Characterisation of a bundling protein involved in the self-assembly of bioelastomeric snail egg cases

Loke, Jun Jie

2017

Loke, J. J. (2017). Characterisation of a bundling protein involved in the self-assembly of bioelastomeric snail egg cases. Master's thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/69474>

<https://doi.org/10.32657/10356/69474>

**CHARACTERISATION OF A BUNDLING PROTEIN
INVOLVED IN THE SELF-ASSEMBLY OF
BIOELASTOMERIC SNAIL EGG CASES**

LOKE JUN JIE

School of Materials Science and Engineering

2016

**CHARACTERISATION OF A BUNDLING PROTEIN
INVOLVED IN THE SELF-ASSEMBLY OF
BIOELASTOMERIC SNAIL EGG CASES**

LOKE JUN JIE

SCHOOL OF MATERIALS SCIENCE AND ENGINEERING

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Master of Engineering

2016

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other University or Institution.

August 3rd, 2016

.....
Date

.....
Loke Jun Jie

Abstract

The carnivorous marine snail *Pugilina cochlidium* is found predominantly on the sea shores of the South-East Asia region. This sea creature lays a string of egg capsules up to a metre in length which houses its embryos. The egg capsule material must be able to withstand and resist the harsh and relentless tidal forces of the ocean in order for the off-springs to develop properly and survive.

The egg capsule was found to be made up of egg capsule protein precursors which domains are predominately filamentous made of coiled-coils. Four laminar protein sheets were also found to be oriented perpendicular to each other. Mechanical testing of the egg capsule had shown that the biomaterial is capable of absorbing shock, where it fully extends upon stretching and is able to recover to its original shape almost instantaneously without significant damage.

Investigations of the egg capsule precursor proteins hierarchical native state were carried out using different electrophoresis techniques. This allows a relative comparison of the molecular weight of cross-linked egg capsule protein at different stages in the egg laying process, and how the egg capsule protein precursor might self-assemble prior to secretion. The proteins self-assemble into 200 kDa bundles which are further assembled into fibres with a protein cross-linker, presumably the *Pugilina* Bundling Protein. *Pugilina cochlidium* Egg Capsule Protein 1, 2 and 3 were found to be forming the 200 kDa bundles but not for Protein 4 and 5.

The genetic sequence of *Pugilina* Bundling Protein was transformed into the *E. coli* recombinant system and expressed using Terrific broth to maximise its yield. Protein purification of the solubilised inclusion bodies was carried out using Reversed Phase High Performance Liquid Chromatography. Electrophoresis and Matrix Assisted Laser Desorption/Ionisation Time-of-Flight gave novel insights into the protein ability to form homo-dimers and oligomers.

The purified *Pugilina* Bundling Protein was refolded with L-Arginine as an aggregate suppressor while cysteine/cystine redox pair aided in the proper reshuffling of disulphide bonds. The refolded protein formed dispersed aggregates when mixed with native egg capsule proteins, suggesting that it is partly responsible for the egg capsule hierarchical assembly and cross-linking.

Acknowledgements

I would like to thank Assistant Professor Ali Miserez for all his support, motivation and encouragement in pursuing my Masters degree, and also allowing me to continue this research project from where my Bachelor degree Final Year Project ended.

I received many valuable advices on different techniques and ideas for my experiments. Firstly to my mentor Mr Gavin Tay for his teamwork on this ambiguous project previously, Mr Fu Tianpei for his patient mentoring on the *Escherichia coli* recombinant system, Ms Clarinda Sutano on chromatography issues, Dr Vitali Lipik for advices on mass spectrometry, Dr Deepankumar for his expertise and experience on microbiology and optimization, Ms Akshita Kumar for her help on EMBL-BLAST protein homology models, Dr Shawn Hoon from A*STAR for the *Pugilina* egg capsule protein glycerol stock, and Dr Paul Guerette for his vision and direction of this project initially.

In addition, I would also like to thank the three Singapore Polytechnic students during their final year project for developing new experiment methodologies and data. To all my other colleagues who had also helped me directly and indirectly. Last but not least to my ever supportive parents and friends who sees me through some tough times on my Masters research programme.

Table of Contents

Abstract	i
Acknowledgements	iii
Table of Contents	v
Table Captions	ix
Figure Captions	xi
Abbreviations	xvii
Chapter 1. Introduction	1
1.1 Background	2
1.2 Hypothesis.....	3
1.3 Objectives and Scope	3
1.4 Dissertation Overview.....	4
References	5
 Chapter 2. Literature Review	 7
2.1 Background of <i>Pugilina cochlidium</i>	8
2.2 Behaviour of <i>Pugilina cochlidium</i>	9
2.3 Egg capsule mechanical properties	11
2.4 Surface Characterisation of <i>Pugilina cochlidium</i> egg capsules	14
2.5 Structure of coiled-coils and egg capsule proteins.....	15
References	19
 Chapter 3. Experimental methodologies	 23
3.1 Preparation of native crude extract from <i>P. cochlidium</i> nidamental gland.....	24
3.2 Native state of egg capsule proteins using electrophoresis techniques.....	24
3.2.1 Tris-Glycine SDS-PAGE.....	24
3.2.2 Tris-Tricine SDS-PAGE	25
3.2.3 Neutral pH SDS-PAGE.....	25
3.2.4 Blue Native Polyacrylamide Gel Electrophoresis.....	25

3.2.5 Non-Reducing Sodium Dodecyl Sulfate Polyacrylamide Gel	
Electrophoresis	26
3.2.6 Two-Dimensional Sodium Dodecyl Sulfate Polyacrylamide Gel	
Electrophoresis	26
3.2.7 Ethylene Glycol Diacrylate cross-linked polyacrylamide gel	26
3.3 Polyacrylamide in-gel protein fixation	27
3.3.1 Blue Silver Staining.....	27
3.3.2 Zinc-Imidazole Negative Stain.....	28
3.4 Transformation of <i>Pugilina</i> Bundling Protein gene.....	28
3.5 Expression Systems using Luria Bertani and Terrific broth.....	28
3.6 Cell Lysis	29
3.7 Protection of cysteines by reversible thiolation with 2,2'-Dithiodipyridine.....	29
3.8 Purification of <i>Pugilina</i> Bundling Protein by High Performance Liquid	
Chromatography	31
3.8.1 Size Exclusion Chromatography	31
3.8.2 Reversed Phase Chromatography	31
3.8.3 Optimisation of Reversed Phase Chromatography.....	33
3.9 Purification by Dialysis.....	35
3.10 Matrix Assisted Laser Desorption Ionisation-Time of Flight Mass	
Spectroscopy	35
3.11 Refolding buffer with redox pair	35
3.12 Characterisation of PcBP – PcECPs aggregates	35
3.12.1 Optical Microscopy	36
References.....	37
 Chapter 4. Oligomeric states of egg capsule proteins.....	 43
4.1 Introduction.....	44
4.2 Experimental Methods.....	45
4.2.1 Preparation of crude extract from <i>P. cochlidium</i> nidamental gland	45
4.2.2 Blue Native Polyacrylamide Gel electrophoresis	45
4.2.3 Tris-Glycine SDS-PAGE.....	47
4.2.4 Tris-Tricine SDS-PAGE.....	47
4.2.5 Two Dimensional SDS-PAGE	48

4.2.6 Ethylene Glycol Diacrylate polyacrylamide gel	49
4.3 Results and Discussion.....	49
4.3.1 Protein concentration of crude extracts.....	49
4.3.2 Blue native polyacrylamide gel electrophoresis	50
4.3.3 Non-reducing SDS-PAGE	52
4.3.4 Two dimensional SDS-PAGE.....	55
4.3.5 Ethylene Glycol Diacrylate cross-linked polyacrylamide gel	56
4.4 Conclusions	59
References	60
 Chapter 5. Expression and Purification of <i>Pugilina</i> Bundling Protein.....	61
5.1 Introduction.....	62
5.2 Protein sequence of <i>Pugilina</i> Bundling Protein	64
5.3 Homology Modelling of <i>Pugilina</i> Bundling Protein	68
5.4 Experimental methods.....	72
5.4.1 Transformation of <i>Pugilina</i> Bundling Protein gene.....	72
5.4.2 Expression Systems using Luria Bertani and Terrific broth.....	72
5.4.3 Cell Lysis	73
5.4.4 Purification by High Performance Liquid Chromatography	73
5.4.5 Purification by Dialysis.....	75
5.4.6 Neutral pH SDS-PAGE.....	75
5.4.7 Matrix Assisted Laser Desorption Ionisation-Time of Flight Mass Spectroscopy	76
5.4 Results and Discussions	76
5.4.1 Efficiency of expression systems.....	76
5.4.2 Purification by Size Exclusion High Performance Liquid Chromatography	77
5.4.3 Purification by Reversed Phase High Performance Liquid Chromatography	79
5.4.4 Optimisation of Reversed Phase High Performance Liquid Chromatography	81
5.4.5 Purity of Reversed Phased Fraction by MALDI-ToF	85
5.4.6 Purification by dialysis	89
5.5 Conclusions	91

References.....	93
 Chapter 6. Refolding of <i>Pugilina</i> Bundling Protein.....	 95
6.1 Introduction.....	96
6.2 Experimental Methods.....	97
6.2.1 Refolding Buffers	97
6.2.2 Neutral pH Polyacrylamide Gel Electrophoresis.....	98
6.2.3 Refolded PcBP - PcECPs aggregation.....	99
6.2.4 Optical Microscopy of Refolded PcBP - PcECPs aggregates	99
6.3 Results and Discussion	99
6.3.1 Neutral PAGE Electrophoresis of refolded PcBP	99
6.3.2 Optical Microscopy of Refolded PcBP - PcECPs aggregates	100
6.4 Conclusions.....	101
References.....	102
 Chapter 7. Discussion and Future Work.....	 105
7.1 General discussion	106
7.2 Circular Dichroism.....	107
7.3 Fourier Transform Infrared Spectroscopy	108
7.4 Rheology	110
7.5 Surface Characterisation	110
7.6 Surface Sclerotisation	111
References.....	112

Table Captions

Table 3.1 Summary list of different acids, organic solvent and additives formulated to purify PcBP using reversed phase column. An X indicate that the purification is either of poor separation or unstable for the protein, and an O indicate that the method can be optimised further. A dash indicates the combination was not tested.	33
Table 4.1. Blue native gel and running buffers protocol.....	46
Table 4.2. Tris-Glycine gel and running buffers protocol.....	47
Table 4.3. Tris-Tricine gel and running buffers protocol.	48
Table 5.1. Amino acid sequence of PcBP matched in <i>Pugilina cochlidium</i> transcriptome database. The characteristic of PcBP was derived from ExPASy ProtParam.	65
Table 5.2. Predicted disulphide bonds ordered by probability in descending order. ...	70
Table 5.3. Binary solvent system gradient timings and elution buffer composition.	74
Table 5.4. Neutral pH SDS-PAGE protocol for Bis-Tris gel buffer and 10x cathode and anode buffer.	75
Table 5.5. Wet weight of harvested cells per 100 ml with different broth and parameters. Terrific broth was able to harvest up to 1.22 g of wet weight cells compared to LB broth.	77
Table 6.1. Refolding buffer composition containing L-Arginine as aggregation suppressor and cysteine/cystine redox pair for proper disulphide reshuffling.....	98
Table 6.2. Neutral pH SDS-PAGE protocol for Bis-Tris gel buffer and 10x cathode and anode buffer.	98

Figure Captions

Figure 2.1 **A)** Relative size comparison of *Busycotypus canaliculatus* (left), *Pugilina cochlidium* (right), **B)** Relative egg capsule size comparison of *Busycotypus canaliculatus* (left), *Pugilina cochlidium* (right), **C)** *Pugilina cochlidium* laying a series of egg capsule on a rock, **D)** *Pugilina cochlidium* beige coloured nidamental gland with creases. 10

Figure 2.2. Stress-strain hysteresis curve for engineering curve and true tensile curve of *Busycotypus canaliculatus* egg capsule. The true stress-strain's true modulus was obtained from the tangent during partial unloading/reloading cycles on the curve. 12

Figure 2.3. Schematic diagram depicting the transition of α -helix to β -sheet when an external load is applied. The transition is reversible from β -sheet to α -helix when the external load is removed. 12

Figure 2.4. Wide-angle X-Ray Scattering peak of *B. canaliculatus* egg capsule at 0%, 70%, 120% strain. 1.03 nm corresponds to the mean distance between helices and the 0.52 nm peak along the meridian owing to the coiled-coil axis. The β -sheet formation is indicated by the 0.465 nm equatorial reflections and 0.33 nm comes from the translational rise between residues along the β -sheet.^{3,12,13} 13

Figure 2.5. Optical microscopy of the egg capsule first three layers. Images are focused stack for improved depth of field at high magnifications. Dark field image B, C and D are taken at 500x magnification. **A)** An overlay of bright and dark field image at 2.5x magnification of an egg capsule with developing embryos. **B)** Deep and tough surface ridges running parallel to each other, suggesting the aid in diverting water current along the surface away from the egg capsule. **C)** Fine and smooth protein fibres in the second layer oriented perpendicularly to the third layers which consist of thick and coarse bundles of protein fibres. **D)** Threads of protein cross-linked to neighbouring fibre bundles when lightly stretched. 15

Figure 2.6. Schematic representation of a dimeric coiled coil.²⁰ Residue a/a' and d/d' forms the hydrophobic core, e/e' and g/g' forms the ionic interaction and b/b', c/c' and f/f' residues are hydrophilic and exposed to the solvent. 16

- Figure 2.7.** A model of an IF dimer. The middle section consists of an α -helical central "rod" domain and flanked by non- α -helical "head" and "tail" domains. 16
- Figure 2.8.** **A)** An IF representation with an α -helical rod in the middle and non- α -helical "head" and "tail" domains. **B)** Comparison of the predicted coiled-coil domains of BcECP and PcECP with the IF model. BcECP-1a and -1b has a predicted coiled-coil C-terminus which differs from that of PcECP and IF..... 17
- Figure 2.9.** **A)** SDS-PAGE of *B. canaliculatus* and *P. cochlidium* nidamental gland crude extract. **B)** MALDI-ToF spectra of *P. cochlidium* showing 5 distinct peaks with molecular weights from 42.0 kDa to 54.3 kDa. **C)** MALDI-ToF spectra of *B. canaliculatus* showing 3 peaks from 48 kDa to 52.5 kDa. 18
- Figure 3.1.** Reduction and protection of *Pugilina* Bundling Protein sulfhydryl group with reversible thiol reagent 2,2'-Dithiodipyridine. The mechanism of mixed disulphide was achieved with saturating concentration of 2,2'-Dithiodipyridine. 2-Mercaptopyridine can be detected at 343 nm wavelength..... 30
- Figure 3.2.** Solubilised *Pugilina* Bundling Protein inclusion bodies solution in 6 M urea **A)** before addition of 10 mM 2,2'-Dithiodipyridine showing translucent orange-brown colour and **B)** addition of 10 mM 2,2'-Dithiodipyridine showing translucent pale yellow colour after complete reaction. 30
- Figure 3.3.** **A)** Schematic diagram of activated silanols at pH above its pKa. **B)** Protonation of silanol at low pH thus cancel the effects of ionic interaction. **C)** Endcapping of silanols with trimethylsilyl to reduce the detrimental effects of silanol coupling with analytes. The bulky group of trimethylsilyl prevents complete derivatisation of silanols due to steric hindrance, thus the use of ionic liquid can help to reduce the secondary interactions. **D)** Representation of 1-Butyl-3-methylimidazolium hexafluorophosphate. 34
- Figure 4.1.** Schematic diagram of manually excised protein bands for 2D SDS-PAGE as an example. In this schematic, a native protein comprised of 4 major protein bands, red, orange, yellow and green. When denatured, the major protein bands disassociate into its sub-units and monomers. By isolating the major bands in the native lane, the sub-units can be clearly determined in the second SDS-PAGE gel..... 49

- Figure 4.2.** **A)** Blue native PAGE of bovine serum albumin and native crude extract. **B)** Peak detection of native crude extract was done using GelAnalyser 2010a (Istvan Lazar, 2010). **C)** A relative mobility plot against molecular weight reference of bovine serum albumin and the estimated native crude extract molecular weight. Red line represents the trendline for the three data points and the blue line represents the extrapolated trendline. 50
- Figure 4.3.** A reducing 10% SDS-PAGE comparing the protein subunits in both native and denatured crude extract. Lane 1 contains the molecular weight reference. Lane 2 contains the native crude extract treated with 2% v/v β -mercaptoethanol and heat. Lane 3 contains denatured crude sample in 6 M urea and 5% v/v acetic acid..... 51
- Figure 4.4.** Non-denaturing and non-reducing SDS-PAGE gel of 9 native crude extracts from different snails. Crude extract for *Pugilina* (PC) and *Busycotypus* (BC) nomenclature was chosen for the respective species. 52
- Figure 4.5.** **A)** Gradient 4% to 16% SDS-PAGE gel with molecular weight ladder and 9 native crude extract for comparison. **B)** Peak detection of lane 8 native crude extract was done using GelAnalyser 2010a (Istvan Lazar, 2010). **C)** A relative mobility plot against Biorad Precision Plus Protein™ All Blue Prestained standard and the estimated native crude extract molecular weight in Lane 8. 53
- Figure 4.6.** Comparison of native crude extract in both denaturing and non-denaturing conditions using NuPAGE® electrophoresis..... 54
- Figure 4.7.** Two dimensional SDS-PAGE of 200 kDa and PcECP 4 and 5 excised gel bands. The native crude extract was denatured and used as the control in Lane 4. 56
- Figure 4.8.** Zinc-imidazole negative stain of an ethylene glycol diacrylate polyacrylamide gel against a black background. The 200 kDa band across the nine lanes were excised for gel digestion with 500 mM ammonium hydroxide and protein extracted using chloroform-methanol precipitation..... 57
- Figure 4.9.** Large format 4% - 16% Tricine SDS-PAGE gel loaded with native crude extract as reference and EDA extracted 200 kDa band in lanes 6, 7 and 8. 58

- Figure 5.1.** A) *P. cochlidium* native crude extract in 6 M Urea, 5% v/v Acetic acid, 10 mM TCEP, pH 2.4 which formed dispersed aggregates B) *P. cochlidium* native crude extract in native solubilisation buffer, pH 7.0 which formed a dense soft gel. 63
- Figure 5.2.** PcBP amino acid sequence with and without the signal peptide, its theoretical molecular weight and isoelectric point. 66
- Figure 5.3.** MALDI-ToF Mass Spectra of PcBP with 11751 Da and 13690 Da corresponding to the bundling protein with possibly its signal peptide that has yet to be cleaved. 67
- Figure 5.4.** A) I-TASSER model, B) QUARK model, C) EMBL-Blast model. Red and yellow colour represents α -helices and β -sheets respectively..... 69
- Figure 5.5.** Linearized 3D models of I-TASSER, QUARK and EMBL-Blast redrawn in AutoCAD 2010. 71
- Figure 5.6.** A) Size Exclusion column performance of standards. Ribonuclease A was chosen as the closest reference to PcBP. B) Fraction were collect starting from 6th to 9th minute. 78
- Figure 5.7.** Electrophoresis of size exclusion HPLC fraction 6, 7 and 8. Fraction 7 contains most of the eluted PcBP. Purity of the sample is poor due to broad overlapping peaks suggesting secondary interactions with the stationary phase..... 78
- Figure 5.8.** Reversed Phase HPLC chromatogram of PcBP using acetonitrile gradient. Fractions 26, 36 to 42 were collected for electrophoresis to determine the quality of separation. 79
- Figure 5.9.** Reversed Phase fraction 26, 36 to 42 of PcBP on NuPAGE® SDS-PAGE gel. Lane 4 Fraction 37 shows that PcBP is rather pure, but with formation of homo-oligomers suggesting the oxidation of cysteines during sample preparation stage and subsequently catalysed by trifluoroacetic acid. The homo-oligomers could not be reduced back to it monomer..... 80
- Figure 5.10.** A) Chromatogram of PcBP purification on Kromasil C18 reversed phase using 0.1% TFA and acetonitrile gradient. Peak tailing of PcBP is significant which indicates strong secondary interactions with the stationary phase. B) Retention effects

of PcBP was studied using acetonitrile and methanol in 0.1% TFA. Methanol elution strength is about 25% weaker than acetonitrile due to the longer retention times, but shows much better separation. **C)** Retention effects of PcBP using methanol in 0.1% TFA and 0.1% formic acid. The use of different acid has different selectivity on the elution time of different peaks. The use of TFA increases the retention time of PcBP due to the increased hydrophobicity from the anion. **D)** Resolution of peaks with and without the addition of 100 mM sodium chloride. Sodium chloride screens the charges of the protein and unreacted silanol which minimises ionic interaction. The use of salt greatly improves the separation and resolution of the peaks. 83

Figure 5.11. Chromatogram of PcBP using binary solvent system. The low elution strength of methanol was used in the initial purification of PcBP to selectively elute less hydrophobic contaminant proteins while keeping PcBP in the column. Subsequently acetonitrile was used to wash out all the retained proteins. 84

Figure 5.12. Binary solvent system Reversed Phase HPLC fraction 32 to 35 of PcBP on NuPAGE® SDS-PAGE gel. Fraction 32 shows the highly purified PcBP with low amount of its dimer. Fraction 33 to 35 shows majority of the contaminants successfully separated from PcBP. 85

Figure 5.13. MALDI-ToF of Reversed Phase HPLC Fraction 37 using acetonitrile gradient. Values are in Daltons. Multiple peaks indicates the formation of homo-oligomers and multimers possibly due to the oxidation of PcBP cysteine residues and catalysed by TFA. 87

Figure 5.14. MALDI-ToF of Reversed Phase HPLC Fraction 32 using methanol-acetonitrile gradient. Peak at 11.86 kDa show that PcBP is highly pure with minor contaminants between 10kDa to 11.86 kDa. Mild dimerization of PcBP at 23.8 kDa was detected. 88

Figure 5.15. Lyophilised dialysate using **A)** only 5% v/v acetic acid. The hydrophilic solution prevents efficient dialysis of proteins below the pore size of the membrane, thus minimal PcBP was recovered. **B)** Lyophilised dialysate using 50% v/v ACN, which decreases the hydrophobicity of PcBP, thus it is much easier to dialyse through the membrane. 89

- Figure 5.16.** PcBP in the dialysate after lyophilising shows effective dialysis. Lane 2 is the inclusion bodies crude extract before dialysis and Lane 3 is after dialysis with 50% v/v ACN and 5% v/v acetic acid. Purity of the dialysed sample is good with a single contaminant band slightly above PcBP. The dimer of PcBP was also able to dialyse out. 90
- Figure 5.17.** MALDI-ToF of PcBP purified by dialysis. Intense peak at 10.6 kDa shows the presence of contaminants below 25 kDa which had been dialysed out. 91
- Figure 6.1.** Gel electrophoresis of refolded PcBP. Lane 3 is the purified protein from reversed phase chromatography. Lane 4 shows refolded PcBP in the non-reduced state and Lane 5 shows refolded PcBP in the reduced state..... 100
- Figure 6.2. A)** Optical microscope image of native PcECP solution and refolded PcBP mixture at 2.5x magnification. **B)** Optical microscope image of native PcECP solution and refolded PcBP mixture at 10x magnification. 101
- Figure 7.1.** A circular dichroism spectra of different structural conformation. 108
- Figure 7.2.** A typical FTIR spectrum depicting Amide A region $\sim 3500\text{ cm}^{-1}$, Amide I bands between $1,700\text{--}1,600\text{ cm}^{-1}$, Amide II bands between 1510 cm^{-1} and 1580 cm^{-1} , and Amide III $\sim 1330\text{ cm}^{-1}$ 109
- Figure 7.3.** A typical Amide 1 spectrum envelope consisting of α -helix, β -sheet and random coil structure. Deconvolution is applied to quantitatively estimate the amount of each secondary structure component. 109

Abbreviations

μl	Microlitre
2D SDS-PAGE	Two-dimensional Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
AA	Acetic acid
ACN	Acetonitrile
ATR-FTIR	Attenuated Total Reflectance – Fourier Transform Infrared Spectroscopy
BN-PAGE	Blue Native Polyacrylamide Gel Electrophoresis
CBG-250	Coomassie Brilliant Blue G-250
CD	Circular Dichroism
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
Da	Daltons, molecular weight
DTT	Dithiothreitol
<i>E. coli</i>	Escherichia coli
EDA	Ethylene glycol diacrylate
EDTA	Ethylenediaminetetraacetic acid
g	Gram
HCl	Hydrochloric acid
HPLC	High Pressure Liquid Chromatography
IEF	Isoelectric Focusing
IF	Intermediate Filament
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	Kilo-Daltons, molecular weight
LB	Luria-Bertani
M	Molar
mA	Milliamps
mg	Milligrams
ml	Millilitre
mM	Millimolar
MWL	Molecular Weight Ladder

NH ₄ OH	Ammonium hydroxide
nm	Nanometres
nO β DG	n-Octyl- β -D-Glucopyranoside
PcBP	Pugilina cochlidium Bundling Protein
PcECP	Pugilina cochlidium Egg Capsule Proteins
RPM	Revolutions per minute
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SEM	Scanning Electron Microscope
TB	Terrific Broth
TCEP	<i>tris</i> (2-carboxyethyl)phosphine
TEM	Transmission Electron Microscope
TFA	Trifluoroacetic acid
V	Volts
xg	Centrifugal gravitational force
β ME	2-Mercaptoethanol, Beta-Mercaptoethanol

Chapter 1.

Introduction

*The egg capsule of *Pugilina cochlidium* has been found to be a unique biological material which is able to withstand ultraviolet radiation and provide shock absorbing properties to protect its developing embryos. At the same time, the egg capsule also allows nutrients, metabolic waste and oxygen diffusion through the capsule wall. These highly optimised and conserved egg capsule design can be found in almost all gastropods invertebrate species. This chapter briefly introduces the specie *Pugilina cochlidium* and the features of its egg capsule material.*

1.1 Background

The evolution of all living organisms had been going on for billions of years, and will never set to stop. From simplistic cells to complex higher order intelligent beings, the need to adapt and stay resilient is of paramount importance to ensure survival of its own species. Survival of the fittest and the natural balance between prey and predation forces species to adapt quickly or risk being pushed to the brink of extinction. Those that survive will have to evolve with the ever changing hostile environment.

Often organisms find themselves on a two pronged approach, to hunt or deter predators to survive, thus the need to develop offensive and defensive mechanisms for themselves. Naturally, when evolution creates either an offensive or defensive apparatus which had been proven effective for millions of years, it will not want to change. That is, the system is conserved, but yet ready to evolve.

Marine carnivorous snails lay a series of egg capsule which houses thousands of developing embryos. Due to the high mortality rate from environmental factors and predation, the egg capsule's main characteristic remains virtually unchanged across its superfamily of marine snail invertebrates. The egg capsules characterised from *Busycotypus canaliculatus*, *Pugilina cochlidium*, *Buccinum undatum*¹ and *Kelletia kelletii*² etc all have a common α -helical structural motif³ and hierarchical layers of ordered fibrous constituents⁴ in a plywood structure. The highly conserved macrostructure is evident that this material organisation is proven to be highly effective as a defence mechanism for the developing embryos.

From biology to technology, biomimetics is an area of study where ideas from nature are reversed engineered and applied to solve some of society's complex issues.⁵ The properties of this egg capsule is of great interest, where its shock absorbing properties may see itself one day in therapeutic tissue engineering⁶, as a drug encapsulant or even cartilage and ligament replacement.

1.2 Hypothesis

The egg capsule proteins PcECP 1, 2, 3, 4 and 5 identified in previous studies are known to self-assemble into higher order hierarchical structures. However the intermediate self-assembly stages are poorly understood. The aim is to identify the native states of native egg capsule precursor proteins using Blue Native Polyacrylamide Gel Electrophoresis, non-reducing Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis and Two Dimensional Polyacrylamide Gel Electrophoresis.

A new low molecular weight egg capsule protein collectively termed as “*Pugilina* Bundling Protein” was discovered in a polymerised crude extract from a gravid snail. The crude extract which contained this bundling protein formed a soft gel, whereas other crude extracts lacking this protein remained as a solution. It is therefore hypothesized that this protein might hold an important key to the hierarchical self-assembly of the egg capsule.

1.3 Objectives and Scope

The objective of this project is to express the *Pugilina* Bundling Protein in *E. Coli* bacteria host and attempt to refold it. Subsequently, the protein will be mixed with unpolymerised native egg capsule protein in an attempt to self-assemble into the egg capsule biomaterial.

1.3.1 Identify the self-assembled higher hierarchical order of PcECP 1, 2, 3, 4, 5.

1.3.2 Sequencing of *Pugilina* Bundling Protein.

1.3.3 Biochemical characterisation of *Pugilina* Bundling Protein.

1.3.4 Cloning of *Pugilina* Bundling Protein in expression vector.

1.3.5 Expression of *Pugilina* Bundling Protein in *E. Coli*.

1.3.6 Purification of recombinant *Pugilina* Bundling Protein.

1.3.7 Refolding of recombinant *Pugilina* Bundling Protein.

1.3.8 Self-assembly initiation of PcECPs with refolded *Pugilina* Bundling Protein.

1.4 Dissertation Overview

Chapter 1 of this thesis gives an introduction to the *Pugilina cochlidium* Egg Capsule Proteins (ECPs). The bundling protein hypothesized to physically cross-link ECPs, and in turn to drive the assembly of ECP oligomers into a macrostructure is also presented.

Chapter 2 gives an overview of previous studies done on the egg capsule mechanical properties as well as biophysical and biochemical characterisation.

Chapter 3 lays out the methodological approaches of the different experimental techniques to achieve the hypotheses and objectives.

Chapter 4, Chapter 5 and Chapter 6 present the results for the oligomeric state of egg capsules, expression and purification of the bundling protein, and the refolding of purified bundling protein, respectively.

Chapter 7 gives a brief summary of the plausible directions in the research of biomimetic egg capsule material. Different techniques are proposed to characterise the biomimetic egg capsule, such as surface characterisation with scanning electron microscopy.

References

- [1] S. Hunt, *Carbohydrate and Amino-acid Composition of the Egg Capsule of the Whelk *Buccinum undatum* L.*, **1966**, *Nature*, 210, pp. 436 - 437
- [2] H. Scott Rapoport, Robert E. Shadwick, *Reversibly labile, sclerotization-induced elastic properties in a keratin analog from marine snails: whelk egg capsule biopolymer (WECB)*, **2007**, *Journal of Experimental Biology*, 210, pp. 12-26
- [3] Flower, N. E., Geddes, A. J., Rudall, K. M., *Ultrastructure of the fibrous protein from the egg capsules of the whelk *Buccinum undatum**, **1969**, *J. Ultrastruct. Res.* 26, pp. 262 -273
- [4] Rawlings, T. A., *Adaptations to physical stresses in the intertidal zone: the egg capsules of neogastropod mollusks*, **1999**, *Am. Zool.* 39, pp. 230 - 243
- [5] Peter Fratzl, *Biomimetic materials research: what can we really learn from nature's structural materials?*, **2007**, *J R Soc Interface*, 4(15), pp. 637-642
- [6] Shin H, Jo S, Mikos AG, *Biomimetic materials for tissue engineering*, **2003**, *Biomaterials.*, 24(24), pp. 4353-64

Chapter 2.

Literature Review

The conserved design of the egg capsule material ensures maximum survivability of its own specie. In this chapter, the snail and its behaviours are described in detail. The complex structure of the egg capsule material which had been studied extensively previously is presented. The predominantly coiled-coil structure of the egg capsule material was able to achieve instantaneous recovery after stress-relaxation cycle. The wall is made up of four distinct layers, each oriented perpendicular to each other stacked in a plywood fashion. These designs were able to withstand the harsh conditions of the marine environment but at the same time provide protection for developing embryos.

2.1 Background of *Pugilina cochlidium*

Pugilina cochlidium, also known as *Volegalea cochlidium* or *Murex cochlidium* (Linnaeus, 1758, World Register of Marine Species)¹ is a genus of large marine gastropod molluscs in the taxonomic family of *Melongenidae*. It is also known as spiral melongena. The family of *Melongenidae* has three shells forms, namely the Whelks, Melon and Crown conches.

Crown and melon conches are more tropically distributed and typically found in brackish or muddy waters near mangrove forests. *P. cochlidium* belongs to the group of the Melon conch and is dispersed in the Indo-Pacific region from the coasts of Sri Lanka to Northern Australia.² All *Melongenidae* groups are carnivorous, feeding on clams, oysters and barnacles. Whelks are generally much larger in size and like their related species, they use their shell lip to pry open shellfish and insert their proboscis to feed.

Pugilina cochlidium lays a long string of egg capsules in the wild with lengths of up to 40 cm. Each individual capsule is connected to each other at the base and anchored to a hard surface with a water-based adhesive that the snail produces. The developing embryos are protected from the harsh sea conditions by the egg capsule walls, which comprises of multiple perpendicularly oriented elastic protein layers. Apart from the shock absorbing properties and impervious barrier against debris, it also serves as a nutrient and waste exchange membrane, and believed to have some form of ultraviolet protection.

The bioelastomeric egg capsules were first discovered from another family of whelks, known as *Busycotypus canaliculatus* in the taxonomy family of *Buccinidae*.¹ This family houses very large sized sea snails and is often known as “true whelks”. True whelks are populated globally from the tropical oceans to the cold arctic sea. They can be found in the intertidal and bathypelagic zones and prefer a solid or sandy surface.²

Both *Pugilina* and *Busycotypus* come from the same superfamily of *Buccinoidea*¹ that comprises very small to large predatory sea snails. Previous research had found that the *Pugilina* egg capsule’s mechanical properties is similar to that of the *Busycotypus*.³ This is not surprising given that both species share the same *Buccinoidea* superfamily.

The elastomeric property of egg capsules has been observed in a variety of snail species, yet this highly conservative biological material underwent some slight mutations in its protein sequence across different families of snails.

Another interesting feature of the egg capsule system, which is not part of this research thesis, is its water-based adhesive. The snail lays a long string of egg capsules which need to be anchored firmly to underwater solid and stable substrates. The adhesive must also be strong enough to withstand delamination from tidal forces. Bivalves, such as mussels, secrete multiple protein-based threads (byssus) that ends in a sticky plaque to anchor the soft body of the mussel to hard submerged surfaces.^{4,5}

2.2 Behaviour of *Pugilina cochlidium*

Snails of *Pugilina* were collected from the monsoon canal^{6,7} in East Coast Park, Singapore, and were kept in aquaria with artificial seawater. They were observed to be social and gregarious creatures, much like mussels which tend to cluster together. This behaviour is believed to impart an evolutionary advantage to the snail, allowing them to survive longer in large colonies. Isolated individuals, which stray away from the main group, have a higher tendency to die faster.

The dominant food source of *Pugilina* are barnacles, but they will devour live mussels (*Perna viridis*) should there be a scarce supply of barnacles. To date there is no evidence on how the snails pry open the bivalves, due to the lack of any clues on the shell indicting a forced entry.⁶

Just before the gestation period, the snail migrates closer to the aquaria tank water inlet, where it will scour the vicinity for a solid rock. Once a proper spot has been found with the strongest water current flowing across the snail body, it will remain almost stationary for a day or two before laying the eggs.

The egg laying process usually lasts around six to ten hours depending on the size of the snail and the length of the egg capsule. During this period, the gravid snail is insensitive to external stimuli and will not retract its soft and vulnerable body even when touched. However, the gravid snail will retract when forcefully removed from the spot or an invasive sharp object, such as a needle, is inserted for sedation. Once the snail senses that it is safe again, it will continue to deposit the remaining egg capsules.

Other gravid snails will come to the same location and deposit their egg capsules just beside the first string. The snails will also lay the egg capsules on each other since they do not favour the aquaria acrylic material.

These behaviours observed are thought to maximize survivability among its own species; clustering in groups, laying eggs near strong water currents to disperse offspring across a larger area, and having the snail readily sacrifice her own life to predators while still having the determination to finish depositing her eggs.

Figure 2.1 A) Relative size comparison of *Busycotypus canaliculatus* (left), *Pugilina cochlidium* (right), B) Relative egg capsule size comparison of *Busycotypus canaliculatus* (left), *Pugilina cochlidium* (right), C) *Pugilina cochlidium* laying a series of egg capsule on a rock, D) *Pugilina cochlidium* beige coloured nidamental gland with creases.

2.3 Egg capsule mechanical properties

The egg capsules are mostly non-entropic elastomeric biomaterials: the large majority of elastomers are governed by entropic elasticity,^{8,9} namely upon the application of an external force chains extend, and leading to a decrease in conformation entropy. When the load is released, the system (the polymer network) tends to increase the entropy, leading to chain “re-coiling”. In contrast, it has been established that in egg cases, the dominant contribution to elasticity is the change in enthalpy associated with the phase-transition from coiled-coil to β -sheets.¹⁰ In the unloaded state, egg case proteins adopt an α -helical coiled-coil conformation, which consists of at least two α -helices intertwined with each other.¹¹ Upon the application of an external load, the coiled-coils first unravel and then form β -sheets, because the Gibbs free energy of the β -sheets is lower than that of coiled-coils at high load. The driving force for elasticity is thus thermodynamically distinct from entropic elastomers (also called “rubber elasticity”).

The coiled-coil structure allows the egg capsule to exhibit full reversible extensibility. The engineering stress-strain hysteresis curve of *Busycotypus canaliculatus* was first reported by Rapoport.¹² A typical stress-strain curve consists of three distinct regions during mechanical loading and it exhibits a hysteresis loop upon unloading. The first region has an initial stiff region of up to approximately 5% strain with elastic modulus in the range of 100 MPa. The second pseudo-yield region extends approximately from 5% to 70% strain, and has a much lower slope. Beyond 70% strain, stiffening occurs in the third region. The egg capsule can be extended up to 150% without failure.^{12,13} When the stress is removed, the egg capsule is able to almost instantaneously recover to its original dimensions. This stress induced and relaxation cycle can be repeated multiple times without significant degradation to the mechanical properties.

Figure 2.2. Stress-strain hysteresis curve for engineering curve and true tensile curve of *Busycotypus canaliculatus* egg capsule. The true stress-strain's true modulus was obtained from the tangent during partial unloading/reloading cycles on the curve.

The hysteresis, shock absorption and energy dissipation characteristics can be attributed to the behaviour of the coiled-coils under stress. The initial slope is high as energy is a reflection of the stiff nature of coiled-coil rods. With increasing stress, the coils unravel into random coils, which allows extensibility. As the random coils further straighten out, it forms elongated β -sheets like conformation. β -sheets are more rigid, giving rise to increased stiffness.

Figure 2.3. Schematic diagram depicting the transition of α -helix to β -sheet when an external load is applied. The transition is reversible from β -sheet to α -helix when the external load is removed.

Wide-angle X-ray scattering (WAXS) analysis was performed on the Busycon egg capsule to support the data from the stress-strain curve.^{3,12,13} At 0% strain, two reflections could be seen which corresponds to α -helical conformation. A broad centred equatorial peak at 1.03 nm corresponds to the mean distance between helices and the 0.52 nm peak along the meridian owing to the coiled-coil axis. At 70% strain, β -sheet formation is indicated by the 0.465 nm equatorial reflections, and the 1.03 nm peak, which is now due to the mean inter-sheet distance.

When stretched further to 120%, the 0.465 nm reflection on the equatorial arc was intensified, as well as the 1.03 nm spots. The faint peak at 0.33 nm comes from the translational rise between residues along the β -sheet. Upon unloading, the diffraction pattern returns to that of an unstretched specimen. In conclusion, the x-ray reflections from WAXS are consistent with the structural and reversible transition between coiled-coil helices and β -sheets during a stretch-relaxation cycle.

Figure 2.4. Wide-angle X-Ray Scattering peak of *B. canaliculatus* egg capsule at 0%, 70%, 120% strain. 1.03 nm corresponds to the mean distance between helices and the 0.52 nm peak along the meridian owing to the coiled-coil axis. The β -sheet formation is indicated by the 0.465 nm equatorial reflections and 0.33 nm comes from the translational rise between residues along the β -sheet.^{3,12,13}

2.4 Surface Characterisation of *Pugilina cochlidium* egg capsules

The egg capsule must be able to withstand heavy impact against debris under strong water currents in all directions, therefore its shock absorbing properties should be anisotropic. Characterisation of the egg capsule revealed that it consists of four layer oriented 90° perpendicular to each other.¹³

The lack of elasticity on the first layer due to sclerotisation was significantly tougher than the other layers. The sclerotised layer which is heavily cross-linked acts as a first line of defence from impending debris strike and particles abrasion. Deep parallel ridges may help to channel the water currents away from the capsule, and also to reduce wear and tear. The second layer is made up of fine protein fibres closely packed together, whereas the third layer consist of coarse bundled fibres. It is postulated that this layer gives the main structural integrity to the egg capsule. Characterisation of the final layer shows a smooth and porous surface, which suggest that it might have mimicked as a dialysis membrane, which diffuses out metabolic waste from the developing embryos and takes in minerals and oxygen from the environment.

Figure 2.5. Optical microscopy of the egg capsule first three layers. Images are focused stack for improved depth of field at high magnifications. Dark field image B, C and D are taken at 500x magnification. **A)** An overlay of bright and dark field image at 2.5x magnification of an egg capsule with developing embryos. **B)** Deep and tough surface ridges running parallel to each other, suggesting the aid in diverting water current along the surface away from the egg capsule. **C)** Fine and smooth protein fibres in the second layer oriented perpendicularly to the third layers which consist of thick and coarse bundles of protein fibres. **D)** Threads of protein cross-linked to neighbouring fibre bundles when lightly stretched.

2.5 Structure of coiled-coils and egg capsule proteins

The structure of the coiled-coil structure of α -keratin was first proposed by Crick in 1952^{14,15} and later by Pauling & Corey¹⁶ using meridional reflections in X-ray diffraction patterns when investigating the properties of natural wool.

Coiled coils consist of two to five helices wrapped around each other “in a knobs-into-holes manner” fashion, forming a left-handed helix that subsequently forms a supercoil. Each complete turn of the helix for a regular α -helices is 3.6 residues, but the distortion within a left-handed coiled coil on each helix lowers this value to around 3.5, therefore the heptad repeat occurs every two turns of the helix.^{17,18,19}

The heptad repeat pattern is made up of seven amino acid residue, **HPPHCPC**, where hydrophobic (**H**), charged (**C**) and polar (**P**) amino-acid residues fill in the respectively positions to stabilize the formation of heterodimers. The positions in the heptad repeat are usually labeled $(a-b-c-d-e-f-g)_n$ in one helix and the other as $(a'-b'-c'-d'-e'-f'-g')_n$.²⁰ Critically, the hydrophobic residues (positions *a* and *d*) are located within the inner core of the supercoil and stabilize it by hydrophobic interactions.

Figure 2.6. Schematic representation of a dimeric coiled coil.²⁰ Residue a/a' and d/d' forms the hydrophobic core, e/e' and g/g' forms the ionic interaction and b/b', c/c' and f/f' residues are hydrophilic and exposed to the solvent.

Intermediate filaments (IFs) are cytoskeletal components found in the cells of many animal species. The role of IFs range from cells signal induction to mechanical integrity. Shown in Figure 2.7, the three subdomains of all Ifs are: an α -helical central "rod" domain a non- α -helical "head" and "tail" domains on its either side. The rod domain is primarily α -helical with a heptad repeat pattern in segments 1A and 1B forming coil 1, as well as in segments 2A and 2B constituting coil 2. Short non-helical linkers (L1, L12, and L2) join these segments together. The most conserved regions are found at the N-terminal end of the rod (1A) and the C-terminal end of the rod domain (2B). A four-residue insertion near the middle of segment 2B is a stutter that is highly conserved in all IF.^{21,22,23}

Figure 2.7. A model of an IF dimer. The middle section consists of an α -helical central "rod" domain and flanked by non- α -helical "head" and "tail" domains.

Previous studies on the ECPs sequences from *B. canaliculatus* and *P. cochlidium* have shown that ECPs contain extensive heptad repeats. The location of the coiled coil domains within the protein were predicted by SCORER 2.0. The comparison of an IF representation and the domain architecture of ECPs is shown in Figure 2.8. The N-terminal region of all the ECP did not strongly predict for coiled coil structure, whereas the C-terminal region of BcECP-1a and -1b does. The domain structures for IFs are known to be highly conserved across diverse IF classes, however it varies quite significantly across the ECPs of different species.^{3,24,25}

Figure 2.8. A) An IF representation with an α -helical rod in the middle and non- α -helical "head" and "tail" domains. B) Comparison of the predicted coiled-coil domains of BcECP and PcECP with the IF model. BcECP-1a and -1b has a predicted coiled-coil C-terminus which differs from that of PcECP and IF.

The nidamental gland crude extracts retrieved from both *B. canaliculatus* and *P. cochlidium* were subjected to SDS-PAGE. The gel revealed three major bands for *B. canaliculatus* and five major bands for *P. cochlidium*. MALDI-TOF was conducted on the samples and the molecular weights are shown in Figure 2.9. The additional proteins seen on the gel for *P. cochlidium* were investigated by searching in the RNA sequence database but did not yield any results, therefore Edman sequencing was conducted and the results were matched against the RNA sequence database. The discrepancy of the results suggested that the extra bands could arise from isoforms

and/or splicing of PcECP which are not easily distinguishable by the RNA sequence assembly, or by post-translational modifications, which has however not been substantiated to date.¹³

Figure 2.9. A) SDS-PAGE of *B. canaliculatus* and *P. cochlidium* nidamental gland crude extract. B) MALDI-ToF spectra of *P. cochlidium* showing 5 distinct peaks with molecular weights from 42.0 kDa to 54.3 kDa. C) MALDI-ToF spectra of *B. canaliculatus* showing 3 peaks from 48 kDa to 52.5 kDa.

To the author's understanding, there is currently no literature detailing the egg capsule self-assembly into the soft gel state, which is a process occurring between the viscous liquid (precursor PcECPs) and solid state (sclerotized) of the egg capsule, thus this thesis aims to bridge this gap. This step is crucial in the complete understanding of the hierarchical assembly of PcECP precursor proteins.

References

- [1] *Pugilina cochlidium*_(Linnaeus, 1758). Retrieved through: World Register of Marine Species on 20 June 2015
- [2] Food and Agriculture Organisation of the United Nations, *Gastropods Melongenidae*, pp. 584
- [3] Paul A. Guerette, Gavin Z. Tay, Shawn Hoon, Jun Jie Loke, Arif F. Hermawan, Clemens N. Z. Schmitt, Matthew J. Harrington, Admir Masic, Angelo Karunaratne, Himadri S. Gupta, Koh Siang Tan, Andreas Schwaighofer, Christoph Nowak and Ali Miserez, *Integrative and comparative analysis of coiled-coil based marine snail egg cases – a model for biomimetic elastomers*, **2014**, Biomater. Sci., 2, pp. 710-722
- [4] J. Herbert Waite, *Surface chemistry: Mussel power*, **2008**, Nature Materials, 7, pp. 8-9
- [5] Bruce P. Lee, P.B. Messersmith, J.N. Israelachvili, J.H. Waite, *Mussel-Inspired Adhesives and Coatings*, **2011**, Annual Review of Materials Research, Vol. 41, pp. 99-132
- [6] K.S. Tan and C.L. Phuah, *Diet and feeding habits of Pugilina cochlidium (Neogastropoda: Melongenidae) in Singapore*, **1999**, J. Moll. Stud., 65, pp. 499-501
- [7] S.Y. Chan, *The Melongenidae (Mollusca: Gastropoda) Of Singapore*, **2009**, Nature in Singapore, 2, pp. 63-67
- [8] Mary C. Boyce, Ellen M. Arruda, *Constitutive Models of Rubber Elasticity: A Review*, **2000**, Rubber chemistry and technology, 73(3), pp. 504-523
- [9] L.R.G Treloar, *The Physics of Rubber Elasticity*, **1975**, Oxford University Press, USA
- [10] Ali Miserez, Paul A. Guerette, *Phase transition-induced elasticity of α -helical bioelastomeric fibres and networks*, **2013**, Chem. Soc. Rev., 42, pp. 1973-1995

- [11] Burkhard P, Stetefeld J, Strelkov SV, *Coiled coils: a highly versatile protein folding motif*, **2011**, Trends Cell Biol., 11(2), pp. 82-88
- [12] H. Scott Rapoport, Robert E. Shadwick, *Mechanical Characterization of an Unusual Elastic Biomaterial from the Egg Capsules of Marine Snails (*Busycon spp.*)*, **2002**, Biomacromolecules, 3(1), pp. 42-50
- [13] A. Miserez, S. Wasko, C. Carpenter and J.H. Waite, *Non-Entropic and Reversible Long Range Deformation of an Encapsulating Bioelastomer*, **2009**, Nature Materials, 8, pp. 910-916
- [14] F. H. C. Crick, *Is α -Keratin a Coiled Coil?*, **1952**, Nature, 170, pp. 882-883
- [15] F. H. C. Crick, *The Packing of α -Helices: Simple Coiled-Coils*, **1953**, Acta Crystallogr., 6, pp. 689-697
- [16] Pauling L, Corey RB, Branson HR, *The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain*, **1951**, Proceedings of the National Academy of Sciences of the United States of America, 37(4), pp. 205-11
- [17] Landschulz WH, Johnson PF, McKnight SL., *The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins*, **1988**, Science, 240(4860), pp. 1759-64
- [18] A. Lupas, *Coiled coils: new structures and new functions*, **1996**, Trends Biochem. Sci., 21(10), pp. 375-382
- [19] Bojana Apostolovic, Maarten Daniala, Harm-Anton Klok, *Coiled coils: attractive protein folding motifs for the fabrication of self-assembled, responsive and bioactive materials*, **2010**, Chem. Soc. Rev., 39, pp. 3541-3575
- [20] Jody M. Mason and Katja M. Arndt, *Coiled Coil Domains: Stability, Specificity, and Biological Implications*, **2004**, ChemBioChem., 5, pp. 170-176

- [21] Herrmann H, Bär H, Kreplak L, Strelkov SV, Aebi U, *Intermediate filaments: from cell architecture to nanomechanics*, **2007**, Nat. Rev. Mol. Cell Biol., 8(7), pp. 562-73
- [22] Karabinos, Anton, Dieter Riemer, Andreas Erber, Klaus Weber, *Homologues of Vertebrate Type I, II and III Intermediate Filament (IF) Proteins in an Invertebrate: The IF Multigene Family of the Cephalochordate Branchiostoma*, **1998**, FEBS Letters, 437(1-2), pp. 15-18
- [23] Qin Z, Kreplak L, Buehler MJ, *Hierarchical structure controls nanomechanical properties of vimentin intermediate filaments*, **2009**, PLoS ONE 4(10): e7294
- [24] Guerette P.A., Hoon S., Seow Y., Raida M., Masic A., Wong F.T., Ho V.H., Kong K.W., Demirel M.C., Pena-Francesch A., Amini S., Tay G.Z., Ding D., Miserez A., *Accelerating the design of biomimetic materials by integrating RNA-seq with proteomics and materials science*, **2013**, Nature Biotechnology, 31, pp. 908-915
- [25] S. Wasko, G. Z. Tay, A. Schwaighofer, C. Nowak, J. H. Waite, A. Miserez, *Structural proteins from whelk egg capsule with long range elasticity associated with a solid-state phase transition*, **2014**, Biomacromolecules, 15(1), pp. 30-42

Chapter 3.

Experimental methodologies

All the experimental and characterisation techniques employed to study the egg capsule protein are described in this chapter. The principles and theory behind each experiment are laid out methodologically. The major experiment techniques includes electrophoresis, High Pressure Liquid Chromatography and refolding of purified protein. Optimisations of Reversed Phase High Pressure Liquid Chromatography are also discussed as purification of Pugilina Bundling Protein proved to be challenging.

3.1 Preparation of native crude extract from *P. cochlidium* nidamental gland

Native crude extract and denatured crude extract 1 ml each were prepared from *P. cochlidium* nidamental gland. Randomised samples from non-gravid and gravid snails were prepared. This was to ensure that any similarities or differences in the samples during analysis are consistent for the non-gravid and gravid nidamental gland.

3.2 Native state of egg capsule proteins using electrophoresis techniques

The study of native state protein using proteomic techniques such as electrophoresis can be a challenge, despite the apparently simplicity. Native proteins are folded into a variety of shapes such as globular, helical and rods.¹ Proteins can be made up with different combination of amino acids. Depending on how the amino acids are arranged, a protein can incorporate an array of random coils¹, coiled-coils², α -helices³ and β -sheets⁴. Isoelectric point of the protein is one of the important factors for electrophoresis.

Since native proteins behave differently from a fully denatured and linearized protein, several electrophoresis systems were used and customised to determine the oligomeric state of the *Pugilina cochlidium* native crude extract. Electrophoresis was carried out using Biorad Mini-PROTEAN handcast system (0.75 mm thick casting plates) and large format gels using Biorad PROTEAN[®] II xi Cell.

3.2.1 Tris-Glycine SDS-PAGE

Tris-Glycine SDS-PAGE^{5,6} is the most commonly used technique to separate proteins between 30 kDa and 200 kDa. When electrophoresis starts, glycine in the cathode buffer enters the pH 6.8 stacking gel where it becomes its zwitterionic form with zero net charge. The slow mobility of glycine trails behind the strong negatively charged chloride ions, thus a region of low conductivity with high voltage drop is formed between them. This zone compresses the proteins through into a narrow band.

Glycine becomes negatively charged when it reaches the pH 8.8 running gel and migrates much faster than the proteins due to a higher charge/mass ratio. The proteins now becomes the main carrier of current and separate according to their molecular mass by the sieving effect of pores in the gel.

Due to the slow mobility of glycine, proteins below 20 kDa co-migrate with SDS in between glycine and chloride ions. This causes poor separation and low resolution of low molecular weight proteins.

3.2.2 Tris-Tricine SDS-PAGE

Tris-Tricine SDS-PAGE was developed to better separate low molecular weight proteins ranging from 1 kDa to 100 kDa.^{7,8,9} Tricine has a lower pKa value than glycine, thus its ionised form is much stronger and trails much closer to the leading chloride ions. As such, the SDS layer is compressed further than that of glycine-chloride layer. This enables low molecular weight proteins between 1 kDa and 20 kDa to be separate from the SDS stack, improving the separation and resolution.

3.2.3 Neutral pH SDS-PAGE

Neutral pH Polyacrylamide Gel Electrophoresis (Invitrogen NuPAGE® Novex®) simulates the denaturing conditions of the traditional Laemmli Tris-Glycine system, but is closer to neutral pH and has been shown to be much superior in terms of separation and resolution for a wide range of proteins.¹⁰ The alkaline pH of Tris-Glycine and Tris-Tricine SDS-PAGE can catalyse the reaction of residual unpolymerized acrylamide with cysteine and lysine residues of the proteins during the run.¹¹ The reducing agent in the sample does not co-migrate with the proteins, thus reoxidation causes broad proteins bands and lack of resolution. The use of antioxidant comprising of sodium metabisulfite and dimethylformamide keeps the cathode buffer in a reducing environment, thus giving high resolution of the separated proteins.

3.2.4 Blue Native Polyacrylamide Gel Electrophoresis

Clear native electrophoresis^{12,13} is similar to that of conventional SDS-PAGE without the addition of anionic sodium dodecyl sulfate detergent. This allows the proteins to keep their native conformational state and enzyme activities. However proteins will migrate according to their isoelectric points. A native crude protein complex with differing isoelectric points is extremely challenging as positively charged proteins will migrate out of the gel should the electrophoresis current run from the anode to cathode in a neutral pH gel.

Blue native electrophoresis^{14,15,16} makes use of CBG-250 anionic dye as a charge shift. Similar to SDS, CBG-250 binds tightly to the proteins and masks the protein intrinsic charge, but does not denature the native conformation state. The proteins now migrate according to their molecular weight. Depending on the shape of the folded protein, it might not migrate accurately with respect to the calculated relative mobility. A relative mobility plot is usually done for different polyacrylamide gel cross-link densities with respect to known molecular weights such as bovine serum albumin monomer of 66 kDa and its oligomers.

3.2.5 Non-Reducing Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Non-reducing SDS-PAGE¹⁷ does not involve the use of heat and thiol reducing agents such as β ME, DTT and TCEP. Although SDS is a strong anionic surfactant, it is unable to penetrate deep into the protein structure to disrupt the covalent bonds. Weak protein-protein interactions such as van der Waals will be disrupted, but hydrogen and covalent bonding will remain intact, thus chaotropes such as urea and guanidine hydrochloride were used to disrupt the hydrogen bonds. Using the non-reducing method for electrophoresis will still disassociate some protein subunits, thus careful analysis of protein bands is necessary.

3.2.6 Two-Dimensional Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Two-dimensional electrophoresis¹⁸ was originally dedicated to Iso Electric Focusing (IEF). IEF uses very high voltage to focus proteins of the same isoelectric point into a very thin band. An ampholyte solution is copolymerized with polyacrylamide gel to form an immobilized pH gradient. The strip with the respective focused proteins is then loaded into an SDS-PAGE gel and the proteins are fully denatured into their subunits.

3.2.7 Ethylene Glycol Diacrylate cross-linked polyacrylamide gel

Conventional polyacrylamide gel uses N,N'-Methylenebisacrylamide as the cross-linking agent. Bisacrylamide reacts with long linear chains of polyacrylamide filaments in the presence of ammonium persulfate to form a dense porous hydrogel.¹⁹ The amide bonding is stable in both alkaline and acidic buffers, especially during storage

and electrophoresis in pH 8.8 gel buffer and during Coomassie blue staining where acidity drops to pH 2.5.

To determine the molecular weight of the protein band using mass spectrometry, in-gel tryptic digestion of the protein is typically performed. This enzymatic process cleaves the protein into shorter peptide fragments²⁰ which then diffuse from the gel into the supernatant. The extraction of native proteins using this method can be complicated. An alternative would be to dissolve the polyacrylamide while leaving the proteins intact.

Ethylene glycol diacrylate polyacrylamide gel attains an ester linkage which is easily reducible in alkali solutions such as sodium hydroxide or ammonium hydroxide.²¹ Ethylene glycol diacrylate gels using pH 8.8 running gel buffers are made fresh prior to the electrophoresis run. Gels stored overnight or excessive heat during the run can cause poor resolution and casting plates removal problems due to the degradation of the ester bonds. Bisacrylamide 3.3% w/v was substituted with ethylene glycol diacrylate 3.3% v/v. Acrylamide solution was prepared separately from ethylene glycol diacrylate.

3.3 Polyacrylamide in-gel protein fixation

After the electrophoresis run has ended, the gel was incubated in sensitisation buffer containing 30% v/v ethanol, 10% v/v acetic acid and 10% v/v methanol to fix and denature the proteins, as well as to remove excess detergent and gel buffer which can interfere with CBG-250 staining, thereby reducing band sensitivity. This fixing step was repeated until the gel tracking dye is not visible. It was then stained with Blue Silver staining solution²² overnight. The stained gels were washed with 10% v/v ethanol briefly multiple times until the gel background is transparent.

3.3.1 Blue Silver Staining

Blue Silver staining technique²² is a highly sensitive protein stain using coomassie blue dye. The high amount of orthophosphoric acid 10% v/v further protonates Asp and Glu residue, favouring the ionic bonding of the dye. Coomassie Blue G-250 forms colloidal particles when the pH is highly acidic and the colour of the solution turns from blue to green. These colloidal particles are micron sized and are unable to penetrate deeply into the porous network of the polyacrylamide gel, thus minimising background staining. The sensitivity of this stain is approximately 10 ng of protein, as compared to conventional coomassie blue staining of 100 ng. Silver stain²³ is the most sensitive at 1 ng,

however due to the use of toxic chemicals such as silver nitrate and formaldehyde coupled with multiple developing step, this method was not favoured.

3.3.2 Zinc-Imidazole Negative Stain

Zinc-Imidazole negative stain²⁴ is an alternative to the conventional coomassie brilliant blue staining technique. Being a negative stain, the gel is stained white while the proteins remained transparent. The gel was incubated in imidazole and SDS buffer to exchange with the Tris gel buffer. The reaction of zinc ions with imidazole-SDS complex forms a white precipitate in the acrylamide gel. Zwitterionic detergent such as CHAPS was used for native as well as non-reduced protein samples. The detection limit for this reversible stain can be as high as 10 ng.

3.4 Transformation of *Pugilina* Bundling Protein gene

The gene of PcBP was purchased from GenScript (New Jersey, USA), cloned in vector pET-26b(+) and transformed into One Shot Top10 chemically competent *E. coli* (Invitrogen, California, USA) and Miniprep (Qiagen, Germany) was done to extract larger amount of PcBP plasmids. The plasmids were subsequently transformed into BL21 (DE3) competent *E. coli* (Agilent, California, USA) for large scale expression.

3.5 Expression Systems using Luria Bertani and Terrific broth

The bundling protein was expressed in two different medium, the widely used Luria Bertani Broth and Terrific Broth. Most expression system uses LB broth²⁵ for the cultivation of *E. coli* in plasmid DNA and recombinant protein applications. It would be advantageous for the broth to sustain the bacterial growth, but with the increase of cell count, the acidity of the broth increases as well which leads to cell death, which in turn decreases the overall protein expressed yield.

Tartoff Hobbs Terrific broth (TB)^{26,27} is a nutritionally rich medium for the growth of bacteria. The formulation of TB medium contains increased concentrations of peptone, yeast extract, and glycerol as a carbohydrate source. Potassium phosphate at pH 7.2 was added as a pH buffer.

3.6 Cell Lysis

Ten millilitres of ice cold lysis buffer at pH 8.0 with protease inhibitor, 1 mM PMSF and 10 mM DTT was added to every 100 ml of pelleted cells. Sonicator probe was used for the small volume of suspended cells. For larger volume, micro-fluidiser with 5 passes was used for lysis.

3.7 Protection of cysteines by reversible thiolation with 2,2'-Dithiopyridine

Cysteines are easily reoxidised after thiol reduction during sample preparation, thus vacuum degassing of buffers is necessary to facilitate the removal of reactive molecular oxygen which can produce sulfenic acid under oxidative stress^{28,29}. Such oxidation results in the formation of dimers and its oligomers which can either irreversibly denature or inhibit full protein activity due to misfolding.³⁰

Thiol quantification and detection is commonly carried out using 5,5'-Dithiobis(2-nitrobenzoic acid) (DTNB), also known as Ellman's reagent.³¹ It provides a reliable method to quantify reduced cysteines and free sulfhydryl groups in the protein at 412 nm wavelength. Based on the same principle as DTNB, another thiol reagent 2,2'-Dithiopyridine was selected due to its lower molecular size, which might be better for small proteins with high number of cysteines, such as *Pugilina* Bundling Protein. The sensitivity is higher than that of DTNB due to less steric hindrance in accessing the free protein thiols.³²

A study by Kazunori Maruyama *et al.* propose that 2,2'-Dithiopyridine was able to oxidise cysteines into an intramolecular S-S bond.³³ Two differences in the methodology were explored. Firstly, oxidation of the cysteine were carried out at alkaline pH and secondly, the molarity of 2,2'-Dithiopyridine was less than 1 mM.

Two possible mechanisms of the intramolecular S-S bond was suggested: (1) 2,2'-Dithiopyridine is reduced by a free SH group in the protein to form an S-S pyridyl derivative and 2-mercaptopyridine³³; and (2) the S-S pyridyl side chain of cysteine is attacked by another free SH group in the same protein to complete the S-S exchange reaction, liberating another free 2-mercaptopyridine. The latter reaction could be due to the low concentration of 2,2'-Dithiopyridine.

To ensure that the reformation of disulphides in the protein are suppressed, the oxidative environment was minimized. The reduced cysteines are biased to the R-SH form with 5% v/v acetic acid at pH 2.4 and a saturating concentration of 2,2'-Dithiopyridine.

Stock solution of 2,2'-dithiopyridine 1 M concentration was dissolved in pure ethanol. After solubilisation of the washed inclusion bodies in 1 ml 6 M urea, 10 μ l of the 1 M stock was added and incubated for at least 1 hour at 37°C, followed by addition of 100 μ l glacial acetic acid. Formation of free 2-Thiopyridone from reduced 2,2'-Dithiopyridine can be measured at 343 nm wavelength. The protein solution takes on a translucent pale yellow after the reaction is completed.

Figure 3.1. Reduction and protection of *Pugilina* Bundling Protein sulfhydryl group with reversible thiol reagent 2,2'-Dithiopyridine. The mechanism of mixed disulphide was achieved with saturating concentration of 2,2'-Dithiopyridine. 2-Mercaptopyridine can be detected at 343 nm wavelength.

Figure 3.2. Solubilised *Pugilina* Bundling Protein inclusion bodies solution in 6 M urea **A)** before addition of 10 mM 2,2'-Dithiopyridine showing translucent orange-brown colour and **B)** addition of 10 mM 2,2'-Dithiopyridine showing translucent pale yellow colour after complete reaction.

3.8 Purification of *Pugilina* Bundling Protein by High Performance Liquid Chromatography

3.8.1 Size Exclusion Chromatography

Size exclusion chromatography separates molecules based on their molecular weight. The porous silica stationary phase allows particles size smaller than the pore size to travel in and out, thereby retaining it longer in the column, whereas larger particles would flow straight through the column. Chaotropic agents such as urea can affect a molecule's hydrodynamic radius.³⁴ In such cases, the molecule may have its retention time varied significantly.³⁵ Size exclusion chromatography might be a good choice to separate PcBP by its size given the low molecular weight of the protein.

3.8.2 Reversed Phase Chromatography

The purification of *Pugilina* Bundling Protein had proven to be extremely difficult and challenging. The protein is rich in reactive cysteine and contains many hydrophobic residues with a net charge of +1. The use of reversed phase column using conventional protocol (6 M Urea + 5% v/v acetic acid sample buffer in 0.1% TFA ACN mobile phase) did not successfully purify the protein.

The mode of purification in reversed phase chromatography is via the decrease in the polarity of the solvent. In a polar solvent mobile phase, the affinity of the protein exposed hydrophobic core is strong, thus binding occurs. When the polarity of the mobile is weaker than that of the stationary phase, the protein is displaced from the ligands and eluted with the mobile phase.

Theoretically reversed phase chromatography utilizes only one purification mode. Most protein however operates in mix-mode, which means that it will have non-specific binding to the reversed phase column.³⁶ In the case of PcBP, three main mode has been identified, namely hydrophobic, electrostatic and hydrogen bonding.

Reversed phase columns are made of silica spheres covalently bonded to an alkyl chain.³⁷ The length of the alkyl chain dictates the level of hydrophobicity and the carbon length ranges from 1 to 18, generally denoted as C1, C3, C4, C8 and C18. Strong hydrophobic column such as C18 are generally used for more polar molecules and

peptide separation,³⁸ whereas weaker hydrophobic column C1 to C8 and Cyanopropyl ligands are more suited for very hydrophobic molecules and proteins.³⁹

During manufacturing, not all alkyl ligands are covalently attached to the silica surface successfully. This leaves the silica surface with a reactive silanol functional group, Si-O-H.^{40,41} Most silanols only have one OH group known as trimethylsilanol and it is responsible in engaging hydrogen bonding.^{42,43} Silanol interaction can be suppressed by lowering the acidity of the mobile phase to pH 3.0 and below, resulting in the protonation of silanols.⁴⁴ In most modern columns, reduction of unreacted silanol involves the derivatization of the silanol groups. An extra process of bonding with trimethylsilyl groups is referred to as endcapping.^{45,46}

Potential ion exchange and hydrogen bonding secondary interaction are unwanted in reversed phase chromatography. While both cannot be totally eliminated, acid additives such as trifluoroacetic acid are added to the mobile phase to significantly reduce the effects. The challenge and difficulty with purifying *Pugilina* Bundling Protein is that the protein exhibit severe peak tailing caused by secondary interactions. By understanding the characteristic of PcBP, choosing appropriate length of carbon ligand chain, acid modifiers and choice of solvent can help reduce peak tailing and greatly improve purification.

In the case of PcBP, the protein in acidic mobile phase will have the cysteine in the R-SH form. Since TFA is not protonated at pH 1.9, it will compete with the protonated silanols. This causes some silanols to be ionized again (SiO⁻). *Pugilina* Bundling Protein having a pI of 6.7 will be protonated at pH 1.9, and thus the ionized silanol will interact with the protonated PcBP as well as the cysteine R-SH, causing ionic exchange, thus increased retention and peak tailing.

Methanol and water both have an OH hydroxyl group and interact more strongly with the free silanol group through hydrogen bonding. Acetonitrile has much weaker interactions with the free silanol due to its low hydrogen-bond acceptor ability.^{47,48} As acetonitrile gradient increases, the amount of water interacting with the free silanols decreases. This in turn increases the overall amount of free silanols which has higher affinity for PcBP amide and reduced thiol group, leading to hydrogen bonding secondary interactions.

3.8.3 Optimisation of Reversed Phase Chromatography

Prior to the development of the binary solvent system, a range of acids and organic solvents were used in an attempt to purify the highly hydrophobic PcBP. A summary of different combinations of solvent, acids and additives that were tested are shown in Table 3.1. An 'O' between organic solvents and acid indicates that the separation and peak shape is good, and different additives are added to further optimise the purification as listed in the additives column. An 'X' indicate that the combination is unstable and gave very poor separation and resolution, whereas a '-' was not tested.

To use the additives, a minimum of 10% v/v water is required for salts and TCEP, and a minimum of 10% v/v organic solvent is required for 1-Butyl-3-methylimidazolium hexafluorophosphate (BMIM PF₆). As an example, Acetonitrile in 5% v/v acetic acid gave acceptable separation and resolution, and the additive on the same row can be added to further improve the separation.

Table 3.1 Summary list of different acids, organic solvent and additives formulated to purify PcBP using reversed phase column. An X indicate that the purification is either of poor separation or unstable for the protein, and an O indicate that the method can be optimised further. A dash indicates the combination was not tested.

Organic Solvents	Acids							Additives						
	0.1% TFA	0.1% H ₃ PO ₄	Formic Acid (v/v %)				5% v/v Acetic Acid	TCEP (mM)			NaCl 0.1 M	BMIM PF ₆ (mM)		
			5	2	1	0.1		1	5	10		5	10	20
Methanol	O	X	X	X	X	O	O	O	-	O	O	O	O	
Acetonitrile	O	X	X	X	X	O	O	O	O	O	O	O	O	
2-Propanol	-	-	X	X	X	O	-	O	-	-	O	-	-	
2-Methyl-2-Butanol	-	-	-	-	-	O	-	O	-	-	O	-	-	

Ionic liquids such as BMIM PF₆ are salts that have melting points typically below 100°C. The liquid state is due to its bulky and non-symmetrical shape of the cations. The thermal stability and non-volatility makes it a good alternative to common salts for ionic charge screening.⁴⁹

Although the use of ionic liquids in reversed phase chromatography is not widely use, its performance as an additive had not gone unnoticed. At low concentration between (0 - 50) mM, BMIM PF₆ was found to improve peak symmetry and resolution in separating co-eluting compounds. However, with increasing concentration of the ionic liquid, retention time of the compounds increases. This suggest that the imidazolium cation had formed an ionic bond with the underivatised silanols, thus increasing the surface hydrophobicity of the stationary phase.^{50,51,52}

As the concentration increases further, the retention time of the compounds begins to decrease. It is proposed that the ionic liquid saturates the silanols and produces a weak bilayer. The electrostatic repulsion repels the basic analytes and interacts with alkyl group through hydrophobic interaction, thus compound retention decreases due to these repulsive and hydrophobic interactions.⁵³

Figure 3.3. **A)** Schematic diagram of activated silanols at pH above its pK_a. **B)** Protonation of silanol at low pH thus cancel the effects of ionic interaction. **C)** Endcapping of silanols with trimethylsilyl to reduce the detrimental effects of silanol coupling with analytes. The bulky group of trimethylsilyl prevents complete derivatisation of silanols due to steric hindrance, thus the use of ionic liquid can help to reduce the secondary interactions. **D)** Representation of 1-Butyl-3-methylimidazolium hexafluorophosphate.

3.9 Purification by Dialysis

Dialysis could be a more efficient way to pre-purify the low molecular weight PcBP using a suitable membrane cutoff. Solutes and analytes below the membrane cutoff size are diffused out into the dialysate and the overall concentration is diluted after several dialysis buffer change. Compatibility of membranes with the buffer must be taken into consideration to prevent membrane degradation. The analytes that had been diffused out were re-concentrated by lyophilisation or ultrafiltration.

3.10 Matrix Assisted Laser Desorption Ionisation-Time of Flight Mass Spectroscopy

Shimadzu Axima MALDI-TOF was able to provide a qualitative analysis of the molecular weight of the target protein and other contaminants which co-eluted with it. Also, this sensitive technique was able to determine the oxidation state of proteins, if any.

The power of the laser was determined based on the signal to noise ratio of the sample. One hundred profiles and ten repetitive shots were selected for every sample. Detection mode was set to Linear.

3.11 Refolding buffer with redox pair

Cysteine rich proteins requires a redox buffer system, generally with glutathione (GSH) and oxidised glutathione (GSSH). Due to its high cost, alternative redox pairs such as cysteine/cystine pair had also been proven useful in refolding such proteins. L-Arginine at high molarity had shown to suppress aggregation. Thus the addition of cysteine/cystine redox pair and L-Arginine was selected to assist in the refolding of PcBP.

3.12 Characterisation of PcBP – PcECs aggregates

To ensure PcBP is correctly refolded and active, 20 µl of the refolded PcBP was added to 20 µl of non-polymerized native crude extract and gently mixed with the pipette. A negative control was also carried out to ensure that the refolding buffer did not cause the aggregation of the native crude extract.

3.12.1 Optical Microscopy

Optical microscopy images of the aggregates were captured using a Canon EOS 6D camera attached to Zeiss microscope. Transmission mode was selected. Twenty microliters of native unpolymerised PcECP crude extract was spotted onto a microscope glass slide, followed by adding 20 μ l of refolded PcBP. The mixture was gently swirled in a circular motion. Magnification at 2.5x and 10x were used.

References

- [1] Charles. Tanford , Kazuo. Kawahara, Savo. Lapanje, *Proteins as Random Coils. I. Intrinsic Viscosities and Sedimentation Coefficients in Concentrated Guanidine Hydrochloride*, **1969**, J. Am. Chem. Soc., 89(4), pp. 729-736
- [2] F. H. C. Crick, *The Packing of α -Helices: Simple Coiled-Coils*, **1953**, Acta Cryst., 6, pp. 689
- [3] Pauling, L. & Corey, R. B., *The Structure Of Proteins: Two Hydrogen-Bonded Helical Configurations Of The Polypeptide Chain*, **1951**, Proc. Natl. Acad. Sci. USA, 37, pp. 729-740
- [4] Jane S. Richardson, *β -Sheet topology and the relatedness of proteins*, **1977**, Nature, 268, pp. 495-500
- [5] David E. Garfin, *Gel Electrophoresis of Proteins*, **2003**, Essential Cell Biology Vol 1: Cell Structure, A Practical Approach, pp. 197-268
- [6] Laemmli UK, *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*, **1970**, Nature, 227(5259), pp. 680-685
- [7] Haider SR, Sharp BL, Reid HJ, *A Comparison Of Tris-Glycine And Tris-Tricine Buffers For The Electrophoretic Separation Of Major Serum Proteins*, **2011**, J Sep Sci., 34(18), pp. 2463-2467
- [8] Hermann Schagger, Gebhard Von Jagow, *Tricine-Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis for the Separation of Proteins in the Range from 1 to 100 kDa*, **1987**, Analytical Biochemistry, 166, pp. 368-379
- [9] Hermann Schagger, *Tricine–SDS–PAGE*, 2006, **Nature**, Protocols 1, pp. 16-22
- [10] Kinoshita E, Kinoshita-Kikuta E, *Improved Phos-tag SDS-PAGE under neutral pH conditions for advanced protein phosphorylation profiling*, **2011**, Proteomics., 11(2), pp. 319-323

- [11] Hachmann JP, Amshey JW, *Models of protein modification in Tris-glycine and neutral pH Bis-Tris gels during electrophoresis: effect of gel pH*, **2005**, *Anal Biochem.*, 342(2), pp. 237-245
- [12] Ilka Wittig and Hermann Schagger, *Advantages and limitations of clear-native PAGE*, **2005**, *Proteomics*, 5, pp. 4338-4346
- [13] Ilka Wittig, Michael Karas and Hermann Schagger, *High Resolution Clear Native Electrophoresis for In-gel Functional Assays and Fluorescence Studies of Membrane Protein Complexes*, **2007**, *Mol Cell Proteomics.*, 6(7), pp. 1215-25
- [14] Hermann Schagger and Gebhard von Jagow, *Blue Native Electrophoresis for Isolation of Membrane Protein Complexes in Enzymatically Active Form*, **1991**, *Analytical Biochemistry*, 199, pp. 223-231
- [15] Hermann Schagger, William A. Cramer and Gebhard von Jagow, *Analysis of Molecular Masses and Oligomeric State of Protein Complexes by Blue Native Electrophoresis and Isolation of Membrane Protein Complexes by Two-Dimensional Native Electrophoresis*, **1994**, *Analytical Biochemistry*, 217, pp. 220-230
- [16] Ilka Wittig, Hans-Peter Braun and Hermann Schagger, *Blue native PAGE*, **2006**, *Nature Protocols* 1, pp. 418-428
- [17] Stern BD, Wilson M, Jagus R, *Use Of Nonreducing SDS-PAGE For Monitoring Renaturation Of Recombinant Protein Synthesis Initiation Factor, eIf-4 Alpha*, **1993**, *Protein Expr Purif.*, 4(4), pp. 320-327
- [18] Bjellqvist B, Ek K, Righetti PG, Gianazza E, Gorg A, Westermeier R, Postel W, *Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications*, **1982**, *J Biochem Biophys Methods*, 6(4), pp. 317-339
- [19] Andrews, A.T., *Electrophoresis: Theory, Techniques, And Biochemical And Clinical Applications*, **1986**, (2nd edn). Oxford University Press, Oxford.

- [20] Andrej Shevchenko, Henrik Tomas, Jan Havliš, Jesper V Olsen & Matthias Mann, *In-gel digestion for mass spectrometric characterization of proteins and proteomes*, **2007**, Nature Protocols 1, pp. 2856-2860
- [21] Dimitrios G. Papasotiriou, Stavroula Markoutsas, Jan Gorka, Enrico Schleiff, Michael Karas, Bjoern Meyer, *MALDI analysis of proteins after extraction from dissolvable ethylene glycol diacrylate cross-linked polyacrylamide gels*, **2013**, Electrophoresis, 34, pp. 2484-2494
- [22] Candiano G, Bruschi M, Musante L, Santucci L, Ghiggeri GM, Carnemolla B, Orecchia P, Zardi L, Righetti PG, *Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis*, **2004**, Electrophoresis., 25(9), pp. 1327-1333
- [23] Mireille Chevallet, Sylvie Luche, Thierry Rabilloud, *Silver staining of proteins in polyacrylamide gels*, **2006**, Nature Protocols 1, pp. 1852-1858
- [24] Maria Luisa Ortiz, Miguel Calero, Carlos Fernandez Patron, Lila Castellanos and Enrique Mendez, *Imidazole-SDS-Zn reverse staining of proteins in gels containing or not SDS and microsequence of individual unmodified electroblotted proteins*, **1992**, FEBS Letters, 296(3), pp. 300-304
- [25] Guennadi Sezonov, Danièle Joseleau-Petit, and Richard D'Ari, *Escherichia coli Physiology in Luria-Bertani Broth*, **2007**, J Bacteriol., 189(23), pp. 8746-8749
- [26] Tartoff K. D. and Hobbs C. A., *Improved Media for Growing Plasmid and Cosmid Clones*, **1987**, Bethesda Res. Lab. Focus, pp. 9-12
- [27] Sambrook J., Fritsch, E. E., and Maniatis T., *Molecular Cloning: A Laboratory Manual*, 2nd Ed., **1989**, Cold Spring Harbor Lab., Cold Spring Harbor, N.Y.
- [28] Vinayak Gupta, Kate S. Carroll, *Sulfenic acid chemistry, detection and cellular lifetime*, **2014**, Biochimica et Biophysica Acta, 1840(2), pp. 847-875
- [29] Mauro Lo Conte, Kate S. Carroll, *The Chemistry Of Thiol Oxidation And Detection*, Department of Chemistry, The Scripps Research Institute, USA

- [30] Xian Huang, James Barnard, Thomas M. Spitznagel, Rajesh Krishnamurthy, *Protein Covalent Dimer Formation Induced by Reversed-Phase HPLC Conditions*, Human Genome Sciences, A GlaxoSmithKline Company, USA
- [31] George L. Ellman, *Tissue sulfhydryl groups*, **1959**, Archives of Biochemistry and Biophysics, 82(1), pp. 70-77
- [32] Jill S. Dewey, Douglas K. Struck, Ry Young, *Thiol protection in membrane protein purifications: A study with phage holins*, **2009**, Analytical Biochemistry, 390, pp. 221-223
- [33] Anders Overgaard PEDERSEN and Jørgen JACOBSEN, *Reactivity of the Thiol Group in Human and Bovine Albumin at pH 3 - 9, as Measured by Exchange with 2,2'-Dithiodipyridine*, **1980**, Eur. J. Biochem., 106, pp. 291-295
- [34] Deborah K. Wilkins, Shaun B. Grimshaw, Véronique Receveur, Christopher M. Dobson, Jonathan A. Jones, and Lorna J. Smith, *Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques*, **1999**, Biochemistry, 38(50), pp. 16424-16431
- [35] Harold P. Erickson, *Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy*, **2009**, Biol ProcedOnline., 11, pp. 32-51
- [36] Kelly Zhang, Xiaodong Liu, *Mixed-mode chromatography in pharmaceutical and biopharmaceutical applications*, **2016**, Journal of Pharmaceutical and Biomedical Analysis, 128, pp. 73-88
- [37] Bogusław Buszewski, Renata M. Gadzała-Kopciuch, Michał Markuszewski, Roman Kaliszan, *Chemically Bonded Silica Stationary Phases: Synthesis, Physicochemical Characterization, and Molecular Mechanism of Reversed-Phase HPLC Retention*, **1997**, Anal. Chem., 69(16), pp. 3277-3328

- [38] David A. Barrett, Vikki A. Brown, P. Nicholas Shaw, and Martyn C. Davies, Harald Ritchie, Paul Ross, *Characterization of a Range of Alkyl-Bonded Silica HPLC Stationary Phases: Chromatographic Behavior of Neutral, Acidic, and Basic Test Solutes*, **1996**, Journal of Chromatographic Science, 34, pp. 146-156.
- [39] Marchand DH, Croes K, Dolan JW, Snyder LR, *Column selectivity in reversed-phase liquid chromatography. VII. Cyanopropyl columns*, **2005**, J Chromatogr A., 1062(1), pp. 57-64
- [40] Bocian S, Buszewski B, *Residual silanols at reversed-phase silica in HPLC-a contribution for a better understanding*, **2012**, J Sep Sci., 35(10-11), pp. 1191-200
- [41] L.T. Zhuravlev, *The surface chemistry of amorphous silica. Zhuravlev model*, **2000**, Colloids and Surfaces A: Physicochemical and Engineering Aspects, 173, pp. 1-38
- [42] Ian Baxter, Lisa D. Cother, Carole Dupuy, Paul D. Lickiss, Andrew J. P. White, David J. Williams, *Hydrogen Bonding to Silanols*, Department of Chemistry, Imperial College of Science, Technology and Medicine, UK
- [43] Jacek Nawrocki, *The silanol group and its role in liquid chromatography*, **1997**, J Chromatography A, 779, pp. 29-71
- [44] Méndez A, Bosch E, Rosés M, Neue UD, *Comparison of the acidity of residual silanol groups in several liquid chromatography columns*, **2003**, J Chromatogr A., 986(1), pp. 33-44
- [45] C. Dewaele, P. Mussche, M. Verzele, *Endcapping of high performance liquid chromatography packings derived from silica gel*, **1982**, Journal of High Resolution Chromatography, 5(11), pp. 616-620
- [46] Kenneth D. Mc Murtrey, *Reaction of Silica Gel with Trimethylsilyl Donors Using Conditions Useful for End-Capping HPLC Bonded Phase Packings*, **1988**, Journal of Liquid Chromatography, 11(16), pp. 3375-3384

- [47] Jared Anderson, Alain Berthod, Veronica Pino, Apryll M. Stalcup, *Analytical Separation Science*, pp. 187-189
- [48] Izaak M. Kolthoff, Miran K. Chantooni, *Effect of water and some other hydrogen bond donors and acceptors on p_aH in acetonitrile of mixtures of acids and their salts*, **1967**, *Anal. Chem.*, 39(10), pp. 1080-1084
- [49] Alesia A. Tietze, Pascal Heimer, Annegret Stark, Diana Imhof, *Ionic Liquid Applications in Peptide Chemistry: Synthesis, Purification and Analytical Characterization Processes*, **2012**, *Molecules*, 17, pp. 4158-4185
- [50] Ye Wang, Minglei Tian, Wentao Bi, Kyung Ho Row, *Application of Ionic Liquids in High Performance Reversed-Phase Chromatography*, **2009**, *Int. J. Mol. Sci.*, 10, pp. 2591-2610
- [51] Jolanta Flieger, *Application of Ionic Liquids in Liquid Chromatography*, Medical University of Lublin Poland
- [52] Yi Huang, Shun Yao, Hang Song, *Application of Ionic Liquids in Liquid Chromatography and Electrodriven Separation*, **2013**, *Journal of Chromatographic Science*, 51, pp. 739-752
- [53] J. Flieger & A. Czajkowska-Żelazko, *Ionic Liquids As Mobile Phase Additives In Reversed-Phase High Performance Liquid Chromatography*, **2011**, *Journal of Liquid Chromatography & Related Technologies*, 34(18), pp. 2224-2242

Chapter 4.

Oligomeric states of egg capsule proteins

The oligomeric states of egg capsule proteins were investigated by the use of different electrophoresis techniques and described in more details in this chapter. The egg capsule proteins were known to form oligomers, but which of the PcECPs actively participated in the oligomerization was unknown. Out of the five PcECPs, PcECP -1, -2 and -3 were found to be part of a higher molecular weight oligomer at approximately 200 kDa.

4.1 Introduction

In previous studies, *Pugilina* Egg Capsule Proteins PcECP -1, -2, -3, -4 and -5 were found to constitute the basic monomeric building blocks for the self-assembly of the egg capsule into fibres, and subsequently into a 4-ply layered macrostructure, with each layer oriented perpendicular to each other.

In this chapter, different electrophoresis techniques were used to identify the intermediate oligomeric state of the egg capsule protein. The use of non-reducing SDS-PAGE can be a powerful tool in tandem with Blue Native PAGE. Despite SDS being a strong denaturant, the tightly folded proteins without being subjected to heat was still able to retain most, if not all, native structural conformation.

The use of 2D SDS-PAGE on single excised gel band of the non-reduced protein were incubated in 6 M urea and 10 mM DTT. This method was able to individually analyse the protein subunit from its non-reduced higher oligomeric state, but the tedious handling techniques calls for a novel approach to extract the target protein of interest. The use of dissolvable ethylene glycol diacrylate cross-linked acrylamide gel was dissolved by ammonium hydroxide to extract the non-reduced protein in the supernatant.

With the extraction of the protein, concentration and purity of the sample can be much easier to regulate and handle. Manual insertion of excised gel pieces for use in 2D-PAGE were cumbersome to handle and full denaturation of the protein may not be complete. Using such techniques, this chapter aims to identify which PcECPs are responsible for the intermediate oligomer.

4.2 Experimental Methods

4.2.1 Preparation of crude extract from *P. cochlidium* nidamental gland

A gravid snail in midst of the egg laying process was sedated and anaesthetised in 7.5% w/v magnesium chloride¹ and 5% v/v ethanol² in artificial seawater. The nidamental gland was surgically removed and the viscous egg capsule precursor protein was extracted. It was stored without any additives in -80°C.

The native crude extract was prepared according to the Blue Native Electrophoresis protocol³ with the exception of the detergent. One hundred milligrams of frozen *P. cochlidium* nidamental gland was incubated in 1 ml of native solubilisation buffer containing 50 mM Imidazole, 500 mM 6-Aminohexanoic acid, 1 mM disodium EDTA, 2% w/v CHAPS zwitterionic detergent, pH 7.0 at 4°C. The gland was subsequently manually homogenised with a glass rod on ice and vortex. The crude extract was left in 4°C refrigerator overnight for complete solubilisation.

In the denatured crude extract, 10 mg of *P. cochlidium* nidamental gland was solubilised in 1 ml of 5% v/v acetic acid, 6 M Urea and 10 mM TCEP. It was sonicated and vortex for 10 minutes each. Both native and denatured crude extracted were centrifuged at 21,500 xg (Hitachi CT 15E) for 5 minutes and the supernatant was filtered through a sterile 0.2 microns syringe filter into a new sterile 2 ml Eppendorf tube. Zwitterionic detergent CHAPS and n-Octyl- β -D-Glucopyranoside were known to be quite efficient in solubilizing membrane proteins and protein-protein interacted structures. Both detergents are also easily dialyzable to remove the excess concentration.

4.2.2 Blue Native Polyacrylamide Gel electrophoresis

Blue native gel electrophoresis buffers were prepared as described by Schägger and von Jagow in Chapter 3 Section 3.2.5, but adapted to Invitrogen NativePAGE™ Bis-Tris gel system. The acrylamide/bis concentration was formulated to 30% T, 3.3% C. All BN-PAGE gels were hand-casted 4% to 16% gradient gels unless otherwise stated. Lyophilised bovine serum albumin was used as the native molecular weight ladder reference at a concentration of 1 mg/ml. Coomassie Brilliant Blue G-250 1% w/v was diluted in 150 ml of cathode buffer to 0.02% (0.3 ml 1% w/v CBG-250). Blue native PAGE gel and buffers composition are listed in Table 4.1.

Table 4.1. Blue native gel and running buffers protocol

Blue native polyacrylamide gel protocol					
Gel	Acrylamide/Bis 30%T, 3.3%C	3x gel buffer 150 mM BisTris-HCl, 1.5 M 6-Aminohexanoic acid*, pH 7.0	Degassed Milli-Q Water	10% APS	TEMED
4% Stacking	240 µl	605 µl	937 µl	15 µl	1.5 µl
10% Running	1,200 µl	1,210 µl	1,154 µl	30 µl	3 µl
12% Running	1,460 µl	1,210 µl	894 µl	30 µl	3 µl
16% Running	1,940 µl	1,210 µl	414 µl	30 µl	3 µl

* 1.5 M 6-Aminohexanoic acid is optional

Blue native electrode buffers			
1x Concentration	BisTris (mM)	Tricine (mM)	pH**
Cathode and Anode	50	50	6.8

** No adjustment of pH

Blue native sample buffer					
4x Concentration	BisTris (mM)	NaCl (mM)	Glycerol (%)	Ponceau S (%)	pH*
Sample buffer	50	50	10	0.001	7.2

*Titrate with 6 M HCl to pH 7.2 (107 µl for 10 ml sample buffer)

One microliter of CBG-250 sample loading buffer additive (5% w/v CBG-250, 500 mM 6-Aminohexanoic acid) was used for all BN-PAGE samples with detergent. Bovine serum albumin protein 1 µl in volume and 5 µl of native crude extract were loaded. The total well volume was brought up to 10 µl with native solubilisation buffer. The electrophoresis was ran at 100 V for the first 30 minutes and increased to 300 V until the tracking dye reaches the bottom of the polyacrylamide gel. The electrophoresis unit is kept at 4°C at all times during the run.

4.2.3 Tris-Glycine SDS-PAGE

The Laemmli Tris-Glycine system protocol for resolving mid to high molecular weights. A constant voltage of 150 V was used for SDS-PAGE. Biorad Precision Plus Protein™ All Blue Prestained standard was used as the SDS-PAGE molecular weight reference. Gel and running buffers composition are listed in Table 4.2. Sample buffer 2x concentration was purchased from Biorad.

Table 4.2. Tris-Glycine gel and running buffers protocol.

Tris-Glycine polyacrylamide gel protocol						
Gel	Acrylamide/Bis 30%T, 3.3%C	4x gel buffer 1.5 M Tris-HCl pH 8.8	Degassed Milli-Q Water	10% SDS	10% APS	TEMED
4% Stacking	240 µl	605 µl	937 µl	18 µl	15 µl	1.5 µl
10% Running	1,200 µl	1,210 µl	1,154 µl	36 µl	30 µl	3 µl
12% Running	1,460 µl	1,210 µl	894 µl	36 µl	30 µl	3 µl
16% Running	1,940 µl	1,210 µl	414 µl	36 µl	30 µl	3 µl

Tris-Glycine electrode buffers				
1x Concentration	Tris (mM)	Glycine (mM)	SDS (%)	pH*
Cathode and Anode	25	250	0.1	8.3

*No adjustment of pH

4.2.4 Tris-Tricine SDS-PAGE

Tris-Tricine²⁰ system protocol for resolving low molecular weight proteins. A constant voltage of 150 V was used for SDS-PAGE. Biorad Precision Plus Protein™. All Blue Prestained standard was used as the SDS-PAGE molecular weight reference. Gel and running buffers composition are listed in Table 4.3. Sample buffer was purchased from Biorad.

Table 4.3. Tris-Tricine gel and running buffers protocol.

Tris-Tricine polyacrylamide gel protocol							
Gel	Acrylamide/Bis 49.5%T, 6%C	3x gel buffer 3 M Tris-HCl pH 8.45 ^b	Degassed Milli-Q Water	50% Glycerol	10% SDS	10% APS	TEMED
4% Stacking	150 μ l	605 μ l	907 μ l	180 μ l	18 μ l	15 μ l	1.5 μ l
16% Running	1,110 μ l	1,210 μ l	884 μ l	360 μ l	36 μ l	30 μ l	3 μ l

Tris-Tricine electrode buffers				
1x Concentration	Tris (mM)	Tricine (mM)	SDS (%)	pH
Cathode	25	250	0.1	8.25 ^a
Anode	50	-	-	8.9 ^b

^a No adjustment of pH^b Adjustment of pH with hydrochloric acid

4.2.5 Two Dimensional SDS-PAGE

The protein bands of interest for 2D SDS-PAGE analysis were carefully excised and incubated in 2D equilibration buffer⁴ containing SDS and 6 M urea. Reducing agent 2% v/v β -mercaptoethanol and alkylating agent 2.5% w/v iodoacetamide was subsequently added to reduce and irreversibly prevent the reformation of any disulphide bonds.

The excised bands were carefully pushed to the bottom of the well using the stacking gel comb, starting with the thinnest, then to the thickest band. A five well 10% SDS gel using the Tris-Tricine system was used. Tricine sample loading buffer was added into each well until it fills up to the height of the excised gel piece. An example on how the experiment was carried out is shown in Figure 4.1.

Figure 4.1. Schematic diagram of manually excised protein bands for 2D SDS-PAGE as an example. In this schematic, a native protein comprised of 4 major protein bands, red, orange, yellow and green. When denatured, the major protein bands disassociate into its sub-units and monomers. By isolating the major bands in the native lane, the sub-units can be clearly determined in the second SDS-PAGE gel.

4.2.6 Ethylene Glycol Diacrylate polyacrylamide gel

Ethylene Glycol Diacrylate polyacrylamide gel was prepared accordingly to either of the Blue Native, Tris-Glycine, Tris-Tricine or Neutral pH electrophoresis system used. Bisacrylamide was replaced with Ethylene Glycol Diacrylate of the same cross-link percentage (3.3% C or 6% C).

4.3 Results and Discussion

4.3.1 Protein concentration of crude extracts

Protein concentration of the crude extracts was determined using both nanodrop and Bradford assay. At 0.5% w/v CHAPS and nO β DG, the native protein concentration was about 4 mg/ml and 3.2 mg/ml respectively. At 2% w/v detergent concentration, values are about 6 mg/ml for CHAPS and 5 mg/ml for nO β DG. The denatured crude extract with 10 mg gland has a protein concentration of about 1 mg/ml. Detergent concentration above 2% w/v did not further increase the protein concentration. The typical concentration of native crude extract and denatured crude extract used in following experiments are 6 mg/ml and 1 mg/ml respectively.

4.3.2 Blue native polyacrylamide gel electrophoresis

The most prominent protein band appears intensively above 198 kDa (Figure 4.2A), indicating one of the more stable protein oligomer states. Four other protein bands were also seen around 66 kDa. The three molecular weight references with their corresponding relative mobility (R_f) were curve fitted with a quadratic equation to generate a calibration curve (Figure 4.2C). Subsequently, the molecular weights of the nine detected peaks in the native crude extract were calculated.

Figure 4.2. A) Blue native PAGE of bovine serum albumin and native crude extract. B) Peak detection of native crude extract was done using GelAnalyser 2010a (Istvan Lazar, 2010). C) A relative mobility plot against molecular weight reference of bovine serum albumin and the estimated native crude extract molecular weight. Red line represents the trendline for the three data points and the blue line represents the extrapolated trendline.

A comparison was done to show that the native high molecular weight proteins disassociated into lower molecular weights. The native crude extract of 2.5 μ l in volume was denatured by heating at 90°C for 10 minutes in 5 μ l Laemmli sample loading buffer, 2% v/v β -mercaptoethanol and compared against a denatured crude extract of 1 μ l in volume. The electrophoresis was done using the Tris-Glycine system with a 10% polyacrylamide gel.

Figure 4.3. A reducing 10% SDS-PAGE comparing the protein subunits in both native and denatured crude extract. Lane 1 contains the molecular weight reference. Lane 2 contains the native crude extract treated with 2% v/v β -mercaptoethanol and heat. Lane 3 contains denatured crude sample in 6 M urea and 5% v/v acetic acid.

The denatured crude extract with 6 M urea and 5% v/v acetic acid shows more intense protein bands than the denatured native crude with molecular weights between 37 kDa and 50 kDa. Lowering the concentration of the denatured crude did not help resolve the bands further. The degradation of urea might have decreased the sharpness and resolution of the bands.

A side by side comparison of the native and denatured crude would give a more complete picture on the native oligomeric state. However the CBG-250 anionic dye binds strongly to the hydrophobic site in the protein. With the denatured crude having its hydrophobic core exposed, CBG-250 will immediately aggregate the protein upon contact. Therefore BN-PAGE cannot be used for the comparison.

4.3.3 Non-reducing SDS-PAGE

Non-reducing SDS-PAGE avoids the use of heat and thiol reducers. This electrophoresis system is able to keep most of the native protein structure conformation. The use of standard SDS-PAGE molecular weight ladder is possible as well. A 10% polyacrylamide gel was casted and Tris-Tricine system was used to separate a broad range of molecular weights ranging from 10 kDa to 250 kDa. A total of 9 native crude extract from both *Pugilina* (PC) and *Busycotypus* (BC) nidamental gland were loaded. Each lane contains 2.5 µl of protein, 5 µl tricine sample loading buffer and 2.5 µl native solubilisation buffer.

Figure 4.4. Non-denaturing and non-reducing SDS-PAGE gel of 9 native crude extracts from different snails. Crude extract for *Pugilina* (PC) and *Busycotypus* (BC) nomenclature was chosen for the respective species.

High molecular weight proteins could be seen around 250 kDa as well as some heavier oligomers in the stacking gel. Many fainter proteins disassociated throughout the entire molecular weight range from the use of SDS detergent. There were no significant protein bands lower than 20 kDa. Using the same non-denaturing Tris-Tricine SDS-PAGE system, a 4% to 16% hand-casted gradient gel was made, which allowed higher molecular weight proteins to separate better while keeping low molecular weight proteins in the gel.

Figure 4.5. **A)** Gradient 4% to 16% SDS-PAGE gel with molecular weight ladder and 9 native crude extract for comparison. **B)** Peak detection of lane 8 native crude extract was done using GelAnalyser 2010a (Istvan Lazar, 2010). **C)** A relative mobility plot against Biorad Precision Plus Protein™ All Blue Prestained standard and the estimated native crude extract molecular weight in Lane 8.

The molecular weight ladder was calibrated with respect to the height of the gradient gel. An Exponential Associate Function was curve-fitted in OriginPro 9.0 with ten relative data points, producing a R^2 value of 0.99948. Lane 8 which is of interest was highlighted in red (Figure 4.5A) and the R_f values were substituted. An estimate of the molecular weight protein was determined in Figure 4.5C.

While the molecular weight values are different from those obtained using BN-PAGE, a predicted guess based on the pattern was hypothesized. Band 4 and 5 from the blue native gel had similar patterns with band 2 and 3 from the non-denatured SDS-PAGE gel. This high intensity band of interest is now collectively termed as ‘200 kDa band’ for easy reference.

The native crude extract was denatured and compared against the non-reduced sample in a 4% to 16% gradient gel using NuPAGE® Polyacrylamide Gel Electrophoresis. Five microliters of the protein was loaded into the well for both native and denatured crude extract.

Figure 4.6. Comparison of native crude extract in both denaturing and non-denaturing conditions using NuPAGE® electrophoresis.

The reduced 200 kDa band disassociated into lower molecular weight subunits ranging between 37 kDa and 50 kDa, which corresponded very well to the known PcECPs. Further analysis of the 200 kDa band and PcECP 4 and 5 which shows up in Lane 2 needs to be isolated and individually denatured with the use of 2D SDS-PAGE. *Pugilina* Bundling Protein will be discussed in detail in Chapter 5.

4.3.4 Two dimensional SDS-PAGE

After the SDS-PAGE gel was destained, the 200 kDa, PcECP 4 and 5 band were carefully excised and incubated in 2D equilibration buffer containing SDS and 6 M urea. Reducing agent 2% v/v β -mercaptoethanol and alkylating agent 2.5% w/v iodoacetamide was subsequently added to reduce and irreversibly prevent the reformation of any disulphide bonds.

The 200 kDa band disassociated into similar molecular weights to that of the denatured crude extract control. This clearly suggests that the 200 kDa band is made up of PcECPs. The denatured '200 kDa' has lesser bands compared to the denatured crude in the 40 kDa region. The missing bands could be assembled into other oligomers other than 200 kDa or isoforms of PcECPs. The PcECP 4 and 5 which was extracted remained in the same position and no relation of this protein to higher oligomer state can be concluded. The two proteins were also not seen in the 200 kDa band.

Excision of protein bands for 2D-PAGE proved to be an effective method for isolating protein bands of interest and determining their respective subunits. Extra care was needed using this method as the excised gels pieces are delicate to handle after swelling in solution. It is easily broken into pieces by mishandling or excessive force picking up and sliding it into the wells. To overcome the challenge of excision and handling of the gel pieces, the use of degradable polyacrylamide gel to extract the proteins was used.

Figure 4.7. Two dimensional SDS-PAGE of 200 kDa and PcECP 4 and 5 excised gel bands. The native crude extract was denatured and used as the control in Lane 4.

4.3.5 Ethylene Glycol Diacrylate cross-linked polyacrylamide gel

The PcECP 200 kDa protein gel slices were carefully excised and collected across all nine lanes and dissolved in 2 ml of 500 mM ammonium hydroxide with 1% w/v CHAPS detergent. The original protocol used SDS, however the use of CHAPS had given similar results as it was able to prevent any further denaturation of the non-reduced protein. Chloroform-methanol precipitation method⁵ of the protein was subsequently used. Four millilitres of methanol was added to precipitate the polyacrylamide filaments. This step was repeated again with additional 4 ml of methanol. After centrifugation, the supernatant was transferred to a new 50 ml centrifuge tube and 2 ml of 500 mM HCl was added to neutralise the NH_4OH .

Two millilitres of chloroform was added and after vortexing, an additional 4 ml of Milli-Q water was added, then vortex and centrifuged at 21,500 xg for 10 minutes. The protein layer at the chloroform-aqueous solution interface was carefully pipetted and inserted into a 10 kDa cut-off dialysis bag. The protein was dialysed against water

overnight and subsequently lyophilised. The lyophilised 200 kDa native protein was resolubilised in Milli-Q water and centrifuged to pellet aggregates. It was then loaded into a large format 4% to 16% gradient gel using the Tris-Tricine system.

Figure 4.8. Zinc-imidazole negative stain of an ethylene glycol diacrylate polyacrylamide gel against a black background. The 200 kDa band across the nine lanes were excised for gel digestion with 500 mM ammonium hydroxide and protein extracted using chloroform-methanol precipitation.

The extracted PcECP 200 kDa protein from ethylene glycol diacrylate dissolvable gel using the modified chloroform-methanol precipitation protocol gave a rather clean band at 200 kDa in Lane 6. No lower molecular weight proteins were detected and this demonstrates that the method works quite well for native protein extraction. There was also control over the protein concentration to give the gel bands sufficient resolution.

The higher molecular weights above 200 kDa in Lane 6 suggest that the protein is still active in oligomerisation. All the bands seen from the 200 kDa protein in Lane 6, 7 and 8 corresponded nicely in comparison with the native crude extract in Lane 2, 3 and 4. The PcECP 4 and 5 were not present in the 200 kDa band after denaturing and alkylating it.

Lane 1: Molecular weight ladder
Lane 2: Non-reducing native crude extract
Lane 3: Reducing native crude extract with β -mercaptoethanol
Lane 4: Reducing native crude extract with β -mercaptoethanol and iodoacetamide
Lane 5: Molecular weight ladder
Lane 6: Non-reducing extracted PcECP 200 kDa band from EDA gel
Lane 7: Reducing extracted PcECP 200 kDa band from EDA gel with β -mercaptoethanol
Lane 8: Reducing extracted PcECP 200 kDa band from EDA gel with β -mercaptoethanol and iodoacetamide
Lane 9: Molecular weight ladder

Figure 4.9. Large format 4% - 16% Tricine SDS-PAGE gel loaded with native crude extract as reference and EDA extracted 200 kDa band in lanes 6, 7 and 8.

4.4 Conclusions

Determining the molecular weights of native proteins using Blue native gel electrophoresis has proven to be challenging. Protein mobility is affected by its size, intrinsic charge, isoelectric point, shape and even buffer pH. Electrophoretic analysis of protein relative mobility could only give an estimation of the molecular weights with curve fitting of the molecular weight ladder.

An alternative method to BN-PAGE is the use of non-denaturing and non-reducing SDS-PAGE. While SDS is a strong denaturing detergent, it is still unable to unfold tightly folded proteins, although caution should still be exercised when analysing the protein bands, especially for weak protein-protein interactions. Non-denaturing and non-reducing SDS-PAGE is very useful for side-by-side comparison between the native crude extract and its denatured derivative.

Two-dimensional SDS PAGE with manual gel excision of the target protein was able to isolate and determine the 200 kDa protein sub-units that made up its oligomer. To further improve the handling capability of the target protein, ethylene glycol diacrylate dissolvable polyacrylamide gel coupled with zinc-imidazole negative stain and chloroform-methanol precipitation proved to be very effective in retrieving native proteins.

Using the above methods, it was identified that the 200 kDa was still active and forming oligomers of higher molecular weight, and this was made up of PcECP 1, 2, 3. Both PcECP 4 and 5 were found to be independent from any self-assembly role in these experiments.

References

- [1] J. B. Messenger, M. Nixon, K. P. Ryan, *Magnesium Chloride as an anaesthetic for cephalopods*, **1985**, Comparative Biochemistry and Physiology Part C: Comparative Pharmacology and Toxicology, 82(1), pp. 203-205
- [2] John E. Cooper, *Anesthesia, Analgesia, and Euthanasia of Invertebrates*, **2011**, Institute For Laboratory Animal Research Journal, 52(2)
- [3] Ilka Wittig, Hans-Peter Braun, Hermann Schägger, *Blue native PAGE*, **2006**, Nature Protocols, 1, pp. 418-428
- [4] Stochaj WR, Berkelman T, Laird N, *Preparative 2D Gel Electrophoresis with Immobilized pH Gradients: IPG Strip Equilibration*, **2006**, CSH Protoc..
- [5] Wessel D, Flügge UI, *A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids*, **1984**, Anal Biochem., 138(1), pp. 141-143

Chapter 5.

Expression and Purification of *Pugilina* Bundling Protein

The newly discovered Pugilina Bundling Protein which might be responsible for the cross-linking of PcECPs into a soft gel was expressed in E. Coli and purified using High Pressure Liquid Chromatography. Biochemical characterisation of the protein revealed that the protein highly hydrophobic and contains many cysteines. Protection of cysteine sulfhydryl group is critical as oxidation leads to non-specific oligomerisation which may be irreversible. The stable state of this protein is in its dimeric form.

5.1 Introduction

After the preparation of every native crude extract, it was subjected to electrophoresis to determine whether the proteins were properly extracted. Snails which are not sexually mature will have very low or even none of the egg capsule precursor proteins. An attempt to separate the native protein using size exclusion HPLC and FPLC was previously carried out (latter not shown in this report).

Nine crude extract were analysed with electrophoresis as shown in Chapter 4 Figure 4.4. *In Lane 8 of that gel, a concentrated protein band of approximately 21.9 kDa as determined by its relative mobility, stood out among the other eight native crude extract.* Polyacrylamide gel electrophoresis that used Tris-Glycine system had proteins with 20 kDa and lower run out of the gel since the focus had previously been on PcECPs in the 37 kDa to 50 kDa range.

The nidamental gland that was used to make the native crude extract in Lane 8 of Figure 4.4 was solubilised in 6 M Urea, 5% v/v acetic acid and 10 mM TCEP. Both native and denatured crude extracts were left in 4°C refrigerator for a few days. It was noticed that the viscosity of the native crude extract had increased over time. A gel-like material eventually starts to form and slowly grew larger in size over the next few days. The other eight samples which lacked the 21.9 kDa protein did not exhibit any polymerisation occurring in the extract.

The polymerised crude extract was obtained from the nidamental gland of a gravid snail in midst of laying egg capsules. Nidamental gland extracted before and after the egg laying process did not have this low molecular protein present. This new low molecular weight protein was thus hypothesized to be responsible for the cross-linking process as thus it was termed as “*Pugilina* Bundling Protein”, abbreviated as PcBP in this report.

The crude extract in 6 M Urea, 5% v/v acetic acid and 10 mM TCEP formed aggregates that are dispersed in solution, thin fibres could be seen entangled together and easily broke apart upon agitation, whereas the crude extract in native solubilisation buffer at pH 7.0 formed a massive soft gel. The gel polymerization volume was initially only about 5% and when left untouched for the next few weeks, the soft gel grew bigger and denser. It ultimately encapsulated the nidamental gland pellet at the bottom.

Figure 5.1. A) *P. cochlidium* native crude extract in 6 M Urea, 5% v/v Acetic acid, 10 mM TCEP, pH 2.4 which formed dispersed aggregates B) *P. cochlidium* native crude extract in native solubilisation buffer, pH 7.0 which formed a dense soft gel.

The acidity and alkaline values of the solution might give a clue on how the protein is stabilised in the nidamental gland and in contact with seawater. The pH value in the nidamental gland was measured to be in the range of pH 6.5 to pH 7.5 using a 1 mm microelectrode probe. The acidic condition might have helped the proteins to remain soluble, and when in contact with seawater, the alkaline pH catalyse the formation and polymerisation of the protein through the thiol group. Formation of dispersed aggregates in 6 M urea suggest that covalent bonding might have taken place between PcECP and PcBP, although the sclerotisation of the egg capsule surface has yet to be fully understood.

5.2 Protein sequence of *Pugilina* Bundling Protein

The native crude extract containing PcBP was loaded into all nine wells in a 12% SDS-PAGE gel using the non-denaturing and non-reducing Tris-Tricine system. After the run had completed, the protein on the gel was blotted to a polyvinylidene difluoride (PVDF) membrane (Amersham Bioscience TE 77 PWR) for 40 minutes at 20 mA. The entire row containing the protein was cut and sent to Iowa State University for Edman sequencing. The result of the PcBP Edman sequence with the first ten amino acid is listed below.

I V D P E I ? E G Q

Edman sequencing is the cyclic degradation of peptides based on the reaction of phenylisothiocyanate with the free amino group of the N-terminal residue. The amino acids are removed one at a time and identified as their phenylthiohydantoin (PTH) derivatives. It is known that Cysteine (Cys) is typically difficult to identify by Edman sequencing because its PTH derivative is unstable,¹ thus the amino acid will appear as an unknown.

In this case, the seventh amino acid of the PcBP Edman sequence is most probably a cysteine. A match against the transcriptome was found for the other nine known amino acid. The first nineteen amino acid code represent the signal peptide and the subsequent ten amino acid was from the Edman sequencing. Both sequence with and without the signal peptide was entered into ExPASy ProtParam to determine its theoretical isoelectric point and molecular weight.

Table 5.1. Amino acid sequence of PcBP matched in *Pugilina cochlidium* transcriptome database. The characteristic of PcBP was derived from ExPASy ProtParam.

With Signal Peptide			Without Signal Peptide		
MAGLHFALHLLVLIVCARA IVDPEICEGQ VCSSGDV			IVDPEICEGQ VCSSGDVKKVVYEYGKCQYKLVSEPTS		
KKVVYEYGKCQYKLVSEPTSAFVDLSACIELCMNN			AFVDLSACIELCMNNCNGLTCPIMNGGAQKVVYEYHGQ		
NCNGLTCPIMNGGAQKVVYEYHGQCKYKRVSPLCPTA			CKYKRVSPLCPTAFDYIDDCKKACTNRKY		
FDYIDDCKKACTNRKY					
Number of amino acids: 124			Number of amino acids: 105		
Molecular weight: 13669.0			Molecular weight: 11638.4		
Theoretical pI: 7.43			Theoretical pI: 6.70		
<u>Code</u>	<u>Amount</u>	<u>Percentage</u>	<u>Code</u>	<u>Amount</u>	<u>Percentage</u>
Ala (A)	9	7.3%	Ala (A)	5	4.8%
Arg (R)	3	2.4%	Arg (R)	2	1.9%
Asn (N)	6	4.8%	Asn (N)	6	5.7%
Asp (D)	6	4.8%	Asp (D)	6	5.7%
Cys (C)	13	10.5%	Cys (C)	12	11.4%
Gln (Q)	4	3.2%	Gln (Q)	4	3.8%
Glu (E)	6	4.8%	Glu (E)	6	5.7%
Gly (G)	8	6.5%	Gly (G)	7	6.7%
His (H)	3	2.4%	His (H)	1	1.0%
Ile (I)	6	4.8%	Ile (I)	5	4.8%
Leu (L)	10	8.1%	Leu (L)	5	4.8%
Lys (K)	10	8.1%	Lys (K)	10	9.5%
Met (M)	3	2.4%	Met (M)	2	1.9%
Phe (F)	3	2.4%	Phe (F)	2	1.9%
Pro (P)	5	4.0%	Pro (P)	5	4.8%
Ser (S)	6	4.8%	Ser (S)	6	5.7%
Thr (T)	4	3.2%	Thr (T)	4	3.8%
Trp (W)	0	0.0%	Trp (W)	0	0.0%
Tyr (Y)	7	5.6%	Tyr (Y)	7	6.7%
Val (V)	12	9.7%	Val (V)	10	9.5%

Figure 5.2. PcBP amino acid sequence with and without the signal peptide, its theoretical molecular weight and isoelectric point.

The isoelectric point of PcBP (with signal peptide) of 7.4 suggest that the protein synthesized *in-vivo* at physiological pH might be in the form of aggregates. It is hypothesized that the redox environment, cleavage of the signal peptide and muscular contraction mixes, actively refolds and knead the egg capsule proteins together. Subsequently, when the immature egg capsule comes in contact with the seawater at pH 8.5, PcBP at pI 6.7 becomes positively charged and PcECPs approximately at pI 9.0 becomes negatively charged. Ionic and disulphide interactions might have assembled and interlock the fibrils and layers together.

The cleavage of the signal peptide sees the largest reduction in the amount of Histidine and Leucine, with a reduction of 66.7% and 50% respectively. Histidine is an active and versatile amino acid that plays multiple roles in protein interactions, and is often a key component of catalytic enzyme reactions², whereas leucine is the trigger for the synthesis of proteins.³ Once the protein has been secreted from the cell and reaches its final destination, such as with the interaction of other proteins, the signal peptide is finally cleaved off.

Cysteine is responsible for the formation of disulphide bridges, in which a pair is oxidised to form a covalent bond which helps to stabilise the protein structure. Cysteine is also very common in the protein active and binding sites. The binding to metals is important in enzymatic function such as the function of a nucleophile.

Lysine often plays an important role in the overall protein structure by being actively involved in forming salt-bridges. Valine is a hydrophobic amino acid which is often found to be buried in the protein hydrophobic cores, and with the preference of forming β -sheets. Being non-reactive, it is rarely involved in protein function, but it can play a role in substrate recognition such as binding of hydrophobic ligands.⁴

Queries into the composition of amino acid can give a few clues on how the protein behaves and its structural stability. This low molecular weight protein contains high amounts of reactive cysteine and lysine, which most probably possesses multiple reactive sites for interaction with the main egg capsule protein.

The native liquid protein, which was directly extracted from the nidamental gland during the egg laying process was analysed using MALDI-TOF to confirm the molecular weight. Half a microliter of the viscous protein was mixed with two microliters of sinapinic acid matrix and spotted onto the target plate.

Figure 5.3. MALDI-ToF Mass Spectra of PcBP with 11751 Da and 13690 Da corresponding to the bundling protein with possibly its signal peptide that has yet to be cleaved.

Two peaks were detected with molecular mass of 11,751 Da and 13,690 Da respectively and it corresponded very closely with the theoretical molecular mass. The presence of the 13.6 kDa protein probably indicates that some residual protein has yet to have its signal peptide cleaved off. However, it is barely above the background and the 11.7 kDa protein has a much higher intensity.

5.3 Homology Modelling of *Pugilina* Bundling Protein

3D models of PcBP shown in Figure 5.4 were generated using standard homology modelling methods. Three different algorithms were used, namely I-TASSER^{5,6,7}, Quark⁸ and Modeller (EMBL-BLAST) for the generation of the models. I-TASSER is a hierarchical method that identifies structural templates from Protein Data Base (PDB) by iterative template fragment assembly. QUARK does not use the global template information, but rather it builds a model from small fragments (1-20 residues long, up to 200 residues) by replica-exchange Monte Carlo simulation under the guide of an atomic-level knowledge-based force field. Modeller implements comparative protein structure modelling by satisfaction of spatial restraints based on the global sequence alignment provided.

I-TASSER (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>) and Quark model (<http://zhanglab.ccmb.med.umich.edu/QUARK/>) were generated using their respective servers. To generate the structure using Modeller pipeline, a full length sequence of X protein with X residues was used initially as a query to retrieve structural templates from the Protein Data Bank (PDB) using EMBL-BLAST. This method retrieves similar sequences in the structural database that have high sequence homology to the query sequence (in this case PcBP). The protein displayed high sequence similarity to *Plasmodium falciparum* erythrocyte membrane protein 1 (PfEMP1), a family of adhesins (PDB ID: 2Y8D). Using this protein as a template, the ten atomistic models of X protein were hence generated. The model used for further evaluation was selected based on the lowest DOPE score.

Figure 5.4. A) I-TASSER model, B) QUARK model, C) EMBL-Blast model. Red and yellow colour represents α -helices and β -sheets respectively.

All three models were remodelled in AutoCAD 2010 to linearize the folded protein shown in Figure 5.5. This helps to simplify and identify the location of the respective amino acid and their possible functions given the conformation they are in,

such as free cysteine or as part of a helix which has a high propensity to form disulphide bonds.

The overall structural fold from the three simulation methods are similar and displayed the presence of α -helical regions, except the QUARK model which showed the presence of additional β -sheets. Comparative analysis of the 3D model generated using I-TASSER and QUARK displayed secondary structure similarity at residue positions 34 to 51 and 87 to 102 as α -helical, however the EMBL-BLAST Modeller structure shows a main α -helix from residue position 54 to 73.

I-TASSER and QUARK are based on fragment-based modelling and local sequence alignment, whereas EMBL-BLAST uses global structural features from the template, thus the discrepancies in secondary structure of the models are noted. Since all three structures are predicted models, further structural refinement of these three models will be subjected to molecular dynamics simulations to get a consensus and a structure with a high confidence level. The structural fold and percentage of secondary structure elements will also be validated experimentally using CD and ATR-FTIR.

A majority of the cysteine residues could be found in the secondary structure of all the models, which could possibly play the role in forming disulphide bonds and hence stabilize the structural fold. To predict the connectivity pattern of the disulphide bonds, DIpro server (<http://scratch.proteomics.ics.uci.edu/>) was used. Out of the 12 cysteines present in the sequence, 5 cysteine pairs were predicted as shown in Table 5.2 below. A further characterization of the reactive binding sites of free cysteine is essential for better understanding of this protein. Cysteines at the following positions are predicted to form disulphide bonds: 12, 27, 36, 46, 50, 55, 60, 77, 86, 96.

Table 5.2. Predicted disulphide bonds ordered by probability in descending order.

Bond Index	Cys1_Position	Cys2_Position
1	86	96
2	55	77
3	12	27
4	36	50
5	46	60

Figure 5.5. Linearized 3D models of I-TASSER, QUARK and EMBL-Blast redrawn in AutoCAD 2010.

5.4 Experimental methods

5.4.1 Transformation of *Pugilina* Bundling Protein gene

The gene of PcBP was purchased from GenScript, cloned in vector pET-26b(+) and transformed into One Shot Top10 chemically competent *E. coli*. The cells were spread on LB agar plates with kanamycin antibiotics and incubated at 37°C overnight. Single colonies were selected and inoculated in 5 ml LB and kanamycin at 37°C with shaking at 250 RPM overnight. Miniprep (Qiagen, Germany) was done to extract larger amount of PcBP plasmids. The working concentration of kanamycin is 30 µg/ml.

The plasmids were subsequently transformed into 500 µl of BL21 (DE3) competent *E. coli*. The cells were spread on LB agar plates with kanamycin and incubated at 37°C overnight. Single colonies were selected and inoculated in 5 ml LB with kanamycin at 37°C with shaking at 250 RPM overnight.

5.4.2 Expression Systems using Luria Bertani and Terrific broth

Tartoff Hobbs Terrific broth contains 12 g of tryptone, 24 g of yeast extract, 4 ml of glycerol, 0.17 M KH₂PO₄ and 0.72 M K₂HPO₄ for every 1 L of culture. Tryptone, yeast and glycerol was added to Milli-Q water up to 900 ml in a baffled shaking flask and autoclaved for 30 minutes at 121°C. One hundred millilitres of potassium phosphate was filtered through a 0.2 µm sterile filter directly into the autoclaved solution after it has cooled to room temperature. One millilitre of kanamycin antibiotics was added before the addition of the overnight culture.

The bacteria culture was incubated at 37°C with shaking at 250 RPM for 5 hours to reach at least OD₆₀₀ 1.5. At 2.5 hours and 5 hours intervals, the culture was removed from the incubator for the removal of excess foam. After five hours, the culture was cooled to 18°C before induction with 1 mM IPTG overnight. Induction at low temperatures helps to minimise the expression of other bacterial proteins, but at the expense of lower expression yield. The cells were aliquot into ten 50 ml centrifuge tubes, centrifuged at 20,700 xg at 4°C (Hitachi CR 22N) for 10 minutes and the supernatant was discarded. This step was repeated again, with each tube having the quantity of 100 ml of cells. The aliquots were stored in -80°C.

5.4.3 Cell Lysis

Ten millilitres of ice cold lysis buffer containing 50 mM of Tris and 200 mM of sodium chloride at pH 8.0 was added for every 100 ml of pelleted wet weight cells. Protease inhibitor (Roche cOmplete), 1 mM PMSF and 10 mM DTT were added to prevent proteolysis and aid the reduction of disulphide bonds. The solution was sonicated on ice and pulse set to 10 seconds 50% duty cycle at 50% power for 10 minutes. For lysis of large volume cell suspension, five passes through the microfluidizer was carried out.

The solution was centrifuged at 10,700 xg at 4°C for 10 minutes and the supernatant discarded. Another 10 ml of lysis buffer with protease inhibitor was added for washing of the inclusion bodies and centrifuged again. Zwitterionic detergent CHAPS 2% w/v and Triton X 2% v/v was added to 10 ml of lysis buffer, sonicated and centrifuged. This step was repeated once. The final washing step was performed thrice to remove the excess detergent. The washed inclusion bodies were solubilised in 2 ml of 6 M Urea, 10 mM DTT and centrifuged at 21,500 xg at room temperature to pellet debris. For protection of cysteines with 10 mM 2,2'-Dithiodipyridine, DTT was omitted. The supernatant was carefully pipetted to a new Eppendorf tube and 100 µl of glacial acetic acid was slowly added to prevent localised aggregation of the proteins. The solubilised inclusion bodies were dialysed against 5% v/v acetic acid overnight to remove the urea and subsequently 0.2 µm filtered.

5.4.4 Purification by High Performance Liquid Chromatography

Purification of the expressed and solubilised inclusion bodies of PcBP was carried out using Size Exclusion and Reversed Phased chromatography. These two methods were chosen due to the low molecular weight of the protein and the highly hydrophobic characteristic of the protein.

Size Exclusion chromatography was performed with an Agilent Bio SEC-3 3 µm 100 Å column on Agilent 1260 Affinity. The mobile phase is in 5% v/v acetic acid, set to a flow rate of 1 ml/min in isocratic mode and fractions collected at 1 minute interval for 15 minutes. Detector absorbance is set to UV 280 nm.

Reversed Phase chromatography was performed with an Agilent Zorbax 300SB-C8 semi-preparative column on Agilent 1260 Affinity. The mobile phase is Milli-Q water in 0.1% v/v TFA and Acetonitrile in 0.1% v/v TFA, set to a flow rate of 1 ml/min in gradient mode for a total time of 60 minutes, from (5% to 70%) ACN in 35 minutes and ending at 100% for the remaining 25 minutes. Fractions are collected at 1 minute interval for 60 minutes. Detector absorbance is set to UV 220 nm and 280 nm.

With a better understanding on how different acids, organic solvents and additives affects the elution pattern of PcBP, a new purification method was developed to significantly improve the purity of PcBP using Reversed Phase HPLC. The binary solvent system comprises of methanol, acetonitrile and 5% v/v acetic acid in place of TFA and formic acid.

To overcome the problem of possible silanol ionic interaction on bonded alkyl phase silica, polymeric reversed phase column (Agilent analytical PLRP-S 300 Å 5 µm) was used. The buffers were kept in a reducing environment with 1 mM TCEP as it is stable in acidic conditions and very stable against oxidation. Dithiothreitol and β-mercaptoethanol was not used as the former is only efficient in an alkaline solution and the latter has very high absorbance with the UV detector (20 mM concentration, 350 mAU at 254 nm; 80 mAU at 280 nm). Flow rate was set to 1 ml/min for 40 minutes and fractions collected every minute.

Table 5.3. Binary solvent system gradient timings and elution buffer composition.

Time (minutes)	Percentage (%)		
	Solvent A	Solvent B	Solvent C
0.00	80.0	20.0	0.0
15.00	30.0	70.0	0.0
20.00	30.0	70.0	0.0
20.01	100.0	0.0	0.0
40.00	0.0	0.0	100

Solvent A: 85% H₂O, 10% Methanol, 5% Acetic Acid, 1 mM TCEP

Solvent B: 90% Methanol, 5% H₂O, 5% Acetic Acid, 1 mM TCEP

Solvent C: 100% ACN only

5.4.5 Purification by Dialysis

The low molecular weight of PcBP can be dialysed out into the dialysate with a membrane molecular cutoff of 25 kDa. While higher cutoff dialysis membranes increases the recovered yield, it also dialyse out more contaminants. Regenerated cellulose dialysis membranes were purchased from SpectrumLabs (Spectra/Por™ 6, 25 kDa). The solubilised inclusion bodies were dialysed against two different dialysate, 200 ml 5% v/v acetic acid only and 200 ml 50% v/v ACN with 5% v/v acetic acid for every 1 ml of protein solution. The dialysate were lyophilised for two days. Lyophilised protein were resolubilised in 1 ml degassed 6 M urea or 6 M guanidine hydrochloride with 10 mM TCEP and briefly sonicated. Insoluble aggregates were centrifuged at 21,500 xg for 30 minutes. The supernatant was saved for further HPLC purification and structural refolding.

5.4.6 Neutral pH SDS-PAGE

The neutral pH acrylamide gel was casted according to Table 5.4. A 12% running gel, 4% stacking gel and MOPS buffer was used for electrophoresis.

Table 5.4. Neutral pH SDS-PAGE protocol for Bis-Tris gel buffer and 10x cathode and anode buffer.

Neutral pH polyacrylamide gel protocol					
Gel	Acrylamide/Bis 30%T, 3.3%C	3X Gel buffer 1 M Bis-Tris pH 6.4	Degassed Milli-Q Water	10% APS	TEMED
4% Stacking	240 µl	605 µl	955 µl	15 µl	1.5 µl
12% Running	1,460 µl	1,210 µl	930 µl	30 µl	3 µl
Neutral pH Electrophoresis cathode and anode buffer					
10x Concentration	MOPS	TRIS	SDS	EDTA	pH
MOPS Cathode & Anode Buffer	500 mM	500 mM	1% w/v	0.3% w/v	7.7

To prevent sample reoxidation, 500 µl of antioxidant comprising of 10% w/v sodium metabisulfite and 15% v/v dimethylformamide was added for every 200 ml of running buffer. Non-reduced sample were mixed with NuPAGE® 4x LDS sample loading buffer and loaded into the wells immediately. Reduced samples were added with 0.5 µl of 1 M DTT and heated at 80°C for 10 minutes. It was briefly centrifuged before loading. The running voltage was set to 150 V constant for 80 minutes.

5.4.7 Matrix Assisted Laser Desorption Ionisation-Time of Flight Mass Spectroscopy

Sinapinic acid matrix was prepared to a concentration of 20 mg/ml in 50% v/v Milli-Q water, 50% v/v acetonitrile and 0.1% v/v Trifluoroacetic acid. Using the sandwich layer method to improve sensitivity, 2 µl of the matrix was spotted onto the target plate and air-dried. Half a microlitre of protein solution and 0.5 µl of matrix were mixed and spotted on top of the first layer and air-dried. Lastly 1 µl of matrix was spotted on top of the protein mixture.

5.4 Results and Discussions

5.4.1 Efficiency of expression systems

Empty 50 ml centrifuge tubes were weighed and their weight recorded. The cells were centrifuged, the supernatant discarded and the wet weight cells were weighed. Luria Bertani Broth cultivated a total volume of 340 mg of cells per 100 ml, while Terrific Broth cultivated a total volume of 630 mg of cells per 100 ml.

To further improve the density of the cultured cells, the growth time, temperature and foam removal parameters were tweaked. This resulted in the cells weight wet of up to 1.22 g per 100 ml of culture. The use of Terrific Broth had been proven to be more effective in the growth of *E. coli* bacteria given the same incubation time than that of LB Broth. Despite the more tedious preparation, the Terrific Broth had been chosen for all expression experiment in this project.

Table 5.5. Wet weight of harvested cells per 100 ml with different broth and parameters. Terrific broth was able to harvest up to 1.22 g of wet weight cells compared to LB broth.

Wet weight of harvested cells per 100 ml with different broth and parameters				
Temperature	LB broth		TB broth	
	3 hours growth and 3 hours induction		Baffled shaking flask, 3 hours growth and overnight induction	Baffled shaking flask, 5 hours growth and overnight induction, foam removal
37°C	340 mg	630 mg	1000 mg	-
18°C	-	-	860 mg	1220 mg

5.4.2 Purification by Size Exclusion High Performance Liquid Chromatography

The size exclusion chromatogram of PcBP separation is shown in Figure 5.6A. A reference chromatogram of the column's performance was used to determine the elution time of the expressed protein. Based on the elution time of Ribonuclease A in the reference chromatogram, PcBP of 11.6 kDa eluted slightly slower, at around 7.5 minutes. The broad overlapping peaks suggest that the protein could be having secondary interaction with the stationary phase. Ionic modifiers such as sodium chloride or addition of organic solvent could help improve the peak shape. The three fractions highlighted in Red, Green and Blue were lyophilised and resolubilised in 5% v/v acetic acid for SDS-PAGE.

Figure 5.6. A) Size Exclusion column performance of standards. Ribonuclease A was chosen as the closest reference to PcBP. B) Fraction were collect starting from 6th to 9th minute.

Lane 1: Blank
 Lane 2: Molecular weight ladder
 Lane 3: Cell culture after IPTG induction
 Lane 4: IB after lysis
 Lane 5: IB washed & filtered for HPLC
 Lane 6: Fraction 6
 Lane 7: Fraction 7
 Lane 8: Fraction 8
 Lane 9: Native crude control
 Lane 10: Blank

Figure 5.7. Electrophoresis of size exclusion HPLC fraction 6, 7 and 8. Fraction 7 contains most of the eluted PcBP. Purity of the sample is poor due to broad overlapping peaks suggesting secondary interactions with the stationary phase.

Elution of PcBP occurred at around 7.5 minutes as anticipated, however the poor resolution of the chromatogram resulted in the poor purification method of the protein. Higher molecular weight contaminants above 25 kDa seen in Lane 6 Fraction 6 were significantly lesser than that in Fraction 7 and 8. Fraction 8 gave the best purity but the recovered yield was low. Size Exclusion HPLC was not the best choice for purifying crude inclusion bodies. For small scale analysis, the separation time of the run and the purity of Fraction 8 would be sufficient but at the expense of yield.

5.4.3 Purification by Reversed Phase High Performance Liquid Chromatography

The reversed phase chromatogram of PcBP is shown in Figure 24. Eight fractions were isolated based on the peaks of the chromatogram. Fraction 26, 36, 37, 38, 39, 40, 41 and 42 were lyophilised and resolubilised in 6 M urea. Fraction 11 to 15 were not analysed since it was known to be urea and other polar molecules.

Figure 5.8. Reversed Phase HPLC chromatogram of PcBP using acetonitrile gradient. Fractions 26, 36 to 42 were collected for electrophoresis to determine the quality of separation.

Electrophoresis using NuPAGE® system was performed on the collected fractions as shown in Figure 24. PcBP was eluted from Fraction 36 onwards. Lane 4 Fraction 37 shows that the majority of the contaminants are removed as indicated by the absence of a contaminant band above PcBP at 11.6 kDa. The concentration and purity

is also significantly higher as compared to the rest of the samples by visually comparing its band intensity. The presence of four higher molecular weight bands suggest that PcBP is possibly forming homo-oligomers in presence of trifluoroacetic acid and acetonitrile during the reversed phase run. This was not observed in size exclusion HPLC using 5% v/v acetic acid. The oligomers present in Lane 3 Fraction 36 could not be reduced to its monomer, indicating irreversible oxidation of the protein.

The dimerization of PcBP suggests that the cysteine had been mildly oxidised to its sulfenic acid form (R-S-OH) catalysed by the acidic conditions of the solvents used in Reversed Phase HPLC, especially with TFA or formic acid up to 0.3% v/v. While this occurrence is absent with the use of 5% v/v acetic acid, this probably indicated that the protein might had been oxidised with oxygen rich buffers, or prolonged incubation in such buffers during the sample preparation stage.

Figure 5.9. Reversed Phase fraction 26, 36 to 42 of PcBP on NuPAGE® SDS-PAGE gel. Lane 4 Fraction 37 shows that PcBP is rather pure, but with formation of homo-oligomers suggesting the oxidation of cysteines during sample preparation stage and subsequently catalysed by trifluoroacetic acid. The homo-oligomers could not be reduced back to its monomer.

5.4.4 Optimisation of Reversed Phase High Performance Liquid Chromatography

Methanol and acetonitrile show different selectivity in the elution pattern of proteins, therefore the elution peaks may differ from the use of either of the single solvent systems. Previous purification methods showed that methanol gave more defined peaks as compared to acetonitrile, probably due to its participation in hydrogen bonding, but at the expense of slower elution time and lower elution strength.

Methanol also caused significant increase in base drift due to the refractive index difference against water. The baseline drift had been corrected in OriginPro 9.0 for Figure 5.10. Acetonitrile on the other hand, gave poorer resolution and much worst selectivity, which co-eluted contaminants with PcBP. Due to the severe peak tailing issue with Kromasil C18 reversed phase column as shown in Figure 5.10A, Zorbax Cyanopropyl column (C3 equivalent) was used in the chromatograms for Figure 5.10B, C and D. Retention effects of PcBP was studied using acetonitrile and methanol in 0.1% TFA. Methanol elution strength is about 25% weaker than acetonitrile due to the increased retention times, but shows much better separation as a peak at 27th minute manage to separate itself from the main peak close to the 30th minute.

As methanol had been shown to be much better than acetonitrile in terms of resolution and separation of PcBP due to hydrogen bonding, retention effects of PcBP using methanol in 0.1% TFA and 0.1% formic acid was studied. The use of different acid has different selectivity on the elution time of different peaks. Firstly, the use of TFA increases the retention time of PcBP due to the increased hydrophobicity from the anion, secondly a contaminant peak eluting behind PcBP using 0.1% TFA is eluted first using 0.1% formic acid. The use of sodium chloride screened the charges of the protein and unreacted silanol which minimises ionic interaction, thus greatly improves the separation and resolution of the peaks.

Ionic liquid 1-Butyl-3-methylimidazolium hexafluorophosphate at 10 mM concentration shows similar resolution of peaks as compared to 100 mM sodium chloride (not shown in this report), however sodium chloride is not miscible in organic solvents, therefore 1-Butyl-3-methylimidazolium hexafluorophosphate is a good alternative for charge screening. Addition of 1 mM to 5 mM TCEP to reduce cysteines also helped to further reduce tailing and improve resolution.

Figure 5.10. **A)** Chromatogram of PcBP purification on Kromasil C18 reversed phase using 0.1% TFA and acetonitrile gradient. Peak tailing of PcBP is significant which indicates strong secondary interactions with the stationary phase. **B)** Retention effects of PcBP was studied using acetonitrile and methanol in 0.1% TFA. Methanol elution strength is about 25% weaker than acetonitrile due to the longer retention times, but shows much better separation. **C)** Retention effects of PcBP using methanol in 0.1% TFA and 0.1% formic acid. The use of different acid has different selectivity on the elution time of different peaks. The use of TFA increases the retention time of PcBP due to the increased hydrophobicity from the anion. **D)** Resolution of peaks with and without the addition of 100 mM sodium chloride. Sodium chloride screens the charges of the protein and unreacted silanol which minimises ionic interaction. The use of salt greatly improves the separation and resolution of the peaks.

With the aid of previous chromatographic studies, the chromatogram of the new binary solvent system used to purify PcBP shown in Figure 5.11. Methanol gradient from 20% to 70% for 20 minutes was used to elute the less hydrophobic contaminant proteins. Immediately after the initial elution, ACN gradient was introduced for the remaining 20 minutes. An intense peak of approximately 550 mAU (254 nm wavelength) at 31st minute was recorded.

Figure 5.11. Chromatogram of PcBP using binary solvent system. The low elution strength of methanol was used in the initial purification of PcBP to selectively elute less hydrophobic contaminant proteins while keeping PcBP in the column. Subsequently acetonitrile was used to wash out all the retained proteins.

Fraction 32, which corresponds to the most intense peak on the chromatogram, and its subsequent elution fractions 33 to 35 were analysed using NuPAGE® electrophoresis system. Lane 3 Fraction 32 of the acrylamide gel shown in Figure 5.12 shows a very pure PcBP as well as the presence of its dimer and insignificant amounts of higher molecular weight contaminants. Fraction 33 to 35 are predominantly contaminants with minor amount of PcBP. The use of degassed buffers, polymeric reversed phase column and the new binary solvent system gave a stark difference in the purity and minor dimerization of PcBP as compared to the generic acetonitrile gradient method. The use of 1-Butyl-3-methylimidazolium hexafluorophosphate in the acetonitrile buffer could potentially further separate the closely eluting contaminants. Purity of Fraction 32 was further determined using MALDI-ToF.

Lane 1: Blank
Lane 2: Molecular weight ladder
Lane 3: Fraction 32
Lane 4: Fraction 33
Lane 5: Fraction 34
Lane 6: Fraction 35
Lane 7: Inclusion bodies crude extract
Lane 8: Blank
Lane 9: Blank
Lane 10: Blank

Figure 5.12. Binary solvent system Reversed Phase HPLC fraction 32 to 35 of PcBP on NuPAGE® SDS-PAGE gel. Fraction 32 shows the highly purified PcBP with low amount of its dimer. Fraction 33 to 35 shows majority of the contaminants successfully separated from PcBP.

5.4.5 Purity of Reversed Phased Fraction by MALDI-ToF

To determine the molecular weight of the homo-oligomers, the amount of contaminants removed and the purity of the sample, MALDI-ToF was used. Fraction 37 was dialysed against 5% v/v acetic acid overnight to remove urea. It was lyophilised and reconcentrated in 5% v/v acetic acid.

The mass spectrograph of Fraction 37 shows the highest peak intensity at 12.1 kDa and its doubly charged species at 6 kDa. Although the molecular differs from that of the theoretical isoelectric point and the relative position on SDS-PAGE, the

formation of homo-oligomers could be deduced with dimers up to decamers and beyond. This suggests that the protein had been heavily oxidised. Multimers of PcBP which interacted with an unknown protein at 18 kDa formed 30 kDa complexes. Purity and separation of PcBP is relatively good apart from the oligomeric formations.

The new binary solvent system method that was developed was highly efficient in purifying PcBP from the solubilised inclusion bodies. The deviation of the molecular mass from the theoretical 11.6 kDa could presumably be non-specific cysteine oxidation with molecular oxygen or metal ions. Nonetheless, the purity was acceptable for refolding despite having very minor level of contaminants between 10 kDa to 11.86 kDa.

Figure 5.13. MALDI-ToF of Reversed Phase HPLC Fraction 37 using acetonitrile gradient. Values are in Daltons. Multiple peaks indicates the formation of homo-oligomers and multimers possibly due to the oxidation of PcBP cysteine residues and catalysed by TFA.

Figure 5.14. MALDI-ToF of Reversed Phase HPLC Fraction 32 using methanol-acetonitrile gradient. Peak at 11.86 kDa show that PcBP is highly pure with minor contaminants between 10kDa to 11.86 kDa. Mild dimerization of PcBP at 23.8 kDa was detected.

5.4.6 Purification by dialysis

The amount of PcBP which was dialysed against 5% v/v acetic acid dialysate was significantly lower than that dialysed against 5% v/v acetic acid with 50% v/v ACN. The use of higher molecular weight cutoff membrane of 50 kDa did not affect the final recovered yield. This could be due to the hydrophobic interaction of the protein with the hydrophilic regenerated cellulose membrane, which caused repulsion from the membrane surface.

It was postulated that the addition of 50% v/v acetonitrile helps to shield the hydrophobic sites on the protein, thus making it more hydrophilic for the protein to diffuse through the membrane and into the dialysate. After lyophilising, the dialysate with acetonitrile was found to have effectively dialyse out approximately 90% of the protein overnight. This approximation was done using HPLC with the dialysed protein concentration compared against the control sample of known concentration before dialysis by peak integration (data not shown).

Lyophilised samples that had been protected with 10 mM 2,2'-Dithiodipyridine appeared yellow in colour and more difficult to solubilise even in 6 M guanidine hydrochloride. The sample was heated to 60°C for 1 hour to assist solubilisation and aggregates were centrifuged at 21,500 xg. The supernatant was subjected to SDS-PAGE to determine the efficiency of the dialysis method.

Figure 5.15. Lyophilised dialysate using **A)** only 5% v/v acetic acid. The hydrophilic solution prevents efficient dialysis of proteins below the pore size of the membrane, thus minimal PcBP was recovered. **B)** Lyophilised dialysate using 50% v/v ACN, which decreases the hydrophobicity of PcBP, thus it is much easier to dialyse through the membrane.

Figure 5.16. PcBP in the dialysate after lyophilising shows effective dialysis. Lane 2 is the inclusion bodies crude extract before dialysis and Lane 3 is after dialysis with 50% v/v ACN and 5% v/v acetic acid. Purity of the dialysed sample is good with a single contaminant band slightly above PcBP. The dimer of PcBP was also able to dialyse out.

The purified fraction of PcBP from the dialysis method was analysed using MALDI-ToF. The level of low molecular weight contaminants with the dialysis method is significantly higher than that purified by reversed phase HPLC as seen with the presence of an intense 10.6 kDa peak, as well as two other minor contaminant peaks flanked to its left and right. As the dialysis method is similar to that of size exclusion chromatography, all proteins below the membrane cutoff molecular weight of 25 kDa would be dialysed out into the dialysate. The PcBP dialysate sample will be much easier to purify since there are only three other contaminants to separate as detected by MALDI-ToF. This method of purification by dialysis coupled with acetonitrile is a good alternative for initial mass purification when protein concentration is high, or when high molecular mass contaminants co-elute very closely with the target protein using different HPLC techniques.

Figure 5.17. MALDI-ToF of PcBP purified by dialysis. Intense peak at 10.6 kDa shows the presence of contaminants below 25 kDa which had been dialysed out.

5.5 Conclusions

The discovery of the new low molecular weight protein termed as *Pugilina* Bundling Protein has led to new insights about the functionality of the cross-linking mechanism that is occurring during the egg capsule formation and laying process. This protein was found in the polymerised crude extract that belonged to a gravid snail in midst of egg capsule deposition. Other crude extracts which lacked this protein remained as a solution.

The protein sequencing of PcBP revealed that it contains high amounts of Cysteine (11.4%), Lysine (9.5%) and Valine (9.5%). Cysteine could be playing an active role as reactive sites for the formation of disulphide bridges, Lysine for salt-bridges and Valine for β -sheets. Structural prediction of PcBP had proved to be challenging with conflicting results from I-Tasser, QUARK and EMBL-Blast due to differences in the location of α -helices and β -sheets availability.

The gene of PcBP was transformed in *E. coli* BL21 (DE3) strain for expression. The use of Terrific Broth as the growth media had proven to be very efficient in

cultivating up to 1.22 g of wet weight cells. This was optimised by growing the cells for 5 hours and inducing overnight, together with excess foam removal.

Conventional reversed phase chromatography was able to purify PcBP from cell contaminants through the analysis of SDS-PAGE and MALDI-ToF, but not without the formation of homo-oligomers and multimers. This was probably due to the mild oxidation of PcBP and oligomerisation during sample preparation catalysed by TFA in the elution buffer. These formations were not seen in Size Exclusion fractions with the use of acetic acid. Therefore switching the acid modifier in the elution buffer may solve the problem of unwanted dimerization.

The elution pattern of PcBP using different acids, organic solvent and additives were studied. The use of methanol and a less hydrophobic acid such as formic acid and acetic acid gave better separation and resolution. The use of sodium chloride to screen the protein charges and underivatised silanols can be effectively replaced by BMIM PF₆ for use with organic solvents.

The new binary solvent system was able to effectively purify PcBP as validated by Neutral PAGE and MALDI-ToF. The methanol gradient selectively elute contaminant proteins while the acetonitrile gradient eluted the remaining PcBP. The use of BMIM PF₆ in the acetonitrile buffer could potentially purify PcBP to a higher degree.

References

- [1] Kiyoshi Nokihara, Naoki Morita, Minoru Yamaguchi & Toshihiro Watanabe, *Sequence Analysis of Cysteine and Cystine Containing Peptides and Proteins by a GasPhase Sequencer with Isocratic Separation of PTH-Amino Acids*, **1992**, Analytical Letters, 25(3), pp. 513-533
- [2] Si-Ming Liao, Qi-Shi Du, Jian-Zong Meng, Zong-Wen Pang, Ri-Bo Huang, *The multiple roles of histidine in protein interactions*, **2013**, Chemistry Central Journal, 7, pp. 44
- [3] Peter J. Garlick, *The Role of Leucine in the Regulation of Protein Metabolism*, **2009**, J. Nutr., 139, pp. 1103-1109
- [4] M.J. Betts, R.B. Russell, *Amino acid properties and consequences of substitutions*, **2003**, In Bioinformatics for Geneticists, M.R. Barnes, I.C. Gray eds, Wiley
- [5] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, *The I-TASSER Suite: Protein structure and function prediction*, **2015**, Nature Methods, 12, pp. 7-8
- [6] A. Roy, A. Kucukural, Y. Zhang, *I-TASSER: a unified platform for automated protein structure and function prediction*, **2010**, Nature Protocols, 5, pp. 725-738
- [7] Y. Zhang, *I-TASSER server for protein 3D structure prediction*, **2008**, BMC Bioinformatics, 9, pp. 40
- [8] D. Xu, Y. Zhang, *Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field*, **2012**, Proteins, 80, pp. 1715-1735

Chapter 6.

Refolding of *Pugilina* Bundling Protein

Refolding of purified PcBP inclusion bodies is necessary to make the protein bioactive and execute its intended function. In this chapter, the role of L-Arginine and redox pair are discussed. L-Arginine had been shown to suppress aggregation during refolding and redox pair were necessary to ensure that the disulphide bonds are reshuffled into the correct bonding site. The refolded PcBP were mixed together with native PcECPs to investigate the occurrence on any form of soft-gel or aggregation.

6.1 Introduction

Proteins overexpressed in BL21 (DE3) *E. coli* are directed and deposited in the periplasmic space as inclusion bodies.¹ The Origami™ 2 strain has mutations in both the glutathione reductase (*gor*) and thioredoxin (*trxB*) gene which greatly enhances disulphide bonds by having an oxidising cytoplasm. Therefore the expressed protein is more likely to be in the soluble form.² The intermediate misfolding of inclusion bodies tends to act as a nucleator for proteins to aggregate further.³ Exposed hydrophobic surfaces on the protein competes with refolding, but at high concentrations, aggregation by hydrophobic interaction tends to dominate.⁴

In Chapter 5, PcBP was found to be stable in its dimeric state through disulphide bonds. Refolding of the protein was a challenge as it is highly hydrophobic and a non-stabilised refolding buffer caused aggregation quickly during the early stages of refolding. Nevertheless, refolding of the protein is necessary to ensure its defined activity as well as to investigate the role of self-assembly with PcECPs.

Common refolding techniques such as step wise dialysis and drop wise dilution among others had shown varying levels of successful refolded yield on different protein, but each techniques has its pros and cons.⁵ While there are no single standard protocol for refolding, experimental trial and error with different additives and concentrations need to be determined.

Cysteine rich proteins like that of PcBP require a redox buffer system, generally with glutathione (GSH) and oxidised glutathione (GSSH) to mimick the *in-vivo* conditions of secretory pathways.^{6,7,8} Alternative redox pairs such as cysteine/cystine pair had also been proven useful in refolding such proteins.⁹ Redox pairs aids in the reshuffling of disulphide bonds by breaking and oxidising free sulfhydryl groups until a thermodynamically stable disulphide bond with the least free energy is achieved.^{10,11,12} Without redox pair, disulphide bonds may form between cysteines which are not intended for each other. This often leads to aggregation shortly after refolding begins.

Bakthisaran *et al.* suggested that cysteine/cystine redox pair had a much higher efficiency at refolding denatured lysozyme as compared to GSH/GSSH. This was probably attributed to lower net ionic charge as well as the smaller molecular size of

cysteine and cystine, which has less steric hindrance, thus enabling these molecules to access the cysteine residues in the partially folded protein.¹³

L-Arginine at high concentration had also been found to stabilise active as well as denatured proteins¹⁴, but the mechanisms of stabilisation are not well understood. Tsumoto *et al.* proposed that the guanidinium group and tryptophan side chains may have an active role in the suppression of protein aggregation.^{15,16} In another study by Chen *et al.*, the use of 0.5 M L-Arginine prevented the formation of soluble oligomers produced by hydrophobic interaction in recombinant human granulocyte colony-stimulating factor (rhG-CSF).^{17,18}

Other studies on refolding using both high concentration of L-Arginine and cysteine/cystine redox pair had consistently refolded various proteins successfully at respectable yields. The refolding buffer for PcBP will thus include these two important and widely used additives.

6.2 Experimental Methods

6.2.1 Refolding Buffers

Purified and lyophilised protein was solubilised in pre-chilled 6 M urea and 10 mM DTT to a concentration of 0.5 mg/ml. Pre-chilled dialysis buffer A was added to the solution with a 1:1 ratio to a final concentration of 0.25 mg/ml and centrifuged to pellet any aggregates.

The protein solution was dialysed against Buffer A, B and C at 4°C overnight respectively using SpectrumLabs (Spectra/Por™, 1 kDa) regenerated cellulose membrane. Due to the highly hydrophobic nature of denatured PcBP, 500 mM L-Arginine was added as an additive to suppress aggregation. L-Arginine has a guanidine-like structure which is believed to aid in stabilising the hydrophobic areas in the protein. Chelating agent EDTA scavenges for free metal ions which is known to catalyse disulphide formation.

Table 6.1. Refolding buffer composition containing L-Arginine as aggregation suppressor and cysteine/cystine redox pair for proper disulphide reshuffling.

	Buffer A (500 mL)	Buffer B (500 mL)	Buffer C (500 mL)
HEPES (mM)	50	50	50
L-Arginine (mM)	500	500	50
NaCl (mM)	500	500	50
Disodium EDTA (mM)	1	1	1
L-Cysteine (mM)	5	5	-
L-Cystine (mM)	0.5	0.5	-
pH	Adjust to pH 8.0	Adjust to pH 8.0	Adjust to pH 8.0
DTT (mM)	0.1	Omit DTT	-

Note. All buffers were vacuum degassed and 0.2 microns filtered after the addition of DDT. The use of EDTA free acid may cause itself to precipitate at 4°C.

6.2.2 Neutral pH Polyacrylamide Gel Electrophoresis

The neutral pH acrylamide gel was casted according to Table 6.2. A 12% running gel, 4% stacking gel and MOPS buffer was used for electrophoresis.

Table 6.2. Neutral pH SDS-PAGE protocol for Bis-Tris gel buffer and 10x cathode and anode buffer.

Neutral pH polyacrylamide gel protocol					
Gel	Acrylamide/Bis 30%T, 3.3%C	3X Gel buffer 1 M Bis-Tris pH 6.4	Degassed Milli-Q Water	10% APS	TEMED
4% Stacking	240 µl	605 µl	955 µl	15 µl	1.5 µl
12% Running	1,460 µl	1,210 µl	930 µl	30 µl	3 µl
Neutral pH Electrophoresis cathode and anode buffer					
10x Concentration	MOPS	TRIS	SDS	EDTA	pH
MOPS Cathode & Anode Buffer	500 mM	500 mM	1% w/v	0.3% w/v	7.7

To prevent sample reoxidation, 500 μ l of antioxidant comprising of 10% w/v sodium metabisulfite and 15% v/v dimethylformamide was added for every 200 ml of running buffer. Non-reduced sample were mixed with NuPAGE® 4x LDS sample loading buffer and loaded into the wells immediately. Reduced samples were added with 0.5 μ l of 1 M DTT and heated at 80°C for 10 minutes. It was briefly centrifuged before loading. The running voltage was set to 150 V constant for 80 minutes.

6.2.3 Refolded PcBP - PcECs aggregation

To ensure PcBP is correctly refolded and active, 20 μ l of the refolded PcBP was added to 20 μ l of non-polymerized native crude extract and gently mixed with the pipette. Negative controls were carried out to ensure that (1) the refolding buffer did not caused the aggregation of the native crude extract. (2) The native solubilisation buffer used in the native crude extract did not caused the aggregation of refolded PcBP.

6.2.4 Optical Microscopy of Refolded PcBP - PcECs aggregates

Optical microscopy images of the aggregates were captured using a Canon EOS 6D camera attached to Zeiss microscope. Transmission mode was selected. Twenty microliters of native unpolymerised PcECP crude extract was spotted onto a microscope glass slide, followed by adding 20 μ l of refolded PcBP. The mixture was gently swirled in a circular motion. Magnification at 2.5x and 10x were used.

6.3 Results and Discussion

6.3.1 Neutral PAGE Electrophoresis of refolded PcBP

In Lane 3 of the SDS-PAGE, HPLC fraction 32 was used as the control in the denatured state. Lane 4 shows the refolded PcBP which correspond to the molecular mass seen in Figure 6.1. A band at the top of lane 4 indicates that minor amount of contaminants was present, and the cysteines on PcBP were also forming disulphide bonds with it. Below the refolded PcBP band, two other bands were also seen. These two contaminants were detected in MALDI-ToF in Chapter 5 Figure 5.14. Lane 5 is the reduced dimer of the refolded PcBP.

The refolded PcBP shows a discrepancy against the reference molecular mass (~23.2 kDa vs ~18 kDa) which might be due to the way the protein is folded. It is known that a rod shaped and a globular shaped protein migrates differently in electrophoresis gel. PcBP in Lane 4 might be folded into a globular shaped and its structure held tightly together by disulphide bonds. This possibly prevented SDS from completely linearizing the protein, thus the further migration down the lane.

Lane 1: Blank
 Lane 2: MWL
 Lane 3: PcBP Fraction 32
 Lane 4: Refolded PcBP from Buffer C (non-reduced sample)
 Lane 5: Refolded PcBP from Buffer C (reduced sample)

Figure 6.1. Gel electrophoresis of refolded PcBP. Lane 3 is the purified protein from reversed phase chromatography. Lane 4 shows refolded PcBP in the non-reduced state and Lane 5 shows refolded PcBP in the reduced state.

6.3.2 Optical Microscopy of Refolded PcBP - PcECPs aggregates

In the left image (2.5x magnification), aggregates of the refolded PcBP and PcECPs formed instantaneously when mixed together. The turbidity and opaqueness increased overtime and can be catalysed by swirling the solution. The solution did not form a soft gel probably due to the low concentration of the refolded PcBP. In the right image (10x magnification), micro-dispersed and rod-like aggregates was present. The aggregates were also insoluble in 6 M guanidine hydrochloride and 1 mM TCEP.

Figure 6.2. **A)** Optical microscope image of native PcECP solution and refolded PcBP mixture at 2.5x magnification. **B)** Optical microscope image of native PcECP solution and refolded PcBP mixture at 10x magnification.

6.4 Conclusions

The refolding and aggregation experiments came as an initial breakthrough in the self-assembly and cross-linking of PcECPs. The refolding buffer formulated was successful in aggregating the native PcECPs extract. Omission of L-arginine and cysteine/cystine redox pair caused aggregation of PcBP, suggesting that the hydrophobic nature and proper disulphide bonding is crucial to the stability and bioactivity of PcBP.

The induced aggregation of PcECP with the correct refolded PcBP proved the hypothesis that this protein might be responsible for the self-assembly and cross-linking of PcECP. The negative control was tested with the refolding buffer to verify that it did not aggregate the PcECPs. The unfolded PcBP also did not cause PcECP to aggregate. Although this preliminary small scale experiment looks promising, more characterisation on the aggregates is needed to be carried out. Protein concentration might also be a critical factor as both the native PcECPs and refolded PcBP amount is low, which prevents the aggregate from forming a soft gel, instead become dispersed aggregates.

References

- [1] S. Barth, M. Huhn, B. Matthey, A. Klimka, E. A. Galinski, A. Engert, *Compatible-Solute-Supported Periplasmic Expression Of Functional Recombinant Proteins Under Stress Conditions*, **2000**, Applied And Environmental Microbiology, 66(4), pp. 1572–1579
- [2] Sheetij Dutta, Lisa A. Ware, Arnaldo Barbosa, Christian F Ockenhouse, David E. Lanar, *Purification, Characterization, and Immunogenicity of a Disulfide Cross-Linked Plasmodium vivax Vaccine Candidate Antigen, Merozoite Surface Protein 1, Expressed in Escherichia coli*, **2001**, Infect. Immun., 69(9), pp. 5464-5470
- [3] Anthony L Fink, *Protein aggregation: folding aggregates, inclusion bodies and amyloid*, **1998**, Folding and Design, 3(1), pp. 9-23
- [4] Marcus Mayer, Johannes Buchner, *Refolding of Inclusion Body Proteins*, **2004**, Methods Mol Med., 94, pp. 239-54
- [5] Kouhei Tsumoto, Daisuke Ejima, Izumi Kumagai, Tsutomu Arakawa, *Practical considerations in refolding proteins from inclusion bodies*, **2003**, Protein Expression and Purification 28, pp. 1-8
- [6] Eliana De Bernardez Clark, *Refolding of recombinant proteins*, **1998**, Current Opinion in Biotechnology, 9, pp. 157-163
- [7] Seema Chakravarthi, Catherine E Jessop, Neil J Bulleid, *The role of glutathione in disulphide bond formation and endoplasmic-reticulum-generated oxidative stress*, **2006**, EMBO Rep., 7(3), pp. 271-275
- [8] Smith GK, D'Alessio G, Schaffer SW, *Glutathione-facilitated refolding of reduced, denatured bovine seminal ribonuclease: kinetics and characterization of products*, 1978, Biochemistry, 17(13), pp. 2633-2638

- [9] Sinan Ketena,b , Chia-Ching Choua , Adri C.T. van Duinc , Markus J. Buehler, *Tunable Nanomechanics Of Protein Disulfide Bonds In Redox Microenvironments*, **2012**, Journal Of The Mechanical Behaviour Of Biomedical Materials, 5, pp. 32-40
- [10] Deborah Fass, *Disulfide Bonding in Protein Biophysics*, **2012**, Annual Review of Biophysics, 41, pp. 63-79
- [11] S. F. Betz, *Disulfide bonds and the stability of globular proteins*, **1993**, Protein Sci., 2(10), pp. 1551-1558
- [12] Tao Zhang, Eric Bertelsen, Tom Alber, *Entropic effects of disulphide bonds on protein stability*, **1994**, Nature Structural Biology, 1, pp. 434-438
- [13] Bakthisaran Raman, Tangirala Ramakrishna and Ch. Mohan Rao, *Refolding of Denatured and Denatured/Reduced Lysozyme at High Concentrations*, The Journal of Biological Chemistry, 271, pp. 17067-17072
- [14] Brian M. Baynes, Daniel I. C. Wang, Bernhardt L. Trout, *Role of Arginine in the Stabilization of Proteins against Aggregation*, **2005**, Biochemistry, 44, pp. 4919-4925
- [15] Tsumoto K, Umetsu M, Kumagai I, Ejima D, Philo JS, Arakawa T, *Role of arginine in protein refolding, solubilization, and purification*, **2004**, Biotechnol Prog., 20(5), pp. 1301-1308
- [16] Arakawa T, Tsumoto K, *The effects of arginine on refolding of aggregated proteins: not facilitate refolding, but suppress aggregation*, **2003**, Biochem Biophys Res Commun., 304(1), pp. 148-52
- [17] Chen J, Liu Y, Li X, Wang Y, Ding H, Ma G, Su Z, *Cooperative effects of urea and L-arginine on protein refolding*, **2009**, Protein Expr Purif., 66(1), pp. 82-90
- [18] Thomson CA, Olson M, Jackson LM, Schrader JW, *A Simplified Method for the Efficient Refolding and Purification of Recombinant Human GMCSF*, **2012**, PLoS ONE 7(11), e49891

Chapter 7.

Discussion and Future Work

*Chapter 5 and 6 investigated the native state of PcECPs using different electrophoresis techniques and the expression, purification and refolding of PcBP. The aggregation formation of native PcECPs with refolded PcBP needs to be further characterised to further determine the cross-linking mechanism and mechanical properties of the egg capsule. In this chapter, future characterisation techniques are proposed to study the biomimetic material of cross-linked PcECPs. Glutaldehyde cross-linking had shown that the mechanical property such as stiffness is much higher than that of the natural egg capsule, but lower in elasticity. The discovery of Pugilina Bundling Protein as a hypothesized cross-linker derived from the snail's nidamental gland directly brings this project closer to a biomimetic egg capsule which closely represents its counterpart found in nature. The scleroprotein which was proposed to be responsible for the final stage of egg capsule cross-linking and hardening, was investigated in many of the elastic structural proteins such as collagen, elastin and resilin. The published studies of fluorescent scleroprotein coincides with the observation of *P. cochlidium* egg capsules under ultraviolet light, suggesting either fluorescent tyrosine derivatives or aldehydes groups may be present.*

7.1 General discussion

The roles of PcECP -1, -2, -3, -4 and -5 had not been fully investigated previously and the cross-linking of purified recombinant PcECPs were done based on gene sequence matching and conjecture. In Chapter 3, the use of electrophoresis with ethylene glycol diacrylate gel and 2D SDS-PAGE was able to precisely identify the oligomeric states of each PcECP band.

The bands of PcECP -1, -2 and -3 were found to be present in the higher oligomer state with a molecular weight of 200 kDa, whereas PcECP -4 and -5 were found to be independent from any self-assembly role in these experiments. More studies are required to determine the exact role of PcECP -4 and -5.

The expression of PcBP was done using Terrific Broth as the growth media had proven to be very efficient in cultivating an additional 360% of wet weight cells as compared to conventional Luria Bertani broth. The stable state of PcBP is in its dimeric form and due to the reactivity of reduced cysteines, not all the PcBP monomers are specifically homo-dimeric. The formation of disulphide bonding with contaminants resulted in a loss of yield and potentially affected the bioactivity of PcBP. Therefore the level of contaminants from the cell contaminants needs to be significantly reduced.

Conventional reversed phase chromatography was able to purify PcBP from cell contaminants based on the analysis through SDS-PAGE and MALDI-ToF, but not without the formation of homo-oligomers and multimers. This was probably due to the mild oxidation of PcBP and oligomerisation during sample preparation catalysed by TFA in the elution buffer. The optimisation was carried out with different acids, additives and organic solvents.

The new reversed phase HPLC binary solvent system was able to stabilise PcBP in its monomeric form. While purity is acceptable for refolding, it can still be further optimised with proper concentrations of BMIM PF₆. Another ionic liquid salt 1-Butyl-3-Methylimidazolium Chloride (BMIM Cl) with its cosmotropic anion might be a better alternative to the more hydrophobic BMIM PF₆. The charge screening of the ionic liquid might help in further stabilising the monomeric form of PcBP.

The successful refolding conditions for PcBP was seen by the aggregation of native PcECPs. Negative controls were carried out to ensure that the refolding buffer did not cause the aggregation of the proteins. An additional negative control was done with the monomeric PcBP stabilised in 6 M urea. This method was not reliable as the chaotropic reagent can denature and solubilise the PcECPs. Solubility is one of the factors to determine whether a highly hydrophobic protein such as PcBP was properly refolded, but this is not an exclusive factor. The hydrophobic core of the protein might be hidden away from the aqueous environment but it is not an indicator for proper disulphide formation.

More characterisation of PcBP is still required to further understand its role in stabilising and self-assembling of PcECPs. Proposed characterisation techniques and their principles in brief are listed in the subsequent subheadings which might be useful.

Given the unique elastomeric properties of the egg capsule, the biomimetic material might be suitable for a wide range of shock absorbing applications. Most current shock absorbers are made of synthetic rubbers and silicone compounds, however these materials are not entirely environmentally friendly and over time they lose their properties due to hardening or degradation.

The biodegradable biomimetic egg capsule material might be a better option in minimising negative environment impacts. For highly shock sensitive equipment or other applications, periodic replacement of the dampers ensure consistent performance at more affordable costs. The material might also potentially be used in the medical industry as a drug encapsulant *in-vivo*, support for skin grafting or knee cartilage and ligament replacement after an injury by exploiting the elastomeric nature of the material.

7.2 Circular Dichroism

The use of Circular Dichroism can be carried out on the refolded PcBP to determine the amount of secondary, tertiary structure and folding properties such as α -helices, β -sheets, polyproline and random coils. It can also be used to study the effects of mutation which affects the protein stability or conformation, as well as protein interactions.^{1,2}

The conformational stability of a protein under stress can be further studied, such as thermal stability, pH stability, and stability towards denaturants, and how all of these are altered by different buffer compositions. It is also an excellent technique for finding the right solvent conditions that help to increase the melting temperature and/or the reversibility of thermal unfolding.^{3,4}

This experiment will be very useful to tally with the generated 3D structure models of PcBP and compare it to the proposed secondary structure. Preliminary studies of the refolded PcBP was done but the signal to noise ratio was poor probably due to the very low concentration.

Figure 7.1. A circular dichroism spectra of different structural conformation.

Based on the 3D structural model derived from I-Tasser, QUARK and EMBL-Blast, PcBP shows predominantly α -helix structure. The expected Circular Dichroism spectra on PcBP is expected to be around 190 nm, with reference from Figure 7.1. Circular Dichroism is complementary to that of Fourier transform infrared spectroscopy.

7.3 Fourier Transform Infrared Spectroscopy

The non-destructive Fourier transform infrared spectroscopy technique is useful for the structural characterization of proteins and polypeptides. The 9 characteristic IR absorption bands (amides A, B and I–VII) are due to the repeat units in proteins. The most prominent Amide I bands between $1,700\text{--}1,600\text{ cm}^{-1}$, associated with the C=O and C-N stretching vibration are sensitive vibrational bands of the protein backbone structure and hydrogen bonding pattern. The Amide A and B band, $\sim 3500\text{ cm}^{-1}$ and $\sim 3100\text{ cm}^{-1}$ respectively, originate from a Fermi resonance between the first overtone of

Amide II and the N-H stretching vibration. Amide II results from the N-H bending vibration and from the C-N stretching vibration and is conformationally sensitive. The mixture of several coordinate displacements results in the complex bands of Amide III and IV. The out-of-plane motions are found in Amide V, VI and VIII.^{5,6,7,8}

Figure 7.2. A typical FTIR spectrum depicting Amide A region $\sim 3500\text{ cm}^{-1}$, Amide I bands between $1,700\text{--}1,600\text{ cm}^{-1}$, Amide II bands between 1510 cm^{-1} and 1580 cm^{-1} , and Amide III $\sim 1330\text{ cm}^{-1}$.

The bands of different secondary structures conformation in the sample overlap each other and cannot be distinguished in the broad amide envelope. Fourier self-deconvolution is therefore applied to separate the individual peaks that make up this envelope. A curve fitting procedure is subsequently applied to quantitatively estimate the area under the curve of each component representing the respective type of secondary structure.⁹

Figure 7.3. A typical Amide I spectrum envelope consisting of α -helix, β -sheet and random coil structure. Deconvolution is applied to quantitatively estimate the amount of each secondary structure component.

7.4 Rheology

Rheology is the study on how matter flows and responds in a liquid, semi-solid and solid state under an applied force. It is a combination of viscous, elastic and plastic behaviour.¹⁰

Rheology measurements were done on the native polymerised crude extract in the form of a hydrated soft-gel but the results were inclusive. The soft-gel was fully hydrated in the native solubilisation buffer, but the liquid content evaporated in the midst of rheology experiments. An attempt was made to rehydrate the sample but the deviation in the results were significant.

As there had been success in the refolding of PcBP and aggregation with native PcECPs, the reproduction of aggregates in a large scale experiment will be subjected to rheology measurements. Properties of the intermediate cross-linking stage of the egg capsule such as apparent yield stress, high shear rate viscosity and interfacial viscoelasticity can be derived.

7.5 Surface Characterisation

Surface characterisation of the aggregates with scanning electron microscope, transmission electron microscope and atomic force microscope are useful techniques to investigate the alignment of fibres in the aggregates and their banding periodicity, if any. The egg capsule was found to have a banding periodicity of 103 nm using scanning electron microscope and was suggested to be intercalated with its neighbouring fibres with the aid of atomic force microscope.

The dispersed and rod-like shape of the aggregates might have a different nano-structure from that of the egg capsule. The formation of amyloid fibres can be further studied as to how intercalation between non-specific fibres interact with each other. The use of transmission electron microscopy can be used to study the mechanism of self-assembly from nano-fibrils to fibres.

7.6 Surface Sclerotisation

Surface sclerotisation of the egg capsule is still not fully understood. The ventral pedal gland of the snail is believed to be the place where scleroproteins are secreted and coated around the semi-matured egg capsule. Through muscular contractions, the proteins were kneaded together to form their distinctive shape.

The egg capsule and operculum of *Buccinum undatum* was one of the first marine snail model when the study of the scleroprotein properties was carried out. It was believed that scleroproteins were responsible for the inter-chain bridges that confer stability on the egg capsule structural protein, and making it generally insoluble.¹¹

One of the characteristic of some structural proteins is the emission of fluorescence under ultraviolet light. La Bella¹² demonstrated that elastin had this fluorescence property and had described a non-extractable yellow fluorophore bound tightly to the protein. Thornhill¹³ suggested that the intrinsic fluorescence of fluorophore and chromophore in elastin may be part of the cross-linking mechanism.

Resilin which is a rubber-like protein from insects, contains both di-tyrosine and tri-tyrosine derivatives which fluoresces strongly. Andersen proposed that the fluorescence of the protein and the cross-linking are connected to these tyrosine derivatives.^{14,15,16,17} The fluorescent properties of collagen was only partially attributed to tyrosine.^{18,19}

Non-tyrosine derivatives responsible for fluorescence of the elastic protein abductin were found in some molluscs internal hinge ligament which also exhibits a fluorescent characteristic.²⁰ Instead, halogenated amino acids especially that of tyrosine, was discovered in the skeletal structure of marine demosponge *Aplysina cavernicola* and *Ianthella basta*.²¹

Like elastin and other elastic proteins mentioned, the egg cases of *P. cochlidium* were both yellow in appearance and exhibited fluorescence under ultraviolet light, suggesting either fluorescent tyrosine derivatives or fluorescent aldehydes groups may be present. Therefore with these two leads, the study into the final stage of egg capsule cross-linking is worth investigating.

References

- [1] Sharon M. Kelly, Thomas J. Jess, Nicholas C. Price, *How to study proteins by circular dichroism*, **2005**, Biochimica et Biophysica Acta, 1751, pp. 119-139
- [2] Norma J. Greenfield, *Using circular dichroism spectra to estimate protein secondary structure*, **2009**, Nat Protoc., 1(6), pp. 2876-2890
- [3] Norma J. Greenfield, *Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions*, **2009**, Nat Protoc., 1(6), pp. 2527-2535
- [4] Sangeeta Benjwal, Shikha Verma, Klaus-Heinrich Röhme, Olga Gursky, *Monitoring protein aggregation during thermal unfolding in circular dichroism experiments*, **2006**, Protein Sci., 15(3), pp. 635-639
- [5] Huayan Yang, Shouning Yang, Jilie Kong, Aichun Dong, Shaoning Yu, *Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy*, **2015**, Nature Protocols, 10, pp. 382-396
- [6] Jilie Kong and Shaoning Yu, *Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures*, **2007**, Acta Biochimica et Biophysica Sinica, 39(8), pp. 549-559
- [7] Andreas Barth, *Infrared spectroscopy of proteins*, **2007**, Biochimica et Biophysica Acta, 1767(9), pp. 1073-1101
- [8] Witold K. Surewicz, Henry H. Mantsch, Dennis Chapman, *Determination of protein secondary structure by Fourier transform infrared spectroscopy: A critical assessment*, **1993**, Biochemistry, 32(2), pp. 389-394
- [9] Payne KJ, Veis A, *Fourier transform IR spectroscopy of collagen and gelatin solutions: deconvolution of the amide I band for conformational studies*, **1988**, Biopolymers, 27(11), pp. 1749-1760

- [10] Daniel T.N. Chen, Qi Wen, Paul A. Janmey, John C. Crocker, Arjun G. Yodh, *Rheology of Soft Materials*, **2010**, Annu. Rev. Condens. Matter Phys., 1, pp. 301-322
- [11] N. R. Price and S. Hunt, *Fluorescent Chromophore Components From The Egg Capsules Of The Gastropod Mollusc Buccinum Undatum (L.), And Their Relation To Fluorescent Compounds In Other Structural Proteins*, **1974**, Comp. Biochem. Physiol., 47, pp. 601-616
- [12] La Bella F. S., *Elastin--A metabolically active lipoprotein*, **1957**, Nature, 180, pp. 1360-1361
- [13] Thornhill D., *Elastin. Locus and characteristics of chromophore and fluorophore*. **1972**, Conn. Tissue Res., 1, pp. 21-30
- [14] Andersen S. O., *Cross links in resilin*, **1963**, Biochem. Biophys., 69, pp. 249-262
- [15] Andersen S. O., *Cross links in resilin identified as di and tri tyrosine*, **1964**, Biochem. Biophys., 93, pp. 213-215
- [16] Andersen S. O., *Covalent cross links in a structural protein-resilin*, **1966**, Acta physiol., 66, pp. 1-81
- [17] Andersen S. O., *Resilin*, **1971a**, In Comprehensive Biochemistry, 26C, pp. 633-656
- [18] Hoerman K. C., Balekjien A. Y., *On the probability that the cross link of collagen is its prime luminescence centre*, **1966a**, J. Dent. Res. (Suppl). I.A.D.R., 86, pp. 59
- [19] Deyl Z., Praus R., Sulcova H., Goldman J. N., *Fluorescence of collagen-properties of tyrosine residues and another fluorescent element in calf skin collagen*, **1969**, FEBS Letters, 5, pp. 187-191
- [20] Thornhill D., *Abductin. Locus and spectral characteristics of a brown fluorescent chromophore*, **1971**, Biochemistry, 10, pp. 2644-2649

- [21] Susanne Ueberlein, Susanne Machill, Hendrik Niemann, Peter Proksch, Eike Brunner, *The Skeletal Amino Acid Composition of the Marine Demosponge Aplysina cavernicola*, **2014**, Mar Drugs, 12(8), pp. 4417–4438