

**DEVELOPING SOMATIC HYPERMUTATION  
AS A PROTEIN ENGINEERING TOOL TO  
STUDY ANGIOPOIETIN BINDING**

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Shikha Sharma

Department of Cardiovascular Sciences

University of Leicester

September 2010

# **Developing Somatic Hypermutation as a Protein Engineering Tool to Study Angiopoietin binding**

Shikha Sharma PhD Thesis

## **Abstract**

The angiopoietin (Ang) and Tie families play an important role in the latter stages of vascular development and in adult vasculature. A variety of studies on Ang1 ligand have provided compelling evidence of its therapeutic potential and along with Ang2 plays a major role in various protective and pathological conditions. Understanding the Ang-Tie interaction by manipulation of the angiopoietins is a very desirable prospect for developing it as a protein therapeutic drug. The aim of this project was to examine the molecular basis of Ang binding using a combination of rational mutagenesis and directed evolution. Directed evolution is a powerful strategy for protein engineering. A new method of directed evolution using mammalian surface display combined with mutagenesis driven by somatic hypermutation (SHM) was investigated for engineering Ang proteins with altered binding characteristics (to Tie1 and Tie2). The Ang receptor binding domains (RBD) were cloned linked to the asialoglycoprotein receptor (ASGPR) transmembrane domain. The fusion protein produced was capable of expression on the extracellular cell surface. Activation-induced (cytidine) deaminase (AID) expressing B cells were used for expression of this fusion protein. SHM driven by AID is capable of mutating highly expressed transgenes; this was used to generate the mutant library. The library was screened for enhanced Tie2 binding affinity and acquired Tie1 binding function by Fluorescence Activated Cell Sorting (FACS). Unstable expression of AID in the cell line and non-specific binding of Tie1 to the cell surface proteins in the binding screen prevented efficient selection of desired mutants. However, Ang2 FReD (fibrinogen-related domain) mutants were generated by SHM and rational mutagenesis. These were analysed and compared to the wild-type Ang2 FReD protein for expression by Western blotting and binding affinity determined by flow cytometry.

## Acknowledgements

I owe my deepest gratitude to Prof. Nick Brindle for giving me the wonderful opportunity to do this PhD. His continuous guidance and encouragement throughout the project made this a great learning experience.

I would also like to thank my internal examiner Dr. Dave Lodwick for his guidance, constructive comments on my thesis and support for my thesis viva defense.

A special thanks to Harprit and Tariq for advice with experiments, Julie and Jonathan for assistance with my cells, Hash for his advice and kindness, Eyad and Kathryn for all the crazy fun times, Nisha for being the most fabulous friend ever and coffee-friend Ankur for assistance with flow cytometry.

I thank everyone on level 2 for all their help and friendship. I would also like to thank everyone in the Department of Cardiovascular Sciences.

Most importantly, I thank my parents for supporting, encouraging and believing in me. I also thank my lovely family especially my brother Dev, my grandparents, my mausijis' and my friends Mobina and Manasvi. An extra special thanks to my husband Sandeep for all the love and support.

*To you I dedicate this, dearest momma and papa.*

# Contents

	<b>Page No</b>
<b>Title page</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>Acknowledgements</b>	<b>iii</b>
<b>Contents</b>	<b>v</b>
<b>List of Abbreviations</b>	<b>xii</b>
<b>List of Figures</b>	<b>xiv</b>
<b>Chapter One</b>	
<b>Introduction</b>	<b>1</b>
1.1 Angiopoietin-Tie System	2
1.1.1 Tie Family	2
1.1.2 Angiopoietin Family	4
1.1.3 Angiopoietin-Tie2 binding interactions	8
1.2 Angiopoietin Variants	14
1.3 Clinical applications	17
1.4 Manipulation of angiopoietins for understanding Ang:Tie interactions	19
1.5 Protein Engineering	20
1.5.1 Rational Design	21
1.5.2 Computational Protein Design	24
1.5.3 Directed protein evolution	26
1.6 Somatic Hypermutation (SHM)	29

1.6.1 Features of SHM	30
1.6.2 SHM as a mutagenesis tool	32
1.7 Expression of Mutant Libraries: Cell Surface Display	36
1.8 Binding and Selection Assay	38
1.9 Hypothesis to be tested	39
<b>Chapter Two</b>	
<b>Materials and Methods</b>	<b>40</b>
General Materials	40
2.1 Polymerase Chain Reaction (PCR)	46
2.2 Restriction Digest	47
2.3 Agarose gel electrophoresis	48
2.4 Purification of PCR and restriction digest products	49
2.5 Dephosphorylation	50
2.6 Ligation	50
2.7 Site Directed Mutagenesis	52
2.8 RNA Extraction	53
2.9 RT PCR	53
2.10 Plasmid preparation	53
2.11 Sequencing	55
2.12 Transformation	56
2.13 Storage of Bacterial Culture	56
2.14 Cell Culture	57
2.15 Transfection	58

2.16 Generating stable cell line	60
2.17 Collection of culture medium and cell lysate	61
2.18 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis	62
2.19 Western Blotting and antibody probing	63
2.20 Immunofluorescence	64
2.21 Magnetic sorting on ASGPR-Ang2FReD expressing stable DT40 cells.	67
2.22 Flow Cytometry	68
2.23 Fluorescence Activated Cell Sorting (FACS)	70
2.24 Controls for Flow Cytometry analysis and Fluorescence Activated Cell Sorting (FACS) of cell populations.	70

### **Chapter Three**

<b>Introduction: Determination of Minimal Receptor Binding Domain in Angiopoietin-1</b>	<b>73</b>
3.1 Cloning of FLAG tagged FReD (fibrinogen domain) in the pSecTag/ FRT/ V5-His TOPO expression vector.	75
3.2 Cloning of FLAG tagged P domain in the pSecTag/FRT/V5-His TOPO expression vector.	78
3.3 Expression of secreted Ang1 FReD and P domains in mammalian cells.	82
3.4 Binding of Tie2 to secreted FReD and P domains.	85
Discussion	88

## **Chapter Four**

<b>Introduction: Rational Mutagenesis of Ang2 Receptor Binding Interface</b>	<b>92</b>
4.1 Rationalisation of mutants and visualisation in PyMOL.	97
4.2 Site directed mutagenesis and cloning to generate mutants Gly428Thr, Ser480Pro and Asn467Ser.	102
4.3 Western blot analysis for mutants Gly428Thr, Ser480Pro and Asn467Ser in HEK293 cell line.	103
4.4 Influence of protein expression levels on $K_d$ determination by flow cytometry.	104
4.5 $K_d$ determination of mutants Gly428Thr, Ser480Pro and Asn467Ser by flow cytometry.	107
Discussion	112

## **Chapter Five**

<b>Introduction: Designing and Cloning of the Cell Surface Display Vector for Expression of Ang1 and Ang2 FReD domains in Ramos Cells</b>	<b>118</b>
5.1 Cloning ASGPR in pcDNA3.1/V5-His-TOPO vector.	119
5.2 Modifications to clone the ASGPR into the pcDNA3.1/V5-His-TOPO vector.	125
5.2.1 Changes in overlap PCR conditions.	125
5.2.2 Changes by site directed mutagenesis.	125
5.3 Designing the ASGPR-pc3.1 surface expression plasmid (pSE2) for commercial synthesis.	126

5.4 Expression of surface display vector ASGPR-pc3.1 (pSE2).	130
5.5 Cloning of FLAG tagged FReD in the pSE2 surface display vector.	134
5.6 Expression of cell bound FReD (Ang1 and Ang2) domains in CHO cells.	136
5.7 Expression of cell bound FReD (Ang1 and Ang2) domains in Ramos cells.	137
5.8 Flow cytometry analysis of Ramos cells transfected with ASGPR linked-FReD domain constructs <i>FAr</i> (Ang1) and <i>F8</i> (Ang2).	140
5.9 Expression of ASGPR linked FReD (Ang1 and Ang2) mRNA in Ramos stable cell lines.	144
5.10 Western analysis of HEK cells transfected with the FReD- ASGPR-pc3.1 constructs.	145
5.11 Immunofluorescence analysis of HEK cells transfected with the ASGPR linked FReD domain (Ang1 and Ang2) constructs.	147
5.12 Detection of FReD (Ang1 and Ang2) mRNA in HEK stable cell lines.	150
Discussion	152

## **Chapter Six**

### **Introduction: Expression of Angiopoietin FReD domain on DT40 Cell**

#### **Surface for Directed Evolution by Somatic Hypermutation and Iterative Selection by Fluorescence Activated Cell Sorting** 158

6.1 Cloning of FLAG tagged ASGPR linked Ang1 and Ang2 FReD (fibrinogen related domain) in the pHypermut2 vector and expression in HEK293 cells.	163
6.2 Expression and localization of Ang2-FReD domain on DT40 cell surface.	168
6.3 Immunofluorescence staining to show binding of Tie2 to cell surface	171

expressed FReD domain in DT40 stable cells.	
6.4 Frequency of mutation in the starting stable populations of DT40 cells expressing Ang2-FReD domain.	175
6.5 Magnetic bead sorting.	176
6.5.1 Selection of high affinity Tie2 binders using magnetic bead sorting.	178
6.5.2 Flow cytometry analysis and comparison of <i>Unsorted</i> population of cells with the magnetically sorted <i>1<sup>st</sup> Sorted</i> , <i>2<sup>nd</sup> Sorted</i> and <i>3<sup>rd</sup> Sorted</i> population of cells.	180
6.5.3 Frequency of mutation in the <i>1<sup>st</sup> sorted</i> population obtained from magnetic sorting.	185
6.6 Selection screen of high affinity binders of Tie2 by Fluorescence Activated Cell Sorting (FACS).	186
6.6.1 AID expression in the ASGPR-Ang2FReD stable cell population.	188
6.6.2 Frequency of mutation in the population of cells used for Fluorescence Activated Cell Sorting (FACS).	189
6.6.3 Determination of <i>Kd</i> for wild type Ang2 FReD domain transfected stable DT40 by Tie2 concentration curve.	190
6.7 FACS selection	191
6.8 Comparison by flow cytometry of FACS sorted population of cells for high affinity binders to Tie2.	195
6.9 Characterisation of mutants <i>10c</i> , <i>FACSD</i> and <i>FACSi</i> obtained by SHM.	201
6.9.1 Western blot analysis for mutants <i>10c</i> , <i>FACSD</i> , <i>FACSi</i> in HEK293 cell line.	202

6.9.2 <i>Kd</i> determination of mutants <i>10c</i> , <i>FACSD</i> , <i>FACSi</i> in HEK293 cell line by flow cytometry.	203
Discussion	208
<b>Chapter Seven</b>	
<b>Introduction: Directed Evolution of Ang2- FReD domain by Somatic Hypermutation to Obtain Tie1 Binding Function</b>	<b>217</b>
7.1 Somatic Hypermutation (SHM) and Fluorescence Activated Cell Sorting (FACS) to evolve Ang2-FReD variants with Tie1 binding function.	218
7.1.1 AID expression in the ASGPR-Ang2FReD stable cell population.	219
7.1.2 FACS sorting for selecting Tie1 binders.	220
7.1.3 Comparison by flow cytometry of FACS sorted population of cells for Tie1 binding.	224
7.2 Analysing non-transfected cells by iterative rounds of FACS selection.	226
Discussion	228
<b>Chapter Eight</b>	
General Discussion	230
<b>Appendix</b>	
Appendix One	240
Appendix Two	242
Appendix Three	246
Appendix Four	248
<b>References</b>	<b>257</b>

## List of Abbreviations

- $A_{260}/A_{280}$  – Ratio of absorbance at the wavelengths of 260 nm and 280 nm
- AID-Activation-Induced (Cytidine) Deaminase
- Ang - Angiopoietin
- ASGPR - Asialoglycoprotein receptor
- BSA - Bovine Serum Albumin
- bp – Base pair
- CHO - Chinese Hamster Ovary
- CIAP - Calf Intestinal Alkaline Phosphatase
- COMP-Ang1- Cartilage Oligomeric Matrix Protein Angiopoietin1
- Cy2-Cyanine 2
- DABCO - 1,4-diazabicyclo[2.2.2]octane
- dH<sub>2</sub>O - distilled water
- DMSO- Dimethyl sulfoxide
- DNA-Deoxyribonucleic acid
- dNTP - Deoxyribonucleotide triphosphate
- DT40- B cell line derived from an avian leukosis virus induced bursal lymphoma in a white leghorn chicken
- DTT - Dithiothreitol
- E.coli*-Escherichia coli
- EDTA – Ethylenediaminetetraacetic acid
- FACS-Fluorescence Activated Cell Sorting
- FCS – Foetal Calf Serum
- FMO-Fluorescence Minus One
- FReD – Fibrinogen domain
- HEK293-Human Embryonic Kidney 293
- Kb - Kilobase
- KCl-Potassium chloride
- kDa – Kilodaltons
- KH<sub>2</sub>PO<sub>4</sub>- Potassium dihydrogen phosphate

LB - Luria Bertani  
LBD-Ligand Binding Domain  
MCS – Multiple Cloning Site  
MgCl<sub>2</sub>- Magnesium Chloride  
MPA- Mycophenolic Acid  
NaCl - Sodium chloride  
NaF-Sodium Fluoride  
NaOH- Sodium hydroxide  
NaPO<sub>4</sub>- Sodium Phosphate  
PBS - Phosphate Buffered Saline  
PCR- Polymerase Chain Reaction  
P domain - P domain  
pmol - Picomoles  
RBD- Receptor Binding Domain  
rpm – Revolutions per minute  
R-PE- R-Phycoerythrin  
RT - Room Temperature  
RTK -Receptor Tyrosine Kinases  
RT-PCR- Reverse transcription polymerase chain reaction  
SB - Sample Buffer  
SDS - Sodium Dodecyl Sulphate  
SHM-Somatic Hypermutation  
SPR- Surface Plasmon Resonance  
TAE - Tris Acetate-EDTA  
TBS - Tris Buffered Saline  
TE - Trypsin-EDTA  
TX-100 – Triton X-100  
UV- Ultraviolet  
V- Volts

# List of Figures

	<b>Page No</b>
<b>Chapter 1</b>	
Figure 1.1: Schematic comparison between human Ang1, Ang2 and Ang4.	5
Table 1.1: A summary of binding studies performed for Ang1, Ang2 and associated truncated and recombinant fragments.	12
Figure 1.2: Schematic outline of a typical directed evolution experiment.	27
<b>Chapter 2</b>	
Table 2.1: Components in a typical PCR reaction.	46
Table 2.2: General cycling parameters of a typical PCR reaction.	46
Table 2.3: General outline of the components of a typical ligation reaction.	51
Table 2.4: General cycling parameters of a site directed mutagenesis reaction.	52
Table 2.5: General outline of components in the resolving and stacking gel used for SDS-PAGE.	62
Figure 2.1: Gating to distinguish dead cells from viable cells prior to obtaining flow cytometric data.	71
Figure 2.2: Non stained cells, used to account for autofluorescence of the cells prior to obtaining flow cytometric data.	71
Figure 2.3: Compensation to correct the detected "spillover" from the emission of the Cy2 fluorochrome into the detector designed to collect the emission from R-phycoerythrin fluorochrome and vice a versa.	72
<b>Chapter 3</b>	
Table 3.1: Description of the constructs cloned in Chapter 3.	75
Figure 3.1: Schematic representation for cloning FLAG tagged FReD (fibrinogen domain) in the pSecTag/FRT/V5-His TOPO expression vector.	76
Figure 3.2: Amplification of the Ang1-FReD domain by PCR.	77
Figure 3.3: PCR screen for Ang1-FReD domain in TOPO-FReD constructs.	78
Figure 3.4 - Schematic representation for cloning FLAG tagged P domain in the pSecTag/FRT/V5-His TOPO expression vector.	79

Figure 3.5: Amplification of the Ang1- P domain by PCR.	80
Figure 3.6: PCR screen for Ang1-P domain in TOPO-P domain constructs.	81
Figure 3.7: Western blot to analyse <i>FReD4 Sec</i> construct for expression of soluble Ang1-FReD protein in mammalian cells.	83
Figure 3.8: Western blot to analyse <i>Pe Sec</i> construct for expression of soluble Ang1-P domain protein in mammalian cells.	84
Figure 3.9: Immunoprecipitation of Ang1 secreted FReD and P domain.	86
Figure 3.10: Immunoprecipitation of Tie2 ectodomain bound to Ang1 FReD and P domains.	87

## Chapter 4

Figure 4.1: Structure of the Tie2-Ang2 complex visualised in the PyMOL imaging software.	93
Figure 4.2: Ribbon representation of the Ang2-Tie2 binding interface.	94
Figure 4.3: Sequence alignment of the Ang1 and Ang2 P domain.	95
Table 4.1: Summary of the rational mutants N467S, S480P and G428T obtained by site directed mutagenesis of Ang2-FReD domain.	97
Figure 4.4: Residues which interact and form H bonds with Ang2-Asn467.	99
Figure 4.5: Residues which interact and form H bonds with Ang2-Ser480.	100
Figure 4.6: Residues which interact and form H bonds with Ang2-Gly428.	102
Figure 4.7: Western blot analysis of mutants, Gly428Thr, Ser480Pro and Asn467Ser in HEK293.	103
Table 4.2: Three sets of WT Ang2 (WT A, WT B and WT C) transfected with varying transfection levels as determined by percentage transfection and Geometric mean (Gm) for FLAG tag expression.	105
Figure 4.8: Tie2 concentration curve to show influence of WT Ang2-FReD protein expression levels on $K_d$ determination in HEK293 cell line by flow cytometry.	106
Figure 4.9: Representative Tie2 concentration curve for WT Ang2-FReD domain and mutant Asn467Ser in HEK cell line.	108
Figure 4.10: Representative Tie2 concentration curve for WT Ang2-FReD	109

domain and mutant Ser480Pro in HEK cell line.

Figure 4.11: Tie2 concentration curve for WT Ang2-FReD domain and mutant Gly428Thr in HEK cell line. 111

## **Chapter 5**

Figure 5.1: Schematic representation showing the key features of the surface display vector (pSE1). 120

Figure 5.2: Schematic representation of the first stage of overlap PCR to amplify transmembrane domain of asialoglycoprotein receptor (ASGPR). 121

Figure 5.3: The sequence of asialoglycoprotein receptor (ASGPR) transmembrane domain amplified by overlap PCR. 122

Figure 5.4: Schematic representation of the cloning strategy for the ASGPR fragment in the pcDNA3.1V5-His-TOPO vector. 123

Figure 5.5: Amplification of the ASGPR DNA obtained by overlap PCR. 123

Figure 5.6: PCR screen for ASGPR-pc3.1 constructs (pSE1). 124

Figure 5.7: The sequence of asialoglycoprotein receptor (ASGPR) transmembrane domain designed for commercial synthesis into the pc3.1 mammalian expression vector (pSE2). 127

Figure 5.8: Incorporation of EagI restriction enzyme site in linker A of the surface display vector (pSE2). 128

Figure 5.9: Incorporation of NotI restriction enzyme site in linker B of the surface display vector. 129

Figure 5.10: Expression of the surface display vector pSE2. 130

Figure 5.11: Localization of expressed ASGPR on CHO cell surface by immunofluorescence staining. 132

Figure 5.12: Schematic representation for cloning FReD (fibrinogen domain) in the pSE2 surface display expression vector. 134

Figure 5.13: Amplification of the Ang1 and Ang2 FReD domains for cloning into pSE2 vector. 135

Figure 5.14: Expression of the ASGPR-FReD domains in mammalian (CHO) cells. 136

Figure 5.15: Expression of the ASGPR lined FReD domains in transiently transfected Ramos cells.	138
Figure 5.16: Expression of the ASGPR lined FReD domains in stably transfected Ramos cells.	139
Figure 5.17: Flow cytometry profiles of Ramos cells transiently transfected with Ang1 and Ang2 FReD-ASGPR constructs ( <i>FAr</i> , <i>F8</i> ).	141
Figure 5.18: Flow cytometry profiles of Ramos cells stably transfected with Ang1 and Ang2 FReD-ASGPR constructs ( <i>FAr</i> , <i>F8</i> ).	143
Figure 5.19: Detection of ASGPR linked FReD (Ang1 and Ang2) mRNA in Ramos stable cells.	144
Figure 5.20: Expression of the ASGPR-FReD domains in transiently transfected HEK293 cells.	146
Figure 5.21: Immunofluorescent localization of cell bound ASGPR-FReD domains on stable HEK cell surface.	149
Figure 5.22: To detect the expression of the FReD-ASGPR (Ang1 and Ang2) mRNA in stable HEK cell line.	151
Figure 5.23: Prediction of net charge and pI values for the ASGPR anchor and the ASGPR-Ang1FReD and ASGPR-Ang2FReD fusion proteins.	155
 <b>Chapter 6</b>	
Figure 6.1: Schematic map of pHypermut2 vector.	160
Figure 6.2: Physical map of the rearranged Ig light chain locus, the pHypermut2-Ang2FReD construct and the rearranged locus after targeted integration of the Ang2-FReD domain coding sequence into the DT40 Ig light chain locus.	161
Table 6.1: List of constructs used in Chapter 6.	162
Figure 6.3: Schematic representation for cloning FLAG tagged FReD (fibrinogen domain) and the ASGPR-surface anchor protein in the pHypermut2 vector.	164
Figure 6.4: Colony PCR screen for pHypermut2-ASGPR-FReD constructs.	166
Figure 6.5: Protein expression of the Ang1-FReD-pHypermut2 and Ang2-	167

FRcD-pHypermut constructs in HEK293 cell line.	
Figure 6.6: Localization of cell bound ASGPR-FRcD (Ang2) domain on stable DT40 cell surface by immunofluorescence.	170
Figure 6.7: Binding of surface expressed ASGPR-FRcD (Ang2) domain to Tie2 ectodomain by immunofluorescence.	172
Figure 6.8: Non-transfected DT40 cells served as a negative control against the transfected cells which demonstrated Tie2 binding.	173
Figure 6.9: Schematic representation of magnetic sorting approach to enrich for high affinity binders from a stable DT40 cell population expressing Ang2 FRcD variants.	177
Table 6.2: Summary of the rounds of magnetic affinity selection, indicating the number of cells put through each round and the percentage of cells recovered after each round.	179
Figure 6.10: Histograms overlay of the various magnetically sorted populations of cells, indicating the protein expression level detected the Cy2 (green) fluorescence.	181
Table 6.3: Tie2 concentrations used for analysing, <i>Unsorted</i> , <i>1<sup>st</sup> Sorted</i> , <i>2<sup>nd</sup> Sorted</i> and <i>3<sup>rd</sup> Sorted</i> population of cells.	182
Figure 6.11: Histogram overlaid for comparison of Tie2 binding (red fluorescence) of various magnetically sorted population of cells.	184
Figure 6.12: Schematic representation of the iterative rounds of FACS selection for Tie2 binding affinity.	187
Figure 6.13: Detection of AID mRNA in Ang2-FRcD expressing stable DT40 cells.	189
Figure 6.14: Concentration curve to determine <i>K<sub>d</sub></i> of WT Ang2-FRcD domain (in DT40 cell line) for binding Tie2.	191
Figure 6.15: Representative dot plots showing the population of cells sorted in the 1 <sup>st</sup> round of FACS for Tie2 binding.	193
Table 6.4: Description of the iterative rounds of selection, for high affinity binders to Tie2 by Fluorescence Activated Cell Sorting (FACS).	194
Figure 6.16: Comparison of population of cells sorted by FACS for Tie2	197

binding.	
Figure 6.17: Contour plot display of flow cytometry data for <i>Unsorted</i> , <i>A5</i> and <i>A7c</i> cells populations stained at 0.1nM Tie2.	199
Table 6.5: List of mutant constructs used in Chapter 6.	201
Figure 6.18: Western blot analysis for mutants, <i>10c</i> , <i>FACSD</i> , <i>FACSi</i> in HEK293.	203
Figure 6.19: Representative Tie2 concentration curve for WT Ang2-FReD domain and mutant <i>10c</i> in HEK293 cell line.	205
Figure 6.20: Representative Tie2 concentration curve for WT Ang2-FReD domain and mutant <i>FACSD</i> in HEK293 cell line.	206
Figure 6.21: Histogram overlay of the <i>FACSi</i> transfected cells and the non-transfected HEK293 cells.	207
Table 6.6: Summary the mutants ( <i>Hm5</i> , <i>10c</i> , <i>FACSD</i> , <i>FACSi</i> ) obtained by AID induced SHM of Ang2-FReD domain.	212
Figure 6.22: Molecular visualization of the <i>Hm5</i> and <i>10c</i> mutant in PyMOL.	214
Figure 6.23: Molecular visualization of the <i>FACSD</i> and <i>FACSi</i> mutant in PyMOL.	215
 <b>Chapter 7</b>	
Figure 7.1: Schematic representation of the iterative rounds of FACS for Tie1 binding.	219
Figure 7.2: Detection of AID mRNA in Ang2-FReD expressing stable DT40 cells prior to selection for Tie1 binding.	220
Figure 7.3: Representative dot plot showing the population of cells sorted in the 1 <sup>st</sup> round of FACS for Tie1 binding.	222
Table 7.1: Description of the iterative rounds of FACS to select Tie1 binders.	223
Figure 7.4: Contour plots of flow cytometry data showing enrichment of a Tie1 binding population.	225
Figure 7.5: Dot plots showing flow cytometry analyses of non-transfected cells after iterative selection for Tie1 binding.	227



# Chapter 1

## Introduction

The vasculature is one of the earliest organs to develop in the mammalian body (Rossant, Howard 2002). Vascular development occurs by formation of blood vessels during embryogenesis called vasculogenesis followed by angiogenesis in which the primary vasculature is organized to form the mature vasculature (Drake 2003).

Regulation of angiogenesis at the spatial and temporal level is crucial as excessive angiogenesis is associated with cancer, retinopathies, hemangioma and chronic inflammatory disorders. Inadequate angiogenesis results in impaired wound healing of burns, impaired embryonic development, impaired organ repair, myocardial or critical limb ischemia (Presta et al. 2005, Milkiewicz et al. 2006).

The angiopoietin and Tie families play an important role in the latter stages of vascular development and in adult vasculature. Signalling through Tie2 appears to depend on the balance between Ang1 and Ang2. The Ang:Tie system is crucial in several clinical contexts, thus understanding and manipulation of Ang protein is desirable for developing it as a protein therapeutic drug.

## **1.1 Angiopoietin-Tie System**

Angiopoietin (Ang) family of growth factors are involved in blood vessel formation during developmental and pathological angiogenesis and its significance has been demonstrated by numerous transgenic mouse models (Thurston et al. 2000, Nykanen et al. 2003, Baffert et al. 2006). The Ang–Tie ligand–receptor system plays a primary role in development in adult vasculature, where they control remodelling and stabilization of vessels and regulating vascular integrity and quiescence. Besides its role in angiogenesis, it is an important regulator of diseases such as inflammation and tumour progression (Fiedler, Augustin 2006, Thomas, Augustin 2009).

### **1.1.1 Tie Family**

The members of the Tie family were identified in the early nineteen nineties. The family comprises of two members; Tie 1 whose signalling pathways are not well understood and a better characterised Tie 2 receptor also known as Tek (tunica interna endothelial cell kinase) in mouse. Tie1 and Tie2 are vascular and lymphatic endothelial cell-specific receptors with similar molecular weight of approximately 135 and 150 kDa, respectively (Dumont et al. 1992, Partanen et al. 1992, Sato et al. 1993).

Tie-1 and Tie-2 display a high degree of similarity in their basic structural features. The protein sequence reveals that the first methionine residue is followed by a stretch of hydrophobic sequence about 20aa long which may encode the signal peptide sequence. The extracellular region of both Tie receptors has N-linked glycosylation sites and a hydrophobic site (Sato et al. 1993). The extracellular domain of Tie receptors consists of a globular head domain and a short rod-like stalk that probably forms a spacer between the cell surface and the angiopoietin-binding site at the receptor N-terminus (Macdonald et al. 2006). Mutational analysis demonstrates that the head

domain consists of the three immunoglobulin-like (Ig) domains and the three epidermal growth factor-like (EGF) modules and that the stalk is formed by the three  $\beta$ -sheet fibronectin type III repeats adjacent to the transmembrane domain (Macdonald et al. 2006). The crystal structure of the three Ig and three EGF domains suggests that the extracellular domain can only be stable upon formation of the compact and rigid three-dimensional architecture (Barton et al. 2006). The intracellular domain contains a tyrosine kinase domain interrupted by a kinase insert and a carboxy-terminal tail (Partanen et al. 1992). The amino acid sequence identity between Tie1 and Tie2 is 76%, within the intracellular domain primarily contributing to the sequence identity (Fiedler et al. 2003) and 33% similarity with the extracellular domain (Schnurch, Risau 1993).

The Tie2 receptor is found to interact with ligand members of the angiopoietin family (discussed later). The ligand-binding region of Tie2 comprises of amino-terminal of the receptor ectodomain encoding the extracellular Ig1 and Ig2 domains and the EGF regions (Macdonald et al. 2006). The crystal structure reveals the Ig2 domain and EGF3 region to be most important of the three Ig and three EGF regions (Barton et al. 2006). Angiopoietins have been shown to bind the Tie2 receptor causing receptor dimerization and activation by autophosphorylation which triggers intracellular signalling pathways. On ligand binding Tie2 inhibits endothelial apoptosis and inflammation in addition to promoting migration and stability. Tie2 gene disruption has primarily been associated with defective remodelling of the vasculature (Puri et al. 1995, Sato et al. 1995).

No ligands have been identified which bind directly to Tie1 and hence the cellular pathways are not well understood. Studies in mice lacking the Tie1 functional gene showed that Tie1 promotes survival and maintenance of the microvasculature in the later stages of angiogenesis (Puri et al. 1995). Full length Tie1 is thought to

heterodimerize with Tie2 and is able to phosphorylate via Tie-2 transphosphorylation (Marron et al. 2000). Proteolytically cleaved Tie1 regulates Tie2 ligand binding and activation, thus acting as a regulator in Ang-Tie2 signalling (Marron et al. 2007).

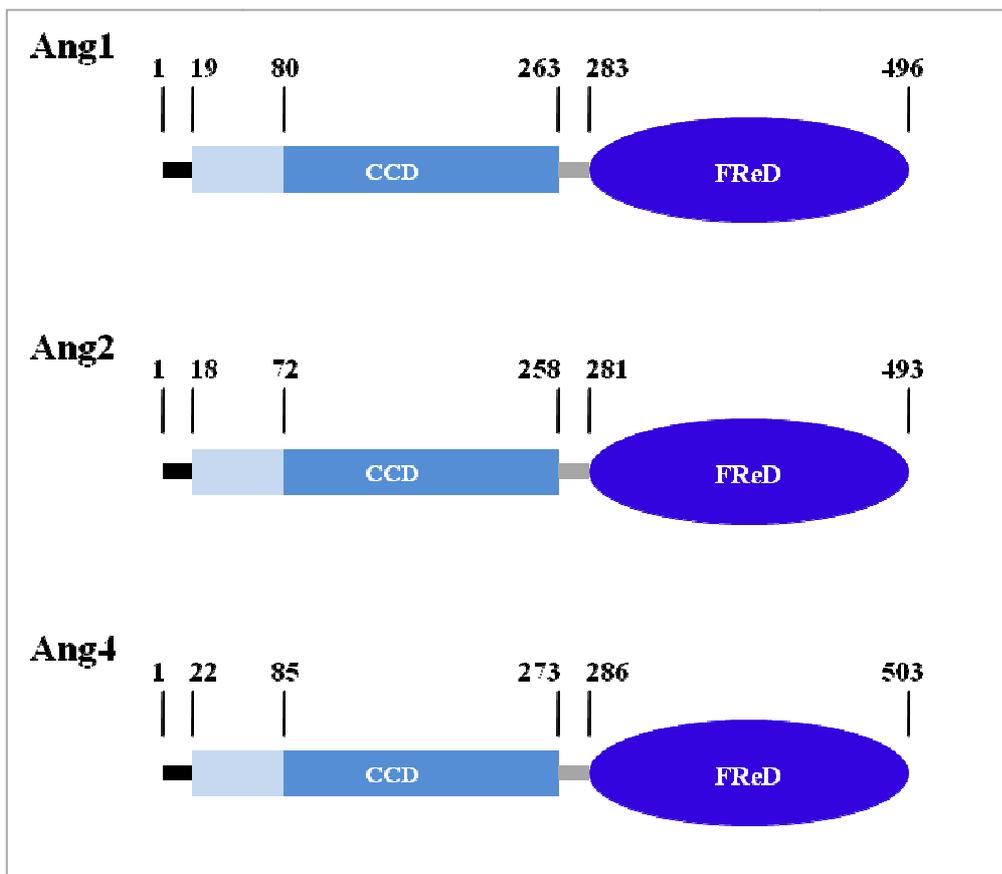
### **1.1.2 Angiopoietin Family**

The angiopoietins are identified as a family of ligands for Tie2 (Davis et al. 1996, Maisonpierre et al. 1997). This family of growth factors comprises of four members, namely, angiopoietin-1 (Ang1), angiopoietin-2 (Ang2), angiopoietin-3 (Ang3) and angiopoietin-4 (Ang4). Among these members, Ang1 and Ang2 are characterized in most detail. Ang3 is the mouse ortholog of human Ang4 (Valenzuela et al. 1999).

Ang1 is constitutively expressed by many different cell types such as pericytes, smooth muscle cells, fibroblasts and some tumour cells (Stratmann, Risau & Plate 1998, Suri et al. 1996, Zagzag et al. 1999). In contrast, Ang2 is almost exclusively expressed by endothelial cells and stored in Weibel-Palade bodies (WPB), and is also detectable in Kaposi's sarcoma cells and in Müller cells in the retina (Stratmann, Risau & Plate 1998, Yao et al. 2007). Ang2 mRNA is almost undetectable in the quiescent vasculature; however, it is induced dramatically at sites of endothelial-cell activation. Ang2 expression is induced by various cytokines, including vascular endothelial growth factor (VEGF) and fibroblast growth factor (FGF)-2, and by microenvironmental factors (e.g. hypoxia) (Fiedler, Augustin 2006, Fiedler et al. 2004, Pichiule, Chavez & LaManna 2004).

The four angiopoietins have a highly conserved structure, which consists of a small amino terminal region responsible for angiopoietin superclustering, followed by a coiled-coil domain coded by residues 100-280 responsible for dimerization. The coiled-

coil domain is separated from the fibrinogen domain by a short linker. The fibrinogen domain encoded by residues 280-498 functions as the receptor recognition and binding region. The fibrinogen domain consists of a variable P domain located near a calcium binding site. The P domain may serve as a ligand-recognition region (Figure 1.1) (Barton, Tzvetkova & Nikolov 2005).



**Figure 1.1: Schematic comparison between human Ang1, Ang2 and Ang4.** The amino terminal region is involved in superclustering of the angiopoietins. The coiled coil domain (CCD) lies between 70 to 270 amino acids across all the angiopoietins followed by the short linker and then the fibrinogen-related domain (FReD) approximately 200 amino acid long.

The first 360 amino acids (Ig-like domain plus EGF-like repeats) of the Tie2 receptor are necessary and sufficient to bind both Ang1 and Ang2, which suggests that differential receptor binding is not likely to be responsible for the different functions of

Ang1 and Ang2 (Fiedler et al. 2003). The N-terminal two immunoglobulin-like domains of Tie2 harbour the angiopoietin-binding site (Macdonald et al. 2006).

### **Ang1-Tie2 signalling**

Ang1 activates Tie2 inducing rapid receptor auto-phosphorylation resulting in endothelial cell migration and maturation (Fiedler et al. 2003). Ang-1-mediated Tie2 phosphorylation signals primarily through the protein kinase B (PKB)–Akt pathway that transduces survival signals (DeBusk, Hallahan & Lin 2004). Ang-1-mediated PI 3-kinase activation results in the phosphorylation and activation of Akt. Akt signalling inactivates the forkhead transcription factor FKHR-1, which is a potent inducer of Ang-2 expression (Daly et al. 2004). Thus, Ang1-mediated PKB–Akt signalling directly inhibits endothelial-cell apoptosis and prevents activation of the endothelium by inhibiting Ang2 expression and secretion. Tie2 activation also results in the recruitment of A20-binding inhibitor of nuclear factor (NF)- $\kappa$ B (ABIN)-2, which inhibits the NF- $\kappa$ B pathway (Hughes, Marron & Brindle 2003). This protects endothelial cells from apoptosis and inhibits inflammatory responses (Tadros et al. 2003). Thus, Tie2 phosphorylation and signalling involves several signalling pathways that are collectively anti-apoptotic and maintain the quiescent state of the resting endothelium.

### **Ang2-Tie2 signalling**

In spite of showing approximately 60% identity to Ang1 and binding to Tie2 with the same affinity as Ang1, Ang2 has a context dependant antagonist effect on Tie2 (Procopio et al. 1999, Tsigkos, Koutsilieris & Papapetropoulos 2003). It acts as a facilitator of vascular morphogenesis and remodelling, acting proangiogenically in the presence of angiogenic stimulators, such as VEGF, and antiangiogenically in the

absence of proangiogenic activity (Fiedler et al. 2003). Ang2 is released from endothelial storage pools (Weibel–Palade bodies) and binds to Tie2, interfering negatively with Ang1-mediated Tie2 signalling resulting in destabilization, thereby rendering the endothelium responsive to stimulation by inflammatory and angiogenic cytokines (Fiedler, Augustin 2006).

Furthermore, Ang2 overexpression results in a phenotype similar to that of Ang1 or Tie2 deletion transgenic mice and results in embryonic lethality. Ang2 is able to activate Tie2 only when present in high concentrations or for a sufficient period. In contrast to the lethal embryonic phenotype of Ang2-transgenic mice, Ang2-deficient mice develop normally. These mice seem phenotypically normal at birth but die within 14 days as a consequence of chylous ascites (on a C129 genetic background) or develop normally throughout adulthood (on a C57/Bl6 genetic background) (Gale et al. 2002). Thus, these findings suggest that Ang2 is dispensable for proper embryonic development but, strong systemic Ang2 elevation is potentially dangerous. The genetic data have also strongly established Ang2 as the functional antagonist of the constitutively functioning Ang1–Tie2 axis (Fiedler, Augustin 2006). *In vivo* Ang2 has the potential to allow vessel growth or regression (Brindle, Saharinen & Alitalo 2006). Ang3 and Ang4 function both as species-specific agonists and antagonists of Tie2 (Lee et al. 2004). The contextuality of the Ang2 action has been conferred to several factors like cell type specific effects, the degree of endothelial confluence, the duration of Ang2 stimulation, concentration-dependent effects, as well as the presence of co-receptors such as Tie1 (Daly et al. 2006, Kim et al. 2000, Teichert-Kuliszewska et al. 2001).

### 1.1.3 Angiopoietin-Tie2 binding interactions

The Angiopoietin-Tie2 interactions can be described and distinguished at four levels: the nature of the binding interface determined from mutational studies, the molecular structure determined by crystallography studies, the oligomerization states determined by rotary metal-shadowing transmission electron microscopy (RMSTEM) or SDS-PAGE and binding affinities determined by ELISA or Surface Plasmon Resonance (SPR).

The binding site of Ang1 and Ang2 in the extracellular domain of the Tie2 receptor spans residues 1-360 encoding for the Ig-like domains and EGF-like repeats of the receptor (Fiedler et al. 2003). Binding analysis of truncated Tie2 domain mutants revealed that neither the Ig domains nor EGF repeats alone are capable of effectively binding the angiopoietins (Fiedler et al. 2003). Crystal structure of Ang2-Tie2 complex reveals that Ang2 P domain binds at the tip of the arrowhead-shaped Tie2, interacting with only the Ig2 domain (residues 122-209) of Tie2. Consequently, the Ang2-Tie2 complex has an elongated shape, with overall dimensions of 130 × 65 × 50 Å (Barton et al. 2006). The N terminus of Tie2, anchored on the extracellular membrane of the Tie2-expressing cell, connects to the C terminus of Ang2, to form the elongated Ang2-Tie2 complex (Barton et al. 2006).

The crystal structure of the Ang2-Tie2 complex reveals no major conformational changes or domain rearrangements on ligand binding. The recognition of Tie2 by its Ang2 ligand proceeds by a lock-and-key mechanism, where molecular surfaces with the 'best-fit' in shape and in chemical nature recognize and bind without marked rearrangements in either binding partner. Thus the Tie2-Ang2 interaction is analogous to the antibody-antigen interaction (Barton et al. 2006). Enthalpic forces drive the

binding reaction, with the vast majority of the interacting residues being engaged in either hydrophobic, van der Waals contacts or hydrogen bonds (Barton et al. 2006). The binding of the ligands to Tie2 is mediated by the fibrinogen-related domain of the ligand. The precise residues for Ang-1/Tie2 interaction are likely to be in the carboxy-terminal half of the fibrinogen domain termed as the P domain as determined by the crystal structure of Ang2 (Brindle, Saharinen & Alitalo 2006). The interacting regions of Ang2 and Tie2, in the absence of the Ang2 coiled-coil and superclustering motifs, bind each other with a 1:1 stoichiometry and the crystal structure of their complex reveals a heterodimeric ligand-receptor assembly (Barton et al. 2006, Barton, Tzvetkova & Nikolov 2005). Similarly the binding stoichiometry of Ang1 receptor binding domain (RBD) and Tie2-Fc in *in vitro* studies is also predicted to be 1:1 (Davis et al. 2003).

The structure of the Ang2–Tie2 complex predicts that Ang1 would bind Tie2 in a similar way as Ang2, indicated by high conservation of the receptor binding domain between both the angiopoietins. Six of the thirteen contact residues are conserved between Ang2 and Ang1 (Barton et al. 2006). In addition, two residues contain conservative substitutions in Ang1 (I434M and F469L), and two others (K432N and Y475H) are also likely to have little or no effect on the ligand-receptor interactions (Barton et al. 2006). Thus, there are only three Tie2-contacting residues different between Ang1 and Ang2 (N468G, S417I and S480P) that might affect the ligand-receptor interactions (Barton et al. 2006). Also on mutation of crucial Tie2 residues (F161A/S164E) both Ang1 and Ang2 were unable to bind or activate the mutated receptors, suggesting that the ligands bind Tie2 in very similar manners (Barton et al. 2006).

Homo-oligomeric ligand binding has the potential to bridge receptor molecules, resulting in dimerization/multimerization of the receptors, thus activating the receptor tyrosine kinases (Heldin et al. 1989). It has been demonstrated that both soluble Ang1 and Ang2 exist in an array of disulfide-linked homo-oligomeric complexes, which may be an important determinant in modulating Tie2 activity (Procopio et al. 1999). The coiled-coil domain (CCD) but not the fibrinogen-related domain (FReD) contributes to distinct oligomerization patterns of native Ang1 and Ang2. The CCD is sufficient to mediate formation of disulfide-linked homo-oligomers (Procopio et al. 1999).

Native Ang1 exists as trimeric, tetrameric, and pentameric homo-oligomers that cluster into multimeric structures via cysteines within the amino-terminal motif. Ang1 must present as a tetramer or higher-order structure to stimulate the receptor (Kim et al. 2005, Cho et al. 2004a). However, Tie2 transfected into non-endothelial (fibroblast) cells can be activated by dimeric Ang1, suggesting that co-receptors or regulators present in endothelial cells require the ligand to assume higher-order multimers to be activating (Davis et al. 2003). Native Ang2 protein mainly forms disulfide-linked dimers (Procopio et al. 1999, Kim et al. 2005, Kim et al. 2009).

Recombinant Ang1 exists as heterogeneous multimers made up of basic trimeric, tetrameric, or pentameric oligomers, whereas Ang2 exists mainly as trimeric, tetrameric, or pentameric oligomers with only a few multimers (Procopio et al. 1999, Kim et al. 2005). Thus it is probable that the inability of Ang-2 to form higher order multimers is a factor which limits it from activating Tie2 but promotes antagonistic action (Brindle, Saharinen & Alitalo 2006). Contextuality of ligand presentation to the Tie-2 receptor may cause differential activities and may be controlled at the level of the target cell, presence of specific co-receptor or modulators or differential oligomerizations states.

Tie-1 and VE-PTP (Vascular Endothelial Phospho-Tyrosine Phosphatase) have been proposed as candidate molecules to serve as modulators of the angiopoietin/Tie-2 signalling (Fiedler et al. 2003). Thus a combinative effect of differential target cell properties and differential presentation of the angiopoietins to the target cells and not the differences in ligand binding to the receptor may be responsible for the differential effects of the angiopoietins (Fiedler et al. 2003).

The affinities of Ang1 and Ang2 examined *in vitro* to the extracellular domain of soluble Tie2 were reported to be the same in early studies (Davis et al. 1996, Maisonpierre et al. 1997). Davis S et al. (1996) showed that the binding of Ang1 was specific for Tie2 with a binding affinity ( $K_d$ ) of  $\sim 3.7$ nM (Davis et al. 1996). Maisonpierre et al. (1997) determined the binding affinity of Ang1\* variant (which has a mutation at Cys<sup>245</sup> and modified NH<sub>2</sub>-terminus in Ang1) and Ang2. The  $K_d$  for both was similar at  $\sim 3$ nM (Maisonpierre et al. 1997). Since then several truncated fragments, recombinant and native forms of the angiopoietins have been examined by binding to Tie2 receptor using a varied range of methods, including ELISA, SPR, SDS-PAGE. A summary of all binding studies performed with angiopoietins is presented in Table 1.1. Recently, Yuan et al. (2009) compared the binding affinity of Ang1 and Ang2 and found that Ang2 binds Tie2 with  $\sim 20$ -fold-lower affinity than does Ang1. The concentrations of ligands bound to achieve 50% saturation were 106ng/ml for Ang1 and 2,018ng/ml for Ang2. They also demonstrated by competitive binding assays that Ang2 binds to Tie2 with a lower affinity compared with Ang1 (Yuan et al. 2009). The difference in affinities may account for different action of Ang1 and Ang2 (Yuan et al. 2009).

Table 1.1: A summary of binding studies performed for Ang1, Ang2 and associated truncated and recombinant fragments.

Paper	Methods	Results														
(Davis et al. 1996)	<ul style="list-style-type: none"> <li>To measure binding affinity of recombinant Ang1 immobilised on nitrocellulose paper.</li> <li>Incubated with varying amounts of Tie2-Fc. Detected I<sup>25</sup> radio-labelled antibody.</li> </ul>	<ul style="list-style-type: none"> <li>No binding of Ang1 to Tie1-Fc.</li> <li>Binding affinity of Ang1 to Tie2-Fc.</li> <li><math>Kd \sim 3.7\text{nM}</math> by Scatchard Analysis</li> </ul>														
(Maisonpierre et al. 1997)	<ul style="list-style-type: none"> <li>Tie2-Fc fusion and Tie1-Fc fusion immobilised to BIAcore sensor chips.</li> <li>Measure binding of purified recombinant Ang1* (variant of Ang1, see Davis et al. 1996) and Ang2</li> </ul>	<ul style="list-style-type: none"> <li>No binding of Ang1* and Ang2 to Tie1-Fc.</li> <li>Similar binding affinity of Ang1* and Ang2 to Tie2-Fc.</li> <li><math>Kd \sim 3\text{nM}</math> by Scatchard Analysis</li> </ul>														
(Fiedler et al. 2003)	<ul style="list-style-type: none"> <li>Binding of immobilised myc-tagged human Ang1 and Ang2 to truncated soluble Tie2 (sTie2) ectodomains by Elisa.</li> </ul>	<table border="1"> <thead> <tr> <th>Variant</th> <th>IC<sub>50</sub> of Ang-1</th> <th>IC<sub>50</sub> of Ang-2</th> </tr> </thead> <tbody> <tr> <td>sTie2-Fc</td> <td>2.2 pmol +/- 1.1</td> <td>3.3 pmol +/- 1.1</td> </tr> <tr> <td>sTie2-(1-440)</td> <td>9.1 pmol +/- 1.3</td> <td>27.5 pmol +/- 1.4</td> </tr> <tr> <td>sTie2-(1-360)</td> <td>4.2 pmol +/- 1.3</td> <td>3.8 pmol +/- 1.2</td> </tr> </tbody> </table>	Variant	IC <sub>50</sub> of Ang-1	IC <sub>50</sub> of Ang-2	sTie2-Fc	2.2 pmol +/- 1.1	3.3 pmol +/- 1.1	sTie2-(1-440)	9.1 pmol +/- 1.3	27.5 pmol +/- 1.4	sTie2-(1-360)	4.2 pmol +/- 1.3	3.8 pmol +/- 1.2		
Variant	IC <sub>50</sub> of Ang-1	IC <sub>50</sub> of Ang-2														
sTie2-Fc	2.2 pmol +/- 1.1	3.3 pmol +/- 1.1														
sTie2-(1-440)	9.1 pmol +/- 1.3	27.5 pmol +/- 1.4														
sTie2-(1-360)	4.2 pmol +/- 1.3	3.8 pmol +/- 1.2														
(Davis et al. 2003)	<ul style="list-style-type: none"> <li>Tie2-Fc fusion immobilised to BIAcore sensor chips.</li> <li>Varying amounts of Ang1RBD (mono, di, tetra, multimers) used to measure binding.</li> </ul>	<table border="1"> <thead> <tr> <th>Variant domains</th> <th>IC<sub>50</sub></th> </tr> </thead> <tbody> <tr> <td>Ang1-N<sub>1</sub>-C<sub>1</sub></td> <td>-</td> </tr> <tr> <td>Ang1-F<sub>1</sub> (monomer)</td> <td>435nM</td> </tr> <tr> <td>Ang1 - C<sub>1</sub>-F<sub>1</sub> (dimer)</td> <td>94nM</td> </tr> <tr> <td>Ang1 (tetramer+)</td> <td>1.5nM</td> </tr> <tr> <td>Ang1-F<sub>1</sub>-F<sub>c</sub> (dimer)</td> <td>71nM</td> </tr> <tr> <td>Ang1-F<sub>1</sub>-F<sub>c</sub> -F<sub>1</sub> (tetramer)</td> <td>9.7nM</td> </tr> </tbody> </table>	Variant domains	IC <sub>50</sub>	Ang1-N <sub>1</sub> -C <sub>1</sub>	-	Ang1-F <sub>1</sub> (monomer)	435nM	Ang1 - C <sub>1</sub> -F <sub>1</sub> (dimer)	94nM	Ang1 (tetramer+)	1.5nM	Ang1-F <sub>1</sub> -F <sub>c</sub> (dimer)	71nM	Ang1-F <sub>1</sub> -F <sub>c</sub> -F <sub>1</sub> (tetramer)	9.7nM
Variant domains	IC <sub>50</sub>															
Ang1-N <sub>1</sub> -C <sub>1</sub>	-															
Ang1-F <sub>1</sub> (monomer)	435nM															
Ang1 - C <sub>1</sub> -F <sub>1</sub> (dimer)	94nM															
Ang1 (tetramer+)	1.5nM															
Ang1-F <sub>1</sub> -F <sub>c</sub> (dimer)	71nM															
Ang1-F <sub>1</sub> -F <sub>c</sub> -F <sub>1</sub> (tetramer)	9.7nM															

Continued.....Table 1.1: A summary of binding studies performed for Ang1, Ang2 and associated truncated and recombinant fragments.												
Paper	Methods	Results										
(Cho et al. 2004a)	Affinity values ( <i>K<sub>d</sub></i> ) examined by biosensor SPR to soluble Tie2-Fc	<table border="1"> <thead> <tr> <th>Oligomeric Ang1 Variants</th> <th><i>K<sub>d</sub></i> (nM)</th> </tr> </thead> <tbody> <tr> <td>native Ang1</td> <td>7.5</td> </tr> <tr> <td>GCN4-Ang1</td> <td>158.5</td> </tr> <tr> <td>MAT-Ang1</td> <td>67</td> </tr> <tr> <td>COMP-Ang1</td> <td>20.5</td> </tr> </tbody> </table>	Oligomeric Ang1 Variants	<i>K<sub>d</sub></i> (nM)	native Ang1	7.5	GCN4-Ang1	158.5	MAT-Ang1	67	COMP-Ang1	20.5
Oligomeric Ang1 Variants	<i>K<sub>d</sub></i> (nM)											
native Ang1	7.5											
GCN4-Ang1	158.5											
MAT-Ang1	67											
COMP-Ang1	20.5											
(Weber et al. 2005)	Solid phase binding assay: Tie2-Fc immobilised on ELISA plates to measure binding at 405nm.ABS.	Soluble ΔAng1 (RBD Ang1 FReD) is the minimal variant of Ang1 that binds Tie2										
(Kim et al. 2009)	<i>In vitro</i> binding analysis to soluble Tie2-Fc by SDS-PAGE	<table border="1"> <thead> <tr> <th>Oligomeric Ang2 Variants</th> <th>% Binding of total input</th> </tr> </thead> <tbody> <tr> <td>native Ang2</td> <td>95</td> </tr> <tr> <td>GCN4-Ang2</td> <td>25</td> </tr> <tr> <td>MAT-Ang2</td> <td>33</td> </tr> <tr> <td>COMP-Ang2</td> <td>99</td> </tr> </tbody> </table>	Oligomeric Ang2 Variants	% Binding of total input	native Ang2	95	GCN4-Ang2	25	MAT-Ang2	33	COMP-Ang2	99
Oligomeric Ang2 Variants	% Binding of total input											
native Ang2	95											
GCN4-Ang2	25											
MAT-Ang2	33											
COMP-Ang2	99											
(Yuan et al. 2009)	Binding studies for recombinant human Ang1 and Ang2 with sTie2 immobilized on the solid phase of ELISA strips	50% saturation was achieved at 106ng/ml (~1.9nM) for Ang1 and 2,018ng/ml (~40.2nM) for Ang2. Ang2 binds Tie2 with ~20-fold lower affinity than Ang1.										

The interaction of Ang1 and Ang2 with the endothelial cell surface may be more complex, because several reports have shown that Ang1 can bind to other surface receptors such as integrins (Carlson et al. 2001, Saharinen et al. 2005) and Ang2 can also bind to integrins and activate intracellular signalling pathways (Hu et al. 2006). Ang2 is released faster after binding compared with Ang1 which suggests that Ang1 may bind with higher affinity to other receptors and therefore remain on the cell surface longer compared with Ang2 (Bogdanovic, Nguyen & Dumont 2006).

Ang1 has an agonistic while Ang2 has an antagonistic action on binding the same binding site in the Tie2 protein (Barton et al. 2006, Fiedler et al. 2003, Yuan et al. 2009). All binding studies have suggested that Ang1 and Ang2 share the same binding site on Tie2 receptor but none have been able to suggest the precise molecular basis for the opposing effects of Ang1 and Ang2.

## **1.2 Angiopoietin Variants**

There are significant problems with production and use of native Ang1. Large-scale production of recombinant Ang1 is hindered by aggregation and insolubility resulting from disulfide-linked higher-order structures. The generation of native Ang1 is difficult and the activity of the purified native Ang1 varies after purification, possibly as a result of its unique structural characteristics with a tendency to form variably sized multimers. Furthermore, the superclustering domain of Ang1 is the major cause of the protein's insolubility. Multimerization of Ang1 by the superclustering domain results in large aggregates over time, and these aggregates show a strong tendency to precipitate and thereby losing activity and failing to stimulate an angiogenic response (Cho et al.

2004a). Thus there are several problems associated with native Ang1 ligand; difficulty in production, aggregation, poor bioavailability, potency and inhibition by Ang2.

To try and overcome poor solubility and production problems a number of variant Ang molecules have been constructed with alterations in the coiled-coil domain. They are Ang1\* and COMP-Ang1 and COMP-Ang2.

In Ang1\*, the 73 amino-terminal amino acids of hAng2 are fused to hAng1 at residue 77. In addition to this, the cysteine corresponding to position 265 of human Ang1 is replaced by a serine. It has been used in several angiopoietin bioassays (Maisonpierre et al. 1997). Ang1\* is used primarily because it is much easier to produce in sufficient quantities for experimental use and is not susceptible to inhibition with Ang2 (Koblizek et al. 1998). However it still has some problems observed in native Ang1.

*Cho et al.* designed oligomeric Ang1 variants in which the N-terminal portion of Ang1 was replaced with the short coiled-coil domains of yeast transcriptional activator GCN4, matrilin 1, and COMP. With respect to elution solubility and stability all three forms were found to be better than native Ang1 which progressively lost its activity and aggregated over time (Cho et al. 2004a). The affinity values (dissociation constant,  $K_d$ ) between the Tie2 receptor and native Ang1, GCN4-Ang1, MAT-Ang1, and COMP-Ang1 were 7.5, 158.5, 67, and 20.5 nM, respectively.

The most successful attempt has been an engineered form called Cartilage Oligomeric Matrix Protein Angiopoietin-1 (COMP-Ang1). The coiled-coil domain has been replaced with the short coil of Cartilage Oligomeric Matrix Protein which functions effectively in oligomerization and prevents aggregation. COMP-Ang1 has

been demonstrated to have potent stable activity. It is easy to generate and does not form insoluble aggregates. Furthermore, COMP-Ang1 has greater biological activity than native Ang1, having the potential for enhanced positive vascular effects. Tie2 phosphorylation induced by COMP-Ang1 is greater and significantly more persistent than that of Ang1\*. Native Ang1 does not induce significant angiogenesis whereas COMP-Ang1 induces angiogenesis comparable with VEGF-induced angiogenesis (Cho et al. 2004a).

COMP-Ang1 has been administered by adenoviral vector to induce long-lasting vascular enlargement and increased tracheal blood flow. However, short-term administration of COMP-Ang1 induced transient vascular enlargement that spontaneously reversed within a month (Cho et al. 2005). Despite this advance in improving protein solubility COMP-Ang1 is still antagonized by Ang2.

As with Ang1, Kim et al. (2009) designed oligomeric Ang2 variants in which the N-terminal portion of Ang2 was replaced with the short coiled-coil domains of yeast transcriptional activator GCN4, matrilin 1, and COMP. The relative binding to the Tie2 receptor of native Ang2, GCN4-Ang2, MAT-Ang2, and COMP-Ang2 were 95%, 25%, 33%, and 99%, of the total input respectively (Kim et al. 2009).

In COMP-Ang2 the coiled-coil domain was replaced with the short coil of Cartilage Oligomeric Matrix Protein and the COMP-Ang2 thus produced was shown to function like COMP-Ang1 in *in vitro* and *in vivo* angiogenesis studies (Kim et al. 2009).

COMP-Ang2 strongly activated Tie2 when compared with native Ang2. Although native Ang2 binds almost as strongly as COMP-Ang2, its ability to phosphorylate Tie2 was much reduced. Thus as opposed to native Ang2, COMP-Ang2 was shown to

promote endothelial cell survival, migration, and tube formation in a Tie2-dependent manner, and also induced angiogenesis in wound healing. The potency of COMP-Ang2 was found to be identical to that of COMP-Ang1 *in vivo* demonstrated by enhanced angiogenesis in the wound healing region by both the variants (Kim et al. 2009).

### **1.3 Clinical applications**

#### **Vascular Protective effects of Ang1**

Angiopoietin-1 (Ang1) has a number of protective effects *in vivo*, including, inhibition of vascular leakage, suppression of vessel inflammation and mediating vessel survival (Brindle, Saharinen & Alitalo 2006).

##### **a) Anti-apoptotic**

Ang1 ligand suppresses apoptosis in a variety of different endothelial cell types, including human umbilical vein endothelial cells (HUVECs), microvessels of the intestinal villi, lung endothelial cells, aortic endothelial cells in response to stimuli ranging from mannitol treatment, irradiation, serum deprivation, hyperosmolarity to tumour necrosis factor- $\alpha$  (TNF- $\alpha$ ) treatment (Brindle, Saharinen & Alitalo 2006, Cho et al. 2004b, Kwak et al. 2000, Papapetropoulos et al. 2000, Zhao et al. 2003). In HUVECs Ang1 activates Tie2 which interferes with NF- $\kappa$ B signalling mediated by ABIN-2 recruitment. This prevents endothelial cells from undergoing apoptosis (Hughes, Marron & Brindle 2003).

##### **b) Anti-leakage**

Acute administration of Ang1 in adult transgenic mouse models has showed that Ang1 could promote vascular integrity and block plasma leakage induced by VEGF and

inflammatory agents (Thurston et al. 2000). Ang1 also significantly reduced VEGF-induced retinal vascular permeability in retinal ischemia (Nambu et al. 2004). Experiments using mouse models show that Ang-1 reduces leakage from inflamed venules by restricting the number and size of gaps at endothelial cell junctions through effects on intracellular signalling, cytoskeleton, and junction-related molecules (Baffert et al. 2006).

Extensive vascular leakage leads to oedema and swelling which has serious implications in pathologies such as brain tumours, diabetic retinopathy, following strokes, during sepsis and also in inflammatory conditions such as rheumatoid arthritis and asthma. This anti-permeability action of Ang1 may extend its potential from vascular therapeutic opportunities to application in numerous diseases (Thurston et al. 1999).

### **c) Anti-Inflammatory**

Ang1 acts as an anti-inflammatory cytokine on the vasculature. Ang1 suppresses adhesion of leukocytes to VEGF-stimulated HUVECs by inhibiting the expression of inflammation and thrombosis-associated adhesion molecules such as ICAM-1, VCAM-1, and E-selectin (Kim et al. 2001). It prevents microvascular leakage against endotoxic shock induced by lipopolysaccharide in mice by suppressing expression of E-selectin, ICAM1, and VCAM1 (Witzenbichler et al. 2005).

Rat model transplanted with cardiac allograft and overexpressing Ang1 showed anti-inflammatory effects *in vivo*, thus preventing allograft arteriosclerosis (Nykanen et al. 2003).

TNF- $\alpha$  is shown to induce Ang-1 expression in Rheumatoid arthritis synovial fibroblasts (RA-SF) primarily through NF- $\kappa$ B activation. Synovial fibroblasts are a key player during rheumatoid arthritis and intervention in this pathway may be relevant in the therapy of the angiogenesis that occurs in chronic inflammation (Scott et al. 2005).

### **Clinical Application of Ang2: Retinal Neovascularization**

The role of Ang2 assessed in transgenic mice model with therapeutic application in ocular neovascularisation (NV) has been studied. It was demonstrated that Ang2 was able to cause very substantial regression of retinal or choroidal NV under some circumstances, and without regression in normal retinal and choroidal vessels (Oshima et al. 2005).

#### **1.4 Manipulation of angiopoietins for understanding Ang:Tie interactions**

Given that Ang1 shows clear therapeutic potential and together with Ang2 plays a major role in various protective and pathological conditions, understanding the Ang-Tie2 interaction by manipulation of the angiopoietins is a very desirable prospect.

Better understanding of the Ang:Tie2 interactions could elucidate reasons for their differential action, facilitate direct administration of Ang1 or potent Ang1 mimetics as well as improve Ang-1 functionally by decreasing or blocking competitive antagonists such as Ang2 (Brindle, Saharinen & Alitalo 2006). Angio-manipulatory drugs may be designed to extend beyond angiogenesis to include non-neoplastic indications, including inflammation or atherosclerosis (Fiedler, Augustin 2006).

The angiopoietin variants generated to date have approached only the stability aspect of angiopoietins by manipulating and replacing the oligomerization domain in

the protein. However, the role of the Tie2 binding interface, the sole point of contact between the receptor and ligand has not been modified or studied. There is significant potential for further engineering of angiopoietins, particularly to generate higher affinity ligands. The influence of the binding strength of angiopoietins could lead to better understanding of its interactions and possibly broaden the scope in design of therapeutic drugs. Production of engineered receptor binding domain of angiopoietins could produce ligands with a range of binding affinities which in turn could induce varying, possibly novel downstream signalling effects.

### **1.5 Protein Engineering**

Protein engineering attempts to alter protein sequence to tailor proteins with a desired function. The ability to modify specific sites in the DNA by mutagenesis has enabled the replacement of any amino acid in a protein. The field of protein engineering has evolved rapidly leading to synthesis and manipulation of proteins with applications in basic research, biotechnology, material sciences and therapy.

Modified proteins with useful and novel characteristics of substrate specificity and affinity have wide range of applications in the chemical, pharmaceutical, and agricultural sciences. Protein engineering holds great promise for the development of new biosensors, diagnostics, therapeutics, and agents for bioremediation (Glasner, Gerlt & Babbitt 2007). From the perspective of therapeutics, protein modification promises a tremendous opportunity.

There are two general strategies for protein engineering; rational design and directed protein evolution. Rational design attempts to understand protein structure and function at a complete mechanistic level so that any desired change can be effected by

calculation from first principles. A closely related bioinformatics approach to design protein by rational design is computational protein design (CPD). CPD is different from traditional rational design because it first selects a desired structure for the protein to be designed and then finds sequences able to fold into it. Directed protein evolution is a powerful algorithm using iterative cycles of random mutagenesis and screening for tailoring protein properties. In principle it mimics natural evolution but over a shorter timescale.

Researchers often employ a combination of these strategies for protein engineering. Proteins are tailored to display an extension or enhancement in an already inherent function, or display a diverse range of function, or display a novel function.

### **1.5.1 Rational Design**

The amino acid sequence of a protein encodes its three-dimensional structure which determines its biological function. Rational design utilises this structural information of the protein to determine which changes in the amino acid sequence would result in the desired function. The structure of the protein can be studied using protein sequence data, protein structural data (crystal structures or NMR data) if available, utilising structural information of its close relatives in the protein family and using molecular dynamic simulation to generate a theoretical protein structure. This 3-D structure of proteins is used as a start point to select amino acid residues that are most promising for generation of desired function. There are several forces of interaction in the protein structure that determine its functions. They include non-covalent interactions within the polypeptide chain, interaction with its solvent, protein interaction with ligands based on van der Waals forces, electrostatic interactions, hydrogen bonds, the

hydrophobic effect, and the favourable packing interactions associated with the condensed state of protein interiors (Dill 1990).

The molecular model is defined based on the protein structural and functional information which is combined with an algorithm for selection and mutation of the residues which would generate the desired function in the protein optimally. Either the “negative design” approach (raising the free energy of competing states) or the “target state optimization” (energy minimization of the desired state) based simulations can be performed to select important residues and mutational changes likely to affect the function (Hellinga 1997). Information obtained from the molecular model is used to generate protein mutants whose function is characterised in laboratory based experimental analysis. The mutants are assessed in a functional assay and depending on the outcome of this first round of designing, a new approach may be required or additional complexity may introduced in the molecular model. Each round of rational design iteration requires defining or refining of new rules, parameters and algorithms (Hellinga 1997).

### **Limitations of Rational Design**

The process of modifying protein sequence for engineering the desired function is limited by the molecular level of understanding of how individual amino acids contribute to the structure and function of a protein. Proteins are highly complex with very little difference in the stability of active and inactive unfolded forms (Tokuriki, Tawfik 2009). The protein structural variables of backbone topology, sequence and side chain conformation pose an additional challenge since enormous numbers of variations of the amino acid sequence are possible. The immense combinatorial complexity,

infinite possible outcomes for every interaction and subtle energetic differences make it difficult to understand even the simplest and basic interactions (Hellinga 1997).

Even though several algorithms with improved accuracy are able to predict globally correct folds, there is a lack of universal applicability to the wide variety of proteins with numerous incompletely defined functions. This can be attributed to the fact that non-covalent interactions are difficult to quantify and free energy associated with mutations are difficult to predict accurately. Hence it is difficult to get correct local details for each unique protein (Morgan, Massi 2010).

Protein stability is determined by interactions among individual amino acids, but not all residues contribute equally. The mutation in the interior of designed proteins may lead to a high degree of disorder, which does not resemble the tightly packed, unique arrangement of natural systems. The side chains in a disordered core adopt many alternative conformations, instead of assuming a single, specific arrangement. Determining the contributions of individual residues to the relative stability of even a small protein requires predicting and characterizing a large number of variant sequences based on simulation. The accuracy in determining the effect of mutation of individual residues on stability is complex and tedious to achieve by simulation models (Hellinga 1997).

Most algorithms have been developed and tested to rebuild the surface of existing binding sites redecorating a fixed protein backbone with amino acids (Pabo 1983). There have been some successes in designing metalloproteins by this approach (Benson, Wisz & Hellinga 1998, Wilson, Mace & Agard 1991). However, several studies have indicated that activating mutations for some proteins may be outside the active/binding site and may not be predicted using computational methods (Sarkar et al. 2007,

Shimotohno et al. 2001, Spiller et al. 1999). Focusing on too few residues may decrease the possibility of identifying the global optima and may result in discarding of interesting protein functions.

The ability of a computational algorithm to explore this vast landscape and seek out preferred solutions that have to be distinguished from closely related inferior possibilities is a crucial and a difficult to achieve component of any rational design approach.

### **1.5.2 Computational Protein Design**

In recent years, an alternative approach to rational design that designs protein based on structural information has matured called computational protein design (CPD). In one sense it first defines the goal structure and then finds alternative sequences that fit best to the goal outcome. The number of natural protein folds is estimated to be approximately 1000, which is much smaller than the number of folds we may think (Orengo, Jones & Thornton 1994). In computational protein design, the space of amino acid sequences is searched and ranked according to their folding free energy towards a predefined folded structure. Favourable low energy substitutions are retained and unfavourable high energy substitutions are discarded (Van der Sloot et al. 2009).

It has been suggested that a combination of CPD and directed evolution could provide best results. The designed library contains only favourable sequence folds selected by CPD. These designed libraries showed greater diversity and greater preservation of function than the randomly generated libraries. This combination approach would rapidly screen *in silico* the space of sequences and eliminate sequences incompatible with the fold, generating a combinatorial library with a degree of

complexity low enough to be experimentally screened and without a relevant loss in its diversity (Suarez, Jaramillo 2009).

### **Limitations of Computational Protein Design**

Many drawbacks of rational design stand true for computational protein design and stem from the basic problem; the lack of understanding comprehensively the nature of protein structure and function. It is expected that similar sequences usually share the same structure. However, there are many cases of nearly identical structures sharing no sequence similarity (Brenner, Levitt 2000, Devos, Valencia 2000, Koppensteiner et al. 2000, Tian, Skolnick 2003).

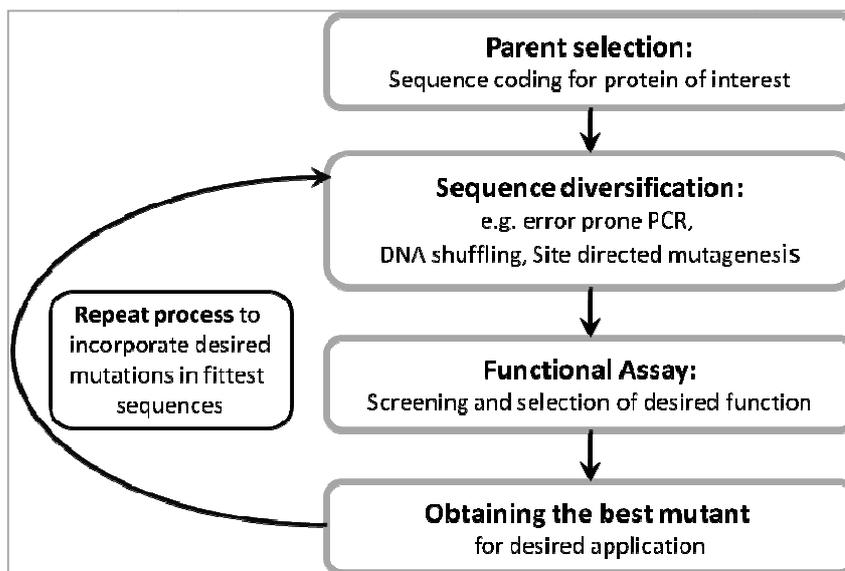
Also, it is assumed that free energies (folding or binding free energies) represent some of the desired characteristics of the protein. Solvent-related effects and the search for an adequate approximation for the solvation energies are other limitations that need to be addressed. Sampling to perform a suitable exploration of the space of available sequences is another challenge in this approach.

For the computational redesign of improved protein-binding affinity, accuracy of energy function within a few orders of magnitude is critical; making redesign from nanomolar to picomolar affinities more challenging. Predictions from visualization-based expert design are more successful in affinity-enhancing mutations and exhibit a higher success rate than the computational methods (Lippow, Tidor 2007).

### 1.5.3 Directed protein evolution

Directed protein evolution is a general term used to describe a method by which the natural process of evolution by mutation and selection of mutants with desired phenotype is performed *in vitro* to modify protein characteristics. In principle it mimics natural evolution but over a shorter timescale. Natural evolution selects only for biologically relevant sequences, however, directed protein evolution is a biological optimization process which modifies sequences to obtain diverse, novel or efficient functions (Romero, Arnold 2009). Over the last three decades, directed protein evolution has emerged as a powerful technology platform in protein engineering (Yuan et al. 2005). Directed protein evolution uses iterative cycles of random mutagenesis and screening for tailoring protein properties to our needs in industrial and therapeutic applications and for elucidating proteins' structure function relationships (Wong, Zhurina & Schwaneberg 2006).

Directed evolution involves diversification of a parent functional protein sequence, expression of the mutants and selection of improved variants (Figure 1.2). The desired variant is then further diversified and selected iteratively until the desired level of improvement is achieved by accumulation of multiple beneficial mutations.



**Figure 1.2: Schematic outline of a typical directed evolution experiment.** The start point is the gene encoding the protein of interest. This parent gene is randomly mutagenised to generate a library of mutants. Mutants with desired properties are assayed by a functional screen and selected. These selected improved mutants are used for iterative rounds of diversification, screening and selection till the desired property is obtained.

Some proteins are said to be more evolvable than others, and display high stability to tolerate mutations and natural functional diversity (Tokuriki, Tawfik 2009, Aharoni et al. 2005, Bloom, Raval & Wilke 2007, England, Shakhnovich 2003). However the complexity of proteins makes it difficult to predict protein evolvability.

Diversification can be achieved by a range of techniques such as error prone PCR (Mullis, Faloona 1987), DNA shuffling (Stemmer 1994), sequence saturation mutagenesis (Wong et al. 2004), chemical mutagenesis (such as ethyl methane sulfonate) (Lai et al. 2004), random oligonucleotide mutagenesis (Hill, Oliphant & Struhl 1987), use of mutator strains such as *E. coli* XL1-RED (Coia et al. 1997). However, the number of possible protein variants increases due to numerous combination and permutations of mutations possible at every position in the sequence and increases proportionally with size of the parent protein. The size of the library generated is restricted compared with the possible theoretical library. Most library sizes

range from  $10^9$ - $10^{10}$  mutants. A library with a high frequency of improved mutants is desired, but improvement is not dependent on mutation rate. Increasing the mutation rate may increase the number of variant sequences but it is not necessary this will ensure functional improvement. High mutations rate can be deleterious and multiple mutations can lead to complete loss of function. Most libraries contain 1-2 amino acid substitutions per gene, which corresponds to 40-50% inactive clones in the library (Romero, Arnold 2009, Arnold, Georgiou 2003a, Arnold, Georgiou 2003b).

Once the library is generated, a good screen is designed to ensure that even small improvements get accurately selected. Proteins with the desired function are then selected and improved further by iterative diversification and selection until no further improvement is obtained.

### **Limitations of Directed Evolution**

The ability to obtain the best novel mutant depends on performing multiple laborious techniques of repeated rounds of mutation, transfection and selection. Generation of mutants is time consuming and cumbersome and may subsequently result in majority of non-functional mutants. Furthermore, random mutation of 30-50% single amino acid are strongly deleterious, 50-70% mutations are neutral and only 0.01-1% are beneficial (Axe, Foster & Fersht 1998, Guo, Choe & Loeb 2004, Shafikhani et al. 1997). There are many ways to reach a neutral or deleterious mutation and very few to reach the desired functional fitness (Romero, Arnold 2009).

Perhaps the biggest constraint of directed evolution and protein engineering in general is that an exhaustive search of sequences is feasible only if a small number of residues are allowed to vary or if the number of allowed amino acids is reduced.

Mutating a 100 residue protein in every combination and permutation would generate at least 1.5 million double mutant sequences. Library size restricts the analysis of all possible mutated sequences and thus the probability of obtaining the ideal variant. To decrease the library size and increase the possibility of obtaining improved variants, structural information can be applied to focus directed evolution in particular regions of the protein (Aharoni et al. 2005, Park et al. 2005). However, activating mutations away from the active site of proteins have been reported, so it is not always possible to choose the most appropriate position to mutate (Shimotohno et al. 2001, Spiller et al. 1999).

There is no rule of thumb by which the parameters of directed evolution can be set. Each protein requires a custom-made program of different rates of mutations, different sensitivity of screening and different rounds of iteration to obtain optimal improvement.

### **1.6 Somatic Hypermutation (SHM)**

The success of directed evolution strategy depends on adequate diversification process, efficient expression system and sensitive screens. However, the problems of current approaches of directed protein evolution are compounded when evolving mammalian proteins. Also, diversification with low error rates may yield few unique clones, high error rates may destabilize proteins, and random mutational techniques may introduce mutational bias (Drummond et al. 2005). Mammalian protein expression is influenced in non-mammalian hosts because they are limited by appropriate post-translational modifications and folding environments. Mammalian DNA is difficult to target for single genomic integration and displays a varied expression range in mammalian host systems (Majors, Chiang & Betenbaugh 2009). Only a small fraction of the all possible mutants are screened due to constraints of the restricted size of the

mutant library and the average number of mutations in each clone. The traditional directed evolution approach is laborious, requiring several stages of DNA diversification, transfection/expression and screening to reach a desired end point.

An ideal alternative would be to enable diversification, expression, screening and selection in the mammalian cell. One way to simplify directed protein evolution is to utilise somatic hypermutation (SHM) as an 'in-cell' mutagenesis tool which uses the endogenous mutation capability of the cell to mutate any desired protein sequence (Wang et al. 2004, Wang, Yang & Wabl 2004).

### **1.6.1 Features of SHM**

Prior to exposure to antigen, the initial generation of a diverse antibody repertoire is achieved in the early B-lymphocyte. The initial response to antigen is provided by IgM antibodies whose binding sites usually exhibit a relatively low affinity for antigen (Peled et al. 2008). The functional variable region gene unit formed through the immunoglobulin rearrangement process (Tonegawa 1983) is not enough to provide high affinity antibodies against the wide range of antigens encountered. The generation of antibody diversity depends in a major way on a second wave of diversification by SHM (Bernard, Hozumi & Tonegawa 1978). Over time, the antibody response matures to yield IgG class antibodies that display greatly increased affinity for antigen (Li et al. 2004b). This affinity maturation is achieved during B cell expansion in germinal centres by an iterative alternation of somatic hypermutation (SHM) and antigen-mediated selection (Kelsoe 1996).

The ability to generate diversity is one of the hallmarks of acquired immunity and this is achieved by the process of SHM. SHM is triggered by Activation-induced

deaminase (AID) enzyme which deaminates deoxycytidine residue to generate a (uracil) U:G mismatch (Muramatsu et al. 1999). This mismatch may be processed by three possible routes to generate mutations. DNA replication of the uracil causes transition mutations at C:G base pairs. Alternatively, Uracil-DNA glycosylase may remove the U to create abasic site which on DNA replication gives rise to transversion or transition mutations. Recognition of the mismatch by mismatch repair enzymes (MMR) can generate mutations at A:T base pairs near the U:G lesion (Di Noia, Neuberger 2007).

SHM is characterized by high frequency of single nucleotide substitutions in the DNA encoding the Ig V<sub>H</sub> and V<sub>L</sub> segments (Neuberger, Rada 2007). SHM occurs at rates of 10<sup>-5</sup> to 10<sup>-3</sup> mutations per base pair per generation (Peled et al. 2008) which is ~ 1 million-fold higher than the spontaneous rate of mutation in most other genes. The mutations are mainly single base substitutions, with occasional insertions and deletions (Li et al. 2004a). Although mutations occur throughout the rearranged V regions and its immediate flanking sequences, there is preferential targeting to RGYW/WRCY motif (G:C is a mutable position; R=A/G, Y=C/T, and W=A/T) and TAA or TA (A is the mutable position) motifs that are therefore referred to as hot spots (Goyenechea, Milstein 1996, Rogozin, Kolchanov 1992). Although the RGYW/WRCY motif is largely accepted to be the mutable motif for AID targeting, recently DGYW/WRCH motif (G:C is the mutable position; D=A/G/T, H=T/C/A) has been identified as a better descriptor of the Ig mutation hotspot (Rogozin, Diaz 2004). This directed *hypermutation* allows for the selection of B cells that express immunoglobulin receptors possessing an enhanced ability to recognize and bind a specific foreign antigen (Papavasiliou, Schatz 2002).

### 1.6.2 SHM as a mutagenesis tool

The use of SHM as a general mutagenesis tool depends on the ability of AID to target non-Ig genes. Several factors of SHM mechanism are not well understood, hence the understanding of any other influencing factor such as Ig enhancers, suitable B cell lines, action of AID are important to determine if a transgene can be predisposed to SHM mutation and hence directed evolution.

Until recently, SHM had been reported to be confined to the 2kb V(D)J segment of the Ig genes in B cells for affinity maturation of antibodies (Lebecque, Gearhart 1990). In 1995, Neuberger et al. reported SHM driven hypermutation of non-Ig genes at the V segment loci. The mutations of non-Ig genes were shown to have features similar to SHM of Ig genes; similar rate of mutation, preference of transitions over transversions, absence of deletions or insertions and typical hotspots (Yelamos et al. 1995). This finding made it possible to explore SHM as a means to mutate non-Ig genes.

Since the exact mechanism of SHM has not been fully elucidated, it was important to determine the essential factors that were sufficient and essential for SHM. The demonstration of AID mediated SHM in non B cells (fibroblasts), hybridomas and *E.coli* suggested that AID was the single most crucial factor in driving SHM and was sufficient to initiate SHM driven mutations in a transgene (Yoshikawa et al. 2002, Martin et al. 2002, Petersen-Mahrt, Harris & Neuberger 2002). Thus, activation induced cytidine deaminase (AID) enzyme was identified as sufficient to activate SHM.

A key determinant in subjecting a transgene to AID driven mutation is the rate of transcription of the transgene of interest. Bachl et al. (2001) have reported that high

transcription rates predispose genes to mutation by AID (Bachl et al. 2001). This predisposition to AID mediated SHM of highly transcribed genes is because AID deamination is targeted specifically to ssDNA and to dsDNA only when it presents itself as ssDNA substrate as observed during DNA transcription (Chaudhuri et al. 2003, Ramiro et al. 2003). However, it should be noted that frequency of mutation is not evenly distributed on a highly transcribed gene, and initial ~150bp of the DNA sequence shows a virtual absence of mutations. This property of AID is suggested to be due to its inability to associate with the transcription complex until transcription has entered the elongation phase (Rada, Milstein 2001, Longerich et al. 2005). SHM is shown to repair premature stop codons deliberately introduced in non-Ig genes, provided that they are transcribed at a high enough rate (Bachl et al. 2001, Sale, Neuberger 1998, Wang, Harper & Wabl 2004). These studies showed that a combination of AID and high transcription can induce SHM in transgenes.

Another consideration for using SHM as a mutagenesis tool was the role of Ig enhancers in the mechanism and whether their presence was essential for optimal mutation targeting. Several studies have suggested that Ig enhancers are required to target hypermutation to the Ig loci and their function indirectly protects non-Ig genes in the genome from indiscriminate AID mutation (Bachl et al. 2001, Bachl, Olsson 1999). This paradigm was challenged when Wabl et al. (2004) reported that SHM was not just targeted to the Ig loci but can target transfected transgenes genome wide (Wang, Harper & Wabl 2004). This was shown using GFP as an inactive reporter with a premature stop codon which was reverted to the GFP activated form. Sequencing of the GFP positive clones revealed that clones did not show integration of the GFP transgene into the Ig loci (Wang, Harper & Wabl 2004). Although SHM does not require Ig gene elements or integration of the transgene at Ig loci, the SHM mutation rate of transgenes at the Ig loci

have been reported to be comparatively higher (1-3 mutations in transgene at non-Ig loci, 5-7 mutations in transgene at Ig loci between IgHV7-34-1 and IgHV4-34) (Wang et al. 2004).

These studies showed that AID functions as a general mutator to mutate over-expressed transgenes and can target transcribed transgenes genome wide but the mutation rate is enhanced when transcribed transgenes are integrated at the Ig locus. With these features of SHM to autonomously diversify a chosen gene of interest, it became possible to evolve proteins *in situ* and explore larger sequence space for protein engineering. Two independent groups simultaneously demonstrated the ability of SHM as a tool to mutagenize exogenous genes and also its utility in producing a desired phenotype (Wang et al. 2004, Wang, Yang & Wabl 2004).

Experiments using a mutant form of green fluorescent protein with low fluorescence were performed by Wang et al. (2004). B cells transfected with the mutant GFP gene were expanded and the cells obtained after undergoing SHM showed marked increase in fluorescence. The increase in fluorescence was attributed to mutations by which the mutant had reverted to wild type forms as well as variants with additional suitable mutations (Wang, Yang & Wabl 2004).

In the second study, Monomeric Red Fluorescent (mRFP) protein was evolved using this 'in cell' SHM mediated mutation principle. Over 23 rounds of iterative selection using SHM and fluorescence-activated cell sorting of Ramos B cell line were performed. Mutant mPlum and mRaspberry with enhanced photostability and far-red emissions were obtained. Parallel experiments using random mutagenesis or rational design based on crystal structure failed to generate mRFP mutants with emission maxima >632nm. Saturating mutagenesis of some of the mutant residues in mPlum

resulted in neutral shifts, loss of fluorescence or blue shift of the protein further emphasising the optimization process of the SHM (Wang et al. 2004).

This Monomeric Red Fluorescent (mRFP) protein was evolved in the Ramos cell line which is a human B cell line that hypermutates its Ig V genes constitutively during culture. SHM in Ramos cells uses activation-induced cytidine deaminase (AID) and error-prone DNA repair to introduce point mutations into the rearranged V regions of Ig at a rate of mutation between  $0.2-1 \times 10^{-4} \text{ bp}^{-1} \text{ generation}^{-1}$  (Sale, Neuberger 1998).

SHM directed protein evolution has also been demonstrated in the chicken B cell line DT40. The rate of mutation in the DT40 cells was studied using a disruption mutant and has been calculated to be  $0.4 \times 10^4 \text{ bp}^{-1} \text{ generation}^{-1}$ . A majority of hypermutation targeting resulted in C:G/G:C transversions and closely resemble SHM of murine and human germinal centre B cells (Arakawa, Buerstedde 2009). Arakawa et al. (2002) generated a new GFP variant in engineered DT40 cells (Arakawa, Hauschild & Buerstedde 2002) that displayed more than 3-fold higher fluorescence activity than the best GFP currently available for bio-imaging of vertebrate cells. This was achieved only after three rounds of FACS sorting of DT40 cells (Arakawa et al. 2008).

Thus SHM of non-Ig genes can be used to accumulate multiple reinforcing mutations to produce new and desirable phenotypes which are impossible to generate by conventional mutagenesis (Wang et al. 2004, Wang, Yang & Wabl 2004, Wang, Harper & Wabl 2004). It should be noted that SHM mediated genetic diversity created by the immune system is able to generate antibodies that bind to nearly every antigen encountered, suggesting the ability of this system to explore the sequence space robustly. SHM-mediated protein evolution in live cells is a suitable alternative to overcome the limitations of *in vitro* mutagenesis and screening. Mammalian proteins

targeted to the Ig loci in B cells allow integration and expression of a single unique cDNA sequence, which can be mutated, expressed (with appropriate post-translational modifications) and screened iteratively within the mammalian cell environment. It can be used to sample a large protein space, and directly link genotypes to cell phenotypes (Wang et al. 2004). To date, this approach of SHM for directed protein evolution has been demonstrated for evolving only fluorescent proteins. This artificial evolution system may be applied for any gene sequence using a sensitive and rapid strategy for selection and screening of improved function.

### **1.7 Expression of Mutant Libraries: Cell Surface Display**

There are 3 basic methods of expression of protein variants for efficient screening, selection and recovering mutated sequences of interest. Both spatially addressable methods and compartmentalization methods have drawbacks when screening for binding function (Lin, Cornish 2002). The most straightforward way to screen for binding function of a mutant library is by the physical linkage method. The easiest way to physically link a protein to an anchor while maintaining its native conformational fold is by displaying the protein with a distinct function on the surface of a living cell; this forms the basis of any cell surface display strategy.

Cell surface display method requires a protein, or protein library of interest be fused to a membrane protein, which serves as an anchor to present the proteins on the cell surface. There are three components to a successful cell surface display system: the anchor protein to hoist the protein on the extracellular membrane, the host cell for expression of the fusion protein and the screening assay that can be performed on the surface display system (Lee, Choi & Xu 2003).

The anchor used in cell surface display expression of the protein should fulfil essential criteria to be used effectively. Firstly the anchor sequence should be amenable to fusion with foreign sequences and should be stable and resistant to the action of proteases. The anchor protein should be physiologically compatible to maintain host cell membrane integrity in which the fusion proteins are expressed. Finally, it should allow expression and appropriate orientation of the protein of interest so that it can be detected sensitively by the screening assay (Lee, Choi & Xu 2003).

In choosing a host for this method the primary consideration must be the biosynthetic needs of the protein to be engineered. Mammalian cell surface expression systems are ideal for expression of engineered mammalian proteins. Mammalian cell surface display allows appropriate post translational modifications and folding for mammalian proteins and is efficient in fast generation and reliable downstream characterization of the variants (Lofblom et al. 2007, Kieke et al. 1997). Human embryonic kidney 293T cells that are widely used for transient protein expression can be used for cell surface display of single-chain Fv antibodies for affinity maturation. Also a random peptide library was expressed on the surface of a mammalian cell by applying retroviral vectors. Expression of proteins on the mammalian cell surface is an attractive tool for biological screenings in more contextual and physiological environments (Wolkowicz, Jager & Nolan 2005).

Selection or screening of the mammalian cell surface display library of mutants, by fluorescence-activated cell sorting (FACS) is a simple and fast way to assess the cells displaying the variant mutant proteins. The advantage of cell surface display is that the FACS detection tool can be used in combination to allow screening of more than 50,000 events per second (Matsuura, Yomo 2006).

## 1.8 Binding and Selection Assay

Cultured cells expressing the library of the protein of interest on the extracellular surface of the B cells can be analysed by binding assays. Since the protein expressed on the cell surface is in the correct orientation and functional conformation in a contextual cellular environment, the binding assay can be performed using a population of live cells. The epitope tag on the expressed protein and the epitope tag on the bound protein can be labelled with a fluorophore. Dual labelling of the library population of variants can account for expression levels which can be normalised to detect improved binding affinity (Kieke et al. 1997). Screens based on equilibrium binding of cell surface expressed proteins can be quantitated accurately using the fluorescent labels.

Flow cytometry screening allows the selection of clones using two-channel FACS sorting; one for expression and second for binding. A sorting window is set based on the fluorescence of the protein bound to the surface expressed ligand. Desired population of mutants can be selected at high speed by FACS sorting technique. Together with the high polyvalency on the cell surface it enables real-time quantitative screening. The enrichment after each sort cycle is easily monitored and used to determine optimized sorting parameters for the next round (Kronqvist et al. 2008). The selected mutant protein expressing B cells are grown in culture and subjected subsequently to a more stringent next round of FACS selection. Iterative rounds of FACS selection of the cell surface expressed mutant library are rapid. This approach does not require tedious repetitive rounds of subcloning and transfections and purification of each clone (Lofblom et al. 2007, Feldhaus et al. 2003).

The selected mutants expressed on the cell surface can be characterized based on binding parameters such as  $K_d$  determined by flow cytometry and compared with the wild-type protein binding affinity.

In this work, the approach of SHM mediated mutagenesis combined with mammalian cell surface display and analysis by FACS was used to explore Ang:Tie binding by iterative rounds of directed evolution.

### **1.9 Hypothesis to be tested**

In this thesis the following points were investigated.

- 1) Ang-1 P-domain is the minimal binding domain capable of Tie2 binding.
- 2) Rational substitution of key residues in the receptor binding interface of Ang2 influence binding affinity to Tie2.
- 3) SHM combined with surface display can be used to evolve angiopoietins with new binding characteristics.

## Chapter 2: Materials and Methods

### Materials

#### Molecular Biology

1. 10X PCR buffer containing 1.5mM MgCl<sub>2</sub> (New England Biolabs)
2. dNTP: mix containing four deoxyribonucleotides: dATP, dCTP, dGTP and dTTP (10 mM, Qiagen)
3. Phusion High-Fidelity DNA polymerase (2units/  $\mu$ l, New England Biolabs)
4. BioTaq DNA polymerase (5units/ $\mu$ l, Bionline)
5. Template pcr2.1 Ang (10ng/ml)
6. All oligodeoxyribonucleotides (primers) used were 1pmol/  $\mu$ l, purchased from MWG- Biotech AG. For detailed list of primers see Appendix 2.
7. Distilled water DNase-RNase free (Gibco, Invitrogen)
8. Mineral Oil (Sigma).
9. Restriction enzymes (10 units/  $\mu$ l), Restriction enzyme Buffers (Roche)
10. TAE buffer- 40mM Tris, 1mM EDTA, 2.85ml Glacial acetic acid diluted in 50ml dH<sub>2</sub>O.
11. Agarose (Melford Molecular Biology Grade Agarose)
12. Ethidium Bromide (10 mg/ml, Invitrogen)
13. Loading Dye - 0.05% bromophenol blue, 10mM Tris-HCl, 10mM EDTA, 65% sucrose.
14. TrackIt 1 Kb Plus DNA ladder (0.1  $\mu$ g/  $\mu$ l, Invitrogen)
15. Phenol-chloroform-isoamyl alcohol 25:24:1 saturated with 10mM Tris, pH8, 1mM EDTA (Sigma)
16. Chloroform stabilized with 100ppm amylene (Fisons Analytical Reagent)

17. Ethanol 99.7-100% (v/v)
18. 3M Sodium Acetate (pH5.2)
19. Shrimp Alkaline Phosphatase (1 unit/  $\mu$ l, Promega)
20. Shrimp Alkaline Phosphatase 10X buffer (When diluted to 1X working concentration contains 50 mM Tris-HCl and 10 mM  $MgCl_2$ )
21. T4 DNA ligase (1U/  $\mu$ l , Roche)
22. Ligase buffer contains 66 mM Tris-HCl and 5 mM  $MgCl_2$ , 5 mM DTT, 1 mM ATP
23. Site directed mutagenesis reagents (Stratagene)- 10X reaction buffer, dNTP (10mM mix), *PfuTurbo* DNA polymerase (2.5units/ $\mu$ l)

### **Microbiology**

1. XL 10-Gold Ultracompetent Cells (Stratagene) used for site directed mutagenesis as per manufacturers' instructions.
2. TOP10 competent cells (Invitrogen) used for TA cloning and transforming ligation products as per manufacturers' instructions.
3. XL-Blue competent *E.coli* cells used for transforming plasmids
4. BD Falcon polypropylene round bottom tubes
5. Luria Bertani (LB) Broth

LB broth was prepared using 10g tryptone, 5g yeast extract, 10g NaCl and distilled water to 1 litre, pH 7.0. Solution was thoroughly mixed and autoclaved for 20 min at 120°C and stored at room temp. Prior to use, ampicillin was added to a final concentration of 100 $\mu$ g/ml. After addition of ampicillin, the broth was stored at 4°C.

6. Luria Bertani (LB) Agar Plates

LB agar plates were prepared using 10g tryptone, 5g yeast extract, 10g NaCl, 15g agar and distilled water to 1litre, pH 7.0. Solution was thoroughly mixed and autoclaved for 20 minutes at 120°C. Solution was allowed to cool to approx 50°C

and ampicillin was added at a final concentration of 100µg/ml. Approximately 25 ml of agar was poured into a 100 mm petridishes (Bibby Sterillin). The LB ampicillin plates were stored at 4°C.

7. Glycerol >99% (Sigma-Aldrich)

### **Tissue Culture**

1. CHO and HEK Complete media - 500ml  $\alpha$  Minimal Essential Media (Lonza BioWhittaker), 10% FCS (Harlan Sera Laboratories), 100 units/ml penicillin, 100 ug/ml streptomycin and 10 µg/ml L-glutamine.
2. Ramos Complete media- RPMI 1640 medium, 10µg/ml L-glutamine, 10% FCS (Lonza BioWhittaker).
3. DT40 Complete Media- Chicken Medium: Dulbecco's Modified Eagle's Medium, Ham's F-12 (DMEM/F-12) with L-Glutamine, 10% Foetal Calf Serum, 1% Chick Serum (Gibco), 100µM beta-mercaptoethanol. (Lonza BioWhittaker)  
Freezing Medium: 50%FCS, 40% complete media, 10% DMSO.
4. Phosphate Buffered Saline (PBS) – 140mM NaCl, 2.7 mM KCl, 10 mM NaPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4
5. 10X Trypsin-EDTA (TE)- 0.05% trypsin, 0.02% EDTA diluted in PBS
6. All plasticware was obtained from Nalge Nunc International.
7. OptiMEM I Reduced Serum Medium (Gibco)
8. Lipofectamine<sup>TM</sup> 2000 (Invitrogen)
9. Nucleofector solution V - 0.5ml supplement in 2.25ml of nucleofector solution V.  
100 µl of this was used per nucleofection. (Amaxa Biosystems from Lonza)
10. Plastic pipettes, Certified cuvettes (Amaxa Biosystems from Lonza)
11. Nucleofector solution T- 0.5ml supplement in 2.25ml of nucleofector solution T.  
100 µl of this was used per nucleofection. (Amaxa Biosystems from Lonza)

12. Geneflow Electroporation Cuvette 4mm (Cell Projects)
13. G418 sulphate salt solution concentration 50mg/ml (Sigma)
14. Puromycin 10 $\mu$ g/ $\mu$ l (Sigma)

### **Protein Studies**

1. 3X Sample Buffer(SB)- 50 mM Tris, pH 6.8, 10% glycerol, 2% SDS, 0.1% bromophenol blue, 5mM EDTA.
2. 2X Reducing SB-600  $\mu$ l 3X SB, 300  $\mu$ l distilled water, 30 mg DL- Dithiothrietol (Sigma).
3. 10X Tris Buffered Saline – 25 mM Tris, pH 7.4, 144 mM NaCl.
4. Lysis Buffer – 1 protease tablet/10ml solution (Roche Biosciences), 50mM Tris-HCl (pH7.4), 50mM NaCl, 1mM NaF, 1% triton TX-100 and 1mM EGTA
5. Precision Plus Protein Kaleidoscope Standard (BioRad)
6. 10X Running Buffer- 192 mM glycine, 25 mM Tris, 30 mM SDS made upto 1 litre in dH<sub>2</sub>O. This was diluted in dH<sub>2</sub>O to give 1X Running buffer
7. 1X Transfer Buffer- 192 mM glycine, 25 mM Tris containing 20% [v/v] methanol.
8. Blocking solution- 5% (w/v) non-fat dried milk powder (Marvel) in TBS 0.1%TX-100
9. anti-FLAG M2 primary antibody – 2 $\mu$ g/ $\mu$ l (Sigma)
10. anti-mouse IgG secondary antibody HRP linked(Amersham Biosciences)
11. anti- hTie2 ectodomain primary antibody (100  $\mu$ g /ml R and D Systems)
12. anti-goat IgG secondary antibody HRP linked
13. Hybond-ECL nitrocellulose membrane (Amersham Biosciences)
14. ECL solution for chemiluminescent detection, DMSO was supplemented with 250mM luminol and 90 mM p-Coumaric acid and 3  $\mu$ l of H<sub>2</sub>O<sub>2</sub>. (Sigma)

**Immunohistochemistry**

1. TBS 2% BSA- 2% BSA diluted in TBS (w/v)
2. Block solution 1 (BSA- TX100-TBS) - 2% BSA diluted in 0.1% TX-100 TBS (w/v)
3. Block solution 2 - 10%FCS-PBS solution
4. Anti- FLAG M2 primary antibody (1mg/ml) (Sigma)
5. Mouse Ig Cy2 linked (Amersham Biosciences, GE Healthcare) 1mg/ml
6. Rabbit polyclonal 6X His tag (Biotin) at 1mg/ml (Abcam)
7. Streptavidin R-Phycoerythrin (Sigma)
8. Recombinant Human Tie2/Fc Chimera 500ng/μl (R&D Biosystems)
9. 4% paraformaldehyde- 4 g paraformaldehyde dissolved in 100ml PBS at 80°C. Add 6 drops of 1 M NaOH. Filter and store at 4°C for a maximum of 2 weeks in dark.
10. Microscope Slides (76X26mm), Coverslip (Menzel-Glaser)
11. DABCO (1,4-diazabicyclo[2.2.2]octane) -220 mM DABCO dissolved in 90% glycerol, 10% PBS at pH 8.6
12. Mounting Oil

**Magnetic Sorting**

1. Rabbit polyclonal 6X His tag (Biotin) at 1mg/ml (Abcam)
2. Recombinant Human Tie2/Fc Chimera 500ng/μl (R&D Biosystems)
3. MACS Separation columns (MS columns) (Miltenyi Biotec)
4. Streptavidin Microbeads (MACS Miltenyi Biotec)
5. MACS Multistand magnet (MACS Miltenyi Biotec)

### **Flow Cytometry and FACS sorting**

1. Block solution 2 - 10%FCS-PBS solution
2. Rabbit polyclonal 6X His tag (Biotin) at 1mg/ml (abcam)
3. Streptavidin R-Phycoerythrin (Sigma)
4. Anti- FLAG M2 primary antibody (1mg/ml) (Sigma)
5. Mouse Ig Cy2 linked (Amersham Biosciences, GE Healthcare) 1mg/ml
6. Recombinant Human Tie2/Fc Chimera 500ng/ $\mu$ l (R&D Biosystems)
7. BD Falcon 5ml Polystyrene Round Bottom Tube (BD Biosciences)
8. Primocin 50mg/ml (InvivoGen)
9. BD FACS Clean Solution
10. 1 X PBS sterile as sheath fluid
11. Accudrop Beads (BD Biosciences)

## Methods

### Molecular Biology

#### 2.1 Polymerase Chain Reaction (PCR)

A typical 50 $\mu$ l PCR reaction contained 10X PCR buffer containing MgCl<sub>2</sub>, dNTP, Phusion High-Fidelity DNA polymerase, template (pcr2.1 Ang is an Ang1 plasmid or Ang2 plasmid available in the lab, unless otherwise stated) and template appropriate primers (For detailed list of primers see Appendix 2). Each reaction tube was then covered in an overlay of 30 $\mu$ l mineral oil. General method is outlined in Table 2.1 below.

Component	Final Concentration or Amount
Autoclaved dH <sub>2</sub> O	39 $\mu$ l
PCR buffer	1X
dNTP	10 mM dNTP
Template	10 ng
5' primer	1 pmol
3' primer	1 pmol
Enhancer	1 $\mu$ l (optional)
DNA polymerase	1 unit
Mineral Oil overlay	30 $\mu$ l

**Table 2.1: Components in a typical PCR reaction.**

The Perkin Elmer thermocycler was used for the amplification cycles. Cycling parameters (Unless stated otherwise) are outlined in Table 2.2 below.

Steps	Temperature	Time	Cycle
	92°C	3 min	1
Denaturation	94°C	1 min	25
Annealing	45-60°C	1 min	25
Extension	72°C	1 min	25
	72°C	10 min	1
Soak	4°C	$\infty$	-

**Table 2.2: General cycling parameters of a typical PCR reaction.**

For subsequent ligation reactions the PCR products were gel extracted and purified.

### **2.1.1 Colony PCR amplification of products for screening positive clones**

Colony PCR was performed on *E.coli* colonies (obtained on LB agar plates by transformation (Section 2.12) to quickly screen for plasmid inserts directly from the colonies. After cloning a particular insert into a vector the ligation reaction was transformed into competent *E.coli* cells. The colonies obtained after cloning were inoculated in the PCR reaction as template. The PCR reaction was setup as described above. Positive clones obtained for particular insert were taken forward for sequence analysis of the cloned insert.

### **2.1.2 PCR amplification of products for TA cloning**

For addition of A overhangs on the PCR product for TA cloning, the amplified PCR product was incubated at 72 °C for 10min with 5units/μl BioTaq DNA polymerase. The Phusion High-Fidelity DNA polymerase is a proof-reading polymerase which produces a blunt end PCR product for cloning. BioTaq DNA polymerase is an error prone polymerase, which is added to the PCR reaction after the error free amplification of the PCR product is complete. The error prone DNA polymerase is used only for the addition of A overhangs on the PCR product obtained from using the proof-reading polymerase.

## **2.2 Restriction Digest**

Restriction enzymes and their respective appropriate buffers were used as per manufacturers' instruction (Roche).

Typically 2 $\mu$ g DNA was digested in the reaction volume with 20 units of enzyme and 1X buffer concentration. The volume of the reaction was made up to 10 $\mu$ l for the vector and 30 $\mu$ l for the insert (unless stated otherwise) using autoclaved distilled RNase/DNase free water. The digestion was incubated at 37°C for 180 min.

For subsequent ligation reactions the digested products were extracted from an agarose gel and purified (Section 2.4.1).

Diagnostic restriction digests were done using 200ng DNA in the reaction volume with 2 units of enzyme and 1X buffer concentration. The volume of the reaction was made up to 10 $\mu$ l using autoclaved distilled RNase/DNase free water. The digestion was incubated at 37°C for 60 min. Digested products were analysed by electrophoresis (Section 2.3).

### **2.3 Agarose gel electrophoresis**

Electrophoresis of DNA samples (namely PCR products, restricted DNA required for subsequent ligation reactions and DNA samples for gel purification) was carried out using 0.5-1% [w/v] agarose TAE gels (Unless otherwise stated). The agarose was heated with 1X TAE buffer in a microwave and the dissolved agarose allowed to cool to approximately 50 °C. Ethidium bromide was added to the gel to a final concentration 0.2 $\mu$ g/ml. This was poured into an appropriate gel tray with comb taking care the gel was uniform and without air bubbles. The gel was allowed to cool and solidify before DNA samples were loaded. The DNA samples were mixed with 2 $\mu$ l loading dye and loaded in the wells. Electrophoresis was carried out at 150V, in gel running buffer 1X TAE. The DNA fragments were separated and then visualized using ultra violet light

trans-illuminator. The sample DNA fragment size was determined by comparison with 1 Kb Plus DNA ladder (Invitrogen) run alongside the samples on the same gel.

Electrophoresis of DNA samples from diagnostic restriction digests, diagnostic PCR and for TA cloning was performed using 1% [w/v] agarose TAE gels (unless stated otherwise). Ethidium bromide was added to the gel at a final concentration of 1 µg/ml. The images were captured using Multi Image Light Cabinet (Flowgen-Alpha Innotech Corporation).

## **2.4 Purification of PCR and restriction digest products**

**2.4.1 For ligation reactions**, the purification of PCR products and digested DNA was performed using MinElute Gel Extraction kit (Qiagen). The DNA fragment to be purified was separated on an agarose gel by electrophoresis (Section 2.3) and excised from the gel using a scalpel blade. Further purification was carried out as per manufacturers' instructions. The dephosphorylated vector ASGPR-pc3.1 was also purified using the MinElute Gel Extraction kit (Qiagen). Briefly, the excised agarose piece containing the DNA was dissolved in buffer QG provided in the kit. Isopropanol was added and the mix was applied to the purification column. The DNA bound to the membrane in the column and the contaminants were washed away. Purified DNA was eluted in 30 µl EB buffer.

## **2.4.2 Ethanol precipitation of linearised plasmid for electroporation**

To 500 µl of linearised plasmid DNA an equal volume of phenol-chloroform-isoamyl alcohol was added and mixture vortexed for 10sec to form emulsion. The emulsion was centrifuged for 1min at 13400g and the upper aqueous phase was collected in a fresh microfuge tube. To this an equal volume of chloroform was added

and mixture vortexed for 10sec to form an emulsion. The emulsion was centrifuged for 1min at 13400g and the upper aqueous phase was collected in a fresh microfuge tube to which 0.1 volume of 3M NaAcetate (pH5.2) and 2.5 volume of 100% ethanol was added. The mixture was vortexed and frozen at -20 °C overnight. The mix was then centrifuged at 13400g for 20min at 4 °C. The pellet was washed in 200 µl of cold 70% ethanol and centrifuged at 13400g for 10min at 4 °C. The pellet obtained was resuspended in 40µl sterile water to obtain DNA concentration of approximately 1µg/µl.

## **2.5 Dephosphorylation**

Purified and restricted vector ASGPR pc3.1 was dephosphorylated with 1U/µg DNA Shrimp Alkaline Phosphatase (SAP). Following digestion the purified vector was incubated for 20 min at 37°C with SAP and SAP 10X buffer diluted to 1X concentration in the reaction mixture. SAP was heat deactivated at 65°C for 15 min and dephosphorylated vector purified by MinElute Gel Extraction kit (Section 2.4.1).

## **2.6 Ligation**

### **2.6.1 Ligation of Ang1 and Ang2 FReD domains in the ASGPR-pc3.1 surface display vector (Geneart) and ligation of ASGPR-Ang1FReD and ASGPR-Ang2FReD sequence in pHypermute2 expression vector.**

The purified vector and the insert were quantitated by agarose gel electrophoresis alongside a standard DNA ladder. The ligation reaction contained the vector and insert in 1:3 ratio along with 1X Ligase Reaction buffer and 5 units of T4 DNA ligase. The reaction was carried out overnight at 8°C. A typical reaction is shown in the table 2.3 below.

	Reaction ( $\mu$ l)	Control1 ( $\mu$ l)	Control 2 ( $\mu$ l)
Insert (purified and eluted in 30 $\mu$ l buffer)	10	0	10
Vector (purified and eluted in 30 $\mu$ l buffer)	3	3	3
T4 DNA Ligase (1unit/ $\mu$ l)	5	5	0
10X Ligase buffer	3	3	3
Distilled water	9	19	14

**Table 2.3: General outline of the components of a typical ligation reaction.**

### **2.6.2 Ligation of ASGPR in the pcDNA3.1V5-His-TOPO expression vector (Invitrogen).**

Ligation was performed using the manufacturers' instructions provided with the pcDNA3.1V5-His-TOPO TA Expression Kit. Briefly, 4 $\mu$ l of the PCR insert was mixed with 1 $\mu$ l of vector and 1 $\mu$ l of salt solution. The reaction was carried out for 10 min at 23°C.

### **2.6.3 Ligation of FReD and P domain in the pSecTag/FRT/V5-His TOPO TA expression vector (Invitrogen).**

Ligation was done using the manufacturers' instructions provided with the pSecTag/FRT/V5-His TOPO TA Expression Kit. Briefly, 4 $\mu$ l of the PCR insert was mixed with 1 $\mu$ l of vector and 1 $\mu$ l of salt solution. TA cloning bench ligation was performed for 60 min at RT.

### **2.6.4 Ligation of DNA into the pCR4 TOPO sequencing vector (Invitrogen).**

Ligation was done using the manufacturers' instructions provided with the TOPO TA Cloning Kit for Sequencing. Briefly, 4 $\mu$ l of the PCR insert was mixed with 1 $\mu$ l of sequencing vector and 1 $\mu$ l of salt solution. The reaction was carried out for 10 min at 23°C.

## 2.7 Site Directed Mutagenesis

Site Directed Mutagenesis was performed as per the manufacturers' instructions provided with the QuikChange Site Directed Mutagenesis Kit (Stratagene). The basic procedure utilized a dsDNA vector (50ng) with the desired insert, 10X reaction buffer (5  $\mu$ l), dNTP (2.5mM each NTP), 1 $\mu$ l of *PfuTurbo* DNA polymerase (2.5units/ $\mu$ l) and two synthetic oligonucleotide primers (125ng each) containing the desired mutation. The total reaction volume was made up to 50 $\mu$ l with ddH<sub>2</sub>O. The oligonucleotide primers, each complement opposite strands of the plasmid vector, and were extended during temperature cycling with *PfuTurbo* DNA polymerase. Cycling parameters (Unless stated otherwise) are outlined in Table 2.4 below.

Steps	Temperature	Time	Cycle
	95°C	30 sec	1
Denaturation	95°C	30 sec	16
Annealing	55-60°C	1 min	16
Extension	68°C	8min	16
Soak	4°C	$\infty$	-

**Table 2.4: General cycling parameters of a typical site directed mutagenesis reaction.**

*PfuTurbo* replicates both plasmid strands with high fidelity without displacing the mutant oligonucleotides. The mutant plasmid containing staggered nicks was obtained. Following temperature cycling the product was treated with DpnI at 37°C for 1hour, in which the methylated parental plasmid (DNA from most *E.coli* is dam methylated) was digested whereas the mutant plasmid was not. 4 $\mu$ l of the nicked mutant plasmid product was then transformed into competent cells in which the nicks in the mutant plasmid were repaired.

## 2.8 RNA Extraction

RNA was extracted from cells as per the manufacturers' instructions provided with the RNeasy Mini Kit (Qiagen). Briefly,  $10^6$  cells were centrifuged and resuspended in RLT buffer containing  $\beta$ -mercaptoethanol to disrupt the cells. The homogenized lysate was mixed with 1 volume of 70% ethanol and the mix applied to RNeasy mini column. The column was centrifuged and the flow through discarded. The column was washed and the silica-gel membrane bound RNA eluted with 30 $\mu$ l RNase-free water.

## 2.9 RT PCR

RT PCR was done on the extracted RNA as per the manufacturers' instructions provided with the RETROscript Kit (Ambion).

Briefly, approximately 2 $\mu$ g of total RNA was heat denatured at 85 °C in the presence of oligo (dT). RT buffer, dNTP mix, RNase Inhibitor and Reverse Transcriptase were added and the components were incubated at 55°C for 1hr (1 cycle). The transcriptase was inactivated at 92°C for 10 min.

## 2.10 Plasmid preparation

Bacterial colonies obtained after transformation (Section 2.12) were picked with a sterile pipette tip and inoculated into 5 ml LB broth with ampicillin (100  $\mu$ g/ml) selection and grown overnight at 37 °C on a shaker at 200 rpm. 4 ml of the bacterial culture was centrifuged at 3500g for 5min and harvested using the QIAprep Mini Prep kit (Qiagen) according to manufacturers' protocol. Briefly, the cell pellet was suspended in buffers P1, P2 and N3 supplied with the kit and the precipitate formed was discarded.

The flow through was applied to the QIAprep spin column to which the DNA bound. The plasmid sample was eluted in 30  $\mu$ l elution buffer.

### **2.10.1 Endotoxin-free plasmid Maxi prep.**

Endotoxin free plasmid preparation for electroporation was performed as per the manufacturers' instructions provided with the EndoFree Plasmid Maxi Kit (Qiagen). Briefly, the cell pellet was suspended in endotoxin free buffers P1, P2 and P3 supplied with the kit and the lysate formed was poured onto the QIAfilter Cartridge. The filtered lysate was applied to the QIAGEN-tip 500 column and allowed to flow through by gravity. The DNA was eluted with Buffer QN and precipitated with isopropanol and washed in ethanol. The plasmid sample was eluted in 300 $\mu$ l TE buffer.

DNA concentration for the plasmid preparation was determined by measuring the absorbance at 260 nm on a spectrophotometer.

### **2.10.2 Spectrophotometric analysis of plasmid DNA**

1  $\mu$ l of the eluate was diluted in 500 $\mu$ l of sterile distilled water. The DNA concentration was determined by recording the absorbance at 260 nm using the UV-Visible spectrophotometer (Shimadzu). The concentration was calculated assuming 50  $\mu$ g/ml of DNA measured in a cuvette with 1 cm path length to have an absorbance of 1. The purity of the DNA preparation was determined by the  $A_{260}/A_{280}$  ratio, which is 1.8-2.0 for pure DNA (Sambrook, Fritsch & Maniatis 1989).

## 2.11 Sequencing

Plasmid samples with FReD domain cloned into the surface display vector ASGPR-pc3.1 and all Ang2-FReD mutants were sent for sequencing at the Protein and Nucleic Acid Chemistry Laboratory (PNACL), University of Leicester. Plasmid samples were supplied at a concentration of 0.5 $\mu$ g in 8 $\mu$ l pure distilled DNase free water and sequenced using the T7 primer at a concentration of 1 pmol/ $\mu$ l provided at PNACL.

Plasmid samples with FReD domain and P domain cloned into the secretory pSecTag/FRT/V5-His TOPO expression vector was sent for sequencing at Macrogen, South Korea. Plasmid samples were supplied at a concentration of 100ng/ $\mu$ l in pure distilled DNase free water and sequenced using the T7 primer (1 pmol/ $\mu$ l) provided at Macrogen.

Plasmid samples with ASGPR-Ang1 FReD domain and ASGPR-Ang2 FReD domain cloned into the pHyperm2 vector were sent for sequencing at the Protein and Nucleic Acid Chemistry Laboratory (PNACL), University of Leicester. Plasmid samples were supplied at a concentration of 0.5 $\mu$ g in 8 $\mu$ l pure distilled DNase free water and sequenced using the ASGPRNheIfwd and Ang1EcoRVrev or Ang2EcoRVrev primers (1 pmol/ $\mu$ l).

Clones from the iterative sorts cloned into the pCR4 TOPO sequencing vector (Invitrogen) were sent for sequencing at Macrogen, South Korea. Plasmid samples were supplied at a concentration of 100ng/ $\mu$ l pure distilled DNase free water and sequenced using the M13 fwd and M13 rev primers at a concentration of 1 pmol/ $\mu$ l provided at Macrogen.

## **Microbiology**

### **2.12 Transformation**

50µl competent cells were thawed on ice and added to pre-chilled 14 ml BD Falcon polypropylene round bottom tubes. 5µl of ligation reaction or 50-500ng plasmid DNA was used for transforming the cells. This was incubated with cells on ice for 30 min followed by a heat shock for 90 sec at 42 °C. The reaction was then incubated on ice for 2min. 200µl of pre-heated (42°C) LB broth (without ampicillin) was added and the reaction incubated at 37°C for 1 hr on a shaker at 225rpm. The cells were centrifuged and resuspended in 60µl LB broth. This volume was plated onto pre-warmed ampicillin selective LB agar plates. The control ligation reactions were also transformed and plated onto another ampicillin selective LB agar plates. The plates were incubated at 37°C for 18 hours.

### **2.13 Storage of Bacterial Culture**

Glycerol stocks were made for long term storage of bacterial cultures transformed with plasmid of interest. 850µl of culture grown overnight in LB broth was added in a labelled freezing vial under sterile condition. 150µl of sterile glycerol was added to the vial and the mixture was vortexed. The culture was snap freezed in liquid nitrogen and stored at -80 °C.

## **Tissue Culture**

### **2.14 Cell Culture**

#### **2.14.1 Chinese Hamster Ovary cells (CHO)** [originally from ECACC]

CHO cells were grown in 6 well plates and complete media to 60% confluency at 37°C /5%CO<sub>2</sub>. For passaging the CHO cells monolayer were washed once with PBS and incubated with 2 ml 2X TE for 5 min at 37°C and the flask was tapped gently to facilitate detachment of the cells from the flask surface. The action of TE was arrested by addition of complete media to a final volume of 10 ml. The suspension was centrifuged at 200g for 7 min and the sedimented cells were resuspended in 1ml of complete media. An appropriate volume of this suspension was seeded in flasks or 6 well plates according to requirement.

#### **2.14.2 Human Embryonic Kidney 293 Cells (HEK)** [originally from ECACC]

HEK cells were grown in 6 well plates and complete media to 60% confluency at 37°C /5%CO<sub>2</sub>. For passaging the HEK cells monolayer was incubated with 1 ml 1X TE for 1 min at 37°C and the flask was tapped gently to facilitate detachment of the cells from the flask surface. The action of TE was arrested by addition of complete media to a final volume of 10 ml. The suspension was centrifuged at 200g for 7 min and a pellet of the cells obtained was resuspended in 1ml of complete media. An appropriate volume of this suspension was seeded in flasks or 6 well plates according to requirement.

#### **2.14.3 Ramos (Human Burkitt's lymphoma cell line)** [originally from ECACC]

Ramos cells were grown in T80 flasks in suspension at 37°C /5%CO<sub>2</sub>. The cells were maintained by subculturing every 3 days by diluting 1 ml of cell suspension in

20ml of complete media. The cells were subcultured by diluting 1ml of cell suspension in 10ml of complete media a day before nucleofection for optimized transfection.

**2.14.4 DT40 chicken B cell line** [This engineered DT40 cell line was kindly supplied by Dr. Jean-Marie Buerstedde]

DT40 cells were grown in T80 flasks in suspension at 41°C/5%CO<sub>2</sub>. The cells were maintained by subculturing every 3 days by diluting 1 ml of cell suspension in 30ml of complete media. The cells were subcultured by diluting 1 ml of cell suspension in 10ml of complete media a day before electroporation for optimized transfection.

## **2.15 Transfection**

### **2.15.1 Lipofection for CHO and HEK cell lines**

The cells were plated to obtain a confluency of 60-70% on the day of transfection. Transfection efficiency was determined by Green Fluorescent Protein (GFP) transfection carried out using varying amounts of GFP DNA and Lipofectamine Transfection Reagent. The most efficient conditions were determined by quantifying fluorescent cells from the GFP transfection. Accordingly 2µg of DNA and 5µl of lipofectamine 2000 in 100µl of OptiMEM I was used in making the transfection complex which was incubated at room temperature for 30 min. This complex was added to a 60% confluent 35mm well containing 600µl of serum free medium. This was incubated for 3 hrs at 37°C in 5%CO<sub>2</sub> incubator. The complex was then replaced by 2 ml of complete media and incubated overnight at 37°C/5%CO<sub>2</sub>. The media obtained post transfection was collected for analysis of extracellular secreted forms of protein or cell lysates made to analyze extracellular bound forms of protein. Alternatively stable cell lines were generated by applying appropriate antibiotic selection pressure.

For transfecting 80cm<sup>2</sup> flasks, 10µg of DNA and 25 µl of lipofectamine 2000 in 1.5ml of OptiMEM I was used in making the transfection complex and applied to T80 flask with 5.5ml of complete media.

### **2.15.2 Nucleofection of Ramos cells**

Ramos cells are sub-cultured a day before transfection. Approximately 2 million cells were centrifuged at 200g and re-suspended in 100µl nucleofector solution and 2µg plasmid was added. Nucleofection was performed using the Cell Line Nucleofector Kit V program O-06. The procedure was carried out according to manufacturers' protocol. 24 hours post nucleofection, the GFP control cells were analyzed by fluorescence microscopy and transfection efficiency of 10% was observed.

### **2.15.3 Nucleofection of DT40 for transient transfection**

The protocol was similar to that used for Ramos cells. Nucleofection was performed using the Cell Line Nucleofector Kit T program B-09. The procedure was carried out according to manufacturers' protocol. 24 hours post nucleofection, the GFP control cells were analyzed by fluorescence microscopy and transfection efficiency of 80% was observed.

### **2.15.4 Electroporation of DT40 cells to generate stable transfectants**

10 million DT40 cells were centrifuged at 200g for 5min and resuspended in 800 µl PBS (RT). 40 µg linearised ethanol precipitated plasmid was added and the mix transferred to electroporation cuvette and electroporated at 700V and 25 µF. The electroporated cells were transferred to 19.5ml complete medium pre-incubated to 41 °C.

## **2.16 Generating stable cell line.**

### **2.16.1 Ramos and HEK cells**

The cells transfected with the Ang1FReD- ASGPR-pc3.1 and Ang2FReD- ASGPR-pc3.1 construct which codes for the G418 resistance gene in the construct vector sequence can be used to obtain a stable cell line. Accordingly G418 selection pressure at a concentration of 600 $\mu$ g/ml was applied 48 hours post-transfection to the transfected cells and a batch of non transfected cells which served as a control. 10 days after addition of G418 antibiotic all the cells in the non transfected cell population died but the transfected cells which incorporated the FReD- ASGPR-pc3.1 construct by genetic recombination were G418 resistant and continued to grow in the presence of the selection pressure.

### **2.16.2 Generating DT40 stable cell line.**

The cells transfected with the Ang2 FReD- ASGPR-pc3.1-Hypermut construct which encodes for the puromycin resistance gene in the construct vector sequence can be used to obtain a stable cell line. Accordingly 24 hours post transfection the selective antibiotic puromycin was applied at a concentration of 1 $\mu$ g/ml of cell suspension. The transfected cells were diluted to theoretically obtain 1 cell per well of a 96 well plate and incubated at 41 °C, 5%CO<sub>2</sub> for up to 10 days without changing the media. 10 days post transfection drug resistant colonies were visible. Drug resistant stable clones were picked and cultured. Stable cells are maintained under the selective pressure 0.5 $\mu$ g/ml puromycin. No colonies appeared in the non-transfected control cells. Transfected cells which incorporated the Ang2 FReD- ASGPR-pc3.1-Hypermut construct by genetic

recombination were puromycin resistant and continued to grow in the presence of the selection pressure.

## **2.17 Collection of culture medium and cell lysate**

### **2.17.1 Immunoprecipitation to concentrate epitope tagged secreted proteins**

For immunoprecipitation of FLAG-tagged soluble proteins, monoclonal anti-FLAG M2 covalently attached to agarose beads (Sigma) was used. 20 $\mu$ l of the slurry was centrifuged and washed thrice with 1ml of PBS solution and resuspended in 20 $\mu$ l PBS before addition of the (1ml or 5ml as stated) media containing soluble protein. The reaction was incubated in the cold room at 4°C for 2 hours on the spiramix (Denley). It was then centrifuged at 13400g for 2mins and the bead pellet was washed thrice with 500 $\mu$ l of PBS and washings discarded. The pellet was then resuspended in 20 $\mu$ l of reducing Laemmli Sample Buffer (SB) containing 100mM DTT. The sample was heated for 6 min at 95°C to dissociate the immune-complex from the beads and centrifuged at 13400g for 1min. The proteins in the supernatant were analyzed by SDS-PAGE (Section 2.18).

### **2.17.2 Preparation of cell lysate to analyze extracellular cell bound protein in HEK293 and CHO cells**

The medium was removed and cells washed in sterile PBS. The cell were lysed in 70  $\mu$ l lysis buffer and the debris-free lysate was prepared in 50 $\mu$ l 2X Reducing SB for a 35 mm tissue culture well. The attached cells were scraped off with a sterile tip. The lysate was heated at 95°C for 6 min, sonicated for 30 sec and centrifuged at 13400g for 1min. The proteins in the supernatant were analyzed by SDS-PAGE (Section 2.18).

### 2.17.3 Preparation of Ramos cell lysate

A million cells were centrifuged at 200g. The medium was removed and cells washed in sterile PBS. The cell lysate was prepared by addition of 70µl cold lysis buffer by vigorous mixing and was incubated on ice for 5 min. This was mixed again by vortexing and incubated on ice for additional 5 min. The cell debris was separated by centrifuging at 13400g for 2 min and the supernatant was collected in a fresh microfuge tube. 20µl 2X Reducing SB was added and the lysate was sonicated for 30 sec and heated at 95°C for 6 min. The lysate was centrifuged at 13400g for 1min and the proteins analyzed by SDS-PAGE (Section 2.18).

### 2.18 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE was done using Biorad apparatus. The acrylamide gel was prepared as shown in the Table 2.5.

Resolving gel	12 %	15%
Components	Volume	Volume
30% Acrylamide/0.8%Bisacrylamide	8ml	10ml
2 M Tris pH 8.8	3.7ml	3.7ml
dH <sub>2</sub> O	7.9ml	6.3ml
20% SDS	100 µl	100 µl
10 % Ammonium Per Sulphate (APS)	134 µl	134 µl
N,N,N',N'-Tetramethylethylenediamine (TEMED)	14 µl	14 µl

#### 5 % Stacking gel

Components	Volume
30% Acrylamide/0.8%Bisacrylamide	3.3ml
1 M Tris pH 6.8	2.5ml
dH <sub>2</sub> O	13.7ml
20% SDS	100 µl
10 % Ammonim Per Sulphate (APS)	200 µl
N,N,N',N'-Tetramethylethylenediamine (TEMED)	20 µl

**Table 2.5: General outline of components in the resolving and stacking gel used for SDS-PAGE.**

The cell lysate and the media samples prepared in SB were run on SDS gel in 1X running buffer at 150V for 1.5 hours alongside 10 µl Protein Kaleidoscope Standard molecular mass markers.

## **2.19 Western Blotting and antibody probing**

The separated proteins were transferred from SDS-PAGE gel onto a Hybond-ECL nitrocellulose membrane by electrotransfer as described by Towbin et al. 1979 (Towbin, Staehelin & Gordon 1979). Proteins were transferred at 230-250mA for 2 hr in 1X transfer buffer. The nitrocellulose membrane was then probed as described.

### **2.19.1 Detection of FLAG-tagged proteins**

The membrane was blocked with 5% blocking solution (non-fat dry milk in TBS) for 1 hr at RT on rocker. The membrane was then probed with anti-FLAG M2 antibody diluted to 1:1000 (10µg/ml) in 1X TBS 0.1%TX-100 and incubated for 30min at RT on rocker. The membrane was washed thrice (10min per wash) with the 1X TBS 0.1%TX-100 solution and labelled by incubating for 1 hr at RT on rocker with the anti-mouse horseradish peroxidase-linked conjugated secondary antibody at a dilution of 1:2500 in 1X TBS 0.1%TX-100. The blots were again washed thrice (15min per wash) in the 1X TBS 0.1%TX-100 solution. Conjugated secondary antibody was visualized by chemiluminescent detection. The membrane was soaked for 2 min in this solution at room temperature. The membrane was then wrapped in cling film and exposed to Chemiluminescence detection film (Kodak) for 1-5mins depending on the intensity of the bands. The film was developed and the protein markers were aligned to the membrane and marked.

To re-probe a membrane, the membrane was stripped of antibody in 1X Re-blot solution (Chemicon International) for 15min. The membrane was then washed in 1X TBS 0.1%TX-100 solution, blocked and probed with appropriate primary and secondary antibodies (as described above).

### **2.19.2 Detection of Tie2 ectodomain**

The membrane was blocked with 5% blocking solution (non-fat dry milk in TBS) for 1 hr at RT on rocker. The membrane was then probed with anti-Tie2 primary antibody diluted to 1:1000 in 1X TBS-0.1%TX-100 and incubated for 16 hr at 4°C on rocker. Blot was washed thrice (10min per wash) with the 1X TBS 0.1%TX-100 solution and labelled anti-goat horseradish peroxidase-linked conjugated secondary antibody at a dilution of 1:1000 in 1X TBS 0.1%TX-100 (1 hr at RT on rocker). The blots were again washed thrice (15min per wash) in the 1X TBS 0.1%TX-100 solution. Conjugated secondary antibody was visualized by chemiluminescent detection. The membrane was soaked for 2 min in this solution at room temperature. The membrane was then wrapped in cling film and exposed to Chemiluminescence detection film (Kodak) for 1-5mins depending on the intensity of the bands. The film was developed and the protein markers were aligned to the membrane and marked.

## **Immunohistochemistry**

### **2.20 Immunofluorescence**

#### **2.20.1 Immunofluorescence to detect cell surface FLAG-tagged ASGPR protein in CHO cell line.**

Cells were added to wells of a 6-well plate containing autoclaved glass coverslips to allow them to adhere and grow on the coverslips. These wells were transfected with

2µg/well plasmid DNA (Section 2.15.1) and kept in complete media overnight. The transfected cells grown on the coverslips were used for immunofluorescence. The media was removed and the cells were washed thoroughly with PBS. The cells on the coverslip were fixed by addition of 2ml of 4% paraformaldehyde in each well. This reaction was allowed to incubate at RT for 20min. Once the cells were fixed, the coverslips were transferred to a new labelled plate with the cells facing up. The coverslips were rinsed gently with TBS and non specific sites blocked with 2% BSA-TBS solution. The blocking reaction was carried out for 10min on a rocker. The cells were incubated for 1 hr at RT with the primary anti- FLAG M2 primary antibody diluted to 1:50 in TBS 2%BSA blocking solution. The cells are washed three times with TBS before incubation with the secondary Cy2 conjugated anti-mouse antibody. The secondary antibody was diluted 1:100 in blocking solution1 and incubated with the cells at 37 °C for 45 min. The coverslips were washed twice with 0.1%TX-100-TBS solution and once with TBS. The cells on the coverslips were mounted in 10µl DABCO on a glass slide with the cell side facing down and sealed with nail polish. The cells were visualized under Nikon Microscope. The excitation wavelength is 492nm and emission wavelength is 510nm for Cy2.

### **2.20.2 Immunofluorescence to detect cell surface FLAG-tagged ASGPR-Ang-FReD protein in transient or stable DT40, Ramos or HEK293 (in suspension) cell lines.**

Cells transfected to express FLAG-tagged Ang-FReD protein on the cell surface were stained. Approximately  $0.5-1.0 \times 10^6$  cells were centrifuged at 200g for 5 min. The medium was removed and the cells were washed thoroughly with PBS. The cells were resuspended in primary anti- FLAG M2 primary antibody (final concentration 20 µg/ µl) diluted to 1:50 in block solution 2 and incubated on the shaker at 4°C for 30 min.

The suspension was centrifuged at 200g and washed thrice with 500  $\mu$ l block solution 2. The cells were resuspended in secondary antibody anti-mouse Cy2 (final concentration 20  $\mu$ g/  $\mu$ l) diluted to 1:50 in block solution 2 and incubated on the shaker at 4°C for 30 min. The suspension was centrifuged at 200g and washed thrice with 500 $\mu$ l block solution 2. The cells were fixed in 100 $\mu$ l of 4% paraformaldehyde for 20 min on ice. The suspension was centrifuged and resuspended in 30 $\mu$ l PBS and 5 $\mu$ l of the suspension mounted on a glass slide. A coverslip was gently placed on this and sealed with nail polish. The cells were visualized under the oil immersion lens of Nikon Microscope. The excitation wavelength for Cy2 is 492 nm and emission is at 510 nm in the green region of the visible spectrum.

### **2.20.3 Immunofluorescence staining for observing Tie2 binding in stable DT40 cells expressing cell surface FLAG-tagged ASGPR-Ang-FReD protein**

Approximately  $0.5-1.0 \times 10^6$  cells DT40 cells expressing cell surface FLAG-tagged ASGPR-Ang-FReD protein were centrifuged at 200g for 5 min. The medium was removed and the cells were washed thoroughly with PBS. The cells were incubated with 5nM of soluble Tie2 ectodomain diluted to 50 $\mu$ l in block solution 2 for 1 hr at 4 °C on a shaker. The cells were washed thoroughly with block solution 2. The cells were resuspended in primary anti- FLAG M2 (final concentration 20  $\mu$ g/  $\mu$ l) and biotinylated anti-His (final concentration 20 $\mu$ g/  $\mu$ l) primary antibodies diluted in 50 $\mu$ l block solution 2 and incubated on the shaker at 4 °C for 30 min. The suspension was centrifuged at 200g and washed thrice with 500 $\mu$ l block solution 2. The cells were resuspended in 1  $\mu$ l secondary antibody anti-mouse Cy2 and 2  $\mu$ l streptavidin R-PE secondary antibody diluted in 50 $\mu$ l block solution 2 and incubated on the shaker at 4 °C for 30 min. The suspension was centrifuged at 200g and washed thrice with 500 $\mu$ l block solution 2. The

cells were fixed in 100 $\mu$ l of 4% paraformaldehyde by incubation for 20min on ice. The suspension was centrifuged and resuspended in 30 $\mu$ l water. 5 $\mu$ l of the suspension was mounted on a glass slide. A coverslip was gently placed on this and sealed with nail polish. The cells were visualized under the oil immersion lens of Nikon Microscope. The excitation wavelength for Cy2 is 492 nm and fluoresces (510 nm) in the green region of the visible spectrum. The excitation wavelength for R-PE is 565 nm and emission wavelength is 575 nm in the red region of the visible spectrum.

### **2.21 Magnetic sorting on ASGPR-Ang2FReD expressing stable DT40 cells.**

Magnetic sorting was performed under sterile conditions as per manufacturers' instructions. Briefly, 10 million cells were counted and collected by centrifugation at 200g for 5min. The pellet was washed once in 1ml PBS. Cells were centrifuged again at 200g and resuspended in 50 $\mu$ l PBS and incubated in the appropriate concentration of Tie2 for 1 hr at 4°C on shaker. The cells were washed and then incubated for 30min at RT with biotinylated anti-His<sub>6</sub> diluted 1:50 in PBS solution. Cells were washed and centrifuged at 200g and resuspended in 90  $\mu$ l PBS /10<sup>7</sup> cells and 10  $\mu$ l magnetic beads /10<sup>7</sup> cells. The magnetic bead-cell mixture was mixed and incubated at 4 °C for 15 min. The MS column was prepared on the magnet by rinsing the column with 500 $\mu$ l of PBS. The magnetic bead-cell mixture was applied to the column and allowed to flow through on the magnet for 1min. The column was washed 6 times with 500 $\mu$ l PBS while the column was still on the magnet. The column was removed from the magnet and 1ml PBS was flushed through the column using the plunger. Flushed cells were counted on a haemocytometer and these sorted cells were cultured in 10ml complete media.

## **Flow Cytometry and FACS sorting**

### **2.22 Flow Cytometry**

#### **2.22.1 Flow cytometry of transfected Ramos cells**

For flow cytometry analysis the suspension of cells were stained as in Section 2.20.2 but the cells were not fixed in paraformaldehyde, instead they were resuspended in RPMI complete medium after completion of the staining protocol. The cells were then analyzed by flow cytometer (BD FAC Scan).

#### **2.22.2 Flow cytometry of stable DT40 cells to determine $K_d$**

DT40 stable cells expressing Ang2-FReD protein were pelleted by centrifugation at 200g for 5min. The pellet was washed once in 1ml PBS. Cells were centrifuged again at 200g and resuspended in 50 $\mu$ l block solution 2 (10%FCS-PBS) and incubated in the various concentration of soluble Tie2 for 1 hr at 4°C on shaker (volumes were adjusted to obtain a 10 times molar excess of Tie2 for 50000 cells, specially at lower concentrations). All subsequent steps were performed keeping the sample in the dark to avoid photo-bleaching. The cells were then washed and then incubated for 30min at 4 °C with 1:50 biotinylated anti-His in block solution 2. The cells were then washed and incubated for 30min at 4 °C with 1:25 Streptavidin R-Phycoerythrin in PBS. The cells were then washed and resuspended in 500  $\mu$ l block solution 2 and transferred to FACS tubes. Ten thousand events for each concentration were analysed on a BD FACS Aria sorter. The 488nm Argon laser was used for excitation and the emission wavelength for R-PE is 575nm. The stained cells were corrected against non transfected controls to account for non specific sticking of chromophore to the cell surface. The FCS files were analysed in WinMDI and the geometric mean for the R-Phycoerythrin chromophore

was obtained. To determine  $K_d$  values non-linear regression analysis was performed of plots of geometric mean fluorescence versus varying concentrations of Tie2 using Graphpad Prism 5.0.

### **2.22.3 Flow cytometry with HEK cell line to determine $K_d$ for mutants N467S, S480P, G428T (Chapter 4) and mutants *10c*, *FACSD*, *FACSi* (Chapter 6)**

HEK cells transfected with wild-type and mutant plasmids (Section 2.15.1) were detached from the wells and pelleted by centrifugation at 200g for 5min. The pellet was washed once in 1ml PBS. Cells were centrifuged again at 200g and blocked in 100 $\mu$ l block solution<sup>3</sup> for 30min at 4 °C. Cells were centrifuged and incubated with various concentrations of soluble Tie2 for 1 hr at 4 °C on shaker (volumes were adjusted to obtain a 10 times molar excess of Tie2 for 50000 cells, specially at lower concentrations). All subsequent steps were performed keeping the sample in dark to avoid photo-bleaching. The cells were washed and then incubated for 15min at 4 °C with 1:500 biotinylated anti-His in PBS. The cells were washed and then incubated for 15min at 4 °C with 1:25 Streptavidin R-Phycoerythrin in PBS. The cells were washed and resuspended in 500  $\mu$ l PBS-FCS and transferred to FACS tubes. 10000 events for each concentration were analysed on a BD FACS Aria sorter. The 488nm Argon laser was used for excitation and the emission wavelength for R-PE is 575nm. The stained cells were corrected against non transfected controls to account for non specific sticking of chromophore to the cell surface. The FCS files were analysed in WinMDI and the geometric mean for the red fluorescence of R-Phycoerythrin chromophore was obtained. To determine  $K_d$  values non-linear regression analysis was performed of plots of geometric mean fluorescence versus varying concentrations of Tie2 using Graphpad Prism 5.0.

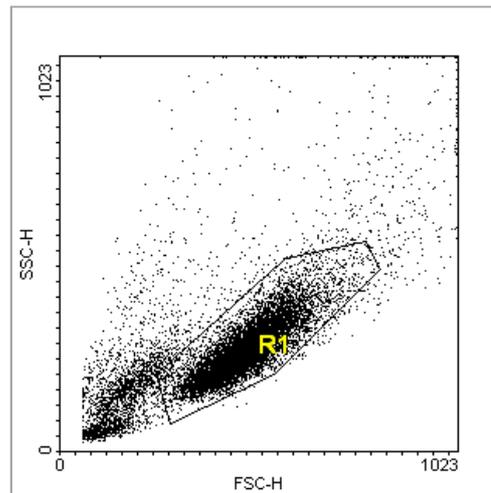
### **2.23 Fluorescence Activated Cell Sorting (FACS)**

DT40 stable cells were stained as described in section 2.20.2 or 2.20.3 but under sterile conditions. After labelling the cells with chromophore, they were resuspended in block solution 2 at a density of  $25 \times 10^6$  cells/ml. The cells were then analyzed and sorted by Fluorescence Activated Cell Sorter (BDFACS Aria SORP II). The machine was rinsed thoroughly with FACS Clean solution and sterile 1 X PBS was used as sheath fluid. The stream was adjusted and the machine calibrated by drop delay. A 100 $\mu$ m nozzle was used and the 488nm Argon laser was used for excitation of the chromophores Cy2 and phycoerythrin. The dyes were compensation corrected for spill-over and live cells were gated prior to each sorting (Section 2.24). The sort was performed at 500-1000 events/sec and the sorted cells were cultured in complete media containing antibacterial anti fungal agent Primocin used at a concentration of 100 $\mu$ g/ml.

### **2.24 Controls for Flow Cytometry analysis and Fluorescence Activated Cell Sorting (FACS) of cell populations.**

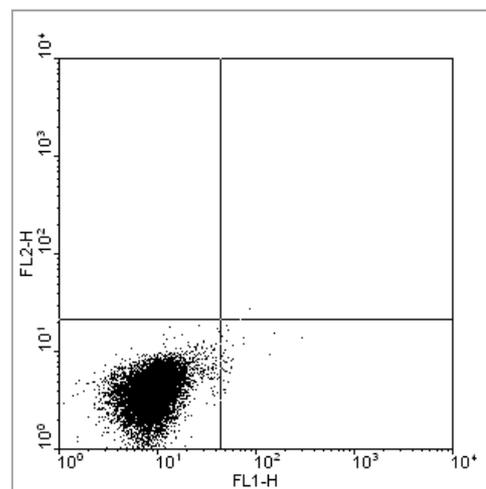
Prior to flow cytometry and FACS of cell populations, controls were used for setting up the machines.

To obtain meaningful data, the cell population was gated to exclude the dead cells from the viable cells. The dead cells can be distinguished from the viable ones, by analyzing the sample through the flow cytometer under the forward and side scatter settings which distinguish the size and complexity of the cells in the sample. Figure 2.1 is a representation of a typical dot plot obtained when gating of viable cells is performed.



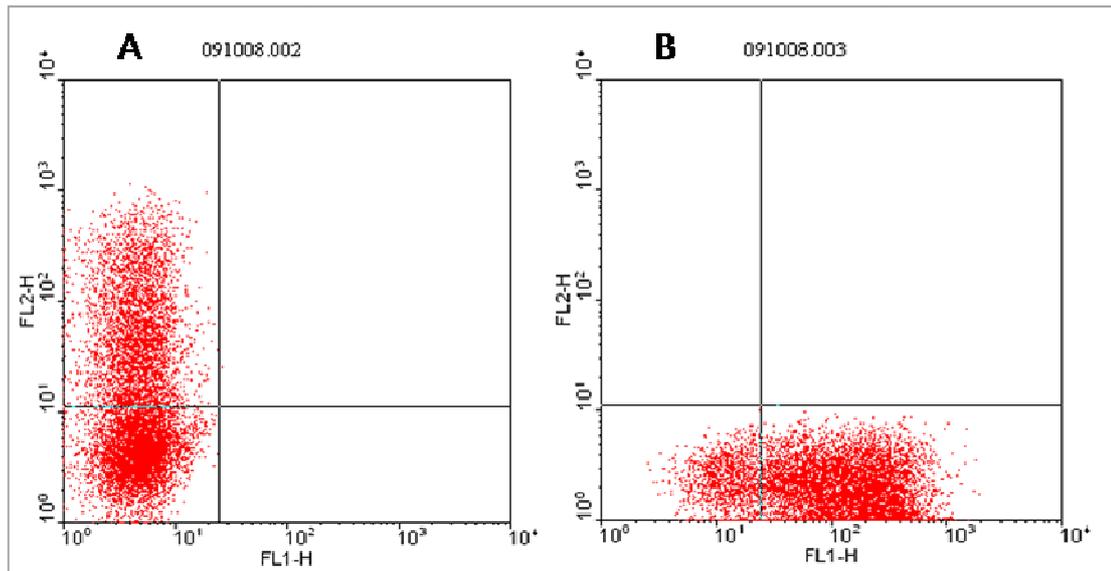
**Figure 2.1: Gating to distinguish dead cells from viable cells prior to obtaining flow cytometric data.** The region enclosed in the polygon R1 represents the live cells. The data shown in all histograms and dotplots are representative of the cells gated in region R1.

Once the viable cells were gated, four controls samples were put through the flow cytometer. The non stained cells served as control to account for autofluorescence from cells (Figure 2.2). Even when no fluorescent tag or stain is added to cells, they fluoresce. This autofluorescence comes from normal cell components which fluoresce, such as riboflavin and flavoproteins (Aubin 1979).



**Figure 2.2: Non stained cells, used to account for autofluorescence of the cells prior to obtaining flow cytometric data.** The flow cytometric parameters were adjusted to compensate for the autofluorescence so that the non stained cells were positioned only in the lower left quadrant.

Single stain controls were used as fluorescence compensation to account for "spillover" of the emission of one fluorochrome into the emission from the other fluorochrome. The single stained controls were cells stained for FLAG epitope tag by Cy2 fluorochrome and for Tie2 by the R-phycoerythrin fluorochrome. Figure 2.3 is a representation of the compensation controls for flow cytometric analysis.



**Figure 2.3: Compensation to correct the detected "spillover" from the emission of the Cy2 fluorochrome into the detector designed to collect the emission from R-phycoerythrin fluorochrome and vice a versa.** A) Compensation control for the red spillover from the R-phycoerythrin fluorochrome were cells stained for bound Tie2. B) Compensation control for the green spillover from the Cy2 fluorochrome were cells stained for FLAG tagged FReD protein.

In addition to the single stain compensation controls, cells were also analysed with FMO (Fluorescence Minus One) controls. FMO controls leave out one reagent at a time of all the reagents used in a flow cytometry experiment. This was performed using non-transfected cells and transfected cell populations. The reagents in the FMO controls did not display any fluorescence with the non-transfected cells. Also they did not show any cross-reactivity between themselves.

## Chapter 3

# Determination of Minimal Receptor Binding Domain in Angiopoietin-1

### Introduction

As discussed in Section 1.1.3 the angiopoietin structure consists of the N-terminus superclustering domain, the coiled coil domain (CCD) and the C-terminal fibrinogen related domain (FReD), first characterised by Davis et al. (1996). The Ang1 structure was further characterised by Davis et al. (2003) where they analysed binding of Tie2 to Ang1 variants. This study was done using Ang receptor binding domain (RBD) with surrogate multimerization motifs, to conclude that the FReD domain was sufficient for Tie2 binding but not for activation of the receptor.

Neither the crystal structure of Ang1 nor Ang1-Tie2 complex has been elucidated. However, Ang1 and Ang2 have a highly conserved structure showing approximately 60% identity. Ang1 and Ang2 have the same binding site in the Tie2 protein (Fiedler et al. 2003). Thus the crystal structure of Ang2-Tie2 complex can be used as a start point to elucidate Ang1-Tie2 interactions (Barton et al. 2006).

It is suggested that the precise residues for Ang1:Tie2 interaction and ligand-recognition are in the carboxy-terminal half of the fibrinogen domain termed as the P domain as determined by the crystal structure of Ang2 (Barton, Tzvetkova & Nikolov 2005, Brindle, Saharinen & Alitalo 2006). The Ang1 and Ang2 P domains share approximately 70% amino acid identity. The structure of the Ang2-Tie2 complex indicates that Ang1 would bind Tie2 in the same way as Ang2 (Barton et al. 2006). Six

of the thirteen contact residues are conserved between Ang2 and Ang1 P domains, further suggesting that both may bind Tie2 in a similar manner (Barton et al. 2006).

The crystal structure of the Ang2-Tie2 complex indicates that the ligand-receptor interface is confined to the top of the Tie2 Ig2 domain, which interacts with the P domain of Ang2 near the Ca<sup>2+</sup> binding site (Barton et al. 2006). To investigate the role of the Ang2 P domain in binding Tie2, mutagenesis study was done by Barton et al. (2005) within the P domain conserved patch. Both the mutants K467E/K472E and F468A/Y474A/Y475A generated proteins incapable of Tie2 binding. Furthermore two Ang2 mutants, I296A/T298A and R336A/E340A which alter conserved regions outside the P domain were also investigated by *in vitro* binding experiments. These mutant Ang2-RBD proteins were still able to form stable complexes with Tie2, indicating that the conserved molecular surfaces outside the P domain are not involved in receptor recognition (Barton, Tzvetkova & Nikolov 2005). These *in vitro* binding assays for Ang2-RBD mutants suggest that the P domain is essential for Tie2 receptor binding. However, binding of P domain by itself to the Tie2 ectodomain has not been demonstrated.

The aim of this chapter was to determine whether the Ang1-P domain is sufficient to bind to Tie2. To do this, the Ang1 fibrinogen related domain (FReD) and the P domain of Ang1 were cloned into a mammalian expression vector in frame with an epitope tag and secretory leader sequence. These constructs were transfected into mammalian CHO cells and secreted proteins recovered from the medium were tested for receptor binding.

Table 3.1 shows summary of the constructs cloned in this chapter.

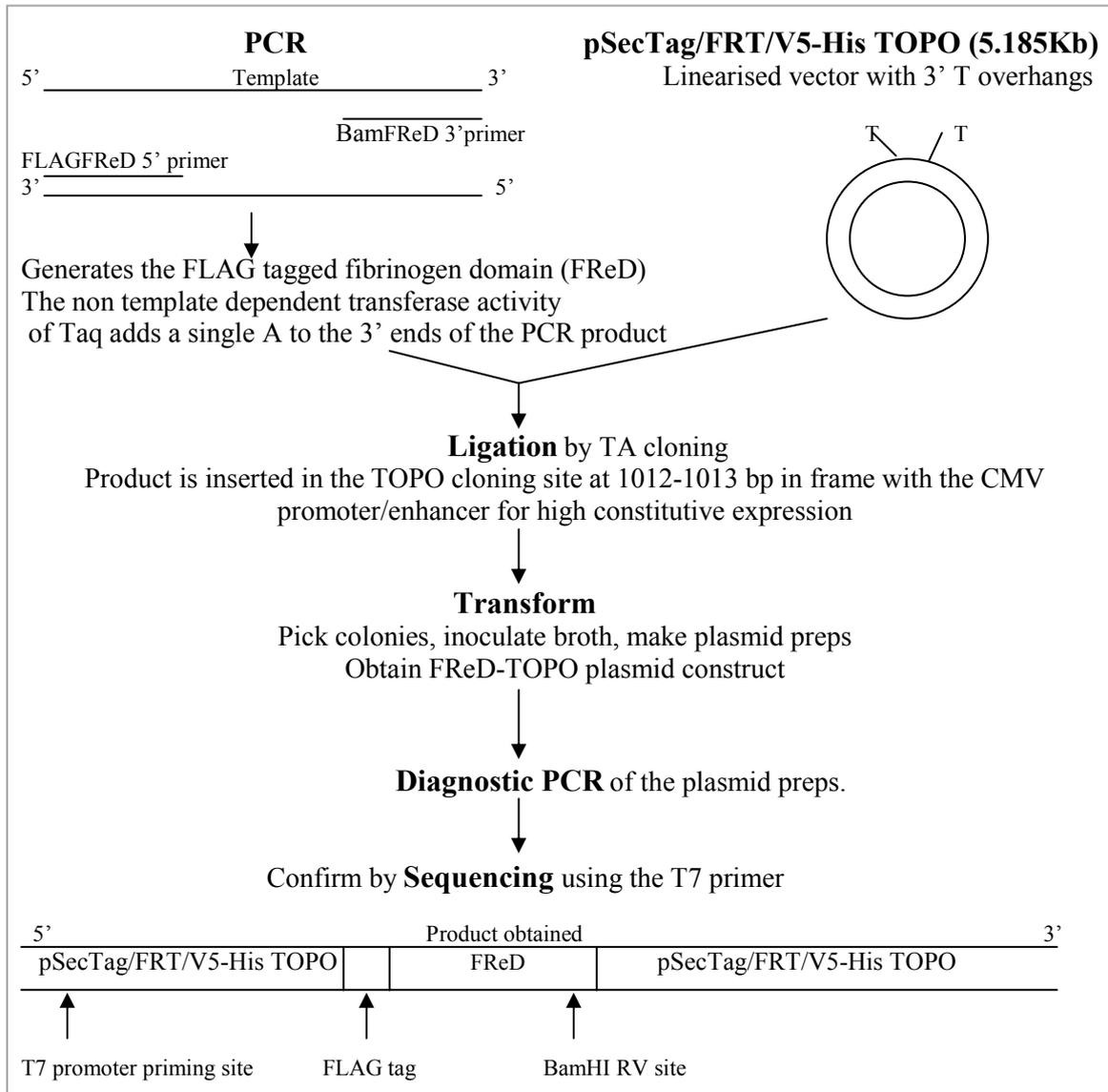
<b>Construct Name</b>	<b>Protein encoded by the construct</b>	<b>Tag</b>	<b>Features</b>	<b>Used for</b>
<i>Fred4Sec</i>	Ang1 Fibrinogen domain (Ang1 FReD) Residue Phe283- Phe498	N-terminus FLAG epitope tag	Soluble FLAG-tagged Ang1 FReD protein secreted into the media.	Binding to Tie2 receptor ectodomain
<i>Pe Sec</i>	Ang1 P domain (Ang1 P domain) Residue Ala412- Arg487		Soluble FLAG-tagged Ang1 P domain protein secreted into the media.	

**Table 3.1: Description of the constructs cloned in Chapter 3.** *Fred4* and *Pe* constructs were cloned into the pSecTag/ FRT/ V5-His TOPO expression vector for expressing soluble FLAG-tagged Ang1 FReD domain and FLAG-tagged Ang1 P domain proteins.

### **3.1 Cloning of FLAG tagged FReD (fibrinogen domain) in the pSecTag/ FRT/ V5-His TOPO expression vector.**

To generate secreted epitope tagged FReD (fibrinogen domain) of Ang1, the cDNA encoding isolated FReD was amplified from template (pcr2.1 Ang previously cloned Ang1 plasmid available in the lab) using FLAG tagged 5' primer and cloned in frame with the pSecTag/ FRT/ V5-His TOPO expression vector. This vector contains a secretion signal peptide sequence and cytomegalovirus (CMV) enhancer/promoter. The secretion signal directs the expression of a secreted protein. The CMV promoter allows high level constitutive expression of a FLAG tagged form of Ang1 FReD domain.

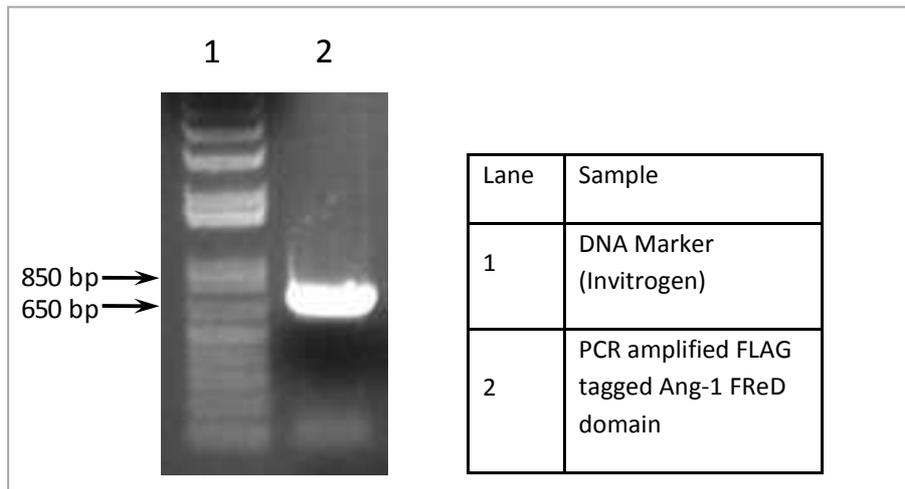
The strategy for cloning truncated fragment of Ang1, the fibrinogen domain is outlined in Figure 3.1.



**Figure 3.1: Schematic representation for cloning FLAG tagged FReD (fibrinogen domain) in the pSecTag/FRT/V5-His TOPO expression vector.** The FReD domain was amplified by PCR using the FLAGFReD 5' and BamFReD 3'primers and the amplified product was ligated to the pSecTag/FRT/V5-His TOPO expression vector. Competent cells were transformed with the ligation product and plasmid was prepared from the colonies on the transformation plate. Plasmid was analyzed by diagnostic PCR to screen for the presence of the desired insert. Sequence data was obtained for the positive clones.

The FReD domain was amplified by PCR using the FLAGFReD 5' and BamFReD 3'primers at annealing temperature 60 °C. The PCR primers incorporated the FLAG epitope tag and the BamHI restriction site in the Ang1 FReD PCR product. The FLAG tag is small (8 amino acid) and hence unlikely to interfere with subsequent analysis. The size of the amplified product which includes the FLAG tag and the

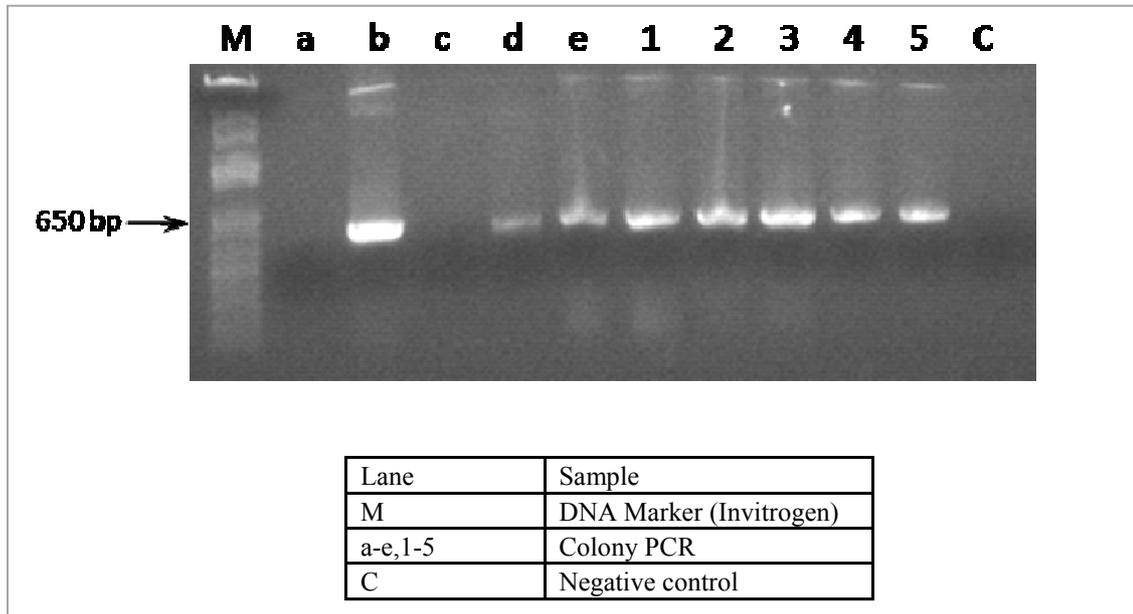
BamHI restriction site should be 689bp. The size of the FReD domain by itself is 655bp. The PCR product size was confirmed by running 10  $\mu$ l of the PCR product on a 1% agarose gel and was found to be of the expected size between 650 and 850bp, as seen in lane 2 of the agarose gel shown in Figure 3.2.



**Figure 3.2: Amplification of the Ang1-FReD domain by PCR.** The FReD domain was amplified using the FLAGFReD 5' and BamFReD 3' primers at annealing temperature 60 °C. The amplified product of 689bp expected size was analysed on an agarose gel. Lane 2 shows a band for the amplified FReD domain just above the 650 bp marker which is at the expected size for the FReD domain.

The amplified product of approximately 689bp contains A overhangs which were added by the transferase activity of Taq polymerase. This product was ligated to the pSecTag/FRT/V5-His TOPO expression vector by TA cloning. Competent cells were transformed with the ligation reaction and plated onto selective agar plates. Ten colonies labelled FReD a-e and FReD 1-5 were picked from the transformation plate and inoculated in selective medium. Plasmid preps were made from these and the plasmid was used for diagnostic PCR using the 5'FReD and 3'FReD primers at annealing temp 55°C to screen for the presence of the desired insert (Figure 3.3). The primers amplify only the Ang1 FReD domain, hence the expected size of the positive

clones should be 655bp. Bands of the expected size of ~655bp were seen for plasmids *b*, *d*, *e*, *1*, *2*, *3*, *4* and *5*, indicating these plasmids were positive clones for FReD domain.

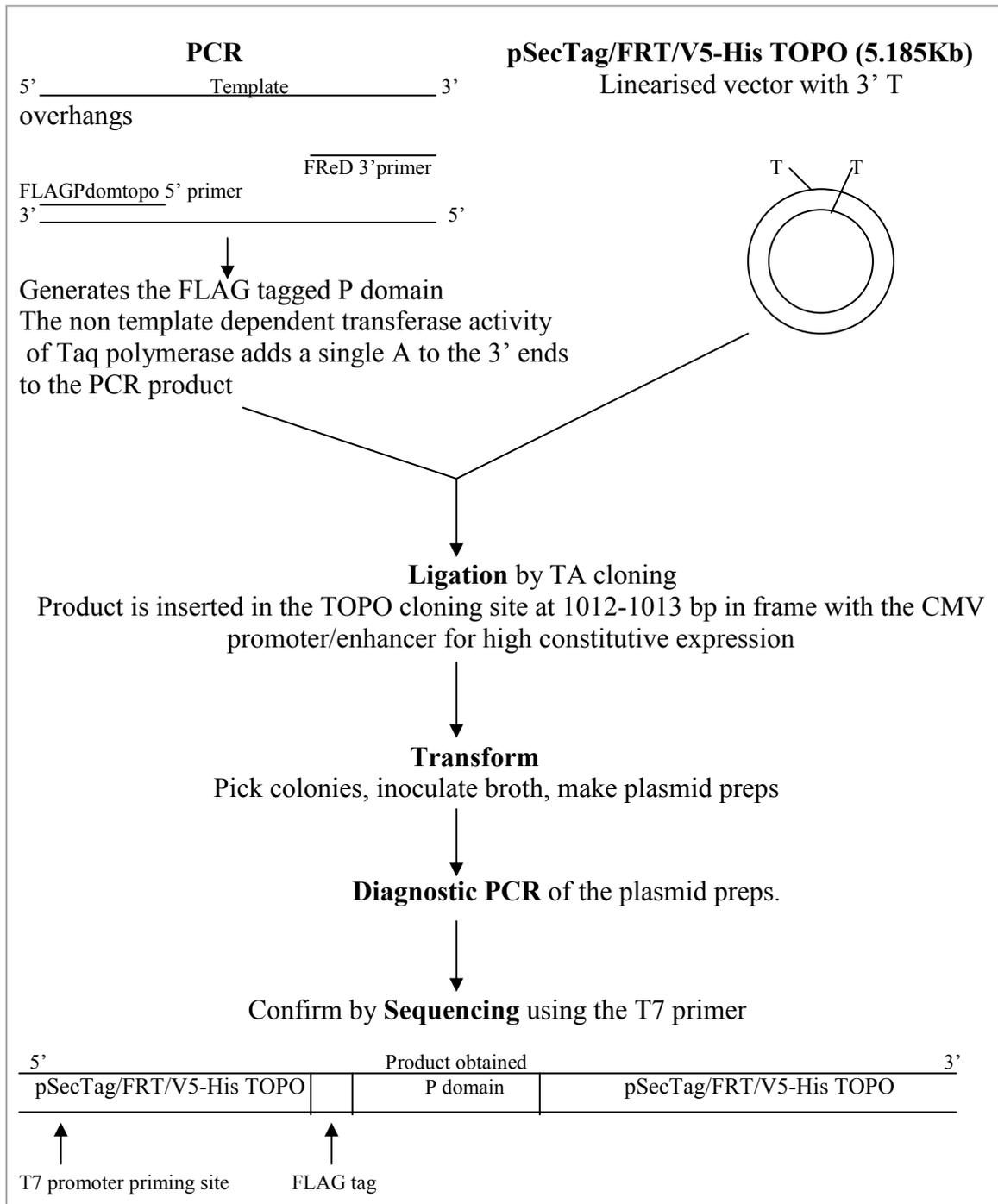


**Figure 3.3: PCR screen for Ang1-FReD domain in TOPO-FReD constructs.** Ten colonies labelled FReD a-e and FReD 1-5 were picked from the transformation plate and inoculated in selective medium. Plasmid preps were made from these and the plasmid was used for diagnostic PCR using the 5'FReD and 3'FReD primers to screen for the presence of the desired insert. Bands (approx 650bp) obtained for plasmid *b*, *d*, *e*, *1*, *2*, *3*, *4* and *5* which indicate positive clones.

The plasmids were sequenced using the T7 primer. A plasmid clone (labelled *FReD4 Sec*) was obtained with the correct sequence (See Appendix 4A for sequence result) and was used subsequently for expression of the secreted forms of the Ang1 FReD domain (See Section 3.3).

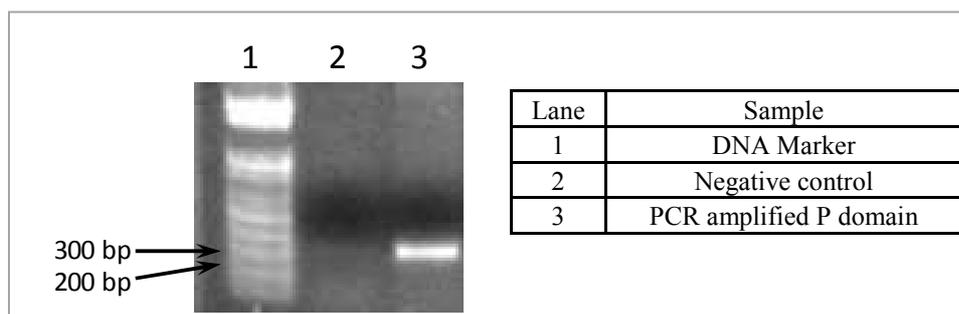
### 3.2 Cloning of FLAG tagged P domain in the pSecTag/FRT/V5-His TOPO expression vector.

To generate secreted epitope tagged P domain of Ang1 the cDNA encoding isolated P domain was amplified from template pcr2.1 Ang using a strategy similar to that used in cloning the fibrinogen domain (See section 3.1). This strategy for cloning truncated fragment of Ang1, the P domain is outlined in Figure 3.4.



**Figure 3.4 - Schematic representation for cloning FLAG tagged P domain in the pSecTag/FRT/V5-His TOPO expression vector.** The P domain was amplified by PCR using the FLAGPdomtopo 5' and bamFReD 3'primers and the amplified product was ligated to the pSecTag/FRT/V5-His TOPO expression vector. Competent cells were transformed with the ligation product and plasmid was prepared from the colonies on the transformation plate. Plasmid was analyzed by diagnostic PCR to screen for the presence of the desired insert. Sequence data was obtained for the positive clones.

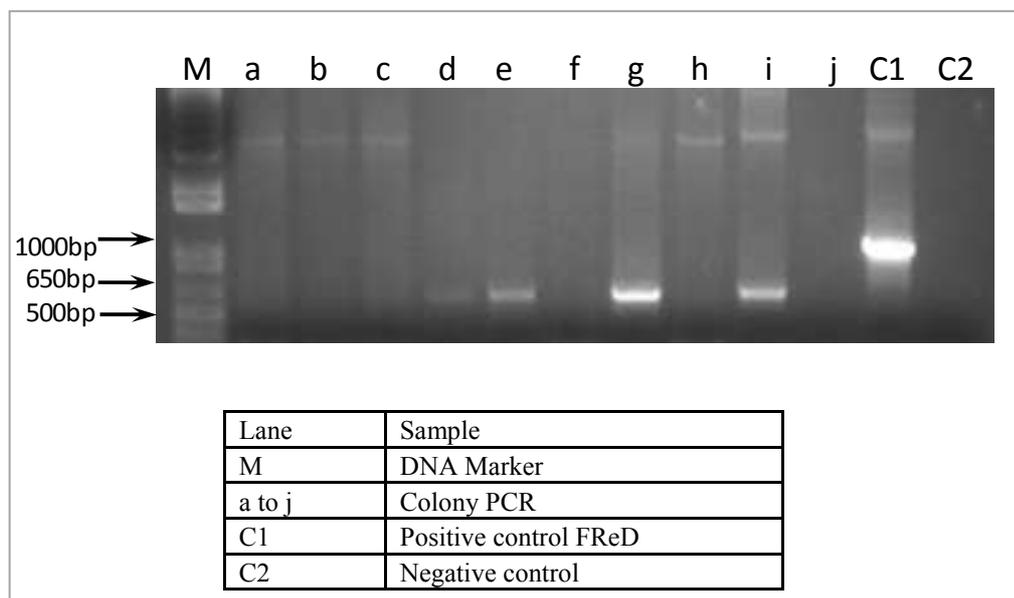
The P domain was amplified by PCR using the FLAGpdomtopo 5' and bamFReD 3'primers at annealing temperature 62 °C. The PCR primers incorporated the FLAG epitope tag at the 5' end of the Ang1 P domain PCR product. The size of the amplified FLAG-tagged P domain product was expected to be 295bp and was confirmed by running 5 µl of the PCR product on a 1% agarose gel as seen in lane 3 (Figure 3.5).



**Figure 3.5: Amplification of the Ang1- P domain by PCR.** The P domain was amplified by PCR using the FLAGpdomtopo 5' and bamFReD 3'primers at annealing temperature 62 °C. The amplified product P domain is analysed on a gel. Lane 3 shows a band for the amplified P domain at approximately 295 bp length. Lane 2 is the negative control PCR in which no template DNA was added and hence no amplified P domain obtained.

The amplified product FLAG-tagged Ang1 P domain was of the expected size at 295bp and was ligated to the pSecTag/FRT/V5-His TOPO expression vector by TA cloning. Competent cells were transformed with the ligation reaction which was plated onto selective agar plates. Eleven colonies labelled *Pa-Pi* were picked from the transformation plate and inoculated in selective medium. Plasmid preps were made from these and the plasmid was used for diagnostic PCR at annealing temperature 42 °C using the T7 5'(Tm=50 °C) and BGH reverse 3'(Tm=58 °C) primers. These primers allow confirmation that the Ang1 P domain PCR product has been cloned into the vector between the T7 and BGH primers sequence. The positive clones will yield a PCR product which includes 295bp of the P domain and 294bp of the vector sequence. Thus P domain positive clones are expected to give a band at ~589 bp and the positive control

*FReD 4 clone* will give a band at 939 bp (Figure 3.6). As seen in figure 3.6, lanes for plasmid *Pd*, *Pe*, *Pg* and *Pi* were positive clones for Ang1 P domain.



**Figure 3.6: PCR screen for Ang1-P domain in TOPO-P domain constructs.** Eleven colonies labelled *Pa-Pi* were picked from the transformation plate and inoculated in selective medium. Plasmid preps were made from these and the plasmid was used for diagnostic PCR at annealing temperature 42 °C using the T7 5'(Tm=50) and BGH reverse 3'(Tm=58) primers. This allows confirmation that the product has been cloned into the vector between the T7 and BGH primers sequence. Thus expected size of positive clones is 589 bp and the expected size of positive control *FReD 4 clone* is 939 bp. Bands at approximately 600bp obtained for plasmid *Pd*, *Pe*, *Pg* and *Pi*.

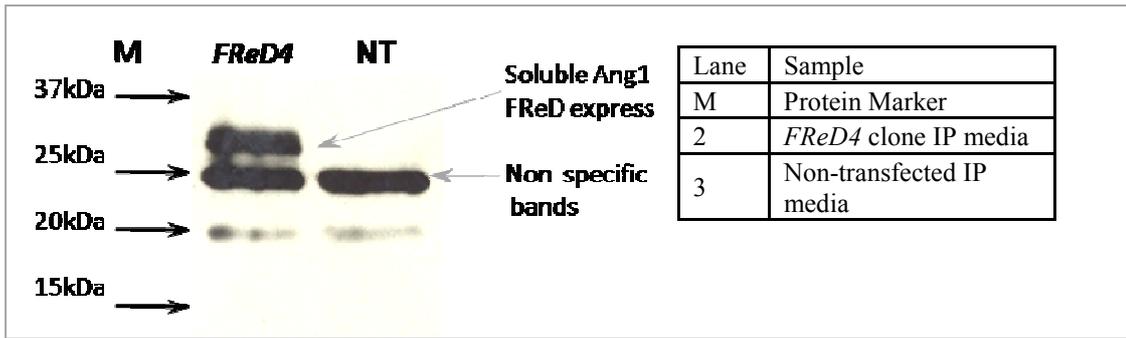
Plasmids from the positive clones were sequenced using the T7 primer and a clone (labelled *Pe Sec*) with the correct sequence (See Appendix 4B for sequence data) was used further for expression of secreted forms of Ang1 P domain in mammalian cells (See Section 3.3).

### 3.3 Expression of secreted Ang1 FReD and P domains in mammalian cells.

The construct obtained in Section 3.1 was analyzed to confirm expression of soluble form of the Ang1 FReD domain in mammalian cells.

HEK 293 cells were transfected when 60% confluent in 6 well plate using 2 $\mu$ g *Fred4 Sec* construct and Lipofectamine 2000 as described in Section 2.15.1. 1ml complete media from HEK 293 cell transfection was collected 24 hours post-transfection. Secreted FLAG-tagged FReD domain from the media was concentrated by immunoprecipitation using Anti-FLAG agarose beads as described in Section 2.17.1. Western blot was performed to analyze the immunoprecipitated media from the transfection.

The expected size for Ang1 FReD domain is approximately 25 kDa. A band between 25 and 37 kDa was obtained in lane 2, showing that the FReD domain was concentrated from the media by immunoprecipitation. The size of the band obtained for the Ang1 FReD domain is more than the expected size 25 kDa. It is possible that this increase in size is a consequence of glycosylation of the FReD domain. Lane 3 served as a non-transfected control media which was concentrated using the Anti-FLAG beads, but did not show the band seen in the transfected lane 2. A band at 25kDa is seen in the transfected lane 2; however a band of the same size was also seen in the non-transfected sample (lane 3). It was concluded that this band is a non specific band due to the anti-FLAG antibody detecting media proteins (Figure 3.7).

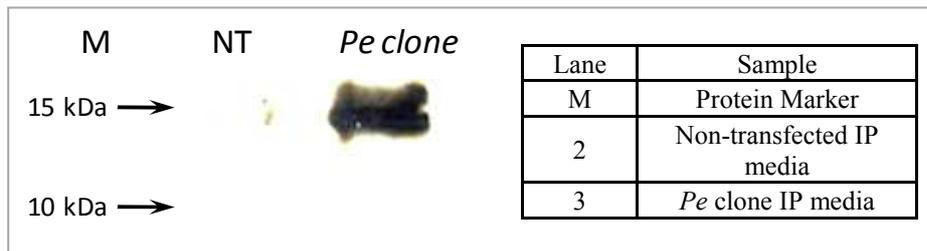


**Figure 3.7: Western blot to analyse *FReD4 Sec* construct for expression of soluble Ang1-FReD protein in mammalian cells.** HEK 293 cells were transfected with *FReD4 Sec* construct (described in Section 3.1). Complete media from HEK 293 cell transfection was collected and secreted FReD domain in the media was concentrated by immunoprecipitation. A Western blot was performed to analyze the concentrated immunoprecipitated media and cell lysate. The expected size for FReD domain is 25 kDa. Band was observed between 25-37 kDa in lane 2. The obtained size of the FReD domain was more than the expected size possibly due to glycosylation of the protein. Lane 3 was non-transfected control concentrated media. The anti-FLAG antibody used for Western blot gave a non-specific band at 25kDa as observed in both the transfected and non-transfected sample (lane2 and 3) The band above the non-specific band of 25kDa in lane 2 is specific for FLAG-tagged Ang1 FReD domain expressed and secreted into the media when mammalian cells were transfected with the *FReD4 Sec* construct.

Thus when the *FReD4 Sec* construct (cloned in section 3.1) was transfected into mammalian HEK cells, FLAG-tagged Ang1 FReD domain was expressed and secreted into the media and was concentrated using Anti-FLAG agarose beads.

The *Pe Sec* construct (obtained in Section 3.2) was analyzed to confirm expression of the Ang1 P domain in mammalian cells. HEK 293 cells were transfected in 6 well plate using 2 $\mu$ g *Pe Sec* construct and Lipofectamine 2000 as described in Section 2.15.1. 5ml complete media from HEK 293 cell transfection was collected 24hours post-transfection. Secreted FLAG-tagged P domain in the media was concentrated by immunoprecipitation using Anti-FLAG agarose beads as described in Section 2.17.1. Only the FLAG-tagged P domain from the media bound to Anti-FLAG beads, thus concentrating the Ang1 P domain. Western blot was performed to analyze the immunoprecipitated media and cell lysate from the transfection.

The expected size for Ang1 P domain is approximately 11 kDa. A band was obtained at just under the 15 kDa marker in lane *Pe*, showing that the P domain was concentrated from the media by immunoprecipitation. The size of the band obtained for the Ang1 P domain is more than the expected size 11 kDa. It is possible that this increase in size is a consequence of glycosylation of the P domain. Lane 2 served as a non-transfected control media which was concentrated using the Anti-FLAG beads, but did not show any band (Figure 3.8).



**Figure 3.8: Western blot to analyse *Pe Sec* construct for expression of soluble Ang1-P domain protein in mammalian cells.** HEK 293 cells were transfected with *Pe Sec* construct (described in Section 3.2). Complete media from HEK 293 cell transfection was collected and secreted P domain in the media was concentrated by immunoprecipitation. A Western blot was performed to analyze the concentrated immunoprecipitated media and the non-transfected control media. The expected size for Ang1 P domain is ~11 kDa. Band between 10 and 15 kDa of approximately 14 kDa was obtained in lane 3, whereas no band was seen in the non-transfected control lane 2, indicating that FLAG-tagged Ang1 P domain was expressed and secreted into the media when mammalian cells were transfected with the *Pe Sec* construct.

Thus when *Pe Sec* construct (cloned in section 3.2) was transfected into mammalian HEK cells, FLAG-tagged Ang1-P domain was expressed and secreted into the media and was concentrated using Anti-FLAG agarose beads.

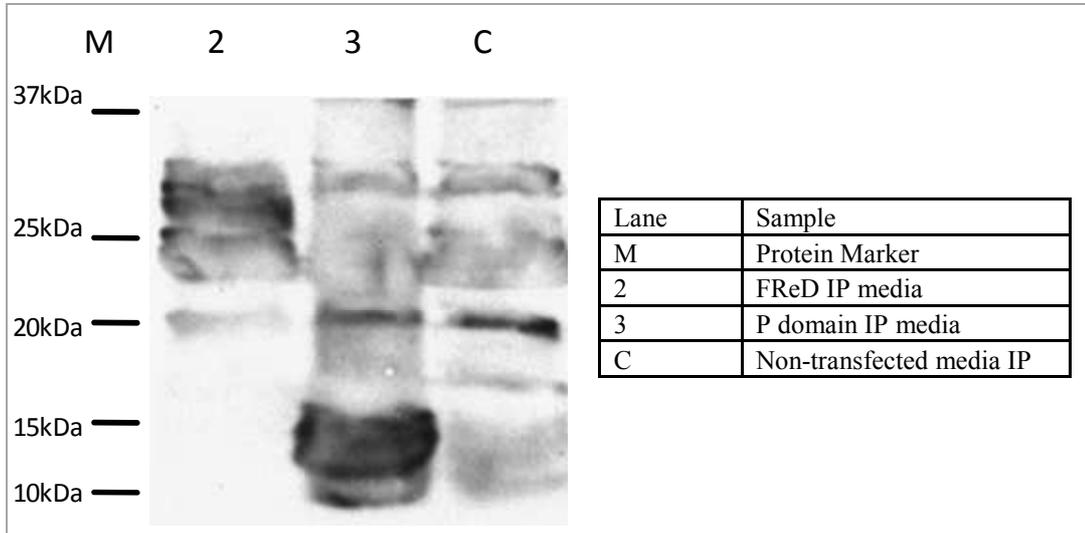
### 3.4 Binding of Tie2 to secreted FReD and P domains.

To determine whether the P domain (Ang1) is sufficient for Tie2 binding, a non quantitative binding assay was performed using the secreted forms of FReD and P domain and soluble Tie2 ectodomain.

CHO cells were transfected in T80 flask with 10 $\mu$ g *FReD4 Sec* and *Pe Sec* constructs (obtained and analysed in sections 3.1, 3.2 and 3.3) encoding soluble Ang1-FReD domain and soluble Ang1-P domain using lipofectamine2000 transfection reagent. Complete media containing soluble forms of the domains was collected 20 hrs post-transfection. 1ml of complete media containing soluble FLAG-tagged FReD domain was incubated with 200ng rhTie2 ectodomain at 4°C for 2hrs. As compared to FReD domain the expression of the P domain was lesser, hence 5 ml of complete media containing soluble FLAG-tagged P domain was incubated with 200ng rhTie2 ectodomain at 4°C for 2hrs. The incubation with Tie2 ectodomain allowed binding of the FLAG-tagged soluble domains to the Tie2 receptor. 5 ml of non-transfected complete media was also incubated with Tie2 to serve as a control to account for any non-specific Tie2 binding to other proteins in the media. This mix of ligand domain and receptor protein was immunoprecipitated with Anti-FLAG agarose beads at 4°C overnight and analyzed by Western blot. Immunoprecipitated proteins were first probed using anti-FLAG primary antibody and anti-mouse secondary antibody to confirm the expression of FReD and P domain.

As shown in the blot in Figure 3.9 similar amounts of FReD and P domain were detected from 1ml and 5ml complete media concentrated by immunoprecipitation and probed by Anti-FLAG antibody (Lane 2 and 3). As observed in section 3.3 both the

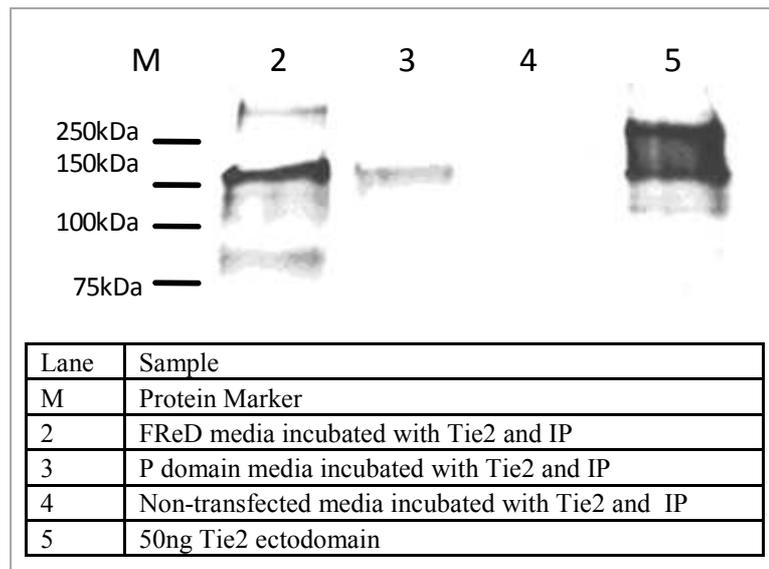
domains were detected at a molecular size more than the expected size possibly due to glycosylation of the protein.



**Figure 3.9: Immunoprecipitation of Ang1 secreted FReD and P domain.** CHO cells were transfected with *FReD4 Sec* and *Pe Sec* constructs encoding secreted Ang1 FReD and Ang1 P domains. Complete media containing soluble forms of the domains was incubated with 200ng rhTie2 ectodomain at 4°C for 2hrs. Non-transfected complete media was also incubated with Tie2 to serve as a control for non-specific Tie2 binding to media proteins. This mix of ligand domain and receptor protein was immunoprecipitated with Anti FLAG agarose beads and analyzed by Western blot. Immunoprecipitated proteins were probed using anti-FLAG primary antibody and anti-mouse secondary antibody to confirm the expression of FReD and P domain. The expected size for FReD domain is ~25 kDa. Band between 25 to 37 kDa of approximately 30 kDa is observed in lane 2. The expected size for P domain is ~11 kDa. Band between 10 to 15 kDa of approximately 14 kDa is observed in lane 3. The obtained size of the secreted domains was more than the expected size possibly due to glycosylation of the protein. Lane 3, the non-transfected control served as a negative control.

To determine whether Tie2 ectodomain bound the immunoprecipitated FReD and P domains, the blot in Figure 3.9 was stripped and reprobed for Tie 2 receptor using anti-Tie2 ectodomain primary antibody and anti-goat secondary antibody. As seen in figure 3.10, both FReD and P domain incubated with Tie2 pulled down Tie2 receptor during the immunoprecipitation at the expected size for glycosylated rhTie2 at 165 kDa. Hence both FReD and P domain are sufficient to bind to the Tie2 ectodomain. However the amount of Tie2 pulled down with P domain (Lane 3) is lesser than the amount

pulled down with FReD domain (Lane 2), indicating that the binding of P domain to Tie2 may be weaker compared to the FReD domain. Lane 4 shows the non-transfected negative control media which does not pull down any Tie2 indicating that binding of Tie2 is specific for the FLAG-tagged domains and no other proteins in the media interfere with this binding to give any non-specific background binding or immunoprecipitation with the beads.



**Figure 3.10: Immunoprecipitation of Tie2 ectodomain bound to Ang1 FReD and P domains.** The blot in Figure 3.9 which showed expression of soluble FReD and P domain in the concentrated media was stripped and reprobed for Tie 2 using anti-Tie2 ectodomain primary antibody. If the FReD or P domain bound to the Tie2 receptor, then the domain bound Tie2 receptor would also be pulled down with the FLAG-tagged domains during immunoprecipitation with the Anti-FLAG beads. Lanes 2 and 3 show that both FReD and P domain incubated with Tie2 pulled down Tie2 receptor during the immunoprecipitation at the expected size for Tie2 at 165 kDa. Band at approximately 150 kDa was observed in lane 2 and 3. Lane 4 of non-transfected cell media shows no binding to Tie2. Lane 5 is a positive control in which 50ng of rhTie2 ectodomain was loaded. This experiment was performed independently five times for FReD and two times for P domain with similar results.

Thus when approximately equal amount of FReD and P domain were expressed (Figure 3.9) FReD shows stronger binding to Tie2 when compared with the weak Tie2 binding of the P domain transfected cells (Figure 3.10). Thus, P domain by itself also

binds to Tie2 (Figure 3.10). This experiment was performed five times for FReD and two times for P domain with similar results.

## **Discussion**

Many proteins are large complex molecules composed of multiple domains which serve different purposes in the complete protein function. The size and complexity of proteins can often make them difficult to produce due to problems associated with conformational instability, aggregation and improper *in vitro* folding. In addition, proteins may be prone to enzymatic and physical barriers such as proteolytic cleavage, instability in the gastro-intestinal tract, poor membrane penetration, and also unfavourable pharmacokinetics immunogenicity (Frokjaer, Otzen 2005). These poor ‘druggability’ characteristics of the entire protein molecule make their bioavailability low and insufficient for effective systemic affect without repeated administration. However, most proteins exert their activity through relatively small regions of their folded surfaces. Thus minimization of proteins involves retaining the most important regions for pharmacokinetic properties and thus solving the problems of the large unstable protein molecule (Fairlie, West & Wong 1998).

Minimization of protein domains can be performed by the use of sequence randomization, functional selection within sequence, combined with structural information. These newly engineered mini-proteins containing minimized protein domain of smaller size with specific binding affinities and well-defined three-dimensional structure, represent novel tools in drug design, biotechnology and structural studies. Minimization of proteins by using oligomeric scaffolds bound to binding domains serves as an option for generating therapeutic mini-protein drugs that retain the function of much larger proteins, are easier to produce, have slower clearance *in vivo* as

opposed to small molecule peptide drugs (<5kDa) and some of them can also be directly used in therapy or present high potential when appropriately formulated as drugs (Vita et al. 1998, McGregor 2008).

Ang1 shows potential as a therapeutic agent at the intersection of angiogenesis, inflammation and homeostasis processes involved in several critical illnesses as discussed in chapter 1, section 1.3. However, native Ang1, a big molecule of ~70kDa is not only difficult to produce due to problems associated with aggregation but also displays variable activity on purification (Kim et al. 2005). Furthermore, native Ang1 can activate Tie2 receptor only on multimerization, which further increases the size of the molecule. The size of the doorway diameters of endothelial caveolae where most Tie2 receptors are located typically range from 10–150 nm (Parton 2003) whereas the sizes of higher order multimeric Ang1 molecules are more than 200 nm (Kim et al. 2005). For uses as a recombinant therapeutic ligand a small binding domain linked to oligomerization motifs could reduce molecular mass. This could allow easier production and purification of the recombinant protein as compared to the native forms as well as ease in formulation and *in vivo* accessibility in clinical administrations.

The crystal structure of Ang2 indicates that the P domain is the interacting interface, with residues in the P domain directly interacting by H bond formation with Tie2 residues (Barton et al. 2006). Mutational studies of some residues outside the P domain did not affect binding and the mutants showed binding similar to wild-type (WT) proteins (Barton et al. 2006). Although the crystal structure of Ang1 is not available the high homology between Ang1 and Ang2 indicate that they are likely to share similar 3-D structure. Until now the Ang1-FReD domain has been demonstrated to be the minimal receptor binding domain (Davis et al. 2003, Weber et al. 2005).

Smaller truncated fragments of Ang1 have not been investigated for binding. A small binding domain of Ang1 could serve as an alternative to the 25kDa FReD domain for binding the receptor. The aim of this chapter was to test whether the Ang1 P domain is capable of binding the Tie2 receptor.

The P domain and FReD domains of Ang1 were cloned into a secretory mammalian expression vector. The secreted proteins were analysed for expression by Western blotting. It was found that P domain had comparatively lower expression levels than FReD. A much higher volume of media was required to be immunoprecipitated to obtain P domain protein for binding assays. The binding assay of P and FReD domains were performed by immunoprecipitation and SDS-PAGE analysis. P domain was able to bind to Tie2 albeit weakly and also required higher amounts of the protein for binding Tie2 compared with FReD protein. It is important to note that higher binding does not always translate into increase in biological activity (Pearce et al. 1999). Thus, it is possible that although Ang1 P domain binds very weakly to Tie2, it is reduced in size and on oligomerization may be able to elicit the same or increased level of response *in vivo* as oligomerized FReD domain.

The binding of the Ang1 P domain to Tie2 could be further investigated by purifying, concentrating and quantifying the P domain followed by a binding assay. High amounts of protein production can be achieved by using specialized *E.coli* strains, however, the post-translational modification systems, codon bias, and protein overproduction in bacterial systems could lead to mRNA stability, protein accumulation levels, misfolding and segregation into insoluble aggregates (Baneyx 1999). Thus, it may not be ideal to support the production of a properly folded Ang1 P domain. The FLAG- tagged protein could be purified and concentrated using the anti-FLAG antibody

(Brizzard, Chubet & Vizard 1994). A sensitive binding assay may be performed using the purified protein by ELISA or surface plasmon resonance (SPR) techniques.

Whether the P domain could be minimalized further was not investigated in this chapter. It is possible to create truncation mutants of P domain to test this. However, it has been suggested that decreasing the size of the therapeutic molecule to <40kDa may increase its systemic clearance rate and thus require repeated administration (Torchilin, Lukyanov 2003). The ideal size of a therapeutic protein binding site largely depends on the total size of the molecule after linkage to oligomerization scaffolds, the size of the active multimeric form, the size post chemical modification required for formulation of stable drugs and the site of clinical application.

## Chapter 4

### Rational Mutagenesis of Ang2 Receptor Binding Interface

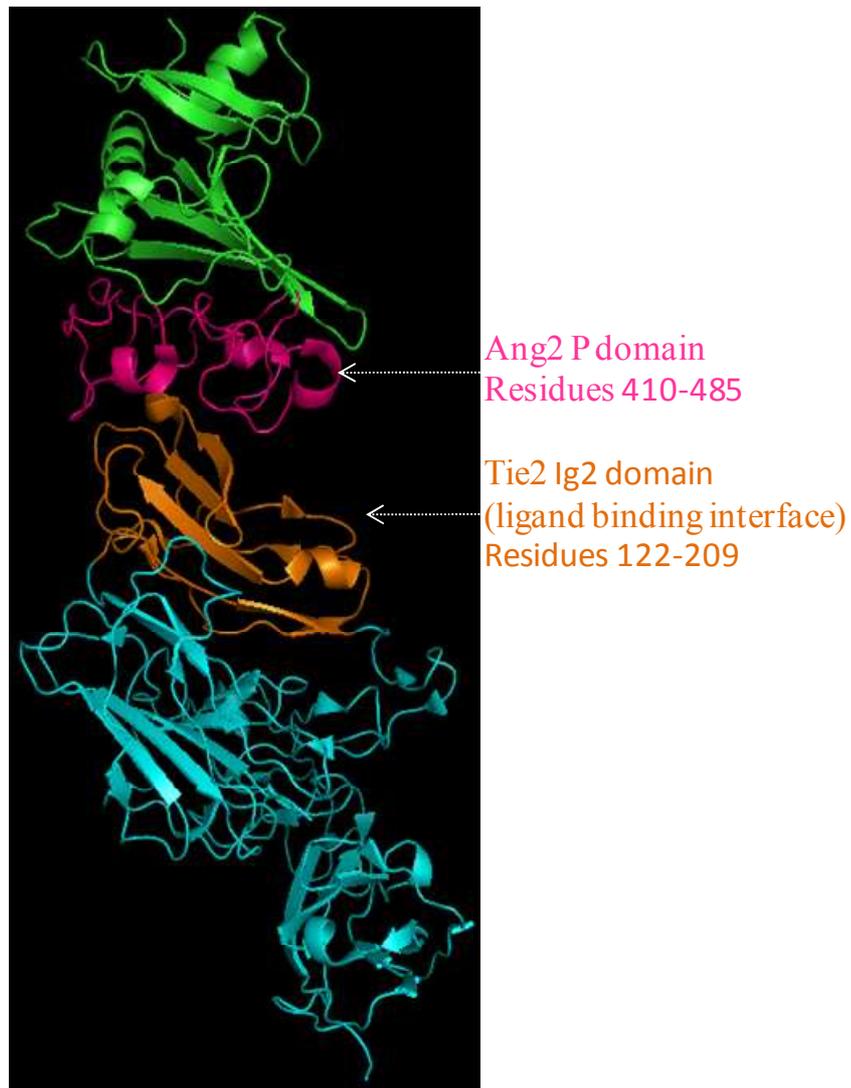
#### Introduction

The angiopoietin family of ligands interact with the Tie receptor family to mediate angiogenesis. The Ang family comprises of four members, namely; Ang1, Ang2, Ang3 and Ang4. Among these, Ang1 and Ang2 are characterized in most detail. The Tie family comprises of two members; Tie 1 whose signalling pathways are not well understood and a well studied Tie 2 receptor.

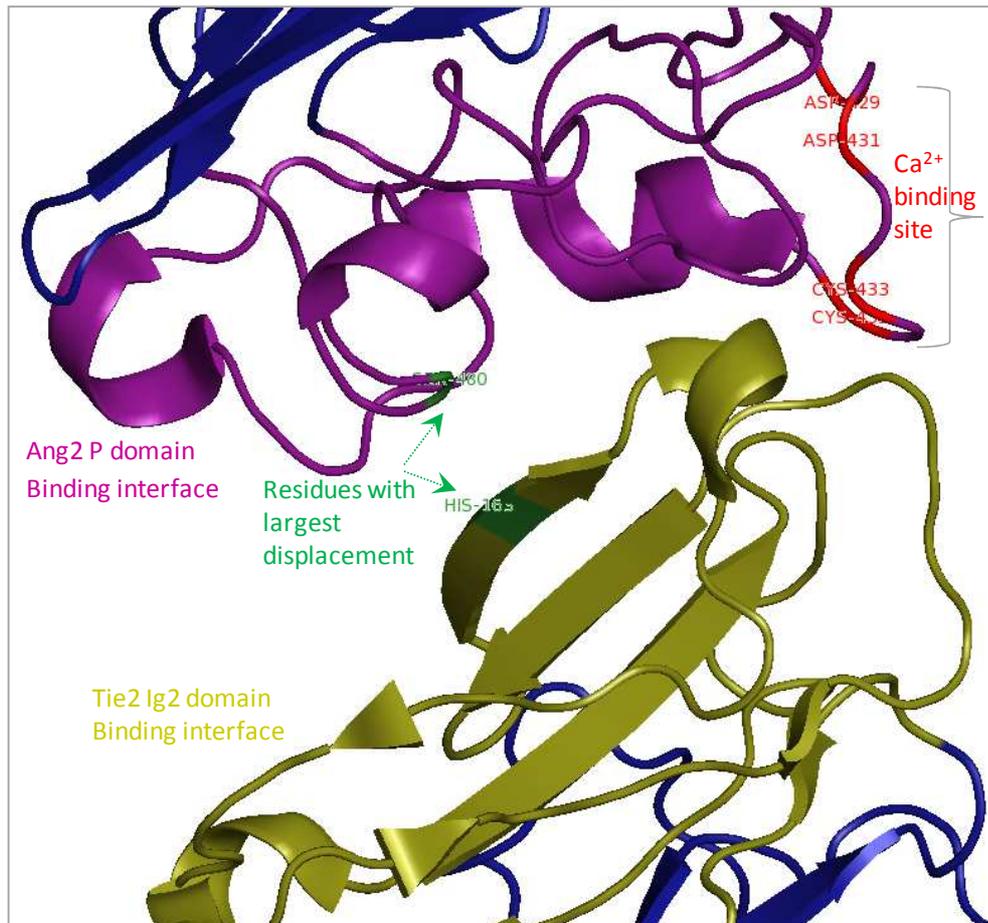
The four angiopoietins have a highly conserved structure, which consists of a small amino terminal responsible for angiopoietin superclustering, followed by a coiled-coil domain coded by residues 100-280 responsible for dimerization. The coiled-coil domain is separated from the fibrinogen domain by a short linker. The fibrinogen domain encoded by residues 280-498 functions as the receptor recognition and binding region. The fibrinogen domain consists of the P domain which is implicated to serve as a ligand-recognition and binding region (Barton, Tzvetkova & Nikolov 2005).

The crystal structure of the Ang2-Tie2 complex has been elucidated. Structure of the Ang2-Tie2 complex shows that the Tie2 Ig2 domain interacts with Ang2 P domain near the Ca<sup>2+</sup> binding site (Figure 4.1 and 4.2) (Barton et al. 2006). The crystal structure of the Ang2-Tie2 complex reveals no major conformational changes or domain rearrangements on ligand binding. Figure 4.2 shows the Ang2-Tie2 binding interface. The Tie2 and Ang molecular surfaces that are complementary both in shape and in chemical nature recognize and bind without marked rearrangements in either binding partner. The free and bound Ang2 structure can be superimposed with a root mean

square (rms) deviation of 0.55 Å, which is comparable to experimental error. In the Ang2-Tie2 complex, the largest displacement in Tie2 is around Ig2 residue His163 (2.6 Å) and in Ang2 at residue Ser480 (1.8 Å). The Ang2 receptor binding interface is rich in aromatic residues, particularly tyrosines and tryptophans (Barton et al. 2006).



**Figure 4.1:** Structure of the Tie2-Ang2 complex visualised in the PyMOL imaging software. The Ang2 P domain spans residues 410-485 (coloured pink) and contacts the Tie2 binding interface (residues 122-209, coloured orange).



**Figure 4.2: Ribbon representation of the Ang2-Tie2 binding interface.** The Ang2 P domain (pink) and Tie2 Ig2 domain (olive) form the interface of the complex structure. The residues Asp429, Asp431, Cys433 and Cys435 in the Ca<sup>2+</sup> binding site are indicated (red). Residues Ang2-Ser480 and Tie2-His163 are the largest displaced residues (green) in the Ang2-Tie2 complex.

When crucial Tie2 residues (F161A/S164E) were mutated, both Ang1 and Ang2 were unable to bind or activate the mutated receptor, suggesting that these residues were important for interaction with both ligands and the ligands may bind Tie2 in very similar manners (Barton et al. 2006). The first 360 amino acids (Ig-like domain plus EGF-like repeats) of the Tie-2 receptor are necessary and sufficient to bind both Ang1 and Ang2, which further suggests that differential receptor binding is not likely to be responsible for the different functions of Ang1 and Ang2 (Fiedler et al. 2003). Furthermore the affinities of Ang1 and Ang2 for Tie2 binding were reported to be the same when binding of either ligand to the extracellular domain of Tie2 was examined *in*

*vitro* (Maisonpierre et al. 1997). However, a recent paper by Yuan et al. (2009) states that the affinity of Ang2 is 20-fold lower than Ang1 (Yuan et al. 2009).

The structure of the Ang2–Tie2 complex and structure based mutagenesis studies suggest that the Ang2 P domain is solely responsible for mediating the interactions with Tie2 (Barton, Tzvetkova & Nikolov 2005, Barton et al. 2006). Residues outside the P domain are not likely to have an effect on binding but may be involved in recognition.

The structure of the Ang2–Tie2 complex, affinity studies, mutational studies on Tie2 and sequence homology between Ang1 and Ang2 indicates that Ang1 is likely to bind Tie2 in the same way as Ang2. Six of the thirteen contact residues at the complex interface are conserved between Ang2 and Ang1. Of the differing residues, two residues contain conservative substitutions in Ang1 (I434M and F469L), and two others (K432N and Y475H) with similar amino acid properties are also likely to have little or no effect on the ligand-receptor interactions (Barton, Tzvetkova & Nikolov 2005). Thus, there are only three Tie2-contacting residues different between Ang1 and Ang2 (N467G, S417I and S480P) that might affect the ligand-receptor interactions (Barton et al. 2006).

Figure 4.3 shows the sequence alignment of the Ang1 and Ang2 P domain, indicating the important contact residues.

```

Ang1-412   AGKQSSLILHGADFSTKDA DNDNCMCKCALMLTGGWWFDACGPSNL
Ang2-410   AGKISSISQPGNDFSTKDG DNDKICCKCSQMLTGGWWFDACGPSNL

Ang1-458   NGMFYTAGQNHGKLNIGIKW HYFKGPSYSLRSTTMMIRPLDF
Ang2-456   NGMYYPQRQNTNKFNIGIKW YYWKGSGYSLKATTMMIRPADF

```

**Figure 4.3: Sequence alignment of the Ang1 and Ang2 P domain.** The residues differing between the Ang1 and Ang2 P domain are highlighted in the sequence alignment. The residues contacting Tie2 and likely to not affect binding are coloured blue. The residues contacting Tie2 and likely to affect binding are coloured red.

Of the N467, S417 and S480 Ang2 residues likely to affect the ligand-receptor interactions, none have been previously investigated by mutagenesis studies. But residue K468 has been analysed by mutagenesis as a double mutant K467E/K472E for Tie2 binding studies. It should be noted that Barton et al. (2005) refer to Lysine (K) residue at position 467 throughout the paper, which is likely to be a nomenclature error and residue corresponds to position 468 as reported in their consequent paper Barton et al. (2006). Mutant K467E/K472E was concluded to be incapable of Tie2 binding as detected by gel filtration binding assays (Barton, Tzvetkova & Nikolov 2005). However, a close examination of the data reported by Barton et al. (2005) (Structure of the angiopoietin-2 receptor binding domain and identification of surfaces involved in Tie2 recognition) suggests that the K467E/K472E mutant does show binding to Tie2 albeit at a much lower level than that observed for WT Ang2-RBD. Crystal structure of mutant K467E/K472E also showed similar structure to WT protein with only the  $\alpha 6$  adopting a slightly different conformation (Barton, Tzvetkova & Nikolov 2005). N467 is adjacent to K468 and is implicated to being important for binding it is not conserved with Ang1 (Gly467).

The Ang2 S480 residue differs from the Ang1 residue P480, which is an amino acid with very different properties. The Ang2 S480 residue displays the largest displacement in the Ang2-Tie2 complex and the side chain of Ang2 S480 makes an H bond with the main chain hydroxyl of Tie2 S164. Thus Ang2 S480 has been implicated as an Ang2 residue likely to affect the ligand-receptor interactions (Barton et al. 2006).

The Ang2 G428 lies within the P domain but does not contact any Tie2 residues. Ang2 G428 itself is not a conserved residue with Ang1 (A428) however, Ang2 G428

lies within a region of residues (D422-D431) conserved with Ang1. Ang2 G428 also lies close to the Ca<sup>2+</sup> binding site which is indicated to stabilise the structure.

This chapter aims to investigate importance of residues N467, S480 and G428 within the receptor binding domain (P domain) of Ang2 in binding to Tie2. A summary of the rational mutants is outlined in Table 4.1. To do this mutant Ang2-RBD were constructed by site directed mutagenesis and utilized in binding assays. Binding was measured by flow cytometry assay using surface expressed Ang2 WT RBD and Ang2 mutant RBD. Details of the surface expression system used are discussed in detail in Chapter 5. Mammalian cell surface expression system holds several advantages for studying affinity of proteins in terms of on-cell characterization of candidate clones in a physiological context without the need for adequate protein expression, quantification and purification. Mammalian cells are large enough to be analyzed and sorted using flow cytometry in a truly quantitative manner (Boder, Wittrup 1998, Wittrup 2001).

<b>Mutant Name</b>	<b>Mutated Residue Number</b>	<b>Wild Type Ang2 amino acid</b>	<b>Mutated amino acid</b>	<b>Wild Type Ang1 amino acid</b>
<i>N467S</i>	467 in P domain	Asparagine	Serine	Glycine
<i>S480P</i>	480 in P domain	Serine	Proline	Proline
<i>G428T</i>	428 in FReD domain	Glycine	Threonine	Alanine

**Table 4.1: Summary of the rational mutants N467S, S480P and G428T obtained by site directed mutagenesis of Ang2-FReD domain.** The wild-type and mutated amino acid residues in Ang2-FReD domain are shown in the table. Also the corresponding residues in the WT Ang1 at the mutated positions are indicated.

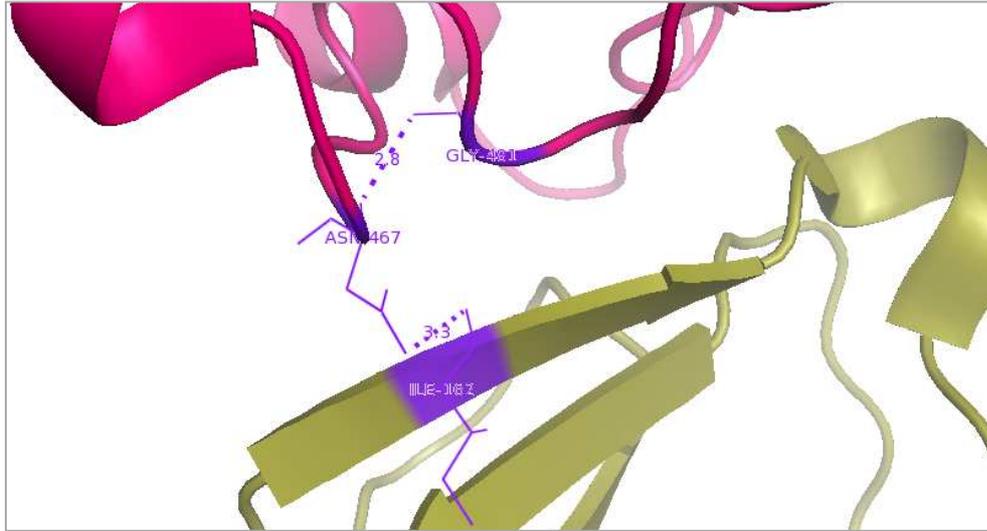
#### **4.1 Rationalisation of mutants and visualisation in PyMOL.**

Three mutants were generated for this study and were chosen based on their position and importance in the Ang2-Tie2 complex. Mutations were made not to disrupt the structure but chosen so as to have an affect or alter binding affinity to Tie2.

Visualisation of Ang2:Tie2 interface was performed using PyMOL (DeLano 2002). It should be noted that PyMOL is not a structure prediction tool, but imaging software to visualise protein 3-D structure.

Crystal structure has suggested N467 as a residue important in binding. Mutation of a residue K468 (as a double mutant K468/K472) adjacent to N467 renders the protein incapable of Tie2 binding (Barton, Tzvetkova & Nikolov 2005, Barton et al. 2006). Hence, it is possible that N467 may not be able to tolerate any mutation all. However, closer analysis of the gel filtration binding assay suggests that K468 mutant may still be capable of binding Tie2 at lower affinity than the WT protein (Barton, Tzvetkova & Nikolov 2005). Furthermore, N467 residue has been implicated to be crucial in binding as determined from the crystal structure. This Ang2 residue N467 differs in Ang1 with glycine at position 467, which is an amino acid with very different physicochemical properties that asparagine. Hence, mutation at N467 residue may have an effect on binding.

Figure 4.4 shows visualisation of the residues which may interact and form H bonds with Asn467 within the Ang2 protein. Ang2 Asn467 forms an H bond with Ang2 Gly481 and an H bond with Tie2 Ile162 and also contacts Tie2-Phe161 (Barton et al. 2006).



**Figure 4.4: Residues which interact and form H bonds with Ang2-Asn467.** Ang2 Asn467 may form an H bond (2.8) with Ang2 Gly481 and an H bond (3.3) with Tie2 Ile162. Approximate H bond strengths are indicated for each bond. The Ang2 P domain is coloured pink and the Tie2 binding interface is coloured olive.

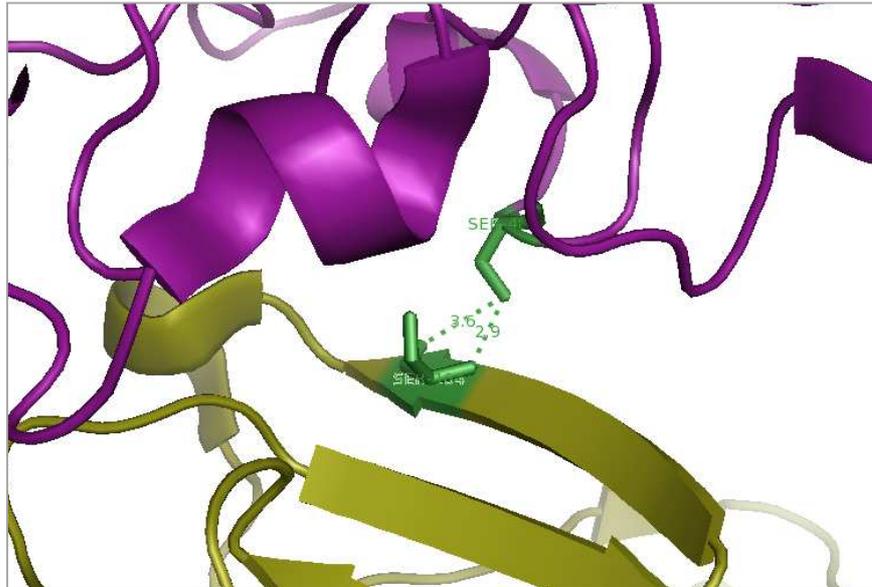
Mutation of Asn467 to serine is unlikely to alter the H bond with Tie2-Ile162, However Ser being slightly less polar than Asn may add rigidity to its interaction with Tie2-Ile162 and possibly affect binding.

Thus Asn467Ser mutant was selected to be analysed to explore the Ang2-Tie2 interface.

Residue Ser480 has previously been discussed as being important in the formation of Ang2-Tie2 complex (Barton et al. 2006). It is a residue which on mutation is likely to have an impact on binding affinity because of its position on the Ang2 binding interface, H bond formation with Tie2 Ser164 and it being a residue which differs from Ang1 (Pro480). Hence residue 480 was selected for mutagenesis. Though Ser480 has been implicated at being an important residue in complex formation, it has not been studied previously by mutagenesis studies. Ang1 has a Proline at position 480 which has a very non-reactive side chain and is unable to adopt many protein main-chain

conformations as compared with Ang2 Ser480. Different properties of Ser and Pro may contribute differential binding affinity of Ang2 and Ang1 to Tie2.

Figure 4.5 shows visualisation (in PyMOL) of the residues which may interact and form H bonds with Ser480 within the Ang2 protein. Ang2 Ser480 may be capable of forming two strong H bonds with Tie2-Ser164.



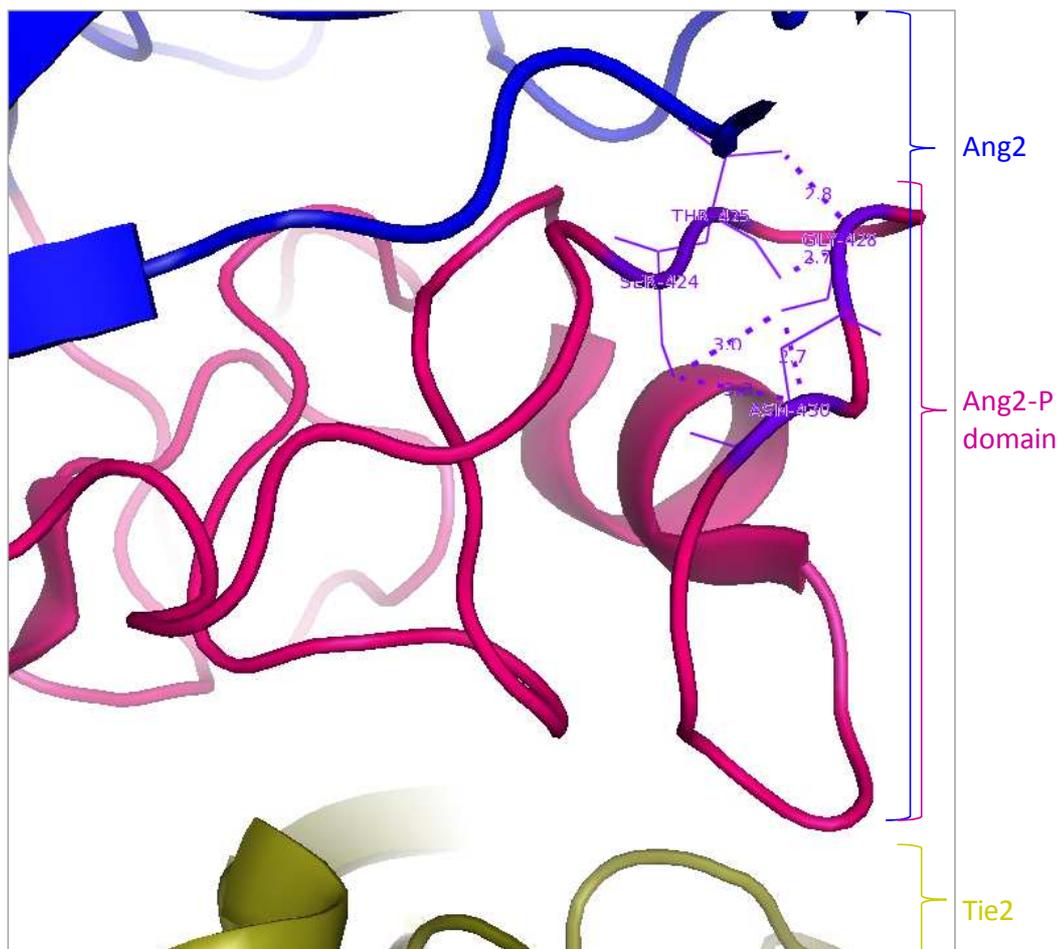
**Figure 4.5: Residues which interact and form H bonds with Ang2-Ser480.** Ang2 Ser480 may be capable of forming two strong H bonds (3.6 and 2.9) with Tie2-Ser164. Approximate H bond strengths are indicated for each bond. The Ang2 P domain is coloured pink and the Tie2 binding interface is coloured olive.

Ang1 residue 480 is a proline, an amino acid very different from any other with respect to its structure. Structurally, the amino nitrogen cyclised with the side chain terminal carbon makes proline a very rigid amino acid contributing to tight turns or kinks in the protein structure. Mutation of Ser480 to proline is likely to contribute structural changes in the region which may alter binding affinity to Tie2. Ser480Pro was therefore selected as the second mutant to be analysed to explore the Ang2-Tie2 interface and the impact of substitution with proline as occurs for Ang1.

Angiopoietins contain a conserved  $\text{Ca}^{2+}$  binding site. Residue Asp429, Asp431, Cys433 and Cys435 are important residues in the  $\text{Ca}^{2+}$  binding site of the Ang2-RBD. This site may not be involved in binding but have been suggested to stabilize the  $\beta 6$ - $\alpha 5$  loop in the structure (Barton et al. 2006).

Residue Gly428 lies near the  $\text{Ca}^{2+}$  binding site of the Ang2-RBD, but has not been implicated to contact any Tie2 residue (Barton et al. 2006). Glycine is the smallest amino acid and its single H side chain allows it to tolerate hydrophilic and hydrophobic environment in a protein structure. Its' structure allows conformational flexibility. However, this Ang2 Gly428 residue is substituted in Ang1 with Ala428 which is an amino acid with similar properties to glycine and hence this residue is unlikely to tolerate mutations without destabilizing the fold.

Figure 4.6 shows residues which may interact and form H bonds with Gly428 within the Ang2 protein. Ang2 Gly428 may be capable of forming an H bond with Ang2 Ser424, two H bond with Ang2 Thr425 and an H bond with Ang2 Asn430. Residues 424, 425 and 430 with which Gly428 may interact are conserved between all the angiopoietins. Thus mutation at this position should potentially be interesting to study the interface. Mutation of Gly428Thr was selected for analysing its impact on Ang2:Tie2 complex. Threonine is a slightly polar amino acid, and can reside within or on the surface of proteins. Most amino acids contain an H and non-H group on their C- $\beta$  carbon. The C- $\beta$  carbon in threonine contains a methyl and hydroxyl group which makes this amino acid bulky and restricted in the conformational flexibility of its main chain. Thus Gly428Thr was selected as the third mutant to be analysed to explore the Ang2-Tie2 interface.



**Figure 4.6: Residues which interact and form H bonds with Ang2-Gly428.** Ang2 Gly428 may form H bond (3.0) with Ang2 Ser424, two H bonds (2.7 and 2.8) with Ang2 Thr425 and an H bond (2.7) with Ang2 Asn430. Approximate H bond strengths are indicated for each bond. The Ang2 P domain is coloured pink and the Tie2 binding interface is coloured olive.

#### 4.2 Site directed mutagenesis and cloning to generate mutants Gly428Thr, Ser480Pro and Asn467Ser.

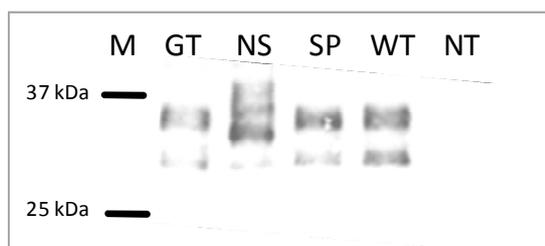
The mutants Gly428Thr, Ser480Pro and Asn467Ser were generated by site directed mutagenesis and cloned into a mammalian surface expression vector allowing expression of the mutated FReD domain on the cell surface of mammalian cells (like HEK293) anchored by the ASGPR transmembrane domain under the influence of a CMV promoter. Details of the surface expression system employed are described in Chapter 5.

Plasmids were sequenced (Sequence data Appendix 4C, 4D and 4E) to confirm that the required mutation had been incorporated in each of the mutant plasmids. These mutants were characterised by Western blotting and determination of dissociation constants. The  $K_d$  of each mutant expressed on the surface of HEK cells was compared with the WT Ang2-FReD.

### 4.3 Western blot analysis for mutants Gly428Thr, Ser480Pro and Asn467Ser in HEK293 cell line.

In order to confirm that mutant forms of Ang2-FReD were expressed in the cells, Gly428Thr, Ser480Pro and Asn467Ser mutant plasmids together with wild-type plasmid were transfected in HEK293 cells and the cell lysate analyzed by Western blotting (Figure 4.7).

The FLAG-tagged mutant and WT Ang2-FReD protein was detected by probing the nitrocellulose membrane with anti-FLAG primary and anti-mouse HRP linked secondary antibody as described in section 2.19.1.



**Figure 4.7: Western blot analysis of mutants, Gly428Thr, Ser480Pro and Asn467Ser in HEK293.** HEK293 cells were transfected with the Wild type Ang2-FReD (WT) and mutant Gly428Thr (GT), Ser480Pro (SP) and Asn467Ser (NS) plasmids. Cell lysate was analyzed 24hours post-transfection by Western blotting. The correct size for ASGPR linked FLAG-tagged WT Ang2-FReD is ~29.5 kDa. Bands between 25 to 37 kDa are observed in lane GT, NS, SP and WT. No bands were observed in the non-transfected (NT) lane. This experiment was performed independently three times with similar results.

As shown in Figure 4.7 immunoreactive bands were observed in lysates from WT and mutants transfected cells but not for non-transfected cells. For Gly428Thr (GT),

Ser480Pro (SP) mutants and WT forms of Ang2-FReD proteins two forms of each protein were detected at approximately 30kDa and 35kDa. Based on the amino acid sequences the expected size of wild-type and mutant proteins is approximately 29kDa. The lower molecular mass bands observed in lysates for transfected cells is close to the expected size of 29kDa and may represent non-glycosylated forms of the proteins. The high molecular mass bands observed for WT and mutant transfected cells may correspond to glycosylated forms of the protein. It is interesting that the Asn467Ser (NS) mutant exhibits multiple forms of the protein approximately between 32 and 37kDa. These bands may indicate that this protein is differentially processed; either differentially glycosylated or folded. Whether the highest molecular form of the Asn467Ser mutant protein is active on the cell surface cannot be concluded.

#### **4.4 Influence of protein expression levels on $K_d$ determination by flow cytometry.**

Flow cytometry combined with cell surface display is a robust technique which produces reproducible output that can accurately rank a large set of candidates by  $K_d$  determinations in a quantitative manner (Lofblom et al. 2007). However, it is possible that measurement of binding affinity could be influenced by the transfection efficiency of the plasmids and corresponding expression level of the fraction of transfected cells. A preliminary experiment was therefore conducted to test this.

Three sets of WT transfected cells (WT A, WT B and WT C) having a range of transfection efficiency were used. Cells were stained for the FLAG tag using Cy2 fluorophore as described in section 2.22.3 and the geometric mean (Gm) for FLAG expression level was determined.

Table 4.2 outlines the %transfection and the Gm values for each set of WT transfected cells.

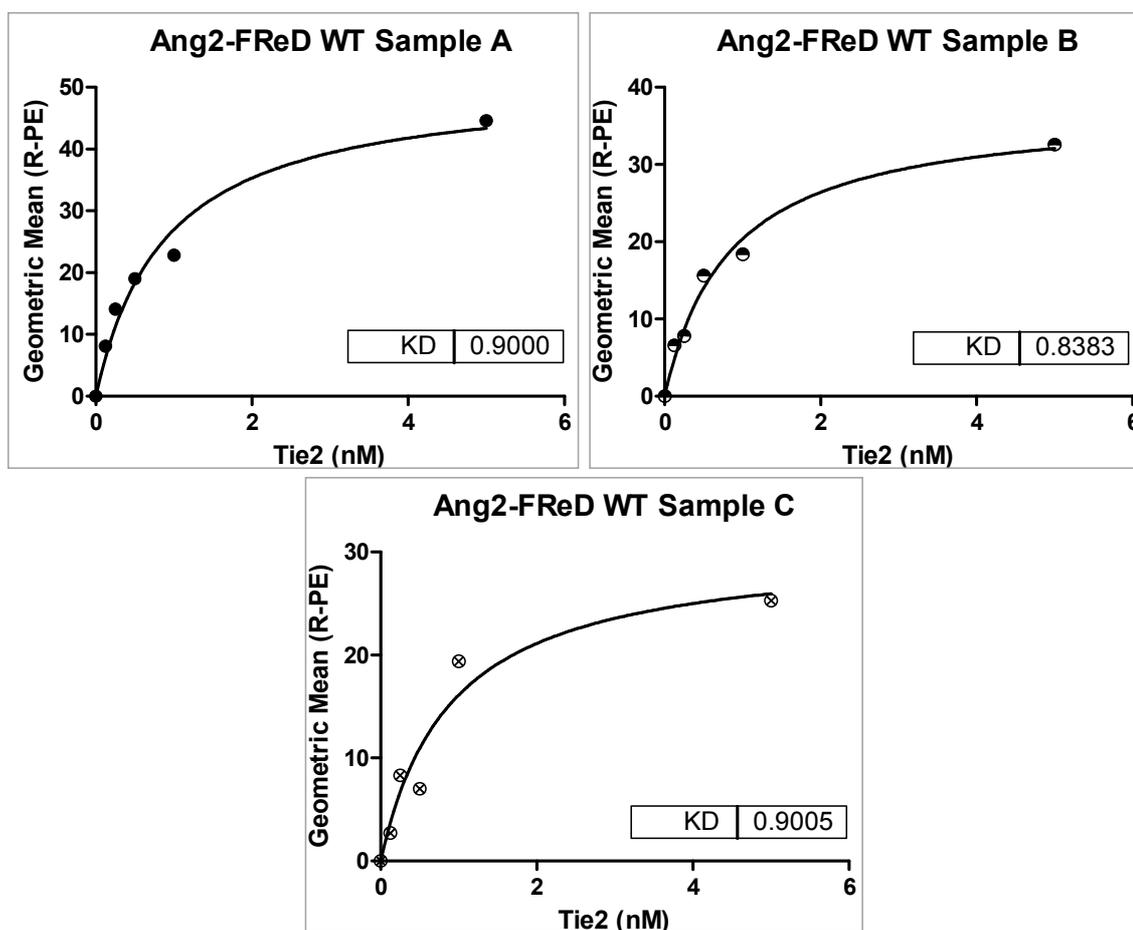
Sample	%Transfection	Gm for FLAG tag expression
WT A	47.28	244.2
WT B	20.49	206.7
WT C	2.01	161.7

**Table 4.2: Three sets of WT Ang2 (WT A, WT B and WT C) transfected with varying transfection levels as determined by percentage transfection and Geometric mean (Gm) for FLAG tag expression.**

Each sample WT A, WT B and WT C display a different percentage transfection and different levels of expression of Ang2-FReD protein. These samples were tested to determine whether the expression levels and % transfection influenced their  $K_d$  values.

The  $K_d$  for samples WT A, WT B and WT C was determined by flow cytometry as described in section 2.22.3. To obtain a concentration curve of binding for the samples of WT protein to the soluble *rhTie2* ectodomain, cells were incubated with a range of soluble *rhTie2* concentrations. These were then washed to remove unbound Tie2 and incubated with biotin-linked anti-His primary antibody to detect the His-tagged Tie2 bound on the cell surface of the WT Ang2-FReD domain. The cells were washed and then incubated using fluorescently labelled Streptavidin-R-Phycoerythrin. The red fluorescence of the phycoerythrin fluorophore was detected on the flow cytometer for each of the Tie2 concentrations for the samples of WT Ang2-FReD with varying expression levels. The geometric mean for the phycoerythrin fluorescence (red) for each Tie2 concentration was obtained. To determine  $K_d$  values non-linear regression analysis was performed of the geometric mean fluorescence value versus concentration of *rhTie2* Fc chimera using Graphpad Prism.

Figure 4.8 shows concentration curves for each of the WT Ang2-FReD samples A, B and C.



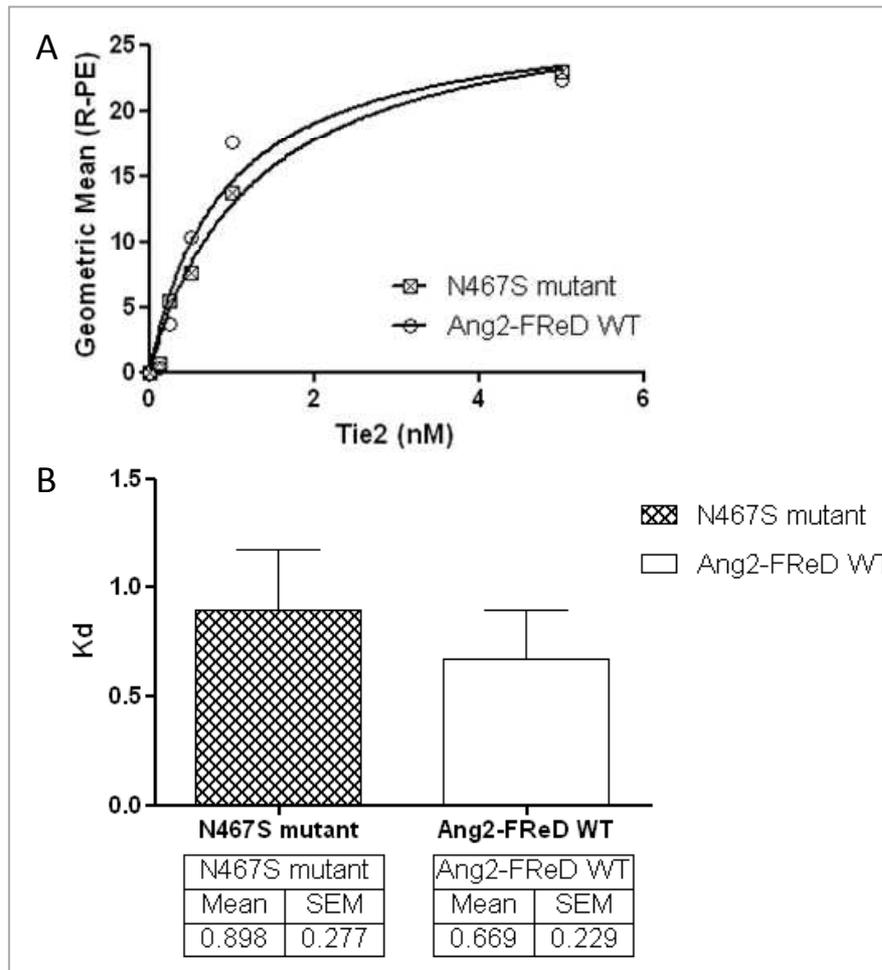
**Figure 4.8:** Tie2 concentration curve to show influence of WT Ang2-FReD protein expression levels on  $K_d$  determination in HEK293 cell line by flow cytometry. HEK cells were transfected with the WT plasmids to obtain cells with varying transfection and expression levels. Aliquots of these cell samples A, B and C were incubated with varying concentrations of soluble Tie2 ectodomain. His-tagged Tie2 bound on the cell surface of the WT FReD was detected using biotin-linked anti-His primary antibody and Streptavidin-R-Phycoerythrin. The geometric mean for red fluorescence for each Tie2 concentration was obtained. To determine  $K_d$  values non-linear regression analysis was performed of the geometric mean fluorescence value (Y axis) versus concentration of Tie2 (X axis) using Graphpad Prism. Concentration curves obtained for the WT Ang2-FReD sample A, B and C are shown with their  $K_d$  values.

Although samples A, B and C transfected with WT Ang2-FReD plasmid display large variation in percentage transfection and expression levels described in Table 4.2, the concentration curves and the  $K_d$  values obtained were similar for all samples A (0.9nM), B (0.8383nM) and C (0.9005nM). Thus  $K_d$  is not influenced by expression levels of the protein in the surface display system and flow cytometry assay used.

#### 4.5 $K_d$ determination of mutants Gly428Thr, Ser480Pro and Asn467Ser by flow cytometry.

Plasmids encoding mutants Gly428Thr, Ser480Pro and Asn467Ser were characterised for their phenotype by  $K_d$  determination using flow cytometry as described in section 2.22.3. The mutant plasmids Gly428Thr, Ser480Pro, Asn467Ser and the WT plasmid were transfected into HEK293 cells and 24hours post-transfection cells were collected. To obtain a concentration curve of binding for the mutants and WT protein to the soluble *rh*Tie2 ectodomain the transfected cells were incubated with a range of soluble *rh*Tie2 concentrations. These were then washed to remove unbound Tie2 and incubated with biotin-linked anti-His primary antibody to detect the His-tagged Tie2 bound on the cell surface of mutant and the WT Ang2-FReD domain. The cells were washed and then incubated using fluorescently labelled Streptavidin-R-Phycoerythrin. The red fluorescence of the phycoerythrin fluorophore was detected on the flow cytometer for each of the Tie2 concentrations for the mutant and the WT Ang2-FReD domain. The geometric mean for the phycoerythrin fluorescence (red) for each Tie2 concentration was obtained. To determine  $K_d$  values non-linear regression analysis was performed of the geometric mean fluorescence value versus concentration of *rh*Tie2 Fc chimera using Graphpad Prism. The  $K_d$  for each mutant was determined by three independent experiments.

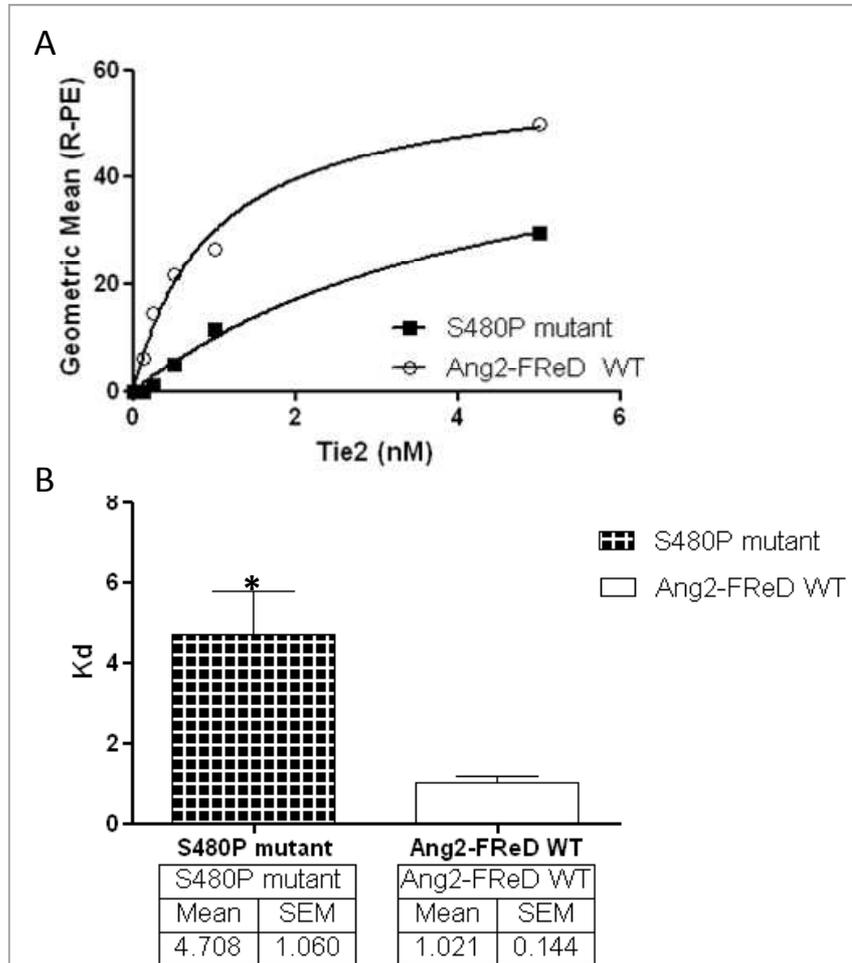
Figure 4.9 is a representative concentration curve obtained for the Asn467Ser mutant and WT Ang2-FReD. The bar graph shows the mean  $K_d$  value of mutant Asn467Ser and WT domain for three independent experiments.



**Figure 4.9: Representative Tie2 concentration curve for WT Ang2-FReD domain and mutant Asn467Ser in HEK cell line.** HEK cells were transfected with the WT and mutant plasmids and incubated with varying concentrations of soluble Tie2 ectodomain. His-tagged Tie2 bound on the cell surface of Asn467Ser FReD mutant domain and the WT FReD domain was detected using biotin-linked anti-His primary antibody and Streptavidin-R-Phycoerythrin. The geometric mean for red fluorescence for each Tie2 concentration was obtained. To determine  $K_d$  values non-linear regression analysis was performed of the geometric mean fluorescence value (Y axis) versus concentration of Tie2 (X axis) using Graphpad Prism. A) Representative concentration curve obtained for the Asn467Ser mutant and WT Ang2-FReD. B) Bar graph shows the mean  $K_d$  value of mutant Asn467Ser of 0.90nM with SEM of +/- 0.277 and  $K_d$  value of WT domain was 0.67nM with SEM of +/- 0.229. (n=3).  $K_d$  value for Asn467Ser mutant were found to be similar to that of WT Ang2-FReD domain.

The  $K_d$  value of the mutant Asn467Ser was obtained and found to be similar to that of WT Ang2-FReD domain. This experiment was performed three times and a  $K_d$  mean of 0.90nM with a SEM of +/- 0.277 obtained for Asn467Ser. The  $K_d$  mean of 0.67nM with a SEM of +/- 0.229 was obtained for WT FReD.

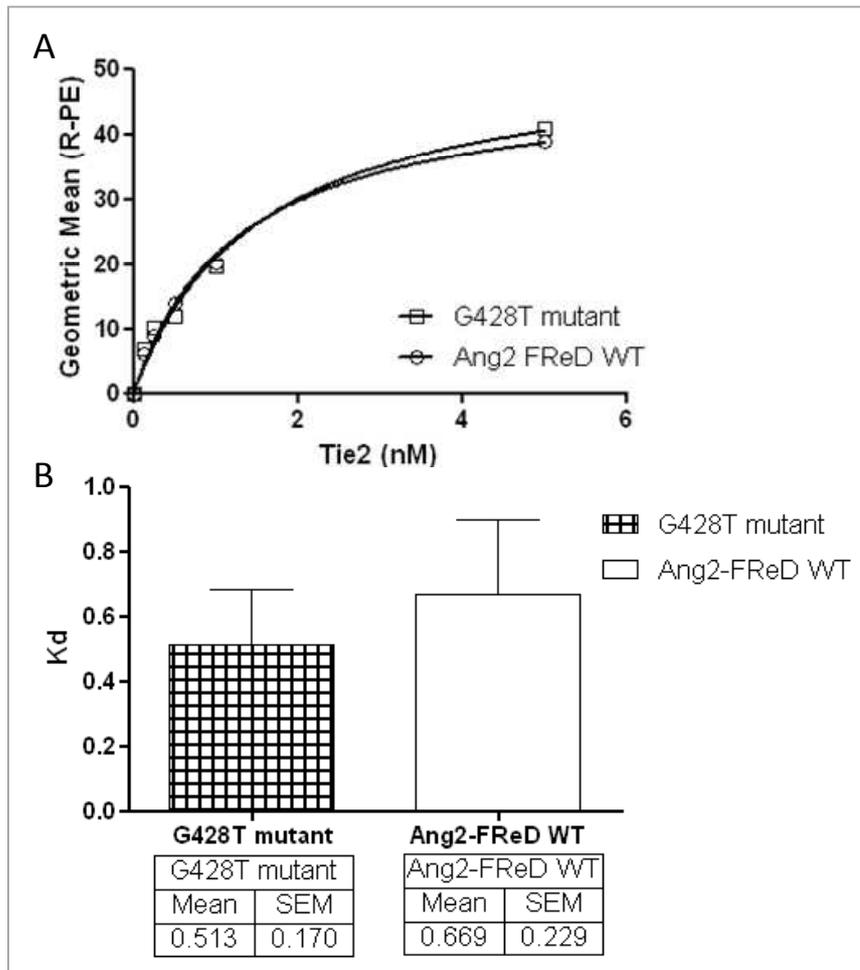
Figure 4.10 is a representative concentration curve obtained for the Ser480Pro mutant and WT Ang2-FReD. The bar graph shows the mean  $K_d$  value of mutant Ser480Pro and WT domain for three independent experiments.



**Figure 4.10: Representative Tie2 concentration curve for WT Ang2-FReD domain and mutant Ser480Pro in HEK cell line.** HEK cells were transfected with the WT and mutant plasmids and incubated with varying concentrations of soluble Tie2 ectodomain. His-tagged Tie2 bound on the cell surface of Ser480Pro FReD mutant domain and the WT FReD domain was detected using biotin-linked anti-His primary antibody and Streptavidin-R-Phycoerythrin. The geometric mean for red fluorescence for each Tie2 concentration was obtained. To determine  $K_d$  values non-linear regression analysis was performed of the geometric mean fluorescence value (Y axis) versus concentration of Tie2 (X axis) using Graphpad Prism. A) Representative concentration curve obtained for the Ser480Pro mutant and WT Ang2-FReD. B) Bar graph shows the mean  $K_d$  value of mutant Ser480Pro of 4.71nM with SEM of +/- 1.060 and  $K_d$  value of WT domain was 1.02nM with SEM of +/- 0.144. (n=3).  $K_d$  value for Ser480Pro mutant were found to be approximately 5 times more than that of WT Ang2-FReD domain. The binding affinity of the Ser480Pro mutant was significantly different from the WT Ang2-FReD domain. \*  $t_{(4)} = 3.4470$ ,  $p < 0.05$ , >95% confidence level.

The  $K_d$  value of the mutant Ser480Pro was obtained and found to be approximately 5 times more than that of WT Ang2-FReD domain. This experiment was performed three times and a  $K_d$  mean of 4.71nM with a SEM of +/- 1.060 obtained for Ser480Pro. The  $K_d$  mean of 1.02nM with a SEM of +/- 0.144 was obtained for WT FReD. Mutant Ser480Pro had a significantly higher  $K_d$  than WT.

Figure 4.11 is a representative concentration curve obtained for the Gly428Thr mutant and WT Ang2-FReD. The bar graph shows the mean  $K_d$  value of mutant Gly428Thr and WT domain for three independent experiments.



**Figure 4.11: Tie2 concentration curve for WT Ang2-FReD domain and mutant Gly428Thr in HEK cell line.** HEK cells were transfected with the WT and mutant plasmids and incubated with varying concentrations of soluble Tie2 ectodomain. His-tagged Tie2 bound on the cell surface of Gly428Thr FReD mutant domain and the WT FReD domain was detected using biotin-linked anti-His primary antibody and Streptavidin-R-Phycoerythrin. Following flow cytometry the geometric mean for red fluorescence for each Tie2 concentration was obtained. To determine  $K_d$  values non-linear regression analysis was performed of the geometric mean fluorescence value (Y axis) versus concentration of Tie2 (X axis) using Graphpad Prism. A) Representative concentration curve obtained for the Gly428Thr mutant and WT Ang2-FReD. B) Bar graph shows the mean  $K_d$  value of mutant Gly428Thr of 0.51nM with SEM of +/- 0.170 and  $K_d$  value of WT domain was 0.67nM with SEM of +/- 0.229. (n=3).  $K_d$  value for Gly428Thr mutant were found to be similar to that of WT Ang2-FReD domain.

The  $K_d$  value of the mutant Gly428Thr was obtained and found to be similar to that of WT Ang2-FReD domain. This experiment was performed three times and a  $K_d$  mean of 0.51nM with a SEM of +/- 0.170 obtained for Gly428Thr. The  $K_d$  mean of 0.67nM with a SEM of +/- 0.229 was obtained for WT FReD.

Significant difference was observed for Ser480Pro mutant, but Asn467Ser and Gly428Thr mutants had a  $K_d$  similar to WT.

## Discussion

The Ang-Tie2 system is involved in a varied range of vascular protective roles and a balance between the agonistic action of Ang1 and largely opposing action of Ang2 holds together the intricate homeostasis of this system. The importance of the angiopoietin ligands has already been proven in various clinical and animal studies. Accordingly, Ang1 and Ang2 ligands have been studied extensively to determine their roles in the cellular context. However, the molecular bases by which these ligands show distinct actions on binding to the same receptor is yet unknown. Like many proteins, the complexity of structure at the receptor binding domain may determine their differential actions.

Ang1 and Ang2 have been suggested to bind to the Tie2 receptor in a similar manner. This view has been shared by several studies based on their shared high degree of homology in the receptor binding interface, crystal structural studies on the Ang2:Tie2 and affinity studies (Barton, Tzvetkova & Nikolov 2005, Barton et al. 2006, Maisonpierre et al. 1997, Davis et al. 1996, Davis et al. 2003). Recently, the affinity of Ang2 to Tie2 was suggested to be 20-fold lower than Ang1, and it has been suggested that this may be one of the factors making its action on Tie2 signalling less potent than Ang1 and that this difference may contribute to the differential actions of the angiopoietins (Yuan et al. 2009). However, this is an isolated report suggesting differential binding affinity between Ang1 and Ang2 and it is largely accepted that they both bind to Tie2 with similar affinities (Maisonpierre et al. 1997, Davis et al. 1996).

The aim of this chapter was to explore the Ang2:Tie2 interface residues by rational substitutions. Thus, three important residues in the receptor binding domain of the Ang2 ligand were analysed by mutational studies. All three residues analysed are not conserved between both the angiopoietins and hence may be responsible for the differential binding and hence differential action of the proteins. The mutants were characterised by protein expression and binding assays.  $K_d$  determinations were performed using cell surface display of the mutant proteins combined with flow cytometry.

Mammalian cell surface display holds an advantage over using secreted proteins which need to be extracted, purified, and immobilized, and there is a risk that at least a portion of the proteins have lost some binding functionality (Lofblom et al. 2007).

Ang2 residues Ser480 and Asn467 in the P domain of the protein have been suggested to be crucial for binding by forming hydrogen bonds with Tie2 residues as determined by the Ang2:Tie2 crystal structure (Barton et al. 2006). These residues are not conserved in Ang1 where they are replaced as Pro480 and Gly467. Both proline and glycine are very different amino acids with chemical properties different to serine and asparagine. To test whether the Pro480 of Ang1 contributed to stronger or weaker binding, Ang2 Ser480 was mutated to the residue found in Ang1 at the 480 position; a proline. To test the importance of Ang2 Asn467, it was mutated to a serine, a residue different from residue at this position in Ang1Gly467.

Ang2 residue Gly428 lies in the P domain but is not implicated in direct interaction with Tie2 residues. Its corresponding residue in Ang1 is an alanine, which is similar to glycine, is a relatively small and largely neutral amino acid. Ang2 Gly428 lies in a region surrounded by residues which are conserved between Ang1 and Ang2 at the

Ca<sup>2+</sup> binding region involved in stabilisation of the complex structure. Mutation of Gly428 residue to a bulky and rigid amino acid, threonine was performed to test if this destabilised the structure or affected binding.

The Ser480Pro, Asn467Ser and Gly428Thr mutants generated by site directed mutagenesis were expressed as a fusion protein linked to the transmembrane anchor protein of the human asialoglycoprotein receptor (ASGPR) protein. This anchor protein was employed to express the Ang2-FReD protein oriented with its receptor binding interface contained at the C-terminal on the extracellular cell surface. The proteins were FLAG epitope tagged to detect expression.

The Ser480Pro, Asn467Ser and Gly428Thr mutant proteins were characterised for expression in mammalian (HEK293) cells by Western blot analyses of the transfected cell lysates. All mutant Ang2-FReD proteins, as well as WT Ang2-FReD protein showed a low molecular band probably representing a non-glycosylated form of the protein on the Western blot. Mutants Ser480Pr and Gly428Thr showed similar sized immunoreactive bands as the WT protein between 25-37kDa. The Asn467Ser mutant showed more than one (possibly 2-3) high molecular mass bands. The higher molecular bands may indicate that this mutant protein is differentially processed, possibly differentially glycosylated in the golgi apparatus. Glycosylation has an effect on the pharmacokinetic property of a potential therapeutic protein. The heterogeneity is contributed by the variable occupancy of potential glycosylation sites, referred as glycosylation macroheterogeneity and also by the diversity in the type of carbohydrate structures that occur at individual glycosylation sites referred as glycosylation microheterogeneity. The heterogeneity affects the mass, charge and hydrodynamic volume of proteins. This effect may have novel action *in vivo* as seen for a novel engineered form of therapeutic recombinant human erythropoietin (rHuEPO). The non-

glycosylated form of EPO demonstrated 3-fold activity *in vitro* but no activity *in vivo* (Higuchi et al. 1992, Yuen et al. 2003). Whether the highest molecular forms of the Asn467Ser mutant protein is active on the cell surface or has different properties *in vivo* was not tested here. The heterogeneity of glycosylation can be analysed by capillary electrophoresis or mass spectrometry.

Before determination of the binding affinity of the mutant proteins, the effect of transfection efficiency and protein expression on  $K_d$  determination was analysed. The cell surface expression of mutants may depend on the folding and translocation of each mutant protein to the cells surface. This variation in expression may influence the binding detected for each mutant protein or may be a limiting factor in the accuracy of  $K_d$  determination. It is possible that the cell surface expression system used may have a threshold limit in transfection efficiency and expression level, below which the protein is in depleting amount on the cell surface for the range of Tie2 concentrations it is assayed for. Hence it was crucial to determine to what extent the binding of Tie2 was dependent upon the levels of expression and transfection of the FLAG-tagged protein. This was assessed by performing a Tie2 binding concentration curve for cells transfected with WT Ang2-FReD plasmid to obtain different expression and transfection efficiencies. It was found that for transfection efficiency, ranging from as low as 2% to 47% and for expression level (determined by geometric mean for FLAG tag detection) ranging from 161 to 244, the  $K_d$  value could be accurately determined by flow cytometry. The expression levels and transfection efficiency of all the mutant proteins was found to lie within this range and hence the  $K_d$  for all could be determined accurately using one-colour flow cytometry.

Each mutant ligand was characterized for its specific affinity to the Tie2 receptor by a binding assay quantitated by flow cytometry. Changes at important residues on the binding interface can result in structural changes which may alter binding of Ang2-RBD to the Tie2 receptor. Changes in binding affinity of the mutants were measured by determining the Dissociation Constant ( $K_d$ ).  $K_d$  is the free receptor concentration at which 50% of ligand is bound. The Ser480Pro, Asn467Ser and Gly428Thr mutant proteins were characterised for binding by determination of Dissociation Constant ( $K_d$ ) in mammalian (HEK293) cells by flow cytometry.

Mutant Asn467Ser protein displayed  $K_d$  value similar to the WT Ang2-FReD protein. Given that Asn467 has been implicated to be important for binding, one would expect any mutation at this residue would affect binding. However, Ang1 Gly467 is a very small and chemically different amino acid compared to Ang2 N467. Also mutation of neighbouring Ang2 K468E did not result in complete loss of binding to Tie2 (Barton, Tzvetkova & Nikolov 2005). Hence it is likely that Asn467 position in Ang2 is tolerant to rational substitutions and can be explored by mutagenesis without destabilising the fold or loss in binding function.

Mutant Ser480Pro displayed  $K_d$  value approximately 5-fold more than the WT Ang2-FReD protein. The SP mutant was generated by mutagenizing residue Ser480 of Ang2 to proline of Ang1. It is possible that the binding affinity in Ang1 is contributed by other amino acids at the interface to compensate for the 5-fold lower affinity of the Ser480Pro mutant than that of the WT protein. It is also possible that the interaction of Pro480 within the Ang1 protein is favourable and does not cause lower binding affinity.

Mutant Gly428Thr protein displayed  $K_d$  value similar to the WT Ang2-FReD protein. Given that Gly428 lies in a conserved region responsible for stabilizing the structure,

one would expect a mutation to a bulky amino acid like threonine was very likely to destabilize the structure. However, this was not observed and the mutant displayed similar binding affinity to the wild type protein domain.

It is evident from these mutants that the theoretical and practical outputs from rational substitutions are unpredictable. To explore the binding interface conclusively and rapidly a new method of protein engineering is required.

## Chapter 5

# Designing and Cloning of the Cell Surface Display Vector for Expression of Ang1 and Ang2 FReD domains in Ramos Cells

### Introduction

As discussed in section 1.5 there are several limitations to protein engineering by current methods of rational design or directed protein evolution. Somatic hypermutation (SHM) has been suggested as an alternative approach for directed protein evolution (Wang, Yang & Wabl 2004, Wang et al. 2004).

SHM driven mutagenesis within growing cells would avoid the need of laborious repetitive transfection and re-isolation of genes. Autonomously generated diversity *in situ* is able to explore the sequence space efficiently by accumulation of multiple reinforcing mutations throughout the gene (Wang et al. 2004). High-throughput selection and screen for improved phenotype would not be limited by the library size of conventional directed evolution methods. Current directed evolution approaches for engineering mammalian proteins employ bacterial and yeast systems wherein processes of post-translational modification, folding and cellular environment may affect mammalian proteins. Directed evolution by SHM allows evolution of mammalian proteins within mammalian cells.

However, to date SHM driven directed evolution has only been described for intracellular fluorescent proteins (Wang, Yang & Wabl 2004, Wang et al. 2004, Arakawa et al. 2008). The aim of experiments described in this chapter was to develop SHM driven directed protein evolution to evolve angiopoietins (Ang1, Ang2) receptor

binding domains (FReD) by combining mammalian surface display with mutagenesis driven by SHM.

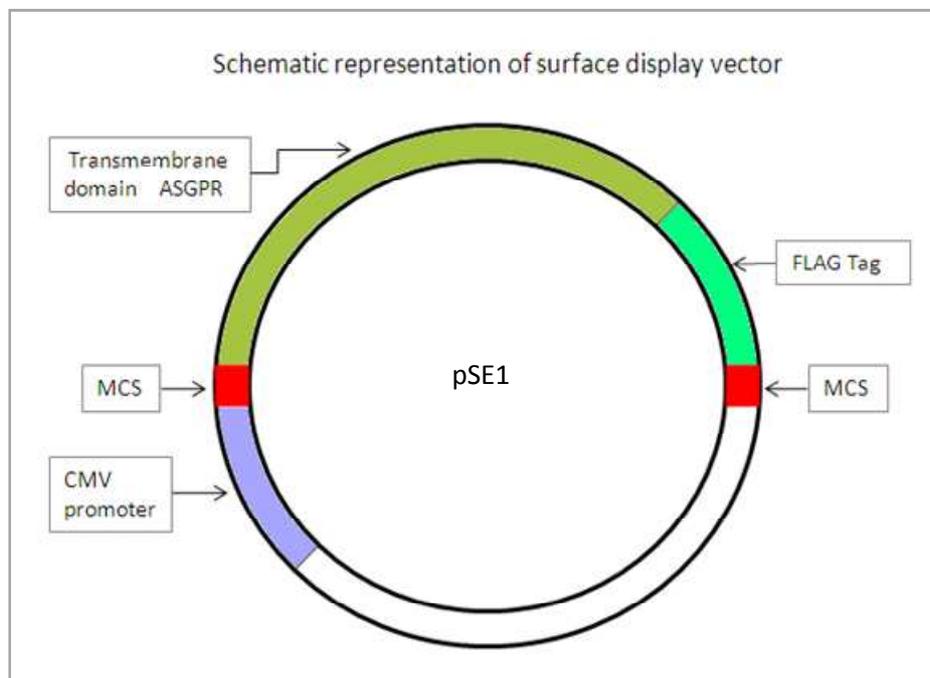
For establishment of this directed protein evolution method initial experiments were performed to construct a fusion protein that would allow expression of an extracellular membrane-tethered form of Ang1 and Ang2 FReD domain. Once the protein of interest was expressed on cellular surface, Tie2 ectodomain binding conditions were established. Somatic hypermutation (SHM) would be initiated to generate mutants with improved binding ability to the Tie2 receptor. Finally, improved mutants were selected iteratively by FACS.

The Burkitt lymphoma Ramos cell line previously used by Wang et al (2004) was utilised for generating mutants in the cell surface displayed angiopoietin fibrinogen domain (FReD). (Wang et al. 2004)

### **5.1 Cloning ASGPR in pcDNA3.1/V5-His-TOPO vector.**

To establish a system to express ligands on the cell surface it was necessary to construct an expression vector capable of expressing the protein of interest on the extracellular surface of the cell and anchored to the membrane by a transmembrane sequence. Since both the ligands; Ang1 and Ang2 have their receptor binding domain (RBD) at the carboxy-terminal it was necessary for the expressed protein to be oriented with a free carboxy-terminal. This can be achieved by using a transmembrane domain of type II membrane protein which has an internal secretory leader sequence and membrane 'stop transfer' sequence which retains the amino terminal domain on the intracellular side of the membrane and carboxy-terminus on the extracellular side of the plasma membrane (Holland, Drickamer 1986).

The surface display vector (pSE1) was designed to include a FLAG epitope tag to enable detection and a multiple cloning site for insertion of Ang domains (Figure 5.1).

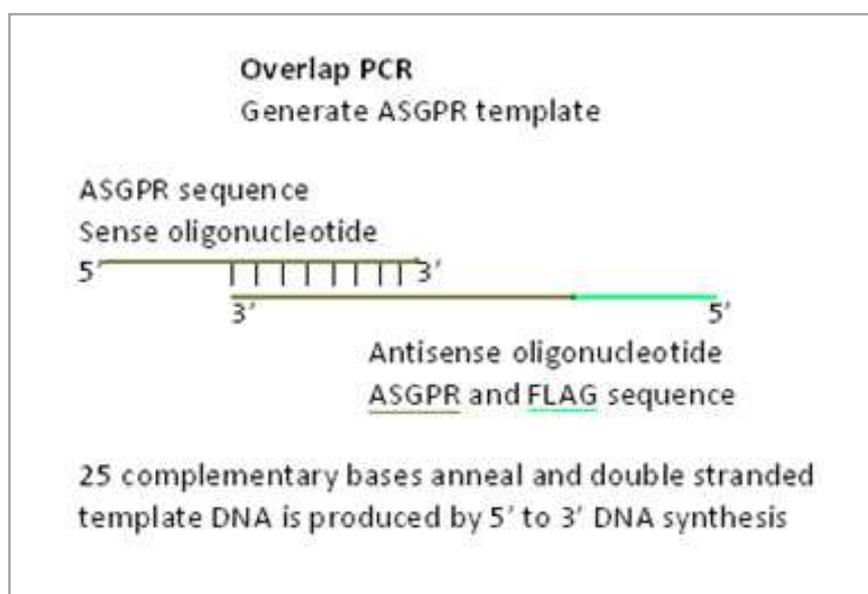


**Figure 5.1: Schematic representation showing the key features of the surface display vector (pSE1).** Cytomegalovirus promoter/enhancer allows high level constitutive expression of protein of interest; the FLAG tag allows the detection of the fusion protein and the multiple cloning site (MCS) enables ease in cloning proteins of interest.

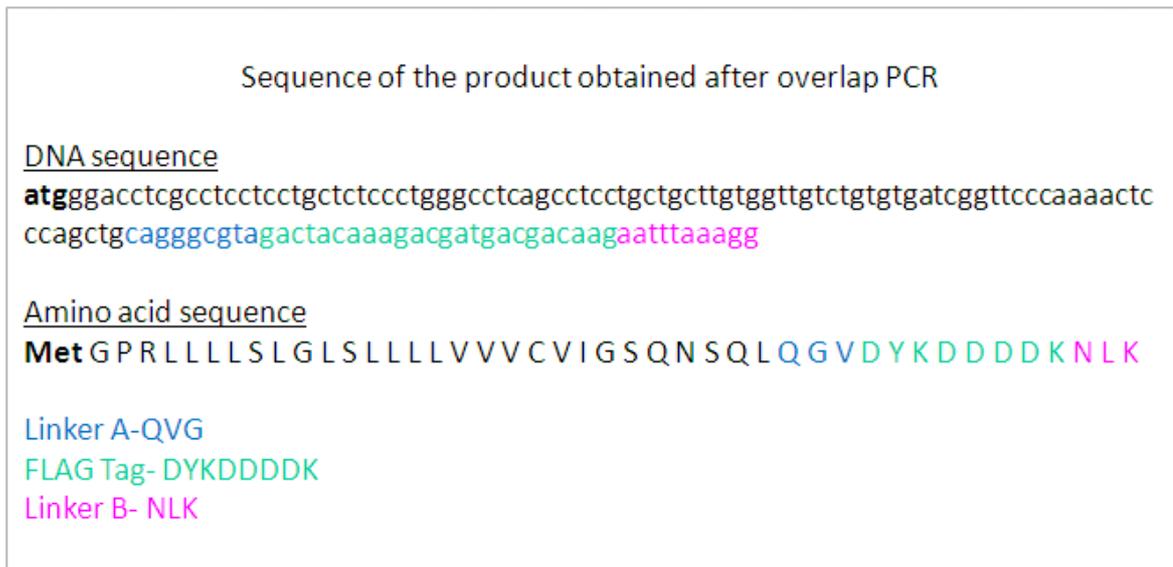
The asialoglycoprotein receptor (ASGPR) is a cell surface protein (291 residues) of human hepatocytes. The protein consists of 3 domains: a short amino terminal cytoplasmic domain of 37 residues, a hydrophobic membrane spanning segment of 20 amino acids and a large carboxy terminal extra-cytoplasmic domain of 225 residues (Spiess, Schwartz & Lodish 1985). The membrane spanning segment of asialoglycoprotein receptor (ASGPR) necessary and sufficient for insertion of the fusion protein in the membrane was chosen; specifically residues 38 to 65 which have been shown to direct protein to the extracellular face of the cell (Spiess, Lodish 1986).

Thus the ASGPR membrane protein can be used to anchor the various domains of the Ang-1 and Ang-2 protein. Cells expressing the ASGPR with the Ang domains can be used directly in cellular receptor binding assays.

The strategy for making the surface display vector involves two stages. First, the ASGPR template was generated by overlap PCR and the template was amplified (Figure 5.2). Next it was cloned into a mammalian expression vector. The ASGPR template was generated by overlap PCR in 7 PCR cycles and annealing temperature 60°C. The first round of overlap PCR requires sense and anti sense oligonucleotides with complementary overlapping ends to generate a double stranded template (Figure 5.3). In a second round of PCR this template is amplified with ASGPR 5' and 3' primers (Figure 5.4).

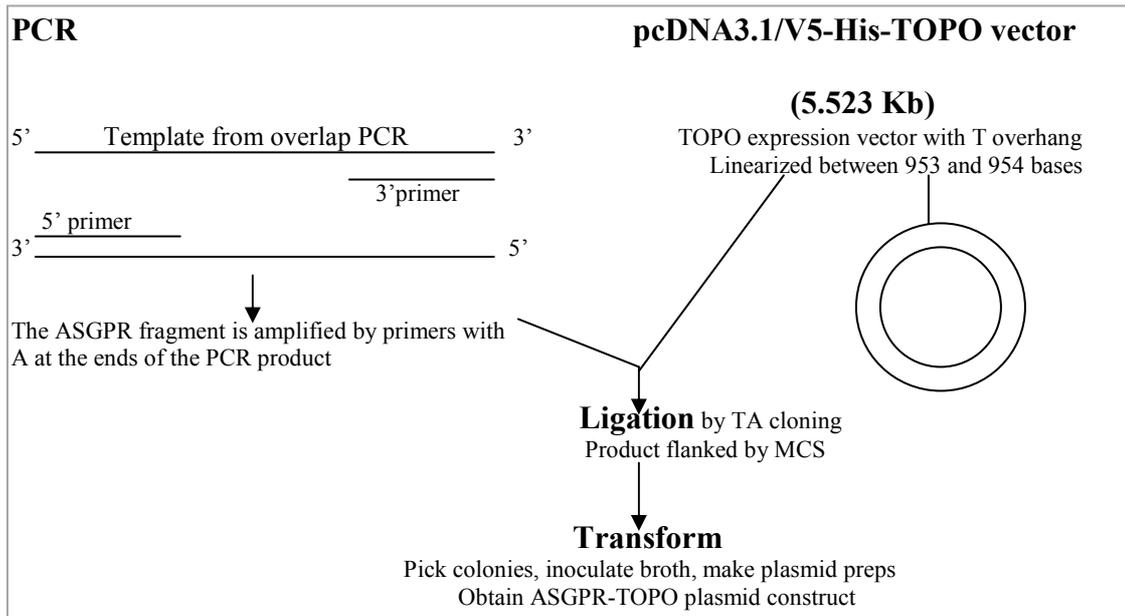


**Figure 5.2: Schematic representation of the first stage of overlap PCR to amplify transmembrane domain of asialoglycoprotein receptor (ASGPR).** Sense and anti sense oligonucleotides with complementary overlapping ends generate a double stranded template by overlap PCR. The transmembrane domain of ASGPR together with a 3' FLAG epitope tag was obtained.



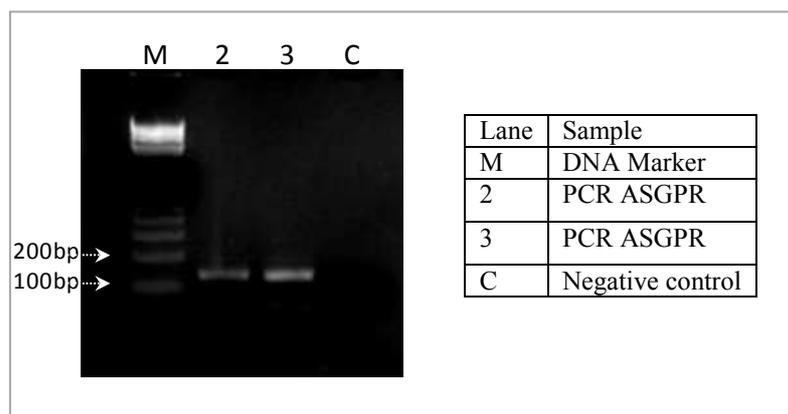
**Figure 5.3: The sequence of asialoglycoprotein receptor (ASGPR) transmembrane domain amplified by overlap PCR.** ASGPR construct consists of the transmembrane domain of ASGPR and a FLAG epitope tag linked to each other by the linker A coding for amino acids QVG. The FLAG epitope tag is followed by another linker B coding for amino acids NLK.

Linker A was incorporated to enable the FLAG tag to be expressed on the surface such that it is not tethered too close to the membrane of the cells and hence enable detection. Linker B was incorporated to enable cloning and expression of protein of interest such that it is not too close to the membrane and is presented on the surface of the cells available for binding to the receptor.



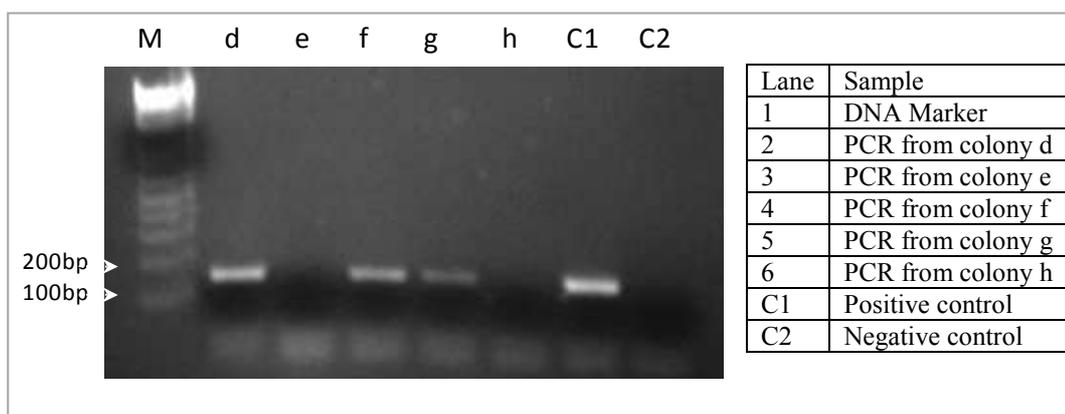
**Figure 5.4: Schematic representation of the cloning strategy for the ASGPR fragment in the pcDNA3.1V5-His-TOPO vector.** The template obtained by overlap PCR (figure 5.3) was amplified using 5' and 3' ASGPR primers at annealing temperature 60°C. The PCR product was cloned by TA cloning into a mammalian expression vector.

The PCR product from the overlap PCR was amplified in 30 PCR cycles using 5' and 3' ASGPR primers at annealing temperature 60°C. The expected size of the ASGPR PCR product is 132bp. The size of the amplified product was confirmed by running 1µl of the product on a 2% agarose gel (Figure 5.5).



**Figure 5.5: Amplification of the ASGPR template DNA obtained by overlap PCR.** Amplification is performed in 30 PCR cycles using 5' and 3' ASGPR primers. The expected size of ASGPR is 132bp. The required amplified product ASGPR in lane 2 and 3 is obtained between 100- 200bp.

The PCR product obtained was then ligated to the pcDNA3.1V5-His-TOPO vector by TA cloning. After completion of ligation, competent cells were transformed with the ligated DNA. Five colonies were picked and PCR screened using 5' and 3' ASGPR primers (Figure 5.6). Colonies d, f and g were chosen and inoculated in LB broth and overnight cultures were used to make plasmid preparations of the ASGPR-TOPO construct. The plasmid constructs were sequenced, however the sequencing data obtained showed that the ASGPR was cloned with point mutations or point deletions (See Appendix 4F for sequence data).



**Figure 5.6: PCR screen for ASGPR-pc3.1 constructs (pSE1).** Five colonies labelled ASGPR d-h were picked from the transformation plate and inoculated in selective medium. Plasmid preps were made from these and used for diagnostic PCR using the 5'ASGPR and 3'ASGPR primers to screen for the presence of the desired insert. Bands between 100- 200bp were obtained for plasmid d, f, and g.

Furthermore 15 colonies from the transformation plate were picked and analyzed by sequencing. All of the clones sequenced had a minimum of two to a maximum of five point mutations. Some clones had point deletions. These mutations contributed significant changes in the protein sequence generating stop codons or major amino acid changes. In some cases the deletions altered the reading frame.

## **5.2 Modifications to clone the ASGPR into the pcDNA3.1/V5-His-TOPO vector.**

### 5.2.1 Changes in overlap PCR conditions.

Several attempts were made to modify the conditions for cloning ASGPR in pc3.1 vector. The number of overlap PCR cycles was dropped to 1 from the initial 7 cycles. Also 25 cycles were performed to amplify the PCR product further. In an attempt to decrease the formation of secondary structure during amplification 1%, 6% and 10% DMSO (dimethylsulfoxide) concentrations were added in the PCR reaction. DMSO in the reaction helps to avoid secondary structures forming within a primer or the ssDNA template (Winship 1989) and 10% DMSO may also lower the melting temperature by 5-6°C. Alternatively PCR was also performed using an enhancer (Molzym) containing Betaine (N, N, N-trimethylglycine) (Henke et al. 1997) which also decreases secondary structure formation without decreasing *Taq* polymerase activity. However, mutations occurred consistently in all the positive clones sequenced. Hence it was attempted to correct the point mutations. The clone (*Clone f*) with the minimal mutations was chosen to be corrected by site directed mutagenesis.

### 5.2.2 Changes by site directed mutagenesis.

PCR site-directed method allows site-specific mutations to be incorporated in a double-stranded plasmid. Site directed mutagenesis was performed according to manufacturers' protocol (Stratagene). The PCR reaction was set to 2 cycles and two temperatures 60°C and 65°C were used for annealing. *Pfu* DNA polymerase was used in the PCR reaction followed by a DpnI reaction.

The process of site directed mutagenesis successfully corrected the two point mutations, but in addition introduced another point mutation elsewhere in the plasmid (See Appendix 4F for sequence data). It was recognized that the mutations were concentrated in the GC rich regions. The high degree of GC content made this amplification and cloning extremely difficult. Hence, it was decided to have the ASGPR sequence synthesized commercially.

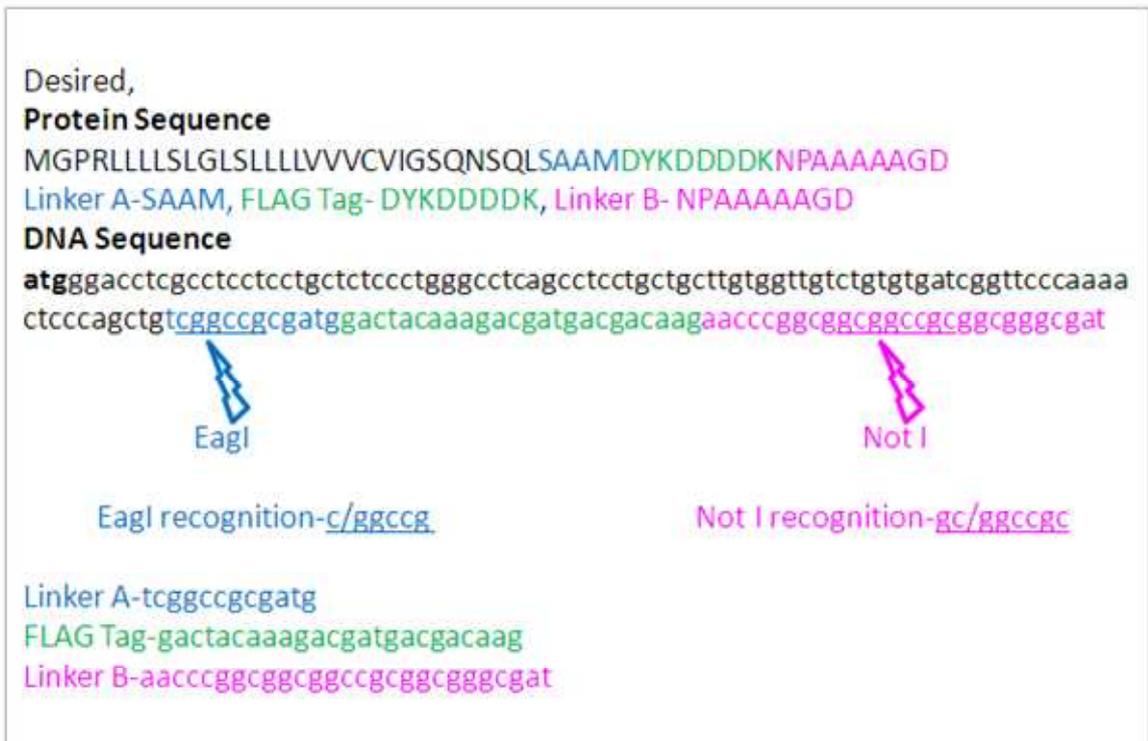
### **5.3 Designing the ASGPR-pc3.1 surface expression plasmid (pSE2) for commercial synthesis.**

The construction of functional fusion proteins often requires a linker sequence that adopts an extended conformation to allow for maximal flexibility. Linker sequences are generally selected based on intuition. Without a reliable selection criterion, the design of such linkers is often difficult, particularly in situations where longer linker sequences are required (Craeto, Feng 2000).

It was decided to include the transmembrane domain of ASGPR with a FLAG epitope tag linked to each other by the linker A. This linker A would enable the FLAG tag to be expressed on the surface such that it is not tethered too close to the membrane of the cells and hence enable detection. The FLAG epitope tag is followed by another linker B which would allow the cloning and expression of protein of interest such that it is not too close to the membrane and is presented on the surface of the cells available for binding to the receptor (Figure 5.7).

In addition to the favourable orientation ability of the linker it is useful to design it so as to include restriction enzyme sites which could allow easy manipulation of the construct and thus easy cloning of DNA of interest. It should be noted that the

restriction enzyme sites included in the linkers should be rare restriction enzymes which do not cut into the DNA of interest or the ASGPR transmembrane anchor sequence. Accordingly, restriction enzyme site Eag I was included into linker A and Not I restriction site was included in linker B.



**Figure 5.7: The sequence of asialoglycoprotein receptor (ASGPR) transmembrane domain designed for commercial synthesis into the pc3.1 mammalian expression vector (pSE2).** In addition to the ASGPR transmembrane domain anchor sequence, the synthesized plasmid included a Linker A which contains the Eag I restriction site (blue), followed by a FLAG epitope tag (green) and a Linker B which contains the Not I restriction site (pink).

The linkers were designed using the designing tool called LINKER (Crauto, Feng 2000, Xue, Gu & Feng 2004). The sequences suggested by the program are those that are likely to adopt an extended conformation in the engineered fusion protein (Crauto, Feng 2000). To include Eag I recognition sequence 5' c/ggccg 3' was translated using Expassy translational tool. The translated frames were obtained (Figure 5.8).

```
5'3' Frame 1: cggccg
                R  P
5'3' Frame 2: cggccg
                G
5'3' Frame 3: cggccg
                A
Now, acg = T (threonine), tcg = S (serine), ccg = P (proline), gcg = A (alanine)
And codon starting with g _ _ could be Valine, Alanine, Aspartate, Glutamate or
Glycine.
```

**Figure 5.8: Incorporation of EagI restriction enzyme site in linker A of the surface display vector (pSE2).** The DNA sequence of the EagI recognition site was translated. The obtained translated frames indicate the residues which should be incorporated into Linker A to incorporate EagI restriction site.

Inclusion of arginine(R) residues in the linker sequence is not favourable since it is a charged residue which could interfere with orientating the protein of interest. Hence it was required to use translated Frame 3 for designing a linker which should start with a T (threonine) or S (serine) or P (proline) or A (alanine).

This could be followed by an alanine (A) residue which could be followed by a Valine (V) or Alanine (A) or Aspartate (D) or Glutamate (E) or Glycine (G) and coded by a DNA sequence cggccg. This DNA sequence cggccg includes the Eag I restriction enzyme site (Figure 5.8).

The Linker database was set to obtain linkers (SAAM and SAGQ) with 4 residues, avoiding all protease cut sites. The hydrophobicity plot indicated SAAM sequence to be more hydrophilic (than SAGQ) and hence was likely to be more flexible

and extend itself in the aqueous environment. On the basis of the hydrophobicity plot SAAM sequence was chosen to be included as the Linker A.

To include Not I recognition sequence (in the second linker) 5'gc/ggccgc 3' was translated using Expassy translational tool. The following translated frames were obtained (Figure 5.9):

5'3' Frame 1: gcggccgc A  A
5'3' Frame 2: gcggccgc R  P
5'3' Frame 3: gcggccgc G  R
Now gct, gca, gcg, gcc = A (alanine)

**Figure 5.9: Incorporation of NotI restriction enzyme site in linker B of the surface display vector.** The DNA sequence of the NotI recognition site was translated. The obtained translated frames indicate the residues which should be incorporated into Linker B to incorporate NotI restriction site.

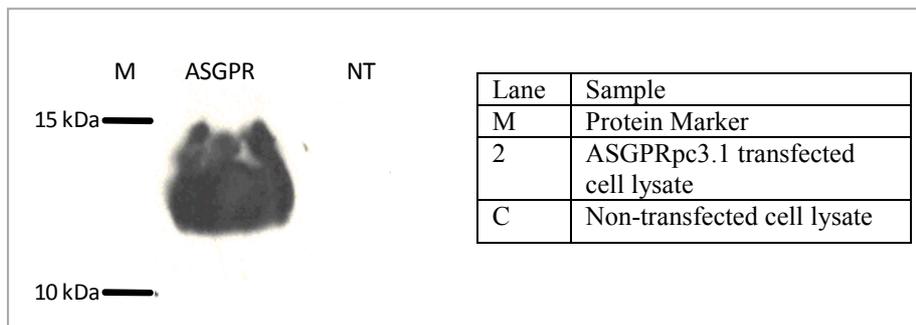
Inclusion of arginine (R) residues in the linker sequence is not favourable since it is a charged residue which interferes with orientating the protein of interest. Hence it was required to use translated Frame 1 for designing a linker which should include three, consecutive alanine (A) residues, coded by a DNA sequence gcggccgc. This DNA sequence gcggccgc includes the Not I restriction enzyme site (Figure 5.9).

The Linker database was set to obtain linkers with 9 residues, avoiding all protease cut sites. AAADGDDSL and NPAAAAAGD linker sequences were obtained containing, three consecutive alanine (A) residues. On the basis of the hydrophobicity plot NPAAAAAGD sequence was chosen to be included as the Linker B.

Accordingly the ASGPR sequence was synthesized commercially to include the linkers A and B interrupted by a FLAG sequence tag and the vector obtained was called (pSE2). See appendix 3 for vector map.

#### 5.4 Expression of surface display vector ASGPR-pc3.1 (pSE2).

To confirm cellular expression of the pSE2 plasmid it was transfected in CHO cells and the cell lysate analyzed by Western blotting (Figure 5.10). The FLAG-tagged ASGPR protein was detected by probing the nitrocellulose membrane with anti-FLAG primary and anti-mouse HRP linked secondary antibody as described in section 2.19.1. The expected size of the FLAG-tagged ASGPR anchor domain protein is ~6 kDa.

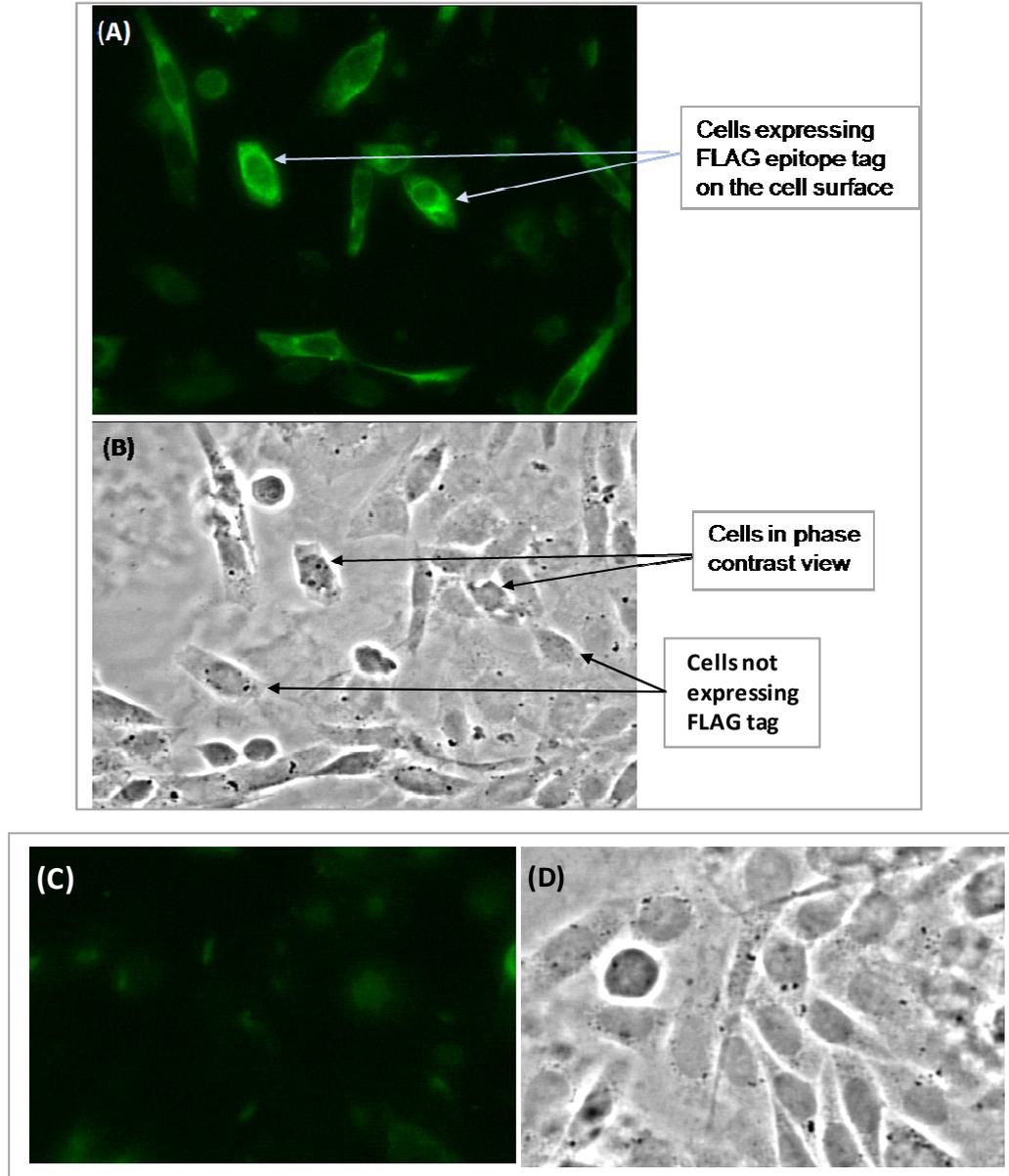


**Figure 5.10: Expression of the surface display vector pSE2.** CHO cells were transfected with ASGPR surface display vector pSE2. Cell lysate was analyzed 24 hours post-transfection. The expected size for ASGPR is 6 kDa. Band between 10 to 15 kDa approximately 12 kDa is observed in lane 2. No band is observed in the non-transfected cell lysate (NT) lane.

The molecular mass of the expressed ASGPR protein is more than the expected size. This may be due to glycosylation of the protein and inaccuracy in detection of low molecular weight proteins on SDS-PAGE.

In order to utilise the vector for surface display pSE2 it is necessary to confirm that the expressed protein is localized to the extracellular face of the plasma membrane. To test this, the ASGPR vector was transfected into CHO cells. 24 hrs post transfected

cells were fixed and probed for the presence of FLAG epitope tag by immunofluorescence staining as described in section 2.20.1. These cells transfected with the pSE2 vector were incubated with monoclonal anti-FLAG primary antibody and this was detected using fluorescently labelled anti-mouse Cy2 antibody. Cells were visualised at 40X magnification under a fluorescence microscope. Transfected cells exhibit clear membrane staining. As shown in Figure 5.11(A) a proportion of cells transfected with ASGPR showed strong cell surface expression of the epitope tag. It is important to note that the cells were not permeabilized before probing for the FLAG epitope tag, this was to ensure that only epitope tags expressed on the extracellular surface were detected. Figure 5.11C shows non-transfected CHO cells labelled with anti-FLAG and Cy2 antibodies. Figure 5.11D is a phase contrast view of the field shown in 5.11C. These non-transfected cells serve as negative control in this experiment and none of the cells show any staining with the fluorophore indicating that the staining seen in figure 5.11A is specific for the cell surface expressed FLAG tag.



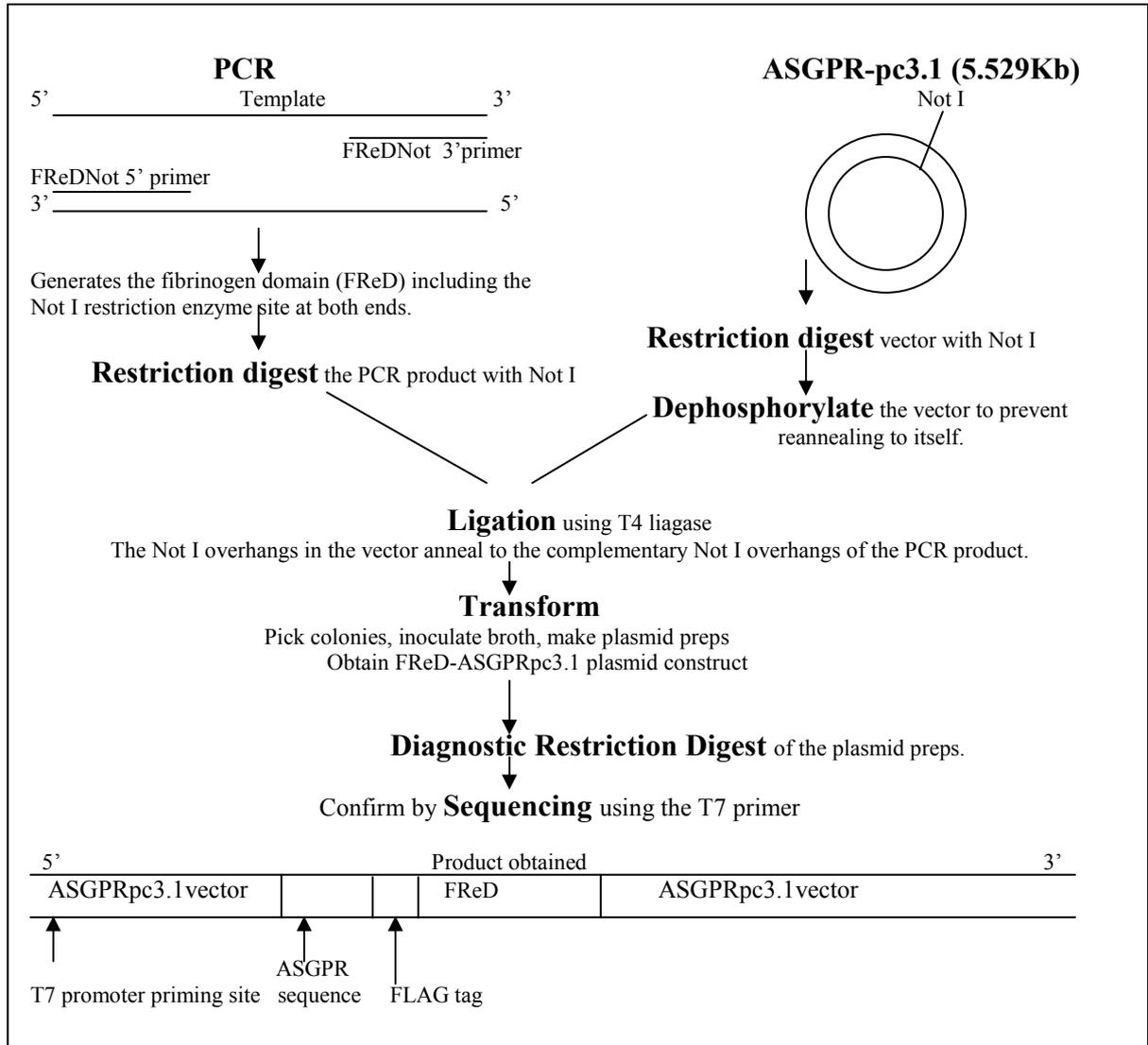
**Figure 5.11: Localization of expressed ASGPR on CHO cell surface by immunofluorescence staining.** CHO cells were transfected with the pSE2 construct. The FLAG-tagged ASGPR protein was expected to be expressed on the cell membrane in the transfected cells. The protein localisation was confirmed by immunohistochemistry. (A) The cells expressing FLAG-tagged ASGPR protein were incubated with mouse monoclonal anti-FLAG primary antibody. The anti-FLAG antibody bound to the FLAG tagged ASGPR protein on the cell surface. The cells were then fluorescently labelled with anti-mouse Cy2 antibody which bound to the primary antibody. The green fluorescing cells were visualised at 40X magnification under a fluorescence microscope. Cell surface protein expression was confirmed by observed green fluorescence on the cell membrane of the transfected cells. (B) Phase contrast of the field shown in figure 5.11A, showing all the cells in the field view under white light. (C) Non-transfected CHO cells served as a negative control in this experiment and were stained with anti-FLAG primary and secondary Cy2 fluorophore. (D) Phase contrast of the field shown in figure 5.11 C.

Thus the FLAG tagged ASGPR was localized to the extracellular face of the cell membrane as detected by immunofluorescence staining.

Having confirmed the expression of ASGPR domain cloned in the pSE2 surface display plasmid the next step was to clone Ang1 and Ang2 truncated fragments; FReD (fibrinogen related domain) into this vector, express the ligand at the cell surface and analyze the ability of surface expressed FReD domains to bind soluble Tie2 receptor ectodomain.

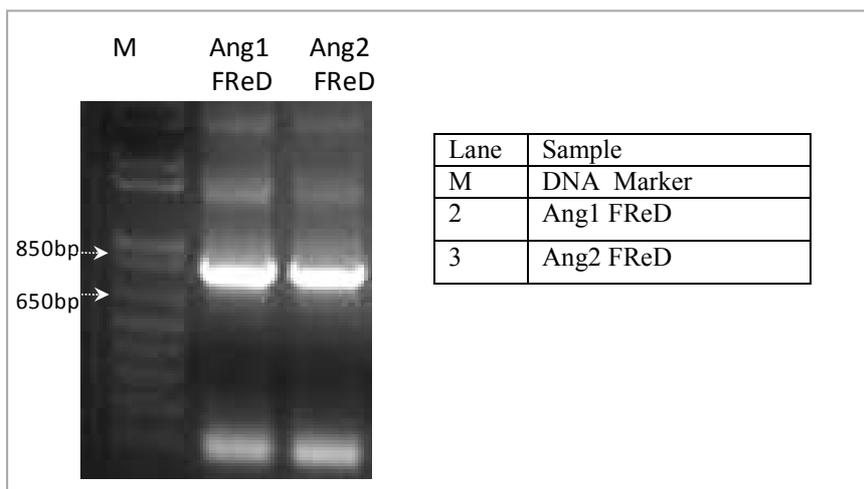
### 5.5 Cloning of FLAG tagged FReD (fibrinogen related domain) in the pSE2 surface display vector.

The truncated fragment of Ang1 and Ang2, the fibrinogen domain was cloned by the strategy outlined in Figure 5.12.



**Figure 5.12: Schematic representation for cloning FReD (fibrinogen domain) in the pSE2 surface display expression vector.** The FReD domains (Ang1 and Ang2) were amplified by PCR using their respective FReDNot 5' and FReDNot 3'primers. The amplified product was restricted with NotI and ligated into the pSE2 plasmid. Competent cells were transformed with the ligation product and plasmid was prepared from the colonies on the transformation plate. Plasmid was analyzed by diagnostic PCR to screen for the presence of the desired insert. Sequence data was obtained for the positive clones.

The fibrinogen domain was amplified from template (Ang1 and Ang2 plasmid available in the lab) by PCR using the FReDNot 5' and FReDNot 3'primers for Ang1 and Ang2 at annealing temperature 45°C. The expected size for the Ang-FReD domain is 655bp. The size of the amplified product was confirmed by running 10µl of the product on a 1% agarose gel (Figure 5.13) with a band obtained between 650-850bp.



**Figure 5.13: Amplification of the Ang1 and Ang2 FReD domains for cloning into pSE2 vector.** Amplification is performed using FReDNot 5' and FReDNot 3'primers for Ang1 and Ang2. The expected size of FReD DNA is 655bp. The required amplified product FReD in lane 2 and 3 is obtained between 650- 850bp.

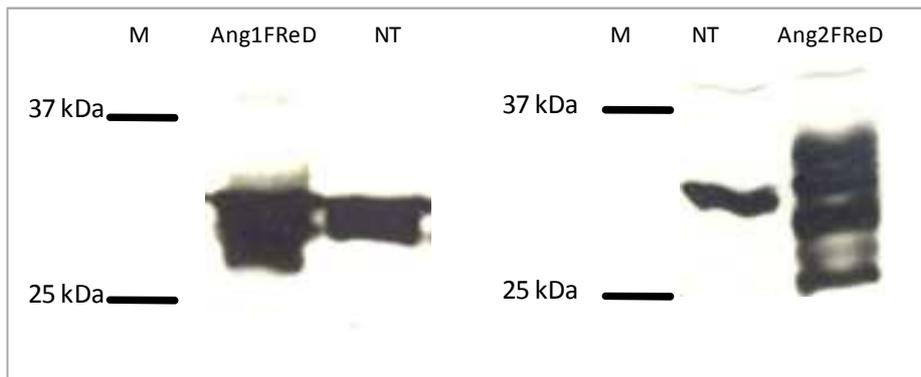
The amplified product was gel extracted and purified. The PCR product and the vector was digested with Not I enzyme purified and ligated. 5µl of the ligation reaction was used to transform competent cells. 25 colonies were picked from the transformation plate and plasmid was prepared. Diagnostic restriction digestion was performed using Not I enzyme to detect insert.

Positive clones were sequenced using the T7 primers. Ang1 FReD plasmid *FAr* and Ang2-FReD plasmid *F8* with the correct sequence were obtained (See Appendix 4G and 4H for sequence data).

### 5.6 Expression of cell bound FReD (Ang1 and Ang2) domains in CHO cells.

The construct obtained in Section 5.5 was analyzed to confirm expression of the ASGPR linked-FReD domains in mammalian cells.

CHO cells were transfected with plasmids *FAr* (Ang1-FReD) and *F8* (Ang2-FReD). 24 hrs post transfection cell lysate was prepared and analyzed by Western blot. Briefly, FLAG-tagged WT FReD protein was detected by probing the nitrocellulose membrane with anti-FLAG primary and anti-mouse HRP linked secondary antibody as described in section 2.19.1. The expected size of the ASGPR linked FLAG-tagged WT Ang-FReD domain protein is ~29.5 kDa.



**Figure 5.14: Expression of the ASGPR-FReD domains in mammalian (CHO) cells.** CHO cells were transfected with the *FAr* (Ang1-FReD) and *F8* (Ang2-FReD) plasmids. Western blot was performed to analyze cell lysate. The FLAG tagged proteins were detected using anti-FLAG antibody and HRP linked anti-mouse. Size of cell bound FReD domain is expected to be ~29.5kDa. Bands between 25 and 37 kDa are observed in Ang1FReD and Ang2FReD lanes. Different bands indicate various forms of the expressed proteins. The non transfected lane showed a non-specific band.

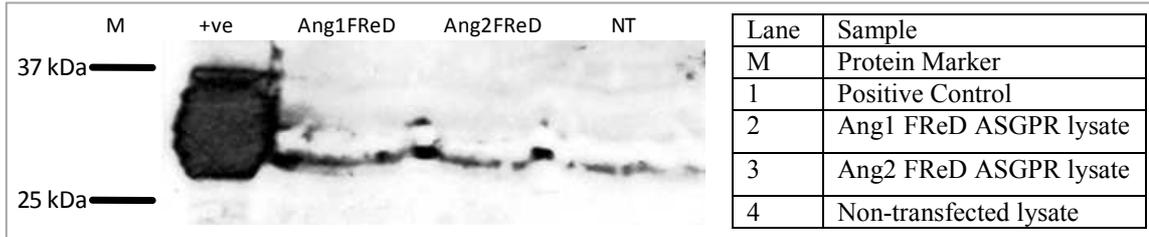
The non transfected lanes showed a non-specific band at approximately 27kDa (Figure 5.14). Ang1FReD showed a band approximately between 27-30kDa. Ang2 showed bands approximately between 27-35kDa (Figure 5.14). Thus expression of both Ang1 and Ang2 FReD domains are observed. The multiple bands for *FAr* (Ang1-FReD) and *F8* (Ang2-FReD) proteins may indicate presence of degraded, cleaved,

differentially or intermediately glycosylated and cell surface expressed forms of the protein.

### **5.7 Expression of cell bound FReD (Ang1 and Ang2) domains in Ramos cells.**

On establishing the expression of the ASGPR linked-FReD domain constructs *FAr* (Ang1-FReD) and *F8* (Ang2-FReD) in CHO cells, the next step was to express this construct in a B cell line capable of mutating the angiopoietin FReD domains using its somatic hypermutation machinery. The Burkitt's lymphoma Ramos cell line is known to hypermutate transgenes (Wang et al. 2004). Ramos cells were therefore transfected (as described in section 2.15.2) with the Ang1 FReD- ASGPR-pc3.1 and Ang2 FReD- ASGPR-pc3.1 constructs to confirm expression in this cell line.

The cell lysate from the transfected cells was analysed by Western blotting (Figure 5.15). The FLAG-tagged ASGPR-FReD protein was detected by probing the nitrocellulose membrane with anti-FLAG primary and anti-mouse HRP linked secondary antibody as described in section 2.19.1. The expected size of the ASGPR linked FLAG-tagged WT Ang2-FReD domain protein is ~29.5 kDa.

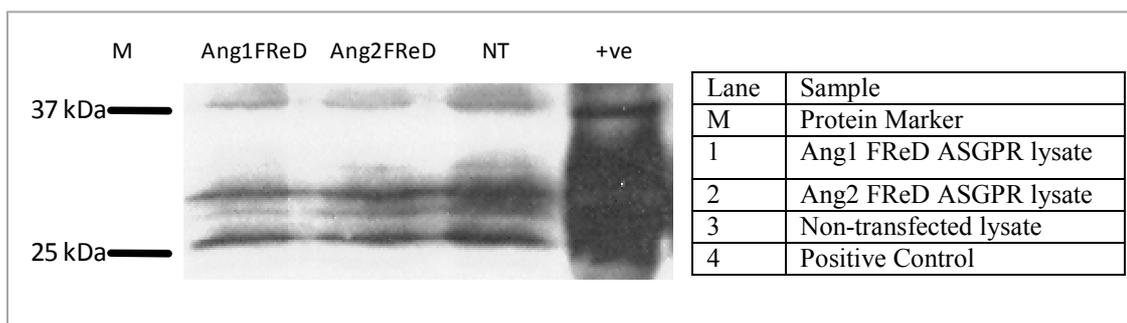


**Figure 5.15: Expression of the ASGPR lined FReD domains in transiently transfected Ramos cells.** Ramos cells were transfected with constructs (*FAr*, *F8*) encoding ASGPR linked FReD domains (Ang1 and Ang2). The cell lysate was analysed by Western blotting by detecting FLAG-tagged proteins with anti-FLAG primary and anti-mouse HRP linked secondary antibody. The expected size for ASGPR-FReD protein product is ~29.5 kDa. The positive control is the cell lysate from CHO cells transfected with the *F8* construct and it shows bands between 25 and 37kDa. No bands observed in Ramos transfected cells, lanes 2 and 3.

As seen in figure 5.15, a specific band of the correct size was not observed in lanes 2 and 3 for Ang1 or Ang2 FReD when compared with the negative control (non transfected cell lysate) in lane 4. The positive control (lane 1) is the cell lysate from CHO cells transfected with the *F8* construct and it shows bands between 25 and 37kDa.

The inability to detect the protein expression in transiently transfected Ramos cells could be due to inability to transfect the Ramos cell line. Hence stable cell lines were made by applying G418 selection pressure for 2 weeks (as described in section 2.16.1). Cells transfected with the *FAr* and *F8* constructs harbouring the antibiotic resistance gene in the vector backbone grew in the presence of G418 whereas the non-transfected Ramos cells did not survive in G418.

The stable cell line lysate was analyzed again by Western blotting and probed for the FLAG epitope tag (Figure 5.16).



**Figure 5.16: Expression of the ASGPR lined FReD domains in stably transfected Ramos cells.** Cell lysate was analysed after establishing stable cell lines with constructs (*FAr*, *F8*) encoding the ASGPR linked Ang1 or Ang2 FReD domains. The expected size for ASGPR FReD protein product is ~29.5 kDa. No band of the correct size is observed in lanes 1 and 2 for Ang1 or Ang2 FReD when compared with the negative control (non transfected cell lysate) in lane 3. The positive control is the cell lysate from CHO cells transfected with the *F8* construct and it shows bands between 25 and 37kDa.

As seen in figure 5.16, a specific band of the correct size was not observed in lanes 1 and 2 for Ang1 or Ang2 FReD when compared with the negative control (non transfected cell lysate) in lane 3. The positive control (lane 4) is the cell lysate from CHO cells transfected with the ASGPR-FReD construct and it shows bands between 25 and 37kDa.

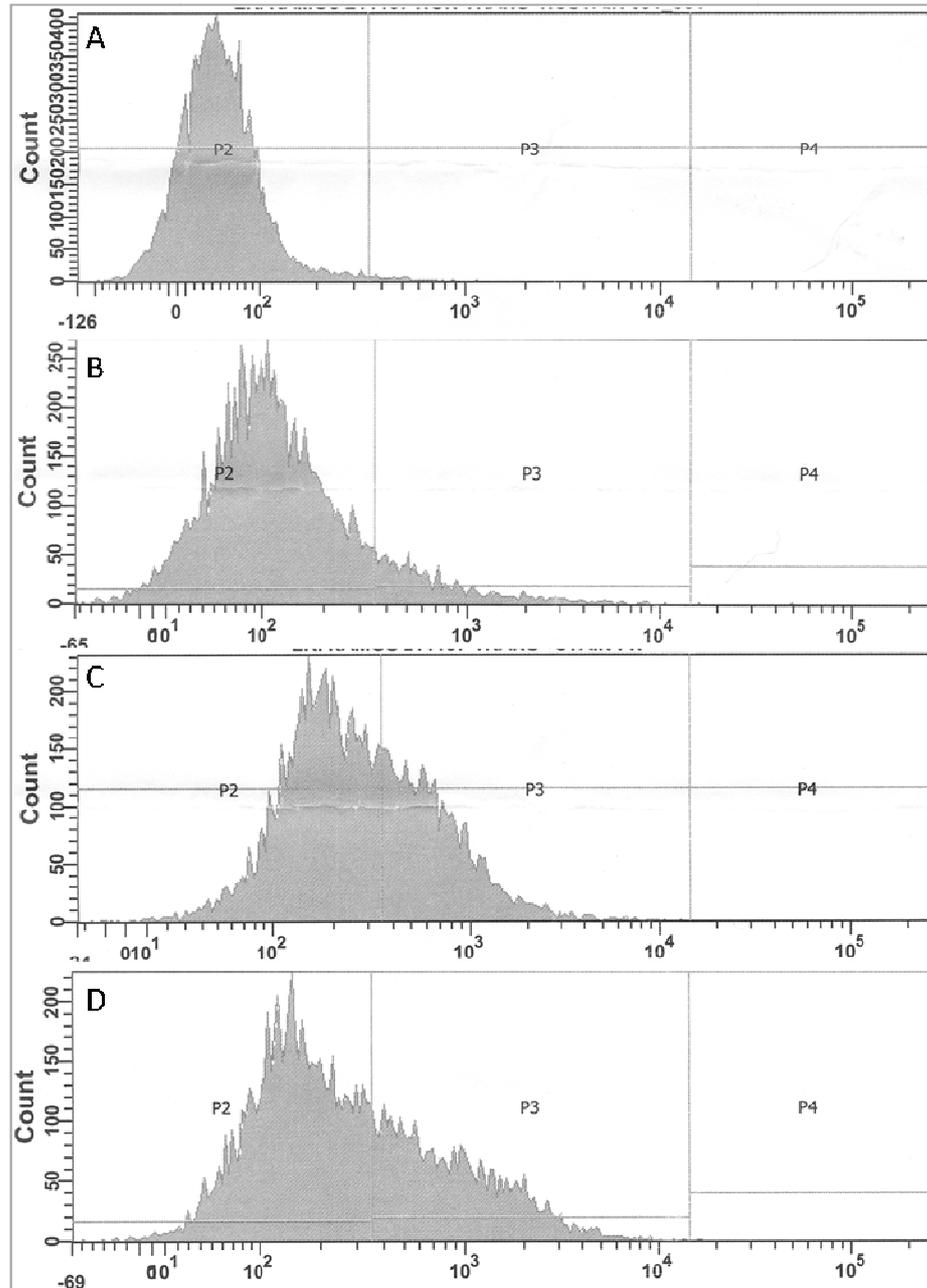
Thus, the protein of interest could not be detected in neither the transiently nor stably transfected Ramos cells by Western blotting. Hence a more sensitive method of flow cytometry analysis was sought for detection of the protein of interest.

### **5.8 Flow cytometry analysis of Ramos cells transfected with ASGPR linked-FReD domain constructs *FAr* (Ang1) and *F8* (Ang2).**

Ramos cells, transiently transfected with Ang1 and Ang2 FReD-ASGPR constructs (*FAr*, *F8*) were stained as described in section 2.22.1 and then analysed on a flow cytometer. Briefly the cells incubated with mouse monoclonal anti-FLAG followed by anti-mouse Cy2 fluorophore. The fluorescent intensity corresponding to FLAG tag was detected on the flow cytometer.

Figure 5.17A is the flow cytometry profile of non-transfected unstained cells and majority of the cells lie in P2 section of the profile. Figure 5.17B is the flow cytometry profile of non-transfected stained with the Cy2 fluorophore. Most of the cells in are obtained in the P2 section of the profile representing unstained population of non-transfected cells. Some cells obtained in the P3 section (Figure 5.17B) of the flow cytometry profile represent non-transfected Ramos cells displaying a background fluorescent signal from the fluorophore.

Figure 5.17C is the flow cytometry profile of Ang1-FReD transfected Ramos cells stained with the Cy2 fluorophore. Figure 5.17D is the flow cytometry profile of Ang2-FReD transfected Ramos cells stained with the Cy2 fluorophore. Both Ang1 and Ang2 transfected Ramos cells show a population emerge in the P3 section of their flow cytometry profiles when compared with the controls in Figure 5.17A and B. Also since the cells stained for the flow cytometry analysis have not been permeablized during the staining process, it also confirms that the protein of interest is localized on the cell surface.

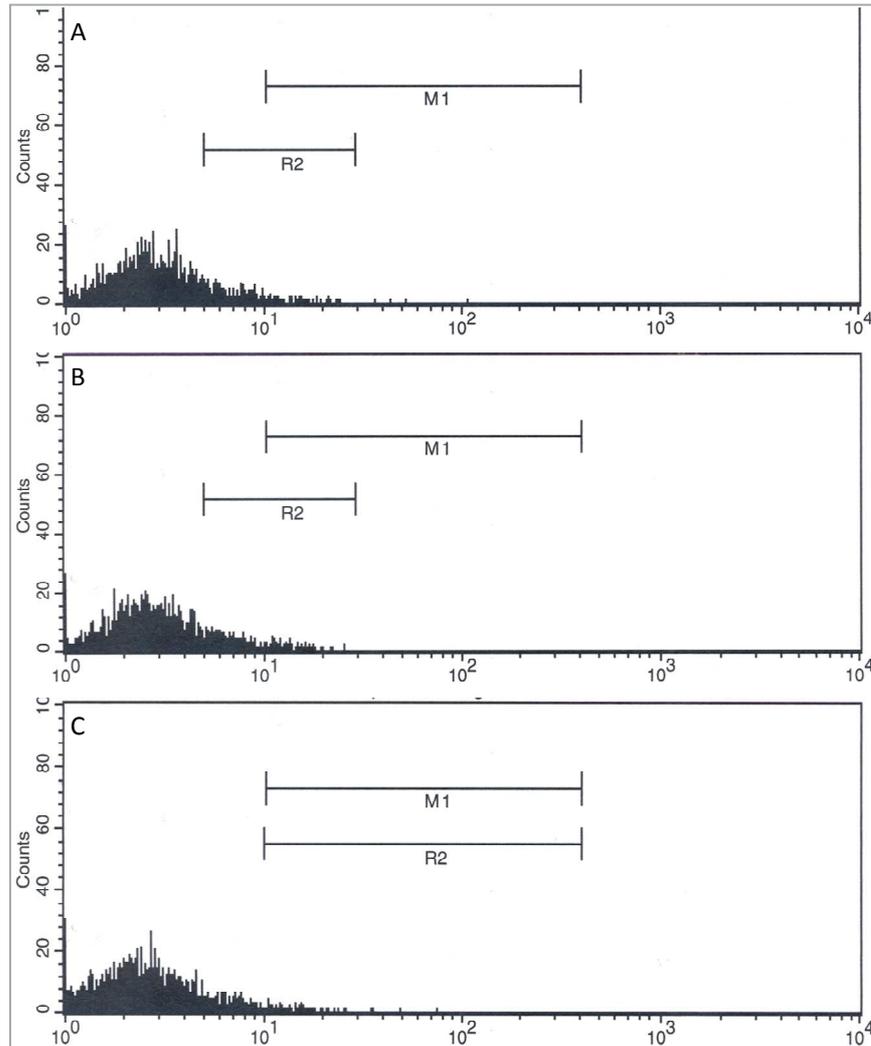


**Figure 5.17: Flow cytometry profiles of Ramos cells transiently transfected with Ang1 and Ang2 FReD-ASGPR constructs (*FAr*, *F8*).** Ramos cells transiently transfected with Ang1 and Ang2 FReD-ASGPR constructs were fluorescently labelled using Cy2 fluorophore and analysed on a flow cytometer. (A) Flow cytometry profile of non- transfected Ramos cells which were not stained. Majority of the cells lie in P2 section of the profile. (B) Flow cytometry profile of non- transfected Ramos cells which were stained in parallel to the transiently transfected cells. Most of the cells in are obtained in the P2 section of the profile representing unstained population of non-transfected cells. (C) Flow cytometry profile of Ramos cells transiently transfected with Ang1 FReD ASGPR construct. (D) Flow cytometry profile of Ramos cells transiently transfected with Ang2 FReD ASGPR construct.

Thus flow cytometry analysis of the transiently transfected Ramos cells confirms the expression and localization of the FReD-ASGPR protein on the cell surface.

The flow cytometry profile of Ramos cells stably transfected with Ang1 and Ang2 FReD-ASGPR constructs was also determined. Stable cells cultured in the G418 antibiotic were stained as described in section 2.22.1 and then analysed on a flow cytometer.

Figure 5.18A is the flow cytometry profile of non-transfected cells stained with the Cy2 fluorophore. This profile represents the unstained population of non-transfected cells. Figure 5.18B and 5.18C are the flow cytometry profiles of Ang1-FReD and Ang2-FReD transfected stable Ramos cells stained with the Cy2 fluorophore. Both Ang1 and Ang2 transfected stable Ramos cells show a flow cytometry profile similar to that observed for the non-transfected cell population in figure 5.18A. Thus, there was found to be no difference between the flow cytometry profiles of the stably transfected cells and the non-transfected cells.

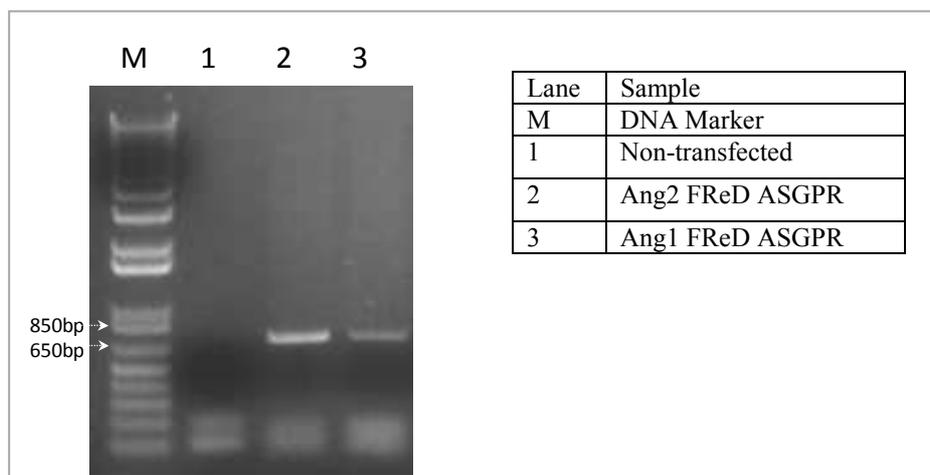


**Figure 5.18: Flow cytometry profiles of Ramos cells stably transfected with Ang1 and Ang2 FReD-ASGPR constructs (*FAr*, *F8*).** Stable Ramos cells transfected with Ang1 and Ang2 FReD-ASGPR constructs were fluorescently labelled using Cy2 fluorophore and analysed on a flow cytometer. (A) Flow cytometry profile of non-transfected Ramos cells which were stained in parallel to the transiently transfected cells. Most of the cells represent unstained population of non-transfected cells. (B) Flow cytometry profile of stable Ramos cells transfected with Ang1-FReD ASGPR construct. (C) Flow cytometry profile of stable Ramos cells transfected with Ang2-FReD ASGPR construct.

Flow cytometry analysis of the transfected Ramos cells was able to detect the protein expression in transiently transfected cells but not in stable cell line. Hence it was decided to check the Ramos stable cells for mRNA expression of the Ang1 and Ang2 FReD-ASGPR proteins.

### 5.9 Expression of ASGPR linked FReD (Ang1 and Ang2) mRNA in Ramos stable cell lines.

The stable Ramos cells were analysed to determine whether they were transcribing the FReD proteins. Messenger RNA was extracted from the Ramos stable cell lines and reverse transcription PCR was performed. The cDNA thus obtained was amplified by PCR. 15µl of the PCR product was analyzed on an agarose gel (as described in section 2.3).



**Figure 5.19: Detection of ASGPR linked FReD (Ang1 and Ang2) mRNA in Ramos stable cells.** mRNA was extracted from stable Ramos cell line (transfected with *FAr* and *F8* constructs). Extracted mRNA was used as template for RT PCR reaction and the cDNA thus obtained was amplified by PCR. The product was checked on an agarose gel. The expected size for ASGPR FReD protein product is ~800bp. Band of the correct size is observed in lanes 2 and 3 for Ang2 and Ang1 FReD-ASGPR stable cells. The negative control (non transfected mRNA) in lane 1 shows no band.

The expected size of the PCR product is 800bp. Figure 5.19 shows the expression of protein mRNA for both Ang2 and Ang1 FReD-ASGPR protein in lanes 2 and 3 respectively. The negative control (non transfected mRNA) in lane 1 shows no band.

The data indicated Ang2 and Ang1 FReD were being transcribed (Figure 5.19) but the protein was not expressed on the stable Ramos transfectants (Figure 5.18). A

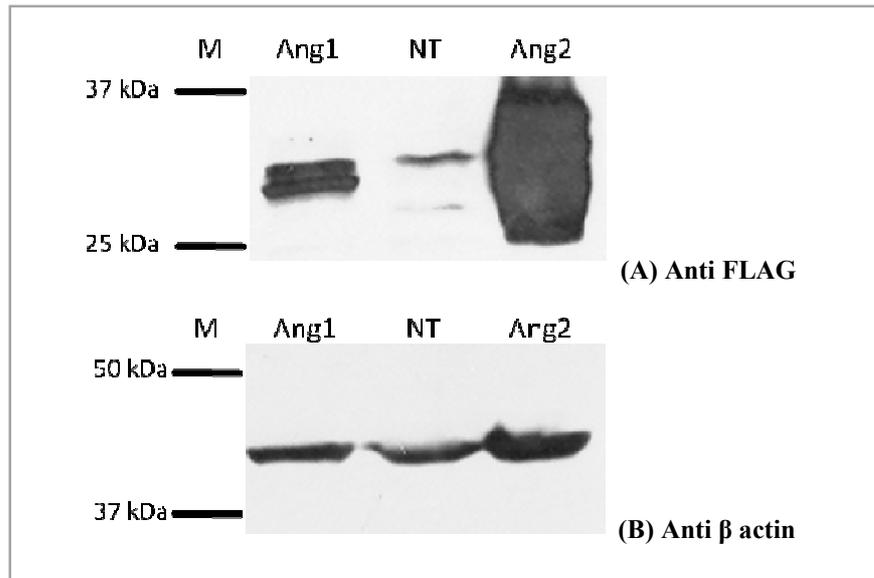
new transfection of Ramos cells was performed and analysed by flow cytometry. Unlike the data obtained in section 5.8, figure 5.17 these transient transfectants were unable to show distinct expression and localization of the FReD-ASGPR protein.

To further understand the discrepancy and inconsistency in expression of protein in the Ramos transient and stable cells, HEK cells were transfected with the plasmid constructs to check if a similar trend was observed in a different mammalian cell line.

#### **5.10 Western analysis of HEK cells transfected with the FReD- ASGPR-pc3.1 constructs.**

HEK cells were transfected with the Ang1 and Ang2 FReD- ASGPR-pc3.1 constructs. The transfectants were analysed by Western blotting as described in sections 2.17.3 and 2.19.1 detecting the FLAG tag on the expressed proteins.

Figure 5.20A shows a Western blot of the cell lysate obtained from the transient expression of Ang1 and Ang2 FReD-ASGPR constructs (*FAr*, *F8*) in HEK cells. Multiple bands are observed for the Ang2-FReD domain (figure 5.20A lane 3) but only one specific band is observed for Ang1-FReD. It was observed that the expression of the Ang1-FReD construct is much lesser than the expression of the Ang2-FReD construct. Differential loading in the wells could be accounted for this drastic difference between the two constructs. To rule out this possibility the blot was reprobated for the  $\beta$ -actin house-keeping gene (Figure 5.20B).



**Figure 5.20: Expression of the ASGPR-FReD domains in transiently transfected HEK293 cells.** (A) Cell lysate was analysed post-transfection to detect the FLAG-tagged Ang1 and Ang2 FReD ASGPR protein. The expected size for ASGPR FReD protein was ~29.5 kDa. Bands between 25 and 37 kDa are observed in lanes 1 and 3; however expression of Ang2 FReD is much more than Ang1 FReD. Also multiple bands are observed for the Ang2-FReD domain but only one specific band is observed for Ang1-FReD. (B) Expression of the  $\beta$  actin house keeping gene in HEK's transiently transfected with Ang1 and Ang2 FReD- ASGPR-pc3.1 constructs. Blot (A) was reprobbed for the  $\beta$ -actin housekeeping gene to confirm equal loading in all wells. A band of the expected size was observed in all lanes between 37-50kDa.

Although Ang2-FReD (lane 3) shows more loading as compared to Ang1FReD lane 1, it cannot account for the large difference in transient expression observed between the Ang1 and Ang2 FReD- ASGPR-pc3.1 constructs.

The reprobe for the housekeeping gene shows that differential loading cannot be the only reason for the low detection of expressed Ang1-FReD protein compared to the Ang2-FReD protein. It is possible that Ang1 is expressed poorly in the cells or the translated protein is unstable as compared to the Ang2-FReD-ASGPR protein.

### **5.11 Immunofluorescence analysis of HEK cells transfected with the ASGPR linked FReD domain (Ang1 and Ang2) constructs.**

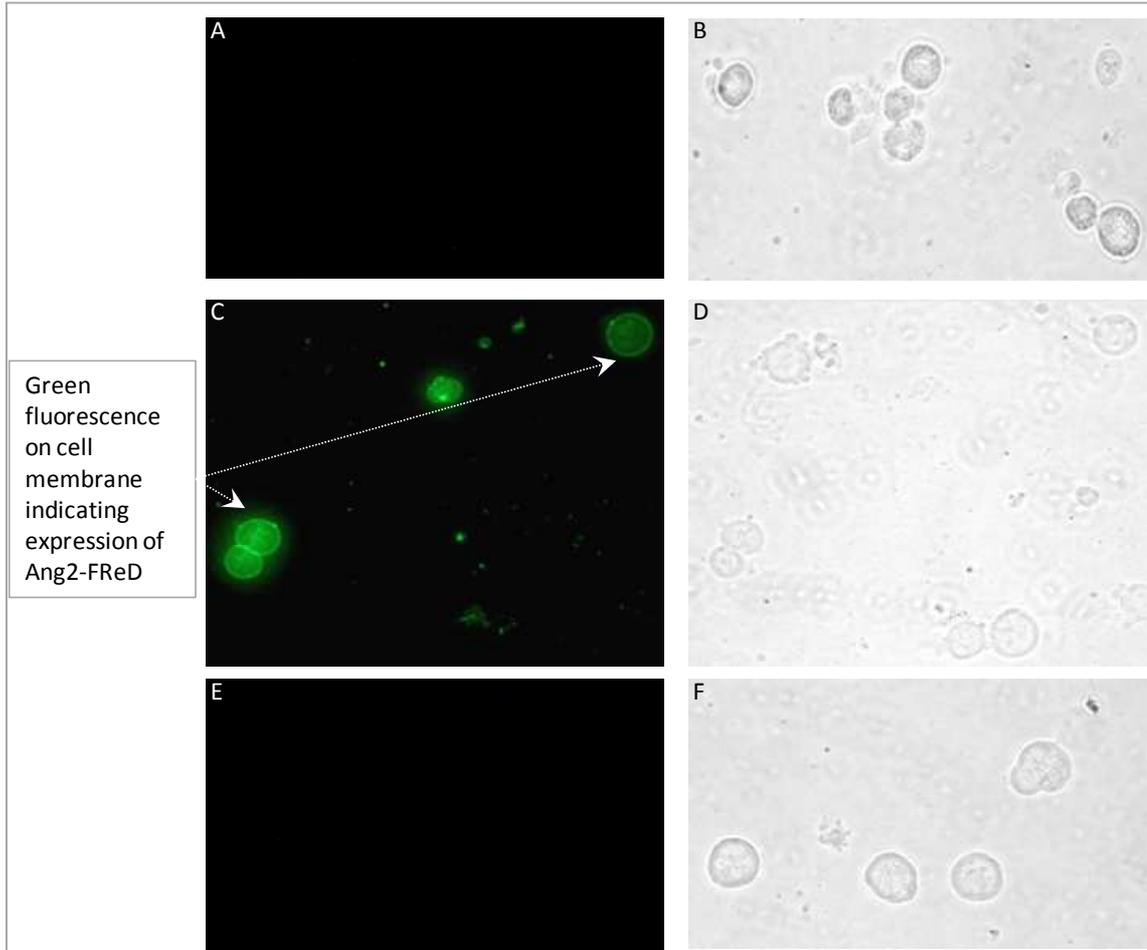
Protein localisation was verified by immunofluorescence staining of HEK cells transfected (as described in section 2.20.2) with the *FAr* and *F8* constructs.

HEK cells transfected with the FReD-ASGPR constructs (*FAr*, *F8*). Stable transfectant cells were generated by applying G418 antibiotic selection pressure for 2 weeks as described in section 2.16.1. Transfected cells were stained for FLAG epitope tag in suspension. Briefly, cells expressing FLAG-tagged ASGPR-FReD domain protein was incubated with monoclonal anti-FLAG primary antibody and this were detected using fluorescently labelled anti-mouse Cy2 antibody. Cells were visualised at 40X magnification under a fluorescence microscope.

Figure 5.21A and C shows stable HEK cells transfected with ASGPR-FReD (Ang1 and Ang2) construct (obtained in Section 5.5) and immunostained for FLAG tag. Ang1-FReD protein (Figure 5.21A) does not exhibit membrane staining when compared to the non-transfected control (Figure 5.21E).

Ang2-FReD protein expressed exhibit clear membrane staining (Figure 5.21C). It should be noted that the fluorescence on the surface of the cells varies. This may be due to the variable expression level of the protein in each cell. The fluorescence variation may also be due to differences in the amount of protein expressed on the cell surface and the rate of internalisation of the protein for each cell which in turn may depend on site of integration of the FReD DNA in the HEK genome. Figure 5.21B and D are the phase contrast views of the same field as 5.21 A and C. The cells in figure 5.21C shows almost all cells are positively stained for FLAG tagged Ang2-FReD.

Figure 5.21E shows non-transfected HEK cells immunostained for FLAG tag and 5.21F is the phase contrast view for the non-transfected cells. These non-transfected cells serve as negative control in this experiment and none of the cells show any staining with the fluorophore indicating that the staining seen in figure 5.21C is specific for the cell surface expressed Ang2-FReD domain.



**Figure 5.21: Immunofluorescent localization of cell bound ASGPR-FReD domains on stable HEK cell surface.** HEK cells were transfected with the *FAr* and *F8* constructs. The FLAG-tagged ASGPR-FReD domain protein was expected to be expressed on the cell membrane anchored by the ASGPR transmembrane domain. The protein localisation was confirmed by immunohistochemistry. The HEK cells expressing FLAG-tagged ASGPR-FReD domain protein (Ang1 and Ang2) were incubated with mouse monoclonal anti-FLAG primary antibody. The cells were then fluorescently labelled with anti-mouse Cy2 antibody which bound to the primary antibody. The cells were visualised at 40X magnification under a fluorescence microscope. (A) No cell surface protein expression was observed on the cell membrane for ASGPR-Ang1FReD protein. (B) Phase contrast of the field shown in figure 5.21A, showing all the cells in the field view under white light. (C) Cell surface protein expression of ASGPR-Ang2FReD protein was confirmed by observed green fluorescence on the cell membrane. Arrows indicate the fluorescence on the cell surface where the FLAG epitope tag is expressed. Variation in expression (fluorescence) was observed for the stable cells which may be due to differential integration of the DNA into the genome. (D) Phase contrast of the field shown in figure 5.21C, showing all the cells in the field view under white light. (E) Non-transfected HEK cells served as a negative control in this experiment and were stained with anti-FLAG primary and secondary Cy2 fluorophore. Non-transfected cells did not show any fluorescence, indicating that the staining of the stable cells in 5.21C was specific for the surface expressed Ang2-FReD domain. (F) Phase contrast of the field shown in figure 5.21E.

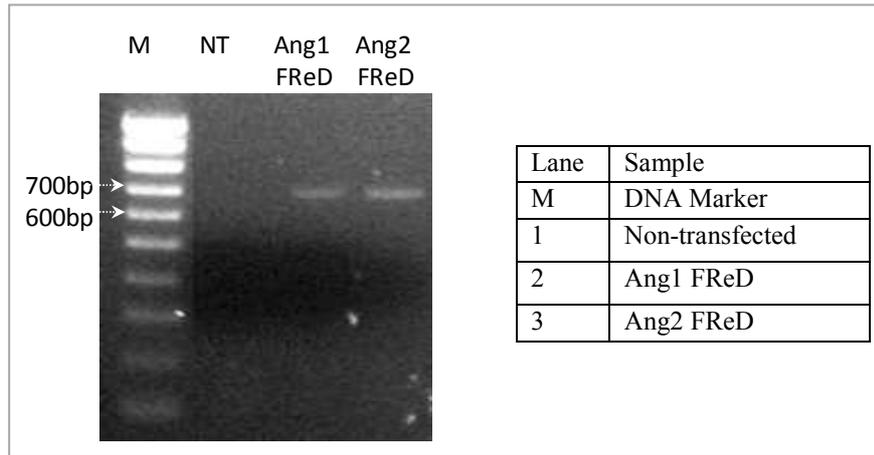
Thus only the ASGPR anchored FReD (Ang2) domain was localized to the extracellular face of the plasma membrane as detected by immunofluorescence staining in stably transfected HEK cells.

No distinct detectable expression of Ang1 FReD was observed in the HEK cell line by immunostaining. Several variations in the immunofluorescence staining protocol were performed. Incubation temperatures of 4°C, RT and 37°C were used, incubation periods were varied in several permutations and combinations, and concentrations of antibodies were varied to optimize the staining. However, expression of Ang1 FReD could not be localised to the extracellular cell surface.

The stable FReD ASGPR HEK cells were also analyzed for the presence of the mRNA for the protein.

#### **5.12 Detection of FReD (Ang1 and Ang2) mRNA in HEK stable cell lines.**

HEK cells were transfected with construct encoding the ASGPR anchor with the Ang1 or Ang2 FReD domain and stable cell line generated using G418 selection. RNA was extracted from the HEK stable cell lines and RT PCR was performed. The cDNA thus obtained was amplified by PCR using 5' and 3' FReD primers. The product was analyzed on an agarose gel (Figure 5.22). The expected size of the Ang-FReD domain is ~655bp. A band between 600-700bp was observed in lanes 2 and 3 for both Ang1 and Ang2 FReD.



**Figure 5.22: To detect the expression of the FReD-ASGPR (Ang1 and Ang2) mRNA in stable HEK cell line.** HEK cells were transfected with construct encoding the ASGPR anchor with the Ang1 or Ang2 FReD domain and stable cell line generated using G418 selection. RNA from these stable cells was extracted and used as template in RT PCR reaction. The cDNA thus obtained was amplified by PCR. The product was checked on an agarose gel. The expected size for Ang FReD product is 655bp. Band between 600-700bp was observed in lanes 2 and 3 when compared with the negative control (non transfected mRNA) in lane 1.

Messenger RNA for both the Ang-1 and Ang-2 FReD domains were detected in the stably transfected HEK cells.

Expression of the ASGPR-Ang1FReD fusion protein was detectable at the transcription level in HEK 293 and Ramos cells (Figure 5.19 and 5.22). Ang1 FReD domain expression was barely detectable by Western blot of CHO and HEK 293 transfected cell lysates (Figure 5.14 and 5.20) and undetectable by immunofluorescence in HEK293 cells (Figure 5.21). Expression of the ASGPR-Ang2FReD fusion protein was detected in CHO and HEK cells by Western blotting (Figure 5.14 and 5.20) and immunofluorescence (Figure 5.21); its expression in the Ramos stable cell line was undetectable (Figure 5.18).

## Discussion

The aim of this chapter was to establish a cell surface display system for expression and orientation of the Ang1 and Ang2 FReD domains on the extracellular surface of Ramos B cell line. This surface display system could be utilised for analysis of the expressed protein by flow cytometry. The Ramos B cell line was employed to support SHM driven hypermutation for directed evolution of the FReD domains.

There were three problems encountered in achieving these aims. Firstly, the cloning of the ASGPR transmembrane domain was extremely difficult, due to the high GC content (64%) in the DNA sequence. Secondary structure formations lead to inaccurate amplification and sequencing. Several efforts to enhance PCR conditions, using high fidelity polymerases, change cycling conditions and adding reagents to prevent secondary structure formation were unable to obtain correctly sequenced plasmid pSE1. It is possible that the required correct sequence may have been amplified by the PCR conditions, but the sequencing for it was consistently inaccurate. Hence, the surface display vector (pSE2) incorporating linker sequences and restriction enzyme sites was designed. This vector; pSE2 was synthesized commercially (Geneart). As observed by immunofluorescence, the transmembrane domain of the ASGPR protein was able to insert in the cell membrane. The linker in the native angiopoietins consists of 27 amino acids and a similar number of residues were desired in the designed anchor. Thus the final linker containing linker A, the FLAG tag and linker B was of 21 amino acids and were predicted to adopt extended conformation. This allowed the FLAG-epitope tag linked to the ASGPR to be detected by immunofluorescent staining.

A second problem encountered was that, the expression of the fusion protein in the Ramos stable cell line was undetectable by Western blotting and immunofluorescence. Although the expression of the Ang1 and Ang2 FReD fusion protein were detected in transient transfected Ramos cells by flow cytometry, this result was not reproducible. The transfection efficiency of the Ramos cells was optimized; however, the expression was low and variable.

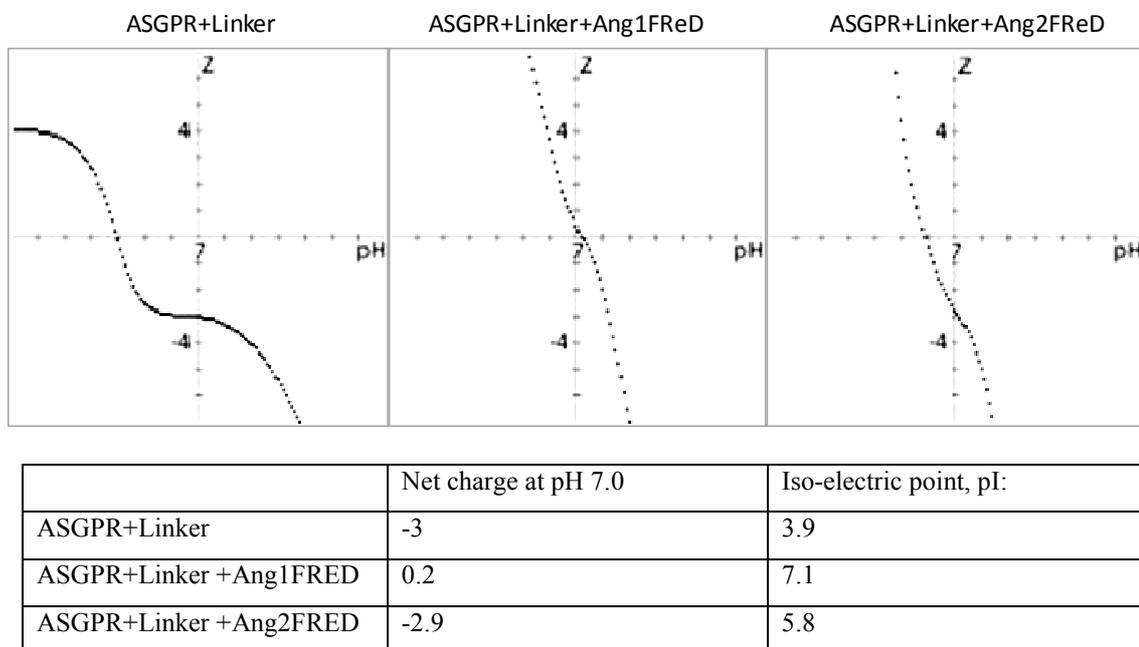
Stable transfections in Ramos cells were performed to obtain positive clones by applying G418 selection pressure. However, the antibiotic resistant cells were not expressing the protein of interest. Since the integration in the cell line was random, it is possible that integration at a poorly transcribed chromatic region lead to undetectable expression. It is also possible that the cell line retained the antibiotic resistant gene by integration into the genome but discarded the ASGPR-FReD DNA sequence. Although the CMV promoter is known to work with high efficiency in a broad range of host cells, including B cell lines, some studies indicate that efficiency of the CMV promoter in some B cell lines may be poor and a EF-1 $\alpha$  promoter or UbC promoter may be more desirable (Zarrin et al. 1999) (<http://tools.invitrogen.com>).

A third problem was the expression of Ang1 Fred was much lower and inconsistent in all cell lines (CHO, HEK293, Ramos) in which expression was analysed. Compared to the Ang1 fusion protein, the Ang2 fusion protein expressed well in HEK 293 and CHO cells, indicating that, the Ang1 expression problem was intrinsic to the Ang1 protein. The ASGPR linked Ang1 FReD protein was detected at the mRNA level in both HEK293 and Ramos cells, indicating that the transcription of the protein occurred normally, under the CMV promoter.

The expression of this fusion protein depends on its folding stability, post-translational modification, translocation and integration into the cell membrane. The expressed protein should be resistant to cellular proteases and lack proteolytic cleavage sites. When the surface display vector pSE2 was designed, it was ensured that the fusion protein is resistant to proteases. Since the expression of Ang2FReD fusion protein cloned into the same vector as the Ang1FReD is observed, it is unlikely that the expression of ASGPR-Ang1 FReD protein is hindered by proteolytic cleavage at least of the regions outside the Ang1 FReD.

Proper folding of the fusion protein into a 3-D structure depends on its interaction with the fused amino acid sequence, its overall net charge, stability at the physiological pH and interaction with the cellular environment. It is possible that the interaction of the Ang1 FReD domain with the ASGPR protein sequence or the linker sequence leads to improper folding. The protein net charge depends on the number, identity and location of amino acids and the solvent pH. An algorithm based program used for calculating net charge revealed that both the ASGPR protein and the ASGPR-Ang2FReD fusion protein were predicted to have a net charge of approximately -3 at pH 7 (Figure 5.23). This charge may allow both proteins to interact with the aqueous cellular environment on the cellular surface and hence be detected by immunofluorescence. Although the amino acid sequence of the Ang1FReD is similar to Ang2FReD the calculated net charge of the ASGPR-Ang1FReD fusion protein was predicted to be neutral at pH 7 (Figure 5.23). This may influence interactions with the cellular environment and possibly cause improper folding. Improperly folding in turn, may result in the cellular machinery to degrade the Ang1 fusion protein. The low molecular weight Ang1FReD

fusion protein detected in Western blotting of transfected CHO cells may be this degraded intracellular form of the protein.



**Figure 5.23: Prediction of net charge and pI values for the ASGPR anchor and the ASGPR-Ang1FReD and ASGPR-Ang2FReD fusion proteins.** The input amino acid sequence is used to calculate a predicted value of net charge at pH 7.0 and the iso-electric point of the proteins. (<http://www.innovagen.se/custom-peptide-synthesis/peptide-property-calculator/peptide-property-calculator.asp>) (Compute pI/Mw: [http://expasy.org/tools/pi\\_tool.html](http://expasy.org/tools/pi_tool.html))

Another possibility is that the low molecular mass band observed for Ang1FReD fusion protein on a Western blot represents protein forms not processed accurately. The band may represent an unprocessed polypeptide or intermediately modified forms of the protein. Again inappropriate processing of a protein sequence could be accounted to amino acid interactions between the ASGPR and FReD sequence of the fusion protein. In some cases the processing of the protein sequence is influenced by the host cell processing machinery which may be appropriate for one protein but not for another (Hossler, Khattak & Li 2009, Durocher, Butler 2009).

It is also possible that the Ang1 fusion protein was expressed, folded and modified accurately; however, translocation of the protein was hindered. It has been shown that the regions flanking the apolar signal peptide sequence may contribute to the orientating of the translocon (Beltzer et al. 1991, Goder, Bieri & Spiess 1999). The ASGPR transmembrane domain, a type II membrane protein is naturally known to contain a signal-anchor sequence which is responsible for both insertion and anchoring, to orient the free C-terminus on the extracellular surface of the cell. In mammalian cells, protein synthesis is coupled to translocation and membrane insertion. The topogenesis of the fusion protein may be influenced by effect of charges in the flanking residues, protein folding and glycosylation of an individual protein sequence. It has been demonstrated that the more positive flanking sequence is generally cytoplasmic and changing the charge difference between the regions flanking the hydrophobic core can change the orientation of insertion (Goder, Spiess 2001). It is possible that the interaction of the Ang1FReD domain with the ASGPR sequence is different as compared to the Ang2FReD domain, leading to inaccurate orientation or inability of insertion, ultimately leading to no detection by immunofluorescence.

The expression of ASGPR-Ang2FReD fusion protein was consistent in both CHO and HEK cell line and the protein was localized by immunofluorescence on the extracellular surface of the cell. Multiple low and high molecular bands possibly representing modified and unmodified forms of the Ang2FReD protein was detected between 25-37 kDa by Western blotting. However, the expression of ASGPR-Ang2FReD fusion protein in the Ramos cell line was undetectable. These problems of poor transfection, uncontrolled random integration and inability to generate stable protein expression in Ramos cells made this cell line unsuitable for establishing the

SHM mutation system to evolve the FReD domain. Use of SHM to generate a mutant library requires stable transfections in which the DNA of interest is incorporated into the genome of the hypermutating B cell line. Poor and variable expression of the protein of interest makes the Ramos cells unsuitable for proceeding with mutagenesis and makes detection of improved mutants difficult. Since SHM occurs in any AID expressing B-cell line, an alternative cell line could be used to overcome the expression and detection problem encountered in the Ramos cells. Due to the problems encountered in expression of active forms of Ang1-FReD further studies in the following chapters were continued with Ang2-FReD.

## **Chapter 6: Expression of Angiopoietin FReD domain on DT40 Cell Surface for Directed Evolution by Somatic Hypermutation and Iterative Selection by Fluorescence Activated Cell Sorting**

### **Introduction**

In Chapter 5, efforts were made to perform somatic hypermutation (SHM) mediated mutagenesis in the Ramos cell line. However, low expression levels of the FReD protein were observed in the Ramos cells. Studies with hypermutating cell lines have suggested the role of transcription in the mutation mechanism. Transcription allows the exposure of the single-stranded DNA on which Activation-Induced Deaminase (AID) acts and causes point mutations (Bachl et al. 2001, Ramiro et al. 2003). Since it has been suggested that transcription enhances SHM mediated mutagenesis, low expression levels in Ramos cells would have posed a problem resulting in sub-optimal SHM of the protein of interest.

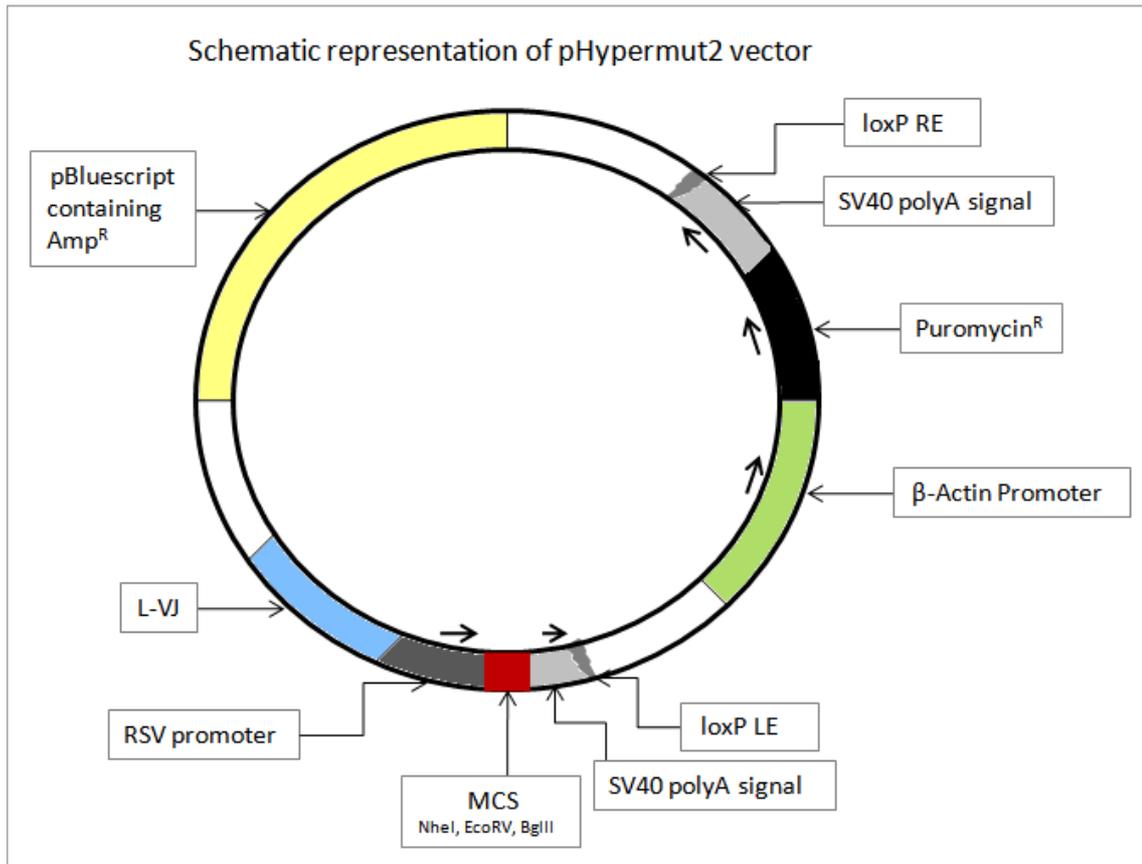
Although genome wide SHM has been reported of non-Ig genes (Wang, Harper & Wabl 2004) studies suggest that the Ig loci supports a higher rate of SHM. The Ig locus is influenced by cis-acting sequences that allow the chromatin structure to expose the DNA sequence more readily to the AID mutator (Martin, Scharff 2002). In Ramos cells, transgenes usually integrate at random chromosomal sites outside the hypermutating loci and this may result in lower mutagenesis in the transgene of interest.

To overcome the problems encountered with the Ramos cells, a clone of the bursal B cell line DT40 was chosen to further the study in SHM mediated mutagenesis of the FReD domains. These cells are suitable to support SHM and they were engineered for higher expression of the AID gene as compared to the wild-type DT40

cells (Arakawa, Hauschild & Buerstedde 2002). This engineered cell line expresses the Cre recombinase as a MerCreMer fusion protein which is inactive in the absence of an estrogen derivative like hydroxytamoxifen. Activation of the Cre recombinase by addition of hydroxytamoxifen to the cells, allows excision of the AID gene when required. The excision of the AID gene by background activity of Cre recombinase in prolonged culture can be kept low by addition of 0.5µg/ml of mycophenolic acid (Arakawa et al. 2008, Zhang et al. 1996). The DT40 cells have a doubling time of 10 hours, which is less than the Ramos cell doubling time of 24 hours, reducing the experimental time scale (Arakawa et al. 2008, Sale, Neuberger 1998).

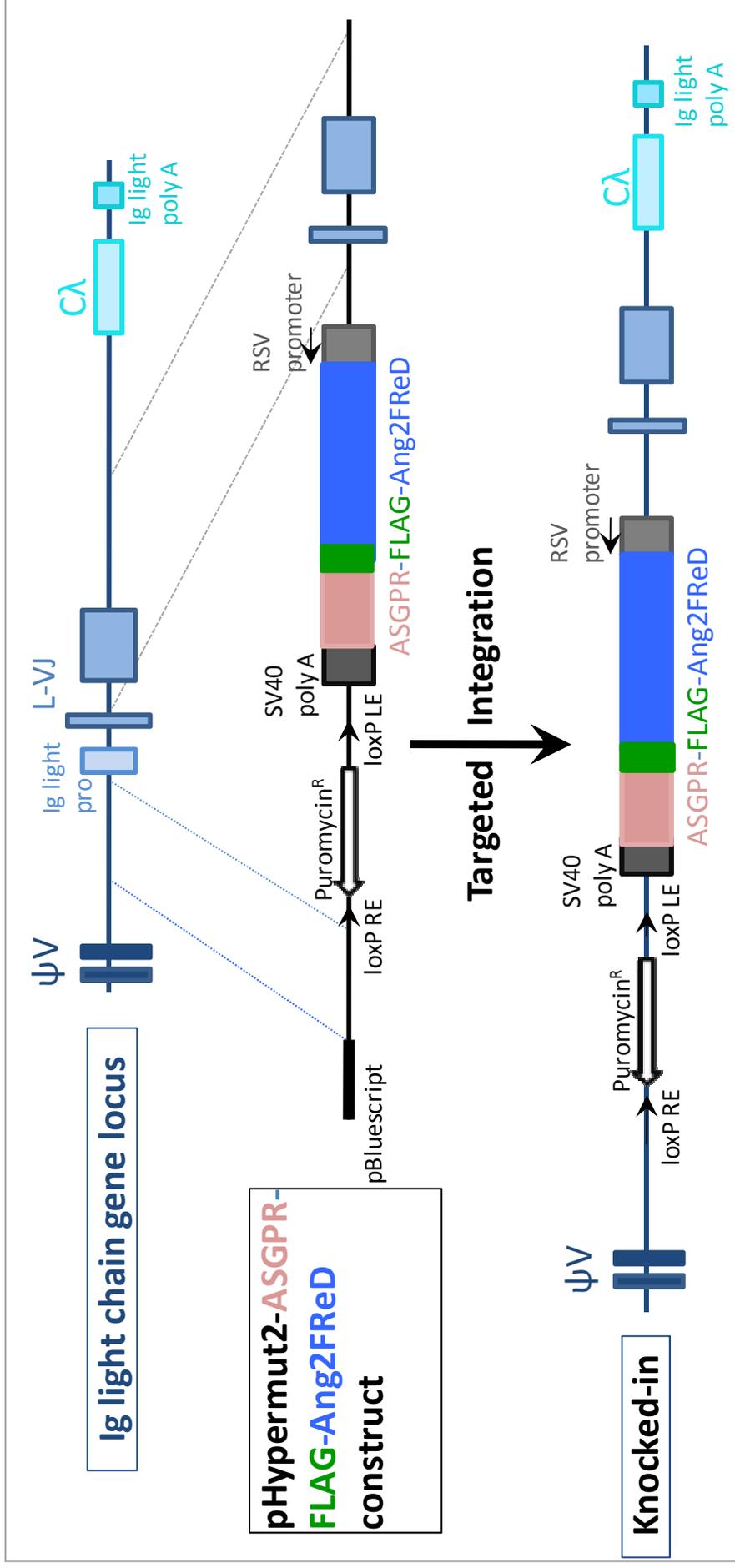
To express the FReD domain in this engineered DT40 cell line a suitable vector pHypermut2 was used (This vector together with the DT40 cell line was kindly supplied by Dr. Jean-Marie Buerstedde). The vector contains a portion of the pBluescript vector which harbours the ampicillin resistance gene ( $Amp^R$ ), the Rous Sarcoma Virus (RSV) promoter to allow high level constitutive expression of the protein of interest and a multiple cloning site (MCS) for cloning of the transgene (desired insert) into the vector. The pHypermut2 vector also included puromycin antibiotic resistance gene ( $puro^R$  cassette) downstream of the transgene for selection of stable transfected cells.

Figure 6.1 outlines key features of the pHypermut2 vector used for cloning of the Ang1-FReD and Ang2-FReD domains.



**Figure 6.1: Schematic map of pHypermut2 vector.** Key features include Amp<sup>R</sup> gene (from pBluescript vector), L-VJ recombination sequence for targeted integration into the DT40 genome, RSV promoter for constitutive expression, Multiple cloning site (MCS) for cloning the Ang2FReD domain, puromycin antibiotic resistance gene for selection of transfected (stable) cells.

Cloned into the vector is the L-VJ recombination sequence for targeted integration of the desired insert (Ang-FReD domain coding sequence) into the rearranged Ig light chain locus in the DT40 genome by recombination as outlined in Figure 6.2. Integration at this locus maximizes the effect of SHM driven mutations.



**Figure 6.2: Physical map of the rearranged Ig light chain locus, the pHypermut2-Ang2FReD construct and the rearranged locus after targeted integration of the Ang2-FReD domain coding sequence into the DT40 Ig light chain locus.** The homologous recombination occurs between Ig light chain gene locus and the similar DNA sequence encoded by the pHypermut2-Ang2FReD construct. The targeted integration results in knocking-in of the FLAG tagged ASGPR-Ang2FReD sequence into the Ig light chain gene locus. [Figure modified from (Arakawa et al. 2008)].

This chapter aims to improve the binding affinity of Ang1 fibrinogen related domain (Ang1-FReD) and Ang2 fibrinogen related domain (Ang2-FReD) to the Tie2 ectodomain by introducing mutations into the domains driven by somatic hypermutation (SHM). The SHM induced mutation mechanism was established in the chicken B-cell line DT40. The FLAG-tagged, ASGPR (asialoglycoprotein receptor transmembrane domain) anchored domains were cloned into pHypermut2 vector. The vector targets the incorporation of the receptor binding domain (RBD) into the DT40 genome to allow for SHM induced mutation of the domain sequence. The screening for affinity improvement was carried out using Fluorescence Activated Cell Sorting (FACS) to sort cells expressing high affinity mutated proteins on the DT40 cell surface. Table 6.1 shows summary of the constructs cloned in this chapter.

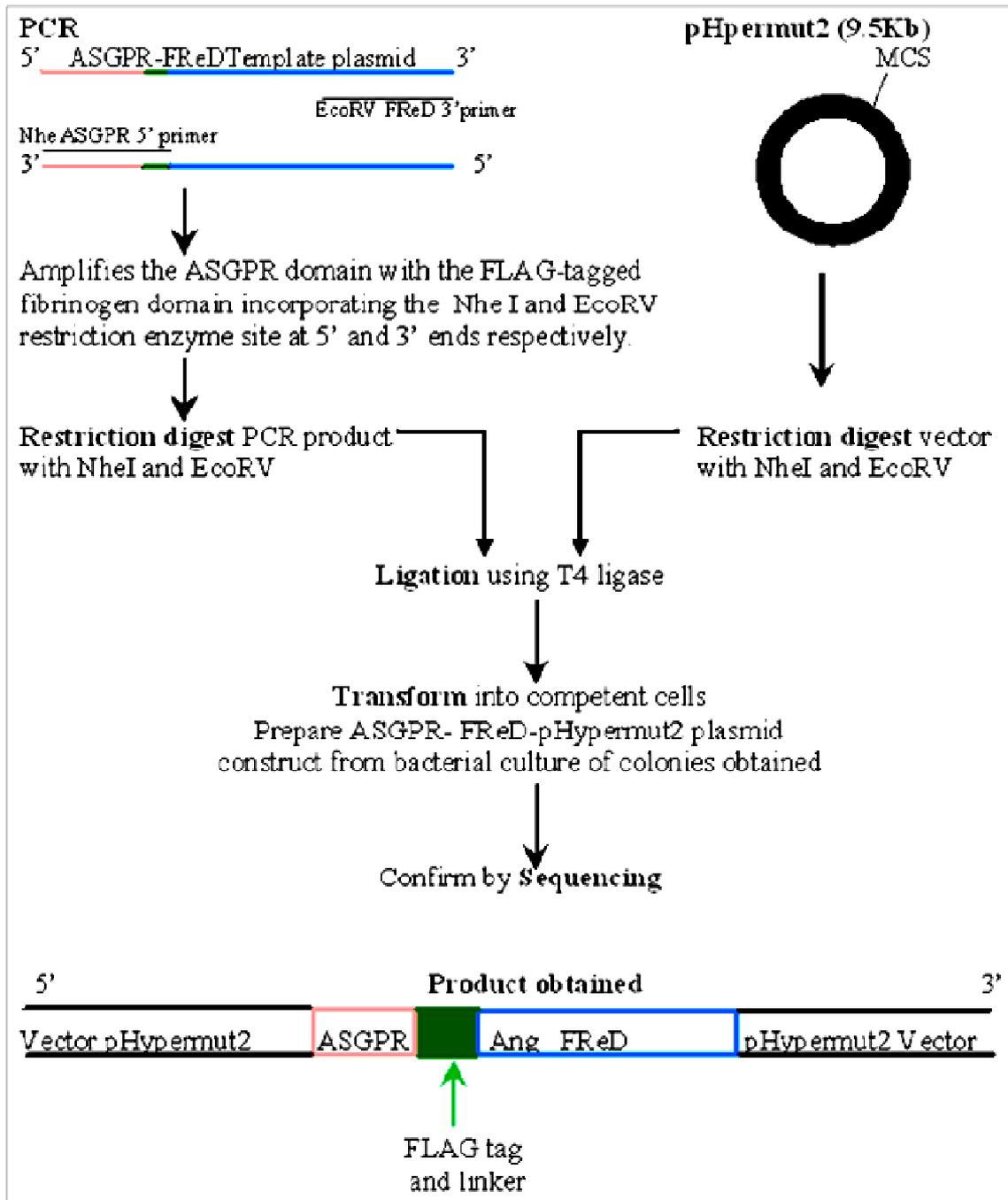
<b>Construct Name</b>	<b>Protein encoded for by the construct</b>	<b>Tag</b>	<b>Features</b>	<b>Used for</b>
<i>FArHe</i>	Ang1 Fibrinogen domain (Ang-1 FReD) Residues 282-498	N-terminus, ASGPR anchor protein and FLAG epitope tag	ASGPR anchored FLAG-tagged Ang1 FReD protein displayed on the cell surface.	Binding to Tie2 receptor ectodomain and selected for
<i>F8Hd</i>	Ang2 Fibrinogen domain (Ang-2 FReD) Residues 280-496		ASGPR anchored FLAG-tagged Ang2 FReD protein displayed on the cell surface.	improved affinity by FACS of DT40 cells

**Table 6.1: List of constructs used in Chapter 6.** *FArHe* and *F8Hd* constructs were obtained by cloning FLAG-tagged ASGPR linked Ang1 FReD domain and FLAG-tagged ASGPR linked Ang2 FReD domain into the pHypermut2 vector. This vector was then used for expressing FLAG-tagged ASGPR linked Ang1 FReD domain and FLAG-tagged ASGPR linked Ang2 FReD domain proteins on the DT40 cell surface.

## **6.1 Cloning of FLAG tagged ASGPR linked Ang1 and Ang2 FReD (fibrinogen related domain) in the pHyperm2 vector and expression in HEK293 cells.**

The truncated fragment of Ang1 and Ang2, namely the fibrinogen domain, linked in frame with the ASGPR surface anchor protein and FLAG tag was cloned into the pHyperm2 vector by the strategy outlined in figure 6.3.

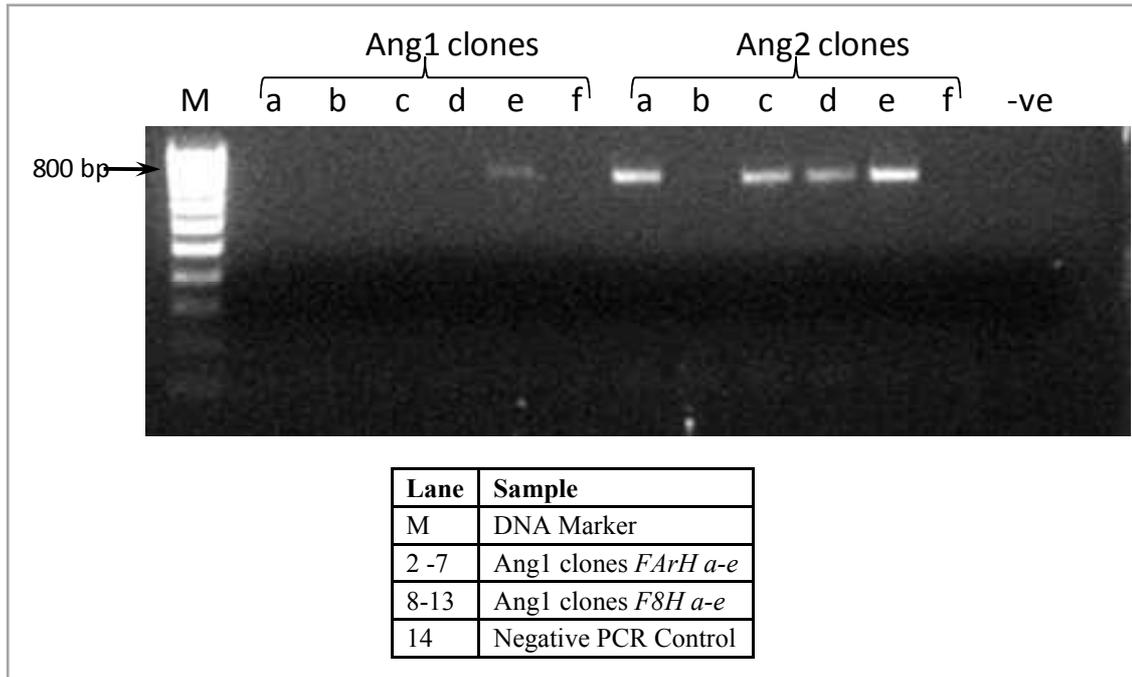
The expression of Ang1 FReD cloned in the pSE2 vector was undetectable on CHO, HEK293 and Ramos cell surface (Chapter 5). It is possible that Ang1 FReD cloned in pHyperm2 vector may not behave similarly when expressed in the chicken B cell line DT40. Hence in spite of previously observed poor expression, Ang1 FReD was attempted to be cloned in pHyperm2 and expressed in DT40 cells in this chapter.



**Figure 6.3: Schematic representation for cloning FLAG tagged FReD (fibrinogen domain) and the ASGPR-surface anchor protein in the pHypermut2 vector.** The ASGPR-linked FLAG-tagged FReD domains were amplified by PCR. The ASGPR Nhe 5' primer and EcoRV Ang1-FReD3' primer for amplifying Ang1 FReD were used. The ASGPR Nhe 5' primer and EcoRV Ang2-FReD3' primer for amplifying Ang2 FReD were used. The PCR product and the pHypermut2 vector were restricted with NheI and EcoRV restriction enzymes. The restricted products were ligated and competent cells were transformed with the ligation product. Colonies were picked from the transformation plate and plasmid was prepared from the colonies. Colonies were also analyzed by colony PCR to screen for the presence of the desired insert. Sequence data were obtained for the positive clones.

The ASGPR linked fibrinogen domain was amplified by PCR from the plasmids *FAr* and *F8* obtained in section 5.5. Briefly, the ASGPR Nhe 5' primer and EcoRV Ang1-FReD3' primer for amplifying Ang1 FReD or EcoRV Ang2-FReD3' primer for amplifying Ang2 FReD were used at annealing temperature 50 °C. The amplified product contained a NheI restriction enzyme site at its 5' end, followed by the ASGPR anchoring domain sequence, the FLAG epitope tag, the FReD domain (Ang1 FReD or Ang2 FReD) and EcoRV restriction enzyme site at its 3' end. The size of the PCR product was expected to be 795 bp. This PCR product was separated on a 1% agarose gel and a product of the correct size of approximately 800 bp was obtained. The DNA was extracted and purified from the gel as described in section 2.4.1. Both the purified PCR product and the pHypermut2 vector were digested with NheI and EcoRV restriction enzymes. To remove the restriction enzyme and its buffers from the digestion, both the restricted PCR product and the restricted vector were purified by gel extraction. The purified restricted insert and vector were ligated in a ratio of 3:1 as described in section 2.6.1. The ligation reaction was transformed into competent cells.

Six colonies each labelled *FArH a-e* and *F8H a-e* for Ang1 FReD and Ang2 FReD were picked from the transformation plate. Plasmids were prepared from bacterial cultures of these colonies. Diagnostic colony PCR was performed to detect insert. The product obtained from the diagnostic colony PCR was analyzed on a 1% agarose gel as shown in figure 6.4. Clone *Ang1-FReD FArH e* and clones *Ang2-FReD F8H a, c, d* and *f* were positive from the diagnostic colony PCR showing a band of the expected size at approximately 800 bp.



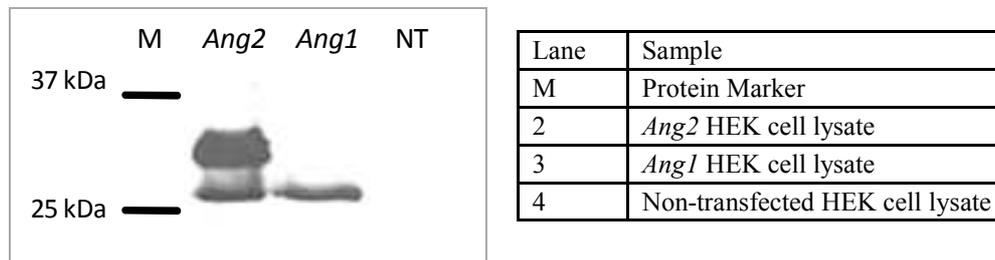
**Figure 6.4: Colony PCR screen for pHypermut2-ASGPR-FReD constructs.** The Ang1 and Ang2 ASGPR-linked FLAG-tagged FReD domain was amplified by PCR. The PCR product and the pHypermut2 vector were restricted and the restricted products were ligated. Competent cells were transformed with the ligation product. Colonies were also analyzed by colony PCR to screen for the presence of the desired insert. Six colonies labelled *FArHa-f* (Ang1-FReD) and *F8H a-f* (Ang2-FReD) were picked from the transformation plate and analysed by diagnostic colony PCR at annealing temperature 50°C using the Nhe ASGPR 5' and EcoRV Ang2-FReD3' primers. The positive clones should give a band at 795 bp. Bands at approximately 800bp were obtained for plasmid *F8H a, b, c d* and *e* were obtained. Band at approximately 800bp was obtained for plasmid *FArH e* was obtained.

Plasmid was prepared for the positive clones and sequenced using 5' ASGPR NheI and 3'Ang1EcoRV primers or 5' ASGPR NheI and 3'Ang2EcoRV primers for Ang1-FReD and Ang2-FReD domains respectively. The colony PCR positive clones were obtained with the correct sequence. Sequence data shown in Appendix 4I and 4J.

Plasmid constructs *Ang1-FReD FArHe* and *Ang2-FReD F8Hd* were analysed for protein expression by Western blotting. HEK293 cells were transfected with the *Ang1-FReD* and *Ang2-FReD* constructs. 24hours post-transfection, the transfected cells were lysed in sample buffer and the lysate separated by SDS-PAGE (as described in section 2.18). Both the Ang1 and Ang2 FReD domains cloned into the pHypermut2 vector are

FLAG-tagged. Western blot was performed using anti-FLAG primary antibody and HRP linked anti-mouse secondary antibody to confirm the expression of FReD domain.

The expected size of the protein for Ang1-FReD and Ang2-FReD was approximately 29kDa. Figure 6.5 (lane 3), shows that the lysate obtained from non-transfected cells does not show any band between 25 and 37 kDa. Figure 6.5 (lane 1), showing lysate obtained from HEK cells transfected with *Ang2-FReD* express two forms of the Ang2-FReD protein between 25 and 37 kDa. These two forms of proteins possibly represent glycosylated and non-glycosylated forms of the Ang2-FReD protein or active and degraded/cleaved fragment of the Ang2-FReD protein. Lane 2 is the lysate from the *Ang1-FReD* transfected HEK cells which shows a single band between 25 and 37 kDa.



**Figure 6.5: Protein expression of the Ang1-FReD-pHypermut and Ang2-FReD-pHypermut2 constructs in HEK293 cell line.** HEK cells were transfected with the *Ang1-FReD FArHe* and *Ang2-FReD F8Hd* constructs. 24-hours post transfection the cells were lysed and the lysate was analysed for protein expression by Western blotting. Both the Ang1 and Ang2 FReD domains cloned into the pHypermut2 vector are FLAG-tagged. These FLAG-tagged proteins were probed with anti-FLAG primary antibody and HRP linked anti-mouse secondary antibody. The expected size of the protein for *Ang1-FReD FArH e* and *Ang2-FReD F8H d* was ~29.5 kDa. Lysate from non-transfected cells (lane 3) does not show any band between 25 and 37 kDa. Lane 1 shows two forms of the expressed Ang2-FReD protein between 25 and 37 kDa. Lane 2 is the lysate from the *Ang1-FReD FArH e* transfected HEK cells which shows a single band between 25 and 37 kDa. This experiment was performed 4 times with similar results for *Ang2-FReD F8H d* protein expression and protein expression of a single band for *Ang1-FReD FArH e* as detected by Western blotting.

This experiment was performed 4 times for Ang1-FReD and Ang2-FReD protein expression. The *Ang1-FReD* construct consistently showed only one band (1 experiment) or no protein expression (3 experiments). The *Ang2-FReD* construct showed consistent double band expression in 4 experiments. Expression studies of the *Ang1-FReD* construct showed that the protein had undetectable or poor expression level in the HEK293 cell line; hence further experiments were continued only with the *Ang2-FReD* construct. The *Ang2-FReD* construct with its FLAG-tagged ASGPR anchored Ang2-FReD protein in the pHypermut2 vector will henceforth be referred to as cell surface expressed Ang2-FReD domain throughout this chapter.

## **6.2 Expression and localization of Ang2-FReD domain on DT40 cell surface.**

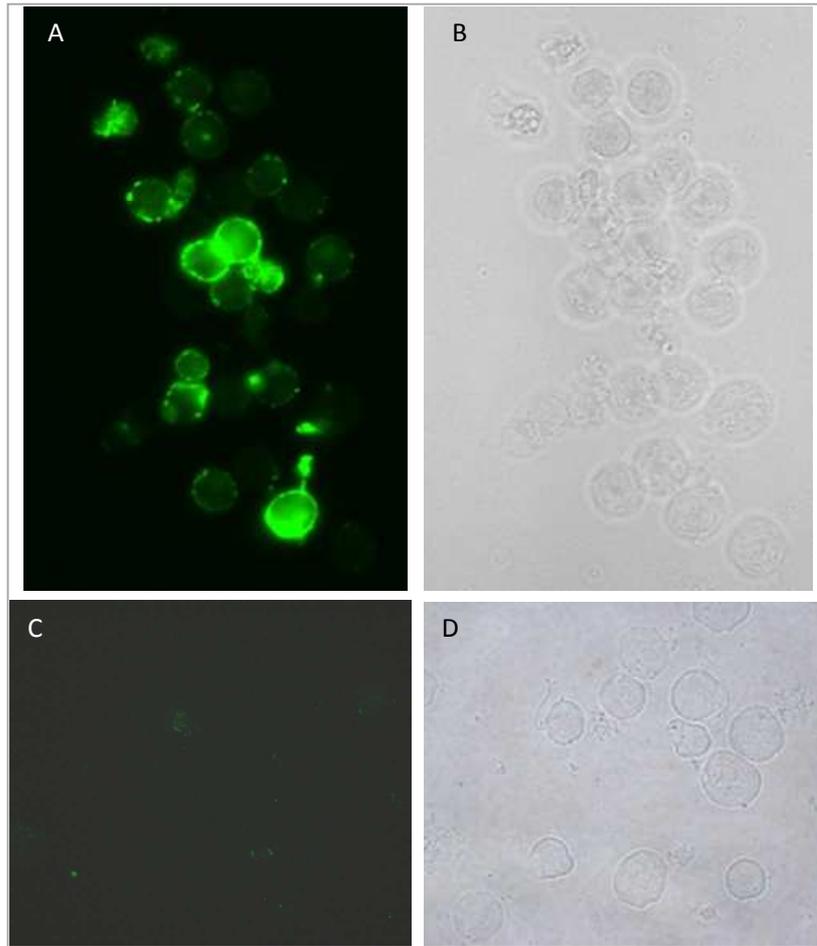
In order to utilize the ASGPR-Ang2-FReD pHypermut2 construct (*F8Hd*) cloned and expressed in section 6.1 it is required that the Ang2-FReD domain is localised to the extracellular face of the plasma membrane. Localisation of the domain on the cell surface allows for easy selection and screening of affinity variants. Protein localisation was therefore verified by immunofluorescence staining of stable DT40 cells transfected with the ASGPR-Ang2-FReD pHypermut2 (*F8Hd*) construct. Immunofluorescence staining also ensures that the cloning of the FReD domain into the surface display vector has not hindered the detection of the FLAG epitope tag on the cell surface.

DT40 cells were transfected with the ASGPR-Ang2-FReD pHypermut2 construct by electroporation (Section 2.15.3). The transfected cells were grown in selective antibiotic puromycin until colonies of stable clones were obtained (Section 2.16.2). These cells expressing FLAG-tagged Ang2-FReD domain protein were incubated with monoclonal anti-FLAG primary antibody and this was detected using fluorescently

labelled anti- mouse Cy2 antibody. Cells were visualised at 40X magnification under a fluorescence microscope.

Figure 6.6A shows DT40 cells transfected with ASGPR-FReD-pHypermut2 construct (obtained in section 6.1) and immunostained for FLAG tag. Cells exhibit clear membrane staining. It should be noted that the fluorescence on the surface of the cells varies. This may be due to the variable expression level of the protein in each cell. Expression may vary due to differences in integration of the DNA in the genome. The fluorescence variation may also be due to differences in the amount of protein expressed on the cell surface and the rate of internalisation of the protein for each cell which in turn may depend on stage in cell cycle of each cell. Figure 6.6B is the phase contrast view of the same field as 6.6A. On comparing the cells in figure 6.6A and B shows almost all cells are positively stained for FLAG tagged Ang2-FReD.

Figure 6.6C shows non-transfected DT40 cells labelled with anti-FLAG and Cy2 antibodies. Figure 6.6D is a phase contrast view of the field shown in 6.6C. These non-transfected cells serve as negative control in this experiment and none of the cells show any staining with the fluorophore indicating that the staining seen in figure 6.6A is specific for the cell surface expressed FReD domain.



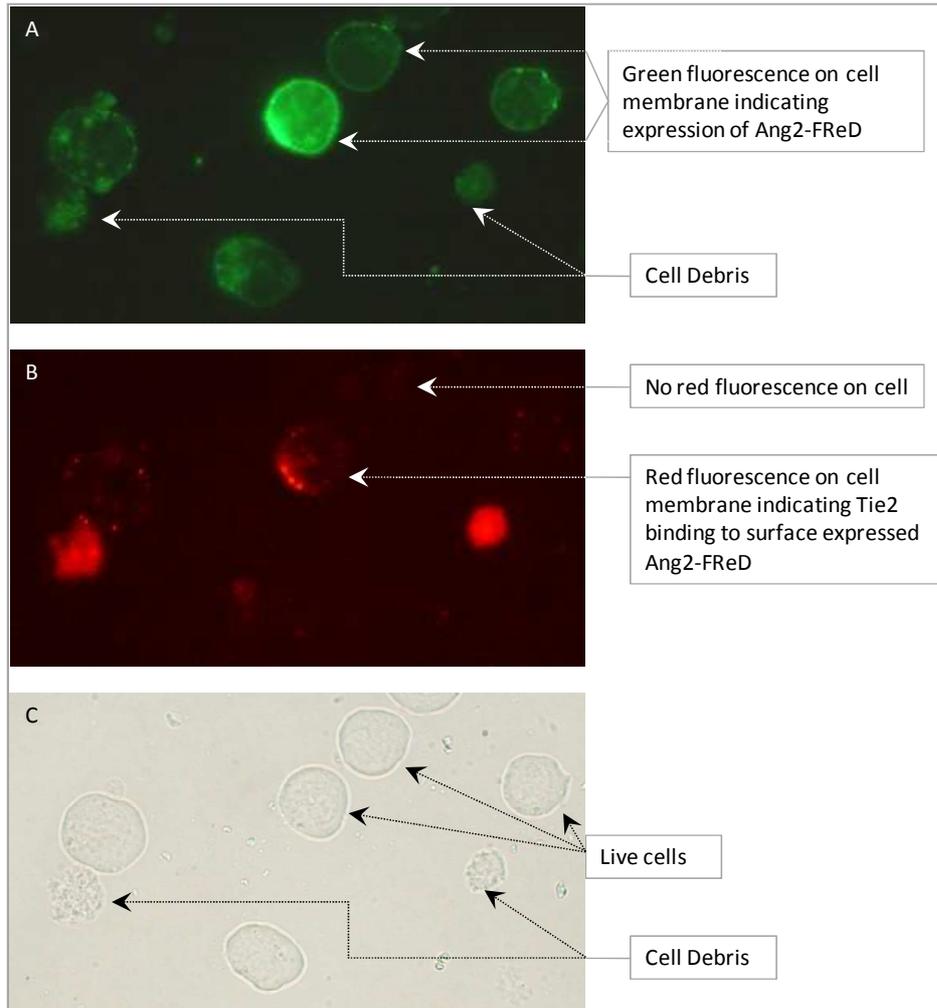
**Figure 6.6: Localization of cell bound ASGPR-FReD (Ang2) domain on stable DT40 cell surface by immunofluorescence.** DT40 cells were transfected with the ASGPR-Ang2-FReD pHypermut2 construct (*F8Hd*) and stable cells were obtained by selective puromycin antibiotic pressure. The FLAG-tagged FReD domain protein was expected to be expressed on the cell membrane anchored by the ASGPR transmembrane domain. The protein localisation was confirmed by immunohistochemistry. (A) The stable cells expressing FLAG-tagged Ang2-FReD domain protein were incubated with mouse monoclonal anti-FLAG primary antibody. The anti-FLAG antibody bound to the FLAG epitope tag at the N terminus of the Ang2-FReD domain on the cell surface of the cells. The cells were then fluorescently labelled with anti- mouse Cy2 antibody which bound to the primary antibody. The green fluorescing cells were visualised at 40X magnification under a fluorescence microscope. Cell surface protein expression was confirmed by observed green fluorescence on the cell membrane. Variation in fluorescence was observed for the stable cells which may be due to differential protein expression or variation of amount of protein translocated onto the cell surface. (B) Phase contrast of the field shown in figure 6.6A, showing all the cells in the field view under white light. (C) Non-transfected DT40 cells served as a negative control in this experiment and were stained with anti-FLAG primary and secondary Cy2 fluorophore. Non-transfected cells did not show any fluorescence, indicating that the staining of the stable cells in 6.6A was specific for the surface expressed FReD domain. (D) Phase contrast of the field shown in figure 6.6C.

Thus the ASGPR anchored FReD (Ang2) domain cloned in pHypermut2 vector was localised to the extracellular face of the DT40 cell membrane as detected by immunofluorescence staining.

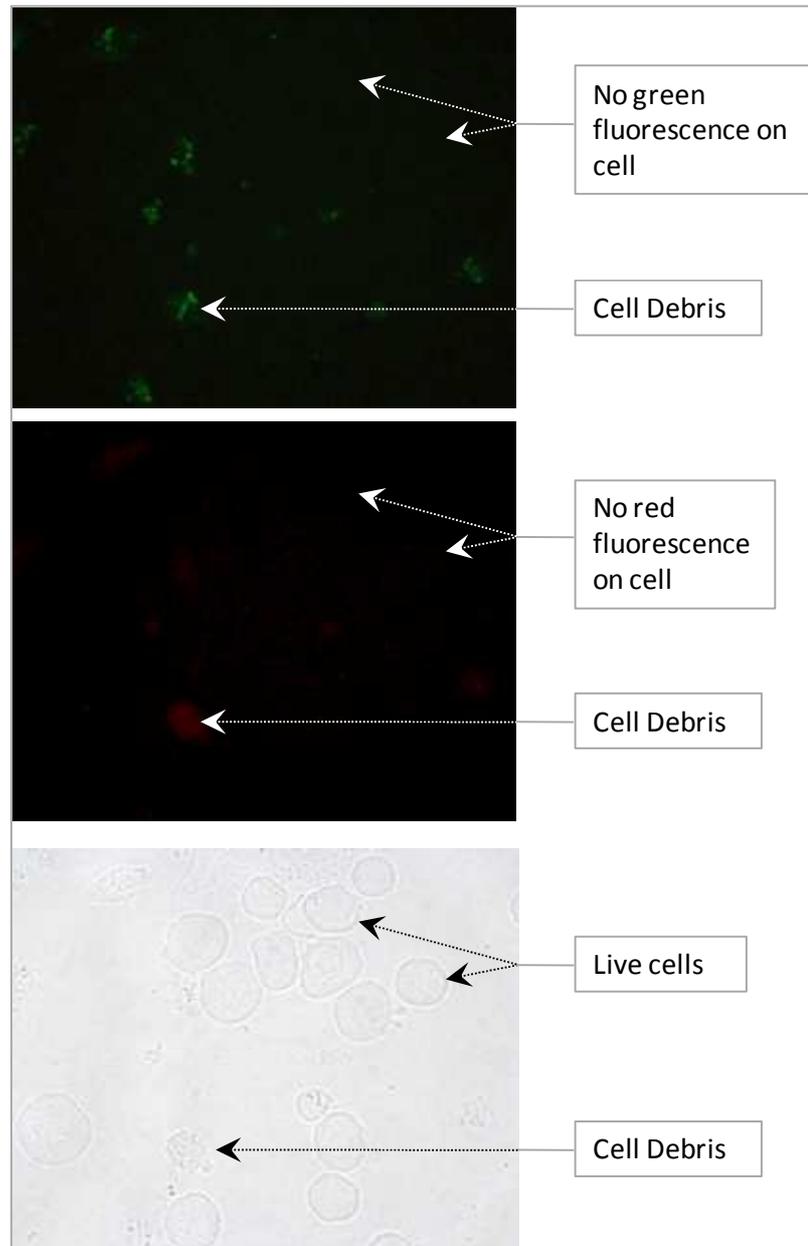
### **6.3 Immunofluorescence staining to show binding of Tie2 to cell surface expressed FReD domain in DT40 stable cells.**

In section 6.2 the Ang2-FReD domain protein expressed on the cell membrane anchored by the ASGPR transmembrane domain was shown to be localised to the extracellular surface of the DT40 cell by immunostaining for FLAG epitope tag. In this section the ability of the cell surface FReD domain to bind to soluble *rh*Tie2 ectodomain was studied by dual immunofluorescence staining (as described in section 2.20.3).

Briefly, the Ang2-FReD expressing stable DT40 cells were incubated with 5nM soluble *rh*Tie2 ectodomain (His-tagged). The cells were further incubated with anti-FLAG and biotin-linked anti-His primary antibodies to detect the FLAG-tagged FReD domain and the Tie2 ectodomain (bound to the FReD domain), respectively. The primary antibodies were washed off and the cells were further incubated with Cy2 and streptavidin-R-phycoerythrin fluorescent secondary antibody. The green (Cy2) and red (Phycoerythrin) fluorescence of the dual stained cells were visualised at 40X magnification by fluorescent microscopy.



**Figure 6.7: Binding of surface expressed ASGPR-FReD (Ang2) domain to Tie2 ectodomain by immunofluorescence.** Ang2-FReD expressing stable DT40 cells were incubated with 5nM soluble *rh*Tie2 ectodomain. The cells were further incubated with anti-FLAG and biotin-linked anti-His primary antibodies to detect the FLAG-tagged FReD domain and the Tie2 ectodomain, respectively. The cells were further incubated with Cy2 and streptavidin-R-phycoerythrin fluorescent secondary antibodies. Cy2 binds to the FReD bound anti-FLAG primary and the streptavidin-R-phycoerythrin binds to the Tie2 bound biotinylated anti-His antibody. This experiment was performed 4 times and similar results were obtained. (A) The dual stained cells were observed at 40X magnification under a fluorescent microscope to detect green fluorescence of the Cy2 fluorophore. Cell surface protein expression was confirmed by green immunofluorescence, indicating detection of the FLAG epitope tag on the cell surface expressed FReD protein. (B) The field visualised in 6.7A was also observed for red fluorescence of the Phycoerythrin fluorophore. Receptor binding was confirmed by immunofluorescence of the bound His tagged Tie2 receptor, in the same field as shown in figure A. (C) Phase contrast of the fields shown in figure 6.7 A and B, showing all the cells in the field under white light.



**Figure 6.8: Non-transfected DT40 cells served as a negative control against the transfected cells which demonstrated Tie2 binding.** Non transfected cells were also stained with Cy2 and Phycoerythrin fluorophores in the same experiment as shown in figure 6.7 for transfected cells. (A) Non-transfected cells did not show any green fluorescence, indicating that the staining of the stable cells in 6.7A was specific for the surface expressed FReD domain. (B) Non-transfected cells did not show any red fluorescence, indicating that the staining of the stable cells in 6.7B was specific for the surface expressed FReD domain. (C) Phase contrast of the field shown in figure 6.8D and E, showing all the cells in the field under white light.

As shown in Figure 6.7A cells transfected with ASGPR-FReD-pHypermut2 construct (obtained in section 6.1) fluoresce green with Cy2 fluorescent antibody

indicating cell surface expression of the FLAG-tagged FReD domain. Figure 6.7B is the view of field shown in Figure 6.7A visualizing red fluorescence of the expressing cells. Some of the cells which show cell surface expression of FReD domain also show staining by the streptavidin-R-Phycoerythrin antibody which detects Tie2 ectodomain bound to FReD domain. A low concentration of Tie2 used in this experiment may account for selective staining of some cells for cell bound Tie2. Figure 6.7C is the phase contrast view of the fields shown in figure 6.7A and 6.7B.

Figure 6.8A and B show non-transfected cells also stained with Cy2 and Phycoerythrin fluorophores in the same experiment and visualised at 40X magnification under a fluorescent microscope. The non-transfected cells did not stain green or red, indicating that the fluorescence observed in 6.7 A, B was specific for the stable Ang2-FReD expressing cells.

Immunostaining in figure 6.7A demonstrates the expression of FReD domain on the extracellular surface and binding of it to soluble Tie2 as seen in figure 6.7B. Once the expression and binding of the FReD domain had been established in the stable DT40 population of cells, the cells undergoing SHM could now be subjected to affinity selection of high affinity binders. The frequency of mutation in these cell surface Ang2-FReD expressing stable cells was assessed before starting affinity selection.

#### **6.4 Frequency of mutation in the starting stable populations of DT40 cells expressing Ang2-FReD domain.**

Prior to putting the stable cell population through affinity selection, a random sample of stably transfected cells were collected and 10 clones from this random sample sequenced. Of the 10 clones sequenced 1 variant labelled *Hm5* was obtained (Appendix 4K). The mutation was at Gly318 in the Ang2-FReD domain. The 318 glycine residue was mutated to arginine in the *Hm5* mutant. This confirms the population contains mutant forms of Ang2-FReD.

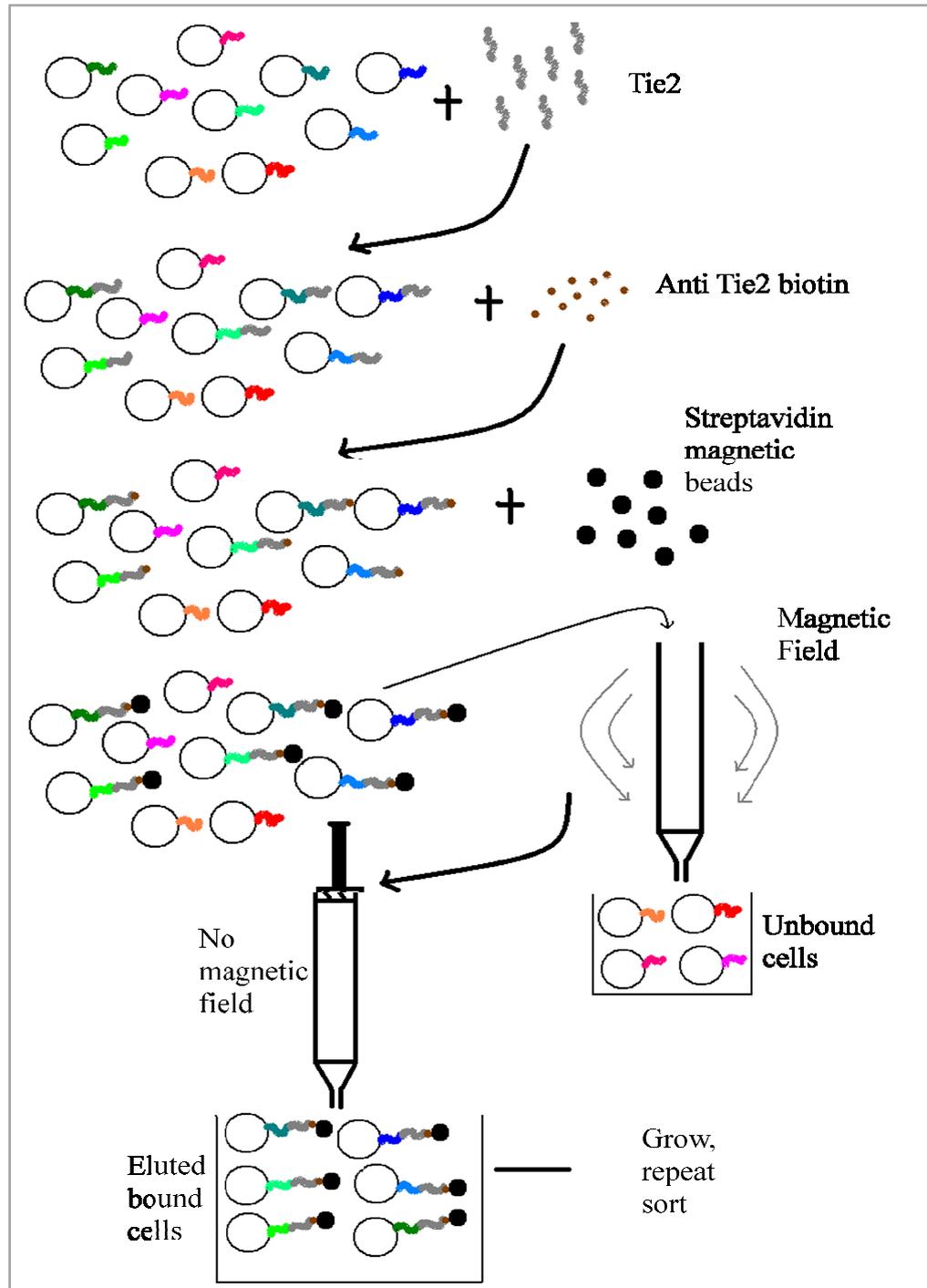
The frequency of mutation in the starting stable populations of DT40 cells expressing Ang2-FReD domain was 0.1. More clones were analysed by sequencing in sections 6.5.3 and 6.6.2. The number of mutations in the GFP DNA sequence obtained by Arakawa et al. (2008) in an unsorted population of cells in a 6 week culture was 13 in 39 sequences, giving a resulting frequency of mutation at  $\sim 0.33$  (Arakawa et al. 2008). Frequency of mutation in the Ang2-FReD domain is one third of that obtained for the GFP sequence. This may be due to fewer hotspots in the Ang2-FReD domain sequence compared with the GFP sequence. At least  $10^7$  cells were routinely used for selection and these would be expected to contain approximately a million mutants at the 0.1 frequency obtained, giving a sufficiently large library of mutants to select from. Once the frequency of mutation for the starting population of cells had been established the cells were then subjected to magnetic sorting for affinity selection.

## 6.5 Magnetic bead sorting.

Once the expression of Ang2-FReD domain and binding of Tie2 to the Ang2-FReD domain was confirmed in section 6.3 the DT40 stable cells were enriched for Tie2 binding using magnetic bead sorting.

Ten million stable DT40 cells characterized for expression of Ang2 FReD and binding to Tie2 were sorted using magnetic beads as described in section 2.21. Figure 6.9 is a schematic illustration of the magnetic sorting experiment performed.

Briefly, the method involves binding of the soluble *rh*Tie2 receptor ectodomain to the variants in the stable cells. The Tie2 binds to a selected number of cells which may express FReD domain variant having high affinity for the receptor. The Tie2 bound cells are then incubated in a biotinylated anti-His antibody. Magnetic beads coated with streptavidin bind to the biotinylated cells. Thus cells expressing FReD domain variants which bind to Tie2 receptor are bound and selected on the magnetic beads. These mix, of bead bound and unbound cells are put through a column under the influence of a magnetic field. The magnetic bead bound cells remain in the column whereas the unbound cells flow through the column. The column bound cells are plunged out from the column in the absence of the magnetic field. These eluted cells are cultured in media until they have grown sufficiently ( $>10^7$ ) to undergo another round of magnetic affinity selection.



**Figure 6.9: Schematic representation of magnetic sorting approach to enrich for high affinity binders from a stable DT40 cell population expressing Ang2 FReD variants.** The stable cell population expressing different variants of Ang2-FReD were incubated with limited amounts of Tie2 receptor ectodomain. Only the high affinity Ang2-FReD can bind at such limited concentration. The mix was then incubated with anti-Tie2 biotin conjugate. Finally streptavidin coated magnetic beads were added. In the presence of a magnetic field the magnetic bead bound high affinity Ang2-FReD expressing cells were not eluted. The unbound cells were washed out from the column. In the absence of the magnetic field the high affinity Ang2-FReD expressing cells were eluted using a plunger. The sorting was repeated thrice and the cells analysed by flow cytometry to determine improvement.

### **6.5.1 Selection of high affinity Tie2 binders using magnetic bead sorting.**

Several binding studies have been performed using Ang and its various recombinant and truncated forms (See Table 1.1, Chapter 1). The most recent report on Ang2 and Ang1 binding to sTie2 suggests that Ang1 has 20-fold higher affinity for Tie2 than Ang2. Concentrations of ligands bound to achieve 50% saturation was 106ng/ml (~1.9nM) for Ang1 and 2,018ng/ml (40.2nM) for Ang2 (Yuan et al. 2009). Based on *in vitro* studies by surface plasmon resonance (SPR) the binding stoichiometry of Tie2 to Ang1 fibrinogen related domain (FReD) gave an IC<sub>50</sub> value of 435nM which was much lower than that of native Ang1 (IC<sub>50</sub> 1.5nM). This difference in binding could be attributed to the avidity effects of higher order native Ang1 making the binding several fold stronger (Davis et al. 2003). However, several studies have also indicated that Ang1 and Ang2 bind with the same affinity (Maisonpierre et al. 1997, Fiedler et al. 2003, Davis et al. 1996).

Magnetic bead sorting described above was used to iteratively select for high affinity binders to Tie2. For each round of iterative selection of Tie2 binders, decreasing concentration of Tie2 was selected and the cells incubated with it. The first round of magnetic sorting was under non-stringent conditions of 10nM Tie2. Second and third sorts were performed using 0.45nM Tie2. The population of cells obtained after 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> rounds of magnetic sorting were then analyzed by flow cytometry as described in section 2.22. Table 6.2 summarizes the rounds of magnetic affinity selection.

Sort Round	Initial number of cells	Population fluorescently stained	Tie2 conc used for sort (nM)	Number of cells recovered	% (of the total) population recovered	Population recovered labelled
0		---	---	---	---	<i>Unsorted</i>
1	$1.6 \times 10^7$	<i>Unsorted</i>	10	$111 \times 10^4$	7%	<i>1<sup>st</sup> Sorted</i>
2	$1.08 \times 10^7$	<i>1<sup>st</sup> Sorted</i>	0.45	$17 \times 10^4$	1.6%	<i>2<sup>nd</sup> Sorted</i>
3	$0.5 \times 10^7$	<i>2<sup>nd</sup> Sorted</i>	0.45	$5 \times 10^4$	1%	<i>3<sup>rd</sup> Sorted</i>

**Table 6.2: Summary of the rounds of magnetic affinity selection, indicating the number of cells put through each round and the percentage of cells recovered after each round.** Stable cells were incubated with the indicated concentration of Tie2 for each round of selection. The variant cells were then incubated with a biotinylated anti-Tie2 antibody which bound to cell bound Tie2 ectodomain. Magnetic beads coated with streptavidin bind to the biotinylated cells. Thus cells expressing FReD domain variant having high affinity for the Tie2 receptor bind to the magnetic beads. The first round of magnetic sorting was under non-stringent conditions of 10nM Tie2 to select for a wide range of binders. Second and third sorts were performed using 0.45nM Tie2. The population of cells obtained after 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> rounds of magnetic sorting were labelled *1<sup>st</sup> sorted*, *2<sup>nd</sup> sorted* and *3<sup>rd</sup> sorted* respectively.

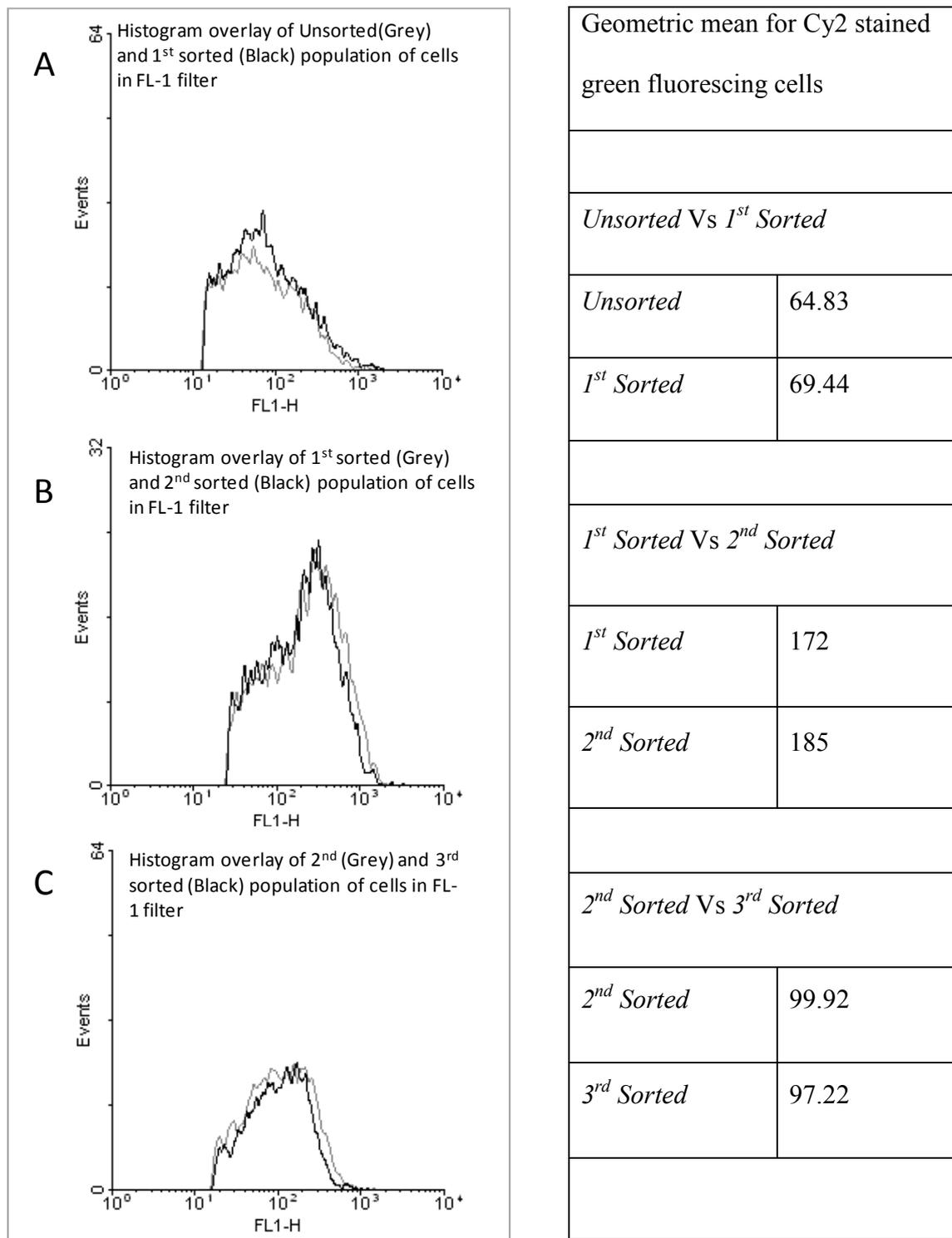
Non-transfected control samples were also analysed simultaneously by magnetic sorting. Non-transfected cells were incubated with Tie2, biotinylated anti-His antibody and the streptavidin coated magnetic beads exactly like the transfected stable cells during the iterative rounds of sorting. The percentage of non-transfected cell population recovered as a background through each round was ~0.1% cells of total number of cells in the starting population applied to the column.

### **6.5.2 Flow cytometry analysis and comparison of *Unsorted* population of cells with the magnetically sorted *1<sup>st</sup> Sorted*, *2<sup>nd</sup> Sorted* and *3<sup>rd</sup> Sorted* population of cells.**

The *1<sup>st</sup> sorted*, *2<sup>nd</sup> sorted* and *3<sup>rd</sup> sorted* population of cells were compared to assess their binding to Tie2 ectodomain using flow cytometry. Approximately  $10^5$  cells for each, the *unsorted* and *1<sup>st</sup> Sorted* population of cells were collected and incubated with 10nM Tie2 and dual stained for flow cytometry analysis as described in section 2.22. Briefly, the cells incubated with Tie2 were washed to remove unbound Tie2 and incubated with mouse monoclonal anti-FLAG and biotin-linked anti-His primary antibodies followed by anti-mouse Cy2 and Streptavidin-R-Phycoerythrin fluorophores.

Varying expression levels of the Ang2-FReD were observed by immunofluorescence previously in Section 6.2. To test whether the sorting procedure selected for cells expressing different levels of FReD domain, the fluorescent intensity corresponding to FLAG tag was compared between cell populations.

Figure 6.10A compares the expression levels between the *Unsorted* and *1<sup>st</sup> Sorted* population of cells. As seen from the geometric mean (Gm) values represented for each population the expression levels for the FReD domain is similar between populations of sorted cells. Similar expression levels of FReD domain for *1<sup>st</sup> Sorted* and *2<sup>nd</sup> Sorted* population of cells (Figure 6.10B) and also *2<sup>nd</sup> Sorted* and *3<sup>rd</sup> Sorted* population of cells (Figure 6.10C) was also observed.



**Figure 6.10: Histograms overlay of the various magnetically sorted populations of cells, indicating the protein expression level detected the Cy2 (green) fluorescence.** As observed (for each population set) from the histogram overlays and the geometric mean indicated in the table, both populations of cells (in each set) have similar levels of expression of cell- surface Ang2-FReD. (A) Histograms overlay of *Unsorted* (Grey) and *1<sup>st</sup> Sorted* (Black) population of cells. (B) Histograms of *1<sup>st</sup> Sorted* (Grey) and *2<sup>nd</sup> sorted* (Black) population of cells. (C) Histograms of *2<sup>nd</sup> sorted* (Grey) and *3<sup>rd</sup> Sorted* (Black) population of cells.

The geometric mean values for each population were similar indicating similar expression levels of the FReD domain on the cell surface in all pairs of populations. Thus the comparison of populations for affinity improvement (red fluorescence of labelled bound Tie2) was not influenced by protein expression.

The Tie2 concentrations used for comparison of magnetically sorted populations are listed in Table 6.3.

Populations Analysed	Tie2 concentration used for binding
<i>Unsorted Vs 1<sup>st</sup> Sorted</i>	10nM
<i>1<sup>st</sup> Sorted Vs 2<sup>nd</sup> Sorted</i>	1nM
<i>2<sup>nd</sup> Sorted Vs 3<sup>rd</sup> Sorted</i>	0.45nM

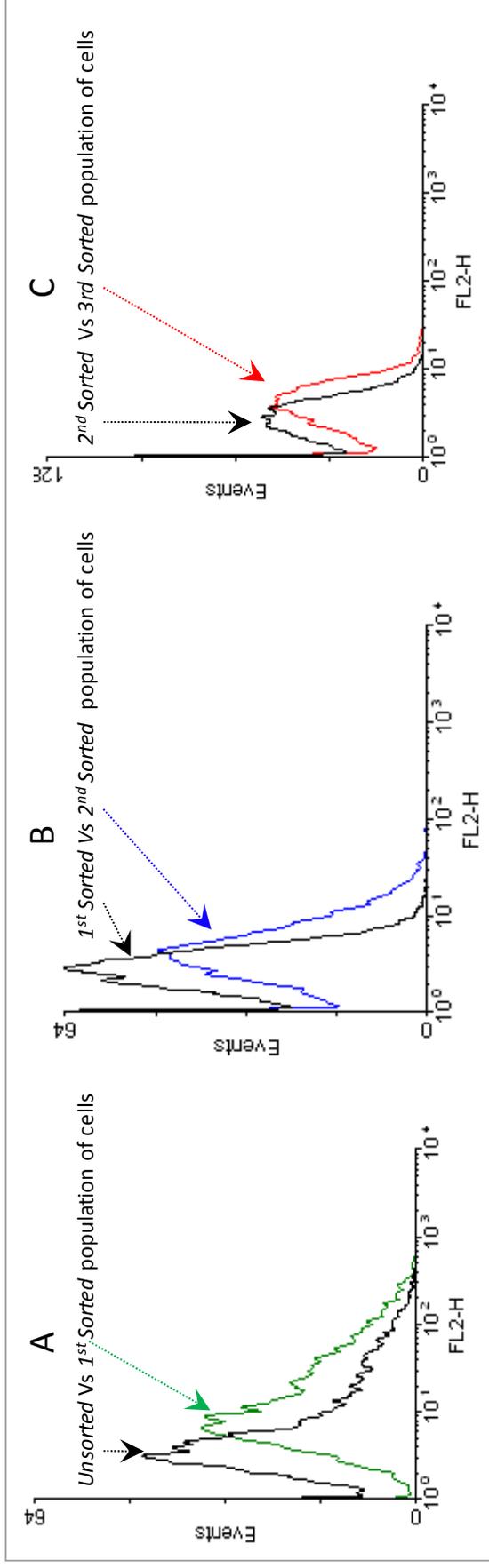
**Table 6.3: Tie2 concentrations used for analysing, *Unsorted*, *1<sup>st</sup> Sorted*, *2<sup>nd</sup> Sorted* and *3<sup>rd</sup> Sorted* population of cells.**

To examine Tie2 binding the red fluorescence was compared between cell populations (Figure 6.11). Overlay of histograms shows a distinct peak for the *1<sup>st</sup> Sorted* cells more towards the right compared to the *Unsorted* cells population peak. This movement of the cell population towards the right on the X-axis is an indicator of the population of cells having a higher red fluorescence. This higher red fluorescence reflects increased binding of Tie2 receptor ectodomain to the FReD domain variants expressed in *1<sup>st</sup> Sorted* population.

For comparing the *1<sup>st</sup> Sorted* and *2<sup>nd</sup> Sorted* population of cells approximately  $10^5$  cells for each population were used for incubation with 1nM Tie2. Subsequently both populations were dual stained for flow cytometry analysis as described previously. Data obtained from flow cytometry show the histogram for *1<sup>st</sup> sorted* cells over-layered with the *2<sup>nd</sup> sorted* cells in Figure 6.11B. As observed in comparison of *Unsorted* and *1<sup>st</sup>*

*Sorted* (Figure 6.11A), figure 6.11B also showed a similar movement of the histogram peak for the  $2^{nd}$  *sorted* population towards the right on the X-axis. This again indicates the selection of high affinity binders to the Tie2 receptor in the  $2^{nd}$  *sorted* population.

The comparison for  $2^{nd}$  *Sorted* and  $3^{rd}$  *Sorted* population of cells was performed at 0.45nM Tie2 and cells subsequently dual stained for flow cytometry analysis as described in section 2.23. Data obtained from flow cytometry show the histogram for  $2^{nd}$  *sorted* cells over-layed with the  $3^{rd}$  *sorted* cells in Figure 6.11C. Again the histogram peak for the  $3^{rd}$  *sorted* population was shifted to the right as compared with the  $2^{nd}$  *sorted* population.



Geometric mean for R-phycoerythrin stained red fluorescing cells			
	<i>1<sup>st</sup> Sorted Vs 1<sup>st</sup> Sorted</i>	<i>1<sup>st</sup> Sorted Vs 2<sup>nd</sup> Sorted</i>	<i>2<sup>nd</sup> Sorted Vs 3<sup>rd</sup> Sorted</i>
<i>Unsorted</i>	8.23	2.83	2.66
<i>1<sup>st</sup> Sorted</i>	14.93	3.98	3.55

**Figure 6.11: Histogram overlaid for comparison of Tie2 binding (red fluorescence) of various magnetically sorted population of cells.** Geometric mean for phycoerythrin (red) fluorescence was obtained for each population. Increase in Geometric mean indicates increased binding of the population of cells to the Tie2 ectodomain. (A) **Comparison of Unsorted population of cells with the magnetically sorted 1<sup>st</sup> sorted population of cells.** Histograms for *Unsorted* (coloured black) cells overlaid with *1<sup>st</sup> Sorted* (green) cells. The shift in peak of the histogram for the *1<sup>st</sup> Sorted* cells compared to the *Unsorted* cells indicates improved binding of the *1<sup>st</sup> Sorted* cell population to Tie2. (B) **Comparison of 1<sup>st</sup> sorted population of cells with the magnetically sorted 2<sup>nd</sup> sorted population of cells.** Histograms for *1<sup>st</sup> sorted* (coloured black) cells overlaid with the *2<sup>nd</sup> Sorted* (coloured blue) cells compared to the *1<sup>st</sup> Sorted* cells indicates improved binding of the *2<sup>nd</sup> Sorted* cell population to Tie2. (C) **Comparison of 2<sup>nd</sup> sorted population of cells with the magnetically sorted 3<sup>rd</sup> sorted population of cells.** Histograms for *2<sup>nd</sup> sorted* (coloured black) cells overlaid with the *3<sup>rd</sup> Sorted* (coloured red) cells. The shift in peak of the histogram for the *3<sup>rd</sup> Sorted* (coloured red) cells compared to the *2<sup>nd</sup> Sorted* cells indicates improved binding of the *3<sup>rd</sup> Sorted* cell population to Tie2.

Each histogram overlay (representing 10,000 events) in figure 6.11 shows that with every subsequent round of magnetic sorting the population appears to have improved binding to Tie2 as compared with the previous populations. These data suggest that magnetic sorting was able to produce a gradual improvement in sorted cell populations (obtained in Section 6.5.1). However, magnetic sorting was unable to sort a pure population of high binders, as seen by the significant degree of overlapping observed for each histogram pair. An alternative sorting method of Fluorescence Activated Cell Sorting (FACS) was therefore used.

### **6.5.3 Frequency of mutation in the 1<sup>st</sup> sorted population obtained from magnetic sorting.**

The 1<sup>st</sup> round of magnetic sorting yielded a population of cells with the maximum shift in the histogram peak when compared with the *Unsorted* population. Hence, it was decided to check for the frequency of mutation in the 1<sup>st</sup> sorted population.

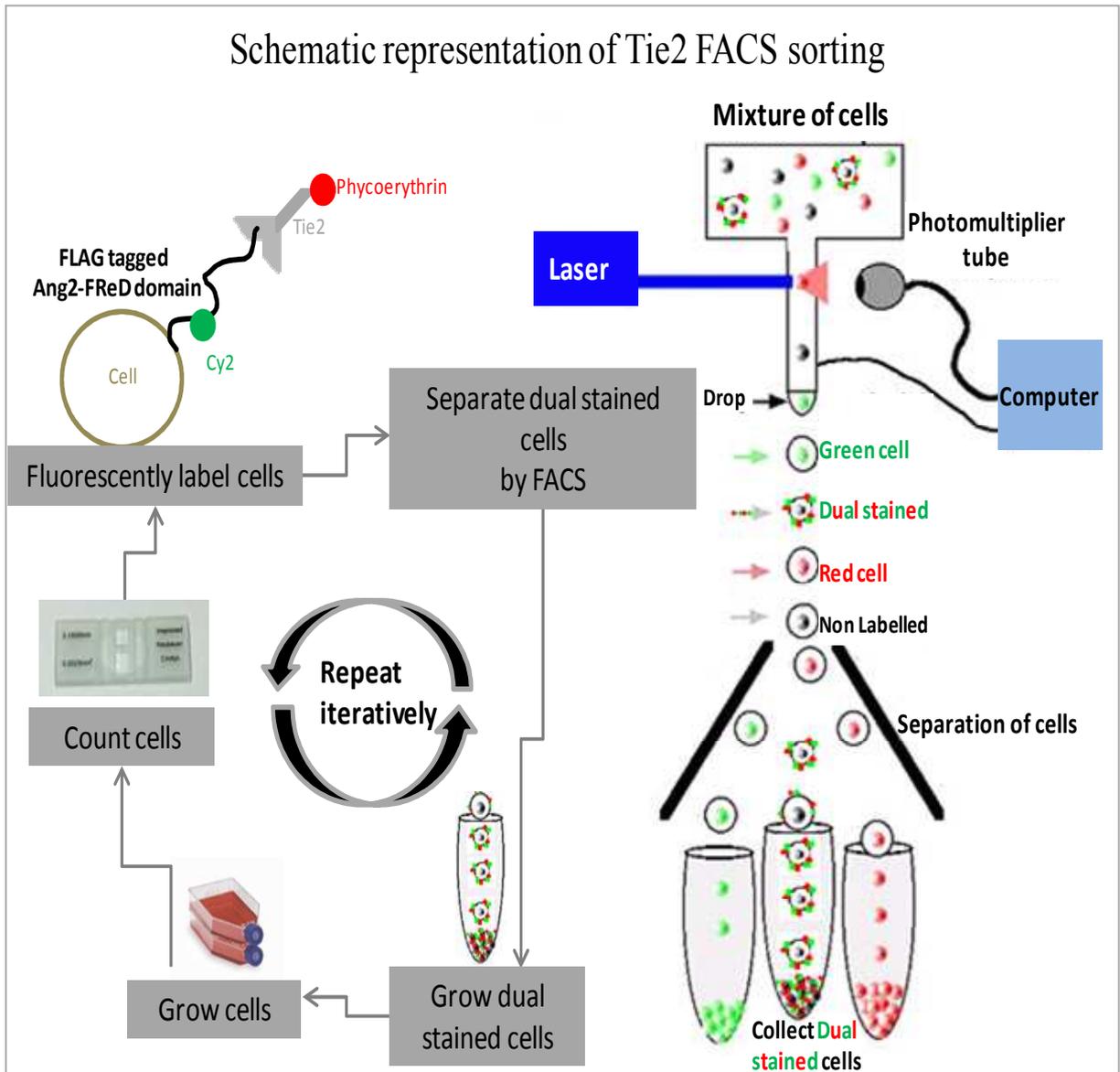
If the frequency of mutation in this population was improved as compared to the frequency of mutation in the starting stable population as described in section 6.4 then it could suggest that magnetic sorting was able to some extent select for high affinity Tie2 binding variants.

A sample of the magnetically sorted 1<sup>st</sup> sorted population was collected and RNA extracted. This RNA was used to reverse transcribe the FReD domain sequence and the DNA obtained was used to generate plasmids for sequencing as described in section 2.6.4. Five clones were picked and plasmids obtained and sequenced using the T7 primer. Of the 5 clones sequenced 1 variant labelled *I0c* was obtained. The frequency of mutation in this stable population of DT40 cells expressing Ang2-FReD domain was

0.2. This population of cells were grown and more clones from it sequenced in section 6.6.2. Further analysis of this mutant *10c* is performed in section 6.9 where it was characterised for its phenotype by  $K_d$  determination.

### **6.6 Selection screen of high affinity binders of Tie2 by Fluorescence Activated Cell Sorting (FACS).**

The stable DT40 cells characterized for expression of Ang2-FReD domain and binding to Tie2 in section 6.3 were screened and selected by FACS. Figure 6.12 is a schematic summary of the iterative rounds of FACS used for selection of high affinity binders to Tie2 receptor ectodomain. The population of cells is a mixture of non mutated FReD and new variants of FReD with affinities higher or lower than the native FReD domain which can be analyzed for their ability to bind soluble Tie2 receptor ectodomain.



**Figure 6.12: Schematic representation of the iterative rounds of FACS selection for Tie2 binding affinity.** A typical sort would involve growing up Ang2FReD expressing stable cells in culture, counting and collecting 10 million cells. Fluorescently labelling the cells for the expressed protein and the protein bound Tie2 receptor with Cy2 (green fluorescence) and R-PE (red fluorescence) dyes respectively. Analysing the labelled cells by FACS and sorting a suitable selected population of cells under sterile conditions. Growing up the selected cells and repeating the round of labelling and selection iteratively.

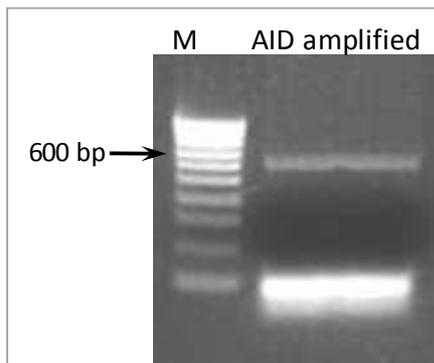
Before selection by FACS, the expression of the key enzyme involved in SHM; AID (Activation Induced [cytidine] Deaminase) was confirmed in the starting stable population of cells. Also the frequency of mutation in the starting stable population of cells was determined by sequencing clones from a representative random sample of cells.

Prior to FACS selection, experiments were also performed to determine the binding affinity (by  $K_d$  determination) of cell bound wild type Ang2 FReD to soluble *rhTie2* receptor ectodomain by flow cytometry.

#### **6.6.1 AID expression in the ASGPR-Ang2FReD stable cell population.**

To confirm that the low percentage of mutants obtained from magnetic enrichment was not due to absence of AID expression, the DT40 Ang2-FReD expressing stable cells were analysed for AID expression.

The stable Ang2-FReD expressing cells were collected and RNA extracted from them. Messenger RNA from the extracted RNA was transcribed into cDNA by RT-PCR as described in section 2.9. The cDNA was used for amplification of chick AID using 5'AIDchickfwd and 3'AIDchickrev at annealing temperature 55 °C. Figure 6.13 shows 5µl of the PCR product visualized on a 1% agarose gel, and a product of approximately 600bp (expected size 597bp) was obtained (lane2), indicating that the expression of AID in the cells.



**Figure 6.13: Detection of AID mRNA in Ang2-FReD expressing stable DT40 cells.** RNA was prepared from the Ang2-FReD expressing stable cells. mRNA was transcribed to cDNA by RT-PCR. AID was amplified from the cDNA by PCR. The size of the AID transcribed segment is expected to be 597bp. Lane 2 shows an amplified product at approximately 600bp indicating expression of AID in the stable cells.

These data confirm AID expression in the Ang2-FReD expressing DT40 stable cell population.

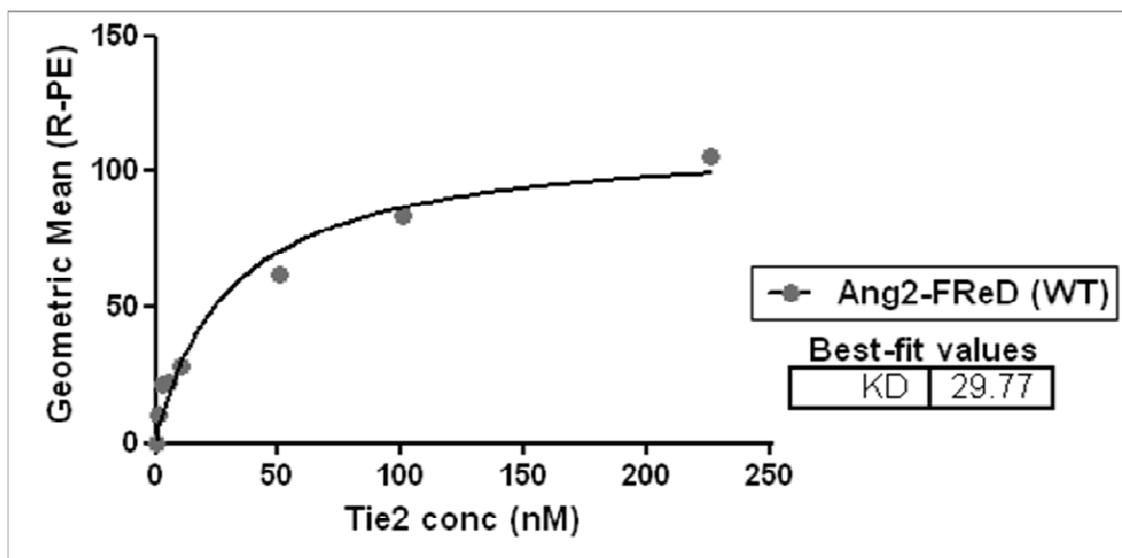
### 6.6.2 Frequency of mutation in the population of cells used for Fluorescence Activated Cell Sorting (FACS).

Prior to selection of the stable cell population using FACS, a random sample of stably transfected cells were collected and 10 clones from this random sample sequenced. Of the 10 clones sequenced 2 variants labelled *FACSD* and *FACSi* were obtained. These mutants were further characterised for their phenotype by  $K_d$  determination in section 6.9. The frequency of mutation in the stable populations of DT40 cells expressing Ang2-FReD domain prior to FACS selection was 0.2.

For each round of FACS sorting at least  $10^7$  cells were analysed, resulting in approximately two million mutants at the 0.2 frequency. This would give a sufficiently large library of mutants to select from.

### **6.6.3 Determination of $K_d$ for wild type Ang2 FReD domain transfected stable DT40 by Tie2 concentration curve.**

Before enriching cells by FACS, it was important to determine the dissociation constant ( $K_d$ ) to screen effectively for the cell surface display system. The affinities for FReD domain reported in the literature have been discussed (See Table 1.1, Chapter1). However, it is possible that in the cell surface expression system of display, the affinity may be different for the ASGPR anchored Ang2 FReD domain expressed in stable DT40 cells. Hence dissociation constant ( $K_d$ ) of wild-type Ang2-FReD domain binding to soluble Tie2 was determined in the cell surface display system used in DT40 cell line. Stable cells were incubated with a range of soluble *rh*Tie2 concentrations and fluorescently labelled for flow cytometry analysis as described in section 2.22.2. The red fluorescence of the phycoerythrin fluorophore was detected on the flow cytometer for each of the Tie2 concentrations. The geometric mean corresponding to binding for various Tie2 concentrations was obtained. To determine  $K_d$  values non-linear regression analysis was performed of the geometric mean fluorescence versus varying concentrations of *rh*Tie2 Fc chimera using Graphpad Prism. This experiment was performed three times and a  $K_d$  mean of 31.25nM with a SEM (Standard Error of the Mean) of +/-0.745 was obtained for cell surface expressed Ang2-FReD domain in DT40 cells. Figure 6.14 is a representative concentration curve for determination of  $K_d$ .



**Figure 6.14: Concentration curve to determine  $K_d$  of WT Ang2-FReD domain (in DT40 cell line) for binding Tie2.** Stable Ang2 FReD expressing DT40 cells were incubated with a range of soluble *rhTie2* ectodomain concentrations. Biotin-linked anti-His primary antibody was used to detect the His-tagged Tie2 bound on the cell surface of FReD domain. The cells were fluorescently labelled with Streptavidin-R-Phycoerythrin. The red fluorescence of the phycoerythrin fluorophore was detected for each of the Tie2 concentrations. The geometric mean corresponding to the various Tie2 concentrations was obtained. To determine  $K_d$  values non-linear regression analysis was performed of the geometric mean fluorescence versus varying concentrations of *rhTie2* Fc chimera using Graphpad Prism. The  $K_d$  value of the FReD domain on the stable cells was obtained. This experiment was performed three times and a  $K_d$  mean of 31.25 was obtained. The curve shown in this figure is a representative curve from one experiment. The bar graph for the three experiments gave a SEM of  $\pm 0.745$ ,  $n=3$ .

The  $K_d$  value of  $\sim 31$ nM for Ang2-FReD expressing DT40 cells was used as a guide to design further experiments to select high affinity binders for Tie2 using Fluorescence Activated Cell Sorting (FACS).

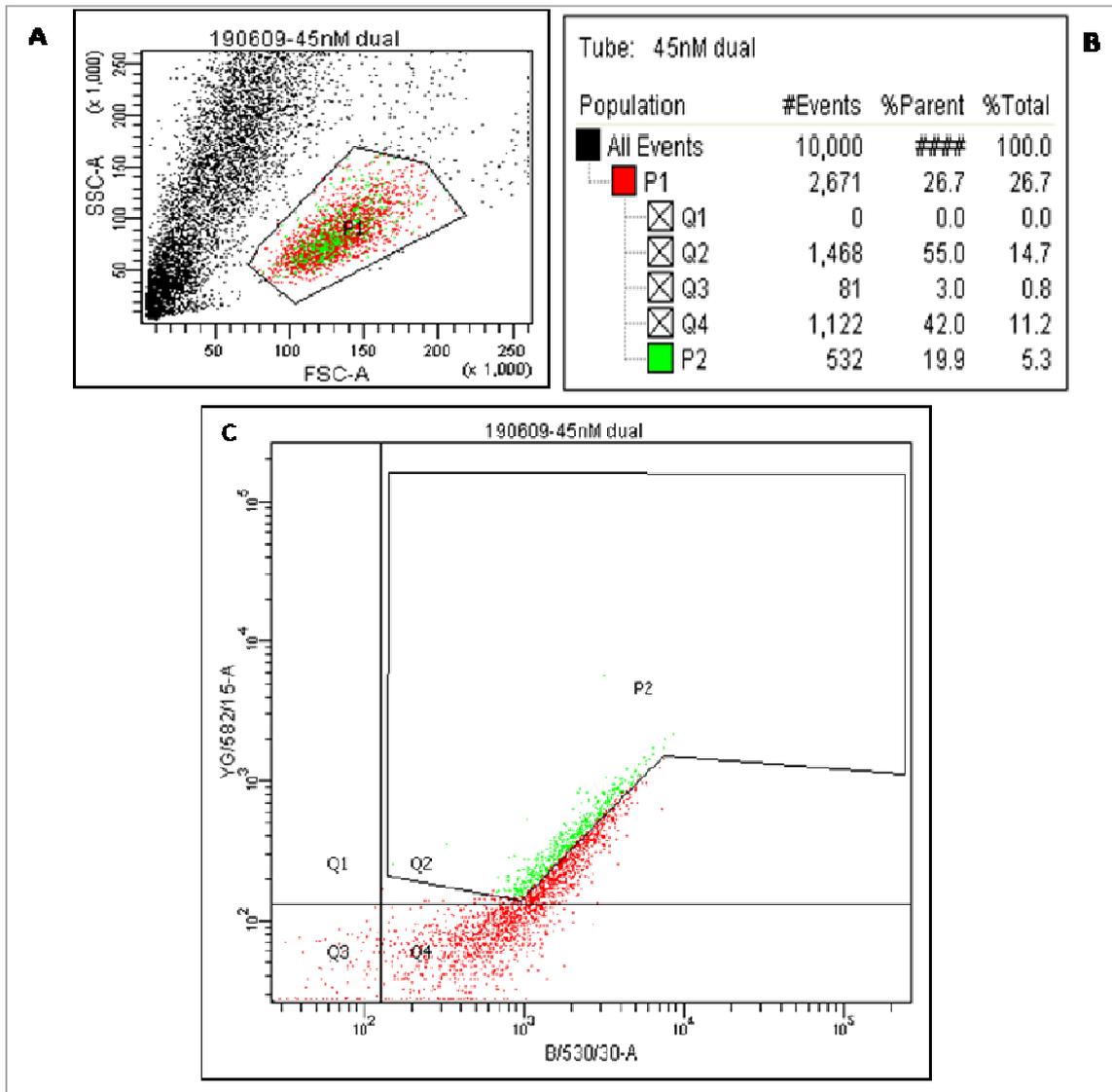
## 6.7 FACS selection

10 million cells of the stable cell population (labelled *Unsorted*) were collected. The cells were then washed with 10% FCS-PBS solution and incubated with soluble Tie2 ectodomain at a concentration of 45nM (non-stringent concentration) for 1hr at 4°C. The primary antibodies anti-FLAG and biotin-linked anti-His were used to detect cell surface (FLAG-tagged) Ang2-FReD and Ang2-FReD bound (His-tagged) Tie2

respectively. Subsequently these cells were fluorescently labelled with Cy2 and R-PE antibodies as described in section 2.23.

Figure 6.15A shows the forward and side scatter plot for the fluorescently labelled stable cell population. The cells contained in polygon P1 are live cells which have different complexity, granularity and size as compared to the dead cells or debris. Only the live cells in polygon P1 were analysed in figure 6.15C.

Figure 6.15C is a dot plot showing the Tie2 ectodomain bound (Y axis) versus the level of FReD expression on each cell (X axis). The cells in the Q3 quadrant are unstained cells. The cells in the Q4 quadrant are green fluorescent cells expressing FReD domain. The cells in the Q2 quadrant are green and red fluorescent cells expressing FReD domain with bound Tie2 receptor ectodomain. Of these fluorescently labelled cells approximately 20% were selected by FACS from the gated live cell population (non-stringent selection to select all binders) indicated in the polygon P2 (Figure 6.15C). This sorted population was labelled *Sort A*. Figure 6.15B shows the statistical information for the 1<sup>st</sup> round of sorting, recorded for a subset of 10000 cells. The total population of 10 million cells were analysed and sorted.



**Figure 6.15: Representative dot plots showing the population of cells sorted in the 1<sup>st</sup> round of FACS for Tie2 binding.**  $10^7$  cells of the stable cell population (labelled *Unsorted*) were incubated with 45nM sTie2. Cy2 and R-PE fluorescent antibodies were used to detect cell surface FLAG-tagged Ang2-FReD and bound His-tagged Tie2. Labelled cells were sorted by FACS and data were represented as dot plots. (A) Forward and side scatter dot plot showing the complexity and size of the cells. Polygon P1 contains the live cell population that were analysed and sorted based on fluorescence. The black dots represent dead cells or cellular debris. (B) The table shows the statistics for 1<sup>st</sup> round of sorting, recorded for a subset of 10000 cells. (C) The level of protein expression for each cell is represented by the Cy2 dye on the X axis. The Tie2 ectodomain bound to the surface Ang2 FReD domain is represented by the R-PE dye on the Y axis. The cells in the Q3 quadrant are unstained cells. The cells in the Q4 quadrant are green fluorescing cells expressing FReD domain. The cells in the Q2 quadrant are green and red fluorescing cells expressing FReD domain with bound Tie2 receptor ectodomain. Polygon P2 drawn in the Q2 quadrant are dual stained cells and select the cells (each dot represents a single cell, shown in green colour) that were collected in the 1<sup>st</sup> round of sorting and labelled *Sort A*.

The cells collected after the 1<sup>st</sup> round of sorting were collected under sterile conditions and grown in culture media. For the 2<sup>nd</sup> sort 10 million *Sort A* cells which were labelled and selected by FACS as described for the 1<sup>st</sup> round of sorting above. This FACS sorting was repeated iteratively using the protocol outlined for the 1<sup>st</sup> round. Table 6.4 summarizes the details of subsequent sort rounds performed.

Sort Round	Population fluorescently stained	Tie2 conc used for sort (nM)	% population collected	Population sorted
0	---	---		<i>Unsorted</i>
1	<i>Unsorted</i>	45	19.9	<i>Sort A</i>
2	<i>Sort A</i>	45	0.1	<i>Sort A2</i>
3	<i>Sort A2</i>	45	0.1	<i>Sort A3</i>
4	<i>Sort A3</i>	10	0.1	<i>Sort A4</i>
5	<i>Sort A4</i>	10	0.2	<i>Sort A5</i>
6	<i>Sort A5</i>	1	0.5	<i>Sort A6b</i>
7	<i>Sort A6b</i>	1	2.1	<i>Sort A7c</i>

**Table 6.4: Description of the iterative rounds of selection, for high affinity binders to Tie2 by Fluorescence Activated Cell Sorting (FACS).** The table indicates the population of cells stained, the Tie2 concentration used for selection, the percentage of the gated population collected and cultured for each round of sorting.

Subsets of cells obtained after each round were grown in culture for comparison with subsequent sorted populations as discussed in section 6.8.

## **6.8 Comparison by flow cytometry of FACS sorted population of cells for high affinity binders to Tie2.**

The populations of cells obtained after each round of sorting in section 6.7 were analysed against populations obtained from previous sorts by comparative flow cytometry analysis. *Sort A* and *Sort A2* populations were compared for binding to 45nM Tie2 concentration. *Sort A2*, *A3* and *A4* were analysed at 10nM Tie2. *Sort A4*, *A5* and *A61b* were analysed at 1nM Tie2. Finally the *Unsorted*, *A5* and *A7c* populations were compared at 0.1nM Tie2.

Briefly,  $2 \times 10^5$  cells of the sorted stable cells population were collected. The cells were washed with 10% FCS-PBS solution and incubated with soluble Tie2 ectodomain at a suitable Tie2 concentration for 1hr at 4°C. The primary antibodies anti-FLAG and biotin-linked anti-His were used to detect cell surface (FLAG-tagged) Ang2-FReD and Ang2-FReD bound (His-tagged) Tie2. Subsequently these cells were fluorescently labelled with Cy2 and R-PE antibodies as described in section 2.22. The Cy2 label (green fluorescence) represents the expression level of the surface expressed Ang2 FReD domain. The R-PE label (red fluorescence) reflects the Tie2 ectodomain bound to the surface Ang2 FReD domain. The cells were then analysed by flow cytometry and geometric mean for each sorted population was obtained. The geometric mean for the green fluorescence (Cy2) represents the expression levels of the Ang2-FReD on the DT40 cells surface. The geometric mean for the red fluorescence(R-PE) represents the level of Tie2 bound to the Ang2-FReD on the DT40 cells surface.

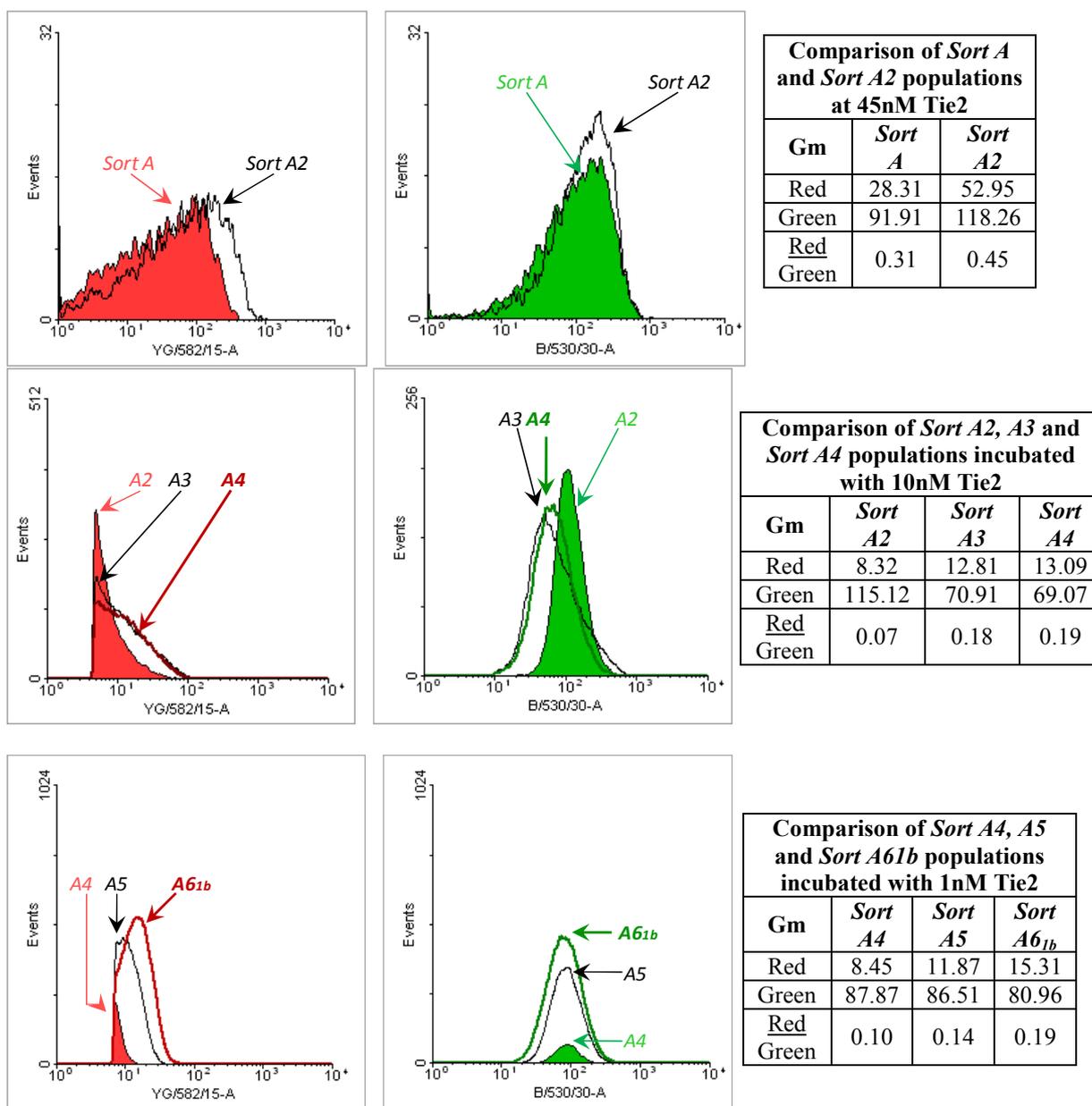
All the comparisons are displayed as histogram overlays for sorted populations compared with previous populations and determined in the same experiment. Histogram

overlays are shown for both green fluorescence (reflecting expression levels) and red fluorescence (reflecting Tie2 binding). Geometric mean (Gm) fluorescence is also determined for each cell population. To account for any effect of expression level when comparing binding between populations, comparisons were made of  $\text{Red}_{\text{Gm}}/\text{Green}_{\text{Gm}}$ .

Figure 6.16A shows a histogram overlay for *Sort A* and *Sort A2* populations. The geometric mean values for red and green fluorescence for each population is presented in the table. As seen from the ratio (red/green geometric mean ratio), an improvement is observed in *Sort A2* (obtained after 2 iterative rounds of sorting). Thus, at 45nM Tie2 a distinct difference between *Sort A* and *Sort A2* populations was observed.

Figure 6.16B is a histogram overlay for *Sort A2*, *A3* and *A4* populations determined in the same experiment at 10nM Tie2. The geometric mean values for red and green fluorescence for each population is presented in the table. Red/Green ratio of 0.18 for *Sort A3* indicates an improvement as compared to *Sort A2* ratio of 0.07. Thus, at 10nM Tie2 a distinct difference between populations was observed.

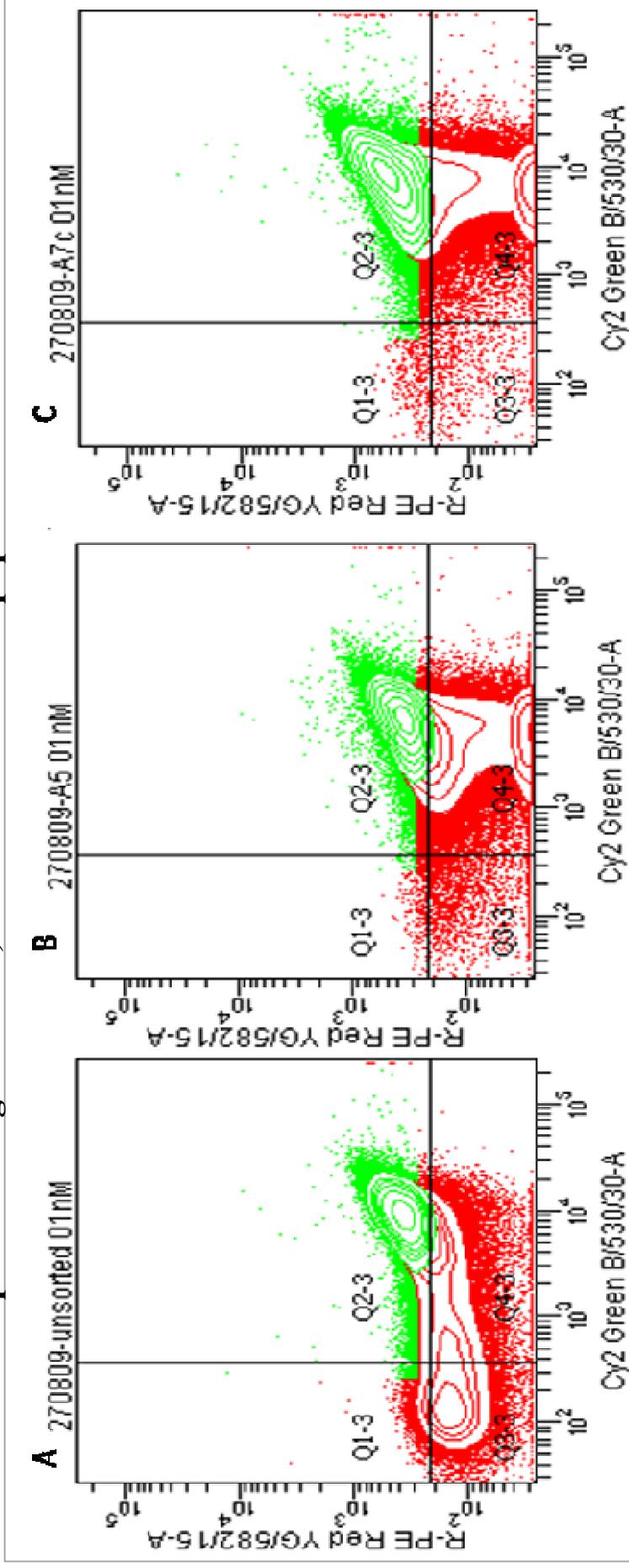
Figure 6.16C is a histogram overlay for *Sort A4*, *A5* and *Sort A61b* populations determined in the same experiment at 1nM Tie2. Red/Green ratio of 0.19 for *Sort A61b* indicates further improvement as compared to *Sort A4* ratio of 0.10. As seen from the ratio an improvement is observed in *Sort A61b* (obtained after 6 iterative rounds of sorting). Thus, at 1nM Tie2 a distinct difference between populations was observed.



**Figure 6.16: Comparison of population of cells sorted by FACS for Tie2 binding.** The sorted populations obtained after each round of sorting in section 5.7 were analysed against populations obtained from previous sorts by comparative flow cytometry analysis. The data was displayed as histogram overlays for red and green fluorescence. The green fluorescence values are used for normalising the expression levels in each population. The red to green ratio is the true reflection of improvement between sorted populations. After normalising the expression levels in each population the Red/Green fluorescence value is obtained. (A) *Sort A* and *Sort A2* populations were compared for binding to 45nM Tie2 concentration. As seen from the Red/Green ratio an improvement is observed in *Sort A*. (B) *Sort A2*, *A3* and *A4* were analysed at 10nM Tie2. Red/Green ratio of 0.18 for *Sort A3* indicates a huge improvement as compared to *Sort A2* ratio of 0.07. (C) *Sort A4*, *A5* and *A61b* were analysed at 1nM Tie2. Red/Green ratio of 0.19 for *Sort A61b* indicates further improvement as compared to *Sort A4* ratio of 0.10.

Figure 6.17 is a dot plot contour display of the flow cytometry data comparing the *Unsorted*, *A5* and *A7c* cells populations stained at 0.1nM Tie2 in the same experiment. As described earlier, cells were stained for normalizing expression using Cy2 fluorophore. Streptavidin-R-Phycoerythrin was used for detecting cell bound Tie2. Population *A7c* obtained after 7 iterative rounds of FACS shows a distinct population of cells emerging (coloured green) when compared with the unsorted population of cells stained at 0.1nM Tie2.

### Dotplot showing *Unsorted*, *SortA5* and *SortA7c* populations at 0.1nM Tie2



**Figure 6.17: Contour plot display of flow cytometry data for *Unsorted*, *A5* and *A7c* cells populations stained at 0.1nM Tie2.** The cell populations *Unsorted*, *A5* and *A7c* were incubated at 0.1nM Tie2 concentration and dual stained using Cy2 and streptavidin-R-phycoerythrin fluorophores. (A) Shows the *Unsorted* population displayed in a contour plot. (B) Shows the *A5* population displayed in a contour plot where a distinct population of cells (coloured green) emerge when compared with the unsorted cells. (C) A distinct population of cells is seen to emerge in cell population *A7c* and the density of cells in this separated population is more when compared with the *A5* or *Unsorted* populations.

Thus, apparent improvement in Tie2 binding is seen with iteratively selected populations. In Figure 6.17A two populations of cells, one which binds and other not binding to Tie2 are observed on the contour plot for *Unsorted* population. In 6.17B an increase in the binding population is observed for *A7c* cells and the density of cells in the binding population is increased as compared to *A5* population. Thus *A7c* population was analysed by sequencing.

### **Sequencing of Sorted cells**

Cells shown in quadrant Q2-3 (Figure 6.17C) were sorted by FACS and collected into PCR tubes. Messenger RNA from these cells was transcribed to cDNA by RT-PCR as described in section 2.9. Ang2-FReD DNA was amplified by PCR as described in section 2.1.2. This DNA was cloned into a TA sequencing vector (pCR 4-TOPO, Invitrogen) and 20 clones sequenced. No mutants were obtained, with all the sequenced clones containing WT ang2-FReD DNA sequence.

These data demonstrate that the variation in binding in the sorted cell population could not be explained by affinity improvements due to mutations in Ang2-FReD, since all the cells in this population were wild type Ang2-FReD expressers. It was possible that the lack of mutants was possibly due to integration of Ang2-FReD domain at a position in the DT40 genome which does not support adequate action of AID enzyme for SHM. It was also possible that the sorting conditions used were not adequate and the Tie2 concentration was too high resulting in enrichment of non-specific binders, through 7 rounds of FACS.

In order to confirm that the data described above are representative of this approach, a new transfection was performed and single stable clones were reanalyzed

for improving binding. The cells were regularly checked for AID expression and cultured in MPA. The sort rounds were repeated, using a lower concentration of Tie2 at 6nM and analysing cells with fewer rounds of FACS. However, no mutants were obtained.

### 6.9 Characterisation of mutants *10c*, *FACSD* and *FACSi* obtained by SHM.

Iterative selection using FACS failed to generate mutant forms of Ang2-FReD with altered binding affinity. However, mutant Ang2-FReD forms were obtained from sequencing clones from the stable population. Table 6.5 lists the mutant clones obtained by SHM.

Construct Name	Protein encoded for by the construct	Population of cells from which the mutant was obtained	Used for
<i>10c mutant</i>	Ang-2 FReD mutant Pro303Arg	After one round of magnetic sorting.	ASGPR anchored FLAG-tagged Ang-2 FReD mutant protein displayed on the cell surface. Binding to Tie2 receptor ectodomain to determine $K_d$ in mammalian HEK cells
<i>FACSD mutant</i>	Ang-2 FReD double mutant Asn295Asp, Asp495Asn	After one round of magnetic sorting grown in culture and sequenced prior to FACS analysis.	
<i>FACSi mutant</i>	Ang-2 FReD mutant Gly411Asp		

**Table 6.5: List of mutant constructs used in Chapter 6.** *10c*, *FACSD* and *FACSi* mutants were generated by site directed mutagenesis in the ASGPR-pcDNA3.1/V5-His-TOPO vector. This vector allows expression in mammalian HEK293 cell line under the influence of CMV promoter and the ASGPR transmembrane domain anchors the FLAG-tagged mutant protein on the extracellular surface.

As these mutants were isolated after one round of magnetic sorting, it is possible they exhibit some increased binding to Tie2. This was tested by measuring  $K_d$  of each mutant expressed on the surface of HEK cells and compared with the WT Ang2-FReD.

In addition, characterisation of these mutants could provide some information about the effect of the SHM driven mutation at the residue positions.

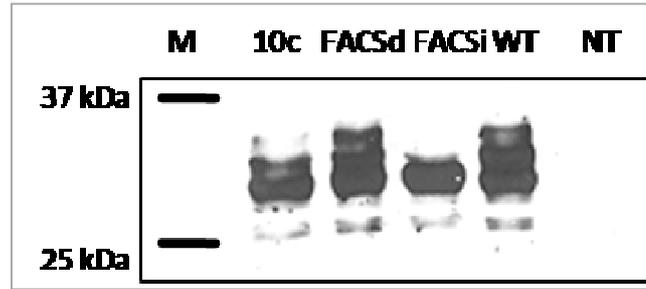
To do this, the mutants were generated by site directed mutagenesis and cloned into a mammalian surface expression vector allowing expression of the mutated FReD domain on the cell surface of mammalian cells (such as HEK293) anchored by the ASGPR transmembrane domain under the influence of a CMV promoter.

Plasmids were sequenced (Sequence data Appendix 4L, 4M, 4N) to confirm the required mutation had been incorporated in each of the mutant plasmids. These mutants were characterised by Western blotting and by determination of dissociation constants.

#### **6.9.1 Western blot analysis for mutants *10c*, *FACSD*, *FACSi* in HEK293 cell line.**

In order to confirm mutant form of Ang2-FReD were expressed in the cells, the *10c*, *FACSD*, *FACSi* mutant plasmids were transfected in HEK293 cells and the cell lysate analyzed by Western blotting (Figure 6.18).

The FLAG-tagged mutant and WT Ang2-FReD protein was detected by probing the nitrocellulose membrane with anti-FLAG primary and anti-mouse HRP linked secondary antibody as described in section 2.19.1. The expected size of the ASGPR linked FLAG-tagged WT Ang2-FReD domain protein is ~29.5 kDa.



**Figure 6.18: Western blot analysis for mutants, *10c*, *FACSd*, *FACSi* in HEK293.** HEK cells were transfected with the Wild type Ang2-FReD and mutant *10c*, *FACSd*, *FACSi* plasmids. Cell lysate was analyzed post-transfection on a Western blot. The correct size for ASGPR linked FLAG-tagged WT Ang2-FReD is ~29.5 kDa. Bands between 25 to 37 kDa were observed in lane 2, 3, 4 and 5. Except for *FACSi* mutant, all other lanes showed up to four bands between 25 and 37 kDa. However the expression of the four bands observed varied between the mutants *10c*, *FACSd* and WT protein. It is possible *FACSi* mutant protein is expressed in the cell but does not fold properly on the cell surface. No bands were observed in the non-transfected (NT) lane.

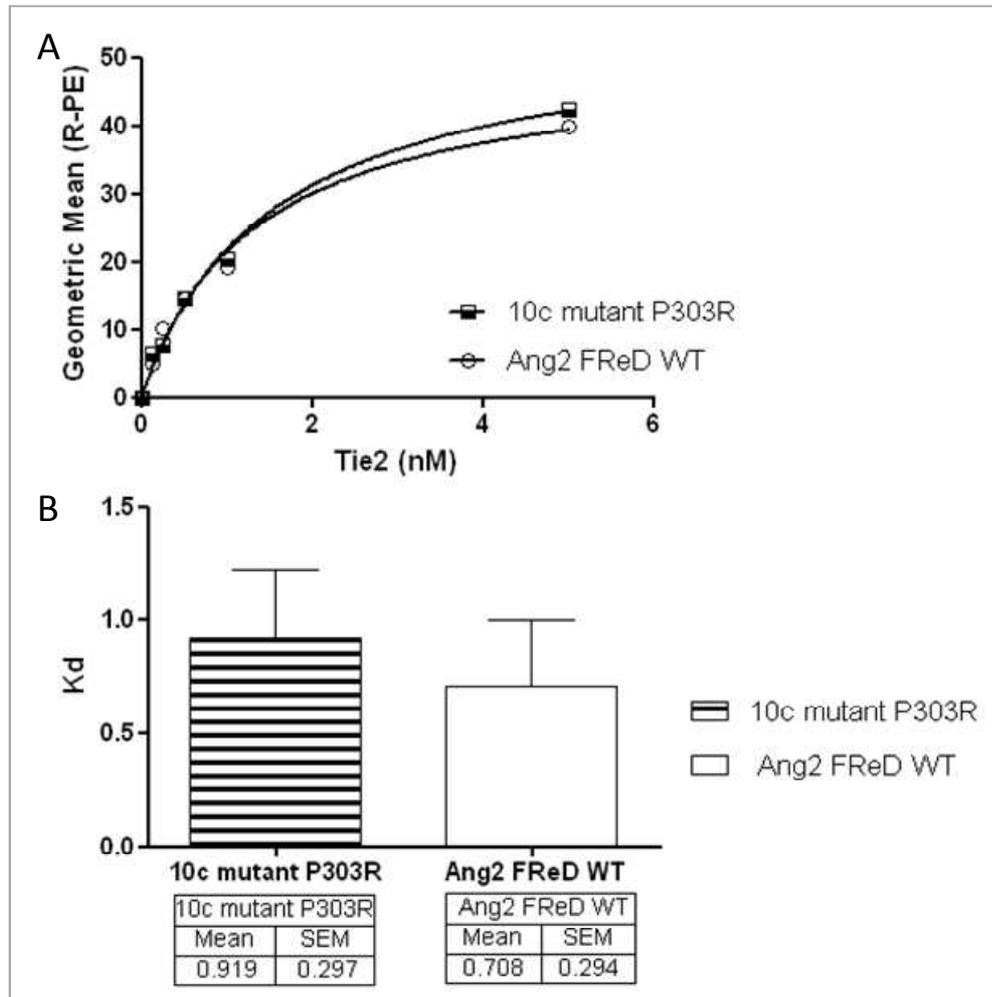
For *10c*, *FACSd* and WT forms of Ang2-FReD proteins four different forms of each protein were detected. The four bands may indicate presence of degraded, cleaved, differentially glycosylated and active cell surface expressed forms of the protein. *FACSi* mutant protein expressed only the lowest molecular mass forms of the protein.

### **6.9.2 *K<sub>d</sub>* determination of mutants *10c*, *FACSd*, *FACSi* in HEK293 cell line by flow cytometry.**

The mutants *10c*, *FACSd*, *FACSi* were characterised for their phenotype by *K<sub>d</sub>* determination. The mutant plasmids *10c*, *FACSd*, *FACSi* and the WT plasmid were transfected into HEK cells and 24hours post-transfection cells were collected. To obtain a concentration curve for binding of the mutants and WT protein to the soluble *rhTie2* ectodomain the transfected cells were incubated with a range of soluble *rhTie2* concentrations. These were then washed to remove unbound Tie2 and incubated with biotin-linked anti-His primary antibody to detect the His-tagged Tie2 bound on the cell surface of mutant and the WT Ang2-FReD domain. The cells were washed and then

incubated using fluorescently labelled Streptavidin-R-Phycoerythrin. The red fluorescence of the phycoerythrin fluorophore was detected on the flow cytometer for each of the Tie2 concentrations for the mutant and the WT Ang2-FReD domain. The geometric mean for the phycoerythrin fluorescence (red) for each Tie2 concentration was obtained. To determine  $K_d$  values non-linear regression analysis was performed of the geometric mean fluorescence value versus concentration of *rhTie2* Fc chimera using Graphpad Prism. The  $K_d$  for each mutant was determined by three independent experiments.

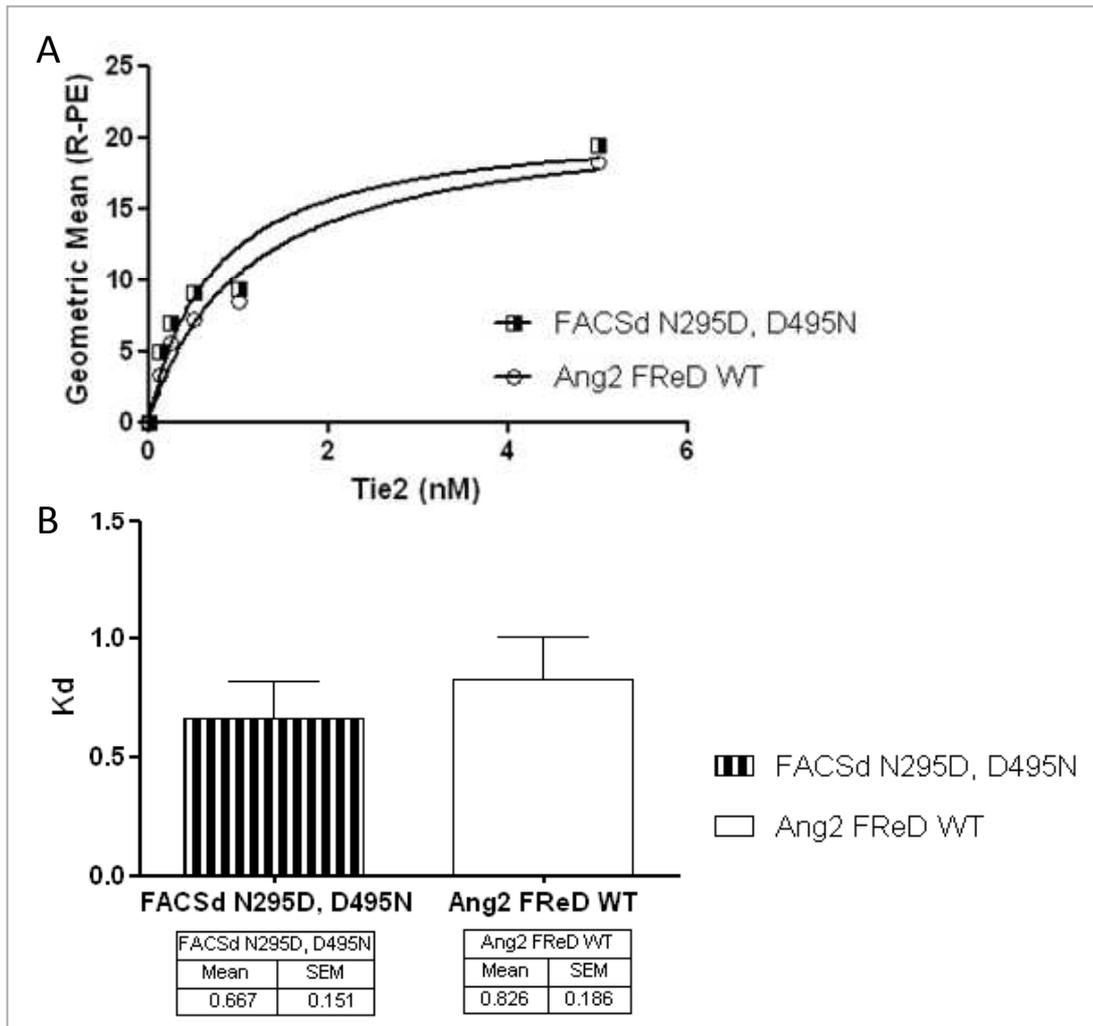
Figure 6.19 is a representative concentration curve obtained for the *10c* mutant and WT Ang2-FReD. The bar graph shows the mean  $K_d$  value of mutant *10c* and WT domain for three independent experiments.



**Figure 6.19: Representative Tie2 concentration curve for WT Ang2-FReD domain and mutant *10c* in HEK293 cell line.** HEK cells were transfected with the WT and mutant plasmids and incubated with varying concentrations of soluble Tie2 ectodomain. The His-tagged Tie2 bound on the cell surface of *10c* FReD mutant domain and the WT FReD domain was detected using biotin-linked anti-His primary antibody and Streptavidin-R-Phycoerythrin. The geometric mean for red fluorescence for each Tie2 concentration was obtained. To determine  $K_d$  values non-linear regression analysis was performed of the geometric mean fluorescence value (Y axis) versus concentration of Tie2 (X axis) using Graphpad Prism. (A) Representative concentration curve obtained for the *10c* mutant and WT Ang2-FReD. (B) Bar graph shows the mean  $K_d$  value of mutant *10c* of 0.92nM with SEM of +/- 0.297 and mean  $K_d$  value of WT domain was 0.71nM with SEM of +/- 0.294. (n=3).  $K_d$  value for *10c* mutant were found to be similar to that of WT Ang2-FReD domain.

The  $K_d$  value of the mutant *10c* was obtained and found to be similar to that of WT Ang2-FReD domain. This experiment was performed three times and a mean  $K_d$  of 0.92nM with a SEM of +/- 0.297 obtained for *10c*. The mean  $K_d$  of 0.71nM with a SEM of +/- 0.294 was obtained for WT Ang2 FReD.

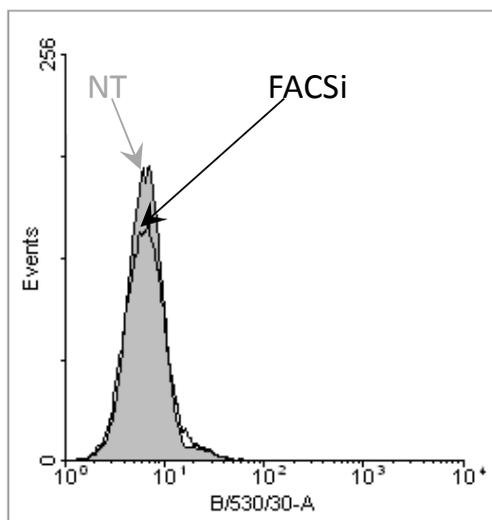
Figure 6.20 is a representative concentration curve obtained for the *FACSD* mutant and WT Ang2-FReD. The bar graph shows the mean  $K_d$  value of mutant *FACSD* and WT domain with error bars.



**Figure 6.20: Representative Tie2 concentration curve for WT Ang2-FReD domain and mutant *FACSD* in HEK293 cell line.** HEK cells were transfected with the WT and mutant plasmids and incubated with varying concentrations of soluble Tie2 ectodomain. The His-tagged Tie2 bound on the cell surface of *FACSD* FReD mutant domain and the WT FReD domain which were detected using biotin-linked anti-His primary antibody and Streptavidin-R-Phycoerythrin. The geometric mean for red fluorescence for each Tie2 concentration was obtained. To determine  $K_d$  values non-linear regression analysis was performed of the geometric mean fluorescence value (Y axis) versus concentration of Tie2 (X axis) using Graphpad Prism. (A) Representative concentration curve obtained for the *FACSD* mutant and WT Ang2-FReD. (B) Bar graph shows the mean  $K_d$  value of mutant *FACSD* of 0.67nM with SEM of +/- 0.151 and mean  $K_d$  value of WT domain was 0.83nM with SEM of +/- 0.186. (n=3).  $K_d$  value for *FACSD* mutant were found to be similar to that of WT Ang2-FReD domain.

The  $K_d$  value of the mutant *FACSD* was obtained and found to be similar to that of WT Ang2-FReD domain. This experiment was performed three times and a mean  $K_d$  of 0.67nM with a SEM of +/- 0.151 was obtained for *FACSD*. The mean  $K_d$  of 0.83nM with a SEM of +/-0.186 was obtained for WT Ang2FReD.

The  $K_d$  value for the *FACSi* mutant could not be determined in three independent experiments, as cells transfected with the *FACSi* mutant displayed no Tie2 binding even at the highest concentration in the concentration curve (data not shown). Hence the *FACSi* transfected cells were labelled with anti-FLAG and Cy2 antibodies to determine level of expression of the mutant protein on the cell surface. Figure 6.21 is a histogram overlay of the *FACSi* transfected cells and the non-transfected cells labelled with the Cy2 fluorophore.



**Figure 6.21: Histogram overlay of the *FACSi* transfected cells and the non-transfected HEK293 cells.** The *FACSi* transfected cells (coloured black line) showed no green fluorescence when compared with the non-transfected cells (coloured grey fill). No expression for mutant *FACSi* was detected by flow cytometry in HEK cells.

As observed from the overlay, the *FACSi* transfected cells showed no green fluorescence when compared with the non-transfected cells. Thus, the *FACSi* mutant protein was not detectable at the cell surface, possibly due to misfolding of the protein

which could lead to cleavage and degradation or inability to translocate to the cell surface or inability to form the appropriate secondary structure on the cell surface. Expression or binding of the *FACSi* mutant protein was not detected by flow cytometry.

## **Discussion**

The aim of this chapter was to establish a cell surface display system for FReD protein expression in DT40 cells in which, a single copy of the DNA of interest was targeted to the Ig locus of the bursal genome where it could be subjected to somatic hypermutation (SHM). Eventually the variant proteins expressed on the DT40 extracellular surface could be sorted for Tie2 binding affinity by iterative rounds of Fluorescence Activated Cell Sorting (FACS).

The Ang1 and Ang2 FReD domain linked to the ASGPR transmembrane domain was successfully cloned into the pHypermut2 vector. As demonstrated previously in Chapter 5, the ASGPR transmembrane domain was fused to the protein of interest to orient the carboxy-terminal (Tie2 binding region) of the FReD domain freely on the extracellular surface of the cell. The cell surface expressed protein could be studied by *in vitro* binding assays and detected using fluorescent labels by immunofluorescence and flow cytometry. The vector was designed to facilitate the targeting of the FReD domain DNA sequence into the Ig locus of DT40 genomic DNA and was used to generate stable clones post transfection using puromycin antibiotic selective pressure.

As observed previously, the expression of the Ang1 FReD domain in the stable DT40 cells showed a trend similar to that observed in CHO, HEK293 and Ramos cells (Chapter 5). Only a low molecular mass band was observed by Western blotting, indicating that the expression of the active cell surface form of the ASGPR linked Ang1

FReD protein is absent in the DT40 cells and is probably due to the inherent nature of the protein sequence. Hence further studies for establishing SHM mediated directed evolution was continued with only the ASGPR linked Ang2 FReD protein. The Ang2 FReD protein expressed successfully in the DT40 stable cells, and both the unmodified polypeptide and higher molecular weight band of active protein was detected by Western blotting. The FLAG-tagged ASGPR linked Ang2 FReD protein was localised on the extracellular cell surface of the stable DT40 cells by immunofluorescent labelling of the FLAG epitope tag. The binding of the cell surface expressed ASGPR linked Ang2 FReD protein to His-tagged Tie2 ectodomain was also demonstrated by immunofluorescent labelling.

The mutation frequency of the protein in the stable DT40 cells was determined and ranged from 0.1 to 0.2 within 25 randomly sequenced clones. A mutation frequency of 0.1-0.2 would produce 1-2 million mutants in the SHM generated library of  $10^7$  cells which were analysed for each iterative round of FACS. This frequency is comparable to the ~0.3 frequency of mutation obtained for evolution of GFP protein in DT40 cells (Arakawa et al. 2008). It has been observed that the first ~150bp of the gene are infrequently mutated (Longerich et al. 2005). Since the ASGPR transmembrane domain constitutes the first ~150bp of the highly transcribed fusion transgene, it is likely to be spared SHM. The DGYW/WRCH (G:C is the mutable position; D=A/G/T, H=T/C/A) has been suggested as the mutable motif for AID (Rogozin, Diaz 2004). 'A' hotspots are targeted when preceded by 'T' (Goyenechea, Milstein 1996). Thus, the frequency of mutation obtained by SHM depends not only on the host B cell specific mutagenesis machinery, but also the DNA sequence subjected to mutagenesis.

The Ang2 FReD protein mutant library generated by SHM was sorted by magnetic selection over three magnetic sort rounds. The cells selected, displayed an apparent improvement in binding affinity to Tie2. However, the degree of improvement within the sorted population was very low and a large number of wild-type protein expressing cells continued to be selected. Magnetic sorting allowed labelling and selection based on one parameter, i.e. the amount of Tie2 bound to the cell population. It was thought that the affinity improvement by magnetic sorting was limited due to the varying expression level of the FLAG-tagged FReD protein which could not be normalised for by single parameter labelling and selection. For this reason a more sensitive approach of FACS was used subsequently.

FACS selection allowed dual labelling of cells and expression of the protein of interest could be traced simultaneously while selecting for improved binding affinity. Expression normalization helps to account for cell to cell variability of cell surface amounts of the fusion protein. Prior to FACS selection, the dissociation constant for Tie2 binding of the WT-Ang2 FReD protein in the DT40 cell surface display system was determined. Although several affinity studies have been performed with the angiopoietins (Discussed in Chapter 1, table 1.1), the absolute binding affinity value of a protein determined by different approaches (immobilization ELISA, SPR) and different forms of the protein (fusion protein, truncated fragments, secreted/membrane bound) often varies (Kronqvist et al. 2008, Kieke et al. 1997). The determined  $K_d$  value of ~30nM for the WT protein gave an approximate concentration range for performing iterative rounds of FACS selection.

In the first round of FACS sorting, a Tie2 concentration 1.5 times the  $K_d$  was used and sort was performed to collect cells non-stringently by FACS. This first round was

performed to collect all variant FReD proteins which bound to Tie2. A concentration higher than  $K_d$  would ensure collection of weaker binders of Tie2. As discussed in Chapter 1, the approach of directed evolution of a protein is dynamic and not unidirectional with respect to the mutations accumulated. Some mutations may lead to some degree of loss in function initially, but in combination with other assisting mutations lead to enhancement of function in further rounds (Romero, Arnold 2009). Thus selection for binding at a concentration higher than  $K_d$  was performed to retain any such mutants. The sort window was big to collect ~20% cells and FACS sorter settings were non-stringent for the first round. In the next two rounds, although the concentration of Tie2 was retained, the percentage of cells collected was dropped selecting only the brightest 0.1-2% cells. For subsequent sort rounds, the concentration of Tie2 used for selection was also decreased until no further improvement at a given concentration was observed. Seven iterative rounds of selection were performed and the final enriched population was obtained at 1nM Tie2 concentration. However, sequencing of this population revealed that the cells selected were expressing WT form of the Ang2 FReD protein. A new transfection, analyses of single clones, iterative sorting at decreased Tie2 concentration yielded similar results; no mutant clones. Since the selection of sorted population of cells was Tie2 concentration dependent, it was thought that some other factor contributed to the selection of apparent improved binder cells. It should be noted expression level of the Ang2 FReD protein was taken into account in all the selections. However, the cell to cell variability in distribution of the expressed FReD protein on the cells surface cannot be determined by the FLAG tag. Clusters of FLAG-tagged FReD protein may expose itself poorly to the Anti-FLAG antibody, resulting in a false estimation of expression level of the protein. So although the cell populations showed improved binding, it is possible that the iterative rounds of

sorting continuously selected for WT higher expresser cells contributing to improved higher affinity to Tie2.

In the final section of this chapter, the SHM mediated mutants (*10c*, *FACSD*, *FACSi*) obtained from a population magnetically sorted once, were analysed in HEK 293 cell line by Western blotting and their determined  $K_d$  compared to the WT  $K_d$  value. Since the mutants *10c*, *FACSD* and *FACSi* obtained by SHM had not been subjected to iterative rounds of selection, they were expected to show no improvement in binding affinity. However, the analyses of the mutants gave some information on the effect of mutations at their respective residue positions.

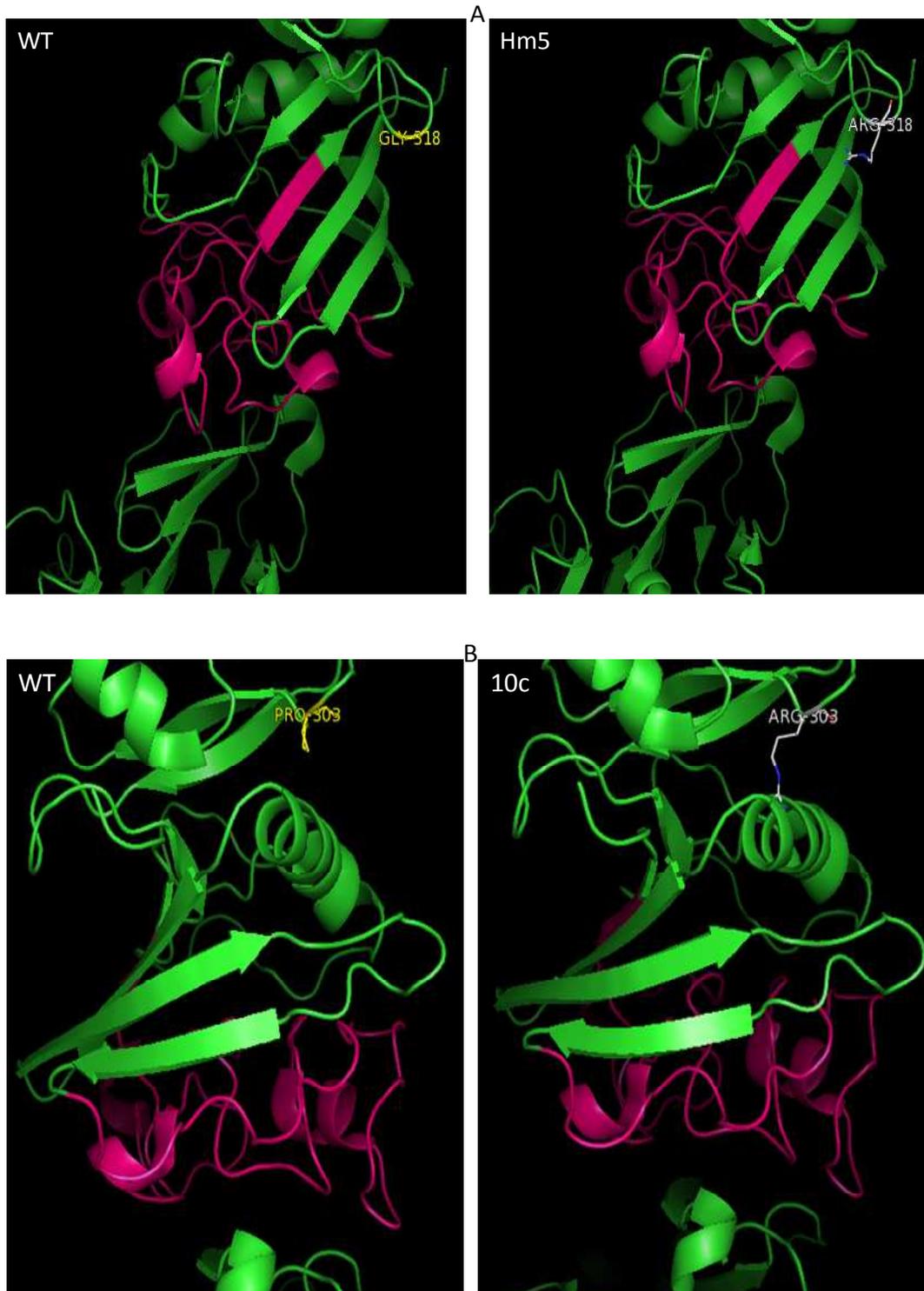
Table 6.6 is a summary of these mutants; *10c*, *FACSD*, *FACSi*. The table also describes mutant *Hm5* (discussed in section 6.4, appendix 4K) which was obtained from a random population of cells not subjected to any sort process, neither FACS nor magnetic sorting.

Mutant Name	Mutated Residue Number	Wild Type Ang2 amino acid	Mutated amino acid	Wild Type Ang1 amino acid
<i>Hm5</i>	318 in FReD domain	Glycine	Arginine	Asparagine
<i>10c</i>	303 in FReD domain	Proline	Arginine	Asparagine
<i>FACSD</i> (Double mutant)	295 in FReD domain	Asparagine	Aspartic acid	Serine
	495 after P domain	Aspartic acid	Asparagine	Aspartic acid
<i>FACSi</i>	411 in P domain	Glycine	Aspartic acid	Glycine

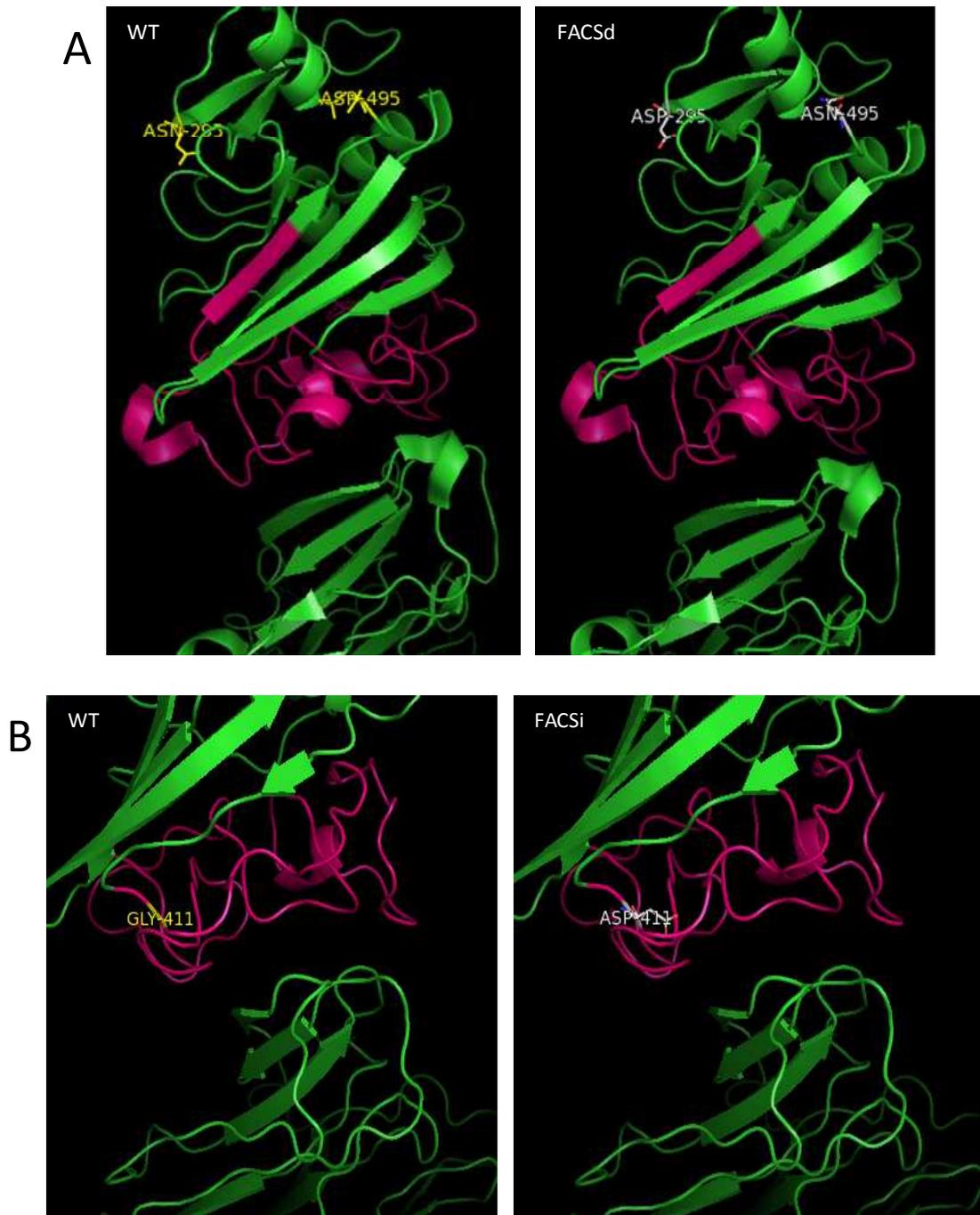
**Table 6.6: Summary the mutants (*Hm5*, *10c*, *FACSD*, *FACSi*) obtained by AID induced SHM of Ang2-FReD domain.** Description of the mutants, indicating the mutated residues and the codon changes in each mutant. Also the corresponding residues in the WT Ang1 at the mutated positions are indicated.

The mutants were visualised in 3-D using PyMOL molecular imaging software (DeLano 2002), which provides a visual representation of the position of the mutated residues with respect to the entire complex structure. Figure 6.22 shows the molecular visualisation of *Hm5* and *10c* mutants. Figure 6.23 shows the molecular visualisation of *FACSD* and *FACSi* mutants.

The *Hm5* and *10c* mutations (Gly318Arg, Pro303Arg) are in residue positions not conserved between Ang1 and Ang2 and do not lie in the P domain of the protein. Mutagenesis of residues in the RBD but outside the P domain is not expected to affect the binding ability of the Ang2 FReD protein (Barton, Tzvetkova & Nikolov 2005). As expected the binding affinity of *10c* mutant to Tie2 was similar to the WT Ang2 FReD protein. As observed from the 3-D structure both mutant residues are not on any beta sheet or alpha helix structure but on the connecting loops between, exposed to the aqueous environment on the outer surface (Figure 6.22). Hence mutation to a large residue like arginine which prefers to interact with the aqueous solvent has not hindered the folding of the *10c* protein nor affected the binding affinity.



**Figure 6.22: Molecular visualization of the *Hm5* and *10c* mutant in PyMOL.** (A) Ribbon structural representation of WT and *Hm5* (Gly318Arg) Ang2-Tie2 complex showing the relative position of the mutated amino acid with respect to the P domain (coloured pink). (B) Ribbon structural representation of WT and *10c* (Pro303Arg) Ang2-Tie2 complex showing the relative position of the mutated amino acid with respect to the P domain (coloured pink).



**Figure 6.23: Molecular visualization of the *FACSd* and *FACSi* mutant in PyMOL.** (A) Ribbon structural representation of WT and *FACSd* double mutant (Asn295Asp, Asp495Asn) Ang2-Tie2 complex showing the relative close proximity of the mutated amino acids with respect to each other. (B) Ribbon structural representation of WT and *FACSi* (Gly411Asp) Ang2-Tie2 complex showing the relative position of the mutated amino acid located within the P domain (coloured pink). The mutation is in the  $\beta 6$ - $\alpha 5$  loop on the binding interface which is key to the Tie2-Ang2 complex formation.

*FACSD* double mutant has a mutation Asn295Asp lying outside the P domain at a residue position not conserved with Ang1 and hence is unlikely to affect binding. The second mutation Asp495Asn is at the far end at the C terminus of the RBD (Figure 6.23). It is in close proximity to residue 295 and farther away from the binding interface at the P domain (specifically the alpha-6 chain) than expected. Furthermore, the mutated residues aspartic acid and asparagine are amino acids with similar polar nature and structure. As expected the binding affinity of the *FACSD* double mutant to Tie2 was similar to the WT Ang2 FReD protein.

The *FACSi* mutant could not be detected on the cell surface and probably failed to fold properly on the cell surface. *FACSi* mutation Gly411Asp is in the P domain just after the  $\beta 6$ - $\alpha 5$  structure which is crucial to the Tie2-Ang2 complex formation. Also the residue 411 is conserved between Ang1, Ang2, Ang3, Ang4 and the two angiopoietin-like protein 1 and 2 (Angptl1, Angptl2), suggesting it may be important for binding or formation of properly folded structure. Glycine is a small amino acid with an H as its side chain as opposed to C and this could contribute flexibility (Betts, Russell 2003). Although the P domain is most divergent sequence stretch between the angiopoietins, it contains 10 glycine residues of which 8 are conserved between the two angiopoietins and hence, may play a crucial role in stabilizing the structure.

In the next chapter 7, iterative sorting was performed to select for Tie1. There is no known binding observed between Tie1 and any angiopoietins, hence any cells sorted against Tie1 would be for selection of new function for the Ang2 FReD protein and the expression of the protein would not hold relevance. Any other factors which may influence selection are also discussed in the next chapter.

## **Chapter 7: Directed Evolution of Ang2- FReD domain by Somatic Hypermutation to Obtain Tie1 Binding Function**

### **Introduction**

Tie1, like Tie2 is a type 1 transmembrane protein receptor tyrosine kinase (RTK) of the Tie family. The *in vivo* ligands for Tie1 are yet to be identified (Maisonpierre et al. 1997). The extracellular region for both Ties' has N-linked glycosylation sites and a hydrophobic site (Sato et al. 1993). The amino acid sequence identity between Tie1 and Tie2 is 76%, with the intracellular domain primarily contributing the sequence identity (Fiedler et al. 2003) and 33% similarity with the extracellular domain (Schnurch, Risau 1993). Although the crystal structure of Tie1 has not been elucidated, homology modelling using the experimentally determined 2.5 Å Tie2 structure (Barton et al. 2006) and the program MODELER, suggests a structure of Tie1 receptor (Seegar et al. 2010). The comparison of hydrophobic surface features of Tie1 and Tie2, although very similar overall, reveal two patches of exposed hydrophobic residues present at the tip of Tie2, but are absent in the equivalent Tie1 region (Seegar et al. 2010). These hydrophobic patches absent in the Tie1 (modelled structure) overlap with the Ang2 (ligand) binding domain of Tie2 as observed in the Ang2-Tie2 complex crystal structure. This may be a possible structural explanation for the distinct ligand-binding properties of the two Tie receptors (Seegar et al. 2010).

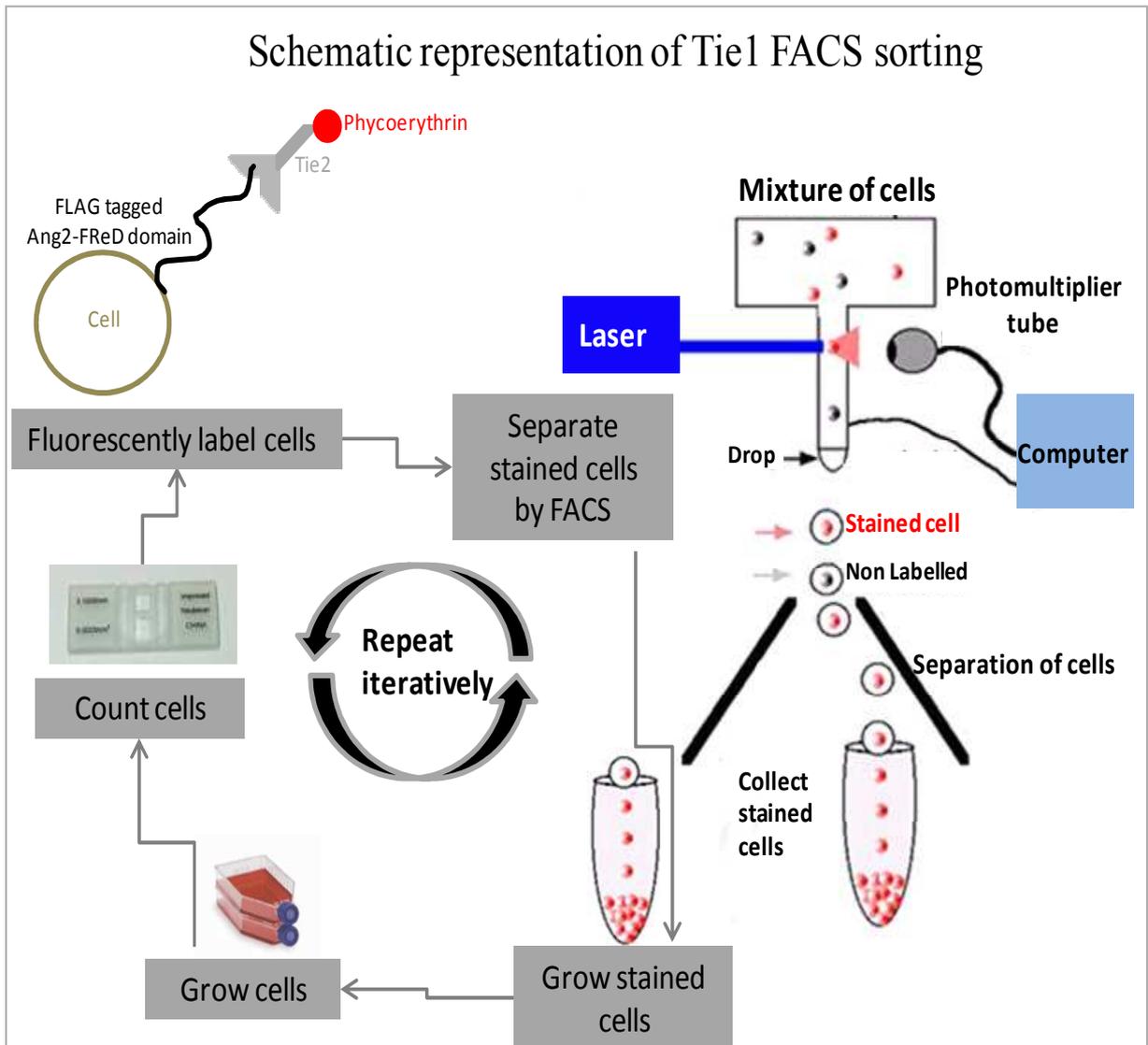
Attempts to evolve a variant of Ang2 with increased affinity for Tie2 using somatic hypermutation were not successful (Chapter 6). It is possible that the options for improving the affinity between Ang2 and Tie2 are limited, and the wild-type Ang2 already has optimal or near optimal structure for Tie2 binding. Tie1 is not considered a receptor for Ang1 or Ang2 and there have been no reports of binding between Ang and

Tie1. However, COMP-Ang1 has been shown to activate Tie1 in cells over-expressing Tie1 but lacking Tie2. Native Ang1 and Ang4 also have been shown to weakly phosphorylate Tie1 (Saharinen et al. 2005). It has been suggested that this may be because over-expressed membrane associated Tie1 has a very weak affinity to the Ang1 receptor-binding domain and Ang1 may induce a conformational change in Tie1, thereby enhancing its affinity for Ang1, which could not be mimicked by the soluble Tie1 extracellular domain (Saharinen et al. 2005).

It is possible, therefore, that there is greater scope for modifying Ang2:Tie1 binding by mutagenesis of Ang2. This is tested in the studies described in this chapter.

### **7.1 Somatic Hypermutation (SHM) and Fluorescence Activated Cell Sorting (FACS) to evolve Ang2-FReD variants with Tie1 binding function.**

The stable cells DT40 cells characterized for expression of Ang2-FReD domain in Chapter 6, section 6.2 were screened and selected by FACS. The stable cells were analysed by FACS to select cells expressing Ang2-FReD that bind to Tie1. Figure 7.1 shows a schematic summary of the iterative rounds of FACS for selection of Tie1 binders. The population of cells is a mixture of non mutated FReD and new variants of FReD with none or some affinity to Tie1. These cell surface expressed variants of Ang2 FReD can be analyzed for their ability to bind soluble Tie1 receptor ectodomain.

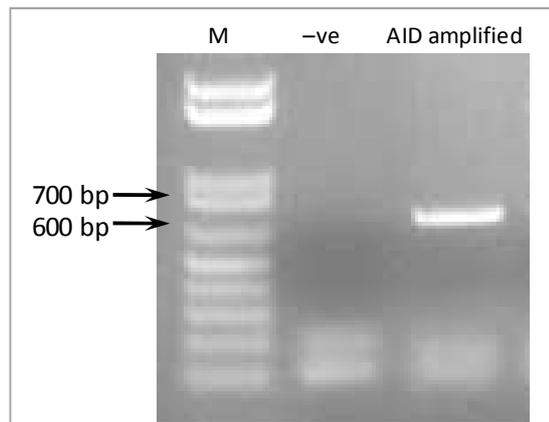


**Figure 7.1: Schematic representation of the iterative rounds of FACS for Tie1 binding.** A typical sort would involve growing Ang2FReD expressing stable DT40 cells in culture, counting and collecting 10 million cells. Fluorescently labelling the cells for the expressed protein and the protein bound Tie2 receptor with R-PE (red fluorescence) dyes. Analysing the labelled cells by FACS and sorting a suitable selected population of cells under sterile conditions. Growing up the selected cells and repeating the round of labelling and selection iteratively.

### 7.1.1 AID expression in the ASGPR-Ang2FReD stable cell population.

Prior to any selection rounds the Ang2-FReD domain expressing stable cells were tested for AID expression (as described previously in section 6.6.1). Briefly, the stable Ang2-FReD expressing cells were collected and RNA extracted from them. Messenger RNA was transcribed into cDNA by RT-PCR as described in section 2.9. The cDNA

was used for amplification of chick AID using 5' AIDchickfwd and 3' AIDchickrev at annealing temperature 55°C. Figure 7.2 shows 5µl of PCR product visualized on a 1% agarose gel, and a product of approximately 600bp (expected size 597bp) was obtained (lane3), indicating expression of AID in the cells. Lane 2 is a negative PCR control and does not show any band. The stable cell population was found to be expressing AID (Figure 7.2) and was used for all subsequent analysis.



**Figure 7.2: Detection of AID mRNA in Ang2-FReD expressing stable DT40 cells prior to selection for Tie1 binding.** RNA was prepared from the Ang2-FReD expressing stable cells. mRNA was transcribed to cDNA by RT-PCR. AID was amplified from the cDNA by PCR. The size of the AID transcribed segment is expected to be 597bp. Lane 3 shows an amplified product at approximately 600bp indicating expression of AID in the stable cells. Lane 2 is a negative PCR control and does not show any band.

For prolonged culture cells were also maintained in 0.5µg/ml mycophenolic acid (MPA) to decrease the chance event of AID excision due to background Cre recombinase expression in the DT40 cell line.

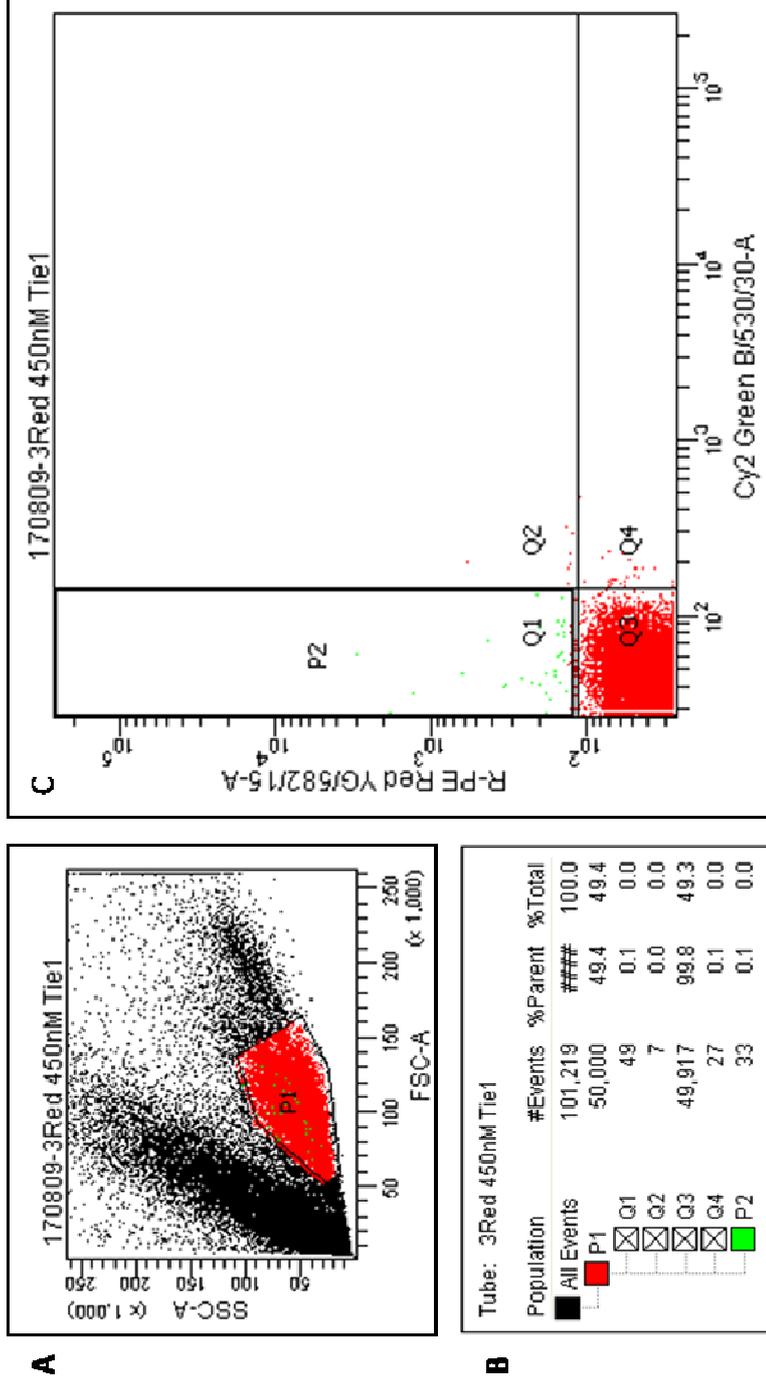
### 7.1.2 FACS sorting for selecting Tie1 binders.

Since Tie1 is not known to bind to any of the angiopoietin ligands a high Tie1 concentration of 450nM was used for 4 iterative rounds of FACS selection (as described in section 2.23). Briefly,  $10^7$  cells of the stable cell population (labelled *Unsorted*) were collected. They were washed with 10% FCS-PBS solution and incubated with soluble

Tie1 ectodomain at a concentration of 450nM (non-stringent concentration) for 1hr at 4°C. The primary antibody biotin linked anti-His was used to detect Ang2-FReD bound (His-tagged) Tie1. Subsequently these cells were fluorescently labelled with R-PE antibody. The R-PE label (red fluorescence) reflects the Tie1 ectodomain bound to the surface Ang2 FReD domain. Cy2 label (used previously) to detect the expression level of the surface expressed Ang2 FReD domain was not used, since selection was not for high affinity binders (which may be influenced by expression levels) but for any acquired binding observed for Tie1 irrespective of expression of the cell surface protein.

The stained cells were analysed and selected on a BDFACS Aria SORP II FACS sorter. Figure 7.3A shows the forward and side scatter plot for the fluorescently labelled stable cell population. The cells contained in polygon P1 are live cells which have different complexity, granularity and size as compared with the dead cells (black dots on the plot) or debris. Only the live cells in polygon P1 were analysed in figure 7.3C.

Figure 7.3C is a dot plot showing the level of Tie1 ectodomain bound (to the surface Ang2 FReD domain) represented by the R-PE dye on the Y axis. The cells in the Q3 quadrant are unstained cells. The cells in the Q1 quadrant are red fluorescing cells bound to Tie1 receptor ectodomain. These fluorescently labelled cells were selected by FACS machine ~0.1% of gated live cell population (very stringent machine settings applied which would abort any conflicting events and cells were sorted with an efficiency of >95%) indicated in the polygon P2 were sorted (Figure 7.3C). This population was labelled *C1*. Figure 7.3B shows the statistical information for the 1<sup>st</sup> round of sorting, recorded for a subset of 50000 cells.



**Figure 7.3: Representative dot plot showing the population of cells sorted in the 1<sup>st</sup> round of FACS for Tie1 binding.** 10<sup>7</sup> cells of the stable cell population (*Unsorted*) were incubated with 450nM sTie1 ectodomain. Biotin-linked anti-His antibody was used to detect Ang2-FReD bound (his-tagged) Tie1. These cells were fluorescently labelled with R-PE antibody. Labelled cells were sorted by FACS and data represented as dot plots. (A) Forward and side scatter dot plot showing the complexity and size of the cells. Polygon P1 contains the live cell population that were analysed and sorted based on fluorescence. The black dots represent dead cells or cellular debris. (B) The table shows the statistics for 1<sup>st</sup> round of sorting, recorded for a subset of 50000 cells. (C) The level of Tie1 ectodomain bound to the surface Ang2 FReD domain represented by the R-PE dye on the Y axis. The cells in the Q3 quadrant are unstained cells. The cells in the Q1 quadrant are red fluorescing cells with bound Tie1 receptor ectodomain. Polygon P2 drawn in the Q1 quadrant selects the cells (each dot represents a single cell) that were collected in the 1<sup>st</sup> round of sorting and labelled *C1*.

The cells collected after the 1<sup>st</sup> round of sorting were collected under sterile conditions from the FACS machine and grown in culture media. Subsets of cells were kept in culture for comparison with subsequent sorted populations as discussed in section 7.1.3. The 2<sup>nd</sup> sort used 10 million *CI* cells which were labelled and selected by FACS as described for the 1<sup>st</sup> round of sorting above.

The FACS sorting was repeated iteratively using the protocol outlined for the 1<sup>st</sup> round. Table 7.1 summarizes the details of subsequent sort rounds performed.

Sort Round	Population fluorescently stained	Tie1 conc used for sort (nM)	% population collected	Population sorted
0	---	---	---	<i>Unsorted</i>
1	<i>Unsorted</i>	450	0.1	<i>Sort C1</i>
2	<i>Sort C1</i>	450	1.5	<i>Sort C2</i>
3	<i>Sort C2</i>	450	33.3	<i>Sort C3</i>
4	<i>Sort C3</i>	450	52.0	<i>Sort C4</i>

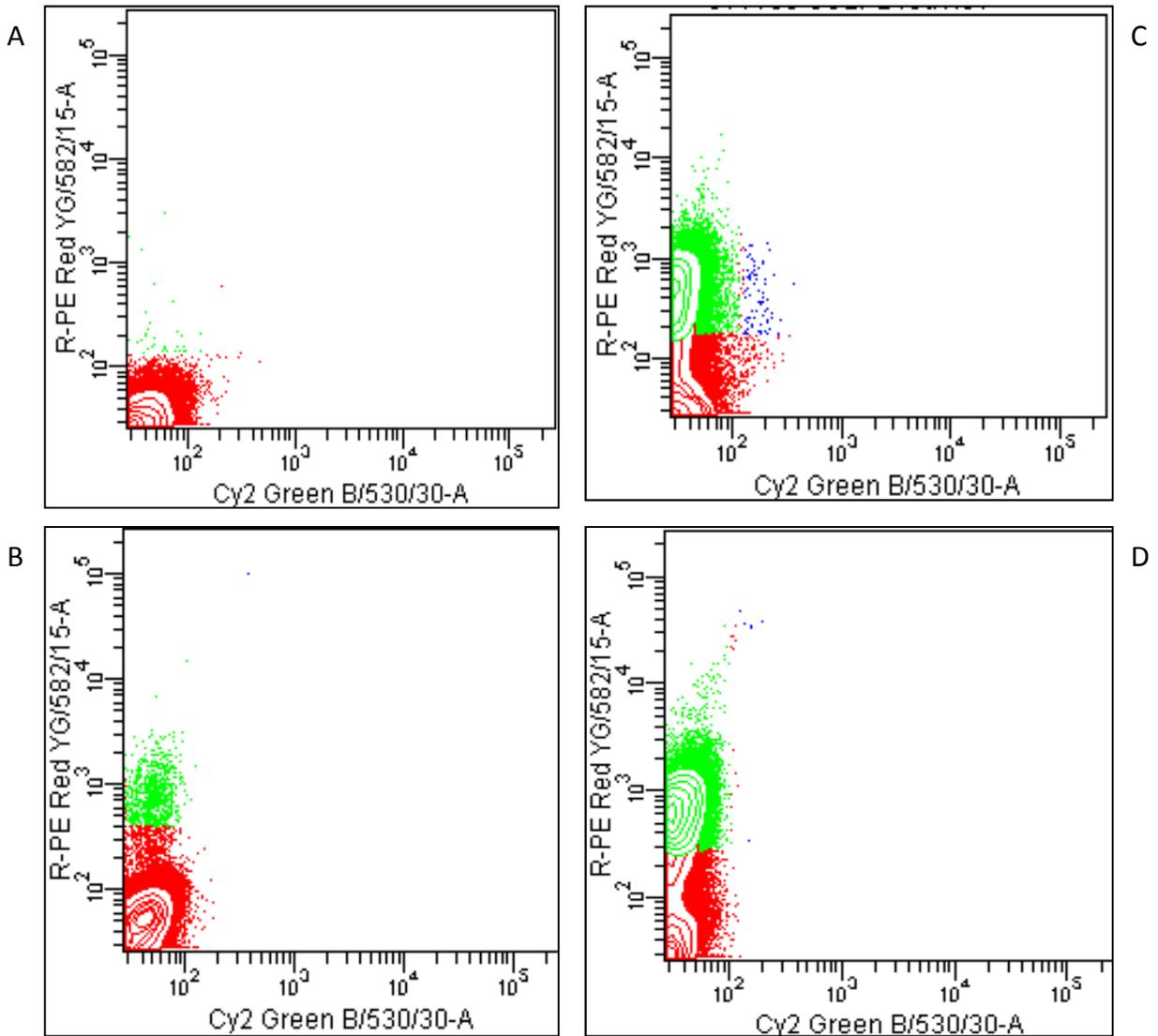
**Table 7.1: Description of the iterative rounds of FACS to select Tie1 binders.** The table indicates the population of cells stained, the Tie1 concentration used for selection, the percentage of the gated population collected and cultured for each round of sorting.

Population of cells from each round of sorting were collected under sterile conditions and grown in culture media. The populations obtained were labelled *Unsorted*, *C1*, *C2* and *C3*.

### **7.1.3 Comparison by flow cytometry of FACS sorted population of cells for Tie1 binding.**

The sorted populations obtained after each round of sorting in section 7.1.2 were analysed against populations obtained from previous sorts by comparative flow cytometry analysis within the same experiment.

Figure 7.4 is an example of the comparison of various sort populations *Unsorted*, *C1*, *C2* and *C3* obtained in section 7.1.2. For comparison, the cell populations were counted and an equal number of cells from each population were incubated with 450nM Tie1. These were then fluorescently labelled using with streptavidin R-PE antibody as described in section 2.22. The R-PE label (red fluorescence) reflects the Tie1 ectodomain bound to the surface Ang2 FReD domain. The various populations are demonstrated in contour plots (Figure 7.4) obtained by flow cytometry analysis.



**Figure 7.4: Contour plots of flow cytometry data showing enrichment of a Tiel binding population.** The cell populations *Unsorted*, *C1*, *C2* and *C3* were incubated at 450nM Tiel1 concentration and stained with streptavidin-R-phycoerythrin fluorophores. A distinct population of Tiel1 binder cells was observed to emerge in cell population *C1* (figure 7.4B). The density of cells in this emerging population were increased with every round as seen in *C2* (figure 7.4C) and *C3* (figure 7.3D) as compared to the *Unsorted* (figure 7.4A) as observed in the contour plots.

As seen in the contour plot in figure 7.4B there is a distinct population of cells emerging. The density of cells in this emerging population increases with every round as seen in figure 7.4C and D. The final enriched population for Tiel1 binding (Figure 7.4D) was analysed by sequencing. These cells were sorted through the FACS machine and collected into PCR tubes. Messenger RNA from these cells was transcribed into cDNA and Ang2-FReD DNA was amplified by PCR as described in section 2.1. This

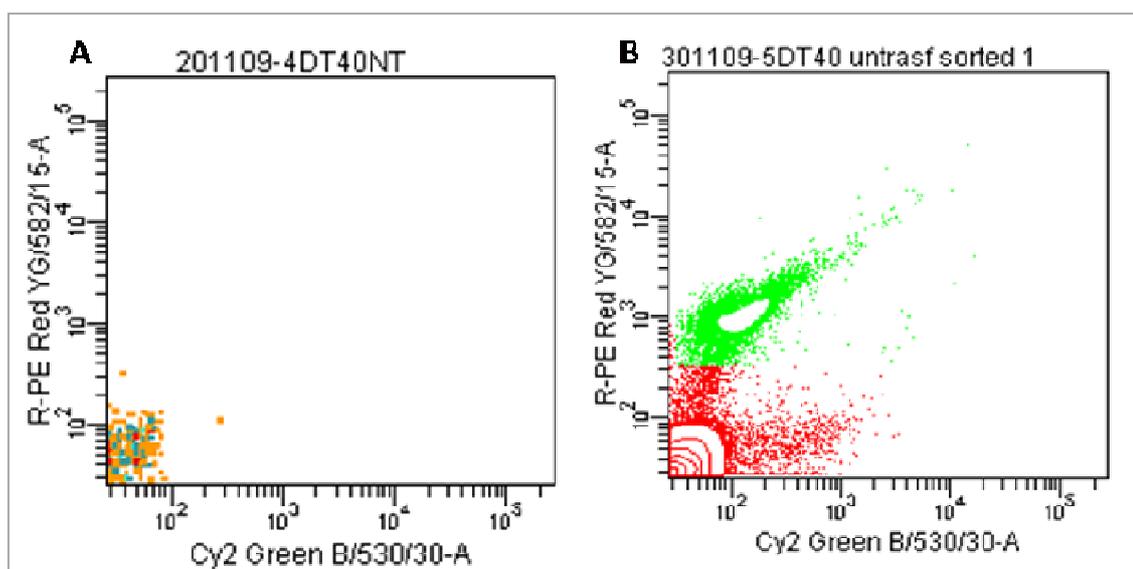
DNA was cloned into a TA sequencing vector and 10 clones sequenced. No mutants were obtained.

The emergence of a distinct population of cells with apparent Tie1 binding could not be explained by presence of SHM generated mutant Ang2-FReD, since all the cells in this population were wild type Ang2-FReD expressers. It was surprising to find that non-specific binding of Tie1 was so strong and consistent that it was able to enrich for a population of cells that were wild type yet displayed distinct consistent binding to Tie1. The explanation for this could be that within the stable population of cells was a subpopulation of cells which displayed Tie1 non specific affinity consistently, that was not a function of the Ang2-FReD expressed on the cells.

## **7.2 Analysing non-transfected cells by iterative rounds of FACS selection.**

To check whether there was a subpopulation of cells within the DT40 stable population which displayed Tie1 non specific affinity consistently, non-transfected cells were analysed by iterative rounds of sorting for Tie1 binding.

The non-transfected cells were sorted through two iterative rounds of FACS using conditions identical in section 7.1.2. Briefly the non-transfected cells were incubated with 450nM Tie1 and fluorescently labelled with R-PE. As seen in figure 7.5, the non-transfected cells without any iterative selection (figure 7.5A) do not display any red fluorescence along the Y-axis. Figure 7.5B confirms that after iterative selection a cell population of Tie1 binders emerge (cells along the Y-axis, coloured green).



**Figure 7.5: Dot plots showing flow cytometry analyses of non-transfected cells after iterative selection for Tie1 binding.** The cell populations *Unsorted NT* and *NT cells sorted through two rounds* were incubated at 450nM Tie1 concentration and fluorescently labelled using streptavidin-R-phycoerythrin fluorophore. (A) Contour plot showing *unsorted non-transfected* cells, incubated with Tie1 at 450nM and stained with streptavidin-R-phycoerythrin. No red fluorescence (along the Y-axis, indicating Tie1 binding) was observed for this population. (B) Contour plot showing *NT cells sorted through two rounds*, incubated with Tie1 at 450nM and stained with streptavidin-R-phycoerythrin. A distinct population of cells (coloured green) is seen to emerge after two rounds of sorts. This population of cells was 10% of the parent population stained.

Thus it can be concluded that the cell surface of DT40 cells was non-specifically sticky for Tie1 binding which lead to inability to select for binders. The iterative improvements in non-transfected cells are comparable to that observed for Tie1 binding in Ang2-FReD stables.

The experiment of iterative sorting of non transfected cells was not performed previously. Non transfected cell and stable cells had been analysed as FMO (fluorescence minus one) controls and also by dual staining (Section 2.24). It should be noted that the non-transfected cells were stained with various concentrations of primary and secondary antibodies to detect any background binding. On flow cytometry analyses none of the controls for non-transfected cells stained for neither Cy2 nor

streptavidin-R-phycoerythrin. The number of non-transfected, non-specific binders for Tie1 was approximately 0.06% (of the gated live cells) in the first round of staining. These few cells seem to have drastically got enhanced in 4 rounds of FACS sorts as observed in figure 7.4. The sorting parameters on the FACS machine were set to the most stringent settings possible to abort all conflicting events, analysing not more than 500-1000 events/second and sorting cells at the highest efficiency possible of at least 95%. However, the interaction of Tie1 to the cell surface proteins could not be prevented.

## **Discussion**

In this chapter it was attempted to evolve the Ang2-FReD protein to bind to the Tie1 receptor for which there is no known ligand. FACS sorting was used to select binders from the mutant library generated by SHM in the DT40 cell line. It was thought that the problems encountered previously while enhancing binding of the Ang2-FReD protein for Tie2 would not be encountered when selecting for a new function. It could be argued that it was difficult to increase binding to Tie2 as it has (possibly) already evolved optimally for Ang2 binding. However, there would be greater potential to increase binding of a poor non-binder; this was tested for Tie1.

A population of cells was enriched for Tie1 binding in just four rounds of FACS sorting. However, sequencing of this enriched population yielded no mutants, suggesting that the non-specific interaction was due to DT40 cell surface proteins. Indeed, iterative sorting of non-transfected DT40 cells, yielded similar results. The unusual high degree of enrichment for apparent Tie1 binding was obtained rapidly with just 4 sort rounds, suggesting that the non-specific interaction of Tie1 to the DT40 cell surface protein was very consistent to the DT40 cell surface protein.

Non-transfected cells had been assessed for non-specific binding using Fluorescence Minus One (FMO) controls (Chapter 2), and none of the components (antibodies or receptor ectodomains or labelling dyes) used for FACS showed any significant background binding. The non-transfected cells sorted due to the non-specific Tie1-cell interaction were reanalysed using FMO controls. As expected, they showed no cross reactivity to the reagents used in FACS and that the enrichment selection was indeed due to the Tie1 interaction with cell surface proteins (data not shown).

Thus, the FACS assay used for selection of binding function was inadequate to screen for the sought binding between Ang2 FReD and Tie1 due to an unidentified but improvable feature of DT40 cells. What protein on the DT40 cell surface may be capable of this high non-specific interaction with Tie1 was not investigated.

Evolving protein to obtain a new function requires sampling a large number of variants to achieve success. The frequency of generating a new functional protein is approximately 1 in  $10^{11}$  (Zhao 2007). To achieve success would require the screening assay to be impeccable, and be able to pick out the only needle from the haystack.

In addition to the inability to detect binding activity there was a problem with generating mutants. Although, the stable cells were cultured under selective pressure (MPA) to prevent excision of the AID DNA, they were found to excise AID DNA unexpectedly; the reason for this was unexplainable. This could result in sub-optimal level of SHM which in turn could drastically decrease the library size and hence the chance of obtaining any variant to evolve new protein function.

## Chapter 8: General Discussion

Protein engineering to manipulate known primary proteins has developed rapidly in the past 30 years due to the advances in biotechnological tools making it easier to alter DNA sequences and link the genotypic change to a desired phenotypic trait. Protein engineering is usually aimed at designing proteins with potential therapeutic or biotechnological applications. The past few decades of research on developing protein based drugs are beginning to show encouraging potential and offer several advantages like specificity in function, lower interference with other biological processes, lower immunogenicity, when compared to small organic molecule or small molecule peptide drugs (Leader, Baca & Golan 2008, Walsh 2005).

The angiopoietin group of ligands were discovered less than two decades ago and are exciting candidates for protein engineering (Davis et al. 1996, Maisonpierre et al. 1997). Ever since the discovery of the ligands, numerous studies have cited their involvement in angiogenesis and associated pathological conditions. Several animal studies have shown successes of angiopoietins (Ang) as a therapeutic molecule but production of active purified native Ang protein is difficult. Hence Ang would be an ideal candidate for protein engineering (van Meurs et al. 2009).

Engineered angiopoietin variants, have been generated previously by manipulation and minimalization of the oligomerization motif of the protein (Kim et al. 2009, Cho et al. 2004). These recombinant angiopoietin proteins possess increased stability and solubility as compared to the native forms. However, manipulation of the receptor binding interface has not been studied extensively. Manipulation of the binding interface could lead to new binding affinity variants of the angiopoietins. For example,

Ang1 with higher affinity may be used to treat pathological conditions in which Ang2 is expressed.

Current methods of protein engineering involve mutation of the gene sequence and selecting the effective phenotypic change in function. These methods have been broadly divided into two subsets; one is the rational design approach which uses protein structural information combined with computational tools to target mutagenesis to a focussed loci. The other approach is directed evolution, which uses random mutations in the gene sequence and emphasizes iterative accumulation of mutations until a desired level of function is achieved. Both approaches have had their successes. However, the complexity of protein structure combined with the limitations of current approaches has held back the practical realization of the true potential of protein engineering. Indeed in the present study, rational substitution of residues in the Ang2 RBD to generate mutants Asn467Ser, Ser480Pro and Gly429Thr were unable to predict the exact outcome of the mutation at the residue position. Some mutations which were expected to affect binding affinity had no effect while others showed the opposite outcome. However, the study is limited with the number of residues explored and the amino acid changes performed using rational substitutions in Ang2FReD.

In the work presented in this thesis, protein engineering of the angiopoietin ligand has been attempted using somatic hypermutation as an alternative tool for directed evolution. Somatic hypermutation (SHM) is a process by which random mutations are introduced in the rearranged immunoglobulin gene (Bernard, Hozumi & Tonegawa 1978). SHM is efficient and the rapid diversification enables the production of high affinity antibodies to counteract numerous antigens in the environments. It was assumed that this process by which random mutations are introduced is targeted only to the Ig

locus, since indiscrete mutation of genes could be deleterious to the organism. The AID mutator enzyme plays a key role in SHM however; the exact mechanism of SHM targeting is unknown (Muramatsu et al. 1999). It was recently discovered that SHM is capable of introducing mutations in highly transcribed non-Ig transgenes. The discovery of this feature of SHM has been exploited successfully for evolution of fluorescent proteins (Wang, Yang & Wabl 2004, Wang et al. 2004, Arakawa et al. 2008). These experiments employed AID positive B cell lines into which fluorescent transgenes were introduced and transcribed under the influence of a strong promoter. Simply by culturing these cells and iterative selection for high fluorescing mutants by FACS, improved fluorescent proteins were obtained.

In this thesis I attempted to extend the SHM driven mutagenesis to evolve the Ang-FReD for improved binding. The strategy involved cloning the Ang-FReD domain fused to the ASGPR transmembrane domain so as to enable the anchoring of Ang-FReD domain on the B cell (AID positive) surface. The ASGPR transmembrane domain is known to orient protein on the extracellular surface with a free C-terminal. Cell surface display combined with fluorescent labelling was used for iterative rounds of FACS selection.

The expression of the fusion protein was difficult to detect in initial SHM experiments performed in the Ramos B cell line. This could be attributed to the fact that the integration of the fusion gene was either at a poorly transcribed genomic locus or that the cell line was unsuitable for the expression of these processed proteins. Furthermore, the Ramos cell line may be unsuitable for SHM experiments wherein the cells are required to be cultured for prolonged periods. This is because the hypermutation phenotype in Ramos cells was reported to be unstable in cultures grown

for more than 1 month, and this was linked to low AID expression levels (Zhang et al. 2001).

To overcome the problem of poor expression and random integration in the Ramos cells, an alternative cell line was used. The engineered DT40 bursal cell line displays high expression of AID enzyme and was used for expression of the ASGPR-AngFReD fusion protein. For this, the fused DNA (ASGPR-AngFReD) was cloned into a vector designed for targeting the DNA sequence to the Ig locus in the genomic DNA. The Ig locus is known to support high rate of SHM (Wang et al. 2004).

The ASGPR-Ang2FReD fusion protein was successfully expressed and localised to the extracellular surface in the stable DT40 cells. The ASGPR-Ang1FReD fusion protein expression could not be localised on the extracellular surface. This discrepancy in expression between Ang1 and Ang2 proteins could be attributed to their interaction with the ASGPR transmembrane domain possibly while translation and folding. Further experiments for SHM mutagenesis were continued with Ang2FReD protein.

Two types of iterative selection were performed; in chapter 6, for enhancement of Tie2 binding affinity and in chapter 7, for acquiring new binding function to Tie1. DT40 cells were transfected with vectors designed to enable targeting of the DNA into the genome. The cells were cultured under selective antibiotic pressure to obtain a stable cell population expressing Ang2-FReD protein. The stable cells were analysed for AID expression and the mutation frequency in the cells was determined. The expressed protein was characterised by immunofluorescence for localisation to the extracellular cell surface and binding to Tie2. The expressed protein was analysed by flow cytometry to determine the  $K_d$  for Tie2 binding in the DT40 cell surface expression system. The FACS selection was performed initially using non-stringent concentrations

of Tie2 and non-stringent sorting. With every round of FACS the conditions were altered to gradually become more and more stringent.

Distinct population of cells were enriched for both Tie2 and Tie1 iterative sorts. However, sequencing of the final sorted population did not yield any mutants. Prior to any round of FACS sorting the stable cell population had been analysed for AID expression and the mutation frequency of ~0.1-0.2 calculated. Yet, sequencing of the final sorted population did not yield any mutants. Analysis of this population showed that the AID expression was lost. It is possible that AID gene was excised from the DT40 cells by the background Cre recombinase activity in prolonged culture. This excision could not be prevented by usage of mycophenolic acid (MPA). Another possibility may be that, like the Ramos cells, the DT40 cell line also displayed instability in undergoing hypermutation due to loss in AID expression. Another explanation for this loss in AID expression may be because the highly transcribed AID may have been predisposed to SHM mutation. It has been reported that AID mutator is capable of self mutating when the rate of transcription is high (Martin, Scharff 2002). The DT40 cell line used was an engineered cell line with higher expression compared to the wild-type DT40 cells (Arakawa, Hauschild & Buerstedde 2002), this may have resulted in self-hypermutation and hence loss of expression. Irrespective of the reason for loss of AID, the iterative selections led to enrichment of the AID negative sub-population, which displayed enhanced binding to Tie2, possibly due to cell surface binding activity independent of Ang2FReD.

The enrichment for Tie1 binding function was concluded to be due to non-specific binding of Tie1 to cell surface proteins. Although the screening assay for FACS had been optimised so that none of the labelling reagents gave any background, it was not

anticipated that the non-specific binding of Tie1 to the cells surface proteins would be enhanced so dramatically by iteration. This non-specific binding in the assay was stable and consistent to be enriched over any real binding between Ang2-Tie1.

Several parameters should be optimized to ensure the success of this SHM approach for selection of binding function. The maintenance, expression and activity of the AID enzyme in the chosen cell line should be stable, since it is crucial for diversification and generation of mutants. The screening assay should be thoroughly checked for false positive cross reactions and should be sensitive to pick out the desired function.

Furthermore, it is difficult to evolve proteins which have a well defined function in important cellular pathways. Proteins which play critical cellular functions may be less evolvable than proteins which may be dispensable for survival; hence evolution of fluorescent proteins may in comparison be easier. As with all protein engineering approaches, one should consider the evolvability of the protein in question. Some proteins are more evolvable and various factors seem to influence the evolvability of protein (Aharoni et al. 2005, Bloom, Raval & Wilke 2007, England, Shakhnovich 2003, Tokuriki, Tawfik 2009). Proteins with naturally promiscuous function may be more evolvable than a protein with a well defined specific function. Proteins which are part of a large protein family containing closely related members with respect to sequence, structure or function may be quicker to manipulate (Romero, Arnold 2009).

Compared to enhancement of function, gain of new function (like Tie1 binding) is far more difficult to introduce and usually requires a higher mutation frequency. Directed evolution for generating a non-existent function in a protein is overwhelming and success is usually achieved for proteins with high evolvability and after several

enrichment rounds. Success in novel function acquisition is achieved for highly evolvable genes like that of ampicillin resistance which involved 50 rounds of DNA shuffling (Yano, Kagamiyama 2001). ATP-binding function was also achieved by screening library size of  $>10^{12}$  proteins with 80 continuous random amino acids (Keefe, Szostak 2001). The instability of the AID gene expression in the cultured DT40 cell line poses an additional challenge to utilise SHM mediated directed evolution for evolving proteins with new function, requiring prolonged stable culturing.

Developing existing proteins to acquire new function requires accumulation of multiple synergistic and sometimes neutral mutations, since individual mutations may have no effect (Bershtein, Goldin & Tawfik 2008, Bolon, Voigt & Mayo 2002, Gupta, Tawfik 2008). Thus the screening assay for novel function would require maintaining neutral drifting variants, while the SHM mediated mutagenesis should occur at higher frequency to generate sufficient, multiple and useful mutations. However, a high mutation frequency does not ensure success as multiple mutations can sometimes be destabilizing (Romero, Arnold 2009).

In spite of the difficulty in AID expression and screening assay some mutants obtained by SHM were analysed. Four mutants *Hm5*, *10c*, *FACSD* and *FACSi* were obtained by random sequencing of the stable DT40 cell population. Three mutants had a single mutation in the FReD DNA and one mutant had 2 mutations in the FReD DNA. Mutants *10c*, *FACSD* and *FACSi* were characterised by Western blotting and  $K_d$  determination. The *10c* and *FACSD* mutants did not appear to be different to the WT Ang2FReD protein with respect to protein expression and binding affinity. The *FACSi* mutant failed to express on the cell surface and hence could not be analysed for binding.

Redesign of Ab-Ag interfaces is not always predictable from 3-D structure and the natural process of SHM may give insight into the optimal trends required for affinity improvement. This is because SHM of Ig genes is influenced by; biases for positions that commonly contact antigen, location of mutation hotspots, accessibility from single base permutations and functional pressures (Clark et al. 2006). Of the four main factors influencing SHM of antibodies, not all apply for the SHM of non-Ig genes. Non-Ig genes mutate in the absence of functional pressure. However, mutants *Hm5*, *10c*, *FACSD* and *FACSi* generated by SHM driven mutagenesis may show some resemblance to the trends observed for natural SHM. Natural SHM displays ~3:1 ratio of transitions: transversions (Betz et al. 1993), but mutants *Hm5*, *10c*, *FACSD* and *FACSi* displayed 1:3 ratio favouring transversion mutations. However, to get an accurate ratio a larger sample size of mutants need to be analysed.

Analysis of a large number of Ab chain sequences revealed that solvent exposed surface positions which are more likely to contact the Ag are also more likely to be targets of productive mutations (Clark et al. 2006). The 3-D position of the mutated residues in mutants *Hm5*, *10c*, *FACSD* and *FACSi* shows that the mutations lie on the surface of the protein and not in the core regions which are more likely to destabilize the fold. Furthermore, aspartate and asparagine are favoured residues generated by SHM (Clark et al. 2006). *FACSD* and *FACSi* mutants have mutated three residue positions to aspartate or asparagine. It is possible that some trends observed for improving antigen binding affinity are also applied to changes in binding affinity of other protein-protein interactions performed by SHM.

No matter how efficient the process of mutagenesis is, in generating novel mutants, the success critically depends on the screen to isolate the novel variant. As a

general rule, kinetic screen should be employed for  $K_d < 10\text{nM}$  and equilibrium screen when  $K_d > 10\text{nM}$  (Boder, Wittrup 1998). When setting up a surface display system for a protein in a cell line, the  $K_d$  of the WT protein in that environment should be determined. Absolute values of  $K_d$  obtained for the same proteins in different assays may vary (Lofblom et al. 2007). Similarly the cell line in which the assay is performed also seems to influence the absolute  $K_d$  value. As observed with the  $K_d$  values for WT Ang2FR<sub>2</sub>D in HEK293 cell line which was ~30 times less than that observed in the DT40 cell line. Once the assay is established the desired mutants can be successfully and sensitively isolated by cell surface display and FACS (Boder, Wittrup 1997, Kieke et al. 1997). The screening process should be customized not only for the protein functional assay, but also for the cell line in which the assay is performed. As seen with the DT40 cells, the cell surface proteins were too sticky for Tie1 and may not be ideal for obtaining binding function in a cell surface display assay. Although none of the reagents produced any background when tested against each other, the interaction of Tie1 with cell surface molecules after iterative selection was not performed. In retrospect, it may be better to perform iterative rounds of selection with non-transfected cells; to rule out the possibility of any enrichment by iteration of non-specific interaction of the binder for which the affinity improvement or binding assay is being performed.

In spite of these problems encountered, SHM mediated mutagenesis holds promise for improvement of affinity in proteins. Engineering a more stable AID expressing cell line may be useful, given the unexpected loss of expression of AID in the DT40 cells. It may be useful to make the AID expression controlled or inducible by external stimuli. Improvement in screening technique holds the key to success of this approach. Although FACS selection has proven to be extremely accurate, interference of non-specific binding of cell surface proteins has hindered success. It may be useful to

analyse non-transfected cells by iterative selection to rule out non-specificity or optimise the assay to exclude non-specificity or use an alternative cell line in which non-specificity is not observed.

Rational design and directed protein evolution processes to evolve function are time consuming and often unpredictable. Natural evolution has a well defined algorithm which has achieved the desired functions. Somatic hypermutation like natural evolution is a natural process, but performed over a shorter time-scale. The success of the SHM process is evident from the near perfect accuracy of antibodies produced to combat antigens. Utilising SHM as a mutagenesis tool to alter non-Ig genes may be an interesting alternative to traditional approaches.

## Appendix One

### List of Analysis Software, Cell Lines, Equipment, Kits and Vectors

#### Analysis Software

1. PyMOL 1.2
2. WinMDI 2.8
3. CellQuestPro
4. FACS DIVA 6
5. GraphPad Prism 5.0

#### Cell Lines

1. Chinese Hamster Ovary cells (CHO) [originally from ECACC]
2. Human Embryonic Kidney 293 Cells (HEK) [originally from ECACC]
3. Ramos (Human Burkitt's lymphoma cell line) [originally from ECACC]
4. DT40 chicken B cell line (This engineered DT40 cell line was kindly supplied by Dr. Jean-Marie Buerstedde)

#### Equipment

1. Perkin Elmer thermocycler
2. Multi Image Light Cabinet (Flowgen-Alpha Innotech Corporation)
3. UV-Visible spectrophotometer (Shimadzu)
4. Bio-Rad Gene Pulser Xcell electroporator
5. Nucleofector (Amaza Biosystems)
6. Spiramix (Denley)
7. Image Capture Camera: NikonTE2000U

8. MACS Separation columns (MS columns) and MACS Multistand magnet (Miltenyi Biotec)
9. Flow Cytometer (BD FACScan)
10. Fluorescence Activated Cell Sorter (BDFACS Aria SORP II)

### **Kits**

1. MinElute Gel Extraction kit (Qiagen)
2. pcDNA3.1V5-His-TOPO TA Expression Kit (Invitrogen)
3. pSecTag/FRT/V5-His TOPO TA Expression Kit (Invitrogen)
4. TOPO TA cloning Kit for sequencing (Invitrogen)
5. QuikChange Site Directed Mutagenesis Kit (Stratagene)
6. RNeasy Mini Kit (Qiagen)
7. RETROscript Kit (Ambion)
8. QIAprep Mini Prep kit (Qiagen)
9. EndoFree Plasmid Maxi Kit (Qiagen)
10. Cell Line Nucleofector Kit T (Amaxa Biosystems)
11. Cell Line Nucleofector Kit V (Amaxa Biosystems)

### **Vectors**

1. pSecTag/FRT/V5-His TOPO for mammalian secreted protein expression
2. pcDNA3.1V5-His-TOPO for mammalian cellular expression
3. pHyperm2 Vector for recombination into DT40 genome and expression in the DT40 bursal cell line. (This vector was kindly supplied by Dr. Jean-Marie Buerstedde)
4. pCR4 TOPO vector (3956bp)

## Appendix Two

### List of Primers

#### Chapter 3

##### PCR Ang1 FReD

Primers 5'FLAGFReD (Tm=73.9) and 3'bamFReD (Tm=68.3) were used for PCR amplification and annealing temp of 60°C.

FLAGFReD      5' -GACTACAAAGACGATGACGACAAGCCATTTAGAGACTGTGCAGATG-3'

BamFReD        5' -CGGGATCCCGTTCAAAAATCTAAAGGTCGAATCA-3'

##### PCR Ang1 P domain

Primers 5'pdomtopo (Tm>75, this primer has a 5' FLAG tag) and 3'bamFReD (Tm=68.3) were used for PCR amplification at annealing temp 62°C.

pdomtopo        5' -GACTACAAAGACGATGACGACAAGGGGACAGCAGGAAAACAGAG-3'

BamFReD        5' -CGGGATCCCGTTCAAAAATCTAAAGGTCGAATCA-3'

#### Chapter 4

##### Primers for site directed mutagenesis to generate *N467S*, *S480P* and *G428T* mutants.

Primers Asn467Ser Fwd (Tm=74), Asn467Ser Rev were used for site directed mutagenesis to generate mutant *N467S*. Primers Ser480Pro Fwd (Tm>75), Ser480Pro Rev were used for site directed mutagenesis to generate mutant *S480P*. Primers Gly428Thr Fwd (Tm=69), Gly428Thr Rev were used for site directed mutagenesis to generate mutant *G428T*.

Asn467Ser Fwd    5' -CCACAGAGGCAGAACACAAGCAAGTTCAACGGCATTAAATGG-3'

Asn467Ser Rev    5' -CCATTTAATGCCGTTGAACTTGCTTGTGTTCTGCCTCTGTGG-3'

Ser480Pro Fwd    5' -GGTACTACTGGAAAGGCCCGGCTATTTCGCTCAAGGCC-3'

Ser480Pro Rev 5' -GGCCTTGAGCGAATAGCCGGGGCCTTTCCAGTAGTACC-3'  
 Gly428Thr Fwd 5' -GCAAATACATTTGTCGTTGTCTGTATGGTTTGTGCTAAAATC-3'  
 Gly428Thr Rev 5' -GATTTTAGCACAAAGGATACAGACAACGACAAATGTATTTGC-3'

## **Chapter 5**

### **PCR ASGPR**

ASGPRF( $T_m > 75$ ) and ASGPRR( $T_m > 75$ ) primers were used to do the overlap PCR. ASGPR5'( $T_m = 72.7$ ) and ASGPR3'( $T_m = 65.3$ ) primers were used to amplify the template from the overlap PCR.

ASGPRF 5' -ATGGGACCTCGCCTCCTCCTGCTCTCCCTGGGCCTCAGCCTCCTGCTGCT  
 TGTGGTTGTCTGTGTGATCGGTTCCCAAAA-3'  
 ASGPRR 5' -CCTTTAAATTCTTGTGTCATCGTCTTTGTAGTCTACGCCCTGCAGCTGG  
 GAGTTTTGGGAACCGATCACACAGACAA-3'  
 ASGPR5' 5' -GGGGTACCATGGGACCTCGCCTCCTC-3'  
 ASGPR3' 5' -GGGGATCCTTTAAATTCTTGTGTCATCG-3'

### **Site directed mutagenesis of the incorrect ASGPR clone TA a**

To correct the first mutation in the clone TAa the StartForward ( $T_m = 72.6$ ) and StartReverse ( $T_m = 72.6$ ) primers were used. Two sets of PCR were done with annealing temperatures 60 °C and 65 °C degrees.

StartFoward 5' -GCCCTTATGGGACCTCGCCTCCTCCTG-3'  
 StartReverse 5' -CAGGAGGAGGCGAGGTCCCATAAGGGC-3'

### **PCR Ang1 FReD Not**

Primers 5'FReDNot Fwd( $T_m = 73.7$ ) and 3'FReDNot rev( $T_m = 69.5$ ) were used for PCR amplification at annealing temp 45°C.

AlFrNotI\_for 5' -ATAAGAATGCGGCCGCCCATTTAGAGACTGTGCAGATG-3'  
 AlNotI\_rev 5' -ATAGTTTAGCGGCCGCTTCAAAAATCTAAAGGTCAATCA-3'

**PCR Ang2 FReD Not**

Primers 5'Ang2FReDNot Fwd( $T_m > 75$ ) and 3'Ang2FReDNot rev( $T_m > 75$ ) were used for PCR amplification at annealing temp 45°C.

Ang2FrNotI\_for 5'-GCAGCAGCAGCGGCCGCCAGCTTCAGAGACTGTGCTGAAGTATTC-3'

Ang2NotI\_rev 5'-TGCTGCTGCGCGGCCGCGTCAGAAATCTGCTGGTCGGATCATCATGGT-3'

**Chapter 6****PCR Ang1 FReD Nhe-EcoRV**

Primers 5'ASGPRNheIfwd( $T_m > 75$ ) and 3'Ang1EcoRV rev( $T_m = 69$ ) were used for PCR amplification at annealing temp 50°C.

AngNhe1\_for 5'-TTATTATTAGCTAGCATGGGCCCCAGGCTGCTGCTGCTGTCC-3'

A1EcoRV\_rev 5'-ATAGTTTAGGATATCTTCAAAAATCTAAAGGTGGAATCA-3'

**PCR Ang2 FReD Nhe-EcoRV**

Primers 5'ASGPRNheIfwd( $T_m > 75$ ) and 3'Ang2EcoRV rev( $T_m = 72.4$ ) were used for PCR amplification at annealing temp 50°C.

AngNhe1\_for 5'-TTATTATTAGCTAGCATGGGCCCCAGGCTGCTGCTGCTGTCC-3'

A2EcoRV\_rev 5'-TGCTGCTGCGATATCTTAGAAATCTGCTGGTCGGATCATCAT-3'

**Primers for Chick AID amplification**

Primers 5'AIDchickfwd( $T_m = 68.2$ ) and 3'AIDchickrev( $T_m = 67$ ) were used for PCR amplification at annealing temp 55°C.

5'AIDchickfwd 5'-AGAAGATCTATGGACAGCCTCTTGATGAAGAGG-3'

3'AIDchickrev 5'-GCAGAATTCTCAAAGTCCCAGAGTTTTAAAGGC-3'

**Primers for amplifying Ang2FReD from clones obtained by FACS sorted stable F8Hd populations.**

Primers 5'Ang2FReDfwd(Tm=66) and 3'Ang2FReDrev(Tm>75) were used for PCR amplification at annealing temp 50°C.

Ang2Fred\_forw 5'-AGCTTCAGAGACTGTGCTGAAGTATTC-3'  
 Ang2Fred-Rev 5'-TCAGAAATCTGCTGGTCGGATCATCATGGTTGT-3'

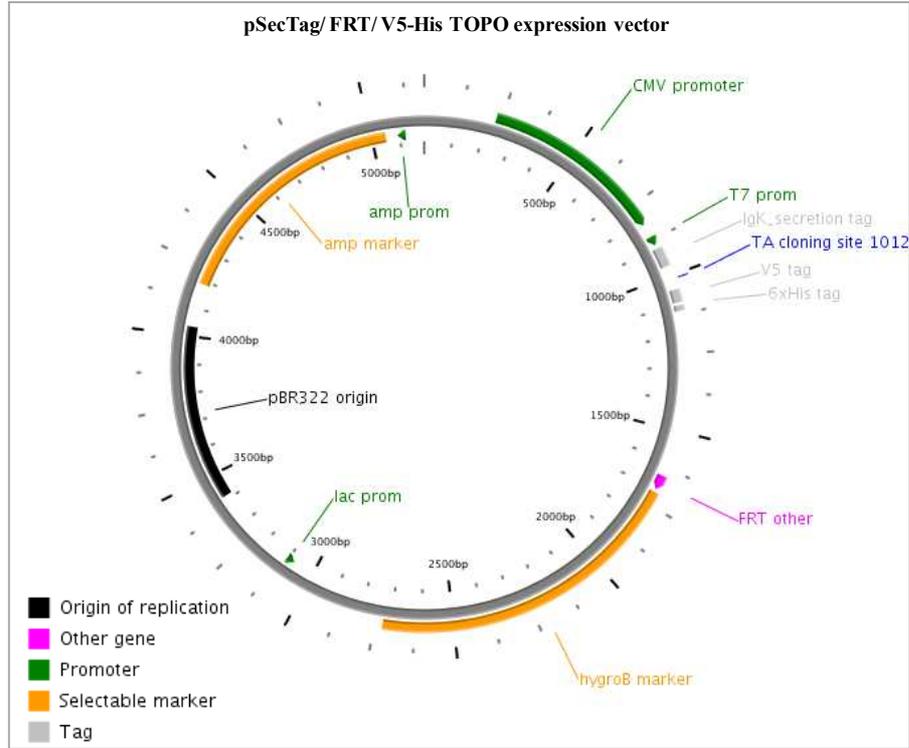
**Primers for site directed mutagenesis to generate *10c*, *FACSD*, *FACSi*, mutants.**

Primers Fwd 10c (Tm=70), Rev10c were used for site directed mutagenesis to generate mutant *10c*. Primers Fwd *FACSi* (Tm=71), Rev *FACSi* were used for site directed mutagenesis to generate mutant *FACSi*. Two pairs of primers Fwd1FACSD (Tm=71), Rev1FACSD and Fwd2FACSD (Tm>75), Rev2FACSD were used for site directed mutagenesis to generate double mutant *FACSD*.

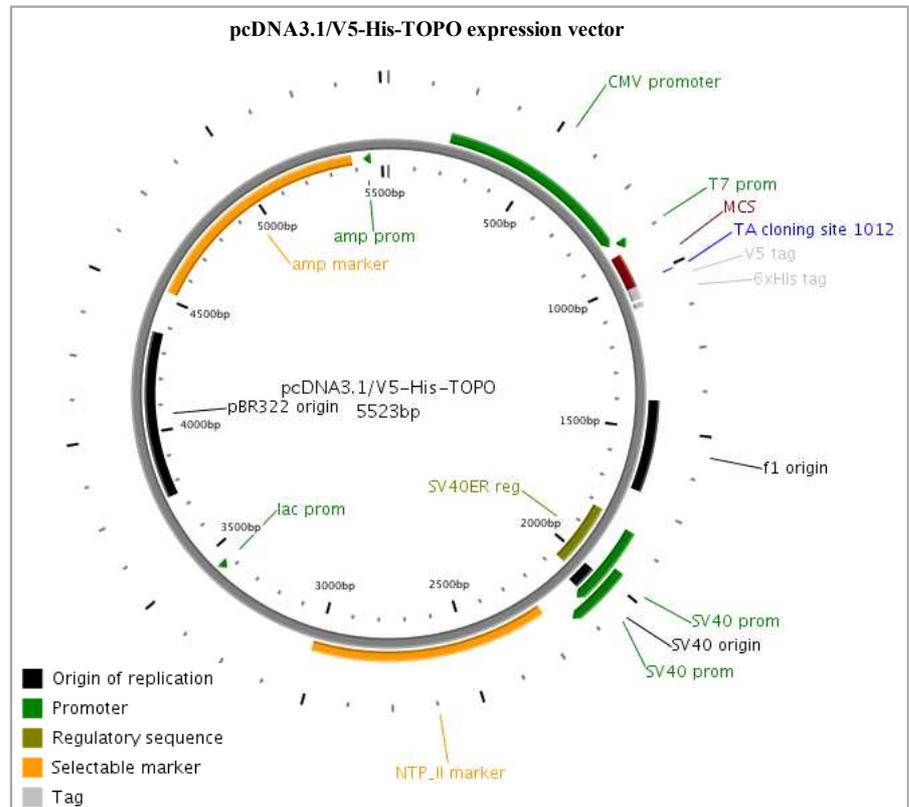
Fwd10c 5'-GGCATCTACACGTAAACATTCCGTAATTCTACAGAAGAG-3'  
 Rev10c 5'-CTCTTCTGTAGAATTACGGAATGTTAACGTGTAGAT-3'  
 FwdFACSi 5'-GGACTTACAGGGACAGCCGACAAAATAAGCAGCAT-3'  
 RevFACSi 5'-ATGCTGCTTATTTTGTCTGGCTGTCCCTGTAAGTCC-3'  
 Fwd1FACSD 5'-AATCAGGACACACCACAGATGGCATCTACACGTAAACA-3'  
 Rev1FACSD 5'-TGTTAACGTGTAGATGCCATCTGTGGTGTGTCCTGATT-3'  
 Fwd2FACSD 5'-GATGATCCGACCAGCAAATTTCTAAGCGGCCGCCGGCG-3'  
 Rev2FACSD 5'-CGCCGGCGGCCGCTTAGAAATTTGCTGGTCGGATCATC-3'

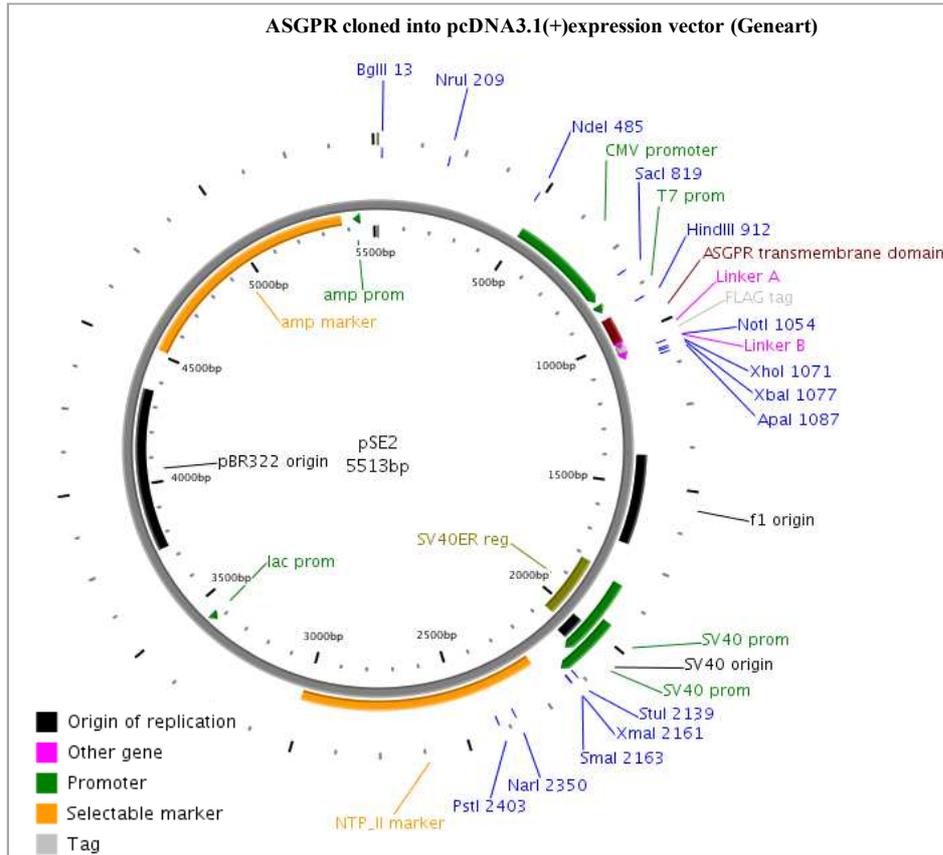
## Appendix Three: Vector Maps (Dong et al. 2004)

### Chapter 3

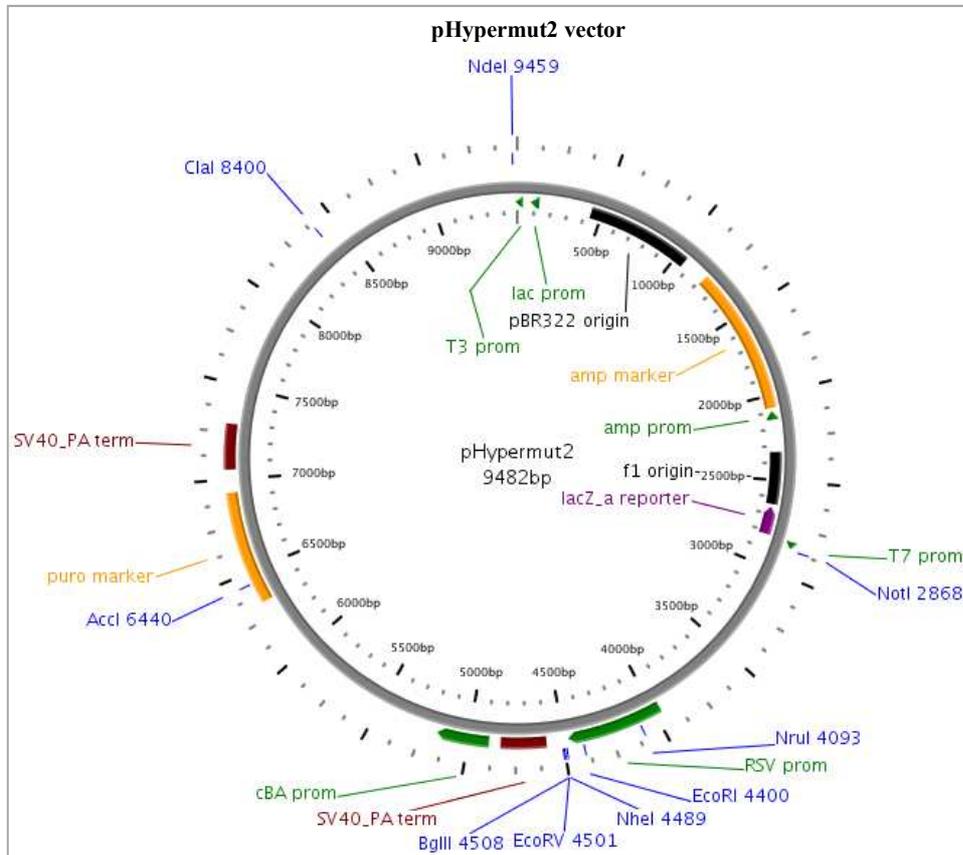


### Chapter 5





**Chapter 6**

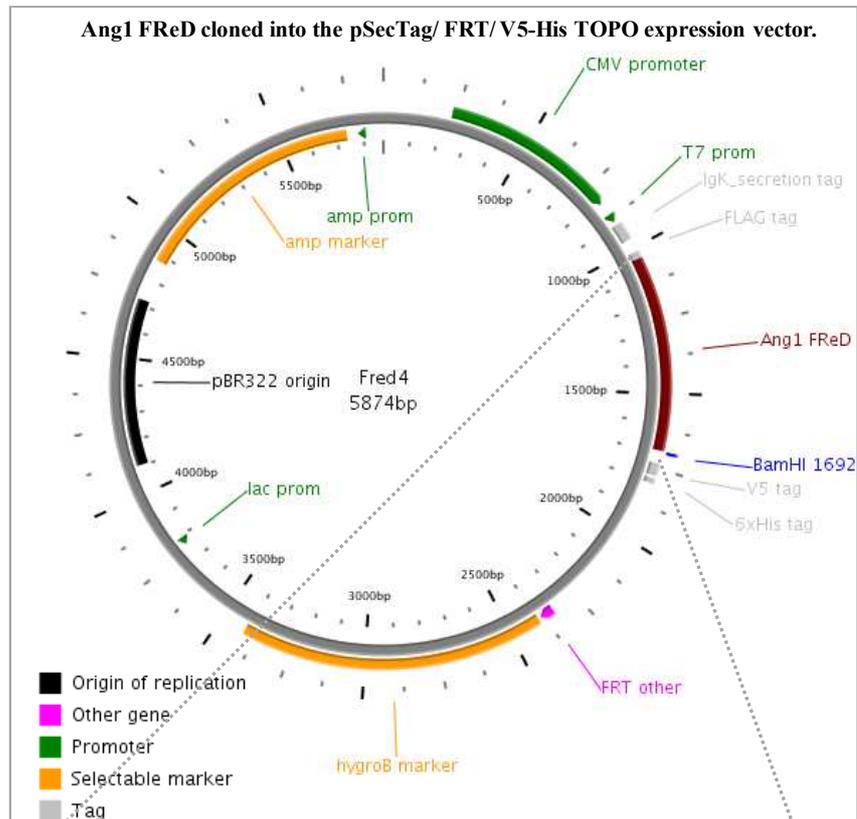


## Appendix Four

### Sequence Alignments

#### Chapter 3

#### A) *Fred 4 Sec*: Ang1 FReD cloned into the pSecTag/ FRT/ V5-His TOPO expression vector.



```

Query 1      CCATTTAGAGACTGTGCAGATGTATATCAAGCTGGTTTTAATAAAAAGTGAATCTACACT 60
Sbjct 1312   CCATTTAGAGACTGTGCAGATGTATATCAAGCTGGTTTTAATAAAAAGTGAATCTACACT 1371

Query 61      ATTTATATTAATAATATGCCAGAACCCAAAAGGTGTTTTGCAATATGGATGTCAATGGG 120
Sbjct 1372   ATTTATATTAATAATATGCCAGAACCCAAAAGGTGTTTTGCAATATGGATGTCAATGGG 1431

Query 121     GGAGGTTGGACTGTAATACAACATCGTGAAGATGGAAGTCTAGATTTCCAAAGAGGCTGG 180
Sbjct 1432   GGAGGTTGGACTGTAATACAACATCGTGAAGATGGAAGTCTAGATTTCCAAAGAGGCTGG 1491

Query 181     AAGGAATATAAAATGGGTTTTGGAAATCCCTCCGGTGAATATTGGCTGGGGAATGAGTTT 240
Sbjct 1492   AAGGAATATAAAATGGGTTTTGGAAATCCCTCCGGTGAATATTGGCTGGGGAATGAGTTT 1551

Query 241     ATTTTGGCCATTACCAGTCAGAGGCAGTACATGCTAAGAATTGAGTTAATGGACTGGGAA 300
Sbjct 1552   ATTTTGGCCATTACCAGTCAGAGGCAGTACATGCTAAGAATTGAGTTAATGGACTGGGAA 1611

Query 301     GGGAAACCGAGCCTATTCACAGTATGACAGATTCCACATAGGAAATGAAAAGCAAACACTAT 360
Sbjct 1612   GGGAAACCGAGCCTATTCACAGTATGACAGATTCCACATAGGAAATGAAAAGCAAACACTAT 1671

Query 361     AGGTTGTATTTAAAAGGTCACACTGGGACAGCAGGAAAACAGAGCAGCCTGATCTTACAC 420
Sbjct 1672   AGGTTGTATTTAAAAGGTCACACTGGGACAGCAGGAAAACAGAGCAGCCTGATCTTACAC 1731

```

```

Query 421 GGTGCTGATTTTCAGCACTAAAGATGCTGATAATGACAACGTATGTGCAAATGTGCCCTC 480
          |||
Sbjct 1732 GGTGCTGATTTTCAGCACTAAAGATGCTGATAATGACAACGTATGTGCAAATGTGCCCTC 1791

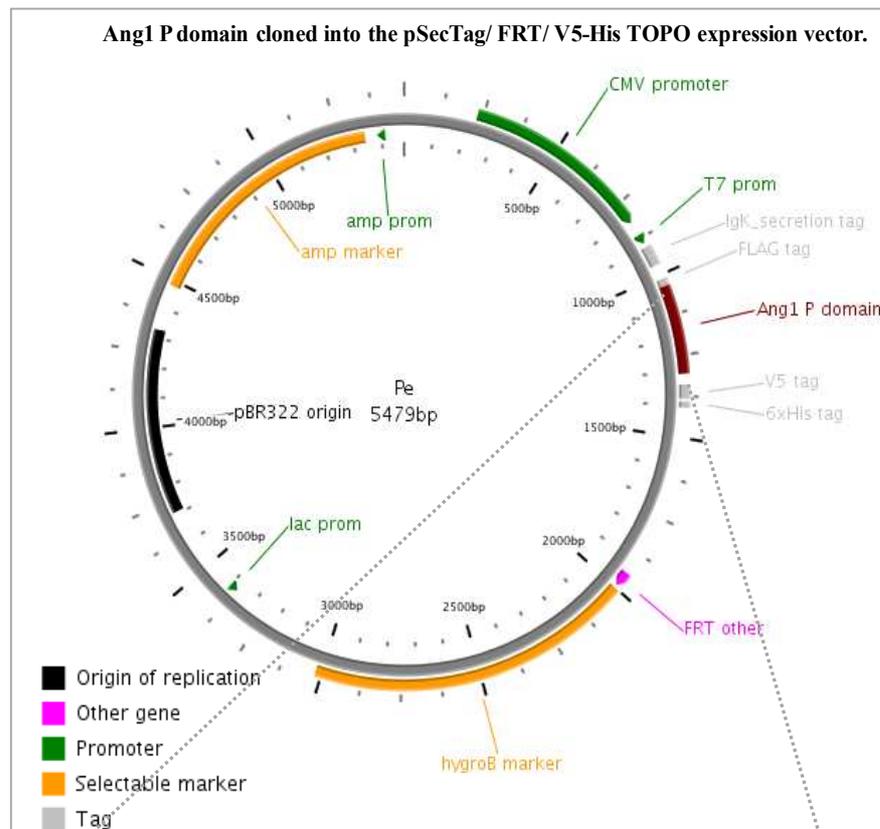
Query 481 ATGTTAACAGGAGGATGGTGGTTTGTATGCTTGTGGCCCCCTCCAATCTAAATGGAATGTTT 540
          |||
Sbjct 1792 ATGTTAACAGGAGGATGGTGGTTTGTATGCTTGTGGCCCCCTCCAATCTAAATGGAATGTTT 1851

Query 541 TATACTGCGGGACAAAACCATGGAAAACGAATGGGATAAAGTGGCACTACTTCAAAGGG 600
          |||
Sbjct 1852 TATACTGCGGGACAAAACCATGGAAAACGAATGGGATAAAGTGGCACTACTTCAAAGGG 1911

Query 601 CCCAGTTACTCCTTACGTTCCACAACATATGATGATTCGACCTTTAGATTTTGGAA 655
          |||
Sbjct 1912 CCCAGTTACTCCTTACGTTCCACAACATATGATGATTCGACCTTTAGATTTTGGAA 1966

```

**B) *Pe Sec*: Ang1 P domain cloned into the pSecTag/ FRT/ V5-His TOPO expression vector.**



```

Query 141 GGGACAGCAGGAAAACAGAGCAGCCTGATCTTACACGGTCTGATTTTCAGCACTAAAGAT 200
          |||
Sbjct 1696 GGGACAGCAGGAAAACAGAGCAGCCTGATCTTACACGGTCTGATTTTCAGCACTAAAGAT 1755

Query 201 GCTGATAATGACAACGTATGTGCAAATGTGCCCTCATGTTAACAGGAGGATGGTGGTTT 260
          |||
Sbjct 1756 GCTGATAATGACAACGTATGTGCAAATGTGCCCTCATGTTAACAGGAGGATGGTGGTTT 1815

Query 261 GATGCTTGTGGCCCCCTCCAATCTAAATGGAATGTTCTATACTGCGGGACAAAACCATGGA 320
          |||
Sbjct 1816 GATGCTTGTGGCCCCCTCCAATCTAAATGGAATGTTCTATACTGCGGGACAAAACCATGGA 1875

Query 321 AAATGAATGGGATAAAGTGGCACTACTTCAAAGGGCCCAGTTACTCCTTACGTTCCACA 380
          |||
Sbjct 1876 AAATGAATGGGATAAAGTGGCACTACTTCAAAGGGCCCAGTTACTCCTTACGTTCCACA 1935

Query 381 ACTATGATGATTCGACCTTTAGATTTTGGAAA 412
          |||
Sbjct 1936 ACTATGATGATTCGACCTTTAGATTTTGGAAA 1967

```

## Chapter 4

NOTE: Ang2 FReD cloned into pSE2 surface display vector (*F8* plasmid) was used for site directed mutagenesis to generate mutants *Asn467Ser*, *Ser480Pro* and *Gly428Thr*.

### Amino Acid sequence alignments

#### C) *Asn467Ser* mutant

```

Query  180  SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGWTIIQRREDGSVDFQRTW  359
      SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGWTIIQRREDGSVDFQRTW
Sbjct  280  SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGWTIIQRREDGSVDFQRTW  339

Query  360  KEYKVGFGNPSGEYWLGNEFVSQLTNQQRVYVKIHLKDWEGNEAYSLEYEHFYLSSEELNY  539
      KEYKVGFGNPSGEYWLGNEFVSQLTNQQRVYVKIHLKDWEGNEAYSLEYEHFYLSSEELNY
Sbjct  340  KEYKVGFGNPSGEYWLGNEFVSQLTNQQRVYVKIHLKDWEGNEAYSLEYEHFYLSSEELNY  399

Query  540  RIHLKGLTGTAGKISSISQPGNDFSTKDGNDKICICKCSQMLTGGWWFDACGPSNLNGMY  719
      RIHLKGLTGTAGKISSISQPGNDFSTKDGNDKICICKCSQMLTGGWWFDACGPSNLNGMY
Sbjct  400  RIHLKGLTGTAGKISSISQPGNDFSTKDGNDKICICKCSQMLTGGWWFDACGPSNLNGMY  459

Query  720  YPQRQNTSKFNGIKWYYWKGSGYSLKATTMMIRPADF  830
      YPQRQNT KFNGIKWYYWKGSGYSLKATTMMIRPADF
Sbjct  460  YPQRQNTNKFNGIKWYYWKGSGYSLKATTMMIRPADF  496

```

#### D) *Ser480Pro* mutant

```

Query  183  SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGWTIIQRREDGSVDFQRTW  362
      SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGWTIIQRREDGSVDFQRTW
Sbjct  280  SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGWTIIQRREDGSVDFQRTW  339

Query  363  KEYKVGFGNPSGEYWLGNEFVSQLTNQQRVYVKIHLKDWEGNEAYSLEYEHFYLSSEELNY  542
      KEYKVGFGNPSGEYWLGNEFVSQLTNQQRVYVKIHLKDWEGNEAYSLEYEHFYLSSEELNY
Sbjct  340  KEYKVGFGNPSGEYWLGNEFVSQLTNQQRVYVKIHLKDWEGNEAYSLEYEHFYLSSEELNY  399

Query  543  RIHLKGLTGTAGKISSISQPGNDFSTKDGNDKICICKCSQMLTGGWWFDACGPSNLNGMY  722
      RIHLKGLTGTAGKISSISQPGNDFSTKDGNDKICICKCSQMLTGGWWFDACGPSNLNGMY
Sbjct  400  RIHLKGLTGTAGKISSISQPGNDFSTKDGNDKICICKCSQMLTGGWWFDACGPSNLNGMY  459

Query  723  YPQRQNTNKFNGIKWYYWKGPGYSLKATTMMIRPADF  833
      YPQRQNTNKFNGIKWYYWKG GYSLKATTMMIRPADF
Sbjct  460  YPQRQNTNKFNGIKWYYWKGSGYSLKATTMMIRPADF  496

```

#### E) *Gly428Thr* mutant

```

Query  183  SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGWTIIQRREDGSVDFQRTW  362
      SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGWTIIQRREDGSVDFQRTW
Sbjct  280  SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGWTIIQRREDGSVDFQRTW  339

Query  363  KEYKVGFGNPSGEYWLGNEFVSQLTNQQRVYVKIHLKDWEGNEAYSLEYEHFYLSSEELNY  542
      KEYKVGFGNPSGEYWLGNEFVSQLTNQQRVYVKIHLKDWEGNEAYSLEYEHFYLSSEELNY
Sbjct  340  KEYKVGFGNPSGEYWLGNEFVSQLTNQQRVYVKIHLKDWEGNEAYSLEYEHFYLSSEELNY  399

Query  543  RIHLKGLTGTAGKISSISQPGNDFSTKDTDNDKICICKCSQMLTGGWWFDACGPSNLNGMY  722
      RIHLKGLTGTAGKISSISQPGNDFSTK DNDKICICKCSQMLTGGWWFDACGPSNLNGMY
Sbjct  400  RIHLKGLTGTAGKISSISQPGNDFSTKGDNDKICICKCSQMLTGGWWFDACGPSNLNGMY  459

Query  723  YPQRQNTNKFNGIKWYYWKGS  785
      YPQRQNTNKFNGIKWYYWKGS
Sbjct  460  YPQRQNTNKFNGIKWYYWKGS  480

```

**Chapter 5****F) Cloning ASGPR in the pcDNA3.1/V5-His-TOPO vector.*****Clone d***

```

Query 71  GGACCTCGCCTCCTCCTGCTCTCCCTCGGCCTCAGCCTCCTGCTGCTTGTGGTTGTCTG 130
      ||| |||||
Sbjct 522  GGA-CCTCGCCTCCTCCTGCTCTCCCTCGGCCTCAGCCTCCTGCTGCTTGTGGTTGTCTG 580

Query 131  TGTGATCGGTTCCCAAACTCCCAGCTGCAGG 162
      |||||
Sbjct 581  TGTGATCGGATCCCAAACTCCCAGCTGCAGG 612

```

***Clone f (before site directed)***

```

Query 70  GGACCTCGCCTCCTCCTGCTCTCCCTGGGCCTCAGCCTCCTGCTGCTTGTGGTTCTCTGT 129
      |||||
Sbjct 522  GGACCTCGCCTCCTCCTGCTCTCCCTGGGCCTCAGCCTCCTGCTGCTTGTGGTTGTCTGT 581

Query 130  GTGATCGGTTCCCAAACTCCCAGCTGCAGG 159
      |||||
Sbjct 582  GTGATCGGATCCCAAACTCCCAGCTGCAGG 612

```

***Clone g***

```

Query 5  GGACCTCGCCTCCTCCTGCTCCTCCCTGGGCCTCAGCCTCCTGCTGCTTGTGGTTGTC 64
      |||||
Sbjct 522  GGACCTC-GCCTCCTCCTGCT-CT-CCCTGGGCCTCAGCCTCCTGCTGCTTGTGGTTGTC 578

Query 65  TGTGATCGGTTCCCAAACTCCCAGCTGCAGG 98
      |||||
Sbjct 579  TGTGATCGGATCCCAAACTCCCAGCTGCAGG 612

```

***Clone f (after site directed mutagenesis to correct the first two mutations)***

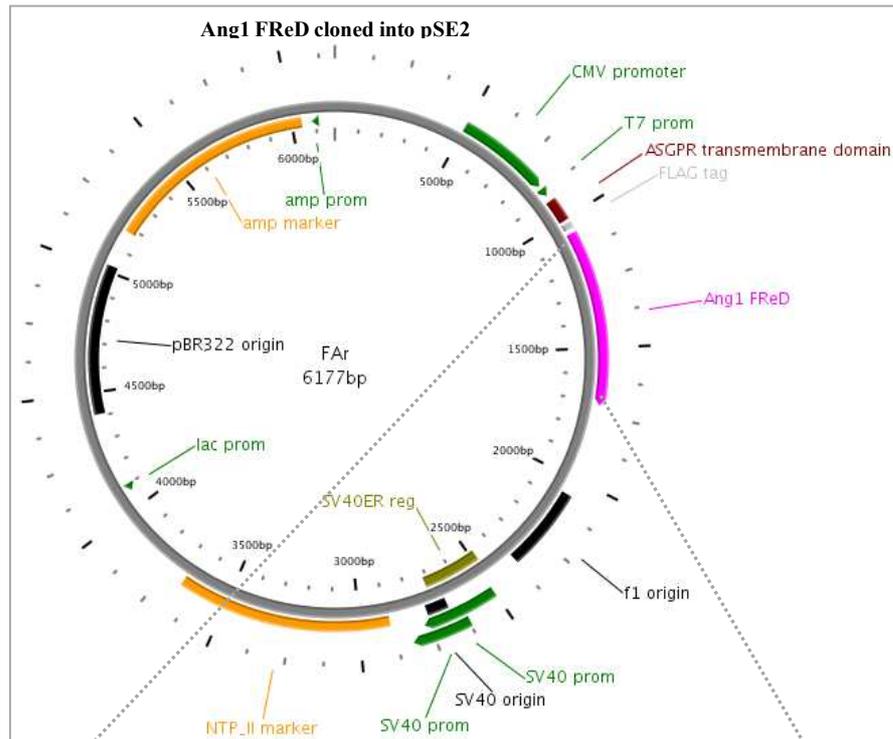
```

Query 70  GGACCTCGCCTCCTGCTGCTCTCCCTGGGCCTCAGCCTCCTGCTGCTTGTGGTTCTCTGT 129
      |||||
Sbjct 522  GGACCTCGCCTCCTGCTGCTCTCCCTGGGCCTCAGCCTCCTGCTGCTTGTGGTTGTCTGT 581

Query 130  GTGATCGGATCCCAAACTCCCAGCTGCAGG 159
      |||||
Sbjct 582  GTGATCGGATCCCAAACTCCCAGCTGCAGG 612

```

### G) *FAr*: Ang1 FReD cloned into pSE2 surface expression vector.



```

Query 1      CCATTTAGAGACTGTCAGATGTATATCAAGCTGGTTTTAATAAAAGTGAATCTACACT 60
Sbjct 1312   CCATTTAGAGACTGTCAGATGTATATCAAGCTGGTTTTAATAAAAGTGAATCTACACT 1371

Query 61      ATTTATATTAATAATATGCCAGAACCCAAAAGGTGTTTGCATATGGATGTCAATGGG 120
Sbjct 1372   ATTTATATTAATAATATGCCAGAACCCAAAAGGTGTTTGCATATGGATGTCAATGGG 1431

Query 121     GGAGTTGGACTGTAATACAACATCGTGAAGATGGAAGTCTAGATTTCCAAAGAGGCTGG 180
Sbjct 1432   GGAGTTGGACTGTAATACAACATCGTGAAGATGGAAGTCTAGATTTCCAAAGAGGCTGG 1491

Query 181     AAGGAATATAAAATGGGTTTTGGAAATCCCTCCGGTGAATATTGGCTGGGAATGAGTTT 240
Sbjct 1492   AAGGAATATAAAATGGGTTTTGGAAATCCCTCCGGTGAATATTGGCTGGGAATGAGTTT 1551

Query 241     ATTTTGGCATTACCAGTCAGAGGCAGTACATGCTAAGAATTGAGTTAATGGACTGGGAA 300
Sbjct 1552   ATTTTGGCATTACCAGTCAGAGGCAGTACATGCTAAGAATTGAGTTAATGGACTGGGAA 1611

Query 301     GGAACCGAGCCTATTCACAGTATGACAGATTCCACATAGGAAATGAAAAGCAAACCTAT 360
Sbjct 1612   GGAACCGAGCCTATTCACAGTATGACAGATTCCACATAGGAAATGAAAAGCAAACCTAT 1671

Query 361     AGGTGTATTTAAAAGGTCACACTGGGACAGCAGGAAAACAGAGCAGCCTGATCTTACAC 420
Sbjct 1672   AGGTGTATTTAAAAGGTCACACTGGGACAGCAGGAAAACAGAGCAGCCTGATCTTACAC 1731

Query 421     GGTGCTGATTTCAGCACTAAAGATGCTGATAATGACAACGTATGTGCAAATGTGCCCTC 480
Sbjct 1732   GGTGCTGATTTCAGCACTAAAGATGCTGATAATGACAACGTATGTGCAAATGTGCCCTC 1791

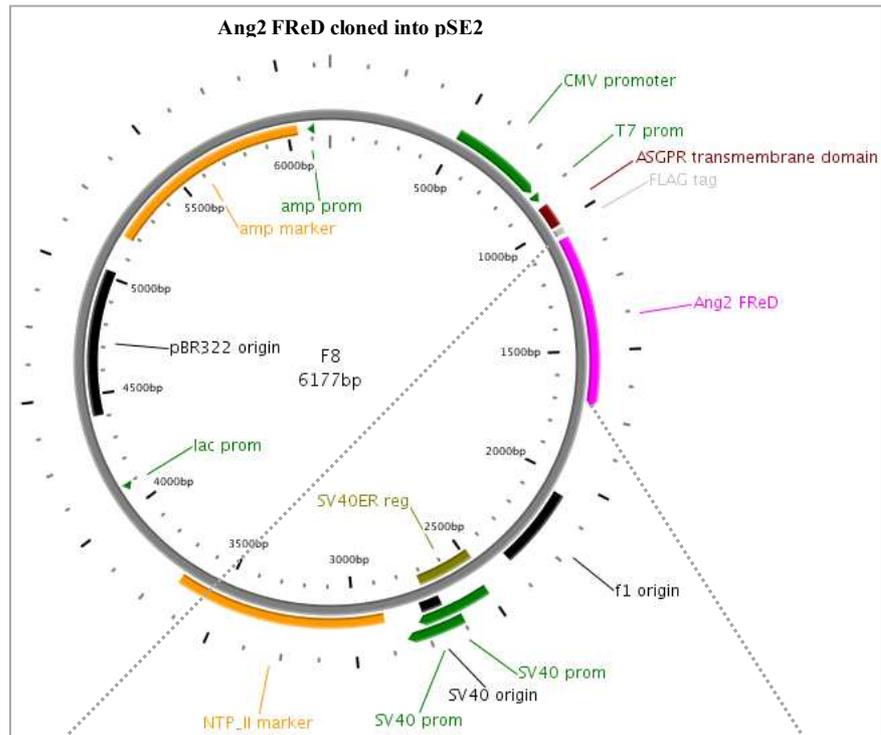
Query 481     ATGTTAACAGGAGGATGGTGGTTTGATGCTTGTGGCCCCCAATCTAAATGGAATGTTT 540
Sbjct 1792   ATGTTAACAGGAGGATGGTGGTTTGATGCTTGTGGCCCCCAATCTAAATGGAATGTTT 1851

Query 541     TATACTGCGGGACAAAACCATGGAACCTGAATGGGATAAAGTGGCACTACTTCAAAGGG 600
Sbjct 1852   TATACTGCGGGACAAAACCATGGAACCTGAATGGGATAAAGTGGCACTACTTCAAAGGG 1911

Query 601     CCCAGTTACTCCTTACGTTCCACAACATGATGATTCGACCTTTAGATTTTGGAA 655
Sbjct 1912   CCCAGTTACTCCTTACGTTCCACAACATGATGATTCGACCTTTAGATTTTGGAA 1966

```

### H) F8: Ang2 FReD cloned into pSE2 surface expression vector.



```

Query 1      CAGCTTCAGAGACTGTGCTGAAGTATTCAAATCAGGACACACCAGAATGGCATCTACAC 60
             |||
Sbjct 1010   CAGCTTCAGAGACTGTGCTGAAGTATTCAAATCAGGACACACCAGAATGGCATCTACAC 1069

Query 61     GTTAACATTCCTAATTCTACAGAAGAGATCAAGGCCTACTGTGACATGGAAGCTGGAGG 120
             |||
Sbjct 1070   GTTAACATTCCTAATTCTACAGAAGAGATCAAGGCCTACTGTGACATGGAAGCTGGAGG 1129

Query 121    AGGCGGGTGACAATTATTTCAGCGACGTGAGGATGGCAGCGTTGATTTTCAGAGGACTTG 180
             |||
Sbjct 1130   AGGCGGGTGACAATTATTTCAGCGACGTGAGGATGGCAGCGTTGATTTTCAGAGGACTTG 1189

Query 181    GAAAGAATATAAAGTGGGATTTGGTAACCCTTCAGGAGAATATTGGCTGGGAAATGAGTT 240
             |||
Sbjct 1190   GAAAGAATATAAAGTGGGATTTGGTAACCCTTCAGGAGAATATTGGCTGGGAAATGAGTT 1249

Query 241    TGTTTCGCAACTGACTAATCAGCAACGCTATGTGCTTAAAATACACCTTAAAGACTGGGA 300
             |||
Sbjct 1250   TGTTTCGCAACTGACTAATCAGCAACGCTATGTGCTTAAAATACACCTTAAAGACTGGGA 1309

Query 301    AGGGAATGAGGCTTACTCATTGTATGAACATTTCTATCTCTCAAGTGAAGAACTCAATTA 360
             |||
Sbjct 1310   AGGGAATGAGGCTTACTCATTGTATGAACATTTCTATCTCTCAAGTGAAGAACTCAATTA 1369

Query 361    TAGGATTCACCTTAAAGGACTTACAGGGACAGCCGGCAAATAAGCAGCATCAGCCAACC 420
             |||
Sbjct 1370   TAGGATTCACCTTAAAGGACTTACAGGGACAGCCGGCAAATAAGCAGCATCAGCCAACC 1429

Query 421    AGGAAATGATTTTAGCACAAAGGATGGAGACAACGACAAATGTATTTGCAAATGTTTACA 480
             |||
Sbjct 1430   AGGAAATGATTTTAGCACAAAGGATGGAGACAACGACAAATGTATTTGCAAATGTTTACA 1489

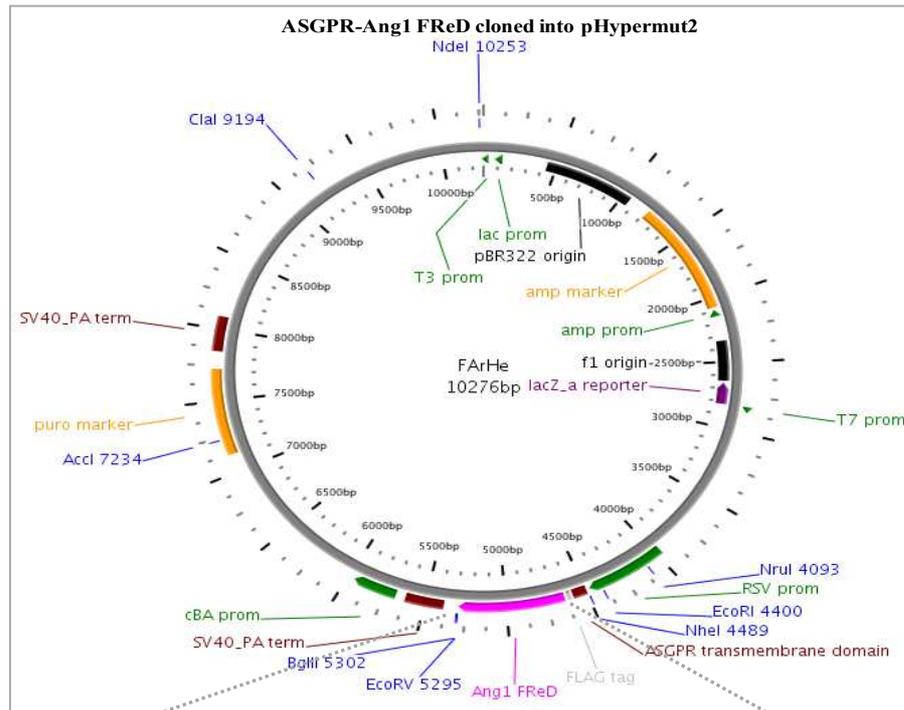
Query 481    AATGCTAACAGGAGGCTGGTGGTTTGTATGCATGTGGTCCTTCCAACCTGAAACGGAATGTA 540
             |||
Sbjct 1490   AATGCTAACAGGAGGCTGGTGGTTTGTATGCATGTGGTCCTTCCAACCTGAAACGGAATGTA 1549

Query 541    CTATCCACAGAGGCAGAACACAAATAAGTTCAACGGCATTAATGGTACTACTGGAAAGG 600
             |||
Sbjct 1550   CTATCCACAGAGGCAGAACACAAATAAGTTCAACGGCATTAATGGTACTACTGGAAAGG 1609

Query 601    CTCAGGCTATTCGCTCAAGGCCACAACCATGATGATCCGACCAGCAGATTTCTAA 655
             |||
Sbjct 1610   CTCAGGCTATTCGCTCAAGGCCACAACCATGATGATCCGACCAGCAGATTTCTAA 1664

```

### I) *FArHe*: ASGPR-Ang1 FReD cloned into pHypermut2 vector.



```

Query 1      CCATTTAGAGACTGTGCAGATGTATATCAAGCTGGTTTTAATAAAAGTGAATCTACACT 60
Sbjct 1312    CCATTTAGAGACTGTGCAGATGTATATCAAGCTGGTTTTAATAAAAGTGAATCTACACT 1371

Query 61      ATTTATATTAATAATATGCCAGAACCCAAAAGGTGTTTTGCAATATGGATGTCAATGGG 120
Sbjct 1372    ATTTATATTAATAATATGCCAGAACCCAAAAGGTGTTTTGCAATATGGATGTCAATGGG 1431

Query 121     GGAGGTTGGACTGTAATACAACATCGTGAAGATGGAAGCTAGATTTCCAAAGAGGCTGG 180
Sbjct 1432    GGAGGTTGGACTGTAATACAACATCGTGAAGATGGAAGCTAGATTTCCAAAGAGGCTGG 1491

Query 181     AAGGAATATAAAATGGGTTTTGGAATCCCTCCGGTGAATATTGGCTGGGGAATGAGTTT 240
Sbjct 1492    AAGGAATATAAAATGGGTTTTGGAATCCCTCCGGTGAATATTGGCTGGGGAATGAGTTT 1551

Query 241     ATTTTGGCATTACCAGTCAGAGGCAGTACATGCTAAGAATTGAGTTAATGGACTGGGAA 300
Sbjct 1552    ATTTTGGCATTACCAGTCAGAGGCAGTACATGCTAAGAATTGAGTTAATGGACTGGGAA 1611

Query 301     GGAACCGAGCCTATTCACAGTATGACAGATTCCACATAGGAAATGAAAAGCAAACTAT 360
Sbjct 1612    GGAACCGAGCCTATTCACAGTATGACAGATTCCACATAGGAAATGAAAAGCAAACTAT 1671

Query 361     AGGTGTATTTAAAAGGTCACACTGGGACAGCAGGAAAACAGAGCAGCCTGATCTTACAC 420
Sbjct 1672    AGGTGTATTTAAAAGGTCACACTGGGACAGCAGGAAAACAGAGCAGCCTGATCTTACAC 1731

Query 421     GGTGCTGATTTCAGCACTAAAGATGCTGATAATGACAACGTATGTGCAAATGTGCCCTC 480
Sbjct 1732    GGTGCTGATTTCAGCACTAAAGATGCTGATAATGACAACGTATGTGCAAATGTGCCCTC 1791

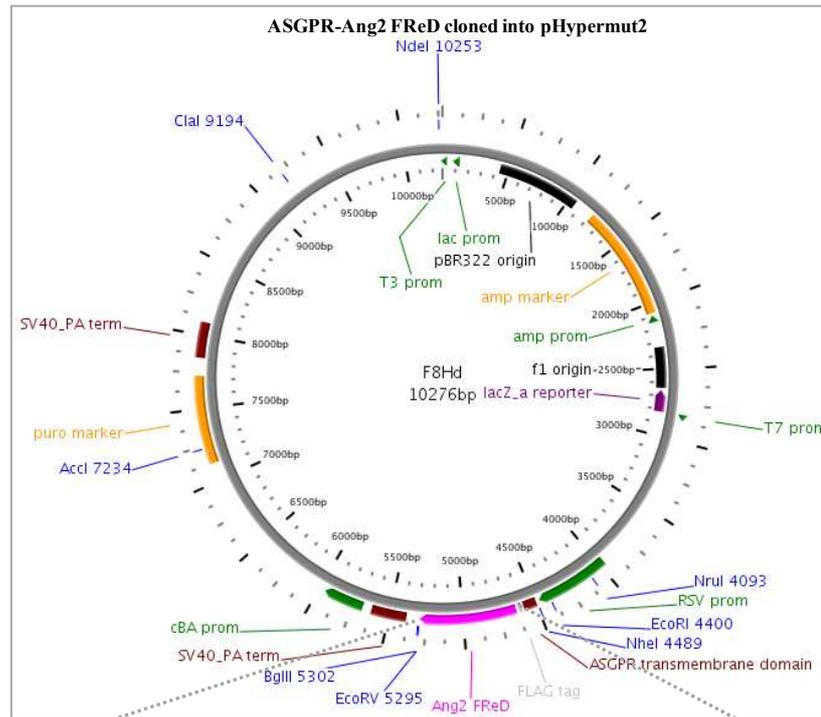
Query 481     ATGTTAACAGGAGGATGGTGGTTTGATGCTTGTGGCCCCCAATCTAAATGGAATGTTT 540
Sbjct 1792    ATGTTAACAGGAGGATGGTGGTTTGATGCTTGTGGCCCCCAATCTAAATGGAATGTTT 1851

Query 541     TATACTGCGGGACAAAACCATGGAAGTGAATGGGATAAAGTGGCACTACTTCAAAGGG 600
Sbjct 1852    TATACTGCGGGACAAAACCATGGAAGTGAATGGGATAAAGTGGCACTACTTCAAAGGG 1911

Query 601     CCCAGTTACTCCTTACGTTCCACAACATGATGATTCGACCTTTAGATTTTGGAA 655
Sbjct 1912    CCCAGTTACTCCTTACGTTCCACAACATGATGATTCGACCTTTAGATTTTGGAA 1966

```

### J) *F8Hd*: ASGPR-Ang2 FReD cloned into pHypermut2



```

Query 102 CAGCTTCAGAGACTGTGCTGAAGTATTCAAATCAGGACACACCACGAATGGCATCTACAC 161
Sbjct 1010 CAGCTTCAGAGACTGTGCTGAAGTATTCAAATCAGGACACACCACGAATGGCATCTACAC 1069

Query 162 GTTAACATTCCTAATTCTACAGAAGAGATCAAGGCCTACTGTGACATGGAAGCTGGAGG 221
Sbjct 1070 GTTAACATTCCTAATTCTACAGAAGAGATCAAGGCCTACTGTGACATGGAAGCTGGAGG 1129

Query 222 AGGCGGGTGGACAATTATTCAGCGACGTGAGGATGGCAGCGTTGATTTTCAGAGGACTTG 281
Sbjct 1130 AGGCGGGTGGACAATTATTCAGCGACGTGAGGATGGCAGCGTTGATTTTCAGAGGACTTG 1189

Query 282 GAAAGAATATAAAGTGGGATTTGGTAACCCTTCAGGAGAATATTGGCTGGGAAATGAGTT 341
Sbjct 1190 GAAAGAATATAAAGTGGGATTTGGTAACCCTTCAGGAGAATATTGGCTGGGAAATGAGTT 1249

Query 342 TGTTTCGCAACTGACTAATCAGCAACGCTATGTGCTTAAAATACACCTTAAAGACTGGGA 401
Sbjct 1250 TGTTTCGCAACTGACTAATCAGCAACGCTATGTGCTTAAAATACACCTTAAAGACTGGGA 1309

Query 402 AGGGAATGAGGCTTACTCATTGTATGAACATTTCTATCTCTCAAGTGAAGAACTCAATTA 461
Sbjct 1310 AGGGAATGAGGCTTACTCATTGTATGAACATTTCTATCTCTCAAGTGAAGAACTCAATTA 1369

Query 462 TAGGATTCACCTTAAAGGACTTACAGGGACAGCCGGCAAAAATAAGCAGCATCAGCCAACC 521
Sbjct 1370 TAGGATTCACCTTAAAGGACTTACAGGGACAGCCGGCAAAAATAAGCAGCATCAGCCAACC 1429

Query 522 AGGAAATGATTTTAGCACAAAGGATGGAGACAACGACAAATGTATTTGCAAATGTTTACA 581
Sbjct 1430 AGGAAATGATTTTAGCACAAAGGATGGAGACAACGACAAATGTATTTGCAAATGTTTACA 1489

Query 582 AATGCTAACAGGAGGCTGGTGGTTTGATGCATGTGGTCCTTCCAACCTTGAACGGAATGTA 641
Sbjct 1490 AATGCTAACAGGAGGCTGGTGGTTTGATGCATGTGGTCCTTCCAACCTTGAACGGAATGTA 1549

Query 642 CTATCCACAGAGGCAGAACACAAATAAGTTCAACGGCATTAAATGGTACTACTGGAAAGG 701
Sbjct 1550 CTATCCACAGAGGCAGAACACAAATAAGTTCAACGGCATTAAATGGTACTACTGGAAAGG 1609

Query 702 CTCAGGCTATTCGCTCAAGGCCACAACCATGATGATCCGACCAGCAGATTTCTAA 756
Sbjct 1610 CTCAGGCTATTCGCTCAAGGCCACAACCATGATGATCCGACCAGCAGATTTCTAA 1664

```

**K) *Hm5* amino acid sequence alignment**

Query	2	SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEARGGGGWTIIQRREDGSVDFQRTW	181
		SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEA GGGWTIIQRREDGSVDFQRTW	
Sbjct	280	SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGGWTIIQRREDGSVDFQRTW	339
Query	182	KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	361
		KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	
Sbjct	340	KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	399
Query	362	RIHLKGLTGTAGKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	541
		RIHLKGLTGTAGKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	
Sbjct	400	RIHLKGLTGTAGKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	459
Query	542	YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPADF	652
		YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPADF	
Sbjct	460	YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPADF	496

**L) *10c* amino acid sequence alignment**

Query	2	SFRDCAEVFKSGHTTNGIYTLTFRNSTEEIKAYCDMEAGGGGGWTIIQRREDGSVDFQRTW	181
		SFRDCAEVFKSGHTTNGIYTLTF NSTEEIKAYCDMEAGGGGGWTIIQRREDGSVDFQRTW	
Sbjct	280	SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGGWTIIQRREDGSVDFQRTW	339
Query	182	KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	361
		KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	
Sbjct	340	KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	399
Query	362	RIHLKGLTGTAGKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	541
		RIHLKGLTGTAGKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	
Sbjct	400	RIHLKGLTGTAGKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	459
Query	542	YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPADF	652
		YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPADF	
Sbjct	460	YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPADF	496

**M) *FACSD* amino acid sequence alignment**

Query	2	SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGGWTIIQRREDGSVDFQRTW	181
		SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGGWTIIQRREDGSVDFQRTW	
Sbjct	280	SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGGWTIIQRREDGSVDFQRTW	339
Query	182	KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	361
		KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	
Sbjct	340	KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	399
Query	362	RIHLKGLTGTAGKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	541
		RIHLKGLTGTAGKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	
Sbjct	400	RIHLKGLTGTAGKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	459
Query	542	YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPANF	652
		YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPA F	
Sbjct	460	YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPADF	496

**N) *FACSi* amino acid sequence alignment**

Query	2	SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGGWTIIQRREDGSVDFQRTW	181
		SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGGWTIIQRREDGSVDFQRTW	
Sbjct	28	SFRDCAEVFKSGHTTNGIYTLTFPNSTEEIKAYCDMEAGGGGGWTIIQRREDGSVDFQRTW	87
Query	182	KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	361
		KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	
Sbjct	88	KEYKVGFGNPSGEYWLGNFVSQLTNQQRVVLKIHLLKDWEGNEAYSLEYEHFYLSSSEELNY	147
Query	362	RIHLKGLTGTADKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	541
		RIHLKGLTGTAKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	
Sbjct	148	RIHLKGLTGTAGKISSISQPGNDFSTKGDNDKCIKCSQMLTGGWWFDACGPSNLNGMY	207
Query	542	YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPADF	652
		YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPADF	
Sbjct	208	YPQRQNTNKFNGIKWYYWKGSGYSLKATMMIRPADF	244

## References

- Aharoni, A., Gaidukov, L., Khersonsky, O., McQ Gould, S., Roodveldt, C. & Tawfik, D.S. 2005, "The 'evolvability' of promiscuous protein functions", *Nature genetics*, vol. 37, no. 1, pp. 73-76.
- Arakawa, H. & Buerstedde, J.M. 2009, "Activation-induced cytidine deaminase-mediated hypermutation in the DT40 cell line", *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 364, no. 1517, pp. 639-644.
- Arakawa, H., Hauschild, J. & Buerstedde, J.M. 2002, "Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion", *Science (New York, N.Y.)*, vol. 295, no. 5558, pp. 1301-1306.
- Arakawa, H., Kudo, H., Batrak, V., Caldwell, R.B., Rieger, M.A., Ellwart, J.W. & Buerstedde, J.M. 2008, "Protein evolution by hypermutation and selection in the B cell line DT40", *Nucleic acids research*, vol. 36, no. 1, pp. e1.
- Arnold, F.H. & Georgiou, G. (eds) 2003, *Directed Enzyme Evolution: Screening and Selection methods.*, Humana Press, Totowa, New Jersey.
- Arnold, F.H. & Georgiou, G. (eds) 2003, *Directed Evolution: Library Creation: Methods and Protocols*, Humana Press, Totowa, New Jersey.
- Aubin, J.E. 1979, "Autofluorescence of viable cultured mammalian cells", *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society*, vol. 27, no. 1, pp. 36-43.
- Axe, D.D., Foster, N.W. & Fersht, A.R. 1998, "A search for single substitutions that eliminate enzymatic function in a bacterial ribonuclease", *Biochemistry*, vol. 37, no. 20, pp. 7157-7166.
- Bachl, J., Carlson, C., Gray-Schopfer, V., Dessing, M. & Olsson, C. 2001, "Increased transcription levels induce higher mutation rates in a hypermutating cell line", *Journal of immunology (Baltimore, Md.: 1950)*, vol. 166, no. 8, pp. 5051-5057.
- Bachl, J. & Olsson, C. 1999, "Hypermutation targets a green fluorescent protein-encoding transgene in the presence of immunoglobulin enhancers", *European journal of immunology*, vol. 29, no. 4, pp. 1383-1389.
- Baffert, F., Le, T., Thurston, G. & McDonald, D.M. 2006, "Angiopoietin-1 decreases plasma leakage by reducing number and size of endothelial gaps in venules", *American journal of physiology. Heart and circulatory physiology*, vol. 290, no. 1, pp. H107-18.
- Baneyx, F. 1999, "Recombinant protein expression in Escherichia coli", *Current opinion in biotechnology*, vol. 10, no. 5, pp. 411-421.

- Barton, W.A., Tzvetkova, D. & Nikolov, D.B. 2005, "Structure of the angiopoietin-2 receptor binding domain and identification of surfaces involved in Tie2 recognition", *Structure (London, England : 1993)*, vol. 13, no. 5, pp. 825-832.
- Barton, W.A., Tzvetkova-Robev, D., Miranda, E.P., Kolev, M.V., Rajashankar, K.R., Himanen, J.P. & Nikolov, D.B. 2006, "Crystal structures of the Tie2 receptor ectodomain and the angiopoietin-2-Tie2 complex", *Nature structural & molecular biology*, vol. 13, no. 6, pp. 524-532.
- Beltzer, J.P., Fiedler, K., Fuhrer, C., Geffen, I., Handschin, C., Wessels, H.P. & Spiess, M. 1991, "Charged residues are major determinants of the transmembrane orientation of a signal-anchor sequence", *The Journal of biological chemistry*, vol. 266, no. 2, pp. 973-978.
- Benson, D.E., Wisz, M.S. & Hellinga, H.W. 1998, "The development of new biotechnologies using metalloprotein design", *Current opinion in biotechnology*, vol. 9, no. 4, pp. 370-376.
- Bernard, O., Hozumi, N. & Tonegawa, S. 1978, "Sequences of mouse immunoglobulin light chain genes before and after somatic changes", *Cell*, vol. 15, no. 4, pp. 1133-1144.
- Bershtein, S., Goldin, K. & Tawfik, D.S. 2008, "Intense neutral drifts yield robust and evolvable consensus proteins", *Journal of Molecular Biology*, vol. 379, no. 5, pp. 1029-1044.
- Betts, M.J. & Russell, R.B. 2003, "Bioinformatics for Geneticists. Amino acid properties and consequences of substitutions." in , eds. M.R. Barnes & I.C. Gray, Wiley, .
- Betz, A.G., Rada, C., Pannell, R., Milstein, C. & Neuberger, M.S. 1993, "Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hot spots", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 6, pp. 2385-2388.
- Bloom, J.D., Raval, A. & Wilke, C.O. 2007, "Thermodynamics of neutral protein evolution", *Genetics*, vol. 175, no. 1, pp. 255-266.
- Boder, E.T. & Wittrup, K.D. 1998, "Optimal screening of surface-displayed polypeptide libraries", *Biotechnology progress*, vol. 14, no. 1, pp. 55-62.
- Bogdanovic, E., Nguyen, V.P. & Dumont, D.J. 2006, "Activation of Tie2 by angiopoietin-1 and angiopoietin-2 results in their release and receptor internalization", *Journal of cell science*, vol. 119, no. Pt 17, pp. 3551-3560.
- Bolon, D.N., Voigt, C.A. & Mayo, S.L. 2002, "De novo design of biocatalysts", *Current opinion in chemical biology*, vol. 6, no. 2, pp. 125-129.
- Brenner, S.E. & Levitt, M. 2000, "Expectations from structural genomics", *Protein science : a publication of the Protein Society*, vol. 9, no. 1, pp. 197-200.

- Brindle, N.P., Saharinen, P. & Alitalo, K. 2006, "Signaling and functions of angiopoietin-1 in vascular protection", *Circulation research*, vol. 98, no. 8, pp. 1014-1023.
- Brizzard, B.L., Chubet, R.G. & Vizard, D.L. 1994, "Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution", *BioTechniques*, vol. 16, no. 4, pp. 730-735.
- Carlson, T.R., Feng, Y., Maisonpierre, P.C., Mrksich, M. & Morla, A.O. 2001, "Direct cell adhesion to the angiopoietins mediated by integrins", *The Journal of biological chemistry*, vol. 276, no. 28, pp. 26516-26525.
- Chaudhuri, J., Tian, M., Khuong, C., Chua, K., Pinaud, E. & Alt, F.W. 2003, "Transcription-targeted DNA deamination by the AID antibody diversification enzyme", *Nature*, vol. 422, no. 6933, pp. 726-730.
- Cho, C.H., Kammerer, R.A., Lee, H.J., Steinmetz, M.O., Ryu, Y.S., Lee, S.H., Yasunaga, K., Kim, K.T., Kim, I., Choi, H.H., Kim, W., Kim, S.H., Park, S.K., Lee, G.M. & Koh, G.Y. 2004, "COMP-Ang1: a designed angiopoietin-1 variant with nonleaky angiogenic activity", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 15, pp. 5547-5552.
- Cho, C.H., Kammerer, R.A., Lee, H.J., Yasunaga, K., Kim, K.T., Choi, H.H., Kim, W., Kim, S.H., Park, S.K., Lee, G.M. & Koh, G.Y. 2004, "Designed angiopoietin-1 variant, COMP-Ang1, protects against radiation-induced endothelial cell apoptosis", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 15, pp. 5553-5558.
- Cho, C.H., Kim, K.E., Byun, J., Jang, H.S., Kim, D.K., Baluk, P., Baffert, F., Lee, G.M., Mochizuki, N., Kim, J., Jeon, B.H., McDonald, D.M. & Koh, G.Y. 2005, "Long-term and sustained COMP-Ang1 induces long-lasting vascular enlargement and enhanced blood flow", *Circulation research*, vol. 97, no. 1, pp. 86-94.
- Clark, L.A., Ganesan, S., Papp, S. & van Vlijmen, H.W. 2006, "Trends in antibody sequence changes during the somatic hypermutation process", *Journal of immunology (Baltimore, Md.: 1950)*, vol. 177, no. 1, pp. 333-340.
- Coia, G., Ayres, A., Lilley, G.G., Hudson, P.J. & Irving, R.A. 1997, "Use of mutator cells as a means for increasing production levels of a recombinant antibody directed against Hepatitis B", *Gene*, vol. 201, no. 1-2, pp. 203-209.
- Crasto, C.J. & Feng, J.A. 2000, "LINKER: a program to generate linker sequences for fusion proteins", *Protein engineering*, vol. 13, no. 5, pp. 309-312.
- Daly, C., Pasnikowski, E., Burova, E., Wong, V., Aldrich, T.H., Griffiths, J., Ioffe, E., Daly, T.J., Fandl, J.P., Papadopoulos, N., McDonald, D.M., Thurston, G., Yancopoulos, G.D. & Rudge, J.S. 2006, "Angiopoietin-2 functions as an autocrine protective factor in stressed endothelial cells", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 42, pp. 15491-15496.

- Daly, C., Wong, V., Burova, E., Wei, Y., Zabski, S., Griffiths, J., Lai, K.M., Lin, H.C., Ioffe, E., Yancopoulos, G.D. & Rudge, J.S. 2004, "Angiopoietin-1 modulates endothelial cell function and gene expression via the transcription factor FKHR (FOXO1)", *Genes & development*, vol. 18, no. 9, pp. 1060-1071.
- Davis, S., Aldrich, T.H., Jones, P.F., Acheson, A., Compton, D.L., Jain, V., Ryan, T.E., Bruno, J., Radziejewski, C., Maisonpierre, P.C. & Yancopoulos, G.D. 1996, "Isolation of angiopoietin-1, a ligand for the TIE2 receptor, by secretion-trap expression cloning", *Cell*, vol. 87, no. 7, pp. 1161-1169.
- Davis, S., Papadopoulos, N., Aldrich, T.H., Maisonpierre, P.C., Huang, T., Kovac, L., Xu, A., Leidich, R., Radziejewska, E., Rafique, A., Goldberg, J., Jain, V., Bailey, K., Karow, M., Fandl, J., Samuelsson, S.J., Ioffe, E., Rudge, J.S., Daly, T.J., Radziejewski, C. & Yancopoulos, G.D. 2003, "Angiopoietins have distinct modular domains essential for receptor binding, dimerization and superclustering", *Nature structural biology*, vol. 10, no. 1, pp. 38-44.
- DeBusk, L.M., Hallahan, D.E. & Lin, P.C. 2004, "Akt is a major angiogenic mediator downstream of the Ang1/Tie2 signaling pathway", *Experimental cell research*, vol. 298, no. 1, pp. 167-177.
- DeLano, W.L. 2002, , *The PyMOL Molecular Graphics System* [Homepage of DeLano Scientific, San Carlos, CA, USA.], [Online]. Available: <http://www.pymol.org>.
- Devos, D. & Valencia, A. 2000, "Practical limits of function prediction", *Proteins*, vol. 41, no. 1, pp. 98-107.
- Di Noia, J.M. & Neuberger, M.S. 2007, "Molecular mechanisms of antibody somatic hypermutation", *Annual Review of Biochemistry*, vol. 76, pp. 1-22.
- Dill, K.A. 1990, "Dominant forces in protein folding", *Biochemistry*, vol. 29, no. 31, pp. 7133-7155.
- Dong, X., Stothard, P., Forsythe, I.J. & Wishart, D.S. 2004, "PlasMapper: a web server for drawing and auto-annotating plasmid maps", *Nucleic acids research*, vol. 32, no. Web Server issue, pp. W660-4.
- Drake, C.J. 2003, "Embryonic and adult vasculogenesis", *Birth defects research.Part C, Embryo today : reviews*, vol. 69, no. 1, pp. 73-82.
- Drummond, D.A., Iverson, B.L., Georgiou, G. & Arnold, F.H. 2005, "Why high-error-rate random mutagenesis libraries are enriched in functional and improved proteins", *Journal of Molecular Biology*, vol. 350, no. 4, pp. 806-816.
- Dumont, D.J., Yamaguchi, T.P., Conlon, R.A., Rossant, J. & Breitman, M.L. 1992, "Tek, a Novel Tyrosine Kinase Gene Located on Mouse Chromosome 4, is Expressed in Endothelial Cells and their Presumptive Precursors", *Oncogene*, vol. 7, no. 8, pp. 1471-1480.

- Durocher, Y. & Butler, M. 2009, "Expression systems for therapeutic glycoprotein production", *Current opinion in biotechnology*, vol. 20, no. 6, pp. 700-707.
- England, J.L. & Shakhnovich, E.I. 2003, "Structural determinant of protein designability", *Physical Review Letters*, vol. 90, no. 21, pp. 218101.
- Fairlie, D.P., West, M.L. & Wong, A.K. 1998, "Towards protein surface mimetics", *Current medicinal chemistry*, vol. 5, no. 1, pp. 29-62.
- Feldhaus, M.J., Siegel, R.W., Opresko, L.K., Coleman, J.R., Feldhaus, J.M., Yeung, Y.A., Cochran, J.R., Heinzelman, P., Colby, D., Swers, J., Graff, C., Wiley, H.S. & Wittrup, K.D. 2003, "Flow-cytometric isolation of human antibodies from a nonimmune *Saccharomyces cerevisiae* surface display library", *Nature biotechnology*, vol. 21, no. 2, pp. 163-170.
- Fiedler, U. & Augustin, H.G. 2006, "Angiopoietins: a link between angiogenesis and inflammation", *Trends in immunology*, vol. 27, no. 12, pp. 552-558.
- Fiedler, U., Krissl, T., Koidl, S., Weiss, C., Koblizek, T., Deutsch, U., Martiny-Baron, G., Marme, D. & Augustin, H.G. 2003, "Angiopoietin-1 and angiopoietin-2 share the same binding domains in the Tie-2 receptor involving the first Ig-like loop and the epidermal growth factor-like repeats", *The Journal of biological chemistry*, vol. 278, no. 3, pp. 1721-1727.
- Fiedler, U., Scharpfenecker, M., Koidl, S., Hegen, A., Grunow, V., Schmidt, J.M., Kriz, W., Thurston, G. & Augustin, H.G. 2004, "The Tie-2 ligand angiopoietin-2 is stored in and rapidly released upon stimulation from endothelial cell Weibel-Palade bodies", *Blood*, vol. 103, no. 11, pp. 4150-4156.
- Frokjaer, S. & Otzen, D.E. 2005, "Protein drug stability: a formulation challenge", *Nature reviews. Drug discovery*, vol. 4, no. 4, pp. 298-306.
- Gale, N.W., Thurston, G., Hackett, S.F., Renard, R., Wang, Q., McClain, J., Martin, C., Witte, C., Witte, M.H., Jackson, D., Suri, C., Campochiaro, P.A., Wiegand, S.J. & Yancopoulos, G.D. 2002, "Angiopoietin-2 is required for postnatal angiogenesis and lymphatic patterning, and only the latter role is rescued by Angiopoietin-1", *Developmental cell*, vol. 3, no. 3, pp. 411-423.
- Glasner, M.E., Gerlt, J.A. & Babbitt, P.C. 2007, "Mechanisms of protein evolution and their application to protein engineering", *Advances in Enzymology and Related Areas of Molecular Biology*, vol. 75, pp. 193-239, xii-xiii.
- Goder, V., Bieri, C. & Spiess, M. 1999, "Glycosylation can influence topogenesis of membrane proteins and reveals dynamic reorientation of nascent polypeptides within the translocon", *The Journal of cell biology*, vol. 147, no. 2, pp. 257-266.
- Goder, V. & Spiess, M. 2001, "Topogenesis of membrane proteins: determinants and dynamics", *FEBS letters*, vol. 504, no. 3, pp. 87-93.

- Goyenechea, B. & Milstein, C. 1996, "Modifying the sequence of an immunoglobulin V-gene alters the resulting pattern of hypermutation", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 24, pp. 13979-13984.
- Guo, H.H., Choe, J. & Loeb, L.A. 2004, "Protein tolerance to random amino acid change", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 25, pp. 9205-9210.
- Gupta, R.D. & Tawfik, D.S. 2008, "Directed enzyme evolution via small and effective neutral drift libraries", *Nature methods*, vol. 5, no. 11, pp. 939-942.
- Heldin, C.H., Ertlund, A., Rorsman, C. & Ronnstrand, L. 1989, "Dimerization of B-type platelet-derived growth factor receptors occurs after ligand binding and is closely associated with receptor kinase activation", *The Journal of biological chemistry*, vol. 264, no. 15, pp. 8905-8912.
- Hellinga, H.W. 1997, "Rational protein design: combining theory and experiment", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 19, pp. 10015-10017.
- Henke, W., Herdel, K., Jung, K., Schnorr, D. & Loening, S.A. 1997, "Betaine improves the PCR amplification of GC-rich DNA sequences", *Nucleic acids research*, vol. 25, no. 19, pp. 3957-3958.
- Higuchi, M., Oh-eda, M., Kuboniwa, H., Tomonoh, K., Shimonaka, Y. & Ochi, N. 1992, "Role of sugar chains in the expression of the biological activity of human erythropoietin", *The Journal of biological chemistry*, vol. 267, no. 11, pp. 7703-7709.
- Hill, D.E., Oliphant, A.R. & Struhl, K. 1987, "Mutagenesis with degenerate oligonucleotides: an efficient method for saturating a defined DNA region with base pair substitutions", *Methods in enzymology*, vol. 155, pp. 558-568.
- Holland, E.C. & Drickamer, K. 1986, "Signal recognition particle mediates the insertion of a transmembrane protein which has a cytoplasmic NH<sub>2</sub> terminus", *The Journal of biological chemistry*, vol. 261, no. 3, pp. 1286-1292.
- Hossler, P., Khattak, S.F. & Li, Z.J. 2009, "Optimal and consistent protein glycosylation in mammalian cell culture", *Glycobiology*, vol. 19, no. 9, pp. 936-949.
- Hu, B., Jarzynka, M.J., Guo, P., Imanishi, Y., Schlaepfer, D.D. & Cheng, S.Y. 2006, "Angiopoietin 2 induces glioma cell invasion by stimulating matrix metalloprotease 2 expression through the  $\alpha$ v $\beta$ 1 integrin and focal adhesion kinase signaling pathway", *Cancer research*, vol. 66, no. 2, pp. 775-783.
- Hughes, D.P., Marron, M.B. & Brindle, N.P. 2003, "The antiinflammatory endothelial tyrosine kinase Tie2 interacts with a novel nuclear factor- $\kappa$ B inhibitor ABIN-2", *Circulation research*, vol. 92, no. 6, pp. 630-636.

- Keefe, A.D. & Szostak, J.W. 2001, "Functional proteins from a random-sequence library", *Nature*, vol. 410, no. 6829, pp. 715-718.
- Kelsoe, G. 1996, "The germinal center: a crucible for lymphocyte selection", *Seminars in immunology*, vol. 8, no. 3, pp. 179-184.
- Kieke, M.C., Cho, B.K., Boder, E.T., Kranz, D.M. & Wittrup, K.D. 1997, "Isolation of anti-T cell receptor scFv mutants by yeast surface display", *Protein engineering*, vol. 10, no. 11, pp. 1303-1310.
- Kim, H.Z., Jung, K., Kim, H.M., Cheng, Y. & Koh, G.Y. 2009, "A designed angiopoietin-2 variant, pentameric COMP-Ang2, strongly activates Tie2 receptor and stimulates angiogenesis", *Biochimica et biophysica acta*, vol. 1793, no. 5, pp. 772-780.
- Kim, I., Kim, J.H., Moon, S.O., Kwak, H.J., Kim, N.G. & Koh, G.Y. 2000, "Angiopoietin-2 at high concentration can enhance endothelial cell survival through the phosphatidylinositol 3'-kinase/Akt signal transduction pathway", *Oncogene*, vol. 19, no. 39, pp. 4549-4552.
- Kim, I., Moon, S.O., Park, S.K., Chae, S.W. & Koh, G.Y. 2001, "Angiopoietin-1 reduces VEGF-stimulated leukocyte adhesion to endothelial cells by reducing ICAM-1, VCAM-1, and E-selectin expression", *Circulation research*, vol. 89, no. 6, pp. 477-479.
- Kim, K.T., Choi, H.H., Steinmetz, M.O., Maco, B., Kammerer, R.A., Ahn, S.Y., Kim, H.Z., Lee, G.M. & Koh, G.Y. 2005, "Oligomerization and multimerization are critical for angiopoietin-1 to bind and phosphorylate Tie2", *The Journal of biological chemistry*, vol. 280, no. 20, pp. 20126-20131.
- Koblizek, T.I., Weiss, C., Yancopoulos, G.D., Deutsch, U. & Risau, W. 1998, "Angiopoietin-1 induces sprouting angiogenesis in vitro", *Current biology : CB*, vol. 8, no. 9, pp. 529-532.
- Koppensteiner, W.A., Lackner, P., Wiederstein, M. & Sippl, M.J. 2000, "Characterization of novel proteins based on known protein structures", *Journal of Molecular Biology*, vol. 296, no. 4, pp. 1139-1152.
- Kronqvist, N., Lofblom, J., Jonsson, A., Wernerus, H. & Stahl, S. 2008, "A novel affinity protein selection system based on staphylococcal cell surface display and flow cytometry", *Protein engineering, design & selection : PEDS*, vol. 21, no. 4, pp. 247-255.
- Kwak, H.J., Lee, S.J., Lee, Y.H., Ryu, C.H., Koh, K.N., Choi, H.Y. & Koh, G.Y. 2000, "Angiopoietin-1 inhibits irradiation- and mannitol-induced apoptosis in endothelial cells", *Circulation*, vol. 101, no. 19, pp. 2317-2324.
- Lai, Y.P., Huang, J., Wang, L.F., Li, J. & Wu, Z.R. 2004, "A new approach to random mutagenesis in vitro", *Biotechnology and bioengineering*, vol. 86, no. 6, pp. 622-627.

- Leader, B., Baca, Q.J. & Golan, D.E. 2008, "Protein therapeutics: a summary and pharmacological classification", *Nature reviews.Drug discovery*, vol. 7, no. 1, pp. 21-39.
- Lebecque, S.G. & Gearhart, P.J. 1990, "Boundaries of somatic mutation in rearranged immunoglobulin genes: 5' boundary is near the promoter, and 3' boundary is approximately 1 kb from V(D)J gene", *The Journal of experimental medicine*, vol. 172, no. 6, pp. 1717-1727.
- Lee, H.J., Cho, C.H., Hwang, S.J., Choi, H.H., Kim, K.T., Ahn, S.Y., Kim, J.H., Oh, J.L., Lee, G.M. & Koh, G.Y. 2004, "Biological characterization of angiotensin-converting enzyme 2 and angiotensin-converting enzyme 2-related receptor-1", *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, vol. 18, no. 11, pp. 1200-1208.
- Lee, S.Y., Choi, J.H. & Xu, Z. 2003, "Microbial cell-surface display", *Trends in biotechnology*, vol. 21, no. 1, pp. 45-52.
- Li, Z., Woo, C.J., Iglesias-Ussel, M.D., Ronai, D. & Scharff, M.D. 2004, "The generation of antibody diversity through somatic hypermutation and class switch recombination", *Genes & development*, vol. 18, no. 1, pp. 1-11.
- Lin, H. & Cornish, V.W. 2002, "Screening and selection methods for large-scale analysis of protein function", *Angewandte Chemie (International ed.in English)*, vol. 41, no. 23, pp. 4402-4425.
- Lippow, S.M. & Tidor, B. 2007, "Progress in computational protein design", *Current opinion in biotechnology*, vol. 18, no. 4, pp. 305-311.
- Lofblom, J., Sandberg, J., Wernerus, H. & Stahl, S. 2007, "Evaluation of staphylococcal cell surface display and flow cytometry for postselectional characterization of affinity proteins in combinatorial protein engineering applications", *Applied and Environmental Microbiology*, vol. 73, no. 21, pp. 6714-6721.
- Longerich, S., Tanaka, A., Bozek, G., Nicolae, D. & Storb, U. 2005, "The very 5' end and the constant region of Ig genes are spared from somatic mutation because AID does not access these regions", *The Journal of experimental medicine*, vol. 202, no. 10, pp. 1443-1454.
- Macdonald, P.R., Progiatis, P., Ciani, B., Patel, S., Mayer, U., Steinmetz, M.O. & Kammerer, R.A. 2006, "Structure of the extracellular domain of Tie receptor tyrosine kinases and localization of the angiotensin-binding epitope", *The Journal of biological chemistry*, vol. 281, no. 38, pp. 28408-28414.
- Maisonpierre, P.C., Suri, C., Jones, P.F., Bartunkova, S., Wiegand, S.J., Radziejewski, C., Compton, D., McClain, J., Aldrich, T.H., Papadopoulos, N., Daly, T.J., Davis, S., Sato, T.N. & Yancopoulos, G.D. 1997, "Angiotensin-converting enzyme 2, a natural antagonist for Tie2 that disrupts in vivo angiogenesis", *Science (New York, N.Y.)*, vol. 277, no. 5322, pp. 55-60.

- Majors, B.S., Chiang, G.G. & Betenbaugh, M.J. 2009, "Protein and genome evolution in Mammalian cells for biotechnology applications", *Molecular biotechnology*, vol. 42, no. 2, pp. 216-223.
- Marron, M.B., Hughes, D.P., Edge, M.D., Forder, C.L. & Brindle, N.P. 2000, "Evidence for heterotypic interaction between the receptor tyrosine kinases TIE-1 and TIE-2", *The Journal of biological chemistry*, vol. 275, no. 50, pp. 39741-39746.
- Marron, M.B., Singh, H., Tahir, T.A., Kavumkal, J., Kim, H.Z., Koh, G.Y. & Brindle, N.P. 2007, "Regulated proteolytic processing of Tie1 modulates ligand responsiveness of the receptor-tyrosine kinase Tie2", *The Journal of biological chemistry*, vol. 282, no. 42, pp. 30509-30517.
- Martin, A., Bardwell, P.D., Woo, C.J., Fan, M., Shulman, M.J. & Scharff, M.D. 2002, "Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas", *Nature*, vol. 415, no. 6873, pp. 802-806.
- Martin, A. & Scharff, M.D. 2002, "AID and mismatch repair in antibody diversification", *Nature reviews.Immunology*, vol. 2, no. 8, pp. 605-614.
- Martin, A. & Scharff, M.D. 2002, "Somatic hypermutation of the AID transgene in B and non-B cells", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 19, pp. 12304-12308.
- Matsuura, T. & Yomo, T. 2006, "In vitro evolution of proteins", *Journal of bioscience and bioengineering*, vol. 101, no. 6, pp. 449-456.
- McGregor, D.P. 2008, "Discovering and improving novel peptide therapeutics", *Current opinion in pharmacology*, vol. 8, no. 5, pp. 616-619.
- Milkiewicz, M., Ispanovic, E., Doyle, J.L. & Haas, T.L. 2006, "Regulators of angiogenesis and strategies for their therapeutic manipulation", *The international journal of biochemistry & cell biology*, vol. 38, no. 3, pp. 333-357.
- Morgan, B.R. & Massi, F. 2010, "Accurate Estimates of Free Energy Changes in Charge Mutations", *Journal of Chemical Theory and Computation*, vol. 6, no. 6, pp. 1884-1893.
- Mullis, K.B. & Faloona, F.A. 1987, "Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction", *Methods in enzymology*, vol. 155, pp. 335-350.
- Muramatsu, M., Sankaranand, V.S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O. & Honjo, T. 1999, "Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells", *The Journal of biological chemistry*, vol. 274, no. 26, pp. 18470-18476.

- Nambu, H., Nambu, R., Oshima, Y., Hackett, S.F., Okoye, G., Wiegand, S., Yancopoulos, G., Zack, D.J. & Campochiaro, P.A. 2004, "Angiopoietin 1 inhibits ocular neovascularization and breakdown of the blood-retinal barrier", *Gene therapy*, vol. 11, no. 10, pp. 865-873.
- Neuberger, M.S. & Rada, C. 2007, "Somatic hypermutation: activation-induced deaminase for C/G followed by polymerase eta for A/T", *The Journal of experimental medicine*, vol. 204, no. 1, pp. 7-10.
- Nykanen, A.I., Krebs, R., Saaristo, A., Turunen, P., Alitalo, K., Yla-Herttuala, S., Koskinen, P.K. & Lemstrom, K.B. 2003, "Angiopoietin-1 protects against the development of cardiac allograft arteriosclerosis", *Circulation*, vol. 107, no. 9, pp. 1308-1314.
- Orengo, C.A., Jones, D.T. & Thornton, J.M. 1994, "Protein superfamilies and domain superfolds", *Nature*, vol. 372, no. 6507, pp. 631-634.
- Oshima, Y., Oshima, S., Nambu, H., Kachi, S., Takahashi, K., Umeda, N., Shen, J., Dong, A., Apte, R.S., Duh, E., Hackett, S.F., Okoye, G., Ishibashi, K., Handa, J., Melia, M., Wiegand, S., Yancopoulos, G., Zack, D.J. & Campochiaro, P.A. 2005, "Different effects of angiopoietin-2 in different vascular beds: new vessels are most sensitive", *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, vol. 19, no. 8, pp. 963-965.
- Pabo, C. 1983, "Molecular technology. Designing proteins and peptides", *Nature*, vol. 301, no. 5897, pp. 200.
- Papapetropoulos, A., Fulton, D., Mahboubi, K., Kalb, R.G., O'Connor, D.S., Li, F., Altieri, D.C. & Sessa, W.C. 2000, "Angiopoietin-1 inhibits endothelial cell apoptosis via the Akt/survivin pathway", *The Journal of biological chemistry*, vol. 275, no. 13, pp. 9102-9105.
- Papavasiliou, F.N. & Schatz, D.G. 2002, "Somatic hypermutation of immunoglobulin genes: merging mechanisms for genetic diversity", *Cell*, vol. 109 Suppl, pp. S35-44.
- Park, S., Morley, K.L., Horsman, G.P., Holmquist, M., Hult, K. & Kazlauskas, R.J. 2005, "Focusing mutations into the P. fluorescens esterase binding site increases enantioselectivity more effectively than distant mutations", *Chemistry & biology*, vol. 12, no. 1, pp. 45-54.
- Partanen, J., Armstrong, E., Makela, T.P., Korhonen, J., Sandberg, M., Renkonen, R., Knuutila, S., Huebner, K. & Alitalo, K. 1992, "A novel endothelial cell surface receptor tyrosine kinase with extracellular epidermal growth factor homology domains", *Molecular and cellular biology*, vol. 12, no. 4, pp. 1698-1707.
- Parton, R.G. 2003, "Caveolae--from ultrastructure to molecular mechanisms", *Nature reviews.Molecular cell biology*, vol. 4, no. 2, pp. 162-167.

- Pearce, K.H., Jr, Cunningham, B.C., Fuh, G., Teeri, T. & Wells, J.A. 1999, "Growth hormone binding affinity for its receptor surpasses the requirements for cellular activity", *Biochemistry*, vol. 38, no. 1, pp. 81-89.
- Peled, J.U., Kuang, F.L., Iglesias-Ussel, M.D., Roa, S., Kalis, S.L., Goodman, M.F. & Scharff, M.D. 2008, "The biochemistry of somatic hypermutation", *Annual Review of Immunology*, vol. 26, pp. 481-511.
- Petersen-Mahrt, S.K., Harris, R.S. & Neuberger, M.S. 2002, "AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification", *Nature*, vol. 418, no. 6893, pp. 99-103.
- Pichiule, P., Chavez, J.C. & LaManna, J.C. 2004, "Hypoxic regulation of angiopoietin-2 expression in endothelial cells", *The Journal of biological chemistry*, vol. 279, no. 13, pp. 12171-12180.
- Presta, M., Dell'Era, P., Mitola, S., Moroni, E., Ronca, R. & Rusnati, M. 2005, "Fibroblast growth factor/fibroblast growth factor receptor system in angiogenesis", *Cytokine & growth factor reviews*, vol. 16, no. 2, pp. 159-178.
- Procopio, W.N., Pelavin, P.I., Lee, W.M. & Yeilding, N.M. 1999, "Angiopoietin-1 and -2 coiled coil domains mediate distinct homo-oligomerization patterns, but fibrinogen-like domains mediate ligand activity", *The Journal of biological chemistry*, vol. 274, no. 42, pp. 30196-30201.
- Puri, M.C., Rossant, J., Alitalo, K., Bernstein, A. & Partanen, J. 1995, "The receptor tyrosine kinase TIE is required for integrity and survival of vascular endothelial cells", *The EMBO journal*, vol. 14, no. 23, pp. 5884-5891.
- Rada, C. & Milstein, C. 2001, "The intrinsic hypermutability of antibody heavy and light chain genes decays exponentially", *The EMBO journal*, vol. 20, no. 16, pp. 4570-4576.
- Ramiro, A.R., Stavropoulos, P., Jankovic, M. & Nussenzweig, M.C. 2003, "Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand", *Nature immunology*, vol. 4, no. 5, pp. 452-456.
- Rogozin, I.B. & Diaz, M. 2004, "Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process", *Journal of immunology (Baltimore, Md.: 1950)*, vol. 172, no. 6, pp. 3382-3384.
- Rogozin, I.B. & Kolchanov, N.A. 1992, "Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis", *Biochimica et biophysica acta*, vol. 1171, no. 1, pp. 11-18.
- Romero, P.A. & Arnold, F.H. 2009, "Exploring protein fitness landscapes by directed evolution", *Nature reviews.Molecular cell biology*, vol. 10, no. 12, pp. 866-876.

- Rossant, J. & Howard, L. 2002, "Signaling pathways in vascular development", *Annual Review of Cell and Developmental Biology*, vol. 18, pp. 541-573.
- Saharinen, P., Kerkela, K., Ekman, N., Marron, M., Brindle, N., Lee, G.M., Augustin, H., Koh, G.Y. & Alitalo, K. 2005, "Multiple angiopoietin recombinant proteins activate the Tie1 receptor tyrosine kinase and promote its interaction with Tie2", *The Journal of cell biology*, vol. 169, no. 2, pp. 239-243.
- Sale, J.E. & Neuberger, M.S. 1998, "TdT-accessible breaks are scattered over the immunoglobulin V domain in a constitutively hypermutating B cell line", *Immunity*, vol. 9, no. 6, pp. 859-869.
- Sambrook, J., Fritsch, E.F. & Maniatis, T. 1989, *Molecular Cloning: A Laboratory Manual*, 2nd edn, Cold Spring Harbour Laboratory Press, New York.
- Sarkar, I., Hauber, I., Hauber, J. & Buchholz, F. 2007, "HIV-1 proviral DNA excision using an evolved recombinase", *Science (New York, N.Y.)*, vol. 316, no. 5833, pp. 1912-1915.
- Sato, T.N., Qin, Y., Kozak, C.A. & Audus, K.L. 1993, "Tie-1 and tie-2 define another class of putative receptor tyrosine kinase genes expressed in early embryonic vascular system", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 20, pp. 9355-9358.
- Sato, T.N., Tozawa, Y., Deutsch, U., Wolburg-Buchholz, K., Fujiwara, Y., Gendron-Maguire, M., Gridley, T., Wolburg, H., Risau, W. & Qin, Y. 1995, "Distinct roles of the receptor tyrosine kinases Tie-1 and Tie-2 in blood vessel formation", *Nature*, vol. 376, no. 6535, pp. 70-74.
- Schnurch, H. & Risau, W. 1993, "Expression of tie-2, a member of a novel family of receptor tyrosine kinases, in the endothelial cell lineage", *Development (Cambridge, England)*, vol. 119, no. 3, pp. 957-968.
- Scott, B.B., Zaratina, P.F., Gilmartin, A.G., Hansbury, M.J., Colombo, A., Belpasso, C., Winkler, J.D. & Jackson, J.R. 2005, "TNF-alpha modulates angiopoietin-1 expression in rheumatoid synovial fibroblasts via the NF-kappa B signalling pathway", *Biochemical and biophysical research communications*, vol. 328, no. 2, pp. 409-414.
- Seegar, T.C., Eller, B., Tzvetkova-Robev, D., Kolev, M.V., Henderson, S.C., Nikolov, D.B. & Barton, W.A. 2010, "Tie1-Tie2 interactions mediate functional differences between angiopoietin ligands", *Molecular cell*, vol. 37, no. 5, pp. 643-655.
- Shafikhani, S., Siegel, R.A., Ferrari, E. & Schellenberger, V. 1997, "Generation of large libraries of random mutants in *Bacillus subtilis* by PCR-based plasmid multimerization", *BioTechniques*, vol. 23, no. 2, pp. 304-310.
- Shimotohno, A., Oue, S., Yano, T., Kuramitsu, S. & Kagamiyama, H. 2001, "Demonstration of the importance and usefulness of manipulating non-active-site residues in protein design", *Journal of Biochemistry*, vol. 129, no. 6, pp. 943-948.

- Spiess, M. & Lodish, H.F. 1986, "An internal signal sequence: the asialoglycoprotein receptor membrane anchor", *Cell*, vol. 44, no. 1, pp. 177-185.
- Spiess, M., Schwartz, A.L. & Lodish, H.F. 1985, "Sequence of human asialoglycoprotein receptor cDNA. An internal signal sequence for membrane insertion", *The Journal of biological chemistry*, vol. 260, no. 4, pp. 1979-1982.
- Spiller, B., Gershenson, A., Arnold, F.H. & Stevens, R.C. 1999, "A structural view of evolutionary divergence", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 22, pp. 12305-12310.
- Stemmer, W.P. 1994, "Rapid evolution of a protein in vitro by DNA shuffling", *Nature*, vol. 370, no. 6488, pp. 389-391.
- Stratmann, A., Risau, W. & Plate, K.H. 1998, "Cell type-specific expression of angiopoietin-1 and angiopoietin-2 suggests a role in glioblastoma angiogenesis", *The American journal of pathology*, vol. 153, no. 5, pp. 1459-1466.
- Suarez, M. & Jaramillo, A. 2009, "Challenges in the computational design of proteins", *Journal of the Royal Society, Interface / the Royal Society*, vol. 6 Suppl 4, pp. S477-91.
- Suri, C., Jones, P.F., Patan, S., Bartunkova, S., Maisonpierre, P.C., Davis, S., Sato, T.N. & Yancopoulos, G.D. 1996, "Requisite role of angiopoietin-1, a ligand for the TIE2 receptor, during embryonic angiogenesis", *Cell*, vol. 87, no. 7, pp. 1171-1180.
- Tadros, A., Hughes, D.P., Dunmore, B.J. & Brindle, N.P. 2003, "ABIN-2 protects endothelial cells from death and has a role in the antiapoptotic effect of angiopoietin-1", *Blood*, vol. 102, no. 13, pp. 4407-4409.
- Teichert-Kuliszewska, K., Maisonpierre, P.C., Jones, N., Campbell, A.I., Master, Z., Bendeck, M.P., Alitalo, K., Dumont, D.J., Yancopoulos, G.D. & Stewart, D.J. 2001, "Biological action of angiopoietin-2 in a fibrin matrix model of angiogenesis is associated with activation of Tie2", *Cardiovascular research*, vol. 49, no. 3, pp. 659-670.
- Thomas, M. & Augustin, H.G. 2009, "The role of the Angiopoietins in vascular morphogenesis", *Angiogenesis*, vol. 12, no. 2, pp. 125-137.
- Thurston, G., Rudge, J.S., Ioffe, E., Zhou, H., Ross, L., Croll, S.D., Glazer, N., Holash, J., McDonald, D.M. & Yancopoulos, G.D. 2000, "Angiopoietin-1 protects the adult vasculature against plasma leakage", *Nature medicine*, vol. 6, no. 4, pp. 460-463.
- Thurston, G., Suri, C., Smith, K., McClain, J., Sato, T.N., Yancopoulos, G.D. & McDonald, D.M. 1999, "Leakage-resistant blood vessels in mice transgenically overexpressing angiopoietin-1", *Science (New York, N.Y.)*, vol. 286, no. 5449, pp. 2511-2514.

- Tian, W. & Skolnick, J. 2003, "How well is enzyme function conserved as a function of pairwise sequence identity?", *Journal of Molecular Biology*, vol. 333, no. 4, pp. 863-882.
- Tokuriki, N. & Tawfik, D.S. 2009, "Protein dynamism and evolvability", *Science (New York, N.Y.)*, vol. 324, no. 5924, pp. 203-207.
- Tonegawa, S. 1983, "Somatic generation of antibody diversity", *Nature*, vol. 302, no. 5909, pp. 575-581.
- Torchilin, V.P. & Lukyanov, A.N. 2003, "Peptide and protein drug delivery to and into tumors: challenges and solutions", *Drug discovery today*, vol. 8, no. 6, pp. 259-266.
- Towbin, H., Staehelin, T. & Gordon, J. 1979, "Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 76, no. 9, pp. 4350-4354.
- Tsigkos, S., Koutsilieris, M. & Papapetropoulos, A. 2003, "Angiopoietins in angiogenesis and beyond", *Expert opinion on investigational drugs*, vol. 12, no. 6, pp. 933-941.
- Valenzuela, D.M., Griffiths, J.A., Rojas, J., Aldrich, T.H., Jones, P.F., Zhou, H., McClain, J., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Huang, T., Papadopoulos, N., Maisonpierre, P.C., Davis, S. & Yancopoulos, G.D. 1999, "Angiopoietins 3 and 4: diverging gene counterparts in mice and humans", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 5, pp. 1904-1909.
- Van der Sloot, A.M., Kiel, C., Serrano, L. & Stricher, F. 2009, "Protein design in biological networks: from manipulating the input to modifying the output", *Protein engineering, design & selection : PEDS*, vol. 22, no. 9, pp. 537-542.
- van Meurs, M., Kumpers, P., Ligtenberg, J.J., Meertens, J.H., Molema, G. & Zijlstra, J.G. 2009, "Bench-to-bedside review: Angiopoietin signalling in critical illness - a future target?", *Critical care (London, England)*, vol. 13, no. 2, pp. 207.
- Vita, C., Vizzavona, J., Drakopoulou, E., Zinn-Justin, S., Gilquin, B. & Menez, A. 1998, "Novel miniproteins engineered by the transfer of active sites to small natural scaffolds", *Biopolymers*, vol. 47, no. 1, pp. 93-100.
- Walsh, G. 2005, "Biopharmaceuticals: recent approvals and likely directions", *Trends in biotechnology*, vol. 23, no. 11, pp. 553-558.
- Wang, C.L., Harper, R.A. & Wabl, M. 2004, "Genome-wide somatic hypermutation", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 19, pp. 7352-7356.

- Wang, C.L., Yang, D.C. & Wabl, M. 2004, "Directed molecular evolution by somatic hypermutation", *Protein engineering, design & selection : PEDS*, vol. 17, no. 9, pp. 659-664.
- Wang, L., Jackson, W.C., Steinbach, P.A. & Tsien, R.Y. 2004, "Evolution of new nonantibody proteins via iterative somatic hypermutation", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 48, pp. 16745-16749.
- Weber, C.C., Cai, H., Ehrbar, M., Kubota, H., Martiny-Baron, G., Weber, W., Djonov, V., Weber, E., Mallik, A.S., Fussenegger, M., Frei, K., Hubbell, J.A. & Zisch, A.H. 2005, "Effects of protein and gene transfer of the angiotensin-1 fibrinogen-like receptor-binding domain on endothelial and vessel organization", *The Journal of biological chemistry*, vol. 280, no. 23, pp. 22445-22453.
- Wilson, C., Mace, J.E. & Agard, D.A. 1991, "Computational method for the design of enzymes with altered substrate specificity", *Journal of Molecular Biology*, vol. 220, no. 2, pp. 495-506.
- Winship, P.R. 1989, "An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide", *Nucleic acids research*, vol. 17, no. 3, pp. 1266.
- Wittrup, K.D. 2001, "Protein engineering by cell-surface display", *Current opinion in biotechnology*, vol. 12, no. 4, pp. 395-399.
- Witzenbichler, B., Westermann, D., Knueppel, S., Schultheiss, H.P. & Tschöpe, C. 2005, "Protective role of angiotensin-1 in endotoxic shock", *Circulation*, vol. 111, no. 1, pp. 97-105.
- Wolkowicz, R., Jager, G.C. & Nolan, G.P. 2005, "A random peptide library fused to CCR5 for selection of mimetopes expressed on the mammalian cell surface via retroviral vectors", *The Journal of biological chemistry*, vol. 280, no. 15, pp. 15195-15201.
- Wong, T.S., Tee, K.L., Hauer, B. & Schwaneberg, U. 2004, "Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution", *Nucleic acids research*, vol. 32, no. 3, pp. e26.
- Wong, T.S., Zhurina, D. & Schwaneberg, U. 2006, "The diversity challenge in directed protein evolution", *Combinatorial chemistry & high throughput screening*, vol. 9, no. 4, pp. 271-288.
- Xue, F., Gu, Z. & Feng, J.A. 2004, "LINKER: a web server to generate peptide sequences with extended conformation", *Nucleic acids research*, vol. 32, no. Web Server issue, pp. W562-5.
- Yano, T. & Kagamiyama, H. 2001, "Directed evolution of ampicillin-resistant activity from a functionally unrelated DNA fragment: A laboratory model of molecular

- evolution", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 3, pp. 903-907.
- Yao, D., Taguchi, T., Matsumura, T., Pestell, R., Edelstein, D., Giardino, I., Suske, G., Rabbani, N., Thornalley, P.J., Sarthy, V.P., Hammes, H.P. & Brownlee, M. 2007, "High glucose increases angiotensin-2 transcription in microvascular endothelial cells through methylglyoxal modification of mSin3A", *The Journal of biological chemistry*, vol. 282, no. 42, pp. 31038-31045.
- Yelamos, J., Klix, N., Goyenechea, B., Lozano, F., Chui, Y.L., Gonzalez Fernandez, A., Pannell, R., Neuberger, M.S. & Milstein, C. 1995, "Targeting of non-Ig sequences in place of the V segment by somatic hypermutation", *Nature*, vol. 376, no. 6537, pp. 225-229.
- Yoshikawa, K., Okazaki, I.M., Eto, T., Kinoshita, K., Muramatsu, M., Nagaoka, H. & Honjo, T. 2002, "AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts", *Science (New York, N.Y.)*, vol. 296, no. 5575, pp. 2033-2036.
- Yuan, H.T., Khankin, E.V., Karumanchi, S.A. & Parikh, S.M. 2009, "Angiotensin 2 is a partial agonist/antagonist of Tie2 signaling in the endothelium", *Molecular and cellular biology*, vol. 29, no. 8, pp. 2011-2022.
- Yuan, L., Kurek, I., English, J. & Keenan, R. 2005, "Laboratory-directed protein evolution", *Microbiology and molecular biology reviews : MMBR*, vol. 69, no. 3, pp. 373-392.
- Yuen, C.T., Storrington, P.L., Tiplady, R.J., Izquierdo, M., Wait, R., Gee, C.K., Gerson, P., Lloyd, P. & Cremata, J.A. 2003, "Relationships between the N-glycan structures and biological activities of recombinant human erythropoietins produced using different culture conditions and purification procedures", *British journal of haematology*, vol. 121, no. 3, pp. 511-526.
- Zagzag, D., Hooper, A., Friedlander, D.R., Chan, W., Holash, J., Wiegand, S.J., Yancopoulos, G.D. & Grumet, M. 1999, "In situ expression of angiotensin-2 in astrocytomas identifies angiotensin-2 as an early marker of tumor angiogenesis", *Experimental neurology*, vol. 159, no. 2, pp. 391-400.
- Zarrin, A.A., Malkin, L., Fong, I., Luk, K.D., Ghose, A. & Berinstein, N.L. 1999, "Comparison of CMV, RSV, SV40 viral and Vlambda1 cellular promoters in B and T lymphoid and non-lymphoid cell lines", *Biochimica et biophysica acta*, vol. 1446, no. 1-2, pp. 135-139.
- Zhang, W., Bardwell, P.D., Woo, C.J., Poltoratsky, V., Scharff, M.D. & Martin, A. 2001, "Clonal instability of V region hypermutation in the Ramos Burkitt's lymphoma cell line", *International immunology*, vol. 13, no. 9, pp. 1175-1184.
- Zhang, Y., Riesterer, C., Ayrall, A.M., Sablitzky, F., Littlewood, T.D. & Reth, M. 1996, "Inducible site-directed recombination in mouse embryonic stem cells", *Nucleic acids research*, vol. 24, no. 4, pp. 543-548.

Zhao, H. 2007, "Directed evolution of novel protein functions", *Biotechnology and bioengineering*, vol. 98, no. 2, pp. 313-317.

Zhao, Y.D., Campbell, A.I., Robb, M., Ng, D. & Stewart, D.J. 2003, "Protective role of angiotensin-1 in experimental pulmonary hypertension", *Circulation research*, vol. 92, no. 9, pp. 984-991.

#### **Other Web Pages**

[http://tools.invitrogen.com/Content/SFS/ProductNotes/F\\_051025\\_MammalianExpressionVectors-TS-TL-MKT-HL.pdf](http://tools.invitrogen.com/Content/SFS/ProductNotes/F_051025_MammalianExpressionVectors-TS-TL-MKT-HL.pdf)

<http://www.innovagen.se/custom-peptide-synthesis/peptide-property-calculator/peptide-property-calculator.asp>

Compute pI/Mw: [http://expasy.org/tools/pi\\_tool.html](http://expasy.org/tools/pi_tool.html)

Webcutter: <http://www.firstmarket.com/cutter/cut2.html>

Translation tool: <http://www.expasy.ch/tools/dna.html>

Alignment tool: <http://blast.ncbi.nlm.nih.gov/Blast.cgi>