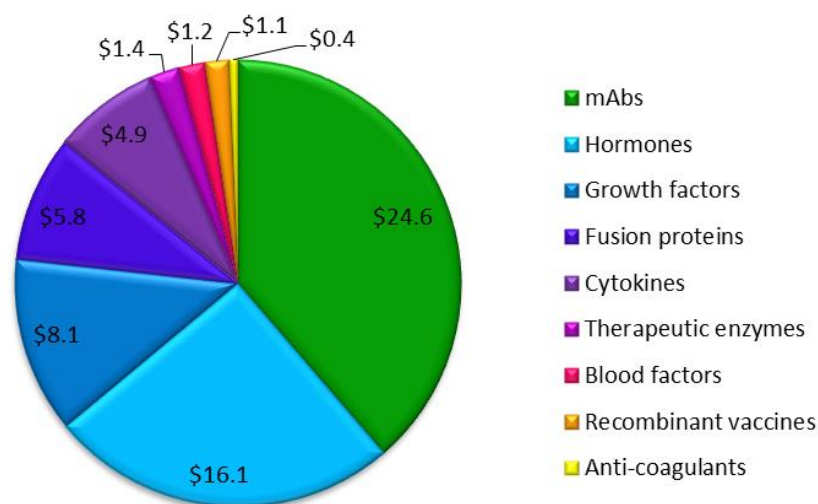
Production of biopharmaceuticals against non-bacterial diseases took longer to advance due to the complexity of isolating the constituents from animal sources. The first example of this was insulin. Its discovery and isolation from dogs in 1921 by Frederick Banting and Charles Best<sup>322, 323</sup> led to the treatment of the first human patient only a year later<sup>324</sup>. The mass production of purified insulin required enormous quantities of porcine and bovine pancreas, a waste product from the meat industry. Unfortunately the availability of animal pancreases was limited, the process time-consuming and costly, and the treatment itself was dogged by compatibility issues. The single amino acid difference between human and porcine insulin, and the three amino acid difference between human and bovine insulin, led to a significant number of patients developing immune reactions over time<sup>325</sup>. A number of animal serum-derived products were developed over the next fifty years, however, with the advent of recombinant DNA technology in the late 1970s biopharmaceutical production was revolutionised. The utilisation of bacteria to express human genes resulted in abundant, inexpensive and low immunogenic proteins free from

other animal substances. The first licensed drug generated using this technology was human insulin in 1982<sup>326, 327</sup>, followed quickly by human growth hormone<sup>317</sup>. The 30 years since their approval have seen an enormous expansion in the therapeutic applications of proteins, ranging from enzymes and regulatory proteins to protein diagnostics (see **Table 1.4** for select examples).

Functional Classification	Mechanism of Action	Therapeutic (Trade Name)	Function	Clinical Use
Protein therapeutics with enzymatic or regulatory activity	Replacing a protein that is deficient or abnormal	Insulin (Novolin) <sup>322, 326, 327</sup>	Regulates blood glucose	Diabetes mellitus
		Lactase (Lactaid) <sup>328</sup>	Recombinant enzyme that digests lactose	Lactose intolerance
	Augmenting an existing pathway	Follicle-stimulating hormone (Follistim) <sup>329</sup>	Augments ovulation	Assisted reproduction
		Factor VIIa (NovoSeven) <sup>330</sup>	Initiates the coagulation cascade	Haemorrhage in haemophilic patients
	Providing a novel function or activity	Collagenase (Santyl) <sup>331</sup>	Digests collagen in necrotic wounds	Severe ulcers and burns
		Deoxyribonuclease (Pulmozyme) <sup>332</sup>	Degrades DNA in purulent secretions	Respiratory tract infections (cystic fibrosis)
Protein therapeutics with specific targeting activity	Interfering with a molecule or organism	Adalimumab (Humira) <sup>333, 334</sup>	Humanised mAb that binds TNF $\alpha$	Rheumatoid arthritis
		Enfuvirtide (Fuzeon) <sup>335, 336</sup>	Inhibits HIV entry into host cells	Patients with advanced HIV infection
	Delivering other compounds or proteins	Gemtuzumab ozogamicin (Mylotarg) <sup>337, 338</sup>	Humanised mAb conjugated chemotherapeutic agent	Acute myeloid leukaemia
Protein vaccines	Protecting against a deleterious foreign agent	Hepatitis B surface antigen (Engerix)	Non-infectious protein on surface of hepatitis B virus	Hepatitis B vaccination
	Treating cancer	Ipilimumab (Yervoy) <sup>339, 340</sup>	Amplifies T cell response	Late-stage melanoma
Protein diagnostics	<i>In vivo</i> diagnostics	Growth hormone releasing hormone (Geref) <sup>341, 342</sup>	Stimulates growth hormone release	Diagnosis of defective growth hormone secretion
	Cancer imaging	Satumomab pendetide (OncoScint) <sup>343</sup>	Labelled mAb specific for tumour-associated glycoprotein	Colon and ovarian cancer detection

**Table 1.4.** Functional classification of protein therapeutics. Data from Leader, 2008<sup>325</sup>

The biologics market today is worth over \$50 billion<sup>344, 345</sup> and is growing rapidly. Monoclonal antibodies (mAbs) are ranked as the highest selling class of biopharmaceutical, with four out of the top ten selling biologics of 2012 - 2013 (in the United States of America) belonging to this category (**Figure 1.18**). Small numbers of protein therapeutics are, however, still purified from human blood due to economic reasons. For example human serum albumin (HSA) can be produced in large quantities relatively inexpensively by direct extraction at a price no recombinant HSA product is likely to be able to compete with<sup>346</sup>. Despite this, the vast majority of therapeutic biologics on the market today are from recombinant sources, requiring efficient, reliable and high throughput cell-based production processes<sup>314</sup>.



Drug	US 2012 Sales (\$ Billions)	Category	Clinical Use
Humira	4.6	mAb	Rheumatoid arthritis
Lantus	4.5	Hormone	Diabetes
Enbrel	3.9	Fusion protein	Rheumatoid arthritis and psoriasis
Remicade	3.6	mAb	Rheumatoid arthritis and psoriasis
Neulasta	3.5	Growth factor	Neutropenia in breast cancer patients
Rituxan	3.5	mAb	Blood cancer and rheumatoid arthritis
NovoLog	2.8	Hormone	Diabetes
Avastin	2.8	mAb	Various cancers
Humalog	2.1	Hormone	Diabetes
Epogen	1.9	Growth factor	Chemotherapy-related anaemia

**Figure 1.18.** Top nine categories of biologic drugs in terms of US sales 2012 - 2013. The chart shows US sales of these biologics in billions of US dollars. The table ranks the top ten selling drugs of 2012 - 2013 and their clinical use. Data obtained from Aggarwal<sup>347</sup>.

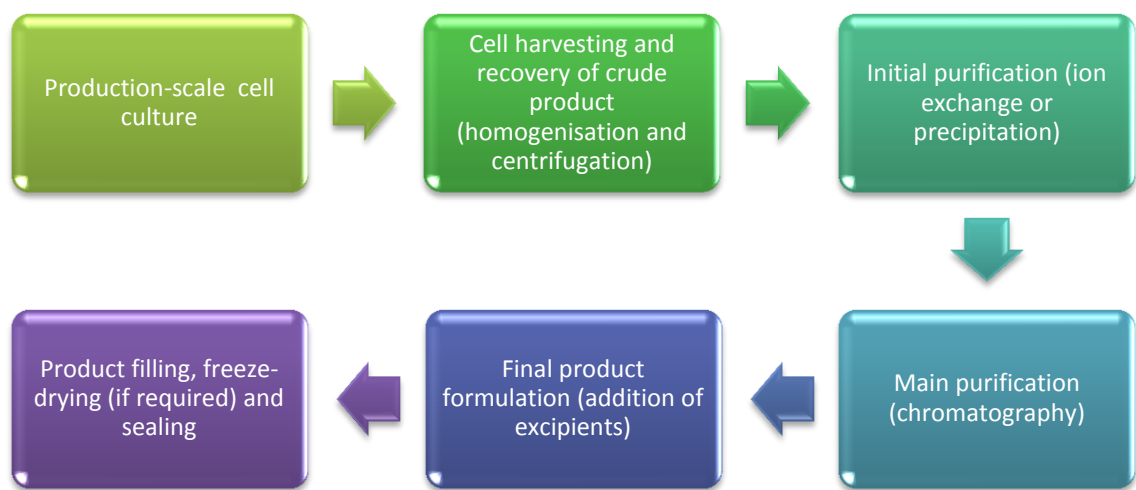
## 1.8.2 Protein biopharmaceutical production

The expression of recombinant proteins in cells in which they do not naturally occur is termed 'heterologous protein production'<sup>348</sup>. By far the most common microbial species used for recombinant protein production is *E. coli*, due to the high level of expression achieved and inexpensive fermentation technology required. A key feature of intracellular protein production is the inevitable accumulation of heterologous proteins in the cell cytoplasm as insoluble inclusion bodies. The separation of the inclusion bodies from homologous *E. coli* proteins is facilitated by their density; they sediment more rapidly than cellular debris during low-speed centrifugation<sup>348</sup>. The inclusion bodies are usually solubilised in strong denaturant (e.g. detergents, urea, solvents), after which the denaturant is removed by dialysis to enable refolding of the protein. The yield of natively-refolded protein depends on the biophysical properties of the protein and a significant portion of time is spent optimising this process to maximise yields. Numerous other production systems are used for recombinant human pharmaceutical proteins, depending on the characteristics of the protein to be produced. A comparison of these systems is given in **Table 1.5**.

Production System	Overall Cost	Production Timescale	Glycosylation	Comments
Bacteria	Low	Short	None	Rapid and cheap production
Yeast	Medium	Medium	Usually incorrect	Expression levels usually remain less than 5 % of total cellular protein
Fungi	Medium	Medium	Usually incorrect	High expression levels, many proteins are secreted extracellularly
Insect cell culture	High	Medium	Minor differences	Moderate expression levels and relatively rapid growth
Mammalian cell culture	High	Long	Correct	Efficient folding of mammalian proteins
Transgenic plants	Low	Long	Minor differences	Cheap production but low yields
Transgenic animals	High	Very long	Correct	Biologic produced in milk of transgenic animal enables easy harvesting

**Table 1.5** Comparison of production systems for recombinant biopharmaceuticals. Information from Walsh<sup>348</sup> and Ma *et al.*<sup>349</sup>.

The downstream processing of the protein biologic involves the recovery, purification, formulation and finish of the product (**Figure 1.19**). The biopharmaceutical must pass through each of these processing steps before the finished product is ready to enter the market. These steps, including freezing, thawing, formulation, filtration, filling, freeze-drying and inspection, are known as "formulation and fill-finish operations". Formulation into the final drug product usually involves the addition of various excipients to stabilise the product, filtration of the product through a 0.22  $\mu\text{m}$  filter for sterilisation, and lyophilisation if the product is to be marketed in a powder format<sup>348</sup>. Each step poses challenges for protein stability which must be rigorously addressed.



**Figure 1.19** Overview of the production process for protein biopharmaceuticals. Information from Walsh<sup>348</sup>.

### 1.8.3 Biopharmaceutical stability and aggregation

The manufacture of biopharmaceutical products is one of the most highly regulated and rigorously controlled manufacturing processes<sup>350</sup>. Owing to the biological origin and macromolecular structure of biologics (compared to small molecule drugs), much attention is focused on preventing contamination of biologics with other biological impurities (such as viruses), as well as conformational changes introduced during the production process<sup>351</sup>. As the therapeutic activity of protein biologics is highly dependent on their conformational structure, changes in this structure caused by external conditions throughout the production and formulation process must be minimised. The product

must not only survive the manufacturing process undamaged and biologically active, it must be stable enough in a given packaging system for its entire shelf life (an 18-month shelf life is usually the minimum requirement for protein-based drugs<sup>352</sup>). Furthermore, drugs destined for clinical trials require accurate quantification of protein aggregates throughout the manufacturing process to show they meet drug product specifications<sup>35</sup>.

Each step in the purification and formulation of protein biologics poses challenges for protein stability. Almost every conceivable environmental factor (e.g. temperature, light, water, pH, shear, detergents, freeze-thawing, freeze-drying, presence of glass, rubber, plastic, or presence of salts and other solutes) can adversely affect protein drugs<sup>353</sup>. For example, freezing samples minimises the risk of microbial growth and eliminates agitation and therefore diffusive collisions during transportation. However, it also introduces complex physical and chemical changes. A matrix of protein and ice can result in protein adsorption on the ice surface and the alteration of protein conformation near the ice surface due to weakening of hydrophobic bonds<sup>354</sup>. Freezing can also change the pH of the solution through selective precipitation of buffer components. This can have a strong influence on the aggregation rate as the pH determines the electrostatic interactions through charge distribution on the protein's surface<sup>355</sup>. Minute changes in conditions can also lead to aggregation. Small single domain proteins often require significant environmental changes to destabilise, however large multidomain proteins can contain 'weak links' between domains that unravel under even the mildest of conditions<sup>351</sup>.

After the drug has been formulated, it must go through filter sterilisation. This can introduce the protein to stress whilst it is pushed through the filter, risking alteration of the protein structure. Furthermore, stabilising components could get adsorbed onto the membrane surface, resulting in a decrease in their concentration and subsequent protein instability<sup>355</sup>. Once the drug product is ready, it is filled into the primary drug product containers. Much care is taken to limit the shear the drugs are subjected to, and also any interactions with, or leachables from, the container surface and components. Finally, to achieve an acceptable shelf life, the drugs are often lyophilised. This prevents most covalent degradation and also minimises diffusion, thereby decreasing the risk of aggregation. Although these general techniques are understood to reduce the protein's propensity to aggregate, each and every protein biopharmaceutical is unique and each step in the manufacturing process must be individually tailored to suit their needs.



## **1.8.4 Techniques employed to prevent biopharmaceutical aggregation**

Most of the antibody-based biopharmaceuticals entering clinical trials are derived from phage display technology<sup>356, 357</sup> or transgenic mice<sup>358</sup>. Lead optimisation during phage display involves cycles of selection to generate high affinity binders to a target antigen. Antigen affinity is improved through directed evolution and most often results in the accumulation of mutations in the complementary determining regions (CDRs) of the antibody fragments. These mutations may confer improved binding through increased hydrogen bonding, electrostatics or even hydrophobic interactions with the target antigen. Consequently, after isolation of the antibody fragments with the highest affinity, they need to be optimised for properties such as high stability, solubility and minimal aggregation propensity, particularly if the affinity is increased through greater hydrophobic interactions as this may result in patches of the antibody fragment that are prone to self-associate.

### **1.8.4.1 Protein engineering**

Historically, lead molecules were developed with very little pre-assessment into stability and aggregation, consequently resulting in significant issues further down the process. More recently, high throughput protein engineering techniques are utilised to dramatically improve protein aggregation.

*In silico* computational techniques are used to predict protein aggregation<sup>359, 360</sup> and are useful for predicting and preventing problems in protein production. For example, aggregation prediction models have been used to predict aggregation rates and the effect of different mutations, depending on hydrophobicity, patterns of alternating hydrophobic-hydrophilic residues, pH and ionic strength<sup>351, 361 362, 363</sup>. By identifying aggregation hot spots, key 'gatekeeper' residues can be incorporated to reduce the aggregation propensity and maintain the molecule into the monomeric form<sup>361</sup>. Nevertheless, it remains a challenge to incorporate all the environmental factors that impinge protein stability during the manufacturing process, and as such, the complexity of the mechanisms leading to the aggregation has prevented other parameters, such as stirring, to as yet be incorporated into these models<sup>351</sup>.

Conversely, structure-based engineering is utilised to stabilise fragile or aggregation-prone regions of a protein by either inserting extra stabilising interactions, such as salt bridges<sup>364</sup> or disulfide bonds<sup>365</sup>, or reducing unfavourable interactions such as charged groups being buried in hydrophobic regions<sup>366</sup>. The new constructs are most often purified and assessed using *in vitro* biophysical techniques. Optimisation of thermal stability has largely benefited from the high throughput screening techniques such as differential scanning fluorimetry<sup>367</sup>. However, a similar high throughput approach has not yet been established for screening mutations affecting aggregation<sup>368</sup>. In these instances, low throughput HPLC technology is foremost utilised for quantifying protein aggregation<sup>368</sup>.

#### **1.8.4.2 Formulation screening**

Mutagenesis to reduce aggregation can be problematic, as it requires detailed analysis to certify no loss of function has occurred, and also extensive clinical trials to ensure no adverse effects develop. Consequently, a complementary approach is to alter the formulation of the drug product. Formulation of the biologic involves the addition of excipients (pharmacologically inactive carriers) to protect the product during the fill-finish steps of the production process. Initially, molecular chaperones from bacteria (such as DnaK-DnaJ-GrpE and GroEL-GroES) were used to inhibit protein aggregation during refolding, by alternately binding and releasing folding intermediates<sup>352</sup>. However, the cost of production, and the extra purification steps required, led to the exploration for more cost-effective additives such as stabilisers and cryoprotectants<sup>354</sup>. Stabilisers such as polyols, polyethylene glycols (PEGs) and other polymers are often used to sterically hinder protein-protein interactions and limit diffusion. As many therapeutic proteins are stored in very high doses, a key role for the excipients is to keep the biologics stably dissolved for extended periods of time at concentrations of tens of mg per mL<sup>351</sup>. Cyclodextrans, circular polymers of typically five to seven glucose rings, can also significantly reduce aggregation of therapeutic proteins, such as insulin<sup>351, 369-371</sup> and growth hormone<sup>372</sup>. They function by binding to aromatic residues and stabilising the unfolded state<sup>351, 373</sup>, however its ability to extract biomembrane components has limited its use to only certain modified forms<sup>351</sup>. Stabilisers that are also effective cryoprotectants include disaccharides such as sucrose and trehalose<sup>374</sup>. Unlike some stabilisers that readily form crystals when frozen, these cryoprotectants remain in the same phase as the protein and are therefore able to exert their stabilising effects.

Small-scale characterisation studies are often conducted to evaluate protein stability under formulation conditions, however this typically involves time-consuming incubation of proteins under various conditions during long-term stability studies. Although numerous approaches for high-throughput techniques for formulation screening have been developed<sup>354</sup>, these still require purified protein.

## **1.9 An *in vivo* platform for assessing protein aggregation and identifying inhibitors**

Utilising phage display, it is possible to obtain very large collections of antibodies for a particular target (such as the more than 1000-member group of antibodies that are specific for B-lymphocyte stimulator<sup>375</sup>), however *in vitro* analyses for aggregation propensity is both costly and time-consuming. For these reasons, the  $\beta$ -lactamase tripartite fusion system (described in Section 0) was also applied in this thesis to the assessment of biopharmaceutical aggregation. By enabling *in vivo* identification of aggregation-prone sequences, *in vitro* purification and biophysical analysis would not be required for a large selection pool of proteins. This would significantly decrease both the cost and time required for biopharmaceutical lead development; a major concern for the biopharmaceutical industry.

To this end, the *in vivo* tripartite  $\beta$ -lactamase system was also utilised in this thesis to identify excipients that prevent protein aggregation, demonstrating the diverse applications of the periplasmic-based enzyme assay for analysing and inhibiting aggregation relevant to both disease and biopharmaceutical aggregation.

## 1.10 Aims of the current work

The tendency of proteins to aggregate has severe consequences in both disease and biopharmaceutical production. Current screening techniques utilised in the search for inhibitors to prevent such aggregation are costly and time consuming, and require large quantities of these difficult-to-purify proteins.

There is an urgent need for the rapid generation of a robust, sensitive, and cost-effective assay for inhibitors of protein aggregation. Therefore, the aim of this study is the development of an *in vivo*, periplasmic-based screen for inhibitors of protein aggregation, to be applicable to both disease-related and biopharmaceutically-relevant proteins.

The system should:

- quantitatively measure the extent of the aggregation of different constructs *in vivo*
- enable the selection of small molecule inhibitors of protein aggregation
- allow identification of excipients that can prevent protein aggregation
- be applicable to a variety of proteins of different size, structure, function and aggregation propensity
- have a low false positive and false negative rate (< 5 - 10 %)
- be amenable to a diverse range of experimental conditions to suit each protein's requirements

These requirements are all been met using the  $\beta$ -lactamase tripartite fusion system for identifying protein aggregation inhibitors. In Chapter 3 the approach quantitatively correlates the extent of aggregation to a simple and easily selectable readout: bacterial growth in the presence of increasing concentrations of antibiotic. Chapter 4 describes the development of the system to differentiate between inhibitors and non-inhibitors of protein aggregation, with particular emphasis placed on assay optimisation. Finally, Chapter 5 highlights the application of the tripartite fusion system to biopharmaceutically relevant proteins and excipients, thereby demonstrating the diverse applications of the *in vivo* assay to both academic and industrial avenues of research.

## 2 Materials and methods

### 2.1 Materials

#### 2.1.1 Technical equipment

##### Equipment

##### Manufacturer

##### Centrifuges

Avanti J-26 XP Centrifuge	Beckman Coulter, Brea, CA, USA
GenFuge 24D Centrifuge	Progen Scientific, London, UK
MiniSpin plus F-45-12-11	Eppendorf, Hauppauge, NY, USA
Contifuge Stratos Continuous-Flow Rotor	Heraeus, Hanau, Germany

##### Incubators, mixers & shakers

Gallenkamp Economy Incubator Size 1	Sanyo, Watford, UK
ORBISAFE Orbital Incubator	Sanyo, Watford, UK
Stuart Magnetic Stirrer SB161	CamLab, Cambridge, UK
Stuart Orbital Incubator S150	Bibby Scientific, Stone, UK
Stuart Vortex Mixer SA8	Bibby Scientific, Stone, UK

##### Gel electrophoresis equipment

Vari-Gel midi system	CamLab, Madingley, UK
Slab Gel Electrophoresis Chamber AE-6200	ATTO, Tokyo, Japan
Standard Power Pack P25	Biometra, Goettingen, Germany

##### Protein purification equipment

ÄKTAprime plus	GE healthcare, Little Chalfont, UK
HiTrap Q HP 5 mL anion exchange column	GE healthcare, Little Chalfont, UK
HiTrap SP HP 5 mL cation exchange column	GE healthcare, Little Chalfont, UK
Superdex™ 75 GL 10/300 gel filtration column	GE healthcare, Little Chalfont, UK
Superdex™ 75 HR 10/30 gel filtration column	GE healthcare, Little Chalfont, UK
HiLoad Superdex™ 75 26/60 gel filtration column	GE healthcare, Little Chalfont, UK

**Fluorometer**

Photon Technology International fluorometer Ford, West Sussex, UK

**Spectrophotometer**

UltraSpec 2100 pro UV/Visible Spectrophotometer GE healthcare, Little Chalfont, UK

**Microplate readers**

FLUOstar OPTIMA plate reader BMG Labtech, Aylesbury, Bucks, UK  
NEPHELOstar Galaxy laser-based nephelometer BMG Labtech, Aylesbury, Bucks, UK

**Electron microscope**

JEOL JEM-1400 transmission electron microscope JEOL Ltd., Tokyo, Japan

**Mass spectrometer**

Synapt high definition mass spectrometry (HDMS) quadrupole-time-of-flight mass spectrometer (Micromass UK Ltd., Waters Corpn., Manchester, UK), equipped with a Triversa automated nano-electrospray ionisation (ESI) interface (Advion Biosciences, Ithaca, USA).

**Other equipment**

0.5 – 10  $\mu$ L StarPet E 12-channel electronic pipette STARLAB, Milton Keynes, UK  
100 – 1200  $\mu$ L StarPet E 8-channel electronic pipette STARLAB, Milton Keynes, UK  
Büchi Vac V-500 Vacuum Pump Sigma Life Sciences, St. Louis, USA  
Corning Costar 3915 96-well plate Corning Life Sci, The Netherlands  
Grant JB1 Unstirred Waterbath Grant Instruments, Shepreth, UK  
InGenius Gel Documentation System Syngene, Cambridge, UK  
Jenway 3020 Bench pH Meter Bibby Scientific, Stone, UK  
PTC-100 Programmable Thermal Controller GMI-Inc, Minneapolis, MN, USA  
Series 2100 Media Autoclave Prestige Medical, Minworth, UK  
SnakeSkin Pleated Dialysis Tubing; 3,500 MWCO Thermo Scientific, Surrey, UK  
Techne Dri-Block Heater DB-2A Bibby Scientific, Stone, UK  
BIO-RAD T100 Thermocycler BIORAD Laboratories, Hercules, USA

## 2.1.2 Chemicals

### A

Acetic acid, glacial	Fisher Scientific, Loughborough, UK
Acid fuchsin	Fisher Scientific, Loughborough, UK
Acridine orange	Fisher Scientific, Loughborough, UK
Acrylamide, 30%	Severn Biotech, Kidderminster, UK
Agar	Melford Laboratories, Suffolk, UK
Agarose	Melford Laboratories, Suffolk, UK
Ampicillin	Formedium, Norfolk, UK
Arabinose	Sigma Life Sciences, St. Louis, USA
D-arginine	Sigma Life Sciences, St. Louis, USA
L-arginine	Sigma Life Sciences, St. Louis, USA
Aspirin	Sigma Life Sciences, St. Louis, USA
Azure A	Sigma Life Sciences, St. Louis, USA
Azure C	Alfa Aesar, Heysham, UK

### B

1 <i>H</i> benzimidazole-2-sulfonic acid	Sigma Life Sciences, St. Louis, USA
$\beta$ -mercaptoethanol	Stratagene, Cambridge, UK
Benzimidazole	Sigma Life Sciences, St. Louis, USA
Bromophenol blue	Sigma Life Sciences, St. Louis, USA

### C

Caffeic acid	Sigma Life Sciences, St. Louis, USA
Chloramphenicol	Sigma Life Sciences, St. Louis, USA
Congo red	Sigma Life Sciences, St. Louis, USA
Curcumin	Sigma Life Sciences, St. Louis, USA

### D

Dimethyl sulphoxide, DMSO	Sigma Life Sciences, St. Louis, USA
1,2-Dithiothreitol, DTT	Formedium, Norfolk, UK

### E

(-)-Epigallocatechin-3-gallate (EGCG)	Sigma Life Sciences, St. Louis, USA
Ethanol	Fisher Scientific, Loughborough, UK
Ethidium bromide, EtBr	Sigma Life Sciences, St. Louis, USA
Ethylenediamine tetra acetic acid, EDTA	Sigma Life Sciences, St. Louis, USA

### F

Fast Green FCF	Fisher Scientific, Loughborough, UK
----------------	-------------------------------------

**G**

Glucose	Fisher Scientific, Loughborough, UK
L-glutamate	Sigma Life Sciences, St. Louis, USA
Glycerol	Fisher Scientific, Loughborough, UK

**H**

Hemin	Sigma Life Sciences, St. Louis, USA
Hydrochloric acid, HCl	Fisher Scientific, Loughborough, UK

**I**

Instant Blue Stain	Expedeon, Harston, UK
Isopropyl $\beta$ -D-1-thiogalactopyranoside, IPTG	Melford Laboratories, Suffolk, UK

**K**

Kanamycin	Formedium, Norfolk, UK
-----------	------------------------

**L**

LB broth, granulated	Melford Laboratories, Suffolk, UK
----------------------	-----------------------------------

**M**

Magnesium sulphate, $MgSO_4$	Fisher Scientific, Loughborough, UK
Magnesium chloride, $MgCl_2$	Sigma Life Sciences, St. Louis, USA
Morin hydrate	Sigma Life Sciences, St. Louis, USA
Myricetin	Sigma Life Sciences, St. Louis, USA

**N**

Nitrocefin	Calbiochem, CA, USA
------------	---------------------

**O**

Orange G	Sigma Life Sciences, St. Louis, USA
Orcein	Sigma Life Sciences, St. Louis, USA

**P**

Phenol red	Fisher Scientific, Loughborough, UK
Polymyxin B sulfate	Sigma Life Sciences, St. Louis, USA
Potassium phosphate dibasic, $K_2HPO_4$	Sigma Life Sciences, St. Louis, USA
Potassium phosphate monobasic, $KH_2PO_4$	Sigma Life Sciences, St. Louis, USA

**R**

Resveratrol	Santa Cruz Biotech, Middlesex, UK
-------------	-----------------------------------



**S**

Silibinin	Sigma Life Sciences, St. Louis, USA
Sodium chloride, NaCl	Fisher Scientific, Loughborough, UK
Sodium dodecyl sulphate, SDS	Sigma Life Sciences, St. Louis, USA
Sodium hydroxide, NaOH	Fisher Scientific, Loughborough, UK
Sodium phosphate dibasic, Na <sub>2</sub> HPO <sub>4</sub>	Sigma Life Sciences, St. Louis, USA
Sodium phosphate dibasic, NaH <sub>2</sub> PO <sub>4</sub>	Sigma Life Sciences, St. Louis, USA
D-Sorbitol	Sigma Life Sciences, St. Louis, USA
Sucrose	Fisher Scientific, Loughborough, UK
Super Optimal broth, granulated	Merck, Darmstadt, Germany

**T**

Tetracycline	Formedium, Norfolk, UK
Tetramethylethylenediamine (TEMED)	Sigma Life Sciences, St. Louis, USA
Thiabendazole	Sigma Life Sciences, St. Louis, USA
Thioflavin T	
Tramiprosate	Santa Cruz Biotech, Middlesex, UK
Tris-(hydroxymethyl)-aminomethane, Tris	Melford Laboratories, Suffolk, UK
Tryptone	Melford Laboratories, Suffolk, UK

**U**

Uranyl acetate	Sigma Life Sciences, St. Louis, USA
----------------	-------------------------------------

**Y**

Yeast extract	Melford Laboratories, Suffolk, UK
---------------	-----------------------------------

### 2.1.3 Antibiotics, markers and dyes

Antibiotic	Solvent	Stock Solution (mg/mL)	Working Concentration ( $\mu\text{g/mL}$ )	Sterilisation
Ampicillin	Purite 18 M $\Omega$ H <sub>2</sub> O	100	100	Filter-sterilised through 0.2 $\mu\text{m}$ filter
Kanamycin	Purite 18 M $\Omega$ H <sub>2</sub> O	20	50	
Tetracycline	70 % (v/v) ethanol	5	12.5	

**Table 2.1.** Antibiotics used in this study.

Marker	Company
Mark12™ Protein Standard	Invitrogen, Paisley, UK
Precision Plus Protein™ Dual Xtra Standards	Bio-Rad, Hemel Hempstead, UK
100 bp DNA Ladder	Promega, Southampton, UK
1 kb DNA Ladder	Promega, Southampton, UK

**Table 2.2.** DNA and protein markers used in this study.

Dye	Company
Gel Loading Dye, Blue (6x) for DNA	Promega, Southampton, UK
Orange G Loading Dye (10x) for DNA	Sigma Life Sciences, St. Louis, MO, USA
Instant Blue Stain	Expedeon Protein Solutions, UK

**Table 2.3.** Dyes used in this study.

### 2.1.4 Kits

Kit	Company
QIAprep Spin Miniprep Kit	QIAGEN, Crawley, UK
QIAquick Gel Extraction Kit	QIAGEN, Crawley, UK
QIAquick PCR Purification Kit	QIAGEN, Crawley, UK
QuikChange II Mutagenesis Kit	Agilent Technologies, Wokingham, UK

**Table 2.4.** Kits used in this study.

### 2.1.5 Enzymes for molecular biology

Enzyme	Company
Antarctic phosphatase (5 U/ $\mu$ L)	New England Biolabs, Hitchin, UK
<i>Bam</i> HI restriction endonuclease (20 U/ $\mu$ L)	New England Biolabs, Hitchin, UK
<i>Bam</i> HI-HF restriction endonuclease (20 U/ $\mu$ L)	New England Biolabs, Hitchin, UK
<i>Nde</i> I restriction endonuclease (20 U/ $\mu$ L)	New England Biolabs, Hitchin, UK
T4 DNA ligase (Quick Ligation™ Kit)	New England Biolabs, Hitchin, UK
Taq DNA polymerase (5 U/ $\mu$ L)	New England Biolabs, Hitchin, UK
Vent DNA polymerase (2 U/ $\mu$ L)	New England Biolabs, Hitchin, UK
<i>Xho</i> I restriction endonuclease (20 U/ $\mu$ L)	New England Biolabs, Hitchin, UK
Lysozyme	Sigma Life Sciences, St. Louis, MO, USA

**Table 2.5.** Enzymes for molecular biology used in this study. The buffers and additives of the corresponding enzymes, supplied by the manufacturer stated, were used.

### 2.1.6 Buffers

Buffer	Reagents
2× reducing SDS loading buffer	50 mM Tris-HCl pH 6.8
	100 mM DTT
	2 % (w/v) SDS
	0.1 % (w/v) bromophenol blue
	10 % (v/v) glycerol
Tris-acetate-EDTA (TAE) buffer	40 mM Tris-HCl, pH 8
	20 mM acetic acid (glacial)
	1 mM EDTA, pH 7.5
Tris-EDTA (TE buffer)	10 mM Tris-HCl, pH 8
	1 mM EDTA, pH 8
Electrophoresis cathode buffer	200 mM Tris-HCl, pH 8.25
	200 mM tricine
	0.2 % (w/v) SDS
Electrophoresis anode buffer	400 mM Tris-HCl, pH 8.8
Polymyxin buffer	50 mM Tris-HCl, pH 7.5
	5 mM EDTA
	50 mM NaCl
	1 mg/mL polymyxin B sulphate

**Table 2.6.** Buffers used in this study.

## 2.1.7 Media

Media	Reagent	Weight / Volume
Luria-Bertani (LB) medium	Bacto-tryptone	10 g
	Yeast extract	5 g
	NaCl	10 g
	Purite 18 MΩ H <sub>2</sub> O	Make to 1 L
	Autoclave	20 min at 121 °C, 15 psi
SOC medium	Bacto-tryptone	2 g
	Yeast extract	0.5 g
	1 M NaCl	1 mL
	1 M KCl	0.25 mL
	Purite 18 MΩ H <sub>2</sub> O	98 mL
	Autoclave	20 min at 121 °C, 15 psi
	1 M Mg <sup>2+</sup> stock (filter-sterilised through 0.2 µm filter)	1 mL
	1 M Glucose (filter-sterilised through 0.2 µm filter)	1 mL
Terrific Broth (TB) medium	Bacto-tryptone	12 g
	Yeast extract	24 g
	Glycerol	4 mL
	KH <sub>2</sub> PO <sub>4</sub> (final concentration of 17 mM)	2.31 g
	K <sub>2</sub> HPO <sub>4</sub> (final concentration of 72 mM)	12.54 g
	Purite 18 MΩ H <sub>2</sub> O	Make to 1 L
	Autoclave	20 min at 121 °C, 15 psi

**Table 2.7.** Media used in this study.

Solid medium was prepared by adding 1.5 % (*w/v*) agar (Melford laboratories, UK) to the liquid medium prior to autoclaving. Filter-sterilised antibiotics were added to media once cooled to < 50 °C.

## 2.2 Molecular biology methods

### 2.2.1 Bacterial strains

***E. coli* XL1-Blue** (Stratagene, Cambridge, UK)

*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI<sup>q</sup>ZΔM15 Tn10* (Tetr)]

***E. coli* SCS1** (Stratagene, Cambridge, UK)

*recA1 endA1 gyrA96 thi-1 hsdR17* (r<sub>K</sub><sup>-</sup> m<sub>K</sub><sup>+</sup>) *supE44 relA1*

***E. coli* DH5α** (Invitrogen, Paisley, UK)

F<sup>-</sup> Φ80*lacZΔM15 Δ(lacZYA-argF)* U169 *recA1 endA1 hsdR17* (r<sub>K</sub><sup>-</sup>, m<sub>K</sub><sup>+</sup>) *phoA supE44 λ-thi-1 gyrA96 relA1*

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

e14<sup>-</sup>(McrA<sup>-</sup>) *recA1 endA1 gyrA96 thi-1 hsdR17* (r<sub>K</sub><sup>-</sup>m<sub>K</sub><sup>+</sup>) *supE44 relA1 Δ(lac-proAB)* [F' *traD36 proAB lacI<sup>q</sup>ZΔM15*]

***E. coli* BL21 (DE3)** (Stratagene, Cambridge, UK)

F<sup>-</sup> *dcm ompT hsdS*(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) *gal λ*(DE3)

***E. coli* BL21 (DE3) pLysS** (Stratagene, Cambridge, UK)

F<sup>-</sup> *dcm ompT hsdS*(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) *gal λ*(DE3) [pLysS Cam<sup>r</sup>]

### 2.2.2 Polymerase chain reaction

The polymerase chain reaction (PCR) was performed to selectively amplify modified DNA sequences *in vitro*. The sequences and purpose of the oligonucleotide primers designed to amplify the desired genes from select plasmids are shown in (Table 2.8).

Primer	Sequence	Use
IAPP Forward	CGC ATT ACT TGT <b>CTC GAG</b> AAA ATG CAA CAC CGC GAC	Addition of <i>XhoI</i> restriction site 5' of hIAPP gene to clone it into $\beta$ -lactamase linker
IAPP Reverse	CGC ATT ACT GTA <b>GGA TCC</b> ATA GGT GTT GCT GCC CAC	Addition of <i>BamH1</i> restriction site 3' of hIAPP gene to clone it into $\beta$ -lactamase linker
A $\beta$ 42 Forward	CGC ATT ACT TGT <b>CTC GAG</b> A GAT GCG GAG TTC CGT CAT G	Addition of <i>XhoI</i> restriction site 5' of A $\beta$ 42 gene to clone it into $\beta$ -lactamase linker
A $\beta$ 42 Reverse	CGC ATT TCT GTA <b>GGA TCC</b> CGC TAT GAC AAC ACC ACC	Addition of <i>BamH1</i> restriction site 3' of A $\beta$ 42 gene to clone it into $\beta$ -lactamase linker
A $\beta$ 40 Forward	GGA TCC GGG AGC GGT TCC	Convert $\beta$ -lactamase-A $\beta$ 42 gene into $\beta$ -lactamase-A $\beta$ 40
A $\beta$ 40 Reverse	GAC AAC ACC ACC CAC CAT G	Convert $\beta$ -lactamase-A $\beta$ 42 gene into $\beta$ -lactamase-A $\beta$ 40
HEL4/Dp47d Forward	CGC ATT ACT TGT <b>CTC GAG</b> A GAA GTG CAG CTG CTG GAA AGC	Addition of <i>XhoI</i> restriction site 5' of HEL4 or Dp47d gene to clone it into the $\beta$ -lactamase linker
HEL4/Dp47d Reverse	CGC ATT AAT ATA <b>GGA TCC</b> GCT GCT CAC GGT CAC CAG	Addition of <i>BamH1</i> restriction site 3' of HEL4 or Dp47d gene to clone it into the $\beta$ -lactamase linker
$\beta$ -lactamase Linker Forward	CGG AGC TGA ATG AAG CCA TAC C	Sequence the linker region of $\beta$ -lactamase to ensure correct insertion of guest protein
$\beta$ -lactamase Linker Reverse	TCA CCG GCT CCA GAT TTA TCA GC	

**Table 2.8.** Oligonucleotide primers used in this thesis. The restriction enzyme recognition sites are highlighted in blue (*XhoI*) and red (*BamH1*).

A typical polymerase chain reaction contained the following components:

dsDNA template	100 ng
Primers	100 pmol
dNTPs	0.25 mM
DMSO	1 $\mu$ L
MgSO <sub>4</sub>	2, 4 or 6 mM
Vent DNA polymerase (2000 U mL <sup>-1</sup> )	1 U
Vent DNA polymerase buffer	1x
Nuclease-free deionised H <sub>2</sub> O	to 100 $\mu$ L

A solution lacking the dsDNA template was used as a negative control to determine if any non-specific interactions occur between the primers.

The theoretical melting temperature ( $T_m$ ) of the primers was calculated from **Equation 2.1**, where  $n_{AT}$  corresponds to the number of AT nucleotide base pairs, and  $n_{GC}$

corresponds to the number of GC nucleotide base pairs. The temperature cycle for a typical reaction is shown in **Table 2.9**.

$$t_m = (n_{AT} \times 2) + (n_{GC} \times 4) \quad \text{Equation 2.1}$$

Step	Temperature (°C)	Time (s)
Initial denaturation	95	300
Denaturation	95	30
Annealing	5 below $T_m$	30
Elongation	72	90
Repeat denaturation, annealing and elongation (x 29)		
Final elongation	72	300

**Table 2.9.** Temperature cycle for a typical PCR reaction.

The PCR products were visualised by agarose gel electrophoresis (Section 2.2.3) and excised from the gel using a scalpel. DNA extraction from the gel was performed using the QIAquick Gel Extraction Kit (QIAGEN, UK) as described in the manufacturer's instructions.

### 2.2.3 Agarose gel electrophoresis

Agarose gel electrophoresis was carried out in Tris-acetate-EDTA (TAE) buffer (**Table 2.6**). Gels were made by dissolving 1.5 % (w/v) agarose in 1 x TAE buffer and heating the solution using a microwave until the agarose had dissolved fully. Once cooled to < 50 °C, 0.5 µg/mL of ethidium bromide was added and the solution mixed. The gel was then poured into a 12 x 15 cm gel tray with a comb and allowed to set before use. DNA samples were diluted in 6 x blue gel loading buffer or 10 x orange G loading buffer. 30 µL samples were loaded, and lanes containing 5 µL of 1 kb and 100 bp DNA ladders (Promega, UK) were included to allow size determination. Electrophoresis was carried out in 1 x TAE buffer at 100 V until DNA fragments were suitably resolved. Gels were visualised using ultra violet (UV) transillumination and photographed using a Syngene InGenius gel documentation system (Syngene, UK).

## 2.2.4 Restriction digest of plasmid DNA

The site-specific restriction digestion of plasmids or PCR products was carried out using enzymes and buffers from New England Biolabs, USA (NEB) and typically contained the following components:

Plasmid DNA or purified PCR product	1 µg
10× NEB buffer 3 or 4	2.5 µL
10× NEB Bovine Serum Albumin	2.5 µL
Enzyme 1 (20 U/µL)	20 U
Enzyme 2 (20 U/µL)	20 U
Nuclease-free deionised H <sub>2</sub> O	to 25 µL

All reactions were accompanied by the appropriate single enzyme and enzyme-free control samples. Reactions were incubated at 37 °C for 1 h, followed by enzyme inactivation at 65 °C for 20 min.

The restriction enzymes, buffer components and unwanted by-products of the digestion were subsequently removed by separating the DNA fragments by agarose gel electrophoresis (Section 2.2.3). The required DNA fragments were excised using a scalpel and extracted from the agarose gel using the QIAquick Gel Extraction Kit (QIAGEN, UK) according to the manufacturer's instructions.

## 2.2.5 Dephosphorylation of restriction endonuclease digests

In order to prevent plasmid re-ligation, the 5'-ends were dephosphorylated with Antarctic Phosphatase (NEB, USA). 1/10 volume of 10× Antarctic Phosphatase Reaction Buffer (NEB, USA) and 0.01 U of enzyme per pmol DNA ends was added to the completed digest. The reaction mixture was incubated for 15 min at 37 °C. The enzyme was then heat inactivated at 65 °C for 5 min.



### 2.2.6 Ligation of DNA

Ligation of DNA fragments was performed using the NEB Quick Ligation Kit. A typical ligation reaction contained the following components:

Digested and dephosphorylated vector	100 ng
2× NEB Quick Ligation Buffer	10 µL
NEB T4 DNA Ligase (2000 U/µL)	2000 U
Digested insert	to 20 µL

Appropriate control reactions containing nuclease-free deionised water instead of a digested insert were carried out. Reactions were incubated at 25 °C for 15 min then kept on ice prior to transformation into SCS1 supercompetent cells (see Section 2.2.7).

### 2.2.7 Transformation and cultivation of *E. coli* cells

Plasmid DNA (50 – 200 ng) was added to 100 µL of competent cells (Section 2.2.1). If the DNA to be transformed was from a ligation reaction, the supercompetent cells were pre-incubated with 0.8 µL of 0.42 M β-mercaptoethanol for 10 min. After incubation on ice for 30 min, the mixture was heat shocked at 42 °C for 45 sec. The cells were incubated on ice for a further 2 min before the addition of 500 µL of sterile medium (SOC medium for supercompetent SCS1 cells and XL1-Blue cells, LB for all other strains) pre-warmed to 37 °C. The cultures were incubated at 37 °C, 200 rpm, for 1 h to allow the transformed cells to express antibiotic resistance. 20 µL or 200 µL of cell culture was spread onto LB agar plates containing the appropriate antibiotic and incubated overnight at 37 °C.

### 2.2.8 Preparation of plasmids

Single colonies from transformation reactions were picked and grown overnight (37 °C, 200 rpm) in 10 mL LB containing the appropriate antibiotic. Plasmid DNA was extracted using the QIAprep Spin Miniprep Kit (QIAGEN, UK), according to the manufacturer's instructions. DNA required for subsequent biological experiments was eluted in sterile H<sub>2</sub>O. DNA for stock maintenance was eluted in TE buffer (**Table 2.6**). To determine DNA

concentrations, the absorbance was measured by spectrophotometry at 260 nm ( $A_{260}$ ). An  $A_{260}$  of 1 was assumed to be equivalent to 50  $\mu\text{g}/\text{mL}$  of double stranded DNA (dsDNA).

## 2.2.9 Sequencing, stock maintenance and storage of DNA

DNA was sequenced by Beckman Coulter Genomics, UK. Sequencing of genes in pET vectors was carried out using universal T7 promoter and T7 terminator primers. Sequencing of the  $\beta$ -lactamase linker region in the pBR322 vectors was carried out using the primers in **Table 2.8**. Further stocks of the various constructs were produced in DH5 $\alpha$  *E. coli* cells (Invitrogen, UK) (see Section 2.2.7 and Section 2.2.8) and the DNA stored at -80 °C in TE buffer.

## 2.2.10 Plasmids

Plasmids obtained during this project are summarised in **Table 2.10**. Plasmids containing the genes encoding wild-type  $\beta$ -lactamase (pET28a- $\beta$ la-WT),  $\beta$ -lactamase with a 28-residue GS linker (pMB1- $\beta$ la-linker<sub>SHORT</sub>) and  $\beta$ -lactamase with a 64-residue GS linker (pMB1- $\beta$ la-linker) were kindly provided by Professor J. Bardwell (Department of Biological Chemistry, University of Michigan, USA).

Plasmids pTXB1-hIAPP and pTXB1-rIAPP, encoding human IAPP (hIAPP) and rat IAPP (rIAPP) respectively, were a generous gift from Associate Professor A. Miranker (Department of Molecular Biophysics and Biochemistry, Yale University, USA). Plasmid pRSET-A $\beta$ 40/42, containing the gene encoding Alzheimer's  $\beta_{1-40/42}$  (A $\beta$ 40/A $\beta$ 42), was a kind gift from Professor Rudolf Glockshuber (Swiss Federal Institute of Technology, Zurich, Switzerland).

Plasmids pEX-HEL4 and pEX-Dp47d, containing the genes encoding HEL4 and Dp47d respectively, were synthesised by Eurofins MWG Operon, Ebersburg, Germany using the sequence from Jespers *et al.*, 2004<sup>376</sup>.

All DNA and protein sequences are given in **Appendices 7.1** and **7.2**. All plasmid maps are located in **Appendix 7.3**.

Plasmid	Insert	Promoter	Vector Backbone	Antibiotic Resistance	Ref.
pET28a- $\beta$ la-WT	WT $\beta$ la	T7	pBR322	Kanamycin	308
pMB1- $\beta$ la-link-28	$\beta$ la-linker <sub>SHORT</sub>	pBAD	pMB1	Tetracycline	308
pMB1- $\beta$ la-link-64	$\beta$ la-linker	pBAD	pMB1	Tetracycline	308
pTXB1-hIAPP	hIAPP	T7	pMB1	Ampicillin	377
pTXB1-rIAPP	rIAPP	T7	pMB1	Ampicillin	377
pRSET-A $\beta$ 40/42	A $\beta$ 40/42	T7	pRSET	Ampicillin	377

**Table 2.10.** Plasmids obtained for this thesis.

### 2.2.11 Summary of plasmids created

The plasmids created during this thesis were made using the techniques described in Sections 2.23 – 2.2.10 and are summarised in **Table 2.11**.

Plasmid	Insert	Promoter	Vector Backbone	Antibiotic Resistance
$\beta$ la-28-hIAPP	$\beta$ la-hIAPP	pBAD	pMB1	Tetracycline
$\beta$ la-28-rIAPP	$\beta$ la-rIAPP	pBAD	pMB1	Tetracycline
$\beta$ la-28-A $\beta$ 40	$\beta$ la-A $\beta$ 40	pBAD	pMB1	Tetracycline
$\beta$ la-28-A $\beta$ 42	$\beta$ la-A $\beta$ 42	pBAD	pMB1	Tetracycline
$\beta$ la-28-HEL4	$\beta$ la-HEL4	pBAD	pMB1	Tetracycline
$\beta$ la-28-Dp47d	$\beta$ la-Dp47d	pBAD	pMB1	Tetracycline
pET-HEL4	HEL4	T7	pET23a	Ampicillin
pET-Dp47d	Dp47d	T7	pET23a	Ampicillin
pEX-HEL4	HEL4	T7	pEX-A	Ampicillin
pEX-Dp47d	Dp47d	T7	pEX-A	Ampicillin

**Table 2.11.** Plasmids created in this thesis.

## 2.3 Protein expression and purification

### 2.3.1 Purification of $\beta$ -lactamase constructs

#### 2.3.1.1 Small-scale expression trials

Small-scale expression trials were carried out to identify the optimal concentration of arabinose for expression of the  $\beta$ -lactamase constructs. BL21 (DE3) cells were transformed with the relevant plasmid (see **Table 2.11**) as described in Section 2.2.7. A single colony from a fresh transformation was used to inoculate 100 mL of LB medium containing 10  $\mu$ g/mL tetracycline. Cultures were incubated overnight at 37 °C with shaking at 200 rpm. 1 mL of overnight culture was used to inoculate 100 mL sterile LB containing 10  $\mu$ g/mL tetracycline. The culture was incubated at 37 °C with shaking (200 rpm) until an OD<sub>600</sub> of 0.6 was reached. A 50 mL sample was taken before induction and stored at -20 °C. Protein expression in the remaining 50 mL culture was induced by the addition of filter-sterilised arabinose to a final concentration (*w/v*) of 0.002 %, 0.02 %, 0.2 %, or 2 % for the  $\beta$ la-linker constructs, or 10, 100 or 1000 mM IPTG for the WT  $\beta$ -lactamase construct. Cultures were grown for a further 3 h (37 °C, 200 rpm). The uninduced and induced cell cultures were harvested by centrifugation (10 min, 4000 rpm, JS 5.3 Beckman-Coulter rotor, 4 °C) and the periplasmic fractions of the remaining cell pellet extracted as described in Section 2.3.1.2.

#### 2.3.1.2 Extraction of periplasmic fraction

To assess the expression levels of the  $\beta$ -lactamase constructs in the periplasm of *E. coli*, a periplasmic extraction was performed. All steps were carried out in the cold room using chilled reagents. The cell pellet was carefully resuspended in 20 mM Tris-HCl pH 8.0, 20 % (*w/v*) sucrose (4 mL of solution per gram of wet cell pellet). 1/100 volume of 100 $\times$  protease inhibitor cocktail (100 mM phenylmethylsulphonyl fluoride, 200 mM benzamidine, dissolved in EtOH) was added to prevent any unwanted degradation. 40  $\mu$ L 0.5 M EDTA and 40  $\mu$ L hen egg white lysozyme (10 mg/mL) were added per gram of wet cell pellet, and the reaction left for 20 min on a shaker. 80  $\mu$ L of 1 M MgCl<sub>2</sub> (per gram of wet cell pellet) was added to the solution, and the cell suspension was centrifuged

(20 min, 12,000 rpm, F-45-12-11 Eppendorf USA rotor, or JA 25.50 Beckman-Coulter rotor, 4 °C) before the supernatant was collected.

The following protocol was used if polymyxin was utilised in place of the lysozyme for the periplasmic extraction. The cell pellet was carefully resuspended in ice cold polymyxin buffer (**Table 2.6**) (4 mL of solution per gram of wet cell pellet) and left to shake for 60 min at 4 °C. 1/100 volume of 100× protease inhibitor cocktail (100 mM phenylmethylsulphonyl fluoride, 200 mM benzamidine, dissolved in EtOH) was added to prevent any unwanted degradation. The cell suspension was centrifuged (20 min, 10,000 rpm, F-45-12-11 Eppendorf USA rotor, or JA 25.50 Beckman-Coulter rotor, 4 °C) before the supernatant was collected. The resulting supernatant fractions were diluted two-fold in 2× reducing SDS loading buffer (**Table 2.6**), and analysed by SDS PAGE (Section 2.3.1.3).

### 2.3.1.3 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Tris-tricine buffered sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was used to separate proteins according to their molecular weight to monitor over-expression and purification of the recombinantly expressed proteins. Two glass plates were assembled according to the manufacturer's instructions using a 1.5 mm spacer. A two-layered gel system consisting of a stacking and resolving gel was made using the components in **Table 2.12**. The resolving gel mixture was rapidly poured to within 2 cm of the top of the glass plates. The required volume of ammonium persulphate was added to the stacking gel and immediately poured on top of the resolving gel. A comb was inserted to create wells for sample loading. The gels were left for a minimum of 1 h to set.

Solution Component	Resolving Gel (mL)	Stacking Gel (mL)
30 % (w/v) Acrylamide: 0.8 % (w/v) bis-acrylamide	7.50	0.83
3 M Tris-HCl, 0.3 % (w/v) SDS pH 8.45	5.00	1.55
H <sub>2</sub> O	0.44	3.72
Glycerol	2.00	-
10 % (w/v) ammonium persulphate	0.05	0.10
Tetramethylethylenediamine (TEMED)	0.01	0.01

**Table 2.12.** Components of a Tris-tricine buffered SDS-PAGE gel. The volumes stated allow casting of two mini-gels (8 cm × 10 cm) using a 1.5 mm spacer.

Protein samples were diluted two-fold in 2× reducing loading buffer (**Table 2.6**) and boiled (5 min) prior to loading. Mark12™ Unstained Protein Standards (Invitrogen, UK) or Precision Plus Protein™ Dual Xtra Standards (Bio-Rad, UK) were loaded into one lane to allow size determination and thus identification of protein bands. Gels were electrophoresed with the inner reservoir of the gel tank buffered with cathode buffer (**Table 2.6**) and the outer reservoir buffered with anode buffer (**Table 2.6**). A constant current of 30 mA was applied until the samples entered the resolving gel, and then the current was adjusted to 65 mA until the dye front reached the bottom of the gel. Gels were stained using Instant Blue stain (Expedeon, UK) and visualised and recorded using a Syngene InGenius gel documentation system (Syngene, UK).

#### **2.3.1.4 Large-scale expression of $\beta$ -lactamase constructs**

Large-scale expression of all constructs was carried out for *in vitro* analysis of aggregation. BL21 (DE3) cells were transformed with the relevant plasmid (see **Table 2.11**) as described in Section 2.2.7. A single colony from a fresh transformation was used to inoculate 250 mL of LB medium containing the relevant antibiotic and incubated overnight at 37 °C with shaking at 200 rpm. 25 mL of overnight culture was then used to inoculate 10× 1 L sterile LB containing antibiotic. The cultures were incubated at 37 °C with shaking (200 rpm) until an approximate OD<sub>600</sub> of 0.6 was reached. Protein expression was induced by the addition of filter-sterilised arabinose to a final concentration of 0.2 % (w/v). Cultures were incubated at 37 °C, 200 rpm, for 3 h. Cells were harvested in a Stratos continuous-flow rotor centrifuge at 15,000 rpm (HCT 22.300 Heraeus rotor) and the periplasmic fraction extracted as described in Section 2.3.1.2.

#### **2.3.1.5 Removal of sucrose from periplasmic preparation**

The periplasmic fraction containing the  $\beta$ -lactamase construct was dialysed overnight (3,500 MWCO) against 5 L of 20 mM Tris-HCl, pH 7.5 at 4 °C with stirring. This process was repeated 3 times to remove all sucrose from the periplasmic preparation. The solution was filtered through a 0.2  $\mu$ m syringe filter (Sartorius Stedim Biotech, UK) to remove large contaminants.

### 2.3.1.6 Anion exchange chromatography

The dialysed periplasmic fraction obtained from 10 L of cells was loaded onto a 5 mL Hi Trap Q HP anion exchange column connected to an ÄKTA prime (GE Healthcare, UK), equilibrated with five column volumes of 20 mM Tris-HCl, pH 7.5 (buffer A) at 4 °C. The column was washed with three column volumes of buffer A before the protein was eluted with a linear gradient of 0 - 0.5 M NaCl, 20 mM Tris-HCl, pH 7.5 (Buffer B), maintaining the flow rate at 5 mL/min for all steps (see **Table 2.13** for programme details). 1.5 mL fractions were collected and the elution of protein monitored by absorbance at 280 nm. Fractions corresponding to the various peaks were analysed by SDS PAGE (see Section 2.2.3). The fractions were also tested for enzyme activity (see Section 2.5.1) to confirm the location of  $\beta$ -lactamase. The fractions that contained the  $\beta$ -lactamase construct were pooled for further purification.

Breakpoint (mL)	Flow Rate (mL/min)	Line	Percentage Line B (%)	Fraction Size (mL)	Auto Zero
0	5	A1	0	0	No
15	5	A2	0	10	Yes
50	5	A1	0	10	No
60	5	A1	0	1.5	No
100	5	A1	70	1.5	No
102	5	A1	100	1.5	No
120	5	A1	100	0	No

**Table 2.13.** ÄKTA prime program for  $\beta$ -lactamase constructs and HEL4 purification. Purification was carried out using a Hi Trap Q HP anion exchange column. Line A = Buffer A, Line A2 = protein sample, Line B = Buffer B.

### 2.3.1.7 Size exclusion column chromatography

The fractions containing the desired  $\beta$ -lactamase construct were concentrated to a final volume of 5 mL using an Amicon Ultra Centrifugal filter (10,000 MWCO). The 5 mL sample was loaded onto a Superdex™ 75 GL 10/300 gel filtration column connected to an ÄKTA prime (GE Healthcare, UK), which had been equilibrated with 250 mL of 20 mM Tris-HCl, pH 7.5 at 4 °C. The protein was eluted from the column at a flow rate of 2 mL/min (see **Table 2.14** for program details). 1.5 mL fractions were collected and the elution of protein monitored by absorbance at 280 nm. Fractions corresponding to the main peak were analysed by SDS PAGE (Section 2.2.3).

Breakpoint (mL)	Flow Rate (mL/min)	Fraction Size (mL)	Injection Valve Position	Auto Zero
0	2	0	Load	No
10	2	0	Inject	Yes
20	2	0	Load	No
90	2	1.5	Load	No
320	2	0	Load	No

**Table 2.14.** ÄKTA prime program for  $\beta$ -lactamase constructs and HEL4 purification. Purification was carried out using a Superdex™ 75 GL 10/300 gel filtration column.



## 2.3.2 Purification of HEL4 and Dp47d antibody domains

### 2.3.2.1 Large-scale expression of antibody domains

The antibody domains HEL4 and Dp47d were over-expressed using the following method adapted from Jespers *et al.*, 2004<sup>376</sup>. BL21 (DE3) cells were transformed with either pET-HEL4 or pET-Dp47d (see **Table 2.11**) as described in Section 2.2.7. A single colony from a fresh transformation was used to inoculate 250 mL of LB medium containing 100 µg/mL ampicillin and incubated overnight at 37 °C with shaking at 200 rpm. 25 mL of overnight culture was then used to inoculate 10× 1 L of sterile LB containing 100 µg/mL ampicillin. The cultures were incubated at 37 °C with shaking (200 rpm) until an approximate OD<sub>600</sub> of 0.6 was reached. Protein over-expression was induced by the addition of filter-sterilised IPTG to a final concentration of 100 µM. After 3 h further incubation (37 °C, 200 rpm), cells were harvested in a Stratos continuous-flow rotor centrifuge at 15,000 rpm (HCT 22.300 Heraeus rotor).

### 2.3.2.2 Isolation of antibody domain inclusion bodies

Cell pellets were resuspended in 50 – 100 mL lysis buffer (25 mM Tris-HCl, pH 8.0, or 10 mM sodium phosphate, pH 7.0, lysozyme (100 µg/mL), PMSF (50 µg/mL), DNase (20 µg/mL), and 1 mM EDTA) and incubated at room temperature for 30 minutes with stirring. As the isoelectric point (pI) of Dp47d is 8.02, the lysis buffer was made using 10 mM sodium phosphate buffer, pH 7.0 instead of the 25 mM Tris-HCl, pH 8.0 used for HEL4. The insoluble fraction was pelleted by centrifugation (15,000 rpm, Beckman JLA 16.250 rotor, 4 °C, 30 min). The cell pellet containing the protein as inclusion bodies was washed five times in 100 mL of 25 mM Tris-HCl, pH 8.0 (HEL4) or 10 mM sodium phosphate, pH 7.0 (Dp47d) with the inclusion bodies being pelleted by centrifugation after each resuspension.

### 2.3.2.3 Purification of HEL4 antibody domain

Inclusion bodies containing the HEL4 antibody domain were solubilised in 200 mL of 8 M urea, 25 mM Tris-HCl, pH 8.0 overnight, then sealed in a piece of SnakeSkin pleated dialysis tubing (Thermo Scientific, UK; 3,500 MWCO). The protein was refolded by dialysis against 5 L of 25 mM Tris-HCl, pH 8.0. The buffer was exchanged for fresh buffer five times over two days, after which any precipitate was removed by filtering the solution through a 0.2 µm syringe filter (Sartorius Stedim Biotech, UK).

The dissolved HEL4 inclusion bodies were loaded onto a 5 mL Hi Trap Q HP anion exchange column connected to an ÄKTA prime (GE Healthcare, UK), equilibrated with five column volumes of 20 mM Tris-HCl, pH 8.0 (buffer A) at 4 °C. The column was washed with one column volume of buffer A before the protein was eluted with a linear gradient of 0 - 0.5 M NaCl in 20 mM Tris-HCl, pH 8.0 (see **Table 2.13** for program details). The flow rate was maintained at 5 mL/min for all steps. 1.5 mL fractions were collected and the elution of protein monitored by measuring the absorbance at 280 nm. Fractions corresponding to the various peaks were analysed by SDS PAGE (see Section 2.3.1.3). The fractions that contained HEL4 were pooled together and concentrated to a final volume of 5 mL using an Amicon Ultra Centrifugal filter (3,000 MWCO). The protein was further purified by size exclusion chromatography using a Superdex™ 75 GL 10/300 gel filtration column connected to an ÄKTA prime chromatography system (GE Healthcare, UK). The column was equilibrated by washing with two column volumes 25 mM Tris-HCl, pH 8.0. For each gel filtration run, 3 mL of solubilised HEL4 was loaded into a 5 mL loop at a concentration of up to 10 mg/mL and eluted using the program detailed in **Table 2.14**. 1.5 mL fractions were collected and the protein elution monitored by absorbance at 280 nm. Fractions corresponding to the main peak were analysed by SDS PAGE (Section 2.3.1.3) and purity was confirmed by mass spectrometry. Purified HEL4 was freeze-dried and stored at -80 °C.

Analytical size exclusion chromatography was performed using a Superdex™ 75 HR/10/30 analytical gel column connected to an ÄKTA prime chromatography system (GE Healthcare, UK). The column was equilibrated by washing with two column volumes 25 mM Tris-HCl, pH 8.0. For each gel filtration run, 90 µL of solubilised HEL4 (50 µM) was loaded into a 100 µL loop and eluted at 0.5 mL/min. Protein elution was monitored by absorbance at 280 nm.

### 2.3.2.4 Purification of Dp47d antibody domain

Inclusion bodies of the Dp47d antibody domain were solubilised in 10 mL of 4 M GuHCl, 50 mM glycine, pH 9.5, overnight at 4 °C. Any precipitate was removed by filtering the solution through a 0.2 µm syringe filter (Sartorius Stedim Biotech, UK). 3 mL was loaded onto a Superdex™ 75 GL 10/300 connected to an ÄKTA prime (GE Healthcare, UK), equilibrated with two column volumes of 4 M GuHCl, 50 mM glycine, pH 9.5. The protein was eluted following the program detailed in **Table 2.15**. Fractions corresponding to the main peak were pooled and trichloroacetic acid (TCA) precipitation (Section 2.3.2.5) was used to remove GuHCl from protein samples to facilitate analysis by SDS-PAGE (see Section 2.3.1.3). Purified Dp47d was refolded by dialysis (3,000 MWCO membrane) against 50 mM glycine, pH 9.5 and concentrated to a final concentration of 2 mg/mL using an Amicon Ultra Centrifugal filter (3,000 MWCO). Protein was frozen using liquid nitrogen and stored at – 80 °C.

Analytical size exclusion chromatography was performed using a Superdex™ 75 HR 10/30 analytical column connected to an ÄKTA prime chromatography system (GE Healthcare, UK). The column was equilibrated by washing with two column volumes 25 mM Tris-HCl, pH 8.0. For each gel filtration run, 90 µL of solubilised HEL4 (50 µM) was loaded into a 100 µL loop and eluted at 0.5 mL/min. Protein elution was monitored by absorbance at 280 nm.

Breakpoint (mL)	Flow Rate (mL/min)	Fraction Size (mL)	Injection Valve Position	Auto Zero
0	1	0	Load	No
10	1	0	Inject	Yes
20	1	0	Load	No
90	1	1.5	Load	No
320	1	0	Load	No

**Table 2.15.** ÄKTA prime programme parameters for purification of Dp47d. Gel filtration was carried out using a Superdex™ 75 GL 10/300 in 4 M GuHCl, 50 mM glycine, pH 9.5.

### 2.3.2.5 Trichloroacetic acid precipitation

TCA precipitation was used to remove GuHCl from Dp47d protein samples prior to SDS-PAGE. An equal volume of 12.5 % (v/v) TCA was added to the sample, which was then incubated on ice for 20 min. The precipitated protein was collected by centrifugation in a microfuge (13,000 rpm, JS 5.3 Beckman-Coulter rotor, 10 min, 4 °C) and the supernatant discarded. The pellet was washed with 1 mL ice cold EtOH and collected by centrifugation as before. The supernatant was again discarded and any residual solvent was removed from the protein pellet by drying in a heat block at 95 °C. The dried pellet was resuspended in SDS-PAGE loading buffer and boiled for 5 min before loading onto the gel.

### 2.3.3 Purification of amyloid- $\beta_{40}$

The amyloid- $\beta_{40}$  purification protocol was developed and performed by Rachel A. Mahood (The Astbury Centre for Structural Molecular Biology, Faculty of Biological Sciences, University of Leeds, UK).

The plasmid containing amyloid- $\beta_{1-40}$  (A $\beta_{40}$ ) was kindly donated by Dominic Walsh (Brigham & Women's Hospital, Boston, USA) and Sara Linse (Lund University, Sweden)<sup>378</sup>. *E. coli* BL21 (DE3) pLysS cells expressing A $\beta_{40}$  were grown in LB medium containing 100  $\mu$ g/mL ampicillin and 25  $\mu$ g/mL chloramphenicol at 37 °C with shaking (200 rpm) until an approximate OD<sub>600</sub> of 0.5 was reached. Protein expression was induced by the addition of filter-sterilised IPTG to a final concentration of 0.5 mM. Cells were incubated at 37 °C for 3 h before being collected by centrifugation (6000 rpm, JLA 8.1000 Beckman-Coulter rotor, 15 min, 4 °C). Purification was performed using a modified version of a protocol (provided by Walsh *et al.*, 2009<sup>378</sup>) published in Saunders & Young *et al.*, 2014<sup>379</sup>. Cells were disrupted in 10 mM Tris-HCl, 1 mM EDTA, pH 8.5, 20  $\mu$ g/mL DNase, 1 mM PMSF, 2 mM benzamide. The suspension was stirred at 4 °C for 1 h before homogenisation and sonication (Soniprep 150 sonicator, 9.5 nm probe (MSE, UK), 30 s, 4 W). The extract was centrifuged (20,000 rpm, JA 25.50 Beckman-Coulter rotor, 15 min, 4 °C) and the pellet resuspended in 10 mM Tris-HCl, pH 8.5, 8 M urea. The suspension was sonicated to dissolve the inclusion bodies, and then centrifuged again. The supernatant was collected and diluted 1:4 in 10 mM Tris-HCl, pH 8.5 buffer and agitated gently with Q Sepharose Fast Flow resin (GE Healthcare, UK.) for 30 min. The resin was washed with 50 mL of 10 mM Tris-HCl, pH 8 and then again with 10 mM Tris-HCl, 25 mM NaCl, pH 8.

Peptide-enriched fractions were then eluted with 4× 50 mL of 125 mM NaCl, dialysed into 50 mM ammonium bicarbonate and lyophilised. Partially purified A $\beta$ 40 (~10 mg/mL) was resolubilised in 50 mM Tris-HCl, 7 M GuHCl, pH 8.5 and purified by size exclusion chromatography on a HiLoad Superdex™ 75 26/60 column (GE Healthcare, UK) into 50 mM ammonium bicarbonate before being lyophilised and stored at -20 °C.

### **2.3.4 Acquisition of IAPP and A $\beta$ 42**

Human and rat islet amyloid polypeptide were synthesised and purified as described in Marek, 2010<sup>380</sup> and kindly provided by Cynthia Tu and Professor Daniel Raleigh (Stony Brook University, New York, USA). Synthetic A $\beta$ 42 was purchased from Invitrogen (Paisley, UK), catalogue number 03-112.

## 2.4 Minimal inhibitory concentration (MIC) of antibiotic assays

### 2.4.1 MIC assay in absence of small molecules

Spot titre tests were performed to determine the level of antibiotic resistance of cells expressing the  $\beta$ -lactamase tripartite fusions. A single colony from fresh *E. coli* BL21 (DE3) cells, transformed with the appropriate plasmid (**Table 2.11**), was used to inoculate 100 mL sterile LB containing 10  $\mu\text{g}/\text{mL}$  tetracycline (Section 2.2.7). Cultures were incubated overnight at 37 °C with shaking (200 rpm). 1 mL of overnight culture was used to inoculate 100 mL sterile LB containing 10  $\mu\text{g}/\text{mL}$  tetracycline and grown at 37 °C (shaking at 200 rpm) until an  $\text{OD}_{600}$  of 0.6 was reached. Expression of the  $\beta$ -lactamase construct was induced by the addition of filter-sterilised arabinose to a final concentration of 0.02 % (*w/v*). Cultures were incubated for a further 1 h, when the  $\text{OD}_{600}$  of the cells was adjusted to 1.0 using sterile LB, and serially diluted in 10-fold increments into sterile 170 mM NaCl solution. 2  $\mu\text{L}$  of each dilution was then spotted onto LB agar plates supplemented with 10  $\mu\text{g}/\text{mL}$  tetracycline, 0.02 % (*w/v*) arabinose, and increasing concentrations of ampicillin (0.0 – 140  $\mu\text{g}/\text{mL}$ ). The plates were incubated at 37 °C for 18 h and the maximal cell dilution allowing cell growth was determined for each ampicillin concentration.

### 2.4.2 MIC assay in presence of increasing small molecule concentrations

Circular agar plates containing the small molecule of interest were prepared prior to the assay using the following technique. Tetracycline (final concentration of 10  $\mu\text{g}/\text{mL}$ ), filter sterilised arabinose (0.02 % *w/v*) and increasing volumes of 10 mM small molecule (dissolved in DMSO,  $\text{H}_2\text{O}$ , or EtOH) or solvent were added to 25 mL of sterile LB agar cooled to < 50 °C, to give final concentrations of small molecule of 0, 0.2, 2, 20 or 200  $\mu\text{M}$ . Ampicillin stock was added to give final ampicillin concentrations from 0 – 0.5 mg/mL. The plates were poured in a sterile environment.

The same culture growth conditions as described above (Section 2.4.1) were used until cultures had reached an  $\text{OD}_{600}$  of 0.6.  $\beta$ -lactamase construct expression was induced by the

addition of filter-sterilised arabinose to a final concentration of 0.02 % (w/v). 10 mM small molecule stock (or solvent) was added to the cultures to give final concentrations of 0, 0.2, 2, 20 or 200  $\mu$ M. These cultures were incubated for a further 1 h after which the OD<sub>600</sub> of the cells was adjusted to 1.0 using sterile LB and serially diluted in 10-fold increments into sterile 170 mM NaCl solution. 3  $\mu$ L of each dilution was then spotted onto the agar plates supplemented with small molecule, 10  $\mu$ g/mL tetracycline, 0.02 % (w/v) arabinose, and increasing concentrations of ampicillin (0.0 – 0.5 mg/mL). The plates were incubated at 37 °C for 18 h and the maximal cell dilution allowing cell growth was determined for each ampicillin concentration.

### 2.4.3 MIC assay with bacterial lawns

To reduce the number of agar plates required to carry out the MIC assay over a wide range of small molecule concentrations, bacterial lawns in 8-well plates were utilised. **Figure 4.5** shows a summary schematic of this technique. Bottom agar was prepared by pipetting 4 mL of 1.5 % (w/v) LB agar containing 10  $\mu$ g/mL tetracycline, 0.02 % (w/v) arabinose and 10 mM small molecule (to give the final concentration required) into each well of a sterile 8-well dish (Nunc International). Single colonies of BL21 (DE3) cells transformed with the relevant plasmid (**Table 2.11**) were used to inoculate 100 mL LB overnight starter cultures, 1 mL of which in turn was used to inoculate 100 mL LB in 250 mL flasks. The cultures were grown at 37 °C until an OD<sub>600</sub> of 0.6 was reached. 0.5 mL of culture was then used to inoculate 25 mL of sterile 0.7 % (w/v) LB agar containing 0.02 % (w/v) arabinose at < 40 °C. 400  $\mu$ L of this culture was added to prepared small molecule aliquots (containing equal final concentrations of solvent: DMSO, H<sub>2</sub>O, or EtOH) ranging from 0.0 – 200  $\mu$ M. 350  $\mu$ L of this was pipetted on top of the bottom agar in a single well in the 8-well dish. Top agar was allowed to set before 1  $\mu$ L of increasing concentrations of ampicillin (0.0 – 0.5 mg/mL) was pipetted on top of each well. Plates were incubated for 18 h at 37 °C. The maximum concentration of ampicillin at which the cells could grow was scored for each construct and each concentration of small molecule. Inhibition of growth was scored by the identification of zones of clearance in the bacterial lawn.

### 2.4.4 MIC assay in presence of constant small molecule

Assays were performed in 48-well agar plates (Greiner Bio-One, Germany) using a constant concentration of small molecule (100  $\mu$ M). **Figure 4.7** shows a summary schematic of the technique.

48-well agar plates containing the small molecule of interest were prepared prior to the assay using the following technique. Two plates were prepared for each small molecule of interest, one for bacteria expressing the  $\beta$ la-linker construct and one for bacteria expressing the  $\beta$ la-hIAPP construct. 3  $\mu$ L of 10 mM small molecule (dissolved in DMSO, H<sub>2</sub>O, or EtOH) were added to each well of two 48-well plates. Tetracycline (10  $\mu$ g/mL final concentration) and filter sterilised arabinose (0.02 % w/v) were added to 100 mL of sterile agar cooled to < 50 °C in a beaker. 297  $\mu$ L of this solution was pipetted into each of the first 6 wells (first row) of both 48-well plates. Plates were shaken (manually) to ensure homogenous mixing of the agar and small molecule. Ampicillin (10 mg/mL stock) was added to the beaker of agar to give the required volume for the next row of wells (see **Table 2.16**). 297  $\mu$ L of this agar was then pipetted into row two of each 48-well agar plate. This procedure was repeated until the plate contained 8 rows of agar containing concentrations of ampicillin from 0 – 140  $\mu$ g/mL. Agar plates were left to set in a sterile environment. The same protocol was followed to produce an additional two plates containing no small molecule (3  $\mu$ L of solvent was added to the wells in place of small molecule).

[Ampicillin] required ( $\mu$ g/mL)	Agar stock volume (mL) (X)	Small molecule required in X (mg)	Small molecule already in X (mg)	Additional small molecule required in X (mg)	10 mg/mL ampicillin to add to X ( $\mu$ L)
0	100.0	0.00	0.00	0.00	0
20	96.4	1.93	0.00	1.93	193
40	92.8	3.71	1.86	1.86	186
60	89.2	5.35	3.57	1.78	178
80	85.6	6.85	5.14	1.71	171
100	82.0	8.20	6.56	1.64	164
120	78.4	9.41	7.84	1.57	157
140	74.8	10.5	8.98	1.50	150

**Table 2.16.** Protocol for titration of ampicillin into agar stock.



The same culture growth conditions as described above (Section 2.4.1) were used until cultures had reached an  $OD_{600}$  of 0.6.  $\beta$ -lactamase construct expression was induced by the addition of filter-sterilised arabinose to a final concentration of 0.02 % (w/v). 396  $\mu$ L were removed and added to a 1.5 mL Eppendorf containing 4  $\mu$ L of 10 mM small molecule (to give a final concentration of 100  $\mu$ M). These cultures were incubated for a further 1 h when the  $OD_{600}$  of the cells was adjusted to 1.0 with LB media and serially diluted in 10-fold increments into sterile 170 mM NaCl solution. 3  $\mu$ L of each dilution was then spotted onto the agar plates supplemented with 100  $\mu$ M small molecule, 10  $\mu$ g/mL tetracycline, 0.02 % (w/v) arabinose, and increasing concentrations of ampicillin (0.0 – 140  $\mu$ g/mL). The plates were incubated at 37 °C for 18 h and the maximal cell dilution allowing cell growth was determined for each ampicillin concentration.

## 2.4.5 MIC assay in presence of excipients

### 2.4.5.1 Sorbitol and glycerol

48-well agar plates containing the excipient of interest were prepared prior to the assay using the following technique. A 5 M stock solution of sorbitol and a 5 M stock solution of glycerol were prepared using LB media and sterilised by autoclave (20 min at 121 °C, 15 psi). The solutions were cooled to < 50 °C then mixed, according to **Table 2.17**, with 2× strength sterile agar (3 % w/v, cooled to < 50 °C) to give the required final concentration of excipient. Tetracycline (final concentration of 10  $\mu$ g/mL) and filter sterilised arabinose (0.02 % w/v) were added to the agar solution, and the 48-well agar plates were prepared as described in Section 2.4.4 to give 8 rows of agar containing concentrations of ampicillin from 0 – 140  $\mu$ g/mL (20  $\mu$ g/mL increments). The plates were poured in a sterile environment.

The same culture growth conditions as described in Section 2.4.1 were used, however the 1.5 mL Eppendorfs contained 200  $\mu$ L of sorbitol or glycerol stock solutions to give the final excipient concentrations required in a final volume of 400  $\mu$ L. When cultures had reached an  $OD_{600}$  of 0.6,  $\beta$ -lactamase construct expression was induced by the addition of filter-sterilised arabinose to a final concentration of 0.02 % (w/v). 200  $\mu$ L were removed and added to the 1.5 mL Eppendorf tubes. The same protocol as described in Section 2.4.4 for addition of cultures to plates and incubation were followed.

[Excipient] Required (M)	5 M Excipient Stock (mL)	Sterile LB Medium (mL)	2× Sterile Agar (3 % w/v) (mL)
0.000	0.0	0.0	50.0
0.250	5.0	45.0	50.0
0.375	7.5	42.5	50.0
0.500	10.0	40.0	50.0
1.000	20.0	30.0	50.0
1.500	30.0	20.0	50.0
2.000	40.0	10.0	50.0

**Table 2.17.** Protocol for preparation of sorbitol and glycerol agar plates.

### 2.4.5.2 Arginine and glutamate

The agar plates for the excipients arginine and glutamic acid were prepared in the same manner as described in Section 2.4.2, however the pH of the agar had to be adjusted prior to pouring the agar plates. A stock solution containing 200 mM L-arginine and 200 mM L-glutamate was prepared using LB medium and the pH reduced to 7 using 5 mM HCl. The solution was sterilised using a 0.2 µm filter. The solutions were cooled to < 50 °C then mixed, according to **Table 2.18**, with 2× strength sterile agar (3 % w/v, cooled to < 50 °C) to give the required final concentration of excipient. Stock solutions of 200 mM L-arginine or 200 mM D-arginine were utilised for the plates requiring only these excipients. The same protocol as described for the excipients sorbitol and glycerol in Section 2.4.5.2 was followed.

[Excipient] Required (M)	200 mM Arg:Glu Stock (mL)	Sterile LB Medium (mL)	2× Sterile Agar (3 % w/v) (mL)
0	0.0	50.0	50.0
25	12.5	37.5	50.0
50	25.0	25.0	50.0
100	50.0	0.0	50.0

**Table 2.18.** Protocol for preparation of mixed L-arginine and L-glutamate agar plates.

## 2.5 *In vitro* techniques

### 2.5.1 Identification of fractions containing $\beta$ -lactamase during purification

Nitrocefin was dissolved in DMSO (5 % (w/v), 0.1 M sodium phosphate buffer, pH 7.0 to make a stock solution of 1 mM. The solution was protected from light and stored in 10  $\mu$ L aliquots at  $-20$  °C for a maximum of 2 months. 5  $\mu$ L samples of each fraction were diluted 200 $\times$  or 400 $\times$  in 20 mM Tris-HCl, pH 7.5. 90  $\mu$ L of this dilution was added to 10  $\mu$ L of nitrocefin stock solution and the solution left for 2 min. 90  $\mu$ L of 20 mM Tris-HCl pH 7.5 containing 10  $\mu$ L of nitrocefin was used as a blank. The absorbance at 486 nm was measured after the 2 min incubation.

### 2.5.2 Isoelectric point precipitation of WT $\beta$ -lactamase

The pH-dependence of wild-type (WT)  $\beta$ -lactamase was assessed at pH 3 – 10, using 0.1 M citrate- $\text{Na}_2\text{HPO}_4$  for the pH range 3 – 8, and 0.1 M glycine-NaOH for pH 10. A 50  $\mu$ M solution of WT  $\beta$ -lactamase was buffer exchanged into the required buffers and the enzyme activity measured immediately using the following protocol. 5  $\mu$ L of nitrocefin stock solution (see Section 2.5.1) was added to 45  $\mu$ L WT  $\beta$ -lactamase solution (to give a final protein concentration of 10  $\mu$ M) and the absorbance at 486 nm was measured over time (150 sec).

The Beer-Lambert law (**Equation 2.2**) using a molar extinction coefficient ( $\epsilon$ ) of hydrolysed nitrocefin ( $20,500 \text{ M}^{-1}\text{cm}^{-1}$ )<sup>381</sup> was used to calculate the change in concentration over time. A plot of hydrolysed nitrocefin vs. time was created using the where  $A$  is the absorbance,  $\epsilon$  corresponds to the molar extinction coefficient of nitrocefin ( $\text{M}^{-1}\text{cm}^{-1}$ ),  $c$  is the concentration of nitrocefin (M), and  $l$  is the length of the light path (cm).

$$A = \epsilon cl \quad \text{Equation 2.2}$$

Percent activity was calculated by taking the activity at pH 7 as 100 %.

The activity of WT  $\beta$ -lactamase samples aggregated by isoelectric point precipitation was measured using the following protocol. The pH of a WT  $\beta$ -lactamase solution in 20 mM Tris-HCl, pH 7.5 was adjusted to 5.4 with concentrated hydrochloric acid. The solution was diluted to 50  $\mu$ M and incubated at 37 °C, 200 rpm, for 0, 12 or 24 h before the activity was measured as described above. Sample turbidity at the end of each incubation time was measured as described in Section 2.5.5.

### **2.5.3 Thermal denaturation of WT $\beta$ -lactamase**

Tryptophan fluorescence emission spectra were acquired using a Photon Technology International fluorimeter (Ford, West Sussex, UK). Spectra were recorded using an excitation wavelength of 280 nm, and recording emission from 290 nm to 400 nm with excitation and emission slit widths set at 3 nm. A step size of 1 nm and resolution time of 1 s was used. Blank spectra were subtracted from the sample. 1  $\mu$ M WT  $\beta$ -lactamase in 50 mM sodium phosphate, pH 7 was incubated in a water bath at 20 °C for 20 min before measurements were taken. Thermal denaturation was carried out by raising the temperature from 20 °C – 80 °C at 3 °C increments at a ramp rate of 3 °C/min.

### **2.5.4 Freeze-thaw precipitation of WT $\beta$ -lactamase**

100  $\mu$ L samples of 50  $\mu$ M  $\beta$ -lactamase in 20 mM Tris-HCl buffer, pH 7.5 was subjected to 0, 5 or 10 freeze-thaw cycles. Samples were placed in a –20 °C freezer for 20 min, before being thawed at room temperature. Samples were incubated at 44 °C for 1 h between each cycle. The enzyme activity of the samples was measured as described in Section 2.5.2.

### **2.5.5 Nephelometry**

Nephelometry was used to determine the extent of aggregation in protein samples. 100  $\mu$ L of 50  $\mu$ M protein in appropriate buffer was added to the wells of a flat-bottomed 96-well plate (CoStar, UK). To avoid evaporation and condensation, the plates were sealed with transparent, hydrophobic and gas permeable plastic films (Breathe Easy, Sigma Aldrich) during incubation and measurement. Aggregation endpoint and time profiles were recorded by monitoring the turbidity of the sample using a NEPHELOstar Galaxy laser-

based microplate nephelometer (BMG LABTECH, Germany). The laser intensity was adjusted to 90 % and the laser beam focus to 2.5 mm. For aggregation endpoint assays, single readings were taken after the plate was orbitally shaken with a shake width of 2 mm for 30 s. For aggregation time profiles, incubation temperature was 37 °C, and between measurement cycles the plate was orbitally shaken with a shake width of 2 mm. These values were determined to be optimal by performing the assay in the presence of pre-aggregated protein. The turbidity in the absence of protein was measured and subtracted from all the protein measurements.

### **2.5.6 Transmission electron microscopy**

Transmission electron microscope images were acquired on a *JEOL JEM-1400* transmission electron microscope (JEOL Ltd., Japan) equipped with a Gatan Orius camera after incubating 32 µM protein ( $\beta$ 1a-linker,  $\beta$ 1a-hIAPP, rIAPP, hIAPP, A $\beta$ 40, or A $\beta$ 42) solutions in the presence or absence of 320 µM small molecule for 5 days at 25 °C. Carbon grids were prepared by irradiating under UV light for 30 min and staining with 4 % (w/v) uranyl acetate solution.

### **2.5.7 Thioflavin T fluorometry**

100 µL samples containing 100 µM thioflavin T and 32 µM protein in 200 mM ammonium acetate, pH 6.8 and a 1 % (v/v) final concentration of DMSO were prepared in a 96-well plate and sealed with clear sealing film. Plates were incubated in a FLUOstar OPTIMA plate reader for 5 days at 25 °C without agitation. Fluorescence was excited using a 440 ± 5 nm filter, and emission intensity was measured using a 485 ± 5 nm filter.

### **2.5.8 Mass spectrometry**

All mass spectrometry was carried out by Lydia M. Young, under the supervision of Professor Alison A. Ashcroft (Astbury Centre for Structural Molecular Biology, Faculty of Biological Sciences, University of Leeds, UK).

### 2.5.8.1 Sample preparation

Lyophilised hIAPP and rIAPP samples were dissolved in DMSO to a concentration of 3.2 mM. After 24 h incubation at 25 °C, stock solutions were diluted in 200 mM ammonium acetate, pH 6.8, to give a final peptide concentration of 32 µM for mass spectrometry analysis. The final concentration of DMSO was 1 % (v/v). Lyophilised Aβ40 and Aβ42 were dissolved at 32 µM in 200 mM ammonium acetate, pH 6.8, 1 % DMSO (v/v). The Aβ40 and Aβ42 peptide samples were centrifuged at 13,000 rpm, 4 °C for 10 min before analysis. All samples were incubated at 25 °C in 96-well plates without agitation.

### 2.5.8.2 ESI-(IMS)-MS analysis

A Synapt HDMS quadrupole-time-of-flight mass spectrometer (Micromass UK Ltd., Waters Corp., Manchester, UK), equipped with a Triversa (Advion Biosciences, Ithaca, NY, USA) automated nano-ESI interface, was used for the analyses. The instrument has a travelling-wave IMS device situated between the quadrupole and the time-of-flight analysers, and has been described in detail elsewhere<sup>382</sup>. hIAPP, Aβ40 or Aβ42 samples were analysed using positive ionisation nanoESI (nESI) with a capillary voltage of 1.7 kV and a nitrogen nebulising gas pressure of 0.8 psi. The following instrumental parameters were used: cone voltage 30 V; source temperature 60 °C; backing pressure 1.6 mBar; ramped travelling wave height 7–20 V; travelling wave speed 300 m/s; IMS nitrogen gas flow 20 mL/min; IMS cell pressure 0.55 mBar. Data were processed by use of MassLynx v4.1 and Driftscope software supplied with the mass spectrometer. The *m/z* scale was calibrated with aq. CsI cluster ions.

For ESI-IMS-MS time course experiments, 50 µM peptide samples were incubated in 200 mM ammonium acetate buffer, pH 6.8, 1 % DMS) for 2 min or 24 h. 10 µL volumes were removed from each solution and infused into the mass spectrometer for analysis.

### 2.5.8.3 Analysis of ligand binding to monomeric hIAPP

hIAPP or Aβ40 (32 µM) were dissolved in 200 mM ammonium acetate (pH 6.8) containing 32 µM or 320 µM of small molecule. For analysis of these samples by nESI-MS, a sampling cone voltage of 30 V was used to preserve protein-ligand interactions, and a backing pressure of 1.6 mbar was applied. Data were acquired over the range *m/z* 200–6,000, and

processed by use of MassLynx v4.1 and Driftscope software supplied with the mass spectrometer. The  $m/z$  scale was calibrated with aq. CsI cluster ions.

### 2.5.9 Virtual screening for small molecules

Virtual screening was performed by Charlotte H. Revill and Dr. Richard J. Foster (University of Leeds). The structure of each of the five query molecules (vanillin, resveratrol, curcumin, chloronaphthoquinine-tryptophan and epigallocatechin-3-gallate (EGCG)) was minimised to the lowest energy conformer using LigPrep<sup>383</sup>. The minimised conformers were used as the query scaffold for virtual screening of an in-house library of 50,000 structurally diverse, novel small molecules using Rapid Overlay of Chemical Structures (ROCS)<sup>384</sup>. ROCS is a 3D method that matches the shape of a molecule to the shape of the query molecule. ROCS also incorporates pharmacophoric features in assessing overlays such that the *ROCS Combiscore* in ROCS measures the similarity of the matched shapes as well as the matched pharmacophoric features. Virtual hits were pooled and ranked according to the *ROCS Combiscore* parameter and 20 of the top 100 compounds selected for screening based on a qualitative assessment of structural diversity.





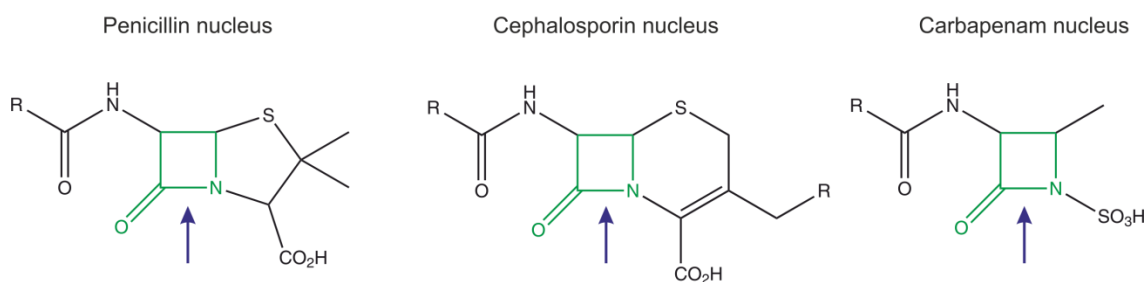
### 3 Demonstrating assay feasibility *in vivo*

#### 3.1 Objectives

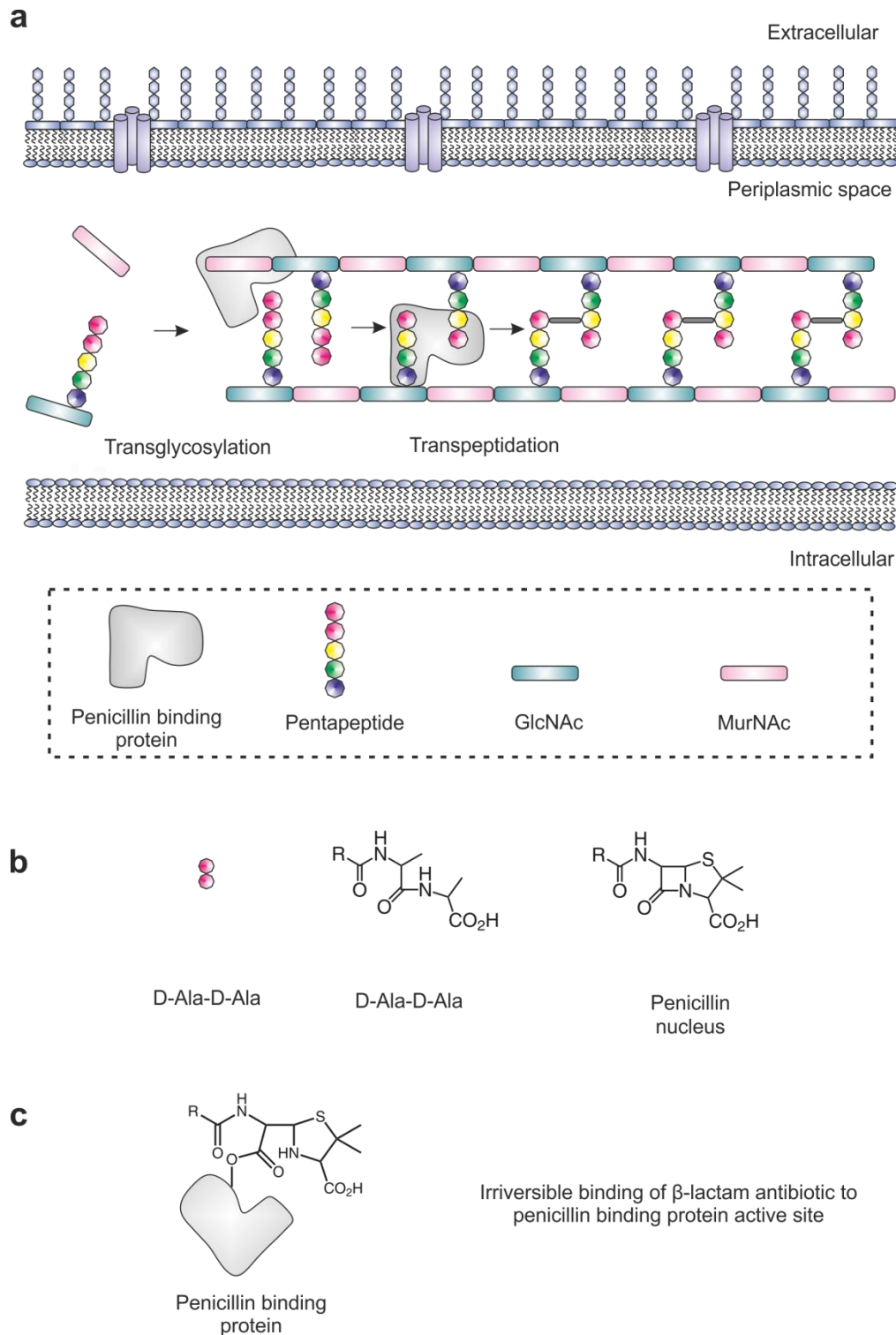
The preliminary objective of this thesis was to develop a periplasmic-based enzyme assay to identify protein aggregation *in vivo*. The assay requires the activity of the reporter protein to be modulated when in an aggregated state, induced by the insertion of an aggregation-prone sequence into a loop region separating the two enzyme domains. The reporter protein used in this assay is the periplasmic enzyme  $\beta$ -lactamase, inspired from previous studies using a tripartite  $\beta$ -lactamase construct<sup>308, 313</sup>.

#### 3.2 The reporter protein $\beta$ -lactamase

The production of  $\beta$ -lactamases is the most widespread defence mechanism of bacteria to  $\beta$ -lactam antibiotics such as the penams, cephalosporins and carbapenems (**Figure 3.1**). These chemical compounds interfere with the final step in the synthesis of peptidoglycan, a major component of bacterial cell walls. Peptidoglycan is a strong, yet flexible, cross-linked polymer which gives structural strength to the bacteria and protects them from the osmotic pressure of the cytoplasm<sup>385</sup>. It is composed of linear glycan strands of repeating disaccharide units of *N*-acetyl glucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc), cross-linked via short peptides (**Figure 3.2a**).



**Figure 3.1.** The core structures of  $\beta$ -lactam antibiotics. They all share a common  $\beta$ -lactam ring, shown in green. The arrows indicate where  $\beta$ -lactamase hydrolyses the amide bond of the  $\beta$ -lactam ring.

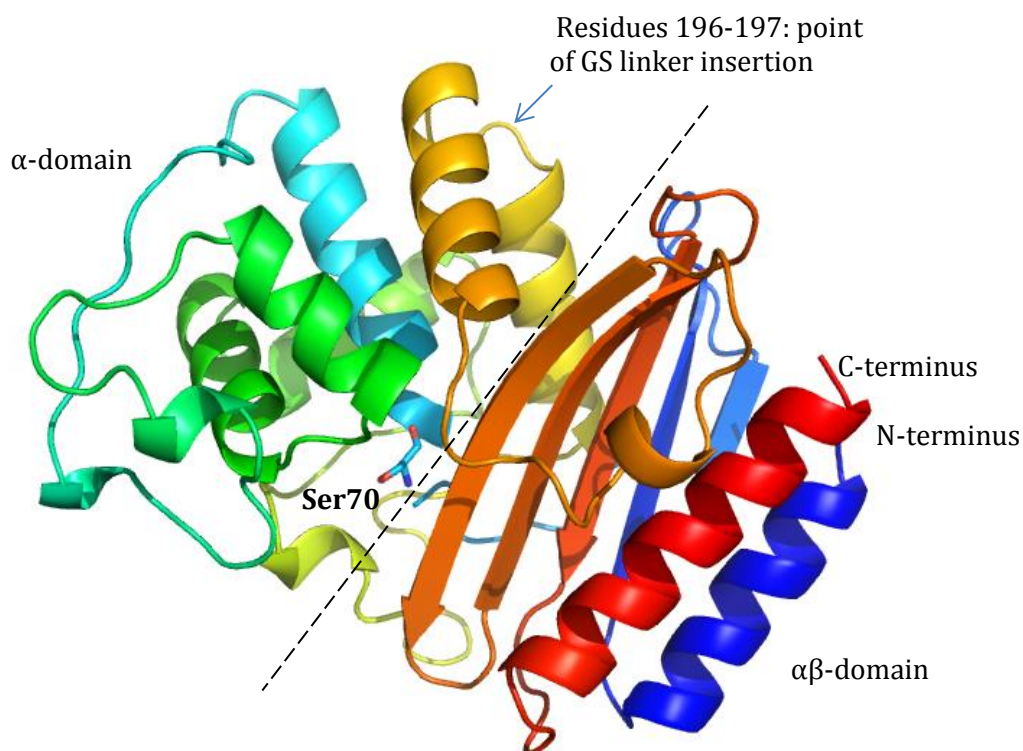


**Figure 3.2.** Biosynthesis of peptidoglycan and its inhibition by  $\beta$ -lactam antibiotics. **(a)** Schematic of peptidoglycan (PG) biosynthesis in gram negative bacteria. The PG precursors GlcNAc and MurNAc-pentapeptide are incorporated into the polymer by transglycosylation and transpeptidation, carried out by the activity of penicillin binding proteins (PBPs). **(b)**  $\beta$ -lactam antibiotics are structural analogues of the dipeptide terminus (D-Ala-D-Ala). **(c)**  $\beta$ -lactam antibiotics irreversibly bind to PBP active sites, thereby rendering the enzymes inactive and unable to synthesise further peptidoglycan.

The bactericidal effect of  $\beta$ -lactam antibiotics results from their inhibition of the penicillin binding proteins (PBPs) responsible for the cross-linking (via transpeptidation) of nascent peptidoglycan chains<sup>386</sup>.  $\beta$ -lactams are structural analogues of D-Ala-D-Ala, the dipeptide terminus of the peptide component of the peptidoglycan unit (**Figure 3.2b**). This structural similarity enables the  $\beta$ -lactam antibiotic to irreversibly bind the catalytic site of the transpeptidase PBP (**Figure 3.2c**)<sup>386</sup>. This process disrupts cell wall synthesis and, at high enough concentrations of  $\beta$ -lactam antibiotic, results in cell lysis<sup>387</sup>.

$\beta$ -lactamases prevent the irreversible inhibition of PBPs by catalysing the irreversible hydrolysis of the amide bond in  $\beta$ -lactam antibiotics (**Figure 3.1**). The enzymes are secreted either to the periplasm (Gram negative bacteria) or the outer membrane (Gram positive bacteria) and can be produced constitutively or induced by the presence of  $\beta$ -lactam agents<sup>388</sup>. The molecular classification of  $\beta$ -lactamases is based on the similarities in their amino acid sequences and the nature of their catalytic mechanism. There are four classes (A to D), three of which contain a serine as an active site residue (classes A, C and D). A small number (class B) require the presence of one or two  $Zn^{2+}$  ions in their active site for catalysis<sup>389</sup>. Interestingly, the enzymes that contain serine as an active site residue show significant structural similarities to the PBPs, suggesting that they are derived from the same ancestral protein<sup>390</sup>.

The  $\beta$ -lactamase used in this study is the class A enzyme TEM-1  $\beta$ -lactamase (E.C. 3.5.2.6)<sup>312</sup>. It is composed of 263 residues organised into two domains: the  $\alpha\beta$ -domain and the  $\alpha$ -domain (**Figure 3.3**). The  $\alpha\beta$ -domain is formed from a large 5-stranded  $\beta$ -sheet and 3  $\alpha$ -helices. The  $\alpha$ -domain contains 8  $\alpha$ -helices and several loops. Together, the domains form a substrate binding cleft on the surface<sup>391</sup>.



**Figure 3.3.** Ribbon diagram of TEM-1  $\beta$ -lactamase from *E. coli*. Elements of secondary structure are coloured from the N-terminal to the C-terminal end in the order violet, blue, green, yellow, orange, red. The  $\alpha\beta$ -domain and the  $\alpha$ -domain are labelled and are situated either side of the dashed line. The C terminus and N terminus are labelled, as is the active site Ser70. Protein DataBank (PDB) entry 1BTL, Jelsch *et al.*, 1993<sup>392</sup>. Drawn using PyMOL Molecular Graphics System (Schrödinger, LLC).

### 3.3 Wild-type $\beta$ -lactamase is rendered inactive when aggregated

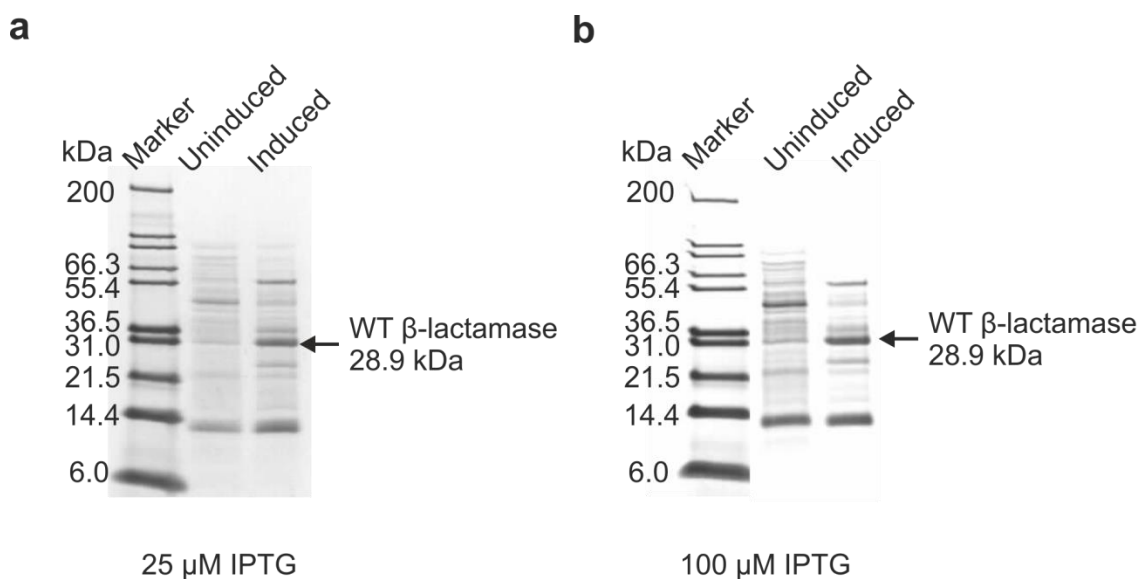
The split  $\beta$ -lactamase aggregation assay is based on the successful application of a  $\beta$ -lactamase tripartite fusion system to study protein stability *in vivo*<sup>308, 313</sup> (see Section 1.7.1). The hypothesis of the assay is that if the test protein aggregates, the  $\beta$ -lactamase domains will either be prevented from associating, or drawn into the aggregate, leading to a loss of activity and a loss of bacterial resistance to the antibiotics. Consequently, a requirement to convert the split  $\beta$ -lactamase system from an assay of protein stability to one for protein aggregation is that the  $\beta$ -lactamase enzyme is rendered inactive when in an aggregated state. Previous studies have shown TEM-1  $\beta$ -lactamase to be unable to function when denatured and aggregated<sup>393, 394</sup>. Furthermore,  $\beta$ -lactamase inclusion bodies have been shown to lack enzymatic activity until careful renaturation of the enzyme has occurred<sup>395</sup>. These results illustrate the potential for using the enzymatic activity of  $\beta$ -lactamase as a measure of its aggregated state. To verify these observations, it was necessary to first purify wild-type (WT) TEM-1  $\beta$ -lactamase ( $\beta$ la-WT). The purification protocol, described below, was also used as the basis for future purification of the  $\beta$ la-test protein constructs.

#### 3.3.1 Purification of wild-type $\beta$ -lactamase

##### 3.3.1.1 Protein expression trials

The plasmid containing the gene encoding WT  $\beta$ -lactamase (pET28a- $\beta$ la-WT, **Table 2.11**, **Appendix 7.23**) was transformed into *E. coli* strain BL21 (DE3) and grown on agar plates containing 50  $\mu$ g/mL kanamycin (Section 2.2.7). Single colonies were then used to inoculate 100 mL LB or TB media overnight starter cultures, 1 mL of which in turn were used to inoculate 100 mL cultures in 250 mL flasks. The cultures were grown at either 25 °C or 37 °C. Once an OD<sub>600</sub> of 0.6 was reached, protein expression was induced with IPTG to final concentrations of 25  $\mu$ M, 100  $\mu$ M, or 1 mM. The cells were then grown for a further 3 h before being pelleted by centrifugation (Section 2.3.1.1). As  $\beta$ -lactamase is a periplasmic protein, its expression was difficult to visualise on an SDS-PAGE gel due to the high abundance of cytoplasmic proteins (data not shown). Therefore, the periplasm was separated from the cytoplasm of the bacteria before analysis by SDS-PAGE (Section 2.3.1.3). The uninduced and induced cell pellets were resuspended in 20 mM

Tris-HCl pH 8.0, 20 % (w/v) sucrose. The outer membrane of the cells was disrupted by treatment with lysozyme and the fraction containing the soluble periplasmic proteins was collected by centrifugation. The samples were analysed by SDS PAGE, with comparative loading guaranteed by the addition of 100  $\mu$ L SDS-loading buffer for each OD<sub>600</sub> unit. The results revealed that the highest protein expression was observed in LB medium at 37 °C, with an IPTG concentration of 100  $\mu$ M (**Figure 3.4**).



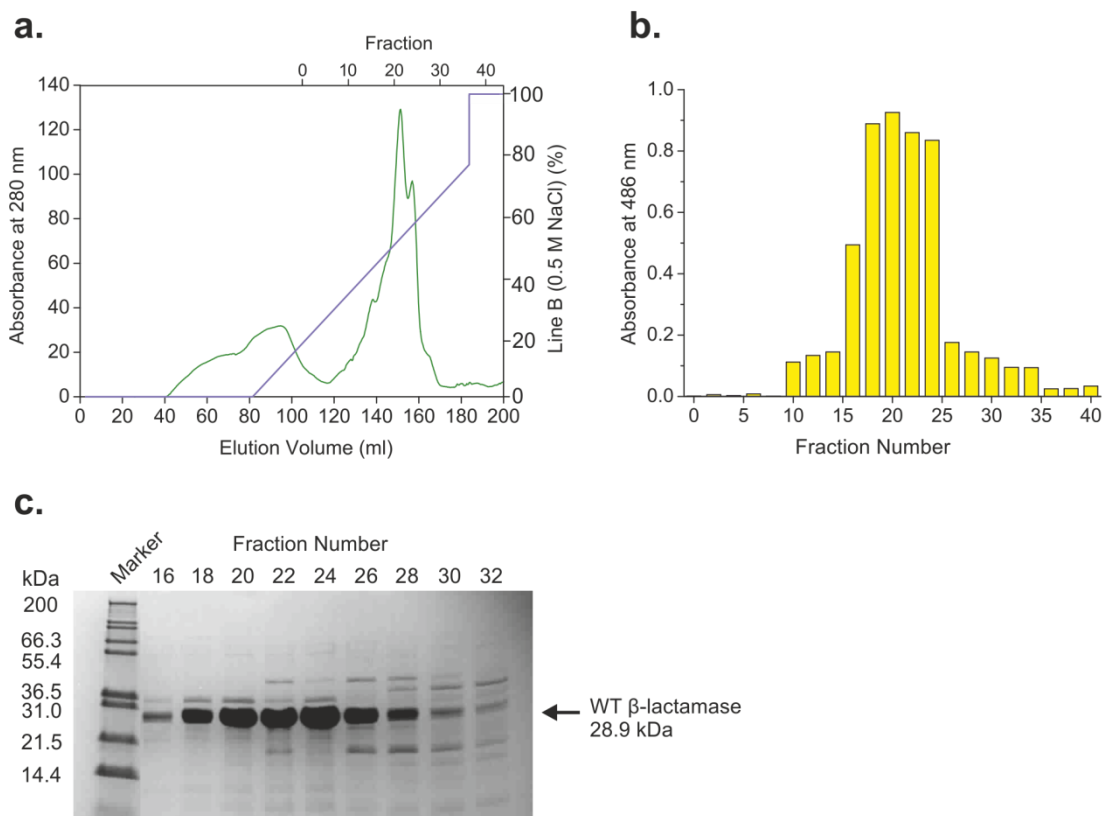
**Figure 3.4.**  $\beta$ la-WT protein expression under pessimal and optimal conditions. **(a)** Lowest expression of  $\beta$ la-WT was observed in LB medium at 37 °C, with an IPTG concentration of 25  $\mu$ M. **(b)** Highest expression of  $\beta$ la-WT was observed in LB medium at 37 °C, with an IPTG concentration of 100  $\mu$ M. Periplasmic fractions extracted before and after induction were analysed. The size in kDa of the protein markers is indicated on the left.

### 3.3.1.2 WT $\beta$ -lactamase purification

Once the growth conditions had been optimised for maximal expression of wild-type  $\beta$ -lactamase, the protein preparation was scaled up to produce enough protein for the subsequent aggregation analyses *in vitro*.

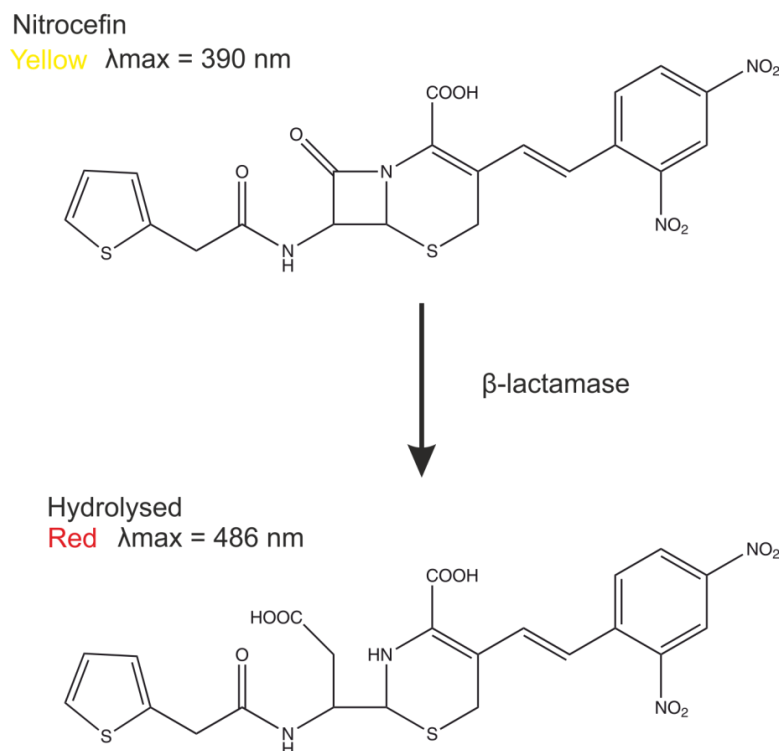
Briefly, single colonies of BL21 (DE3) competent cells transformed with plasmid pET28a- $\beta$ la-WT were used to inoculate 250 mL of LB containing 50  $\mu$ g/mL kanamycin in 500 mL flasks. These starter cultures were incubated overnight before 25 mL was used to inoculate each of ten 1 L solutions of LB in 2 L flasks, containing sterile filtered antibiotic (50  $\mu$ g/mL kanamycin). The flasks were incubated at 37 °C until the OD<sub>600</sub> reached 0.6. Protein expression was then induced by the addition of IPTG to a final concentration of 100  $\mu$ M. The cultures were allowed to grow for 3 h after induction before being harvested in a continual action centrifuge. As the use of lysozyme during the periplasmic extraction adds another protein to the solution that requires purifying out, Polymyxin B was used instead to disrupt the outer membrane of the *E. coli*, as described in Section 2.3.1.2.

In order to purify  $\beta$ -lactamase, the sucrose from the periplasmic preparation was first removed by dialysis. The periplasmic fraction was then loaded onto a Hi Trap anion exchange column. The column was washed with 20 mM Tris-HCl, pH 8.0, before protein elution using a linear gradient of 0.0 – 0.5 M NaCl over 300 mL (Section 2.3.1.6) (**Figure 3.5a**). 1.5 mL fractions were collected during elution, and those containing  $\beta$ -lactamase were identified by enzyme activity assay (**Figure 3.5b**) and SDS PAGE (**Figure 3.5c**). The activity assay uses a chromogenic  $\beta$ -lactamase substrate, nitrocefin. Nitrocefin undergoes a distinctive colour change from yellow ( $\lambda_{\max}$  = 390 nm at pH 7.0) to red ( $\lambda_{\max}$  = 486 nm at pH 7.0) as the amide bond in the  $\beta$ -lactam ring is hydrolysed by  $\beta$ -lactamase (**Figure 3.6**)<sup>381</sup>. 10  $\mu$ L of 1 mM nitrocefin (in 5 % (w/v) DMSO, 0.1 M sodium phosphate buffer, pH 7.0) was added to 85  $\mu$ L of 0.1 M sodium phosphate buffer, pH 7.0, and 5  $\mu$ L of each fraction. The absorbance at 486 nm (relative to a control reaction containing 10  $\mu$ L nitrocefin and 90  $\mu$ L buffer) was recorded after 30 sec (Section 2.5.1) (**Figure 3.5b**). The fractions confirmed to contain  $\beta$ -lactamase were pooled prior to further purification using size exclusion chromatography.



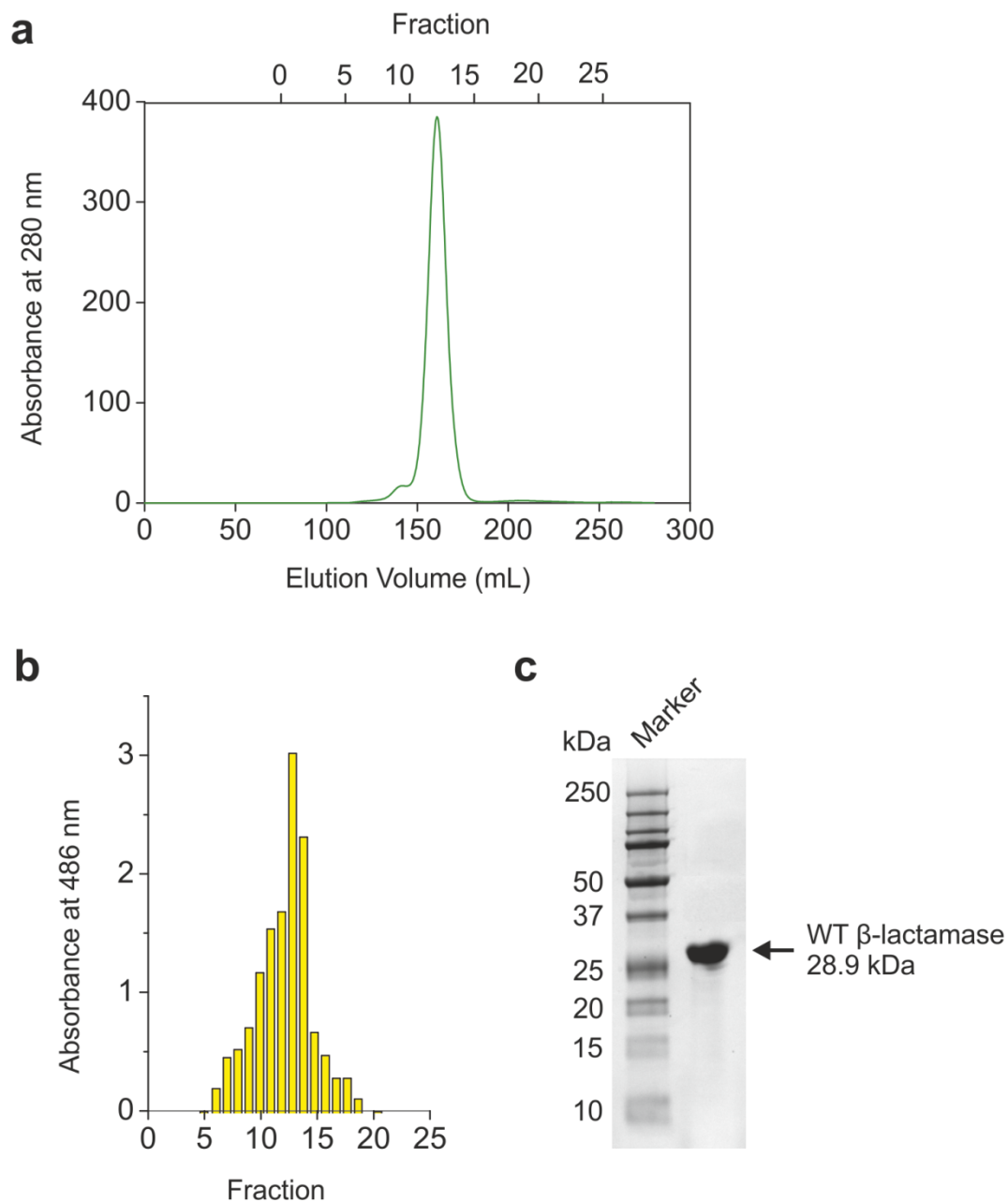
**Figure 3.5.** Anion exchange purification of WT  $\beta$ -lactamase. **(a)** Typical WT  $\beta$ -lactamase elution profile from the anion exchange column. The 0.0 – 0.5 M NaCl gradient is indicated in purple, fractions collected are shown on the top x-axis. **(b)**  $\beta$ -lactamase enzyme activity assay of fractions collected during anion exchange chromatography. The colour change upon hydrolysis of nitrocefin was measured at 486 nm (see **Figure 3.5**). **(c)** SDS-PAGE gel of WT  $\beta$ -lactamase showing fractions collected during anion exchange (numbers denote fractions collected). The size in kiloDaltons (kDa) of the protein marker is indicated.





**Figure 3.6.** Schematic of nitrocefim hydrolysis by  $\beta$ -lactamase. The  $\beta$ -lactam ring in the centre is hydrolysed by  $\beta$ -lactamase, causing a distinctive colour change from yellow to red.

The fractions containing the desired  $\beta$ -lactamase protein were concentrated to a final volume of 5 mL and then loaded onto a Superdex™ 75 GL 10/300 gel filtration column. The protein was eluted from the column in 20 mM Tris-HCl, pH 7.5 at a flow rate of 2 mL/min. 1.5 mL fractions were collected and the elution of protein monitored by absorbance at 280 nm (**Figure 3.7a**). Fractions corresponding to the main peak were assessed for enzyme activity (**Figure 3.7b**), and were subsequently pooled and analysed by SDS-PAGE (**Figure 3.7c**). Protein identity was confirmed by mass spectrometry (observed mass: 28906.8 Da, expected mass: 28906.0 Da).

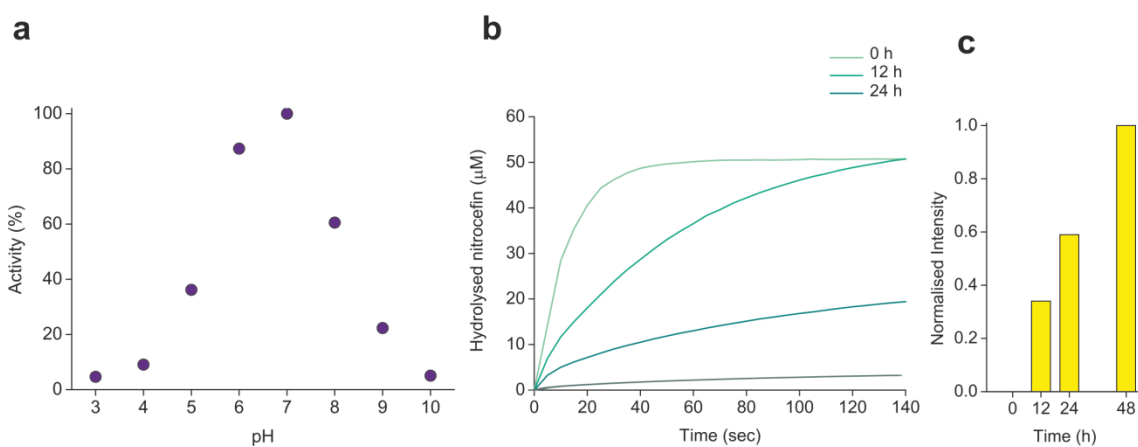


**Figure 3.7.** Size exclusion chromatography of WT  $\beta$ -lactamase. **(a)** Typical WT  $\beta$ -lactamase elution profile from the size exclusion column. Fractions collected are shown on the top x-axis. **(b)**  $\beta$ -lactamase enzyme activity assay of fractions collected during size exclusion chromatography. The colour change upon hydrolysis of nitrocefin was measured at 486 nm. **(c)** SDS-PAGE gel of WT  $\beta$ -lactamase after size exclusion chromatography. The size in kiloDaltons (kDa) of the protein marker is indicated.

### 3.3.2 Reduction in WT $\beta$ -lactamase enzyme activity correlates with aggregation

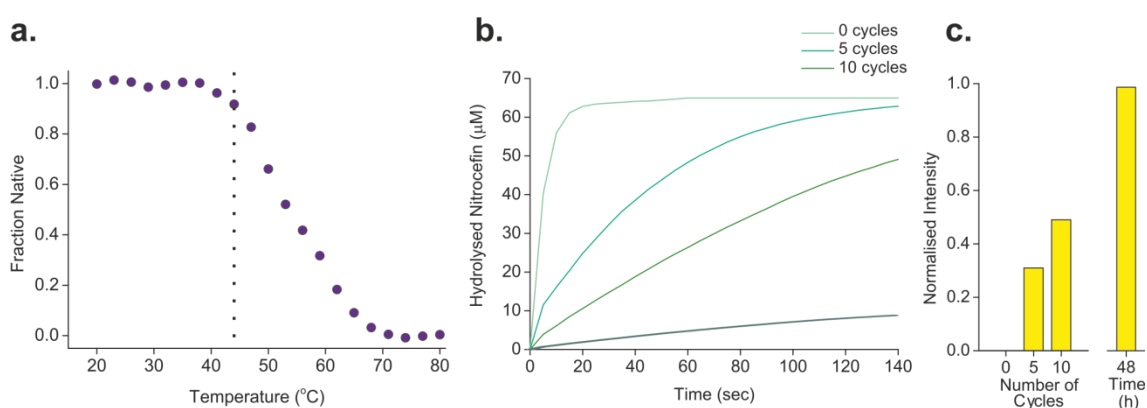
Two methods (isoelectric point (pI) precipitation and repeated freeze-thaw cycles) were employed to aggregate WT  $\beta$ -lactamase so that the utility of using the enzymatic activity of  $\beta$ -lactamase as an indication of its aggregated state could be assessed.

The pI of a protein is the pH at which its net charge is zero. By placing a protein in a solution with a pH the same as its pI, protein solubility is at a minimum<sup>396</sup>. To determine whether the nitrocefin assay could be used to assay  $\beta$ -lactamase activity at pH 5.4 (the pI of WT  $\beta$ -lactamase is 5.4), the pH activity profile of the enzyme was assessed (**Figure 3.8a**).  $\beta$ -lactamase has ~ 40 % activity at pH 5.4, when compared to its activity at pH 7.0, confirming the assay can be used at this pH. The activity of 50  $\mu$ M WT  $\beta$ -lactamase after various time periods at pH 5.4 (200 rpm, 37 °C) was measured (Section 2.5.2) (**Figure 3.8b**). The results indicate that  $\beta$ -lactamase activity decreases over time, in conjunction with increasing turbidity of the sample (Section 2.5.5) (**Figure 3.8c**).



**Figure 3.8.** WT  $\beta$ -lactamase activity after isoelectric point precipitation. **(a)** pH activity profile of nitrocefin hydrolysis by WT  $\beta$ -lactamase. Activity at pH 7.0 is 100 %. **(b)** Enzyme activity assay of 50  $\mu$ M WT  $\beta$ -lactamase after 0 h, 12 h, and 24 h incubation in 100  $\mu$ M citrate- $\text{Na}_2\text{HPO}_4$ , pH 5.4, at 37 °C, 200 rpm. Hydrolysis of nitrocefin was measured at 486 nm over time. **(c)** Endpoint turbidity of the WT  $\beta$ -lactamase after isoelectric point precipitation. Turbidity was normalised to a sample left for 48 h at 37 °C, 200 rpm, pH 5.4.

Analogous results were obtained during aggregation via repeated freeze-thaw-partial thermal denaturation cycles. Freeze-thaw cycles introduce physical and chemical changes, such as the creation of ice/solution interfaces, adsorption of the protein to the container surface and pH changes due to crystallisation of buffer components<sup>35</sup>. Briefly, 50  $\mu\text{M}$  WT  $\beta$ -lactamase was frozen to  $-20\text{ }^{\circ}\text{C}$  in a laboratory freezer for 20 min and then thawed at room temperature. Care was taken not to mix the sample during the thawing stage to ensure the creation of concentration and thermal gradients. Between each cycle, the protein was subjected to partial thermal denaturation, to aid the formation of non-native aggregates. A thermal denaturation curve of WT  $\beta$ -lactamase (following tryptophan fluorescence) was performed to find the temperature at which the protein begins to become unstable and unfolding begins (**Figure 3.9a**).  $44\text{ }^{\circ}\text{C}$  was chosen, and the protein was incubated for 1 h at this temperature between each freeze-thaw cycle. As with isoelectric point precipitation, the experiment revealed a decrease in  $\beta$ -lactamase enzyme activity concurred with increased sample turbidity (**Figure 3.9b, c**).



**Figure 3.9.** WT  $\beta$ -lactamase activity after freeze-thaw-thermal denaturation cycling. **(a)** Thermal denaturation curve of 1  $\mu\text{M}$  WT  $\beta$ -lactamase in 50 mM sodium phosphate, pH 7.0. Black dotted line indicates temperature selected for partial thermal denaturation. **(b)** Enzyme activity assay of 50  $\mu\text{M}$  WT  $\beta$ -lactamase after 0, 5, and 10 cycles of freeze-thaw-thermal denaturation in 50 mM sodium phosphate, pH 7.0. Hydrolysis of nitrocefin was measured at 486 nm over time. **(c)** Endpoint turbidity of the WT  $\beta$ -lactamase after each cycle. Turbidity was normalised to a sample left for 48 h at  $37\text{ }^{\circ}\text{C}$ , 200 rpm.

It can be concluded that  $\beta$ -lactamase enzyme activity decreases in conjunction with aggregation, supporting the hypothesis that this may be used as a read-out of aggregation propensity of an inserted test protein. Furthermore, a purification protocol has been established which can be used as the basis for  $\beta$ -lactamase-test protein purification.

## 3.4 $\beta$ -lactamase system allows identification of aggregation-prone proteins

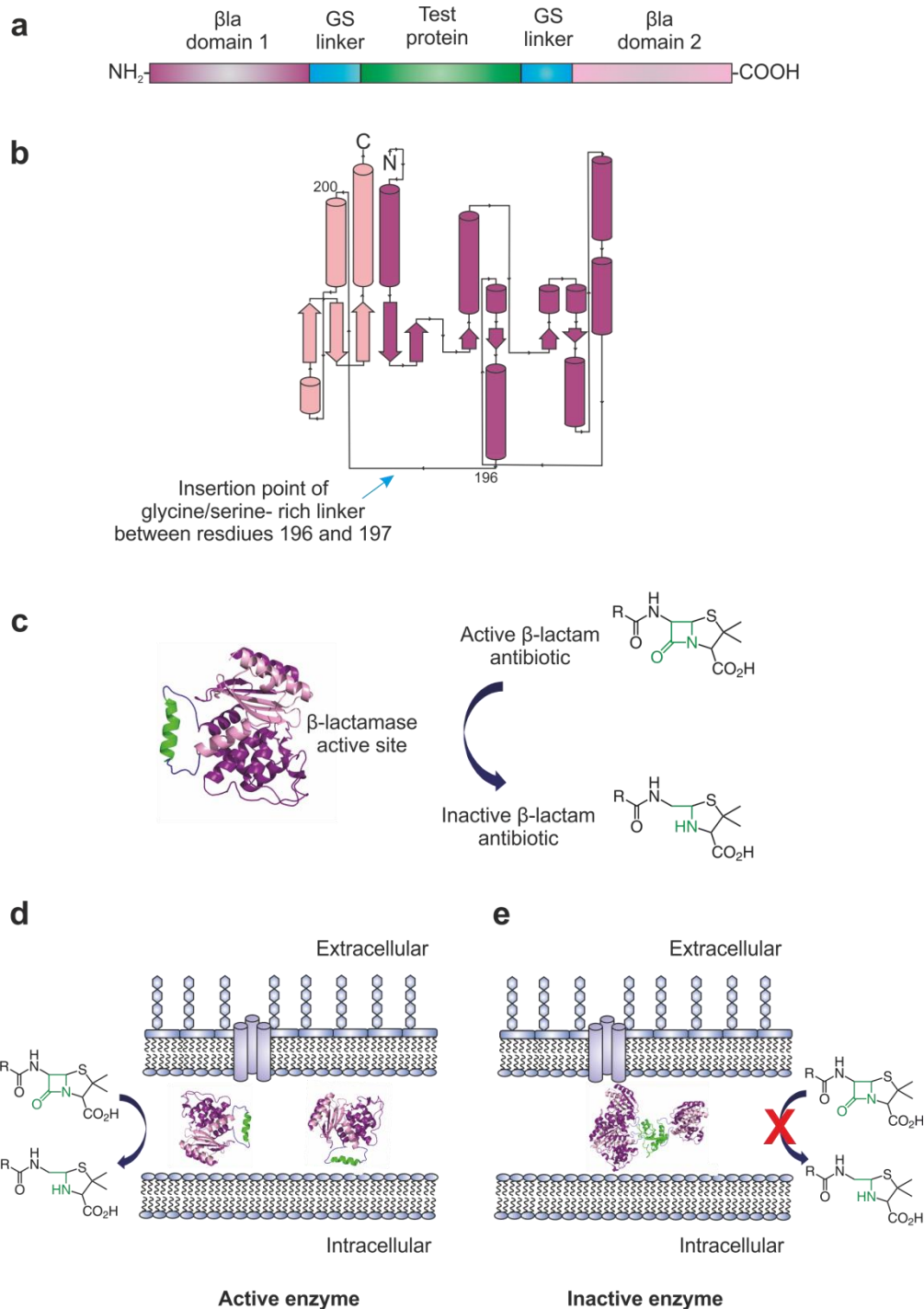
The split  $\beta$ -lactamase aggregation assay exploits the tolerance of the periplasmic antibiotic resistance enzyme TEM-1  $\beta$ -lactamase to the insertion of a test protein into a glycine-serine-rich loop on its surface<sup>308</sup> (**Figure 3.10a**). The assumption of the assay described here is that upon correct folding of the test protein, the two halves of  $\beta$ -lactamase are brought close enough together to associate such that the bacteria will be resistant to  $\beta$ -lactam antibiotics (**Figure 3.10b, c**). If the test protein aggregates, the activity of the  $\beta$ -lactamase is reduced and the bacteria become more sensitive to  $\beta$ -lactam antibiotics (**Figure 3.10d**).

### 3.4.1 Selection of test proteins

To test whether the tripartite fusion system allows antibiotic resistance to be used to monitor protein aggregation *in vivo*, four test proteins with varying degrees of aggregation propensity were selected: hIAPP (human islet amyloid polypeptide), rIAPP (rat islet amyloid polypeptide), A $\beta$ 40 (amyloid  $\beta$  residues 1-40) and A $\beta$ 42 (amyloid  $\beta$  residues 1-42).

#### 3.4.1.1 Human and rat islet amyloid polypeptide

Human islet amyloid polypeptide (hIAPP, also known as amylin) is a 37-residue hormone peptide belonging to the calcitonin gene-related peptide (CGRP) family<sup>397</sup>. It is produced and stored together with insulin in pancreatic islet  $\beta$ -cells<sup>398</sup>. hIAPP has a variety of physiological functions, including the regulation of glucose metabolism in conjunction with insulin and glucagon<sup>397</sup>. Under physiological concentrations of glucose, insulin and hIAPP are secreted by the islet  $\beta$ -cells. However, at elevated glucose concentrations, hIAPP is secreted at much higher levels and can form extensive extracellular  $\beta$ -sheet aggregates<sup>119</sup>. These proteinaceous deposits are found in approximately 95 % of patients suffering from type II diabetes mellitus<sup>399</sup>, a disease characterised by a failure to secrete sufficient insulin due to decreased  $\beta$ -cell mass and function. While the exact role of hIAPP in type II diabetes mellitus is unclear, there is strong evidence that the amyloid fibrils disrupt the integrity of the  $\beta$ -cell membranes, leading to cell death<sup>400</sup>. The amyloid formed

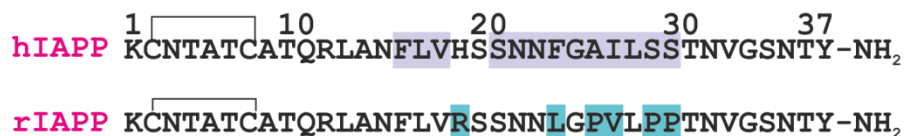


**Figure 3.10.** Schematic of the *in vivo* split  $\beta$ -lactamase assay for protein aggregation. **(a)** The test protein (green) is inserted into a glycine/serine rich linker (blue) within the loop region separating the two domains of the periplasmic enzyme TEM-1  $\beta$ -lactamase (pink and purple). **(b)** Topology of TEM1  $\beta$ -lactamase highlighting insertion point of the glycine/serine-rich linker between residues 196 and 197. Created using online database PDBsum<sup>401</sup> and PDB entry 1BTL<sup>391</sup>. **(c)** Association of the two  $\beta$ -lactamase domains results in the formation of the enzyme active site. **(d)** If the test protein does not aggregate within the periplasm of *E. coli*, the bacteria will be resistant to  $\beta$ -lactam antibiotics. **(e)** If the test protein aggregates, the activity of  $\beta$ -lactamase is reduced and the bacteria become more sensitive to  $\beta$ -lactam antibiotics.

by hIAPP is restricted to the islet area of the pancreas and is not found systemically. The fibrils are not organised within the deposits, but close to  $\beta$ -cells there is often an orientation of bundles of fibrils towards the cell membrane<sup>397</sup>. It is believed that the very first amyloid formed is intracellular which leads to apoptosis of the cell<sup>37</sup>. Once the amyloid is in the extracellular milieu, it seeds further amyloid formation from exocytosed hIAPP<sup>397</sup>.

Full length (37-residue) hIAPP is composed of multiple functional regions, including an N-terminal region (residues 1-19) involved in membrane binding, a primary amyloidogenic region (residues 20-29), and a C-terminal region (residues 30-37) that enhances amyloid formation<sup>402</sup>. Soluble hIAPP has been shown to have an essentially disordered conformation<sup>403</sup>, however it may also assume compact structures and transiently sample an  $\alpha$ -helical conformation<sup>404</sup>. In contrast to other protein aggregation diseases such as Alzheimer's disease, aberrant cleavage of hIAPP is not necessary for fibrillogenesis as only full length hIAPP has been identified in the pancreatic amyloid deposits<sup>37</sup>. However, evidence is mounting that aberrant cleavage of the hIAPP precursor protein, pro-IAPP, may start amyloid deposition by the formation of a nucleus<sup>405</sup>.

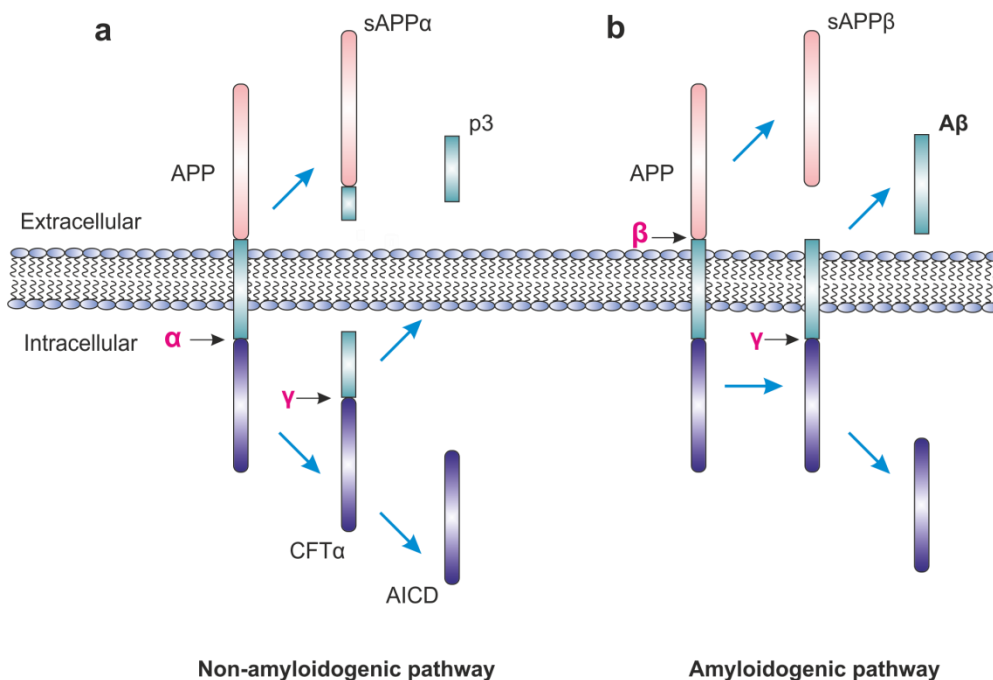
hIAPP was selected for the assays described here primarily because of its rapid aggregation rates compared to other aggregation-prone proteins<sup>37</sup>. Furthermore, IAPP has the advantage of differing slightly from organism to organism<sup>406</sup>. For example, unlike hIAPP (human), rat IAPP (rIAPP) is non-amyloidogenic<sup>37</sup>, thereby providing a convenient control for the prospective aggregation screen. A comparison of the peptide sequences of human and rat IAPP shows six amino acid substitutions (**Figure 3.11**). As five of these occur in region 20-29, this segment of the peptide is known as the amyloidogenic region. Further studies have identified residues 15-17 as also being important for the amyloidogenesis of hIAPP<sup>407</sup>.



**Figure 3.11.** The amino acid sequences of human and rat IAPP. The residues in rIAPP that differ in hIAPP are highlighted in blue. The residues of hIAPP known to play an important role in amyloid fibril formation *in vitro* are highlighted in purple<sup>407</sup>. The intramolecular disulfide bond is indicated by a black line, and the amidated C terminal is shown.

### 3.4.1.2 Amyloid $\beta$

The accumulation of amyloid  $\beta$  ( $A\beta$ ) peptides into amyloid plaques is a central feature of the neurodegenerative disorder Alzheimer's disease<sup>408, 409</sup>.  $A\beta$  is formed via the sequential proteolytic processing of the integral membrane amyloid precursor protein (APP) by the  $\beta$ - and  $\gamma$ -secretases<sup>410</sup>. APP is expressed in many cell types, however the full function of the protein and its subsequent cleavage products are not known. Within the central nervous system, evidence for functions associated with synaptogenesis, cell adhesion, neuronal mobility and transcriptional regulation has been found<sup>411, 412</sup>. APP processing can occur via the non-amyloidogenic pathway or the amyloidogenic pathway (**Figure 3.12**). The predominant pathway is non-amyloidogenic, whereby APP is initially cleaved by  $\alpha$ -secretase and subsequently cleaved by  $\gamma$ -secretase leading to the formation of sAPP $\alpha$ , CTF $\alpha$ , AICD and p3 (**Figure 3.12a**). Conversely, the alternative amyloidogenic route involves initial cleavage by the  $\beta$ -secretase BACE1 enzyme (aspartyl protease  $\beta$ -site APP cleaving enzyme) (**Figure 3.12b**). A shorter variant of the extracellular domain is released



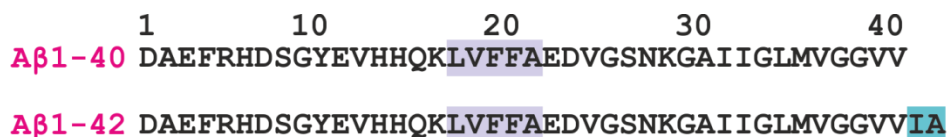
**Figure 3.12.** The non-amyloidogenic and amyloidogenic pathways of APP processing. **(a)** The majority of APP is processed via the non-amyloidogenic pathway. APP is sequentially cleaved by  $\alpha$ - then  $\gamma$ -secretase, resulting in the formation of the non-amyloidogenic peptide fragments sAPP $\alpha$ , p3 and AICD. **(b)** During the amyloidogenic pathway, sequential cleavage by  $\beta$ -secretase then  $\gamma$ -secretase leads to the production of the toxic A $\beta$  peptide. Figure redrawn and adapted from Yates, 2008<sup>413</sup>.



(sAPP $\beta$ ), and the longer transmembrane domain is cleaved by  $\gamma$ -secretase. The  $\gamma$ -secretase can cleave the resulting fragment at several sites, with a preference at positions 40 and 42<sup>410</sup>. The resulting family of A $\beta$  peptides are amphipathic, containing a central hydrophobic region (residues 17-21), a long hydrophobic C terminus, and a charged N terminus (**Figure 3.13**). The central hydrophobic region is known to be critical for amyloid formation<sup>414-417</sup>. The physiological role of the A $\beta$ 40 and A $\beta$ 42 peptides remains unclear, however A $\beta$ 40 has been proposed to regulate the activity of potassium ion channels<sup>418</sup> and to modulate synaptic transmission<sup>419</sup>.

Although A $\beta$ 40 is the predominant form produced during APP cleavage by  $\gamma$ -secretase, A $\beta$ 42 is the most abundant species in the neuritic plaques of Alzheimer's disease. A $\beta$ 42 is much more prone to aggregation than A $\beta$ 40, and is also significantly more toxic to neuronal cells<sup>420, 421</sup>. *In vitro* forms of both A $\beta$  peptides aggregate into well-ordered  $\beta$ -sheet rich structures<sup>422</sup>, in which the dimers and trimers are significantly more toxic than monomeric species<sup>423</sup>. Evidence suggests that the peptides form fibrils via different pathways, demonstrated by the production of distinctly larger oligomeric species by A $\beta$ 42 compared to A $\beta$ 40<sup>424</sup>. Furthermore, A $\beta$ 42 is able to sample significantly more  $\beta$ -rich states en-route to fibril formation<sup>425</sup>. Molecular dynamics simulations have also shown that the C terminus of A $\beta$ 42 is constrained to a  $\beta$ -hairpin structure, thereby increasing the amyloid propensity of the peptide<sup>426</sup>.

The toxicity of A $\beta$  peptides has been attributed to numerous mechanisms, including the production of reactive oxygen species<sup>427</sup>, the alteration of energy metabolism due to damaged mitochondria<sup>428, 429</sup>, the elevation of intracellular calcium levels via membrane damage<sup>144, 430</sup>, the modification of metal homeostasis<sup>431</sup> and the ultimate loss of neurons via apoptotic pathways<sup>413, 432, 433</sup>.



**Figure 3.13.** The amino acid sequences of amyloid  $\beta$  residues 1-40 and 1-42. The extra two residues in A $\beta$ 42 are highlighted in blue. The central hydrophobic region (residues 17-21) is highlighted in purple.

### 3.4.2 Cloning test proteins into $\beta$ -lactamase linker

The four test proteins were cloned into the 28-residue glycine/serine-rich linker that had previously been inserted between residues 196 and 197 of TEM-1  $\beta$ -lactamase<sup>308</sup> (**Figure 3.10a,b**). PCR was performed to amplify the hIAPP, rIAPP and A $\beta$ 42 genes from the pTXB1-hIAPP, pTXB1-rIAPP and pRSET-A $\beta$ 40/42 plasmids respectively (**Table 2.10**) in Section 2.2.10). Primers were designed (**Table 2.8**) to include an *Xho*I restriction site 5' and a *Bam*HI restriction site 3' to the gene, as these are the restriction sites in the  $\beta$ -lactamase 28-residue glycine/serine-rich linker (denoted  $\beta$ la-linker<sub>SHORT</sub>). The plasmid containing  $\beta$ la-linker<sub>SHORT</sub> was digested with the same enzymes, and the PCR product encoding the test protein was ligated into it. The ligation products were transformed into *E. coli* JM109 cells (Section 2.2.7) and the cells grown on agar plates containing 10  $\mu$ g/mL tetracycline. Successful ligation was identified by the resistance to tetracycline obtained from the  $\beta$ -lactamase vector. DNA was purified from a selection of colonies and was sent for sequencing to confirm that they contained the correct sequences (Section 2.2.9). Primers for the sequencing reactions were designed to bind upstream and downstream of the GS linker of  $\beta$ -lactamase (**Table 2.8**). The newly synthesised plasmids were named  $\beta$ la-28-hIAPP,  $\beta$ la-28-rIAPP and  $\beta$ la-28-A $\beta$ 42. The plasmid  $\beta$ la-28-A $\beta$ 40 was created using QuikChange mutagenesis (Agilent Technologies, UK) by deleting the last two residues of the A $\beta$ 42 sequence that had been inserted into the  $\beta$ -lactamase linker (primers in **Table 2.8**). Full gene and protein sequences are provided in **Appendices 7.15 – 7.21**.

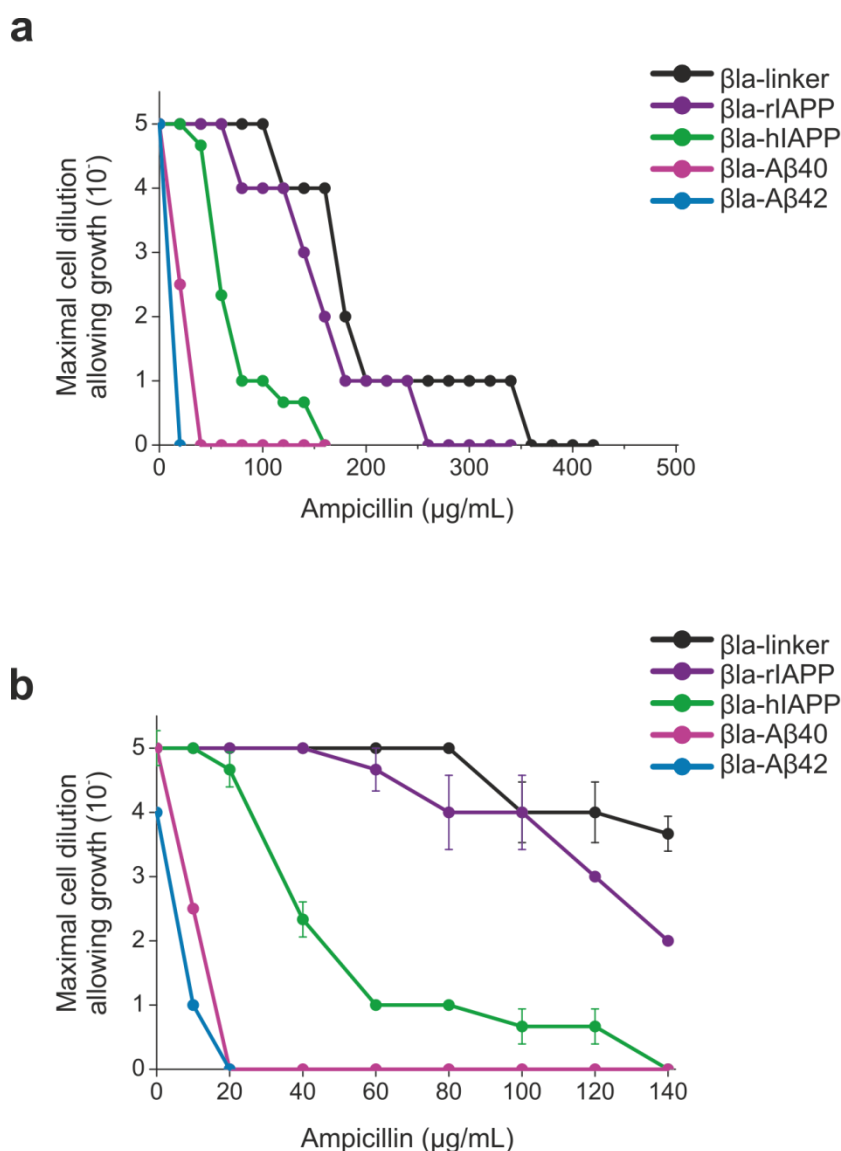
### 3.4.3 Split $\beta$ -lactamase system is amenable to a variety of culture conditions

A large number of parameters were examined for the optimisation of the minimal inhibitory concentration of antibiotic (MIC) assay; they are summarised in **Table 3.1**. The ability of the split  $\beta$ -lactamase system to discriminate between aggregating and non-aggregating test proteins was found to be relatively insensitive to parameters such as cultivation media, level of induction, incubation time, or  $\beta$ -lactam antibiotic, thereby allowing flexible adaption of the system for any specific conditions required depending on the test protein. Cells were pre-incubated for a range of times in the presence of 0.02 % (*w/v*) arabinose to induce protein expression prior to spot titrations on the agar plates. Plates were incubated for a variety of incubation times (12 h, 16 h, 18 h, or 24 h) before the maximal cell dilution at which growth could occur was scored. The  $\beta$ -lactam

antibiotic penicillin V was also tested, along with the cultivation medium terrific broth. It was immediately evident that the cells required arabinose to be present in both the culture and agar medium for colonies to grow. Furthermore, a pre-incubation time of 1 h in the presence of arabinose was also essential for bacterial growth. This suggests  $\beta$ -lactamase induction is essential prior to bacterial exposure to  $\beta$ -lactam antibiotics. Different ampicillin concentrations were also investigated with the aim to find the concentration range that gave the clearest distinction between the aggregating and non-aggregating constructs, while using the minimal number of agar plates (**Figure 3.14**). In summary, for the four test proteins selected, a pre-incubation time of 1 h in the presence of 0.02 % (w/v) arabinose in LB medium prior to incubation on agar plates containing 0-140  $\mu\text{g}/\text{mL}$  ampicillin (20  $\mu\text{g}/\text{mL}$  increments) was chosen.

Parameter	Condition	Score	Conclusion
Cultivation media	Terrific broth	***	Both produced positive results, however colonies are easier to visualise with LB.
	LB	****	
Location of arabinose	Culture only	**	Arabinose in the agar plates was essential for any bacterial growth. Addition of arabinose to both culture and the plate enables more colonies to grow, thereby aiding colony-visualisation for scoring.
	Plate only	*	
	Culture and plate	****	
$\beta$ -lactam antibiotic	Ampicillin	****	Both produced positive, reproducible results, however ampicillin is more economical to purchase.
	Penicillin V	****	
Range of ampicillin concentrations ( $\mu\text{g}/\text{mL}$ )	0 – 500	**	0–140 $\mu\text{g}/\text{mL}$ enables clear resolution between the growth of bacteria expressing the aggregating and non-aggregating constructs. 0-500 $\mu\text{g}/\text{mL}$ required a large number of agar plates, and the bacteria expressing the highly aggregating constructs could not grow above $\sim 150 \mu\text{g}/\text{mL}$ ampicillin.
	0 – 140	****	
Sample incubation time with arabinose prior to addition to agar plate	0 h	*	Pre-incubation in the presence of arabinose was required for colonies to grow under selected growth conditions. Optimal time was 1 h.
	1 h	****	
	2 h	***	
	3 h	**	

**Table 3.1.** Parameters optimised for MIC assay. Conditions are scored from the most (four stars) to the least (one star) favourable.



**Figure 3.14.** Comparison of MIC assays over two ranges of ampicillin concentration. Maximal cell dilution allowing growth ( $\text{MCD}_{\text{GROWTH}}$ ) over (a) a wide range (0-500  $\mu\text{g/mL}$ ), or (b) a restricted range (0-140  $\mu\text{g/mL}$ ), of ampicillin concentrations for bacteria expressing  $\beta\text{la-linker}$  (●),  $\beta\text{la-riIAPP}$  (●),  $\beta\text{la-hIAPP}$  (●),  $\beta\text{la-A}\beta\text{40}$  (●) or  $\beta\text{la-A}\beta\text{42}$  (●) tripartite fusion constructs. Error bars represent the standard error from a minimum of 3 replicates.

### 3.4.4 Antibiotic resistance correlates with aggregation-propensity

To validate the hypothesis that the activity of  $\beta$ -lactamase will be reduced *in vivo* when fused to an aggregation-prone protein, the minimal inhibitory concentration of antibiotic was assessed for the split  $\beta$ -lactamase-test protein constructs. The concept of this assay was that on induction of expression of the  $\beta$ -lactamase-test protein, the bacteria would fail to grow in the presence of the  $\beta$ -lactam antibiotic ampicillin in a quantitative manner, depending on the aggregation-propensity of the test protein.

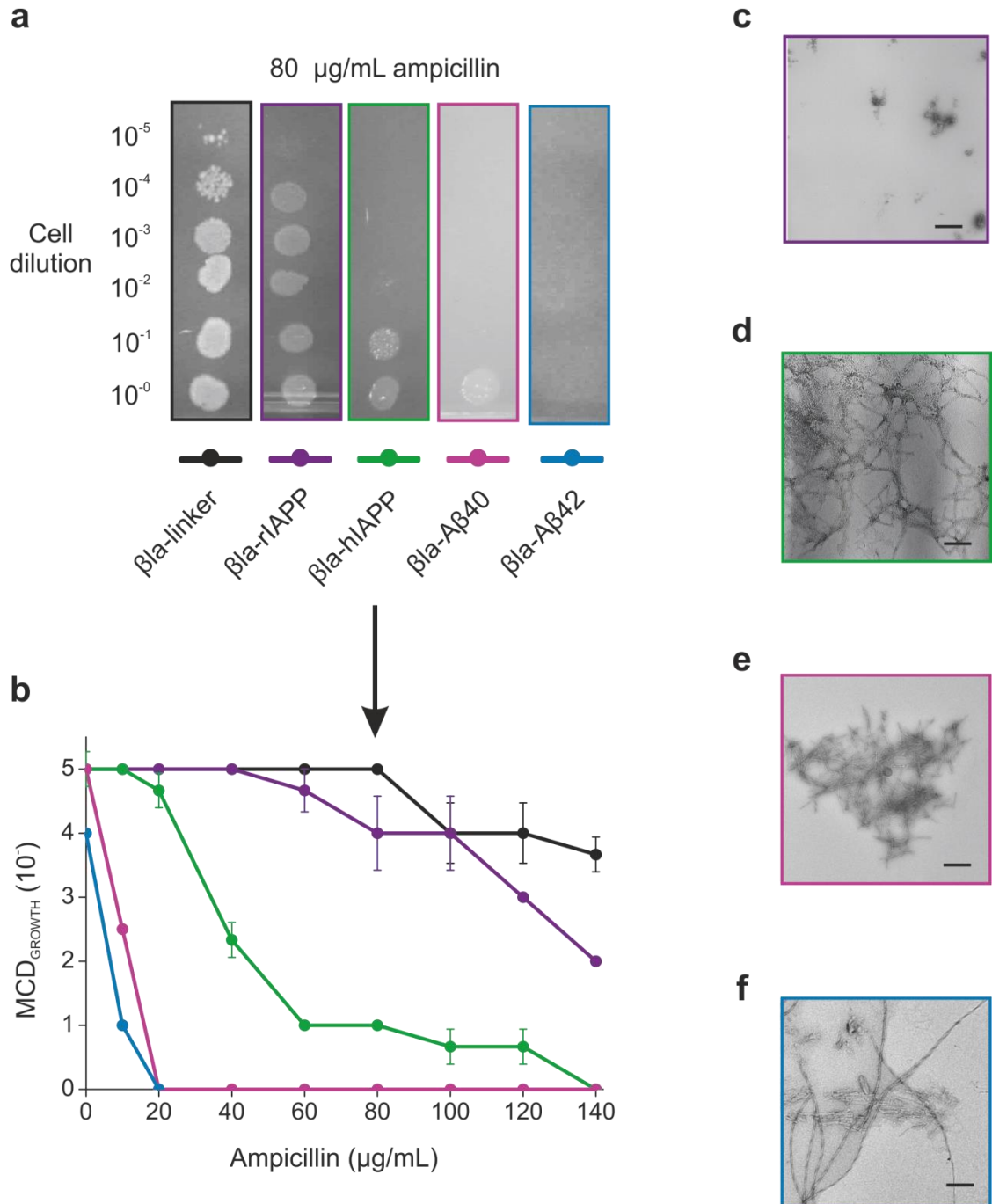
In this assay, the optimal conditions as described in Section 3.4.3 were utilised. Briefly, cultures of BL21 (DE3) cells expressing each construct were grown until the  $OD_{600} = 0.6$  (37 °C, 200 rpm). Protein expression was then induced by the addition of 0.02 % (w/v) arabinose and bacteria were allowed to grow for a further 1 h (37 °C, 200 rpm), before sequential cell dilutions were pipetted onto six circular agar plates containing increasing concentrations of the  $\beta$ -lactam antibiotic ampicillin (0–0.5 mg/mL; 0.1 mg/mL increments). Plates were incubated at 37 °C for 18 h, after which they were scored for the maximal cell dilution at which cells could grow at each antibiotic concentration. To enable comparison of the test protein constructs to a non-aggregating control, growth of *E. coli* expressing a  $\beta$ -lactamase construct containing only a 64-residue glycine/serine linker (a similar length to the inserted test proteins, **Table 3.2**) was also analysed ( $\beta$ la-linker). The spot titre comparing the growth of the strains expressing the different constructs at 80  $\mu$ g/mL ampicillin is shown in **Figure 3.15a**. Over a wide range of antibiotic concentrations, the bacteria expressing the non-aggregating glycine-serine linker ( $\beta$ la-linker) and rIAPP ( $\beta$ la-rIAPP) constructs could grow at significantly higher concentrations of ampicillin than the strains expressing the aggregating constructs ( $\beta$ la-hIAPP,  $\beta$ la-A $\beta$ 40 or  $\beta$ la-A $\beta$ 42) (**Figure 3.15b**). As hIAPP, A $\beta$ 40 and A $\beta$ 42 are intrinsically disordered in solution before aggregation<sup>425, 426, 434</sup>, these results demonstrate that the reduced enzymatic activity of  $\beta$ -lactamase must be due to aggregation and not solely the presence of an extended loop region separating the two domains of  $\beta$ -lactamase.

Construct	Length of Test Protein (Residues)	Length of GS-Linker (Residues)	Total Length (Residues)
$\beta$ la-linker	-	64	64
$\beta$ la-rIAPP	37	28	65
$\beta$ la-hIAPP	37	28	65
$\beta$ la-A $\beta$ 40	40	28	68
$\beta$ la-A $\beta$ 42	42	28	70

**Table 3.2.** Comparison of  $\beta$ -lactamase construct linker lengths.

The *in vitro* aggregation propensity of the inserted test proteins in the absence of the tripartite fusion was assessed using transmission electron microscopy (TEM) and electrospray ionisation ion mobility mass spectrometry (ESI-IMS). Rachel A. Mahood purified the A $\beta$ 40 peptide and performed the transmission electron microscopy on the sample. The hIAPP and rIAPP peptides were kindly provided by Cynthia Tu and Professor Daniel Raleigh (Stony Brook, NY). Synthetic A $\beta$ 42 was purchased from Invitrogen, Paisley, UK (Catalogue Number 03-112). Lydia M. Young performed the transmission electron microscopy of hIAPP, rIAPP and A $\beta$ 42.

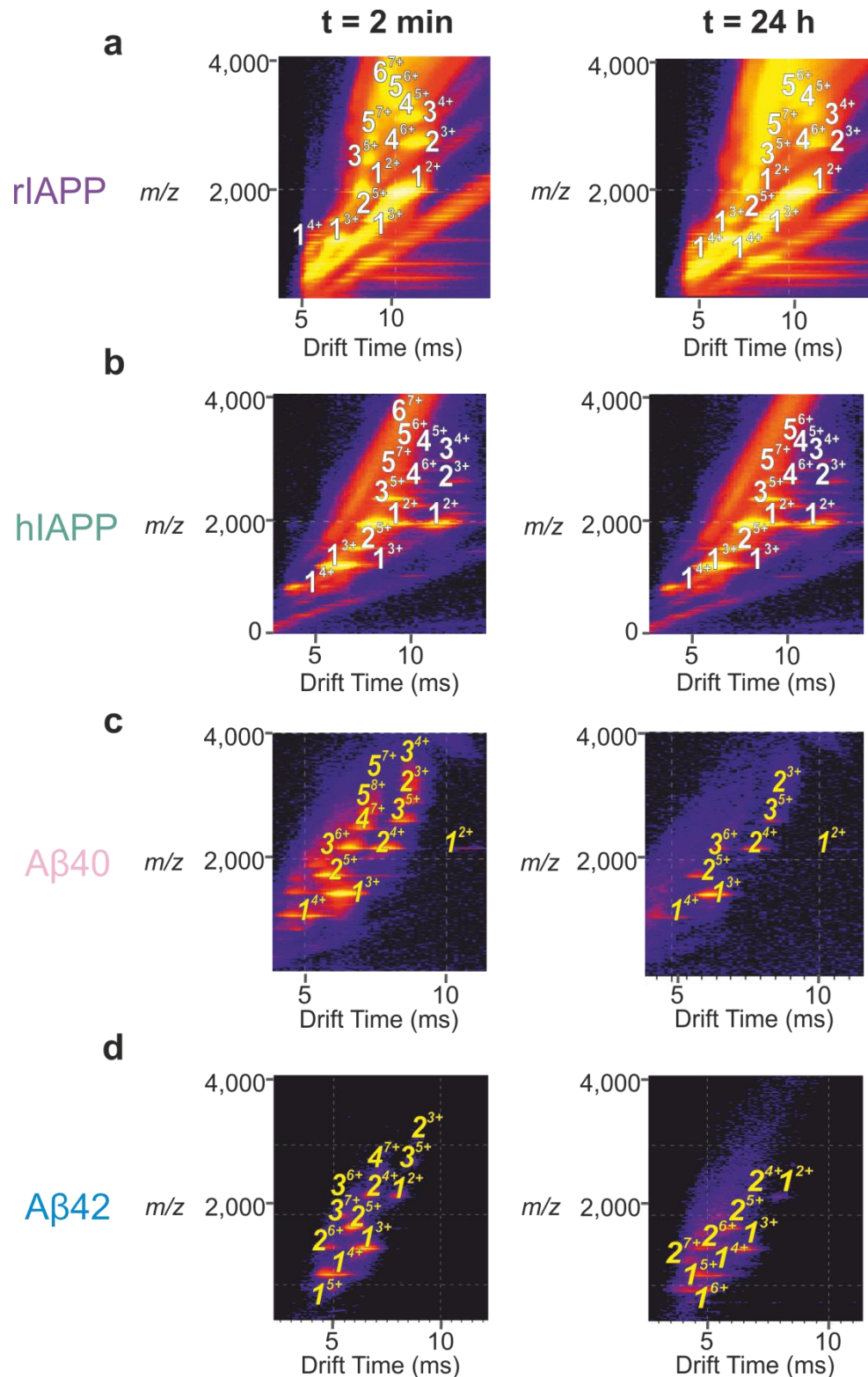
TEM images of aggregates formed after 5 days incubation (quiescent, 33  $\mu$ M peptide, 200 mM ammonium acetate buffer, pH 6.8, 1 % (w/v) DMSO) (Section 2.5.6) show a striking correlation between the ability to form amyloid fibrils observed by TEM and the level of antibiotic resistance *in vivo*: the more prevalent the aggregates (**Figure 3.15c-f**), the lower the resulting antibiotic resistance (**Figure 3.15b**). Importantly rIAPP, which is known to not form amyloid fibrils<sup>237</sup>, only forms small oligomeric species, as observed by TEM (**Figure 3.15c**). Conversely, the three amyloidogenic proteins hIAPP, A $\beta$ 40 and A $\beta$ 42 form long fibrils and amorphous aggregates (**Figure 3.15d-f**).



**Figure 3.15.** *In vitro* aggregation correlates with *in vivo* antibiotic resistance. Each tripartite fusion protein is denoted by a different colour. **(a)** Spot titre from a single plate containing 80  $\mu\text{g}/\text{mL}$  ampicillin. **(b)** Maximal cell dilution allowing growth over a wide range of ampicillin concentrations for each of the tripartite fusion constructs. Error bars represent the standard error from a minimum of 3 replicates. Negative stain TEM images represent the aggregates formed by 33  $\mu\text{M}$  rIAPP **(c)**, hIAPP **(d)**, A $\beta$ 40 **(e)** or A $\beta$ 42 **(f)** after 5 days incubation at 25  $^{\circ}\text{C}$ , quiescent (scale bar = 100 nm). TEM images acquired by Lydia M. Young and Rachel A. Mahood.

The progress of aggregation of each of the four peptides was also compared using an ESI-IMS-MS time-course experiment. All mass spectrometry was performed by Lydia M. Young. Peptide samples were incubated in parallel (50  $\mu$ M peptide, 200 mM ammonium acetate buffer, pH 6.8, 1 % (w/v) DMSO) and at specified time points (2 min and 24 h), 10  $\mu$ L volumes were removed from each solution and infused into the mass spectrometer for analysis (see Section 2.5.8). The Driftscope plots represent a snap shot of the species present in solution at any given point during the aggregation pathway. At  $t = 2$  min, monomer through hexamer were observed for both hIAPP and rIAPP (**Figure 3.16a, b**). These same oligomers, albeit at a slightly lower intensity, were observed after 24 h, suggesting that although aggregation into fibrils is occurring, the reaction has not yet reached completion. In the case of A $\beta$ 40, at  $t = 2$  min, monomer through pentamer were observed (**Figure 3.16c**). After 24 h, aggregation was occurring, resulting in the consumption of higher order species and all that remained observable by ESI-IMS-MS were the lower order species, monomer through trimer (**Figure 3.16c**). For A $\beta$ 42, the most aggregation prone of the four peptide sequences, aggregation occurred at such a rate that it was difficult to observe by ESI-IMS-MS. As early as  $t = 2$  min (**Figure 3.16d**), no higher order oligomers were observed, suggesting that these higher order species exist briefly en route to aggregation into large oligomeric species.





**Figure 3.16.** *In vitro* aggregation correlates with *in vivo* antibiotic resistance. ESI-IMS-MS Driftscope plot of rIAPP (a), hIAPP (b), A $\beta$ 40 (c) and A $\beta$ 42 (d) present 2 min and 24 h after diluting the monomer to a final peptide concentration of 50  $\mu$ M in 200 mM ammonium acetate, pH 6.8. Monomers (1) through to hexamers (6) are labelled. ESI-IMS-MS Driftscope plots show IMS drift time versus  $m/z$  versus intensity ( $z =$  square root scale). Experiment performed by Lydia M. Young.

### 3.5 Discussion

Understanding the underlying reasons why proteins aggregate requires challenging and costly purification prior to biophysical analyses *in vitro*. It is often necessary to purify a selection of mutational variants to begin to understand the extent to which each individual mutation affects protein stability and aggregation. Using an *in vivo* system to investigate aggregation propensity would circumvent the necessity to purify a large number of variants, thereby reducing both experimental time and costs.

The decision to perform the assay on solid media rather than in liquid cultures was reached after extensive discussion with the collaborators who designed and performed the initial *in vivo* assay for protein stability<sup>308</sup>. The authors investigated monitoring bacterial growth (in the presence of the  $\beta$ -lactam antibiotic penicillin V) in liquid growth medium by measuring the change in sample optical density (OD) over time<sup>312</sup>. They found that a number of destabilising mutants caused an increase in lag time, when compared to a stable variant. However, the correlation between stability and the length of the lag time was weak. Furthermore, some variants increased the lag time, but after a sufficient time of incubation reached approximately the same final OD. As the correlation between variant and stability was so strong using the solid medium MIC assay, their advice was to develop the proof-of-principle aggregation assay on a solid medium.

Performing the assay over a wide range of cell dilutions was necessary to obtain the sensitivity required to differentiate between the different aggregating constructs. Although the use of a single cell dilution would simplify the work required to perform the assay, it would not produce the fine distinction between proteins with varying aggregating propensities (see Figure 3.14). However, in attempts to produce a semi-high throughput screen, in Chapter 4 an alternative approach is described using a lawn of bacteria (in essence, a single cell dilution) to investigate whether removing a variable from the assay would aid in assay optimisation (see Section 4.2.2).

Development of the *in vivo*  $\beta$ -lactamase tripartite fusion system progressed as a cycle. Variables were tested and the parameters that give the optimal 'reading window' were fixed, after which further parameters were tested under the prior fixed conditions. As the overall aim of this project was to develop a simple, usable assay, a criterion at each optimisation step was to utilise minimal reagents and reduce the number of steps required to perform the assay. One of the most critical steps was controlling protein induction. It was quickly apparent that protein induction must occur a minimum of 1 h

before the bacteria were exposed to antibiotic. Once this had been established, all other parameters were varied until conditions were found with which reproducible results could be obtained.

The correlation of *in vivo* antibiotic resistance and *in vitro* aggregation extent was striking. The three test proteins that rapidly form amyloid fibrils *in vitro* all significantly reduced the bacteria's ability to grow in the presence of antibiotics, when they were expressed as part of a  $\beta$ -lactamase tripartite fusion. Notably, however, the presence of rIAPP in the linker region of  $\beta$ -lactamase did not markedly reduce bacterial growth, compared to the non-aggregation  $\beta$ la-linker. This was vital for assay validation, as although only six residues distinguish rIAPP from hIAPP, the fact that it did not impede bacterial growth suggests the assay is sensitive to the aggregation-extent of the inserted test protein. Interestingly, the extent of bacterial growth impedance correlated with the initial progression of aggregation over 24 h, as observed by ESI-IMS-MS. A $\beta$ 42 oligomers were the fastest to disappear in the progression to larger oligomers, followed by A $\beta$ 40 then hIAPP. rIAPP oligomers, however, were still present after 24 h. This appears to correlate with the observation that growth of bacteria expressing  $\beta$ la-A $\beta$ 42 was lowest, followed by  $\beta$ la-A $\beta$ 40,  $\beta$ la-hIAPP then  $\beta$ la-rIAPP. Conversely, however, fibril formation, as observed by ThT, is fastest for hIAPP<sup>37, 237</sup>. Furthermore, when hIAPP and A $\beta$ 42 are mixed in equimolar ratios, the rate of fibril formation and membrane permeabilisation occurs at a rate intermediate of that observed for hIAPP or A $\beta$ 42 alone<sup>435</sup>. hIAPP-A $\beta$ 42 hetero-aggregates adsorb, aggregate, and permeabilise membranes significantly more slowly than pure hIAPP, but at a much faster rate than observed for pure A $\beta$ 42<sup>435</sup>.

The disparity between the rate of oligomer disappearance (as observed by ESI-IMS-MS) and the appearance of amyloid fibrils (as observed by ThT fluorescence) of hIAPP and A $\beta$ 42 may be due to competing pathways en-route to fibril formation. A vast array of pre-amyloid oligomeric species can be formed on- or off-pathway during amyloid formation, each of which is thermodynamically distinct<sup>436</sup>. The formation of more off-pathway oligomeric species during A $\beta$ 42 fibril formation may account for its longer fibrillation time as compared to hIAPP<sup>437</sup>. Conversely, differences in the rate of fibril formation may be due to the rate at which the high-energy oligomeric nucleus forms. The critical nucleus, defined as the least thermodynamically stable species, is the smallest oligomer capable of initiating further growth<sup>437, 438</sup>. Formation of this high-energy species acts as a bottleneck that limits the rate of fibril formation<sup>439, 440</sup>. Consequently, although initial A $\beta$ 42 oligomerisation occurs much more rapidly than hIAPP, subsequent oligomerisation, including the formation of the A $\beta$ 42 critical oligomeric nucleus required for initialising

fibril formation, may be less thermodynamically favourable than formation of the nucleus for hIAPP fibril formation.

In context of the *in vivo* assay for protein aggregation, the correlation between *in vivo* bacterial resistance and initial aggregation rates (but not fibrillation rates) *in vitro* suggests the assay is an accurate indicator for the initial stages of aggregation, but not the later formation of amyloid fibrils. However, comparisons between aggregation rates *in vivo* and *in vitro* are fraught with difficulties, as the formation of amyloid *in vivo* is complicated by molecular crowding effects, the presence of membrane surfaces, and even molecular chaperones. Furthermore, all the *in vitro* experiments were performed on the isolated aggregation-prone sequences. *In vivo*, these peptides are constrained within the linker of  $\beta$ -lactamase. As such, the question arises: can these peptides still form fibrils when expressed as a tripartite-fusion protein? It is possible that the disparity between the rate of fibril formation observed *in vitro* and the *in vivo* bacterial resistance may simply be due to the inability of the amyloidogenic peptides to fibrillate *in vivo* under the conditions they are subjected to in this assay. As such, the work in Chapter 4 aims to address this question by investigating whether fibrils are formed by a  $\beta$ -lactamase tripartite fusion protein *in vitro*. The purification protocol developed in this chapter was utilised as the basis for purifying the  $\beta$ -lactamase-linker constructs *in vitro*.

A key difference between the *in vivo* screen described here, and previously developed *in vivo* screens for aggregation propensity, is the location of the assay in the periplasm of *E. coli*. The oxidising environment of the periplasm allows the formation of disulfide bonds, unlike the reducing environment of the cytosol of bacteria or the cytoplasm of eukaryotes. Previous cytoplasmic assays, for example using A $\beta$ 42 fused to GFP<sup>302</sup>, have successfully identified aggregation-enhancing mutations; however this assay would be redundant for proteins that require disulfide bond formation for successful folding. Disulfide bonds are present in 15 % of the human proteome and they are enriched in secreted proteins (65 %) <sup>441</sup>. This is presumed to be due to the need for greater protein stability in the absence of the quality control systems present within the cell. Notably, however, they are present in 55 % of proteins involved in pathologic amyloid formation <sup>442</sup>. Consequently, the cytoplasmic-based assays described previously <sup>299, 301, 303</sup> would be limited to screening only those 45 % of amyloidogenic proteins that do not contain disulfide bonds.

The reproducibility of the data demonstrates the application of the split  $\beta$ -lactamase system as a robust *in vivo* screen for aggregation propensity. As the test proteins used are

of a similar length and all are intrinsically disordered in solution, for this system to be a generic indicator of protein aggregation it would need to be applicable to a range of different proteins, varying in both structure and size. The previous studies using this system have demonstrated the ability to insert large proteins (maltose binding protein, 43 kDa) into the linker region of  $\beta$ -lactamase. This suggests that screening for the aggregation-propensity of large and/or structured proteins would also be possible using the  $\beta$ -lactamase tripartite fusion system. This possibility is investigated in Chapter 5, when the aggregation-propensity of two human derived antibody domains is assessed.

In summary, the data demonstrate the application of the split  $\beta$ -lactamase system as a robust *in vivo* screen for aggregation propensity of a variety of test proteins. The correlation of *in vivo* antibiotic resistance and *in vitro* aggregation rates of the aggregation-prone sequences validate the suitability of this system to accurately assess the extent of protein aggregation. Furthermore, the assay works under a variety of conditions, thereby enabling flexibility in the experimental conditions, as required.



## 4 Developing a phenotypic small molecule screen

### 4.1 Objectives

In Chapter 3 it was demonstrated that, by using the  $\beta$ -lactamase tripartite fusion protein, bacterial resistance to antibiotics could be used to identify protein aggregation propensity. The next objective, therefore, was to develop an *in vivo* screen using this system to identify small molecule inhibitors of protein aggregation. The assay needs to be robust, reproducible, simple to perform and have potential to be converted to high throughput methodology.

Developing a cell-based phenotypic assay for protein aggregation inhibitors requires a number of factors to be taken into account, compared with more simple *in vitro*-based assays.

Questions to be addressed:

- How robust is the assay?
- Can off-phenotype effects be accounted for (e.g. cytotoxicity)?
- How robust is the data analysis?
- Is it amenable to miniaturisation for high throughput screening?

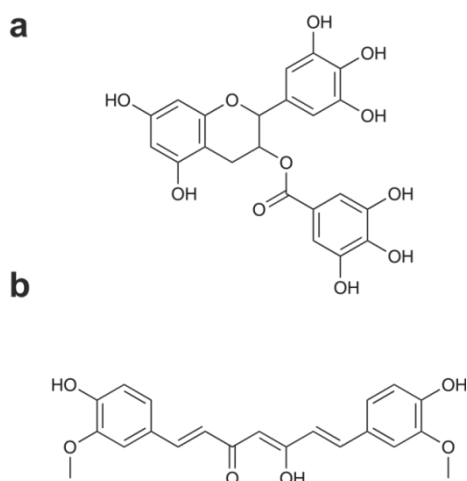
### 4.2 Assay development

#### 4.2.1 Identification of positive and negative reference controls

One of the first steps in any assay development is the identification of suitable positive and negative reference controls. Positive controls must give an assay signal which would be considered a 'hit', whereas the negative control must give no, or minimal, signal. For this assay, the ideal positive reference control would be a potent small molecule inhibitor of aggregation, while the negative control would be the assay in the absence of an inhibitor, or a small molecule known not to have an effect.

To examine the ability of the  $\beta$ -lactamase tripartite system as a screen to identify protein aggregation inhibitors, the assay was first carried out using the known inhibitors of A $\beta$ 42

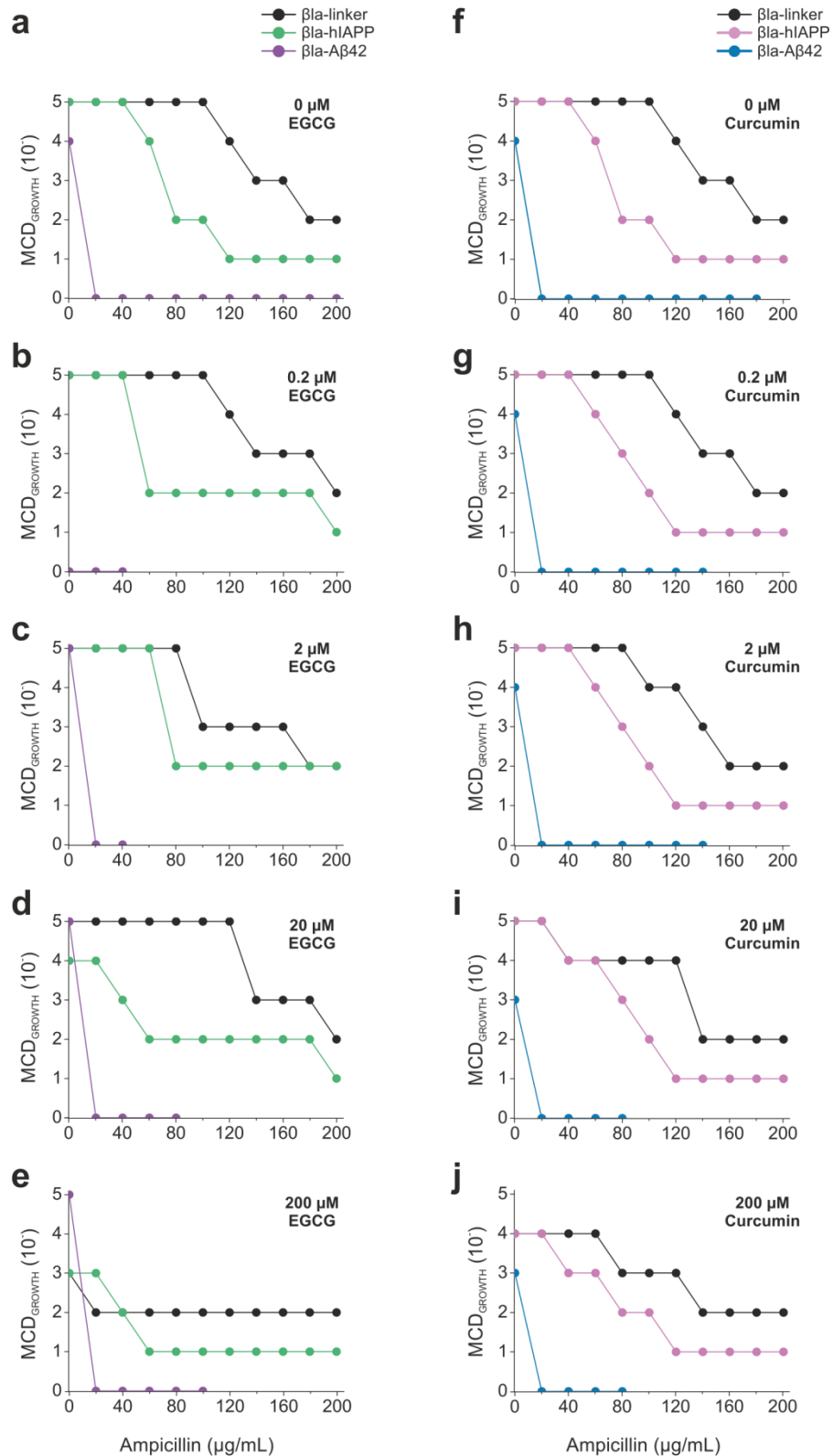
and hIAPP aggregation: (-)-epigallocatechin gallate (EGCG) (**Figure 4.1a**) and curcumin (**Figure 4.1b**). Systematic IUPAC names, structures, molecular weights and LogP values of all small molecules used in this chapter are also given in **Appendix 7.4**. EGCG has been demonstrated to prevent aggregation of many amyloidogenic and aggregation-prone proteins, including A $\beta$ , hIAPP,  $\kappa$ -casein, calcitonin, insulin, huntingtin and transthyretin<sup>227, 234, 235, 237, 239, 240, 255, 256, 443</sup>. The small molecule binds to native disordered A $\beta$ 42 and promotes the formation of non-toxic oligomers<sup>240</sup>. EGCG inhibits hIAPP aggregation *in vitro*, and even disaggregates pre-formed fibrillar material<sup>235</sup>. Similarly, curcumin has also been shown to prevent the *in vitro* aggregation of multiple proteins, including A $\beta$ , transthyretin,  $\alpha$ -synuclein, lysozyme, and hIAPP<sup>238, 243-248, 444-446</sup>. NMR studies indicate that curcumin binds to the  $\beta$ -sheet structure of A $\beta$  fibrils and interrupts the aggregation process<sup>446</sup>. Conversely, curcumin modulates hIAPP self-assembly by binding to, and disassembling, non-native  $\alpha$ -helices<sup>246</sup>.



**Figure 4.1.** Structures of (-)-epigallocatechin gallate (EGCG) (a) and curcumin (b).

Preliminary investigations using the *in vivo* assay determined that the small molecules were required to be present in the agar, and not only the cultures, prior to plating. If they were absent from the agar, no beneficial effect on bacterial growth was observed. Increasing concentrations of the compounds (0 – 200  $\mu$ M) were mixed into agar plates containing ampicillin concentrations from 0 – 200  $\mu$ g/mL. Sequential dilutions of cells expressing the aggregating constructs were then spotted on top (Section 2.4.2). The plates were incubated for 18 h at 37  $^{\circ}$ C and the maximal cell dilution at which cells could grow ( $MCD_{GROWTH}$ ) was assessed at each ampicillin concentration for each concentration of small molecule (**Figure 4.2**).

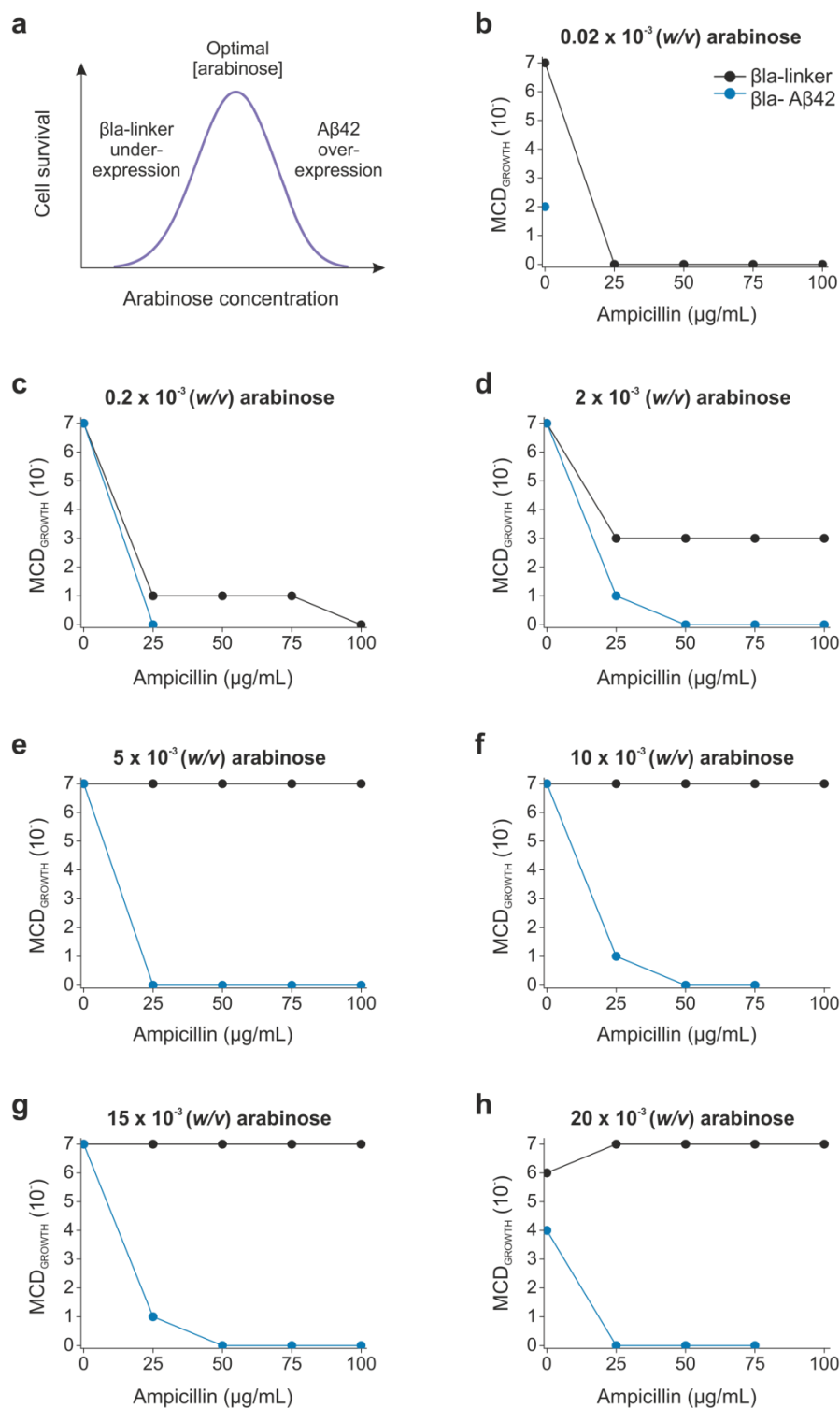




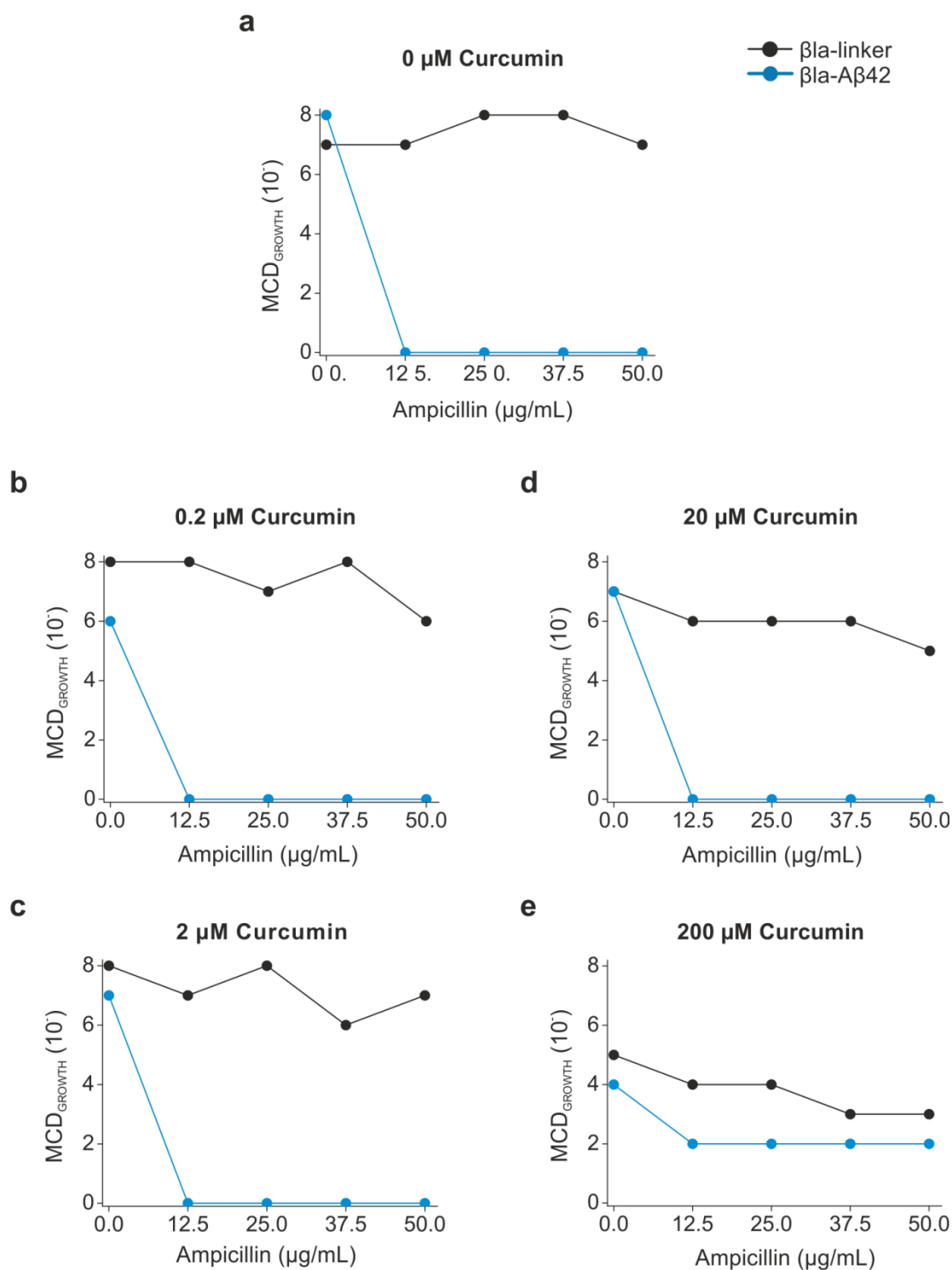
**Figure 4.2.** Bacterial growth in presence of small molecule inhibitors of aggregation. The maximal cell dilution allowing growth ( $\text{MCD}_{\text{GROWTH}}$ ) was scored at each ampicillin concentration for each concentration of (a-e) EGCG ( $\beta\text{la-linker}$  (●),  $\beta\text{la-hIAPP}$  (●),  $\beta\text{la-A}\beta\text{42}$  (●)) and (f-j) curcumin ( $\beta\text{la-linker}$  (●),  $\beta\text{la-hIAPP}$  (●),  $\beta\text{la-A}\beta\text{42}$  (●)).  $\beta\text{la-A}\beta\text{42}$  experiments were performed with undergraduate student Joseph Casson.

An increase in EGCG concentration led to a significant decrease in the difference between the growth of bacteria expressing  $\beta$ la-hIAPP and bacteria expressing  $\beta$ la-linker. (**Figure 4.2a-e**). However at 200  $\mu$ M EGCG, the compound begins to exhibit significant toxicity to cells, as shown by its effect on  $\beta$ la-linker (**Figure 4.1e**). A similar result was observed in the presence of increasing concentrations of curcumin (**Figure 4.2f-j**). Conversely, EGCG and curcumin appeared to have no positive effect (increased MIC) when administered to cells expressing the  $\beta$ la-A $\beta$ 42 construct (**Figure 4.2a-j**).

It was hypothesised that extensive aggregation of the A $\beta$ 42 insert may have severely hindered cell growth. If the concentration of arabinose was reduced, A $\beta$ 42 levels would decrease and therefore cell survival should increase. This improvement would continue until the lack of  $\beta$ -lactamase enzyme becomes detrimental for cellular growth with antibiotic selection (**Figure 4.3a**). To examine whether this was the case, varied arabinose concentrations were used to induce construct expression to different levels and the relationship between protein expression level and cell survival was assessed. A much lower range of ampicillin concentrations was also used. As expected, at the lowest concentration of arabinose added ( $0.02 \times 10^{-3}$  % (w/v), **Figure 4.3b**), neither bacteria expressing the  $\beta$ la-A $\beta$ 42 or  $\beta$ la-linker construct could survive. As the arabinose concentration was increased, growth of cells expressing the  $\beta$ la-linker construct increased in a titratable manner (**Figure 4.3c-g**), until a concentration of arabinose was reached at which cell survival decreased (**Figure 4.3h**). A similar, but less dramatic, effect was observed for cells expressing  $\beta$ la-A $\beta$ 42 (**Figure 4.3c-h**). Based on these results, a concentration of 0.015 % (w/v) arabinose was considered to produce the optimal growth conditions (**Figure 4.3g**), and this concentration was used subsequently to assay the effect of increasing concentrations of the small molecule curcumin on cell survival (**Figure 4.4**). In the presence of 200  $\mu$ M curcumin, bacteria expressing  $\beta$ la-A $\beta$ 42 could grow at significantly higher concentrations of ampicillin when compared to growth in the absence of the curcumin (compare **Figure 4.4a** and **4.4e**). All subsequent assays using the  $\beta$ la-A $\beta$ 42 construct were performed, therefore, using 0.015 % (w/v) arabinose.



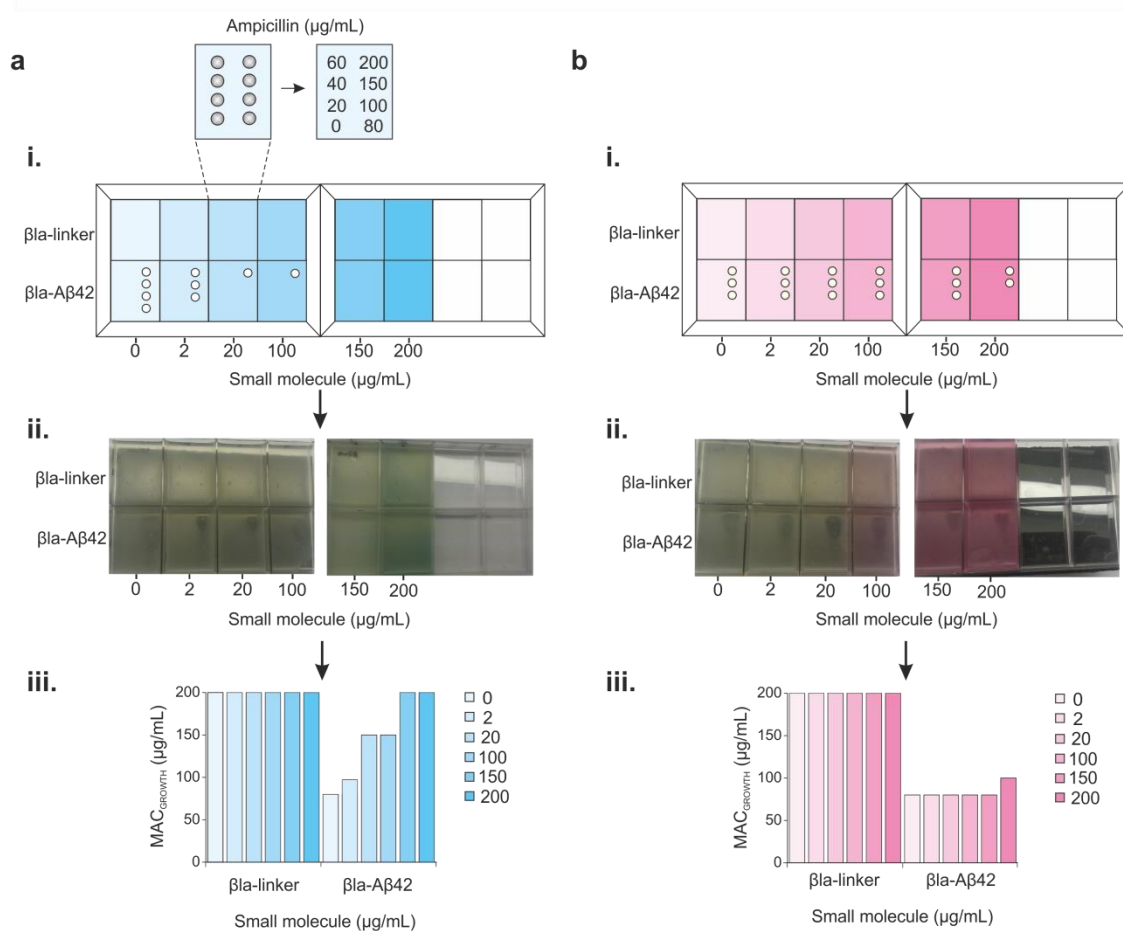
**Figure 4.3.** Relationship between protein expression and bacterial growth. (a) Hypothesised relationship between cell survival and expression of  $\beta$ -lactamase induced by different concentrations of arabinose. At lower [arabinose], the lack of  $\beta$ -lactamase expression hinders cell growth. Conversely, at higher [arabinose], the over-expression of A $\beta$ 42 leads to extensive aggregation and cell death. (b-h) MIC assays performed in the presence of increasing arabinose concentrations with  $\beta$ -lactamase- (●) and  $\beta$ -lactamase-A $\beta$ 42-expressing cells (●). This experiment was performed with undergraduate student Joseph Casson.



**Figure 4.4.** Bacterial growth in presence of curcumin with optimised conditions. (a-e) MIC assay performed in the presence of increasing concentrations of the small molecule curcumin.  $\beta\text{la-A}\beta\text{42}$  expression was induced with the optimised concentration of arabinose of 0.015 % (w/v). This experiment was performed by undergraduate student Joseph Casson.

## 4.2.2 Reducing the number of assay variables

Screening eight concentrations of a single compound, at six concentrations of ampicillin, would require 48 circular agar plates. To circumvent the large number of plates used (which would limit any development toward medium or high throughput studies), a bacterial lawn approach was investigated. Various sample plates were considered (8-, 12-, 48-, 96-well), however a sterile 8-well rectangular petri provided the greatest ease for producing bacterial lawns within the wells, and also in identification of zones of clearance in the bacterial lawns (**Figure 4.5a**).

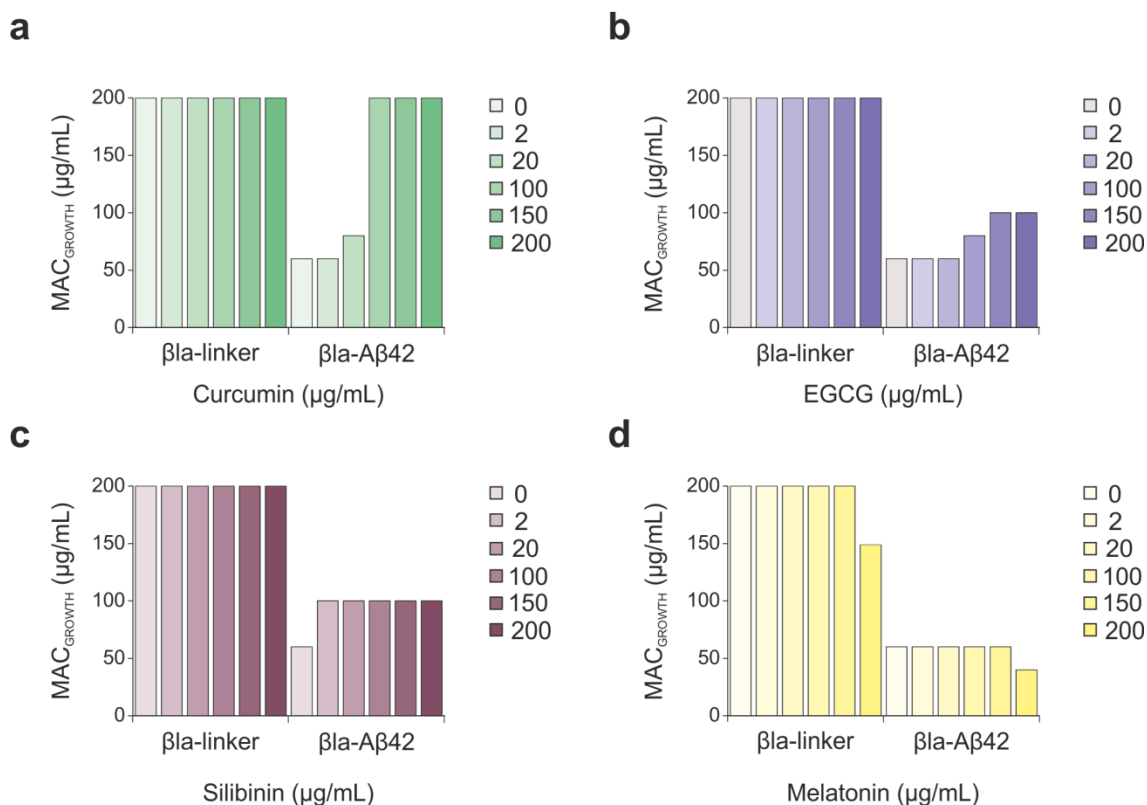


**Figure 4.5.** Schematic of bacterial lawn-based assay. **(a)** Example of positive inhibitor of A $\beta$ 42 aggregation (azure C). **(i)** Agar plates containing increasing concentrations of the small molecule of interest are created. A bacterial culture is mixed with agar and spread on top of the plate. Increasing concentrations of ampicillin (0 – 200  $\mu\text{g/mL}$ ) are spotted on top of each well of the plate. **(ii-iii)** Plates are incubated at 37  $^{\circ}\text{C}$  for 18 h and the maximum ampicillin concentration at which growth occurs ( $\text{MAC}_{\text{GROWTH}}$ ) is scored for each concentration of small molecule. **(bi-iii)** Example of a small molecule that does not inhibit A $\beta$ 42 aggregation (rhodamine B).

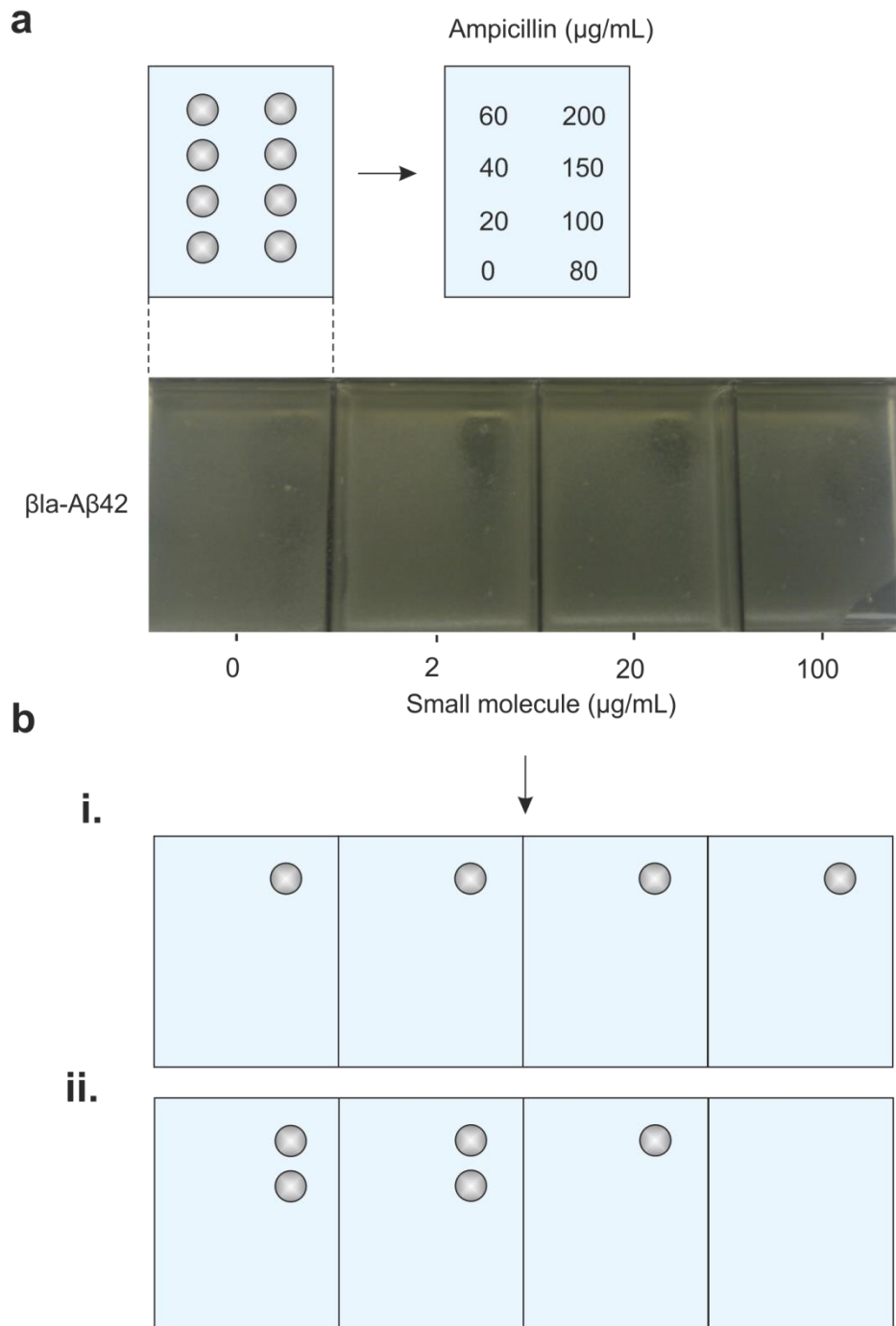
Utilisation of this format reduced the number of plates required to screen a single small molecule to just two. In the bacterial lawn methodology, a bacterial culture was mixed into a layer of soft agar (0.7 % w/v) on top of an agar plate containing increasing concentrations of the small molecule of interest (**Figure 4.5ai, bi**). Ampicillin was then spotted on top of the agar and the plates incubated at 37 °C for 18 h (Section 2.4.3). To score the plates, zones of clearance in the bacterial lawns were identified (**Figure 4.5aii, bii**). A zone of clearance means that concentration of ampicillin has prevented bacterial growth. **Figure 4.5a** gives an example of a positive inhibitor of A $\beta$ 42 aggregation (azure C), and **Figure 4.5b** gives an example of a negative inhibitor (rhodamine B).

The top agar technique was used to assay numerous known inhibitors and small molecules known not to inhibit A $\beta$ 42 aggregation (examples are given in **Figure 4.6a-d**). However, despite the initial success of the bacterial lawn assay, there were a number of disadvantages to this technique. Firstly, the methodology was difficult to carry out. The bacterial culture had to be pre-mixed into molten agar prior to pouring the plates. This required the agar to be at < 40 °C to ensure the cells did not experience heat shock<sup>447</sup>. At this temperature the agar begins to solidify, making stirring, pipetting, and production of a homogenous solution containing the small molecules difficult. These issues led to numerous failed assays or inconsistent scores (data not shown). Secondly, the amount of small molecule required for the assay is excessive, when compared to the amounts used for standard medium or high throughput assays (microgram quantities). Originally the small molecule was only mixed in to the layer of top agar containing the bacterial lawn. However, it was found that many of the compounds dissipated into the bottom agar during the 18 h incubation (as shown by diffusion of the coloured small molecules). This led to a decrease in small molecule concentration experienced by the bacteria. To counter this issue, the small molecule was also added to the layer of bottom agar, resulting in a requirement of ~ 2 mg of small molecule per assay. This quantity was deemed too high for a successful screening assay. The final, but considerable, issue was the large error observed during scoring the zones of clearance in the bacterial lawns. As the clearance occurred over a gradient, rather than complete clearance, judgement of what constituted a clearance fell to the observer (**Figure 4.7**). Obtaining consistency over different experiments, and days, was not possible. Although this issue could possibly be circumvented by automating scoring, this would be difficult to implement and validate. It was decided, therefore, that the best course of action would be to return to the original plate set-up, where colony scoring was utilised. As the objective had originally been to

reduce the number of variables, and hence the number of agar plates required, the protocol was changed such that only one concentration of small molecule would be tested in the first instance. More detailed, concentration dependent assays could then be performed if required later, once initial hits had been identified.



**Figure 4.6.** Identifying inhibitors of Aβ42 aggregation using bacterial lawn assay. Three known inhibitors of Aβ42 aggregation, curcumin (**a**), EGCG (**b**) and silibinin (**c**), were assayed using the top agar bacterial lawn technique. A small molecule known not to inhibit Aβ42 aggregation, melatonin (**d**), was also assayed as a negative control. The maximal ampicillin concentration allowing growth (MAC<sub>GROWTH</sub>) was scored for both the βla-linker and βla-Aβ42 constructs at each concentration of small molecule (0 – 200 µg/mL).



**Figure 4.7.** Example of scoring discrepancies in the bacterial lawn assay. **(a)** Eight concentrations of ampicillin are spotted onto each well of the plate containing the bacterial culture suspended in agar. After 18 h incubation, the plates are recorded for analysis. **(b)** Example of the discrepancies in scoring by two individuals within the same laboratory. Scorer one **(i)** records no difference in the amount of clearance in the bacterial lawns for each concentration of small molecule. Scorer two **(ii)** records a decrease in lawn clearance as the concentration of small molecule increases.



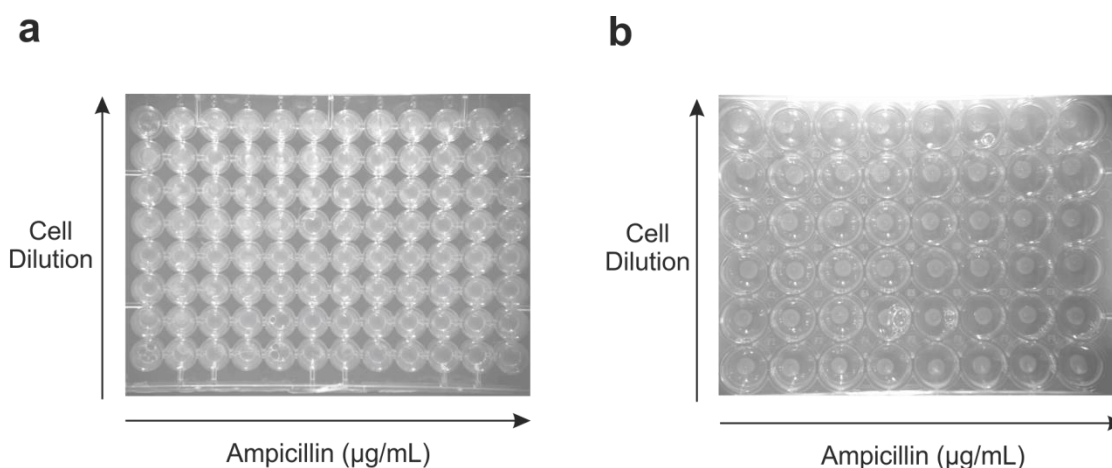
### 4.2.3 Assay optimisation

A high quality assay that required the minimal amount of time and reagents to perform (summarised in **Table 4.1**) was developed over time by optimising assay parameters. The final assay is described in Section 2.4.4 and illustrated in the schematic in **Figure 4.10**.

#### 4.2.3.1 Optimising choice of agar plate

The primary task was to find a suitable plate that required the minimal amount of agar (and therefore small molecule) to perform the assay, while still allowing colony growth to be scored. The 8-well plates utilised for the bacterial lawn assay (Section 2.4.3) required ~ 2 mg of small molecule (a 200 Da compound) per assay. Utilising a 12-well plate reduced the requirement of small molecule to ~ 1 mg per assay, however this was still deemed too high for a successful assay. Investigations with a 96 well plate (one well per cell dilution of culture) successfully reduced the necessity for large quantities of small molecule (only 144 µg/assay) by providing the absolute minimal volume required for agar. However, the small well size made scoring the presence of colonies near-impossible (**Figure 4.8a**). By increasing the well size a small fraction to 48-well agar plates, colonies could be visualised easily and recorded (**Figure 4.8b**). Investigations into the volume of agar to use in each well revealed that 300 µL was optimal. This was the lowest volume of agar that could be used that gave full well-bottom coverage and did not dry out during the 18 h incubation at 37 °C (by contrast with 200 µL agar per well). This plate set-up required only 288 µg/assay for a 200 Da small molecule.

Screening a small library of known inhibitors using the optimised 48-well plate methodology at a single concentration of ampicillin would allow the assay to be miniaturised further, thereby promoting further reduction in the quantity of small molecule required (to 48 µg of a 200 Da compound) (see Section 4.5).



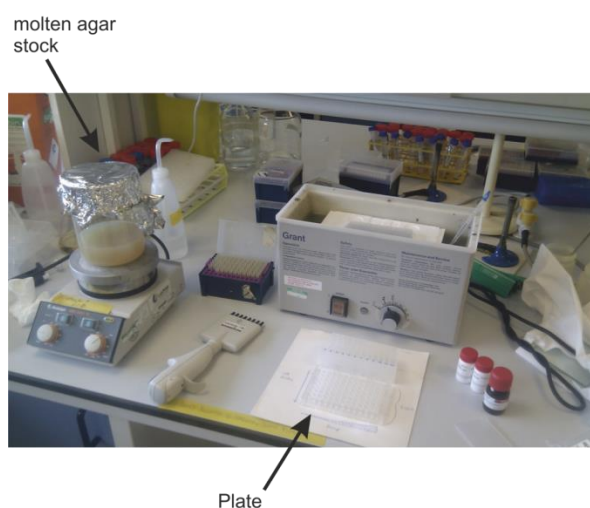
**Figure 4.8.** Comparison of scoring growth within a 96-well and a 48-well plate. Photographs illustrating the difficulties imaging colonies within the wells of a 96-well plate (**a**) compared to the larger wells of a 48-well plate (**b**).

#### 4.2.3.2 Sample incubation with small molecule

Initial experiments (Section 4.2.1) showed that the best results were obtained when the bacterial cultures were pre-incubated with the small molecule prior to exposure to antibiotic. If pre-incubation did not occur, and the bacteria were exposed to the small molecule and ampicillin in the agar plate simultaneously, no bacteria grew at ampicillin concentrations higher than 40 µg/mL (data not shown). For rapidly aggregating proteins such as hIAPP and Aβ42, these observations suggest that the small molecule must be present prior to the initial onset of aggregation for them to exert their most potent anti-aggregation effects. This is supported by literature on *in vitro* aggregation in the presence of small molecules; only a small group of compounds have been demonstrated to disaggregate pre-formed fibrils or amorphous aggregates<sup>235, 448, 449</sup>. Extending the pre-incubation time in the presence of small molecules from 1 h to 3 h had no significant effect on the ability of bacteria to grow in the presence of antibiotics. It may be that even longer incubation times would have had beneficial effects on bacterial growth, however it was decided to limit pre-incubation to 1 h to minimise the time taken to perform the assay.

### 4.2.3.3 Point of small molecule addition

The initial assay optimisation experiments were performed using a large beaker of 100 mL molten 'stock' agar (**Figure 4.9**). The large volume was initially required to maintain the agar in a molten state and enable full submersion of the multichannel pipette tips into the liquid. The small molecule was mixed into this prior to pouring the plates, and although only a small amount of compound was required in each well, a large quantity of compound was needed to produce 100 mL of agar containing 100  $\mu\text{M}$  small molecule. Two techniques were investigated to circumvent this problem: addition of small molecule to the agar plate after the agar had set, and addition of small molecule to the plate prior to agar pouring. The first technique involved spotting the small molecule on top of each agar well and leaving it to dissipate into the agar. This technique worked well for small molecules dissolved in water, however molecules in DMSO or ethanol did not fully dissipate throughout the agar (as shown by dissipation of coloured compounds). Furthermore, the successful dissipation could only be confirmed visually for coloured small molecules. The second technique attempted was the addition of the small molecule to the well prior to agar addition. The plate was then shaken to ensure homogenous dispersion of the small molecule throughout the agar. This method was much more successful, resulting in complete mixing of molecules dissolved in water, ethanol and DMSO in the agar. It also reduced the amount of small molecule required to 576  $\mu\text{g}/\text{assay}$  (for a 200 Da compound). This quantity is further reduced during miniaturisation (see Section 4.5) to 72  $\mu\text{g}/\text{assay}$ .



**Figure 4.9.** Assay setup for pre-mixing of compound into the agar. A large beaker of molten agar (containing 100  $\mu\text{M}$  of small molecule) acts as the stock while preparing the agar plates for the assay.

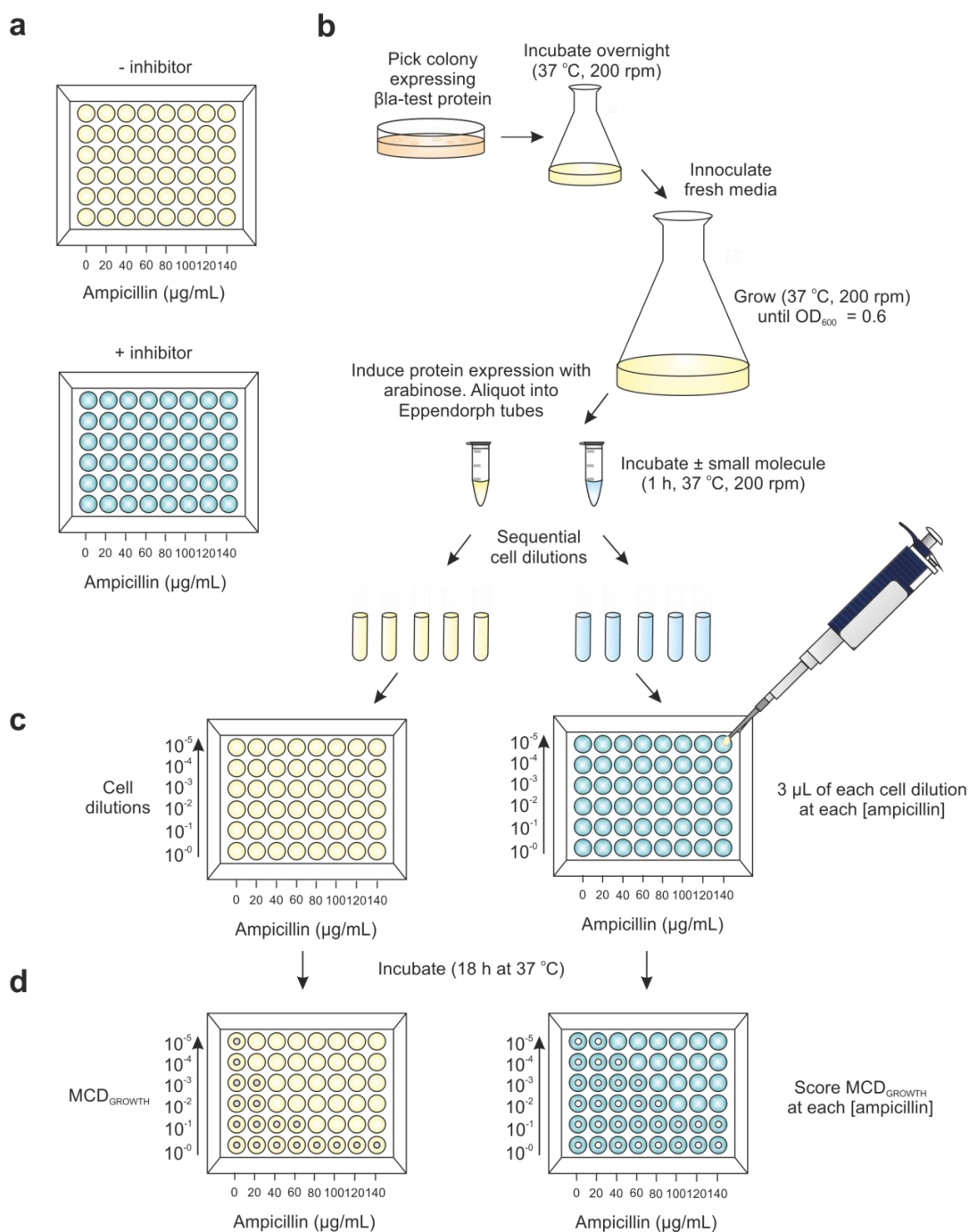
#### 4.2.3.4 Concentration of small molecule

Numerous concentrations of small molecule were investigated with the aim of reducing the cytotoxicity of the small molecules to the cells, without forfeiting their anti-aggregation properties. It was found for a number of small molecules, that a concentration of 500  $\mu\text{M}$  small molecule significantly rescued cellular growth, however a considerable number of small molecules were either partially or completely toxic at this concentration. Reducing the small molecule concentration to  $< 100 \mu\text{M}$  prevented any growth rescue, suggesting that this concentration of small molecule is below the concentration required to inhibit protein aggregation. For further screening of small molecules, therefore, a concentration of 100  $\mu\text{M}$  small molecule was chosen to warrant minimal cell toxicity, while maintaining the chance of observing any anti-aggregation activity of the compound. Furthermore, it was believed this concentration would guarantee the selection of non-colloidal inhibitors, often identified during other screening techniques using high concentrations of small molecule<sup>450, 451</sup>.

4.2.3.5 Summary of optimised *in vivo* assay

Parameter	Condition	Score	Conclusion
Agar plate to be used to minimise volume of agar/small molecule	8-well	*	8- and 12- well plates required too much agar/small molecule. 96-well plates used a suitable amount of agar however the colonies were too difficult to visualise. 48-well plates circumvented both these issues successfully.
	12-well	**	
	48-well	****	
	96-well	**	
Volume of agar/well	500 $\mu$ L	**	300 $\mu$ L agar/well ensured the minimal amount of small molecule would be used. 200 $\mu$ L was not enough to cover the base of the wells in a 48-well plate.
	400 $\mu$ L	***	
	300 $\mu$ L	****	
	200 $\mu$ L	*	
Sample incubation with small molecule prior to addition to agar plate	0 h	*	Pre-incubation in the presence of small molecule inhibitors was required for colonies to grow under selected growth conditions. Optimal time was 1 h.
	1 h	****	
	2 h	****	
	3 h	****	
Point of small molecule addition	Spotted on top of agar	*	Small molecules could not dissipate fully into set agar. Pre-mixing into agar was easier to perform, however it required 2 mg of small molecule (a 200 Da compound). Addition of small molecule directly to well prior to addition of agar only required 288 $\mu$ g of compound.
	In well prior to agar addition	****	
	Pre-mixed into stock agar	*	
Concentration of small molecule	50 $\mu$ M	**	A small molecule concentration of 100 $\mu$ M was most optimal, combining the highest hit rates with the lowest cytotoxicity.
	100 $\mu$ M	****	
	250 $\mu$ M	**	
	500 $\mu$ M	*	

**Table 4.1.** Parameters optimised for small molecule screen. Conditions are scored from the most (four stars) to the least (one star) favourable.

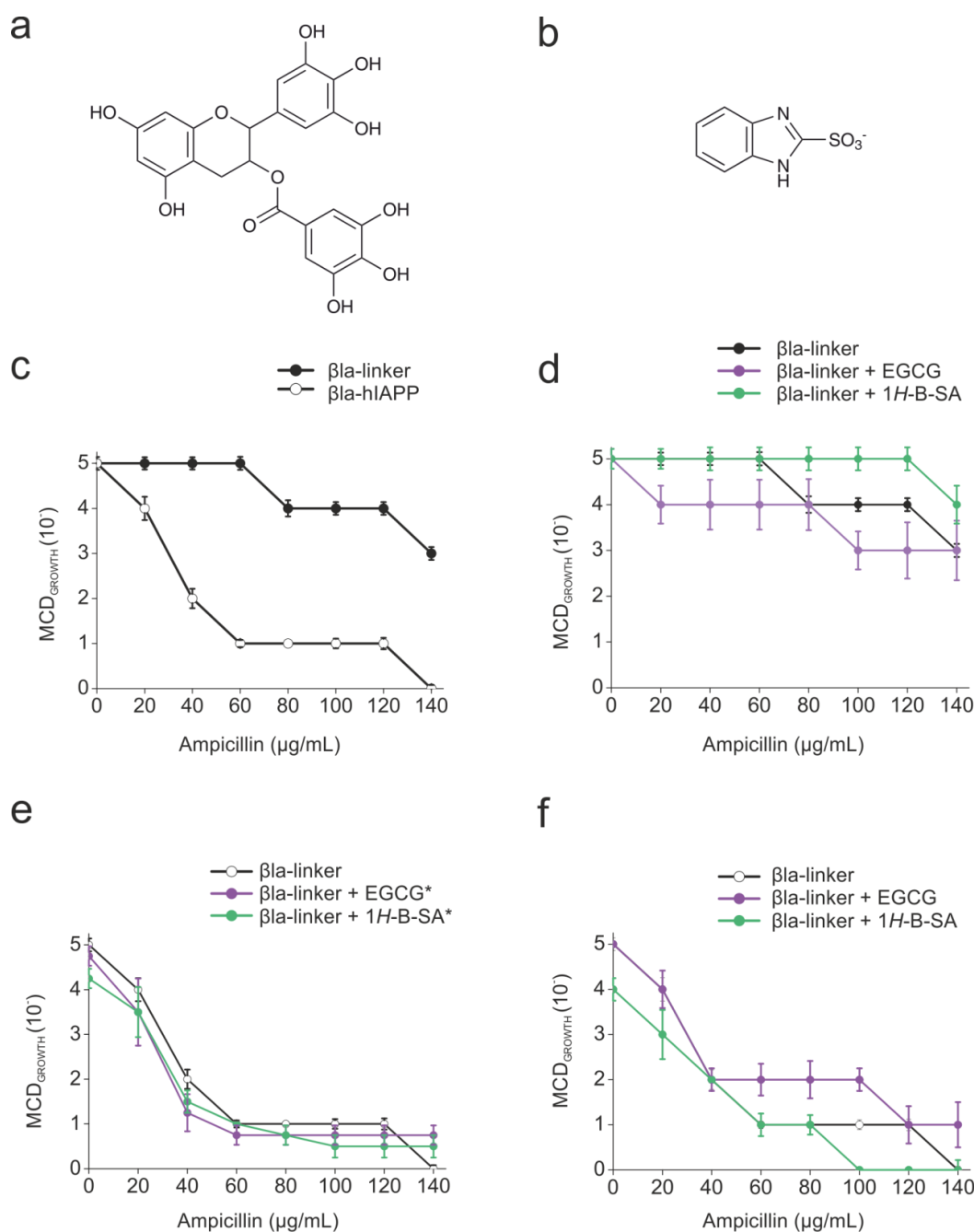


**Figure 4.10.** Schematic of optimised 48-well plate assay. **(a)** 48-well agar plates, containing the small molecule of interest or the relevant solvent, are prepared prior to performing the assay. **(b)** Colonies transformed with the required plasmid are selected and grown until an  $\text{OD}_{600}$  of 0.6 is reached.  $\beta$ -lactamase-fusion protein expression is induced with arabinose and cultures are pre-incubated in the presence or absence of small molecule for 1 h. **(c)** Cultures are serially diluted and 3  $\mu\text{L}$  pipetted into each well of the prepared agar plates. Plates are incubated for 18 h at 37 °C. **(d)** The maximal cell dilution at which growth occurs ( $\text{MCD}_{\text{GROWTH}}$ ) in the presence and absence of small molecule is scored for each concentration of ampicillin. Control plates of bacteria expressing  $\beta$ la-linker are also examined for intrinsic effects of small molecule on growth.

#### 4.2.4 Accounting for off-target effects

The addition of small molecules to an *in vivo* system risks unwanted off-target effects. These could range from non-specific cytotoxicity to cell cycle arrest<sup>452</sup>. To account for these effects, a control plate was created in each experiment using bacteria expressing the  $\beta$ la-linker construct. Any aberrant effect a small molecule has on cellular growth can therefore be identified in these bacteria, and accounted for during the assay. As an example, the effects of the small molecules EGCG (**Figure 4.11a**) and 1*H*-benzimidazole-2-sulfonic acid (1*H*-B-SA) (**Figure 4.11b**) on bacterial growth are calculated as described below.

Any effect of the small molecule on bacterial growth attributed to the presence of the small molecule alone is first identified by comparing the growth of the  $\beta$ la-linker expressing bacteria in the presence and absence of the small molecule of interest (**Figure 4.11c, d**). For example, EGCG has a slightly adverse effect on bacterial growth, whereas 1*H*-B-SA promotes cell growth (**Figure 4.11d**). These effects are deducted from the  $\beta$ la-hIAPP assay (**Figure 4.11e**) to give a true representation of the small molecule's ability to prevent  $\beta$ la-hIAPP aggregation (**Figure 4.11f**). 100  $\mu$ M EGCG enabled greater cell dilutions of bacteria to grow at higher concentrations of ampicillin; at most concentrations of ampicillin a 10-fold improvement in growth is observed (**Figure 4.11f**). The presence of 100  $\mu$ M 1*H*-B-SA did not aid bacterial growth in the presence of ampicillin (**Figure 4.11f**), in agreement with *in vitro* analysis (see Section 4.3.3). The slight toxicity observed for the negative control 1*H*-B-SA at the highest concentrations of ampicillin could result from the synergistic toxic effects of protein aggregation and the presence of small molecule, although this remains speculation at this point in time.



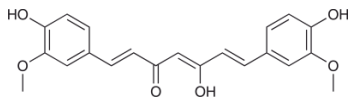
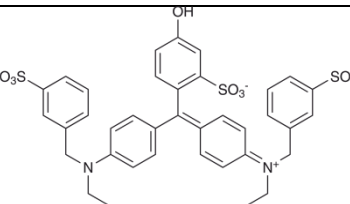
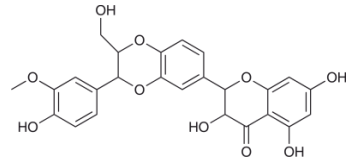
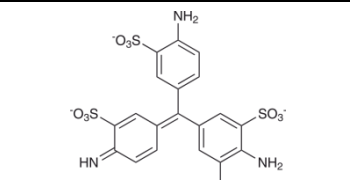
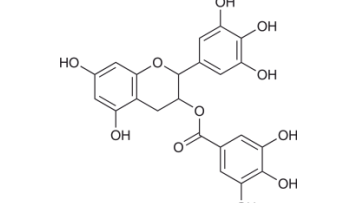
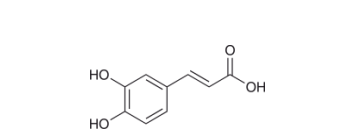
**Figure 4.11.** Correcting for intrinsic effects of small molecules on bacterial growth. **(a)** Structure of (-)-epigallocatechin gallate (EGCG). **(b)** Structure of 1H-benzimidazole-2-sulfonic acid (1H-B-SA) **(c)** Assay in absence of small molecules. Maximal cell dilution allowing growth ( $MCD_{GROWTH}$ ) was assessed at each ampicillin concentration. Bacteria expressing the non-aggregating  $\beta$ la-linker (●) have a higher minimal inhibitory concentration (MIC) of antibiotic than bacteria expressing  $\beta$ la-hIAPP (○). **(d)** The small molecule effect on cell growth (independent of aggregation) is assessed. 100  $\mu$ M of EGCG (●) is detrimental to growth, whereas 100  $\mu$ M of 1H-B-SA (●) promotes cell growth. **(e)** The ability of EGCG and 1H-B-SA to prevent  $\beta$ la-hIAPP aggregation is assessed, uncorrected for intrinsic effects of the small molecule on growth (\*) (Section 4.2.4). **(f)** Assay corrected for intrinsic effects of the small molecules on bacterial growth. Error bars represent the standard error from a minimum of 4 replicate experiments.



## 4.3 Assay validation

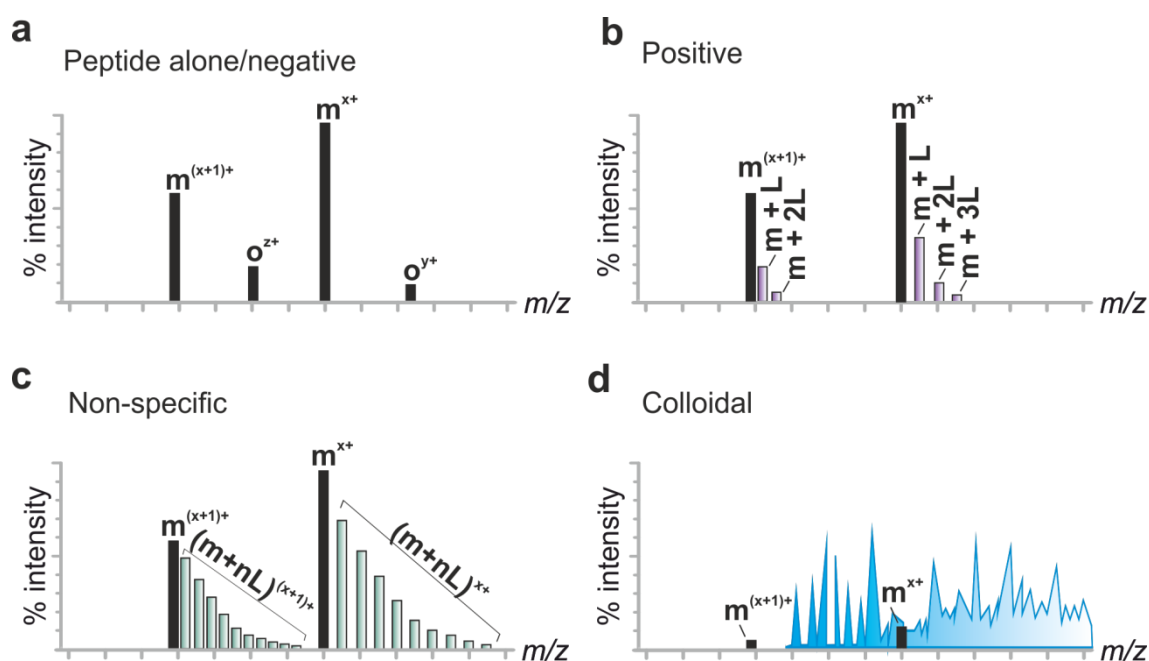
### 4.3.1 Selection of known small molecule inhibitors

After assay optimisation, the next stage was to validate the strategy by screening a small library of small molecules whose activity against protein aggregation is known. Six well-characterised inhibitors of hIAPP aggregation were selected from the literature for screening (**Table 4.2**)<sup>234-237, 245, 284-287, 453</sup>.

Small Molecule	Mode of Inhibition	Molecular Weight (Da)	Structure
Curcumin	Significantly reduces hIAPP aggregation <i>in vitro</i> and alleviates some toxicity of pancreatic $\beta$ -cells <i>in vivo</i> <sup>245</sup>	368.4	
Fast Green FCF	10:1 molar ratio of Fast Green FCF:hIAPP inhibits all aggregation <sup>284</sup>	765.9	
Silibinin	Results in amorphous aggregates at 5:1 molar ratio of silibinin:hIAPP <sup>287</sup> , and complete inhibition of aggregation at 10:1 molar ratio <sup>237</sup>	482.4	
Acid fuchsin	Inhibits all amyloid formation at 10:1 molar ratio of acid fuchsin:hIAPP <sup>284</sup> . Arrests amyloid formation by trapping intermediate species <sup>285</sup>	585.3	
EGCG	Potent inhibitor of hIAPP aggregation <sup>234, 236, 237, 453</sup> ; can disaggregate amyloid fibrils <sup>235</sup>	458.4	
Caffeic acid	5:1 molar ratio of caffeic acid:hIAPP inhibits all aggregation <sup>286</sup>	180.2	

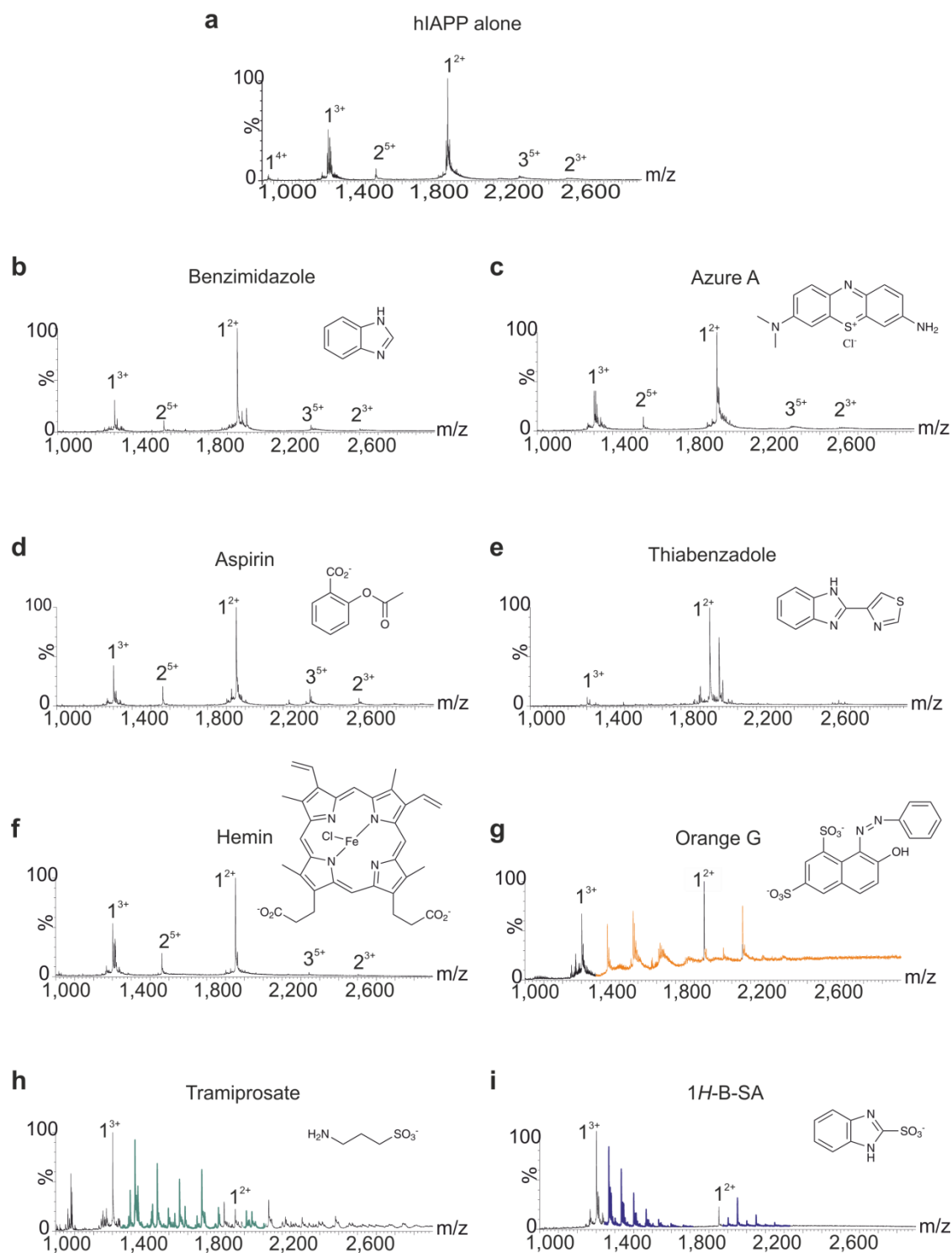
**Table 4.2.** Known inhibitors of hIAPP aggregation.

As only one small molecule with no effect on hIAPP aggregation had been published to date (tramiprosate<sup>285</sup>), a selection of small molecules were assessed by a technique under development in the laboratory using electrospray ionisation-ion mobility spectrometry-mass spectrometry (ESI-IMS-MS) to characterise the mode of protein aggregation inhibition by small molecules. ESI-MS spectra can be used to classify compounds that do not bind, and those that bind specifically, non-specifically or colloiddally to the protein of interest (**Figure 4.12**). The mass spectrometry experiments were all performed by Lydia M. Young (University of Leeds) and have been submitted for publication<sup>379</sup>.



**Figure 4.12.** Schematic of ESI-MS spectra for the various small molecule binding modes. Peptide monomer peaks are denoted ‘m’, oligomer peaks ‘o’ and bound ligand ‘ $m+xL$ ’. Charge states are in superscript. (a) A non-interacting small molecule will produce a spectrum the same as the peptide alone. (b) A small molecule that specifically interacts with the peptide will produce a binomial distribution of bound peaks (purple)<sup>454</sup>. (c) A non-specific ligand will bind but result in a Poisson distribution of bound peaks (green)<sup>454</sup>. (d) A colloidal inhibitor will produce overlapping peaks due to self-association of the small molecule (blue). Figure adapted from Saunders & Young, 2014<sup>379</sup>.

Eight small molecules were identified from their ESI-MS spectra (**Figure 4.13**) and selected as negative reference controls for the *in vivo* screen validation. Five of the small molecules (benzimidazole, azure A, aspirin, thiabendazole and hemin) did not bind hIAPP (**Figure 4.13b-f**), one (Orange G) was colloidal (**Figure 4.13g**), and two (tramiprosate and 1H-B-SA) bound non-specifically to hIAPP (**Figure 4.13h** and **Figure 4.13i**).



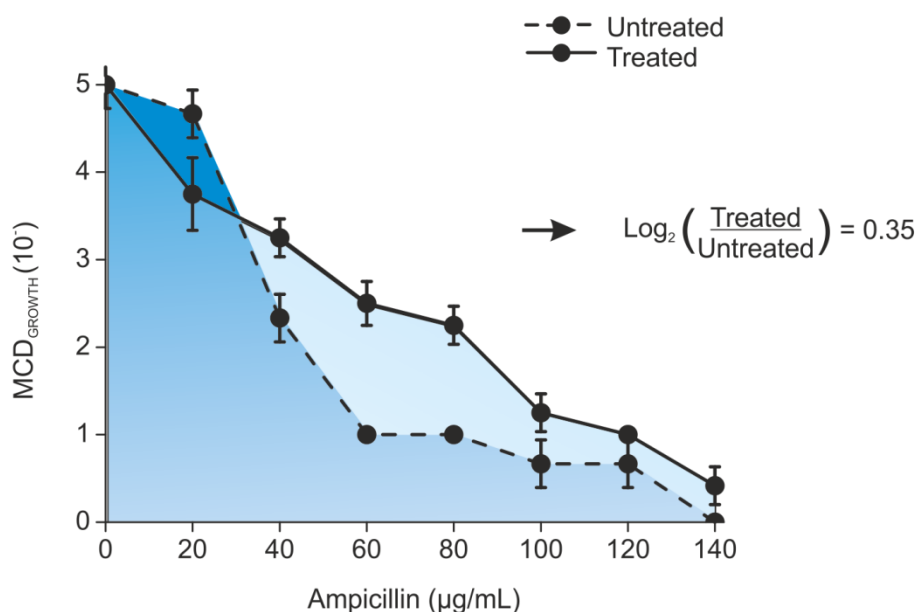
**Figure 4.13.** ESI mass spectra of small molecule/hiAPP interaction. ESI-MS mass spectra of 32  $\mu$ M hiAPP alone (**a**), or in the presence of 320  $\mu$ M (**b**) benzimidazole, (**c**) azure A, (**d**) aspirin, (**e**) thiabendazole, (**f**) hemin, (**g**) Orange G, (**h**) tramiprosate or (**i**) 1*H*-B-SA. Numbers above peaks denote oligomer order, with the positive charge state of ions in superscript. Peaks attributed to colloidal Orange G molecules are denoted in orange (**g**). Peaks attributed to non-specific binding of tramiprosate (**h**) and 1*H*-B-SA (**i**) are denoted in green and blue respectively. Experiments were performed in collaboration with Lydia M. Young (University of Leeds).

### 4.3.2 *In vivo* screening of known inhibitors and non-inhibitors

The maximal cell dilution allowing growth at the various ampicillin concentrations was assessed for each of the 14 selected small molecules (in all cases, any intrinsic effect on cell growth by the small molecule alone was corrected for as described in Section 4.2.4). In order to obtain a single value from each MIC assay, illustrative of the small molecule's effect on bacterial growth, the area under the MIC assay curves was calculated as a sum of the areas of 7 trapezia (**Equation 4.1**), where  $A_{curve}$  is the total area under the curve, and  $x_i$  and  $y_i$  are the  $x$ -axis and  $y$ -axis values at each concentration of ampicillin (see **Figure 4.14**).

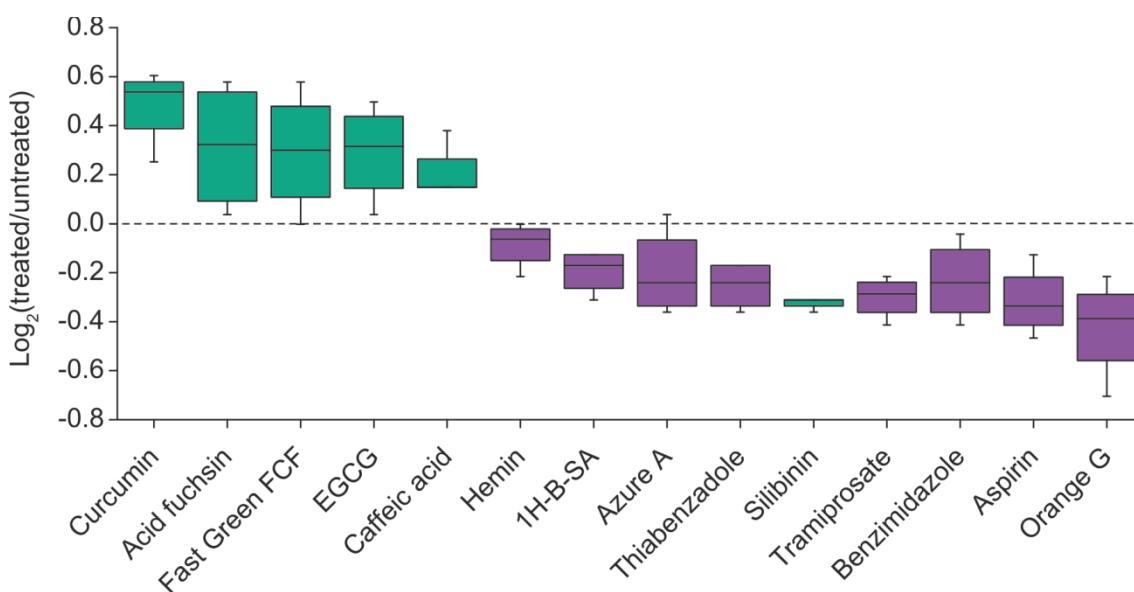
$$A_{curve} = \sum_{i=1}^7 \frac{y_i + y_{i+1}}{2} \times (x_{i+1} - x_i)$$

**Equation 4.1**



**Figure 4.14.** Growth quantification of bacteria in the presence or absence of a small molecule. The area under the curve of the treated and untreated cells is calculated using Equation 4.1.

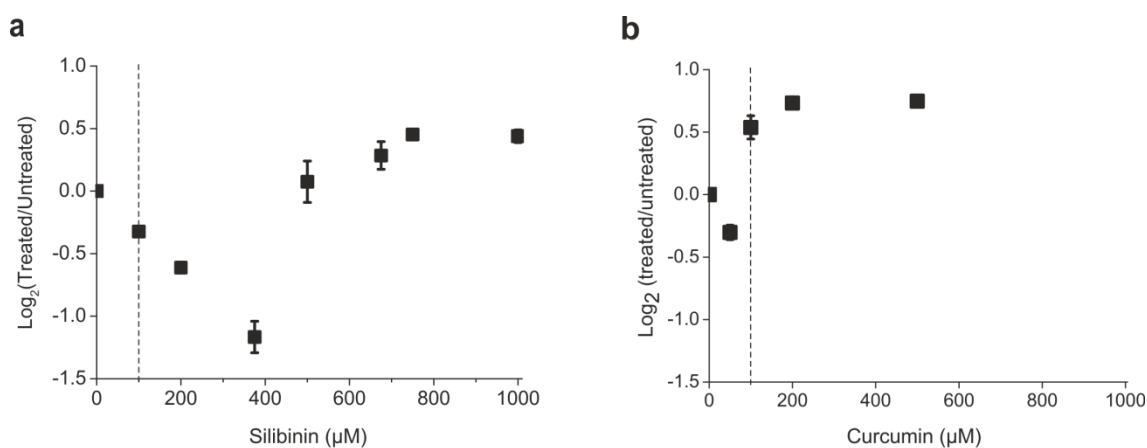
The area under the curve of cells treated with small molecule was compared with that of an untreated control (the absence of small molecule) (example in **Figure 4.14**). This was repeated for each small molecule on four different days. The fraction was plotted on a logarithm to base 2 graph, producing a graph that is symmetrical about zero (**Figure 4.15**). Any small molecule that has a positive effect on cell growth therefore produces a score greater than zero and is classified as a 'hit'.



**Figure 4.15.** *In vivo* classification of small molecule inhibitors of  $\beta$ Ia-hIAPP aggregation. Box plot (n=4) showing the effect on growth of each small molecule. Centre line = median; box limits indicate the 25<sup>th</sup> and 75<sup>th</sup> percentiles, with whiskers extending to  $\pm 1.5$  times the interquartile range. Dashed line through graph indicates separation of small molecules that have an effect ( $> 0$ ), and those that do not ( $\leq 0$ ). Compounds are coloured as known inhibitors (●) or as molecules known to have no effect (●) on hIAPP aggregation *in vitro*.

The *in vivo* classification of small molecules as 'hits' or as no positive effect on cell growth showed a striking correlation with the published data (**Table 4.2**) and the mass spectrometry analysis (**Figure 4.13**). Importantly, all the molecules that were shown by mass spectrometry to not bind, or to bind non-specifically or colloiddally to hIAPP (**Figure 4.13**), were classified as negatives in the *in vivo* screen (**Figure 4.15**). Furthermore, five out of six known inhibitors of hIAPP aggregation consistently aided

bacterial growth in the presence of antibiotics, resulting in their classification as 'hits' in the *in vivo* screen. The only inconsistency between the published data and the *in vivo* screen was for the small molecule silibinin (**Figure 4.15**). To investigate this anomaly, a dose-response experiment was performed which showed that when the concentration of silibinin was increased to  $> 500 \mu\text{M}$  (as compared to the  $100 \mu\text{M}$  used previously), a positive score in the *in vivo* assay was achieved (**Figure 4.16a**). This is in contrast to the strongest inhibitor identified in the assay, curcumin, which induces a significant improvement in bacterial growth at  $100 \mu\text{M}$  (**Figure 4.16b**). By performing the assay at  $100 \mu\text{M}$  small molecule, only strong, specific and non-colloidal inhibitors will be identified. The toxicity observed at  $50 \mu\text{M}$  curcumin and  $50\text{-}375 \mu\text{M}$  silibinin is most likely due to synergistic toxic effects of protein aggregation and the presence of the small molecule as, at this concentration,  $\beta\text{Ia-hIAPP}$  aggregation is not being prevented.



**Figure 4.16.** *In vivo* dose-response relationship between bacterial growth and small molecule concentration. Effect on bacterial growth of increasing concentrations of (a) silibinin and (b) curcumin. Error bars represent the standard error from a minimum of 4 replicates, dashed line indicates small molecule concentration utilised in standard MIC assay ( $100 \mu\text{M}$ ).

### 4.3.3 Validating classification of hIAPP aggregation inhibitors

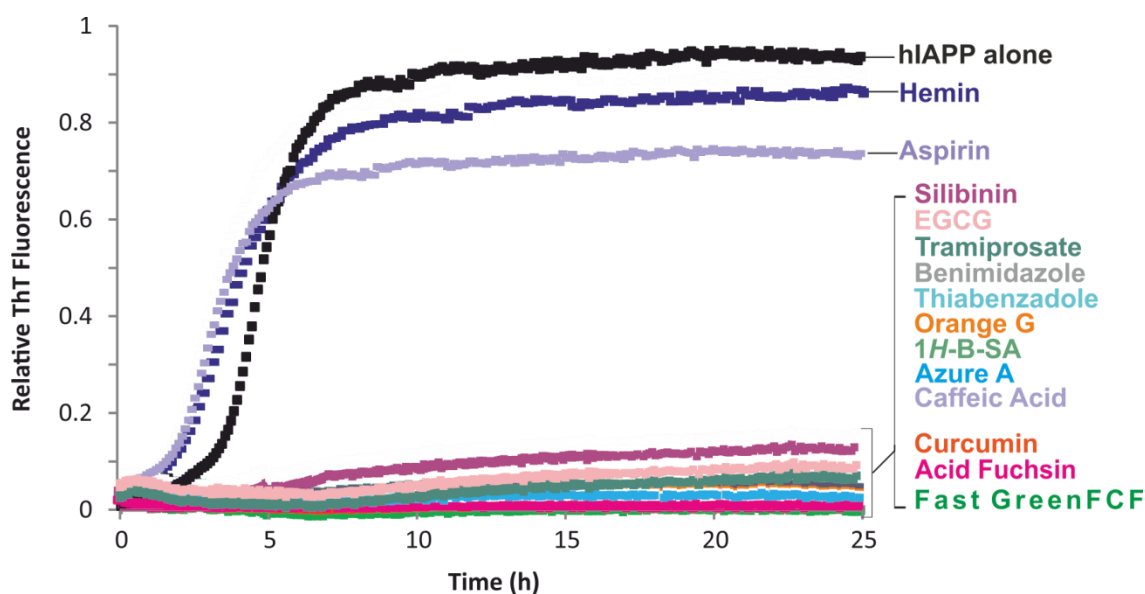
*In vitro* analysis of the aggregation of hIAPP peptide in the presence and absence of the small molecule inhibitors was performed to validate the *in vivo* classification of 'hits'. All data are summarised in **Tables 4.3-4.7**.

#### 4.3.3.1 Transmission electron microscopy

Transmission electron microscopy (TEM) was carried out to analyse the nature of aggregates formed by hIAPP in the presence of the 14 small molecules screened. A 1:10 molar ratio of peptide:small molecule was incubated for five days before TEM images were captured (Section 2.5.6.2). The five small molecules classified as 'hits' in the *in vivo* screen (curcumin, acid fuchsin, Fast green FCF, EGCG and caffeic acid) also prevented hIAPP amyloid formation *in vitro*, resulting either in very small oligomers or amorphous aggregates (**Table 4.3, Table 4.4**). Furthermore, eight of the nine small molecules that failed to prevent  $\beta$ la-hIAPP aggregation *in vivo* (at a concentration of 100  $\mu$ M), also failed to prevent amyloid formation by hIAPP *in vitro*, resulting in amyloid fibrils or large amorphous aggregates (**Table 4.5, Table 4.6, Table 4.7**). The single false-negative in the *in vivo* screen, silibinin, resulted in the formation of small amorphous aggregates of hIAPP *in vitro* (**Table 4.6**).

### 4.3.3.2 Thioflavin T fluorescence

The increase in extrinsic fluorescence of the benzothiazole dye thioflavin T (ThT), upon binding to  $\beta$ -sheet-rich structures, is commonly used to quantify the extent of fibrillation *in vitro*<sup>455</sup>. However, as discussed in detail in Section 1.6.1, as many potential small molecule inhibitors are intrinsically fluorescent and share the same binding site on the fibril as ThT<sup>277-279</sup>, there is a significant risk in using this technique as the sole method to assess fibril formation inhibition. To illustrate this issue further, the ThT fluorescence of hIAPP samples incubated in the presence and absence of each of the 14 small molecules was monitored over time (**Figure 4.17**). Using ThT fluorescence data alone, 12 molecules would be classified as 'hits' (preventing fibril formation) and only two small molecules (hemin and aspirin) would be classified as compounds that do not inhibit hIAPP fibril formation (**Figure 4.17**). However, when the ThT data is compared to TEM data (**Tables 4.3-4.7**), only six small molecules actually prevent hIAPP fibril formation *in vitro*, confirming that six small molecules create false-positive results in the ThT fluorescence assay (tramiprosate, benzimidazole, thiabendazole, Orange G, 1H-B-SA and azure A) (**Figure 4.17**).



**Figure 4.17.** ThT fluorescence intensity of hIAPP in presence or absence of small molecules. hIAPP alone (32  $\mu$ M peptide, 200 mM ammonium acetate buffer, pH 6.8, 25  $^{\circ}$ C, quiescent) or in the presence of each of the 14 small molecules screened (1:10 molar ratio of hIAPP:small molecule) was incubated at 25  $^{\circ}$ C for 25 h.



### 4.3.3.3 Mass spectrometry

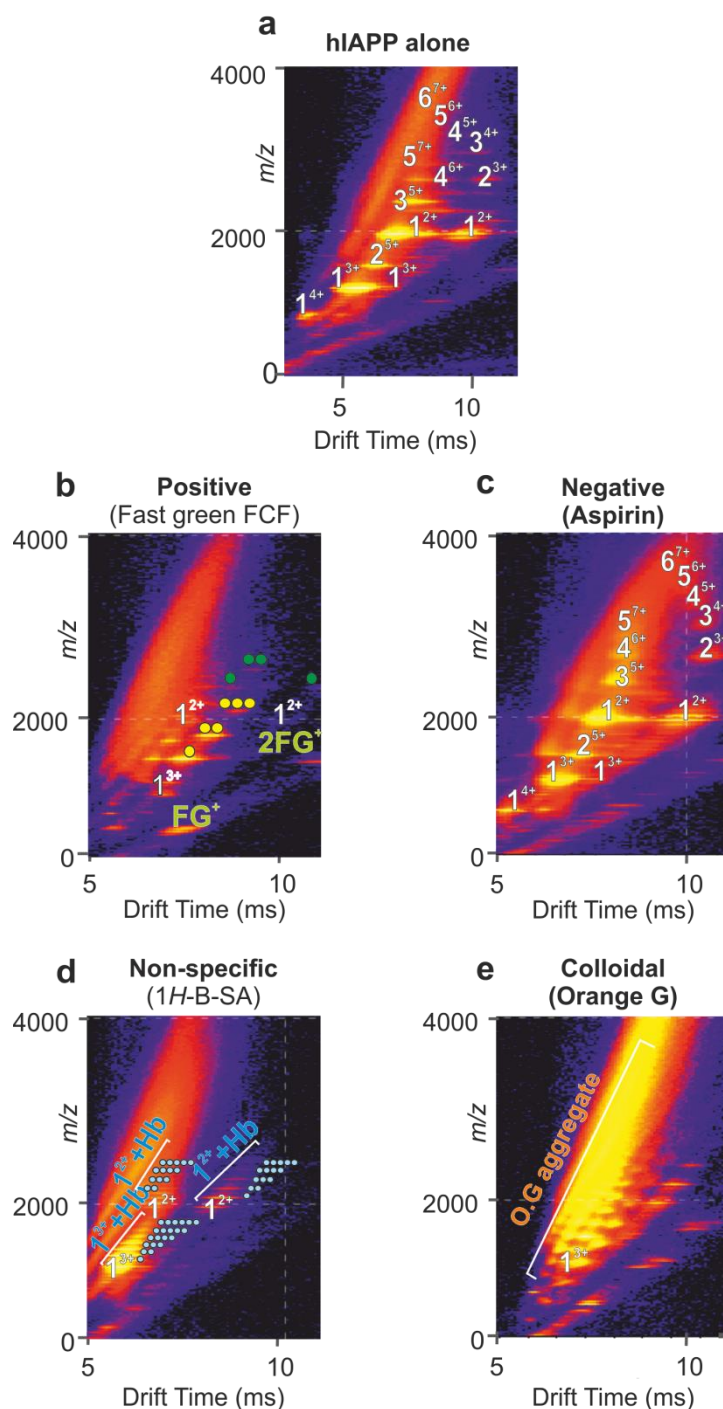
For a more quantitative analysis of hIAPP aggregation in the presence of small molecules, electrospray ionisation-ion mobility spectrometry-mass spectrometry (ESI-IMS-MS) was utilised (performed by Lydia M. Young). As described previously (Section 4.3.1; **Figure 4.13**), ESI-MS spectra can be used to classify compounds that do not bind, and those that bind specifically, non-specifically or colloiddally<sup>379</sup>. Further information can be obtained from ESI-IMS-MS Driftscope plots to determine the oligomeric distribution within the sample<sup>237, 379, 456</sup>. Example Driftscope plots of specific (positive), non-specific, negative and colloidal binding are shown in **Figure 4.18**.

In the absence of small molecule, hIAPP forms oligomers up to, and including, hexamers within two minutes of dilution into buffer (**Figure 4.18a, Table 4.3**). When the 'hits' from the *in vivo* screen were added at a 10-fold molar excess over hIAPP (32  $\mu$ M hIAPP, 320  $\mu$ M small molecule), the mass spectra revealed the *in vivo* 'hits' were, overall, specific binders to hIAPP *in vitro*, with trimers being the maximum size oligomers observed at a 1:10 molar excess of hIAPP:small molecule (**Table 4.3, Table 4.4**). Caffeic acid, previously published as a hIAPP aggregation inhibitor<sup>286</sup>, was the only 'hit' from the *in vivo* screen not identified to bind to hIAPP by ESI-MS, with oligomers up to pentamers being observed. Analysis of the TEM data (**Tables 4.3-4.7**) revealed the formation of small amorphous aggregates, and short fibrillar material. The previously published report on caffeic acid also only found small amorphous aggregates, formed at a 1:5 molar ratio, with no observable fibrils<sup>286</sup>. As the compound's effect on the hIAPP aggregation pathway is not due to non-specific or colloidal inhibition (**Table 4.4**), the data suggest that caffeic acid must interact with higher ordered oligomeric species and not monomeric hIAPP. The positive result in the *in vivo* assay suggests that the oligomers formed must be of small enough size to allow sufficient  $\beta$ la-hIAPP activity. This illustrates, therefore, that not all aggregates of  $\beta$ la-hIAPP are sufficient to reduce antibiotic resistance.

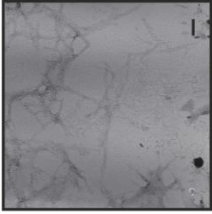
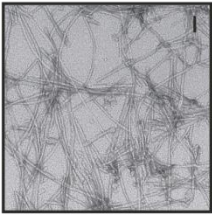
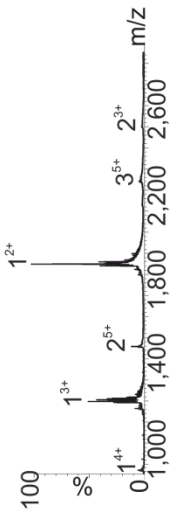
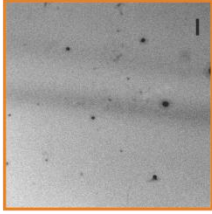
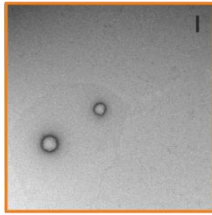
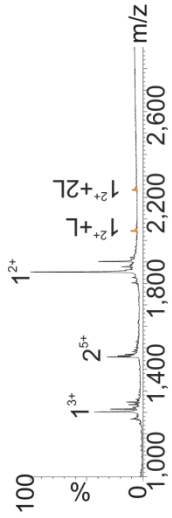
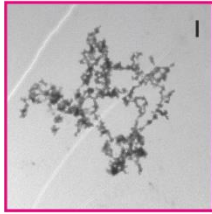
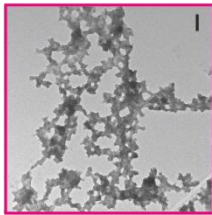
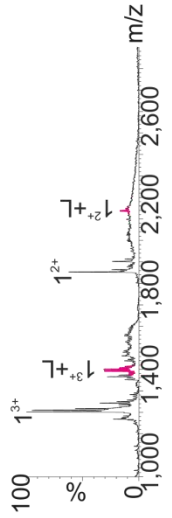
The small molecules classified as negatives in the *in vivo* screen were either non-binding, non-specific binding, or colloidal inhibitors as observed by mass spectrometry. Hemin, azure A, thiabendazole, benzimidazole and aspirin did not bind to hIAPP, resulting in the formation of tetrameric to hexameric oligomers (**Table 4.5, Table 4.6 and Table 4.7**). Conversely, tramiprosate and 1*H*-B-SA both bound to hIAPP, but in a non-specific manner. Although the hIAPP oligomers were reduced to monomer and tetramer, in the presence of tramiprosate and 1*H*-B-SA respectively, an interaction of this type is largely based on charge and is not overly sensitive to structure<sup>457</sup>. Non-specific interactions are overcome

during the aggregation pathway, leading to a final formation of fibrillar species (**Table 4.5** and **Table 4.6**). The negative score in the *in vivo* assay demonstrates that initial non-specific interactions with the aggregating protein are not enough to enable the bacteria to grow in the presence of antibiotics.

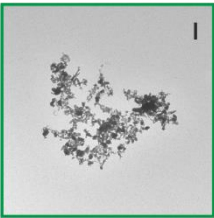
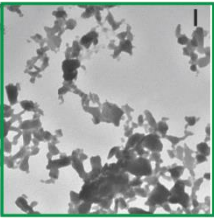
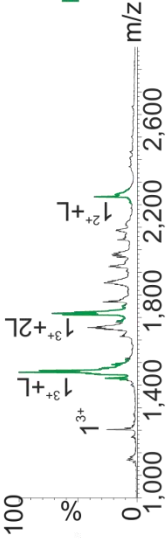
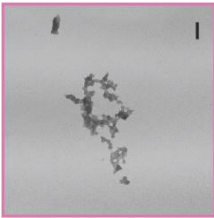
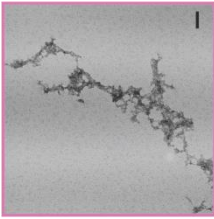
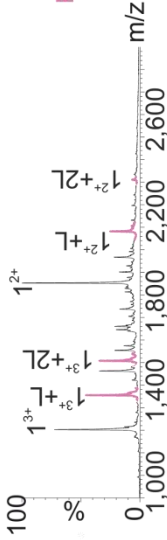
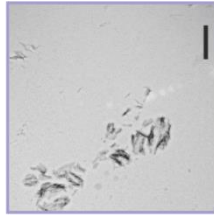

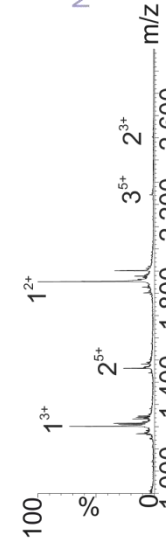
ESI-IMS-MS analysis show that the single false negative observed in the *in vivo* screen, silibinin, binds specifically to hIAPP, with monomer being the only species observed (**Table 4.6**) at a 1:10 molar ratio of hIAPP:silibinin. Previous reports on the effects of a 1:1 molar ratio of hIAPP:silibinin show a reduction in fibril formation (TEM and ThT data), albeit with no significant change on the lag time<sup>237 287</sup>. A 1:5<sup>287</sup> or 1:10<sup>237</sup> molar ratio of hIAPP:silibinin resulted in a dramatic reduction in ThT fluorescence (not abolition), with few or no fibrils observable by TEM. The formation of fibrils even at high molar ratios of hIAPP:silibinin suggests that the small molecule retards the aggregation process, rather than halting it or driving the formation of off-pathway species. This could potentially explain the negative result observed in the *in vivo* assay, as, although the small molecule specifically binds to hIAPP *in vitro*, the aggregates en route to fibril formation may potentially disrupt  $\beta$ -lactamase activity enough to inhibit bacterial growth in the presence of antibiotics.



**Figure 4.18.** ESI-IMS-MS demonstrates the mode of inhibition (specific/colloidal/non-specific) or lack of inhibition of hIAPP amyloid formation by small molecules. ESI-IMS-MS Driftscope plots of hIAPP alone (a), or in the presence of 10-fold molar excess (320  $\mu\text{M}$  small molecule to 32  $\mu\text{M}$  hIAPP) of (b) Fast Green FCF (bound peaks denoted  $\bullet$  ( $1^{3+}$  bound) or  $\bullet$  ( $1^{2+}$  bound), number of circles represents number of ligands bound), (c) aspirin, (d) 1H-B-SA (bound peaks denoted  $\bullet$ ) or (e) Orange G. An example of a (b) positive, (c) negative (d) non-specific and (e) colloidal inhibitor are illustrated. The numbers on the Driftscope plots indicate the oligomer order and the adjacent superscripted numbers show the charge state of those ions. Experiment performed by Lydia M. Young (University of Leeds).

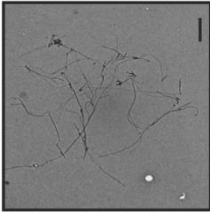
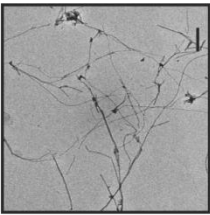
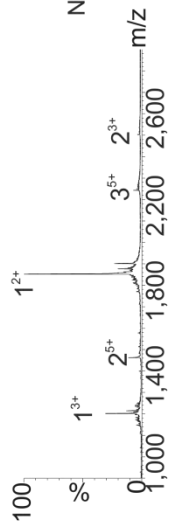
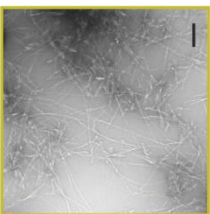

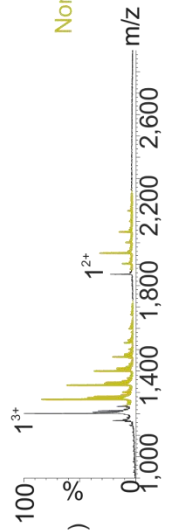
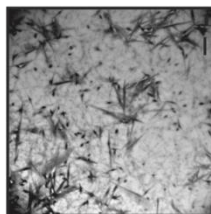
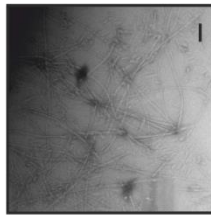
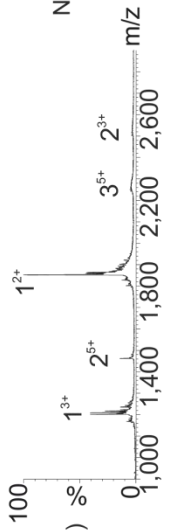
Small molecule	In vivo assay classification		TEM		ThT fluorescence	Mass spectrometry	
			500 nm	100 nm		Spectrum	Binding mode and max oligomers observed
-	-	-			Yes		-
Curcumin	Hit	-			No		Positive
Acid fuchsin	Hit	-			No		Positive

**Table 4.3.** *In vivo* classification compared to *in vitro* hIAPP aggregation in the absence of small molecule, or in the presence of curcumin or acid fuchsin. Negative stain TEM was performed after five days incubation (25 °C, quiescent) of a 1:10 molar ratio of protein:small molecule. Scale bars = 500 and 100 nm. ThT fluorescence of hIAPP alone or in the presence of 1:10 molar ratio of protein:small molecule was observed for 25 h. Positive ion ESI mass spectra. Labels indicate number of ligands (L) bound to each charge state. Binding mode as determined from the mass spectra is denoted as positive, negative, non-specific or colloidal. Maximum number of oligomers observed in IMS-MS is indicated. Mass spectrometry and TEM performed by Lydia M. Young.

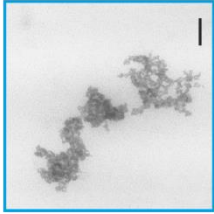
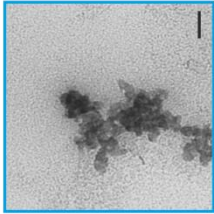

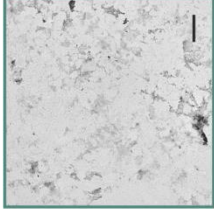
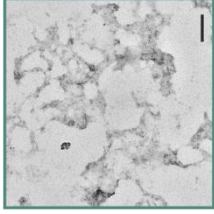

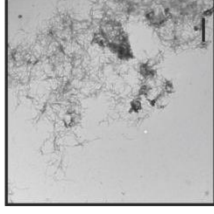
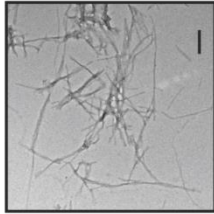
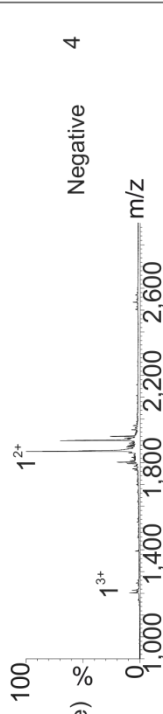
Small molecule	TEM		ThT fluorescence	Mass spectrometry		
	<i>In vivo</i> assay classification	500 nm		100 nm	Spectrum	Binding mode and max oligomers observed
Fast Green FCF	Hit			No		Positive 1
EGCG	Hit			No		Positive 1
Caffeic acid	Hit			No		Negative 5

**Table 4.4** *In vivo* classification compared to *in vitro* hIAPP aggregation in the presence of Fast green FCF, EGCG or caffeic acid.

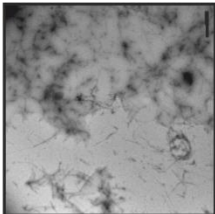
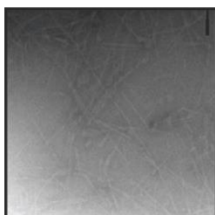
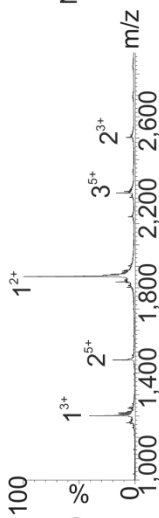
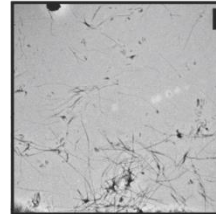
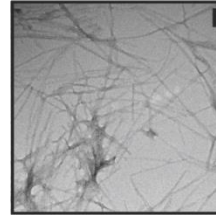
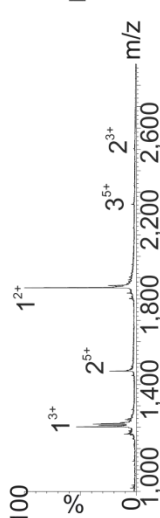
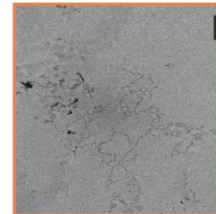
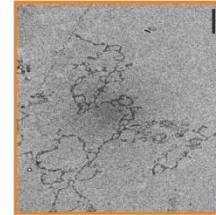
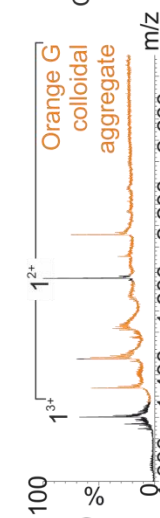
Negative stain TEM was performed after five days incubation (25 °C, quiescent) of a 1:10 molar ratio of protein:small molecule. Scale bars = 500 and 100 nm. ThT fluorescence of hIAPP alone or in the presence of 1:10 molar ratio of protein:small molecule was observed for 25 h. Positive ion ESI mass spectra. Labels indicate number of ligands (L) bound to each charge state. Binding mode as determined from the mass spectra is denoted as positive, negative, non-specific or colloidal. Maximum number of oligomers observed in IMS-MS is indicated. Mass spectrometry and TEM performed by Lydia M. Young.

Small molecule	TEM		ThT fluorescence	Mass spectrometry	
	In vivo assay classification	500 nm		100 nm	Spectrum
Hemin	Negative			Yes	 1 <sup>2+</sup> 1 <sup>3+</sup> 2 <sup>5+</sup> 3 <sup>5+</sup> 2 <sup>3+</sup> Negative 6
1H-B-SA	Negative			No (false positive)	 1 <sup>3+</sup> 1 <sup>2+</sup> Non-specific 1
Azure A	Negative			No (false positive)	 1 <sup>2+</sup> 1 <sup>3+</sup> 2 <sup>5+</sup> 3 <sup>5+</sup> 2 <sup>3+</sup> Negative 5

**Table 4.5.** *In vivo* classification compared to *in vitro* hIAPP aggregation in the presence of hemin, 1H-B-SA or Azure A. Negative stain TEM was performed after five days incubation (25 °C, quiescent) of a 1:10 molar ratio of protein:small molecule. Scale bars = 500 and 100 nm. ThT fluorescence of hIAPP alone or in the presence of 1:10 molar ratio of protein:small molecule was observed for 25 h. Positive ion ESI mass spectra. Labels indicate number of ligands (L) bound to each charge state. Binding mode as determined from the mass spectra is denoted as positive, negative, non-specific or colloidal. Maximum number of oligomers observed in IMS-MS is indicated. Mass spectrometry and TEM performed by Lydia M. Young.

Small molecule	TEM		ThT fluorescence	Mass spectrometry		
	<i>In vivo</i> assay classification	500 nm		100 nm	Spectrum	Binding mode and max oligomers observed
Silibinin	Negative			No		Positive 1
Tramiprosate	Negative			No (false positive)		Non-specific 1
Thiabendazole	Negative			No (false positive)		Negative 4

**Table 4.6.** *In vivo* classification compared to *in vitro* hIAPP aggregation in the presence of silibinin, tramiprosate or thiabendazole. Negative stain TEM was performed after five days incubation (25 °C, quiescent) of a 1:10 molar ratio of protein:small molecule. Scale bars = 500 and 100 nm. ThT fluorescence of hIAPP alone or in the presence of 1:10 molar ratio of protein:small molecule was observed for 25 h. Positive ion ESI mass spectra. Labels indicate number of ligands (L) bound to each charge state. Binding mode as determined from the mass spectra is denoted as positive, negative, non-specific or colloidal. Maximum number of oligomers observed in IMS-MS is indicated. Mass spectrometry and TEM performed by Lydia M. Young.

Small molecule	<i>In vivo</i> assay classification	TEM		ThT fluorescence	Mass spectrometry	
		500 nm	100 nm		Spectrum	Binding mode and max oligomers observed
Benimidazole	Negative			No (false positive)		Negative 6
Aspirin	Negative			Yes		Negative 6
Orange G	Negative			No (false positive)		Colloidal 1

**Table 4.7.** *In vivo* classification compared to *in vitro* hIAPP aggregation in the presence of benzimidazole, aspirin or Orange G. Negative stain TEM was performed after five days incubation (25 °C, quiescent) of a 1:10 molar ratio of protein:small molecule. Scale bars = 500 and 100 nm. ThT fluorescence of hIAPP alone or in the presence of 1:10 molar ratio of protein:small molecule was observed for 25 h. Positive ion ESI mass spectra. Labels indicate number of ligands (L) bound to each charge state. Binding mode as determined from the mass spectra is denoted as positive, negative, non-specific or colloidal. Maximum number of oligomers observed in IMS-MS is indicated. Mass spectrometry and TEM performed by Lydia M. Young.

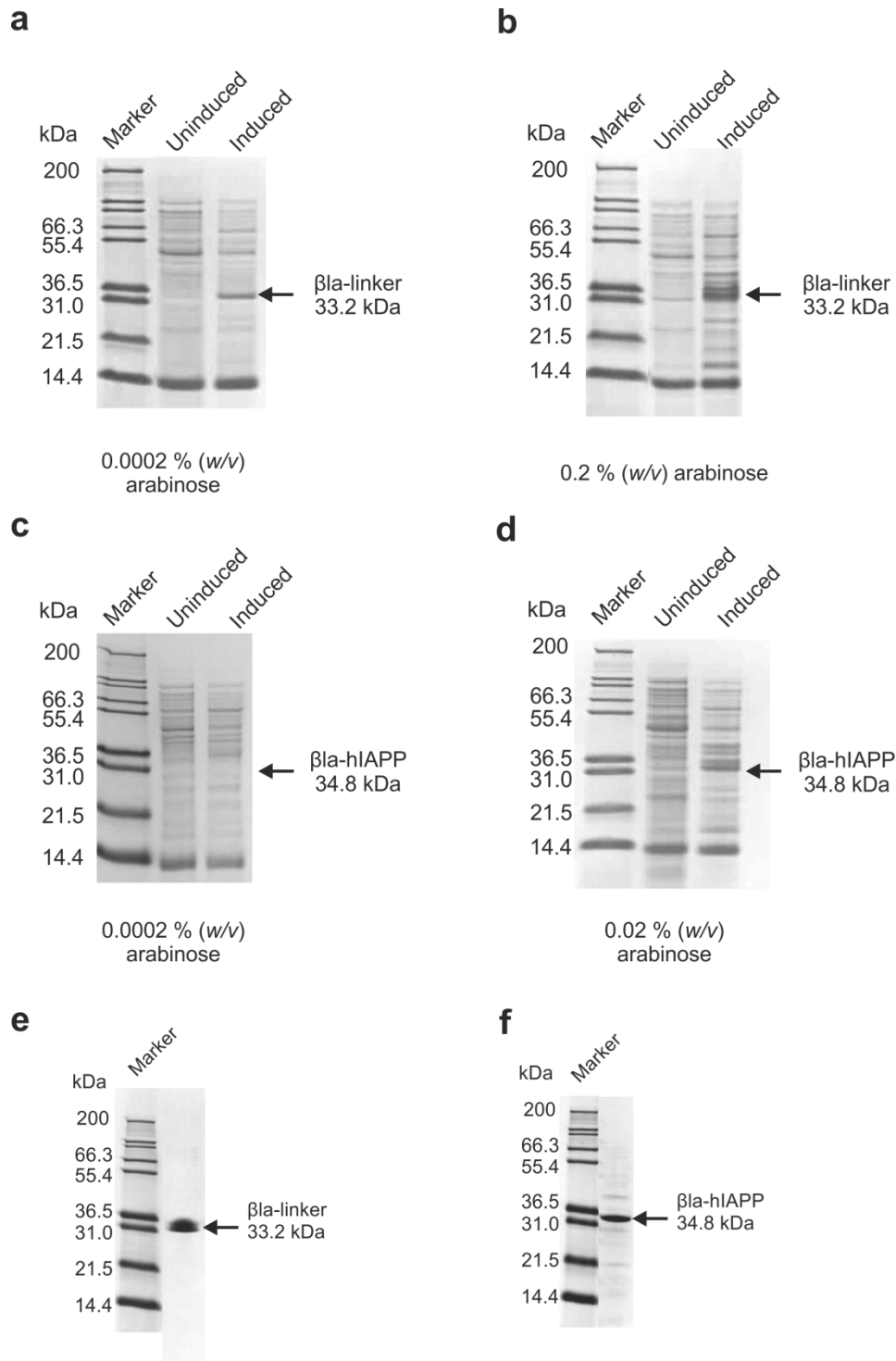


#### 4.3.4 *In vitro* analysis of $\beta$ la-hIAPP aggregation

An assumption made during the development of the *in vivo* assay for inhibitors of protein aggregation was that the aggregating insert will still aggregate when in the  $\beta$ -lactamase construct. To validate this hypothesis, the  $\beta$ la-hIAPP and  $\beta$ la-linker constructs were purified for *in vitro* aggregation analysis.

##### 4.3.4.1 Purification of $\beta$ la-linker and $\beta$ la-hIAPP

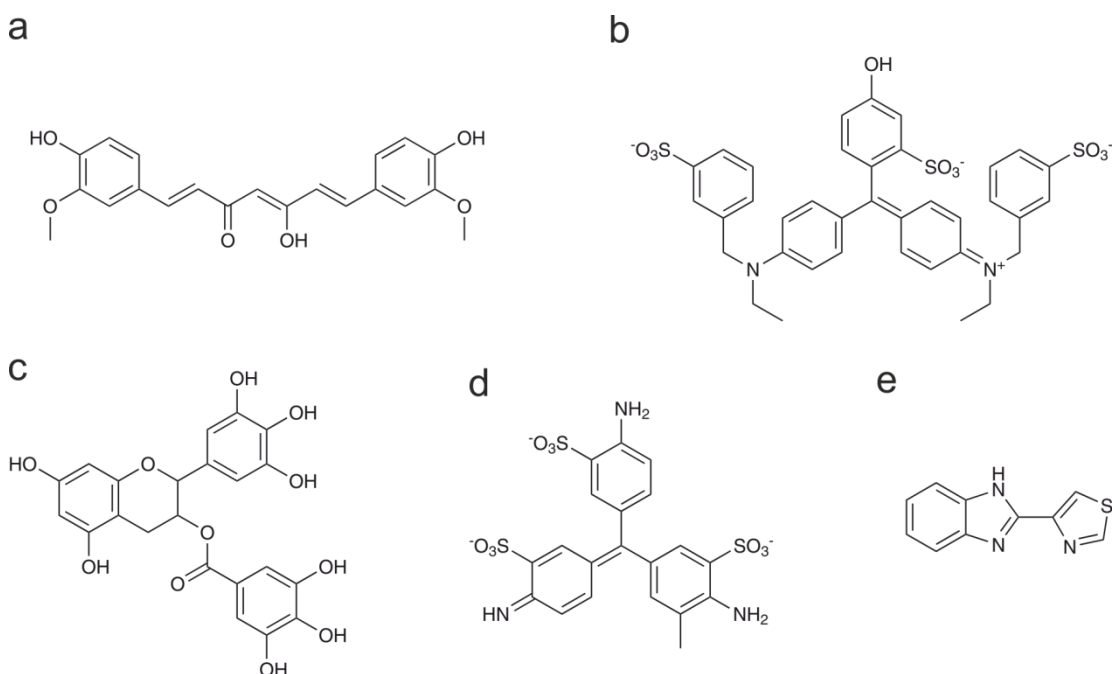
The purification protocol created for wild-type  $\beta$ -lactamase (Section 2.3.1) was used as a basis for the purification of both  $\beta$ la-linker and  $\beta$ la-hIAPP. As the  $\beta$ la-linker and  $\beta$ la-hIAPP constructs are expressed from a pBAD vector, expression trials using 2, 0.2, 0.02, or 0.0002 % (*w/v*) arabinose were performed (Section 2.3.1.1). The results revealed that the highest protein expression was observed in LB medium at 37 °C with an arabinose concentration of 0.2 % (*w/v*) for the  $\beta$ la-linker construct (**Figures 4.19a, b**) and 0.02 % (*w/v*) for the  $\beta$ la-hIAPP construct (**Figures 4.19c, d**). Periplasmic purification of  $\beta$ la-linker and  $\beta$ la-hIAPP was carried out as described in Sections 2.4.14 – 2.3.17. Protein purity was confirmed by SDS-PAGE (**Figures 4.19e, f**) and protein identity was confirmed by mass spectrometry (observed mass of  $\beta$ la-linker: 33219.2 Da, expected mass of  $\beta$ la-linker: 33219.5 Da; observed mass of  $\beta$ la-hIAPP: 34789.7 Da, expected mass of  $\beta$ la-hIAPP: 34789.1 Da).



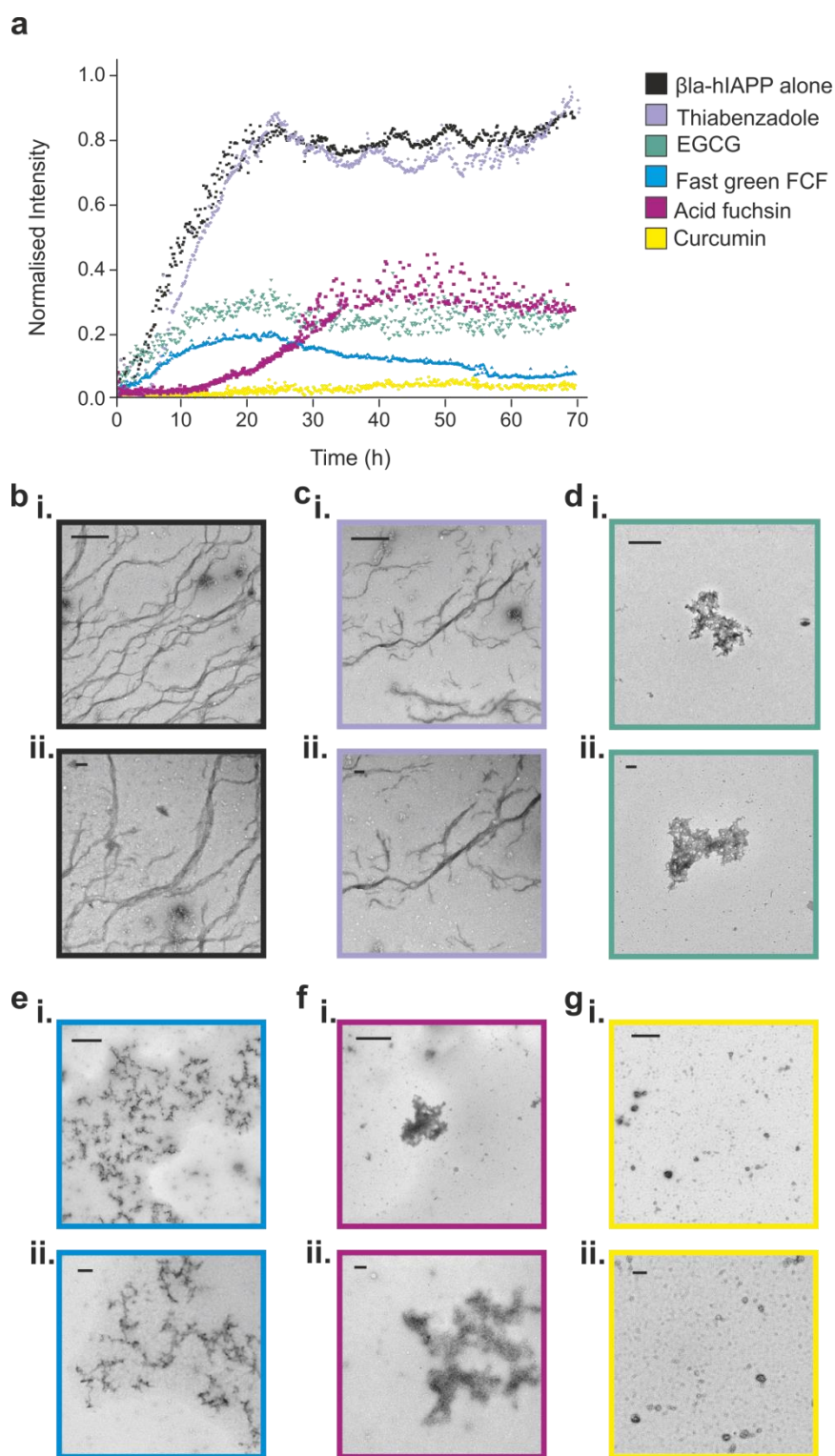
**Figure 4.19.**  $\beta$ la-linker and  $\beta$ la-hIAPP protein expression and purification.  $\beta$ la-linker expression was minimal during induction with 0.0002 % (w/v) arabinose (a) and optimal with 0.2 % (w/v) arabinose (b).  $\beta$ la-hIAPP expression was also minimal under 0.0002 % (w/v) arabinose induction (c) but optimal under 0.2 % (w/v) arabinose (d). Periplasmic fractions extracted before and after induction were analysed. SDS-PAGE gels of (e)  $\beta$ la-linker and (f)  $\beta$ la-hIAPP after size exclusion chromatography. The size in kDa of the protein markers are indicated on the left.

#### 4.3.4.2 *In vitro* aggregation of $\beta$ la-linker and $\beta$ la-hIAPP

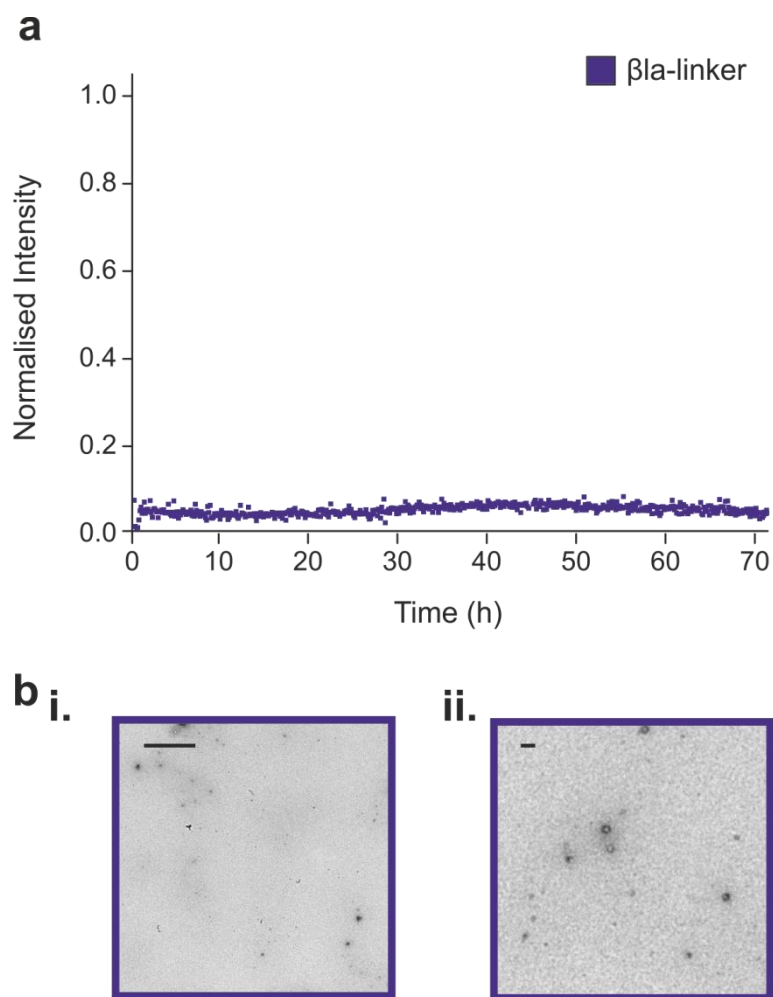
The  $\beta$ la-linker and  $\beta$ la-hIAPP constructs were incubated for 72 h in the presence or absence of four 'hits' from the *in vivo* screen (curcumin, Fast green FCF, EGCG and acid fuchsin) or the non-inhibitor thiabenzadole (**Figure 4.20**). Briefly, a 1:10 molar ratio of protein:small molecule (33:330  $\mu$ M) was incubated at 37 °C for 72 h (Section 2.5.5), after which TEM images were recorded. Nephelometry revealed a clear difference between protein samples that contained inhibitors of hIAPP aggregation and those that did not (**Figure 4.21a**). TEM analysis corroborates this observation, with fibrils only being observed in the sample not containing any small molecule or the sample containing the non-inhibitor of hIAPP aggregation thiabenzadole (**Figure 4.21b, c**). The observation of fibrillar structures confirms that hIAPP incorporation as a  $\beta$ -lactamase tripartite fusion protein can proceed along an aggregation pathway at least similar to that observed by the hIAPP peptide alone. Furthermore, the aggregates produced by  $\beta$ la-hIAPP in the presence of the inhibitors of hIAPP aggregation (**Figure 4.21d-g**) are remarkably similar to those formed by hIAPP peptide in the presence of the same inhibitors (**Table 4.3, Table 4.4**). Conversely, when the  $\beta$ la-linker construct was subjected to the same conditions, turbidity was lower than samples of  $\beta$ la-hIAPP (**Figure 4.22a**), and TEM data revealed the presence of only sparse, small amorphous aggregates (**Figure 4.22b i. ii.**).



**Figure 4.20.** Structures of small molecules used for *in vitro* analysis of  $\beta$ la-hIAPP aggregation. Curcumin (**a**), Fast green FCF (**b**), EGCG (**c**), acid fuchsin (**d**) and thiabenzadole (**e**).



**Figure 4.21.** *In vitro* aggregation of βIa-hIAPP in the presence of different small molecules. (a) Aggregation reactions monitored using turbidity at 635 nm. hIAPP alone, (■) or a 1:10 molar ratio (33:330 μM) of hIAPP:small molecule (thiabendazole (■), EGCG (■), Fast green FCF (■), acid fuchsin (■) or curcumin (■)), was incubated for 72 h at 37 °C. TEM data of aggregates present after 72 h in the absence (b) or presence of thiabendazole (c), EGCG (d), Fast green FCF (e), acid fuchsin (f) or curcumin (g). Scale bars = 500 nm (i) and 100 nm (ii).

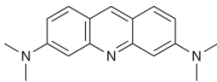
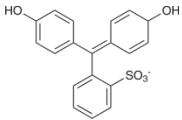
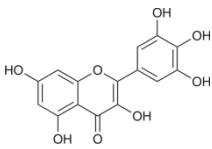
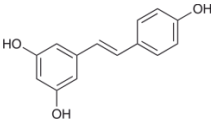
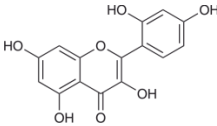
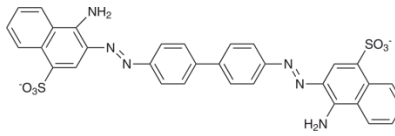


**Figure 4.22.** *In vitro* aggregation of  $\beta$ la-linker. (a) Aggregation reaction of  $33 \mu\text{M}$   $\beta$ la-linker over 72 h at  $37^\circ\text{C}$  monitored using turbidity at 635 nm. One example is shown (three replicates were measured). (b) TEM data of aggregates present after 72 h, scale bar = 500 nm (i) and 100 nm (ii).

## 4.4 Investigating ambiguous inhibitors

To demonstrate the effectiveness of the *in vivo* assay, it was utilised to investigate the effects of six small molecules identified as inhibitors of hIAPP aggregation, but with either incomplete or inconclusive data due to conflicting results (**Table 4.8**)<sup>266, 288, 290, 458-462</sup>. The identification of specific small molecule inhibitors of protein aggregation is complicated by the behaviour of the compounds within the assays used. A traditional method for identifying inhibitors of fibrillation is the ThT assay. Several reports have given evidence of the high false-positive rates obtained using this assay, due to competitive binding of the small molecule to the ThT binding pocket on the amyloid fibril<sup>245, 283, 461</sup>. Indeed, as demonstrated here, ThT data alone cannot be relied on to provide an accurate account of the level of cross  $\beta$ -sheet structure of a sample in the presence of small molecules (**Tables 4.3-4.7**). Despite these well documented disadvantages, some small molecules have been proposed as inhibitors of hIAPP aggregation using this technique alone<sup>266, 460</sup>.

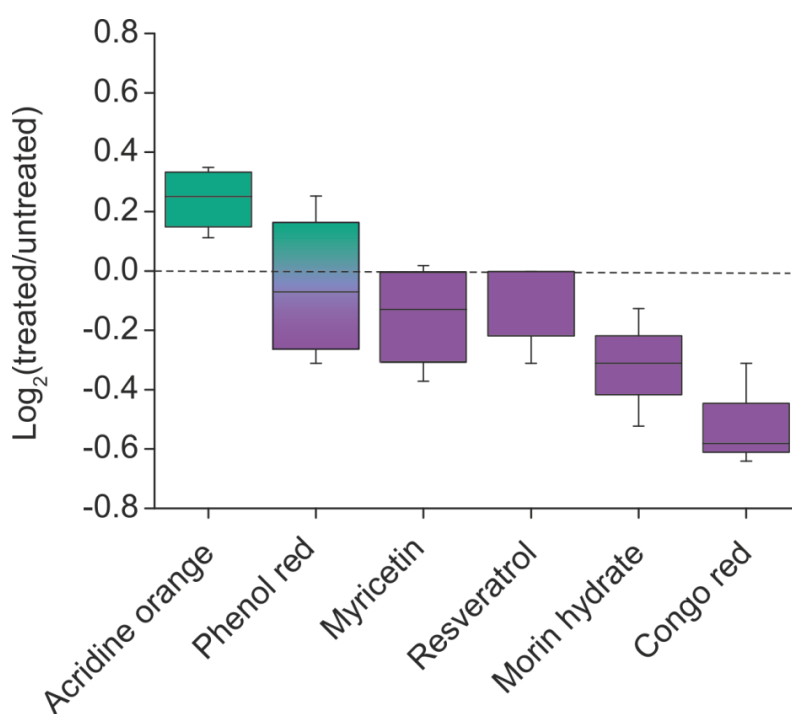
Identification of specific inhibitors of protein aggregation is also complicated by the high number of 'colloidal' inhibitors represented in this group<sup>450, 451</sup>. These promiscuous chemical aggregates are able to sequester proteins in a non-specific manner to prevent aggregation<sup>450, 451</sup>. Colloidal inhibitors are often differentiated from specific inhibitors by investigating their anti-aggregation properties against other target proteins<sup>450</sup>. However, as demonstrated in Sections 4.3.2 and 4.3.3, the *in vivo* screen, in conjunction with ESI-IMS-MS, is able to identify colloidal inhibitors as 'negatives'.

Small Molecule	Mode of Inhibition	MW (Da)	Structure
Acridine orange	20-fold molar excess required for inhibition. Only ThT fluorescence data shown, no TEM data <sup>266</sup>	265.4	
Phenol red	10:1 molar ratio of phenol red:hiAPP leads to small reduction in fibril formation <sup>458</sup> , potentially binds and improves solubility of early protofibrils <sup>459</sup>	354.4	
Myricetin	Published as inhibitor due to lower ThT fluorescence observed in presence of myricetin, but no analyses of aggregates performed <sup>460</sup> . Another study found inhibition occurred for 45 min at a 10:1 molar ratio of myricetin:hiAPP by AFM imaging <sup>288</sup> , however no effect found in another study <sup>461</sup>	318.2	
Resveratrol	Formation of hiAPP amyloid is claimed to be suppressed by a 1:1 <sup>462</sup> and 2:1 <sup>290</sup> molar ratio of resveratrol:hiAPP, however AFM images still show fibrils present	228.2	
Morin hydrate	10:1 molar ratio of morin hydrate:hiAPP leads to formation of short fibrils and amorphous aggregates <sup>461</sup>	302.2	
Congo red	Previously published as an inhibitor of hiAPP aggregation <sup>266</sup> , however recent work has demonstrated it inhibits protein aggregation via a colloidal mechanism <sup>233, 450</sup>	696.7	

**Table 4.8.** Published ambiguous inhibitors of hiAPP aggregation.

In agreement with the previous report<sup>266</sup>, the *in vivo* assay confirmed the small molecule acridine orange as an inhibitor of hiAPP aggregation (**Figure 4.23**). ESI-IMS-MS data reveal acridine orange to be specific inhibitor of hiAPP aggregation, with trimers being the maximum number of oligomers observed at the beginning of the reaction (**Table 4.9**). When a 10-fold molar ratio of acridine orange was incubated with hiAPP for 5 days, only sparse amorphous aggregates are visible by TEM (**Table 4.9**).

Four of the ambiguous inhibitors of hIAPP aggregation were negative in the *in vivo* screen: myricetin, resveratrol, morin hydrate and Congo red (**Figure 4.23**). The small molecule phenol red was classified as a hit in only 50 % of the experiments (n=4). ESI-MS spectra confirm that phenol red, myricetin, resveratrol and morin hydrate bind hIAPP non-specifically, whereas Congo red binds by a colloidal mechanism (**Table 4.9** and **Table 4.10**). TEM data support the *in vivo* data, with fibrillar material visible after 5 days incubation of hIAPP with a 10-fold molar excess of phenol red, myricetin, resveratrol and morin hydrate (**Table 4.9**, **Table 4.10**). Similar to the aggregates of a 1:10 molar ratio of hIAPP:Orange G viewed previously (**Table 4.7**), incubation of hIAPP with the colloidal inhibitor Congo red also results in very large amorphous aggregates (**Table 4.10**).



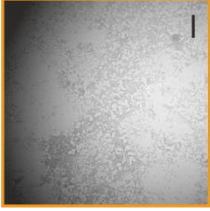
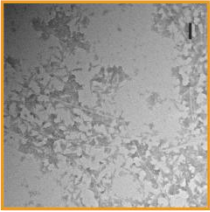
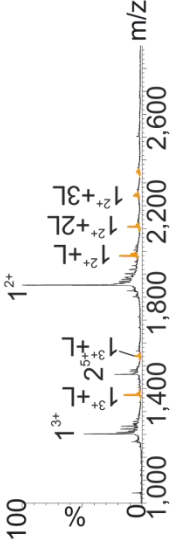
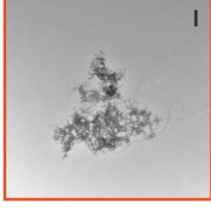
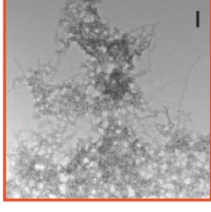
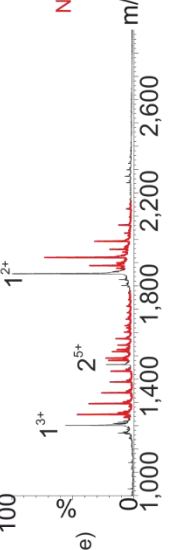
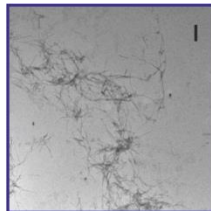
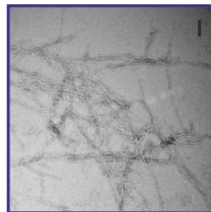
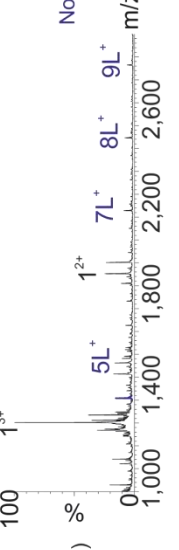
**Figure 4.23.** *In vivo* analysis of six ambiguous inhibitors of hIAPP aggregation. Box plot (n=4) showing the effect on growth of each small molecule as a comparison of treated versus untreated cell growth. Centre line = median; box limits indicate the 25<sup>th</sup> and 75<sup>th</sup> percentiles, with whiskers extending to  $\pm 1.5$  times the interquartile range. Dashed line through graph indicates separation of small molecules that have an effect ( $> 0$ ), and those that do not ( $\leq 0$ ). Compounds are coloured as positive inhibitors (●) or negative-inhibitors (●) of hIAPP aggregation *in vitro*.



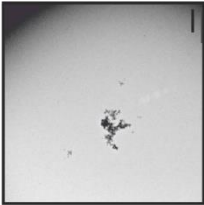
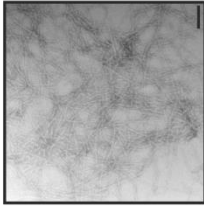
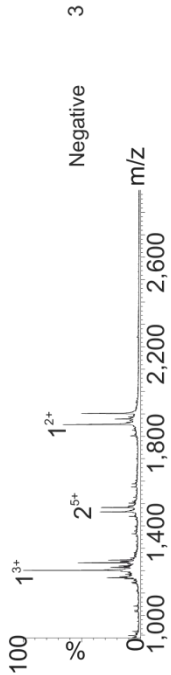
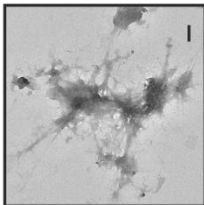
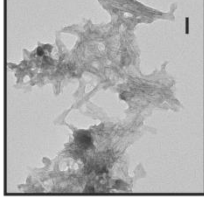
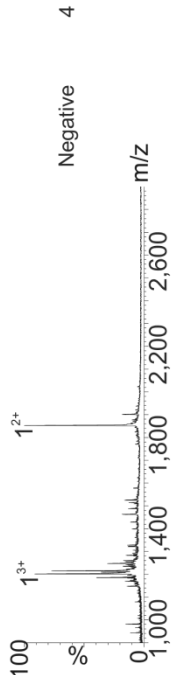
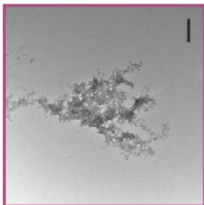
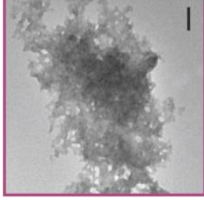
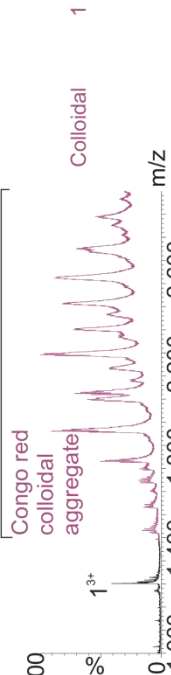
A twenty-fold molar excess of Congo red was previously reported to prevent hIAPP aggregation *in vitro*<sup>266</sup>, however the data presented here clearly show that Congo red acts via a colloidal mechanism. The previous study only provided ThT data, with no visualisation of the aggregates by TEM or atomic force microscopy (AFM)<sup>266</sup>. In a similar vein, the small molecule myricetin was reported as an inhibitor of hIAPP aggregation based solely on ThT data<sup>460</sup>. A different study combined ThT data with AFM imaging to analyse the aggregates formed by hIAPP in the presence of a 10-fold molar ratio of myricetin, however the images were captured after only 45 min<sup>288</sup>. There is a slight decrease in aggregate load when compared to hIAPP in the absence of myricetin<sup>288</sup>, however, this is most likely due to non-specific interactions of myricetin with hIAPP. As clearly shown in the work presented here (**Table 4.9**), and in another published study<sup>461</sup>, myricetin does not prevent hIAPP fibril formation when incubated at a 10-fold molar ratio of small molecule.

Morin hydrate was classified as a non-inhibitor of hIAPP both *in vivo* and *in vitro*. This is in stark contrast to a recent publication of morin hydrate as an inhibitor of hIAPP aggregation<sup>461</sup>. TEM data presented here show large amorphous aggregates, together with distinct fibrillar material (**Table 4.10**). Furthermore, ESI-IMS-MS shows that morin hydrate does not bind to hIAPP, with oligomers up to tetramer observed at the beginning of the incubation. The published study found morin hydrate reduced fibrillar sample load at 1:5 and 1:10 molar ratio of hIAPP: morin hydrate (as observed by TEM), however small plaques of shorter fibrils were still present<sup>461</sup>.

Interestingly, resveratrol was also classified as a non-inhibitor of hIAPP aggregation, in disagreement with a number of published studies<sup>289, 290, 463</sup>. One study<sup>290</sup> reported a decrease in oligomer number (at 18 h) and fibrillar load (at 96 h) when a 1:0.5 molar ratio of hIAPP:resveratrol was incubated at 25 °C, as visualised by AFM. However, the decrease is negligible, and the fibril morphology appears identical. Conversely, other published studies demonstrated that the presence of resveratrol prevents hIAPP from forming fibrils on the surface of INS-1E (rat pancreas) cells for up to 48 h; small amorphous aggregates are all that are visible<sup>289, 463</sup>. It may be that resveratrol is able to slow hIAPP fibrillation significantly by some currently unknown method. The presence of dense fibrillar material after five days incubation presented here may be the result of slower accumulation of amyloid fibril in the presence of resveratrol. This hypothesis could be tested via a time-point assay, with aggregated material removed at 12 h intervals and the aggregation process monitored by TEM.

Small molecule	<i>In vivo</i> assay classification	TEM		ThT fluorescence	Mass spectrometry	
		500 nm	100 nm		Spectrum	Binding mode and max oligomers observed
Acridine Orange	Hit			No		Positive
Phenol red	Inconclusive			No (false positive)		Non-specific
Myricetin	Inconclusive			No (false-positive)		Non-specific

**Table 4.9.** *In vivo* classification compared to *in vitro* hIAPP aggregation in the presence of acridine orange, phenol red or myricetin. Negative stain TEM was performed after five days incubation (25 °C, quiescent) of a 1:10 molar ratio of protein:small molecule. Scale bars = 500 and 100 nm. ThT fluorescence of hIAPP alone or in the presence of 1:10 molar ratio of protein:small molecule was observed for 25 h. Positive ion ESI mass spectra. Labels indicate number of ligands (L) bound to each charge state. Binding mode as determined from the mass spectra is denoted as positive, negative, non-specific or colloidal. Maximum number of oligomers observed in IMS-MS is indicated. Mass spectrometry and TEM performed by Lydia M. Young.

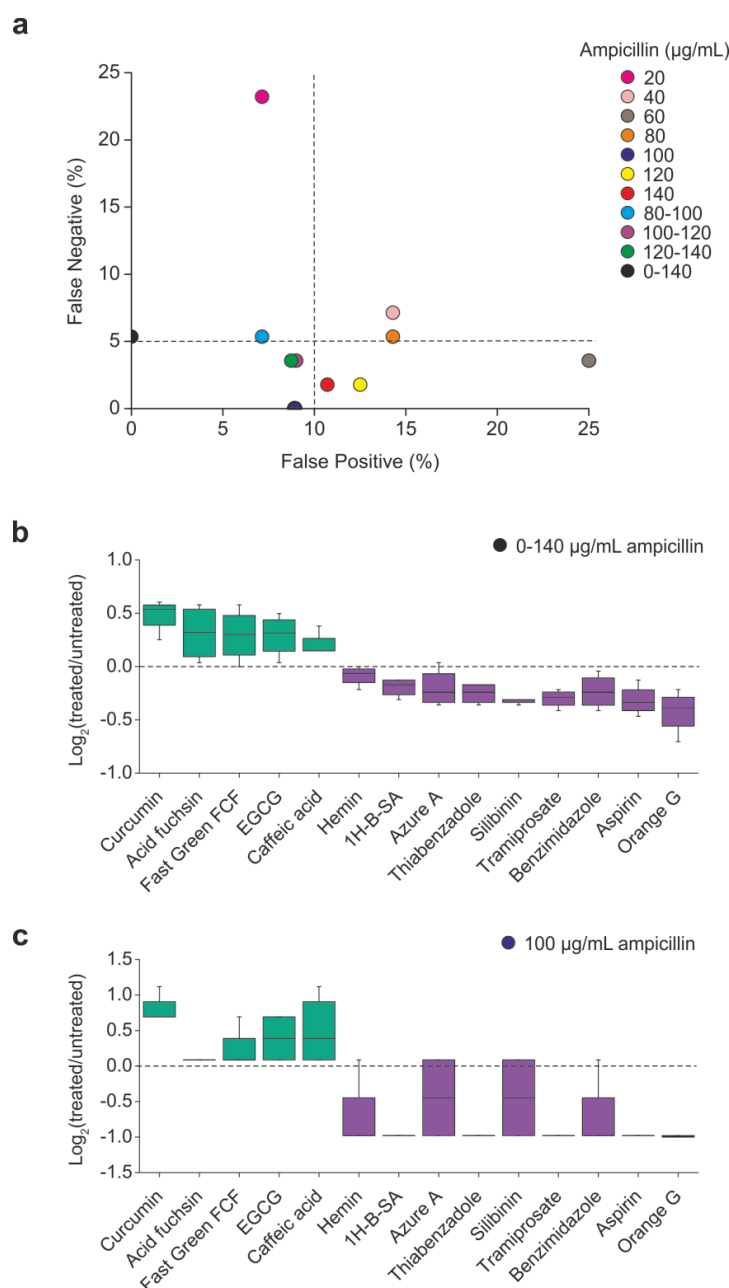
Small molecule	TEM		ThT fluorescence	Mass spectrometry		
	In vivo assay classification	500 nm		100 nm	Spectrum	Binding mode and max oligomers observed
Resveratrol	Negative			No (false-positive)		Negative 3
Morin hydrate	Negative			No (false positive)		Negative 4
Congo red	Colloidal			No (false positive)		Colloidal 1

**Table 4.10.** *In vivo* classification compared to *in vitro* hIAPP aggregation in the presence of resveratrol, morin hydrate or Congo red. Negative stain TEM was performed after five days incubation (25 °C, quiescent) of a 1:10 molar ratio of protein:small molecule. Scale bars = 500 and 100 nm. ThT fluorescence of hIAPP alone or in the presence of 1:10 molar ratio of protein:small molecule was observed for 25 h. Positive ion ESI mass spectra. Labels indicate number of ligands (L) bound to each charge state. Binding mode as determined from the mass spectra is denoted as positive, negative, non-specific or colloidal. Maximum number of oligomers observed in IMS-MS is indicated. Mass spectrometry and TEM performed by Lydia M. Young.

## 4.5 Assay miniaturisation for high throughput screening

One of the key stages in the development of a high throughput assay is miniaturisation, with the aim to reduce screening costs and increase speed, while not compromising data quality. In this context, as the current *in vivo* assay setup requires an excessive amount of small molecule (288  $\mu\text{g}$  for a 200 Da compound), and requires two 48-well plates, miniaturisation was deemed a further important development. To achieve this, a single, or minimum of two, concentrations of ampicillin was searched for that could give the same reliable and reproducible results as the full plate of eight concentrations of ampicillin. The methodology as before was utilised, however instead of calculating the difference between the areas under the MIC assay curve in the presence and absence of small molecules, the difference between single scores was calculated.

Using the data presented in **Figure 4.15** (presented again in **Figure 4.24b**), for each of the 14 original test molecules the number of false positive and false negative scores was tallied at each ampicillin concentration (rather than over all ampicillin concentrations used as described in Section 4.3.2). The total number of false positives and negatives, if only a single concentration of ampicillin would have been used, was determined (**Figure 4.24a**). For the full MIC assay, a false positive rate of 0 % and a false negative rate of 5.1 % was observed (**Figure 4.24a**). In any screen, a false positive is an unwanted result, as it increases both costs and time of the assay through determination of the nature of the false hit in subsequent analyses. However, false positive results can be detected and readily nullified by a second assay. For this reason a threshold maximum of 10 % false positives was set. The data obtained for concentrations of ampicillin from 2 ampicillin concentrations (20 and 100  $\mu\text{g}/\text{mL}$ ) and three ampicillin concentration ranges (80-100, 100-120 and 120-140  $\mu\text{g}/\text{mL}$ ) fell below this threshold, with false positive rates ranging from 7.1 % - 8.2 % (**Figure 4.24a**). Reducing the number of false negative results in an assay is imperative as these cannot be determined at a later stage, therefore, the false negative rate threshold was set at 5 %. Six concentrations of ampicillin produced a false negative rate below the threshold, however only three simultaneously had a false positive rate of below 10 %: 100  $\mu\text{g}/\text{mL}$ , 100-120  $\mu\text{g}/\text{mL}$  and 120-140  $\mu\text{g}/\text{mL}$ . An ampicillin concentration of 100  $\mu\text{g}/\text{mL}$  produced no false negatives, and as this condition only requires one concentration of ampicillin (and therefore only one row of wells in a 48-well plate), it was selected as the concentration at which the miniaturised assay should be performed. **Figures 4.24b-c** show the comparison of the results of the full MIC assay (**Figure 4.24b**) and the miniaturised MIC assay at 100  $\mu\text{g}/\text{mL}$  ampicillin (**Figure 4.24c**).



**Figure 4.24.** Comparison of false-positive and false-negative rates during assay miniaturisation. **(a)** The number of false-positive and false-negative rates were calculated at each concentration of ampicillin, and also over three ranges of ampicillin. The rates for the full MIC assay (●) are included for comparison. 14 small molecules were assayed ( $n=4$ ), with the rates calculated as a percentage of the total number of assays ( $n=56$ ). The threshold maximum percentage, 5 % for false-negatives and 10 % for false-positives, is indicated as a dashed line. Box plots ( $n=4$ ) show the effect on growth of each small molecule as a comparison of treated versus untreated cell growth. **(b)** Full MIC assay at eight concentrations of ampicillin. **(c)** MIC assay at only 100  $\mu\text{g/mL}$  ampicillin. Centre line = median; box limits indicate the 25<sup>th</sup> and 75<sup>th</sup> percentiles, with whiskers extending to  $\pm 1.5$  times the interquartile range. Dashed line through graph indicates separation of small molecules that have an effect ( $> 0$ ), and those that do not ( $\leq 0$ ). Compounds are coloured as positive inhibitors (●) or non-inhibitors (●) of hIAPP aggregation *in vitro*.

## 4.6 Investigating unknown inhibitors

Miniaturisation of the  $\beta$ -lactamase tripartite fusion assay for identifying inhibitors of hIAPP *in vivo* enabled a screen of a library of novel molecules (with structural similarity to the aggregation inhibitors previously reported) to be performed. Focused screening is a well-versed method to improve the hit-rate of a high throughput screen by seeding a screening library with compounds which have a higher probability to inhibit, or bind to, the target compared with random screening<sup>464</sup>. The screening method uses the structural information from known bio-active ligands to identify novel compounds with similar structure, and hence biological activity. Five known inhibitors of hIAPP and/or A $\beta$ 42 aggregation (vanillin<sup>465</sup>, resveratrol<sup>466</sup>, curcumin<sup>445</sup>, chloronaphthoquinine-tryptophan<sup>467</sup> and EGCG<sup>468</sup>) were selected as queries to seed a focussed library of compounds for screening (**Section 2.5.9**). The seeding process involved an assessment of each of the inhibitors for structural similarity to an in-house, 50,000-member, structurally diverse library of lead-like small molecules using the programme Rapid Overlay of Chemical Structures (ROCS)<sup>384</sup>, followed by cherry-picking a subset of 20 compounds for 'wet' screening based on the default comparator (ROCS Combiscore) with consideration to maximal structural diversity of the proposed screening set. All seeding and molecule selection was performed by Dr Charlotte H. Revill and Dr Richard J. Foster (University of Leeds).

The 20 novel compounds were first screened using ESI-IMS-MS (Section 2.5.8). One compound was found to specifically bind to hIAPP, three demonstrated non-specific binding to hIAPP, and the remainder did not bind *in vitro* (**Table 4.11**; data not shown). The three non-specific binders and one specific binder were blind-tested *in vivo* for their anti-aggregation properties at a single concentration of ampicillin (100  $\mu\text{g}/\text{mL}$ ). In agreement with the mass spectrometry data, only the specific binder (as determined by ESI-MS) prevented  $\beta$ la-hIAPP aggregation *in vivo* (**Table 4.11**). TEM data of hIAPP incubated for five days in the presence of the compounds corroborated the mass spectrometry and  $\beta$ -lactamase tripartite fusion system findings; only in the presence of a 10-fold molar excess of the *in vivo* 'hit' molecule was fibril formation inhibited and the formation of amorphous aggregates resulted (**Table 4.11**).

Small molecule	Structure	Mass spectrum	ESI-MS binding mode classification	<i>In vivo</i> assay score $\text{Log}_2(\text{treated}/\text{untreated})$	TEM
5,7-dihydroxy-2-[(4-methoxyphenyl)amino]thieno[3,2-b]pyridine-3-carbonitrile			Non-specific	-0.21	
N-(2,3-dihydro-1,4-benzodioxin-6-yl)-2-[4-[(4-methyl-6-oxo-1,6-dihydropyrimidin-2-yl)amino]piperidin-1-yl]acetamide			Non-specific	-0.36	
1-(Adamantan-1-ylcarbonyl)-1H-spiro[piperidine-4,2'-quinazolin]-4(3H)-one			Non-specific	-0.28	
6-[(4-(2-fluorophenyl)-1-piperazinyl)carbonyl]-3-methyl-5H-[1,3]thiazolo[3,2-b]pyrimidin-5-one			Positive	0.36	

**Table 4.11.** *In vivo* and *in vitro* analysis of four unknown compounds for their ability to bind to hIAPP, and/or inhibit hIAPP aggregation.

Positive ion ESI mass spectra. Labels indicate number of ligands (L) bound to each charge state (green: non-specific, red: specific). Binding mode as determined from the mass spectra is denoted as non-specific or positive. The *in vivo* effect on growth of each small molecule is given as a comparison of treated versus untreated cell growth. A score > 0 is classified as a 'hit'. Negative stain TEM was performed after five days incubation (25 °C, quiescent) of a 1:10 molar ratio of protein:small molecule. Scale bars = 500 and 100 nm. Mass spectrometry and TEM performed by Lydia M. Young.

## 4.7 Discussion

Screening for small molecule inhibitors of protein aggregation often requires the laborious, and costly *in vitro* purification of protein for the initial rounds of screening. The use of an *in vivo* system circumvents the difficulties faced with the purification of these challenging proteins, and as such, it significantly reduces the time and cost of the analyses. In Chapter 3 it was demonstrated that the  $\beta$ -lactamase tripartite fusion system could be used successfully to differentiate between aggregating and non-aggregating test proteins. It was hypothesised that if the bacteria expressing the tripartite fusion protein were exposed to small molecule inhibitors of protein aggregation, the bacteria would grow at higher concentrations of the antibiotic ampicillin.

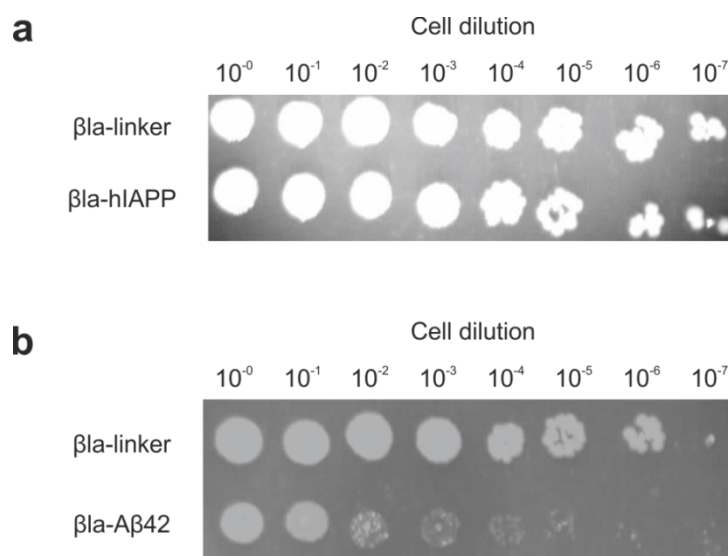
### 4.7.1 Assay development

The addition of another variable to the *in vivo* system (the small molecule) required significant consideration and optimisation to ensure the minimal level of unwanted side effects, false positives and false negatives occurred. Initial experiments with the  $\beta$ la-hIAPP construct were positive, with rescue of bacterial growth observed in the presence of select known inhibitors of hIAPP aggregation (curcumin and EGCG) (**Figure 4.2**). However, the assay was not successful under the same conditions using the  $\beta$ la-A $\beta$ 42 construct (**Figure 4.2**). It was hypothesised that the level of aggregation exhibited by A $\beta$ 42 was severely toxic to the cells, so the level of protein induction was optimised to reduce the concentration of A $\beta$ 42 present in the cell, without going below the concentration threshold of  $\beta$ -lactamase that the cells required to grow (**Figure 4.3**).

One potential complication of the *in vivo* assay is the apparent toxicity of amyloidogenic proteins to bacteria. Initially it was believed that the significant growth difficulties experienced by bacteria expressing either the  $\beta$ la-A $\beta$ 40 (Chapter 3) or  $\beta$ la-A $\beta$ 42 constructs could be attributable to the severe aggregation of the A $\beta$  proteins hindering  $\beta$ -lactamase activity. Although this toxicity was alleviated somewhat by increasing  $\beta$ -lactamase activity (by increasing protein expression levels) while limiting A $\beta$  production, even in the presence of numerous well-known inhibitors of A $\beta$  aggregation, bacterial growth was not significantly improved. Furthermore, it became apparent that even when the bacteria were not subjected to high antibiotic concentrations, the presence of either  $\beta$ la-A $\beta$  construct was deleterious. For example, **Figure 4.25** shows the colonies



observed on a plate containing 20  $\mu\text{g}/\text{mL}$  ampicillin and 100  $\mu\text{M}$  of the known small molecule inhibitor of hIAPP and A $\beta$ 42 aggregation, curcumin<sup>244, 245, 446</sup>. Although for both the  $\beta\text{la-hIAPP}$  and  $\beta\text{la-A}\beta$ 42 constructs the maximal cell dilution at which colonies can be observed is  $10^{-7}$ , the size of the colonies differs drastically for the  $\beta\text{la-A}\beta$ 42 expressing bacteria. These significantly smaller colonies were universally observed in all MIC assays performed with both the  $\beta\text{la-A}\beta$ 40 and  $\beta\text{la-A}\beta$ 42 constructs even in the presence of potent anti-aggregation small molecules.



**Figure 4.25.** Example of disparity in colony size produced by different tripartite fusion proteins. Cell dilutions of bacteria expressing either (a)  $\beta\text{la-linker}$  or  $\beta\text{la-hIAPP}$ , or (b)  $\beta\text{la-linker}$  or  $\beta\text{la-A}\beta$ 42, were spotted onto an agar plate containing 20  $\mu\text{g}/\text{mL}$  ampicillin and 100  $\mu\text{M}$  curcumin. After 18h incubation at 37  $^{\circ}\text{C}$ , the maximal cell dilution at which growth occurs ( $\text{MCD}_{\text{GROWTH}}$ ) for all construct-expressing cells is  $10^{-7}$ . Colonies of bacteria expressing  $\beta\text{la-A}\beta$ 42 are significantly smaller than colonies of bacteria expressing either  $\beta\text{la-linker}$  or  $\beta\text{la-hIAPP}$ .

A literature review revealed accumulating evidence that there may be a relationship between amyloid formation and antimicrobial activity<sup>469, 470</sup>. Host defence peptides (HDPs) are antimicrobial agents of the innate immune system, many of which are able to form amyloid fibrils<sup>469</sup>. Many of these small peptides are intrinsically disordered in solution but adopt  $\alpha$ -helical structure at membrane interfaces, which leads to self-assembly and membrane insertion. These  $\alpha$ -helical species undergo further conformational changes into  $\beta$ -sheet-rich structures, leading to further aggregation and amyloid-like fibril formation<sup>470</sup>.

These species disrupt cellular membranes (bacteria, fungi, enveloped viruses etc.), leading to membrane leakage and destruction of the microbe. The striking similarity between antimicrobial membrane damage and membrane damage attributed to disease-related amyloidogenic proteins raises the question: if antimicrobial peptides can form amyloid fibrils, can amyloid peptides also act as antimicrobials? The answer appears to be yes. A number of studies have shown that both A $\beta$ 42 and A $\beta$ 40 have potent antimicrobial activities<sup>471, 472</sup>. Brain tissue from patients with Alzheimer's disease exhibits antimicrobial activity levels significantly higher than tissue from healthy subjects<sup>471</sup>. Furthermore, A $\beta$ 40 and A $\beta$ 42 (at micromolar levels) can impede the growth of *Streptococcus pneumoniae*, the organism responsible for the majority of cases of bacterial meningitis<sup>472</sup>. It has been suggested, therefore, that these peptides may actually serve as HDPs for the innate immune system<sup>469</sup>. Consistent with the mechanism by which HDPs interact with lipid membranes<sup>473, 474</sup>, A $\beta$ 40 and A $\beta$ 42 have been found to bind to and penetrate lipid mimics of bacterial membranes<sup>471</sup>. The peptides convert from an unstructured conformation to a  $\beta$ -sheet structure with amphiphilic characteristics when they interact with anionic membranes<sup>471, 475, 476</sup>. So although the mechanism by which these peptides exert their antimicrobial activity remains unconfirmed, it can be suggested that it most probably involves membrane permeabilisation.

In the context of the *in vivo* assay, the potential antimicrobial activity of A $\beta$ 40 and A $\beta$ 42 raises a number of interesting questions. If the peptides are indeed antimicrobial, and interact with the anionic outer layer of bacterial membranes, how would they exert this activity from within the periplasm? The inner leaflet of the outer membrane of Gram negative bacteria does not contain the layer of negatively charged lipopolysaccharide (LPS) found on the outer leaflet. However, by expressing the A $\beta$  peptides within the periplasmic space, the close proximity of peptide and membrane may be sufficient to trigger peptide:membrane interactions. Alternatively, lysis of any bacteria within the culture would result in the externalisation of the  $\beta$ la-A $\beta$  constructs. This would enable the A $\beta$  peptides to interact with the exterior membranes of other bacteria within the culture. Membrane disruption of these bacteria would result in the release of more  $\beta$ la-A $\beta$ , potentially leading to a toxic cascade.

A second question raised is how do the A $\beta$  peptides *insert* into lipid membranes? It is well established that A $\beta$ 40 and A $\beta$ 42 oligomers form pores and ion-conducting channels in lipid membranes<sup>477, 478</sup>, but what is the peptide's orientation during insertion? Molecular dynamics simulations suggest A $\beta$  peptides insert into the hydrophobic region of the membrane via their C-terminal or central hydrophobic region<sup>479</sup>. Indeed, this is supported

by studies demonstrating that A $\beta$ 42 exhibits much higher levels of binding to anionic membranes and antimicrobial activity when compared with A $\beta$ 40<sup>471, 475</sup>. The additional C-terminal hydrophobic residues isoleucine and alanine in A $\beta$ 42, it seems, aid in this interaction. How, then, would these peptides insert or disrupt a lipid membrane when they are constrained as a tripartite fusion in the *in vivo* assay? It may be that full insertion is not necessary to disrupt the membrane. Interacting with, or aggregating on, the inner surface of the outer membrane may be sufficient for bacterial toxicity.

To answer these questions, *in vitro* analysis of the effects of  $\beta$ la-A $\beta$  on lipid micelle structure could be performed. If the addition of the tripartite fusion protein caused membrane leakage (monitored by dye release), whereas  $\beta$ la-linker did not, it would indeed suggest that the insertion of A $\beta$ 40 or A $\beta$ 42 into  $\beta$ -lactamase produces a membrane-disrupting entity. Furthermore, small molecule A $\beta$  channel blockers could be utilised to investigate whether the  $\beta$ la-A $\beta$  constructs can form pores from within the periplasm<sup>480</sup>. Nevertheless, the ability of small molecule inhibitors of A $\beta$  aggregation to rescue bacterial growth in the assay, albeit to a less extent to the level of rescue observed against  $\beta$ la-hIAPP, suggests that the assay is still applicable to potentially antimicrobial peptides.

Assay setup and conditions were investigated to obtain an assay which was both simple and rapid to perform without sacrificing assay accuracy. Initially a bacterial lawn approach was investigated, however, despite some promising positive results, the setup and scoring method was deemed too laborious and error-prone. By returning to the original MIC assay plates, scoring difficulties were eradicated, and the assay protocol was significantly simplified. Numerous assay conditions were optimised before a full screen of known positive and negative inhibitors of protein aggregation could be performed; **Figure 4.10** shows a schematic of the final methodology developed. A significant step was accounting for any off-target effects of the small molecule on bacterial growth. By performing the assay with the non-aggregating  $\beta$ la-linker construct, any effects of the small molecule on growth could be quantified and accounted for. Interestingly, it was found that, particularly for the non-inhibitors of protein aggregation, there were additional toxic effects experienced by the bacteria expressing the aggregating constructs. This suggests that toxic effects originating from the presence of the small molecules is not independent of the toxicity induced by the presence of an aggregating protein. Together, they act synergistically, leading to a much more severe toxic profile than predicted. This synergistic toxicity is minimised in the presence of inhibitors of  $\beta$ la-hIAPP aggregation *in vivo*, and therefore does not interfere with the purpose of the assay, which is to identify positive inhibitors of protein aggregation.

#### 4.7.2 The *in vivo* system as a screen for small molecules inhibitors of protein aggregation

A primary stage in assay development is confirmation that the assay can differentiate between known positive and negative reference controls. To this end, six published inhibitors of hIAPP aggregation were selected from the literature as positive controls. A combination of TEM and ESI-IMS-MS was utilised to identify eight small molecules that do not inhibit hIAPP aggregation to use as negative reference controls in the *in vivo* screen. The *in vivo* results corroborated with the *in vitro* analyses; small molecules that did not bind, bound non-specifically or bound colloiddally were all negative in the *in vivo* assay. The positive reference controls were all 'hits' except for the published inhibitor of hIAPP aggregation, silibinin. As discussed previously, the presence of low levels of ThT fluorescence even after incubation of 5- and 10-fold molar ratio of silibinin over hIAPP<sup>237</sup>,<sup>287</sup> suggests the small molecule is only retarding the aggregation process, not inhibiting it. Increasing the concentration of silibinin to 500  $\mu\text{M}$  in the *in vivo* assay produces a 'hit'. Thus, by performing the assay at only 100  $\mu\text{M}$  small molecule, only potent, specific inhibitors of aggregation will be selected for.

The  $\beta\text{la}$ -linker and  $\beta\text{la}$ -hIAPP constructs were purified for *in vitro* analysis of the aggregates formed.  $\beta\text{la}$ -hIAPP formed fibrillar structures after 72 h incubation at 37 °C. As no fibrillar or amorphous aggregates were present when  $\beta\text{la}$ -linker was subjected to the same conditions, it can be concluded that hIAPP is able to fibrillate even when contained as a tripartite fusion protein. When  $\beta\text{la}$ -hIAPP was incubated separately with four inhibitors of hIAPP aggregation (curcumin, acid fuchsin, Fast green FCF or EGCG), sample turbidity was significantly reduced when compared to a sample of  $\beta\text{la}$ -hIAPP alone. Analysis of TEM images after 72 h showed a distinct lack of fibrillar material, with small amorphous aggregates being the only visible entities. In contrast, incubation with a small molecule that does not inhibit hIAPP aggregation (thiabendazole) had no measurable effect. These data support the validity and utility of the *in vivo* tripartite fusion system as a screen for inhibitors of aggregation. As aggregation of the hIAPP test protein is able to progress along a fibrillation pathway, identity of both fibrillar and non-fibrillar aggregation inhibitors is possible.

The success of the assay in differentiating between known inhibitors and compounds known to have no effect on hIAPP aggregation led to the assessment of six small molecule inhibitors with ambiguous effects. The *in vivo* assay, in conjunction with ESI-IMS-MS, identified five of these small molecules as negative, that is, they did not prevent protein

aggregation *in vivo* or *in vitro*. The disparity between the published work and the data presented here is due to the techniques employed to assess the ability of small molecules to inhibit aggregation. By only assaying low concentrations of the small molecule *in vivo*, the  $\beta$ -lactamase tripartite fusion system does not identify small molecules that prevent hIAPP aggregation in a colloidal manner. This is a key requirement of any small molecule screen as this type of interaction is undesirable as a therapeutic.

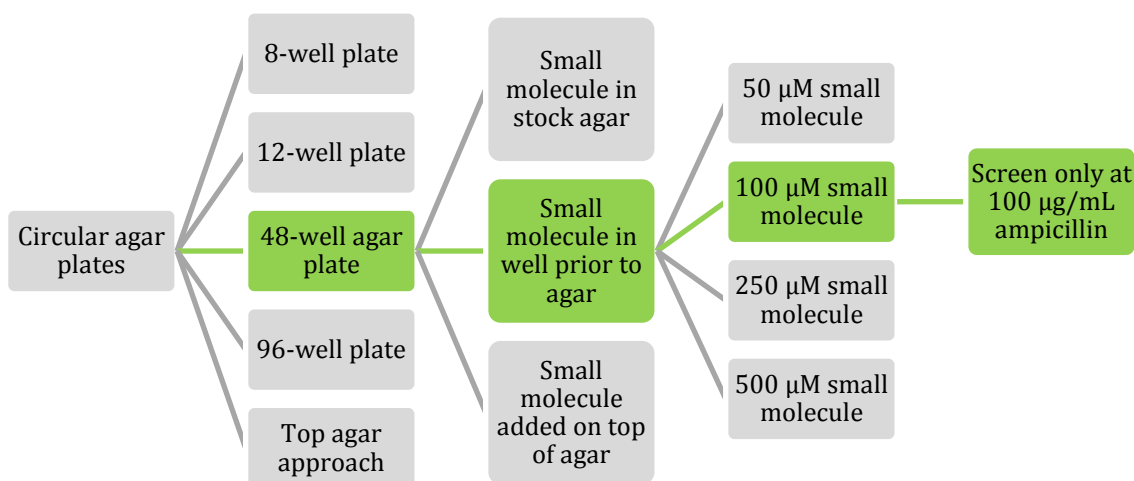
### 4.7.3 Assay miniaturisation

An imperative step in assay development is miniaturisation to enable high throughput screening. The advantage of miniaturisation is that it allows rapid, inexpensive scale-up to large scale assay formats. The original format of the *in vivo* assay used two 48-well plates per small molecule screened. One plate was utilised to assess any intrinsic off-target effects that the small molecule had on bacterial growth, while the other was used to assess the anti-aggregation properties of the small molecule. The proof-of-principle of the assay required the analysis of bacterial growth at eight concentrations of antibiotic. However, the volume of small molecule and number of plates required is a disadvantage for development into a high throughput screen. A solution to this issue was to reduce the number of concentrations of ampicillin at which the assay was performed. The number of false positives and false negatives observed at each concentration of ampicillin was quantified, and it was found that a concentration of 100  $\mu\text{g}/\text{mL}$  ampicillin only produced 8.1 % false-positive and 0 % false-negative results. By performing the screen at a single ampicillin concentration, the amount of small molecule required is significantly reduced (48  $\mu\text{g}$  of a 200 Da compound). This reduced requirement of small molecule was taken advantage of for screening a selection of small molecules whose effect on hIAPP aggregation was unknown. The *in vivo*  $\beta$ -lactamase tripartite fusion system, in conjunction with ESI-MS, successfully identified a novel inhibitor of hIAPP amyloid formation, using a fraction of the small molecule that would be required for other screens.

The assay in the current form is still in its infancy, however industrial development and automation would significantly reduce the workload and increase the screening capacity of the assay. If the assay were maintained in the current agar plate format, a number of stages could be automated. For example, several automated agar plate-pouring systems are available, thereby reducing production time significantly. Furthermore, an integrated colony picking robotic system could be utilised to automate plate scoring (via the

identification of wells containing colonies). Alternatively, the bacterial strain used could be genetically modified to contain the gene for a fluorescent protein (e.g. green fluorescent protein, GFP) in the chromosomal DNA<sup>481</sup>. This would change the readout of the assay from the presence or absence of colonies to quantitative fluorescence.

In summary, a robust and reproducible *in vivo* assay for small molecule inhibitors of protein aggregation has been successfully developed. A variety of conditions and experimental procedures were investigated to create an assay that was amenable to semi-high throughput format. **Figure 4.26** illustrates this timeline and highlights the successful conditions selected for the final screen. Small molecules that bind to proteins via a non-specific or colloidal mechanism are negative in the screen, thereby enabling the selection of specific inhibitors of protein aggregation. Combined with the ESI-IMS-MS screening technique developed in parallel to the *in vivo* assay, inhibitors of protein aggregation can be rapidly identified and their binding mode characterised using a fraction of the protein required for traditional screening techniques. Assay optimisation has produced a miniaturised assay which, through industrial collaborations, has the potential to be converted into a successful high throughput screen. The final objective in this body of work was to investigate the potential application of the *in vivo* system to biopharmaceutical aggregation and its inhibition. Chapter 5 describes the progress made toward this goal.



**Figure 4.26.** Summary of small molecule inhibitor of protein aggregation assay optimisation. The selected conditions are highlighted in green.

## 5 Applying the *in vivo* assay to biopharmaceutical aggregation

### 5.1 Objectives

The work described in Chapters 3 and 4 demonstrates that the *in vivo*  $\beta$ -lactamase tripartite fusion system can be used to identify amyloidogenic proteins and small molecule inhibitors of protein aggregation. The final phase of the project was to investigate the potential applications of the system developed to biopharmaceutical aggregation. The primary objective was to use the *in vivo* system to differentiate between aggregation-prone and non-aggregation-prone biopharmaceutically-relevant proteins. Once the system had been validated for this use, known excipients (chemical chaperones) were screened for their anti-aggregation properties. Together, these data validate the *in vivo* tripartite system as a platform for the identification of protein aggregation inhibitors, and for use against disease- and biopharmaceutical-related protein aggregation. The test proteins selected were antibody domains (the aggregation-prone Dp47d, and the non-aggregating HEL4 domains<sup>376, 482</sup>), as antibodies constitute one of the fastest growing areas of biopharmaceuticals.

### 5.2 Antibody biopharmaceuticals

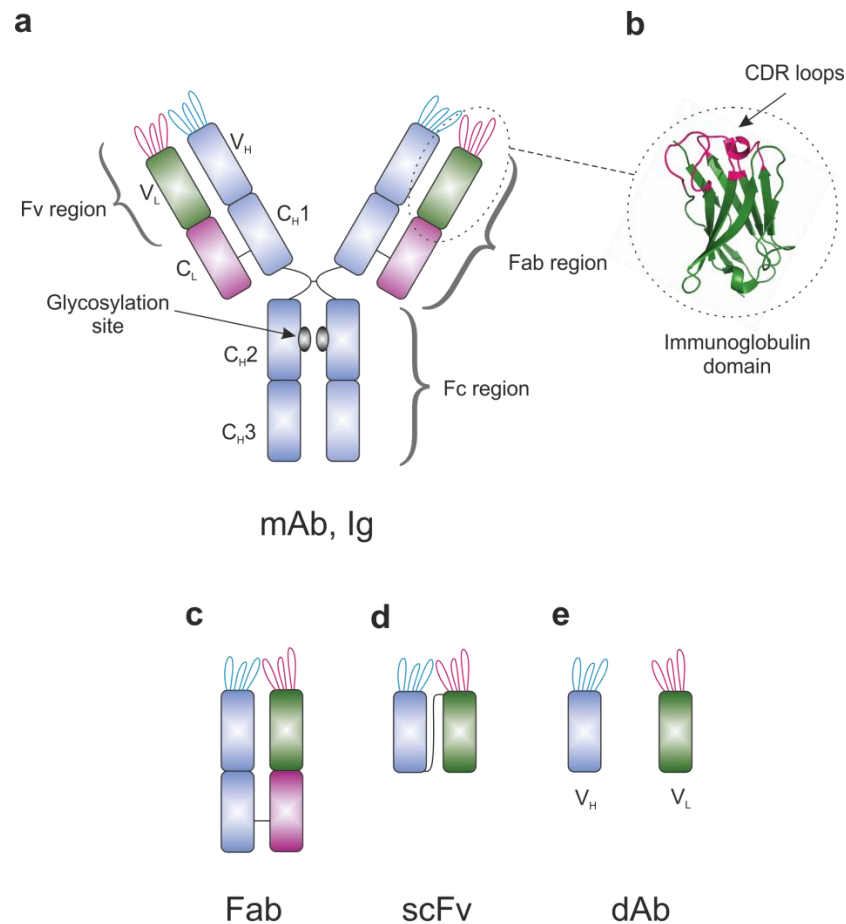
In 1975, César Milstein and Georges Köhler introduced monoclonal antibody (mAb) technology, in which mouse cells lines could be immortalised to secrete a single type of antibody<sup>483</sup>. This led to the first therapeutic antibody for human use (muromonab) being approved by the US Food and Drug Administration (FDA) for the treatment of acute transplant rejection in 1986<sup>320</sup>. In the nearly 40 years since the advent of mAbs, the market for mAbs and mAb fragments has grown exponentially<sup>484</sup>. mAbs are now ranked as the highest selling class of biopharmaceutical, with four out of the top ten selling biologics of 2012 - 2013 (in the United States of America) belonging to this category (see Section 1.8.1 and **Figure 1.18**).

### 5.2.1 Monoclonal antibodies

Whole mAbs, (also known as conventional immunoglobulins, Ig) are large (~ 150 kDa) glycosylated globular plasma proteins. A mAb monomer is a 'Y' shaped molecule, consisting of four polypeptide chains connected via disulfide bonds: two identical 'heavy chains' and two identical 'light chains' (**Figure 5.1a**). Each polypeptide chain is arranged into consecutive structural domains (immunoglobulin domains), consisting of two  $\beta$ -sheets arranged in an immunoglobulin fold (**Figure 5.1b**). A full length mAb consists of the constant Fc (crystallisable fragment) domain and an antigen binding domain, the latter comprising the Fv (variable fragment) and the Fab (antibody binding fragment)<sup>485</sup> (**Figure 5.1a**). The Fv of an antibody is the region responsible for the interaction with antigens, specifically via their complementary determining region (CDR) loops. There are three CDRs in a variable domain (**Figure 5.1b**), and as the antigen binding site comprises two variable domains, there are six CDRs that can come into contact with a specific antigen. The variation within the sequences of the CDRs is responsible for the vast diversity in antigen-recognition by mAbs.

The use of conventional mAbs in therapeutic and diagnostic applications was initially limited by a number of issues. Firstly, the necessity for the mAbs to be glycosylated restricted the constructs to production solely in mammalian expression systems. This vastly increased the cost (when compared with simpler bacterial expression systems) and limited yield<sup>486</sup>. Furthermore, there were significant immunogenicity issues when the mouse antibodies were administered to humans<sup>487</sup>. Advances in antibody design technology, combined with a fuller understanding of the action of therapeutic antibodies, has led to new and improved next-generations of mAb therapeutics<sup>320</sup>. Now, humanised smaller fragments of antibodies (**Figure 5.1c-e**) can be manufactured in high-volume bacterial or yeast cultures, thereby circumventing the previous limitations on this technology.





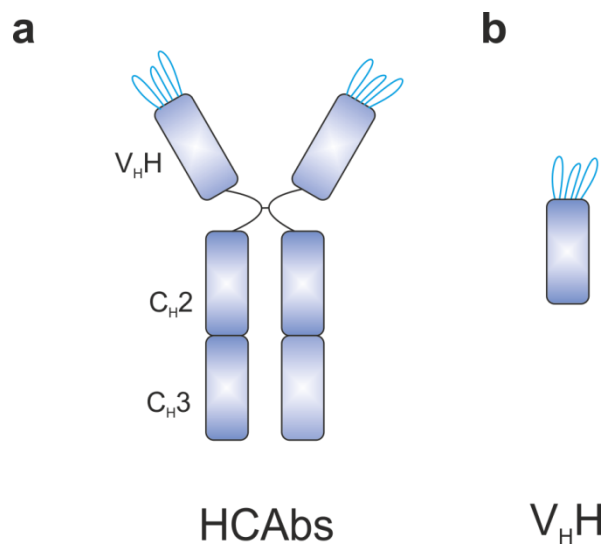
**Figure 5.1.** Structure of antibodies and their fragment-based therapeutics. **(a)** Structure of a monoclonal antibody (mAb, Ig). **(b)** Example of a single immunoglobulin domain. The complementary-determining region (CDR) loops are highlighted in pink. **(c-e)** Therapeutic fragments derived from mAbs, **(c)** antibody binding fragment (Fab), **(d)** single-chain variable fragment (scFv) and **(e)** single antibody domains (dAbs).

### 5.2.2 Antibody fragments

It was determined that a full-length glycosylated mAb was not necessary for antigen recognition, as paired N-terminal domains of the V<sub>H</sub> and V<sub>L</sub> chains (Fv region) were sufficient for antigen binding<sup>488</sup>. These fragments could be produced by joining the single V<sub>H</sub> and V<sub>L</sub> domains by a polypeptide linker<sup>489</sup> (single-chain Fv, scFv, ~ 27 kDa<sup>490</sup>) (**Figure 5.1d**), or simply as isolated Fab regions (~ 57 kDa<sup>490</sup>) (**Figure 5.1c**). Despite this progress, major technical hurdles remained in implementing this technology on an industrial scale<sup>486</sup>. For example, the scFv molecules were found to have a reduced affinity

when compared with the parent antibody<sup>491, 492</sup>. Furthermore, the linker joining the two domains was susceptible to aggregation and was easily degraded by proteolysis<sup>493</sup>. These limitations led researchers to explore the use of single domain antibodies (**Figure 5.1e**). Initial research mainly focused on using the  $V_H$  domain as, in contrast to the  $V_L$  domain,  $V_H$  domains often retain the antigen-specificity of the parental antibody, since their CDR3 is the major contributor to antigen binding<sup>486</sup>. In the late 1980s, the first functional murine  $V_H$  domain, for binding to lysozyme, was isolated<sup>494</sup>. However, it quickly became clear that removal of the  $V_L$  domain led to the exposure of a hydrophobic surface on the  $V_H$  domain (at the  $V_H$ - $V_L$  interface). These isolated  $V_H$  molecules were therefore 'sticky' and extremely difficult to produce in a soluble form. A further issue with single  $V_H$  domains was the significant loss of affinity when compared to the intact scFv domain (one to three orders of magnitude<sup>491</sup>).

In 1993 it was discovered that camels, dromedaries and llamas produce a unique type of antibody devoid of the entire light chain, and the first heavy chain constant region ( $C_{H1}$ ) (**Figure 5.2a**)<sup>495</sup>. These heavy chain antibodies (HCAbs) are able to bind their antigen via the single variable domain, referred to as  $V_{HH}$  to distinguish it from classic  $V_H$  domains<sup>496</sup> (**Figure 5.2b**). Impressively, the absence of the light chain does not limit the diversity of the epitopes<sup>497</sup>. Antibodies similar to camelid HCAbs have also been found in wobbegong (carpet sharks), nurse sharks and spotted ratfish<sup>498</sup>.



**Figure 5.2.** Structure of heavy chain antibodies and fragments derived from camelids. (a) The structure of a heavy chain antibody (HCAb) from a camelid. The light chain and the first heavy chain constant region ( $C_{H1}$ ) found in traditional antibodies are missing. (b) A single heavy chain variable fragment from a camelid antibody ( $V_{HH}$ ).

It was immediately apparent that V<sub>H</sub>H domains exhibit superior solubility and stability when compared to classical V<sub>H</sub> domains<sup>499</sup>. When compared to human V<sub>H</sub> domains, the V<sub>H</sub>H framework regions show a high sequence homology (> 80 %), and their three-dimensional structures can be superimposed<sup>486, 500</sup>. The most distinctive difference in V<sub>H</sub>H domains is amino acid substitutions (V37Y, G44E, L45R, W47G) at four positions that are conserved in conventional V<sub>H</sub> domains<sup>496, 501</sup>. This tetrad of conserved residues is involved in the formation of the hydrophobic interface with V<sub>L</sub> domains<sup>499</sup>. The substitutions increase the hydrophilicity of V<sub>H</sub>H domains by replacing large hydrophobic or non-polar residues with small hydrophobic or charged residues. Another key difference between V<sub>H</sub> and V<sub>H</sub>H domains is in the variable loop regions. The CDR3 loop of camelid V<sub>H</sub>H domains is on average much longer (17 residues) than human (12 residues) CDR3 loops<sup>496, 502</sup>. This extended loop masks the hydrophobic interface that would normally form with a V<sub>L</sub> domain<sup>503, 504</sup>. The large variable region of the V<sub>H</sub>H enlarges the surface of the antigen binding site, thereby compensating for the absence of the antigen-binding surface provided by the V<sub>L</sub> domain in variable fragments<sup>505</sup>. Indeed, the CDR3 contributes between 50 and 100 % of the antigen interacting surface<sup>486</sup>. Furthermore, the long CDR3 loop is often connected by a disulfide bond to the CDR1<sup>506</sup>. This bond restricts the conformational flexibility of the long CDR3 loop in the antigen-free form, so that the immobilisation of the loop upon antigen binding minimises the entropic penalty<sup>486</sup>. Some examples of the advantages of V<sub>H</sub>H domains over traditional antibody fragments are summarised in **Table 5.1**.

Advantage	Molecular Basis
High physicochemical stability	Efficient refolding due to increased hydrophilicity and single-domain nature <sup>507</sup>
High solubility	Increased hydrophilicity <sup>508</sup>
Recognition of hidden antigenic sites	Small size and extended flexible CDR3 <sup>509</sup>
Rapid tissue penetration, fast clearance	Small size <sup>510</sup>
Well expressed	Efficient folding due to increased hydrophilicity and single-domain nature <sup>485</sup>

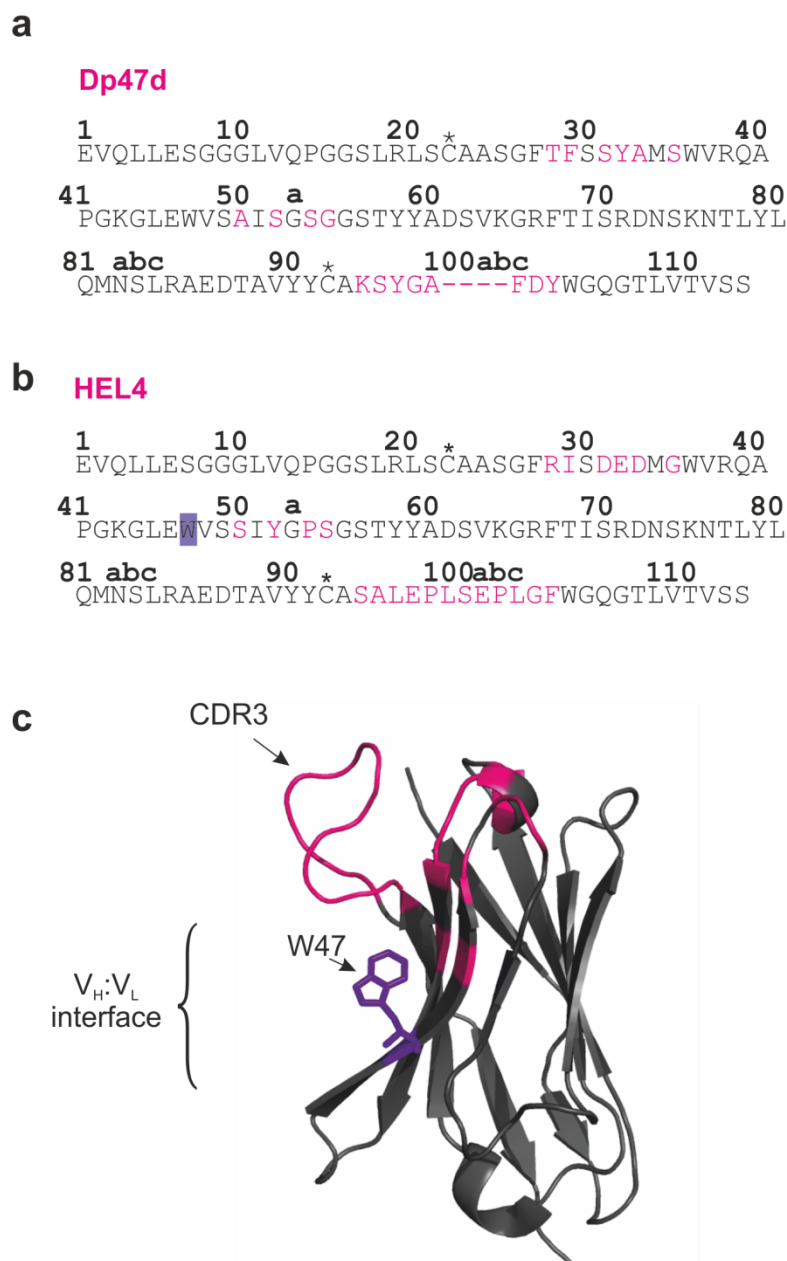
**Table 5.1.** Advantages of single-domain antibody fragments compared to conventional antibody fragments. Table modified from Harmsen, 2007<sup>499</sup>.

### 5.2.3 Reducing the aggregation propensity of dAbs

The discovery of highly stable and soluble V<sub>H</sub>H domains led to the idea of ‘camelisation’, in which the hydrophobic residues on the traditional V<sub>H</sub>:V<sub>L</sub> interface of dAbs were replaced with the hydrophilic amino acids frequently found in V<sub>H</sub>H domains<sup>490, 511</sup>. Similarly, extension and/or sequence variation within the CDRs has also been shown to increase dAb solubility and reduce aggregation-propensity<sup>512-514</sup>. These alterations could be either rationally designed<sup>515</sup>, or discovered through phage selection methods. For example, a phage library of human V<sub>H</sub> dAbs (10<sup>10</sup> clones) was selected for binding to hen-egg lysozyme (HEL)<sup>376</sup>. From a total of 21 HEL-specific clones, one was selected (HEL4) that was significantly less aggregation-prone *in vitro* (determined by thermal unfolding and solubility experiments) compared with the highly aggregation prone Dp47d, a human V<sub>H</sub> dAb encoded by the same germline gene (**Figure 5.3a, b**). X-ray crystallography revealed that although the hydrophobic framework residues of the V<sub>H</sub>:V<sub>L</sub> interface were maintained, the side-chain of framework residue W47 is flipped into a cavity, thereby increasing the hydrophilicity of the V<sub>H</sub>:V<sub>L</sub> interface domain<sup>376</sup> (**Figure 5.3c**).

A significant technological advancement was made when a powerful technique, linking dAb thermal stability and phage display, was developed<sup>482</sup>. Multiple copies of individual dAbs were displayed on the surface of phage, which were subsequently heated to 80 °C to induce dAb unfolding and aggregation. After cooling the phage, any thermally unstable dAbs would not refold correctly, leaving any dAbs that could reversibly unfold to be selected and characterised. By simultaneously combining selection for thermal stability and antigen specificity, thermal-based aggregation-prone specific dAbs could be rapidly identified. However, a key disadvantage of this technique was that it involved full denaturation of dAbs. This therefore restricted the selection only to mutations that reduced the tendency of the unfolded state to aggregate. Indeed, it was found that the thermodynamic stability of the selected dAbs was no better than a typical aggregation-prone human dAb; although the selected dAbs could avoid irreversible aggregation at 80 °C, their average free energy of folding ( $\Delta G^\circ$ ) at 25 °C was only -20 kJ/mol, whereas that of Dp47d and other human dAbs was -35 kJ/mol<sup>514, 516</sup>. As the aggregation-resistance of these domains was therefore due to the properties of the denatured state, this selection technique gave little advantage for dAbs of greater thermodynamic stability. To counter this issue, the same group modulated the technique to only lead to partial unfolding of most of the dAb members on the phage surface<sup>517</sup>. By keeping the temperature below the melting temperature of the majority of dAbs, and

inducing partial acid unfolding (at pH 3.2), the selection process favoured domains that are thermodynamically stable ( $\Delta G^\circ = 27\text{-}34$  kJ/mol), as well having aggregation-resistant denatured states ( $T_m = 58\text{-}68$  °C)<sup>517</sup>.

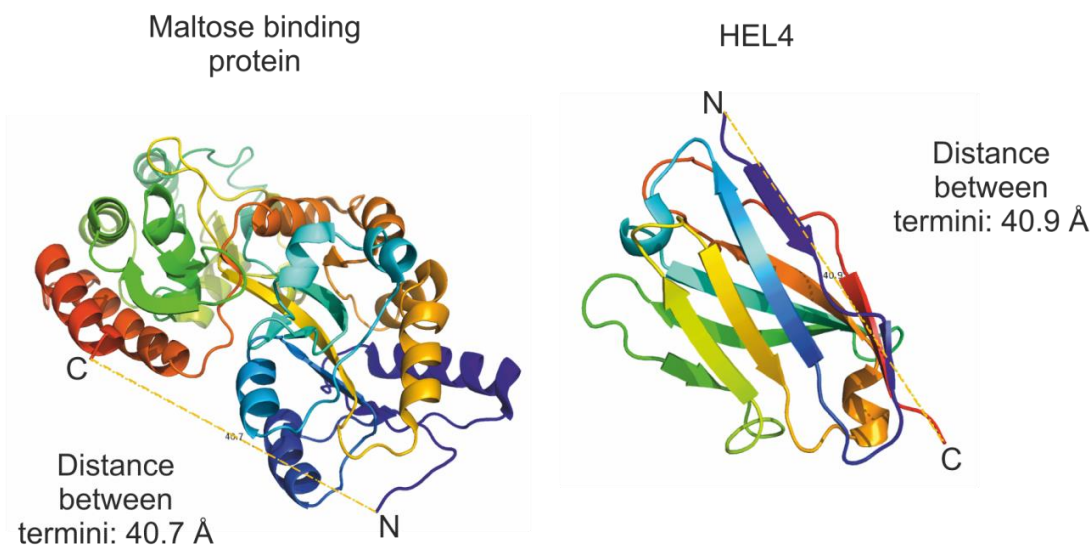


**Figure 5.3.** Sequence and structural comparisons of the  $V_H$  dAbs Dp47d and HEL4. Sequence of (a) Dp47d and (b) HEL4. Residues numbered according to Kabat *et al.*<sup>518</sup> (standardised numbering of residues in an antibody). Residues that differ between the two dAbs are coloured pink. Residue W47 that is flipped into a hydrophobic cavity of HEL4 is highlighted in purple. Asterisk marks residues involved in the disulfide bond. (c) Crystal structure of HEL4 coloured according to sequences in (b). The key solubility-determining factors (CDR3 and W47) are labelled. Created using Protein DataBank (PDB) entry 1OHQ<sup>376</sup> and PyMOL Molecular Graphics System (Schrödinger, LLC).

### 5.3 Using the $\beta$ -lactamase tripartite fusion system to identify dAb aggregation propensity

Previous studies using the  $\beta$ -lactamase tripartite fusion system demonstrated its ability to identify stability-enhancing mutations and chemical additives that increase stability *in vivo*<sup>308, 313</sup>. As the data presented in Chapters 3 and 4 of this thesis established that the  $\beta$ -lactamase tripartite fusion system can determine the extent of aggregation of 37- to 42-residue peptides, it was hypothesised that the *in vivo*  $\beta$ -lactamase tripartite fusion system could also be used to differentiate soluble, stable dAbs from their aggregation-prone counterparts. By linking aggregation propensity to enzyme activity, a direct readout of dAb aggregation could be obtained without resorting to denaturing conditions such as increased temperature<sup>482</sup> or low pH<sup>517</sup>.

HEL4 and Dp47d were chosen as test proteins for the *in vivo* screen. A comparison of HEL4 and Dp47d sequences revealed 22 residue differences (HEL4 is 120 and Dp47d is 116 residues in length) (**Figure 5.3a, b**). Analysis of the X-ray crystal structure of HEL4<sup>376</sup> shows the N- and C-termini to be 40.9 Å apart, a similar length to the distance between the termini of the maltose binding protein (MBP, 40.7 Å)<sup>519</sup> (**Figure 5.4**). MBP was utilised as a test protein in the original study using the  $\beta$ -lactamase tripartite system to evolve protein stability *in vivo*<sup>308</sup>, suggesting that incorporation of the dAbs into the  $\beta$ -lactamase GS-rich linker would be possible.



**Figure 5.4.** N- and C-termini distance between maltose-binding protein (MBP) and HEL4. The distance between the termini of (a) MBP is 40.7 Å and (b) HEL4 is 40.9 Å. Created using Protein DataBank (PDB) entries 1OMP (MBP)<sup>519</sup> and 1OHQ (HEL4)<sup>376</sup>, and PyMOL Molecular Graphics System (Schrödinger, LLC).

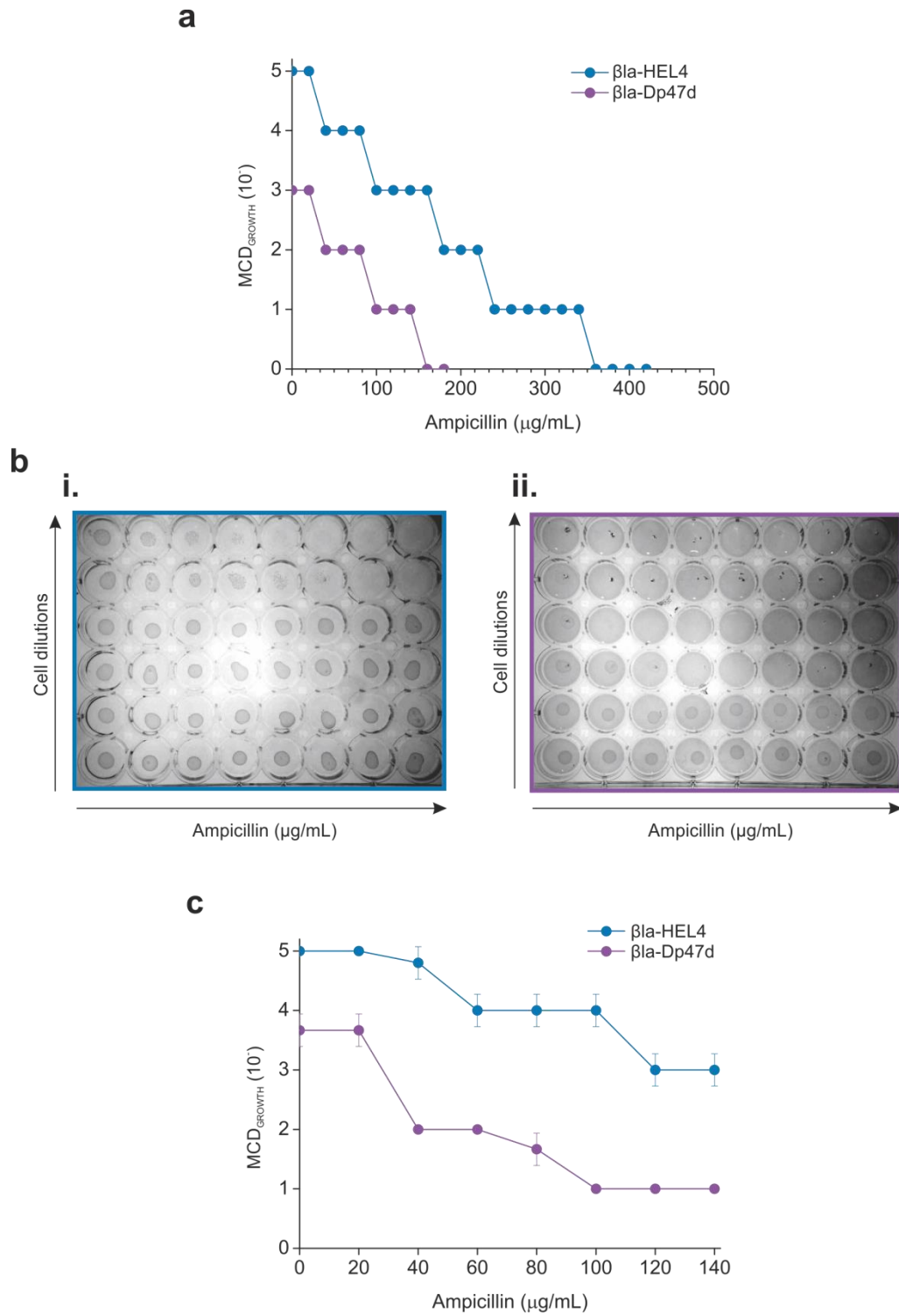
### 5.3.1 Cloning dAbs into $\beta$ -lactamase linker

Using the protein sequences stated in Jaspers *et al.*, 2004<sup>376</sup>, plasmids containing the genes encoding HEL4 and Dp47d (pEX-HEL4 and pEX-Dp47d, respectively) were synthesised by Eurofins MWG Operon (Ebersburg, Germany) (**Table 2.11**). Plasmid maps are shown in **Appendices 7.3.4** and **7.3.5**. HEL4 and Dp47d were cloned into the 28-residue glycine/serine-rich linker between residues 196 and 197 of TEM-1  $\beta$ -lactamase<sup>308</sup>, using primers designed to include an *Xho*I restriction site 5' and a *Bam*HI restriction site 3' to the gene (**Table 2.8**). The plasmid containing  $\beta$ la-linker<sub>SHORT</sub> was digested with the same enzymes, and the PCR product encoding the test protein was ligated into it. The ligation products were transformed into *E. coli* JM109 cells (Section 2.2.7) and the cells grown on agar plates containing 10  $\mu$ g/mL tetracycline. Successful ligation was identified by the resistance to tetracycline obtained from the  $\beta$ -lactamase vector. DNA was purified from a selection of colonies and was sent for sequencing to confirm that they contained the correct sequences (Section 2.2.9). Primers for the sequencing reactions were designed to bind upstream and downstream of the GS-linker of  $\beta$ -lactamase (**Table 2.8**). The newly synthesised plasmids were named  $\beta$ la-28-HEL4 and  $\beta$ la-28-Dp47d; full DNA and protein sequences are provided in **Appendices 7.1.8** and **7.1.9**.

### 5.3.2 Antibiotic resistance correlates with dAb aggregation-propensity

The minimal inhibitory concentration of antibiotic (MIC) was determined for bacteria expressing the  $\beta$ la-HEL4 or the  $\beta$ la-Dp47d constructs, over a large range of ampicillin concentrations in order to determine the optimal antibiotic range to use for subsequent assays. Briefly, cultures of BL21 (DE3) cells expressing each construct were grown until the  $OD_{600} = 0.6$  (37 °C, 200 rpm). Protein expression was then induced by the addition of 0.02 % (w/v) arabinose and bacteria were allowed to grow for a further 1 h (37 °C, 200 rpm), before sequential cell dilutions were pipetted onto 48-well agar plates containing increasing concentrations of the  $\beta$ -lactam antibiotic ampicillin (0 – 500  $\mu$ g/mL; 20  $\mu$ g/mL increments). Plates were incubated at 37 °C for 18 h, after which they were scored for the maximal cell dilution at which cells could grow at each antibiotic concentration (**Figure 5.5a**). As bacteria expressing the  $\beta$ la-Dp47d construct were unable to grow at ampicillin concentrations greater than 180  $\mu$ g/mL, an ampicillin concentration range of 0 – 140  $\mu$ g/mL ampicillin (20  $\mu$ g/mL increments) was used for all subsequent MIC assays, to allow one assay/48-well plate (**Figure 5.5b**). The *in vivo* assay consistently differentiated between the aggregation-prone  $\beta$ la-Dp47d and the more soluble  $\beta$ la-HEL4 (**Figure 5.5c**). These findings confirm that the  $\beta$ -lactamase tripartite fusion system is amenable to a variety of test proteins with different structures and sizes. Furthermore, it is applicable for disulfide bond-containing immunoglobulin domains, something that has not been shown before, either in this study or the previous studies<sup>308, 313</sup>.





**Figure 5.5.** *In vivo* antibiotic resistance correlates with dAb aggregation propensity. **(a)** Maximal cell dilution allowing growth ( $MCD_{GROWTH}$ ) over a wide range of ampicillin concentrations for  $\beta la-HEL4$  (●) and  $\beta la-Dp47d$  (●) tripartite fusion constructs. **(b)** Example agar plates of  $MCD_{GROWTH}$  over the restricted ampicillin range of 0-140  $\mu g/mL$  ( $\beta la-HEL4$  (■) and  $\beta la-Dp47d$  (■)). **(c)**  $MCD_{GROWTH}$  over the restricted ampicillin range of 0-140  $\mu g/mL$ . Error bars represent the standard error from a minimum of 3 replicates.

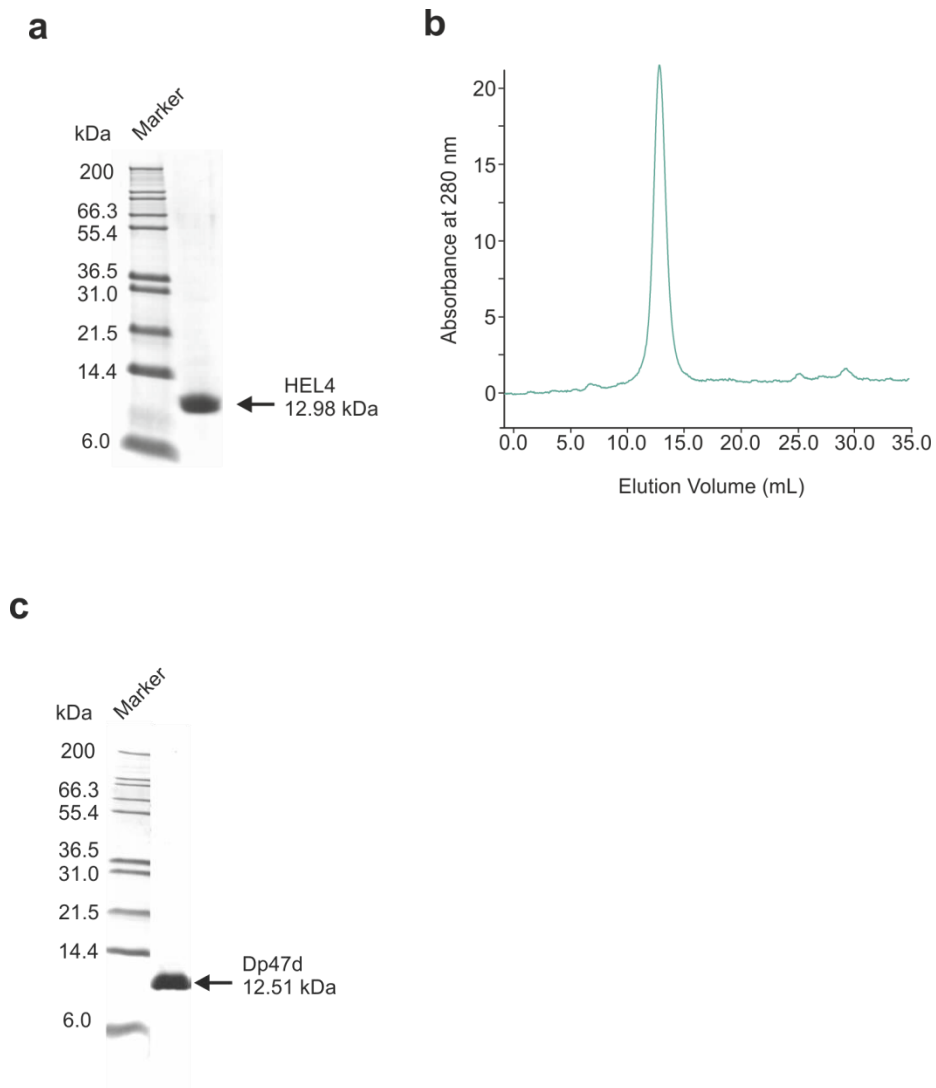
## 5.4 *In vivo* screening of excipients for anti-aggregation properties

The stability and solubility of proteins are routinely enhanced during biopharmaceutical production and storage by the addition of excipients (pharmacologically inactive chemical chaperones)<sup>351, 354, 516</sup>. Assessing the suitability of various excipients often requires the *in vitro* purification and incubation of the protein of interest (POI) in various solution conditions under several stresses<sup>303, 520-522</sup>. Furthermore, if the POI is inherently unstable and aggregation-prone, initial purification can be costly and laborious to carry out. Whether the *in vivo*  $\beta$ -lactamase tripartite fusion system could be used to screening excipients for the ability to stabilise and/or prevent the aggregation of the inserted test protein was next tested, therefore, with a view to providing valuable information to aid the selection of co-solvents that benefit purification.

### 5.4.1 Dp47d purification in the absence of co-solvents

To demonstrate the aggregation-propensity of Dp47d in the absence of any co-solvent (as compared to its non-aggregating counterpart HEL4), purification of both domains was attempted. The plasmids containing HEL4 or Dp47d were digested with the restriction enzymes *XhoI* and *NdeI* and the genes ligated into a pET23a vector digested with the same enzymes (Section 2.2). Plasmids, named pET-HEL4 and pET-Dp47d are detailed in **Table 2.11** and plasmid maps are shown in **Appendices 7.3.7** and **7.3.8**. HEL4 and Dp47d were over-expressed using a method adapted from Jespers *et al.*, 2004<sup>376</sup> (Section 2.3.2.1) and the inclusion bodies isolated as described in Section 2.3.2.2.

As expected, HEL4 purification was a relatively simple process; the inclusion bodies were solubilised and dialysed into 25 mM Tris-HCl, pH 8.0 (see Section 2.3.2.3), before purification by anion exchange chromatography followed by size exclusion chromatography (full details in Section 2.3.2.3). Protein purity was confirmed by SDS PAGE (**Figure 5.6a**) and identity was confirmed by mass spectrometry (observed mass: 12979.2 Da, expected mass: 12979.5 Da). The monomeric nature of purified HEL4 was confirmed by analytical gel filtration (**Figure 5.6b**). The yield was 56.3 mg per litre of culture and the protein was lyophilised and stored at -80 °C.



**Figure 5.6.** Purification of the dAbs HEL4 and Dp47d. **(a)** SDS-PAGE gel of purified HEL4. **(b)** HEL4 elution profile from the Superdex™ 75 HR 10/30 analytical gel filtration column. Column calibration confirms peak position corresponds to a ~ 13 kDa protein. **(c)** SDS-PAGE gel of Dp47d purified in a denatured state. The GuHCl in the protein sample was removed by trichloroacetic acid precipitation prior to SDS-PAGE analysis. The size in kDa of the protein marker is indicated.

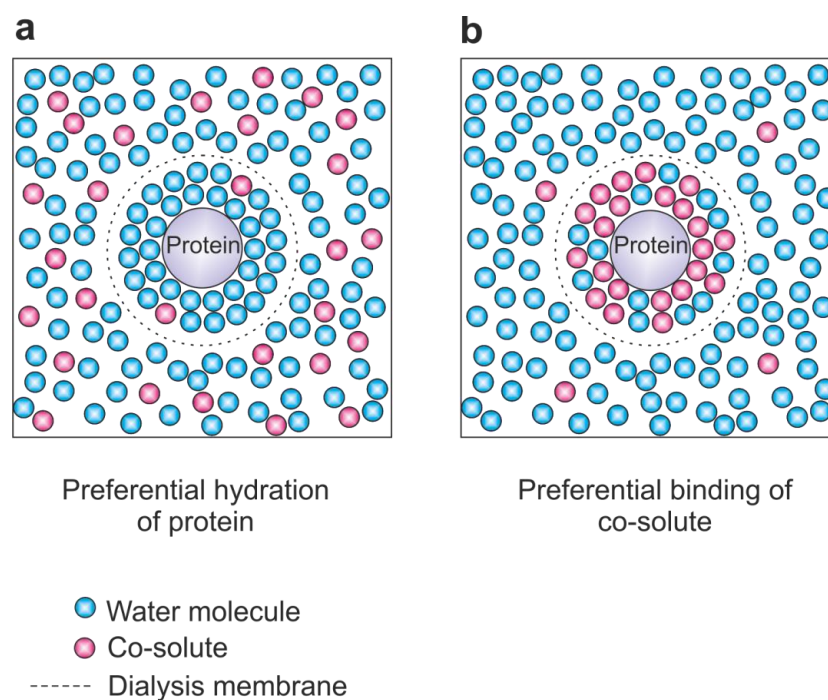
Consistent with previous publications<sup>376, 517</sup>, the Dp47d V<sub>H</sub> domain was significantly more difficult to purify in soluble form. Using the same protocol as used for HEL4 purification (Section 2.3.2.1) resulted in complete aggregation of Dp47d on the size-exclusion gel filtration column, and recovery of no soluble protein. Purifying Dp47d in a denatured form, using 4 M guanidine hydrochloride (GuHCl) before refolding into buffer, was also tested as a purification route. Briefly, the Dp47d inclusion bodies from 10 L of culture were solubilised in 10 mL of 4 M GuHCl, 50 mM glycine, pH 9.5, overnight at 4 °C. Any precipitate was removed by filtering the solution through a 0.2 µm syringe filter. 3 mL/run

was then loaded onto a Superdex™ 75 GL 10/300 gel filtration column equilibrated with two column volumes of 4 M GuHCl, 50 mM glycine, pH 9.5. The protein was eluted from the column in 20 mM Tris-HCl, pH 7.5 at a flow rate of 1 mL/min. 1.5 mL fractions were collected and protein elution monitored by absorbance at 280 nm (Section 2.3.2.4). Trichloroacetic acid (TCA) precipitation (Section 2.3.2.5) was used to remove GuHCl from an aliquot of the protein sample prior to purity analysis by SDS-PAGE (**Figure 5.6c**). The sample was confirmed to be pure, so subsequent refolding of the denatured Dp47d was attempted. Dialysis (3,000 MWCO membrane) against 50 mM glycine, pH 9.5 was attempted, however this resulted in rapid precipitation of the protein. Refolding by rapid dilution (dropwise addition of solubilised Dp47d into 50 mM glycine, pH 9.5 with rapid stirring) also resulted in complete precipitation of Dp47d.

The data highlight the difficulties faced when dealing with an aggregation-prone protein *in vitro*. As such, a number of excipients were next screened to assess their ability to prevent  $\beta$ la-Dp47d aggregation *in vivo* (Sections 5.4.4 and 5.4.6), to allow selection of those that aid in the purification of soluble Dp47d *in vitro* (Section 5.4.7).

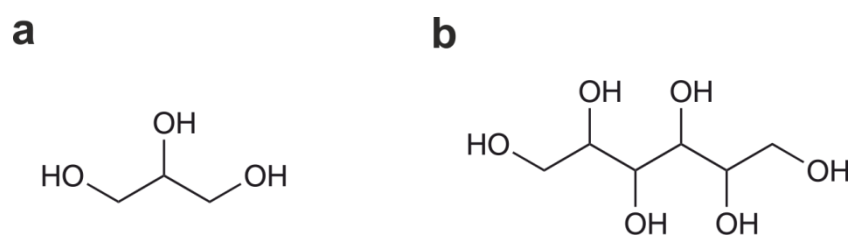
## 5.4.2 Selection of protein stabilisers

Osmolytes are ancient members of *in vivo* stress responses, used by plants and microorganisms to protect proteins in response to heat, cold, water or salt stress<sup>516, 520</sup>. These low molecular weight, uncharged or zwitterionic molecules can be present at high concentrations (> 1 M) *in vivo* in order to raise intracellular osmotic pressure<sup>523</sup>, yet they do not interfere with protein function<sup>516</sup>. This compatibility with proteins is attributable to the fact that osmolytes do not physically interact with proteins<sup>524-526</sup>. Rather, they are excluded from the protein surface in preference for water (**Figure 5.7a**). This 'excluded-volume effect' leads to the preferential hydration of the protein surface, and as the unfolded or partially unfolded state has a greater solvent-exposed surface area than that of the folded state, more osmolyte molecules are excluded in the unfolded state<sup>520</sup>. As this would lead to greater hydration of the protein by the water molecules, and therefore more ordered water molecules, it is entropically more favourable for the protein to be in its folded state. By limiting the access to partially or fully unfolded states, protein aggregation, attributed to exposure of hydrophobic regions normally buried within the protein interior, is minimised<sup>516</sup>.



**Figure 5.7.** Schematic representation of preferential hydration of protein versus preferential binding of co-solute to protein. **(a)** Preferential hydration of the protein. More water molecules are in the protein solution than in the bulk phase (outside the dialysis membrane). **(b)** Preferential binding of co-solute. The co-solute is present in excess in the protein solution relative to its concentration in the bulk phase. Figure redrawn from Arakawa *et al.*, 2006<sup>516</sup>.

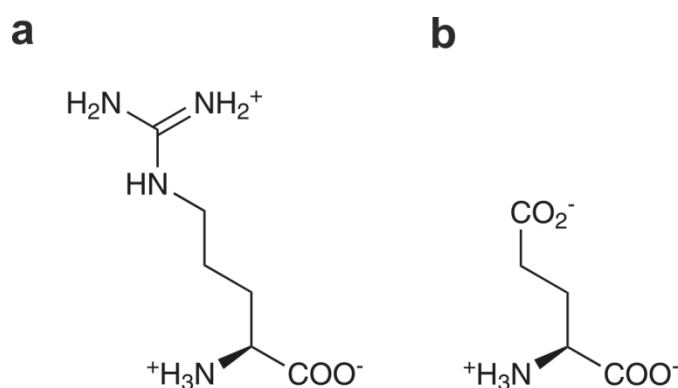
Polyols are some of the most widely used co-solvents used to stabilise proteins and prevent protein aggregation during purification and refolding of proteins via the excluded volume effect<sup>527-531</sup>. Two of the best studied are glycerol<sup>352, 527</sup> and sorbitol<sup>531-534</sup> (**Figure 5.8**). These molecules were selected for analysis of their ability to rescue the growth of bacteria expressing the aggregation-prone  $\beta$ la-Dp47d construct (Section 5.4.4).



**Figure 5.8.** Structure of the polyols (a) glycerol and (b) sorbitol.

### 5.4.3 Selection of aggregation suppressors

Increasing the thermodynamic stability of the native state can prevent protein unfolding which can lead to aggregation (Section 5.4.1). However, aggregation can still occur, even when conditions thermodynamically favour the native conformation, if electrostatic interactions between protein molecules are strong<sup>535</sup>. Co-solutes that prevent these interactions function by preferentially binding to the protein (unlike co-solutes that lead to preferential hydration) (**Figure 5.7b**), thereby masking areas of the protein prone to inter-molecular interactions. A combination of free L-arginine (Arg) and L-glutamate (Glu) (**Figure 5.9**) has been demonstrated to prevent the aggregation of six unrelated proteins by a combination of neutralising opposite charges (which would otherwise lead to intermolecular association) and masking adjoining hydrophobic areas with their aliphatic tails<sup>536</sup>. Indeed, one study has demonstrated that L-arginine strongly binds to tryptophan (Trp) residues during refolding assays and decreases the extent of aggregation<sup>537</sup>. As the improved solubility of HEL4 (as compared to Dp47d) is attributed to the burial of a Trp residue (W47) in a hydrophobic pocket in HEL4, this suggests addition of L-Arg to could prevent  $\beta$ la-Dp47d aggregation *in vivo*. Due to the reported success of a combination of L-arginine and L-glutamate in preventing protein aggregation<sup>536</sup>, these were selected to investigate the ability to screen for the anti-aggregation ability of these free amino acids in the tripartite *in vivo* system.

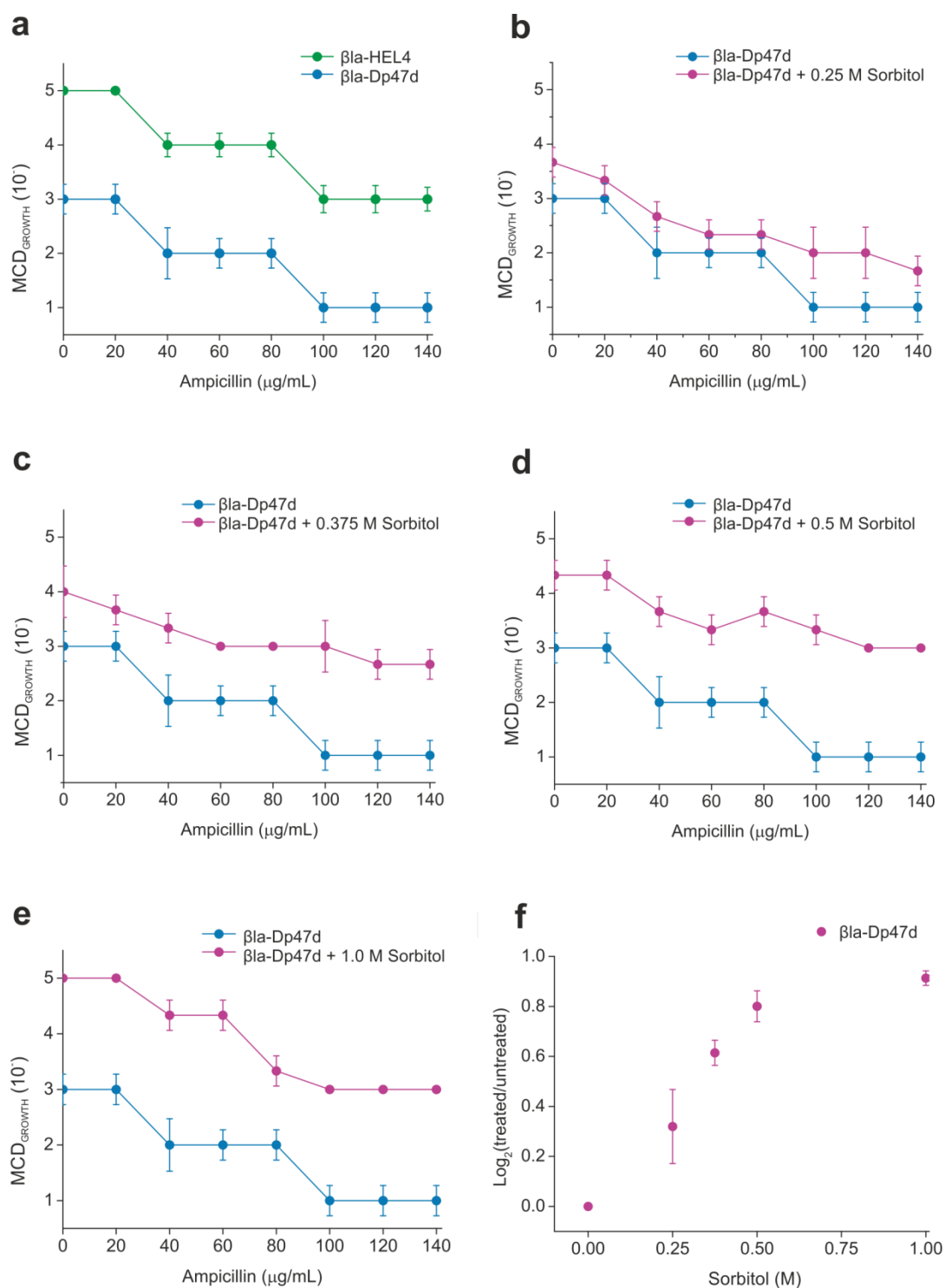


**Figure 5.9.** Structure of (a) L-arginine and (b) L-glutamate at neutral pH.

#### 5.4.4 Glycerol and sorbitol rescue bacterial growth in the presence of ampicillin

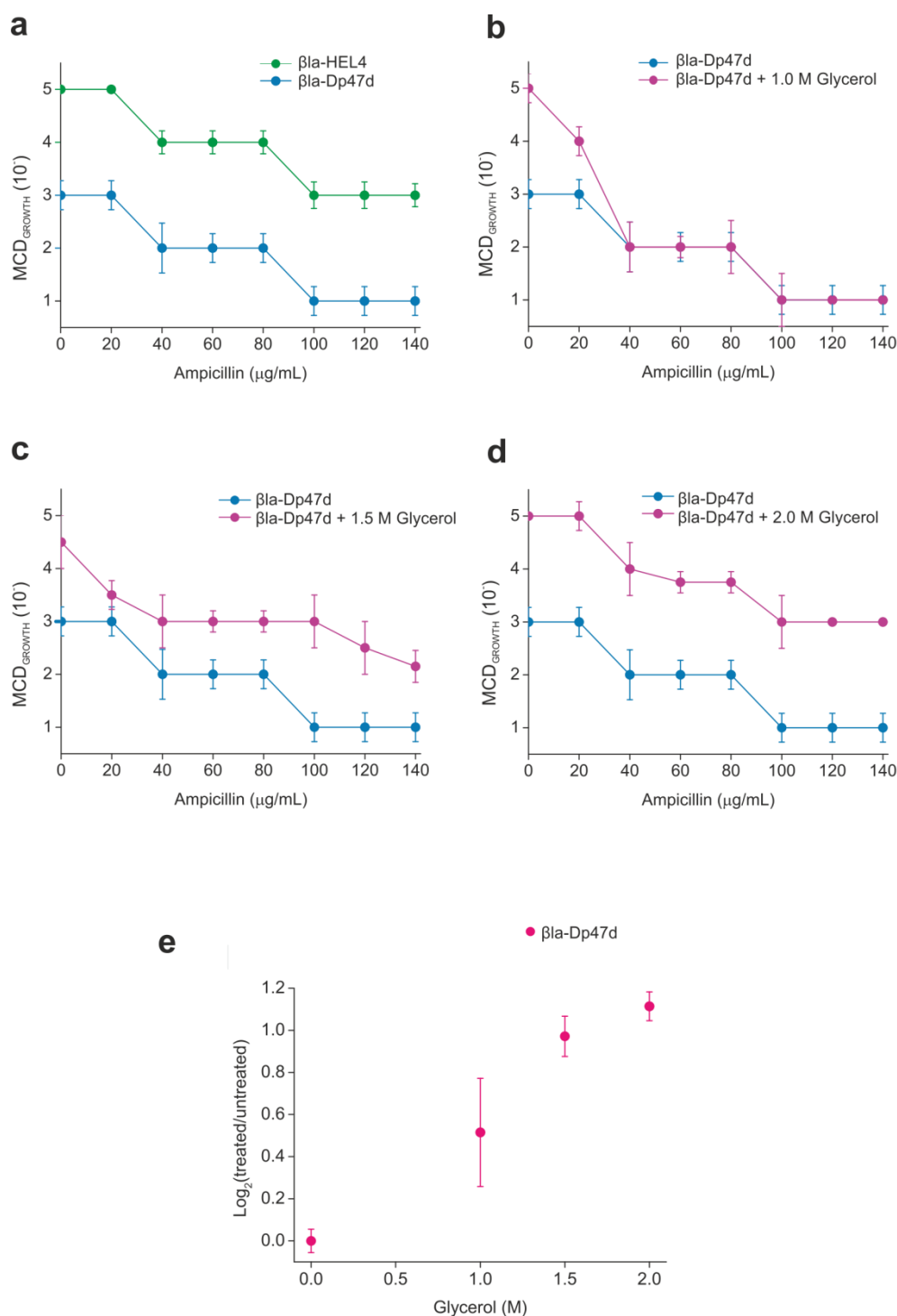
48-well agar plates were created containing 0, 0.25, 0.375, 0.5 or 1 M sorbitol, or 0, 1, 1.5 or 2 M glycerol as described in Section 2.4.5.1. A similar protocol as used for screening of small molecule inhibitors of protein aggregation was performed. Briefly, cultures of BL21 (DE3) cells expressing each construct were grown until the  $OD_{600} = 0.6$  (37 °C, 200 rpm). Protein expression was then induced by the addition of 0.02 % (w/v) arabinose. 200  $\mu$ L were removed and added to 1.5 mL Eppendorf tubes containing 200  $\mu$ L of sorbitol or glycerol stock (warmed to 37 °C) to give the final concentration of excipient required (see Section 2.4.5.1). These cultures were incubated for a further 1 h (37 °C, 200 rpm), before sequential cell dilutions were pipetted onto the prepared 48-well agar plates containing excipients and increasing concentrations of ampicillin (0 – 140  $\mu$ g/mL; 20  $\mu$ g/mL increments). Plates were incubated at 37 °C for 18 h, after which they were scored for the maximal cell dilution at which cells could grow at each antibiotic concentration.

In all experiments, any intrinsic effect that the excipients had on bacterial growth was accounted for as described previously (Section 4.2.4). The addition of sorbitol and glycerol to the growth medium was found to rescue bacterial growth in a quantitative manner (**Figure 5.10** and **Figure 5.11** respectively). In fact, the presence of 1 M sorbitol in the agar enabled  $\beta$ la-Dp47d-expressing bacteria to grow at the same concentration of ampicillin as bacteria expressing the non-aggregating construct  $\beta$ la-HEL4 (compare **Figure 5.10a** and **Figure 5.10e**). Conversely, 2 M of glycerol was required for a similar effect to be observed with this excipient (**Figure 5.11d**). The data confirm that osmolyte excipients can be assayed *in vivo* for their anti-aggregation properties. The data suggest that sorbitol is superior to glycerol in its ability to prevent Dp47d aggregation *in vivo*. Consequently, the ability of sorbitol and glycerol to prevent Dp47d aggregation *in vitro* is assessed in Section 5.4.7.



**Figure 5.10.** Bacterial growth in presence of sorbitol. **(a)** MIC assay in the absence of sorbitol.  $\beta$ la-HEL4 (●) is included as an example of a non-aggregating construct. The maximal cell dilution allowing growth ( $MCD_{GROWTH}$ ) was scored at each ampicillin concentration in the presence of **(b)** 0.25 M sorbitol, **(c)** 0.375 M sorbitol, **(d)** 0.5 M sorbitol and **(e)** 1 M sorbitol. In each case, any intrinsic effect that sorbitol had on bacterial growth was accounted for as described in Section 4.2.4. **(f)** Bacterial growth in the presence (treated) versus the absence (untreated) of a range of sorbitol concentrations (calculated using **Equation 4.1**). Error bars represent the standard error from a minimum of 3 replicates. Experiments were performed with undergraduate student Oliver Holman.

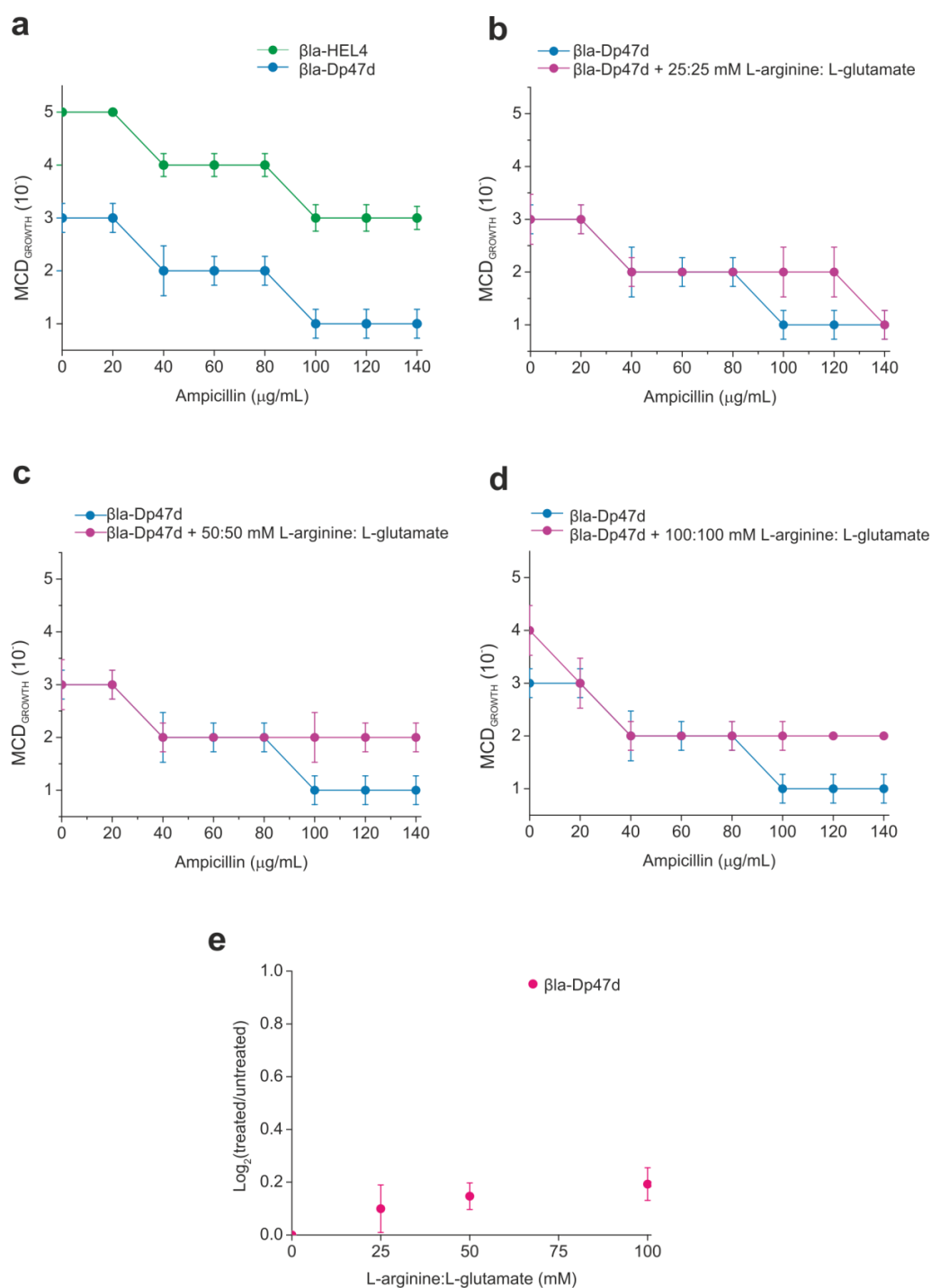




**Figure 5.11.** Bacterial growth in presence of glycerol. **(a)** MIC assay in the absence of glycerol.  $\beta$ la-HEL4 (●) is included as an example of a non-aggregating construct. The maximal cell dilution allowing growth ( $MCD_{GROWTH}$ ) was scored at each ampicillin concentration in the presence of **(b)** 1.0 M glycerol, **(c)** 1.5 M glycerol and **(d)** 2.0 M glycerol. In each case, any intrinsic effect that glycerol had on bacterial growth was accounted for as described in Section 4.2.4. **(e)** Bacterial growth in the presence (treated) versus the absence (untreated) of a range of glycerol concentrations (calculated using **Equation 4.1**). Error bars represent the standard error from a minimum of 3 replicates. Experiments were performed with undergraduate student Oliver Holman.

### 5.4.5 L-Arginine and L-glutamate elicit minimal rescue of bacterial growth

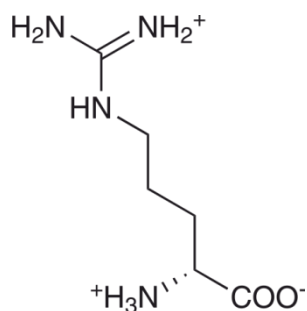
Golovanov *et al.*<sup>536</sup> demonstrated that an equimolar (50 mM) combination of L-arginine (L-Arg) and L-glutamate (L-Glu) could increase the solubility limit of six unrelated proteins by up to a factor of 8.7 (from 0.1 to 1.3 mM). Consequently, equimolar concentrations of L-Arg and L-Glu were utilised to investigate whether these free amino acids could also improve the solubility of Dp47d in the  $\beta$ -lactamase tripartite fusion system. Agar plates were created containing 25:25 mM, 50:50 mM or 100:100 mM L-Arg:L-Glu (Section 2.4.5.2). The pH of the agar was corrected to pH 7 prior to pouring the plates (see Section 2.4.5.2). The same protocol utilised to investigate the effects of the excipients glycerol and sorbitol was applied to assay the effects of these amino acids, however the Eppendorf tubes contained 200  $\mu$ L of L-Arg or L-Glu stock (prepared with LB medium, warmed to 37 °C) to give the final concentration of excipient required. After 18 h incubation at 37 °C the plates were scored for the maximal cell dilution at which growth could occur (**Figure 5.12**). The ability of these excipients to rescue growth of the bacteria expressing the aggregating construct  $\beta$ la-Dp47d was limited, especially in comparison to the significant growth improvement in the presence of the excipients sorbitol and glycerol (**Figure 5.10** and **Figure 5.11**). As L-amino acids are key substituents and metabolites for bacterial growth, it was hypothesised that the addition of high concentrations of these free amino acids may induce toxicity by interfering with normal cellular metabolic pathways. It was hypothesised that D-amino acids may circumvent these issues as they are minimally utilised by *E. coli* (as compared to L-amino acids).



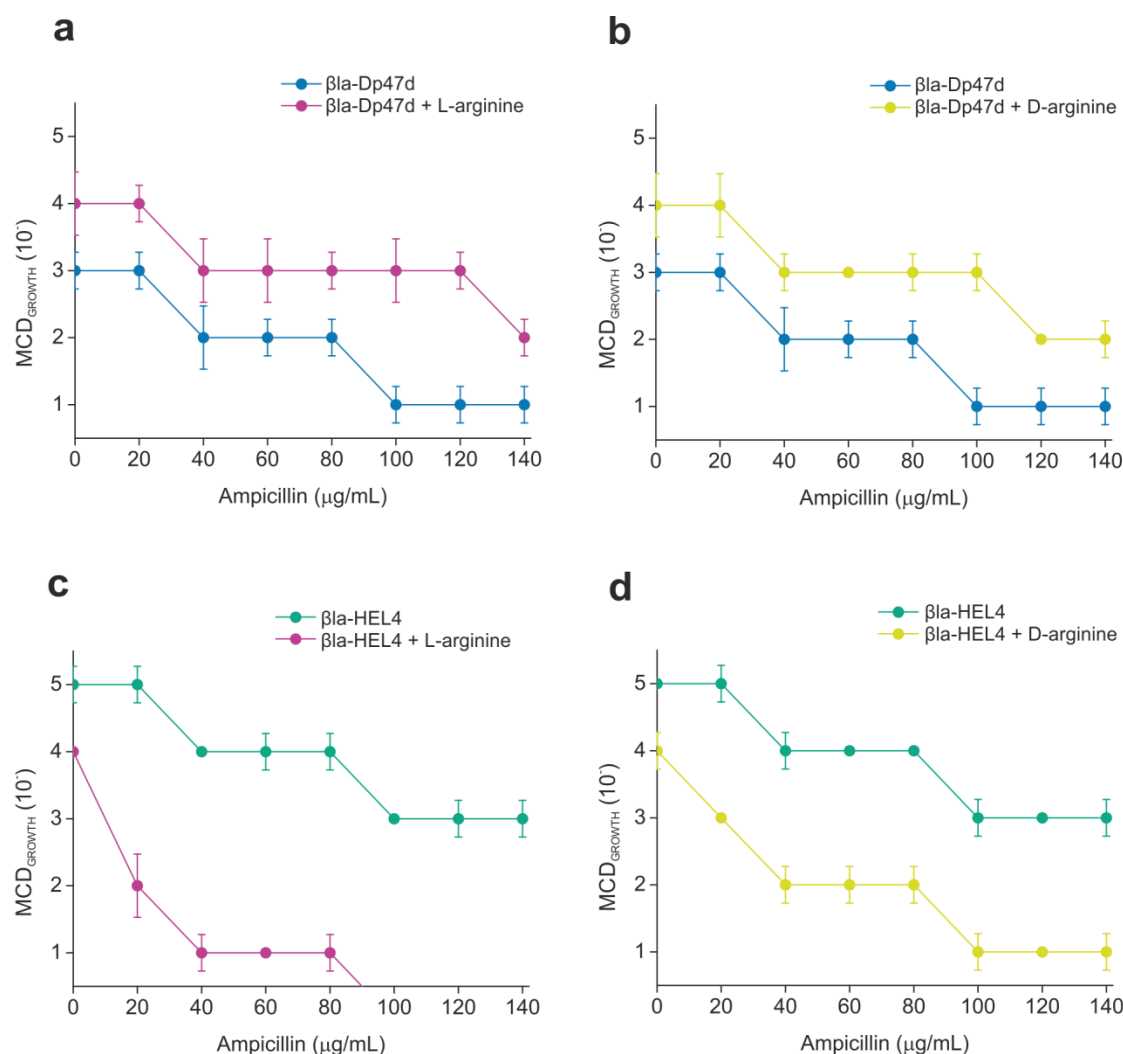
**Figure 5.12.** Bacterial growth in presence of equimolar L-Arg and L-Glu. **(a)** MIC assay in the absence of free amino acids.  $\beta\text{la-HEL4}$  (●) is included as an example of a non-aggregating construct. The maximal cell dilution allowing growth ( $\text{MCD}_{\text{GROWTH}}$ ) was scored at each ampicillin concentration in the presence of **(b)** 25:25 mM, **(c)** 50:50 mM and **(d)** 100:100 mM L-Arg:L-Glu. In each case, any intrinsic effect that addition of L-Arg and L-Glu had on bacterial growth was accounted for as described in Section 4.2.4. **(e)** Bacterial growth in the presence (treated) versus the absence (untreated) of equimolar mixtures of L-Arg:L-Glu (calculated using **Equation 4.1**). Error bars represent the standard error from a minimum of 3 replicates. Experiments were performed with undergraduate student Oliver Holman.

### 5.4.6 D-arginine is less toxic than L-arginine to bacterial growth

As substantial research demonstrates that exogenic addition of D-Glu has adverse effects on bacterial growth<sup>538-540</sup>, the possibility of utilising D-Arg (**Figure 5.13**) as a protein aggregation inhibitor was investigated. The aim was to reduce the toxicity, as compared to the addition of L-Arg, while retaining the anti-aggregation properties. 48-well agar plates were created containing either 50 mM L-Arg or 50 mM D-Arg, and the same MIC assay technique as described for the addition of L-Arg/L-Glu mixtures was followed (Section 2.4.5.2). After the intrinsic toxicity attributed to the addition of L-Arg or D-Arg was accounted for, the ability of the two isomers to rescue growth of bacterial expressing the aggregating construct  $\beta$ la-Dp47d was remarkably similar (**Figure 5.14a, b**). However, the toxicity induced by these free amino acids was significantly different. L-Arg (**Figure 5.14c**) was considerably more toxic to bacteria expressing the non-aggregating  $\beta$ la-HEL4 construct than D-Arg (**Figure 5.14d**). The bacterial colonies in the presence of L-Arg were small and faint, making them very difficult to score. Conversely, in the presence of D-Arg, the bacterial colonies were of a comparative size and shape to colonies in the absence of excipient addition. These data suggests that use of D-amino acids in place of their L- counterparts may be beneficial in screening for their anti-aggregation properties *in vivo*.



**Figure 5.13.** Structure of D-arginine at neutral pH.

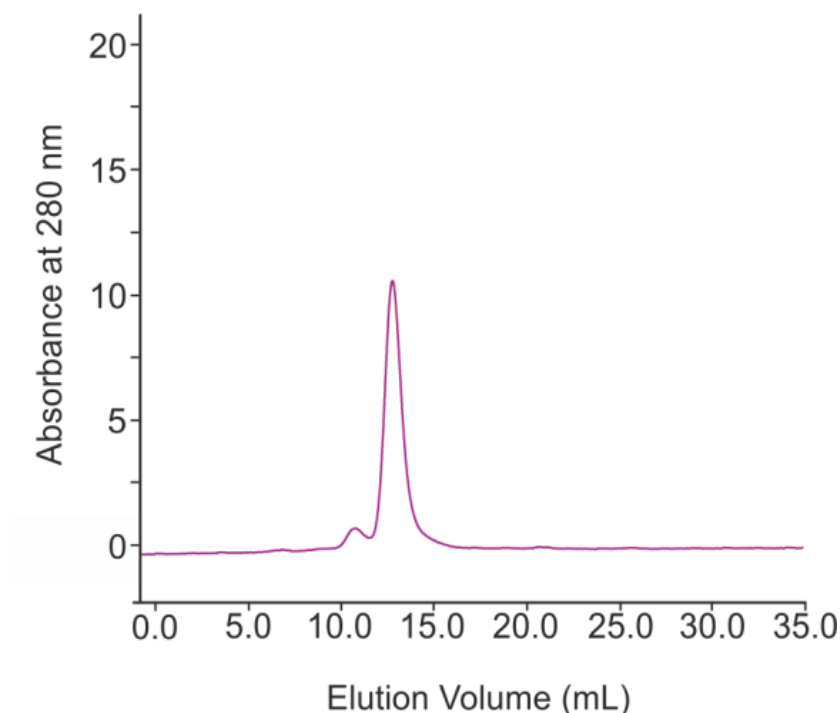


**Figure 5.14.** Comparison of bacterial growth in presence of L-Arg or D-Arg. **(a)** Growth of bacteria expressing the  $\beta$ la-Dp47d in the absence (●) or presence of 50 mM **(a)** L-Arg (●) or **(b)** D-Arg (●). The maximal cell dilution allowing growth ( $MCD_{GROWTH}$ ) was scored at each ampicillin concentration. In each case, any intrinsic effect that addition of 50 mM L-Arg or D-Arg had on bacterial growth was accounted for as described in Section 4.2.4. **(c-d)** Toxic effect on the growth of bacteria expressing the non-aggregating  $\beta$ la-HEL4 construct in the absence (●) or presence of 50 mM **(c)** L-Arg (●) or **(d)** D-Arg (●). Error bars represent the standard error from a minimum of 3 replicates. Experiments were performed with undergraduate student Oliver Holman.

### 5.4.7 Addition of *in vivo* 'hit' excipients aid purification of Dp47d

As demonstrated in Section 5.4.1, the dAb Dp47d cannot be purified in a soluble form *in vitro*. Due to the success of the MIC assay in the presence of sorbitol and glycerol, the ability of these 'hit' excipients to prevent aggregation of Dp47d *in vitro* was investigated.

The same purification protocol as detailed above (Section 5.4.1) was utilised to purify Dp47d in a denatured state. Refolding by dialysis was attempted in the presence of 0.5 M sorbitol (50 mM glycine, pH 9.5), and 0.5 M glycerol (50 mM glycine, pH 9.5). The presence of sorbitol prevented precipitation of Dp47d, and analytical size exclusion chromatography revealed the presence of mainly monomeric Dp47d (**Figure 5.15**). The identity of the protein was confirmed by mass spectrometry (observed mass: 12509.2 Da, expected mass: 12508.9 Da). 0.5 M glycerol was not able to solubilise Dp47d, with protein precipitation visibly identified. Due to time constraints, no further analyses of the effects of the excipients on Dp47d aggregation *in vitro* were performed.



**Figure 5.15.** Refolding of Dp47d using the excipient sorbitol. Dp47d elution profile from the Superdex™ 75 HR 10/30 analytical gel filtration column in the presence of 0.5 M sorbitol (50 mM glycine, pH 9.5). Column calibration confirms the main peak position corresponds to a ~ 13 kDa protein and the small peak position corresponds to a ~ 26 kDa protein.

## 5.5 Discussion

Protein-based pharmaceuticals are among the fastest growing category of therapeutic agents<sup>541</sup>, with significant emphasis on monoclonal antibodies and their isolated variable domains (dAbs). The small, monomeric size of dAbs makes them ideal candidates for therapeutics, imaging agents and diagnostics<sup>542</sup>. Unlike their larger parent molecules (mAbs), dAbs can cross the blood-brain-barrier<sup>543</sup>, penetrate tumour cells<sup>510, 544</sup>, and even inactivate target enzymes by binding within the active site<sup>509, 545</sup>. However, removing the native interacting partner domain exposes a large hydrophobic surface of the V<sub>H</sub> domain to solvent, so that these isolated molecules become severely aggregation-prone. As with all biopharmaceutical aggregation, this increases the risk of reduced efficacy, bioavailability and stability, leading to unwanted and potentially dangerous host immune responses<sup>546, 547</sup>.

Researchers are able to engineer antibodies on a rational basis, resulting in the production of more stable molecules with enhanced specificity and solubility<sup>320</sup>. However, protein design is much more complex than simply eliminating aggregation-prone sequences; the elimination of large stretches of hydrophobic residues, or even single residues, can significantly disrupt folding<sup>541</sup>. As such, *in vitro* purification of proteins containing various mutations can be required for biophysical analysis of their stability and solubility. A technique was developed to link phage display with the selection of thermally<sup>482</sup> (or thermodynamically<sup>517</sup>) stable dAbs, thereby providing a rapid and high throughput method to identify promising lead candidates.

It was hypothesised that the *in vivo*  $\beta$ -lactamase tripartite fusion system could also be used to identify aggregation-prone dAbs, and as such, may be of potential use to the biopharmaceutical industry. By linking the *in vivo* aggregation of the dAb to the enzymatic activity of  $\beta$ -lactamase, aggregation propensity could be characterised without the need for destabilising conditions such as increased temperature<sup>482</sup> or low pH<sup>517</sup>. As shown in Section 5.3.2 above, the *in vivo* assay successfully differentiated between the aggregation-prone human V<sub>H</sub> domain Dp47d and the highly stable dAb HEL4<sup>376</sup>. As the  $\beta$ -lactamase tripartite fusion system has successfully been used before to identify stabilising mutations in a number of test proteins<sup>308</sup>, the potential for this system to identify stabilising and/or aggregation-preventing mutations in biopharmaceutically relevant dAbs is promising.

As only a proof-of-principle has been demonstrated in this body of work, developing this system for practical use would require a number of considerations. Firstly, the inhibition

of *in vivo* aggregation must be linked to dAb function, as any mutation that decreases aggregation cannot significantly reduce antigen affinity or specificity. This could be achieved by coupling the MIC assay with a detection method such as the enzyme-linked immunosorbent assay (ELISA). For example, random mutagenesis could be used to create a library of mutants. These mutants would be cloned into the linker region of  $\beta$ -lactamase and transformed into *E. coli* cells. In a similar manner to phage display, each *E. coli* cell would then express a different mutant, thereby increasing the throughput of the assay. As the final readout of the assay is a highly sensitive immunodetection assay (ELISA), theoretically a traditional full MIC assay would not be required. Single colonies (or even liquid cultures) could be grown in the presence of one (optimised) concentration of antibiotic and any colonies present after 18 h incubation at 37 °C could be resuspended in a gentle lysis buffer to remove the outer membrane of the cells<sup>548</sup>. ELISA would allow direct detection of antigen binding by any non-aggregated constructs. Identical control plates could be used to extract and sequence the required mutant. In summary, the mutants would first be screened for aggregation propensity (by the MIC) and then for antigen affinity and specificity (by ELISA).

The approach described above could be tested using published Dp47d mutants with well characterised aggregation and stability properties<sup>376, 482</sup>. For example, mutating a key tryptophan residue residing in the  $V_H:V_L$  interface (W47R) results in an increase in protein recovery during purification from 5 % to > 80 %<sup>376</sup>. *In vitro* analyses showed that over 90 % of the protein was now in monomeric form, in contrast to WT Dp47d with which the authors could not obtain any data due to the rapid aggregation<sup>376</sup>. Theoretically, this mutation would result in more monomeric  $\beta$ la-Dp47d in the *in vivo* assay, which would be identifiable by ELISA (using hen-egg lysozyme<sup>376</sup>).

The ability of the *in vivo* assay to identify excipients, known to increase protein stability and prevent aggregation *in vitro*, was also investigated. Purification of soluble Dp47d in the absence of any additives that aid in maintaining protein solubility was not possible (Section 5.4.1). Consequently, a number of excipients were screened for their ability to prevent  $\beta$ la-Dp47d aggregation *in vivo*. The well-known protein stabilising osmolytes sorbitol and glycerol were able to rescue bacterial growth in a quantitative manner (Section 5.4.4). When grown in the presence of  $\geq 0.5$  M sorbitol, bacteria expressing the aggregating construct  $\beta$ la-Dp47d grew at the same level of antibiotic as bacteria expressing the non-aggregating  $\beta$ la-HEL4. A similar extent of growth recovery was obtained using 1.5 M glycerol. *In vitro* purification of Dp47d in the presence of these osmolytes corroborated the *in vivo* findings, with monomeric Dp47d only obtained when



0.5 M sorbitol was present in the refolding buffer. 0.5 M glycerol did not prevent the rapid precipitation of Dp47d.

The effects of charged amino acids on the stability and aggregation propensities of Dp47d were subsequently assessed *in vivo*. In contrast to the powerful stabilising properties observed with the osmolytes sorbitol and glycerol, L-Arg and L-Glu had only a small positive effect on the aggregation of  $\beta$ la-Dp47d *in vivo*. The concentrations chosen for the experiments (25, 50 and 100 mM) may have been too low to prevent aggregation; Arg concentrations greater than 500 mM are one of the most common additives known to assist refolding of recombinant proteins from inclusion bodies<sup>549-552</sup>. However, the properties of the colonies observed suggested that the presence of L-Arg and L-Glu were detrimental to the cells. In all cases, even in the absence of ampicillin, the colonies were small and faint. The presence of exogenous amino acids is known to regulate a number of genes via the enhancement or reversal of transcription factor expression<sup>553-555</sup>. These transcription factor responses make these amino acids not only nutrients but also signalling molecules<sup>556</sup>. Consequently, the presence of high concentrations of these amino acids is most likely interfering with cellular homeostasis.

Charged amino acids cannot diffuse across cellular membranes, and therefore must be internalised via specific transporter systems. *E. coli* cells have two L-Arg transporters, one with high affinity but low specificity for L-Arg, and one with low affinity but high specificity for L-Arg<sup>557</sup>. The low affinity transporter is down-regulated by the presence of excess L-Arg, however, conversely, the high affinity transporter, also responsible for the transport of other basic amino acids into the cell, is not inhibited by the presence of excess exogenous concentrations of L-Arg<sup>557, 558</sup>. Due to the significant importance of glutamate as a substrate in *E. coli* (~ 88 % of cellular nitrogen is obtained from glutamate<sup>559</sup>), there are three glutamate transporters. These transporters are regulated by the level of L-Glu present, however, similarly to L-Arg transporters, a high excess of L-Glu does not lead to complete down-regulation. In the context of the *in vivo* screen, the high concentrations of L-Arg and L-Glu utilised (25 – 100 mM) most likely resulted in significantly high concentrations of these amino acids within the cytoplasm.

The excess of charged amino acids can induce toxicity via a number of mechanisms. High concentrations of substrate can interfere with metabolic pathways. For example, metabolism of both L-Arg and L-Glu result in the eventual production of ammonia, which is reported, at concentrations above 100 mM, to inhibit cytosolic enzymes of some methanogenic bacteria<sup>559</sup>. A second mechanism is via pH-based growth inhibition. L-Arg

and L-Glu are key substrates utilised in the protection of bacteria from low environmental pH. Gastric acidity is a barrier through which all microbial food-borne pathogens must pass, with a pH value as low as 1.5 – 2.5, the mammalian stomach is one of the most inhospitable environments mammalian pathogens face<sup>560 561</sup>. Surprisingly, *E. coli* can survive at pH 2 for hours<sup>562</sup>, thanks to a number of acid-resistance systems. Two of these systems involve the decarboxylation of Arg and Glu, in which the  $\alpha$ -carboxyl group of the amino acid is replaced with a proton that is recruited from the cytoplasm<sup>563</sup>. The decarboxylated products are expelled from the cell in exchange for the import of new amino acid substrates<sup>560</sup>. The removal of intracellular protons by this mechanism causes the internal pH to increase to levels that are less toxic to the bacteria. The pH optima of the decarboxylases that carry out this function is pH 4 (for glutamate decarboxylases) and pH 5 (for arginine decarboxylases)<sup>564</sup>. As such, the enzyme's activity decreases as the internal pH of the bacteria is returned to normal. Critically, however, these enzymes are still able to function at neutral and higher pH, albeit with a slower activity<sup>565</sup>. Consequently, the addition of very high exogenous concentrations of L-Arg and L-Glu could be increasing the internal pH of the *E. coli* through the slow decarboxylation of the amino acids. A third possible route of toxicity could be via the inhibition of protein-protein interactions by L-Arg and L-Glu. As discussed previously, these charged amino acids can prevent protein aggregation by masking and neutralising opposite charges. However, this could pose a significant risk for endogenous protein-protein interactions within *E. coli* as the presence of such high concentrations of charged amino acids could prevent key cellular interactions from occurring.

It was hypothesised that the use of D-amino acids may circumvent at least some of the toxicity issues attributed to L-amino acids. L-amino acids are the predominant building blocks of proteins, as D-amino acids cannot be incorporated into proteins via ribosomal synthesis<sup>538</sup>. The majority of enzymes display marked L-amino acid substrate selectivity, and as a consequence, D-amino acids are not substrates in the majority of cellular processes. It was hypothesised that D-amino acids, therefore, may not interfere with cellular processes to the extent of their L-counterparts. However, even though L-amino acids are the dominant substrates for ribosome-based protein synthesis, D-amino acids do play a role in some other biological processes, in particular in bacteria<sup>538</sup>. In most bacteria, only D-Ala and D-Glu, which are incorporated into the peptidoglycan layer of the cell wall, are produced in significant quantities<sup>539</sup>. It was therefore hypothesised that addition of D-Glu to the growth media in excess would be equally detrimental to bacterial growth as excess addition of L-Glu. For this reason, MIC assays were only performed in the presence

or absence of 50 mM L-Arg or D-Arg. The level of growth rescue attributed to the presence of these two amino acids was consistently similar. This was to be expected, as the aggregation-inhibiting property of these molecules, specifically their positively charged side chain, is identical. Conversely, the toxicity exhibited by these two amino acids was markedly different. The presence of L-Arg consistently produced weak and faint colonies of bacteria, whereas the presence of D-Arg produced colonies with similar properties to those in the absence of any additive.

On the surface, the reduced toxicity from D-Arg compared to L-Arg appears to corroborate the possibility of utilising D-amino acids in place of L-amino acids during the *in vivo* screen. However, data from other studies suggest caution with this assumption. Racemases are enzymes responsible for the conversion of L-amino acid substrate into their D-counterparts. It has recently been discovered that, although D-Ala and D-Glu are by far the most commonly used D-amino acids in bacteria, most bacterial genomes encode one or more amino acid racemases in addition to the enzymes required to produce D-Glu and D-Ala<sup>539</sup>. The conservation of such systems across bacterial species suggests an important physiological role. Remarkably, it has been discovered that bacteria convert L-amino acids into their D-isomers during stationary growth phase<sup>539</sup>. The accumulation of D-amino acids leads to down-regulation of peptidoglycan synthesis (by a currently unknown mechanism) and overall metabolic slowing of the bacteria. It has been suggested that the rapid expression and diffusion of these small molecule regulators enables a quick and synchronised response from the whole bacterial population in times of scarce nutrients (when they enter stationary phase)<sup>538</sup>. Furthermore, studies have reported that exogenous addition of D-amino acids in combination with  $\beta$ -lactam antibiotic leads to an even stronger and synergistic reduction in bacterial growth<sup>539, 566</sup>.

Although some inhibition of  $\beta$ la-Dp47d aggregation was observed in the presence of each of the amino acids screened, the findings discussed above, along with other studies which demonstrate that D-amino acids can also be used as nutrients under some conditions<sup>567, 568</sup>, are evidence that further attempts to assess D-amino acid propensity as anti-aggregation excipients *in vivo* would be unwise. Within the biopharmaceutical industry, protein aggregation is controlled or minimised by optimising solution conditions such as pH, ionic strength, and/or added excipients. Optimisation of each of these parameters is often complicated due to the interactions between them. As such, high throughput screening approaches are used to evaluate large matrices of possible combinations of conditions<sup>521, 522, 569</sup>. Optical density, extrinsic and intrinsic fluorescence, dynamic and static light scattering and particle count can be assessed in 96- or even 384-well plate

format<sup>521, 522, 569</sup>. Screening such complex combinations of additives *in vivo* would be fraught with difficulties, due to the unknown effects each of the additives would have on bacterial growth. Furthermore, excipients are often only required to aid the stability of biopharmaceuticals during production, as biopharmaceuticals are expected to be of high enough stability to resist aggregation/degradation *in vivo* long enough to exert their therapeutic effects<sup>355</sup>. As such, the identification of excipients that prevent aggregation in the *in vivo* assay may not be relevant or suitable to prevent *in vitro* aggregation during production. For example, some biopharmaceuticals are produced and formulated at a pH drastically different to the optimal growth pH of *E. coli*.

In conclusion, the *in vivo*  $\beta$ -lactamase tripartite fusion system developed here was shown successfully to differentiate between aggregating and non-aggregating dAbs inserted into the  $\beta$ -lactamase linker. This provides a potential new method for assessing biopharmaceutical aggregation propensity without the necessity to purify large numbers of potential lead compounds for *in vitro* analysis. Furthermore, the system is also able to identify excipients that can stabilise proteins and prevent unwanted protein-protein interactions during purification *in vitro*.

## 6 Concluding remarks and future directions

Preventing protein aggregation is of paramount importance in the mission to alleviate some of the prevalent diseases in the developed world, from the neurodegenerative disorders Alzheimer's disease and Parkinson's disease, to the systemic disease type II diabetes mellitus. The key pathological hallmark of these amyloid diseases is the accumulation of aggregated proteins into large fibrillated structures known as amyloid plaques<sup>397, 413</sup>. The majority of studies of these aggregation-based diseases posit that low molecular weight soluble oligomers, or high molecular weight prefibrillar intermediates en-route to amyloid fibrils, account for their toxicity<sup>119, 217, 570, 571</sup>.

One approach to prevent protein oligomerisation is the use of small molecules to bind specifically to the protein of interest, and inhibit the initial stages of misfolding or aggregation<sup>290, 572, 573</sup>. Unfortunately, as many aggregation-prone proteins are intrinsically disordered, and we currently lack understanding of the intricate details of the aggregation process, screening against the monomer is not always possible<sup>574</sup>. Identification of aggregation-prone proteins and inhibitors which can prevent their aggregation often requires the demanding purification of these proteins, to permit *in vitro* biophysical analyses. However, their hydrophobic nature and tendency to aggregate result in laborious and expensive purification protocols. *In vivo* systems have recently been utilised to circumvent the issues associated with purifying these challenging proteins<sup>301, 303, 575, 576</sup>, however these have been limited to experiments within the cytosol or extracellular milieu. As such, any inhibitor screen is limited to those molecules that can traverse the restrictive nature of the inner membrane. The work described herein differs from these approaches by basing the screen within the periplasmic space of *E. coli*. The presence of outer membrane porins allows diffusive access to any molecule smaller than  $\sim 600$  Da<sup>577</sup>. Furthermore, the oxidising milieu of the periplasm allows the formation of disulfide bonds, something that is prevented in the reducing environment of the cytoplasm of eukaryotes, or the cytosol of prokaryotes, unless they are specifically engineered to be oxidising<sup>578</sup>.

This thesis details the development of a generally applicable system that directly links the aggregation-propensity of a test protein to a simple phenotypic readout: antibiotic resistance. The test protein is inserted between two domains of the periplasmic-based reporter enzyme TEM-1  $\beta$ -lactamase. Upon correct folding of the test protein in the periplasm of *E. coli*, the two halves of  $\beta$ -lactamase are brought close enough together to associate, such that the bacteria are resistant to  $\beta$ -lactam antibiotics. If the test protein

aggregates, however, the  $\beta$ -lactamase domains will be prevented from associating and the bacteria lose their resistance to the antibiotics. In Chapter 3, four test proteins with varying degrees of aggregation-propensity were cloned into the linker region of  $\beta$ -lactamase: rIAPP, hIAPP, A $\beta$ 40 and A $\beta$ 42. In remarkable agreement with previous *in vitro* analyses<sup>237, 407, 414, 415</sup>, the level of antibiotic resistance conferred to the bacteria expressing each construct correlated with the susceptibility of each test protein to aggregate. Furthermore, an ESI-IMS-MS time-course experiment revealed that the oligomeric species present 2 min into the aggregation reaction of the three aggregation-prone peptides (hIAPP, A $\beta$ 40 and A $\beta$ 42) are rapidly converted into higher order species as the reaction proceeds. Conversely, the small number of low molecular weight oligomers of rIAPP, the non-amyloidogenic construct, are still present after 24 h. Despite the extensive breadth of work on the sequence<sup>54, 262, 361, 363</sup> and structural<sup>41, 150, 402, 404, 411, 579</sup> causes of aggregation, there is still a significant amount to learn about these complex relationships. In this vein, the ability to distinguish aggregation-prone variants from their non-aggregating counterparts *in vivo* offers the opportunity to study the aggregation propensity of mutational variants of aggregation-prone proteins, without the requirement to purify each variant. Consequently, mutational work on mouse  $\beta_2$ -microglobulin ( $\beta_2m$ , another disulfide bond-containing immunoglobulin domain<sup>580, 581</sup>) is soon to be undertaken within the laboratory to investigate why a protein that is 70 % identical and shares 90 % sequence homology with human  $\beta_2m$  is so resistant to aggregation<sup>582-584</sup>, whereas amyloid formation by human  $\beta_2m$  is the underlying basis of the disease haemodialysis-related amyloidosis<sup>58, 583-585</sup>.

The ability to use the  $\beta$ -lactamase tripartite fusion system as an *in vivo* screen for small molecule inhibitors of protein aggregation was demonstrated in Chapter 4. Notably, not only did the assay correctly identify known inhibitors of hIAPP aggregation, it also refuted a number of published studies<sup>288, 290, 458, 459, 461, 462</sup>. The discrepancies in conclusions from this thesis and the literature are due to the choice of experimental techniques used to study protein aggregation. As extensively discussed in Chapter 1, a number of techniques are inappropriate to use alone in the search for small molecules with the capability to disrupt the aggregation pathway. By relying too heavily on methods such as dye binding studies, a number of studies presumptively identified 'inhibitors' of aggregation, whereby in fact, they were only inhibitors of the dye binding to the fibril<sup>266, 459</sup>. The tripartite fusion system developed here, coupled with ESI-IMS-MS (Chapter 4), clearly shows that a number of these claimed inhibitors did not, in fact, have any effect on protein aggregation. Importantly, the *in vivo* assay is also able to discriminate between specific inhibitors of

protein aggregation, and non-specific and colloidal interacting molecules (as determined by ESI-MS). This is a key requirement of an inhibitor screen, as any molecule that erroneously and non-specifically interacts with the protein of interest is undesirable as a therapeutic. Some studies that, albeit it correctly, reported inhibitors of protein aggregation, used upward of 20-fold molar excess of inhibitor to protein ratios<sup>266, 458, 459</sup>. The presence of such high concentrations of small molecule physically prevents protein from aggregating, however only in a colloidal manner. These types of inhibitors are ruled out in the *in vivo* screen, due to the low concentration of small molecule used, thereby minimising further downstream analyses of all potential 'hits'.

The inherent ability of proteins to aggregate underlies more than fifty human diseases<sup>45</sup> and, with an increasingly ageing population, the social and economic burdens on patients and society are expected to grow dramatically over the coming years. As such, it is imperative to be able to screen large pools of potential aggregation inhibitors against a diverse range of protein targets. Importantly, as the readout of the *in vivo* assay developed is independent of the activity of the inserted test protein, it obviates the need to develop individual assays for different test proteins to determine their aggregation-levels. Furthermore, the data in Chapter 4 illustrate that the *in vivo* assay is amenable to miniaturisation, which results in reduced small molecule requirements (~ 50 µg) and increased throughput. Further automation of the *in vivo* assay is crucial to exploiting its full potential, in particular for screening large libraries of protein variants for aggregation propensity. As such, this is to be investigated within the laboratory, in conjunction with industrial partners, over the coming year.

The assay described herein has advantages beyond the study of disease-related protein aggregation and its inhibition. The work in Chapter 5 is concerned with the control of protein-based pharmaceutical aggregation, one of the most demanding and costly elements in biopharmaceutical production and formulation. The ability of the assay to differentiate between an aggregation-prone and non-aggregating human dAb was investigated, and in accordance with the prior success utilising disease-related proteins (Chapter 3), the *in vivo* β-lactamase tripartite fusion system successfully distinguished the proteins by their aggregation-propensity. Furthermore, the addition of the well-known protein-stabilising excipients, sorbitol and glycerol, to the assay rescued bacterial growth in a titratable manner. Although similar results were not achieved using L/D-Arg or L-Glu, the success using polyols demonstrates that some classes of excipients are amenable to screening in such a manner.

The discovery that many, if not all, proteins could potentially form amyloid fibrils *in vitro*, indistinguishable from those extracted from patients, suggests that the ability to form amyloid is a generic, inherent property of all proteins<sup>42</sup>. Conversion of soluble proteins into amyloid fibrils is accelerated by a number of factors; mutation, transmissibility or even intracellular loss of environmental controls can lead to the accumulation of these molecular structures in a large variety of diseases<sup>47, 84, 111, 586</sup>. Indeed, that so many of these diseases are associated with old age suggests that as the population ages, not only will more cases of these diseases emerge, but more diseases caused by the aberrant aggregation of proteins may yet emerge. To date, only one successful small molecule therapeutic targeting a protein aggregation disease has been developed<sup>189-191</sup>. Nevertheless, the resources and effort currently being directed towards the detection of compounds means that new therapeutics, with the potential to arrest some of the most debilitating diseases in the world, may become available in the future. However, the future challenges of developing these drugs are not limited solely to the identification of molecules with the ability to prevent the underlying aggregation. Once identified, these compounds must be developed into potent, stable drug formulations that, in the case of neurodegenerative diseases, have the formidable task of crossing the blood-brain-barrier. Furthermore, many anti-amyloid therapies disappoint in clinical trials because the patients are given them too late<sup>161</sup>. If the diseases could be detected before symptoms manifest, these same therapies could be administered before irreversible damage has occurred, and perhaps the disease could be prevented. The complex multifaceted nature of these diseases demands a multi-pronged therapeutic attack. Rapid identification of potential protein aggregation inhibitors may not be the silver bullet, however it is a first, and important step to therapeutic alleviation of an ever growing class of devastating diseases. The system developed here may prove to be useful in these goals, either by direct identification of novel small molecules, or through its potential benefits in the context of biopharmaceuticals and the biopharmaceutical industry.



## 7 Appendices

### 7.1 DNA and protein sequences of all peptides and constructs

#### 7.1.1 Wild-type TEM-1 $\beta$ -lactamase

ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT  
GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTGGGTGCA  
 CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC  
 GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC  
 CGTGTTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG  
 GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA  
 TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC  
 GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT  
 GATCGTTGGG AACC GGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG  
 CCTGCAGCAA TGGCAACAAC GTTGCGCAAA CTATTAAGCTG GCGAACTATT GACTCTAGCT  
 AGCCGGCAGC AGCTCATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC  
 TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT  
 CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC  
 ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC  
 TCACTGATTA AGCATTGGTA A

**Appendix 7.1.** DNA sequence of wild-type TEM-1  $\beta$ -lactamase. The periplasmic signal sequence is in purple. The start and stop codons are underlined

MSIQHFRVAL IPFFAAFLP VFAHPETLVK VKDAEDQLGA RVGYIELDLN  
 SGKILESFRP EERFPMSTF KVLLCGAVLS RVDAGQEQLG RRIHYSQNDL  
 VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTI GGPKELTFL  
 HNMGDHVRTL DRWEPELNEA IPNDERDTM PAAMATTLRK LLTGELLTLA  
 SRQQLIDWME ADKVAGPLLK SALPAGWFIA DKSGAGERGS RGIIAALGPD  
 GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW

Molecular weight after signal sequence cleavage: 28.9 kDa

**Appendix 7.2.** Protein sequence of wild-type TEM-1  $\beta$ -lactamase. Signal sequence is highlighted in purple. The signal sequence is cleaved off after translocation into the periplasm.

## 7.1.2 $\beta$ la-linker<sub>SHORT</sub>

```

ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT
GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA
CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC
GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC
CGTGTTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG
GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA
TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC
GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT
GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG
CCTGCAGCAA TGGCAACAAC GTTGCGCAA CTATTAAC TGCGAACTAGG TGGTGGTGGT
TCAGGCGGAG GTGGCTCGAG CTCAGGATCC GGGAGCGGTT CCGGAAGCGG AGGAGGTGGT
TCAGGCGGAG GTGGAAGCTT GACTCTAGCT AGCCGGCAGC AGCTCATAGA CTGGATGGAG
GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT
GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT
GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC TATGGATGAA
CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACTGATTA AGCATTGGTA A

```

**Appendix 7.3.** DNA sequence of  $\beta$ la-linker<sub>SHORT</sub>. The 28 residue GS linker is shown in bold. The restriction sites for *Xho*I and *Bam*HI restriction enzymes are shown in blue and red respectively. The periplasmic signal sequence is in purple. The start and stop codons are underlined.

```

MSIQHFRVAL IPFFAAFLCLP VFAHPETLVK VKDAEDQLGA RVGYIELDLN
SGKILESFRP EERFPMSTF KVLICGAVLS RVDAGQEQLG RRIHYSQNDL
VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLLTTI GPKELTAFL
HNMGDHVRTL DRWEPELNEA IPNDERDTM PAAMATTLRK LLTGELGGGG
SGGGSSSSGS GSGSGSGGGG SGGGGSLTLA SRQQLIDWME ADKVAGPLLR
SALPAGWFIA DKSGAGERGS RGIIAALGPD GKPSRIVVIY TTGSQATMDE
RNRQIAEIGA SLIKHW

```

Molecular weight after signal sequence cleavage: 30.9 kDa

**Appendix 7.4.** Protein sequence of  $\beta$ la-linker<sub>SHORT</sub>. The signal sequence is shown in purple and the GS linker is in bold. The signal sequence is cleaved off after translocation into the periplasm.

### 7.1.3 $\beta$ la-linker

```

ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT
GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA
CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC
GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC
CGTGTTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG
GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA
TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC
GGAGGACCGA AGGAGCTAAC CGCTTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT
GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG
CCTGCAGCAA TGGCAACAAC GTTGCGCAA CTATTAAC TGCGAACTAGG TGGTGGTGGT
TCTGGTGGTG GTGGTTCTTC CTCAGGTTCA GGTGGCGGGG GATCTGGTGG TGGTGGCTCA
GGATCCGGTG GCTCGAGTTC CGGGAGCGGG AGCTCTTCTG GTTCCGGAGG CGGTGGAGGA
TCAGGCGGTG GCGGATCAGG AAGTGGGAGC GGAGGCGGGC GATCAGGCGG AGGTGGAAGC
TTGACTCTAG CTAGCCGGCA GCAGCTCATA GACTGGATGG AGGCGGATAA AGTTGCAGGA
CCACTTCTGC GCTCGGCCCT TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT
GAGCGTGGGT CTCGCGGTAT CATTCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC
GTAGTTATCT ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT
GAGATAGGTG CCTCACTGAT TAAGCATTGG TAA

```

**Appendix 7.5.** DNA sequence of  $\beta$ la-linker. The signal sequence is highlighted in purple. The 64-residue GS linker is shown in bold. The restriction sites for *Bam*H1 and *Xho*I restriction enzymes are shown in red and blue respectively. The start and stop codons are underlined.

```

MSIQHFRVAL IPFFAAFLCP VFAHPETLVK VKDAEDQLGA RVGYIELDLN
SGKILESFRP EERFPMSTF KVLICGAVLS RVDAGQEQLG RRIHYSQNDL
VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLLTII GGPKELTAFL
HNMGDHVTRL DRWEPELNEA IPNDERDTM PAAMATTLRK LLTGELGGGG
SGGGSSSSGS GGGGSGGGGS GSGGSSSSGS SSSGSGGGGG SGGGSGSGS
GGGGSGGGGS LTLASRQLI DWMEADKVAG PLLRSALPAG WFIADKSGAG
ERGSRGIIAA LGPDGKPSRI VVIYTTGSQA TMDERNRQIA EIGASLIKHW

```

Molecular weight after signal sequence cleavage: 33.2 kDa

**Appendix 7.6.** Protein sequence of  $\beta$ la-linker. The signal peptide is highlighted in purple. And the 64-residue GS linker is shown in bold. The signal sequence is cleaved off after translocation into the periplasm.

### 7.1.4 $\beta$ la-hIAPP

```

ATGAGTATTC AACATTTCCG TGTGCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT
GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA
CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC
GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC
CGTGTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG
GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA
TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC
GGAGGACCGA AGGAGCTAAC CGTTTTTTTT CACAACATGG GGGATCATGT AACTCGCCTT
GATCGTTGGG AACC GGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG
CCTGCAGCAA TGGCAACAAC GTTGC GCAA CTATTA ACTG CGGA ACTAGG TGGTGGTGGT
TCTGGTGGTG GTGGCTCGAG AAAATGCAAC ACCGCGACCT GCGCGACCCA GCGCCTGGCG
AACTTTCTGG TGCATAGCAG CAACAAC TTT GGCGCGATT C TGAGCAGCAC CAACGTGGGC
AGCAACACCT ATGGATCCGG GAGCGGTTC GGAAGCGGAG GAGGTGGTTC AGGCGGAGGT
GGAAGCTTGA CTCTAGTAG CCGGCAGCAG CTCATAGACT GGATGGAGGC GGATAAAGTT
GCAGGACCAC TTCTGCGCTC GGCCTTCCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA
GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC
CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG
ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTGGTAA

```

**Appendix 7.7.** DNA sequence of  $\beta$ la-hIAPP. The signal peptide is shown in purple and the GS linker is shown in bold. The restriction sites for *Xho*I and *Bam*HI restriction enzymes are shown in blue and red respectively. hIAPP is indicated in orange. The start and stop codons are underlined.

```

MSIQHFRVAL IPFFAAFLCLP VFAHPETLVK VKDAEDQLGA RVGYIELDLN
SGKILESFRP EERFPMMSTF KVL LCGAVLS RVDAGQEQLG RRIHYSQNDL
VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLLTTI GPKELTAFL
HNMGDHVTRL DRWEPELNEA IPNDERDTTM PAAMATTLRK LLTGELGGGG
SGGGGSRKCN TATCATQRLA NFLVHSSNNF GAILSSTNVG SNTYGS GSGS
GSGGGGSGGG GSLTLASRQQ LIDWMEADKV AGPLLR SALP AGWFIADKSG
AGERGSRGII AALGPDGKPS RIVVIYTTGS QATMDERNRQ IAEIGASLIK
HW

```

Molecular weight after signal sequence cleavage: 34.8 kDa

**Appendix 7.8.** Protein sequence of  $\beta$ la-hIAPP. The signal peptide is shown in purple and the GS linker is shown in bold. hIAPP is indicated in orange. The signal sequence is cleaved off after translocation into the periplasm.

### 7.1.5 $\beta$ la-rIAPP

```

ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT
GTTTTTGGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA
CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC
GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC
CGTGTTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG
GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA
TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC
GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT
GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG
CCTGCAGCAA TGGCAACAAC GTTGCGCAAA CTATTAAGT GCGAACTAGG TGGTGGTGGT
TCTGGTGGTG GTGGCTCGAG AAAATGCAAC ACCGCGACCT GCGCGACCCA GCGCCTGGCG
AACTTTCTGG TCGCAGCAG CAACAACCTG GGCCCGGTGC TGAGCAGCAC CAACGTGGGC
AGCAACACCT ATGGATCCGG GAGCGGTTCC GGAAGCGGAG GAGGTGGTTC AGGCGGAGGT
GGAAGCTTGA CTCTAGCTAG CCGGCAGCAG CTCATAGACT GGATGGAGGC GGATAAAGTT
GCAGGACCAC TTCTGCGCTC GGCCCTTCCG GCTGGCTGGT TTATTGCTGA TAAATCTGGA
GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCCAGATGG TAAGCCCTCC
CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG
ATCGCTGAGA TAGGTGCCTC ACTGATTAAG CATTGGTAA

```

**Appendix 7.9.** DNA sequence of  $\beta$ la-rIAPP. The signal peptide is shown in purple and the GS linker is shown in bold. The restriction sites for *Xho*I and *Bam*H1 restriction enzymes are shown in blue and red respectively. rIAPP is indicated in orange. The start and stop codons are underlined.

```

MSIQHFRVAL IPFFAAFLCP VFAHPETLVK VKDAEDQLGA RVGYIELDLN
SGKILESFRP EERFPMSTF KVLICGAVLS RVDAGQEQLG RRIHYSQNDL
VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTI GGPKELTAFL
HNMGDHSVRL DRWEPELNEA IPNDERDTM PAAMATTLRK LLTGELGGGG
SGGGSRKCN TATCATQLA NFLVRSNNL GPVLPPTNVG SNTYGSYGS
GSGGGSGGG GSLTLASRQQ LIDWMEADKV AGPLLSALP AGWFIADKSG
AGERGSRGII AALGPDGKPS RIVVIYTTGS QATMDERNRQ IAEIGASLIK
HW

```

Molecular weight after signal sequence cleavage: 34.8 kDa

**Appendix 7.10.** Protein sequence of  $\beta$ la-rIAPP. The signal peptide is shown in purple and the GS linker is shown in bold. rIAPP is indicated in orange. The signal sequence is cleaved off after translocation into the periplasm.

## 7.1.6 $\beta$ la-A $\beta$ 40

```

ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT
GTTTTTGCTC ACCCAGAAA CGCTGGTGA AAGTAAAAG ATGCTGAAG ATCAGTTGG
GTGCACGAG TGGGTTACA TCGAACTGG ATCTCAACA GCGGTAAGA TCCTTGAGA
GTTTTTCGCC CCGAAGAAC GTTTTCCAA TGATGAGCA CTTTTAAAG TTCTGCTAT
GTGGCGCGG TATTATCCC GTGTTGACG CCGGGCAAG AGCAACTCG GTCGCCGCA
TACACTATT CTCAGAATG ACTTGGTTG AGTACTCAC CAGTCACAG AAAAGCATC
TTACGGATG GCATGACAG TAAGAGAAT TATGCAGT CTGCCATAA CCATGAGTG
ATAACACTG CGGCCAACT TACTTCTGA CAACGATCG GAGGACCGA AGGAGCTAA
CCGCTTTTT TGCACAACA TGGGGGATC ATGTA ACTC GCCTTGATC GTTGGGAAC
CGGAGCTGA ATGAAGCCA TACCAAACG ACGAGCGTG ACACCACGA TGCCTGCAG
CAATGGCAA CAACGTTGC GCAA ACTAT TAACTGGCG AACTAGGTG GTGGTGGTT
CTGGTGGTG GTGGCTCGA GAGATGCGG AGTTCCGTC ATGATTCAG GCTATGAAG
TCCACCATC AAAAACTGG TGTTCTTTG CAGAAGATG TGGGTTCAA ACAAAGGTG
CCATCATTG GACTCATGG TGGGTGGTG TTGTCGGAT CCGGGAGCG GTTCCGGAA
GCGGAGGAG GTGGTTCAG GCGGAGGTG GAAGCTTGA CTCTAGCTA GCCGGCAGC
AGCTCATAG ACTGGATGG AGGCGGATA AAGTTGCAG GACCACTTC TGCGCTCGG
CCCTTCCGG CTGGCTGGT TTATTGCTG ATAAATCTG GAGCCGGTG AGCGTGGGT
CTCGCGGTA TCATTGCAG CACTGGGGC CAGATGGTA AGCCCTCCC GTATCGTAG
TTATCTACA CGACGGGGA GTCAGGCAA CTATGGATG AACGAAATA GACAGATCG
CTGAGATAG GTGCCTCAC TGATTAAGC ATTGGTAA

```

**Appendix 7.11.** DNA sequence of  $\beta$ la-A $\beta$ 40. The signal peptide is shown in purple and the GS linker is shown in bold. The restriction sites for *Xho*I and *Bam*HI restriction enzymes are shown in blue and red respectively. A $\beta$ 40 is indicated in pink. The start and stop codons are underlined.

```

MSIQHFRVAL IPFFAAFCPL VFAHPETLVK VKDAEDQLGA RVGYIELDLN
SGKILESFRP EERFPMSTF KVLICGAVLS RVDAGQEQLG RRIHYSQNDL
VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTI GGPKELTAFL
HNMGDHVTRL DRWEPELNEA IPNDERDTM PAAMATTLRK LLTGELGGG
SGGGSRDAE FRHDSGYEVH HQKLVFFAED VGSNKGAIIG LMVGGVVGSG
SGSGSGGGS GGGGSLTLAS RQQLIDWMEA DKVAGPLLR ALPAGWFIAD
KSGAGERGSR GIIAALGPDG KPSRIVVIYTT GSQATMDERN RQIAEIGAS
LIKHW

```

Molecular weight after signal sequence cleavage: 35.2 kDa

**Appendix 7.12.** Protein sequence of  $\beta$ la-A $\beta$ 40. The signal peptide is shown in purple and the GS linker is shown in bold. A $\beta$ <sub>40</sub> is indicated in pink.

### 7.1.7 $\beta$ la-A $\beta$ 42

ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT  
GTTTTTGCTC ACCCAGAAA CGCTGGTGA AAGTAAAAG ATGCTGAAG ATCAGTTGG  
 GTGCACGAG TGGGTTACA TCGAACTGG ATCTCAACA GCGGTAAGA TCCTTGAGA  
 GTTTTCGCC CCGAAGAAC GTTTTCCAA TGATGAGCA CTTTTAAAG TTCTGCTAT  
 GTGGCGCGG TATTATCCC GTGTTGACG CCGGGCAAG AGCAACTCG GTCGCCGCA  
 TACACTATT CTCAGAATG ACTTGGTTG AGTACTCAC CAGTCACAG AAAAGCATC  
 TTACGGATG GCATGACAG TAAGAGAAT TATGCAGTG CTGCCATAA CCATGAGTG  
 ATAACACTG CGGCCAACT TACTTCTGA CAACGATCG GAGGACCGA AGGAGCTAA  
 CCGCTTTTT TGCACAACA TGGGGGATC ATGTAACTC GCCTTGATC GTTGGGAAC  
 CGGAGCTGA ATGAAGCCA TACCAAACG ACGAGCGTG ACACCACGA TGCCTGCAG  
 CAATGGCAA CAACGTTGC GCAAACAT TAACTGGCG AACTAGGTG **GTGGTGGTT**  
**CTGGTGGTG** **GTGGCTCGA** **GAGATGCGG** **AGTTCGTC** **ATGATTCAG** **GCTATGAAG**  
**TCCACCATC** **AAAAACTGG** **TGTTCTTTG** **CAGAAGATG** **TGGGTTCAA** **ACAAAGGTG**  
**CCATCATTG** **GACTCATGG** **TGGGTGGTG** **TTGTCATAG** **CGGGATCCG** **GGAGCGGTT**  
**CCGGAAGCG** **GAGGAGGTG** **GTTCAGGCG** **GAGGTGGAA** **GC**TTGACTC TAGCTAGCC  
 GGCAGCAGC TCATAGACT GGATGGAGG CCGATAAAG TTGCAGGAC CACTTCTGC  
 GCTCGGCC TTCCGGCTG GCTGGTTTA TTGCTGATA AATCTGGAG CCGGTGAGC  
 GTGGGTCTC GCGGTATCA TTGCAGCAC TGGGGCCAG ATGGTAAGC CCTCCCGTA  
 TCGTAGTTA TCTACACGA CGGGGAGTC AGGCAACTA TGGATGAAC GAAATAGAC  
 AGATCGCTG AGATAGGTG CCTCACTGA TTAAGCATT GGTAA

**Appendix 7.13.** DNA sequence of  $\beta$ la-A $\beta$ 42 The signal peptide is shown in purple and the GS linker is shown in bold. The restriction sites for *Xho*I and *Bam*HI restriction enzymes are shown in blue and red respectively. A $\beta$ 42 is indicated in pink. The start and stop codons are underlined.

MSIQHFRVAL IPFFAAFCPLP VFAHPETLVK VKDAEDQLGA RVGYIELDLN  
 SGKILESFRP EERFPMSTF KVLCCGAVLS RVDAGQEQLG RRIHYSQNDL  
 VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLTTI GGPKELTAFI  
 HNMGDHVTRL DRWEPELNEA IPNDRDSTM PAAMATTLRK LLTGEL**GGGG**  
**SGGGSR**DAE **FRHDSGYEVH** **HQKLVFFAED** **VGSNKGAIIG** **LMVGGVVIAG**  
**SGSGSGSGGG** **GSGGGG**SLTL ASRQQLIDWM EADKVAGPLL RSALPAGWFI  
 ADKSGAGERG SRGIIAALGP DGKPSRIVVIY TTGSQATMDE RNRQIAEIGA  
 SLIKHW

Molecular weight after signal sequence cleavage: 35.4 kDa

**Appendix 7.14.** Protein sequence of  $\beta$ la-A $\beta$ 42. The signal peptide is shown in purple and the GS linker is shown in bold. A $\beta$ 42 is indicated in pink.

7.1.8  $\beta$ la-HEL4

ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT  
 GTTTTTGCTC ACCCAGAAA CGCTGGTGA AAGTAAAAG ATGCTGAAG ATCAGTTGG  
 GTGCACGAG TGGGTTACA TCGAACTGG ATCTCAACA GCGGTAAGA TCCTTGAGA  
 GTTTTTCGCC CCGAAGAAC GTTTTCCAA TGATGAGCA CTTTTAAAG TTCTGCTAT  
 GTGGCGCGG TATTATCCC GTGTTGACG CCGGGCAAG AGCAACTCG GTCGCCGCA  
 TACACTATT CTCAGAATG ACTTGGTTG AGTACTCAC CAGTCACAG AAAAGCATC  
 TTACGGATG GCATGACAG TAAGAGAAT TATGCAGTG CTGCCATAA CCATGAGTG  
 ATAACACTG CGGCCAACT TACTTCTGA CAACGATCG GAGGACCGA AGGAGCTAA  
 CCGCTTTTTT TGCACAACA TGGGGGATC ATGTAACTC GCCTTGATC GTTGGGAAC  
 CGGAGCTGA ATGAAGCCA TACCAAACG ACGAGCGTG ACACCACGA TGCCTGCAG  
 CAATGGCAA CAACGTTGC GCAAACATAT TAACTGGCG AACTAGGTG **GTGGTGGTT**  
**GTGGTGGTT** **CTGGTGGTG** **GTGGCTCGA** **GAGAAGTGC** **AGCTGCTGG** **AAAGCGGCG**  
 GCGGCCTGG TGCAGCCGG GCGGCAGCC TGCGCCTGA GCTGCGCGG CGAGCGGCT  
 TTCGCATTA GCGATGAAG ATATGGGCT GGGTGCGCC AGGCGCCGG GCAAAGGCC  
 TGAATGGG TGAGCAGCA TTTATGGCC CGAGCGGCA GCACCTATT ATGCGGATA  
 GCGTGAAAG GCCGCTTTA CCATTAGCC GCGATAACA GCAAAAACA CCCTGTATC  
 TGCAGATGA ACAGCCTGC GCGCGGAAG ATACCGCGG TGTATTATT GCGCGAGCG  
 CGCTGGAAC CGCTGAGCG AACCGCTGG GCTTTTGGG GCCAGGGCA CCCTGGTGA  
 CCGTGAGCA **GC**GGATCC**** **GGAGCGGTT** **CCGGAAGCG** **GAGGAGGTG** **GTT**CAGGCC****  
**GAGGTGGAA** **GC**TTGACTC TAGCTAGCC GGCAGCAGC TCATAGACT GGATGGAGG  
 CGGATAAAG TTGCAGGAC CACTTCTGC GCTCGGCC TTCCGGCTG GCTGGTTTA  
 TTGCTGATA AATCTGGAG CCGGTGAGC GTGGGTCTC GCGGTATCA TTGCAGCAC  
 TGGGGCCAG ATGGTAAAG CCTCCCGTA TCGTAGTTA TCTACACGA CGGGGAGTC  
 AGGCAACTA TGGATGAAC GAAATAGAC AGATCGCTG AGATAGGTG CCTCACTGA  
 TTAAGCATT GGTAA

**Appendix 7.15.** DNA sequence of  $\beta$ la-HEL4. The signal peptide is shown in purple and the GS linker is shown in bold. The restriction sites for *Xho*I and *Bam*HI restriction enzymes are shown in blue and red respectively. HEL4 is indicated in green. The start and stop codons are underlined.

MSIQHFRVAL IPFFAAFCPL VFAHPETLVK VKDAEDQLGA RVGYIELDLN  
 SGKILESFRP EERFPMSTF KVLICGAVLS RVDAGQEQLG RRIHYSQNDL  
 VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLLLTTI GGPKELTAFL  
 HNMGDHVTRL DRWEPELNEA IPNDERDTM PAAMATTLRK LLTGEL**GGGG**  
**SGGGGSRE**EVQ LLESGGGLVQ PGGSLRLSCA ASGFRISDED MGWVRQAPGK  
 GLEWVSSIYG PSGSTYYADS VKGRFTISR D NSKNTLYLQM NSLRAEDTAV  
 YYCASALEPL SEPLGFWGQG TLVTVSS**GSG** **SGSGSGGGGS** **GGGGS**LTLAS  
 RQQLIDWMEA DKVAGPLLR ALPAGWFIAD KSGAGERGSR GIIAALGPDG  
 KPSRIVVIYT TGSQATMDER NRQIAEIGAS LIKHW

Molecular weight after signal sequence cleavage: 43.7 kDa

**Appendix 7.16.** Protein sequence of  $\beta$ la-HEL4. The signal peptide is shown in purple and the GS linker is shown in bold. HEL4 is indicated in green.



### 7.1.9 $\beta$ la-Dp47d

```

ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT
GTTTTTGCTC ACCCAGAAA CGCTGGTGA AAGTAAAAG ATGCTGAAG ATCAGTTGG
GTGCACGAG TGGGTTACA TCGAACTGG ATCTCAACA GCGGTAAGA TCCTTGAGA
GTTTTCGCC CCGAAGAAC GTTTTCCAA TGATGAGCA CTTTTAAAG TTCTGCTAT
GTGGCGCGG TATTATCCC GTGTTGACG CCGGGCAAG AGCAACTCG GTCGCCGCA
TACACTATT CTCAGAATG ACTTGGTTG AGTACTCAC CAGTCACAG AAAAGCATC
TTACGGATG GCATGACAG TAAGAGAAT TATGCAGTG CTGCCATAA CCATGAGTG
ATAAACTG CGGCCAACT TACTTCTGA CAACGATCG GAGGACCGA AGGAGCTAA
CCGCTTTTT TGCACAACA TGGGGGATC ATGTAACTC GCCTTGATC GTTGGGAAC
CGGAGCTGA ATGAAGCCA TACCAAACG ACGAGCGTG ACACCACGA TGCCTGCAG
CAATGGCAA CAACGTTGC GCAAACAT TAACTGGCG AACTAGGTG GTGGTGGTT
GTGGTGGTT CTGGTGGTG GTGGCTCGA GAGAAGTGC AGCTGCTGG AAAGCGGCG
GCGGCCTGG TGCAGCCGG GCGGCAGCC TGCGCCTGA GCTGCGCGG CGAGCGGCT
TTACCTTTA GCAGCTATG CGATGAGCT GGGTGCGCC AGGCGCCGG GCAAAGGCC
TGGAATGGG TGAGCGCGA TTAGCGGCA GCGGCGGCA GCACCTATT ATGCGGATA
GCGTGAAAG GCCGCTTTA CCATTAGCC GCGATAACA GCAAAAACA CCCTGTATC
TGCAGATGA ACAGCCTGC GCGCGGAAG ATACCGCGG TGTATTATT GCGCGAAAA
GCTATGGCG CGTTTGATT ATTGGGGCC AGGGCACCC TGGTGACCG TGAGCAGCG
GATCCGGGA GCGGTTCCG GAAGCGGAG GAGGTGGTT CAGGCGGAG GTGGAAGCT
TGACTCTAG CTAGCCGGC AGCAGCTCA TAGACTGGA TGGAGGCGG ATAAAGTTG
CAGGACCAC TTCTGCGCT CGGCCCTTC CGGCTGGCT GGTTTATTG CTGATAAAT
CTGGAGCCG GTGAGCGTG GGTCTCGCG GTATCATTG CAGCACTGG GGCCAGATG
GTAAGCCCT CCCGTATCG TAGTTATCT ACACGACGG GGAGTCAGG CAACTATGG
ATGAACGAA ATAGACAGA TCGCTGAGA TAGGTGCCT CACTGATTA AGCATTGGT
AA

```

**Appendix 7.17.** DNA sequence of  $\beta$ la-Dp47d. The signal peptide is shown in purple and the GS linker is shown in bold. The restriction sites for *XhoI* and *BamHI* restriction enzymes are shown in blue and red respectively. Dp47d is indicated in green. The start and stop codons are underlined.

```

MSIQHFRVAL IPFFAAFLCLP VFAHPETLVK VKDAEDQLGA RVGYIELDLN
SGKILESFRP EERFPMSTF KVLICGAVLS RVDAGQEQLG RRIHYSQNDL
VEYSPVTEKH LTDGMTVREL CSAAITMSDN TAANLLLLTTI GGPKELTAFL
HNMGDHSVTRL DRWEPELNEA IPNDERDTM PAAMATTLRK LLTGELGGGG
SGGGGSREVQ LLESGGGLVQ PGGSLRLSCA ASGFTFSSYA MSWVRQAPGK
GLEWVSAISG SGGSTYYADS VKGRFTISRDN SKNTLYLQM NSLRAEDTAV
YYCAKSYGAF DYWGQGLVLT VSSGSGSGSG SGGGGSSLTLASRQQL
IDWMEADKVA GPLLRALPA GWFIADKSGA GERGSRGIIA ALGPDGKPSR
IVVIYTTGSQ ATMDERNRQI AEIGASLIKH W

```

Molecular weight after signal sequence cleavage: 43.3 kDa

**Appendix 7.18.** Protein sequence of  $\beta$ la-Dp47d. The signal peptide is shown in purple and the GS linker is shown in bold. Dp47d is indicated in green.

## 7.2 Genes from which test proteins were cloned

### 7.2.1 hIAPP

```

ATGAAAATTG AAGAAGGCAA CGCGAACCCG AACGCGAACC CGAACGCGAA CCCGGAAAAA
TGCAACACCG CGACCTGCGC GACCCAGCGC CTGGCGAACT TTCTGGTGCA TAGCAGCAAC
AACTTTGGCG CGATTCTGAG CAGCACCAAC GTGGGCAGCA ACACCTATTG CATCACGGGA
GATGCACTAG TTGCCCTACC CGAGGGCGAG TCGGTACGCA TCGCCGACAT CGTGCCGGGT
GCGCGGCCCA ACAGTGACAA CGCCATCGAC CTGAAAGTCC TTGACCGGCA TGGCAATCCC
GTGCTCGCCG ACCGGCTGTT CCACTCCGGC GAGCATCCGG TGTACACGGT GCGTACGGTC
GAAGGTCTGC GTGTGACGGG CACCGCGAAC CACCCGTTGT TGTGTTTGGT CGACGTGCGC
GGGGTGCCGA CCCTGCTGTG GAAGCTGATC GACGAAATCA AGCCGGGCGA TTACGCGGTG
ATTCAACGCA GCGCATTCAG CGTCGACTGT GCAGGTTTTG CCCGCGGAAA ACCCGAATTT
GCGCCCACAA CCTACACAGT CGGCGTCCCT GGACTGGTGC GTTTCTTGGA AGCACACCAC
CGAGACCCGG ACGCCCAAGC TATCGCCGAC GAGCTGACCG ACGGGCGGTT CTACTACGCG
AAAGTCGCCA GTGTCACCGA CGCCGGCGTG CAGCCGGTGT ATAGCCTTCG TGTCGACACG
GCAGACCACG CGTTTATCAC GAACGGGTTT CAGCCACCG CTACTGGCCT CACCGGTCTG
AACTCAGGCC TCACGACAAA TCCTGGTGTA TCCGCTTGGC AGGTCAACAC AGCTTATACT
GCGGGACAAT TGGTCACATA TAACGGCAAG ACGTATAAAT GTTTCAGCC CCACACCTCC
TTGGCAGGAT GGAACCATC CAACGTTCC TGCCTGTGGC AGCTTCAATG A

```

**Appendix 7.19.** DNA sequence of the gene from which hIAPP is cloned for insertion into  $\beta$ -lactamase. The signal peptide is shown in purple, hIAPP is in orange, the Mxe GyrA intein is blue and the chitin binding domain is red. Nucleotide differences between hIAPP and rIAPP are indicated in bold. The start and stop codons are underlined.

### 7.2.2 rIAPP

```

ATGAAAATTG AAGAAGGCAA CGCGAACCCG AACGCGAACC CGAACGCGAA CCCGGAAAAA
TGCAACACCG CGACCTGCGC GACCCAGCGC CTGGCGAACT TTCTGGTGCG CAGCAGCAAC
AACTTGGGCG CGGTGCTGCC GCCGACCAAC GTGGGCAGCA ACACCTATTG CATCACGGGA
GATGCACTAG TTGCCCTACC CGAGGGCGAG TCGGTACGCA TCGCCGACAT CGTGCCGGGT
GCGCGGCCCA ACAGTGACAA CGCCATCGAC CTGAAAGTCC TTGACCGGCA TGGCAATCCC
GTGCTCGCCG ACCGGCTGTT CCACTCCGGC GAGCATCCGG TGTACACGGT GCGTACGGTC
GAAGGTCTGC GTGTGACGGG CACCGCGAAC CACCCGTTGT TGTGTTTGGT CGACGTGCGC
GGGGTGCCGA CCCTGCTGTG GAAGCTGATC GACGAAATCA AGCCGGGCGA TTACGCGGTG
ATTCAACGCA GCGCATTCAG CGTCGACTGT GCAGGTTTTG CCCGCGGAAA ACCCGAATTT
GCGCCCACAA CCTACACAGT CGGCGTCCCT GGACTGGTGC GTTTCTTGGA AGCACACCAC
CGAGACCCGG ACGCCCAAGC TATCGCCGAC GAGCTGACCG ACGGGCGGTT CTACTACGCG
AAAGTCGCCA GTGTCACCGA CGCCGGCGTG CAGCCGGTGT ATAGCCTTCG TGTCGACACG
GCAGACCACG CGTTTATCAC GAACGGGTTT CAGCCACCG CTACTGGCCT CACCGGTCTG
AACTCAGGCC TCACGACAAA TCCTGGTGTA TCCGCTTGGC AGGTCAACAC AGCTTATACT
GCGGGACAAT TGGTCACATA TAACGGCAAG ACGTATAAAT GTTTCAGCC CCACACCTCC
TTGGCAGGAT GGAACCATC CAACGTTCC TGCCTGTGGC AGCTTCAATG A

```

**Appendix 7.20.** DNA sequence of the gene from which rIAPP is cloned for insertion into  $\beta$ -lactamase. The signal peptide is shown in purple, rIAPP is in orange, the Mxe GyrA intein is blue and the CBD is red. Nucleotide differences between hIAPP and rIAPP are indicated in bold. The start and stop codons are underlined.

### 7.2.3 A $\beta$ 40 and A $\beta$ 42

```

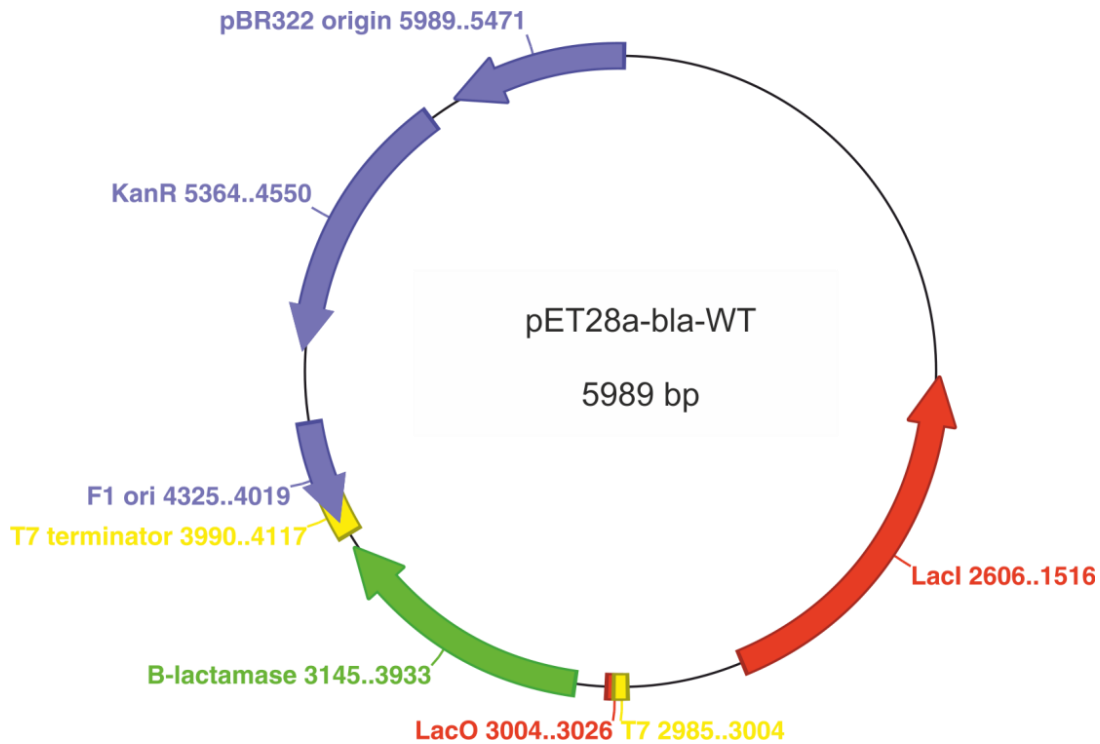
ATGAGAGGAT CGCATCACCA TCACCATCNC GGNTCTAACG CGAACCCGAA CGCGAACCCG
AACGCGAACC CGAACGCGAA CCCGAACGCG AACCCGAACG CGAACCCGAA CGCGAACCCG
AACGCGAACC CGAACGCGAA CCCGAACGCG AACCCGAACG CGAACCCGAA CGCGAACCCG
AACGCGAACC CGAACGCGAA CCCGAACGCG AACCCGAACG CGAACCCGAA CGCGAACCCG
AACGCGAACC CGAACGCGAA CCCGAGATCT GAAAACCTGT ATTTCCAGGA TGCGGAGTTC
CGTCATGATT CAGGCTATGA AGTCCACCAT CAAAAACTGG TGTTCTTTGC AGAAGATGTG
GGTTCAAACA AAGGTGCCAT CATTGGACTC ATGGTGGGTG GTGTTGTCAT AGCGTAA

```

**Appendix 7.21.** DNA sequence of the gene from which A $\beta$ 40 and A $\beta$ 42 were cloned. The beginning of the gene encodes His<sub>6</sub>-(NANP)<sub>19</sub>-TEV: a His-tag and a self-cleavage site to aid in purification. The A $\beta$ 40 and A $\beta$ 42 sequences are in blue, with the extra 6 nucleotides for A $\beta$ 42 highlighted in bold. The start and stop codons are underlined.

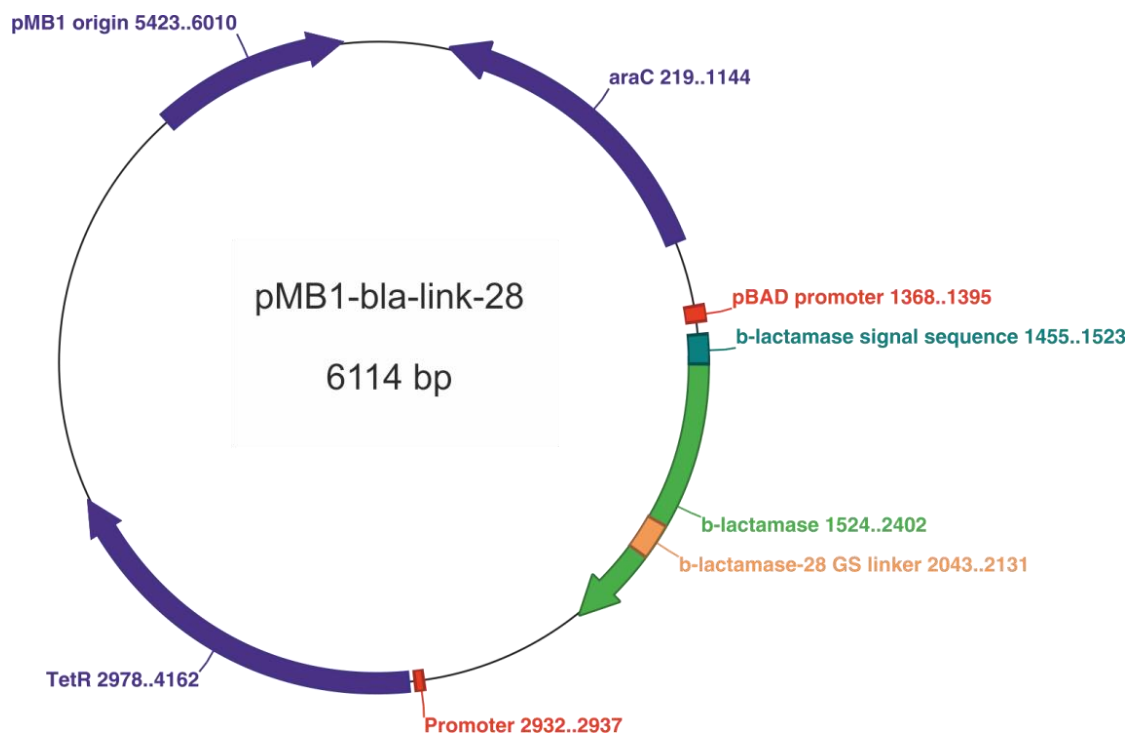
## 7.3 Plasmid maps of all plasmids obtained for this project

### 7.3.1 pET28a- $\beta$ la-WT



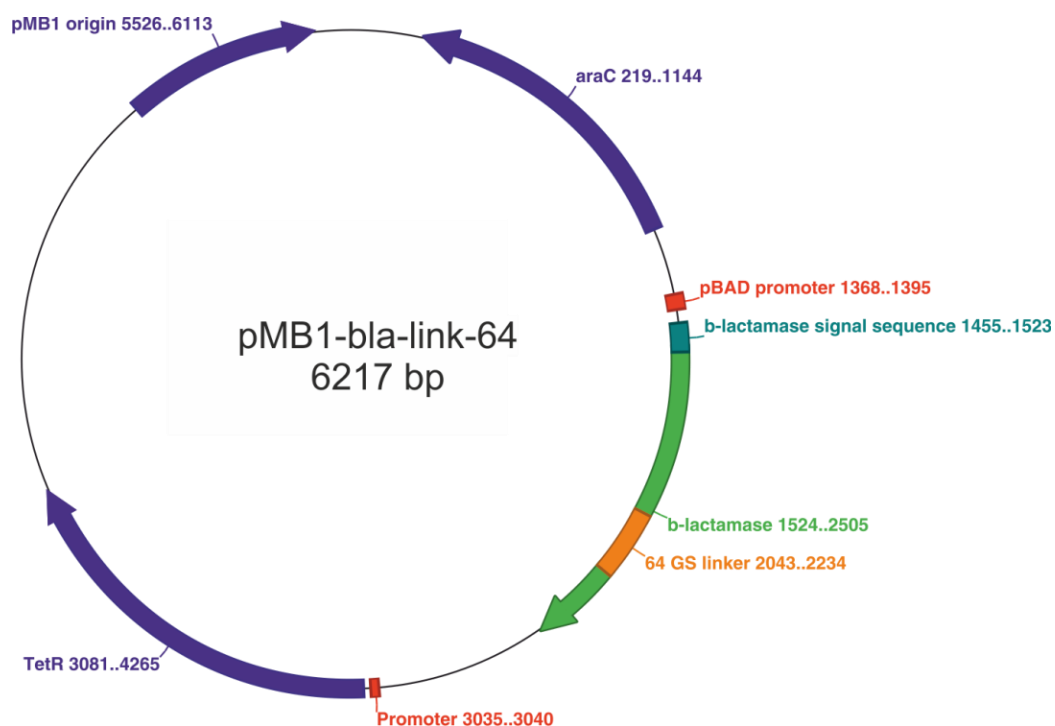
**Appendix 7.22.** Plasmid map of pET28a- $\beta$ la-WT, kindly provided by Professor J. Bardwell (Department of Biological Chemistry, University of Michigan, USA)<sup>308</sup>.

### 7.3.2 pBM1- $\beta$ la-link-28



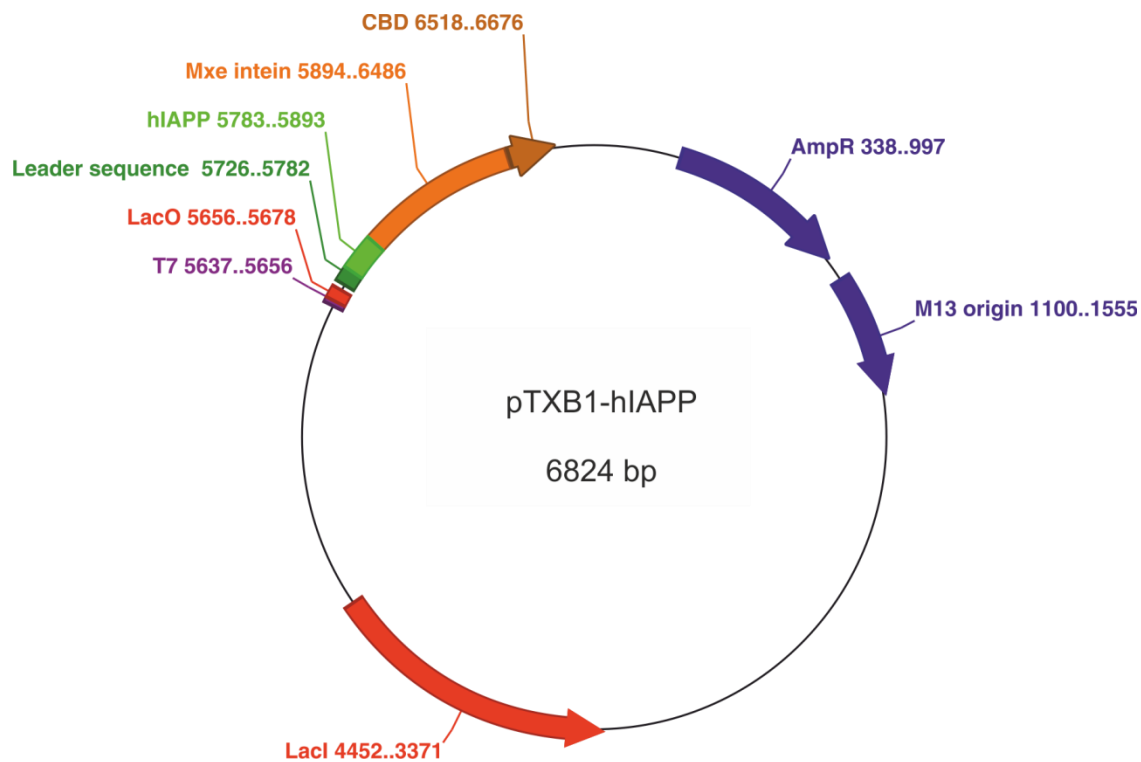
**Appendix 7.23.** Plasmid map of pMB1- $\beta$ la-link-28, kindly provided by Professor J. Bardwell (Department of Biological Chemistry, University of Michigan, USA)<sup>308</sup>.

### 7.3.3 pBM1- $\beta$ la-link-64



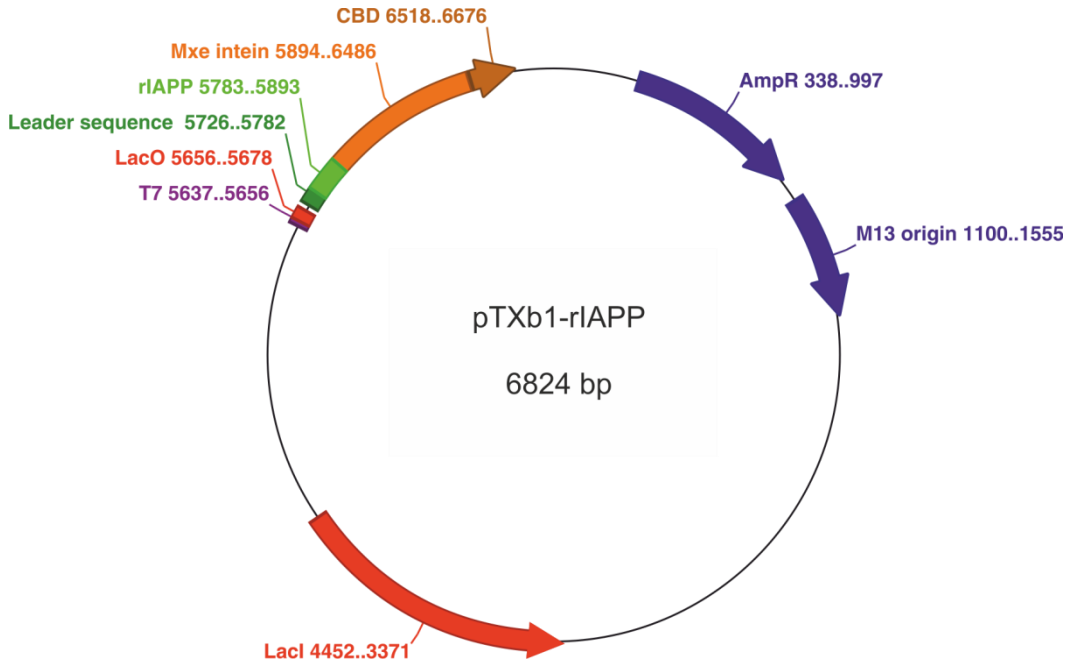
**Appendix 7.24.** Plasmid map of pMB1- $\beta$ la-link-64, kindly provided by Professor J. Bardwell (Department of Biological Chemistry, University of Michigan, USA)<sup>308</sup>.

### 7.3.4 pTXB1-hIAPP



**Appendix 7.25.** Plasmid map of pTXB1-hIAPP, kindly donated by Associate Professor A. Miranker (Department of Molecular Biophysics and Biochemistry, Yale University, USA)<sup>377</sup>.

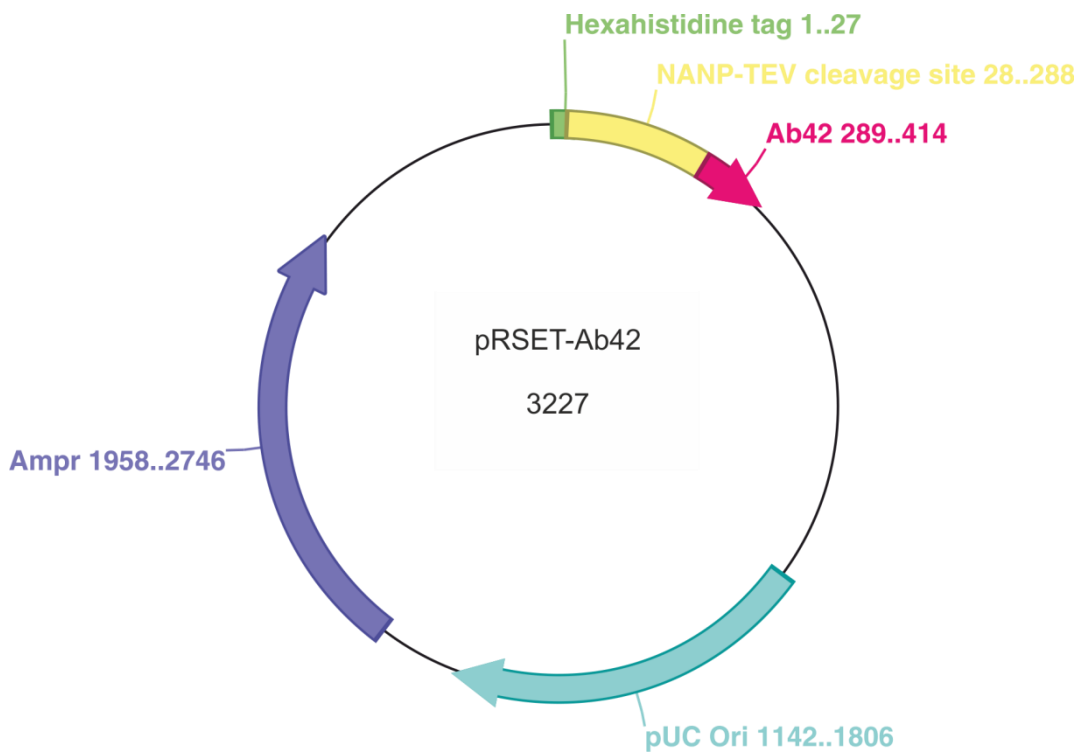
### 7.3.5 pTXB1-rIAPP



**Appendix 7.26.** Plasmid map of pTXB1-rIAPP, kindly donated by Associate Professor A. Miranker (Department of Molecular Biophysics and Biochemistry, Yale University, USA)<sup>377</sup>.

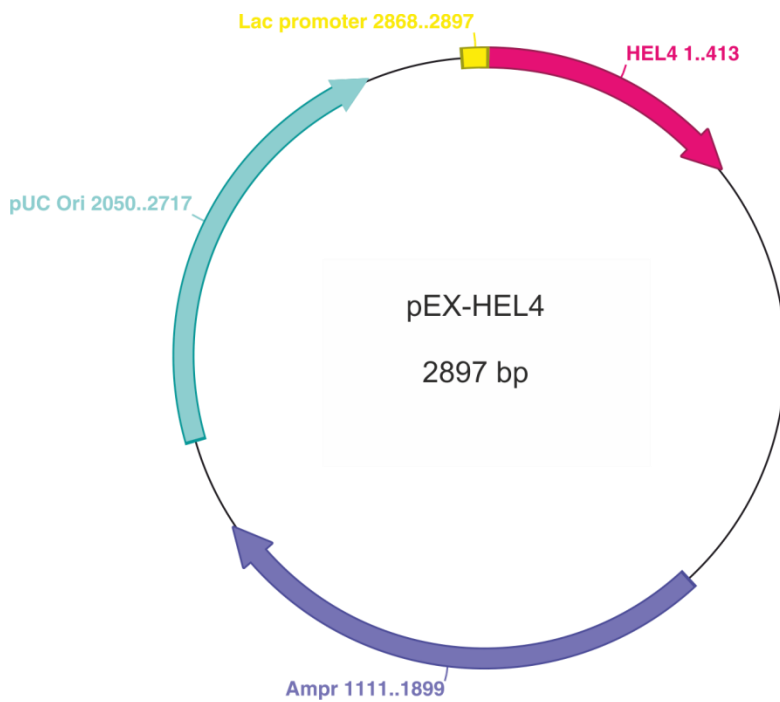


### 7.3.6 pRSET-A $\beta$ 42



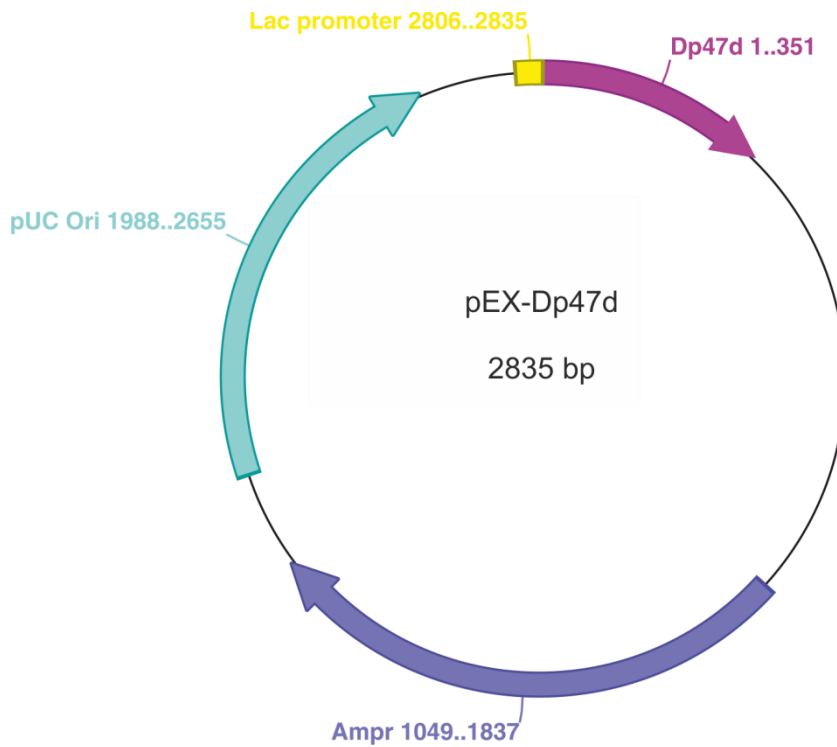
**Appendix 7.27.** Plasmid map of pRSET-A $\beta$ 40/42, kindly donated by Dominic Walsh (Brigham & Women's Hospital, Boston, USA) and Sara Linse (Lund University, Sweden)<sup>378</sup>.

### 7.3.7 pEX-HEL4



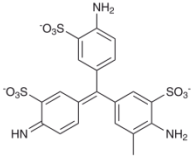
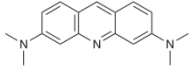

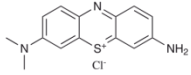
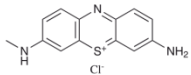
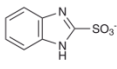

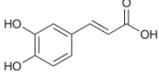
**Appendix 7.28.** Plasmid map of pEX-Hel4, synthesised by Eurofins MWG Operon (Ebersburg, Germany) using the sequence from Jaspers *et al.*, 2004<sup>376</sup>.

### 7.3.8 pEX-Dp47d

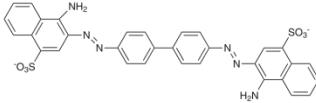
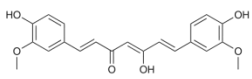
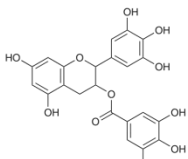
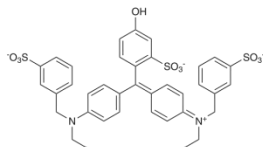
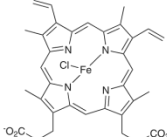
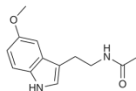
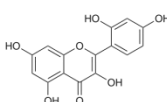
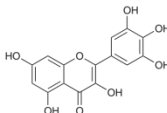


**Appendix 7.29.** Plasmid map of pEX-Dp47d, synthesised by Eurofins MWG Operon (Ebersburg, Germany) using the sequence from Jespers *et al.*, 2004<sup>376</sup>.

## 7.4 Small molecule compounds used in this study

Small Molecule	IUPAC Name	MW (Da)	LogP	Structure
Acid fuchsin	2-amino-5-({4-amino-3-methyl-5-[(sodiooxy)sulfonyl]phenyl}[(1E)-4-imino-3-[(sodiooxy)sulfonyl]cyclohexa-2,5-dien-1-ylidene]methyl)benzene-1-sulfonic acid	585.5	-3.9	
Acridine orange	N,N,N',N'-Tetramethylacridine-3,6-diamine	265.4	3.3	
Aspirin	2-Acetoxybenzoic acid	180.2	-2.1	
Azure A	N',N'-dimethylphenothiazin-5-ium-3,7-diamine chloride	291.8	0.0	
Azure C	3-Amino-7-(methylamino) phenothiazin-5-ium chloride	277.8	-0.2	
1H-B-SA	1H-Benzimidazole-2-sulfonic acid	198.2	-1.1	
Benzimidazole	1H-Benzimidazole	118.1	1.4	
Caffeic acid	3-(3,4-Dihydroxyphenyl)-2-propenoic acid 3,4-Dihydroxy-cinnamic acid trans-Caffeate	180.2	0.9	

**Table 7.1.** Molecular weight, LogP value and structure of small molecules used in this thesis. LogP values (the log of the aqueous/hydrophobic partition coefficient) were calculated using [www.molinspiration.com](http://www.molinspiration.com) software, which determines the hydrophobic parameters of the substituents. Molecules with high LogP values have high hydrophobicity.

Small Molecule	IUPAC Name	MW (Da)	LogP	Structure
Congo red	disodium 4-amino-3-[4-[4-(1-amino-4-sulfonato-naphthalen-2-yl)diazenylphenyl]phenyl]diazenyl-naphthalene-1-sulfonate	696.7	3.1	
Curcumin	1E,6E)-1,7-Bis(4-hydroxy-3-methoxyphenyl)-1,6-heptadiene-3,5-dione	368.4	3.0	
EGCG	[(2R,3R)-5,7-dihydroxy-2-(3,4,5-trihydroxyphenyl)chroman-3-yl] 3,4,5-trihydroxybenzoate	458.4	2.2	
Fast green FCF	Disodium 2-[(E)-{4-[ethyl(3-sulfonatobenzyl)amino]phenyl} {(4E)-4-[ethyl(3-sulfonatobenzyl)iminio]-2,5-cyclohexadien-1-ylidene} methyl]-5-hydroxybenzenesulfonate	765.9	-4.4	
Hemin	Chloro[3,7,12,17-tetramethyl-8,13-divinylporphyrin-2,18-dipropanoato(2-)]iron(III)	651.9	6.8	
Melatonin	N-[2-(5-methoxy-1H-indol-3-yl)ethyl]acetamide	232.3	1.4	
Morin hydrate	2-(2,4-dihydroxyphenyl)-3,5,7-trihydroxychromen-4-one	302.2	1.9	
Myricetin	3,5,7-Trihydroxy-2-(3,4,5-trihydroxyphenyl)-4-chromenone	318.2	1.4	

**Table 7.2.** Molecular weight, LogP value and structure of small molecules used in this thesis continued. LogP values (the log of the aqueous/hydrophobic partition coefficient) were calculated using [www.molinspiration.com](http://www.molinspiration.com) software, which determines the hydrophobic parameters of the substituents. Molecules with high LogP values have high hydrophobicity.

Small Molecule	IUPAC Name	MW (Da)	LogP	Structure
Orange G	Disodium 7-hydroxy-8-[( <i>E</i> )-phenyldiazenyl]-1,3-naphthalenedisulfonate	452.4	-0.3	
Phenol red	4,4'-(1,1-Dioxido-3 <i>H</i> -2,1-benzoxathiole-3,3-diyl)diphenol	354.4	-0.6	
Resveratrol	5-[( <i>E</i> )-2-(4-Hydroxyphenyl)-vinyl]-1,3-benzenediol	228.2	3.0	
Rhodamine B	9-(2-Carboxyphenyl)-6-(diethylamino)- <i>N,N</i> -diethyl-3 <i>H</i> -xanthen-3-iminium chloride	479.0	2.7	
Silibinin	(2 <i>R</i> ,3 <i>R</i> )-3,5,7-Trihydroxy-2-[(2 <i>R</i> ,3 <i>R</i> )-3-(4-hydroxy-3-methoxyphenyl)-2-(hydroxymethyl)-2,3-dihydro-1,4-benzodioxin-6-yl]-2,3-dihydro-4 <i>H</i> -chromen-4-one	482.4	1.5	
Thiabenzadole	2-(1,3-Thiazol-4-yl)-1 <i>H</i> -benzimidazole	201.2	2.3	
Tramiprosate	3-Amino-1-propanesulfonic acid	139.2	-3.5	

**Table 7.3.** Molecular weight, LogP value and structure of small molecules used in this thesis continued. LogP values (the log of the aqueous/hydrophobic partition coefficient) were calculated using [www.molinspiration.com](http://www.molinspiration.com) software, which determines the hydrophobic parameters of the substituents. Molecules with high LogP values have high hydrophobicity.

## References

1. Reynaud, E. Protein misfolding and degenerative diseases. *Nat. Educ.*, 2010, **3**(9), p.28.
2. Anfinsen, C.B., Haber, E., Sela, M. and White, F.H. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.*, 1961, **47**(9), pp.1309-1314.
3. Levinthal, C. How to fold graciously. In: *Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House, Monticello, Illinois*: University of Illinois Press, 1969, pp.22-24.
4. Wetlaufer, D.B. Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl. Acad. Sci. U. S. A.*, 1973, **70**(3), pp.697-701.
5. Fersht, A.R. From the first protein structures to our current knowledge of protein folding: delights and scepticisms. *Nat. Rev. Mol. Cell. Biol.*, 2008, **9**(8), pp.650-654.
6. Tsong, T.Y., Baldwin, R.L. and Elson, E.L. Properties of the refolding and unfolding reactions of Ribonuclease A. *Proc. Natl. Acad. Sci. U. S. A.*, 1972, **69**(7), pp.1809-1812.
7. Fersht, A.R. Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.*, 1997, **7**(1), pp.3-9.
8. Kim, P.S. and Baldwin, R.L. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.*, 1982, **51**(1), pp.459-489.
9. Karplus, M. and Weaver, D.L. Protein-folding dynamics. *Nature*, 1976, **260**(5550), pp.404-406.
10. Karplus, M. and Weaver, D. Protein folding dynamics: the diffusion-collision model and experimental data. *Protein Sci.*, 1994, **3**(4), pp.650-668.
11. Tanford, C. Contribution of hydrophobic interactions to the stability of the globular conformation of proteins. *J. Am. Chem. Soc.*, 1962, **84**(22), pp.4240-4247.
12. Baldwin, R.L. How does protein folding get started? *Trends Biochem. Sci.*, 1989, **14**(7), pp.291-294.
13. Jackson, S.E. and Fersht, A.R. Folding of chymotrypsin inhibitor 2. Evidence for a two-state transition. *Biochem.*, 1991, **30**(43), pp.10428-10435.
14. Otzen, D.E., Itzhaki, L.S., Elmasry, N.F., Jackson, S.E. and Fersht, A.R. Structure of the transition state for the folding/unfolding of the barley chymotrypsin inhibitor 2 and its implications for mechanisms of protein folding. *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**(22), pp.10422-10425.
15. Viguera, A.R., Martinez, J.C., Filimonov, V.V., Mateo, P.L. and Serrano, L. Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition. *Biochem.*, 1994, **33**(8), pp.2142-2150.
16. Schindler, T., Herrler, M., Marahiel, M.A. and Schmid, F.X. Extremely rapid protein folding in the absence of intermediates. *Nat. Struct. Mol. Biol.*, 1995, **2**(8), pp.663-673.
17. Fersht, A.R. Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**(24), pp.10869-10873.
18. Itzhaki, L.S., Otzen, D.E. and Fersht, A.R. The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J. Mol. Biol.*, 1995, **254**(2), pp.260-288.
19. Daggett, V. and Fersht, A.R. Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.*, 2003, **28**(1), pp.18-25.
20. Gianni, S., Guydosh, N.R., Khan, F., Caldas, T.D., Mayor, U., White, G.W.N., Demarco, M.L., Daggett, V. and Fersht, A.R. Unifying features in protein-folding mechanisms. *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**(23), pp.13286-13291.
21. Dill, K. and Chan, H. From Levinthal to pathways to funnels. *Nat. Struct. Mol. Biol.*, 1997, **4**(1), pp.10-19.
22. Jahn, T.R. and Radford, S.E. Folding versus aggregation: polypeptide conformations on competing pathways. *Arch. Biochem. Biophys.*, 2008, **469**(1), pp.100-117.

23. Hartl, F.U., Bracher, A. and Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature*, 2011, **475**(7356), pp.324-332.
24. Bartlett, A.I. and Radford, S.E. An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.*, 2009, **16**(6), pp.582-588.
25. Dobson, C.M. Principles of protein folding, misfolding and aggregation. *Semin. Cell. Dev. Biol.*, 2004, **15**(1), pp.3-16.
26. Chiti, F. and Dobson, C.M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.*, 2006, **75**(1), pp.333-366.
27. Knowles, T.P.J., Vendruscolo, M. and Dobson, C.M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell. Biol.*, 2014, **15**(6), pp.384-396.
28. Baldwin, A.J., Knowles, T.P.J., Tartaglia, G.G., Fitzpatrick, A.W., Devlin, G.L., Shammass, S.L., Waudby, C.A., Mossuto, M.F., Meehan, S., Gras, S.L., Christodoulou, J., Anthony-Cahill, S.J., Barker, P.D., Vendruscolo, M. and Dobson, C.M. Metastability of native proteins and the phenomenon of amyloid formation. *J. Am. Chem. Soc.*, 2011, **133**(36), pp.14160-14163.
29. Gazit, E. The "correctly folded" state of proteins: is it a metastable state? *Angew. Chem. Int. Ed.*, 2002, **41**(2), pp.257-259.
30. Hartl, F.U. and Hayer-Hartl, M. Converging concepts of protein folding *in vitro* and *in vivo*. *Nat. Struct. Mol. Biol.*, 2009, **16**(6), pp.574-581.
31. Kunitani, M., Wolfe, S., Rana, S., Apicella, C., Levi, V. and Dollinger, G. Classical light scattering quantitation of protein aggregates: off-line spectroscopy versus HPLC detection. *J. Pharm. Biomed. Anal.*, 1997, **16**(4), pp.573-586.
32. Livney, Y.D. and Dalgleish, D.G. Specificity of disulfide bond formation during thermal aggregation in solutions of  $\beta$ -lactoglobulin B and  $\kappa$ -casein A. *J. Agric. Food Chem.*, 2004, **52**(17), pp.5527-5532.
33. Arakawa, T., Prestrelski, S.J., Kenney, W.C. and Carpenter, J.F. Factors affecting short-term and long-term stabilities of proteins. *Adv. Drug Delivery Rev.*, 2001, **46**(1-3), pp.307-326.
34. Ross, C. and Poirier, M. Protein aggregation and neurodegenerative disease. *Nat. Med.*, 2004, **10**, pp.S10-S17.
35. Mahler, H.C., Friess, W., Grauschopf, U. and Kiese, S. Protein aggregation: pathways, induction factors and analysis. *J. Pharm. Sci.*, 2009, **98**(9), pp.2909-2934.
36. Ho, M.R., Lou, Y.C., Lin, W.C., Lyu, P.C., Huang, W.N. and Chen, C. Human pancreatitis-associated protein forms fibrillar aggregates with a native-like conformation. *J. Biol. Chem.*, 2006, **281**(44), pp.33566-33576.
37. Khemtémourian, L., Killian, J., Höppener, J., Engel, M.,. Recent insights in islet amyloid polypeptide-induced membrane disruption and its role in  $\beta$ -cell death in type 2 diabetes mellitus. *Exp. Diabetes Res.*, 2008, **421287**, pp.1-9.
38. Idicula-Thomas, S. and Balaji, P.V. Protein aggregation: a perspective from amyloid and inclusion-body formation. *Curr. Sci.*, 2007, **92**(6), pp.758-767.
39. Turoverov, K.K., Kuznetsova, I.M. and Uversky, V.N. The protein kingdom extended: ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. *Prog. Biophys. Mol. Biol.*, 2010, **102**(2-3), pp.73-84.
40. Howie, A.J., Brewer, D.B., Howell, D. and Jones, A.P. Physical basis of colors seen in Congo red-stained amyloid in polarized light. *Lab. Invest.*, 2007, **88**(3), pp.232-242.
41. Eisenberg, D., Nelson, R., Sawaya, M.R., Balbirnie, M., Sambashivan, S., Ivanova, M.I., Madsen, A. and Riek, C. The structural biology of protein aggregation diseases: some fundamental questions and some answers. *Acc. Chem. Res.*, 2006, **39**(9), pp.568-575.
42. Dobson, C.M. Protein-misfolding diseases: getting out of shape. *Nature*, 2002, **418**(6899), pp.729-730.
43. Stefani, M. Protein misfolding and aggregation: new examples in medicine and biology of the dark side of the protein world. *BBA - Mol. Basis Dis.*, 2004, **1739**(1), pp.5-25.



44. Mcparland, V.J. Investigations into different conformations of beta-2-microglobulin and their role in fibril formation. PhD thesis, *University of Leeds*, 2001.
45. Sipe, J.D., Benson, M.D., Buxbaum, J.N., Ikeda, S.-I., Merlini, G., Saraiva, M.J.M. and Westermark, P. Amyloid fibril protein nomenclature: 2012 recommendations from the Nomenclature Committee of the International Society of Amyloidosis. *Amyloid*, 2012, **19**(4), pp.167-170.
46. Guijarro, J.I.A., Sunde, M., Jones, J.A., Campbell, I.D. and Dobson, C.M. Amyloid fibril formation by an SH3 domain. *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**(8), pp.4224-4228.
47. Stefani, M. and Dobson, C. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *J. Mol. Med.*, 2003, **81**(11), pp.678-699.
48. Monsellier, E. and Chiti, F. Prevention of amyloid-like aggregation as a driving force of protein evolution. *EMBO Rep.*, 2007, **8**, pp.737-742.
49. Fitzpatrick, A.W.P., Debelouchina, G.T., Bayro, M.J., Clare, D.K., Caporini, M.A., Bajaj, V.S., Jaroniec, C.P., Wang, L., Ladizhansky, V., Müller, S.A., Macphee, C.E., Waudby, C.A., Mott, H.R., De Simone, A., Knowles, T.P.J., Saibil, H.R., Vendruscolo, M., Orlova, E.V., Griffin, R.G. and Dobson, C.M. Atomic structure and hierarchical assembly of a cross- $\beta$  amyloid fibril. *Proc. Natl. Acad. Sci.*, 2013, **10**, pp.5468-5473.
50. Kopito, R.R. Aggresomes, inclusion bodies and protein aggregation. *Trends Cell Biol.*, 2000, **10**(12), pp.524-530.
51. Morris, A.M., Watzky, M.A. and Finke, R.G. Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literature. *BBA - Prot. Proteo.*, 2009, **1794**(3), pp.375-397.
52. Dobson, C.M. Protein folding and misfolding. *Nature*, 2003, **426**(6968), pp.884-890.
53. Uversky, V.N. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.*, 2002, **11**(4), pp.739-756.
54. Iidicula-Thomas, S. and Balaji, P.V. Understanding the relationship between the primary structure of proteins and their amyloidogenic propensity: clues from inclusion body formation. *Protein Eng. Des. Sel.*, 2005, **18**(4), pp.175-180.
55. Merlini, G. and Bellotti, V. Lysozyme: a paradigmatic molecule for the investigation of protein structure, function and misfolding. *Clin. Chim. Acta.*, 2005, **357**(2), pp.168-172.
56. Hammarstrom, P., Jiang, X., Hurshman, A.R., Powers, E.T. and Kelly, J.W. Sequence-dependent denaturation energetics: a major determinant in amyloid disease diversity. *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**(Suppl 4), pp.16427-16432.
57. Platt, G.W. and Radford, S.E. Glimpses of the molecular mechanisms of  $\beta$ 2-microglobulin fibril formation *in vitro*: aggregation on a complex energy landscape. *FEBS Lett.*, 2009, **583**(16), pp.2623-2629.
58. Jahn, T.R., Parker, M.J., Homans, S.W. and Radford, S.E. Amyloid formation under physiological conditions proceeds via a native-like folding intermediate. *Nat. Struct. Mol. Biol.*, 2006, **13**(3), pp.195-201.
59. Balch, W.E., Morimoto, R.I., Dillin, A. and Kelly, J.W. Adapting proteostasis for disease intervention. *Science*, 2008, **319**(5865), pp.916-919.
60. Goldberg, A.L. Protein degradation and protection against misfolded or damaged proteins. *Nature*, 2003, **426**(6968), pp.895-899.
61. Ellgaard, L. and Helenius, A. Quality control in the endoplasmic reticulum. *Nat. Rev. Mol. Cell Biol.*, 2003, **4**(3), pp.181-191.
62. Meusser, B., Hirsch, C., Jarosch, E. and Sommer, T. ERAD: the long road to destruction. *Nat. Cell Biol.*, 2005, **7**(8), pp.766-772.
63. Rubinsztein, D.C. The roles of intracellular protein-degradation pathways in neurodegeneration. *Nature*, 2006, **443**(7113), pp.780-786.
64. Pincus, D. and Walter, P. A first line of defense against ER stress. *J. Cell Biol.*, 2012, **198**(3), pp.277-279.
65. Bukau, B., Weissman, J. and Horwich, A. Molecular chaperones and protein quality control. *Cell*, 2006, **125**(3), pp.443-451.

66. Richter, K., Haslbeck, M. and Buchner, J. The heat shock response: life on the verge of death. *Mol. Cell*, **40**(2), pp.253-266.
67. Courgeon, A.-M., Maisonhaute, C. and Best-Belpomme, M. Heat shock proteins are induced by cadmium in *Drosophila* cells. *Exp. Cell Res.*, 1984, **153**(2), pp.515-521.
68. Heikkila, J.J., Schultz, G.A., Iatrou, K. and Gedamu, L. Expression of a set of fish genes following heat or metal ion exposure. *J. Biol. Chem.*, 1982, **257**(20), pp.12000-12005.
69. Yura, T., Tobe, T., Ito, K. and Osawa, T. Heat shock regulatory gene (*htpR*) of *Escherichia coli* is required for growth at high temperature but is dispensable at low temperature. *Proc. Natl. Acad. Sci.*, 1984, **81**(21), pp.6803-6807.
70. Rothman, James e. and Schekman, R. Molecular mechanism of protein folding in the cell. *Cell*, **146**(6), pp.851-854.
71. Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M.K. and Hartl, F.U. Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature*, 1992, **356**(6371), pp.683-689.
72. Frydman, J., Nimmesgern, E., Ohtsuka, K. and Hartl, F.U. Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature*, 1994, **370**(6485), pp.111-117.
73. Horwich, A.L., Farr, G.W. and Fenton, W.A. GroEL–GroES-mediated protein folding. *Chem. Rev.*, 2006, **106**(5), pp.1917-1930.
74. Braig, K., Simon, M., Furuya, F., Hainfeld, J.F. and Horwich, A.L. A polypeptide bound by the chaperonin GroEL is localized within a central cavity. *Proc. Natl. Acad. Sci.*, 1993, **90**(9), pp.3978-3982.
75. Fenton, W.A., Kashi, Y., Furtak, K. and Norwich, A.L. Residues in chaperonin GroEL required for polypeptide binding and release. *Nature*, 1994, **371**(6498), pp.614-619.
76. Dunn, A.Y., Melville, M.W. and Frydman, J. Review: cellular substrates of the eukaryotic chaperonin TRiC/CCT. *J. Struct. Biol.*, 2001, **135**(2), pp.176-184.
77. Vabulas, R.M., Raychaudhuri, S., Hayer-Hartl, M. and Hartl, F.-U. Protein folding in the cytoplasm and the heat shock response. *Cold Spring Harb. Perspect. Biol.*, 2010, **2**(12), pp.a004390-a004390.
78. Schubert, U., Antón, L.C., Gibbs, J., Norbury, C.C., Yewdell, J.W. and Bennink, J.R. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature*, 2000, **404**(6779), pp.770-774.
79. Cohen, F.E. and Kelly, J.W. Therapeutic approaches to protein-misfolding diseases. *Nature*, 2003, **426**(6968), pp.905-909.
80. Braakman, I. and Bulleid, N.J. Protein folding and modification in the mammalian endoplasmic reticulum. *Annu. Rev. Biochem.*, 2011, **80**(1), pp.71-99.
81. Liu, Y., Gotte, G., Libonati, M. and Eisenberg, D. A domain-swapped RNase A dimer with implications for amyloid formation. *Nat. Struct. Mol. Biol.*, 2001, **8**(3), pp.211-214.
82. Betts, S. and King, J. There's a right way and a wrong way: *in vivo* and *in vitro* folding, misfolding and subunit assembly of the P22 tailspike. *Structure*, 1999, **7**(6), pp.R131-R139.
83. Woehlbier, U. and Hetz, C. Modulating stress responses by the UPRosome: a matter of life and death. *Trends Biochem. Sci.*, 2011, **36**(6), pp.329-337.
84. Tyedmers, J., Mogk, A. and Bukau, B. Cellular strategies for controlling protein aggregation. *Nat. Rev. Mol. Cell. Biol.*, 2010, **11**(11), pp.777-788.
85. Arrasate, M., Mitra, S., Schweitzer, E.S., Segal, M.R. and Finkbeiner, S. Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature*, 2004, **431**(7010), pp.805-810.
86. Tanaka, M., Kim, Y.M., Lee, G., Junn, E., Iwatsubo, T. and Mouradian, M.M. Aggresomes formed by  $\alpha$ -synuclein and synphilin-1 are cytoprotective. *J. Biol. Chem.*, 2004, **279**(6), pp.4625-4631.
87. Douglas, P.M., Treusch, S., Ren, H.-Y., Halfmann, R., Duennwald, M.L., Lindquist, S. and Cyr, D.M. Chaperone-dependent amyloid assembly protects cells from prion toxicity. *Proc. Natl. Acad. Sci.*, 2008, **105**(20), pp.7206-7211.

88. Cohen, E., Bieschke, J., Perciavalle, R.M., Kelly, J.W. and Dillin, A. Opposing activities protect against age-onset proteotoxicity. *Science*, 2006, **313**(5793), pp.1604-1610.
89. Saudou, F., Finkbeiner, S., Devys, D. and Greenberg, M.E. Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. *Cell*, 1998, **95**(1), pp.55-66.
90. Taylor, J.P., Tanaka, F., Robitschek, J., Sandoval, C.M., Taye, A., Markovic-Plese, S. and Fischbeck, K.H. Aggresomes protect cells by enhancing the degradation of toxic polyglutamine-containing protein. *Hum. Mol. Genet.*, 2003, **12**(7), pp.749-757.
91. Kaganovich, D., Kopito, R. and Frydman, J. Misfolded proteins partition between two distinct quality control compartments. *Nature*, 2008, **454**(7208), pp.1088-1095.
92. Lindner, A.B., Madden, R., Demarez, A., Stewart, E.J. and Taddei, F. Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc. Natl. Acad. Sci.*, 2008, **105**(8), pp.3076-3081.
93. Liu, B., Larsson, L., Franssens, V., Hao, X., Hill, Sandra m., Andersson, V., Höglund, D., Song, J., Yang, X., Öling, D., Grantham, J., Winderickx, J. and Nyström, T. Segregation of protein aggregates involves actin and the polarity machinery. *Cell*, **147**(5), pp.959-961.
94. Winkler, J., Seybert, A., König, L., Pruggnaller, S., Haselmann, U., Sourjik, V., Weiss, M., Frangakis, A.S., Mogk, A. and Bukau, B. Quantitative and spatio-temporal features of protein aggregation in *Escherichia coli* and consequences on protein quality control and cellular ageing. *EMBO J.*, 2010, **29**(5), pp.910-923.
95. Rujano, M.A., Bosveld, F., Salomons, F.A., Dijk, F., Van Waarde, M.a.W.H., Van Der Want, J.J.L., De Vos, R.a.I., Brunt, E.R., Sibon, O.C.M. and Kampinga, H.H. Polarised asymmetric inheritance of accumulated protein damage in higher eukaryotes. *PLoS Biol.*, 2006, **4**(12), p.e417.
96. Aguilaniu, H., Gustafsson, L., Rigoulet, M. and Nyström, T. Asymmetric inheritance of oxidatively damaged proteins during cytokinesis. *Science*, 2003, **299**(5613), pp.1751-1753.
97. Gragerov, A.I., Martin, E.S., Krupenko, M.A., Kashlev, M.V. and Nikiforov, V.G. Protein aggregation and inclusion body formation in *Escherichia coli* rpoH mutant defective in heat shock protein induction. *FEBS Lett.*, 1991, **291**(2), pp.222-224.
98. Laskowska, E., Bohdanowicz, J., Kuczyńska-Wiśnik, D., Matuszewska, E., Kędzierska, S. and Taylor, A. Aggregation of heat-shock-denatured, endogenous proteins and distribution of the IbpA/B and Fda marker-proteins in *Escherichia coli* WT and grpE280 cells. *Microbiology*, 2004, **150**(1), pp.247-259.
99. Rokney, A., Shagan, M., Kessel, M., Smith, Y., Rosenshine, I. and Oppenheim, A.B. *E. coli* transports aggregated proteins to the poles by a specific and energy-dependent process. *J. Mol. Biol.*, 2009, **392**(3), pp.589-601.
100. Kirstein, J., Strahl, H., Molière, N., Hamoen, L.W. and Turgay, K. Localization of general and regulatory proteolysis in *Bacillus subtilis* cells. *Mol. Microbiol.*, 2008, **70**(3), pp.682-694.
101. García-Mata, R., Bebök, Z., Sorscher, E.J. and Sztul, E.S. Characterization and dynamics of aggresome formation by a cytosolic GFP-chimera. *J. Cell Biol.*, 1999, **146**(6), pp.1239-1254.
102. Glover, J.R. and Lindquist, S. Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. *Cell*, 1998, **94**(1), pp.73-82.
103. Kopito, R.R. and Sitia, R. Aggresomes and Russell bodies. *EMBO Rep.*, 2000, **1**(3), pp.225-231.
104. Haslberger, T., Weibezahn, J., Zahn, R., Lee, S., Tsai, F.T.F., Bukau, B. and Mogk, A. M domains couple the ClpB threading motor with the DnaK chaperone activity. *Mol. Cell*, 2007, **25**(2), pp.247-260.
105. Haslberger, T., Zdanowicz, A., Brand, I., Kirstein, J., Turgay, K., Mogk, A. and Bukau, B. Protein disaggregation by the AAA+ chaperone ClpB involves partial threading of looped polypeptide segments. *Nat. Struct. Mol. Biol.*, 2008, **15**(6), pp.641-650.
106. Weibezahn, J., Tessarz, P., Schlieker, C., Zahn, R., Maglica, Z., Lee, S., Zentgraf, H., Weber-Ban, E.U., Dougan, D.A., Tsai, F.T.F., Mogk, A. and Bukau, B. Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. *Cell*, **119**(5), pp.653-665.

107. Lum, R., Tkach, J.M., Vierling, E. and Glover, J.R. Evidence for an unfolding/threading mechanism for protein disaggregation by *Saccharomyces cerevisiae* Hsp104. *J. Biol. Chem.*, 2004, **279**(28), pp.29139-29146.
108. Lelouard, H., Gatti, E., Cappello, F., Gresser, O., Camosseto, V. and Pierre, P. Transient aggregation of ubiquitinated proteins during dendritic cell maturation. *Nature*, 2002, **417**(6885), pp.177-182.
109. Yamamoto, A., Lucas, J.J. and Hen, R. Reversal of neuropathology and motor dysfunction in a conditional model of Huntington's disease. *Cell*, **101**(1), pp.57-66.
110. Powers, E.T., Morimoto, R.I., Dillin, A., Kelly, J.W. and Balch, W.E. Biological and chemical approaches to diseases of proteostasis deficiency. *Annu. Rev. Biochem.*, 2009, **78**(1), pp.959-991.
111. Soto, C. Unfolding the role of protein misfolding in neurodegenerative diseases. *Nat. Rev. Neurosci.*, 2003, **4**(1), pp.49-60.
112. Cattaneo, E., Rigamonti, D., Goffredo, D., Zuccato, C., Squitieri, F. and Sipione, S. Loss of normal huntingtin function: new developments in Huntington's disease research. *Trends Neurosci.*, 2001, **24**(3), pp.182-188.
113. Reaume, A.G., Elliott, J.L., Hoffman, E.K., Kowall, N.W., Ferrante, R.J., Siwek, D.R., Wilcox, H.M., Flood, D.G., Beal, M.F., Brown, R.H., Scott, R.W. and Snider, W.D. Motor neurons in Cu/Zn superoxide dismutase-deficient mice develop normally but exhibit enhanced cell death after axonal injury. *Nat. Genet.*, 1996, **13**(1), pp.43-47.
114. Norrby, E. Prions and protein folding diseases. *J. Internal Med.*, 2011, **270**(1), pp.1-14.
115. Nomura, Y. Molecular approaches to the treatment, prophylaxis, and diagnosis of Alzheimer's disease. *J. Pharmacol. Sci.*, 2012, **118**(3), pp.317-318.
116. Jain, S., Wood, N. and Healy, D. Molecular genetic pathways in Parkinson's disease: a review. *Clin. Sci.*, 2005, **109**, pp.355-364.
117. Kiernan, M.C., Vucic, S., Cheah, B.C., Turner, M.R., Eisen, A., Hardiman, O., Burrell, J.R. and Zoing, M.C. Amyotrophic lateral sclerosis. *Lancet*, **377**(9769), pp.942-955.
118. Ross, C.A. and Tabrizi, S.J. Huntington's disease: from molecular pathogenesis to clinical treatment. *Lancet Neurol.*, 2011, **10**(1), pp.83-98.
119. Westermark, P., Andersson, A. and Westermark, G.T. Islet amyloid polypeptide, islet amyloid, and diabetes mellitus. *Physiol. Rev.*, 2011, **91**(3), pp.795-826.
120. Shikama, Y., Kitazawa, J.-I., Yagihashi, N., Uehara, O., Murata, Y., Yajima, N., Wada, R. and Yagihashi, S. Localized amyloidosis at the site of repeated insulin injection in a diabetic patient. *Intern. Med.*, 2010, **49**(5), pp.397-401.
121. Teoh, C., Griffin, M.W. and Howlett, G. Apolipoproteins and amyloid fibril formation in atherosclerosis. *Protein Cell*, 2011, **2**(2), pp.116-127.
122. Merlini, G., Comenzo, R.L., Seldin, D.C., Wechalekar, A. and Gertz, M.A. Immunoglobulin light chain amyloidosis. *Exp. Rev. Hematol.*, 2014, **7**(1), pp.143-156.
123. Pettersson, T., Kontinen, Y. and Maury, C. Treatment strategies for amyloid A amyloidosis. *Exp. Opin. Pharmacother.*, 2008, **9**(12), pp.2117-2128.
124. Smith, D., Ashcroft, A. and Radford, S. Hemodialysis-related amyloidosis. In: M. RAMIREZ-ALVARADO, J. KELLY and C. DOBSON, eds. *Protein misfolding diseases: current and emerging principles and therapies*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2010.
125. Granel, B., Valleix, S., Serratrice, J., Chérin, P., Texeira, A., Disdier, P., Weiller, P. and Grateau, G. Lysozyme amyloidosis: report of 4 cases and a review of the literature. *Medicine*, 2006, **85**(1), pp.66-73.
126. Bucciantini, M., Giannoni, E., Chiti, F., Baroni, F., Formigli, L., Zurdo, J., Taddei, N., Ramponi, G., Dobson, C.M. and Stefani, M. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*, 2002, **416**(6880), pp.507-511.
127. Yan, S.D., Zhu, H., Zhu, A., Golabek, A., Du, H., Roher, A., Yu, J., Soto, C., Schmidt, A.M., Stern, D. and Kindy, M. Receptor-dependent cell stress and amyloid accumulation in systemic amyloidosis. *Nat. Med.*, 2000, **6**(6), pp.643-651.

128. Engel, M.F.M. Membrane permeabilization by islet amyloid polypeptide. *Chem. Phys. Lipids*, 2009, **160**(1), pp.1-10.
129. Engel, M.F.M., Khemtémourian, L., Kleijer, C.C., Meeldijk, H.J.D., Jacobs, J., Verkleij, A.J., De Kruijff, B., Killian, J.A. and Höppener, J.W.M. Membrane damage by human islet amyloid polypeptide through fibril growth at the membrane. *Proc. Natl. Acad. Sci.*, 2008, **105**(16), pp.6033-6038.
130. Narayan, P., Ganzinger, K.A., Mccoll, J., Weimann, L., Meehan, S., Qamar, S., Carver, J.A., Wilson, M.R., St. George-Hyslop, P., Dobson, C.M. and Klenerman, D. Single molecule characterization of the interactions between amyloid- $\beta$  peptides and the membranes of hippocampal cells. *J. Am. Chem. Soc.*, 2013, **135**(4), pp.1491-1498.
131. Pieri, L., Madiona, K., Bousset, L. and Melki, R. Fibrillar  $\alpha$ -synuclein and huntingtin exon 1 assemblies are toxic to the cells. *Biophys. J.*, **102**(12), pp.2894-2905.
132. Cummings, C.J., Mancini, M.A., Antalfy, B., Defranco, D.B., Orr, H.T. and Zoghbi, H.Y. Chaperone suppression of aggregation and altered subcellular proteasome localization imply protein misfolding in SCA1. *Nat. Genet.*, 1998, **19**(2), pp.148-154.
133. Cook, C. and Petrucelli, L. A critical evaluation of the ubiquitin-proteasome system in Parkinson's disease. *BBA - Mol. Basis Dis.*, 2009, **1792**(7), pp.664-675.
134. Freeman, D., Cedillos, R., Choyke, S., Lukic, Z., Mcguire, K., Marvin, S., Burrage, A.M., Sudholt, S., Rana, A., O'connor, C., Wiethoff, C.M. and Campbell, E.M. Alpha-synuclein induces lysosomal rupture and cathepsin dependent reactive oxygen species following endocytosis. *PLoS ONE*, 2013, **8**(4), p.e62143.
135. Lee, J.-H., Hong, C.-S., Lee, S., Yang, J.-E., Park, Y.I., Lee, D., Hyeon, T., Jung, S. and Paik, S.R. Radiating amyloid fibril formation on the surface of lipid membranes through unit-assembly of oligomeric species of  $\alpha$ -synuclein. *PLoS ONE*, 2012, **7**(10), p.e47580.
136. Soto, C., Kindy, M.S., Baumann, M. and Frangione, B. Inhibition of Alzheimer's amyloidosis by peptides that prevent  $\beta$ -sheet conformation. *Biochem. Biophys. Res. Commun.*, 1996, **226**(3), pp.672-680.
137. Dobson, C.M. Protein misfolding, evolution and disease. *Trends Biochem. Sci.*, 1999, **24**(9), pp.329-332.
138. Janson, J., Ashley, R.H., Harrison, D., McIntyre, S. and Butler, P.C. The mechanism of islet amyloid polypeptide toxicity is membrane disruption by intermediate-sized toxic amyloid particles. *Diabetes*, 1999, **48**(3), pp.491-498.
139. Koo, E.H., Lansbury, P.T. and Kelly, J.W. Amyloid diseases: abnormal protein aggregation in neurodegeneration. *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**(18), pp.9989-9990.
140. Conway, K.A., Lee, S.-J., Rochet, J.-C., Ding, T.T., Williamson, R.E. and Lansbury, P.T. Acceleration of oligomerization, not fibrillization, is a shared property of both  $\alpha$ -synuclein mutations linked to early-onset Parkinson's disease: Implications for pathogenesis and therapy. *Proc. Natl. Acad. Sci.*, 2000, **97**(2), pp.571-576.
141. Zhu, Y.J., Lin, H. and Lal, R. Fresh and nonfibrillar amyloid  $\beta$  protein(1-40) induces rapid cellular degeneration in aged human fibroblasts: evidence for A $\beta$ P-channel-mediated cellular toxicity. *FASEB J.*, 2000, **14**(9), pp.1244-1254.
142. Nilsberth, C., Westlind-Danielsson, A., Eckman, C.B., Condron, M.M., Axelman, K., Forsell, C., Stenh, C., Luthman, J., Teplow, D.B., Younkin, S.G., Naslund, J. and Lannfelt, L. The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced A $\beta$  protofibril formation. *Nat. Neurosci.*, 2001, **4**(9), pp.887-893.
143. Mendes Sousa, M., Cardoso, I., Fernandes, R., Guimarães, A. and Saraiva, M.J. Deposition of transthyretin in early stages of familial amyloidotic polyneuropathy: evidence for toxicity of nonfibrillar aggregates. *Am. J. Pathol.*, 2001, **159**(6), pp.1993-2000.
144. Lin, H., Bhatia, R. and Lal, R. Amyloid  $\beta$  protein forms ion channels: implications for Alzheimer's disease pathophysiology. *FASEB J.*, 2001, **15**(13), pp.2433-2444.
145. Lin, M.-C., Mirzabekov, T. and Kagan, B.L. Channel formation by a neurotoxic prion protein fragment. *J. Biol. Chem.*, 1997, **272**(1), pp.44-47.

146. Soua, V., Parker, M.S., Williams, T.L., Ratnayaka, A., Atherton, J., Gorringer, K., Tuffin, J., Darwent, E., Rambaran, R., Klein, W., Lacor, P., Staras, K., Thorpe, J. and Serpell, L.C. Visualization of co-localization in A $\beta$ 42-administered neuroblastoma cells reveals lysosome damage and autophagosome accumulation related to cell death. *Biochem. J.*, 2012, **441**(2), pp.579-590.
147. Behl, C., Davis, J.B., Lesley, R. and Schubert, D. Hydrogen peroxide mediates amyloid  $\beta$  protein toxicity. *Cell*, **77**(6), pp.817-827.
148. Hsu, L.J., Sagara, Y., Arroyo, A., Rockenstein, E., Sisk, A., Mallory, M., Wong, J., Takenouchi, T., Hashimoto, M. and Masliah, E.  $\alpha$ -Synuclein promotes mitochondrial deficit and oxidative stress. *Am. J. Pathol.*, 2000, **157**(2), pp.401-410.
149. Olzscha, H., Schermann, S.M., Woerner, A.C., Pinkert, S., Hecht, M.H., Tartaglia, G.G., Vendruscolo, M., Hayer-Hartl, M., Hartl, F.U. and Vabulas, R.M. Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell*, 2011, **144**(1), pp.67-78.
150. Cheon, M., Chang, I., Mohanty, S., Luheshi, L.M., Dobson, C.M., Vendruscolo, M. and Favrin, G. Structural reorganisation and potential toxicity of oligomeric species formed during the assembly of amyloid fibrils. *PLoS Comput. Biol.*, 2007, **3**(9), p.e173.
151. Kelly, J.W. and Balch, W.E. Amyloid as a natural product. *J. Cell Biol.*, 2003, **161**(3), pp.461-462.
152. Otzen, D. Functional amyloid: turning swords into plowshares. *Prion*, 2010, **4**(4), pp.256-264.
153. Hammer, N.D., Schmidt, J.C. and Chapman, M.R. The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. *Proc. Natl. Acad. Sci.*, 2007, **104**(30), pp.12494-12499.
154. Elliot, M.A., Karoonuthaisiri, N., Huang, J., Bibb, M.J., Cohen, S.N., Kao, C.M. and Buttner, M.J. The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in *Streptomyces coelicolor*. *Genes Develop.*, 2003, **17**(14), pp.1727-1740.
155. Newby, G.A. and Lindquist, S. Blessings in disguise: biological benefits of prion-like mechanisms. *Trends Cell Biol.*, **23**(6), pp.251-259.
156. Saupe, S.J. and Daskalov, A. The [Het-s] prion, an amyloid fold as a cell death activation trigger. *PLoS Pathog.*, 2012, **8**(5), p.e1002687.
157. Kenney, J.M., Knight, D., Wise, M.J. and Vollrath, F. Amyloidogenic nature of spider silk. *Eur. J. Biochem.*, 2002, **269**(16), pp.4159-4163.
158. Hamodrakas, S.J., Hoenger, A. and Iconomidou, V.A. Amyloid fibrillogenesis of silkworm chorion protein peptide-analogues via a liquid-crystalline intermediate phase. *J. Struct. Biol.*, 2004, **145**(3), pp.226-235.
159. Borman, S.T.U. Mammalian amyloid has a useful role. *Chem. Eng. News. Arch.*, 2005, **83**(49), p.13.
160. Hamley, I.W. The amyloid beta peptide: a chemist's perspective. Role in Alzheimer's and fibrillization. *Chem. Rev.*, 2012, **112**(10), pp.5147-5192.
161. Miller, G. Stopping Alzheimer's before it starts. *Science*, 2012, **337**(6096), pp.790-792.
162. Selkoe, D.J. Preventing Alzheimer's disease. *Science*, 2012, **337**(6101), pp.1488-1492.
163. Bateman, R.J., Xiong, C., Benzinger, T.L.S., Fagan, A.M., Goate, A., Fox, N.C., Marcus, D.S., Cairns, N.J., Xie, X., Blazey, T.M., Holtzman, D.M., Santacruz, A., Buckles, V., Oliver, A., Moulder, K., Aisen, P.S., Ghetti, B., Klunk, W.E., Mcdade, E., Martins, R.N., Masters, C.L., Mayeux, R., Ringman, J.M., Rossor, M.N., Schofield, P.R., Sperling, R.A., Salloway, S. and Morris, J.C. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N. Engl. J. Med.*, 2012, **367**(9), pp.795-804.
164. Mckay, R. Stem cells in the central nervous system. *Science*, 1997, **276**(5309), pp.66-71.
165. Alvarez-Buylla, A., Herrera, D.G. and Wichterle, H. The subventricular zone: source of neuronal precursors for brain repair. *Prog. Brain Res.*, 2000, **127**, pp.1-11.
166. Gage, F.H. Mammalian neural stem cells. *Science*, 2000, **287**(5457), pp.1433-1438.

167. Park, K.I., Teng, Y.D. and Snyder, E.Y. The injured brain interacts reciprocally with neural stem cells supported by scaffolds to reconstitute lost tissue. *Nat. Biotechnol.*, 2002, **20**(11), pp.1111-1117.
168. Galimberti, D. and Scarpini, E. Inflammation and oxidative damage in Alzheimer's disease: friend or foe? *Schol. Ed.*, 2011, **3**, pp.252-266.
169. Perry, V.H., Nicoll, J.a.R. and Holmes, C. Microglia in neurodegenerative disease. *Nat. Rev. Neurol.*, 2010, **6**(4), pp.193-201.
170. Huang, Y. and Mucke, L. Alzheimer mechanisms and therapeutic strategies. *Cell*, 2012, **148**(6), pp.1204-1222.
171. Lee, H.P., Zhu, X., Casadesus, G., Castellani, R.J., Nunomura, A., Smith, M.A., Lee, H.-G. and Perry, G. Antioxidant approaches for the treatment of Alzheimer's disease. *Expert Rev. Neurother.*, 2010, **10**(7), pp.1201-1208.
172. Cole, G. and Frautschy, S. Mechanisms of action of non-steroidal anti-inflammatory drugs for the prevention of Alzheimer's disease. *CNS Neurol. Disord. Drug Targets*, 2010, **9**, pp.140-148.
173. Bezprozvanny, I. The rise and fall of Dimebon. *Drug News Perspect.*, 2010, **23**, pp.518-523.
174. Mallucci, G. and Collinge, J. Rational targeting for prion therapeutics. *Nat. Rev. Neurosci.*, 2005, **6**(1), pp.23-34.
175. Blömer, U., Naldini, L., Kafri, T., Trono, D., Verma, I.M. and Gage, F.H. Highly efficient and sustained gene transfer in adult neurons with a lentivirus vector. *J. Virol.*, 1997, **71**(9), pp.6641-9.
176. Rubinson, D.A., Dillon, C.P., Kwiatkowski, A.V., Sievers, C., Yang, L., Kopinja, J., Rooney, D.L., Zhang, M., Ihrig, M.M., Mcmanus, M.T., Gertler, F.B., Scott, M.L. and Van Parijs, L. A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. *Nat. Genet.*, 2003, **33**(3), pp.401-406.
177. Davidson, B.L. and Breakefield, X.O. Viral vectors for gene delivery to the nervous system. *Nat. Rev. Neurosci.*, 2003, **4**(5), pp.353-364.
178. Xia, H., Mao, Q., Paulson, H.L. and Davidson, B.L. siRNA-mediated gene silencing *in vitro* and *in vivo*. *Nat. Biotechnol.*, 2002, **20**(10), pp.1006-1010.
179. Caplen, N.J., Taylor, J.P., Statham, V.S., Tanaka, F., Fire, A. and Morgan, R.A. Rescue of polyglutamine-mediated cytotoxicity by double-stranded RNA-mediated RNA interference. *Hum. Mol. Genet.*, 2002, **11**(2), pp.175-184.
180. Davidson, B.L. and Paulson, H.L. Molecular medicine for the brain: silencing of disease genes with RNA interference. *Lancet Neurol.*, 2004, **3**(3), pp.145-149.
181. Daude, N., Marella, M. and Chabry, J. Specific inhibition of pathological prion protein accumulation by small interfering RNAs. *J. Cell Sci.*, 2003, **116**(13), pp.2775-2779.
182. Miller, V.M., Xia, H., Marrs, G.L., Gouvion, C.M., Lee, G., Davidson, B.L. and Paulson, H.L. Allele-specific silencing of dominant disease genes. *Proc. Natl. Acad. Sci.*, 2003, **100**(12), pp.7195-7200.
183. Sacchettini, J.C. and Kelly, J.W. Therapeutic strategies for human amyloid diseases. *Nat. Rev. Drug Discov.*, 2002, **1**(4), pp.267-275.
184. Levine Iii, H. and Scholten, J.D. Screening for pharmacologic inhibitors of amyloid fibril formation. In: W. RONALD, ed. *Methods in Enzymology*. Academic Press, 1999, pp.467-476.
185. Weggen, S., Eriksen, J.L., Das, P., Sagi, S.A., Wang, R., Pietrzik, C.U., Findlay, K.A., Smith, T.E., Murphy, M.P., Bulter, T., Kang, D.E., Marquez-Sterling, N., Golde, T.E. and Koo, E.H. A subset of NSAIDs lower amyloidogenic A $\beta$ 42 independently of cyclooxygenase activity. *Nature*, 2001, **414**(6860), pp.212-216.
186. Golde, T.E., Schneider, L.S. and Koo, E.H. Anti-A $\beta$  therapeutics in Alzheimer's disease: the need for a paradigm shift. *Neuron*, 2011, **69**(2), pp.203-213.
187. Mason, J.M., Kokkoni, N., Stott, K. and Doig, A.J. Design strategies for anti-amyloid agents. *Curr. Opin. Struct. Biol.*, 2003, **13**(4), pp.526-532.

188. Hammarström, P., Jiang, X., Hurshman, A.R., Powers, E.T. and Kelly, J.W. Sequence-dependent denaturation energetics: a major determinant in amyloid disease diversity. *Proc. Natl. Acad. Sci.*, 2002, **99**(suppl 4), pp.16427-16432.
189. Johnson, S.M., Connelly, S., Fearn, C., Powers, E.T. and Kelly, J.W. The transthyretin amyloidoses: from delineating the molecular mechanism of aggregation linked to pathology to a regulatory-agency-approved drug. *J. Mol. Biol.*, 2012, **421**(2–3), pp.185-203.
190. Razavi, H., Palaninathan, S.K., Powers, E.T., Wiseman, R.L., Purkey, H.E., Mohamedmohaideen, N.N., Deechongkit, S., Chiang, K.P., Dendle, M.T.A., Sacchettini, J.C. and Kelly, J.W. Benzoxazoles as transthyretin amyloid fibril inhibitors: synthesis, evaluation, and mechanism of action. *Angew. Chem. Int. Ed.*, 2003, **42**(24), pp.2758-2761.
191. Hammarström, P., Wiseman, R.L., Powers, E.T. and Kelly, J.W. Prevention of transthyretin amyloid disease by changing protein misfolding energetics. *Science*, 2003, **299**(5607), pp.713-716.
192. Dumoulin, M., Last, A.M., Desmyter, A., Decanniere, K., Canet, D., Larsson, G., Spencer, A., Archer, D.B., Sasse, J., Muyldermans, S., Wyns, L., Redfield, C., Matagne, A., Robinson, C.V. and Dobson, C.M. A camelid antibody fragment inhibits the formation of amyloid fibrils by human lysozyme. *Nature*, 2003, **424**(6950), pp.783-788.
193. Tatzelt, J., Prusiner, S.B. and Welch, W.J. Chemical chaperones interfere with the formation of scrapie prion protein. *EMBO J.*, 1996, **15**(23), pp.6363-6373.
194. Salomon, A.R., Marcinowski, K.J., Friedland, R.P. and Zagorski, M.G. Nicotine inhibits amyloid formation by the  $\beta$ -peptide. *Biochem.*, 1996, **35**(42), pp.13568-13578.
195. Burrows, J.a.J., Willis, L.K. and Perlmutter, D.H. Chemical chaperones mediate increased secretion of mutant  $\alpha$ 1-antitrypsin ( $\alpha$ 1-AT) Z: a potential pharmacological strategy for prevention of liver injury and emphysema in  $\alpha$ 1-AT deficiency. *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**(4), pp.1796-1801.
196. Winkhofer, K.F., Reintjes, A., Hoener, M.C., Voellmy, R. and Tatzelt, J. Geldanamycin restores a defective heat shock response *in vivo*. *J. Biol. Chem.*, 2001, **276**(48), pp.45160-45167.
197. Jones, D. Modifying protein misfolding. *Nat. Rev. Drug Discov.*, 2010, **9**(11), pp.825-827.
198. Medina, D.X., Caccamo, A. and Oddo, S. Methylene blue reduces A $\beta$  levels and rescues early cognitive deficit by increasing proteasome activity. *Brain Pathol.*, 2011, **21**(2), pp.140-149.
199. Kisilevsky, R., Lemieux, L.J., Fraser, P.E., Kong, X., Hultin, P.G. and Szarek, W.A. Arresting amyloidosis *in vivo* using small-molecule anionic sulphonates or sulphates: implications for Alzheimer's disease. *Nat. Med.*, 1995, **1**(2), pp.143-148.
200. Bieschke, J., Herbst, M., Wiglenda, T., Friedrich, R.P., Boeddrich, A., Schiele, F., Kleckers, D., Lopez Del Amo, J.M., Grüning, B.A., Wang, Q., Schmidt, M.R., Lurz, R., Anwyl, R., Schnoegl, S., Fändrich, M., Frank, R.F., Reif, B., Günther, S., Walsh, D.M. and Wanker, E.E. Small-molecule conversion of toxic oligomers to nontoxic  $\beta$ -sheet-rich amyloid fibrils. *Nat. Chem. Biol.*, 2012, **8**(1), pp.93-101.
201. Pallitto, M.M., Ghanta, J., Heinzelman, P., Kiessling, L.L. and Murphy, R.M. Recognition sequence design for peptidyl modulators of  $\beta$ -amyloid aggregation and toxicity. *Biochem.*, 1999, **38**(12), pp.3570-3578.
202. Arosio, P., Vendruscolo, M., Dobson, C.M. and Knowles, T.P.J. Chemical kinetics for drug discovery to combat protein aggregation diseases. *Trends. Pharmacol. Sci.*, 2014, **35**(3), pp.127-135.
203. Cohen, S.I.A., Linse, S., Luheshi, L.M., Hellstrand, E., White, D.A., Rajah, L., Otzen, D.E., Vendruscolo, M., Dobson, C.M. and Knowles, T.P.J. Proliferation of amyloid- $\beta$ <sub>42</sub> aggregates occurs through a secondary nucleation mechanism. *Proc. Natl. Acad. Sci.*, 2013, **110**(24), pp.9758-9763.
204. Findeis, M.A., Musso, G.M., Arico-Muendel, C.C., Benjamin, H.W., Hundal, A.M., Lee, J.-J., Chin, J., Kelley, M., Wakefield, J., Hayward, N.J. and Molineaux, S.M. Modified-peptide inhibitors of amyloid  $\beta$ -peptide polymerization. *Biochem.*, 1999, **38**(21), pp.6791-6800.



205. Hughes, E., Burke, R.M. and Doig, A.J. Inhibition of toxicity in the  $\beta$ -amyloid peptide fragment  $\beta$ -(25–35) using *N*-methylated derivatives: a general strategy to prevent amyloid formation. *J. Biol. Chem.*, 2000, **275**(33), pp.25109-25115.
206. Gordon, D.J., Sciarretta, K.L. and Meredith, S.C. Inhibition of  $\beta$ -amyloid(40) fibrillogenesis and disassembly of  $\beta$ -amyloid(40) fibrils by short  $\beta$ -amyloid congeners containing *N*-methyl amino acids at alternate residues. *Biochem.*, 2001, **40**(28), pp.8237-8245.
207. Poduslo, J.F., Curran, G.L., Kumar, A., Frangione, B. and Soto, C.  $\beta$ -sheet breaker peptide inhibitor of Alzheimer's amyloidogenesis with increased blood–brain barrier permeability and resistance to proteolytic degradation in plasma. *J. Neurobiol.*, 1999, **39**(3), pp.371-382.
208. Chabry, J., Caughey, B. and Chesebro, B. Specific inhibition of *in vitro* formation of protease-resistant prion protein by synthetic peptides. *J. Biol. Chem.*, 1998, **273**(21), pp.13203-13207.
209. El-Agnaf, O.M.A., Paleologou, K.E., Greer, B., Abogreïn, A.M., King, J.E., Salem, S.A., Fullwood, N.J., Benson, F.E., Hewitt, R., Ford, K.J., Martin, F.L., Harriott, P., Cookson, M.R. and Allsop, D. A strategy for designing inhibitors of  $\alpha$ -synuclein aggregation and toxicity as a novel treatment for Parkinson's disease and related disorders. *FASEB J.*, 2004, **18**(11), pp.1315-1317.
210. Scrocchi, L.A., Chen, Y., Waschuk, S., Wang, F., Cheung, S., Darabie, A.A., McLaurin, J. and Fraser, P.E. Design of peptide-based inhibitors of human islet amyloid polypeptide fibrillogenesis. *J. Mol. Biol.*, 2002, **318**(3), pp.697-706.
211. Nagai, Y., Tucker, T., Ren, H., Kenan, D.J., Henderson, B.S., Keene, J.D., Strittmatter, W.J. and Burke, J.R. Inhibition of polyglutamine protein aggregation and cell death by novel peptides identified by phage display screening. *J. Biol. Chem.*, 2000, **275**(14), pp.10437-10442.
212. Sievers, S.A., Karanicolas, J., Chang, H.W., Zhao, A., Jiang, L., Zirafi, O., Stevens, J.T., Munch, J., Baker, D. and Eisenberg, D. Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature*, 2011, **475**(7354), pp.96-100.
213. Luheshi, L.M., Hoyer, W., De Barros, T.P., Van Dijk Härd, I., Brorsson, A.-C., Macao, B., Persson, C., Crowther, D.C., Lomas, D.A., Ståhl, S., Dobson, C.M. and Härd, T. Sequestration of the A $\beta$  peptide prevents toxicity and promotes degradation *in vivo*. *PLoS Biol.*, 2010, **8**(3), p.e1000334.
214. Cheng, B., Gong, H., Xiao, H., Petersen, R.B., Zheng, L. and Huang, K. Inhibiting toxic aggregation of amyloidogenic proteins: a therapeutic strategy for protein misfolding diseases. *BBA - Gen. Subjects*, 2013, **1830**(10), pp.4860-4871.
215. Perchiacca, J.M., Ladiwala, A.R.A., Bhattacharya, M. and Tessier, P.M. Structure-based design of conformation- and sequence-specific antibodies against amyloid  $\beta$ . *Proc. Natl. Acad. Sci.*, 2011.
216. Solomon, B., Koppel, R., Hanan, E. and Katzav, T. Monoclonal antibodies inhibit *in vitro* fibrillar aggregation of the Alzheimer beta-amyloid peptide. *Proc. Natl. Acad. Sci.*, 1996, **93**(1), pp.452-455.
217. Kaye, R., Head, E., Thompson, J.L., Mcintire, T.M., Milton, S.C., Cotman, C.W. and Glabe, C.G. Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science*, 2003, **300**(5618), pp.486-489.
218. Schenk, D., Barbour, R., Dunn, W., Gordon, G., Grajeda, H., Guido, T., Hu, K., Huang, J., Johnson-Wood, K., Khan, K., Kholodenko, D., Lee, M., Liao, Z., Lieberburg, I., Motter, R., Mutter, L., Soriano, F., Shopp, G., Vasquez, N., Vandeventer, C., Walker, S., Wogulis, M., Yednock, T., Games, D. and Seubert, P. Immunization with amyloid- $\beta$  attenuates Alzheimer-disease-like pathology in the PDAPP mouse. *Nature*, 1999, **400**(6740), pp.173-177.
219. Morgan, D., Diamond, D.M., Gottschall, P.E., Ugen, K.E., Dickey, C., Hardy, J., Duff, K., Jantzen, P., Dicarlo, G., Wilcock, D., Connor, K., Hatcher, J., Hope, C., Gordon, M. and Arendash, G.W. A $\beta$  peptide vaccination prevents memory loss in an animal model of Alzheimer's disease. *Nature*, 2000, **408**(6815), pp.982-985.
220. McLaurin, J., Kierstead, M.E., Brown, M.E., Hawkes, C.A., Lambermon, M.H.L., Phinney, A.L., Darabie, A.A., Cousins, J.E., French, J.E., Lan, M.F., Chen, F., Wong, S.S.N., Mount, H.T.J., Fraser, P.E., Westaway, D. and George-Hyslop, P.S. Cyclohexanehexol inhibitors of A $\beta$  aggregation prevent and reverse Alzheimer phenotype in a mouse model. *Nat. Med.*, 2006, **12**(7), pp.801-808.

221. Frautschy, S. and Cole, G. Why pleiotropic interventions are needed for Alzheimer's disease. *Mol. Neurobiol.*, 2010, **41**(2), pp.392-409.
222. Aisen, P.S. The development of anti-amyloid therapy for Alzheimer's disease: from secretase modulators to polymerisation inhibitors. *CNS Drugs*, 2005, **19**(12), pp.989-996.
223. Blazer, L.L. and Neubig, R.R. Small molecule protein-protein interaction inhibitors as CNS therapeutic agents: current progress and future hurdles. *Neuropsychopharmacology*, 2008, **34**(1), pp.126-141.
224. Masuda, M., Hasegawa, M., Nonaka, T., Oikawa, T., Yonetani, M., Yamaguchi, Y., Kato, K., Hisanaga, S.-I. and Goedert, M. Inhibition of  $\alpha$ -synuclein fibril assembly by small molecules. *FEBS Lett.*, 2009, **583**(4), pp.787-791.
225. Gestwicki, J.E., Crabtree, G.R. and Graef, I.A. Harnessing chaperones to generate small-molecule inhibitors of amyloid- $\beta$  aggregation. *Science*, 2004, **306**(5697), pp.865-869.
226. Mishra, R., Bulic, B., Sellin, D., Jha, S., Waldmann, H. and Winter, R. Small-molecule inhibitors of islet amyloid polypeptide fibril formation. *Angew. Chem. Int. Ed.*, 2008, **47**(25), pp.4679-4682.
227. Porat, Y., Abramowitz, A. and Gazit, E. Inhibition of amyloid fibril formation by polyphenols: structural similarity and aromatic interactions as a common inhibition mechanism. *Chem. Biol. Drug Des.*, 2006, **67**(1), pp.27-37.
228. Lorenzo, A. and Yankner, B.A. Beta-amyloid neurotoxicity requires fibril formation and is inhibited by Congo red. *Proc. Natl. Acad. Sci.*, 1994, **91**(25), pp.12243-12247.
229. Lee, V.M.Y. Amyloid binding ligands as Alzheimer's disease therapies. *Neurobiol. Aging*, 2002, **23**(6), pp.1039-1042.
230. Poli, G., Ponti, W., Carcassola, G., Ceciliani, F., Colombo, L., Dall'ara, P., Gervasoni, M., Giannino, M.L., Martino, P.A., Poller, C., Villa, S. and Salmona, M. *In vitro* evaluation of the anti-prionic activity of newly synthesized Congo red derivatives. *Arzneim.-Forsch.*, 2003, **53**(12), pp.875-888.
231. Wu, C., Scott, J. and Shea, J.-E. Binding of Congo red to amyloid protofibrils of the Alzheimer  $A\beta_{9-40}$  peptide probed by molecular dynamics simulations. *Biophys. J.*, 2012, **103**(3), pp.550-557.
232. Turnell, W.G. and Finch, J.T. Binding of the dye Congo red to the amyloid protein pig insulin reveals a novel homology amongst amyloid-forming peptide sequences. *J. Mol. Biol.*, 1992, **227**(4), pp.1205-1223.
233. Kim, Y.-S., Randolph, T.W., Manning, M.C., Stevens, F.J. and Carpenter, J.F. Congo red populates partially unfolded states of an amyloidogenic protein to enhance aggregation and amyloid fibril formation. *J. Biol. Chem.*, 2003, **278**(12), pp.10842-10850.
234. Palhano, F.L., Lee, J., Grimster, N.P. and Kelly, J.W. Toward the molecular mechanism(s) by which EGCG treatment remodels mature amyloid fibrils. *J. Am. Chem. Soc.*, 2013, **135**(20), pp.7503-7510.
235. Meng, F., Abedini, A., Plesner, A., Verchere, C.B. and Raleigh, D.P. The flavanol (-)-epigallocatechin 3-gallate inhibits amyloid formation by islet amyloid polypeptide, disaggregates amyloid fibrils, and protects cultured cells against IAPP-induced toxicity. *Biochem.*, 2010, **49**(37), pp.8127-8133.
236. Cao, P. and Raleigh, D.P. Analysis of the inhibition and remodeling of islet amyloid polypeptide amyloid fibers by flavanols. *Biochem.*, 2012, **51**(13), pp.2670-2683.
237. Young, L.M., Cao, P., Raleigh, D.P., Ashcroft, A.E. and Radford, S.E. Ion mobility spectrometry-mass spectrometry defines the oligomeric intermediates in amylin amyloid formation and the mode of action of inhibitors. *J. Am. Chem. Soc.*, 2013, **136**(2), pp.660-670.
238. Ferreira, N., Saraiva, M.J. and Almeida, M.R. Natural polyphenols inhibit different steps of the process of transthyretin (TTR) amyloid fibril formation. *FEBS Lett.*, 2011, **585**(15), pp.2424-2430.
239. Hudson, S.A., Ecroyd, H., Dehle, F.C., Musgrave, I.F. and Carver, J.A. (-)-Epigallocatechin-3-gallate (EGCG) maintains  $\kappa$ -casein in its pre-fibrillar state without redirecting its aggregation pathway. *J. Mol. Biol.*, 2009, **392**(3), pp.689-700.

240. Ehrnhoefer, D.E., Bieschke, J., Boeddrich, A., Herbst, M., Masino, L., Lurz, R., Engemann, S., Pastore, A. and Wanker, E.E. EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nat. Struct. Mol. Biol.*, 2008, **15**(6), pp.558-566.
241. Taniguchi, S., Suzuki, N., Masuda, M., Hisanaga, S.-I., Iwatsubo, T., Goedert, M. and Hasegawa, M. Inhibition of heparin-induced Tau filament formation by phenothiazines, polyphenols, and porphyrins. *J. Biol. Chem.*, 2005, **280**(9), pp.7614-7623.
242. Howlett, D., Cutler, P., Heales, S. and Camilleri, P. Hemin and related porphyrins inhibit  $\beta$ -amyloid aggregation. *FEBS Lett.*, 1997, **417**(2), pp.249-251.
243. Ono, K., Hasegawa, K., Naiki, H. and Yamada, M. Curcumin has potent anti-amyloidogenic effects for Alzheimer's  $\beta$ -amyloid fibrils *in vitro*. *J. Neurosci. Res.*, 2004, **75**(6), pp.742-750.
244. Jiang, T., Zhi, X.-L., Zhang, Y.-H., Pan, L.-F. and Zhou, P. Inhibitory effect of curcumin on the Al(III)-induced A $\beta$ <sub>42</sub> aggregation and neurotoxicity *in vitro*. *BBA - Mol. Basis Dis.*, 2012, **1822**(8), pp.1207-1215.
245. Daval, M., Bedrood, S., Gurlo, T., Huang, C.-J., Costes, S., Butler, P.C. and Langen, R. The effect of curcumin on human islet amyloid polypeptide misfolding and toxicity. *Amyloid*, 2010, **17**(3-4), pp.118-128.
246. Sparks, S., Liu, G., Robbins, K.J. and Lazo, N.D. Curcumin modulates the self-assembly of the islet amyloid polypeptide by disassembling  $\alpha$ -helix. *Biochem. Biophys. Res. Commun.*, 2012, **422**(4), pp.551-555.
247. Pandey, N., Strider, J., Nolan, W., Yan, S. and Galvin, J. Curcumin inhibits aggregation of  $\alpha$ -synuclein. *Acta Neuropathol.*, 2008, **115**(4), pp.479-489.
248. Borana, M.S., Mishra, P., Pissurlenkar, R.R.S., Hosur, R.V. and Ahmad, B. Curcumin and kaempferol prevent lysozyme fibril formation by modulating aggregation kinetic parameters. *BBA - Proteins Proteomics*, 2014, **1844**(3), pp.670-680.
249. Harris, J.R. *In vitro* fibrillogenesis of the amyloid  $\beta$ 1-42 peptide: cholesterol potentiation and aspirin inhibition. *Micron*, 2002, **33**(7-8), pp.609-626.
250. Jiang, P., Li, W., Shea, J.-E. and Mu, Y. Resveratrol inhibits the formation of multiple-layered  $\beta$ -sheet oligomers of the human islet amyloid polypeptide segment 22-27. *Biophys. J.*, 2011, **100**(6), pp.1550-1558.
251. Ahmad, E., Ahmad, A., Singh, S., Arshad, M., Khan, A.H. and Khan, R.H. A mechanistic approach for islet amyloid polypeptide aggregation to develop anti-amyloidogenic agents for type-2 diabetes. *Biochimie*, 2011, **93**(5), pp.793-805.
252. Jiang, L., Liu, C., Leibly, D., Landau, M., Zhao, M., Hughes, M.P., Eisenberg, D.S. and Kuriyan, J. Structure-based discovery of fiber-binding compounds that reduce the cytotoxicity of amyloid beta. *eLife*, 2013, **2**, p.e00857.
253. Landau, M., Sawaya, M.R., Faull, K.F., Laganowsky, A., Jiang, L., Sievers, S.A., Liu, J., Barrio, J.R. and Eisenberg, D. Towards a pharmacophore for amyloid. *PLoS Biol.*, 2011, **9**(6), p.e1001080.
254. Armstrong, A.H., Chen, J., Mckoy, A.F. and Hecht, M.H. Mutations that replace aromatic side chains promote aggregation of the Alzheimer's A $\beta$  peptide. *Biochem.*, 2011, **50**(19), pp.4058-4067.
255. Bieschke, J., Russ, J., Friedrich, R.P., Ehrnhoefer, D.E., Wobst, H., Neugebauer, K. and Wanker, E.E. EGCG remodels mature  $\alpha$ -synuclein and amyloid- $\beta$  fibrils and reduces cellular toxicity. *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**(17), pp.7710-7715.
256. Ferreira, N., Saraiva, M.J. and Almeida, M.R. Epigallocatechin-3-gallate as a potential therapeutic drug for TTR-related amyloidosis: "*in vivo*" evidence from FAP mice models. *PLoS ONE*, 2012, **7**(1), p.e29933.
257. Choi, Y.-T., Jung, C.-H., Lee, S.-R., Bae, J.-H., Baek, W.-K., Suh, M.-H., Park, J., Park, C.-W. and Suh, S.-I. The green tea polyphenol (-)-epigallocatechin gallate attenuates  $\beta$ -amyloid-induced neurotoxicity in cultured hippocampal neurons. *Life Sci.*, 2001, **70**(5), pp.603-614.
258. Cegelski, L., Pinkner, J.S., Hammer, N.D., Cusumano, C.K., Hung, C.S., Chorell, E., Å...Berg, V., Walker, J.N., Seed, P.C., Almqvist, F., Chapman, M.R. and Hultgren, S.J. Small-molecule inhibitors target *Escherichia coli* amyloid biogenesis and biofilm formation. *Nat. Chem. Biol.*, 2009, **5**(12), pp.913-919.

259. Arkin, M.R. and Wells, J.A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.*, 2004, **3**(4), pp.301-317.
260. Bogan, A.A. and Thorn, K.S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.*, 1998, **280**(1), pp.1-9.
261. Wells, J.A. Systematic mutational analyses of protein-protein interfaces. In: J.L. JOHN, ed. *Methods in Enzymology*. Academic Press, 1991, pp.390-411.
262. Darnell, S.J., Legault, L. and Mitchell, J.C. KFC Server: interactive forecasting of protein interaction hot spots. *Nucleic Acids Res.*, 2008, **36**(suppl 2), pp.W265-W269.
263. Shulman-Peleg, A., Shatsky, M., Nussinov, R. and Wolfson, H.J. MultiBind and MAPPIS: webservers for multiple alignment of protein 3D-binding sites and their interactions. *Nucleic Acids Res.*, 2008, **36**(suppl 2), pp.W260-W264.
264. Klabunde, T., Petrassi, H.M., Oza, V.B., Raman, P., Kelly, J.W. and Sacchettini, J.C. Rational design of potent human transthyretin amyloid disease inhibitors. *Nat. Struct. Mol. Biol.*, 2000, **7**(4), pp.312-321.
265. Petrassi, H.M., Klabunde, T., Sacchettini, J. and Kelly, J.W. Structure-based design of *N*-phenyl phenoxazine transthyretin amyloid fibril inhibitors. *J. Am. Chem. Soc.*, 2000, **122**(10), pp.2178-2192.
266. Aitken, J.F., Loomes, K.M., Konarkowska, B. and Cooper, G.J.S. Suppression by polycyclic compounds of the conversion of human amylin into insoluble amyloid. *Biochem. J.*, 2003, **374**(3), pp.779-784.
267. Tomiyama, T., Shoji, A., Kataoka, K.-I., Suwa, Y., Asano, S., Kaneko, H. and Endo, N. Inhibition of amyloid protein aggregation and neurotoxicity by rifampicin. *J. Biol. Chem.*, 1996, **271**(12), pp.6839-6844.
268. Lashuel, H.A., Hartley, D.M., Balakhaneh, D., Aggarwal, A., Teichberg, S. and Callaway, D.J.E. New class of inhibitors of amyloid- $\beta$  fibril formation. *J. Biol. Chem.*, 2002, **277**(45), pp.42881-42890.
269. Jameson, L.P., Smith, N.W. and Dzyuba, S.V. Dye-binding assays for evaluation of the effects of small molecule inhibitors on amyloid (A $\beta$ ) self-assembly. *ACS Chem. Neurosci.*, 2012, **3**(11), pp.807-819.
270. Hudson, S.A., Ecroyd, H., Kee, T.W. and Carver, J.A. The thioflavin T fluorescence assay for amyloid fibril detection can be biased by the presence of exogenous compounds. *FEBS J.*, 2009, **276**(20), pp.5960-5972.
271. Biancalana, M., Makabe, K., Koide, A. and Koide, S. Molecular mechanism of thioflavin-T binding to the surface of  $\beta$ -rich peptide self-assemblies. *J. Mol. Biol.*, 2009, **385**(4), pp.1052-1063.
272. Stsiapura, V.I., Maskevich, A.A., Kuzmitsky, V.A., Turoverov, K.K. and Kuznetsova, I.M. Computational study of thioflavin T torsional relaxation in the excited state. *J. Phys. Chem. A.*, 2007, **111**(22), pp.4829-4835.
273. Conway, K.A., Rochet, J.-C., Bieganski, R.M. and Lansbury, P.T. Kinetic stabilization of the  $\alpha$ -synuclein protofibril by a dopamine- $\alpha$ -synuclein adduct. *Science*, 2001, **294**(5545), pp.1346-1349.
274. Byeon, S.R., Lee, J.H., Sohn, J.-H., Kim, D.C., Shin, K.J., Yoo, K.H., Mook-Jung, I., Lee, W.K. and Kim, D.J. Bis-styrylpyridine and bis-styrylbenzene derivatives as inhibitors for A $\beta$  fibril formation. *Bioorg. Med. Chem. Lett.*, 2007, **17**(5), pp.1466-1470.
275. Reinke, A.A. and Gestwicki, J.E. Structure-activity relationships of amyloid  $\beta$ -aggregation inhibitors based on curcumin: influence of linker length and flexibility. *Chem. Biol. Drug Des.*, 2007, **70**(3), pp.206-215.
276. Byun, J.H., Kim, H., Kim, Y., Mook-Jung, I., Kim, D.J., Lee, W.K. and Yoo, K.H. Aminostyrylbenzofuran derivatives as potent inhibitors for A $\beta$  fibril formation. *Bioorg. Med. Chem. Lett.*, 2008, **18**(20), pp.5591-5593.
277. Falkovskaia, E., Sengupta, P.K. and Kasha, M. Photophysical induction of dual fluorescence of quercetin and related hydroxyflavones upon intermolecular H-bonding to solvent matrix. *Chem. Phys. Lett.*, 1998, **297**(1-2), pp.109-114.

278. Kunwar, A., Barik, A., Pandey, R. and Priyadarsini, K.I. Transport of liposomal and albumin loaded curcumin to living cells: an absorption and fluorescence spectroscopic study. *BBA - Gen. Subjects*, 2006, **1760**(10), pp.1513-1520.
279. Banwell, C.N. and Mccash, E.M. *Fundamentals of Molecular Spectroscopy*. 4th ed. London: McGraw-Hill, 1994.
280. Daval, M., Bedrood, S., Gurlo, T., Huang, C.J., Costes, S., Butler, P.C. and Langen, R. The effect of curcumin on human islet amyloid polypeptide misfolding and toxicity. *Amyloid*, 2010, **17**(3-4), pp.118-28.
281. Noor, H., Cao, P. and Raleigh, D.P. Morin hydrate inhibits amyloid formation by islet amyloid polypeptide and disaggregates amyloid fibers. *Protein Sci.*, 2012, **21**(3), pp.373-82.
282. Necula, M., Kayed, R., Milton, S. and Glabe, C.G. Small molecule inhibitors of aggregation indicate that amyloid  $\beta$  oligomerization and fibrillization pathways are independent and distinct. *J. Biol. Chem.*, 2007, **282**(14), pp.10311-10324.
283. Meng, F., Marek, P., Potter, K.J., Verchere, C.B. and Raleigh, D.P. Rifampicin does not prevent amyloid fibril formation by human islet amyloid polypeptide but does inhibit fibril thioflavin-T interactions: implications for mechanistic studies of  $\beta$ -cell death. *Biochem.*, 2008, **47**(22), pp.6016-6024.
284. Meng, F. and Raleigh, D.P. Inhibition of glycosaminoglycan-mediated amyloid formation by islet amyloid polypeptide and proIAPP processing intermediates. *J. Mol. Biol.*, 2010, **406**(3), pp.491-502.
285. Meng, F., Abedini, A., Plesner, A., Middleton, C.T., Potter, K.J., Zanni, M.T., Verchere, C.B. and Raleigh, D.P. The sulfated triphenyl methane derivative acid fuchsin is a potent inhibitor of amyloid formation by human islet amyloid polypeptide and protects against the toxic effects of amyloid formation. *J. Mol. Biol.*, 2010, **400**(3), pp.555-566.
286. Cheng, B., Liu, X., Gong, H., Huang, L., Chen, H., Zhang, X., Li, C., Yang, M., Ma, B., Jiao, L., Zheng, L. and Huang, K. Coffee components inhibit amyloid formation of human islet amyloid polypeptide *in vitro*: possible link between coffee consumption and diabetes mellitus. *J. Agric. Food Chem.*, 2011, **59**(24), pp.13147-13155.
287. Cheng, B., Gong, H., Li, X., Sun, Y., Zhang, X., Chen, H., Liu, X., Zheng, L. and Huang, K. Silibinin inhibits the toxic aggregation of human islet amyloid polypeptide. *Biochem. Biophys. Res. Commun.*, 2012, **419**(3), pp.495-499.
288. Zelusa, C., Foxa, A., Calciano, A., Faridianb, B., Nogajb, L. and Moffeta, D. Myricetin inhibits islet amyloid polypeptide (IAPP) aggregation and rescues living mammalian cells from IAPP toxicity. *Open Biochem. J.*, 2012, **6**, pp.66-70.
289. Evers, F., Jeworrek, C., Tiemeyer, S., Weise, K., Sellin, D., Paulus, M., Struth, B., Tolan, M. and Winter, R. Elucidating the mechanism of lipid membrane-induced IAPP fibrillogenesis and its inhibition by the red wine compound resveratrol: a synchrotron X-ray reflectivity study. *J. Am. Chem. Soc.*, 2009, **131**(27), pp.9516-9521.
290. Wei, L., Jiang, P., Xu, W., Li, H., Zhang, H., Yan, L., Chan-Park, M.B., Liu, X.-W., Tang, K., Mu, Y. and Pervushin, K. The molecular basis of distinct aggregation pathways of islet amyloid polypeptide. *J. Biol. Chem.*, 2011, **286**(8), pp.6291-6300.
291. Engel, M.F.M., Vandenakker, C.C., Schleegeer, M., Velikov, K.P., Koenderink, G.H. and Bonn, M. The polyphenol EGCG inhibits amyloid formation less efficiently at phospholipid interfaces than in bulk solution. *J. Am. Chem. Soc.*, 2012, **134**(36), pp.14781-14788.
292. Pryor, N.E., Moss, M.A. and Hestekin, C.N. Unraveling the early events of amyloid- $\beta$  protein (A $\beta$ ) aggregation: techniques for the determination of A $\beta$  aggregate size. *Int. J. Mol. Sci.*, 2012, **13**(3), pp.3038-3072.
293. Streets, A.M., Sourigues, Y., Kopito, R.R., Melki, R. and Quake, S.R. Simultaneous measurement of amyloid fibril formation by dynamic light scattering and fluorescence reveals complex aggregation kinetics. *PLoS ONE*, 2013, **8**(1), p.e54541.
294. Hamrang, Z., Rattray, N.J.W. and Pluen, A. Proteins behaving badly: emerging technologies in profiling biopharmaceutical aggregation. *Trends Biotechnol.*, 2013, **31**(8), pp.448-458.

295. Nicholson, R.L., Welch, M., Ladlow, M. and Spring, D.R. Small-molecule screening: advances in microarraying and cell-imaging technologies. *ACS Chem. Biol.*, 2007, **2**(1), pp.24-30.
296. Chen, J., Armstrong, A.H., Koehler, A.N. and Hecht, M.H. Small molecule microarrays enable the discovery of compounds that bind the Alzheimer's A $\beta$  peptide and reduce its cytotoxicity. *J. Am. Chem. Soc.*, 2010, **132**(47), pp.17015-17022.
297. Butterfield, S., Hejjaoui, M., Fauvet, B., Awad, L. and Lashuel, H.A. Chemical strategies for controlling protein folding and elucidating the molecular mechanisms of amyloid formation and toxicity. *J. Mol. Biol.*, 2012, **421**(2-3), pp.204-236.
298. Sabate, R., De Groot, N. and Ventura, S. Protein folding and aggregation in bacteria. *Cell. Mol. Life Sci.*, 2010, **67**(16), pp.2695-2715.
299. Mckoy, A.F., Chen, J., Schupbach, T. and Hecht, M.H. A novel inhibitor of amyloid  $\beta$  (A $\beta$ ) peptide aggregation: from high throughput screening to efficacy in an animal model of Alzheimer disease. *J. Biol. Chem.*, 2012, **287**(46), pp.38992-39000.
300. Navarro, S., Villar-Piqué, A. and Ventura, S. Selection against toxic aggregation-prone protein sequences in bacteria. *BBA - Mol. Cell. Res.*, 2014, **1843**(5), pp.866-874.
301. Lee, L.L., Ha, H., Chang, Y.T. and Delisa, M.P. Discovery of amyloid-beta aggregation inhibitors using an engineered assay for intracellular protein folding and solubility. *Protein Sci.*, 2009, **18**(2), pp.277-286.
302. Wurth, C., Guimard, N.K. and Hecht, M.H. Mutations that reduce aggregation of the Alzheimer's A $\beta_{42}$  peptide: an unbiased search for the sequence determinants of A $\beta$  amyloidogenesis. *J. Mol. Biol.*, 2002, **319**(5), pp.1279-1290.
303. Kim, W., Kim, Y., Min, J., Kim, D.J., Chang, Y.-T. and Hecht, M.H. A high-throughput screen for compounds that inhibit aggregation of the Alzheimer's peptide. *ACS Chem. Biol.*, 2006, **1**(7), pp.461-469.
304. Wehrman, T., Kleaveland, B., Her, J.-H., Balint, R.F. and Blau, H.M. Protein-protein interactions monitored in mammalian cells via complementation of  $\beta$ -lactamase enzyme fragments. *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**(6), pp.3469-3474.
305. Galarneau, A., Primeau, M., Trudeau, L.-E. and Michnick, S.W.  $\beta$ -Lactamase protein fragment complementation assays as *in vivo* and *in vitro* sensors of protein-protein interactions. *Nat. Biotechnol.*, 2002, **20**(6), pp.619-622.
306. Gerlach, R.G. and Hensel, M. Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. *Int. J. Med. Microbiol.*, 2007, **297**(6), pp.401-415.
307. Decad, G.M. and Nikaido, H. Outer membrane of gram-negative bacteria. XII. Molecular-sieving function of cell wall. *J. Bacteriol.*, 1976, **128**(1), pp.325-336.
308. Foit, L., Morgan, G.J., Kern, M.J., Steimer, L.R., Von Hacht, A.A., Titchmarsh, J., Warriner, S.L., Radford, S.E. and Bardwell, J.C.A. Optimizing protein stability *in vivo*. *Mol. Cell*, 2009, **36**(5), pp.861-871.
309. Barany, F. Single-stranded hexameric linkers: a system for in-phase insertion mutagenesis and protein engineering. *Gene*, 1985, **37**(1-3), pp.111-123.
310. Zebala, J. and Barany, F. Mapping catalytically important regions of an enzyme using two-codon insertion mutagenesis: a case study correlating  $\beta$ -lactamase mutants with the three-dimensional structure. *Gene*, 1991, **100**(0), pp.51-57.
311. Mathonet, P., Deherve, J., Soumillon, P. and Fastrez, J. Active TEM-1  $\beta$ -lactamase mutants with random peptides inserted in three contiguous surface loops. *Protein Sci.*, 2006, **15**(10), pp.2323-2334.
312. Foit, L. Optimising protein stability and folding *in vivo*, PhD Thesis. thesis, 2010.
313. Hailu, T.T., Foit, L. and Bardwell, J.C.A. *In vivo* detection and quantification of chemicals that enhance protein stability. *Anal. Biochem*, 2013, **434**(1), pp.181-186.
314. Strohl, W.R. and Knight, D.M. Discovery and development of biopharmaceuticals: current issues. *Curr. Opin. Biotechnol.*, 2009, **20**(6), pp.668-672.
315. Shukla, A.A. and Thömmes, J. Recent advances in large-scale production of monoclonal antibodies and related proteins. *Trends Biotechnol.*, 2010, **28**(5), pp.253-261.

316. Chennamsetty, N., Voynov, V., Kayser, V., Helk, B. and Trout, B.L. Design of therapeutic proteins with enhanced stability. *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**(29), pp.11937-11942.
317. Rader, R.A. (Re)defining biopharmaceutical. *Nat. Biotechnol.*, 2008, **26**(7), pp.743-751.
318. Walsh, G. Biopharmaceuticals and biotechnology medicines: an issue of nomenclature. *Eur. J. Pharm. Sci.*, 2002, **15**(2), pp.135-138.
319. Rader, R.A. What is a biopharmaceutical? Part 1: (bio)technology-based definitions. *BioExecutive Internat.*, 2005, **March**, pp.60-65.
320. Leavy, O. Therapeutic antibodies: past, present and future. *Nat. Rev. Immunol.*, 2010, **10**(5), pp.297-297.
321. The College of Physicians of Philadelphia. *The history of vaccines* [online]. 2014. [Accessed 17<sup>th</sup> January 2014]. Available from: <http://www.historyofvaccines.org/content/timelines/all>.
322. Banting, F.G., Best, C.H., Collip, J.B., Cambell, W.R. and Gfletcher, A.A. Pancreatic extracts in the treatment of diabetes mellitus: preliminary report. *Can. Med. Assoc. J.*, 1922, **145**, pp.1281-1286.
323. Scott, D.A. and Best, C.H. The Preparation of Insulin. *Ind. Eng. Chem.*, 1925, **17**(3), pp.238-240.
324. Simoni, R.D., Hill, R.L. and Vaughan, M. The discovery of insulin: the work of Frederick Banting and Charles Best. *J. Biol. Chem.*, 2002, **277**(26), pp.e15-e15.
325. Leader, B., Baca, Q.J. and Golan, D.E. Protein therapeutics: a summary and pharmacological classification. *Nat. Rev. Drug Discov.*, 2008, **7**(1), pp.21-39.
326. Goeddel, D.V., Kleid, D.G., Bolivar, F., Heyneker, H.L., Yansura, D.G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K. and Riggs, A.D. Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *Proc. Natl. Acad. Sci.*, 1979, **76**(1), pp.106-110.
327. Clark, A.L., Knight, G., Wiles, P., Keen, H., Ward, J., Cauldwell, J., Adeniyi-Jones, R., Leiper, J., Jones, R., Maccuish, A., Watkins, P., Glynne, A. and Scotton, J. Biosynthetic human insulin in the treatment of diabetes. *Lancet*, 1982, **320**(8294), pp.354-357.
328. Rosado, J.L., Solomons, N.W., Lisker, R. and Bourges, H. Enzyme replacement therapy for primary adult lactase deficiency. Effective reduction of lactose malabsorption and milk intolerance by direct addition of beta-galactosidase to milk at mealtime. *Gastroenterology*, 1984, **87**(5), pp.1072-1082.
329. Out, H.J., Driessen, S.G.a.J., Mannaerts, B.M.J.L. and Coelingh Bennink, H.J.T. Recombinant follicle-stimulating hormone (follitropin beta, Puregon) yields higher pregnancy rates in *in vitro* fertilization than urinary gonadotropins. *Fertil. Steril.*, 1997, **68**(1), pp.138-142.
330. Shapiro, A.D., Gilchrist, G.S., Keith Hoots, W., Cooper, H.A. and Gastineau, D.A. Prospective, randomised trial of two doses of rFVIIa (NovoSeven) in haemophilia patients with inhibitors undergoing surgery. *Thromb. Haemostasis*, 1998, **80**(11), pp.773-778.
331. Rao, D.B., Sane, P.G. and Georgiev, E.L. Collagenase in the treatment of dermal and decubitus ulcers. *J. Am. Geriatr. Soc.*, 1975, **1**, pp.22-30.
332. Fuchs, H.J., Borowitz, D.S., Christiansen, D.H., Morris, E.M., Nash, M.L., Ramsey, B.W., Rosenstein, B.J., Smith, A.L. and Wohl, M.E. Effect of aerosolized recombinant human DNase on exacerbations of respiratory symptoms and on pulmonary function in patients with cystic fibrosis. *N. Engl. J. Med.*, 1994, **331**(10), pp.637-642.
333. Olsen, N.J. and Stein, C.M. New drugs for rheumatoid arthritis. *N. Engl. J. Med.*, 2004, **350**(21), pp.2167-2179.
334. Weinblatt, M.E., Keystone, E.C., Furst, D.E., Moreland, L.W., Weisman, M.H., Birbara, C.A., Teoh, L.A., Fischkoff, S.A. and Chartash, E.K. Adalimumab, a fully human anti-tumor necrosis factor  $\alpha$  monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. *Arthritis Rheum.*, 2003, **48**(1), pp.35-45.
335. Matthews, T., Salgo, M., Greenberg, M., Chung, J., Demasi, R. and Bolognesi, D. Enfuvirtide: the first therapy to inhibit the entry of HIV-1 into host CD4 lymphocytes. *Nat. Rev. Drug Discov.*, 2004, **3**(3), pp.215-225.
336. Lazzarin, A., Clotet, B., Cooper, D., Reynes, J., Arastéh, K., Nelson, M., Katlama, C., Stellbrink, H.-J., Delfraissy, J.-F., Lange, J., Huson, L., Demasi, R., Wat, C., Delehanty, J., Drobnes, C. and Salgo,

- M. Efficacy of Enfuvirtide in patients infected with drug-resistant HIV-1 in Europe and Australia. *N. Engl. J. Med.*, 2003, **348**(22), pp.2186-2195.
337. Sievers, E.L., Larson, R.A., Stadtmauer, E.A., Estey, E., Löwenberg, B., Dombret, H., Karanes, C., Theobald, M., Bennett, J.M., Sherman, M.L., Berger, M.S., Eten, C.B., Loken, M.R., Van Dongen, J.J.M., Bernstein, I.D., Appelbaum, F.R. and Group, F.T.M.S. Efficacy and safety of Gemtuzumab Ozogamicin in patients With CD33-positive acute myeloid leukemia in first relapse. *J. Clin. Oncol.*, 2001, **19**(13), pp.3244-3254.
338. Giles, F., Estey, E. and O'brien, S. Gemtuzumab ozogamicin in the treatment of acute myeloid leukemia. *Cancer*, 2003, **98**(10), pp.2095-2104.
339. Hodi, F.S., O'day, S.J., Mcdermott, D.F., Weber, R.W., Sosman, J.A., Haanen, J.B., Gonzalez, R., Robert, C., Schadendorf, D., Hassel, J.C., Akerley, W., Van Den Eertwegh, A.J.M., Lutzky, J., Lorigan, P., Vaubel, J.M., Linette, G.P., Hogg, D., Ottensmeier, C.H., Lebbé, C., Peschel, C., Quirt, I., Clark, J.I., Wolchok, J.D., Weber, J.S., Tian, J., Yellin, M.J., Nichol, G.M., Hoos, A. and Urba, W.J. Improved survival with Ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.*, 2010, **363**(8), pp.711-723.
340. Sliwkowski, M.X. and Mellman, I. Antibody therapeutics in cancer. *Science*, 2013, **341**(6151), pp.1192-1198.
341. Ranke, M., Gruhler, M., Roskamp, R., Brüggmann, G., Attanasio, A., Blum, W. and Bierich, J. Testing with growth hormone-releasing factor (GRF(1-29)NH<sub>2</sub>) and somatomedin C measurements for the evaluation of growth hormone deficiency. *Eur. J. Pediatr.*, 1986, **145**(6), pp.485-92.
342. Ghigo, E., Aimaretti, G., Gianotti, L., Bellone, J., Arvat, E. and Camanni, F. New approach to the diagnosis of growth hormone deficiency in adults. *Eur. J. Endocrinol.*, 1996, **134**(3), pp.352-356.
343. Maguire, R., Pascucci, V., Maroli, A. and Gulfo, J. Immunoscintigraphy in patients with colorectal, ovarian, and prostate cancer. Results with site-specific immunoconjugates. *Cancer*, 1993, **72**(11), pp.3453-3462.
344. Pavlou, A.K. The market of therapeutic recombinant proteins to 2010. *J. Commer. Biotechnol.*, 2004, **10**, pp.363-367.
345. Meyer, H.P., Brass, J., Jungo, C., Klein, J., Wenger, J. and Mommers, R. An emerging star for therapeutic and catalytic protein production. *BioProcess Int.*, 2008, **6**(S6), pp.10-21.
346. Walsh, G. Blood products and therapeutic enzymes. In: *Biopharmaceuticals: biochemistry & biotechnology*. First ed. Chichester: J. Wiley & sons, 1998, pp.298-299.
347. Aggarwal, S. What's fueling the biotech engine - 2012 to 2013. *Nat. Biotechnol.*, 2014, **32**(1), pp.32-39.
348. Walsh, G. The drug manufacturing process. In: *Biopharmaceuticals: biochemistry & biotechnology*. First ed. Chichester: J. Wiley & sons, 1998, pp.95-115.
349. Ma, J.K.C., Drake, P.M.W. and Christou, P. The production of recombinant pharmaceutical proteins in plants. *Nat. Rev. Genet.*, 2003, **4**(10), pp.794-805.
350. Walsh, G. The drug-manufacturing process. In: *Biopharmaceuticals: biochemistry and biotechnology*. Chichester: J. Wiley & Sons, 1998, p.75.
351. Frokjaer, S. and Otzen, D.E. Protein drug stability: a formulation challenge. *Nat. Rev. Drug Discov.*, 2005, **4**(4), pp.298-306.
352. Wang, W. Protein aggregation and its inhibition in biopharmaceutics. *Int. J. Pharm.*, 2005, **289**(1-2), pp.1-30.
353. Akers, M.J., Vasudevan, V. and Stickelmeyer, M. Formulation development of protein dosage forms. In: S.L. NAIL and M.J. AKERS, eds. *Development and Manufacture of Protein Pharmaceuticals*. New York: Plenum Publishers, 2002, p.50.
354. Shire, S.J. Formulation and manufacturability of biologics. *Curr. Opin. Biotechnol.*, 2009, **20**(6), pp.708-714.
355. Rathore, N. and Rajan, R.S. Current perspectives on stability of protein drug products during formulation, fill and finish operations. *Biotechnol. Prog.*, 2008, **24**(3), pp.504-514.
356. Carter, P.J. Potent antibody therapeutics by design. *Nat. Rev. Immunol.*, 2006, **6**(5), pp.343-357.



357. Hoogenboom, H.R. Selecting and screening recombinant antibody libraries. *Nat. Biotechnol.*, 2005, **23**(9), pp.1105-1116.
358. Lonberg, N. Human antibodies from transgenic animals. *Nat. Biotechnol.*, 2005, **23**(9), pp.1117-1125.
359. Patro, S.Y. and Przybycien, T.M. Simulations of kinetically irreversible protein aggregate structure. *Biophys. J.*, 1994, **66**(5), pp.1274-1289.
360. Istrail, S., Schwartz, R. and King, J. Lattice simulations of aggregation funnels for protein folding. *J. Comput. Biol.*, 1999, **6**, pp.143-162.
361. Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J. and Serrano, L. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.*, 2004, **22**(10), pp.1302-1306.
362. Chiti, F., Stefani, M., Taddei, N., Ramponi, G. and Dobson, C.M. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature*, 2003, **424**(6950), pp.805-808.
363. Dubay, K.F., Pawar, A.P., Chiti, F., Zurdo, J., Dobson, C.M. and Vendruscolo, M. Prediction of the absolute aggregation rates of amyloidogenic polypeptide chains. *J. Mol. Biol.*, 2004, **341**(5), pp.1317-1326.
364. Miller, B.R., Demarest, S.J., Lugovskoy, A., Huang, F., Wu, X., Snyder, W.B., Croner, L.J., Wang, N., Amatucci, A., Michaelson, J.S. and Glaser, S.M. Stability engineering of scFvs for the development of bispecific and multivalent antibodies. *Protein Eng. Des. Sel.*, 2010, **23**(7), pp.549-557.
365. Zhao, J.-X., Yang, L., Gu, Z.-N., Chen, H.-Q., Tian, F.-W., Chen, Y.-Q., Zhang, H. and Chen, W. Stabilization of the single-chain fragment variable by an interdomain disulfide bond and its effect on antibody affinity. *Int. J. Mol. Sci.*, 2010, **12**(1), pp.1-11.
366. Langedijk, A.C., Honegger, A., Maat, J., Planta, R.J., Van Schaik, R.C. and Plückthun, A. The nature of antibody heavy chain residue H6 strongly influences the stability of a V<sub>H</sub> domain lacking the disulfide bridge. *J. Mol. Biol.*, 1998, **283**(1), pp.95-110.
367. King, A.C., Woods, M., Liu, W., Lu, Z., Gill, D. and Krebs, M.R.H. High-throughput measurement, correlation analysis, and machine-learning predictions for pH and thermal stabilities of Pfizer-generated antibodies. *Protein Sci.*, 2011, **20**(9), pp.1546-1557.
368. Zhi-Jian Lu, Su-Jun Deng, Da-Gang Huang, Yun He, Ming Lei, Zhou, L. and Jin., P. Frontier of therapeutic antibody discovery: the challenges and how to face them. *World J. Biol. Chem.*, 2012, **3**(12), pp.187-196.
369. Banga, A. and Mitra, R. Minimization of shaking-induced formation of insoluble aggregates of insulin by cyclodextrins. *J. Drug Target.*, 1993, **1**(4), pp.341-345.
370. Dotsikas, Y. and Loukas, Y.L. Kinetic degradation study of insulin complexed with methyl- $\beta$  cyclodextrin. Confirmation of complexation with electrospray mass spectrometry and  $^1\text{H}$  NMR. *J. Pharm. Biomed. Anal.*, 2002, **29**(3), pp.487-494.
371. Tokihiro, K., Irie, T., Uekama, K. and Pitha, J. Potential use of maltosyl- $\beta$ -cyclodextrin for inhibition of insulin self-association in aqueous solution. *Pharm. Pharmacol. Commun.*, 1995, **1**(2), pp.49-53.
372. Otzen, D.E., Knudsen, B.R., Aachmann, F., Larsen, K.L. and Wimmer, R. Structural basis for cyclodextrins' suppression of human growth hormone aggregation. *Protein Sci.*, 2002, **11**(7), pp.1779-1787.
373. Cooper, A. Effect of cyclodextrins on the thermal stability of globular proteins. *J. Am. Chem. Soc.*, 1992, **114**(23), pp.9208-9209.
374. Carpenter, J., Pikal, M., Chang, B. and Randolph, T. Rational design of stable lyophilized protein formulations: some practical advice. *Pharm. Res.*, 1997, **14**(8), pp.969-975.
375. Edwards, B.M., Barash, S.C., Main, S.H., Choi, G.H., Minter, R., Ullrich, S., Williams, E., Du Fou, L., Wilton, J., Albert, V.R., Ruben, S.M. and Vaughan, T.J. The remarkable flexibility of the human antibody repertoire; isolation of over one thousand different antibodies to a single protein, BlyS. *J. Mol. Biol.*, 2003, **334**(1), pp.103-118.

376. Jespers, L., Schon, O., James, L.C., Veprintsev, D. and Winter, G. Crystal structure of HEL4, a soluble, refoldable human V<sub>H</sub> single domain with a germ-line scaffold. *J. Mol. Biol.*, 2004, **337**(4), pp.893-903.
377. Williamson, J.A. and Miranker, A.D. Direct detection of transient  $\alpha$ -helical states in islet amyloid polypeptide. *Protein Sci.*, 2007, **16**(1), pp.110-117.
378. Walsh, D.M., Thulin, E., Minogue, A.M., Gustavsson, N., Pang, E., Teplow, D.B. and Linse, S. A facile method for expression and purification of the Alzheimer's disease-associated amyloid  $\beta$ -peptide. *FEBS J.*, 2009, **276**(5), pp.1266-1281.
379. Young, L.M., Saunders, J.C., Mahood, R.A., Revill, C.H., Foster, R.J., Tu, L.-H., Raleigh, D.P., Radford, S.E. and Ashcroft, A.E. Screening and classifying small molecule inhibitors of amyloid formation using ion mobility spectrometry-mass spectrometry. *Nature Chemistry (accepted)*, 2014.
380. Marek, P., Woys, A.M., Sutton, K., Zanni, M.T. and Raleigh, D.P. Efficient microwave-assisted synthesis of human islet amyloid polypeptide designed to facilitate the specific incorporation of labeled amino acids. *Org. Lett.*, 2010, **12**(21), pp.4848-51.
381. O'callaghan, C.H., Morris, A., Kirby, S.M. and Shingler, A.H. Novel method for detection of  $\beta$ -lactamases by using a chromogenic cephalosporin substrate. *Antimicrob. Agents Chemother.*, 1972, **1**(4), pp.283-288.
382. Giles, K., Pringle, S.D., Worthington, K.R., Little, D., Wildgoose, J.L. and Bateman, R.H. Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun. Mass Spectrom.*, 2004, **18**(20), pp.2401-14.
383. *Maestro* [CD-ROM]. LLC, NY, USA: Schrödinger, 2014-2.
384. *Rapid Overlay of Chemical Structures (ROCS)* [CD-ROM]. Santa Fe, NM, USA: OpenEye, Scientific Software.
385. Frère, J.M., Joris, B. and Shockman, G.D. Penicillin-sensitive enzymes in peptidoglycan biosynthesis. *Crit. Rev. Neurobiol.*, 1984, **11**(4), pp.299-396.
386. Hermann, J.C., Ridder, L., Holtje, H.-D. and Mulholland, A.J. Molecular mechanisms of antibiotic resistance: QM/MM modelling of deacylation in a class A  $\beta$ -lactamase. *Org. Biomol. Chem.*, 2006, **4**(2), pp.206-210.
387. Tipper, D.J. Mode of action of  $\beta$ -lactam antibiotics. *Pharmacol. Ther.*, 1985, **27**(1), pp.1-35.
388. Sanders, C.C. and Sanders, W.E. Microbial resistance to newer generation  $\beta$ -lactam antibiotics: clinical and laboratory implications. *J. Infect. Dis.*, 1985, **151**(3), pp.399-406.
389. Felici, A., Amicosante, G., Oratore, A., Strom, R., Ledent, P., Joris, B., Fanuel, L. and Frère, J.M. An overview of the kinetic parameters of class B  $\beta$ -lactamases. *Biochem. J.*, 1993, **291**(1), pp.151-155.
390. Ghuysen, J.-M. Molecular structures of penicillin-binding proteins and  $\beta$ -lactamases. *Trends Microbiol.*, 1994, **2**(10), pp.372-380.
391. Jelsch, C., Mourey, L., Masson, J.-M. and Samama, J.-P. Crystal structure of *Escherichia coli* TEM1  $\beta$ -lactamase at 1.8 Å resolution. *Proteins: Struct., Funct., Bioinf.*, 1993, **16**(4), pp.364-383.
392. Jelsch, C., Lenfant, F., Masson, J.M. and Samama, J.P.  $\beta$ -lactamase TEM1 of *E. coli* crystal structure determination at 2.5 Å resolution. *FEBS Lett.*, 1992, **299**(2), pp.135-142.
393. Sideraki, V., Huang, W., Palzkill, T. and Gilbert, H.F. A secondary drug resistance mutation of TEM-1  $\beta$ -lactamase that suppresses misfolding and aggregation. *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**(1), pp.283-288.
394. Georgiou, G., Valax, P., Ostermeier, M. and Horowitz, P.M. Folding and aggregation of TEM  $\beta$ -lactamase: analogies with the formation of inclusion bodies in *Escherichia coli*. *Protein Sci.*, 1994, **3**(11), pp.1953-1960.
395. Valax, P. and Georgiou, G. Molecular characterization of  $\beta$ -lactamase inclusion bodies produced in *Escherichia coli*. *Biotechnol. Prog.*, 1993, **9**(5), pp.539-547.
396. Janson, J.-C. *Protein purification: principles, high-resolution methods and applications*. New York, : John Wiley & Sons, 1998.

397. Westermark, P. Amyloid in the islets of Langerhans: thoughts and some historical aspects. *Uppsala J. Med. Sci.*, 2011, **116**(2), pp.81-89.
398. Pillay, K. and Govender, P. Amylin uncovered: a review on the polypeptide responsible for type II diabetes. *BioMed Res. Int.*, 2013, **2013**, p.17.
399. Kahn, S.E., Andrikopoulos, S. and Verchere, C.B. Islet amyloid: a long-recognized but underappreciated pathological feature of type 2 diabetes. *Diabetes*, 1999, **48**(2), pp.241-253.
400. Brender, J.R., Lee, E.L., Hartman, K., Wong, P.T., Ramamoorthy, A., Steel, D.G. and Gafni, A. Biphasic effects of insulin on islet amyloid polypeptide membrane disruption. *Biophys. J.*, 2011, **100**(3), pp.685-692.
401. Laskowski, R. *PDBsum: Pictorial database of 3D structures in the Protein Data Bank* [online]. 2014. [Accessed 16th August]. Available from: <http://www.ebi.ac.uk/pdbsum/>.
402. Zhao, J., Yu, X., Liang, G. and Zheng, J. Structural polymorphism of human islet amyloid polypeptide (hIAPP) oligomers highlights the importance of interfacial residue interactions. *Biomacromolecules*, 2010, **12**(1), pp.210-220.
403. Goldsbury, C., Goldie, K., Pellaud, J., Seelig, J., Frey, P., Müller, S.A., Kistler, J., Cooper, G.J.S. and Aebi, U. Amyloid fibril formation from full-length and fragments of amylin. *J. Struct. Biol.*, 2000, **130**(2-3), pp.352-362.
404. Andrews, M.N. and Winter, R. Comparing the structural properties of human and rat islet amyloid polypeptide by MD computer simulations. *Biophys. Chem.*, 2011, **156**, pp.43-50.
405. Paulsson, J.F. and Westermark, G.T. Aberrant processing of human proislet amyloid polypeptide results in increased amyloid formation. *Diabetes*, 2005, **54**(7), pp.2117-2125.
406. Betsholtz, C., Christmansson, L., Engström, U., Rorsman, F., Svensson, V., Johnson, K.H. and Westermark, P. Sequence divergence in a specific region of islet amyloid polypeptide (IAPP) explains differences in islet amyloid formation between species. *FEBS Lett.*, 1989, **251**(1-2), pp.261-264.
407. Fox, A., Snollaerts, T., Errecart Casanova, C., Calciano, A., Nogaj, L.A. and Moffet, D.A. Selection for nonamyloidogenic mutants of islet amyloid polypeptide (IAPP) identifies an extended region for amyloidogenicity. *Biochem.*, 2010, **49**(36), pp.7783-7789.
408. Glenner, G.G. and Wong, C.W. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem. Biophys. Res. Commun.*, 1984, **120**(3), pp.885-890.
409. Jonsson, T., Atwal, J.K., Steinberg, S., Snaedal, J., Jonsson, P.V., Bjornsson, S., Stefansson, H., Sulem, P., Gudbjartsson, D., Maloney, J., Hoyte, K., Gustafson, A., Liu, Y., Lu, Y., Bhangale, T., Graham, R.R., Huttenlocher, J., Bjornsdottir, G., Andreassen, O.A., Jonsson, E.G., Palotie, A., Behrens, T.W., Magnusson, O.T., Kong, A., Thorsteinsdottir, U., Watts, R.J. and Stefansson, K. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature*, 2012, **488**(7409), pp.96-99.
410. Zhang, Y.-W., Thompson, R., Zhang, H. and Xu, H. APP processing in Alzheimer's disease. *Mol. Brain*, 2011, **4**(3).
411. Reinhard, C., Hébert, S.S. and De Strooper, B. The amyloid- $\beta$  precursor protein: integrating structure with biological function. *EMBO J.*, 2005, **24**(23), pp.3996-4006.
412. Selkoe, D.J. The molecular pathology of Alzheimer's disease. *Neuron*, 1991, **6**(4), pp.487-498.
413. Yates, D. and Mcloughlin, D.M. The molecular pathology of Alzheimer's disease. *Psychiatry*, 2008, **7**(1), pp.1-5.
414. Esler, W.P., Stimson, E.R., Ghilardi, J.R., Lu, Y.-A., Felix, A.M., Vinters, H.V., Mantyh, P.W., Lee, J.P. and Maggio, J.E. Point substitution in the central hydrophobic cluster of a human  $\beta$ -amyloid congener disrupts peptide folding and abolishes plaque competence. *Biochem.*, 1996, **35**(44), pp.13914-13921.
415. Tjernberg, L.O., Callaway, D.J.E., Tjernberg, A., Hahne, S., Lilliehöök, C., Terenius, L., Thyberg, J. and Nordstedt, C. A molecular model of Alzheimer amyloid  $\beta$ -peptide fibril formation. *J. Biol. Chem.*, 1999, **274**(18), pp.12619-12625.

416. Hilbich, C., Kisters-Woike, B., Reed, J., Masters, C.L. and Beyreuther, K. Substitutions of hydrophobic amino acids reduce the amyloidogenicity of Alzheimer's disease  $\beta$ A4 peptides. *J. Mol. Biol.*, 1992, **228**(2), pp.460-473.
417. De Groot, N.S., Aviles, F.X., Vendrell, J. and Ventura, S. Mutagenesis of the central hydrophobic cluster in  $\text{A}\beta$ 42 Alzheimer's peptide. *FEBS J.*, 2006, **273**(3), pp.658-668.
418. Ramsden, M., Plant, L.D., Webster, N.J., Vaughan, P.F.T., Henderson, Z. and Pearson, H.A. Differential effects of unaggregated and aggregated amyloid  $\beta$  protein (1–40) on  $\text{K}^+$  channel currents in primary cultures of rat cerebellar granule and cortical neurones. *J. Neurochem.*, 2001, **79**(3), pp.699-712.
419. Kamenetz, F., Tomita, T., Hsieh, H., Seabrook, G., Borchelt, D., Iwatsubo, T., Sisodia, S. and Malinow, R. APP processing and synaptic function. *Neuron*, 2003, **37**(6), pp.925-937.
420. Barrow, C.J., Yasuda, A., Kenny, P.T.M. and Zagorski, M.G. Solution conformations and aggregational properties of synthetic amyloid  $\beta$ -peptides of Alzheimer's disease: analysis of circular dichroism spectra. *J. Mol. Biol.*, 1992, **225**(4), pp.1075-1093.
421. Jarrett, J.T., Berger, E.P. and Lansbury, P.T. The C-terminus of the  $\beta$  protein is critical in amyloidogenesis. *Ann. NY Acad. Sci.*, 1993, **695**(1), pp.144-148.
422. Cukalevski, R., Boland, B., Frohm, B., Thulin, E., Walsh, D. and Linse, S. Role of aromatic side chains in amyloid  $\beta$ -protein aggregation. *ACS Chem. Neurosci.*, 2012, **3**(12), pp.1008-1016.
423. Roher, A.E., Chaney, M.O., Kuo, Y.-M., Webster, S.D., Stine, W.B., Haverkamp, L.J., Woods, A.S., Cotter, R.J., Tuohy, J.M., Krafft, G.A., Bonnell, B.S. and Emmerling, M.R. Morphology and toxicity of  $\text{A}\beta$ -(1-42) dimer derived from neuritic and vascular amyloid deposits of Alzheimer's disease. *J. Biol. Chem.*, 1996, **271**(34), pp.20631-20635.
424. Bitan, G., Kirkitadze, M.D., Lomakin, A., Vollers, S.S., Benedek, G.B. and Teplow, D.B. Amyloid  $\beta$ -protein ( $\text{A}\beta$ ) assembly:  $\text{A}\beta$ 40 and  $\text{A}\beta$ 42 oligomerize through distinct pathways. *Proc. Natl. Acad. Sci.*, 2003, **100**(1), pp.330-335.
425. Fisher, Charles k., Ullman, O. and Stultz, Collin m. Comparative studies of disordered proteins with similar sequences: application to  $\text{A}\beta$ 40 and  $\text{A}\beta$ 42. *Biophys. J.*, **104**(7), pp.1546-1555.
426. Sgourakis, N.G., Yan, Y., Mccallum, S.A., Wang, C. and Garcia, A.E. The Alzheimer's peptides  $\text{A}\beta$ 40 and 42 adopt distinct conformations in water: a combined MD / NMR study. *J. Mol. Biol.*, 2007, **368**(5), pp.1448-1457.
427. Emerit, J., Edeas, M. and Bricaire, F. Neurodegenerative diseases and oxidative stress. *Biomed. Pharmacother.*, 2004, **58**(1), pp.39-46.
428. Barnham, K.J., Haeffner, F., Ciccotosto, G.D., Curtain, C.C., Tew, D., Mavros, C., Beyreuther, K., Carrington, D., Masters, C.L., Cherny, R.A., Cappai, R. and Bush, A.I. Tyrosine gated electron transfer is key to the toxic mechanism of Alzheimer's disease  $\beta$ -amyloid. *FASEB J.*, 2004, **18**(12), pp.1427-1429.
429. Byrne, E. Does mitochondrial respiratory chain dysfunction have a role in common neurodegenerative disorders? *J. Clin. Neurosci.*, 2002, **9**(5), pp.497-501.
430. Arispe, N., Rojas, E. and Pollard, H.B. Alzheimer disease amyloid beta protein forms calcium channels in bilayer membranes: blockade by tromethamine and aluminum. *Proc. Natl. Acad. Sci.*, 1993, **90**(2), pp.567-571.
431. Bush, A.I., Masters, C.L. and Tanzi, R.E. Copper,  $\beta$ -amyloid, and Alzheimer's disease: tapping a sensitive connection. *Proc. Natl. Acad. Sci.*, 2003, **100**(20), pp.11193-11194.
432. Masters, C.L., Cappai, R., Barnham, K.J. and Villemagne, V.L. Molecular mechanisms for Alzheimer's disease: implications for neuroimaging and therapeutics. *J. Neurochem.*, 2006, **97**(6), pp.1700-1725.
433. Hardy, J. and Selkoe, D.J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science*, 2002, **297**(5580), pp.353-356.
434. Qiao, Q., Bowman, G.R. and Huang, X. Dynamics of an intrinsically disordered protein reveal metastable conformations that potentially seed aggregation. *J. Am. Chem. Soc.*, 2013, **135**(43), pp.16092-16101.
435. Seeliger, J., Weise, K., Opitz, N. and Winter, R. The effect of  $\text{A}\beta$  on IAPP aggregation in the presence of an isolated  $\beta$ -cell membrane. *J. Mol. Biol.*, 2012, **421**(2–3), pp.348-363.

436. Glabe, C.G. Structural classification of toxic amyloid oligomers. *J. Biol. Chem.*, 2008, **283**(44), pp.29639-29643.
437. Uversky, V.N. Mysterious oligomerization of the amyloidogenic proteins. *FEBS J.*, 2010, **277**(14), pp.2940-2953.
438. Morris, A.M., Watzky, M.A. and Finke, R.G. Protein aggregation kinetics, mechanism, and curve-fitting: A review of the literature. *BBA - Proteins Proteomics*, 2009, **1794**(3), pp.375-397.
439. Ferrone, F.A. Nucleation: the connections between equilibrium and kinetic behavior. In: K. INDU and W. RONALD, eds. *Methods in Enzymology*. Academic Press, 2006, pp.285-299.
440. Powers, E.T. and Powers, D.L. Mechanisms of protein fibril formation: nucleated polymerization with competing off-pathway aggregation. *Biophys. J.*, **94**(2), pp.379-391.
441. Mossuto, M.F. Disulfide bonding in neurodegenerative misfolding diseases. *International Journal of Cell Biology*, 2013, **2013**(Article ID 318319), p.7
442. Mossuto, M.F., Bolognesi, B., Guixer, B., Dhulesia, A., Agostini, F., Kumita, J.R., Tartaglia, G.G., Dumoulin, M., Dobson, C.M. and Salvatella, X. Disulfide bonds reduce the toxicity of the amyloid fibrils formed by an extracellular protein. *Angew. Chem. Int. Ed.*, 2011, **50**(31), pp.7048-7051.
443. Ehrnhoefer, D.E., Duennwald, M., Markovic, P., Wacker, J.L., Engemann, S., Roark, M., Legleiter, J., Marsh, J.L., Thompson, L.M., Lindquist, S., Muchowski, P.J. and Wanker, E.E. Green tea (-)-epigallocatechin-gallate modulates early events in huntingtin misfolding and reduces toxicity in Huntington's disease models. *Hum. Mol. Genet.*, 2006, **15**(18), pp.2743-2751.
444. Guerrero-Muñoz, M.J., Castillo-Carranza, D.L. and Kaye, R. Therapeutic approaches against common structural features of toxic oligomers shared by multiple amyloidogenic proteins. *Biochem. Pharmacol.*, 2014, **88**(4), pp.468-478.
445. Yang, F., Lim, G.P., Begum, A.N., Ubeda, O.J., Simmons, M.R., Ambegaokar, S.S., Chen, P.P., Kaye, R., Glabe, C.G., Frautschy, S.A. and Cole, G.M. Curcumin inhibits formation of amyloid  $\beta$  oligomers and fibrils, binds plaques, and reduces amyloid *in vivo*. *J. Biol. Chem.*, 2005, **280**(7), pp.5892-5901.
446. Masuda, Y., Fukuchi, M., Yatagawa, T., Tada, M., Takeda, K., Irie, K., Akagi, K.-I., Monobe, Y., Imazawa, T. and Takegoshi, K. Solid-state NMR analysis of interaction sites of curcumin and 42-residue amyloid  $\beta$ -protein fibrils. *Bioorg. Med. Chem.*, 2011, **19**(20), pp.5967-5974.
447. Straus, D.B., Walter, W.A. and Gross, C.A. The heat shock response of *E. coli* is regulated by changes in the concentration of  $\sigma^{32}$ . *Nature*, 1987, **329**(6137), pp.348-351.
448. Durairajan, S.S.K., Yuan, Q., Xie, L., Chan, W.-S., Kum, W.-F., Koo, I., Liu, C., Song, Y., Huang, J.-D., Klein, W.L. and Li, M. Salvianolic acid B inhibits A $\beta$  fibril formation and disaggregates preformed fibrils and protects against A $\beta$ -induced cytotoxicity. *Neurochem. Int.*, 2008, **52**(4-5), pp.741-750.
449. Ono, K., Hasegawa, K., Yoshiike, Y., Takashima, A., Yamada, M. and Naiki, H. Nordihydroguaiaretic acid potently breaks down pre-formed Alzheimer's  $\beta$ -amyloid fibrils *in vitro*. *J. Neurochem.*, 2002, **81**(3), pp.434-440.
450. Feng, B.Y., Toyama, B.H., Wille, H., Colby, D.W., Collins, S.R., May, B.C.H., Prusiner, S.B., Weissman, J. and Shoichet, B.K. Small-molecule aggregates inhibit amyloid polymerization. *Nat. Chem. Biol.*, 2008, **4**(3), pp.197-199.
451. Owen, S.C., Doak, A.K., Ganesh, A.N., Nedyalkova, L., Mclaughlin, C.K., Shoichet, B.K. and Shoichet, M.S. Colloidal drug formulations can explain "bell-shaped" concentration-response curves. *ACS Chem. Biol.*, 2014, **9**(3), pp.777-784.
452. Eggert, U.S. The why and how of phenotypic small-molecule screens. *Nat. Chem. Biol.*, 2013, **9**(4), pp.206-209.
453. Kamihira-Ishijima, M., Nakazawa, H., Kira, A., Naito, A. and Nakayama, T. Inhibitory mechanism of pancreatic amyloid fibril formation: formation of the complex between tea catechins and the fragment of residues 22-27. *Biochem.*, 2012, **51**(51), pp.10167-10174.
454. Daubenfeld, T., Bouin, A.-P. and Van Der Rest, G. A deconvolution method for the separation of specific versus nonspecific interactions in noncovalent protein-ligand complexes analyzed by ESI-FT-ICR mass spectrometry. *J. Am. Soc. Mass Spectrom.*, 2006, **17**(9), pp.1239-1248.

455. Levine Iii, H. Quantification of  $\beta$ -sheet amyloid fibril structures with thioflavin T. *In: W. RONALD*, ed. *Methods in Enzymology*. Academic Press, 1999, pp.274-284.
456. Woods, L.A., Platt, G.W., Hellewell, A.L., Hewitt, E.W., Homans, S.W., Ashcroft, A.E. and Radford, S.E. Ligand binding to distinct states diverts aggregation of an amyloid-forming protein. *Nat. Chem. Biol.*, 2011, **7**(10), pp.730-739.
457. Sun, N., Sun, J., Kitova, E.N. and Klassen, J.S. Identifying nonspecific ligand binding in electrospray ionization mass spectrometry using the reporter molecule method. *J. Am. Soc. Mass Spectrom.*, 2009, **20**(7), pp.1242-50.
458. Porat, Y., Mazor, Y., Efrat, S. and Gazit, E. Inhibition of islet amyloid polypeptide fibril formation: a potential role for heteroaromatic interactions. *Biochem.*, 2004, **43**(45), pp.14454-14462.
459. Wu, C., Lei, H., Wang, Z., Zhang, W. and Duan, Y. Phenol red interacts with the protofibril-like oligomers of an amyloidogenic hexapeptide NFGAIL through both hydrophobic and aromatic contacts. *Biophys. J.*, 2006, **91**(10), pp.3664-3672.
460. Aarabi, M.H. and Mirhashemi, S.M. The role of two natural flavonoids on human amylin aggregation. *Afr. J. Ph. Pharmacol.*, 2012, **6**(31), pp.2374-2379.
461. Noor, H., Cao, P. and Raleigh, D.P. Morin hydrate inhibits amyloid formation by islet amyloid polypeptide and disaggregates amyloid fibers. *Protein Sci.*, 2012, **21**(3), pp.373-382.
462. Mishra, R., Sellin, D., Radovan, D., Gohlke, A. and Winter, R. Inhibiting islet amyloid polypeptide fibril formation by the red wine compound resveratrol. *ChemBioChem*, 2009, **10**(3), pp.445-449.
463. Radovan, D., Opitz, N. and Winter, R. Fluorescence microscopy studies on islet amyloid polypeptide fibrillation at heterogeneous and cellular membrane interfaces and its inhibition by resveratrol. *FEBS Lett.*, 2009, **583**(9), pp.1439-1445.
464. Valler, M.J. and Green, D. Diversity screening versus focussed screening in drug discovery. *Drug Discovery Today*, 2000, **5**(7), pp.286-293.
465. De Felice, F.G., Viera, M.N.N., Saraiva, L.M., Figueroa-Villar, J.D., Garcia-Abreu, J., Liu, R., Chang, L., Klein, W.L. and Ferreira, S.T. Targeting the neurotoxic species in Alzheimer's disease: inhibitors of A $\beta$  oligomerization. *The FASEB Journal*, 2004, **18**(12), pp.1366-1372.
466. Ladiwala, A.R.A., Lin, J.C., Bale, S.S., Marcelino-Cruz, A.M., Bhattacharya, M., Dordick, J.S. and Tessier, P.M. Resveratrol selectively remodels soluble oligomers and fibrils of amyloid A $\beta$  into off-pathway conformers. *J. Biol. Chem.*, 2010, **285**(31), pp.24228-24237.
467. Scherzer-Attali, R., Pellarin, R., Convertino, M., Frydman-Marom, A., Egoz-Matia, N., Peled, S., Levy-Sakin, M., Shalev, D.E., Caflish, A., Gazit, E. and Segal, D. Complete phenotypic recovery of an Alzheimer's disease model by a quinone-tryptophan hybrid aggregation inhibitor. *PLoS ONE*, 2010, **5**(6), p.e11101.
468. Young, L.M., Cao, P., Raleigh, D.P., Ashcroft, A.E. and Radford, S.E. Ion mobility spectrometry-mass spectrometry defines the oligomeric intermediates in amylin amyloid formation and the mode of action of inhibitors. *J. Am. Chem. Soc.*, 2014, **136**(2), pp.660-70.
469. Harris, F., Dennison, S.R. and Phoenix, D.A. Aberrant action of amyloidogenic host defense peptides: a new paradigm to investigate neurodegenerative disorders? *FASEB J.*, 2012, **26**(5), pp.1776-1781.
470. Kinnunen, P., K. Amyloid formation on lipid membrane surfaces. *Open Biochem. J.*, 2009, **2**, pp.163-175.
471. Soccia, S.J., Kirby, J.E., Washicosky, K.J., Tucker, S.M., Ingelsson, M., Hyman, B., Burton, M.A., Goldstein, L.E., Duong, S., Tanzi, R.E. and Moir, R.D. The Alzheimer's disease-associated amyloid  $\beta$ -protein is an antimicrobial peptide. *PLoS ONE*, 2010, **5**(3), p.e9505.
472. Brouwer, M.C., Tunkel, A.R. and Van De Beek, D. Epidemiology, diagnosis, and antimicrobial treatment of acute bacterial meningitis. *Clin. Microbiol. Rev.*, 2010, **23**(3), pp.467-492.
473. Brogden, K.A. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat. Rev. Microbiol.*, 2005, **3**(3), pp.238-250.
474. Dennison, S.R., Wallace, J., Harris, F. and Phoenix, D., A. Amphiphilic  $\alpha$ -helical antimicrobial peptides and their structure / function relationships. *Protein Pept. Lett.*, 2005, **12**, pp.31-39.

475. Chauhan, A., Ray, I. and Chauhan, V.S. Interaction of amyloid beta-protein with anionic phospholipids: possible involvement of Lys28 and C-terminus aliphatic amino acids. *Neurochem. Res.*, 2000, **25**(3), pp.423-429.
476. Wong, P.T., Schauerte, J.A., Wissner, K.C., Ding, H., Lee, E.L., Steel, D.G. and Gafni, A. Amyloid- $\beta$  membrane binding and permeabilization are distinct processes influenced separately by membrane charge and fluidity. *J. Mol. Biol.*, 2009, **386**(1), pp.81-96.
477. Eckert, G.P., Wood, W.G. and Muller, W.E. Lipid membranes and  $\beta$ -amyloid: a harmful connection. *Curr. Protein. Pept. Sci.*, 2010, **11**, pp.319-325.
478. Sakono, M. and Zako, T. Amyloid oligomers: formation and toxicity of A $\beta$  oligomers. *FEBS J.*, 2010, **277**(6), pp.1348-1358.
479. Tsai, H.-H.G., Lee, J.-B., Tseng, S.-S., Pan, X.-A. and Shih, Y.-C. Folding and membrane insertion of amyloid-beta (25–35) peptide and its mutants: implications for aggregation and neurotoxicity. *Proteins: Struct., Funct., Bioinf.*, 2010, **78**(8), pp.1909-1925.
480. Kawahara, M., Ohtsuka, I., Yokoyama, S., Kato-Negishi, M. and Sadakane, Y. Membrane incorporation, channel formation, and disruption of calcium homeostasis by Alzheimer's &  $\beta$ -amyloid protein. *Int. J. Alzheimer. Dis.*, 2011, **2011**, p.17.
481. Pinheiro, L.B., Gibbs, M.D., Vesey, G., Smith, J.J. and Bergquist, P.L. Fluorescent reference strains of bacteria by chromosomal integration of a modified green fluorescent protein gene. *Appl. Microbiol. Biotechnol.*, 2008, **77**(6), pp.1287-1295.
482. Jespers, L., Schon, O., Famm, K. and Winter, G. Aggregation-resistant domain antibodies selected on phage by heat denaturation. *Nat. Biotechnol.*, 2004, **22**(9), pp.1161-1165.
483. Kohler, G. and Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, 1975, **256**(5517), pp.495-497.
484. Beck, A., Wurch, T., Bailly, C. and Corvaia, N. Strategies and challenges for the next generation of therapeutic antibodies. *Nat. Rev. Immunol.*, 2010, **10**(5), pp.345-352.
485. Spadiut, O., Capone, S., Krainer, F., Glieder, A. and Herwig, C. Microbials for the production of monoclonal antibodies and antibody fragments. *Trends Biotechnol.*, 2014, **32**(1), pp.54-60.
486. Muyldermans, S. Single domain camel antibodies: current status. *Rev. Mol. Biotechnol.*, 2001, **74**(4), pp.277-302.
487. Kuus-Reichel, K., Grauer, L.S., Karavodin, L.M., Knott, C., Krusemeier, M. and Kay, N.E. Will immunogenicity limit the use, efficacy, and future development of therapeutic monoclonal antibodies? *Clin. Diagn. Lab. Immunol.*, 1994, **1**(4), pp.365-72.
488. Sundberg, E.J. and Mariuzza, R.A. Molecular recognition in antibody-antigen complexes. *Ad. Protein Chem.*, 2002, **61**, pp.119-160.
489. Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.S., Novotný, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E. and Crea, R. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proc. Natl. Acad. Sci.*, 1988, **85**(16), pp.5879-5883.
490. Holt, L.J., Herring, C., Jespers, L.S., Woolven, B.P. and Tomlinson, I.M. Domain antibodies: proteins for therapy. *Trends Biotechnol.*, 2003, **21**(11), pp.484-490.
491. Borrebaeck, C.a.K., Malmberg, A.-C., Furebring, C., Michaelsson, A., Ward, S., Danielsson, L. and Ohlin, M. Kinetic analysis of recombinant antibody-antigen interactions: relation between structural domains and antigen binding. *Nat. Biotechnol.*, 1992, **10**(6), pp.697-698.
492. Glockshuber, R., Malia, M., Pfizinger, I. and Plueckthun, A. A comparison of strategies to stabilize immunoglobulin Fv-fragments. *Biochem.*, 1990, **29**(6), pp.1362-1367.
493. Whitlow, M., Bell, B.A., Feng, S.-L., Filpula, D., Hardman, K.D., Hubert, S.L., Rollence, M.L., Wood, J.F., Schott, M.E., Milenic, D.E., Yokota, T. and Schlom, J. An improved linker for single-chain Fv with reduced aggregation and enhanced proteolytic stability. *Protein Eng.*, 1993, **6**(8), pp.989-995.
494. Ward, E.S., Gussow, D., Griffiths, A.D., Jones, P.T. and Winter, G. Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. *Nature*, 1989, **341**(6242), pp.544-546.

495. Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hammers, C., Bajyana Songa, E., N., B. and Hammers, R. Naturally occurring antibodies devoid of light chains. *Nature*, 1993, **363**, pp.446-448.
496. Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J.a.R.G. and Hamers, R. Sequence and structure of V<sub>H</sub> domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. *Protein Eng.*, 1994, **7**(9), pp.1129-1135.
497. Sircar, A., Sanni, K.A., Shi, J. and Gray, J.J. Analysis and modeling of the variable region of camelid single-domain antibodies. *J. Immunol.*, 2011, **186**(11), pp.6357-6367.
498. Roovers, R., Van Dongen, G. and Van Bergen En Henegouwen, P. Nanobodies in therapeutic applications. *Curr. Opin. Mol. Ther.*, 2007, **9**(4), pp.327-335.
499. Harmsen, M.M. and De Haard, H.J. Properties, production, and applications of camelid single-domain antibody fragments. *Appl. Microbiol. Biotechnol.*, 2007, **77**(1), pp.13-22.
500. Holliger, P. and Hudson, P.J. Engineered antibody fragments and the rise of single domains. *Nat. Biotechnol.*, 2005, **23**(9), pp.1126-1136.
501. Nguyen, V.K., Hamers, R., Wyns, L. and Muyldermans, S. Camel heavy-chain antibodies: diverse germline V<sub>H</sub>H and specific mechanisms enlarge the antigen-binding repertoire. *EMBO J.*, 2000, **19**(5), pp.921-930.
502. Spinelli, S., Frenken, L.G.J., Hermans, P., Verrips, T., Brown, K., Tegoni, M. and Cambillau, C. Camelid heavy-chain variable domains provide efficient combining sites to haptens. *Biochem.*, 2000, **39**(6), pp.1217-1222.
503. Desmyter, A., Transue, T.R., Ghahroudi, M.A., Dao Thi, M.-H., Poortmans, F., Hamers, R., Muyldermans, S. and Wyns, L. Crystal structure of a camel single-domain V<sub>H</sub> antibody fragment in complex with lysozyme. *Nat. Struct. Mol. Biol.*, 1996, **3**(9), pp.803-811.
504. Desmyter, A., Spinelli, S., Payan, F., Lauwereys, M., Wyns, L., Muyldermans, S. and Cambillau, C. Three camelid V<sub>H</sub>H domains in complex with porcine pancreatic  $\alpha$ -amylase: inhibition and versatility of binding topology. *J. Biol. Chem.*, 2002, **277**(26), pp.23645-23650.
505. Muyldermans, S., Baral, T.N., Retamozzo, V.C., De Baetselier, P., De Genst, E., Kinne, J., Leonhardt, H., Magez, S., Nguyen, V.K., Revets, H., Rothbauer, U., Stijlemans, B., Tillib, S., Wernery, U., Wyns, L., Hassanzadeh-Ghassabeh, G. and Saerens, D. Camelid immunoglobulins and nanobody technology. *Vet. Immunol. Immunopathol.*, 2009, **128**(1-3), pp.178-183.
506. Vu, K.B., Ghahroudi, M.A., Wyns, L. and Muyldermans, S. Comparison of llama V<sub>H</sub> sequences from conventional and heavy chain antibodies. *Mol. Immunol.*, 1997, **34**(16-17), pp.1121-1131.
507. Van Der Linden, R.H.J., Frenken, L.G.J., De Geus, B., Harmsen, M.M., Ruuls, R.C., Stok, W., De Ron, L., Wilson, S., Davis, P. and Verrips, C.T. Comparison of physical chemical properties of llama V<sub>H</sub>H antibody fragments and mouse monoclonal antibodies. *BBA - Protein Struct. Mol. Enzymol.*, 1999, **1431**(1), pp.37-46.
508. Davies, J. and Riechmann, L. Single antibody domains as small recognition units: design and *in vitro* antigen selection of camelized, human V<sub>H</sub> domains with improved protein stability. *Protein Eng.*, 1996, **9**(6), pp.531-537.
509. De Genst, E., Silence, K., Decanniere, K., Conrath, K., Loris, R., Kinne, J., Muyldermans, S. and Wyns, L. Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**(12), pp.4586-4591.
510. Cortez-Retamozo, V., Lauwereys, M., Hassanzadeh, G., Gobert, M., Conrath, K., Muyldermans, S., De Baetselier, P. and Revets, H. Efficient tumor targeting by single-domain antibody fragments of camels. *Int. J. Cancer*, 2002, **98**(3), pp.456-462.
511. Davies, J. and Riechmann, L. 'Camelising' human antibody fragments: NMR studies on V<sub>H</sub> domains. *FEBS Lett.*, 1994, **339**(3), pp.285-290.
512. Lee, C.C., Perchiacca, J.M. and Tessier, P.M. Toward aggregation-resistant antibodies by design. *Trends Biotechnol.*, 2013, **31**(11), pp.612-620.
513. Ewert, S., Huber, T., Honegger, A. and Plückthun, A. Biophysical properties of human antibody variable domains. *J. Mol. Biol.*, 2003, **325**(3), pp.531-553.
514. Ewert, S., Cambillau, C., Conrath, K. and Plückthun, A. Biophysical properties of camelid V<sub>H</sub>H domains compared to those of human V<sub>H</sub>3 domains. *Biochem.*, 2002, **41**(11), pp.3628-3636.



515. Dudgeon, K., Rouet, R., Kokmeijer, I., Schofield, P., Stolp, J., Langley, D., Stock, D. and Christ, D. General strategy for the generation of human antibody variable domains with increased aggregation resistance. *Proc. Natl. Acad. Sci.*, 2012, **109**(27), pp.10879-10884.
516. Arakawa, T., Ejima, D., Kita, Y. and Tsumoto, K. Small molecule pharmacological chaperones: from thermodynamic stabilization to pharmaceutical drugs. *BBA - Prot. Proteo.*, 2006, **1764**(11), pp.1677-1687.
517. Famm, K., Hansen, L., Christ, D. and Winter, G. Thermodynamically stable aggregation-resistant antibody domains through directed evolution. *J. Mol. Biol.*, 2008, **376**(4), pp.926-931.
518. Kabat, E.A., Wu, T.T., Perry, H.M., Gottesman, K.S. and Foeller, C. *Sequences of proteins of immunological interest*. 5th ed. Bethesda: U.S. Department of Health and Human Services, Public Health Service, National Institute of Health, 1991.
519. Sharff, A.J., Rodseth, L.E., Spurlino, J.C. and Quiocho, F.A. Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochem.*, 1992, **31**(44), pp.10657-10663.
520. Qoronfleh, M.W., Hesterberg, L.K. and Seefeldt, M.B. Confronting high-throughput protein refolding using high pressure and solution screens. *Protein Expression Purif.*, 2007, **55**(2), pp.209-224.
521. Li, Y., Mach, H. and Blue, J.T. High throughput formulation screening for global aggregation behaviors of three monoclonal antibodies. *J. Pharm. Sci.*, 2011, **100**(6), pp.2120-2135.
522. Dasnoy, S., Dezutter, N., Lemoine, D., Le Bras, V. and Pr eat, V. High-throughput screening of excipients intended to prevent antigen aggregation at air-liquid interface. *Pharm. Res.*, 2011, **28**(7), pp.1591-1605.
523. Lentzen, G. and Schwarz, T. Extremolytes: natural compounds from extremophiles for versatile applications. *Appl. Microbiol. Biotechnol.*, 2006, **72**(4), pp.623-634.
524. Arakawa, T. and Timasheff, S.N. Preferential interactions of proteins with solvent components in aqueous amino acid solutions. *Arch. Biochem. Biophys.*, 1983, **224**(1), pp.169-177.
525. Arakawa, T. and Timasheff, S.N. The mechanism of action of Na glutamate, lysine HCl, and piperazine-*N,N'*-bis(2-ethanesulfonic acid) in the stabilization of tubulin and microtubule formation. *J. Biol. Chem.*, 1984, **259**(8), pp.4979-4986.
526. Arakawa, T. and Timasheff, S.N. The stabilization of proteins by osmolytes. *Biophys. J.*, 1985, **47**(3), pp.411-414.
527. Vagenende, V., Yap, M.G.S. and Trout, B.L. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. *Biochem.*, 2009, **48**(46), pp.11084-11096.
528. Gekko, K. and Morikawa, T. Thermodynamics of polyol-induced thermal stabilization of chymotrypsinogen. *J. Biochem*, 1981, **90**(1), pp.51-60.
529. Kaushik, J.K. and Bhat, R. Thermal stability of proteins in aqueous polyol solutions: role of the surface tension of water in the stabilizing effect of polyols. *J. Phys. Chem. B.*, 1998, **102**(36), pp.7058-7066.
530. Haque, I., Singh, R., Moosavi-Movahedi, A.A. and Ahmad, F. Effect of polyol osmolytes on  $\Delta G(D)$ , the Gibbs energy of stabilisation of proteins at different pH values. *Biophys. Chem.*, 2005, **117**(1), pp.1-12.
531. Mishra, R., Seckler, R. and Bhat, R. Efficient refolding of aggregation-prone citrate synthase by polyol osmolytes: how well are protein folding and stability aspects coupled? *J. Biol. Chem.*, 2005, **280**(16), pp.15553-15560.
532. Street, T.O., Bolen, D.W. and Rose, G.D. A molecular mechanism for osmolyte-induced protein stability. *Proc. Natl. Acad. Sci.*, 2006, **103**(38), pp.13997-14002.
533. Petersen, S.B., Jonson, V., Fojan, P., Wimmer, R. and Pedersen, S. Sorbitol prevents the self-aggregation of unfolded lysozyme leading to an up to 13 °C stabilisation of the folded form. *J. Biotechnol.*, 2004, **114**(3), pp.269-278.
534. Bahrami, H., Zahedi, M., Moosavi-Movahedi, A., Azizian, H. and Amanlou, M. Theoretical investigation of interaction of sorbitol molecules with alcohol dehydrogenase in aqueous solution using molecular dynamics simulation. *Cell Biochem. Biophys.*, 2011, **59**(2), pp.79-88.

535. Chi, E., Krishnan, S., Kendrick, B., Chang, B., Carpenter, J. and Randolph, T. Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor. *Protein Sci.*, 2003, **12**(5), pp.903-913.
536. Golovanov, A.P., Hautbergue, G.M., Wilson, S.A. and Lian, L.-Y. A simple method for improving protein solubility and long-term stability. *J. Am. Chem. Soc.*, 2004, **126**(29), pp.8933-8939.
537. Arakawa, T. and Tsumoto, K. The effects of arginine on refolding of aggregated proteins: not facilitate refolding, but suppress aggregation. *Biochem. Biophys. Res. Commun.*, 2003, **304**(1), pp.148-152.
538. Cava, F., Lam, H., De Pedro, M. and Waldor, M. Emerging knowledge of regulatory roles of D-amino acids in bacteria. *Cell. Mol. Life Sci.*, 2011, **68**(5), pp.817-831.
539. Lam, H., Oh, D.-C., Cava, F., Takacs, C.N., Clardy, J., De Pedro, M.A. and Waldor, M.K. D-amino acids govern stationary phase cell wall remodeling in bacteria. *Science*, 2009, **325**(5947), pp.1552-1555.
540. Zhang, G. and Sun, H.J. Racemization in reverse: evidence that D-amino acid toxicity on Earth is controlled by bacteria with racemases. *PLoS ONE*, 2014, **9**(3), p.e92101.
541. Roberts, C.J. Therapeutic protein aggregation: mechanisms, design, and control. *Trends Biotechnol.*, 2014, **32**(7), pp.372-380.
542. De Marco, A. Biotechnological applications of recombinant single-domain antibody fragments. *Microb. Cell Fact.*, 2011, **10**, pp.44-44.
543. Muruganandam, A., Tanha, J., Narang, S. and Stanimirovic, D. Selection of phage-displayed llama single-domain antibodies that transmigrate across human blood-brain barrier endothelium. *FASEB J.*, 2001, **16**(2), pp.240-242.
544. Kijanka, M., Warnders, F.-J., El Khattabi, M., Lub-De Hooge, M., Van Dam, G., Ntziachristos, V., De Vries, L., Oliveira, S. and Van Bergen En Henegouwen, P.P. Rapid optical imaging of human breast tumour xenografts using anti-HER2 V<sub>H</sub> site-directly conjugated to IRDye 800CW for image-guided surgery. *Eur. J. Nucl. Med. Mol. Imaging*, 2013, **40**(11), pp.1718-1729.
545. Lauwereys, M., Arbabi Ghahroudi, M., Desmyter, A., Kinne, J., Hölzer, W., De Genst, E., Wyns, L. and Muyldermans, S. Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *EMBO J.*, 1998, **17**(13), pp.3512-3520.
546. Maggio, E. Use of excipients to control aggregation in peptide and protein formulations. *J. Excipients Food Chem.*, 2010, **1**(2), pp.40-49.
547. Purohit, V., Middaugh, C. and Balasubramanian, S. Influence of aggregation on immunogenicity of recombinant human factor VIII in hemophilia A mice. *J. Pharm. Sci.*, 2006, **95**(2), pp.358-371.
548. Rouet, R., Lowe, D., Dudgeon, K., Roome, B., Schofield, P., Langley, D., Andrews, J., Whitfeld, P., Jermutus, L. and Christ, D. Expression of high-affinity human antibody fragments in bacteria. *Nat. Protocols*, 2012, **7**(2), pp.364-373.
549. Bondos, S.E. and Bicknell, A. Detection and prevention of protein aggregation before, during, and after purification. *Anal. Biochem*, 2003, **316**(2), pp.223-231.
550. Misawa, S. and Kumagai, I. Refolding of therapeutic proteins produced in *Escherichia coli* as inclusion bodies. *Biopolymers Peptide Sci.*, 1999, **51**(4), pp.297-307.
551. Clark, E.D.B. Protein refolding for industrial processes. *Curr. Opin. Biotechnol.*, 2001, **12**(2), pp.202-207.
552. Suenaga, M., Ohmae, H., Tsuji, S., Itoh, T. and Nishimura, O. Renaturation of recombinant human neurotrophin-3 from inclusion bodies using a suppressor agent of aggregation. *Biotechnol. Appl. Biochem.*, 1998, **28**(2), pp.119-124.
553. Cho, B.-K., Federowicz, S., Park, Y.-S., Zengler, K. and Palsson, B.Ø. Deciphering the transcriptional regulatory logic of amino acid metabolism. *Nat. Chem. Biol.*, 2012, **8**(1), pp.65-71.
554. Hershberg, R. and Margalit, H. Co-evolution of transcription factors and their targets depends on mode of regulation. *Genome Biol.*, 2006, **7**, p.R62.

555. Cho, B.-K., Barrett, C.L., Knight, E.M., Park, Y.S. and Palsson, B.Ø. Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. *Proc. Natl. Acad. Sci.*, 2008, **105**(49), pp.19462-19467.
556. Tani, T.H., Khodursky, A., Blumenthal, R.M., Brown, P.O. and Matthews, R.G. Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. *Proc. Natl. Acad. Sci.*, 2002, **99**(21), pp.13471-13476.
557. Celis, T.F.R., Rosenfeld, H.J. and Maas, W.K. Mutant of *Escherichia coli* K-12 defective in the transport of basic amino acids. *J. Bacteriol.*, 1973, **116**(2), pp.619-626.
558. Celis, R.T.F. Repression and activation of arginine transport genes in *Escherichia coli* K 12 by the ArgP protein. *J. Mol. Biol.*, 1999, **294**(5), pp.1087-1095.
559. Wang, L., Lai, L., Ouyang, Q. and Tang, C. Flux balance analysis of ammonia assimilation network in *E. coli* predicts preferred regulation point. *PLoS ONE*, 2011, **6**(1), p.e16362.
560. Foster, J.W. *Escherichia coli* acid resistance: tales of an amateur acidophile. *Nat. Rev. Microbiol.*, 2004, **2**(11), pp.898-907.
561. Smith, J.L. The role of gastric acid in preventing foodborne disease and how bacteria overcome acid conditions. *J. Food Prot.*, 2003, **66**(7), pp.1292-1303.
562. Small, P., Blankenhorn, D., Welty, D., Zinser, E. and Slonczewski, J.L. Acid and base resistance in *Escherichia coli* and *Shigella flexneri*: role of rpoS and growth pH. *J. Bacteriol.*, 1994, **176**(6), pp.1729-1737.
563. Castanie-Cornet, M.-P., Penfound, T.A., Smith, D., Elliott, J.F. and Foster, J.W. Control of Acid Resistance in *Escherichia coli*. *J. Bacteriol.*, 1999, **181**(11), pp.3525-3535.
564. Capitani, G., Biase, D.D., Aurizi, C., Gut, H., Bossa, F. and Grütter, M.G. Crystal structure and functional analysis of *Escherichia coli* glutamate decarboxylase. *EMBO J.*, 2003, **22**(16), pp.4027-4037.
565. Richard, H. and Foster, J.W. *Escherichia coli* glutamate- and arginine-dependent acid resistance systems increase internal pH and reverse transmembrane potential. *J. Bacteriol.*, 2004, **186**(18), pp.6032-6041.
566. Tuttle, A. and Gest, H. Induction of morphological aberrations in *Rhodospirillum Rubrum* by D-amino acids. *J. Bacteriol.*, 1960, **79**(2), pp.213-216.
567. Chang, Y.-F. and Adams, E. D-lysine catabolic pathway in *Pseudomonas putida*: interrelations with L-lysine catabolism. *J. Bacteriol.*, 1974, **117**(2), pp.753-764.
568. Roesch, P.L., Redford, P., Batchelet, S., Moritz, R.L., Pellett, S., Haugen, B.J., Blattner, F.R. and Welch, R.A. Uropathogenic *Escherichia coli* use D-serine deaminase to modulate infection of the murine urinary tract. *Mol. Microbiol.*, 2003, **49**(1), pp.55-67.
569. Goldberg, D.S., Bishop, S.M., Shah, A.U. and Sathish, H.A. Formulation development of therapeutic monoclonal antibodies using high-throughput fluorescence and static light scattering techniques: role of conformational and colloidal stability. *J. Pharm. Sci.*, 2011, **100**(4), pp.1306-1315.
570. Kim, H.-J., Chae, S.-C., Lee, D.-K., Chromy, B., Lee, S.C., Park, Y.-C., Klein, W.L., Krafft, G.A. and Hong, S.-T. Selective neuronal degeneration induced by soluble oligomeric amyloid beta protein. *FASEB J.*, 2003, **17**(1), pp.118-120.
571. Last, N.B., Rhoades, E. and Miranker, A.D. Islet amyloid polypeptide demonstrates a persistent capacity to disrupt membrane integrity. *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**(23), pp.9460-9465.
572. Kroth, H., Ansaloni, A., Varisco, Y., Jan, A., Sreenivasachary, N., Rezaei-Ghaleh, N., Giriens, V., Lohmann, S., Lopez-Deber, M.P., Adolfsson, O., Pihlgren, M., Paganetti, P., Froestl, W., Nagel-Steger, L., Willbold, D., Schrader, T., Zweckstetter, M., Pfeifer, A., Lashuel, H.A. and Muhs, A. Discovery and structure activity relationship of small molecule inhibitors of toxic  $\beta$ -amyloid<sub>42</sub> fibril formation. *J. Biol. Chem.*, 2012, **287**(41), pp.34786-34800.
573. Nie, Q., Du, X.-G. and Geng, M.-Y. Small molecule inhibitors of amyloid  $\beta$  peptide aggregation as a potential therapeutic strategy for Alzheimer's disease. *Acta Pharmacol. Sin.*, 2011, **32**(5), pp.545-551.

- 
574. Bulawa, C.E., Connelly, S., Devit, M., Wang, L., Weigel, C., Fleming, J.A., Packman, J., Powers, E.T., Wiseman, R.L., Foss, T.R., Wilson, I.A., Kelly, J.W. and Labaudinière, R. Tafamidis, a potent and selective transthyretin kinetic stabilizer that inhibits the amyloid cascade. *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, pp.9629-9634.
575. Espargaro, A., Sabate, R. and Ventura, S. Thioflavin-S staining coupled to flow cytometry. A screening tool to detect *in vivo* protein aggregation. *Mol. Biosyst.*, 2012, **8**(11), pp.2839-2844.
576. Romero, D., Sanabria-Valentín, E., Vlamakis, H. and Kolter, R. Biofilm inhibitors that target amyloid proteins. *Chem. Biol.*, 2013, **20**(1), pp.102-110.
577. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.*, 2003, **67**(4), pp.593-656.
578. De Marco, A. Strategies for successful recombinant expression of disulfide bond-dependent proteins in *Escherichia coli*. *Microb. Cell Fact.*, 2009, **8**(1), pp.1-18.
579. Ma, B. and Nussinov, R. Simulations as analytical tools to understand protein aggregation and predict amyloid conformation. *Curr. Opin. Chem. Biol.*, 2006, **10**(5), pp.445-452.
580. Becker, J.W. and Reeke, G.N. Three-dimensional structure of beta 2-microglobulin. *Proc. Natl. Acad. Sci.*, 1985, **82**(12), pp.4225-4229.
581. Trinh, C.H., Smith, D.P., Kalverda, A.P., Phillips, S.E.V. and Radford, S.E. Crystal structure of monomeric human  $\beta$ -2-microglobulin reveals clues to its amyloidogenic properties. *Proc. Natl. Acad. Sci.*, 2002, **99**(15), pp.9771-9776.
582. Eichner, T., Kalverda, A.P., Thompson, G.S., Homans, S.W. and Radford, S.E. Conformational conversion during amyloid formation at atomic resolution. *Mol. Cell*, 2011, **41**(2), pp.161-172.
583. Ivanova, M.I., Sawaya, M.R., Gingery, M., Attinger, A. and Eisenberg, D. An amyloid-forming segment of  $\beta$ 2-microglobulin suggests a molecular model for the fibril. *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**(29), pp.10584-10589.
584. Karamanos, Theodoros k., Kalverda, Arnout p., Thompson, Gary s. and Radford, Sheena e. Visualization of transient protein-protein interactions that promote or inhibit amyloid assembly. *Mol. Cell*, 2014, **55**(2), pp.214-226.
585. Gejyo, F., Odani, S., Yamada, T., Honma, N., Saito, H., Suzuki, Y., Nakagawa, Y., Kobayashi, H., Maruyama, Y., Hirasawa, Y., Suzuki, M. and Arakawa, M.  $\beta$ 2-microglobulin: A new form of amyloid protein associated with chronic hemodialysis. *Kidney Int.*, 1986, **30**(3), pp.385-390.
586. Westermark, G.T. and Westermark, P. Prion-like aggregates: infectious agents in human disease. *Trends Mol. Med.*, 2010, **16**(11), pp.501-507.