

Aptamer selection against GFR α 1 for its application in the prognosis of breast cancer



UNIVERSITY *of the*
WESTERN CAPE

Lauren Taryn Swartz
UNIVERSITY *of the*
WESTERN CAPE

A thesis submitted in fulfilment of the requirements for the degree of
Philosophiae Doctor in the Faculty of Science, Department of Biotechnology,
University of the Western Cape.

Supervisor: Prof Mervin Meyer

Co-supervisor: Prof Ciara O' Sullivan

August 2019

Abstract

Breast cancer is the second most common cancer amongst South African women. Despite ongoing efforts to combat breast cancer, current prognostic and/or therapeutic monitoring methods are limited since very little improvement, in the rate of long term recurrence of breast cancer, has been observed. Considering this, developing novel strategies to detect breast cancer recurrence – at an early onset – is crucial for monitoring the disease and potentially preventing disease progression. Methods currently used for the detection of BC are costly and can also be very uncomfortable for the patient. These methods are also too costly to use as a routine test, following surgery or treatment to assess disease progression. Thus, developing a cost-effective detection method appears to be an appealing alternative. Serum/blood-based biomarkers are ideal targets for the development of low cost detection assays. Two candidate biomarkers, unique ligand binding protein 2 (ULBP2) and glial cell line-derived neurotrophic factor family receptor alpha 1 (GFR α 1) were identified using bioinformatics and proteomics, respectively. These biomarkers have demonstrated to be useful prognostic biomarkers for breast cancer. The selection of aptamers against these biomarkers can facilitate the development of cost-effective detection methods. Aptamers are short DNA or RNA oligonucleotides that have very high affinity and specificity for its targets and can potentially replace antibodies as tools for molecular recognition in detection systems, such as the enzyme-linked immunosorbent assay (ELISA), lateral flow assays and electrochemical biosensors.

The aim of this study was to develop DNA aptamers against ULBP2 and GFR α 1, respectively, for its application in the prognosis of breast cancer. Candidate DNA aptamer libraries against ULBP2 and GFR α 1 were developed using systematic evolution of ligands by exponential enrichment (SELEX). The evolution of the candidate DNA aptamer libraries was monitored using polymerase chain reaction (PCR). Next generation sequencing (NGS) was then employed

to isolate and identify six candidate aptamer sequences. The electrophoretic mobility shift assay (EMSA), surface plasmon resonance (SPR), microscale thermophoresis (MST) and flow cytometry was used to investigate the binding affinity and specificity of the candidate aptamers towards the target biomarker. The best characterized candidate aptamer, G_Apta1, showed a K_D value in the nano-molar range and further demonstrated binding specificity towards the MCF7 breast cancer cell line, relative to a panel of human cell lines. This study showed that using *in vitro* SELEX, G_Apta1, an aptamer against GFR α 1, was successfully selected, identified and characterized. G_Apta1 also demonstrated the ability to bind, with specificity to the MCF7 breast cancer cell line and should further be investigated for its application in the development of prognostic detection systems for breast cancer.



Keywords

Aptamers

Biomarkers

Breast cancer

GFR α 1

SELEX

ULBP2



UNIVERSITY *of the*
WESTERN CAPE

Declaration

I declare that **Aptamer selection against GFR α 1 for its application in the prognosis of breast cancer** is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged by complete references.



Miss Lauren Taryn Swartz

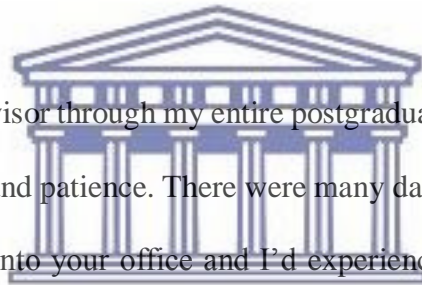
Signed.....*Swartz*.....

Date: 12 August 2019

Acknowledgements

God, words fail me. All that I am and have the ability to do was born the moment you breathed your breath of life into me. Before the world began you already saw this day and oh, what a glorious day it is. I will spend the rest of my life thanking You.

Shucks! Then there's my *family*. My first and strongest imagery of support came from you. Thank you for believing in me. Thank you for encouraging me. For praying me through every season of my life. Your presence in my life is invaluable. Thank you, daddy, mommy, Godfrey, Rochelle, Ashson, Leticia, Michaela, Seth, Abigail and Zachary. And thank you to my extended family, julle is yster. A special thanks to Ashson and Leticia for always opening your home as a quiet place for me to work and exploit your Wi-Fi subscription haha.

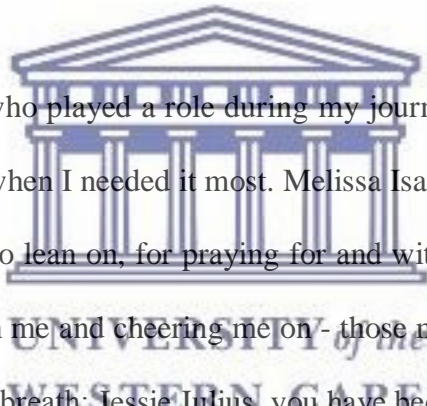


Prof Mervin Meyer, my supervisor through my entire postgraduate journey. Man, I am grateful for your academic leadership and patience. There were many days I expected you to wring my neck haha, but then I'd walk into your office and I'd experience a consistently strong leader and mentor – who always carried a smile even when days were pretty gloomy for me. I hope to be even half the scientist you are. To my co-supervisor, Prof Ciara O' Sullivan, thank you for your fervent leadership and all the opportunities I was afforded, simply because you believed in my ability as a scientist. Thank you for extending an invitation and then welcoming me, for 6 months, to the Universitat Rovira I Virgili. I learned so much under your guidance. Thank you.

Dr. Marketa Svobodova, during my 6-month research visit you diligently introduced me to Aptamer technology. You were a tremendous help for a very crucial part of my PhD project. I am now regarded as an aptamer expert back home ☺, thank you for empowering me to be able to teach others. Not only did you serve me well academically, but you also befriended me –

which made a world of difference in a city where very few people spoke English. And in the same breath, I'd like to express my gratitude to the Nanobiotechnology and Bioanalysis research group.

To all my fellow students in the Nanotechnology Innovation Centre Biolabels research group, I have been stretched as a scientist by each of you. Thank you for being folk that I could laugh with, share frustrations with, be encouraged by and spend my 7hr day with. I'd like to extend a special thanks to Dr. Nicole Sibuye, Dr. Mustafah Drah, Dr. Abdulrahman Elbagory (otherwise known as Abdulla ☺), Riziki Martin and Boitumelo Maobelo. Thank you to the Biotechnology Department.



I am so grateful to everyone who played a role during my journey. Tarryn Oliver, thank you for encouraging me in a time when I needed it most. Melissa Isaacs, thank you for listening to me nag, for being a shoulder to lean on, for praying for and with me. Megan Schilder, thank you for always checking up on me and cheering me on - those moments are deeply treasured. To my boyfriend, wow...deep breath; Jessie Julius, you have been a true champ, a shoulder to lean on, a sound board, a disciplinarian and my best friend. Your support and sweet notes really motivated me towards the end.

Thank you to my church family. I often, literally, felt your prayers strengthening me. Thank you for celebrating me – it means more than I can put into words. Thank you to the National Research Foundation and DST/Mintek NIC for funding this project. Thank you Dr. Thomas Schubert for assisting with the MST assays and analysis.

I am so, incredibly, grateful. Thank you. Dankie. Nkosi, Gracias.

Dedication

For my family and everyone who journeyed with me.



UNIVERSITY *of the*
WESTERN CAPE

Table of Contents

Abstract.....	ii
Keywords	iv
Declaration.....	v
Acknowledgements	vi
Dedication	viii
Table of Contents	ix
List of Abbreviations	1
List of Figures.....	6
List of Tables	8
Chapter One: Literature review.....	10
1.1. Breast cancer in the developing world	10
1.2. Cancer: The Disease	12
1.2.1. Angiogenesis	13
1.2.2. The metastatic cascade	14
1.2.2.1. Local invasion.....	14
1.2.2.2. Intravasation and survival in the circulatory system	15
1.2.2.3. Extravasation	15
1.2.2.4. Micro- and macro-metastasis.....	16
1.3. Breast cancer	17
1.3.1. The basal subtype.....	18
1.3.2. The Her2 ⁺ subtype	19
1.3.3. The luminal subtype	21
1.3.4. Limitations of current breast cancer testing	23
1.3.5. Current multi-analyte prognostic and predictive applications	24
1.3.1. Current status of biomarkers for breast cancer	24
1.3.1.1. Urokinase plasminogen activator and plasminogen activator inhibitor	28
1.3.1.2. Carcinoembryonic antigen.....	29
1.3.1.3. Cancer antigen 15-3.....	29
1.3.1.1. Breast cancer carcinoma-associated antigen	30
1.3.1.1. Soluble Her2	30
1.4. Biomarker discovery.....	31
1.5. Bioinformatics as a tool for biomarker discovery.....	31
1.6. Proteomics as a tool for biomarker discovery	32
1.7. The identification of ULBP2 using bioinformatics	33
1.8. The role of ULBP2 in breast cancer	34
1.8.1. NK cells and immune surveillance	34
1.8.2. NKG2D ligands	35
1.8.2.1. MHC class I chain-related molecule A/B.....	35
1.8.2.2. UL16 binding proteins.....	36
1.8.2.3. ULBP2 as a prognostic cancer biomarker	36
1.9. The identification of GFR α 1 using proteomics	38
1.10. The role of GDNF family receptor alpha 1 in breast cancer	39

1.10.1.	RET signalling	39
1.10.2.	GDNF/GFR α 1 signalling	40
1.10.3.	Implications of GDNF-RET signalling in breast cancer	41
1.11.	Limitation of current methods of detection for breast cancer	45
1.12.	Advantages of aptamers over antibodies	46
1.13.	Aptamers.....	47
1.13.1.	What are aptamers?	47
1.13.2.	SELEX libraries	48
1.13.2.1.	Library complexity.....	48
1.13.2.2.	Library chemistry	48
1.13.3.	SELEX: The process.....	50
1.13.3.1.	Adsorption.....	51
1.13.3.2.	Separation methods	52
1.13.3.3.	Enrichment and ssDNA generation.....	53
1.13.4.	Modifications of the selection process	55
1.13.5.	Aptamer processing	56
1.13.5.1.	Identification of full-length aptamer sequence.....	56
1.13.5.2.	Determining the minimal binding sequence.....	57
1.13.6.	Characterisation of aptamer-target interaction.....	60
1.13.6.1.	Electrophoretic mobility shift assay.....	61
1.13.6.2.	Surface plasmon resonance	62
1.13.6.3.	MicroScale thermophoresis.....	62
1.13.7.	Aptamers in application	63
1.13.7.1.	Aptamers in therapeutics.....	64
1.13.7.2.	FDA approved and aptamers in clinical trials.....	65
1.13.7.3.	Aptamers in diagnostics	66
1.14.	Problem Statement and Rationale.....	70
1.15.	The aim of this study.....	72
1.16.	The objectives of this study	72
1.17.	References	73
Chapter Two: Large scale expression and purification of ULBP2		93
2.1.	Introduction.....	93
2.1.1.	The aim of this chapter	94
2.1.1.	Objectives of this chapter	94
2.2.	Materials	95
2.2.1.	Chemicals and reagents	95
2.2.2.	Buffers and solutions.....	96
2.2.3.	Oligonucleotide sequences	98
2.2.4.	Bacterial expression vectors and strains	98
2.2.5.	Assay kits and equipment	98
2.3.	Experimental methodology	99
2.3.1.	Cloning of ULBP2 into pGEX-6P-2.....	99
2.3.1.1.	Primer design.....	99
2.3.1.2.	PCR amplification of ULBP2.....	100
2.3.1.3.	Agarose gel electrophoresis	100
2.3.1.4.	Gel purification of the ULBP2 PCR amplified product	101
2.3.1.5.	Double digestion of pGEX-6P-2 and ULBP2 PCR product.....	101
2.3.2.	The pGEX-6P-2 expression system.....	102

2.3.2.1. Ligation of ULBP2 into the pGEX-6P-2 vector.....	104
2.3.2.2. Transformation of competent <i>E. coli</i> XL10-Gold cells with the pGEX-6P-2:ULBP2 construct.....	104
2.3.2.3. Colony PCR	105
2.3.2.4. Preparation of glycerol stocks.....	105
2.3.2.5. Isolation of plasmid DNA	105
2.3.3. Small-scale protein expression of ULBP2	106
2.3.3.1. Transformation of recombinant clones into <i>E.coli</i> BL21(DE3) cells.....	106
2.3.3.2. Protein expression screen of ULBP2	107
2.3.3.3. SDS polyacrylamide gel electrophoresis	108
2.3.4. Large-scale protein expression of ULBP2	109
2.3.4.1. Protein expression of ULBP2	109
2.3.4.2. Preparation of the cell lysate.....	109
2.3.4.3. Solubilisation of the cell lysate	110
2.3.5. Purification of the GST-tagged recombinant ULBP2 protein	110
2.3.5.1. Preparation of glutathione-agarose column	110
2.3.5.2. Purification of the recombinant ULBP2 protein.....	110
2.4. Results and discussion	112
2.4.1. PCR amplification of ULBP2	112
2.4.2. Cloning of ULBP2 into PGEX-6P2	113
2.4.3. Sequence analysis of positive clones.....	114
2.4.4. Small-scale protein expression of recombinant GST-ULBP2 fusion protein	115
2.4.5. Large-scale protein expression of recombinant GST-ULBP2 fusion protein.....	116
2.4.6. Solubility screen of recombinant GST-ULBP2 fusion protein.....	117
2.4.7. Purification of the recombinant GST-ULBP2 fusion protein	118
2.5. Conclusions and future recommendations	122
2.6. References	123
Chapter Three: Aptamer selection against ULBP2 and GFR α 1	126
3.1. Introduction.....	126
3.1.1. The aim of this chapter:.....	127
3.1.2. The objectives of this chapter.....	127
3.2 Materials and Methods.....	128
3.2.1. Chemicals and reagents	128
3.2.2. Buffers	129
3.2.3. Oligonucleotide sequences	129
3.2.4. Proteins and antibodies.....	129
3.2.5. Kits and Equipment	130
3.3. Experimental methodology	130
3.3.1. SELEX.....	130
3.3.1.1. Conjugation of LibIII to the target protein.....	130
3.3.1.2. Immobilization of the LibIII/protein complex to the magnetic beads	130
3.3.1.3. Pilot PCR	131
3.3.1.4. Scaled up PCR.....	132
3.3.1.5. Generation of ssDNA	132
3.3.1.6. Purification of ssDNA	133
3.3.1.7. Preparation of ssDNA for SELEX.....	133
3.3.1.8. Negative SELEX.....	133
3.3.1.9. Counter SELEX	134
3.3.1.10. Evolution monitoring by PCR	134
3.3.2. Identification of candidate aptamers.....	136

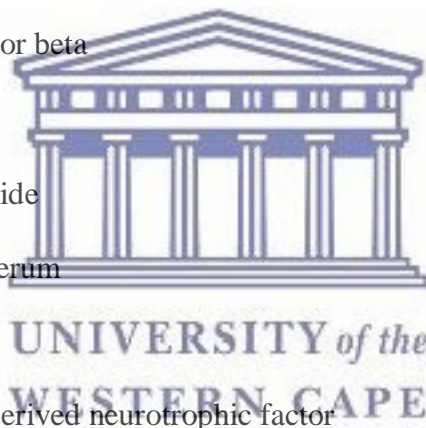
3.3.2.1. Sample preparation: NGS	136
3.3.2.2. Next generation sequencing	136
3.3.2.3. Sequence analysis using Geneious	137
3.3.3.4. Secondary structure prediction	138
3.3.3.5. Candidate aptamer selection and synthesis	138
3.4. Results and discussion	139
3.4.1. <i>In vitro</i> aptamer selection against the soluble ULBP2 ecto-domain	139
3.4.2. <i>In vitro</i> aptamer selection against the GFR α 1 protein	143
3.4.3. Identification of candidate aptamers using Next Generation Sequencing analysis.	147
3.4.4. Secondary structure prediction.....	152
3.5. Conclusion and future recommendations	155
3.6. References	156
Chapter Four: Characterisation of the binding affinity and specificity of the candidate aptamers against GFRα1.	160
4.1. Introduction.....	160
4.1.1. The aim of this chapter	162
4.1.2. The objectives of this chapter	162
4.2. Materials and equipment	163
4.2.1. Chemicals and reagents	163
4.2.2. Buffers	163
4.2.3. Oligonucleotide sequences	164
4.2.4. Recombinant proteins and antibodies	164
4.2.5. Assay kits and Equipment	164
4.3. Experimental methodology	166
4.3.1. Electrophoretic mobility shift assay	166
4.3.1.1. EMSA sample preparation.....	166
4.3.1.2. Non-denaturing Poly-Acrylamide Gel Electrophoresis	166
4.3.2. Surface plasmon resonance analysis	167
4.3.2.1. Immobilization of GFR α 1 onto CM5 sensor chips	167
4.3.3. Microscale thermophoresis.....	168
4.3.4. Flow cytometry	169
4.3.4.1. Cell culture.....	169
4.3.4.2. Cell harvesting.....	169
4.3.4.3. Cell counting/viability	170
4.3.4.4. G_Apta1 binding assay using flow cytometry	170
4.4. Results and discussion	172
4.4.1. EMSA: do the candidate aptamers bind GFR α 1?	172
4.4.2. SPR: Investigating the binding affinity and specificity of the candidate aptamers	175
4.4.3. MST: Investigating the binding interaction of G_Apta1 and the GFR α family.	188
4.4.4. Evaluating the binding and specificity of G_Apta1 to cell surface GFR α 1.	193
4.5. Conclusions and future recommendations	197
4.6. References	200
Chapter Five: Thesis Overview	202

List of Abbreviations

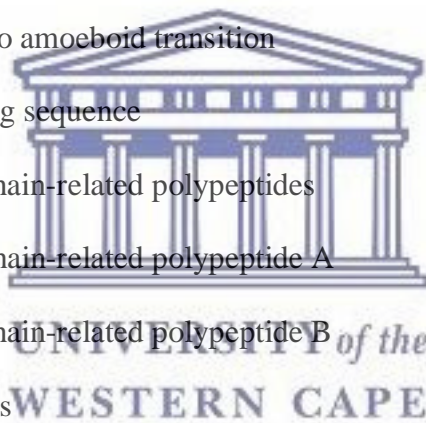
2D-PAGE:	Two-dimensional polyacrylamide gel electrophoresis
μM:	Micromolar
ALISA:	aptamer-linked immobilized sorbent assay
APS:	Ammonium persulfate
AMD:	Age-related macular degeneration
Amp:	Ampicillin
ARTN:	Artemin
ASCO:	American Society of Clinical Oncology
BSA:	Bovine serum albumin
bp:	Base pair
CA 15-3:	Cancer antigen 15-3
CAT:	Collective to amoeboid transition
CE:	Capillary electrophoresis
CEA:	Carcinoembryonic antigen
CTC's:	Circulating tumour cells
dH₂O:	Distilled water
DMEM:	Dulbecco Modified Eagle Medium
DMEM-F12:	Dulbecco Modified Eagle Medium Nutrient mixture F12
DNA:	Deoxyribonucleic acid
dsDNA:	Double-stranded DNA
DTT:	Dithiothreitol
ECL:	electro-chemiluminescence
ECM:	Extracellular matrix
<i>E.coli:</i>	<i>Escherichia coli</i>



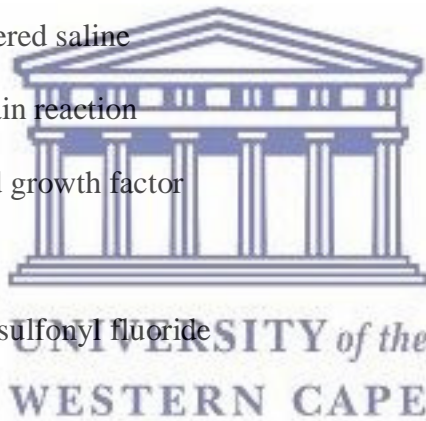
EDC/NHS:	1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide/N-hydroxysuccinimide
EDTA:	Ethylenediaminetetraacetic acid
EGFR:	Epidermal growth factor receptor
EGTM:	European group of tumour markers
ELISA:	Enzyme-linked immunosorbent assay
ELONA:	Enzyme-linked immunosorbent oligonucleotide assay
EMSA:	Electrophoretic mobility shift assay
EMT:	Epithelial to mesenchymal transition
ER:	Estrogen receptor
ERα:	Estrogen receptor alpha
ERβ:	Estrogen receptor beta
EtOH:	Ethanol
EtBr:	Ethidium Bromide
FBS:	Foetal bovine serum
FL:	Flow through
GDNF:	Glial cell line-derived neurotrophic factor
GFRα1:	GDNF family receptor alpha 1
GFRα2:	GDNF family receptor alpha 2
GFRα3:	GDNF family receptor alpha 3
GST:	Glutathione S-transferase
HEK cells:	Human embryonic kidney 293 cells
Her2:	Human epidermal growth factor 2
HSA:	Human Serum Albumin
IgG:	Immunoglobulin G
IHC:	Immunohistochemistry



IMPAKT:	Improving care and knowledge through translational research in breast cancer
IPTG:	Isopropyl β -D-1 thiogalactopyranoside
JNK:	c-Jun-NH ₂ -kinase
Kb:	Kilobases
K_D:	Equilibrium dissociation constant
kDa:	Kilodalton
ltr:	Litre
LB:	Luria broth
LH-RH:	Luteinizing hormone-releasing hormone
LibIII:	SELEX Library three
MAT:	Mesenchymal to amoeboid transition
MBS:	Minimal binding sequence
MIC:	MHC Class I chain-related polypeptides
MICA:	MHC Class I chain-related polypeptide A
MICB:	MHC Class I chain-related polypeptide B
Mg⁺⁺:	Magnesium ions
MgCl₂:	Magnesium chloride
mM:	Millimolar
MS:	Mass spectrometry
MST:	Microscale Thermophoresis
MUC-1:	Mucin-1
Na⁺:	Sodium ions
NaAc:	Sodium Acetate
NaCl:	Sodium chloride
NCCN:	National Comprehensive Cancer Network



NKG2D:	Natural killer group 2, member D
NKG2DL:	Natural killer group 2, member D ligand
NGS:	Next Generation Sequencing
nM:	Nanomolar
NRTN:	Neurturin
NT:	Nucleotides
OD:	Optical Density
O/N:	Overnight
PAGE:	Polyacrylamide gel electrophoresis
PAI-1:	Plasminogen activator inhibitor 1
PBS:	Phosphate buffered saline
PCR:	Polymerase chain reaction
PDGF:	Platelet-derived growth factor
pM:	Picomolar
PMSF:	Phenylmethanesulfonyl fluoride
POC:	Point-of-care
PR:	Progesterone receptor
PSPN:	Persephin
qRT-PCR:	Quantitative real time polymerase chain reaction
RET:	Rearranged during transformation
RNA:	Ribonucleic acid
RPMI 1640:	Roswell Park Memorial Institute
RT:	Room temperature
RTK:	Receptor tyrosine kinase
RT-PCR:	Real-time polymerase chain reaction



SELEX:	Systematic Evolution of Ligands by Exponential enrichment
SDS:	Sodium Dodecyl Sulfate
SDS-PAGE:	Sodium Dodecyl Sulfate polyacrylamide gel electrophoresis
SEOM:	Spanish Society of Medical Oncology
sHER2:	Soluble human epidermal growth factor2
SILAC:	Stable isotope labelling with amino acids in cell culture
SPR:	Surface Plasmon Resonance
SSA:	Sub-Saharan Africa
ssDNA:	Single-stranded DNA
sULBP2:	Soluble unique ligand binding protein 2
TBE:	Tris/Boric acid/EDTA
T-DM1:	Trastuzumab emtansine
TEMED:	Tetramthylethylenediamine
Tris:	Trisaminomethane
ULBP1:	Unique ligand binding protein 1
ULBP2:	Unique ligand binding protein 2
ULBP3:	Unique ligand binding protein 3
ULBP6:	Unique ligand binding protein 6
uPA:	Urokinase plasminogen activator
UV:	Ultraviolet
VEGF:	Vascular endothelial growth factor



List of Figures

- Figure 1.1:** Tumour growth from a single cell.
- Figure 1.2:** Histological classification of breast cancers.
- Figure 1.3:** Activation of the various RET signalling pathways.
- Figure 1.4:** A schematic diagram of a typical SELEX process.
- Figure 1.5:** Stepwise aptamer truncation using a rational approach.
- Figure 1.6:** A lateral flow assay (LFA) developed for the detection of Vaspin.
- Figure 2.1:** NCBI nucleotide sequence for the coding region of ULBP2.
- Figure 2.2:** Circularized map and sequence referencing points of the pGEX-6P-2 expression vector
- Figure 2.3:** PCR amplification of ULBP2 resolved on a 1% agarose gel.
- Figure 2.4:** Colony PCR screen for the presence of ULBP2.
- Figure 2.5:** Pairwise sequence alignment of the translated DNA sequence obtained for Clone 1 (query sequence) against the protein sequence for the human ULBP2 protein (subject sequence).
- Figure 2.6:** Small-scale protein expression screen of the recombinant ULBP2.
- Figure 2.7:** Large-scale protein expression screen of the recombinant ULBP2.
- Figure 2.8:** Solubilisation of the recombinant ULBP2 protein.
- Figure 2.9:** Affinity purification of the solubilised sarkosyl recombinant GST-ULBP2 fusion protein.
- Figure 3.1:** Typical workflow for a round of SELEX.
- Figure 3.2:** SELEX against the ULBP2 protein.
- Figure 3.3:** SELEX against the GFR α 1 protein.
- Figure 3.4:** Monitoring the evolution of the *in vitro* selection against GFR α 1.

Figure 3.5: Multiple sequence alignment of top 16 overrepresented candidate aptamers against GFR α 1.

Figure 4.1: Electrophoretic mobility shift assay of six candidate aptamers against GFR α 1.

Figure 4.2: BIAcore sensorgram for the immobilization of GFR α 1 and control proteins on a CM5 sensor chip, using EDC chemistry.

Figure 4.3: BIAcore sensorgram validating the immobilization of GFR α 1, using Anti-GFR α 1.

Figure 4.4: Screening of six candidate aptamers for binding to GFR α 1 and cross-reactivity to control proteins.

Figure 4.5: Multiple alignment (A) and percentage similarity matrix (B) of the GFR α protein family, using Clustal Omega.

Figure 4.6: BIAcore sensorgram for the immobilization of GFR α family of proteins on a CM5 sensor chip, using EDC chemistry.

Figure 4.7: BIAcore sensorgram validating the immobilization of GFR α 1, using Anti-GFR α 1.

Figure 4.8: Evaluating the specificity of the six candidate aptamers against the GFR α protein family.

Figure 4.9: Quality check of capillaries.

Figure 4.10: Normalized fluorescence of all the protein-aptamer complexes.

Figure 4.11: MST analysis of the interaction of the G_Apta1 with the GFR α protein family.



Figure 4.12: Evaluation of G_Apta1 binding to the cell surface GFR α 1 protein, using flow cytometry.

List of Tables

- Table 1.1:** Recommendations for different genetic tests as prognostic tools or to establish the benefit of adding chemotherapy to hormone therapy in the management of breast cancer.
- Table 1.2:** Separation methods used for SELEX.
- Table 1.3:** Different types of SELEX methodologies.
- Table 1.4:** FDA approved and aptamers in various stages of clinical trials.
- Table 1.5:** Commercialized aptamer-based products
- Table 2.1:** PCR parameters for the amplification of ULBP2.
- Table 2.2:** Double digestion using BamHI and XhoI.
- Table 2.3:** Reagents used in the ligation reaction of pGEX-6P-2 and ULBP2.
- Table 2.4:** Reagents used for the preparation of two SDS polyacrylamide gels.
- Table 2.5:** Pilot PCR parameters for the amplification of LibIII against the recombinant ULBP2 protein.
- Table 2.6:** Scaled-up PCR parameters for the amplification of LibIII against the recombinant ULBP2 protein.
- Table 3.1:** Pilot PCR parameters.
- Table 3.2:** Scaled up PCR parameters.
- Table 3.3:** Summary of the SELEX conditions for GFR α 1.
- Table 3.4:** Next Generation Sequencing analysis of the tenth selection pool against GFR α 1.
- Table 3.5:** Sequences of final candidate aptamer list.
- Table 4.1:** Preparation of an 8% EMSA gel.
- Table 4.2:** Preparation of proteins for MST assay.

Table 4.3: Cell lines used to screen for G_Apta1 binding.



Chapter One: Literature review

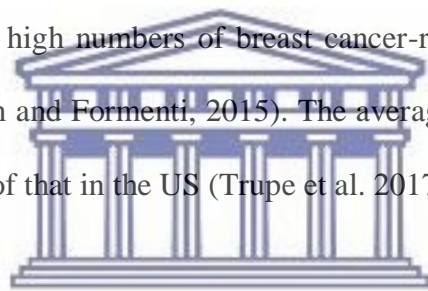
1.1. Breast cancer in the developing world

Globally, breast cancer is one of the leading causes of cancer-related deaths among women (Krombein and De Velliers 2006; Sharma et al. 2012; Cumber et al. 2017). In 2012, approximately 14.1 million cases of women were diagnosed with cancer; 1.7 million were diagnosed with breast cancer and 56.8% represented cases from low-income countries (Cumber et al. 2017). In that same year, 522 000 deaths were recorded as a result of breast cancer, with the majority from Sub-Saharan Africa (SSA) (Cumber et al. 2017). By 2025, it is estimated that more than 19.3 million women will be suffering from breast cancer; of which the majority is expected to be from SSA (Cumber et al. 2017). Since most SSA countries lack national cancer registries and limited published data is available, an accurate account of the prevalence of breast cancer in SSA has not been recorded (Cumber et al. 2017; Adeloje et al. 2018).

The latest (2014) national cancer registry for South Africa showed that breast cancer is the most commonly diagnosed cancer amongst South African women; accounting for 21.78 % of all cancers present in South African women (South African National Cancer Registry, 2014). A study carried out by Rambau et al. (2011) demonstrated that more than 70 % of breast cancer patients in Eastern Africa only presented the disease at clinical stages III or IV (Rambau et al. 2011). In these patients, it is not the primary breast cancer, but rather the metastatic tumours, that have spread to vital organs that will result in mortality (Weigelt et al. 2005). What comes forth, pressingly, in developing countries – as a result of late diagnosis, is that approximately 50 % of the patients have tumour sizes above 6 cm at the time of diagnosis (Coughlin and Ekwueme, 2009 and Abdulrahman and Rahman, 2012). Consequently, most patients of African origin are left with mastectomy and adjuvant therapy as their best options; leaving a great

number of patients to receive palliative care since the disease progressed too far by the time of diagnosis (Coughlin and Ekwueme, 2009 and Abdulrahman and Rahman, 2012).

In the US, over the past 30 years, a decline in breast cancer deaths have been observed. This is largely attributed to increased routine mammography screening, subsequently allowing for improved and earlier treatment of breast cancer (Kalager et al. 2009; Laronga 2016). Despite the reported reduction in breast cancer mortalities, implementing routine screening using current methods, such as mammography is implausible in developing countries due to significantly high associated costs (Shulman et al. 2010). At this point, even though incidence rates of breast cancer are still higher in developed countries, survival rates are higher in developed countries; whereas high numbers of breast cancer-related deaths are observed in developing countries (Balogun and Formenti, 2015). The average 5-year survival in Africa is said to be half (50% vs 98%) of that in the US (Trupe et al. 2017).



In trying to answer the reasons why low survival rates are observed in developing countries, Adeloje and colleagues sought to perform a systemic review of breast cancer incidences in Africa; they found that some of the reasons why low survival rates are observed in low- and middle-income countries are due to poor healthcare infrastructures; a lack of registries; lack of population awareness and low levels of education and empowerment around the topic of breast cancer – which then leads to delayed health seeking behaviours; and furthermore the challenge of improving the public health response in these particular health settings (Adeloje et al. 2018). In particular, the issue of undocumented prevalence places these countries in a position where the exact burden of the disease is unquantified and as a result the required healthcare for breast cancer incidences are neglected, since the government places a significant emphasis on combating those diseases regarded as healthcare priorities (Cumber et al. 2017). Even in some

isolated SSA countries where relatively good breast cancer services are available, issues such as inequity and affordability still prevents the general public from accessing these services, resulting in overall low survival rates (Cumber et al. 2017).

In view of this, it is clear that a more feasible alternative is required to help alleviate the mortality burden caused by breast cancer in developing countries. In order to efficiently make recommendations in this regard, a better understanding of the disease is required.

1.2. Cancer: The Disease

Cancer may be described as the abnormal growth of cells (Gotter 2009). Normally, when a cell recognises any sort of impairment to its DNA, repair mechanisms are initiated (Gotter 2009). However, failure to execute this gives rise to genomic instabilities; consequently, a cancerous cell is formed (Gotter 2009). In the process of maintaining normal biological processes, cancerous cells along with healthy cells grow and divide to form new cells; and when any abnormal cells are identified, there are regulatory signalling pathways that eliminate these cells through programmed cell death/apoptosis (Steichen et al. 2013). However, cancer cells can acquire mutations resulting in malfunctioning apoptotic signalling and consequently, these cells escape apoptosis, multiplying unhindered to form tumour masses (**Figure 1.1**) (Brannon-peppas and Blanchette 2004). Of note, is that the replication rates of cancerous cells outcompete that of healthy cells and in so doing, prompt a higher demand for nutrient supply; and as a result displaces healthy cells (Steichen et al. 2013). Initial tumour growth stages rely on diffusion of nutrients, restricting the tumour size to approximately 2 mm³ (Brannon-peppas and Blanchette 2004). To overcome this diffusion limit, the tumour requires a new blood vessel system; and this is accomplished through a process called angiogenesis (Danhier et al. 2010; Steichen et al. 2013).

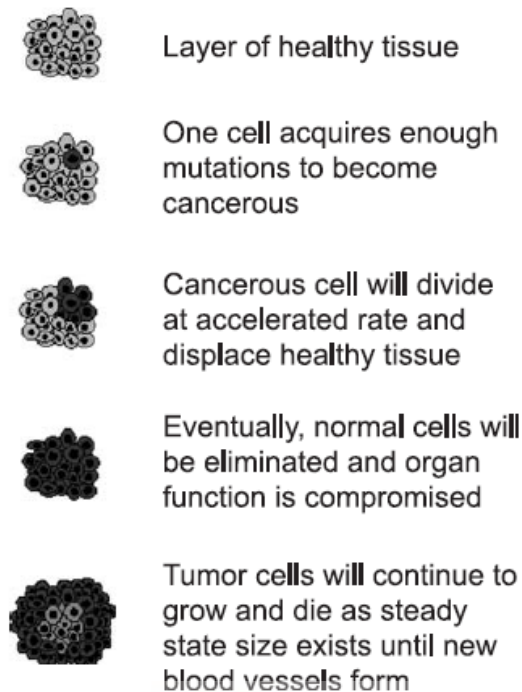
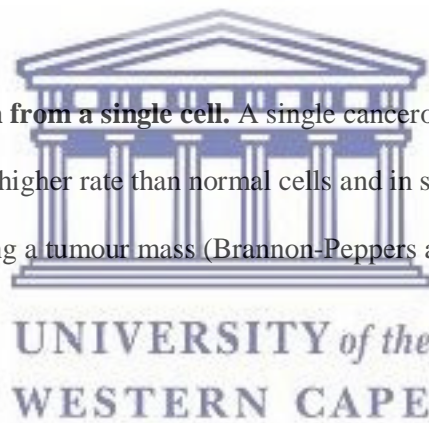


Figure 1.1. Tumour growth from a single cell. A single cancerous cell around healthy tissue undergoes replication at a much higher rate than normal cells and in so doing, displaces healthy tissue and eventually forming a tumour mass (Brannon-Peppers and Blanchette 2004).



1.2.1. Angiogenesis

Angiogenesis comprises the construction of new blood vessels from pre-existing vessels and once the tumour reaches its solid size state (2 mm^3), hypoxia is initiated; subsequently inducing angiogenesis (Danhier et al. 2010). Another point to note is that the number of endothelial cells, that line the blood vessels, usually exist in a relatively dormant state, with about 1 in every 10 000 endothelial cells subjected to cell division. During angiogenesis, however, the turnover increases up to 50 times, thus promoting tumour growth (Brannon-peppas and Blanchette 2004). Since an adequate blood supply is absolutely necessary for tumour growth, blood vessels then rapidly grow, generating a highly disorganised and irregular vasculature with some regions rich in blood supply and others low in blood supply (Byrne et al. 2008; Danhier et al.

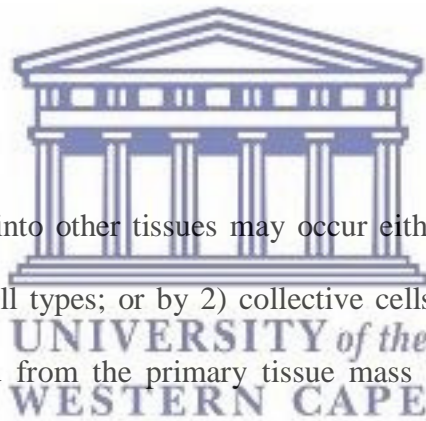
2010). Those cancerous cells in closer proximity to the rich blood supply have the potential to spread and invade distant organs by a process known as metastasis (Brannon-peppas and Blanchette 2004).

1.2.2. The metastatic cascade

The invasion and metastasis of cancer cells, from the primary tumour site, is a complex sequence of events, termed the metastatic cascade. The stages of this cascade are listed as follow: 1) local invasion, 2) intravasation and survival in the circulatory system, 3) extravasation, 4) micrometastasis and macrometastasis (colonization) (Hanahan and Weinberg 2011; Zijl et al. 2011).

1.2.2.1. Local invasion

The invasion of cancer cells into other tissues may occur either by 1) single cells, such as mesenchymal or amoeboid cell types; or by 2) collective cells, such as epithelial sheets or clusters that have fragmented from the primary tissue mass (Zijl et al. 2011). As cancer progresses, various cancer cells change their plasticity, demonstrated by morphological and phenotypical adaptations. These occur in one of three ways, namely, epithelial to mesenchymal transition (EMT); mesenchymal to amoeboid transition (MAT) or collective to amoeboid transition (CAT). Since EMT has been recognized as the most prominent of the three, it will be the focus of this discussion. It is a well-preserved process that is initiated upon cancer progression, among other cell activities. There are three types of EMT; of which type III is indicative of cancer progression to a more aggressive carcinoma (Hanahan and Weinberg 2011; Zijl et al. 2011). Molecular characteristics of EMT are illustrated by the loss of epithelial functions and gain of mesenchymal gene expression patterns. One of the well-defined characteristics of EMT is the loss of E-cadherin expression; which is an important cell to cell



adhesion molecule required for the assembly and integrity of adjacent epithelial cells within epithelial sheets (Hanahan et al. 2000; Hanahan and Weinberg 2011). Therefore, functional loss of E-cadherin promotes the detachment of cell clusters or single cells that have the ability to migrate with mesenchymal properties (Zijl et al. 2011).

1.2.2.2. Intravasation and survival in the circulatory system

Depending on the cancer type, intravasation is the process by which various changes at the molecular level facilitates the migration of the locally invasive cells, as described in Section 1.2.2.1, across the endothelial barriers and into the lumina of blood or lymph vessels (Valastyan and Weinberg, Robert 2011). At this point, the cells are referred to as circulating tumour cells (CTC's) and have to survive several processes that act to eliminate them. One key example is that these CTC's do not possess the ability to adhere to the extracellular matrix (ECM), which is otherwise important for cell survival. And as a result, the lack of this type of anchorage, epithelial cells are subjected to cell death (Valastyan and Weinberg, Robert 2011). However, cancer cells have found ways to evade cell death during circulation; these include the expression of receptors that stimulate anti-apoptotic signalling and thus promote cell survival (Zijl et al. 2011). In breast cancer, specifically, it has been reported that due to their brief circulatory times, the CTC's manage to escape the circulation before cell death signals is activated (Valastyan and Weinberg, Robert 2011).

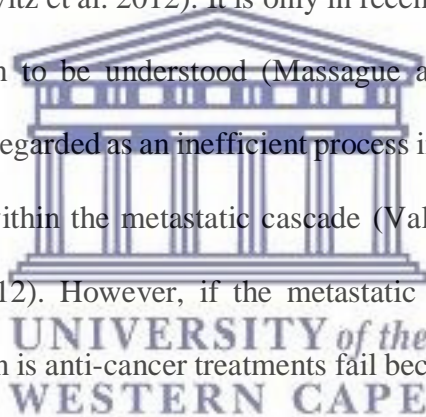
1.2.2.3. Extravasation

During the extravasation process, the CTC's escape from the lumina of the lymph or blood vessels into the parenchymal tissues distant from the primary tumour site (Hanahan and Weinberg 2011). Bearing in mind that the CTC's are required to survive the circulatory system

during extravasation, one of the well-characterised mechanisms that CTC's employ is the formation of thrombi that surround the tumour cells. Thrombi formation leads to platelet activation. This, in turn, acts as a physical shield against the pressure of blood flow and enable CTC's to evade immune surveillance and successfully extravasate (Labelle and Hynes 2012; Lou et al. 2015).

1.2.2.4. Micro- and macro-metastasis

Metastatic colonization is the final event in the metastatic cascade and the point at which the cancerous cells, following intra- and extravasation, initiate the growth of a tumour mass at a secondary tissue site (Szmulewitz et al. 2012). It is only in recent years that the mechanism, of this specific event, has begun to be understood (Massague and Obenauf 2016). As often reported, overall metastasis is regarded as an inefficient process in that the cancerous cells often fail to carry out some steps within the metastatic cascade (Valastyan and Weinberg, Robert 2011; Labelle and Hynes 2012). However, if the metastatic cascade successfully reaches colonization, often what is seen is anti-cancer treatments fail because current treatments do not yet show sustained effects at this point of the disease (Massague and Obenauf 2016). It is important to note, that even though it would make sense to target earlier points within the metastatic cascade, the swift dissemination of cells often has already occurred by the time of diagnosis. Whereas re-initiation of tumour growth, colonization and tumour survival, at the secondary site, occurs over a longer period and therefore, therapeutic intervention at these points can be made and are thus also regarded as more clinically feasible targets. This approach is more beneficial in retarding the tumour growth and, in turn providing more time to manage the treatment of the disease (Labelle and Hynes 2012).



This section provided a general summary of cancer, however, depending on the origin of the tumour, very different responses to treatment options and diagnostic capabilities are observed. Thus, a further look into breast cancer, specifically, will be explored in this review.

1.3. Breast cancer

It is well recognized that breast cancer is a heterogeneous group of neoplasms originating from the ductal epithelium (Place et al. 2011; Polyak 2011; Arango et al. 2013). Its' heterogeneity is attributed to the morphology, immune-histology as well as molecular events associated with cancer (Bhargava et al. 2011). There are two major histological types of breast cancer, namely, in situ carcinoma and invasive carcinoma. These can further be categorized as schematically illustrated in **Figure 1.2** (Malhotra et al. 2010).

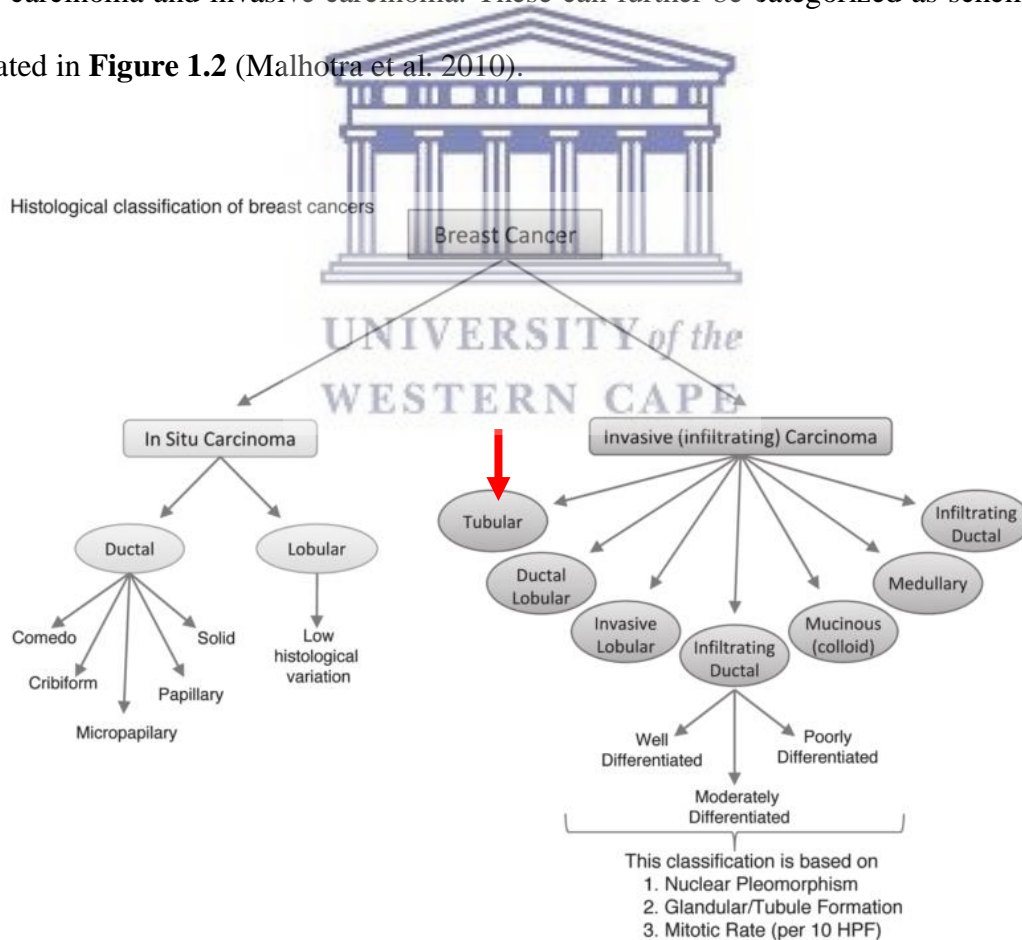


Figure 1.2. Histological classification of breast cancers. The red arrow indicates the most commonly diagnosed breast cancer type. (Malhotra et al. 2010).

Invasive ductal carcinoma (IDC), indicated by the red circle in **Figure 1.2**, is the most common type of breast cancer, accounting for approximately 80 % of all breast cancers. It originates in the milk ducts of the breast and further infiltrates the duct wall, thus invading fatty tissue and potentially other organs in the body (Sharma et al. 2010). Once the patient has been diagnosed with IDC, surgical and therapeutic intervention is required. And although the above-mentioned classification provides a distinction between the types of cancer, being able to identify and classify the molecular markers presented by the type of breast cancer provides a platform to guide appropriate treatment options and predict the relative risk of recurrence (Malhotra et al. 2010). This has led to the classification of three major subtypes of breast cancer, as defined by both clinical practice and gene expression profiling; namely, the human epidermal growth factor 2 (Her2⁺) subtype, the luminal subtype and the basal subtype (Hu et al. 2009; Higgins and Baselga 2011; Polyak 2011). The three genes associated with these three subtypes are Her2 receptor, a membrane tyrosine kinase that when activated affects cell proliferation and survival (Krishnamurti and Silverman, 2014); the estrogen receptor (ER), a nuclear hormone receptor that is activated by estrogen and once activated affects cellular proliferation and differentiation (Gross and Yee, 2002) and the progesterone receptor (PR), a nuclear receptor that is activated by progesterone and once activated affects signal transduction pathways involved in proliferative signalling in the breast tissue (Daniel, Hagan and Lange, 2012).

1.3.1. The basal subtype

The basal subtype, also known as the triple negative subtype, accounts for approximately 15-25 % of all breast cancers and is characterised by the low expression of ER, PR and Her2; overexpression of proliferative genes and a high expression of what is known as the basal gene cluster (P-cadherin; epidermal growth factor (EGFR); KIT and MET proto-oncogene receptor

tyrosine kinases; β 4-integrin; caveolin 1 and 2; α -basic crystalline) (Bertucci, F. and Birnbaum 2008; Bertucci et al. 2008). Even though the basal subtype is relatively sensitive to treatment, it is strongly associated with poor prognosis; which is corroborated by the expression of some genes in the basal cluster, such as P-cadherin, EGFR, α -basic crystalline and moesin – genes that are independently associated with poor prognosis in breast cancer (Bertucci et al. 2008). Despite extensive research, no clinically validated molecular prognostic tests have been developed for the basal subtype (von Wahlde et al. 2016).

1.3.2. The Her2⁺ subtype

The Her2⁺ subtype (Her2-positive breast cancer) accounts for approximately 30 % of all breast cancers and is characterized by Her2 gene amplification/protein overexpression as well as low expression of the luminal genes [Her2 (+); ER/PR-receptor (-)] (Polyak 2011; Kittaneh and Montero 2013). The Her2 receptor is a membrane tyrosine kinase, that when overexpressed promotes tumour growth, invasion and survival, by activating signalling cascades, such as the MAPK and P13K/Akt pathways (Krishnamurti and Silverman, 2014; Schettini et al. 2016). Clinically, Her2 measurement serves, primarily, as a predictive biomarker for the response of breast cancer patients to anti-Her2 therapies in the neo-adjuvant (treatment administered prior to the main treatment/surgery, for the purpose of shrinking the tumour), adjuvant (additional treatment after the primary treatment, for the purpose of reducing risk of recurrence) and advanced disease (treatment of metastatic tumours) settings (Duffy et al. 2017). Currently, there are four FDA approved anti-Her2 therapies for the treatment of Her2 positive breast cancer, namely, trastuzumab, lapatinib, pertuzumab and trastuzumab emtansine (T-DM1) (Duffy et al. 2017). The Her2 receptor is currently the only approved predictive biomarker for the selection of patients that may benefit from the above-mentioned anti-Her2 therapies; however, this treatment option does not always result in a positive response to treatment. In

fact, the overexpression of Her2 is associated with aggressive breast tumours, reduced responses to traditional anti-Her2 therapies and decreased survival and thus better serves as a prognostic biomarker for breast cancer recurrence (Bhargava et al. 2011; Polyak 2011).

Due to the complex biology of Her2 positive breast cancers, various interactions in several signalling pathways contribute to the development of resistance of the Her2 positive breast cancers to trastuzumab (Baselga et al. 2017). However, this lead to the introduction of other combinational treatment options that has reported to reduce the relapse rate in Her2 positive early breast cancer (Baselga et al. 2017). Additionally, several promising neo-adjuvant, adjuvant and metastatic clinical trials (phase II-III), for the use of new anti-Her2 agents, are currently underway and are demonstrating good pathological response, disease free survival and overall survival of Her2 positive breast cancer patients (Wahlde et al. 2016; Baselga et al. 2017). Despite the ongoing efforts to develop new treatment strategies, in accordance with the most recently published guidelines on the clinical use of biomarkers in breast cancer, established by the European Group of Tumour Markers (EGTM), further research with respect to additional biomarkers for Her2 positive breast cancer is recommended as follows:

- *“Identify additional markers to increase the positive predictive value of Her2. This should focus on Her2-positive patients who do not benefit from trastuzumab or other forms of anti-Her2 therapy, as well as the identification of patients who derive long-term benefit from anti-Her2 therapy.”* (Duffy et al. 2017).
- *“Identify biomarkers for selecting the most appropriate form of anti-Her2 therapy for a given patient.”* (Duffy et al. 2017).
- *“Markers should be identified for selecting patients likely to particularly benefit from dual anti-Her2 therapies... in the neo-adjuvant and advanced disease settings.”* (Duffy et al. 2017).

- ‘Establish whether patients with equivocal scores should or should not receive anti-Her2 therapy. This question might be addressed by evaluating the potential predictive value of other assays for Her2, such as enzyme-linked immunosorbent assay (ELISA) for Her2 protein or real-time PCR (RT-PCR) for Her2 mRNA.’ (Duffy et al. 2017).
- ‘..The potential biomarker value of Her2 mutations in predicting the response or resistance to specific anti-Her2 therapies should be explored...’ (Duffy et al. 2017).

1.3.3. The luminal subtype

The luminal subtype is a hormonal breast cancer which may be further categorized into two subtypes, namely, luminal A and luminal B (Kittaneh and Montero 2013). Luminal A tumours account for approximately 40 % of all breast cancers and are characterised by the overexpression of ER-related genes and the low expression of Her2 [ER (+) and/or PR (+); Her2 (-)]. Additionally, this subtype is well associated with positive prognosis (Kittaneh and Montero 2013). Luminal B tumours account for approximately 20 % of all breast cancers and are characterized by a lower expression of ER-related genes, variable expression of Her2 and relatively high expression of proliferative genes [ER (+) and/or PR (+); Her2 (+)] (Creighton 2012; Kittaneh and Montero 2013). Furthermore, this subtype has been described as the more aggressive luminal tumour, enhanced by the overexpression of proliferative genes and consequently, associated with a significantly higher risk of relapse (Creighton 2012). The luminal B subtype is also less common than the luminal A subtype and is associated with a poorer prognostic outcome (Creighton 2012; Kittaneh and Montero 2013; Ades et al. 2014). The luminal subtypes are identified based on two immunohistochemistry (IHC) biomarkers, 1) the ER, a nuclear hormone receptor that is activated by estrogen and once activated affects cellular proliferation and differentiation (Gross and Yee, 2002). When ER is overexpressed in breast tissues, it drives the growth and survival of breast tumours (Schettini et al. 2016) and 2)

the PR, a nuclear receptor that is activated by progesterone and once activated affects signal transduction pathways involved in pro-proliferative signalling in breast (Daniel, Hagan and Lange, 2012). It is important to note that independent PR overexpression provides no treatment beneficial predictive information; however; PR overexpression together with ER overexpression in breast tumours predicts the increased probability of the breast tumour to respond positively to tamoxifen treatment (Lim et al. 2016; Duffy et al. 2017). It has been reported that ‘less than 1% of tumours in all breast cancer cases express the PR, but not the ER’ (Wahid et al. 2017). In all other cases, the ER overexpression status assists in selecting patients with invasive breast tumours that are likely to respond to endocrine therapy, in the neoadjuvant, adjuvant and advanced disease settings (Duffy et al. 2017). Several FDA approved endocrine therapies exist, these include estrogen receptor modulators (tamoxifen), third generation aromatase inhibitors (anastrozole, letrozole, or exemestane), luteinizing hormone-releasing hormone (LH-RH) agonists (leuprolide and goserelin), pure estrogen receptor down-regulators (fulvestrant), oophorectomy and others (Duffy et al. 2017).



There exist two forms of ER, namely, ER α and ER β , however, only a clinically validated role for ER α has been established to date (Duffy et al. 2017). Furthermore, authorised IHC assays for breast cancer testing only detect ER α . In approximately 70% of diagnosed breast cancer cases, ER α expression is observed and therefore serves as a good predictive biomarker for response to endocrine therapy and it also offers prognostic value by confirming resistance to certain endocrine therapies (Duffy et al. 2017). The PR also exists in two forms, PR α and PR β ; both can be detected in authorised IHC assays (Duffy et al. 2017).

The fundamental principle behind endocrine therapy is to reduce ER activity or reduce the expression levels of the ER within breast cancer cells (Lim et al. 2017). Although the ER-

positive breast cancer subtype is more responsive to treatment, favourable prognosis is only observed in the first 5-7 years after initial diagnosis (Duff et al. 2017). Thereafter, the risk of recurrence increases and a great proportion of ER-positive breast cancer patients begin to develop resistance to endocrine therapy (Duff et al. 2017; Lim et al. 2017). Thus, focus has shifted to combining endocrine therapy with novel therapeutic agents that target the biology around the resistance mechanism and essentially aims to enhance the usefulness of endocrine therapies that are currently available (Lim et al. 2017). Additionally, the EGTM recommends further research with respect to additional biomarkers for ER positive breast cancer, as follows:

- *“Development of biomarkers for increasing the positive predictive value of ER.”*
- *“Identify biomarkers for selecting patients that preferentially benefit from an aromatase inhibitor, vis-á-vis tamoxifen or vice versa.”* (Duffy et al. 2017).
- *“Validate biomarkers for selecting patients who do not need extended adjuvant endocrine therapy.”* (Duffy et al. 2017).



UNIVERSITY of the
WESTERN CAPE

1.3.4. Limitations of current breast cancer testing

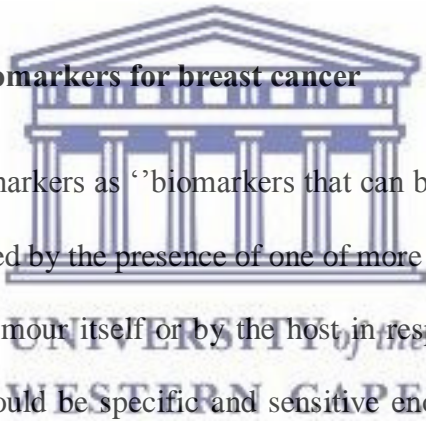
Three validated IHC breast cancer biomarkers (ER, PR and Her2) have played an invaluable role in increasing the efficacy of currently available treatments as well as the detection of tumorous lesions in breast tissues (Wahid et al. 2017). Both the American Society of Clinical Oncology (ASCO) and the EGTM recommend that all newly diagnosed breast cancer patients undergo the evaluation of Her2, ER and PR expression in tissue samples that are obtained from biopsies or surgery (Ulaner et al. 2016; Duffy et al. 2017). Thereafter, the breast tumours are classified according to the four breast cancer subtypes, based on the Her2, ER and PR expression status; which in turn determines the prognostic outcome and treatment predictive responses (Ulaner et al. 2016). Although very effective in clinical practice for many years, as a result of patients developing resistance to currently available treatment options, it is quite

clear that there is a growing need for new biomarkers to compensate for the limitations presented by Her2, ER and PR (Ulaner et al. 2016; Duffy et al. 2017).

1.3.5. Current multi-analyte prognostic and predictive applications

To advance the accuracy of traditional prognostic and predictive markers, the application of multi-analyte testing was introduced and have now become routine testing platforms to aid in guiding the appropriate treatment (Hagemann, 2016). **Table 1.1** describes the application of the various multi-analyte platforms and their recommendations as provided by different expert panels.

1.3.1. Current status of biomarkers for breast cancer



Kabel 2017 describes cancer markers as “biomarkers that can be found in the blood, urine or body tissues that can be elevated by the presence of one or more types of cancer (Kabel 2017). It is produced either by the tumour itself or by the host in response to the tumour” (Kabel 2017). Ideally, biomarkers should be specific and sensitive enough to detect tumours while they are still small to allow for early diagnosis or screening of recurrent tumours; which in turn allows for early therapeutic intervention (Kabel 2017). It is important to note that although most biomarkers are produced by different tumours, they may still be produced by the same tissue types, which is the reason why it is often observed that some biomarkers are not specific to a single tumour (Kabel 2017). However, they are very useful because the expression of tumour markers, in either the blood or tissue, are clearly distinguishable between cancerous and healthy subjects (Kabel 2017). This information gives insight into the progression of the disease status after initial treatments, such as chemotherapy or radiotherapy and plays a vital

part in assisting medical practitioners to make therapeutic interventions when the patient is very likely still to respond to adjuvant therapies (Kabel 2017).

Table 1.1. Recommendations for different genetic tests as prognostic tools or to establish the benefit of adding chemotherapy to hormone therapy in the management of breast cancer.

	Oncotype DX®	Mammaprint®	Prosigna®	EndoPredict®
ASCO	Guides the decision to prescribe adjuvant systematic chemotherapy. Evidence: High Recommendation: Strong	Should not be used for decision making about adjuvant systematic chemotherapy use. Evidence: Intermediate Recommendation: moderate	Guides the decision to prescribe adjuvant systematic chemotherapy together with other clinical and pathological variables. Evidence: High Recommendation: strong	Guides the decision to prescribe adjuvant systematic chemotherapy. Evidence: Intermediate Recommendation: moderate
NCCN	The only test recommended for patients with >0.5cm tumour. Oncotype DX® can be considered for selecting patients with 1-3 ipsilateral lymph nodes involved. The only test validated for predicting chemotherapy response.	Prognostic value, but not validated for predicting chemotherapy response.		
St Gallen 2015	Prognostic value and predictive of the benefit of adjuvant chemotherapy.			
SEOM	5 year recurrence risk prognosis: 1A/1B 10 year recurrence risk prognosis: 1B Chemotherapy benefit prediction: 1A/1B	5 year recurrence risk prognosis: 1B 10 year recurrence risk prognosis: - Chemotherapy benefit prediction: -	5 year recurrence risk prognosis: 1B 10 year recurrence risk prognosis: 1B Chemotherapy benefit prediction: -	5 year recurrence risk prognosis: 1B 10 year recurrence risk prognosis: 1B Chemotherapy benefit prediction: -
IMPAKT	Little, but significant prognostic information above and beyond clinical and pathological parameters. No evidence of clinical usefulness for modifying the treatment decision.			



ASCO: American Society of Oncology, *IMPAKT*: Improving Care and Knowledge Through Translational Research in Breast Cancer, *NCCN*: National Comprehensive Cancer Network. *SEOM*: Spanish Society of Medical Oncology (Colomer et al. 2018).

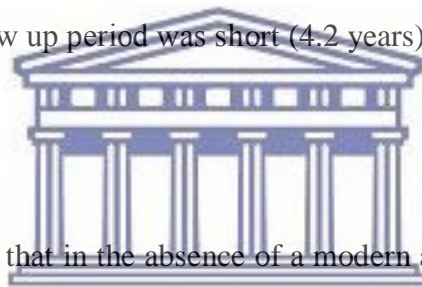
To date, tissue-based biomarkers, such as hormone receptors (ER and PR), Her2 and Ki67 have been recommended for determining prognostic and predictive information (Kabel 2017; Duffy et al. 2018). Whereas, blood-based or serum biomarkers, such as carcinoembryonic antigen

(CEA) and cancer antigen 15-3 (CA-15-3) are mainly used to monitor the response to endocrine or chemotherapy of patients in advanced disease settings (Duffy et al. 2018).

Although several studies have demonstrated the ability of the blood-based or serum biomarkers to be used as a measurement tool for the detection of asymptomatic recurrences of cancer, most expert panels recommend against it (Duffy et al. 2018). Duffy and colleagues argues that this recommendation is based on the “lack of high level evidence demonstrating that an intensive follow-up strategy improves outcome or improves quality of life compared to a minimal follow-up strategy” (Duffy et al. 2018). And this lack of evidence is, primarily, based on two outdated findings from Italian randomized prospective trials that were published in the 1990’s, in which both studies concluded that a follow-up strategy consisting of regular physical examinations and an annual mammography is the most effective approach with respect to overall survival (Duffy et al. 2018). Duffy and colleagues further explain that because these trials took place more than 30 years ago, their relevance to modern management of breast cancer patients is questionable for several reasons, such as 1) currently proposed blood-based or serum biomarkers were not included; 2) older methods used as imaging instruments were less accurate in the detection of early recurrences; 3) at the time, only tamoxifen was available as a systematic endocrine therapy; and 4) presently, multiple targeting agents for Her2+ tumours are available – which is particularly of note because prior to the use of trastuzumab, Her2+ breast cancer patients demonstrated much worse survival outcomes as compared to Her2- patients (Duffy et al. 2018).

It is, however, noted that a subsequent prospective study (1991 – 1995) that included the use of CA15-3 was performed (Kokko et al. 2005; Duffy et al. 2018). CA 15-3 is a blood-based biomarker that has been shown to be a useful prognostic indicator of breast cancer and it is also

used for monitoring treatment efficacy in breast cancer (Kabel 2017). This study consisted of 472 patients with newly diagnosed breast cancer, who were randomized to a 3 versus 6-month follow-up strategy, together with standard investigations at every visit (full blood count, CA 15-3 and standard biochemical testing) versus no routine testing (Kokko et al. 2005; Duffy et al. 2018). The authors concluded that it was just as effective to have a follow up every 6 months as it is having it every 3 months and furthermore, that in asymptomatic breast cancer patients, it is unnecessary to routinely perform blood tests and imaging in the follow up, in fact it doubled the cost of the follow up visit (Kokko et al. 2005; Duffy et al. 2018). Duffy and colleagues puts forth that “it is unclear whether this trial had sufficient power to address whether intensive follow up with CA 15-3 testing resulted in a superior outcome to that of a minimalist follow-up strategy” because the follow up period was short (4.2 years) and a small cohort of patients were used (Duffy et al. 2018).

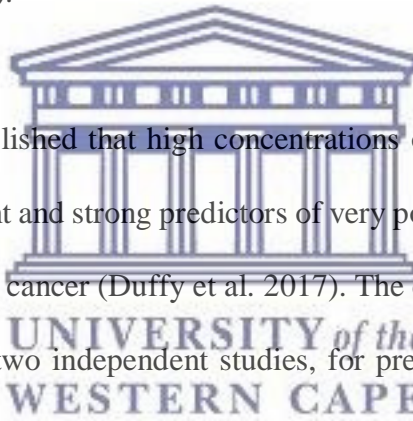


It now becomes very apparent that in the absence of a modern and well-designed prospective randomized clinical trial, there exists an insufficient amount of evidence to make recommendation either *for* or *against* the use of blood-based or serum biomarkers to monitor asymptomatic patients after being diagnosed with breast cancer (Duffy et al. 2018). Until then, the best characterised role for blood-based or serum biomarkers in breast cancer is monitoring the response to endocrine or chemotherapy of patients in advanced disease settings, of which CEA and CA 15-3 are most widely used in clinical practice (Duffy et al. 2018).

It is therefore clear that there is still a need for alternative blood-based or serum biomarkers for patients with breast cancer. Recently updated recommendations as made by ASCO and EGTM for breast tumour biomarkers, that are clinically useful, include the following as enlisted and described below:

1.3.1.1. Urokinase plasminogen activator and plasminogen activator inhibitor

To date, although not yet widely used, the urokinase plasminogen activator (uPA) and its inhibitor, plasminogen activator inhibitor 1 (PAI-1) are the most strongly validated prognostic biomarkers for breast cancer (Gouri et al. 2016; Duffy et al. 2017). uPA is a proteolytic enzyme that converts plasminogen into a proteolytically active form (plasmin); which then partakes in physiological and pathophysiological processes involved in tumour growth and metastasis (Lampelj et al. 2014). The naturally inhibiting action of PAI-1 on uPA is important for cancer progression (Lampelj et al. 2014). Roles for uPA and PAI-1 in fundamental cellular processes, such as chemotaxis, migration, invasion, adhesion, proliferation and angiogenesis, have been described (Lampelj et al. 2014).



Furthermore, it has been established that high concentrations of uPA and PAI-1 proteins, in breast tumours, are independent and strong predictors of very poor prognosis in patients newly diagnosed with invasive breast cancer (Duffy et al. 2017). The clinical value of uPA and PAI-1 has been demonstrated, by two independent studies, for predicting the outcome in lymph node-negative breast cancer (Duffy et al. 2017). Additionally, as prognostic biomarkers, high concentrations of uPA and PAI-1 are also associated with beneficial outcome of early stage breast cancer patients who have undergone adjuvant chemotherapy (Harbeck et al. 2013; Duffy et al. 2017). Finally, analytical validation of uPA/PAI-1 measurements by ELISA have been performed and as recommended by the EGTM, ‘*further research should aim to establish, validate and standardise a method for measuring uPA and PAI-1 by IHC or other techniques using formalin-fixed and paraffin embedded tumour tissue*’ (Duffy et al. 2017).

1.3.1.2. Carcinoembryonic antigen

CEA is a cell surface glycoprotein that is overexpressed in gastrointestinal, colorectal, lung and breast cancer (Bhatt et al. 2010; Kabel 2017). It is also found in normal foetal gastrointestinal tract epithelial cells (Bhatt et al. 2010; Kabel 2017). When CEA levels are elevated in breast cancer, it may be indicative of the cancer not responding to treatment or recurrence after treatment (Kabel 2017). It has been observed that elevated levels of CEA are associated with patients who have advanced or metastatic cancer as opposed to patients with localised cancer (Kabel 2017). Unfortunately, its presence in multiple cancers limits its applicability in breast cancer (Bhatt et al. 2010; Kabel 2017). CEA is being replaced by more specific tumour markers (Kabel 2017). Although it has been proposed that CEA, and CA 15-3, together with clinicopathological parameters would be more useful to accurately diagnose patients with metastatic breast cancer (Kabel 2017).



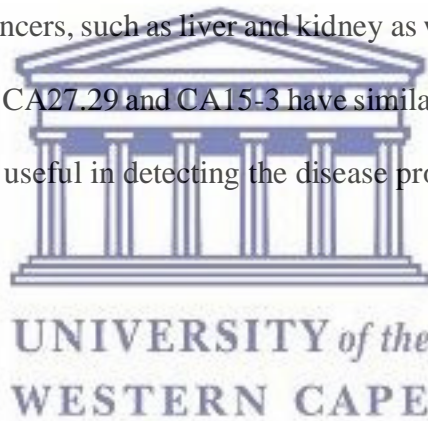
1.3.1.3. Cancer antigen 15-3

CA15-3, also known as MUC-1, belongs to the mucin-1 family of glycosylated transmembrane proteins with predicted roles in cell adhesion and signalling (Duffy et al. 2010). Under normal physiological conditions, it is expressed at the apical plasma membrane, however, during malignant transformation, the MUC-1 expression is extended to the entire membrane and the cytoplasm (Duffy et al. 2010). This is accompanied by changes in glycosylation patterns, leading to shorter carbohydrate chains that encourage shedding by proteolytic cleavage of the extracellular domain (Duffy et al. 2010). Elevated levels of CA 15-3 in the blood can be used to detect breast, ovarian, lung, pancreatic, colon and liver cancer (Kabel 2017). It has been shown to be more useful as a prognostic indicator of breast cancer and for monitoring treatment efficacy (Kabel 2017). Additionally, Darlix and colleagues demonstrated that serum CA 15-3

can act as an independent prognostic marker for metastatic breast cancer patients (Darlix et al. 2016; Kabel 2017). CA15-3, however, still lacks sensitivity for the detection of small tumours; and what this means is that the disease would be at an advanced stage at the point of detection (Duffy et al. 2010; Kabel 2017).

1.3.1.1. Breast cancer carcinoma-associated antigen

Breast cancer carcinoma-associated antigen, otherwise known as CA27.29 is a carbohydrate-containing protein antigen, produced by the MUC-1 gene (Kabel 2017). Approximately 80% of women diagnosed with breast cancer have elevated levels of CA27.29; however, this marker can also be detected in other cancers, such as liver and kidney as well as in patients with ovarian cysts (Kabel 2017). Clinically, CA27.29 and CA15-3 have similar performance in breast cancer patients and similarly, is more useful in detecting the disease progression (Kabel 2017).



1.3.1.1. Soluble Her2

Soluble Her2 (sHer2) has also, successfully, been demonstrated as a useful prognostic and predictive biomarker for breast cancer treatment, however, it has not yet shown clinical significance (Moreno-Aspitia et al. 2013).

Overall, several biomarkers have demonstrated the ability to detect tumour lesions, however, only two serum biomarkers (CEA and CA 15-3) for patients with breast cancer have been accepted in clinical use, yet their specificity and sensitivity are still in question. As such the need for more reliable prognostic and predictive breast cancer biomarkers is realised. As a result, herein this study, the use of two putative breast cancer biomarkers are proposed. These

two biomarkers, namely, Unique ligand binding protein 2 (ULBP2) and Glial cell line-derived neurotrophic factor (GDNF) family receptor alpha 1 (GFR α 1), were discovered at the University of the Western Cape, Department of Biotechnology, using bioinformatics and proteomics biomarker discovery pipelines, respectively (Ngcoza, 2012, University of the Western Cape, Department of Biotechnology; Drah, 2015, University of the Western Cape, Department of Biotechnology).

1.4. Biomarker discovery

The completion of the human genome sequencing project paved the way for a better understanding of human diseases, such as cancer (Hood and Rowen 2013). In particular, it has created a platform for which the process of biomarker discovery has greatly improved over the years, through means of functional genomics, bioinformatics and proteomics (Srinivas et al. 2002; Hood and Rowen 2013). The process of biomarker discovery involves the initial discovery, followed by validation and finally, clinical implementation (Goossens et al. 2015). And ultimately, the main objective of this process is to establish clinically useful biomarkers that enable medical practitioners to make therapeutic interventions, that leads to improved patient outcomes (Goossens et al. 2015). Several tools for cancer biomarker discovery have been described in the literature, of which the bioinformatic and proteomic approaches are highlighted in this study.

1.5. Bioinformatics as a tool for biomarker discovery

Bioinformatics can be defined as an interdisciplinary scientific field used to develop methods and software tools for storing, retrieving, organising and analysing biological data stored in biological databases (Luscombe et al. 2001). The high volume of data generated by gene

expression analysis is best-studied using Bioinformatic computational methods (Kulasingam and Diamandis 2008). These *in silico* methods can be used to analyse data generated by DNA microarray, proteomics and metabolomics expression profiling studies, which is stored in databases such as Gene Expression Atlas (<http://www.ebi.ac.uk>), ONCOMINE (<http://www.oncomine.org>), ArrayExpress (<http://www.ebi.ac.uk>) and Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>). This becomes relevant because bioinformatics is able to extract information from these high-throughput datasets and translate it into information that is useful in clinical settings (Kulasingam and Diamandis 2008).

1.6. Proteomics as a tool for biomarker discovery

Proteomics is another route that can be undertaken for biomarker discovery and can be defined as the study of the proteome (the entire set of proteins) of cells, tissue or an organism that is expressed at a given time, under defined conditions (Cho 2007). The proteome is a functional representation of almost all biological processes. Its abundance and distinct interactions are highly regulated and bridge the gap between the genome and the corresponding phenotypical presentation (Mesri 2014). Therefore, the proteome is regarded as a more accurate account of the state of the cell, tissue or organism, subsequently, proteomics is expected to be a promising strategy for identifying disease biomarkers (Cho 2007). Traditionally, two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) was used due to its ability to resolve large numbers of protein, based on the two parameters, namely, size and charge (Li et al. 2002). However, it is a labour intensive technique that requires large amounts of protein and has demonstrated a lack in reproducibility (Li et al. 2002). The limitation of gel-based approaches then lead to the use of mass spectrometry (MS), a gel-free approach, which then became the core of proteomic technology (Brewis and Brennan 2010; Hudler et al. 2014). MS is an analytical technique that ionises molecules and separates the ions based on their mass-

dependent velocities. A detector transfers this information to a computer through which an analysis can be made (Li et al. 2002; Cho 2007). The progression of proteomic technologies has created an accelerative platform for cutting-edge research to provide clinically useful information (Karley et al. 2011; Honda et al. 2013).

1.7. The identification of ULBP2 using bioinformatics

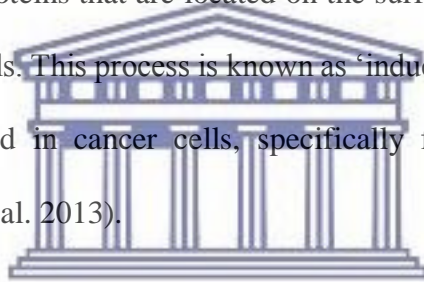
In a study by Ngcoza, bioinformatics computational methods were used to identify novel serum biomarkers for breast cancer (Ngcoza, 2012, University of the Western Cape, Department of Biotechnology). The transformation of healthy cells to cancerous cells involve changes in the expression pattern and activity of a number of distinct genes and proteins (Futsher, 2013). Therefore, the aim was to identify genes that encode for cell surface proteins that are differentially expressed in cancer cells (Ngcoza, 2012, University of the Western Cape, Department of Biotechnology). The rationale is that these proteins are exposed on the surface of cancer cells, which can potentially become circulating tumour cells (CTCs) that enter the bloodstream, after breaking away from a tumour mass during metastasis. And it is these conditions that make cell surface proteins favorable and suitable targets for identification of serum biomarkers. After the application of bioinformatic tools, a refined list of 23 genes encoding proteins up-regulated in breast cancer was identified. One of the 23 genes were found to be an FDA approved biomarker for breast cancer, while the other 22 genes were not validated for use in the detection of breast cancer. However, there is evidence in the literature that supports the non-validated genes as potential biomarkers for breast cancer (Ngcoza, 2012, University of the Western Cape, Department of Biotechnology). One of these genes, namely, Unique ligand binding protein 2 (ULBP2) was selected for further study as a biomarker for breast cancer. Subsequent validation by quantitative real-time polymerase chain reaction (qRT-PCR) showed that this gene is up-regulated in human breast cancer cell lines as compared to

non-cancerous human cell lines and some other human cancer cell lines (Ngcoza, 2012, University of the Western Cape, Department of Biotechnology).

1.8. The role of ULBP2 in breast cancer

1.8.1. NK cells and immune surveillance

Natural Killer (NK) cells play a critical role in tumour immune surveillance (Sutherland et al. 2001). NK cells kill cancer cells through selective induction of apoptosis (Levy et al. 2011). These cells use several molecular mechanisms to distinguish cancer cells from healthy cells. One of these molecular mechanisms involves NK activator ligands. NK cells are able to recognize highly expressed proteins that are located on the surface of cancer cells, but absent from the surface of healthy cells. This process is known as ‘induced self-recognition’, in which these proteins are upregulated in cancer cells, specifically for interaction with NK cell activating receptors (Raulet et al. 2013).



UNIVERSITY of the
WESTERN CAPE

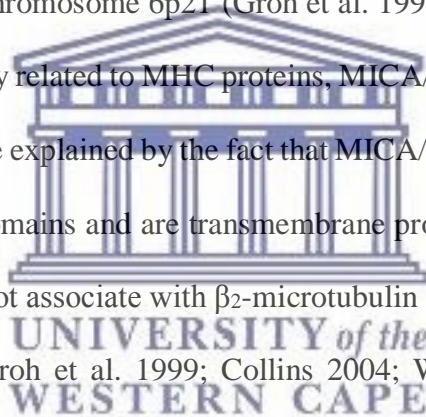
In cancer, more specifically, the NK cell molecular recognition of these upregulated proteins is attributed to the NK group 2, member D (NKG2D) activating receptor and its interaction with its ligand counterparts (Waldhauer and Steinle 2008). Upon interaction with the NKG2D ligands, the NKG2D conveys an activating signal by means of its interaction with the adaptor transmembrane DAP10. The DAP10 protein structure consists of a YINM-motif which is important for co-stimulation of pathways that promote the cytotoxic elimination of the target cells (Chitadze et al. 2013; Raulet et al. 2013). Of note, NKG2D signaling, alone, appears to be insufficient for effective cytotoxic effects prompted by primary human NK cells. Further research into the cytotoxic mechanism presented by the DAP10-YINM interaction is needed (Bryceson et al. 2009; Chitadze et al. 2013). Nonetheless, several studies have demonstrated

the importance of the NKG2D and its' ligands in immune surveillance in cancer (Chitadze et al. 2013).

1.8.2. NKG2D ligands

1.8.2.1. MHC class I chain-related molecule A/B

In humans, there are two NKG2D ligand families that resemble the structure of major histocompatibility complex (MHC) class I chain-related (MIC) molecules, namely, MICA and B (MICA/B); and the Unique ligand 16 binding proteins (ULBPs) (Waldhauer and Steinle 2008; Champsaur, M. and Lanier 2011; Raulet et al. 2013). MICA/B are encoded by genes mapped to the human MHC chromosome 6p21 (Groh et al. 1999; Collins 2004; Sutherland et al. 2001). Although structurally related to MHC proteins, MICA/B are distinct in their function (Groh et al. 1999). This may be explained by the fact that MICA/B have an ectodomain derived from the MHC class I α 1-3 domains and are transmembrane proteins. However, unlike MHC class I proteins, MICA/B do not associate with β 2-microtubulin and has neither been shown to present antigenic peptides (Groh et al. 1999; Collins 2004; Waldhauer and Steinle 2008). Although MICA/B encode for different transcripts, these two proteins are functionally indistinguishable; which is further confirmed by the fact that they share 83 % amino acid homology (Groh et al. 1999; Collins 2004). There has, however, been evidence of MICB, but not MICA, binding to the UL16 glycoprotein. The literature suggests that this interaction may be due to functional differences or the nature of MICB's polymorphism (Sutherland et al. 2001; Cosman et al. 2001).



1.8.2.2. UL16 binding proteins

The ULBP binding protein family consists of six ULBP members, namely, ULBP1 – 6. They are encoded within a gene cluster mapped to chromosome 6 (6q24.2 – 6q25.3) (Champsaur, M. and Lanier 2011). They are distantly related to the MHC molecules but structurally possess the $\alpha 1$ and $\alpha 2$ extracellular domains. ULBPs, however, lack of the $\alpha 3$ domain and as with the MICA/B proteins, they do not associate with $\beta 2$ -microtubulin (Sutherland et al. 2001). Furthermore, in contrast to MICA/B, ULBP 1-3 are cell surface glycosylphosphatidylinositol (GPI) anchored, rather than transmembrane proteins (Song et al. 2006).

A very interesting phenomenon is seen upon interaction of NK cells with cell surface NKG2D ligands versus their soluble counterparts. Although high expression of the cell surface NKG2D ligands induces oncogenesis, its expression favours NK cell-induced cell death (Zhang et al. 2015). However, cleavage of these cell surface ligands result in increased expression of their soluble forms that directly bind to the NKG2D receptor on NK cells; thereby, preventing the cytotoxic action against tumour cells presenting cell surface NKG2D ligands (Sutherland et al. 2002; Song et al. 2006; Zhang et al. 2015).

Overexpression of the soluble forms of both MICA/B and ULBPs have been reported in the serum of cancer patients. Moreover, higher expression of the soluble NKG2D ligands are indicative of tumour progression and thus poor prognosis (Huergo-Zapico et al. 2014).

1.8.2.3. ULBP2 as a prognostic cancer biomarker

The ULBP2 protein is a member of the NKG2D ligand family that is expressed in response to stressful conditions, such as infections or tumour transformation (Raulet et al. 2013). Several

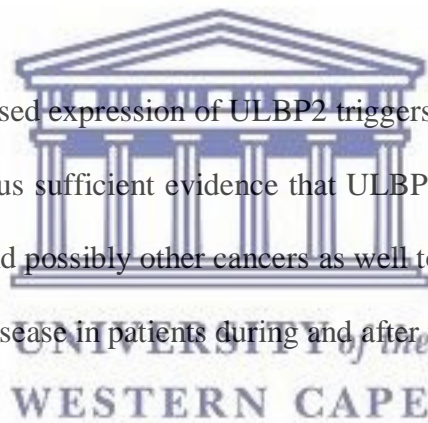
studies suggest that in an effort to evade tumour immune surveillance, ULBP2 is cleaved from the cell surface of cancer cells by various proteases, which are also overexpressed in cancer cells (Groh et al. 2002; Waldhauer and Steinle 2006). ULBP2, together with other NKG2DL's have been implicated as serum biomarkers, demonstrating increased sensitivity and accuracy of detecting cancer (Chang et al. 2011). Chang and colleagues reported on the development of a high-throughput biosensor for the detection of diluted soluble ULBP2 (sULBP2) in pancreatic cancer (Chang et al. 2011). The biosensor was able to detect sULBP2 at 16-18pg/ml in both 1% BSA-PBS (bovine serum albumin – phosphate buffered saline) and a 10-fold diluted human serum (Chang et al. 2011). Furthermore, they demonstrated that under the same conditions, the biosensor developed in their study was 100-fold more sensitive than limit of detection observed by the traditional enzyme-linked immunosorbent assay (ELISA) method (Chang et al. 2011).



A study by Li and colleagues investigated the clinical significance of MICA/B and ULBP2 expression in ovarian cancer (Li et al. 2008). The study population consisted of n=82 patients with ovarian cancer and n=6 patients without ovarian cancer, who were enrolled in the study between 1993 – 2003 (Li et al. 2008). Using immuno-histochemical staining, the expression of MICA/B and ULBP2 and were evaluated in both cancerous and normal ovarian tissue (Li et al. 2008). It was found that significantly higher expression levels of MICA/B and ULBP2 could be detected in the cancerous ovarian tissue as compared to the normal tissue (Li et al. 2008). In particular, high expression of ULBP2 was correlated to a decline in intraepithelial infiltration of T cells, which is important for prolonged survival of ovarian cancer patients, linking ULBP2 to potentially to T cell dysfunction in the tumour microenvironment and poor prognosis in ovarian cancer and (Li et al. 2008).

de Kruijf and colleagues then went on to investigate the prognostic potential of NKG2D ligands on breast cancer patients (de Kruijf et al. 2012). Their study consisted of a population size of n=677 breast cancer patients that had undergone surgery between 1985 and 1994 (de Kruijf et al. 2012). Formalin-fixed paraffin-embedded tumour tissues were immuno-histochemically stained, using antibodies directed at all the NKG2DL's, except ULBP6 (de Kruijf et al. 2012). The results demonstrated that the increased expression of MICA/B and ULBP2, resulted in a statistically significant beneficial outcome for breast cancer patients with regards to a longer relapse-free period (de Kruijf et al. 2012). And furthermore, the co-expression of MICA/B and ULBP2 has demonstrated to have prognostic value that can be used, independently, from known clinico-pathological parameters (de Kruijf et al. 2012).

The rationale is that the increased expression of ULBP2 triggers immune responses that result in poor prognosis. There is thus sufficient evidence that ULBP2 can be used as a prognostic biomarker for breast cancer and possibly other cancers as well to inform medical practitioners about the progression of the disease in patients during and after therapeutic intervention.



1.9. The identification of GFR α 1 using proteomics

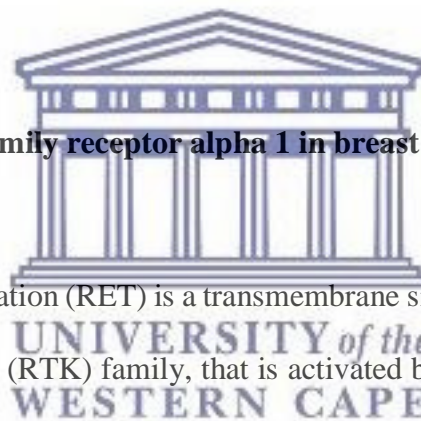
In a study by Drah, 2015, a quantitative proteomic approach was used to identify novel biomarkers for breast cancer (Drah, 2015, University of the Western Cape, Department of Biotechnology). The aim was to identify cell surface proteins that are overexpressed in breast cancer cells. This was accomplished using stable isotope labelling by amino acids in culture (SILAC), together with liquid chromatography tandem mass spectrometry (LC-MS/MS), to perform a comparative analysis of the expression levels of membrane proteins present in a human breast cancer cell line (MCF7) and a non-cancerous human breast cell line (MCF12A). This led to the identification of the glial cell line-derived neurotrophic factor (GDNF) family

receptor alpha 1 (GFR α 1) as a potentially new biomarker for breast cancer. In this study, the proteomics data revealed that GFR α 1 is significantly overexpressed in MCF7 cells as compared to MCF12A cells (Drah, 2015, University of the Western Cape, Department of Biotechnology). This was further validated by using both qRT-PCR and IHC, representing the gene and protein encoded product, respectively. In both cases, the expression of the GFR α 1 corroborated with the proteomics data. The IHC data was performed using tissue from three breast cancer patients at stage I, II and III, respectively and revealed that the expression of GFR α 1 was upregulated in all three stages and its' expression was observed in the myoepithelial layer of the milk duct. Furthermore, the results suggest that the expression of GFR α 1, in the cancerous breast tissues, increased from stage I to III (Drah 2015).

1.10. The role of GDNF family receptor alpha 1 in breast cancer

1.10.1. RET signalling

REarranged during Transformation (RET) is a transmembrane signaling receptor, which is part of the receptor tyrosine kinase (RTK) family, that is activated by the GDNF family of growth factors (Boulay et al. 2008). The GDNF family consists of neurturin (NRTN), artemin (ARTN), persephin (PSPN) and GDNF (Bordeaux et al. 2000). To activate RET, these ligands are first required to bind the GPI-linked receptors (GFR α 1-4) which in turn recruits RET to the lipid rafts, resulting in an active macromolecular signaling complex (e.g. GDNF/GFR α 1/RET) (Morandi et al. 2011; Fleming et al. 2015; Groot et al. 2016). Activation of the RET receptor results in its dimerization and subsequent autophosphorylation, leading to a cascade of signaling events (**Figure 1.3**) that are essential for controlling cell survival, differentiation, proliferation and migration (Esseghir et al. 2007; Morandi et al. 2011).



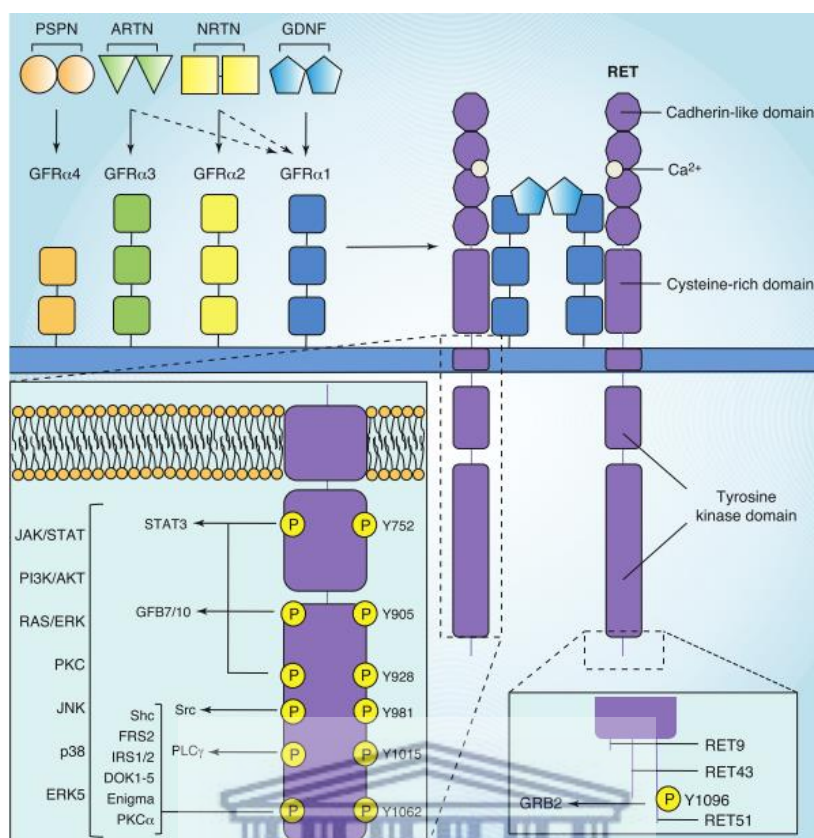


Figure 1.3. Activation of the various RET signalling pathways.

1.10.2. GDNF/GFRα1 signalling

The GDNF ligands preferentially interact with their corresponding GFRα's, as depicted in **Figure 1.3**. Even though the GFRα's can interact with more than one GDNF ligand, GDNF mainly interacts with GFRα1 with high affinity and thus the GDNF induced RET signaling pathway will be the focus of this discussion. The GFRα1 protein is an encoded gene product mapped to chromosome 10q25 (Angrist et al. 1998). The structure of the mature GFRα1 consists of three homologous cysteine-rich domains (D1, D2 and D3), having C-termini extending at different lengths. D1 acts as a stabilizing structure for the interaction between GDNF and GFRα1, whereas D2 and D3, collectively acts as the binding site for GDNF ligand and stabilizes the interaction of GDNF/GFRα1 with RET (Leppa et al. 2004; Virtanen et al. 2005; Parkash et al. 2008). Furthermore, the GFRα1 protein can activate RET through two

mechanisms, the first is when GFR α 1, bound by GDNF, is present on the same cell as RET (cis signalling) or when soluble GFR α 1, generated by shedding, is released into the extracellular space, binds GDNF and subsequently interacts with RET on neighbouring cells (trans signalling) (Fleming et al. 2015; Maheu et al. 2015). Although GFR α 1 is able to bind other GDNF ligands, neither GFR α 1 nor GDNF is able to activate RET independently (Groot et al. 2016).

The GDNF protein is a potent neurotrophic factor secreted by neurons and is known to promote the survival, neurite outgrowth and differentiation of neurons (Ledda et al. 2007). A study by Esseghir and colleagues used MCF-7 (human breast adenocarcinoma) cells, known to overexpress both GFR α 1 and RET receptors, as a model for breast cancer and demonstrated that treatment with GDNF promotes proliferation, survival and scattering (a phenotypical indication of cell motility) in MCF-7 cells (Esseghir et al. 2007). Furthermore, in neurons, it has been shown that inflammatory cytokines induce the expression of GDNF and since GDNF was observed to act as a trophic factor in MCF-7 cells, its source in breast tumour tissue was investigated. It was found that under normal culture conditions, significantly low amounts of GDNF was detected in infiltrating fibroblasts in MCF-7 xenografts. However, upon treatment with TNF- α or IL-1 β , GDNF expression was significantly upregulated in MCF-7 cells (Esseghir et al. 2007). This study revealed a role for the GDNF induced RET (GDNF-RET) signalling complex in breast cancer.

1.10.3. Implications of GDNF-RET signalling in breast cancer

The driving force of tumour growth in breast carcinomas is best characterised by the overexpression of tyrosine kinases, such as Her2 and EGFR (Morandi et al. 2011; Morandi and Isacke 2014). There is an increasing body of evidence implicating GDNF-RET signalling in

breast cancer, in which RET is another tyrosine kinase that is shown to be highly expressed in the ER⁺ subset of breast cancers (Morandi et al. 2011; Morandi and Isacke 2014).

In a study by Boulay and colleagues, an ER α ⁺ breast cancer model was used to demonstrate that the GDNF-RET signalling pathway, in both the *cis* and *trans* mode, activates RET autophosphorylation and subsequent activation of the ERK and c-Jun-NH₂-kinase (JNK) pathways (Boulay et al. 2008). MCF-7 cells (GFR α 1⁺/RET⁺) were used to model the *cis*-activating mode and T47D (Human metastatic breast) cells (GFR α 1⁻/RET⁺) were used to model the *trans*-activating mode, by supplemental treatment with soluble GFR α 1 (Boulay et al. 2008). Both cell lines were stimulated with GDNF and a time-dependent activation of ERK phosphorylation was observed for both MCF-7 (after 5 min) and T47D (after 30 min) cell lines. The ERK pathway is known to facilitate proliferation and migratory properties, in association with RET, in thyroid tumour (Boulay et al. 2008). In a similar manner, JNK phosphorylation was observed in both cell lines. The JNK pathway has been implicated in RET-associated mitogenic and motility properties (Boulay et al. 2008). Therefore, together, the results of this study suggests a role for GDNF-RET signalling to confer mitogenic, proliferation and migratory properties in an ER⁺ subset of breast cancer (Boulay et al. 2008).

Endocrine therapy is the most effective route of treatment for ER⁺ breast carcinomas, however, the endocrine resistant phenotype is often induced by activation of ER α in an estrogen-independent manner (Plaza-Menacho et al. 2010). In a study by Plaza-Menacho and colleagues, estrogen-independent expression of the ER α receptor, in MCF-7 cells, was shown to be upregulated by activation of the GDNF-RET signalling pathway (Plaza-Menacho et al. 2010). Even though this implicates GDNF-RET signalling with the endocrine resistant phenotype, RET was identified as a potential therapeutic target in tamoxifen-resistant (TAM^R-1) ER α ⁺

breast tumours (Plaza-Menacho et al. 2010). Two RET-targeted siRNA oligonucleotides were used to down-regulate RET expression in TAM^R-1 MCF-7 cells and as expected, upon GDNF stimulation the cells were sensitised to tamoxifen treatment; whereas the mock knockdown cells displayed resistance. This study suggests that down-regulation of the GDNF-RET signalling confers sensitization of ER α ⁺ breast cancer cells to tamoxifen treatment (Plaza-Menacho et al. 2010). GFR α 1 is required for GDNF-RET signalling to confer an endocrine resistant phenotype and therefore, ER⁺ breast cancer patients displaying GFR α 1 upregulation, together with upregulation of the RET co-receptor could provide medical practitioners with critical information regarding treatment response to tamoxifen.

In 2013, with the advent of investigating RET as a potential therapeutic target, an extensive study by Gattelli and colleagues demonstrated that the inhibition of GDNF-RET signalling decreases growth and metastatic potential of ER α ⁺ breast cancer cells (Gattelli et al. 2013). In this study, it was shown that RET activation, via the GDNF-RET signalling pathway, upregulates pro-inflammatory cytokines during endocrine treatment (Gattelli et al. 2013). Implying that despite endocrine therapy in ER α ⁺ breast carcinoma, the metastatic potential still exists if the GDNF-RET signalling is active (Gattelli et al. 2013). Interleukin 6 (IL6) was one of the cytokines identified in the transcriptomic profile of aromatase expressing MCF-7 cells stimulated by GDNF and estradiol precursors; and subsequent treatment fulvestrant (Gattelli et al. 2013). Fulvestrant exerts its inhibitory action on the estrogen receptor, resulting in the activation of downstream events that accelerate estrogen degradation (Osborne et al. 2004). This specific treatment conditions revealed that IL6 was associated with genes involved in cellular movement and lead to the investigation of GDNF and IL6 stimulation in migration assays (Gattelli et al. 2013). It was shown that GDNF and IL6 promote migration, both *in vitro* and *in vivo* (Gattelli et al. 2013). Additionally, knockdown of RET resulted in reduced breast

tumour outgrowth and metastatic potential (Gattelli et al. 2013). What this means is that the GDNF-RET signalling pathway has shown to induce ER⁺ breast cancer progression by promoting its the metastatic potential. This is in turn suggests that monitoring ER⁺ breast cancer patients treated with fulvestrant by examining the expression of proteins in the GDNF-RET signalling complex, such as GFR α 1, RET and IL6 could be indicative of poor prognostic outcome.

Bhakta and colleagues looked at the development of an anti-GFR α 1 antibody drug conjugate (ADC) for the targeted treatment of the luminal A breast cancer subtype (hormone-receptor positive) (Bhakta et al. 2018). Both RNAseq and IHC (n=40 specimens) showed significant upregulation of GFR α 1 expression in the luminal A breast cancer tissues as compared to minimal or no GFR α 1 expression in the non-cancerous tissue (Bhakta et al. 2018). In this study both *in vitro* and *in vivo* models demonstrated a target-dependent toxicity of the anti-GFR α 1 ADC in two luminal A breast cancer models (MCF7 and KPL-1 cell line derived tumour xenografts) (Bhakta et al. 2018). The authors of this study recommend the potential clinical utility of the anti-GFR α 1 ADC as a target-specific therapeutic for the large number of patients relapsing from the luminal A type breast cancer (Bhakta et al. 2018).

The rationale is that the increased expression of GFR α 1 promotes the progression of ER⁺ breast cancers that result in poor prognosis. Suggesting that GFR α 1 can be used as a prognostic biomarker for ER⁺ breast cancer to inform medical practitioners about the progression of the disease in patients during and after therapeutic intervention. Furthermore, GFR α 1 has also demonstrated the potential to be used together with GDNF and RET as predictive biomarkers for response of ER⁺ breast cancer patients to tamoxifen and fulvestrant.

1.11. Limitation of current methods of detection for breast cancer

For hormone receptor positive breast cancer patients, the recommended methods of detection for prognosis and selection of a suitable treatment regime are IHC and the enzyme-linked immunosorbent assay (ELISA) (Duffy et al. 2018). IHC can be described as a technique used to detect the expression of proteins, using an antibody that binds to specific molecules present in disease cells (Zaha 2014). ELISA is a biochemical assay that uses antibodies to detect the presence of a protein in a liquid sample (Selman et al. 2011). In particular, IHC, has become the gold standard for tumour diagnosis, however, due to the use of antibodies several challenges exist for this method (Bauer et al. 2016). It is a technique that, conventionally, uses specific antibodies to stain and detect antigens in cell or tissue (de Matos et al. 2010). These antigens ranges from amino acids to proteins, infectious agents and specific cellular population (de Matos et al. 2010). IHC is performed in two phases, 1) slide preparation and stages involved in the reaction and 2) the interpretation and quantification of the obtained results (de Matos et al. 2010). This is a very powerful tool as it can provide insight into the subtyping of neoplasias, determination of the primary site of malignant neoplasias and furthermore, it is able to discriminate between benign and malignant neoplasias (de Matos et al. 2010).

In a clinical setting, IHC has diagnostic, prognostic and therapeutic significance; which in turn has a direct impact on managing the patients' treatment decision and strategies (Bauer et al. 2016). Currently, antibodies are the only commercially recognised and clinically validated molecular probe used for IHC applications, however, several limitations still exist in their use, particularly, for application in paraffin-embedded tissues (Bauer et al. 2016).

Antibodies have, routinely, been applied to IHC for more than 40 years (Bauers et al. 2016). Both monoclonal and polyclonal antibodies are employed, although monoclonal antibodies are

more preferred (Bauers et al. 2016). This preference arises from the fact that monoclonal antibodies bind to a single epitope in an antigen and are thus more specific than their polyclonal counterparts that will react with a variety of epitopes in a single antigen, resulting in lower specificity (Bauer et al. 2016; Kim et al. 2016). In particular, the specificity of the primary antibodies is crucial for performing a successful IHC assay and although routinely used, it has been reported that poor antibody selection and cross-reactivity results in ~42% of the diagnostic inconsistencies observed ((Bauer et al. 2016). Thus, it is clear that either alternatives to increase the specificity of the antibody production itself is necessitated or an alternative bio-recognition moiety should be employed. In this regard, aptamers have emerged as novel bio-recognition molecules with several advantages over antibodies (Bauer et al. 2016).

1.12. Advantages of aptamers over antibodies

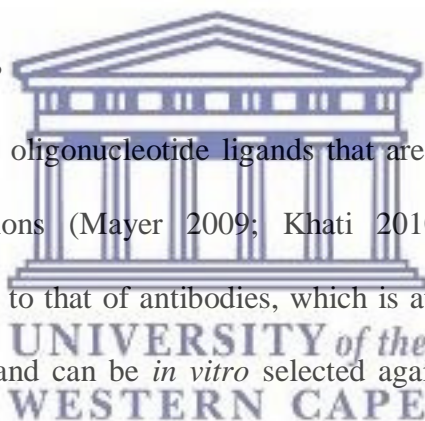
Aptamers are small, synthetic oligonucleotide ligands that are able to bind its target with high affinity and specificity (Mayer 2009; Khati 2010; Zeng et al. 2010) Specifically, as an alternative bio-recognition moiety to antibodies, aptamers are demonstrating increasing application on the biotechnological front; which can be ascribed to the key features that show clear advantages over antibodies (Thiviyanathan and Gorenstein 2012). These include increased stability, since aptamers are able to maintain their structural integrity even after a series of denaturation/renaturation cycles (Thiviyanathan and Gorenstein 2012). Consequently, aptamers are still able to bind their target once their native conformation is recovered and this attribute make aptamers applicable in a wider range of assay conditions (Thiviyanathan and Gorenstein 2012). Whereas antibodies are irreversibly denatured at elevated temperatures, which, in turn, results in its inability to recognise and bind its target (Thiviyanathan and Gorenstein 2012). Aptamers, therefore, have a longer shelf life (Bruno 2015). Aptamers are significantly smaller than antibodies, generally ranging between 6 to 30 kDa whereas

antibodies are generally about 150 kDa. This feature is important because its' small size is advantageous for tissue penetration (Santosh and Yadava 2014; Ku et al. 2015). Because aptamer sequences can be chemically synthesised, batch-to-batch variation, which is often experienced with antibodies, is eliminated (Song et al. 2012; Santosh and Yadava 2014). Additionally, chemical synthesis makes aptamers more readily available through cost effective large scale production (Song et al. 2012; Santosh and Yadava 2014). On the premise of the above-mentioned advantages, aptamers show great potential as an alternative biorecognition moiety that can be applied to IHC and immunoassays used for breast cancer detection.

1.13. Aptamers

1.13.1. What are aptamers?

Aptamers are small, synthetic oligonucleotide ligands that are able to take on well-defined three-dimensional conformations (Mayer 2009; Khati 2010). They exhibit molecular recognition properties, similar to that of antibodies, which is attributed to their high affinity and specificity (Khati 2010) and can be *in vitro* selected against a large variety of targets (Hamula et al. 2006), such as proteins (Khati 2010), drugs (Mairal et al. 2008) and other molecules (Melorose et al. 2015). Systematic evolution of ligands by exponential enrichment (SELEX) (Sampson 2003) is a stringent *in vitro* selection process by which aptamers, having a high affinity for a particular target, are progressively isolated from random oligonucleotide libraries under specific conditions. (Srisawat and Engelke 2001; Gopinath 2007; Mairal et al. 2008; Chen et al. 2011; Khati 2010).



1.13.2. SELEX libraries

1.13.2.1. Library complexity

The most commonly used SELEX library consists of 10^{13} - 10^{15} nucleic acid sequences and these sequences contain a central random region, usually 60 – 90 nucleotides in length; which is further flanked by fixed primer regions on each side, usually around 20 nucleotides in length (Blind and Blank 2015). To ensure that the selection has an optimal start, the design of the library is carefully considered based on the length of the random region, the type of randomization and the chemical modifications of the library (Gopinath 2007; Wang 2008; Blind and Blank 2015). It has been demonstrated that a subsequence or motif within the aptamer generally facilitates the binding interaction with the target molecule and from this premise the rationale is that the random region should contain an equal distribution of all four nucleotides (NT) bases (Blind and Blank 2015). In doing so, an equal distribution of the subsequence or motif increases the maximum diversity, also known as the sequence space, which in turn could potentially generate aptamers with favourable binding capacities (Sampson 2003; Wang 2008; Blind and Blank 2015). However, in practice, limitations exist to screen more than 10^{16} random sequences in one library. Consequently, the most successfully documented and frequently employed SELEX generally begins with 10^{13} - 10^{15} single stranded DNA/RNA molecules (Hamula et al. 2006; Gopinath 2007; Svobodova et al. 2012).

1.13.2.2. Library chemistry

The interaction of the aptamers with the target is dependent on the chemical stability of the DNA or RNA, which translates into the structural stability of the aptamers (Mairal et al. 2008; Sampson 2003). Oligonucleotides (ssDNA and RNA), formed by natural bases, are generally susceptible to enzymatic and chemical degradation. RNA, in particular, at pH greater than

neutral is highly reactive in the presence of nucleases due to the presence of the 2' hydroxyl group on the ribose moiety. Although RNA libraries can potentially provide a more structurally diverse starting library for the generation of aptamers, due to its' instability, DNA libraries are employed more often (Sampson 2003; Svobodova 2012). Some strategies have been developed in an attempt to address this issue; these include 1) modifications of the nucleotide bases, for example, modification of the 5'-position of pyrimidines with Br, Cl, I, N₂ and NH₃ and the 2' position with OCH₃, F and NH₂ (Sampson 2003). Although these modifications provide stability to the oligonucleotides, it is important to consider the compatibility of these modifications with the SELEX protocol (Mairal et al. 2008).

Alternatively, 2) the phosphodiester backbone can be modified via the introduction of alpha-thio substituted deoxynucleoside triphosphates. This approach, however, has demonstrated to offer better results for DNA as compared to RNA (Mairal et al. 2008). And finally, 3) another approach is the use of mirror-images of the aptamers, also known as Spiegelmers; which is based on the concept of chirality presented in the centre of naturally occurring deoxyribose sugar moieties of nucleotides (D-nucleotides) (Mairal et al. 2008). This approach begins with the generation of a mirror-image of the target (D-amino acids), for which aptamers (D-oligonucleotides) are selected against; to which end spiegelmers (L-oligonucleotides) are then synthesised (Mairal et al. 2008). As described by the law of symmetry and further verified by studies using this approach, these chiral converted aptamers (spiegelmers) are able to exhibit comparable binding to the natural target (L-amino acids), but are not recognised by nucleases and consequently the spiegelmers are not susceptible to degradation (Sampson 2003; Kulbachinskiy 2007; Mairal et al. 2008; Vater and Klussmann 2015). An example of this approach was demonstrated by Purschke and colleagues, in which DNA spiegelmer aptamers were developed against the domain of the Staphylococcal enterotoxin B protein (Purschke et

al. 2003). In a recent review by Vater and Klussmann 2015, further examples of successfully developed spiegelmers are enlisted. The major drawback of this approach is that the generation of large/complex D-peptides may be challenging, which is a crucial step in the aptamer selection process (Purschke et al. 2003; Vater and Klussmann 2015).

Depending on the application of the aptamers, modified RNA aptamers or ssDNA aptamers may be used. More specifically, Rotherham and colleagues report that aptamers selected for diagnostic applications have favoured DNA aptamers due to intrinsic stability, a simpler process of enrichment and lower in production costs (Orava et al. 2010; Rotherham et al. 2012). Thus, for the purpose of this study, the focus is directed at DNA aptamers.

1.13.3. SELEX: The process

In principle, SELEX comprises of three steps (Figure 1.4), that is, (1) adsorption: selecting ligand sequences, from a random oligonucleotide library, that binds to the target, (2) partitioning: separation of the bound from the unbound sequences using affinity techniques and (3) exponential enrichment: amplification of those sequences that bind the target (Hamula et al. 2006; Gopinath 2007). These steps are repeated until the affinity saturation is reached, usually between 8 to 16 cycles. At this point, the ligand sequences with the highest affinity for the target have been enriched and are isolated for sequence identification (Mairal et al. 2008; Wang 2008; Lakhin et al. 2013).



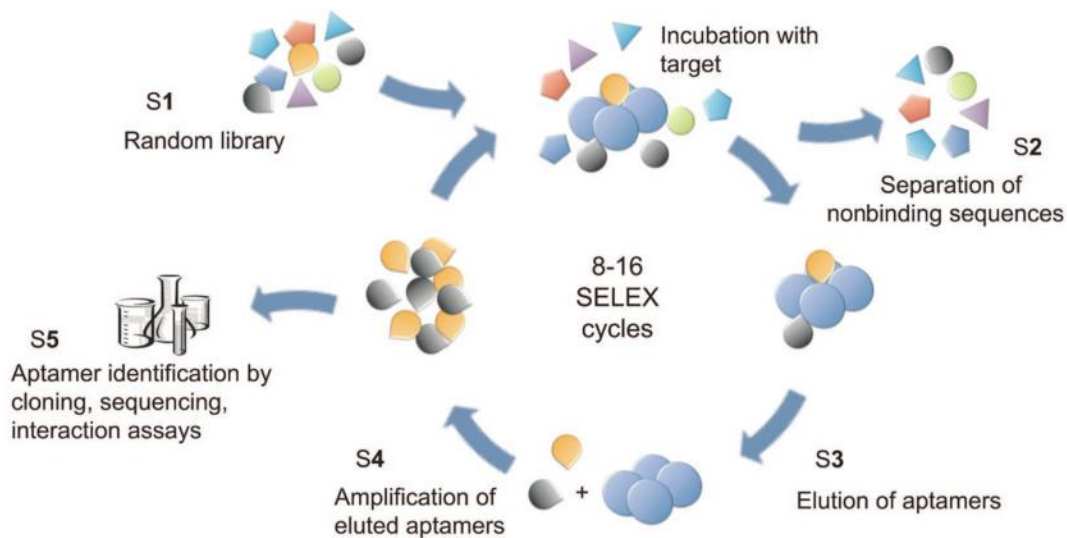


Figure 1.4. A schematic diagram of a typical SELEX process. (Blind and Blank 2015).

1.13.3.1. Adsorption

Once an appropriate library has been designed, the library is incubated with the target under a specific set of conditions, that is, an appropriate buffer and temperature (Sampson 2003; Gopinath 2007). At this point, usually, a small percentage of the sequences within the library population interact with the target and are representative of those sequences that have the strongest affinity for the target (Jayasena 1999; Sampson 2003). This step does not usually present any challenges, however, the conditions under which the selection begins is important and may vary in later rounds to strategically create conditions that would favour the enrichment of high-affinity sequences (Sampson 2003; Gopinath 2007; Stoltenburg et al. 2007). The bound sequences generated during this step is then separated from the unbound sequences using a partitioning method of choice (Santosh and Yadava 2014).

1.13.3.2. Separation methods

Separating the target bound sequences from the library is one of the critical steps during the selection process. Several methods exist to accomplish, as reviewed in detail by Gopinath, 2007 and showed in **Table 1.2** (Gopinath 2007).

The most commonly used method is separation by means of nitrocellulose membranes, which typically involves the incubation of the library with the target prior to applying the mixture to the membrane and the unbound sequences are, thoroughly, washed from the membrane and the bound sequences are amplified by PCR and used for subsequent rounds of selection. However, this approach has demonstrated low separation efficiencies and requires large quantities of the library and its' target in the incubation step. The aptamers generated using this method have K_D values within the micromolar and nanomolar range (Hamula et al. 2006; Gopinath 2007).

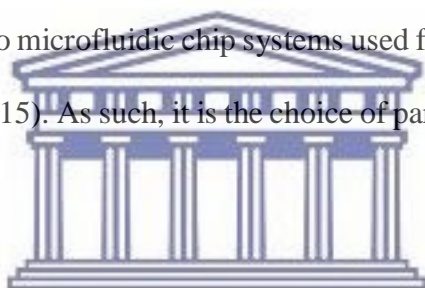
Table 1.2. Separation methods used for SELEX.

Target	Separation method	Aptamer	Reference
Lysozyme	Affinity beads	RNA	Yamamoto et al. 1998
A/Panama influenza virus	Surface plasmon resonance	RNA	Gopinath et al. 2006
YPEN-1 rat endothelial cells	Flow cytometry	DNA	Blank et al. 2001
Hepatitis B virus RNA encapsidation	Antibody-based	RNA	Rieger and Nassal, 1995
IgE	Capillary electrophoresis	DNA	Mendonsa and Bowser, 2004
Neutrophil elastase	Gel mobility shift	RNA	Smith et al. 1995

Capillary electrophoresis (CE) has shown the best improvement, with a high selection efficiency, low non-specific interactions of the DNA sequences to the capillary and a significant reduction in the number of selection rounds (as little as four) required to produce

high affinity aptamers and K_D values in the low nanomolar to low picomolar range. The only shortfall is that injection volumes are very low (nanolitres) which limits the initial amount of the library and translates into a lower number of possible sequences available to be selected from (Hamula et al. 2006; Gopinath 2007; Blind and Blank 2015).

Even though the use of magnetic beads an alternative partitioning method has not illustrated the same capacity of CE, it has, nonetheless, shown an improvement in comparison to nitrocellulose membranes; with aptamers having K_D values in the low nanomolar range and the additional advantage of being able to use much less of the target; which is directly immobilized onto the beads using affinity tags. It is also a very simple and rapid method that may even be miniaturised and integrated into microfluidic chip systems used for SELEX-on-a-chip (Hamula et al. 2006; Blind and Blank 2015). As such, it is the choice of partitioning the bound sequences for the purpose of this study.



1.13.3.3. Enrichment and ssDNA generation

To effectively complete SELEX, PCR is used to generate size specific double-stranded (dsDNA) of the target bound sequences, which is then used as the template to produce ssDNA for the successive rounds of SELEX. This is another crucial step in the SELEX process since the purity and quantity of the ssDNA significantly affect the enrichment of sequences against the target (Citartan et al. 2012; Svobodova 2012). Additionally, aptamers as ssDNA molecules are implicated in its folding proficiency, as previously described, which is necessary for high affinity and specific interactions with the target (Citartan et al. 2012; Tabarзад et al. 2014). A few methods exist for generating ssDNA, these include 1) magnetic separation by means of streptavidin coated beads, 2) DNA strand separation under denaturing-urea electrophoresis, 3)

asymmetric PCR and 4) lambda enzymatic digestion (Stoltenburg et al. 2007; Citartan et al. 2012; Ozer et al. 2014; Tabarzad et al. 2014).

The streptavidin-bead based method is achieved by performing PCR using a biotinylated primer that becomes incorporated into one strand of the dsDNA which is then captured on streptavidin-coated beads. Following thermal or chemical denaturation, the unlabeled strand is eluted and used in the subsequent rounds of selection (Stoltenburg et al. 2007). Another approach is size separation using polyacrylamide gel electrophoresis (PAGE). PCR amplification of the bound sequences is performed using a specifically designed extended primer that results in ssDNA products with different lengths. These ssDNA products are then subjected to PAGE, viewed under UV light and the strand size of interest is purified from the gel (Svobodova et al. 2012). Asymmetric PCR is another alternative, which consists of using skewed concentrations of the forward and reverse primers during amplification resulting in ssDNA (Tabarzad et al. 2014). Another method is the use of exonuclease digestion. One example of this approach is modifying the 5' end of one of the primers with phosphorothioates. Once amplified, the resultant PCR product constitutes one DNA strand that is hydrolytically protected and as such, when the dsDNA PCR product is subjected to digestion by the T7 Gene 6 exonuclease, only the unmodified strand is hydrolysed in the 5'-3' direction and ssDNA is produced (Svobodova et al. 2012). Moreover, Svobodova and colleagues demonstrated that the combination of asymmetric PCR and exonuclease digestion can achieve increased efficiency in the generation of ssDNA. Therefore, for the purpose of this study, this approach will be employed.

1.13.4. Modifications of the selection process

Above, the conventional SELEX process was described, however, it is important to note that since its discovery, modifications at specific points within the process have been implemented to improve the selection of aptamers with higher affinities and specificity to the target (Stoltenburg et al. 2007; Aquino-Jarquín and Toscano-Garibay 2011; Sun and Zu 2015). **Table 1.3** summarizes the variations in the SELEX process and the advantages the application of each of these method offers (Stoltenburg et al. 2007). Of the list described, negative and counter-SELEX was applied in the SELEX protocol in this study.

Table 1.3. Different types of SELEX methodologies.

	Description	Reference
Negative SELEX	Minimises the co-selection unwanted nucleic acids.	Blank et al. 2001
Counter SELEX	Generate aptamers that are able to discriminate between closely related structures.	Lee and Lee, 2006
Blended SELEX	To give additional properties to the aptamers beyond the binding capability.	Radrizzani et al. 1999
Photo SELEX	Aptamers bearing photo reactive groups that bind and photo cross-link with targets and/or photo activate a target.	Golden et al. 2000
Genomic SELEX	Construction of a SELEX library of an organisms' genome.	Shimada et al. 2005
Toggle SELEX	Switching between targets during alternating selection rounds.	White et al. 2001
CE SELEX	The use of capillary electrophoresis for separation.	Tang et al. 2006
Non-SELEX	A process that involves repetitive steps of partitioning with no amplification in between	Berezovski et al. 2006
EMSA SELEX	The use of electrophoretic mobility shift for separation during each round of SELEX.	Tsai and Reed, 1998
Tailored SELEX	Integrated method to identify aptamers with only 10 fixed nucleotides through ligation and removal of primer binding sites within the SELEX process.	Vater et al. 2003
Deconvolution SELEX	To generate aptamers for complex targets.	Blank et al. 2001

1.13.5. Aptamer processing

Screening for potential aptamers with the highest binding affinity may be categorised in two phases, that is, 1) identifying the full-length aptamer and then 2) identifying the minimal binding sequence (MBS) of the aptamer, that is required to bind the target, while maintaining the integrity of the binding affinity (Le et al. 2014; Ku et al. 2015).

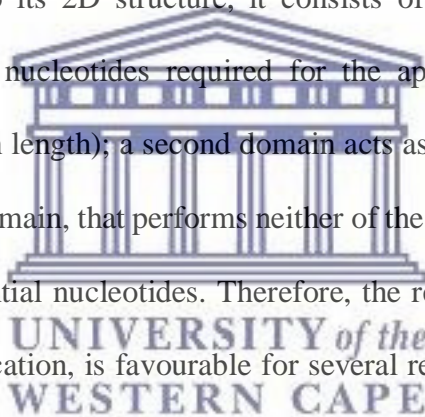
1.13.5.1. Identification of full-length aptamer sequence

Once the selection process has been completed, the next step would comprise sequencing the relevant DNA selection pools to identify the aptamer sequences. Conventionally, this is achieved by cloning, picking a few colonies and subjecting them to Sanger nucleotide sequencing. The most frequently occurring clones are identified, counted and aligned to group similar sequences. However, Sanger sequencing is limited to the number of sequences that can be produced per run and therefore, the complete sequence space (diversity) of the selection pool is not analysed. As a result, many sequences that may, potentially, have the highest affinity and selectively toward the target are not accounted for nor identified (Guo et al. 2014; Blind and Blank 2015). To overcome this limitation and in an effort to better understand the evolutionary nature of the selection process, Next Generation sequencing (NGS) has been employed. The NGS platforms have enabled the generation of high-throughput datasets that allows for a significantly more comprehensive *in silico* analysis of the data that could otherwise not be offered using an end-point Sanger sequencing approach (Blind and Blank 2015; Blank 2016). An example of this is the ability to now comparatively assess several rounds of SELEX, by describing the full sequence space, to monitor the enrichment or abundance of sequences, against the target, throughout the selection process (Ditzler et al. 2013). This translates into improved sequence coverage of the aptamer selection pool diversity (Beier et al. 2014). Although the data is analysed in a similar manner, that is, counting and subsequently sorting

the sequences by their frequency, a significantly more accurate and greater number of potential aptamer sequences may be identified (Blind and Blank 2015).

1.13.5.2. Determining the minimal binding sequence

After the first phase of screening, full-length aptamers (usually 40-90 NT in length) are identified, however, all the nucleotides in the full-length aptamer are generally not involved in the binding interaction between the aptamer and the target. In fact, in some cases, it was found that the redundant nucleotides may actually interfere with the aptamer-target interaction (Ku et al. 2015; Hasegawa et al. 2016). When considering this, it is important to note that when the full-length aptamer folds into its 2D structure, it consists of three domains. One domain constitutes the indispensable nucleotides required for the aptamer-target interaction (this domain is usually 10-15 NT in length); a second domain acts as a supportive structure for the binding activity; and a third domain, that performs neither of the two functions, makes up what is referred to as the non-essential nucleotides. Therefore, the removal of these non-essential nucleotides, by means of truncation, is favourable for several reasons (Nadal et al. 2013; Xie 2013).



Perhaps the first and most obvious is the cost implications when considering the scale up production of aptamers longer than 60nt. Secondly, as previously mentioned, the non-essential domain can often result in steric hindrance and subsequently affecting the binding activity of the domain directly involved in binding the target (Xie 2013). Another reason for using this approach, is that identifying potential binding subsequences of the selected aptamers may be accomplished by extracting conserved or common sequence and/or structural motifs from full-length aptamers, however, this is not always possible and in these instances, sequence truncation is employed as a better alternative (Le et al. 2014). One of the most commonly used

truncation methods and one of the more recently introduced alternatives to determine the MBS are described below.

1.13.5.2.1. Stepwise truncation

Since some of the nucleotides within the fixed priming region of the aptamer are sometimes involved in the aptamer-target interaction, it is not so apparent then to truncate the entire priming regions. Also, it has been shown that not all the nucleotides within the variable region are involved in the aptamer-target interaction. So this approach of stepwise fragmentation to produce a number of truncated variants has been adopted (Le et al. 2014). Although this approach requires the synthesis and screening of all the variants, it has, nonetheless, shown to produce high affinity MBS, sometimes to the extent of generating sequences with significantly higher affinities as compared to the full-length aptamer (Kaur and Yung 2012; Le et al. 2014; Ku et al. 2015; Jauset Rubio et al. 2016). There are also examples in which this approach results in the aptamers completely losing its binding activity, however, the rate of its successful application exceeds such cases (Le et al. 2014; Stoltenburg et al. 2015). This stepwise method is an *in silico* truncation of the candidate aptamers and can be employed in one of two ways, 1) by means of 'rational truncation', that is, generating 2D structure predictions of the full length aptamer, followed by a stepwise truncation, while maintaining an overall similar structure to that of the full-length aptamer as shown in **Figure 1.5** (Rockey et al. 2011).

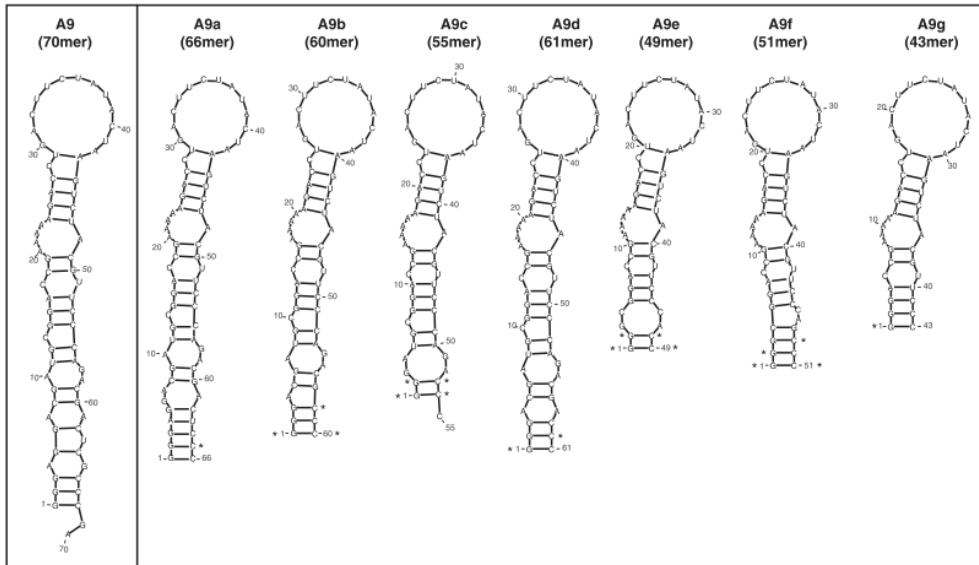


Figure 1.5. Stepwise aptamer truncation using a rational approach. Secondary structural predictions of truncated aptamers generated using RNAStructure 4.6 algorithm (Rockey et al. 2011).

Alternatively, the truncation can be performed using a ‘trial and error’ approach, that is, truncation of the 5’ or 3’ end, or truncation of both the 5’ and 3’ ends. Further trial and error truncation of one or both of the primer ends, together with a portion of the variable region may be attempted. This technique typically involves the removal of 3-5 bases at one time (Le et al. 2014; Stoltenburg et al. 2015).

1.13.5.2.2. Hybridisation inhibition

In this approach, the MBS is determined using a hybridization inhibition assay. Based on the results obtained, the respective interfering segments, identified within the full-length aptamer sequence, are eliminated to produce the truncated variants that are further characterized (Ku et al. 2015). The hybridization inhibition assay involves the synthesis of competitive complimentary probes against the full-length aptamer. These probes vary in length and are designed to hybridize to either the 5’ or 3’, or dual 5’ and 3’ ends of the full-length aptamer. The respective probes are incubated with the full-length aptamer and the K_D values are

determined using enzyme-linked immunosorbent oligonucleotide assay (ELONA), in the presence of the target. If no change occurs in K_D values of the hybridized aptamer after being exposed to the target, it suggests that those nucleotides involved in the hybridization are not essential for the aptamer-target interaction, since no inhibition of the aptamer binding activity to target occurred (Le et al. 2014).

Le and colleagues reported on the comparison of the truncation approach versus the hybridization inhibition method and concluded that both techniques yield biologically active truncated aptamers. The hybridization inhibition has an added advantage in that short probes are required for determining the MBS. But, truncated aptamers are advantageous because higher affinities were observed. This was attributed to the fact that in this approach, direct removal of the sequence results in a significantly low likelihood of steric hindrance. Furthermore, an improvement in the K_D values of these truncated aptamers was observed, mainly due to the elimination of unproductive intra-molecular interactions which, in the case of full-length aptamers, would have to be broken to promote productive intra-molecular interactions between the aptamer and the target. This means, in the case of full-length aptamers used in the hybridization inhibition assay, the double stranded nature of the hybrid results in more rigid 5' and 3' ends, subsequently introducing steric hindrances and ultimately interfering with the aptamer-target interaction (Le et al. 2014).

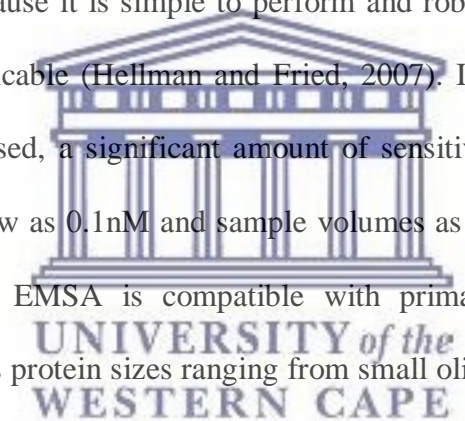
1.13.6. Characterisation of aptamer-target interaction

With the aim of applying high affinity and target-specific aptamers in a clinical setting, it is imperative to understand the interaction between the aptamer and its target (Jing and Bowser 2011). In doing so, information such as the affinity, binding kinetics upon interaction with the target, dissociation constants and identifying the binding region can be evaluated and used to

assess and/or improve the aptamer design (Jing and Bowser 2011). This can be accomplished by using characterisation techniques applicable to the aptamer/target interaction, of which the most commonly used methods are discussed below.

1.13.6.1. Electrophoretic mobility shift assay

The electrophoretic mobility shift assay (EMSA) is a semi-quantitative method that relies on a change in the electrophoretic mobility of the nucleic acids (Holden and Tacon 2010). The principle is based on the premise that a protein bound by nucleic acids has a reduced mobility, through the gel matrix, as compared to unbound nucleic acids (Hellman and Fried, 2007). EMSA is advantageous because it is simple to perform and robust because a wide range of binding conditions are applicable (Hellman and Fried, 2007). In cases where radioisotope-labelled nucleic acids are used, a significant amount of sensitivity is added; consequently, detection limits can be as low as 0.1nM and sample volumes as small as 20 μ l (Hellman and Fried, 2007). Furthermore, EMSA is compatible with primary, secondary and tertiary structures of DNA as well as protein sizes ranging from small oligo-peptides to complexes as large as $M_r \geq 10^6$ (Hellman and Fried, 2007). Purified or crude extracts of proteins may be used (Hellman and Fried, 2007). Some limitations do exist, these include the fact that the samples are not in chemical equilibrium, and thus, rapid dissociation during electrophoresis may prevent the detection of the complexes (Hellman and Fried, 2007). Additionally, EMSA does not provide any information regarding the binding kinetics. It is, nonetheless, useful as a complementary method (McKeague and De Rosa 2012; Ozer et al. 2014).



1.13.6.2. Surface plasmon resonance

Surface plasmon resonance (SPR) is a well-established and effective method that has been employed to characterize aptamer/target interactions for quite a number of years (Ruscito and DeRosa 2016). It is a label-free optical based detection method that measures changes in the refractive index, on the metal surface layer, in response to ligand/analyte interaction (Jerabek-willemsen et al. 2011; Nguyen et al. 2015). SPR is an invaluable technique used to provide information about binding kinetics and can also be used to determine K_D values in low mM to the nM range (Jerabek-willemsen et al. 2011; Jerabek-willemsen et al. 2014). One of the limitations of the method is that it labor intensive to set up a new assay and since it's a sensor surface-based method, it requires the immobilization of the ligand to the surface of the sensor chip (Jerabek-willemsen et al. 2011; Jerabek-willemsen et al. 2014). In spite of the limitations, it is still the most widely used method used to characterize aptamer/target interactions (Jerabek-willemsen et al. 2011; Jerabek-willemsen et al. 2014).



1.13.6.3. MicroScale thermophoresis

MicroScale thermophoresis (MST) is a highly sensitive quantitative method used to characterize, virtually, any kind of molecular interactions based on the principle of thermophoresis, under microscopic temperature gradients (Jerabek-willemsen et al. 2014). Upon a binding event, even the slightest change in conformation, charge or size of the molecule-solvent interface is directly translated into the thermophoretic effect detection for each interaction. Detection is monitored using either labeled or intrinsic fluorophores present in the sample. This method has a detection limit in the pM range, which is particularly applicable for aptamer/target interactions (Jerabek-willemsen et al. 2011; Ruscito and DeRosa 2016). Furthermore, it has an added advantage over SPR, in that it is a solution-based

technique, thus eliminating the need for surface immobilization. The interaction can be performed in almost any buffer, even serum and cell lysates; and as a result, mimicking an environment within a biological system (Wienken et al. 2010; Jerabek-willemsen et al. 2014; Stoltenburg et al. 2015). Due to the versatility and several advantages presented by MST, it has become more widely applied in aptamer/target characterisation and in so doing, present aptamers that have potential for different biological applications (Jerabek-willemsen et al. 2014).

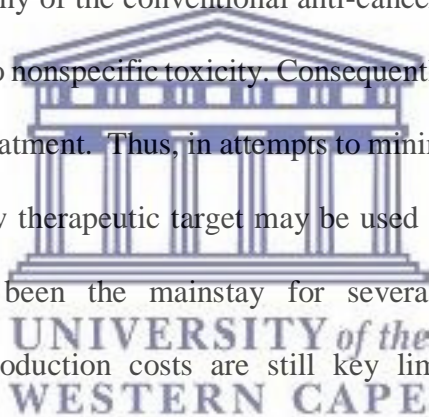
1.13.7. Aptamers in application

When taking into account the advantages demonstrated by aptamers – in contrast to antibodies – it would appear that several aptamers should have been commercialized already; however, this is not yet the case (Morita et al. 2018). Some of the challenges, to some extent, include aptamer stability, renal excretion, aptamer safety and to a larger extent, the reluctance of shifting the substantial financial investment from well-established antibody research, development and manufacturing process to a new technology, such as aptamers (Morita et al. 2018). In the past 28 years, since aptamers have emerged, to date only one aptamer, known as Macugen®, (further described in Section 1.13.7.2) has been approved by the Food and Drug Administration for therapeutic application (Morita et al. 2018). However, due to the desirable properties offered by aptamers as bio-recognition moieties, such as its high affinity and specificity towards a diverse array of targets, an increase in research for its application in the field of biotechnology has been observed (Wu et al. 2015). Considering the several advantages, especially in clinical applications, aptamers could offer, the prospect of commercialisation appears very promising (Sun and Zu 2015). Furthermore, since its inception, an ever increasing adoption of aptamers into several research diagnostic and therapeutic applications has occurred (Kong and Byun 2013; Sun and Zu 2015; Kaur et al. 2018). The aptamer industry is expected

to grow at such a rapid pace, with estimations of the industry reaching 244.93 million US dollars, by 2020 (Kaur et al. 2018). Currently, there are 40 aptamer-based companies actively engaged in the global commercialisation of aptamers in both therapeutic and diagnostic platforms (Kaur et al. 2018). Some of the aptamer-based products are commercialized, some have the potential to make immediate translation into the market and others may reach the market in the next 5-10 years (Kaur et al. 2018). A highlight of some of these advancements is discussed in this study.

1.13.7.1. Aptamers in therapeutics

In the area of therapeutics, many of the conventional anti-cancer drugs are very effective, but have adverse side effects due to nonspecific toxicity. Consequently, patients experience a lower quality of life at the time of treatment. Thus, in attempts to minimise the side effects, targeted therapy or alternatively, a new therapeutic target may be used (Sun et al. 2014). Antibody-based targeted drugs have been the mainstay for several years, however, potential immunogenicity and high production costs are still key limitations. And as previously discussed, aptamers are emerging as an alternative bio-recognition moiety to address these limitations (Santosh and Yadava 2014; Sun et al. 2014). Several studies and ongoing clinical trials have illustrated aptamers as therapeutic targets as well as a powerful instrument for targeted delivery of a diverse range of therapeutic agents for various diseases, including peptides, small molecules and anti-cancer drugs (Santosh and Yadava 2014). Examples of these aptamers are those listed in **Table 1.4**.



1.13.7.2. FDA approved and aptamers in clinical trials

The first FDA approved aptamer, Macugen, was launched in December 2004 (Ng et al. 2006; Sun et al. 2014). It is used for the treatment of wet age-related macular degeneration (AMD), a disease that results in the loss of vision (Ng et al. 2006). Macugen exerts its action as an antagonist against the vascular endothelial growth factor (VEGF), known to be a key regulator of aggressively abnormal growth and permeability of blood vessels in the eye (Ng et al. 2006). Other examples of aptamers in clinical trials, to date, are shown in **Table 1.4**.

Another example that has shown significant promise is Fovista, an aptamer synergistically used with anti-VEGF agents for the treatment of AMD (Maier and Levy 2016). Fovista targets the platelet-derived growth factor (PDGF), a vital cell survival growth factor required for stabilisation of VEGF induced blood vessel growth (Maier and Levy 2016). A phase 11b, randomized, clinical trial has just been completed, in which the efficacy and safety of Fovista, when administered with an anti-VEGF agent was evaluated in a cohort of 449 patients (Maier and Levy 2016). A favourable safety profile was achieved and has led to the initiation of phase III (Maier and Levy 2016).

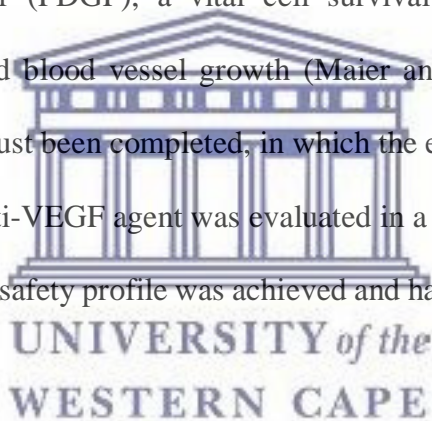


Table 1.4. FDA approved and aptamers in various stages of clinical trials.

Aptamer name	Target	Condition	Equilibrium dissociation constant (K_D)	Status	Reference
Macugen	Vascular endothelial growth factor	Age-related macular degeneration	200pM	Approved	Ruckman et al. 1998
Fovista	Platelet-derived growth factor	Age-related macular degeneration	0.1nM	Phase III	Floege et al. 1999
ARC1905	Human complement C5	Age-related macular degeneration	2-5nM	Phase III	Biesecker et al. 1999
AS1411	Nucleolin	Metastatic renal cell carcinoma	54.8nM ± 7.3nM	Terminated	Bates et al. 2009
NOX-A12	Stromal cell-derived factor-1	Multiple myeloma	200pM	Phase II	Roccaro et al. 2014
NOX-E36	Monocyte chemoattractant protein-1	Type II diabetes	1.32nM	Phase II completed	Oberthür et al. 2015
NOX-H94	Hepcidin	Anemia of chronic disease	0.65nM ± 0.06nM	Phase II completed	Schwoebel et al. 2013
ARC1779	A1 domain of von Willebrand Factor	von Willebrand disease	2nM	Phase II	Diener et al. 2009
REG1 system	Coagulation factor IXa	Acute coronary syndrome	2.83nM ± 0.4nM	Phase III	Rusconi et al. 2002
BX499	Tissue factor pathway inhibitor	Hemophilia	2.8nM ± 0.3nM	Terminated	Waters et al. 2011

1.13.7.3. Aptamers in diagnostics

The aptamers that have, this far, been investigated in clinical trials are primarily for therapeutic applications. However, several studies demonstrating the use of aptamers as a diagnostic tool are also reported in the literature (Maier and Levy 2016; Kaur et al. 2018). Over the past ten years, aptamers have made significant contributions on the diagnostic front, particularly in point-of-care (POC) platforms, such as aptamer-linked immobilized sorbent assay (ALISA), dot-blot, electro-chemiluminescence (ECL) assays, fluorescence based assays, nanoparticle-

based assays, lateral flow assays (LFA's), upconverting fluorophore-based assays, electrochemical sensors and glucometer-based assays to detect analytes (Kaur et al. 2018).

One such example was demonstrated by Raston and colleagues, who developed a pair of cognate aptamers against Vaspin, an insulin resistance biomarker that is involved in diabetes-related diseases (Raston and Gu, 2015; Raston et al. 2016). These aptamers were used to develop a sandwiched type LFA for the detection of Vaspin (Raston et al. 2016). The LFA consisted of 1) a biotinylated aptamer (capture aptamer) that was bound to streptavidin, which was in turn immobilized on the membrane – this is known as the test zone, 2) a AuNP-conjugated aptamer (signalling probe) and 3) a biotinylated DNA oligonucleotide that was complimentary to the AuNP-conjugated aptamer which was bound to streptavidin on the membrane – this is known as the control zone (**Figure 1.6A**) (Raston et al. 2016). The signaling probe was incubated with Vaspin and then added to the LFA. Upon doing this, the capture aptamer bound to the other site on Vaspin and the accumulation of AuNP-conjugated aptamers at the test zone caused the AuNPs to agglomerate, resulting in the observation of a red colour in the test zone (Raston et al. 2016). In the case of the control experiment, the complimentary aptamer captures the remaining AuNP-conjugated aptamers, by means of DNA hybridisation and which should always result in the observation of a red colour in the control zone (**Figure 1.6B**) (Raston, et al 2016). To demonstrate the specificity of the aptamers to Vaspin, counter target proteins, adinopectin and human serum albumin (HSA) were tested as controls (**Figure 1.6B**).

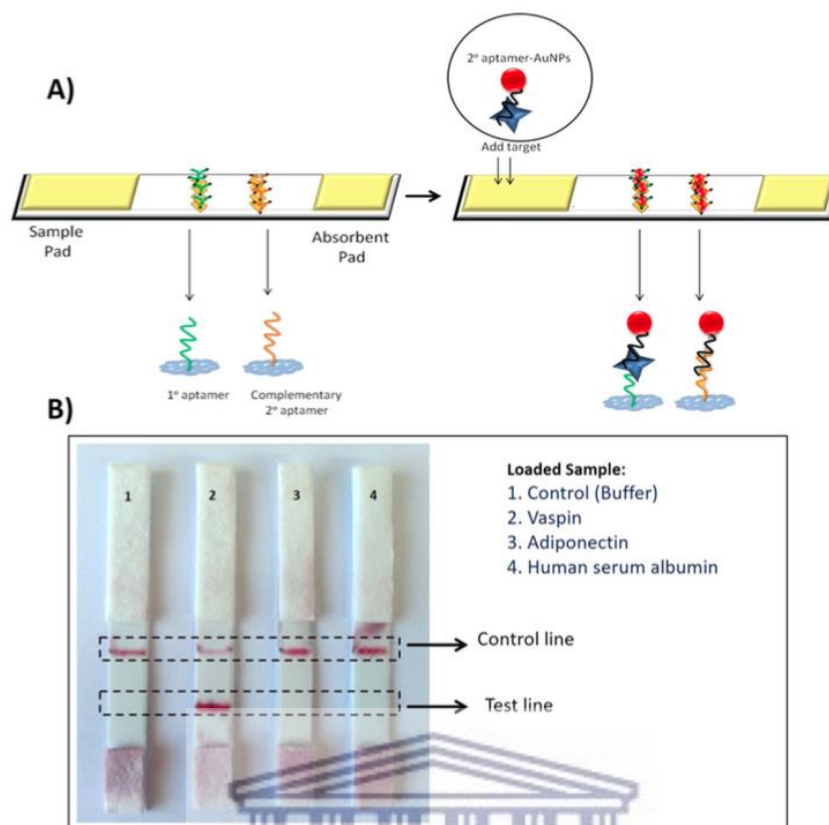


Figure 1.6. A lateral flow assay (LFA) developed for the detection of Vaspin. A) A schematic diagram of vaspin detection, using an aptamer-based LFA. B) Specific detection of Vaspin using an aptamer-based sandwiched type LFA.

UNIVERSITY of the
WESTERN CAPE

Of note, a vast number of the aptamer-based diagnostic assays have shown the potential to meet the ‘ASSURED’ criteria for POC platforms, which is affordable, sensitive, specific, user-friendly, robust and can always be performed by unskilled laborers outside the confines of a laboratory or hospital (Kaur et al. 2018). Although an extensive amount of research for diagnostic applications of aptamers exist, only a handful of companies have evolved these applications into successfully commercialized products (Kaur et al. 2018). Examples of these companies are Neo Ventures Biotechnology Inc., SomaLogic Inc., AptasSci Inc. and Base Pair (Kaur et al. 2018) and the commercialized products are enlisted in **Table 1.5**.

Table 1.5. Commercialized aptamer-based products.

Product	Company	Mode of detection	Application	Limitation
OTA-Sense and AflaSense	Neo Ventures Biotechnology Inc.	Fluorescence based assay	Detection of Mycotoxins (produced by <i>Aspergillus</i> and <i>Penicillium</i> species) in food	Require extraction of toxin from the sample
AptoCyto	AptaSci Inc.	Aptamer-based flow cytometry	Aptamer and magnetic bead-based isolation of biomarker positive cell	Dependence on expensive instrument (flow cytometer) to visualise the outcome
AptoPrep	AptaSci Inc.	Fluorescence-based assay and PAGE	Aptamer-magnetic-bead-based pulldown of biomarker positive cells and protein isolation	4°C storage is recommended for long-term storage
SOMAscan	SomaLogic	SOMAmer-based detection and quantification of biomarkers	A highly efficient platform with multiplex capability for biomarker discovery and diagnostics	Multistep process
CibusDx	CibusDx	Electrochemical sensing	Detection of food-borne pathogens	Instrument-based technology
OLIGOBIND	Sekisui diagnostics	Fluorogenic activity assay	Detection of active thrombin	Platelet contamination in plasma sample may interfere with assay
Hot Start Taq DNA Polymerase	NEB	Aptamer-based reversible inhibition of Taq polymerase	Hot start PCR	In comparison to normal polymerase this is relatively expensive

Kaur et al. 2018

From these examples and several others in the literature, the flexibility of aptamers in therapeutic and diagnostic systems demonstrate great potential for application in human diseases, such as cancer.

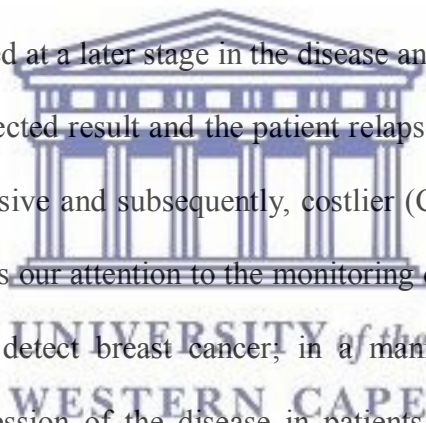
1.14. Problem Statement and Rationale

As cancer is one of the world's most unrelenting health burdens, the need to prevent, control and reduce cancer-related deaths remains a priority (World Health Organization, 2018). Despite current global efforts to combat the disease, exponential growth in mortality rates is projected for the coming years and thus, vigorous focus on developing new strategies and efficient resolutions must be actioned (Adeloye et al. 2018). At the 70th World Health Assembly, held in Geneva (May 2017), and also as reported by the Cancer Association of South Africa (CANSA), an estimate of the cost of cancer to the world economy is approximately 1.16 trillion US dollars, annually (World Health Organization, 2018). Sadly, the most significant financial and social impact of cancer is experienced in low- and middle-income countries, such as South Africa, where *only* a staggering 5% of the worlds' resources for cancer prevention and control are consumed (World Health Organization, 2018).

Among the top five cancers affecting women in South Africa, breast cancer is the most commonly diagnosed cancer in women and has, therefore, been identified, together with cervical cancer, as a national priority due to increasing incidence rates (CANSA, 2018). A report compiled by the Council for medical schemes (CMS) - which is an independent statutory body created by Parliament to regulate medical schemes in South Africa - showed that although a significant percentage of diagnosed breast cancer cases occur in developed countries as well, the survival rates differ greatly in various geographical regions; for example, North America has survival rates at ~80%, Sweden, Japan and some middle-income countries ~60% and down to as little as 40% in low-income countries (CMS, 2013). The CMS explains that the reason for

these skewed survival rates can be attributed to the lack of early detection, adequate diagnosis and treatment facilities in lower income countries; and subsequently yielding a high number of cases at late stage diagnosis (CMS, 2013).

Mammography is the recommended screening method for breast cancer, however, there is not yet compelling evidence that demonstrates a reduction in mortality as a result of using this technique (Bleyer et al. 2015). The aftermath of that prohibits mammography from being a prescribed minimum benefit level of care (CMS, 2013). Subsequently, the accessibility and cost of having mammograms as a routine screening method is not feasible in the context of South Africa (CMS, 2013). Therefore, while there is ongoing research in the area of early diagnosis to reduce the incidence of the disease, the reality for South Africa is that we have a high number of cases diagnosed at a later stage in the disease and in some cases, the treatment regime does not yield the expected result and the patient relapses – translating into repetitive biopsies, which is highly invasive and subsequently, costlier (Carlos et al. 2017). Therefore, the purpose of this study draws our attention to the monitoring of breast cancer, by proposing a cost-effective approach to detect breast cancer, in a manner that can inform medical practitioners about the progression of the disease in patients during and after therapeutic intervention.



1.15. The aim of this study

The aim of this study was to develop DNA aptamers against putative breast cancer biomarkers for its application in breast cancer prognosis.

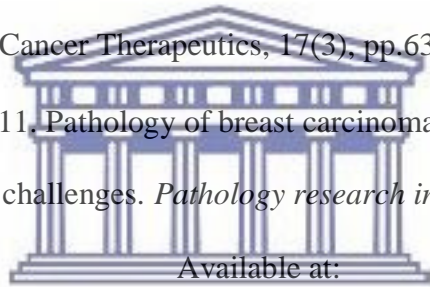
1.16. The objectives of this study

- To express and purify recombinant human ULBP2 protein.
- To develop DNA aptamers against ULBP2 and GFR α 1, respectively, using *in vitro* systematic evolution of ligands by exponential enrichment.
- To identify candidate aptamers for both ULBP2 and GFR α 1, using Next Generation Sequencing.
- To characterise the candidate aptamers using the EMSA, SPR and MST.
- To evaluate the application of the selected candidate aptamers using flow cytometry.



1.17. References

- Abratt, R.P. and Vorobiof, D.A., 2003. Cancer in Africa. *Lancet Oncology*, 4(7), pp.394–396.
- Adeloje, et al. 2018. Estimating the incidence of breast cancer in Africa: a systematic review and meta-analysis. *Journal of Global Health*, 8(1), pp.1-21.
- Andreotti, P.E. et al. 2003. Immunoassay of infectious agents. *BioTechniques Euro Edition*, 859, pp.850–859.
- Angrist, M. et al. 1998. Human GFRA1: cloning, mapping, genomic structure, and evaluation as a candidate gene for Hirschsprung disease susceptibility. *Genomics*, 48(3), pp.354–362.
- Aquino-Jarquín, G. and Toscano-Garibay, J.D., 2011. RNA aptamer evolution: Two decades of selection. *International Journal of Molecular Sciences*, 12(12), pp.9155–9171.
- Arango, B.A., Rivera, C.L. and Gl??ek, S., 2013. Gene expression profiling in breast cancer. *American Journal of Translational Research*, 5(2), pp.132–138.
- Assadollahi, S. et al. 2009. From lateral flow devices to a novel nano-color microfluidic assay. *Sensors*, 9(8), pp.6084–6100.
- Bantscheff, M. et al. 2007. Quantitative mass spectrometry in proteomics : a critical review. *Analytical and Bioanalytical Chemistry*, 404(4), pp.939–965.
- Bates, P.J. Laber, D.A. Miller, D.M. Thomas, S.D. and Trent, J.O. 2009. Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer. *Exp Mol Pathol*, 86, pp.151-64.
- Bauer, M., MacDonald, J., Henri, J., Duan, W. and Shigdar, S. 2016. The application of aptamers for immunohistochemistry. *Nucleic Acid Therapeutics*, 26(3), pp.120-126.

- Beier, R., Boschke, E. and Labudde, D., 2014. SELEX Experiments. *BioMed Research International*, 2014.
- Berezovski, M. Musheev, M. Drabovich, A. and Krylov, S.N. 2006. Non-SELEX selection of aptamers. *J. Am. Chem. Soc.* 128, pp.1410–1411.
- Bertucci, F. and Birnbaum, D., 2008. Reasons for breast cancer heterogeneity. *Journal of biology*, 7(2), p.6.
- Bertucci, F., Finetti, P. and Birnbaum, D., 2008. Basal Breast Cancer : A Complex and Deadly Molecular Subtype. *Journal of biology*, 7(6).
- Bhakta et al. 2018. An anti-GDNF family receptor alpha 1 (GFRA1) antibody-drug conjugate for the treatment of hormone receptor-positive breast cancer. *Molecular Cancer Therapeutics*, 17(3), pp.638-649.
- Bhargava, R. et al. 2011. Pathology of breast carcinoma: diagnostic, prognostic, and therapeutic issues and challenges. *Pathology research international*, 2011, p.731470.


Available at:
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3180077&tool=pmcentrez&rendertype=abstract>
- Bhatt, A.N. et al. 2010. Cancer biomarkers - current perspectives. *The Indian journal of medical research*, 132, pp.129–149.
- Bhattacharyya, D., Singh, S. and Satnalika, N. 2009. Nanotechnology , Big things from a Tiny World : a Review. , 2(3), pp.29–38.
- Blank, M. Weinschenk, T. Priemer, M. and Schluesener, H. 2001. Systematic evolution of a DNA aptamer binding to rat brain tumor microvessels— selective targeting of endothelial regulatory protein pigpen. *J. Biol. Chem.* 276, pp16464–16468.
- Blind, M. and Blank, M., 2015. Aptamer Selection Technology and Recent Advances.

Molecular Therapy—Nucleic Acids, 4(1), p.e223.

- Biesecker, G. Dihel, L. Enney, K. and Bendele, R.A. 1999. Derivation of RNA aptamer inhibitors of human complement C5. *Immunopharmacology*, 1999, 42, pp.219-30.
- Bohunicky, B. and Mousa, S.A., 2011. Biosensors: The new wave in cancer diagnosis. *Nanotechnology, Science and Applications*, 4(1), pp.1–10.
- Bordeaux, M.C. et al. 2000. The RET proto-oncogene induces apoptosis: a novel mechanism for Hirschsprung disease. *Embo J*, 19(15), pp.4056–4063. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10921886> \ <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC306592/pdf/cdd384.pdf>.
- Boulay, A. et al. 2008. The ret receptor tyrosine kinase pathway functionally interacts with the ER?? pathway in breast cancer. *Cancer Research*, 68(10), pp.3743–3751.
- Brannon-peppas, L. and Blanchette, J.O., 2004. Nanoparticle and targeted systems for cancer therapy. *Advanced Drug Delivery Reviews*, 56, pp.1649–1659.
- Bruno, J.G., 2015. Predicting the uncertain future of aptamer-based diagnostics and therapeutics. *Molecules*, 20(4), pp.6866–6887.
- Bryceson, Y.T., Ljunggren, H.G. and Long, E.O., 2009. Minimal requirement for induction of natural cytotoxicity and intersection of activation signals by inhibitory receptors. *Blood*, 114(13), pp.2657–2666.
- Byrne, J.D., Betancourt, T. and Brannon-peppas, L., 2008. Active targeting schemes for nanoparticle systems in cancer therapeutics. *Advanced Drug Delivery Reviews*, 60(15), pp.1615–1626. Available at: <http://dx.doi.org/10.1016/j.addr.2008.08.005>.
- Champsaur, M. and Lanier, L.L., 2011. Effects of NKG2D ligand expression on host immune responses. *Immunological reviews*, 235(1), pp.267–285.
- Chang, Y.T. et al. 2011. Secretome-based identification of ULBP2 as a novel serum

marker for pancreatic cancer detection. *PLoS ONE*, 6(5).

- Chen, T. et al. 2011. Aptamer-conjugated nanomaterials for bioanalysis and biotechnology applications. *Nanoscale*, 3(2), pp.546–556. Available at: <Go to ISI>://000287363500017\http://pubs.rsc.org/en/content/articlepdf/2011/nr/c0nr00646g.
- Chitadze, G. et al. 2013. Generation of Soluble NKG2D Ligands: Proteolytic Cleavage, Exosome Secretion and Functional Implications. *Scandinavian Journal of Immunology*, 78(2), pp.120–129.
- Cho, W.C.S., 2007. Proteomics Technologies and Challenges From Genomics to Proteomics. *Genomics, Proteomics and Bioinformatics*, 5(2), pp.77–85. Available at: [http://dx.doi.org/10.1016/S1672-0229\(07\)60018-7](http://dx.doi.org/10.1016/S1672-0229(07)60018-7).
- Ciccimaro, E. et al. 2010. Stable-isotope dilution LC – MS for quantitative biomarker analysis. *Bioanalysis*, 2(2), pp.311–341.
- Citartan, M. et al. 2012. Asymmetric PCR for good quality ssDNA generation towards DNA aptamer production. *Songklanakarit Journal of Science and Technology*, 34(2), pp.125–131.
- Collins, R.W.M., 2004. Human MHC class I chain related (MIC) genes: Their biological function and relevance to disease and transplantation. *European Journal of Immunogenetics*, 31(3), pp.105–114.
- Cosman, D. et al. 2001. ULBPs, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor. *Immunity*, 14(2), pp.123–133.
- Creighton, C.J., 2012. The molecular profile of luminal B breast cancer. *Biologics: Targets and Therapy*, 6, pp.289–297.
- Cumber, S.N., Nchanji, K.N and Tsoka-Gwegweni, J.M. 2017. Breast cancer among

women in sub-Saharan Africa: prevalence and a situational analysis. *South African Journal of Gynaecological Oncology*, 9(2), pp.35-37.

- Danhier, F., Feron, O. and Pr at, V., 2010. To exploit the tumour microenvironment : Passive and active tumour targeting of nanocarriers for anti-cancer drug delivery. *Journal of Controlled Release*, 148(2), pp.135–146. Available at: <http://dx.doi.org/10.1016/j.jconrel.2010.08.027>.
- de Matos, L.L., Truffelli, D.C., de Matos, M.G.L. and Pinhal, M.A. 2010. Immunohistochemistry as an important tool in biomarkers detection and clinical practice. *Biomarker Insights*, 5(2010), pp.9-20.
- Diener, J.L. et al. 2009. Inhibition of von Willebrand factor-mediated platelet activation and thrombosis by the anti-von Willebrand factor A1-domain aptamer ARC1779. *J Thromb Haemost*, 7, pp.1155-1162.
- Ditzler, M.A. et al. 2013. High-throughput sequence analysis reveals structural diversity and improved potency among RNA inhibitors of HIV reverse transcriptase. *Nucleic Acids Research*, 41(3), pp.1873–1884.
- Drah, M., 2015. The Development of Nanotechnology- based Detection Systems for the Diagnosis of Breast Cancer. *Ph.D. Thesis*. University of the Western Cape.
- Duffy, M.J., Evoy, D. and McDermott, E.W., 2010. CA 15-3: Uses and limitation as a biomarker for breast cancer. *Clinica Chimica Acta*, 411(23–24), pp.1869–1874. Available at: <http://dx.doi.org/10.1016/j.cca.2010.08.039>.
- Duffy et al. 2017. Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumour Markers (EGTM). *European Journal of Cancer*, 75(2017), pp.284-298.
- Duffy, M.J., McDermott, E.W. and Crown, J. 2018. Blood-based biomarkers in breast

cancer: From proteins to circulating tumour DNA. *Tumour Biology*, (2018), pp.1-11.

- Elmore, J.G. et al. 2011. Screening for Breast Cancer. *Health Care*, 293(10), pp.1245–1256. Available at: <http://dx.doi.org/10.1016/B978-1-4557-4007-9.00116-3>.
- Essegir, S. et al. 2007. A role for glial cell derived neurotrophic factor induced expression by inflammatory cytokines and RET/GFR alpha 1 receptor up-regulation in breast cancer. *Cancer Research*, 67(24), pp.11732–11741.
- Fleming, M.S. et al. 2015. Cis and trans RET signaling control the survival and central projection growth of rapidly adapting mechanoreceptors. *eLife*, 2015(4), pp.1–26.
- Floege, J. et al. 1999. Novel approach to specific growth factor inhibition in vivo: Antagonism of platelet-derived growth factor in glomerulonephritis by aptamers. *Am J Pathol*, 154, pp.169-79.
- Gao, X. et al. 2014. Recent Advances in Nanoparticles-based Lateral Flow Biosensors. *American Journal of Biomedical Sciences*, 6(1), pp.41–57.
- Gattelli, A. et al. 2013. Ret inhibition decreases growth and metastatic potential of estrogen receptor positive breast cancer cells. *EMBO Molecular Medicine*, 5(9), pp.1335–1350.
- Golden, M.C. Collins, B.D. Willis, M.C. and Koch, T.H. 2000. Diagnostic potential of PhotoSELEX-evolved ssDNA aptamers. *J. Biotechnol.* 81, pp.167–178.
- Goossens, N., Nakagawa, S., Sun, X. and Hoshida, Y. 2015. Cancer biomarker discovery and validation. *Translational Cancer Research*, 4(3), pp. 256-269.
- Gopinath, S.C.B., 2007. Methods developed for SELEX. *Analytical and Bioanalytical Chemistry*, 387(1), pp.171–182.
- Grebe, S.K.G. and Singh, R.J., 2011. LC-MS / MS in the Clinical Laboratory – Where

to From Here ? *Clinical Biochemistry*, 32, pp.5–31.

- Groh, V. et al. 1999. Broad tumour-associated expression and recognition by tumour-derived gamma delta T cells of MICA and MICB. *Proceedings of the National Academy of Sciences of the United States of America*, 96(12), pp.6879–6884.
- Groh, V. et al. 2002. Tumour-derived soluble MIC ligands impair expression of NKG2D and T-cell activation. *Nature*, 419(6908), pp.734–738.
- Groot, J.W.B. De et al. 2016. RET as a Diagnostic and Therapeutic Target in Sporadic and Hereditary Endocrine Tumours. , 27(June), pp.535–560.
- Guo, W.M. et al. 2014. Identification and Characterization of an eIF4e DNA Aptamer That Inhibits Proliferation With High Throughput Sequencing. *Molecular therapy. Nucleic acids*, 3, pp.2162–2531. Available at: <http://dx.doi.org/10.1038/mtna.2014.70>.
- Hamula, C.L.A. et al. 2006. Selection and analytical applications of aptamers. *TrAC - Trends in Analytical Chemistry*, 25(7), pp.681–691.
- Hanahan, D. and Weinberg, R.A., 2011. Hallmarks of Cancer : The Next Generation. *Cell*, 144(5), pp.646–674. Available at: <http://dx.doi.org/10.1016/j.cell.2011.02.013>.
- Hanahan, D., Weinberg, R.A. and Francisco, S., 2000. The Hallmarks of Cancer. *Cell Press*, 100, pp.57–70.
 - Hanash, S., 2003. Disease proteomics. *Nature*, 422, pp.226–232.
 - Hasegawa, H. et al. 2016. Methods for Improving Aptamer Binding Affinity. *Molecules*, 21(421), pp. 1-15.
 - He, A. et al. 2013. A novel immunoassay for the quantization of CYFRA 21-1 in human serum. *Journal of Clinical Laboratory Analysis*, 27(4), pp.277–283.
 - Henry, N.L. and Hayes, D.F., 2012. Cancer biomarkers 5. *Molecular Oncology*, 6(2), pp.140–146. Available at: <http://dx.doi.org/10.1016/j.molonc.2012.01.010>.
 - Higgins, M.J. and Baselga, J., 2011. Review series Targeted therapies for breast

cancer. *Journal of clinical investigation*, 121(10), pp.3797–3803.

- Honda et al. 2013. Proteomics approaches to discovery of cancer biomarkers for early detection and personalised medicine. *Japanese Journal of Clinical Oncology*, 43(2), pp.103-109.
- Hong, P. et al. 2012. Applications of Aptasensors in Clinical Diagnostics. , pp.1181–1193.
- Hood, L. and Rowen, L., 2013. The Human Genome Project : big science transforms biology and medicine. *Genome Medicine*, 5(79), pp.1–8.
- Hu, X. et al. 2009. Genetic alterations and oncogenic pathways associated with breast cancer subtypes. *Mol Cancer Res*, 7(4), pp.511–522. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/19372580>.
- Hudler, P., Kocevar, N. and Komel, R. 2014. Proteomic approaches in biomarker discovery: New Perspectives in cancer diagnostics. *The Scientific World Journal*, 2014, pp.1-18.
- Huergo-Zapico, L. et al. 2014. Molecular bases for the regulation of NKG2D ligands in cancer. *Frontiers in Immunology*, 5(106), pp.1–7.
- Jakesz, R., 2008. Breast Care Breast Cancer in Developing Countries : Challenges for Multidisciplinary Care. *Breast Care*, pp.4–5.
- Jauset Rubio, M. et al. 2016. β -Conglutin dual aptamers binding distinct aptatopes. *Analytical and Bioanalytical Chemistry*, 408(3), pp.875–884.
- Jayasena, S.D., 1999. Aptamers: An emerging class of molecules that rival antibodies in diagnostics. *Clinical Chemistry*, 45(9), pp.1628–1650.
- Jerabek-willemsen, M. et al. 2014. MicroScale Thermophoresis : Interaction analysis and beyond q. *Journal of Molecular Structure*, 1077, pp.101–113. Available at:
<http://dx.doi.org/10.1016/j.molstruc.2014.03.009>.

- Jerabek-willemsen, M. et al. 2011. Molecular Interaction Studies Using Microscale Thermophoresis. *Assay and Drug Development Technologies*, pp.342–353.
- Jing, M. and Bowser, M., 2011. A Review of Methods for Measuring Aptamer-Protein Equilibria. *Analytica Chimica Acta*, 686(1–2), pp.9–18.
- Kabel, A.M. 2017 Tumour biomarkers of breast cancer: New Perspectives. *Journal of Oncological Sciences*, 3(2017), pp.5-11.
- Kalager, M. et al. 2009. Improved breast cancer survival following introduction of an organized mammography screening program among both screened and unscreened women: a population-based cohort study. *Breast cancer research : BCR*, 11(4), p.R44. Available at:
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2750103&tool=pmcentrez&rendertype=abstract>.
- Kalogerakos, K., Sofoudis, C. and Baltayiannis, N., 2008. Early breast cancer : A review Review Article. *Therapy*, 6, pp.463–476.
- Karley, D., Gupta, D. and Tiwari, A. 2011. Biomarker for cancer: A great promise for future. *World Journal of Oncology*, 2(4), pp.151-157.
- Kaur, H. and Yung, L.Y.L., 2012. Probing high affinity sequences of DNA aptamer against VEGF 165. *PLoS ONE*, 7(2), pp.19–26.
- Kaur, H., Bruno, J.G., Kumar, A. and Sharma, T.K. 2018. Aptamers in therapeutics and diagnostics pipelines. *Theranostics*, 8(15), pp. 4016-4032.
- Khati, M., 2010. The Future of Aptamers in Medicine. *Journal of clinical pathology*, 63(6), pp.480–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20360137>.
- Kittaneh, M. and Montero, A.J., 2013. Biomarkers in Cancer Molecular Profiling for Breast Cancer : A Comprehensive Review. *Biomarkers in Cancer*, 5, pp.61–70.
- Kong, H.Y. and Byun, J., 2013. Nucleic acid aptamers: New methods for selection,

stabilization, and application in biomedical science. *Biomolecules and Therapeutics*, 21(6), pp.423–434.

- Krombein, I. and De Velliers, P., 2006. Breast cancer – early detection and screening in South African women from the Bonteheuwel township in the Western Cape : Knowledge , attitudes and practices. *South African Family Practice*, 48(5), p.14a–14f.
- de Kruijf, E.M. et al. 2012. NKG2D ligand tumour expression and association with clinical outcome in early breast cancer patients: an observational study. *BMC cancer*, 12(1), p.24. Available at: <http://bmccancer.biomedcentral.com/articles/10.1186/1471-2407-12-24>.
- Ku, T.H. et al. 2015. Nucleic acid aptamers: An emerging tool for biotechnology and biomedical sensing. *Sensors*, 15(7), pp.16281–16313.
- Kulasingam, V. and Diamandis, E.P., 2008. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. *Nature Clinical Practice*, 5(10), pp.588–599.
- Kulbachinskiy, a V, 2007. Methods for selection of aptamers to protein targets. *Biochemistry. Biokhimiia*, 72(13), pp.1505–1518.
- Labelle, M. and Hynes, R.O., 2012. The initial hours of metastasis: the importance of cooperative host-tumour cell interactions during hematogenous dissemination. *Cancer Discovery*, 2(12), pp.1091–1099.
 - Lakhin, A. V, Tarantul, V.Z. and Gening, L. V, 2013. Aptamers : Problems , Solutions and Prospects. *Acta Naturae* , 5(19), pp.34–43.
- Laronga, C., 2016. Patient education : Breast cancer guide to diagnosis and treatment (Beyond the Basics). , pp.1–15.
- Le, T.T., Chumphukam, O. and Cass, A.E.G., 2014. Determination of minimal sequence for binding of an aptamer. A comparison of truncation and hybridization

inhibition methods. *RSC Adv.*, 4(88), pp.47227–47233. Available at:

<http://dx.doi.org/10.1039/C4RA08243E>.

- Ledda, F. et al. 2007. GDNF and GFR α 1 promote formation of neuronal synapses by ligand-induced cell adhesion. *Nature neuroscience*, 10(3), pp.293–300.
- Lee, Y.J. and Lee, S.W. 2006. In vitro selection of cancer-specific RNA aptamers. *J. Microbiol. Biotechnol.* 16, pp.1149–1153.
- Leppa, V. et al. 2004. The structure of GFR α 1 domain 3 reveals new insights into GDNF binding and RET activation. *European Molecular Biology Organisation*, 23(7), pp.1452–1462.
- Levy, E.M., Roberti, M.P. and Mordoh, J., 2011. Natural killer cells in human cancer: from biological functions to clinical applications. *Journal of biomedicine and biotechnology*, 2011, p.676198. Available at:
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3085499&tool=pmcentrez&rendertype=abstract>.
- Li, J. et al. 2002. Proteomics and Bioinformatics Approaches for Identification of Serum Biomarkers to Detect Breast Cancer. *Clinical chemistry*, 48(8), pp.1296–1304.
- Li et al. 2008. Clinical significance of the NKG2D ligands, MICA/B and ULBP2 in ovarian cancer: high expression of ULBP2 is an indicator of poor prognosis. *Cancer Immunology Immunotherapy*, 58(2009), pp.641-652.
- Liu, J., Mazumdar, D. and Lu, Y., 2006. A simple and sensitive “dipstick” test in serum based on lateral flow separation of aptamer-linked nanostructures. *Angewandte Chemie - International Edition*, 45(47), pp.7955–7959.
- Lou, X. et al. 2015. Interaction between circulating cancer cells and platelets : clinical implication. *chinese journal of cancer research*, 27(5), pp.450–460.
- Luscombe, N.M., Greenbaum, D. and Gerstein, M. 2001. What is Bioinformatics? A

proposed definition and overview of the field. *Methods of Information in Medicine*, 40(2001), pp. 346-358.

- Maheu, M. et al. 2015. MicroRNA regulation of central glial cell line-derived neurotrophic factor (GDNF) signalling in depression. *Translational psychiatry*, 5(2), p.e511. Available at: <http://dx.doi.org/10.1038/tp.2015.11>.
- Maier, K.E. and Levy, M., 2016. From selection hits to clinical leads : progress in aptamer discovery. *MolecularTherapy - Methods and Clinical Development*, 5, pp.1–10.
- Mairal, T. et al. 2008. Aptamers: Molecular tools for analytical applications. *Analytical and Bioanalytical Chemistry*, 390(4), pp.989–1007.
- Malhotra, G.K. et al. 2010. Histological , molecular and functional subtypes of breast cancers. *Cancer biology and therapy*, 10(10), pp.955–960.
- Mathelin, C. et al. 2006. Serum biomarkers for detection of breast cancers: A prospective study. *Breast Cancer Research and Treatment*, 96(1), pp.83–90.
- Mayer, G., 2009. The chemical biology of aptamers. *Angewandte Chemie - International Edition*, 48(15), pp.2672–2689.
- McNerney, R. and Daley, P., 2011. Towards a point-of-care test for active tuberculosis: obstacles and opportunities. *Nature reviews. Microbiology*, 9(3), pp.204–213. Available at: <http://dx.doi.org/10.1038/nrmicro2521>.
- Melorose, J., Perroy, R. and Careas, S., 2015. *The Aptamer Handbook*,
- Mesri, M., 2014. Advances in Proteomic Technologies and Its Contribution to the Field of Cancer. *Advances in Medicine*, pp.1–25.
- Morandi, A. and Isacke, C.M., 2014. Targeting RET – interleukin-6 crosstalk to impair metastatic dissemination in breast cancer. *Brea*, 16(301), pp.2–4.
- Morandi, A., Plaza-Menacho, I. and Isacke, C.M., 2011. RET in breast cancer:

Functional and therapeutic implications. *Trends in Molecular Medicine*, 17(3), pp.149–157.

- Moreno-Aspitia, A. et al. 2013. Soluble HER2 (sHER2) levels in HER2-positive breast cancer patients receiving chemotherapy ± trastuzumab : Results from the NCCTG adjuvant trial N9831. *Cancer*, 119(15), pp.2675–2682.
- Ng, E.W.M. et al. 2006. Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nature Reviews. Drug Discovery*, 5, pp.123–132.
- Ngcoza, N. 2015. Identification of biomarkers in breast cancer as potential diagnostic and therapeutic agents using a combined in silico and molecular approach. *M.Sc. thesis*. University of the Western Cape.
- Nguyen, H.H. et al. 2015. Surface plasmon resonance: A versatile technique for biosensor applications. *Sensors (Switzerland)*, 15(5), pp.10481–10510.
- Ni, X. et al. 2011. Nucleic acid aptamers : clinical applications and promising new horizons. *Curr Med Chem*, 18(27), pp.4206–4214.
- Oberthur, D. et al. 2015. Crystal structure of a mirror-image L-RNA aptamer (Spiegelmer) in complex with the natural L-protein target CCL2. *Nat Commun*, 6, pp.6923.
- Ogunbode, A.M. et al. 2013. Breast examination as a cost-effective screening tool in a clinical practice setting in Ibadan, Nigeria. *African Journal of Primary Health Care and Family Medicine*, 5(1), p.7 pages.
- Ong, S. and Mann, M., 2006. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nature protocols*, 1(6), pp.2650–2660.
- Orava, E.W., Cicmil, N. and Gariépy, J., 2010. Biochimica et Biophysica Acta Delivering cargoes into cancer cells using DNA aptamers targeting internalized surface portals. *BBA - Biomembranes*, 1798(12), pp.2190–2200.

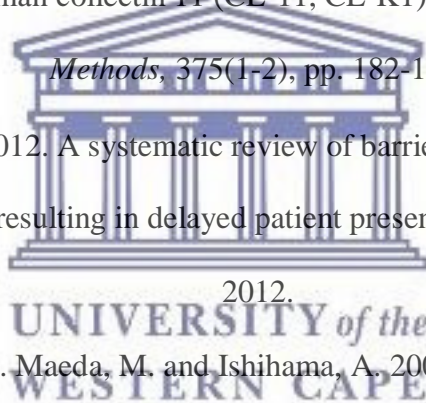
- Osborne, C.K., Wakeling, A. and Nicholson, R.I., 2004. Fulvestrant : an oestrogen receptor antagonist with a novel mechanism of action. *British journal of cancer*, 90, pp.S2–S6.
- Ozer, A., Pagano, J.M. and Lis, J.T., 2014. New Technologies Provide Quantum Changes in the Scale, Speed, and Success of SELEX Methods and Aptamer Characterization. *Molecular therapy. Nucleic acids*, 3, pp.1–18.
- Parkash, V. et al. 2008. The structure of the glial cell line-derived neurotrophic factor-coreceptor complex: Insights into RET signaling and heparin binding. *Journal of Biological Chemistry*, 283(50), pp.35164–35172.
- Pei, X., Zhang, J.U.N. and Liu, J.I.E., 2014. Clinical applications of nucleic acid aptamers in cancer (Review). *Molecular and Clinical Oncology*, 2, pp.341–348.
- Perkins, G.L. et al. 2003. Serum tumour markers. *American Family Physician*, 68(6), pp.1075–1082.
- Place, A.E., Jin Huh, S. and Polyak, K., 2011. The microenvironment in breast cancer progression: biology and implications for treatment. *Breast cancer research : BCR*, 13(6), p.227.
- Plaza-Menacho, I. et al. 2010. Targeting the receptor tyrosine kinase RET sensitizes breast cancer cells to tamoxifen treatment and reveals a role for RET in endocrine resistance. *Oncogene*, 29(33), pp.4648–4657.
- Polyak, K., 2011. Review series introduction Heterogeneity in breast cancer. *J.Clin.Invest.*, 121(10), pp.2011–2013.
- Popiel, M. et al. 2012. Mammary carcinoma – current diagnostic methods and symptomatology in imaging studies. *Polish Journal of Radiology*, 77(4), pp.35–44.
- Purschke, W.G. et al. 2003. A DNA Spiegelmer to staphylococcal enterotoxin B. *Nucleic acids research*, 31(12), pp.8–13.

- Radrizzani, M. et al. 1999. Oligobodies: bench made synthetic antibodies. *Med.-Buenos Aires* 59, pp.753–758.
- Rambau, P.F. et al. 2011. Pathological features of Breast Cancer seen in Northwestern Tanzania: a nine years retrospective study. *BMC research notes*, 4(1), p.214.
- Raullet, D.H. et al. 2013. *Regulation of Ligands for the NKG2D Activating Receptor*, Available at: <http://www.annualreviews.org/doi/abs/10.1146/annurev-immunol-032712-095951>.
- Roccaro, A.M. et al. 2014. SDF-1 inhibition targets the bone marrow niche for cancer therapy. *Cell Rep*, 9, pp.118-28.
- Rockey, W.M. et al. 2011. Rational truncation of an RNA aptamer to prostate-specific membrane antigen using computational structural modeling. *Nucleic acid therapeutics*, 21(5), pp.299–314.
- Rotherham, L.S. et al. 2012. Selection and Application of ssDNA Aptamers to Detect Active TB from Sputum Samples. *PLoS ONE*, 7(10), pp.1–11.
- Ruckman, J. et al. 1998. 2'-Fluoropyrimidine RNA-based aptamers to the 165-amino acid form of vascular endothelial growth factor (VEGF165). Inhibition of receptor binding and VEGF-induced vascular permeability through interactions requiring the exon 7-encoded domain. *J Biol Chem*, 273, pp.20556-20567.
- Ruscito, A. and DeRosa, M.C., 2016. Small-Molecule Binding Aptamers: Selection Strategies, Characterization, and Applications. *Frontiers in Chemistry*, 4(May), pp.1–14.
- Rusconi, C.P. et al. 2002. RNA aptamers as reversible antagonists of coagulation factor IXa. *Nature*, 419, pp.90-94.
- Sajid, M. and Daud, M., 2014. Designs , formats and applications of lateral flow

assay : A literature review. *Journal of Saudi Chemical Society*, 19(6), pp.689–705.

Available at: <http://dx.doi.org/10.1016/j.jscs.2014.09.001>.

- Sampson, T., 2003. Aptamers and SELEX: The technology. *World Patent Information*, 25(2), pp.123–129.
- Santosh, B. and Yadava, P.K., 2014. Nucleic Acid Aptamers : Research Tools in Disease Diagnostics and Therapeutics. , 2014.
- Schwoebel, F. et al. 2013. The effects of the anti-hepcidin Spiegelmer NOX-H94 on inflammation-induced anemia in cynomolgus monkeys. *Blood*, 121, pp.2311-2315.
- Selman et al. 2012. An enzyme-linked immunosorbent assay (ELISA) for quantification of human collectin 11 (CL-11, CL-K1), *Journal of Immunological Methods*, 375(1-2), pp. 182-188.
- Sharma, K. et al. 2012. A systematic review of barriers to breast cancer care in developing countries resulting in delayed patient presentation. *Journal of Oncology*, 2012.
- Shimada, T. Fujita, N. Maeda, M. and Ishihama, A. 2005. Systematic search for the Cra-binding promoters using genomic SELEX system. *Genes Cells* 10, pp.907–918.
- Shulman, L.N. et al. 2010. Breast cancer in developing countries: Opportunities for improved survival. *Journal of Oncology*, 2010.
- Song, H. et al. 2006. Soluble ULBP suppresses natural killer cell activity via down-regulating NKG2D expression. *Cellular Immunology*, 239(1), pp.22–30.
- Song, K.M., Lee, S. and Ban, C., 2012. Aptamers and their biological applications. *Sensors*, 12(1), pp.612–631.
- Srinivas, P.R., Verma, M., Zhao, Y. and Srivastava, S. 2002. Proteomics for cancer biomarker discovery. *Clinical Chemistry*, 48(8), pp. 1160-1169.



- Srisawat, C. and Engelke, D.R., 2001. Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. *RNA (New York, N.Y.)*, 7(4), pp.632–641.
- Steichen, S.D., Caldorera-moore, M. and Peppas, N.A., 2013. European Journal of Pharmaceutical Sciences A review of current nanoparticle and targeting moieties for the delivery of cancer therapeutics. *European Journal of Pharmaceutical Sciences*, 48(3), pp.416–427. Available at: <http://dx.doi.org/10.1016/j.ejps.2012.12.006>.
- Stoltenburg, R., Reinemann, C. and Strehlitz, B., 2007. SELEX-A (r)evolutionary method to generate high-affinity nucleic acid ligands. *Biomolecular Engineering*, 24(4), pp.381–403.
- Stoltenburg, R., Schubert, T. and Strehlitz, B., 2015. In vitro Selection and Interaction Studies of a DNA Aptamer Targeting Protein A. *PLoS ONE*, 10(7), pp.1–23.
- Sun, H. et al. 2014. Oligonucleotide Aptamers : New Tools for Targeted Cancer Therapy. , (April).
- Sun, H. and Zu, Y., 2015. A Highlight of recent advances in aptamer technology and its application. *Molecules*, 20(7), pp.11959–11980.
- Sutherland, C.L. et al. 2002. UL16-binding proteins, novel MHC class I-related proteins, bind to NKG2D and activate multiple signaling pathways in primary NK cells. *Journal of immunology (Baltimore, Md. : 1950)*, 168, pp.671–679.
- Sutherland, C.L., Chalupny, N.J. and Cosman, D., 2001. The UL16-binding proteins, a novel family of MHC class I-related ligands for NKG2D, activate natural killer cell functions. *Immunological reviews*, 181, pp.185–92. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11513139>.
- Svobodova, M., 2012. *SELECTION OF APTAMERS AGAINST PROSTATE*.
Universitat Rovira i Virgili.
- Svobodova, M. et al. 2012. Svobodova et al 2012_Comparison of different methods

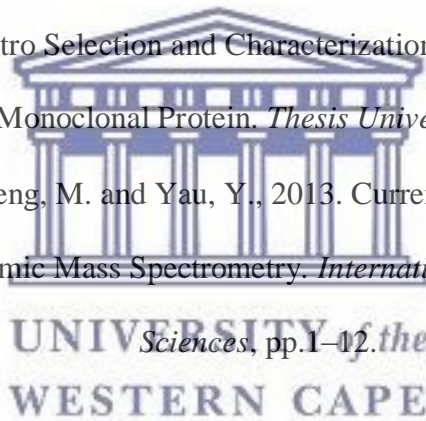
for generation of single-stranded DNA for SELEX processes - ResearchGate.pdf.

Analytical and Bioanalytical Chemistry, 404, pp.835–842.

- Tabarzad, M. et al. 2014. Challenges to design and develop of DNA aptamers for protein targets. I. optimization of asymmetric PCR for generation of a single stranded DNA library. *Iranian Journal of Pharmaceutical Research*, 13, pp.133–141.
- Tang, J.J. Xie, J.W. Shao, N.S. and Yan, Y. 2006. The DNA aptamers that specifically recognize ricin toxin are selected by two in vitro selection methods. *Electrophoresis*, 27, pp.1303–1311.
- Thiviyanathan, V. and Gorenstein, D.G., 2012. Aptamers and the next generation of diagnostic reagents. *Proteomics - Clinical Applications*, 6(11–12), pp.563–573.
- Tsai, T. et al. 2015. LC-MS / MS based Serum Proteomics for Identification of Candidate Biomarkers for Hepatocellular Carcinoma. *Proteomics*, 15(13), pp.2369–2381.
- Tsai, R.Y.L. and Reed, R.R. 1998. Identification of DNA recognition sequences and protein interaction domains of the multiple-Zn-finger protein Roaz. *Mol. Cell. Biol.* 18, pp.6447–6456.
- Valastyan, S. and Weinberg, Robert, A., 2011. Tumour Metastasis : Molecular Insights and Evolving Paradigms. *Cell*, 147(2), pp.275–292.
- Vater, A. Jarosch, F. Buchner, K. and Klussmann, S. 2003. Short bioactive Spiegelmers to migraine-associated calcitonin gene-related peptide rapidly identified by a novel approach: tailored-SELEX. *Nucleic Acids Res.* 31, e130.
- Vater, A. and Klussmann, S., 2015. Turning mirror-image oligonucleotides into drugs : the evolution of Spiegelmer 1 therapeutics §. *Drug Discovery Today*, 20(1), pp.147–155. Available at: <http://dx.doi.org/10.1016/j.drudis.2014.09.004>.
- Veigas, B., Doria, G. and Baptista, P., 2012. Nanodiagnosics for Tuberculosis.

Understanding Tuberculosis - Global Experiences and Innovative Approaches to the Diagnosis, pp.257–276. Available at: http://cdn.intechopen.com/pdfs/28542/InTech-Nanodiagnosics_for_tuberculosis.pdf.

- Virtanen, H. et al. 2005. The first cysteine-rich domain of the receptor GFR α 1 stabilizes the binding of GDNF The first cysteine-rich domain of the receptor GFR α 1 stabilizes the binding of GDNF. *Biochemical Society*, pp.1–8.
- Waldhauer, I. and Steinle, A., 2008. NK cells and cancer immunosurveillance. *Oncogene*, 27(45), pp.5932–43. Available at: <http://dx.doi.org/10.1038/onc.2008.267>.
- Waldhauer, I. and Steinle, A., 2006. Proteolytic release of soluble UL16-binding protein 2 from tumour cells. *Cancer Research*, 66(5), pp.2520–2526.
- Wang, Y., 2008. In Vitro Selection and Characterization of DNA Aptamers against a Multiple Myeloma Monoclonal Protein. *Thesis University of Goettingen*, p.119.
- Wasinger, V.C., Zeng, M. and Yau, Y., 2013. Current Status and Advances in Quantitative Proteomic Mass Spectrometry. *International Journal of Molecular Sciences*, pp.1–12.
- Waters, E.K. et al. 2011. Aptamer ARC19499 mediates a procoagulant hemostatic effect by inhibiting tissue factor pathway inhibitor. *Blood*, 117, pp.5514-5522.
- Weigel, M.T. and Dowsett, M., 2010. Current and emerging biomarkers in breast cancer: Prognosis and prediction. *Endocrine-Related Cancer*, 17(4).
- Weigelt, B., Peterse, J.L. and van 't Veer, L.J., 2005. Breast cancer metastasis: markers and models. *Nature reviews. Cancer*, 5, pp.591–602.
- White, R. et al. 2001. Generation of species cross-reactive aptamers using “toggle” SELEX. *Mol. Ther.* 4, pp.567–574.
- Wienken, C.J. et al. 2010. Protein-binding assays in biological liquids using



microscale thermophoresis. *Nat Commun*, 1(7), p.100. Available at:

<http://www.ncbi.nlm.nih.gov/pubmed/20981028>.

- Wu, X. et al. 2015. Aptamers : Active Targeting Ligands for Cancer Diagnosis and Therapy. *Theranostics*, 5(4), pp.322–344.
- Zaha, D.C. 2014. Significance of immunohistochemistry in breast cancer. *World Journal of Clinical Oncology*, 5(3), pp.382-392.
- Zhang, J., Basher, F. and Wu, J.D., 2015. NKG2D ligands in tumour immunity: Two sides of a coin. *Frontiers in Immunology*, 6(97), pp.1–7.
- Zijl, F. Van, Krupitza, G. and Mikulits, W., 2011. Initial steps of metastasis : Cell invasion and endothelial transmigration. *Mutation Research-Reviews in Mutation*

Research, 728(1–2), pp.23–34. Available at:

<http://dx.doi.org/10.1016/j.mrrev.2011.05.002>.

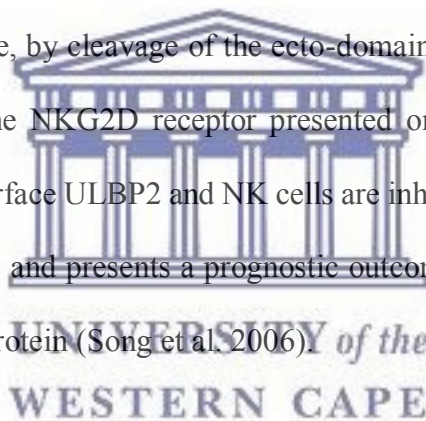


UNIVERSITY of the
WESTERN CAPE

Chapter Two: Large scale expression and purification of ULBP2

2.1. Introduction

It is the overexpression of the soluble form of ULBP2 (sULBP2) that has been reported to be found in the serum of breast cancer patients (Huergo-Zapico et al. 2014). In breast cancer cells, the mechanism leading up to the presentation of a poor prognostic outcome is first elicited by the overexpression of the cell surface ULBP2; which acts as a signal for the removal of cancerous cells, by means of cytotoxic action of NK cells (Sutherland et al. 2002; Song et al. 2006; Zhang et al. 2015). Therefore, high expression of the cell surface ULBP2 protein is associated with good prognosis. However, the cancerous cells have shown the ability to evade the tumor immune surveillance, by cleavage of the ecto-domain of ULBP2. It does this when the sULBP2 directly binds the NKG2D receptor presented on NK cells; subsequently the interaction between the cell surface ULBP2 and NK cells are inhibited, which in turn promotes the survival of cancerous cells and presents a prognostic outcome. Therefore, the interests of this study lie in the sULBP2 protein (Song et al. 2006).



Herein, we propose the use of a putative breast cancer serum marker, namely, ULBP2 that was identified at The University of the Western Cape, Department of Biotechnology. ULBP2 is a glycosylphosphatidyl-inositol (GPI) anchored cell surface protein, within a family of proteins that are recognized by the NKG2D receptor, expressed on NK cells and T cells, for apoptotic removal of cancerous cells (Leung et al. 2015). The aim of this study is to develop aptamers against the ecto-domain of ULBP2 that can be used in prognostic applications for breast cancer detection. Aptamers are short DNA/RNA oligonucleotides that can bind to a variety of targets, with high affinity and specificity (Khatri et al. 2010). In this study, the aptamers will serve as recognition moieties that bind to different aptatopes on the ecto-domain of ULBP2. High

quantities of the protein are required to perform the aptamer selection process and downstream characterization techniques; therefore, the aim of this chapter was to perform the large-scale expression and purification of the human recombinant ULBP2; using the pGEX-6P2 vector within the inducible *Escherichia coli* expression system. The ULBP2 protein was expressed as the soluble ecto-domain, with a c-terminal glutathione sepharose transferase (GST) tag; which would be required for partitioning and purification during the aptamer selection process.

2.1.1. The aim of this chapter

The aim of this chapter was to express and purify the human recombinant ULBP2 protein.

2.1.1. Objectives of this chapter

- To clone ULBP2 into pGEX-6P2.
- To transform *E.coli* XL1- Gold cells with pGEX-6P2-ULBP2.
- To isolate and purify pGEX-6P2-ULBP2.
- To transform *E.coli* BL21(DE3) cells with pGEX-6P2-ULBP2.
- To express and purify the ULBP2 protein from *E.coli* BL21(BE3) cells.



2.2. Materials

2.2.1. Chemicals and reagents

1 kb DNA ladder	Invitrogen
10 bp DNA ladder	Invitrogen
40% Acrylamide	Promega
Ammonium persulfate	Sigma-Aldrich
Ampicillin	Sigma-Aldrich
Agarose	WhiteHead Scientific
Bacteriological Agar	Merck
BamHI FastDigest Enzyme	Fermentas
Boric acid	Merck
Bromophenol blue	Sigma-Aldrich
cOmplete™, EDTA-free protease inhibitor cocktail tablet	Roche
Dithiothreitol	Sigma-Aldrich
Ethanol (99.9%)	Merck
Ethylenediaminetetraacetic acid	Merck
GelRed nucleic acid stain	Biotium
Glycerol	Merck
Isopropyl β-D-1 thiogalactopyranoside	Sigma-Aldrich
KAPA Taq Extra Hotstart ReadyMix	Kapa Biosystems
Lysozyme from chicken egg white	Sigma-Aldrich
Magnesium chloride	Roche
Phenylmethanesulfonyl fluoride	Sigma-Aldrich
Phosphate buffered saline powder	Sigma-Aldrich
Rattler Plating beads	Zymo Research



Recombinant ULBP2 clone	GenScript
Sodium Chloride	Sigma-Aldrich
Sodium Dodecyl Sulfate	Sigma-Aldrich
T4 DNA Ligase	Fermentas
T7 Gene 6 exonuclease	Thermo Fisher
Tetramthylethylenediamine	Sigma-Aldrich
Tfi DNA Polymerase	Invitrogen
Trisaminomethane	Merck
Triton-X-100	Sigma-Aldrich
Tryptone	Merck
XhoI FastDigest Enzyme	Fermentas
Yeast extract	Merck



UNIVERSITY of the
WESTERN CAPE

2.2.2. Buffers and solutions

4 X Protein sample buffer: 100 mM Tris-Cl (pH 6.8); 200 mM DTT; 4% SDS; 0.2% Bromophenol blue and 20% Glycerol and made up to a final volume of 8 ml, using dH₂O.

10 X TE: 10 mM Tris; 1 mM EDTA; adjusted to pH 7.4 and made up to a final volume of 100 ml using dH₂O. Autoclaved.

10 X FastDigest Green Buffer (Fermentas).

10 X SDS Running buffer: 4 mM Tris; 0.5 mM Glycine and 29 mM SDS; adjusted pH to 8.3 and made up to a final volume of 1 ltr, using dH₂O.

10 X T4 DNA Ligase Buffer (Fermentas).

10 X TBE: 1 M Tris; 1 M Boric acid; 0.02 M EDTA (pH 8); adjusted pH to 8.3 and made up to a final volume of 1 ltr using dH₂O.

10% APS: 0.1 g APS dissolved in 1 ml of dH₂O.

10% Triton-X-100: 1 ml of Triton-X-100 was dissolved in 10 ml of STE buffer.

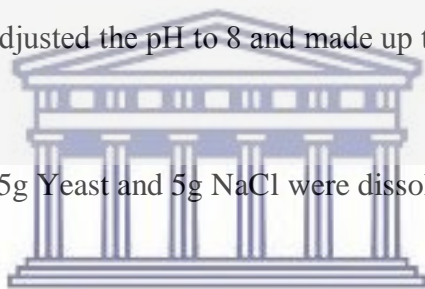
Coomassie Stain: 0.02% Coomassie blue; 40% Methanol; 10% Acetic acid and made to a final volume of 500 ml using dH₂O.

Destaining solution: 30% Methanol; 10% Acetic acid and made to a final volume of 500 ml using dH₂O.

Elution buffer: 20 mM Tris, pH 8; 10 mM NaCl; 10% glycerol; 0.1% Triton-X; 5 mM DTT; 50 µg/ml PMSF and 20 mM glutathione; adjusted the pH to 8 and made up to a final volume of 100 ml using dH₂O.

Equilibration buffer: 20 mM Tris, pH 8; 10 mM NaCl; 10% glycerol; 0.1% Triton-X; 5 mM DTT; 50 µg/ml PMSF; adjusted the pH to 8 and made up to a final volume of 100 ml using dH₂O.

Luria Broth: 10g Tryptone; 5g Yeast and 5g NaCl were dissolved in 1 ltr of dH₂O. Autoclaved.



STE buffer: 150 mM NaCl; 10 mM Tris; and 1 mM EDTA; adjusted to pH 8 and made up to a final volume of 100 ml using dH₂O.

SELEX buffer: PBS (0.138M NaCl and 0.0038M KCl), pH 7.4 (Sigma), 1.5mM MgCl₂, made to a final volume of 1litre using milli-Q water.

Tris, pH 6.8: 500 mM Tris was dissolved in 100 ml dH₂O with pH adjusted to 6.8.

Tris pH 8.8: 1.5M Tris was dissolved in 100 ml dH₂O with pH adjusted to 8.8.

2.2.3. Oligonucleotide sequences

Forward primer for the ecto-domain of ULBP2 (Fp_ULBP2) Inqaba Biotech

5'-GATTCGGATCCATGGACCCTCACTCTCTTT-3'

Reverse primer for the ecto-domain of ULBP2 (Rp_ULBP2) Inqaba Biotech

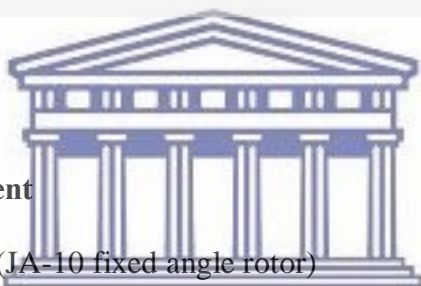
5'-TAACTCTCGAGCTATCCTGCACTTGGCTCC-3'

2.2.4. Bacterial expression vectors and strains

pGEX-6P2 expression vector Sigma-Aldrich

E. coli BL21(DE3) New England Biolabs

E. coli XL10 Gold New England Biolabs



2.2.5. Assay kits and equipment

Beckman Coulter Centrifuge (JA-10 fixed angle rotor) Beckman Coulter

Biospectrum Imaging System UVP

PCR and SV gel clean up kit Promega

NanoDrop ND-1000 Spectrophotometer
NanoDrop
Technologies

SDS-PAGE gel apparatus Bio-Rad

Thermocycler Bio-Rad

Wizard® Plus SV Minipreps DNA purification system kit Promega

2.3. Experimental methodology

2.3.1. Cloning of ULBP2 into pGEX-6P-2

2.3.1.1. Primer design

The overexpression of the soluble (ecto-domain) form of ULBP2 (sULBP2) has been reported to be found in the serum of breast cancer patients (Huergo-Zapico et al. 2014). Therefore, gene specific primers ([Section 2.2.3](#)) for the ecto-domain of ULBP2 (**Figure 2.1**) were designed from the coding sequence, obtained from NCBI, for the ULBP2 precursor (**Accession number: NM_025217.3**). The forward primer was designed to include the BamHI restriction site (G[^]GATCC) at the 5' end and the reverse primer was designed to include the XhoI restriction site (C[^]TCGAG) at its 3' end. These primers were purchased from Inqaba Biotech.



```
>NM_025217.3:136-876 Homo sapiens UL16 binding protein 2 (ULBP2), mRNA
ATGGCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTGCTGTCCGGC
TGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGTCATCCCTAAG
TTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGACTTTTCTTCAC
TATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAACTAAATGTCAC
AACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACTTACAGAGCAA
CTGCGTGACATTCAGCTGGAGAATTAGACACCCAAGGAACCCCTCACCTGCAGGCAAGGAT
GTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGTTTCGATGGGC
AGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCCTGGAGCCAGA
AAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATTACTTCTCAAT
GGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACCCCTGGAGCCAA
GTGCAGGAGCACCCTCGCCATGTCCTCAGGCACAACCCAACCTCAGGGCCACAGCCACCACC
CTCATCCTTTGCTGCCTCCTCATCATCCTCCCTGCTTCATCCTCCCTGGCATCTGA
```

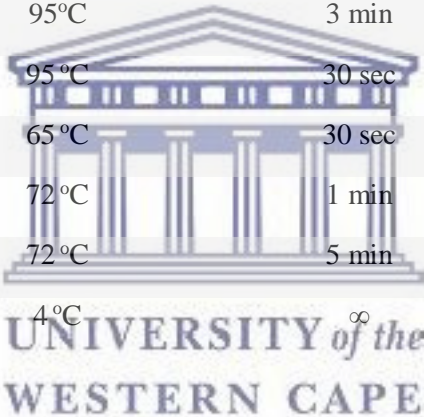
Figure 2.1. NCBI nucleotide sequence for the coding region of ULBP2. The ecto-domain of ULBP2 is indicated in yellow.

2.3.1.2. PCR amplification of ULBP2

PCR amplification was carried out using the KAPA Taq Extra Hotstart ReadyMix PCR kit (KAPA Biosystems). Reactions were setup according to the manufacturer's instructions. Each reaction mixture consisted of 1X KAPA Taq Extra Hotstart ReadyMix, 0.4 μ M Fp_ULBP2, 0.4 μ M Rp_ULBP2 and 200 ng of cDNA converted from mRNA that was extracted from MCF7 cells (kindly provided by Dr. Habeeb Bankole, Department of Biotechnology, University of Western Cape). The PCR reactions were performed in a final volume of 25 μ l.

Table 2.1: PCR parameters for the amplification of ULBP2

Parameters	Temperature	Time	No. of cycles
Initial denaturation	95°C	3 min	1
Denaturation	95°C	30 sec	
Annealing	65°C	30 sec	30
Elongation	72°C	1 min	
Final elongation	72°C	5 min	1
Incubation	4°C	∞	-



2.3.1.3. Agarose gel electrophoresis

A 1% (w/v) agarose gel was prepared by boiling the agarose in 1 X TBE until it dissolved. This solution was then cooled to \sim 55 $^{\circ}$ C, followed by the addition of 5X GelRed nucleic acid stain. At this point the agarose was poured into a casting gel tray and allowed to solidify. Once the gel solidified, a DNA ladder (5 μ l) was loaded in the first well, followed by loading the respective PCR amplicons (20 μ l) in the remaining wells. The samples were electrophoresed in 1 X TBE buffer at 100 V for 45 min. The gel was then visualised and images were captured using the Biospectrum Imaging system (UVP).

2.3.1.4. Gel purification of the ULBP2 PCR amplified product

The DNA fragment of interest electrophoresed, as described in **Section 2.3.1.3**, was excised using a sterile blade. The gel slice, containing the DNA fragment of interest, was macerated and transferred to a clean 1.5 ml micro-centrifuge tube. The DNA was isolated and purified using the PCR and SV gel clean up kit (Promega) according to the manufacturer's instructions:

The Binding Solution (700 μ l) was added to the macerated gel slice and heated at 55 °C for 10-15 min (with vortexing every 2-3 min) until the agarose was completely melted. This DNA solution was then added to the SV column, then placed in a collection tube and incubated for 1 min at 25 °C. The DNA solution was centrifuged (13200 x g; 1 min; 25 °C) and the flow through was discarded. The Membrane Wash Solution (700 μ l) was added to the column and the tube was centrifuged (13200 x g; 1 min; 25 °C). The flow through was discarded, followed by the addition of the Membrane Wash Solution (500 μ l) to the column and the tube was centrifuged (13200 x g; 5 min; 25 °C). The flow through was discarded. The column was placed into a clean 1.5 ml micro-centrifuge tube; 1X TB (50 μ l) was added to the column and incubated for 1 min at 25 °C. The latter was centrifuged (13200 x g; 2 min; 25 °C). The DNA collected in the 1.5 ml micro-centrifuge tube was quantified using the Nanodrop ND-1000 Spectrophotometer (NanoDrop Technologies) and stored at -20 °C.

2.3.1.5. Double digestion of pGEX-6P-2 and ULBP2 PCR product

Circularised pGEX-6P-2 plasmid DNA (kindly provided by Ms Tephney Hutchinson, Department of Biotechnology, University of the Western Cape) and the purified ULBP2 PCR product were, separately, digested with BamHI and XhoI FastDigest enzymes (Fermentas) as described in **Table 2.2**.

Table 2.2: Double digestion using BamHI and XhoI.

Reagents	Purified ULBP2 PCR product
Nuclease free water	11 μ l
10X FastDigest Green Buffer	2 μ l
DNA	500 ng
FastDigest BamHI	1 μ l
FastDigest XhoI	1 μ l
Total volume	20 μ l

The reagents were added in the order indicated in **Table 2.3** and incubated at 37 °C for 30 min. The digested samples were incubated at 80 °C for 5 min to inactivate the enzymes and were subsequently electrophoresed as described in **Section 2.3.1.3**.

2.3.2. The pGEX-6P-2 expression system

The pGEX-6P-2 protein expression vector is a bacterial expression system, designed for high-level expression of GST fusion proteins (Harper and Speicher 2011). The target gene is cloned into the multiple cloning site. The expression of the GST fusion protein is under the control of the *tac* promoter, which is chemically inducible by IPTG, a lactose analog. The vector contains an ampicillin resistance gene, used for the selection of successfully transformed *E. coli* cells.

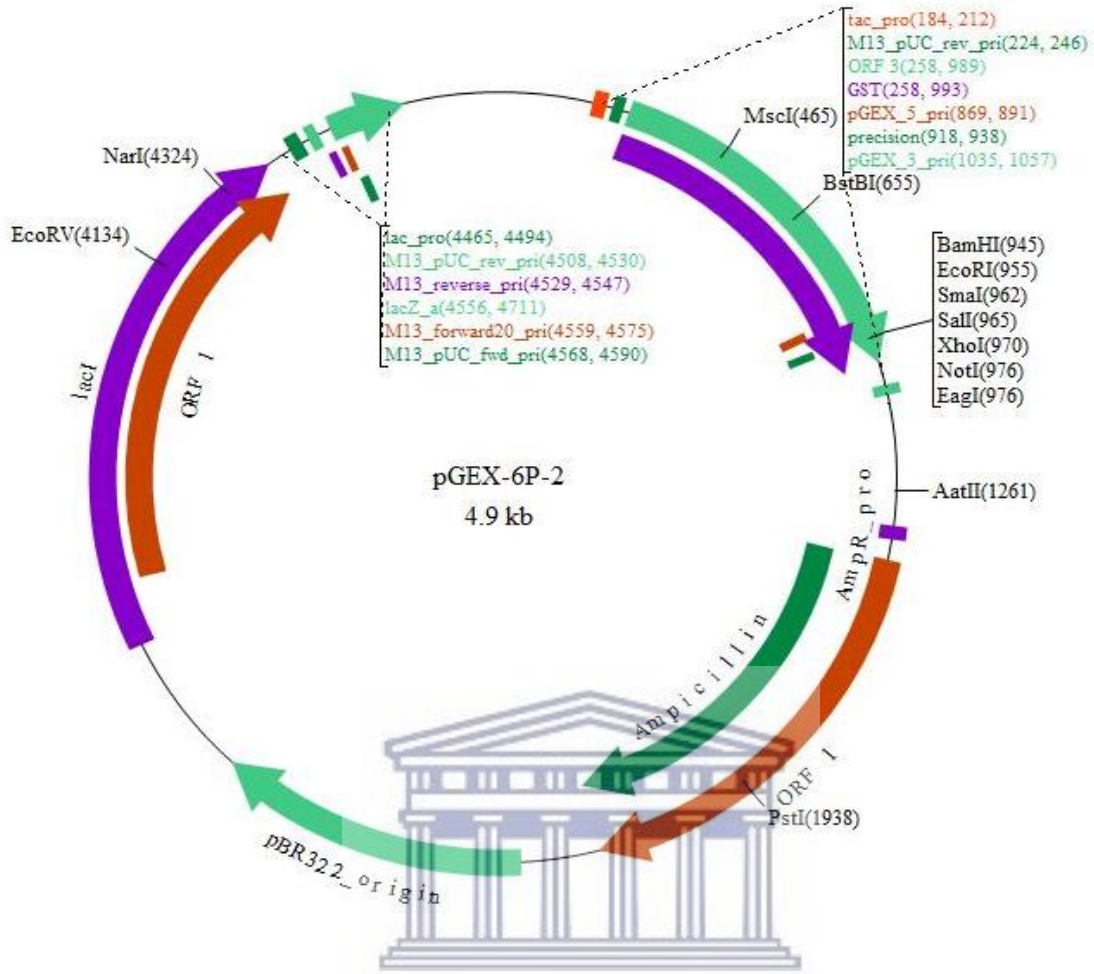


Figure 2.2: Circularised map and sequence referencing points of the pGEX-6P-2 expression vector (Harper and Speicher 2011).

2.3.2.1. Ligation of ULBP2 into the pGEX-6P-2 vector

The ligation reactions were set up, using an insert:vector ratio of 3:1, as described below:

Table 2.3: Reagents used in the ligation reaction using the pGEX-6P-2 vector

Reagent	Experimental	Vector only control
ULBP2 insert (200ng)	5 μ l	-
PGEX-6P2 vector (50ng)	2 μ l	2 μ l
10x T4 DNA ligase buffer (Fermentas)	2 μ l	2 μ l
Nuclease free water	10 μ l	15 μ l
T4 DNA ligase (5 U/ μ l)	1 μ l	1 μ l
Total volume	20 μ l	20 μ l

All reagents listed in **Table 2.4** were added to a 0.6 ml micro-centrifuge tube. The reaction tube was vortexed gently and incubated O/N at 4 °C.



2.3.2.2. Transformation of competent *E. coli* XL10-Gold cells with the pGEX-6P-2:ULBP2 construct

The competent *E. coli* XL10-Gold cells (kindly provided by Dr. Andrew Farou, Department of Biotechnology, University of the Western Cape) were thawed on ice. From the ligation reaction mixture prepared in **Section 2.3.2.1**, 2 μ l was added into a micro-centrifuge tube containing competent *E. coli* XL10-Gold cells (100 μ l), this was followed by gentle mixing. The mixture was placed on ice for 20 min, and subsequently heat shocked for 45 sec, at 42 °C. The mixture was, immediately, transferred to ice and incubated for a further 5 min. Thereafter, pre-warmed LB medium (900 μ l) was added to the mixture and incubated at 37 °C for 1 hour, while shaking. After incubation, 100 μ l of the bacterial culture was plated on LB-agar plates containing 100 μ g/ μ l of ampicillin (LB-amp). The remaining culture was centrifuged (13200 x g; 2 min; 25

°C). The pellet was re-suspended in 100 µl of the LB medium and plated on LB-amp agar plates. The plates were incubated O/N at 37 °C.

2.3.2.3. Colony PCR

Escherichia coli colonies (Section 2.3.2.2) were screened for the presence of the insert (ULBP2) using colony PCR as described below:

Two random colonies were selected from the plate (Section 2.3.2.2) and re-suspended in 5 µl of sterile dH₂O, respectively. The colony PCR was carried out as described in Section 2.3.1.2, with the amendment that the re-suspended colony (1 µl) was used as a template, instead of the cDNA. The PCR products were analysed by agarose gel electrophoresis, as described in Section 2.3.1.3.



2.3.2.4. Preparation of glycerol stocks

The remaining 4 µl (Section 2.3.2.3) for each positive clone was inoculated into 10 ml of LB-amp media, respectively. The cultures were incubated O/N at 37 °C, whilst shaking. Duplicate glycerol stocks, for each clone, were prepared by transferring 600 µl of the respective cultures to clean 1.5 ml micro-centrifuge tubes, followed by the addition of 400 µl of 80 % glycerol and inverting the tube a few times to mix. The glycerol stocks were stored at -80 °C for long term storage. The remaining culture was used to isolate the plasmid DNA.

2.3.2.5. Isolation of plasmid DNA

The Wizard® Plus SV Minipreps DNA purification system kit (Promega) was used for the purification of the pGEX-6P2-ULBP2 (p_{ULBP2}). The protocol was carried out according to the manufacturer's instructions:

The remaining O/N cultures prepared in **Section 2.3.2.4** were centrifuged (4800 x g; 15 min; 4°C). The cell pellets were re-suspended in the cell re-suspension solution (300 µl). The cells were then lysed by the addition of the cell lysis solution (500 µl). The cell suspensions were gently mixed and incubated at 25 °C for 5 min. The alkaline protease solution (20 µl) was added to the lysed cells and gently mixed, followed by 5 min incubation at 25 °C. The neutralisation solution (500 µl) was added to the cell lysate, mixed and incubated at 25 °C for 5 min. The cell lysates were centrifuged (13200 x g; 5 min; 25 °C) and the clear lysate (supernatant) was transferred into a column which was placed into a collection tube and incubated at 25 °C for 1 min. The flow through was discarded. The columns were washed by the addition of the column wash solution (750 µl), followed by centrifugation (13200 x g; 1 min; 25 °C). The flow through was discarded and the column was washed by the addition of column wash solution (250 µl). The tubes were centrifuged (13200 x g; 2 min; 25 °C). The flow through was discarded. The columns were then transferred to sterile 1.5 ml micro-centrifuge tubes. The plasmid DNA was eluted by the addition of 1 X TE (50 µl) to the column, which was then and incubated for 1 min at 25 °C. The tubes, containing the plasmid DNA, were centrifuged (13 200 x g; 2 min; 25 °C). The plasmid DNA collected in the 1.5 ml micro-centrifuge tubes were quantified using the Nanodrop ND-1000 Spectrophotometer (NanoDrop Technologies) and stored at -20 °C.

2.3.3. Small-scale protein expression of ULBP2

2.3.3.1. Transformation of recombinant clones into *E.coli* BL21(DE3) cells

Competent *E.coli* BL21(DE3) cells were thawed on ice. Once the cells were thawed, 100 µl of the cells were added to the p_ULBP2 recombinant clone, prepared in **Section 2.3.2.5**. A negative control (without DNA) was included. The cells and DNA mixture was left on ice for 20 min and subsequently subjected to heat shock at 42 °C for 45 sec; followed by immediate incubation on ice for 5 min. LB medium (900 µl), pre-warmed at 37 °C, was added to the cells

and DNA mixture; which was then incubated at 37 °C for 1.5 hrs. The respective cultures (100 µl) were plated onto LB-amp agar plates, using plating beads (Zymo Research). The remaining cultures (900 µl) were then centrifuged (13200 x g; 5 min; 25 °C). Approximately 800 µl of the supernatant was discarded, followed by re-suspending the cell pellet in the remaining supernatant (100 µl). These cells were then plated onto LB-amp agar plates, using plating beads. All the plates were incubated O/N, at 37 °C.

2.3.3.2. Protein expression screen of ULBP2

Four separate tubes containing LB-amp media (5 ml) were inoculated with 4 randomly selected colonies, respectively, from the plates prepared in **Section 2.3.3.1**. The cultures were grown O/N at 37 °C, whilst shaking. Four separate tubes containing LB-amp media (3 ml) was inoculated with 2 ml of the respective O/N cultures and incubated for 2 hrs at 37 °C, whilst shaking (this was carried out in duplicate). IPTG, at a final concentration of 1 mM, was added to one culture of each duplicate. Both the un-induced and IPTG-induced cultures were cultured for 4 hrs at 37 °C, whilst shaking. The cultures were transferred as aliquots (2 ml) into clean 2 ml micro-centrifuge tubes. The cells were harvested by centrifugation (13200 x g; 5 min, 25 °C). The supernatant was discarded and the cell pellets were re-suspended in 200 µl of protein sample buffer. The samples were boiled for 2 min at 95 °C and vigorously vortexed for 30 sec. This step was repeated for three cycles. The samples were then centrifuged (13200 x g; 5 min, 25 °C). The top layer (soluble proteins) for each sample was analysed by a 12 % (v/v) SDS polyacrylamide gel electrophoresis as described in **Section 2.3.3.3**. The clone indicative of the highest level of recombinant protein expression was selected for large-scale expression.

2.3.3.3. SDS polyacrylamide gel electrophoresis

Table 2.4. Reagents used for the preparation of two SDS polyacrylamide gels.

	12 % (v/v) Separating gel	5 % (v/v) Stacking gel
40% Acrylamide	3 ml	0.5 ml
0.5 M Tris, pH 8	-	0.63 ml
1.5 M Tris, pH 8	2.5 ml	-
10 % SDS	0.1 ml	0.05 ml
10 % Ammonium persulfate	0.2 ml	0.05 ml
dH₂O	4.3 ml	3.65 ml
TEMED	0.02 ml	0.01 ml
Total	10 ml	5 ml

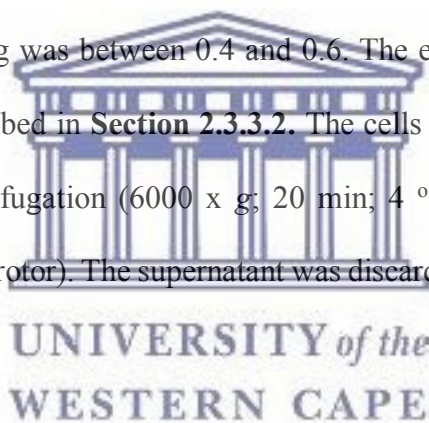
Two gels were prepared using the Bio-rad mini-gel apparatus (Bio-rad) according to the Laemmli SDS-PAGE protocol (Laemmli, 1970). The separating gel was prepared in a 15 ml Greiner tube by combining the reagents in the order listed in Table 2.4. The solution was mixed by inverting the tube a few times and then immediately poured between the gel plates. Ethanol was added to ensure that the shape of the gel was level when solidified. The ethanol was poured off and the gel was rinsed with dH₂O. The stacking gel was then prepared in a 15 ml Greiner tube and immediately poured on top of the separating gel. A ten-well comb, for each gel, was placed into the stacking gel solution. The stacking gel was then allowed to solidify. The gel plates, containing the gel, were removed from the casting trays and assembled in the electrophoresis system (Bio-rad). A final volume of 15 µl containing one part sample and one part sample loading buffer was heated at 95 °C for 2 min, followed by centrifugation (13200 x g; 5min, 25 °C). The top layer of the supernatant was loaded into the respective wells as well as the protein ladder (7 µl). The gel was electrophoresed in 1 X SDS running buffer at 110 V for 90 min. The gels were stained for 30 min with the coomassie staining solution; followed by

destaining O/N with the destaining solution. Thereafter, the gel was placed between two plastic sheets and scanned to a computer. Thereafter, the gel was removed from the plastic sheets and stored in distilled water. The gel was then visualised using white light and images were captured using the CanoScan LiDE 210 scanner (Canon).

2.3.4. Large-scale protein expression of ULBP2

2.3.4.1. Protein expression of ULBP2

The glycerol stock (100 μ l) of the best expressing clone in **Section 2.3.3.2** was used to inoculate 50 ml of LB-amp media. The cultures were incubated O/N at 37 °C, whilst shaking. The primary culture (50 ml) was scaled up to 2 ltr using LB-amp media; followed by incubation at 37 °C until the OD600 reading was between 0.4 and 0.6. The expression of the recombinant proteins was induced as described in **Section 2.3.3.2**. The cells were transferred to propylene tubes and harvested by centrifugation (6000 x g; 20 min; 4 °C) using a Beckman Coulter Centrifuge (JA-10 fixed angle rotor). The supernatant was discarded and the pellets were stored at -20 °C.



2.3.4.2. Preparation of the cell lysate

The cell pellet, prepared in **Section 2.3.4.1**, was homogenised in 10 ml of lysis buffer (STE buffer) by vigorous vortexing. This was followed by incubation on ice for 20 min. A final concentration of 1% Triton-X-100, 50 μ g/ml PMSF and a cOmplete™, EDTA-free protease inhibitor cocktail tablet was added to the cell lysate; followed by brief vortexing. The cell lysate was then subjected to 3 cycles of sonication (1 min sonication; 2 min on ice); followed by centrifugation (3750 x g; 15 min; 4°C). The supernatant was transferred to a clean 10 ml Greiner tube. The different fractions were analysed by SDS-PAGE as described in **Section 2.3.3.3** and stored at -20°C.

2.3.4.3. Solubilisation of the cell lysate

The cell lysate was prepared in the same manner as described in **Section 2.3.4.2** with the following amendment: N-laurylsarcosine (sarkosyl) or Triton-X-100, to a final concentration of 2% and 10%, respectively, was added prior to sonication. Each fraction (un-induced, soluble and insoluble fractions) were analysed by SDS-PAGE as described in **Section 2.3.3.3** and stored at -20°C.

2.3.5. Purification of the GST-tagged recombinant ULBP2 protein

2.3.5.1. Preparation of glutathione-agarose column

For the preparation of a 5 ml column, 0.35 mg of the glutathione-agarose powder was added to 70 ml of dH₂O and incubated O/N, at 4°C. The glutathione agarose slurry was carefully poured into a column. The resin was allowed to pack at the bottom of the column. Without letting the resin run dry at any point, it was equilibrated by washing with 10 X the resin bed volume (50 ml) of equilibration buffer.



2.3.5.2. Purification of the recombinant ULBP2 protein

The supernatant of the cell lysate prepared in **Section 2.3.4.3**, was loaded onto the glutathione resin prepared in **Section 2.3.5.1** and incubated on an end-over-end rotator for 1 hr, at 4°C. The cell lysate (10 ml) was passed through the column under gravity flow. This fraction was labelled flow through 1 (FL1) and stored on ice. The resin was washed with 2 X the resin bed volume of STE buffer. This was repeated for four cycles. Each fraction (10 ml), that passed through the column, was labelled as follows: flow through 2 (FL2), wash 1 (W1), wash 2 (W2) and wash 3 (W3) and stored on ice. Elution buffer (15 ml) was loaded onto the resin and the GST-tagged protein was eluted in 3 x 5 ml fractions and labelled as follows: elute 1 (E1), elute 2 (E2) and elute 3 (E3) and stored on ice. An aliquot (20 µl) of the resin was transferred to a

clean 1.5 ml micro-centrifuge tube and labelled 'resin'. Each fraction was analyzed by SDS-PAGE as described in **Section 2.3.3.3**.



2.4. Results and discussion

In order to accomplish the first objective of this chapter, the soluble domain of the ULBP2 gene was required to be cloned into a protein expression vector; and upon achieving this, the recombinant ULBP2 protein can be generated.

2.4.1. PCR amplification of ULBP2

Using gene specific primers, the soluble ecto-domain of ULBP2 (hereon referred to as ULBP2) was PCR amplified, as described in **Section 2.3.1.2**. A negative control (no DNA) was included to ensure that no non-specific products were amplified. As expected, no amplification was observed for the negative control, as demonstrated in Lane 1 (**Figure 2.3**). Lane 2, however, representing the amplification of the ULBP2, shows evidence of a PCR amplicon between 500 – 700 bp. The expected PCR product for ULBP2 is 552bp, therefore, the results obtained in **Figure 2.3**, suggests that ULBP2 was in fact amplified.

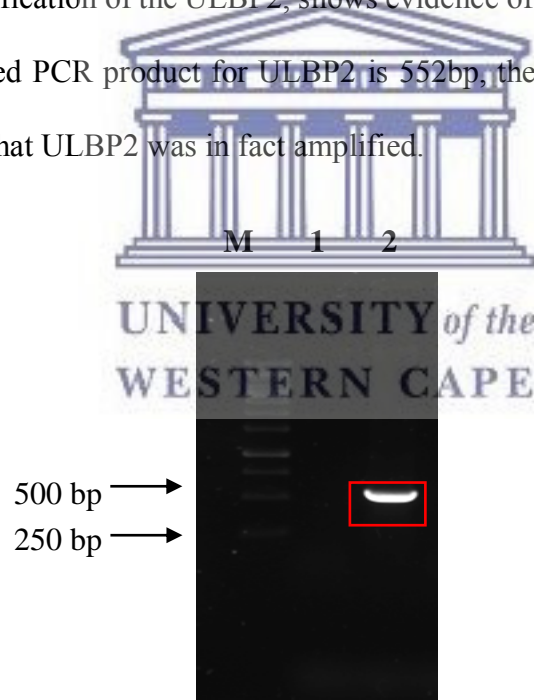


Figure 2.3: PCR amplification of ULBP2 resolved on a 1% agarose gel. Lane M: 1kb GeneRuler DNA ladder, Lane 1: Negative control (no DNA), Lane 2: PCR amplicon of the ULBP2. The expected PCR amplicon (552 bp) is indicated in the red block.

2.4.2. Cloning of ULBP2 into PGEX-6P2

After purification, the PCR amplicon and the expression vector (pGEX-6P2) were subjected to a double digestion, as described in Sections 2.3.1.4 and 2.3.1.5. Two factors were considered in trying to ensure that the ULBP2 gene would, efficiently, be taken up by vector: 1) The specific enzyme restriction sites used for this experiment are present in the multiple cloning site of the pGEX-6P2 vector, but not in the gene sequence for ULBP2 and were, therefore, included in the design of the primers and 2) following the digestion, a ligation experiment was setup, as described in Section 2.3.2.1, in which the insert (ULBP2) was present in excess concentrations, to promote the insertion of ULBP2 into the digested pGEX-6P2 vector. To confirm that ULBP2 was successfully ligated into pGEX-6P2, *E.coli* XL-Gold cells were transformed with the ligation product and screened for the presence of ULBP2 using colony PCR, as described in Sections 2.3.2.2 and 2.3.2.3. Lane 1, which represents the negative control (without colony), showed no amplification; whereas a distinct band between 500 – 700 bp was observed for both colonies that were screened (Lanes 2 and 3), suggesting and confirming the presence of ULBP2 (Figure 2.4) in these clones.

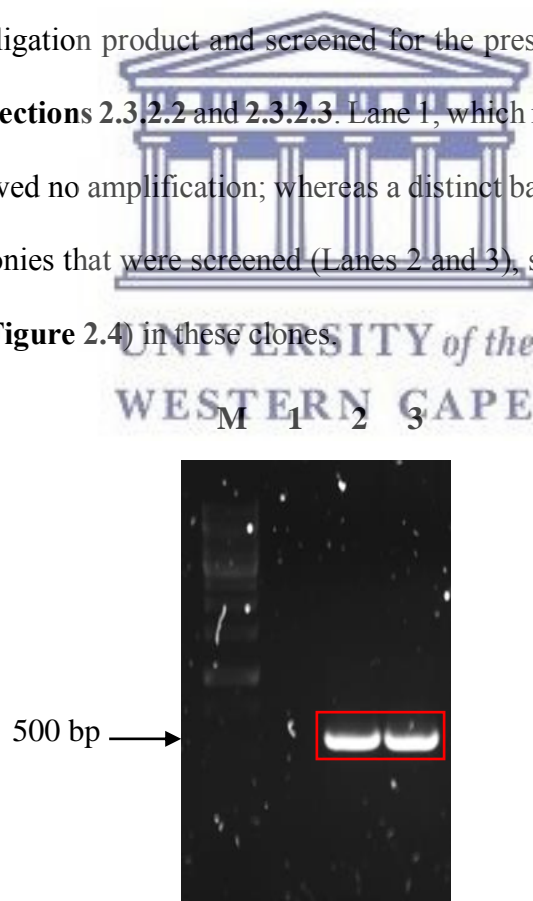
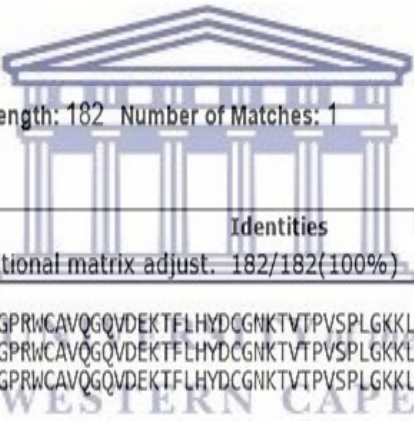


Figure 2.4: Colony PCR screen for the presence of ULBP2. Lane M: 1kb GeneRuler DNA ladder, Lane 1: Negative control (in absence of colony), Lanes 2 and 3: Colony PCR performed on colonies A and B, respectively. The expected PCR amplicon (552 bp) is indicated by the red block.

2.4.3. Sequence analysis of positive clones

To further validate that the positive clones did in fact contain the gene of interest and was free of mutations, plasmid DNA was extracted from these clones and sequenced at Inqaba Biotech. The DNA sequence information was analysed using FinchTV software. The DNA sequence was translated into a protein sequence using the online ExPASy translate tool; and finally, subjected to a pairwise sequence alignment, against the ULBP2 protein sequence, using the online NCBI protein BLAST software. The sequence alignment, for both clones, showed a 100% match to the expected protein sequence and this, in turn, verifies then that ULBP2 was present and no mutations were introduced during the amplification processes. A sequence alignment for one of the clones is shown below (**Figure 2.5**).



Sequence ID: lc|Query_46743 Length: 182 Number of Matches: 1

Range 1: 1 to 182 [Graphics](#) ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps
385 bits(988)	4e-142	Compositional matrix adjust.	182/182(100%)	182/182(100%)	0/182(0%)
Query 1	DPHSLCYDITVIPKFRPGPRWCAVQGVDEKTFLLHYDCGNKTVTPVSPLGKKLNVTTAWK			60	
Sbjct 1	DPHSLCYDITVIPKFRPGPRWCAVQGVDEKTFLLHYDCGNKTVTPVSPLGKKLNVTTAWK			60	
Query 61	AQNPVLRVVDILTEQLRDIQLENYTPKEPLTLQARMSCEQKAEGHSSGSWQFSFDGQIF			120	
Sbjct 61	AQNPVLRVVDILTEQLRDIQLENYTPKEPLTLQARMSCEQKAEGHSSGSWQFSFDGQIF			120	
Query 121	LLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFSMGDCIGWLEDFLMGMDSTLEPS			180	
Sbjct 121	LLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFSMGDCIGWLEDFLMGMDSTLEPS			180	
Query 181	AG 182				
Sbjct 181	AG 182				

Figure 2.5: Pairwise sequence alignment of the translated DNA sequence obtained for Clone A (query sequence) against the protein sequence for the human ULBP2 protein (subject sequence).

2.4.4. Small-scale protein expression of recombinant GST-ULBP2 fusion protein

To prepare sufficient quantities of the pGEX-6P2/ULBP2 construct for upstream experiments, a large-scale plasmid DNA isolation was performed. Thereafter, *E.coli* BL21(DE3) cells were transformed with the isolated plasmid DNA, as described in **Section 2.3.3.1**. A small-scale (5 ml) protein expression screen, induced using IPTG (**Section 2.3.3.2**), was then performed to identify the clones that express the recombinant GST-ULBP2 fusion protein. A negative control for each clone was included, which is basically a culture of the positive clone in the absence of IPTG. The clones were screened in duplicate, that is, two colonies per plate. The total bacterial lysates extracted from the clones expressed in *E.coli* BL21(DE3) cells were then subjected to SDS-PAGE, as described in **Section 2.3.3.3** to visualise the respective protein profiles for each clone.

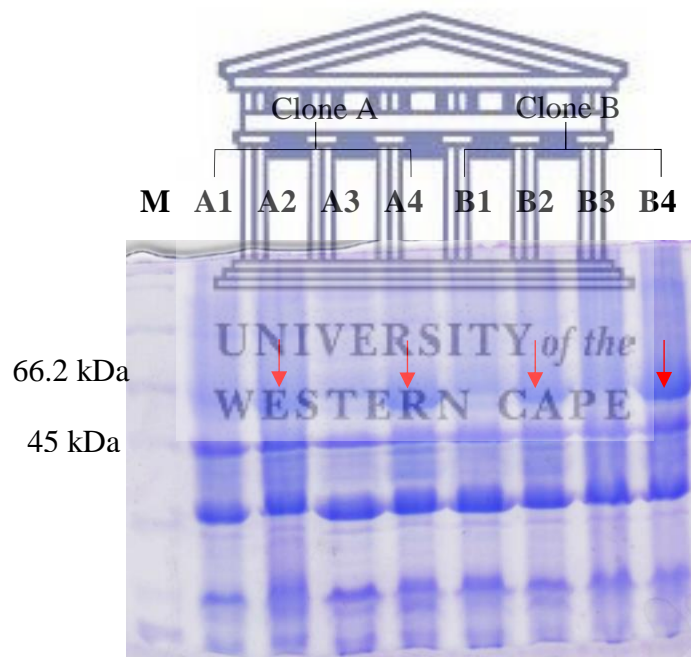


Figure 2.6: Small-scale protein for expression screen of the recombinant ULBP2. Lane M: PageRuler Protein ladder. Lanes A1 and A3 represents the un-induced colonies for clone A. Lanes A2, and A4 represents the induced colonies for clone A. Lanes B1 and B3 represents the un-induced colonies for clone B. Lanes B2, and B4 represents the induced colonies for clone B. The band of interest at ~59 kDa for ULBP2 is indicated by the red arrows. The experiment was performed in triplicate.

As shown in **Figure 2.6**, the expression of a protein, between 45 – 66.2 kDa was observed in all the clones. Although this protein is also present in the uninduced samples, the expression levels of the protein is much lower. This was expected because the IPTG inducible tac promoter, present in the pGEX-6P2 vector, resulted in the over-expression of the GST fusion protein. The GST tag is ~26 kDa and the expected molecular of ULBP2 is ~20 kDa; however, as a result of glycosylation, under reducing conditions the apparent molecular weight of ULBP2 is expected at ~33 kDa; therefore, the expected fusion protein is ~59 kDa. The observed result suggests that the recombinant GST-ULBP2 fusion protein was expressed in all of the clones.

2.4.5. Large-scale protein expression of recombinant GST-ULBP2 fusion protein

Based on **Figure 2.6**, colony A4 expressed much higher levels of the GST-ULBP2 fusion protein and was therefore selected for large-scale (2 ltr) protein expression, as described in **Sections 2.3.4.1**. In addition to scaling up the expression, the solubility of the recombinant protein was investigated by looking at both the soluble (proteins in the supernatant) and insoluble (proteins in the pellet) fractions. A negative control (expression in the absence of IPTG), as previously described in **Section 2.4.4**, was included. As seen with the small-scale expression, a higher expression of proteins was observed between 45 – 66.2 kDa, in the IPTG induced sample as compared to the un-induced sample (**Figure 2.7**); further suggesting that the recombinant GST-ULBP2 fusion protein was over-expressed. What was also observed is that the fusion protein was almost entirely expressed in the insoluble fraction as compared to the soluble fraction. These results suggest that the GST-ULBP2 fusion protein may be entrapped in inclusion bodies. This is supported by the fact that although *E.coli* is a reliable, highly efficient and well-established bacterial expression strain, its use for the over-expression of recombinant proteins can, in some cases, result in substantial aggregation of proteins, from

which inclusion bodies are formed (Palmer and Wingfield, 2004; Sørensen and Mortensen, 2005). One way to combat this is to solubilise the proteins within the insoluble fraction by employing an anionic detergent, such as sarkosyl or a non-ionic detergent, such as Triton-X, at low concentrations (Kim et al. 2013; Singh et al. 2015; Wingfield et al. 2015). Therefore, an attempt was made, using the above-mentioned detergents, to solubilise the protein of interest.

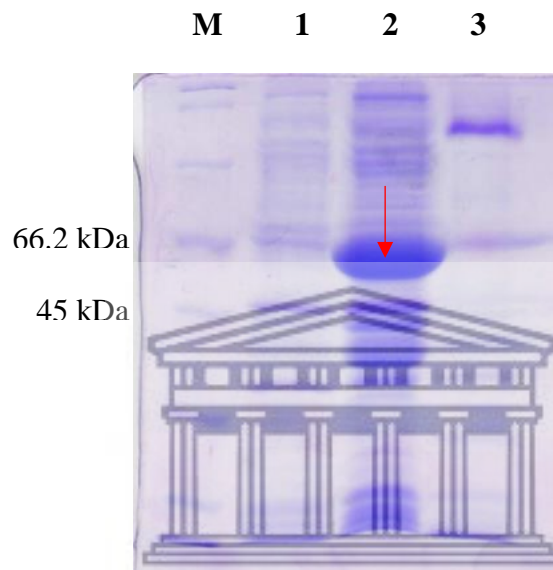


Figure 2.7.: Large-scale protein expression screen of the recombinant ULBP2. Lane M: PageRuler Protein ladder. Lanes 1: Un-induced culture. Lane 2: Insoluble fraction. Lane 3: Soluble fraction. The band of interest at ~59 kDa for ULBP2 is indicated by the red arrow. The experiment was performed in triplicate.

2.4.6. Solubility screen of recombinant GST-ULBP2 fusion protein

A large-scale protein expression was performed and the resulting protein extracts were further subjected to solubilisation using the respective detergents, as described in **Section 2.3.4.3**. The protein profile for the protein lysate, treated with Triton-X, yielded no detectable solubilisation of the protein of interest (**Figure 2.8**). In contrast, the protein profile for the protein lysate, treated with sarkosyl, showed a significant amount of the protein of interest was solubilised,

suggesting that the putative GST-ULBP2 recombinant protein may have been successfully solubilised from the inclusion bodies.

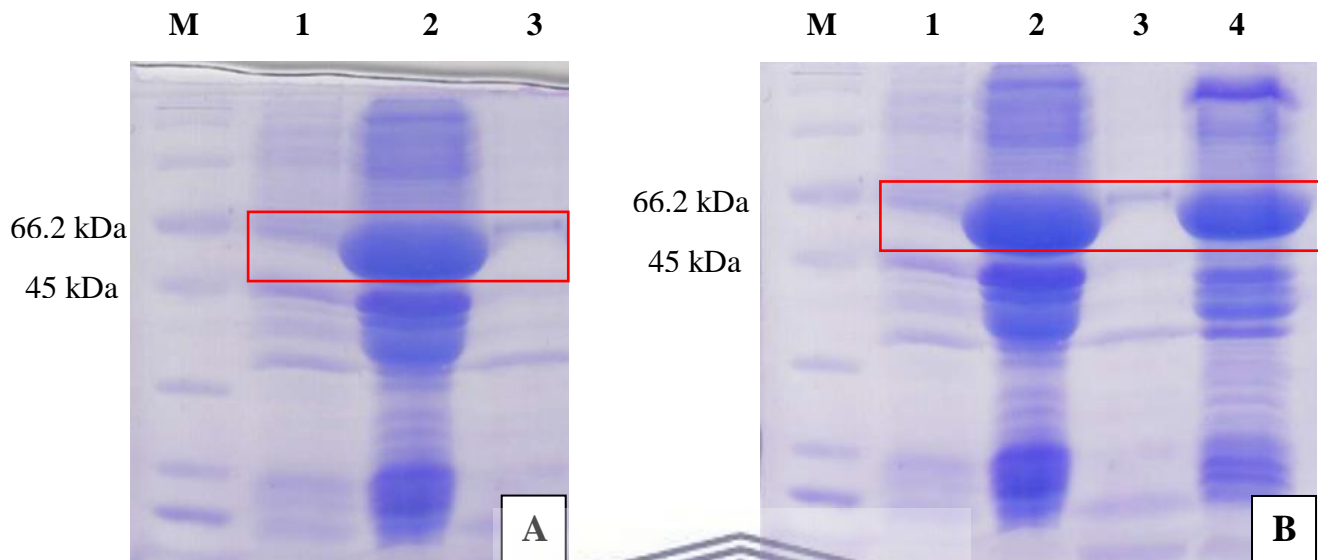


Figure 2.8: Solubilisation of the recombinant ULBP2 protein. Lane M: PageRuler Protein ladder. **A** represents the protein solubilised using 10% Triton X-100. Lane 1: un-induced culture. Lane 2: insoluble fraction. Lane 3: soluble fraction. **B**: represents the protein solubilised using 2% sarkosyl. Lane 1: un-induced culture. Lane 2: insoluble fraction. Lane 3: soluble fraction (without sarkosyl). Lane 4: soluble fraction (with sarkosyl). The band of interest at ~59 kDa for ULBP2 is indicated by the red block. The experiment was performed in triplicate.

UNIVERSITY of the
WESTERN CAPE

2.4.7. Purification of the recombinant GST-ULBP2 fusion protein

Since the recombinant protein contained the C-terminal GST tag, it could facilitate affinity purification using glutathione agarose beads. The GST fused to the protein of interest binds, with high affinity, to the glutathione molecules immobilised on the agarose bead purification column. Any proteins deficient of the GST tag are washed away and the GST fusion protein can be eluted by the addition of glutathione, at higher concentrations than what is present on the agarose beads. Consequently, competitive binding for the GST fusion protein occurs, which in turn, results in the elution of the protein of interest (Harper and Speicher, 2011). Taking this into account, the solubilised recombinant protein generated in **Section 2.4.6** was subjected to

the GST affinity purification, as described in **Section 2.3.5.2**. What was observed is that the putative GST-ULBP2 fusion proteins, seen at ~59 kDa, was present in all the flow through and some wash fractions, that is, lanes 1-4; however, no detectable protein was observed in the elution fractions (**Figure 2.9**). This suggests that no binding occurred between the GST-ULBP2 fusion protein and the glutathione column. Consequently, the protein of interest was not purified, but rather washed off the column prior to elution.

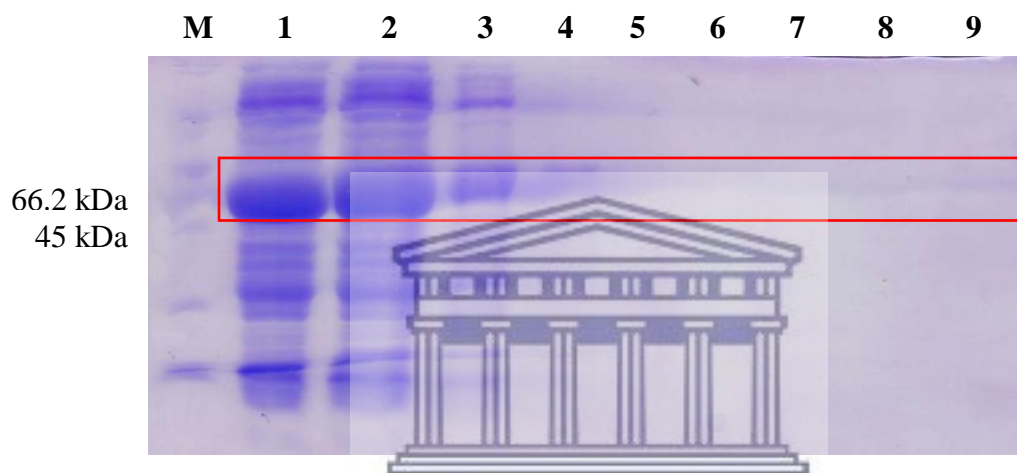


Figure 2.9. Affinity purification of the sarkosyl solubilised recombinant GST-ULBP2 fusion protein. Lane M: PageRuler Protein ladder. Lane 1: Total protein lysate. Lane 2: Soluble fraction. Lane 3: Flow through 1. Lane 4: Flow through 2. Lane 5-7: Wash 1-3. Lane 8: Elute. Lane 9: Uncleaned agarose beads. The band of interest at ~59 kDa for ULBP2 is indicated by the red block. The experiment was performed in triplicate.

In an attempt to improve the solubility of the GST-ULBP2 fusion protein, several alternative methods of expression and purification were evaluated. These include, 1) using a lower concentration of the sarkosyl; 2) increasing concentrations of glutathione in the elution buffer; 3) altering the pH of the elution buffer and 4) using a combination of sarkosyl, Triton-X and CHAPS as demonstrated by Tao *et al.* 2010 to significantly increase the solubility of GST fusion proteins (Tao et al. 2010); 5) lowering the IPTG concentration; 6) lowering the temperature at which the expression was performed and 7) unfolding and refolding the protein

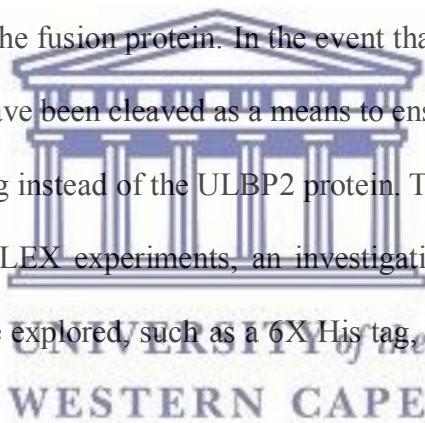
using 6M urea. However, none of these attempts improved the binding of the protein of interest to the column nor the solubility of the protein of interest.

Although the recombinant protein was successfully solubilised, it is possible that the concentration and the anionic nature of sarkosyl, required for efficient solubilisation, was too high and thus resulted in the denaturation or incorrect folding of the recombinant GST-ULBP2 fusion protein. Consequently, inhibiting the affinity purification process as observed with studies by Palmer and Wingfield, 2004; Schlager et al. 2012; Kimple et al. 2015. Another reason could be due to the fact that the ULBP family of proteins are, generally, expressed in inclusion bodies using *E.coli* as the expression host (Cao et al. 2008; Müller et al. 2010; Watson et al. 2011; Mao et al. 2012). However, as previously mentioned, although the most widely used recombinant protein expression host, a major drawback for *E.coli* is that the resultant protein of interest can be entrapped in inclusion bodies (Bernaudat et al. 2011). This can occur because the expression rate is significantly high and places stress on the host cell, resulting in protein misfolding and ultimately, aggregation and the formation of inclusion bodies (Costa et al. 2014). A eukaryotic expression host, may be a more plausible approach as successfully demonstrated by the use 293T cells (commonly known as HEK cells), which was used for the expression and purification of ULBP2 by Strandmann et al. (2006) and Rothe et al. (2013) (Strandmann et al. 2006; Rothe et al. 2013). It is the most commonly used eukaryotic expression host, specifically for the expression of heterologous membrane proteins (Ooi et al. 2016).

If the possibility of using sarkosyl as a means of solubilisation is ruled out, then perhaps an entirely different strategy must be employed to recover the functional rULBP2 fusion protein. The most commonly used method is the refolding of proteins expressed in inclusion bodies.

And although this approach may be cumbersome, since inclusion bodies contain a reasonable quantity of pure and intact proteins, it offers the possibility of producing high concentrations of good quality and bio-active recombinant proteins (Yamaguchi and Miyazaki, 2014; Singh et al 2015).

Because GST fusion proteins equate to high expression levels in *E.coli*, typically ~10mg/litre, simply by employing the GST tag, a probability exists for the accumulation of insoluble aggregated fusion proteins (Kimple et al. 2015). Furthermore, a GST tag is quite large and its dimerization in solution may also influence the properties of the fusion protein (Kimple et al. 2015). Effectively, in this study, the GST tag may also be responsible for the insolubility and potentially, functional loss of the fusion protein. In the event that the protein was successfully purified, the GST tag would have been cleaved as a means to ensure that the aptamer selection was not directed at the GST tag instead of the ULBP2 protein. Therefore, since GST tag is not necessary for downstream SELEX experiments, an investigation into the employment of a different affinity tag should be explored, such as a 6X His tag, which is also commonly used (Kimple et al. 2015).



A limitation recognised in this study is that Western blot was not employed to ensure that the protein in fact corresponded to the expected ULBP2 protein. It could be that what appeared as the over-expression of ULBP2 may have been overloading of the sample, which again could have been clarified using Western blot screening.

The objective of using the recombinant protein expression technology was to produce high quantities of the GST-ULBP2 fusion protein; such that it could be used for downstream experiments in SELEX. However, since it required further optimisation, that could potentially

be time consuming – in addition to the time already spent attempting to express the GST-ULBP2 fusion protein – at this point it was decided to acquire a commercially available ULBP2 ecto-domain protein to proceed with the aptamer selection.

2.5. Conclusions and future recommendations

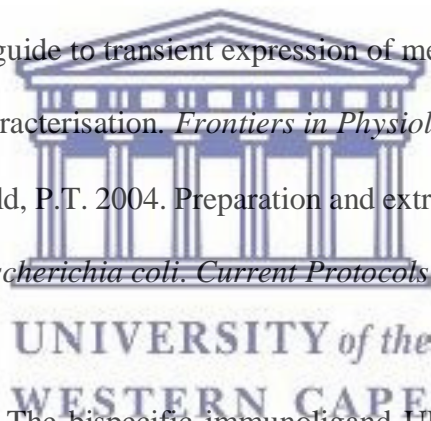
The cloning and large-scale expression of the ecto-domain of ULBP2, in *E.coli*, appeared successful, however, may strongly have been validated by the employment of a western blot screen. The study was significantly hindered by the solubility problems encountered upon attempting to purify ULBP2. The use of *E.coli* as an expression host, largely, contributed to the resultant protein expression in inclusion bodies (Bernaudat et al. 2011). Even though solubilisation of ULBP2 was achieved, the minimal concentration of sarkosyl required to solubilise ULBP2 may have hindered the binding of the GST tag, present on the fusion GST-ULBP2 protein, to the glutathione column (Palmer and Wingfield, 2004; Schlager et al. 2012; Kimple et al. 2015). Therefore, an alternative approach to purify ULBP2 is required. These include and are not limited to: 1) purifying the inclusion bodies and refolding the GST-ULBP2 fusion protein; 2) investigating whether a different expression vector/affinity tag may result in the expression of soluble ULBP2 and 3) exploring the use of a eukaryotic expression host.



2.6. References

- Adeloje, D. et al. 2018. Estimating the incidence of breast cancer in Africa: a systematic review and analysis. *Journal of Global Health*, 8(1), pp1-21.
- Aquino-Jarquin, G. and Toscano-Garibay, J.D. 2011. RNA aptamer evolution: Two decades of selection. *International Journal of Molecular Science*, 12(12), pp. 9155-9171.
- Bernaudat, F. et al. 2011. Heterologous expression of membrane proteins: choosing the appropriate host. *PLOS ONE*, 6(12), pp. 1-17.
- Bleyer, A., Baines, C. and Miller, A.B. 2015. Impact of screening mammography on breast cancer mortality. *International Journal of Cancer*, 138(2015), pp.2003-2012.
- Cao, W. et al. 2008. Four novel ULBP splice variants are ligands for human NKG2D. *International Immunology*, 20(8), pp. 981-991.
- Costa, S., Almeida, A., Castro, A. and Domingues, L. 2014. Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel FH8 system. *Frontiers in Biotechnology*, 5(63), pp. 1-20.
- Harper, S. and Speicher, D.W. 2011. Purifications of proteins fused to glutathione S-transferase. *Methods in Molecular Biology*, 681, pp. 259-280.
- Huergo-Zapico, L. et al. 2014. Molecular bases for the regulation of NKG2D ligands in cancer. *Frontiers in Immunology*, 5(106), pp.1-7.
- Khati, M., 2010. The Future of Aptamers in Medicine. *Journal of clinical pathology*, 63(6), pp.480-7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20360137>.
- Kim, M.J et al. 2013. Complete solubilisation and purification of recombinant human growth hormone produced in *Escherichia coli*. *PLOS ONE*, 8(2), pp. 1-8.
- Kimple, M.E. et al. 2013. Overview of affinity tags for protein purification. *Current Protocols in Protein Science*, 24(73), pp. 1-9.

- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227(5259), 680-685.
- Leung, W. H. et al. 2015. PRL-3 mediates the protein maturation of ULBP2 by regulating the tyrosine phosphorylation of HSP60. *The Journal of Immunology*, 194, pp. 2930 – 2941.
- Mao, C. et al. 2012. A monoclonal antibody against human UL16-binding protein 3. *Hybridoma*, 31(3), pp. 203-208.
- Müller, S., Zocher, G., Steinle, A. and Stehle, T. 2010. Structure of the HCMV UL16-MICB complex elucidates select binding of a viral immunoevasin to diverse NKG2D ligands. *PLOS Pathogens*, 6(1), pp. 1-12.
- Ooi, A. et al. 2016. A guide to transient expression of membrane proteins in HEK-293 cells for functional characterisation. *Frontiers in Physiology*, 7 (300), pp. 1-15.
- Palmer, I. and Wingfield, P.T. 2004. Preparation and extraction of insoluble (insoluble-body) proteins from *Escherichia coli*. *Current Protocols in Protein Science*, Chapter 6 (Unit 6.3), pp 1-25.
- Rothe, A. et al. 2014. The bispecific immunoligand ULBP2-aCEA redirects natural killer cells to tumor cells and reveals potent anti-tumor activity against colon carcinoma. *International Journal of Cancer*, 134, pp. 2829-2840.
- Singh, A. et al. 2015. Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilisation process. *Microbial Cell Factories*, 14(41), pp. 1-10.
- Schlager, B., Straessle, A. and Hafen, E. 2012. Use of anionic denaturing detergents to purify insoluble proteins after overexpression. *BMC Biotechnology*, 19(95), pp. 1-7.
- Song, H. et al. 2006. Soluble ULBP suppresses natural killer cell activity via down-regulating NKG2D expression. *Cellular Immunology*, 239(1), pp.22–30.



- Sørensen, H.P. and Mortensen, K.K. 2005. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial Cell Factories*, 4(1), pp. 1-8.
- Sutherland, C.L. et al. 2002. UL16-binding proteins, novel MHC class I-related proteins, bind to NKG2D and activate multiple signaling pathways in primary NK cells. *Journal of immunology (Baltimore, Md. : 1950)*, 168, pp.671–679.
- Tao, H. et al. 2010. Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100 and CHAPS. *Biotechniques*, 48, pp. 61-64.
- von Strandmann, E. et al. 2006. A novel bispecific protein (ULBP2-B4) targeting the NKG2D receptor on natural killer cells and CD138 activates NK cells and has potent antitumor activity against human multiple myeloma in vitro and in vivo. *Immunobiology*, 107(5), pp. 1955-1962.
- Watson, A.A., Christou, C.M. and O'Callaghan, C.A. 2011. Expression, purification and crystallisation of the human UL16-binding protein ULBP1. *Protein Expression and Purification*, 79, pp. 44-48.
- Wingfield, P. T. 2015. Overview of the purification of recombinant proteins. *Current Protocol in Protein Science*, 80, pp. 1-50.
- Yamaguchi, H. and Miyazaki, M. 2014. Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules*, 4(1), pp. 235-251.
- Zhang, J., Basher, F. and Wu, J.D. 2015. NKG2D ligands in tumor immunity: two sides of a coin. *Frontiers in Immunology*, 6(97), pp. 1-7.



Chapter Three: Aptamer selection against ULBP2 and GFR α 1

3.1. Introduction

The aim of this chapter was to develop aptamers against two protein targets, namely, ULBP2 and GFR α 1, respectively; which can be applied to prognostic platforms for the detection of breast cancer. The aptamers will serve as a recognition moiety that binds to the protein target. Aptamers are short DNA/RNA oligonucleotides or peptides that can fold into three dimensional structures and bind a variety of targets with high affinity and specificity (Mayer, 2009; Khati, 2010; Jauset-Rubio et al. 2016). Aptamers have emerged as an alternative recognition moiety to antibodies, due to several advantages (Liu et al. 2006; Khati, 2010; Jauset-Rubio et al. 2016). Aptamers are selected using a stringent *in vitro* process, known as Systematic Evolution of Ligands by Exponential enrichment (SELEX); which consists of three fundamental steps, that is, 1) adsorption to the target, 2) partitioning of target-bound from the unbound sequences and 3) exponential enrichment of target-bound sequences (Gopinath et al. 2007; Mairal et al. 2008; Wang et al., 2008; Lakhin et al. 2013). There are several types of SELEX, of which a more detailed description of is provided in Chapter 1.

Herein, classical SELEX, with the incorporation of negative and counter SELEX, was employed, to select for aptamers against the recombinant 6X His-tagged ULBP2 protein and the 6X His-tagged GFR α 1 protein, respectively. Subsequent analysis, using the Illumina next generation sequencing (NGS) platform, was then employed to identify potential candidate aptamers.

3.1.1. The aim of this chapter:

The aim of this chapter was to develop aptamers against ULBP2 and GFR α 1, respectively.

3.1.2. The objectives of this chapter

- To select for DNA aptamers against the recombinant 6X His-tagged ULBP2, using *in vitro* systematic evolution of ligands by exponential enrichment.
- To select for DNA aptamers against the recombinant 6X His-tagged GFR α 1 protein, using *in vitro* systematic evolution of ligands by exponential enrichment.
- To identify the candidate aptamers sequences for ULBP2 and GFR α 1, respectively, using Next Generation Sequencing.



3.2 Materials and Methods

3.2.1. Chemicals and reagents

2X Kapa Taq Extra HotStart Ready Mix with dye	Kapa Biosystems
6X DNA loading dye	Thermo Fisher Scientific
10bp DNA ladder	Invitrogen
Bovine Serum Albumin	Sigma-Aldrich
Certified™ Low Range Ultra Agarose	Bio-Rad
Ethidium Bromide	Thermo Fisher Scientific
Human Serum Albumin	Sigma-Aldrich
Hydrochloric acid	Merck
Immunoglobulin G	Sigma-Aldrich
Magnesium Chloride	Sigma-Aldrich
NEBuffer 4	New England Biolabs
Nuclease free water	Sigma-Aldrich
O'GeneRuler 100bp DNA ladder	Thermo Fisher Scientific
Phosphate buffered saline sachets	Sigma-Aldrich
SiMAG-IDA-Nickel beads	Chemicell
T7 gene 6 exonucelase	New England Biolabs
Trisaminomethane	Merck



3.2.2. Buffers

10X PCR buffer: 500mM KCl, 100mM Tris-HCL (pH 8.3), 1% Triton X-100, 25mM MgCl₂, made to a final volume of 50ml. Filter sterilise using a 0.2 micron filter.

10 X TBE: 1 M Tris; 1 M Boric acid; 0.02 M EDTA (pH 8); adjusted pH to 8.3 and made up to a final volume of 1 ltr using dH₂O.

EDC/NHS solution: 400mM EDC; 100mM

Regeneration buffer: 2M NaCl; 50mM NaOH buffer

SELEX buffer: PBS (0.138M Nacl and 0.0038M KCl), pH 7.4 (Sigma), 1.5mM MgCl₂, made to a final volume of 1litre using milli-Q water.

Sodium acetate: (NaAc), pH 4.5:



3.2.3. Oligonucleotide sequences

Library III (LibIII) Biomers

5'-AGCTCCAGAAGATAAATTACAGG – N(53) – CAACTAGGATACTATGACCCC-3'

Forward Primer, LibIII, 5' phosphorothiate modified (F_pto) Biomers

5'CCCAAGCTTAATACGACTCACTATAGGGAGCTCCAGAAGATAAATTACA-3'

Reverse Primer, LibIII (R_LibIII) Biomers

5'-GGGGTCATAGTATCCTAGTTG-3'

3.2.4. Proteins and antibodies

Recombinant human GFR α 1 His-tagged protein Sino Biological Inc

Recombinant human ULBP2 His-tagged protein Sino Biological Inc

3.2.5. Kits and Equipment

Oligo Clean and Concentrator™	Zymo Research
BIAcore 3000	GE Healthcare
UVP Transilluminator	UVP
Thermocycler	Bio-rad

3.3. Experimental methodology

3.3.1. SELEX

3.3.1.1. Conjugation of LibIII to the target protein

The SELEX protocol was adapted from Svobodová, et al. 2012.

The conjugation of LibIII to the target was carried out as follows: 300 pmol of LibIII was added to 100 µl of SELEX buffer (PBS, pH 7.4; 1.5 mM MgCl₂). The LibIII was then denatured at 95°C for 5 min and gradually cooled, in increments of 10°C every 30 sec, down to 25°C. The first round of SELEX was then carried out by incubating the denatured LibIII with 30 pmol of the respective target protein (Sino Biological Inc.), for 30 min at 25°C while shaking at 500 rpm.

3.3.1.2. Immobilization of the LibIII/protein complex to the magnetic beads

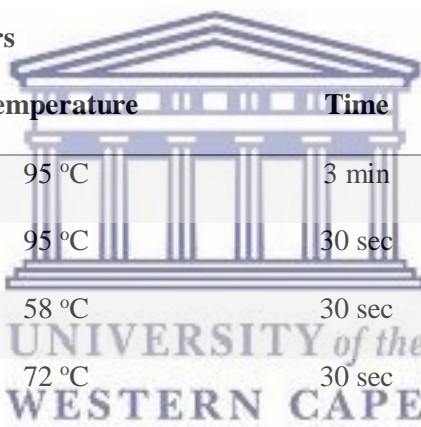
The immobilization was carried out by incubating ~0.15 mg of the SiMAG-IDA-Nickel beads with the LibIII/protein complex, for 30 min at 25°C, whilst shaking at 500rpm. To remove any unbound DNA, the immobilized LibIII/protein complex was isolated by incubation on a magnet for 1-2 min. The bead complex was washed, three times, by re-suspension in SELEX buffer (100 µl), followed by isolation using the magnet. The bead complex was then re-suspended in 25 µl of nuclease free water.

3.3.1.3. Pilot PCR

To determine the optimal number of cycles, a pilot PCR was carried out as follows:

A PCR master mix was prepared for six samples by addition of the following reagents at final concentration of 1 X Kapa Taq Extra HotStart Readymix with dye, 3 mM MgCl₂, 0.5 mg/ml BSA, 0.2 μM Forward primer (F_pto), 0.2 μM Reverse primer (R_LibIII) and made up to a final volume of 125 μl. For the negative control, a 25 μl aliquot was transferred to a thin-walled PCR tube. To the remaining master mix, 5 μl of the bead complex, prepared in **Section 3.3.1.2** was added and further aliquoted (25 μl) into 5 x thin-walled PCR tubes and subjected to PCR with the parameters as described in the table below:

Table 3.1: Pilot PCR parameters



	Temperature	Time	No of cycles
Initial denaturation	95 °C	3 min	1
Denaturation	95 °C	30 sec	(5, 10, 15, 20 and 25, respectively)
Annealing	58 °C	30 sec	
Elongation	72 °C	30 sec	

3.3.1.3.1. Agarose gel electrophoresis

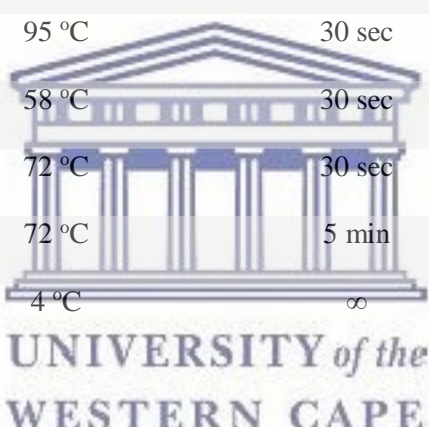
A 3 % (w/v) agarose gel was prepared by boiling the agarose in 1 X TBE until it dissolved. This solution was then cooled to ~55 °C, followed by the addition of 0.4 μg/ml UltraPure™ Ethidium Bromide. At this point the agarose was poured into a casting gel tray and allowed to solidify. Once the gel solidified, the DNA ladder (2 μl) and each sample (10 μl) was loaded into the respective wells and electrophoresed in 1 X TBE buffer at 110 V for 35 min. The gel was then visualised and images were captured using the Biospectrum Imaging System (UVP).

3.3.1.4. Scaled up PCR

A PCR reaction (in duplicate) was prepared by addition of the following reagents to a final concentration of 1 X Kapa Taq Extra HotStart Readymix with dye, 3 mM MgCl₂, 0.5 mg/ml BSA, 0.2 μM Forward primer (F_pto), 0.2 μM Reverse primer (R_LibIII) in a final volume of 100 μl. A negative control (no DNA) was included and setup in a similar manner. All the samples were subjected to the PCR parameters as described in the table below:

Table 3.2: Scaled up PCR parameters.

	Temperature	Time	No of cycles
Initial denaturation	95 °C	3 min	1
Denaturation	95 °C	30 sec	Optimised no of cycles
Annealing	58 °C	30 sec	
Elongation	72 °C	30 sec	
Final elongation	72 °C	5 min	1
Incubation	4 °C	∞	-



3.3.1.5. Generation of ssDNA

The ssDNA was prepared by digesting the antisense strand of the dsDNA prepared in **Section 3.3.1.3**. It was carried out as follows:

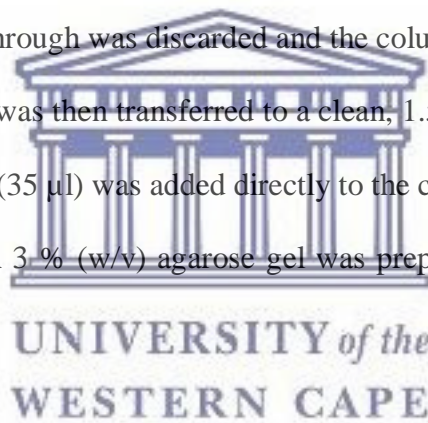
The NEBuffer 4 was added to each PCR amplicon to achieve a final NEBuffer 4 concentration of 1 X. This was followed by the addition of 30U of the T7 Gene 6 exonuclease. The mixture was vortexed briefly and incubated at 25 °C for 60 min. The enzyme was then heat inactivated for 5 min at 65 °C and immediately transferred to ice.

3.3.1.6. Purification of ssDNA

The ssDNA was purified using the Oligo Clean and Concentrator™ kit (Zymo Research), according to the manufacturer's instruction:

The digested ssDNA, prepared in **Section 3.3.1.4**, was transferred to a clean, 2 ml Eppendorf tube; to which twice the volume of the *Oligo Binding Buffer* was added. This was followed by the addition of absolute ethanol (400 µl). The DNA solution was mixed by pipetting a few times and transferred to a *Zymo-Spin™ Column* that was suspended in a *Collection tube*. The DNA solution was centrifuged (10 000 x g; 30 sec; 25°C) and the flow-through was discarded. The *DNA Wash Buffer* (750 µl) was added to the column, followed by centrifugation (10 000 x g; 30 sec; 25°C). The flow-through was discarded and the column was centrifuged (10 000 x g; 60 sec; 25°C). The column was then transferred to a clean, 1.5 ml Eppendorf tube. To elute the DNA, nuclease free water (35 µl) was added directly to the column matrix and centrifuged (10 000 x g; 30 sec; 25°C). A 3% (w/v) agarose gel was prepared, as described in **Section**

3.3.1.3.1.



3.3.1.7. Preparation of ssDNA for SELEX

The remaining ssDNA (25 µl) was prepared by addition of 2 X SELEX buffer (25 µl), followed by brief pipetting. The ssDNA was then denatured at 95 °C for 5 min and gradually cooled as described in **Section 3.3.1.1**.

3.3.1.8. Negative SELEX

After the first round of selection, negative SELEX was introduced to eliminate any sequences that may bind the beads non-specifically. This was carried out by incubating the purified ssDNA, prepared in **Section 3.3.1.6**, with ~0.15 mg of the SiMAG-IDA-Nickel beads, for 30

min at room temperature, whilst shaking at 500rpm. To remove the unbound DNA, the DNA-bound beads (denoted as negative SELEX) were isolated by incubation on a magnet for 1-2 min. The unbound DNA, present in the supernatant, was transferred to a clean, 1.5 ml Eppendorf tube and used as the template for the next round of positive SELEX. The DNA-bound beads were washed, three times, by re-suspension in SELEX buffer (100 μ l), followed by isolation using the magnet. The DNA-bound beads were then re-suspended in 25 μ l of nuclease free water and stored at -20 °C until further use.

3.3.1.9. Counter SELEX

After the seven rounds of selection, counter SELEX was introduced to increase the specificity of the sequences that bind the target protein of interest. This was carried out by incubating the supernatant containing the unbound DNA, prepared in **Section 3.3.16**, with the protein of interest (15 pmol) for 10 min, at 25°C, whilst shaking at 500rpm. After the incubation, the counter proteins (Fc region of IgG and human serum albumin (HSA)) at a final concentration of 15 pmol were added and incubated for a further 10 min, at 25°C, whilst shaking at 500 rpm. At this point, ~0.15 mg of the SIMAG-IDA-Nickel beads were added and incubated, isolated and re-suspended as described in **Section 3.3.1.2**.

3.3.1.10. Evolution monitoring by PCR

Equal amounts of the ssDNA pool from rounds 2, 4, 6, 8 and 10 were 1) directly incubated with naked beads (negative control) and 2) with GFR α 1, followed by incubation with the beads. The bound DNA was then subjected to PCR amplification and resolved on an agarose gel as described in **Section 3.3.1.3.1**. The intensity of each band was calculated using ImageJ software.

Table 3.3: Summary of the SELEX conditions.

	Negative SELEX	Counter SELEX	Positive SELEX
Round 1	-	-	300 pmol LibIII + 30 pmol target; 30 min
Round 2-4	Purified ssDNA + SiMAG-IDA-Ni beads; 30 min; 500 rpm	-	Unbound ssDNA + 30 pmol target; 30 min; 500 rpm
Round 5-7	Purified ssDNA + SiMAG-IDA-Ni beads; 20 min; 500 rpm	-	Unbound ssDNA + 15 pmol target; 20 min; 500 rpm
Round 8-10	Purified ssDNA + SiMAG-IDA-Ni beads; 20 min; 500 rpm	Unbound ssDNA + 15 pmol target; 10 min; 15 pmol IgG + 15 pmol HSA; 10 min	Counter SELEX; 20 min; 500 rpm



UNIVERSITY *of the*
WESTERN CAPE

3.3.2. Identification of candidate aptamers

3.3.2.1. Sample preparation: NGS

The positive SELEX bead complexes for the relevant rounds were amplified, at the optimal number of cycles, using PCR as described in **Section 3.3.1.4**. The PCR amplicons were purified as described in **Section 3.3.1.5**, except that the DNA was eluted in 10 mM Tris-HCl, as recommended by Inqaba Biotec. Each purified amplicon was quantified using the NanoDrop™ 2000 UV-Vis spectrophotometer (Thermo Fisher). The PCR amplicons were then electrophoresed on a 3 % ultra-agarose gel as described as in **Section 3.3.1.3.1**. The samples were sealed in parafilm, stored on ice and sent to Inqaba Biotec for sequencing, using the Illumina MiSeq sequencing platform.

3.3.2.2. Next generation sequencing

Next Generation was performed as described by Inqaba Biotec:

Briefly, PCR products, provided in **Section 3.3.2.1**, were purified (size selected), end repaired and an Illumina specific adapter sequence was ligated to all fragments using the NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (New England BioLabs Inc.).

Following quantification, the samples were individually indexed and a second size selection step was performed, using AMPure XPBeads (Agilent Technologies). Libraries were quality controlled on a DNA chip (Agilent 2100 Bioanalyzer) and then sequenced on Illumina's MiSeq platform, using a MiSeq V3 (600 cycle) kit (Illumina) according to the manufacturer's protocol.

50 Mb of data (2x300bp long paired end reads) were produced for each sample and adapter sequences were trimmed off the raw data.

3.3.2.3. Sequence analysis using Geneious

To identify the full-length candidate aptamer sequences, the sequencing data (**Section 3.3.2.2**) were subjected to analysis using the Geneious software (“...an easy-to-use and flexible desktop software application framework for the organization and analysis of biological data...”) (Kearse et al. 2012) as follows for each SELEX round that was submitted:

3.3.2.3.1. Sequence merge and mapping

To create an overlapped single 5’ to 3’ sequence read, the forward (F1) and reverse (R1) reads that were uploaded onto Geneious (Version 10.0.4) were grouped together as a paired end sequencing set. This was followed by merging the grouped paired end sequences using BBMerge (Version 36.32) at a normal merge rate. Only reads at the expected sequence length (97mer) were extracted into a new file for further processing. A reference sequence was then created to produce an alignment of the sequences to ensure that the candidate aptamer sequences were identified in the correct orientation. In this case, the forward primer (**Section 3.2.3**) was used as a reference sequence. After mapping the sequences to the reference, only those sequences containing both the forward and reverse sequence sub-sections were extracted into a new file for further processing.

3.3.2.3.2. Identification of unique sequences

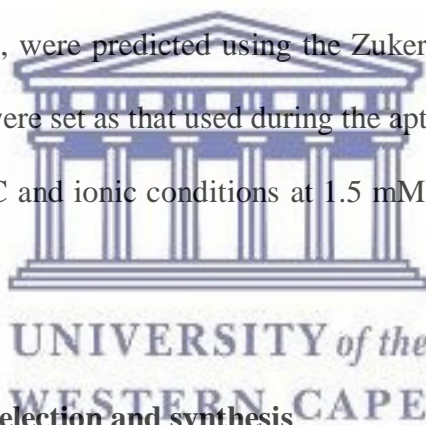
All duplicates from the final sequence set derived in **Section 3.3.2.3.1**, with identical residues, were extracted into a new file as unique sequences. These would represent all candidate aptamer sequences that were potentially enriched against the protein target. Once this was achieved, any unique sequences that occurred more than 1000 X – within the set of unique sequences – were selected for to screen as potential candidate aptamers.

3.3.2.3.3. Aptamer similarity analysis using sequencing alignment

To identify any potential consensus motifs, the fixed primer regions were trimmed from the final set of unique candidate aptamer sequences generated in **Section 3.3.2.3.2**. The trimmed sequences were extracted into a new file as the variable regions of the candidate aptamer sequences. At this point, the variable regions were subjected to a multiple sequence alignment using ClustalW, with default parameters (cost matrix: IUB, gap opening cost: 15 and gap extend cost: 6.66).

3.3.3.4. Secondary structure prediction

Theoretical secondary structures of the final set of unique candidate aptamer sequences, generated in **Section 3.3.2.3.2**, were predicted using the Zuker algorithm in Mfold (Version 3.4). The folding parameters were set as that used during the aptamer selection process, which is, folding temperature at 25°C and ionic conditions at 1.5 mM and 138 mM, for [Mg⁺⁺] and [Na⁺], respectively.



3.3.3.5. Candidate aptamer selection and synthesis

After the sequence analysis, a total of six candidate aptamers were selected for characterisation, three from each criterion, that is, 1) sequence abundance and 2) lowest minimal free energy. The six candidate aptamers were synthesised by Biomers (Germany).

3.4. Results and discussion

3.4.1. *In vitro* aptamer selection against the soluble ULBP2 ecto-domain

The first objective of this chapter was to select for aptamers against the soluble ecto-domain of the ULBP2 recombinant protein (ULBP2). In order to accomplish this, a stringent *in vitro* selection process (SELEX) was employed, as described in **Section 3.3.1**. A SELEX round essentially consists of three steps, namely, 1) adsorption: selecting ligand sequences that binds to the target, 2) separation: the partitioning of target bound ligand sequences from the unbound sequences and 3) enrichment: PCR amplification of those ligand sequences that have bound the target (Hamula et al. 2006; Gopinath, 2007). The PCR amplicon is then subjected to an enzyme treatment to produce ssDNA, which is then used as the template for the next round of SELEX. This is repeated until affinity saturation is reached and at this point, those ligand sequences that have high affinity and specificity towards the target would have been isolated (Mairal et al. 2008; Wang, 2008; Lakhin et al. 2013). This typical workflow is demonstrated by agarose gel images (**Figure 3.1**), as a representation for all the rounds of SELEX performed against the proteins of interest (ULBP2 and GFR α 1), in this study.

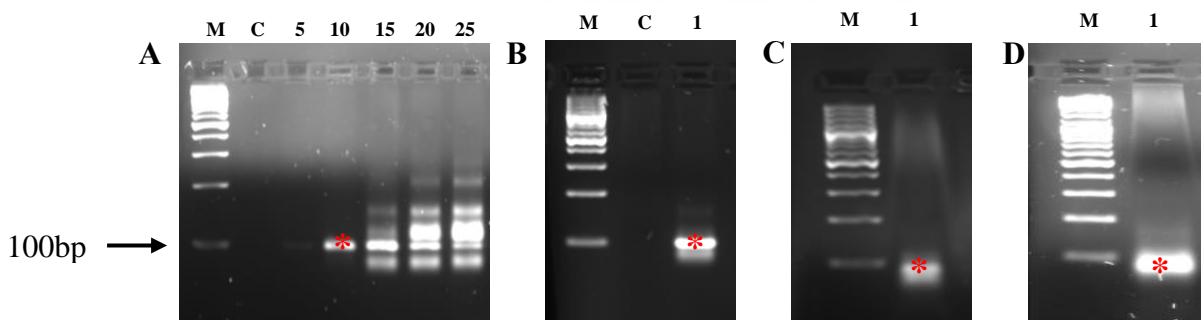


Figure 3.1: Typical workflow for a round of SELEX. After incubation of the ssDNA library with the GFR α 1, the bound sequences were subjected to **A)** Pilot PCR (Lane M: Marker, Lane C: Control, Lanes 5, 10, 15, 20 and 25 represent the no of PCR cycles); **B)** Scaled up PCR (Lane M: Marker, Lane C: Control, Lane 1: PCR product), **C)** T7 enzyme digestion (Lane M: Marker, Lane 1: Digested PCR product) and **D)** ssDNA purification (Lane M: Marker, Lane 1: Purified ssDNA). The regions of interest are indicated by the red asterisk.

The SELEX procedure for the selection of DNA aptamers against ULBP2 is summarized into seven rounds, as shown in **Table 3.3**, beginning with a highly diverse ssDNA library. The library consisted of a randomized region, flanked by fixed primer binding regions on both termini. In the event that an effective selection process occurred, it is expected that the diversity of the library would, significantly, reduce with each successive round of SELEX, resulting in the enrichment of the strongest target binding sequences from the initial library. This would be observed as an increase in the size and intensity of the PCR amplicon band for the positive SELEX (incubation of target with the SELEX library and SiMAG-IDA-Nickel beads) as compared to the negative SELEX (incubation of the SELEX library with the SiMAG-IDA-Nickel beads) (Mencin et al. 2014). A pilot PCR was used to determine the optimal number of PCR cycles required to amplify the bound sequences, for each round of SELEX. The optimum amplification herein was considered as the PCR amplicon band produced from the least amount of PCR cycles performed and was free from by-products, indicated by the red astericks. When evaluating this phenomenon, it was found that almost no changes in the PCR amplicon size and intensity for the positive SELEX, from rounds 1 through 7, could be observed (**Figure 3.2**). On the contrary, the negative SELEX demonstrated noticeable changes in the PCR amplicon size and intensity, from rounds 1 through 7. This can occur due to sequences that, non-specifically, bind to the partitioning matrix, in this case the SiMAG-IDA-Nickel beads. Since negative SELEX is a modification to the classic SELEX, intended to eliminate those sequences with low specificity to the target, the observation (**Figure 3.2**) that very little changes in the level of amplification occurred suggests that the sequences in the SELEX library, potentially, have low affinity and specificity to the target (ULBP2) (Djordjevic, 2007; Ray and White, 2017). Alternatively, it is possible that the interaction and/or the affinity of the SELEX library to the SiMAG-IDA-Nickel beads were much stronger, relative to that of the ULBP2 protein.

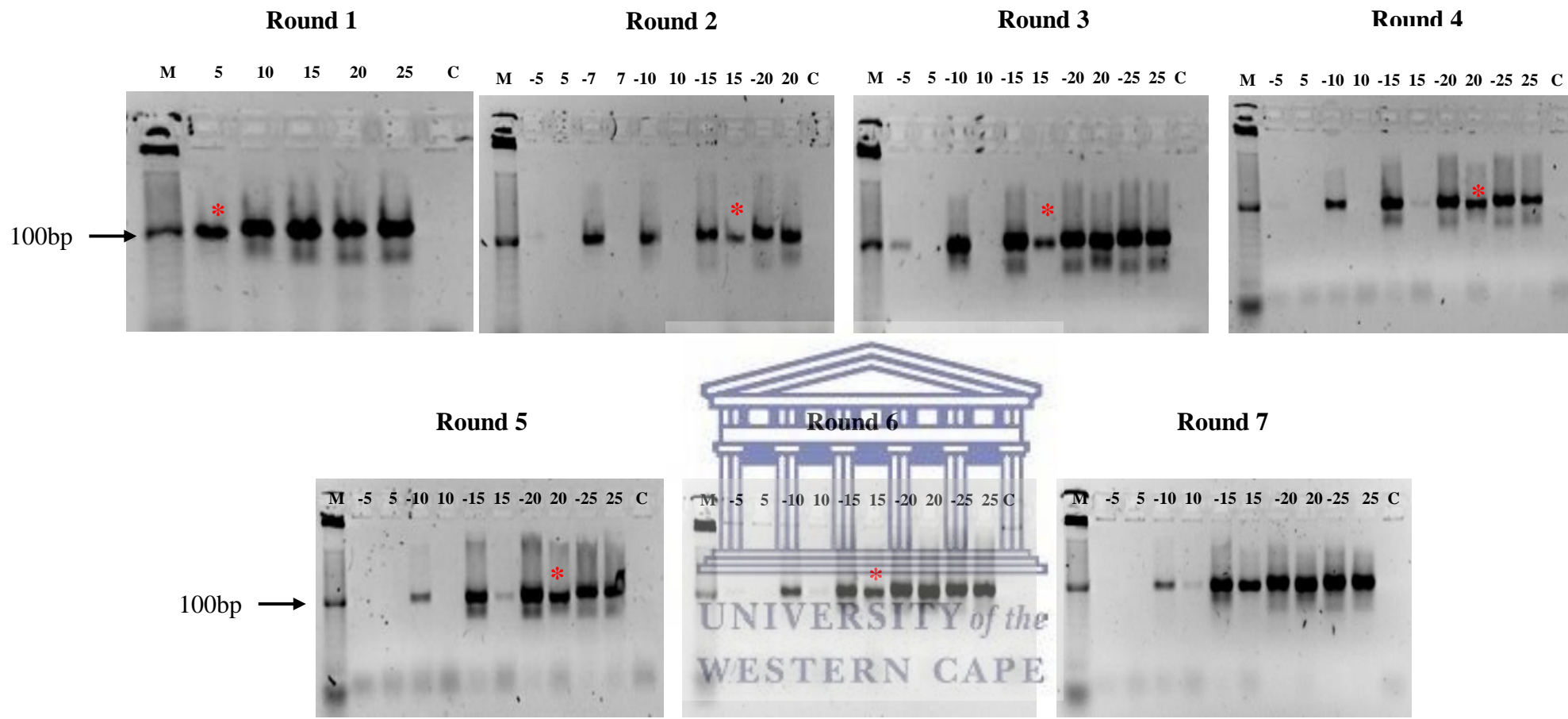
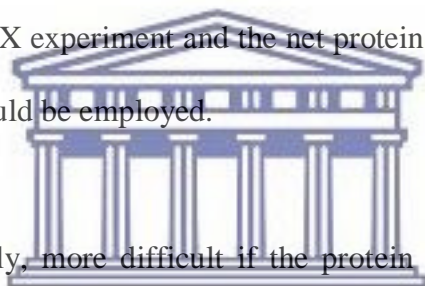


Figure 3.2. SELEX against ULBP2 protein. Each agarose gel image represents a round of SELEX. Lane M: 10bp DNA ladder. Lanes -5, -7, -10, -15, -20, -25: No of cycles used for the Negative SELEX. Lanes: 5, 7, 10, 15, 20 and 25: No of cycles used for the Positive SELEX. Lane C: Negative control (no DNA). The optimal no of cycles used to proceed to the next round of SELEX is indicated by the red asterisks.

A study carried out by Kowalska and colleagues investigated the impact of various metal ion and resins on DNA aptamer selection; and it was found that it is important to screen multiple metal ions/resins, to ensure that the most suitable metal ions/resins are used because the SELEX library could, potentially, contain a large fraction of sequences that are ion-specific (Kowalska et al. 2014). Thus, in the case of the SiMAG-IDA-Nickel beads used in this study, the affinity of the nickel-binding sequences may have exceeded the affinity of the ULBP2 binding sequences. Furthermore, another study by Ahmad and colleagues demonstrated that the overall net charge of the protein may have an effect on the selection affinity (Ahmad et al. 2011). They found that there was an inverse correlation of between the net protein charge and observed aptamer K_D values (Ahmad et al. 2011). Therefore, either different metal ions/resins should be investigated prior to the SELEX experiment and the net protein charge must be considered or a different SELEX library should be employed.



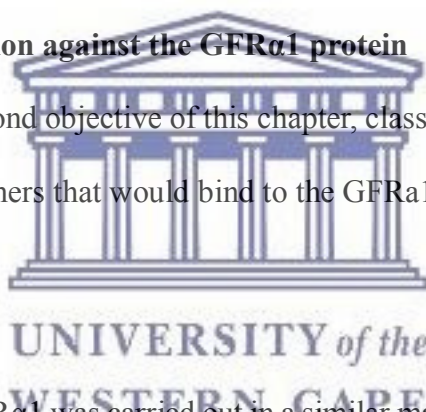
Aptamer selection is, generally, more difficult if the protein target possesses hydrophobic regions because oligonucleotides are hydrophilic, thus resulting in limited interactions (Stoltenburg et al. 2007; Hasegawa et al. 2016). The ULBP2 protein, used for this study, was provided as the soluble domain; however, it includes amino acids at its C-terminus, usually required for GPI anchorage (Fernández-Messina et al. 2011) and this region specifically, is hydrophobic in nature (Müller et al. 2010). Taking this into consideration, it may possibly be the reason for the poor selection and evolution observed against the ULBP2 protein (**Figure 3.2**). The SELEX library consists of hydrophilic oligonucleotides, thereby limiting hydrophobic interactions between the SELEX library and protein targets (Hasegawa et al. 2016). Therefore, an approach to overcome this limitation is to incorporate a modification of the SELEX library using hydrophobic moieties, consequently, changing the nature of the binding properties exhibited by the SELEX library (Hasegawa et al. 2016). This, in turn, may

translate into an improved aptamer selection against the ULBP2 protein. Alternatively, an investigation into the removal of the hydrophobic residues on ULBP2, prior to SELEX, can be performed.

A limitation recognized in this study is that the SELEX library could potentially have an affinity towards the 6X His tag; therefore, an additional control would be to employ a purified 6X His tag as a protein during the counter selection. This would favour the selection towards the protein of interest (ULBP2) and furthermore, facilitate the elimination of sequences that may bind to the 6X His tag.

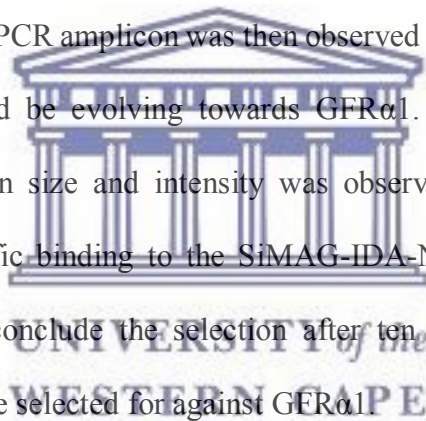
3.4.2. *In vitro* aptamer selection against the GFR α 1 protein

In order to accomplish the second objective of this chapter, classical SELEX was performed to select for potential DNA aptamers that would bind to the GFR α 1 protein. This was carried out as follows:



The SELEX procedure for GFR α 1 was carried out in a similar manner as described for ULBP2, but with a few amendments, these include: 1) the gradual decrease in target protein concentration and incubation times with increasing rounds of SELEX (**Table 3.3**) and 2) the incorporation of counter SELEX, in the last 3 rounds. The full protocol is described in **Section 3.3.1**. By incorporating the above-mentioned steps, favourable conditions for enrichment of the strongest target binding sequences were created; because the affinity and specificity of the aptamers, to the target, is directly affected by the stringency of the conditions used during the selection (Stoltenburg et al. 2007; Thiviyathan and Gorenstein 2012).

Figure 3.3 shows the 10 rounds of SELEX performed for GFR α 1, of which an indication of a successful SELEX library evolution was observed by an increase in the size and intensity of the PCR amplicons for the positive SELEX, as the rounds of SELEX were progressing. Round 4 displayed a slight decrease in intensity of the PCR amplicon, which was expected because by round 4, negative SELEX had already been iterated three times. This means that several sequences that bound, non-specifically, to the SiMAG-IDA-Nickel beads were eliminated from the SELEX library and less sequences were amplified. However, the intensity of the PCR amplicon band increased again from rounds 5 through 7. Since counter SELEX was introduced from round 8, a decrease in the intensity of the PCR amplicon band occurred again, accounting for the elimination of sequences that bound to the counter proteins (IgG and HSA). A steady increase in the intensity of the PCR amplicon was then observed in rounds 9 and 10, suggesting that the SELEX library could be evolving towards GFR α 1. And furthermore, noticeable changes in the PCR amplicon size and intensity was observed for the negative SELEX, suggesting that no non-specific binding to the SiMAG-IDA-Nickel beads occurred. These results were satisfactory to conclude the selection after ten rounds and hypothesise that successful DNA aptamers were selected for against GFR α 1.



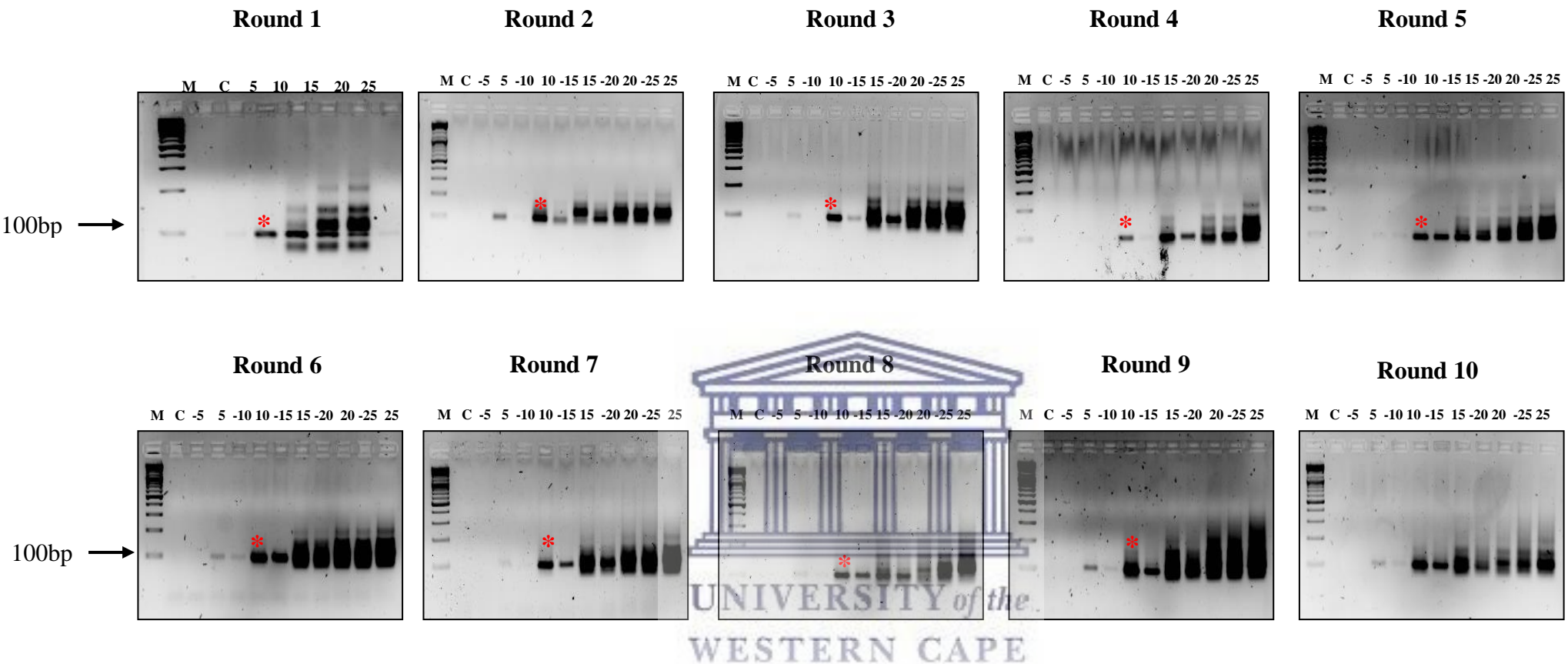


Figure 3.3. SELEX against the GFR α 1 protein. Each agarose gel image represents a round of SELEX. Lane M: 100bp DNA ladder. Lanes -5, -7, -10, -15, -20, -25: No of cycles used for the Negative SELEX. Lanes: 5, 7, 10, 15, 20 and 25: No of cycles used for the Positive SELEX. Lane C: Negative control (no DNA). The optimal no of cycles used to proceed to the next round of SELEX is indicated by the red asterisks.

Since equal concentrations of the starting SELEX library pool, for each selection round, was not used in the pilot PCR, the evolution was confirmed by assessing the negative and positive SELEX of 5 rounds (round 2, 4, 6, 8 and 10), using PCR as described in **Section 3.3.1.10** and then further analyzed using ImageJ software (Schindelin et al. 2012). Almost no interaction was observed between the respective SELEX library pools (round 2, 4, 6, 8 and 10) and the naked SiMAG-IDA-Nickel beads (negative SELEX); however, a steady increase in the amplification of sequences bound to the GFR α 1 protein (positive SELEX) was observed (**Figure 3.4A**), corroborating the evidence of a successful aptamer selection against the GFR α 1 protein as observed during the pilot PCR (**Figure 3.3**). Only round two showed some discrepancy; however, this could be explained by the fact that the introduction of negative SELEX may have eliminated a significant amount of low affinity sequences that initially bound the GFR α 1 protein. Since the tenth SELEX library pool showed evidence of the highest enrichment of sequences that bound to the GFR α 1, it was selected for sequencing.

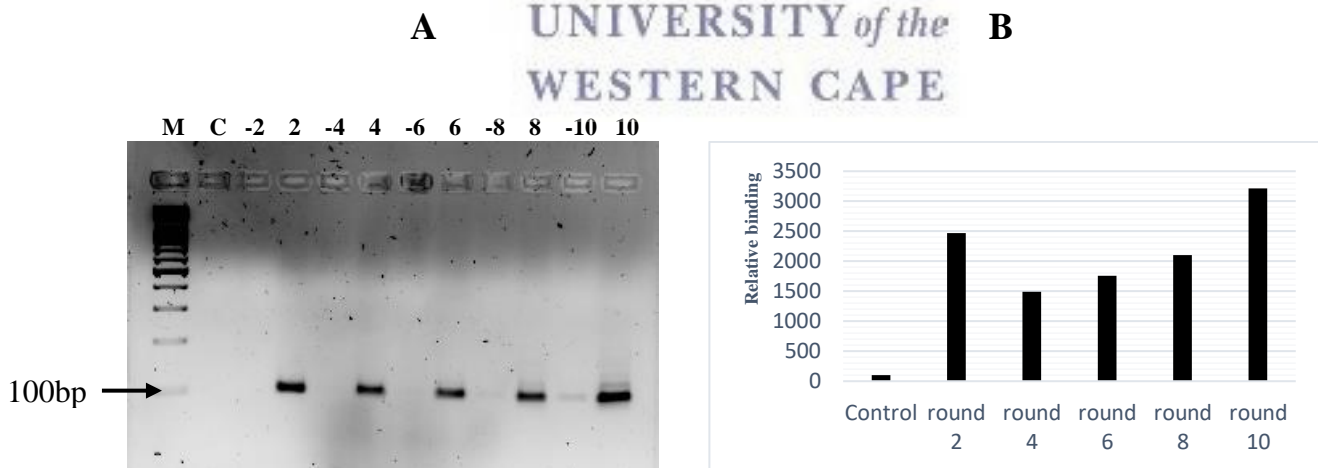


Figure 3.4. Monitoring the evolution of the *in vitro* selection against GFR α 1. (A). Lane M:100bp DNA ladder. Lane C: Negative control (no DNA). Lanes -2, -4, -6, -8 and -10: Negative SELEX DNA library pool for the respective rounds of SELEX. Lanes 2, 4, 6, 8 and 10: Positive SELEX DNA library pool for the respective rounds of SELEX. **(B).** Image J analysis of the results shown in image A.

3.4.3. Identification of candidate aptamers using Next Generation Sequencing analysis.

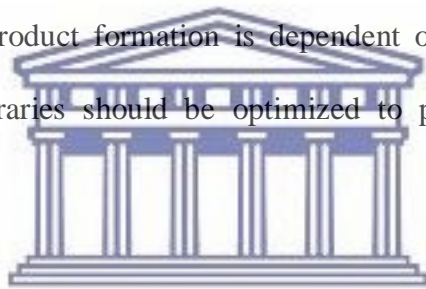
Using the Next Generation Sequencing (NGS) platform, the tenth SELEX library pool was sequenced and analyzed as described in **Section 3.3.2**. The selection of the final aptamer sequences is summarized in **Table 3.4**. The starting SELEX library, used in this study, contained sequences of 97 nucleotides (NT) in length, which constitutes the variable region (53NT) and the fixed primer regions at the 5' and 3' ends, respectively. It was therefore expected that the sequencing data would reflect 97mer long sequences, however, when looking at the initial total sequences acquired from the NGS, a varying sequence length distribution was observed. These included sequences shorter than 97NT, sequences of 97NT in length and sequences longer than 97NT. The shorter sequences were made up of either forward primer regions only or forward primer regions with fluctuating lengths of the variable region. Whereas the longer sequences were made of multiple repeats of the primer regions. Since SELEX involves repetitive PCR amplification steps the varying sequence length distribution could be attributed to PCR by-product formation over time (Musheev et al. 2006).



Table 3.4: Next Generation Sequencing analysis of the tenth selection pool against GFR α 1.

	Total no. of sequences
Total sequences	912837
Total sequences after merging the paired end sequences	707652
Extraction of 97mer length sequences	586565
Total overrepresented sequences	256634
Total overrepresented sequences with copy number >1000	16

Contrary to conventional PCR, which involves the PCR amplification of a homogenous DNA template - where product formation ceases when the primers are depleted – Musheev et al. 2006, for the first time, demonstrated that in the case of PCR amplification of a random library as the template, product formation ceases while the primers are still in excess of the products (Musheev et al. 2006). Consequently, at a very high rate, the products are converted to by-products (product-product hybridization) and the starting library template can, effectively, disappear after another 5 PCR cycles (Musheev et al. 2006). Furthermore, it was elucidated that the by-product formation was in no way affected by the initial number of DNA molecules in the PCR mixture; however, an increase in the concentration of Taq DNA polymerase, in the PCR mixture, can selectively favor the yield of PCR products (Musheev et al. 2006). And finally, the product and by-product formation is dependent on library and primer design. Therefore, SELEX DNA libraries should be optimized to perform an efficient aptamer selection.



In a study by Tolle et al. 2014, aptamer selection was performed against the Activated Protein C (APC) and it was demonstrated that PCR amplification reactions can yield two types of artificial PCR by-product formation during SELEX experiments (Tolle et al. 2014). In the first instance, the sequencing data showed evidence of the forward primer linked by random stretches of the variable region; and in the second instance, a fusion of multiple reverse primer regions was observed (Tolle et al. 2014). They explain the mechanism of both types of PCR by-product formation as a result of the 3' end of a random sequence that anneals to a complimentary section in the variable region, on another strand. This strand then acts as a primer and subsequent elongation of the complimentary strand occurs (Tolle et al. 2014). Consequently, PCR amplification of this elongated strand in the subsequent PCR cycle may then lead to by-products consisting of double-stranded DNA with doubled reverse primer

binding sites (Tolle et al. 2014). Over time, these by-products dominate the library composition, becoming the new template which then results in the formation of strands with multiple copies of the reverse primer binding sites (Tolle et al. 2014). It is therefore crucial to carefully consider the primer design, but also be mindful that some library/target pairs are problematic and therefore, optimization of the library design and PCR method should be performed.

Since the use of complex random libraries are necessary for SELEX and the problem of by-product formation can significantly hinder the generation of aptamer selection, by essentially resulting in the loss of ~50% of the library diversity with each PCR cycle; Witt et al. 2017 endorsed the use of emulsion PCR (ePCR) as an alternative to overcome the limitations presented by conventional PCR and real-time PCR methods during aptamer selection (Witt et al. 2017). The advantage of ePCR is that this method relies on the formation of individual droplets containing a single species of the oligonucleotide template. From this study, Witt et al. 2017 established that the starting concentration of the template is crucial to prevent the formation of by-products. When the template concentration is too high, multiple templates are entrapped per droplet, consequently, by-products are formed (Witt et al. 2017). Together these three studies explain the reason for the varying sequence length distribution obtained from the sequencing data and provided ways to circumvent the loss of the library composition, which ultimately lowers the probability of acquiring the best high affinity candidate aptamers.

Although the above discussed limitation existed in this study, as shown in **Table 3.4**, ~65% (586565) of the 97mer sequences were able to be retrieved, using the Geneious software. The overrepresented sequences were then extracted and filtered down to the top 16 sequences that occurred in the data more than 1000 times. This methodology is supported by conventional

candidate aptamer identification which involves analyzing the large NGS dataset in a way that counts the sequences, followed by arranging them in order of frequency (Blind and Blank, 2015). Because SELEX experiments involves the use of repetitive PCR amplification, there is a possibility that the shift of the evolution pressure can move from high affinity sequences to sequences that are simply amplifying well, therefore, the selection of candidate aptamers based only on frequency isn't plausible (Blind and Blank, 2015). Another strategy to select candidate aptamers is by identifying frequent sub-sequences, referred to as motifs, that occur within the variable regions of the aptamer sequence (Blind and Blank, 2015). Taking this into consideration, the fixed primer binding regions were removed from the top 16 overrepresented sequences in order to obtain the variable region only; which was then subjected to a multiple alignment to search for any conserved motifs (Figure 3.4). To which end no conserved motifs were observed and as a result, the top 3 overrepresented sequences (G_Apta1, G_Apta2 and G_Apta3) were selected as candidate aptamers (Table 3.5).

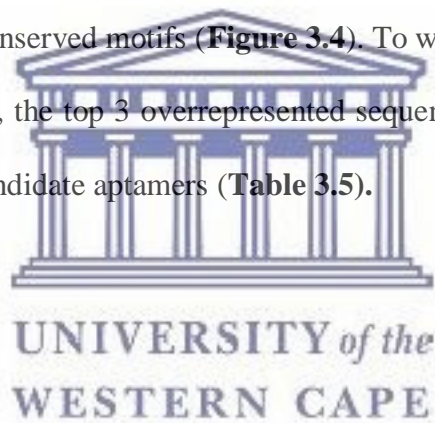




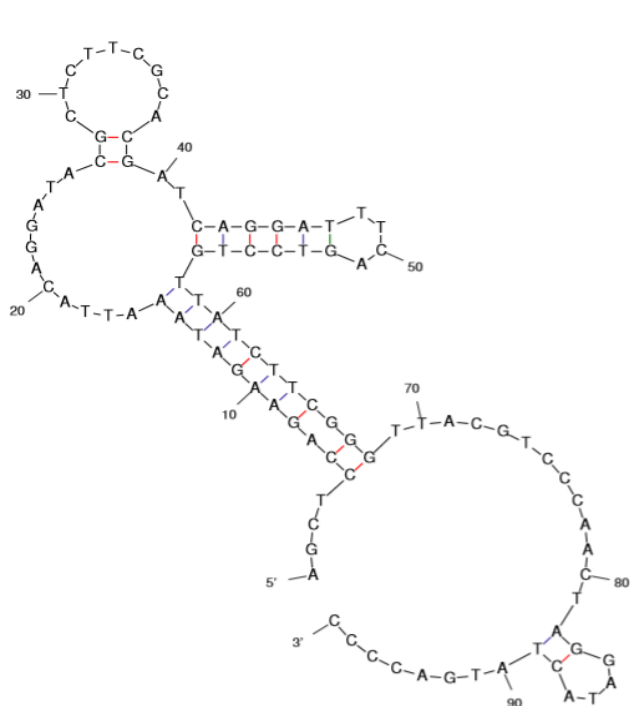
Figure 3.5. Multiple sequence alignment of top 16 overrepresented candidate aptamers against GFR α 1. The sixteen candidate aptamers, filtered from NGS data, were screened for conserved motifs using Clustal W.



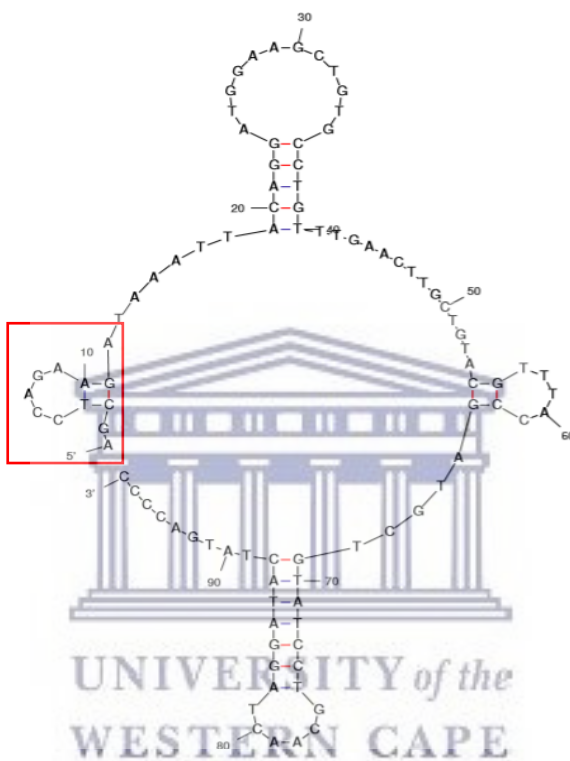
3.4.4. Secondary structure prediction

As observed with the multiple alignment (**Section 3.4.3**), no obvious conserved motif could be identified and as such, the clustering of sequences by conserved or similar structural motifs instead could also be investigated to provide information into the potential evolutionary lineage of the aptamer sequences (Ditzler et al. 2012; Blind and Blank, 2015). This can be achieved using tools based on secondary structure predictions, such as Mfold (Zuker, 2003; Le et al. 2014). The initial full length aptamer, generally, consists of three structural domains, that is, 1) a domain that constitutes the indispensable nucleotides required for the aptamer-target interaction; 2) another domain that acts as a supportive structure for the binding activity and 3) a final domain that performs neither of the two functions and makes up what is referred to as the non-essential nucleotides (Xie et al. 2013). The third domain can be removed by means of follow up truncation studies (Xie et al. 2013).

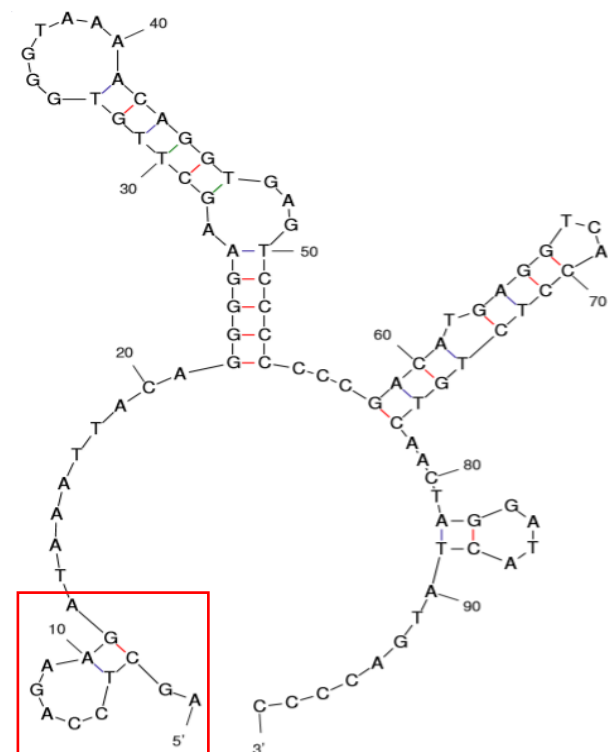
In this study, the respective domains have not yet been elucidated, therefore it was important to analyze the top 16 overrepresented full length aptamers candidates (**Section 3.4.3**). This was carried out as described in **Section 3.3.3.4**. No obvious conserved structural motifs were identified, except for the stem loop fold observed at the 5' region (**Figure 3.5**), indicated in the red blocks, for G_Apta13 and G_Apta15. Since it is expected that the stability of the secondary structure contributes to the overall interaction of the candidate aptamers with the GFR α 1 protein, three predicted aptamer structures (G_Apta9, G_Apta13 and G_Apta15) with the lowest minimal free energy values (**Figure 3.5**) were selected as candidate aptamers for further characterization. The sequences of the candidate aptamers are enlisted in **Table 3.5**.



G-Apta9: $\delta G = -10.72$



G-Apta13: $\delta G = -12.23$



G-Apta15: $\delta G = -11.26$

Figure 3.5. Secondary structure prediction of three candidate aptamers with the lowest free minimal energy. The candidate aptamers were subjected to secondary structure predictions using Mfold, under the same conditions used during SELEX, to screen for structurally conserved motifs. The minimal free energy values were extracted for each folded aptamer. Similarities in structural configurations are highlighted in the red boxes.

Table 3.5. Sequences of final candidate aptamer list.

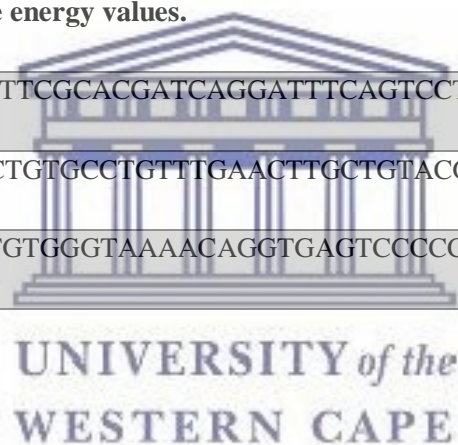
A. Top three overrepresented sequences selected as candidate aptamers.

G_Apta1	AGCTCCAGAAGATAAAATTACAGGGAACGTGTTGGTTGCGGTTCTTCCGATCTGCTGTGTTCTCTATCTGTGCCATGCAACTAGGATACTATGACCCC
G_Apta2	AGCTCCAGAAGATAAAATTACAGGGTCCAGCGGTCTCACTTGTGGTAATACATCACTTCAGACGGGTTTCATTGCACAAGTAGGATACTATGACCCC
G_Apta3	AGCTCCAGAAGATAAAATTACAGGAATGAAGCTTTGAGCCTGTTTGAAGTCATGTAAAGTATCGGTTTCTGCCATACAAGTAGGATACTATGACCCC

B. Top three candidate aptamers with lowest minimal free energy values.

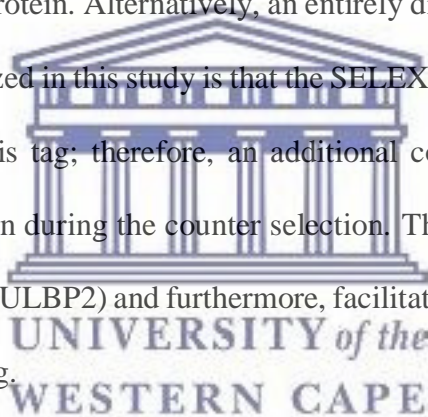
G_Apta9	AGCTCCAGAAGATAAAATTACAGGATACGCTCTTCGCACGATCAGGATTTTCAGTCCTGTTATCTTCGGGTACGTCCCAACTAGGATACTATGACCCC
G_Apta13	AGCTCCAGAAGATAAAATTACAGGATGGAAGCTGTGCCTGTTTGAAGTTGCTGTACGTTTACCGATGCTGTATCCTGCAACTAGGATACTATGACCCC
G_Apta15	AGCTCCAGAAGATAAAATTACAGGGGAAGCTTGTGGGTAAAACAGGTGAGTCCCCCGACATGAGGTCACCTCTGTCAACTAGGATACTATGACCCC

*Variable regions are indicated in red.



3.5. Conclusion and future recommendations

The aptamer selection process, used in this study, against the commercially available ULBP2 protein was unsuccessful as no indication of evolution was observed, after seven rounds of selection. What was observed is that the affinity of the SELEX library pool towards the SiMAG-IDA-Nickel beads exceeded the affinity of the SELEX library pool towards the ULBP2 protein; which could be translated as significantly high, non-specific interactions (Djordjevic, 2007; Ray and White, 2017). This may be due to the sequences in the SELEX library pool being ion-specific in nature (Kowalska et al. 2014). Therefore, the first point of action, prior to SELEX, would be to screen and optimize the metal ion/resin used in the aptamer selection against the ULBP2 protein. Alternatively, an entirely different SELEX library can be explored. A limitation recognized in this study is that the SELEX library could potentially have an affinity towards the 6X His tag; therefore, an additional control would be to employ a purified 6X His tag as a protein during the counter selection. This would favour the selection towards the protein of interest (ULBP2) and furthermore, facilitate the elimination of sequences that may bind to the 6X His tag.



In the case of the aptamer selection against the GFR α 1 protein, a successful evolution was observed. However, as observed with NGS analysis, approximately 35% of the sequences were invalid due to incorrect sequence length. This may be as a result of by-product formation that occurs during PCR; which can be overcome by careful design of the library and primers or alternatively the employment of ePCR (Musheev et al. 2006; Witt et al. 2017). Even though some sequences were lost due to potential by-product formation, from the remaining ~65% of sequence reads, 6 potential candidate aptamers were identified using NGS analysis and secondary structure prediction. These 6 candidate aptamers will be characterized in Chapter 4.

3.6. References

- Blind, M. and Blank, M., 2015. Aptamer Selection Technology and Recent Advances. *Molecular Therapy—Nucleic Acids*, 4(1), p.e223.
- Boulay, A. et al. 2008. The ret receptor tyrosine kinase pathway functionally interacts with the ER α pathway in breast cancer. *Cancer Research*, 68(10), pp.3743–3751.
- Ditzler, M.A. et al. 2013. High-throughput sequence analysis reveals structural diversity and improved potency among RNA inhibitors of HIV reverse transcriptase. *Nucleic Acids Research*, 41(3), pp.1873–1884.
- Drah, M., 2015. *The Development of Nanotechnology- based Detection Systems for the Diagnosis of Breast Cancer*. University of the Western Cape.
- Duffy et al. 2017. Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumour Markers (EGTM). *European Journal of Cancer*, 75(2017), pp.284-298.
- Esseghir, S. et al. 2007. A role for glial cell derived neurotrophic factor induced expression by inflammatory cytokines and RET/GFR alpha 1 receptor up-regulation in breast cancer. *Cancer Research*, 67(24), pp.11732–11741.
- Fleming, M.S. et al. 2015. Cis and trans RET signaling control the survival and central projection growth of rapidly adapting mechanoreceptors. *eLife*, 2015(4), pp.1–26.
- Gatelli et al. 2013. Ret inhibition decreases growth and metastatic potential of estrogen receptor positive breast cancer cells. *EMBO Molecular Medicine*, 5(9), pp. 1335-1350.
- Gopinath, S.C.B., 2007. Methods developed for SELEX. *Analytical and Bioanalytical Chemistry*, 387(1), pp.171–182.
- Hamula, C.L.A. et al. 2006. Selection and analytical applications of aptamers. *TrAC - Trends in Analytical Chemistry*, 25(7), pp.681–691.
- Groot, J.W.B. De et al. 2016. RET as a Diagnostic and Therapeutic Target in Sporadic

and hereditary endocrine tumors. *Endocrine Reviews*, 27(5), pp.535-560.

- Hasegawa, H. et al. 2016. Methods for Improving Aptamer Binding Affinity. *Molecules*, 21(421), pp. 1-15.
- He, A. et al. 2013. A novel immunoassay for the quantization of CYFRA 21-1 in human serum. *Journal of Clinical Laboratory Analysis*, 27(4), pp.277–283.
- Hereditary Endocrine Tumours. , 27(June), pp.535–560.
- Jauset Rubio, M. et al. 2016. β -Conglutin dual aptamers binding distinct aptatopes. *Analytical and Bioanalytical Chemistry*, 408(3), pp.875–884.
- Kazarian, A. et al. 2017. Testing breast cancer serum biomarkers for early detection and prognosis in pre-diagnosis samples. *British Journal of Cancer*, 116, pp.501-508.
- Kearse, M. et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organisation and analysis of sequence data. *Bioinformatics*, 28(12), 1647 – 1649.
- Khati, M., 2010. The Future of Aptamers in Medicine. *Journal of clinical pathology*, 63(6), pp.480–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20360137>.
- Kowalska, E., Bartnicki, F., Pels, K. and Strzalka, W. 2014. The impact of immobilized metal affinity chromatography IMAC resins on DNA aptamer selection. *Analytical Bioanalysis Chemistry*, 406(2014), pp. 5495-5499.
- Lakhin, A. V, Tarantul, V.Z. and Gening, L. V, 2013. Aptamers : Problems , Solutions and Prospects. *Acta Naturae* , 5(19), pp.34–43.
- Liu, J., Mazumdar, D. and Lu, Y., 2006. A simple and sensitive “dipstick” test in serum based on lateral flow separation of aptamer-linked nanostructures. *Angewandte Chemie - International Edition*, 45(47), pp.7955–7959.
- Mairal, T. et al. 2008. Aptamers: Molecular tools for analytical applications. *Analytical and Bioanalytical Chemistry*, 390(4), pp.989–1007.

- Mencin, N. et al. 2014. Optimization of SELEX: comparison of different methods for monitoring the progress of in vitro selection of aptamers. *Journal of Pharmaceutical and Biomedical analysis*, 91, pp.151-159.
- Morandi, A., Plaza-Menacho, I. and Isacke, C.M., 2011. RET in breast cancer: Functional and therapeutic implications. *Trends in Molecular Medicine*, 17(3), pp.149–157.
- Morandi, A. and Isacke, C.M. 2014. Targeting RET-interleukin-6 crosstalk to impair metastatic dissemination in breast cancer. *Breast Cancer Research*, 16(2014), pp. 1-3.
- Mirabelli, P. and Incoronato, M. Usefulness of traditional serum biomarkers for management of breast cancer patients. *Biomedical Research International*, 2013, pp. 1-9.
- Musheev, M.U and Krylov, S.N. 2006. Selection of aptamers by systematic evolution of ligands by exponential enrichment: addressing the polymerase chain reaction issue. *Analytica Chimica Acta*, 564, pp.91–96.
- Plaza-Menacho, I. et al. 2010. Targeting the receptor tyrosine kinase RET sensitizes breast cancer cells to tamoxifen treatment and reveals a role for RET in endocrine resistance. *Oncogene*, 29(33), pp.4648–4657. Available at: <http://dx.doi.org/10.1038/onc.2010.209>.
- Schindelin, J. et al. 2012, "Fiji: an open-source platform for biological-image analysis", *Nature methods* 9(7), pp. 676-682
- Stoltenburg, R., Schubert, T. and Strehlitz, B., 2015. In vitro Selection and Interaction Studies of a DNA Aptamer Targeting Protein A. *PLoS ONE*, 10(7), pp.1–23.
- Svobodová, M., Pinto, A. and O' Sullivan, C.K. 2012. Comparison of different methods for generation of single-stranded DNA for SELEX processes. *Analytical and Bioanalytical Chemistry*, 404(3), pp.835 – 842.



UNIVERSITY of the
WESTERN CAPE

- Thiviyanathan, V. and Gorenstein, D.G., 2012. Aptamers and the next generation of diagnostic reagents. *Proteomics - Clinical Applications*, 6(11–12), pp.563–573.
- Tolle, F., Wilke, J., Wengel, J. and Mayer, G. 2014. By-product formation in repetitive PCR amplification of DNA libraries during SELEX. *PLOS One*, 9(12), pp. 1-12.
- Wang, D.L., Xiao, C., Fu, G., Wang, X. and Li, L. 2017. Identification of potential serum biomarkers for breast cancer using a functional proteomics technology. *Biomarker Research*, 5(11), pp.1-10.
- Witt, M. et al. 2017. Comparing two conventional methods of emulsion PCR and optimizing of Tegosoft-based emulsion PCR. *Engineering in Life Science*, 17(2017), pp. 953-958.
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, 31(13), pp.3406 – 3415.



Chapter Four: Characterisation of the binding affinity and specificity of the candidate aptamers against GFR α 1.

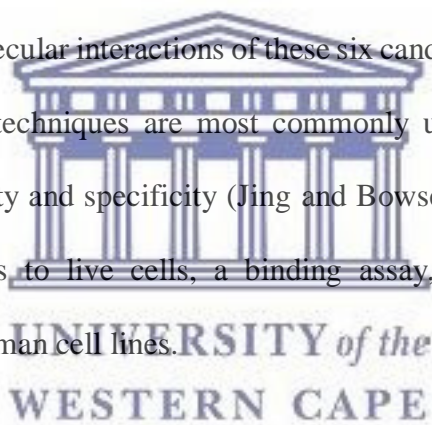
4.1. Introduction

In view of the potential application of aptamer technologies in clinical settings, it is imperative to understand the aptamer-target interaction (Jing and Bowser 2011). In this study, aptamer-protein interactions, in particular, were investigated and this is important because when these interactions occur, it results in the subsequent activation or regulation of biochemical pathways that can determine the fate of a cell, which in turn, can contribute towards disease management or eradication (Ritzefeld and Sewald 2011). Furthermore, understanding the aptamer-protein interaction also gives insight into the binding mechanisms of the aptamers and in doing so, information such as the binding ability, the binding affinity, binding kinetics upon interaction with the target, dissociation constants and identification of the binding region on the aptamer/target can be investigated and used to assess and/or improve the aptamer design and application (Jing and Bowser 2011). This can be accomplished by using an array of techniques applicable to the aptamer-protein interaction, of which EMSA, SPR, MST and flow cytometry will be discussed.

In Chapter 1, a detailed explanation of the principles for each of the characterization techniques are described (**Section 1.13.6**). The rationale behind using the combination of these techniques builds a narrative which asks, 1) using EMSA, do the candidate aptamers bind to the isolated recombinant GFR α 1 protein? 2) using SPR, what is the binding affinity for each of the candidate aptamers and do these aptamers demonstrate specificity to GFR α 1? 3) using MST, can the questions asked using SPR be validated? And finally, 4) using flow cytometry, can the

candidate aptamers bind to the surface of those cells overexpressing GFR α 1 and show discriminative binding ability between cells that have high or low level expression of GFR α 1?

Chapter 3 described the identification of 6 possible candidate aptamers (G_Apta1, -2, -3, -9, -13 and -15) that can bind to GFR α 1. Using classical SELEX, aptamers were selected for against the recombinant His-tagged GFR α 1 protein. After ten rounds of selection, with increasing stringency as outlined in **Section 3.3.1**, the final selection pool was subjected to NGS. Thereafter, the NGS data was interrogated and refined as described in **Section 3.3.2.3**, to which end six candidate aptamers were identified. The first three aptamers were identified based on sequence frequency and another three aptamers based on their secondary structure predictions. Herein, is detailed the bio-molecular interactions of these six candidate aptamers, using EMSA, SPR and MST. These three techniques are most commonly used methods to characterize aptamer binding ability, affinity and specificity (Jing and Bowser, 2011). To demonstrate the application of these aptamers to live cells, a binding assay, using flow cytometry was performed using a panel of human cell lines.



4.1.1. The aim of this chapter

To determine the binding affinity and specificity of 6 candidate aptamers developed against GFR α 1.

4.1.2. The objectives of this chapter

- To determine if the candidate aptamers binds to GFR α 1, using EMSA.
- To determine the binding affinity and specificity of the candidate aptamers to immobilised GFR α 1, using SPR.
- To determine the binding affinity of the candidate aptamers to GFR α 1 in solution, using MST.
- To evaluate the binding specificity of the candidate aptamers towards the cell surface GFR α 1 protein, using flow cytometry.



4.2. Materials and equipment

4.2.1. Chemicals and reagents

40% Acrylamide	Promega
Ammonium Persulfate	Sigma-Aldrich
Ethanolamine, pH 8	Sigma-Aldrich
Ethidium Bromide	Thermo Fisher Scientific
Human Serum Albumin	Sigma-Aldrich
Immunoglobulin G	Sigma-Aldrich
Magnesium Chloride	Sigma-Aldrich
Nuclease free water	Sigma-Aldrich
O'GeneRuler 100bp DNA ladder	Thermo Fisher Scientific
Tetramthylethylenediamine	Sigma-Aldrich

4.2.2. Buffers

10 X TBE: 1 M Tris; 1 M Boric acid; 0.02 M EDTA (pH 8); adjusted pH to 8.3 and made up to a final volume of 1 ltr using dH₂O.

10% APS: 0.1 g APS dissolved in 1 ml of dH₂O.

EDC/NHS solution: 400mM EDC; 100mM

EMSA loading buffer: 60mM Potassium chloride; 10mM Tris, pH 7.6; 5% Glycerol; 0.01% Xylene cyanol; 0.01% Bromophenol; made to a final volume of 5 ml using dH₂O.

Regeneration buffer: 2M NaCl; 50mM NaOH buffer.

SELEX buffer: PBS (0.138M NaCl and 0.0038M KCl), pH 7.4 (Sigma), 1.5mM MgCl₂, made to a final volume of 1 ltr using milli-Q water.

Sodium acetate: 100mM NaAc (pH 4.5)

4.2.3. Oligonucleotide sequences

G_Apta1 Inqaba Biotech

GCTCCAGAAGATAAAATTACAGGGAACGTGTTGGTTGCGGTTCTTCCGATCTGCTGTGTTCTCTAT
CTGTGCCATGCAACTAGGATACTATGACCCC

G_Apta2 Inqaba Biotech

AGCTCCAGAAGATAAAATTACAGGGTTCCCAGCGGTCTCACTTGTGGTAAT
ACATCACTTCAGACGGGTTCAATTGCACAACACTAGGATACTATGACCCC

G_Apta3 Inqaba Biotech

AGCTCCAGAAGATAAAATTACAGGAATGAAGCTTTGAGCCTGTTTGAAGT
CATGTAAAGTATCGGTTTCTGCCATAACAACACTAGGATACTATGACCCC

G_Apta9 Inqaba Biotech

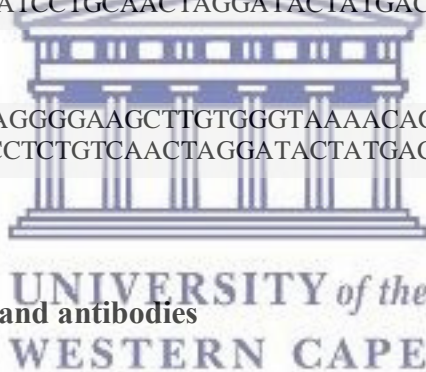
AGCTCCAGAAGATAAAATTACAGGATACGCTCTTCGCACGATCAGGATTTTC
AGTCTGTATCTTCGGGTTACGTCCCAACTAGGATACTATGACCCC

G_Apta13 Inqaba Biotech

AGCTCCAGAAGATAAAATTACAGGATGGAAGCTGTGCCTGTTTGAACCTTG
CTGTACGTTTACCGATGCTGTATCCTGCAACTAGGATACTATGACCCC

G_Apta15 Inqaba Biotech

AGCTCCAGAAGATAAAATTACAGGGGAAGCTTGTGGGTAAAAACAGGTGAG
TCCCCCGACATGAGGTCACCTCTGTCAACTAGGATACTATGACCCC



4.2.4. Recombinant proteins and antibodies

Anti-GDNF Receptor alpha 1 antibody Abcam

Human GFR α 1 His-tagged protein Sino Biological Inc.

Human GFR α 2 His-tagged protein Sino Biological Inc.

Human GFR α 3 His-tagged protein Sino Biological Inc.

4.2.5. Assay kits and Equipment

BD Accuri™ C6 BD Biosciences

BIAcore 3000 GE HealthCare

Biospectrum Imaging System UVP

Monolith NT.115 Pico

NanoTemper Tech

T25 cm² culture flasks

Sigma-Aldrich

Countess® Automated Cell Counter

Invitrogen



UNIVERSITY *of the*
WESTERN CAPE

4.3. Experimental methodology

4.3.1. Electrophoretic mobility shift assay

To investigate the potential binding of the aptamers to the protein of interest, the six candidate aptamers, identified in **Section 3.4.3**, were synthesized and subjected to EMSA as follows:

4.3.1.1. EMSA sample preparation

The aptamers, independently, were conditioned by adding 250 nM of the respective aptamer to 100 μ l of SELEX buffer (PBS, pH 7.4; 1.5 mM MgCl₂). The aptamers were then denatured at 95 °C for 5 min and gradually cooled (renatured), in increments of 10 °C every 30 sec, down to 25 °C. The binding reactions were setup by incubating (60min, at 25 °C, 500rpm) the respective aptamers (250 nM) with increasing concentrations (0 nM; 250 nM; 500nM and 750 nM) of the protein of interest in a final volume of 30 μ l.

4.3.1.2. Non-denaturing Poly-Acrylamide Gel Electrophoresis

The 8% non-denaturing gel was prepared as described in the table below:

Table 4.2. Preparation of an 8% EMSA gel

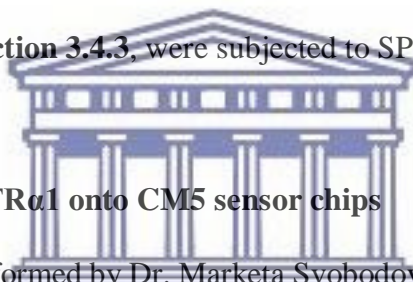
	Final concentration
Acrylamide (40%)	8 %
TBE, pH 8.6 (10x)	1X
APS (10%)	0.16 %
DH ₂ O	8.54ml
TEMED (100%)	0.1%
Total volume	12.5ml

The gel solution was prepared in a 15 ml Greiner tube by combining the reagents, in the order listed in **Table 4.1**. The solution was mixed by inverting the tube a few times and immediately poured between the gel plates, followed by inserting a ten-well comb between the gel plates.

The gel was then allowed to solidify. The gel plates were then transferred to an assembled PAGE system (Biorad). The combs were removed from the gels and the samples (prepared by mixing a 1:5 ratio of sample:EMSA loading buffer), at a final volume of 24 μ l, were loaded in the respective wells. The gel was electrophoresed in 1X TBE running buffer at 100V for 90 min. Each gel was stained for 30 min with EtBr (20 μ g/ μ l prepared in 50 ml of dH₂O), followed by de-staining for 15 min with dH₂O. Thereafter, the gels were visualized under UV light and images were taken using the Biospectrum imaging system (UVP).

4.3.2. Surface plasmon resonance analysis

To investigate the binding interaction of the aptamers to the GFR α 1, the six candidate aptamers, selected and synthesized in Section 3.4.3, were subjected to SPR as follows:



4.3.2.1. Immobilization of GFR α 1 onto CM5 sensor chips

All SPR experiments were performed by Dr. Marketa Svobodova, using the Biacore 3000 (GE Healthcare) at the Centre for Omic Sciences (Reus, Spain). After insertion of the CM5 (carboxylated) sensor chips, the sensor chips were primed using SELEX buffer. This was followed by activating all the channels on the sensor chip, using EDC/NHS buffer (400mM:100mM), at a flow rate of 10 μ l/min. The human GFR α 1 recombinant protein (Sino Biological Inc.), at a final concentration of 200 μ g/ml, was prepared in NaAc (100 mM, pH 4.5) and then injected into channel two, at a flow rate of 5 μ l/min. IgG and HSA, were prepared and injected in a similar manner, into channels three and four, respectively. To remove excess buffer, the sensor chip was then regenerated, three times, by injecting NaCl/NaOH buffer (2M:50mM), at a flow rate of 5 μ l/min. The EDC crosslinking on the sensor chip was inactivated by injecting ethanolamine (1M, pH 8). The sensor chip was then regenerated, three

times, as previously described; and the baseline was allowed to stabilize by allowing the selection buffer to flow over the sensor chip for approximately 1-2 hours.

4.3.3. Microscale thermophoresis

To validate the binding interaction of the G_Apta1 to the protein of interest, a fluorescent MST binding assay was performed as follows:

All MST experiments were performed by Dr. Thomas Schubert, at 2Bind Molecular Interactions (Germany). Serial dilutions (from 153pM – 9.09 μ M) of the GFR α 1, GFR α 2 and GFR α 3 proteins, respectively, were prepared in SELEX buffer.

Table 4.2. Preparation of proteins for MST assay.

	Maximum concentration	Minimum concentration
GFRα1	9.09 μ M	277 pM
GFRα2	8.33 μ M	254 pM
GFRα3	5 μ M	153 pM

Each dilution step (5 μ l) was mixed with 5 μ l of Cy5-labelled G_Apta1. The final reaction mixture, which were loaded in capillaries, contained a respective amount of the ligand (GFR α 1, GFR α 2 and GFR α 3), i.e. maximum concentration of 4.55 μ M, minimum concentration 139 pM and a constant concentration (2.5 nM) of the Cy5-labelled G_Apta1. The samples were analyzed using the Monolith NT.115 Pico, at 25 °C, with 15% LED power and 60% laser power.

4.3.4. Flow cytometry

4.3.4.1. Cell culture

The cell lines used in this study were obtained from the American Type Culture Collection (**Table 4.3**). The respective cell lines were regularly cultured in T25 cm² culture flasks at 37 °C, in a 5% CO₂ humidified incubator, in the respective culture media (**Table 4.3**). The media was supplemented with 10% heat-inactivated foetal bovine serum (FBS) and 1% penicillin-streptomycin. This media mix will be referred to as complete media from hereon. The media were changed and replaced with fresh complete media every three days. The procedures were performed in the Biological Safety Cabinet (Laminar flow) using aseptic techniques and all the reagents to be used were sprayed with 70% ethanol before introducing them into the Laminar flow.

Table 4.3. Cell lines used to screen for G_Apta1 binding.

Cell line	Acronym	Culture medium
Breast cancer (adenocarcinoma)	MCF7	Dulbecco Modified Eagle Medium (DMEM)
Breast (mammary gland)	MCF12A	DMEM Nutrient mixture F12 (DMEM-F12) + 20ng/ml EGFR, 500ng/ml hydrocortisone and 0.01mg/ml
Colon cancer (colorectal adenocarcinoma)	CaCo2	DMEM
Cervical cancer (epidermoid carcinoma)	Caski	Roswell Park Memorial Institute Medium (RPMI 1640)
Liver cancer (hepatocellular carcinoma)	HepG2	DMEM
Embryonic kidney	Hek293	DMEM
Prostate cancer (adenocarcinoma)	PC3	DMEM-F12
Keratinocyte	HaCat	DMEM

4.3.4.2. Cell harvesting

The cells were harvested when they reached $\geq 70\%$ confluency. The media was aspirated from the flask using a 10 ml serological pipette. The cells were then washed with 5 ml of PBS and discarded into a waste beaker. This was followed by the addition of 2ml of 2X trypsin to the

cells, in the flask, and then incubated for 1-3 minutes at 37 °C. The detachment of the cells from the flask was inspected using a light microscope. As soon as the cells start to round up, the detachment was facilitated by gently tapping the flask on the surface to dislodge the cells. The trypsinization reaction was terminated by adding 3 ml of complete medium. The cells that were still attached to the flask were removed by pipetting. The cells were then transferred to pre-labeled 15 ml conical tubes and centrifuged at 3000 rpm for 5 min. The supernatant was aspirated from the flask and the cell pellet were then re-suspended in 1-5ml complete media. A cell count was performed and an appropriate number of cells were sub-cultured.

4.3.4.3. Cell counting/viability

A cell count/viability was performed using trypan blue exclusion assay, using the Countess® Automated Cell Counter, according to the manufacturer's instructions. Briefly, a 10 µl aliquot of the trypsinized cell suspension was mixed with 10 µl of 0.4% trypan blue dye reagent. The cells were mixed by pipetting and 10 µl of the cell-dye mixture was then loaded into a Countess™ chamber slide. Following a cell count, the data readout was given in total cells/ml, live cells/ml, dead cells/ml, and percent viability; of which 2×10^4 live cells/ml were sub-cultured in 24-well plates for 24 hrs, at 37 °C, in a 5% CO₂ humidified incubator.

4.3.4.4. G_Apta1 binding assay using flow cytometry

The culture medium from the cells cultured in **Section 4.3.4.3** was aspirated, followed by the addition of 500 µl of PBS. The PBS was discarded into a waste beaker and the cells were trypsinized as described in **Section 4.3.4.2**. The cell pellet was re-suspended in PBS containing 31.25 nM of either the Cy5 labeled G_Apta1 or unlabeled G_Apta1. The cells were then incubated for 1 hr at 37 °C, in a 5% CO₂ humidified incubator. Unstained cells were included

as a control. Another control was included in which the cells were blocked with PBS containing 31.25 nM of the unlabeled G_Apt1 for 1hr, at 37 °C, in a 5% CO₂ humidified incubator. The PBS was then discarded into a waste beaker and fresh PBS containing 31.25 nM of the Cy5 labeled G_Apt1 was added to the cells and incubated for a further hour at 37 °C, in a 5% CO₂ humidified incubator. The respective incubated cell mixtures were then centrifuged at 3000 rpm for 3 min. The supernatant was discarded into a waste beaker and the cell pellet was re-suspended in 500 µl of PBS. The cells were centrifuged at 3000 rpm for 2 min. The supernatant was carefully discarded into a waste beaker and a wash with PBS was repeated. The cell pellet was then re-suspended in 200 µl of PBS and analysed using the BD Accuri™ C6 flow cytometer.



4.4. Results and discussion

In order to accomplish the objectives of this chapter, three separate characterization techniques, that is, EMSA, SPR and MST and a flow cytometry assay were performed to investigate the binding interaction between the candidate aptamers and GFR α 1.

4.4.1. EMSA: do the candidate aptamers bind GFR α 1?

As previously mentioned, EMSA is a qualitative technique and therefore limited to providing information about whether the aptamer is, potentially, binding to the protein of interest; and therefore, for the purpose of this study was used as a first line screening method. A detailed description of the methodology is provided in **Section 4.3.1**. EMSA is a simple method which only requires prior incubation of the aptamer with the target protein, followed by electrophoresis under non-denaturing conditions, using polyacrylamide (Hellman and Fried, 2007). The principle is that the bound aptamer-protein complex, because of its larger size, is expected to migrate slower than the corresponding free unbound aptamer and this differential migration provides insight into whether the aptamer potentially binds the protein (Hellman and Fried, 2007; Holden and Tacon, 2010).

For each candidate aptamer, a negative control, that is, free unbound aptamer, was included. And as expected for the negative control (lane labelled "C"), only one band, corresponding to the aptamer (97NT) was observed, in each case (**Figure 4.1**). To demonstrate the potential binding of the six candidate aptamers; increasing concentrations, ranging between 0 – 750nM (labelled on each lane), of the recombinant GFR α 1 were incubated with a constant concentration (250nM) of the respective aptamers.

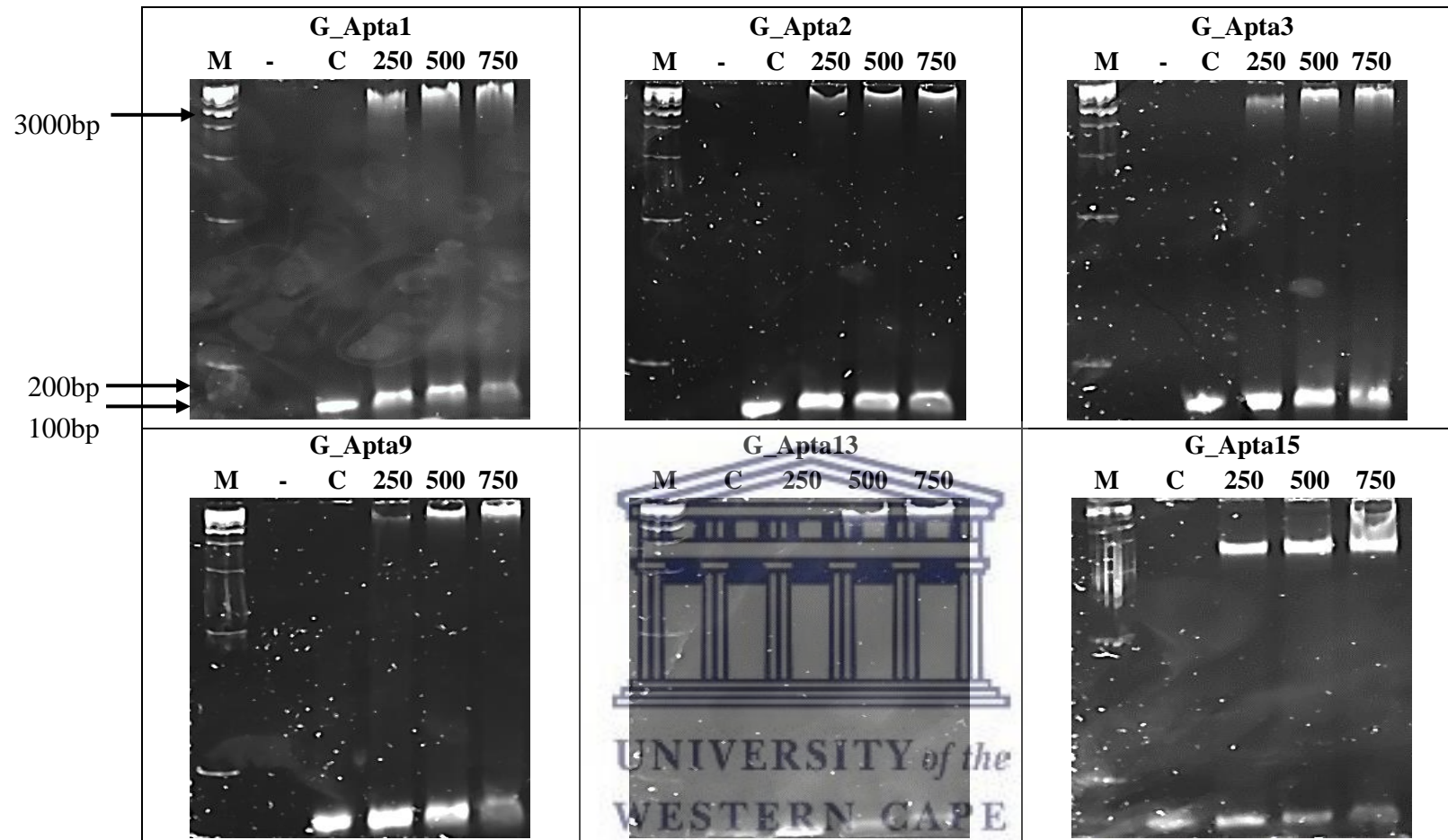
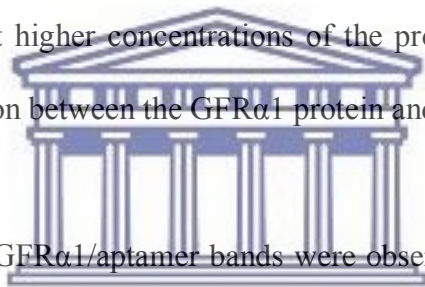


Figure 4.1. Electrophoretic mobility shift assay of six candidate aptamers against GFR α 1. The respective aptamers (250nM) were incubated with increasing concentrations (0, 250, 500 and 750 nM) of GFR α 1 and resolved on a 3% non-denaturing PAGE gel.


The presence of a band with a molecular weight higher than that of the free unbound aptamer suggests that there is interaction between the aptamer and the protein. All 6 aptamers (G_Apta1-6) showed the capacity to bind to the GFR α 1 protein, as shown by the high molecular weight band present in **Figure 4.1**. G_Apta13 was the only aptamer where no complex formation was observed at the lowest (250nM) GFR α 1 concentration. The intensity of the high molecular weight band increased as the concentration of GFR α 1 increased from 250nM to 750nM. This suggests that as the concentration of the protein increases, so too does the concentration of complex (GFR α 1/aptamer). The intensity of the low molecular weight band, which represents the aptamers, decreased as the concentration of the GFR α 1 was increased. This is due to the increasing amounts of aptamers binding to the GFR α 1 as more GFR α 1 becomes available at higher concentrations of the protein. This also confirms that there was indeed an interaction between the GFR α 1 protein and the aptamers.



In some cases, very distinct GFR α 1/aptamer bands were observed (G_Apta2, 9, 13 and 15) and in other cases, smears were observed (G_Apta1 and 3). It is possible that the GFR α 1-aptamer complexes dissociated either prior to electrophoresis and as a result, the co-migration of the aptamer and protein caused smearing effects; or potentially, the heat produced during electrophoresis resulted in dissociation of the complex. These can be managed easily by limiting the time taken to load samples, optimizing the time and voltage required to electrophorese the gel and controlling the temperature during electrophoresis using a cooling system or performing the electrophoresis in a cold room (Holden et al. 2011).

4.4.2. SPR: Investigating the binding affinity and specificity of the candidate aptamers

SPR is the most commonly employed method to characterize aptamer/target interactions (Ruscito and DeRosa, 2016). It provides information about binding, binding kinetics and can also be used to determine the binding affinity (K_D) values in low mM to the nM range (Jerabek-willemsen et al. 2011; Jerabek-willemsen et al. 2014). A detailed description of the methodology is provided in **Section 4.3.2**. In principle, it is a label-free optical based detection method that measures the change in the refractive index, on a metal surface layer of a sensor, in response to ligand/analyte interactions (Jerabek-willemsen et al. 2011; Nguyen et al. 2015). The binding kinetics, consists of four phases, that is, 1) association, 2) steady state, 3) dissociation and 4) regeneration (Ritzefeld and Sewald, 2011).



During the association phase, the analyte is injected and binding may occur when the ligand and analyte collide at the sensor surface, due to diffusion and the point at which the correct orientation and a sufficient amount of energy is achieved (Ritzefeld and Sewald, 2011). The steady state phase represents the point at which the association and dissociation of the ligand/analyte complex occurs at equal rates (Ritzefeld and Sewald, 2011). Hereafter, the analyte injection ends and the analyte dissociates from the ligand – typically resulting in the reduction of the response units (Ritzefeld and Sewald, 2011). And finally, the regeneration phase, if necessary, is the point at which the regeneration buffer, usually a high salt concentration or detergents such as SDS, is used to wash off any remaining analyte molecules (*in the case of covalent immobilization*) or ligand molecules (*in the case of non-covalent immobilization*) from the sensor surface (Ritzefeld and Sewald, 2011).

To facilitate the SPR interaction, the ligand must be immobilized on the sensor chip surface (Ritzefeld and Sewald, 2011). For DNA-protein interactions, either one of the biomolecules

can be immobilized first, depending on the nature of the experiment; for example, in this study, several candidate aptamers were screened against one protein and therefore, the most cost effective approach was to immobilize the GFR α 1 protein. **Figure 4.2** shows the successful immobilization of GFR α 1 and the control proteins (HSA and IgG) on a CM5 sensor chip (BIAcore). The CM5 sensor chip is a carboxymethylated dextran matrix that is attached to the gold surface and is suitable for covalent coupling via amine, thiol, aldehyde or carboxyl groups (Muguruma et al. 2000).

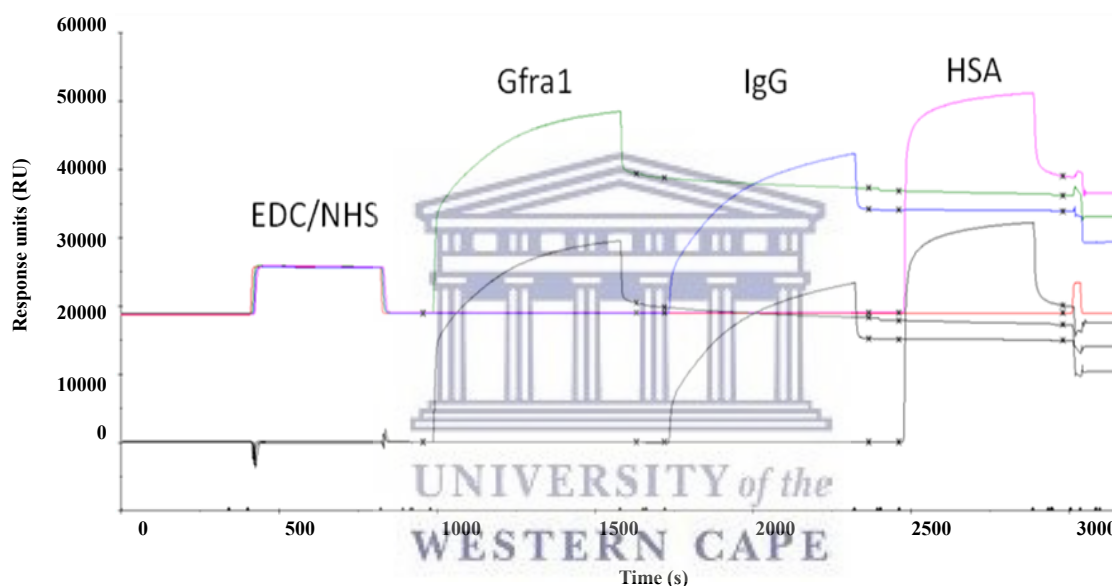


Figure 4.2. BIAcore sensorgram for the immobilization of GFR α 1 and control proteins on a CM5 sensor chip, using EDC chemistry. Each protein was bound to a separate channel. Channel 1: No protein (EDC/NHS) (red). Channel 2: GFR α 1 (green). Channel 3: Immunoglobulin G (IgG) (blue) and Channel 4: Human serum albumin (HSA) (pink).

In this study, the proteins were immobilized using 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) chemistry. This is observed by the first peak on the sensorgram, which indicates the activation of all the channels in the CM5 matrix. Each line on the sensorgram represents a channel and it is clear that all the channels were equally activated, based on the response units (on average >50 000 RU) that were similar throughout the injection and

stabilization after injection. Following activation, 200 ug/ml of the respective proteins were immobilized onto designated channels, that is, channel 2: GFR α 1 (green); channel 3: IgG (blue) and channel 4: HSA (pink). Channel 1 (red) was used as a blank control (buffer only), with no protein immobilized. This is performed to provide information about any non-specific interactions of the aptamers with the sensor surface (Ritzefeld and Sewald, 2011). After blank subtraction, all three proteins had relative response units at approximately 20 000 RU. Additionally, only one peak was observed for each protein injection, indicating that the proteins were only immobilized in the channels as expected and no leakage occurred between channels. The results obtained for the immobilization were satisfactory to proceed with the interaction analysis. At this point, the sensor surface was blocked with 1M ethanolamine.

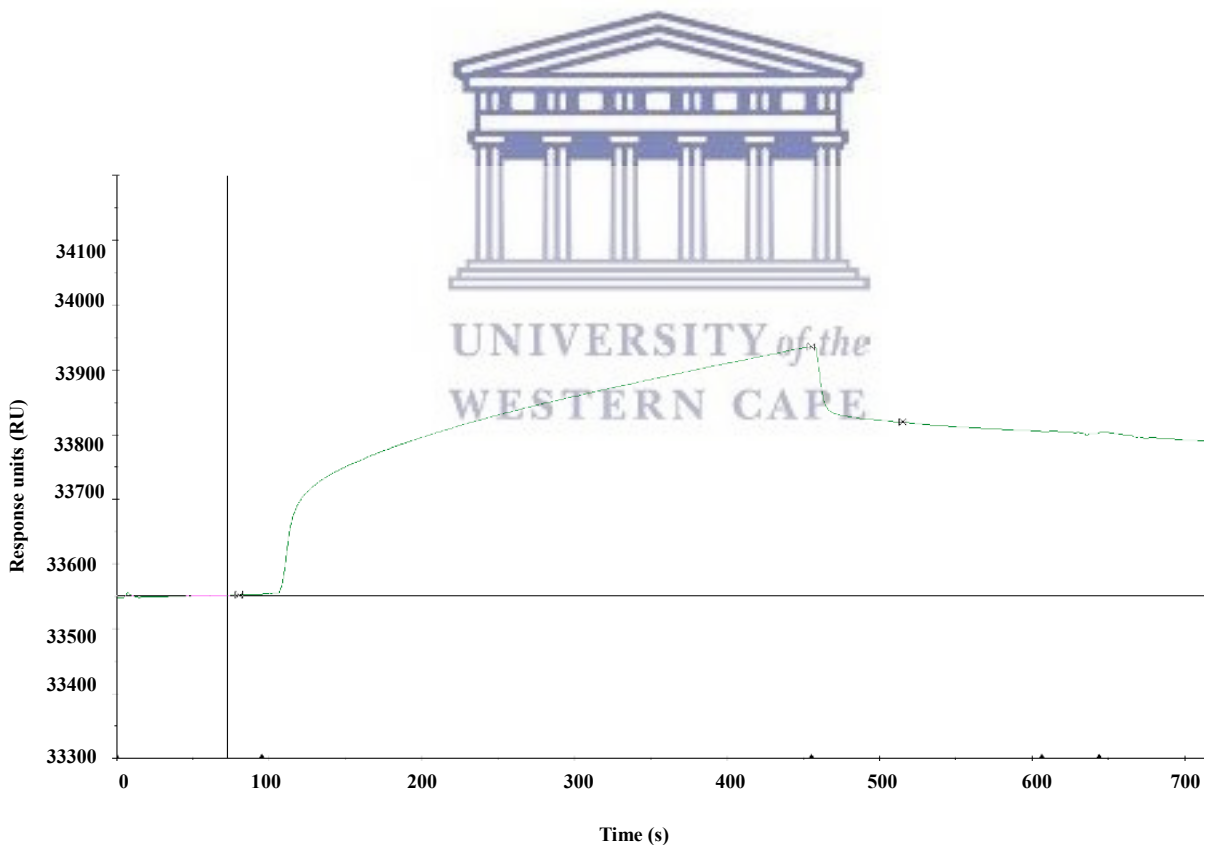



Figure 4.3. BIAcore sensorgram validating the immobilization of GFR α 1, using Anti-GFR α 1. Anti-GFR α 1 antibody (0.33 μ M) was injected into channel 2 (GFR α 1) at a flow rate of 5 ul/min.

To validate that the immobilization of GFR α 1 in fact occurred and that it was localized to the expected channel, 330 nM of anti-GFR α 1 antibody (Abcam) was injected into channel 2 (GFR α 1). To which end, an interaction with a relative response unit at approximately 33000 RU, before blank subtraction, was only observed in channel 2 (GFR α 1) (**Figure 4.3**). These results were satisfactory to proceed with the aptamer screening.

To establish if the aptamers indeed bind GFR α 1, the six candidate aptamers, respectively, were injected at 3.3 μ M and 10 μ M, to flow over all four channels, at a flow rate of 5 μ l/min. All the experiments were performed in duplicate to demonstrate reproducibility. Several controls were included; these were as follows:



1) A blank channel (SELEX buffer only) on the sensor chip to ensure that no interactions were occurring with the sensor surface and erroneously translated as binding. As a double reference, SELEX buffer was introduced as an analyte injection to further validate that any binding curves observed were not in fact due to buffer.

2) Both IgG and HSA, which were used in the counter selection process, during SELEX, were immobilized onto the sensor chip to establish that no cross-reactivity occurred with these proteins. This is also important because the future application of the aptamers aim to detect GFR α 1 in serum samples, which is generally rich in other unrelated proteins, such as HSA. If the aptamers are not able to discriminate between proteins, the detection test may yield false positive results.

3) The use of three unrelated aptamer sequences (C1, C2 and C3), to ensure that the interaction observed with the candidate aptamers are a true reflection of their binding and that it is not attributed to the intrinsic protein properties that may, otherwise, be favourable for DNA interactions, in general. This aspect was not evaluated during the analysis since an unrelated aptamer was not employed to test the specificity of the aptamers.



Graph A

Graph B

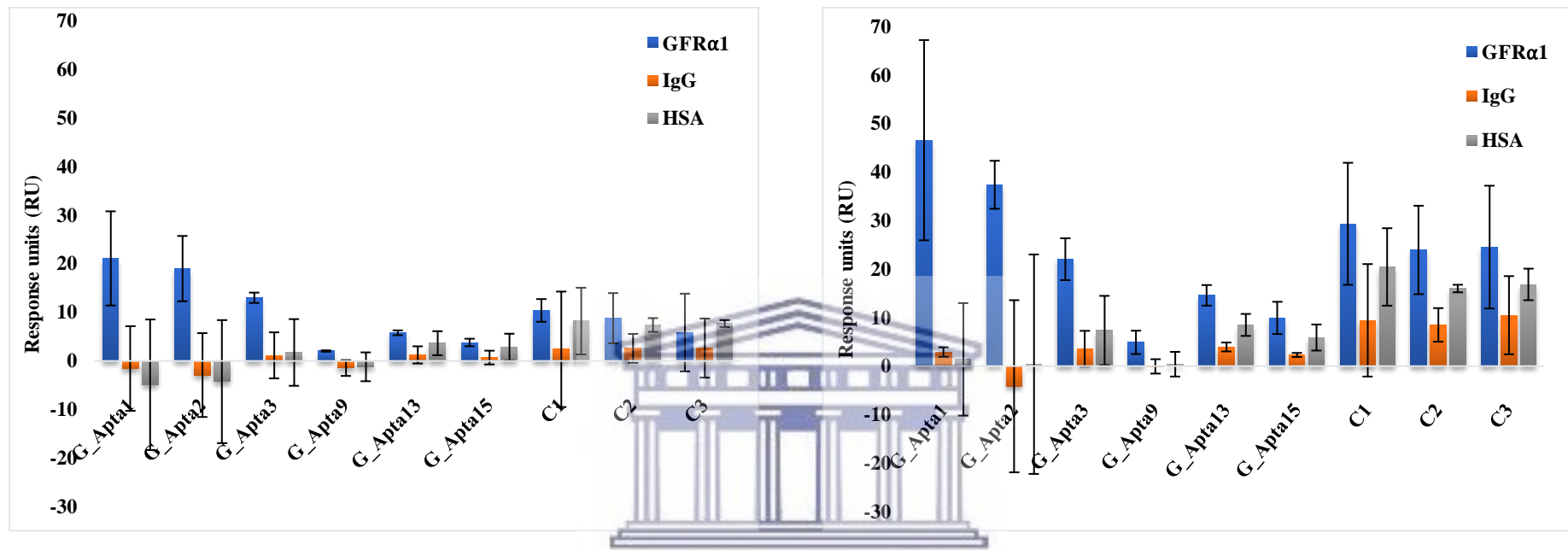
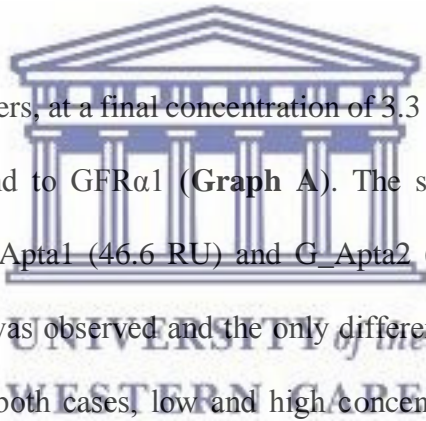


Figure 4.4. Screening of six candidate aptamers for binding to GFRα1 and cross-reactivity to control serum proteins. Binding measurements of the six candidate aptamers (G_Apta1, 2, 3, 9, 13 and 15), in all sensor channels, were performed at 3.3 μM (Graph A) and 10 μM (Graph B), with buffer subtractions. Three unrelated aptamers were used as negative controls, denoted as C1, C2 and C3. Error bars represents standard deviations from duplicate measurements.

The results as shown in **Figure 4.4** are expressed as values after the buffer subtraction. Graph A and B represents the aptamer binding measurements performed at 3.3 μM and 10 μM , respectively. One RU is equivalent to the binding of approximately 1pg protein/ mm^2 of the sensor surface and in practice, >50 pg/ mm^2 (50 RU's) of analyte binding is required to produce good reproducible results (Joshi 2006). However, to achieve this level of analyte binding can be challenging because it would require high density immobilization of the ligand (Joshi 2006). To an extent, Biacore sensor chips have compensated for this difficulty by generating 100-200 nm thick dextran matrices, which adds a third dimension to the surface and in turn allows for higher ligand immobilization (Joshi 2006). Thus, RU values between 30-50 are also acceptable and indicative of successful binding (Joshi 2006).



From the six candidate aptamers, at a final concentration of 3.3 μM , only G_Apta1 (21.1 RU) and G_Apta2 (19 RU) bound to GFR α 1 (**Graph A**). The same trend follows at a final concentration of 10 μM , G_Apta1 (46.6 RU) and G_Apta2 (37.4 RU) (**Graph B**) where selectively towards GFR α 1 was observed and the only difference, as expected, were higher relative RU's at 10 μM . In both cases, low and high concentrations, G_Apta3, G_Apta9, G_Apta13 and G_Apta15 did not show any evidence of binding to GFR α 1. Further evaluation using the serum proteins, IgG and HSA, demonstrated that all the aptamers do not bind non-specifically to these proteins; which is indicative of the specificity and potential discriminative power of the aptamers to bind GFR α 1 in serum.

Due to the fact that GFR α 1 has a theoretical pI of approximately 8, it is possible that the positively charged nature of the protein, may in fact, favour an electrostatic interaction with DNA, in general. It was therefore important to determine if the unrelated aptamer sequences

(C1, C2 and C3) may bind to GFR α 1. These unrelated aptamers didn't show any selectivity to GFR α 1 (**Figure 4.4**).

A point to consider is that the hypothesis used in the selection of G_Apta9, G_Apta13 and G_Apta15, as candidate aptamers, based on the secondary structure prediction and lowest minimal free energy values, proved to be null and void. Perhaps a better approach would be to employ a program such as Vfold, a software tool that can be used to predict the secondary structure of aptamers, which contains an algorithm that minimises the free energy and is optimised to predict all the stem candidate combinations, including the G-Quartet structures (Akitomi et al. 2011). Since aptamers are generally known to bind their target by means of the single-stranded bulge or loop regions, Vfold is designed to extract these regions from the predicted secondary aptamer structure and use this to calculate scores for conserved motifs (Akitomi et al. 2011). If the conserved motifs are found in either the bulge or loop region, further processing by truncation should aim to retain these motifs and be evaluated by binding assays (Akitomi et al. 2011). In contrast, G_Apta1, 2 and 3, which were selected based on sequence frequency, proved to favour the hypothesis which stated that as the selection library evolves to bind the target protein, the most abundant occurring sequence may potentially have the highest affinity towards the target. Although reported in literature that this is not always the case, it was observed that from all the aptamers candidates, G_Apta1 and 2; in fact, bound GFR α 1 with the highest affinity in the order listed.

It is important to note that GFR α 1 belongs to a family of proteins and in some cases, despite the high specificity, the aptamers can also bind biomolecules with similar structures. Bearing this in mind, a multiple alignment of the GFR α protein family was generated using Clustal Omega and as expected the percentage similarity ranged from 32.32% (GFR α 4) up to 50.56%

(GFR α 2) (Figure 4.5). To this end, a comparative analysis of the binding of the 6 aptamers to GFR α 1, 2 and 3 was performed using SPR.

A

```

sp|060609|GFRA3_HUMAN      MVRPLNPRPLPPVWLMLLLLLPPSPLPL----AAGDPLPTESRLMNSCLQARRKQADPT
sp|Q9GZ27|GFRA4_HUMAN      -----MFLATLYFALPLLDL-----LLSAEVS-GGDRLDCVKASDQCLKEQS
sp|P56159|GFRA1_HUMAN      -----MILANVFLCFFLDETLRSLASPSLQGPGLHGWRPPVDCVRANELCAAESN
sp|000451|GFRA2_HUMAN

sp|060609|GFRA3_HUMAN      CSAAYHLDSDCTSSISTPLPS-EEPSVPADCLEAAQQLRNSSLIGCMCHRRMKNVQVACLD
sp|Q9GZ27|GFRA4_HUMAN      -----
sp|P56159|GFRA1_HUMAN      CSTKYRTLROCVAGKETNFSLASGLEAKDECRSAMEALKQKSLYNCRCKRGMKKEKNCLR
sp|000451|GFRA2_HUMAN      CSSRYRTLROCLAGDRNTML----ANKECQAALVLEQESPLYDCRCKRGMKKEQLCQLQ

sp|060609|GFRA3_HUMAN      IWYTVHRARSLGNYELDVSPYEDVTSTKPW-----KMNLSKLNMLKPPSDLCLKFAMIL
sp|Q9GZ27|GFRA4_HUMAN      -----MVRCLGPAL-----LLLLLLGSASSVGGNRCVDAAEA
sp|P56159|GFRA1_HUMAN      IWWSMYQSL-QGNDLLEDSPYEPVNSRLSDIFRVVPIISDVFQQVEHIPKGNCLDAAKA
sp|000451|GFRA2_HUMAN      IWWSIHLGLTEGEEFYEASPYEPVTSRLSDIFRLASIFSGTGADPVVSAKSNHCLDAAKA
                                : : * : *

sp|060609|GFRA3_HUMAN      CTLNDKCDRLRKAYGEACSG----PHCQRHVCLRQLLTFEKAAPHAQGLLLCPCAPN
sp|Q9GZ27|GFRA4_HUMAN      CTADARCQRLRSEYVAQCLGRAA-QGGCPRARCRRALRRFFARGPPALHALLFCPCA--
sp|P56159|GFRA1_HUMAN      CNLDDICKKYRSAYITPCTTSVS-NDVCNRRKCHKALRQFFDKVPAKHSYGMFCSCR--
sp|000451|GFRA2_HUMAN      CNLNDNCKKLRSYSISICNREISPTERCNRRKCHKALRQFFDRVPSEYTYRMLFCSCQ--
                                * : * : * : * * * * * * * * * * * * * * * * * * * * * * *

sp|060609|GFRA3_HUMAN      DRGCGERRRNTIAPNCLP---PVAPNCLLRLRCLFSDPLCRSRLVDFQT----HCHP
sp|Q9GZ27|GFRA4_HUMAN      GPACAERRRQTFVPSCAFSGPGAPPSCLEPLNFCERSRVCRARAAAGPWRGNGRGLSP
sp|P56159|GFRA1_HUMAN      DIACTERRRQTIIVPVSYEER--EKPNCLNLQDSCKTMYICRSRLADFFT----NCQP
sp|000451|GFRA2_HUMAN      DQACAERRRQTIILPSCSYEDK--EKPNCLDLRGVCRTHLCSRSLADFFHA----NCRA
                                * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

sp|060609|GFRA3_HUMAN      MD-----ILGTCA--TEQSRCLRAYLGLIGTAMTPNFVSNV
sp|Q9GZ27|GFRA4_HUMAN      AHRPPAAQASPPGLSGLVHPSAQRPRRLPAGPGRPLPARLRGPRGVPAGTAVTPNYVDNV
sp|P56159|GFRA1_HUMAN      ES-----RSVSSCLKENYADC--LLAYSGLIGTVMTPNYIDSS
sp|000451|GFRA2_HUMAN      SY-----QTVTSCPADNYQAC-LGSYAGMIGFDMTPNYVDSS
                                : : * : * * * : :

sp|060609|GFRA3_HUMAN      N--TSVALSCTCRGSGNLQEECEMLEGFFSHNPCLTEAIAAKMRFHSQFLSQDHPHPTFA
sp|Q9GZ27|GFRA4_HUMAN      S--ARVAPWCDGASGNRRBDCEAFRGLFTRNRCLDGAIQ-----FASGNPPVLLD
sp|P56159|GFRA1_HUMAN      S--LSVAPWCDGASGNNDLEECLKFLNFFKDNTCLKNAIQ-----FNGSDVTVWQ
sp|000451|GFRA2_HUMAN      PTGIVVSPWCSGRSGNHEECEEKFLRDFTENPCLRNAIQ-----FNGTDMVWSP
                                * : * * * * * * * * * * * * * * * * * * * * * * * * * * *

sp|060609|GFRA3_HUMAN      -----VMAHQENPAVRPQPIWPSLFC--TLPLILLLSLW-----
sp|Q9GZ27|GFRA4_HUMAN      QLNPQGDPEHSLQLQVSTGRALERRSL-----LSILPVLALPALL-----
sp|P56159|GFRA1_HUMAN      PAFPVQT---TTATTTALRVKNKPLGPA--GSENEIPTHVLPCCANLQAQKLSNVS
sp|000451|GFRA2_HUMAN      KGPSFQA----TQA-----PRVEKTPSLPDDLSSTLSLGTSTVITTTCTSVQEGLKANNSK
                                : : :

sp|060609|GFRA3_HUMAN      -----
sp|Q9GZ27|GFRA4_HUMAN      -----
sp|P56159|GFRA1_HUMAN      NTHLCISNGNYEKEGLGASSHIT--TKSMAAPPSCGLSPLLVLVVTALSTLLSLTETS
sp|000451|GFRA2_HUMAN      ELSMCFTELTTNII-PGSNKVIKPNSGPSRARPAAALTVLSVLMK-----LAL----

```

B

Uniprot ID	Protein	% Similarity			
		GFR α 3	GFR α 4	GFR α 1	GFR α 2
1: sp 060609 GFRA3_HUMAN		100.00	33.06	33.24	33.69
2: sp Q9GZ27 GFRA4_HUMAN		33.06	100.00	32.32	32.56
3: sp P56159 GFRA1_HUMAN		33.24	32.32	100.00	50.56
4: sp 000451 GFRA2_HUMAN		33.69	32.56	50.56	100.00

Figure 4.5. Multiple alignment (A) and percentage similarity matrix (B) of the GFR α protein family, using Clustal Omega.

Figure 4.6 shows the immobilization of the GFR α family of proteins, that is, Channel 2: GFR α 1, Channel 3: GFR α 2 and Channel 4: GFR α 3. Channel 1 was used as a blank control (buffer only). As before, the CM5 sensor chip was first activated using EDC/NHS, followed by immobilization of the respective proteins, blocking with ethanolamine and then regeneration using NaOH/NaCl. After the blank subtraction, immobilization was achieved at 6816 RU, 2887 RU and 4802 RU, for GFR α 1, 2 and 3, respectively.

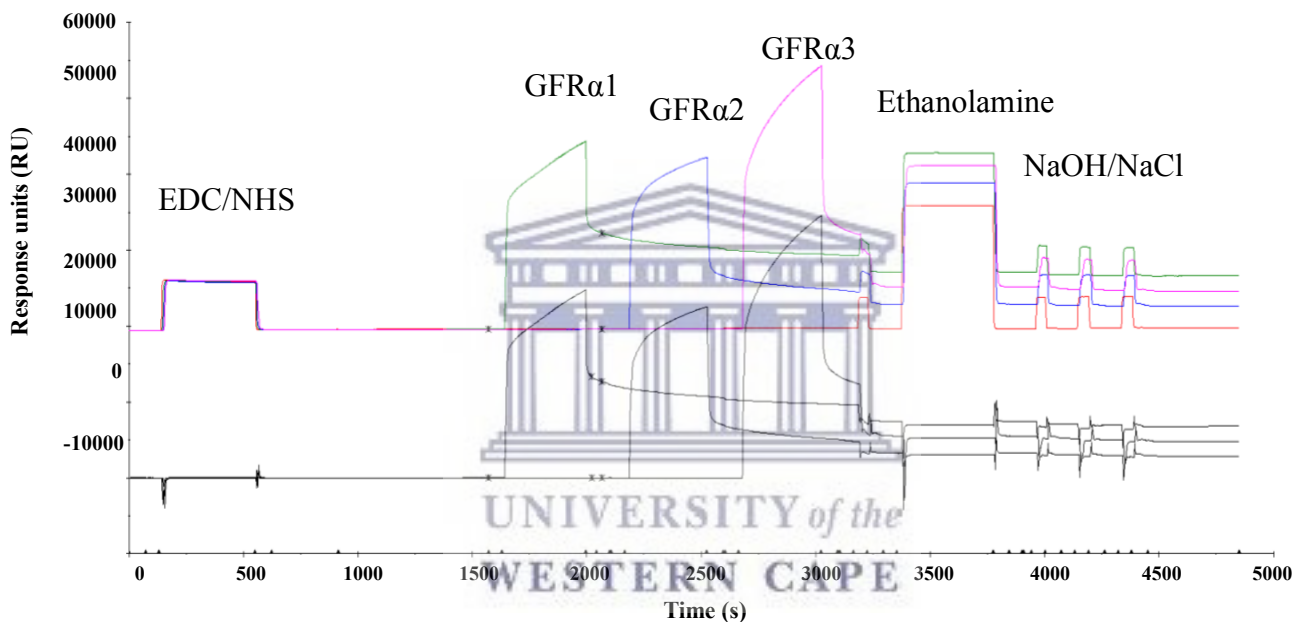


Figure 4.6. BIAcore sensorgram for the immobilization of GFR α family of proteins on a CM5 sensor chip, using EDC chemistry. Each protein was bound to a separate channel. Channel 1: Blank. Channel 2: GFR α 1. Channel 3: GFR α 2 and Channel 4: GFR α 4

All the channels were equally activated, based on the response units that were similar throughout the injection and stabilization after injection. Additionally, only one peak was observed for each protein injection, indicating that the proteins were only immobilized in the channels as expected and no leakage occurred between channels. The results obtained for the immobilization were satisfactory to proceed with the interaction analysis.

To validate that the immobilization of GFR α 1 in fact occurred and that it was localized to the expected channel, 330 nM of anti-GFR α 1 (Abcam) was injected into all four channels. To which end, an interaction was observed in all the channels, indicative of cross-reactivity of anti-GFR α 1 to the GFR α family of proteins. Although a signal was detected in the blank channel, the curve does not imitate a typical SPR curve, instead it shows evidence of a sharp increase, immediate plateau, sharp decrease and finally resolving back to the baseline. This trend is suggestive of full mass transfer. So, the analyte (anti-GFR α 1) transfers from the bulk solution to the sensor surface and no actual binding to the ligand (buffer) occurs, resulting in the immediate dissociation.

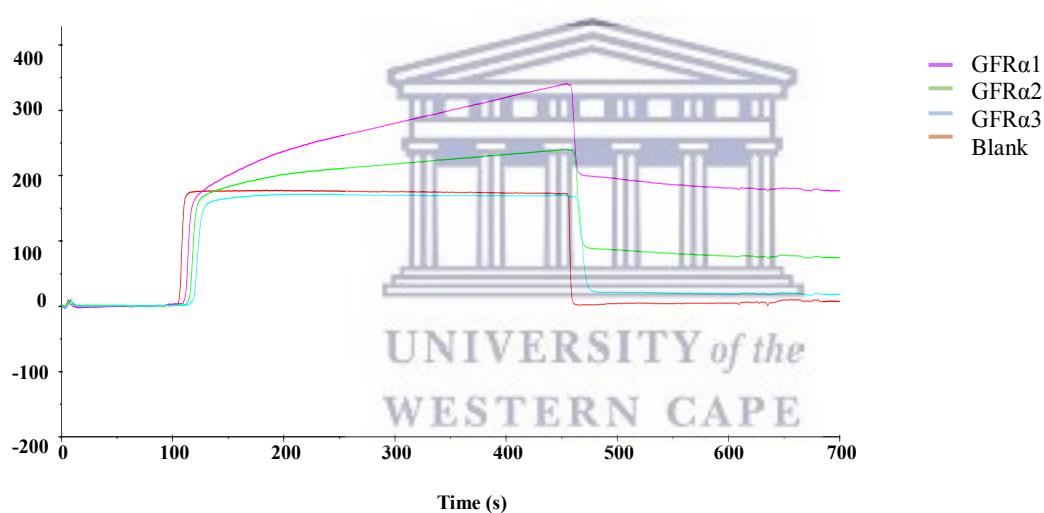
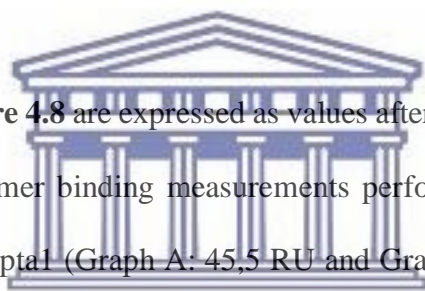


Figure 4.7. BIAcore sensorgram validating the immobilization of GFR α 1, using Anti-GFR α 1. 0.33 μ M Anti-GFR α 1 was injected into all channels at a flow rate of 5 μ l/min.

As with GFR α 3, a similar trend as that which occurred with the buffer control was observed and since GFR α 3 has a lower sequence similarity to that of GFR α 1 (**Figure 4.5B**), the region at which anti-GFR α 1 antibody binds is possibly not present on GFR α 3. Whereas with GFR α 2, the SPR curve depicts an association, followed by a steady increase towards an equilibrium and then a dissociation with the baseline equaling out above zero. Therefore, if the alignment

(**Figure 4.5A**) and structural similarity between GFR α 1 and 2 are taken into consideration, then it is not so unlikely that anti-GFR α 1 could show cross-reactivity because there is a possibility that GFR α 2 may be similar in sequence and structure in the same region where anti-GFR α 1 binds GFR α 1.

Another experiment was setup in a similar manner as with the previous SPR experiments. These included: 1) a blank channel (SELEX buffer only) on the sensor chip as well as a SELEX buffer injection over all the channels. 2) GFR α 2 and 3 were immobilized on channel 3 and 4, respectively, to determine the potential discriminative power and specificity of the aptamers developed against GFR α 1. 3) Two unrelated aptamer sequences were included as negative controls.



The results as shown in **Figure 4.8** are expressed as values after the buffer subtraction. Graph A and B represents the aptamer binding measurements performed at 3.3 μ M and 10 μ M, respectively. As before, G_Apta1 (Graph A: 45,5 RU and Graph B: 58,7 RU) and G_Apta2 (Graph A: 22,55 RU and Graph B: 57,55 RU) were the only aptamers that bound to GFR α 1. What was observed is that G_Apta1 and G_Apta2 only bound to GFR α 2 and GFR α 3 at a higher concentration (10 μ M). Although the GFR α protein family share sequence and structural similarities (**Figure 4.5**), the cross-reactivity of G_Apta1 and G_Apta2 towards GFR α 2 and GFR α 3 was almost absent at a lower concentration. Therefore, at 10 μ M, the binding observed could be accounted for as non-specific interactions as a result of the high concentration of aptamers. In the case of the unrelated aptamer sequences C1 and C2, it was expected that no binding should occur since these aptamers were not developed against the GFR α family. As expected, the control aptamers didn't show any selectivity to any of the GFR α proteins. Only C2 (Graph B: 34,5 RU) bound non-specifically to GFR α 1, at a high concentration.

Graph A

Graph B

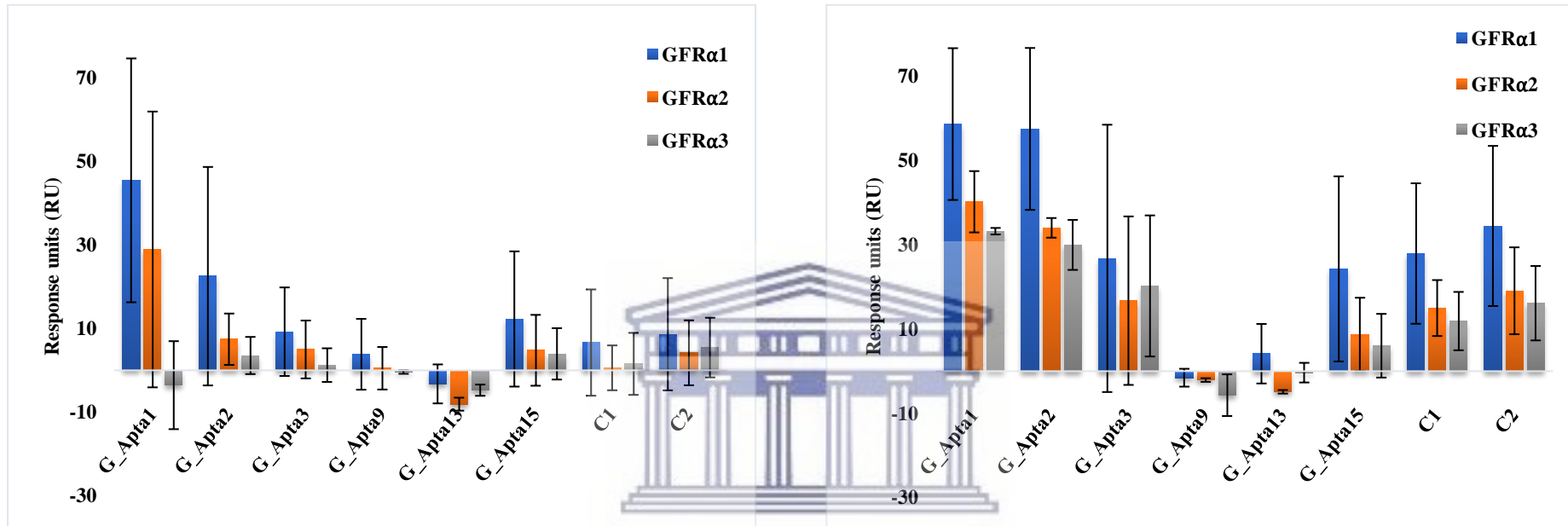


Figure 4.8. Evaluating the specificity of the six candidate aptamers against the GFRα protein family. Binding measurements of the six candidate aptamers (G_Apta1, 2, 3, 9,13 and 15), in all four channels, were performed at 3.3μM (Graph A) and 10μM (Graph B), with buffer subtractions. Two unrelated aptamers were used as negative controls, denoted as C1 and C2. Error bars represents standard deviations from duplicate measurements.

The SPR data (**Figure 4.8**) still supports G_Apta1 and G_Apta2 as potential candidates to take forward; of which G_Apta1 shows the strongest affinity to GFR α 1. This, however, should be validated by another method. For the purpose of this study, G_Apta1 was further investigated.

4.4.3. MST: Investigating the binding interaction of G_Apta1 and the GFR α family.

MST is a highly sensitive quantitative method used to characterize, virtually, any kind of molecular interaction based on the principle of thermophoresis, under microscopic temperature gradients (Jerabek-Willemsen et al. 2014). Upon a binding event, even the slightest change in conformation, charge or size of the molecule-solvent interface is directly translated into the thermophoretic effect detection for each interaction. Detection is monitored using either labeled or intrinsic fluorophores present in the sample. This method has a detection limit in the pM range, which is particularly applicable for aptamer/target interactions (Jerabek-willemsen et al. 2011; Ruscito and DeRosa 2016). In this study, G_Apta1 was labelled with a Cy5 fluorophore in order to investigate its binding affinity and specificity to GFR α 1.

In order to validate the results obtained from the SPR analysis, further evaluation of the binding affinity was performed using MST. In SPR, the ligand was immobilized, whereas with MST another dimension is added because the binding capability and affinity of the protein-aptamer interaction can be evaluated in a free solution. The MST process begins with a quality control check of the capillaries to ensure that the fluorescent aptamer does not stick to the walls of the capillary, due to aggregation, precipitation or adsorption, subsequently interfering or preventing the thermophoretic movement of the aptamers (Jerabek-Willemsen et al. 2011; Plach et al. 2017). In total, 16 capillaries with 16 different concentrations can be

analysed in one experiment. Each peak, observed in **Figure 4.9**, represents a capillary and in the event that any sticking occurred, the shape of the peak (U-shaped peaks, bumps or spikes) would alter. Furthermore, if large differences in the fluorescence read outs between the capillaries are observed, it can also be indicative of sample aggregation. This is because aggregates are usually characterised by having significantly higher fluorescent densities, subsequently exhibiting high fluorescent read outs. The capillaries were scanned for each interaction, that is, Cy5-labelled G_Apta1 with GFR α 1, 2 and 3; showed evidence of adequate quality. In each case, the response for the quality control check was the same, therefore, one capillary scan was included as a representation of all the interactions (**Figure 4.10**).

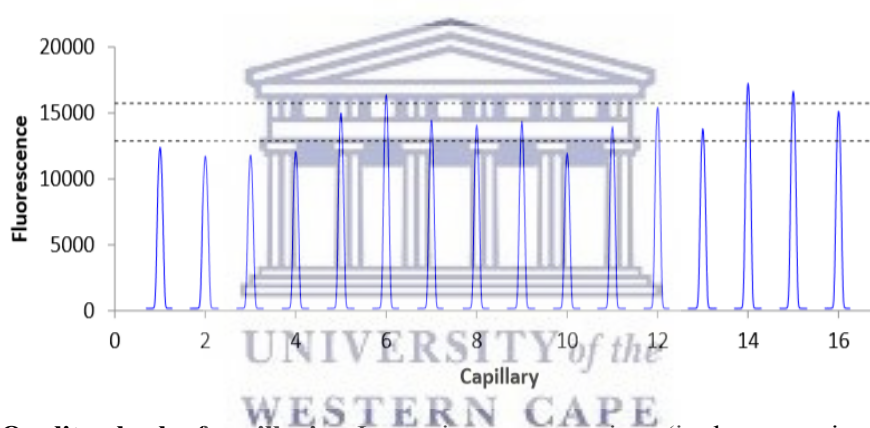


Figure 4.9. Quality check of capillaries. Increasing concentrations (in the respective capillaries) of the respective GFR α 1 proteins were incubated with the respective candidate aptamers. This was followed by loading all 16 capillaries and inducing a temperature gradient, using infrared light.

The focal temperature gradient, within the glass capillary, required for the thermophoretic movement of the molecules was then generated using infrared (IR) light. The target molecules, in this case, the Cy5-labelled G_Apta1, located in the IR light, were monitored by its fluorescence and movement along the temperature gradient (**Figure 4.10**). At first, in the absence of a temperature gradient, G_Apta1's fluorescence was detected to ensure homogeneity. Then after ~5sec, activation of the IR light induced the temperature gradient; which resulted in a step initial drop in the fluorescent signal, known as the Temperature jump

(T-jump). Hereafter, a slower thermophoretic depletion of the fluorophore, in the optic focus, occurred until the IR light was switched off and results in a reverse T-jump, to which end a back-diffusion of the Cy5 aptamers was observed. Again, any form of aggregation would have been noted as a scrambled and irregular graph, however, a smooth lined graph of the thermophoretic movement of the protein-aptamer complex was observed.

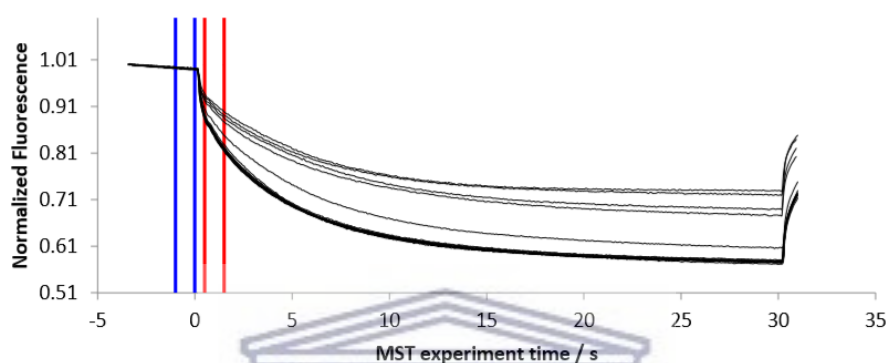
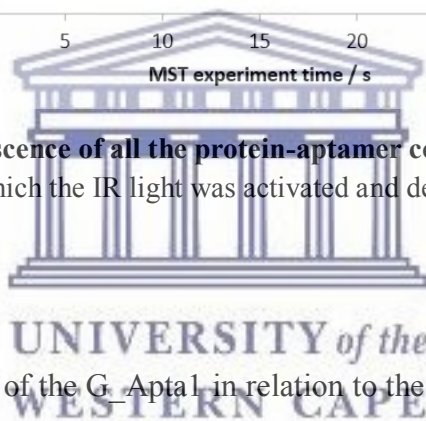


Figure 4.10. Normalized fluorescence of all the protein-aptamer complexes. The blue and red lines represent the point at which the IR light was activated and deactivated, respectively.



Towards determining the K_D of the G_i Aptal in relation to the GFR α protein family, several MST measurements, using increasing concentrations of the ligand and a fixed concentration of the analyte must be recorded (Plach et al. 2017). The ligand concentration range is prepared in a way that the lowest concentration, essentially, does not bind – typically, at a concentration that is 10-fold below the anticipated K_D . Conversely, the highest concentration should result in the ligand being fully bound (saturated), usually at a concentration that is 10-fold above the anticipated K_D . Additionally, the concentration of the analyte is kept below the anticipated K_D , to allow for accurate extrapolation of the equilibrium affinity.

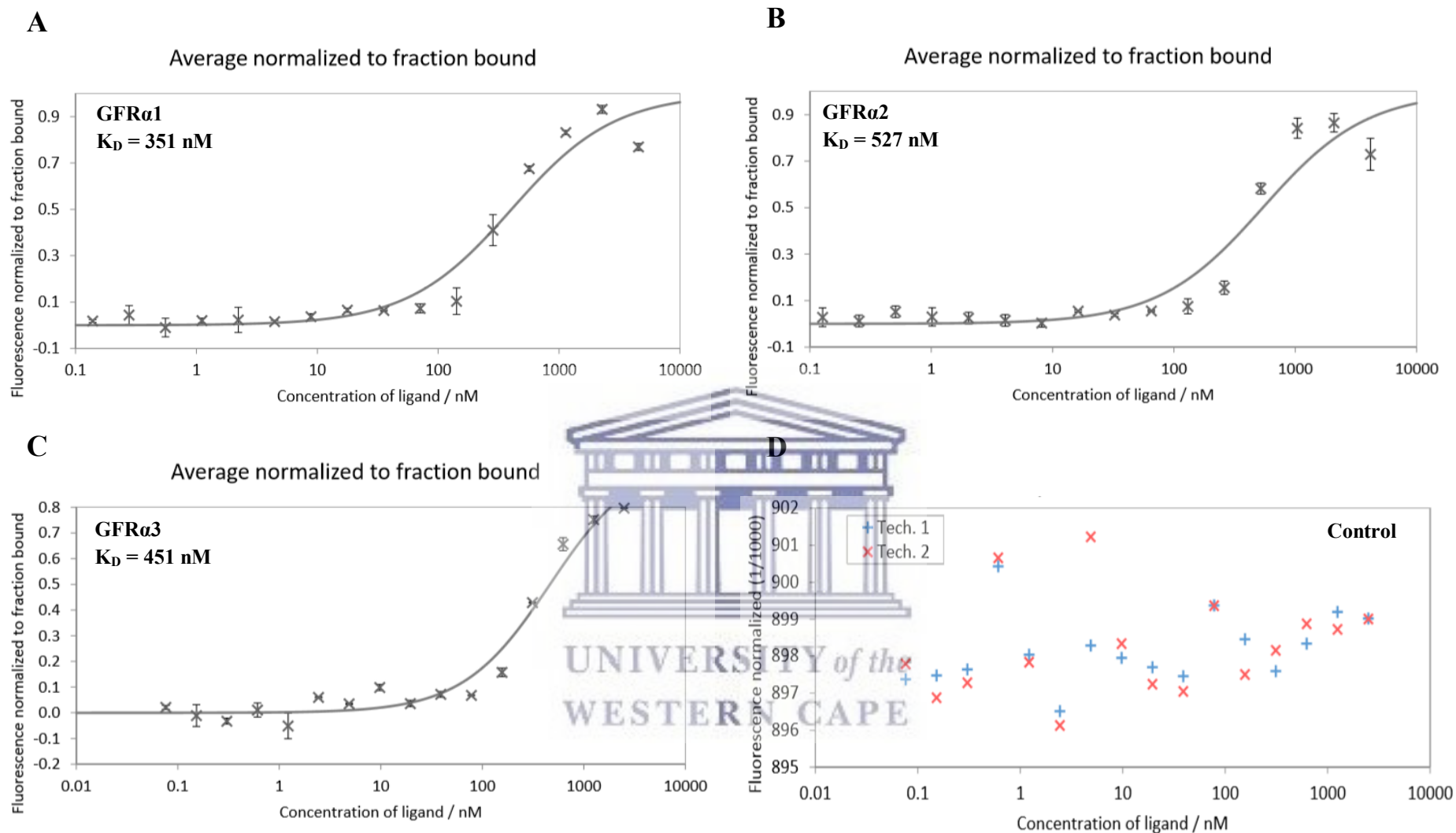
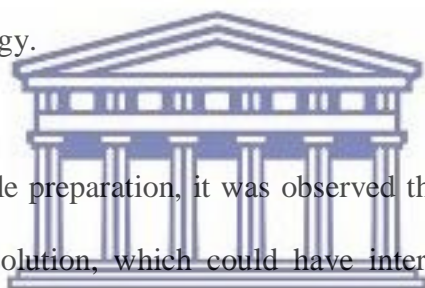


Figure 4.11. MST analysis of the interaction of the G_Apta1 with the GFR α protein family. The proteins (A, B and C: GFR α 1, 2 and 3, respectively) served as the unlabelled ligand and was added in increasing concentrations to the Cy5 labelled G_Apta1 that served as the analyte. An unrelated aptamer (D) was used as a control. The raw fluorescence was normalized to the bound fraction of G_Apta1. Error bars represent standard deviations from duplicate measurements. The K_D was determined from a 1:1 binding fit.

To comparatively evaluate the different interactions, the data was normalized to the fraction of bound molecules. The MST assay (**Figure 4.11**) demonstrated that G_Apta1 binds to GFR α 1 with high affinity. The experiment was performed in duplicate and the average calculated K_D value of this interaction was determined to be 351 nM. G_Apta1, also bound GFR α 2 with an average K_D value of 527 nM. In the case of GFR α 3, evidence of binding was detected and a K_D value of 451 nM was calculated; which may not be reliable because a proper curve fit could not be established since no bound plateau was reached. Hence, either higher concentrations of the protein are required or there is a possibility that non-specific binding occurred. The latter is a more likely assumption, since with the SPR analysis, poor binding to GFR α 3 was observed and from the alignment data (**Figure 4.5**), it has been shown that GFR α 3 and GFR α 1 only share ~30 % sequence homology.



Furthermore, during the sample preparation, it was observed that at high concentrations, the proteins became viscous in solution, which could have interfered with the movement of G_Apta1. Subsequently, an unrelated aptamer control was included to evaluate if the protein viscosity influenced the MST measurement. No binding was observed for this aptamer (**Figure 4.11D**) and as a result, the data could not be normalized to the bound fraction. The results of the control confirmed that the viscosity effect had no impact on the interaction and subsequent movement of G_Apta1. The MST data validated that G_Apta1 binds with high affinity towards GFR α 1 and that cross-reactivity occurs against the GFR α 2; which can be attributed that the structural similarity between the family of proteins.

4.4.4. Evaluating the binding and specificity of G_Apta1 to cell surface GFR α 1.

Flow cytometry is a tool used to measure optical and fluorescence characteristics of single cells (Brown and Wittwer 2000). Every cell that passes through the flow cytometer and is detected is regarded as an event (Brown and Wittwer 2000). When labeled cells pass through a light source in the flow cytometer, the fluorescent molecules (fluorophores) are excited by a laser to a higher energy state and as these fluorophores return to their resting energy state, they emit light (colour) at higher wavelengths (Nilsson et al. 2013). In this study, G_Apta1 was labelled with a Cy5 fluorophore in order to investigate 1) if G_Apta1 is able to bind to GFR α 1 on MCF7 cells and 2) to investigate if G_Apta1 shows specificity to GFR α 1 expressing cells. In principle, after incubating the cells with Cy5 labelled G_Apta1, what was expected is that G_Apta1 would bind to GFR α 1 on the surface of the cells and as these cells pass through the light source, the Cy5 fluorophore attached to G_Apta1 would generate a fluorescent signal that is measured to generate a histogram. From the histogram, the Cy5 fluorescence intensity (horizontal axis) vs the number of detected events (vertical axis) will be displayed. This allows for the identification of distinct cellular population subsets, that is either a negative population as exhibited by lower fluorescence intensity (a peak at low number of events) or a positive population as exhibited by a higher fluorescence intensity (a peak that shifts towards the right or higher number of events) for Cy5 labeled G_Apta1 staining.

To establish if G_Apta1 indeed binds to the cell surface GFR α 1, MCF7 cells were incubated with the Cy5 labelled G_Apta1 as described in **Section 4.3.4**. All the experiments were performed in triplicate to demonstrate reproducibility and several controls were included as follows:

- 1) MCF12A cells were included as a negative control cell line, as demonstrated by Drah (2015) that MCF7 cells expresses significantly higher protein levels of GFR α 1 as compared to MCF12A cells (Drah 2015, University of the Western Cape, Department of Biotechnology). This was important to determine staining specificity.
- 2) MCF7 and MCF12A cells that were not incubated with the Cy5 labelled G_Apta1 (unstained cells). These were included to identify the negative population, of which background fluorescence could be observed due to cellular auto-fluorescence; therefore, the negative population was expected to be displayed as a low number of detected events on the histogram (Nilsson et al. 2013).



UNIVERSITY of the
WESTERN CAPE

- 3) MCF7 and MCF12A cells incubated with unlabelled G_Apta1 were included to ensure that the fluorescence observed were not artefacts caused by the interaction between the cells and the aptamer.
- 4) MCF7 and MCF12A cells that were blocked with unlabelled G_Apta1, followed by incubation with the Cy5 labelled G_Apta1. This was done to establish that G_Apta 1 in fact bound to GFR α 1 and no non-specific interactions between the fluorophore and the cells occurred.
- 5) Finally, in the second experiment, a panel of human cancer cell lines were evaluated to determine the specificity of G_Apta1 towards GFR α 1 in MCF7 cells.

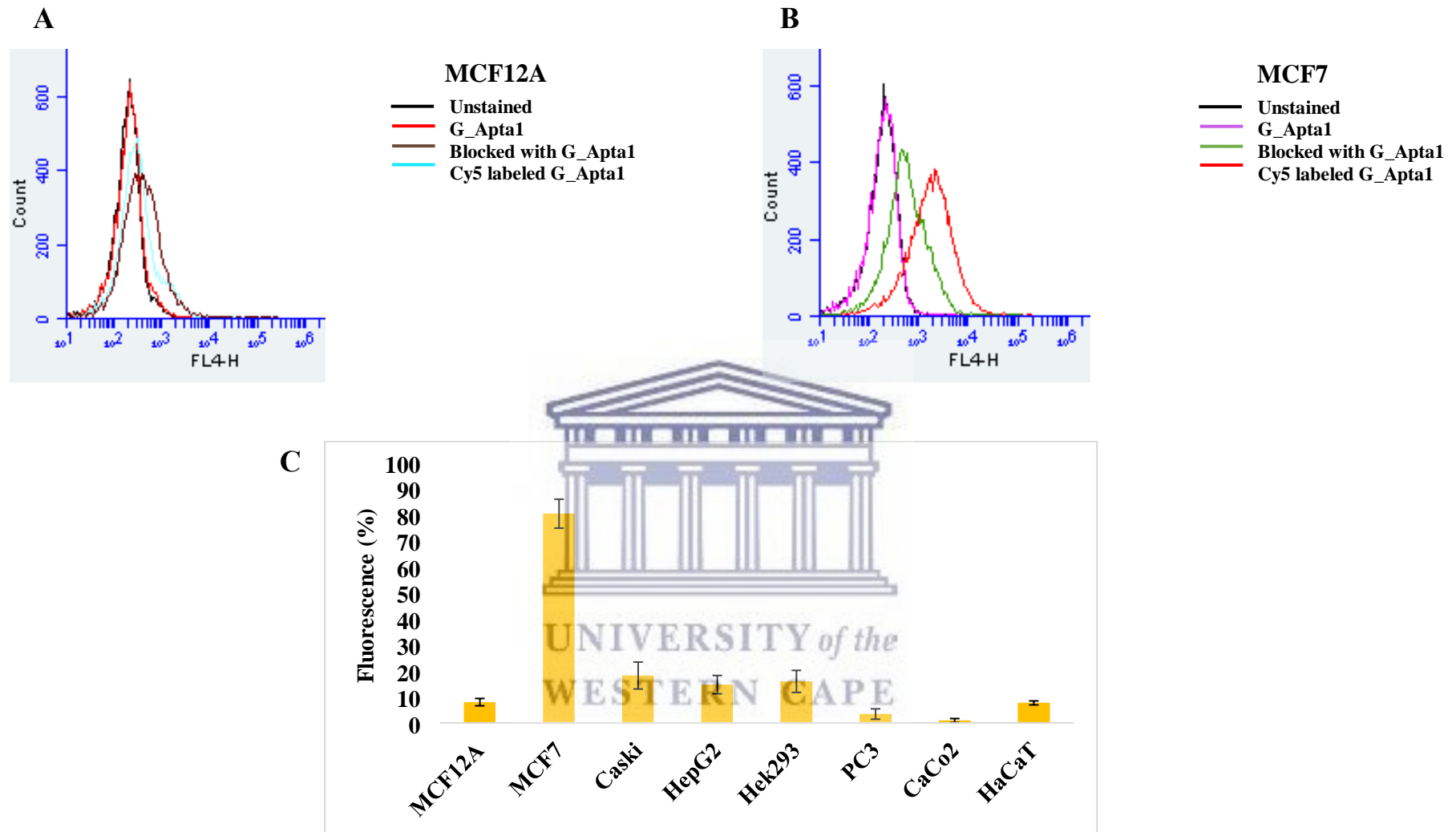


Figure 4.12. Evaluation of G_Apta1 binding to the cell surface GFR α 1 protein, using flow cytometry. **A** and **B** shows a histogram of the control MCF12A and MCF7 cells, respectively, incubated with or without Cy5 labelled G_Apta1 (31.25nM) as shown in the legend. **C**) shows the percentage fluorescence for a panel of human cell lines incubated with Cy5 labelled G_Apta1 (31.25nM). Error bars represent standard deviations from triplicate measurements.

The histogram generated for MCF7 (**Figure 4.12B**) shows the fluorescence intensity of events detected on FL4, for two distinct cell populations as observed by the black and red peaks. FL4, in the BD Accuri™ C6 flow cytometer used in this study, is a detector for fluorophores that have wavelength emissions near infrared; which where Cy5 emits light. The red peak (**Figure 4.12B**), represents the cells that were stained with the Cy5 labelled G_Apta1 (positive population), since the peak formed at a high number of detected events and showed higher fluorescence intensity on FL4. This confirms that G_Apta1 bound to GFR α 1 expressed on the surface of the MCF7 cells. The black peak, represents the cells that were unstained (negative population) and since the peak formed at a low number of detected events, it is indicative of low fluorescence intensity. The weak fluorescence was therefore accounted for as background fluorescence, possibly emitted from intracellular fluorescent components. What was observed is that the cells that were incubated with the unlabelled G_Apta1 had a cell population peak (pink) in the exact region as the unstained cells. Confirming that the fluorescence was in fact coming from the cells.



Furthermore, those cells that were blocked with unlabeled G_Apta1 formed a peak (green) that overlapped between the negative and positive populations. Although it was expected that these cells should only be observed in the negative population, it is possible that the concentration of the unlabeled G_Apta1 was not sufficient to saturate the surface of all the cells or the incubation period may have been too short; consequently, the Cy5 labeled G_Apta1 was able to bind to any GFR α 1 that was not blocked by the unlabeled G_Apta1. It was evident that G_Apta1 demonstrated specificity to the MCF7 cells because the Cy5 labelled G_Apta1 did not bind to MCF12A cells (**Figure 4.12A**). To further interrogate the specificity of G_Apta1, it was incubated with a panel of human cell lines (**Figure 4.12C**), to which end it was very

clearly demonstrated that the interaction between G_Apta1 and GFR α 1 was specific to MCF7 cells.

4.5. Conclusions and future recommendations

The characterization of the six candidate aptamers, developed against GFR α 1, were successfully performed using EMSA, SPR, MST and flow cytometry. The combination of these techniques provided information about the binding ability, binding affinity and specificity of the aptamers to GFR α 1.

In the case of EMSA, some of the limitations identified were that more controls could have been used to substantiate the EMSA findings. These include the use of 1) a positive control, that is, a known protein-aptamer complex to validate the technique; 2) the use of the control proteins (IgG and HSA) used during the counter selection process and 3) the use of an unrelated control aptamer or a scrambled sequence of the aptamers to demonstrate the specificity of the aptamers that bind to GFR α 1. Despite the limitations, EMSA provided insight into the potential binding of all the aptamers to GFR α 1.

An extensive SPR analysis was performed to answer three pertinent questions, 1) Do the candidate aptamers bind GFR α 1? 2) If they do, how well do they bind? Finally, 3) how specific is the interaction between the candidate aptamers and GFR α 1? From the data obtained, it was established that out of the six candidate aptamers, only two (G_Apta1 and 2) bound to GFR α 1, of which G_Apta1 exhibited the highest affinity. At this point, two serum proteins, IgG and HSA, were used as controls. To which end the potential discriminative power of the candidate aptamers were recognized because no cross-reactivity was observed. Since GFR α 1 is a part of a family of proteins, it was possible that some cross-reactivity may occur. Therefore, to

establish how specific the aptamer-protein interaction was, the candidate aptamers were tested against GFR α 2 and GFR α 3. As before, G_Apta1 had the highest affinity and selectivity towards GFR α 1, from the two aptamers that did show evidence of binding. In this case, however, at a high concentration (10 μ M), G_Apta1 also bound to GFR α 2, which could be as a result of GFR α 2 that shares sequence and structural similarities, however, it is more likely that it was due to non-specific interactions as a result of the high concentration of the aptamer.

To further validate the binding and binding affinity, MST assays were performed using G_Apta1 against GFR α 1, 2 and 3. The data obtained corroborated what was established with SPR, that is, the strength of the interaction between G_Apta1 and GFR α 1 and that G_Apta1 also binds GFR α 2, possibly due to sequence and structural similarity. Perhaps, to have eliminated this challenge with cross-reactivity, a better approach would have been to incorporate the GFR α family during the selection process. This way, any sequences that bound the other GFR α proteins would have been eliminated. However, since the ssDNA library consists of approximately 10^{15} molecules, the likelihood of repeating the selection and pipetting the same set of sequences used for SELEX in this study, is very low. Therefore, it would not be the best approach to take at this point, however, the limitation was recognized.

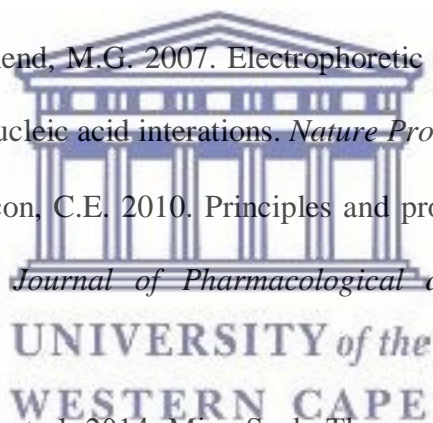
Alternatively, following through on a stepwise truncation study might lead to the establishment of sequences with higher affinity for GFR α 1 and in the process, eliminate those regions within the aptamer sequence that may, specifically, recognize GFR α 2. This may prove to be true as it's often been shown that all the nucleotides in the aptamer sequence is not necessarily involved in the protein-aptamer interaction. Additionally, using the stepwise truncation approach would provide useful information about both the aptamer sequence and structure.

In looking at the application of G_Apta1, using flow cytometry, the binding assay very strongly demonstrated the binding and specificity of G_Apta1 towards GFR α 1 on the surface of MCF7 cells. In conclusion, one high affinity aptamer, namely, G_Apta1 was successfully developed against GFR α 1 and future work would entail investigating the use of G_Apta1 in prognostic applications for the detection of breast cancer.

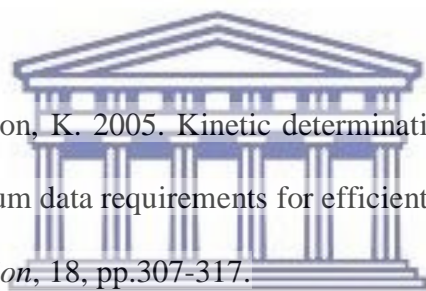


4.6. References

- Akitomi, J. et al. 2011. Valfold: Program for the aptamer truncation process. *Bioinformatics*, 6(7), pp. 38-40.
- Brown, M. and Wittwer, C. 2000. Flow cytometry: Principles and clinical applications in hematology. *Clinical Chemistry*, 46(8), pp. 1221-1229.
- Drah, M. 2015. The Development of Nanotechnology- based Detection Systems for the Diagnosis of Breast Cancer. *Ph.D. Thesis*. University of the Western Cape.
- Hahnefeld, C., Drewianka, S and Herberg, F.W. 2004. Determination of kinetic data using surface plasmon resonance biosensors, *Methods in Molecular Medicine*, 94, pp. 299-320.
- Hellman, L.M. and Friend, M.G. 2007. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature Protocols*, 2(8), pp. 1849-1861.
- Holden, N. S. and Tacon, C.E. 2010. Principles and problems of the electrophoretic mobility shift assay. *Journal of Pharmacological and Toxicological Methods*, 63(2011), pp.7-14
- Jerabek-willemsen, M. et al. 2014. MicroScale Thermophoresis : Interaction analysis and beyond q. *Journal of Molecular Structure*, 1077, pp.101–113. Available at: <http://dx.doi.org/10.1016/j.molstruc.2014.03.009>.
- Jerabek-willemsen, M. et al. 2011. Molecular Interaction Studies Using Microscale Thermophoresis. *Assay and Drug Development Technologies*, pp.342–353.
- Jing, M. and Bowser, M., 2011. A Review of Methods for Measuring Aptamer-Protein Equilibria. *Analytica Chimica Acta*, 686(1–2), pp.9–18.
- Lakhin, A. V, Tarantul, V.Z. and Gening, L. V, 2013. Aptamers : Problems , Solutions and Prospects. *Acta Naturae* , 5(19), pp.34–43.




- Muguruma, H. et al 2000. Sensor chip using a plasma-polymerized film for surface plasmon resonance biosensors: reliable analysis of binding kinetics. *Analytical Sciences*, 16, pp. 347-348.
- Ngcoza, N. 2015. Identification of biomarkers in breast cancer as potential diagnostic and therapeutic agents using a combined in silico and molecular approach. *M.Sc. thesis*. University of the Western Cape.
- Nguyen, H.H. et al. 2015. Surface plasmon resonance: A versatile technique for biosensor applications. *Sensors (Switzerland)*, 15(5), pp.10481–10510.
- Nilsson, A.R., Bruder, D. and Pronk, C.J.H. 2013. Frequency determination of rare populations by flow cytometry: a hematopoietic stem cell prospective. *Cytometry Part A*, 83A, pp.721-727.
- Onell, A. and Andersson, K. 2005. Kinetic determinations of molecular interactions using Biacore - minimum data requirements for efficient experimental design. *Journal of Molecular Recognition*, 18, pp.307-317.
- Plach, M.G. Grasser, K. and Schubert, T. 2017. Microscale thermophoresis as a tool to study protein-peptide interactions in the context of large eukaryotic protein complexes. *Bio-Protocol*, 7(23), pp.1-10.
- Ritzefeld, M. and Sewald, N. 2011. Real-time analysis of specific protein-DNA interactions with surface plasmon resonance. *Journal of Amino Acids*, 2012, pp.1-19.
- Ruscito, A. and DeRosa, M.C., 2016. Small-Molecule Binding Aptamers: Selection Strategies, Characterization, and Applications. *Frontiers in Chemistry*, 4(May), pp.1–14.



UNIVERSITY of the
WESTERN CAPE

Chapter Five: Thesis Overview

Among the top five cancers affecting women in South Africa, breast cancer is the most commonly diagnosed cancer in women and has, therefore, been identified, together with cervical cancer, as a national priority due to increasing incidence rates (CANSA, 2018). Despite ongoing efforts to combat breast cancer, current prognostic and/or therapeutic monitoring methods are limited since very little improvement, in the rate of long term recurrence of breast cancer, has been observed. Considering this, developing novel strategies to detect breast cancer recurrence – at an early onset – is crucial for monitoring the disease and potentially preventing disease progression.



This study proposed the development of aptamers with the objective of applying it in the development of a cost-effective detection method for breast cancer. Serum/blood-based biomarkers are ideal targets for the development of low cost detection assays. Two candidate biomarkers, Unique Ligand Binding Protein 2 (ULBP2) and Glial cell line-derived neurotrophic Factor family Receptor alpha 1 (GFR α 1) were previously identified in the studies by Ngcoza (MSc thesis, UWC) and Drah (PhD thesis, UWC) using bioinformatics and proteomics, respectively. In **Chapter One**, a review is provided in which a strong case for these biomarkers as useful and independent, prognostic biomarkers for breast cancer are demonstrated. Furthermore, this chapter describes how the selection of aptamers against these biomarkers can facilitate the development of cost-effective detection methods. Therefore, the aim of this study was to develop DNA aptamers against putative breast cancer biomarkers for its application in breast cancer prognosis.

In **Chapter Two**, the aim was to express and purify the human recombinant ULBP2 protein for downstream use in aptamer selection. The cloning and large-scale expression of the ecto-domain of ULBP2, in *E.coli*, appeared successful; however, the objective of this chapter was significantly hindered by the solubility problems encountered upon attempting to purify ULBP2. In **Section 2.5**, several reasons and/or limitations for this outcome were addressed; followed by providing the appropriate recommendations for future work. For the purpose of this study, it was best suited to acquire the commercially available human recombinant proteins to proceed with the aptamer selection.

In **Chapter Three**, the aim was to develop DNA aptamers against the human recombinant ULBP2 and GFR α 1 proteins, respectively. A classical *in vitro* SELEX approach, including negative and counter SELEX was employed. In the case of ULBP2, the aptamer selection was unsuccessful as no indication of evolution was observed; instead the SELEX library pool showed a stronger affinity towards the SIMAG-IDA-Nickel beads used to facilitate the selection process. Therefore, it is recommended that the appropriate metal ion/resin used to facilitate the bead-based *in vitro* SELEX should be screened and optimised. Further limitations and recommendations are extensively discussed in **Section 3.4.1** and **Section 3.5**.

In the case of GFR α 1, a successful evolution was observed, of which six potential candidate aptamers were identified, namely, G_Apta1, G_Apta2, G_Apta3, G_Apta9, G_Apta13 and G_Apta15. As a result, **Chapter Four** speaks to the characterization (binding affinity and specificity) of these six aptamer candidates against GFR α 1, using EMSA, SPR, MST and flow cytometry. The combination of these techniques provided information about the binding ability, binding affinity and specificity of the aptamers to GFR α 1. EMSA provided insight into the ability of all the aptamers to interact with GFR α 1, but more controls could have been

included to substantiate the results obtained from this method. The SPR analysis spoke to three pertinent questions, 1) Do the candidate aptamers bind GFR α 1? 2) If they do, how well do they bind? Finally, 3) how specific is the interaction between the candidate aptamers and GFR α 1? And from the data obtained, it was established that out of the six candidate aptamers, only two (G_Apta1 and 2) bound to GFR α 1, of which G_Apta1 exhibited the highest affinity. No cross-reactivity was observed for IgG and HSA, however G_Apta1 showed cross-reactivity to GFR α 2 at high concentrations of the aptamer. This could be as a result of GFR α 2 sharing >50% sequence similarity, however, it may also be due to non-specific interactions as a result of the high concentration of the aptamer.

To further validate the binding affinity and specificity of G_Apta1, MST assays were performed using GFR α 1, 2 and 3. The data obtained corroborated what was established with SPR, that is, G_Apta1 shows cross-reactivity towards GFR α 2, but exhibited the strongest affinity toward GFR α 1. Furthermore, preliminary results indicate K_D values in the nano-molar range. Perhaps, to have eliminated this challenge with cross-reactivity, a better approach would have been to incorporate the GFR α family during the selection process. Alternatively, following through on a stepwise aptamer truncation study might lead to the establishment of sequences with higher affinity for GFR α 1 and in the process, eliminate those regions within the aptamer sequence that may, specifically, recognize GFR α 2.

To strengthen the case for G_Apta1, flow cytometry was used to evaluate 1) the ability of the G_Apta1 to bind to the cell surface GFR α 1 and 2) the specificity of the G_Apta1 to MCF7 cells. The binding assay very strongly demonstrated the binding and specificity of G_Apta1 towards GFR α 1 on the surface of MCF7 cells.

In conclusion, at least one high affinity aptamer, namely, G_Apta1 was successfully developed against GFR α 1 and future work would entail:

- 1) Screening a panel of human breast cancer cell lines, using flow cytometry, to further evaluate the specificity of G_Apta1 to MCF7 cells.
- 2) Performing comparative studies between G_Apta1 and anti-GFR α 1 antibodies, using techniques such as SPR and MST.
- 3) Another route to explore is the ability/specificity of G_Apta1 to bind to breast cancer tissue samples, using IHC.
- 4) It would also be of value to further characterise G_Apta2, using SPR and MST.
- 5) Investigating the use of G_Apta1 in prognostic applications for the detection of breast cancer in patient samples, in platforms such as lateral flow devices, ELONA and biosensors.

