

OPCML-GPI MIMETICS AS POTENTIAL BIOLOGICAL THERAPEUTICS FOR OVARIAN CANCER

Eloise Victoria Norton Morecroft

Imperial College London

Department of Chemistry

A thesis submitted to Imperial College London in
candidature for the degree of Doctor of Philosophy

May 2019

Declaration of Originality

This thesis and all work and results is my original research unless otherwise stated. Work carried out by others or collaborators has been clearly acknowledged in the text and references.

E Morecroft May 2019

Acknowledgment of Collaboration

- Dr Mostafa Jamshidiha expressed and purified the Sortase A used in the majority of the sortagging reactions.
- Dr Markus Ritzefeld ran all PMF samples and provided raw data for analysis.
- YnTB was synthesised and purified by the Tate group.
- Dr Chiara Recchi took confocal images.
- Dr Valerie Good passaged sf9 and high5 insect cells.
- OPCML WT pLenti plasmid and OPCML WT lentivirus was made by the Gabra group.

Copyright Declaration

'The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

Firstly I would like to thank my supervisors Ed Tate, Chiara Recchi, Keith Willison and Hani Gabra for all their help and guidance, particularly Ed and Chiara.

I would like to thank all members of the Tate and Barnard group for their help and advice. In particular I would like to thank Markus Ritzefeld, Fiona Rowan and Thomas Lanyon-Hogg for teaching me the ways of proteins and how to use an AKTA machine. I also want to thank everyone in the Tate group and the Barnard group for being so supportive, friendly and kind throughout my time at Imperial and for making the lab and office a fun place to be.

I would also like to thank the Gabra group particularly Evdoxia Karali for taking me under her wing and helping teach me biology.

My thanks also go to Valerie Good who helped me learn baculovirus expression, and who passaged and grew insect cells in the BV lab.

I would like to thank Anna Barnard, Wouter Kallemijn and Anna Neodo for their proof reading chapters of this thesis. My thanks also go to my friends who answered generic questions about thesis writing including Gavin, Sophie and Ruhina.

I also want to acknowledge Caitlin, Charlie, Becky, Jon, Kwang and Theo who started in the ICB CDT the same time as me and who have been great friends throughout the PhD years.

Special thanks go to James for being an awesome human and being hugely supportive throughout my PhD! Also to my parents for being supportive, understanding and encouraging throughout it all.

Abstract

Opioid binding Protein/Cell adhesion Molecule-Like (OPCML) is a tumour suppressor gene (TSG) in epithelial ovarian cancer (EOC). EOC is one of the most common cancers affecting women and has a very poor prognosis due to the asymptomatic nature of the disease, with the added issue of extremely high recurrence. OPCML functions as a tumour suppressor but the tumour suppressor function is epigenetically inactivated in EOC. Previous studies have demonstrated OPCML could be a viable therapeutic to treat EOC patients. Studies have shown re-introduction of functional OPCML to cancer cell lines causes a reduction in growth and proliferation. Therefore, a novel biotherapeutic of synthetic OPCML could have potential as a treatment post-surgery to prevent further spreading of the cancer and recurrence.

The focus of this thesis has been to create a novel biotherapeutic with a semi-synthetic OPCML targeted to the membrane of cancer cells using a GPI anchor mimetic, and to validate this protein anchor complex as a potential drug in EOC.

The design and synthesis of a highly engineered OPCML construct, which has the functionality required to introduce complex chemical prosthetic groups to mimic the GPI anchor, was successfully expressed and purified. The construct included an HA tag for visualisation, an LPETG recognition site for sortase mediated transpeptidation (sortagging), an Fc protein sequence for purification, and a thrombin cleavage site so the Fc sequence could be removed post-purification.

The GPI anchor mimetic was synthesised by solid phase peptide synthesis (SPPS) on a biotin resin and included myristoyl groups, a side chain with a photo cross-linkable moiety, and three glycines suitable for acting as a sortagging nucleophile. The protein was consequently modified by attachment of a GPI anchor mimetic via sortagging successfully. However, purification post-sortagging was challenging.

Table of Contents

DECLARATION OF ORIGINALITY.....	2
ACKNOWLEDGMENT OF COLLABORATION	2
COPYRIGHT DECLARATION	2
ACKNOWLEDGEMENTS	3
ABSTRACT	4
TABLE OF CONTENTS	5
TABLE OF FIGURES	9
PAPERS AND PRESENTATIONS FROM THIS THESIS	14
ORAL PRESENTATIONS:	14
POSTER PRESENTATIONS:	14
PAPERS:	14
ABBREVIATIONS.....	15
CHAPTER 1	18
INTRODUCTION.....	18
1.1 CANCER	18
1.1.1 <i>Ovarian Cancer</i>	18
1.1.2 <i>Diagnostics and Therapeutics in Ovarian Cancer</i>	19
1.2 OPCML	20
1.2.1 <i>OPCML and Ovarian Cancer</i>	20
1.2.2 <i>The Secretory Pathway</i>	21
1.2.3 <i>Protein Structure</i>	23
1.2.4 <i>Biological Function</i>	24
1.3 GPI ANCHORS AND LIGATION TECHNIQUES	25
1.3.1 <i>GPI Anchor Structure and Function</i>	25
1.3.2 <i>GPI Anchor Mimetics</i>	26
1.3.3 <i>Approaches for C-terminal Modification of Proteins</i>	29
1.4 SUMMARY	33
1.5 THESIS OBJECTIVES.....	35
CHAPTER 2	37
TESTING SAR OF OPCML.....	37

2.1	DESIGN OF CONSTRUCTS FOR SAR STUDIES	37
2.2	GENERATION OF STABLY TRANSDUCED CELL LINES	38
2.2.1	<i>Lentiviral transduction</i>	38
2.2.2	<i>Characterisation to confirm SKOV3 transduction with OPCML constructs</i>	40
2.3	CONCLUSION	43
CHAPTER 3		45
THE DESIGN AND SYNTHESIS OF A HIGHLY ENGINEERED AND FUNCTIONAL OPCML CONSTRUCT		45
3.1	CONSTRUCT DESIGN AND PROTEIN EXPRESSION IN MAMMALIAN AND INSECT CELL LINES	45
3.1.1	<i>Mammalian Expression of rOPCML in HEK293 cells</i>	47
3.1.2	<i>Insect Cell Expression of OPCML using the Baculovirus Expression System</i>	49
3.2	ROPCL PURIFICATION AND CHARACTERISATION	50
3.2.1	<i>Affinity purification to isolate rOPCML from insect cell supernatant and confirmation of protein identity</i>	50
3.2.2	<i>Cleaving the affinity sequence from rOPCML in preparation for assay development and future experiments</i>	51
3.2.3	<i>Anion Exchange to remove impurities after Thrombin Cleavage</i>	56
3.2.4	<i>Size Exclusion Chromatography to confirm aggregation state of rOPCML</i>	57
3.3	GENERATING A ROPCL+Fc CONSTRUCT WITH NO FREE CYSTEINE TO PREVENT AGGREGATION OF THE PROTEIN IN FUTURE RESEARCH	60
3.4	GENERATING THE INACTIVE R65L MUTANT CONSTRUCT FOR ROPCL+Fc AS A CONTROL FOR FUNCTIONAL ASSAYS	62
3.5	TESTING THE FUNCTION OF ROPCL	65
3.6	CONCLUSION	66
CHAPTER 4		67
FUNCTIONALISING ROPCL WITH A GPI ANCHOR MIMETIC		67
4.1	INTEIN-MEDIATED PROTEIN LIGATION	67
4.2	SORTASE MEDIATED TRANSPEPTIDATION – SORTAGGING	72
4.2.1	<i>GPI Anchor Mimetic Design Considerations and Synthesis</i>	72
4.3	PURIFICATION OF THE PROTEIN ANCHOR COMPLEX POST-SORTAGGING	84
4.4	CONCLUSION	90
CHAPTER 5		91
CONCLUSIONS AND OUTLOOK		91
5.1	CONCLUSIONS	91
5.2	FUTURE WORK	92
CHAPTER 6		94

METHODS	94
6.1 GENERAL BIOLOGY METHODS	94
6.1.1 <i>Plasmids/Constructs and Cloning</i>	94
6.1.2 <i>Transformations</i>	95
6.1.3 <i>PCR</i>	95
6.2 MAMMALIAN CELL CULTURE.....	95
6.2.1 <i>Cell Lines</i>	95
6.2.2 <i>Antibodies</i>	96
6.2.3 <i>Immunofluorescence Microscopy</i>	96
6.2.4 <i>SDS-PAGE</i>	96
6.2.5 <i>Western Blot</i>	96
6.2.6 <i>Dot Blot</i>	97
6.2.7 <i>Lentivirus Production</i>	97
6.2.8 <i>Producing Stably Transduced Cell lines</i>	97
6.2.9 <i>Transfection</i>	98
6.2.10 <i>Cell Proliferation Assay</i>	98
6.2.11 <i>Spheroid Invasion Assay</i>	98
6.3 BACULOVIRUS PRODUCTION METHODS.....	99
6.3.1 <i>Cell Lines</i>	99
6.3.2 <i>Generating Recombinant Bacmid</i>	99
6.3.3 <i>Generating OPCML Baculovirus</i>	99
6.4 OPCML PURIFICATION FROM INSECT CELLS AND PROTEIN MODIFICATIONS.....	100
6.4.1 <i>Affinity Purification from Cell Supernatant</i>	100
6.4.2 <i>Anion Exchange</i>	100
6.4.3 <i>Size Exclusion Chromatography</i>	101
6.4.4 <i>Peptide Mass Fingerprint Analysis</i>	101
6.4.5 <i>ZipTip De-salting</i>	102
6.4.6 <i>Free Cysteine Capping</i>	102
6.4.7 <i>Intein-mediated Protein Ligation (IPL) Cloning</i>	102
6.4.8 <i>Bacteria Expression system for Intein mediated Protein Ligation</i>	103
6.4.9 <i>Sortagging</i>	103
6.4.10 <i>Protein Precipitation</i>	104
6.4.11 <i>CuAAC reaction on Anchor Mimetics</i>	104
6.4.12 <i>Pulldown</i>	104
6.5 SYNTHESIS METHODS	104
6.5.1 <i>General SPPS Protocols</i>	104
6.5.2 <i>Automated SPPS</i>	105
6.5.3 <i>Cleavage</i>	106

6.5.4	<i>IvDde deprotection</i>	106
6.5.5	<i>BOC Deprotection of commercial BOC-GGG-OH</i>	106
6.5.6	<i>SPPS of anchor 3</i>	106
6.5.7	<i>SPPS of anchor 4</i>	108
BIBLIOGRAPHY		110
APPENDIX A		124
APPENDIX B		127
APPENDIX C		130
APPENDIX D		144

Table of Figures

FIGURE 1 SCHEMATIC SHOWING THE SECRETORY PATHWAY OPCML UNDERTAKES TO GET TO THE PLASMAL MEMBRANE. IMAGE TAKEN FROM PREVIOUS REPORT ¹¹⁹ .	22
FIGURE 2 THE STRUCTURE OF OPCML WITH SITES OF GLYCOSYLATION (ASPARAGINE 44, 70, 140, 285, 293 AND 306) SHOWN IN YELLOW. DOMAIN 1 IS DEPICTED IN RED, DOMAIN 2 BLUE AND DOMAIN 3 GREEN. SHORT EXTENDED LINKERS BETWEEN DOMAINS ARE DEPICTED IN GREY. (A) OPCML HOMODIMER SHOWS AS A SCHEMATIC ON THE SURFACE ON THE PLASMA MEMBRANE. GPI ANCHOR ATTACHMENT OCCURS AT ASP322. (B) KEY RESIDUES FOR THE DOMAIN 1 HOMODIMER INTERFACE; ARG65, LEU69, ILE74, ASP80, TRP82 AND THR114 (CYAN FOR ONE OPCML AND MAGENTA FOR THE SECOND OPCML). (C) RESIDUES THAT UNDERGO MUTATION TO GIVE LOSS OF TUMOUR SUPPRESSOR FUNCTION R65L, N70H AND P95R.	23
FIGURE 3. SCHEMATIC REPRESENTATION OF THE BIOLOGICAL MECHANISM OF OPCML ^{31,172} .	24
FIGURE 4 (A) SCHEMATIC OF THE GENERIC STRUCTURE OF GPI ANCHORS ⁵⁶ . (B) ⁶⁹ , (C) ⁵⁵ , (D) ⁷² , (E) ⁷⁵ AND (F) ⁷⁴ GPI ANCHOR MIMETICS REPORTED IN THE LITERATURE.	27
FIGURE 5 (A) COPPER CATALYSED AZIDE ALKYNE CYCLOADDITION (CUAAC) ⁷⁸⁻⁸¹ . (B) STRAIN-PROMOTED AZIDE-ALKYNE CYCLOADDITION (SPAAC) ⁸² .	30
FIGURE 6 SCHEMES SHOWING THE MECHANISM OF NATIVE CHEMICAL LIGATION AND INTEIN-MEDIATED PROTEIN LIGATION. SCHEMES ADAPTED/REPRODUCED FROM PAPERS ^{91,100} .	31
FIGURE 7 SCHEMATIC SHOWING REACTION SCHEME FOR SORTAGGING AND THE ATTACHMENT OF THE ANCHOR TO THE PROTEIN. IMAGE ADAPTED FROM ANTOS ET AL ¹⁰⁹ .	32
FIGURE 8 SCHEMATIC REPRESENTATION OF THREE MAIN PARTS OF THIS RESEARCH.	35
FIGURE 9 SCHEMATICS OF OPCML WT AND FL, D1, D2 AND D3 CONSTRUCTS DESIGNED FOR TESTING FUNCTIONAL IMPORTANCE OF THE SIGNAL DOMAINS ¹¹⁹ . IMAGE TAKEN FROM REPORT ¹²⁰ .	37
FIGURE 10 (A) DNA GEL ELECTROPHORESIS OF OPCML FL, D1, D2, D3 AND PENTR4 DOUBLE DIGESTION TO SHOW INSERTS AND PENTR BACKBONE. (B) DNA GEL ELECTROPHORESIS OF DIGESTED FL, D1, D2 AND D3 CONSTRUCTS IN PENTR VECTOR: THE DOUBLE DIGESTION (BOTH) CONFIRMS THE PRESENCE OF THE CONSTRUCT IN THE ENTRY VECTOR. (A) AND (B) FL IS SEEN AROUND 1000 BP AND D1, D2 AND D3 AROUND 500 BP. DIGESTION WAS DONE USING XBAI AND SALI. U = UNDIGESTED, D = DIGESTED. IMAGES TAKEN FROM REPORT ¹²⁰ .	39
FIGURE 11 (A) SDS- PAGE ANALYSIS OF TRANSDUCE SKOV3 CELLS TO CONFIRM EXPRESSION. IMAGE ADAPTED FROM REPORT ¹²⁰ . (B) CONFOCAL IMMUNOFLUORESCENCE MICROSCOPY SHOWING THE EXPRESSION AND LOCATION OF PROTEINS IN TRANSDUCE SKOV3 CELLS: SKOV3 CELLS TRANSDUCE	

WITH FL, D1, D2, D3 OR OPCML WT CONTROLS WERE FIXED USING PFA (SEE METHODS) AND STAINED WITH RAT ANTI-HA (GREEN), MOUSE ANTI-OPCML (RED) AND DAPI (BLUE). SCALE BARS 20 μ M.....	41
FIGURE 12 MTS CELL GROWTH ASSAY OF LENTIVIRALLY TRANSDUCED SKOV3 CELLS. RESULTS WERE TAKEN IN TRIPLICATE.	43
FIGURE 13 SCHEMATIC SHOWING THE STAGES IN PRODUCING A BACULOVIRUS EXPRESSION SYSTEM ¹³⁴ . IMAGE TAKEN FROM PREVIOUS REPORT ¹²⁰	46
FIGURE 14(A) SCHEMATIC REPRESENTATION OF THE OPCML CONSTRUCT DESIGNED FOR A MAMMALIAN EXPRESSION SYSTEM.(B) AND (C) SDS-PAGE SHOWING EXPRESSION LEVEL OF OPCML IN TRANSFECTED HEK293 CELLS: HEK293 CELLS WERE TRANSFECTED WITH OPCML (60 kDA) IN A PCDNA3.1+ VECTOR OR A CONTROL EMPTY PCDNA3.1+ VECTOR. COMMERCIAL ROPCML (PURCHASED FROM CREATIVE BIOMART) WAS USED AS A CONTROL. (B) WESTERN BLOT (C) COOMASSIE STAINING.	48
FIGURE 15(A) SCHEMATIC REPRESENTATION OF THE OPCML CONSTRUCT DESIGNED FOR AN INSECT EXPRESSION SYSTEM. (B) AGAROSE GEL ELECTROPHORESIS OF RECOMBINANT BACMID. IMAGE ADAPTED FROM REPORT ¹²⁰ . (C) WESTERN BLOT SHOWING EXPRESSION LEVEL OF OPCML IN HIGH 5 CELL SUPERNATANT OVER TIME: HIGH 5 CELLS WERE INFECTED WITH OPCML BACULOVIRUS AND ALIQUOTS TAKEN EVERY 24 HOURS.	50
FIGURE 16 SDS-PAGE SHOWING STAGES OF ROPCML PROTEIN PRODUCTION.	51
FIGURE 17 SDS-PAGE OF ON-RESIN THROMBIN CLEAVAGE OVER TIME. (A) WESTERN BLOT. (B) COOMASSIE STAINED GEL.	52
FIGURE 18 SDS-PAGE SHOWING STAGES OF PURIFICATION AFTER THROMBIN CLEAVAGE.	53
FIGURE 19 THROMBIN CLEAVAGE OPTIMISATION USING THROMBIN CLEANCLEAVE KIT. (A) OVERNIGHT INCUBATION AT 4 °C WITH THROMBIN RESIN. (B) AND (C) THROMBIN CLEAVAGE USING THE RESIN BOUND THROMBIN AT 25 °C AND 37 °C RESPECTIVELY.	54
FIGURE 20 THROMBIN CLEAVAGE OPTIMISATION AT 37 °C WITH IMMOBILISED THROMBIN ON AGAROSE RESIN, 'LIQUID' THROMBIN AND A ROPCML+Fc CONTROL REACTION OF JUST BUFFER AND PROTEIN SHAKING AT 37 °C.	55
FIGURE 21 ANION EXCHANGE OF ROPCML AFTER THROMBIN CLEAVAGE TO REMOVE IMPURITIES. (A) TABLE SHOWING THE DIFFERING PI VALUES FOR BIOMOLECULES IN THE THROMBIN CLEAVAGE REACTION. (B) UV TRACE FROM ANION EXCHANGE COLUMN. (C) SDS-PAGE STAINED WITH COOMASSIE SHOWING THE IDENTITY OF THE PEAKS ELUTED FROM THE IEX COLUMN. BC BEFORE THROMBIN CLEAVAGE. AC AFTER THROMBIN CLEAVAGE.	57
FIGURE 22 SEC TRACES OF ROPCML (A) S75 16/600 SEC. (B) S200 10/300 SEC. (C) S200 16/600 SEC. (D) SDS-PAGE OF S200 16/600 SEC OF ROPCML.....	58

FIGURE 23 SDS-PAGE OF ROPCML WITH AND WITHOUT REDUCING AGENT. (A) SDS-PAGE STAINED WITH COOMASSIE. (B) WESTERN BLOT.....	59
FIGURE 24 SEC S200 10/300 OF ROPCML AFTER REDUCING AND CAPPING.	60
FIGURE 25 (A) AGAROSE GEL ELECTROPHORESIS OF PCR PRODUCT OF SAMPLE AND CONTROL AFTER AMPLIFYING WITH PRIMERS TO REMOVE FREE CYSTEINE. (B) AGAROSE GEL ELECTROPHORESIS OF RECOMBINANT BACMID. (C) DOT BLOT TO SHOW EXPRESSION OF ROPCML-C295DEL IN INSECT CELLS AFTER 72 HOURS.....	61
FIGURE 26 SDS-PAGE STAINED WITH COOMASSIE SHOWING STAGES OF OPCML PROTEIN PURIFICATION.....	61
FIGURE 27 (A) SDS-PAGE OF ANION EXCHANGE ANX FF COLUMN STAINED WITH COOMASSIE, AC = AFTER THROMBIN CLEAVAGE. (B) UV, CONDUCTIVITY AND CONCENTRATION TRACE OF ANION EXCHANGE. (C) SEC S200 10/300 ROPCML-C295DEL.	62
FIGURE 28 (A) AGAROSE GEL ELECTROPHORESIS OF PCR PRODUCT OF SAMPLE AND CONTROL AFTER AMPLIFYING WITH PRIMERS TO REMOVE FREE CYSTEINE. (B) AGAROSE GEL ELECTROPHORESIS TO ANALYSE PRODUCTION OF RECOMBINANT BACMID.	63
FIGURE 29 SDS-PAGE STAINED WITH COOMASSIE SHOWING STAGES OF R56L MUTANT PROTEIN PURIFICATION.	63
FIGURE 30 (A) SDS-PAGE OF ANION EXCHANGE ANX FF COLUMN STAINED WITH COOMASSIE, AC = AFTER THROMBIN CLEAVAGE. (B) UV TRACE OF ANION EXCHANGE. (C) SEC S200 10/300 ROPCML-R56L..	64
FIGURE 31 (A) SCHEMATIC SHOWING STAGES OF THE SPHEROID INVASION ASSAY. (B) IMAGES OF U87MG CELLS BEFORE AND AFTER TREATMENT WITH ROPCML.	66
FIGURE 32 SCHEMES FOR THE LIGATION OF RECOMBINANT OPCML PROTEIN TO THE GPI ANCHOR MIMETIC. PART 1) P TXB1-OPCML IS EXPRESSED WITH AN INTEIN CONTAINING A CHITIN BINDING DOMAIN (CBD), FIRST THE CBD BINDS TO A CHITIN COLUMN, AND AN S N SHIFT WILL OCCUR. MESNA IS USED TO CLEAVE THE PROTEIN FROM THE COLUMN AND LEAVES THE OPCML PROTEIN WITH REACTIVE THIOESTER ⁹¹ . PART 2) THE C-TERMINAL THIOESTER OF OPCML WILL UNDERGO TRANSESTERIFICATION WITH A TERMINAL CYSTEINE RESIDUE IN THE GPI ANCHOR MIMETIC. AN S N ACYL SHIFT THEN OCCURS LEADING TO A PROTEIN-GPI ANCHOR COMPLEX ⁹¹ . IMAGE TAKEN FROM INITIAL RESEARCH PLAN ¹⁴⁶	68
FIGURE 33 AGAROSE GEL ELECTROPHORESIS CONFIRMING SUCCESSFUL GENERATION OF INTEIN DNA CONSTRUCTS U= UNDIGESTED VECTOR, D=DIGESTED VECTOR WITH NDEI AND SAPI. (A) PCR WITH PRIMERS TO INTRODUCE RESTRICTION ENZYME SITES. (B) GEL INDICATING PTXB1 FULLY DIGESTED AND FL AND D3 INSERTS CONTAIN NO IMPURITIES PRIOR TO LIGATION. (C) LIGATION OF FL AND D3 INTO PTXB1 WITH THE DIGESTED FL AND D3 IN PTXB1 SHOWING A SMALLER BAND AT 1000 BP AND 400 BP RESPECTIVELY. IMAGE TAKEN FROM REPORT ¹²⁰	69

FIGURE 34 SDS-PAGE TO CONFIRM OPTIMUM EXPRESSION OF D3 AND FL PTXB1 CONSTRUCTS. (A) D3 COOMASSIE AND WESTERN BLOT. (B) FL COOMASSIE AND WESTERN BLOT.	70
FIGURE 35 (A) PTXB1-FL AND PTXB1-D3 BOTH HAVE OPCML PROTEIN PRESENT IN SOLUBLE (S) AND INSOLUBLE (IN) FRACTIONS (B) STAGES OF PURIFICATION OF THE PTXB1 PROTEIN CONSTRUCTS. IMAGE TAKEN FROM REPORT ¹²⁰	71
FIGURE 36 (A) SCHEMATIC OF CONSTRUCT DESIGNED FOR SORTAGGING ³² . (B) SCHEMATIC SHOWING PROCESS OF SORTAGGING ¹⁷³ . (C) SCHEMATIC DEMONSTRATING HOW OPCML WILL BE ANCHORED TO THE CELL MEMBRANE ⁷²	72
FIGURE 37 GPI ANCHOR MIMETIC DESIGNS. RED INDICATES GROUP FOR ANCHORING WITHIN CELL MEMBRANE. GREEN ARE GROUPS THAT CAN BE USED FOR CUAAC. THE BLUE MOIETY IS BIOTIN WHICH CAN BE USED FOR ENRICHMENT AND VISUALISATION. THE PINK BOX OUTLINES THE HYDROPHILIC CYTOTOPIC PEPTIDE WHICH WILL ASSOCIATE WITH THE CELL MEMBRANE. THE ORANGE GROUP IS A DIAZARINE PHOTO REACTIVE CROSSLINKER.	73
FIGURE 38 REACTION SCHEME TO FORM COMPOUND 1.	73
FIGURE 39 REACTION SCHEME TO FORM ANCHOR 1.	74
FIGURE 40 REACTION SCHEME TO FORM CHOLESTEROL ANCHOR MIMETIC, COMPOUND 5.	75
FIGURE 41 SCHEME TO SHOW STEPS IN SYNTHESIS OF ANCHOR 4.	77
FIGURE 42 FULL STRUCTURE OF ANCHOR 4.	77
FIGURE 43 (A) SDS -PAGE TO CONFIRM THE SUCCESSFUL ANCHORING OF THE PROTEIN VIA FLUORESCENCE IMAGING OF THE WET SLAB GEL AND COOMASSIE. (B) SCHEMATIC OF CONSTRUCT DESIGNED FOR SORTAGGING WITH SORTASE REACTION CLEAVAGE SITE INDICATED.	78
FIGURE 44 SDS -PAGE OF OPTIMISING SORTAGGING REACTION. (A) OPCML+Fc. (B) OPCML. WITH NO Fc. FOR BOTH (A) AND (B) AS IN FIGURE 10; 1=OPCML + ANCHOR, 2=OPCML + SORTASE, 3=OPCML, 4=OPCML + ANCHOR + SORTASE, 4= SORTASE + ANCHOR.	80
FIGURE 45 (A) SDS-PAGE OF PRECIPITATION AND CLICK OF ROPCML AFTER SORTAGGING. Fc = ROPCML+Fc, C = ROPCML CAPPED WITH IODOACETAMIDE, UC = ROPCML UNCAPPED. (B) SDS-PAGE OF PULLDOWN WITH MAGNETIC STREPTAVIDIN BEADS POST-SORTAGGING.	81
FIGURE 46 SDS-PAGE OF ROPCML+Fc POST-SORTAGGING TO TEST SOLUBILITY OF ANCHORED PROTEIN. SAMPLES WERE TAKEN AFTER SORTAGGING PRIOR TO THE CLICK REACTION WITH YNTB AND AFTER THE CLICK REACTION. SR = SORTASE REACTION, CR = CONTROLS REACTION WITH NO SORTASE PRESENT.	83
FIGURE 47 (A) ANION EXCHANGE AFTER SORTAGGING. (B) TABLE OF PI VALUES. (C) UV, CONDUCTIVITY AND CONCENTRATION TRACE OF ANION EXCHANGE.	85

FIGURE 48 SDS-PAGE OF POST-SORTAGGING PURIFICATION ATTEMPTS. (A) PURIFICATION USING PROTEIN A AGAROSE AND ANTI-HA MAGNETIC BEADS. (B) PURIFICATION USING PROTEIN A AGAROSE AND SPIN CONCENTRATORS.....	86
FIGURE 49 (A) AND (B) SDS-PAGE OF POST-SORTAGGING PURIFICATION ATTEMPT WITH MONOMERIC AVIDIN AGAROSE.	87
FIGURE 50 SDS- PAGE OF POST-SORTAGGING PURIFICATION ATTEMPT WITH MONOMERIC AVIDIN AGAROSE.....	88
FIGURE 51 (A) SDS-PAGE OF POST-SORTAGGING PURIFICATION ATTEMPT WITH (A) MONOMERIC AVIDIN AGAROSE (B) PROTEIN A AGAROSE.....	89
FIGURE 52 SDS-PAGE OF POST-SORTAGGING PURIFICATION ATTEMPT WITH MONOMERIC AVIDIN AGAROSE AND PROTEIN A AGAROSE.	90
FIGURE 53 ANCHOR 3 LC TRACE AND MASS SPEC	107
FIGURE 54 ANCHOR 4 LC TRACE AND MASS SPEC	109

Papers and Presentations from this Thesis

Oral Presentations:

E. Morecroft et al. 'OPCML-GPI Mimetics as Potential Biological Therapeutics for Ovarian Cancer' Imperial College London, Chemistry Department Postgraduate Symposium, London, UK, July 2018.

Poster Presentations:

E. Morecroft et al. 'OPCML-GPI Mimetics as Potential Biological Therapeutics for Ovarian Cancer', Joint Doctoral Training Centre Conference, Warwick, UK, June 2016.

E. Morecroft et al. 'OPCML-GPI Mimetics as Potential Biological Therapeutics for Ovarian Cancer', Precision Medicine Conference, London, UK, July 2016.

E. Morecroft et al. 'OPCML-GPI Mimetics as Potential Biological Therapeutics for Ovarian Cancer' Imperial College London, Chemistry Department Postgraduate Symposium, London, UK, July 2017.

Papers:

Birtley JR, Alomary M, Zanini E, Antony E, Maben Z, Weaver GC, Von Arx C, Mura M, Marinho AT, Lu H, Morecroft E, Karali E, Chayen NE, Tate E, Jurewicz M, Stern LJ, Recchi C, Gabra H. Clinical inactivating mutations of the tumour suppressor OPCML reveal new cancer-associated functions. Nat Comm, in press.

Abbreviations

AMBIC	Ammonium bicarbonate
APS	Ammonium persulfate
Boc	Tert-butyloxycarbonyl
bp	Base pair
BSA	Bovine serum albumin
BV	Bed volume
CBD	Chitin binding domain
CHCA	α -Cyano-4-hydroxycinnamic acid
CO-IP	Co-Immunoprecipitation
CYT	Cytotopic Anchor
D1	Domain 1
D2	Domain 2
D3	Domain 3
DAPI	4',6-diamidino-2-phenylindole
DCM	Dichloromethane
DIC	N,N'-Diisopropylcarbodiimide
DIPEA	N,N-Diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
DRM	Detergent resistant membrane
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
EOC	Epithelial ovarian cancer
FA	Formic acid
FAM	5(6)-Carboxyfluorescein
FL	Full length
Fmoc	Fluorenylmethyloxycarbonyl
FT	Flow through
GPI	Glycophosphatidylinositol
GPI-AP	Glycophosphatidylinositol-anchored protein
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate

HBTU	N,N,N',N'-Tetramethyl-O-(1H-benzotriazol-1-yl)uranium hexafluorophosphate
HOBt	Hydroxybenzotriazole
HRP	Horseradish peroxidase
IEX	Ion exchange chromatography
IFM	Immunofluorescence microscopy
IP	Immunoprecipitation
IPL	Intein mediated protein ligation
IPTG	Isopropyl- β -D-thiogalactoside
ivDde	1-(4,4-Dimethyl-2,6-dioxocyclohex-1-ylidene)-3-methylbutyl
LB	Lysogeny broth
LC-MS	Liquid chromatography – mass spectrometry
MALDI	Matrix Assisted Laser Desorption/Ionization
MESNA	2-mercaptoethanesulfonic acid
Myr	Myristoyl
NMM	N-methylmorpholine
NMP	N-methylpyrrolidine
OPCML	Opioid binding Protein/Cell adhesion Molecule-Like
PAGE	Polyacrylamide gel electrophoresis
Pbf	2,2,4,6,7-Pentamethyl-2,2-dihydrobenzofuran-5-sulfonyl
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
pI	Isoelectric point
PMSF	Phenylmethanesulfonyl fluoride
rOPCML	Recombinant OPCML
rOPCML+Fc	Recombinant OPCML with Fc portion still attached
rOPCML-CYS	Recombinant OPCML with no free cysteine
RT	Room temperature
RTK	Receptor tyrosine kinase
SAR	Structure activity relationship
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SPPS	Solid Phase peptide synthesis
TBTA	Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine

TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TFA	Trifluoroacetic acid
TIS	Triisopropylsilane
Tris	Tris(hydroxymethyl)aminomethane
TSG	Tumour suppressor gene
W	Wash
WB	Western blot
X-gal	5-Bromo-4-chloro-3-indolyl β -D-galactopyranoside
YnTB	Tamra and biotin containing capture reagent ¹

Chapter 1

Introduction

The focus of this thesis is on the generation of a novel biotherapeutic for ovarian cancer, based on a protein-anchor mimetic. Therefore the following chapter introduces background information including ovarian cancer, the role of the tumours suppressor Opioid binding Protein Cell adhesion Molecule-Like (OPCML), glycoposphatidylinositol (GPI) anchors and protein ligation techniques that are relevant to this research.

1.1 Cancer

Cancer is a significant worldwide health concern; it is the second most common cause of death in the US². Mutations causing genetic and epigenetic alterations can drive cancer leading to uncontrolled proliferation in defiance of normal controls and migration of cells with invasion into surrounding tissues the cells normally would not inhabit. Epigenetic changes are those that do not alter the nucleotide sequence of deoxyribonucleic acid (DNA), such as DNA methylation or histone modifications^{3,4}. These mutations can be due to errors in cell replication, exposure to radiation or carcinogens, hereditary dispositions or viral infections³. Cancer inducing mutations can be activating or inactivating, and affect the signalling pathways for example by inactivating tumour suppressors³.

Tumour suppressor genes (TSG) are genes that when inactivated by loss-of-functions mutations can lead to cancer³. When a TSG undergoes an inactivating mutation uncontrolled cell proliferation, reduced growth factor requirements and genetic instability leading to further mutations can occur. OPCML is one such TSG that is inactivated in epithelial ovarian cancer (EOC)³⁻⁵.

1.1.1 Ovarian Cancer

Ovarian cancer is the 6th most common cancer to affect females in the UK with 80-90 % of these cases being epithelial ovarian cancer (EOC)^{6,7}. EOC is caused by genetic modifications in ovarian surface epithelium (OSE)^{5,8}. The OSE is a layer of flat-to-cuboidal epithelial cells that covers the ovary^{8,9}.

The molecular basis of ovarian cancer is not fully understood but often involves the p53 TSG that is mutated in many cancers¹⁰, BRCA1 and 2 which are breast cancer susceptibility genes¹¹,

the phosphoinositide 3-kinase pathway (an intracellular signalling pathway)¹², growth-factor signalling pathways that stimulate cells to grow, survive or proliferate^{13–15}, and angiogenesis (the growth of new blood vessels sprouting from existing blood vessels)¹⁶.

If EOC is diagnosed in its early stage there is a 90 % chance of survival for 5 or more years⁶. However, if diagnosis occurs in the late stage of the disease that chance of survival falls to just 5 %⁶. Due to the asymptomatic nature of EOC, diagnosis in the early stages is difficult and 60 % of ovarian cancer cases are diagnosed at late stage in England^{6,17}. Late stage diagnosis results in a poor prognosis for the patient due to dissemination of the cancer within the peritoneal cavity¹⁷. Lengthy surgical procedures to try and remove all the cancerous tissue within the peritoneal cavity are undertaken post-diagnosis but it has proven to be difficult to remove all cancerous tissue and therefore the rate of recurrence is high^{17,18}. This high relapse rate causes negative psychological and physical impacts on patients and greatly reduces their quality of life¹⁸.

1.1.2 Diagnostics and Therapeutics in Ovarian Cancer

The current methods of diagnosing a patient with EOC relies on physical examination findings and radiographic imaging¹⁹. It is also possible to screen for EOC; CA125 is a widely used tumour marker for EOC, it is large glycoprotein that is overexpressed in 80 % of EOC and secreted into the bloodstream^{17,20,21}. However, CA125 has low sensitivity for the early stage of the disease and there are elevated CA125 levels in benign states^{17,20}.

Ovarian cancer tumours are heterogeneous. Different histological types of EOC have different prognosis, and there are also distinct molecular subtypes²². Current treatment for EOC consists of surgical debulking to remove the cancerous tissue, this is followed with cytotoxic platinum and taxane chemotherapy^{19,23}. In advanced stages of EOC complete cytoreduction is often not possible. These patients may be treated with chemotherapy initially and if they respond to this treatment surgical debulking is then performed²². After initial therapy if EOC recurs within 6 months of treatment it is known as platinum resistant, and if it recurs after 6 months it is considered platinum sensitive²³.

New treatments are being developed and investigated such as anti-angiogenic factors^{24,25}, inhibitors of growth factor signalling^{24,25}, polyADP-ribose polymerase (PARP) inhibitors^{26,27}, and folate receptor inhibitors²⁸. As of yet these therapies haven't provided a cure but might aid in the delay of recurrence or stabilise the disease²².

Introduction

Recently approved drugs for ovarian cancer treatment include bevacizumab and PARP inhibitors²². Inhibiting angiogenesis is another therapeutic target in EOC. An increase in vascular endothelial growth factor (VEGF) expression has been linked to ovarian cancer¹⁴. Bevacizumab is a humanized anti-vascular endothelial growth factor (VEGF) monoclonal antibody that binds to the VEGF-receptor ligand blocking VEGF and preventing growth of new blood vessels¹⁶. Clinical trials have looked at adding bevacizumab to the existing carboplatin and taxane chemotherapy but findings led to the treatment not being used routinely as a first-line therapy^{23–25,29}.

Olaparib and niraparib are PARP inhibitors that have been approved for the treatment of ovarian cancer^{22,26,30}. PARP has a role in the repair of DNA breaks and BRCA1 and 2 are tumour suppressor genes which also have a role in DNA repair and have been found to be mutated in some cases of ovarian³⁰. Treating with a PARP inhibitor kills cells with BRCA1 or BRCA2 mutations due to the cells being unable to repair DNA breaks through PARP or BRCA1/2 and is a standard treatment for women with BRCA-mutated ovarian cancer²².

1.2 OPCML

1.2.1 *OPCML and Ovarian Cancer*

OPCML is a tumour suppressor gene (TSG) in EOC^{5,31}. Opioid binding Protein Cell adhesion Molecule-Like (OPCML) is a member of the IgLON family of cell adhesion molecules. The IgLON family is an Ig (immunoglobulin) domain-containing subfamily of glycosphosphatidylinositol (GPI)-anchored cell adhesion molecules^{5,8}. OPCML has three IgG type immunoglobulin (IgG) domains and six sites of N-linked glycosylation. OPCML is also a glycosphosphatidylinositol anchored protein (GPI-AP) with a GPI anchor attaching at Asp 322 in the protein^{32,33}.

It has been reported that OPCML is expressed in normal OSE but expression is lost in 83 % of the primary ovarian tumours and 88 % of the ovarian cancer cell lines that were tested⁵. OPCML gene silencing is due to epigenetic inactivation by CpG island methylation⁵. CpG islands are regulatory and promotor regions in a DNA sequence where cytosine is next to guanine³⁴. OPCML has been found to be inactivated via methylation in numerous cancers including EOC^{5,31,35–44}. Loss of heterozygosity (LOH) also occurs in OPCML at chromosome 11q25 due to chromosomal deletions resulting in the loss of an allele thought to drive inactivation of the tumour suppressor gene⁴⁵.

Methylation of the CpG island by DNA methyltransferases (DNMT) causes inactivation of OPCML with 83 % of ovarian cancer tumours and 82 % of ovarian cancer cell lines tested having

methylation of CpG islands⁵. OPCML silencing is due to promotor methylation^{31,36}. A promotor is a region of DNA responsible for initiating the transcription of the gene of interest upstream and containing the regulatory sequences. Studies have demonstrated that treating cancer cells in which OPCML expression is silenced by methylation with the DNMT inhibitory 5-aza-2'-deoxycytidine (5-AZA), results in the re-expression of OPCML^{31,36}. Thus treating patients with a DNMT inhibitor could prevent methylation and subsequent silencing of OPCML but there are concerns over the selectivity of such inhibitors and their effect on healthy cells^{46,47}.

The reintroduction of functional OPCML is another approach to treat cancer induced by OPCML silencing. Reintroduction of functional OPCML has been shown to have potential as a therapeutic. Studies have indicated that cell proliferation is reduced in cells expressing OPCML when compared to cells without OPCML expression³². Furthermore, bacterially expressed OPCML without the signal peptide or GPI anchor site caused growth inhibition of EOC cell lines³². An *in vivo* test with mice injected with ovarian cancer cell lines and a week later treated with rOPCML was also carried out. The protein treated mice remained healthy whilst a control group had to be terminated early due to the extent of tumour growth. In this research the expressed rOPCML was present in inclusion bodies and thus solubilised prior to refolding suggesting problems with solubility when OPCML is expressed bacterially. Furthermore, when this research was published the crystal structure of OPCML was unknown and an N-terminal his₆-thioredoxin tag was included in the rOPCML construct. However, the crystal structure was subsequently elucidated (see section 1.2.4) and this N-terminal tag would most likely cause disruption of the protein homodimerisation (see section 1.2.4)³³. Therefore, although this research showed promising results, future research using rOPCML without N-terminal tags and with a GPI anchor present might provide a construct more targeted to the cell membrane with greater activity as a tumour suppressor.

1.2.2 The Secretory Pathway

The secretory pathway is the way in which proteins get transported to the cell membrane. The protein is synthesised on an ER membrane bound ribosome. Translocation into the ER occurs during synthesis due to the presence of a signal sequence which is a part of the protein that directs it to its required location. The addition of the glycoposphatidylinositol anchor occurs in the ER lumen (a cavity enclosed by the ER) as a posttranslational modification⁴⁸. Posttranslational modifications are enzyme catalysed changes to a protein after it is synthesised such as glycosylation and lipidation that can regulate the degradation, localisation, and functional interactions of proteins^{3,49}.

Introduction

GPI anchors are lipids that anchor proteins at their C-terminus to the cell membrane⁴⁸ (discussed in more detail in section 1.3.1). GPI-anchored proteins (GPI-APs) serve many functions within the cell such as signalling and cell adhesion^{50,51}.

Once in the ER there are preassembled GPI anchor intermediates that are added to the protein after it is cleaved from its propeptide at the C-terminus of the protein. The protein remains membrane bound and is transported to the Golgi apparatus in a transport vesicle. The Golgi apparatus is an organelle in the cells where proteins are stored and modified before being transported to the plasma membrane of the cell. Further post-translational modifications occur in the Golgi apparatus before the protein is packaged into a secretory vesicle and dispatched to the cell membrane³.

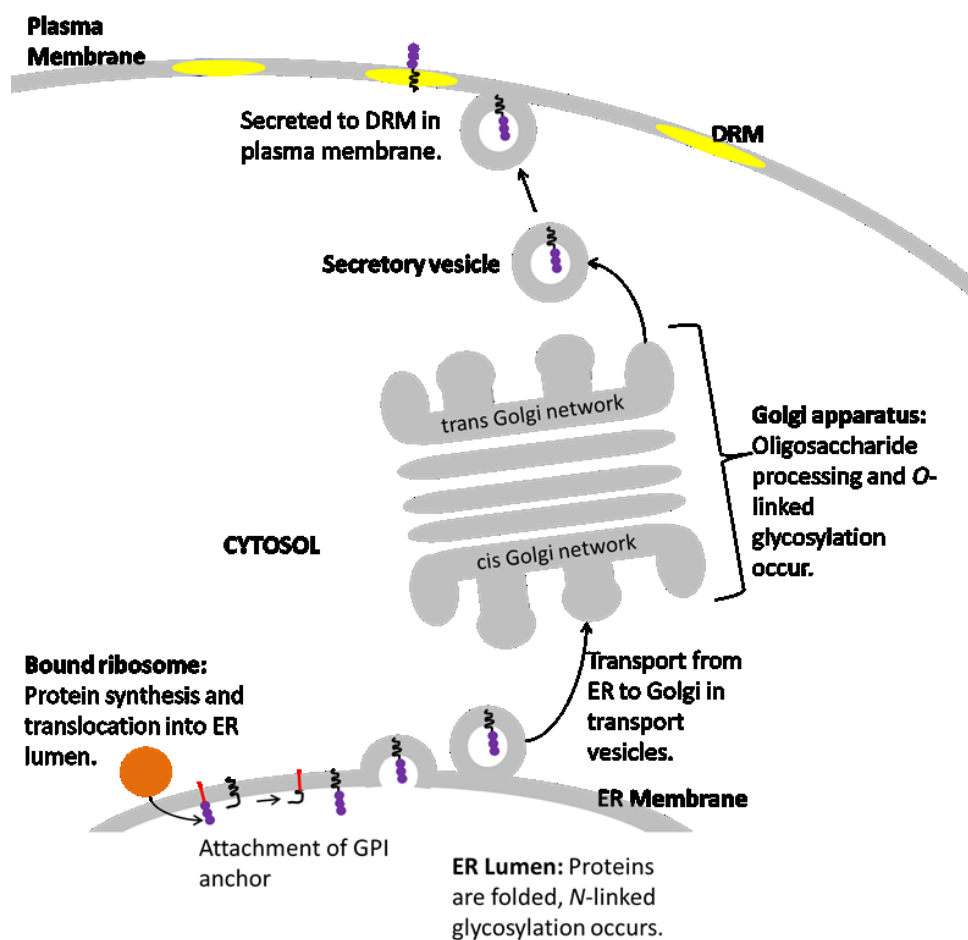


Figure 1 Schematic showing the secretory pathway OPCML undertakes to get to the plasmal membrane. Image taken from previous report¹²⁰.

1.2.3 Protein Structure

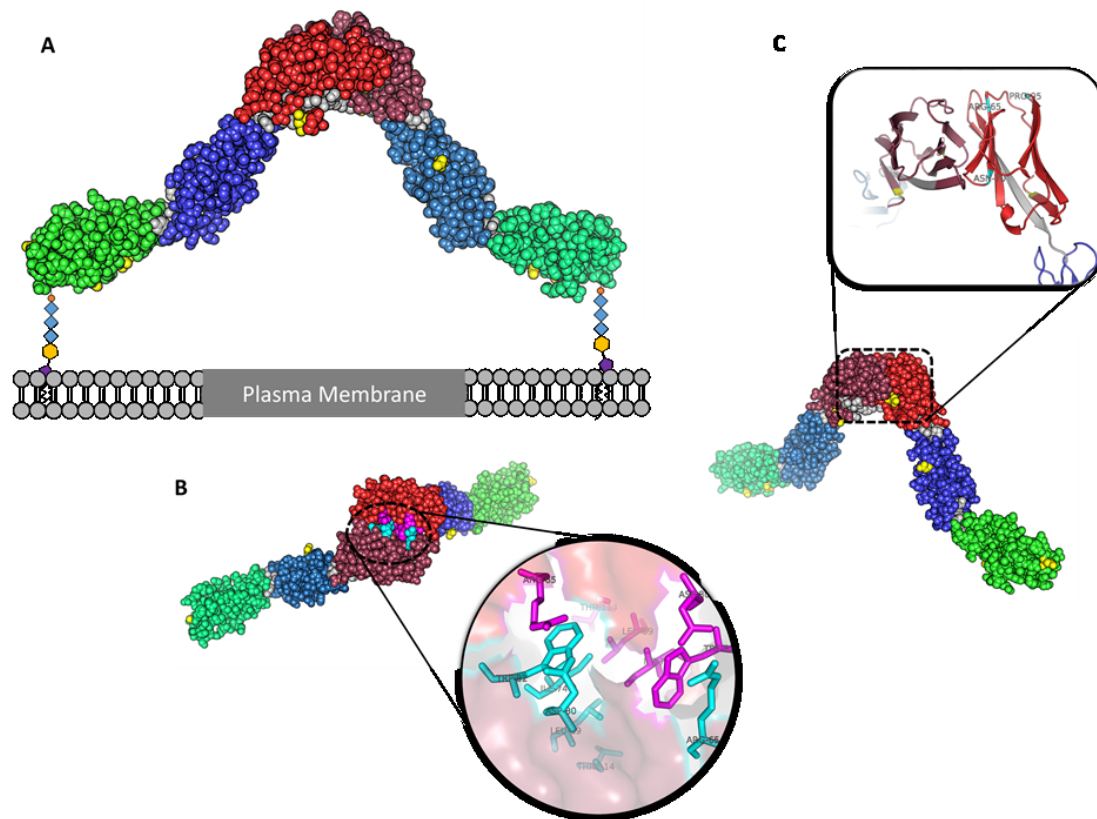


Figure 2 The structure of OPCML with sites of glycosylation (asparagine 44, 70, 140, 285, 293 and 306) shown in yellow. Domain 1 is depicted in red, domain 2 blue and domain 3 green. Short extended linkers between domains are depicted in grey. (A) OPCML homodimer shows as a schematic on the surface on the plasma membrane. GPI anchor attachment occurs at Asp322. (B) Key residues for the Domain 1 homodimer interface; Arg65, Leu69, Ile74, Asp80, Trp82 and Thr114 (cyan for one OPCML and magenta for the second OPCML). (C) Residues that undergo mutation to give loss of tumour suppressor function R65L, N70H and P95R.32

The structure of OPCML has been recently elucidated by Birtley et al³³ (**Figure 2**). They reported that OPCML forms a homodimer on the surface of cells through hydrogen bonds, and hydrophobic contacts and salt bridges between residues in domain 1 (D1) (**Figure 2A**). Arg65, Leu69, Ile74, Asp80, Trp82 and Thr114 are key residues in the homodimer interface (**Figure 2B**). Arg65 forms a reciprocal salt bridge with Asp80. The D1-D1 interface is composed of hydrophobic contacts, hydrogen bonds and salt bridges. A reciprocal salt bridge is formed between the Arg65 of one monomer and Asp80 of the other and hydrophobic contacts form through Leu69, Ile74, Trp82, Thr114 and Arg127 that stabilise the D1-D1 interface (**Figure 2B**).

D1 is at the N-terminus of the protein and forms 26.9 % of the total protein³³. R71C is the most common mutation found in D1 and is found on the underside of the D1 homodimer interface.

Introduction

This mutation caused OPCML to lose ability to inhibit migration, invasion and anchorage independent growth.

R65L, N70H and P95R mutations all demonstrated loss of tumour suppressor function (**Figure 2C**)³³. R65L is located at the dimer interface in D1 and in OPCML WT Arg65 forms a reciprocal salt bridge with Asp80. This mutation caused a disruption in homodimerisation and subsequent loss of function with the mutant protein being unable to bind to RTKs after stimulation with ligands. The N70H mutation is also within D1 at a position where there is an N-linked glycan. This mutant protein disrupted glycosylation and protein stability causing loss of function. Pro95 is in a type I β -turn of D1 and P95R, P95L and P95S mutants were still expressed but were only able to interact with some of the RTKs which interact with OPCML wild type (WT), thus suggesting the protein has different methods of interaction for different RTKs.

Mutations were also carried out in D2 and D3 in this study with mutants exhibiting loss of function in migration and invasion assays³³. However it was found that while mutation in D1 caused a decrease in growth of cancer cells, D2 and D3 mutation did not inhibit cell proliferation. These mutations all show the importance of D1 for homodimerisation and tumour suppressor function and informed the design of OPCML constructs for producing a novel biotherapeutic.

1.2.4 Biological Function

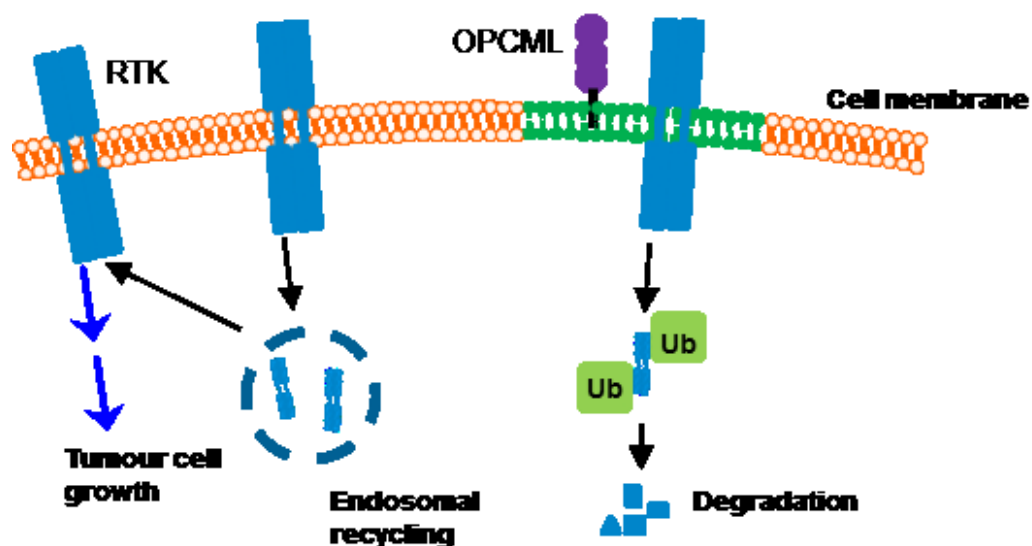


Figure 3. Schematic representation of the biological mechanism of OPCML^{32,172}.

The GPI anchor localises OPCML to the detergent resistant membrane (DRM) which is an area of the cell membrane with a high concentration of cholesterol and (glyco)sphingolipids^{32,52}. OPCML acts as a tumour suppressor by negatively regulating a subset of receptor tyrosine kinases (RTKs) including EPHA2, FGFR1, FGFR3, HER2 and HER4^{32,53}. RTKs are transmembrane proteins that signal from the outside to the inside of cells influencing processes such as cell growth and proliferation. It is hypothesised that OPCML sequesters selected RTKs to the DRM and causes OPCML binding dependent polyubiquitination and subsequent degradation of the RTKs arresting their signalling cascade, preventing cell proliferation and growth³².

AXL is an RTK OPCML interacts with: OPCML has a high affinity for AXL when it is bound to AXL's ligand Gas6. Upon OPCML binding, AXL is sequestered into the DRM and interacts with the phosphatase PTPRG which dephosphorylates pAXL, causing its inactivation⁵⁴. These interactions take place in the DRM where OPCML is localised due to the GPI anchor, highlighting the importance of this moiety for OPCML function.

1.3 GPI Anchors and ligation techniques

1.3.1 GPI Anchor Structure and Function

GPI anchors have a general structure which consists of a glycan core with three mannose sugars and a glucosamine, a phosphoethanol amine linker to the anchored protein and a phosphatidylinositol lipid anchor with a phosphodiester-linked inositol headgroup and two fatty acid chains that insert into the cell membrane (**Figure 4A**)⁵⁵⁻⁵⁷.

GPI anchors are attached posttranslationally after being synthesised in the ER^{3,55,58}. The GPI anchor is transferred to the protein when removal of the protein's secretion signal peptide occurs⁵⁵. GPI anchors themselves undergo further modifications in the Golgi apparatus where the anchor undergoes maturation steps such as fatty acid removal^{3,48,55}.

Although the general structure of GPI anchors stays the same, they vary with differing side groups on the glycan core and varying lengths and saturation of the fatty acid chains^{56,59}. The fatty acid chain length is of importance to function of GPI-AP with it being reported that the chain length influences the association with the DRM, with shorter chain length (C14) having no significant association with the DRM whereas longer chain length (C18) showed association⁶⁰. Additionally, Malherbe et al⁶¹ reported removing the GPI anchor lipid moieties from GPI-APs prevents association with lipid bilayers, indicating the importance of GPI anchors for membrane insertion, and prevented recognition by antibodies raised against GPI-APs.

GPI anchors have been shown to have more purpose than just anchoring proteins to the plasma membrane of cells, with GPI anchors directing proteins to specific areas of the cell membrane such as detergent resistant membranes (DRM)^{52,55,58}. Furthermore, when associated with cell membranes it has been found that GPI-APs can be present on the cell surface as monomers but also as nanoscale cholesterol sensitive clusters of a maximum of four molecules per cluster⁶².

A further structural feature of GPI anchors is the glycan core. While glycans have been found to be unnecessary for membrane insertion, they do affect the diffusion within the membrane^{56,63}.

GPI-APs have also been found to have therapeutic value with the GPI-AP CD44 being suggested as a strategy for HIV mediated gene therapy⁶⁴. Furthermore, GPI-anchored variants of proteins have also been investigated for use in tumour therapy^{65–68}. However, there are reservations about used GPI-APs in biomedicine. Concerns include the difficulty in purification and solubilisation of these proteins, as well as the anchors being unable to differentiate between tumour cells and normal cells when targeting the cell membrane⁵⁹.

1.3.2 GPI Anchor Mimetics

Synthesis of a naturally occurring GPI anchor requires complex synthetic routes with a moiety for protein ligation needing to be considered^{56,69}. Therefore GPI anchor mimetics are a valuable tool. An advantage to using GPI anchor mimetics is the ability to design anchors for the specific purpose with functionality added dependent on need. Studies have been undertaken that highlight the potential of using GPI anchor mimetics with phospholipid tails as a replacement for naturally occurring GPI anchors^{69–72}.

Many GPI anchor mimetics utilise lipid moieties (lipidation) to insert into the cell or model membranes. Ball et al⁷⁰ used myristic acid as a potentially simplified GPI anchor mimetic (**Figure 4B**). In their research, the presence of an orthogonally protected lysine (Lys) at the C-terminal of polypeptides of interest was coupled to myristic acid, enabling binding to the surface of cells. Binding was lost after exchanging the myristoyl group to an acetyl, highlighting the importance of the lipid extension⁷⁰.

27

Introduction

A further example of lipidated anchor mimetics is Paulick et al.^{56,63} and their synthesis of GPI anchor mimetics with varying amounts of glycans (**Figure 4C**). The mimetics consisted of; phosphoethanolamine linker attached to a free cysteine for ligation to a protein, a glycan core of one, two or no mannose residues, a poly(ethylene glycol) linker instead of phosphoinositol and glucosamine, and a phospholipid tail for anchoring to the cell membrane. The simplest anchor with no glycan core was synthesised using common reactions and building blocks via a 9-step synthetic route with varying yields at each stage. They ligated these anchor mimetics to GFP (green fluorescent protein) and found that the glycan core of GPI anchors is not necessary for proteins to insert in the cell membrane. However, the absence of monosaccharides does affect the mobility of protein in lipid bilayers with decreased diffusion within the cell membrane^{56,63}. Hill et al.⁷³ reported attaching a protein to cell membranes using a GPI anchor mimetic consisting of two myristoyl groups for insertion into the plasma membrane and a charged hydrophilic peptide to associate with the phospholipid headgroups in the plasma membrane (**Figure 4D**)⁷³.

GPI anchor mimetics have also been generated that use a cholesterol moiety rather than a phospholipid to insert into the cell membrane^{74,75}. One study from Teruya et al.⁷⁶ reports the addition of a GPI anchor mimetic comprised of cholesterol, a hydrophilic diethylene glycol-based diamine linker and a reactive cysteine for ligation to GFP with the protein remaining functional and localising to cell membrane when added exogenously (**Figure 4E**)⁷⁴. A further study by Schafer et al.⁷⁵ synthesises a GPI anchor mimetic using a cholesterol with a fluorescent moiety included to aid in visualisation of anchored proteins (**Figure 4F**) and hypothesise that cholesterol has the potential to target the attached protein to the DRM⁷⁵. Previous research has reported exogenously added cholesterol modified fluorescein was observed in cholesterol rich membrane domains⁷⁷. However, this targeting to the DRM was not evident in the case of Teruya et al.⁷⁶ or Schafer et al.⁷⁵.

OPCML is localised to the DRM of cells by a GPI anchor. In the DRM OPCML functions by sequestering RTKs to the DRM and arresting their signalling cascade^{32,52–54}. Therefore, the GPI anchor is important for localising OPCML to the cell membrane where it can function as a tumour suppressor. As localising to the DRM is important an anchor mimetic that could target the DRM would be beneficial. Furthermore, GPI anchor mimetics have the added benefit of adding functionality by inclusion of fluorescent groups or tags. However, it is important that if included such tags do not adversely affect the function of the protein^{75,77}. Most of the GPI anchor mimetics mentioned in this section lack the glycan core found in naturally occurring GPI anchors which, though not essential for membrane insertion, affects the diffusion within the

cell membrane^{56,63}. Therefore, this needs to be kept in mind as its membrane diffusion may be important for OPCML function and protein-protein interactions. The purification required after GPI-anchor modification of the protein must also be considered as this may affect the protein function.

1.3.3 Approaches for C-terminal Modification of Proteins

In order to attach GPI anchor mimetics to proteins of interest the anchor must be functionalised correctly depending on the method chosen to modify the protein. As OPCML has a GPI anchor present at the C-terminus (Asp 322) the following highlights various approaches for C-terminal modification of proteins⁷⁸.

Modification utilising cysteine reactivity

Free cysteines are found in somewhat low abundance in proteins. For example in OPCML there are only six cysteines in the mature protein all of which are present as disulfide bonds³³. Therefore it is possible to introduce free cysteines as point mutations at the desired location within the protein for subsequent modification.

One method by which to modify free cysteines is to utilise the natural nucleophilicity of the cysteine thiol/thiolate with an electrophile. One such example of this is using a thiol reactive maleimide on the anchor mimetic and to conjugate this to a Michael donor thiol group on the protein via Michael addition as shown in a paper by Schafer et al⁷⁵. However, this method can only be used when the only free cysteine present is the one required for conjugation otherwise unwanted conjugation could occur.

Ligation utilising Bio-orthogonal Chemistry

Functional groups not present in nature such as alkynes and azides can be integrated into proteins and subsequently used for ligation⁷⁸. Copper catalysed azide alkyne cycloaddition (CuAAC) is a reaction that can be used in this way⁷⁹⁻⁸². Strain-promoted azide-alkyne cycloaddition (SPAAC) is a further method⁸³. However, before ligation can occur azide or alkyne functionality must be added to the protein of interest.

Strategies to introduce azide functionality to the C-terminus of proteins investigated by Wang et al⁸⁴ include sortase mediated transpeptidation (discussed later in this section) and by expressing the protein with an unnatural amino acid replacing methionine with azidohomoalanine⁸⁴. A further approach by Khatwani et al⁸⁵ was to enzymatically add azidofarnesyl or azidogeranyl groups to C-terminally tagged proteins by using the enzyme farnesyltransferase to catalyse attachment of azidoprenyl molecules onto the cysteine of

proteins with a C-terminal CVIA tag^{86,87}. Khatwani et al⁸⁵ then went on to use CuAAC or copper free SPAAC to attached azide-functionalised protein to an oligonucleotide.

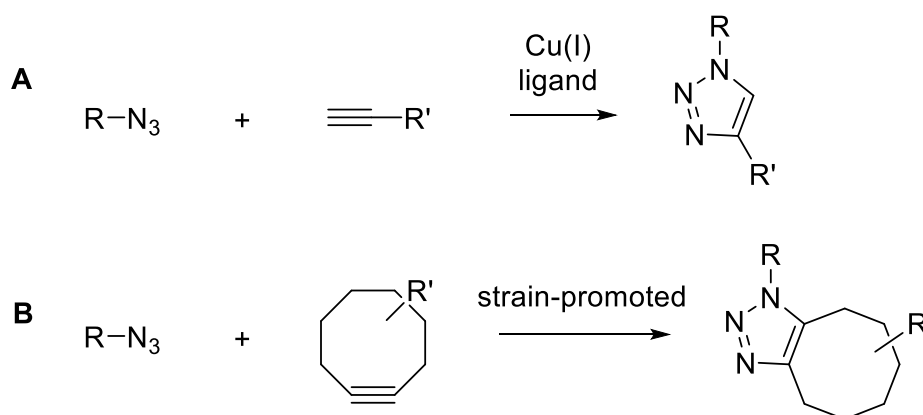


Figure 5 (A) Copper catalysed azide alkyne cycloaddition (CuAAC)^{79–82}. **(B) Strain-promoted azide-alkyne cycloaddition (SPAAC)**⁸³.

Intein-mediated Protein Ligation

A further approach used is Intein-mediated Protein Ligation (IPL) (**Figure 6**) also known as expressed protein ligation^{88–92}. Numerous papers report using a fused a self-cleavable intein containing an affinity binding domain for chitin to the C-terminus of a protein to purify recombinant proteins in a single step^{88,93–95}. An intein (*intervening protein*) is a length of protein present in a precursor protein that excises itself via peptide bond cleavage and ligates flanking exteins (*external proteins*) generating a mature protein and stable intein^{96–98}. This intein fusion technology was developed to enable fusion of the intein at the N or C-terminus of the desired protein^{88,89,94,99}. If the recombinant protein can be expressed with a reactive C-terminal thioester after cleavage from the intein, ligation can occur by reaction with a substrate containing an N-terminal cysteine⁸⁸.

The IMPACT™ kit from New England Biolabs (NEB) uses these techniques to enable ligation in a methods known as IPL. IPL has been successfully used to site specifically ligate various peptides and lipid moieties to proteins^{90,100}. IPL utilises native chemical ligation (NCL) (**Figure 6**), a method by which two unprotected peptides can be conjugated via a chemoselective reaction to give a thioester-linked intermediate¹⁰¹. This thioester species spontaneously rearranges to give a product with a native peptide bond¹⁰¹. One drawback of NCL is that it requires the efficient synthesis of peptide thioesters, which can be difficult on proteins¹⁰². NCL was utilised by Paulick et al⁵⁶ to ligate a GPI anchor mimetic to a protein.

Introduction

The IMPACT™ kit from NEB enables a protein to be expressed with an intein and a chitin binding domain (CBD) when cloned into a pTXB1 vector⁸⁸. Once expressed the protein can be purified by affinity with a chitin column and eluted from the column using 2-mercaptoethanesulfonic acid (MESNA) to produce a protein with a C-terminal thioester⁸⁸. The protein can then undergo NCL with anchor mimetic that has an N-terminal cysteine (**Figure 6**)^{88–92}. However this approach does require the protein to be successfully expressed as an intein fusion.

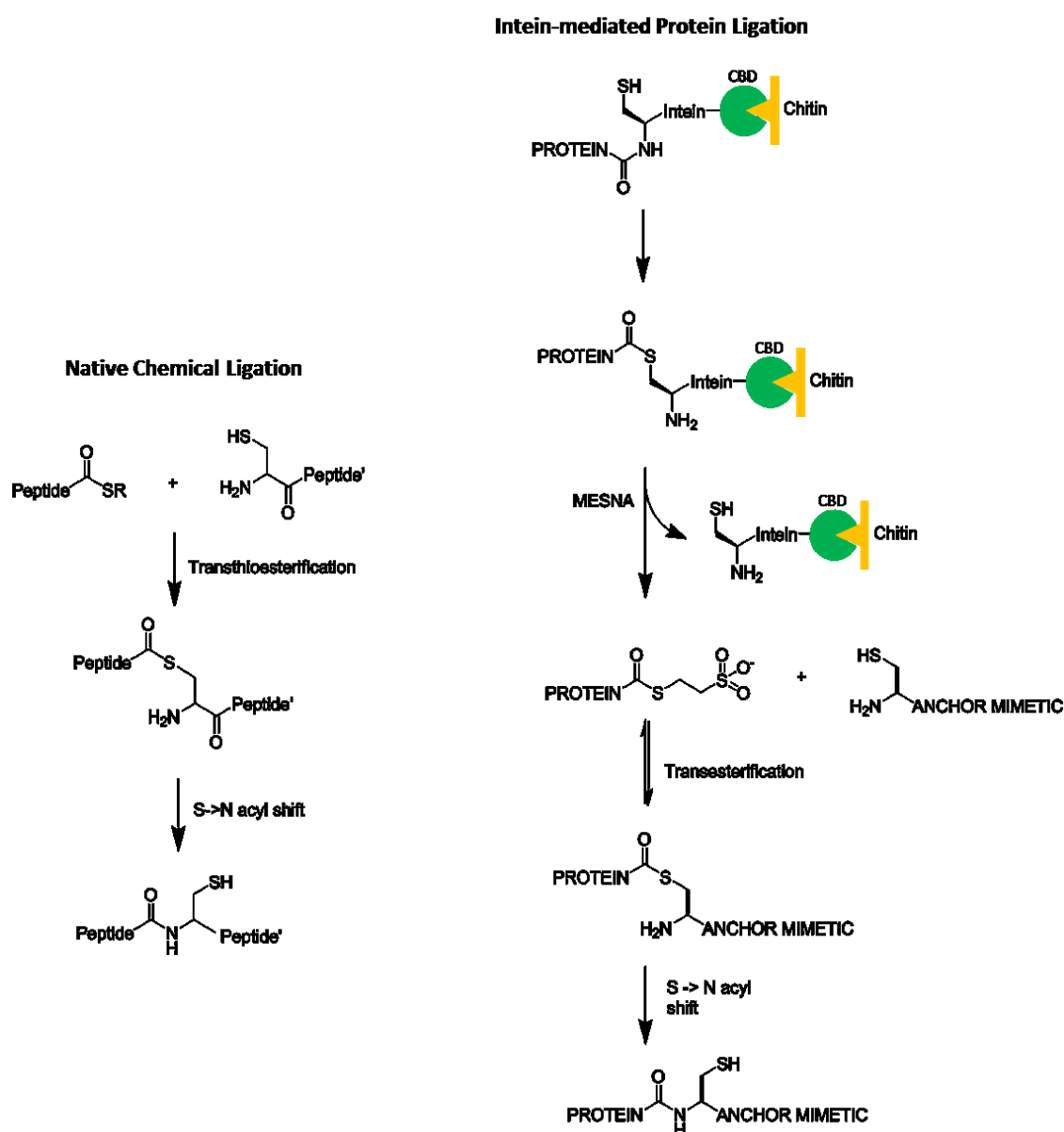


Figure 6 Schemes showing the mechanism of native chemical ligation and intein-mediated protein ligation. Schemes adapted/reproduced from papers^{92,101}.

Sortase Mediated Transpeptidation – Sortagging

Sortase is a membrane bound transpeptidase found on the surface of Gram positive bacteria that anchors proteins to the bacteria cell wall^{102,103,104}. Transpeptidases are enzymes that catalyse the transfer of a peptide or amino acid from one amino compound to another. Sortase from *Staphylococcus aureus* is known as sortase A (SrtA) and was purified and characterised by Ton-That et al¹⁰³ in the late nineties. To aid purification of the protein they included a six-histidine tag in their sortase construct.

Sortase A recognises the sequence LPXTG where X can be any amino acid but most commonly in natural sortase A substrates, X is glutamic acid. Cleavage occurs between the threonine and glycine creating an acyl-enzyme intermediate^{103,105}. The acyl-enzyme intermediate reacts with a pentaglycine which is a feature of Staphylococcal peptidoglycan that forms part of the chemical structure of the bacterial cell, wall forming an amide bond between the pentaglycine and the cell wall^{103,106}.

The sortase transpeptidation reaction can also be used for ligation termed “sortagging” with the first example of this demonstrated by purified sortase reacting with an LPXTG containing peptide and a H₂N-Gly₃ substrate¹⁰⁴. The rate of LPETG cleavage was also investigated with nucleophiles containing one, two or three glycines and it was determined that the rate of cleavage increases with increasing numbers of glycines¹⁰⁴. Furthermore, in the absence of a nucleophile hydrolysis of the LPXTG substrate occurs at the same site as the cleavage between the threonine and glycine^{104,107}.

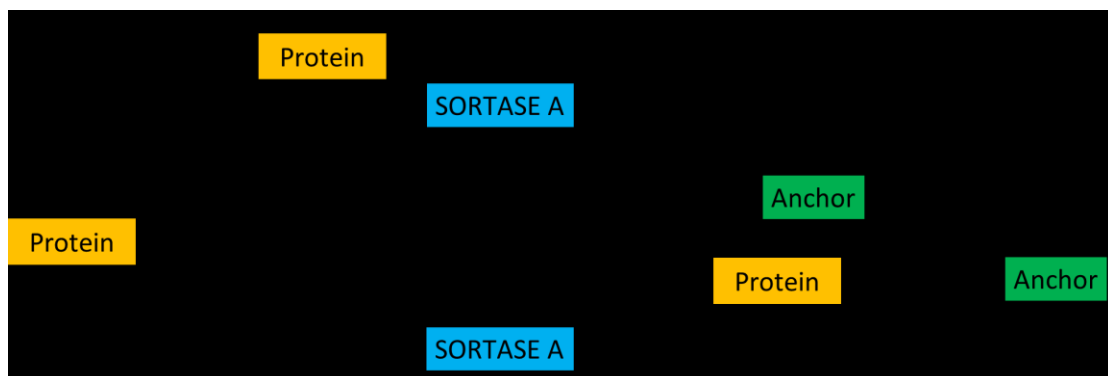


Figure 7 Schematic showing reaction scheme for sortagging and the attachment of the anchor to the protein. Image adapted from Antos et al ¹⁰⁹.

Sortase from *Staphylococcus aureus* SrtA is calcium dependent. Ca^{2+} ions are needed because the calcium ions binds to Glu residues which are distal to the active site of sortase and stabilise the substrate binding site¹⁰⁸. However, sortases have been engineered to overcome this dependence^{108,109}. A further limitation of sortagging is the reversible nature of the reaction, with nucleophiles being able to attack the LPXTG site present in the original protein substrate, as well as the LPETG site created from ligation of the protein and anchor^{110,111}. To overcome this, an excess of glycine as nucleophile can be used to push the reaction equilibrium towards the product¹¹¹. Another approach to prevent the reversibility of the sortagging reaction, is to prevent the formation of a glycine nucleophile during the transpeptidation. Replacing the C-terminal glycine in the LPXTG recognition site with a glycolic acid or a methyl ester are two ways in which the formation of a competing glycine nucleophile have been prevented^{112,113}.

Sortases have been utilised for applications such as anchoring proteins to surfaces, producing antibody drug conjugates and protein ligation in living cells^{109,114–117}. Mao et al¹⁰² investigated the possibility of using sortase for protein-peptide and protein-protein ligation. In their study they successfully demonstrated the ability of sortase to ligate GFP containing the LPETG recognition site with a glycine peptide nucleophile. Ligation was also successful between two GFP proteins one containing the recognition site, and one with N-terminal glycine showing that sortase can be a useful tool in protein engineering¹⁰². Sortagging has also been undertaken for the selective labelling of proteins in solution, cell lysates and on living cells¹¹⁸. Sortagging can be used as a way of attaching a GPI anchor mimetic to a protein with Guo et al¹¹⁹, showing this by successful linking of a GPI anchor mimetic containing a glycine residue with a peptide containing the LPXTG recognition site. Sortagging can also be used to cyclise or polymerise proteins¹¹⁰. **Figure 7** shows how sortagging can be utilised for ligation of anchor mimetics to proteins.

1.4 Summary

Cancer is a worldwide problem, EOC is a one of the most lethal gynaecological malignancies with a high rate of recurrence after treatment^{6,7,17,18}. OPCML is a tumour suppressor gene which is epigenetically inactivated in EOC with high incidence^{5,31}. OPCML is localised to the DRM of cells by a GPI anchor where it functions as a tumour suppressor by sequestering RTKs to the DRM and arresting their signalling cascade^{32,52–54}. Naturally occurring GPI anchors are synthetically challenging to synthesise, therefore anchor mimetics might be a viable option for ligating to recombinant proteins such as OPCML for localisation of proteins to the plasma

Introduction

membranes^{56,69–72}. There are various techniques for ligation such as IPL and sortagging that can be used to ligate GPI anchor mimetics to proteins^{75,90,100,104}.

1.5 Thesis Objectives

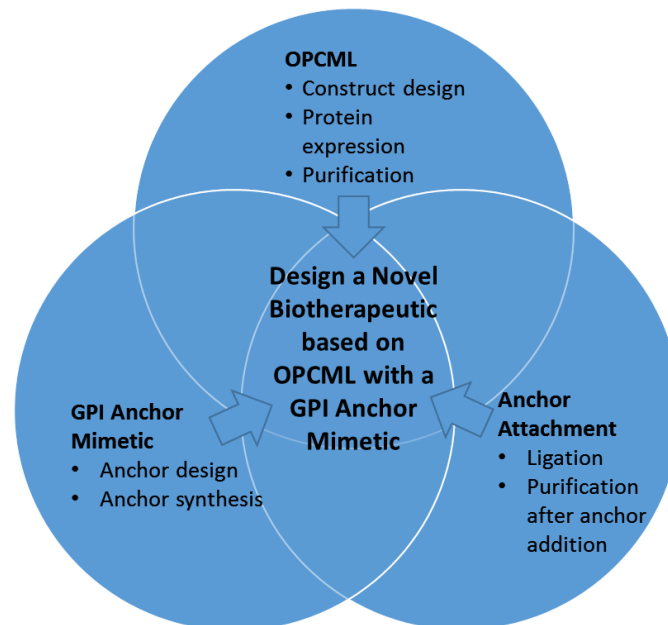


Figure 8 Schematic representation of three main parts of this research.

OPCML is epigenetically inactivated in EOC^{5,31}. Reintroduction of functional OPCML has been shown to have potential as a therapeutic with cancer cell growth and proliferation reduced upon addition of functional OPCML³². Recurrence of EOC is high with survival rates post EOC dramatically decreasing with time^{6,7,17,18}. If a therapeutic could be designed that prevented further spreading and dissemination of cancerous tissue after patients have undergone surgery the recurrence rate could be greatly decreased. This would benefit ovarian cancer patients by improving prognosis after surgery and give them a better quality of life post cancer. Thus the main aim of this research is to design and validate a novel biotherapeutic based on an OPCML-GPI anchor mimetic.

There are three parts of the research which work towards this main aim:

1. To produce recombinant OPCML.
 - Recombinant OPCML needs to be expressed such that it is functional and can be easily purified in high enough yield for downstream experiments. The protein must also be engineered in a way that a GPI anchor mimetic can be introduced.

Introduction

2. To design a GPI anchor mimetic.

- For the GPI anchor design the method of anchor attachment must be considered. The anchor must then be synthesised and purified ensuring all the required functionality is present for ligation as well as the possibility of including extra functionality in the form of tags for purification or visualisation.

3. To ligate the anchor to the recombinant OPCML.

- Ligation needs to be successful whilst not affecting the activity of the protein. A suitable method of purifying the protein-anchor complex must also be investigated.

Once these three objectives have been met and a novel OPCML-GPI anchor mimetic has been obtained in reasonable yields the final aim is to test the functionality of the protein-anchor in cell based assays.

Chapter 2

Testing SAR of OPCML

Knowledge of the structure activity relationship (SAR) of OPCML is important when designing a novel biotherapeutic for EOC using the tumour suppressor protein. Before knowledge of the crystal structure of OPCML was available it was decided to elucidate the SAR of OPCML via the generation of stably transduced cell lines. In particular the importance of each domain was to be investigated, with a view to finding out whether all domains are necessary for OPCML to function as a tumour suppressor. Some of the work presented in this chapter follows on from research previously reported please refer to Appendix D^{120,121}.

2.1 Design of constructs for SAR studies

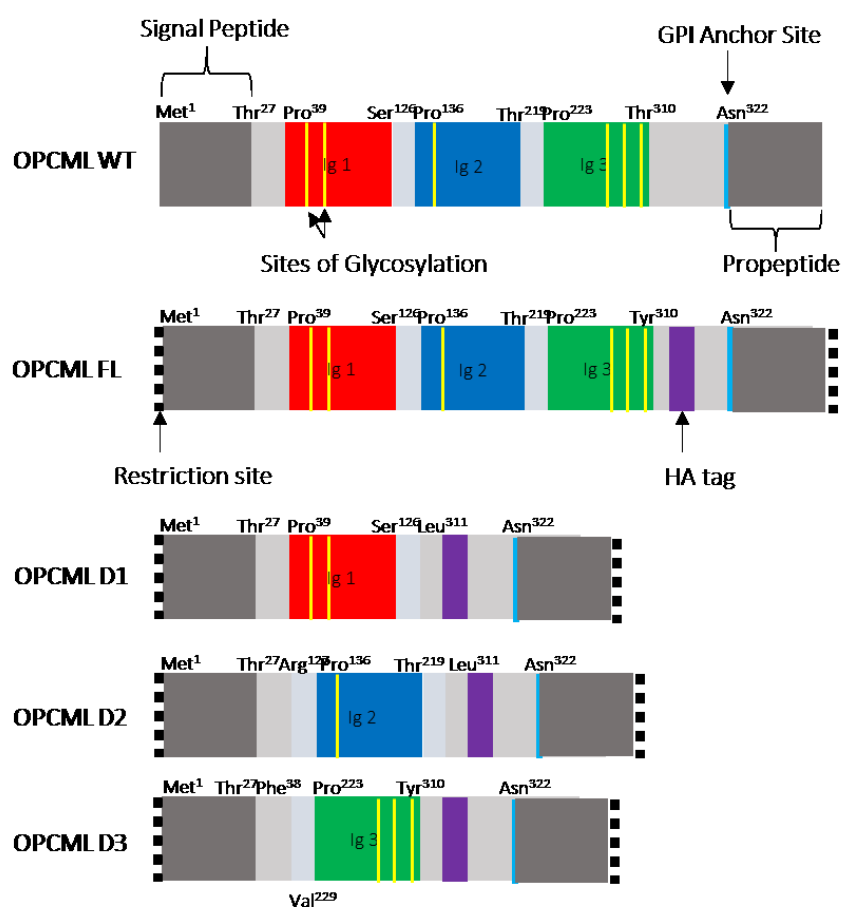


Figure 9 Schematics of OPCML WT and FL, D1, D2 and D3 constructs designed for testing functional importance of the signal domains¹²⁰. Image taken from report¹²¹.

Constructs for investigating the importance of each OPCML domain for tumour suppressor function were previously designed for lentiviral transduction in SKOV3 cells¹²⁰. This epithelial ovarian cancer cell line was chosen for transduction and functional assays as RT-PCR has shown the cells express no OPCML, and the cell line has been previously validated for OPCML functional assays^{5,32,54}. **Figure 9** shows schematics of wild type OPCML (OPCML WT) and the four constructs designed for testing functionality. All constructs tested for functionality included an HA tag at the C-terminus to aid in detection and visualisation of the protein in cells. They also all had restriction sites XbaI and Sall flanking them to enable digestion and ligation into a vector backbone. The full length construct, OPCML FL, included all three domains and all six sites of glycosylation. The constructs designed to test signal domains OPCML D1, D2 and D3 preserved the glycosylation sites and the sequences between the defined domains where possible. The OPCML constructs also have a signal sequence to translocate the protein to the cell membrane, and a propeptide sequence that is cleaved upon maturation or activation of the protein. As the constructs were designed for transduction into mammalian cells the signal sequence and propeptide were not altered. These constructs were ordered as plasmids using Invitrogen GeneArt® Gene Synthesis.

2.2 Generation of stably transduced cell lines

2.2.1 *Lentiviral transduction*

Lentiviral transduction is a suitable way of producing cell lines that stably express OPCML constructs endogenously. Lentiviral vectors can accommodate larger transgenes (~10 kilobases) and stably integrate into the cell genome due to their genetic alterations being maintained in dividing cells^{122,123}.

Lentiviral transduction works by first creating a lentivirus in cells from three plasmids: a lentiviral transfer plasmid containing the construct of interest, an envelope vector and a packaging vector¹²³. These plasmids are co-transfected into cells such as HEK293 or HEK293T, from which after a short incubation period the viral supernatant can be harvested¹²⁴. HEK293T are used in preference to HEK293 because they have the SV40 T-antigen making them more efficient at vector production, and they also have better transfection efficiency and growth¹²⁵. The viral supernatant is then filtered before being used to transduce the cell line of interest. The virus is internalised by interaction with the viral envelope protein and a receptor on the surface of the cells. Once internalised the virus RNA is reverse transcribed into proviral dsDNA which is then transported to the cell nucleus and integrated into the host genome¹²⁶. In this research the lentiviral vector used was pLenti CMV Puro DEST which expresses cDNA under the control of a

Testing SAR of OPCML

CMV promotor¹²⁴. pMDG is a plasmid that contains the VSVG gene encoding for the envelope, and p8.9 encodes for the polyproteins gag and pol¹²⁶.

2.2.1.1 Gateway® cloning to create a lentiviral transport vector

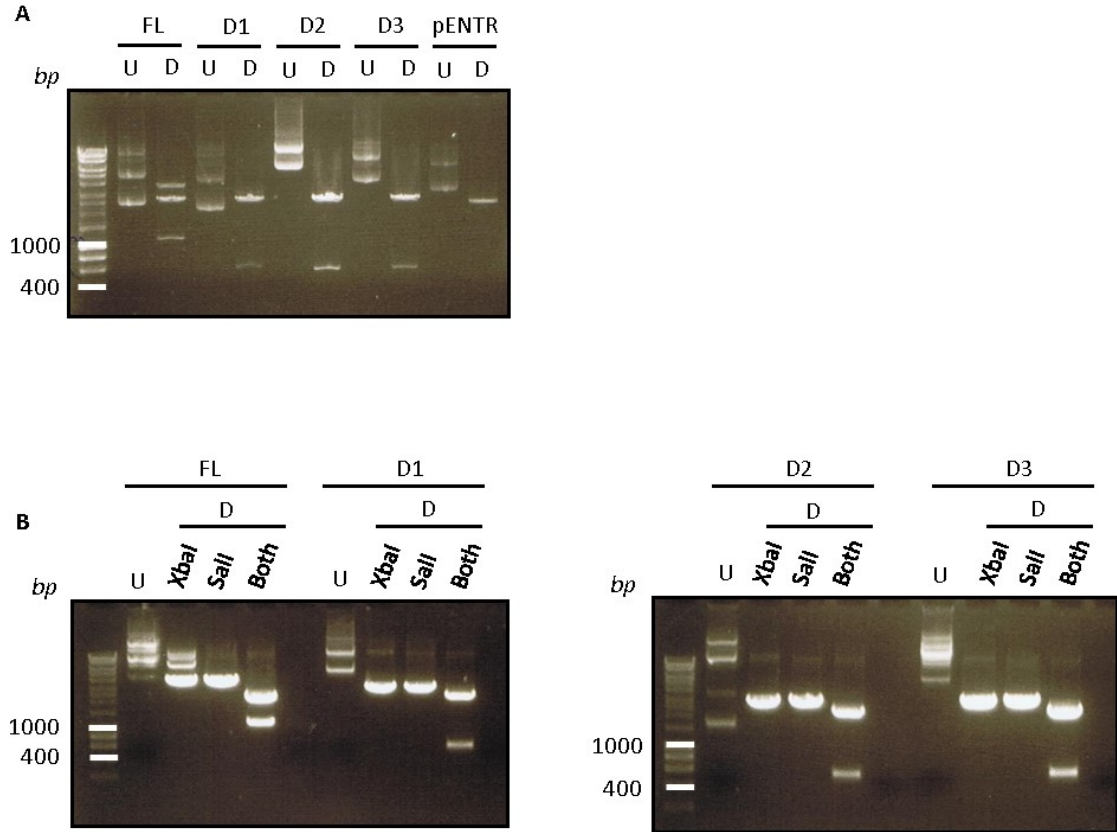


Figure 10 (A) DNA gel electrophoresis of OPCML FL, D1, D2, D3 and pENTR4 double digestion to show inserts and pENTR backbone. (B) DNA gel electrophoresis of digested FL, D1, D2 and D3 constructs in pENTR vector: The double digestion (both) confirms the presence of the construct in the entry vector. (A) and (B) FL is seen around 1000 bp and D1, D2 and D3 around 500 bp. Digestion was done used XbaI and SalI. U = undigested, D = digested. Images taken from report¹²¹.

Gateway® cloning (Life Technologies) was the technique used to get the construct of interest into the pLenti CMV Puro DEST vector¹²⁷. Gateway® cloning (Life Technologies) is based on the ability of bacteriophage λ to undergo site-specific recombination¹²⁸. This system allows site specific recombination between DNA recombination sequences (att sites) present in an entry vector and att sites in a destination vector in the presence of a Clonase™ enzyme mix^{127,129}. Once cloned into the entry vector the construct of interest is flanked by attL sites. An LR recombination reaction can occur upon incubation with LR Clonase™ between the attL sites in the entry vector and the attR sites in the destination vector^{127,129}. The product of this recombination reaction is the construct of interest in the destination vector which, in the case of this work, was pLenti CMV Puro DEST¹²⁴. The LR recombination reaction does not create any by-products due to the entry and destination vectors having different antibiotic resistances to

allow antibiotic selection of colonies. Furthermore, there is a *ccdB* gene present between the att sites in the destination vector, and this prevents cell growth by inhibiting DNA gyrase from correctly functioning (DNA gyrase is an enzyme that relieves strain when ds DNA is being unwound)¹³⁰. If Gateway® cloning has occurred successfully the *ccdB* gene should no longer be present in the destination vector, thus the bacteria will be able to grow¹²⁷. The differing antibiotic resistances ensures only the bacteria containing the destination vector will grow, and the *ccdB* gene means only bacteria containing successfully cloned destination vector are viable^{124,129}.

The construct design was previously reported for use with Gateway® cloning and production of stable cell lines¹²⁰. This previous attempt aimed to use the constructs as entry vectors directly performing Gateway® cloning with the ordered constructs into the destination vector, but this proved unsuccessful with low transduction efficiencies observed¹²⁰. Therefore, in this research the same constructs were used but with traditional cloning to get the construct into a pENTR4 entry vector before Gateway® cloning was used for generation of the lentiviral destination vector. All constructs were digested with XbaI and SalI as was the pENTR4 vector and ligated with ligation Ready-To-Go T4 DNA Ligase following the manufacturer's protocol. The double digestion of the OPCML constructs was successful as shown by the presence of inserts corresponding to the correct size in **Figure 10A**. This figure also shows the successful linearization of the pENTR4 vector. The DNA inserts were purified from the agarose gel prior to ligation. Sequencing and DNA gel electrophoresis (**Figure 10B**) confirmed ligation had occurred. After successful generation of the OPCML pLenti vectors HEK293T cells were co-transfected to produce virus that was collected and used to transduce SKOV3 cells (See methods).

2.2.2 Characterisation to confirm SKOV3 transduction with OPCML constructs

Confirmation of protein expression

To confirm the SKOV3 cells were expressing the protein constructs SDS PAGE and immunofluorescence microscopy (IFM) were carried out. Three anti-OPCML antibodies were tested due to the possibility that the single domains of OPCML D1, D2 and D3 may not be recognised if they do not contain the correct epitope for the specific antibody. Controls of SKOV3 cells transduced with a pLenti empty vector and parental SKOV3 cells were also tested to ensure no endogenous expression of OPCML was present. **Figure 11A** confirms D1, D2, D3 and FL are being expressed. The positive control OPCML WT was identified with all the anti-OPCML

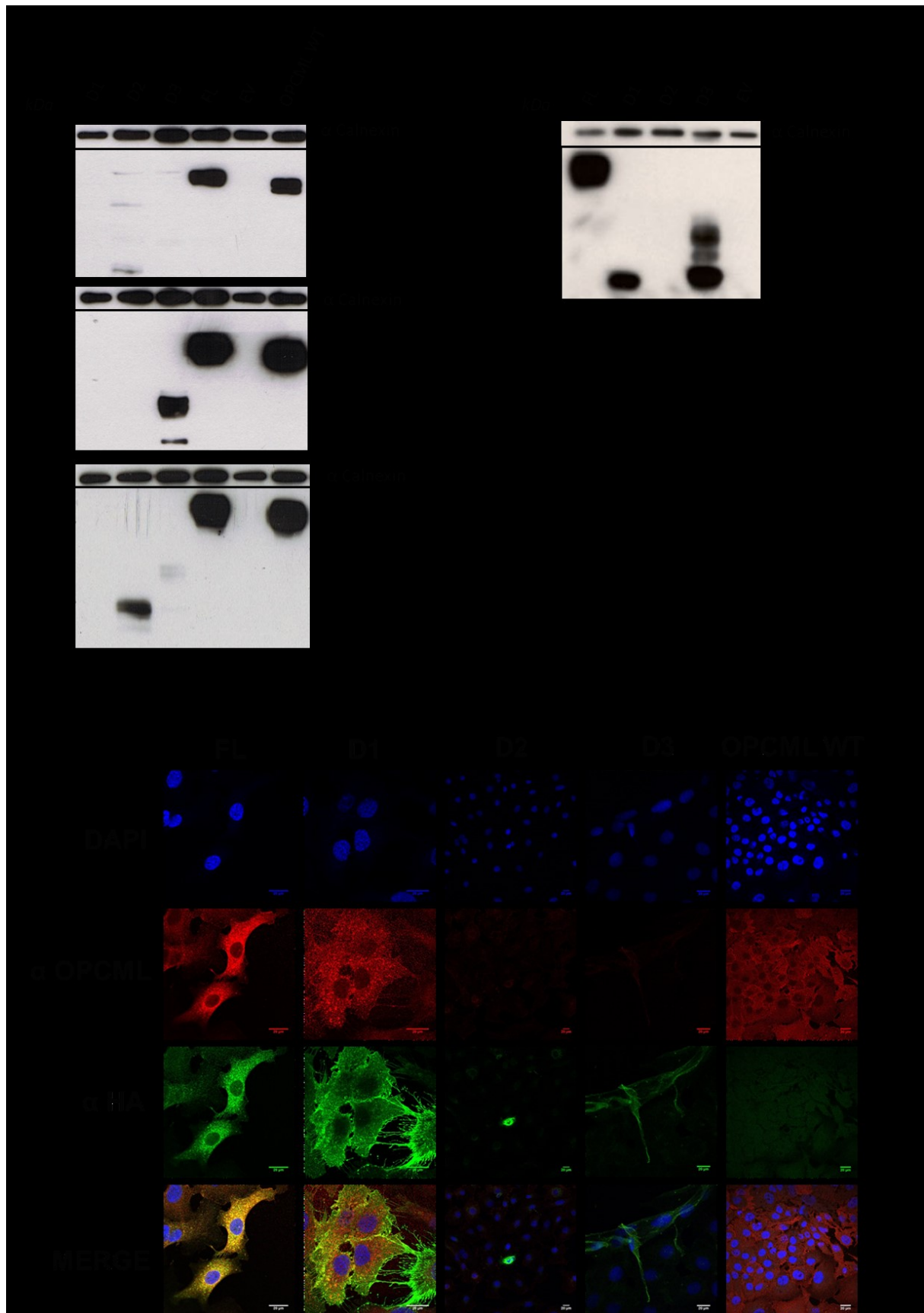


Figure 11 (A) SDS- PAGE analysis of transduced SKOV3 cells to confirm expression. Image adapted from report¹²¹. (B) Confocal immunofluorescence microscopy showing the expression and location of proteins in transduced SKOV3 cells: SKOV3 cells transduced with FL, D1, D2, D3 or OPCML WT controls were fixed using PFA (See Methods) and stained with rat anti-HA (green), mouse anti-OPCML (red) and DAPI (blue). Scale bars 20 μ m.

antibodies and the EV negative control was not identified at all. The monoclonal mouse antibody only shows a signal for D3, FL and OPCML WT indicating that D3 is the region of the protein this antibody recognises. The rabbit and goat antibodies are polyclonal. However, the rabbit only recognises FL and OPCML WT and the goat only recognises D2. The anti-HA antibody recognises FL, D1 and D3. Therefore, combining the results from all anti-OPCML antibodies and the anti-HA antibody it is shown that the transduced cell lines are expressing the desired protein segments at the expected sizes with FL and OPCML WT around 55 kDa and the single domains at 20 kDa.

Confocal IFM also confirms the expression of FL and D1 in SKOV3 cells. **Figure 11B** shows IFM images of cells after fixation with PFA (See methods). FL and D1 are stained clearly with both the anti-mouse and anti-HA antibodies. OPCML WT is successfully stained with anti-OPCML but there is also background signal from the HA antibody. Combining the results obtained from the SDS-PAGE and IFM it can be seen that the transduced SKOV3 cells are expressing the desired constructs.

2.2.2.1 Confirmation of expected cell phenotype

Previous studies have shown the expression of OPCML in cells suppresses cell growth, with transfected cells expressing differing levels of OPCML having growth inhibition when compared to a cell line transfected with a control vector³². Therefore, to test the transduced cell lines are behaving as expected MTS assays to survey cell growth were undertaken (see Methods).

It is expected that the SKOV3 cells expressing functional OPCML would show less cell growth over time than the empty vector control. However, the results from this assay differed greatly between repeats. **Figure 12** shows the results of the cell proliferation assays. In two of the three repeats the SKOV3 cells expressing OPCML WT show the most cell growth with the D1 transduced cell line showing the least growth. This is not the case in the third repeat where the SKOV3 FL cells grow the most and SKOV3 D3 the least. The large variation in cell growth for the transduced cell lines upon each repeat suggests that transduced cell lines are not a suitable model for testing the SAR of OPCML.

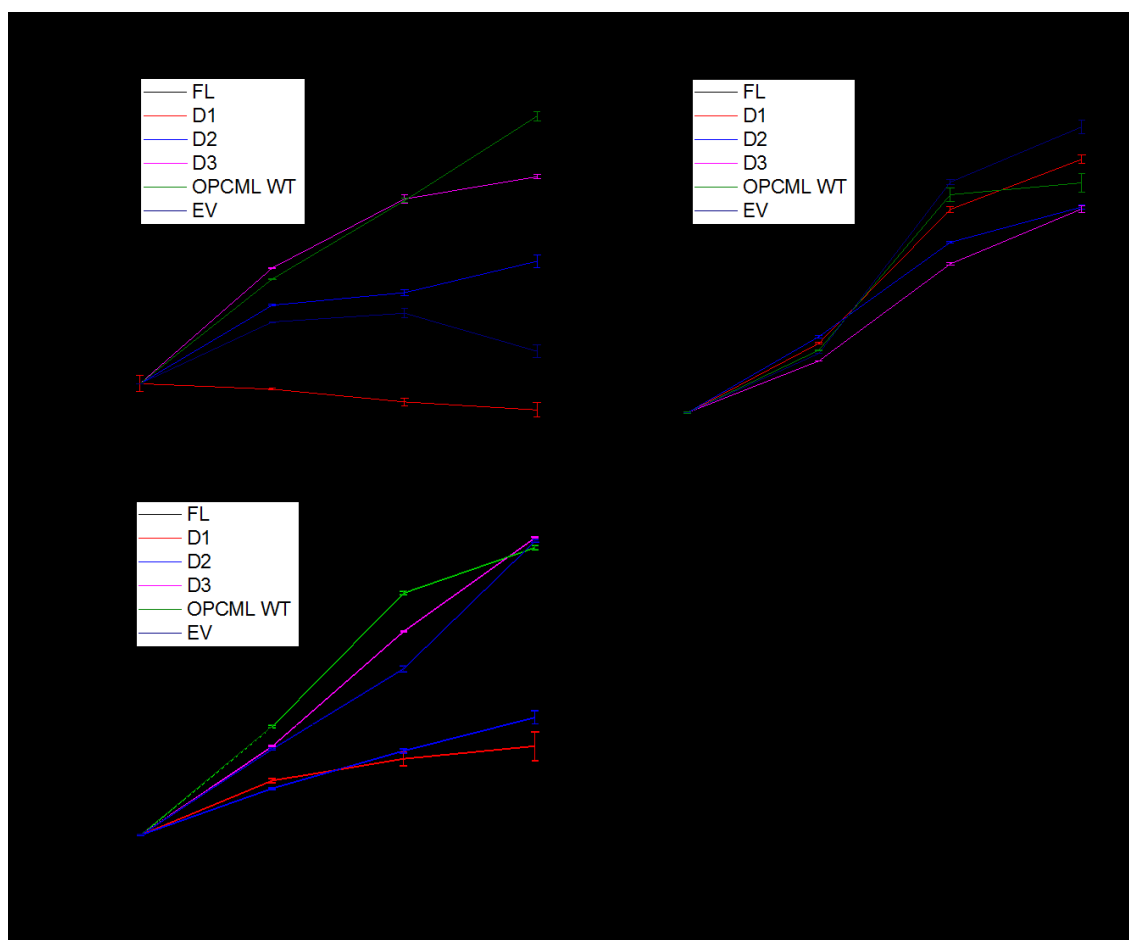


Figure 12 MTS cell growth assay of lentivirally transduced SKOV3 cells. Results were taken in triplicate.

2.3 Conclusion

The initial aim of the research presented in this chapter was to determine the SAR of OPCML, to ultimately produce a novel biotherapeutic based on the generation of a protein-GPI anchor mimetic. The results presented in this chapter led to the decision not to take transduced cell lines any further in SAR studies. It is not possible to control how much OPCML is present in each cell line as the protein is inherently expressed. The IFM images indicate non-uniform levels of protein expression throughout the transduced cell population, and the MTS assays indicate problems with reproducibility. There were also unsuccessful attempts with these transduced cell lines at immunoprecipitation and gap closure assays. These initial experiments were designed with the aim of elucidating the SAR of OPCML, specifically which domains are required for OPCML tumour suppressor function. As previous unpublished research had shown expressing and purifying recombinant OPCML was very challenging and thus neither the

Testing SAR of OPCML

structure nor SAR was known. However, the structure of OPCML was subsequently elucidated by collaborators and this allowed SAR information to be gathered and provided a method of reliable protein production³³. Thus the remaining chapters focus on the production a novel biotherapeutic based on the generation of a protein-GPI anchor mimetic

Chapter 3

The design and synthesis of a highly engineered and functional OPCML construct

3.1 Construct design and protein expression in mammalian and insect cell lines

The main objective of this research was to create a novel biotherapeutic with a semi-synthetic OPCML targeted to the membrane of cancer cells using a GPI anchor mimetic. A construct for OPCML was designed to enable the required functionality to be added to meet this aim whilst ensuring the function of the protein was not adversely affected. The crystal structure of OPCML was elucidated and it was determined that the protein forms a homodimer through hydrogen bonds and hydrophobic contacts in domain 1 (D1)³³. Furthermore mutation in D1 caused loss of tumour suppressor gene function. D1 is vital for dimerisation of OPCML on the cell surface; the R65L mutation disrupted homodimerisation and this mutant protein was unable to bind to partner RTKs. Moreover an N70H mutation was found to disrupt glycosylation and protein stability.

Therefore to ensure the construct designed retains function there were no changes to D1 in the construct design and no tags were introduced at the N terminus, which is present in D1, to avoid disrupting dimerisation. To replicate the native GPI anchor attachment, GPI anchor mimetics were attached to the C-terminus and it was decided to keep as many of the six glycosylation sites conserved as possible to avoid disrupting protein stability through lack of glycosylation.

The methods of protein expression and anchor ligation also needed to be considered during construct design. Some of the work presented in this chapter has been previously reported¹²¹.

There are a variety of methods for protein production all with associated strengths and weaknesses. Bacterial expression is commonly used as it is generally high yielding, but it is not capable of introducing post translational modifications. For example bacteria lack the cellular organelles needed for glycosylation^{131,132}. Mammalian systems provide post translational modifications but have much lower yields^{131,132}. Insect cells are another system that can be used for producing recombinant proteins. Insect cells have the advantage of providing higher

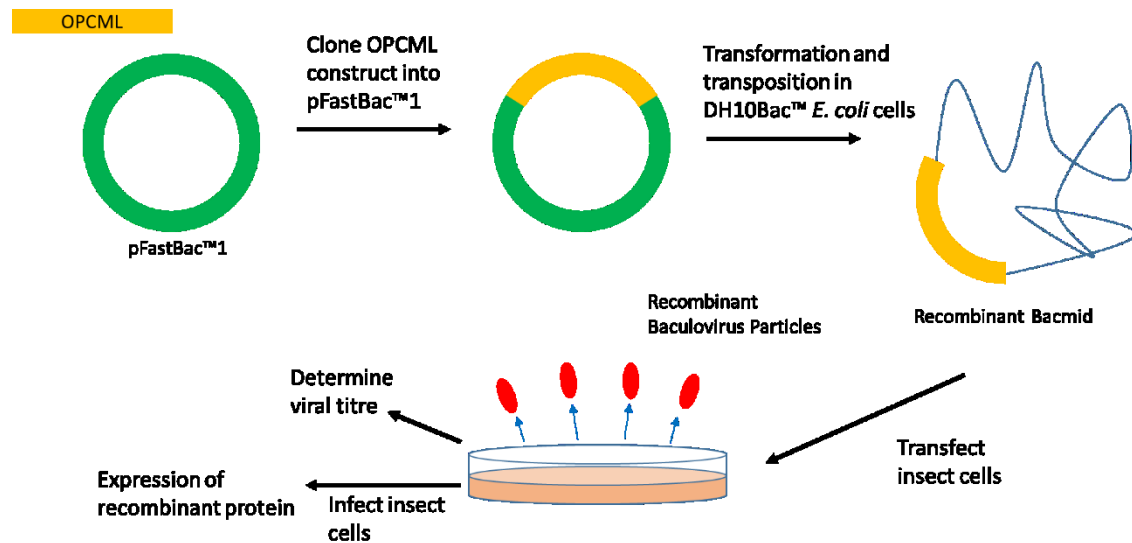


Figure 13 Schematic showing the stages in producing a baculovirus expression system ¹³⁵. Image taken from previous report ¹²¹.

expression levels when compared to mammalian cultures whilst still being able to carry out post translational modifications unlike bacterial systems^{133,134}. **Figure 13** shows a schematic of the process to make a baculovirus using the Bac-to-Bac® expression system from Invitrogen^{135,136}. The first step is to clone the construct of interest into a FastBac™ plasmid. This plasmid contains a baculovirus specific promotor, an antibiotic resistance gene, a mini F replicon, LacZα gene and a mini-attTn7 site which is the target site for the bacterial transposon Tn7.^{135,136} Transposons are mobile sections of DNA and during transposition, a transposase enzyme cleaves the DNA that flanks transposons and then inserts this DNA into another area³. During transformation of the pFastBac™1 vector into DH10Bac™ competent cells the transposition occurs. In the case of the baculovirus expression system the transposition occurs between mini-attTn7 in the DH10Bac™ competent cells (which contains a helper plasmid with all the proteins necessary for a transposition reaction to occur) and the mini-Tn7 in the plasmid to form the bacmid^{135,137}. The success of the transposition is evident by the colour of the colonies formed when the transformed bacteria is spread on a LB agar plate. If transposition is successful the expression of LacZα is disrupted and colonies will appear white. Unsuccessful transposition will result in colonies that appear blue. The bacmid then undergoes transfection into insect cells to form a

The design and synthesis of a highly engineered and functional OPCML construct

baculovirus viral stock. This stock can then be used to infect insect cells which will product recombinant protein¹³⁵.

As OPCML needs to retain function for the latter stages of the project it was decided not to use a bacterial expression system due to the six glycosylation sites and three disulphide bridges, as expression in bacteria can cause abnormal disulphide bonds which could affect structure and stability of the protein. A secretion signal peptide is present in many secreted proteins and indicates to the cell to translocate the protein to its membrane. Thus the secretion signal is important and needs to be optimised for the chosen expression system.

The method of attaching an artificial anchor also needed to be considered. First the GPI anchor attachment site was removed so unanchored OPCML could be expressed³². Sortase mediated transpeptidation or sortagging (see Introduction) was selected as a suitable technique therefore the sortase recognition sequence LPETG was included in the construct at the C terminus. The recognition site needs to be accessible to the sortase enzyme therefore a GGGGS precedes the recognition site in the construct. **Figure 14A** shows the first amino acid after the initial methionine in OPCML was changed from glycine to alanine. This change was done to prevent OPCML polymerising or cyclising during sortagging – when the sortase reaction takes place the glycine nucleophile attacks the sortase enzyme intermediate and if the initial glycine remained it would compete with the anchor mimetic during this step¹⁰⁵. A free cysteine was also included near where the native anchor attachment site would be as this could enable anchor attachment by Michael Addition to a maleimide functionalised anchor mimetic if sortagging proved unsuccessful⁷⁵.

A way of visualising the protein other than with OPCML antibodies is also desirable so an HA tag was included in the construct design. Furthermore, an affinity tag for isolating protein from cell supernatant was included in the construct as was a thrombin cleavage site to enable tag cleavage following purification³³.

3.1.1 Mammalian Expression of rOPCML in HEK293 cells

Initial expression was attempted in HEK 293 cells taking all the design points from above into consideration. As the expression system is mammalian and OPCML is normally expressed in eukaryotic cells the native secretion signal was conserved. A Twin-Strep-Tag® was included as the affinity tag for purification **Figure 14A**.

HEK293 cells were transfected using Lipofectamine® 2000 and the cell supernatant was harvested, concentrated and proteins were separated by SDS-PAGE. The western blot (**Figure**

The design and synthesis of a highly engineered and functional OPCML construct

14B) shows a faint band around 60 kDa in the OPCML transfected lane that is not present in the control lane. This band corresponds with the size of the commercial rOPCML (Creative BioMart) control showing that OPCML protein has been produced and secreted into the cell supernatant.

Figure 14C shows the coomassie stain of the SDS-PAGE and there is no clear band that corresponds to OPCML. The band at 80 kDa is due to BSA: Although the cells were washed after plating and before transfection residual BSA remained.

The lack of clearly defined signal shows that, although OPCML can be produced successfully using this mammalian expression system, the yields are low and there would a significant scale up before a suitable quantity of OPCML could be obtained for later modification and testing. At this time research from a collaborator had shown that functional rOPCML can be produced using the baculovirus expression system with reasonable yields³³. Therefore, it was decided to

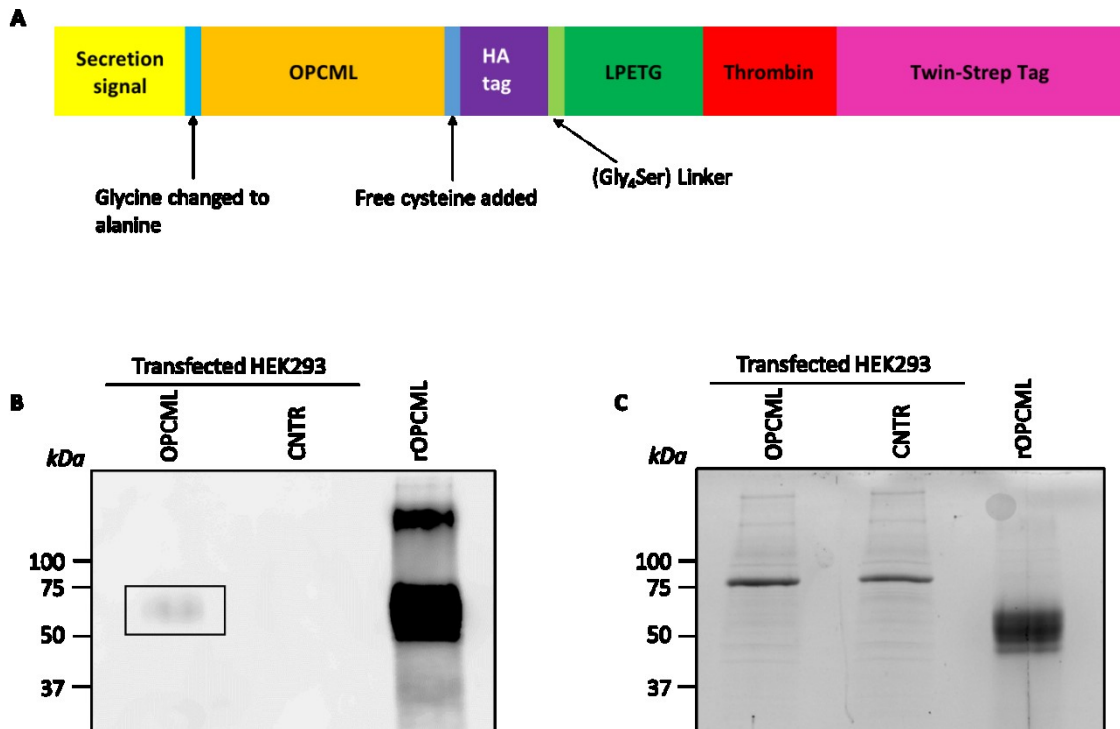


Figure 14(A) Schematic representation of the OPCML construct designed for a mammalian expression system. **(B)** and **(C)** SDS-PAGE showing expression level of OPCML in transfected HEK293 cells: HEK293 cells were transfected with OPCML (60 kDa) in a pcDNA3.1+ vector or a control empty pcDNA3.1+ vector. Commercial rOPCML (purchased from Creative BioMart) was used as a control. **(B)** Western blot **(C)** Coomassie staining.

The design and synthesis of a highly engineered and functional OPCML construct

switch to an insect cell expression system. Insect cell systems are capable of post translational modifications and tolerate disulphide bonds better than bacteria¹³⁴.

3.1.2 Insect Cell Expression of OPCML using the Baculovirus Expression System

The construct for insect cell expression was designed taking into account the considerations for the mammalian construct with a few alterations based on a previous report of OPCML production using insect cells by Birtley et al³³. Insufficient secretion is a problem associated with insect cell expression of proteins however with the addition of an insect secretion signal this problem can be resolved¹³⁴. Following the design and expression of OPCML in the paper by Birtley et al³³ the protein was expressed as an Fc-fusion protein, a ceruloplasmin secretion signal was added (MKILILGIFLFLCSTPLQ) and a thrombin site after the LPETG sortase recognition site. After the thrombin cleavage site two glycines were included followed by residues 120 to 330 of human IgG1 (Uniprot entry P01857) as an affinity tag ³³ (**Figure 15A**). Briefly, the construct was cloned into a pFastBac™1 vector (GeneArt Invitrogen). The pFastBac™1-OPCML vector was then transformed into DH10Bac™ competent cells to generate the recombinant bacmid. The production of the recombinant bacmid was confirmed via PCR. Analysis of the production of recombinant bacmid by PCR was necessary as recombinant bacmid DNA is too large for restriction analysis to be successful. The pUC/M13 primers flank either side of the mini-*att*Tn7 site where the gene of interest should insert. Therefore, the PCR product after agarose gel electrophoresis should correspond to the size of the bacmid transposed with pFastBac™1. **Figure 15B** shows three lanes of bacmid DNA (isolated from three separate colonies) and a control PCR reaction which had no DNA present. The size of the bacmid alone is around 300 bp, when successfully transposed the size of the PCR product will correspond to the bacmid with pFastBac™1 (2300 bp) plus the size of the insert. The size of the bacmid lanes in **Figure 15B** show the PCR product is about 4000 bp which corresponds to a successful transposition.

Sf9 insect cells were transfected with the bacmid to generate the baculovirus and the titre of this virus increased by a further infection of Sf9 cells before the viral supernatant was harvested. **Figure 15C** shows the production of OPCML begins after 24 h with a band around 70 kDa for OPCML+Fc. However, the protein seems to undergo cleavage with two bands appearing one at 70 kDa for OPCML+Fc and one around 45 kDa. As the Fc sequence has a calculated weight of 24.4 kDa the weight difference between the bands suggests that this lower band is OPCML without the Fc sequence. The test expression proved successful with OPCML being secreted into

The design and synthesis of a highly engineered and functional OPCML construct

the supernatant after 24 h. When left for longer than 72 h the Fc affinity tag appears to cleave therefore supernatant was harvested after 72 h.

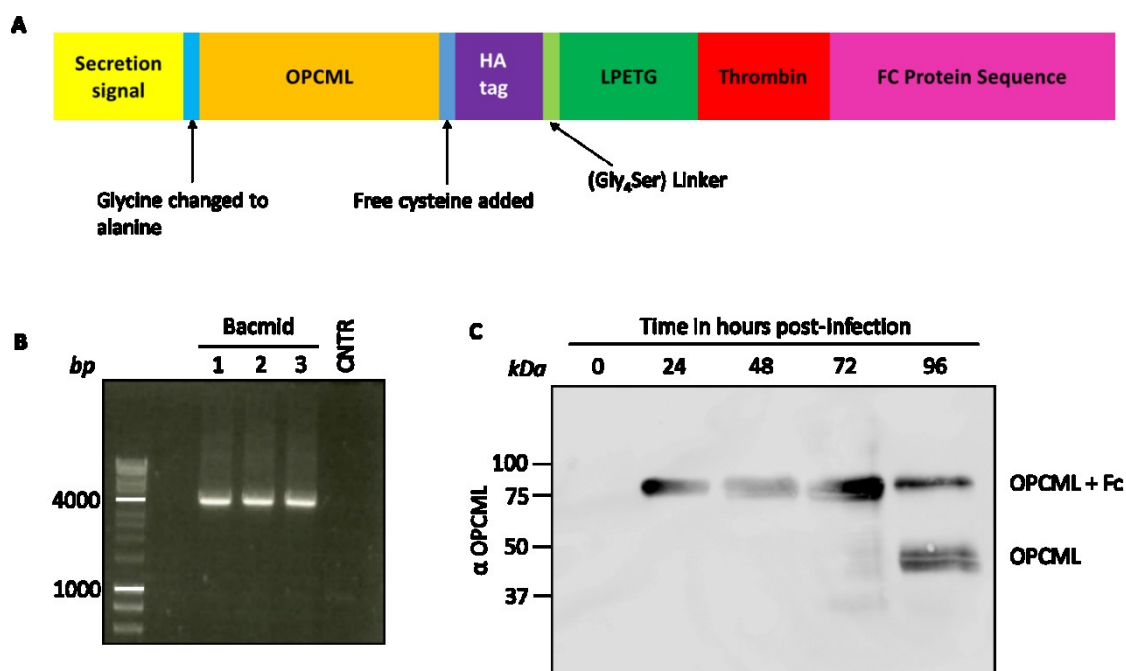


Figure 15(A) Schematic representation of the OPCML construct designed for an insect expression system. (B) Agarose Gel Electrophoresis of recombinant bacmid. Image adapted from report¹²¹. (C) Western blot showing expression level of OPCML in High 5 cell supernatant over time: High 5 cells were infected with OPCML Baculovirus and aliquots taken every 24 hours.

3.2 rOPCML Purification and Characterisation

3.2.1 Affinity purification to isolate rOPCML from insect cell supernatant and confirmation of protein identity

As expression was successful larger scale production and purification was initiated. **Figure 16** shows the SDS-PAGE gel of the different stages of the affinity purification of rOPCML from the cell supernatant. Batches of 1 L were prepared with 40 mL of baculovirus used to infect each litre of HighFive cells. After 72 h the supernatant was harvested, PMSF was added to a final concentration of 20 mM to inhibit serine proteases and pH was altered to 7.0 with NaOH. The supernatant was then filtered and 2 mL Protein A agarose (Abcam) was added per litre of supernatant and left overnight at 4 °C to bind to the Fc sequence in the protein. The next day the supernatant with resin was shaken at 700 rpm for 1 h at 18 °C to facilitate further binding of the protein to the resin, and then the supernatant was allowed to pack into a fritted column by gravity at 4 °C. Once packed the resin was washed with 30 BV of wash buffer then eluted

The design and synthesis of a highly engineered and functional OPCML construct

with 15 BV of elution buffer in 1 mL aliquots. The aliquots were neutralised with neutralisation buffer and the protein containing aliquots determined using a Nanodrop (MW = 58147.1 g/mol $\epsilon = 71120 \text{ M}^{-1}\text{cm}^{-1}$) before being combined and concentrated to a volume of 2.5 mL. The protein was then buffer exchanged using PD-10 Desalting Columns into wash buffer for long term storage and concentrated using a spin concentrator.

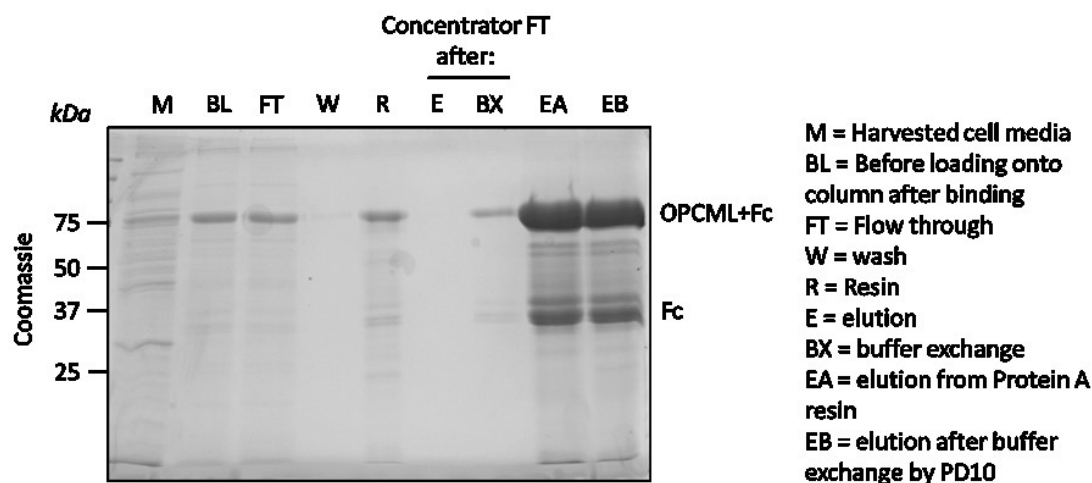


Figure 16 SDS-PAGE showing stages of rOPCML protein production.

To confirm the identity of the bands at 75 kDa and 30 kDa peptide mass fingerprint (PMF) analysis was carried out (Appendix C). The bands were cut from the EA lane of the gel (**Figure 16**) and digested with trypsin into peptides whose masses were measured by MALDI measurements. These peptides were then compared to a database of known peptide sequences to identify the sample using mMass software. The data identified the upper band as OPCML+Fc and the lower band as Fc confirming the earlier hypothesis that over time the Fc sequence cleaves from the rest of the protein.

3.2.2 Cleaving the affinity sequence from rOPCML in preparation for assay development and future experiments

As mentioned previously the OPCML construct was designed with a thrombin cleavage site so that after affinity purification the Fc sequence could be cleaved off to leave a protein that is more closely related to the native protein (**Figure 15A**). As thrombin is 37 kDa and after thrombin cleavage the rOPCML is 33.7 kDa (calculated using PepCalc) the size difference between cleaved rOPCML and thrombin is not large enough for separation by size exclusion chromatography (gel filtration).

3.2.2.1 Using resin bound rOPCML-Fc to test thrombin cleavage conditions

To test that thrombin cleavage would occur successfully a small scale on-resin cleavage of OPCML+Fc with thrombin was carried out. Insect cell supernatant was harvested as mentioned previously (3.2.1) and 100 μ L of the resin bound OPCML+Fc (after overnight incubation) was diluted into 490 μ L wash buffer containing 1 unit of thrombin (10 μ L). The reaction was incubated for 4 h at 10 °C and aliquots of supernatant taken after 1, 2 and 4 h for analysis via SDS-PAGE.

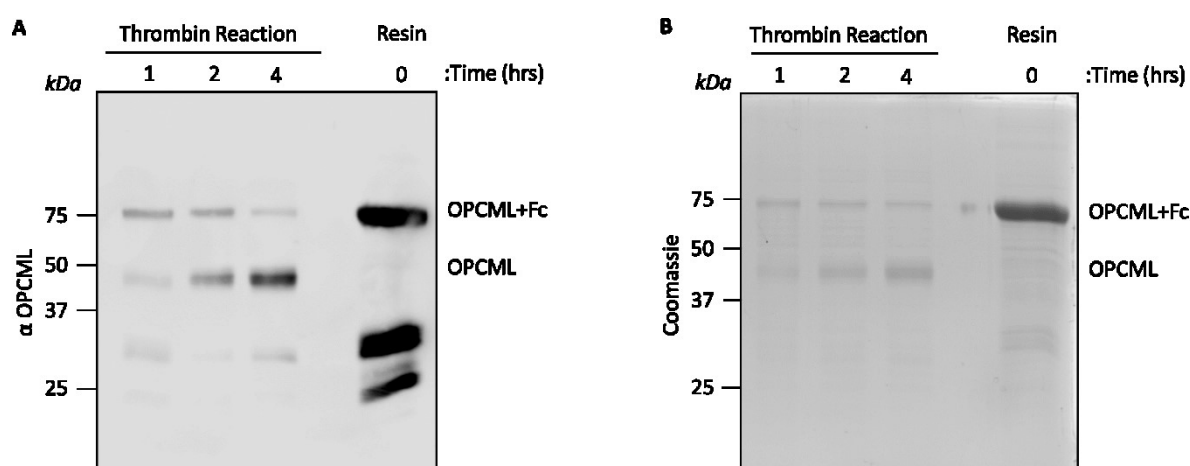


Figure 17 SDS-PAGE of on-resin thrombin cleavage over time. (A) Western Blot. (B) Coomassie stained gel.

Figure 17A and **B** have a band around 40 kDa showing there is an increase of OPCML over time. In both the western blot and the coomassie stained gel there is still a band present at 75 kDa corresponding to OPCML+Fc which hasn't been cleaved by the thrombin. At each time point the reaction was centrifuged to pellet the resin without breaking it and enable a resin free sample to be taken. However, some resin beads were still present in the aliquots which explains the uncleaved OPCML+Fc present in the figure. It can be seen from **Figure 17A** that over time, as well as an increase in the amount of OPCML that has been successfully cleaved, there is a decrease in the amount of OPCML+Fc present on the resin. The upper band present in the T=24 h lane was confirmed to be rOPCML+Fc by PMF analysis (Appendix C). The lower band corresponds to the size of cleaved OPCML. This shows that the thrombin cleavage site is accessible in the designed protein construct and the protein can be cleaved successfully. It would be possible to undertake the thrombin cleavage of the Fc sequence while the protein is still bound to the Protein A agarose which would prevent the need to purify out any OPCML+Fc which hasn't been cleaved from cleaved OPCML. However, the problem of removing thrombin from the cleaved rOPCML would still remain.

3.2.2.2 Using a HiTrap™ Benzamidine column to remove thrombin and prevent protein degradation

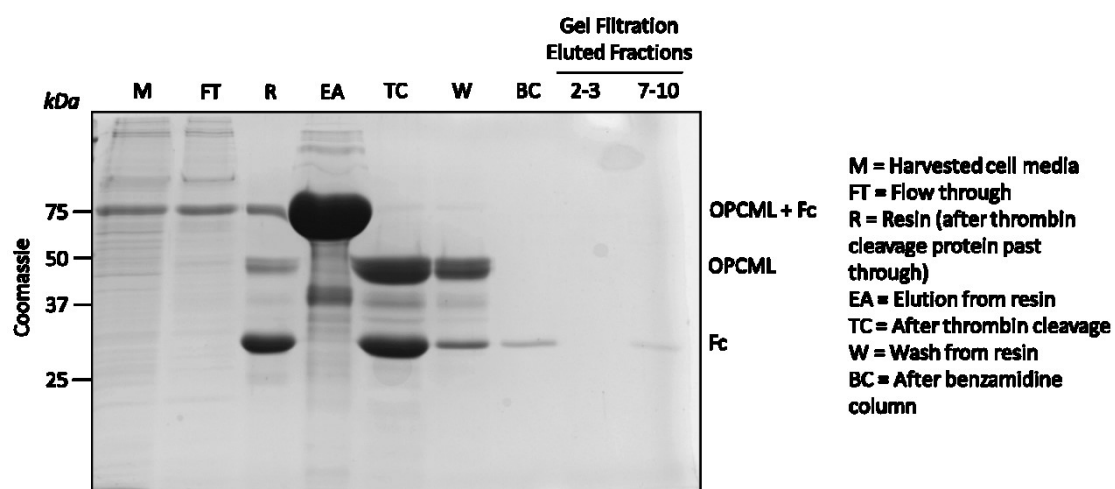


Figure 18 SDS-PAGE showing stages of purification after thrombin cleavage.

HiTrap™ Benzamidine FF columns (GE Healthcare) can be used to remove thrombin after thrombin cleavage to prevent further protein degradation. Thrombin cleavage was performed in wash buffer overnight at 4 °C. **Figure 18** shows the stages of thrombin cleavage purification. The first two lanes (BF and FT) are the cell media harvested from insect cells before and after filtering. The thrombin cleavage site is positioned between the OPCML and the Fc sequence in the construct. After thrombin cleavage the protein is passed through Protein A agarose to remove any un-cleaved OPCML+Fc and Fc. Lane 'R' is the boiled Protein A agarose and shows bands which correspond the Fc and rOPCML+Fc. The next two lanes show the protein before (EA) and after (TC) thrombin cleavage. Briefly, the column was equilibrated with binding buffer (0.05 M Tris-HCl, 0.5 M NaCl, pH 7.4), the sample was added and the column washed with binding buffer to collect the unbound recombinant protein. The column wash was checked by a NanoDrop™ spectrophotometer for the presence of recombinant protein however none was detected. This could be because the column has some ionic bonding characteristics thus the protein might get stuck to the column. Therefore an increase in salt concentration to 1 M was tried to encourage protein to elute. The lane 'BC' in **Figure 18** shows that after using the HiTrap™ Benzamidine FF column there is only one band present and this band corresponds to the Fc sequence, there is no rOPCML eluted from the column. This shows the cleaved rOPCML stayed bound to the column even after increasing salt concentration. Therefore it was decided to try a different method to remove thrombin after the cleavage reaction.

3.2.2.3 Cleaving rOPCML+Fc with resin bound thrombin using the Thrombin CleanCleave™ Kit

The Thrombin CleanCleave™ kit contains thrombin immobilised on agarose resin which enables the cleaved protein to be washed off the solid support, therefore resolving the issue of removing thrombin. The previous thrombin cleavage (using the Using a HiTrap™ Benzamidine column) went to completion after reacting overnight at 4 °C therefore the first attempt using the thrombin agarose was undertaken using these conditions. **Figure 19A** shows the recombinant OPCML protein before and after incubation with the thrombin resin following the manufacturer's instructions (40 µL resin was used with 500 µL OPCML+Fc at 0.33 mg/ml). The figure shows that overnight incubation at 4 °C did not cause the cleavage to go to completion, thus further optimisation was needed. **Figure 19B** also shows that the Fc sequence after cleavage further degrades into two bands above 25 kDa as the Fc appears as one band at 0 h and two bands from 1 h.

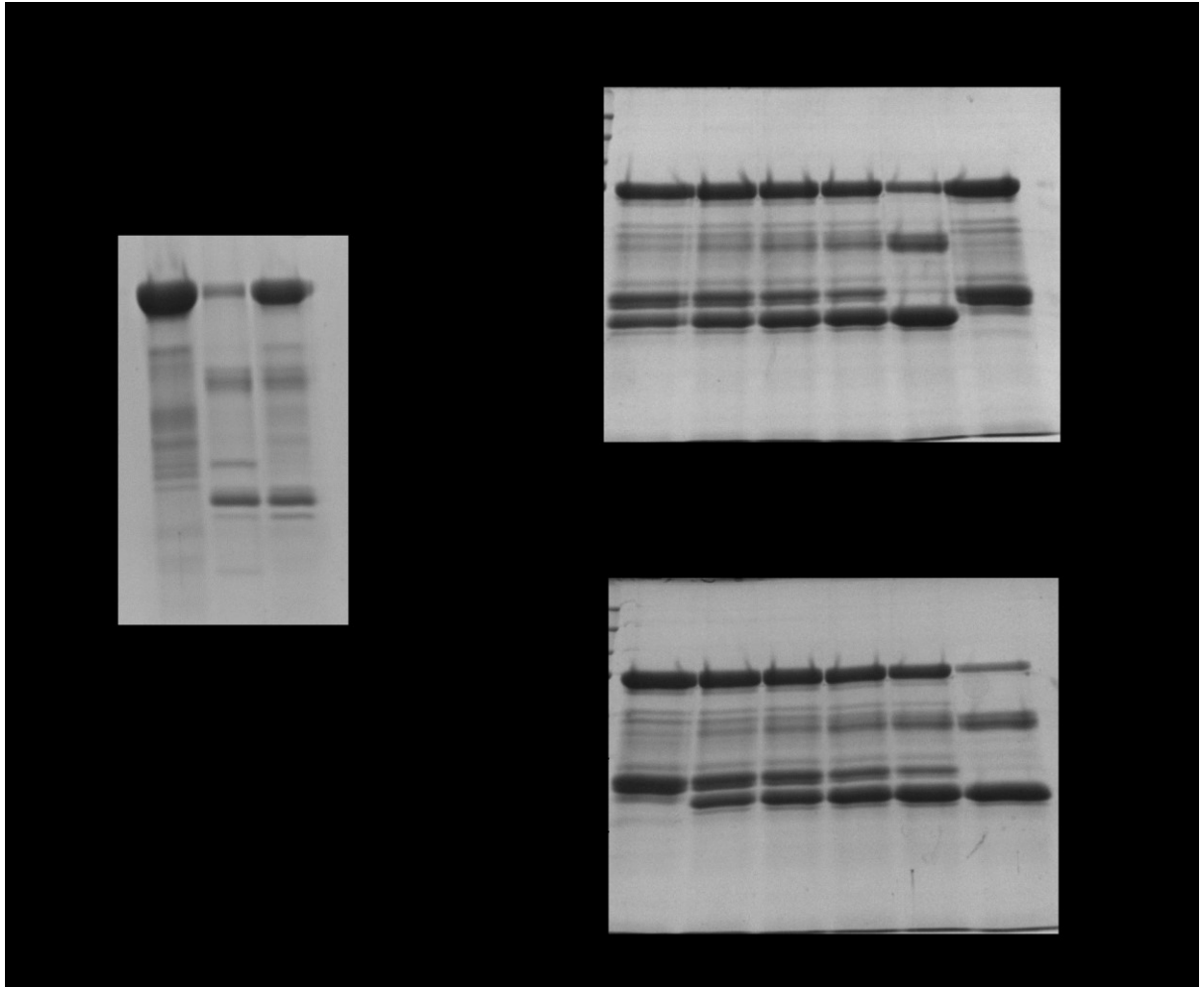


Figure 19 Thrombin cleavage optimisation using Thrombin CleanCleave Kit. (A) Overnight incubation at 4 °C with thrombin resin. (B) and (C) Thrombin cleavage using the resin bound thrombin at 25 °C and 37 °C respectively.

The design and synthesis of a highly engineered and functional OPCML construct

The Thrombin CleanCleave™ Kit manual suggests higher temperatures for optimal cleavage thus the next attempts were undertaken at 25 °C and 37 °C. 40 µL of the thrombin resin was added to 500 µL 0.33 mg/ml OPCML+Fc in wash buffer and shaken for 24 h with aliquots taken at regular time intervals for analysis. **Figure 19A and B** show the second attempt at thrombin cleavage using the resin bound thrombin at 25 °C and 37 °C respectively. There is more complete cleavage at 37 °C but it is possible that keeping OPCML at high temperatures for a prolonged time may adversely affect the protein. In an attempt to optimise the thrombin cleavage further a test reaction with liquid thrombin vs. resin bound thrombin was performed. A reaction that

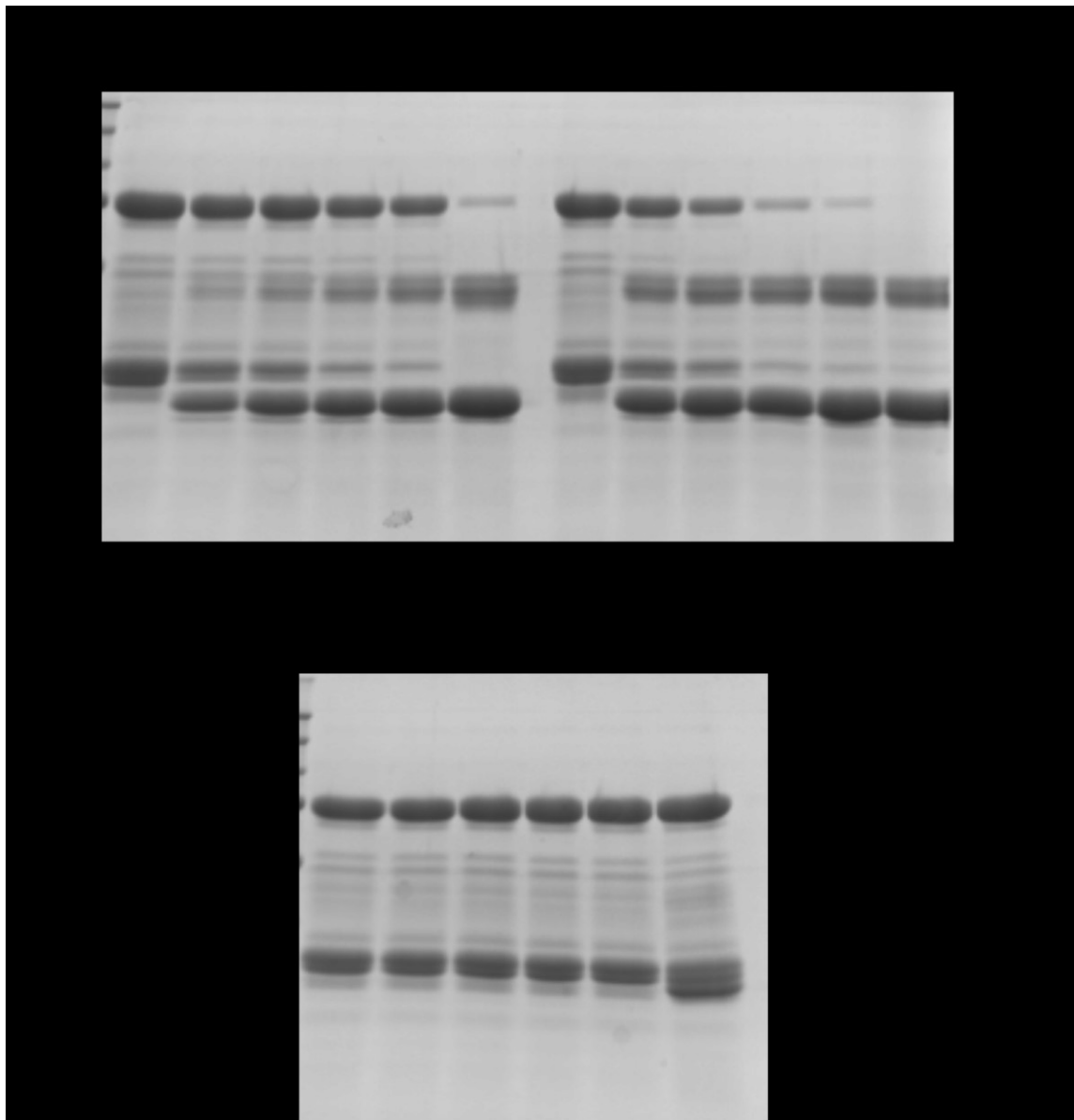


Figure 20 Thrombin cleavage optimisation at 37 °C with immobilised thrombin on agarose resin, 'liquid' thrombin and a rOPCML+Fc control reaction of just buffer and protein shaking at 37 °C.

The design and synthesis of a highly engineered and functional OPCML construct

was just rOPCML+Fc shaking at 37 °C was also carried out as a control to see if any obvious degradation was observed. The amount of thrombin resin was also increased in an attempt to get thrombin cleavage to completion with 40 µL resin incubated with 250 µL rOPCML+Fc (0.33 mg/ml). The liquid thrombin reaction was carried out using 30 µL thrombin (100 units/mL) with 250 µL rOPCML+Fc (0.33 mg/ml).

Figure 20 shows an SDS-PAGE following thrombin cleavage and the control OPCML reaction. There was no complete cleavage of rOPCML+Fc with the increased amount of thrombin agarose resin, whereas with the 'liquid' thrombin complete cleavage occurred, thus for following experiments the liquid thrombin will be used. Furthermore, the double band for Fc appeared in the rOPCML control blot after 24 h which suggests it was due to degradation of the Fc sequence. Moreover, as the reaction takes 24 h to go to completion future thrombin cleavage reactions will be carried out at 4 °C to prevent rOPCML from degrading or becoming inactive.

3.2.3 Anion Exchange to remove impurities after Thrombin Cleavage

The recombinant OPCML can now be successfully cleaved with thrombin however there needs to be a way to remove the thrombin from the cleaved protein. Cleaved rOPCML and thrombin are a similar size thus size exclusion cannot be used. Ion exchange chromatography (IEX) can be used to separate molecules based on their net surface charge. Anion exchange columns bind negatively charged biomolecules. When the pH of the buffer is above a protein's isoelectric point (pI) it will have a negative charge and bind to the column and an increasing salt gradient will elute any bound protein. The pI values of all the molecules to be separated from the cleaved rOPCML are sufficiently different for separation by IEX (**Figure 21A**).

rOPCML+Fc was cleaved with thrombin overnight at 4 °C in wash buffer (50 mM Tris, 150 mM NaCl, pH 7.5). For each cleavage reaction 10 units of thrombin were used per mg of rOPCML+Fc. After the overnight incubation the reaction mixture was diluted with 50 mM Tris to 50 mM NaCl so the injected protein had the same buffer composition as the start buffer. A gradient was run up to a 50 % elution buffer (50 mM Tris, 1 M NaCl) over 50 CV (50 mL).

Figure 21B shows the UV trace from the IEX. Peaks can be seen at elution volumes of 0-10 mL and 65-80 mL. The first peak is the unbound fraction which has a higher pI than that of the starting buffer. The second peak elutes during the salt gradient and is the cleaved OPCML.

The design and synthesis of a highly engineered and functional OPCML construct

Figure 21C confirms the first peak is due to the unbound thrombin and Fc sequence. The latter peak is cleaved OPCML with both running at the expected molecular weight on the gel.

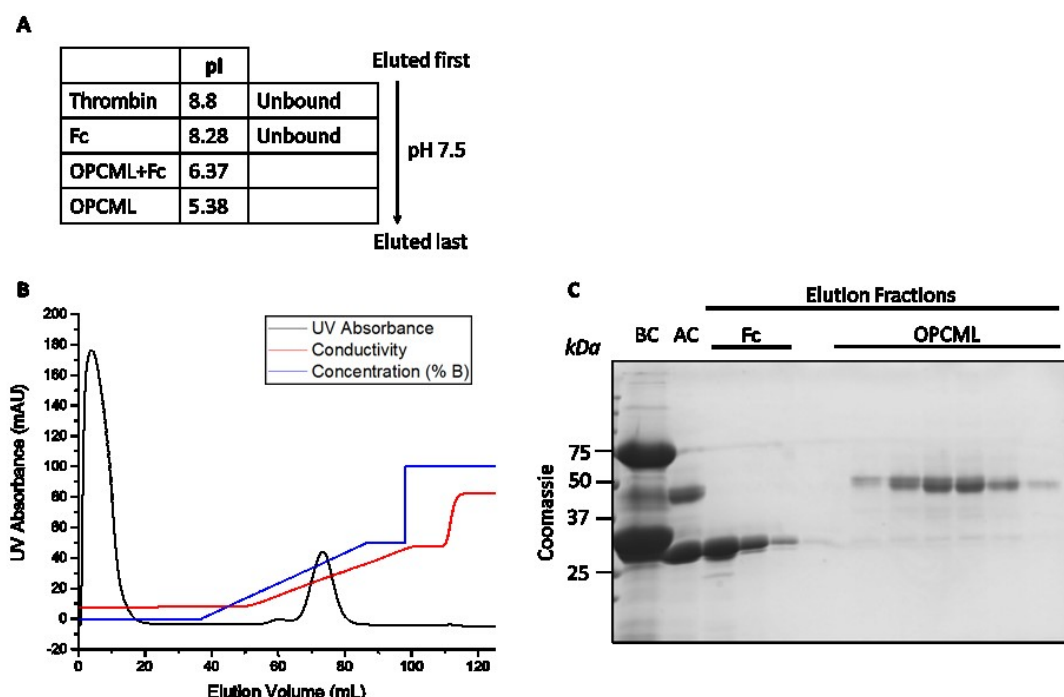


Figure 21 Anion exchange of rOPCML after thrombin cleavage to remove impurities. (A) Table showing the differing pI values for biomolecules in the thrombin cleavage reaction. (B) UV trace from anion exchange column. (C) SDS-PAGE stained with coomassie showing the identity of the peaks eluted from the IEX column. BC before thrombin cleavage. AC after thrombin cleavage.

3.2.4 Size Exclusion Chromatography to confirm aggregation state of rOPCML

After anion exchange to remove the cleaved Fc sequence and thrombin, SEC was used to remove any potential aggregates and remaining contaminants and to approximate the protein molecular weight. However, the shape of a protein can affect how it interacts with the pores of the SEC column and the crystal structure shows that OPCML is a long rod shaped protein thus will run different to globular proteins so any calculated size would be a vast approximation³³. The first attempt at SEC was done with an S75 column which is suitable for proteins between 3 and 70 kDa. The cleaved rOPCML should be approximately 33 kDa so is within this range.

Figure 22A shows the UV trace of the first SEC after thrombin cleavage with protein loaded directly after anion exchange and spin concentration. The protein eluted just after the void

The design and synthesis of a highly engineered and functional OPCML construct

volume of the column (40 mL) which is very early and the peak was not symmetrical and had a slight shoulder. The early elution suggests the protein is larger than expected and should be run on a S200 SEC column which is more suitable for proteins between 10 and 600 kDa. The early elution and the shoulder on the peak could be due to the rod like shape of the protein.

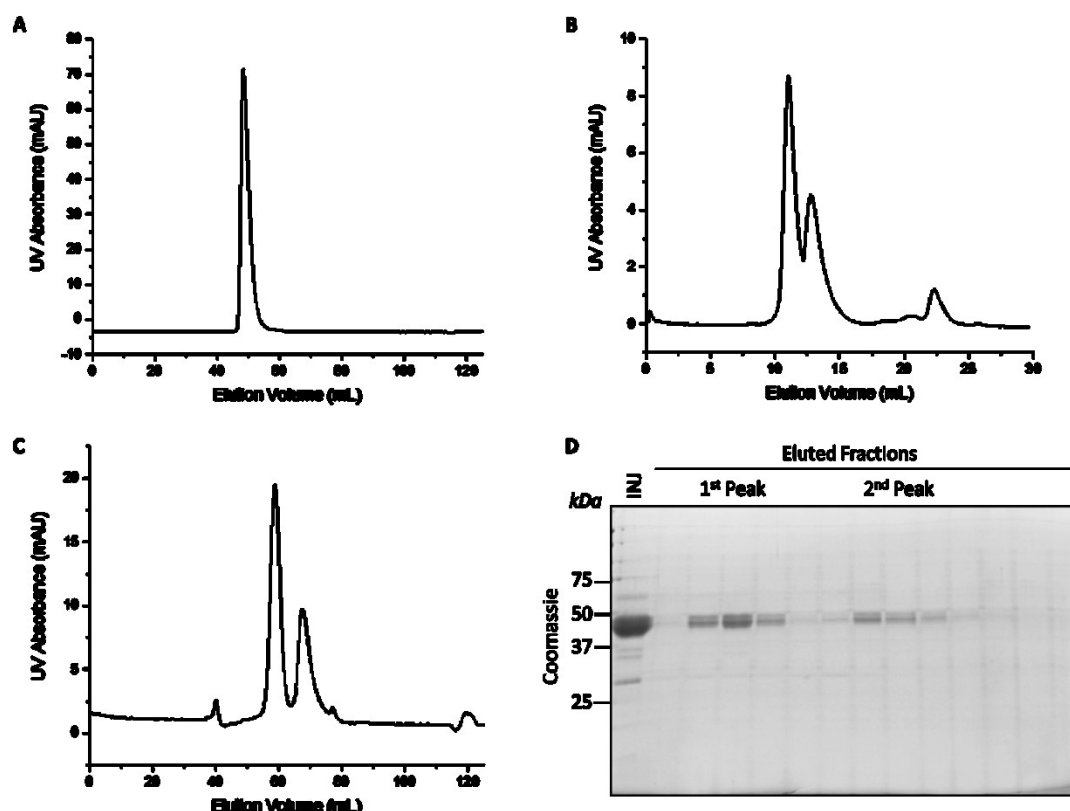


Figure 22 SEC traces of rOPCML (A) S75 16/600 SEC. (B) S200 10/300 SEC. (C) S200 16/600 SEC. (D) SDS-PAGE of S200 16/600 SEC of rOPCML.

700 μ L rOPCML at 0.9 mg/mL was recovered from the S75 column and this was run on a S200 10/300 column and two peaks were observed in the UV trace (**Figure 22B**). This suggest that there are two different species of OPCML present. A S200 16/600 column was then run using the remaining protein to try and isolate and identify the two species. Both peaks were run on an SDS-PAGE and are of the expected size for OPCML.

The rOPCML construct was designed with a free cysteine so Michael Addition could be used for attaching anchor mimetics if no other methods work (section 3.1). It was hypothesised that this free cysteine could be forming disulphide bonds which would account for the two OPCML species observed.

3.2.4.1 Confirming a free cysteine was causing secondary rOPCML species

To confirm whether the free cysteine is forming disulphide bonds rOPCML was separated by SDS-PAGE both with and without prior reduction. Previous gels had all used reducing agent in the loading buffer so the two species resolved into one band. The western blot and coomassie stained SDS-PAGE gel (**Figure 23**) show that without reduction two bands appear at differing sizes for rOPCML. The identity of these bands was further confirmed using PMF analysis: The bands from 1st and 2nd peak lanes were taken as well as the two bands from the last lane of the coomassie stained gel and were digested with trypsin before all being identified as OPCML.

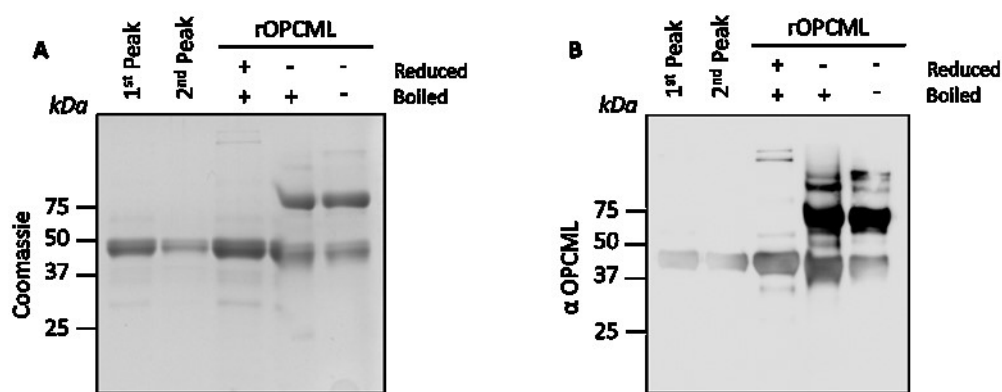


Figure 23 SDS-PAGE of rOPCML with and without reducing agent. (A) SDS-PAGE stained with coomassie. (B) Western blot.

To further confirm the free cysteine as the cause of the multiple species of rOPCML the protein was reduced with DTT to make the free cysteine accessible then the protein was reacted with iodoacetamide to cap the free cysteine. Once capped the 0.1 mg of protein in 250 μ L was run on an S200 10/300 size exclusion column to see if the two peaks resolved into one. **Figure 24** is the SEC after capping and shows that capping does prevent the presence of multiple OPCML species.

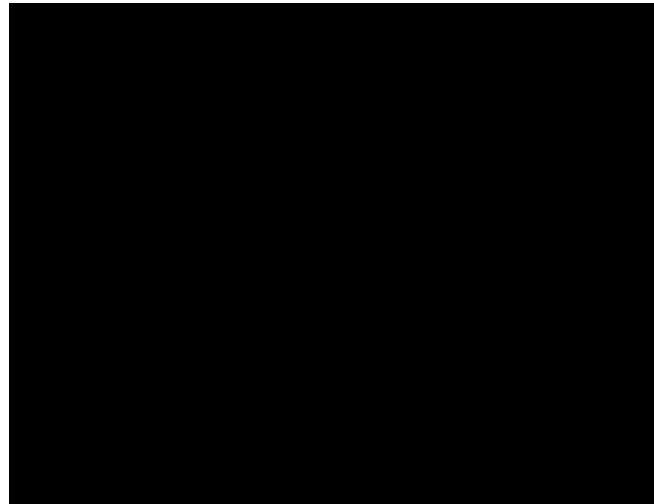


Figure 24 SEC S200 10/300 of rOPCML after reducing and capping.

The presence of a free cysteine, which is binding to other rOPCML molecules, could cause the efficiency of the sortase reaction to decrease by make the sortase recognition site inaccessible. Therefore, it was decided to use site directed mutagenesis to remove the free cysteine and use rOPCML without a free cysteine for all future experiments.

3.3 Generating a rOPCML+Fc construct with no free cysteine to prevent aggregation of the protein in future research

The free cysteine was removed from the OPCML protein to prevent protein aggregation and any problems that may arise due to the inaccessibility of the sortase recognition site. The Q5® Site-Directed mutagenesis kit from NEB was used for the deletion of the free cysteine following the manufacturer's instructions. Briefly primers were designed (Appendix B) that flanked either side of the cysteine to be removed. These were amplified by PCR using a Q5® polymerase with an annealing temperature of 55 °C, the PCR product (**Figure 25A**) was then used directly in a KLD (kinase, ligase and DpnI) reaction and incubated for an hour before being transformed into competent E.coli.

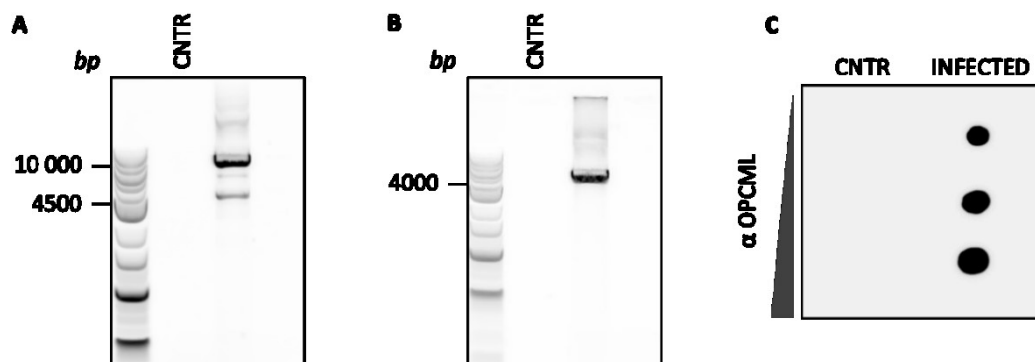


Figure 25 (A) Agarose gel electrophoresis of PCR product of sample and control after amplifying with primers to remove free cysteine. (B) Agarose gel electrophoresis of recombinant bacmid. (C) Dot blot to show expression of rOPCML-C295del in insect cells after 72 hours.

Recombinant bacmid was generated as before (3.1.2) and the production of the bacmid confirmed by agarose gel electrophoresis. The size of the bacmid alone is 300 bp and the size of the OPCML insert is 1662 bp. The size of the successfully transposed bacmid is approximately 3962 bp and a band this size can be seen on **Figure 25B** confirming the bacmid was successfully transposed. The expression was tested (**Figure 25C**) on a small scale to ensure rOPCML-C295del production was occurring before larger scale expressions were undertaken.

The production of the protein was carried out as before using the baculovirus expression system in insect cells followed by harvesting the supernatant. The protein was isolated using affinity chromatography with Protein A Agarose which binds to the Fc sequence still present (**Figure 26**).

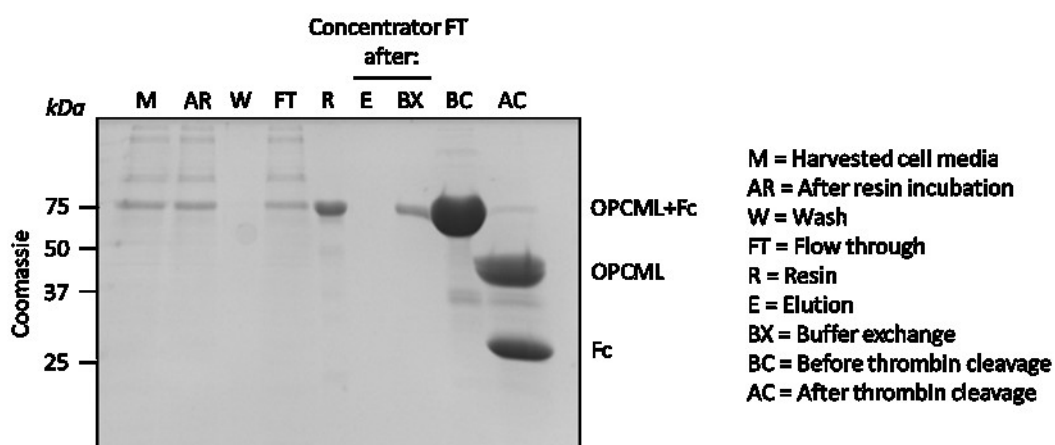


Figure 26 SDS-PAGE stained with coomassie showing stages of OPCML protein purification.

The design and synthesis of a highly engineered and functional OPCML construct

Following this thrombin cleavage was carried out as before and anion exchange used to purify the cleaved rOPCML-C295del from the Fc and thrombin (**Figure 27A and B**). Finally SEC was performed. Pleasingly, one symmetrical peak in the UV trace was observed indicative of a single species of protein (**Figure 27C**).

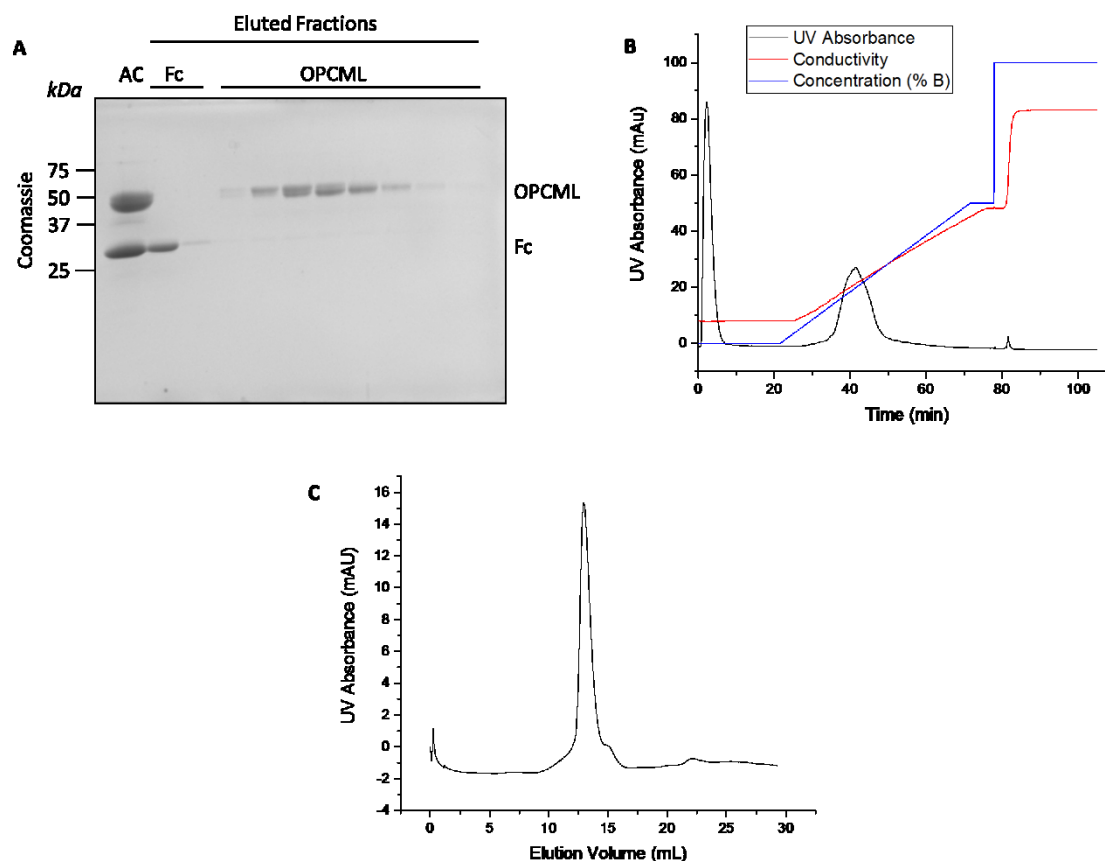


Figure 27 (A) SDS-PAGE of Anion exchange ANX FF column stained with coomassie, AC = after thrombin cleavage. (B) UV, conductivity and concentration trace of anion exchange. (C) SEC S200 10/300 rOPCML-C295del.

3.4 Generating the inactive R65L mutant construct for rOPCML+Fc as a control for functional assays

The R65L mutation causes loss of function via the disruption of homodimerisation and inability to bind to partner RTKs³³. For functional assays having an inactive mutant as a control would provide a more complete picture of the function of OPCML, therefore the R65L mutant was

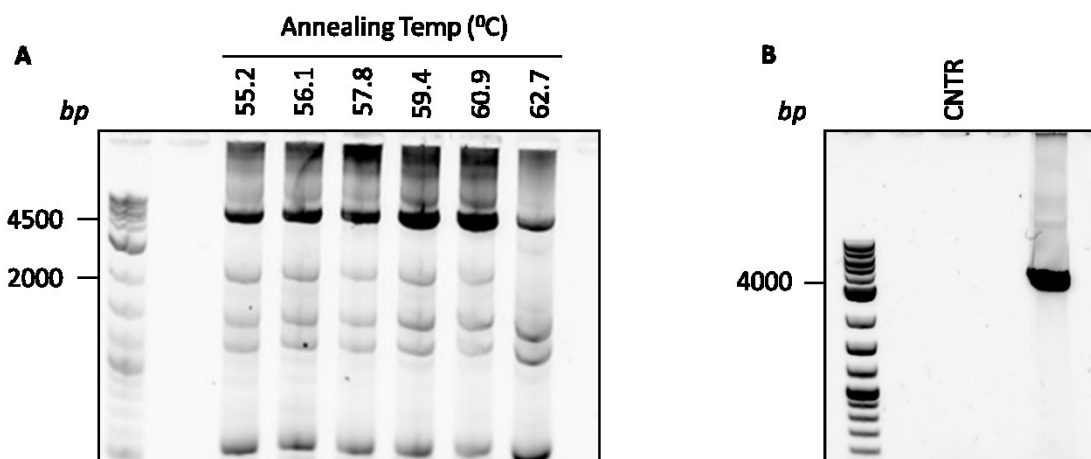


Figure 28 (A) Agarose gel electrophoresis of PCR product of sample and control after amplifying with primers to remove free cysteine. (B) Agarose gel electrophoresis to analyse production of recombinant bacmid.

made from the previously designed construct (**Figure 15A**). In this construct due to the secretion signal being one suitable for insect cell expression, and not the native mammalian secretion signal, the R65L mutation is actually R56L. The mutation was carried out using the Q5® Site-Directed mutagenesis kit from NEB as previously described (section 3.3). Briefly, primers were designed with one as a mutagenic primer containing the substitution required in the centre of the primer (5'- CCGGGTAACCTTGGTGGCCTGGC -3'). The second primer is entirely complimentary to the sequence and begins at the base next to the 5' end of the first mutagenic primer (5'- TCATCTATGGTACACCTGAGG -3'). A gradient PCR was undertaken and the PCR product used directly in a reaction prior to transformation (**Figure 28A**).

The baculovirus expression system was used to produce the protein and formation of the bacmid confirmed by agarose gel electrophoresis (**Figure 28B**). The size of the bacmid alone is 300 bp, the size of the OPCML insert is 1662 bp thus the size of the successfully transposed

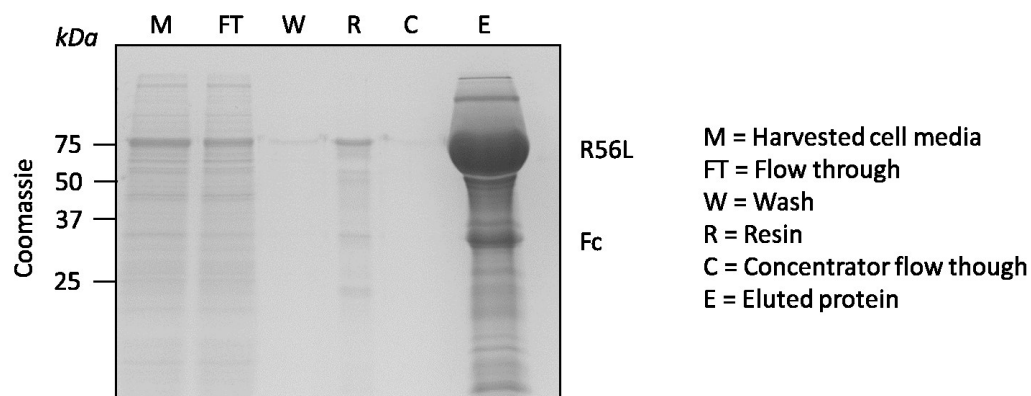


Figure 29 SDS-PAGE stained with coomassie showing stages of R56L mutant protein purification.

The design and synthesis of a highly engineered and functional OPCML construct

bacmid is approximately 3962 bp and a band of this size can be seen on **Figure 28B** confirming the bacmid is successfully transposed. The expression was tested on a small scale to ensure rOPCML-R56L production and the protein was shown to express after 72 h. **Figure 29** shows the mutant protein was successfully expressed and isolated using Protein A Agarose. Following this thrombin cleavage was carried out as previously and anion exchange used to purify the cleaved mutant protein from the Fc and thrombin (**Figure 30A and B**). SEC was also carried out to determine if the protein was a single species or aggregated (**Figure 30C**).

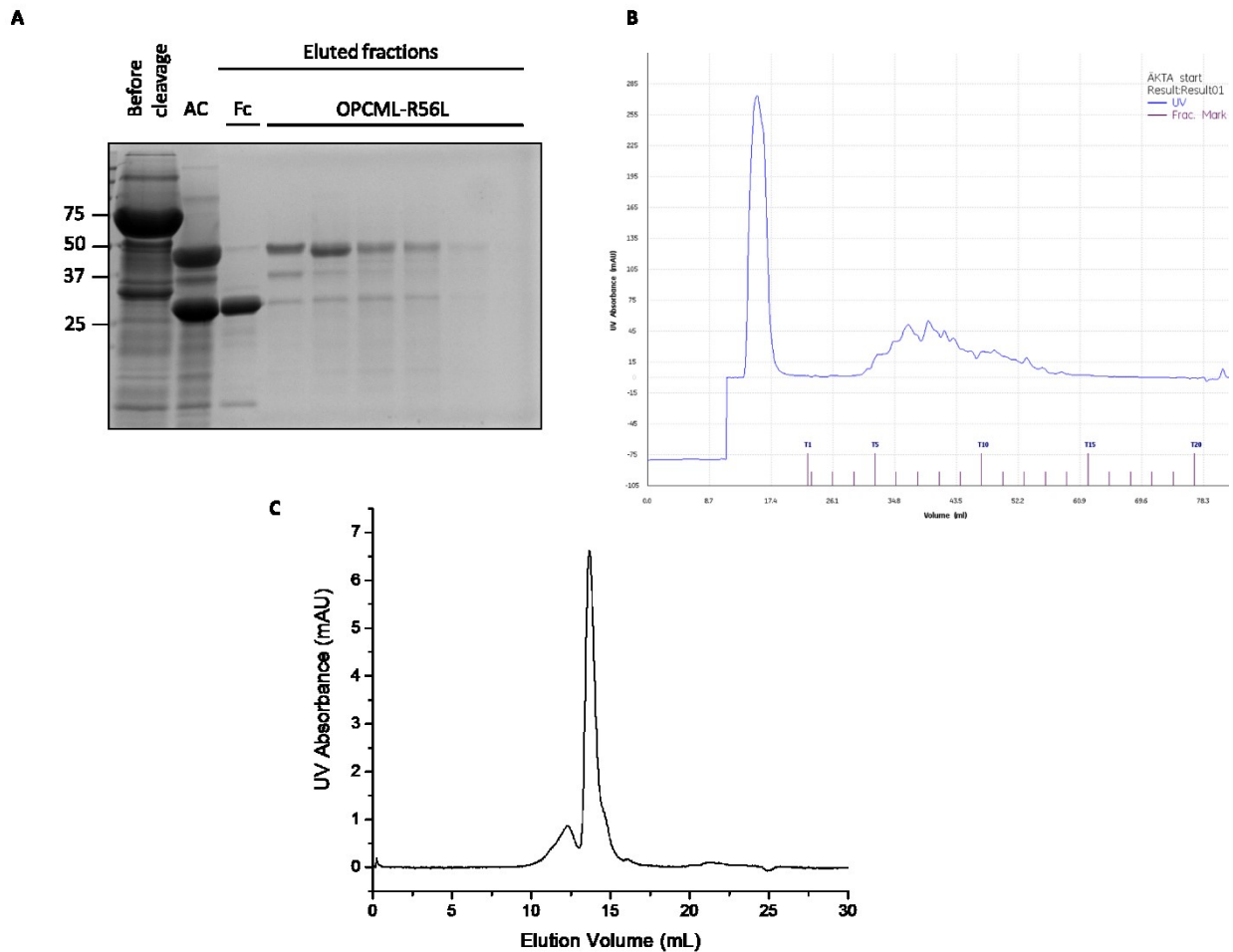


Figure 30 (A) SDS-PAGE of Anion exchange ANX FF column stained with coomassie, AC = after thrombin cleavage. (B) UV trace of anion exchange. (C) SEC S200 10/300 rOPCML-R56L..

The SEC trace shows the presence of two peaks in the UV trace this suggest there could be a small impurity present after the thrombin cleavage and anion exchange or that the protein is not a single species. However, as there are a few bands present in the eluted fractions in **Figure 30A** after anion exchange it is most likely that this is the cause of the second peak. Therefore after anion exchange SEC should be undertaken to remove any further impurities and give a pure protein.

3.5 Testing the function of rOPCML

The OPCML construct designed and synthesised is highly engineered and capable of attachment of complex chemical prosthetic groups to mimic the GPI anchor. However, the modifications made to the protein must still enable the protein to function as a tumour suppressor. Therefore, it is important to test the function of this novel protein construct. Birtley et al³³ have demonstrated the tumour suppressor function of OPCML through spheroid invasion assays with OPCML reducing the amount of spheroid invasion when compared to controls.

Spheroids are heterogeneous aggregates of cells with hypoxic regions and necrotic centres^{138,139}. Spheroid invasion assays enable treatments to be tested in 3D environment, which is useful as the assay can mimic the tumour microenvironment and thus have been used in cancer research^{140–144}. **Figure 31A** shows a schematic of the stages of a spheroid invasion assay. Firstly the chosen cells are seeded in an ultra-low adherence 96 well plate. Over time these cells aggregate to form a spheroid. Once the spheroid has reached the desired diameter the spheroid is treated with compound and a layer of Matrigel® placed on top. Corning® Matrigel® matrix is an extracellular matrix and helps mimic the *in vivo* environment for cell culture. The spheroid is then imaged over time for invasion. The invasion occurs through actin-rich protrusions from the plasma membrane, known as invadopodia, that are due to degradation of the extracellular matrix and are associated with cancer progression^{140,145,146}.

As mentioned in the introduction OPCML is inactivated via methylation in numerous cancers^{5,31,35–44}. OPCML expression is reduced or absent in 83 % of brain tumours and is not present in U87MG cells⁴⁰. U87MG cells are a glioblastoma cell line that has been shown to invade well during spheroid invasion assays¹⁴⁰. Therefore, due to the robustness of U87MG spheroid invasion assays this cell line was used to test the function of rOPCML. There were two buffer treated controls; one which was serum starved overnight like the treated spheroids, and one control that was left in full serum media for the duration of the experiment. The results show that the control spheroids increased the most in volume and the spheroids treated with rOPCML had a much lesser increase in volume. This preliminary experiment suggests that the rOPCML is functional and is preventing invasion of the U87MG cells.

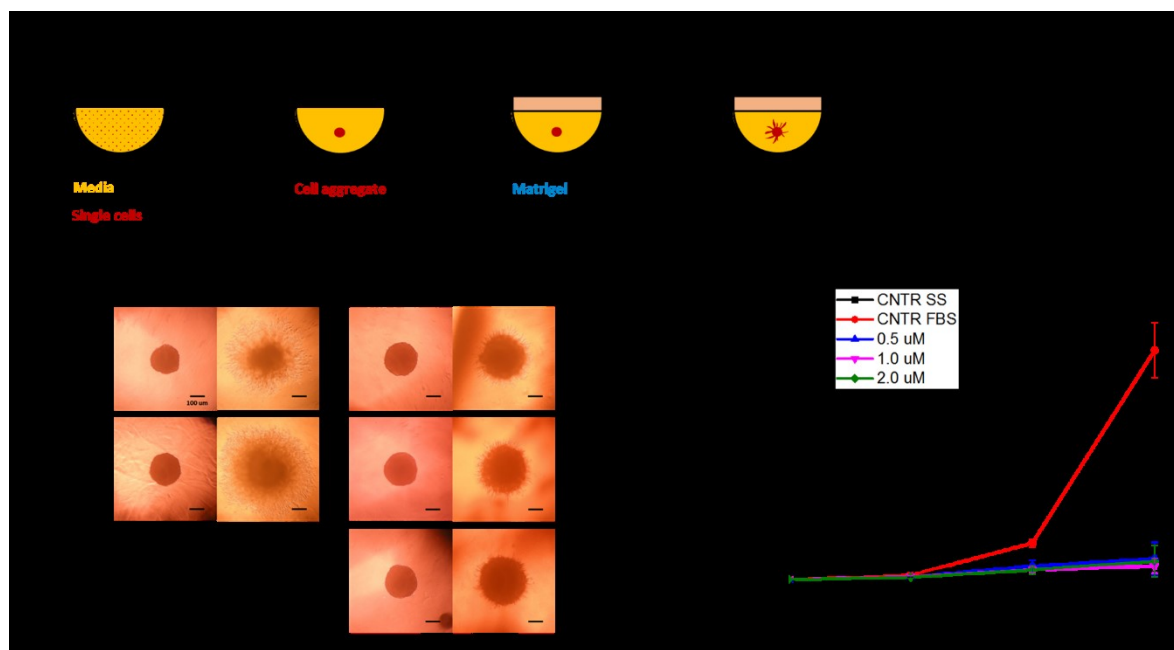


Figure 31 (A) Schematic showing stages of the spheroid invasion assay. (B) Images of U87MG cells before and after treatment with rOPCML.

3.6 Conclusion

The overall objective of this research is to design a novel bio-therapeutic based on OPCML with a GPI anchor mimetic and to validate this protein anchor complex as a potential drug in EOC. To do this it was necessary to design and synthesise a highly engineered and functional OPCML construct capable of attachment of complex chemical prosthetic groups to mimic the GPI anchor. rOPCML was successfully expressed using both a mammalian expression system and a baculovirus expression system. The construct design was influenced by one reported by Birtley et al³³ with added functionality for two methods of anchor attachment, and a tag for visualisation. Optimisation of cleavage conditions to remove the affinity sequence was undertaken followed by further purification to remove impurities. SEC was carried out to further characterise and purify the protein. SEC highlighted multiple species of rOPCML which proved to be due to the presence of a free cysteine added for anchor attachment via Michael Addition. Thus, a further construct was made with this cysteine removed to give a single rOPCML species for all future uses. Finally, a mutant of the construct was made (also without a free cysteine) to use as an inactive control based on the loss-of-function R56L mutant reported by Birtley et al³³.

Chapter 4

Functionalising rOPCML with a GPI anchor mimetic

OPCML is localised to the DRM of cells by a GPI anchor. In the DRM OPCML functions as a tumour suppressor by sequestering RTKs to the DRM and arresting their signalling cascade^{32,52–54}. Therefore, the GPI anchor is important for localising OPCML to the cell membrane where it can function as a tumour suppressor. There are various methods that can be used for the ligation of anchors to the recombinant protein as mentioned in the Introduction. One method is including a free cysteine in the protein enabling Michael Addition with an anchor containing a thiol reactive maleimide. However, this method was not pursued due to protein aggregation upon inclusion of a free cysteine (see section 3.2.4.1) ⁷⁵. Alternative approaches of anchor attachment include using a kit such as The IMPACT™ kit (New England Biolabs) to do Intein-mediated Protein Ligation (IPL) (**Figure 32**)^{88–92}, or sortase-mediated transpeptidation (sortagging)^{105,111}. These latter two approaches were attempted and will be discussed in this chapter. Some of the work presented in this chapter has been previously reported¹²¹.

4.1 Intein-mediated Protein Ligation

The IMPACT™ kit involves the fusing of proteins to an intein tag and a chitin binding domain (CBD), then purifying the protein by affinity chromatography with a chitin column. When 2-mercaptoethanesulfonic acid (MESNA) is used to elute the bound protein a C-terminal thioester is formed. The modified protein is then amenable for ligation to a GPI anchor mimetic provided there is an N- terminal cysteine^{88–92}.

The first strategy for anchor attachment to be explored was IPL using the IMPACT™ kit (New England Biolabs). The first step in IPL is to clone the protein construct of interest into a pTXB1 vector. This vector is used for producing the rOPCML as it allows fusion of the intein tag to the C-terminal of OPCML at the site where the native GPI anchor would be located. Digestion and ligation of OPCML into the pTXB1 vector was done using the restriction enzymes NdeI and SapI. At this stage of experiments the crystal structure of OPCML was not yet known therefore IPL was initially attempted with full length OPCML (FL) and just domain 3 (D3) to test whether all the protein is needed for function or just a single domain would suffice. The OPCML construct

was generated following the instructions in the IMPACT™ manual. Briefly the primers were designed to add NdeI and SapI restriction sites into the OPCML construct as these would be used for ligation into the pTXB1 vector (see Methods). The SapI primer is designed with an overhang of GCA as when ligated with the pTXB1 vector this creates an in-frame fusion to the N-terminus of an intein. The NdeI primer is designed such that the NdeI site directly precedes the start of the protein of interest. The SapI primer precedes the site at which the native GPI anchor would attach. The plasmids used in the PCR reaction were those designed for lentiviral production from Geneart™ and were used as delivered (section 2.1).

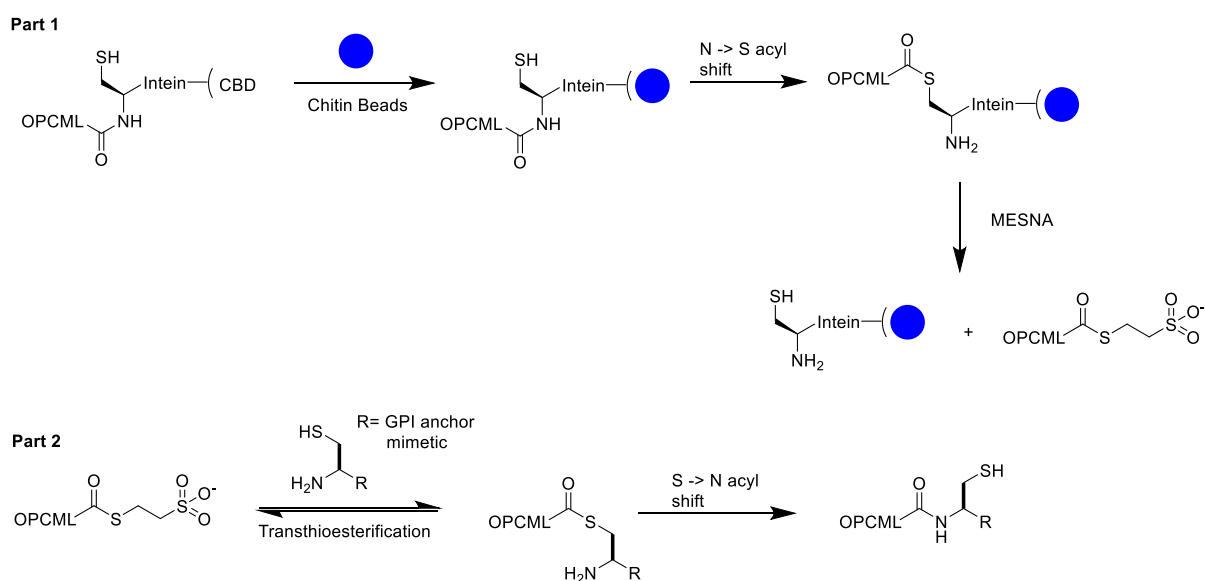


Figure 32 Schemes for the ligation of recombinant OPCML protein to the GPI anchor mimetic. Part 1) p TXB1-OPCML is expressed with an intein containing a chitin binding domain (CBD), first the CBD binds to a chitin column, and an S N shift will occur. MESNA is used to cleave the protein from the column and leaves the OPCML protein with reactive thioester⁹². Part 2) The C-terminal thioester of OPCML will undergo transesterification with a terminal cysteine residue in the GPI anchor mimetic. An S N acyl shift then occurs leading to a protein-GPI anchor complex⁹². Image taken from initial research plan ¹⁴⁷.

After confirming the PCR had been successful via agarose gel electrophoresis (**Figure 33A**) the DNA was purified using a Gel Extraction Kit (Qiagen) following the manufacturer's instructions. The pTXB1 vector was digested for 4 h at 37 °C, Shrimp Alkaline Phosphatase was added to prevent ligation of the linearised DNA plasmid and subsequently removed by heat activation at 65 °C for 20 min. Ready-To-Go T4 DNA Ligase (GE Healthcare, Little Chalfont, UK) was used to ligate the OPCML construct into the pTXB1 vector, and the success of the ligation confirmed by the presence of a smaller band in DNA gel electrophoresis corresponding to the size of the insert when of the pTXB1-FL and pTXB1-D3 were digested (**Figure 33C**). The FL insert is 885 bp and the D3 insert is 378 bp. The correct ligation of pTXB1 FL and D3 plasmids was also confirmed by sequencing.

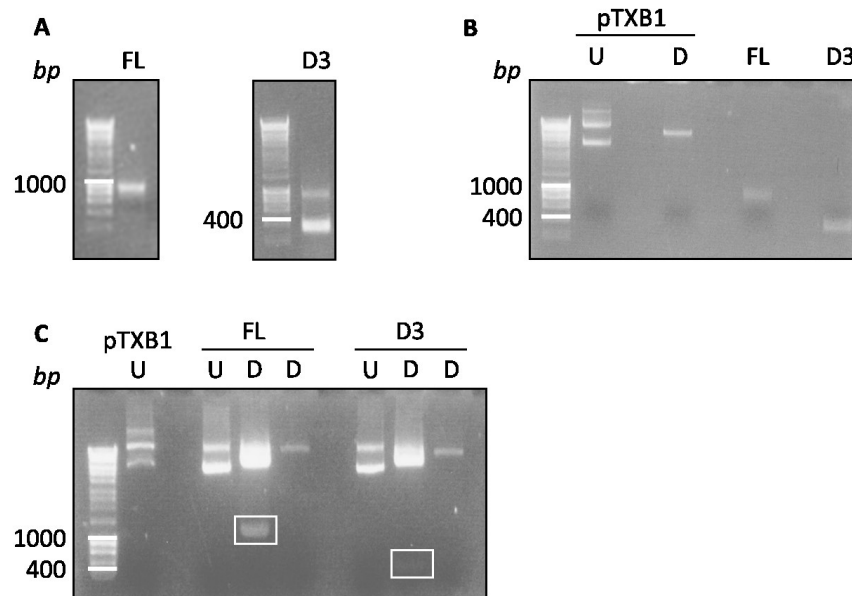


Figure 33 Agarose gel electrophoresis confirming successful generation of intein DNA constructs U= undigested vector, D=digested vector with *NdeI* and *SapI*. (A) PCR with primers to introduce restriction enzyme sites. (B) Gel indicating pTXB1 fully digested and FL and D3 inserts contain no impurities prior to ligation. (C) Ligation of FL and D3 into pTXB1 with the digested FL and D3 in pTXB1 showing a smaller band at 1000 bp and 400 bp respectively. Image taken from report¹²¹.

After cloning had been carried out the pTXB1-FL and pTXB1-D3 constructs were expressed in competent cells (BL21 competent cells). Overnight starter cultures were grown at 37 °C and 20 mL of the starter culture used to inoculate larger 1 L cultures. Once the larger culture had reached an OD₆₀₀ between 0.5 and 0.7 IPTG (final concentration 0.5 mM) was added to induce protein expression. After induction, to determine the optimum conditions for protein expression, one flask was incubated at 37 °C with samples taken for SDS-PAGE analysis every hour for four hours and one flask incubated at 18 °C overnight (18 h). **Figure 34A** and **B** show the expression of D3 and FL respectively. These coomassie stained SDS-PAGE gels show the D3 protein construct is clearly expressed under both conditions but the expression is strongest after 4 h at 37 °C. The FL construct is not as strongly expressed as the D3 with a faint band appearing between 50 and 75 kDa corresponding to the construct. However, the western blot for FL expression confirms protein is produced successfully after 4 h at 37 °C. Therefore, for both pTXB1 constructs it was decided to express the proteins at 37 °C.

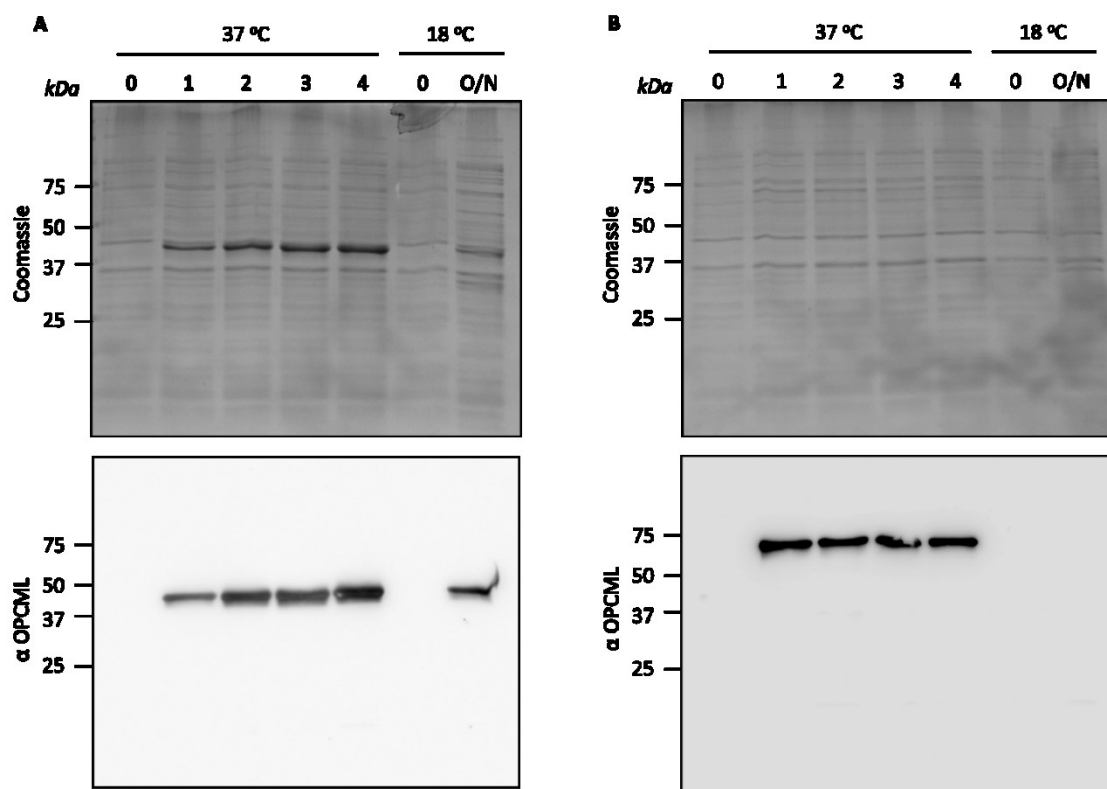


Figure 34 SDS-PAGE to confirm optimum expression of D3 and FL pTXB1 constructs. (A) D3 coomassie and Western blot. (B) FL coomassie and Western blot.

The pTXB1 protein constructs were also tested for solubility: After expression was induced, the cells were grown for 4 h at 37 °C and pelleted. The pellet was then re-suspended in lysis buffer (pH 8.5, 20 mM HEPES, 500 mM NaCl, 1.0 mM EDTA, 0.2 % TWEEN and protease inhibitor), sonicated to lyse, and the lysate centrifuged to give the soluble and insoluble fractions. Samples for insoluble and soluble fractions were then run on SDS-PAGE gels and visualised by western blot. **Figure 35A** shows that the FL and D3 construct proteins both have protein present in soluble (S) and insoluble (IN) fractions. As for both constructs there is protein in the soluble fraction it was decided to continue with IPL using the soluble proteins.

The pTXB1 vector enables the protein to be expressed with an intein-containing chitin binding domain (CBD). As mentioned previously the protein can be purified by affinity chromatography with a chitin column and eluted with MESNA forming a C-terminal thioester on the protein. The first step was to bind the CBD to the beads in the column using column buffer (pH 8.5, 20 mM HEPES, 500 mM NaCl, 1.0 mM EDTA), after which an S N shift occurs. Cleavage buffer containing MESNA (0.1 M Tris, pH 8.5, 10 mM MESNA) was then used to cleave the OPCML-Intein complex from the column and give the OPCML protein with reactive thioester (

Figure 32)⁹³.

Figure 35B shows the purification of pTXB1-FL and pTXB1-D3. For both constructs there was no elution of the protein from the column with all the protein remaining bound to the chitin beads. The intein tag is 28 kDa and the CBD is 6 kDa thus the multiple bands present in the FL construct western blot could be due to the FL construct with and without the intein tag and CBD. The lack of any proteins in the elution is most likely down to the protein with intein tag binding to the column after elution. The IMPACT manual suggests that this could be resolved by longer incubation with cleavage buffer at higher temperatures. Another hypothesis is that the protein-intein fusion may be poorly folded causing less efficient cleavage by MESNA. However it was decided not to continue with IPL as, as well as poor expression, increasing the temperatures could degrade the protein and expressing the proteins in bacteria means there are no post-translational modifications. Furthermore, functional rOPCML was then shown to be expressed successfully in insect cells with a high yield therefore this method of production was adopted and bacterial expression was no longer pursued³³.

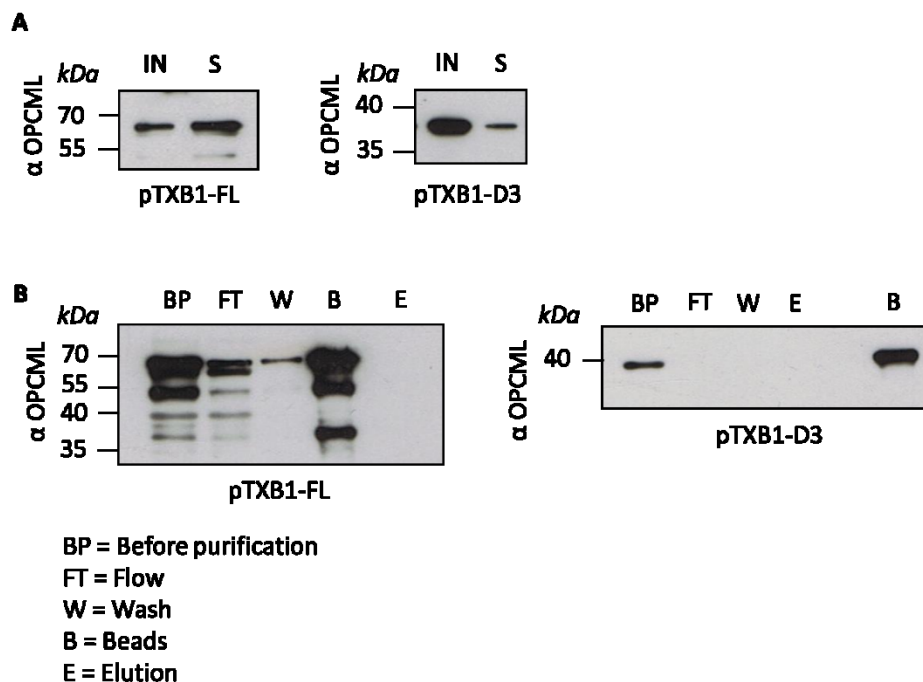


Figure 35 (A) pTXB1-FL and pTXB1-D3 both have OPCML protein present in soluble (S) and insoluble (IN) fractions (B) Stages of purification of the pTXB1 protein constructs. Image taken from report¹²¹.

4.2 Sortase Mediated Transpeptidation – Sortagging

Sortases are transpeptidases that bind proteins to the cell wall in gram-positive bacteria via the recognition of the peptide sequence LPXTG (where X is any amino acid). The sortase recognises the LPXTG site, cleaves between the threonine and glycine leaving an acyl-enzyme intermediate where a lipid nucleophile can attack^{105,111}. Including the LPXTG recognition sequence at the C-terminus of a protein enables the targeted addition of an anchor with N-terminal glycines to the protein. The production of a new protein construct suitable for sortagging is discussed in section 3.1.

4.2.1 GPI Anchor Mimetic Design Considerations and Synthesis

It has been shown that OPCML sequesters a subset of RTKs to the detergent resistant membrane (DRM) abrogating their signalling and preventing cell migration and proliferation^{32,54}. It is the GPI anchor that localises OPCML to the DRM³². To ensure any OPCML biotherapeutic made is targeted to the membrane of cancer cells, where it is found natively, a GPI anchor mimetic must be designed and attached to the recombinant OPCML. The GPI mimetic needs to ensure OPCML is located on the cell surface. Furthermore, as the OPCML-GPI anchor complex will need to be visualised, clickable handles such as alkyne and azide moieties can be incorporated into the design to allow for attachment of fluorescent markers. The anchor also needs to contain the functional moiety required for ligation to the protein.

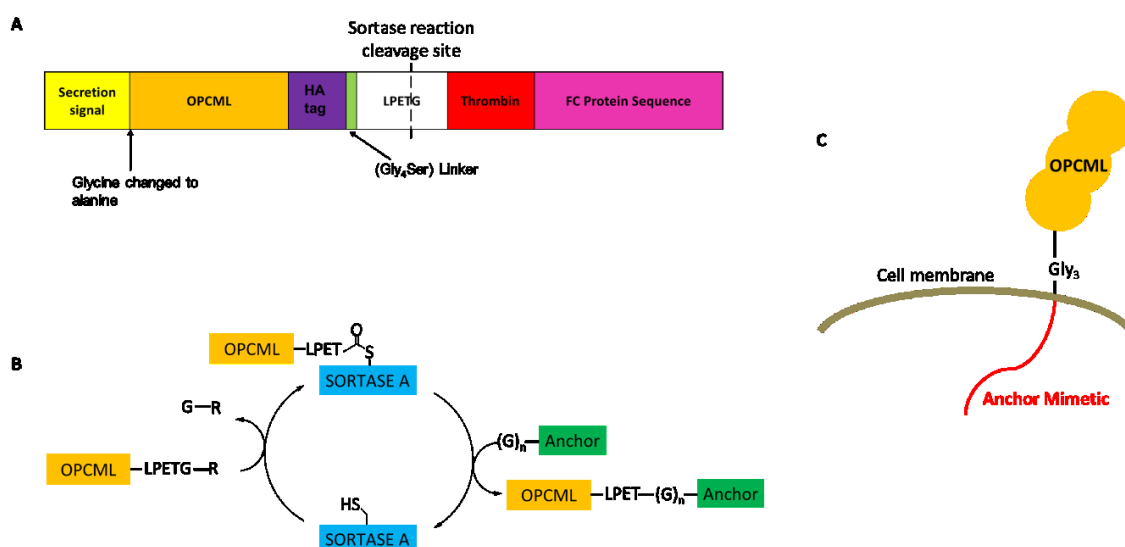


Figure 36 (A) Schematic of construct designed for sortagging³³. (B) Schematic showing process of sortagging¹⁷³. (C) Schematic demonstrating how OPCML will be anchored to the cell membrane⁷³.

Functionalising rOPCML with a GPI anchor mimetic

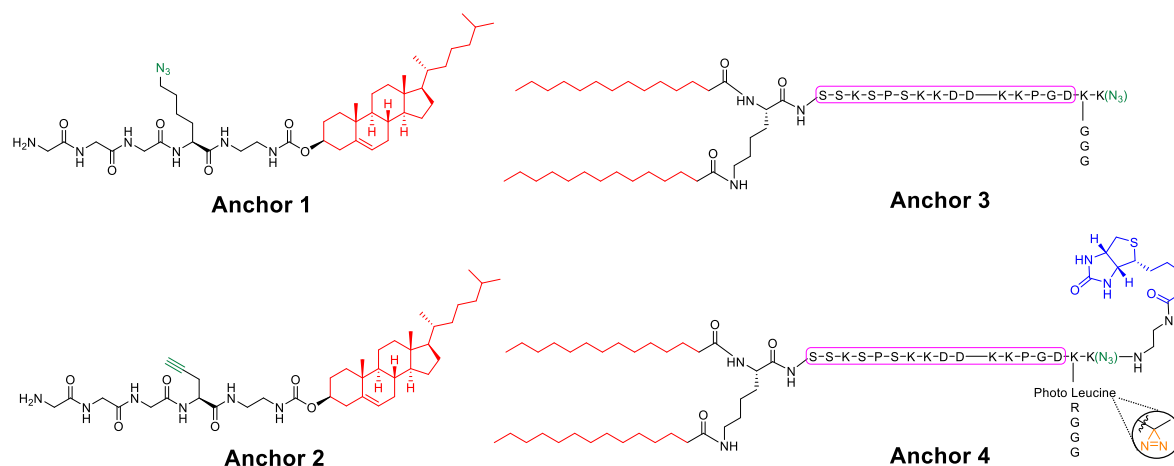


Figure 37 GPI anchor mimetic designs. Red indicates group for anchoring within cell membrane. Green are groups that can be used for CuAAC. The blue moiety is biotin which can be used for enrichment and visualisation. The pink box outlines the hydrophilic cytotopic peptide which will associate with the cell membrane. The orange group is a diazine photo reactive crosslinker.

4.2.1.1 GPI Anchor Mimetic based on Cholesterol Localisation to the Cell Membrane

When designing a GPI anchor mimetic the need for localisation to the cell surface must be taken into account. The current model suggests the DRM is a cholesterol rich lipid domain⁵². As such to encourage recombinant anchored OPCML to localise to this area a cholesterol moiety would be ideal as it would direct the anchored protein to these domains, as suggested previously by Schäfer et al⁴¹. The first anchor design (**anchor 1, Figure 39**) was based upon a cholesterol moiety to direct the protein to the DRM, and a peptide suitable to act as a nucleophile for Sortagging.

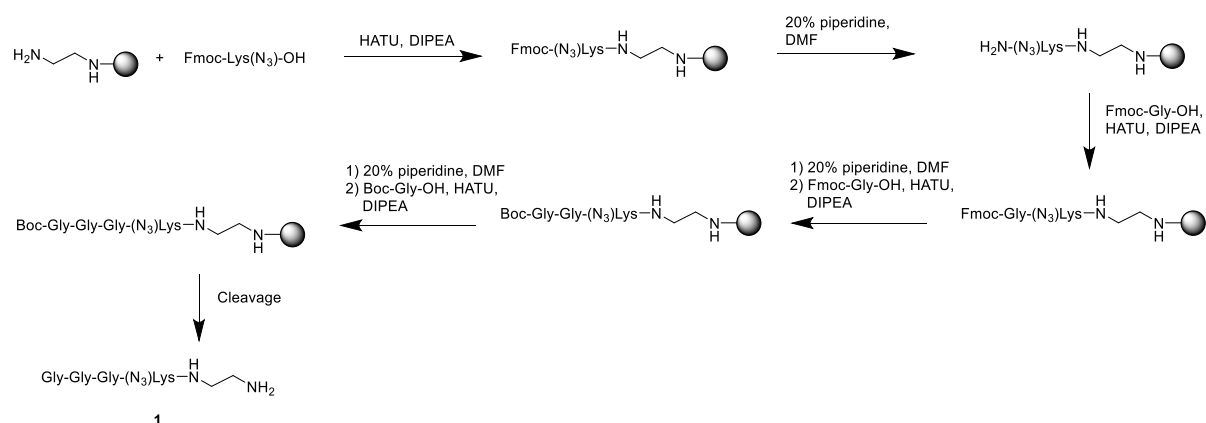


Figure 38 Reaction scheme to form compound 1.

Figure 38 shows the scheme for the first step in the synthesis of the anchor. Solid phase peptide synthesis (SPPS) was used to make the peptide moiety of the anchor (**compound 1**). Ethylene diamine titryl resin was used so after cleavage from the resin the free amine could be coupled to cholesteryl chloroformate. The diamine acts as a spacer between the cholesterol and the peptide. The three glycines enable nucleophilic attack of the acyl-enzyme intermediate during the sortagging reaction for anchor attachment (1.3.3). Three glycines are used as although one glycine may be enough for nucleophilic attack having more has been shown to increase the efficiency of the reaction^{104,105}. An azido lysine was included to allow further functionalisation of the anchor with fluorescent moieties or biotin via click chemistry. After SPPS the peptide was cleaved from the resin with HFIP in DCM (3:7) for 1 h at RT and dried. The crude peptide was then reacted with cholesteryl chloroformate overnight at RT. Briefly the peptide was dissolved in 1 mL of dry DCM and DMF, to this DIPEA (650 μ L, 3.9 mmol, 10 eq) and cholesteryl chloroformate (505 mg, 1.12 mmol, 3 eq.) were added. The reaction was left at RT overnight. Following this the product was washed with water, 5% lithium chloride and brine. The product was then dried with magnesium sulphate and the DCM removed by rotary evaporation. The resulting product was then purified using a Biotage Isolera system equipped with Biotage SNAP KP-sil flash purification cartridge. After purification the yield was low (not enough for mass measurements) and NMR showed none of the characteristics peaks expected for cholesterol (such as C-H₃, C-H₂ and C-H around 0.6-1.8 ppm). Furthermore none of the hydrogens expected from the ethylene diamine or peptide were visible. This lead to the conclusion that the cholesterol may have stuck to the column used during purification or the cleavage of the peptide may not have occurred successfully.

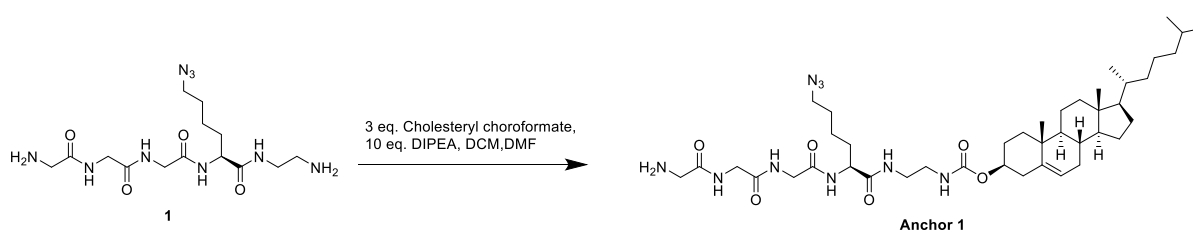


Figure 39 Reaction scheme to form Anchor 1.

A second attempt to synthesis anchor 1 was carried out following the same reaction scheme with a change of cleavage cocktail to 1 % TFA in DCM for 1 hour at RT. The coupling was carried out and the product purified as before with two products isolated in low yield (not enough for mass measurements). A mini cleavage was carried out on the resin left from the earlier cleavage

and this showed that the peptide had not been successfully cleaved from the resin prior to coupling. Therefore, the next cleavage attempt was undertaken using 5 % TFA in DCM. The cholesteryl chloroformate reaction was carried out as previously. However, there was no product after purification. As this reaction proved capricious it was decided to try a different method to synthesise a cholesterol anchor.

The next approach to synthesise a cholesterol anchor mimetic was to couple ethylene diamine with cholesteryl chloroformate prior to reacting with a peptide (**compound 3**, **Figure 40**) to form **anchor 2**. **Compound 3** was designed to include three glycine molecules necessary for sortagging. It also included a propargyl glycine for click reactions to functionalise the anchor. The synthesis of compound 3 was attempted first and undertaken using commercially bought Boc-(Gly)₃-OH and Pra-OMe. Briefly DIPEA (1.034 mL, 6.32 mmol, 4 eq.), Boc-(Gly)₃-OH (457 mg, 1.58 mmol, 1 eq.) and EDC (454 mg, 62.37 mmol, 1.5 eq.) were dissolved in dry DCM. After 1 min Pra-OMe (100 mg, 0.79 mmol, 0.5 eq.) was added and the reaction left at RT overnight. The product was then washed with water, 5 % (w/v) lithium chloride and brine before purification with a Biotage SNAP KP-sil flash purification cartridge. Unfortunately after repeated attempts all yields were zero. Due to the difficult nature of the cholesterol reactions it was decided to take a different approach to anchor synthesis.

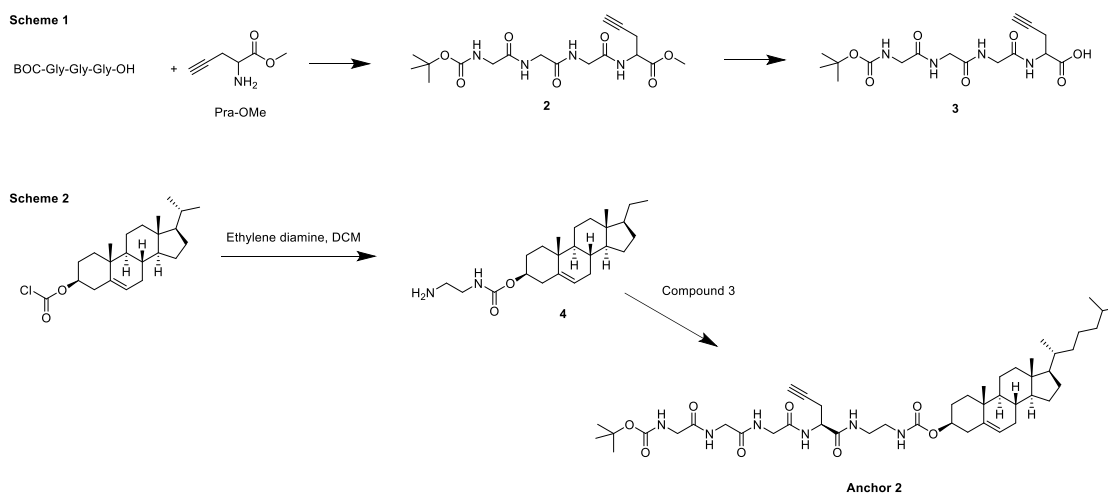


Figure 40 Reaction scheme to form cholesterol anchor mimetic, compound 5.

4.2.1.2 *GPI Anchor Mimetics based on Di-myristoylated Cytotopic Peptide with Myristoyl Group Localisation to Cell Membrane*

Synthesis of a GPI anchor mimetic with cholesteryl chloroformate proved particularly challenging and in light of this it was decided to design a new anchor mimetic without a cholesterol moiety based on previous research by Hill et al⁷³. Hill et al⁷³ have synthesised a GPI anchor mimetic based on a charged hydrophilic cytotopic peptide and two myristoyl groups linked to an N-terminal lysine. The charged peptide should associate with the phospholipid headgroups of the cell membrane and the myristoyl groups into the hydrophobic core of the membrane thus anchoring to the cell surface⁷³. The next anchor mimetic design utilised this approach with both the cytotopic peptide and myristoyl groups. However, an azide and glycines were added to facilitate visualisation and sortagging.

4.2.1.2.1 1st Generation Cytotopic Myristoylated **Anchor 3**

The cytotopic peptide (Fmoc)K(Fmoc)SSKSPSKDDKKPGDK(ivDde)K(N₃) was made using SPPS on a rink amide resin (see Methods). Following SPSS the orthogonal ivDde protecting group was removed to allow the growth of the glycine side chain. Following this the N-terminal lysine was deprotected and both amino groups myristoylated to give the final product anchor 3.

4.2.1.2.2 2nd Generation Cytotopic Myristoylated **Anchor 4**

After initial sortagging experiments with the CYT anchor it was found that purification after sortagging was challenging as discussed in further detail in section 4.3. Therefore, a second anchor (**anchor 4, Figure 37**) was designed informed by the success and limitations of **anchor 3**. The peptide was synthesised using SPPS as previously but using a biotinylated resin enabling purification of anchored protein with monomeric avidin agarose (section 4.3). Furthermore, there is a photo leucine and an arginine added to the glycine side chain. The arginine is present so there is an additional cleavage site available for trypsin digestion providing smaller fragments for analysis by mass spectrometry. The photo leucine provides extra functionalisation; the diazirine upon irradiation with UV light forms a reactive carbene species that forms covalent bonds with amino acid side chains and/or peptide backbones. This photocrosslinking will enable the protein-protein interactions occurring between OPCML and other proteins to be elucidated via proteomics. **Anchor 4** was used in preference to the first generation in all future experiments.

Functionalising rOPCML with a GPI anchor mimetic

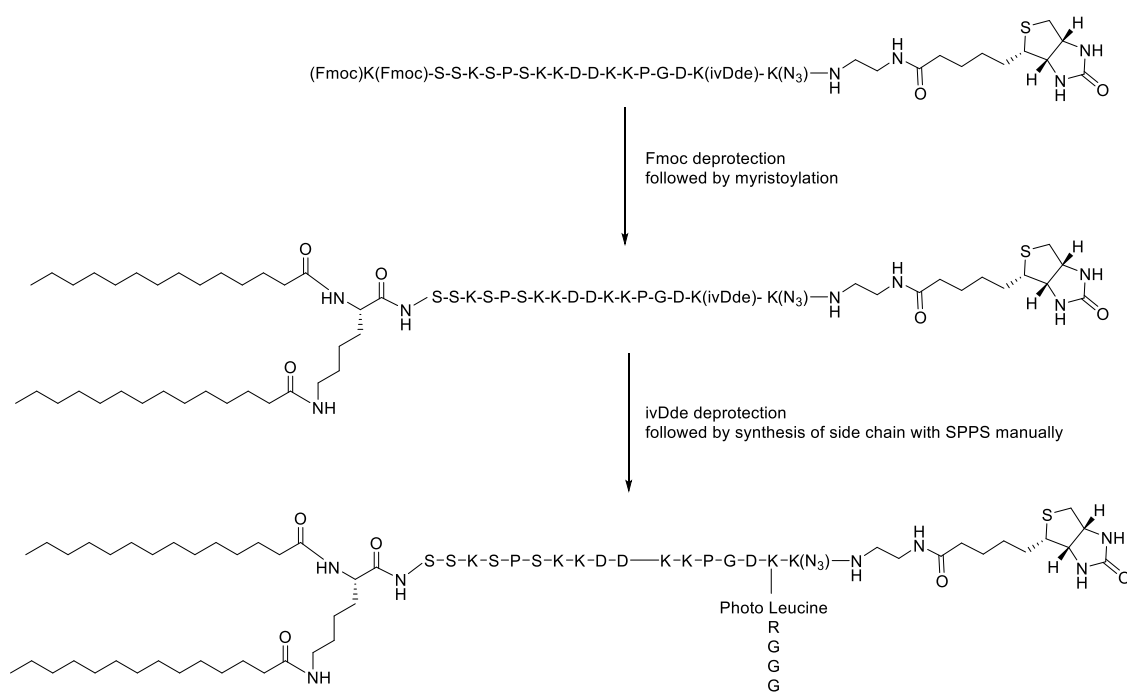


Figure 41 Scheme to show steps in synthesis of anchor 4.

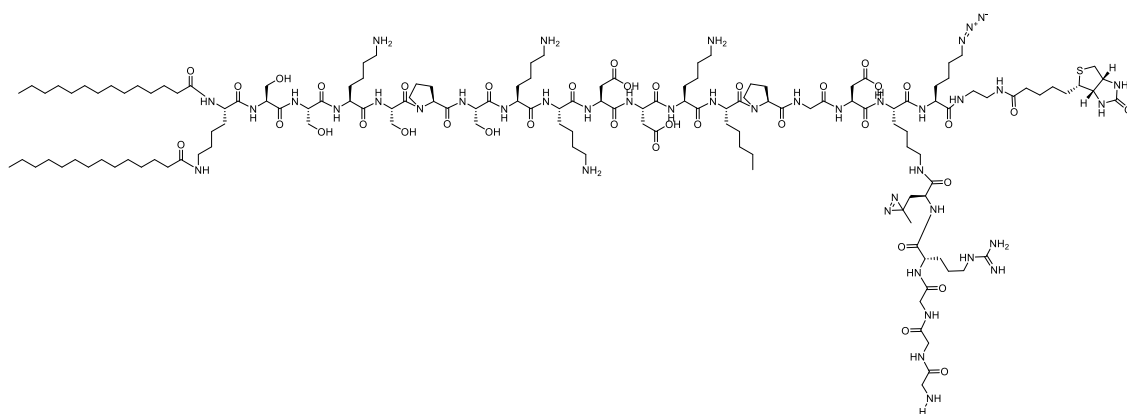


Figure 42 Full structure of anchor 4.

4.2.1.3 Confirming ability of sortagging to attach anchor mimetic

Initial sortagging test reactions were carried out using **anchor 3** (discussed in 4.2.1.2.1) to see if the sortagging reaction would be successful. The first attempt was carried out as outlined in Methods with the reaction being incubated at 37 °C for 2 h with sortase 5M^{148,149}. Briefly, Sortase 5M was reacted with rOPCML+Fc (this protein was discussed in 3.1.2) in 0.05 % DDM to ensure the anchor remained in solution (attempts at this reaction at without DDM yielded no result due to anchor precipitation). After 2 h the reaction was stopped by precipitation (see Methods) which also served as a way of removing any excess anchor mimetic. A copper(I)-catalysed azide alkyne cycloaddition (CuAAC) was then undertaken with YnTB^{81,82}. YnTB is a trifunctional capture reagent with an alkyne for undergoing cycloaddition and TAMRA and biotin groups for added functionality¹. The capture reagent was used to click to the azide group present in anchor 3. The TAMRA present on the YnTB allows visualisation of the successfully anchored OPCML on an SDS-PAGE gel visualised by fluorescence scanning the wet slab-gel.

The SDS-PAGE analysis of the test reaction is shown in **Figure 43A** shows the SDS-PAGE analysis

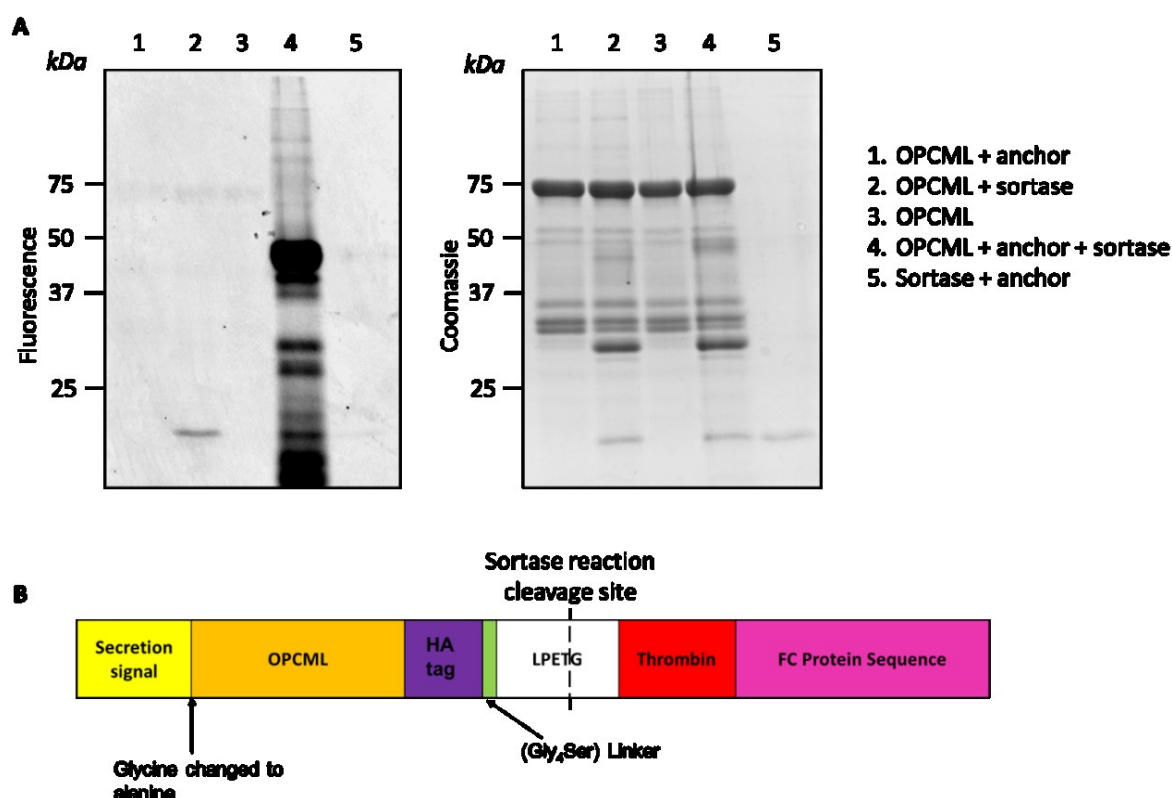


Figure 43 (A) SDS -PAGE to confirm the successful anchoring of the protein via fluorescence imaging of the wet slab gel and coomassie. (B) Schematic of construct designed for sortagging with sortase reaction cleavage site indicated.

of the test reaction. In each gel lanes 1, 2, 3, and 5 are control reactions. In the Coomassie

stained gel, the lanes show all protein present thus rOPCML+Fc, rOPCML-anchor, Fc and sortase can also be seen in various lanes depending on the controls.

Figure 43B shows the schematic of the protein construct that is undergoing sortagging and marks the cleavage site where the sortase forms the acyl enzyme intermediate prior to anchor attachment. Therefore, if the sortagging has occurred there should be a shift in size of rOPCML. Lane 4 is the only lane which has all components required for sortagging to proceed and the coomassie stained gel shows there is unreacted rOPCML+Fc at around 75 kDa, but there is also a smaller sized band at 50 kDa due to the LPETG site cleavage upon sortagging. In the fluorescence visualised gel the only band seen should be anchored protein, as only the anchored rOPCML will have a fluorescent moiety. There is also a band present at around in lanes 2, 4 and 5 which is the correct size for sortase (~20 kDa), as this band is only seen in lanes where sortase is present in the reaction it can be assumed that the sortase is the cause of this extra band. Lane 4 is also the only one that has a fluorescence signal and the size of the band corresponds to this smaller band on the coomassie stained gel indicating the sortagging is successful. However, the presence of unmodified rOPCML+Fc in lane 4 shows that this reaction has not gone to completion and there is a significant amount of unreacted protein. Therefore the reaction needs to be optimised.

Optimising the conditions for Anchor Attachment

The sortagging reaction is reversible so there will always be some unreacted protein^{111,150}. However, it could be possible that the Fc protein sequence is blocking the availability of the recognition site for the sortase enzyme. Therefore, the sortagging was repeated using rOPCML with and without the Fc protein sequence to compare results and see if the Fc is causing inaccessibility issues for the sortase. The sortagging reaction was set up as previously with one reaction undertaken using rOPCML+Fc, and a second reaction with rOPCML. The reactions were left for 24 h at 37 °C with samples taken at 0, 6 and 24 h prior to precipitation and CuAAC ligation to the YnTB capture reagent.

Figure 44A shows a slight decrease in band intensity between 6h and 24 h of the unmodified protein at 75 kDa in the coomassie stained gel for rOPCML+Fc. This result suggests that longer reaction times increase the amount of protein that is modified by sortagging. **Figure 44B** gives the results of sortagging with rOPCML. However, it is not possible to determine with this experiment whether the sortagging reaction proceeds better without the Fc protein sequence as there is no size difference between un-reacted and reacted rOPCML on the SDS-PAGE. Unfortunately all the sample was loaded onto the SDS-PAGE but in the future it would be advisable to retain sample so that a pulldown with streptavidin beads can be used to indicate the amount of reacted protein.

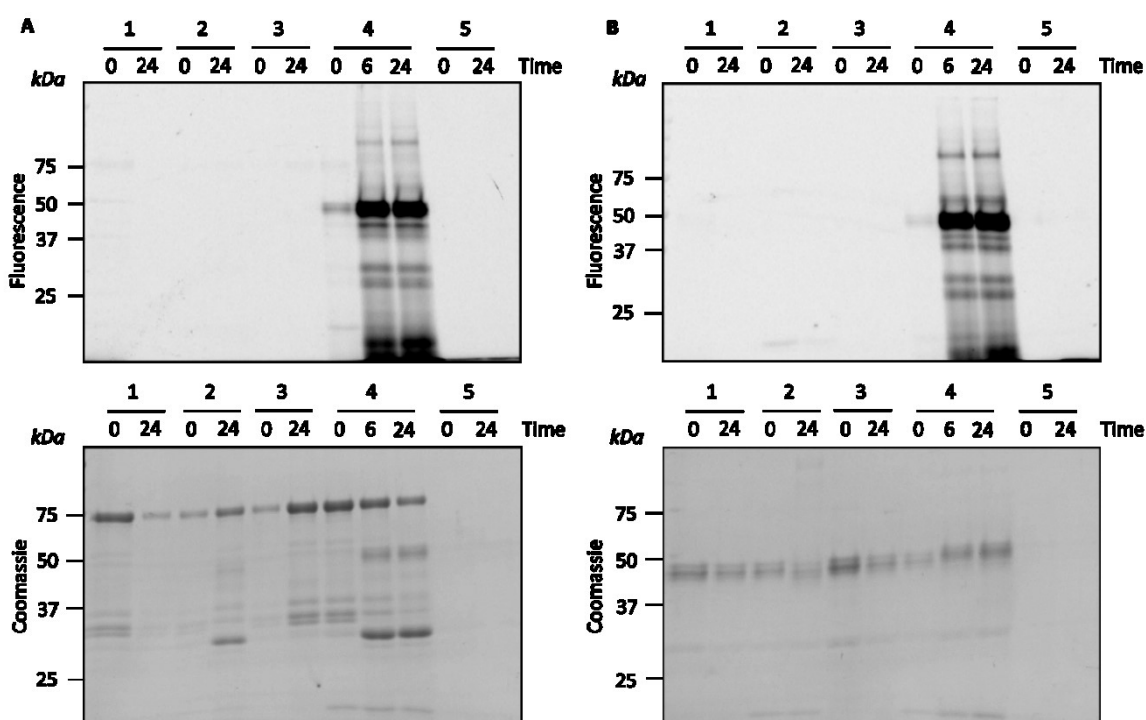


Figure 44 SDS -PAGE of optimising sortagging reaction. (A) OPCML+Fc. (B) OPCML. With no Fc. For both (A) and (B) as in Figure 10; 1=OPCML + anchor, 2=OPCML + sortase, 3=OPCML, 4=OPCML + anchor + sortase, 5= Sortase + anchor.

It was at this stage of experimentation that the free cysteine in rOPCML was ascertained to be causing protein aggregation (3.2.4.1). Therefore, for the next sortagging attempts the possibility that sortagging efficiency is being hindered by aggregation causing LPETG recognition site inaccessibility was investigated. Moreover, due to long reaction times being preferable for more complete anchoring of the protein (determined from **Figure 44A**) all future sortagging experiments were carried out at 4 °C based on the assumption that this will ensure the protein stays functional and folded.

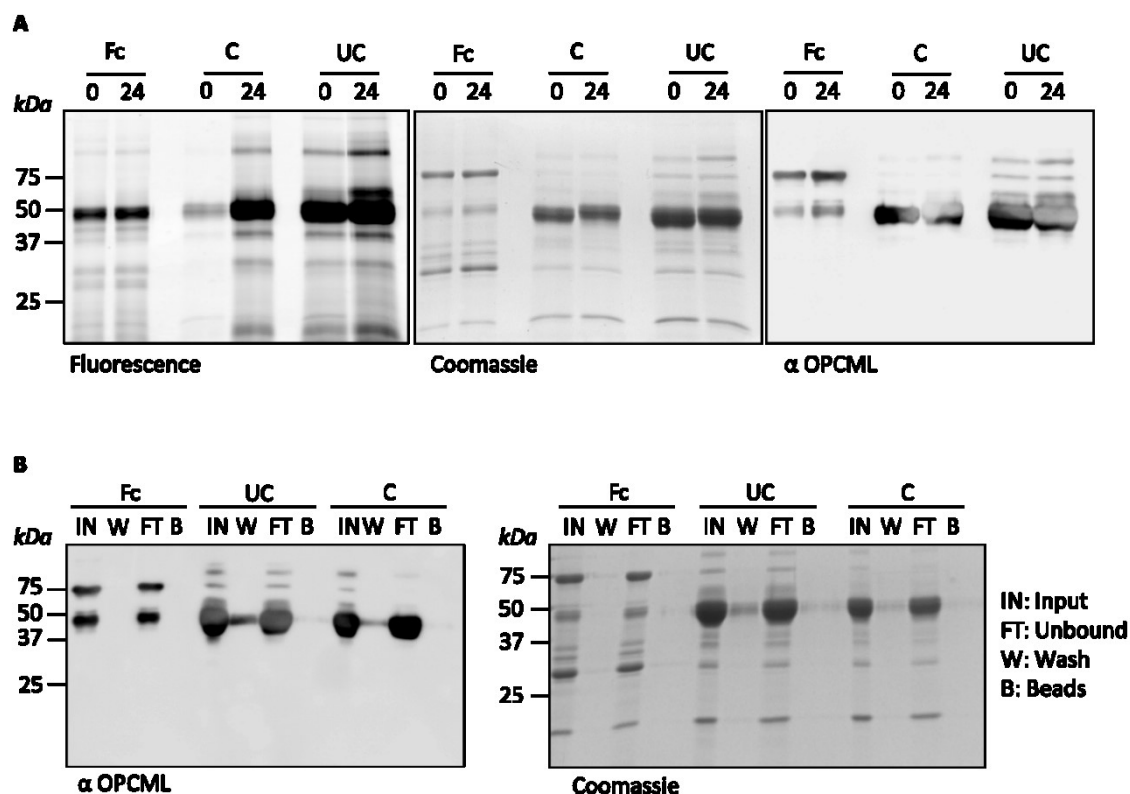


Figure 45 (A) SDS-PAGE of precipitation and click of rOPCML after sortagging. Fc = rOPCML+Fc, C = rOPCML capped with iodoacetamide, UC = rOPCML uncapped. (B) SDS-PAGE of pull-down with magnetic streptavidin beads post-sortagging.

The efficiency of the sortagging reaction after removing the free cysteine was tested after reducing the free cysteine in rOPCML+Fc with 10 mM DTT for 15 min at RT and subsequent capping with iodoacetamide (20 mM for 30 min at RT protected from light). After sortagging the samples were precipitated (see Methods) and CuAAC was carried out with YnTB as previously. The samples tested were rOPCML+Fc, and rOPCML capped (C) and un-capped (UC). **Figure 45A** shows sortagging was successful as there is a fluorescence band which would not be present if the anchor had not attached to the protein and clicked to the YnTB capture reagent. This band is also at 50 kDa which corresponds to the size of rOPCML. As before although there is a shift in band size when sortagging has occurred with rOPCML+Fc there is no such size shift with the capped or uncapped protein due to the Fc already being removed. Therefore, to see the proportion of anchored protein a pull-down was undertaken with magnetic streptavidin beads to enrich only anchored OPCML which has been clicked to YnTB (see Methods).

Figure 45B shows the SDS-PAGE from the pull-down post sortagging. Unfortunately for Fc, UC and C the protein did not bind to the beads and all the protein appears in the flow through supernatant collected from the beads after a 1 h incubation. It was decided to continue with sortagging experiments even though the conversion could not be determined from this

pulldown due to time constraints and lack of sample available for further repeats. Additionally, as the free cysteine had been removed the possibility of the aggregation causing lower reaction rates was no longer relevant. Therefore, it was decided that the Fc protein sequence should be kept prior to sortagging to aid in purification post-sortagging (section 4.3). Furthermore, the process of removing the Fc sequence causes protein loss (due to anion exchange post thrombin cleavage followed by buffer exchange and protein concentration) thus it was decided future sortagging experiments would use rOPCML+Fc rather than the cleaved rOPCML. All future sortagging experiments were also to be undertaken overnight at 4 °C to get more complete conversion.

4.2.1.4 Testing the solubility of anchored rOPCML post sortagging

Unmodified OPCML is soluble but the anchor requires detergent to be soluble for the sortagging reaction. The need for detergent is problematic for future processing. However, as rOPCML is soluble it is possible that after anchor addition the protein-anchor complex may stay in solution without detergent or with reduced levels of detergent. Therefore BioBeads™ can be used to test the solubility post sortagging. Bio-Beads™ are adsorbent macroporous polymeric beads that can be used to remove detergents from solution. To test the solubility of the anchored protein after sortagging the reaction was carried out as before. After the reaction half the product was precipitated (see Methods) and CuAAC was carried out with YnTB. The other half was incubated with Bio-Beads™ SM-2 for 1 h at RT, then the beads pelleted and the supernatant collected prior to precipitation and CuAAC. **Figure 46** shows the SDS-PAGE of the reactions. Samples were taken for SDS-PAGE analysis after the sortagging reaction prior to clicking with YnTB and after the click reaction. The Bio-Beads™ themselves were also boiled with loading buffer and the supernatant analysed. The attachment of the anchor onto the protein was successful as shown by the band at 50 kDa on the fluorescence gel (**Figure 46**).

Furthermore, the removal of detergent by the Bio-Beads™ did not cause the product to crash out as both with and without Bio-Bead™ incubation the click reaction worked. There is a small decrease in band intensity when the detergent is removed, but this could be due to protein sticking to the beads shown by the presence of protein in the boiled beads lane of the gels. This result shows the anchored protein is soluble therefore the concentration of detergent does not need to be kept constant in downstream use of the protein-anchor complex.

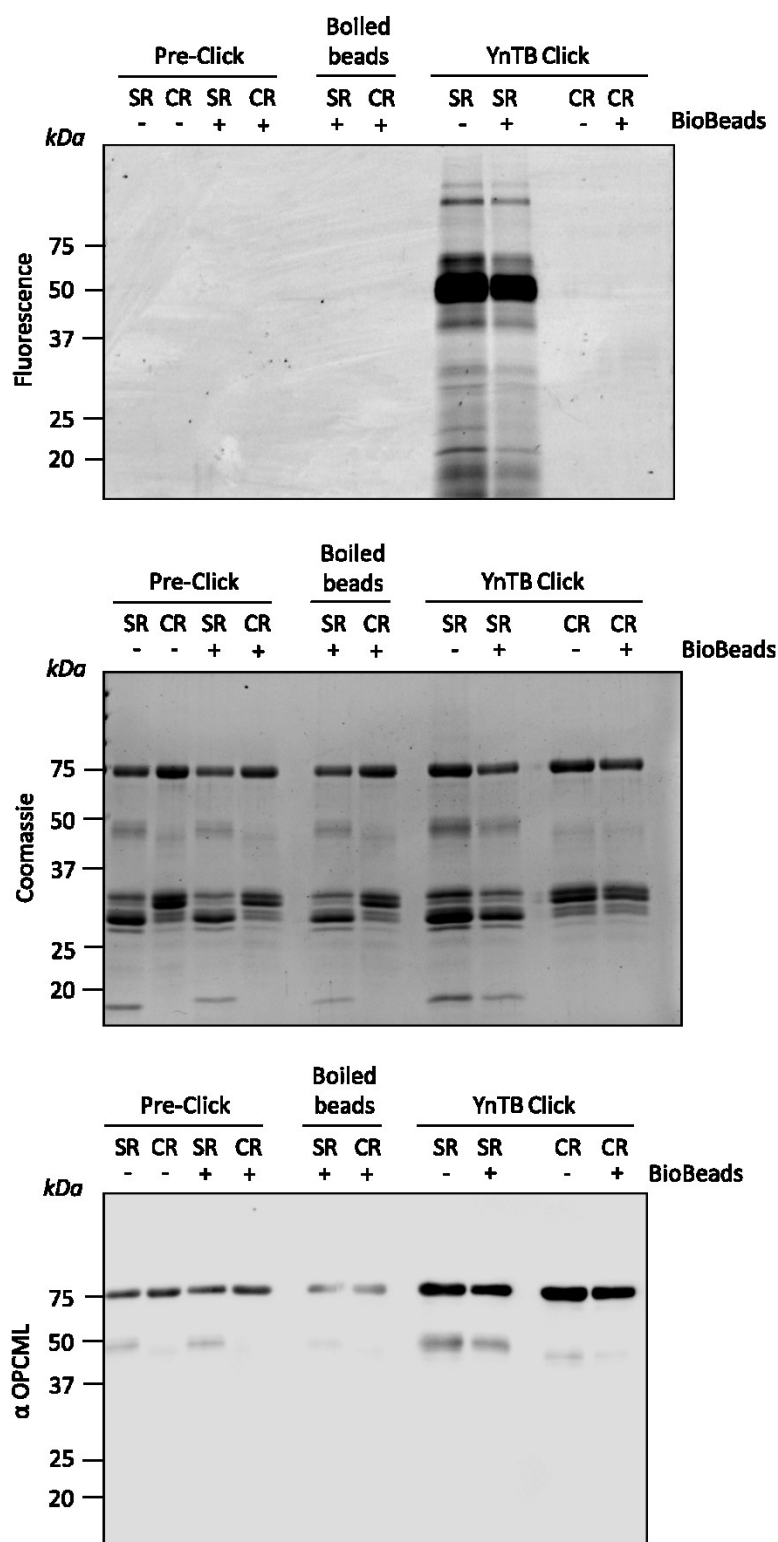


Figure 46 SDS-PAGE of rOPCML+Fc post-sortagging to test solubility of anchored protein. Samples were taken after sortagging prior to the click reaction with YnTB and after the click reaction. SR = sortase reaction, CR = controls reaction with no sortase present.

The critical micellar concentration (CMC) is the minimum concentration required for micelle formation¹⁵¹. Therefore the concentration of detergent used for solubilising the anchor for sortagging is above this CMC. If the detergent concentration were below the CMC micelles would not form and the hydrophobic anchor would not be able to insert into the micelle to be stabilised. However, upon the removal of detergent by BioBeads™ the concentration of detergent will decrease to below the CMC and micelles will no longer be formed. Thus no stabilisation for the anchor and if the anchor or protein-anchor complex is insoluble it will precipitate. However, it could be possible that the hydrophilicity of the protein overcomes the hydrophobicity of the anchor once they are ligated.

4.3 Purification of the Protein Anchor Complex Post-Sortagging

The sortagging reaction has been proven to be successful and the conditions optimised therefore the next step was to undertake larger scale sortagging reactions. However, all previous attempts at sortagging have concluded with precipitating the product followed by a click reaction, which is not a suitable way of purifying the protein for functional downstream assays. Therefore, a suitable way of purifying the recombinant protein post sortagging was needed. As mentioned previously sortagging can be carried out on the protein whilst the Fc is still present but there will still be cleaved un-anchored protein present after the sortagging reaction due to its reversible nature. The difference in size between the anchored and non-anchored protein is approximately 2 kDa and this size difference is not enough for SEC to be used to separate out reacted and unreacted protein. Therefore, SEC is not a suitable technique for purification post sortagging. During purification of rOPCML anion exchange (3.2.3) was used after thrombin cleavage to remove impurities. It was decided to try and use anion exchange to purify post-sortagging as there is a small difference in the pI between rOPCML+Fc (pI = 6.37), anchored rOPCML (pI = 6.0) and rOPCML (pI = 5.38) that may make the separation of the different proteins species possible (**Figure 47B** pI values were calculated using PepCalc.com, copyright © 2015 Innovagen AB). Anion exchange was attempted with a HiTrap ANX column (GE Healthcare) using a starting buffer of 50 mM Tris, 50 mM NaCl, pH 7.5 containing 0.05 % DDM to avoid any protein precipitation which would result in blocking the anion exchange column. A gradient was set up to 50 % elution buffer (1 M NaCl, 50 mM Tris pH 7.5) over 100 CVs.

Figure 47A shows rOPCML+Fc was eluted during the anion exchange and only a small amount of this protein is present as can be seen by the UV trace in **Figure 47C**. Due to the lack of eluted protein it was decided to try a different method for purifying post-sortagging. Many examples of purification after using sortase to modify proteins use affinity purification^{105,149,152}. Therefore, affinity purification was used for the following attempts, with resins being prepared for use following the manufacturer's protocols unless otherwise stated.

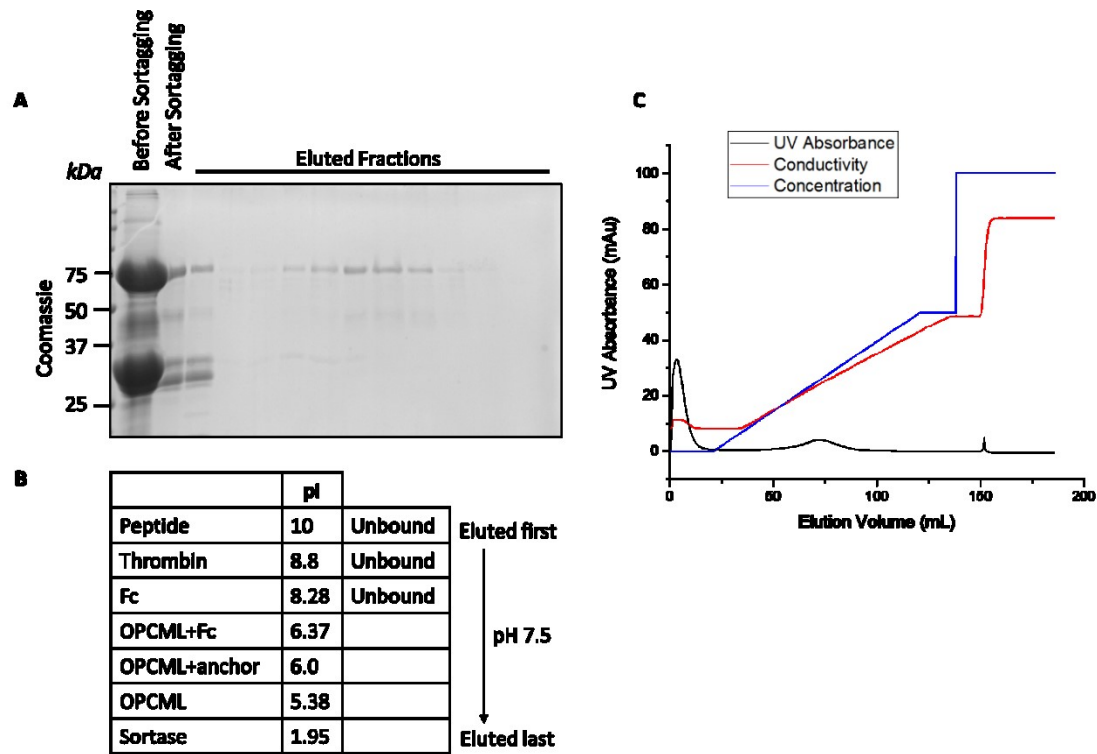


Figure 47 (A) Anion exchange after sortagging. (B) Table of pI values. (C) UV, conductivity and concentration trace of anion exchange.

The first attempt at affinity purification utilised the HA tag which is present in rOPCML. Sortagging was carried out overnight at 4 °C where after it was isolated with Pierce™ Protein A agarose for 1 h at 4 °C. Anchored rOPCML should stay unbound whereas any Fc or rOPCML+Fc should be bound to the Protein A agarose. Following this, the supernatant was incubated with Pierce™ Anti-HA agarose for 2 h at 4 °C whilst rotating. The agarose beads were then washed with wash buffer (50 mM Tris, 150 mM NaCl, pH 7.5) three times prior to elution with 50 mM glycine, pH 3.0. The eluted fractions were neutralised with 100 µL neutralisation buffer (2 M Tris, pH 8.5). The protein did not bind to the HA beads as can be seen from lane 5 in **Figure 48A** as the protein is present in the supernatant from the HA bead incubation. However, the

incubation with the Protein A agarose did successfully remove the majority of the unreacted rOPCML+Fc and Fc. Therefore a new approach was tried using incubation with Pierce™ Protein A agarose as before but using a spin concentrator to remove the unreacted peptide (2.6 kDa) and sortase (21 kDa) as these are smaller than rOPCML. Unfortunately the spin concentrator concentrated all components of the reaction with only a small amount of sortase present in the flow through **Figure 48B**.

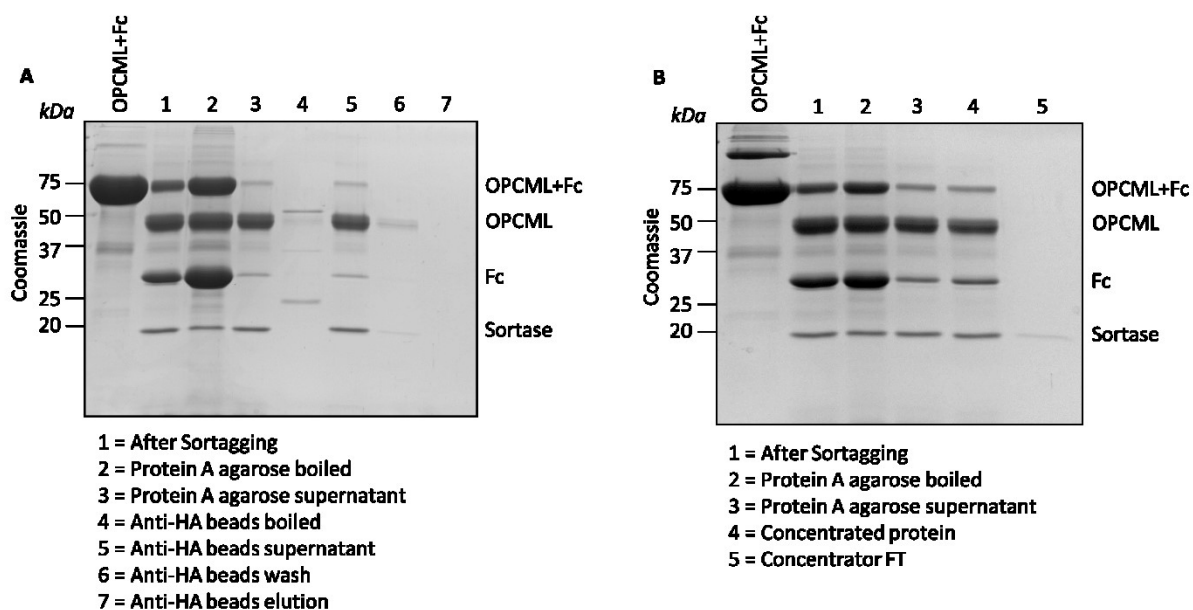


Figure 48 SDS-PAGE of post-sortagging purification attempts. (A) Purification using protein A agarose and Anti-HA magnetic beads. (B) Purification using protein A agarose and spin concentrators.

At this stage due to the difficulties purifying anchored rOPCML and the need for two resin incubation steps a new anchor mimetic was designed (section 4.2.1.2.2). Anchor 4 included a biotin molecule at the C terminus of the peptide allowing the anchored protein to be isolated from the sortase reaction mixture by incubation with monomeric avidin agarose. Avidin is a protein that binds biotin and can be used for enriching proteins with a biotin tag. The affinity between biotin and avidin is quite high. Therefore monomeric avidin, which has a lower affinity for biotin, is used when proteins need to be eluted from avidin and remain functional¹⁵³.

The post-sortagging purification with anchor 4 was carried out by incubating the sortagging reaction mixture with the monomeric avidin agarose overnight at 4 °C. Pierce™ Monomeric Avidin agarose was prepared for use as per the manufacturer's instructions unless otherwise stated. For this purification attempt after the sortagging reaction the reaction mixture was bound to the monomeric avidin for 3 h at 4 °C. The protein was then eluted using 1.0 M glycine pH 2.8 prior to dialysis in wash buffer (50 mM Tris, 150 mM NaCl, pH7.5). **Figure 49A** shows that

rOPCML+Fc and Fc were present after eluting from the monomeric avidin suggesting that further washing steps are needed prior to elution. Furthermore, a band corresponding in size to cleaved rOPCML is present in the washes. This band could be due to not all the anchored protein binding to the monomeric avidin, but could also be rOPCML that has been hydrolysed but has not reacted due to the reversible nature of the sortagging reaction. **Figure 49A** also indicates that the monomeric avidin agarose was still bound to protein after elution.

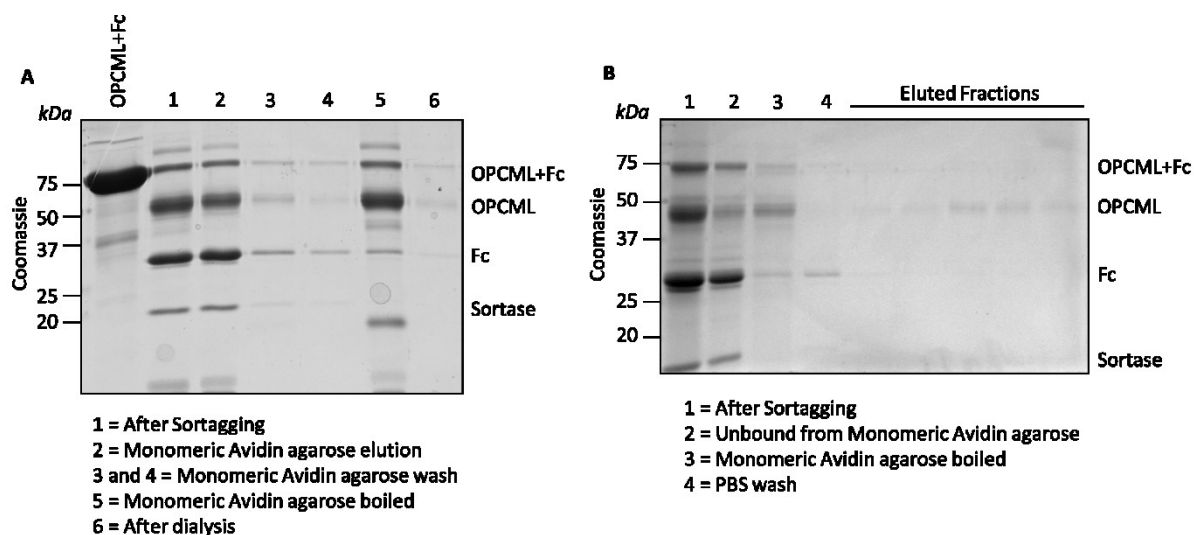


Figure 49 (A) and (B) SDS-PAGE of post-sortagging purification attempt with monomeric Avidin agarose.

To overcome these issues a further attempt at purification was undertaken with the sortagging reaction mix being incubated with the monomeric avidin resin overnight at 4 °C to facilitate more complete binding of anchored protein. Furthermore, the anchored protein was eluted from the agarose with 2mM biotin in PBS and 1 mL aliquots were monitored for presence of protein via UV detection. **Figure 49B** shows the result of this later purification attempt. From lane 2 it is evident that incubation overnight with the agarose still leaves unbound protein. There is also anchored protein still on the resin after elution. However, anchored protein has been successfully eluted as can be seen in the latter lanes of elution fractions.

To attempt complete elution from the monomeric avidin resin a further attempt at purification was carried out. In this attempt the sortagging reaction was carried out overnight at 4 °C as was the binding to the monomeric avidin agarose. The resin was washed with PBS to remove any unbound species and then incubated with 2 mM biotin in PBS overnight at 4 °C to elute the anchored protein. After this the elution was dialysed to remove any unreacted anchor and biotin, followed by concentrating the dialysed protein with a spin concentrator. In **Figure 50** it

can be seen that overnight elution with 2 mM biotin does not release all anchored OPCML from the avidin beads (lane 2). This suggests that modification of the monomeric avidin agarose may be required to reduce the binding efficiency to facilitate more complete elution. Furthermore, lane 6 shows that after concentrating down there are still impurities (OPCML+Fc, Fc and sortase) present. This suggests the earlier eluted protein in **Figure 49B** was not free from impurities, they were just dilute enough not to be visible in SDS-PAGE.

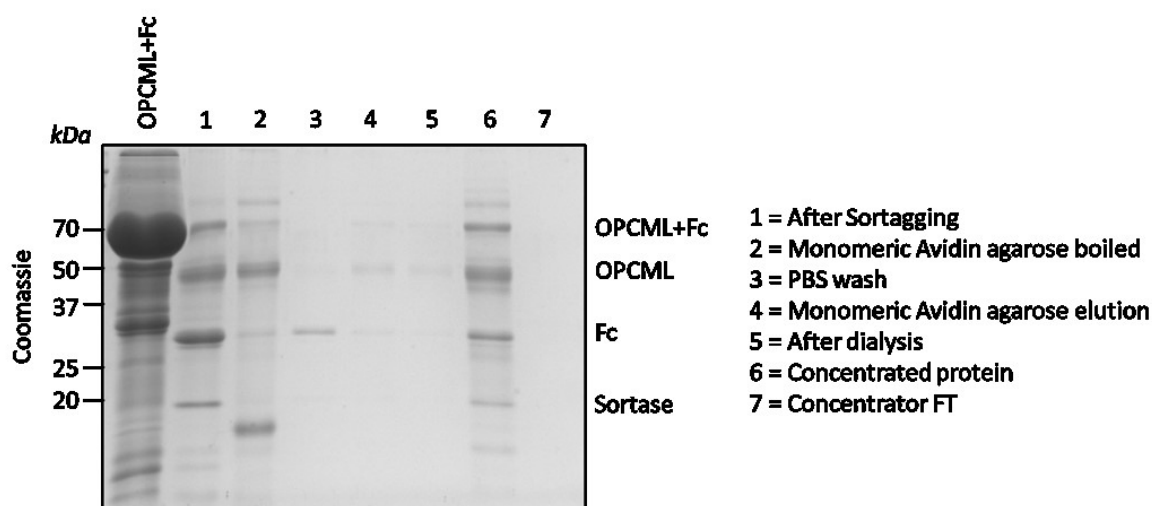


Figure 50 SDS- PAGE of post-sortagging purification attempt with monomeric Avidin agarose.

To overcome the problem of protein staying bound to the monomeric avidin resin after elution with 2 mM biotin the concentration of biotin was increased to 5 mM and the elution was carried out at RT rather than 4 °C (**Figure 51A**). However, there was still protein bound to the monomeric avidin resin after the 3 h incubation with elution buffer. The bands present in lanes 5-7 do appear relatively clean but there are still trace amounts of Fc and OPCML+Fc present.

A further attempt at purification was then undertaken taking advantage of the Fc tag present in unreacted protein and cleaved Fc side product, and the 6x his tag present in the sortase. Thus after the sortagging reaction the reaction mixture was incubated with Protein A agarose overnight at 4 °C to remove unreacted OPCML+Fc and Fc. Subsequently the supernatant was removed and incubated with Ni-NTA resin in an attempt to remove the sortase. **Figure 51B** shows the result of this purification attempt with the sortase still being present after purification.

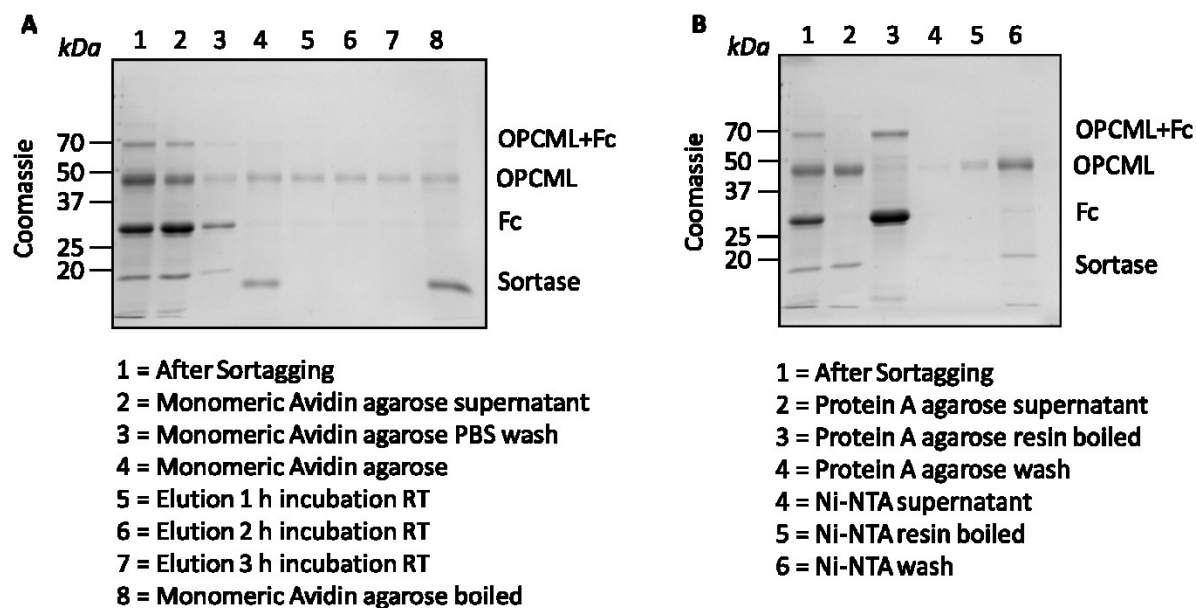


Figure 51 (A) SDS-PAGE of post-sortagging purification attempt with (A) Monomeric Avidin agarose (B) Protein A agarose.

As the monomeric avidin elution was successful at RT and the protein A agarose successfully removed OPCML+Fc and Fc impurities it was decided to combine these two resin steps for the next purification. After sortagging the reaction was dialysed for 12 h in an attempt to remove unreacted anchor then the sample incubated with monomeric avidin agarose. Elution was carried out for 3 h at RT and the eluted mixture dialysed in the presence of Protein A agarose to remove any remaining impurities. **Figure 52** shows the first dialysis after sortagging has no obvious effect on the smaller impurities visible on the coomassie stained gel. However, lane 6 does seem to have no visible bands present for impurities and from concentration readings from Nanodrop spectrophotometer and there was 0.08 mg of anchored protein obtained from a starting amount of 1.2 mg of OPCML+Fc indicating 6.6 % of the initial amount of protein was anchored. Therefore, in future this method will be used for purification of the recombinant protein after modification with the GPI anchor mimetic with further optimisation of the sortase reaction this yield could be increased.

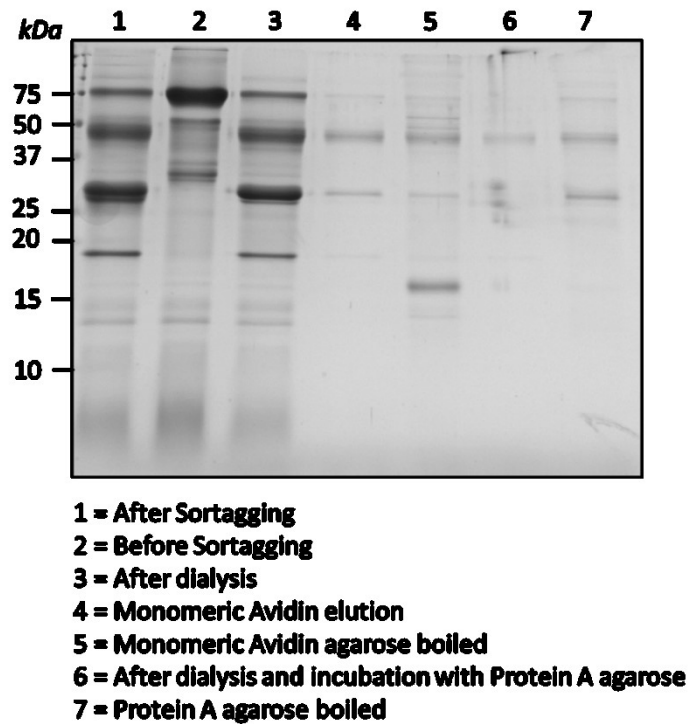


Figure 52 SDS-PAGE of post-sortagging purification attempt with monomeric Avidin agarose and protein A agarose.

4.4 Conclusion

The main objective of this research is to create a novel OPCML-GPI anchor mimetic that functions as a biotherapeutic in epithelial ovarian cancer. To that end a GPI anchor mimetic that anchors OPCML to the cell membrane and facilitates the protein-protein interactions necessary for OPCML tumour suppressor function is needed. Different GPI anchor designs were investigated as well as methods of anchor attachment. Initial attempts at anchor attachment using IPL did not yield promising results with low protein expression observed. Different GPI anchor mimetics were investigated with cholesterol containing anchors proving challenging to synthesis. However, later anchor mimetics utilising myristoyl groups and a cytotopic peptide influenced by Hill et al⁷³ have proven promising with useful anchoring of the protein observed by SDS-PAGE with sortagging. Further optimisation of sortagging could be carried out in future to increase yields of anchored protein. Purification of the protein post-sortagging was also carried out with various methods, ultimately ending with an improved anchor design to facilitate a more direct purification process. This purification was also optimised giving anchored rOPCML with few impurities thus making it suitable for functional assays in cells.

Chapter 5

Conclusions and Outlook

5.1 Conclusions

EOC is one of the most common cancers affecting women and the prognosis for patients with this disease is very poor^{6,7}. After treatment the chance of cancer recurrence is extremely high^{6,7,17,18}. OPCML is a TSG epigenetically inactivated in EOC^{5,31}. The function of OPCML as a tumour suppressor is due to the protein negatively regulating a subset of RTKs via sequestering them into the DRM where OPCML is localised due to a GPI anchor^{32,52–54}. Native GPI anchors require complex synthetic routes to synthesise, therefore simplified mimetics are often utilised for anchoring recombinant proteins^{56,69–72}. Approaches such as IPL and sortagging can be used to ligate GPI anchor mimetics to recombinant OPCML^{75,90,100,104}.

Previous research has demonstrated that reintroduction of functional OPCML into cancer cells reduces cell migration and proliferation³². Therefore, it was hypothesised that rOPCML could be a valid therapeutic for treating patients with EOC after surgery to prevent further dissemination and improve quality of life for patients post-surgery. To this end the main aim of this research was to design and validate a novel biotherapeutic based on an OPCML-GPI anchor mimetic and this was undertaken in three main parts; design, production and purification of an rOPCML protein suitable for ligation of an anchor mimetic, design and synthesis of a GPI anchor mimetic, ligation of the anchor mimetic to the protein.

Initially the SAR of OPCML was to be determined through the use of lentiviral transduction to generate cell lines stably expressing truncated forms of OPCML. However, although SKOV3 cells were generated that stably expressed single domains of OPCML in MTS assays reproducible results were not obtained. This could be due to the inability to control the levels of protein expression in each cell line.

The structure of OPCML was subsequently elucidated by Birtley et al³³ providing valuable structural information and a method of functional recombinant protein production. Thus the focus of this research moved towards the generation of an OPCML construct suitable for use as a novel biotherapeutic based on the generation of a protein-GPI anchor mimetic.

A highly engineered functional OPCML construct capable of attachment of complex chemical prosthetic groups to mimic the GPI anchor was designed, expressed using a baculovirus expression system and purified. The first construct included two methods of anchor attachment; a free cysteine and a sortase recognition sequence. However, SEC showed the presence of the free cysteine was causing aggregation so this cysteine was removed in the final protein. A mutant protein R56L was also produced to act as an inactive control in future functional assays.

For the GPI anchor mimetic different GPI anchor designs were investigated as well as methods of anchor attachment. Cholesterol containing anchors were initially explored as a suitable way of anchoring OPCML to the DRM but the synthesis proved capricious. The first attempts at ligation with IPL were unsuccessful with low protein expression observed therefore sortagging was explored as an alternative approach. Anchor mimetics utilising myristoyl groups and a cytotopic peptide influenced by Hill et al⁷³ proved promising with anchor attachment observed by SDS-PAGE with sortagging. Sortagging was optimised in an attempt to improve the yield of modified protein. Purification of the protein post-sortagging was also carried out with various methods, ultimately ending with an improved anchor design to facilitate a more direct purification process.

5.2 Future work

To overcome the issue of purifying the protein after anchor addition re-designing the protein construct to include a 6x his-tag between the thrombin cleavage site and the Fc sequence could be a viable option. This would allow sortase, unreacted protein and cleaved Fc sequence to be removed in one step via affinity purification. However, any unreacted anchor would still remain and would need to be removed via dialysis.

There is a possibility that after all the steps required to purify the protein after anchor attachment that the protein is no longer folded correctly and thus not functional. Therefore, future work will need to be carried out to characterise the modified anchored OPCML.

Important next steps following on from this research include demonstrating the semi-synthetic anchored OPCML can be targeted to the membrane of cancer cells via a GPI anchor mimetic. This could be achieved via addition to cells of interest followed by co-localisation microscopy and fractionation. A photo-crosslinkable amino acid is present in the anchor mimetic therefore this could be utilised to identify interacting proteins at the membrane and demonstrate physical

Conclusions and Outlook

interaction. To undertake this photo-crosslinking could be done on live cells followed by pull down with biotin and digestion to enable analysis with proteomics.

Once the anchored protein has been shown to be functional in cell assays a method of delivery would need to be investigated.

Chapter 6

Methods

6.1 General Biology Methods

6.1.1 Plasmids/Constructs and Cloning

FL, D1, D2 and D3 entry clone constructs were purchased from Life Technologies (Carlsbad, California, USA). Restriction enzymes were purchased from New England Biolabs and used as per the manufacturer's instructions.

pet30b-5M SrtA was a gift from Hidde Ploegh (Addgene plasmid # 51140) (Cambridge, Massachusetts, USA). The OPCML sortagging construct was design and ordered from Life Technologies (Carlsbad, California, USA).

The OPCML pFastBac1 plasmid was design and ordered from Life Technologies (Carlsbad, California, USA).

Q5® Site-Directed Mutagenesis Kit (New England Biolabs, Hitchin, UK) was used following the manufacturers' instruction.

Ready-To-Go T4 DNA Ligase (GE Healthcare, Little Chalfont, UK) was used to transfer constructs into pENTR4 as per the manufacturer's instructions and QIAquick® Gel Extraction Kit was used to purify the DNA. Gateway LR Clonase II Enzyme mix kit from Life Technologies was used to transfer the constructs from entry into destination vectors following the manufacturer's protocol, optimised with a change from a 1 hour incubation at 25 °C to 3 h. p8.9 and pMDG plasmids were provided by Dr Ari Fassati, University College London UK. pLenti CMV Puro DEST was purchased from Addgene.

pLenti CMV Puro DEST (w118-1) was a gift from Eric Campeau & Paul Kaufman (Addgene plasmid # 17452 ; <http://n2t.net/addgene:17452> ; RRID:Addgene_17452)

6.1.2 Transformations

Transformations were undertaken using Alpha-Select Efficiency Gold Competent Cells from Bioline, dam-/dcm- Competent E. Coli from New England Biolabs, and One Shot OmniMAX™ 2 T1R Chemically Competent E. coli from ThermoFisher Scientific (Waltham, Massachusetts, USA) following the manufacturer's instructions.

Competent cells were thawed on ice for 20 minutes. After thawing, the cells were gently mixed and 50 µL aliquots (one per each transformation) were transferred to 1.5 mL eppendorf, which were pre-chilled on ice. 1 µL of DNA (100 ng/mL) was added to the competent cell aliquot and mixed gently. The cells and DNA were incubated on ice for 30 min, then heat shocked for 45 sec at 42 °C and placed on ice for 2 minutes. 950 µL of pre-warmed YT media containing no antibiotic was then added to the cells and the eppendorf shaken for 1 h at 30 °C. The transformation mixture was then spread on agar plates containing the appropriate antibiotic, allowed to dry and incubated at 30 °C overnight.

Purification was undertaken using QIAprep® Spin Miniprep Kit or Highspeed Plasmid Maxi kit from Qiagen both used as per the manufacturer's instructions.

6.1.3 PCR

PCR was carried out on a MJ Research PTC-200 peltier thermal cycler or an Agilent SureCycle 8800. All reagents used were from New England Biolabs apart from AccuPrime Taq DNA Polymerase from Thermo Fisher Scientific. Samples were then analysed using DNA electrophoresis. Agarose gel electrophoresis was carried out using 1 % agarose gels and visualised using a UV light chamber.

6.2 Mammalian Cell Culture

6.2.1 Cell Lines

HEK293T and SKOV3 were purchased from ATCC. HEK293T cells were grown in DMEM supplemented with 10 % FCS, and 2 mM L-glutamine. SKOV3 cells were grown in RPMI supplemented with 10 % FCS, 100 units/mL penicillin, 0.1 mg/mL streptomycin, and 2 mM L-glutamine. U87MG cells were obtained from Public Health England. Cell culture reagents were purchased from Gibco and Sigma Aldrich. Both cells lines were grown in a humidified environment with 5 % CO₂ at 37°C.

6.2.2 *Antibodies*

Monoclonal mouse anti-OPCML and polyclonal goat anti-OPCML were purchased from R&D Systems. Polyclonal rabbit anti-OPCML and monoclonal rat anti-HA Tag were purchased from Sigma Aldrich. Polyclonal rabbit anti-calnexin was purchased from Enzo Life Sciences. Horseradish peroxidase-conjugated secondary antibodies were bought from Dako. Alex Fluor 488 goat anti-rat and Alexa Fluor 546 goat anti-mouse were purchased from Thermo Fisher Scientific. Anti-6X His tag[®] antibody (HRP) (ab1187) was purchased from Abcam. Rabbit anti-goat IgG-HRP (sc-2768) was purchased from Santa Cruz. NeutrAvidin (A2664) was purchased from Thermo Fisher Scientific.

6.2.3 *Immunofluorescence Microscopy*

SKOV3 cells were grown on glass coverslips and fixed in ice cold methanol for 5 min at -20 °C. The coverslips were then washed with PBS and blocked for 30 min with 3 % BSA in PBS. The coverslips were incubated with primary antibodies diluted in 1 % BSA for 1 h at RT in a dark wet chamber, washed, then incubated with secondary antibodies (diluted in the same mixture) for 30 min under the same conditions.

All antibodies were diluted according to the manufacturer's instructions. All coverslips were then mounted onto slides using ProLong Gold antifade reagent with DAPI (Molecular Probes Life Technologies) and left to dry at RT before imaging on a Leica SP5 confocal.

6.2.4 *SDS-PAGE*

12 % acrylamide gels were made using reagents from National Diagnostics, ProtoGel (30%), ProtoGel Resolving Buffer (4X), ProtoGel Stacking Buffer. Gels were cast and run using Protein Electrophoresis and Blotting equipment from Bio-Rad. Gels were run at 80 volts until the sample front passed into the running gel then the voltage was changed to 180 until the gel had finished running. For coomassie stained gels, gels were washed for 5 min then stained in coomassie solution overnight, destained with water then imaged.

6.2.5 *Western Blot*

Lysates were prepared using RIPA (Radio-Immunoprecipitation assay) buffer (150 mM NaCl, 1.0 % IGEPAL[®] CA-630, 0.5 % sodium deoxycholate, 0.1 % SDS, 50 mM Tris, pH 8.0) and protease and phosphatase inhibitors. Protein concentration was determined using Thermo Scientific Pierce BCA protein assay kit as per the manufacturer's instructions. Absorbance in 96 well plates was measured using Molecular devices OPTImax tuneable microplate reader. SDS-PAGE was used to separate the proteins using 12% acrylamide gel. Proteins were transferred to a

Methods

nitrocellulose membrane, blocked for 30 min (5 % milk or BSA in TBST) before incubating overnight at 4 °C with primary antibody. Antibodies were diluted in 5% milk or 5 % BSA in TBST based on manufacturers' instructions. The membrane was then washed with TBST (3 times for 10 minutes each time) and incubated in secondary antibody for 45 min. Membranes were then incubated with Immobilon Western Chemiluminescent HRP (Millipore) or Luminata Crescendo Western HRP substrate (Millipore) for 1 min before being developed on Amersham Hyperfilm ECL films (GE Healthcare) or imaged on a ImageQuant LAS 500 (GE Healthcare). For fluorescence visualisation of wet slab gels a Typhoon FLA laser scanner (GE Healthcare)

6.2.6 Dot Blot

Samples for analysis were spotted onto dry nitrocellulose membrane and left to dry. The membrane was blocked with 5 % milk for 30 min then incubated with primary antibody overnight at 4 °C. Antibodies were diluted in 5% milk or 5 % BSA in TBST based on manufacturers' instructions. The membrane was then washed with TBST (3 times for 10 minutes each time) and incubated in secondary antibody for 45 min before washing with TBST. Membranes were then incubated with Luminata Crescendo Western HRP substrate (Millipore) for 1 min before being imaged on an ImageQuant LAS 5000 (GE Healthcare).

6.2.7 Lentivirus Production

HEK293T cells were plated 24 h prior to transfection in 10 cm Petri dishes in order to achieve 90 -95 % confluency in RPMI media with 10 % FCS and L-glutamine.

12 µg of each pLenti plasmid, 10.2 µg packaging plasmid (p8.9), and 1.8 µg VSV-G (pMDG) were mixed and diluted up to 1500 µL OptiMEM, vortexed and incubated at RT for 5 min. 60 µL Lipofectamine 2000 (Invitrogen) was diluted up to 1500 µL in OptiMEM, vortexed and incubated at RT for 5 min. Both the plasmid mixture and the Lipofectamine 2000 mixture were mixed, vortexed and incubated for 20 min at RT. The media of the HEK293T cells was replaced with OptiMEM. The mixture was then added dropwise to cells with swirling. After 5 h the cell media was replaced with RPMI (10 % FCS and L-glutamine). The cells were then incubated for 48 h before the virus was collected in the cell medium, centrifuged for 4 min at 1500 rpm and filtered (0.45 µL filter). The virus was then aliquoted and stored at -80 °C. Polybrene (Sigma Aldrich) was added to cell media at a concentration of 8 µg/mL 30 min before transduction.

6.2.8 Producing Stably Transduced Cell lines

SKOV3 cells were plated in 24 well plates and allowed to grow to confluence prior to transduction with the virus. 24 h after transduction, transduced cells were selected by using

Methods

medium containing puromycin (10 µg/mL). Once selected cells were confluent, they were expanded into a larger flask and the process continued until stocks of stable cell lines had been produced.

6.2.9 Transfection

HEK293 cells were plated 24 h prior to transfection in 10 cm Petri dishes in order to achieve 70-90 % confluency in RPMI media with 10 % FCS.

pcDNA3.1+ plasmid was mixed and diluted up to 1500 µL OptiMEM, vortexed and incubated at RT for 5 min. 60 µL Lipofectamine 2000 (Invitrogen) was diluted up to 1500 µL in OptiMEM, vortexed and incubated at RT for 5 min. Both the plasmid mixture and the Lipofectamine 2000 mixture were mixed, vortexed and incubated for 20 min at RT. The media of the HEK293 cells was replaced with OptiMEM. The mixture was then added dropwise to cells with swirling. After 5 h the cell media was replaced with RPMI (10 % FCS and L-glutamine). The cells were then incubated for 48 h before being harvested and lysed.

6.2.10 Cell Proliferation Assay

Cell proliferation assays were carried out using CellTiter 96® AQueous One Solution Cell Proliferation Assay (Promega). The assay was carried out in triplicate or more with 2000 cells seeded per well in a 96 well plate. 10 µL MTS reagent was added to each well and incubate for 1 h before absorbance was recorded at 490 nm. The cells were allowed to adhere for 4 h before the Day 0 reading was taken.

6.2.11 Spheroid Invasion Assay

5000 cells per well were seeded in a Corning® Ultra-Low adherence 96 well plate (100 µL media per well) and left to grow for 5 days. On day 5 the spheroids were serum starved overnight by removing 50 µL of media from the well and replacing with 50 µL serum free media, this wash was repeated six times. The following day (experimental day 0) rOPCML in full serum media was added to the spheroids and left to incubate for 3 h. 50 µL matrigel was added to each well and left for 30 min. 50 µL full serum media was then added on top of the matrigel. Images were then taken every 24 h on a Nikon COOLPIX P5100 camera attached to a microscope on a 4 X objective. The volume of spheroids was calculated manually using Image J. Each spheroid treatment was repeated in triplicate.

6.3 Baculovirus Production Methods

6.3.1 Cell Lines

Sf9 and HighFive cells were purchased from Life Technologies and adapted to Lonza InsectXpress media purchased from Lonza (Slough, UK). Cells were grown in media supplemented with sodium benzylpenicillin (30 µg/mL) and streptomycin sulphate (50 µg/mL). All medium for insect cells was serum free.

6.3.2 Generating Recombinant Bacmid

Baculovirus production was done as per the Bac-to-Bac Manual from Invitrogen. Briefly the FastBac vector containing OPCML was transformed as per the manufacturer's instructions into MAX Efficiency™ DH10Bac™ Competent Cells (Invitrogen) and bacmid was obtained by PureLink™ HiPure Plasmid Maxiprep Kit (Invitrogen). The recombinant bacmid was analysed via PCR using pUC/M13 Forward and Reverse primers. Initial denaturation was for 3 min at 93 °C for 1 cycle. Samples were denaturation for 45 sec at 94 °C, annealed for 45 sec at 55 °C and extension occurred for 5 min at 72 °C for 30 cycles. The final extension was 7 min at 72 °C. AccuPrime Pfx polymerase and Q5 High Fidelity polymerase was used following the manufacturer's instructions. 10 following the manufacturer's instructions. 10 µM of each primer was used.

6.3.3 Generating OPCML Baculovirus

Sf9 cells were seeded at 0.5×10^6 cell/mL in a 6 well plate and left to adhere before a medium caged into antibiotic free medium. 5 µL bacmid (200 ng/mL) was added to 100 µl medium. In a second tube 10 µL Cellfectin™ II Reagent (Gibco) was added to 100 µL medium. The bacmid solution and Cellfectin™ solution were then mixed together and incubated for 20 min before being added dropwise to the Sf9 cells. The transfection was left for 72 h then the viral supernatant was harvested and used directly to infect a 100 mL suspension culture of Sf9 cells at 2×10^6 cells/mL to amplify the virus. The viral supernatant was harvested after 72 h and kept refrigerated until further use.

6.4 OPCML Purification from Insect Cells and protein modifications

Thrombin Active (High Activity) from bovine plasma was purchased from Sigma-Aldrich, UK. Protein A agarose (ab193254) was purchased from Addgene. Thrombin CleanCleave™ Kit from Sigma Aldrich, UK. Pierce™ Monomeric Avidin Agarose was purchased from Thermo Fisher.

Protein concentration was determined using a NanoDrop One spectrophotometer and molecular weight and extinction coefficients calculated from Peptide Property Calculator (PepCalc) Innovagen.

6.4.1 Affinity Purification from Cell Supernatant

The method was adapted from Birtley et al³³. HighFive cells were used for protein production. HighFive cells were grown in suspension and infected with virus at 2×10^6 cells/mL (20 mL virus for 500 mL cell suspension). HighFive supernatant was harvested after 72 h by centrifugation at 3000 rpm for 20 min. PMSF was added to a final concentration of 0.2 mM. The pH was adjusted to 7.0 with NaOH then the supernatant was filtered to remove insoluble material (0.22 micron). Protein A Agarose (Abcam) was added to the filtered supernatant (1 mL resin per litre) and left to bind overnight at 4 °C. The resin was then bound for a further hour with shaking at RT before being added to a gravity flow column. The column was washed with 30 BV wash buffer (50 mM Tris, 150 mM NaCl, pH 7.5) and the protein eluted in 1 mL fractions with 15 BV elution buffer (50 mM glycine, pH 3.0). The eluted fractions were neutralised with 100 µL neutralisation buffer (2 M Tris, pH 8.5). The fractions were concentrated then buffer exchanged into wash buffer, with PD-10 Desalting Columns contain Sephadex G-25 resin, 1.0–2.5 mL samples (GE Healthcare). Following buffer exchange the protein was concentrated again, 10 % glycerol was added prior to freezing at -80 °C.

6.4.2 Anion Exchange

Anion exchange was used to purify OPCML after thrombin cleavage. Thrombin cleavage was carried out in wash buffer (50 mM Tris, 150 mM NaCl, pH 7.5) with 10 units of Thrombin Active for every 1 mg of OPCML, the cleavage was left mixing overnight at 4°C. The following day the cleavage mixture was diluted with 50 mM Tris to a buffer concentration of 50 mM Tris, 50 mM NaCl. Anion exchange was carried out using a HiTrap ANX Sepharose FF 1 mL column and an ÄKTA system (GE Healthcare). Briefly the column was washed with 20 mL degassed filtered water, then washed with 20 mL elution buffer (50 mM Tris, 1 M NaCl) to ensure no impurities

Methods

were present in the column from previous use. The column was then equilibrated with 20 mL start buffer (50 mM Tris, 50 mM NaCl). The cleavage reaction mix was then added to the column using a 10 mL loop and 2.5 mL fractions collected. Unbound protein eluted first and after this the elution gradient was applied over 50 CV (50 mL) to a final concentration of 50 % elution buffer. Followed by a final wash of 100 % elution buffer. All washes were carried out at a flow rate of 1 mL/min. Fractions were assessed for purity via SDS-PAGE gel and stained with Coomassie for visualisation, pooled then concentrated before freezing at -80°C .

6.4.3 Size Exclusion Chromatography

Preparative runs were carried out using HiLoad 16/600 Superdex 75 pg or HiLoad 16/600 Superdex 200 pg with an ÄKTA system (GE Healthcare). A flow rate of 1 mL/min was used in each case with a pressure limit set to 0.5 MPa. Analytical runs were run on a Superdex 200 Increase 10/300 GL from GE Healthcare all used with an ÄKTA system (GE Healthcare), with a variable flow rate of 0.5 mL/min to keep a set pressure of 0.9 MPa.

For both preparative and analytical runs protein was concentrated to a volume of 10 mL or less in 50 mM Tris, 150 mM NaCl buffer. The column was washed with 1 CV filtered and degassed water followed by 2 CV filtered and degassed gel filtration buffer (50 mM Tris, 150 mM NaCl). Protein sample was then loaded onto the column using a loop of the appropriate size (ranging from 500 μL to 10 mL depending on protein volume). Protein was eluted using 1.1 CV gel filtration buffer and fractions collected. Fractions were assessed for purity via SDS-PAGE gel and stained with coomassie for visualisation, pooled then concentrated before freezing at -80°C .

6.4.4 Peptide Mass Fingerprint Analysis

Samples for PMF analysis were run on an SDS-PAGE gel and stained with coomassie. The gels were stained for at least an hour then washed with water overnight before the samples were excised using scalpels. Excised samples were placed in Low Protein Binding microcentrifuge tubes and incubated for 5 min at RT with 150 μL buffer B1 (v/v 50 % acetonitrile, 50 % milliQ water). The supernatant was then removed and the sample incubated at RT for 30 min with buffer B2 (v/v 50 % acetonitrile, 50 % 50 mM AMBIC). The supernatant was removed and the sample was incubated at RT for 30 min with buffer B3 (v/v 50 % acetonitrile, 50 % 10 mM AMBIC). The supernatant was removed and gel pieces dried in a SpeedVac. Samples were then stored at -20°C until ready for digestion. To digest the samples they were first incubated for 10 min at RT in 10 μL buffer containing 15 μL 10 mM AMBIC with 1 μL trypsin solution (20 μg trypsin in 100 μL 50 mM AMBIC). Samples were then incubated overnight at 37°C in 20 μL 10 mM AMBIC. The supernatant was then used for MALDI-TOF measurements using 1 μL supernatant

Methods

with 1 μ L α -Cyano-4-hydroxycinnamic acid matrix on a 4800 MALDI-TOF-TOFs (AB Sciex) The data was analysed using mMass.

6.4.5 ZipTip De-salting

ZipTip C18 were used with wetting solution (acetonitrile 100 %), wash solution (0.1 % TFA in MilliQ water) and elution solution (60 % acetonitrile, 40 % MilliQ with 0.1 % TFA). The ZipTip was washed once with wetting solution followed by three washes with washing solution. Sample was added to the Ziptip followed by three more washed with washing buffer. The sample was then eluted with 4 μ L elution solution.

6.4.6 Free Cysteine Capping

Proteins were reduced with DTT (10 mM) for 15 min. Reducing agent was then removed using a disposable PD-10 Desalting Column. The free cysteines were then capped with iodoacetamide (20 mM) for 30 min at RT protected from light. Iodoacetamide was removed using a disposable PD-10 Desalting Column.

6.4.7 Intein-mediated Protein Ligation (IPL) Cloning

IPL was carried out using the IMPACT™ kit from New England Biolabs following the manufacturer's instructions. Briefly primers were designed to add NdeI and SapI restriction sites into the FL, and D3 entry clone constructs purchased from Life Technologies (Carlsbad, California, USA).

Restriction Site	Primer Sequence
NdeI FL	5' – T GGT <u>CAT ATG</u> GGA GTG CCC GTG CGC AGC GGA – 3'
NdeI D3	5' – T GGT <u>CAT ATG</u> GTA AAC TAT CCT CCC TAT ATC – 3'
SapI	5' – GGT <u>TGC TCT TCC</u> GCA GTT TAC ACC ATC AAT GAC TGC – 3'

Table 1 Primers designed following instructions from the IMPACT manual. The underlined bases correspond to the restriction enzyme. The nucleotides at the 5' end before the site are random and are to facilitate efficient cleavage of the DNA by restriction enzyme digest.

Accuprime Polymerase (Thermo Fisher) was used to clone the restriction sites into the constructs. Gradient PCR was used to determine the successful annealing temperatures for each primer.

Step	Temp (°C)	Time (min)
Initial Denaturation	95	5
30 Cycles	95	1
	Annealing Temp (FL = 52, D3 = 60)	1
	68	1 min per kb (110)
Final Extension	68	10
Hold	4	∞

Table 2 PCR conditions for generating OPCML constructs for IPL.

After PCR the DNA was purified using a Gel Extraction Kit (Qiagen) following the manufacturer's instructions. The pTXB1 vector was digested for 4 h at 37 °C, Shrimp Alkaline Phosphatase was added to prevent ligation of the linearised DNA plasmid and subsequently removed by heat activation at 65 °C for 20 min. Ready-To-Go T4 DNA Ligase (GE Healthcare, Little Chalfont, UK) was used to ligate the OPCML constructs into the pTXB1 vector.

6.4.8 Bacteria Expression system for Intein mediated Protein Ligation

Intein Mediated Protein Ligation was carried out using the IMPACT™ Kit from New England Biolabs as per the manufacturer's instructions. Briefly, FL and D3 constructs were cloned into a pTXB1 vector using custom designed primers (New England Biolabs) and transformed into BL21(DE3) competent cells and glycerol stocks were made. After initial tests to work out the optimal time for protein expression, 20 mL of culture in LB media grown overnight from glycerol stock was added to 1 litre LB media with ampicillin (100 mg/mL) and grown until OD₆₀₀ reached between 0.5-0.7. The culture was then induced with 0.5 mM IPTG (Sigma Aldrich) and incubated at 37 °C for 4 h, then pelleted, re-suspended and lysed by sonication on ice. Deoxyribonuclease I from bovine pancreas (Sigma Aldrich) was added to the suspension and mixed for 15 mins prior to centrifugation. The supernatant was added to the column to undergo affinity purification and on column cleavage as the manual instructs. All column steps were undertaken at 4 °C.

6.4.9 Sortagging

The sortagging reaction was set up with 10 µM OPCML protein, 1 µM sortase, 100 µM anchor mimetic and 0.05 % DDM in the sortase reaction buffer (150 mM NaCl, 20 mM Tris, 10 mM CaCl₂). The CaCl₂ was added as a solution of 100 µM to the final volume to ensure that the concentration of CaCl₂ was not reduced by the other reaction components. The reaction was incubated at various temperatures and times (see Results chapters). The reaction was stopped by addition of EDTA (5.5 mM final concentration) or by precipitation.

6.4.10 Protein Precipitation

A 2:0.5:1 mixture of MeOH:CHCl₃:H₂O was added to the protein sample and centrifuge for 2 min at maximum speed on a benchtop centrifuge. The upper phase of liquid was then removed to leave a protein disk which was re-suspended in an excess of MeOH. The sample was centrifuged again at maximum speed for 5 min. All the supernatant was then removed and the protein pellet left to dry before further use.

6.4.11 CuAAC reaction on Anchor Mimetics

After protein samples were prepared by precipitation they were re-suspended in 0.1 % SDS in PBS. The reaction was carried out with final concentrations of 100 µM YnTB, 1 mM CuSO₄, 1 mM TCEP, 100 µM TBTA and left shaking for 1 h at RT. The sample was then precipitated (6.4.10) and re-suspended in sample buffer for later analysis.

6.4.12 Pulldown

50 µL bead suspension was used per 1 mg of protein. The beads centrifuged at 3000 g for 2 min to pellet them then vortexed for 1 min to wash with 10 bed volumes of 0.2 % SDS in PBS three times. The sample was added to the beads and incubated for 1 hour at RT with gentle shaking. The supernatant was then removed and kept for analysis and the beads washed with 0.2 % SDS in PBS three times as before and kept for analysis by SDS-PAGE.

6.5 Synthesis Methods

All chemicals were purchased from Sigma-Aldrich Ltd (Gillingham, UK), Apollo Scientific (Stockport, UK), Acros Organics (Geel, Belgium) and Alfa Aesar (Heysham, UK) and used without further purification. Compounds were purified and analysed on an LC-MS system equipped with both an XBridge prep C18 5 µm, 19 × 100 mm OBD column and an XBridge C18 5 µm, 4.6 × 100 mm column. The LC-MS comprises of Waters 2767 Sample Manager, Waters SFO System Fluids Organiser, Waters 2545 Quaternary Gradient Module and a Waters 3100 Mass Detector from Waters (Elstree, Herts, UK). Unless specified otherwise, all compounds were separated over a gradient of acetonitrile in water (5–98% over 12 minutes then 98% acetonitrile for 3 minutes), both containing 0.1% formic acid. Alternative gradient elutions started from 20 or 50% acetonitrile for the same time periods.

6.5.1 General SPPS Protocols

Peptides were synthesis via SPPS using an Fmoc protecting strategy on an automated peptide synthesiser ResPep SL from Intavis (Cologne, Germany) using a mini column module on a scale

Methods

of 20 μmol per column using the Multipep software. Fmoc- $\epsilon\text{Ahx-OH}$ and Biotin-PEG NovatagTM resin were bought from Novabiochem, Merck Millipore (Massachusetts, USA). Fmoc-Lys(N_3)-OH, N-Boc-2-propargyl-L-glycine, Fmoc-Lys(ivDde) and Boc-Gly-Gly-Gly were bought from Fluorochem (Hadfield, UK). Fmoc-Photo-Leucine was purchased from Iris Biotech (Kortrijk, Belgium). All other Fmoc protected amino acids, NMP, HATU and HBTU were obtained from Cambridge Reagents (Hessle, UK) or AGTC Bioproducts (Hessle, UK). TentaGel rink amide resin was bought from Rapp Polymere (Tübingen, Germany).

6.5.2 Automated SPPS

Low-water DMF was used throughout. Natural amino acids were dissolved in NMP, unnatural amino acids were dissolved in DMF. Per mini column 20 μmol resin was used with the total number of columns used per peptide dependent on the scale of synthesis. Reagent stocks were made to 0.5 M concentration. Natural amino acids used HBTU for activation, unnatural amino acids used HATU. Each coupling reaction contained resin (20 μmol , 1 eq), Fmoc-amino acid (4.4 eq), HBTU or HATU (4.25 eq) and NMM (10 eq). Prior to automated synthesis the resins were swelled in DMF for 30 min. The remainder of the synthesis was automated. The first step was Fmoc deprotection using 20 % piperidine (v/v) in DMF (400 μL , 2 x 5 min). The deprotected amino acid was activated with HBTU and NMM before coupling. Each coupling step was undertaken twice and took between 30-45 min. Unreacted N-terminal amino acids were 'capped' by acetylating with 5 % acetic anhydride (v/v) in DMF (400 μL , 10 min). DMF washes occurred between each deprotection, coupling and 'capping'. The final amino acid was Fmoc deprotected.

After synthesis resins were transferred into a fritted syringe fitted and stored in a vacuum desiccator until further modification. One modifications were complete peptides were purified via LCMS then freeze dried on an Aplpha2.4 LDplus from Christ (Osterode, Germany).

The side chain functionalities of amino acids used are;

Lys(Boc), Ser(tBu), Asp(OtBu), Arg(pbf)

Unnatural amino acids;

Lys(N_3), Lys(ivDde), N-Boc-2-propargyl-L-glycine, Fmoc-Photo-Leucine

6.5.3 Cleavage

6.5.3.1 Mini-Cleavage of crude peptides

Mini-cleavage was undertaken after each peptide modification before proceeding to the next step. 100 μ L of cleavage cocktail (95 % TFA, 2.5 % TIS, 2.5 % distilled water) was added to a small amount of resin in a 1.5 mL centrifuge tube. The resin was shaken at RT for 1 h. For MALDI 10 μ L of the mixture was taken and added to 10 μ L distilled water. The rest of the cocktail was dried under a nitrogen stream then analysed by LC-MS.

6.5.3.2 Cleavage

After analysis to ensure all peptide has been modified all peptide was cleaved from the resin. The resin was left in the fritted syringe. 4 mL cleavage cocktail (95 % TFA, 2.5 % TIS, 2.5 % distilled water) was added to the syringe and the mixture shaken at RT for 2 h. The cleavage cocktail was reduced to around 1 mL under a nitrogen stream then ice cold diethyl ether was added in excess (20 mL). The solution was then centrifuged for 15 min at 3113 x g (4000 rpm) on a table top Centrifuge 5810 R from eppendorf (Hamburg, Germany). The supernatant was discarded and the pellet dried *in vacuo* for at least 10 min before purification by LC-MS.

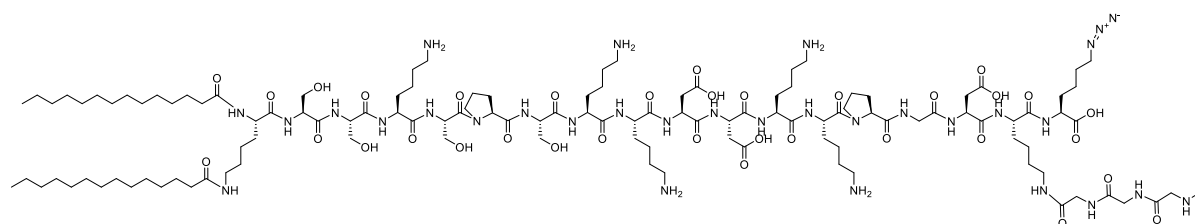
6.5.4 IvDde deprotection

IvDde deprotection was undertaken using 4 mL 5 % Hydrazine Monohydrate in DMF. The hydrazine was added to the resin in a fritted syringe and the mixture shaken at RT for 20 min. This step was repeated four times. The resin was then washed with DMF, DCM, MeOH and DCM. After deprotection a mini-cleavage and LC-MS analysis was undertaken to ensure complete deprotection before proceeding.

6.5.5 BOC Deprotection of commercial BOC-GGG-OH

Deprotection of BOC-GGG-OH was carried out with 50% TFA in DCM for 30 min at RT.

6.5.6 SPPS of anchor 3



Methods

Anchor 3 was synthesised on an automated peptide synthesiser on a scale of 120 μmol . Rink amide resin (0.24 mmole/g, 498 mg, 120 μmol) was swollen in DMF for 30 min prior to automated peptide synthesis. From N-terminal to C-terminal the sequence synthesised was Fmoc-Lys(Fmoc)-Ser-Ser-Lys-Ser-Pro-Ser-Lys-Lys-Asp-Asp-Lys-Lys-Pro-Gly-Arg-Lys(ivDde)-Lys(N_3). Following deprotection of the Fmoc groups on the N-terminal lysine myristic acid (82 mg, 0.36 mmol, 3 eq), HOBt (47 mg, 0.36 mmol, 3 eq), and DIC (45 mg, 0.36 mmol, 3 eq) were dissolved in DMF (4 mL) before being added to the resin bound peptide (120 μmol , 1eq) and the reaction mixture microwaved at 60°C for 20 min then washed with DMF, DCM, MeOH and DCM. This step was done twice. After the second myristoylation a mini-cleavage was performed to ensure di-myristoylation of the peptide. The peptide was then ivDde deprotected using hydrazine monohydrate. Boc-Gly₃-OH (104 mg, 0.36 mmol, 3 eq) and HATU (137 mg, 0.36 mmol, 3 eq) were dissolved in DMF (4 mL) then DIPEA (118 μL , 0.72 mmol, 6 eq) was added to the solution before the reaction mix was added to the resin bound peptide (120 μmol , 1 eq). The reaction was shaken at RT for 2 h. The peptide was washed with DMF, DCM, MeOH and DCM dried and analysed by LSMS after mini-cleavage. The peptide was then cleaved from the resin with 95 % TFA, 2.5 % TIS, 2.5 % distilled water then purified and analysed on an LC-MS system with gradient elutions starting from 20 % acetonitrile. The product was obtained by lyophilisation as a white powder (13 mg, 4 % yield), calculated for $\text{C}_{118}\text{H}_{209}\text{N}_{31}\text{O}_{34}$ $m/z = 2605.56$.

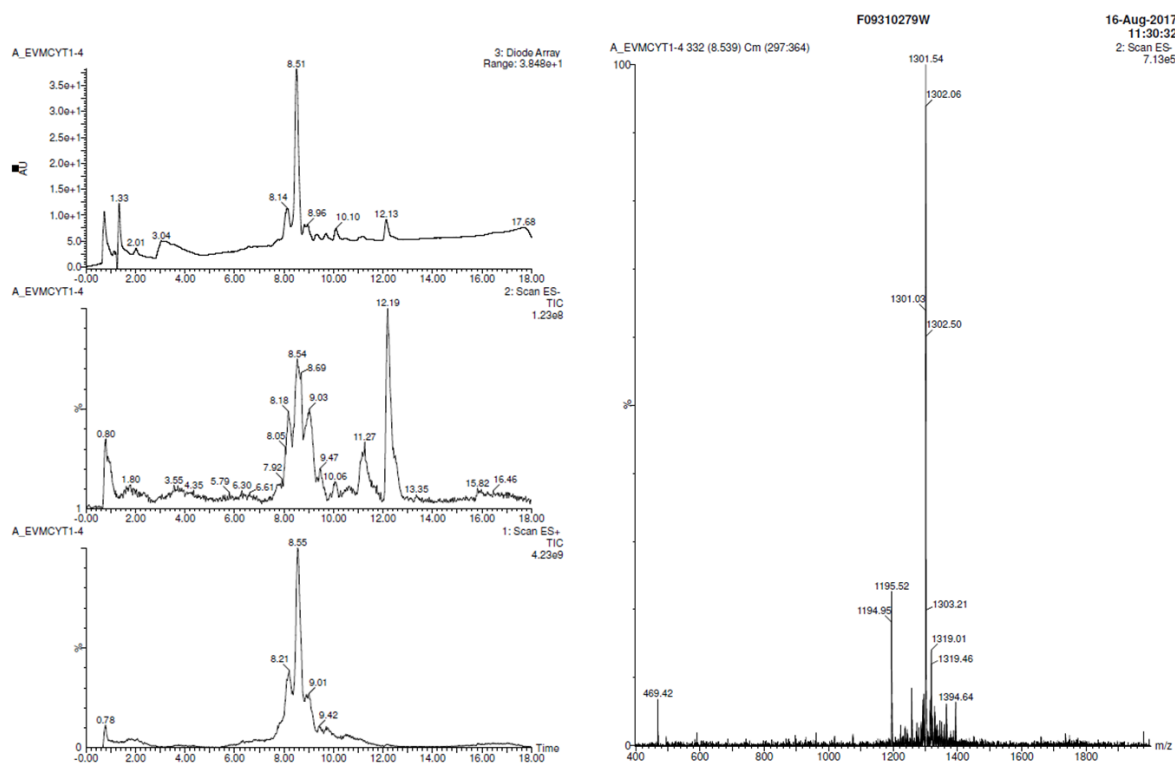
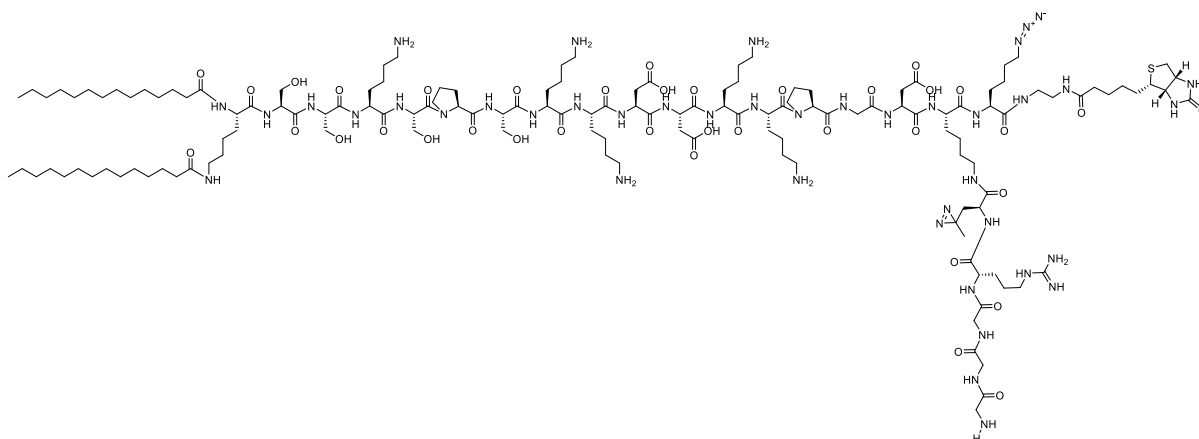


Figure 53 Anchor 3 LC trace and mass spec



Boc-Gly₃-OH (104 mg, 0.36 mmol, 3 eq) and HATU (137 mg, 0.36 mmol, 3 eq) were dissolved in DMF (4 mL) then DIPEA (118 µL, 0.72 mmol, 6 eq) was added to the solution before the reaction mix was added to the resin bound peptide (120 µmol, 1 eq). The reaction was shaken at RT for 2 h. The peptide was washed with DMF, DCM, MeOH and DCM dried and analysed by LSMS after mini-cleavage. The peptide was then cleaved from the resin with 95 % TFA, 2.5 % TIS, 2.5 % distilled water then purified and analysed on an LC-MS system with gradient elutions starting from 20 % acetonitrile. The product was obtained by lyophilisation as a white powder (74 mg, 24 % yield), calculated for C₁₂₉H₂₂₈N₃₈O₃₆ m/z = 3154.86.

Methods

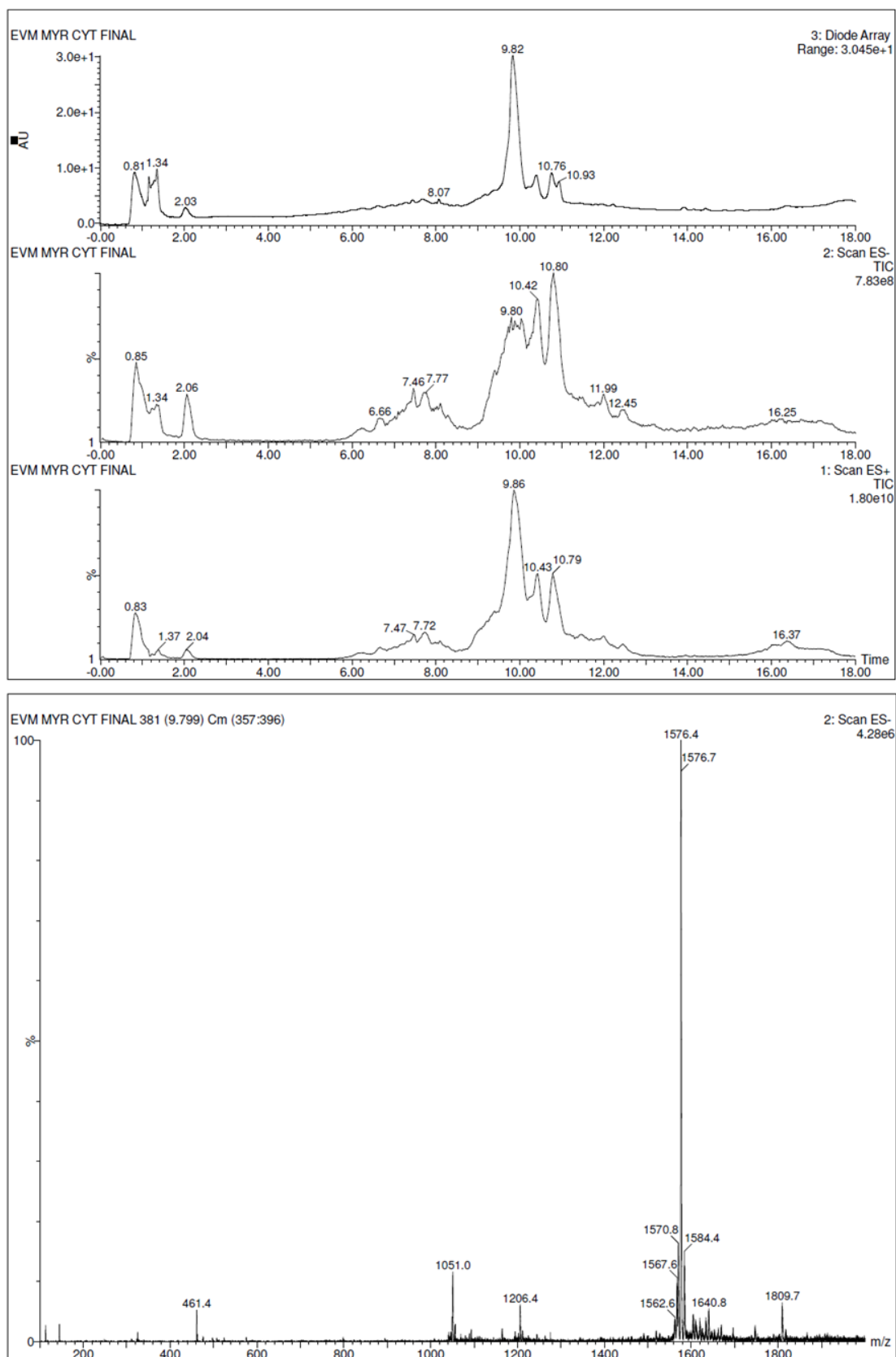


Figure S4 Anchor 4 LC trace and mass spec

Bibliography

1. Heal, W. P., Wright, M. H., Thinon, E. & Tate, E. W. Multifunctional protein labeling via enzymatic N-terminal tagging and elaboration by click chemistry. *Nat. Protoc.* **7**, 105–17 (2012).
2. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2018. *CA. Cancer J. Clin.* **68**, 7–30 (2018).
3. Alberts, B. *et al.* *Molecular Biology of the Cell*, 5th Ed.
4. Sever, R. & Brugge, J. S. Signal Transduction in Cancer. *Mutat. Res. Mol. Mech. Mutagen.* **555**, 1–2 (2004).
5. Sellar, G. C. *et al.* OPCML at 11q25 is epigenetically inactivated and has tumor-suppressor function in epithelial ovarian cancer. *Nat. Genet.* **34**, 337–343 (2003).
6. Cancer Research UK. <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer>, Accessed January 2019
7. Cancer Research UK. <https://Cancer Research UK, www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer/incidence#heading-Four>, Accessed January 2019
8. Auersperg, N., Wong, A. S. T., Choi, K., Kang, S. K. & Leung, P. C. K. Ovarian Surface Epithelium : Biology , Endocrinology ,. *Endocr. Rev.* **22**, 255–288 (2001).
9. Auersperg, N., Ota, T. & Mitchell, G. W. E. Early events in ovarian epithelial carcinogenesis: Progress and problems in experimental approaches. *Int. J. Gynecol. Cancer* **12**, 691–703 (2002).
10. Hall, J., Paul, J. & Brown, R. Critical evaluation of p53 as a prognostic marker in ovarian cancer. *Expert Rev. Mol. Med.* **6**, 1–20 (2004).
11. Paul, A. & Paul, S. The breast cancer susceptibility genes (BRCA) in breast and ovarian cancers. *Front. Biosci.* **19**, 605–618 (2014).
12. Meng, Q., Xia, C., Fang, J., Rojanasakul, Y. & Jiang, B. H. Role of PI3K and AKT specific isoforms in ovarian cancer cell migration, invasion and proliferation through the p70S6K1 pathway. *Cell. Signal.* **18**, 2262–2271 (2006).

Bibliography

13. Le Page, C. *et al.* From gene profiling to diagnostic markers: IL-18 and FGF-2 complement CA125 as serum-based markers in epithelial ovarian cancer. *Int. J. Cancer* **118**, 1750–1758 (2006).
14. Trinh, X. B. *et al.* The VEGF pathway and the AKT/mTOR/p70S6K1 signalling pathway in human epithelial ovarian cancer. *Br. J. Cancer* **100**, 971–978 (2009).
15. Lafky, J. M., Wilken, J. A., Baron, A. T. & Maihle, N. J. Clinical implications of the ErbB/epidermal growth factor (EGF) receptor family and its ligands in ovarian cancer. *Biochim. Biophys. Acta - Rev. Cancer* **1785**, 232–265 (2008).
16. Monk, B. J., Minion, L. E. & Coleman, R. L. Anti-angiogenic agents in ovarian cancer: Past, present, and future. *Ann. Oncol.* **27**, i33–i39 (2016).
17. Gloss, B. S. & Samimi, G. Epigenetic biomarkers in epithelial ovarian cancer. *Cancer Lett.* **342**, 257–263 (2014).
18. Colombo, N., Lorusso, D. & Scollo, P. Impact of recurrence of ovarian cancer on quality of life and outlook for the future. *Int. J. Gynecol. Cancer* **27**, 1134–1140 (2017).
19. Orr, B. & Edwards, R. P. Diagnosis and Treatment of Ovarian Cancer. *Hematol. Oncol. Clin. North Am.* **32**, 943–964 (2018).
20. M.J., D. *et al.* CA125 in ovarian cancer: European Group on Tumor Markers guidelines for clinical use. *Int. J. Gynecol. Cancer* **15**, 679–691 (2005).
21. Bast, R. C., Hennessey, B. & Mills, G. B. The biology of ovarian cancer: New opportunities for translation. *Nat. Rev. Cancer* **9**, 415–428 (2009).
22. Cortez, A. J., Tudrej, P., Kujawa, K. A. & Lisowska, K. M. Advances in ovarian cancer therapy. *Cancer Chemother. Pharmacol.* **81**, 17–38 (2018).
23. Moufarrij, S. *et al.* Epigenetic therapy for ovarian cancer: Promise and progress. *Clin. Epigenetics* **11**, 1–11 (2019).
24. Perren, T. J. *et al.* A Phase 3 Trial of Bevacizumab in Ovarian Cancer. *N. Engl. J. Med.* **365**, 2484–2496 (2011).
25. Aghajanian, C. *et al.* OCEANS: A randomized, double-blind, placebo-controlled phase III trial of chemotherapy with or without bevacizumab in patients with platinum-sensitive recurrent epithelial ovarian, primary peritoneal, or fallopian tube cancer. *J. Clin. Oncol.* **30**, 2039–2045 (2012).

Bibliography

26. Audeh, M. W. *et al.* Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: A proof-of-concept trial. *Lancet* **376**, 245–251 (2010).
27. Fong, P. C. *et al.* Inhibition of Poly(ADP-Ribose) Polymerase in Tumors from BRCA Mutation Carriers Peter. *N. Engl. J. Med.* **359**, 229–241 (2004).
28. Walters, C. L., Arend, R. C., Armstrong, D. K., Naumann, R. W. & Alvarez, R. D. Folate and folate receptor alpha antagonists mechanism of action in ovarian cancer. *Gynecol. Oncol.* **131**, 493–8 (2013).
29. Graybill, W., Sood, A. K., Monk, B. J. & Coleman, R. L. State of the science: Emerging therapeutic strategies for targeting angiogenesis in ovarian cancer. *Gynecol. Oncol.* **138**, 223–226 (2015).
30. Franzese, E. *et al.* PARP inhibitors in ovarian cancer. *Cancer Treat. Rev.* **73**, 1–9 (2019).
31. Cui, Y. *et al.* OPCML is a broad tumor suppressor for multiple carcinomas and lymphomas with frequently epigenetic inactivation. *PLoS One* **3**, e2990 (2008).
32. McKie, A. B. *et al.* The OPCML tumor suppressor functions as a cell surface repressor-adaptor, negatively regulating receptor tyrosine kinases in epithelial ovarian cancer. *Cancer Discov.* **2**, 156–171 (2012).
33. Birtley, J. *et al.* Clinical inactivating mutations of the tumour suppressor OPCML reveal new cancer associated-functions. *Nat Comms, In Press* (2019).
34. Jones, P. A. & Baylin, S. B. The Fundamental Role of Epigenetic Events in Cancer. *Nat. Rev. Genet.* **3**, 415–428 (2002).
35. Zhang, C. *et al.* Meta-analysis of DNA methylation biomarkers in hepatocellular carcinoma. *Oncotarget* **7**, 81255–81267 (2016).
36. Zhou, F. *et al.* Methylation of OPCML promoter in ovarian cancer tissues predicts poor patient survival. *Clin. Chem. Lab. Med.* **52**, 735–742 (2014).
37. Amornpisutt, R., Proungvitaya, S., Jearanaikoon, P. & Limpaiboon, T. DNA methylation level of OPCML and SFRP1: a potential diagnostic biomarker of cholangiocarcinoma. *Tumor Biol.* **36**, 4973–4978 (2015).
38. Li, C. *et al.* OPCML is frequently methylated in human colorectal cancer and its restored expression reverses EMT via downregulation of smad signaling. *Am. J. Cancer Res.* **5**,

- 1635–1648 (2015).
39. Mei, F. C., Young, T. W., Liu, J. & Cheng, X. RAS-mediated epigenetic inactivation of OPCML in oncogenic transformation of human ovarian surface epithelial cells. *FASEB J.* **20**, 497–499 (2006).
40. Reed, J. E. *et al.* Expression of cellular adhesion molecule ‘OPCML’ is down-regulated in gliomas and other brain tumours. *Neuropathol. Appl. Neurobiol.* **33**, 77–85 (2007).
41. Wang, B. *et al.* Detection of OPCML methylation, a possible epigenetic marker, from free serum circulating DNA to improve the diagnosis of early-stage ovarian epithelial cancer. *Oncol. Lett.* **14**, 217–223 (2017).
42. Wu, Y. *et al.* Methylation profiling identified novel differentially methylated markers including OPCML and FLRT2 in prostate cancer. *Epigenetics* **11**, 247–258 (2016).
43. Xing, X. *et al.* Down-regulated expression of OPCML predicts an unfavorable prognosis and promotes disease progression in human gastric cancer. *BMC Cancer* **17**, 1–10 (2017).
44. Ye, F., Zhang, S.-F., Xie, X. & Lu, W.-G. OPCML gene promoter methylation and gene expression in tumor and stroma cells of invasive cervical carcinoma. *Cancer Invest.* **26**, 569–74 (2008).
45. Thiagalingam, S. *et al.* Loss of heterozygosity as a predictor to map tumor suppressor genes in cancer: Molecular basis of its occurrence. *Curr. Opin. Oncol.* **14**, 65–72 (2002).
46. Lyko, F. & Brown, R. DNA methyltransferase inhibitors and the development of epigenetic cancer therapies. *J. Natl. Cancer Inst.* **97**, 1498–1506 (2005).
47. Dhanak, D. & Jackson, P. Development and classes of epigenetic drugs for cancer. *Biochem. Biophys. Res. Commun.* **455**, 58–69 (2014).
48. Kinoshita, T. Glycosylphosphatidylinositol (GPI) Anchors: Biochemistry and Cell Biology: Introduction to a Thematic Review Series. *J. Lipid Res.* **57**, 4–5 (2015).
49. Lee, M. J. & Yaffe, M. B. Protein regulation in signal transduction. *Cold Spring Harb. Perspect. Biol.* **8**, 1–20 (2016).
50. Saha, S., Anilkumar, A. A. & Mayor, S. GPI-anchored protein organization and dynamics at the cell surface. *J. Lipid Res.* **2**, 159–175 (2015).
51. Fujita, M. & Kinoshita, T. GPI-anchor remodeling: Potential functions of GPI-anchors in

- intracellular trafficking and membrane dynamics. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1821**, 1050–1058 (2012).
52. Simons, K. & Ikonen, E. Functional rafts in cell membranes. *Nature* **387**, 569–572 (1997).
 53. Zanini, E. *et al.* The tumor suppressor protein OPCML potentiates anti-EGFR and anti-HER2 targeted therapy in HER2-positive ovarian and breast cancer. *Mol. Cancer Ther.* **16**, 2246–2256 (2017).
 54. Antony, J. *et al.* The tumour suppressor OPCML promotes AXL inactivation by the phosphatase PTPRG in ovarian cancer. *EMBO Rep.* e45670 (2018). doi:10.15252/embr.201745670
 55. Nosjean, O., Briolay, A. & Roux, B. Mammalian GPI proteins : sorting , membrane residence and functions. *Biochem. Biophys. Acta* **1331**, 153–186 (1997).
 56. Paulick, M. G., Wise, A. R., Forstner, M. B., Groves, J. T. & Bertozzi, C. R. Synthetic Analogues of Glycosylphosphatidylinositol-Anchored Proteins and Their Behavior in Supported Lipid Bilayers Molecular Modeling Studies of the GPI Anchor Analogues. *J. Am. Chem. Soc.* **129**, 11543–11550 (2007).
 57. Chatterjee, S. & Mayor, S. The GPI-anchor and protein sorting. *Cell. Mol. Life Sci.* **58**, 1969–1987 (2001).
 58. Brown, D. A. & Rose, J. K. Sorting of GPI-Anchored Proteins to Membrane Subdomains during Transport to the Apical Cell Surface. **66**, 533–544 (1992).
 59. Heider, S., Dangerfield, J. A. & Metzner, C. Biomedical applications of glycosylphosphatidylinositol-anchored proteins. *J. Lipid Res.* **57**, 1778–1788 (2016).
 60. Rietveld, A., Ansorge, I., Simons, K. & Benting, J. Acyl and alkyl chain length of GPI-anchors is critical for raft association in vitro. *FEBS Lett.* **462**, 47–50 (1999).
 61. Malherbe, T., Boschung, M., Bu, P. & Roditi, I. GPI-anchored proteins : now you see ' em , now you don ' t. *Faseb* **15**, 545–548 (2000).
 62. Sharma, P. *et al.* Nanoscale Organization of Multiple GPI-Anchored Proteins in Living Cell Membranes. *Cell* **116**, 577–589 (2004).
 63. Paulick, M. G., Forstner, M. B., Groves, J. T. & Bertozzi, C. R. A chemical approach to unraveling the biological function of the glycosylphosphatidylinositol anchor. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 20332–20337 (2007).

Bibliography

64. Brodsky, R. A. *et al.* Purified GPI-Anchored CD4DAF as a Receptor for HIV-Mediated Gene Transfer. **1239**, 1231–1239 (1994).
65. Muenchmeier, N. *et al.* A Novel CXCL10-Based GPI-Anchored Fusion Protein as Adjuvant in NK-Based Tumor Therapy. *PLoS One* **8**, 1–12 (2013).
66. McHugh, R. S. *et al.* Construction, purification, and functional incorporation on tumor cells of glycolipid-anchored human B7-1 (CD80). *Immunology* **92**, 8059–8063 (1995).
67. Nagarajan, S. & Selvaraj, P. Glycolipid-anchored IL-12 expressed on tumor cell surface induces antitumor immune response. *Cancer Res.* **62**, 2869–2874 (2002).
68. Djafarzadeh, R. *et al.* GPI-anchored TIMP-1 treatment renders renal cell carcinoma sensitive to FAS-mediated killing. *Oncogene* **25**, 1496–1508 (2006).
69. Musiol, H. & Dong, S. Toward Semisynthetic Lipoproteins by Convergent Strategies Based on Click and Ligation Chemistry. *ChemBioChem* **6**, 625–628 (2005).
70. Ball, H. L., King, D. S., Cohen, F. E., Prusiner, S. B. & Baldwin, M. A. Engineering the prion protein using chemical synthesis. *J. Pept. Res.* **58**, 357–374 (2001).
71. Grogan, M. J., Kaizuka, Y., Conrad, R. M., Groves, J. T. & Bertozzi, C. R. Synthesis of Lipidated Green Fluorescent Protein and Its Incorporation in Supported Lipid Bilayers. *J. Am. Chem. Soc.* **127**, 14383–14387 (2005).
72. Hicks, M. R. *et al.* Synthesis and structural characterization of a mimetic membrane-anchored prion protein. *FEDS J.* **273**, 1285–1299 (2006).
73. Hill, A. *et al.* Protection of erythrocytes from human complement-mediated lysis by membrane-targeted recombinant soluble CD59: A new approach to PNH therapy. *Blood* **107**, 2131–2137 (2006).
74. Teruya, K., Nishizawa, K. & Doh-ura, K. Semisynthesis of a Protein with Cholesterol at the C-Terminal , Targeted to the Cell Membrane of Live Cells. *Protein J* **29**, 493–500 (2010).
75. Schafer, B. *et al.* Preparation of Semisynthetic Lipoproteins with Fluorescent Cholesterol Anchor and Their Introduction to the Cell Membrane with Minimal Disruption of the Membrane. *Bioconjug. Chem.* **24**, 1684–1697 (2013).
76. Teruya, K., Nishizawa, K. & Doh-ura, K. Semisynthesis of a protein with cholesterol at the C-terminal, targeted to the cell membrane of live cells. *Protein J.* **29**, 493–500 (2010).

Bibliography

77. Sato, S. B. *et al.* Distribution and transport of cholesterol-rich membrane domains monitored by a membrane-impermeant fluorescent polyethylene glycol-derivatized cholesterol. *J. Biol. Chem.* **279**, 23790–23796 (2004).
78. Webb, M. E., Bon, R. S. & Wright, M. H. in *Chemical and Biological Synthesis: Enabling Approaches for Understanding Biology* 313–356 (2018).
79. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes. *Angew. Chemie - Int. Ed.* **41**, 2596–2599 (2002).
80. Tornøe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-Triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **67**, 3057–3064 (2002).
81. Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem. Int. Ed. Engl.* **40**, 2004–2021 (2001).
82. Huisgen, R. 1,3-Dipolar Cycloadditions Past and Future. *Angew. Chem. Int. Ed. Engl.* **2**, 565–598 (1963).
83. Agard, N. J., Prescher, J. A. & Bertozzi, C. R. A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **126**, 15046–15047 (2004).
84. Wang, L., Jiang, R., Wang, L., Liu, Y. & Sun, X. L. Preparation of chain-end clickable recombinant protein and its bio-orthogonal modification. *Bioorg. Chem.* **65**, 159–166 (2016).
85. Khatwani, S. L. *et al.* Covalent protein-oligonucleotide conjugates by copper-free click reaction. *Bioorganic Med. Chem.* **20**, 4532–4539 (2012).
86. Duckworth, B. P., Xu, J., Taton, T. A., Guo, A. & Distefano, M. D. Site-specific, covalent attachment of proteins to a solid surface. *Bioconjug. Chem.* **17**, 967–974 (2006).
87. Duckworth, B. P., Zhang, Z., Hosokawa, A. & Distefano, M. D. Selective labeling of proteins by using protein farnesyltransferase. *ChemBioChem* **8**, 98–105 (2007).
88. Evans, T. C., Benner, J. & XU, M.-Q. Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* **7**, 2256–2264 (1998).
89. Southworth, M. W., Amaya, K., Evans, T. C., Xu, M. Q. & Perler, F. B. Purification of

Bibliography

- proteins fused to either the amino or carboxy terminus of the Mycobacterium xenopi gyrase A intein. *Biotechniques* **27**, 110–120 (1999).
90. Muir, T. W., Sondhi, D. & Cole, P. A. Expressed protein ligation: A general method for protein engineering. *PNAS* **95**, 6705–6710 (1998).
 91. Muir, T. W. Semisynthesis of Proteins by Expressed Protein Ligation. *Annu. Rev. Biochem.* **72**, 249–289 (2003).
 92. Xu, M.-Q. & Evans, T. C. Intein-Mediated Ligation and Cyclization of Expressed Proteins. *Methods* **24**, 257–277 (2001).
 93. Chong, S. *et al.* Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **192**, 271–281 (1997).
 94. Chong, S. *et al.* Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. *Nucleic Acids Res.* **26**, 5109–5115 (1998).
 95. Watanabe, T. *et al.* The roles of the C-terminal domain and type III domains of chitinase A1 from *Bacillus circulans* WL-12 in chitin degradation. *J. Bacteriol.* **176**, 4465–4472 (1994).
 96. Liu, X. Protein-Splicing Intein : Genetic Mobility , Origin , and Evolution. *Biochemistry* **34**, 61–76 (2000).
 97. Perler, F. B. & Allewell, N. M. Evolution, mechanisms, and applications of Intein-mediated protein splicing. *J. Biol. Chem.* **289**, 14488–14489 (2014).
 98. Shah, N. H. & Muir, T. W. Inteins: Nature’s Gift to Protein Chemists. *Chem. Sci.* **5**, 446–461 (2014).
 99. Chong, S., Williams, K. S., Wotkowicz, C. & Xu, M. Modulation of Protein Splicing of the *Saccharomyces cerevisiae* Vacuolar Membrane ATPase Intein. *J. Biol. Chem.* **273**, 10567–10577 (1998).
 100. Rak, A. *et al.* Structure of Rab GDP-Dissociation Inhibitor in Complex with Prenylated YPT1 GTPase. *Science (80-.).* **302**, 646–651 (2003).
 101. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. H. Synthesis of Proteins by Native Chemical Ligation. *Science (80-.).* **266**, 776–779 (1994).

Bibliography

102. Mao, H., Hart, S. a., Schink, A. & Pollok, B. a. Sortase-mediated protein ligation: a new method for protein engineering. *J. Am. Chem. Soc.* **126**, 2670–1 (2004).
103. Ton-That, H., Liu, G., Mazmanian, S. K., Faull, K. F. & Schneewind, O. Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 12424–9 (1999).
104. Ton-That, H., Mazmanian, S. K., Faull, K. F. & Schneewind, O. Anchoring of Surface Proteins to the Cell Wall of *Staphylococcus aureus*. *J. Biol. Chem.* **275**, 9876–9881 (2000).
105. Guimaraes, C. P. *et al.* Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. *Nat. Protoc.* **8**, 1787–1799 (2013).
106. Giesbrecht, P., Kersten, T. & Maidhof, H. Staphylococcal Cell Wall : Morphogenesis and Fatal Variations in the Presence of Penicillin. *Microbiol. Mol. Biol. Rev.* **62**, 1371–1414 (1998).
107. Huang, X. *et al.* Kinetic Mechanism of *Staphylococcus aureus* Sortase SrtA. *Biochemistry* **42**, 11307–11315 (2003).
108. Hirakawa, H., Ishikawa, S. & Nagamune, T. Design of Ca²⁺-independent *Staphylococcus aureus* sortase A mutants. *Biotechnol. Bioeng.* **109**, 2955–2961 (2012).
109. Stribis, K., Spooner, E. & Ploegh, H. L. Protein Ligation in Living Cells Using Sortase. *Traffic* **13**, 780–789 (2012).
110. Antos, J. M. *et al.* A straight path to circular proteins. *J. Biol. Chem.* **284**, 16028–16036 (2009).
111. Ritzeveld, M. Sortagging: a robust and efficient chemoenzymatic ligation strategy. *Chemistry* **20**, 8516–8529 (2014).
112. Williamson, D. J., Webb, M. E. & Turnbull, W. B. Dipeptide substrates for sortase-mediated N-terminal protein ligation. *Nat. Protoc.* **9**, 253–62 (2014).
113. Antos, J. M. *et al.* Site-Specific N- and C-Terminal Labeling of a Single Polypeptide Using Sortases of Different Specificity. *J. Am. Chem. Soc.* 10800–10801 (2009).
114. Chan, L. *et al.* Covalent attachment of proteins to solid supports and surfaces via sortase-mediated ligation. *PLoS One* **2**, 1–5 (2007).

Bibliography

115. Clow, F., Fraser, J. D. & Proft, T. Immobilization of proteins to biacore sensor chips using *Staphylococcus aureus* sortase A. *Biotechnol. Lett.* **30**, 1603–1607 (2008).
116. Tanaka, T., Yamamoto, T., Tsukiji, S. & Nagamune, T. Site-specific protein modification on living cells catalyzed by sortase. *ChemBioChem* **9**, 802–807 (2008).
117. Stefan, N. *et al.* Highly Potent, Anthracycline-based Antibody–Drug Conjugates Generated by Enzymatic, Site-specific Conjugation. *Mol. Cancer Ther.* **16**, 879–892 (2017).
118. Popp, M. W., Antos, J. M., Grotenbreg, G. M., Spooner, E. & Ploegh, H. L. Sortagging: a versatile method for protein labeling. *Nat. Chem. Biol.* **3**, 707–708 (2007).
119. Guo, X., Wang, Q., Swarts, B. M. & Guo, Z. Sortase-Catalyzed Peptide - Glycosylphosphatidylinositol Analogue Ligation. *J. Am. Chem. Soc.* 9878–9879 (2009).
120. Morecroft, E. *MRes Research Manuscript: The Tumour Suppressor Protein OPCML as an Anticancer drug : A Biological Therapeutic enabled through Self-splicing MRes Research Manuscript The Tumour Suppressor Protein OPCML as an Anticancer Drug : A Biological Therapeutic enab.* (2015).
121. Morecroft, E. V. N. *Early-Stage Assessment Report, The Tumour Suppressor Protein OPCML as an Anti-Cancer Drug : A Biological Therapeutic enabled through Self-Splicing GPI Mimetics.* (2017).
122. Thomas, C. E., Ehrhardt, A. & Kay, M. a. Progress and problems with the use of viral vectors for gene therapy. *Nat. Rev. Genet.* **4**, 346–58 (2003).
123. Mátrai, J., Chuah, M. K. L. & VandenDriessche, T. Recent advances in lentiviral vector development and applications. *Mol. Ther.* **18**, 477–490 (2010).
124. Campeau, E. *et al.* A Versatile Viral System for Expression and Depletion of Proteins in Mammalian Cells. *PLoS One* **4**, e6529 (2009).
125. Merten, O. W., Hebben, M. & Bovolenta, C. Production of lentiviral vectors. *Mol. Ther. - Methods Clin. Dev.* **3**, 16017 (2016).
126. Robbins, P. D. & Ghivizzani, S. C. Viral Vectors for Gene Therapy. *Pharmacol. Ther.* **80**, 35–47 (1998).
127. Liang, X., Peng, L., Baek, C. H. & Katzen, F. Single step BP/LR combined Gateway reactions. *Biotechniques* **55**, 265–268 (2013).

Bibliography

128. Landy, A. Dynamic, Strctural, and Regulatory Aspects Of Lambda Site-specific Recombination. *Annu. Rev. Biochem.* **58**, 913–949 (1989).
129. Hartley, J. L., Temple, G. F. & Brasch, M. A. DNA cloning using in vitro site-specific recombination. *Genome Res.* **10**, 1788–1795 (2000).
130. Bernard, P. & Couturier, M. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J. Mol. Biol* **226**, 735–745 (1992).
131. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in Escherichia coli: Advances and challenges. *Front. Microbiol.* **5**, 1–17 (2014).
132. Sahdev, S., Khattar, S. K. & Saini, K. S. Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies. *Mol. Cell. Biochem.* **307**, 249–264 (2008).
133. Ikonomou, L., Schneider, Y. J. & Agathos, S. N. Insect cell culture for industrial production of recombinant proteins. *Appl. Microbiol. Biotechnol.* **62**, 1–20 (2003).
134. Demain, A. L. & Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* **27**, 297–306 (2009).
135. Donahue, R. Rapid Generation of Recombinant Baculovirus and Expression of Foreign Genes using the Bac-To-Bac™ Baculovirus Expression System. *Focus (Madison).* **17**, 53–58 (1996).
136. Luckow, V. A., Lee, S. C., Barry, G. F. & Olins', P. O. Efficient Generation of Infectious Recombinant Baculoviruses by Site-Specific Transposon-Mediated Insertion of Foreign Genes into a Baculovirus Genome Propagated in Escherichia coli. *J. Virol.* **67**, 4566–4579 (1993).
137. Invitrogen. Bac-to-Bac® Baculovirus Expression System. 1–78 (2010).
138. Sutherland, R. M. *et al.* Oxygenation and Differentiation in Multicellular Colon Carcinoma1 Spheroids of Human Colon Carcinoma. *Cancer Res.* **46**, 5320–5329 (1986).
139. Hirschhaeuser, F. *et al.* Multicellular tumor spheroids : An underestimated tool is catching up again. *J. Biotechnol.* **148**, 3–15 (2010).
140. Vinci, M. *et al.* Advances in establishment and analysis of three-dimensional tumor spheroid-based functional assays for target validation and drug evaluation. *BMC Biol.* **10**, 29 (2012).

Bibliography

141. Berens, E. B., Holy, J. M., Riegel, A. T. & Wellstein, A. A Cancer Cell Spheroid Assay to Assess Invasion in a 3D Setting. *J. Vis. Exp.* 1–6 (2015). doi:10.3791/53409
142. Stein, A. M., Demuth, T., Mobley, D., Berens, M. & Sander, L. M. A mathematical model of glioblastoma tumor spheroid invasion in a three-dimensional in vitro experiment. *Biophys. J.* **92**, 356–365 (2007).
143. Sodek, K. L., Ringuelette, M. J. & Brown, T. J. Compact spheroid formation by ovarian cancer cells is associated with contractile behavior and an invasive phenotype. *Int. J. Cancer* **124**, 2060–2070 (2009).
144. Kramer, N. *et al.* In vitro cell migration and invasion assays. *Mutat. Res.* **752**, 10–24 (2013).
145. Ayala, I., Baldassarre, M., Caldieri, G. & Ā, R. B. Invadopodia : A guided tour. *Eur. J. Cell Biol.* **85**, 159–164 (2006).
146. Stylli, S. S., Kaye, A. H. & Lock, P. Invadopodia : At the cutting edge of tumour invasion. *J. Clin. Neurosci.* **15**, 725–737 (2008).
147. Morecroft, E. *Initial Research Plan - The Tumour Suppressor Protein OPCML as an Anti-Cancer Drug : A Biological Therapeutic enabled through Self-splicing Glycophosphatidylinositol Mimetics.* (2015).
148. Shi, J. *et al.* Engineered red blood cells as carriers for systemic delivery of a wide array of functional probes. *Proc. Natl. Acad. Sci.* **111**, 10131–10136 (2014).
149. Sarpong, K. & Bose, R. Efficient sortase-mediated N-terminal labeling of TEV protease cleaved recombinant proteins. *Anal. Biochem.* **521**, 55–58 (2017).
150. Antos, J. M., Truttmann, M. C. & Ploegh, H. L. Recent advances in sortase-catalyzed ligation methodology. *Curr. Opin. Struct. Biol.* **38**, 111–118 (2016).
151. Israelachvili, J. N. *Intermolecular and Surface Forces, Third Edition.*
152. Antos, J. M., Miller, G. M., Grotenbreg, G. M. & Ploegh, H. L. Lipid modification of proteins through sortase-catalyzed transpeptidation. *J. Am. Chem. Soc.* **130**, 16338–16343 (2008).
153. Spro, J. & Sinz, A. Monolithic columns with immobilized monomeric avidin: Preparation and application for affinity chromatography. *Anal. Bioanal. Chem.* **402**, 2395–2405 (2012).

Bibliography

154. Ntougkos, E. *et al.* The IgLON Family in Epithelial Ovarian Cancer : Expression Profiles and Clinicopathologic Correlates The IgLON Family in Epithelial Ovarian Cancer : Expression Profiles and Clinicopathologic Correlates. *Clin. cancer Res.* **11**, 5764–5768 (2005).
155. Lodge, A. P., Howard, M. R., McNamee, C. J. & Moss, D. J. Co-localisation, heterophilic interactions and regulated expression of IgLON family proteins in the chick nervous system. *Mol. Brain Res.* **82**, 84–94 (2000).
156. Teodoridis, J. M. *et al.* CpG Island Methylation of DNA Damage Response Genes in Advanced Ovarian Cancer Advanced Ovarian Cancer. *Cancer Res.* **65**, 8961–8967 (2005).
157. Jones, P. A. & Laird, P. W. Cancer epigenetics comes of age. *Nat. Genet.* **21**, 163–167 (1999).
158. Goffin, J. & Eisenhauer, E. DNA methyltransferase inhibitors - State of the art. *Ann. Oncol.* **13**, 1699–1716 (2002).
159. Rankin, E. B. *et al.* AXL is an essential factor and therapeutic target for metastatic ovarian cancer. *Cancer Res.* **70**, 7570–7579 (2010).
160. Linger, R., Keating, A., Shelton Earp, H. & Graham, D. TAM Receptor Tyrosine Kinases: Biologic Functions, Signaling, and Potential Therapeutic Targeting in Human Cancer. *Adv Cancer Res* **100**, 35–83 (2008).
161. Sasaki, T. *et al.* Structural basis for Gas6-Axl signalling. *EMBO J.* **25**, 80–87 (2006).
162. Gabra, H. Unpublished Research.
163. Paulick, M. G. & Bertozzi, C. R. The Glycosylphosphatidylinositol Anchor : A Complex Membrane-Anchoring. *Biochemistry* **47**, 6991–7000 (2008).
164. Baeschlin, D. K. *et al.* 1,2-Diacetals in Synthesis: Total Synthesis of a Glycosylphosphatidylinositol Anchor of Trypanosoma brucei. *Chem. - A Eur. J.* **6**, 172–186 (2000).
165. Kelley, L. A. & Sternberg, M. J. E. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* **4**, 363–371 (2009).
166. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling , prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
167. The Uniprot Consortium. UniProt: a hub for protein information. *Nucleic Acids Res.* **43**,

Bibliography

- D204–D212 (2014).
168. Somia, N. & Verma, I. M. Gene therapy: trials and tribulations. *Nat. Rev. Genet.* **1**, 91–99 (2000).
 169. Indraccolo, S. *et al.* Gene Transfer in Ovarian Cancer Cells : A Comparison between Retroviral and Lentiviral Vectors 1. *Cancer Res.* **62**, 6099–6107 (2002).
 170. Davis, H. E., Rosinski, M., Morgan, J. R. & Yarmush, M. L. Charged polymers modulate retrovirus transduction via membrane charge neutralization and virus aggregation. *Biophys. J.* **86**, 1234–1242 (2004).
 171. Davis, H. E., Morgan, J. R. & Yarmush, M. L. Polybrene increases retrovirus gene transfer efficiency by enhancing receptor-independent virus adsorption on target cell membranes. *Biophys. Chem.* **97**, 159–172 (2002).
 172. Wu, S. Y. & Sood, A. K. New roles opined for OPCML. *Cancer Discov.* **2**, 115–116 (2012).
 173. Antos, J. M., Miller, G. M., Grotenbreg, G. M. & Ploegh, H. L. Lipid modification of proteins through sortase-catalyzed transpeptidation. *J. Am. Chem. Soc.* **130**, 16338–16343 (2008).

Appendix A

This appendix provides the full DNA sequences and primers for each OPCML construct designed for lentiviral production of transduced SKOV3 cells pENTR FL, D1, D2, D3.

KEY
SIGNAL PEPTIDE
PROPEPTIDE
GLYCOSYLATION SITES
DOMAIN 1
DOMAIN 2
HA TAG
SAL1 RESTRICTION SITE
XBA1 RESTRICTION SITE

	Primers 5' to 3'
OPCML sequencing	Forward: GTCTGTGGGTACCTGTTCTG Reverse: AAGAAGTGGGCTAAGAGGGTC

OPCML Domain 1 (D1)

GTCGACGCCACCATGGGGGTCTGTGGGTACCTGTTCTGCCCTGGAAGTGCCTCGTGGTCTGTCTCT
 CAGGCTGCTGTTCTTGTACCCACAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCCCAAAGCTA
 TGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGTACCATAGATGACCGGGTAAC
 CCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTACGCTGGGAATGACAAGTGGTCCATAGACCCTC
 GTGTGATCATCCTGGTCAATACCAACCCAGTACAGCATCATGATCCAAAATGTGGATGTGTATGACG
 AAGGTCCGTACACCTGCTCTGTGCAGACAGACAATCATCCAAAACGTCCGGGTTACCTAATAGTGC
 AAGTTCCTTTGTATTACCATACGATGTTCCAGATTACGCTGGGCCTGGAGCAGTCATTGATGGTGTA
 ACTCGGCCTCCAGAGCACTGGCTTGTCTCTGGCTATCAGGGACCCTCTTAGCCCACTTCTTCATCAAGTT
 TTGATCTAGA

OPCML Domain 2 (D2)

GTCGACGCCACCATGGGGGTCTGTGGGTACCTGTTCTGCCCTGGAAGTGCCTCGTGGTCTGTCTCT
 CAGGCTGCTGTTCTTGTACCCACAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCCGGGTTACCC

Appendix A

TAATAGTGCAAGTTCCTCCTCAGATCATGAATATCTCCTCAGACATCACTGTGAATGAGGGAAGCAGTG
TGACCCTGCTGTGTCTTGCTATTGGCAGACCAGAGCCAACTGTGACATGGAGACACCTGTCAGTCAAG
GAAGGCCAGGGCTTTGTAAGTGAGGATGAGTACCTGGAGATCTCTGACATCAAGCGAGACCAGTCCG
GGGAGTACGAATGCAGCGCGTTGAACGATGTCGCTGCGCCCGATGTGCGGAAAGTAAAAATCACTGT
AAACTATTTGTATTACCCATACGATGTTCCAGATTACGCTGGGCCTGGAGCAGTCATTGATGGTGTAAA
CTCGGCCTCCAGAGCACTGGCTTGCTCTGGCTATCAGGGACCCTCTTAGCCCACTTCTTCATCAAGTTT
TGATCTAGA

OPCML Domain 3 (D3)

GTCGACGCCACCATGGGGGTCTGTGGGTACCTGTTCTGCCCTGGAAGTGCCTCGTGGTCTGTCTCT
CAGGCTGCTGTTCTTGTAACCCACAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCGTAAACTATC
CTCCCTATATCTCAAAAGCCAAGAACAACACTGGTGTTTCAGTCGGTCAGAAGGGCATCCTGAGCTGTGAA
GCCTCTGCAGTCCCCATGGCTGAATTCCAGTGGTTCAAGGAAGAAACCAGGTTAGCCACTGGTCTGGA
TGGAATGAGGATTGAAAACAAAGGCCGCATGTCCACTCTGACTTTCTTCAATGTTTCAGAAAAGGATTAA
TGGGAACATATACTTGTGTGGCCACGAACAAGCTTGGAACACC AATGCCAGCATCACATTGTATTACCC
ATACGATGTTCCAGATTACGCTGGGCCTGGAGCAGTCATTGATGGTGTAAACTCGGCCTCCAGAGCAC
TGGCTTGCTCTGGCTATCAGGGACCCTCTTAGCCCACTTCTTCATCAAGTTTTGATCTAGA

OPCML Full length (FL)

GTCGACGCCACCATGGGGGTCTGTGGGTACCTGTTCTGCCCTGGAAGTGCCTCGTGGTCTGTCTCT
CAGGCTGCTGTTCTTGTAACCCACAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCCCAAAGCTA
TGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGTACCATAGATGACCGGGTAAC
CCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTACGCTGGGAATGACAAGTGGTCCATAGACCCTC
GTGTGATCATCCTGGTCAATACCAACCCAGTACAGCATCATGATCCAAATGTGGATGTGTATGACG
AAGGTCCGTACACCTGCTCTGTGCAGACAGACAATATCCAAAACGTCCGGGTTACCTAATAGTGC
AAGTTCCTCCTCAGATCATGAATATCTCCTCAGACATCACTGTGAATGAGGGAAGCAGTGTGACCCTGC
TGTGTCTTGCTATTGGCAGACCAGAGCCAACTGTGACATGGAGACACCTGTCAGTCAAGGAAGGCCAG
GGCTTTGTAAGTGAGGATGAGTACCTGGAGATCTCTGACATCAAGCGAGACCAGTCCGGGGAGTACG
AATGCAGCGCGTTGAACGATGTCGCTGCGCCCGATGTGCGGAAAGTAAAAATCACTGTAAACTATCCT
CCCTATATCTCAAAAGCCAAGAACAACACTGGTGTTTCAGTCGGTCAGAAGGGCATCCTGAGCTGTGAAGC
CTCTGCAGTCCCCATGGCTGAATTCCAGTGGTTCAAGGAAGAAACCAGGTTAGCCACTGGTCTGGATG
GAATGAGGATTGAAAACAAAGGCCGCATGTCCACTCTGACTTTCTTCAATGTTTCAGAAAAGGATTATG
GGAACTATACTTGTGTGGCCACGAACAAGCTTGGAACACC AATGCCAGCATCACATTGTATTACCCAT
ACGATGTTCCAGATTACGCTGGGCCTGGAGCAGTCATTGATGGTGTAAACTCGGCCTCCAGAGCACTG
GCTTGCTCTGGCTATCAGGGACCCTCTTAGCCCACTTCTTCATCAAGTTTTGATCTAGA

OPCML Wild Type (WT)

ATGGGGGTCTGTGGGTACCTGTTCTGCCCTGGAAGTGCCTCGTGGTCGTGTCTCTCAGGCTGCTGTTCTTGTACCCACAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGTACCATAGATGACCGGGTAACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTACGCTGGGAATGACAAGTGGTCCATAGACCCTCGTGTGATCATCCTGGTCAATACACCAACCCAGTACAGCATCATGATCCAAAATGTGGATGTGTATGACGAAGGTCCGTACACCTGCTCTGTGCAGACAGACAATCATCCAAAACGTCCCGGGTTCACCTAATAGTGCAAGTTCCTCCTCAGATCATGAATATCTCCTCAGACATCACTGTGAATGAGGGAAGCAGTGTGACCCTGCTGTGTCTTGCTATTGGCAGACCAGAGCCAACCTGTGACATGGAGACACCTGTCAGTCAAGGAAGGCCAGGGCTTTGTAAGTGAGGATGAGTACCTGGAGATCTCTGACATCAAGCGAGACCAGTCCGGGGAGTACGAATGCAGCGCGTTGAACGATGTCGCTGCGCCCGATGTGCGGAAAGTAAAAATCACTGTAAACTATCCTCCCTATATCTCAAAGCCAAGAACACTGGTGTTTCAGTCGGTCAGAAGGGCATCCTGAGCTGTGAAGCCTCTGCAGTCCCATGGCTGAATTCCAGTGGTTCAAGGAAGAAACCAGGTTAGCCACTGGTCTGGATGGAATGAGGATTGAAAACAAAGGCCGCATGTCCACTCTGACTTTCTTCAATGTTTCAGAAAAGGATTATGGGAACATACTTGTGTGGCCACGAACAAGCTTGGGAACACCAATGCCAGCATCACATTGTATTGGGCCTGGAGCAGTCA TTGATGGTGTAACCTCGGCCTCCAGAGCACTGGCTTGTCTCTGGCTATCAGGGACCCTCTAGCCCACTCTTCATCAAGTTTTG

Appendix B

This appendix provides the full sequences and primers for each OPCML construct designed for protein production.

KEY
SIGNAL PEPTIDE
PROPEPTIDE
GLYCOSYLATION SITES
DOMAIN 1
DOMAIN 2
HA TAG
FREE CYSTEINE
THROMBIN CLEAVAGE SITE
Fc PROTEIN SEQUENCE
TWIN STREP TAG
ECORV RESTRICTION SITE
XHOI RESTRICTION SITE
SORTASE RECOGNITION SITE

	Primers 5' to 3'
OPCML sequencing	Forward: GTCTGTGGGTACCTGTTCTCTG Reverse: AAGAAGTGGGCTAAGAGGGTC
For removing free cysteine in OPCML sortagging construct	Forward: TACCCATACGATGTTCCAGATTACGC Reverse: TACACCATCAATGACTGCTCCAGG
For making R65L mutant	Forward: CCGGGTAACCTTGGTGGCCTGGC Reverse: TCATCTATGGTACACCTGAGG

OPCML sequence for mammalian expression of rOPCML for sortagging

GATATCGCCACCATGGGGGTCTGTGGGTACCTGTTCTCTGCCCTGGAAGTGCCTCGTGGTCTGTCTCTC
AGGCTGCTGTTCTTGTACCCACAGCAGTGCCCGTGCGCAGCGGAGATGCCACCTTCCCAAAGCTAT
GGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGTACCATAGATGACCGGGTAACC
CGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTACGCTGGGAATGACAAGTGGTCCATAGACCCTCG
TGTGATCATCCTGGTCAATACCAACCCAGTACAGCATCATGATCCAAAATGTGGATGTGTATGACGA

AGGTCCGTACACATGCTCTGTGCAGACAGACAATCATCCAAAACGTCCCGGGTTCACCTAATAGTGCA
 AGTTCCTCCTCAGATCATGAATATCTCCTCAGACATCACTGTGAATGAGGGAAGCAGTGTGACCCTGCT
 GTGTCTTGCTATTGGCAGACCAGAGCCAACCTGTGACATGGAGACACCTGTCAAGGAAGGCCAG
 GGCTTTGTAAGTGAGGATGAGTACCTGGAGATCTCTGACATCAAGCGAGACCAGTCCGGGGAGTACG
 AATGCAGCGCGTTGAACGATGTCGCTGCGCCCGATGTGCGGAAAGTAAAAATCACTGTAACTATCCT
 CCCTATATCTCAAAAGCCAAGAACACTGGTGTTCAGTCGGTCAGAAGGGCATCCTGAGCTGTGAAGC
 CTCTGCAGTCCCCATGGCTGAATTCCAGTGGTTCAAGGAAGAAACCAGGTTAGCCACTGGTCTGGATG
 GAATGAGGATTGAAAACAAAGGCCGCATGTCCACTCTGACTTTCTTCAATGTTTCAGAAAAGGATTATG
 GGAATATACTTGTGTGGCCACGAACAAGCTTGGGAACACCAATGCCAGCATCACATTGTATGGGCCT
 GGAGCAGTCATTGATGGTGTATGCTTACCCATACGATGTTCCAGATTACGCTGGAGGAGGAGGATCGCT
CCCAGAGACAGGTCTGGTCCGCGTGGTAGTTGGAGCCACCCGCAGTTCGAAAAAGGTGGAGGTTCT
GGCGGTGGATCGGGAGGTTCAAGCTGGAGCCACCCGCAGTTCGAAAAATAACTCGAC

OPCML sequence for Insect cell expression of rOPCML for sortagging

ATGAAAAATTTGATCTTGCGGTATATTTTGTTCCTTTGCTCAACGCCGCTCCAAGCAGTGCCCGTGCGCAGCGGAG
 ATGCCACCTTCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGTACCATA
 GATGACCGGGTAACCCGGGTGGCTGGCTAAACCGCAGCACCATCCTCTACGCTGGGAATGACAAGTGGTCCAT
 AGACCCTCGTGTGATCATCCTGGTCAATACCAACCCAGTACAGCATCATGATCCAAAATGTGGATGTGTATGAC
 GAAGGTCCGTACACATGCTCTGTGCAGACAGACAATCATCCAAAACGTCCCGGGTTCACCTAATAGTGCAAGTT
 CCTCCTCAGATCATGAATATCTCCTCAGACATCACTGTGAATGAGGGAAGCAGTGTGACCCTGCTGTGTCTTGCTA
 TTGGCAGACCAGAGCCAACCTGTGACATGGAGACACCTGTCAAGGAAGGCCAGGGCTTTGTAAGTGAGGAT
 GAGTACCTGGAGATCTCTGACATCAAGCGAGACCAGTCCGGGGAGTACGAATGCAGCGCGTTGAACGATGTCGC
 TGCGCCCGATGTGCGGAAAGTAAAAATCACTGTAACTATCCTCCTATATCTCAAAAGCCAAGAACACTGGTGT
 TCAGTCGGTCAGAAGGGCATCCTGAGCTGTGAAGCCTCTGCAGTCCCCATGGCTGAATTCCAGTGGTTCAAGGAA
 GAAACCAGGTTAGCCACTGGTCTGGATGGAATGAGGATTGAAAACAAAGGCCGCATGTCCACTCTGACTTTCTTC
 AATGTTTCAGAAAAGGATTATGGGAATATACTTGTGTGGCCACGAACAAGCTTGGGAACACCAATGCCAGCATC
 ACATTGTATGGGCCTGGAGCAGTCATTGATGGTGTATGCTTACCCATACGATGTTCCAGATTACGCTGGAGGAGGA
 GGATCGCTCCAGAGACAGGTCTGGTCCGCGTGGTAGTGGTGGCTCAACTAAGGGTCCCTCTGTATTTCTGTTTC
 CCCCAGGCAAGAGACACCTTATGATTTCAGGACCCCTGAAGTTACGTGTGTCGTGGTTGACGTATCTCACGA
 AGATCCTGAGGTGAAGTTCAACTGGTACGTGATGGCGTGAAGTGCATAACGCGAAAAACCAAGCCGCGCGAG
 GAACAGTATAATTCCACGTATAGAGTGGTGAGCGTACTTACGGTCTCCACCAGGATTGGCTTAATGGTAAGGAA
 TACAAATGTAAGGTCTCTAATAAGGCTCTCCCTGCACCGATAGAAAAGACGATCAGTAAAGCCAAAGGCCAACCT
 AGAGAACCTCAGGTCTATACCCTCCACCTAGTCGCGATGAAGTACAAAAAATCAAGTAAGTCTCACTTGTCTCG
 TCAAAGGCTTTTATCCCTCCGACATCGCTGTGAGTGGGAGTCAAACGGTCAGCCCGAAAAATAACTATAAACGA
 CACCGCCGGTACTCGATTGATGATCTTTTTTTCTTATAGTAAATTGACTGTGACAAGAGTCGCTGGCAACA

AGGTAATGCTTTAGTTGTTTCAGTTATGCATGAAGCCCTTCACAACCACTACACGCAAAAGTCCCTCTCACTTTCTC
CGGGAAAGTAA

R65L mutant OPCML sequence for Insect cell expression of rOPCML for
sortagging

ATGAAAATTTTGATCTTGGGTATATTTTGTTCCTTTGCTCAACGCCGCTCCAAGCAGTGCCCGTGCGCAGCGGAG
ATGCCACCTTC~~CCCAAAGCTATGGACAAC~~GTGACGGTCCGGCAGGGGGAGAGCGCCAC~~CTCAGGTGTACCATA~~
~~GATGA~~CCGGGTAA~~CC~~GGGTGGCCTGGCTAAACCGCAGCACCATCCTCTACGCTGGGAATGACAAGTGGTCCAT
AGACCCCTCGTGTGATCATCCTGGTCAATACCAACCCAGTACAGCATCATGATCCAAAATGTGGATGTGTATGAC
GAAGGTCCGTACACATGCTCTGTGCAGACAGACAATCATCCCAAACGTCCCGGGTTCACCTAATAGTGCAAGTT
CCTCCTCAGATCATGAATATCTCCTCAGACATCACTGTGAATGAGGGAAGCAGTGTGACCCTGCTGTGTTGCTA
TTGGCAGACCAGAGCCAACCTGTGACATGGAGACACCTGTCAAGGAAGGCCAGGGCTTTGTAAGTGAGGAT
GAGTACCTGGAGATCTCTGACATCAAGCGAGACCAGTCCGGGGAGTACGAATGCAGCGCGTTGAACGATGTGCG
TGCGCCCGATGTGCGGAAAGTAAAATCACTGTAACTATCCTCCCTATATCTCAAAGCCAAGAACTGGTGT
TCAGTCGGTCAGAAGGGCATCCTGAGCTGTGAAGCCTCTGCAGTCCCATGGCTGAATTCAGTGGTTCAAGGAA
GAAACCAGGTTAGCCACTGGTCTGGATGGAATGAGGATTGAAAACAAAGGCCGCATGTCCACTCTGACTTTCTTC
AATGTTTCAGAAAAGGATTATGGGAACATACTTGTGTGGCCACGAACAAGCTTGGGAACACCAATGCCAGCATC
ACATTGTATGGGCCTGGAGCAGTCATTGATGGTGTATACCCATACGATGTTCCAGATTACGCTGGAGGAGGAGG
ATCGCTCCAGAGACAGGTCTGGTTCCGCGTGGTAGTGGTGGCTCAACTAAGGGTCCCTCTGTATTTCTGTTTCCC
CCAAAGCCAAAAGACACCCTTATGATTTCAAGGACCCCTGAAGTTACGTGTGTCGTGGTTGACGTATCTCACGAA
GATCCTGAGGTGAAGTTCAACTGGTACGTCGATGGCGTGGAAGTGATAACGCGAAAACCAAGCCGCGCGAGG
AACAGTATAATTCCACGTATAGAGTGGTGAGCGTACTTACGGTCCTCCACCAGGATTGGCTTAATGGTAAGGAAT
ACAAATGTAAGGTCTCTAATAAGGCTCTCCCTGCACCGATAGAAAAGACGATCAGTAAAGCCAAAGGCCAACCTA
GAGAACCTCAGGTCTATACCCTCCACCTAGTCGCGATGAACTGACAAAAAATCAAGTAAGTCTCACTTGTCTCGT
CAAAGGCTTTTATCCCTCCGACATCGCTGTGAGTGGGAGTCAAACGGTCAGCCCGAAAATAACTATAAACGAC
ACCGCCGGTACTCGATTGAGATGGATCTTTTTTCTTTATAGTAAATTGACTGTCGACAAGAGTCGCTGGCAACAA
GGTAATGTCTTTAGTTGTTTCAGTTATGCATGAAGCCCTTCACAACCACTACACGCAAAAGTCCCTCTCACTTTCTCC
GGGAAAGTAA

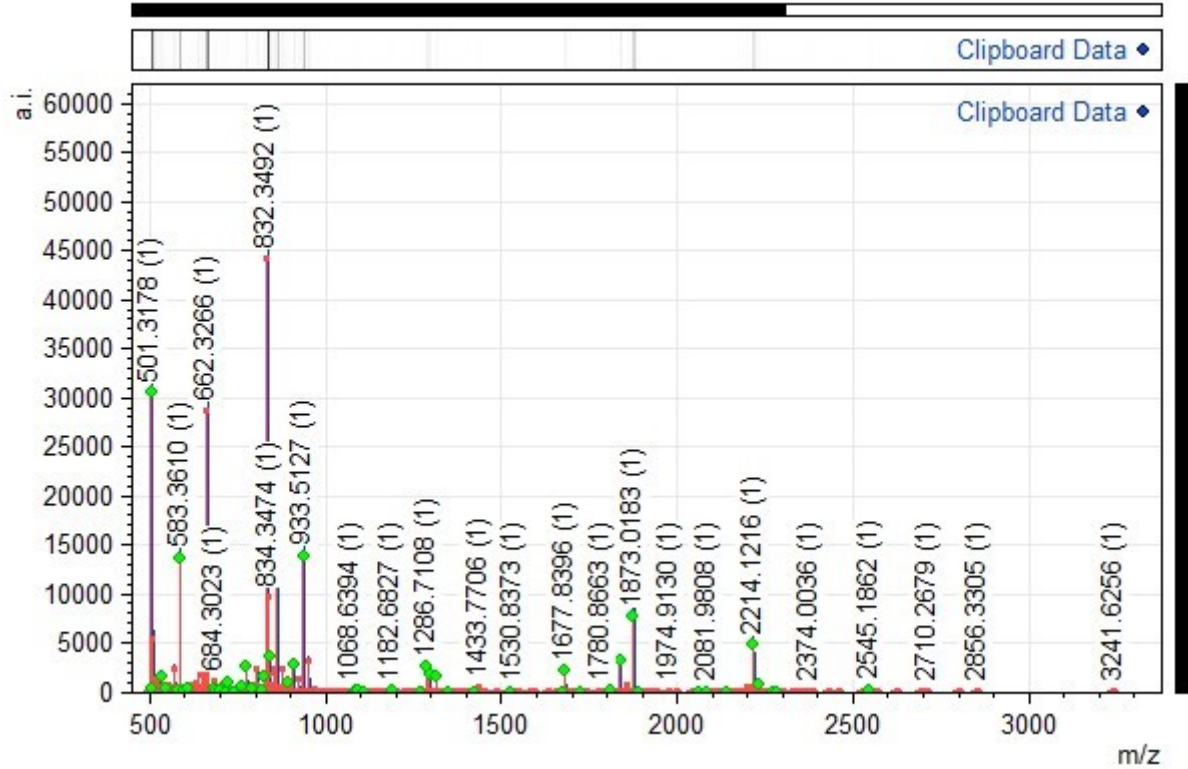
Appendix C

This appendix provides the peptide mass fingerprint (PMF) data that was used to confirm the identity of proteins produced. The data was analysed using mMass software.

mMass analysis/Mascot search of protein band in Figure 16 after trypsin digest

mMass Report: *Clipboard Data*

Date		Scan Number	
Operator		Retention Time	
Contact		MS Level	
Institution		Precursor m/z	
Instrument		Polarity	unknown
		Spectrum Points	157210
		Peak List	663



Appendix C

Sequence - OPCML+FC Purification Upper Band

Accession	Length	Mo. Mass	Av. Mass	Coverage	Matched Int.
	527	58287.1570	58323.2020	62.8 %	38.4 %
AVPVRSGDAT FPKAMDNVTV RQGESATLR C TIDDRVTRVA WLNRSITLYA GNDKWSIDPR VIILVNTPTQ YSIMIQNV DV YDEGPYT CSV QTDNHPKTSR VHLIVQVPPQ IMNISSDITV NEGSSVTLL C LAIGRPEPTV TWRHLSVKEG QGFVSEDEYL EISDIKRDQS GEYEC SALND VAAPDVRKVK ITVNYPYIS KAKNTGVS VG QKGIL SC EAS AVPMAEFQWF KEETRLATGL DGMRIENKGR MSTLTFFNVS EKDYGN YTCV ATNKLGNNTA SITLYGPGAV IDGV CG GGGS LPETGLVPRG SGGSTKGPSV FLFPKPKDT LMISRTPEVT CV VVDVSHED PEVKFNWYVD GVEVHNAKTK PREEQYNSTY RVVSVLTVLH QDWLNGKEYK CK VS NKALPA PIEKTISKAK GQPREPQVYT LPPSRDELTK NQVSLT CL VK GFYPSDIAVE WESNGQPENN YKTPPVLD S DGSFFLYSKL TVDKSRWQQG NVFS C SVMHE ALHNHYTQKS LSLSPGK					

Meas. m/z	Calc. m/z	δ (Da)	δ (ppm)	Rel. Int. (%)	z	Annotation	Formula
501.3178	501.3144	0.0035	6.9	69.08	1	[369-372] k.TKPR.e	C21H40N8O6
503.1507	503.2824	-0.1317	-261.7	0.70	1	[245-248] r.IENK.g	C21H38N6O8
534.2545	534.2518	0.0027	5.1	3.82	1	[232-235] k.EETR.l	C20H35N7O10
541.2109	541.3457	-0.1348	-249.0	0.99	1	[1-5] .AVPVR.s	C24H44N8O6
575.1824	575.3399	-0.1575	-273.8	0.27	1	[490-494] k.LTVDK.s	C25H46N6O9
575.3411	575.3399	0.0012	2.1	0.64	1	[490-494] k.LTVDK.s	C25H46N6O9
583.3610	583.3562	0.0048	8.3	30.96	1	[144-148] r.HLSVK.e	C26H46N8O7
593.2956	593.2889	0.0067	11.3	0.52	1	[310-316] r.GSGGSK.g	C22H40N8O11
605.3401	605.3141	0.0260	43.0	0.79	1	[436-440] r.DELTK.n	C25H44N6O11
656.3988	656.3838	0.0149	22.7	29.19	1	[419-424] k.AKGQPR.e	C27H49N11O8
678.3263	678.3603	-0.0340	-50.1	0.89	1	[401-406] k.CKVS NK.a	C27H51N9O9S
686.3007	686.3178	-0.0171	-24.9	0.39	1	[398-402] k.EYKCK.v [1xOxidation]	C29H47N7O10S
694.2213	694.3552	-0.1340	-193.0	0.26	1	[401-406] k.CKVS NK.a [1xOxidation]	C27H51N9O10S
716.4153	716.4050	0.0103	14.4	0.18	1	[245-250] r.IENKGR.m	C29H53N11O10
722.3321	722.3138	0.0183	25.4	2.26	1	[30-35] r.CTIDDR.v	C27H47N9O12S
738.2946	738.3087	-0.0140	-19.0	0.13	1	[30-35] r.CTIDDR.v [1xOxidation]	C27H47N9O13S
758.4483	758.4308	0.0176	23.2	1.21	1	[39-44] r.VAWLNR.s	C35H55N11O8
773.4094	773.3941	0.0153	19.8	6.22	1	[55-60] k.WSIDPR.v	C35H52N10O10
788.4595	788.4512	0.0082	10.4	0.87	1	[520-527] k.SLSLSPGK.	C34H61N9O12
818.3242	818.4730	-0.1488	-181.8	0.53	1	[490-496] k.LTVDKSR.w	C34H63N11O12
822.4154	822.3992	0.0162	19.7	3.68	1	[6-13] r.SGDATFPK.a	C36H55N9O13
838.5161	838.5033	0.0128	15.3	8.10	1	[407-414] k.ALPAPIEK.t	C39H67N9O11
861.4669	861.4425	0.0244	28.3	27.01	1	[22-29] r.QGESATLR.c	C34H60N12O14
889.4863	889.4738	0.0125	14.1	2.11	1	[204-212] k.NTGVS VGQK.g	C36H64N12O14
905.4735	905.4509	0.0225	24.9	6.37	1	[14-21] k.AMDNVTVR.q	C36H64N12O13S
933.5127	933.4822	0.0305	32.7	31.53	1	[236-244] r.LATGLDGMR.i	C38H68N12O13S
1078.5696	1078.5310	0.0386	35.8	0.08	1	[30-38] r.CTIDDRVTR.v	C42H75N15O16S
1081.5759	1081.5524	0.0235	21.7	0.34	1	[45-54] r.STILYAGNDK.w	C47H76N12O17

Appendix C

1088.6304	1088.6058	0.0246	22.6	0.28	1	[202-212] k.AKNTGVSVGQK.g	C45H81N15O16
1104.6185	1104.6082	0.0104	9.4	0.12	1	[441-450] k.NQVSLTCLVK.g	C47H85N13O15S
1189.5549	1189.5120	0.0429	36.1	0.25	1	[373-381] r.EEQYNSTYR.v	C50H72N14O20
1267.7261	1267.7620	-0.0360	-28.4	0.04	1	[407-418] k.ALPAPIEKTISK.a	C58H102N14O17
1286.7108	1286.6739	0.0369	28.6	5.94	1	[425-435] r.EPQVYTLPPSR.d	C58H91N15O18
1294.7253	1294.7042	0.0212	16.4	4.02	1	[191-201] k.ITVNYPPYISK.a	C62H95N13O17
1313.7914	1313.7616	0.0298	22.7	3.57	1	[317-328] k.GPSVFLFPPKPK.d	C66H100N14O14
1348.6338	1348.5838	0.0501	37.1	0.04	1	[263-274] k.DYGNYTCVATNK.I	C57H85N15O21S
1417.7917	1417.7468	0.0449	31.7	0.06	1	[236-248] r.LATGLDGMRIENK.g	C59H104N18O20S
1521.8805	1521.8675	0.0130	8.5	0.03	1	[189-201] k.VKITVNYPPYISK.a	C73H116N16O19
1671.8604	1671.8085	0.0519	31.1	0.09	1	[369-381] k.TKPREEQYNSTYR.v	C71H110N22O25
1677.8396	1677.8020	0.0376	22.4	5.03	1	[355-368] k.FNWYVDGVEVHNAK.t	C78H108N20O22
1724.7846	1724.9078	-0.1232	-71.4	0.03	1	[421-435] k.GQPREPQVYTLPPSR.d	C76H121N23O23
1808.0580	1808.0065	0.0515	28.5	0.64	1	[382-397] r.VVSVLTVLHQDWLNGK.e	C83H134N22O23
1835.9841	1835.9286	0.0554	30.2	7.25	1	[45-60] r.STILIYAGNDKWSIDPR.v	C82H126N22O26
1873.0183	1872.9702	0.0481	25.7	17.51	1	[425-440] r.EPQVYTLPPSRDELTK.n	C83H133N21O28
1888.0175	1888.0327	-0.0152	-8.1	0.06	1	[310-328] r.GSGGSTKGPSVFLFPPKPK.d	C88H138N22O24
2058.0011	2057.9550	0.0461	22.4	0.06	1	[149-166] k.EGQGFVSEDEYLEISDIK.r	C90H136N20O35
2081.9808	2082.0060	-0.0251	-12.1	0.12	1	[336-354] r.TPEVTCVVVDVSHEDPEVK.f	C89H144N22O33S
2139.0030	2138.9295	0.0735	34.4	0.18	1	[168-187] r.DQSGEYECALNDVAAPDVR.k	C87H135N25O36S
2214.1216	2214.0561	0.0656	29.6	10.92	1	[149-167] k.EGQGFVSEDEYLEISDIK.d	C96H148N24O36
2228.1214	2228.2074	-0.0859	-38.6	2.03	1	[382-400] r.VVSVLTVLHQDWLNGKEYK.c	C103H162N26O29
2267.0723	2267.0245	0.0478	21.1	0.04	1	[168-188] r.DQSGEYECALNDVAAPDVRK.v	C93H147N27O37S
2283.0958	2283.0194	0.0764	33.5	0.03	1	[168-188] r.DQSGEYECALNDVAAPDVRK.v [1xOxidation]	C93H147N27O38S
2544.1945	2544.1314	0.0632	24.8	0.37	1	[451-472] k.GFYPSDIAVEWESNGQPENNYK.t	C114H158N28O39



Mascot Search Results

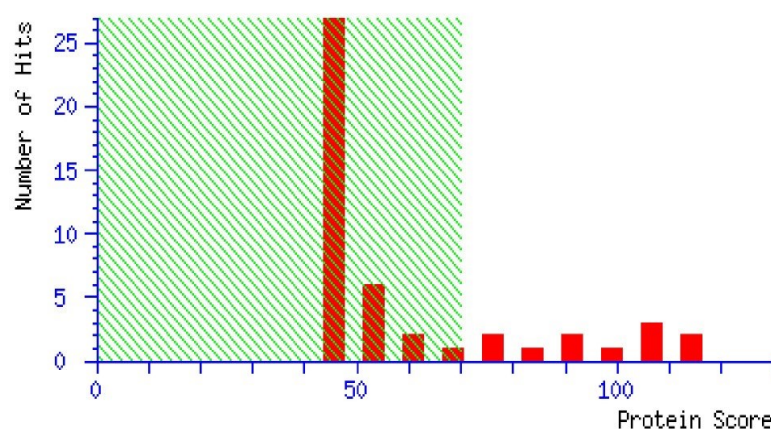
Database : SwissProt 2018_10 (558681 sequences; 200603466 residues)
 Timestamp : 22 Nov 2018 at 15:49:26 GMT
 Top Score : 114 for **Mixture 1**, OPCM_BOVIN

+IGHG1_HUMAN

Mascot Score Histogram

Protein score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event.

Protein scores greater than 70 are significant ($p < 0.05$).



Concise Protein Summary Report

Significance threshold $p < 0.05$

Max. number of hits: AUTO

- Mixture 1** **Total score: 114** **Expect: 2.2e-06** **Matches:**
 45 Components (only one family member shown for each component):
[OPCM_BOVIN](#) **Mass:** 37890 **Score: 89** **Expect:** 0.00066 **Matches:** 26
 Opioid-binding protein/cell adhesion molecule OS=Bos taurus OX=9913 GN=OPCML
 PE=1 SV=1

[IGHG1_HUMAN](#) **Mass:** 36083 **Score: 79** **Expect:** 0.0064 **Matches:** 20
 Immunoglobulin heavy constant gamma 1 OS=Homo sapiens OX=9606 GN=IGHG1 PE=1
 SV=1
- Mixture 2** **Total score: 114** **Expect: 2.2e-06** **Matches:**
 48 Components (only one family member shown for each component):
[OPCM_HUMAN](#) **Mass:** 37983 **Score: 89** **Expect:** 0.00066 **Matches:** 29

Opioid-binding protein/cell adhesion molecule OS=Homo sapiens OX=9606 GN=OPCML
PE=1 SV=1

[IGHG1 HUMAN](#) Mass: 36083 Score: **79** Expect: 0.0064 Matches: 20

Immunoglobulin heavy constant gamma 1 OS=Homo sapiens OX=9606 GN=IGHG1 PE=1
SV=1

3. **Mixture 3** Total score: **110** Expect: 5.6e-06 Matches:

50 Components (only one family member shown for each component):

[OPCM BOVIN](#) Mass: 37890 Score: **89** Expect: 0.00066 Matches: 26

Opioid-binding protein/cell adhesion molecule OS=Bos taurus OX=9913 GN=OPCML
PE=1 SV=1

[IGG1 HUMAN](#) Mass: 49298 Score: **76** Expect: 0.013 Matches: 25

Immunoglobulin gamma-1 heavy chain OS=Homo sapiens OX=9606 PE=1 SV=2

4. **Mixture 4** Total score: **110** Expect: 5.6e-06 Matches:

53 Components (only one family member shown for each component):

[OPCM HUMAN](#) Mass: 37983 Score: **89** Expect: 0.00066 Matches: 29

Opioid-binding protein/cell adhesion molecule OS=Homo sapiens OX=9606 GN=OPCML
PE=1 SV=1

[IGG1 HUMAN](#) Mass: 49298 Score: **76** Expect: 0.013 Matches: 25

Immunoglobulin gamma-1 heavy chain OS=Homo sapiens OX=9606 PE=1 SV=2

5. **Mixture 5** Total score: **104** Expect: 2.2e-05 Matches:

45 Components (only one family member shown for each component):

[OPCM RAT](#) Mass: 38043 Score: **80** Expect: 0.0053 Matches: 26

Opioid-binding protein/cell adhesion molecule OS=Rattus norvegicus OX=10116
GN=Opcml PE=1 SV=2

[IGHG1 HUMAN](#) Mass: 36083 Score: **79** Expect: 0.0064 Matches: 20

Immunoglobulin heavy constant gamma 1 OS=Homo sapiens OX=9606 GN=IGHG1
PE=1 SV=1

6. **Mixture 6** **Total score: 100** **Expect: 5.7e-05** **Matches: 50** Components (only one family member shown for each component):
[OPCM RAT](#) **Mass: 38043** **Score: 80** **Expect: 0.0053** **Matches: 26**
 Opioid-binding protein/cell adhesion molecule OS=Rattus norvegicus OX=10116
 GN=Opcml PE=1 SV=2

[IGG1 HUMAN](#) **Mass: 49298** **Score: 76** **Expect: 0.013** **Matches: 25**
 Immunoglobulin gamma-1 heavy chain OS=Homo sapiens OX=9606 PE=1 SV=2
-
7. [OPCM BOVIN](#) **Mass: 37890** **Score: 89** **Expect: 0.00066** **Matches: 26**
 Opioid-binding protein/cell adhesion molecule OS=Bos taurus OX=9913 GN=OPCML
 PE=1 SV=1
-
8. [OPCM HUMAN](#) **Mass: 37983** **Score: 89** **Expect: 0.00066** **Matches: 29**
 Opioid-binding protein/cell adhesion molecule OS=Homo sapiens OX=9606 GN=OPCML
 PE=1 SV=1

[OPCM PANTR](#) **Mass: 37983** **Score: 89** **Expect: 0.00066** **Matches: 29**
 Opioid-binding protein/cell adhesion molecule OS=Pan troglodytes OX=9598
 GN=OPCML PE=2 SV=1
-
9. [OPCM RAT](#) **Mass: 38043** **Score: 80** **Expect: 0.0053** **Matches: 26**
 Opioid-binding protein/cell adhesion molecule OS=Rattus norvegicus OX=10116
 GN=Opcml PE=1 SV=2
-
10. [IGHG1 HUMAN](#) **Mass: 36083** **Score: 79** **Expect: 0.0064** **Matches: 20**
 Immunoglobulin heavy constant gamma 1 OS=Homo sapiens OX=9606 GN=IGHG1
 PE=1 SV=1
-
11. [IGG1 HUMAN](#) **Mass: 49298** **Score: 76** **Expect: 0.013** **Matches: 25** Immunoglobulin
 gamma-1 heavy chain OS=Homo sapiens OX=9606 PE=1 SV=2
-
12. [CPXS PORPU](#) **Mass: 17235** **Score: 66** **Expect: 0.13** **Matches: 14**
 Chromophore lyase CpcS/CpeS homolog OS=Porphyra purpurea OX=2787 GN=cpcS
 PE=3 SV=1
-

Search Parameters

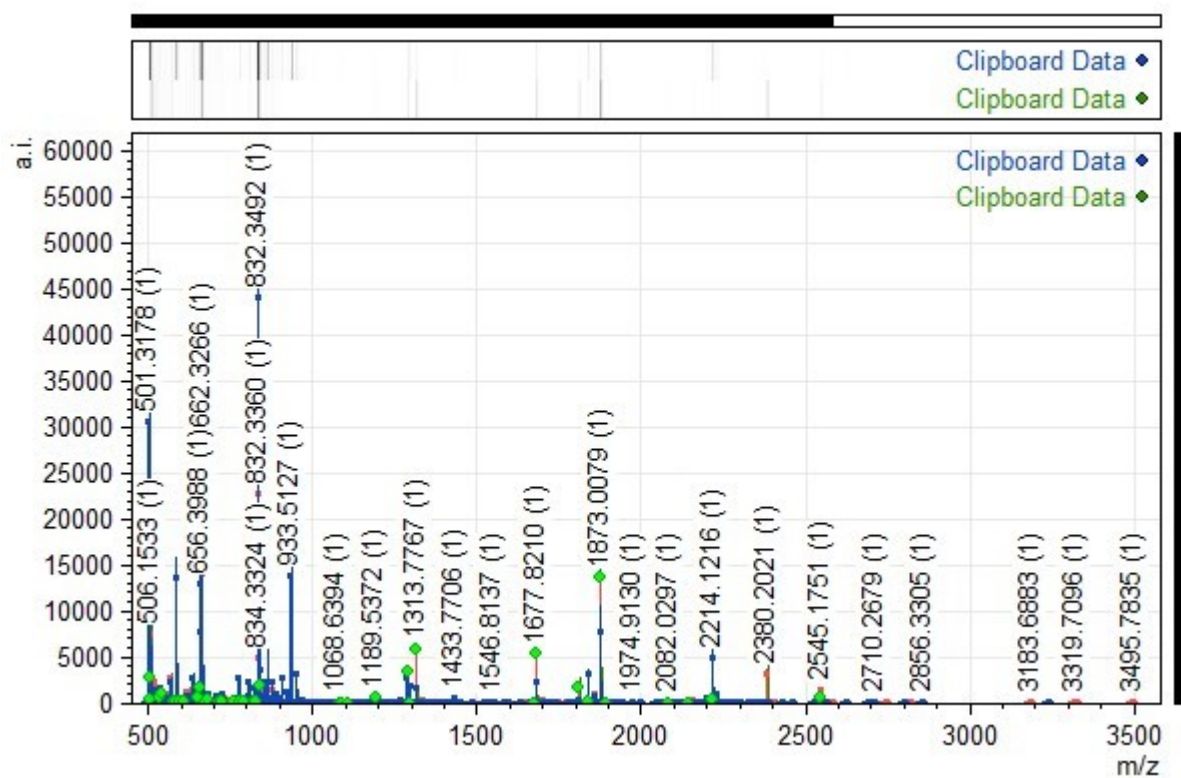
Type of search : Peptide Mass Fingerprint
Enzyme : Trypsin
Variable modifications : [Oxidation\(M\)](#)
Mass values : Monoisotopic
Protein Mass : Unrestricted
Peptide Mass Tolerance : ± 0.1 Da
Peptide Charge State : 1+
Max Missed Cleavages : 1
Number of queries : 663
Selected for scoring : 134

Mascot:<http://www.matrixscience.com/>

mMass analysis/Mascot search of protein band in Figure 16 after trypsin digest

mMass Report: *Clipboard Data*

Date		Scan Number	
Operator		Retention Time	
Contact		MS Level	
Institution		Precursor m/z	
Instrument		Polarity	unknown
		Spectrum Points	157210
		Peak List	703



Appendix C

Sequence - OPCML+FC Purification Lower Band

Accession	Length	Mo. Mass	Av. Mass	Coverage	Matched Int.
	527	58111.2129	58147.2085	55.2 %	20.8 %
<u>AVPVRSGDAT</u> <u>FPK</u>AMDNVTV RQGESATLRC <u>TIDDRVTRVA</u> <u>WLN</u>RSTILYA <u>GNDK</u>WSIDPR VIILVNTPTQ YSIMIQNV DV YDEGPYTC SV QTDNHPKTSR VHLIVQVPPQ IMNISSDITV NEGSSVTL LC LAIGRPEPTV TWR<u>HLSV</u>KEG <u>QGFV</u>SEDEYL <u>EISDIK</u>RQDS <u>GEYEC</u>SALND <u>VAAPD</u>VRKVK <u>ITV</u>NYPPYIS <u>KAKNT</u>GVSVG <u>QKGIL</u>SCEAS AVPMAEFQWF <u>KEETR</u>LATGL DGMRIENKGR MSTLTFFNVS EKDYGN YTCV ATNKLGN TNA SITLYGPGAV IDGVC GGGGS LPETGLVPRG <u>SGGST</u>KGPSV <u>FLFP</u>PKP KDT LMISRTPEVT <u>CVVVD</u>VSHED <u>PEVKF</u>NWYVD <u>GVEVH</u>NAKTK <u>PREEQ</u>YNSTY <u>RVVSV</u>LTVLH <u>QDWL</u>NGKEYK <u>CKVSN</u>KALPA <u>PIEKT</u>ISKAK <u>GQPRE</u>QVYT <u>LPPSR</u>DELTK <u>NQVSL</u>TCLVK <u>GFYPS</u>DIAVE <u>WESNG</u>Q PENN <u>YKTP</u>PVLDS DGSFFLYSKL <u>TVDK</u>SRWQQG NVFSCSV MHE ALHNHYTQK <u>LSL</u>SPGK					

Meas. m/z	Calc. m/z	δ (Da)	δ (ppm)	Rel. Int. (%)	z	Annotation	Formula
501.3106	501.3144	-0.0037	-7.4	12.52	1	[369-372] k.TKPR.e	C21H40N8O6
503.1506	503.2824	-0.1318	-261.9	1.80	1	[245-248] r.IENK.g	C21H38N6O8
534.1562	534.2518	-0.0956	-179.0	2.77	1	[232-235] k.EETR.l	C20H35N7O10
541.2059	541.3457	-0.1397	-258.1	4.46	1	[1-5] .AVPVR.s	C24H44N8O6
575.3312	575.3399	-0.0087	-15.2	1.27	1	[490-494] k.LTVDK.s	C25H46N6O9
583.1808	583.3562	-0.1754	-300.7	1.07	1	[144-148] r.HLSVK.e	C26H46N8O7
593.2735	593.2889	-0.0154	-26.0	1.16	1	[310-316] r.GSGGSTK.g	C22H40N8O11
605.1952	605.3141	-0.1189	-196.4	0.85	1	[436-440] r.DELTK.n	C25H44N6O11
647.2788	647.4087	-0.1298	-200.5	2.73	1	[415-420] k.TISKAK.g	C28H54N8O9
656.3892	656.3838	0.0053	8.1	7.48	1	[419-424] k.AKGQPR.e	C27H49N11O8
670.2198	670.3229	-0.1031	-153.8	0.50	1	[398-402] k.EYKCK.v	C29H47N7O9S
678.3141	678.3603	-0.0463	-68.2	1.50	1	[401-406] k.CKVS NK.a	C27H51N9O9S
716.2484	716.4050	-0.1566	-218.6	1.16	1	[245-250] r.IENKGR.m	C29H53N11O10
722.2960	722.3138	-0.0178	-24.6	0.51	1	[30-35] r.CTIDDR.v	C27H47N9O12S
758.4503	758.4308	0.0195	25.7	0.52	1	[39-44] r.VAWLNR.s	C35H55N11O8
773.4064	773.3941	0.0123	15.9	1.03	1	[55-60] k.WSIDPR.v	C35H52N10O10
788.4521	788.4512	0.0008	1.1	0.53	1	[520-527] k.SLSLSPGK.	C34H61N9O12
818.3038	818.4730	-0.1692	-206.8	0.69	1	[490-496] k.LTVDKSR.w	C34H63N11O12
822.3015	822.3992	-0.0977	-118.8	0.61	1	[6-13] r.SGDATFPK.a	C36H55N9O13
838.5067	838.5033	0.0035	4.1	8.69	1	[407-414] k.ALPAPIEK.t	C39H67N9O11
1078.3687	1078.5310	-0.1622	-150.4	0.05	1	[30-38] r.CTIDDRVTR.v	C42H75N15O16S
1081.3894	1081.5524	-0.1630	-150.7	0.07	1	[45-54] r.STILYAGNDK.w	C47H76N12O17
1088.6285	1088.6058	0.0227	20.8	0.17	1	[202-212] k.AKNTGVSVGQK.g	C45H81N15O16
1104.6258	1104.6082	0.0177	16.0	0.38	1	[441-450] k.NQVSLTCLVK.g	C47H85N13O15S
1189.5372	1189.5120	0.0252	21.2	2.43	1	[373-381] r.EEQYNSTYR.v	C50H72N14O20
1286.7002	1286.6739	0.0263	20.4	15.64	1	[425-435] r.EPQVYTLPPSR.d	C58H91N15O18
1294.7277	1294.7042	0.0236	18.2	0.35	1	[191-201] k.ITVNYPPYISK.a	C62H95N13O17

Appendix C

1313.7767	1313.7616	0.0151	11.5	25.93	1	[317-328] k.GPSVFLFPPKPK.d	C66H100N14O14
1671.8491	1671.8085	0.0406	24.3	0.41	1	[369-381] k.TKPREEQYNSTYR.v	C71H110N22O25
1677.8210	1677.8020	0.0190	11.4	23.81	1	[355-368] k.FNWYVDGVEVHNAK.t	C78H108N20O22
1808.0317	1808.0065	0.0252	13.9	7.82	1	[382-397] r.VVSVLTVLHQDWLNGK.e	C83H134N22O23
1835.9728	1835.9286	0.0441	24.0	0.81	1	[45-60] r.STILYAGNDKWSIDPR.v	C82H126N22O26
1873.0079	1872.9702	0.0377	20.1	60.37	1	[425-440] r.EPQVYTLPPSRDELTK.n	C83H133N21O28
1887.9931	1888.0327	-0.0396	-21.0	0.10	1	[310-328] r.GSGGSTKGPSVFLFPPKPK.d	C88H138N22O24
2082.0297	2082.0060	0.0238	11.4	0.44	1	[336-354] r.TPEVTCVVVDVSHEDPEVK.f	C89H144N22O33S
2139.0377	2138.9295	0.1082	50.6	0.06	1	[168-187] r.DQSGEYECALNDVAAPDVR.k	C87H135N25O36S
2214.1039	2214.0561	0.0478	21.6	1.85	1	[149-167] k.EGQGFVSEDEYLEISDIKR.d	C96H148N24O36
2544.1771	2544.1314	0.0457	18.0	3.18	1	[451-472] k.GFYPSDIAVEWESNGQPENNYK.t	C114H158N28O39

Generated by mMass • Open Source Mass Spectrometry Tool • www.mmass.org

MASCOT Search Results

Database : SwissProt 2018_10 (558681 sequences; 200603466 residues)

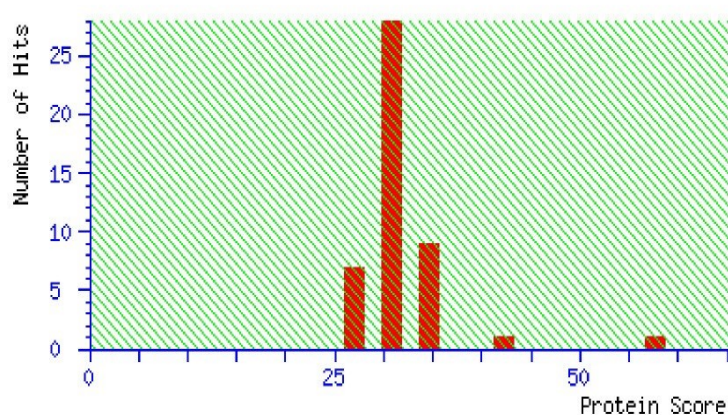
Timestamp : 22 Nov 2018 at 16:18:55 GMT

Top Score : 58 for **IGHG1_HUMAN**, Immunoglobulin heavy constant gamma 1 OS=Homo sapiens OX=9606 GN=IGHG1 PE=1 SV=1

Mascot Score Histogram

Protein score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event.

Protein scores greater than 70 are significant ($p < 0.05$).



Concise Protein Summary Report

Significance threshold: $p < 0.05$

Max. number of hits: AUTO

Preferred taxonomy: All entries

1. [IGHG1_HUMAN](#) Mass: 36083 Score: 58 Expect: 0.95 Matches: 16

Immunoglobulin heavy constant gamma 1 OS=Homo sapiens OX=9606 GN=IGHG1 PE=1 SV=1

Search Parameters

Type of search	: Peptide Mass Fingerprint
Enzyme	: Trypsin
Variable modifications	: Oxidation(M)
Mass values	: Monoisotopic
Protein Mass	: Unrestricted
Peptide Mass Tolerance	: ± 0.1 Da

Appendix C

Peptide Charge State : 1+
Max Missed Cleavages : 1
Number of queries : 703
Selected for scoring : 213

Mascot: <http://www.matrixscience.com/>

Mascot search of protein band in Figure 17 after trypsin digest

Mascot Search Results

Database : SwissProt 2018_10 (558681 sequences; 200603466 residues)

Timestamp : 23 Nov 2018 at 11:26:51 GMT

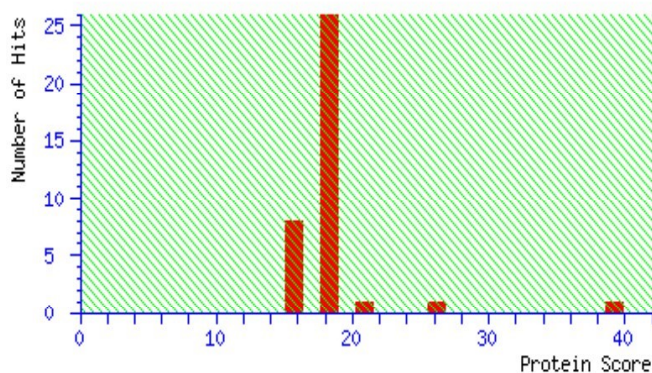
Top Score : 39 for **OPCM_HUMAN**, Opioid-binding protein/cell adhesion molecule OS=Homo sapiens

OX=9606 GN=OPCML PE=1 SV=1

Mascot Score Histogram

Protein score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event.

Protein scores greater than 70 are significant ($p < 0.05$).



Concise Protein Summary Report

Significance threshold: $p < 0.05$ Max number of hits: AUTO

Preferred taxonomy: All entries

1. [OPCM_HUMAN](#) Mass: 37983 Score: 39 Expect: 67 Matches: 15

Opioid-binding protein/cell adhesion molecule OS=Homo sapiens OX=9606 GN=OPCML PE=1 SV=1

[OPCM_PANTR](#) Mass: 37983 Score: 39 Expect: 67 Matches: 15

Opioid-binding protein/cell adhesion molecule OS=Pan troglodytes OX=9598 GN=OPCML PE=2 SV=1

[OPCM_BOVIN](#) Mass: 37890 Score: 33 Expect: 3.1e+02 Matches: 14

Opioid-binding protein/cell adhesion molecule OS=Bos taurus OX=9913 GN=OPCML PE=1 SV=1

[OPCM_RAT](#) Mass: 38043 Score: 27 Expect: 1.1e+03 Matches: 12

Opioid-binding protein/cell adhesion molecule OS=Rattus norvegicus OX=10116 GN=Opcml PE=1 SV=2

[UBACT_NITMO](#) Mass: 7607 Score: 17 Expect: 1.1e+04 Matches: 3

Appendix C

Prokaryotic ubiquitin-like protein UBact OS=Nitrospira moscoviensis OX=42253 GN=ubact PE=3 SV=1

[TAT_HV1A2](#) Mass: 11550 Score: 16 Expect: 1.3e+04 Matches: 3

Protein Tat OS=Human immunodeficiency virus type 1 group M subtype B (isolate ARV2/SF2) OX=11685
GN=tat PE=3 SV=1

[RL21_COXBN](#) Mass: 12738 Score: 16 Expect: 1.3e+04 Matches: 3

50S ribosomal protein L21 OS=Coxiella burnetii (strain Dugway 5J108-111) OX=434922 GN=rplU PE=3 SV=1

[RL21_COXBR](#) Mass: 12752 Score: 16 Expect: 1.3e+04 Matches: 3

50S ribosomal protein L21 OS=Coxiella burnetii (strain RSA 331 / Henzerling II) OX=360115 GN=rplU PE=3
SV=1

[RL21_COXBU](#) Mass: 12752 Score: 16 Expect: 1.3e+04 Matches: 3

50S ribosomal protein L21 OS=Coxiella burnetii (strain RSA 493 / Nine Mile phase I) OX=227377 GN=rplU
PE=3 SV=2

Search Parameters

Type of search	: Peptide Mass Fingerprint
Enzyme	: Trypsin
Variable modifications	: Oxidation(M)
Mass values	: Monoisotopic
Protein Mass	: Unrestricted
Peptide Mass Tolerance	: ± 0.1 Da
Peptide Charge State	: 1+
Max Missed Cleavages	: 1
Number of queries	: 343
Selected for scoring	: 139

Mascot: http://www.matrixscience.com/

MRes Report: The Tumour Suppressor Protein OPCML as an Anticancer Drug: A Biological Therapeutic enabled through Self-splicing Glycophosphatidylinositol mimetics.

Eloise V. N. Morecroft¹, Keith Willison¹, Hani Gabra², Ed Tate¹, Chiara Recchi².

¹Department of Chemistry, South Kensington Campus, ²Ovarian Cancer Action Research Centre, Department of Surgery and Cancer, Hammersmith Campus, Imperial College London.

KEYWORDS: OPCML, epithelial ovarian cancer, GPI anchor, mimetic, DNA methylation, dissemination, structure activity relationship.

ABSTRACT: Epithelial ovarian cancer (EOC) is the most lethal gynaecological malignancy and is almost undetectable until it reaches its later stages, when it disseminates within the peritoneal cavity. Opioid binding protein/cell adhesion molecule-like (OPCML) is a tumour suppressor gene (TSG) that is epigenetically inactivated in EOC. OPCML has three immunoglobulin (Ig) domains, is highly glycosylated and has a glycophosphatidylinositol (GPI) anchor. Previous studies from our group have shown that re-expressing OPCML in cancer cells can prevent their growth in both *in vitro* and *in vivo* models. This is due to OPCML binding to receptor tyrosine kinases (RTKs) and negatively regulating their signalling pathways. Therefore OPCML could represent a valid therapeutic agent in patients with EOC to prevent cancer growth. The aim of this research is to develop OPCML as a potential therapeutic, by first studying the structure activity relationship (SAR) to determine which moieties of OPCML are required for an active TSG function and whether glycosylation is necessary. Elucidating a simplified structure of OPCML would enable a more effective and efficient expression/purification of the protein whilst still having all the benefits of a TSG.

Introduction

Epithelial ovarian cancer (EOC) is caused by genetic changes in the ovarian surface epithelium (OSE), the layer of flat-to-cuboidal epithelial cells that cover ovaries. It is believed that the rupture and repair cycle during ovulation can contribute to such genetic changes, as can the epithelial to mesenchymal transition OSE cells undergo during such cycle. This causes migration to the stroma, leading to the formation of inclusion cysts and ultimately tumorigenesis⁸.

EOC is asymptomatic until the later stages of the malignancy, when dissemination into the peritoneal cavity has begun. This dissemination is caused by epithelial cells losing the adhesion between them to become migratory, a process known as “epithelial to mesenchymal transition” or EMT. Late detection of the cancer makes extensive surgical

procedures a necessity, though recurrence after surgery remains high^{5,17,32}.

Appendix D

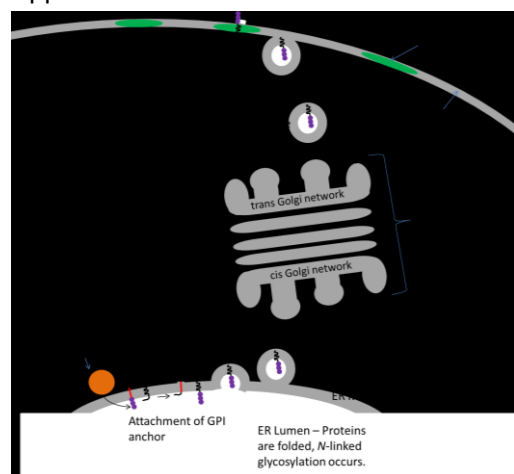


Figure 55 Simplified schematic of the secretory pathway of GPI anchored proteins such as OPCML.

Opioid binding protein/cell adhesion molecule- like (OPCML) is a tumour suppressor gene (TSG) in various cancers including EOC³¹. OPCML is a non-transmembrane protein that is part of the IgLON family of cell adhesion molecules, which has three C2 like Ig (immunoglobulin-like) domains and is highly glycosylated with six glycosylation sites^{31,154,155}. OPCML is associated via a glycosphosphatidylinositol (GPI) anchor to the cell membrane, where it is localised within cholesterol rich regions of the membrane, known as detergent resistant membranes (DRMs)³².

Many TSGs undergo epigenetic inactivation: CpG islands are areas with a high density of cytosine nucleotides next to guanine, and these areas can be methylated at the cytosine causing inactivation of the gene via transcriptional repression^{5,156, 157}. Research has shown that OPCML has CpG islands that are normally unmethylated, but has CpG island methylation in 82 % of ovarian cancer cells⁵.

DNA methyltransferases (DNMT) are the enzymes that catalyse the methylation of CpG islands. Methyltransferase inhibitors can be used as therapeutics to prevent methylation and subsequent silencing of TSGs. However, there is concern over the effect such inhibitors could have on normal cells, with any epigenetic therapy needing to be highly selective^{46,47,158}.

OPCML mediates its tumour suppressor capacity through the regulation of a number of receptor tyrosine kinases (RTKs). OPCML binds the extracellular domains of a specific set of RTKs, sequesters the RTK to the detergent resistant membranes and alters RTK signalling via proteasomal degradation³².

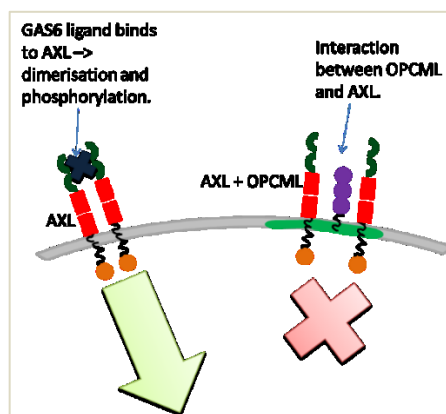


Figure 56 Schematic of OPCML interacting with AXL. The figure shows OPCML binding AXL and interrupting the resultant signaling, causing loss of EMT factors and preventing migration.

AXL is one RTK OPCML interacts with in this way. AXL has two Ig domains and two FNIII (fibronectin type 3 like) domains¹⁵⁹, and it is part of the TAM family of RTKs. Growth arrest specific gene-6 (GAS6) is the ligand for AXL. Upon GAS6 binding to AXL, AXL dimerisation and phosphorylation of tyrosine residues occurs allowing intracellular signalling molecules to dock, leading to a signalling pathway triggering EMT^{159,160,161}. It has been shown that the expression of AXL is higher in ovarian cancer tissues compared to normal and that this increases tumour cell migration^{159,160}. Recent data from our lab have shown that OPCML interacts with AXL, causing sequestering into DRMs; the presence of OPCML causes the loss of AXL phosphorylation and the arrest of the signalling cascade preventing induction of Slug and EMT¹⁶². This data suggests that the re-expression of OPCML into patients with EOC could possibly act to prevent spreading via reduction of EMT.

Due to the extracellular nature of OPCML, the possibility of a direct tumour suppressor therapy rather than gene therapy can be explored³².

As mentioned, OPCML has a GPI anchor. GPI anchors are posttranslational modifications that anchor proteins to the cell membrane. The structure of GPI anchors can be seen in figure 3. Since GPI anchors have complex structures, which causes difficulties with studying and synthesis, GPI mimetics may represent a valid alternative^{56,63,75,163,164}.

Appendix D

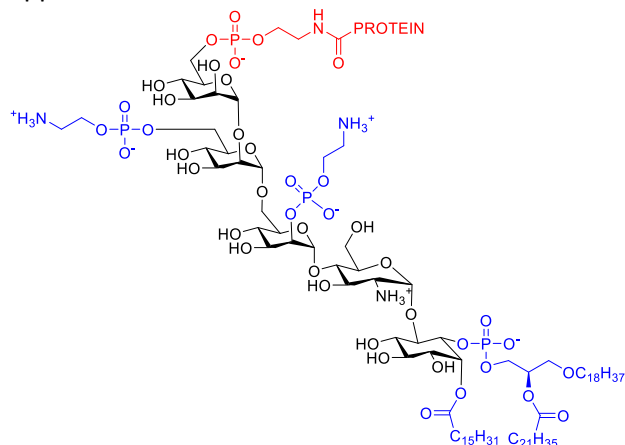


Figure 57 Example of a GPI anchor from human erythrocyte acetylcholinesterase showing the structure¹⁶³. Phospholipids are shown in blue, a

phosphoethanolamine linker in red, and a glycan core in black.

If a functional, simplified OPCML protein could be expressed and purified, a GPI-anchor mimetic could then be attached and this complex used as a direct tumour suppressor therapy.

The functional moieties of OPCML will be investigated via testing each Ig domain using migration assays. The final aim will be to find a functional simplified OPCML structure, which can then be used to generate a direct tumour suppressor therapy.

In this study the design and structure of four OPCML constructs is reported, one full length and three expressing different single OPCML domains, with the aim to show whether these constructs function in the same way as wild type OPCML.

Results and Discussion

Design of Constructs

Protein modelling was used to predict the structure of the OPCML protein. Structure prediction using Phyre2 (see Supplementary Information) enabled the reported location of the Ig domains to be confirmed¹⁶⁵⁻¹⁶⁷. Phyre 2 predicts the structure of a protein via comparison of the given protein sequence with the sequences of proteins with known structures, rather than simulating the folding of proteins. The structure is generated by comparing the similarity of the sequences to give alignments and then these alignments are used to form the model structure¹⁶⁶. Full length wild type OPCML, each single OPCML domain, and OPCML with an HA tag were all run on the Phyre2 system. The predicted structures were able to confirm the areas listed as the start and end of each domain in the UniProt database¹⁶⁷. This aided in the design of the four constructs. However there are some limitations to Phyre2, which are addressed in the paper by Kelley et al¹⁶⁶. Firstly, if the input protein sequence is found not to be homologous with any known structure the results generated are very unreliable if generated at all. Furthermore it is difficult to predict the effect point mutations have on structure. Therefore the structure prediction using Phyre2 was only used to confirm the already stated position of the domains and not relied on solely.

Figure 4 shows the schematic structure of wild type OPCML with the location of glycosylation sites and Ig domains (see Supplementary Information for full sequences). The figure also shows a schematic of the four constructs tested for functionality. The first step is to determine the relative importance of the three Ig domains of OPCML, thus the constructs were designed to test each of the domains independently. There are four constructs:

full length OPCML (FL), domain 1 (D1), domain 2 (D2) and domain 3 (D3). Glycosylation sites and sequences between domains were preserved where possible in all the constructs. An HA-tag was included in each sequence to aid in detection of the protein during western blots and immunofluorescence. Each construct was designed to act as an entry clone for Gateway cloning. Figure 5 shows a schematic of the Gateway cloning procedure (see methods). Each construct was recombined with the destination vector (pLenti CMV Puro Dest from Addgene) using Gateway LR Clonase II Enzyme Mix (Invitrogen) to form the desired expression vector and any by-products.

Gateway cloning uses att recombination to enable chosen DNA sequences (which are present between the recombination sequences) to be swapped from one vector into another via an enzyme. As Gateway LR Clonase kit would be used for cloning, the attL1 and attL2 recombination sequences were added to flank the construct sequences. As a precaution, unique restriction sites XbaI and SalI were added between the recombination sequences and OPCML construct so that, in case of difficulties with Gateway cloning, the sequence could be removed and a different approach explored. The full sequences of each construct can be found in the supplementary information.

Cloning the Constructs into Lentiviral vectors

After the DNA constructs had been designed and received, the first step was to amplify the plasmids in bacteria and then purify them. Following this, the Gateway cloning kit was used.

Initial attempts to verify the identity of the received DNA did not yield the expected result. When running the digested DNA on agarose gels to confirm the presence of the restriction sites it was found that the XbaI site was not

Appendix D

present. Upon further investigation it was discovered that the DNA was being methylated by Dam methylase in the bacteria (Alpha-Select Silver Competent Cells from Biotin) and the presence of a GA upstream of the XbaI site was blocking cleavage by the enzyme. Additionally, sequencing the first attempt at Gateway cloning yielded a sequence that did not include the DNA constructs, suggesting that there had been no swapping between recombination sites.

The destination vector contains a gene that confers resistance to the antibiotic ampicillin as well as a *ccdB* gene. The *ccdB* gene is a “suicide” gene that targets DNA gyrase preventing it from functioning correctly. After cloning, any by-products and any un-swapped destination vectors that still have the *ccdB* gene present will prevent the bacteria from growing. Destination vectors that have been swapped will not have the *ccdB* gene anymore, as the *ccdB* gene is between the att sites that are swapped during successful Gateway cloning. However the Alpha Select Silver Efficiency competent cells were resistant to the *ccdB* gene because they carry a mutation in the gyrase. So, all the bacteria transformed with by-products and unreacted destination vectors grew and there was no selection.

For subsequent cloning methyltransferase deficient *E. Coli* cells were used (*dam/dcm* Competent *E. Coli* from New England Biolabs) preventing blocking of the restriction site. This bacteria also had no resistance to the *ccdB* gene.

The Gateway cloning procedure was attempted twice more using these new competent bacteria. The first pLenti constructs made were sent for sequencing (Beckman) but only the FL construct could be sequenced. Thus the cloning was repeated with a positive control of a known entry vector, and as a negative control no entry vector. The transformation of the negative control should not have given any colonies due to no swapping having taken place, thus the *ccdB* gene should ensure the destination vector does not grow. However, after transforming the products of the LR Clonase reaction the negative control still gave positive colonies perhaps due to a mutation in the *ccdB* gene of plasmid.

Therefore colony polymerase chain reaction (PCR) was carried out using primers designed to work for all the constructs (see Supplementary Information) to identify the positive colonies that had swapped successfully.

An adequate number of colonies were picked for each condition. Several colony PCRs were carried out in order to obtain at least one positive colony for each construct. Figure 6 shows DNA agarose gels of the PCR products. FL is approximately 1000 base pairs and D₁, D₂, and D₃ all around 550 base pairs. The results clearly show at least one PCR product of the right size for each construct.

One colony for each construct was selected except from D₂ where two colonies were found and these were labelled D_{2.1} and D_{2.2}. All the colonies were amplified, purified and sent for sequencing. The results showed that all constructs apart from pLenti FL in this second attempt were

successfully sequenced. This suggests that the primers could not recognise the binding site in the pLenti FL construct indicating the Gateway cloning reaction did not work. The sequence received for the pLenti FL during the first attempt at Gateway cloning and the sequences for the remaining constructs from this second attempt were then further analysed. This sequence analysis defined the presence of each construct between att sites within the pLenti backbone.

Verification of OPCML Protein Expression via Transient Transfection

In order to confirm the expression of the new pLenti OPCML constructs, cells were transiently transfected. HEK293T cells were chosen as a model system. 48 hours after the transfection, the cells were fixed with methanol and immunofluorescence staining was carried out. Cells were stained using a mouse monoclonal anti-OPCML antibody and analysed by confocal microscopy (Figure 7). However, only FL and D₃ proteins could be visualised. The mouse anti-OPCML antibody is monoclonal, meaning it binds the protein specifically in one epitope only. These results indicate that the sequence recognised by this primary antibody is in domain 3.

Transient transfections were repeated in the ovarian cancer cell line SKOV3. Cells were transfected with the pLenti constructs and after two days cells were fixed using PFA (see Methods). Cells were then stained using a mouse anti-HA in order to get a signal for each construct as seen in figure 8. The results clearly indicated that all proteins were expressed in SKOV3 cells (Figure 8).

Figure 7 and figure 8 show the transfection efficiency of each construct in each cell line. In HEK293T cells the transfection efficiency was higher than in the SKOV3 cell line as can be seen by the immunofluorescence images. As the transfections were transient and undertaken using the pLenti constructs, which are large in size (around 9000 bp compared to mammalian expression vectors which are around 5000 bp), it is not unexpected that the transfection efficiency would be low. Moreover it was observed that all the constructs are localised to the membrane of the cells as with the wild-type (WT) OPCML. This indicates that each construct, though truncated, may be following the same secretory pathway as WT OPCML to the cell surface. Figure 9 shows OPCML D₃ in more detail. From this image it is clear that the construct is localised at the cell membrane similarly to WT OPCML.

In parallel to immunofluorescence experiments, cells were lysed and western blotting was performed for both cell lines as an alternative method to detect the presence of each protein. As mentioned, when compared to the HEK293T cells, the SKOV3 cell line did not have as good a transfection efficiency and western results were not analysable due to background noise and low signals. However, the western of HEK293T lysates (Figure 10) showed the expression of the FL construct, which can be

Appendix D

detected as a band at around 50 kilo daltons, the expected molecular weight of the FL construct. The expressions of the D₁, D₂ and D₃ constructs, which have an expected molecular weight of approximately 20 kilo daltons, cannot be clearly identified from the western due to the presence of numerous unspecific bands from the HA antibody as seen in figure 10.

Most likely, the cause of the lack of signal for the protein is due to the low transfection efficiency, which could be seen from the scarcity of positive cells in the immunofluorescence images. As such, it was decided to prepare the lentivirus.

Production of OPCML Lentivirus

The constructs expressing each of the three domains and the full length OPCML were used to produce lentiviruses in HEK293T cells using a well established technique. The lentiviral system was selected due to the high transduction efficiency it provides. It also overcomes some of the limitations of other methods, such as the use of adenoviruses, because lentiviruses enable the integration of the mutated OPCML constructs into the genome whilst systems such as adenoviruses do not^{123,168,169}. For each

construct, the genomic plasmid (the pLenti constructs), a second generation packaging plasmid (p8.9), and an envelope plasmid (pMDG containing the VSV-G gene encoding for the envelope) were used to transfect HEK293T cells (see Methods). The viruses produced by HEK293T cells were then used to infect SKOV₃ cells (infection performed by Dr E Karali) in order to produce stable cells lines expressing either full length OPCML or the truncated constructs of the protein.

Cells were infected post Polybrene treatment. Polybrene was used to enhance the transduction of cells with the lentiviruses (see Methods). When cells are transduced with virus the first step is the viral particles reaching the cell surface, next non-specific binding to the cell, and finally binding to a specific receptor on the cell surface¹⁷⁰. Numerous papers have looked at the mechanism through which Polybrene increases the transduction efficiency and it is believed to occur through the enhancement of virus adsorption and charge shielding^{170,171}.

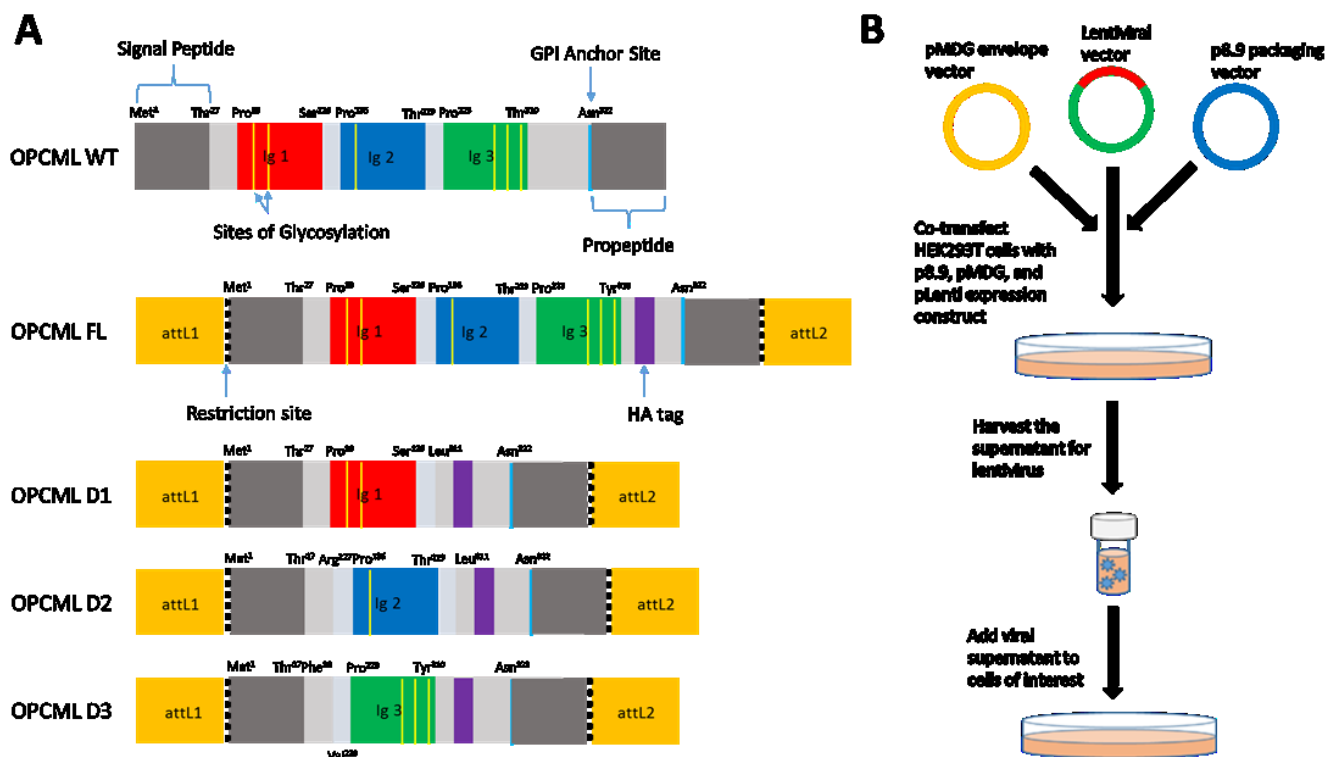


Figure 4: A) Schematic representation of the four constructs designed to examine the function of each Ig domain of OPCML: A) The structure of wild type OPCML with glycosylation sites denoted by yellow lines and areas of importance such as signal peptide, GPI anchor, and Propeptide labelled. The full length (FL) construct includes the full length sequence of wild type OPCML with added sequences for the attL sites and an HA tag, plus restriction sites between the attL and the rest of the sequence to enable sub-cloning of the gene if needed. Schematics of D1, D2, and D3 have the added attL sites, HA tag and restriction sites with only one domain. B) Schematic representation of the steps required to produce a stable cell line using a lentiviral approach.

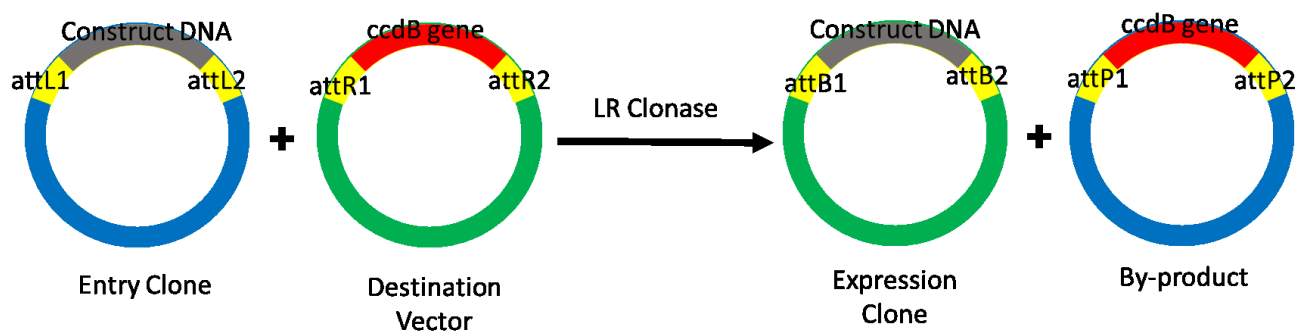


Figure 5 Schematic representation of Gateway cloning using the LR Clonase reaction (see Methods).

Appendix D

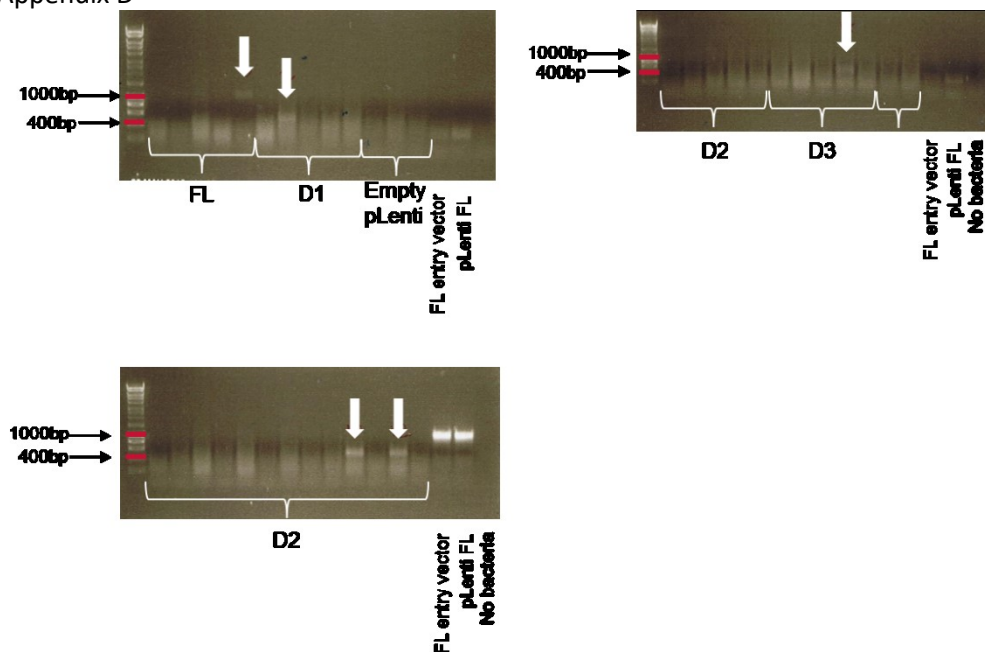


Figure 6 DNA gel electrophoresis results indicating the positive DNA colonies for each construct: Colony PCR was performed with results showing the presence of pLenti FL around 1000 base pairs and pLenti D1, 2, and D3 constructs around 500 base pairs.

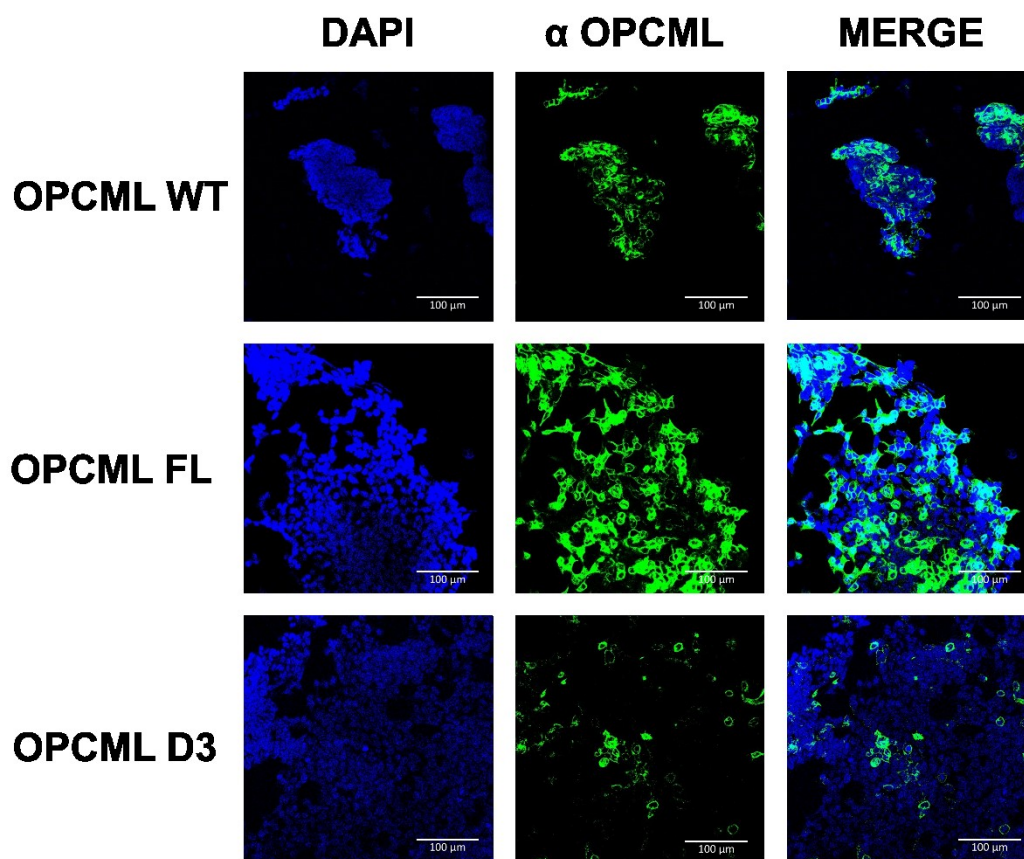


Figure 7 Confocal immunofluorescence microscopy showing OPCML expression (green): HEK293T cells were transfected using one of the FL, D1, D2 or D3 constructs. After 48 hours cells were fixed using methanol fixation and stained with mouse anti-OPCML (green) and DAPI for nuclear staining (blue). Confocal microscopy using LEICA SP5 images show the transfection

Appendix D

efficiency in HEK293T. Only OPCML WT, FL and D3 could be visualised using this antibody indicating the monoclonal antibody recognises the domain 3 of the OPCML protein. Scale bars 100 μ m. Confocal images taken by Dr Evdoxia Karali.

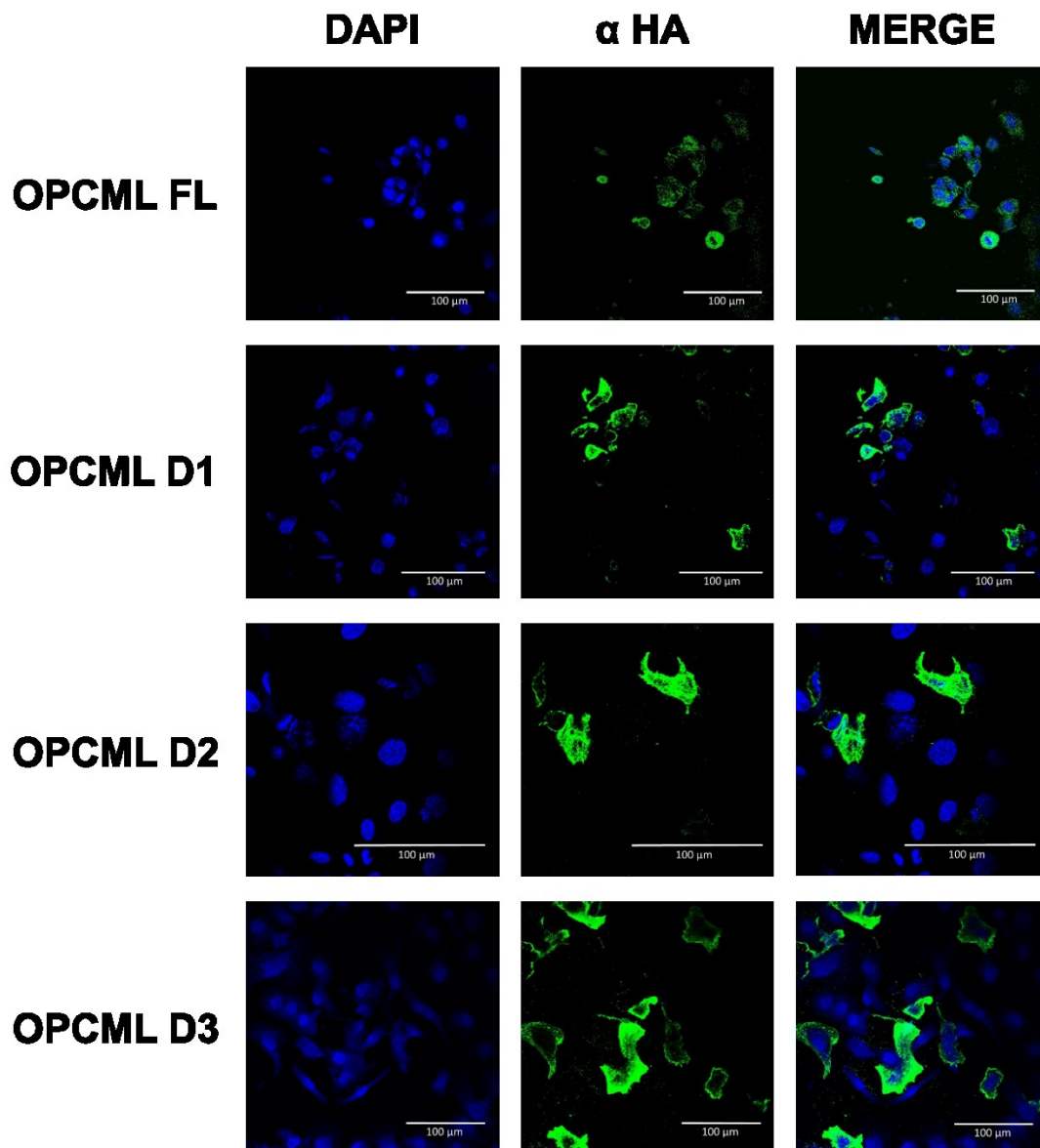
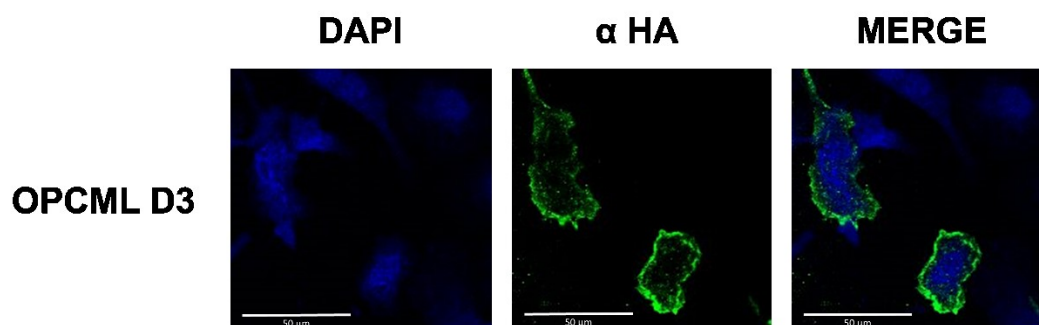


Figure 8 Confocal immunofluorescence microscopy showing the expression of HA tagged protein (green): SKOV₃ cells were transfected using one of the FL, D1, D2 or D3 constructs. After 48 hours, cells were fixed using PFA fixation and stained with monoclonal mouse anti-HA (green) and DAPI (blue). Confocal microscopy using LEICA SP5 images show the transfection efficiency in SKOV₃ cells. Scale bars 100 μ m. Confocal images taken by Dr Evdoxia Karali.



Appendix D

Figure 9 Confocal Immunofluorescence microscopy showing protein expression through use of anti-HA tag antibody (green): The images show the localisation of the D₃ OPCML protein to the cell membrane. Scale bars 50 µm. Confocal images taken by Dr Evdokia Karali.

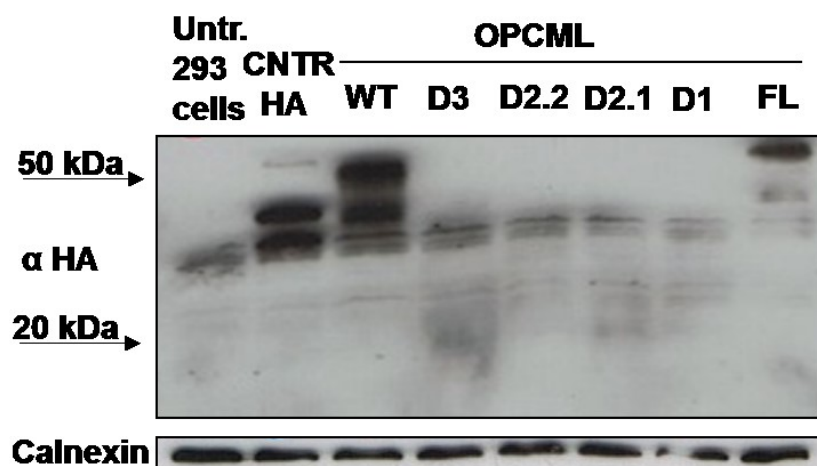


Figure 10 Western blot to demonstrate the expression of construct proteins: HEK 293T cells were transfected using FL, D₂, D₂, or D₃ constructs with OPCML WT and untreated cells as controls. After 24 hours cells were lysed, the protein concentration was calculated using the BCA method and SDS PAGE was performed. Calnexin was used as a loading control. Mouse monoclonal anti-HA tag antibody was used. FL OPCML construct can be seen around 50 kDa. Unspecific bands from the HA antibody mask any signals due to constructs D₁, D₂ and D₃, which would be present at around 20 kDa.

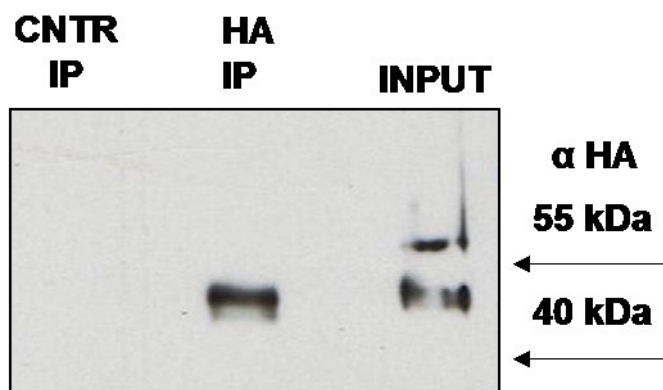


Figure 11 Western blot analysis of immunoprecipitation of FL protein: SKOV₃ cells were transiently transfected using pLenti FL. After 24 hours the cells were lysed and cross-linked immunoprecipitation was performed and SDS PAGE. Mouse monoclonal anti-HA antibody was used. FL OPCML protein can be seen around 50 kDa. A control IP of the procedure carried out using only beads was used to confirm specificity.

Immunoprecipitation to Confirm Successful Transient Transfection of SKOV₃ cells

Cross-linked Immunoprecipitation was used as a technique to further confirm the success of transient transfection of SKOV₃ cells with the pLenti FL construct. The FL protein was immunoprecipitated using mouse anti-HA antibody as can be seen in figure 11. In the input lane the presence of the transfected HA tagged FL protein around 50 kDa in the initial SKOV₃ lysate is confirmed. The band around 50 kDa is also present in the HA IP

column confirming the successful immunoprecipitation of the FL protein. The control beads IP (last column) gives no signal, indicating the IP signal is specific.

Conclusion

The aim of this research was to investigate the structure activity relationship of OPCML, which has not been done previously, by studying the importance of the three OPCML Ig domains separately. In this study, entry clones

Appendix D

have been designed that are able to swap into destination vectors. After confirmation that the right plasmids had been obtained following recombination, the FL and the truncated OPCML constructs D1, D2 and D3 were observed in cell lines upon transient transfections. Immunofluorescence further proved that the proteins were correctly localised to the cell membrane, which is expected as OPCML is anchored to the external cell membrane via a GPI anchor³².

Research is still being undertaken to determine the relative importance of the domains, with future experiments including migration assays for testing protein function, further immunoprecipitation to investigate protein interactions, and mouse studies to see *in vivo* effects of any functional single domains.

Once current research has concluded, the domains will be further characterised by analysing whether the six glycosylation sites are necessary for TSG function. Depending on the results, the function of various combinations of domains might also be investigated. This will be performed using similar experimental techniques to those described.

Due to OPCML having the potential to act as a direct tumour suppressor therapy³², the wider purpose of this research would be to identify the structure needed for active tumour suppressor function and to develop this into a therapeutic for clinical trials.

To complete this overall aim, future research needs to include the synthesis and development of GPI anchor mimetics. A GPI mimetic could then be attached to the simplified OPCML structure to create a protein anchor complex. This complex would then be able to be tested as a direct tumour suppressor therapy. A technique to produce this protein anchor complex would then need to be developed to enable large scale synthesis.

The research being undertaken in this paper is the beginning of a wider aim to develop a therapeutic to prevent the spreading of EOC. This is of importance due to the high number of women affected by EOC and the high mortality rates even after surgery.

Methods

Plasmids/Constructs

FL, D1, D2 and D3 entry clone constructs were purchased from Life Technologies designed as entry clones. Constructs were amplified in competent bacteria (*dam*⁻/*dcm*⁻ Competent *E. Coli* from New England Biolabs). Gateway LR Clonase II Enzyme mix kit from Life Technologies was used to transfer the constructs from entry into destination vectors following the manufacturers protocol.

P8.9 and pMDG plasmids were provided by Dr Ari Fassati, University College London UK. pLenti CMV Puro DEST was purchased from Addgene.

Transformations

Transformations were undertaken using Alpha-Select Silver Competent Cells from Bioline and *dam*⁻/*dcm*⁻ Competent *E. Coli* from New England Biolabs following the manufacturer's instructions.

Competent cells were thawed on ice for 20 minutes. After thawing the cells were gently mixed and 50 µL aliquots (one per each transformation) were transferred to 1.5 mL eppendorf, which were pre-chilled on ice. 1 µL of DNA (100 ng/mL) was added to the competent cell aliquot and mixed gently. The cells and DNA were incubated on ice for 30 minutes, then heat shocked for 45 seconds at 42 °C and placed on ice for 2 minutes. 950 µL of pre-warmed YT media containing no antibiotic was then added to the cells and the eppendorf shaken for 1 hour at 30 °C. The transformation mixture was then spread on agar plates containing the appropriate antibiotic, allowed to dry and incubated at 30 °C overnight.

Purification was undertaken using GeneJET Plasmid Miniprep kit from Thermo Scientific, or Highspeed Plasmid Maxi kit from Qiagen both used as per the manufacturer's instructions.

Colony PCR

Colony PCR was carried out using standard reagents from New England Biolabs on a MJ Research PTC-200 peltier thermal cycler. Samples were denatured for 10 minutes at 96°C and then cycled for 1 minute at 95°C, 1 minute at 59°C, and 1 minute 68°C for 30 cycles. Then they were kept at 68°C for 10 minutes. Samples were then analysed using DNA electrophoresis. Agarose gel electrophoresis was carried out using 1 % agarose gels and visualised using a UV light chamber.

Cell Lines

HEK293T and SKOV3 were purchased from ATCC.

Cell Culture

HEK293T cells were grown in DMEM supplemented with 10 % FCS, and 2 mM L-glutamine. SKOV3 cells were grown in RPMI supplemented with 10 % FCS, 100 units penicillin/0.1 mg/mL streptomycin, and 2 mM L-glutamine. Cell culture reagents were purchased from Gibco and Sigma Aldrich. Both cells lines were grown in a humidified environment with 5 % CO₂ at 37°C. Cells were plated 24 hours before transfection.

Transient Transfection

HEK293T and SKOV3 cells were plated 24 hours prior to transfection in 6 well plates. 2.5 µg of each DNA construct was made up in 200 µL OptiMEM (Gibco). 87.5 µL Lipofectamine 2000 (Invitrogen) was added to 2000 µL

Appendix D

OptiMEM. Each OptiMEM mixture was vortexed, spun down, and incubated at room temperature for 5 minutes. 200 μ L of the Lipofectamine 2000/OptiMEM mixture was added to each DNA construct mixture to give a total volume of 400 μ L. This was vortexed, spun down, and incubated at room temperature for 20 minutes. Whilst incubating, cell media was removed and replaced with OptiMEM. After the incubation the lipid/DNA mixture was slowly added to the cells dropwise with swirling. 5 hours after addition, the cell media was removed and replaced with the appropriate media for the cell line.

Antibodies

Monoclonal mouse anti-OPCML and polyclonal goat anti-OPCML were purchased from R&D Systems. Polyclonal rabbit anti-calnexin was purchased from Enzo Life Sciences. Horseradish peroxidase-conjugated secondary antibodies were bought from Dako. Alex Fluor 488 goat anti-mouse and Alexa Fluor 546 donkey anti-goat were purchased from Invitrogen.

Immunofluorescence Microscopy

SKOV3 cells were grown on glass coverslips and fixed in PBS, 4% paraformaldehyde for 20 minutes, washed with PBS, 50mM NH_4Cl , then incubated for 30 minutes with PBS containing 0.05 % saponin and 2 % FCS at room temperature. The coverslips were incubated with primary antibodies diluted in PBS containing 0.05 % saponin and 2 % FCS for 1 hour at room temperature in a dark wet chamber, washed then incubated with secondary antibodies (diluted in the same mixture) for 30 minutes under the same conditions.

HEK293T cells were grown on glass coverslips and fixed in ice cold methanol for 5 minutes at -20°C . The coverslips were then washed with PBS and blocked for 30 minutes with 3 % BSA in PBS. The coverslips were incubated with primary antibodies diluted in 1 % BSA for 1 hour at room temperature in a dark wet chamber, washed, then incubated with secondary antibodies (diluted in the same mixture) for 30 minutes under the same conditions.

All antibodies were diluted according to the manufacturer's instructions. All coverslips were then mounted onto slides using ProLong Gold antifade reagent with DAPI (Molecular Probes Life Technologies) and left to dry at room temperature before imaging on a Leica SP5 confocal.

Western Blot

Lysates were taken using RIPA (Radio-Immunoprecipitation assay) buffer (150 mM NaCl, 1.0 % IGEPAL[®] CA-630, 0.5 % sodium deoxycholate, 0.1 % SDS, 50 mM Tris, pH 8.0) and protease and phosphatase inhibitors. Protein concentration was determined using Thermo Scientific Pierce BCA protein assay kit as per the manufacturer's instructions. Absorbance in 96 well plates

was measured using Molecular devices OPTImax tuneable microplate reader. SDS-PAGE was used to separate the proteins using 15% acrylamide gel. Proteins were transferred to a nitrocellulose membrane, blocked for 30 minutes (5 % milk in TBST) before incubating overnight with primary antibody. The membrane was then washed (3 times for 10 minutes each time) and incubated in secondary antibody for 45 minutes. Membranes were then incubated with Immobilon Western Chemiluminescent HRP (Millipore) for 1 minute before being developed on Amersham Hyperfilm ECL films (GE Healthcare). Antibodies were diluted in 5% milk in TBST.

Lentivirus Production

HEK293T cells were plated 24 hours prior to transfection in 10 cm Petri dishes in order to achieve 90 -95 % confluency in RPMI media with 10 % FCS and L-glutamine.

12 μ g of each pLenti plasmid, 10.2 μ g packaging plasmid (p8.9), and 1.8 μ g VSV-G (pMDG) were mixed and diluted up to 1500 μ L OptiMEM, vortexed and incubated at room temperature for 5 minutes. 60 μ L Lipofectamine 2000 (Invitrogen) was diluted up to 1500 μ L in OptiMEM, vortexed and incubated at room temperature for 5 minutes. Both the plasmid mixture and the Lipofectamine 2000 mixture were mixed, vortexed and incubated for 20 minutes at room temperature. The media of the HEK293T cells was replaced with OptiMEM. The mixture was then added dropwise to cells with swirling. After 5 hours the cell media was replaced with RPMI (10 % FCS and L-glutamine). The cells were then incubated for 48 hours before the virus was collected in the cell medium, centrifuged for 4 minutes at 1500 rpm and filtered (0.45 μ L filter). The virus was then aliquoted and stored at -80°C .

Polybrene (Sigma Aldrich) was added to cell media at a concentration of 8 μ g/mL 30 minutes before transduction.

Producing Stably Transduced Cell lines

SKOV3 cells were plated in 24 well plates and allowed to grow to confluence prior to transduction with the virus. 24 hours after transduction, transduced cells were selected by using medium containing puromycin (10 μ g/mL). Once selected cells were confluent, they were amplified into a larger flask and the process continued until stocks of stable cell lines had been produced.

Immunoprecipitation

Pierce[®] Crosslink Immunoprecipitation kit (Thermo Scientific) was used as per the manufacturer's instructions. All centrifuge steps were carried out at 1000 x g for 1 minute at 4°C .

SKOV3 cells were grown in a 10 cm petri dish for 24 hours after transient transfection with pLenti FL construct. The cells were then washed with ice cold PBS before 1 mL IP Lysis/Wash Buffer with Protease and Phosphatase

Appendix D

Inhibitors was added and the dish left to incubate for two hours shaking at 4°C. Lysates were then taken. 20 µL Pierce Protein A/G Plus Agarose was added to a spin column and washed twice with 200 µL 1X Coupling Buffer. The column was incubated with 5 µL 20X Coupling Buffer, 10 µL HA antibody (1 mg/mL), and 85 µL distilled water at room temperature for 1 hour. After incubation the column was washed once with 100 µL 1X Coupling Buffer and twice with 300 µL. 9 µL 2.5 mM DSS (disuccinimidyl suberate), 2.5 µL 20X Coupling Buffer and 38.5 µL distilled water were then added to the column and the reaction incubated for 1 hour on a rotator, at room temperature. After incubation the column was centrifuged. 50 µL of Elution buffer was then added and the column centrifuged. The column was then washed twice more with 100 µL Elution buffer and twice with 200 µL IP Lysis/Wash Buffer. Then incubated overnight on a rotator at 4°C with pre-cleared lysate (400 mg) obtained via use of Pierce Control Agarose Resin Slurry. After incubation the column was washed five times with 200 µL Lysis/Wash Buffer and once with 100 µL 1X Conditioning Buffer and the antigen eluted using Elution Buffer. After elution samples were then heated in sample buffer and western blotting was used to visualise the proteins.

ACKNOWLEDGMENTS

I would like to acknowledge the support of the Engineering and Physics Research council (EPSRC), and the Ovarian Cancer Action Research Centre Core for lab funding.

I would like to thank Professor Ed Tate, Dr Chiara Recchi, Professor Keith Willison, and Professor Hani Gabra for their supervision and guidance throughout the project. I would like to thank Dr Evdoxia Karali for taking the images on the confocal, and collecting and adding viral supernatant to the cell lines. I would also like to thank the rest of the Gabra research group.

REFERENCES

1. Heal, W. P., Wright, M. H., Thinon, E. & Tate, E. W. Multifunctional protein labeling via enzymatic N-terminal tagging and elaboration by click chemistry. *Nat. Protoc.* **7**, 105–17 (2012).
2. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2018. *CA. Cancer J. Clin.* **68**, 7–30 (2018).
3. Alberts, B. *et al.* *Molecular Biology of the Cell*, 5th Ed.
4. Sever, R. & Brugge, J. S. Signal Transduction in Cancer. *Mutat. Res. Mol. Mech. Mutagen.* **555**, 1–2 (2004).
5. Sellar, G. C. *et al.* OPCML at 11q25 is epigenetically inactivated and has tumor-suppressor function in epithelial ovarian cancer. *Nat. Genet.* **34**, 337–343 (2003).
6. Cancer Research UK. <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer>, Accessed January 2019

7. Cancer Research UK. <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer/incidence#heading-Four>, Accessed January 2019
8. Auersperg, N., Wong, A. S. T., Choi, K., Kang, S. K. & Leung, P. C. K. Ovarian Surface Epithelium : Biology , Endocrinology ,. *Endocr. Rev.* **22**, 255–288 (2001).
9. Auersperg, N., Ota, T. & Mitchell, G. W. E. Early events in ovarian epithelial carcinogenesis: Progress and problems in experimental approaches. *Int. J. Gynecol. Cancer* **12**, 691–703 (2002).
10. Hall, J., Paul, J. & Brown, R. Critical evaluation of p53 as a prognostic marker in ovarian cancer. *Expert Rev. Mol. Med.* **6**, 1–20 (2004).
11. Paul, A. & Paul, S. The breast cancer susceptibility genes (BRCA) in breast and ovarian cancers. *Front. Biosci.* **19**, 605–618 (2014).
12. Meng, Q., Xia, C., Fang, J., Rojanasakul, Y. & Jiang, B. H. Role of PI3K and AKT specific isoforms in ovarian cancer cell migration, invasion and proliferation through the p70S6K1 pathway. *Cell. Signal.* **18**, 2262–2271 (2006).
13. Le Page, C. *et al.* From gene profiling to diagnostic markers: IL-18 and FGF-2 complement CA125 as serum-based markers in epithelial ovarian cancer. *Int. J. Cancer* **118**, 1750–1758 (2006).
14. Trinh, X. B. *et al.* The VEGF pathway and the AKT/mTOR/p70S6K1 signalling pathway in human epithelial ovarian cancer. *Br. J. Cancer* **100**, 971–978 (2009).
15. Lafky, J. M., Wilken, J. A., Baron, A. T. & Maihle, N. J. Clinical implications of the ErbB/epidermal growth factor (EGF) receptor family and its ligands in ovarian cancer. *Biochim. Biophys. Acta - Rev. Cancer* **1785**, 232–265 (2008).
16. Monk, B. J., Minion, L. E. & Coleman, R. L. Anti-angiogenic agents in ovarian cancer: Past, present, and future. *Ann. Oncol.* **27**, i33–i39 (2016).
17. Gloss, B. S. & Samimi, G. Epigenetic biomarkers in epithelial ovarian cancer. *Cancer Lett.* **342**, 257–263 (2014).
18. Colombo, N., Lorusso, D. & Scollo, P. Impact of recurrence of ovarian cancer on quality of life and outlook for the future. *Int. J. Gynecol. Cancer* **27**, 1134–1140 (2017).
19. Orr, B. & Edwards, R. P. Diagnosis and Treatment of Ovarian Cancer. *Hematol. Oncol. Clin. North Am.* **32**, 943–964 (2018).
20. M.J., D. *et al.* CA125 in ovarian cancer: European Group on Tumor Markers guidelines for clinical use. *Int. J. Gynecol. Cancer* **15**, 679–691 (2005).
21. Bast, R. C., Hennessy, B. & Mills, G. B. The biology of ovarian cancer: New opportunities for translation. *Nat. Rev. Cancer* **9**, 415–428 (2009).
22. Cortez, A. J., Tudrej, P., Kujawa, K. A. & Lisowska, K. M. Advances in ovarian cancer therapy. *Cancer Chemother. Pharmacol.* **81**, 17–38 (2018).
23. Moufarrij, S. *et al.* Epigenetic therapy for ovarian cancer: Promise and progress. *Clin. Epigenetics* **11**, 1–11 (2019).
24. Perren, T. J. *et al.* A Phase 3 Trial of Bevacizumab in Ovarian Cancer. *N. Engl. J. Med.* **365**, 2484–2496 (2011).
25. Aghajanian, C. *et al.* OCEANS: A randomized, double-blind, placebo-controlled phase III trial of chemotherapy with or without bevacizumab in patients with platinum-sensitive recurrent epithelial ovarian, primary peritoneal, or fallopian

Appendix D

- tube cancer. *J. Clin. Oncol.* **30**, 2039–2045 (2012).
26. Audeh, M. W. *et al.* Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: A proof-of-concept trial. *Lancet* **376**, 245–251 (2010).
27. Fong, P. C. *et al.* Inhibition of Poly(ADP-Ribose) Polymerase in Tumors from BRCA Mutation Carriers Peter. *N. Engl. J. Med.* **359**, 229–241 (2004).
28. Walters, C. L., Arend, R. C., Armstrong, D. K., Naumann, R. W. & Alvarez, R. D. Folate and folate receptor alpha antagonists mechanism of action in ovarian cancer. *Gynecol. Oncol.* **131**, 493–8 (2013).
29. Graybill, W., Sood, A. K., Monk, B. J. & Coleman, R. L. State of the science: Emerging therapeutic strategies for targeting angiogenesis in ovarian cancer. *Gynecol. Oncol.* **138**, 223–226 (2015).
30. Franzese, E. *et al.* PARP inhibitors in ovarian cancer. *Cancer Treat. Rev.* **73**, 1–9 (2019).
31. Cui, Y. *et al.* OPCML is a broad tumor suppressor for multiple carcinomas and lymphomas with frequently epigenetic inactivation. *PLoS One* **3**, e2990 (2008).
32. McKie, A. B. *et al.* The OPCML tumor suppressor functions as a cell surface repressor-adaptor, negatively regulating receptor tyrosine kinases in epithelial ovarian cancer. *Cancer Discov.* **2**, 156–171 (2012).
33. Birtley, J. *et al.* Clinical inactivating mutations of the tumour suppressor OPCML reveal new cancer associated-functions. *Nat Comms, Press* (2019).
34. Jones, P. A. & Baylin, S. B. The Fundamental Role of Epigenetic Events in Cancer. *Nat. Rev. Genet.* **3**, 415–428 (2002).
35. Zhang, C. *et al.* Meta-analysis of DNA methylation biomarkers in hepatocellular carcinoma. *Oncotarget* **7**, 81255–81267 (2016).
36. Zhou, F. *et al.* Methylation of OPCML promoter in ovarian cancer tissues predicts poor patient survival. *Clin. Chem. Lab. Med.* **52**, 735–742 (2014).
37. Amornpisutt, R., Proungvitaya, S., Jearanaikoon, P. & Limpaboon, T. DNA methylation level of OPCML and SFRP1: a potential diagnostic biomarker of cholangiocarcinoma. *Tumor Biol.* **36**, 4973–4978 (2015).
38. Li, C. *et al.* OPCML is frequently methylated in human colorectal cancer and its restored expression reverses EMT via downregulation of smad signaling. *Am. J. Cancer Res.* **5**, 1635–1648 (2015).
39. Mei, F. C., Young, T. W., Liu, J. & Cheng, X. RAS-mediated epigenetic inactivation of OPCML in oncogenic transformation of human ovarian surface epithelial cells. *FASEB J.* **20**, 497–499 (2006).
40. Reed, J. E. *et al.* Expression of cellular adhesion molecule ‘OPCML’ is down-regulated in gliomas and other brain tumours. *Neuropathol. Appl. Neurobiol.* **33**, 77–85 (2007).
41. Wang, B. *et al.* Detection of OPCML methylation, a possible epigenetic marker, from free serum circulating DNA to improve the diagnosis of early-stage ovarian epithelial cancer. *Oncol. Lett.* **14**, 217–223 (2017).
42. Wu, Y. *et al.* Methylation profiling identified novel differentially methylated markers including OPCML and FLRT2 in prostate cancer. *Epigenetics* **11**, 247–258 (2016).
43. Xing, X. *et al.* Down-regulated expression of OPCML predicts an unfavorable prognosis and promotes disease progression in human gastric cancer. *BMC Cancer* **17**, 1–10 (2017).
44. Ye, F., Zhang, S.-F., Xie, X. & Lu, W.-G. OPCML gene promoter methylation and gene expression in tumor and stroma cells of invasive cervical carcinoma. *Cancer Invest.* **26**, 569–74 (2008).
45. Thiagalingam, S. *et al.* Loss of heterozygosity as a predictor to map tumor suppressor genes in cancer: Molecular basis of its occurrence. *Curr. Opin. Oncol.* **14**, 65–72 (2002).
46. Lyko, F. & Brown, R. DNA methyltransferase inhibitors and the development of epigenetic cancer therapies. *J. Natl. Cancer Inst.* **97**, 1498–1506 (2005).
47. Dhanak, D. & Jackson, P. Development and classes of epigenetic drugs for cancer. *Biochem. Biophys. Res. Commun.* **455**, 58–69 (2014).
48. Kinoshita, T. Glycosylphosphatidylinositol (GPI) Anchors: Biochemistry and Cell Biology: Introduction to a Thematic Review Series. *J. Lipid Res.* **57**, 4–5 (2015).
49. Lee, M. J. & Yaffe, M. B. Protein regulation in signal transduction. *Cold Spring Harb. Perspect. Biol.* **8**, 1–20 (2016).
50. Saha, S., Anilkumar, A. A. & Mayor, S. GPI-anchored protein organization and dynamics at the cell surface. *J. Lipid Res.* **2**, 159–175 (2015).
51. Fujita, M. & Kinoshita, T. GPI-anchor remodeling: Potential functions of GPI-anchors in intracellular trafficking and membrane dynamics. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1821**, 1050–1058 (2012).
52. Simons, K. & Ikonen, E. Functional rafts in cell membranes. *Nature* **387**, 569–572 (1997).
53. Zanini, E. *et al.* The tumor suppressor protein OPCML potentiates anti-EGFR and anti-HER2 targeted therapy in HER2-positive ovarian and breast cancer. *Mol. Cancer Ther.* **16**, 2246–2256 (2017).
54. Antony, J. *et al.* The tumour suppressor OPCML promotes AXL inactivation by the phosphatase PTPRG in ovarian cancer. *EMBO Rep.* e45670 (2018). doi:10.15252/embr.201745670
55. Nosjean, O., Briolay, A. & Roux, B. Mammalian GPI proteins : sorting , membrane residence and functions. *Biochem. Biophys. Acta* **1331**, 153–186 (1997).
56. Paulick, M. G., Wise, A. R., Forstner, M. B., Groves, J. T. & Bertozzi, C. R. Synthetic Analogues of Glycosylphosphatidylinositol-Anchored Proteins and Their Behavior in Supported Lipid Bilayers Molecular Modeling Studies of the GPI Anchor Analogues. *J. Am. Chem. Soc.* **129**, 11543–11550 (2007).
57. Chatterjee, S. & Mayor, S. The GPI-anchor and protein sorting. *Cell. Mol. Life Sci.* **58**, 1969–1987 (2001).
58. Brown, D. A. & Rose, J. K. Sorting of GPI-Anchored Proteins to Membrane Subdomains during Transport to the Apical Cell Surface. **66**, 533–544 (1992).
59. Heider, S., Dangerfield, J. A. & Metzner, C. Biomedical applications of glycosylphosphatidylinositol-anchored proteins. *J. Lipid Res.* **57**, 1778–1788 (2016).
60. Rietveld, A., Anson, I., Simons, K. & Bunting, J. Acyl and alkyl chain length of GPI-anchors is critical for raft association in vitro. *FEBS Lett.* **462**, 47–50 (1999).

Appendix D

61. Malherbe, T., Boschung, M., Bu, P. & Roditi, I. GPI-anchored proteins : now you see 'em , now you don ' t. *Faseb* **15**, 545–548 (2000).
62. Sharma, P. *et al.* Nanoscale Organization of Multiple GPI-Anchored Proteins in Living Cell Membranes. *Cell* **116**, 577–589 (2004).
63. Paulick, M. G., Forstner, M. B., Groves, J. T. & Bertozzi, C. R. A chemical approach to unraveling the biological function of the glycosylphosphatidylinositol anchor. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 20332–20337 (2007).
64. Brodsky, R. A. *et al.* Purified GPI-Anchored CD4DAF as a Receptor for HIV-Mediated Gene Transfer. **1239**, 1231–1239 (1994).
65. Muenchmeier, N. *et al.* A Novel CXCL10-Based GPI-Anchored Fusion Protein as Adjuvant in NK-Based Tumor Therapy. *PLoS One* **8**, 1–12 (2013).
66. McHugh, R. S. *et al.* Construction, purification, and functional incorporation on tumor cells of glycolipid-anchored human B7-1 (CD80). *Immunology* **92**, 8059–8063 (1995).
67. Nagarajan, S. & Selvaraj, P. Glycolipid-anchored IL-12 expressed on tumor cell surface induces antitumor immune response. *Cancer Res.* **62**, 2869–2874 (2002).
68. Djafarzadeh, R. *et al.* GPI-anchored TIMP-1 treatment renders renal cell carcinoma sensitive to FAS-mediated killing. *Oncogene* **25**, 1496–1508 (2006).
69. Musiol, H. & Dong, S. Toward Semisynthetic Lipoproteins by Convergent Strategies Based on Click and Ligation Chemistry. *ChemBioChem* **6**, 625–628 (2005).
70. Ball, H. L., King, D. S., Cohen, F. E., Prusiner, S. B. & Baldwin, M. A. Engineering the prion protein using chemical synthesis. *J. Pept. Res.* **58**, 357–374 (2001).
71. Grogan, M. J., Kaizuka, Y., Conrad, R. M., Groves, J. T. & Bertozzi, C. R. Synthesis of Lipidated Green Fluorescent Protein and Its Incorporation in Supported Lipid Bilayers. *J. Am. Chem. Soc.* **127**, 14383–14387 (2005).
72. Hicks, M. R. *et al.* Synthesis and structural characterization of a mimetic membrane-anchored prion protein. *FEDS J.* **273**, 1285–1299 (2006).
73. Hill, A. *et al.* Protection of erythrocytes from human complement-mediated lysis by membrane-targeted recombinant soluble CD59: A new approach to PNH therapy. *Blood* **107**, 2131–2137 (2006).
74. Teruya, K., Nishizawa, K. & Doh-ura, K. Semisynthesis of a Protein with Cholesterol at the C-Terminal , Targeted to the Cell Membrane of Live Cells. *Protein J* **29**, 493–500 (2010).
75. Schafer, B. *et al.* Preparation of Semisynthetic Lipoproteins with Fluorescent Cholesterol Anchor and Their Introduction to the Cell Membrane with Minimal Disruption of the Membrane. *Bioconjug. Chem.* **24**, 1684–1697 (2013).
76. Teruya, K., Nishizawa, K. & Doh-ura, K. Semisynthesis of a protein with cholesterol at the C-terminal, targeted to the cell membrane of live cells. *Protein J.* **29**, 493–500 (2010).
77. Sato, S. B. *et al.* Distribution and transport of cholesterol-rich membrane domains monitored by a membrane-impermeant fluorescent polyethylene glycol-derivatized cholesterol. *J. Biol. Chem.* **279**, 23790–23796 (2004).
78. Webb, M. E., Bon, R. S. & Wright, M. H. in *Chemical and Biological Synthesis: Enabling Approaches for Understanding Biology* 313–356 (2018).
79. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective 'ligation' of azides and terminal alkynes. *Angew. Chemie - Int. Ed.* **41**, 2596–2599 (2002).
80. Tornøe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-Triazoles by regioselective copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **67**, 3057–3064 (2002).
81. Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew. Chem. Int. Ed. Engl.* **40**, 2004–2021 (2001).
82. Huisgen, R. 1,3-Dipolar Cycloadditions Past and Future. *Angew. Chem. Int. Ed. Engl.* **2**, 565–598 (1963).
83. Agard, N. J., Prescher, J. A. & Bertozzi, C. R. A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **126**, 15046–15047 (2004).
84. Wang, L., Jiang, R., Wang, L., Liu, Y. & Sun, X. L. Preparation of chain-end clickable recombinant protein and its bio-orthogonal modification. *Bioorg. Chem.* **65**, 159–166 (2016).
85. Khatwani, S. L. *et al.* Covalent protein-oligonucleotide conjugates by copper-free click reaction. *Bioorganic Med. Chem.* **20**, 4532–4539 (2012).
86. Duckworth, B. P., Xu, J., Taton, T. A., Guo, A. & Distefano, M. D. Site-specific, covalent attachment of proteins to a solid surface. *Bioconjug. Chem.* **17**, 967–974 (2006).
87. Duckworth, B. P., Zhang, Z., Hosokawa, A. & Distefano, M. D. Selective labeling of proteins by using protein farnesyltransferase. *ChemBioChem* **8**, 98–105 (2007).
88. Evans, T. C., Benner, J. & XU, M.-Q. Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* **7**, 2256–2264 (1998).
89. Southworth, M. W., Amaya, K., Evans, T. C., Xu, M. Q. & Perler, F. B. Purification of proteins fused to either the amino or carboxy terminus of the Mycobacterium xenopi gyrase A intein. *Biotechniques* **27**, 110–120 (1999).
90. Muir, T. W., Sondhi, D. & Cole, P. A. Expressed protein ligation: A general method for protein engineering. *PNAS* **95**, 6705–6710 (1998).
91. Muir, T. W. Semisynthesis of Proteins by Expressed Protein Ligation. *Annu. Rev. Biochem.* **72**, 249–289 (2003).
92. Xu, M.-Q. & Evans, T. C. Intein-Mediated Ligation and Cyclization of Expressed Proteins. *Methods* **24**, 257–277 (2001).
93. Chong, S. *et al.* Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **192**, 271–281 (1997).
94. Chong, S. *et al.* Utilizing the C-terminal cleavage activity of a protein splicing element to purify recombinant proteins in a single chromatographic step. *Nucleic Acids Res.* **26**, 5109–5115 (1998).
95. Watanabe, T. *et al.* The roles of the C-terminal domain and type III domains of chitinase A1 from *Bacillus circulans* WL-12 in chitin degradation. *J. Bacteriol.* **176**, 4465–4472 (1994).
96. Liu, X. Protein-Splicing Intein : Genetic Mobility , Origin , and Evolution. *Biochemistry* **34**, 61–76 (2000).
97. Perler, F. B. & Allewell, N. M. Evolution, mechanisms, and applications of Intein-mediated protein splicing. *J. Biol.*

Appendix D

- Chem.* **289**, 14488–14489 (2014).
98. Shah, N. H. & Muir, T. W. Inteins: Nature's Gift to Protein Chemists. *Chem. Sci.* **5**, 446–461 (2014).
 99. Chong, S., Williams, K. S., Wotkowicz, C. & Xu, M. Modulation of Protein Splicing of the *Saccharomyces cerevisiae* Vacuolar Membrane ATPase Intein. *J. Biol. Chem.* **273**, 10567–10577 (1998).
 100. Rak, A. *et al.* Structure of Rab GDP-Dissociation Inhibitor in Complex with Prenylated YPT1 GTPase. *Science* (80-.). **302**, 646–651 (2003).
 101. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. H. Synthesis of Proteins by Native Chemical Ligation. *Science* (80-.). **266**, 776–779 (1994).
 102. Mao, H., Hart, S. a., Schink, A. & Pollok, B. a. Sortase-mediated protein ligation: a new method for protein engineering. *J. Am. Chem. Soc.* **126**, 2670–1 (2004).
 103. Ton-That, H., Liu, G., Mazmanian, S. K., Faull, K. F. & Schneewind, O. Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 12424–9 (1999).
 104. Ton-That, H., Mazmanian, S. K., Faull, K. F. & Schneewind, O. Anchoring of Surface Proteins to the Cell Wall of *Staphylococcus aureus*. *J. Biol. Chem.* **275**, 9876–9881 (2000).
 105. Guimaraes, C. P. *et al.* Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions. *Nat. Protoc.* **8**, 1787–1799 (2013).
 106. Giesbrecht, P., Kersten, T. & Maidhof, H. Staphylococcal Cell Wall: Morphogenesis and Fatal Variations in the Presence of Penicillin. *Microbiol. Mol. Biol. Rev.* **62**, 1371–1414 (1998).
 107. Huang, X. *et al.* Kinetic Mechanism of *Staphylococcus aureus* Sortase SrtA. *Biochemistry* **42**, 11307–11315 (2003).
 108. Hirakawa, H., Ishikawa, S. & Nagamune, T. Design of Ca²⁺-independent *Staphylococcus aureus* sortase A mutants. *Biotechnol. Bioeng.* **109**, 2955–2961 (2012).
 109. Strijbis, K., Spooner, E. & Ploegh, H. L. Protein Ligation in Living Cells Using Sortase. *Traffic* **13**, 780–789 (2012).
 110. Antos, J. M. *et al.* A straight path to circular proteins. *J. Biol. Chem.* **284**, 16028–16036 (2009).
 111. Ritzeveld, M. Sortagging: a robust and efficient chemoenzymatic ligation strategy. *Chemistry* **20**, 8516–8529 (2014).
 112. Williamson, D. J., Webb, M. E. & Turnbull, W. B. Depsipeptide substrates for sortase-mediated N-terminal protein ligation. *Nat. Protoc.* **9**, 253–62 (2014).
 113. Antos, J. M. *et al.* Site-Specific N- and C-Terminal Labeling of a Single Polypeptide Using Sortases of Different Specificity. *J. Am. Chem. Soc.* 10800–10801 (2009).
 114. Chan, L. *et al.* Covalent attachment of proteins to solid supports and surfaces via sortase-mediated ligation. *PLoS One* **2**, 1–5 (2007).
 115. Clow, F., Fraser, J. D. & Proft, T. Immobilization of proteins to biacore sensor chips using *Staphylococcus aureus* sortase A. *Biotechnol. Lett.* **30**, 1603–1607 (2008).
 116. Tanaka, T., Yamamoto, T., Tsukiji, S. & Nagamune, T. Site-specific protein modification on living cells catalyzed by sortase. *ChemBioChem* **9**, 802–807 (2008).
 117. Stefan, N. *et al.* Highly Potent, Anthracycline-based Antibody–Drug Conjugates Generated by Enzymatic, Site-specific Conjugation. *Mol. Cancer Ther.* **16**, 879–892 (2017).
 118. Popp, M. W., Antos, J. M., Grotenbreg, G. M., Spooner, E. & Ploegh, H. L. Sortagging: a versatile method for protein labeling. *Nat. Chem. Biol.* **3**, 707–708 (2007).
 119. Guo, X., Wang, Q., Swarts, B. M. & Guo, Z. Sortase-Catalyzed Peptide - Glycosylphosphatidylinositol Analogue Ligation. *J. Am. Chem. Soc.* 9878–9879 (2009).
 120. Morecroft, E. *MRes Research Manuscript: The Tumour Suppressor Protein OPCML as an Anticancer drug: A Biological Therapeutic enabled through Self-splicing MRes Research Manuscript The Tumour Suppressor Protein OPCML as an Anticancer Drug: A Biological Therapeutic enab.* (2015).
 121. Morecroft, E. V. N. *Early-Stage Assessment Report, The Tumour Suppressor Protein OPCML as an Anti-Cancer Drug: A Biological Therapeutic enabled through Self-Splicing GPI Mimetics.* (2017).
 122. Thomas, C. E., Ehrhardt, A. & Kay, M. a. Progress and problems with the use of viral vectors for gene therapy. *Nat. Rev. Genet.* **4**, 346–58 (2003).
 123. Mátrai, J., Chuah, M. K. L. & VandenDriessche, T. Recent advances in lentiviral vector development and applications. *Mol. Ther.* **18**, 477–490 (2010).
 124. Campeau, E. *et al.* A Versatile Viral System for Expression and Depletion of Proteins in Mammalian Cells. *PLoS One* **4**, e6529 (2009).
 125. Merten, O. W., Hebben, M. & Bovolenta, C. Production of lentiviral vectors. *Mol. Ther. - Methods Clin. Dev.* **3**, 16017 (2016).
 126. Robbins, P. D. & Ghivizzani, S. C. Viral Vectors for Gene Therapy. *Pharmacol. Ther.* **80**, 35–47 (1998).
 127. Liang, X., Peng, L., Baek, C. H. & Katzen, F. Single step BP/LR combined Gateway reactions. *Biotechniques* **55**, 265–268 (2013).
 128. Landy, A. Dynamic, Structural, and Regulatory Aspects Of Lambda Site-specific Recombination. *Annu. Rev. Biochem.* **58**, 913–949 (1989).
 129. Hartley, J. L., Temple, G. F. & Brasch, M. A. DNA cloning using in vitro site-specific recombination. *Genome Res.* **10**, 1788–1795 (2000).
 130. Bernard, P. & Couturier, M. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J. Mol. Biol.* **226**, 735–745 (1992).
 131. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Front. Microbiol.* **5**, 1–17 (2014).
 132. Sahdev, S., Khattar, S. K. & Saini, K. S. Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies. *Mol. Cell. Biochem.* **307**, 249–264 (2008).
 133. Ikonomou, L., Schneider, Y. J. & Agathos, S. N. Insect cell culture for industrial production of recombinant proteins. *Appl. Microbiol. Biotechnol.* **62**, 1–20 (2003).
 134. Demain, A. L. & Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* **27**, 297–306 (2009).
 135. Donahue, R. Rapid Generation of Recombinant Baculovirus and Expression of Foreign Genes using the Bac-To-Bac™

Appendix D

- Baculovirus Expression System. *Focus (Madison)*. **17**, 53–58 (1996).
136. Luckow, V. A., Lee, S. C., Barry, G. F. & Olins, P. O. Efficient Generation of Infectious Recombinant Baculoviruses by Site-Specific Transposon-Mediated Insertion of Foreign Genes into a Baculovirus Genome Propagated in *Escherichia coli*. *J. Virol.* **67**, 4566–4579 (1993).
137. Invitrogen. Bac-to-Bac® Baculovirus Expression System. 1–78 (2010).
138. Sutherland, R. M. *et al.* Oxygenation and Differentiation in Multicellular Colon Carcinoma Spheroids of Human Colon Carcinoma. *Cancer Res.* **46**, 5320–5329 (1986).
139. Hirschhaeuser, F. *et al.* Multicellular tumor spheroids: An underestimated tool is catching up again. *J. Biotechnol.* **148**, 3–15 (2010).
140. Vinci, M. *et al.* Advances in establishment and analysis of three-dimensional tumor spheroid-based functional assays for target validation and drug evaluation. *BMC Biol.* **10**, 29 (2012).
141. Berens, E. B., Holy, J. M., Riegel, A. T. & Wellstein, A. A Cancer Cell Spheroid Assay to Assess Invasion in a 3D Setting. *J. Vis. Exp.* 1–6 (2015). doi:10.3791/53409
142. Stein, A. M., Demuth, T., Mobley, D., Berens, M. & Sander, L. M. A mathematical model of glioblastoma tumor spheroid invasion in a three-dimensional in vitro experiment. *Biophys. J.* **92**, 356–365 (2007).
143. Sodek, K. L., Ringuette, M. J. & Brown, T. J. Compact spheroid formation by ovarian cancer cells is associated with contractile behavior and an invasive phenotype. *Int. J. Cancer* **124**, 2060–2070 (2009).
144. Kramer, N. *et al.* In vitro cell migration and invasion assays. *Mutat. Res.* **752**, 10–24 (2013).
145. Ayala, I., Baldassarre, M., Caldieri, G. & R. B. Invadopodia: A guided tour. *Eur. J. Cell Biol.* **85**, 159–164 (2006).
146. Styli, S. S., Kaye, A. H. & Lock, P. Invadopodia: At the cutting edge of tumour invasion. *J. Clin. Neurosci.* **15**, 725–737 (2008).
147. Morecroft, E. *Initial Research Plan - The Tumour Suppressor Protein OPCML as an Anti-Cancer Drug: A Biological Therapeutic enabled through Self-splicing Glycophosphatidylinositol Mimetics.* (2015).
148. Shi, J. *et al.* Engineered red blood cells as carriers for systemic delivery of a wide array of functional probes. *Proc. Natl. Acad. Sci.* **111**, 10131–10136 (2014).
149. Sarpong, K. & Bose, R. Efficient sortase-mediated N-terminal labeling of TEV protease cleaved recombinant proteins. *Anal. Biochem.* **521**, 55–58 (2017).
150. Antos, J. M., Truttmann, M. C. & Ploegh, H. L. Recent advances in sortase-catalyzed ligation methodology. *Curr. Opin. Struct. Biol.* **38**, 111–118 (2016).
151. Israelachvili, J. N. *Intermolecular and Surface Forces, Third Edition.*
152. Antos, J. M., Miller, G. M., Grotenbreg, G. M. & Ploegh, H. L. Lipid modification of proteins through sortase-catalyzed transpeptidation. *J. Am. Chem. Soc.* **130**, 16338–16343 (2008).
153. Spro, J. & Sinz, A. Monolithic columns with immobilized monomeric avidin: Preparation and application for affinity chromatography. *Anal. Bioanal. Chem.* **402**, 2395–2405 (2012).
154. Ntougkos, E. *et al.* The IgLON Family in Epithelial Ovarian Cancer: Expression Profiles and Clinicopathologic Correlates The IgLON Family in Epithelial Ovarian Cancer: Expression Profiles and Clinicopathologic Correlates. *Clin. cancer Res.* **11**, 5764–5768 (2005).
155. Lodge, A. P., Howard, M. R., McNamee, C. J. & Moss, D. J. Co-localisation, heterophilic interactions and regulated expression of IgLON family proteins in the chick nervous system. *Mol. Brain Res.* **82**, 84–94 (2000).
156. Teodoridis, J. M. *et al.* CpG Island Methylation of DNA Damage Response Genes in Advanced Ovarian Cancer Advanced Ovarian Cancer. *Cancer Res.* **65**, 8961–8967 (2005).
157. Jones, P. A. & Laird, P. W. Cancer epigenetics comes of age. *Nat. Genet.* **21**, 163–167 (1999).
158. Goffin, J. & Eisenhauer, E. DNA methyltransferase inhibitors - State of the art. *Ann. Oncol.* **13**, 1699–1716 (2002).
159. Rankin, E. B. *et al.* AXL is an essential factor and therapeutic target for metastatic ovarian cancer. *Cancer Res.* **70**, 7570–7579 (2010).
160. Linger, R., Keating, A., Shelton Earp, H. & Graham, D. TAM Receptor Tyrosine Kinases: Biologic Functions, Signaling, and Potential Therapeutic Targeting in Human Cancer. *Adv Cancer Res* **100**, 35–83 (2008).
161. Sasaki, T. *et al.* Structural basis for Gas6-Axl signalling. *EMBO J.* **25**, 80–87 (2006).
162. Gabra, H. Unpublished Research.
163. Paulick, M. G. & Bertozzi, C. R. The Glycosylphosphatidylinositol Anchor: A Complex Membrane-Anchoring. *Biochemistry* **47**, 6991–7000 (2008).
164. Baeschlin, D. K. *et al.* 1,2-Diacetals in Synthesis: Total Synthesis of a Glycosylphosphatidylinositol Anchor of *Trypanosoma brucei*. *Chem. - A Eur. J.* **6**, 172–186 (2000).
165. Kelley, L. A. & Sternberg, M. J. E. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* **4**, 363–371 (2009).
166. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
167. The Uniprot Consortium. UniProt: a hub for protein information. *Nucleic Acids Res.* **43**, D204–D212 (2014).
168. Somia, N. & Verma, I. M. Gene therapy: trials and tribulations. *Nat. Rev. Genet.* **1**, 91–99 (2000).
169. Indraccolo, S. *et al.* Gene Transfer in Ovarian Cancer Cells: A Comparison between Retroviral and Lentiviral Vectors 1. *Cancer Res.* **62**, 6099–6107 (2002).
170. Davis, H. E., Rosinski, M., Morgan, J. R. & Yarmush, M. L. Charged polymers modulate retrovirus transduction via membrane charge neutralization and virus aggregation. *Biophys. J.* **86**, 1234–1242 (2004).
171. Davis, H. E., Morgan, J. R. & Yarmush, M. L. Polybrene increases retrovirus gene transfer efficiency by enhancing receptor-independent virus adsorption on target cell membranes. *Biophys. Chem.* **97**, 159–172 (2002).
172. Wu, S. Y. & Sood, A. K. New roles opined for OPCML. *Cancer Discov.* **2**, 115–116 (2012).
173. Antos, J. M., Miller, G. M., Grotenbreg, G. M. & Ploegh, H. L. Lipid modification of proteins through sortase-catalyzed

Appendix D

transpeptidation. *J. Am. Chem. Soc.* **130**, 16338–16343 (2008).

Supplementary Information

Phyre2 Structure Predictions

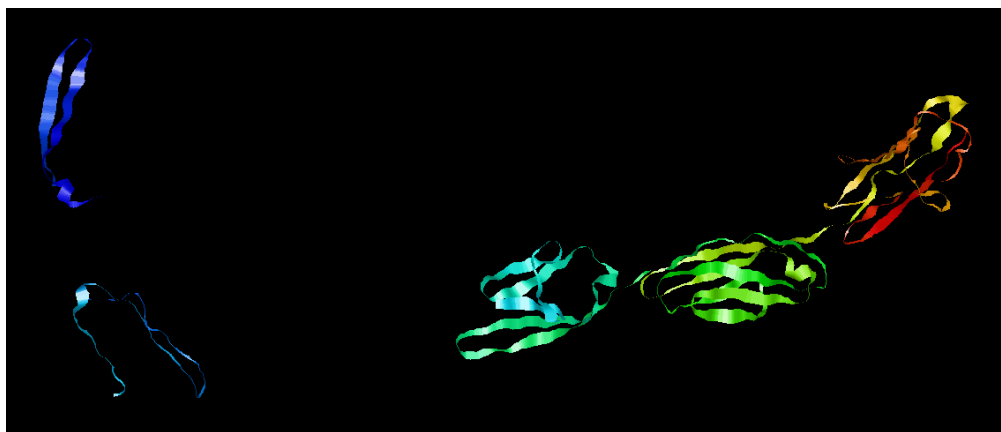


Figure S1: Full length OPCML structure prediction using Phyre2

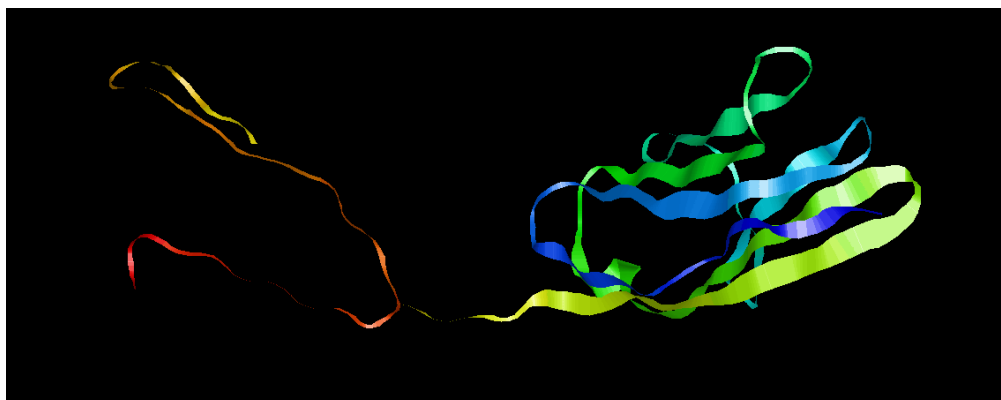


Figure S2: OPCML Domain 1 structure prediction using Phyre2

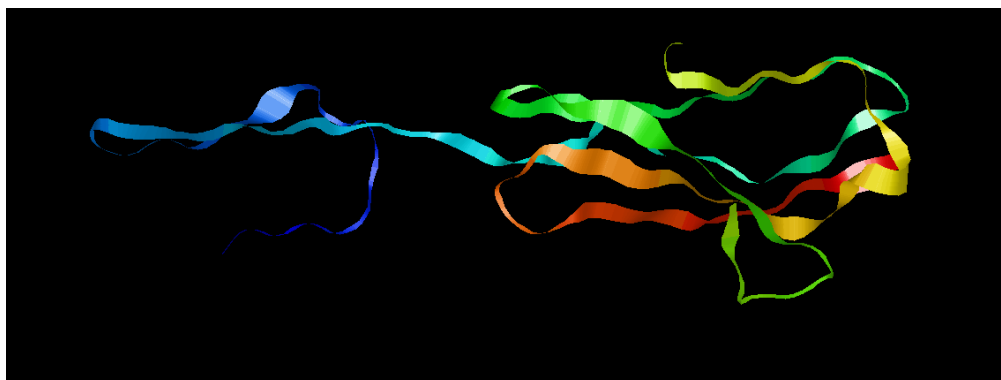


Figure S3: OPCML Domain 2 structure prediction using Phyre2

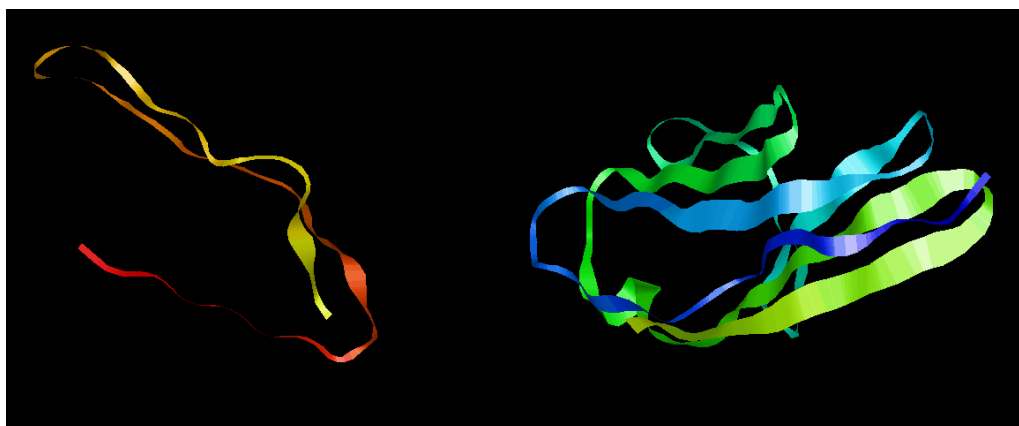


Figure S4: OPCML Domain 3 structure prediction using Phyre2

Primers 5' to 3'

Forward

GTCTGTGGGTACCTGTTTCCTG

Reverse

AAGAAGTGGGCTAAGAGGGTC

Full entry clone sequences

KEY
Propeptide
Signal peptide
Glycosylation sites
attL sites
Domain 1
Domain 2
Domain 3
HA tag
Sall restriction site
XbaI restriction site

OPCML Domain 1

GGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTTCGTTGCA
 ACAAATTGATAAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAG
 CAGGCTCCGTCGACGCCACCATGGGGGTCTGTGGGTACCTGTTTCCTGCCC
 TGGAAGTGCCTCGTGGTCTGTCTCTCAGGCTGCTGTTTCCTTGTACCCAC
 AGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCCCAAAGCTATGGACA
 ACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGTACCATAGAT
 GACCGGGTAACCCGGGTGGCCTGGCTAACCGCAGCACCATCCTCTACGC
 TGGGAATGACAAGTGGTCCATAGACCCTCGTGTGATCATCCTGGTCAATA
 CACCAACCCAGTACAGCATCATGATCCAAATGTGGATGTGTATGACGAA
 GGTCCGTACACCTGCTCTGTGCAGACAGACAATCATCCCAAACGTCCCG
 GGTTACCTAATAGTGCAAGTTCCTTTGTATTACCATACGATGTTCCAG
 ATTACGCTGGGCCTGGAGCAGTCATTGATGGTGTAAACTTCGGCCTCCAGA
 GCACTGGCTTGTCTCTGGCTATCAGGGACCCTCTTAGCCCACTTCTTCAT
 CAAGTTTTGATCTAGAGACCCAGCTTTCTTGTACAAAGTTGGCATTATAA
 GAAAGCATTGCTTATCAATTTGTTGCAACGAACAGGTCACCTATCAGTCAA
 AATAAAATCATTATTTG

OPCML Domain 2

GGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTTCGTTGCA
 ACAAATTGATAAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAG
 CAGGCTCCGTCGACGCCACCATGGGGGTCTGTGGGTACCTGTTTCCTGCCC
 TGGAAGTGCCTCGTGGTCTGTCTCTCAGGCTGCTGTTTCCTTGTACCCAC
 AGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCCGGGTTCACCTAATAG
 TGCAAGTTCCTCCTCAGATCATGAATATCTCCTCAGACATCACTGTGAAT
 GAGGGAAGCAGTGTGACCCTGCTGTGTCTTGCTATTGGCAGACCAGAGCC
 AACTGTGACATGGAGACACCTGTCAGTCAAGGAAGGCCAGGGCTTTGTAA
 GTGAGGATGAGTACCTGGAGATCTCTGACATCAAGCGAGACCAGTCCGGG
 GAGTACGAATGCAGCGCGTTGAACGATGTCGCTGCGCCCGATGTGCGGAA
 AGTAAAAATCACTGTAACTATTTGTATTACCATACGATGTTCCAGATT
 ACGCTGGGCCTGGAGCAGTCATTGATGGTGTAAACTTCGGCCTCCAGAGCA
 CTGGCTTGTCTCTGGCTATCAGGGACCCTCTTAGCCCACTTCTTCATCAA
 GTTTTGA TCTAGAGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAA
 AGCATTGCTTATCAATTTGTTGCAACGAACAGGTCACCTATCAGTCAAAAT
 AAAATCATTATTTG

OPCML Domain 3

GGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTTCGTTGCA
 ACAAATTGATAAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAG
 CAGGCTCCGTCGACGCCACCATGGGGGTCTGTGGGTACCTGTTTCCTGCCC
 TGGAAGTGCCTCGTGGTCTGTCTCTCAGGCTGCTGTTTCCTTGTACCCAC

Appendix D

AGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCGTAAACTATCCTCCCT
ATATCTCAAAAGCCAAGAACACTGGTGTTCAGTCGGTCAGAAGGGCATC
CTGAGCTGTGAAGCCTCTGCAGTCCCCATGGCTGAATTCCAGTGGTTCAA
GGAAGAAACCAGGTTAGCCACTGGTCTGGATGGAATGAGGATTGAAAACA
AAGGCCGCATGTCCACTCTGACTTTCTTC AATGTTTCAGAAAAGGATTAT
GGG AAC TATACTTGTGTGGCCACGAACAAGCTTGGGAACACC AATGCCAG
CATCACATTGTATTACCCATACGATGTTCCAGATTACGCTGGGCCTGGAG
CAGTCATTGATGGTGTA AAC TCGGCCTCCAGAGCACTGGCTTGTCTCTGG
CTATCAGGGACCCTCTTAGCCCACTTCTTCATCAAGTTTGTCTAGAGA
CCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAA
TTTGTGCAACGAACAGGTCAGTCAAAATAAAATCATTATTTG

OPCML Full length

GGGCCCCAAATAATGATTTTATTTTGACTGATAGTGACCTGTTTCGTTGCA
ACAAATTGATAAGCAATGCTTTTTTATAATGCCAACTTTGTACAAAAAAG
CAGGCTCC GTCGAC GCCACC ATGGGGGTCTGTGGGTACCTGTTCCCTGCCC
TGGAAGTGCCCTCGTGGTCTGTCTCTCAGGCTGCTGTTCCCTTGTACCCAC
AGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTC CCCAAAGCTATGGACA
ACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGTACCATAGAT
GACCGGGTAACCCGGGTGGCTAAACCGCAGCACCATCCTCTACGC
TGGGAATGACAAGTGGTCCATAGACCCTCGTGTGATCATCCTGGTCAATA
CACCAACCCAGTACAGCATCATGATCCAAAATGTGGATGTGTATGACGAA
GGTCCGTACACCTGCTCTGTGCAGACAGACAATCATCCCAAACGTCCCG
GGTTCACCTAATAGTGCAAGTTCCT CCTCAGATCATG AATATCTCCTCAG
ACATCACTGTGAATGAGGGAAGCAGTGTGACCCTGCTGTGTCTTGCTATT
GGCAGACCAGAGCCAACTGTGACATGGAGACACCTGTCAGTCAAGGAAGG
CCAGGGCTTTGTAAGTGAAGATGAGTACCTGGAGATCTCTGACATCAAGC
GAGACCAGTCCGGGGAGTACGAATGCAGCGCGTTGAACGATGTCGCTGCG
CCCGATGTGCGGAAAGTAAAAATCACTGTAAACTATCCTCCCTATATCTC
AAAAGCCAAGAACACTGGTGTTCAGTCGGTCAGAAGGGCATCCTGAGCT
GTGAAGCCTCTGCAGTCCCCATGGCTGAATTCCAGTGGTTCAAGGAAGAA
ACCAGGTTAGCCACTGGTCTGGATGGAATGAGGATTGAAAACAAAGGCCG
CATGTCCACTCTGACTTTCTTC AATGTTTCAGAAAAGGATTATGGG AAC
ATACTTGTGTGGCCACGAACAAGCTTGGGAACACC AATGCCAGCATCACA
TTGTATTACCCATACGATGTTCCAGATTACGCTGGGCCTGGAGCAGTCAT
TGATGGTGTA AAC TCGGCCTCCAGAGCACTGGCTTGTCTCTGGCTATCAG
GGACCCTCTTAGCCCACTTCTTCATCAAGTTTGTCTAGAGACCCAGCT
TTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTG
CAACGAACAGGTCAGTCAAAATAAAATCATTATTTG

