An Investigation of the Binding Capacities of Recombinant Domain Mutants of the Human Polymeric Immunoglobulin Receptor (pIgR)

Earl Adin Gerard Prinsloo

Submitted in fulfilment of the requirements for the degree of

*Philosophiae Doctor*

in the Faculty of Science

at the Nelson Mandela Metropolitan University

January 2006

Promoter: Dr Vaughan Oosthuizen
Co-Promoters: Dr Maryna Van De Venter
Prof Ryno J Naudé
Dedicated to Daniel and Muriel Prinsloo

Thank You
Acknowledgements

The Author wishes to acknowledge and thank the following for their individual contributions.

Dr Vaughan Oosthuizen, for insight, encouragement and allowing me the opportunity to continue the work he initiated.

Dr Maryna van de Venter and Prof Ryno J Naudé for their imparted knowledge.

The staff and students of the Department of Biochemistry and Microbiology; especially Mr Swannie Swanepoel.

Dr Jeannie Fehrsen, Dr Dion du Plessis and the Onderstepoort Veterinary Institute for the use of the BIAcore biosensor and their specialist knowledge of surface plasmon resonance.

Dr David Pugh and the University of Western Cape for the use of the BIAcore biosensor.

Dr Arnoud Marquart and Dr David Myszka for their specialist knowledge of surface plasmon resonance.

Prof Koji Muramoto for amino terminal sequencing and MALDI-TOF MS analysis.

Dr Peter Sondermann (Glycart Biotech, Switzerland) and Katja Wenig (Max Planck Institute, Germany) for the kind donation of recombinant CD89.

The National Research Foundation (GUN2053597 and GUN2050294) and Medical Research Council for funding the study.

The National Research Foundation, Department of Labour, NMMU, and Vodacom for financial support during the duration of the study.

The South African Society of Biochemistry and Molecular Biology (SASBMB) for the travel award allowing me to attend the 30th FEBS Congress and 9th IUBMB Conference, Budapest, Hungary, 2-7July 2005.

My fellow lab rats Brodie, Sandi and Wayne for the ability to bounce ideas off you guys.

My family for support and understanding.

Leisl, for all your encouragement and always believing in me.
Summary

The membrane bound glycoprotein, polymeric immunoglobulin receptor (pIgR) is the primary transport molecule of the polymeric immunoglobulins, dimeric IgA and pentameric IgM, across epithelial cells. This process, known as transcytosis, is essential in order to establish immunity at mucosal surfaces. Typically, pIgR binds to the polymeric immunoglobulin at the basolateral surface of the epithelial cell, via five homologous immunoglobulin-like domains of the ectodomain. Binding is covalent to IgA and non-covalent to IgM; the IgM binding varying among species. The pIgR-bound complex is released at the apical surface of the cell after cleavage of pIgR at Arg^{585}, thereafter referred to as secretory component (SC). SC confers protective and immunologic functions to the polymeric immunoglobulin. Free SC, i.e. not complexed with polymeric immunoglobulins, is also known to be released into mucosal secretions; and binds to pathogenic bacteria and bacterial products. It is known that domain I of the ectodomain is the primary domain in the interaction with polymeric immunoglobulins, while domain V is involved in a covalent linkage with IgA. However, little is known of domains II-IV and their role in immunoglobulin binding, particularly to IgM. This study aimed to characterize the binding of recombinant human pIgR domain mutants to polymeric IgM using immunological, biophysical and cell based techniques; thereby allowing greater insight into the contribution of each of the five domains. The unique domain structure allowed for selective amplification of single and multiple domain mutants from cloned human PIGR ectodomain cDNA. Mutants were cloned and expressed in *Escherichia coli* BL21 (DE3) as inclusion bodies. Recombinant mutant proteins were refolded *in vitro* by equilibrium gradient dialysis and purified to homogeneity. Equilibrium binding data show significant contributions to specific binding as a factor of domain presence. Binding kinetics determined by biophysical surface plasmon resonance measurements show the interplay between association and dissociation rates as defined by individual domains. *In vitro* competitive binding studies using the human intestinal carcinoma, HT29, known to constitutively express pIgR, show that the constructed recombinant domain mutants outcompete native pIgR. The level of competition is shown to be dependant on the domains downstream of domain I. The data also confirm the biological activity of the first *in vitro* refolded recombinant human SC.
**Keywords:** recombinant, human, mucosal immunity, polymeric immunoglobulin receptor, secretory component, IgA, IgM
Table of Contents

List of Figures  IV
List of Tables  X
List of Equations  XI
List of Abbreviations  XII

CHAPTER 1 1
The Polymeric Immunoglobulin Receptor (pIgR), Polymeric Immunoglobulins and Mucosal Immunity 1
1.1 The PIGR Gene-Structure, Regulation and Homology 2
1.2 Structural Features of pIgR 3
  1.2.1 The pIgR Ectodomain 5
    1.2.1.1 The amino terminal domain of pIgR/SC, domain I 6
    1.2.1.2 Domain II, III, IV and V 9
    1.2.1.3 N-glycosylation and Binding Activity 10
1.3 Secretory Immunoglobulins (SIgA and SIgM) 12
  1.3.1 Structural and Biological Features of Secretory Immunoglobulins 12
    1.3.1.1 IgA 12
    1.3.1.2 IgM 13
    1.3.1.3 J Chain 15
1.4 The Short Cytoplasmic Tail, Transcytosis and Mucosal Immunity 15
1.5 Functions of pIgR, Secretory Component 17

CHAPTER 2 19
Aim of Research 19

CHAPTER 3 21
Production, Purification and Characterization of Recombinant Domain Mutants 21
  3.1 Materials and Methods 23
    3.1.1 Reagents 23
    3.1.2 Bacterial Cultures and Media 23
    3.1.3 Amplification and Molecular Cloning 24
    3.1.4 Preparative Expression and Isolation of Inclusion Bodies 25
    3.1.5 In Vitro Refolding of Recombinant Protein 26
    3.1.6 Chromatography 27
    3.1.6.1 Immobilized Metal Affinity Chromatography (IMAC) 27
      3.1.6.1.1 Denaturing IMAC using Ni²⁺ Affinity Resin 27
      3.1.6.1.2 Non-Denaturing IMAC using Ni²⁺ Affinity Resin 28
    3.1.6.2 Gel Filtration Chromatography 28
    3.1.7 Protein Assays 28
    3.1.8 Dot Blot Assay for Epitope Identity 28
    3.1.9 MALDI-TOF MS Analysis of Disulfide Bonds 29
  3.2 Results and Discussion 30
    3.2.1 Amplification and Molecular Cloning of Recombinant SC Mutants 30
      3.2.1.1 Single Domain Mutants 32
      3.2.1.2 Multiple Domain Mutants 32
      3.2.1.3 Sequence Analysis and in silico Plasmid Mapping. 33
    3.2.2 Protein Expression and Purification of Inclusion Bodies 36
      3.2.2.1 Single domain mutants (domains I and V) 37
3.2.2.2 Multiple domain mutants (hSC_rD I-II, hSC_rD I-III, hSC_rD I-IV and hSC_rD I-V [rSC]) 39
3.2.3 Refolding & Purification of Recombinant Domain Mutants 41
3.2.3.1 Single domain mutants 41
3.2.3.2 Multiple Domain Mutants 44
3.2.4 Dot Blot Association Assay: Establishment of “Identity” of recombinant domain mutants to Native human SC 50
3.2.5 MALDI-TOF MS Analysis of Disulfide Bonds 50

CHAPTER 4 52

Binding Activity Studies 52
4.1 Materials and Methods 54
   4.1.1 Reagents and Equipment 54
   4.1.2 Dot Blot Reassociation Assay for IgM Binding 54
   4.1.3 Receptor-Ligand Binding Assay (Direct ELISA) 54
   4.1.4 Biacore™ Surface Plasmon Resonance Measurements 55
      4.1.4.1 Amine Coupling Procedure: Attaching a Ligand to BIAcore Chip Surface 55
      4.1.4.2 Receptor Coupling: Recombinant Domain Mutants as “Ligands” 56
      4.1.4.3 IgM Coupling: Recombinant Domain Mutants and rSC as “Analytes” 56
      4.1.4.4 SPR Data Analysis 56
   4.1.5 Competitive Binding Studies: CELISA (Bound Whole Cell ELISA) 56
      4.1.5.1 Cell Culture and Maintenance 57
      4.1.5.2 Immunohistochemical probe for plgR 57
      4.1.5.3 Induction of plgR Expression 58
      4.1.5.4 Western Blot for Determination of plgR Expression 58
      4.1.5.5 Competitive Binding Studies: CELISA 58
      4.1.5.6 Data Analysis 59
4.2 Results and Discussion 61
   4.2.1 Dot Blot Reassociation Assay (DORA) Analysis 61
   4.2.2 Receptor Ligand Binding Assay: Direct ELISA 62
   4.2.3 Surface Plasmon Resonance Analysis 66
      4.2.3.1 Recombinant Domain Mutants as Ligands 66
      4.2.3.2 Recombinant Domain Mutants as Analyte 77
   4.2.4 Competitive Binding Studies 95
      4.2.4.1 Immunohistochemical and Western Blot Analysis of plgR Expression in HT29 96
      4.2.4.2 Whole Cell ELISA Competitive Binding Assay 97

CHAPTER 5 101

Conclusions 101
5.1 Expression and Purification of Recombinant Domain Mutants 101
5.2 Binding Studies 102
5.3 Future Work and Applications of Recombinant SC mutants 105

REFERENCES 108

Further Reading 120

APPENDIX A 122
   DNA Sequence Alignments, Predicted Protein Sequence Alignments and Automated Sequencing Electropherograms 122

APPENDIX B 155
   MALDI TOF Spectra 155

APPENDIX C 162
Publications Resulting from the Current Study
List of Figures

Figure 1.1: Diagrammatic representation of pIgR showing the extracellular ligand binding domain composed of five smaller domains (I-V), the transmembrane region and the short intracellular tail. Key features highlighted are the disulfide bonds of the extracellular ligand binding domain; domain I’s non-covalent immunoglobulin binding region and V’s reactive disulfide which forms a covalent disulfide with IgA Ca3. Key signalling regions in the cytoplasmic tail are also shown (from Kaetzel, 2001)...................................................................................................................................................4

Figure 1.2: Alignment of pIgR amino acid sequences from representing various species. Dashes indicate identity, dots indicate gaps, * conserved Cys residues. Domain borders are clearly marked.............4

Figure 1.3: Ribbon structure of domain I showing characteristic arrangement of CDR loops 1, 2 and 3 and the 10 anti-parallel β-strands (see text) (from Hamburger et al., 2004).........................................................8

Figure 1.4: Topological diagram of pIgR domain I and positions of the CDR loops (from Hamburger et al., 2004). ...................................................................................................................................................8

Figure 1.5: Overlap of IgVd domain and domain I showing structural similarity and non-canonical CDR3 loop arrangement. The modelled pIgR D1 dimer figure shows the improbability of pIgR dimer formation, a function of the CDR3 loop in immunoglobulins (from Hamburger et al., 2004)....9

Figure 1.6: Histogram comparing translocation of 100nM pIgA and pIgM across MDCK cell monolayers (polarized) expressing wild type pIgR, and the deletion mutants pIgRΔ2,3 and pIgRΔ4,5 (control = MDCK untransfected) (from Norderhaug et al., 1999).........................................................................................................................10

Figure 1.7: Diagrammatic representation of oligosaccharide structures bound to asparagine residues of human SC (from Hughes et al., 1999). (A) Fuc-Hex-HexNAc4; (B) Fuc-Hex-HexNAc4; (C) NeuAcFucHex-HexNAc4; (D) NeuAcFucHex-HexNAc4; (E) NeuAcHex-HexNAc4; (F) NeuAcFucHex-HexNAc4; .................................................11

Figure 1.8: Diagrammatic representations of (a) IgA isotypes and (b) SIgA. The heavy chain (Ca) consists of 3 constant regions Ca1 − Ca3 and one variable domain with light chain composed of 1 variable and 1 constant domain. The relative positions for O and N linked glycosylation sites are represented by | and •, respectively. The position for the docking site of the IgA receptor FcaRI (CD89) is also shown (from van Egmond et al., 2001). .....................................................................................................................................................13

Figure 1.9: Diagrammatic representation of pentameric IgM showing key structural features. Light chains are shown in green and heavy chains in blue. Orange circles show sites of glycosylation (from Benjamini et al., 2000). ...................................................................................................................................................14

Figure 1.10: Transcytosis of polymeric immunoglobulins by pIgR. JC, junctional complex (see text for details) (adapted from Mostov, 1994). .........................................................................................................................................................16

Figure 1.11: Modes of defence mediated by dIgA, pIgR and fSC. (1) transport of polymeric immunoglobulins and release of fSC; (2) transport of pathogens across epithelia; (3) intracellular neutralization, (4) immune exclusion, (5) prevention of binding and entry of pathogens by fSC and (6) pathogen invasion by subversion of pIgR (from Phalipon and Corthésy, 2003). ....................................................................................17

Figure 2.1: Schematic diagram showing the outline of the pIgR ectodomain encoding the 5 domains (I-V). The coloured arrows indicate the desired amplification products as defined by the primers in Table 3.1. The sizes (bp) of desired fragments are shown in brackets. Blue- arrow, hSC rDI (344), Green arrow, hSC rDI-II (665), Grey arrow, hSC rDI-III (986), Red arrow, hSC rDI-IV (1325) and Orange arrow, hSC rDI-V (1658). The double headed brown arrow shows the amplified region encoding domain IV and V (672). The white arrow indicates the desired amplification product encoding domain V (hSC rDV-[333]). Grey/black crosshatch box represents cloned pIgR ectodomain cDNA. .....................................................................................................................................................19

Figure 3.1: Agarose gel (1%) electrophoretic patterns of RT-PCR products showing successful amplification of pIgR ectodomain upon temperature optimization. Arrow indicates band of interest. (a) control RT-PCR; (b) 50°C annealing; (c) 45°C annealing; (d) 53°C annealing (~1.8kb); (e) 60°C annealing; (f) Roche DNA Marker XVI..........................30
Figure 3.2: DNA sequence and three letter in silico translation representing cloned plgR ectodomain showing highlighted regions defining borders of domains I to V. Blue box represents the 5' forward primer of hSC_rDI (Table 3.1). The red boxes represent the C terminal borders as defined by reverse primers shown in Table 3.1. The yellow and green boxes represent the N terminal ends of domains IV and V, respectively. The orange Met signifies the N terminal residue of the cDNA translation product. Black box highlights indicate residues outlined in Table 3.2. ...............................31

Figure 3.3: Agarose gel electrophoretic pattern (1% gel) showing amplified fragments of single domain mutants. Lanes (a) domain I (hSC_rDI ~ 364bp) (b) 1kb Gene Ruler (Fermentas); (c) domain V (hSC_rDV ~ 320bp); and (d) phage λ DNA digested with PstI as a marker (λ.PstI). ..................32

Figure 3.4: Agarose gel electrophoretic patterns (1% gel) showing PCR amplification of multiple domain mutants. Lanes: (a) isolated pET22b(+) hSC_rDI; (b) pET22b(+) hSC_rDI-II; (c) pET22b(+) hSC_rDI-III; (d) pET22b(+) hSC_rDI-IV and (e) pET22b(+) hSC_rDV. Inserted fragments are shown in orange. Dark blue arrows are C terminal His tags conferred by plasmid. Yellow arrows indicate position of Ampicillin resistance gene (amp). ........................................33

Figure 3.5: Section of output from automated DNA sequencing showing the region encoding the N terminal end of multiple domain mutant hSC_rDI-II. Note the Nde I restriction site (highlighted in yellow) immediately upstream of position 70. .................................................................34

Figure 3.6: Representative plasmid maps of cloned recombinant domain mutants constructed in silico. (a) pET22b(+) hSC_rDI; (b) pET22b(+) hSC_rDI-II; (c) pET22b(+) hSC_rDI-III; (d) pET22b(+) hSC_rDI-IV and (e) pET22b(+) hSC_rDV. Inserted fragments are shown in orange. Dark blue arrows are C terminal His tags conferred by plasmid. Yellow arrows indicate position of Ampicillin resistance gene (amp). ........................................35

Figure 3.7: SDS-PAGE patterns (15% gel) of expression trials of single domain mutants. Gel 1 (hSC_rDI), lanes (a) Sigma low molecular weight markers; (b) preinduction E. coli extract; (c) 4 Hour expression upon induction with 1 mM IPTG; Gel 2 (hSC_rDV), lanes (a-c) preinduction and (e-g) expression (4 hours) in E. coli for clone selection purposes upon induction with 1mM IPTG; (d) Sigma low molecular weight markers. Arrows indicate bands of interest. .......................................................38

Figure 3.8: SDS-PAGE patterns (15%) of partially isolated inclusion bodies for single domain mutants. Gel 1, partially purified hSC_rDI I inclusion bodies using sequential washing and centrifugation; lanes (a) Sigma low molecular weight markers; (b) hSC_rDI I (5µg) inclusion bodies (predicted size ~13kDa); (c) hSC_rDI I inclusion bodies (10µg). Gel 2, partially isolated hSC_rDV (predicted size 13kDa) inclusion bodies; lanes (a) 10µg loaded (b) 5µg loaded, respectively (c) Sigma low molecular weight markers. Arrows indicate bands of interest. ........................................38

Figure 3.9: SDS-PAGE patterns (12% gel) of expression of hSC_rDI-II (predicted size ~25kDa). Lanes (a) Sigma low molecular weight markers; (b) preinduction E. coli extract; (c) 4 hours post induction; (d) overnight (18 hours) post induction. Arrow indicates band of interest. .......................................................39

Figure 3.10: SDS-PAGE pattern (15% gel) of expression of hSC_rDI-III (predicted size ~37kDa). Lanes (a) Sigma high molecular weight marker; (b) preinduction E. coli extract; (c) 2 hours uninduced; (d) 2 hours post induction; (e) 4 hours uninduced; (f) 4 hours post induction (g) overnight (16 hours) uninduced; (h) overnight (16 hours) post induction. Arrow indicates band of interest. .........................39

Figure 3.11: SDS-PAGE pattern (12%) of expression of hSC_rDI-IV (predicted size ~50kDa). Lanes (a) Sigma high molecular weight marker; (b) and (c) preinduction E. coli extract; (d) and (e) 2 hour post induction; (f) and (g) 3 hours post induction; (h) and (i) 4 hours post induction and (j) and (k) Overnight (16 hours) post induction of two clones, respectively. Arrow indicates band of interest. .........................40

Figure 3.12: SDS-PAGE pattern (12%) gel of expression of hSC_rDI-V (predicted size ~61kDa). Lanes (a) Sigma high molecular weight marker; (b) and (c) preinduction E. coli extract; (d) and (e) 1 hour post induction; (f) & (g) 2 hours post induction; (h) and (i) 3 hours post induction; (j) and (k) 4 hours post induction and (l) & (m) overnight (16 hours) post induction, of two clones, respectively. Arrow indicates band of interest. ................................................40

Figure 3.13: SDS-PAGE pattern (10%) of partially purified inclusion bodies of recombinant multiple domain mutants. Lanes (a) hSC_rDI-II; (b) hSC_rDI-III; (c) hSC_rDI-IV; (d) hSC_rDI-V. Approximately 15 µg total protein loaded per well........................................41

Figure 3.14: A typical chromatogram representing the purification of hSC_rDI using the Ni²⁺ affinity resin under denaturing conditions. Blue line indicates absorbance at 280nm and the red line indicates
conductivity of buffer. Blue arrow indicates target peak (5ml fractions 8-12), Red arrow indicates start of imidazole addition. ........................................................................................................................................42

Figure 3.15: SDS-PAGE (15%) of single domain mutant inclusion bodies purified using Ni\textsuperscript{2+} affinity chromatography under denaturing conditions. Lanes (a) Sigma low molecular weight markers; (b) fraction 8 hSC_rDI; (c) fraction 9 hSC_rDI; (d) fraction 10 hSC_rDI; (e) fraction 11 hSC_rDI; (f) fraction 12 hSC_rDI; (g) Sigma low molecular weight markers; (h) partially purified inclusion bodies hSC_rDV: (i) fraction 10 hSC_rDV; (j) fraction 11 hSC_rDV; (k) fraction 12 hSC_rDV. Arrows indicate bands of interest.................................................................42

Figure 3.16: SDS-PAGE patterns (15% gel) showing the refolding of hSC_rDI. (a) Sigma low molecular weight markers; (b) 5µg reduced fraction; (c) 5µg non-reduced fraction (white arrow represents the sample). The black filled arrow shows possible aggregates. ..................................................................................43

Figure 3.17: Superdex75 gel filtration chromatography of refolded hSC_rDI. Blue arrow shows refolded hSC_rDI, red arrow shows potential aggregates and unfolded material. Peak bounded by green arrows pooled......................................................................................................................................43

Figure 3.18: SDS-PAGE pattern (15%) of purified single domain mutants (hSC_rDI and hSC_rDV). Gel 1, lane (a) broad range molecular weight markers; (b) hSC_rDI (M~15.3kDa). Gel 2, lane (a) Sigma low molecular weight markers; (b) hSC_rDV (M~12.8kDa). ........................................................................................................44

Figure 3.19: N-terminal sequence alignment of hSC_rDI and hSC_rDV with plgR. Legend: *, consensus; *, mismatches and –, gaps..........................................................................................44

Figure 3.20: SDS-PAGE pattern (15% gel) of elution of multiple domain mutants byNi\textsuperscript{2+} affinity chromatography. Gel 1 (hSC_rDI-II) lane (a) partially purified inclusion bodies; (b) fraction 10; (c) fraction 11; (d) fraction 12. Gel 2 (hSC_rDI-III), lane (a) partially purified inclusion bodies; (b) fraction 10; (c) fraction 11; (d) fraction 12. Gel 3 (hSC_rDI-IV), lane (a) partially purified inclusion bodies; (b) fraction 10; (c) fraction 11; (d) fraction 12. Gel 4 (hSC_rDI-V), lane (a) partially purified inclusion bodies; (b) fraction 10; (c) fraction 11; (d) fraction 12. Arrows indicate bands of interest........................................................................................................45

Figure 3.21: Elution profile of hSC_rDI-V(rSC) separation on a Superdex200. Blue line indicates the absorbance at 280nm of concentrated refolded inclusion bodies and the red line (secondary Y axis) indicates the absorbance at 280nm of concentrated pooled fraction from the initial chromatography (i.e. the rechromatographed sample). .........................................................................................................................................46

Figure 3.22: SDS-PAGE patterns (15% gel) showing purified inclusion bodies and hSC_rDI-V (rSC). Lanes (a) and (b) broad range molecular weight markers; (b) purified inclusion bodies (prior to refold); (c) and (e), pure hSC_rDI-V (reduced [-58kDa] and non-reduced [-54kDa], respectively). .................................................................46

Figure 3.23: SDS-PAGE patterns (15% gel) of purified multiple domain mutants. Lanes: (1b) hSC_rDI-II (~27kDa); (2b) hSC_rDI-III (~36kDa); (3b) hSC_rDI-IV (~51kDa). (1a), (2a) and (3a) broad range molecular weight markers. ..........................................................................................................................46

Figure 3.24: N-terminal sequence alignment of multiple domain mutants (hSC_rDI-II – hSC_rDI-V) with plgR. Legend: *, consensus; *, mismatches and –, gaps..................................................................................47

Figure 3.25: Dot blot analysis showing identity of recombinant domain mutants with native human SC. (a) SlgA positive control; (b) recombinant CD89 negative control; (c) hSC_rDI; (d) hSC_rDI-II; (e) hSC_rDI-III; (f) hSC_rDI-IV; (g) hSC_rDI-V (rSC). .................................................................50

Figure 3.26: In silico digestion of hSC_rDI-V with lysylendopeptidase. Red blocks represent Lys residues, i.e. cleavage sites. Green blocks represent the 20 cysteines involved in internal disulfide bonds. ........................................................................51

Figure 3.27: Overlay of MALDI-TOF MS spectra of lysylendopeptidase digested hSC_rDI and BSA. ...............................................................................................................................51

Figure 4.1: Excitation of surface plasmons at a sensor surface by polarized light allows for measurement of the changes in the angle of reflected light upon interaction events. SPR results in reduced reflected light intensity which is related to the angle of reflection and the wavelength (from BIAcore, 1998).52

Figure 4.2: Dot blot reassociation assay for IgM binding activity of hSC_rDI. (a) Free SC positive control; (b) recombinant CD89 negative control; (c) hSC_rDI-I; (d) hSC_rDI-II; (e) hSC_rDI-III; (f) hSC_rDI-IV; (g) hSC_rDI-V (rSC). .........................................................................................................................................................61

Figure 4.3: Dot blot reassociation assay for IgM binding activity of multiple domain mutants. (a) Free SC positive control; (b) recombinant CD89 negative control; (c) hSC_rDI-II; (d) hSC_rDI-III; (e) hSC_rDI-IV; (f) hSC_rDI-V (rSC). .........................................................................................................................................................62
Figure 4.4: Nonspecific binding of CD89 to IgM. Error bars represent standard deviation (n=3). Data were subtracted from total binding data as determined by equation 4.2. .................................63
Figure 4.5: Non-linear fit for determination of $K_d$ of hSC_rDI for pentameric IgM using ELISA. Error bars represent standard deviation (n=3) ($R^2 = 0.8343$). .................................................................63
Figure 4.6: Non-linear fit for determination of $K_d$ of hSC_rDI-II for pentameric IgM using ELISA. Error bars represent standard deviation (n=3) ($R^2 = 0.8395$). ...........................................................................64
Figure 4.7: Non-linear fit for determination of $K_d$ of hSC_rDI-III for pentameric IgM using ELISA. Error bars represent standard deviation (n=3) ($R^2 = 0.7007$). ...........................................................................64
Figure 4.8: Non-linear fit for determination of $K_d$ of hSC_rDI-IV for pentameric IgM using ELISA. Error bars represent standard deviation (n=3) ($R^2 = 0.8447$). ...........................................................................65
Figure 4.9: Non-linear fit for determination of $K_d$ of hSC_rDI-V for pentameric IgM using ELISA. Error bars represent standard deviation (n=3) ($R^2 = 0.8380$). ...........................................................................65
Figure 4.10: Sensorgrams showing the binding capacity of hSC_rDI and hSC_rDV. The analyte (IgM) injection shown here was 1000nM. .................................................................67
Figure 4.11: Sensorgrams showing the effect of flow rate on the binding of analyte, IgM, to ligand, hSC_rDI. .................................................................68
Figure 4.12: (a) Sensorgrams showing the Langmuir fit of hSC_rDI binding to varying IgM concentrations (analyte) at 30µl/min using local fitting algorithm (see text for details). (b) relative residual curves showing deviation of fit from the data points. .................................................................69
Figure 4.13: Sensorgrams showing the Langmuir fit of hSC_rDI binding to varying IgM concentrations (analyte) at 30µl/min using global fitting algorithm. (b) relative residual curves showing deviation of fit from the data points. .................................................................70
Figure 4.14: (a) Sensorgrams showing the Bivalent model fit to hSC_rDI binding to varying IgM concentrations (analyte) at 30µl/min using global fitting algorithm (see text for details). (b) relative residual curves showing deviation of fit from the data points. ...........................................................................71
Figure 4.15: Component view showing the contribution (at 1000nM IgM) of individual complexes and bulk and drift to total binding response with hSC_rDI, as defined by the bivalent model. .................................................................72
Figure 4.16: Sensorgrams showing the effect of flow rate on the binding of analyte, IgM, to ligand. (a) hSC_rDI-II and (b) hSC_rDI-III. .................................................................72
Figure 4.17: Sensorgrams showing the Langmuir model fitted to varying IgM concentrations binding to (a) hSC_rDI-II and (c) hSC_rDI-III. Relative residual curves showing deviation of fit from data points are shown for (b) hSC_rDI-II and (d) hSC_rDI-III. .................................................................74
Figure 4.18: Sensorgrams showing the bivalent model fitted to varying IgM concentrations binding to (a) hSC_rDI-II and (c) hSC_rDI-III. Relative residual curves showing deviation of fit from data points are shown for (b) hSC_rDI-II and (d) hSC_rDI-III. .................................................................76
Figure 4.19: Component view of (a) hSC_rDI-II and (b) hSC_rDI-III showing the contribution (at 300 nM IgM) of individual complexes and bulk and drift to total binding response as defined by the bivalent model. .................................................................77
Figure 4.20: Sensorgrams of hSC_rDI, as analyte, binding to varying IgM concentrations. The curves show aberrations that prevented reliable global fits and were overcome by using an initial rate analysis. .................................................................79
Figure 4.21: (a) Sensorgram showing a representative saturation curve for hSC_rDI as analyte that allowed for the determination of $R_{max}$ and $k_a$ and (b) the initial portions of the sensorgram curves used for the determination of the association constant $k_a$. ...........................................................................80
Figure 4.22: Linear regression curve used for the determination of the association rate of hSC_rDI for IgM from the initial rates. The slope of the curve is used for the determination of $k_a$ using equation 4.8. Error bars represent standard deviation (n = 3) ($R^2 = 0.9988$). .................................................................80
Figure 4.23: (a) Sensorgrams showing the Langmuir model fitted to varying concentrations of hSC_rDI-II, as analyte, binding to IgM as ligand. (b) Relative residual curve showing deviation of fit ($\chi^2 = 9.232$). .................................................................82
Figure 4.24: (a) Sensorgrams showing the Langmuir model fitted to varying concentrations of hSC_rDI-III, as analyte, binding to IgM, as ligand. (b) Relative residual curve showing deviation of fit ($\chi^2 = 6.215$). .................................................................83
Figure 4.25: (a) Sensorgrams showing the Langmuir model fitted to varying concentrations of hSC_rDI-IV, as analyte, binding to IgM, as ligand. (b) Relative residual curve showing deviation of fit ($\chi^2 = 4.42$). .................................................................84
Figure A.10: Electropherogram representing the sequencing of pET22b hSC_rDI-IV using T7 promoter primer.........................................................................................................................147
Figure A.11: Electropherogram representing the sequencing of pET22b hSC_rDI-IV using T7 terminator primer..............................................................................................................................................148
Figure A.12: Electropherogram representing the sequencing of pET22b hSC_rDI-IV using the internal forward domain III 5’ border region primer..........................................................................................149
Figure A.13: Electropherogram representing the sequencing of pET22b hSC_rDI-V using T7 promoter primer..............................................................................................................................................150
Figure A.14: Electropherogram representing the sequencing of pET22b hSC_rDI-V using T7 terminator primer..............................................................................................................................................151
Figure A.15: Electropherogram representing the sequencing of pET22b hSC_rDI-V using the internal forward primer representing the domain V 5’ border region. ..............................................................152
Figure A.16: Electropherogram representing the sequencing of pET22b hSC_rDI-V using an internal reverse primer. ..............................................................................................................................................153
Figure A.17: Schematic diagram of the structure of the mouse plgR gene. E1 – E11 represent exons. The introns are represented by horizontal lines. The grey hatched boxes represent 5’ and 3’ untranslated regions. The scale is in kilobases. (from Martin et al., 1997) ..............................................................................................................................................154

Figure B.1: MALDI TOF MS spectrum of lysylendopeptidase digestion of hSC_rDI-V.................................156
Figure B.2: MALDI TOF MS spectrum of lysylendopeptidase digestion of hSC_rDI-IV.................................157
Figure B.3: MALDI TOF MS spectrum of lysylendopeptidase digestion of hSC_rDI-III.................................158
Figure B.4: MALDI TOF MS spectrum of lysylendopeptidase digestion of hSC_rDI-II.................................159
Figure B.5: MALDI TOF MS spectrum of lysylendopeptidase digestion of hSC_rDI.................................160
Figure B.6: MALDI TOF MS spectrum of lysylendopeptidase digestion of BSA..............................................161
### List of Tables

**Table 1.1**: Disulfide bond arrangement of plgR ectodomain. “Large” and “Small” refer to spatial arrangement of the bonds and “Extra” refers to the reactive disulfide bond (adapted from Effert et al., 1984).  

**Table 1.2**: Relative affinities of recombinant domain I for IgA and IgM, compared to IgM affinity purified human secretory component. SD, standard deviation; $K_D$, equilibrium dissociation constant (adapted from Bakos et al., 1994).

**Table 3.1**: Primer sequences used in the amplification of domain deletion mutants (see text for details). Nde I sites are denoted by CATATG and Xho I sites are denoted by CTCGAG. Boxes shaded grey denote the primers targeting the 5' portion of individual domain regions. The forward primer for hSC_rDI was used for all multiple domain mutants unless stated otherwise. Forward primers all contain start codon ATG.

**Table 3.2**: Amino and carboxy terminal amino acid residues defining 5 immunoglobulin-like domains of plgR ectodomain.

**Table 3.3**: Statistical evaluation of recombinant protein solubility upon overexpression in E. coli using the method of Wilkinson and Harrison (1991). n.d. = not determined.

**Table 3.4**: Purification table of recombinant domain mutants showing total protein content (mg) and percentage yield from a 100ml refold.

**Table 4.1**: Equilibrium dissociation constants as determined by non-linear regression for the recombinant domain deletion mutants as obtained from receptor-ligand binding assay.

**Table 4.2**: Rate constants obtained by performing local fits on single curve sets collected for 100nM IgM injections at varying flow rates with hSC_rDI as ligand.

**Table 4.3**: Rate constants obtained from local data fitting of IgM binding to hSC_rDI.

**Table 4.4**: Rate constants obtained by performing global fits on single curve sets of hSC_rDI-II and hSC_rDI-III collected for 100nM IgM injections at varying flow rates.

**Table 4.5**: Kinetic constants obtained from performing global fits of the Langmuir model to data sets for IgM binding to hSC_rDI, hSC_rDI-II and hSC_rDI-III.

**Table 4.6**: Kinetic constants obtained from performing global fits of the bivalent analyte model to data sets for IgM binding to hSC_rDI, hSC_rDI-II and hSC_rDI-III.

**Table 4.7**: kinetic constants obtained from performing global fits of the heterogeneous ligand model to data sets for IgM binding to hSC_rDI, hSC_rDI-II and hSC_rDI-III.

**Table 4.8**: Kinetic constants of hSC_rDI-V (rSC) for IgM obtained from data fitting, compared with reported values for native SC. Association ($k_a$ ± SD) and dissociation ($k_d$ ± SD) rates shown are from separate global fits. (n=3).

**Table 4.9**: Kinetic constants derived from the model fits to the data sets of recombinant SC mutants binding to IgM as ligand. $1:1$ ≡ Langmuir model and $HL$ ≡ Heterogeneous ligand model. SD is the standard deviation (n=3).

**Table 4.10**: IC$_{50}$ and $K_i$ values derived from the one site competition model fit.

**Table A.1**: Interpretation of single letter codes used for editing electropherograms.
# List of Equations

<table>
<thead>
<tr>
<th>Equation</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>36</td>
</tr>
<tr>
<td>3.2</td>
<td>36</td>
</tr>
<tr>
<td>3.3</td>
<td>36</td>
</tr>
<tr>
<td>4.1</td>
<td>60</td>
</tr>
<tr>
<td>4.2</td>
<td>62</td>
</tr>
<tr>
<td>4.3</td>
<td>63</td>
</tr>
<tr>
<td>4.4</td>
<td>67</td>
</tr>
<tr>
<td>4.5</td>
<td>70</td>
</tr>
<tr>
<td>4.6</td>
<td>73</td>
</tr>
<tr>
<td>4.7</td>
<td>73</td>
</tr>
<tr>
<td>4.8</td>
<td>78</td>
</tr>
<tr>
<td>4.9</td>
<td>78</td>
</tr>
<tr>
<td>4.10</td>
<td>88</td>
</tr>
<tr>
<td>4.11</td>
<td>99</td>
</tr>
</tbody>
</table>
List of Abbreviations

A
analyte
ATCC
American Type Culture Collection
ATRA
all trans retinoic acid
B
ligand
CbpA
choline binding protein A
CDR
complementarity determining region
CELISA
Whole Cell Enzyme Linked Immunosorbant Assay
CV
canonical variable
DAB
3, 3’-Diaminobenzidine
dIgA
dimeric Immunoglobulin A
DMEM
Dulbecco’s Modified Eagle’s Medium
dNTP
deoxyribonucleotide triphosphate
DORA
dot blot reassociation assay
FCS
Fetal calf serum
FLAG
octapeptide fusion tag Asp Tyr Lys Asp Asp Asp Asp Lys
fSC
Free Secretory Component
GlcNac
N-acetylglucosamine
hSC_rD
human secretory component recombinant domain mutant
IFN
Interferon
IL
Interleukin
$ka$
Association rate constant
$kd$
Dissociation rate constant
$KD$
Equilibrium Dissociation Constant
$Ki$
Inhibitory Constant
LB
Luria-Bertani broth
MDCK
Madin Darby Canine Kidney cells
NeuAc
N-acetylneuramic acid
PBS-E
Phosphate buffered saline containing 1mM EDTA
PBS-T
PBS containing 0.05% Tween 20
plg
Polymeric immunoglobulin
plgA
polymeric Immunoglobulin A
plgM
pentameric Immunoglobulin M
plgR
polymeric immunoglobulin receptor
$PIGR$
polymeric immunoglobulin receptor gene assignment
$RLigand$
level of immobilized ligand
$R_{max}$
Maximum response defined by level of immobilized ligand
rSC
recombinant secretory component
RU
response units
SC
Secretory Component
S1gA, S1gM
Secretory Immunoglobulin A and M
SPR
Surface Plasmon Resonance
TBS
Tris buffered saline
TMB 3, 3’, 5, 5’-Tetramethylbenzidine
TNF Tumour Necrosis Factor
USF upstream stimulatory factor
\( V_{Ligand} \) valency of ligand

All standard abbreviations and biochemical nomenclature according to the conventions as outlined by Biochemical Journal (www.biochemj.org/bji2a) and the Nomenclature Committee of the IUBMB and the IUPAC-IUBMB Joint Commission on Biochemical Nomenclature (www.chem.gmul.ac.uk/iupac/jcbn).
CHAPTER 1

The Polymeric Immunoglobulin Receptor (pIgR), Polymeric Immunoglobulins and Mucosal Immunity

Polymeric immunoglobulin receptor (pIgR) is a membrane bound protein molecule, typically synthesized as a 90 kDa molecule (after glycosylation, this increases to approximately 105kDa) (Mostov et al., 1980; Mostov and Blobel, 1982), consisting of three components: an ectodomain composed of five homologous domains resembling immunoglobulin variable regions, a transmembrane component and a short cytoplasmic tail (Apodaca et al., 1991). The ectodomain is known to be cleaved and released as Secretory Component (SC) bound covalently to polymeric immunoglobulins (pIg).

Polymeric IgR is commonly expressed on the basolateral surface of secretory epithelial cells (Mostov et al., 1980) and mediates the transcytosis of polymeric immunoglobulins (dimeric IgA and pentameric IgM) from the basolateral to the apical surface of the cell (Areoti et al., 1992).

Secretory IgA (SIgA) (secretory component bound to dimeric IgA) is the first line of defence against microbial invasion at the mucosal surfaces, particularly in the gastrointestinal tract, upper and lower respiratory tracts and genitourinary system (Brandtzaeg, 2003); major areas of pathogen entry, specifically air-borne and sexually transmitted pathogens.

Numerous other functions, like eosinophil effector function activation (Lamkhioued et al., 1995; Motegi and Kita, 1998) and prevention of pneumococcal entry (Zhang et al., 2000) have been attributed to pIgR and the production of a recombinant plgR would facilitate elucidation of the function of this molecule in host pathogen relationships, as well as provide an opportunity to solve the 3D structure of this important immunological protein.
1.1 The *PIGR* Gene-Structure, Regulation and Homology

Various groups have described the gene structure, sequence and chromosomal location of the *PIGR* gene from mammalian sources ranging from the mouse (Kushiro and Sato, 1997) through to the human (Martin *et al.*, 1997). Recently the *PIGR* gene was isolated and described in the chicken, *Gallus gallus* (Wieland *et al.*, 2004). Typically the gene, located in similar positions in both the mouse and human i.e. on chromosome 1, consists of 11 exons and spans approximately 32kb (Martin *et al.*, 1997). Figure A.17 (Appendix A) shows the complete murine *PIGR* gene structure.

Every secretory immunoglobulin produced by B lymphocytes requires a plgR molecule for transport, therefore it stands to reason that constitutive expression of the receptor must occur. Using the HT29 cell line, specifically the HT29m3 subclone which exhibits increased expression of secretory component (Kvale *et al.*, 1988), it was shown that constitutive expression of the gene occurs in the presence of the upstream stimulatory factor (USF) and is controlled by a composite DNA element in the promoter region consisting of a so-called E-box, a target for transcription factors with basic helix-loop-helix leucine zipper motifs, and an inverted repeat sequence (IRS) (Johansen *et al.*, 1998).

It has been shown that steady state levels of plgR in human airway and epithelial cells can be increased by exposure to proinflammatory cytokines like interleukin-4 (IL-4), interferon-γ (IFN–γ) and tumour necrosis factor-α (TNF–α) (Ackermann *et al.*, 1999; Blanch *et al.*, 1999; Loman *et al.*, 1999; Bruno and Kaetzel, 2005). Bruno *et al.* (2004) have shown increased transcription of the *PIGR* gene in the presence of USF, IFN–γ, TNF–α and IL-1β. Blanch *et al.* (1999) have shown *in vitro* and *in vivo* that there exists a coordinate regulation of IFN regulatory factor 1 (IRF-1) and plgR. IRF-1 deficient mice show decreased levels of expression of plgR in the intestine (by 47% from the wild type) and liver (by 98%). This work led to a proposed model of plgR expression involving binding of proinflammatory cytokines by cell-surface receptors, followed by production of IRF-1, resulting in the stimulation of *PIGR* gene transcription. Recently Takenouchi-Ohkubo *et al.* (2004) have shown that retinoic acid (vitamin A), specifically
the all trans retinoic acid (ATRA), enhances transcription and expression of pIgR in the presence of TNF–α.

1.2 Structural Features of pIgR

Mostov and Blobel (1982) and Mostov et al. (1984) first described the pIgR molecule by cloning and sequencing cDNA. This revealed the presence of the five homologous immunoglobulin-like domains of the ectodomain (typically denoted by roman numerals I through V), which show a strong similarity to immunoglobulin variable regions. It has been suggested that this is as a result of the receptor’s close evolutionary origin with its target molecule. A “sixth” domain was described but this is not functionally “present” in SC (only 26 amino acids remain after cleavage) and overlaps the transmembrane region into the cytoplasmic tail. Gene expression results in a protein consisting of approximately 771 amino acids, subdivided into 3 distinct regions, namely the ligand binding ectodomain, the transmembrane domain and the cytoplasmic tail (Figure 1.1) (Apodaca et al., 1991).

Amino acid alignments have shown that functional regions are conserved across species. Banting et al. (1989) have shown that a high degree of homology exists among the cytoplasmic regions of the rat, rabbit and human; these have been attributed to sites for intracellular targeting signals. Piskurich et al. (1995) have shown that various functional sites are conserved among the mouse, rat, human, bovine and rabbit pIgR (Figure 1.2). These include a 23 amino acid pIg binding site, 10 intradomain disulfide bonds and several N-glycosylation sites.

The transmembrane domain was shown to contain a highly conserved ten amino acid sequence suggested to be a potential mechanism for transmitting signals to the cytoplasmic domain where the conserved sequences are clustered in motifs involved in polarised sorting, endocytosis and transcytosis. The subsequent sections will describe these features in greater detail.
Figure 1.1: Diagrammatic representation of pIgR showing the extracellular ligand binding domain composed of five smaller domains (I-V), the transmembrane region and the short intracellular tail. Key features highlighted are the disulfide bonds of the extracellular ligand binding domain; domain I’s non-covalent immunoglobulin binding region and V’s reactive disulfide which forms a covalent disulfide with IgA Cα3. Key signalling regions in the cytoplasmic tail are also shown (from Kaetzel, 2001).

Figure 1.2: Alignment of pIgR amino acid sequences from various species. Dashes indicate identity, dots indicate gaps. * conserved Cys residues. Domain borders are clearly marked. Human insertions and deletions are marked in red (from Piskurich et al., 1995).
1.2.1 The pIgR Ectodomain

Eiffert et al. (1984 and 1991) elucidated the complete amino acid sequence and the disulfide bond arrangement of human free secretory component. They described secretory component as an 86000 Da glycoprotein consisting of 558 amino acids with 7 carbohydrate chains bound to asparagine residues; characteristically not containing any methionine. Using peptide mapping of material collected from a single subject Hughes et al. (1997) determined that the C terminal residue of human SC was Arg$^{585}$; expanding the size of SC from 558 amino acids.

Upon enzymatic digestion, partial reduction and sequencing and peptide mapping Eiffert et al. (1984) showed that the five homologous domains vary in size from 104-114 amino acids. These domains contain 20 cysteines allowing for the formation of 10 disulfide bonds; 9 for intradomain bonds and a 10th for an interchain disulfide (Table 1.1). The presence of the tenth reactive cystine was previously shown by Cunningham-Rundles and Lamm (1975) by partial reduction and alkylation with radiolabeled iodoacetic acid. Using cyanogen bromide degradation and sequential enzymatic degradation Fallgreen-Gebauer et al. (1993) showed that the cysteine residue, Cys$^{467}$, is involved in the covalent linkage between the secretory component of pIgR and Cys$^{311}$ of the $\alpha 3$ domain of one of the immunoglobulin molecules present in dimeric IgA (dIgA). The Cys$^{311}$-Cys$^{467}$ disulfide is involved in the stabilization and protection of dIgA against proteolytic cleavage, a function postulated by Lindh (1975). Interesting to note is that IgM does not form a covalent bond with pIgR/SC; this phenomenon is as a result of the C terminal domain of the immunoglobulin controlling the interaction (Braathen et al., 2002). Prior to attachment of the pIgR to dIgA both Cys$^{311}$ residues of dIgA are involved in an inter-heavy chain disulfide bond between the two $\alpha 3$ chains. Residues Cys$^{467}$ and Cys$^{501}$ of pIgR are involved in a reactive disulfide bond prior to linkage (Fallgreen-Gebauer et al., 1993).
Table 1.1: Disulfide bond arrangement of pIgR ectodomain. “Large” and “Small” refer to spatial arrangement of the bonds and “Extra” refers to the reactive disulfide bond (adapted from Eiffert et al., 1984).

<table>
<thead>
<tr>
<th>Disulfide Bond Type</th>
<th>Domain</th>
<th>Disulfide Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Large” bonds</td>
<td>I</td>
<td>Cys22—Cys92</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>Cys134—Cys202</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>Cys238—Cys306</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>Cys352—Cys422</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>Cys463—Cys525</td>
</tr>
<tr>
<td>“Small” bonds</td>
<td>I</td>
<td>Cys38—Cys46</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>not present</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>Cys252—Cys260</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>Cys366—Cys376</td>
</tr>
<tr>
<td></td>
<td>V</td>
<td>Cys472—Cys484</td>
</tr>
<tr>
<td>“Extra” bonds</td>
<td>V</td>
<td>Cys467—Cys501</td>
</tr>
</tbody>
</table>

The five aforementioned homologous domains of pIgR/SC are the key to the formation of secretory antibodies. Studies have shed little light on the individual domains with most of the focus being on domain I.

1.2.1.1 The amino terminal domain of pIgR/SC, domain I

Upon investigation of the individual domains, Frutiger et al. (1986) showed that the amino-terminal domain of rabbit SC (i.e. domain I) was responsible for non-covalent binding to dIgA by investigating the interaction between tryptic cleavage fragments and dIgA. Bakos et al. (1991) were the first to characterize a critical binding site of human pIgR domain I responsible for the non-covalent binding. Upon tryptic digestion of domain I it was found that the fragment containing residues 15-37 (i.e. amino acid residues GNSVSITCYYPPTSVRHRTRKYW) were functionally important for binding. These residues were also shown to have 78% identity between human, rabbit and rat SC. Piskurich et al. (1995) expanded this to include the murine and bovine species.

Coyne et al. (1994) have shown via molecular modelling techniques that the residues elucidated by Bakos et al. (1991) were on an exposed loop known as the complementarity determining region (CDR, in this case CDR1) similar to those found in immunoglobulins.
Two other loops, CDR2 and CDR3, were also described but the residues contained within do not appear to be as homologous among species; although some residues are, however, conserved. Upon mutational analysis by replacement of residues from predicted CDR loops of domain II into domain I it was shown that the three CDR-like loops (CDR1, CDR2 and CDR3) are involved in binding. The presence of these CDR-like regions on exposed loops of domain I suggested that the interaction between pIgR and dIgA is analogous to that of antibody-antigen interactions (Coyne et al., 1994). Mutation of the CDR-like loops of domain I resulted in either slightly decreased binding affinity or almost complete loss of binding to dIgA; particularly CDR1, which is essential. Røe et al. (1999) expanded the knowledge on the functionality of CDR-like loops in domain I for binding of pentameric IgM, particularly CDR2. The CDR1 loop is highly conserved between the rabbit and human and since rabbit pIgR barely binds IgM, other determinants must play a role. Using rabbit/human chimaeras, human/rabbit domain I substitutions were shown to give rabbit pIgR the ability to bind IgM (Røe et al., 1999). The CDR2 loop was shown to be essential for the interaction with IgM, CDR1 was found to contribute substantially to binding and CDR3 was required for maximal binding.

Bakos et al. (1994) managed to express functional recombinant domain I (with both IgA and IgM binding capacity) in Escherichia coli strain BL21 (DE3). Table 1.2 illustrates the relative affinities compared to IgM purified secretory component. The data shows that domain I contributes significantly toward whole molecule binding affinity.

**Table 1.2: Relative affinities of recombinant domain I for IgA and IgM, compared to IgM affinity purified human secretory component.** SD, standard deviation; $K_D$, equilibrium dissociation constant (adapted from Bakos et al., 1994).

<table>
<thead>
<tr>
<th>Secretory Component ($K_D\pm SD$)</th>
<th>Recombinant Domain I ($K_D\pm SD$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=5)</td>
</tr>
<tr>
<td>$dIgA$</td>
<td>4.8 ± 0.8 nM</td>
</tr>
<tr>
<td>$pIgM$</td>
<td>6.4 ± 0.5nM</td>
</tr>
</tbody>
</table>

Hamburger et al. (2004) elucidated the 1.9Å crystal structure of domain I (Figure 1.3), confirming the features proposed by other groups and shedding new light on functional groups. Mostov et al. (1984) showed the immunoglobulin-like structure of the domain
consisting of two anti-parallel β sheets composed of 10 β strands named A, B, E and D, and C'', C', C, F, G and A’, respectively (Figure 1.4).

**Figure 1.3:** Ribbon structure of domain I showing characteristic arrangement of CDR loops 1, 2 and 3 and the 10 anti-parallel β-strands (see text) (from Hamburger et al., 2004).

**Figure 1.4:** Topological diagram of pIgR domain I and positions of the CDR loops (from Hamburger et al., 2004).

The first and most notable feature is that the CDR loops do not conform to the known listed canonical conformations exhibited by these structures as found in immunoglobulins:

- CDR1 exhibits a single helical turn, notably containing the conserved residues 15-37 (as outlined by Bakos et al., 1991);
• CDR2 is very short and exhibits a tight turn;
• CDR3 is tilted toward the $\beta$ sheet, C’’, C’, C, F, G and A, (Figure 1.5).

A disulfide bond, between Cys$^{22}$ and Cys$^{92}$ referred to as the large disulfide (Eiffert et al., 1984), links the $\beta$ sheets and Cys$^{38}$ and Cys$^{46}$ link the C and C’ strands. Another feature includes salt bridges between Arg$^{63}$ and Asp$^{86}$.

Figure 1.5: Overlap of IgV$\lambda$ domain and domain I showing structural similarity and non-canonical CDR3 loop arrangement. The modelled plgR D1 dimer figure shows the improbability of plgR dimer formation, a function of the CDR3 loop in immunoglobulins (from Hamburger et al., 2004)

1.2.1.2 Domain II, III, IV and V

Domains II and III have been implicated in enhancing the interaction of plgR with dIgA and the formation of the covalent link, i.e. the disulfide bridge. Frutiger et al. (1987) showed the presence of two forms of SC in rabbits; a high molecular weight and a low molecular weight form. The low molecular weight form was subsequently shown not to covalently bind dIgA although non-covalent binding did occur; this form was shown to contain a domain II-III deletion. Crottet and Corthesy (1999) investigated interactions between SC and dIgA by replacing predicted loops (CDR-like regions) with a FLAG octapeptide (DYKDDDDK) sequence (Crottet et al., 1999a). They showed that creating FLAG domain II and domain III mutants where the FLAG sequence was buried resulted in abrogation of the molecule’s covalent binding capacity. Furthermore, expression of domain deletion mutants in Madin Darby Canine Kidney (MDCK) cell monolayers revealed differing domain requirements in IgA and IgM binding events (Norderhaug et al., 1999). Simultaneous deletion of domain II and III resulted in almost complete loss of the plgR molecule’s IgA transcytosis ability, seemingly relating the covalent linkage with
its ability to affect IgA transcytosis. Deletion of domains IV and V affected IgA translocation although not significantly (Figure 1.6).

![Histogram comparing translocation of 100nM pIgA and pIgM across MDCK cell monolayers (polarized) expressing wild type pIgR, and the deletion mutants pIgRΔ2,3 and pIgRΔ4,5 (control = MDCK untransfected) (from Norderhaug et al., 1999).](image)

Figure 1.6: Histogram comparing translocation of 100nM pIgA and pIgM across MDCK cell monolayers (polarized) expressing wild type pIgR, and the deletion mutants pIgRΔ2,3 and pIgRΔ4,5 (control = MDCK untransfected) (from Norderhaug et al., 1999).

Transcytosis of IgM, on the other hand, was not significantly affected by either simultaneous deletion of domains II and III, or IV and V. No distinct function has yet been attributed to domain IV in polymeric Ig binding. Conversely it appears that domain IV does have a function in the pathogenicity of *Streptococcus pneumoniae* by mediating (along with domain III) an interaction with the choline binding protein, CbpA (Lu et al., 2003). As previously stated domain V interacts (via Cys\(^{467}\)) with Cys\(^{311}\) of IgA to form a permanent covalent linkage.

All the data seem to point to the fact that the non-covalent nature of the bond between IgM and domain I is the dominant interaction in the SC/IgM interaction.

### 1.2.1.3 N-glycosylation and Binding Activity

Hughes *et al.* (1999) showed that the 7 asparagine residues at positions 65, 72, 117, 168, 403, 451 and 481 were N-linked to several forms of glycosyl residues (Figure 1.7).
Upon investigating the role of N-glycosylation of rabbit pIgR and its role in binding to the ligand, Elkon (1984) and Frutiger et al. (1988) showed, respectively, that sugar composition is unrelated to immunoglobulin binding and that in three allotypes, variation in the degree of glycosylation and the sites did not measurably affect the non-covalent binding of domain I. The non-requirement of glycosylation was reiterated by Bakos et al. (1991 and 1994) and Hamburger et al. (2004) by chemical/enzymatic deglycosylation of SC and recombinant expression of domain I, respectively. Recently, Prinsloo et al. (2006) showed that unglycosylated recombinant SC expressed in E.coli maintains its ability to
bind IgM. Glycosylation, however, does play an active role in the immunological functions of SC, e.g. only glycosylated SC bound to dIgA can mediate localization of SIgA to bronchial mucus (Phalipon et al., 2002).

1.3 Secretory Immunoglobulins (SIgA and SIgM)
IgA and IgM represent approximately 10 and 15% of the total immunoglobulin pool (Prescott et al., 1996), respectively. Secretory immunoglobulins are the first line of defense against microbial infection at the mucosal surface (Brandtzaeg, 2003). Typically the secretory mucosa and exocrine glands contain >80% of immunoglobulin producing cells in the adult human (Brandtzaeg et al., 1989).

1.3.1 Structural and Biological Features of Secretory Immunoglobulins
1.3.1.1 IgA
IgA is the dominant antibody in external secretions (gastric fluids, saliva, tears, etc.) (Benjamini et al., 2000). Monomeric IgA typically has a molecular weight of 160 000 Da which increases to just over 400 000 Da by dimerization and complex formation with the J (joining) chain (~15 kDa) and secretory component (80-90 kDa). Two genetically distinct subclasses exist in humans, IgA1 and IgA2 that are expressed by separate genes in a tissue specific manner (Mestecky and McGhee, 1987). The predominant subclass is IgA2 (produced en mass in gastrointestinal secretions), which characteristically exhibits a 13 amino acid deletion in the α2 hinge region, a distinction which accounts for the resistance to proteolytic cleavage by IgA1 proteases synthesized typically by Neisseria meningitides, Haemophilus influenzae and Streptococcus pneumoniae. IgA1 subclass production does exceed IgA2 in nasal and bronchial secretions (Brandtzaeg et al., 1999). Boehm et al. (1999) showed that IgA1 exists in a T-shaped structure, with Fab tips separated by 23nm, opposed to the traditional Y-shaped immunoglobulin structure (e.g. IgM Fab tips are separated by 13nm [Perkins et al., 1991]). This has been attributed to the hinge region. Furthermore, Boehm et al. (1999) have shown that the T-shape allows for maximal accessibility of the Fab antigen binding sites and the FcαR binding site, and exposure of the hinge region. This allows for a functional link between antigen binding and effector functions. X-ray scattering curve fits were shown to be consistent with a labile Cys471-Cys311 disulfide bridge and that the surface of IgA1 is large enough for
simultaneous complex formation with J chain and pIgR (Boehm et al., 1999). Cys$^{311}$ is known to be involved in the covalent linkage with Cys$^{467}$ of secretory component (Fallgreen-Gebauer et al., 1993). Cys$^{471}$ is involved in dimer formation and disulfide bridge formation with the J chain (Mestecky et al., 1974; Atkin et al., 1996). The solution structure of IgA$_2$ revealed a similar T-shaped structure but more compact than IgA$_1$; this was attributed to the shorter hinge region. The compact T-shape conformation of IgA$_2$ allows access to the Fc$\alpha$R site (Furtado et al., 2004). Furthermore, deduced from the flexible and compact structures; it is suggested that the two subclasses exist in order to give IgA a broader range of antigens (Furtado et al., 2004).

No X-ray or NMR structural analysis of the IgA:J chain:SC complex exists to date. Figure 1.8 highlights essential features of IgA$_1$, IgA$_2$ and SIgA. The functions of SlgA are discussed in section 1.5.

1.3.1.2 IgM

Immunoglobulin M is a large macromolecule produced first in a response to immunization [a function attributed to IgM’s proclivity to activate complement (Taylor et al., 1994)]. The molecular weight of IgM is approximately 970 000 Da (at least in pentameric IgM; Figure 1.9). IgM exists in both pentameric and, to a lesser extent, hexameric forms; the latter lacking the J chain, a feature that prevents complex formation
and transcytosis by pIgR. The heavy chains of IgM exhibit the added feature of an extra Cµ domain. Brewer and Corley (1997) argued that the pentameric J chain containing form of IgM is a function of thermodynamic stability and that J chain is not associated strongly with any IgM assembly intermediates prior to “pentamerization”. Simply put, late assembly events in IgM complex formation prior to maturation in the golgi appear to favour inclusion of the J chain over a sixth monomer. This accentuates the findings of Niles et al. (1995) that pentameric IgM formation is favoured in vitro in the presence of excess J chain. De Lalla et al. (1998) showed that a carboxy terminal cysteine, Cys\textsuperscript{464}, and a glycan attached to Asn\textsuperscript{452} are essential for IgM polymerization and association with J chain. Perkins et al. (1991) showed in their solution structure of IgM that the best fit for J chain in the complex was between two Cµ4 chains. X-ray scattering curve fits also showed the planar structure of IgM (a ring of Fc monomers) (Perkins et al., 1991). The planar structure was shown to be possibly controlled by the µ-tailpiece glycans (De Lalla et al., 1998). Structural analysis allowed for the identification of the complement C1q binding site on the Cµ3 domain. Analysis of the solution structure of C1q showed that binding can induce a conformational change allowing for complement activation (Perkins et al., 1991).

Figure 1.9: Diagrammatic representation of pentameric IgM showing key structural features. Light chains are shown in green and heavy chains in blue. Orange circles show sites of glycosylation (from Benjamini et al., 2000).
1.3.1.3 J Chain

Initially described independently by Halpern and Koshland (1970) and Mestecky et al. (1971), J or joining chain has become a central feature in secretory immunoglobulin structure and function. J chain was shown to be 7.6% glycosylated with a molecular weight of 15.6 kDa and 137 amino acid residues (Neidermeier et al., 1972; Schrohenloher et al., 1973). J chain is synthesized along with IgA and IgM in plasma cells. Secondary structure analysis of human J chain revealed a secondary structure similar to immunoglobulin V_\text{L} domains; probably an evolutionary trait similar to that of pIgR, showing the parallel evolution of receptor and ligand (Zikan et al., 1985). A putative single domain structure of an 8 stranded anti-parallel $\beta$-barrel with 37% $\beta$-sheet (experimental value was 34%) and the rest composed of random coils was proposed (Zikan et al., 1985). Conversely, upon primary structure analysis of the murine J chain, Cann et al. (1982) proposed a two domain model. Zikan et al. (1985), however, argued that a single immunoglobulin like domain model is “more compatible with quaternary structures typical for functional molecules in the immunoglobulin superfamily”. Furthermore, they stated that the single domain model allowed for the inter- and intra-chain domain interactions required for polymeric immunoglobulin assembly and transport mediated by pIgR binding. The preceding sections have highlighted the function of the J chain [postulated initially by Halpern and Koshland (1970) and Mestecky et al. (1971)], i.e. its core role in the formation of polymeric immunoglobulins; most elegantly the formation of pentameric IgM (section 1.3.1.2). J chain containing polymeric immunoglobulins are transcytosed selectively (Brandtzaeg and Prydz, 1984; Hendrickson et al., 1995; Johansen et al., 2000). This hints at an interaction between pIgR ectodomain and J chain; perhaps J chain allows for correct positioning of pIgR, allowing for efficient transcytosis.

1.4 The Short Cytoplasmic Tail, Transcytosis and Mucosal Immunity

Aroeti et al. (1992) summarized the process of polymeric immunoglobulin transcytosis as follows (Figure 1.10): (1) membrane bound pIgR is synthesized in membrane bound polysomes of the rough endoplasmic reticulum; (2) the molecule passes through the Golgi apparatus, and (3) is targeted to the basolateral surface; (4) the pIgR ectodomain is
now exposed and allows for binding of pIg; (5) the ligand/receptor complex is
endocytosed and transported to the apical surface; (6) secretory component/pIg is
proteolytically cleaved and (7) released at the apical surface. The process outlined in
Figure 1.10 is controlled by signal sequences in the cytoplasmic tail of pIgR (Casanova et al., 1991); i.e. a 17 residue sorting signal situated between residues 653-670. Reich et al. (1996) have shown that three residues (His$^{656}$, Arg$^{657}$ and Val$^{660}$) in the sorting signal are responsible for basolateral sorting because Ala mutations of these residues resulted in decreased basolateral targeting and increased apical targeting. These residues form part of two functional domains with the Val$^{660}$ residue apparently being in a $\beta$-turn secondary structure while the other two residues are not in the turn. The cytoplasmic domain also appears to mediate different signals for basolateral sorting with the Val$^{660}$ residue being the critical residue controlling sorting. Chapin et al. (1996) suggested that calmodulin (CaM) may be involved in regulation of transcytosis in the presence of Ca$^{2+}$. A CaM binding site was identified within the 17 residue sorting signal. It was speculated that Ca$^{2+}$ release triggered by a hormonal signalling event, or by dIgA binding, could stimulate CaM binding and therefore prevent basolateral targeting and allow transcytosis to occur.

Figure 1.10: Transcytosis of polymeric immunoglobulins by pIgR. JC, junctional complex (see text for details) (adapted from Mostov, 1994).
Casanova et al. (1990) investigated the signals required for transcytosis and identified that phosphorylation of Ser<sup>664</sup> results in transcytosis. Mutation to Ala and Asp resulted in decreased and increased (due possibly to negative charge mimicking phosphate charge) transcytosis, respectively. The ability of the ligand-binding event to regulate transcytosis was shown by Luton et al. (1998) and Singer and Mostov (1998). This entails the activation of protein tyrosine kinases, concomitant phosphorylation of phospholipase C-γ1 and production of inositol triphosphate; this results in the activation of protein kinase C and signals up-regulation of transcytosis. Deletion of residues 727-736 have been shown to completely negate this process and has been targeted as potentially controlling constitutive expression of pIgR.

1.5 Functions of pIgR, Secretory Component

The initial function of SC was thought to be protection of plg against proteolytic cleavage (Lindh, 1975). Figure 1.11 shows some functions of pIgR and SC elucidated recently. Apart from the polymeric immunoglobulin transport function, pIgR also exhibits non-inflammatory methods of defence. These include immune exclusion of antigens from the epithelial cell (Phalipon et al., 2002), intracellular neutralization of pathogens (Schwatz-Cornil et al., 2002), and transport of pathogens across epithelia (Robinson et al., 2001).

Figure 1.11: Modes of defence mediated by dIgA, pIgR and fSC. (1) transport of polymeric immunoglobulins and release of fSC; (2) transport of pathogens across epithelia; (3) intracellular neutralization, (4) immune exclusion, (5) prevention of binding and entry of pathogens by fSC and (6) pathogen invasion by subversion of pIgR (from Phalipon and Corthésy, 2003).
Phalipon et al. (2002) showed that only glycosylated SC bound to SIgA formed interactions with bronchial mucus, localizing the SIgA and enhancing the protective function of SIgA; in this case against *Shigella flexneri*.

Interaction of pIgR/SC with bacteria and bacterial products can be both beneficial and deleterious. Free unbound secretory component has been postulated to prevent cellular entry of pathogens e.g. enterotoxigenic *E. coli* (de Oliveira et al., 2001). SC has been shown to reduce toxic effects of *Clostridium difficile* toxin A (Dallas and Rolfe, 1998). Although it has been shown that fSC interaction with pneumococci prevents pathogen entry, unfortunately the cell bound pIgR molecule can also be hijacked by *Streptococcus pneumoniae* using surface receptors (pneumococcal choline binding protein A, CbpA) and subverted for their own purposes, specifically entry into epithelia (Zhang et al., 2000; Lu et al., 2003). This has been shown to be a specific interaction between CbpA and domains III and IV of SC/pIgR and is not dependant on glycosylation (Lu et al., 2003).

Free SC (and SIgA) has been shown to stimulate effector functions (degranulation and superoxide production) of eosinophils but not neutrophils. (Motegi and Kita, 1998). This is dependant on β-integrins. The possibility that SC plays a role in a novel function to regulate mucosal tissue inflammation cannot be underestimated. Actually, this function can be considered potentially ancient if one considers that SC production precedes IgA production in foetal development (Moro et al., 1991).

The possibility of SC in the maintenance of pregnancy has also been highlighted by Bennett et al. (1999). Known for years to gynaecologists as gravidin, SC was shown to be a powerful inhibitor of phospholipase A₂, in turn maintaining levels of prostaglandin. Prostaglandin is known to induce labour. The interaction between SC and phospholipase A₂ may prevent premature birth.

It is clear that pIgR/SC has numerous roles. The interaction and mechanisms behind many of these functions are poorly understood and require further study.
CHAPTER 2

Aim of Research

Free unbound secretory component (fSC) does not require \( N \)-glycosylation for polymeric immunoglobulin binding capacity/activity. Recombinant domain deletion mutants of human fSC, expressed in \textit{Escherichia coli} BL21(DE3), will produce unglycosylated recombinant proteins. These recombinant domain deletion mutants would allow for an investigation of the binding affinities of individual or concomitant domains.

The unique domain structure of the pIgR ectodomain allows for specific targeting of N and C terminal borders using standard RT-PCR and PCR techniques. Amplified cloned pIgR ectodomain cDNA, from the mammalian cell line HT29, would permit the targeted amplification of single or multiple domain mutants (Figure 2.1).

![Figure 2.1: Schematic diagram showing the outline of the pIgR ectodomain encoding the 5 domains (I-V). The coloured arrows indicate the desired amplification products as defined by the primers in Table 3.1. The sizes (bp) of desired fragments are shown in brackets. Blue arrow, hSC_rDI (344), Green arrow, hSC_rDI-II (665), Grey arrow, hSC_rDI-III (986), Red arrow, hSC_rDI-IV (1325) and Orange arrow, hSC_rDI-V (1658). The double headed brown arrow shows the amplified region encoding domain IV and V (672). The white arrow indicates the desired amplification product encoding domain V (hSC_rDV [333]). Grey/black crosshatch box represents cloned pIgR ectodomain cDNA.](image)

Primers designed with unique restriction sites would allow for directional insertion into an appropriate expression vector, specifically the pET vector system. Overexpression of
the cloned cDNAs would allow for production of large quantities of recombinant protein. A potential pitfall of this approach is the formation of inclusion bodies instead of soluble protein. This can be overcome by \textit{in vitro} refolding using buffered oxido-shuffling reagent couples and low molecular weight molecules to aid in correct disulfide formation and stabilizing refolding complexes, respectively. The pET vectors often confer unique fusion tags (e.g. the 6x neighbouring histidine tag) allowing for simple, potentially one-step purification of either insoluble or soluble recombinant proteins.

Binding characteristics and affinities of the recombinant domain mutants can be investigated using various immunological (dot blot, receptor-ligand ELISA and whole cell ELISA) and physical (real time biosensors using surface plasmon resonance) techniques.

Our current understanding of pIgR ectodomain binding to plgs is based almost solely on the expression of pIgR in mammalian and insect cell culture, except, of course, the expression of human domain I in \textit{E.coli} by Bakos \textit{et al.} (1994). As it is known that glycosylation is not required for plg binding capacity, a better understanding of protein-protein interaction would most certainly be gained from the production of unglycosylated mutants. Most interaction studies have focused on dIgA binding, traditionally considered to be the favoured binding partner of pIgR. However, pentameric IgM has been shown to actually have a higher affinity for pIgR, yet its interaction with domains II-V is not well understood.

Considering the above statements, the domain deletion mutants expressed in \textit{E.coli} should possess polymeric immunoglobulin binding activity when refolded (\textit{in vivo} or \textit{in vitro}). This is the first time that multiple domain mutants of human pIgR ectodomain would be expressed in a bacterial system. Using refolded recombinant domain deletion mutants of human fSC it would be possible to investigate biochemical and biophysical contributions of single or multiple domains in the interaction with polymeric immunoglobulin M (IgM).
CHAPTER 3

Production, Purification and Characterization of Recombinant Domain Mutants

Bacterial expression systems have allowed for the elucidation of structure and function of numerous proteins as well as the production of medically important products; most notably the production of recombinant human insulin (Goeddel et al., 1979).

Recombinant proteins produced by bacteria are, however, often found as insoluble deposits known as inclusion bodies (Taylor et al., 1986). The major problem with the production of inclusion bodies is clear; renaturation procedures to reinstate a native conformation are required. Misawa and Kumagai (1999), conversely, argue in favour of inclusion bodies stating that the production of inclusion bodies may be favourable. This is true when considering that the native conformation of the recombinant protein may be toxic to the host cell or even unstable in the cytoplasm, prone to proteolysis. Inclusion bodies also offer the distinct advantage of easy purification owing to their density and size (Georgiou and Valax, 1999).

The development of novel tag systems, conferred as fusion tags by expression vectors, exploiting unique interactions, allow for relatively one step purification procedures from complex cell lysates. One such system is the strong affinity observed between neighbouring histidine residues and chelated Ni$^{2+}$ ions (Porath et al., 1975, and Hochuli et al., 1987).

In vitro refolding techniques of inclusion bodies have improved significantly, allowing for production of high levels of refolded material. A major concern is the correct disulfide arrangements in highly complex proteins. Renaturation procedures of inclusion bodies require complete denaturation, typically by the use of high molarity buffered urea
(8M) or guanidine HCl (6M) and β-mercaptoethanol to ensure complete reduction of cysteines. Refolding entails the complete removal of denaturants and reducing agents and the introduction of oxidizing conditions (Misawa and Kumagai, 1999; De Bernardez Clark, 2001). Numerous methods exist to ensure that recombinant proteins are efficiently renatured; these include dilution refolding (De Bernardez Clark, 2001), buffer exchange by dialysis (West et al., 1998) and on column refolding by virtue of Ni\(^{2+}\) affinity being maintained under denaturing conditions (Rogi et al., 1998). As mentioned, problems are encountered when considering complex proteins containing numerous disulfides (e.g. immunoglobulins). Systems to control orderly formation of disulfides include the use of oxido-shuffling reagent couples, e.g. reduced and oxidized glutathione (GSH/GSSG) and cysteine/cystine (De Bernardez Clark et al., 1999). Additional additives to refold buffers include the use of low molecular weight molecules which act as artificial chaperones like β-cyclodextrin and L-arginine (Daugherty et al., 1998; De Bernardez Clark et al., 1999) stabilizing refolding complexes. Baneyx (1999) has highlighted the in vitro use of native foldases and chaperones in refolding buffers.

The potential to retrieve active soluble proteins from bacterial expression strains has been improved considerably by the introduction of oxidizing mutant E. coli strains (e.g. FA113 [Bessette et al., 1999]). These strains typically are suppression mutants which actively suppress the thioredoxin reductase (trxB) and glutathione reductase (gor) genes responsible for redox interconversion between free sulfhydryl and disulfide bonded conformations (Derman et al., 1993; Äshlund et al., 1999). Mutation of these genes would naturally allow for the stable formation of disulfide bonded recombinant proteins.

Using standard RT-PCR and PCR techniques the pIgR ectodomain was amplified and cloned allowing for the targeted amplification of cDNAs representing single and multiple domain mutants of the pIgR ectodomain (basically mutant fSCs). These cDNAs were cloned into expression vectors and transformed into appropriate bacterial expression strains allowing for high level production of recombinant domain mutants. The production of domain mutants as inclusion bodies can be overcome using in vitro refolding.
3.1 Materials and Methods

3.1.1 Reagents
The RNeasy™ Minikit was from QIAgen GmbH (Hilden, Germany); Titan™ One Tube RT-PCR kit, isopropyl-β-D-thiogalactopyranoside (IPTG), X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside), lysozyme, ethidium bromide, PVDF membrane, bovine serum albumin (BSA), and Marker XVI from Roche (Mannheim, Germany); GeneRuler 1kb DNA Ladder from MBI Fermentas (Germany); pCRScript SK (+) Cloning kit from Stratagene (Cedar Creek, Texas, USA); pET22b (+), E. coli BL21 (DE3), ROSETTA (DE3) and ORIGAMI (DE3) from Novagen (Madison, Wisconsin); E. coli DH5(α) from Invitrogen GmbH (Karlsruhe, Germany). Electrophoresis grade agarose from Hispanagar (Spain). EasyjecT® electroporator from Equibio (England). The His-Select™ nickel affinity resin, bacterial protease inhibitor cocktail, Tween 20, Micro-Biotinylation Kit, polyclonal anti-human secretory component, human colostral SIgA, human serum IgM, extravidin-peroxidase conjugate and SigmaFast™ 3, 3-diaminobenzidine were from Sigma Aldrich (St. Louis, Missouri, USA). The Superdex 75 (26/60) FPLC column was from Amersham Pharmacia Biosciences AB (Uppsala, Sweden); Ndel and XhoI were from New England Biolabs (Massachusetts, USA). Microbank™ cryopreservation vials from Pro-Lab Diagnostics (Ontario, Canada); lysylendopeptidase was from Wako Chemicals (USA). All reagents were of the highest analytical grade available.

3.1.2 Bacterial Cultures and Media
Bacterial strains used for cloning and expression were E. coli DH5α, E. coli BL21 (DE3), E. coli BL21 ROSETTA (DE3) and E. coli ORIGAMI (DE3). Untransformed bacterial cultures were grown and maintained in LB (Luria-Bertani [10g Tryptone, 10g NaCl, 5g Yeast extract, if required 15g bacteriological agar]) broth and LB agar. Transformed bacterial strains were maintained in LB broth and agar containing 200µg/ml ampicillin (LB-AMP) as a selection agent. All cultures were grown at 37°C unless stated otherwise. Single pure colonies (grown overnight on LB-AMP agar) of transformed bacterial strains were cryo-preserved in Microbank™ vials at -80°C.
3.1.3 Amplification and Molecular Cloning

Human plgR mRNA was isolated from HT29 cells grown in monolayer using the QIAgen RNeasy Minikit. RT-PCR was performed with the Titan™ One Tube RT-PCR kit with gene specific primers for the plgR ectodomain (5’ primer AAA AAA ACA TAT GAG TCC CAT ATT TGG; 3’ primer AAG AAT TCA GGA GCT GCT TCC ACC TTG).

PCR conditions were 2 minutes at 95°C for denaturation followed by 20 cycles of 95°C/45 seconds; annealing at 53°C/1 minute; extension at 68°C/3 minutes. This was followed with a final extension of 10 minutes at 68°C. PCR product was cloned and screened for in E. coli DH5(α) using the pCRScript SK (+) Cloning kit using blue-white selection on LB-AMP agar plates containing IPTG and X-gal.

Domain deletion mutants were amplified by PCR using domain-specific primers (Table 3.1) for the ectodomain encoding PIGR gene. Typical reaction mixtures contained 1U Pfu Polymerase, 2.5µl 10x Polymerase buffer, 250µM dNTPs, 4µM of each primer, 1.25µl DMSO, 50ng template DNA, made up to 25µl with sterile ddH2O. PCR was performed with 25 reaction cycles composed of a 45 second denaturation at 95°C, annealing for 60 seconds at 60°C and an extension for 3 minutes at 72°C followed by a 10 minute final extension at 72°C. Amplified PCR fragments and an appropriate marker (Marker XVI) were electrophoresed on a 1% (w/v) Tris acetate EDTA (TAE, pH 8.0) - agarose gel containing 5µg/ml ethidium bromide at 120V and 60mA for 45 minutes and visualised on a transilluminator.

Table 3.1: Primer sequences used in the amplification of domain deletion mutants (see text for details). Nde I sites are denoted by CATATG and Xho I sites are denoted by CTCGAG. Boxes shaded grey denote the primers targeting the 5’ portion of individual domain regions. The forward primer for hSC_rDI was used for all multiple domain mutants unless stated otherwise. Forward primers all contain start codon ATG.

<table>
<thead>
<tr>
<th>Forward 5’-3’</th>
<th>Reverse 5’-3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>hSC_rDI</td>
<td>AAA AAA ACA TAT GAG TCC CAT ATT TGG</td>
</tr>
<tr>
<td>hSC_rDI-II</td>
<td>AAA AAA ACA TAT GAA TGA CAC TAA AG</td>
</tr>
<tr>
<td>hSC_rDI-III</td>
<td>AAA AAA ACA TAT GGG CTC AGT GAC</td>
</tr>
<tr>
<td>hSC_rDI-IV</td>
<td>AAA AAA ACA TAT GGA GTC CAC GAT TC</td>
</tr>
<tr>
<td>hSC_rDI-V (rSC)</td>
<td>AAA AAA ACA TAT GGG AGA ACC AAA C</td>
</tr>
</tbody>
</table>
Amplified cDNA bands were excised and purified by gel extraction (QIAGEN gel extraction kit). Fragments were blunt ligated into pCRScript SK (+). Ligation products were transformed by electroporation using an EasyjecT® at 17000V/cm into E.coli DH5α for blue/white selection on LB-AMP. Positive clones were cultivated in 50ml flasks of LB-AMP and plasmids were isolated using the Plasmid Maxi Prep (QIagen), checked for size and purity by gel electrophoresis as previously described and digested with Nde I and Xho I. A typical restriction contained 10U of each enzyme, 1µg plasmid DNA, in a 10µl reaction buffer made up to 10µl with sterile ddH2O and incubated at 37°C overnight to ensure complete digestion. Digested bands were gel extracted and ligated (at 4°C overnight) with 1U T4 DNA ligase into similarly digested pET22b(+). Ligation mixes were transformed into E. coli BL21 (DE3). Other strains also used for transformation and expression included E.coli ROSETTA (DE3) and ORIGAMI (DE3).

Plasmids were isolated and sent for automated DNA sequencing at the DNA Sequencing Service, Central Analytical Facility, University of Stellenbosch, South Africa. The T7 promoter (5’TAATACGACTCACTATAG3’) and T7 terminator (5’CCGCTGGAGCAATAACTAG3’) primers were used for sequencing. Where necessary (i.e. depending on sequence length) the primers in Table 3.1 were used in addition to the above sequencing primers. DNA sequences were edited in Chromas 1.45 and contig overlaps (forward and reverse) were analysed using Vector NTi v9.0 (Invitrogen). DNA sequence alignments were performed using CLUSTAL X 1.83 (Thompson et al., 1997) sequence alignment software. Plasmid maps were drawn using ApE. Clones were screened for protein expression in 50 ml LB-AMP. Protein expression was induced by IPTG (1mM) as the protein expression is under the influence of the lacI promoter/T7 fusion region. Production of recombinant proteins of required size was determined by electrophoresis on a 15% acrylamide sodium dodecyl sulphate polyacrylamide gel (SDS-PAGE). Electrophoresis was performed at 200V, 64mA for 45 minutes in a BIORAD mini-gel system using the method of Laemmli (1970).

3.1.4 Preparative Expression and Isolation of Inclusion Bodies
One litre (4 litres total) preparative protein expressions were performed in 3L borosilicate flasks (Schott-Duran). Five percent inoculations from overnight cultures grown in LB-
AMP were performed and cell growth was monitored at 600nm (Ultrospec 2000, Pharmacia) until optical density reached ~ 0.7 (mid log phase). Expression was induced by addition of 1mM IPTG for 4 hours. After 4 hours the cells were harvested by centrifugation (5000g for 5 minutes) and resuspended in 25ml sonication buffer (50mM Tris HCl, 0.3M NaCl, 1mM EDTA, 0.2% (w/v) NaN₃, pH 8.0). Inclusion bodies were isolated by addition of lysozyme (2mg/ml) to the cell paste and incubation at room temperature for 1 hour. The cells were lysed by sonication at 50% duty cycle, 40% power for 3 bursts of 6 minutes (HD2200 Sonopuls, Bandelin fitted with a UW2200 probe) and centrifuged for 40 minutes at 42000g. The cell pellet was homogenised and washed with inclusion body wash buffer (50mM Tris HCl, 0.3M NaCl, 1mM EDTA, 0.2% (w/v) NaN₃, 0.5% (w/v) N,N-dimethyldodecylamine N-oxide, pH 8.0 [LDAO]) to remove soluble *E. coli* proteins. The final pellet was homogenised in sonication buffer and stored at 4°C in the presence of bacterial protease inhibitor cocktail. Inclusion body purity was checked by SDS-PAGE as previously described. Inclusion bodies were transferred onto PVDF membranes for N-terminal sequencing (Prof. Koji Muramoto, Tohoku University, Sendai, Japan) (Muramoto *et al.*, 1993).

### 3.1.5 In Vitro Refolding of Recombinant Protein

Inclusion bodies were purified by batch purification using NiCAM (Sigma) resin (see Section 3.1.6.1 for method). Proteins were refolded by dialysis to facilitate a slow removal of the denaturant and extended refold kinetics. Equilibrium gradient dialysis was used for refolding of domain deletion mutants to facilitate formation of disulfide bonds. Dialysis was performed in the presence of arginine (acting as a small molecule chaperone) and cysteine/cystine (oxidative/reductive pair) to prevent formation of aggregates and aid formation of disulfide bridges, respectively. Inclusion bodies were resuspended and dissolved in 6M guanidine HCl, 50mM Tris HCl, 1mM EDTA, and 0.2% NaN₃, pH 8.0. The base buffer components of the refold buffer were 50mM Tris HCl, 1mM EDTA, 0.6M arginine, 1mM cysteine/1mM cystine, and 2% (w/v) NaN₃, pH 8.0. Steep equilibrium gradient dialysis was performed from 6M to 0M in 2M increments, whereas shallow equilibrium gradient dialysis was performed from 6M to 0M in 1M increments. Each dialysis step typically occurred overnight (approximately 20 hours), except for the complete removal of guanidine HCl which was performed by two
additional dialysis steps at 4 hours each into the base buffer. Following refolding, proteins were concentrated by exposure to polyethylene glycol 20000 and filtered through 0.22µm acetate filters to remove particulate matter prior to chromatography and gel electrophoresis. Folding efficiency was determined on reduced and non reduced samples by SDS-PAGE on a 15% (w/v) acrylamide gel.

3.1.6 Chromatography
All domain mutants were purified by performing batch denaturing affinity purification on inclusion bodies with IMAC, followed by non-denaturing IMAC and gel filtration after refolding (section 3.1.5). Detailed protocols are described in section 3.1.6.1.

3.1.6.1 Immobilized Metal Affinity Chromatography (IMAC)
The presence of the C-terminal His tag of the recombinant mutants as conferred by the pET22b(+) vector, allowed for their purification from the heterogeneous mix of *E. coli* proteins. Due to the high affinity of this system the mutants could be purified using two published techniques, i.e. batch purification of the denatured proteins, or traditional chromatography of non-denatured proteins.

3.1.6.1.1 Denaturing IMAC using Ni$^{2+}$ Affinity Resin
Typically 50mg inclusion bodies (total protein) were solubilized in 100mM sodium phosphate buffer, pH 8.0, containing 8M urea (prepared freshly to prevent microamidation of amino acid side chains that can occur over time in urea buffers) and 150mM NaCl; 5mM β-mercaptoethanol was included to ensure complete reduction of disulfide bonds. This solution was stirred into 10 ml NiCAM or His-Select, equilibrated in the aforementioned buffer. The resin has a documented binding capacity of 25mg/ml. The mixture was stirred at 4°C for 2 hours, packed into an appropriate glass column, connected to an Äkta FPLC system (Pharmacia) and washed (1ml/min) to baseline absorbance at 280nm with 10 column volumes of 100mM sodium phosphate buffer, pH 8.0, containing 10mM imidazole, 150mM NaCl and 8M urea. Fractions (5 ml) were eluted at 1 ml/min using 100mM sodium phosphate buffer, pH 8.0, containing 250mM imidazole, 150mM NaCl and 8M urea. Pooled fractions were immediately dialysed against the inclusion body solubilization buffer used as the first buffer for refolding described in section 3.1.5.
3.1.6.1.2 Non-Denaturing IMAC using Ni²⁺ Affinity Resin
Samples were applied to 10ml His-Select or NiCAM columns for IMAC on an FPLC system. Typically the 10ml of resin was packed into an appropriate glass column and equilibrated with 5 column volumes of 100mM sodium phosphate buffer, pH 8.0, containing 20mM imidazole and 150mM NaCl. The sample was applied to the column and washed with equilibration buffer until baseline absorbance at 280nm (typically after approximately 10 column volumes). Elution was achieved by 100mM sodium phosphate, pH 8.0, containing 250mM imidazole and 150mM NaCl.

3.1.6.2 Gel Filtration Chromatography
All separations were performed on an FPLC system. Due to molecular weight restrictions of the columns, domains I, I-II and I-III were chromatographed initially on a Superdex75 26/60 column, whereas domains I-IV and I-V were chromatographed on the Superdex200 26/60 column. Gel filtration chromatographic steps were all performed in PBS, pH 7.4, at a flow rate of 2ml/min. Large sample volumes (13ml) were applied to the columns using a 50ml Superloop™ (Pharmacia). Five millilitre fractions collected were conservatively pooled at the peaks and concentrated by centrifugal filtration (15ml Biomax 5K centrifugal filters, Millipore). The fractions were buffer-exchanged into HBS (10mM Hepes, 150mM NaCl), pH7.4, by centrifugal filtration. Protein purity was determined by SDS-PAGE on 15% (w/v) acrylamide reducing and non-reducing gels.

3.1.7 Protein Assays
Protein concentrations were determined by absorbance at 280nm and by the BCA (bicinchoninic acid) determination method (Smith et al., 1985). Samples were stored at 4°C.

3.1.8 Dot Blot Assay for Epitope Identity
Dot blot assays were performed by immobilization of 100ng of purified recombinant mutant in TBS, pH 8.0, on PVDF blotting membrane. Spots were air-dried and the membrane was blocked overnight in TBS containing 1% (w/v) BSA and 0.05% (v/v) Tween 20. Biotinylated polyclonal goat anti-human SC was prepared according to the Microbiotinylation Kit. The biotinylated anti-human SC was diluted in fresh blocking buffer and incubated with the membrane for 1 hour at room temperature. Membranes
were washed in TBS-0.05% (v/v) Tween 20 and incubated with a 1:500 dilution of Extravidin-Peroxidase conjugate in TBS-1% (w/v) BSA. Commercial human SIgA derived from colostrum was used as a positive control. Recombinant CD89, (FcαRI, a kind donation by Katja Wenig, Max Planck Institute for Biochemistry, Martinsried, Germany) was used as a negative control. The membrane was washed as previously described and developed with SigmaFast 3, 3’ diaminobenzidine (DAB). Membranes were air dried and photographed.

3.1.9 MALDI-TOF MS Analysis of Disulfide Bonds
Purified recombinant domain mutants hSC_rDI - hSC_rDI-V were electrophoresed on 10% (w/v) acrylamide non-reducing SDS-PAGE gels and stained with Coomassie Blue. Candidate bands were excised with a sterile scalpel and sent to Prof Koji Muramoto (Tohoku University, Sendai, Japan) for MALDI-TOF MS analysis. Briefly, gel pieces were destained with 50% (v/v) acetonitrile in 25mM ammonium bicarbonate and dehydrated with 1ml acetonitrile. In gel digestion of protein was performed using 0.5µg lysylendopeptidase in 25mM ammonium bicarbonate (10µl) and incubated at 37°C for 20 hours (Prof Koji Muramoto, Pers. Comm). Digested samples were injected into a Voyager DE STR System (PE Biosystems) at an acquisition mass range of 500-15000 Da. The 66 kDa monomer of commercial BSA was used as a negative control. Analysis of digestion patterns was performed using MSight v1.0 (http://www.expasy.org/MSight).
3.2 Results and Discussion

3.2.1 Amplification and Molecular Cloning of Recombinant SC Mutants

Amplified cDNA encoding the plgR ectodomain (confirmed by DNA sequencing) was cloned and transformed into E.coli DH5α. Figure 3.1 shows the results from the cDNA amplification of the ectodomain. The cDNA was cloned into pET22b(+). The cloned cDNA was required as a template for the construction of single and multiple domain mutants spanning various regions of the ectodomain.

![Figure 3.1](image)

Figure 3.1: Agarose gel electrophoretic patterns (1% gel) of RT-PCR products showing successful amplification of plgR ectodomain upon temperature optimization. Arrow indicates band of interest. Lane: (a) control RT-PCR; (b) 50°C annealing; (c) 45°C annealing; (d) 53°C annealing (~1.8kb); (e) 60°C annealing; (f) Roche DNA Marker XVI.

Mostov et al. (1984) and Piskurich et al. (1995) divided the plgR ectodomain into five distinct domain regions based upon their knowledge of immunoglobulin structure. Table 3.2 outlines these domains using N and C terminal amino acid residue borders.

Table 3.2: Amino and carboxy terminal amino acid residues defining 5 immunoglobulin-like domains of plgR ectodomain

<table>
<thead>
<tr>
<th>Domain</th>
<th>Amino Terminal Residue</th>
<th>Carboxy Terminal Residue</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Lys&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Gly&lt;sup&gt;114&lt;/sup&gt;</td>
</tr>
<tr>
<td>II</td>
<td>Leu&lt;sup&gt;115&lt;/sup&gt;</td>
<td>Val&lt;sup&gt;219&lt;/sup&gt;</td>
</tr>
<tr>
<td>III</td>
<td>Leu&lt;sup&gt;220&lt;/sup&gt;</td>
<td>Thr&lt;sup&gt;323&lt;/sup&gt;</td>
</tr>
<tr>
<td>IV</td>
<td>Ile&lt;sup&gt;334&lt;/sup&gt;</td>
<td>Asn&lt;sup&gt;445&lt;/sup&gt;</td>
</tr>
<tr>
<td>V</td>
<td>Leu&lt;sup&gt;447&lt;/sup&gt;</td>
<td>Ser&lt;sup&gt;566&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
Initial PCR conditions were found to amplify the desired fragment lengths (see Figure 3.2), but additional optimization required. This is attributable to the template material being purified plasmid DNA. Desired bands were excised from agarose gels and blunt ligated into pCRScript SK (+) (section 3.3.1). Single colony selection upon transformation of blunt ligated fragments ensured that single copy fragments were recloned into pET 22b(+); thereby eliminating multiple copies of the amplified cDNA generated by the polymerase. The Pfu polymerase used in the PCR reactions should have reduced the probability of error being introduced due to its increased proof-reading ability. Desired primer sequences were designed to capture the distinct border regions of the immunoglobulin-like domains of SC. Figure 3.2 outlines the regions of interest to be amplified from the cloned pIgR ectodomain cDNA. The amplifications were designed in such a manner as to create a single and multiple domain mutants in order to characterize the binding of the parent molecule to IgM.
fragment lengths upon amplification are hSC_rDI ~ 344bp, hSC_rDI-II ~665bp, hSC_rDI-III ~986bp, hSC_rDI-IV ~1325bp and hSC_rDI-V ~1658bp. The amplified region encoding domain IV and V is ~672bp and the desired amplification product encoding domain V, i.e. hSC_rDV is ~333bp.

3.2.1.1 Single Domain Mutants
Figure 3.3 shows the successful amplification by PCR of the regions encoding domains I and V. Domains I and V were chosen for amplification based on the significant roles they play in polymeric immunoglobulin binding. Domain I (hSC_rDI determined size 364bp) was amplified and cloned into pET22b(+) (confirmed by automated sequencing) and transformed into E.coli BL21 (DE3). Border regions specifying the internal start and end of domain V were used in primer design and successful amplification of the domain (hSC_rDV determined size ~ 320bp). It should be noted, however, that domain V has been shown not to interact with polymeric immunoglobulins (Bakos et al., 1991). Determined fragment sizes correlated well with expected sizes, i.e. 344bp for hSC_rDI and 333bp for hSC_rDV.

3.2.1.2 Multiple Domain Mutants
Mutiple domain mutants were designed to all contain domain I at the amino terminal end, with carboxy termini determined by the C-terminal borders of domains II, III, IV and V, respectively. The amplification of multiple domain mutants encoding regions spanning
domains I-II (~667bp), I-III (~1055bp), I-IV (~1331bp) and I-V (~1777bp), as well as the successful amplification of a multiple domain mutant spanning the region encoding domain IV-V (~692bp) are shown in Figure 3.4. The amplified sizes correlated well with the predicted fragment sizes. The cloning and expression of the IV-V fragment was abandoned based on the data presented by Bakos et al. (1991) who produced a tryptic cleavage fragment of native SC containing domains IV and V, which does not possess any plg binding ability.

3.2.1.3 Sequence Analysis and in silico Plasmid Mapping.
Following successful blunt end ligation of cDNA into pCRScript SK(+), Nde I/Xho I restricted fragments were cloned into similarly restricted pET22b(+). Colonies grown on LB-AMP agar were selected by test expression (upon induction with 1mM IPTG), PCR amplification, digestion of isolated plasmids and subjected to automated sequencing at the Central Analytical Facility (University of Stellenbosch). The DNA sequences obtained were of a high integrity, confirmed by forward and reverse sequencing, and required minor editing (Figure 3.5) by visual inspection of the electropherograms in Chromas v2.23 (see Appendix A for complete electropherograms).
Figure 3.5: Section of output from automated DNA sequencing showing the region encoding the N terminal end of multiple domain mutant hSC_rDI-II. Note the Nde I restriction site (highlighted in yellow) immediately upstream of position 70.

Attenuation of signal when sequencing DNA fragments longer than 500 bases was circumvented by using multiple primer pairs. Based on the edited forward and reverse sequences, contigs (sequence fragments) were assembled using Vector Nti v9.1 to construct entire sequences. The constructed sequences were cloned \textit{in silico} into the pET22b(+) sequence in order to draw representative plasmid maps (Figure 3.6). Sequences were also subjected to multiple sequence alignment using Clustal X v1.83 to check integrity and for errors. Appendix A contains the full DNA and predicted protein sequences as well as alignments.
Figure 3.6: Representative plasmid maps of cloned recombinant domain mutants constructed \textit{in silico}. (a) pET22b(+) hSC_rDI; (b) pET22b(+) hSC_rDI-II; (c) pET22b(+) hSC_rDI-III; (d) pET22b(+) hSC_rDI-IV and (e) pET22b(+) hSC_rDI-V. Inserted fragments are shown in \textcolor{orange}{orange}. \textcolor{ForestGreen}{Dark blue} arrows are C terminal His tags conferred by plasmid. Yellow arrows indicate position of Ampicillin resistance gene (\textit{amp}^R).
3.2.2 Protein Expression and Purification of Inclusion Bodies

Typical expression of recombinant eukaryotic proteins in *E. coli* BL21 (DE3) results in the formation of inclusion bodies (Wilkinson and Harrison, 1991; Misawa and Kumagai, 1999; De Bernardez Clark, 2001). In order to test how soluble the recombinant domains would be when overexpressed, *in silico* overexpression was performed using the web applet available at biotech.ou.edu (Wilkinson and Harrison, 1991). The program generates the probability of inclusion body formation (therefore solubility) upon overexpression based on the amino acid content and the sequence of the recombinant protein. Equation 3.1 shows the calculation of a so called canonical variable (*CV*) used in the final calculation of solubility:

\[
CV = \lambda_1 \left( \frac{N + G + P + S}{n} \right) + \lambda_2 \left[ \left( \frac{(R + K) - (D + E)}{n} \right) - 0.03 \right]
\]

*Equation 3.1*

where *n* is the number of amino acids in a protein; *N*,*G*,*P* and *S* are the number of Asn, Gly, Pro, and Ser residues, respectively; *R*,*K*,*D* and *E* are the number of Arg, Lys, Asp and Glu residues, respectively, and *\( \lambda_1 \) & *\( \lambda_2 \) are constant coefficients (15.43 and -29.56, respectively). The overall probability of solubility is based on the parameter

\[
CV - CV^*
\]

*Equation 3.2*

where *CV*\(^*\), a constant, equals 1.71. Simply, if equation 3.2 is positive, the protein is predicted to be insoluble. The probability of solubility is calculated using equation 3.3:

\[
Probability\ of\ solubility = 0.4934 + 0.276 \cdot [CV - CV^*] - 0.0392 \cdot [CV - CV^*]
\]

*Equation 3.3*

This is based on the statistical evaluation of protein solubility upon overexpression in *E.coli* systems (Wilkinson and Harrison, 1991). Table 3.3 shows the results obtained from this experiment.
Table 3.3: Statistical evaluation of recombinant protein solubility upon overexpression in *E. coli* using the method of Wilkinson and Harrison (1991). *n.d.* = not determined

<table>
<thead>
<tr>
<th>Recombinant Protein</th>
<th>Insolubility (%)</th>
<th>CV-CV*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single Domain Mutants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>hSC</em> _rDI*</td>
<td>96.1</td>
<td>2.84</td>
</tr>
<tr>
<td><em>hSC</em> _rDV*</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td><strong>Multiple Domain Mutants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>hSC</em> _rDI-II*</td>
<td>94.9</td>
<td>2.64</td>
</tr>
<tr>
<td><em>hSC</em> _rDI-III*</td>
<td>87.7</td>
<td>1.90</td>
</tr>
<tr>
<td><em>hSC</em> _rDI-IV*</td>
<td>85.4</td>
<td>1.74</td>
</tr>
<tr>
<td><em>hSC</em> _rDI-V*</td>
<td>84.3</td>
<td>1.66</td>
</tr>
</tbody>
</table>

It is interesting to note the decreasing possibility of insolubility as the recombinant protein increases in length. This is probably related to a decreased ratio of specified amino acid content to the length of the peptide. The probabilities are, however, still high and, therefore, overexpression of the domain mutants would probably result in formation of inclusion bodies. It should be noted, however, that production of soluble proteins in *E. coli* systems typically result in low yields. Therefore, it was decided to overexpress the domain mutants and refold them using equilibrium gradient dialysis. This decision was also made after expression trials in the oxidizing mutant strain *E. coli* ORIGAMI (DE3) resulted in no soluble protein production (data not shown).

### 3.2.2.1 Single domain mutants (domains I and V)

From expression trials performed it was concluded that four hour expression times at 37°C resulted in high levels of recombinant protein production (Figure 3.7). The recombinant fractions were, however, largely insoluble (as predicted) as soluble fractions of cell lysates contained no detectable traces of expressed protein (not shown). It was concluded that inclusion bodies of *hSC* _rDI* and *hSC* _rDV* may have been formed. This was confirmed by buffered detergent washes, centrifugation to remove the soluble fraction and SDS-PAGE (Figure 3.8). The predicted sizes of *hSC* _rDI* and *hSC*_rDV are ~13kDa.
As previously discussed, inclusion body expression would result in sufficient material for refolding and subsequent experiments as opposed to expression of soluble proteins where yields are typically low and the native conformation may be toxic to the expressing cell. The single domain mutants were overexpressed for 4 hours and the insoluble inclusion bodies were purified by buffered detergent washes and centrifugation.
3.2.2.2 Multiple domain mutants (hSC_rD I-II, hSC_rD I-III, hSC_rD I-IV and hSC_rD I-V [rSC])

Considering the cysteine content (see Chapter 1, Table 1.1) and the potential for inter-peptide disulfide bond formation, the overexpression of multiple domain mutants as inclusion bodies upon induction was expected (and predicted, see Table 3.3). Figures 3.9-3.12 show the levels of proteins expressed in timed expression trials. Expressions were, however, not performed for longer than 4 hours owing to the potential toxic effects of the overexpressed recombinant proteins upon accumulation. Bands of expected $M_r$ values were observed as indicated in Figures 3.9-3.12: hSC_rDI-II (~25kDa), hSC_rDI-III (~37kDa), hSC_rDI-IV (~50kDa) and hSC_rDI-V (~61kDa).

![Figure 3.9](image1.png)

**Figure 3.9:** SDS-PAGE patterns (12% gel) of expression of hSC_rD I-II (predicted size ~25kDa). Lanes: (a) Sigma low molecular weight markers; (b) preinduction E. coli extract; (c) 4 hours post induction; (d) overnight (18 hours) post induction. Arrow indicates band of interest.

![Figure 3.10](image2.png)

**Figure 3.10:** SDS-PAGE patterns (15% gel) of expression of hSC_rD I-III (predicted size ~37kDa). Lanes: (a) Sigma high molecular weight markers; (b) preinduction E. coli extract; (c) 2 hours uninduced; (d) 2 hours post induction; (e) 4 hours uninduced; (f) 4 hours post induction; (g) overnight (16 hours) uninduced; (h) overnight (16 hours) post induction. Arrow indicates band of interest.
Figure 3.11: SDS-PAGE patterns (12% gel) of expression of hSC_rDI-IV (predicted size ~50kDa). Lanes: (a) Sigma high molecular weight markers; (b) and (c) preinduction E. coli extract; (d) and (e) 2 hour post induction; (f) and (g) 3 hours post induction; (h) and (i) 4 hours post induction and (j) and (k) Overnight (16 hours) post induction of two clones, respectively. Arrow indicates band of interest.

Figure 3.12: SDS-PAGE patterns (12% gel) of expression of hSC_rDI-V (predicted size ~61kDa). Lanes: (a) Sigma high molecular weight markers; (b) and (c) preinduction E. coli extract; (d) and (e) 1 hour post induction; (f) and (g) 2 hours post induction; (h) and (i) 3 hours post induction; (j) and (k) 4 hours post induction and (l) and (m) overnight (16 hours) post induction of two clones, respectively. Arrow indicates band of interest.

Partial purification of inclusion bodies (Figure 3.13) of the recombinant SC mutants was achieved by buffered detergent washing and centrifugation. Further washing of the inclusion bodies was deemed unnecessary due to the presence of the His tag (6x) which would facilitate further purification of the target protein, whether unfolded or folded (Rogi et al., 1998), although the possibility of the His tag being buried in the refolded
protein had to be considered. Hence it was decided to perform denaturing purification of inclusion bodies.

Figure 3.13: SDS-PAGE patterns (10% gel) of partially purified inclusion bodies of recombinant multiple domain mutants. Lanes: (a) hSC rDI-II; (b) hSC rDI-III; (c) hSC rDI-IV; (d) hSC rDI-V. Approximately 15 µg total protein loaded per well.

3.2.3 Refolding & Purification of Recombinant Domain Mutants
Based on the complexity of the disulfide bond arrangement present in native human fSC (similar to the ordered disulfide arrangement present in immunoglobulin variable regions [see Chapter 1]) it was decided to use the equilibrium dialysis refolding technique. The method facilitates extended refold kinetics as a direct result of the gradation of buffer changes. The use of high concentrations of L-Arg and cysteine/cystine would also facilitate increased yields of correctly folded protein with proper disulfide bond arrangement.

3.2.3.1 Single domain mutants
Nickel affinity chromatography was used to purify recombinant proteins from inclusion bodies prior to refolding to maximize correct formation of the disulfide bonds in the single domain mutants hSC_rDI and hSC_rDV. The almost one step purification of the single domain mutants was achieved using denaturing Ni$^{2+}$ affinity chromatography. Figure 3.14 shows a typical chromatogram of imidazole elution of His tagged recombinant proteins, in this case hSC_rDI. The high background is as a result of interference by imidazole. Imidazole that does not absorb at 280nm is commercially
available but in this case was not used. Figure 3.15 shows affinity purified inclusion bodies of hSC_rDI and hSC_rDV using this technique.

Figure 3.14: A typical chromatogram representing the purification of hSC_rDI using the Ni$^{2+}$ affinity resin under denaturing conditions. Blue line indicates absorbance at 280nm and the red line indicates conductivity of buffer. Blue arrow indicates target peak (5ml fractions), Red arrow indicates start of imidazole addition.

Figure 3.15: SDS-PAGE patterns (15% gel) of single domain mutant inclusion bodies purified using Ni$^{2+}$ affinity chromatography under denaturing conditions. Lanes: (a) Sigma low molecular weight markers; (b) fraction 8 hSC_rDI; (c) fraction 9 hSC_rDI; (d) fraction 10 hSC_rDI; (e) fraction 11 hSC_rDI; (f) fraction 12 hSC_rDI; (g) Sigma low molecular weight markers; (h) partially purified inclusion bodies hSC_rDV; (i) fraction 10 hSC_rDV; (j) fraction 11 hSC_rDV; (k) fraction 12 hSC_rDV. Arrows indicate bands of interest.
Figure 3.16 shows the refolded purified hSC_rDI. The result of the refolding is represented by SDS-PAGE of reduced and non-reduced fractions. The non-reduced fraction shows the presence of aggregates of unfolded material. The sample was concentrated and applied to a Superdex75 size exclusion column to remove aggregates.

Figure 3.17 shows the elution profile obtained from gel filtration chromatography of refolded hSC_rDI. Figure 3.18 shows purified hSC_rDI and hSC_rDV. The $M_r$ of pure hSC_rDI ($\sim 15$kDa) is slightly larger, this may be an electrophoretic artefact. The size of hSC_rDV ($\sim 12.8$kDa) is in good accordance with the predicted size (13kDa).

Figure 3.16: SDS-PAGE patterns (15% gel) showing the refolding of hSC_rDI. (a) Sigma low molecular weight markers; (b) 5µg reduced fraction; (c) 5µg non-reduced fraction (white arrow represents the sample). The black filled arrow shows possible aggregates.

Figure 3.17: Superdex75 gel filtration chromatography of refolded hSC_rDI. Blue arrow shows refolded hSC_rDI, red arrow shows potential aggregates and unfolded material. Peak bounded by green arrows pooled.
Protein identity of hSC_rDI and hSC_rDV was confirmed by N-terminal sequencing and alignment with human pIgR in Clustal X (Figure 3.19). The Gly:Asp mismatches in the sequence may represent sequencing errors. It should be noted, however, that the hSC_rDV clone was not subjected to DNA sequencing.

3.2.3.2 Multiple Domain Mutants
Inclusion bodies of expressed multiple domain mutants were purified by Ni$^{2+}$ affinity chromatography as described for the single domain mutants (Figure 3.20).
Figure 3.20: SDS-PAGE patterns (15% gel) of elution of multiple domain mutants from Ni$^{2+}$ affinity chromatography. Gel 1 (hSC_rDI-II); lanes: (a) partially purified inclusion bodies; (b) fraction 10; (c) fraction 11; (d) fraction 12. Gel 2 (hSC_rDI-III); lanes: (a) partially purified inclusion bodies; (b) fraction 10; (c) fraction 11; (d) fraction 12. Gel 3 (hSC_rDI-IV); lanes: (a) partially purified inclusion bodies; (b) fraction 10; (c) fraction 11; (d) fraction 12. Gel 4 (hSC_rDI-V); lanes: (a) partially purified inclusion bodies; (b) fraction 10; (c) fraction 11; (d) fraction 12. Arrows indicate bands of interest.

The refolding and purification of hSC_rDI-V (rSC) will be described here in more detail (all multiple domains were refolded and purified in the same way). This is the first time that recombinant free SC was produced in a bacterial system and refolded in vitro (Prinsloo et al., 2006; Appendix C). Affinity purified inclusion bodies (Figure 3.20, gel 4) were completely solubilized and reduced in 6M guanidine HCl, 50mM Tris HCl, 1mM EDTA, pH 8.0, containing 100mM β-mercaptoethanol. The base buffer components of the refold buffer were 50mM Tris HCl, 1mM EDTA, 0.6M Arg, 1mM cysteine/1mM cystine, 2% (w/v) NaN$_3$, pH 8.0 (as described in section 3.1.5). Gradual equilibrium dialysis was employed to facilitate extended refolding kinetics and to permit formation of the complex disulfide bond arrangement (this holds true for all other multiple domain mutants). After the final buffer change the dialysis mix was gradually concentrated by exposure to PEG 20000. Gel filtration chromatography using a Superdex200 column
(Figure 3.21) facilitated separation from aggregates. Pooled fractions were typically applied to the column for rechromatography. Figure 3.22 shows SDS-PAGE patterns of the pure sample obtained from a second round of chromatography (Figure 3.21, red line).

![Figure 3.21: Elution profile of hSC_rDI-V(rSC) separation on a Superdex200 column. Blue line indicates the absorbance at 280nm of concentrated refolded inclusion bodies and the red line (secondary Y axis) indicates the absorbance at 280nm of concentrated pooled fraction from the initial chromatography (i.e. the rechromatographed sample).](image)

<table>
<thead>
<tr>
<th>M_r (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
</tr>
<tr>
<td>72</td>
</tr>
<tr>
<td>55</td>
</tr>
<tr>
<td>40</td>
</tr>
<tr>
<td>33</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>17</td>
</tr>
<tr>
<td>11</td>
</tr>
</tbody>
</table>

![Figure 3.22: SDS–PAGE patterns (15% gel) showing purified inclusion bodies and hSC_rDI-V (rSC). Lanes: (a) and (d) broad range molecular weight markers; (b) purified inclusion bodies (prior to refold); (c) and (e), pure hSC_rDI-V (reduced ~58kDa) and non-reduced ~54kDa, respectively.](image)
The deviations of the $M_r$ for the reduced (~58kDa) and non-reduced samples (~54kDa) of hSC_rDI-V may be electrophoretic artefacts; alternatively, it could signify a “refolded conformation” by possibly inferring a tightly packed conformation. It should be noted that electrophoresis with a smaller gel pore size (7.5%) would allow for better resolution of recombinant proteins and markers.

Figure 3.23 shows the purified multiple domain mutants. The multiple domain mutants hSC_rDI-II, hSC_rDI-III and hSC_rDI-IV were purified in the same manner as that described for hSC_rDI-V.

The $M_r$ values of hSC_rDI-II (~27kDa), hSC_rDI-III (~36kDa), hSC_rDI-IV (~51kDa) and hSC_rDI-V (~58kDa) are all in good accordance with the predicted sizes.
Figure 3.24 shows the N terminal sequence alignments with pIgR. N terminal results obtained contained numerous peak overlaps. This was probably due to N-terminal processing common in *E. coli*. Recombinant proteins do, however, show identity with human pIgR. The sequenced cDNA of the multiple domain mutants (Appendix A), however, align perfectly with human pIgR ectodomain. Therefore, the errors are likely related directly to N terminal processing.

Table 3.4 shows a purification table with the protein yields of all recombinant domain mutants obtained after Ni$^{2+}$ IMAC and gel filtration. The isolated yields of inclusion bodies from the Ni$^{2+}$ IMAC are quite good. This represents the inclusion body protein contents from 4L preparative expressions in *E. coli* BL21 (DE3) 4 hours post induction. These inclusion bodies were used in 10-15mg protein batches for refolding, using 100ml protein solutions per refold (i.e. 100-150µg protein/ml). Table 3.4 shows the exact protein amount used for each domain per refold and the final protein yields obtained per refold after gel filtration chromatography. Although the protein concentration per refold was low, decreasing the concentrations may result in elevated yields of refolded protein. Low unfolded protein concentration may prevent possible protein-protein interactions between refold intermediates. An alternative to lowering protein concentration may be to increase the concentration of L-Arg in the refold or to use other small molecule artificial chaperones like β-cyclodextrin (Daugherty *et al.*, 1998). The cost:yield ratio should, however, be considered. The yields obtained are probably a direct result of the complex structures of the recombinant domain mutants, i.e. the disulfide arrangements into “large” and “small” disulfide bridges (Eiffert *et al.*, 1984). Further optimization of the oxido-

---

**Figure 3.24: N-terminal sequence alignment of multiple domain mutants (hSC_rDI-II – hSC_rDI-V) with pIgR. ***, consensus; •, mismatches and –, gaps.**

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;hSC_rDI-II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;hSC_rDI-III</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;hSC_rDI-IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;hSC_rDI-V</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;pIgR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Clustal Consensus</strong></td>
<td>*</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

| >hSC_rDI-II          |    |    |    |    | -MS- | PGEVEVNS- |
| >hSC_rDI-III         |    |    |    | -MS- | PGEVEVNS- |    |
| >hSC_rDI-IV          |    |    |    |    | -PGEV- | VN- |
| >hSC_rDI-V           |    |    |    |    | -MS- | PGEVEVNS- |
| >pIgR                |  MLLFVLICLLAVFAISTKSPFGPVEEVSVEGSVSIYFYTPTSVRHRTKYGCRQGA |
shuffling reagent couple (cysteine/cystine) may improve yields by careful control of the oxidizing conditions. The use of GSH/GSSG was investigated but resulted in no significant yields (data not shown).

Table 3.4: Purification table of recombinant domain mutants showing total protein content (mg) and percentage yield from a 100ml refold.

<table>
<thead>
<tr>
<th></th>
<th>Total Protein (mg)</th>
<th>Percentage Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single Domain Mutants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hSC_rDI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inclusion bodies (after IMAC)</td>
<td>118</td>
<td>−</td>
</tr>
<tr>
<td>Total protein per100ml refold</td>
<td>15</td>
<td>100</td>
</tr>
<tr>
<td>Pure protein from 100ml refold*</td>
<td>3.3</td>
<td>22</td>
</tr>
<tr>
<td>hSC_rDV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inclusion bodies (after IMAC)</td>
<td>125</td>
<td>−</td>
</tr>
<tr>
<td>Total protein per100ml refold</td>
<td>15</td>
<td>100</td>
</tr>
<tr>
<td>Pure protein from 100ml refold*</td>
<td>2.6</td>
<td>17.3</td>
</tr>
<tr>
<td><strong>Multiple Domain Mutants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inclusion bodies (after IMAC)</td>
<td>98</td>
<td>−</td>
</tr>
<tr>
<td>Total protein per100ml refold</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Pure protein from 100ml refold*</td>
<td>0.53</td>
<td>5.3</td>
</tr>
<tr>
<td>hSC_rDI-III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inclusion bodies (after IMAC)</td>
<td>105</td>
<td>−</td>
</tr>
<tr>
<td>Total protein per100ml refold</td>
<td>15</td>
<td>100</td>
</tr>
<tr>
<td>Pure protein from 100ml refold*</td>
<td>1.1</td>
<td>7.3</td>
</tr>
<tr>
<td>hSC_rDI-IV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inclusion bodies (after IMAC)</td>
<td>89</td>
<td>−</td>
</tr>
<tr>
<td>Total protein per100ml refold</td>
<td>12</td>
<td>100</td>
</tr>
<tr>
<td>Pure protein from 100ml refold*</td>
<td>0.48</td>
<td>4</td>
</tr>
<tr>
<td>hSC_rDI-V (rSC)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inclusion bodies (after IMAC)</td>
<td>110</td>
<td>−</td>
</tr>
<tr>
<td>Total protein per100ml refold</td>
<td>11</td>
<td>100</td>
</tr>
<tr>
<td>Pure protein from 100ml refold*</td>
<td>0.6</td>
<td>5.5</td>
</tr>
</tbody>
</table>

* after gel filtration.
3.2.4 Dot Blot Association Assay: Establishment of “Identity” of recombinant domain mutants to Native human SC

Structural identity between the recombinant domain mutants and native human SC was established by dot blot assays. The polyclonal antibody, anti-SC (bound and free), was raised in a goat model against native free SC and dIgA bound SC. Human SIgA isolated from colostrum was used as a positive control. Figure 3.25 shows that all the unglycosylated recombinant domain mutants were recognized as containing structural features akin to that of native human SC. The antibody did not recognize the negative control, recombinant CD89, the Fc receptor for IgA. Although not a conclusive proof of correct protein folding, the recombinant mutants revealed structural similarity to native human SC.

Figure 3.25: Dot blot analysis showing identity of recombinant domain mutants with native human SC. (a) SIgA positive control; (b) recombinant CD89 negative control; (c) hSC_rDI; (d) hSC_rDI-II; (e) hSC_rDI-III; (f) hSC_rDI-IV; (g) hSC_rDI-V (rSC).

3.2.5 MALDI-TOF MS Analysis of Disulfide Bonds

In silico digestion of the recombinant domain mutants with lysylendopeptidase shows specified cleavage points at the C terminal side of the 28 Lys residues (Figure 3.26). Theoretically, digestion of the disulfide bonded protein would yield fewer bands than reduced protein and would therefore reveal characteristic MS bands of specified sizes, depending on the disulfide bond arrangement, and a characteristic pattern allowing for an inference of disulfide bond arrangement.

Overlay analysis (against BSA negative control; Figure 3.27) of the digestion patterns showed multiple bands for hSC_rDI; more than the expected 3 peaks ($M_w \sim 2623, 3256$ and 7237 Da, respectively), if disulfide bonds were present. In fact, only 4 peaks are expected for the reduced protein. Peak sizes observed upon overlay of the digestion pattern MS spectra of hSC_rDI with that of BSA (Figure 3.27) showed unique peaks with m/z values of 7260, 3550, 3360, 2388, 2253, 1923, 805 and 673. Although there appears
to be some correlation with the expected m/z values, the smaller peaks observed do not allow for analysis of disulfide arrangement.

Figure 3.26: *In silico* digestion of hSC_rDI-V with lysylendopeptidase. Red blocks represent Lys residues, i.e. cleavage sites. Green blocks represent the 20 cysteines involved in internal disulfide bonds. Black lines indicate disulfide bonds as outlined by Eiffert *et al.* (1984) (Table 1.1).

This trend was observed with all the recombinant domains. This method, therefore, did not provide enough information regarding the disulfide bond arrangement of the recombinant domains. A better approach would probably entail reduced and non-reduced fragment analysis and correlation to purified native human SC. Appendix B contains digestion spectra of all recombinant domain mutants. Also multiple enzyme and chemical cleavages of recombinant proteins and sequencing of resulting fragments would allow for precise analysis of the disulfide arrangement.

Figure 3.27: Overlay of MALDI-TOF MS spectra of lysylendopeptidase digested hSC_rDI and BSA.
CHAPTER 4

Binding Activity Studies

Receptor-ligand interactions form the cornerstone of a healthy immune system, the study of which allows the modern immunologist insight into understanding the immune system and also aids in the development of novel therapies. As discussed in Chapter 1, the polymeric immunoglobulin receptor possesses the ability to bind polymeric immunoglobulins as well as other receptors. Receptor-ligand interactions can be analysed using traditional immunological (dot blot, receptor-ligand ELISA) and cell based binding assays (whole cell ELISA). Surface plasmon resonance allows for a uniquely sensitive method for describing these interactions and characterization of new ligands (especially in drug discovery).

According to Homola et al. (1999) surface plasmon resonance (SPR) “is a charge-density oscillation that may exist at the interface of two media with dielectric constants of opposite signs, for instance, a metal and a dielectric.” SPR biosensors exploit this by measuring changes in the refractive index at a sensor surface, i.e. excitation of surface plasmon waves at a metal liquid interface (McDonnell, 2001). This principle is illustrated by Figure 4.1.

![Diagram of surface plasmon resonance](image)

Figure 4.1: Excitation of surface plasmons at a sensor surface by polarized light allows for measurement of the changes in the angle of reflected light upon interaction events. SPR results in reduced reflected light intensity which is related to the angle of reflection and the wavelength (from BIAcore, 1998).
The most important feature of SPR measurements is that it allows for the study of specific biomolecular interactions without the use of labelled molecules (Leidberg et al., 1995). SPR biosensors allow for both qualitative and quantitative analysis of binding events. The association and dissociation rates can typically be derived by model fits (BIAcore, 1998). These data are often in accordance with that collected by other means.

The ability of the recombinant domain mutants produced in Chapter 3 to bind IgM was investigated using various immunological, cell based and biophysical techniques. It was hoped that the mutants would allow for an investigation into the binding characteristics of individual domains I, II, III, IV and V.
4.1 Materials and Methods

4.1.1 Reagents and Equipment
PVDF membrane and BSA from Roche (Mannheim, Germany); Dulbecco’s Modified Eagles Medium (DMEM), recombinant TNFα, tissue protease inhibitor cocktail, Tween 20, Micro-Biotinylation Kit, polyclonal anti-human secretory component, human IgG, human IgM, SIgA, anti-human Secretory Component (bound and free), extravidin-peroxidase conjugate, SigmaFast™ 3, 3’-diaminobenzidine (DAB), 3,3’,5,5’ tetramethylbenzidine (TMB) liquid substrate supersensitive and TMB stop reagent were from Sigma Aldrich (St. Louis, Missouri, USA). Amicon Ultra-15 5k centrifugal filters from Millipore (Cork, Ireland); CM5 sensor chip, HBS-E (10 mM HEPES, pH 7.4, 150 mM NaCl, 3.4 mM EDTA, 0.005% (v/v) Surfactant P20), ethanolamine, 1-ethyl-3(3-dimethylaminopropyl)-carbodiimide (EDC) and N-hydroxysuccinimide (NHS) from Biacore® AB (Uppsala, Sweden). HT29 intestinal carcinoma from American Type Culture Collection (USA); 96 well MaxiSorp, 96 well Nuncolon and 10cm cell culture dishes from Nunc (Roskilde, Denmark); foetal calf serum from Highveld Biological (Johannesburg, South Africa).

4.1.2 Dot Blot Reassociation Assay for IgM Binding
Dot blot reassociation assays for IgM binding were performed in triplicate as described in Chapter 3 (section 3.1.8). IgM was used in place of anti-human SC. The commercial human SIgA was used as a positive control due to the presence of free SC in the sample, shown by western blot. Briefly, non-reduced samples of SIgA (1µg) were resolved by SDS-PAGE on a 10% acrylamide gel and transferred to PVDF membrane by semi-dry-blotting. Blots were detected as described in section 3.1.8. Recombinant CD89, FcαRI, was used as a negative binding control.

4.1.3 Receptor-Ligand Binding Assay (Direct ELISA)
Recombinant mutants (hSC_rDI, hSC_rDI-II, hSC_rDI-III, hSC_rDI-IV and hSC_rDI-V[rSC]) were coated overnight at 4°C on 96 well MaxiSorp plates at a concentration of 1µg/ml in 100µl PBS pH 7.4. Plates were washed at least three times with PBS-T (PBS, pH 7.4, 0.05% (v/v) Tween 20). Excess binding sites were blocked with 200µl PBS-T
containing 5% (w/v) BSA. Plates were washed as described and incubated with 100 µl of biotinylated IgM (78-1250 nM) at 37°C for 4 hours and at 4°C overnight. Plates were washed with PBS-T and incubated with 200 µl extravidin-peroxidase conjugate at a dilution of 1:1000 in PBS-T containing 1% BSA at room temperature for two hours. Plates were washed with PBS-T and bound extravidin-peroxidase was detected with 100 µl TMB liquid substrate and the reaction was stopped with 50 µl TMB stop reagent after 5 min. The colour change was measured at 450 nm. Microwells coated with recombinant CD89 were used to determine non-specific binding activity. Data were analysed with Prism 4 (Graphpad software) using non-linear regression.

4.1.4 Biacore™ Surface Plasmon Resonance Measurements

4.1.4.1 Amine Coupling Procedure: Attaching a Ligand to Biacore Chip Surface

The procedure for amine coupling of ligands to Biacore CM5 chips will be described here. The subsequent sections will deal with experimental procedures performed with the recombinant domain mutants. The CM5 sensor chip was docked into the Biacore X surface plasmon resonance biosensor at 25°C and the system was primed (equilibrated) with HBS-E, pH 7.4, using the system procedure as outlined by the control software BIAcontrol. Proteins were diluted in various acetate buffers with pHs ranging from 4.5-7. This was for preconcentration; simply, preconcentration tests the optimal buffer pH in order to maximize the amount of sample brought into contact with the chip surface. Typically, the optimal pH is 1 pH unit below the protein pI.

Proteins (recombinant domain mutants and IgM) were concentrated to approximately 1 mg/ml by centrifugal filtration and buffer exchanged into HBS-E (pH 7.4) prior to use on the Biacore system in order to minimize differences in refractive index (Karlsson and Fält, 1997). Upon determination of the optimum conditions for immobilization the chip surface of the first channel (Fc1) was activated at a flow rate of 10 µl/min using a 1:1 mixture of 0.4 M EDC and 0.1 M NHS. The ligand (the molecule to be bound) was diluted in the appropriate coupling buffer (determined by preconcentration) and injected to allow for amine coupling and finally ethanolamine was used to deactivate excess reactive groups (Biacore, 1998). The chip was pulsed with 5 mM glycine to prepare the chip and
counteract matrix effects, e.g. swelling of CM5 matrix. Buffer was allowed to flow over the surface of the chip at 10µl/min to stabilize the baseline.

4.1.4.2 Receptor Coupling: Recombinant Domain Mutants as “Ligands”
Recombinant domain mutants hSC_rDI, hSC_rDV, hSC_rDI-II and hSC_rDI-III were coupled to Fc1 using the procedure described in 4.1.4.1. Fc2 was activated (EDC/NHS) and HBS-E buffer was injected over the flow cell as a mock coupling. Active sites were quenched with ethanolamine and the channel was used as a negative subtraction control. IgM was injected at a concentration range of 10-1000nM. Chip surfaces were regenerated with 1mM NaOH.

4.1.4.3 IgM Coupling : Recombinant Domain Mutants and rSC as “Analytes”
IgM (50nM) was coupled to Fc1 at 13413RU. Chemically treated IgM (1M NaOH, overnight at 37°C) was coupled to Fc2 at 1700RU (~1nM IgM) and used as a negative subtraction control. Although binding did not occur when untreated IgM was coupled at low density, the chemical treatment was deemed necessary to ensure elimination of non-specific interactions. The domain deletion mutants were run as analytes at concentrations ranging from 1.25 -1000nM. Sensorgrams were run at 20, 30 and 50µl/min in triplicate. Chip surfaces were regenerated with 1.5M MgCl₂.

4.1.4.4 SPR Data Analysis
Sensorgrams were edited and analysed in BIAevaluation 3.1. Kinetic models were applied and fit in BIAevaluation and ClampXP (available for download at http://www.hci.utah.edu/groups/interaction).

4.1.5 Competitive Binding Studies: CELISA (Bound Whole Cell ELISA)
Various groups have shown that pIgR upregulation occurs in the presence of proinflammatory cytokines (Ackermann et al., 1999; Blanch et al., 1999; Loman et al., 1999, Takenouchi-Ohkubo et al., 2004; Bruno and Kaetzel, 2005). The mechanisms behind the expression were discussed in Chapter 1. The human intestinal carcinoma, HT29, constitutively expresses pIgR molecules (ATCC). Recently it has been shown that constitutive levels in HT29 can be increased above baseline using TNFα and that long-term exposure can result in continued increased expression (Bruno and Kaetzel, 2005).
Based on this, HT29 cells were induced with TNFα and the question of whether the recombinant domain mutants (hSC_rDI, hSC_rDI-III and hSC_rDI-V) can compete with the native polymeric immunoglobulin receptor in a biological system was investigated by competitive binding analysis using whole-cell ELISA (CELISA).

### 4.1.5.1 Cell Culture and Maintenance

The human intestinal carcinoma, HT29, was cultured and maintained in Dulbecco’s Modified Eagle’s medium (DMEM) (Dulbecco and Freeman, 1959) containing 1.0g/l D(+) glucose, 0.585g/l L-glutamine and 0.11g/l sodium pyruvate supplemented with 10% (v/v) foetal calf serum (FCS). Cells were maintained in a humidified incubator under 5% CO₂ at 37°C in 10cm culture dishes. Cells were subcultured at approximately 80% confluency (monitored microscopically) by trypsinization (1% [w/v] trypsin in PBS-E [containing 1mM EDTA]) of PBS-E washed cells. Cells were incubated for 5 minutes at 37°C to ensure complete trypsinization. Cells were resuspended in 10ml DMEM/10% FCS and plated (after 1 in 10 dilution). Viable cell numbers were determined by haemocytometer counts of unstained cells after trypan blue staining.

### 4.1.5.2 Immunohistochemical probe for pIgR

Cells were plated at a density of approximately 10⁵ cells per ml in 10cm culture dishes containing sterile autoclaved glass microscope coverslips in 10ml DMEM/10% FCS. Growth was monitored microscopically for 48 hours. The medium was aspirated and coverslips were washed 3x with ice cold PBS, pH 7.4. Cells were fixed with 5ml ice cold 4% (w/v) formalin in sterile ddH₂O at 4°C for 30 minutes. Endogenous peroxidases were quenched by incubation with 100µl 10% (v/v) H₂O₂ (in sterile ddH₂O) at room temperature for 30 minutes. Fixed cells were washed (3x) with ice cold sterile PBS and excess binding sites were blocked by incubation with 10ml 1% (w/v) BSA in PBS-0.05% (v/v) Tween 20 (PBS-T) overnight at 4°C. Coverslips were washed 3 times with PBS-T and incubated with biotinylated anti-human SC (1:10000) overnight at 4°C. Coverslips were washed with PBS-T at least 3 times and incubated with 1:1000 dilution of extravidin-peroxidase conjugate for 1 hour at room temperature. Coverslips were washed as previously described and bound extravidin-peroxidase was detected with SigmaFast DAB. Staining was allowed to proceed for maximum of 5 minutes and quenched with
PBS. The cells were dehydrated by transfer through an ethanol series (2x 70% [5 minutes], 2x 95% [5 minutes] and 2x 100% [5 minutes]) and air dried. Slides were fixed in DPX and observed and photographed under an Olympus BX-60 microscope.

4.1.5.3 Induction of pIgR Expression
Expression of pIgR above constitutive levels was induced by incubation of cells in a final concentration of 10ng/ml human recombinant TNFα (Takenouchi-Ohkubo et al., 2004; Bruno and Kaetzel, 2005). Freshly subcultured cells were allowed to grow for 24 hours prior to TNFα exposure. The medium was aspirated and replaced with fresh filter sterilized (0.22µm acetate filter) DMEM containing 10% FCS and 10ng/ml TNFα. Cells were incubated for approximately 72 hours, prior to cell lysis and protein extraction or bound whole cell ELISA (CELISA). Uninduced cells were maintained for negative control purposes.

4.1.5.4 Western Blot for Determination of pIgR Expression
Whole cell lysis and protein extraction were performed by aspiration of media and washing attached cells at least 3 times with ice-cold sterile PBS prior to 5 minute incubation with 1 ml ice-cold lysis buffer (50 mM Tris-HCl, pH 7.4, 1% (v/v) Non-Idet P-40, 0.25% (w/v) sodium deoxycholate, 150 mM NaCl, 1 mM EDTA). Tissue protease inhibitor cocktail was added to lysis buffer immediately prior to use at a 1:200 dilution. Attached cells were scraped using a teflon cell scraper and the suspension was centrifuged at 14000g (4°C) for 15 minutes in a cooled microfuge (Eppendorf). The supernatant (~1ml) was transferred to a fresh sterile 1.5ml Eppendorf microfuge tube. Total protein concentrations were determined using the BCA protein determination method described in section 3.1.7. Isolated protein solutions were stored at -20°C until used. Western blots were probed for pIgR expression (using 5µg total protein from cell lysates) as previously described (section 4.1.2) using biotinylated anti-human SC.

4.1.5.5 Competitive Binding Studies: CELISA
Trypsinized cells were plated at 10⁴ cells per ml in 96 well Nunclon microwell plates, in a volume of 200µl DMEM/10% FCS. Cultures were allowed to grow for 24 hours prior to induction (section 4.1.5.3). Cells were monitored for 3 days (72 hours) microscopically to assess 100% confluence. The medium was aspirated and wells were washed at least 3
times with 200µl ice cold PBS (pH 7.4) prior to fixation in 100µl 4% formalin (in sterile ddH₂O) for 30 minutes. Endogenous peroxidases were quenched by incubation with 100µl 10% H₂O₂ (in sterile ddH₂O) at room temperature for 30 minutes. Fixed cells were washed (3x) with ice cold sterile PBS and excess binding sites were blocked by incubation with 200µl 5% BSA in PBS-0.05% Tween 20 (PBS-T) overnight at 4°C. Wells were washed 3 times with PBS-T. Competitive binding was assessed by co-incubation (total volume 100µl) of increasing competitor (in this case the hSC_rD mutants) with 625nM biotinylated IgM. The concentration ranges for the hSC_rDI, hSC_rDI-III and hSC_rDI-V were 37nM-19µM, 18nM-9.2µM and 3.32nM-1.70µM, respectively. All proteins were buffer exchanged into PBS prior to use by centrifugal filtration using 5K MWCO Amicon filters. Incubations were performed for 4 hours at 37°C. Plates were washed 3 times with PBS-T and bound IgM was probed with 100µl extravidin peroxidase (1:1000) in PBS-T containing 1% BSA at room temperature for 1 hour. Plates were washed at least 3 times with PBS-T prior to detection with 100µl TMB liquid substrate, supersensitive. Detection solution was incubated for 5 minutes at room temperature and the colour change was stabilized by addition of 50µl TMB stop reagent. Absorbance values were determined at 450nm in a microwell plate reader (Metrohm). Total binding was determined by incubation of 625nM IgM (100% IgM) with no competitor and background binding was determined by incubation with PBS-T containing 1% BSA. Non-specific binding was monitored by co-incubation of IgM with recombinant CD89. Experiments were performed in triplicate. Binding was monitored in uninduced cells for comparison.

4.1.5.6 Data Analysis
Data were transformed to percentage inhibition of IgM binding and plotted against log transformed hSC_rD concentration. Non-linear regression curve fitting was performed in Prism 4 (Graphpad software) using the one site competition binding model defined by equation 4.1.
$Y = \text{Minimum} + \left[ \frac{(\text{Maximum} - \text{Minimum})}{1 + 10^{(X - \log IC_{50})}} \right]$  

\textbf{Equation 4.1}

where, \textit{Maximum} and \textit{Minimum} define the maximum and minimum specific binding, respectively. \textit{X} is the competitor concentration and \textit{Y} the percentage inhibition. \textit{IC}_{50} is the inhibitory concentration 50%, i.e. the competitor concentration representing the halfway point between the \textit{Maximum} and \textit{Minimum} points.
4.2 Results and Discussion

4.2.1 Dot Blot Reassociation Assay (DORA) Analysis

The ability of the recombinant domains, hSC_rDI – hSCrDI-V, to bind IgM was determined by dot blot reassociation assay (DORA). Secretory IgA (commercial) was used as a positive binding control after it was determined to contain a fraction of free SC (fSC) by western blot using polyclonal goat anti-human SC (Figure 4.2d). Free SC is often found loosely associated with secretory antibodies in secretions. A positive binding event to fSC was observed using commercial IgM (Figure 4.2a). The single domain mutant, hSC_rDI, showed a positive binding event (Figure 4.2c), while the recombinant FcαRI (CD89) receptor did not bind IgM, as expected (Figure 4.2b). It was therefore primarily concluded that the domain possessed IgM binding activity. DORA was not performed with hSC_rDV.

![Figure 4.2: Dot blot reassociation assay for IgM binding activity of hSC_rDI. (a) Free SC positive control; (b) recombinant CD89 negative control; (c) hSC_rDI and (d) western blot probed for presence of fSC (lane 2) in commercial SIgA (lane 2), lane 1 MWM from the bottom up 75, 100, 130 and 170 kDa.]

Figure 4.3 (a-f) shows the results from a typical DORA using the multiple domain mutants. All mutants exhibit positive binding events although to a different degree of intensity. It should be noted that the technique employed here is largely qualitative and not quantitative. More quantitative binding experiments are required to adequately describe binding of the multiple domain mutants to IgM.
4.2.2 Receptor Ligand Binding Assay: Direct ELISA

A receptor ligand binding assay (direct ELISA) was developed to determine the equilibrium dissociation constant ($K_D$) of hSC_rDI; as well as the multiple domain mutants (hSC_rDI-II – hSC_rDI-V). This was performed to primarily determine the ability of the mutants constructed in this study to bind IgM and whether the binding observed was comparable to that of the recombinant domain I by Bakos et al. (1994) and constants shown in literature for native SC and pIgR (Bakos et al., 1991). The recombinant proteins were immobilised on the inert polystyrene surface of the Nunc Maxisorp microwell plates (as outlined in section 4.1.3). Biotinylated IgM was added at increasing concentrations to saturation (approximately 0.1 to 10$K_D$) and incubated overnight at 4°C to ensure equilibrium was reached. It was established that CD89 did not bind IgM (see dot blot reassociation assay, section 4.2.1); therefore, it was immobilised to check for non specific binding events. Data sets were analysed using non-linear regression. Traditional Scatchard-Rosenthal plots were not used to determine equilibrium dissociation constants because linearization often skews standard deviation (experimental error) resulting in under or overestimation of $K_D$ (Motulsky and Christopoulos, 2003). Figure 4.4 shows the non-specific binding of CD89 to IgM. Consider the following assumption denoted by the relationship:

\[
\text{Total} = \text{Specific} + \text{Nonspecific}
\]

Equation 4.2
Based on equation 4.2 non-specific binding events were subtracted from total binding to yield specific binding of the recombinant domain deletion mutants. Specific binding data were fit using equation 4.3 based on the law of mass action,

\[ \text{Specific} = \frac{B_{\text{max}} \cdot X}{K_D + X} \]  

Equation 4.3

where the ligand is \( X \). \( B_{\text{max}} \) is the maximal binding level at saturation and \( K_D \) is the equilibrium dissociation constant.

![Graph of nonspecific binding of CD89 to IgM](image)

Figure 4.4: Nonspecific binding of CD89 to IgM. Error bars represent standard deviation (n=3). Data were subtracted from total binding data as determined by equation 4.2.

The \( K_D \) of hSC_rDI for IgM obtained from Figure 4.5 was 199.1 ± 57nM (Table 4.1) and compares favourably with that obtained by Bakos et al. (1994) for recombinant domain I using competition ELISA (510nM). Therefore hSC_rDI possesses physiological binding capacity for IgM.

![Graph of non-linear fit for determination of \( K_D \) of hSC_rDI for pentameric IgM using ELISA](image)

Figure 4.5: Non-linear fit for determination of \( K_D \) of hSC_rDI for pentameric IgM using ELISA. Error bars represent standard deviation (n=3) (\( R^2 = 0.8343 \)).
Figures 4.6-4.9 show the non-linear regression fits to data sets representing the specific binding of multiple domain mutants to IgM. Data were fit to specific binding curves, after subtraction of non-specific binding data (according to the assumption represented by equation 4.2), using equation 4.3. Table 4.1 shows the equilibrium dissociation constants from the curve fitting analyses.

![Graph showing specific binding](image1)

Figure 4.6: Non-linear fit for determination of $K_d$ of hSC_rDI-II for pentameric IgM using ELISA. Error bars represent standard deviation (n=3) ($R^2=0.8395$).

![Graph showing specific binding](image2)

Figure 4.7: Non-linear fit for determination of $K_d$ of hSC_rDI-III for pentameric IgM using ELISA. Error bars represent standard deviation (n=3) ($R^2=0.7007$).
Figure 4.8: Non-linear fit for determination of $K_D$ of hSC_rDI-IV for pentameric IgM using ELISA. Error bars represent standard deviation (n=3) ($R^2=0.8447$).

Table 4.1: Equilibrium dissociation constants as determined by non-linear regression for the recombinant domain deletion mutants as obtained from receptor-ligand binding assay.

<table>
<thead>
<tr>
<th></th>
<th>$K_D$ ($\pm$ SD) ($\times 10^8$)(M)</th>
<th>95% confidence ($\times 10^8$)(M)</th>
<th>$K_D$ ($\times 10^8$)(M) (Bakos et al., 1994)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hSC_rDI</td>
<td>19.91 (5.710)</td>
<td>0.000 – 50.10</td>
<td>51.0</td>
</tr>
<tr>
<td>hSC_rDI-II</td>
<td>3.130 (0.366)</td>
<td>2.160 – 4.150</td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-III</td>
<td>5.726 (2.052)</td>
<td>1.376 – 10.08</td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-IV</td>
<td>3.009 (1.006)</td>
<td>0.000 – 6.401</td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-V</td>
<td>3.615 (0.950)</td>
<td>0.976 – 6.253</td>
<td>0.64</td>
</tr>
</tbody>
</table>
It is clear from Table 4.1 that domains II to V of the recombinant domain mutants increase the binding affinity compared to the single domain mutant, hSC_rDI. There seems, however, to be no further contribution towards affinity as additional domains are added. When considering that IgM does not form any covalent bonds with pIgR it may be reasonable to assume that the domains downstream from domain I interact with IgM specifically but in a manner to stabilize the complex. Binding studies conducted by Norderhaug et al. (1999), using domain deletion mutants where simultaneous deletions of either domains II & III or IV & V were used, resulted in no loss of binding to pentameric IgM. The data presented here confirm and strengthen the findings of Norderhaug and coworkers.

Although sensitive, the ELISA method described here does not allow for the description of the individual binding components of the equilibrium dissociation constant. A more sensitive real time technique was used to examine this aspect.

4.2.3 Surface Plasmon Resonance Analysis

SPR analysis allows for more sensitive measurements of the binding kinetics of receptor ligand interactions without molecular labels. The following sections will describe the binding kinetics observed between single and multiple domain mutants and IgM.

4.2.3.1 Recombinant Domain Mutants as Ligands

Real time binding analysis of the single domain mutants, hSC_rDI and hSC_rDV, were performed by immobilization of the recombinant domains on CM5 biosensor chips. Figure 4.10 shows the IgM binding ability of hSC_rDI and hSC_rDV. From Figure 4.10 it is clear that hSC_rDV failed to bind IgM. This clearly illustrates the requirement of domain I for polymeric immunoglobulin binding, as was shown independently by Frutiger et al. (1988) and Bakos et al. (1991). The data shown here corroborate that of Bakos et al. (1991) that single domain regions (in their case tryptic digest fragments) that do not contain domain I do not bind pentameric IgM (or dIgA).
Data sets were collected for hSC_rDI as ligand at 3 flow rates to check for mass transfer limitation/effects. Typically it is suggested that the smaller molecule be immobilised and the larger molecule be used as the analyte (David Pugh, Pers. Comm.) The ligand hSC_rDI was immobilised at 275 RU to further minimize mass transfer limitation (BIAcore, 1998). Figure 4.11 shows the effects of different flow rates on 100nM IgM binding to hSC_rDI. As can be seen there is attenuation in signal as the large IgM molecule flows over the surface. Mass transfer limitations are typically determined during data fitting; the association ($k_a$) and dissociation rate ($k_d$) constants should be relatively similar at different flow rates (Marquart, 2005). This is clearly not the case as seen in Table 4.2. The values in Table 4.2 were obtained by fitting single curves obtained at different flow rates with a constant analyte concentration (100nM IgM). Data were fit using the Langmuir 1:1 binding model as outlined by equation 4.4:

$$A_{analyte} + B_{ligand} \overset{k_a}{\underset{k_d}{\rightleftharpoons}} AB_{complex}$$

Equation 4.4
Figure 4.11: Sensorgrams showing the effect of flow rate on the binding of analyte, IgM, to ligand, hSC_rDI.

Table 4.2: Rate constants obtained by performing local fits on single curve sets collected for 100nM IgM injections at varying flow rates with hSC_rDI as ligand.

<table>
<thead>
<tr>
<th>Flow rate (µl/min)</th>
<th>$k_a (M^{-1} s^{-1})$</th>
<th>$k_d (s^{-1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>$2.52 \times 10^4$</td>
<td>$4.96 \times 10^{-4}$</td>
</tr>
<tr>
<td>30</td>
<td>$2.13 \times 10^4$</td>
<td>$8.55 \times 10^{-4}$</td>
</tr>
<tr>
<td>50</td>
<td>$7.62 \times 10^3$</td>
<td>$1.22 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

The association rate constants vary quite significantly from 10 to 50µl/min (Table 4.2). Data were collected at all flow rates although the set collected at 30µl/min was used for analysis due to attenuation of signal at higher flow rates and low analyte concentrations. Mass transfer limitations are more likely to be present at low flow rates (10µl/min). The flow rate (30µl/min) is suggested for use as a “starting point” in order to overcome mass transfer limitation skewing data fits (Biacore, 1997). Data were fit using the Langmuir model (equation 4.4).

Figure 4.12 shows the model fit of IgM binding to hSC_rDI using the Langmuir 1:1 binding model and the local fitting of parameters $k_a$ and $k_d$. This acts as a test of mass transfer limitation in the system. Table 4.3 shows the minimal variation that occurs among the rate constants as concentration increases at 30µl/min. As mentioned this shows the minimal mass transfer effects present at this flow rate.
Figure 4.12: (a) Sensorgrams showing the Langmuir fit of hSC_rDI binding to varying IgM concentrations (analyte) at 30µl/min using local fitting algorithm (see text for details). (b) relative residual curves showing deviation of fit from the data points.

Table 4.3: Rate constants obtained from local data fitting of IgM binding to hSC_rDI

<table>
<thead>
<tr>
<th>[IgM] nM</th>
<th>$k_a \times 10^4$ (M$^{-1}$s$^{-1}$)</th>
<th>$k_d \times 10^4$ (s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>4.22</td>
<td>5.56</td>
</tr>
<tr>
<td>25</td>
<td>6.02</td>
<td>9.21</td>
</tr>
<tr>
<td>50</td>
<td>4.01</td>
<td>0.24</td>
</tr>
<tr>
<td>100</td>
<td>5.86</td>
<td>8.25</td>
</tr>
<tr>
<td>250</td>
<td>5.63</td>
<td>5.02</td>
</tr>
<tr>
<td>500</td>
<td>4.46</td>
<td>4.45</td>
</tr>
<tr>
<td>1000</td>
<td>3.19</td>
<td>5.22</td>
</tr>
</tbody>
</table>

The global fitting algorithm was applied to the data set using the Langmuir binding model (Figure 4.13).
A visual inspection of the relative residuals (Figure 4.13b) shows the increased deviation of the fit from measured data points when compared to Figure 4.12b. It is important to note that the data fit does not appear to describe the association curve at higher concentrations (1000nM IgM). This may be as a result of the potential of IgM to bind more than one pIgR/SC molecule. It is important to note that a good fit does not necessarily mean that it is a representation of the biological system. The $K_D$ obtained from the global fit was 9.42nM (Table 4.5). This is a much higher affinity for IgM than previously reported (~510nM; Bakos et al., 1994).

Figure 4.13: Sensorgrams showing the Langmuir fit of hSC_rDI binding to varying IgM concentrations (analyte) at 30µl/min using global fitting algorithm. (b) relative residual curves showing deviation of fit from the data points.

IgM was reported to possess multiple binding sites for plgR/fSC owing to its pentameric nature and molar content of J chain (>1) (Lindh and Bjork, 1976). It stands to reason, therefore, that IgM may at least be considered a bivalent analyte and therefore, data fitting with the bivalent analyte model as outlined by equation 4.5 should be more biologically correct:

\[
A_{analyte} + B_{ligand} \rightleftharpoons [AB]_{complex} + B_{ligand} \rightleftharpoons [AB]_{2complex}
\]

Equation 4.5
where $k_a$ and $k_d$ represent association and dissociation rate constants and 1 and 2 represent the first and second reactions, respectively. Figure 4.14 shows the application of the bivalent model in the description of binding by IgM to hSC_rDI. The data appears to fit better at higher concentrations of IgM. This may be as a result of increased analyte binding sites being made available and the model more accurately representing the binding events. Figure 4.15 represents the individual components as defined by the fit. The complex consisting of $AB_2$, i.e. a single analyte bound to two ligands dominates all other binding events. Table 4.6 summarises the rate constant data derived from the model fit.

![Figure 4.14(a)](image-a)

![Figure 4.14(b)](image-b)

**Figure 4.14:** (a) Sensorgrams showing the Bivalent model fit to hSC_rDI binding to varying IgM concentrations (analyte) at 30µl/min using global fitting algorithm (see text for details). (b) relative residual curves showing deviation of fit from the data points.
Figure 4.15: Component view showing the contribution (at 1000nM IgM) of individual complexes and bulk and drift to total binding response with hSC_rDI, as defined by the bivalent model.

The multiple domain mutants hSC_rDI-II and hSC_rDI-III were immobilized at 872 and 936 RU, respectively. As was observed with the analysis of hSC_rDI, responses vary with flow rate. Figure 4.16 shows the response variation with flow rate at 100nM IgM binding to hSC_rDI-II (figure 4.16a) and hSC_rDI-III (figure 4.16b).

Figure 4.16: Sensorgrams showing the effect of flow rate on the binding of analyte, IgM, to ligand. (a) hSC_rDI-II and (b) hSC_rDI-III.
Table 4.4 shows the rate constants obtained from single curve global fits at 100nM IgM for ligands, hSC_rDI-II and hSC_rDI-III. When compared to the single curve analysis of hSC_rDI the variation in rate constants is not as pronounced. This may be as a result of the ratio as defined by equation 4.6 becoming smaller (derived from equation 4.7):

\[
\frac{M_{W_{analyte}}}{M_{W_{ligand}}} \cdot R_{ligand} \cdot V_{ligand} = R_{max}
\]

Equation 4.6

Equation 4.7

where \(M_w\) is the molecular weight (\(x \equiv \text{analyte or ligand}\)). \(R_{max}\) is the maximal response dependent on the level immobilized (\(R_{Ligand}\)) and the valency of the ligand (\(V_{Ligand}\)). The ratio of analyte and ligand sizes is directly proportional to response (RU). Essentially, from the analysis as shown in Table 4.4, mass transport is not as limiting in the systems where hSC_rDI-II and hSC_rDI-III (larger ligands) are immobilized.

Table 4.4: Rate constants obtained by performing global fits on single curve sets of hSC_rDI-II and hSC_rDI-III collected for 100nM IgM injections at varying flow rates.

<table>
<thead>
<tr>
<th>Flow rate (µl/min)</th>
<th>hSC_rDI-II</th>
<th>hSC_rDI-III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(k_a) ((10^{-5}) (M(^{-1}).s(^{-1}))</td>
<td>(k_d) ((10^3) (s(^{-1}))</td>
</tr>
<tr>
<td>10</td>
<td>0.91</td>
<td>0.46</td>
</tr>
<tr>
<td>30</td>
<td>1.74</td>
<td>0.93</td>
</tr>
<tr>
<td>50</td>
<td>3.03</td>
<td>0.99</td>
</tr>
</tbody>
</table>

The data sets collected at 50µl/min were used for data fitting in order to ensure reduction of mass transport limitation, if present. Figure 4.17 (a-d) shows the fitting of the Langmuir model to the data sets obtained for hSC_rDI –II (Figure 4.17a) and hSC_rDI-III (Figure 4.17c) as ligands. The goodness of fits are represented by the residual curves (Figure 4.17b & d). As observed for the fitting of IgM binding to hSC_rDI (as ligand) the Langmuir model does not adequately explain binding at higher concentrations of analyte. The fits are adequate when considering the goodness of fit as denoted by \(\chi^2\) (fits are considered good when \(\chi^2 < 10\), [Arnoux Marquart, Pers. Comm.]). Table 4.5 shows the kinetic constants obtained from the Langmuir model fits. The Langmuir fits do show that
there appears to be not much of an advantage in possessing domain II or III for IgM binding when considering the $K_D$ of hSC_rDI (9.42nM).

As discussed in the analysis of hSC_rDI, the bivalent analyte model is possibly the most biologically correct model to describe the binding of hSC_rDI-II and hSC_rDI-III to IgM.
as analyte. Figure 4.18 (a-d) shows the fitting of the bivalent model to describe IgM binding to the hSC_rDI-II and hSC_rDI-III data sets. Visual inspection of the fits reveal improved fitting to the data at higher analyte concentrations (500nM). Figure 4.19 shows the component view of the model fits at 500nM IgM. From the component view of hSC_rDI-II binding to IgM it can be seen that the AB₂ complex formation dominates initially and then it appears that the single complex (AB) formation dominates the total binding. The AB₂ complex formation completely dominates the total binding contribution in the fitting of hSC_rDI-III to IgM. The question is, does this show the existence of other potential contact points between IgM and hSC_rDI-III or perhaps a synergistic effect between domain II and III? There may be a combined effect of the two domains on binding if one considers the lack of binding to IgA that occurs when the domains are deleted. Although it was shown (Norderhaug et al., 1999) that deletion of the domains resulted in no loss of IgM binding ability, the domains may still play a role in binding by possessing undefined contact points required for stabilization of the complex. Table 4.6 shows the kinetic constants derived from the bivalent model fits. Immediately noticeable are the low \(T\) values for hSC_rDI-III. The \(T\) values shown in the table are relative measures of the standard error defined by dividing the value of \(k_a\) or \(k_d\) by the standard error.

**Table 4.5:** Kinetic constants obtained from performing global fits of the Langmuir model to data sets for IgM binding to hSC_rDI, hSC_rDI-II and hSC_rDI-III.

<table>
<thead>
<tr>
<th></th>
<th>(k_a) ((10^n)) (M⁻¹. s⁻¹)</th>
<th>(T)</th>
<th>(k_d) ((10^n)) (s⁻¹)</th>
<th>(T)</th>
<th>(K_D) (nM)</th>
<th>(\chi^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hSC_rDI</td>
<td>0.56</td>
<td>163</td>
<td>0.53</td>
<td>41.5</td>
<td>9.42</td>
<td>0.92</td>
</tr>
<tr>
<td>hSC_rDI-II</td>
<td>1.98</td>
<td>193</td>
<td>1.27</td>
<td>68.1</td>
<td>6.41</td>
<td>1.64</td>
</tr>
<tr>
<td>hSC_rDI-III</td>
<td>1.23</td>
<td>194</td>
<td>1.37</td>
<td>83.1</td>
<td>11.1</td>
<td>0.726</td>
</tr>
</tbody>
</table>

**Table 4.6:** Kinetic constants obtained from performing global fits of the bivalent analyte model to data sets for IgM binding to hSC_rDI, hSC_rDI-II and hSC_rDI-III.

<table>
<thead>
<tr>
<th></th>
<th>(k_{a1}) ((10^n)) (M⁻¹. s⁻¹)</th>
<th>(T)</th>
<th>(k_{d1}) ((10^n)) (s⁻¹)</th>
<th>(T)</th>
<th>(k_{a2}) ((RU^n)) (s⁻¹)</th>
<th>(T)</th>
<th>(k_{d2}) () (s⁻¹)</th>
<th>(T)</th>
<th>(\chi^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hSC_rDI</td>
<td>1.13</td>
<td>141</td>
<td>48.3</td>
<td>7.51</td>
<td>0.0666</td>
<td>8.28</td>
<td>0.0054</td>
<td>7.61</td>
<td>0.877</td>
</tr>
<tr>
<td>hSC_rDI-II</td>
<td>6.13</td>
<td>110</td>
<td>3.50</td>
<td>45.7</td>
<td>31.5</td>
<td>4470</td>
<td>161</td>
<td>4800</td>
<td>1.2</td>
</tr>
<tr>
<td>hSC_rDI-III</td>
<td>2.25</td>
<td>135</td>
<td>555</td>
<td>1.06</td>
<td>0.864</td>
<td>4.13</td>
<td>0.0441</td>
<td>1.08</td>
<td>0.791</td>
</tr>
</tbody>
</table>
$T$ values are said to be statistically valid when greater than 10 (Biacore, 1997). Therefore, the bivalent model may explain the initial association better but appears not to resolve the dissociation phases.

Figure 4.18: Sensorgrams showing the bivalent model fitted to varying IgM concentrations binding to (a) hSC_rDI-II and (c) hSC_rDI-III. Relative residual curves showing deviation of fit from data points are shown for (b) hSC_rDI-II and (d) hSC_rDI-III.
Sensorgrams were not obtained for hSC_rDI-IV and hSC_rDI-V as ligands due to insufficient material at the time. Recombinant mutants were re-expressed, refolded and purified according to the methods in Chapter 3.

4.2.3.2 Recombinant Domain Mutants as Analyte

Due to the large mass of IgM (~1MDa) it is thought that this affected mass transfer significantly as the large $M_w$ IgM moved over the surface, resulting in non specific contact points. It was, therefore, decided to immobilize IgM to minimize mass transfer. It has been said that detecting a small analyte binding to a large ligand is like detecting a flea on an elephant’s back but due to the sensitivity of the SPR technique it is thought not to pose a significant problem, particularly if the small analyte has a molecular weight of more than 10kDa (Arnoud Marquart, Pers. Comm.).
Immobilization of IgM also removes the requirement of complex analysis in terms of data fitting, i.e. the bivalent analyte has now been immobilized and therefore the monovalent ligand, now the analyte, can be assumed to follow simple 1:1 Langmuir binding kinetics. This is, of course, based on the assumption that all the binding sites are homogeneous. IgM was immobilized at 13413 RU.

Data sets were collected in triplicate at 30µl/min for hSC_rDI as analyte and the data were analysed using the initial rate analysis method of Edwards and Leatherbarrow (1997). An injection of 1000nM recombinant domain I allowed for the determination of the physical $R_{\text{max}}$ (~160RU) which correlated well with the theoretical value (179RU) calculated by equation 4.7. This allowed the determination of the association rate constant ($k_a$) by equation 4.8:

$$\frac{dR}{dt} = R_{\text{max}} \times [\text{Ligand}] \times k_a$$

**Equation 4.8**

The dissociation rate constant ($k_d$) was determined from the saturation curve and performing a separate fit in the BIAevaluation software package and the $K_D$ was determined from the relationship (equation 4.9):

$$K_D = \frac{k_d}{k_a}$$

**Equation 4.9**

Initial analysis using IgM as an analyte resulted in large mass transfer effects resulting in variation in the response with flow rate (section 4.2.3.1). This is largely attributed to the size of IgM (0.97 MDa). Generally, data collected with IgM as ligand proved reliable with no observable mass transfer effects and reproducible results. Unfortunately, some parts of the hSC_rDI curves proved unusable due to aberrations in the curve (Figure 4.20). This was overcome by using initial rate analysis of Edwards and Leatherbarrow (1997). Initial rate analysis (Figures 4.21 and 4.22) allowed us to measure the fast association rate before molecular crowding events occurred (Edwards and Leatherbarrow, 1997). It appears that domain I has a much greater affinity (almost 100 fold) for IgM than previously presented (Table 4.7). This is attributable to the combined effects of a fast
association \( (k_a = 1.29 \times 10^6 \text{M}^{-1}\text{s}^{-1}) \) and a slow dissociation \( (k_d = 4.08 \times 10^{-3} \text{s}^{-1}) \). The slow dissociation is clearly demonstrated by the residual levels of analyte left after the change from injection to running buffer (Figure 4.21a). Our data also show that the IgM:domain I interaction appears stronger than the IgA:domain I interaction reported by Bakos et al. (1994) and Hamburger et al. (2004). The low \( K_D \) reported here allows further insight into the findings of Norderhaug et al. (1999) that deletion of domains II and III or IV and V resulted in no loss of binding to pentameric IgM; the domain II and III deletion resulted in almost complete abrogation of IgA binding. This study highlighted the difference that exists between IgM and IgA binding requirements with respect to domain interaction, i.e. the IgM:domain I interaction is the controlling factor in the IgM:pIgR interaction. Our data underlines these findings by demonstrating that the interaction is in fact in the low nanomolar range (Table 4.7).

![Figure 4.20: Sensorgrams of hSC_rDI, as analyte, binding to varying IgM concentrations. The curves show aberrations that prevented reliable global fits and were overcome by using an initial rate analysis.](image-url)

79
Figure 4.21: (a) Sensorgram showing a representative saturation curve for hSC_rDI as analyte that allowed for the determination of $R_{max}$ and $k_d$, and (b) the initial portions of the sensorgram curves used for the determination of the association constant $k_a$.

Figure 4.22: Linear regression curve used for the determination of the association rate of hSC_rDI for IgM from the initial rates. The slope of the curve is used for the determination of $k_a$ using equation 4.8. Error bars represent standard deviation ($n = 3$) ($R^2 = 0.9988$).
Table 4.7: Kinetic constants for the interaction of pIgR domain I with IgM or IgA. Data were collected using competitive ELISA (in competition with native SC) or SPR. All data from recombinant domain I of pIgR except Bakos et al. (1991) that used tryptic cleavage fragments of pIgR.

<table>
<thead>
<tr>
<th></th>
<th>IgM</th>
<th>IgM</th>
<th>IgA</th>
<th>IgA</th>
<th>IgA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_D$ (M)</td>
<td>$3.16 \times 10^{-9}$</td>
<td>$5.1 \times 10^{-7}$</td>
<td>$4.00 \times 10^{-7}$</td>
<td>$1.6 \times 10^{-7}$</td>
<td>$3.0 \times 10^{-7}$</td>
</tr>
<tr>
<td>$k_a$ (M$^{-1}$s$^{-1}$)</td>
<td>$1.29 \times 10^{6}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$k_d$ (s$^{-1}$)</td>
<td>$4.08 \times 10^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a, Present study, SPR;
b, Bakos et al. (1994), ELISA;
c, Bakos et al. (1991), ELISA;
d, Hamburger et al. (2004), SPR (equilibrium data analysis).

In summary, our data adds significantly to the argument that IgM only requires a non-covalent interaction with domain I of pIgR for recognition and transport and that the interaction is strong enough to maintain the IgM:pIgR complex during transcytosis and release of SIgM at the apical surface.

Figures 4.23-4.27 show the data fits of hSC_rDI-II - hSC_rDI-V binding to the immobilized IgM using the 1:1 Langmuir model. From visual inspection the immediate conclusion drawn is that the binding of hSC_rDI-II and hSC_rDI-III cannot be adequately described using 1:1 binding kinetics. The residual curve fits (Figure 4.23b - 4.24b) show high levels of deviation from the real time data. The model does not appear to describe the association and dissociation curves at high IgM concentrations. The deviation seen at the switch from association to dissociation (Figures 4.24 and 4.25) at approximately 70 seconds for 100nM hSC_rDI-III and hSC_rDI-IV, respectively, is simply a buffer change artefact. This causes the residual curves to deviate from the fit but does not affect the overall fit.
Figure 4.23: (a) Sensorgrams showing the Langmuir model fitted to varying concentrations of hSC_rDI-II, as analyte, binding to IgM as ligand. (b) Relative residual curve showing deviation of fit ($\chi^2=9.232$).
Figure 4.24: (a) Sensorgrams showing the Langmuir model fitted to varying concentrations of hSC_rDI-III, as analyte, binding to IgM, as ligand. (b) Relative residual curve showing deviation of fit ($\chi^2 = 6.215$).
Consistent sensorgrams of multiple domain mutant hSC_rDI-IV, containing domains I through IV, were obtained for an IgM concentration range of 25-500nM (Figure 4.25a). The residual curve (Figure 4.25b) and the $\chi^2$ (4.42) shows that the Langmuir model fits (Figure 4.25a) the binding data with slight deviations observed for the dissociation curves. Table 4.9 shows the kinetic constants derived from the model fit.

Figure 4.25: (a) Sensorgrams showing the Langmuir model fitted to varying concentrations of hSC_rDI-IV, as analyte, binding to IgM, as ligand. (b) Relative residual curve showing deviation of fit ($\chi^2 = 4.42$).
The recombinant multiple domain mutant, hSC_rDI-V, containing domains I through V, is the first recombinant free secretory component (rSC) produced in a bacterium and refolded \textit{in vitro} (Prinsloo \textit{et al.}, 2006;-Appendix C). Published results show the molecule to be physiologically active with a reported $K_D$ of 46.5 nM. Table 4.8 shows the data obtained from the separate global fitting of the association and dissociation phases based on simple Langmuir 1:1 binding kinetics (Figure 4.26). The $K_D$ obtained from the non-linear fit is comparable to that obtained for the ELISA receptor-ligand assay, i.e. 36.15nM (Table 4.1).

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
 & hSC_rDI-V (rSC) & Native SC \\
\hline
$K_D \ (10^8)(M)$ & $4.65 \pm 1.93$ & $3.00^a$ & $0.64^b$ \\
\hline
$k_a \ (10^{-5})(.M^{-1}.s^{-1})$ & $1.25 \pm 0.37$ & & \\
$k_d \ (10^3)(.s^{-1})$ & $5.43 \pm 0.67$ & & \\
\hline
\end{tabular}
\caption{Kinetic constants of hSC_rDI-V (rSC) for IgM obtained from data fitting, compared with reported values for native SC. Association ($k_a \pm SD$) and dissociation ($k_d \pm SD$) rates shown are from separate global fits. (n=3).}
\end{table}

\textit{a, Bakos \textit{et al.} (1991);}
\textit{b, Bakos \textit{et al.} (1994).}

Separate fitting allows an adequate determination of $k_a$ and $k_d$ values but are not ideal. Whole curve fitting is the preferred method of model fitting. Whole curve fitting of the Langmuir model (Figure 4.27 and Table 4.9) shows the incomplete description of the dissociation phase of the curve.
Figure 4.26: (a) Normalised fitted sensorgrams of varying concentrations of recombinant free human SC (hSC_rDI-V) binding to IgM, as ligand. Data fits were performed using the Langmuir 1:1 (simple bimolecular) model. (b) The relative residual plot shows the goodness of fit.
Figure 4.27: (a) Sensorgrams showing the Langmuir model fitted to varying concentrations of hSC_rDI-V (rSC), as analyte, binding to IgM, as ligand. (b) Relative residual curve showing deviation of fit ($\chi^2 = 5.638$).
Let us consider what we know from the data fits of the recombinant domain mutants as ligands and IgM as an analyte. IgM is a potential multivalent ligand of fSC. Clearly immobilization of IgM should, in theory, have allowed for simple 1:1 kinetics. This is based on the assumption that all the binding sites are equal. Unfortunately, this is rarely the case. Considering the bivalent model (equation 4.5) the reaction scheme follows a system where the analyte (IgM) has two independent binding sites for the ligand (pIgR). Immobilization of IgM creates a ligand that has at least two independent binding sites of varying affinity for the receptor, hence it becomes a heterogeneous ligand. Consider the heterogeneous ligand model (equation 4.10), a system defined by simultaneous/parallel kinetic reactions:

$$B_1 + A \xleftrightarrow{k_{a1}} B_1A \ & \ B_2 + A \xleftrightarrow{k_{d2}} B_2A$$

Equation 4.10

where $B_x$ represents the ligand ($x$ is 1 or 2 representing the ligands present, in this case binding sites of IgM), $A$ is the analyte, and $k_{ax}$ and $k_{dx}$ ($x = 1$ or 2) are the association and dissociation rate constants, respectively, as previously described. Assume that the multiple binding sites of IgM be considered independent ligands.

The heterogeneous model was fit to the data sets of hSC_rDI-II (Figure 4.28) and hSC_rDI-III (Figure 4.29). Upon inspection of the curve fits it appears that the heterogeneous ligand model effectively describes the binding of both hSC_rDI-II and hSC_rDI-III. This is clearly apparent from the residual curves (Figure 4.28b and 4.29b) as the decrease in deviation from fit represents an increased goodness of fit represented by the residual sum of squares, $\chi^2$. Table 4.9 summarises the kinetic constants derived from the fit.
Figure 4.28: (a) Sensorgrams showing the Heterogeneous ligand model fitted to varying concentrations of hSC_rDI-II, as analyte, binding to IgM, as ligand. (b) Relative residual curve showing deviation of fit ($\chi^2 = 3.273$).
Figure 4.29: (a) Sensorgrams showing the Heterogeneous ligand model fitted to varying concentrations of hSC_rDI-III, as analyte, binding to IgM, as ligand. (b) Relative residual curve showing deviation of fit ($\chi^2 = 1.968$).

From Figure 4.30 it is clear that fitting the heterogeneous ligand model to hSC_rDI-IV improves curve fitting although only minimally when compared to the Langmuir 1:1 fits (Figure 4.25). Why does hSC_rDI-IV exhibit this type of behaviour? As was explained previously, IgM is a potential multivalent ligand of fSC, hence IgM does allow for flexibility in how many fSC molecules it can associate with. This appears to be the case.
with the hSC_rDI-IV binding described here. An alternative explanation may be simple steric hindrance. The presence of an additional domain (domain IV) may cause a conformational change in the recombinant protein that causes the domain II and III contact points to be partially obstructed or, resulting in decreased binding capacity of the recombinant protein; i.e. favouring simple 1:1 binding kinetics.

Figure 4.30: (a) Sensorgram showing the Heterogeneous ligand model fitted to varying concentrations of hSC_rDI-IV, as analyte, binding to IgM, as ligand. (b) Relative residual curve showing deviation of fit ($\chi^2=4.24$).
Figure 4.31 shows the almost perfect fitting of the heterogeneous model to the binding data of hSC_rDI-V. Application of the heterogeneous ligand model clearly allows for better fitting of the dissociation phase and the association phases after injection of higher analyte concentrations.

Figure 4.31: (a) Sensorgram showing the Heterogeneous ligand model fitted to varying concentrations of hSC_rDI-V (rSC), as analyte, binding to IgM, as ligand. (b) Relative residual curve showing deviation of fit ($\chi^2 = 1.533$).

The complexity of the interaction between IgM and fSC makes it difficult to distil a single model to describe the SPR data for the interactions of the multiple domain mutants.
and IgM. The simple 1:1 model does appear to describe the association at low concentrations when the recombinant proteins are used as ligands or analytes but appears to fail at higher concentrations. The dissociation rates are described to varying degrees with almost complete failure for hSC_rDI-II and hSC_rDI-III. Why is this? Why the requirement for multiple models? If we consider the pIgR/fSC molecule and its function, typically each molecule of pIgR produced binds a polymeric immunoglobulin, representing a high energy cost for the cell, especially considering that the bound pIgR is cleaved to be released as SC bound to IgM. Due to this harsh extracellular environment, binding is necessarily strong (nM range). It has been reported that only complete denaturation releases the SC from IgM (Lindh and Björk, 1976). This is apparent from the dissociation phases of the sensorgrams displayed. The simple bimolecular model describes simple on and off rates of the binding, not taking into account the strong interaction between the recombinant protein and the immunoglobulin. The bivalent and heterogeneous ligand models appear to fit better due to their description of multiple binding sites available on the IgM molecule and potential rebinding of dissociated analytes. Consider that more molecules present may result in molecular crowding effects and weak binding analytes may be displaced by stronger binding analytes, and the displaced analytes would later rebind in the laminar flow over the chip surface.

Table 4.9 shows a comparison of rate and equilibrium constants derived from 1:1 Langmuir and heterogeneous ligand model fits for the multiple domain mutants. A comparison of equilibrium dissociation constants is largely inconclusive due to large variation contradicting the receptor-ligand ELISA assay. It should be mentioned that the receptor-ligand ELISA assay is a measure of equilibrium conditions. Also the 1:1 Langmuir model fits were not ideal. The heterogeneous ligand model describes association of the curves more effectively. The initial association rate ($k_{a1}$) appears to increase >10-fold as the recombinant decreases in length, i.e. as domains are deleted. The $k_{a1}$ rate, therefore, appears to decrease whereas the $k_{d1}$ rates appear to increase as domains are added, especially domains IV-V. Does this signify a potential mechanism within these domains that promotes dissociation? If one considers that the original described function of SC was to prevent proteolytic cleavage of the immunoglobulin, why would some
domains actively promote dissociation when it would be more favourable to remain associated? Firstly, one point to consider is that the increase in dissociation rate is not that considerable; this may simply be a method of regulating binding to IgM. Secondly, more functions have been elucidated for SC bound to immunoglobulins; this increase in dissociation rate may be related to some undescribed function. A better picture of binding affinity of the recombinant domain mutants would most certainly be painted if equilibrium conditions were achieved, but this would not have allowed for kinetic analysis of individual rates (Table 4.9). It is most interesting that association rates decrease accordingly as domains are added; is this related to conformational change and loss of contact points?

Assume that domain I’s CDR loops are considered essential for binding to IgM; do the CDR loops present on the other domains perhaps play a role in binding? All three CDR loops of domain I (especially CDR2) are required for binding or at least initial association with IgM. Mutation of these results in abrogation of binding (Røe et al., 1999). Hamburger et al. (2004) showed the presence of conserved residues in the rest of the domains (i.e. II-V) upon alignment with domain I and the known structural elements of domain I. Among these determinants surely lie contact points for binding to IgM. It is known that domains II and III are required for maximal dIgA binding (Norderhaug et al., 1999). This, therefore, points at potential functions of the CDR loops in domains II and III for IgA binding, but does this also point at interactions with IgM? Piskurich et al. (1995) and Hamburger et al. (2004) have shown, by alignment, that structural elements are conserved in domain I among species (these include human through to the chicken). More variation, however, occurs among domains II-V (Piskurich et al., 1995). Rabbit pIgR undergoes alternate splicing (domains II and III are on a single exon) to yield a SC that only contains domain I, IV and V (Deitcher and Mostov, 1986; Frutiger et al., 1986). Socken and Underdown (1978) showed that the relative ability to displace \(^{125}\)I-labelled human SC from human IgM was as follows: human>bovine>sheep>>rabbit>rat. This clearly points at variable binding to IgM among species. Furthermore rabbit SC was shown to bind IgM with a seven-fold lower affinity than rabbit IgA. Rabbit and rat IgM secretions are observed to be low or not present at all (Eddie et al., 1971; McGhee et al.,
1975). All other species aside, it is clear from the studies of Norderhaug et al. (1999) that domain requirements for human IgA and IgM vary significantly. Exactly what determinants in these domains regulate binding requires further investigation.

Individual domain interaction may also be resolved by performing competitive binding assays specifically in competition with native fSC or of recombinant domain deletion mutants with known affinities. This may actually be a better approach because it was shown that it can avoid rebinding effects (Neiba et al., 1996); thereby simplifying the interaction analysis and model fitting.

Table 4.9: Kinetic constants derived from the model fits to the data sets of recombinant SC mutants binding to IgM as ligand. 1:1, Langmuir model and HL, Heterogeneous ligand model. SD is the standard deviation (n=3).

<table>
<thead>
<tr>
<th></th>
<th>$k_{a1} \pm SD$ $(10^3)(\text{M}^{-1}\text{s}^{-1})$</th>
<th>$k_{d1} \pm SD$ $(10^3)(\text{s}^{-1})$</th>
<th>$k_{a2} \pm SD$ $(10^3)(\text{RU}^{-1}\text{s}^{-1})$</th>
<th>$k_{d2} \pm SD$ $(10^2)(\text{s}^{-1})$</th>
<th>$K_D \pm SD$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hSC_rDI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:1*</td>
<td>12.9 ± 4.08</td>
<td>3.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-II</td>
<td>9.33 ± 0.46</td>
<td>6.28 ± 0.07</td>
<td>6.75 ± 0.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HL</td>
<td>9.55 ± 0.44</td>
<td>2.52 ± 0.22</td>
<td>742 ± 94.3</td>
<td>4.09 ± 0.25</td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-III</td>
<td>1.85 ± 0.63</td>
<td>6.49 ± 0.54</td>
<td>35.1 ± 20.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HL</td>
<td>1.88 ± 0.43</td>
<td>2.16 ± 1.25</td>
<td>30.2 ± 2.57</td>
<td>3.68 ± 0.82</td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-IV</td>
<td>0.635 ± 0.021</td>
<td>5.90 ± 0.46</td>
<td>93.1 ± 10.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HL</td>
<td>0.159 ± 0.164</td>
<td>6.45 ± 0.36</td>
<td>259 ± 47.0</td>
<td>6.94 ± 3.00</td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-V</td>
<td>1.06 ± 0.24</td>
<td>8.52 ± 2.95</td>
<td>80.4 ± 38.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HL</td>
<td>0.68 ± 0.20</td>
<td>5.97 ± 2.75</td>
<td>137 ± 89.0</td>
<td>1.22 ± 0.25</td>
<td></td>
</tr>
</tbody>
</table>

* determined by initial rate analysis

4.2.4 Competitive Binding Studies

Competitive binding would allow a better understanding of individual domain interactions between the recombinant domain mutants and IgM. Due to constraints with
protein quantities the competitive binding studies were only performed using hSC_rDI, hSC_rDI-III and hSC_rDI-V; this does, however, give insight into the synergistic effects of domains II and III, and IV and V.

4.2.4.1 Immunohistochemical and Western Blot Analysis of pIgR Expression in HT29

Immunohistochemical visualization of pIgR expression by HT29 cells was inconclusive. Figure 4.32 shows fixed HT29 cells probed with biotinylated anti-human SC in comparison to the negative control. The results are inconclusive due to the high background staining. It was decided to induce expression of pIgR using TNFα.

![Figure 4.32: Formalin fixed uninduced HT29 cells probed for pIgR expression. (a) biotinylated anti-human SC and (b) negative control extravidin peroxidase conjugate only. The arrows indicate potential stained cells. Magnification 100x](image)

Western blots were performed to determine the presence of pIgR upon induction of HT29 cells with TNFα. Tumour necrosis factor has been used before to successfully induce pIgR expression above constitutive levels (Bruno and Kaetzel, 2005). Figure 4.33 shows a typical blot after DAB exposure. Expression levels do not appear to increase upon TNFα induction. The inability to effectively detect pIgR (~110kDa) may be as a result of epitope distortion by the levels of detergents in the extraction buffers, although they have been optimized for membrane lysis and native protein extraction. The SDS in the electrophoresis and transfer buffers may also have played a role. It should be noted that
the anti-human SC was raised against fSC and SIgA; therefore, the potential that the epitope recognition sites may differ from membrane bound pIgR may be inherent in anti-human SC.

Figure 4.33: Western blot showing induction of pIgR expression in HT29 cells. (a) and (d) PageRuler Prestained Protein ladder (Fermentas) markers; (b) induced and (c) uninduced HT29 cell lysate probed for pIgR expression. Box indicates region of interest. Arrow indicates size range of interest (100-110kDa).

4.2.4.2 Whole Cell ELISA Competitive Binding Assay

The competitive binding assays show definite binding upon induction of HT29 cells and little or no binding in uninduced cells (Figure 4.34-4.36). Model fits, using a one site competition model, were only performed with data showing positive binding, i.e. for induced HT29 cells. This led to the conclusion that TNFα induction was successful, although the western blot showed no specific band. Although the data distribution in Figure 4.35 and 4.36 suggest a slight biphasic nature, fitting with a two site competition model failed to describe the data. The inability of recombinant CD89 to inhibit IgM (Figure 4.37) binding to the cell surface shows that specific binding of the hSC_rD
mutants to IgM occurred, effectively sequestering IgM from its native pIg receptor on the cell surface.

Figure 4.34: Competitive binding assay with hSC_rDI as competitor. (a) uninduced and (b) induced HT29 cells. Error bars represent standard deviation (n=3). Solid lines represent the best fit of the one site competition model ($R^2=0.9874$).

Figure 4.35: Competitive binding assay with hSC_rDI-III as competitor. (a) uninduced and (b) induced HT29 cells. Error bars represent standard deviation (n=3). Solid lines represent the best fit of the one site competition model ($R^2=0.9620$).

Figure 4.36: Competitive binding assay with hSC_rDI-V as competitor. (a) uninduced and (b) induced HT29 cells. Error bars represent standard deviation (n=3). Solid lines represent the best fit of the one site competition model ($R^2=0.9192$).
Figure 4.37: Competitive binding assay with recombinant CD89 as competitor performed using induced HT29 cells. Error bars represent standard deviation (n=3).

Table 4.10 shows the IC$_{50}$ values obtained from model fitting to the data. Initial conclusions from Table 4.10 are that the recombinant domain mutants (at least hSC_rDI, hSC_rDI-III and hSC_rDI-V) possess biological activity or at least an ability to effectively compete with the native plgR molecule in an *in vitro* system using attached whole cells. The results show hSC_rDI-V has approximately 19 and 11 times greater inhibitory potential compared to hSC_rDI and hSC_rDI-III, respectively. Equation 4.11 allows for the calculation of $K_i$; i.e. the equilibrium dissociation constant of the competitor (essentially the $K_D$ of the competitor):

$$K_i = \frac{IC_{50}}{1 + \frac{[\text{Ligand}]}{K_D}}$$

**Equation 4.11**

The $K_D$ referred to in equation 4.11 is the equilibrium dissociation constant of the labelled ligand for the bound receptor. The $K_i$ values shown in Table 4.10 were calculated using equation 4.11. The $K_D$ value used was that reported by Norderhaug *et al.* (1999). The $K_D$ of pentameric IgM for wild type plgR was 37.1nM. The ligand concentration was 625nM. It is clear that affinity of the recombinant domain mutants for IgM increases as domains are added. This may be as a result of increased stability of the hSC_rDI/plgM complex, potentially due to increased, as yet, undefined contact points between SC and plgM. It should be noted that the affinities reported here may be higher due to potential low levels of expression of plgR, even after induction. This may result in data analysis being skewed in favour of the recombinant competitor.
Table 4.10: IC50 and \( K_i \) values derived from the one site competition model fit.

<table>
<thead>
<tr>
<th></th>
<th>IC50 (nM)</th>
<th>95% Confidence (nM)</th>
<th>( R^2 )</th>
<th>( K_i ) (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hSC_rDI</td>
<td>798.9</td>
<td>486.9 – 1311</td>
<td>0.9874</td>
<td>44.8</td>
</tr>
<tr>
<td>hSC_rDI-III</td>
<td>494.3</td>
<td>342.3 – 713.7</td>
<td>0.9620</td>
<td>27.7</td>
</tr>
<tr>
<td>hSC_rDI-V</td>
<td>43.15</td>
<td>23.85 – 78.05</td>
<td>0.9192</td>
<td>2.42</td>
</tr>
</tbody>
</table>

The inhibitory constants determined for the recombinant domain mutants show that the mutants constructed in this study can effectively outcompete native pIgR for IgM. The data compares well with the trend observed for the equilibrium based receptor-ligand ELISA assay. Additional domains appear to add a competitive edge to the recombinant mutants, i.e. a defined increase in the inhibitory constant (~19x) is observed when all 5 domains are present, as opposed to the single domain I mutant. Not much competitive advantage can be attributed to the presence of domains II and III in hSC_rDI-III. Why is this? This may point at determinants within domains II-V that regulate binding, particularly in domains IV and V. The presence of all domains (in the human) must confer a competitive advantage to the molecule. As discussed (section 4.2.3.2) it is known that there is significant species variation in the interaction between SC and IgM (Socken and Underdown, 1978). The structural determinants that regulate the binding, however, remain unknown. Most probably, these determinants are indeed the CDR loop structures that have been shown to be present in the rest of the domains (Hamburger et al., 2004).
CHAPTER 5

Conclusions

The aim of the research presented here was to clone, express and purify recombinant domain mutants of human pIgR. The binding characteristics of the in vitro refolded domain mutants would be investigated in order to gain insight into the association of pIgR with IgM.

5.1 Expression and Purification of Recombinant Domain Mutants

High levels of expression were obtained for all domain mutants (Table 3.4). This is the first bacterial expression and in vitro refolding of unglycosylated recombinant multiple domain mutants of pIgR ectodomain. This work also resulted in the publication of the first in vitro refolded recombinant SC (Prinsloo et al., 2006). Previously, recombinant pIgR and SC had been successfully produced in HeLa cells (Crottet et al., 1999b) and insect cells (Rindisbacher et al., 1995). These methods are costly and require highly specialized equipment. Typically, binding analysis has been performed using purified native SC (Bakos et al., 1991) or recombinant expression in the mammalian MDCK cell line (Norderhaug et al., 1999). Exceptions to this rule was the first recombinant bacterial expression of domain I (Bakos et al., 1994).

This study showed recombinant SC domain mutants were expressed as inclusion bodies (Figures 3.8 and 3.13). Although recombinant proteins were fused with carboxy terminal His-tags (6x) that would have allowed direct purification from E. coli lysate, inclusion bodies were partially purified, by detergent washes and high speed centrifugation, in order to increase the level of purity prior to IMAC. Recombinant proteins were found not to be present, or at least not detectable, in any soluble fraction, even when expression was performed in the oxidizing mutant E. coli ORIGAMI (DE3). This was probably related directly to the complex disulfide arrangement pattern of the immunoglobulin-like domains present in the recombinant mutants. It should be noted, however, that inclusion body formation was favoured because of the potential for toxicity to the expression
strains by native conformations of recombinant proteins (Misawa and Kumagai, 1999). However, the ability of the recombinant mutants to be folded in vivo in an appropriate host strain requires further investigation; e.g. modulation of expression rates by controlling IPTG levels or temperature of growth medium. In vitro refolding by equilibrium gradient dialysis was successful, as observed by reducing and non-reducing SDS-PAGE (Figures 3.16 and 3.22). Epitope identification with goat anti-human SC confirmed that the recombinant mutants have structural similarity to the native human SC (Figure 3.25). MALDI-TOF MS analysis of lysylendopeptidase digested fragments of the recombinant domain mutants, unfortunately, did not resolve disulfide bond arrangement of the mutants. Further analyses of the disulfide bond arrangements of the recombinant mutants are needed; a similar approach could be followed with the addition of the sequencing of digested fragments, by selectively picking spots from 2D PAGE gels. This should allow exact elucidation of disulfide bond arrangement. Conformational analysis could also be performed using circular dichroism or NMR. Ultimately, X-ray crystallography or NMR structural resolution could be performed in order to elucidate the structure of the recombinant mutants and therefore SC (pIgR ectodomain). The X-ray structure of domain I would provide an important primer in the structural resolution of the complete protein (Hamburger et al., 2004). However, these analyses require high amounts of protein. Therefore, optimization of the refolding procedures are required. The equilibrium gradient refolding presented here allowed for the recovery of high amounts (17-22%) of the single domain mutants (hSC_rDI and hSCrDV); yields for the multiple domain mutants, however, never exceeded 10% (Table 3.4). This can be attributed to the high cysteine content of the mutants and requirement for disulfide bond formation. The yields may be improved by increasing the levels of L-Arg (or replacement with some other small molecule chaperone) or by more careful control of the introduction of oxidizing conditions. The potential for in vitro use of native chaperones cannot be excluded.

5.2 Binding Studies
Recombinant domain mutants of the polymeric immunoglobulin receptor produced in vitro possess the ability to interact with IgM with physiological and biological activity. The first in vitro production of unglycosylated SC in a bacterial expression system has
shown unequivocally that glycosylation is not required for immunoglobulin binding. Previous binding studies have used chemically or enzymatically deglycosylated SC to show the non requirement for glycosylation (Frutiger et al., 1988 and Bakos et al., 1991). Bakos et al. (1994) showed that in vitro refolded unglycosylated recombinant domain I retains biological activity toward IgA and IgM. The data presented here for recombinant domain I (hSC_rDI) confirm the published data and is shown to be true for all the multiple domain mutants.

Receptor-ligand ELISA studies (Figures 4.4-4.9 and Table 4.1) have shown that the recombinant domain mutants bind IgM with high and specific affinity. The data shows a defined increase in affinity for the multiple domain mutants compared to the single domain mutant. Although domain I is an absolute requirement for pIgM binding (confirmed by SPR analysis) and contributes greatly to total binding, it appears that the multiple domain mutant confers increased affinity; possibly by increasing the stability of the complex. The equilibrium conditions demonstrate the mutants’ specific binding to IgM and a general trend of increasing affinity, but fail to explain the exact kinetics of individual domain contribution.

As mentioned the SPR data confirmed the requirement of domain I in the interaction with IgM (Figure 4.10). The large disparity in the molecular weight of IgM and the recombinant domain mutants resulted in mass transfer affecting binding analysis when the mutants were immobilized. Immobilization of IgM appeared to negate mass transfer effects and allowed for kinetic analysis of mutant binding. Domain I, hSC_rDI, shows a considerably high affinity (3.16nM) with fast association and slow dissociation rates (Figures 4.21 and 4.22). This illustrates the large contribution of domain I toward whole molecule binding. Analysis of multiple domain mutant binding revealed contributions of the individual domains (II, III, IV and V) in observed differences in the relative association and dissociation rates upon interaction with IgM (Table 4.9). This pointed at structural determinants within these domains that could determine further interaction. It appears that domains I and II are involved in regulating association of the SC molecule with IgM, whereas domains IV-V may be involved in determining dissociation of the
complex. Concurrently, this could contribute to overall maintenance of a stable IgM/SC. Competitive binding experiments can be performed using purified native human SC and the recombinant domain mutants, in order to gain insight into specific domain requirements for optimal binding. Fitting of the SPR data to known models has highlighted the heterogeneous nature of IgM binding to SC, i.e. the requirement of the bivalent and heterogeneous ligand models to describe binding. Considering the data presented in Chapter 3 that N terminal processing may have occurred during expression, affecting N terminal sequencing and MALDI-TOF MS analysis, fitting the data to a heterogeneous analyte model resulted in complete failure of the fit. This showed that these proteins may have only been present in low levels and did not influence interaction to a significant degree. The heterogeneity of IgM as ligand is probably directly related to J chain content of IgM, i.e. the ability to associate with more than one J chain (Lindh and Bjork, 1976; Socken and Underdown, 1978). How do the domains interact with J chain and how does J chain content influence affinity? J chain is in fact one of the key driving forces behind translocation of polymeric immunoglobulins. In native systems J chain containing plgs are selectively transcytosed over non-J chain containing plgs (Brandtzaeg and Prydz, 1984; Hendrickson et al., 1995; Johansen et al., 2000). The recombinant domain mutants can be used to investigate the interaction with J chain more closely and also study the role J chain plays in potential multiple valency of IgM.

The competitive equilibrium binding assay established the ability of hSC_rDI, hSC_rDI-III and hSC_rDI-V to compete with native pIgR and shows that the affinities of the recombinant proteins are of high enough order to outcompete native pIgR, for IgM (Figures 4.34 - 4.37). The data pointed to a role for domains II-V in binding in that the equilibrium dissociation constant decreased significantly when all domains were present. Exactly what structural determinants (e.g. CDR loops) control this interaction and how conformation of the molecule enhances binding remain to be shown. Of particular interest would be the expression of the recombinant domain mutants in a non pIgR expressing mammalian cell, e.g. MDCK. Prior to this, however, the recombinant domain mutant clones have to be ligated to regions encoding the functional transmembrane domain and a cytoplasmic tail. Competitive binding analysis for IgM could then be
performed between, e.g. a domain I-III expressing mutant with a domain I-IV mutant as an antagonist. This would allow for defined description of binding effectors present in individual domains.

It is clear that the recombinant domain mutants produced in this study allow for an insight into the receptor-ligand interaction with IgM but several questions remain partially answered. What is the advantage of producing a multiple domain receptor for IgM? Of course it has a lot to do with the energy cost to the cell, bearing in mind that one copy of pIgR is used to transcytose either IgM or IgA oligomers. It has been shown with measurements using equilibrium based studies (receptor-ligand ELISA and competitive CELISA) that there does in fact appear to be an advantage to possessing multiple domains. In fact it results in an increase of an order of magnitude from the single domain mutant to the five domain mutant (Table 4.10). What role does conformation conferred by the domains play in affinity? This can only be answered using techniques that can monitor conformational change upon binding, e.g. circular dichroism spectroscopy and FTIR (Fourier transformed infrared spectrometry). The energetics of binding can also be measured (microcalorimetry) to give a better understanding of binding energies and hence binding contribution of individual domains. Although this can be performed with SPR measurements, the temperature needed to be altered and this was beyond the scope of the instrument used in this study. Binding studies may also be greatly enhanced by purifying recombinant proteins of high affinity by IgM-Sepharose affinity chromatography (Underdown et al., 1977).

5.3 Future Work and Applications of Recombinant SC mutants

The recombinant domain mutants produced in this study possess physiological and biological activity. The unglycosylated nature of the mutants does, however, prevent the recombinant proteins from possessing certain key immunological functions, e.g. bronchial mucus localization of SIgA for protection against *Shigella flexneri* (Phalipon et al., 2002). The ability of the recombinant domain mutants to bind IgM allow for further study of the interaction of individual domains with IgM and potentially IgA. The IgA binding ability of the mutants remains unknown, although due to the interaction of the mutants with IgM being comparable to that of native SC, it is expected that the mutants
should possess IgA binding ability. The interaction with J chain should also be investigated. Recombinant J chain could be used to perform binding interaction studies similar to those performed in this study, in order to clarify the exact interactions between J chain and pIgR ectodomain.

Conformational change analysis can be performed using circular dichroism to investigate potential conformation changes upon binding to a polymeric immunoglobulin. The single and multiple domain mutants can be used to effectively pinpoint domains contributing towards conformational change, if it occurs at all.

On a broader application scale it is known that the pneumococcal choline binding protein, cbpA, interacts with domains III and IV of SC/pIgR ectodomain. The interaction occurs via direct disulfide linkage and is independent of glycosylation (Lu et al., 2003). Therefore, the interaction of the recombinant domain mutants, hSC_rDI-IV and hSC_rDI-V, with cbpA should be investigated. The SC:cbpA interaction has been implicated in the homeostatic relationship between the host and the microbe, i.e. SC can effectively prevent S. pneumoniae invasion of respiratory epithelia. It has been suggested that S.pneumoniae invasion triggers a cytokine cascade allowing for increased pIgR expression. Increased pIgR molecules are produced and recycled, allowing S. pneumoniae cells on the apical/mucosal surface of respiratory epithelia to invade cells, thus, upsetting homeostasis (Zhang et al., 2000). The potential of the recombinant domain mutants hSC_rDI-IV or hSC_rDI-V as therapeutic agents in the clearing of S.pneumonia from infected individuals cannot be underestimated and requires further study, especially when considering the role pneumonia plays in the aetiology of AIDS. If it is established that the hSC_rD mutants can bind cbpA or S. pneumoniae directly, then the lego-like molecular biology techniques used to construct the recombinant domain mutants can be used to produce hSC_rDI-IV domain mutants to test the ability of the in vitro refolded mutants to interact with cbpA. If this is possible, then it would allow for the production of defined therapeutic agents against S. pneumoniae. The ability to construct fusion proteins could allow for the production of III-IV mutants fused directly
to engineered antibiotics. CbpA, is highly conserved and the interaction with SC III-IV, therefore, presents a definite target that can and must be exploited.

Human secretory component has also been shown to act as an endogenous inhibitor of phospholipase A\textsubscript{2}, which in turn regulates levels of prostaglandins, important mediators of labour during pregnancy. The presence of SC has therefore been implicated in the maintenance of pregnancy, preventing premature birth (Bennett \textit{et al.}, 1999). The interaction between SC and phospholipase A\textsubscript{2} is unknown. Whether this interaction requires glycosylation of SC remains to be investigated. The recombinant domain mutants can be used to probe the interaction with phospholipase A\textsubscript{2} in order to localize specific structural elements required for inhibition. Based on this, the recombinant domain mutants could be used in the development of potential therapeutic agents in difficult pregnancies.

In terms of rational drug design, the crystallization and structural elucidation of the recombinant domain mutants should allow for the design of mucosal targeted therapeutics. White and Capra (2002) have demonstrated that by using C\textalpha3 from IgA fused to green fluorescent protein, transcytosis does occur via pIgR. The data presented here suggest that domain I has a greater affinity for IgM than IgA (also shown by Røe \textit{et al.}, 1999). Competitive binding studies hint at interactions with domains II-V. This interaction needs to be resolved, e.g by X ray analysis for determination of the solution structure of recombinant domain mutants complexed with IgM. This will then allow for rational design of targeted therapeutics to the mucosa.
REFERENCES

- Arnoud Marquart, Personal Communication, 2005, Department of Vascular Medicine, Academic Medical Center, University of Amsterdam, Amsterdam. The Netherlands.


BIAcore (1997) *BiaEvaluation 3.0.* BIAcore AB.


Crottet P, Peitsch M.C., Servis C and Corthésy B (1999b) Covalent homodimers of murine secretory component induced by epitope substitution unravel the capacity of the polymeric Ig receptor to dimerize non-covalently in the absence of IgA ligand. J. Biol. Chem. 274: 31445-31455.


David Pugh, Personal Communication, 2003, Department of Biotechnology, University of the Western Cape, Cape Town, Republic of South Africa. dpugh@uwc.ac.za.


Further Reading

- Hancock K (2001) *Purification of (His)_6-tagged recombinant proteins expressed as inclusion bodies in E.coli using a Ni^{2+}-charged HiTrap Chelating HP column.* Life Science News 8: 14-16


APPENDIX A

DNA Sequence Alignments, Predicted Protein Sequence Alignments and Automated Sequencing Electropherograms
Figure A.1: DNA alignment of recombinant domain mutants with human pIgR. *, consensus; •, mismatch (Piskurich et al., 1995).
human pIgR
TCACCAACTTCCCCGAGGAGCACATTTTGTGGTGAACATTGCCCAGCTG
hSC_rDI
TCACCAACTTCCCCGAGGAGCACATTTTGTGGTGAACATTGCCCAGCTG
hSC_rDI-II
TCACCAACTTCCCCGAGGAGCACATTTTGTGGTGAACATTGCCCAGCTG
hSC_rDI-III
TCACCAACTTCCCCGAGGAGCACATTTTGTGGTGAACATTGCCCAGCTG
hSC_rDI-IV
TCACCAACTTCCCCGAGGAGCACATTTTGTGGTGAACATTGCCCAGCTG
hSC_rDI-V
TCACCAACTTCCCCGAGGAGCACATTTTGTGGTGAACATTGCCCAGCTG
pIgR_cDNA
TCACCAACTTCCCCGAGGAGCACATTTTGTGGTGAACATTGCCCAGCTG
Clustal Consensus
**************************************************
<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1010</td>
<td>human pIgR</td>
</tr>
<tr>
<td>1020</td>
<td>AGGAAGGCTGCTGCCCTATCCAGGCTGGCAACTCTTGGTCAAATGAGGAGTCC</td>
</tr>
<tr>
<td>1030</td>
<td>hSC_rDI</td>
</tr>
<tr>
<td>1040</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1050</td>
<td>hSC_rDI-II</td>
</tr>
<tr>
<td>1060</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1070</td>
<td>hSC_rDI-III</td>
</tr>
<tr>
<td>1080</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1090</td>
<td>hSC_rDI-IV</td>
</tr>
<tr>
<td>1100</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1110</td>
<td>pIgR_cDNA</td>
</tr>
<tr>
<td>1120</td>
<td>AGGAAGGCTGCTGCCCTATCCAGGCTGGCAACTCTTGGTCAAATGAGGAGTCC</td>
</tr>
<tr>
<td>1130</td>
<td>Clustal Consensus</td>
</tr>
<tr>
<td>1140</td>
<td>***************************<em><strong><strong><strong><strong>...</strong></strong></strong></strong></em></td>
</tr>
<tr>
<td>1150</td>
<td></td>
</tr>
<tr>
<td>1160</td>
<td>human pIgR</td>
</tr>
<tr>
<td>1170</td>
<td>AGGATTCCCCGCAGCCCCACTGTGGTGAAAGGAGGTTGACAGGAAGCTCTGT</td>
</tr>
<tr>
<td>1180</td>
<td>hSC_rDI</td>
</tr>
<tr>
<td>1190</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1200</td>
<td>hSC_rDI-II</td>
</tr>
<tr>
<td>1210</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1220</td>
<td>hSC_rDI-III</td>
</tr>
<tr>
<td>1230</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1240</td>
<td>hSC_rDI-IV</td>
</tr>
<tr>
<td>1250</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1260</td>
<td>pIgR_cDNA</td>
</tr>
<tr>
<td>1270</td>
<td>AGGATTCCCCGCAGCCCCACTGTGGTGAAAGGAGGTTGACAGGAAGCTCTGT</td>
</tr>
<tr>
<td>1280</td>
<td>Clustal Consensus</td>
</tr>
<tr>
<td>1290</td>
<td>***************************<em><strong><strong><strong><strong>...</strong></strong></strong></strong></em></td>
</tr>
<tr>
<td>1300</td>
<td></td>
</tr>
<tr>
<td>1310</td>
<td>human pIgR</td>
</tr>
<tr>
<td>1320</td>
<td>GGCCTGCTCTGCCCTAAGAGAAAGCCAAAGCTCAAGTACT</td>
</tr>
<tr>
<td>1330</td>
<td>hSC_rDI</td>
</tr>
<tr>
<td>1340</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1350</td>
<td>hSC_rDI-II</td>
</tr>
<tr>
<td>1360</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1370</td>
<td>hSC_rDI-III</td>
</tr>
<tr>
<td>1380</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1390</td>
<td>hSC_rDI-IV</td>
</tr>
<tr>
<td>1400</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1410</td>
<td>pIgR_cDNA</td>
</tr>
<tr>
<td>1420</td>
<td>GGCCTGCTCTGCCCTAAGAGAAAGCCAAAGCTCAAGTACT</td>
</tr>
<tr>
<td>1430</td>
<td>Clustal Consensus</td>
</tr>
<tr>
<td>1440</td>
<td>***************************<em><strong><strong><strong><strong>...</strong></strong></strong></strong></em></td>
</tr>
<tr>
<td>1450</td>
<td></td>
</tr>
<tr>
<td>1460</td>
<td>human pIgR</td>
</tr>
<tr>
<td>1470</td>
<td>GGTGTCTCTGGGAAGGGGCCCAGAATGGCCGCTGCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>1480</td>
<td>hSC_rDI</td>
</tr>
<tr>
<td>1490</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1500</td>
<td>hSC_rDI-II</td>
</tr>
<tr>
<td>1510</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1520</td>
<td>hSC_rDI-III</td>
</tr>
<tr>
<td>1530</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1540</td>
<td>hSC_rDI-IV</td>
</tr>
<tr>
<td>1550</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1560</td>
<td>pIgR_cDNA</td>
</tr>
<tr>
<td>1570</td>
<td>GGTGTCTCTGGGAAGGGGCCCAGAATGGCCGCTGCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>1580</td>
<td>Clustal Consensus</td>
</tr>
<tr>
<td>1590</td>
<td>***************************<em><strong><strong><strong><strong>...</strong></strong></strong></strong></em></td>
</tr>
<tr>
<td>1600</td>
<td></td>
</tr>
<tr>
<td>1610</td>
<td>human pIgR</td>
</tr>
<tr>
<td>1620</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>1630</td>
<td>hSC_rDI</td>
</tr>
<tr>
<td>1640</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1650</td>
<td>hSC_rDI-II</td>
</tr>
<tr>
<td>1660</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1670</td>
<td>hSC_rDI-III</td>
</tr>
<tr>
<td>1680</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1690</td>
<td>hSC_rDI-IV</td>
</tr>
<tr>
<td>1700</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1710</td>
<td>pIgR_cDNA</td>
</tr>
<tr>
<td>1720</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>1730</td>
<td>Clustal Consensus</td>
</tr>
<tr>
<td>1740</td>
<td>***************************<em><strong><strong><strong><strong>...</strong></strong></strong></strong></em></td>
</tr>
<tr>
<td>1750</td>
<td></td>
</tr>
<tr>
<td>1760</td>
<td>human pIgR</td>
</tr>
<tr>
<td>1770</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>1780</td>
<td>hSC_rDI</td>
</tr>
<tr>
<td>1790</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1800</td>
<td>hSC_rDI-II</td>
</tr>
<tr>
<td>1810</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1820</td>
<td>hSC_rDI-III</td>
</tr>
<tr>
<td>1830</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1840</td>
<td>hSC_rDI-IV</td>
</tr>
<tr>
<td>1850</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1860</td>
<td>pIgR_cDNA</td>
</tr>
<tr>
<td>1870</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>1880</td>
<td>Clustal Consensus</td>
</tr>
<tr>
<td>1890</td>
<td>***************************<em><strong><strong><strong><strong>...</strong></strong></strong></strong></em></td>
</tr>
<tr>
<td>1900</td>
<td></td>
</tr>
<tr>
<td>1910</td>
<td>human pIgR</td>
</tr>
<tr>
<td>1920</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>1930</td>
<td>hSC_rDI</td>
</tr>
<tr>
<td>1940</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1950</td>
<td>hSC_rDI-II</td>
</tr>
<tr>
<td>1960</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1970</td>
<td>hSC_rDI-III</td>
</tr>
<tr>
<td>1980</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>1990</td>
<td>hSC_rDI-IV</td>
</tr>
<tr>
<td>2000</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2010</td>
<td>pIgR_cDNA</td>
</tr>
<tr>
<td>2020</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>2030</td>
<td>Clustal Consensus</td>
</tr>
<tr>
<td>2040</td>
<td>***************************<em><strong><strong><strong><strong>...</strong></strong></strong></strong></em></td>
</tr>
<tr>
<td>2050</td>
<td></td>
</tr>
<tr>
<td>2060</td>
<td>human pIgR</td>
</tr>
<tr>
<td>2070</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>2080</td>
<td>hSC_rDI</td>
</tr>
<tr>
<td>2090</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2100</td>
<td>hSC_rDI-II</td>
</tr>
<tr>
<td>2110</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2120</td>
<td>hSC_rDI-III</td>
</tr>
<tr>
<td>2130</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2140</td>
<td>hSC_rDI-IV</td>
</tr>
<tr>
<td>2150</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2160</td>
<td>pIgR_cDNA</td>
</tr>
<tr>
<td>2170</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>2180</td>
<td>Clustal Consensus</td>
</tr>
<tr>
<td>2190</td>
<td>***************************<em><strong><strong><strong><strong>...</strong></strong></strong></strong></em></td>
</tr>
<tr>
<td>2200</td>
<td></td>
</tr>
<tr>
<td>2210</td>
<td>human pIgR</td>
</tr>
<tr>
<td>2220</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>2230</td>
<td>hSC_rDI</td>
</tr>
<tr>
<td>2240</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2250</td>
<td>hSC_rDI-II</td>
</tr>
<tr>
<td>2260</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2270</td>
<td>hSC_rDI-III</td>
</tr>
<tr>
<td>2280</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2290</td>
<td>hSC_rDI-IV</td>
</tr>
<tr>
<td>2300</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2310</td>
<td>pIgR_cDNA</td>
</tr>
<tr>
<td>2320</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>2330</td>
<td>Clustal Consensus</td>
</tr>
<tr>
<td>2340</td>
<td>***************************<em><strong><strong><strong><strong>...</strong></strong></strong></strong></em></td>
</tr>
<tr>
<td>2350</td>
<td></td>
</tr>
<tr>
<td>2360</td>
<td>human pIgR</td>
</tr>
<tr>
<td>2370</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>2380</td>
<td>hSC_rDI</td>
</tr>
<tr>
<td>2390</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2400</td>
<td>hSC_rDI-II</td>
</tr>
<tr>
<td>2410</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2420</td>
<td>hSC_rDI-III</td>
</tr>
<tr>
<td>2430</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2440</td>
<td>hSC_rDI-IV</td>
</tr>
<tr>
<td>2450</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>2460</td>
<td>pIgR_cDNA</td>
</tr>
<tr>
<td>2470</td>
<td>AGCGAGGGGTGGGTTAAGGCCCAGATGACAGGAGGCTCCCTGGTGGTGGAC</td>
</tr>
<tr>
<td>2480</td>
<td>Clustal Consensus</td>
</tr>
<tr>
<td>2490</td>
<td>***************************<em><strong><strong><strong><strong>...</strong></strong></strong></strong></em></td>
</tr>
</tbody>
</table>
Human pIgR: 1260

hSC_rDI: 1270

hSC_rDI-II: 1280

hSC_rDI-III: 1290

hSC_rDI-IV: 1300

hSC_rDI-V: 1310

pIgR_cDNA: 1320

Clustal Consensus: 1330

1340

1350

1360

1370

1380

1390

1400

Human pIgR: 1410

hSC_rDI: 1420

hSC_rDI-II: 1430

hSC_rDI-III: 1440

hSC_rDI-IV: 1450

pIgR_cDNA: 1460

Clustal Consensus: 1470

1480

1490

1500

Human pIgR: 1510

hSC_rDI: 1520

hSC_rDI-II: 1530

hSC_rDI-III: 1540

hSC_rDI-IV: 1550

pIgR_cDNA: 1560

Clustal Consensus: 1570

1580

1590

1600

Human pIgR: 1610

hSC_rDI: 1620

hSC_rDI-II: 1630

hSC_rDI-III: 1640

hSC_rDI-IV: 1650

pIgR_cDNA: 1660

Clustal Consensus: 1670

1680

1690

1700
<table>
<thead>
<tr>
<th>Human pIgR</th>
<th>Clustal Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGAAGGTGGCTAGACTCTGGTTTTCGGGAGATTGAGAACAAAGCCATTCAG</td>
<td></td>
</tr>
<tr>
<td>Clustal Consensus</td>
<td>**************************************************</td>
</tr>
<tr>
<td>AGAAGGTGGCTAGACTCTGGTTTTCGGGAGATTGAGAACAAAGCCATTCAG</td>
<td></td>
</tr>
<tr>
<td>Clustal Consensus</td>
<td>**************************************************</td>
</tr>
<tr>
<td>AGCCGATGGGAGCAGAGCATCTGTGGATTCCGGCAGCTCTGAGGAACAAG</td>
<td></td>
</tr>
<tr>
<td>Clustal Consensus</td>
<td>**************************************************</td>
</tr>
<tr>
<td>GTGGAAGCTCCAGAGCGCTGGTCTCCACCCTGGTGCCCCTGGGCCTGGTG</td>
<td></td>
</tr>
<tr>
<td>Clustal Consensus</td>
<td>**************************************************</td>
</tr>
<tr>
<td>CTGGCAGTGGGAGCCGTGGCTGTGGGGGTGGCCAGAGCCCGACAGGAA</td>
<td></td>
</tr>
</tbody>
</table>

Clustal Consensus: "*****..*...**."
human pIgR
AGCACCGTGCCGCCGAGGCCCAGGACGGCCCCAGGAAGCTAG
hSC_rDI
-----------------------------------------------------
hSC_rDI-II
-----------------------------------------------------
hSC_rDI-III
-----------------------------------------------------
hSC_rDI-IV
-----------------------------------------------------
hSC_rDI-V
-----------------------------------------------------
pIgR_cDNA
-----------------------------------------------------
Clustal Consensus
Figure A.2: Predicted protein sequence alignment of recombinant domain mutants with human pIgR. *, consensus; •, mismatch (Piskurich et al., 1995).
<table>
<thead>
<tr>
<th></th>
<th>human</th>
<th>hSC_rDI</th>
<th>hSC_rDI-II</th>
<th>hSC_rDI-III</th>
<th>hSC_rDI-IV</th>
<th>hSC_rDI-V</th>
<th>pIgR_cDNA</th>
<th>Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>260</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>290</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>310</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>320</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>330</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>340</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>350</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>360</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>370</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>380</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>390</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>410</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>420</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>430</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>440</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>450</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>460</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>470</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>480</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>490</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>760</td>
<td>....</td>
<td>....</td>
<td>....</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>human</td>
<td>STVAEADGPQEA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hSC_rDI</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-II</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-III</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-IV</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hSC_rDI-V</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pIgR_cDNA</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consensus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table A.1: Interpretation of single letter codes used for editing electropherograms.

<table>
<thead>
<tr>
<th>Code</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Not A</td>
</tr>
<tr>
<td>D</td>
<td>Not C</td>
</tr>
<tr>
<td>H</td>
<td>Not G</td>
</tr>
<tr>
<td>K</td>
<td>G or T</td>
</tr>
<tr>
<td>M</td>
<td>A or C</td>
</tr>
<tr>
<td>N</td>
<td>G, A, T, C</td>
</tr>
<tr>
<td>R</td>
<td>A or G</td>
</tr>
<tr>
<td>S</td>
<td>C or G</td>
</tr>
<tr>
<td>V</td>
<td>Not T</td>
</tr>
<tr>
<td>W</td>
<td>A or T</td>
</tr>
<tr>
<td>Y</td>
<td>C or T</td>
</tr>
</tbody>
</table>
Figure A.3: Electropherogram representing the sequencing of pET22b hSC_rDI using T7 promoter primer.
Figure A.4: Electropherogram representing the sequencing of pET22b hSC_rDI using T7 terminator primer.
Figure A.5: Electropherogram representing the sequencing of pET22b hSC_rDI-II using T7 promoter primer.
Figure A.6: Electropherogram representing the sequencing of pET22b hSC_rDI-II using T7 terminator primer.
Figure A.7: Electropherogram representing the sequencing of pET22b hSC_rDI-III using T7 promoter primer.
Figure A.8: Electropherogram representing the sequencing of pET22b hSC_rDI-III using T7 terminator primer.
Figure A.9: Electropherogram representing the sequencing of pET22b hSC_rDI-III using domain 2 3’ border region reverse primer
Figure A.10: Electropherogram representing the sequencing of pET22b hSC_rDI-IV using T7 promoter primer.
Figure A.11: Electropherogram representing the sequencing of pET22b hSC_rDI-IV using T7 terminator primer.
Figure A.12: Electropherogram representing the sequencing of pET22b hSC_rDI-IV using the internal forward domain III 5’ border region primer.
Figure A.13: Electropherogram representing the sequencing of pET22b hSC_rDI-V using T7 promoter primer.
Figure A.14: Electropherogram representing the sequencing of pET22b hSC_rDI-V using T7 terminator primer.
Figure A.15: Electropherogram representing the sequencing of pET2b hSC_rDI-V using the internal forward primer representing the domain V 5’ border region.
Figure A.16: Electropherogram representing the sequencing of pET22b hSC_rDI-V using an internal reverse primer.
Figure A.17: Schematic diagram of the structure of the mouse pIgR gene. E1 – E11 represent exons. The introns are represented by horizontal lines. The grey hatched boxes represent 5' and 3' untranslated regions. The scale is in kilobases. (from Martin et al., 1997)
APPENDIX B

MALDI TOF Spectra
Figure B.1: MALDI TOF MS spectrum of lysylendopeptidase digestion of hSC_rDI-V.
Figure B.2: MALDI TOF MS spectrum of lysylendopeptidase digestion of hSC_rDI-IV.
Figure B.3: MALDI TOF MS spectrum of lysylendopeptidase digestion of hSC_rDI-III.
Figure B.4: MALDI TOF MS spectrum of lysylendopeptidase digestion of hSC_rDI-II.
Figure B.5: MALDI TOF MS spectrum of lysylendopeptidase digestion of hSC_rDL.
Figure B.6: MALDI TOF MS spectrum of lysylendopeptidase digestion of BSA.
APPENDIX C

Publications Resulting from the Current Study