

Enhanced Antibody and Glycomics-Based Approaches for the Detection of Prostate Cancer

Sarah Gilgunn, M.Sc.

This thesis is submitted to Dublin City University for the degree of
Ph.D.

December 2015

Based on research carried out at
School of Biotechnology,
Dublin City University,
Dublin 9, Ireland.

Supervisor: Professor Richard O’Kennedy
External Supervisor: Professor Pauline M. Rudd,
NIBRT.

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed:

ID No.:

Date:

Acknowledgements

There are a great deal of people whom I need to thank but first and foremost, I would like to express my sincere gratitude to my supervisor Professor Richard O’Kennedy for the continuous support throughout my PhD, for his patience, motivation, enthusiasm, and immense knowledge. Richards’s guidance has helped me greatly through the highs and lows of this research. In addition I would also like to thank my co-supervisor, Professor Pauline Rudd, for her kind words of advice, help and support throughout this body of work.

I also wish to acknowledge the support, guidance, encouragement and direction provided by Dr. Paul Conroy. Paul is always there for me, giving me advice and ideas to accomplish this research. Paul has not only become a scientist I greatly admire and respect but a very dear friend. Dr. Radka Saldoval-Fahey, I owe you immense gratitude for your patience in teaching me the world of glycobiology!! Thank you for some great memories in the GlycoScience Lab in NIBRT and for guiding me so thoughtfully through my PhD.

I want to give a huge shout out to all members past and present of the Applied Biochemistry group in DCU. Thank you all for the great advice given and more importantly so many great nights out (particularly all those cocktail nights in capitol!!). I have made so many friends that I will never forget.

Finally, I would like to say a special word of thanks to thank my parents Marion and Seamus, sister, Siobhan, bother, Stevie, extended family and friends and Richy, for whom this thesis is dedicated, for giving constant support and encouragement to pursue my love of science. None of this would be possible without all of your help. Thank you.

Table of Contents

Declaration.....	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Abbreviations	xii
Units	xvii
List of Figures	xviii
List of Tables	xxiii
Publications, Presentations, Awards, Invention Disclosure and Courses Attended	xxv
Abstract.....	xxvii
Chapter One : Introduction	1
1.1 Thesis outline	2
1.2 Prostate cancer (PCa).....	2
1.3 Prostate cancer diagnosis	4
1.4 Current approaches in PCa detection	7
1.4.1 Biology of PSA	7
1.4.2 Improving PSA specificity for PCa diagnosis	9
1.4.2.1 Exploring the ratio of free PSA (fPSA) to total PSA (tPSA)	10
1.4.2.2 proPSA – The cancer-specific isoform.....	10
1.4.2.3 BPSA- specific for BPH diagnosis.....	12
1.4.2.4 Ratio of cPSA to fPSA	12
1.5. Glycomics and PCa	13
1.5.1 An introduction to glycans	14
1.5.2 Recent advances in the analysis of glycoproteins.....	16
1.5.3 Altered glycosylation patterns in PCa	20
1.6 Introduction to the immune system.....	23

1.6.1 Antibody structure	25
1.6.2 Recombinant antibodies.....	28
1.7 Phage display technology.....	31
1.7.1 Filamentous bacteriophage biology.....	33
1.8 Antibody characterisation by surface plasmon resonance.....	39
1.9 Antibody characterisation by X-ray crystallography	43
1.9.1 Crystal structure determination X-ray diffraction	45
1.9.2 Data collection:	47
1.9.3 Data analysis	48
1.9.3.1 The phase problem	49
1.9.3.2 Assessing the quality of the diffracted data	50
1.9.4 Crystal structure of antibodies.....	51
1.10 Thesis aims	53
Chapter Two: Materials and Methods.....	54
2.1 Materials	55
2.1.1 Equipment list	55
2.1.2 Reagents.....	56
2.1.3 Commercial Kits	58
2.1.4 Bacterial Strains	58
2.1.5 Media Compositions	59
2.1.6 Media Additives	59
2.1.6.1 Antibiotics	59
2.1.6.2 Additives.....	60
2.1.7 Buffers.....	60
2.1.8 Buffers for sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)	60
2.1.9 Protein Purification Buffers.....	62

2.2 Methods	63
2.2.1 General molecular methods	63
2.2.1.1 Agarose gel electrophoresis.....	63
2.2.1.2 Ethanol precipitation of DNA	63
2.2.1.3 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)	64
2.2.1.4 Western Blotting	66
2.2.1.5 Preparation of bacterial cell stocks.....	68
2.2.1.6 Buffer exchange and protein concentration.....	68
2.2.2 Screening and analysis of an existing avian anti-fPSA library	69
2.2.2.1 Identification of anti-fPSA specific antibody fragments by monoclonal ELISA and Biacore 4000	70
2.2.2.2 Optimisation of expression of scFv B8	72
2.2.2.3 Purification of scFv B8 by Immobilised Metal Affinity Chromatography (IMAC).....	73
2.2.2.4 Anti-fPSA scFv B8 Sandwich ELISA Limit of Detection (LOD) Assay	76
2.2.2.5 Kinetic analysis of scFv B8 using Biacore 3000	77
2.2.2.5.1 Instrument /Chip preparation.....	77
2.2.2.5.2 Kinetics analysis of anti-PSA scFv B8 using Biacore 3000	78
2.2.2.6 Crystal Structure of scFv B8	79
2.2.2.7 Site-Directed Mutagenesis of scFv B8.....	80
2.2.2.8 Expression and Purification of scFvB8 Ser MT by Immobilised Metal Affinity Chromatography (IMAC) and Gel Filtration.....	83
2.2.2.9 Determination of the crystal structure of scFv B8 Ser MT.....	85
2.2.2.9.1 Coarse screening for optimal crystal formation conditions.....	85
2.2.2.9.2 Fine-Screening to refine optimal crystal formation conditions	86
2.2.2.9.3 X-ray diffraction of crystals of scFv B8 Ser MT.....	90
2.2.2.10 Kinetic analysis of scFv B8 Ser MT using Biacore 3000	90

2.2.2.11 Expression and Purification of scFvB8 Ala MT by Immobilised Metal Affinity Chromatraphy (IMAC) and Gel Filtration.....	91
2.2.2.12 Crystal Structure of scFv B8 Ala MT	91
2.2.2.13 Mammalian Expression of recombinant PSA.....	92
2.2.2.14 Purification of recombinant PSA by Immobilised Metal Affinity Chromatraphy (IMAC) and Gel Filtration.....	92
2.2.2.15 Formation of recombinant PSA-scFvB8 complex.....	93
2.2.2.16 Determination of the crystal structure of recombinant PSA-scFvB8 complex by X-ray crystallography	93
2.2.2.16.1 Coarse screening for optimal crystal formation conditions.....	94
2.2.2.16.2 Fine screen to refine for optimal crystal formation conditions.....	95
2.2.2.16.3 X-ray diffraction of crystals of recombinant PSA-scFvB8 complex	95
2.2.3 Generation of an avian scFv antibody library	95
2.2.3.1 Immunisation of female Leghorn chickens	95
2.2.3.2 Determination of avian serum antibody titres for immune response to cPSA.....	96
2.2.3.3 Isolation and quantification of total cellular RNA from spleen and bone marrow.....	97
2.2.3.4 Reverse transcription of total RNA to cDNA	98
2.2.3.5 ScFv library construction.....	99
2.2.3.6 PCR amplification and optimisation of the heavy and light-chain variable (V_H and V_L) coding regions	100
2.2.3.7 Generation of scFv-fragment by splice-by-overlap extension (SOE) PCR.....	102
2.2.3.8 Preparation of pComb3xSS vector	103
2.2.3.9 Restriction digestion of the SOE-PCR and phagemid cloning vector pComb3xSS	104
2.2.3.10 Library ligation and transformation	105
2.2.3.11 Selection of cPSA-specific phage particles by panning against immobilised cPSA on an ELISA plate	107
2.2.3.11.1 Phage rescue after each round of panning.....	107

2.2.3.11.2 Library panning on an cPSA immobilised surface	108
2.2.3.12 Polyclonal phage ELISA	110
2.2.3.13 Soluble expression of scFv fragments and subsequent monoclonal ELISA analysis...	111
2.2.3.14 Small-scale expression and purification of anti-cPSA 2D2SG by Immobilised Metal Affinity Chromatraphy (IMAC)	112
2.2.3.15 Examaniation of anti-cPSA 2D2SG cross reactivity	114
2.2.4 <i>N</i> -glycan profiling of avian polyclonal IgY antibodies	114
2.2.4.1 Purification polyclonal IgY antibodies from serum	115
2.2.4.2 Glycan release	115
2.2.4.3 Labelling of released glycans with 2-aminobenzamide (2AB)	117
2.2.4.4 Ultra Performance Liquid Chromatography (UPLC) for the separation of the 2-AB labelled glycans.....	118
2.2.4.5 Weak Anion-Exchange High-Performance Liquid Chromatography (WAX HPLC) for the determination of sialic acid content of IgY.....	119
2.2.4.6 Exoglycosidase digestion for the determination of monosaccharide sequence and linkage of the <i>N</i> -glycans present in the IgY	120
2.2.4.7 Ultra Performance Liquid Chromatography-Fluorescence-Mass Spectrometry (UPLC-FLR-MS)	121
2.2.4.8 Molecular Modelling of IgY.....	122
2.2.5 Differentiation of indolent, significant and aggressive prostate cancer by robotic high-throughput <i>N</i> -glycan profiling.....	122
2.2.5.1 Serum Samples.....	124
2.2.5.2 Glycoprotein denaturation and glycan release.....	125
2.2.5.3 Hydrazide-mediated glycan ‘clean-up’	126
2.2.5.4 Glycan solid phase extraction	127
2.2.5.5 Ultra Performance Liquid Chromatography (UPLC) of the 2-AB labelled glycans	128
2.2.5.6 Statistical analysis	128

Chapter Three: Screening and analysis of an avian anti-fPSA library	129
3.1 Introduction	130
3.2 Results	134
3.2.1 Identification of high affinity anti-fPSA scFv	134
3.2.2 Optimisation of expression of scFv B8	138
3.2.3 Purification of scFv B8 by Immobilised Metal Affinity Chromatography (IMAC)	141
3.2.4 Anti-fPSA scFv B8 Sandwich ELISA Limit of Detection (LOD) Assay	142
3.2.5 Anti-fPSA scFv B8 cross reactivity assay	146
3.2.6 Kinetic analysis of scFv B8 using Biacore 3000	147
3.2.7 Crystal Structure of scFv B8	151
3.2.8 Expression and purification of scFvB8 Ser MT by Immobilised Metal Affinity Chromatography (IMAC) and Gel Filtration (GF)	153
3.2.9 Determination of the crystal structure of anti fPSA scFv B8 Ser MT by X-ray crystallography	154
3.2.10 Kinetic analysis of scFv B8 Ser MT using Biacore 3000	156
3.2.11 Expression and Purification of scFvB8 Ala MT by Immobilised Metal Affinity Chromatography (IMAC) and Gel Filtration	160
3.2.12 Determination of the crystal structure of anti fPSA scFv B8 Ala MT by X-ray crystallography	160
3.2.13 Mammalian Expression of recombinant PSA	163
3.2.14 Purification of recombinant PSA by Immobilised Metal Affinity Chromatography (IMAC) and Gel Filtration	166
3.2.15 Formation of recombinant PSA-scFv B8 complex	167
3.2.16 Determination of the crystal structure of rPSA-anti-fPSA scFv B8 complex	168
3.3 Discussion	170
Chapter Four: Generation of recombinant anti-PSA antibody fragments and N-glycan characterisation of polyclonal antibodies from an avian immune system	180

4.1 Introduction	181
4.2 Results	187
4.2.1 Immunisation of chickens with cPSA antigen	187
4.2.2 Optimisation and amplification of chicken heavy and light chain genes.....	189
4.2.3 SOE-PCR of variable heavy and light chains.....	191
4.2.4 Construction of scFv library in the pComb3xSS vector.....	193
4.2.5 Polyclonal phage ELISA.....	196
4.2.6 Soluble monoclonal ELISA.....	198
4.2.7 Expression and purification of anti-cPSA scFv 2D2SG.....	200
4.2.8 Cross-reactivity assay.....	201
4.2.9.1 IgY Purification	202
4.2.9 N-glycan profiling of avian polyclonal IgY antibodies	204
4.2.9.2 IgY N-glycan profiling	204
4.2.9.2 Molecular Modelling of IgY.....	209
4.3 Discussion.....	209
Chapter Five : Differentiation of indolent, significant and aggressive prostate cancer by robotic high-throughput N-glycan profiling.....	220
5.1 Introduction	221
5.2 Results	224
5.2.1 PCRC patient cohort.....	224
5.2.2 UPLC-HILIC profiling of PCa patients serum.....	225
5.2.2 Statistical analysis	229
5.2.3 Significant changes in whole serum N-glycome in prostate cancer	230
5.2.4 N-glycome biomarker panel.....	233
5.3 Discussion.....	238
Chapter Six: Overall Conclusions.....	247

6.1 Overall Conclusions.....	248
<i>Appendices</i>	252
Appendix One.....	253
Appendix Two	254
Appendix Three	258
Appendix Four	254
Appendix Five.....	271
Bibliography	277

List of Abbreviations

AAL	<i>Aleuria aurentia</i> lectin
Ab	Antibody
Abs	Absorbance
ABS	<i>Arthrobacter ureafaciens</i> sialidase
ACT	Alpha 1-anti-chymotrypsin
Ag	Antigen
AIC	Akaike information criterion
ANOVA	Analysis of variance
APS	Ammonium persulfate
API	Alpha 1-protease inhibitor
AUC	Area under the curve
A2M	Alpha 2-macroglobulin
2-AA	2-anthranilic acid
2AB	2-aminobenzamide
BFK	Bovine kidney alpha-fucosidase
BM	Bone marrow
BPH	Benign prostate hyperplasia
BPSA	Benign prostate specific antigen
BSA	Bovine serum albumin
BTG	Bovine testes beta-galactosidase
cDNA	Complementary deoxyribonucleic acid
cPSA	Complex prostate specific antigen
CDR	Complementary determining region
C_H	Constant heavy chain
C_L	Constant light chain
CM	Carboxymethylated dextran
Cfu	Colony forming units
CV	Column volume
CZE	Capillary zone electrophoresis
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate

DTT	Dithiothreitol
DRE	Digital rectal examination
<i>E. coli</i>	<i>Escherichia coli</i>
EDC	N-ethyl-N'-(dimethylamioethyl) – carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ELLA	Enzyme-linked lectin assay
EGFR	Epidermal growth factor receptor
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmic reticulum
Fab	Antibody binding fragment
Fc	Fragment crystallisable
FCA	Freund's complete adjuvant
FDA	Food and drug administration
FICA	Freund's incomplete adjuvant
FLR-MS	Fluorescence mass spectrometry
fPSA	Free prostate specific antigen
FWR	Framework region
Fv	Fragment variable
GF	Gel filtration
GPI	Glycosylphosphatidylinositol
GNA	<i>Galanthus nivalis</i>
GU	Glucose unit
GUH	<i>N</i> -acetylglucosaminidase
GWAS	Genome-wide association study
HA	Haemagglutinin
HILIC	Hydrophilic interaction liquid chromatography
HPAEC	High-performance anion-exchange chromatography
HPLC	High performance liquid chromatography
HRP	Horse radish peroxidase
HVL	Hypervariable loops
Ig	Immunoglobulin
IMAC	Immobilised metal affinity chromatography

IMGT	Immunogenetic database
IPA	Isopropanol alcohol
iPSA	Intact prostate specific antigen
IPTG	Isopropyl- β -D-galactopyranoside
JBM	Jack bean α -mannosidase
LB	Luria Broth
LoB	Limit of blank
LoD	Limit of detection
MAA	<i>Maackia amurensis</i> agglutinin
MAD	Multi-wavelength anomalous dispersion
MALDI-TOF-MS	Matrix-assisted laser desorption/ionization time of flight mass spectrometry
MgCl₂	Magnesium chloride
MG H₂O	Molecular grade water
MHC	Major histocompatibility complex
MIR	Multiple isomorphous replacement
MOPS	3-(<i>N</i> -morpholino)propanesulfonic acid
MR	Molecular replacement
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MT	Mutant
MWCO	Molecular weight ‘cut-off’
NAN1	<i>Streptococcus pneumoniae</i> sialidase
NHS	<i>N</i> -hydroxysuccinimide
NPHPLC	Normal phase high performance liquid chromatography
NTA	Nitrilotriacetic acid
o/n	Overnight
OSB	Osmotic shock buffer
pAb	Polyclonal antibody
PBS	Phosphate buffered saline
PBSM	Phosphate buffered saline with milk marvel (5%, w/v)
PBST	Phosphate buffered saline with tween (0.05%, v/v)
PBSTM	Phosphate buffered saline with tween milk marvel (1%, w/v)

PCa	Prostate cancer
PCR	Polymerase chain reaction
PCRC	Prostate Cancer Research Consortium
PDB	Protein data bank
PEG	Polyethylene glycol
PEI	Polyethylenimine
pH	Log of the hydrogen ion concentration in solution
<i>Phi</i>	Prostate health index
PK	Pharmacokinetic
PMT	Photomultiplier
PNGase F	Peptide - <i>N</i> -Glycosidase F
PNI	Perineural invasion
PSA	Prostate specific antigen
PTM	Post translational modification
PVA	Polyvinyl alcohol
PVDF	Polyvinylidene fluoride
rAb	Recombinant antibody
R_{max}	Maximal binding response
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
RP	Radical prostatectomy
RP-LC	Reverse phase liquid chromatography
Rpm	Revolutions per minute
RT	Room temperature
RU	Response units
SB	Super broth
scAb	Single chain antibody
scFv	Single chain variable fragment
SDM	Site directed mutagenesis
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
Sia	Sialic acid

SMEG	Semenogelin
SOC	Super optimal catabolite
SOE	Splice by overlap extension
SP	Spleen
SPE	Solid phase extraction
SPFS	Surface plasmon field-enhanced fluorescence spectroscopy
SPG	<i>Streptococcus pneumoniae</i> β -galactosidase
SPR	Surface plasmon resonance
ssDNA	Single stranded deoxyribonucleic acid
TAE	Tris-acetate Ethylenediaminetetraacetic acid
TEMED	Tetramethylethylenediamine
TMB	3, 3', 5, 5'- tetramethylbenzidine
tPSA	Total prostate specific antigen
TNM	Tumour nodes metastasis
UPLC	Ultra performance liquid chromatography
USPSTF	United States preventive services task force
UV	Ultra Violet
V_H	Variable heavy chain
V_L	Variable light chain
WAX	Weak anion exchange
WB	Western blot
WT	Wild type

Units

%	Percentage
χ^2	Chi ² or “goodness of fit”
°C	Degrees Celsius
Å	Angstrom
ASU	Asymmetric unit
AU	Absorbance units
Bp	Base pair
Da	Dalton
G	Grams
GU	Glucose unit
L	Litre
M	Molar
Mg	Milligram
mM	Millimolar
mL	Millilitre
nM	Nanomolar
pM	Picomolar
fM	Femtomolar
Rpm	Revolutions per minute
RU	Response units
v/v	Volume per unit volume
w/v	Weight per unit volume
µg	Microgram
µL	Microlitre
µM	Micromolar
x'g'	relative centrifugal force (rcf)

List of Figures

Chapter One: Introduction

Figure 1.1: The prostate gland.....	4
Figure 1.2: Biosynthesis of PSA.....	9
Figure 1.3: Symbolic diagrams to depict N-linked monosaccharides and linkages of the three general N-linked glycan structures.....	16
Figure 1.4: Methodology for glycan release, fluorescent labeling, NPHPLC profiling, structure assignment and quantitation.....	19
Figure 1.5: (A) Basic structure of the immunoglobulin G (IgG) molecule and (B) Ribbon diagram of IgG, illustrating the Fv, Fab and Fc regions.....	27
Figure 1.6: Recombinant antibody fragment constructs.....	30
Figure 1.7: Selection of antibodies by phage display (panning).....	33
Figure 1.8: Filamentous bacteriophage.....	34
Figure 1.9: Model for the early events in the phage infection of <i>E. coli</i>	37
Figure 1.10: Schematic representation of a Surface Plasmon Resonance (SPR)-based system.....	40
Figure 1.11: Binding interaction for analytes A and B forming the complex AB.....	42
Figure 1.12: History of X-ray crystallography in Nobel prizes.....	44
Figure 1.13: Overview of protein crystallography by X-ray diffraction.....	47
<i>Chapter Two: Materials and Methods</i>	
Figure 2.1: Semi-Dry transfer system.....	67

Figure 3.15: Kinetic interaction analyses of fPSA and scFv B8 Ser MT.....	159
Figure 3.16: Crystal structures of wild type and mutant variants of anti-fPSA scFv B8.....	162
Figure 3.17: Vector map of pSecTag2/PSA.....	164
Figure 3.18: Recombinant PSA expression.....	165
Figure 3.19: Purification of recombinant PSA (rPSA) by (A) IMAC and (B) Gel Filtration.....	166
Figure 3.20: Formation of rPSA-scFv B8 complex.....	167
Figure 3.21: Crystal structure of anti-fPSA scFvB8-rPSA complex.....	169
Figure 3.23: Diagrammatic representation of the surface charge of wild type and mutant variants of anti-fPSA scFv B8.....	176
Figure 3.24: Anti-fPSA scFv B8-recombinat PSA binding contacts.....	178
 <i>Chapter Four: Generation of recombinant anti-PSA antibody fragments and N-glycan characterisation of polyclonal antibodies from an avian immune system</i>	
Figure 4.1: The process of gene conversion in chicken is illustrated for both the heavy and light chain.....	184
Figure 4.2: Structures of Immunoglobulin G and Y and their N-glycosylation sites.....	186
Figure 4.3: cPSA immunised chicken serum titrations.....	188
Figure 4.4: MgCl ₂ optimisation for spleen (SP) and bone marrow (BM) for both for the (A) V _L and (B) V _H amplifications.....	190
Figure 4.5: Optimised large-scale amplifications of (A) V _L for BM and SP and (B) V _H for BM and SP of cPSA immunised chicken.....	191
Figure 4.6: MgSO ₄ optimisation for SOE-PCR.....	192
Figure 4.7: Large-scale amplification of (A) SP SOE and (B) BM SOE.....	192

Figure 4.8: Schematic representation of the pComb3xSS vector.....	194
Figure 4.9: Restriction digests of (A) pComb3xSS vector and (B) scFv inserts.....	195
Figure 4.10: Polyclonal phage ELISA.....	197
Figure 4.11: Soluble monoclonal ELISA of scFv clones from each round of panning.....	199
Figure 4.12: IMAC purification of anti-cPSA 2D2SG.....	201
Figure 4.13: Cross reactive indirect ELISA for scFv 2D2SG.....	202
Figure 4.14: SDS-PAGE and Western Blot analysis of IgY purified from serum.....	203
Figure 4.15: HILIC-UPLC profile of undigested N-glycans from serum IgY.....	205
Figure 4.16: Unfractionated IgY profile was subjected to exoglycosidase digestions.....	206
Figure 4.17: IgY glycans separated according sialic acid content by WAX-HPLC.....	207
Figure 4.18: WAX fraction's separated by HILIC-UPLC.....	208
Figure 4.19: Molecular model of glycosylated IgY.....	210
Figure 4.20: Sequence alignment of anti-cPSA 2D2SG and 2D8SG.....	212
 <i>Chapter Five: Differentiation of indolent, significant and aggressive prostate cancer by robotic high-throughput N-glycan profiling</i>	
Figure 5.1: Robotic automated sample preparation workflow for UPLC-based glycan analysis on a liquid handling platform.....	223
Figure 5.2: Prostate cancer patients sera sourced from the Prostate Cancer Research Consortium (PCRC) BioResource.....	224
Figure 5.3: Typical HILIC-UPLC profile of undigested serum N-glycome from prostate cancer patients.....	226

Figure 5.4: Boxplots of 4 glycosylation peaks which exhibited significant differences across 3 patient groups (indolent, significant and aggressive) from the Irish PCRC cohort.....231

Figure 5.5: ROC curves of the 4 significantly changed peaks as prostate cancer progresses....232

Figure 5.6: Plotted peaks form part of the PCa whole serum N-glycome biomarker.....234

Figure 5.7: ROC curves are plotted with AUC values of the peaks in the biomarker panel in indolent, significant and aggressive PCa.....236

Figure 5.8: Calibration curves of the N-glycome biomarker panel in indolent, significant and aggressive PCa.....232

List of Tables

Chapter One: Introduction

Table 1.1: Comparison of innate and acquired immunity.....	22
---	----

Table 1.2: Definitions of biophysical determinants for biomolecular interactions.....	40
--	----

Chapter Two: Materials and Methods

Table 2.1: Reaction set up for site-directed mutagenesis of scFv B8.....	81
---	----

Table 2.2: PCR conditions for site-directed mutagenesis of scFv B8.....	81
--	----

Table 2.3: Preparation of buffers for fine screening for scFv B8 Ser MT crystal formation.....	87
--	----

Table 2.4: Preparation of buffers for fine screening for scFv B8 Ser MT crystal formation.....	89
--	----

Table 2.5: Small-scale (1x) reaction for magnesium chloride (MgCl ₂) optimisation for V _H /V _L amplifications.....	101
--	-----

Table 2.6: PCR conditions for amplification of the V _H and V _L chain genes.....	101
--	-----

Table 2.7: Small-scale (1x) magnesium sulphate (MgSO ₄) optimisation for SOE-PCR.....	102
--	-----

Table 2.8: SOE-PCR reaction conditions.....	103
--	-----

Table 2.9: <i>Sfi</i> I digest of pComb3xSS vector and scFv inserts.....	104
---	-----

Table 2.10: <i>Xba</i> I and <i>Xho</i> I digestion of pComb3xSS vector.....	105
---	-----

Table 2.11: Antarctic phosphatase treatment of digested pComb3x vector.....	105
--	-----

Table 2.12: Ligation reaction components of scFv insert and pComb3x vector.....	107
--	-----

Table 2.13: Panning conditions employed for four successive rounds of panning for isolation of cPSA-specific scFv antibody.....	110
---	-----

Chapter Three: Screening and analysis of an avian anti-fPSA library

Table 3.1: Conditions for optimisation for recombinant antibody expression.....	139
--	-----

Table 3.2: Determination of anti-fPSA scFv B8 Limit of Detection (LOD) by indirect ELISA...	143
--	-----

Table 3.3: Relative risk of prostate cancer as determined by free:total PSA ratio.....	172
---	-----

Chapter Four: Generation of recombinant anti-PSA antibody fragments and N-glycan characterisation of polyclonal antibodies from an avian immune system

Table 4.1: Specific antibody titres obtained for cPSA immunised chickens.....	188
--	-----

Table 4.2: Panning input and output titres for avian anti-cPSA scFv library.....	196
---	-----

Chapter Five: Differentiation of indolent, significant and aggressive prostate cancer by robotic high-throughput N-glycan profiling

Table 5.1: Summary of identified N-Glycans from prostate cancer sera.....	227
--	-----

Table 5.2: Current clinical model predictability.....	239
--	-----

Table 5.3: Sensitivity, specificity and predictability of glycosylation biomarker panel.....	244
---	-----

Publications, Presentations, Awards, Invention disclosure and Courses Attended

Publications

Gilgunn, S., Conroy, P.J., Saldova, R., Rudd, P.M. and O’Kennedy, R.J. (2013) “Aberrant PSA glycosylation-a sweet predictor of prostate cancer” *Nature Reviews Urology*, 10(2), pp. 99–107.

Conroy, P.J., Law, R.H., **Gilgunn, S.**, Hearty, S., Caradoc-Davies, T.T., Lloyd, G., O’Kennedy, R.J., and Whisstock, J.C. (2014) “Reconciling the Structural Attributes of Avian Antibodies” *Journal of Biological Chemistry* 30;289(22) pp.15384-15392.

Spain, E., **Gilgunn, S.**, Sharma, S., Adamson, K., Carthy, E., O’Kennedy, R.J., and Forster, R.J. (2015) “Detection of Prostate specific Antigen Based on Electrocatalytic Platinum Nanoparticles Conjugated to a Recombinant scFv Antibody”. *Biosensors and Bioelectronics* 2015 77:759-766. [Epub ahead of print]

Publications submitted

Murphy, C., **Gilgunn, S.**, and O’Kennedy, R.J. (2015) “An Overview of Immunoassays” *IN: Murphy, C. and O’Kennedy, R.J. (eds) Immunology-Developments, Applications and Future Trends. PanStanford Publishings Ltd, In Press.*

Gilgunn, S., Martin, S., Wormald, M., Conroy, P.J., O’Kennedy, R.J., Rudd, P.M., and Saldova, R. (2015) “N-glycan profiling of avian immunoglobulin Y”. Manuscript submitted to *Journal of Proteome Research*.

Publications in preparation

Gilgunn, S., Murphy, K., Stöckmann, H., Conroy, P.J., Watson, R.W., O’Kennedy, R.J., Rudd, P.M. and Saldova, R. (2015) “Differentiation of indolent, significant and aggressive prostate cancer by robotic high-throughput N-glycan profiling”. Manuscript in preparation for submission.

Gilgunn, S., Conroy, P.J., Law, R.H., Hearty, S., Caradoc-Davies, T.T., Lloyd, G., O’Kennedy, R.J., and Whisstock, J.C. (2015) “High resolution crystal structure of nanomolar affinity chicken

single-chain Fv antibody -antigen complex". Manuscript in preparation for submission to *Journal of Biological Chemistry*.

Poster Presentations

Gilgunn, S., Byrne, H., Fitzgerald, S., and O’Kennedy, R.J. *‘Antibody Generation’* PROSENSE Workshop 2nd-3rd July 2013, Dublin City University, Dublin, Ireland.

Fitzgerald, S., **Gilgunn, S.,** Byrne, H., and O’Kennedy, R.J. *‘Phage Display’* PROSENSE Workshop 2nd-3rd July 2013, Dublin City University, Dublin, Ireland.

Byrne, H., **Gilgunn, S.,** Fitzgerald, S., and O’Kennedy, R.J. *‘Generation of recombinant avian library’* PROSENSE Workshop 2nd-3rd July 2013, Dublin City University, Dublin, Ireland.

Gilgunn, S., Conroy, P.J., Law, R., Whisstock, J.C. and O’Kennedy, R.J. *‘Investigating the significance of conserved cysteine residues in avian antibodies’* Annual ‘Antibody Engineering and Therapeutics’ Conference, 7th-11th December 2014, Huntington Beach, California, USA.

Gilgunn, S., Conroy, P.J., Law, R., Whisstock, J.C. and O’Kennedy, R.J. *‘Investigating the significance of conserved cysteine residues in avian antibodies’*. School of Biotechnology Research Day, 30th January 2015, Dublin City University, Ireland.

Oral Presentations

Gilgunn, S. *“Generating recombinant antibodies for prostate cancer detection”* 21st October 2012, Cold Spring Harbour Laboratory, Cold Spring Harbour, New York, USA.

Gilgunn, S., Fitzgearld, S., Byrne, H., and Kijanka, G. *“Antibody Expression, Purification and Application”*. 2nd-3rd July 2013 PROSENSE workshop - Practical Session IV, Dublin City University, Dublin, Ireland.

Gilgunn, S. *“Generation and characterisation of anti-prostate specific antigen (PSA) isoform-specific recombinant antibody fragments”* EMBO Funded Laboratory Visit, Monash University, 3rd June 2014, Melbourne, Australia.

Gilgunn, S. ‘Getting to the bottom of prostate cancer diagnostics’ ‘Thesis in Three’ DCU regional trials, 15th Oct 2014 Dublin City University, Ireland. **Awarded First place.**

Gilgunn, S. ‘Getting to the bottom of prostate cancer diagnostics’ ‘Thesis in Three’ National Finals, 20th October 2014, Sugar Club, Dublin, Ireland.

Gilgunn, S. ‘Generating antibodies and shooting crystals’ DCU, School of Biotechnology Research Day, 30th January 2015, Dublin City University, Ireland. **Awarded ‘Best Speaker’.**

Awards

Fully funded scholarship awarded from DCU to attend an international Erasmus intensive course- HUROPOL, which focused on Human Rights, Ethics, Old People and End of Life care. April 23rd- May 4th 2012, Lappeenranta University, Finland.

European Molecular Biology Organization (EMBO) Short-Term Research Fellowship – Fully funded scholarship to travel to Monash University, May-July, 2014, Melbourne, Australia.

The Faculty of Science and Health, ‘Outstanding Graduate Researcher 2014-2015’, Dublin City University.

Invention Disclosure

Gilgunn, S., O’Kennedy R.J., Saldiva, R., and Rudd P.M. (2015). *“Novel glycan signature for the differentiation of indolent, significant and aggressive prostate cancer”* **Invention Disclosure Form (IDF) Filed, July 2015, Patent Pending.**

Courses attended/Invitations

LAST Ireland/UK (Animal Handling Licence) Certified, Sept, 2012

Awarded a place to attend Cold Spring Harbour Laboratory, Cold Spring Harbour, New York, for the “Antibody Engineering & Phage Display Course Oct 16th – Oct 29th, 2012

Invited to return to Cold Spring Harbour Laboratory, Cold Spring Harbour, New York, as a Teaching Assistant (TA) for the “Antibody Engineering & Phage Display Course, November 1st - 21st, 2013 and, November 5th-18th 2014.

Abstract

Enhanced Antibody and Glycomics-Based Approaches for the Detection of Prostate Cancer

Sarah Gilgunn

Prostate Cancer (PCa) is one of the leading medical issues faced by men worldwide and is the most prevalent cancer diagnosed in men in both Europe and the United States. The vast majority of current literature concerning PCa diagnosis concludes that better PCa markers are needed to reduce over-diagnosis of low risk disease and widespread over-treatment.

The primary objective of this research entails the generation and characterisation of anti-prostate specific antigen (PSA) isoform-specific recombinant antibody fragments. Selective screening of avian single chain antibody fragment (scFv) libraries was carried out for the isolation of high-affinity, anti-fPSA and anti-cPSA specific antibody fragments. These antibodies were kinetically evaluated using surface plasmon resonance-based instrumentation and incorporated into an enzyme-linked immunosorbent assay (ELISA) that can detect PSA in the ng/mL range. Anti-fPSA scFvB8 is a nanomolar affinity antibody fragment that contains highly conserved cysteine residues. High resolution X-ray crystallography of this antibody fragment and mutant variants was successfully carried out and revealed interesting structural attributes of avian antibody fragments.

Glycan expression patterns change with the cellular modifications that accompany the onset of tumourigenesis. Hence, as a secondary objective, whole serum *N*-glycan profiling was carried out on 117 prostate cancer patient's serum using a robotised, high-throughput analysis platform for glyco-profiling which utilises ultra-performance liquid chromatography (UPLC) to obtain high resolution separation of *N*-linked glycans released from serum. A specific glycomic signature identified from this analysis can distinguish between indolent and aggressive prostate cancer with high accuracy.

The reagents generated and the results obtained from this body of work provide significant insight into antibody and glycomic-approaches that can potentially deliver the much sought after cancer-specific biomarker analysis for improved PCa diagnosis.

Chapter One

Introduction

1.1 Thesis outline

This body of work endeavours to address the clinically relevant issue of prostate cancer (PCa) diagnosis through antibody and glycomics-based techniques. There is an unmet need for better biomarkers to stratify patients into the most appropriate treatment option for their individual prostatic disease. While using prostate cancer as a target for the development of recombinant antibodies, the structural and glycomic attributes the avian immunoglobulin repertoire was explored, with the ultimate aim generating highly specific, sensitive antibodies for improved prostate cancer detection. It is anticipated that antibodies developed from this research will ultimately be transferred to a microfluidic-based platform for rapid, sensitive and specific diagnosis of prostate cancer. In addition, an insight into glycomic-approaches that can potentially deliver the much sought after cancer-specific biomarker analysis for improved PCa diagnosis is critically assessed.

1.2 Prostate cancer (PCa)

Prostate cancer (PCa) is one of the leading medical issues faced by men worldwide and is the most prevalent cancer diagnosed in men in both Europe and the United States (Ferlay *et al.*, 2010; Siegal *et al.*, 2012). Approximately 1 in 7 men will be diagnosed with PCa within their lifetime and this high prevalence brings with it health-related costs and mortality (American Cancer Society, 2015; Van der Broeck, 2014). Various risk factors for developing PCa include: age (risk increases with aging), family history (5% to 10% of PCa cases are believed to be primarily caused by inherited genetic factors or PCa susceptibility genes), race (African-

American and African-Caribbean men are more at risk than other ethnic groups) and diet (high consumption of red meat and high-fat dairy products with low consumption of green vegetables may have a slightly higher chance of developing PCa) (Irish Cancer Society, 2015).

More than 1.1 million cases of PCa were recorded worldwide in 2012, with a total of 68% of PCa cases occurring in more developed countries. Ireland is ranked 10th in the world for PCa incidence, with over 3,300 men diagnosed in Ireland each year (Irish Cancer Society, 2015; World Cancer Research Fund, 2015).

PCa is generally considered as an adenocarcinoma or glandular cancer of the prostate gland which is located beneath the bladder near the rectal wall. This walnut-sized structure surrounds the urethra and plays an integral role in the male reproductive system (Figure 1.1). One of its principal functions is the production of prostate specific antigen (PSA), a protein which liquefies semen in the coagulum, permitting sperm motility. The development of solid tumours is a multi-step process in which successive genetic insults occur in a normal healthy cell, rendering it increasingly malignant over an extended period of time (Loeb *et al.*, 1974). The tumour penetrates the capsule via the perineural and lymphatic channels to reach the periprostic tissues of the gland (Tabares *et al.*, 2006). However, the exact aetiology of the disease is unknown and the genetic and epigenetic modifications that lead to tumor development remain poorly understood (Holland, 2003). Some prostatic cancers present as an indolent disease with no clinical symptoms during the life-time of the patient, and in many cases, men die with the disease rather than from the disease. Autopsy results have suggested that 30% of men age >50 years and 70% of men >70 years have PCa (Kirk *et al.*, 1997; Hoffman, 2011). However, for some PCa patients the disease takes a much more aggressive route spreading into the seminal vesicles, bladder, and rectum and further metastasising to the

lymph nodes, bone and other organs (Bott, 2003). PCa is a heterogeneous disease and regardless of the manifestation, severely impacts a patient's life.

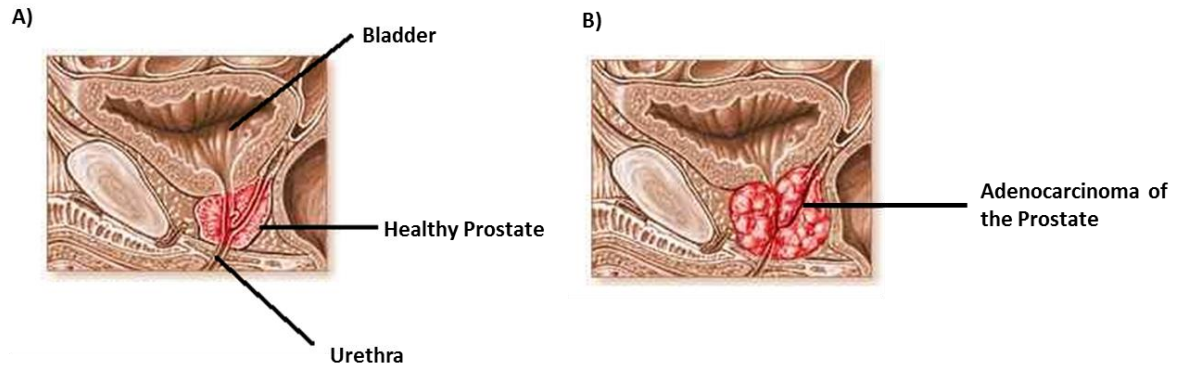


Figure 1.1: *The prostate gland.* The normal prostate (A) architecture becomes disrupted in prostate cancer (B) causing enlargement of the gland, in turn causing pain, difficulty in urinating and/or erectile dysfunction. A malignant tumor of the prostate is characterised by uncontrollable cell division and rapid growth of cancerous cells, which can invade nearby tissues. Figure adapted from Georgia Urology, 2015.

1.3 Prostate cancer diagnosis

The two principal methods used for detection of possible cancerous abnormalities of the prostate are the Digital Rectal Examination (DRE) and serum concentration of PSA (Heidenreich *et al.*, 2011). Disruption of the prostate gland can cause PSA to escape into the blood stream and thus elevated levels of PSA are an indicator for the presence of PCa. Measurement of PSA levels in serum is routinely used for PCa diagnosis and monitoring disease progression (Lilja *et al.*, 1987). Conversely, PSA levels can also rise as a result of Benign Prostate Hyperplasia (BPH)

and, hence, provides little or no insight into the biology of an individual's prostatic disorder (Pinsky *et al.*, 2006). Due to the lack of sensitivity and specificity (20% and 93%, respectively) the use of PSA alone as a biomarker for PCa remains a very controversial topic amongst clinicians and researchers (Thompson *et al.*, 2005). The use of the PSA test in a routine, national prostate screening programme is not currently carried out in Ireland. The current costs of unregulated PSA testing in Ireland (and elsewhere) are poorly understood, but are likely to be substantial. The PICTURE study is currently estimating the cost-effectiveness of PSA testing in Ireland (National Cancer Registry, 2015). PSA levels circulating in blood range from <0.1 ng/mL in a normal healthy individual to 10^4 ng/mL in a PCa patient with advanced metastasised disease. Traditionally, a PSA level of 4 ng/mL or greater indicates the requirement for a biopsy (Lilja *et al.*, 2008). Transrectal ultrasonography-guided biopsy is the preferred method to obtain prostate tissue for analysis (Heidenreich *et al.*, 2011). Once detected, the tumor is clinically staged using the internationally recognised tumor-nodes-metastasis (TNM) staging system which endeavours to identify if the primary tumor remains confined to or has spread beyond the prostate. Another fundamental part of PCa diagnosis and classification is the use of the Gleason Grading system which describes the aggressiveness of the tumour (Gleason and Mellinger, 1974). The Gleason score is based on the glandular architecture of the tumour biopsy sample which is analysed by histopathologists and divided into 5 different patterns of growth and differentiation. The two most common patterns are given a grade in the range of 1 to 5, one being most differentiated or low grade and 5 being least differentiated or high grade. The grades are summed and given as the Gleason score, generally represented as 6 (3+3), where a higher score indicates a more aggressive tumour. To be counted, a pattern (grade) needs to occupy more than 5% of the tissue section, (Heidenreich *et al.*, 2011; Gleason and Mellinger, 1974). However, a biopsy sample represents only a proportion of the pathology

of the tissue of the prostate, and a Gleason score is an inaccurate measure of the level of aggressiveness across the entire tumour.

Many prostate cancers are clinically insignificant, that is, the pathogenicity of the disease is insufficient to threaten the survival of the patient (Johnstone *et al.*, 2007). The common occurrence of these cancers is one of the main challenges in PCa diagnosis and the clinical intervention taken. There is currently no methodology to allow a clinician to specifically distinguish between a potentially fatal and insignificant PCa. It is widely thought that PSA screening can lead to increased detection of insignificant diseases and over-diagnosis. PSA was approved by the Food and Drug Administration (FDA) in 1986, initially to monitor disease progression in men with PCa and later in 1994 for PCa detection and diagnosis (Mikolajczyk *et al.*, 2004). The threshold of 4 ng/mL is frequently subjected to criticism, and lower 'cut-offs' have subsequently been considered. Patients with a PSA level of between 4-10 ng/mL are said to be in the diagnostic grey zone, of which only 25% are found to have PCa. However, approximately 15% of patients with a serum PSA level of between 2-4 ng/mL and negative DRE have PCa (Hoffman, 2011). Conversely, reducing the 'cut-off' level, while detecting more cases of cancer (sensitivity), may result in a higher rate of false positives (specificity) (Thompson *et al.*, 2005). Abnormal PSA tests lead to biopsies which result in pain, discomfort and stress to the patient. It is increasingly evident that a balance must be struck to weigh the harm from over-diagnosis against the benefit from early detection and treatment of potentially fatal PCa (Lilja *et al.*, 2008; Vickers *et al.*, 2012; Morgan *et al.*, 2015).

1.4 Current approaches in PCa detection

1.4.1 Biology of PSA

Physiologically, PSA is synthesised by the prostate ductal and acinal epithelium. Upon secretion into the lumen, PSA proteolyses gel proteins of the seminal fluid such as semenogelin 1 and semenogelin 2 (SMEG 1 and SMEG 2) (Lilja 1985; Lilja *et al.*, 1987). The normal prostate architecture confines PSA to the gland and only a minor fraction leaks into circulation. Elevation of serum PSA in PCa patients is not due to an increased expression of PSA but is considered to be as a result of disruption to the ductal lumen architecture and disordering of the basement membrane (Lilja *et al.*, 2008). PSA is an androgen-regulated serine protease and is also known as hK3, as it is a member of the tissue kallikrein family (Young *et al.*, 1992). Fifteen kallikrein members have been identified in humans and the encoding genes reside within 280kb of genomic DNA at chromosome 19q133-4 (Yousef *et al.*, 1999). Within this family, PSA (hK3) is not the only prostate-specific kallikrein as hK2 and hK4 are also expressed in the prostate and appear to be androgen-regulated (Nelson *et al.*, 1999; Yousef and Diamandis 2001).

Circulating total PSA (tPSA) consists of many molecular forms, of which the most abundant form is complexed PSA (cPSA), which accounts for approximately 84% of tPSA (Lilja *et al.*, 1991; Ozen and Sozen, 2006). Free PSA (fPSA) makes up the remaining 16% of circulating tPSA and includes various precursor isoforms of PSA (proPSA), benign PSA (BPSA) and inactive PSA (iPSA) (Nurmikko *et al.*, 2001; Peter *et al.*, 2001). The different isoforms of PSA and their synthesis are outlined in Figure 1.2. PSA is initially translated as a pre-pro-polypeptide with a 17 amino acid

leader sequence that is co-translationally cleaved during passage through the secretory pathway (Lundwall and Lilja, 1987). This results in removal of the leader sequence yielding an enzymatically inactive 244 amino acid proPSA precursor protein, [-7]proPSA. Cleavage of 7 amino acids at the N-terminal of proPSA generates the mature active protease consisting of 5 intra-chain disulphide bonds, a single *N*-linked glycan, and has a molecular mass of 33kDa (Armbruster, 1993; Kumar *et al.*, 1997). PSA is activated by hK2 as it cleaves proPSA between the arginine at position 7 and isoleucine at position 8, resulting in isoleucine as the N-terminal amino acid of the mature active PSA protein. Another member of the kallikrein family, hK4, was also shown to activate PSA (Lundwall and Lilja, 1987; Takayama *et al.*, 1997; Balk *et al.*, 2003). Truncated forms of proPSA are also generated by cleavage within the propeptide. Isoforms with leader peptide variants of 1, 2, 4 and 5 amino acids, ([-1]proPSA, [-2]proPSA, [-4]proPSA and [-5]proPSA, respectively) have been identified (Peter *et al.*, 2001; Mikolajczyk *et al.*, 1997).

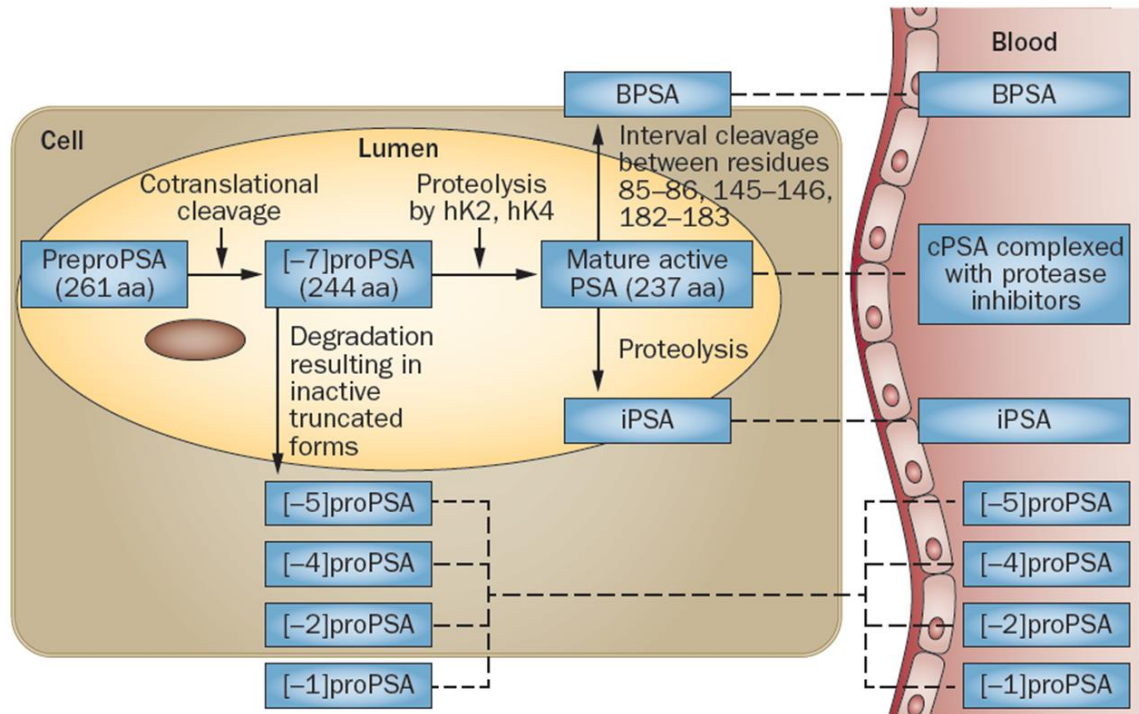


Figure 1.2: Biosynthesis of PSA. PSA is synthesised as a 261 amino acid prepropeptide, that is co-translationally cleaved in the cell to generate [-7]proPSA, an inactive precursor protein. Cleavage within the propeptide yields inactive truncated variants. In the lumen, hK2 acts as the major activating enzyme, cleaving the propeptide into a 237 amino acid active protein. Additional isoforms are cleaved at specific sites generating BPSA and iPSA. The majority of mature active PSA that enters into circulation rapidly binds to protease inhibitors, primarily α_1 -antichymotrypsin (ACT) and the rest of the unbound forms circulate as free PSA. cPSA = complexed prostate specific antigen; iPSA = inactive prostate specific antigen, BPSA = benign prostate specific antigen (Taken from Gilgunn *et al.*, 2013).

1.4.2 Improving PSA specificity for PCa diagnosis

From its initial introduction in 1986 PSA revolutionised oncology testing, yet its clinical specificity is questionable. Currently the gold standard PSA test is complemented with

parameters such as PSA-isoform specific tests, PSA kinetics and PSA density. From an increased, comprehensive, understanding of the molecular subforms of PSA, the specificity and sensitivity of PSA as a diagnostic tool has improved markedly.

1.4.2.1 Exploring the ratio of free PSA (fPSA) to total PSA (tPSA)

Early investigations examined molecular forms of PSA to enhance specificity of the test. Christensson and co-workers (1993) demonstrated that in proportion to tPSA, fPSA is significantly lower in PCa than in BPH patients (Christensson *et al.*, 1993). In current clinical practice, patients in the 'diagnostic grey zone' are screened for percentage fPSA. A cut-off of $\leq 25\%$ fPSA can detect 95% of PCa cases, reducing the number of unnecessary biopsies for patients in the 4-10 ng/mL range, and is now established as the criterion for biopsy (Catalona *et al.*, 1998). However, diagnostic use of % fPSA is prone to false negatives and many other factors influence its reliability including prostate volume, assay variation and specimen handling (Ozen and Sozen, 2006). The fPSA assay measures all forms of fPSA, thus further investigations into molecular subforms of fPSA were explored to ascertain if different isoforms are more significantly correlated with specific aspects of either PCa or BPH (Nurmikko *et al.*, 2001; Mikolajczyk *et al.*, 2002).

1.4.2.2 proPSA – The cancer-specific isoform

Mikolajczyk and co-workers (1997) identified proPSA as the major constituent of fPSA. The propeptide contains 4 amino acids instead of the normally expressed 7 amino acids and was the first of the truncated variants identified in serum (Mikolajczyk *et al.*, 1997). Later, Peter and

colleagues (2001) demonstrated that there is a mixture of pro-PSA forms in PCa sera, namely [-1], [-2], [-4], [-5] and [-7] proPSA, all of which are truncated at various points at the N-terminus (Peter *et al.*, 2001). Several studies have illustrated proPSA as a cancer-specific isoform and highlighted the proPSA forms [-2] and [-4] as particularly relevant for detection of PCa. In fact [-2] proPSA was correlated with more aggressive cancer (Mikolajczyk *et al.*, 2001; Catalona *et al.*, 2004; Sokoll *et al.*, 2010). The distinct biochemical properties of [-2] and [-4] proPSA resulted in both being considered as more stable in cancerous tissues as they cannot be converted to mature PSA by hK2 (Mikolajczyk and Rittenhouse, 2003). In a recent large clinical study, Catalona and co-workers (2011) evaluated [-2] proPSA in combination with fPSA and PSA for PCa detection in the 2-10 ng/mL PSA range for men over 50 years old. They developed a mathematical formula, the prostate health index, $\phi = [-2]\text{proPSA}/\text{fPSA} \times \text{PSA}^{1/2}$ to enhance specificity for detecting aggressive disease. The results showed greater sensitivity and specificity for the ϕ test than PSA and %fPSA. This test is FDA approved and is now commercially available in the U.S., Europe and Australia (Catalona *et al.*, 2011; Loeb *et al.*, 2015).

The U.S. Preventive Services Task Force (USPSTF) published a report in 2012 illustrating the urgent need to identify novel screening tools that can better identify indolent vs aggressive disease. In an attempt to address this issue, Loeb and colleagues (2015) included 658 men in a study to determine if ϕ improves the detection of clinically significant PCa. The authors compared the performance of PSA, %fPSA, [-2]proPSA and ϕ to predict biopsy results and the presence of clinically significant PCa. The results showed that the use of a ϕ score with a cut-off of 28.6 could potentially avoid approximately 30% of biopsies in men with benign or insignificant disease. The conclusion reached from this study was the ϕ test may be useful as

part of a multivariable approach to reduce prostate biopsies, thus limiting over diagnosis (Loeb *et al.*, 2015).

1.4.2.3 BPSA- specific for BPH diagnosis

Benign PSA is a distinct inactive form of PSA in which internal degradation at residues Arg85-Phe86, Lys145-Lys146 and Lys182-Ser183 occurs (Zhang and Leinonen, 1995). BPSA is found predominantly in the transition zone of patients with BPH (Mikolajczyk *et al.*, 2000) hence BPSA and proPSA are often said to be the 'yin and yang' forms of fPSA (Mikolajczyk and Rittenhouse, 2003). BPSA is postulated to arise from post translational cleavages by specific proteases in the hyperplastic tissue of BPH patients (Mikolajczyk *et al.*, 2004). Linton and co-workers (2003) developed an immunoassay that employed anti-BPSA antibodies specific for the internal Lys182 cleavage site. Results from this study showed that BPSA was a significant percentage of fPSA in the sera of biopsy-negative BPH patients with elevated PSA levels but not in control sera, illustrating its potential as a marker for BPH (Linton *et al.*, 2003). In combination with other makers this immunoassay may provide additional information to discriminate BPH from PCa. However, as with the other fPSA isoforms, the BPSA test has its limitations. Notably, although BPSA is generally lower in PCa patients, this correlation is weak as the vast majority of elderly men with PCa also have BPH (Mikolajczyk *et al.*, 2004).

1.4.2.4 Ratio of cPSA to fPSA

cPSA is the major component of tPSA and circulates as mature active PSA complexed with protease inhibitors, predominantly α 1-antichymotrypsin (ACT), a SERPIN type inhibitor (Lilja *et*

al., 1991). Complexes with α 2-macroglobulin (A2M) and α 1-protease inhibitor (API) are less common and are estimated to account for 1-2% of cPSA (Lilja *et al.*, 2008). As with the ratio of fPSA to tPSA, the ratio of cPSA to tPSA was shown to improve the specificity of PCa detection. Several studies have demonstrated that the cPSA immunoassay, which is an FDA approved test for PCa detection, provides better specificity over tPSA, particularly in patients with PSA serum levels of >4 ng/mL (Jung *et al.*, 2000; Okihara *et al.*, 2002,). Partin and co-workers (2003) conducted a large clinical trial which concluded that cPSA had increased specificity in detecting PCa over tPSA, however, it was noted that % fPSA and % cPSA offer no additional benefit in differentiating between indolent and malignant disease (Partin *et al.*, 2003).

1.5. Glycomics and PCa

Although unravelling the molecular subforms of PSA has enhanced the PSA test for PCa detection, there is still no outright test that can distinguish between significant and insignificant disease.

Recent advances in the field of glycobiology show aberrant glycosylation patterns as a fundamental characteristic of tumourgenesis. These observations suggest that modifications of glycoproteins with tumor-specific glycan moieties are viable diagnostic targets for cancer (Meany *et al.*, 2011). This has spurred extensive research exploring the changes in PSA glycosylation and the propensity to improve the cancer specificity of the PSA test, hence, the glycomic era could potentially revolutionise PCa diagnosis.

1.5.1 An introduction to glycans

The entire collection of sugar chains (glycans) synthesized by a cell under specific conditions (the “Glycome”) is at the top of the hierarchy in terms of size and complexity when compared to the genome, transcriptome and proteome (Freeze, 2006; Varki *et al.*, 2009). Many factors influence the glycome creating a highly diverse environment. Given that the glycome can change dramatically in response to subtle changes in the cellular surroundings it is becoming increasingly clear that extensive alterations of these specific glycosylation patterns are present in many diseases (Arnold *et al.*, 2008; Varki *et al.*, 2009). Changes to cellular pathways in cancerous tissue can affect glycan processing and these changes can be monitored by analyzing the individual sugar chains of the glycoprotein (Arnold *et al.*, 2008). It is essential to understand the structure and function of these important glycans to exploit their aberrant expression patterns for the discovery of potential new disease biomarkers. The structural complexity of glycoproteins arises from the huge diversity generated due to the presence of more than one glycosylation site, the sequence in which monosaccharide subunits are linked together, the anomeric position of the linkage (α or β configurations), the presence and degree of branching and also modifications made by non-carbohydrate substituents (for example sulfate, phosphate and acetate) (Varki *et al.*, 2009; Tharmalingam *et al.*, 2010). Glycosylation is the most common post translational modification (PTM) of secreted proteins (Lauc *et al.*, 2010). The process of protein glycosylation involves a series of glycotransferases and glycosidases (glycoenzymes) that act upon a protein as it is processed from the endoplasmic reticulum (ER) through to the Golgi apparatus. The structural modifications made by these glycoenzymes allow for the generation of cell-specific glycan expression patterns (Kattla *et al.*, 2011). Glycans have numerous functions and can be involved in correct conformation and folding of the

protein, effect the stability and activity of the protein, protect the protein from proteases and serve as recognition motifs for lectins (carbohydrate binding proteins) (Varki *et al.*, 2009; Mariño *et al.*, 2010).

The two most commonly understood mechanisms in which glycans may be linked to a protein are **(i)** *N*-Linked glycans, which are a preassembled block of 14 sugars transferred cotranslationally via the amide group of an asparagine residue. The sequence for *N*-Linked glycosylation sites is Asn-Xaa-Ser/Thr, where X is any amino acid but proline (Varki *et al.*, 2009). *N*-linked glycosylation is essential for multi-cellular life and its complete absence is embryonically lethal (Marek *et al.*, 1999). *N*-Linked glycans are classified into three general types, oligomannose, complex and hybrid, as shown in Figure 1.3. **(ii)** *O*-Linked glycosylation occurs in a stepwise fashion where sugars are added incrementally to the hydroxyl oxygen of Ser and Thr residues, and to a lesser extent, to hydroxyproline side chains. *O*-linked glycans play important roles in protein localization and trafficking, solubility, antigenicity and cell to cell interactions (Varki *et al.*, 2009; Hansen *et al.*, 1996).

Other common types of protein glycosylation include phosphoserine glycans, C-linked glycans, a rare form of glycans added to a carbon on the Trp side chain and glypiatedglycans, a glycosylphosphatidylinositol (GPI) anchor that links proteins to lipids through glycan linkages (Kattla *et al.*, 2011).

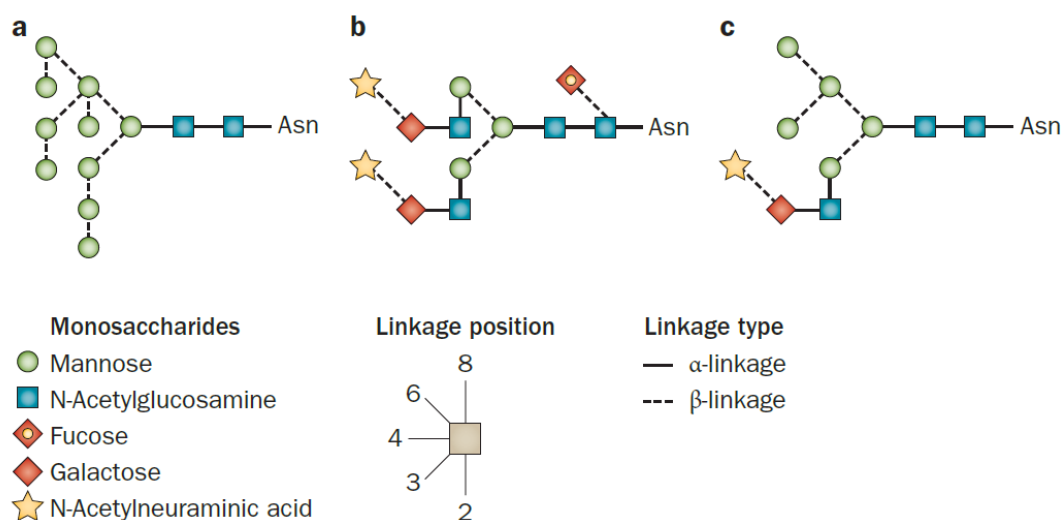


Figure 1.3: Symbolic diagrams to depict N-linked monosaccharides and linkages of the three general N-linked glycan structures (as devised by Harvey and co-workers) (Harvey *et al.*, 2009). All N-linked glycans share the common sugar core sequence $\text{Man}\alpha 1-6(\text{Man}\alpha 1-3) \text{Man}\beta 1-4\text{GlcNAc}\beta 1-4\text{GlcNAc}\beta 1-\text{Asn-X-Ser/Thr}$. (A) Oligomannose glycan. Only mannose residues are attached to the core. (B) Complex glycan. ‘Antennae’ initiated by N-acetylglucosaminyltransferases are attached to the core. (C) Hybrid glycan. Only mannose residues are attached to the $\text{Man}\alpha 1-6$ arm of the core and one or two antennae are on the $\text{Man}\alpha 1-3$ arm (Taken from Gilgunn *et al.*, 2013).

1.5.2 Recent advances in the analysis of glycoproteins

Given the heterogeneity of glycoforms, the characterisation of glycosylation patterns in serum glycoproteins can be problematic. The development of glycoanalytical techniques is to some extent lagging in comparison to the huge technological advances in the areas of genomics and proteomics. However, as our ever expanding knowledge of the biological significance of glycans grows, so does the development of glycoanalytical tools. Techniques employed in

glycan analysis include capillary zone electrophoresis (CZE), mass spectrometry (MS), High Performance Liquid Chromatography (HPLC), Ultra Performance Liquid Chromatography (UPLC), enzyme-linked immunosorbent lectin assay (ELLA) and combinatorial methodologies. Analysis of intact glycoproteins allows elucidation of the glycan structures at individual sites. CZE coupled with MS is commonly used and although efficient for characterization of oligosaccharide structure, detailed information on individual glycans and their site specificities can be difficult to obtain (Kattla *et al.*, 2011).

Lectins, a family of naturally occurring glycan-binding proteins, can discriminate between different glycan structures and are very useful tools for glycoanalysis in assays similar to standard ELISA (Wu *et al.*, 2009). Lectins are grouped based on their specificity for individual monosaccharide units, for example *Aleuriaaurentia* lectin (AAL) preferentially binds α 1-2,-3 or -6 linked fucose whereas *Galanthusnivalis* agglutinin (GNA) specifically binds terminal α 1-3 linked mannose (Varki *et al.*, 2009). Lectins can be conveniently used in ELLAs for glycan analysis but do have some limitations including low affinity (generally in the millimolar range) and high background signals with many blocking agents. Thompson and co-workers (2011) recently described an optimized ELLA method that examined the suitability of a number of blocking agents for a wide number of commonly used lectins and identified the synthetic polymer polyvinyl alcohol (PVA) as the optimal blocking agent for performing ELLAs (Thompson *et al.*, 2011).

The analysis of free glycans (*N*- and *O*-linked structures released from the glycoprotein) is essential for biomarker discovery. The initial step of liberating the glycans is generally based on chemical (e.g. Hydrazinolysis) or enzymatic deglycosylation (e.g. PNGaseF treatment) strategies (Kattla *et al.*, 2011). Once released, glycans with a free reducing end are labelled with

fluorescent molecules such as 2-AB (2-aminobenzamide), 2-AA (anthranilic acid), and APTS (8-aminopyrene-1,3,6-trisulfonate) (Varki *et al.*, 2009). The glycan profile can be elucidated by a variety of glycoanalytical techniques including weak anion exchange (WAX) HPLC, MS, CE, and UPLC-hydrophilic interaction liquid chromatography (HILIC). In combination with such techniques, exoglycosidase digestions can be used to reveal specific sequences and linkages (Royle *et al.*, 2006). The Dublin-Oxford Glycobiology group describe a HPLC-based analysis of serum *N*-glycans on a 96-well plate format. This fully automated technology platform, employing the software GlycoBase, which allows for detailed analysis of femtomolar quantities of *N*-linked sugars released from the glycoprotein (Royle *et al.*, 2008; Campbell *et al.*, 2008). This approach was designed specifically for the identification and screening for disease biomarkers (Figure 1.4).

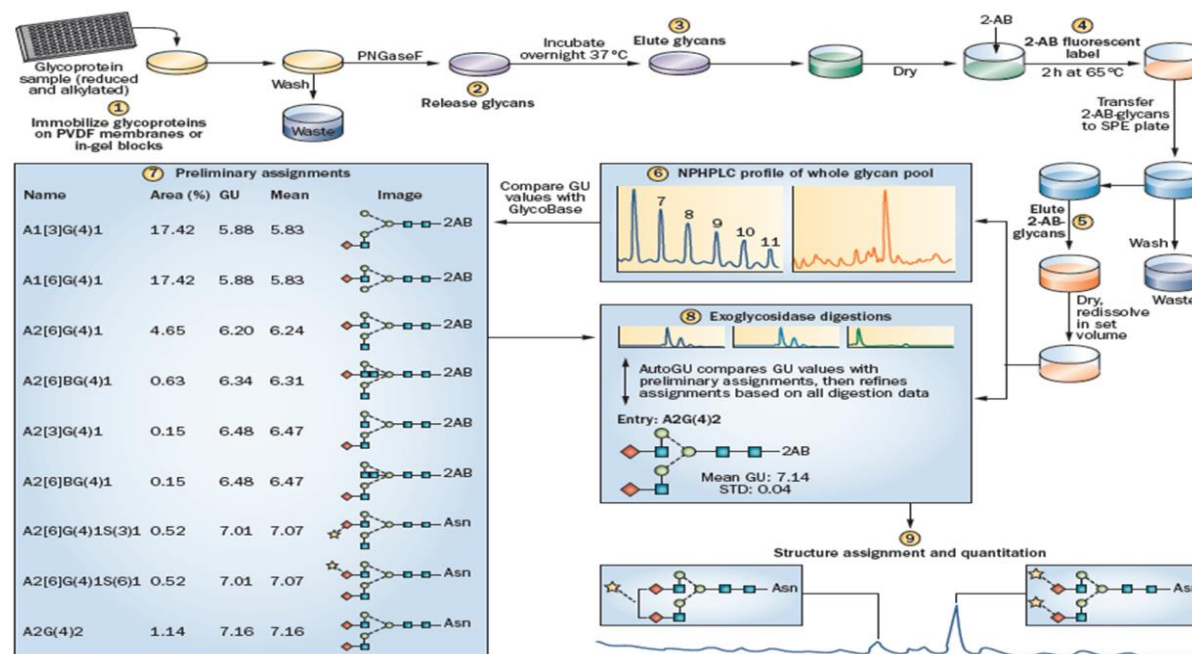


Figure 1.4: Methodology for glycan release, fluorescent labeling, NPHPLC profiling, structure assignment and quantitation. Briefly, the process includes (i) sample immobilization on PVDF membranes or in-gel blocks, sample washing and *N*-glycan release following treatment with PNGaseF, fluorescent labeling of released glycans with 2-aminobenzamide (2-AB) and (ii) glycan profiling by quantitative HPLC analysis preliminary structural assignment using the GlycoBase database of more than 350 *N*-linked glycan structures (iii) exoglycosidase digestions providing highly specific monosaccharide sequence, linkage and arm-specific details with further HPLC profiling to produce a refined list of glycan structures (Taken from Gilgunn *et al.*, 2013).

A recent study by Lauc and co-workers (2010) employed this methodology in the first genome-wide association study (GWAS) of the human N-glycome, comparing different glycosylation patterns in the plasma of 2,705 patients. This comprehensive study of genetic polymorphisms that affect protein glycosylation, identified HIF1 α as a master regulator of plasma protein fucosylation (Lauc *et al.*, 2010). Further development of such techniques focusing on reducing separation times and increasing resolution has enhanced the ability to analyze glycosylation patterns. Bones and colleagues (2010) significantly improved the efficiency and selectivity of Ultra Performance Liquid Chromatography (UPLC) profiling of 2-AB labeled N-Linked glycans by introducing a sub-2 μ m hydrophilic interaction-based stationary phase (Bones *et al.*, 2010). This study demonstrated the speed and selectivity of this technique for identification of altered glycosylation in the serum N-glycome of stomach cancer patients for possible markers for cancer progression.

1.5.3 Altered glycosylation patterns in PCa

Alterations to certain glycan structures are now considered a universal feature of malignant transformation and tumor progression. Aberrant expression can manifest in many forms in malignant cells, including loss of expression or excessive expression of particular structures and patterns (Varki *et al.*, 2009). Many studies have looked at altered glycosylation patterns in whole serum of PCa patients and of the PSA protein itself to discriminate between PCa and BPH, and to develop a more reliable diagnostic than the current PSA test (Tabarés *et al.*, 2006)

The PSA protein is a biantennary *N*-linked oligosaccharide of *N*-acetylglucosamine type (Peracaula *et al.*, 2003; Ohyama *et al.*, 2004). A noted increase in α 2-3 linked sialic acid in PCa patients has been highlighted in several studies. Ohyama *et al.*, (2004) utilized lectin affinity chromatography to analyse the binding affinity of numerous lectins to total and free serum PSA. *Maackia amurensis* agglutinin (MAA), which identifies α 2-3 linked sialic acids, bound to fPSA serum significantly more in PCa patients versus BPH patients (Ohyama *et al.*, 2004). Tajiri and co-workers (2008) also suggested that identification of α 2-3 linked sialic acids on PSA could potentially discriminate between malignant and benign disease when comparing oligosaccharide profiles of both fPSA and cPSA in PCa serum and in seminal plasma PSA (Tajiri *et al.*, 2008).

Over 10 years ago, Chandrasekaran and co-workers reported that α 1, 2- L- fucosyltransferase exhibits increased activity in the PCa cell line, LNCaP, and malignancy has been shown to be associated with increased fucosylation (Chandrasekaran *et al.*, 2002). Changes in fucosylation pathways could affect the level of core-fucosylation, as increase in fucosyltransferases results in higher core-fucosylation of Epidermal Growth Factor Receptor (EGFR), which is commonly overexpressed in carcinomas and correlates to a poor prognosis (Matsumoto *et al.*, 2008). High expression levels of the fucosyltransferases FUT2, FUT5, FUT6, FUT8 and FUT11 lead to an increased level of α 1,3- outer arm and α 1,6-linked core-fucosylated glycans in the lung cancer cell line CL1 (Liu *et al.*, 2011). In comparing 24 metastatic PCa patients with 10 healthy controls Kyselova *et al.*, (2007) demonstrated that the fucosylation of glycan structures was significantly higher in the serum of the PCa patients (Kyselova *et al.*, 2007).

Saldiva and colleagues (2011) examined if differential glycosylation patterns could distinguish between different stages of PCa and BPH by exploiting high-throughput *N*-glycan analysis

techniques (Figure 1.4) to analyze a cohort of 13 BPH patients and 34 PCa patients, of which 17 patients were Gleason score 5 and 17 were Gleason score 7. This study found an increase in both core-fucosylated biantennary glycans and α 2-3 linked sialic acids in the serum N-glycome of the PCa patients compared to the BPH patients. A notable decrease in triantennary trigalactosylated glycans (A3G3, A3G3S2, A3G3S3), and in tetraantennary tetrasialylated outer-arm fucosylated glycans (A4F1G4S4) and increases in tetraantennary tetrasialylated glycans (A4G4S4) in patients with Gleason score 7 in comparison to Gleason score 5 patients was also established. Using these altered expression patterns the authors were able to predict the Gleason score at higher sensitivity and specificity than the current PSA test (Saldova *et al.*, 2011). Patients with a Gleason score of 7 are generally positive for perineural invasion (PNI), found to be at a more advanced stage of metastasis and have an increased chance of re-occurrence. PNI is an important pre-operative indicator of the pathological stage of the tumor. Here, the authors showed that decreases in triantennary trigalactosylated glycans (A3G3) and/or bisected core-fucosylated biantennary monosialylated glycans (FA2BG2S1) and increases in tetraantennary tetrasialylated glycans (A4G4S4) correlate with PNI (Saldova *et al.*, 2011).

Collectively these studies show that the use of glycan patterns has significant potential in discriminating between significant and insignificant disease, and could help diagnose tumor spread and predict patients' survival, increasing confidence in PCa diagnosis.

1.6 Introduction to the immune system

The immune system is a highly adaptable defence mechanism that has evolved to protect multicellular organisms from invading pathogens. Two arms of the immune system, namely, the innate immune system and the adaptive immune system collaborate together to protect our body. Innate immunity is a non-specific genetically derived system that consists of many cellular and molecular mechanisms pre-deployed prior to onset of infection (Kindt *et al.*, 2007). Innate defences encompass surface barriers (e.g. skin) and are characterised by rapid assimilation, which is particularly valued in an emergency situation, and non-specific response, of a limited duration. Adaptive immunity, also referred to as acquired immunity, is highly specific and capable of selectively recognising and eliminating foreign microorganism and antigens. This arm of the immune system involves a cascade of events involving T- and B-lymphocytes. B lymphocytes (B-cells) are derived from the stem cells of the bone marrow and produce antibodies. T lymphocytes (also known as T cells) are produced in the bone marrow but processed in the thymus and constitute the basis of cell-mediated immunity (McCullough and Summerfield, 2005; Kindt *et al.*, 2007). Table 1.1 details the major differences between innate and acquired immunity.

Table 1.1: Comparison of innate and acquired immunity (Taken from Kindt *et al.*, 2007).

	Innate	Acquired
Response time	Minutes- hours	Days
Specificity	Limited and fixed	Vast diversity; improves during the course of immune response
Response to repeat infection	Identical to primary response	Much more rapid than primary response
Major components	Barriers: e.g., skin, mucosal membranes and cellular components; Mast cells, natural killer cells, phagocytes, neutrophils and dendritic cells.	Lymphocytes (B - and T-cells), antigen specific receptors, antibodies

The key difference between the acquired and innate immunity is the acquired immunity's ability to improve and refine its response following exposure to antigens. When a naïve B-cell (one that has not previously encountered an antigen) first encounters an antigen, the antigen binds to the membrane bound antibody located on the surface of the B-cell. This causes the cell to rapidly divide, giving rise to both memory B-cells that immediately recognise an antigen post primary exposure and plasma cells, which are responsible for the secretion of antigen-specific antibodies. The capability of the body to generate a diverse range of antigen-specific antibodies accounts for the remarkable nature of the adaptive immune system (Kindt *et al.*, 2007). Our natural immune repertoire is highly dynamic, with the capacity to generate a repertoire of 10^8 by affinity maturation *in vivo*, which is substantiated by continual exposure or sensitisation to antigens from the environment or by immunisation (Conroy *et al.*, 2014).

1.6 .1 Antibody structure

Antibodies are glycoproteins and are the major secretory products of the adaptive immune system. The terms immunoglobulin and antibody are used widely and interchangeably. The term “antibody” refers to a molecule that binds a known antigen while the term “immunoglobulin” describes a family of antibody proteins that have common structure features. All antibodies are constructed in the same way from paired heavy and light polypeptide chains. There are 5 major classes of immunoglobulins, denoted IgG, IgM, IgA, IgD and IgE; the differences in these lie in their heavy chains (termed γ , μ , α , δ , and ϵ respectively) which lead to differences in their effector functions (Barbas *et al.*, 2001). IgG is by far the most abundant immunoglobulin in humans and exists as many subforms (IgG1, IgG2, IgG3, IgG4).

The basic structure of an IgG antibody molecule (approximately 150 kDa) contains four polypeptides- two identical heavy chains (50 kDa each) and two identical light chains (25 kDa each) joined to form a "Y" shaped molecule (Figure 1.5). The heavy chains are linked to each other and to a light chain each by disulphide bonds, these form the arms of the 'Y' and each arm is known as a fragment antigen binding (Fab). Each Fab arm is composed of 2 variable domains (indicated as V_H for the heavy chain and V_L for the light chain) and 2 constant domains (C_H1 and C_L) (Barbas *et al.*, 2001; Sela-Culang *et al.*, 2013). Domain pairing leads to close interaction of the heavy and light arms of the Fab, known as the Fv region, which makes antigen contact. The specificity of antibodies is dictated by the hypervariable loops- HVL or complementary determining regions – CDRs which forms the sites in the Fv region for contact with its cognate antigen. There are 6 hypervariable loops in the variable domain, 3 in the light chain (L1, L2, L3) and 3 in the heavy (H1, H2, H3, which are supported by a conserved

framework region (FWR) of β -sheets (Wu and Kabat, 1970; Chothia and Lesk, 1987). The FWR provides a structural scaffold for the antigen-binding site and are important for structural diversity and in some instances may make direct antigen contacts (Sela-Culang *et al.*, 2013).

The stem of the 'Y' comes in the form of 2 additional chains of the heavy chain, C_H2 and C_H3, (called the Fc region) which is responsible for mediating the biological activity of the antibody molecule (Sela-Culang *et al.*, 2013). The hinge region is a flexible amino acid stretch, rich in cysteine and proline amino acids, in the central part (top) of these heavy chains, which links the 2 chains by disulfide bonds (Adlersberg, 1976). Disulfide bonds are an important component of the IgG structure, with intrachain disulfide bonds stabilising domain folding and interchain disulfide bonds stabilising both the interaction of heavy and light chains and the interaction of heavy chains (Barbas *et al.*, 2001).

The C_H2 domain of the Fc region also contains a conserved asparagine (N) 297-linked glycosylation site which plays a crucial role in modulating the structural stability and functional activity of the antibody (Millán Martín *et al.*, 2015). The linked carbohydrate moieties in the Fc region of therapeutic antibodies can affect both their thermal stability and physicochemical properties, along with other crucial features like receptor-binding activity, circulating half-life and immunogenicity. Hence, it is unsurprising that the structural characterisation of *N*-glycans present on antibody therapeutics is required under regulatory guidelines (Millán Martín *et al.*, 2015; Venetz *et al.*, 2015).

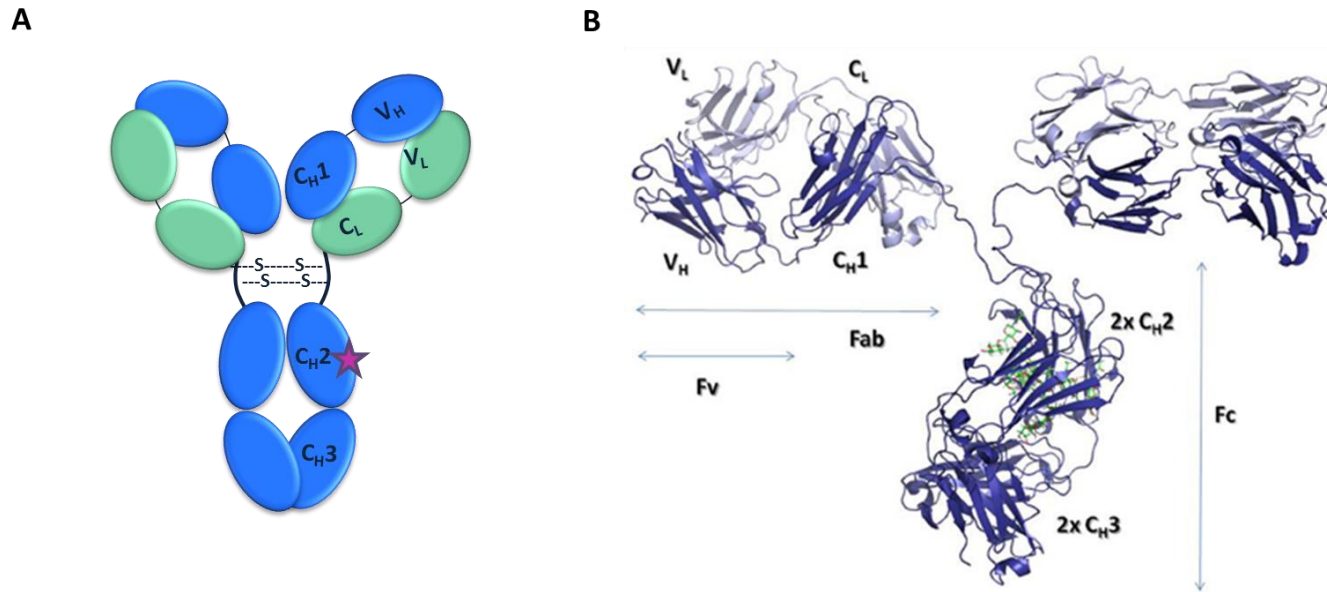


Figure 1.5: (A) Basic structure of the Immunoglobulin G (IgG) molecule and (B) Ribbon diagram of IgG, illustrating the Fv, Fab and Fc regions. The immunoglobulin is composed of two identical heavy chains consisting of constant (C_{H1} , C_{H2} , and C_{H3}) and variable (V_H) domains, associated by a disulphide bond at the hinge region of the molecule. The light chain consist of a constant (C_L) and a variable (V_L) domain that are linked with the heavy chain via a disulphide bond. The molecule has three distinct regions. The variable fragment (Fv) determines the antigen-binding specificity of the antibody. The antigen binding fragment (Fab) contains the Fv and is separated by the hinge region from the crystallisable fragment (Fc). The Fc domain confers effector functions on the antibody and contains a single *N*-glycosylation site at Asn297 (indicated by a purple star).

It was over 100 years ago when Paul Ehrlich realised that antibodies were crucial elements of the immune system and would play a key role in the development of immunotherapies “antibodies are in a way magic bullets that identify their target themselves without harming the organism” (Strebhardt and Ullrich, 2008). Today, antibodies are at the forefront of the field of targeted therapeutics and diagnostics due to their natural high affinity and excellent half-life properties that are well established. These molecules can be readily manipulated using standard molecular biology techniques into customised antibodies that are tailored to perform efficiently in their chosen end-point application. The biopharmaceutical industry has heavily invested in antibody-based diagnostics and therapeutics, which currently represents the largest and fastest growing class of biopharmaceuticals (Finlay and Almagro, 2012; Conroy *et al.*, 2014).

Antibodies are generally represented in three forms (i) Polyclonal (produced from a mixture of various B lymphocyte clones) (ii) Monoclonal (secreted from a single clone of B lymphocytes) and (iii) recombinant antibodies (are the product of genetic manipulation of antibody genes) (Leenaars and Hendriksen, 2005; Zhu *et al.*, 2014). In making the choice between monoclonal, polyclonal or recombinant antibody production, cost, production and time need to be considered. However, the desired end point application is the deciding factor and each form has its own advantages and disadvantages.

1.6.2 Recombinant Antibodies

Recombinant antibodies are ideal biological reagents for diagnostic assays as they are highly specific for a chosen analyte and genetic manipulations can be easily

employed to precisely tailor their specificity and biophysical properties (Conroy *et al.*, 2009; Hayes *et al.*, 2012). The ability to readily generate bulk quantities of recombinant proteins with low production costs in *Escherichia coli* has fuelled the emergence of an assortment of distinct antibody constructs that can be employed in a wide variety of applications within the life sciences, particularly in the field of diagnostics. Some common recombinant antibody constructs are illustrated in Figure 1.6.

The smallest antibody fragment that retains the full monovalent antigen binding capabilities of a human IgG is the Fv, which comprises of the V_H and V_L domains held together by non-covalent interactions. However, these molecules cannot be produced with ease by proteolysis of the intact IgG and recombinant single-chain Fv (~28kDa) constructs, in which the variable heavy (V_H) and variable light (V_L) domains are associated via a polypeptide linker, are commonly used to mimic Fv fragments (Hudson, 1998; Hollinger and Hudson, 2005). The scFv is more stable than the Fv (which is more aggregation prone than the scFv) and typically, a flexible 15 residue (Gly₄Ser)₃ linker is incorporated as it aids in antibody folding and stability for bacterial expression (Barbas *et al.*, 2001).

The antigen-binding fragment (Fab) is a well characterised, larger, more stable fragment of approximately 50kDa. The Fab fragment consists of the variable heavy and light chains and the constant heavy and light chains (V_H & C_{H1} and V_L & C_L). An interchain disulphide bond exists between the C_{H1} and C_L (Joosten *et al.*, 2003).

There are various advantages and disadvantages of the different types of recombinant antibody structures that can be engineered as it is necessary to take the end point into consideration. It should also be noted that antibody fragments

can be relatively easily reformatted for improved characteristics such as expression (Hayhurst *et al.*, 2003), reduction in non-specific binding (Birtalan *et al.*, 2008) and enhanced specificity (Stura *et al.*, 2011). The selection of superlative antigen-specific recombinant fragments can be achieved using a process known as phage display technology.

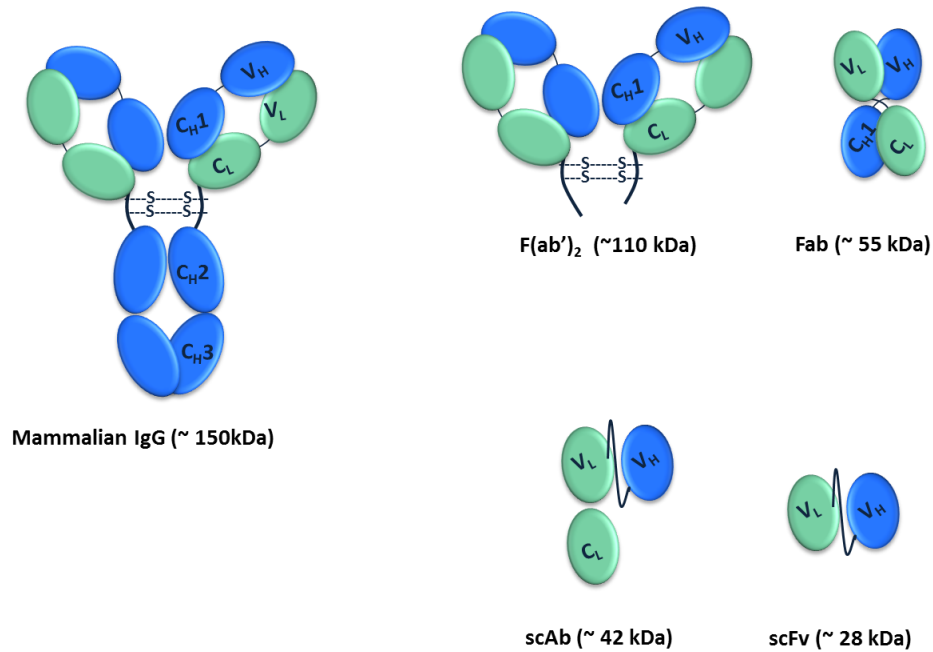


Figure 1.6: Recombinant antibody fragment constructs. The mammalian IgG construct (150kDa) is secreted by B cells of the immune system and consists of two heavy and two light chains linked together by disulphide bonds. The chains have constant (C) and variable regions (V). Diversity in antibody–antigen interactions is achieved through the ‘hyper-variable region’ in the variable light (V_L) and variable heavy (V_H) chains, of approximately 3–25 amino acids, which is known as the complementarity determining regions (CDRs). The scFv consists of the V_H and V_L chains joined together by a flexible linker. A scAb fragment comprises of a scFv with an additional constant light chain added to the end of V_L . Fab contains the full light chain with V_H and C_{H1} . $F(ab')_2$ is comprised of two fab fragments held together by disulphide bond.

1.7 Phage display technology

When George P. Smith first demonstrated, in 1985 that the link between phenotype and genotype could be established in filamentous bacteriophage the advent of phage display technology emerged (Smith, 1985). Over the last few decades an increased understanding of antibody structure and function has allowed phage display to become the most powerful tool in the controlled selection of unique antibodies with properties unattainable from conventional hybridoma methods. Some five years later McCafferty *et al.*, (1990) identified a novel method in which a functional scFv was successfully expressed on the surface of bacteriophage, providing a way to select antibodies from libraries based on the antigen specificity for the individual clones (McCafferty *et al.*, 1990; Hoogenboom, 2005). Phage display can be used with naïve, immunised or synthetic repertoires (Bradbury *et al.*, 2011) and allows for the controlled selection and screening of antibodies against defined antigen conformations from a highly diverse combinatorial library of antibody fragments (Bradbury and Marks, 2004).

Selection involves subjecting the phage library to successive rounds of exposure to antigen, allowing for antigen-specific phage-displayed antibodies to bind to their targets, in a process known as panning. The antigen-specific phage are recovered and subsequently infected in bacteria for re-amplification (Hoogenboom, 2005). Ideally, a high affinity antibody would be selected for in the initial round. However non-specific binding of phage confines the enrichment that can be achieved during each round and hence typically 2-5 rounds of panning are necessary to identify high affinity binders (Bradbury and Marks, 2004). One round of panning is shown in Figure 1.7.

The most common method of selection is panning on immobilised antigen coated onto a solid support, but many other methods exist (Bradbury and Marks, 2004; Conroy *et al.*, 2009). There are many selection considerations that must be taken into account when designing phage selection regimes. In pursuance of a high affinity antibody, the coating antigen concentration should ideally be decreased in each successive panning round. Antigen presentation is also of particular significance as some epitopes may be partially denatured when immobilised on solid surfaces, leading to the selection of antibodies that will not recognise the native antigen. A means of avoiding this issue is by indirect coating, using a commercially available antigen specific antibody to capture the target antigen (Bradbury and Marks, 2004; Conroy *et al.*, 2009). It is also very important to use a suitable blocking agent in order to inhibit some of the non-specific phage binding; some commonly employed blocking agents include BSA and semi-skimmed milk powder. The addition of a detergent such as Tween 20 also helps reduce non-specific phage binding (Bradbury and Marks, 2004).

The washing step is typically performed by rinsing with PBS and/or PBS containing Tween 20 (0.05-0.05%, v/v) between 10-30 times (Barbas *et al.*, 2001). After the first cycle it is very important to recover all phage-binding antigen. Thus, the washing conditions should be less severe. However, in subsequent rounds washing should be more stringent to allow for enrichment of strong binders (Bradbury and Marks, 2004).

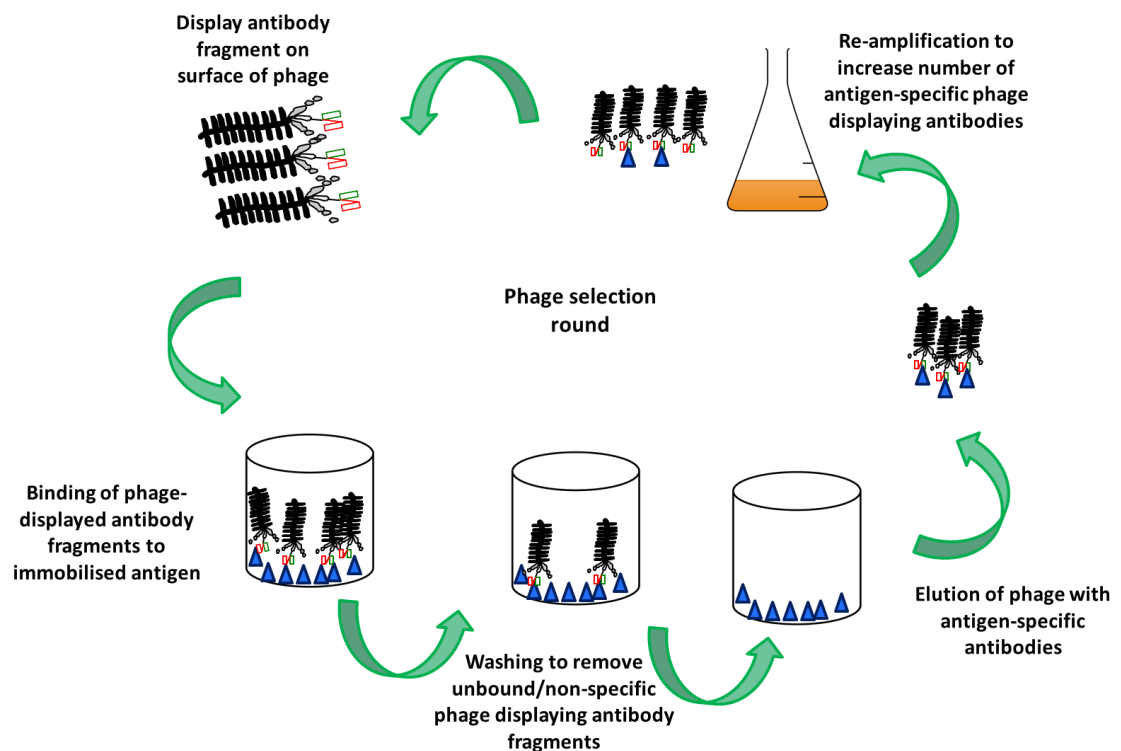


Figure 1.7: Selection of antibodies by phage display (panning). The process of ‘panning’ is a means of selection of defined clones through binding of the phage library to the target, washing unbound phage and elution of the target-specific phage. This process is repeated for a number of ‘rounds’ to enrich for a high affinity binder.

1.7.1 Filamentous bacteriophage biology

Filamentous bacteriophage (Ff) are a structurally and genetically diverse group of viruses capable of infecting *E. coli*. The Ff class of filamentous bacteriophages (f1, fd and M13) are the main tools utilised in phage display technology and are also the most extensively understood (Barbas *et al.*, 2001; Qi *et al.*, 2012). Ff phages, illustrated in Figure 1.8A, are a rod-like shape with a circular single stranded DNA (ssDNA) genome that infect *E. coli* cells and do not kill the host during productive

infection. The 'tube-like' capsid consists of approximately 2,700 helically arranged molecules of the major coat protein pVIII (~50 amino acids), while the ends are composed of two pairs of proteins – pVI & pIX and pIII & pVII. At one end there is approximately 5 molecules of both pVII (33 amino acids) and pIX (32 amino acids), whereas the opposite end is capped by 3-5 copies of the 112 residue long pVI and the 406 amino acid pIII (Barbas *et al.*, 2001; Bratkovic, 2010).

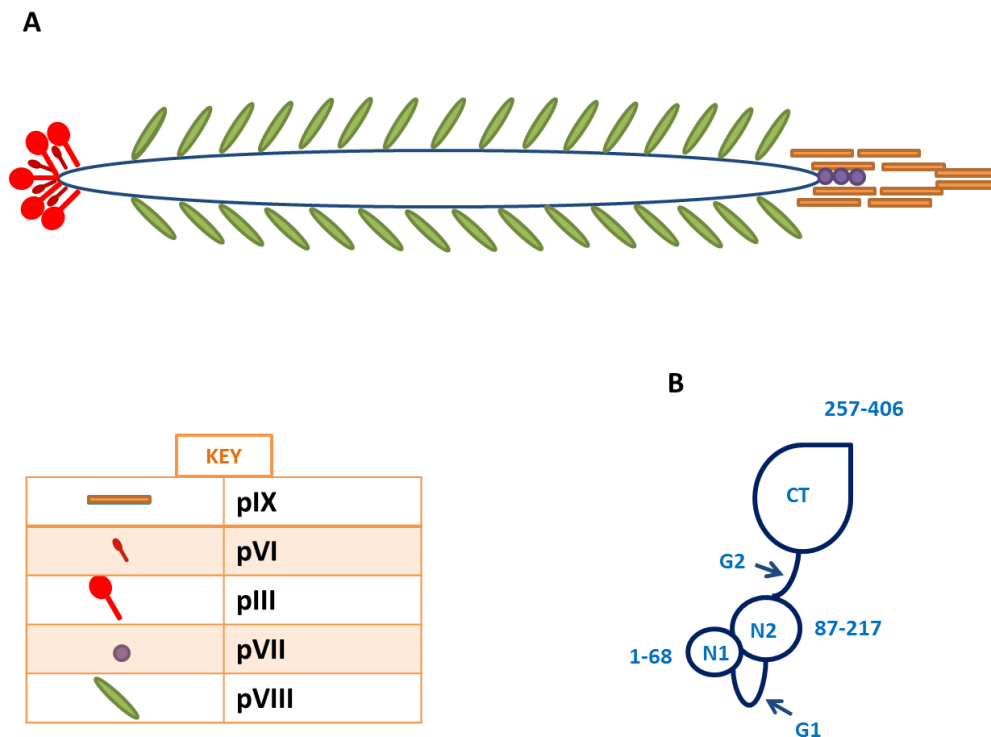


Figure 1.8: *Filamentous bacteriophage*. (A) Schematic representation of the phage particle showing each of the coat proteins. Ff phages are rod-like viruses with a circular ssDNA genome. (B) The minor coat protein pIII is essential for phage infectivity and is comprised of 3 domains separated by glycine-rich regions.

pIII mediates entry of the phage into the cell. It consists of 3 domains, designated N1, N2 and CT, which are separated by glycine rich areas (Figure 1.8B). The N-terminal domains are essential for infectivity – N1 is 68 amino acids in length and

interacts with the bacterial membrane protein TolA, facilitating translocation of the phage DNA into the cytoplasm and insertion of the phage coat proteins into the cytoplasmic membrane (Kehoe and Kay, 2005). The N2 domain (residues 87-217) is in close proximity to the N1 domain, forming a horseshoe structure. Infection begins when the N2 domain of pIII attaches to the F pilus of *E. coli*. The third domain of pIII is the 150 residue carboxy-terminus (CT) which is essential for forming a stable phage particle (Figure 1.9) (Barbas *et al.*, 2001).

Phage display libraries have been successfully generated using the major coat protein pVIII, however, in direct comparison between pIII and pVIII, pIII was shown to be far more efficient, thus is widely accepted as the coat protein of choice for the display of recombinant antibody fragments (Kretzschmar and Geiser, 1995; Bradbury and Marks, 2004). The clear advantage pIII has over the other coat proteins is its ability to package large insertions without interfering with phage integrity and infectivity, its ability to display monovalent copies of phage antibody particles and, also, the wide availability of suitable vectors (Barbas *et al.*, 2001).

Both phage (Cwirla *et al.*, 1990) and phagemid vectors (Barbas *et al.*, 1991) have been used to display recombinant antibody fragments and other proteins successfully. Phage display vectors will readily support presentation of short peptides on all copies of the minor coat protein (pIII), as the gene encoding the recombinant protein is included in the phage genome, leading to polyvalent display. Polyvalent display tends to allow for the selection of lower affinity variants after the panning (selection) process (Bradbury and Marks, 2004; Bratkovic, 2010).

The development of phage-based plasmid vectors, known as phagemid vectors, allows for monovalent display and thus the selection of high affinity antigen-specific antibodies (Barbas *et al.*, 1991). Phagemids bring together features of plasmids (replication of dsDNA) with features of phage vectors (production and packaging of ssDNA into virions) (Bratkovic, 2010). Phagemids contain both an *E. coli* origin of replication and phage origin of replication. They are specifically engineered to express recombinant pIII-fused phage under controlled conditions. However, they are unable to make phage unless the bacterial host is super-infected with helper phage. Helper phage are essentially defective Fφ phage. They carry mutations in their packaging signal and also carry antibiotic resistance genes. The packaging signal is flawed, and hence the helper phage cannot make phage when alone in the bacterium. However, in the presence of a phagemid, which contains an optimal packaging signal, the phagemid is preferentially packaged over helper phage. Thus, phagemid preparations are both phenotypically and genotypically heterogeneous (Bradbury and Marks, 2004).

The three domains of pIII have important functions in infectivity; therefore, when using phagemid vectors that express pIII-fusion proteins, it is important to consider to which end of pIII the fragment will be fused. In order to introduce the phagemid DNA into *F' E. coli* hosts, the amino terminal domain is required (Figure 1.9).

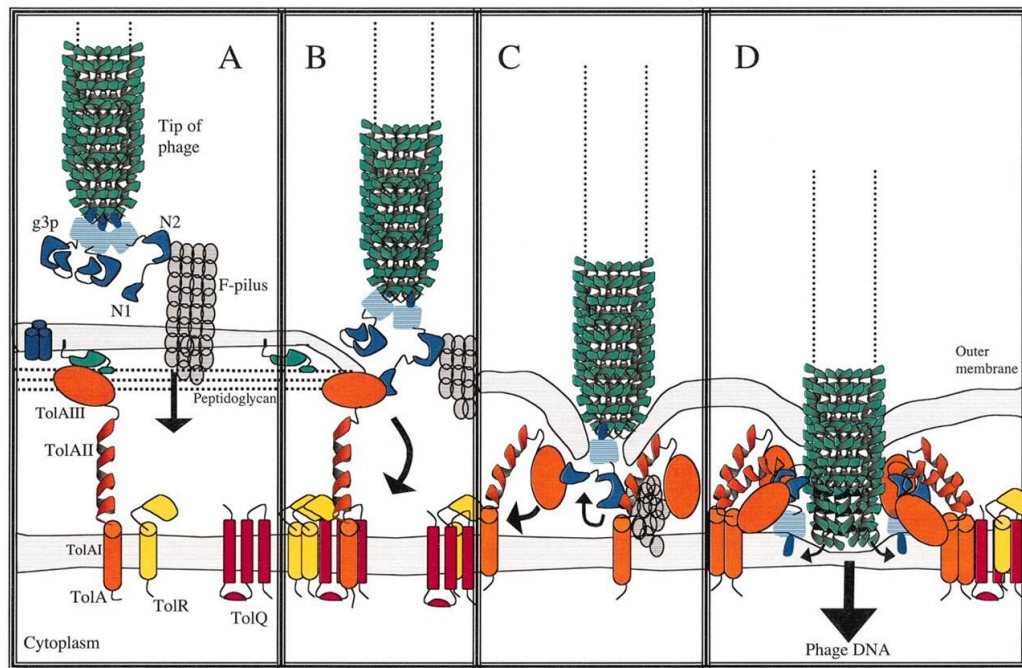


Figure 1.9: Model for the early events in the phage infection of *E. coli*. (A) The phage pIII second N-terminal domain (N2) interacts with the F-pilus on the outside of the bacteria. The outer membrane proteins OmpF (blue cylinders) and Pal lipoprotein (green) are also included, as there have been reports of TolA interacting with these proteins prior to infection. (B) After F-pilus retraction, the first N-terminal domain of pIII (N1) binds to the C-terminal domain of bacterial TolA (TolAIII). (C) The retracting pilus brings phage pIII domains in closer contact with TolA domains, allowing the TolA to assume a more compact state of assembly, bringing the outer and inner membranes of the bacteria closer together. At this stage, the central domain of TolA (TolAII) has the possibility to interact with the N2 domain of phage pIII. (D) The phage pIII is inserted into the inner membrane, and the cap of the phage head is opened to allow phage DNA to enter the bacteria. Taken from Karlsson *et al.*, 2003.

This step is followed by infection with helper phage, which is required for phage production (including packaging of the phagemid genome) and the display of fusion proteins that constitute the library. To allow for infection with helper

phage, the amino terminal of pIII in the fusion expressed by the phagemid should be deleted. Truncation of pIII, leaving residues 198-406, serves as the fusion partner for the displayed recombinant antibody fragment. The resulting packaged phagemid carries both native 406 amino acids long pIII (necessary for infection) and the protein-fused truncated pIII, which is displayed for selection (Barbas *et al.*, 1991; Radar and Barbas, 1997; Barbas *et al.*, 2001).

The pComb series of phagemid vectors control expression under a single *lac* promoter with *ompA* and *pelB* leader sequences to direct expression of the recombinant antibody fragment-pIII fusion into the periplasmic space (Barbas *et al.*, 2001). The pComb3X vector series has several features that make it very attractive for display of antibody fragments (See Chapter Four, Figure 4.8 for schematic representation of pComb3xSS phagemid vector).

This vector facilitates the soluble expression of recombinant antibodies through the addition of an amber (TAG) stop codon at the junction between the gene III and the region encoding the antibody fragment. Propagation of phage during selection in suppressor strains of *E. coli* such as XL1 blue or ERB2537 allows for transcription of gene III fusion antibody for display. However, by incorporating phage into a male non-suppressor strain such as TOP10F', the stop codon is read and soluble expression of the recombinant antibody fragment into the periplasmic space, without continued fusion to the phage particle, is produced (Barbas *et al.*, 2001; Bradbury and Marks, 2004). This vector also includes two peptide tags at the carboxyl terminus of the displayed protein. The hexa-histidine (His₆) tag facilitates purification by immobilised metal affinity chromatography (IMAC) and the influenza haemagglutinin (HA) epitope (YPYDVPDYAS) tag, which facilitates for simple detection (Barbas *et al.*, 2001).

1.8 Antibody characterisation by surface plasmon resonance

Classical methods for characterisation of antibody interactions such as ELISA demonstrate that a binding interaction is taking place. This provides evidence for the formation of the complex (i.e. antibody and antigen). When specific molecular interactions such as this occur, the principles of thermodynamics and biomolecular structure and recognition come into play (Wilson, 2002).

Within the last few decades, surface plasmon resonance (SPR)-based optical biosensors are at the forefront of instruments for the study of biomolecular interactions (deCrescenzo *et al.*, 2008). SPR-based methods are favourable as the binding events can be monitored in real-time and without the requirement of ancillary labels. These key attributes facilitate accurate kinetic measurements, rapid analysis times, reduced additional costs and potential heterogeneities or other complications that can arise when using labelled interactants (Wilson, 2002; Leonard *et al.*, 2011).

SPR is an optical phenomenon that occurs when polarised light is passed through a prism and undergoes total internal reflection at a metal surface (of a sensor chip) and liquid interface. The light energy excites plasmons in the metal surface causing them to resonate and emit electromagnetic waves. A decrease in the intensity of the reflected light is observed by an optical detection unit as these plasmons absorb incident light (Wilson, 2002; Leonard *et al.*, 2011) (Figure 1.10).

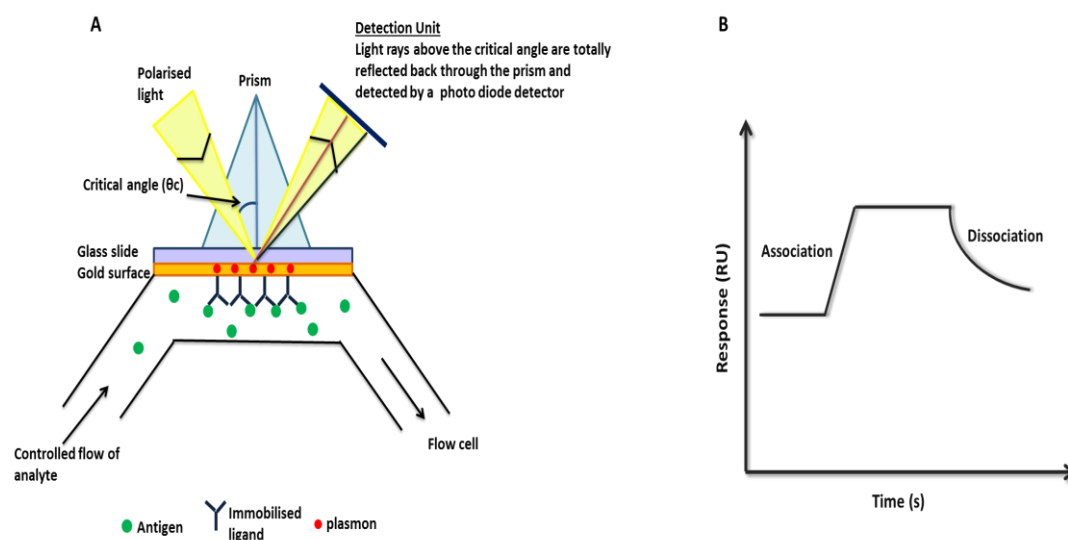


Figure 1.10: Schematic representation of a Surface Plasmon Resonance (SPR)-based system. (A) SPR occurs on a thin layer of gold, which is sandwiched between a glass slide and the sample solution flowing through a specially designed microfluidic cartridge. Polarized light from a light source is passed through the prism at an angle of incidence greater than the critical angle causing total internal reflection to occur. When the light hits the gold at the critical angle (θ_c), not only does it internally reflect, but it also leaks an electromagnetic energy called an evanescent wave which moves exponentially across the gold interface exciting plasmons, this is seen as a drop in the intensity in the reflected light and is measured the detector. When analyte binds to the sensor surface a mass change occurs near the surface of the sensor chip. The refractive index near the sensor chip surface changes, this in turn changes the angle at which the SPR occurs. The photo diode array detector detects the change in SPR angle and displays the change in real-time on a sensorgram. **(B)** Association and dissociation are measured in arbitrary units and displayed in a graph called a sensorgram. The example shown consists of an association, steady state and dissociation phase of an exponential molecular interaction.

Over 40 years ago, seminal work by Otto, Krestschmann and Raether established that the SPR phenomenon occurs when surface plasmon waves, related to collective excitations of electrons in metal, are excited by light at a metal/liquid interface (deCrescenzo *et al.*, 2008). The vast majority of SPR biosensors that are commercially available today still utilise the geometry proposed by Krestschmann and Raether to direct polarised light through a prism and then reflect it from a gold film, of approximately 50 nm, deposited on its surface (deCrescenzo *et al.*, 2008).

BiacoreTM (GE healthcare) developed one of the first instruments that utilises SPR for 'real-time' analysis of biomolecular interactions in a 'label-free' environment. Biacore technology exploits SPR methodology by using a gold sensor chip immobilised with ligand and passing over analyte in solution. As analyte binds to the immobilised ligand a change in mass at the metal surface occurs. This binding interaction perturbs the resonating plasmons and this change is recognised by the optical detection unit as a decrease in the intensity of reflected light. This binding profile is then reported as a sensogram (Wilson, 2002; Leonard *et al.*, 2011).

The key application of Biacore in antibody characterisation is determination of affinity and kinetics characteristics of a biomolecular interaction, which in turn is determined by a number of biophysical parameters (Table 1.2). The simplest interaction is the 1:1 binding event, which can be described for immobilised ligand (A) binding to cognate analyte (B) forming a complex (AB) as illustrated by the equation in Figure 1.11 (Leonard *et al.*, 2011).

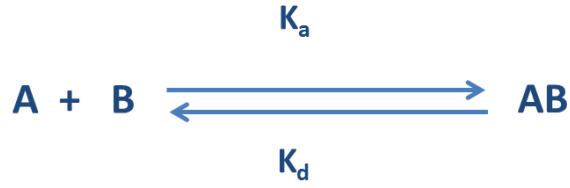


Figure 1.11: Binding interaction for analytes A and B forming the complex AB

Table 1.2: Definitions of biophysical determinants for biomolecular interactions (Taken from Hearty *et al.*, 2010)

Constant	Definition	Description	Proportionality	Unit
k_a	Association rate constant (also referred to as 'on-rate')	No. of AB complexes formed per unit time at unit concentration of A and B	Affinity $\propto K_a$	$M^{-1}s^{-1}$
k_d	Dissociation rate constant (also referred to as 'off-rate')	No of AB complexes dissociating per unit time	Affinity $\propto 1/ K_d$	s^{-1}
K_A	Equilibrium association constant	Affinity propensity to association of A + B	$K_A = K_a/K_d = [AB]/[A][B]$	M^{-1}
K_D	Equilibrium dissociation constant	Affinity contribution to stability of AB complex	$K_D = K_d/K_a = ([A][B])/[AB]$	M

BiacoreTM instruments use surfaces with carboxymethylated dextran layers that form a hydrogel-like polysaccharide matrix. These carboxymethylated dextran chains are tethered to a gold substrate via a nanometric-thin self-assembled alkyl monolayer. Stable anchorage is mediated through a sulphur atom present at one

end of the alkyl chain (deCrescenzo *et al.*, 2008). Ligands can be covalently immobilized via primary amine groups after activation of the carboxymethyl groups on the sensor surface with a mixture of N-hydroxysuccinimide and N-ethyl-N'-(dimethylaminopropyl) carbodiimide (deCrescenzo *et al.*, 2008).

The Biacore™ 3000 provides comprehensive characterisation of binding events. Furthermore kinetic analysis and epitope mapping can be easily conducted using the 3000 system (Biacore, 2015). The Biacore™ 4000 is a powerful tool for larger scale, high through put screening. This system allows 60 hour automated runs and parallel analysis of up to 16 targets or 4800 interactions in 24 hours. The Biacore 4000 is particularly useful for screening and ranking large libraries of antibodies direct from crude *E. coli* extracts whilst also characterising antibody specificity and selectivity (Biacore, 2015).

Yu and co-workers (2004) reported the use of a CM5 Biacore chip as a dextran surface on which a sandwich assay for free PSA was conducted (Yu *et al.*, 2004). The assay combined SPR and fluorescent labelling, i.e. surface plasmon field-enhanced fluorescence spectroscopy (SPFS). This technology used surface plasmons to excite the fluorescent label and produce the signal. The limit of detection observed was 80 fM for a 40 minute contact time, which is many orders of magnitude lower than detection limits of third generation PSA assays commercially available today.

1.9 Antibody characterisation by X-ray crystallography

Understanding protein structure and its relationship to function is of paramount concern to scientists to elucidate the interactions between proteins, for example

between an antibody and an antigen. X-ray crystallography is a powerful, robust method for obtaining high resolution structural information about biological macromolecules. The history of protein crystallography started at the beginning of the 19th century when W.C Roentgen discovered X-rays and since then the major contributions X-ray crystallography has made to the scientific community have been recognised with the awarding of 14 Nobel prizes (Shi, 2014) (Figure 1.12).

Scientist(s)	Year	Field	Research
Roentgen	1901	Physics	Discovery of X-rays
Laue	1914	Physics	Discovery of X-ray diffraction by crystals
W.H Bragg, W. L Bragg	1915	Physics	Use of X-rays to determine crystal structure
Sumner	1946	Chemistry	Crystallization of the enzyme urease
Pauling	1954	Chemistry	Nature of chemical bond and its application in structure determination
Kendrew and Perutz	1962	Chemistry	Pioneering work in protein structure determination
Crick, Watson, and Wilkins	1962	Physiology or Medicine	Discovery of DNA double helix structure
Hodgkin	1964	Chemistry	Structural elucidation of many biochemical substances, including vitamin B12
Deisenhofer, Huber, and Michel	1988	Chemistry	Structure of bacterial photosynthetic reaction centre
Walker	1997	Chemistry	Structure of F1- ATPase
MacKinnon	2003	Chemistry	Structure and function of potassium channel
Kornberg	2006	Chemistry	Structure of RNA polymerases
Ramakrishnan, Steitz, and Yonath	2009	Chemistry	Studies on the structure and function of the ribosome
Kobilka	2012	Chemistry	Structure and function of G-protein-couple protein receptors



Figure 1.12: History of X-ray crystallography in Nobel prizes (Adapted from Shi, 2014)

Some of the key milestones highlighted in Figure 1.12 only tip the iceberg in terms of the immeasurable impact the field of structural biology has had on science and

medicine. The discovery of the structure of the DNA double helix model is likely the famous development (Watson and Crick, 1953). However, owing to crystallography we have also revolutionised the concepts of vaccine design through insights gained from viral structures (Rossmann *et al.*, 1985; Doores *et al.*, 2015), gained significant insights into the understanding of membrane proteins (Lyons *et al.*, 2012) and developed what is now one of the central paradigms in immunology from determination of the crystal structure of the class 1 Major Histocompatibility Complex (MHC) molecule HLA-A2 (Bjorkman *et al.*, 1987).

Protein crystallisation has advanced significantly since the determination of the first protein structure of myoglobin in 1957 (at 6 Å) (Kendrew *et al.*, 1958). Information sourced from X-ray diffraction studies can uncover considerable details regarding the dynamics and heterogeneity of protein structure and function (Acharya and Lloyd, 2005). Much research focuses around the three dimensional structure of proteins and the protein data bank (PDB) was established in 1971. To date 10, 9661 biological macromolecule structures have been archived (as of June 2015). The majority (approximately 89%) of structures housed in the PDB were determined by X-ray crystallography.

1.9.1 Crystal structure determination X-ray diffraction

The three dimensional molecular structure of a protein can be obtained from a crystal that is exposed to an X-ray beam. A protein crystal is a solid composed of repeated structural motifs in a 3-dimensional lattice and the building blocks of a crystal (asymmetric units) are arranged, according to well defined symmetries, (forming 230 space groups) into unit cells that are repeated in 3-dimensions. Due

to their chirality, biological macromolecules do not crystallise in all space groups reducing the number of available space groups to 65 (Wukovitz and Yeates, 1995).

The precise positions/arrangements of atoms in a protein crystal can be determined. When a crystal is exposed to an X-ray beam, diffraction occurs, resulting in the scattering of the X-ray beam in a predictable pattern based on the crystal lattice structure. X-ray diffraction is a reflection of the architecture of the crystal lattice; hence any imperfections in the crystal generally result in a poor diffraction pattern (Acharya and Lloyd, 2005).

In a diffraction pattern, each spot is defined by 3 parameters, (1) amplitude (can be measured from the intensity of the spot), (2) wavelength (set by X-ray source) and (3) 'phase', which all need to be known for all the spots in order to determine the position of the atoms that give rise to the diffracted beams. As 'phases' are not directly measurable (the 'phase-problem' is further discussed in section 1.9.3.1), reasonable estimates for the phases of all diffracted beams can be obtained through indirect methods. Once these three parameters are known, they can be combined to produce an electron-density map. A model can then be progressively built into the observed electron density, producing a relatively accurate protein structure model (Figure 1. 13) (Acharya and Lloyd, 2005).

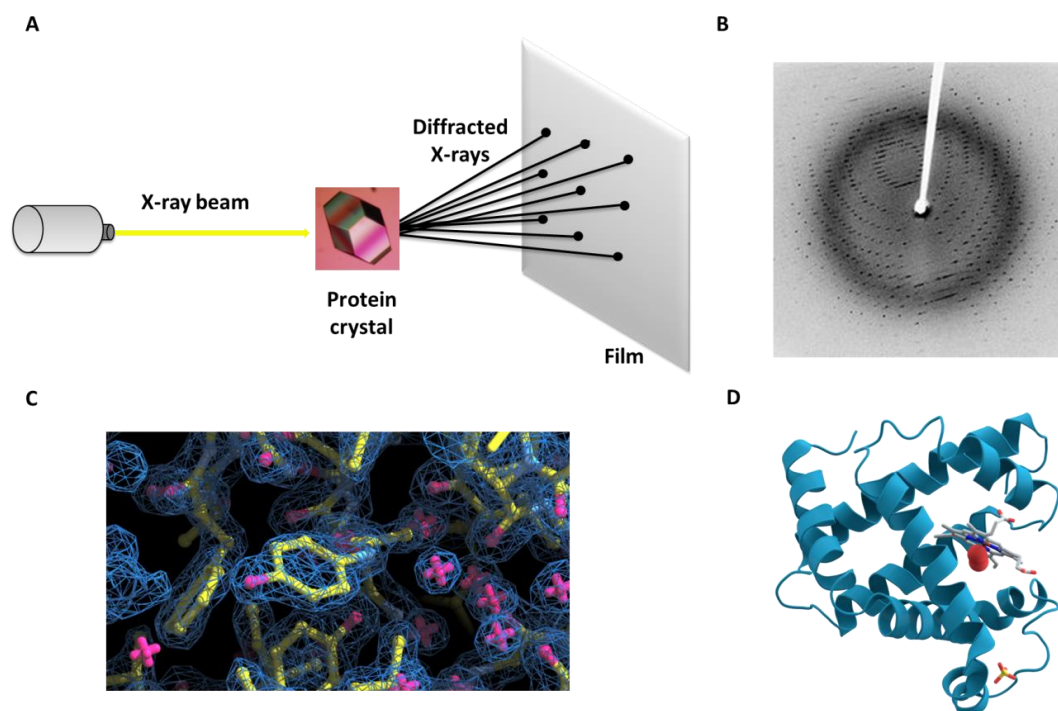


Figure 1.13: Overview of protein crystallography by X-ray diffraction. There are many stages involved in determination of protein structure (A) X-ray diffraction; the scattering of X-rays when they hit a crystal lattice (B) typical X-ray diffraction pattern: The diffraction pattern is a map of the arrangement of the molecules or atoms in the crystal. The pattern of spots (reflections) and the relative strength of each spot (intensities) can be used to generate (C) an electron density map; the electron density map is interpreted in terms of a set of atomic coordinates. This generates the preliminary model which needs to be refined (D) Ribbon diagram of the structure of myoglobin. Amino acid residues are fitted into the electron density map to produce the protein model. The final model is refined using methods such as molecular replacement. Figure adapted from (Acharya and Lloyd, 2005).

1.9.2 Data collection:

Protein crystallography relies almost entirely on the scattering of X-rays by the electrons in the molecules that constitute the sample under investigation. It was

more than a century ago when Roentgen discovered X-rays and in the recent decades the sources of X-rays used for crystallographic experiments and the methods of their detection have undergone dramatic changes and improvements.

Traditional laboratory X-ray sources have a rotating anode X-ray generator, (predominantly with a copper target) which gives a peak of CuK α radiation of 1.54 Å. However, data is now more commonly collected using synchrotron radiation sources. Synchrotrons have existed since the 1940's. However, in 1980 the world's first synchrotron dedicated to producing synchrotron light for experiments was built. Synchrotrons are now considered the single most powerful tool available to X-ray crystallographers (Wlodawer *et al.*, 2013, Diamond Light Source, 2015).

Synchrotron radiation is emitted when electrons, travelling close to the speed of light, are forced to move in a circular orbit under the action of a magnet. Unlike conventional lab sources, this type of X-ray beam has much higher intensity and data can be collected at any wavelength approximately in the range 0.5-3 Å. Synchrotrons are very large facilities (roughly the size of a football field) and contain several dedicated beamlines with charged-coupled detectors (CCDs), making data collection exceptionally accurate and fast (Wlodawer *et al.*, 2013, Diamond Light Source, 2015).

1.9.3 Data analysis

Once protein crystals that diffract X-rays to a reasonable resolution have been obtained, the obstacle to overcome is to obtain an interpretable electron density map. The dataset which is built from many different images contains spots relating to the electron density calculated by Fourier transforming the diffraction

intensities. The data is indexed (matching electron density variations to spots), merged and scaled (matching relative strengths of the spots in various images) and combined to give the total electron density (phasing).

1.9.3.1 The phase problem

The 'phase problem' arises from the lack of ability to directly measure the phases of individual diffracted X-ray waves. Intensities (amplitude) are directly measured during data collection; however, phases must be obtained by other means (Scapin, 2013). There are numerous experimental methods for the determination of phase including multiple isomorphous replacement (MIR), multi-wavelength anomalous dispersion (MAD) and the molecular replacement method (MR). Markedly, it is not just determination of crystal structures that will earn one a Nobel prize; The Nobel Prize in Chemistry 1985 was awarded jointly to Herbert A. Hauptman and Jerome Karle "for their outstanding achievements in the development of direct methods for the determination of crystal structures"- i.e. the development of MAD for gathering phase information. Soulimane and colleagues (2000) used MAD phasing techniques to determine the crystal structure of the *ba₃*-cytochrome *c* oxidase from *Thermus thermophilus* to 2.4 Å and revealed the presence of a novel subunit IIa spanning the membrane. This was the first intrinsic membrane protein to be solved by MAD phasing (Soulimane *et al.*, 2000).

Generally, in the absence of any previous phase information the MIR method is the most important method for solving the phase problem, that said, the most popular method is MR. When the structure of the protein is expected to be

closely related to another known protein structure (e.g. high sequence homology) this similarity may be used to obtain preliminary phase information about the unknown structure, which can then be refined (Smyth and Martin, 2000).

MR involves finding out how to orient and position the known structure such that it corresponds with the unknown structure in the crystal by using rotation and translation functions. Using the rotation function the approximate orientation of the two molecules is calculated and secondly, the translation function generates a superimposition of the two molecules. If more than one molecule is present in the unit cell, another rotation-translation function is performed until all molecules are placed into the unit cell. MR operates on the idea of 'borrowing' phases from known structures to determine structure factor of the unknown and calculate the electron density map. There are numerous well established software packages and program suits available for carrying out MR and the dramatic evolution of both computer hardware and software has resulted in faster, more flexible, and, in many cases, fully automated methodologies (Smyth and Martin, 2000; Scapin, 2013).

1.9.3.2 Assessing the quality of the diffracted data

Once an electron density map is prepared, this can be refined to generate the final protein structure. High quality protein structures arise from properly refined, good quality data. It is vital to correctly assign the space group symmetry during collection of the diffracted data as this dictates how much data is needed to solve the structure. In three dimensions, there are seven lattice systems: cubic, hexagonal, tetragonal, rhombohedral, orthorhombic, monoclinic and triclinic,

which are subdivided into 230 space groups, of which proteins can crystallise in 65 (Acharya and Lloyd, 2005). With a high symmetry crystal such as a cubic system, one needs only to collect diffraction data through as little as 35°. Conversely, in a monoclinic lattice, a lower symmetry crystal, data might need to be collected through 180°. That said, generally as much diffractable data as possible is collected to aid in structure determination (Smyth and Martin, 2000).

Final structure model analysis should ideally include analysis of the geometry (chirality, bond lengths and angles) and folding of the main-chain (Ramachandran plot), accompanied by important quality indicators such as resolution, completeness, signal-to-noise ratio (I/σ) and R_{merge} (Wlodawer *et al.*, 2008; Evans *et al.*, 2011). A glossary of the various parameters/quality measures for determination of the quality of diffraction data is detailed in Appendix One. It should be noted that if a structure contains outliers from 'ideal' conformations, it is not necessarily an indicator that the structure is wrong as sometimes unfamiliar conformations are associated with some biological functions (Acharya and Lloyd, 2005).

1.9.4 Crystal structure of antibodies

The three dimensional structure of antibodies has played a central role in elucidating humoral immune response mechanisms, evolution of the antibody repertoire and optimisation of *in vitro*-generated antibodies. Information generated from knowing the structure of antibodies can be exploited for the design of novel diagnostic and therapeutic reagents as well as for development of

immunogens for improved vaccines with greater breadth and efficacy (Lee and Wilson, 2015).

Mutational hotspots and structure-guided mutation strategies can maximise the diagnostic performance of recombinant antibody fragments (Muller *et al.*, 2011; Stura *et al.*, 2011). Muller and colleagues (2011) specifically mutated an anti-fPSA antibody (Fab – 5D3D11) and the results from this study, showed improved diagnostic performance for patients in the ‘diagnostic grey zone’, allowing for better discrimination between prostate cancer and BPH (Muller *et al.*, 2011).

Recent exploitation of the avian immune system has highlighted its use for the generation of high-quality, high-affinity antibodies to a wide range of antigens for a number of therapeutic and biotechnological applications. The principle mechanisms in which chickens generate antibody diversity and the advantages of using chickens for the development of recombinant antibodies is discussed in detail in Chapter three of this thesis. There is however a limited knowledge of the structural attributes of avian antibodies. During a search of the PDB, using the key words ‘single chain fragment, with the refinement of ‘*Gallus gallus*’ (chicken), the vast majority of the structures relate to avian lysozyme predominantly in complex with antibodies from human sources. As of July 2015, there are three chicken scFv fragment crystal structures in the PDB, one of which was deposited from this body of work (Conroy *et al.*, 2014).

Basic knowledge of chicken antibody structures would be of benefit to the scientific community and would enhance our knowledge of antibody structures, allowing for the development of antibodies with unique binding specificities targeting human proteins that can ultimately address various clinically relevant medical issues, such as prostate cancer diagnostics.

1.10 Thesis aims

The costs of unregulated PSA testing in Ireland are considered to be extensive. These include the economic burden of testing, diagnosis and PCa treatment, which fall not only on the health services, but on patients themselves as lost time and out-of-pocket expenses, and on society in the form of lost productivity. Over-diagnosis and over-treatment are widely recognised consequences of inadequate PSA testing (National Cancer Registry, 2015). The vast majority of current literature concerning PCa diagnosis conclude the article with the same statement- better PCa markers are needed that will reduce diagnosis of low risk disease and widespread overtreatment. The aims of this body of work are to use both antibody and glycomic-based approaches to address this research gap through:

- Generation of recombinant avian anti-fPSA and anti-cPSA-specific antibody fragments using phage display technologies.
- Investigation of the kinetic and structural attributes of these avian antibody fragments to enhance their diagnostic capabilities by SPR and X-ray crystallographic methodologies.
- Development of a further understanding the avian immunoglobulin structure by investigating the detailed *N*-glycan profile of IgY polyclonal antibodies using high-throughput *N*-glycan analysis, based on ultra-performance liquid chromatography (UPLC) separation of released glycans.
- Analysis of the *N*-glycome of PCa patient's sera for the identification of a glycomics based signature that can distinguish between indolent and aggressive prostate cancer with high accuracy.

Chapter Two

Materials and Methods

2.1 Materials

2.1.1 Equipment list

Instrument	Supplier
Tecan Safire™ 2 plate reader	Tecan Group Ltd., Seestrasse 103, 8708 Männedorf, Switzerland.
NanoDrop™ ND1000	NanoDrop Technologies, Inc., 3411 Silverside Road, Wilmington, DE 19810, USA.
Gene Pulser Xcell™ electroporator Bio-rad PowerPac Basic Trans-Blot SD semi-dry transfer cell Bio-rad Min-Protean electrophoresis system Gel Doc™ EZ system and Image Lab™	Bio-Rad Laboratories Inc., 1000 Alfred Nobel Drive, Hercules, California 9454, USA.
Balances (Chyo JK-180, Mettler PJ300) Orion 3 Star pH meter	Medical Supply Company Ltd., Damastown, Mulhuddart, Dublin 15, Ireland.
Heraeus Hera-safe laminar flow cabinet	Thermo Scientific, 12-16 Sedgeway Business Park, Witchford, Cambridgeshire CB6 2HY, UK.
Eppendorf refrigerated centrifuge 5810 R Rotors: Fixed angle rotor: F45-30-11 (max. speed 14,000 rpm/ 18, 500 x g) Swing-bucket rotor: A-4-62 (max. speed: 4,000 rpm/3,220 x g)	Eppendorf UK Ltd., Endurance House, Vision Park Histon, Cambridge CB24 9ZR, UK.
Homogeniser (Ultra- Turrax)	IKA-Werke GmbH & Co. KG, 10 Janke & Kunkel-Str., 79219, Staufen, Germany.
HermLe Z233 MK-2 refrigerated centrifuge Rotar- Hermle 220.87 V05 (15,000 rpm/21, 380 x g, 1.5/2 mL tubes)	HermLe Labortechnik GmbH, Siemensstrasse, Wehingen, 78564, Ireland.
Biacore 3000 and A100 (4000)	GE Healthcare Bio-Sciences AB, Danmarksgratan 41, M5, 75184 Uppsala, Sweden.
Stuart Roller Mixer – SRTI	Lennox Laboratory Supplies Ltd., John F. Kennedy Drive, Naas Road, Dublin 12, Ireland.

Instrument	Supplier
37°C static incubator	Sanyo Europe Ltd., 18 Colonial Way, Watford WD24 4PT, UK.
Waters Acquity UPLC H-Class	Waters, Milford, MA, USA.
Waters Ethylene Bridged Hybrid (BEH) Glycan Column	Waters, Milford, MA, USA.
Vydac 301VHP575 7.5 x 50 mm Column	Anachem Ltd., Anachem House, 1 & 2 Titan Court, Laporte Way, Luton, Bedfordshire, LU4 8EF, UK.
2695 Alliance separations module with a 2475 fluorescence detector (For HPLC)	Waters, Milford, MA, USA.
Waters Xevo G2 QTof (YCA 219) with Acquity® UPLC	Waters, Milford, MA, USA.
Hamilton Robotics StarLet liquid-handling platform	Hamilton Medical AG, Bonaduz, Via Crusch 8, CH-7402 Bonaduz, GR, Switzerland.
AKTA FPLC system	GE Healthcare, Silverwater NSW 1811, Australia.
S75 (16/60) analytical gel filtration column	GE Healthcare, Silverwater NSW 1811, Australia.
Clifton stirred water bath	Nickel – Electro Ltd., Oldmixon Crescent, Weston-super-Mare, North Somerset BS24 9BL, UK.

2.1.2 Reagents

All reagents were of the highest quality (analytical grade) available and purchased from Sigma-Aldrich Ireland Ltd., Vale Road, Arklow, Wicklow, Ireland or Fisher Scientific Ireland., Suite 3, Plaza 212, Blanchardstown Corporate Park 2, Ballycoolin, Dublin 15 Ireland, except where otherwise stated.

Reagent	Supplier
Bacteriological Agar Yeast Extract Tryptone	Cruinn Diagnostics Ltd., Hume Centre, Parkwest Business Park, Nangor Road, Dublin 12, Ireland.
DNA Ligase M13K07 helper phage Restriction enzymes dNTP's	Brennan and Company, 61 Birch Avenue, Stillorgan Industrial Park, Stillorgan, Co. Dublin, Ireland.
GoTaq DNA polymerase	Medical Supply Company Ltd., Damastown, Mulhuddart , Dublin 15, Ireland.

Reagent	Supplier
RNA Later RNase Free Tubes RNase Free Tubes Glycogen 3M Sodium Acetate	Ambion/Bio-sciences, 3 Charlemont Terrace, Crofton Road, Dun Laoghaire, Co Dublin, Ireland.
1Kb DNA Hyperladder InstantBlue Protein stain Amintra Ni-NTA	MyBio, Hebron Business Park, Kilkenny, Ireland.
PCR Primers	Integrated DNA Technologies, BVBA, Interleuvenlaan 12A, B-3001 Leuven, Belgium.
fPSA cPSA ACT	SCIPAC/BBI Solutions, Broad Oak Road, Sittingbourne, Kent, ME9 8AQ, UK.
HRP-labelled donkey anti-chicken- IgY- Fab Donkey anti-chicken- IgY-Fab HRP-labelled donkey anti-chicken-IgY- H+L	Gallus Immunotech, 6570 1st Line West Garafraxa, Fergus, ON N1M 2W4, Canada.
Goat anti-PSA pAb Goat- anti-HA tag pAb	Abcam, 330 Cambridge Science Park, Cambridge, CB4 0F, UK.
Platinum Taq High Fidelity DNA polymerase Trizol SYBR TM Safe DNA gel stain	Bio-sciences, 3 Charlemont Terrace, Crofton Road, Dun Laoghaire, Co Dublin, Ireland.

2.1.3 Commercial Kits

Kit	Supplier
NucleoSpin Plasmid DNA purification kit NucleoSpin Gel and PCR Clean-up kit NucleoBond Xtra Midi Plasmid DNA purification kit	Macherey-Nagel GmbH & Co. KG., Neumann Neander Str. 6-8, D-52355 Düren, Germany.
Pierce Thiophilic Adsorption Kit	ThermoFisher /Fisher Scientific Ireland, Suite 3, Plaza 212, Blanchardstown Corporate Park 2, Ballycoolin, Dublin 15, Ireland.
Superscript III RT-PCR Kit	Bio-sciences, 3 Charlemont Terrace, Crofton Road, Dun Laoghaire, Co Dublin, Ireland.
Amine Coupling Kit	GE Healthcare Life Sciences, Amersham Place, Little Chalfont, Buckinghamshire, HP7 9NA, UK.

2.1.4 Bacterial Strains

Strain	Genotype	Supplier
XL1Blue E. coli	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1</i> <i>lac [F proAB lacI^qΔM15Tn10(Tet^r)]</i>	Agilent Technologies Ireland Ltd., Unit 3, Euro, House, Euro Business, Park, Little Island, Cork, Ireland.
TOP10F' E. coli	<i>F{lacI^qTn10(Tet^R)}mcrAΔ(mrr-hsdRMS-</i> <i>mcrBC)Φ80lacZΔM15ΔlacX74recA1araD139Δ(ar</i> <i>a-leu)7697galUgalKrpsLendA1nupG</i>	Invitrogen/ Bio-sciences, 3 Charlemont Terrace, Crofton Road, Dun Laoghaire, Co Dublin, Ireland.

2.1.5 Media Compositions

Media	Components/1L
Luria Broth (LB)	Tryptone 10 g Yeast Extract 5 g NaCl 10 g
Super Broth (SB)	MOPS 10 g Tryptone 30 g Yeast Extract 20 g
2xTryptone Yeast (2xTY)	Tryptone 16 g Yeast Extract 10 g NaCl 5 g
Super Optimal Carbolite (SOC)	Tryptone 20 g Yeast Extract 5 g KCL 2.5 mM NaCl 0.5 g MgCl ₂ 20 mM Glucose 20 mM

2.1.6 Media Additives

2.1.6.1 Antibiotics

Antibiotics used for bacterial cell culture were prepared as described below and sterile filtered through 0.2µm filter before use. Antibiotic stocks were stored at -20°C in aliquots to reduce the risk of contamination.

Antibiotic	Component	Composition
Carbenicillin 100 mg/mL (w/v)	Carbenicillin salt MG H ₂ O	500 mg 5 mL
Kanamycin 70 mg/mL (w/v)	Kanamycin MG H ₂ O	700 mg 10 mL
Tetracycline 5 mg/mL (w/v)	Tetracycline MG Ethanol	50 mg 10 ml

2.1.6.2 Additives

Media Additive	Component	Composition
1M MgSO₄	Magnesium Sulfate	12.036 g
	MG H ₂ O	100 mL
100x 505	Glycerol	50% (v/v)
	Glucose	5% (w/v)
	dH ₂ O (Autoclaved)	Up to 50 mL

2.1.7 Buffers

Buffer	Component	Composition
1x Phosphate Buffered Saline (PBS)	NaCl	5.84 g/1L
	Na ₂ HPO ₄	4.72 g/1L
	NaH ₂ PO ₄	2.64 g/1L
		Final pH 7.2
PBS-Tween 0.05% (PBST)	PBS	1x
	Tween 20	0.05% (v/v)
PBS-Milk 5% (PBSM)	PBS	1x
	Milk Marvel	5% (w/v)
PBST-Milk 1% (PBSTM)	PBST	1x
	Milk Marvel	1% (w/v)

2.1.8 Buffers for sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

10x Electrophoresis Buffer	
50mM Tris (pH 8.3)	30 g
196mM Glycine	144 g
0.1% (w/v) SDS	10 g
dH₂O	Up to 1L

Preparation of 12.5% Separating Gel	1 gel
1M TrisHCl, pH 8.8	1.5 mL
30% (w/v) acrylamide (AcrylaGel)	2.5 mL
2% (w/v) methylamine bisacrylamide (Bis-AcrylaGel)	1.0 mL
dH ₂ O	934 µL
10% (w/v) SDS	30 µL
10% (w/v) Ammonium persulfate (APS)	30 µL
N, N, N', N'-tetramethylethylenediamine (TEMED)	6 µL

Preparation of 4.5% Stacking Gel	1 gel
1M TrisHCl, pH 8.8	300 µL
30% (w/v) acrylamide (AcrylaGel)	375 µL
2% (w/v) methylamine bisacrylamide (Bis-AcrylaGel)	150 µL
dH ₂ O	1.74 mL
10% (w/v) SDS	24 µL
10% (w/v) Ammonium persulphate (APS)	24 µL
N, N, N', N'-tetramethylethylenediamine (TEMED)	2.5 µL

Transfer Buffer	500 mL
Trizma Base	2.4 g
Glycine	7.2 g
Methanol	100 mL
dH ₂ O	Adjust to 500 mL

4x Loading Buffer dye	
0.5M Tris, pH 8.3	2.5 mL
Glycerol	2.0 mL
2-mecaptoethanol	0.5 mL
20% (w/v) SDS	2.5 mL
Bromophenol blue	20 ppm
dH ₂ O	2.5 mL

2.1.9 Protein Purification Buffers

Buffer	Composition/200mL
Osmotic Shock Buffer A (OSB-A)	1XPBS + 150 mM NaCl, pH 8.0
Osmotic Shock Buffer B (OSB-B)	1XPBS + 150 mM NaCl + 1 M Sucrose + 2 mM EDTA, pH 8.0
Periplasmic Buffer	5 mM MgSO ₄ (prepared in dH ₂ O)
Wash # 1	1XPBS + 150 mM NaCl + 10 mM Imidazole, pH 8.0
Wash # 2	1XPBS + 150 mM NaCl + 15 mM Imidazole, pH 8.0
Elution Buffer	250 mM Imidazole in 0.3 M NaCl and 50 mM Sodium Phosphate, pH 8.0

2.2 Methods

2.2.1 General molecular methods

2.2.1.1 Agarose gel electrophoresis

Gels were prepared by weighing out the desired amount to give you the require percentage gel in grams of agarose (i.e. for 1.5% (w/v) gel, 1.5 g agarose is needed). The weighed agarose was added to 100 mL 1X Tris-acetate-EDTA (TAE) buffer and heated in a conical flask in the microwave for 2-3 minutes until completely dissolved. This was left to cool before adding 10 µL SYBRTM safe DNA gel stain. This was then poured into a gel box containing a comb and allowed to cool. Once set, the comb was removed and gel electrophoresis was carried out using the Bio-Rad PowerPacTM in 1x TAE running buffer at 90V for approximately 50 minutes.

2.2.1.2 Ethanol precipitation of DNA

Ethanol precipitation was used to remove any contaminants, to concentrate the nucleic acid and to change the solution in which a DNA sample was suspended. Recommended temperatures and times for ethanol precipitation with sodium acetate range from -80°C to RT, and times range from 10 minutes to overnight. Ethanol precipitation was carried out at -20°C overnight (o/n) or at -80°C for two hours. To a 1.5 mL tube containing DNA the following was added: 0.1x volume of 3 M sodium acetate (NaOAc), pH 5.2, 2x volumes of ice cold absolute ethanol and 1 µL glycogen (5 mg/mL). The tubes were mixed gently by inverting several times and then placing at the desired temperature for the required period of time. The

precipitates were centrifuged at 19,500 x *g* for 30 minutes at 4°C and the pellet was washed with 200 µL ice cold 70% (v/v) EtOH and again centrifuged at 19,500 x *g* for 5 minutes at 4°C. The pellet was then air dried and resuspended in an appropriate volume of MG H₂O.

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

The separation of macromolecules in an electric field is called electrophoresis. The most commonly used method for separating proteins by electrophoresis uses a polyacrylamide gel as a support medium and sodium dodecyl sulfate (SDS) to denature the proteins, and was originally described in 1970 (Laemmli, 1970). SDS is an anionic detergent and has a net negative charge within a wide pH range. A polypeptide chain binds amounts of SDS in proportion to its relative molecular mass. The protein sample is mixed with SDS which destroys the complex structure of proteins (secondary and non-disulphide linked tertiary structures) and is strongly attracted toward an anode (positively-charged electrode) in an electric field. Heating the protein sample to 95°C further promotes denaturation, allowing SDS to bind.

Polyacrylamide gels restrict larger molecules from migrating as fast as smaller molecules. As the charge-to-mass ratio is approximately the same amongst SDS-denatured polypeptides, the final separation of proteins is dependent almost entirely on the mass differences. If proteins of known mass are run simultaneously with an unknown sample the, masses of unknown proteins can be estimated (Schägger and von Jagow, 1987).

Assessment of protein quality and protein purifications were routinely assessed by SDS-PAGE and Western Blot (WB) analysis using the buffers described in section 2.1.8.

Initially, the casting apparatus was set up as per the manufacturer's guidelines. The mixture for the 12.5% separation gel and 4.5% stacking gel were prepared as detailed in section 2.1.8. Polymerisation occurs on addition of APS and TEMED, hence is not added until the gel was ready to be cast. The 12.5% gel is responsible for separating the polypeptides and was cast first by adding the APS and TEMED and immediately pouring the mixture between two clean glass plates and covering with a layer of isopropanol to promote more rapid polymerisation of the gel. Once the gel polymerises, the isopropanol layer was removed by inverting the apparatus. APS and TEMED was then added to the 4.5% stacking gel gently poured on top of the polymerised 12.5% gel, and a comb was inserted to make wells in preparation for loading of the protein samples.

Once this gel fully polymerised, the plates were placed into the electrophoretic chamber, the chamber was filled with 1x electrophoresis buffer and the comb was removed. The samples to be assessed were prepared by adding 4x loading buffer dye and/or dH₂O to a final concentration of 1x and heating for 5 minutes at 95°C, to further promote denaturation. A volume of 10 µL sample was added to each well and 6 µL PageRuler™ prestained protein ladder was added to the first well. The gels were resolved at 160V until the tracker dye (bromophenol blue) almost reached the bottom of the gel, typically taking 55 minutes. The gels are then removed from the casting apparatus and stained with InstantBlue™ for 10-30 minutes and washed with dH₂O.

2.2.1.4 Western Blotting

Western blotting is a commonly used technique in cell and molecular biology that allows for the identification of specific proteins from a complex mixture of proteins extracted from cells or tissue homogenate. There are three key steps involved in western blotting, namely (1) separation by size (SDS-PAGE), (2) transfer to a solid support, and (3) marking target protein using a specific primary and labelled secondary antibody for visualisation (Mahmood and Yang, 2012).

Initially, proteins were separated by SDS-PAGE as described above (section 2.2.1.3) and once the gel was resolved, it was left to soak in transfer buffer, along with adsorbent tissue paper (4 pieces cut to gel dimensions) and a nitrocellulose membrane (also cut to gel dimensions), for 10 minutes. A sandwich consisting of 2 layers of adsorbent tissue paper, nitrocellulose membrane, SDS-PAGE gel and 2 further layers of adsorbent tissue paper was prepared as shown in Figure 2.1. Air bubbles were removed by carefully rolling each of the layers with a plate spreader. The separated proteins were then transferred to the nitrocellulose membrane using the Trans-Blot® SD Semi-Dry Transfer System (Bio-rad) at 15V for 21 minutes.

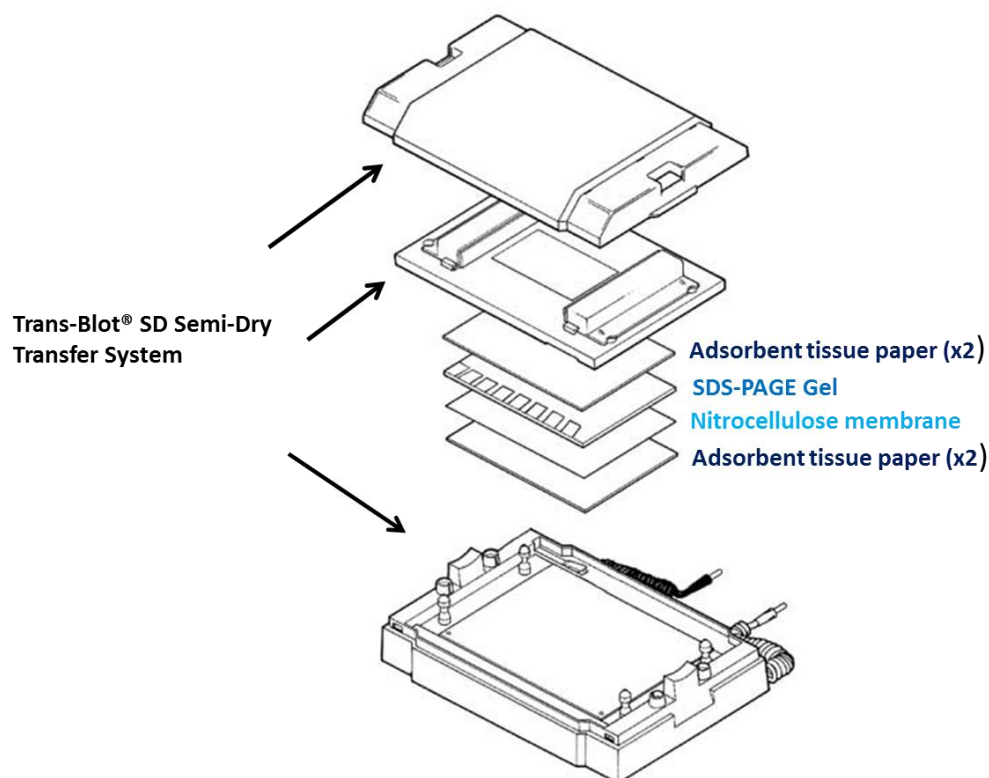


Figure 2.1: *Semi-Dry transfer system.* This system allows for rapid transfer of separated proteins from an SDS-PAGE gel to a nitrocellulose membrane which can then be probed for detection of specific protein. (Figure adapted from Biorad, 2015.)

The nitrocellulose membrane was then blocked for non-specific binding in 20 mL PBST in a large weigh boat for 2 hours at room temperature (RT) or overnight at 4°C with agitation. The membrane was washed briefly three times with PBST and three times with PBS. The membrane was then incubated with 10 mL of 1 in 2,000 dilution of a primary antibody diluted in PBST for 1 hour at RT in a large weigh boat with gentle agitation. The antibody solution was discarded and the membrane washed as before. Next, the membrane was incubated with a specific relevant enzyme labelled secondary antibody (1 in 2,000 dilution), prepared in 10 mL PBST, for 1 hour at RT, with gently agitation. The antibody solution was discarded and the membrane was washed as before. The development of the

specific reaction was achieved by adding (5-10 mL) of the liquid enzyme substrate, 3, 3', 5, 5' -tetramethylbenzidine (TMB) and was stopped by washing with dH₂O (20-30 mL).

2.2.1.5 Preparation of bacterial cell stocks

Cell stocks were prepared in either 1.5 mL tubes or 96-well culture plates and stored at -80°C.

96-well plates: Individual colonies were picked and grown in SB media supplemented with the appropriate antibiotic(s) overnight at 220 rpm at 37°C in a sterile 96 well plate. A sample of 30 µL was taken from each well of the overnight cultures and was inoculated into 370 µL fresh 2xTY media for expression. The 30 µL sample taken was replaced with 30 µL 80% (v/v) glycerol and the plates kept at -80°C for storage.

1.5 mL tubes: An overnight culture was prepared by inoculating 5 mL SB (supplemented with the relevant antibiotic(s)) with 5 µL clone. The cultures were grown overnight at 37°C while shaking at 220 rpm. The cells were collected by centrifugation at 5,500 x *g* for 30 minutes at 4°C. The supernatant was discarded and the cells were resuspended in SB containing 20% (v/v) glycerol and snap frozen in liquid nitrogen and stored at -80°C.

2.2.1.6 Buffer exchange and protein concentration

Protein samples sometimes need to be buffer exchanged to a solution more suitable for long-term storage. Proteins were concentrated and buffer exchanged

using Vivaspin columns (GE healthcare) with an appropriate molecular weight 'cut-off' (Mr 3,000 to 100,000, depending on protein size). A 5 mL sample was concentrated and buffer exchanged against PBS/0.02% (w/v) NaN₃ using a Vivaspin column with a 10,000 molecular weight 'cut-off'. The 6 mL concentrator was washed with 5 mL water by centrifugation at 5,500 x *g* 4°C for ~5 minutes and then washed twice more in the same manner with 5 mL sterile PBS. The protein sample was added to the column and buffer exchanged by centrifuging the sample (5,500 x *g*, 4°C) until the volume reached ~0.5 mL. It was then refilled with sterile PBS/0.02% (w/v) NaN₃ to 5 mL and centrifuged as above until the volume reached ~0.5 mL. This process was repeated 3 times. On the last spin, the volume was brought to 1 mL and the concentrated protein was pipetted from the surface of the Vivaspin column and stored at -20°C in aliquots of 20 µL. Protein was quantified using the NanoDrop (A_{280nm}).

2.2.2 Screening and analysis of an existing avian anti-fPSA library

Anti-fPSA scFv and Fab clones from a previously built avian anti-fPSA library (generated by Dr. Conor Hayes) were screened for the identification of a high affinity anti-fPSA antibody. Briefly, an adult leghorn chicken was immunised with the fPSA molecule according to protocols described by Barbas *et al.* (2001). The animal was sacrificed, the spleen harvested and the RNA extracted with Trizol reagent using a well established protocol (see section 2.2.3.3).

The amplified antibody variable genes were assembled by splice by overlap extension polymerase chain reaction and cloned into the pComb3X vector for subsequent panning (Barbas *et al.*, 2001). Anti-fPSA specific clones were isolated

from a large library (3.10×10^8 cfu/mL) following three rounds of selection from the antibody repertoire displayed on the surface of filamentous phage by immobilised panning against fPSA.

2.2.2.1 Identification of anti-fPSA specific antibody fragments by monoclonal ELISA and Biacore 4000

Glycerol stocks plates of anti-fPSA scFv and Fab clones were used for inoculation in 150 μ L SB supplemented with 50 μ g/mL carbenicillin in sterile 96-well culture plates. The plates were cultured overnight at 37°C while shaking at 220 rpm. The following day, 30 μ L of this overnight culture was infected into 370 μ L fresh SB media supplemented with 50 μ g/mL carbenicillin, 1 mM MgSO_4 (and 1x505 (0.5% (v/v) glycerol & 0.05% (v/v) glucose), in sterile deep-well 96-well plate. The plates were shaken at 220 rpm at 37°C, until an optical density (OD_{600}) of ~ 0.600 was reached. A final concentration of 0.1 mM IPTG was added to each well and the plate was left shaking at 220 rpm overnight at 30°C. An ELISA plate was coated overnight at 4°C with 100 μ L/well 0.5 μ g/mL fPSA and subsequently blocked at 37°C for 1 hour with 200 μ L/well PBSM. The plate was then washed three times with 200 μ L/well PBST, followed by three times with PBS.

The antibodies were extracted from the periplasmic space of the cell by placing the plates at -80°C for 10 minutes and then thawing at 37°C for 20 minutes. One hundred microliters/well of 1 mg/mL lysozyme/PBS was added to disrupt the cell wall. The plate was then subject to two successive freeze-thaw cycles. This commonly used method releases periplasmic proteins by bursting the *E. coli* cells. The cell extracts were cleared by centrifugation at 5,500 x g for 30 minutes and the resultant lysate was diluted 1 in 2 in PBSTM and 100 μ L of this lysate was then

added to the wells of the ELISA plate and incubated for 1 hour at 37°C. The plate was washed as before and the fPSA-specific scFv/Fab clones were then detected with 100 µL/well with a HRP-labelled goat anti-hemagglutinin tag antibody (anti-HA) diluted (1 in 2,000) in PBSTM for 1 hour at 37°C. The plate was washed as before and 100 µL/well TMB substrate enzyme solution added to each well. The plate was incubated at RT for approximately 5 minutes to allow chromophore development, after which the reaction was stopped by addition of 50 µL/well 10% (v/v) HCl. The absorbance was read at 450 nm using the Tecan Safire™ plate reader.

Clones with an absorbance value greater than 0.3 were taken further for kinetic analysis using the Biacore™ 4000, this work was carried out in collaboration with Dr. Paul Conroy. Briefly, the clones were screened using a previously immobilised CM5 dextran chip (Conroy *et al.*, 2012) functionalised with approximately 6,000 RU of anti-HA epitope by EDC-NHS coupling chemistry over spots 1, 2, 4 and 5. Anti-fPSA scFv and Fab lysates were diluted 1:3 (in duplicate) in HBPS-EP⁺ and analysed in a capture approach over the anti-HA surface at 25°C. The diluted lysates were captured on spots 1, 2, 4 and 5 and two concentrations of fPSA (10 nM and 50 nM diluted in HBS-EP⁺) were then passed over the captured antibody fragments and referenced against the unmodified surface on spot 3. The data was evaluated with BiaEvaluation software using the '2 over 2' template and curves were fitted with a 1:1 Langmuir binding equation.

2.2.2.2 Optimisation of expression of scFv B8

The gram negative bacterium *Escherichia coli* is the most widely used host for the expression of recombinant proteins. It is fast and inexpensive to test a wide variety of possible strategies in *E. coli*, its genetics are well characterised and numerous mutant strains are available. Expression in *E. coli* requires 4 key elements; (1) protein of interest (2) a bacterial expression vector (3) an expression cell line, and (4) equipment/materials (i.e. Shaking Incubator, media, etc.). Within these fundamental elements however, there are multiple questions that must be answered – which expression vector should be used? What strain(s) should be chosen? Should the protein of interest be tagged? What affinity tag is the best? What temperature is optimal for induction? Many steps in the production of recombinant proteins must be optimised and there is no ‘right’ answer to any of the above questions, as every protein is different and the ‘right’ expression strategies and purification protocols must be worked out for each individual protein (Francis and Page, 2010).

An overview of the optimisation of expression of the anti-fPSA scFv clone B8 is detailed in Figure 2.2. Briefly, inoculation was carried out in 5 mL SB supplemented with 50 µg/mL carbenicillin with 2 µL of a glycerol stock of scFv B8 and the culture was then grown overnight at 37°C while shaking at 220 rpm. This was then sub-cultured into 50 mL fresh SB media supplemented with 50 µg/mL carbenicillin, 1 mM MgSO₄ and 1x 505 in a 250 mL flask and grown at 37°C until OD₆₀₀ of 0.600 was reached. The culture was then induced. Various concentrations of IPTG were assessed at both 25°C and 30°C. Over different time points (ranging from 2 hours to overnight), 1 mL samples were taken and

centrifuged and the pellet was retained at -80°C for further analysis. The pellets were lysed, as previously described, and loaded onto SDS-PAGE gels for analysis.

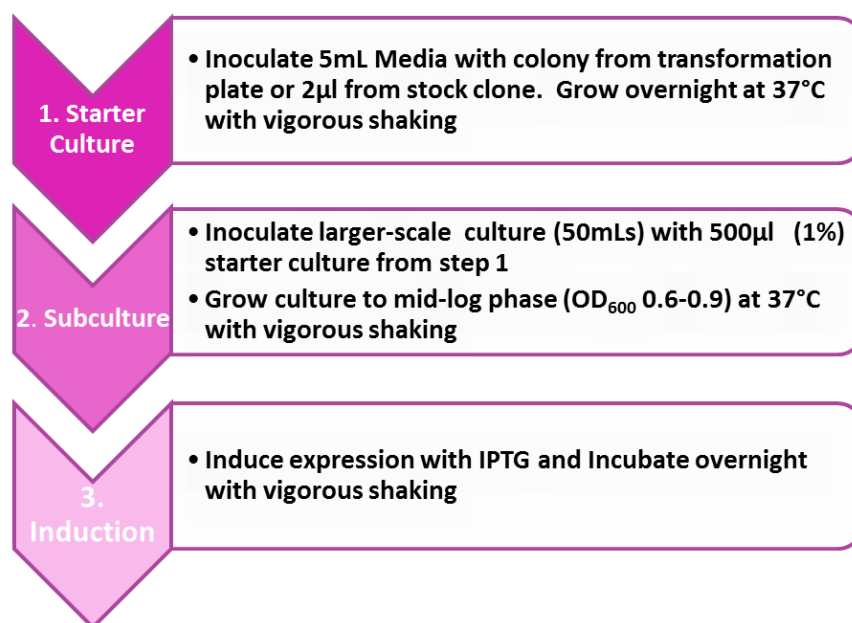


Figure 2.2: Overview of optimisation of expression of recombinant proteins in *E. coli*.

Optimal conditions must be assessed for expression of each individual recombinant antibody clone to identify the best expression strategies. There are 3 basic steps involved in expression of recombinant proteins. A small overnight culture was prepared and grown o/n at 37°C. This was then used to inoculate a fresh starter culture which was cultured until mid-log phase when it was induced. Expressions studies must be carried out where temperature, time and inducer concentration are rigorously optimised.

2.2.2.3 Purification of scFv B8 by Immobilised Metal Affinity Chromatraphy (IMAC)

The DNA sequence specifying a string of six to nine histidine residues is frequently used in vectors for production of recombinant proteins. Recombinantly expressed His-tagged proteins can be purified by a single-step affinity chromatography called immobilized metal ion affinity chromatography (IMAC), which is commercially

available in different kinds of formats, Ni-NTA matrices being the most widely used (Spriestersbach *et al.*, 2015).

Supports such as beaded agarose or magnetic particles can be derivatized with chelating groups to immobilize the desired metal ions, which then function as ligands for binding and purification of biomolecules of interest. This forms the basis IMAC. The chelator most commonly used as ligands for IMAC is nitrilotriacetic acid (NTA). Once the NTA-agarose resin is prepared, it can be "loaded" with the desired divalent metal (e.g., Ni, Co, Cu, Fe). Using nickel as the example metal, the resulting affinity support is usually called Ni-chelate or Ni-NTA resin. IMAC has many advantages including strong, specific binding, mild elution conditions and the ability to control selectivity by including low concentrations of imidazole in chromatography buffers (Sulkowski, 1985; Gräslund *et al.*, 2008).

An overnight culture of scFv was prepared by inoculating 2 μ L glycerol stock scFv into 1 mL SB media with 50 μ g/mL carbenicillin, with shaking at 220 rpm at 37°C. This culture was sub-cultured into five 1 L flasks, each containing 200 mL SB, 50 μ g/mL carbenicillin, 1 mM MgSO₄ and 1x505. The culture was grown at 37°C until an OD₆₀₀ of 0.600 was reached and it was subsequently induced with 0.1 mM IPTG at 25°C for 4 hours. The expressed cultures were centrifuged at 5,500 x *g* for 45 minutes 4°C (four 250 mL bottles) and the supernatant was discarded and the pellets frozen at -80°C.

The following day, the pellets were thawed at 37°C for 10 minutes (in the water bath) and each pellet was then resuspend thoroughly in 7.5 mL *OSB-A* and transferred to two 85 mL Oakridge tubes (i.e. 15 mL in each bottle). *OSB-B* (15 mL)

was added and the tubes were incubated for 10 minutes at RT on a roller. The shocked cells were collected by centrifugation at 14,000 x *g* (4°C) for 20 minutes. The supernatant was discarded and pellets resuspended in 15 mL periplasmic buffer. Samples were placed on ice for 10 minutes, with gently mixing by inverting the tubes every 2 minutes to ensure the solution was homogenous.

The cell debris was collected by centrifugation at 14,000 x *g* (4°C) for 20 minutes. The supernatant was pooled in a 50 mL tube and the protein was sequentially filtered through a 0.45 µm and 0.2 µm filters. A 100 µL sample of the filtered lysate was taken to use for SDS-PAGE analysis. Ni-NTA nickel resin slurry (2 mL) was added to a chromatography column and allowed to settle. The resin bed was washed with 10 mL PBS twice (20 mL total) followed by 10 mL of *OSB-A* buffer. The lysate was then applied to the column, collecting the 'flow-through' in a 50 mL tube (collecting a 100 µL sample and labelling it F1- i.e. 'flow through' 1). This process was repeated (collecting a 100 µL sample again-FT2). The column was then washed with 10 mL of wash buffer #1 (1XPBS + 150 mM NaCl + 10 mM Imidazole, pH 8.0) collecting in a 50 mL tube and again sampling (100 µL- W1). The retained protein was then eluted by adding 5 mL elution buffer (250 mM Imidazole in 0.3 M NaCl and 50 mM Sodium Phosphate, pH 8.0) to the column and collecting (taking 100 µL sample – EP (eluted protein)). The eluted sample (5 mL) was then concentrated and buffer exchanged against PBS/0.02 % (w/v) NaN₃ using a Vivaspinn column with a 10,000 molecular weight 'cut-off' level, as described in section 2.2.1.6.

2.2.2.4 Anti-fPSA scFv B8 Sandwich ELISA Limit of Detection (LOD) Assay

The anti-fPSA scFv B8 Limit of Detection (LOD) was determined in accordance with the Clinical and Laboratory Standards Institute's (CLSI) published guidelines "EP17, Protocols for Determination of Limits of Detection and Limits of Quantification" (Armbruster and Pry, 2008). This ELISA was carried out 24 times to ensure the accuracy of the result.

A 96 well Maxisorb plate was coated with 100 μ l/well 1 μ g/mL goat anti-PSA polyclonal antibody diluted in PBS and incubated for 1 hour at 37°C. The plate was then blocked with PBST (200 μ l/well) and incubated for 1 hour at 37°C. The wells were then washed three times with PBST and three times with PBS (200 μ l/well). One hundred microliters per well of fPSA in varying dilutions (serial dilutions from 32 to 0 ng/mL) was then added to the plate and again incubated at 37°C for 1 hour. The plate was washed as before and anti fPSA scFv B8 was then added in PBST (100 μ l/well) at a 1:1000 dilution and incubated for 1 hour at 37°C. The plate was then washed as before. A 1 in 2,000 dilution of the secondary antibody, HRP-labelled anti-HA, diluted in PBST (100 μ l/well) was added and incubated as before. The plate was washed as before and 100 μ L/well TMB substrate solution was added to each well. The plate was incubated at RT for approximately 5 minutes to allow chromophore development, after which the reaction was stopped with 50 μ L/well 10% (v/v) HCl. The absorbance was read at 450 nm using the Tecan SafireTM plate reader.

2.2.2.5 Kinetic analysis of scFv B8 using Biacore 3000

In the past decade surface plasmon resonance (SPR)-based optical biosensors, such as Biacore systems have evolved as standard tools for antibody characterisation and binding interaction studies. The Biacore 3000 instrument uses SPR technology for biomolecular interaction analysis in 'real-time' without the need for labels. This instrument provides a comprehensive characterisation of binding events and kinetics analysis can also be conducted with relative ease (Wilson, 2002; Leonard *et al.*, 2011).

2.2.2.5.1 Instrument /Chip preparation

The Biacore 3000 was cleaned prior to the study using a 'super desorb' program at 25°C. This is carried out to ensure that the instrument is completely free of any residual proteins, for optimal performance. A maintenance chip was docked and the instrument was primed five times with 0.05% (w/v) SDS, followed by priming once with ultrapure water, and, finally, five times with 50 mM glycine-NaOH, pH 9.5. The maintenance chip was replaced with a research-grade carboxymethylated (CM) dextran sensor chip CM5 sensor chip, and primed three times with running buffer. For this study sterile filtered and degassed HEPES-buffered saline (HBS, pH 7.4: 10 mM 4-(2-hydroxyethyl) piperazine 1 ethanesulfonic acid N-(2-hydroxyethyl) piperazine N'-(2-ethanesulfonic acid)(HEPES), 150 mM NaCl, 3 mM EDTA (ethylenediaminetetraacetic acid) and 0.05% [v/v] Tween 20) as both running and sample preparation buffer. The CM5 chip was then preconditioned with two consecutive injections each of 50 mM NaOH and 0.05% (w/v) SDS for 10 seconds at a flow rate of 100 µL/minute.

A preconcentration study was carried out by passing over 25 µg/mL anti-HA polyclonal antibody in 10 mM NaOAc over a range of pH values to determine the optimal pH to promote efficient immobilisation. The pH which demonstrates the highest RU level of antibody onto the underivatized surface was employed for the immobilisation of the polyclonal antibody to the surface, using amine-coupling chemistry. The surface chip was activated by passing over equal volumes of 400 mM EDC (1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride) and 100 mM NHS (N-hydroxysuccinimide) at a flow rate of 10 µL/minute for 7 minutes. Anti-HA polyclonal antibody was diluted in 10 mM NaOAc to 25 µg/mL, pH 4.4 was passed over the activated surface at a flow rate of 10 µL/minute for 15 minutes. Unreactive NHS groups were capped and deactivated and non-covalently bound antibody was removed by addition of 1 M ethanolamine-hydrochloride solution at a flow rate of 10 µL/minute for 7 minutes. Post-conditioning of the surface was then carried out by pulsing the surface 5 times with 15 µL injections of 20 mM NaOH.

2.2.2.5.2 Kinetics analysis of anti-PSA scFv B8 using Biacore 3000

Kinetic analysis of avian anti-PSA clone B8 was performed using the kinetic wizard functionality and data analysis was evaluated using BIAevaluation 4.1.1 software. Initially, the appropriate dilution for capture of approximately 100 RU of scFv B8 was assessed. Purified scFv B8 was diluted 1 in 3,000 in running buffer and passed over the anti-HA surface at a flow rate of 30 µL/minute for 3 minutes. After capturing the scFv, fPSA was passed over the surface in varying concentrations (12.5, 6.25, 3.12 (in duplicate), 1.6, 0.8, 0.4 and 0 nM) for 3 minutes, and allowed to dissociate for 12 minutes. Regeneration of the surface was carried out by

passing over 20 mM NaOH for 10 seconds at a flow rate of 30 μ L/minute. The data obtained was evaluated using the kinetic evaluation model of BiaEvaluation software version 4.1.1. The dataset from the reference flow cell (underivatised surface) was subtracted to remove any systemic effects, and the 0 nM fPSA concentration binding response (buffer only) was subtracted from each antigen response, thus double referencing subtracting the dataset. Kinetic constants were modelled on the curves with 1:1 Langmuir binding model.

2.2.2.6 Crystal Structure of scFv B8

The interaction of antibodies with their cognate antigen involves a multitude of non-covalent interactions between the epitope (binding site on the antigen) and the paratope (binding site on the antibody). An indepth understanding of the molecular structure of antibodies and antigens can provide insight into the mechanisms of molecular recognition and therefore assist in the rational design of antibodies (Finlay and Almagro, 2012; Sela-Culang *et al.*, 2013).

Through a collaboration with Dr Paul Conroy and Prof. James Whisstock (Monash University, Melbourne, Australia), the crystal structure of scFv B8 was solved, as described by Conroy *et al.*, (2014). Briefly, scFv B8 was expressed, purified and subsequently concentrated to 5.8 mg/ml, with a 3-kDa concentrator (Merck-Millipore). Crystallization was carried out by the hanging drop method with a 1:1 mixture of protein and mother liquor at 4 °C, where crystals formed in 3–7 days. Initial crystal forming conditions were identified from the Basic Crystallization kit for Proteins (Sigma). Crystals of scFv B8 were obtained in 0.1 M NaOAc, pH 4.6, 0.2 M ammonium sulphate. The crystals were flash-cooled in liquid nitrogen using 25% (v/v) glycerol as the cryoprotectant.

Data sets were collected at the Australian Synchrotron MX2 beamline at 100K. The data was merged and processed using XDS , POINTLESS, and SCALA. Molecular Modelling and replacement was carried as described (Conroy *et al.*, 2014).

2.2.2.7 Site-Directed Mutagenesis of scFv B8

Mutant variants of scFv B8 were generated in order to develop a greater understanding of the structural and binding properties of this antibody fragment. Plasmid DNA was prepared by inoculation of scFv B8 in 10 mL SB media supplemented with 50 µg/mL carbenicillin. The culture was grown o/n at 37°C with shaking at 220 rpm. The overnight culture was centrifuged at 5,500 x *g* for 30 minutes at 4°C, the supernatant was discarded and the plasmid was purified using the Nucleospin plasmid prep, as per the manufacturer's guidelines. Plasmid DNA was eluted in 40µL MG H₂O and quantified with the NanoDrop spectrophotometer.

scFv B8 plasmid DNA was then used as a template to introduce a cysteine to serine double mutant at C97 and C100 in CDRH3 (C97/100S; scFv B8 Ser MT) and a cysteine to alanine double mutant at C97 and C100 in CDRH3 (C97/100A; scFv B8 Ala MT). These site specific mutations were generated by polymerase chain reaction (PCR) site directed mutagenesis (SDM) using KOD Hot Start DNA polymerase (Quickchange mutagenesis), with the following forward and reverse primers;

scFv B8 Ser MT(C97/100S) Forward:

5'GATCTCATAGCAGTGGTAGCCGGAATGCTGCTCTCATC3'

scFv B8 Ser MT(C97/100S) Reverse:

5'GATGAGAGCAGCATTCCGGCTACCACTGCTATGAGATC3'

scFv B8 Ala MT(C97/100A) Forward:

5'GATCTCATGCCAGTGGTGCCCGGAATGCTGCTCTCATC3'

scFv B8 Ala MT(C97/100A) Reverse:

5'GATGAGAGCAGCATTCCGGGCACCACTGGCATGAGATC3'

The PCR reaction was set up according to Table 2.1 under the conditions outlined in Table 2.2

Table 2.1: Reaction set up for site-directed mutagenesis of scFv B8

Component	Test (μl)	Control (μl)
Buffer	5	5
MgCl ₂	3	3
dNTP's	1	1
Fwd	1	NA
Rev	1	NA
DNA (50 ng)	1	1
Polymerase	0.2	0.2
dH ₂ O	37.8	39.8
Final Volume	50μL	

Table 2.2: PCR conditions for site-directed mutagenesis of scFv B8

Stage	Step	Temp (°C)	Time (sec)	Number of Cycles
1.	Initial Denature	95	30	1
2.	Denature	95	30	16
	Annealing	56	60	
	Extension	72	420	

Before transformation the original plasmid, which served as a template must be removed from the reaction to minimise background. This was carried out using the restriction digestion enzyme *Dpn* I. To the PCR product, 1 µL of Cut Smart Buffer (10x) and 1 µL *Dpn* I restriction enzyme (10 U/µL) was added and incubated at 37°C for approx. 6 hours. The reaction was then deactivated following incubation at 65°C for 10 minutes. PCR clean-up was performed as per manufacturer's guidelines: Promega, Wizard® SV Gel and PCR Clean-Up System and the DNA was eluted in 20 µL MG H₂O.

Transformation of *E. coli*. was carried out using XL1 blue competent cells for uptake of the PCR product. Two microliters of DNA was added to 50 µL XL1 blue competent cells and this mix was left on ice for 30 minutes. The cells and DNA were then subjected to heat shock for 45 seconds at 42°C followed by cooling on ice for 5 minutes. The transformed material was then added to 100 µL pre-warmed 2xTY media to the tube and allowed to outgrow by incubating for 1 hour at 37°C while shaking at 220 rpm. This mix (150 µL) was then plated out on LB/Carb plates and incubated o/n at 37°C. The control sample was treated in the same manner.

Four clones were picked and grow o/n in 3 mL 2xTY supplemented with 1:1,000 dilution of carbenicillin. The overnight culture was centrifuged at 5,500 x g for 30 minutes at 4°C, the supernatant was discarded and the plasmid was purified using the Nucleospin plasmid prep as per the manufacturer's guidelines. Plasmid DNA was eluted in 40 µL MG H₂O and subsequently sent for sequencing to ensure that the correct mutations were introduced. Sequence analysis was carried out by Source Bioscience Ltd, Dublin, Ireland. The DNA sequences obtained were

translated using the translate tool from Expasy (<http://web.expasy.org/translate/>) and the amino acid sequences were then exported to PowerPoint.

2.2.2.8 Expression and Purification of scFvB8 Ser MT by Immobilised Metal Affinity Chromatography (IMAC) and Gel Filtration

Anti-fPSA scFv B8 Ser MT was expressed and purified according to methods described in Conroy *et al.*, (2014). Briefly, single colonies were selected from LB-Agar supplemented with 25 µg/mL carbenicillin and grown overnight in 5 mL SB supplemented with 25 µg/mL carbenicillin and 1% (w/v) glucose at 37°C, with shaking at 220 rpm. This starter culture was used to inoculate 100 mL SB with 25 µg/mL carbenicillin and was grown to O.D₆₀₀ = 0.6-0.8 before sub-culturing into 10 x 500 mL SB with 25 µg/mL carbenicillin (2 L flasks). At O.D₆₀₀ = 0.6-0.8 the cultures were induced with 0.2 mM IPTG, at 30°C with shaking at 230 rpm o/n.

The bacteria were harvested by centrifugation (3, 220 x *g*) at 4°C for 20 minutes. Soluble scFv were released from the periplasmic space by osmotic shock in a two-step process. The pellet was first thoroughly resuspended in 1X TBS (25 mM Tris, pH 8.0, 150 mM NaCl) and an equal volume of 2X shock buffer (50 mM Tris, pH 8.0, 300 mM NaCl, 1 M sucrose and 2 mM EDTA) was added before incubation at RT for 15 minutes. Shocked cells were recovered by centrifugation at 12, 400 x *g*, 4°C for 20 minutes, followed by resuspension in ice-cold 5 mM MgSO₄ and incubation on ice for 15 minutes.

The periplasmic-stripped cells were collected by centrifugation (27, 200 x *g* at 4°C for 20 minutes) and 0.2 times the volume of 5X binding buffer (125 mM Tris, pH 8.0, 750 mM NaCl, 50 mM imidazole, 0.02% (w/v) NaN₃) was added to the

supernatant. HisBind (Novagen) resin (1 mL equilibrated in 30 mL 1X binding buffer) was added and scFv recovered by batch binding overnight at 4°C on an end-over-end roller. The resin was collected by gravity flow and washed with 30 mL binding buffer followed by a second wash with 30 mL wash buffer (25 mM Tris, pH 8.0, 150 mM NaCl, 20 mM imidazole, 0.05% (v/v) Tween® 20, 0.02% (w/v) NaN₃). Bound protein was eluted with 10 mL elution buffer (1X running buffer with 300 mM Imidazole) in 1 mL fractions.

Eluted protein fractions from the IMAC purification were further purified by chromatography on a size exclusion column. Gel filtration (also referred to as size exclusion chromatography) is a basic chromatography technique that separates proteins based on differences in molecular size and shape as they pass through a gel filtration medium in a packed column, and it is often used as a final polishing step following IMAC purification for isolation of highly pure protein (Kim *et al.*, 2011).

An S75 (16/60) gel filtration column (GE Healthcare) was equilibrated with 25 mM Tris pH 7.4, 150 mM NaCl, 0.02% (w/v) NaN₃ (1CV = 50 mL). Protein-containing fractions (eluates 2-6) of scFv B8 Ser MT (26 mg/mL) previously purified by IMAC were applied to the column at a flow rate of 1 mL/minute. Fractions were collected after 0.4CV until 1CV was reached. Protein was monitored online by A₂₈₀ using the dedicated Unicorn control and evaluation software (GE Healthcare). Protein peaks were analysed by SDS-PAGE.

2.2.2.9 Determination of the crystal structure of scFv B8 Ser MT

Binding capabilities of antibodies are highly diverse given the structural similarities between most antibodies (Sela-Culang *et al.*, 2013). However, in contrast to the mechanism employed by humans, mice and primates, the V-gene repertoire of the chicken does not utilise highly diverse sequences resulting in the generation of unique canonical classes (Finlay and Almagro, 2012). A limited number of avian antibody crystal structures are available in the Protein Data Bank (PDB) which has restricted our understanding of the structural consequences of uniquely chicken features. Investigation of the structural properties of mutant variants of scFv B8 by X-ray crystallography will aid in the understanding of the binding mechanisms employed by chicken antibodies.

2.2.2.9.1 Coarse screening for optimal crystal formation conditions

The process of growing protein crystals has been carried out for well over 150 years. Initially, identification of crystal-forming conditions occurred empirically. However, this has now moved towards more rationality as a result of a better understanding of the physical chemistry of protein crystal growth. Factors such as the solubility of the protein, source and amount of protein, the type of salts used to induce super-saturation, temperature and need for additives are of great importance for identification of conditions that will sustain crystal growth (Giegé, 2013). Screening is now simplified and made easier through the use of commercially available screening kits. These kits provide crystallisation conditions that included buffers with pHs ranging from pH 4.2 to pH 10.5, polymers (PEG 400 to PEG 8,000), salts (such as sodium chloride, ammonium sulphate, calcium, zinc,

and magnesium), alcohols (such as ethanol, isopropanol, 1,4-butanediol, pentaerythritol propoxylate, and 2-methyl-2,4-pentanediol), organics (such as dioxane, polyacrylic acid, and Jeffamine), and many other additives that promote crystal growth (Kim *et al.*, 2011).

Coarse screening was carried out in a 24-well plate format. The antibody fragment was concentrated to 16.12 mg/mL, since a very high protein concentration is necessary for setting crystallization trials. A total of 240 conditions were set using the following kits; Crystal Screen, 2 x 24-well plates (Hampton Research), Crystal Screen II, 2 x 24-well plates (Hampton Research), PEG/Ion, 2 x 24-well plates (Hampton Research), and The JCSG Suite, 4 x 24-well plates (Qiagen). The 24-well plates were prepared by application of a ring of petroleum jelly to the circumference of each well and adding 500 µL of mother liquor (one condition per well) from each kit. A 1 µL drop of mother liquor was placed in the centre of the siliconised glass slides (Hampton Research) and 1 µL of anti-PSA scfvB8 Ser MT (at 16.12 mg/mL) was mixed into the mother liquor drop. A control drop was included for each condition in which a 1 µL drop of mother liquor was mixed with 1 µL buffer and placed in the centre of the siliconised glass slides below the protein drop. The slide was then carefully inverted and sealed. All plates were then placed in a vibration excluded temperature controlled environment at both 20°C and 4°C and monitored daily for crystal formation.

2.2.2.9.2 Fine-Screening to refine optimal crystal formation conditions

Once initial crystal forming conditions were identified refinement of these conditions was carried out in order to increase the size and number of crystalline

outcomes. Fine screening conditions were prepared for the Crystal Screen II condition number 13 and condition number 5. Each buffer was mixed thoroughly by vortexing for 1-2 minutes before setting up the trays for crystallisation.

Refinement of Crystal Screen number 13 (0.2 M ammonium sulphate, 0.1 M sodium acetate trihydrate, pH 4.6, 30 % (w/v) polyethylene glycol monomethyl ether (PEG MME) 2,000)

Variations to the pH of this condition with the amount of PEG MME were assessed to see if crystal growth formation was improved. Investigation for the optimal pH of this condition was performed by adding 50 µL sodium acetate (pH range 4.3-4.8) according to the plate map (Figure 2.3). Stock solutions of PEG MME 2000 (50% (w/v)) and ammonium sulphate (NH₄)₂SO₄ (2M) were used to make up the individual buffers; the amounts added are detailed in Table 2.3. Once the buffers were prepared, 450 µL of each condition was added to the appropriate well (Figure 2.3). Once buffers were prepared and thoroughly mixed, trays were set as for the coarse screen (section 2.2.2.9.1)

Table 2.3: Preparation of buffers for fine screening for scFv B8 Ser MT crystal formation

% PEG MME 2000	Amount of Stock to Add (mL)	Ammonium Sulphate (M) concentration	Amount of Stock to Add (mL)	MG H₂O (mL)
28% (w/v)	3.53	0.2 M	0.63	1.51
29% (w/v)	3.65	0.2 M	0.63	1.39
30% (w/v)	3.78	0.2 M	0.63	1.26
31% (w/v)	3.91	0.2 M	0.63	1.13

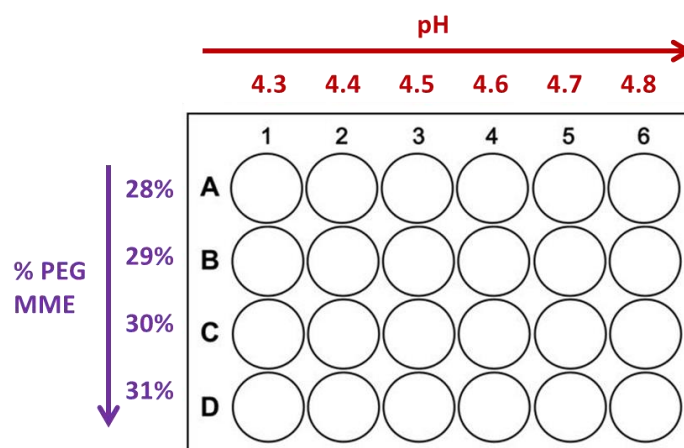


Figure 2.3 Plate map detailing the layout of the fine screening of condition number 13 for scFv B8 Ser MT crystal formation. To see if variations in pH could improve on crystal formation a variety of different pH's (pH 4.3-pH 4.8) was assessed (across-red), as was the % PEG MME (28-31%) (down-purple). Drops (1 μ) were set with a 1:1 ratio of mother liquor: protein.

Refinement of Crystal Screen number 5 (2 M ammonium sulphate, 5% (v/v) 2-propanol)

Varying high pH was performed by adding 50 μ L Bis-Tris pH buffers (pH range 5.3-6.0) whilst varying low pH was carried out with acetate pH buffers (pH range 4.3-4.8) according to the plate map in Figure 2.4. Stock solutions of IPA (25% (v/v)) and $(\text{NH}_4)_2\text{SO}_4$ (2 M) were used to make up the individual buffers as detailed in Table 2.4. Original conditions were also repeated, using freshly prepared 2 M $(\text{NH}_4)_2\text{SO}_4$ and varying % IPA, according to Figure 2.4. Four hundred and fifty microliters of each buffer was added to the appropriate well according to the plate maps (Figure 2.4). Once buffers were prepared and thoroughly mixed, trays were set as for the coarse screen (section 2.2.2.9.1).

Table 2.4: Preparation of buffers for fine screening for scFv B8 Ser MT crystal formation

% 2 Propanol	Amount of Stock to Add (mL)	Ammonium Sulphate (M) concentration	Amount of Stock to Add (mL)	MG H ₂ O (mL)
5% (v/v)	1.26	2 M	3.71	0.704
4% (v/v)	1.01	2 M	3.71	0.954
3% (v/v)	0.76	2 M	3.71	1.204
2% (v/v)	0.5	2 M	3.71	1.464

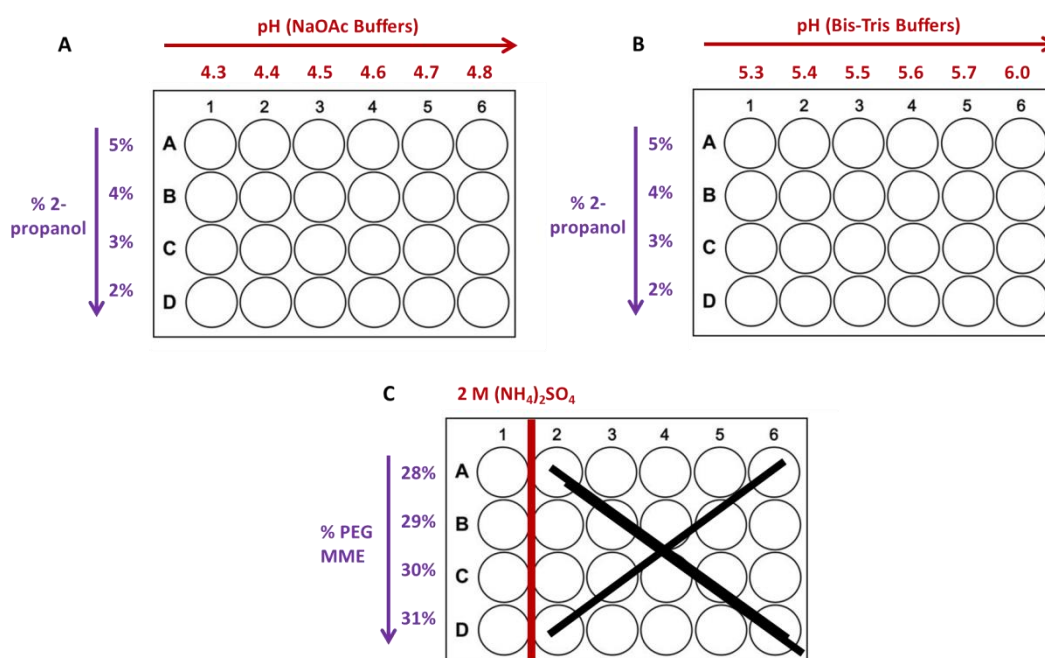


Figure 2.4: Plate map detailing the layout of the fine screening of condition number 5 for scFv B8 Ser MT crystal formation. Fine screening was carried out to see if variations in pH could improve on crystal formation at a variety of different pH's (A) pH 4.3-pH 4.8, using NaOAc buffers), (B) pH 5.3-pH 6.0) was assessed (across-red), (C) The original condition was also repeated and the percentage 2-propanol (5%-2% [v/v]) was also investigated for each new condition (down-purple). Drops (1 μ) were set with a 1:1 ratio of mother liquor: protein.

2.2.2.9.3 X-ray diffraction of crystals of scFv B8 Ser MT

Individual crystals were collected in a goniometer loop (allows the crystal to be positioned accurately within the X-ray beam and rotated) of an appropriate size (decided visually under a microscope) and flash-cooled in liquid nitrogen using 25% (v/v) glycerol as a cryoprotectant. X-ray diffraction data-sets were collected at the Australian Synchrotron MX2 beamline. The data was merged and processed using XDS , POINTLESS, and SCALA. Molecular Modelling and replacement was carried out by Dr. Paul Conroy, as previously described (Conroy *et al.*, 2014).

2.2.2.10 Kinetic analysis of scFv B8 Ser MT using Biacore 3000

Kinetic analysis was performed in a similar manner to the wild type (2.2.2.5). However, fPSA was immobilised to the CM5 chip. The instrument was prepared as described in section 2.2.2.5 and a preconcentration study to identify the optimal buffer conditions for subsequent immobilisation was carried out by preparation of 25 µg/mL fPSA, diluted in 10 mM NaOAc over a range of pH values (4.0, 4.2, 4.4, 4.6, 4.8 and 5.0). This was then passed over the underivatized chip surface (flow cell 4) and the degree of electrostatic binding monitored.

Once the optimal pH for immobilisation was identified fPSA was immobilised to the surface by passing fPSA diluted in 10 mM NaOAc to 25 µg/mL, pH 4.0 over the activated surface at a flow rate of 10 µL/minute for 15 minutes. Following capping and deactivation of any unreactive NHS groups, the surface was post conditioned with 20 mM NaOH.

ScFv B8 Ser MT was kinetically characterised using the kinetic wizard functionality on the Biacore 3000 and the subsequent data generated was evaluated using BIAevaluation software, version 4.1.1. Purified scFv B8 Ser MT was passed over the fPSA surface in varying concentrations (2.5 μ M, 1.25 μ M, 0.625 μ M (x2), 0.312 μ M, and 0 μ M) for 3 minutes, and allowed to dissociate for 12 minutes. The surface was regenerated using a 20 mM solution of NaOH. The dataset from the reference flow cell (flow cell 3 - underivatized surface) was subtracted to remove any systemic effects, and the zero nM scFv B8 Ser MT concentration binding response (buffer only) was subtracted from each antigen response, thus double reference subtracting the dataset. Kinetic constants were modelled on the curves using a 1:1 Langmuir binding model.

2.2.2.11 Expression and Purification of scFvB8 Ala MT by Immobilised Metal Affinity Chromatography (IMAC) and Gel Filtration

Expression and purification of scFv B8 Ala MT was carried out by Dr. Paul Conroy as described in section 2.2.2.8 and in Conroy *et al.*, (2014).

2.2.2.12 Crystal Structure of scFv B8 Ala MT

The crystal structure analysis of scFv B8 Ala MT was carried out by Dr. Paul Conroy. Briefly, scFv B8 Ala MT was expressed, purified and concentrated to 6.5 mg/mL, with a 3 kDa concentrator (Merck Millipore). Crystallisation was carried out using the hanging drop method with a 1:1 mixture of protein and mother liquor. Screening for crystal forming conditions was carried out using Crystal Screen and Crystal Screen II (Hampton Research) at 20°C. Crystals of scFv B8 Ala

MT were obtained in 2 M ammonium sulphate. The crystals were flash-cooled in liquid nitrogen using 25% (v/v) glycerol as the cryoprotectant.

Data sets were collected at the Australian Synchrotron MX2 beamline at 100K. The data was merged and processed using XDS , POINTLESS, and SCALA. Molecular Modelling and replacement was carried as described by Conroy *et al.*, (2014).

2.2.2.13 Mammalian Expression of recombinant PSA

At 24 hours prior to transfection HEK293F cells were seeded to $1-2 \times 10^6$ cells/mL in Freestyle Media (Gibco) in 125 mL shake flasks, in the following conditions; 5% (v/v) carbon dioxide, 37°C atmosphere, at 120 rpm. Recombinant PSA (pSecTag2A/PSA, Invitrogen) was transiently expressed using 1 µg/mL DNA and polyethylenimine (PEI) as transfection reagent, with a DNA:PEI ratio of 1:3. FreeStyle Media was supplemented with 2 mM glutamine, 5 g/L (w/v) lupin and the glucose concentration was adjusted to 33 mmol/L. Six days following transfection, the PSA-containing medium was collected for subsequent purification.

2.2.2.14 Purification of recombinant PSA by Immobilised Metal Affinity Chromatraphy (IMAC) and Gel Filtration

IMAC purification was carried out by adding the PSA-containing medium to HisBind (Novagen) resin (1 mL equilibrated with 1x binding buffer). Batch binding was then carried out overnight at 4°C on an end-over-end roller. The resin was collected by gravitational flow and washed with 10 mL 1x binding buffer. The

recombinant PSA was recovered by eluting in 5 mL elution buffer (300 mM Imidazole), in 0.5 mL fractions. Protein-containing fractions were concentrated to 1 mL in a 3 kDa concentrator. A S75 (16/60) gel filtration column (GE Healthcare) was equilibrated with 25 mM Tris, pH 7.4, 150 mM NaCl, 0.02% (w/v) NaN₃ (1CV = 50 mL). The concentrated 1 mL sample of PSA (1 mg/mL), previously purified by IMAC, was applied to the column at a flow rate of 1 mL/minute. Fractions were collected after 0.4 column volumes until 1 column volume was reached. Protein was monitored online by A₂₈₀ using the dedicated Unicorn control and evaluation software (GE Healthcare). Protein peaks were analysed by SDS-PAGE.

2.2.2.15 Formation of recombinant PSA-scFvB8 complex

Anti-PSA scFvB8 was expressed in *E. coli* and purified by IMAC and gel filtration chromatography as previously described in section 2.2.2.8 and in Conroy *et al.*, (2014). The antibody-antigen complex was formed by adding 1 mg/mL anti-fPSA scFvB8 to 0.4 mg/mL recombinant PSA, mixing gently and leaving static on ice for 2 hours (excess antibody was added to ensure there would be no free antigen in solution). This antibody-antigen complex was then concentrated down to 1 mL and resolved on a S75 (16/60) gel filtration column (GE Healthcare) as before. Purified complexed protein fractions were collected and analysed by SDS-PAGE.

2.2.2.16 Determination of the crystal structure of recombinant PSA-scFvB8 complex by X-ray crystallography

Elucidation of the molecular architecture of antibody-antigen provides a wealth of information across a variety of areas, including immunity (understating the

mechanisms that underlie antibody-antigen interactions), antibody engineering and epitope recognition. Knowledge of the antibody-antigen interface can identify residues that are important for antigen recognition and, hence, guide engineering of better binders (Sela-Culang *et al.*, 2013). There is currently no protein antigen-avian antibody structure available in the PDB. Determination of the crystal structure of recombinant PSA-scFvB8 complex could aid in our understanding of chicken antibody recognition.

2.2.2.16.1 Coarse screening for optimal crystal formation conditions

Coarse screening was carried out in 24-well plate format. The complex was concentrated to 7.3 mg/mL and the plates (24-well) were prepared by application of a ring of petroleum jelly to the circumference of each well and adding 500 μ L of mother from the ProPlex Kit (Molecular Dimensions) for the crystallization of protein complexes. Conditions 1 to 96 were placed in the appropriate number of wells (4 x 24-well plates, per kit). A 0.5 μ L drop of mother liquor was placed in the centre of the siliconised glass slides (Hampton Research) and 0.5 μ L of anti-PSA scfvB8/PSA protein complex (at 7.3 mg/mL) was mixed into the mother liquor drop. Two controls were used, firstly to ensure the crystals formed were complex crystals a drop was included for each condition in which a 0.5 μ L drop of mother liquor was mixed with 0.5 μ L scFv B8 (at 3.25 mg/mL). The second control drop included (for each condition) a 0.5 μ L drop of mother liquor was mixed with 0.5 μ L buffer and placed in the centre of the siliconised glass slides below the protein drop. The slide was then carefully inverted and sealed. All plates were then placed in a vibration excluded temperature controlled environment at 20°C and monitored daily for crystal formation.

2.2.2.16.2 Fine screen to refine for optimal crystal formation conditions

Fine screening for condition 2.8 from the identified in Proplex crystallisation kit, (0.1 M magnesium acetate, 0.1 M MOPS pH 7.5, 12% (w/v) PEG 8000) was carried out by Dr. Paul Conroy by optimising for pH and percentage PEG 8000.

2.2.2.16.3 X-ray diffraction of crystals of recombinant PSA-scFvB8 complex

X-ray diffraction of the recombinant PSA-scFvB8 complex was carried out by Dr. Paul Conroy. Individual crystals were collected in a goniometer loop of an appropriate size (decided visually under a microscope) and flash-cooled in liquid nitrogen using 25% (v/v) glycerol as a cryoprotectant. X-ray diffraction data-sets were collected at the Australian Synchrotron MX2 beamline. The data was merged and processed using XDS, POINTLESS, and SCALA. Molecular Modelling and replacement was carried out as previously described (Conroy *et al.*, 2014).

2.2.3 Generation of an avian scFv antibody library

2.2.3.1 Immunisation of female Leghorn chickens with cPSA

An adult female Leghorn chicken was administered 1 mL, subcutaneous injections, with cPSA for the production of cPSA recombinant antibodies. Namely, on day one, the chicken was immunised with 25 µg/mL cPSA in 550 µL PBS emulsified in a 1:1 ratio with Freund's complete adjuvant (FCA). The mixture was vigorously mixed to ensure a complete homogenous solution. Subsequent 'booster' injections were performed at 4 week intervals until a sufficient titre was achieved.

Booster immunisations were performed with 12.5 µg/mL cPSA mixed in a 1:1 ratio with Freund's incomplete adjuvant (FCIA). One-week post immunisation bleeds were taken from the animal to determine antiserum titre. Once the titre reached an appropriate level, as determined by ELISA, the animals were given a final boost 7 days prior to being sacrificed by lethal injection (EUTHASOL® (pentobarbital sodium and phenytoin sodium) Solution). Both the bone marrow and the spleen were harvested for library building.

2.2.3.2 Determination of avian serum antibody titres for immune response to cPSA

An indirect ELISA was used to determine the serum antibody titres from the immunised animal. A Maxisorp plate was coated with equimolar concentrations (30 nM) of cPSA/fPSA/ACT diluted in PBS for 1 hour at 37°C (100 µL/well). The plate was then blocked with 200 µL/well PBSM for 1 hour at 37°C and washed three times with 200 µL/well PBST, followed by three washes with PBS only. A series of dilutions ranging from 1 in 100 to 1 in 1,100,000 of the chicken serum, diluted in PBSTM was added to the plate (100 µL/well) and incubated for 1 hour at 37°C. The plate was washed as before. Donkey anti-chicken-horse radish peroxidase-conjugated secondary antibody (anti-IgY-Fab-HRP) was added at 1 in 2,000 dilution in PBSTM (100 µL/well) and incubated for 1 hour at 37°C. The plate was washed as before and 100 µL/well TMB substrate solution was added to each well. The plate was incubated at room temperature for approximately 5 minutes to allow chromophore development, after which the reaction was stopped with addition of 50 µL/well 10% (v/v) HCl. The absorbance was read at 450 nm using the Tecan Safire™ plate reader.

2.2.3.3 Isolation and quantification of total cellular RNA from spleen and bone marrow

Total RNA was recovered from the bone marrow (BM) and spleen (SP) of the immunised animals by chloroform-based extraction carried out immediately after sacrifice. All reagents used were of molecular grade and all materials were treated with RNaze™ Zap where possible. This is a decontamination solution that completely removes RNase contamination from glass and plastic surfaces. Minor trace quantities of RNase can lead to degradation during the RNA purification process and hence it is essential to ensure the method can be carried out in an 'RNase-free' environment (Sigma Aldrich, 2015).

Prior to use the laminar unit was cleaned with 70% (v/v) isopropanol alcohol (IPA) and RNaze™ Zap to ensure an uncontaminated environment. The spleen was removed and placed in a 50 mL 'RNase-free tube' and 10 mL of ice-cold TRIzol® reagent was immediately added (TRIzol® reagent is an extremely hazardous substance and Personal Protective Equipment must be worn when handling TRIzol®. Spills must be cleaned up immediately with great care. Hypochlorite's should not be used to clean up spills and TRIzol® must be kept away from heat or any ignition sources at all times). SP and BM were processed separately.

The bone-marrow (from the thigh bone) was extracted by cutting away the joint knuckles at either end of the bone and flushing out as much marrow as possible using a syringe and needle containing 10 mL TRIzol® reagent. Both SP and BM were homogenised using a pre-sterilised (with presept), RNaze™ Zap-washed and autoclaved homogeniser probe. A further 20 mL TRIzol® reagent was added to each tube when both the SP and BM were completely homogenised and allowed to stand at room temperature (RT) for 10 minutes. The mixture was then

centrifuged at 5,000 x *g* for 20 minutes at 4°C (ensuring that the centrifuge was cooled prior to centrifugation) and the resultant supernatant carefully transferred to a 50 mL Oakridge tube. Six mL chloroform was added and the tube was vigorously shaken (by hand) for 15 seconds and then let stand at RT for 15 minutes.

The Oakridge tubes were centrifuged at 15,500 x *g* for 30 minutes at 4°C and the upper aqueous layer was removed and transferred to an 85 mL Oakridge tube. Hence a small residual portion of the aqueous layer was left rather than risking disruption of the interphase. To the aqueous layer, 15 mL IPA was added with vigorous shaking for 15 seconds and then let stand at RT 10 minutes, followed by centrifugation at 15,500 x *g* for 30 minutes at 4°C. The resultant supernatant was removed carefully ensuring the opaque pellet was undisturbed. The pellet was then washed with 30 mL 75% (v/v) MG EtOH and centrifuged at 15,500 x *g* for 30 minutes at 4°C and the pellet was left to air dry for 20 minutes. The pellet was resuspended in 500 µL 'RNase-Free' water and the RNA quantified using the NanoDrop spectrophotometer. A sufficient quantity of RNA was used for cDNA synthesis and the remainder was ethanol-precipitated as described in section 2.2.1.2 and stored at -80°C.

2.2.3.4 Reverse transcription of total RNA to cDNA

The superscript III first stand synthesis system for RT-PCR was used, as per the manufacturer's guidelines to transcribe the total RNA obtained to cDNA. Once successfully synthesised the cDNA was quantified using the NanoDrop spectrophotometer and stored in 10 µL aliquots at -80°C.

2.2.3.5 ScFv library construction

Antibody library construction was carried out as described by Andris-Widhopf and co-workers (Andris-Widhopf *et al.*, 2000; Barbas *et al.*, 2001). The initial steps carried out were the amplification of the V_H and V_L domains from the synthesised cDNA from the BM and SP of the chicken. The primers used for library construction are detailed in Phage Display: A Laboratory Manual, and are used to clone antibody fragments into the pComb3x vector series. Chapter 9 of this book details the construction of a chicken scFv library with a short linker (GGSSRSS) or a long linker (GGSSRSSSGGGGSGGGG) (Barbas *et al.*, 2001). The primers were commercially synthesised by Integrated DNA Technologies and are listed herein.

Variable Heavy (V_H)

CSCVHo-FL: (sense - Long linker)

5' GGT CAG TCC TCT AGA TCT TCC GGC GGT GGT GGC AGC TCC GGT GGT
GGC GGT TCC GCC GTC ACG TTG GAC GAG 3'

CSCG-B (reverse)

5' CTG GCC GGC CTG GCC ACT AGT GGA GGA GAC GAT GAC TTC GGT CC 3'

Variable Light (V_L)

CSCVK (sense)

5' GTG GCC CAG GCG GCC CTG ACT CAG CCG TCC TCG GTG TC 3'

CKJo-B (reverse)

5' GGA AGA TCT AGA GGA CTG ACC TAG GAC GGT CAG G 3'

SOE Overlap

CSC-F (sense)

5' GAG GAG GAG GAG GAG GAG GTG GCC CAG GCG GCC CTG ACT CAG 3'

CSC-B (reverse)

5' GAG GAG GAG GAG GAG GAG GAG CTG GCC GGC CTG GCC ACT AGT
GGA GG 3'

2.2.3.6 PCR amplification and optimisation of the heavy and light-chain variable (V_H and V_L) coding regions

To amplify the V genes, one V_H amplification and one V_L amplification were carried out. Initially magnesium chloride ($MgCl_2$) is optimised for the V_H/V_L amplifications. Identification of the optimal $MgCl_2$ concentration is crucial for successful amplifications. Optimisation reactions were carried out at 1x (Table 2.5) and once the optimal $MgCl_2$ concentration was determined large-scale reactions (10x) were carried out. The PCR reaction(s) were carried out under the following conditions outlined in Table 2.6.

Table 2.5: Small-scale (1x) reaction for magnesium chloride (MgCl₂) optimisation for V_H/V_L amplifications

	1.5mM	2.0mM	3.0mM	4.0mM
GoTaq Buffer (5x)	10 µL (1x)	10 µL	10 µL	10 µL
MgCl₂ (25 mM)	3 µL	4 µL	6 µL	8 µL
MG H₂O	33.75 µL	32.75 µL	30.75 µL	28.75 µL
Fwd. Primer (60 pM)	0.5 µL	0.5 µL	0.5 µL	0.5 µL
Rev. Primer (60 pM)	0.5 µL	0.5 µL	0.5 µL	0.5 µL
dNTPs (100 mM)	1 µL	1 µL	1 µL	1 µL
cDNA	1 µL (1 µg)	1 µL (1 µg)	1 µL (1 µg)	1 µL (1 µg)
Polymerase (5U/(µL))	0.25 µL (1.25 U)	0.25 µL (1.25 U)	0.25 µL (1.25 U)	0.25 µL (1.25 U)
Total Vol	50 µL	50 µL	50 µL	50 µL

Table 2.6: PCR conditions for amplification of the V_H and V_L chain genes

Stage	Step	Temp (°C)	Time (seconds)	Number of cycles
1	Initial Denature	94	120	1
2	Denature	94	15	30
	Annealing	56	30	
	Extension	72	60	
3	Final Extension	72	600	1

In order to ensure there was sufficient PCR product to carry out the second round PCR's, 10x reactions were carried out four times for each of the variable domains for both the BM and SP. The products were pooled and ethanol-precipitated overnight at -20°C, as described in section 2.2.1.2. The pellet was resuspended in 36 µL pre-warmed (in the 80°C dryer) MG H₂O and 4 µL 10x loading buffer dye (for a final concentration of 1x). The products were resolved at 90V for 50 minutes on 1.5% (w/v) agarose gel and bands at approximately 350bp (V_L) and 450bp (V_H) were excised and gel-purified as per the manufacturer's guidelines using

Nucleospin Extract II – Extraction of DNA from agarose gels kit. The purified products were quantified using the NanoDrop spectrophotometer.

2.2.3.7 Generation of scFv-fragment by splice-by-overlap extension (SOE) PCR

For the second round PCR, the purified first round PCR products were mixed in equal ratios to generate the final overlap product. The primers used in the first round PCR produce identical sequences that serve as the overlap for the extension of the full length scFv construct in the V_L - V_H orientation (Barbas *et al.*, 2001). The use of High Fidelity Platinum Taq polymerase was employed for these PCR cycles to ensure optimal yield, fidelity and specificity. Initially the magnesium sulphate ($MgSO_4$) concentration was optimised at 1x and once the optimal concentration was determined large-scale (10x) reactions were carried out. The SOE- PCR reactions were carried out under the conditions set out in Table 2.8

Table 2.7: Small-scale (1x) magnesium sulphate ($MgSO_4$) optimisation for SOE-PCR

	1.5mM	2.0mM	3.0mM	4.0mM
High Fidelity PCR Buffer (10x)	5 μ L	5 μ L	5 μ L	5 μ L
$MgSO_4$ (50 mM)	1.5 μ L	2 μ L	3 μ L	4 μ L
M.G. H_2O	Up to 50 μ L	Up to 50 μ L	Up to 50 μ L	Up to 50 μ L
Fwd. Primer CSC-F	0.5 μ L	0.5 μ L	0.5 μ L	0.5 μ L
Rev. Primer CSC-B	0.5 μ L	0.5 μ L	0.5 μ L	0.5 μ L
dNTPs	1 μ L	1 μ L	1 μ L	1 μ L
V_L	100 ng	100 ng	100 ng	100 ng
V_H	100 ng	100 ng	100 ng	100 ng
Polymerase (5U/ μ L)	0.2 (IU)	0.2 (IU)	0.2 (IU)	0.2 (IU)
Total	50 μ L	50 μ L	50 μ L	50 μ L

Table 2.8: SOE-PCR reaction conditions

Stage	Step	Temp (°C)	Time (seconds)	Number of cycles
1	Initial Denature	94	120	1
2	Denature	94	15	30
	Annealing	56	30	
	Extension	68	60	
3	Final Extension	68	600	1

The SOE products were pooled and ethanol-precipitated, as described in section 2.2.1.2 and gel-purified by resolution on 1.5% (w/v) agarose gel at 90V for 50 minutes. These were then extracted using the Nucleospin Extract II – Extraction of DNA from agarose gels kit, as per the manufacturer’s guidelines. The purified scFv constructs were eluted in 30 µL MG H₂O.

2.2.3.8 Preparation of pComb3xSS vector

The pComb3xSS vector was kindly supplied by Prof. Carlos Barbas III (Scripps Institute, La Jolla, California). An overnight culture (100 mL SB media supplemented with 50 µg/mL carbenicillin) was inoculated with 2 µL of glycerol stock of pComb3xSS and cultured at 37°C with shaking at 220 rpm. The overnight culture was centrifuged at 5,500 x *g* for 30 minutes at 4°C, the supernatant was discarded and the plasmid was purified using the Nucleobond Xtra midi kit, as per the manufacturer’s guidelines. Plasmid DNA was eluted in 300 µL MG H₂O and quantified with the NanoDrop spectrophotometer.

2.2.3.9 Restriction digestion of the SOE-PCR and phagemid cloning vector pComb3xSS

The PCR products for both the SP and BM and the pComb3xSS vector were prepared for cloning by restriction digestions with *Sfi* I. All reactions were set up as detailed in Table 2.5 and incubated at 50°C for 5 hours in the PCR thermocycler. The scFv product was ethanol-precipitated @ -20°C overnight, as previously described in section 2.2.1.2.

Table 2.9: *Sfi* I digest of pComb3xSS vector and scFv inserts

	SP	BM	pComb3xSS
DNA	10 ng	10 ng	20 ng
Buffer 4 (10x)	1x	1x	1x
BSA (100x)	1x	1x	1x
M.G. H₂O	Up to 200 µL	Up to 200 µL	Up to 200 µL
<i>Sfi</i> I (20 U/µL)	36 U/µg	36 U/µg	6 U/µg
Total	200 µL	200 µL	200 µL

The pComb3xSS vector was then subjected to further digestion with *Xho* I and *Xba* I by incubation at 37°C for 1 hour followed by 65°C for 30 minutes (Table 2.6). The vector was then Antarctic phosphatase-treated (Table 2.10) by incubation at 37°C for 30 minutes, followed by deactivation at 65°C for 15 minutes. Post triple digestion and Antarctic phosphatase treatment the vector was ethanol-precipitated, as outlined in section 2.2.1.2.

Table 2.10: *Xba* I and *Xho* I digestion of pComb3xSS vector

	pComb3xSS	pComb3xSS (vol)
<i>Xba</i> I	3 U/μg	1.5 μL
<i>Xho</i> I	3 U/μg	1.5 μL

Table 2.11: Antarctic phosphatase treatment of digested pComb3x vector

	pComb3x	pComb3x (vol)
Antarctic Phosphatase Buffer (10x)	1x	20 μL
Antarctic Phosphatase (5 U/μL)	1 U/μg	4 μL

2.2.3.10 Library ligation and transformation

The purified *Sfi* I digested scFv products for BM and SP were ligated into the pComb3x using T4 DNA ligase. The reaction set up (detailed in Table 2.12) and was carried out overnight at RT.

Table 2.12: Ligation reaction components of scFv insert and pComb3x vector

	SP	BM	Control
Vector	1.4 μg	1.4 μg	1.4 μg
Insert	0.7 μg	0.7 μg	NA
MG H₂O	Up to 100 μL	Up to 100 μL	Up to 100 μL
Ligase Buffer (10x)	1x	1x	1x
T4 DNA Ligase (500 U/μL)	10 U/μg	10 U/μg	10 U/μg
Total	100 μL	100 μL	100 μL

The overnight ligated product was deactivated at 65°C for 20 minutes and then ethanol-precipitated at -20°C overnight as described in section 2.2.1.2. The overnight precipitates were centrifuged at 19,500 x *g* for 30 minutes at 4°C,

washed with 100 μ L ice cold 70% (v/v) EtOH and centrifuged again at 19,500 $\times g$ for 10 minutes at 4°C. The resultant pellet was resuspended in 15 μ L pre-warmed (in an 80°C oven) MG H₂O

Electrocompetent XL1 Blue *E. coli* cells (150 μ L) (Agilent technologies, Ireland) were placed in pre-chilled 2.5 mm electroporation cuvettes (Bio-rad) and allowed to thaw before adding 7.5 μ L of the library ligation. The cells and DNA were let stand on ice for 1 minute and then the cuvette was gently tapped to ensure the suspension was at the base of the cuvette. DNA was electroporated into the cells using the Gene Pulser xCell Electroporation system (Bio-rad) with the following parameters: 2.5 kV, 25 μ F and 200 Ω (τ = 4.0 msec). The cuvette was immediately flushed with 1.25 mL of SOC media and transferred to a pre-warmed 50 mL tube. All the electroporated samples (x2 BM & x2 SP) were combined in one 50 mL tube (approximately 5 mL total volume). To facilitate the recovery of the cells they were then incubated at 37°C for 1 hour with shaking at 220 rpm.

A control ligation (digested pComb3x – no insert) was also electroporated and treated in the same fashion. To both the library and the control, 10 mL of pre-warmed (37°C) SB media and 3 μ L carbenicillin (100 mg/mL) and 30 μ L tetracycline (5 mg/mL) were added and a transformation titre was carried out by taking 2 μ L of this 15 mL culture and adding to 198 μ L SB media. This was then plated on LB agar plates supplemented with 100 μ g/mL carbenicillin (10 μ L and 100 μ L was plated each for control and library). The control was discarded and the library was incubated for 1 hour at 37°C, 220 rpm.

After this incubation, 4.5 μ L carbenicillin (100 mg/mL) was added and the culture was incubated further for 1 hour at 37°C, while shaking at 220 rpm. Two mL of MK1307 (1 $\times 10^{11}$ pfu/mL) helper-phage was then added and the library was

transferred to a 500 mL baffled flask containing 183 mL pre-warmed SB supplemented with 91.5 μ L (100 mg/mL) carbenicillin and 370 μ L tetracycline (5 mg/mL). This culture was incubated for 2 hours with shaking, as before. For the final step, 280 μ L of 70 mg/mL kanamycin was added and the culture was left to shake at 220 rpm o/n at 37°C. At this stage the library was displayed on the phage particle and is ready for panning.

2.2.3.11 Selection of cPSA-specific phage particles by panning against immobilised cPSA on an ELISA plate

2.2.3.11.1 Phage rescue after each round of panning

Phage rescue is the process carried out during each round of panning to recover the selected scFv-displaying phage from the culture supernatant. After incubation of the transformed library overnight (section 2.2.3.10) the phage are isolated by PEG precipitation. The o/n culture was split into four 50 mL aliquots in (new, clean) 85 mL Oakridge tubes and centrifuged at 11,000 x *g* for 15 minutes, 4°C. The supernatant was then transferred to clean (new) 85 mL Oakridge tubes and 12.5 mL 5x PEG (total volume 62.5 mL) was added. The samples were gently inverted several times and left on ice for 30 minutes. The precipitated phage was then centrifuged at 16,000 x *g* for 15 minutes at 4°C, the supernatant discarded and the tubes left inverted on paper towel for 10 minutes. The phage-containing pellet was resuspended in 2 mL 1% (w/v) BSA/PBS/0.02% (w/v) NaN₃ and transferred to a 2 mL micro-centrifuge tube. Residual cell debris was then removed by centrifugation at 19,500 x *g* for 5 minutes at 4°C. This supernatant was then transferred to fresh 2 mL tube and kept at 4°C until ready to add to the ELISA plate for panning (R0).

2.2.3.11.2 Library panning on an cPSA immobilised surface

Panning consist of several rounds of binding phage to an antigen immobilised on the wells of an ELISA plate, washing, elution of the specifically bound phage and reamplification (Figure 1.4). The wells of a 96-well ELISA plate (8 wells) were coated with 50 µg/mL ACT (depletion step) (100 µL/well) diluted in PBS and a separate plate was coated 50 µg/mL cPSA (100 µL/well) diluted in PBS and both were left overnight at 4°C. The ACT/cPSA was removed and the coated wells were blocked with PBSM 200 µL/well, for 1 hour at 37°C. The blocking solution was removed and the plate was washed 3 times with PBST and 3 times with PBS (200 µL/well). The freshly rescued phage (100 µL/well) were then added to each of the 8 wells of the ACT plate and allowed to interact with the ACT for 2 hours at 37°C.

Unbound scFv displaying phage were then removed from this plate and transferred to the cPSA plate and allowed to interact with the cPSA for 2 hours at 37°C. Non-specific phage were then removed by washing 5 times with PBST and 5 times PBS, by pipetting up and down 5 times vigorously followed by leaving static for 5 minutes for each wash step. Specifically bound cPSA displaying scFv phage were then eluted with 100 µL per well of 10 mg/mL trypsin/PBS by pipetting up and down vigorously 10 times and incubating for 30 minutes at 37°C.

These eluted phage were then infected into 2mL of mid-exponential *E. coli* XL-1 blue cells and let stand at RT for 15 minutes. This 2.8 mL was then added to 6 mL pre-warmed SB media with 1.6 µL carbenicillin (100 mg/mL) and 12 µL tetracycline (5 mg/mL) in a 50 mL tube and incubated for 1 hour at 37°C with shaking at 220 rpm. An output titre was performed at this point by taking 2 µL of this 8 mL culture and diluting to 200 µL in SB, and subsequently plating out 10 µL

and 100 μ L of this on LB-carbenicillin (50 μ g/mL)-agar plates and incubating upside down o/n at 37°C. After the hour incubation, a further 2.4 μ L carbenicillin (100 mg/mL) was added and the 50 mL tube was incubated for an additional hour at 37°C, with shaking at 220 rpm.

Helper phage (MK13O7: 1×10^{11} pfu/ml) were then added and the culture was transferred to a 500 mL baffled flask containing 91 mL pre-warmed SB, 46 μ L carbenicillin (100 mg/mL) and 184 μ L tetracycline (5 mg/mL) and incubated at 37°C for 2 hours, while shaking at 230 rpm. The final step was the addition of 140 μ L kanamycin (70 mg/mL), and the culture was left to grow o/n 37 °C, with shaking at 230 rpm.

An input titre was also performed (during the incubation steps) by adding 10 μ L of the rescued phage to 90 μ L SB, mixing gently and then making 10-fold serial dilutions to 10^{-9} . Mid-exponential *E. coli* XL-1 blue (50 μ L) was then infected with a 2 μ L sample of each of the 10^{-7} 10^{-8} and 10^{-9} phage dilutions, and incubated at RT for 15 minutes. These were then plated out on LB-carbenicillin (50 μ g/mL) agar plates and incubated upside down at 37°C overnight.

This panning process was repeated for 3 rounds on successive days with increasing washes and decreasing antigen concentrations (Table 2.13) to enrich for high-affinity cPSA-specific clones.

Table 2.13: Panning conditions employed for four successive rounds of panning for isolation of cPSA-specific scFv antibody

<i>Round</i>	<i>Depletion (ACT)</i>	<i>cPSA</i>	<i>Number of Wells</i>	<i>Washes</i>
R1	50 µg/mL	50 µg/mL	8	3 + 3
R2	50 µg/mL	50 µg/mL	4	5 + 5
R3	50 µg/mL	25 µg/mL	4	5 + 5

2.2.3.12 Polyclonal phage ELISA

After completion of 3 rounds of panning experiments a polyclonal phage ELISA was carried out to assess the success of the panning conditions and to identify which phage pool shows specific binding to cPSA. A 96-well ELISA plate was coated with 100 nM cPSA in PBS (100 µL/well) o/n at 4°C. The plate was then blocked with PBSM (200 µL/well) and incubated for 1 hour at 37°C. The wells were then washed three times with PBST and three times with PBS (200 µL/well). Dilutions (1 in 5) of each round of phage in PBSTM were prepared and 100 µL per well (in duplicate) was added to the plate and incubated for 1 hours at 37 °C.

The plate was then washed as before and 100 µL/well of a 1 in 2,000 dilution of HRP-labelled anti-M13 antibody prepared in PBSTM was added to the plate for 1 hour at 37°C. The plate was washed as before and 100 µL/well TMB enzyme substrate solution was added to each well. The plate was incubated at RT for approximately 5 minutes to allow chromophore development, after which the reaction was stopped with 50 µL/well of 10% (v/v) HCl. The absorbance of the wells was then read at 450 nm using the Tecan Safire™ plate reader.

2.2.3.13 Soluble expression of scFv fragments and subsequent monoclonal ELISA analysis

To express the avian anti-cPSA scFv clones, the scFv-containing phage were directly infected into a non-suppressor strain of *E. coli*. A growing culture (OD_{600} 0.6-0.8) of TOP10F' *E. coli* was infected with 50 μ L phage from each round of panning and incubated for 1 hour at 37°C. Dilutions from 10^{-1} to 10^{-7} were prepared in SB and plated (100 μ L) on LB/carbenicillin (100 μ g/mL) agar plates and incubated o/n at 37°C. Single colonies from each round were picked and grown in 150 μ L SB supplemented with 50 μ g/mL carbenicillin in the inner 60 wells of sterile 96-well culture plates and cultured overnight at 37°C with shaking at 220 rpm. Glycerol stocks were made of these plates, as described in section 2.2.1.5.

To assess the specificity of these scFv clones for cPSA, 30 clones from each round of panning were inoculated in 150 μ L SB supplemented with 50 μ g/mL carbenicillin and cultured overnight at 37°C with shaking at 220 rpm in a sterile 96-well culture plates. The following day, 30 μ L of this overnight culture was infected into 370 μ L fresh SB media supplemented with 50 μ g/mL carbenicillin, 1 mM $MgSO_4$ and 1x505 (0.5% (v/v) glycerol & 0.05% (v/v) glucose), in a sterile deep-well 96-well plate. The plate was shaken at 220 rpm at 37°C, until an optical density (OD_{600}) of ~0.600 was reached. A final concentration of 0.1 mM IPTG was added to each well and the plate was left shaking at 220 rpm overnight at 30°C. An ELISA plate was coated overnight at 4°C with 100 μ L/well 100 nM cPSA and subsequently blocked at 37°C for 1 hour with 200 μ L/well PBSM. The plate was then washed three times with 200 μ L/well PBST, followed by three times with PBS.

The scFv antibodies were extracted from the periplasmic space of the cell by placing the plates at -80°C for 10 minutes and then thawing at 37°C for 20 minutes. One hundred microliters/well of 1 mg/mL lysozyme/PBS were added to disrupt the cell wall. The plate was then subjected to two successive freeze-thaw cycles. This commonly used method releases periplasmic proteins by bursting the *E. coli* cells. The cell extracts were cleared by centrifugation at 5,500 x *g* for 30 minutes and the resultant lysate was diluted 1 in 2 in PBSTM and 100 µL of this scFv-lysate was added to the wells of the ELISA plate and incubated for 1 hour at 37°C. The plate was washed as before and the cPSA-specific scFv clones were then detected with 100 µL/well with a HRP tagged anti-IgY (H+L) antibody diluted (1 in 2,000) in PBSTM for 1 hour at 37°C. The plate was washed as before and 100 µL/well TMB substrate solution was added to each well. The plate was incubated at RT for approximately 5 minutes to allow chromophore development, after which the reaction was stopped with 50 µL/well 10% (v/v) HCl. The absorbance was read at 450 nm using the Tecan Safire™ plate reader.

2.2.3.14 Small scale expression and purification of anti-cPSA 2D2SG by Immobilised Metal Affinity Chromatography (IMAC)

An overnight culture of scFv was prepared by inoculating 2 µL glycerol stock scFv 2D2SG into 10 mL SB media with 50 µg/mL carbenicillin, with shaking at 220 rpm at 37°C. This culture was sub-cultured into a fresh 100 mL culture of SB containing 50 µg/mL carbenicillin, 1 mM MgSO₄ and 1x505. The culture was grown at 37°C until an OD₆₀₀ of 0.6-0.8 was reached and it was subsequently induced with 0.2 mM IPTG at 30°C for overnight. The expressed culture was then centrifuged at

5,500 x *g* for 45 minutes 4°C (50 mL in two 85 mL Oakridge tubes) and the supernatant was discarded.

Each pellet was then resuspended completely in 7.5 mL OSB-A and transferred to a clean 85 mL Oakridge tube (i.e. 15 mL total). OSB-B (15 mL) was then added and the incubated for 10 minutes at RT on the roller. The shocked cells were collected by centrifugation at 14,000 x *g* (4°C) for 20 minutes. The supernatant was discarded and pellets resuspended in 15 mL periplasmic buffer. Samples were place on ice for 10 minutes, with gently mixing by inverting the tubes every 2 minutes to ensure the solution was homogenous.

The cell debris was collected by centrifugation at 14,000 x *g* (4°C) for 20 minutes. The supernatant was pooled in a 50 mL tube and the protein was sequentially filtered through a 0.45 µm and 0.2 µm filters. Ni-NTA nickel resin slurry (2 mL) was added to a chromatography column and allowed to settle. The resin bed was washed with 10 mL PBS twice (20 mL total) followed by 10 mL of OSB-A buffer. The lysate was then applied to the column, collecting the 'flow-through' in a 50 mL tube. This process was repeated. The column was then washed with 10 mL of wash buffer 1 collecting in a 50 mL tube. This wash step was repeated with 10 mL wash buffer 2. The retained protein was then eluted by adding 5 mL elution buffer to the column and collecting the flow through. The eluted sample (5 mL) was then concentrated and buffer exchanged against PBS/0.02 % (w/v) NaN₃ using a Vivaspinn column with a 10,000 molecular weight 'cut-off' level, as described in section 2.2.1.6. Samples were taken throughout and analysed by SDS-PAGE.

2.2.3.15 Examination of anti-cPSA 2D2SG cross reactivity

To ensure anti-cPSA 2D2SG was specific for cPSA and not cross reactive with ACT or fPSA an indirect ELISA was performed. A 96 well Maxisorb plate was coated with 100 µl/well 100 nM cPSA/fPSA/ACT diluted in PBS and incubated for 1 hour at 37°C. The plate was then blocked with PBST (200 µl/well) and incubated for 1 hour at 37°C. The wells were then washed three times with PBST and three times with PBS (200 µl/well). One hundred microliters per well of anti-cPSA scFv 2D2SG (1 in 10 dilution in PBST) was then added to the plate and again incubated at 37°C for 1 hour. The plate was washed as before. A 1 in 2,000 dilution of anti-HA diluted, in PBST (100 µl/well) was added and incubated as before. The plate was washed as before and 100 µL/well TMB enzyme substrate solution was added to each well. The plate was incubated at RT for approximately 5 minutes to allow chromophore development, after which the reaction was stopped with 50 µL/well 10% (v/v) HCl. The absorbance of the wells was read at 450 nm using the Tecan Safire™ plate reader.

2.2.4. N-glycan profiling of avian polyclonal IgY antibodies

IgY is the predominant serum immunoglobulin in birds, reptiles and amphibians and is considered to be the evolutionary ancestor of mammalian IgG and IgE antibodies. IgY is more heavily glycosylated than its mammalian counterpart as it contains two potential *N*-glycosylation sites. One is located in the C_v3 domain, that is absent in mammalian IgG, and the other is located in the C_v3 domain which corresponds to the C_H2 (C_γ2) domain of mammalian IgG. Knowledge of the carbohydrate moieties present on immunoglobulins is essential as certain glycan

structures can adversely impact physicochemical and biological properties (Narat, 2003; Suzuki and Lee, 2004). Glycoproteins consist of a heterogeneous population of glycoforms in which a variety of glycan structures are present at each glycosylation site. Hence, the analysis of glycans requires high resolution separation techniques that can provide detailed structural analysis, including monosaccharide sequence and linkage information (Royle *et al.*, 2006). The polyclonal IgY purification approach and subsequent glycan analysis methods used are described herein.

2.2.4.1 Purification polyclonal IgY antibodies from serum

A Pierce® Thiophilic Adsorption Kit was used to purify polyclonal IgY antibodies from the serum of the immunised chickens. The manufacturer's protocol was used and the eluates were pooled and concentrated as previously described in section 2.2.1.6. Concentrated purified polyclonal IgY was then analysed by SDS-PAGE and Western Blotting (WB), as detailed in section 2.2.1.3 and 2.2.1.4, respectively. The Western Blots were probed with a 1 in 2,000 dilution of a HRP labelled donkey anti-chicken IgY-H+L antibody, diluted in PBSTM.

2.2.4.2 Glycan release

N-glycans were released from the dissolved samples (1 mg/mL) (lyophilised samples were resuspend in water) using PNGase F, (Prozyme, San Leandro, CA, USA) in accordance with methods previously described by Royle *et al.* (2006) and Royle *et al.* (2008). IgY samples were reduced in a polypropylene 96-well flat-bottomed microplate by adding 5 µl sample, 2 µl sample buffer (100 µl 10% SDS +

62.5 µl stacking buffer + 337.5 µl water), 2 µl water, and 1 µl of 0.5 M dithiothreitol (DTT) to the well mixing gently using pipette action. The plate was then were incubated at 65°C for 15 minutes. The sample was then alkylated by adding 1 µl of 100 mM iodoacetamide and incubated for 30 minutes in the dark at room temperature. The sample was then set into a gel block by addition of 22.5 µl of 30% (w/w) acrylamide/0.8% (w/v) bis-acrylamide stock solution (30% ProtoGel, National Diagnostics, Hesse, Hull, UK), 11.25 µl of 1.5 M Tris (pH 8.8), 1 µl of 10% SDS, 1 µl of 10% ammonium peroxodisulfate (APS), and finally, for gel polymerisation, 1 µl of *N,N,N,N'*-tetramethyl-ethylenediamine (TEMED). Reagents were mixed by pipetting and left to set.

The gel blocks were transferred to a filter plate (Whatman Protein Precipitation FF plate) and then washed with 1 ml of acetonitrile with shaking on a plate mixer for 10 min, followed by removal of the liquid (all wash and elution steps were performed on a vacuum manifold). This washing procedure was repeated twice with 1 ml of 20 mM NaHCO₃ (pH 7.2) followed by 1 ml of acetonitrile (Royle *et al.*, 2008).

N- and *O*- linked glycans utilise different chemistries for protein-sugar attachment hence require different approaches for their release. Intact *N*-glycans can be enzymatically released from glycoproteins with an amidase called peptide *N*-glycosidase F, PNGase F (Mariño *et al.*, 2010). The *N*-linked glycans were released from IgY using peptide *N*-glycanase F (Prozyme, San Leandro, CA, USA), as previously described (Royle *et al.*, 2008, Küster *et al.*, 1997). Briefly, after washing the gel pieces, the plate on top of a polypropylene 2ml deep 96-well plate (collection plate) and 25 µl 100 mU/ml PNGaseF solution was added to each gel. This was left for 5 minutes to be soaked up by the gel before adding an additional

25 µl PNGaseF. To the gel pieces 50 µl of 20 mM NaHCO₃ was added and the plate was sealed (with a SealPlate adhesive film) and incubated at 37°C overnight (Royle *et al.*, 2008). The released glycans were collected in the collection plate by washing the gel pieces with 3 x 200 µl of water, 200 µl of acetonitrile, 200 µl of water, and finally 200 µl of acetonitrile (1200 µl total volume). The released glycans were dried completely in vacuum centrifuge (this can take between 5 – 16 hours). Once samples were dry they were subjected to a formic acid treatment by addition of 20 µl 1% formic acid (made up fresh) and incubated at room temperature for 40 minutes, and then re-dried.

2.2.4.3 Labelling of released glycans with 2-aminobenzamide (2AB)

Glycans were then fluorescently labelled with 2-aminobenzamide (2AB) by reductive amination (Bigge *et al.*, 1995). Carbohydrates have no significant UV absorbing/fluorescent elements hence labelling is essential to achieve sensitive detection of the released glycans (Mariño *et al.*, 2010a). Labelling was performed by adding 5 µl of 2AB labeling solution (350 mM 2-aminobenzamide, 1 M sodium cyanoborohydride in acetic acid/dimethyl sulfoxide (30:70)), vortexing for 10 minutes, incubation for 30 minutes at 65°C, vortexing for a further 10 minutes, and a final incubation for 90 minutes at 65°C.

Excess 2AB was removed using solid-phase extraction (SPE) with Whatman 3MM chromatography paper. Prewashed, dried Whatman 3MM chromatography paper (cut into 1-cm square pieces) were folded into quarters and placed into a filter plate (Whatman Protein Precipitation FF previously conditioned by washing with 200 µl of acetonitrile followed by 200 µl of water). The 5 µl of 2AB-labelled

samples were applied to the paper and left to dry/bind for 15 minutes. Any excess 2AB was removed by addition 2 ml of acetonitrile for 15 minutes (with agitation) and then removing the acetonitrile using a vacuum manifold; this procedure was repeated four times. The labelled glycans were eluted from the chromatography paper by paper by addition of 900 µl of water and left to agitate for 30 minutes before vacuuming to a collection block. This was repeated with a further 900 µl water and the labelled eluted glycans (1800 µl total volume) were then left to dry overnight in a vacuum centrifuge (Royle *et al.*, 2008).

2.2.4.4 Ultra Performance Liquid Chromatography (UPLC) for the separation of the 2-AB labelled glycans

A high throughput platform for the separation of *N*-glycans from various samples using UPLC-HILIC with fluorescence detection with high efficiency was recently described, which greatly improves over HPLC based methods (Mariño *et al.*, 2010a; Saldova *et al.*, 2014a). HILIC separates both neutral and charged glycans in a single separation allowing for total glycan characterisation with single-column chemistry. Separation is based on the hydrophilic potential of the glycan, which is affected by many factors, including: size, charge, composition, structure, linkage, and oligosaccharide branching. Notably, the main advantage of HILIC is the ability to separate structural isomers, which is very difficult to achieve with other methodologies (Saldova *et al.*, 2014a).

The 2-AB derivatized *N*-glycans (section 2.2.3.3) were separated by UPLC with fluorescence detection on a Waters Acquity UPLC H-Class instrument consisting of a binary solvent manager, sample manager and fluorescence detector under the control of Empower 3 chromatography workstation software (Waters). The

hydrophilic interaction liquid chromatography (HILIC) separations were performed using a Waters Ethylene Bridged Hybrid (BEH) Glycan column, 150 x 2.1 mm i.d., 1.7 μ m BEH particles, with 50 mM ammonium formate, pH 4.4, as solvent A and acetonitrile (MeCN) as solvent B. This system allows for separation of *N*-glycans with a decrease in run time and greatly increased resolution due to the sub 2 μ m stationary phase. The 30 minute method was used with a linear gradient of 30-47% (v/v) with buffer A at 0.56 mL/minute flow rate for 23 minutes followed by 47-70% (v/v) A and finally reverting back to 30% (v/v) A to complete the run (Salдова *et al.*, 2014). An injection volume of 10 μ L sample prepared in 70% (v/v) MeCN was used throughout. Samples were maintained at 5°C prior to injection, while separation was carried out at 40°C. For detection the fluorescence excitation/emission wavelengths were $\text{ex} = 330 \text{ nm}$ and $\text{em} = 420 \text{ nm}$, respectively. The system was calibrated using an external standard of hydrolyzed and 2AB-labeled glucose oligomers to create a dextran ladder, as described previously (Royle *et al.*, 2006).

2.2.4.5 Weak Anion-Exchange High-Performance Liquid Chromatography (WAX HPLC) for the determination of sialic acid content of IgY

Weak anion exchange (WAX) HPLC was used to separate the *N*-glycans by charge and was carried out as detailed in Royle *et al.* (2006), with a fetuin *N*-glycan standard as reference. The principle behind this method involves the application of labelled glycans to an anion exchange column and elution with a salt gradient. Negatively charged sugars (such as sialic acid) will bind to the anion exchanger, while neutral glycans will pass through the column. The separated glycans (sialylated glycans) are then compared to a fetuin standard and classified as

mono-, di-, tri-, and tetrasialylated glycans (J. J. Kattla et al., 2011). WAX HPLC was performed using a Vydac 301VHP575 7.5 × 50 mm column (Anachem) on a 2695 Alliance separations module with a 2475 fluorescence detector (Waters), which was set with excitation and emission wavelengths of 330 and 420 nm, respectively (Royle *et al.*, 2008). Solvent A was 0.5 M formic acid adjusted to pH 9.0 with ammonia solution and solvent B was 10% (v/v) methanol in water. Gradient conditions were as follows: a linear gradient of 0 to 5% (v/v) A over 12 minutes at a flow rate of 1 mL/minute, followed by 5–21% (v/v) A over 13 minutes and then 2–50% (v/v) A over 25 minutes, 80–100% (v/v) A over 5 minutes followed by 5 minutes at 100% A. Samples were injected in water. A fetuin *N*-glycan standard was used for calibration (Royle *et al.*, 2008).

2.2.4.6 Exoglycosidase digestion for the determination of monosaccharide sequence and linkage of the *N*-glycans present in the IgY

Exoglycosidase digestions were carried out on aliquots of the total labelled *N*-glycan pools in accordance with methods previously described by Royle *et al.* (2006). All exoglycosidase enzymes were obtained from Prozyme, San Leandro, CA, USA. The 2AB-labeled glycans were digested in a volume of 10 µL for 18 hours at 37 °C in 50 mM sodium acetate buffer, pH 5.5 (except in the case of jack bean α-mannosidase (JBM) where the buffer was 100 mM sodium acetate, 2 mM Zn²⁺, pH 5.0), using arrays of the following enzymes: *Arthrobacter ureafaciens* sialidase (ABS), 0.5 U/mL; *Streptococcus pneumoniae* sialidase (NAN1), 1 U/mL; bovine testes β-galactosidase (BTG), 1 U/mL; *Streptococcus pneumoniae* β-galactosidase (SPG), 0.4 U/mL; bovine kidney α-fucosidase (BKF), 1 U/mL; β-N-acetylglucosaminidase cloned from *S. pneumonia*, expressed in *Escherichia coli*

(GUH), 8 U/mL and JBM (EC), 50 U/mL. After incubation, enzymes were removed by filtration through a 10 kDa protein-binding EZ filters (Millipore Corporation). *N*-Glycans were then analyzed by UPLC and WAX HPLC.

2.2.4.7 Ultra Performance Liquid Chromatography-Fluorescence-Mass Spectrometry (UPLC-FLR-MS)

The UPLC-FLR-MS analysis was carried out by Dr. Silvia Millán Martín (NIBRT, Dublin, Ireland) using the following methods.

For UPLC-FLR-QTOF MS analysis lyophilised IgY samples were reconstituted in 3 μ l of water and 9 μ l acetonitrile. Online coupled fluorescence (FLR)-mass spectrometry detection was performed using a Waters Xevo G2 QTof (YCA 219) with Acquity® UPLC (Waters Corporation, Milford, MA, USA) and BEH Glycan column (1.0 x 150 mm, 1.7 μ m particle size).

For MS acquisition data the instrument was operated in negative-sensitivity mode with a capillary voltage of 1.80 kV. The ion source block and nitrogen desolvation gas temperatures were set at 120 °C and 400 °C, respectively. The desolvation gas was set to a flow rate of 600 L/h. The cone voltage was maintained at 50V. Full-scan data for glycans were acquired over *m/z* range of 450 to 2500. Data collection and processing were controlled by MassLynx 4.1 software (Waters Corporation, Milford, MA, USA).

The fluorescence detector settings were as follows ex = 330 nm, em = 420 nm; data rate was 1pts/second and a PMT gain = 10. Sample injection volume was 10 μ L. The flow rate was 0.150 mL/minute and column temperature was maintained at 60 °C; solvent A was 50 mM ammonium formate in water (pH 4.4) and solvent

B was acetonitrile. A 40 minute linear gradient was used and was as follows: 28% (v/v) A for 1 minute, 28-43% (v/v) A for 30 minutes, 43-70% (v/v) A for 1 minute, 70% (v/v) A for 3 minutes, 70-28% (v/v) solvent A for 1 minute and finally 28% (v/v) A for 4 minutes.

Samples were diluted in 75% (v/v) acetonitrile prior to analysis. The weak wash solvent was 80% (v/v) acetonitrile and the strong wash solvent was 20% (v/v) acetonitrile. To avoid contamination of Mass Spec system, flow was sent to waste for the first 1.2 minutes and after 32 minutes.

2.2.4.8 Molecular Modelling of IgY

Molecular modelling of IgY was carried out by Dr. Mark Wormald (Oxford University, England). Modelling was performed on a Silicon Graphics Fuel workstation using InsightII and Discover software (Accelrys Inc., San Diego, USA). Figures were produced using the program Pymol (Schrodinger, LLC, 2010). Protein structures used for modelling were obtained from the Protein Data Bank (PDB) database (Berman *et al.*, 2000). The peptide structure of chicken IgY was based on the crystal structures of human IgE domains Cε2-4 (Wan *et al.*, 2002) and human IgG Fab domain (Stanfield *et al.*, 1990). The *N*-linked glycans to model on IgY were chosen on the basis of the peaks with the largest areas in the glycan profile (Appendix Four).

2.2.5 Differentiation of indolent, significant and aggressive prostate cancer by robotic high-throughput *N*-glycan profiling

To investigate if differential glycosylation patterns could distinguish between 3 cohorts of PCa patients, namely, indolent, significant and aggressive PCa. Whole

serum *N*-glycan profiling was carried out on 117 prostate cancer patients' sera using a robotised, high-throughput analysis platform for glyco-profiling which utilises ultra-performance liquid chromatography (UPLC) to obtain high resolution separation of *N*-linked glycans released from the serum.

On the basis of the technologies used for *N*-glycan profiling of IgY (section 2.2.4), the UPLC analytical workflow was improved on by developing an automatable sample preparation workflow, implemented on a liquid handling platform. The system, originally developed for glycoprofiling of immunoglobulin G antibodies (IgG) (Figure 2.5) uses a robotic liquid-handling workstation, allowing the preparation of 96 samples (or multiples thereof) in 22 hours with excellent reproducibility and can greatly facilitate serum biomarker discovery (Stöckmann *et al.*, 2013, 2015).

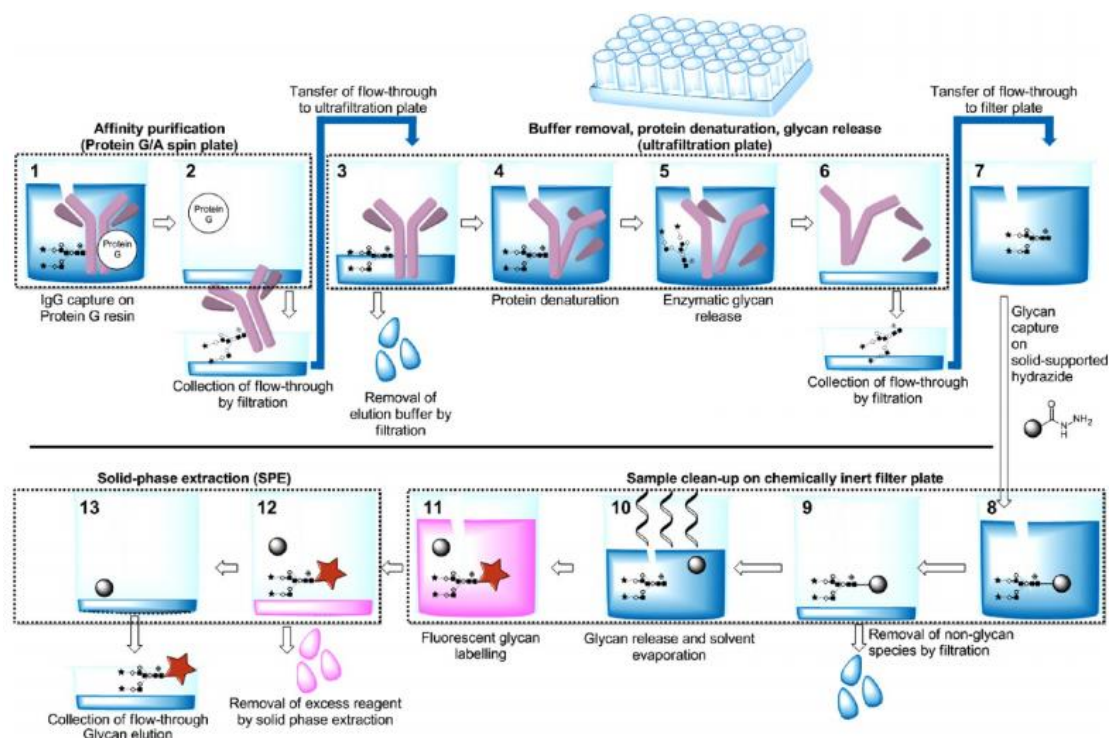


Figure 2.5: Automated high-throughput IgG-antibody glyco-profiling platform. The sample preparation workflow consists of glycoprotein affinity purification, protein denaturation, enzymatic glycan release, glycan immobilisation on solid supports, removal of contaminants, glycan release, labelling, and solid-phase extraction (Taken from Stöckmann *et al.*, 2013).

2.2.5.1 Serum Samples

Samples were collected with consent from prostate cancer patients following a standard operating procedure, which is part of the Prostate Cancer Research Consortium BioResource. Ethical consent was granted from respective Hospital ethics committees of the consortium. Blood samples (10 mL) were collected into anticoagulant-free tubes. Samples were coded and transported on ice to the

laboratory. The tubes were centrifuged at 900 x *g* at 20°C for 10 minutes within a 30 minute time-frame. Serum from each patient sample was then collected, aliquoted and stored at -80°C until time of analysis. Each serum sample underwent no more than three freeze/thaw cycles prior to analysis. A total of 117 patients (indolent [41], significant [32] and aggressive [44]) were included in this study. Epstein's criteria was used to define patients as having indolent, significant or aggressive PCa (Epstein *et al.*, 1994). Indolent PCa was defined as tumour volume <0.5cm³, organ-confined disease and no Gleason patterns 4 or 5. Significant disease was defined as tumour volume >0.5 cm³, organ-confined disease and Gleason pattern 4. Aggressive PCa was defined as Gleason patterns 4 or 5 and non-organ-confined disease.

2.2.5.2 Glycoprotein denaturation and glycan release

Serum samples (5 µL) were added to a 96-well V-bottom plate (in duplicate) and placed in a Hamilton Robotics StarLet liquid-handling platform. Denaturation buffer (55 µL per well, 100 mM ammonium bicarbonate, 12 mM dithiothreitol) was dispensed into the 96-well V-bottom plate (Greiner Bio-One) containing the PCa serum samples, which was placed on a robotic heater shaker, fully covered and insulated with an anti-evaporation lid. This assembly was incubated at 65°C with agitation at 700 rpm for 30 minutes. After cooling to room temperature for 10 minutes, a iodoacetamide solution (120 mM, 10 µL per well) was added, the plate was covered with an anti-evaporation lid and incubated at room temperature with agitation at 700 rpm for 30 minutes. A trypsin solution (40, 000 U/mL) in water, 10 µL per well) was added, the plate was sealed, placed on the robotic heater shaker and fully covered with an anti-evaporation lid. Incubation

was performed at 37°C with agitation at 700 rpm for 120 minutes. The temperature was increased to 105°C and incubation was continued for a further 10 minutes. The plate was left to cool to room temperature, centrifuged briefly (2,700 x *g* for 30 seconds) and the seal was removed. PNGase F (Prozyme R Glyco N-Glycanase R, 10 µL per well, 0.5 mU in 1 M ammonium bicarbonate, pH 8.0) was added and the plate was covered with an anti-evaporation lid and incubated at 40°C with agitation at 700 rpm for 2 hours.

2.2.5.3 Hydrazide-mediated glycan ‘clean-up’

The wells of a 96-well chemically inert filter plate (Millipore Solvinert, hydrophobic polytetrafluoroethylene membrane, 0.45 µm pore size) were washed with 100 µL methanol (MeOH) and ultralink hydrazide resin (40 µL suspension in water, Thermo Scientific) was subsequently dispensed to each well. The resin was sequentially washed with (100 µL per well) MeOH, H₂O and acetonitrile (MeCN) and the plate was placed on a heater (80°C, 8 minutes) to seal the membranes. One hundred and eighty microliters MeCN/acetic acid, (98:2) was added to the resin, followed by 20 µL of the glycan solution. The filter plate was incubated with shaking at 700 rpm at 80°C for 60 minutes. Fifty microlitres MeCN/acetic acid (98:2) was then added and shaking was continued at the same temperature for 10 minutes to disrupt resin aggregates. The resin was washed sequentially with MeOH (3 x 200 µL), 2M guanidine (2 x 200 µL), H₂O (2 x 200 µL), triethylamine/MeOH (1:99, 2 x 200 µL) and MeOH (2 x 200 µL per well). Fresh MeOH (180 µL) and acetic anhydride (20 µL) were added and the plate was incubated for 30 minutes with agitation at 600 rpm, at 21°C. Excess reagent was removed by filtration and the resins were twice washed sequentially with MeOH,

H₂O and MeCN, Acetic acid/MeCN (2:98, 175 µL) and H₂O (25 µL) were sequentially added and the plate was incubated at 70°C with agitation at 700 rpm for 90 minutes. Fluorescent labeling mix (50µL, 350 mM 2-aminobenzamide, 1 M sodium cyanoborohydride in acetic acid/dimethyl sulfoxide (30:70)) was dispensed into each well and the plate was incubated at 65 °C with agitation at 800 rpm for 120 minutes.

2.2.5.4 Glycan solid phase extraction

The labelling reaction was quenched by the addition of 550 µL MeCN/H₂O (95:5). The suspension was transferred to a 2 mL collection plate, the beads were left to settle, 200 µL of the supernatant was aspirated and dispensed back into the filter plate. After extensive mixing the suspension was transferred back into the collection plate. This cycle was repeated to ensure a quantitative transfer of the resins. HyperSep Diol Solid Phase Extraction (SPE) cartridges (ThermoScientific) were washed with 500 µL MeCN/H₂O (95:5), 1 mL H₂O and 1 mL MeCN/H₂O. Next, the quenched reaction supernatant was transferred onto the SPE cartridges. An incubation period of 5 minutes typically led to complete drainage of the solvent by gravity. The SPE cartridges were washed three times with 750 µL MeCN/H₂O (95:5). A collection plate was placed inside the robotic vacuum manifold and the SPE cartridges were washed twice with 200 µL H₂O/MeCN (80:20), with an incubation period of 5 minutes between each wash step. The samples were concentrated to dryness in a vacuum evaporator and then dissolved in 30 µL MeCN/H₂O (70:30). A portion (33%) of the resuspended sample was then analysed by UPLC.

2.2.5.5 Ultra Performance Liquid Chromatography (UPLC) of the 2-AB labelled glycans

The 2-AB derivatized *N*-glycans were separated by HILIC-ULPC, as described in section 2.2.3.4.

2.2.5.6 Statistical analysis

Statistical analysis was carried out by Mr. Keefe Murphy (PCRC biostatistician, University College Dublin). Briefly, all 50 peaks were examined individually to explore baseline differences in expression using one way analysis of variance (ANOVA). Tukey's post-hoc honestly significant difference (HSD) was performed to calculate the p-values. Boxplots were used to illustrate the distribution of each biomarker by significance.

The peaks were then considered in combination (in a multivariate predictive regression context) to see which panel of variables would yield the best predictive performance. This was arrived at using logistic regression analysis and the best model fit was ensured by forward and backwards stepwise AIC selection. Model diagnostics was then carried out to assess the performance of this biomarker panel. The discriminant ability of both the 'significant peaks' and the 'combined biomarker panel' were determined by means of ROC curves and AUC values.

Chapter Three

Screening and analysis of an avian anti-fPSA library

3.1 Introduction

In this chapter screening of pre-existing scFv and fab clones from an avian fPSA library for the isolation of an nM affinity anti-fPSA antibody fragment is detailed. Evaluation of the binding characteristics of this avian anti-fPSA scFv fragment by surface plasmon resonance using Biacore 3000 instrumentation is described in detail. Finally, discovering the unique structural attributes of this avian antibody fragment was explored by X-ray crystallographic techniques.

Antibodies are ideal bio-recognition elements and are likely to remain the affinity molecules of choice for wide variety of analytical, biochemical, and medical approaches. Our advanced understanding of the factors that contribute to antibody specificity arises from the knowledge gained from multiple disciplines including; structural biology, immunogenetics, cellular and molecular biology and immunoinformatics. Once an antibody fragment is successfully isolated, expressed and purified it also needs to be extensively characterised. Binding assays of the antibody to the cognate antigen need to be defined, ideally epitope mapping should be conducted and most importantly, the affinity (dissociation constant) the antibody has for the antigen needs to be determined.

In the past decade surface plasmon resonance (SPR)-based optical biosensors, such as Biacore systems have evolved as standard tools for antibody characterisation and binding interaction studies. The Biacore 3000 instrument uses SPR technology for biomolecular interaction analysis in real time without the need for labels. This instrument provides a comprehensive characterisation of binding events and kinetics analysis can also be conducted with relative ease (Wilson, 2002; Leonard *et al.*, 2011).

Much research is focused on the precise nature of the binding interactions between an antibody and its cognate antigen. There are a numerous different methods that are commonly used to map the fine details of these interactions. The rules that govern antibody-antigen recognition, including the molecular structures within a target antigen that make specific contacts with an antibody (referred to as antibody or B-cell epitopes) and the molecular determinants within the antibody structure that have specific interactions with the antigen epitope (termed paratopes), are only now becoming well characterised due to an increase in the availability of three-dimensional structures (Sela-Culang *et al.*, 2013; Abbott *et al.*, 2014). X-ray crystallography is undoubtedly the method that is most commonly used to determine the precise sites of interaction between an antibody and its antigen.

A deep understanding of the role that each structural element plays in antibody-antigen interactions is crucial both to evolve our understanding of immunity, and is essential for successful engineering of unique high affinity binders. Developments in structure-based design strategies have aided in vaccine design and development of antivirals for the treatment of the human immunodeficiency virus (HIV) and the influenza (HA) virus (Lee and Wilson, 2015). The structural characterisation of both HIV-1 Envelope protein and influenza HA in complex with a variety of diverse antibodies has allowed for the identification of the key sites of vulnerability on these viruses. The structural information describing the interaction of these antibodies with the antigens is now publically available, allowing for rational, structure-guided antibody engineering (Julien *et al.*, 2012).

Crystal structures enable very accurate determination of key interactions between individual amino acids from both side chain and main chain atoms in both the

epitope and paratope. Amino acids that are within 4 Å of each other are generally considered to be contacting residues (Abbott *et al.*, 2014). The number of residues that are in contact with antigens is thought to differ in antibodies recognising peptides, haptens and proteins (Finlay and Almagro, 2012; Raghunathan *et al.*, 2012). The high-quality integrated knowledge resource specialising in immunogenetics data of antibodies-The Immunogenic Database (IMGT)- provides access to sequence, genome and structural data on antibodies and related proteins of the immune system (Lefranc, 2014). Raghunathan and colleagues mined this rich source of information along with the three-dimensional structures available at the PDB and analysed 140 unique high-resolution (<3 Å) antibody structures, including 55 in complex with proteins, 46 with haptens and 39 with peptides to develop and deeper insight into the structural architecture of the antigen-binding site (Raghunathan *et al.*, 2012). These authors successfully determined length variations of the hypervariable loops, number of contacts with antigen, frequency and location of antigen-contacting residues and the type of amino acid residues interacting with proteins, haptens and peptides (Raghunathan *et al.*, 2012).

Information such as that published by Raghunathan and colleagues enhances our understanding of the structure–function relationship in antibodies while providing valuable guidance to design libraries for antibody discovery. There is, however, a limited understanding of the structural attributes of avian antibodies. As highlighted in Chapter Three of this thesis, chicken antibodies are now receiving significant recognition as high-quality, high-affinity antibodies that can be generated to a wide range of antigens for a number of therapeutic, diagnostic and biotechnological applications.

Extensive examination of the amino acid characteristics of the chicken repertoire has provided a significant insight into mechanisms employed by the avian immune system (Wu *et al.*, 2012; Conroy *et al.*, 2014). The lack of avian antibody crystal structures in the PDB has limited our understanding of the structural consequences of these uniquely chicken features. To date only three structures are available for chicken antibody fragments (Shih *et al.*, 2012; Conroy *et al.*, 2014). This chapter focuses on the isolation and characterisation of an avian anti-fPSA scFv in an attempt to further our knowledge of chicken antibody fragments and develop antibodies for improved prostate cancer diagnostics.

3.2 Results

3.2.1 Identification of high affinity anti-fPSA scFv

Over 1,000 anti-fPSA scFv and Fab clones from an avian anti-fPSA library (generated by Dr. Conor Hayes, in DCU) were screened by monoclonal ELISA (section 2.2.2.1) and clones with an absorbance value of greater than 0.3 (Figure 3.1) were taken forward for kinetic characterisation on the Biacore 4000. These clones were screened using a previously immobilised CM5 dextran chip functionalised with approximately 6,000 RU of anti-HA epitope by EDC-NHS coupling chemistry over spots 1, 2, 4 and 5. Anti-fPSA scFv and Fab lysates were diluted 1:3 (in duplicate) in HBPS-EP⁺ and analysed in a capture approach over the anti-HA surface at 25°C. The data was evaluated using the BiaEvaluation software using the '2 over 2' template and curves were fitted with a 1:1 Langmuir binding equation (Figure 3.2).

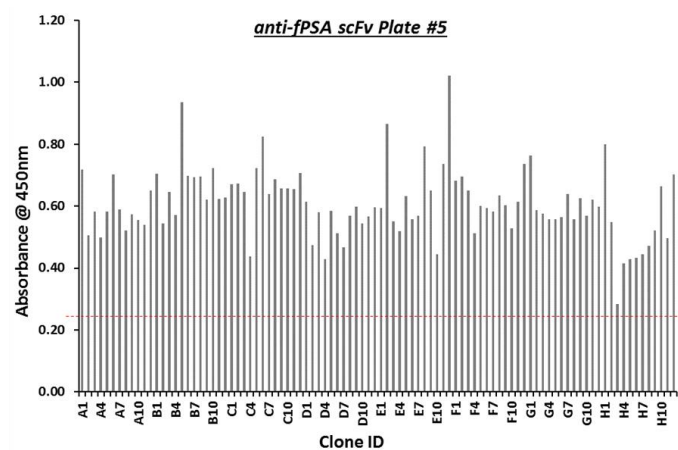
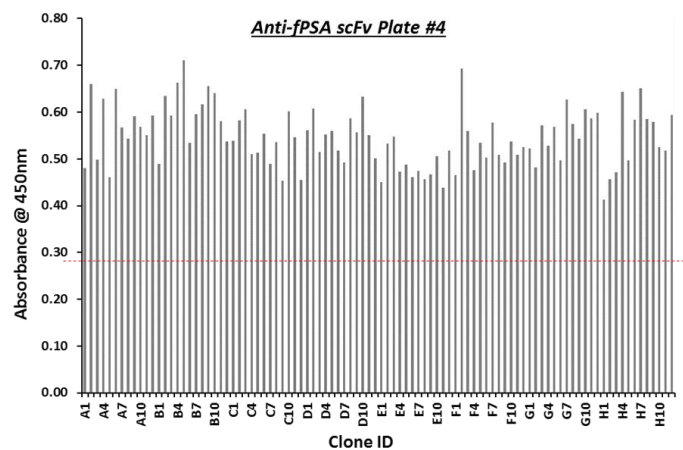
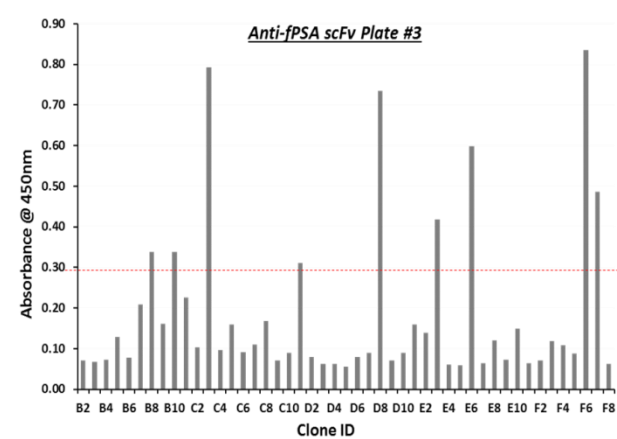
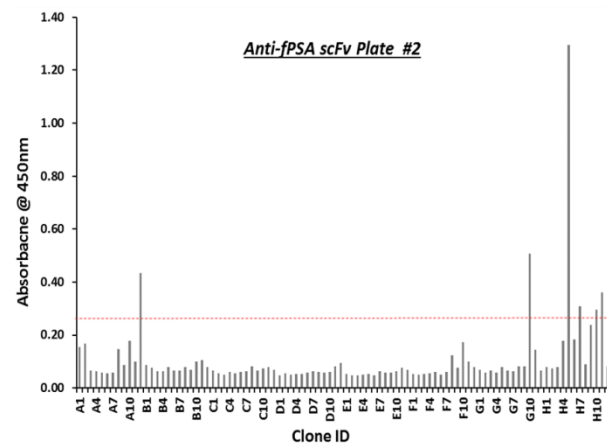
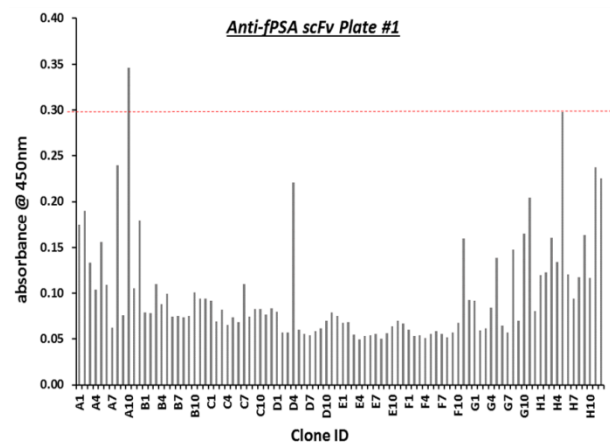


Figure Legend on Page 172

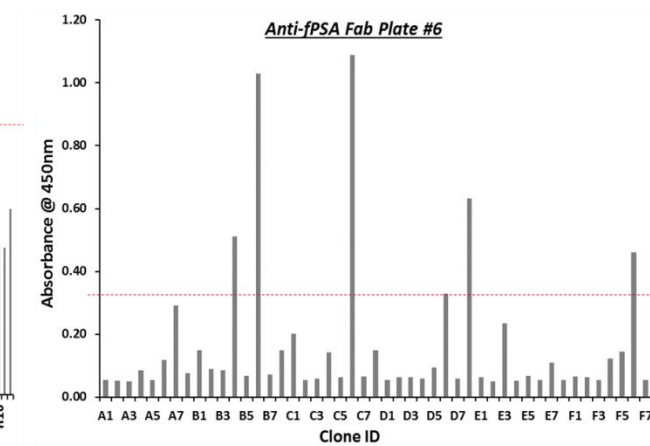
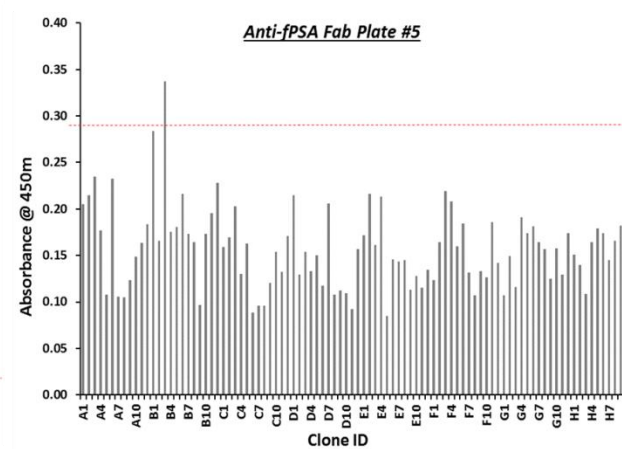
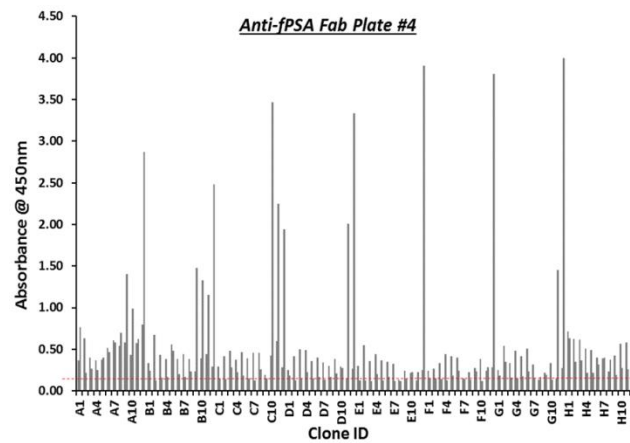
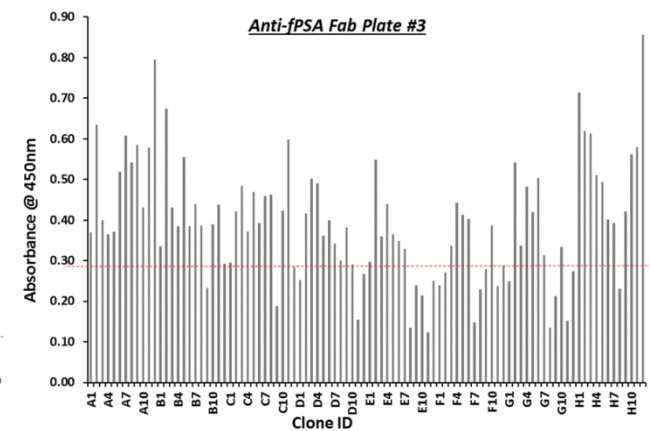
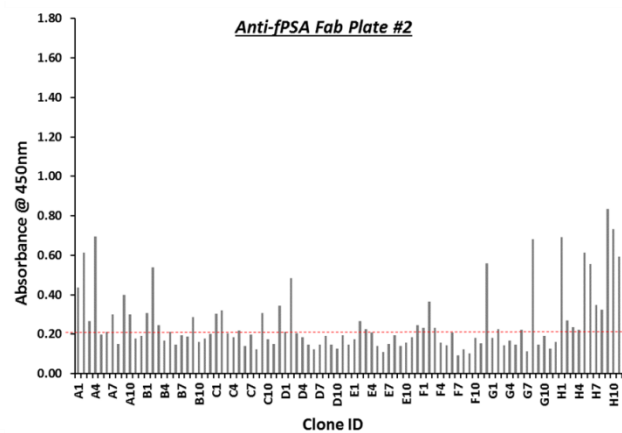
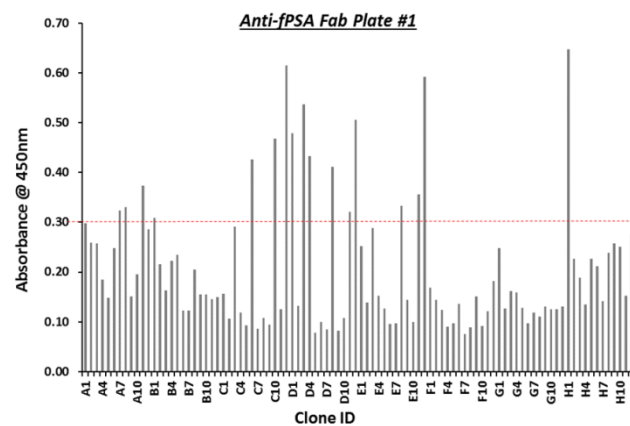


Figure Legend on Page 172

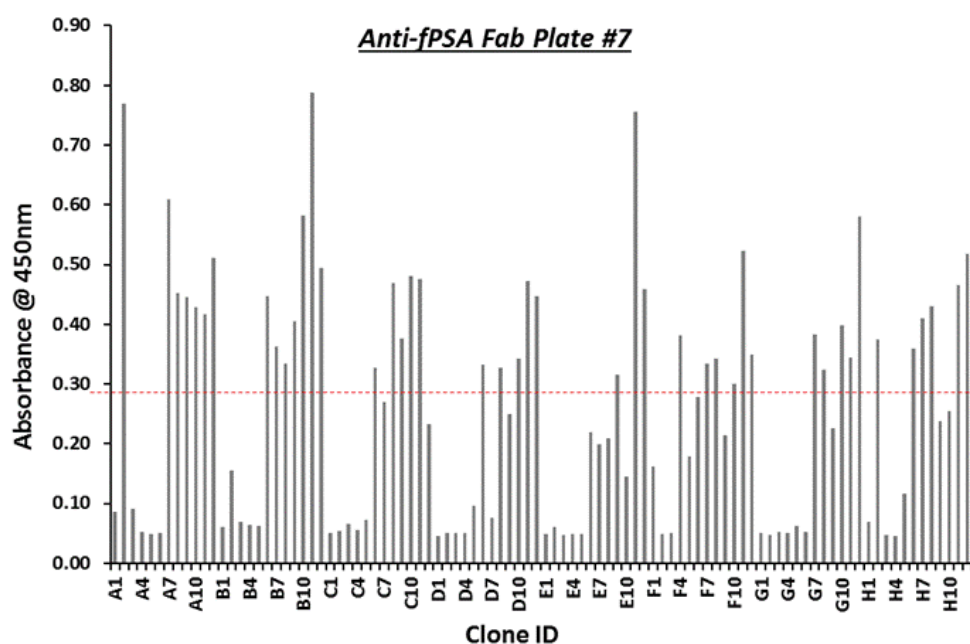


Figure 3.1: *Anti-fPSA scFv monoclonal ELISA's*. Over 1,000 anti-fPSA scFv and Fab antibody clones were assessed for reactivity to fPSA by monoclonal ELISA. Expressed antibody lysate was diluted 1 in 2 in PBSTM and added to a fPSA-coated plate. Specifically bound clones were then detected with an HRP-labelled anti-HA polyclonal antibody, TMB was added, the reaction stopped with 10% (v/v) HCl, and the absorbance was read at 450 nm. All clones above the 'cut-off' of 0.3 were taken for further characterisation on the Biacore 4000.

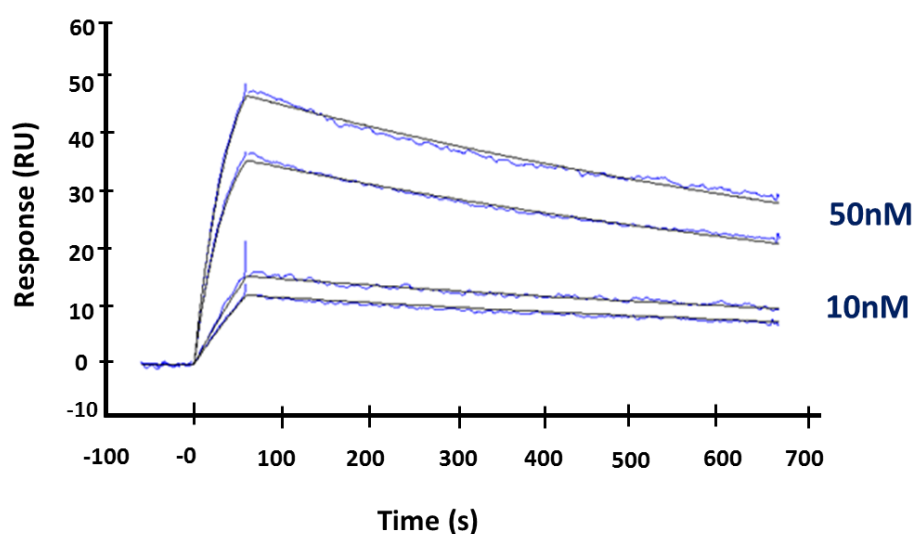


Figure 3.2: '2 over 2' kinetic analysis of scFv B8. The surface was composed of a polyclonal anti-HA antibody on spots 1, 2, 4 and 5. Anti-fPSA scFv and Fab lysates were diluted 1:3 (in duplicate) in HBPS-EP⁺ and analysed in a capture approach over the anti-HA surface at 25°C. The diluted lysates were captured on spots 1, 2, 4 and 5 and two concentrations of fPSA (10 nM and 50 nM diluted in HBS-EP⁺) were then passed over the captured antibody fragments and referenced against the unmodified surface on spot 3. The data was evaluated using the BiaEvaluation software using the '2 over 2' template and curves were fitted with 1:1 Langmuir binding equation. Analysis identified scFv B8 as a single digit nano-molar affinity antibody (K_D 1.25×10^{-9} M).

3.2.2 Optimisation of expression of scFv B8

Once the clone was identified as having high affinity for fPSA it was necessary to express this antibody fragment. An overview of the main conditions that were optimised for the expression of scFv B8 is detailed in Table 3.1.

Table 3.1: Conditions for optimisation for recombinant antibody expression

Temperature	<i>Lowering the expression temperature routinely improves the solubility of recombinant proteins, and can result in a reduction in degradation. Once induced, temperatures of 18-30°C should be assessed.</i>
Time	<i>At lower temperatures the rate of protein production also decreases and, thus, longer induction times are necessary to obtain sufficient protein. Time points should be taken at 2- 16 hours.</i>
Inducer Concentration	<i>In addition to lowering the temperature, a reduction in transcription rate can also be achieved by lowering the concentration of the induction agent. Lowering the IPTG concentration can enhance the production of soluble protein.</i>

Various concentrations of IPTG (0 mM, 0.05 mM and 0.1 mM) were assessed at both 25°C and 30°C and samples were taken at time points at 0, 2, 4 and 16 (overnight) hours for analysis by SDS-PAGE and WB. Once transferred to the nitrocellulose membrane the antibody fragment was detected with a HRP tagged anti-HA antibody and the reaction developed with liquid TMB for membranes. The determination of optimal expression conditions proved very difficult to achieve, as can be seen in Figure 3.3. However, optimal conditions of 0.1 mM IPTG for 4 hours at 25°C were employed for large-scale expression and subsequent purification.

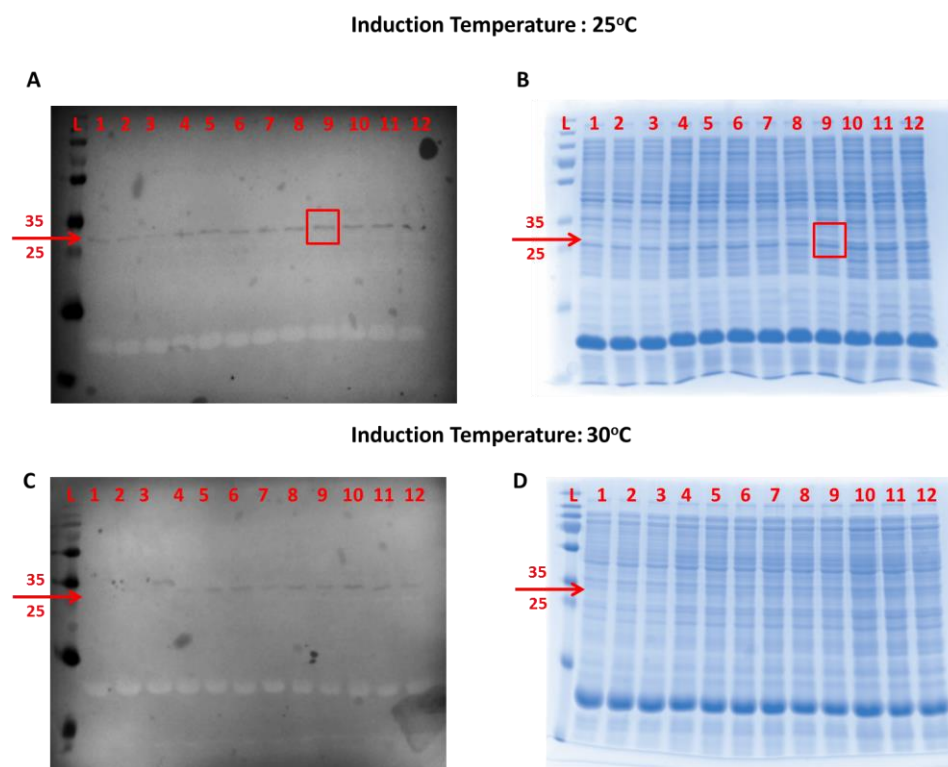


Figure 3.3: Optimisation of expression of scFv B8. Varied conditions were optimised to ascertain the optimal expression conditions for anti-fPSA scFv B8. Although expression levels are poor, 0.1 mM IPTG for 4 hours at 25°C was chosen as the optimal conditions (Red box; A and B). In the SDS (B & D) the presence of extracellular *E. coli* proteins are also expressed at high levels for the overnight sample. Thus, overnight expression was not deemed ideal for scFv B8. Low levels of expression were seen for all IPTG concentrations at each time point in the 30°C study, as there was a very faint band at approximately 25kDa. A: Western blot analysis of scFv B8 expression study at 25°C, probed with anti-HA antibody; B: SDS-PAGE for optimisation at 25°C; C: Western blot analysis of scFv B8 expression study at 30°C, probed with anti-HA antibody; D: SDS-PAGE for optimisation at 30°C. Legend; L: Ladder, 1: Time point = 0 hrs, 0 mM IPTG; 2: Time point = 0 hrs, 0.05 mM IPTG; 3: Time point = 0 hrs, 0.1 mM IPTG; 4: Time point = 2 hrs, 0 mM IPTG; 5: Time point = 2 hrs, 0.05 mM IPTG; 6: Time point = 2 hrs, 0.1 mM IPTG; 7: Time point = 4 hrs, 0 mM IPTG; 8: Time point = 4 hrs, 0.05 mM IPTG; 9: Time point = 4 hrs, 0.1 mM IPTG; 10: Time point = o/n, 0 mM IPTG; 11: Time point = o/n, 0.05 mM IPTG; 12: Time point = o/n, 0.1 mM IPTG. Red arrow indicates scFv fragment at 28 kDa.

3.2.3 Purification of scFv B8 by Immobilised Metal Affinity Chromatography (IMAC)

Using the optimal conditions identified in section 3.2.2 a 1 L culture of scFv B8 was prepared and lysed, as detailed in section 2.2.2.3. Ni-NTA resin was applied to a 10 mL chromatography column and once settled, the resin bed was equilibrated with 2 column volumes of PBS followed by one column volume of OSB-A buffer. The expressed antibody lysate was filtered and applied to the column (twice). Samples were collected at each step of the purification process and analysed by SDS-PAGE and WB. The eluted protein was concentrated and buffer exchanged against PBS/0.02% (w/v) NaN₃ using a 10,000 MWCO vivaspin column. The purified concentrated protein was determined to be 1 mg/mL using the NanoDrop spectrophotometer.

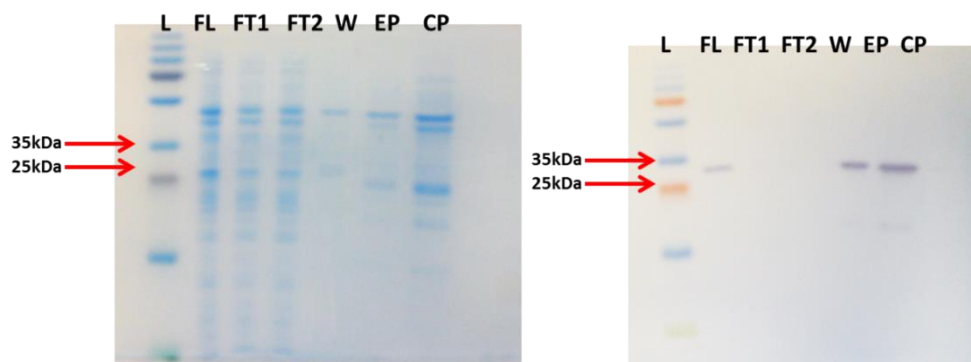


Figure 3.4: IMAC purification of scFvB8. A 1 L culture was prepared and expressed antibody lysate purified by IMAC. Samples were taken throughout the process for analysis by SDS-PAGE (left) and WB (right). Specific detection of the scFv fragment was performed by probing the WB membrane with a HRP-labelled anti-HA antibody. ScFv B8 was successfully purified to 1 mg/mL. L: Ladder; FL: Filtered lysate; FT1: Flow through 1; FT2: flow through 2; W: Wash; EP: Eluted protein and CP: Concentrated protein.

3.2.4 Anti-fPSA scFv B8 Sandwich ELISA Limit of Detection (LOD) Assay

Limit of Detection (LOD) is a well-established parameter used to characterise an analytical method. It is defined as the lowest analyte concentration likely to be reliably distinguished from the limit of the blank (LOB) and at which detection is feasible. Anti fPSA scFv B8 Limit of Detection (LOD) was determined in accordance with the Clinical and Laboratory Standards Institute's (CLSI) published guidelines "EP17, Protocols for Determination of Limits of Detection and Limits of Quantification" (Armbruster and Pry, 2008). The assay was carried out in sandwich ELISA format, with a total of 24 replicates to ensure confidence in the result (Table 3.2).

Table 3.2: Determination of anti fPSA scFv B8 Limit of Detection (LOD) by indirect ELISA

PSA Concentration	32	16	8	4	2	1	0.5	0.25	0.12	0.06	0.01	0
Replicate # 1	2.527	1.884	1.388	0.690	0.258	0.084	0.062	0.054	0.060	0.059	0.000	0.019
Replicate # 2	2.486	1.954	1.204	0.692	0.264	0.075	0.067	0.054	0.051	0.053	0.011	0.008
Replicate # 3	2.416	1.948	1.290	0.504	0.253	0.070	0.062	0.047	0.064	0.045	0.008	0.013
Replicate # 4	2.330	1.941	1.214	0.629	0.253	0.082	0.054	0.048	0.048	0.048	0.004	0.004
Replicate # 5	2.472	1.927	1.601	0.612	0.256	0.092	0.065	0.056	0.056	0.048	0.006	0.008
Replicate # 6	2.423	1.974	1.358	0.571	0.257	0.083	0.070	0.056	0.060	0.046	0.021	0.007
Replicate # 7	2.060	1.991	1.485	0.650	0.207	0.101	0.066	0.051	0.058	0.043	0.008	0.006
Replicate # 8	2.469	1.802	1.478	0.614	0.259	0.073	0.058	0.048	0.055	0.045	0.026	0.040
Replicate # 9	2.461	2.049	1.423	0.731	0.234	0.077	0.067	0.043	0.058	0.048	0.009	0.008
Replicate # 10	2.361	1.912	1.299	0.661	0.228	0.074	0.049	0.060	0.055	0.045	0.017	0.011
Replicate # 11	2.470	2.000	1.311	0.663	0.218	0.068	0.062	0.054	0.060	0.068	0.009	0.008
Replicate # 12	2.444	2.003	1.322	0.640	0.208	0.067	0.063	0.054	0.060	0.039	0.008	0.007
Replicate # 13	2.437	2.028	1.342	0.621	0.201	0.062	0.062	0.040	0.058	0.048	0.005	0.004
Replicate # 14	2.481	2.030	1.368	0.638	0.207	0.072	0.050	0.059	0.056	0.045	0.004	0.005
Replicate # 15	2.477	1.997	1.350	0.648	0.218	0.073	0.065	0.044	0.051	0.049	0.007	0.006
Replicate # 16	2.547	2.112	1.436	0.715	0.238	0.077	0.062	0.060	0.053	0.011	0.017	0.005
Replicate # 17	2.014	1.774	1.311	0.546	0.242	0.083	0.065	0.058	0.047	0.045	0.018	0.008
Replicate # 18	2.415	1.921	1.181	0.557	0.228	0.078	0.062	0.056	0.055	0.045	0.014	0.016
Replicate # 19	2.706	1.943	1.130	0.419	0.114	0.087	0.057	0.050	0.050	0.048	0.009	0.009
Replicate # 20	2.579	1.969	1.682	0.694	0.296	0.086	0.066	0.049	0.049	0.048	0.009	0.008
Replicate # 21	2.552	1.804	1.391	0.625	0.214	0.082	0.054	0.047	0.057	0.046	0.005	0.005
Replicate # 22	2.306	1.365	1.358	0.656	0.211	0.073	0.064	0.057	0.057	0.030	0.005	0.006
Replicate # 23	2.326	2.021	1.304	0.619	0.211	0.087	0.066	0.040	0.059	0.049	0.006	0.007
Replicate # 24	2.445	2.032	1.318	0.521	0.234	0.082	0.067	0.059	0.056	0.045	0.007	0.002
Average	2.425	1.933	1.356	0.621	0.229	0.079	0.062	0.052	0.055	0.046	0.010	0.009
Std. Dev	0.148	0.145	0.124	0.072	0.034	0.009	0.006	0.006	0.004	0.010	0.006	0.007

Determination of LOD

Limit of Blank (LoB) is the highest apparent analyte concentration expected to be found when replicates of a blank sample containing no analyte are tested;

$$\text{LoB} = \text{mean blank} + 1.645(\text{Std. Dev. blank})$$

Limit of Detection (LoD) is the lowest analyte concentration likely to be reliably distinguished from the LoB and at which detection is feasible. LoD is determined

by utilising both the measured LoB and test replicates of a sample known to contain a low concentration of analyte;

$$\text{LoD} = \text{LoB} + 1.645(\text{Std. Dev. low concentration sample}^*)$$

$$\text{LoB} = 0.009^* + 1.645(0.007^*) = 0.021$$

$$\text{LoD} = 0.021 + 1.645(0.006^*) = \mathbf{0.03}$$

$$\text{LoD of anti-fPSA scFv B8} = \mathbf{0.03 \text{ ng/mL}}$$

**Note: Samples figures used highlighted in red in table*

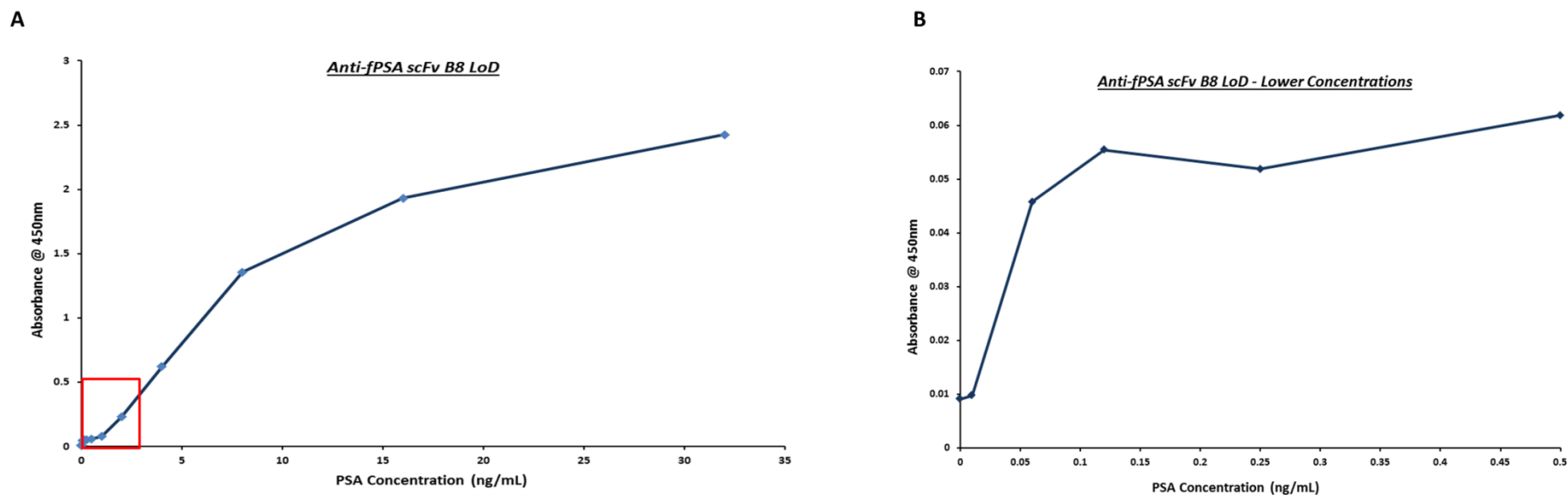


Figure 3.5: Limit of Detection (LoD) sandwich ELISA for scFv B8. Varying concentrations of fPSA (30 - 0 ng/mL) were captured on a goat anti-PSA polyclonal coated ELISA plate. Binding of anti-fPSA scFv B8 was then detected with the HRP-labelled secondary antibody, anti-HA followed by addition of TMB. The LoD of scFv B8 was determined to be 0.03 ng/mL by carrying out 24 replicates of the assay. A: Full graph (PSA concentrations 30 - 0 ng/mL); B: Lower PSA concentrations (0.5 - 0 ng/mL)- zoom in of red section in A.

3.2.5 Anti-fPSA scFv B8 cross reactivity assay

To ensure anti-fPSA scFv B8 was specific for fPSA and not cross reactive with cPSA a sandwich ELISA was performed, as described in section 2.2.2.4. Varying concentrations of cPSA were included in the assay, along with a control antigen (3 % BSA), to provide confidence in the results obtained. Anti-fPSA scFv B8 is highly specific for fPSA and does not cross react with cPSA or the control antigen, as can be seen in Figure 3.6.

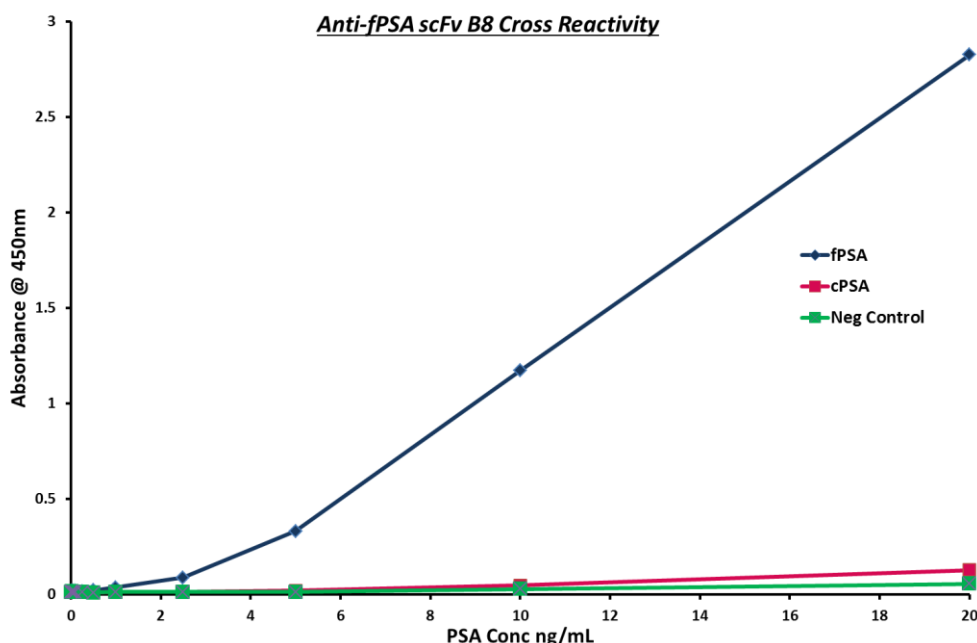


Figure 3.6: *Cross reactive Sandwich ELISA for scFv B8.* Varying concentrations of fPSA/cPSA (20 - 0 ng/mL) were captured with a polyclonal goat anti-PSA antibody coated ELISA plate. Binding of anti-fPSA scFv B8 was then detected with the HRP-labelled secondary antibody. The absorbance was read at 450nm. scFv B8 did not cross react with cPSA. The negative control was 3% (w/v) BSA which was substituted for PSA (Neg Control).

3.2.6 Kinetic analysis of scFv B8 using Biacore 3000

Initially a preconcentration study was carried out to identify the optimal buffer conditions for subsequent immobilisation, to increase the potential yield of immobilised ligand. Preparations of 25 µg/mL anti-HA antibody, diluted in a low ionic strength buffer (10 mM NaOAc) over a range of pH values (pH 4.0, 4.2, 4.4, 4.6, 4.8 and 5.0) was passed over the underivitisised chip surface (flow cell 2) and the degree of electrostatic binding monitored.

The principle behind using an array of pH buffers is that the protein carries a 'net' positive charge at pH values less than its isoelectric point causing an electrostatic interaction between the negatively charged dextran surface and the positively charged protein. This results in a higher coupling efficiency with a lower requirement for protein. The highest pH at which preconcentration of the protein onto the underivitisised surface is observed should be employed for the immobilisation (Leonard *et al.*, 2011).

Maximum potential binding for anti-HA was observed in NaOAc, pH 4.4 (Figure 3.7). Hence, this was utilised for immobilisation of 25 µg/mL anti-HA antibody to the surface of flow cell 2, using amine coupling chemistry.

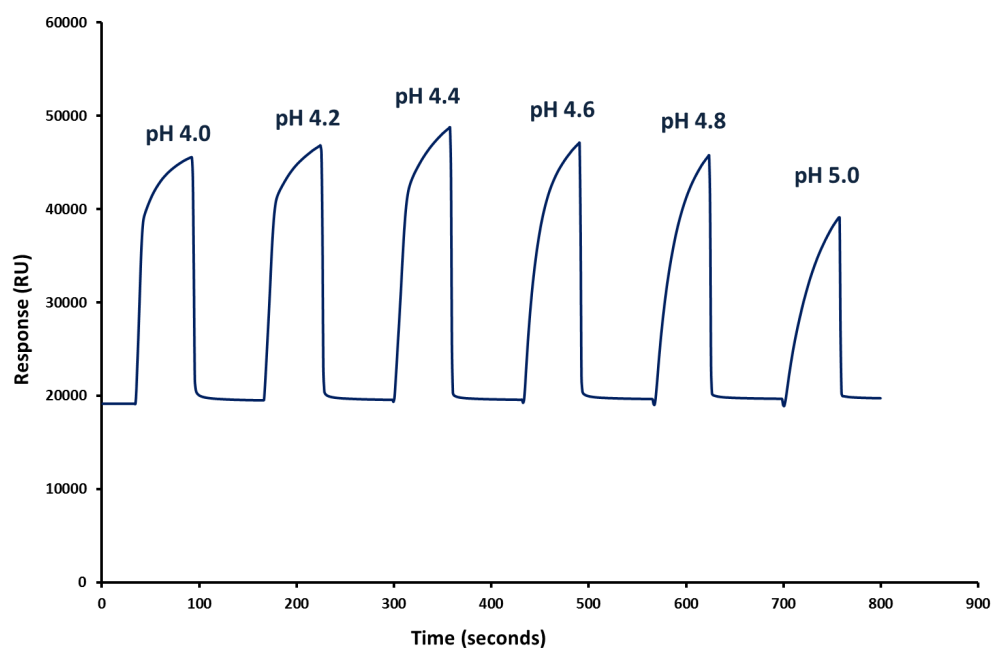


Figure 3.7: Preconcentration study for anti-HA antibody. Anti-HA antibody diluted in 10 mM sodium acetate solutions of pH of 4.0, 4.2, 4.4, 4.6, 4.8 and 5.0 was passed over the surface of a CM5 dextran chip at a flow rate of 10 μ L/minute. During preconcentrations studies there is an electrostatic interaction between the positive charged protein (due to its charge below the isoelectric point) and the negatively charged surface carboxylated groups. The optimal buffer conditions for the immobilisation of anti-HA antibody to the CM5 dextran surface was identified as NaOAc, pH 4.4, as the highest levels of binding was observed at this pH.

The immobilisation of a ligand to a CM5 dextran chip involves a number of steps described in section 2.2.3.4.2. The chip is activated with 400 mM EDC (1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride) and 100 mM NHS (N-hydroxysuccinimide), the ligand is subsequently passed over the surface and the chip surface is then deactivated with 1 M ethanolamine-hydrochloride to block and eliminate any potential non-specific binding (NSB) to the surface.

Immobilisation was performed by passing anti-HA polyclonal antibody diluted in 10 mM NaOAc to 25 µg/mL, pH 4.4, over the activated surface at a flow rate of 10 µL/minute for 15 minutes. Following capping and deactivation of any unreactive NHS groups, the surface was post conditioned with 20 mM NaOH (Figure 3.8). Approximately 19,000 RU were successfully immobilised to the surface of flow cell 2 of the CM5 dextran chip.

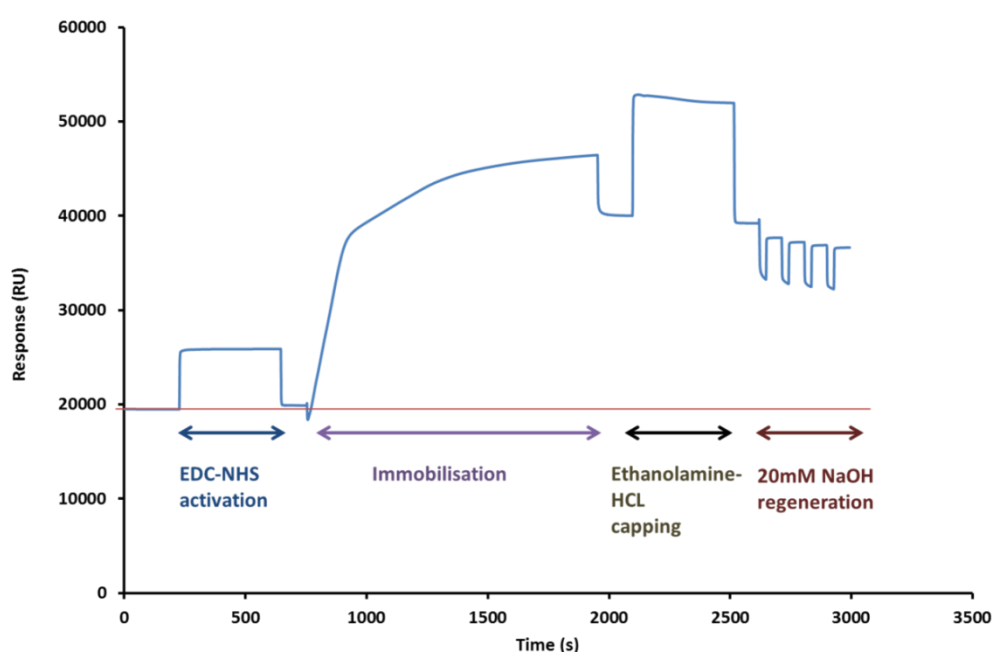


Figure 3.8: Immobilisation of anti-HA polyclonal antibody. The surface of a CM5 dextran chip was activated with an EDC/NHS solution and the anti-HA polyclonal antibody passed over the activated carboxyl group surface. The sample was injected for 15 minutes until the surface was fully saturated. Binding of the ligand occurs via the amino groups on the protein with the activated carboxyl groups on the dextran surface chip. After sufficient capturing of the ligand was achieved the remaining activated groups were deactivated with ethanolamine-HCl solution. The surface was regenerated using a 20 mM solution of NaOH. The level of anti-HA successfully immobilised to the surface was approximately 19,000RU.

ScFv B8 was kinetically characterised using the kinetic wizard functionality on the Biacore 3000 and the subsequent data generated was evaluated using BIAevaluation software, version 4.1.1. Purified scFv B8 was diluted 1 in 3,000 in running buffer and passed over the anti-HA surface at a flow rate of 30 $\mu\text{L}/\text{minute}$ for 3 minutes. fPSA was then passed over the surface in varying concentrations (6.25, 3.12 (in duplicate), 1.6, 0.8, 0.4 and 0 nM) for 3 minutes, and allowed to dissociate for 12 minutes. Regeneration of the surface was carried out by passing over 20 mM NaOH for 10 seconds at a flow rate of 30 $\mu\text{L}/\text{minute}$. The dataset from the reference flow cell (flow cell 1 - underivatized surface) was subtracted to remove any systemic effects, and the 0 nM fPSA concentration binding response (buffer only) was subtracted from each antigen response, thus double reference subtracting the dataset. The simplest 1:1 Langmuir binding model fit the dataset well with a global R_{max} and the curve (Figure 3.9, upper graph) was fitted with a χ^2 value of 0.482. This value gives information about the 'goodness' of the curve fit and for good qualitative data this value should be as low as possible. The residual plot (Figure 3.9, lower graph) shows the data points are distributed randomly above and below the zero line but are within acceptable limits.

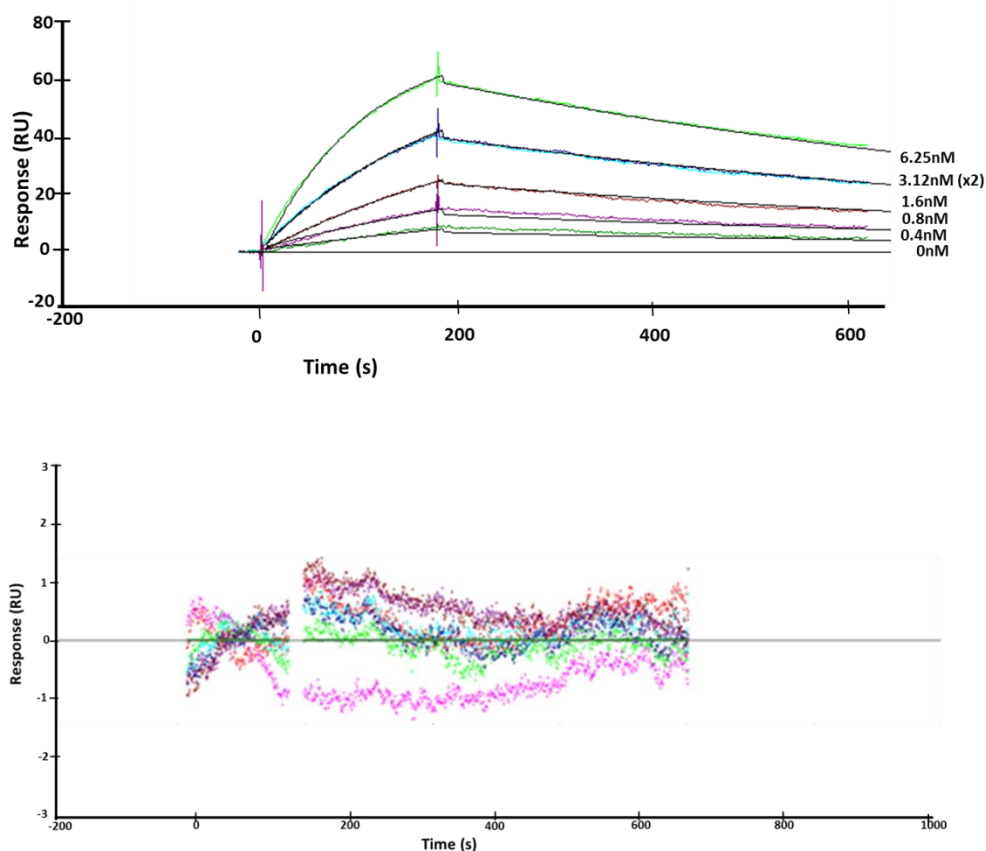


Figure 3.9: Kinetic interaction analyses of fPSA and scFv B8. Six concentrations of PSA (6.25 nM, 3.12 nM (x2) 1.6 nM, 0.8 nM, 0.4 nM and 0 nM) were used to fit the kinetics using 1:1 Langmuir binding model (global fit). The following kinetic parameters were derived; K_a : $1.36 \times 10^6 \text{ Ms}^{-1}$, K_d $1.15 \times 10^{-3} \text{ s}^{-1}$, and K_D ; 0.8×10^{-9} with a χ^2 (χ^2) value of 0.482.

3.2.7 Crystal Structure of scFv B8

In collaboration with the lab of Prof. James Whisstock, Monash University, Dr. Paul Conroy successfully solved the crystal structure of scFv B8 to 1.4 Å. This structure revealed unique binding attributes of avian antibodies, including a

CDRL1 canonical structure that has hitherto to be observed in mammalian repertoires and a rarely captured disulphide constrained CDRH3 (Figure 3.10).

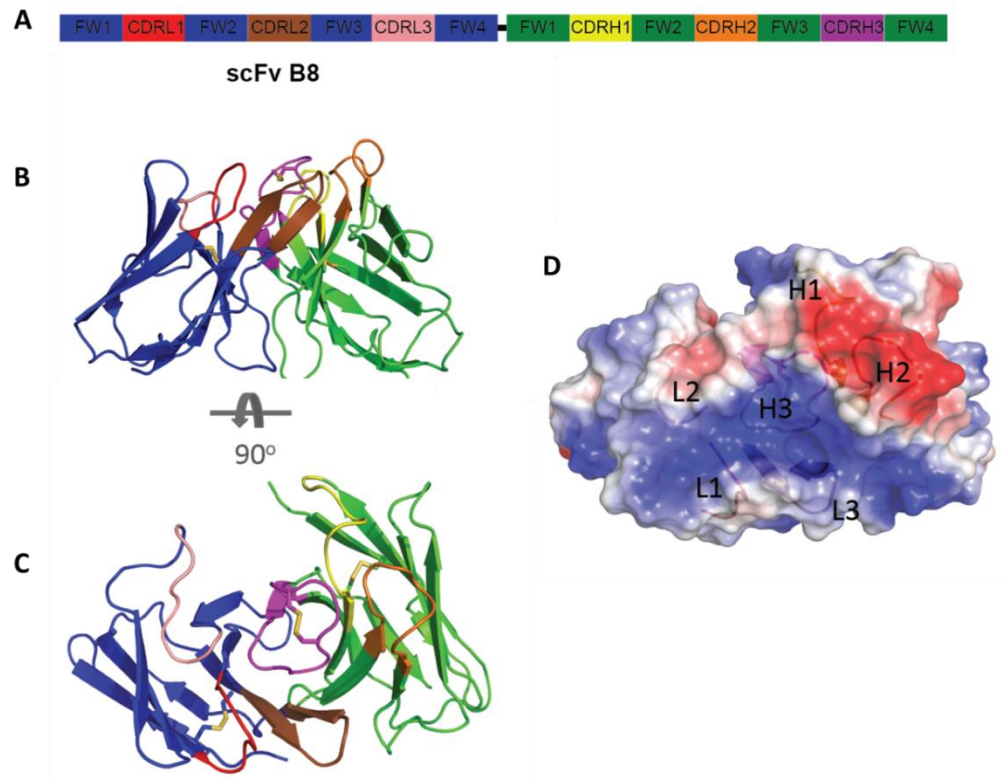


Figure 3.10; The crystal structure of anti fPSA scFv B8. (A) Diagrammatic representation of the scFv format coloured by chain (V_L: blue, V_H: green) with the CDR L1 (red), L2 (salmon), L3 (brown), H1 (yellow), H2 (orange) and H3 (magenta) shown. The flexible linker (black) connects the C-terminus of the V_L to the N terminus of the V_H, but is not modelled in the solved structure. (B) The CDR arrangement of scFv B8 is compact creating a ‘shelf-like’ arrangement for antigen binding. (C) The antibody binding sites as viewed from the antigen perspective. (D) The transparent electrostatic surface view of the scFv, from the antigen perspective, with the CDR loops (L1, L2, L3, H1, H2, H3) visible. The positively and negatively charged areas are indicated in blue and red, respectively. ScFv B8 has a predominantly positively charged surface (Figure adapted from Conroy *et al.*, 2014).

3.2.8 Expression and purification of scFvB8 Ser MT by Immobilised Metal Affinity Chromatography (IMAC) and Gel Filtration (GF)

In order to investigate the structural and binding impact of the conserved cysteine pair observed in scFv B8 a mutant variant was generated. scFv B8 plasmid DNA was used as a template to introduce a cysteine to serine double mutant at C97 and C100 in CDRH3 (C97/100S; scFv B8 Ser MT) (section 2.2.2.7). Successful generation of the mutant sequence was confirmed (Appendix two) and the mutant clone was subsequently expressed and purified as detailed in section 2.2.2.8.

The expressed antibody was purified in a 2-step process. IMAC purification was used initially (Figure 3.11 (A), followed by size exclusion chromatography (also referred to as gel filtration) (Figure 3.11 (B)). “Structural biology-grade” proteins must be produced using a purification process that results in the generation of an abundance of highly pure protein in order to determine the structure using X-ray crystallography (Kim *et al.*, 2011). Gel filtration is a basic chromatography technique that separates proteins with differences in molecular size and shape as they pass through gel filtration medium in a packed column, and is often used as a final polishing step following IMAC purification (Kim *et al.*, 2011). Gel filtration of scFv B8 Ser MT was carried out on an S 75/16/60 column on the AKTA Explorer FPLC system (GE Healthcare). The protein was homogenous, soluble, mono-dispersed, stable and estimated to be greater than 95% pure (as determined by SDS-PAGE, Figure 3.11 (B), hence was concentrated to approximately 16 mg/mL to take forward to coarse screening for crystallisation conditions.

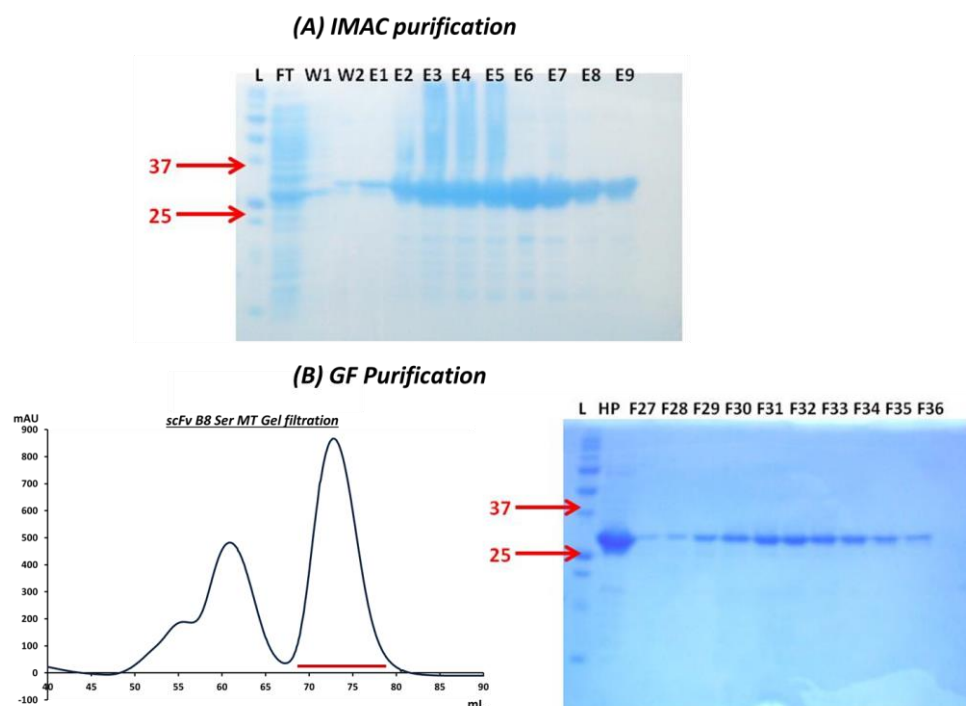


Figure 3.11: Purification of scFv B8 Ser MT by (A) IMAC and (B) Gel Filtration (GF). A 4 L culture of scFv B8 Ser MT was prepared and expressed antibody lysate purified by (A) IMAC and (B) GF. Eluates 2-6 from the IMAC purification were further polished by resolution on a size exclusion column. Fractions 21- 31 (underlined –red) were collected and analysed by SDS-PAGE. ScFv B8 Ser MT was successfully purified and concentrated to 16.12 mg/mL. L: Ladder; FT: Flow through 1; W1: Wash; W2: Wash 2; E1-9: Eluted protein fractions; HP: Histidine (IMAC) purified protein; F21-F31: Gel filtration containing protein fractions.

3.2.9 Determination of the crystal structure of anti fPSA scFv B8 Ser MT by X-ray crystallography

Initial crystal-forming conditions were identified by setting up 240 different conditions from a variety of commonly employed kits that provide highly effective and rapid screening methods for the crystallisation proteins. Each kit provides a differing set of crystallants, buffer composition, pH, salts and other additives to

comprehensively coarse screen for conditions to promote crystallisation of proteins. Within 72 hours, crystals were observed in 14% of the conditions set in the initial screen (34 crystal hits). Of these 34 crystals, 2 were notable similar to the wild type scFv crystals, that is large hexagonal type crystals (Figure 3.12). Refinement of these two conditions (as described in section 2.2.2.9.2) was carried out in order to increase the size and number of crystalline outcomes. Common refinements include varying the concentrations of all components in the crystallisation condition, slight pH changes, using additives and switching to similar buffers or precipitants (Luft *et al.*, 2014).

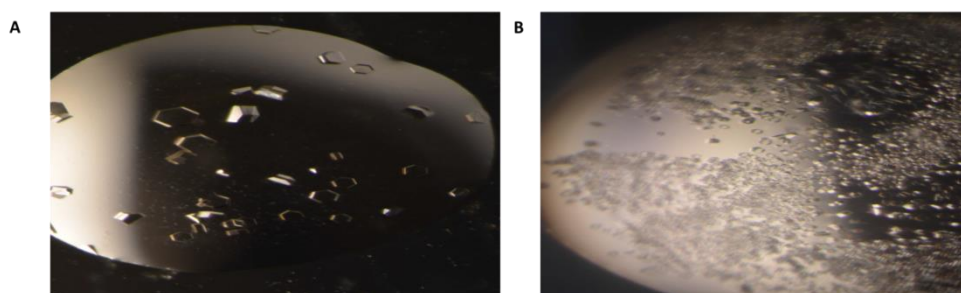


Figure 3.12: Crystals of scFv B8 Ser MT. Hexagonal type crystals were observed from the Crystal screen kit (Hampton Research) (A) condition number 5: 2.0 M Ammonium sulphate, 5 % (v/v) 2-propanol and (B) condition number 13: 0.2 M Ammonium sulphate, 0.1 M sodium acetate trihydrate, pH 4.6, and 30 % (w/v) polyethylene glycol monomethyl ether (PEG MME) 2,000.

Following the fine screening of these 2 conditions numerous crystals were obtained, with condition number 5 providing the largest crystals. Interestingly, many of these conditions generated lots of small crystals in clusters; however, as the % IPA decreased in the original condition (2.0 M Ammonium sulphate, 5 % (v/v) 2-propanol) the crystals were notably more spaced out and quite large in comparison to the high % IPA crystals. Individual crystals obtained from the

condition 2.0 M Ammonium sulphate, 3 % (v/v) 2-propanol were extracted from the drop (collected in a goniometer loop of an appropriate size (decided visually under a microscope) and loaded directly into the puck and stored in liquid nitrogen, using 25% (v/v) glycerol as the cryoprotectant, until the synchrotron visit.

X-ray diffraction was carried out at the Australian Synchrotron MX2 beamline at 100K. The data was merged and processed using XDS , POINTLESS and SCALA. Molecular Modelling and replacement was carried as described Conroy *et al.*, (2014). scFv B8 Ser MT was solved in the same space group as the wild type antibody (P6122), to 1.5Å and appears to retain the compact conformation observed in CDRH3 of scFv B8 (Figure 3.16C).

3.2.10 Kinetic analysis of scFv B8 Ser MT using Biacore 3000

Kinetic analysis was carried out in a similar manner to the wild type, starting with a preconcentration study to identify the optimal buffer conditions for subsequent immobilisation. Preparations of 25 µg/mL fPSA, diluted in 10 mM NaOAc over a range of pH values (pH, 4.0, 4.2, 4.4, 4.6, 4.8 and 5.0) were passed over the underivitisised chip surface (flow cell 4) and the degree of electrostatic binding monitored.

Maximum potential binding for fPSA was observed in NaOAc, pH 4.0 (Figure 3.13), Hence, this was utilised for immobilisation of 25 µg/mL fPSA to the surface of flow cell 4, using amine coupling chemistry.

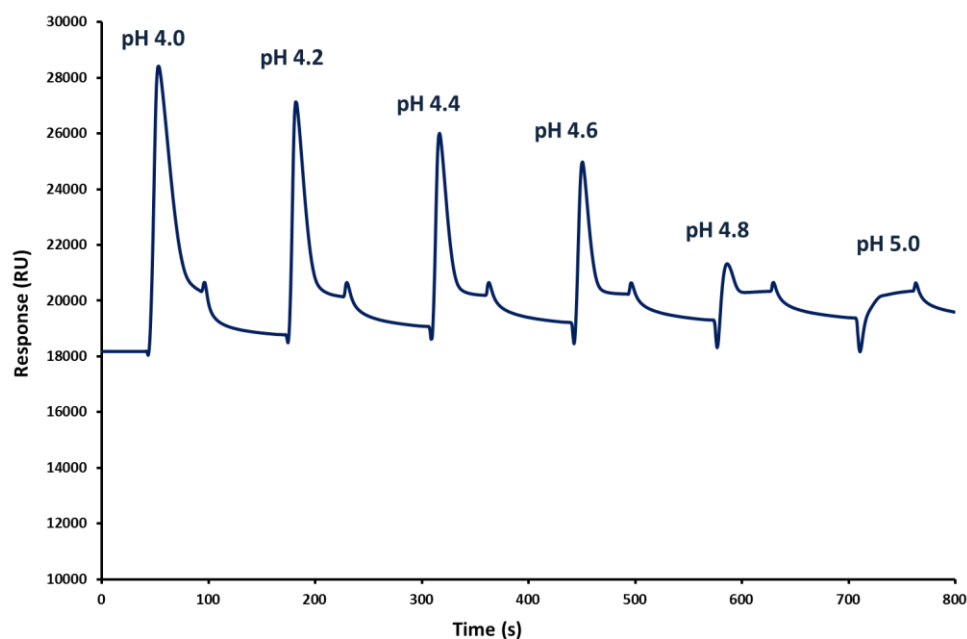


Figure 3.13: Preconcentration study for fPSA. fPSA was diluted in 10 mM sodium acetate solutions of pH, 4.0, 4.2, 4.4, 4.6, 4.8 and 5.0 and passed over the underivatised surface of a CM5 dextran chip at a flow rate of 10 μ L/minute. The optimal buffer conditions for the immobilisation of fPSA to the CM5 dextran surface was identified as NaOAc, pH 4.0, as the highest levels of binding was observed at this pH.

Once the optimal pH for immobilisation was identified, fPSA was immobilised to the surface, as detailed in the methods section (2.2.2.10.1), by passing fPSA diluted in 10 mM NaOAc to 25 μ g/mL, pH 4.0, over the activated surface at a flow rate of 10 μ L/min for 15 minutes. Following capping and deactivation of any unreactive NHS groups, the surface was post conditioned with 20 mM NaOH. Approximately 1000 RU were successfully immobilised to the surface of flow cell 4 of the CM5 dextran chip (Figure 3.14).

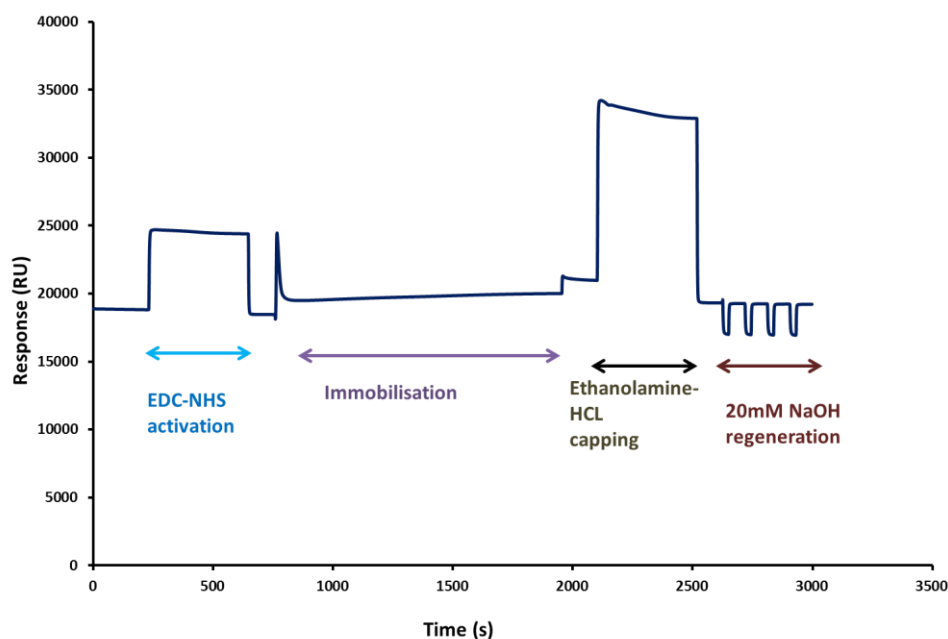


Figure 3.14: Immobilisation of fPSA. The surface of a CM5 dextran chip was activated with an EDC/NHS solution and fPSA passed over the activated carboxyl group surface. The sample was injected for 15 minutes until the surface was fully saturated. Binding of the ligand occurs via the amino groups on the protein with the activated carboxyl groups on the dextran surface chip. After sufficient capturing of the ligand was achieved the remaining activated groups were deactivated with ethanolamine-HCl solution. The surface was regenerated using a 20 mM solution of NaOH. The level of anti-HA successfully immobilised to the surface was approximately 1,000RU.

ScFv B8 Ser MT was kinetically characterised using the kinetic wizard functionality on the Biacore 3000 and the subsequent data generated was evaluated using BIAevaluation software, version 4.1.1. Purified scFv B8 Ser MT was passed over the fPSA surface in varying concentrations (2.5, 1.25, 0.625 (x2), 0.312, and 0 μ M) for 3 minutes, and allowed to dissociate for 12 minutes. The dataset from the reference flow cell (flow cell 3 - underivatized surface) was subtracted to remove any systemic effects, and the 0 nM scFv B8 Ser MT concentration binding response (buffer only) was subtracted from each antigen response, thus double

reference subtracting the dataset. The simplest 1:1 Langmuir binding model fit the dataset well with a global R_{\max} and the curve (Figure 3.15, upper graph) was fitted with a χ^2 value of 7.25. The χ^2 value is a quantitative measure of the closeness of fit, and in an ideal situation will approximate to the square of the noise level. It is difficult to recommend absolute values for acceptance limits for χ^2 ; the values need to be considered in combination with assessment of the shape of the residuals. The residual plot (Figure 3.15, lower graph) shows the data points are distributed randomly above and below the zero line.

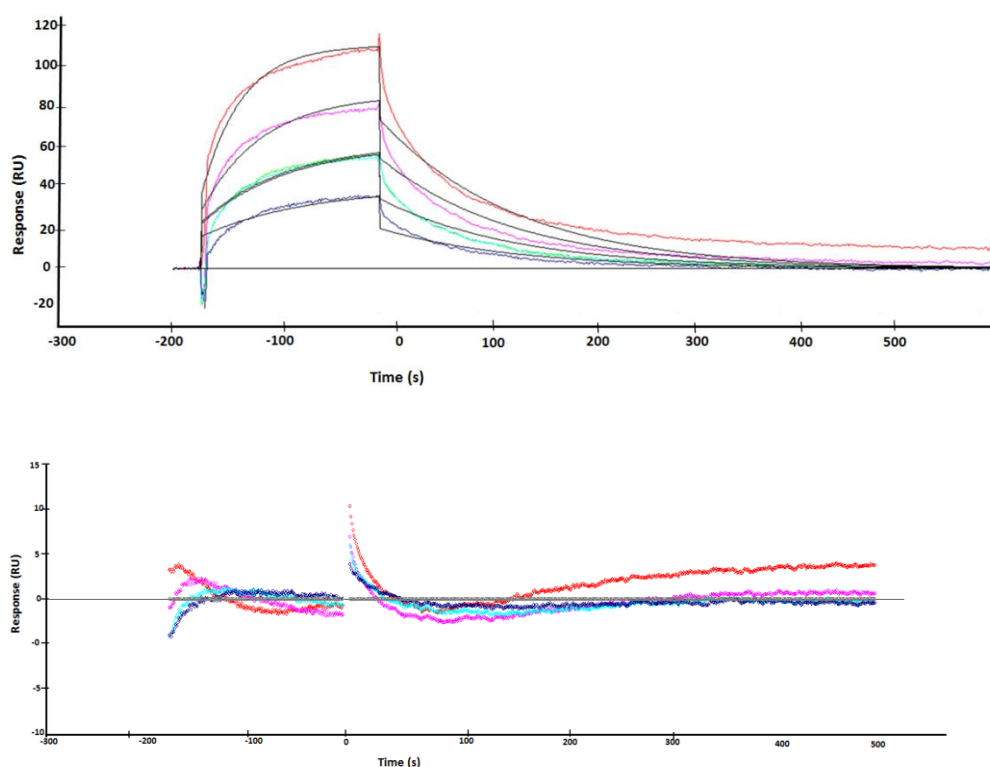


Figure 3.15: Kinetic interaction analyses of fPSA and scFv B8 Ser MT. Five concentrations of scFv B8 Ser MT (2.5 μM , 1.25 μM , 0.625 μM (x2), 0.312 μM , and 0 μM) were used to fit the kinetics using 1:1 Langmuir binding model (global fit) The following kinetic parameters were derived; K_a : $7.18 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$, K_d $7.38 \times 10^{-3} \text{ s}^{-1}$, and K_D ; 1.03×10^{-6} with a χ^2 (χ^2) value of 7.25.

A Loss of binding was observed when the cysteine residues are substituted with serine as scFv B8 Ser MT binding drops from sub nM affinity (wild type scFv B8: K_D ; 0.8×10^{-9} to μ M affinity (scFv B8 Ser MT K_D ; 1.03×10^{-6}).

3.2.11 Expression and Purification of scFvB8 Ala MT by Immobilised Metal Affinity Chromatography (IMAC) and Gel Filtration

Given the structural similarities between scFv B8 and scFv B8 Ser MT yet the huge differences in binding kinetics, a second mutant was created to further investigate the impact of the cysteine pair in CDRH3.

ScFv B8 plasmid DNA was used as a template to introduce a cysteine to alanine double mutant at C97 and C100 in CHRH3 (C97/100A; scFv B8 Ala MT) (section 2.2.2.7). Successful generation of the mutant sequence was confirmed (Appendix two) and the mutant clone was subsequently expressed and purified by Dr. Paul Conroy, as detailed in section 2.2.2.8.

3.2.12 Determination of the crystal structure of anti fPSA scFv B8 Ala MT by X-ray crystallography

Crystal forming conditions for scFv B8 Ala MT were identified from the Crystal Screen and Crystal Screen II (Hampton Research) kit by Dr. Paul Conroy. Crystals of scFv B8 Ala MT were obtained in 2 M ammonium sulphate. Crystals were also obtained in 0.1 M NaOAc, pH 4.6, with 0.2 M ammonium sulphate. These crystals were flash-cooled in liquid nitrogen using 25% (v/v) glycerol as the cryoprotectant.

Data sets were collected at the Australian Synchrotron MX2 beamline at 100K. The data was merged and processed using XDS, POINTLESS, and SCALA. Molecular Modelling and replacement was carried as described by Conroy *et al.*, (2014).

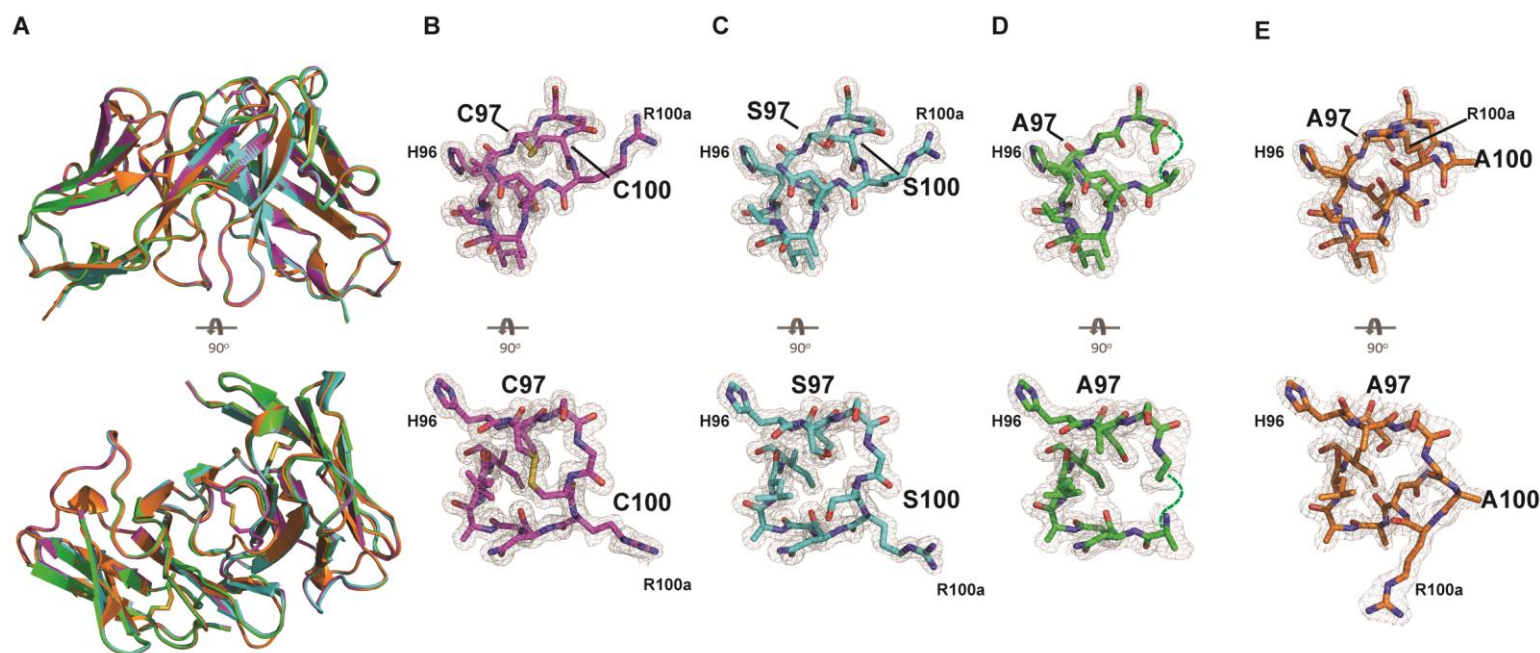
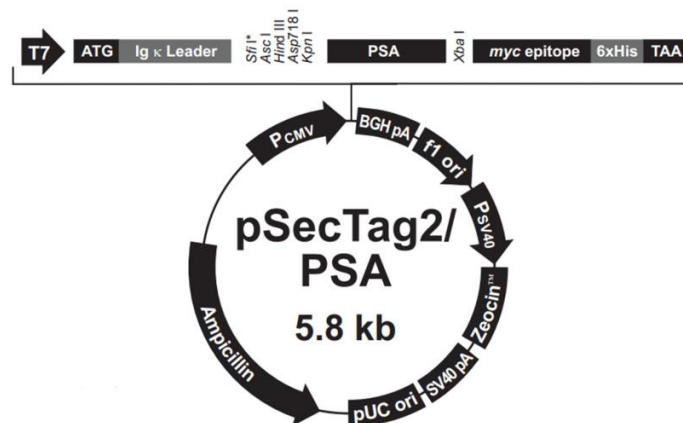


Figure 3.16: Crystal structures of wild type and mutant variants of anti-fPSA scFv B8. (A) Diagrammatic representation of wild type scFv format with each mutant variant overlaid (magenta: wild type, cyan: Ser MT and green/orange: Ala MT). Electron density map of (B) wild-type (C) Ser MT and (D) and (E) Ala MT, with each showing the CDRH3 arrangement. All structures were solved in the same space group –P6122, with the exception of the second Ala MT (E), which was solved in the P212121 space group. The Ser MT (C) contains the same compact configuration of wild type scFv BB due to hydrogen bond formation. The rigid CDRH3 is opened up completely in the Ala MT (D) and in the second Ala structure (E) R100a swings out, widening out CDRH3.

3.2.13 Mammalian Expression of recombinant PSA

The commercial available Invitrogen vector suite, pSecTag2 A, B, and C are specifically designed expression vectors for high-level stable and transient expression in mammalian hosts. This vector suite also includes pSecTag2/PSA, in which PSA is fused to the *c-myc* epitope and the polyhistidine tag. The resulting fusion protein is ~33 kDa which includes the N-terminal secretion signal (Figure 3.17)



pSecTag2/PSA: 5822 nucleotides:

CMV promoter: bases 209-863	BGH polyadenylation sequence: 1829-2043
T7 promoter: bases 863-882	f1 origin: bases 1954-2367
Murine Ig κ-chain V-J2-C signal peptide: bases 905-967	SV40 promoter and origin: bases 2587-2908
PSA gene: bases 1042-1732	EM-7 promoter: bases 2924-2990
c-myc epitope: bases 1745-1777	Zeocin TM resistance gene: bases 2991-3365
Polyhistidine tag: bases 1790-1807	pUC origin: bases 4008-4681
BGH reverse priming site: bases 1830-1847	Ampicillin resistance gene: bases 4826-5686

Figure 3.17: *Vector map of pSecTag2/PSA*. Map details the many features of the pSecTag2A/PSA vector (Invitrogen). Features include: Human cytomegalovirus (CMV) promoter/enhancer, T7 promoter, ATG initiation codon (permits initiation of translation of the pSecTag2 fusion protein) Murine Ig κ-chain leader sequence, Multiple cloning site, c-myc epitope (allows detection of pSecTag2 fusion protein with widely available anti-myc antibodies), Polyhistidine tag (for high affinity binding to Ni²⁺ chelating resin easy purification by IMAC). In addition, Bovine growth hormone (BGH) reverse priming site and BGH polyadenylation signal, f1 origin , SV40 early promoter and origin , ZeocinTM resistance gene (for selection of transformants in *E. coli* and stable transfectants in mammalian cells), SV40 polyadenylation signal (permits efficient transcription termination and polyadenylation of mRNA), pUC origin and ampicillin resistance gene.

Figure adapted from

http://tools.lifetechnologies.com/content/sfs/manuals/psectag2_man.pdf.

HEK293F cells were seeded to $1-2 \times 10^6$ cells/mL in Freestyle media; a chemically defined, 'protein-free' expression medium specifically developed to support the growth and transfection of HEK293F cells in suspension without adaptation. Polyethylenimine (PEI) was used as a transfection agent and 1 mL samples were taken at days 3, 4, 5 and 7 post transfection. Samples were centrifuged at $5,500 \times g$ for 30 minutes and the supernatant analysed by SDS-PAGE (Figure 3.18).

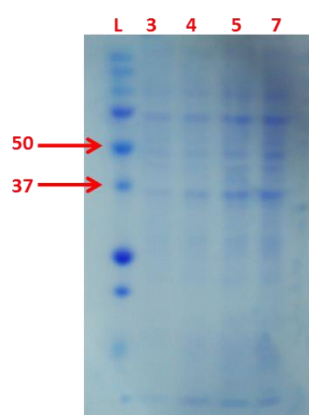


Figure 3.18: Recombinant PSA expression. Samples (1 mL) were taken at days 3, 4, 5, and 7 and the presence of PSA confirmed by resolution on a 15% SDS –PAGE gel. PSA is fused to the *c-myc* epitope and the polyhistidine tag, resulting in a ~33 kDa fusion protein. Expression increases over time and 7 days post transfection was determined to be the optimal time for expression of recombinant PSA. L: Ladder; 3: PSA sample taken at 3 days; 4: PSA sample taken at 4 days; 5: PSA sample taken at 5 days and 7: PSA sample taken at 7 days.

3.2.14 Purification of recombinant PSA by Immobilised Metal Affinity Chromatography (IMAC) and Gel Filtration

Recombinant PSA was purified in the 2-step process detailed in section 2.2.2.14, which yields highly pure, homogenous proteins (Figure 3.19).

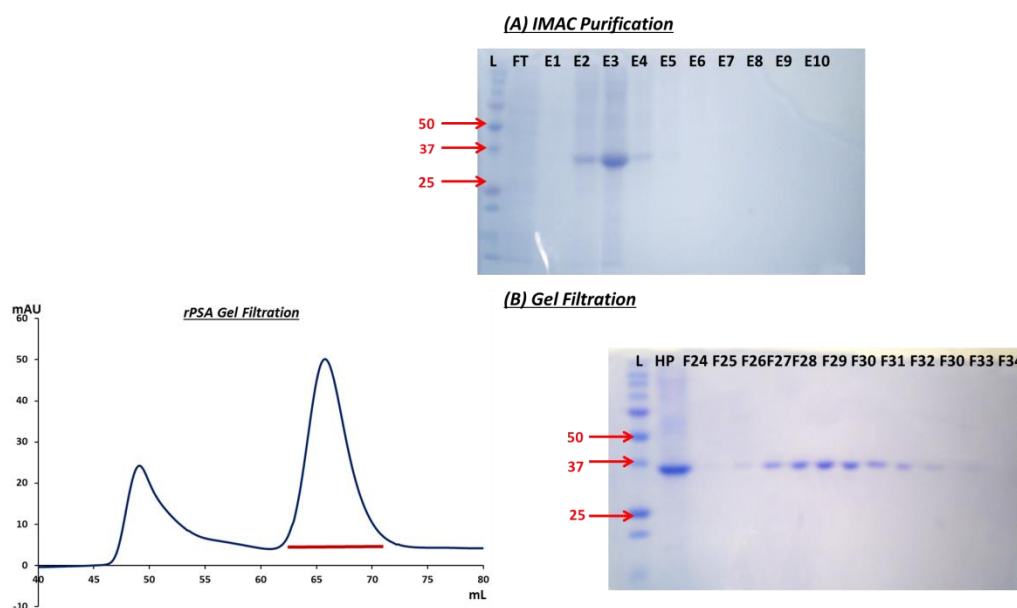


Figure 3.19: Purification of recombinant PSA (rPSA) by (A) IMAC and (B) Gel Filtration (GF). Secreted media (35 mLs) was collected and lysate purified by (A) IMAC and (B) GF. Eluates 2,3 and 4 from the IMAC purification were further polished by resolution on a size exclusion column. Fractions 26- 34 (underlined –red) were collected and analysed by SDS-PAGE. rPSA was successfully purified and concentrated to 0.4 mg/mL. L: Ladder; FT: ‘Flow through’ 1; W1: Wash; W2: Wash 2; E1-10: IMAC-eluted protein fractions; HP: Histidine (IMAC) purified protein and F24-F34: Gel filtration containing protein fractions.

3.2.15 Formation of recombinant PSA-scFv B8 complex

The rPSA-scFv B8 complex was formed by addition of excess scFv B8 (1 mg) to PSA (0.4 mg). This protein mixture was left on ice for 2 hours and then concentrated before resolution on a S75/16/60 analytical gel filtration column (section 2.2.2.15). The complex was successfully formed (Figure 3.20) and complex-containing fractions were isolated and concentrated to 7.3 mg/mL for screening for identification of crystal-forming conditions.

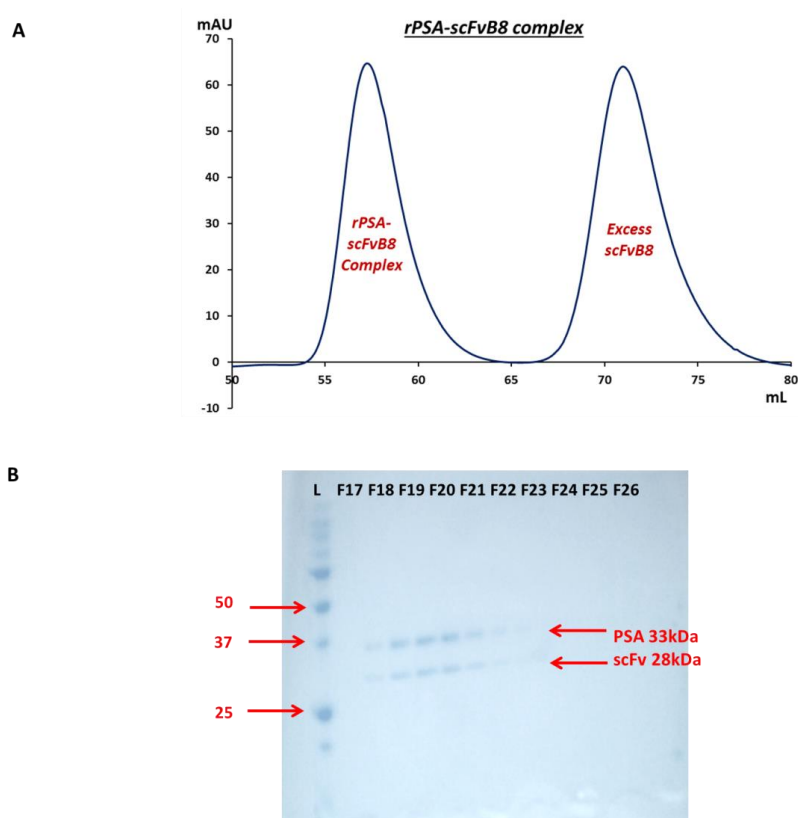


Figure 3.20: Formation of rPSA-scFv B8 complex. Protein complex was allowed to form by adding excess scFv B8 to rPSA and leaving static for 2 hours. (A) The protein complex was confirmed by resolution on a size exclusion column. The first peak contains the protein complex while the 2nd peak contains excess scFv B8. (B) Fractions 17- 26 (1st peak) were collected and analysed by SDS-PAGE to ensure the presence of both PSA (33 KDa) and scFv B8 (28KDa).

3.2.16 Determination of the crystal structure of rPSA-anti-fPSA scFv B8 complex

Initially, crystals were observed in 0.1 M magnesium acetate, 0.1 M MOPS, pH 7.5, with 12% (w/v) PEG 8000. Following fine screening of this condition, an increase in % (w/v) PEG 8000 yield better crystal formation (0.1 M magnesium acetate, 0.1 M MOPS, pH 7.5, 15% (w/v) PEG 8000). Polyethylene glycol (PEG) is a high molecular weight linear polymer that is frequently used as a precipitating agent for generation of protein crystals. PEG is highly effective both in terms of precipitating ability and cost effectiveness and is particularly well suited for macromolecular crystallisation (McPherson, 1985).

Crystals were flash-cooled in liquid nitrogen using 25% (v/v) glycerol as a cryoprotectant. X-ray diffraction data-sets were collected at the Australian Synchrotron MX2 beamline. The data was merged and processed using XDS, POINTLESS, and SCALA. Molecular Modelling and replacement was carried as described by Conroy *et al.*(2014). The resultant anti-fPSA scFv B8-rPSA complex structure was solved to 1.95 Å in the P1 21 1 space group (Figure 3.21).

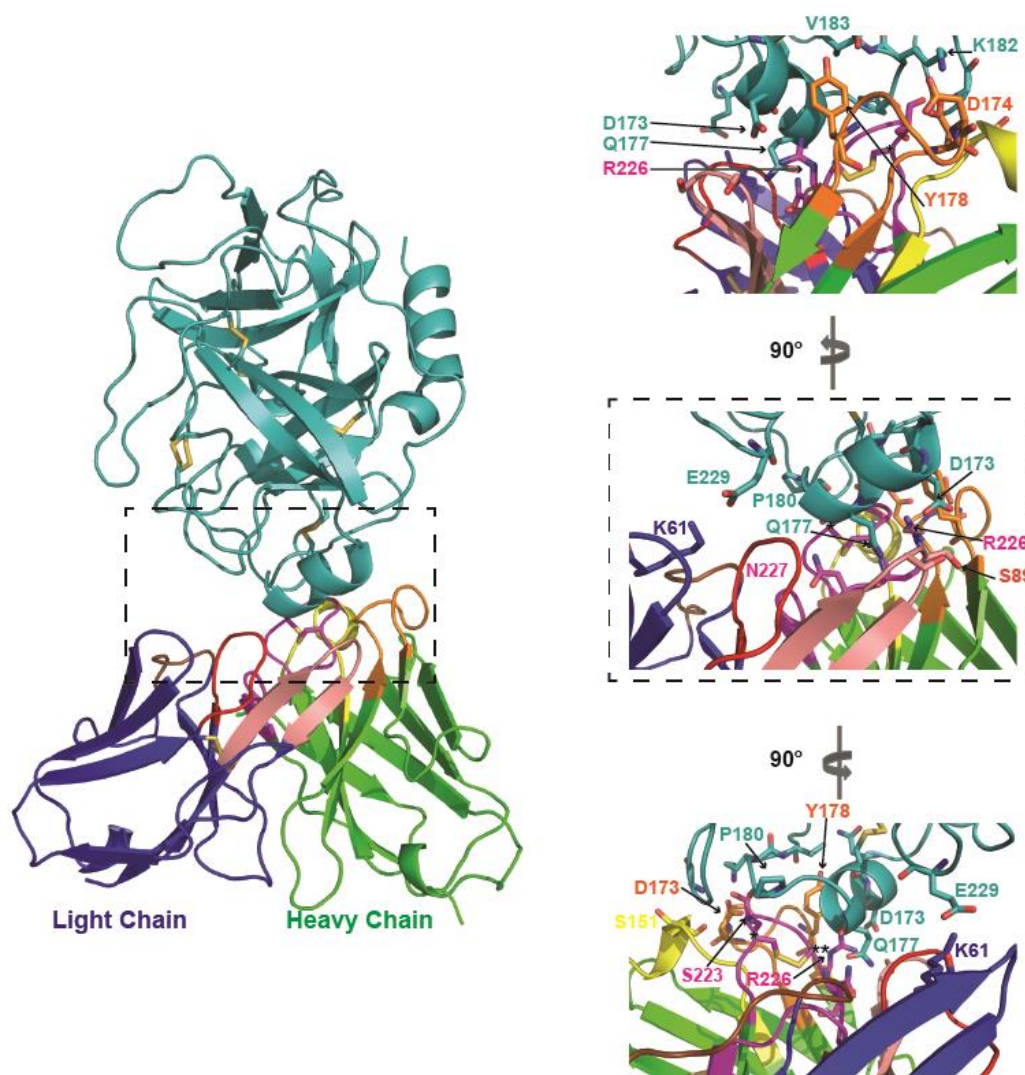


Figure 3.21: Crystal structure of anti-fPSA scFvB8-rPSA complex. ScFv B8 light (blue) and heavy (green) chains with the CDRs shown: L1 (red), L2 (brown), L3 (salmon), H1 (yellow), H2 (orange) and H3 (magenta) in contact with rPSA. Expanded view of the interaction (black dashed box) with residues contributing to the interaction with the Y-axis rotate by 90° to view from the heavy (top) and light (bottom) chain.

3.3 Discussion

This chapter describes the screening of a pre-existing scFv and fab anti-fPSA library for identification of a highly sensitive anti-fPSA-specific antibody fragment and its subsequent characterisation.

An anti-fPSA library was constructed from the spleens of female Leghorn chickens immunised with fPSA was readily available for screening and isolation of an fPSA specific antibody fragment. These antibody fragments were previously isolated from a large library (3.10×10^8 cfu/mL) following three rounds of selection from the antibody repertoire displayed on the surface of filamentous phage by immobilised panning using phage display (carried out by Dr. Conor Hayes).

Here, approximately 800 anti-fPSA fab clones and 280 anti-fPSA scFv clones were screened by monoclonal ELISA and evaluated for binding affinity using the Biacore 4000 and a nanomolar affinity anti-fPSA scFv was identified. Various optimisation steps were taken with this fragment with the aim of expressing and purifying sufficient amounts of antibody for extensive kinetic characterisation with the Biacore 3000. Expression of scFv B8 was optimised and although proving difficult, 1 mg/mL protein was purified by IMAC (Figure 3.4).

One point to note from the optimisation of expression (Figure 3.3) is that the antibody fragment appears to be expressing in the absence of IPTG. Typically, in the absence of the inducer, the target gene will not be transcribed in lac controlled vectors, as pComb3 is under control of the *lac* operon, scFv B8 should not express without IPTG. This issue of uninduced expression is sometimes observed in *E. coli* systems that utilise the lac operon and is often referred to as 'leaky' expression (Grossman *et al.*, 1998). Some strains of *E. coli* such as

BL21(DE3)pLysS and Rosetta(DE3)pLysS are specifically developed to assist in the reduction of leaky expression in lac controlled systems. These strains contain a plasmid encoding T7 lysozyme, which binds and inhibits T7 polymerase (Moffatt and Studier, 1987). The concentration of T7 lysozyme produced in these strains results in low basal transcription activity. The use of these strains of *E. coli* could be used in future experiments to reduce the observed levels of uninduced expression.

The anti-HA tag integrated into the pComb3x vector not only facilitates purification but also assisted in the kinetic evaluation of scFv B8. A polyclonal anti-HA antibody was immobilised to the surface of a CM5 dextran chip via amine coupling chemistry and a kinetic program using the capture binding assay format was custom developed for interaction analysis between anti-fPSA scFv B8 and fPSA. The affinity for fPSA was determined to be 0.8×10^{-9} M from the data generated. Each kinetic evaluation study was carried out 3 times (inter-day and intra-day) to provide confidence in the results obtained.

'Limit of detection' (LoD) is a term used to describe the smallest concentration of a measurand that can be reliably measured by an analytical procedure. It can be defined as the lowest analyte concentration likely to be reliably distinguished from the limit of the blank (LOB) and at which detection is feasible. An approach for determining the LoD was described by Armbruster and Pry (2008) and was employed in this study to determine the LoD of anti-fPSA scFv B8. This method utilises analysis of samples containing small but known concentrations of the substance of interest. The advantage of this empirical approach is that objective data is used to compare the analytical response of blank and low concentration samples to determine conclusively what concentration of analyte is necessary to

distinguish its presence from its absence (Armbruster and Pry, 2008). A total of 24 replicates were carried out (Table 3.2) and the LoD of anti-fPSA scFv B8 was identified as 0.03 ng/mL. The ability to detect (repeatedly) low levels of fPSA is essential for patients in the 'diagnostic grey zone' of 4-10 ng/mL as clinicians use a measure of percentage fPSA (free:total PSA ratio) to determine the percent probability of finding prostate cancer on a needle biopsy by age in years, as detailed in Table 3.3.

Table 3.3: Relative risk of prostate cancer as determined by free:total PSA ratio (Taken from:<http://www.mayomedicallaboratories.com/testcatalog/Clinical+and+Interpretive/81944>).

Free:total PSA ratio	50-59 years	60-69 years	> or =70 years
< or =0.10	49.2%	57.5%	64.5%
0.11-0.18	26.9%	33.9%	40.8%
0.19-0.25	18.3%	23.9%	29.7%
>0.25	9.1%	12.2%	15.8%

Järås and colleagues developed an assay technology platform on which a large number of plasma samples as well as a large number of analytes could be measured. The author's effectively show quantitative duplex detection of free and total PSA on the same porous silicon chip in an antibody microarray. The microarray format was chosen to allow multiplex detection of biomarkers related to prostate cancer. The fPSA antibody used in this assay was a recombinant Fab (rFab 5A10)(Eriksson *et al.*, 2000) measured free PSA with a dynamic range of 0.40-74.9 ng/mL and a LOD: 0.47 ng/mL (Järås *et al.*, 2012). Anti-fPSA scFv B8 has an LoD of 0.03 ng/mL, significantly lower than the antibody employed by the assay developed by Järås and colleagues. This antibody fragment is also highly

specific to fPSA (Figure 3.6) and was not shown to cross-react with the other major component of total PSA, complex PSA (cPSA). Incorporation of anti-fPSA scFv B8, alongside other prostate cancer specific markers, into a multiplex assay such as that developed by Järås and colleagues could greatly enhance PCa diagnosis.

Sequence analysis of anti-fPSA scFv B8 carried out by Source Bioscience Ltd, Dublin, Ireland. The DNA sequences obtained were translated using the translate tool from Expasy (<http://web.expasy.org/translate/>) and the amino acid sequences were then exported to PowerPoint and the significant regions highlighted (appendix two).

Sequencing exposed an unusual cysteine pair in CDRH3. In a recent study, the V_H repertoire of IgY was characterised and compared with V_H CDR diversity of humans, mice and camelids (Wu *et al.*, 2012). Here, the authors show that the chicken CDRH3 repertoire has higher cysteine content than other species (chicken 9.4%, human 1.6% and mice 0.25%). Data from this study suggests that noncanonical disulphides are a structural diversification strategy in the restricted V gene germline repertoire of chickens, and are favourable as they potentially aid in stability and affinity for its target (Shih *et al.*, 2012; Wu *et al.*, 2012).

Following this observation, the crystal structure of scFv B8 was resolved to 1.4 Å (work carried out by Dr. Paul Conroy, Monash University, Melbourne). To precisely understand the molecular mechanisms of proteins high resolution of 2 Å is required. We successfully resolved scFv B8 to 1.4 Å, providing confidence in the level of accuracy of this structure, and exposed a unique shelf-like structure within CDRH3. Until recently, these types of bonds in CDR have never been seen in a crystal structure in chicken antibodies (Shih *et al.*, 2012). Mutational hotspots and structure-guided mutation strategies can maximise the diagnostic performance of

recombinant antibody fragments (Muller *et al.*, 2011; Stura *et al.*, 2011). Muller and colleagues specifically mutated an anti-fPSA antibody (Fab – 5D3D11) and the results from this study, showed improvement to the diagnostic performance for patients in the ‘diagnostic grey zone’, allow better discrimination between prostate cancer and BPH (Muller *et al.*, 2011).

The solved structure revealed a number of interesting canonical structural grouping deviations that are likely distinctly ‘chicken’. At present the lack of chicken structures in the PDB impedes a rational, focused appraisal of the consequences of novel CDRL1 canonical structures or the non-canonical disulphide bonded CDRH3, for the classical modes of protein-, peptide- or hapten-binding. The CDRL1 of scFv B8 mirrors the chicken canonical cluster described by Shih and co-workers (Shih *et al.*, 2012) and its presence is also concurrent with a long, non-canonical disulphide constrained CDRH3, which exhibits a distinct bias towards small amino acids (G/S/A/C/T) (Finlay and Almagro, 2012; Wu *et al.*, 2012; Conroy *et al.*, 2014).

This may suggest that such a binding site arrangement is not antigen binding mode specific, but raises the possibility that it is necessary to have a short CDL1 to facilitate the elongated and disulphide constrained CDRH3. These descriptive structures further support the increasing genetic, functional and now structural evidence that such constrained CDRH3 are an active diversification strategy in the restricted germline repertoire of chicken (Shih *et al.*, 2012; Wu *et al.*, 2012; Conroy *et al.*, 2014). This CDRL1-CDRH3 arrangement forms a common structure in chicken antibodies (2 of the 3 available chicken structures), which is distinctly different from the structures described in mammals to date (North *et al.*, 2011; Chailyan *et al.*, 2011; Shih *et al.*, 2012; Conroy *et al.*, 2014).

Following these observations, mutant variants of scFv B8 were generated by site-directed mutagenesis to assess the impact the two intra-CDRH3 cysteine residues their structural impact played in antigen binding. Mutation of these cysteine residues to serine did not change the 'shelf-like' structure in CDRH3 (Figure 3.17-B) however, it dramatically reduced binding affinity (sub-nanomolar to micromolar) (Figure 3.16). Given that the presence of charged surfaces can often encourage protein-protein interactions (Boswell *et al.*, 2010), it is unsurprising that there is a noticeable reduction of positive charge in the Ser-MT antibody fragment (Figure 3.23-B) in CDRH3 compared to the wild type (Figure 3.23-A). The deliberate modification of antibody charge through direct chemical modification of amino acid side chains with positive, negative or neutral chemical groups is often carried out to permit exploratory study of the relationships between the net molecular charge and binding kinetics, and are also carried out to improve the pharmacokinetic properties of antibodies for the diagnosis or treatment of disease (Boswell *et al.*, 2010).

A second mutant variant of wild type scFv B8, with a double alanine substitution at C97 and C100 was also created. Two crystals were diffracted and the structures were solved in different space groups. One structure diffracted to 1.7 Å in the same space group as wild type and Ser MT (P61 2 2) (Figure 3.17 C) and the second was solved to 1.9 Å in the P 21 21 21 space group (Figure 3.17 D). In both cases there was a single molecule in the asymmetric unit (Appendix two lists full refinement statistics for all structures).

The rigid formation seen in the wild type and Ser MT was disrupted and CDRH3 is disordered in the Ala MT. Kinetics analysis was carried out (data not shown). However, the Ala MT lost all binding capabilities and no interaction was observed

with PSA. On the basis of these results, it is believed this disulfide bond in CDRH3 of scFv B8 is critical for its ultra-specific recognition of fPSA. Selection of antibodies from chicken antibody libraries that contain intra-CDR disulfide bridges is highly favourable and can ultimately lead to identification of high affinity antibodies with unique binding profiles.

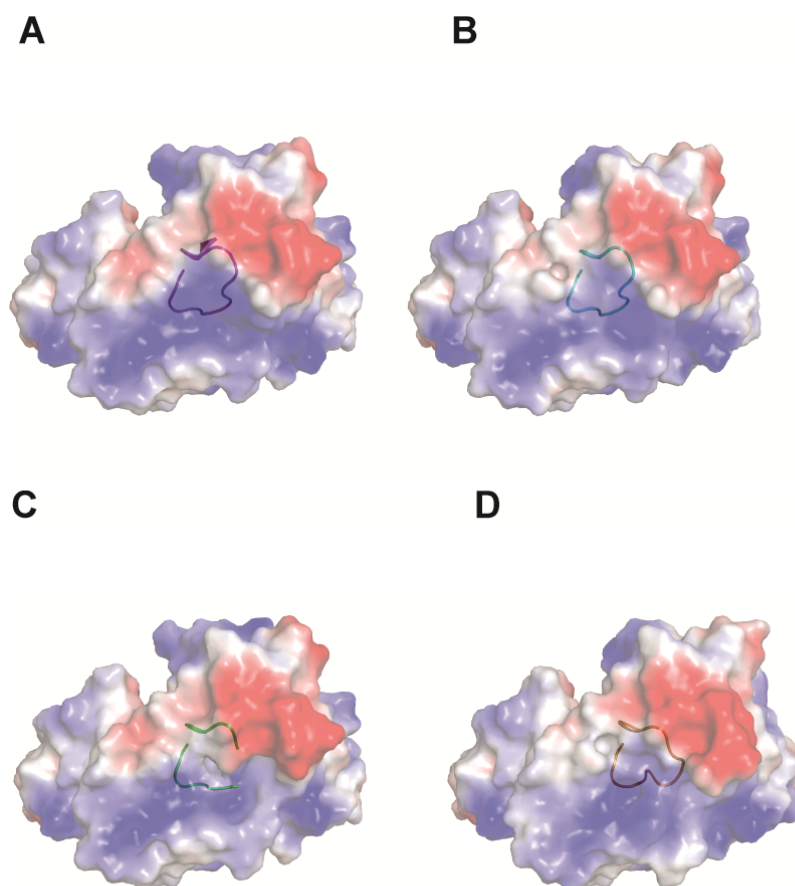


Figure 3.23: Diagrammatic representation of the surface charge of wild type and mutant variants of anti-fPSA scFv B8. Transparent electrostatic surface view of the WT (A) and Ser MT (B) and Ala MT (C, D) from the antigen perspective, with the CDR loops visible. Positively and negatively charged areas are indicated in blue and red, respectively.

While determination of the crystal structure and kinetics analysis of anti-fPSA scFv B8 and its mutant variations strengthened the hypothesis that conserved disulphides in Aves play an important role in creating paratopes of great complexity, to truly understand the binding mechanisms of this antibody-antigen interaction, the crystal structure of anti-fPSA scFv B8 and recombinantly expressed PSA was determined. The high resolution co-complex structure was solved to 1.95 Å in the *P1* space group following diffraction of one crystal at the Australian Synchrotron. This is the first example of a protein antigen-avian antibody structure to our knowledge.

The kinetics studies on the mutant variants clearly showed that the disulphide bridge was essential for antigen recognition; however it is possible that the disulphides do not actually make contact with recombinant PSA (Figure 3.24). Recombinant PSA forms contact with scFv B8 directly between the disulphide bond at S98, R100a and N100b (Kabat numbering) in CDRH3. It is likely that the disulfide bond stabilizes the domain creating a rigid scaffold for interaction with recombinant PSA. Each of the hypervariable loops in the heavy chain form contact with recombinant PSA, whilst two of the light chain CDRs forms part of the paratope. Binding in CDRH1 occurs with Y56 and negatively charged amino acid aspartic acid at D52a and D52b. Interestingly, human and mice antibodies both show a strong preference for the use of Tyrosine residues whereas the amino acid content in the chicken shows very different paratope chemical composition, including an unusually low representation of Y (Finlay and Almagro, 2012; Wu *et al.*, 2012).

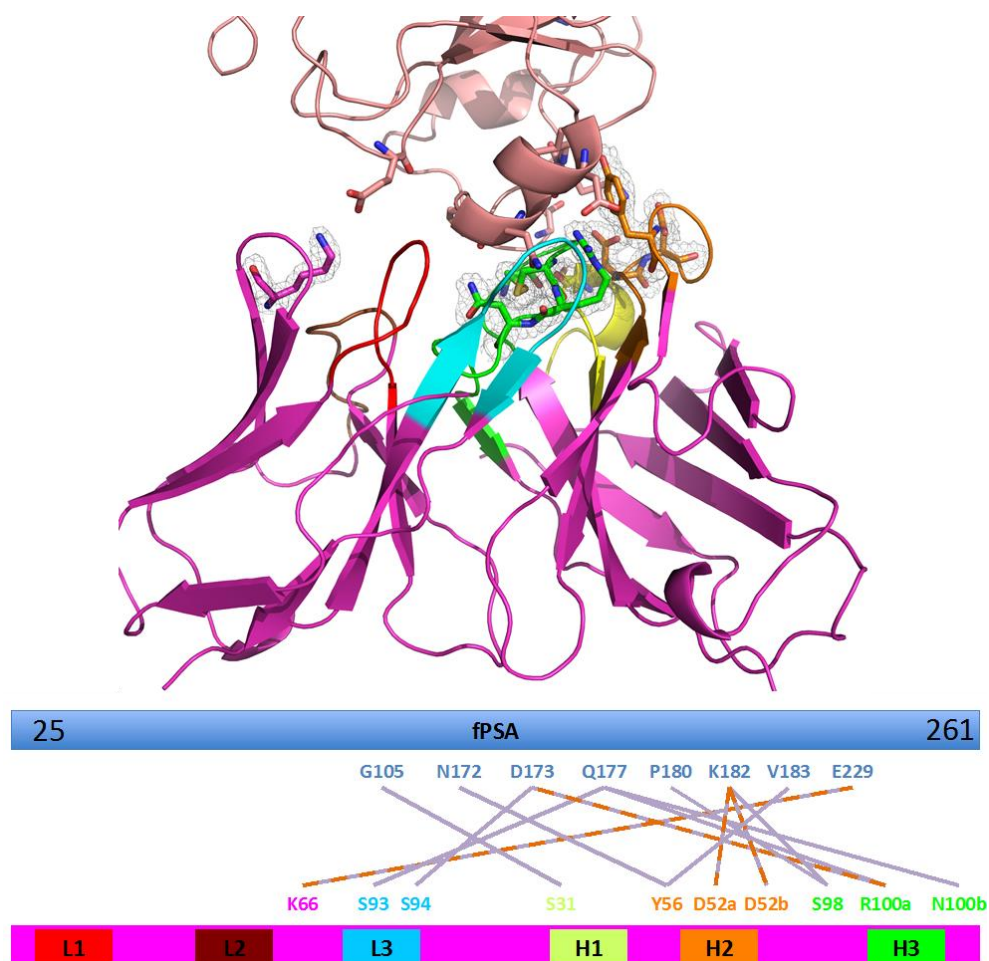


Figure 3.24: Anti-fPSA scFv B8-recombinat PSA binding contacts. The conformational epitope is comprised of 2 major binding segments on recombinant PSA (N172, D173, Q177 and P180, K182, V183), with single residue contacts (G105 and E229) at each end of the major contact segments. Each of the hypervariable regions of the heavy chain of scFv B8 comprises the majority of binding residues with the paratope.

There is a distinct bias toward small amino acids G/S/A/C/T in chicken antibodies (Wu *et al.*, 2012). The prevalence of these small residues in the CDR loops implies that they are important for conformational flexibility (Birtalan, Zhang, Frederic A. Fellouse, et al., 2008). Serine accounts for 40% of the residues in contact with

recombinant PSA, suggesting this flexibility is essential for scFv B8's highly affinity recognition of PSA.

A noteworthy observation of this antibody-antigen complex is that there is a contact residue (K66) in the framework region between CDRL1 and L2. The less variable framework regions form the β sheets and generally provides structural support for the antigen binding site, rather than making contact with the antigen (Finlay and Almagro, 2012). Sequence variability in the framework regions can directly affect CDR loop conformation and the orientation of V_H - V_L pairing. Albeit rarely, binding contacts have been previously observed in the framework region, particularly the heavy chain FR3- the framework region between CDR H2 and H3, which accounts for 1.3% of human antibody-antigen contacts (Sela-Culang *et al.*, 2013). Framework residues in the light chain rarely make contact with the antigen, as was observed with anti-fPSA scFv B8 and recombinant PSA. It is likely that K66 plays a role in maintaining the stability of the overall structure of the antibody-antigen complex.

Chapter conclusions

Overall screening of the anti-fPSA library was successful and a highly specific antibody fragment was isolated, expressed, purified and extensively characterised. The crystal structure of this antibody fragment revealed that the restricted germline repertoires of chicken antibodies can produce antibodies with unique binding mechanisms. The co-complex of this avian antibody fragment in contact recombinant PSA was also solved.

Chapter Four

***Generation of recombinant
anti-PSA antibody fragments
and N-glycan characterisation
of polyclonal antibodies from
an avian immune system***

4.1 Introduction

This chapter describes the development of avian recombinant antibody fragments to complex prostate specific antigen (cPSA) by phage display, the process of library building and the selection processes involved in identifying cPSA-specific antibodies. The isolation and *N*-glycan characterisation of the avian IgY molecule is also detailed.

The use of PSA as a biomarker of prostate cancer (PCa) remains highly controversial as it lacks cancer specificity. There is a nominal level of PSA in serum in healthy males hence the presence of PSA in serum does not necessarily mean a man has prostate cancer. Circulating levels of the protein can fluctuate in response to a variety of circumstances including sexual activity, BPH and chronic prostatitis (Botchorishvili *et al.*, 2009). A PubMed search of ‘prostate cancer diagnosis’ returns over 80,000 articles (as of July 2015) but, in the 25 years since the FDA first approved the serum PSA test for clinical use, no other biomarkers have gained FDA approval that surpass the PSA gold standard. PSA is still the primary tool for detection, risk stratification and monitoring of PCa. A greater understanding of the molecular isoforms of PSA has led to improvements in the cancer-related specificity of the test.

While advances are being made in improving the PSA test, there is an unmet need to clearly distinguish between significant and insignificant PCa. However, it is unlikely—in the near future, at least—that any single biomarker will be capable of definitively distinguishing between significant and insignificant disease. Combining panels of markers (Oon *et al.*, 2011)—for example, serum protein markers (Fan *et al.*, 2011), miRNA expression profiles (Hassan *et al.*, 2012), and epigenetic

modifications such as DNA methylation (Jerónimo *et al.*, 2011)—with novel glycan expression patterns could improve the accuracy of diagnosing and screening for prostate cancer (Gilgunn *et al.*, 2013). The primary aim of this chapter is to develop isoform-specific recombinant antibody fragments to incorporate into a microfluidic-based platform with multiple PCa-specific markers for improved PCa diagnosis.

Phage display is the predominant technique employed in recombinant antibody selection. Random segments of DNA can be readily inserted into the phage coat particle and amplified in bacterial cells. Antibody fragments are frequently inserted into the coat protein of the phage and subsequently expressed on the surface of the phage following superinfection with helper phage. The phage-scFv particle can then be selectively enriched for the cognate antigen using solid phase selection screening of the phage library. The major advantage of phage display is the link between phenotype and the corresponding genotype (Barbas *et al.*, 2001).

Polyclonal and recombinant antibodies are commonly developed in many different species provided the immunoglobulin gene sequences are known (Andris-Widhopf *et al.*, 2000). However, a large number of protein targets are highly conserved in mammalian evolution, thus commonly used mammalian species such as rabbits and mice are inclined to generate a limited immune response due to immunological tolerance invoked during foetal development (Andris-Widhopf *et al.*, 2000). The use of a species more phylogenetically distant from humans such as chickens, who diverged from mammalian genomes some 310 million years ago (International chicken Genome Sequencing Consortium,

2004), are ideal alternatives for immunisation and selection of recombinant antibodies (Andris-Widhopf *et al.*, 2000; Narat *et al.*, 2003).

Recent exploitation of the avian immune system has highlighted its use for the generation of high-quality, high-affinity antibodies to a wide range of antigens for a number of therapeutic and diagnostic applications (Spillner *et al.*, 2012; Wu *et al.*, 2012).

The mechanism in which chickens generate antibody diversity was extensively characterised at the genomic level and is considerably simpler in comparison to the v-gene repertoire of mice and humans, which involves a large set of sequences that are highly diverse in both sequence and structure (Wu *et al.*, 2012). The avian v-gene germline repertoire utilises single functional V-genes for the heavy (V_H3 family) and light chains (exclusively λ light chains) that contain unique V_L-J_L and V_H-D-J_H segments (Finlay and Almagro, 2012). Despite this simplified system, chickens have a broadly adaptable repertoire capable of generating unique, high affinity antibodies to proteins, peptides and hapten antigens. The v-genes are rearranged using conventional V(D)J recombination mechanisms, however in order to generate a large, diverse antibody repertoire, chickens employ a complex V-gene diversification mechanism known as 'gene conversion'. This is a process in which upstream pseudogenes, that lack a recombination signal, are incorporated into the expressed rearranged V-region gene (Wu *et al.*, 2012). Chickens also only have 15 function D-genes, all of which are highly homologous. Some are in fact identical in amino acid content (e.g. D9/12/13 and D4/8/11) (Reynaud *et al.*, 1991; Wu *et al.*, 2012). The limited CDR3 V_H repertoire is hyper-diversified by the use of D-D junctions, somatic mutation and gene conversion. These D-like sequences may replace the entire D segment or

only a small section; this leads to the creation of “mosaic CDRs (Shih *et al.*, 2012; Wu *et al.*, 2012). These D-segments contain cysteine residues at a far higher frequency than humans or mice. This prevalence results in >50% of the chicken repertoire containing non-canonical CDRH3 cysteine residues and potentially plays an important role in functional diversity (Figure 4.1)

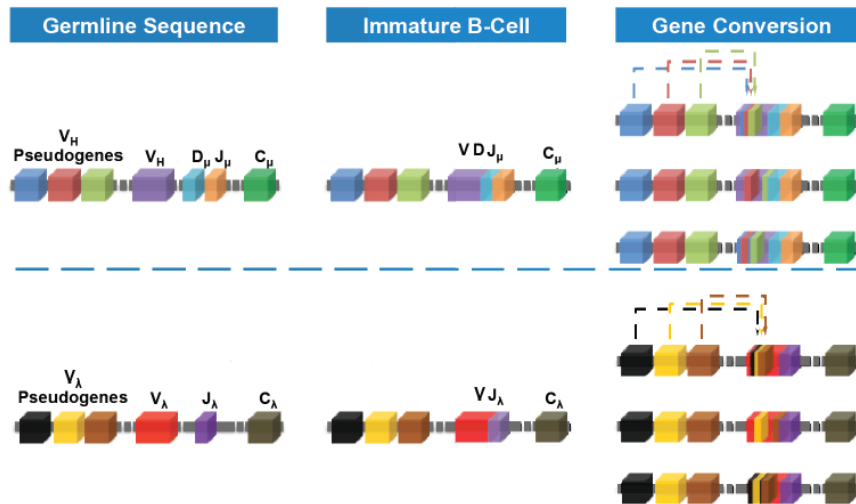
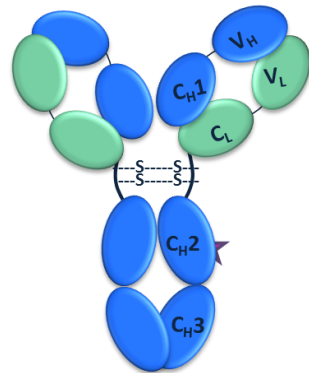


Figure 4.1: The process of gene conversion in chicken is illustrated for both the heavy and light chain. In the H germline, a functional V_H domain is comprised of unique V_H and J_μ gene segments with one of a family of D_μ elements (~15 D-family members). In the L germline, a single light chain exists (λ) and each light chain is composed of the same V_λ - J_λ arrangement, which in itself generates minimal diversity. A process known as ‘gene conversion’ which occurs in immature B-cells gives rise to diversity by translocation of pseudogene sequences into the V-genes. Typically the closest pseudogene is used more frequently in gene conversion; however, in the light chain the pseudogenes are distributed across a 20kb region preceding the V_L gene (Taken from Conroy *et al.*, 2014).

Chickens are an ideal immunological host for the generation of recombinant antibodies for a number of reasons (i) they can produce antibodies against highly conserved human proteins (ii) their relatively high core body temperature (41°C) which results in very stable antibodies in terms of half-life and (iii) there is considerably less germ-line diversity in chickens than in mice and rabbits thus fewer primers are required for amplification of the heavy and light chains (Andris-Widhopf *et al.*, 2000; Barbas *et al.*, 2001; Narat *et al.*, 2003).

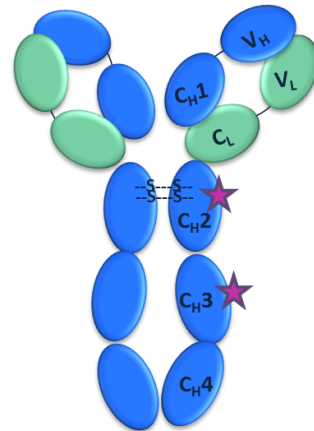
To further understand uniquely chicken attributes, the *N*-glycan content of IgY was also investigated in this body of work, using a fully quantitative high-throughput *N*-glycan analysis-based on ultra-performance liquid chromatographic (UPLC) separation of released glycans. IgY is more heavily glycosylated than its mammalian counterpart as it contains two potential *N*-glycosylation sites. One is located in Cv3 domain, that is absent in the mammalian IgG, and the other is located in the Cv2 domain which corresponds to the C_H2 (C_γ2) domain of mammalian IgG (Figure 4.2) (Narat, 2003; Suzuki and Lee, 2004).

A



Mammalian IgG (~ 150kDa)

B



Avian IgY (~ 180kDa)

Chicken
Duck
Human

```
-SQSPAVRSGS-TYSLSSRVNVSGTDWREGKSYSCRVRHPTNTVVE DHVKGCP-DGAQS
TMTRPQREAGSKTYMATSQT NVSREDWKAGKAFTRVKHPATGGTAQGHARFCPGSGAQS
-----GLYSLSSVVTVPSSSLG-TQTYICNVNHKPSNTKVDKKVEPKSCDKTHT
```

Chicken
Duck
Human

```
CSP-----IQLYAIPSP-GELYISLDAKLRCLVNLPS-DSSLSVTWTREKSGNL
CSP-----IQIFVVPSP-GSLYIRQDAKVHCLVNLPS-DASLSISWTREKSGAL
CPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH
```

Chicken
Duck
Human

```
RPDPMVLQEHFNGTYSASSAVPVSTQDWLSGERFTCTVQHEELPLPLSKSVYRNTGPTTP
RPDPMVLTEHFNGTFTASSSLAISTQDWLAGERFTCTVQHEDLPVPLGKSIAGHAGKVTA
NAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE
```

Chicken
Duck
Human

```
PLIYPFAPRPEELSLSRVTLSCLVRGFRPRDIEIRWLRDHRAVPATEFVTTAVLPEERTA
PYIFTFPPHAEELSIAEVTLTCLVRGFQPEHVEVQWLR-----
PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPE-----NNYKTT
```

Figure 4.2: Structures of Immunoglobulin G and Y and their N-glycosylation sites. (A) IgG (left) is composed of four constant heavy domains with a single carbohydrate site (purple star) on each heavy chain. In contrast, the additional constant domain introduces a second carbohydrate site (right). The flexible hinge region in the evolutionary ancestor (IgY) is absent and thus restricts flexibility in comparison to IgG. (B) Comparison of amino acid sequences around potential N-glycosylation sites on chicken, duck and human heavy chains. Potential N-glycosylation sites are indicated in red.

Structural characterisation of *N*-glycans present on antibody therapeutics is required under regulatory guidelines as the nature of these glycans can decisively influence the therapeutic performance of an antibody (Millán Martín *et al.*, 2015). The linked carbohydrate moieties of antibodies affect both their thermal stability and physicochemical properties, along with other crucial features like receptor-binding activity, circulating half-life and immunogenicity (Szekrényes *et al.*, 2015). *N*-glycan profiling of therapeutic antibodies with good reproducibility is vital to fulfil the needs of the both the biopharmaceutical industry and with national regulatory agency requirements (Hristodorov *et al.*, 2013).

4.2 Results

4.2.1 Immunisation of chickens cPSA antigen

An adult female Leghorn chicken was immunised with cPSA over a period of 3 months and the serum extracted and the development of the antibody response to the antigen was evaluated using an indirect ELISA (section 2.2.2.2). The serum was also assessed for a response to free PSA (fPSA) and ACT (protease inhibitor bound to cPSA). Evaluation of the animal's response is essential to ensure cPSA-specific antibodies were being generated and the titre was sufficiently high for recombinant antibody generation. The specific titres were determined as twice the background (control antigen) and are illustrated in Figure 4.3 and Table 4.1.

Table 4.1: Specific antibody titres obtained for cPSA immunised chickens

Antigen	cPSA immunised chicken
cPSA	1 in 1,000,000
fPSA	1 in 64, 000
ACT	1 in 256, 000

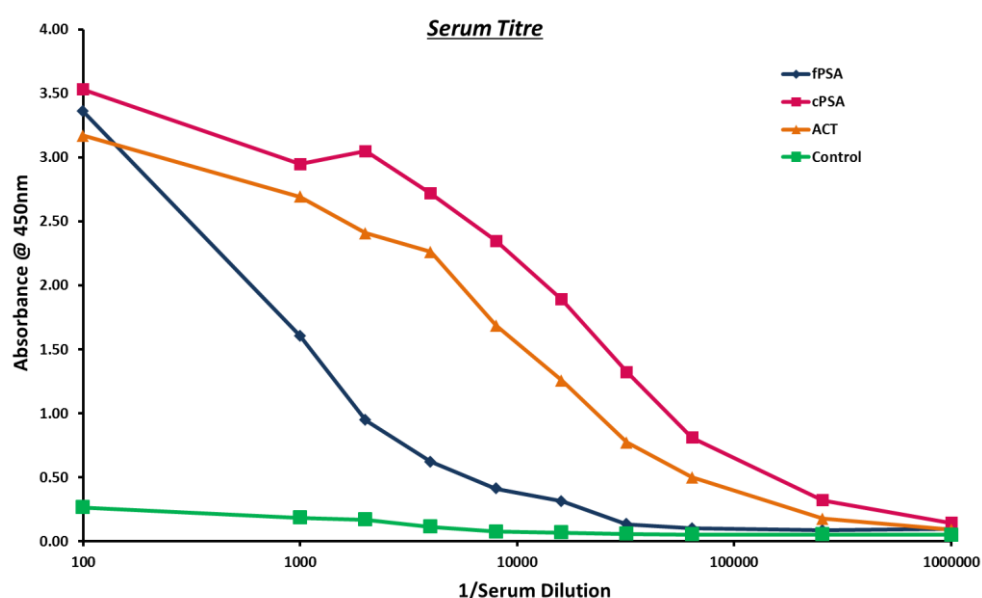


Figure 4.3: cPSA immunised chicken serum titrations. Serum was harvested from the immunised chickens and an antibody serum titre determined by ELISA. Bound avian antibodies were detected using an HRP-labelled anti-IgY-fab antibody followed by addition of the liquid substrate, TMB. The absorbance was read at 450nm. High serum responses for fPSA, cPSA and ACT can be seen, while the predominant response is specific for cPSA with the signal at twice the background at 1 in 1,000,000, there is also a significant response for fPSA (signal is twice the background at 1 in 64,000) and ACT (signal is twice the background at 1 in 256,000). Signal for the chicken was deemed sufficiently high for recombinant antibody library generation for cPSA-specific antibodies.

4.2.2 Optimisation and amplification of chicken heavy and light chain genes

As the chicken exhibited a sufficient response to cPSA for library building, the recombinant antibody library was built using the cDNA generated from RNA isolated and reverse transcribed to cDNA following extraction of the spleen and bone marrow using the methods described in sections 2.2.2.3 and 2.2.2.4.

The primers used for the amplification of the V_H and V_L chain domains with a long linker are listed in section 2.2.2.5. Initially, the $MgCl_2$ concentration was optimised for both the bone marrow (BM) and spleen (SP) for amplification of both the V_H and V_L in order to identify the optimal conditions for the PCR reaction that will give the optimal yield. The PCR products were resolved on 1.5% (w/v) agarose gels and visualised using the Gel Doc™ EZ system and Image Lab™, as illustrated in Figure 4.4.

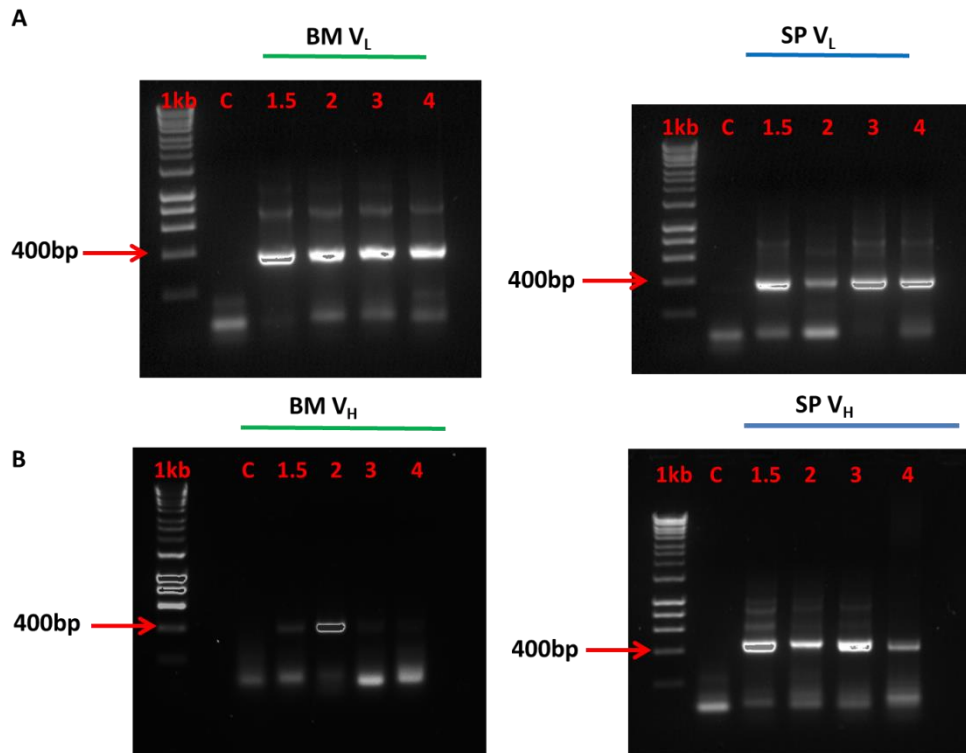


Figure 4.4: $MgCl_2$ optimisation for spleen (SP) and bone marrow (BM) for both for (A) V_L and (B) V_H amplifications. These reactions were carried out small-scale (1x) and the results revealed that (A) 1.5 mM $MgCl_2$ was optimal for both the SP and BM for the V_L amplifications, and (B) while 1.5 mM $MgCl_2$ was also optimal for the SP V_H amplification, 2.0 mM $MgCl_2$ provided the optimal yield for the BM V_H reaction. Negative reactions (Control –C) in the absence of cDNA showed no contaminating bands.

Large-scale (10x) reactions were then carried out using the optimal conditions obtained and the correct size amplicons were isolated by gel extraction and subsequent clean up (Figure 4.5).

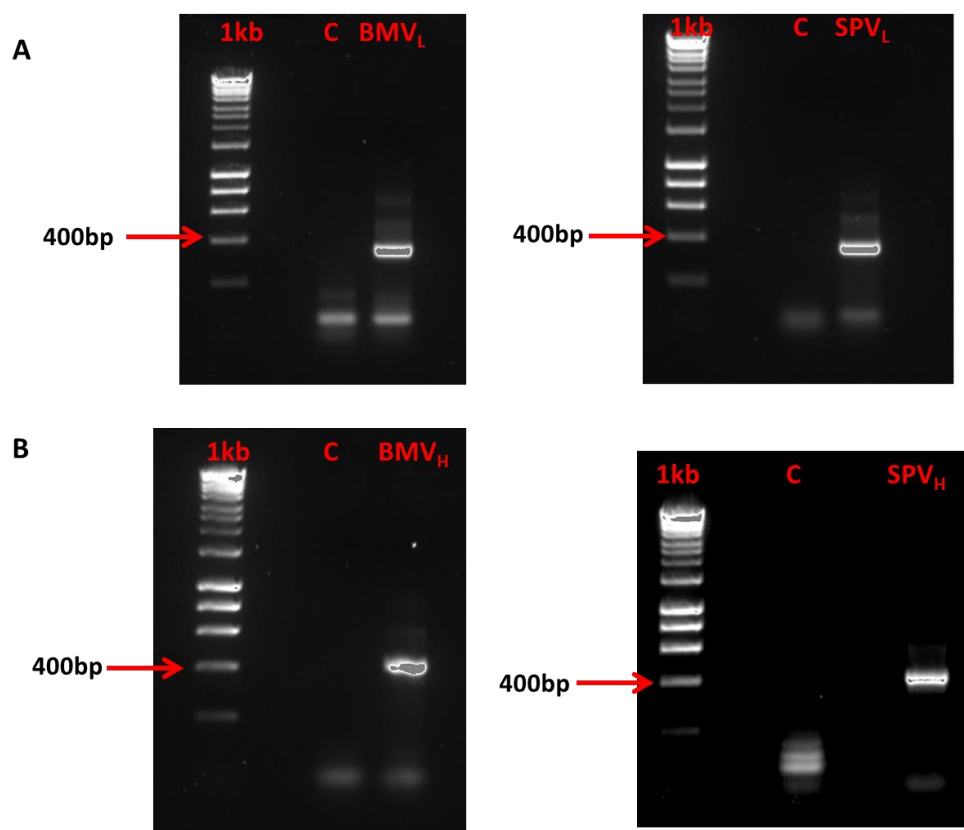


Figure 4.5: Optimised large-scale amplifications of V_L for BM and SP (A) and V_H for BM and SP (B) of cPSA immunised chicken. In both cases successful amplification of V_L (bands at ~ 350 bp) and V_H (bands at ~ 450 bp) were observed. The negative control showed no contaminating non-specific bands. The 1Kb DNA Hyperladder™ allowed approximation of fragment size following resolution on a 1.5% (w/v) agarose gel.

4.2.3 SOE-PCR of variable heavy and light chains

The avian scFv fragment was constructed by fusing equimolar concentrations of the purified variable light and variable heavy chains using the overlap primers, CSC-F and CSC-B, listed in section 2.2.2.5. These primers amplified the fusion product with incorporation of a long serine-glycine linker (GGSSRSSSSGGGGSGGGG). Platinum Taq DNA Polymerase enzyme was used for this step in the library building process as it is a robust system, providing highly specific amplification and a high yield. Platinum Taq requires the inclusion of

MgSO₄, which was initially optimised on a small-scale (1x) (Figure 4.6) to ensure the correct scFv fragment was amplified in a sufficient amount for digestion and subsequent ligation. The optimised reactions identified 2 mM as the best MgSO₄ concentration for both the BM and SP and thus was utilised for the large-scale (10x) reactions (Figure 4.7)

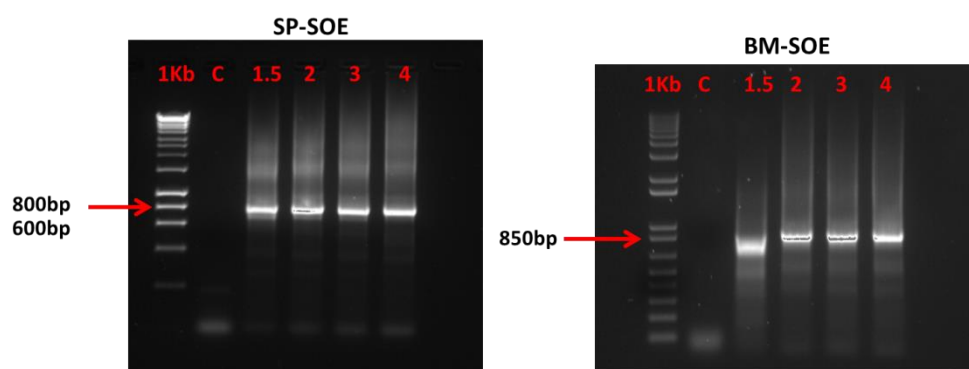


Figure 4.6: MgSO₄ optimisation for SOE-PCR. An equimolar mix of V_L and V_H was overlapped successfully as shown by the formation of a specific band at ~800bp, corresponding to the scFv SOE construct for both the spleen (A) and bone marrow (B). The optimal MgSO₄ concentration was identified as 2 mM for both the BM and SP. Negative control reactions show that no non-specific bands were amplified.

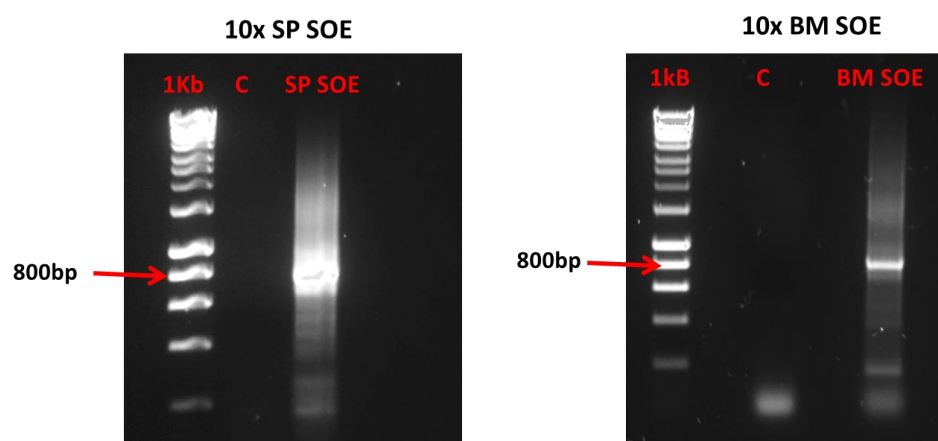


Figure 4.7: Large-scale amplification of (A) SP SOE and (B) BM SOE. The BM and SP scFv constructs were successfully amplified for the large-scale reaction using the optimised MgSO₄ concentrations in sufficient quantities necessary for the next stage of library construction. The negative control reactions show no contaminating bands.

4.2.4 Construction of scFv library in the pComb3xSS vector

Both the BM SOE and SP SOE constructs were cloned into the pComb3xSS by digestion with the restriction enzyme *SfiI*. A culture of the pComb3xSS vector was prepared (section 2.2.2.8) and plasmid DNA was purified using the Nucleospin Xtra midi kit. The pComb3xSS vector contains a double stuffer fragment between the two *SfiI* sites (Figure 4.8), and digestion with *SfiI* alone can result in a significant amount of intact stuffer fragment or undigested vector, which can ultimately lead to library contamination. To overcome this issue, following digestion with *SfiI*, the pComb3xSS vector was further degraded with *XbaI* and *XhoI*. Post triple digestion, the vector was de-phosphorylated via antarctic phosphatase treatment (Figure 4.9A). Antarctic phosphatase catalyses the removal of 5' phosphate from DNA. Phosphatase-treated fragments lack the 5' phosphoryl termini required by ligases, and, thus, cannot self-ligate. The fully digested and treated vector yield a single band at approximately 3,000 bp, which was resolved in a 0.7% (w/v) agarose gel (2.2.2.9). The scFv inserts were *SfiI* digested (Figure 4.9B) and following ethanol precipitation, were gel-purified, as previously described.

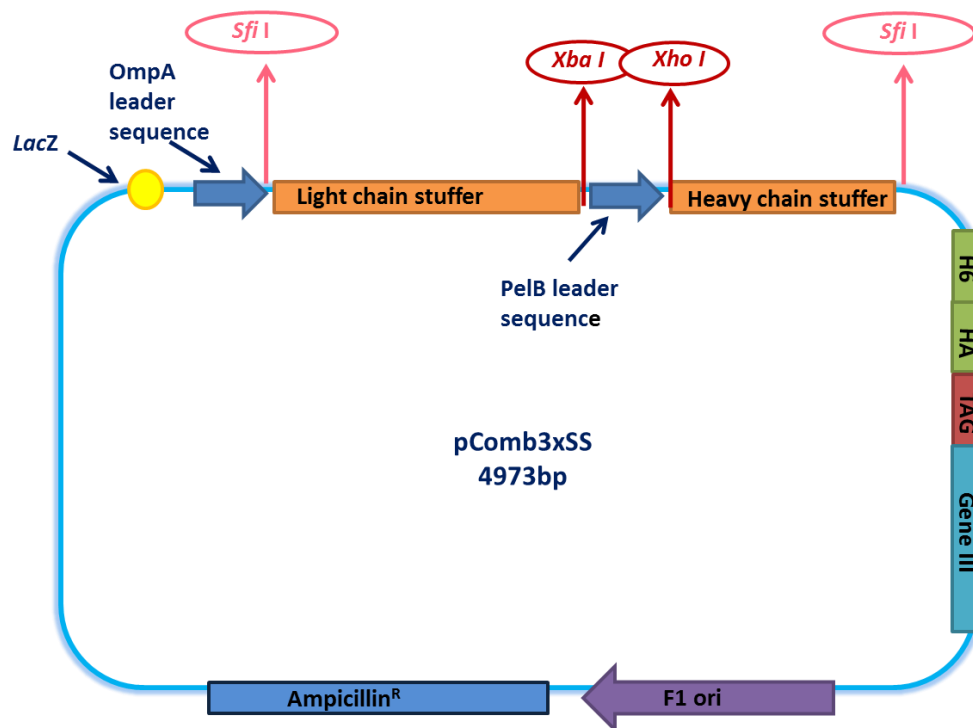


Figure 4.8: Schematic representation of the pComb3xSS vector. The pComb3xSS vector contains both *E. coli* and f1 bacteriophage origins of replication. The stuffer fragment is located between the *sfi*I restriction enzyme sites. The amber stop codon (TAG) inserted between the 3' *sfi*I site and the 5' end of gIII allows for soluble protein expression in non-suppressor strains of *E. coli*. The vector also contains a hexa-histidine (H6) tag and a haemagglutinin (HA) tag to aid purification and detection.

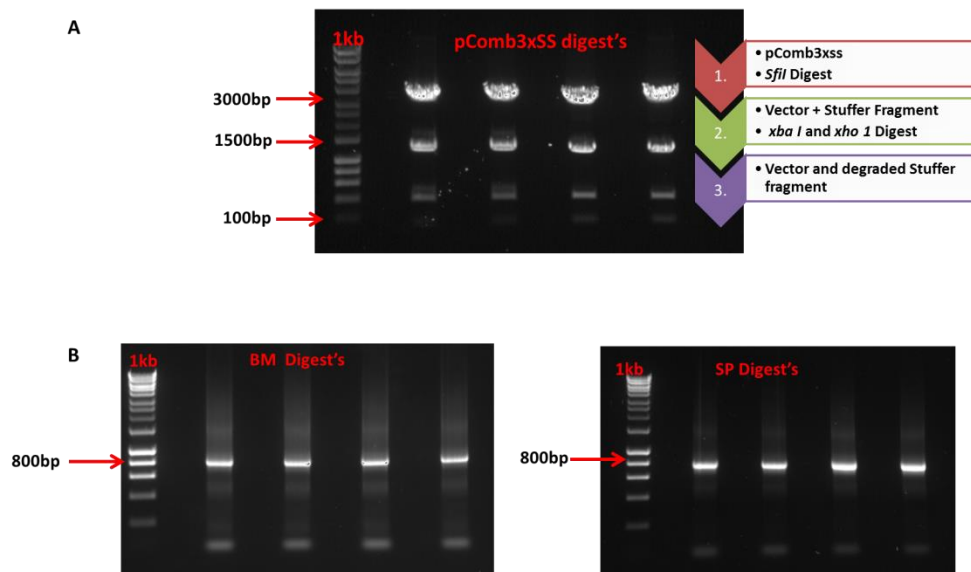


Figure 4.9: Restriction digests of (A) pComb3xSS vector and (B) scFv inserts. (A) The vector was digested initially with the enzyme *SfiI* and then *XbaI* and *XhoI*. This triple digestion ensures the stuffer fragment is fully degraded, reducing the risk of contamination in the later stages of the library building process. The vector was then antarctic phosphatase treated and the fully treated and digested fragment was resolved on a 0.7% (w/v) agarose gel and compared to the 1kb DNA ladder. (B) The BM and SP were successfully digested with *sfiI* and resolved on a 1.5% (w/v) agarose gel. Each digest was carried out four times to ensure there was sufficient product for the library ligation.

Following ligation of the purified digested inserts and vector, the antibody library was transformed into electrocompetent *E. coli*, generating a library size of 1×10^8 cfu/mL. The background (test for self-ligation/contamination) was less than 5%, which was assessed by control ligation (vector-no insert). The library was rescued with helper phage for panning and subsequently enriched for an anti-cPSA specific antibody.

Three rounds of panning were carried out. A negative selection step (subtractive/depletion panning against ACT) was introduced before panning

against the actual target, i.e. cPSA, to further minimise the retention of off-target binders.

Table 4.2: Panning input and output titres for avian anti-cPSA scFv library

Round	Input	Output
1	1.5×10^{13}	3.04×10^6
2	4.1×10^{11}	1.6×10^7
3	1.1×10^{13}	1.3×10^7

4.2.5 Polyclonal phage ELISA

Following three successive rounds of panning the scFv displaying phage from each round were analysed by polyclonal phage ELISA for enrichment against cPSA, as fully described in section 2.2.3.12. To monitor the selection process, phage pools acquired in each round of selection were tested in parallel with phage from the unpanned library. To increase confidence in the result, negative controls such as BSA and helper phage were included to observe the degree of non-specific binding. The phage pools were also assessed for possible enrichment against ACT, as an ACT-depletion step was employed in this panning strategy to enhance the selection of a highly specific cPSA clone.

The scFv-displaying phage particles were detected by an anti-M13 HRP-labelled antibody. Figure 4.10 shows an increase in absorbance from the first round onwards suggesting the presence of cPSA-specific scFv harbouring phage within the panned library. It is possible the library was highly biased towards ACT, as

there was no dramatic enrichment between the rounds. However, cPSA-specific antibodies are evidently present in abundance when compared to ACT. There was minimal or no background binding to any of the included controls.

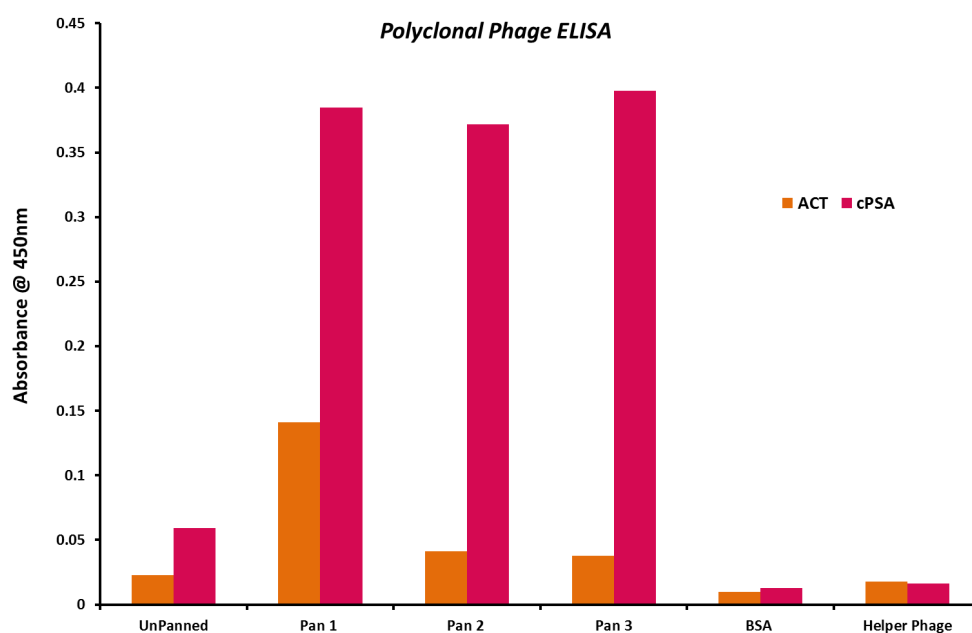


Figure 4.10: Polyclonal phage ELISA. Phage pools from the unpanned library (RO) and phage pools obtained after each round of panning were evaluated for reactivity to both cPSA and ACT by indirect ELISA. Bound phage was detected with a HRP-labelled anti-M13 conjugated antibody and the absorbance read at 450nm. Negative controls were also used including BSA and Helper Phage. No enrichment was observed between each round of panning.

4.2.6 Soluble monoclonal ELISA

Phage from each round was infected into mid-exponential Top 10F' *E. coli* for the soluble expression of scFv fragments and spread over the surface of agar LB agar plates supplemented with carbenicillin. Thirty clones were picked from each round and assessed for binding reactivity with cPSA by monoclonal ELISA. Colonies were randomly picked, cultured in deep-well 96-well sterile culture plates, and lysed as detailed in section 2.2.3.13. The soluble expressed antibody culture lysate was analysed for binding to a cPSA-coated ELISA plate (100nM cPSA). Bound cPSA-specific scFv antibody fragments were detected with a HRP tagged anti-IgY (H+L) antibody (Figure 4.11). Over 95% of the soluble-expressed scFv clones positively bound cPSA, highlighting the success of the panning experiment.

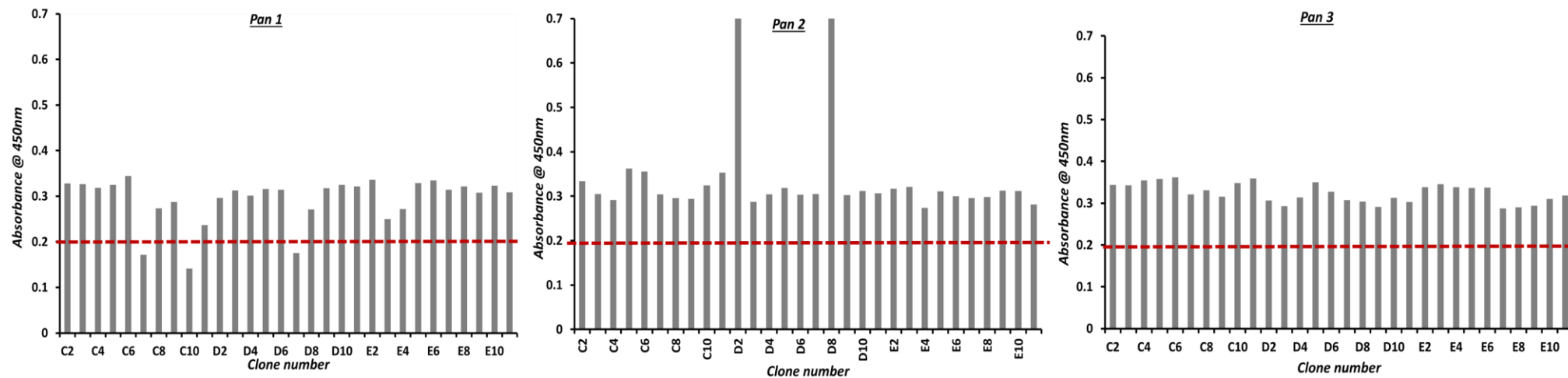


Figure 4.11: Soluble monoclonal ELISA of scFv clones from each round of panning. Sixty clones picked from each round of panning were randomly selected and assessed for binding to cPSA by indirect ELISA. Over 95% of the clones showed significant binding to the cPSA-coated plate, showing that the panning was highly successful. A selected cut-off of 0.2 was used (greater than three times the background- red dotted line).

4.2.7 Expression and purification of anti-cPSA scFv 2D2SG

Sequence analysis was carried out on two anti-cPSA scFv clones randomly selected (2D2SG and 2D8SG) by Source Bioscience Ltd, Dublin, Ireland. The DNA sequences obtained were translated using the translate tool from Expasy (<http://web.expasy.org/translate/>) and the amino acid sequences were then exported to PowerPoint and the significant regions highlighted (Appendix Three). Anti-cPSA clone 2D2SG was taken forward for small-scale expression and purification by Immobilised Metal Affinity Chromatography (IMAC).

A 100 mL culture of scFv 2D2SG was prepared and lysed, as detailed in section 2.2.2.14 (chapter two). Ni-NTA resin was applied to a 10mL chromatography column and once settled, the resin bed was equilibrated with 2 column volumes of PBS followed. The expressed antibody lysate was then applied to the column (twice). Samples were collected at each step of the purification process and analysed by SDS-PAGE and WB (Figure 4.12). The eluted protein was concentrated and buffer exchanged against PBS/0.02% (w/v) NaN_3 using a 10,000 molecular weight 'cut-off' vivaspin column. The purified concentrated protein was determined to be 0.6 mg/mL using the NanoDrop spectrophotometer.

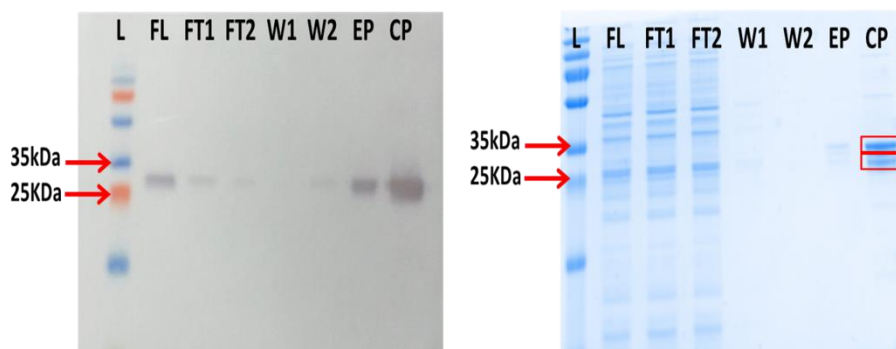


Figure 4.12: IMAC purification of anti-cPSA 2D2SG. A 100 mL culture was prepared and expressed antibody lysate purified by IMAC. Samples were taken throughout the process for analysis by SDS-PAGE (left) and WB (right) Specific detection of the scFv fragment was performed by probing the WB membrane with a HRP-tagged anti-HA antibody. The presence of 2 bands was observed in the SDS-PAGE gel (left), indicating that scFv 2D2SG possibly formed as a dimer. ScFv 2D2SG was successfully purified to 0.6 mg/mL. L: Ladder, FL: Filtered lysate, FT1: 'Flow-through' 1, FT2: 'Flow-through' 2, W1: Wash 1, W2: Wash 2, EP: Eluted protein, CP: Concentrated protein.

4.2.8 Cross-reactivity assay

To ensure anti-cPSA 2D2SG was specific for cPSA and not cross reactive with ACT or fPSA an indirect ELISA was performed, as described in section 2.2.2.15. Varying concentrations (doubling dilution; 100 nM, 50 nM, 25 nM 12.5 nM and 0 nM) of cPSA, ACT and fPSA were coated to the plate and a 1:10 dilution of IMAC-purified scFv 2D2SG was added. Bound antibody was detected using a HRP-labelled anti-HA antibody. From Figure 4.13 it can be seen that anti-cPSA is highly specific for cPSA and binds at low concentrations of 12.5 nM. There is no cross-reactivity with fPSA whatsoever; however there is some cross-reactivity to ACT.

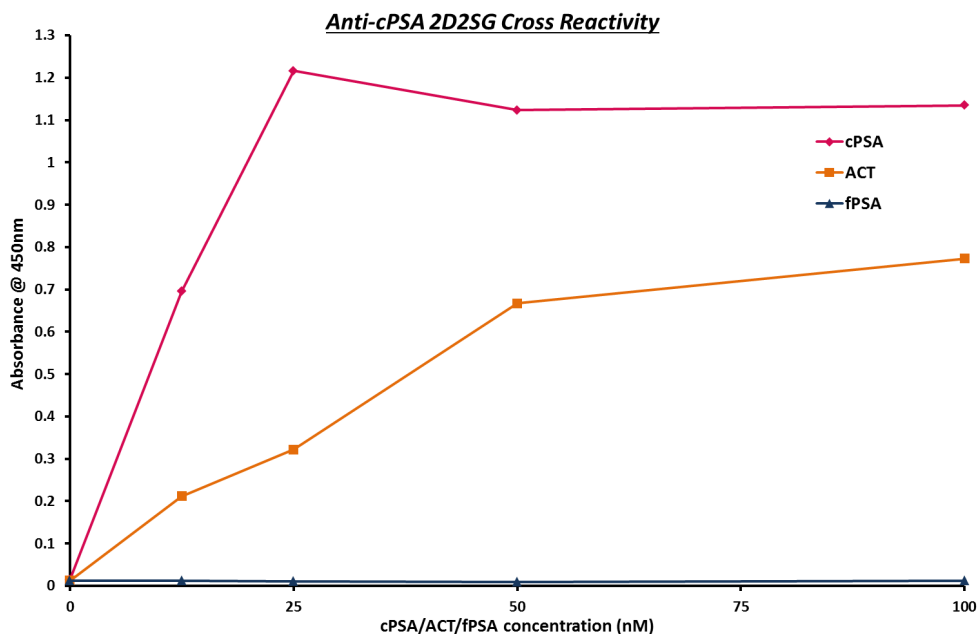


Figure 4.13: Cross reactive indirect ELISA for scFv 2D2SG. Varying concentrations of cPSA/ACT/fPSA (100 - 0 nM) were coated an ELISA plate. Binding of anti-cPSA scFv 2D2SG was then detected with HRP-labelled anti-HA, followed by addition of the liquid substrate TMB. The absorbance was read at 450nm. scFv 2D2G does not cross react with fPSA, however does cross-react with ACT; scFv 2D2SG binds approximately 2-fold less to ACT compared to cPSA.

4.2.9.1 IgY Purification

Immunoglobulin Y differs from most of the other immunoglobulins as it does not bind protein A or protein G (Constantinoiu *et al.*, 2007). IgY was successfully recovered from the serum of chickens using thiophilic adsorption. Total protein concentration was determined to be 11.5 mg/mL by spectrophotometry at 280 nm (NanoDrop 1000). The heavy and light chains of the purified IgY were visualized by reducing SDS-PAGE and Western Blot analysis (Figure 4.14).

The SDS-PAGE analysis revealed the presence of some contaminating bands (above 70kDa) indicating the presence of non-specific avian serum proteins. The IgY was estimated to be approximately 70% pure, which is in line with that expected from use of the Pierce Thiophilic adsorption KitTM. Constantinoiu and colleagues (2007) showed that IgY purity can be increased, without comprising recovery by reduction of the potassium salt from 0.5 M (as used in Pierce Thiophilic adsorption KitTM) to 0.2 M (Constantinoiu *et al.*, 2007).

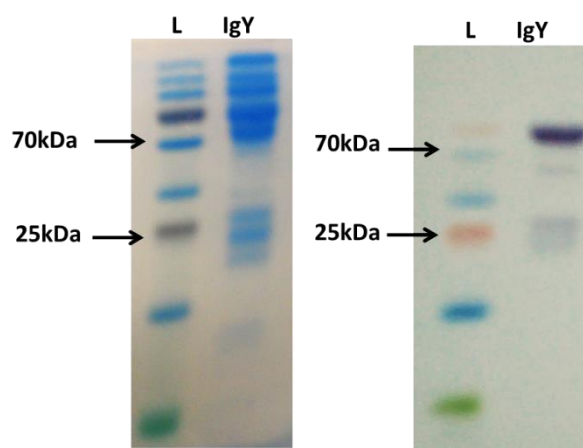


Figure 4.14: SDS-PAGE and Western Blot analysis of IgY purified from serum. Approximately 500 µg of purified IgY was resolved on 12% SDS-PAGE gels and visualised by staining with InstantBlue (left). The resolved proteins were also transferred to nitrocellulose membranes and the presence of the heavy chains at approximately 65-68 kDa and the light chains at 25 kDa can be seen after probing with an anti-IgY H+L-HRP-tagged antibody (right). L: PageRuler Plus Prestained protein ladder.

4.2.9 *N*-glycan profiling of avian polyclonal IgY antibodies

This study describes the detailed *N*-glycan profile of IgY polyclonal antibodies from the serum of leghorn chickens using a fully quantitative high-throughput *N*-glycan analysis approach, based on ultra-performance liquid chromatography (UPLC) separation of released glycans.

4.2.9.2 IgY *N*-glycan profiling

The *N*-glycans released from the IgY purified from avian serum (Figure 4.14) were analysed by WAX-HPLC and UPLC-HILIC in combination with exoglycosidase digestions and structural assignments using established methods (Royle, Campbell, Radcliffe, White, Harvey, Abrahams, Kim, Henry, N. A. Shadick, et al., 2008) and the software tool GlycoBase (<https://glycobase.nibrt.ie>). UPLC-FLR-QTOF MS analysis was also carried out for comparative analysis. Annotation of the *N*-glycans present in each chromatographic peak was based upon the oligosaccharide composition as derived from the *m/z* value.

Over 80 different glycans structures were assigned to 40 peaks (Figure 4.15 and Appendix Four). *N*-glycan structures annotated include high mannose, hybrid and complex glycans with variable degrees of core fucosylation, galactosylation and sialylation. The presence of highly complex glycans that contain more sialic acids than branches (known as polysialic acids) was also observed.

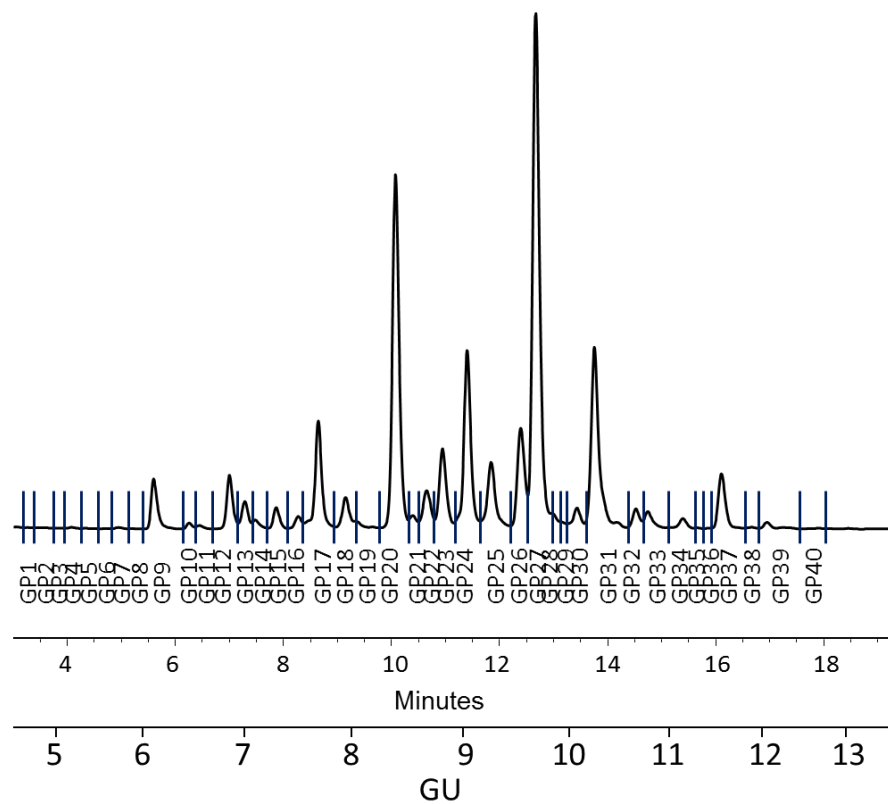


Figure 4.15: HILIC-UPLC profile of undigested N-glycans from serum IgY. Profiles are standardised against a dextran hydrolysate (GU). The HILIC-chromatogram was separated into 40 peaks. (Final IgY structural assignments are listed in Appendix Four).

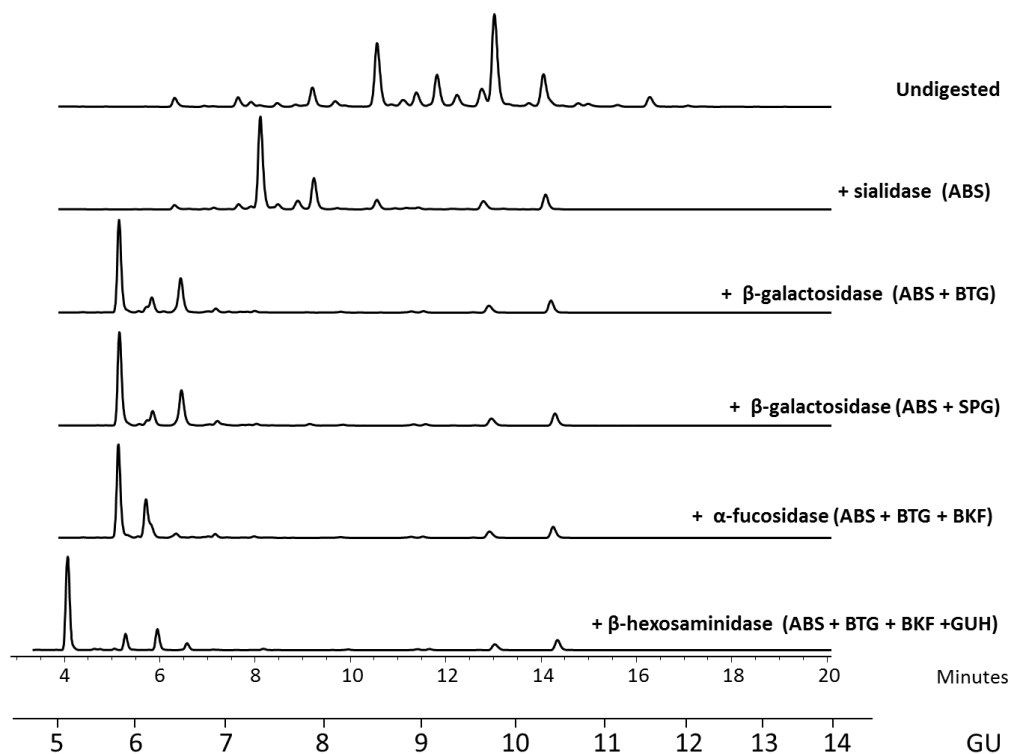


Figure 4.16: *Unfractionated IgY profile was subjected to exoglycosidase digestions.* In order to confirm the structures, the sequence of particular glycan moieties can be elucidated by sequential exoglycosidase digestion. The GU for each glycan is directly related to the number and linkage of its constituent monosaccharides; the larger the glycan, the higher its GU. Therefore, the removal of terminal sugars by exoglycosidase digestion results in a movement of the peak to a lower GU. (Final IgY structural assignments are listed in Appendix Four).

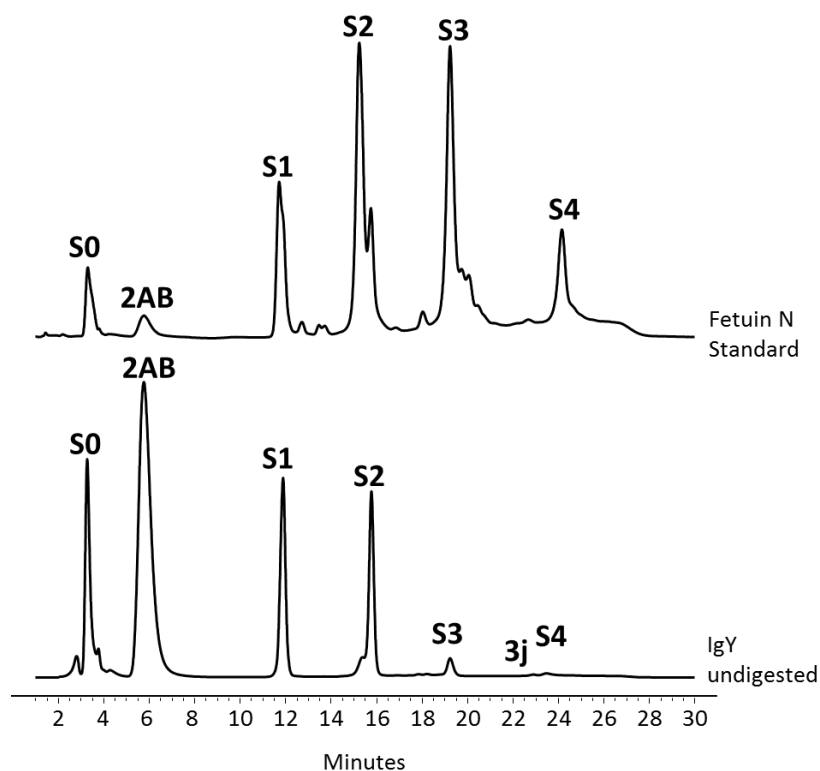


Figure 4.17: IgY glycans separated according sialic acid content by WAX-HPLC. Where a sample contains a large number of structures with varying degrees of sialylation, it is helpful to first separate the pool into fractions by charge (i.e., containing zero, one, two, . . . sialic acids) using WAX-HPLC because several peaks may co-elute on UPLC-HILIC. The WAX-chromatogram was separated into 5 peaks: S1, S2, S3, S3j and S4. (Final IgY structural assignments are listed in Appendix Four).

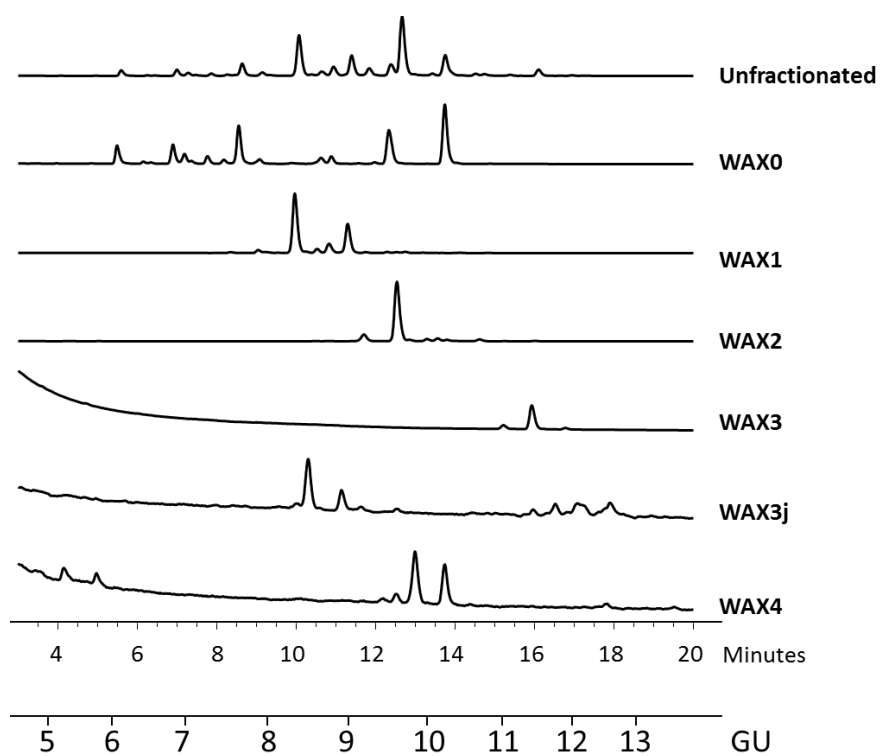


Figure 4.18: WAX fraction's separated by HILIC-UPLC. The differently charged fractions are analysed by UPLC-HILIC following WAX-HPLC to determine all of the structures present in each charged fraction. (Final IgY structural assignments are listed in Appendix Four).

4.2.9.2 Molecular Modelling of IgY

Molecular Modelling of IgY was carried out by Dr. Mark Wormald (Oxford University). The precise glycans chosen for sites N390 (M9Glc/peak 31 and A2G2S1/peak 20) were based on the site specific glycan analysis of IgY (Suzuki and Lee, 2004). The glycans chosen for sites N292 (A2G2S2/peak 27 and FA3G3/peak 24) were representatives from the other largest peaks. Glycan structures were generated using the database of glycosidic linkage conformations (Wormald *et al.*, 2002) and *in vacuo* energy minimisation to relieve unfavourable steric interactions. The Asn-GlcNAc linkage conformations were based on the observed range of crystallographic values (Petrescu *et al.*, 2004), the torsion angles around the Asn C α -C β and C β -C γ bonds then being adjusted to eliminate unfavourable steric interactions between the glycans and the protein surface (Figure 4.19).

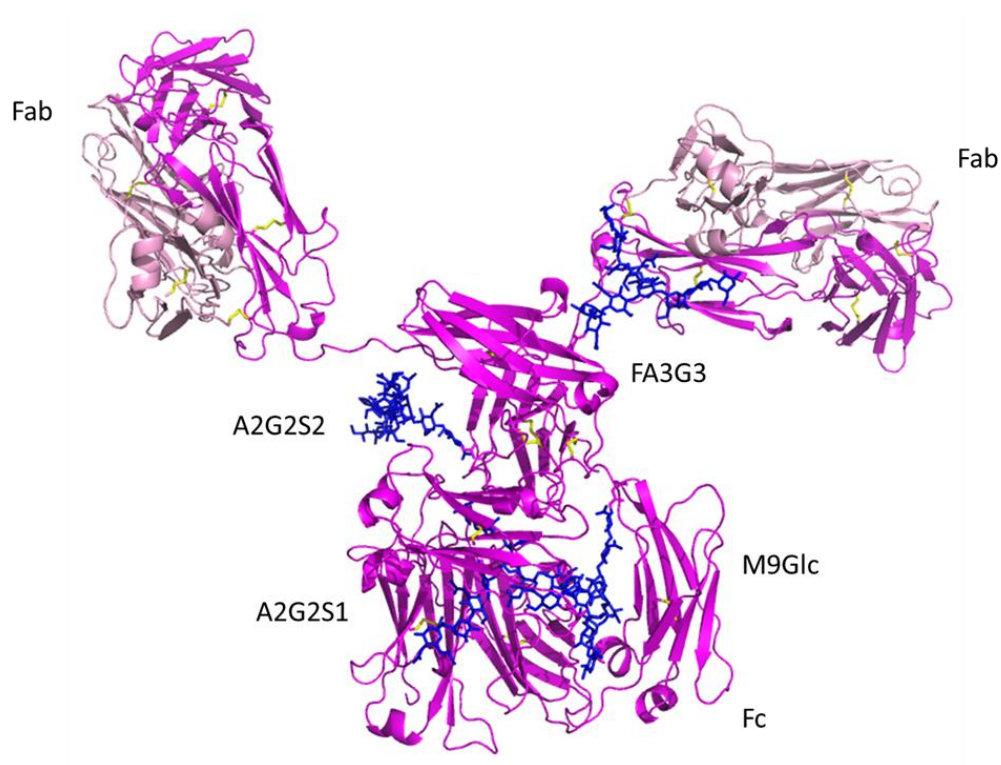


Figure 4.19: Molecular model of glycosylated IgY. Magenta – heavy chain; pink – light chain; yellow–stick representation of cysteine residues; blue – glycans. The glycans are labelled according to the nomenclature in Appendix Four.

4.3 Discussion

This chapter describes the generation of polyclonal and recombinant avian antibodies, and the subsequent characterisation of the glycan content of the avian polyclonal antibodies.

Initially, the construction of a chicken anti-PSA scFv library and subsequent screening of the library for identification of a cPSA-specific antibody was described. Sufficient quantities of V_H and V_L were amplified and the desired single chain Fv fragment was successfully assembled with a flexible glycine-serine linker in the V_L -(Gly₄Ser)₄- V_H orientation. A scFv phage library with a combined (bone marrow and spleen) size of 1×10^8 cfu/mL was generated. The generated library size indicates the number of possible clones which could potentially bind cPSA. A library size of 10^7 after electroporation is deemed sufficient for identification of the target antigen. The library size is dependent on both the concentration and quality of the DNA and the efficiency of the *E. coli* cells (Andris-Widhofp *et al.*, 2000; Barbas *et al.*, 2001).

When developing the panning strategy, a negative selection step (i.e., subtractive panning against ACT) was introduced before the positive one (panning against cPSA) to further minimize the retention of ACT-specific binders. After three rounds of subtractive-based panning against cPSA immobilised on a solid support, polyclonal ELISA confirmed the presence of cPSA-specific antibodies (Figure 4.10). The negative selection step was highly successful in depleting the library of ACT-specific antibody fragments, hence no enrichment for cPSA was observed. Individual clones were analysed by monoclonal ELISA and over 95% of the clones successfully bound to the cPSA-coated wells.

Two clones from pan 2, 2D2SG and 2D8SG (Figure 4.11), showed a high degree of antibody expression when grown in small-scale (96-well plates) and an exceptionally high degree of

binding to immobilised cPSA (in comparison to the other clones analysed in the monoclonal ELISA). These two clones were sent for sequencing to check for inter-clone sequence diversity. The DNA sequences were translated using the translate tool from ExPASy and the acquired amino acid sequences were then aligned and the complementary determining binding (CDR) regions of the antibody evaluated using ClustalW2 software. The CDR regions of the antibodies were highlighted according to the Kabat scheme (Figure 4.20)

		CDRL1		
2D2SG	AGLTQPSSVSANPGETVKIX	<u>C</u> SGS--- <u>SSAYEY</u> GWY QQKSPGSAPVTL IY	47	
2D8SG	AALTQPSSVSANPGETVKITC	<u>S</u> GGGRYAGSY <u>Y</u> GWY QQKSPGSAPVTL IY	50	
		CDRL2		CDRL3
2D2SG	<u>NNDKR</u> P S	DIPSRFSGSKFGSTATLTITGVQADDEAVYF	<u>C</u> GSADSNADGI <u>F</u>	97
2D8SG	<u>ESNKR</u> P S	SNIPSRFSGSTSGSTGTLTITGVQADDEAVYF	<u>C</u> GSRDSTY-AA <u>F</u>	99
2D2SG	<u>GAG</u> TTLTVL GQSSRS -----	S AVTLDESGGGLQTPGGALSLV C K	136	
2D8SG	<u>GAG</u> TTLTV- GSSSRSSSGGSSGGGGS	AVTLDESGGGLQTPGGALSLV C K	147	
		CDRH1		CDRH2
2D2SG	<u>ASGF</u> S FSY AMe t GW VRQAPGK G <u>LEW</u> VAD I-- GS TTWYGA AVK GRATIS	184		
2D8SG	<u>ASGF</u> T FS SHG Me t <u>FW</u> VRQAPGK G <u>LEW</u> VAG IDGGGGVTWYGA AVK GRATIS	197		
			CDRH3	
2D2SG	RDNRQSTMetRLQLNNLRAEDTGVYY	<u>C</u> AKYVVSTGGRWGVYSIDAWGH <u>GT</u>	234	
2D8SG	RDNGQSTV--RLQLNNLRAEDTATYY	<u>C</u> ARESYSSG ----- TID I WGR <u>GT</u>	239	
2D2SG	EVIVSSTSGQAGQ	HHHHHHGA YPYDV PDYAS	265	
2D8SG	EVIVSSTSGQAGQ	HHHHHHGA YPYDV PDYAS	270	

Figure 4.20: Sequence alignment of anti-cPSA 2D2SG and 2D8SG. The CDRs are highlighted using the Kabat numbering system. Conserved CDR flanking regions are highlighted (bold and underlined). Light chain CDRs are indicated in blue with the heavy chain CDRs in green. Other highlighted regions: linker; pink, hexahistidine tag; purple and HA-tag; gold. Sequences translated using ExPASy tool (<http://web.expasy.org/translate/>) and aligned using ClustalW (http://www.ebi.ac.uk/Tools/psa/emboss_needle/)

The diversity between the clones is evident from the sequence variation between the heavy and light chains, in particular CDRH3, which is unsurprising given that highly diverse CDR3 loops are the key determinant of specificity in antigen recognition. The CDRH3 of anti-cPSA 2D2SG is quite long (19 residues) with over 40% comprised of small amino acids (A/G/S), the prevalence

of small residues in the CDR loops confers conformational flexibility which is crucial for effective antigen recognition. This CDRH3 notably also contains a single tryptophan and two tyrosine residues. Tyrosine, tryptophan and arginine residues are capable of mediating a wide-array of intermolecular interactions which are desirable in the composition of the CDRs (Birtalan *et al.*, 2008).

Following sequence alignment, it was evident that the linker in scFv 2D2SG is significantly shorter than expected, being only 7 residues in length and contains a serine to glutamine point mutation at the start (GQSSRS-----S). The presence of this linker is highly unusual given that the primers used to generate the overlapping of the V_L and the V_H domains encoded for a long linker, 18 residues in length (GGSSRSSSSGGGSGGGG). Anti-cPSA 2D2SG was expressed on a small-scale and purified to 0.6 mg/mL by IMAC and, interestingly, the presence of two bands can be seen in the SDS, both at approximately 28kDa (Figure 4.12). Linkers of different lengths can result in levels of oligomerisation and may alter the interface of the V_H and V_L domains (Barbas *et al.*, 2001). scFv fragments can be expressed as a monomer, dimer or multimers depending on the length of the linker. Connections of the light and heavy variable regions with long linkers appear to favour monomeric molecules whereas short linkers tend to express as non-covalent dimers. A scFv molecule with a linker of 3 to 12 residues cannot fold into a functional Fv domain and instead associates with a second scFv molecule to form a bivalent dimer (referred to as a diabody, approximately 60 kDa). Diabodies, have several advantages over monomeric antibody fragments including high binding avidity, which can lead to an increase in affinity constants (higher 'off-rates')(Hudson and Kortt, 1999; Holliger and Hudson, 2005; Wang *et al.*, 2008).

In a diabody, the two pairs of domains can pack together with the antigen-binding sites pointing in opposite directions, resulting in bivalent or bispecific interactions. In figure 4.13, we can see that 2D2SG binds to both ACT and cPSA. This bivalent binding favours the concept that 2D2SG could potentially have formed as a dimer. However, significant further analysis would need to be carried out to confirm this. The comparative avidity of a multivalent format depends on numerous factors, such as the density and accessibility of adjacent epitopes as well as the conformational flexibility of both target antigens (Holliger and Hudson, 2005).

Ideally, non-reducing SDS-PAGE analysis would confirm dimerisation. A bivalent antibody fragment contains a cross-paired scFv resulting in a fragment of ~55 kDa in size. The majority of SDS-PAGE gels run under reducing conditions which incorporates agents to break bonds that potentially exist between multimeric proteins. Thus, the use of non-reducing (non-denaturing) SDS-PAGE conditions maintains the proteins' secondary structure and native charge density (Corley, 2005). The presence of a protein band at approximately 55 kDa would indicate that scFv 2D2SG has formed a dimeric structure and is indeed a diabody. Also, size-exclusion chromatography (SEC) would aid in determining the presence of a dimeric format. SEC is a separation technique based on the molecular size of the components. Separation is achieved by the differential exclusion of the sample molecules from the pores of the packing material as they pass through a bed of porous particles. Molecules of small size will enter the pores whereas larger molecules will flow quickly through the matrix and exit the column first (Corley, 2005). Hence, a dimeric scFv (55 kDa) will elute from the column before a monomeric scFv. Additionally, structural analysis of this antibody fragment alone, and in contact with cPSA and/or ACT would provide the greatest insight into its multivalent binding capabilities.

Avian serum IgY is functionally equivalent to mammalian IgG, as detailed in the introduction to this chapter. However, it contains an additional constant region and is significantly more glycosylated. To clarify the particular structural glycan characteristics of the IgY molecule detailed analysis of the *N*-glycan profile of IgY polyclonal antibodies from the serum of leghorn chickens was also investigated.

Structural assignments revealed that serum IgY contains mainly high mannose, bi-, tri- and tetraantennary glycans with or without core fucose and bisects, all with varying levels of galactosylation and sialylation (See appendix two for full list of structural assignments).

In investigating the site-specific *N*-glycosylation of IgY, Suzuki and Lee (2004) noted the Fc portion of IgY possesses a *N*-glycosylation site which is structurally equivalent to conserved glycosylation sites of other Ig classes in mammals and is composed of predominantly high-mannose type oligosaccharides (Suzuki and Lee, 2004). High-mannose glycans observed in this study are rarely found in mammalian glycoproteins (Kondo *et al.*, 2012). This uniquely avian glycosylation pattern at the conserved *N*-glycosylation site is thought to be attributed to the structural differences between IgG and IgY (IgY lacks the hinge region seen in IgG) (Kondo *et al.*, 2012). The additional *N*-glycosylation site, located in C_v2 domain, was previously shown to contain exclusively complex-type oligosaccharides (Suzuki and Lee, 2004; Kondo *et al.*, 2012). These distinct avian glycosylation patterns and structural differences provide IgY with distinct biochemical properties and behaviour. A model of IgY with the glycans identified from this study was generated following glycan assignment (Figure 4.19). Characterisation of the individual glycans decorated on a protein is essential for detailed understanding of structure/function relationships and the design of potential therapeutic agents. The model generated from this study aims to enhance our understanding of the therapeutic potential of

IgY. Computational modelling methods are universally accepted as central tools in the invention process for many biopharmaceuticals, facilitating drug development areas, such as optimising affinity for a target while minimising cross reactive effects, alongside optimising pharmacokinetic properties (Kuroda *et al.*, 2012).

The oligosaccharide content of therapeutic immunoglobulins plays a significant role in its bioactivity and pharmacokinetic (PK) activity. Raju and colleagues (2000) examined at variations between the glycan content of IgG across several species. These authors highlighted the importance of choosing the right host in generating therapeutic IgG as the terminal sialylation of IgG is species-specific (Raju *et al.*, 2000). In this study, MALDI-TOF-MS analysis of chicken IgY suggested the *N*-linked glycosylation of chicken IgY is considerably more heterogeneous than its human counterpart, IgG.

These results correlate with previous studies (Suzuki and Lee, 2004; Kondo *et al.*, 2012; Raju *et al.*, 2000) and also highlight several previously unidentified structures. A high sialic acid content was observed with many sialic acid isomers (same composition but different sialic acid linkage arrangements resulting in different GU from the original structure). The presence of unusual sialic acids was also noted, and these are likely to be polysialic acids or sialic acid (Sia) is linked on *N* acetylgalactosamine (GalNAc) as well as on Galactose. Sias are most commonly α 2-3 or α 2-6 linked to galactose (Gal) or α 2–6 linked to GalNAc. However, Sia can also be found linked to *N*-acetylglucosamine (GlcNAc) or to another Sia in α 2–8 or α 2–9 linkages (Cohen and Varki, 2010). Polysialic acids occupy internal positions within glycans, the most common being one Sia residue attached to another, often at the C-8 position (Varki *et al.*, 2009b).

The high sialic acid content of IgY is very important when considering IgY as a therapeutic agent as the level of sialic acid can have a significant impact on the PK of therapeutic antibodies. A

lower content of total sialic acids can significantly reduce half-life (Liu, 2015). Hence, the high sialic acid content of IgY that was observed suggests IgY-based biotherapeutics could have potentially extended circulating half-lives and are promising candidates for use against a variety of pathogens.

Certain glycan structures have a direct impact on the immunogenicity of therapeutic proteins i.e., their presence can affect protein structure in such a way that the protein becomes immunogenic. However the glycan structure itself can also induce an immune response. The sialic acid N-glycolylneuraminic acid (Neu5Gc) and terminal galactose- α -1,3-galactose are examples of such structures that are not naturally present in humans and are known to be immunogenic when used as therapeutics (Van Beers and Bardor, 2012). These non-human antigenic structures could promote clearance of a biopharmaceutical preparation from circulation (Tangvoranuntakul *et al.*, 2003; Chung *et al.*, 2008; Padler-Karavani *et al.*, 2008).

The chimeric mouse–human IgG1 monoclonal antibody, Cetuximab, is an anti-human epidermal growth factor receptor (EGFR) antibody used for the treatment several cancers (Chung *et al.*, 2008). High incidences of hypersensitivity reactions to Cetuximab were reported and a study by Chung and colleagues showed that the majority of patients who had a hypersensitivity reaction to Cetuximab also had circulating IgE antibodies against Cetuximab before therapy was initiated. These antibodies were specific for the glycan structure galactose- α -1, 3-galactose, which is present on the Fab portion of the Cetuximab heavy chain (Chung *et al.*, 2008). In order to overcome these severe hypersensitivity reactions which are observed with many immunoglobulin-based biotherapeutic agents it is of primordial importance to ensure the oligosaccharide content will not elicit such reactions. Recently, a glyco-engineered anti-EGFR monoclonal antibody with a lower α -Gal content than Cetuximab was developed (Yi

et al., 2014), highlighting the importance of these structures to the biopharma industry in the development of novel biotherapeutics.

While IgY is heavily decorated with complex glycans, no non-human immunogenic structures were identified. The results from this study, combined with other known advantages of chicken antibodies, such as increased stability over IgG and phylogenetic distance from man (Spillner *et al.*, 2012), makes chickens ideal hosts for the generation of novel oral therapeutic interventions for the treatment of numerous infectious diseases.

Chapter conclusions

Overall, in this chapter, antigen-specific recombinant antibody fragments were successfully selected from the generated immune libraries by phage display. Uniquely, chicken antibody *N*-glycan features were also investigated and it was concluded that IgY is heavily decorated with complex glycans. However, no known non-human or immunogenic glycans were identified.

Future work

Extensive characterisation of the anti-cPSA antibodies generated from the cPSA immunised chicken needs to be carried out including optimisation of the expression and purification conditions to generate high yields of pure protein. Determination of the Limit of Detection (LoD) of these antibodies should be performed to define their limits for use in a working assay. Kinetic characterisation using the Biacore 3000 should also be carried out to determine the on/off rates of these antibodies so they can be effectively incorporated into microfluidic-based

platform for PCa detection. It is anticipated that anti-fPSA scFv B8 (Chapter 3) and anti cPSA 2D2SG and 2D8SG will be utilised, along with other PCa –specific antibodies, in a mutilplex assay for improved PCa diagnosis.

Chapter Five

***Differentiation of indolent,
significant and aggressive prostate
cancer by robotic high-throughput
N-glycan profiling***

5.1 Introduction

In this chapter, differential glycosylation patterns of PCa serum was investigated to see if it could distinguish between 3 cohorts of PCa patients, namely, indolent, significant and aggressive PCa. A total of 117 patients (indolent [41], significant [32] and aggressive [44]) were included in this study (Figure 5.2) (Patient clinical data in Appendix Five). With respect to patient enrolment in the study, all men going for a radical prostatectomy in Ireland were eligible and there were no exclusion criteria as long as they were suitable for surgery. No patients withdraw from the study following their surgery or during their follow up. Patient data was collected, including medical history, current medication and family cancer history but it was not included in the analysis.

Here, whole serum *N*-glycan profiling was carried out on all 117 prostate cancer patients' sera using a highly efficient robotised, high-throughput analysis platform for glyco-profiling with UPLC (Stöckmann *et al.*, 2013, 2015).

The decoration of the vast majority of secreted eukaryotic proteins with glycan structures is one of the most common post translational modifications (Dube and Bertozzi, 2005). The glycome can change dramatically in response to a subtle change in the cellular environment and it is now universally accepted that glycan expression patterns change with the cellular condition modifications that accompany the onset of tumourigenesis (Varki *et al.*, 2009; Freeze, 2006).

The size, diversity and complexity of glycans found in mammalian systems provides a wealth of information and a significant analytical challenge scientists are faced with is to understand in detail which glycans contribute to specific biological functions (Stöckmann *et al.*, 2013).

However, within the past decade, advances in the field of proteomics and glycomics has helped glycobioologists decipher the link between glycan structures and disease progression (Dube and Bertozzi, 2005).

The high throughput methods devised by Professor Pauline Rudd's group in The National Institute for Bioprocessing Research and Training (NIBRT) provide a superbly reliable and robust method that allows for detailed analysis of femtomolar quantities of *N*-linked sugars released from glycoproteins. This approach was designed specifically for the identification and high-throughput screening for disease biomarkers (Figure 1.3) (Campbell et al., 2008b; Royle, Campbell, Radcliffe, White, Harvey, Abrahams, Kim, Henry, N. A. Shadick, et al., 2008).

An automated workflow for sample preparation implemented on a robotic liquid handling platform was developed to improve on the time and efficiency of the methods previously developed by Royle and colleagues (Royle *et al.*, 2008). This system allows for the preparation of 96 samples (or multiples thereof) in 22 hours with excellent reproducibility. It also improves on the HPLC technique by use of hydrophilic interaction chromatography ultra-performance liquid chromatography (HILIC-UPLC) which allows for shorter run times and greatly increased resolution (Figure 5.1).

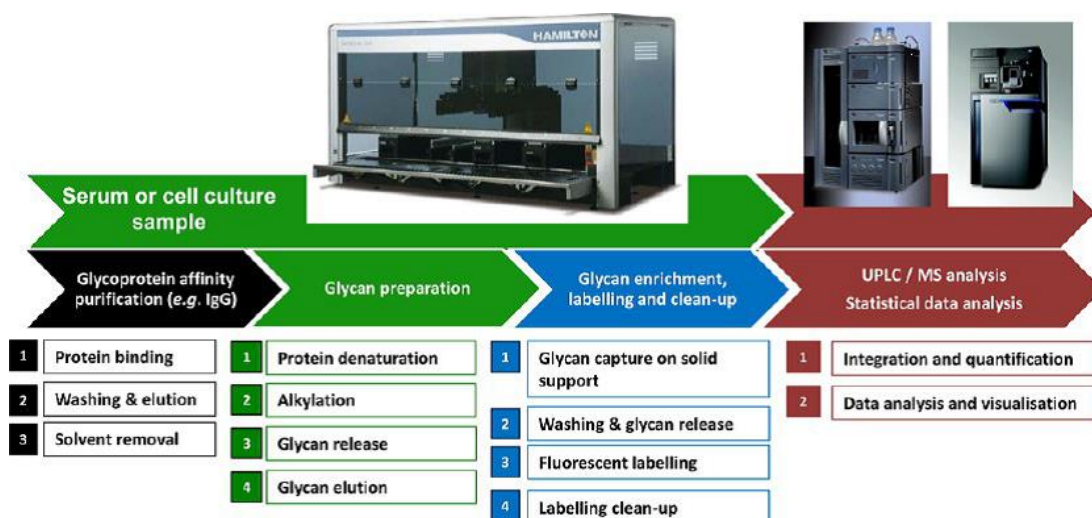


Figure 5.1: Robotic automated sample preparation workflow for UPLC-based glycan analysis on a liquid handling platform. The sample preparation workflow consists of protein denaturation, enzymatic glycan release, glycan immobilization on solid supports, removal of contaminants, glycan release, labelling, and solid-phase extraction (Stöckmann *et al.*, 2013).

HILIC glycan analysis is continuously advancing and has many advantages over other methods such as high-performance anion-exchange chromatography (HPAEC) reversed-phase liquid chromatography (RP-LC) and capillary electrophoresis (CE) (Saldova *et al.*, 2014b).

HILIC separates charged glycan and neutral glycans within a single separation enabling total glycan characterisation in single-column chemistry. In HILIC, separation is based on the hydrophilicity of the glycan, which is affected by its, composition, size, charge structure, linkage, and branching. The separation of structural isomers is difficult to achieve with other methodologies, which is one of main advantages of HILIC (Mariño *et al.*, 2010; Saldova *et al.*, 2014). Saldova and co-workers showed that using HILIC-UPLC greatly improved separation of human serum *N*-glycans as compared with HPLC (Saldova *et al.*, 2014b).

5.2 Results

5.2.1 PCRC patient cohort

A total of 117 patients (indolent [41], significant [32] and aggressive [44]) were included in this study (Figure 5.2). Epstein's criteria was used to define patients as having indolent, significant or aggressive PCa (Epstein *et al.*, 1994).

	Indolent	Significant	Aggressive
Number of Patients	41	32	44
Percentage	35.04%	27.35%	37.61%

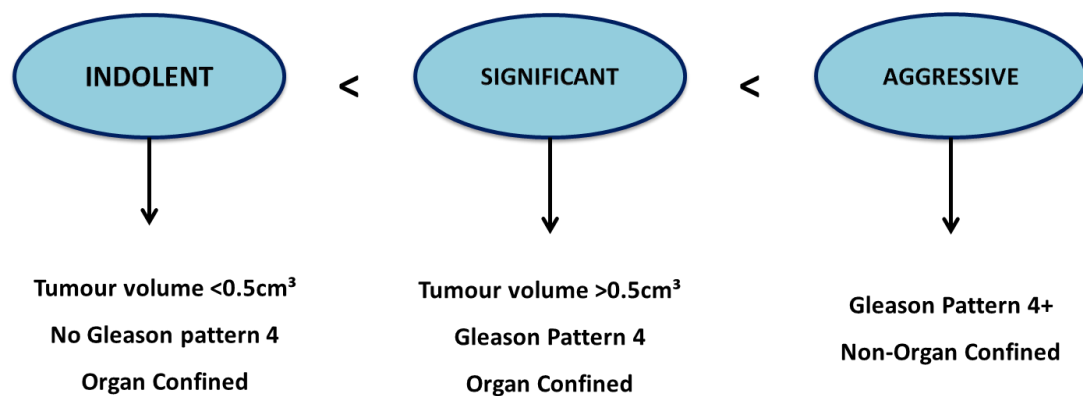


Figure 5.2: Prostate cancer patients sera sourced from the Prostate Cancer Research Consortium (PCRC) BioResource. Indolent PCa was defined as tumour volume $<0.5\text{cm}^3$, organ-confined disease and no Gleason patterns 4 or 5. Significant disease was defined as tumour volume $>0.5\text{cm}^3$, organ-confined disease and Gleason pattern 4. Aggressive PCa was defined as Gleason patterns 4 or 5 and non-organ-confined disease (Patient clinical data available in Appendix Five).

5.2.2 UPLC-HILIC profiling of PCa patients serum

A high-throughput platform, automated in a robotic 96-well plate format, for *N*-glycan analysis of human serum using ultra-performance liquid chromatography (UPLC)–HILIC with fluorescence detection was recently described (Saldova *et al.*, 2014). This high resolution UPLC–HILIC method allows potential biomarkers to be separated without the need for exoglycosidase digestion or pre-separation on WAX-HPLC (Saldova *et al.*, 2014).

The serum *N*-glycome of all 117 PCa patients was analysed in duplicate. The profiles were separated into 50 peaks and each peak was examined individually to see if there was any significantly different expression between indolent, significant and aggressive PCa patients (Figure 5.3). The amounts of total core-, outer-arm fucose, and high-mannosylated glycans were calculated based on composition of these 50 peaks, and the glycans were then summarized into five groups according to sialylation (S0–S4), five groups according to galactosylation (G0–G4), and four groups according to branching (A1–A4) (See glycan features in Table 5.1).

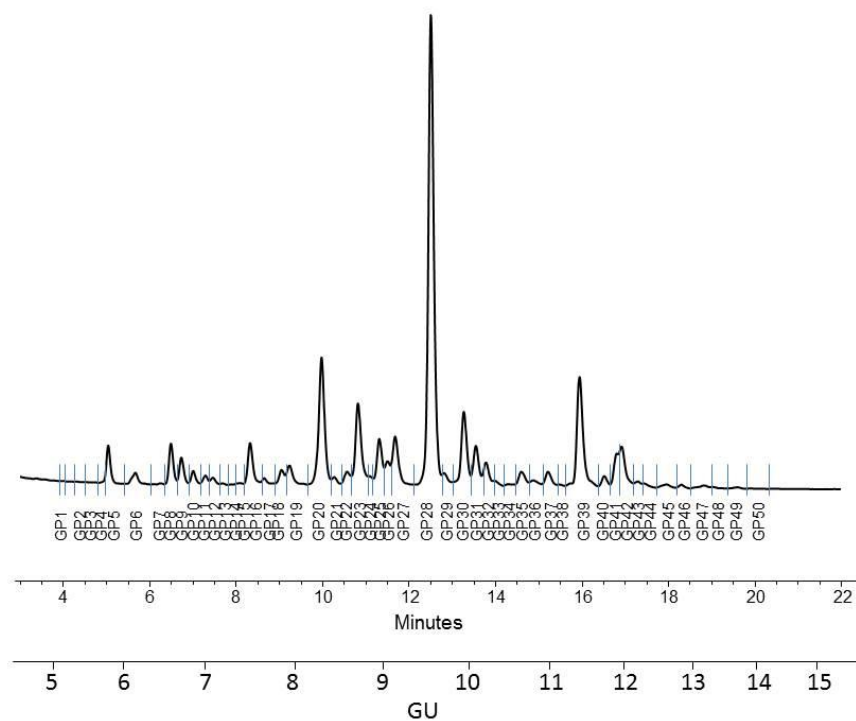


Figure 5.3: Typical HILIC-UPLC profile of undigested serum N-glycome from prostate cancer patients. Profiles were standardised against a dextran hydrolysate (GU). The HILIC-chromatogram was separated into 50 peaks. Structural assignments are listed in Table 5.1.

Table 5.1: Summary of identified *N*-Glycans from prostate cancer sera ^{a,b,c}

<i>N</i> -glycans Total		Features																
Glycan label	GU Glycans	S0	S1	S2	S3	S4	G0	G1	G2	G3	G4	A1	A2	A3	A4	high M	CoreF	Outer arm F
GP1	4.85 A1																	
	5.34 M4																	
GP2	5.41 FA1																	
	5.41 A2																	
GP3	5.61 A1[6]G1																	
GP4	5.78 A2B																	
	5.78 A1[3]G1																	
GP5	5.88 FA2																	
GP6	6.18 M5																	
	6.24 FA2B																	
	6.24 FA1[6]G1						50%						50%			50%	50%	
	6.24 A2[6]G1																	
	6.38 FA1[3]G1																	
	6.38 A2[3]G1																	
GP7	6.55 A2[6]BG1																	
GP8	6.71 A2[3]BG1																	
	6.71 FA2[6]G1																	
GP9	6.71 M4A1G1																	
	6.84 FA2[3]G1																	
	6.95 FA2[6]BG1																	
	7.08 FA2[3]BG1																	
GP10	7.08 M6 D1,D2																	
	7.08 M6 D3																	
GP11	7.20 A1[3]G1S[3]1	50%	50%				50%	50%				50%	50%					
	7.20 A2G2																	
GP12	7.38 A1[3]G1S[6]1																	
GP13	7.38 A2BG2																	
GP14	7.62 A2[3]G1S[3]1																	
	7.62 A2[3]G1S[6]1																	
	7.62 FA1G1S[3]1											50%	50%				50%	
	7.62 FA1G1S[6]1																	
GP15	7.62 M5A1G1																	
	7.62 FA2G2																	
GP16	7.76 FA2BG2																	
	7.76 M7 D3																	
	7.76 A2[6]BG1S[3]1																	
	7.76 A2[6]BG1S[6]1																	
GP17	7.92 A2[3]BG1S[3]1																	
	7.92 A2[3]BG1S[6]1																	
	7.92 FA2[6]G1S[3]1																	
	7.92 FA2[6]G1S[6]1																	
	7.92 M4A1G1S[3]1																	
	7.92 M4A1G1S[6]1																	
GP18	7.92 M7 D1																	
	8.03 FA2[3]G1S[3]1																	
	8.03 FA2[3]G1S[6]1																	
	8.03 FA2[6]BG1S[3]1																	
	8.03 FA2[6]BG1S[6]1																	
	8.20 FA2[3]BG1S[3]1																	
GP19	8.20 FA2[3]BG1S[6]1																	
	8.38 A2G2S[3]1																	
	8.38 A2G2S[6]1																	
	8.38 A3G3																	
GP20	8.53 A2BG2S[3]1																	
	8.53 A2BG2S[6]1																	
	8.63 M5A1G1S[3]1																	
	8.63 M5A1G1S[6]1																	
GP21	8.63 FA3G3																	
	8.63 M8 D2,D3																	
	8.63 A2G2S[3,3]2																	
	8.63 A2G2S[3,6]2																	
GP22	8.80 FA2G2S[3]1																	
	8.80 FA2G2S[6]1																	
GP23	8.80 M8 D1,D3																	
GP24	9.02 FA2BG2S[3]1																	
	9.02 FA2BG2S[6]1																	
GP25	9.21 A2G2S[3,3]2																	
	9.21 A2G2S[3,6]2																	
GP26	9.21 A2F1G2S[3]1																	
	9.21 A2F1G2S[6]1																	
GP27	9.43 A3G3S[3]1																	
	9.43 A3G3S[6]1																	
	9.43 M9																	
	9.62 A3BG3S[3]1																	
	9.62 A3BG3S[6]1																	
	9.62 A2G2S[3,3]2																	
GP28	9.62 A2G2S[3,6]2																	
	9.62 A2G2S[6,6]2																	
	9.79 FA3G3S[3]1																	
	9.79 FA3G3S[6]1																	
	9.79 FA3BG3S[3]1																	
	9.79 A2BG2S[3,3]2																	
GP29	9.79 A2BG2S[3,6]2																	
	9.79 A2BG2S[6,6]2																	

N-glycans Total			Features																
Glycan label	GU	Glycans	S0	S1	S2	S3	S4	G0	G1	G2	G3	G4	A1	A2	A3	A4	high M	CoreF	Outer arm F
GP30	10.04	A3F1G3S[3,1]																	
	10.04	FA2G2S[3,3]2																	
	10.04	FA2G2S[3,6]2																	
	10.04	FA2G2S[6,6]2																	
GP31	10.17	FA2BG2S[3,3]2																	
	10.17	FA2BG2S[3,6]2																	
	10.17	FA2BG2S[6,6]2																	
	10.17	A2F1G2S[3,3]2																	
	10.17	A2F1G2S[3,6]2																	
	10.17	A2F1G2S[6,6]2																	
	10.17	M9Glc																	
GP32	10.31	A3G3S[3,3]2																	
	10.31	A3G3S[3,6]2																	
	10.31	A3G3S[6,6]2																	
GP33	10.43	A3BG3S[3,3]2																	
	10.43	A3BG3S[3,6]2																	
	10.43	A3BG3S[6,6]2																	
GP34	10.60	A4G4S[3,1]																	
GP35	10.77	FA3G3S[3,3]2																	
	10.77	FA3G3S[3,6]2																	
	10.77	FA3G3S[6,6]2																	
	10.77	*A3G3S[3,3]2																	
GP36	10.96	*A3G3S[3,3,3]3																	
	10.96	A3F1G3S[3,3]2																	
	10.96	A4G4S[6,1]																	
GP37	11.14	*A3G3S[3,3,6]3																	
	11.14	*A3G3S[3,6,6]3																	
	11.14	*A3BG3S[3,3,3]3																	
	11.14	*A3BG3S[3,3,6]3																	
	11.14	*A3BG3S[3,6,6]3																	
GP38	11.28	*FA3G3S[3,3,3]3																	
	11.28	*FA3BG3S[3,3,3]3																	
GP39	11.54	A4G4S[3,6]2																	
	11.54	A3G3S[3,3,3]3																	
	11.54	A3G3S[3,3,6]3																	
	11.54	A3G3S[3,6,6]3																	
	11.54	A3G3S[6,6,6]3																	
	11.72	A3BG3S[3,3,3]3																	
	11.72	A3BG3S[3,3,6]3																	
	11.72	A3BG3S[6,6,6]3																	
GP40	11.89	FA3G3S[3,3,3]3																	
	11.89	FA3G3S[3,3,6]3																	
	11.89	FA3G3S[3,6,6]3																	
	11.89	FA3G3S[6,6,6]3																	
GP41	12.03	*A3G3S[3,3,3]3																	
	12.03	*A3G3S[3,3,6]3																	
	12.03	*A3G3S[3,6,6]3																	
	12.03	A3F1G3S[3,3,3]3																	
	12.03	A3F1G3S[3,3,6]3																	
	12.03	FA3BG3S[3,3,3]3																	
GP42	12.15	A4G4S[3,3,3]3																	
	12.15	A3F1G3S[3,6,6]3																	
GP43	12.33	A4G4S[3,3,6]3																	
	12.33	A4G4S[3,6,6]3																	
GP44	12.48	A4F1G3S[3,3,3]3																	
	12.48	A3F2G3S[3,3,3]3																	
	12.48	A4F1G3S[3,3,6]3																	
	12.48	A4F1G3S[3,6,6]3																	
GP45	12.67	A3F2G3S[3,3,6]3																	
	12.67	A4F2G3S[3,3,3]3																	
	12.67	A4F2G3S[3,3,6]3																	
	12.67	*A4G4S[3,3,3,3]4																	
	12.78	A4G4S[3,3,3,3]4																	
GP46	12.96	A4G4S[3,3,3,6]4																	
GP47	13.27	*A4G4S[3,3,3,6]4																	
	13.27	A4G4S[3,3,6,6]4																	
	13.27	A4G4S[3,6,6,6]4																	
GP48	13.47	*A4G4S[3,3,3,3]4																	
	13.47	FA4G4S[3,3,3,3]4																	
	13.47	FA4G4S[3,3,3,6]4																	
	13.47	A4BG4S[3,3,6,6]4																	
GP49	13.82	A4F1G4S[3,3,3,3]4																	
	13.82	A4F1G4S[3,3,3,6]4																	
	13.82	A4F1G4S[3,3,6,6]4																	
	13.82	A4F1G4S[3,6,6,6]4																	
GP50	13.99	A4G4LacS[3,3,3,3]4																	
	13.99	A4G4LacS[3,3,3,6]4																	
	13.99	A4F2G4S[3,3,3,3]4																	
	13.99	A4F2G4S[3,3,6,6]4																	
	14.43	A4F3G4S[3,3,3,3]4																	

^a The HILIC-chromatogram was separated into 50 peaks and structural assignment carried out according to Saldova *et al.*, 2014. ^b *Sialic acids isomers (same composition but different sialic acid linkage arrangements resulting in different GUs from the original structures). ^c Peaks calculated into specific features are highlighted in grey (where there is 33 or 50%; the glycans with the given feature are approximately that abundant in the given peak): Sialylation: S0 (neutral, GP1–11+(GP12/2)+14+16+17+24); S1 (monosialylated, (GP12/2)+GP13+GP15+GP18–23+GP25+GP27+GP34); S2 (disialylated, GP26+28–33+GP35) ;S3 (trisialylated, GP36–44); S4 (tetrasialylated, GP45–50); Galactosylation: G0 (agalactosylated, GP1–2+GP4–5+(GP6/2)+(GP12/2)); G1 (monogalactosylated, GP3+GP7–10+(GP12/2)+GP13+GP15+GP18–19+(GP22/2)); G2 (digalactosylated, GP14+GP16–17+GP20–21+(GP22/2)+GP23+GP25–31); G3 (trigalactosylated, GP32–33+GP35–41); G4 (tetragalactosylated, GP34+GP42–50); Branching: A1 (monoantennary, GP1–3+ (GP12/2) +13+ (GP15/2) + (GP22/2)); A2 (biantennary, GP4–5+(GP6/2)+GP7–10+(GP12/2)+GP14+(GP15/2)+GP16–21+(GP22/2)+GP23+GP25–31); A3 (triantennary, GP32–33+GP35–41); A4 (tetraantennary, GP34+GP42–50); High mannose: (GP6/2) +GP11 ; Fucosylation: Core-fucose: GP2+GP5+ (GP6/2) +GP8–10+ (GP15/2) +GP16–17+GP19+GP23+GP25+GP30–31+GP40+ (GP48/2) ; Outer-arm fucose: GP41+GP44+(GP45/3)+GP49+(GP50/3).Structure abbreviations: all *N*-glycans have two core GlcNAcs; F at the start of the abbreviation indicates a core-fucose α 1,6-linked to the inner GlcNAc; Mx, number (x) of mannose on core GlcNAcs;D1 indicates that the α 1-2 mannose is on the Man α 1-6Man α 1-6 arm, D2 on the Man α 1-3Man α 1-6 arm, D3 on the Man α 1-3 arm of M6 and on the Man α 1-2Man α 1-3 arm of M7 and M8; Ax, number of antenna (GlcNAc) on trimannosyl core; A2, biantennary with both GlcNAcs as β 1,2-linked; A3, triantennary with a GlcNAc linked β 1,2 to both mannose and the third GlcNAc linked β 1,4 to the α 1,3 linked mannose; A4, GlcNAcs linked as A3 with additional GlcNAc β 1,6 linked to α 1,6 mannose; B, bisecting GlcNAc linked β 1,4 to β 1,3 mannose; Gx, number (x) of β 1,4 linked galactose on antenna; F(x), number (x) of fucose linked α 1,3 to antenna GlcNAc; Sx, number (x) of sialic acids linked to galactose; Lac(x), number (x) of lactosamine (Gal β 1-4GlcNAc) extensions.

5.2.2 Statistical analysis

Statistical analysis was carried out by Mr. Keefe Murphy (PCRC biostatistician, University College Dublin). Initially, all 50 peaks were examined individually to explore baseline differences in expression using one way analysis of variance (ANOVA). Tukey's post-hoc honestly significant difference (HSD) was performed to calculate the p-values and boxplots were used to illustrate the distribution of each biomarker by significance. From this analysis a panel of 13 peaks were identified as significant, as follows;

Peak 4 + Peak13 + Peak 14 + Peak15 + Peak17 + Peak 18 + Peak 22 + Peak 27 + Peak 33 + Peak
34 + Peak 37 + Peak38 + Peak40

All 50 peaks were then considered in combination (in a multivariate predictive regression context) to see which panel of variables would yield the best predictive performance, the following panel was identified:

Peak15 + Peak22 + Peak1 + Peak29 + Peak44 + Peak11 +Peak21 + Peak7 + Peak45 + Peak17 +
Peak37 + Peak10 + Peak30

5.2.3 Significant changes in whole serum *N*-glycome in prostate cancer

When considered individually, 13 peaks showed statistically significantly different expression levels across the 3 patient cohorts. Of these 13 peaks, 4 peaks showed statistically significantly different expression levels as disease progresses from indolent to aggressive (Figure 5.4) and will be further discussed.

Peak 4, containing mostly biantennary bisected glycans (A2B) and peak 22 containing mostly biantennary digalactosylated monosialylated glycans (A2G2S1) and high mannose structures (M5A1G1S1, M8) were significantly increased as disease progresses in this study.

Notable decreases were also observed, particularly in triantennary trigalactosylated trisialylated glycans with and without core-fucosylation (Peak 37- A3G3S3, Peak 40-FA3G3S3), as PCa progresses from indolent to aggressive.

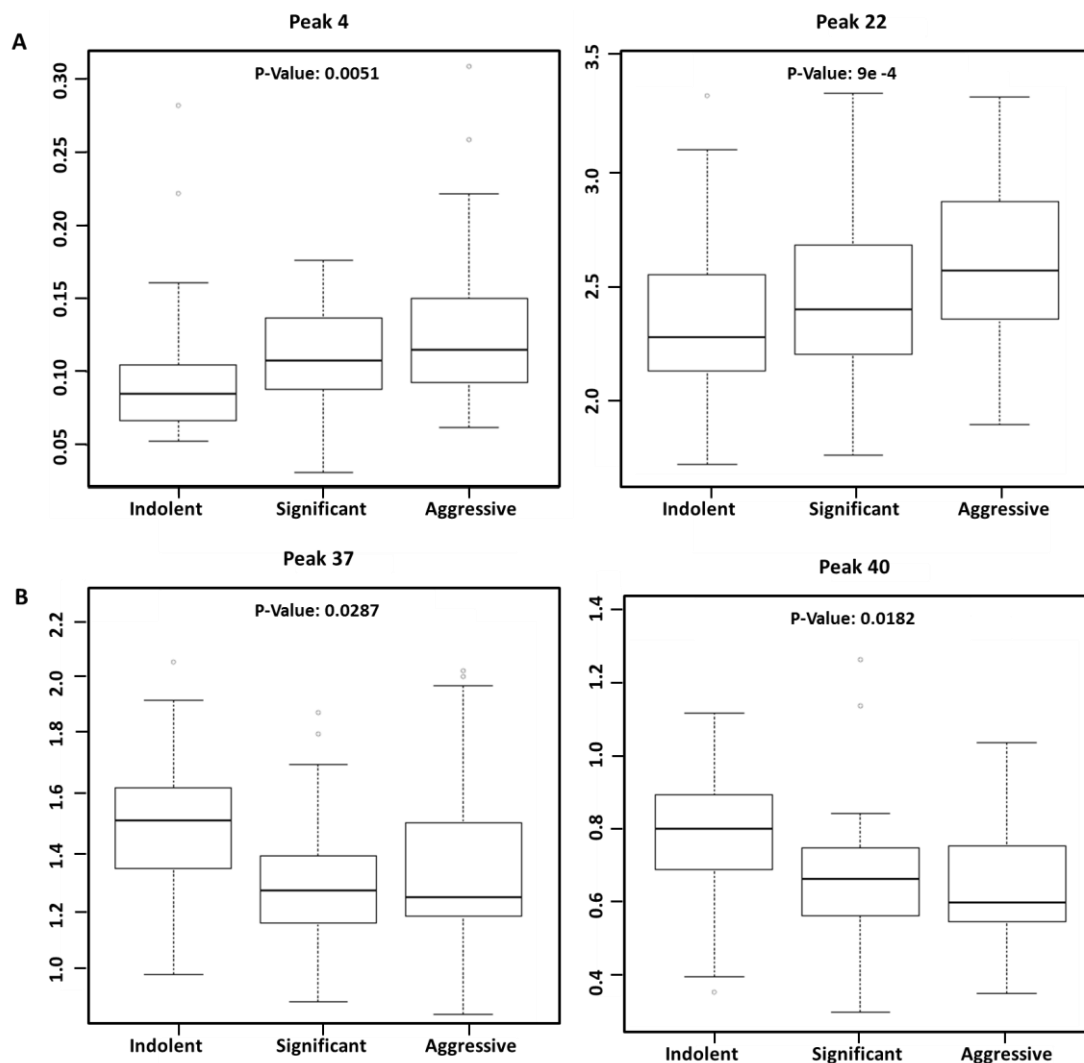


Figure 5.4: Boxplots of 4 glycosylation peaks which exhibited significant differences across 3 patient groups (indolent, significant and aggressive) from the Irish PCRC cohort. A box plot is a simple way of representing statistical data based on the minimum, first quartile, median, third quartile, and maximum. Here, boxes represent the 25th and 75th percentiles with the median indicated. The lower and upper quartiles are shown as horizontal lines either side of the rectangle. (A) Peaks that significantly increase as disease progresses and (B) peaks that significantly decrease as disease progresses.

Model diagnostics was then carried out to assess the predictive performance of these 4 peaks.

The discriminant ability was determined by means of ROC curves and AUC values.

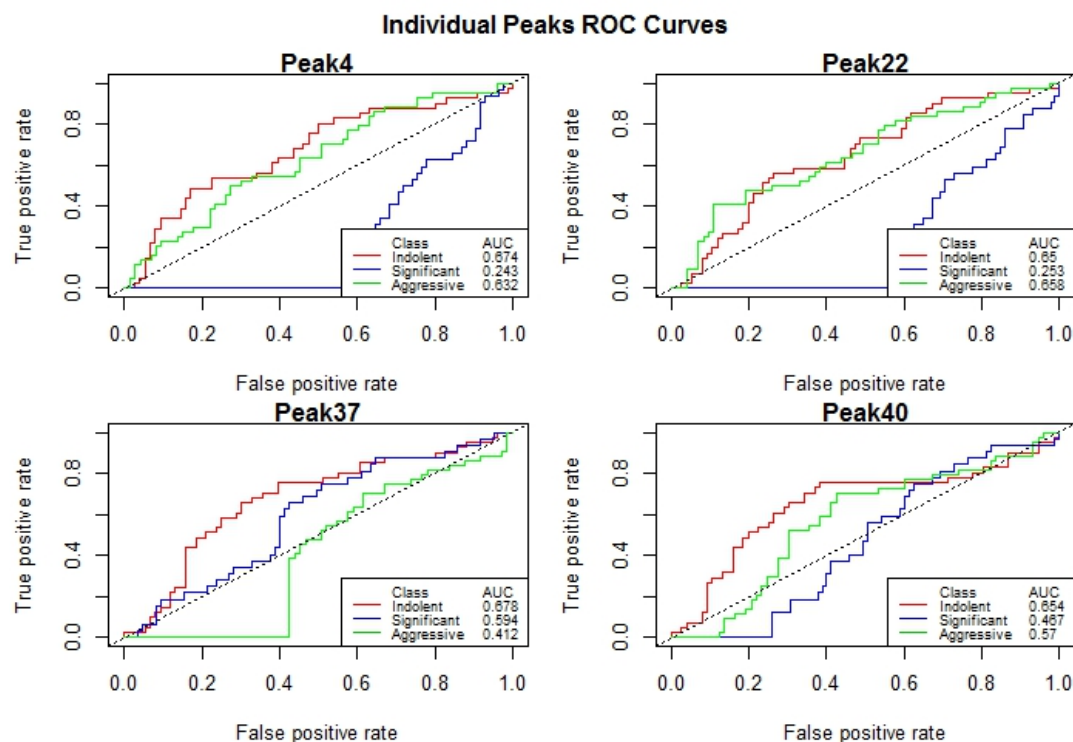


Figure 5.5: ROC curves of the 4 significantly changed peaks as prostate cancer progresses. A ROC curve is a fundamental tool for diagnostic test evaluation where the true positive rate (Sensitivity) is plotted in function of the false positive rate (100-Specificity) for different cut-off points of a parameter. The AUC value indicates how each peak can predict indolent, significant or aggressive PCa. A perfect diagnostic test has an AUC 1.0 whereas a non-discriminating test has an area 0.5. The AUC values for all peaks very close to or less than 0.5, which indicates the glycans in these peaks, alone, have poor discriminatory ability.

5.2.4 *N*-glycome biomarker panel

Within this combination of glycans there are notable increases and decreases in certain glycan structures as disease progresses from the indolent cohort to the more advanced cohort as demonstrated in Figure 5.6.

There is an increase biantennary bisected monogalactosylated glycans (Peak 7-A2[6]BG1) and biantennary digalactosylated monosialylated and high mannose (Peak 22 – mostly A2G2S1, M5A1G1S1) glycans as disease progresses. We also observed increases in core fucosylated biantennary disialylated glycans (Peak 30 – mostly FA2G2S2) and tetraantennary tetrasialylated glycans in peak 45 (predominately A4G4S4).

With disease progression we also identified decreases in bisected biantennary disialylated glycans (mostly A2BG2S2 - peak 29) and certain high mannose glycans (mostly M6 isomers – Peak 11). Peak 37, which is comprised mostly of triantennary trigalactosylated trisialylated glycans (A3G3S3) was also included in this biomarker panel and was significantly decreased with disease progression.

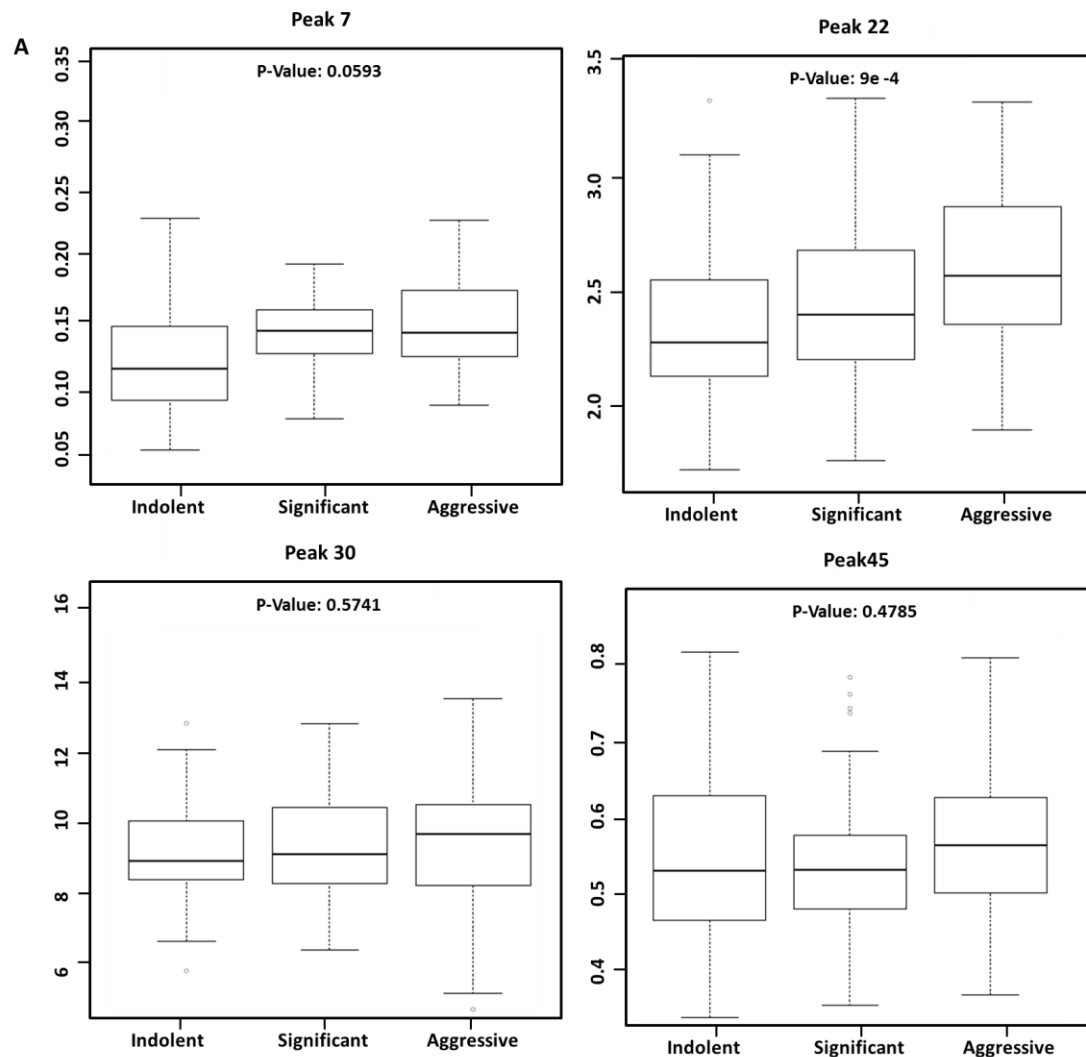


Figure 5.6: Plotted peaks form part of the PCa whole serum N-glycome biomarker panel that show altered expression in indolent, significant and aggressive PCa from the Irish PCRC cohort. Figure

Legend on Page 230

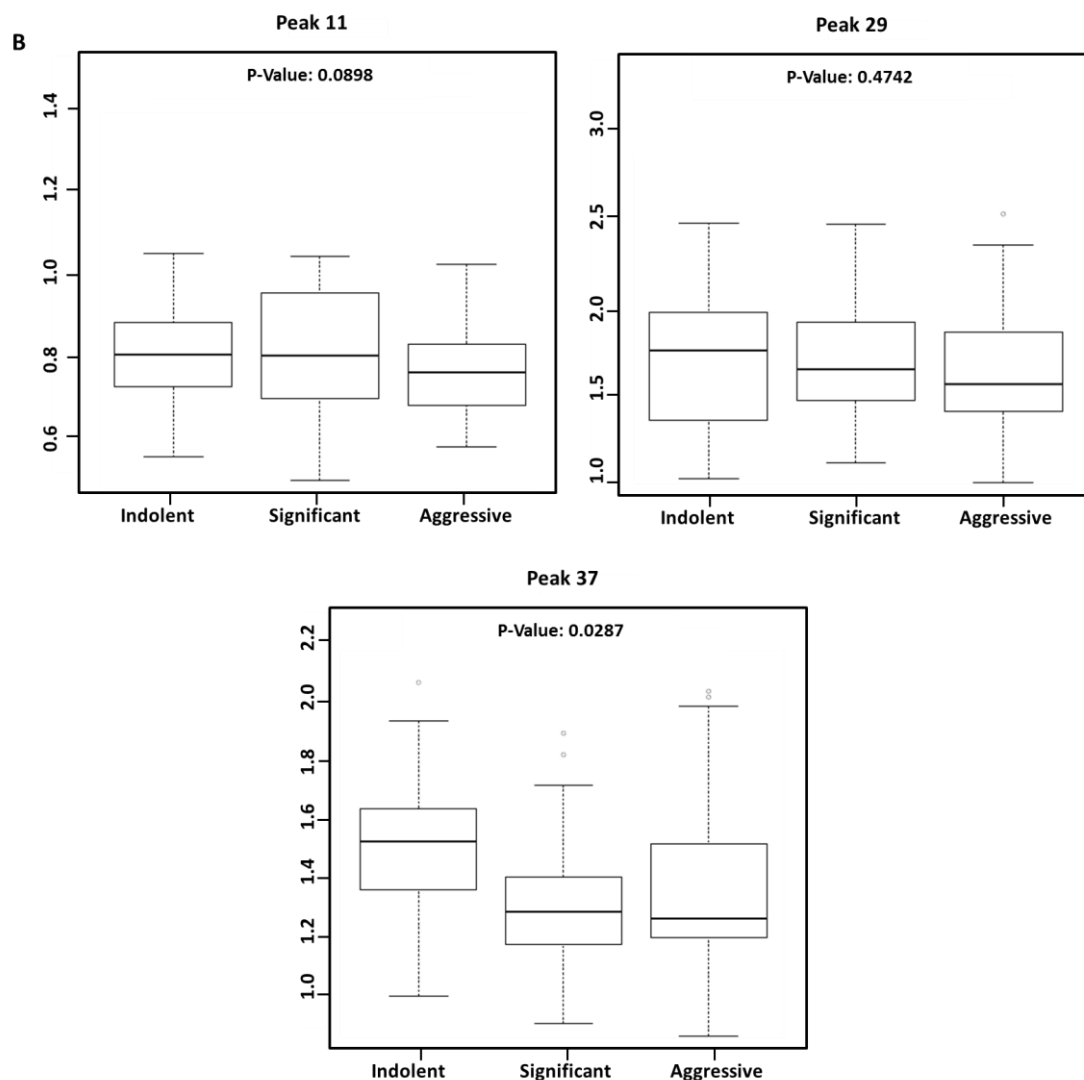


Figure 5.6: *Plotted peaks form part of the PCa whole serum N-glycome biomarker panel that show altered expression in indolent, significant and aggressive PCa from the Irish PCRC cohort. A box plot is a simple way of representing statistical data based on the minimum, first quartile, median, third quartile, and maximum. Here, boxes represent the 25th and 75th percentiles with the median indicated. The lower and upper quartiles are shown as horizontal lines either side of the rectangle (A) Four peaks that increase as disease progresses and (B) three peaks that decrease as disease progresses.*

As for the significant biomarker panel, model diagnostics was then carried out to assess the performance of this biomarker panel (Peak15 + Peak22 + Peak1 + Peak29 + Peak44 + Peak11 + Peak21 + Peak7 + Peak45 + Peak17 + Peak37 + Peak10 + Peak30). The discriminant ability was determined by means of ROC curves and AUC values (Figure 5.7). To further assess the net benefit of this prediction model calibration curves (Figure 5.8) were prepared.

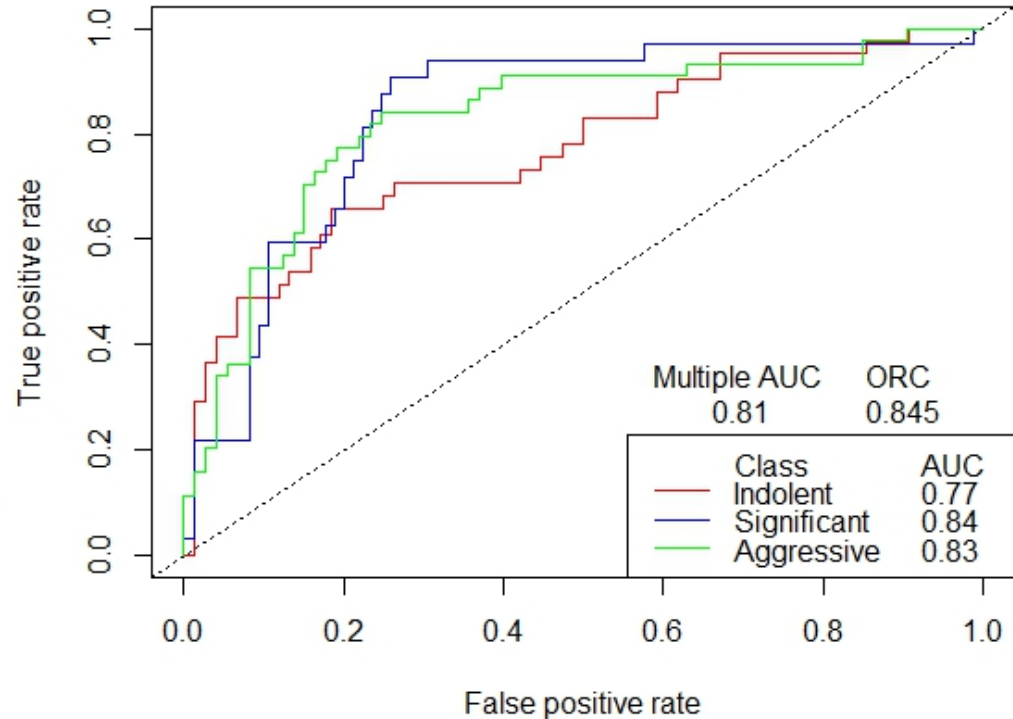


Figure 5.7: ROC curves are plotted with AUC values of the peaks in the biomarker panel in indolent, significant and aggressive PCa. A ROC curve is a fundamental tool for diagnostic test evaluation where the true positive rate (Sensitivity) is plotted in function of the false positive rate (100-Specificity) for different cut-off points of a parameter. The AUC value indicates how each peak can predict indolent, significant or aggressive PCa. A perfect diagnostic test has an AUC 1.0 whereas a non-discriminating test has an area 0.5. The glycans present in *N*-glycome biomarker panel can predict PCa cohorts with high sensitivity and specificity (multiple AUC 0.81). This biomarker has high discriminant ability and can predict indolent cancer to 77%, significant to 84% and aggressive disease to 83%.

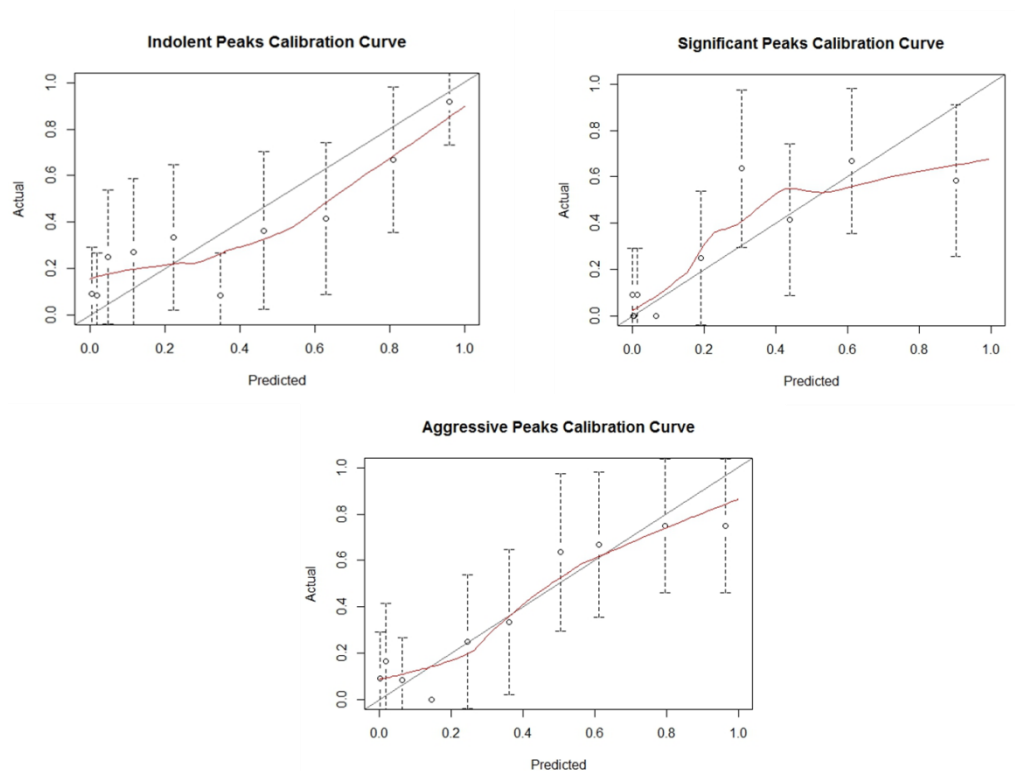


Figure 5.8: Calibration curves of the N-glycome biomarker panel in indolent, significant and aggressive PCa. A calibration curve is a measure of the agreement between observed outcomes and predictions. The resulting curve shows whether the model underestimates or overestimates PCa classification. A perfect prediction should be on the 45° line (Steyerberg *et al.*, 2010). The glycosylation biomarker panel is close to the 45° line for the indolent and significant and is almost equal to the 45° line for aggressive cohort. This result illustrates that the N-glycan biomarker panel is not grossly misclassifying PCa patients i.e. there is strong agreement between the observed (PCRC classified) and predicted (by N-glycome biomarker) results.

5.3 Discussion

Recently, the over-diagnosis of PCa has become one of the major clinical dilemmas faced by urologists. PCa is slow growing in most men, presenting no clinical symptoms during the life-time of the patient, and in many cases, men die with the disease rather than from the disease (Vickers *et al.*, 2012).

For this reason, urologists have had to identify new treatment strategies for men diagnosed with low risk 'indolent' PCa, namely, active surveillance and watchful waiting. Prostate cancer-specific mortality under active surveillance is very low and trials of active surveillance report encouraging results (Venderbos *et al.*, 2013).

The Epstein criteria can be used to determine PCa significance, and thereby identify indolent PCa's which may be eligible for active surveillance. However, the Epstein criteria depends largely on Gleason score, and changes to the Gleason scoring system (Epstein, 2010) have resulted in a major shift in Gleason classification standards and have raised questions as to what Gleason score now constitutes indolent disease.

It is increasingly obvious that there is an unmet need to develop a biomarker that can differentiate between indolent and aggressive prostate cancer, ultimately, maximising the benefits of cancer screening while limiting harm.

In this chapter, whole serum *N*-glycan profiling was carried out on 117 prostate cancer patients' serum using a highly efficient robotised, high-throughput analysis platform for glyco-profiling with UPLC. The aim of this study was to investigate if differential glycosylation patterns could distinguish between the 3 cohorts of PCa patients.

The current clinical model used for determining PCa classification in Ireland focus on 4 variables; serum PSA (ng/mL), clinical stage (DRE), age, biopsy Gleason Score and has limited sensitivity, specificity, and predictability as detailed in Table 5.2.

Table 5.2: Current clinical model predictability

Type	Specificity	Sensitivity	AUCs	95% C.I.	Multi AUC	ORC
Indolent	0.929	0.548	0.776	{0.703, 0.849}		
Significant	0.667	0.452	0.471	{0.372, 0.570}	0.668 {0.618,	0.724 {0.667,
Aggressive	0.788	0.623	0.779	{0.702, 0.855}	0.706}	0.781}

Initially, statistical analysis was performed to determine if significant peaks could differentiate between indolent, significant and aggressive PCa. From these results, 2 peaks significantly increased and 2 peaks significantly decreased as disease progresses. Model diagnostics was carried out for these 4 peaks to determine their predictive performance (Figure 5.5).

A receiver operating characteristic (ROC) curve was generated for each peak (by Mr. Keefe Murphy), which plots the sensitivity (true positive rate) against 1 – (false positive rate) for a range of values for the ability to predict an outcome (Steyerberg *et al.*, 2010). Overall accuracy is sometimes expressed as area under the ROC curve (AUC). The closer the curve is located to upper-left hand corner and the larger the area under the curve, the better the test will be at discriminating between the 3 cohorts (Eusebi, 2013). In Figure 5.5, the AUC values are all very

close to or less than 0.5, indicating the glycans in these peaks, alone, have poor discriminatory ability. A perfect diagnostic test has an AUC 1.0 whereas a non-discriminating test has an area 0.5 (Eusebi, 2013). The ability of these peaks alone to predict between indolent, significant and aggressive PCa, is worse than the current clinical model (Table 5.2).

Given the poor predictability of these peaks alone for discriminating between indolent, significant and aggressive PCa, further statistical analysis was carried out where the peaks were then considered in combination (in a multivariate predictive regression context) to see which panel of variables would yield the best predictive performance.

From this biomarker panel there are noticed increases and decreases in expression patterns of glycosylation in several peaks (Figure 5.6), of which 2 peaks are statistically significantly altered (Peak 22 and Peak 37). The glycans present in these peaks correlate with the most common serum glycome alterations observed in prostate cancer, and are well documented, including sialylation, branching and fucosylation (Salдова *et al.*, 2011).

Increases in core-fucosylation

The ageing process is known to effect serum *N*-glycans and differential patterns between age groups were detected in various studies (Marino *et al.*, 2012), notably, core-fucosylation is known to decrease with age (Knezevic *et al.*, 2009). In this study patient groups were matched for age when they were selected for inclusion in the cohort so there should be no difference in age between the groups.

The observed increase in core-fucosylation (FA2G2S2, Peak 30) in significant and aggressive patients compared to indolent patients is unsurprising as altered fucosylation is now considered a common feature of many cancers (Marino *et al.*, 2012) and has been previously

observed in prostate cancer studies (Kyselova *et al.*, 2007, Saldova *et al.*, 2011). In comparing 24 PCa patients with 10 healthy controls Kyselova *et al.* (2007) noted that the fucosylation of glycan structures was significantly higher in the serum of the PCa patients. Changes in fucosylation pathways could affect the level of core fucosylation. An increase in fucosyltransferases increases the core fucosylation of Epidermal Growth Factor Receptor (EGFR), which is commonly overexpressed in many carcinomas (Masumoto *et al.*, 2008). In an earlier study (Saldova *et al.*, 2011), an increase in core-fucosylation in PCa sera in a study of 13 patients with Benign Prostate Hyperplasia (BPH) and 34 PCa patients was shown. In this study it can be seen that increases in core-fucosylated glycans in this biomarker panel helps distinguish indolent PCa from clinically insignificant and aggressive PCa.

Alterations in high mannose glycans

Complex high mannose structures (M5A1G1S[3]1, M5A1G1S[6]1 and M8 D2,D3; Peak 22) increase as disease progress. *N*-glycans with high mannose, hybrid, or complex type sugar chains contribute in a multitude of cellular processes including cell-cell/cell-matrix/receptor-ligand interaction, cell signalling/growth and differentiation (Banerjee, 2012). Given that high mannose structures participate both in cell survival and cell death this could be an indication for the observed increased abundance of high mannose structures in significant and aggressive disease compared to indolent disease.

An abundance of high mannose structures were observed when the *N*-glycan profiles of membrane proteins across 3 different colorectal cancer cell lines (LIM1215, LIM1899, and LIM2405) was investigated (Sethi *et al.*, 2014). The authors of this study suggest incomplete *N*-glycan processing led to the accumulation of high mannose type structures in the colorectal cancer cell lines. In another recent colorectal cancer study the authors showed high-mannose

N-glycan structures were more common in carcinomas than in adenomas (Kaprio *et al.*, 2015). We did however observe a decrease in isomers of the high mannose glycan-M6 (predominant glycans in peak 11); therefore it is likely that only certain high mannose structures are increased in PCa patients.

A3G3S3/A4G4S4

Alterations in certain glycan structures were noted in a previous prostate cancer study that correlate with the glycans in this biomarker panel (Saldova *et al.*, 2011), in which significant decreases in triantennary trigalactosylated trisialylated glycans (A3G3S3) and increases in tetraantennary tetragalactosylated tetrasialylated (A4G4S4) are observed in advanced aggressive PCa.

Certain glycan structures are well-known markers for tumour progression and through decades of glycomic studies one of the more typical features of disease progression that is consistently observed is increased expression of sialic acid sugars on the surface of cancer cells (Büll *et al.*, 2014). While we can see an increase in tetraantennary tetrasialylated glycans (A4G4S4, peak 45) in aggressive PCa we also observed a significant decrease in triantennary trisialylated glycans (A3G3S3, peak 37).

Patients with a Gleason score of 7 are generally positive for perineural invasion (PNI), found to be at a more advanced stage of metastasis and have an increased chance of re-occurrence. PNI is an important pre-operative indicator of the pathological stage of the tumor. In our initial study we showed that increases in tetraantennary tetragalactosylated tetrasialylated (A4G4S4) correlate with PNI and can help diagnose tumor spread (Saldova *et al.*, 2011)

This study supports previous serum PCa glycomics analysis (Saldova *et al.*, 2011) as our aggressive patients are defined as Gleason patterns greater than 4 and non-organ-confined disease. Differentiation between the more challenging significant and aggressive prostate cancer over clinically insignificant disease could potentially be identified by measurement of increases in tetraantennary tetragalactosylated tetrasialylated in patient's sera.

Various studies have noted excessively high expression of sialic acids and aberrant sialylation in cancer. One reason this is thought to occur is that factors such as oncogenes, hormones, and/or chemotherapy can aid in the increased expression of sialyltransferases and downregulate the expression of sialidases in cancerous cells. Ultimately, this leads to accumulation of uncharacteristic hyper-sialylated structures on the cell membrane (Büll *et al.*, 2014). The tumour microenvironment alters key elements of glycan biosynthesis apparatus, such as glycotransferases and glycosidases, and aberrant glycosylation is now widely accepted as a characteristic of cancer. The altered expression of these enzymes results in the production of glycoproteins with modified glycans (Drake *et al.*, 2010). There is now a common acceptance that increases in branching and sialylation correlate with more advanced cancers, which we have discussed and observed in this study.

However, interestingly it was noticed that there was a decrease in the highly branched and sialylated triantennary trigalactosylated trisialylated glycans (A3G3S3) in peak 37. This structure was also decreased in previous studies, where decreases in A3G3S3 could distinguish between patients with Gleason score 5 and Gleason score 7 significantly better than the currently used PSA assay (Saldova *et al.*, 2011). In this body of work, significant decreases in this structure was observed in the significant and aggressive patient groups compared to the

indolent group. It can be hypothesised that decreases in triantennary trigalactosylated trisialylated glycans could have good predictability in monitoring PCa disease progression.

Table 5.3: Sensitivity, specificity and predictability of glycosylation biomarker panel

Type	Specificity	Sensitivity	AUCs	95% C.I.	Multi AUC	ORC
Indolent	0.811	0.628	0.771	{0.678, 0.864}	0.810 {0.759, 0.861}	0.845 {0.793, 0.898}
Significant	0.844	0.667	0.836	{0.753, 0.919}		
Aggressive	0.857	0.723	0.827	{0.745, 0.909}		

Within this panel we have observed increases in certain high mannose structures (M5A1G1S1 and M8) and tetraantennary tetrasialylated glycans (A4G4S4) and significant decreases in triantennary trigalactosylated trisialylated glycans (A3G3S3). These glycans are highly specific in predicting the outcome of PCa patients and can provide a novel approach to diagnosing PCa.

The AUC values for this biomarker panel (Figure 5.7) are considerably better than the current clinical model and further model diagnostics (Figure 5.8) to confirm the predictive ability of this combination of glycans was also carried out (by Mr. Keefe Murphy). Traditionally used performance measures in medicine include, but are not limited to ROC curves, calibration curves and decision curves. A calibration curve refers to the agreement between observed outcomes and predictions (Steyerberg *et al.*, 2010). A graphical assessment of calibration curve for this biomarker panel is shown in Figure 5.8, with predictions on the x-axis, and the outcome on the y-axis. A perfect prediction should be on the 45° line (Steyerberg *et al.*, 2010), and we

can see the glycosylation biomarker panel is quite close to this, particularly for the aggressive cohort.

This biomarker panel can predict between indolent, significant and aggressive PCa with better sensitivity and specificity than the current clinical methods (Tables 5.2 and 5.3).

The glycosylation biomarker identified from this body of work was also included in analysis which showed that a panel of biomarkers attains significantly higher predictive ability than each of the biomarkers alone. DNA methylation, miRNA expression, proteomic analysis and serum *N*-glycosylation data were pooled and variables analysed using stepwise AIC and multi-block aggregation methods (by Mr. Keefe Murphy). This integrated biomarker panel can predict PCa with exceptionally high sensitivity and specificity; indolent disease with 94%, significant (87%) and aggressive PCa with 92% accuracy.

Chapter Conclusions

In conclusion, whole serum *N*-glycan analysis was successfully carried out on 117 prostate cancer patients, in a high-throughput fashion, using a novel robotized platform for rapid serum-based *N*-glycan sample preparation and quantitation. From the high-quality glycomics data generated, a clinically valuable, novel, glyco-biomarker panel was identified that is considerably better than the current clinical model. Significant changes between three cohorts of PCa were observed, notably, alterations in high mannose structures (M5A1G1S1, M6 and M8) and highly branched glycan structures including increases in tetraantennary tetragalactosylated tetrasialylated glycans (A4G4S4) and decreases in triantennary trisialylated glycans (A3G3S3), and also increases in core-fucosylated glycans. These changes are consistent with altered fucosylation and sialylation pathways observed in prostate cancer (Saldova *et al.*,

2011). The glycans identified in this biomarker panel have better predictability than the current clinical model and therefore may provide a novel, non-invasive approach to diagnosing prostate cancer. Development of an effective diagnostic method for prostate cancer is of primordial clinical importance as it can aid the effectiveness of treatment and increase a patient's chance of survival and quality of life.

Chapter Six

Overall Conclusions

6.1 Overall Conclusions

The aims identified at the outset of this project were to develop an in-depth understanding of avian antibodies in order to produce highly sensitive antibodies for improving prostate cancer diagnostics. It is hoped that antibodies generated from this research will provide clinicians with the highly sought after, enhanced tools for detecting the disease and determining the most effective intervention plan. In addition, investigation of alterations to the serum glycome across different forms of prostate cancer led to the identification of a highly specific, novel glyco-marker.

Here, the strategies employed involved the generation and characterisation of recombinant antibody fragments (scFv) to fPSA and cPSA. An anti-cPSA library was generated and antibody fragments isolated and characterised by ELISA, SDS-PAGE and Western Blotting. An anti-fPSA scFv was also isolated from a fPSA immune repertoire and extensively kinetically and structurally evaluated. Both of these antibody fragments are currently being optimised for incorporation into a multiplexed microfluidics-based assay, in our laboratory.

Anti-fPSA scFv B8 was also reformatted to a single chain antibody fragment (scAb) for improved expression (results not shown). Reformatting antibody fragments can produce more stable, rigid protein structures. Hayhurst and colleagues overcame issues of poor soluble expression of a scFv fragment by converting the scFv into a scAb format and introduced co-expression of the periplasmic chaperone *skp*, to alleviate protein aggregation in the bacterial periplasm (Hayhurst *et al.*, 2003). A modified version of the vector developed by Hayhurst and colleagues for the soluble expression of scAb fragments was used to generate an avian anti-fPSA scAb antibody. This vector was re-engineered by Dr. B. Vijayalakshmi Ayyar (designed by

Dr. Stephen Hearty) by replacing the existing HuCk (human constant domain) with a chicken constant domain (C_{λ}). This scAb fragment retained full binding capabilities of scFv B8 (nM affinity for fPSA) and was used for the development of an amperometric magneto-immunosensor.

Optimisation of the physicochemical properties of an antibody without perturbing the affinity and specificity for its cognate antigen is a challenging endeavour. However, progress in the field of structural biology is rapidly accelerating knowledge in this area. In this research, the structural binding attributes of chicken antibodies were determined through X-ray crystallographic techniques. The interaction surface of antibody-antigen complexes can involve a large number of residues and, thus, knowledge of these interactions can provide a wealth of knowledge for structure-guided design. It is hoped that the research described in this thesis will aid scientists in a rational design-based approach for the development of avian antibodies for therapeutic and diagnostics purposes.

In addition to antibody-based methods for improving PCa diagnostics, glycomic-based analysis was also investigated. Recently, rapid screening methods have been used to analyse glycosylation patterns on glycoproteins in large cohorts of patients, enabling the identification of a new generation of disease biomarkers. In a healthy organism, glycosylation systems produce numerous variants of glycans, but the sequences of glycoforms are relatively stable. The glycosylation pathway is, however, sensitive to even subtle changes of the surrounding intracellular environment. This leads to interruptions in glycosyltransferase activities. Hence in pathological conditions, the final structure of synthesised glycans may be significantly altered. Aberrant glycosylation patterns are a fundamental characteristic of tumorigenesis, suggesting

that modified glycoproteins with tumour-specific glycan moieties are viable diagnostic targets for cancer (Ferens-Sieczkowska *et al.*, 2012; Gilgunn *et al.*, 2013).

At the outset of this project, it was hypothesised that altered glycosylation patterns of PSA isoforms will help distinguish between indolent and aggressive forms of prostate cancer (Saldova *et al.*, 2011). It was initially envisaged that antibodies generated from this research would be transferred to a platform for selective isolation of PSA isoforms and subsequent glycoanalysis. Proof of concept studies were carried out with anti-PSA polyclonal antibodies and anti-fPSA scFv B8 (results not shown), however, a minimum of 30 µg of glycoprotein was needed for the associated glycan analysis. Irrespective of the high affinity anti-fPSA scFv B8 had for fPSA, there was not enough PSA protein in the patient's samples to carry out the analysis. Hence, whole serum *N*-glycan profiling was carried out in an attempt to develop a novel glyco-marker. The methods used to generate this data were specifically designed for the discovery of clinical glyco-biomarkers from human serum.

A combination of glycans was identified that can distinguish between indolent, significant and aggressive prostate cancer with better sensitivity and specificity than current clinical models. This combination of glycans, in particular, increase in A4G4S4 and decreases in A3G3S3 are highly specific in predicting the outcome of PCa patients and can provide a novel, non-invasive approach to diagnosing PCa. There are currently no effective systems which have been clinically proven to decipher clinically insignificant prostate cancer and aggressive prostate cancer that can be successfully applied for improved patient welfare. This novel-glyco-biomarker panel should aid in the reduction of 'over-diagnosis' of clinically insignificant PCa and help stratify patients into the most appropriate treatment option for their individual prostatic disease.

In conclusion, this body of research has the potential to provide clinicians with novel tools for improving PCa diagnostics. It is hoped that the findings presented can be translated into new approaches to greatly improve patient outcomes.

Appendices

Appendix One

Glossary of important X-ray crystallographic terms for determining quality of diffracted data

(Acharya and Lloyd, 2005; Wlodawer *et al.*, 2008)

Parameter	Description
B-factor	Also referred to as temperature factor or thermal vibration factor. This is a measure of the thermal mobility, i.e. much an atom oscillates or vibrates around the position specified in the model. The majority of structures are defined using isotropic B-factors (equal in all directions) but can also be anisotropic (not equal in all directions). Diagrams of the model are depicted by colour where red ("hot") indicates high values of B and blue ("cold") indicates low values of B.
Completeness	The number of observed reflections measured in a data set, expressed as a percentage of the total number of reflections present at the specified resolution. Structures approaching 100% are good quality structures.
Ramachandran Plot	This is a plot of the conformational angles around the polypeptide backbone compared with those most favoured for various protein structural elements
Resolution	Resolution is measured in Å and defines the level of detail, or the minimum distance between structural features that can be distinguished in an electron-density map. The mean resolution for X-ray crystallographic results in the Protein Data Bank is 2.21 Å (June, 2015).
R-factor	A global reliability, usually expressed as percent that measures the goodness-of-fit between the refined structure model and the actual X-ray diffraction data. Well defined biological structures are expected to have an R-factor < 20%.
R_{merge}	A statistical measure of how well the X-ray diffracted data scale together. This is often represented for data from the highest-resolution shell, allowing the model user to evaluate the reliability of data at the highest resolution used.
Signal-to-noise ratio	This is the X-ray diffraction signal-to-noise ratio, denoted (I/σ). Generally values of > 2 are required for good-quality structures.

Appendix Two

X-ray crystallography statistics for each mutant fragment and anti-fPSA scFv B8-rPSA complex.

	C97/100S	C97/100A	C97/100A*	Complex
Resolution Range (Å)	45.14 - 1.5 (1.554 - 1.5)	36.77 - 1.7 (1.761 - 1.7)	41.66 - 1.8 (1.864 - 1.8)	38.62 - 1.95 (2.02 - 1.95)
Space Group	P61 2 2	P61 2 2	P 21 21 21	P1 21 1
a, b, c (Å)	52.702, 52.702, 304.053	52.789, 52.789, 309.381	51.994, 66.79, 69.632	107.051, 52.053, 111.943
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 90	90, 116.496, 90
Number of molecules in ASU	1	1	1	4
Unique Reflections	41690 (4040)	26107 (2518)	23098 (2273)	80772 (7788)
Completeness	100.00 (100.00)	88.27 (87.64)	100.00 (100.00)	99.68 (96.76)
I/σI	16.68 (5.15)	12.86 (2.01)	13.70 (1.76)	16.63 (2.87)
Wilson β-Factor	14.11	21.30	22.96	22.35
R_{work}	0.1728 (0.2120)	0.1726 (0.2487)	0.1852 (0.2799)	0.1568 (0.2121)
R_{free}	0.2023 (0.2727)	0.2083 (0.3103)	0.2037 (0.3325)	0.2041 (0.2699)
RMS (Bonds)	0.010	0.014	0.007	0.007
RMS (Angles)	1.30	1.44	1.21	1.00
Ramachandran Favoured (%)	98	98	96	98
Ramachandran Outliers (%)	0	0	0	0
Average β-factor	19.80	29.10	29.60	29.40
Macromolecule	16.90	27.90	28.00	28.30
Ligands	16.50	40.10	46.20	-
Solvent	33.10	36.90	41.10	37.30
Mol probity score (percentile)	1.20 (97 th)	1.50 (91 st)	-	0.99 (100 th)

Appendix Three

Sequencing analysis of antibody clones

Highlighted in yellow: start and end of complementary determining regions (CDR- light 1, 2, 3 and heavy 1, 2, 3); Highlighted in red: linker; Highlighted in green; Histidine tag (6xHis Tag) and hemagglutinin (HA-Tag).

Anti-cPSA scFv 2D2SG

AGLTQPSSVSANPGETVKIX **CSGSSSA**YEY **WY**QQKSPGSAPVTLY **Y**NN
CDR L1
DKRP **SDIPS**RFSGSKFGSTATLTITGVQADDEAVYF **CG**SADSNADGI **FGA**
CDR L2 CDR L3
GTTLTVL **GQSSRSS**AVTLDES GGG LQTPGGALSLV **CKAS**GFSSFSYAMet
Linker CDR H1
G **WV**RQAPGKG **LEWVAD**IGSSTTWYGA AV **KGRA**TISRDNRRQSTMetRLQ
CDR H2
LNNLRAEDTG VYY **CAKYV**VSTGGRWGVYSIDA **WGHG**TEVIVSSTSGQA
CDR H3
GQ **HHHHHHGAYPDV**PDYAS Stop
6xHis Tag HA-Tag

Anti-cPSA scFv 2D8SG

CDR L1
 AALTQPSSVSANPGETVKITCSGGGRYAGSYYYGWYQQKSPGSAPVTLIY
CDR L2 CDR L3
 ESNKRPSNIPSRFSGSTSGSTGTLTITGVQADDEAVYFCGSRDSTYAAF
Linker
 AGTTLTGSSSRSSSGGGSSGGGSAVTLDESGGGLQTPGGALSLVCKAS
CDR H1 CDR H2
 GFTFSSHGMetFWVRQAPGKGLEWVG GIDGGGGVTWYGAAVDGRATIS
CDR H3
 RDNGQSTVRLQLNNLRAEDTATYYCARESYSSGTIDIWGRGTEVIVSSTS
6xHis Tag HA-Tag
 GQAGQHGGGGGGGAYPDVPDYAS

Anti-fPSA scFv B8

CDR L1 CDR L2
 ALTQPSSVSANPGETVKITCSGSSGSYGWYQKSPDSAPVTVIYQSNQ
CDR L3
 RPSDIPSRFSGSKSGSTGTLTITGVQAEDEAVYYCGGWGSSVGMetFGA
Linker
 GTTLTVLGQSSRSSSGGGSSGGGSAVTLDESGGGLQTPGGALSLVCK
CDR H1 CDR H2
 ASGFTFSSYAMetGWVRQAPGKGLEWVAGISDDGDSYISYATAVKGRA
CDR H3
 TISRDNGQSTVRLQLNNLRAEDTATYYCARSHCSGCRNAALIDAWGHG
6xHis Tag HA-Tag
 TEVIVSSTSGQAGQHGGGGGGGAYPDVPDYAS Stop

Anti-fPSA scFvB8_Ser MT

ALTQPSSVSANPGETVKIT **CDR L1** CSGSSGSYG **WYQ** QKSPDSAPVTVI **CDR L2** IYQSNQ
 RP **SDIPS** RFSGSKSGSTGTLTITGVQAEDEAVYY **CDR L3** C GGWGSSVG Met **FGA**
Linker
G TTLTVL **GQSSRSSGGGGSSGGGGS** AVTLDES GGLQTPGGALSLV **CK**
CDR H1 **A** SGFTFSSYAMetG **WV** RQAPGKG **LEWVAG** **CDR H2** ISDDGDSYISYATAV **KGRA**
T ISRDNGQSTVRLQLNNLRAEDTATYY **CDR H3** CARSH **SSG** RNAALIDA **WGHG**
6xHis Tag **HA-Tag**
 TEVIVSSTSGQAGQ **HHHHHHGAYPYDVPDYAS** Stop

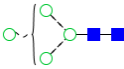





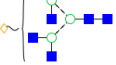



Anti-fPSA scFv B8_Ala MT

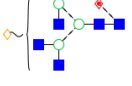
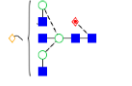
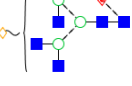
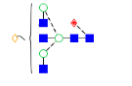
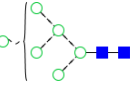
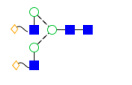
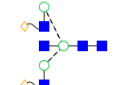

ALTQPSSVSANPGETVKIT **CDR L1** CSGSSGSYG **WYQ** QKSPDSAPVTVI **CDR L2** IYQSNQ
 RP **SDIPS** RFSGSKSGSTGTLTITGVQAEDEAVYY **CDR L3** C GGWGSSVG Met **FGA**
Linker
G TTLTVL **GQSSRSSGGGGSSGGGGS** AVTLDES GGLQTPGGALSLV **CK**
CDR H1 **A** SGFTFSSYAMetG **WV** RQAPGKG **LEWVAG** **CDR H2** ISDDGDSYISYATAV **KGRA**
T ISRDNGQSTVRLQLNNLRAEDTATYY **CDR H3** CARSH **ASG** RNAALIDA **WGHG**
6xHis Tag **HA-Tag**
 TEVIVSSTSGQAGQ **HHHHHHGAYPYDVPDYAS** Stop



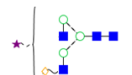
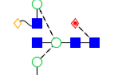
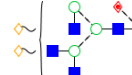

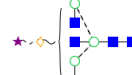
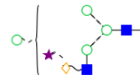
Appendix Four

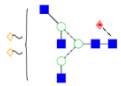
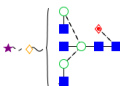
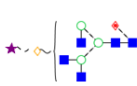
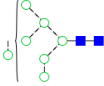
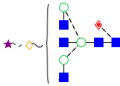
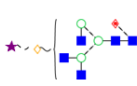
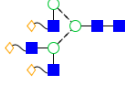
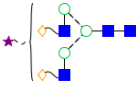
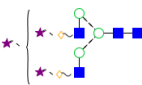
Summary table of N-glycans released from IgY purified from serum (Saldova *et al.*, 2014).


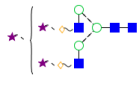
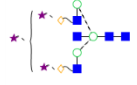

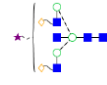


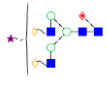
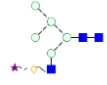
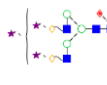
The Hydrophilic interaction liquid chromatography (HILIC)-chromatogram was separated into 40 peaks and structural assignment carried out using established methods (Royle *et al.*, 2008) and the software tool GlycoBase (<https://glycobase.nibrt.ie>). * indicates sialic acid linkages; Weak Anion Exchange (WAX) fractions were separated out into 5 fractions; S1: Monosialylated, S2: disialylated, S3: Trisialylated, S3J: Trisialylated and S4: Tetrasialylated. All monosialylated glycans are linked by α 2-3 and α 2-6; Disialylated glycans have all combinations of α 2-3 and α 2-6 linkages [i.e. (3, 3), (3, 6) and (6, 6)]; Trisialylated glycans have all combinations of α 2-3 and α 2-6 linkages except (6, 6, 6) [i.e. (3, 3, 3), (3, 3, 6) and (3, 6, 6)]; and Tetrasialylated glycans have all combinations of α 2-3 and α 2-6 linkages [i.e. (3,3,3,3), (3,3,3,6), (3,3,6,6) and (6,6,6,6)]. Structure abbreviations: all N-glycans have two core GlcNAcs; F at the start of the abbreviation indicates a core-fucose α 1,6-linked to the inner GlcNAc; Mx, number (x) of mannose on core GlcNAcs; Ax, number of antenna (GlcNAc) on trimannosyl core; A2, biantennary with both GlcNAcs as β 1,2-linked; A3, triantennary with a GlcNAc linked β 1,2 to both mannose and the third GlcNAc linked β 1,4 to the α 1,3 linked mannose; A3', isomer with the third GlcNAc linked β 1-6 to the α 1-6 linked mannose; A4, GlcNAcs linked as A3 with additional GlcNAc β 1,6 linked to α 1,6 mannose; B, bisecting GlcNAc linked β 1,4 to β 1,3 mannose; Gx, number (x) of β 1,4 linked galactose on antenna; F(x), number (x) of fucose linked α 1,3 to antenna GlcNAc; Sx, number (x) of sialic acids linked mostly to galactose (Structures with more sialic acids than galactoses may have some sialic acids linked to GlcNAc or another sialic acid in form of polysialic acid.); GU: Glucose unit.

Peak	Structure		Peak area (%)	GU	Experimental mass (m/z)	Theoretical mass (m/z)	Ion
5	M4		0.02	5.34	---	1191.4421	[M-H] ⁻
9	M5		2.17	6.14	1353.514 676.2448	1353.4949 676.2438	[M-H] ⁻ [M-2H] ⁻²
	FA2B				891.8257	891.8390	[M-2H] ⁻²
	FA3						
	A2[6]G1				798.2899	798.2968	[M-2H] ⁻²
10	A2[3]G1		0.22	6.51	798.2899	798.2968	[M-2H] ⁻²
	A3G1				899.8270	899.8365	[M-2H] ⁻²
	A2BG1						
11	M5A1		0.17	6.61	777.7769	777.7835	[M-2H] ⁻²
	M4A1G1						


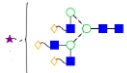
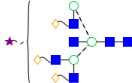
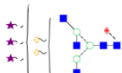
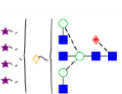

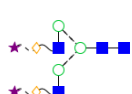
Peak	Structure		Peak area (%)	GU	Experimental mass (m/z)	Theoretical mass (m/z)	Ion
12	FA3G1		2.26	6.90	972.8573	972.8654	$[M-2H]^{-2}$
	FA2BG1						
13	FA3G1		1.22	7.04	972.8509	972.8654	$[M-2H]^{-2}$
	FA2BG1						
14	M6		0.44	7.14	757.2596	757.2702	$[M-2H]^{-2}$
	A2G2				879.3184	879.3232	$[M-2H]^{-2}$
15	A2BG2		0.92	7.32	980.8661	980.8629	$[M-2H]^{-2}$
	A2[6]G1S1*				943.8293	943.8445	$[M-2H]^{-2}$

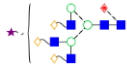

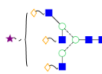






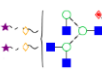
Peak	Structure		Peak area (%)	GU	Experimental mass (m/z)	Theoretical mass (m/z)	Ion
16	FA2G2		0.59	7.52	952.3262	952.3521	$[M-2H]^{-2}$
	M5A1G1				858.7934	858.8099	$[M-2H]^{-2}$
	A2[3]G1S1*				943.8293	943.8445	$[M-2H]^{-2}$
17	FA2BG2		5.14	7.69	1053.8828	1053.8918	$[M-2H]^{-2}$
	FA3G2				---	1045.3842	$[M-2H]^{-2}$
	A3G1S1*				---	1045.3842	$[M-2H]^{-2}$
	A2BG1S1*				---	1045.3842	$[M-2H]^{-2}$
	M4A1G1S1*				923.3214	923.3312	$[M-2H]^{-2}$

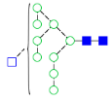
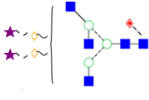
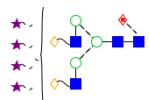
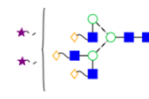
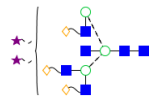
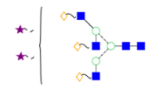
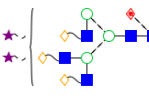
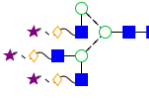
Peak	Structure		Peak area (%)	GU	Experimental mass (m/z)	Theoretical mass (m/z)	Ion
18	FA3'G2		1.60	7.93	1053.8828	1053.8918	$[M-2H]^{-2}$
	FA2BG1S1*				1118.4037	1118.4131	$[M-2H]^{-2}$
	FA3G1S1*						
19	M7		0.51	8.02	838.2908	838.2966	$[M-2H]^{-2}$
	FA2BG1S1*				1118.4243	1118.4131	$[M-2H]^{-2}$
	FA3G1S1*						
20	A3G3		16.38	8.36	---	1061.8893	$[M-2H]^{-2}$
	A2G2S1*				1024.8506	1024.8709	$[M-2H]^{-2}$
	A2G2S3*				---	1315.9663	$[M-2H]^{-2}$

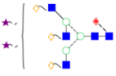
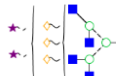

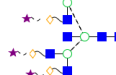



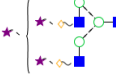

Peak	Structure		Peak area (%)	GU	Experimental mass (m/z)	Theoretical mass (m/z)	Ion
21	A3BG3		0.65	8.51	---	1163.4290	$[M-2H]^{-2}$
	A2G2S3*				---	1315.9663	$[M-2H]^{-2}$
	A2BG2S3*				---	1417.5060	$[M-2H]^{-2}$
22	FA3G3		2.12	8.64	1134.9095	1134.9182	$[M-2H]^{-2}$
	A2BG2S1*				1126.4175	1126.4106	$[M-2H]^{-2}$
	A3'G3				---	1061.8893	$[M-2H]^{-2}$
23	M8		4.00	8.78	919.3203	919.3230	$[M-2H]^{-2}$
	FA2G2S1*				1097.8785	1097.8998	$[M-2H]^{-2}$
	M5A1G1S1*				1004.3476	1004.3576	$[M-2H]^{-2}$
	FA2G2S3*				---	1388.9953	$[M-2H]^{-2}$

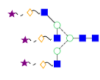
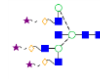
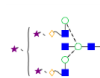
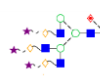
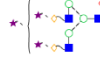
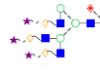
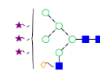
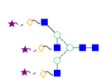
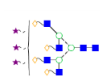
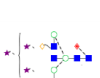
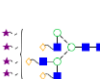
Peak	Structure		Peak area (%)	GU	Experimental mass (m/z)	Theoretical mass (m/z)	Ion
24	FA3'G3		8.68	9.00	1134.9095	1134.9182	$[M-2H]^{-2}$
	FA2BG2S1*				1199.4524	1199.4395	$[M-2H]^{-2}$
	FA3G2S1*						
25	FA4G3		3.93	9.23	---	1236.4579	$[M-2H]^{-2}$
	FA3'G2S1*				1199.4240	1199.4395	$[M-2H]^{-2}$
	A2G2S2*				1170.4113	1170.4186	$[M-2H]^{-2}$
	FA2BG2S3*				---	1490.5349 993.3542	$[M-2H]^{-2}$ $[M-3H]^{-3}$
	M5A1G1S3*				---	1295.4530 863.2996	$[M-2H]^{-2}$ $[M-3H]^{-3}$
	FA3G2S3*				---	1490.5349 993.3542	$[M-2H]^{-2}$ $[M-3H]^{-3}$

Peak	Structure		Peak area (%)	GU	Experimental mass (m/z)	Theoretical mass (m/z)	Ion
26	M9		5.11	9.51	1000.3368	1000.3495	$[M-2H]^{-2}$
	A3G3S1*				1207.4227	1207.4370	$[M-2H]^{-2}$
27	A3BG3S1*		24.92	9.65	---	1308.9767	$[M-2H]^{-2}$
	FA3'G2S3*				---	1490.5349 993.3542	$[M-2H]^{-2}$ $[M-3H]^{-3}$
	FA2BG1S4*				---	1555.0562 1036.3684	$[M-2H]^{-2}$ $[M-3H]^{-3}$
	FA3G1S4*				---	1555.0562 1036.3684	$[M-2H]^{-2}$ $[M-3H]^{-3}$
	A2G2S2*				1170.4043	1170.4186	$[M-2H]^{-2}$

Peak	Structure		Peak area (%)	GU	Experimental mass (m/z)	Theoretical mass (m/z)	Ion
28	FA3G3S1*		0.67	9.82	---	1280.4659	$[M-2H]^{-2}$
	A4G4				1244.4453	1244.4554	$[M-2H]^{-2}$
	A3'G3S1*				1207.4299	1207.4370	$[M-2H]^{-2}$
	A2BG2S2*				---	1271.9583	$[M-2H]^{-2}$
	A4G4S4*				---	1826.6462 1217.4284	$[M-2H]^{-2}$ $[M-3H]^{-3}$
29	FA3'G3S1*		0.24	9.90	---	1280.4659	$[M-2H]^{-2}$
	FA2G2S2*				1243.4315	1243.4476	$[M-2H]^{-2}$
30	FA4G4		1.19	10.06	---	1317.4843	$[M-2H]^{-2}$
	FA2BG2S2*				1344.9840	1344.9872	$[M-2H]^{-2}$
	FA3G2S2*						

Peak	Structure		Peak area (%)	GU	Experimental mass (m/z)	Theoretical mass (m/z)	Ion
31	M9Glc		10.11	10.24	1081.3684	1081.3759	$[M-2H]^{-2}$
	FA3'G2S2*				1344.9840	1344.9872	$[M-2H]^{-2}$
	FA2G2S4*				---	1534.5430 1022.6929	$[M-2H]^{-2}$ $[M-3H]^{-3}$
	A3G3S2*				1352.9702	1352.9847	$[M-2H]^{-2}$
32	A3BG3S2*		0.97	10.67	---	1454.5244	$[M-2H]^{-2}$
33	A3'G3S2*		1.11	10.80	1352.9626	1352.9847	$[M-2H]^{-2}$
34	FA3G3S2*		0.65	11.17	---	1426.0137	$[M-2H]^{-2}$
	A3G3S3*				1498.5284 998.6880	1498.5324 998.6858	$[M-2H]^{-2}$ $[M-3H]^{-3}$

Peak	Structure		Peak area (%)	GU	Experimental mass (m/z)	Theoretical mass (m/z)	Ion
35	FA3'G3S2*		0.06	11.33	---	1426.0137	$[M-2H]^{-2}$
	FA4G3S2*				---	1527.5533	$[M-2H]^{-2}$
	A3G3S3*				1498.5126 998.6880	1498.5324 998.6858	$[M-2H]^{-2}$ $[M-3H]^{-3}$
	A3BG3S3*				---	1673.1010 1115.0649	$[M-2H]^{-2}$ $[M-3H]^{-3}$
36	A3G3S3*		0.10	11.46	1498.5284 998.6880	1498.5324 998.6858	$[M-2H]^{-2}$ $[M-3H]^{-3}$
	A4G4S2*				1535.5466	1535.5508	$[M-2H]^{-2}$
	A2G2S3*				---	1315.9663 876.9757	$[M-2H]^{-2}$ $[M-3H]^{-3}$
37	A2G2S3*		2.81	11.59	---	1315.9663 876.9757	$[M-2H]^{-2}$ $[M-3H]^{-3}$
	A3G3S3*				1498.5284 998.6880	1498.5324 998.6858	$[M-2H]^{-2}$ $[M-3H]^{-3}$

Peak	Structure		Peak area (%)	GU	Experimental mass (m/z)	Theoretical mass (m/z)	Ion
38	A3'G3S3*		0.13	11.92	---	1498.5324 998.6858	[M-2H] ⁻² [M-3H] ⁻³
	A3BG3S3*				---	1600.0721 1066.3790	[M-2H] ⁻² [M-3H] ⁻³
	A2BG2S3*				---	1417.5060 944.6682	[M-2H] ⁻² [M-3H] ⁻³
	FA3G3S3*				---	1571.5614 1047.3718	[M-2H] ⁻² [M-3H] ⁻³
39	FA2G2S3*		0.52	12.09	---	1388.9953 925.6611	[M-2H] ⁻² [M-3H] ⁻³
	FA3G3S3*				---	1571.5614 1047.3718	[M-2H] ⁻² [M-3H] ⁻³
40	M5A1G1S3*		0.10	12.66	---	1295.4530 863.2996	[M-2H] ⁻² [M-3H] ⁻³
	A3'G3S3*				1498.5284 998.6880	1498.5324 998.6858	[M-2H] ⁻² [M-3H] ⁻³
	A4G4S3*				---	1681.0985 1120.3966	[M-2H] ⁻² [M-3H] ⁻³
	FA2BG2S3*				---	1490.5349 993.3542	[M-2H] ⁻² [M-3H] ⁻³
	A3G3S4*				---	1644.0801 1095.7177	[M-2H] ⁻² [M-3H] ⁻³

Appendix Five

Clinical data for prostate cancer patient cohort; serum samples for whole serum *N*-glycome analysis

Source: Prostate Cancer Research Consortium (PCRC) BioResource

RP: Radical Prostatectomy, PSA: Prostate Specific Antigen, DRE: Digital Rectal Examination

PCRC NUMBER	Cohort	Biopsy Gleason A	Biopsy Gleason B	Biopsy Gleason Score	RP Gleason A	RP Gleason B	RP Gleason Score	PSA (ng/mL)	Age at RP	DRE
81	Indolent	2	3	5	3	2	5	9.5	64	
92	Indolent	3	2	5	3	2	5	8.45	66	
93	Indolent				2	3	5	4.4	65	Negative
97	Aggressive				4	3	7	6.1	65	Positive
104	Aggressive	4	4	8	4	3	7	6.5	55	
114	Indolent				3	2	5	4.5	52	
117	Aggressive	3	3	6	3	4	7	6.5	69	Negative
121	Aggressive	4	3	7	3	4	7	11.3	60	Negative
122	Aggressive	4	5	9	4	3	7	8.82	63	Negative
124	Aggressive	3	4	7	4	3	7	6.2	69	Negative
151	Indolent	2	3	5	3	2	5	10.7	55	Positive
161	Indolent	2	3	5	3	2	5	5.3	61	Negative

PCRC NUMBER	Cohort	Biopsy Gleason A	Biopsy Gleason B	Biopsy Gleason Score	RP Gleason A	RP Gleason B	RP Gleason Score	PSA (ng/mL)	Age at RP	DRE
181	Significant	3	4	7	3	4	7	8.8	64	Negative
168	Indolent	3	4	7	3	3	6	5.3	49	Positive
185	Indolent	2	3	5	3	3	6	8.3	59	
189	Aggressive	4	3	7	4	3	7	6.3	59	Positive
198	Indolent	2	3	5	2	3	5	10.4	68	Negative
220	Aggressive	3	4	7	4	3	7	7.3	67	
223	Indolent	3	2	5	2	3	5	6.5	62	
233	Indolent	2	3	5	2	3	5	4.5	59	Negative
234	Indolent	2	3	5	2	3	5	0.7	54	Positive
242	Aggressive	4	3	7	4	4	8	8.7	65	
246	Indolent	2	3	5	3	2	5	7.7	58	Negative
258	Significant	3	4	7	3	4	7	7.8	52	Negative
260	Significant	3	4	7	4	4	8	8.1	66	Positive
261	Aggressive	3	3	6	4	4	8	8.1	67	Negative
278	Indolent	3	3	6	3	3	6	8.3	64	Negative
288	Aggressive	4	5	9	4	3	7	5.3	58	
298	Indolent	3	3	6	3	3	6	9.2	55	
303	Indolent	3	3	6	3	3	6	12	59	Negative
311	Indolent	3	3	6	3	3	6		54	Negative
314	Indolent	3	3	6	3	3	6	2.27	51	Negative
315	Indolent	3	3	6	3	2	5	10.4	56	Negative
317	Indolent	3	3	6	3	3	6	3.9	49	Negative
318	Aggressive		3		3	4	7	4	66	Positive
321	Indolent	3	3	6	3	3	6	7.45	61	
335	Aggressive	3			3	4	7	8.2	62	Negative

PCRC NUMBER	Cohort	Biopsy Gleason A	Biopsy Gleason B	Biopsy Gleason Score	RP Gleason A	RP Gleason B	RP Gleason Score	PSA (ng/mL)	Age at RP	DRE
338	Aggressive	3	4	7	4	4	8	8	56	
341	Indolent	2	3	5	3	2	5	1.2	60	Negative
344	Aggressive	3	3	6	4	3	7	17	54	
363	Indolent	2	3	5	3	2	5	6	52	Negative
365	Significant	3	4		4	3	7	13.4	64	Positive
383	Aggressive	4	3	7	4	3	7	9.4	71	Negative
387	Significant	3	3	6	3	4	7	6.9	52	
391	Aggressive	4	4	8	4	3	7	18.7	60	Negative
396	Aggressive	3	4	7	3	4	7	4.2	52	Negative
405	Indolent	3	3	6	3	3	6	1.3	60	
407	Indolent	3	3	6	3	3	6	5.93	64	
411	Significant	3	3	6	3	4	7	6.1	55	
416	Significant	4	4	8	3	4	7	6.39	67	
419	Significant	3	3	6	4	4	8	6.3	73	Negative
421	Indolent	3	3	6	3	3	6	7.01	56	Negative
424	Indolent	3	4	7	3	3	6	5.92	58	Negative
427	Indolent	3	3	6	3	3	6	12.3	58	Negative
446	Indolent	3	3	6	3	3	6	6.2	63	
447	Significant	3	4	7	3	4	7	5.71	68	Negative
450	Significant	3	3	6	3	4	7	12.3	57	
451	Aggressive				3	4	7	10.2	66	
452	Significant	4	3	7	3	4	7	3.5	59	Negative
464	Indolent	3	3	6	3	3	6	9.8	62	Negative
467	Aggressive	4	4	8	4	3	7	15.6	55	Positive
471	Significant	3	3	6	3	4	7	4.8	60	Positive

PCRC NUMBER	Cohort	Biopsy Gleason A	Biopsy Gleason B	Biopsy Gleason Score	RP Gleason A	RP Gleason B	RP Gleason Score	PSA (ng/mL)	Age at RP	DRE
472	Indolent	3	3	6	3	3	6	3.69	54	Negative
483	Indolent	3	3	6	3	3	6	0.9	61	Positive
479	Aggressive	3	4	7	4	3	7	7.9	63	Positive
487	Indolent	2	3	5	3	3	6		70	
493	Significant	3	3	6	3	4	7	2.8	56	Negative
497	Aggressive	3	4	7	4	3	7	5.5	69	Positive
502	Aggressive	3	4	7	4	3	7		57	
503	Significant	4	3	7	4	4	8	13.1	68	Positive
504	Significant	3	4	7	3	4	7	5.3	65	Negative
514	Significant	3	3	6	3	4	7	11.9	58	Negative
524	Significant	3	4	7	3	4	7	14.7	56	Negative
528	Aggressive	3	3	6	3	4	7	3.6	60	Negative
555	Indolent	3	3	6	3	3	6	4.8	67	Negative
576	Significant	4	3	7	3	4	7	8.3	55	Negative
577	Indolent	3	3	6	3	3	6	9.6	66	
592	Aggressive	4	5	9	3	4	7	7.3	67	Negative
602	Significant	3	3	6	4	3	7	6.79	68	Negative
604	Aggressive	3	4	7	4	4	8	7	63	Negative
605	Significant	3	3	6	3	4	7	5.66	63	
607	Aggressive	4	4	8	3	4	7	13.2	62	Negative
610	Aggressive	4	5	9	4	4	8	4.1	73	Positive
611	Aggressive	4	3	7	4	3	7	8.58	65	Negative
612	Significant	3	4	7	3	4	7	6.3	61	Negative
615	Indolent	3	3	6	3	3	6	5	59	
616	Aggressive	3	4	7	4	4	8	5.2	55	Negative

PCRC NUMBER	Cohort	Biopsy Gleason A	Biopsy Gleason B	Biopsy Gleason Score	RP Gleason A	RP Gleason B	RP Gleason Score	PSA (ng/mL)	Age at RP	DRE
619	Indolent	3	4	7	3	2	5	9.5	56	Negative
621	Significant	3	3	6	3	4	7	5.88	56	Positive
628	Significant	3	4	7	3	4	7	4.8	61	Negative
630	Significant	3	4	7	3	4	7	2.8	51	
635	Aggressive	3	3	3	6	3	4	7	60	
641	Significant	3	4	7	3	4	7	5.7	58	Negative
642	Aggressive	4	3	7	4	3	7	5.1	64	Positive
643	Significant	3	4	7	3	4	7	11.9	59	Negative
644	Aggressive	3	4	7	3	4	7	6.7	62	
647	Significant	4	3	7	3	4	7	3.7	61	Positive
654	Significant	3	4	7	3	4	7	7.2	62	
655	Aggressive	4	3	7	3	4	7	6.5	64	Positive
659	Significant	3	4	7	3	4	7	3.3	49	Negative
664	Indolent	3	3	6	3	3	6	5	60	Negative
665	Aggressive	3	4	7	3	4	7	5	59	
673	Significant	3	4	7	3	4	7	5.2	59	Negative
679	Aggressive	3	3	6	3	4	7	7.2	58	
680	Aggressive	3	3	6	4	3	7	11.2	56	
681	Aggressive	4	4	8	4	3	7	6.6	69	Positive
685	Significant	3	4	7	3	4	7	4.9	55	Negative
704	Significant	3	3	6	3	4	7	7.1	60	Negative
705	Indolent	3	3	6	3	3	6	3.8	57	Negative
706	Aggressive	3	4	7	4	3	7	3.6	68	
713	Aggressive	3	4	7	3	4	7	5.3	67	Negative
717	Aggressive	4	4	8	3	4	7	5.6	59	Negative

PCRC NUMBER	Cohort	Biopsy Gleason A	Biopsy Gleason B	Biopsy Gleason Score	RP Gleason A	RP Gleason B	RP Gleason Score	PSA (ng/mL)	Age at RP	DRE
724	Aggressive				4	4	8	3.1	62	
732	Aggressive	3	4	7	3	4	7	8.4	74	Negative
735	Aggressive	4	3	7	3	4	7	3.2	61	Positive

Bibliography

Bibliography

Abbott, W.M., Damschroder, M.M., and Lowe, D.C. (2014). Current approaches to fine mapping of antigen-antibody interactions. *Immunology*, **142** (4),pp. 526–35.

Acharya, K. R. and Lloyd, M. D. (2005). The advantages and limitations of protein crystal structures. *Trends in Pharmacological Sciences* **26** (1), pp. 10–14.

Adlersberg, J. B. (1976). The immunoglobulin hinge (interdomain) region. *La Ricerca in Clinica e in Laboratorio*, **6** (3), pp. 191–205.

American Cancer Society (2015). Available online at:

<http://www.cancer.org/cancer/prostatecancer/detailedguide/prostate-cancer-key-statistics>
(Accessed June, 2015).

Andris-Widhopf, J., Rader, C., Steinberger, P., Fuller, R. and Barbas, C. F. (2000). Methods for the generation of chicken monoclonal antibody fragments by phage display. *Journal of Immunological Methods*, **242** pp. 159–81.

Armbruster, D. A. and Pry, T. (2008). Limit of blank, limit of detection and limit of quantitation. *Clinical Biochemist Reviews*, **29** Suppl. 1:S49-52.

Armbruster, D. A. (1993). Prostate-specific antigen: biochemistry, analytical methods, and clinical application. *Clinical Chemistry*, **39** pp. 181–95.

Arnold, J. N., Saldova, R., Hamid, U.M., Rudd, P.M. (2008). Evaluation of the serum N-linked glycome for the diagnosis of cancer and chronic inflammation. *Proteomics*, **8** (16), pp. 3284-93.

- Balk, S. P.,** Ko, Y. J., and Bubley, G. J. (2003). Biology of prostate-specific antigen. *Journal of Clinical Oncology*, **21** pp. 383–91.
- Banerjee, D. K.** (2012) N-glycans in cell survival and death: Cross-talk between glycosyltransferases. *Biochimica et Biophysica Acta*, **1820** (9) pp. 1338–46.
- Barbas, C. F.,** Burton, D. R., Scott, J. K. and Silverman, G. J. (2001). *Phage Display: A Laboratory Manual*. (Cold Spring Harbor Laboratory Press, New York, USA).
- Barbas, C. F.,** Kang, A. S., Lerner, R. A. and Benkovic, S. J. (1991). Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proceedings of the National Academy of Sciences (USA)*, **88** pp. 7978–82.
- Berman, H.M.,** Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. *Nucleic acids research*, **28** (1), pp. 235–242.
- Biacore** (2015). Available online at: https://www.biacore.com/lifesciences/products/systems/_overview/index.html (Accessed June 2015).
- Bigge, J. C.,** Patel, T. P., Bruce, J. A., Goulding, P. N., Charles, S. M., and Parekh, R. B.(1995) Nonselective and efficient fluorescent labeling of glycans using 2-amino-benzamide and anthranilic acid. *Analytical Biochemistry*, **230** pp. 229–38.
- Biorad** (2015). Available online at: <http://www.bio-rad.com/en-us/product/trans-blot-sd-semi-dry-transfer-cell> (Accessed June 2015).
- Birtalan, S.,** Zhang, Y., Fellouse, F.A., Shao, L., Schaefer, G., and Sidhu, S.S. (2008). The intrinsic contributions of tyrosine, serine, glycine and arginine to the affinity and specificity of antibodies. *Journal of Molecular Biology*, **377** pp. 1518–28.

Bjorkman, P. J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. (1987). Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* (London), **329** (6139), pp. 506–12.

Bones, J., Mittermayr, S., O'Donoghue, N., Guttman, A., and Rudd, P.M. (2010). Ultra Performance Liquid Chromatographic Profiling of Serum N-Glycans for Fast and Efficient Glycosylation. *Analytical Chemistry*, **82** (24), pp. 10208–215.

Boswell, C. A., Tesar, D.B., Mukhyala, K., Theil, F.P., Fielder, P.J., Khawli, L.A. (2010). Effects of charge on antibody tissue distribution and pharmacokinetics. *Bioconjugate Chemistry*, **21** (12) pp.2153–63.

Botchorishvili, G., Matikainen, M. P. and Lilja, H. (2009). Early prostate-specific antigen changes and the diagnosis and prognosis of prostate cancer. *Current Opinion in Urology*, **19** pp. 221–26.

Bott, S. R. J. (2003). Prostate cancer management: (2) an update on locally advanced and metastatic disease. *Postgraduate Medical Journal*, **79** pp. 643–45.

Bradbury, A. R. M. and Marks, J. D. (2004). Antibodies from phage antibody libraries. *Journal of Immunological Methods*, **290** (1-2) pp. 29–49.

Bradbury, A. R. M., Sidhu, S., Dübel, S. and McCafferty, J. (2011). Beyond natural antibodies: the power of *in vitro* display technologies. *Nature Biotechnology*, **29** pp. 245–54.

Bratkovic, T. (2010). Progress in phage display: evolution of the technique and its application. *Cellular and Molecular Life Sciences*, **67** pp. 749–67.

Büll, C., Stoel, M.A., Brok, M.H.M.G.M. den., and Adema, G.J. (2014) Sialic acids sweeten a tumor's life. *Cancer Research*, **74** (12) pp. 3199–204.

Campbell, M. P., Royle, L., Radcliffe, C.M., Dwek, R.A., and Rudd, P.M. (2008) GlycoBase and autoGU: tools for HPLC-based glycan analysis. *Bioinformatics*, **24** (9), pp. 1214–16.

Catalona, W. J., Bartsch, G., Rittenhouse, H.G., Evans, C.L., Linton, H.J., Horninger, W., Klocker, H., and Mikolajczyk, S.D. (2004). Serum pro-prostate specific antigen preferentially detects aggressive prostate cancers in men with 2 to 4 ng/ml prostate specific antigen. *The Journal of Urology*, **171** pp. 2239–44.

Catalona, W. J., Partin, A.W., Sanda, M.G., Wei, J.T., Klee, G.G., Bangma, C.H., Slawin, K.M., Marks, L.S., Loeb, S., Broyles, D.L., Shin, S.S., Cruz, A.B., Chan, D.W., Sokoll, L.J., Roberts, W.L., van Schaik, R.H., and Mizrahi, I.A. (2011). A multicenter study of [-2]pro-prostate specific antigen combined with prostate specific antigen and free prostate specific antigen for prostate cancer detection in the 2.0 to 10.0 ng/ml prostate specific antigen range. *The Journal of Urology*, **185** (5), pp. 1650–55.

Catalona, W. J., Partin, A.W., Sanda, M.G., Wei, J.T., Klee, G.G., Bangma, C.H., Slawin, K.M., Marks, L.S., Loeb, S., Broyles, D.L., Shin, S.S., Cruz, A.B., Chan, D.W., Sokoll, L.J., Roberts, W.L., van Schaik, R.H., and Mizrahi, I.A. (2011). A Multi-Center study of [-2]pro-prostate specific antigen (PSA) in Combination with PSA and free PSA for Prostate Cancer Detection in the 2.0 to 10.0 ng/ml PSA range. *The Journal of Urology*, **185** pp. 1650–55.

Chailyan, A., Marcatili, P., Cirillo, D., and Tramontano, A. (2011). Structural repertoire of immunoglobulin λ light chains. *Proteins: Structure, Function and Bioinformatics*, **79** (5), pp. 1513–24.

- Chandrasekaran, E. V.,** Chawda, R., Locke, R.D., Piskorz, C.F., and Matta, K.L. (2002). Biosynthesis of the carbohydrate antigenic determinants, Globo H, blood group H, and Lewis b: a role for prostate cancer cell alpha1,2-L-fucosyltransferase. *Glycobiology*, **12** (3), pp. 153–62.
- Chothia, C.** and Lesk, A. M. (1987). Canonical structures for the hypervariable regions of immunoglobulins. *Journal of Molecular Biology*, **196** (4), pp. 901–17.
- Christensson, A.,** Björk, T., Nilsson, O., Dahlén, U., Matikainen, M.T., Cockett, A.T., Abrahamsson, P.A., and Lilja, H. (1993). Serum prostate specific antigen complexed to alpha 1-antichymotrypsin as an indicator of prostate cancer. *The Journal of Urology*, **150** pp. 100–05.
- Chung, C. H.,** Mirakhur, B., Chan, E., Le, Q.T., Berlin, J., Morse, M., Murphy, B.A., Satinover, S.M., Hosen, J., Mauro, D., Slebos, R.J., Zhou, Q., Gold, D., Hatley, T., Hicklin, D.J., and Platts-Mills, T.A. (2008). Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. *New England Journal of Medicine*, **358** (11), pp. 1109–17.
- Cohen, M.** and Varki, A. (2010). The sialome--far more than the sum of its parts. *OMICS: A Journal of Integrative Biology*, **14** (4), pp. 455–64.
- Conroy, P. J.,** Hearty, S., Leonard, P., and O'Kennedy, R.J. (2009). Antibody production, design and use for biosensor-based applications. *Seminars in Cell & Developmental Biology*, **20** (1), pp. 10–26.
- Conroy, P. J.,** Law, R.H., Gilgunn, S., Hearty, S., Caradoc-Davies, T.T., Lloyd, G., O'Kennedy, R.J., and Whisstock, J.C. (2014). Reconciling the structural attributes of avian antibodies. *Journal of Biological Chemistry*, **289** (22) pp. 15384–92.

Conroy, P. J., O’Kennedy, R. J. and Hearty, S. (2012). Cardiac troponin I: a case study in rational antibody design for human diagnostics. *Protein Engineering, Design & Selection*, **25** pp. 295–305.

Constantinoiu, C. C., Molloy, J.B., Jorgensen, W.K., and Coleman, G.T. (2007). Purification of immunoglobulins from chicken sera by thiophilic gel chromatography. *Poultry Science*, **86** (9) pp. 1910-14.

Corley, R. B. (2005). *A Guide to Methods in the Biomedical Science*. Boston: Springer Sciences.

Cwirla, S. E., Peters, E. A., Barrett, R. W. and Dower, W. J. (1990). Peptides on phage: a vast library of peptides for identifying ligands. *Proceedings of the National Academy of Sciences (USA)*, **87** pp. 6378–82.

deCrescenzo, G, Boucher, C, Durocher, Y, and Jolicoeur, M. (2008). Kinetic Characterisation by Surface Plasmon Resonance-Based Biosensors: Principles and Emerging Trends. *Cellular and Molecular Bioengineering*, **1** (4), pp. 204-15.

Diamond Light Source (2015). Available online at: <http://www.diamond.ac.uk/Home.html> (Accessed July 2015)

Doores, K. J., Kong, L., Krumm, S.A., Le, K.M., Sok, D., Laserson, U., Garces, F., Poignard, P., Wilson, I.A., and Burton, D.R. (2015). Two classes of broadly neutralizing antibodies within a single lineage directed to the high-mannose patch of HIV envelope. *Journal of Virology*, **89** (2), pp. 1105–18.

Drake, P. M., Cho, W., Li, B., Prakobphol, A., Johansen, E., Anderson, N.L., Regnier, F.E., Gibson, B.W. and Fisher, S.J. (2010). Sweetening the pot: Adding glycosylation to the biomarker discovery equation. *Clinical Chemistry*, **56** (2) pp. 223–36.

Dube, D. H. and Bertozzi, C. R. (2005). Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nature Reviews Drug Discovery*, **4** (6),pp. 477–88.

Epstein, J. I. (2010). An Update of the Gleason Grading System. *The Journal of Urology*, **183** (2) pp. 433–40.

Epstein, J.I., Walsh, P.C., Carmichael, M., and Brendler, C.B. (1994). Pathologic and clinical findings to predict tumor extent of nonpalpable (stage T1c) prostate cancer. *The Journal of the American Medical Association*, **271** (5) pp. 368-74.

Eriksson, S., Vehniäinen, M., Jansén, T., Meretoja, V., Saviranta, P., Pettersson, K., and Lövgren, T. (2000). Dual-label time-resolved immunofluorometric assay of free and total prostate-specific antigen based on recombinant Fab fragments. *Clinical Chemistry*, **46** (5),pp. 658–66.

Eusebi, P. (2013) Diagnostic accuracy measures. *Cerebrovascular Diseases*, **36** (4), pp.267–72.

Evans, G., Axford, D., and Owen, R.L. (2011). The design of macromolecular crystallography diffraction experiments. *Acta Crystallographica Section D: Biological Crystallography*, **67** (4), pp. 261–70.

Fan, Y., Murphy, T.B., Byrne, J.C., Brennan, L., Fitzpatrick, J.M., and Watson, R.W. (2011). Applying random forests to identify biomarker panels in serum 2D-DIGE data for the detection and staging of prostate cancer. *Journal of Proteome Research*, **10** (3), pp. 1361–73

Ferens-Sieczkowska, M., Kowalska, B., and Kratz, E.M. (2012). Seminal plasma glycoproteins in male infertility and prostate diseases: is there a chance for glyco-biomarkers? *Biomarkers*, **18**(1) pp. 10-22.

Ferlay, J., Parkin, D.M and Steliarova-Foucher, E. (2010). Estimates of incidence and mortality in Europe in 2008. *European Journal of Cancer*, **46** pp. 765-81.

Finlay, W. J. J. and Almagro, J. C. (2012). Natural and man-made V-gene repertoires for antibody discovery. *Frontiers in Immunology*, **3**:342. doi: 10.3389/fimmu.2012.00342. eCollection

Francis, D. M. and Page, R. (2010). Strategies to optimize protein expression in *E. coli*. *Current Protocols in Protein Science*, **61**:5.24.1-5.24.29.

Freeze, H. H. (2006). Genetic defects in the human glycome. *Nature Reviews Genetics*, **7** (7), pp. 537–51.

Georgia Urology. (2015). Available online at: <http://www.gaurology.com/specialties/georgia-urology-prostate-center/prostate-cancer-georgia/> (Accessed July 2015).

Giegé, R. (2013). A historical perspective on protein crystallization from 1840 to the present day. *FEBS Journal*, **280** (24) pp. 6456–97.

Gilgunn, S., Conroy, P.J., Saldova, R., Rudd, P.M., and O'Kennedy, R.J. (2013). Aberrant PSA glycosylation-a sweet predictor of prostate cancer. *Nature Reviews Urology*, **10** (2), pp. 99–107.

Gleason, D. F. and Mellinger, G. T. (1974). Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. *The Journal of Urology*, **111** pp. 58–64.

Grossman, T. H, Kawasaki, E.S., Punreddy, S.R., Osburne, M.S. (1998). Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability. *Gene*, **209** (1-2), pp. 95–103.

Hansen, J. E., Lund, O., Nielsen, J.O., Hansen, J.E., and Brunak, S. (1996). O-GLYCBASE: a revised database of O-glycosylated proteins. *Nucleic Acids Research*, **24** (1), pp. 248–52.

Harvey, D. J., Merry, A.H., Royle, L., Campbell, M.P., Dwek, R.A., and Rudd, P.M. (2009). Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds. *Proteomics*, **9** (15) pp. 3796–801.

Hassan, O., Ahmad, A., Sethi, S., and Sarkar, F.H. (2012). Recent updates on the role of microRNAs in prostate cancer. *Journal of Haematology and Oncology*, **5** (1) pp. 9-19.

Hayes, C. J., Leonard, P., and O'Kennedy, R.J. (2012). Overcoming antibody expression and screening limitations by smart design: applications to PSA immunoassay development. *Protein Expression and Purification*, **83** (1), pp. 84–91.

Hayhurst, A., Happe, S., Mabry, R., Koch, Z., Iverson, B.L., and Georgiou, G. (2003). Isolation and expression of recombinant antibody fragments to the biological warfare pathogen *Brucella melitensis*. *Journal of Immunological Methods*, **276** pp. 185–96.

Hearty, S., Conroy, P. J., Ayyar, B. V., Byrne, B. and O'Kennedy, R. J. (2010). Surface plasmon resonance for vaccine design and efficacy studies: recent applications and future trends. *Expert Review of Vaccines*, **9** pp. 645–64.

Heidenreich, A., Bellmunt, J., Bolla, M., Joniau, S., Mason, M., Matveev, V., Mottet, N., Schmid, H.P., van der Kwast, T., Wiegel, T., and Zattoni F. (2011). EAU guidelines on prostate cancer.

Part 1: screening, diagnosis, and treatment of clinically localised disease. *European Urology*, **59** pp. 61–71.

Hoffman, R. M. (2011). Clinical practice, Screening for prostate cancer. *The New England Journal of Medicine* 365, pp. 2013-19.

Holland, J.F. Cardinal Manifestations of Cancer. (2003). *IN: Kufe, D.W., Pollock, R.E., Weichselbaum, R.R., Bast, R.C., Gansler, T.S., Holland J.F., and Frei E., (eds) Holland-Frei Cancer Medicine. 6th edition. Hamilton (ON): BC Decker; Chapter 1. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK12489/>*

Holliger, P. and Hudson, P. J. (2005). Engineered antibody fragments and the rise of single domains. *Nature Biotechnology*, **23** (9), pp. 1126–36.

Hoogenboom, H. R. (2005). Selecting and screening recombinant antibody libraries. *Nature Biotechnology*, **23** pp. 1105–16.

Hristodorov, D., Fischer, R., and Linden, L. (2013). With or without sugar? (A)glycosylation of therapeutic antibodies. *Molecular Biotechnology*, **54** (3) pp. 1056-68

Hudson, P. J. (1998). Recombinant antibody fragments. *Current Opinion in Biotechnology*, **9** pp. 395–402.

Hudson, P. J. and Kortt, A. A. (1999). High avidity scFv multimers; diabodies and triabodies. *Journal of Immunological Methods*, **231** (1-2), pp. 177–89.

International Chicken Genome Sequencing Consortium. (2004). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* (London), **432** pp. 695–716.

Irish Cancer Society (2015). Available online at: <http://www.cancer.ie/reduce-your-risk/prostate-cancer> (Accessed November 2015)

Järås, K., Adler, B., Tojo, A., Malm, J., Marko-Varga, G., Lilja, H., and Laurell, T. (2012). Porous silicon antibody microarrays for quantitative analysis: Measurement of free and total PSA in clinical plasma samples. *Clinica Chimica Acta*, **414** pp. 76-84.

Jerónimo, C., Bastian, P.J., Bjartell, A., Carbone, G.M., Catto, J.W., Clark, S.J., Henrique, R., Nelson, W.G., and Shariat, S.F. (2011). Epigenetics in prostate cancer: Biologic and clinical relevance. *European Urology*, **60** (4) pp. 753–66.

Johnstone, P. A. S., Rossi, P. J., Jani, A. B. and Master, V. (2007). “Insignificant” prostate cancer on biopsy: pathologic results from subsequent radical prostatectomy. *Prostate Cancer and Prostatic Diseases*, **10** pp. 237–41.

Joosten, V., Lokman, C., Hondel, C. A. Van Den and Punt, P. J. (2003). The production of antibody fragments and antibody fusion proteins by yeasts and filamentous fungi. *Microbial Cell Factories*, **2** (1) pp. 1-15.

Julien, J. P., Lee, P.S., and Wilson, I.A. (2012). Structural insights into key sites of vulnerability on HIV-1 Env and influenza HA. *Immunological Reviews*, **250** (1), pp. 180–98.

Jung, K., Brux, B., Lein, M., Rudolph, B., Kristiansen, G., Hauptmann, S., Schnorr, D., Loening, S.A., and Sinha, P. (2000). Molecular forms of prostate-specific antigen in malignant and benign prostatic tissue: biochemical and diagnostic implications. *Clinical Chemistry*, **46** pp. 47–54.

Kattla, J.J., Struwe, W.B., Doherty, M., Adamczyk, B., Saldova, R., Rudd, P.M., and Campbell, M.P. (2011) Biologics | Protein Glycosylation. *IN: MurrayMoo-Young (ed.), Comprehensive Biotechnology*, Second Edition, volume 3, pp. 467–86. Elsevier Ltd, Oxford, UK.

Karlsson, F., Borrebaeck, C.A.K., Nilsson, N., and Malmborg-Hager, A. (2003). The Mechanism of Bacterial Infection by Filamentous Phages Involves Molecular Interactions between TolA and Phage Protein 3 Domains. *Journal of Bacteriology*, **185** (8), pp. 2628–2634.

Kehoe, J. W. and Kay, B. K. (2005). Filamentous Phage Display in the New Millennium. *Chemical Reviews*, **3824** pp. 4056–72.

Kendrew, J. C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H., and Phillips, D.C. (1958). A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature (London)*, **181** (4610), pp. 662–66.

Kim, Y., Babnigg, G., Jedrzejczak, R., Eschenfeldt, W.H., Li, H., Maltseva, N., Hatzos-Skintges, C., Gu, M., Makowska-Grzyska, M., Wu, R., An, H., Chhor ,G., and Joachimiak, A. (2011). High-throughput protein purification and quality assessment for crystallization. *Methods*, **55** (1) pp. 12–28.

Kindt, T. J., Goldsby, R.A, Osborne, B.A., and Kuby, J. (2007). *Kuby Immunology*. 6th Edition. New York: W.H. Freeman.

Kirk, D. (1997). MRC study: when to commence treatment in advanced prostate cancer. *Prostate Cancer and Prostatic Diseases*, **1** pp. 11–15.

Knezević ,A., Polasek, O., Gornik, O., Rudan, I., Campbell, H., Hayward, C., Wright, A., Kolcic, I., O'Donoghue, N., Bones, J., Rudd, P.M., and Lauc, G.(2009). Variability, heritability and

environmental determinants of human plasma n-glycome. *Journal of Proteome Research*, **8**(2) pp. 694-701.

Kondo, S., Yagi, H., Kamiya, Y., Ito, A., Kuhara, M., Kudoh, A., Takahashi, N., and Kato, K. (2012). N-Glycosylation Profiles of Chicken Immunoglobulin Y Glycoproteins Expressed by Different Production Vehicles. *Journal of Glycomics and Lipidomics*, [Online] **S5:002** pp. 1–5.

Kretzschmar, T. and Geiser, M. (1995). Evaluation of antibodies fused to minor coat protein III and major coat protein VIII of bacteriophage M13. *Gene*, **155** pp. 61–65.

Kumar, A., Mikolajczyk, S. D., Goel, A. S., Millar, L. S. and Saedi, M. S. (1997). Expression of pro form of prostate-specific antigen by mammalian cells and its conversion to mature, active form by human kallikrein 2. *Cancer Research*, **57** pp. 3111–14.

Kuroda, D., Shirai, H., Jacobson, M.P., and Nakamura, H. (2012). Computer-aided antibody design. *Protein Engineering Design and Selection*, **25** (10), pp.507–522

Küster, B., Wheeler, S.F., Hunter, A.P., Dwek, R.A., and Harvey, D.J. (1997). Sequencing of N-linked oligosaccharides directly from protein gels: in-gel deglycosylation followed by matrix-assisted laser desorption/ionization mass spectrometry and normal-phase high-performance liquid chromatography. *Analytical Biochemistry*, **250** (1) pp. 82-101.

Kyselova, Z., Mechref, Y., Al Bataineh, M.M., Dobrolecki, L.E., Hickey, R.J., Vinson, J., Sweeney, C.J., and Novotny, M.V. (2007). Alterations in the Serum Glycome Due to Metastatic Prostate Cancer *Journal of Proteome Research*, **6** (5) pp. 1822–32.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)*, **227** pp.680 - 85

Lauc, G., Essafi, A., Huffman, J.E., Hayward, C., Knežević, A., Kattla, J.J., Polašek, O., Gornik, O., Vitart, V., Abrahams, J.L., Pučić, M., Novokmet, M., Redžić, I., Campbell, S., Wild, S.H., Borovečki, F., Wang, W., Kolčić, I., Zgaga, L., Gyllensten, U., Wilson, J.F., Wright, A.F., Hastie, N.D., Campbell, H., Rudd, P.M., and Rudan, I. (2010) Genomics meets glycomics-the first GWAS study of human N-glycome identifies HNF1A as a master regulator of plasma protein fucosylation. *PLoS Genetics*, **6** (12), pp. 1–14.

Lauc, G., Rudan, I., Campbell, H., and Rudd, P.M. (2010). Complex genetic regulation of protein glycosylation. *Molecular BioSystems*, **6** (2), pp. 329–35.

Lee, P. S. and Wilson, I. A. (2015) 'Structural Characterization of Viral Epitopes Recognized by Broadly Cross-Reactive Antibodies'. *IN: Oldstone, M.B.A. and Compans, R.W. (eds) Influenza Pathogenesis and Control - Volume II.* Switzerland: Springer International Publishing, pp. 323–41.

Leenaars, M. and Hendriksen, C. F. M. (2005). Critical Steps in the Production of Polyclonal and Monoclonal Antibodies: Evaluation and Recommendations. *ILAR Journal*, **46** (3), pp. 269–79.

Lefranc, M.P. (2014). Immunoglobulins: 25 years of immunoinformatics and IMGT-ONTOLOGY. *Biomolecules*, **4**(4) pp. 1102-39.

Leonard, P., Hearty, S. and O’Kennedy, R. (2011). Measuring protein-protein interactions using Biacore *IN; (Walls, D and Loughran, S.T (eds) Protein Chromatography -Methods in Molecular Biology.* Clifton, N.J: Humana Press, pp. 403–18.

Lilja, H. (1985). A ' kallikrein-like' serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. *The Journal of Clinical Investigation*, **76** pp. 1899–1903.

Lilja, H., Christensson, A., Dahlén, U., Matikainen, M.T., Nilsson, O., Pettersson, K., and Lövgren, T. (1991). Prostate-specific antigen in serum occurs predominantly in complex with alpha 1-anti-chymotrypsin. *Clinical Chemistry*, **37** pp. 1618–25.

Lilja, H., Oldbring, J., Rannevik, G. and Laurell, C. B. (1987). Seminal vesicle-secreted proteins and their reactions during gelation and liquefaction of human semen. *The Journal of Clinical Investigation*, **80** pp. 281-85.

Lilja, H., Ulmert, D. and Vickers, A. J. (2008). Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. *Nature Reviews Cancer*, **8** pp. 268–78.

Liu, L. (2015). Antibody Glycosylation and Its Impact on the Pharmacokinetics and Pharmacodynamics of Monoclonal Antibodies and Fc-Fusion Proteins. *Journal of Pharmaceutical Sciences*, **104**(6), pp. 1866-84

Liu, Y., Yen, H.Y., Chen, C.Y., Chen, C.H., Cheng, P.F., Juan, Y.H., Chen, C.H., Khoo, K.H., Yu. C.J., Yang, P.C., Hsu, T.L. and Wong, C.H. (2011). Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. *Proceedings of the National Academy of Sciences (USA)*, **108** (28), pp. 11332–37.

Loeb, L. A., Springgate, C. F. and Battula, N. (1974). Errors in DNA Replication as a Basis of Malignant Changes. *Cancer Research*, **34** pp. 2311–21.

Loeb, S., Sanda, M.G., Broyles, D.L., Shin, S.S., Bangma, C.H., Wei, J.T., Partin, A.W., Klee, G.G., Slawin, K.M., Marks, L.S., van Schaik, R.H., Chan, D.W., Sokoll, L.J., Cruz, A.B., Mizrahi, I.A., and Catalona, W.J. (2015). The Prostate Health Index Selectively Identifies Clinically Significant Prostate Cancer. *The Journal of Urology*, **193** (4), pp. 1163–69.

Luft, J. R., Newman, J., and Snell, E.H. (2014). Crystallization screening: The influence of history on current practice. *Acta Crystallographica Section F: Structural Biology Communications*, **70** (7), pp. 835–53.

Lundwall, A. and Lilja, H. (1987). Molecular cloning of human prostate specific antigen cDNA. *FEBS Letters*, **214** pp. 317–22.

Lyons, J. A., Aragão, D., Slattey, O., Pisiakov, A.V., Soulimane, T., and Caffrey, M. (2012). Structural insights into electron transfer in caa3-type cytochrome oxidase. *Nature* (London), **487** (7408) pp.514–18.

Mahmood, T. and Yang, P. C. (2012). Western blot: Technique, theory, and trouble shooting. *North American Journal of Medical Sciences*, **4** (9), pp. 429–34.

Marek, K. W., Vijay, I.K., and Marth, J.D. (1999). A recessive deletion in the GlcNAc-1-phosphotransferase gene results in peri-implantation embryonic lethality. *Glycobiology*, **9** (11),pp. 1263–71.

Mariño, K., Bones, J., Kattla, J.J. and Rudd, P.M. (2010). A systematic approach to protein glycosylation analysis: a path through the maze. *Nature Chemical Biology*, **6** (10), pp. 713–23.

Mariño, K., Saldoval, R., Adamczyk, B., and Rudd, P.M. (2012) Changes in Serum N-Glycosylation Profiles: Functional Significance and Potential for Diagnostics. *IV: Carbohydrate Chemistry: Chemical and Biological Approaches*; Rauter, A. P., (Ed.); RSC Publishing: Cambridge, U.K., Vol. 37.

Matsumoto, K., Yokote, H., Arao, T., Maegawa, M., Tanaka, K., Fujita, Y., Shimizu, C., Hanafusa, T., Fujiwara, Y., and Nishio, K. (2008). N-Glycan fucosylation of epidermal growth factor

receptor modulates receptor activity and sensitivity to epidermal growth factor receptor tyrosine kinase inhibitor. *Cancer Science*, **99** (8), pp. 1611–17.

McCafferty, J., Griffiths, A. D., Winter, G. and Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature (London)*, **348** pp. 552–54.

McCullough, K. C. and Summerfield, A. (2005). Basic concepts of immune response and defense development. *ILAR Journal*, **46** (3), pp. 230–40.

McPherson, A. (1985) Use of polyethylene glycol in the crystallization of macromolecules. *Methods in Enzymology*, **114** pp. 120–25.

Meany, D. L., Hackler, L Jr., Zhang, H., and Chan, D.W. (2011). Tyramide signal amplification for antibody-overlay lectin microarray: a strategy to improve the sensitivity of targeted glycan profiling. *Journal of Proteome Research*, **10** (3), pp. 1425–31.

Mikolajczyk, S. D. and Rittenhouse, H. G. (2003). Pro PSA: a more cancer specific-form of prostate specific antigen for the early detection of prostate cancer. *The Keio Journal of Medicine*, **52** pp. 86–91.

Mikolajczyk, S. D., Grauer, L.S., Millar, L.S., Hill, T.M., Kumar, A., Rittenhouse, H.G., Wolfert, R.L., and Saedi, M.S., (1997). A precursor form of PSA (pPSA) is a component of the free PSA in prostate cancer serum. *Urology*, **50** pp. 710–14.

Mikolajczyk, S. D., Marker, K.M., Millar, L.S., Kumar, A., Saedi, M.S., Payne, J.K., Evans, C.L., Gasior, C.L., Linton, H.J., Carpenter, P., and Rittenhouse, H.G. (2001). A truncated precursor form of prostate-specific antigen is a more specific serum marker of prostate cancer. *Cancer Research*, **61** pp. 6958–63.

Mikolajczyk, S. D., Marks, L. S., Partin, A. W. and Rittenhouse, H. G. (2002). Free prostate-specific antigen in serum is becoming more complex. *Urology*, **59** pp.797–802.

Mikolajczyk, S. D., Millar, L.S., Wang, T.J., Rittenhouse, H.G., Wolfert, R.L., Marks, L.S., Song, W., Wheeler, T.M., and Slawin, K.M. (2000). “BPSA,” a specific molecular form of free prostate-specific antigen, is found predominantly in the transition zone of patients with nodular benign prostatic hyperplasia. *Urology*, **55** pp. 41–45.

Mikolajczyk, S. D., Song, Y., Wong, J. R., Matson, R. S. and Rittenhouse, H. G. (2004). Are multiple markers the future of prostate cancer diagnostics? *Clinical Biochemistry*, **37** pp. 519–28.

Millán Martín, S., Delporte, C., Farrell, A., Navas Iglesias, N., McLoughlin, N.,and Bones, J. (2015). Comparative analysis of monoclonal antibody N-glycosylation using stable isotope labelling and UPLC-fluorescence-MS. *Analyst*, **40** (5) pp. 1442-47.

Moffatt, B. A. and Studier, F. W. (1987) T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* , **49** (2), pp. 221–27

Morgan, T., Palapattu, G., and Wei, J. (2015). Screening for Prostate Cancer-Beyond Total PSA, Utilization of Novel Biomarkers. *Current Urology Reports*, **16** (9), pp. 537.

Muller, B. H., Savatier, A., L'Hostis, G., Costa, N., Bossus, M., Michel, S., Ott, C., Becquart, L., Ruffion, A., Stura, E.A., and Ducancel F. (2011). *In vitro* affinity maturation of an anti-PSA antibody for prostate cancer diagnostic assay. *Journal of Molecular Biology*, **414** pp. 545–62.

National Cancer Registry (2015). Available online at:

<http://www.ncri.ie/research/projects/cost-effectiveness-psa-testing-secondary-prevention-prostate-cancer> (Accessed November 2015).

Narat, M. (2003). Production of Antibodies in Chickens. *Food Technology and Biotechnology*, **41** (3), pp. 259–67.

Nelson, P. S., Gan, L., Ferguson, C., Moss, P., Gelinas, R., Hood, L., and Wang, K. (1999). Molecular cloning and characterization of prostase, an androgen-regulated serine protease with prostate-restricted expression. *Proceedings of the National Academy of Sciences (USA)*, **96** pp. 3114–19.

North, B., Lehmann, A., and Dunbrack, R.L. Jr. (2011). A new clustering of antibody CDR loop conformations. *Journal of Molecular Biology*, **406** (2), pp. 228–56.

Nurmikko, P., Pettersson, K., Piironen, T., Hugosson, J. and Lilja, H. (2001). Discrimination of prostate cancer from benign disease by plasma measurement of intact, free prostate-specific antigen lacking an internal cleavage site at Lys145-Lys146. *Clinical Chemistry*, **47** pp. 1415–23.

Ohyama, C., Hosono, M., Nitta, K., Oh-eda, M., Yoshikawa, K., Habuchi, T., Arai, Y., and Fukuda, M. (2004). Carbohydrate structure and differential binding of prostate specific antigen to *Maackia amurensis* lectin between prostate cancer and benign prostate hypertrophy. *Glycobiology*, **14** pp. 671–79.

Okihara, K., Cheli, C.D., Partin, A.W., Fritche, H.A., Chan, D.W., Sokoll, L.J., Brawer, M.K., Schwartz, M.K., Vessella, R.L., Loughlin, K.R., Johnston, D.A., and Babaian, R.J. (2002). Comparative analysis of complexed prostate specific antigen, free prostate specific antigen and their ratio in detecting prostate cancer. *The Journal of Urology*, **167** pp. 2017–24.

- Oon, S. F.**, Pennington, S.R., Fitzpatrick, J.M., and Watson, R.W. (2011). Biomarker research in prostate cancer-towards utility, not futility. *Nature Reviews Urology*, **8** (3), pp. 131–38.
- Özen, H.** and Sözen, S. (2006). PSA Isoforms in Prostate Cancer Detection. *European Urology Supplements*, **5** pp. 495–99.
- Padler-Karavani, V.**, Yu, H., Cao, H., Chokhawala, H., Karp, F., Varki, N., Chen, X., and Varki, A. (2008). Diversity in specificity, abundance, and composition of anti-Neu5Gc antibodies in normal humans: Potential implications for disease. *Glycobiology*, **18** (10), pp. 818–30.
- Partin, A. W.**, Brawer, M.K., Bartsch, G., Horninger, W., Taneja, S.S., Lepor, H., Babaian, R., Childs, S.J., Stamey, T., Fritsche, H.A., Sokoll, L., Chan, D.W., Thiel, R.P., and Cheli, C.D. (2003). Complexed prostate specific antigen improves specificity for prostate cancer detection: results of a prospective multicenter clinical trial. *The Journal of Urology*, **170** pp. 1787–91.
- Peracaula, R.**, Tabarés, G., Royle, L., Harvey, D.J., Dwek, R.A., Rudd, P.M., and de Llorens, R. (2003). Altered glycosylation pattern allows the distinction between prostate-specific antigen (PSA) from normal and tumor origins. *Glycobiology*, **13** (6), pp. 457–70.
- Peter, J.**, Unverzagt, C., Krogh, T. N., Vorm, O. and Hoesel, W. (2001). Identification of precursor forms of free prostate-specific antigen in serum of prostate cancer patients by immunosorption and mass spectrometry. *Cancer Research*, **61** pp. 957–62.
- Petrescu, A.J.**, Milac, A.L., Petrescu, S.M., Dwek, R.A., and Wormald, M.R. (2004). Statistical analysis of the protein environment of N-glycosylation sites: Implications for occupancy, structure, and folding. *Glycobiology*, **14** (2), pp.103–114.

- Pinsky, P. F.,** Kramer, B.S., Crawford, E.D., Grubb, R.L., Urban, D.A., Andriole, G.L., Chia, D., Levin, D.L., and Gohagan, J.K., (2006). Prostate volume and prostate-specific antigen levels in men enrolled in a large screening trial. *Urology*, **68** pp. 352–56.
- Qi, H.,** Lu, H., Qiu, H.-J., Petrenko, V. and Liu, A. (2012). Phagemid vectors for phage display: properties, characteristics and construction. *Journal of Molecular Biology*, **417** pp. 129–43.
- Rader, C.** and Barbas, C. F. (1997). Phage display of combinatorial antibody libraries. *Current Opinion in Biotechnology*, **8** pp. 503–08.
- Raghunathan, G.,** Smart, J., Williams, J., and Almagro, J.C. (2012). Antigen-binding site anatomy and somatic mutations in antibodies that recognize different types of antigens. *Journal of Molecular Recognition*, **25** (3), pp. 103–13.
- Raju, T. S.,** Briggs, J.B., Borge, S.M., and Jones, A.J. (2000). Species-specific variation in glycosylation of IgG: evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics. *Glycobiology*, **10**(5) pp. 477-86.
- Reynaud, C. A.,** Anquez, V., and Weill, J.C. (1991). The chicken D locus and its contribution to the immunoglobulin heavy chain repertoire. *European Journal of Immunology*, **21** (11), pp. 2661–70.
- Rossmann, M. G.,** Arnold, E., Erickson, J.W., Frankenberger ,E.A., Griffith, J.P., Hecht, H.J., Johnson, J.E., Kamer, G., Luo, M., Mosser, A.G., Rueckert, R.R., Sherry, B., and Vriend, G. (1985). Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature*, **317** (6033), pp. 145–53.

Royle, L., Campbell, M.P., Radcliffe, C.M., White, D.M., Harvey, D.J., Abrahams, J.L., Kim, Y.G., Henry, G.W., Shadick, N.A., Weinblatt, M.E., Lee, D.M., Rudd, P.M., and Dwek, R.A. (2008). HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software. *Analytical Biochemistry*, **376** pp. 1–12.

Royle, L., Radcliffe, C.M., Dwek, R.A., and Rudd, P.M. (2006). Detailed structural analysis of N-glycans released from glycoproteins in SDS-PAGE gel bands using HPLC combined with exoglycosidase array digestions. *Methods in Molecular Biology*, **347** pp. 125-43

Saldiva, R., Asadi Shehni, A., Haakensen, V.D., Steinfeld, I., Hilliard, M., Kifer, I., Helland, A., Yakhini, Z., Børresen-Dale, A.L., and Rudd, P.M. (2014). Association of N-glycosylation with breast carcinoma and systemic features using high-resolution quantitative UPLC. *Journal of Proteome Research*, **13**(5) pp. 2314-27.

Saldiva, R., Fan, Y., Fitzpatrick, J.M., Watson, R.W., and Rudd, P.M. (2011). Core fucosylation and alpha2-3 sialylation in serum N-glycome is significantly increased in prostate cancer comparing to benign prostate hyperplasia. *Glycobiology*, **21** (2), pp. 195–205.

Scapin, G. (2013) 'Molecular replacement then and now'. *Acta Crystallographica Section D: Biological Crystallography*, pp. 2266–75.

Schägger, H. and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry*, **166** (2), pp. 368–79.

Schrodinger, L.L.C., (2010). The PyMOL Molecular Graphics System

Sela-Culang, I., Vered Kunik, V., and Ofran, Y. (2013). The structural basis of antibody-antigen recognition. *Frontiers in Immunology*, **4** pp. 302.

Sethi, M.K., Thaysen-Andersen, M., Smith, J.T., Baker, M.S., Packer, N.H., Hancock, W.S., and Fanayan, S. (2014). Comparative N-glycan profiling of colorectal cancer cell lines reveals unique bisecting GlcNAc and α -2,3-linked sialic acid determinants are associated with membrane proteins of the more metastatic/aggressive cell lines. *Journal of Proteome Research*, **13**(1) pp. 277-88.

Shi, Y. (2014) Review A Glimpse of Structural Biology through X-Ray Crystallography. *Cell*, **159** (5), pp. 995–1014.

Shih, H. H., Tu, C., Cao, W., Klein, A., Ramsey, R., Fennell, B.J., Lambert, M., Ní Shúilleabháin, D., Autin, B., Kouranova, E., Laxmanan, S., Braithwaite, S., Wu, L., Ait-Zahra, M., Milici, A.J., Dumin, J.A., LaVallie, E.R., Arai, M., Corcoran, C., Paulsen, J.E., Gill, D., Cunningham, O., Bard, J., Mosyak, L., and Finlay, W.J. (2012). An ultra-specific avian antibody to phosphorylated tau protein reveals a unique mechanism for phosphoepitope recognition. *The Journal of Biological Chemistry*, **287** (53), pp. 44425–34.

Siegel, R., Naishadham, D. & Jemal, A. Cancer statistics, 2012. *CA: A Cancer Journal for Clinicians*, **62** pp. 10–29.

Sigma Aldrich (2015). Available online at:

<http://www.sigmaaldrich.com/catalog/product/sigma/r2020?lang=en®ion=IE> (Accessed June 2015).

Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, **228** pp. 1315–17.

- Smyth, M. S.** and Martin, J. H. (2000). X ray crystallography. *Molecular Pathology*, **53** (1), pp. 8–14.
- Sokoll, L. J.,** Sanda, M.G., Feng, Z., Kagan, J., Mizrahi, I.A., Broyles, D.L., Partin, A.W., Srivastava, S., Thompson, I.M., Wei, J.T., Zhang, Z., and Chan, D.W. (2010). A Prospective, Multicenter, National Cancer Institute Early Detection Research Network study of [-2]proPSA: Improving prostate cancer detection and correlating with cancer aggressiveness. *Cancer Epidemiology, Biomarkers & Prevention*, **19** pp. 1193–1200.
- Soulimane, T.,** Buse, G., Bourenkov, G.P., Bartunik, H.D., Huber, R., and Than, M.E. (2000). Structure and mechanism of the aberrant ba(3)-cytochrome c oxidase from thermus thermophilus. *EMBO Journal*, **19** (8), pp. 1766–76.
- Spillner, E.,** Braren, I., Greunke, K., Seismann, H., Blank, S., and du Plessis, D. (2012). Avian IgY antibodies and their recombinant equivalents in research, diagnostics and therapy. *Biologicals: journal of the International Association of Biological Standardization*, **40** (5), pp. 313–22.
- Spriestersbach, A.,** Kubicek, J., Schäfer, F., Block, H., and Maertens, B. (2015). 'Purification of His-Tagged Proteins'. *Methods in Enzymology*, **559** pp. 1-15.
- Steyerberg, E. W.,** Vickers, A.J., Cook, N.R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M.J., and Kattan, M.W. (2010). Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*, **21** (1) pp. 128–38.
- Stöckmann, H.,** Adamczyk, B., Hayes, J., and Rudd, P.M. (2013). Automated, high-throughput IgG-antibody glycoprofiling platform. *Analytical Chemistry*, **85** (18), pp. 8841–49.

Stöckmann, H., O'Flaherty, R., Adamczyk, B., Saldova, R., and Rudd, P.M. (2015). Automated, high-throughput serum glycoprofiling platform. *Integrative Biology Online*, Advance Article.

Strebhardt, K. and Ullrich, A. (2008). Paul Ehrlich's magic bullet concept: 100 years of progress. *Nature Reviews Cancer*, **8** (6), pp. 473–80.

Structural Genomics Consortium, China Structural Genomics Consortium, Northeast Structural Genomics Consortium, **Gräslund, S.**, Nordlund, P., Weigelt, J., Hallberg, B.M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., Ming, J., dhe-Paganon, S., Park, H.W., Savchenko, A., Yee, A., Edwards, A., Vincentelli, R., Cambillau, C., Kim, R., Kim, S.H., Rao, Z., Shi, Y., Terwilliger, T.C., Kim, C.Y., Hung, L.W., Waldo, G.S., Peleg, Y., Albeck, S., Unger, T., Dym, O., Prilusky, J., Sussman, J.L., Stevens, R.C., Lesley, S.A., Wilson, I.A., Joachimiak, A., Collart, F., Dementieva, I., Donnelly, M.I., Eschenfeldt, W.H., Kim, Y., Stols, L., Wu, R., Zhou, M., Burley, S.K., Emtage, J.S., Sauder, J.M., Thompson, D., Bain, K., Luz, J., Gheyi, T., Zhang, F., Atwell, S., Almo, S.C., Bonanno, J.B., Fiser, A., Swaminathan, S., Studier, F.W., Chance, M.R., Sali, A., Acton, T.B., Xiao, R., Zhao, L., Ma, L.C., Hunt, J.F., Tong, L., Cunningham, K., Inouye, M., Anderson, S., Janjua, H., Shastry, R., Ho, C.K., Wang, D., Wang, H., Jiang, M., Montelione, G.T., Stuart, D.I., Owens, R.J., Daenke, S., Schütz, A., Heinemann, U., Yokoyama, S., Büsow, K., and Gunsalus, K.C. (2008). Protein production and purification. *Nature Methods*, **5** (2), pp. 135-46.

Stanfield, R.L., Fieser, T.M., Lerner, R.A., and Wilson, I.A. (1990). Crystal structures of an antibody to a peptide and its complex with peptide antigen at 2.8 Å. *Science* (New York, N.Y.), **248** (4956), pp.712–719.

- Stura, E.,** Muller, B.H., Bossus, M., Michel, S., Jolivet-Reynaud, C., and Ducancel, F. (2011). Crystal structure of human prostate-specific antigen in a sandwich antibody complex. *Journal of Molecular Biology*, **414** pp. 530–44.
- Sulkowski, E.** (1985) Purification of proteins by IMAC. *Trends in Biotechnology*, **3** (1) pp.1–7.
- Suzuki, N.** and Lee, Y. C. (2004). Site-specific N-glycosylation of chicken serum IgG. *Glycobiology*, **14**(3) pp. 275-92.
- Szekrényes, A.,** Partyka, J., Varadi, C., Krenkova, J., Foret, F., and Guttman, A. (2015). Sample Preparation for N-Glycosylation Analysis of Therapeutic Monoclonal Antibodies by Electrophoresis. *Methods in Molecular Biology*, **1274** pp. 183-95.
- Tabarés, G.,** Radcliffe C.M., Barrabés S., Ramírez M., Aleixandre R.N., Hoesel W., Dwek R.A., Rudd P.M., Peracaula R., and de Llorens R. (2006). Different glycan structures in prostate-specific antigen from prostate cancer sera in relation to seminal plasma PSA. *Glycobiology*, **16** pp. 132–45.
- Tabarés, G.,** Radcliffe C.M., Barrabés S., Ramírez M., Aleixandre R.N., Hoesel W., Dwek R.A., Rudd P.M., Peracaula R., and de Llorens R. (2006). Different glycan structures in prostate-specific antigen from prostate cancer sera in relation to seminal plasma PSA. *Glycobiology*, **16** (2) pp. 132–45.
- Tajiri, M.,** Ohyama, C. and Wada, Y. (2008). Oligosaccharide profiles of the prostate specific antigen in free and complexed forms from the prostate cancer patient serum and in seminal plasma: a glycopeptide approach. *Glycobiology*, **18** (1) pp. 2–8.

Takayama, T. K., Fujikawa, K., and Davie, E. W. (1997). Characterisation of the precursor of prostate-specific antigen. Activation by trypsin and by human glandular kallikrein. *The Journal of Biological Chemistry*, **272** pp. 21582–88.

Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., and Muchmore, E. (2003). Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proceedings of the National Academy of Sciences (USA)*, **100** (21), pp. 12045–50.

Tharmalingam, T., Mariño, K., and Rudd, P.M. (2010). Platform technology to identify potential disease markers and establish heritability and environmental determinants of the human serum N-glycome. *Carbohydrate Research*, **345** (10), pp. 1280–82.

Thompson, I. M., Ankerst, D.P., Chi, C., Lucia, M.S., Goodman, P.J., Crowley, J.J., Parnes, H.L., and Coltman, C.A. Jr. (2005). Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/ml or lower. *The Journal of the American Medical Association*, **294** pp. 66–70.

Thompson, R., Creavin, A., O'Connell, M., O'Connor, B., and Clarke, P. (2011). Optimization of the enzyme-linked lectin assay for enhanced glycoprotein and glycoconjugate analysis. *Analytical Biochemistry*, **413** (2), pp. 114–22.

Van Beers, M. M. C. & Bardor, M. (2012) Minimizing immunogenicity of biopharmaceuticals by controlling critical quality attributes of proteins. *Biotechnology Journal*, **7** (12) pp. 1473–84.

Van Den Broeck, T., Joniau, S., Clinckemalie, L., Helsen, C., Prekovic, S., Spans, L., Tosco, L., Van Poppel, H., and Claessens, F. (2014). The role of single nucleotide polymorphisms in predicting prostate cancer risk and therapeutic decision making. *BioMed Research International* Volume: 2014, 627510, pp. 1-16

Varki, A., Kannagi, R., and Toole B. P., (2009). Glycosylation changes in cancer *IN*; Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W., and Etzler, M.E (eds) *Essentials of Glycobiology*, 2nd Edition. New York: Cold Spring Harbor Laboratory Press, Chapter 44.

Venderbos, L. D. F., Bokhorst, L.P., Bangma, C.H., and Roobol, M.J. (2013). Active surveillance: oncologic outcome. *Current Opinion in Urology*, **23** (3), pp. 268–72.

Venetz, D., Hess, C., Lin, C.W., Aebi, M., and Neri, D. (2015). Glycosylation profiles determine extravasation and disease-targeting properties of armed antibodies. *Proceedings of the National Academy of Sciences (USA)*, **112**(7), pp. 2000-5.

Vickers, A. J., Roobol, M.J., and Lilja, H. (2012). Screening for prostate cancer: early detection or overdetetection? *Annual Review of Medicine*, **63** pp. 161–70.

Wan, T., Beavil, R.L., Fabiane, S.M., Beavil, A.J., Sohi, M.K., Keown, M., Young, R.J., Henry, A.J., Owens, R.J., Gould, H.J. and Sutton, B.J. (2002). The crystal structure of IgE Fc reveals an asymmetrically bent conformation. *Nature immunology*, **3** (7), pp.681–686.

Wang, S., Zheng, C., Liu, Y., Zheng, H., and Wang, Z. (2008). Construction of multiform scFv antibodies using linker peptide. *Journal of Genetics and Genomics*, **35** (5), pp. 313–16.

Watson, J. D. and Crick, F. H. C. (1953) Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature (London)*, **171** (4356), pp. 737–38.

Wilson, W. D. (2002). Analyzing biomolecular interactions. *Science*, **295** (5562), pp. 2103–05.

Wlodawer, A., Minor, W., Dauter, Z., and Jaskolski, M. (2008). Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS Journal*, **275** (1) pp. 1–21.

Wlodawer, A., Minor, W., Dauter, Z., and Jaskolski, M. (2013). Protein crystallography for aspiring crystallographers or how to avoid pitfalls and traps in macromolecular structure determination. *FEBS Journal*, **280** (22), pp. 5705–36.

World Cancer Research Fund (2015). Available online at: <http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/prostate-cancer-statistics> (Accessed November 2015)

Wormald, M.R., Petrescu, A.J., Pao, Y.L., Glithero, A., Elliott, T., and Dwek, R.A. (2002). Conformational studies of oligosaccharides and glycopeptides: Complementarity of NMR, X-ray crystallography, and molecular modelling. *Chemical Reviews*, **102** (2), pp.371–386.

Wu, A. M., Lisowska, E., Duk, M., and Yang, Z. (2009). Lectins as tools in glycoconjugate research. *Glycoconjugate Journal*, **26** (8), pp. 899–913.

Wu, L., Oficjalska, K., Lambert, M., Fennell, B.J., Darmanin-Sheehan, A., Ní Shúilleabháin, D., Autin, B., Cummins, E., Tchistiakova, L., Bloom, L., Paulsen, J., Gill, D., Cunningham, O., and Finlay, W.J. (2012). Fundamental characteristics of the immunoglobulin VH repertoire of chickens in comparison with those of humans, mice, and camelids. *Journal of Immunology*, **188** (1), pp. 322–33.

Wu, T. T. and Kabat, E. A. (1970). An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *The Journal of Experimental Medicine*, **132** (2), pp. 211–50.

- Wukovitz, S. W.** and Yeates, T. O. (1995). Why protein crystals favour some space-groups over others. *Nature Structural Biology*, **2** (12), pp. 1062–67.
- Yi, C.**, Ruan, C.P., Wang, H., Xu, X.Y., Zhao, Y.P., Fang, M., Ji, J., Gu, X., and Gao, C.F. (2014). Function characterization of a glyco-engineered anti-EGFR monoclonal antibody *cetuximab* *in vitro*. *Acta Pharmacologica Sinica*, **35** (11), pp. 1439–46.
- Young, C. Y.**, Andrews, P. E., Montgomery, B. T. and Tindall, D. J. (1992). Tissue-specific and hormonal regulation of human prostate-specific glandular kallikrein. *Biochemistry*, **31** pp. 818–24.
- Yousef, G. M.** and Diamandis, E. P. (2001). The new human tissue kallikrein gene family: structure, function, and association to disease. *Endocrine Reviews*, **22** pp. 184–204.
- Yousef, G. M.**, Luo, L. Y. and Diamandis, E. P. (1999). Identification of novel human 'kallikrein-like' genes on chromosome 19q13.3-q13.4. *Anticancer Research*, **19** pp. 2843–52.
- Yu, F.**, Persson, B, Lofas, S, and Knoll, W. (2004). Surface Plasmon Fluorescence Immunoassay of Free Prostate-Specific Antigen in Human Plasma at the Femtomolar Level. *Analytical Chemistry* **76**, pp. 6765-70.
- Zhang, W.** and Leinonen, J. (1995). Purification and Characterization of Different Molecular Forms of Prostate-Specific Antigen in Human Seminal Fluid. *Clinical Chemistry*, **41** pp. 1567–73.
- Zhu, K.**, Dietrich, R., Didier, A., Doyscher, D., and Märtilbauer, E. (2014). Recent Developments in Antibody-Based Assays for the Detection of Bacterial Toxins. *Toxins*, **6** (4), pp. 1325–48.