Structural Characterisation of Antibody/Antigen Interactions: Implications for B-Cell Receptor Signalling

Philip William Addis

BSc(Hons) (Leicester)

Thesis submitted for the Degree of Doctor of Philosophy

Department of Biochemistry

University of Leicester

January 2013

Abstract

Structural Characterisation of Antibody/Antigen Interactions:

Implications for B-Cell Receptor Signalling

Philip William Addis

Therapeutic antibodies are an important class of pharmaceutical molecules contributing a significant portion of the total therapeutic market. Due to their high specificity, modular nature and low toxicity, these proteins provide great potential for the development of novel therapeutics for a wide range of diseases. The continued development of improved and unique therapeutics will benefit from improvements in our understanding of the characteristics of these proteins and their functions, facilitating a more informed approach towards therapeutic research and potentially highlighting new molecular information on poorly understood areas of antibody function, such as the initiation of BCR signalling. NMR provides a useful and highly sensitive tool for the investigation of the structural properties of proteins and the changes which occur upon binding. Smaller antibody fragments also provide a useful model for the study of antibody/antigen interactions in solution. The work described in this thesis describes the development of the scFv format of antibody fragment as an amenable and reliable model system from which high quality NMR data can be collected to investigate both the behaviour of the antigen recognising CDR loops and the movements occurring at the interface between the variable domains. An efficient, high yield protein expression method was adapted and developed for the expression of large quantities of ^{15}N/^{13}C/^{2}H labelled scFv which allowed, for the first time, sequence specific backbone assignments for the majority of the scFv in both the free and bound states and the comparison of chemical shift data. Further to this, based upon the domain movements observed, potential mechanisms for the initiation of BCR signalling were investigated resulting in data suggesting the first antigen modulated molecular basis for the coupling of binding from a structurally diverse antigen repertoire to a conserved signalling response.

Acknowledgements

Table of Contents

Abbreviations

2D	Two dimensional
3D	Three dimensional
A	Amperes
A_{600}	Absorbance at 600 nm
ADCC	Antibody dependant cellular cytotoxicity
AEBSF	4-(2-Aminomethyl) benzenesulphonyl fluoride
APC	Antigen presenting cell
Au	Absorbance units
B-cell	B-lymphocyte
BCR	B-cell receptor
BLNK	B-cell linker
BSF-2	B-cell stimulatory factor 2
Btk	Brutons tyrosine kinase
˚C	Degrees centigrade
C_H	Heavy chain constant domain
C_L	Light chain constant domain

CBM	Cytokine-binding motif
CD	Circular dichroism
CDC	Complement dependant cytotoxicity
CDR	Complementarity determining region
CLSM	Confocal laser scanning microscopy
CV	Column volume
Da	Daltons
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ERK	Extracellular signal-related kinase
ERM	Ezrin-radixin-moesin
Fab	Antibody antigen binding fragment
Fc	Antibody crystallisable fragment
FDA	Food and drug administration
FRET	Fluorescence resonance energy transfer
FPLC	Fast protein liquid chromatography
Fv	Antibody variable fragment

g	Gram
GC	Germinal centre (B-cell)
HAMA	Human anti-mouse antibody
His$_6$ tag	Hexa-histidine tag
HSQC	Heteronuclear single quantum coherence
HRP	Horseradish peroxidase
Hz	Hertz
Ig	Immunoglobulin
Ig$_{NAR}$	Shark new antigen receptor antibody
IL	Interleukin
IL-6R	Interleukin-6 receptor
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITAM	Immuno-tyrosine activation motif
IVIG	Intravenous immune globulin
JAK	Janus kinase
JNK	Jun N-terminal protein kinase
K$_D$	Dissociation constant
l	Litre

LB	Luria Bertani medium
LDS	Lithium dodecyl sulphate
m	Metres
M	Molar
MAPK	Mitogen activated protein kinase
MHC	Major histocompatibility complex
mIg	Membrane immunoglobulin
MWCO	Molecular weight coefficient
NEM	N-ethylmaleimide
Ni-NTA	Nickel nitrilotriacetic acid
NF-κB	Nuclear factor kappa B
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol

PI3K	Phosphinositide 3 kinase
PKC	Protein Kinase C
PLCγ2	Phospholipase-C-gamma-2
PMSF	Phenylmethylsuphonyl fluoride
ppm	Parts per million
PVDF	Polyvinylidine difluoride
RDC	Residual dipolar coupling
rpm	Revolutions per minute
s	Seconds
scFv	Single chain antibody variable fragment
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SHP-2	SH2 containing tyrosine specific protein phosphatase
sIL-6R	Soluble interleukin-6 receptor
SLAM	Selected lymphocyte activation method
SPT	Single particle tracking
STAT	Signal transducer and activator of transcription
T	Tesla

T_0	Time point (with number of hours)
T_2	Spin-spin relaxation time
T_H2	T helper-2 cell
T_H17	T helper-17 cell
T-cell	T-lymphocyte
TEV	Tobacco etch virus
TIRFM	Total internal reflection fluorescence microscopy
TNF-α	Tumour necrosis factor alpha
T_{REG}	Regulatory T-cell
Tris	Tris(hydroxymethyl)aminomethane
TROSY	Transverse relaxation optimised spectroscopy
US	United States
US$	United States dollars
UV	Ultra violet
V	Volts
v/v	Volume per volume
V_H	Heavy chain variable domain
V_{HH}	Variable domain from camelid heavy chain antibody

V_L	Light chain variable domain
V_{NAR}	Variable domain from shark new antigen receptor antibody
V(D)J	Variable (diversity) and joining gene segments
w/v	Weight per volume
WATERGATE	Water suppression through gradient tailored excitation
x g	Times gravity
$\Delta\delta$	Change in chemical shift

Amino Acid One Letter Codes and Abbreviations

Alanine	A	Ala	Leucine	L	Leu
Arginine	R	Arg	Lysine	K	Lys
Asparagine	N	Asn	Methionine	M	Met
Aspartic acid	D	Asp	Phenylalanine	F	Phe
Cysteine	C	Cys	Proline	P	Pro
Glutamine	Q	Gln	Serine	S	Ser
Glutamic acid	E	Glu	Threonine	T	Thr
Glycine	G	Gly	Tryptophan	W	Trp
Histidine	H	His	Tyrosine	Y	Tyr
Isoleucine	I	Ile	Valine	V	Val

Chapter 1 – General Introduction

1.1 – Antibodies

Antibodies are the primary effector molecules of the adaptive immune system. They are glycoproteins able to perform a number of roles that both directly immobilise antigen and aid the innate immune response, by promoting responses such as phagocytosis, lymphocyte localisation, antibody-dependent cellular cytotoxicity and complement activation (Schroeder & Cavacini, 2010). The antibody molecule is a dislulphide linked heterotetramer consisting of two heavy and two light chains (Carter, 2006). These chains contain a number of immunoglobulin domains (Ig), one variable (V) domain and one (light chain) or three to four (heavy chain) constant (C) domains (Davies and Chacko, 1993; Schroeder & Cavacini, 2010). The general structural arrangement of an antibody is outlined in Figure 1.1.1.

These Ig domains are approximately 12 to 13 kDa in size (Schroeder & Cavacini, 2010) and comprise of a structural motif known as an immunoglobulin fold. This structure is formed from a sandwich of two β-sheets with a number of extended flexible loops fixed together with a conserved disulphide bond (Poljak *et al,* 1973; Davies & Chacko, 1993). Ig domains are the building blocks of a large range of immunoglobulin superfamily proteins (including class-I and II major histocompatibility complex proteins, cofactors such as CD4 and CD8 and cytokine receptors such as those for Interleukin-6, amongst others) and represent a common and important class of protein structural motif. The overall structure of an antibody, as shown in Figure 1.1.1, can be divided into three main portions with two distinct functional roles (Porter, 1950; Porter, 1959). These

regions, known as the antibody binding fragments (Fab) and the crystallisable fragment (Fc), allow for antigen recognition (Fab) and the effector functions (Fc) of the molecule.
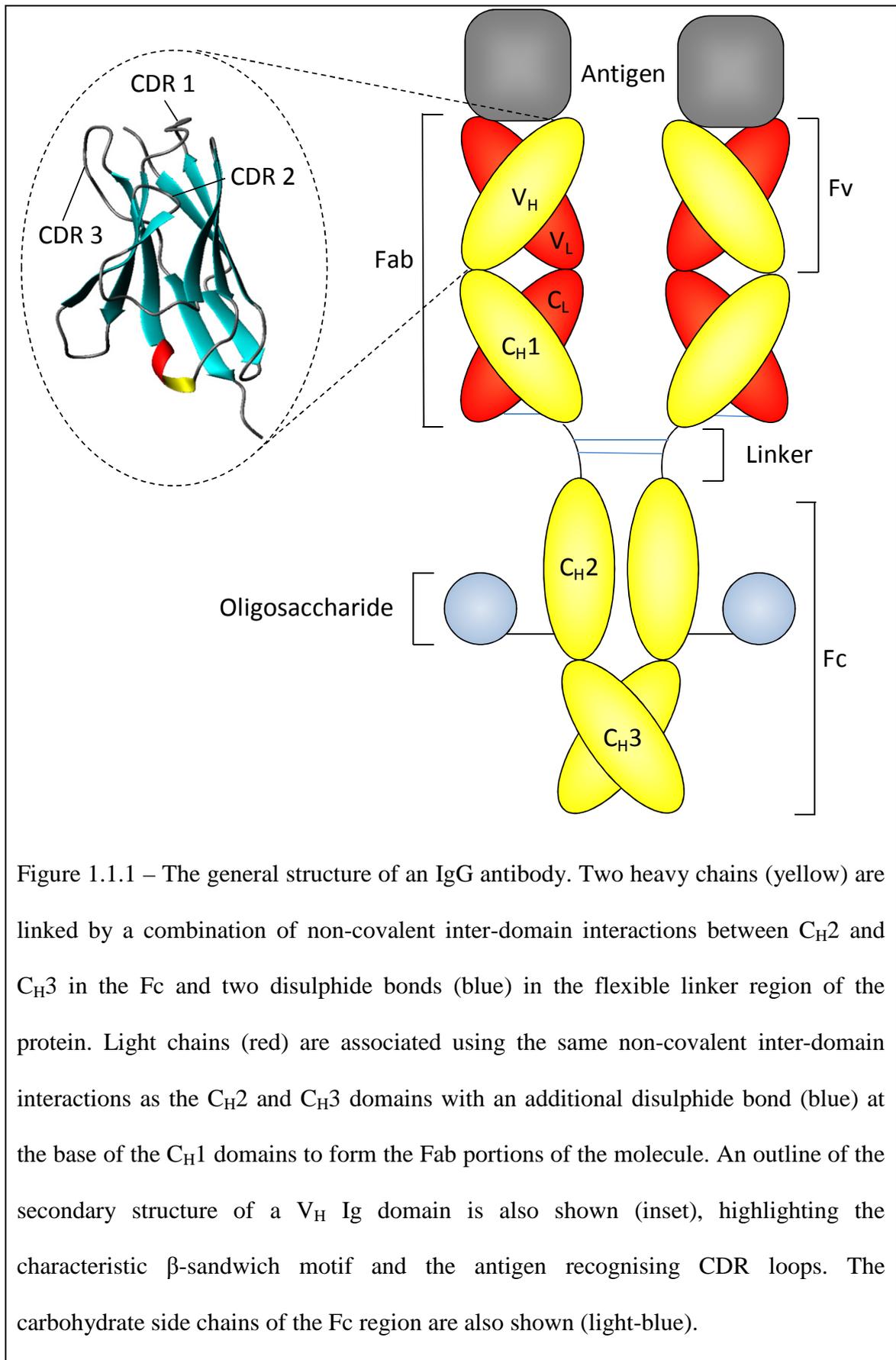
Figure 1.1.1 – The general structure of an IgG antibody. Two heavy chains (yellow) are linked by a combination of non-covalent inter-domain interactions between C_H2 and C_H3 in the Fc and two disulphide bonds (blue) in the flexible linker region of the protein. Light chains (red) are associated using the same non-covalent inter-domain interactions as the C_H2 and C_H3 domains with an additional disulphide bond (blue) at the base of the C_H1 domains to form the Fab portions of the molecule. An outline of the secondary structure of a V_H Ig domain is also shown (inset), highlighting the characteristic β-sandwich motif and the antigen recognising CDR loops. The carbohydrate side chains of the Fc region are also shown (light-blue).

The variability of the V domains provides a wide array of antigen recognising surfaces and is introduced through a complex set of genetic rearrangements throughout the development of the antibody producing B-lymphocytes. These rearrangements primarily affect the antigen binding complementarity determining region (CDR) loops. Both the heavy and light chains are encoded by separate multi-gene families in which the V and C domains are in turn encoded by independent elements. The V domains are encoded by the recombination of variable, diversity (V_H only) and joining (V(D)J) gene segments that are rearranged to form the genetic structure of the domain, as overviewed in Figure 1.1.2 (Oka & Kawaichi, 1995; Schroeder & Cavacini, 2010). For each variable domain locus (V_H, V_κ or V_λ), a number of these elements are imprecisely combined at the region coding for CDR3 using the non-homologous end joining DNA repair machinery in combination with a number of other processes. These include the inversion or deletion of the D gene degments (V_H only), the loss of nucleotides during the recombination process or the addition of a number of random nucleotides by terminal deoxynucleotidyl transferase (Tonegawa, 1983; Oka & Kawaichi, 1995; Schroeder & Cavacini, 2010). The combination of these effects, and the use of two unique V domains in the final structure, provides for a potential pre-immune antibody repertoire of over 10^{16} different specificities. Further to this, the fully recombined antibody is able to undergo the process of affinity maturation when exposed to antigen in combination with interactions from T-helper cells to produce higher affinity antibodies by somatic hyper-mutation (Diaz & Casali, 2002). This process introduces mutations in the antibody genes at a rate of 1 base in 1000 per cell cycle, both by targeting mutation hot spots and by error prone DNA synthesis. The precise order of these processes is well reported and ties in closely with the developmental stages of the

18

B-lymphocyte, the location of the cells within the body and the external signals they receive (Schroeder & Cavacini, 2010; Wang & Clark 2003).
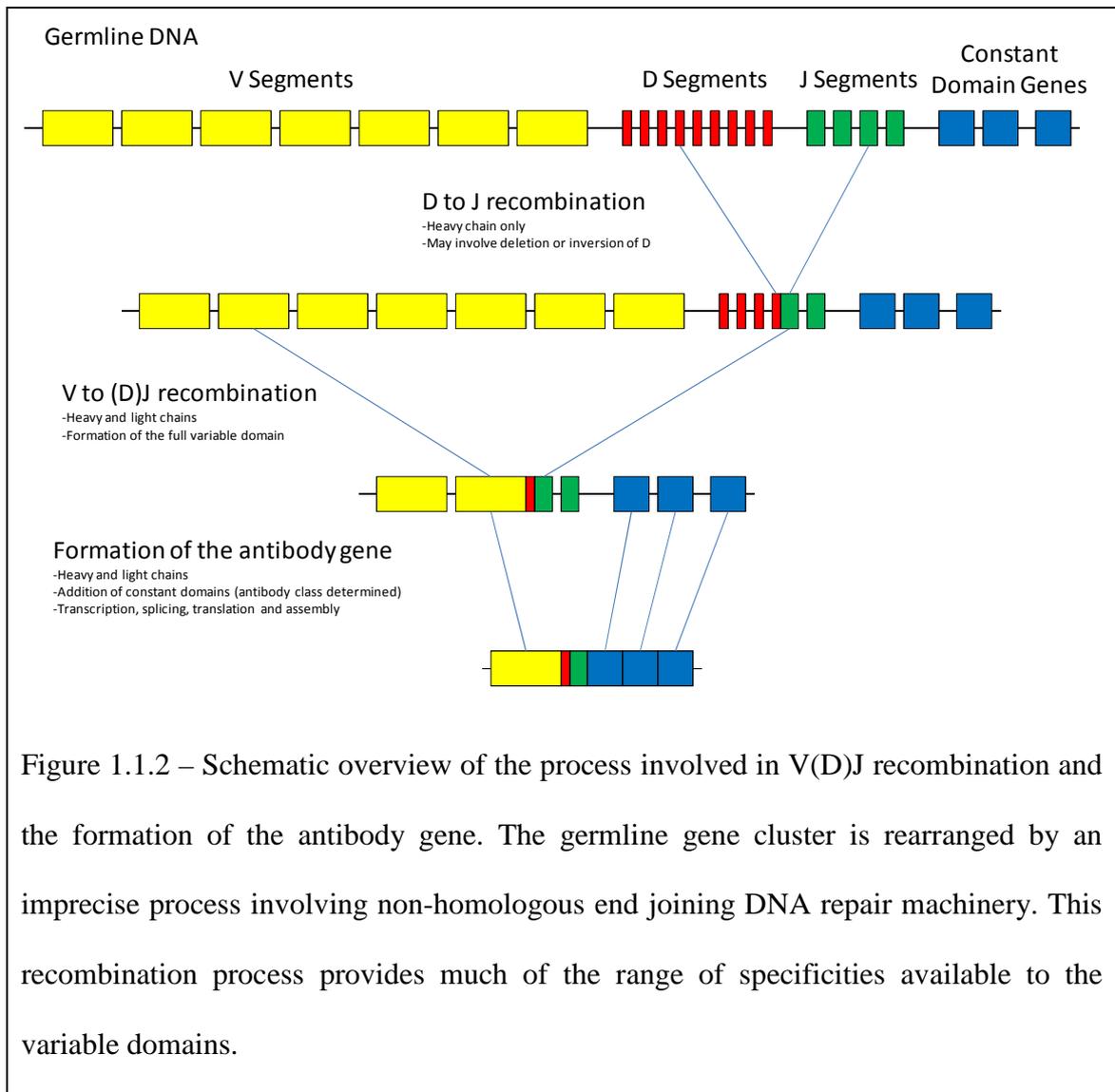
Figure 1.1.2 – Schematic overview of the process involved in V(D)J recombination and the formation of the antibody gene. The germline gene cluster is rearranged by an imprecise process involving non-homologous end joining DNA repair machinery. This recombination process provides much of the range of specificities available to the variable domains.

The C domains for the light chains are determined by which germ line gene cluster was used for the recombination of the V domain (either the λ or κ gene cluster). The C_H domains, however, are subject to recombination events, although on a much less complex scale to the V domains. This process does not introduce mutations into the sequence, but instead selects the appropriate class of C domain depending upon the

external signals received and the local environment of the B-cell (Honjo, 1983; Wabl & Steinberg, 1996). There are nine functional genes encoding for C domains downstream of the V(D)J regions in the heavy chain gene cluster consisting of a series of exons encoding for domains, hinges or termini. All of these C domains confer different properties and effector functions and may be alternatively spliced to include a membrane anchoring terminus, which is essential for forming the B-cell receptor (BCR) complex, or a terminus that allows for the secretion of soluble antibody from the cell (Schroeder & Cavacini, 2010). This process uses recombination via switching regions within the genes and is known as class-switch recombination. The V_H domains formed by V(D)J recombination may be joined to any of the nine different heavy chains (Wabl & Steinberg, 1996).

The different heavy chain genes are arranged into five classes based upon structure and function (Igμ, Igα, Igδ, Igε or Igγ). Class-switch recombination into one of these five classes provides for the number of different roles required of the antibody molecule and the differing environments in which they may be found. These antibody classes, known as IgM, IgA, IgD, IgE or IgG, all differ in role, localisation, valency, affinity and flexibility. Subclasses within IgA and IgG further extend these differences whilst conserving other properties related to those classes. The overall structure of the antibodies, and stages of B-cell development at which each class is produced, is summarised in Figures 1.1.3 and 1.1.4 respectively. Throughout the development process, the surface expressed antibodies, in the form of BCRs, are critical for the progression and regulation of the B-cell maturation process. This will be further discussed in chapter 4.

Figure 1.1.3 – The general structure of the five classes of human antibodies. The antibodies all consist of light chains (red) and heavy chains (yellow). The heavy chains contain C domains specific to the antibody class determining its structural characteristics and functional properties. Disulphide bonds are shown in blue. The antibodies may also be presented on the surface of a B-cell in their monomeric form at various stages of development. This is done by the inclusion of a membrane spanning tail during genetic the genetic rearrangement of the heavy chain (not shown).

Figure 1.1.4 – A basic overview of the B-cell development process. The B-cells are developed using the recombination and mutation events shown through several stages to produce an immune response. This progression is dependent on antigen recognition through the BCR and T-cell help (not shown) in addition to other factors that facilitate the correct development of the cells. The antibody class at each stage is shown indicating that IgM and IgD are important in the development process. Once class switching has occurred, antigen recognition through the BCR will trigger antibody production and further differentiation into memory or plasma cells (adapted from Wang & Clark, 2003; Kurosaki *et al*, 2009; Schroeder & Cavacini 2010). IgE and IgD may also be secreted in to the circulation (not shown) but make up only a small percentage of the total circulating antibody content.

IgM is the first antibody expressed during B-cell development and is associated with the primary immune response. Upon antigenic stimulation and maturation of the B-cell, soluble, typically pentameric, antibody is produced. The IgM antibody contains four C domains and is linked to other IgM monomers by disulphide bonds. A J-chain is also linked to two of the monomers in the pentameric form facilitating secretion at mucosal surfaces. Due to the relative immaturity of the IgM antibody, it tends to have a weaker affinity than other antibody classes and may be more polyreactive. However, due to the nature of the soluble form, it tends to have a much higher avidity and is able to opsonise antigen and fix complement highly efficiently. IgA is the only other antibody class that is able to associate with a J-chain and is the primary antibody found at mucosal surfaces. The secretory form of IgA is a dimer and tends to consist of the IgA2 subclass which contains shorter hinge regions and is less sensitive to proteolysis. The other IgA subclass, IgA1, is primarily found in the serum. IgD is the second least common antibody class found within the body and relatively poorly understood. It has a high sensitivity to proteolysis and has an undetermined function in the low levels in which it is found in the serum. The membrane bound form of the molecule has been the subject of more investigation and is found mainly co-expressed with IgM on B-cells that are leaving the bone marrow and populating secondary lymphoid organs. Although its role in this situation has not been clearly determined, it has been suggested that it may regulate B-cell fate through changes in its activation status. IgE is the least common antibody class found in the serum but is well studied due to its role in allergy and hypersensitivity which is closely related to its very high affinity for its Fc receptor. This class of antibody molecule is also involved in the immune response to parasitic worm infections. The final, and by far the most common and most extensively studied serum antibody, is IgG. This antibody, primarily associated with the secondary immune

response from re-exposure to antigen, is one of the body's main defences against invading pathogens. IgG is high affinity and is able to opsonise and recruit complement, although to a lesser extent than IgM. There are four subclasses of IgG, each with differing structural characteristics and roles. IgG3 has the most flexible hinge region followed by IgG1. Both of these subclasses are mainly associated with responses to protein antigens. IgG2 and IgG4 are less flexible and tend to be associated with responses to polysaccharide antigens. IgG4 is the only IgG which is unable to opsonise antigen or activate complement. In all cases the effector functions of the Fc are mediated by a number of factors of which the most prevalent is binding to the appropriate Fc receptor. This binding is able to activate numerous processes such as antigen-dependant cell cytotoxicity, phagocytosis and antibody recycling, amongst others, depending on heavy chain class and the subclass of the receptors. The Fcs of IgM and IgG1-3 are able to recruit the C1q complement component to activate the classical complement pathway.

The antibody molecule, and the process by which it is formed, provides a versatile and adaptable tool for combating the wide range of potential pathogens that face the immune system. It is this adaptability and specificity that has led to much research into the adaptation of antibodies for a number of purposes.

1.2 – Therapeutic Antibodies

Despite a focus on antibodies for use in modern biology and medicine since the early 20th century, the development of therapeutic antibodies has been, until quite recently, relatively slow. The initial work by Robert Koch, Kitasato Shibasaburo, Emil von

Behring and Paul Ehrlich at the turn of the 20th century, using serum from infected patients to treat infectious diseases, identified that there must be active components within the bloodstream that were able to combat disease and constituted the first use of antibodies in therapy (Winau *et al,* 2004). As time progressed, this serum therapy was refined by isolating antibodies for intravenous immune globulin (IVIG) therapy (Stangel & Pul, 2006). These early steps into the use of antibodies for therapeutics were met with relative success, but it was not until the development of monoclonal antibodies from mouse hybridoma cells (Köhler & Milstein, 1975) that significant advances were able to be made. This new technology allowed the development of antibodies for use in a wide variety of applications, including their use as therapeutic agents. The first therapeutic antibody, the anti-CD3 muromomab, was approved by the United States (US) Food and Drug Administration (FDA) in 1986 for the treatment of acute organ rejection. Since this time, a number of advancements have been made producing treatments with high specificity, avidity and safety. As of 2010, there were 28 antibodies, or antibody derived, therapeutics approved for use by the FDA, with at least nine of these treatments producing sales of over 1 billion US$ (Reichert, 2010). These treatments now constitute a large proportion of total therapeutic sales and, with similar development times and higher success rates compared to small molecule treatments, look set to continue to be a major factor in medicine for some time to come.

Despite the initial success of murine monoclonal antibodies as therapeutics, it was soon clear that a number of obstacles must be overcome. One of the primary problems faced was the immunogenic response that can be developed when injecting non-human proteins into the body. This effect, known as the human anti-mouse antibody (HAMA) response, can lead to a much reduced efficacy of treatment as the body combats the injected therapeutic and promotes rapid clearance from the system. In severe cases this

may also result in side effects such as anaphylactic shock, which may be life threatening. Muronab-CD3, for example, elicited blocking antibodies in roughly 50 % of patients (Norman *et al,* 1993). The human immune system also fails to recognise mouse Fc and produce any antibody mediated responses which, depending on the therapeutic target, would be beneficial. These features of the original treatments led to the development of a number of improvement strategies. Initially, chimeric antibodies were produced (Boulianne *et al,* 1984; Morrison *et al*, 1984), followed by CDR grafted humanised antibodies (Jones *et al,* 1986; Kettleborough *et al*, 1991). More recently, fully human antibodies have been produced using technologies such as phage display (McCafferty *et al*, 1990; Vaughan *et al*, 1996), ribosome display (Hanes *et al*, 1998) and transgenic mice (Russel *et al*, 2000; Lonberg, 2005).

Chimeric antibodies consist of mouse variable domains grafted onto human constant domains (Figure 1.2.1). This approach retains the binding specificity of the murine derived antibody, but significantly reduces the immunogenicity of the molecule. It also provides Fc related effector functions which may be desirable to further boost the effectiveness of the treatment. This process was made possible by the development of molecular biology techniques and provides a relatively straightforward procedure to produce the final antibody molecule. The first therapeutic of this type was abciximab, a chimeric anti-GPIIb/IIIa Fab fragment, that produces a blocking reaction in only 6 % of patients (Faulds & Sorkin, 1994). Humanised antibodies are again recombinant proteins developed as a result of advancements in molecular biology. The CDR loops of a murine monoclonal antibody are grafted onto a human antibody framework (Figure 1.2.2) to create an almost fully human antibody. This process, however, has two main drawbacks. The first is that the molecular biology methods required to produce such an antibody are significantly more complex than those required to produce a chimeric

antibody. The second is a potential reduction in antigen binding affinity due to the human framework residues failing to maintain the appropriate conformation of the mouse CDRs. This must be addressed either by further mutations in the framework residues or by the transfer of some of the mouse framework residues along with the CDRs (Al-Lazikani *et al*, 1997; Chothia and Lesk, 1987; Reichmann *et al*, 1988; Tramontano *et al*, 1990; Verhoeyen *et al*, 1993). Despite the potential drawbacks of producing humanised antibodies, a significant number have been developed and are approved for use. The first antibody of this type was daclizumab (anti-CD25), approved in 1997 for the treatment of kidney transplant rejection (Vincenti *et al,* 1998). In general the humanised antibodies again elicit a reduced blocking response that is generally less than that observed for chimeric antibodies even after repeated doses. The most recent produce immunogenic responses in less than 1 % of patients (Yoon *et al,* 2010).

Figure 1.2.2 – Chimeric and humanised IgG antibodies. The chimeric antibody consists of human (red/yellow) C domains with mouse (purple/green) Fvs grafted onto the antibody structure at the DNA level. In a similar manner, humanised antibodies consist of human C and V domains with mouse derived CDR loops grafted onto the human antibody framework.
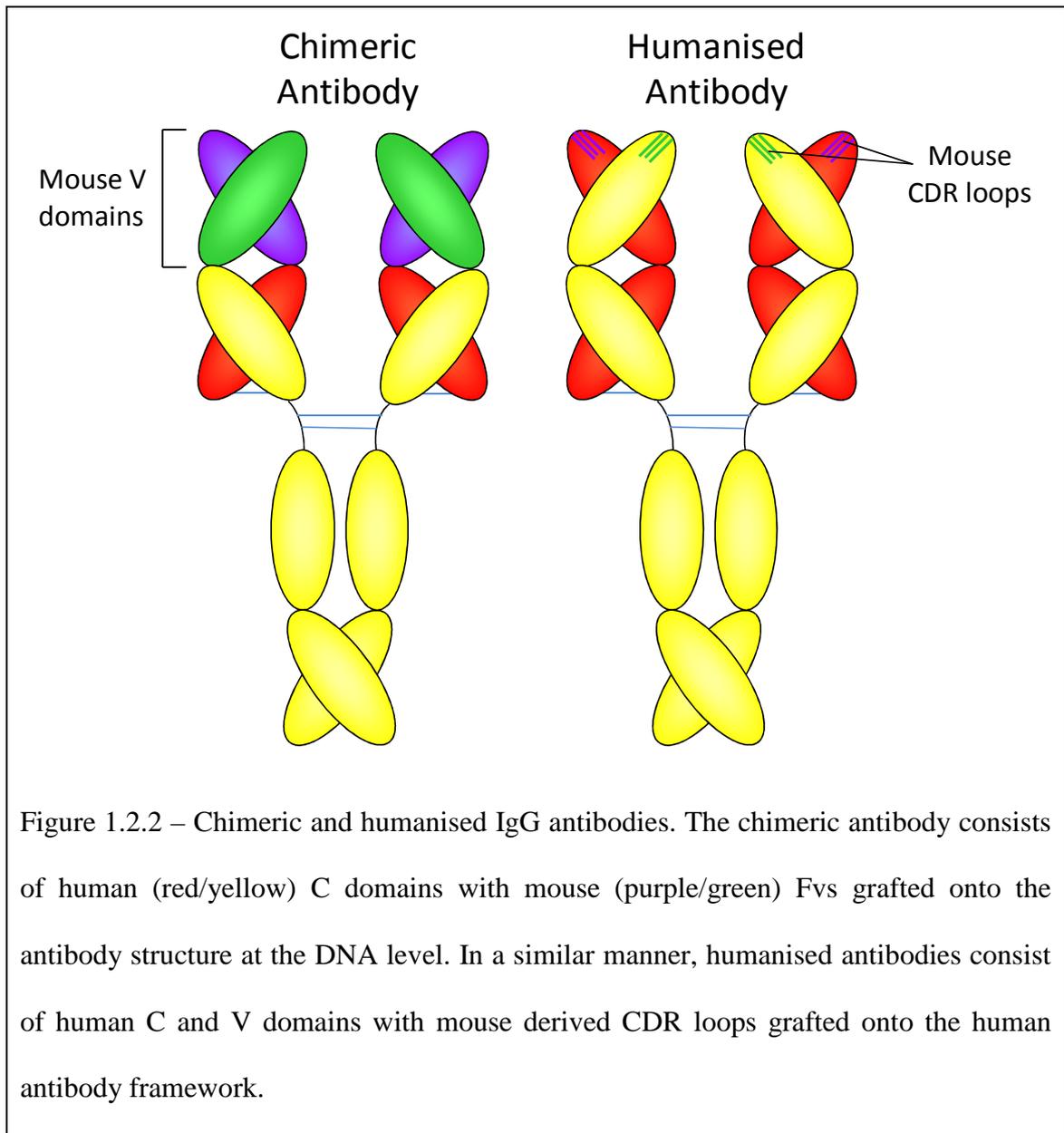
The production of fully human antibodies is the most recent step in improving the molecules generated as therapeutics. Both the phage/ribosome display and transgenic mouse methods (see below) have the potential to avoid the complex methods and potentially poor outcome of the CDR grafting process used in humanisation by the use of newer methods and technologies. The first of the two methods for generating fully human antibodies was the phage display method (Vaughan *et al*, 1996) which expresses

human antibody fragments on the surfaces of bacteriophage molecules to create large libraries of specificities which are used to screen for leads using the target antigen. The first antibody approved by the FDA demonstrating this process was the anti-TNFα antibody adalimumab in 2003 for the treatment of rheumatoid arthritis (Weinblatt *et al*, 2003). The alternative approach, using transgenic mice in which the native antibody genes have been replaced by human versions (Russel *et al*, 2000; Lonberg, 2005), is the most recent published method of generating potential therapeutic antibodies and is advantageous in that the antibodies will be affinity matured by the mice as they are produced. This method, however, does have drawbacks as the antibodies will be non-reactive to murine orthologs of the target gene, ruling out an important testing system for the potential therapeutics. The first antibody produced by this process, approved by the FDA in 2006 for use in the treatment of colorectal cancer, was the anti-EGFR antibody panitumumab. As the technology develops, there is a trend in the industry to move towards fully human antibodies. However, the majority of fully human antibodies are more immunogenic than their most recent humanised counterparts with only very recent examples achieving less than 1 % immunogenicity (Yoon *et al*, 2010). This information suggests that there is still a great deal to be learnt about these proteins, how they function and how they interact with the immune system.

In addition to the published methods, it is highly likely that many companies have produced proprietary methods of producing high affinity monoclonal antibodies for use as therapeutics. One such technology is the selected lymphocyte activation method (SLAM) employed by UCB. This process selects single immune cells from a complete antibody repertoire that are expressing and secreting antibodies that demonstrate a desired specificity, affinity and functional activity. The genes from these cells are extracted and assembled into an expression vector. This process can be used for high

throughput screening of antibodies and has the potential to discover high affinity antibodies with rare functional properties that may be beneficial. It also holds a number of advantages over the display and transgenic mice technologies described previously, as a population of human B-cells is used. This ensures that only natural V domain pairings are used, a characteristic not selected for in display technologies, without the need for an animal system and the potential problems that this produces. The antibody fragments used in this project were produced using this technology.

1.3 – The Future of Antibody Therapeutics

Despite the vast success of antibody derived molecules as therapeutics, there are still a number of drawbacks associated with them. One of the most significant drawbacks is their expense when compared to small molecule drugs. Antibodies are large ~150 kDa protein complexes, containing disulphide bonds and numerous post translational modifications, which require complex mammalian expression systems for their production. In addition to this, the average dose of an antibody is very high, requiring grams of protein per patient over the course of a treatment (Chames *et al,* 2009). It is clear that improvements are required to both the manufacturing processes and the proteins themselves to make them cheaper to produce. Another potential drawback of the antibody molecule as a therapeutic is the relatively poor tissue penetration observed for such a large molecule. This characteristic may pose a particular problem in certain clinical situations, such as the treatment of solid tumours, where altered vasculature and tissue density heavily affect the efficacy of a treatment (Carter, 2006; Holliger & Hudson, 2005). The Fc mediated effector functions of the antibodies may also cause undesirable effects depending on the clinical application.

A number of the properties of antibodies may be either desirable or undesirable depending upon the required application. Properties such as the large size and poor tissue penetration that are a result of mechanisms that allow the antibody to persist in the circulation for a half life of up to 23 days (IgG1) (Chames *et al,* 2009; Yoon *et al*, 2010) or the Fc mediated effector responses are useful only in certain situations. If targeting cancers or pathogenic cells for example, antibody dependant cellular cytotoxicity (ADCC) and complement dependant cytotoxicity (CDC) may be advantageous as they help lead to the destruction of the cancerous or invading cells. However, if targeting molecules within the body, to alleviate auto-immune or inflammatory diseases for example, an Fc mediated response may be detrimental (Carter, 2006).

A number of potential drawbacks may be removed by tailoring the antibody molecule to suit the required application. In its simplest form this constitutes including or excluding the Fc from the final therapeutic, most commonly resulting in either a Fab fragment or a full antibody. This approach has been available for some time and currently there are three Fab fragments approved for therapeutic use by the FDA (Riechert, 2010). The effectiveness of this approach has led to a wide range of fragments being produced during the research and development process for generating therapeutic molecules (Figure 1.3.1) that aim to address certain qualities of the antibody and modify them for a specific purpose. The most common of these fragments are the Fab and single chain Fv (scFv) fragments which are discussed in more detail in chapter two. Other constructs which provide features such as multiple specificity (e.g. diabodies, triabodies, tetrabodies, Fab$'_2$, Fab$'_3$, scFv$_2$ and bi-specific IgG), a lack of Fc (e.g. Fab, scFv), or the inclusion of an Fc and an attempt to reduce the overall size of the molecule (e.g. minibody, scFv-Fc) have also been developed (Chames *et al,* 2009;

Holliger & Hudson, 2005; Nelson, 2010). Any number of these fragments, and indeed whole antibodies, may be conjugated to other proteins or functional groups such as radioisotopes to add additional features that are useful in the treatment of cancers and other diseases. The molecules may even be useful for delivering enzyme/pro-drug combinations to direct potentially toxic drugs to specific sites (Brekke & Sandlie, 2003; Holliger & Hudson, 2005; Scott *et al,* 2012). Despite much research, and a large number of different treatments being researched, no fragments other than Fabs are currently approved for use in the clinic.

Figure 1.3.1 – A graphical overview of a selection of the potential antibody derived fragments currently being investigated for therapeutic use. The classical antibody fragments are shown along with multi-specific and highly engineered molecules such as the scFv-Fc and minibody.
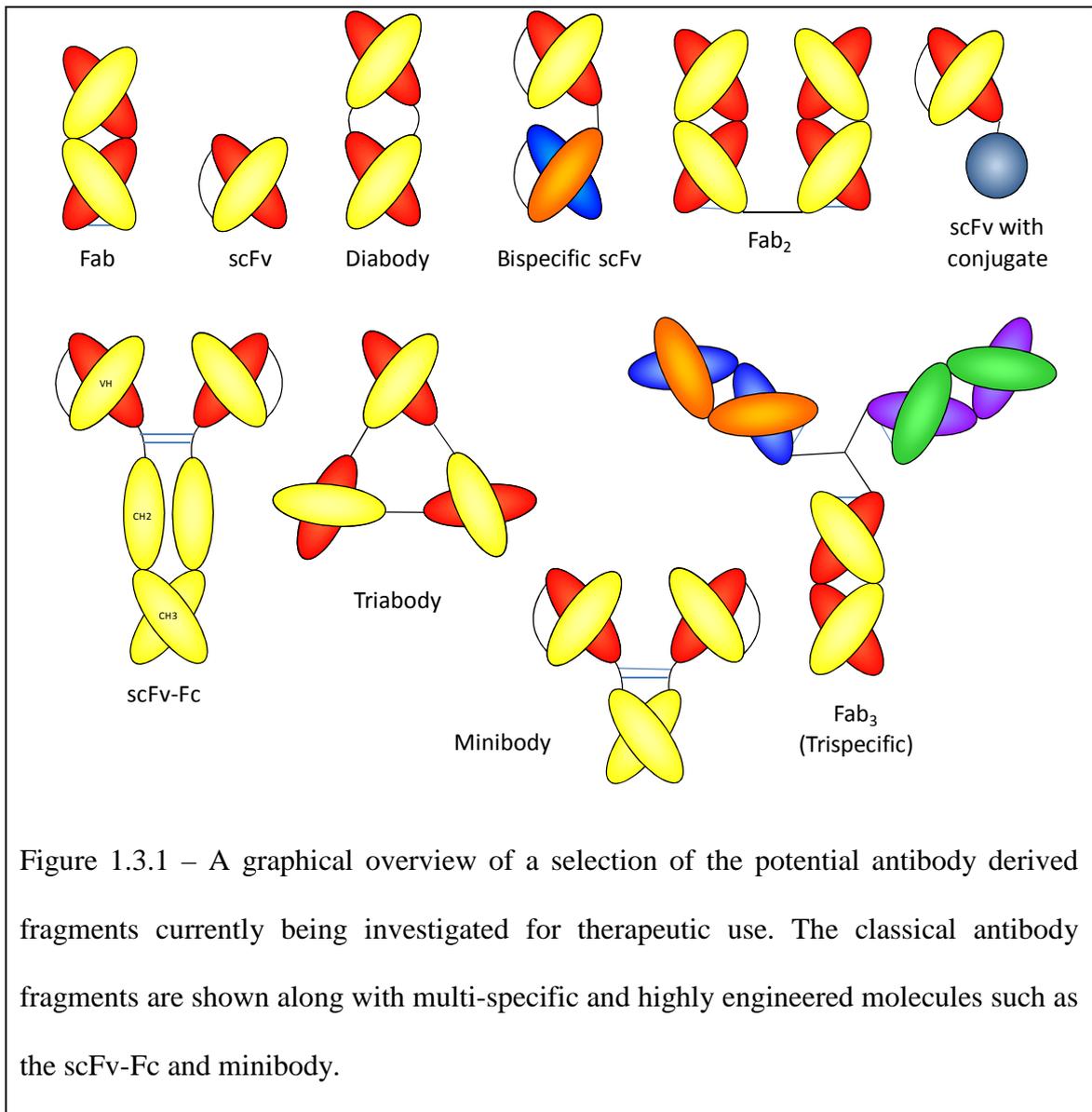
In addition to using antibody fragments and conjugates, a number of other interesting research areas are also active in attempting to improve antibody therapeutics. It has been shown that the exact nature of the glycosylation on the Fc portion of the antibody is highly important for proper Fc mediated responses, with differences in glycosylation showing an altered affinity for various Fc receptors that may affect the efficacy of the therapeutic (Chames *et al*, 2009). These differences in glycosylation are generally a

result of the expression system used and may require further research to fully understand the role that they play.

Most therapeutic antibodies are derived from IgG1, with a small number from IgG2 and IgG4. This use of subclasses in itself allows a certain tailoring of the response required. Additional research into using other antibody classes, particularly IgA, may offer alternatives to IgG with different characteristics (Bakema & van Egmond, 2011).

One of the most recent research areas for the development of antibody based therapeutics is the development of V_{HH} (variable domain of camelid heavy chain antibody) and V_{NAR} (variable domain of the shark novel antigen receptor) domains. These heavy chain only antibodies are found only in certain camelid (dromedary, camel, llama and alpaca) and cartilaginous fish species (such as sharks) (Saerens *et al*, 2008; Wesolowski *et al*, 2009) (Figure 1.3.2). These single domain antibodies show much potential for a large range of applications from experimental reagents to diagnostics and therapeutics (Wesolowski *et al,* 2009). From a therapeutic perspective, these molecules provide a number of interesting options due to their small size, high solubility, stability and longer CDRs (particularly CDR3) than conventional V domains that are able to penetrate into binding sites or crevices within an antigen. These features give the potential for high yields of expressed proteins that have good pharmokinetic properties, and are able to target a novel array of binding sites which may be inaccessible to other antibody derived therapeutics. Potential applications include the use of these single domains to target proteins and block binding sites in much the same way as for conventional antibody derived fragments. They are ideal for any application that requires deep tissue penetration and rapid clearance, such as radionuclide targeting in tumour imaging or cytotoxic drug or enzyme/pro-drug delivery in cancer therapy.

Other more elaborate applications have also been tested. It is possible to form multi-valency or multi-specificity molecules with tethered single domains that either have a higher avidity (Conrath *et al*, 2001) or are able to improve serum half life and efficacy by binding to molecules such as albumin (Copieters *et al*, 2006). It is also possible to attach these domains to a human Fc to elicit effector functions (Himla *et al*, 2008). Due to their small size and stability, it is also possible to generate commensal gut bacteria that express the domains on their surface. This may be particularly useful for the treatment of diseases such as severe diarrhea (Pant *et al*, 2006). It is obvious that the potential for these single domains and their related constructs is large. A number of drawbacks, however, still exist that limit their usefulness until further breakthroughs are made. Currently, the immunisation of llamas and dromedaries is less than ideal in terms of the cost and awkwardness of handling large animals. Immunogenicity is also a large concern as the proteins are non-human in nature. Currently, this does not appear to be a major issue for the single domains as they are minimally immunogenic, although this may change after repeated doses or as the proteins become more complex with the addition of other components. Work to resolve issues such as these using transgenic mice and humanisation schemes are still in the early stages of development, but show promise for the future and provide a potential path for developing novel antibody therapeutics (Saerens *et al*, 2008; Wesolowski *et al*, 2009).

Figure 1.3.2 – An overview of the heavy chain antibodies found in camels and sharks. Both the camelid heavy chain antibody and shark novel antigen receptor (IgNAR) lack light chains and do not contain additional C domains in their Fab sections. The antibodies also contain long flexible linkers and larger CDR loops to allow the binding of novel epitopes.
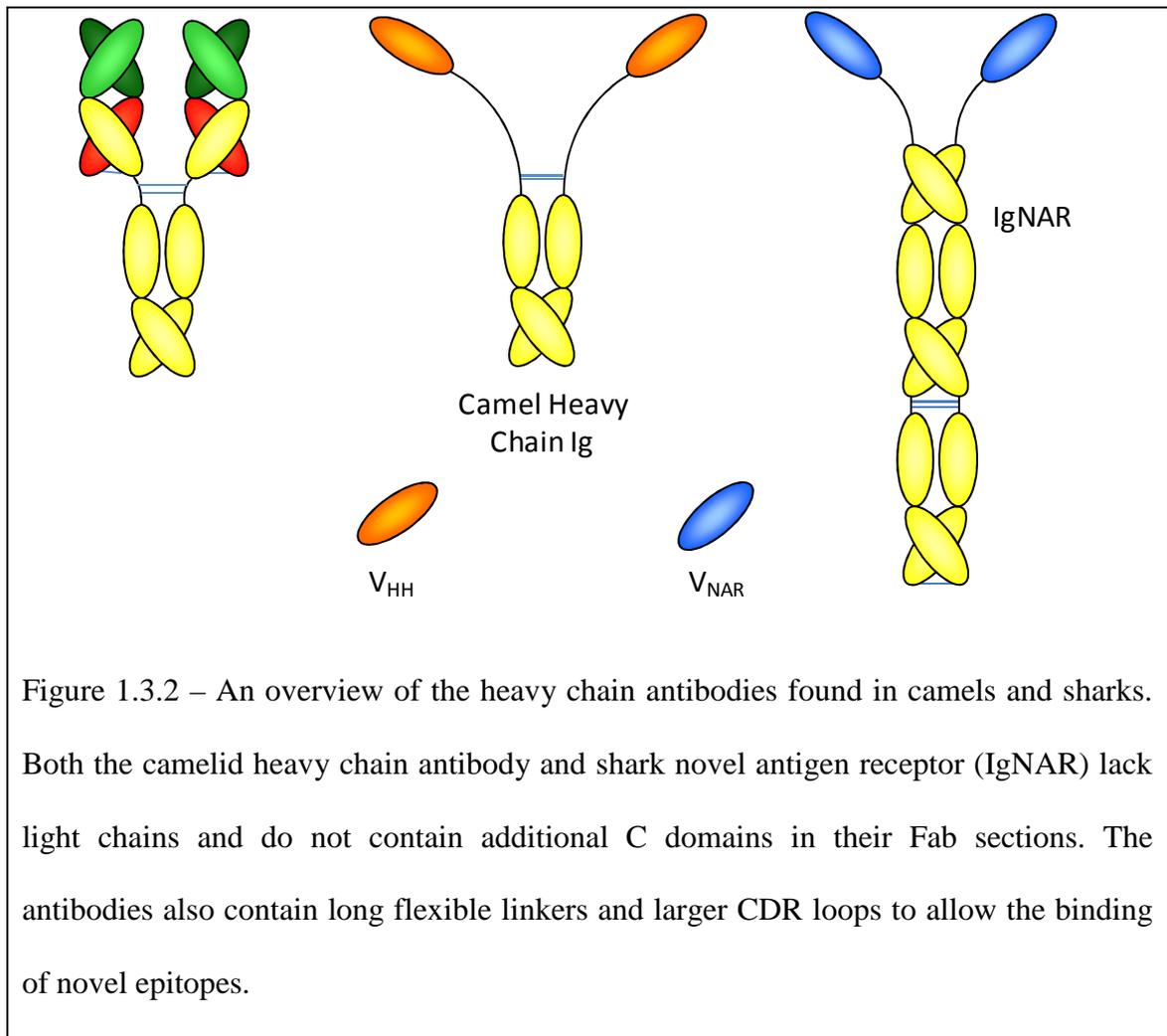
It is clear to see that there is significant work to still be done in developing antibodies for use as therapeutics. As technologies become more advanced, antibodies and their derived molecules will become more tailored to specific applications. It is this increase in rational development and property selection that paves the way for the future of therapeutic antibodies.

1.4 – Interleukin-6

Interleukin-6 (IL-6) was first cloned in 1986 and identified as a differentiation factor, progressing B-cells into an immunoglobulin producing state. As more detail into functional properties was determined in the initial research, a number of names were used. Initially known as B-cell stimulatory factor-2 (BSF-2) when discovered by Hirano *et al* in 1986, and subsequently as 26 kDa protein (Haegeman *et al*, 1986), interferon-β2 (Sehgal *et al*, 1987), hybridoma/plasmocytoma growth factor and hepatocyte stimulating factor (Gauldie *et al*, 1987), IL-6 has been linked to an ever expanding set of functions, both in normal and disease states. IL-6 is a 21-28 kDa glycosylated protein of the IL-6 family of cytokines that all share a similar four helix bundle structure (Xu *et al* 1996, pdb accession code 2IL6, Figure 1.4.1) and, bar interleukin-31 (IL-31), share a common signal transducer as part of their receptor complex (Scheller *et al*, 2011). IL-6 itself has long been known to be a key factor in the maturation of B-cells and is one of the main proteins driving the acute inflammatory response along with interleukin-1 (IL-1) and tumour necrosis factor-α (TNF-α). These three proteins are the primary regulators of the transfer from the innate to the adaptive immune response (Naugler & Karin, 2008). It is now known that, along with playing a crucial role in B-cell differentiation, IL-6 is a key factor in the differentiation of T-cells. Both T-helper-2 and T-helper-17 (T_H2 and T_H17) cells are promoted whilst differentiation into regulatory T-cells (T_{reg}) is down-regulated along with their recruitment and survival at sites of inflammation (Mihara *et al*, 2012; Naugler & Karin, 2008; Scheller *et al*, 2011). More recently, IL-6 has also been shown to have a role in processes such as metabolic control, bone metabolism and the pain response (Scheller *et al*, 2011). Perhaps unsurprisingly for a protein with such a wide range of roles, the aberrant regulation of IL-6 is associated with a large number of inflammation associated diseases such as

inflammatory arthritis, inflammatory bowel disease, sepsis, obesity, insulin resistance, anaemia of chronic diseases, various cancers and ageing amongst others (Mihara *et al,* 2012; Naugler & Karin, 2008; Scheller *et al*, 2011). This varied pathological role of IL-6 makes it a valuable potential therapeutic target for a wide range of diseases, particularly as there is currently only one approved therapy targeting it (Tocilizumab, Mihara *et al*, 2005) with previous development of drugs targeting inflammatory diseases focussing more on TNF-α.
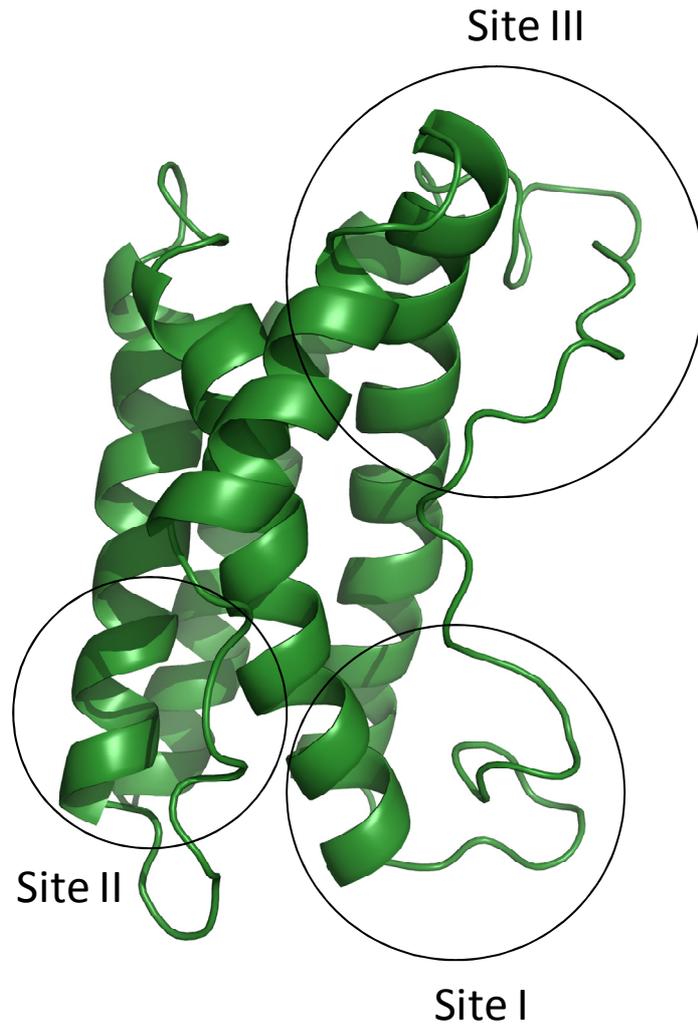
Figure 1.4.1 – The NMR derived IL-6 structure as determined by Xu *et al* 1996. The secondary structure of the protein is shown displaying the characteristic α-helical nature of the IL-6 family of cytokines. Site I, II and III indicate the IL-6R/gp130 interaction sites which are required for signalling.

IL-6 signals through a multi-component receptor complex that involves both a specific receptor (IL-6R), found mainly on T-cells, B-cells, monocytes and neutrophils, and a signal transducing component that can be found on the surface of the majority of cells (gp130). Initially IL-6 binds to its membrane bound IL-6R with a high affinity (0.1-1.0

nM) (Mihara *et al*, 2012). The signalling active receptor complex then forms from two gp130 proteins (which may be in a pre-associated homodimer (Scheller *et al*, 2006)) binding with very high affinity (~10 pM) to one (Scheller *et al,* 2006) or two (Skiniotis *et al*, 2008; Wang *et al*, 2009) IL-6/IL-6R complexes to properly orient and position the intracellular components of the gp130 protein for signalling. The exact number of IL-6/IL-6R complexes involved in this receptor complex, and the manner in which they associate with gp130, is a matter of some dispute within the field. The IL-6 protein has three binding sites for this receptor complex that are referred to as site I (IL-6R binding), site II (binding to gp130 between domain 2 and 3) and site III (contacting the Ig like domain 1 of gp130). These sites are highlighted on the structure of IL-6 shown in Figure 1.4.1, and a schematic of the potential signalling complexes is shown in Figure 1.4.2. It has been proposed by Scheller *et al* (2006) that in situations where there is a particularly high level of IL-6, a hexameric arrangement of the IL-6 receptor complex may be favoured which is signalling inactive (Figure 1.4.2). This characteristic may contribute to the bell shaped signalling response curve that sees signalling attenuated at high levels of IL-6, but is disputed by other research into this area that gives evidence only for an active hexameric complex (Skiniotis *et al*, 2008; Wang *et al*, 2009). In addition to this, IL-6 signalling can be triggered by an alternate form of receptor activation involving a soluble IL-6R (sIL-6R) that is either cleaved from cells by enzymes such as zinc metalloprotease enzymes of the ADAM (a disintegrin and metalloprotease) family or produced by alternative splicing, and is able to activate cells which do not express the IL-6R (Figure 1.4.3) (Scheller *et al*, 2006). As relatively few cells express the IL-6R, this trans-signalling allows IL-6 to signal to any cell expressing the gp130 component. As the majority of cells within the body express gp130, this expands the scope of IL-6 initiated responses dramatically. This form of signalling

(known as trans-signalling) is particularly important for regulating lymphocyte trafficking to sites of inflammation by the regulation of chemokines (Chalaris *et al*, 2007; Rabe *et al*, 2008), the differentiation of T-cells (Dominitzki *et al*, 2007) and the regeneration of the liver (Scheller *et al*, 2011). It is also involved in inducing T-cell proliferation during colon cancer development and regulating adhesion molecule expression on endothelial cells (Becker *et al*, 2004; Chen *et al*, 2006). Under normal circumstances, this form of signalling is regulated by a soluble form of the gp130 signal transducer that binds to the soluble IL-6/IL-6R complex and prevents association with the membrane bound, signalling active form of gp130 (Jostock *et al*, 2001). Soluble gp130 levels are regulated mainly by alternative splicing and can be found in fairly high concentrations in the serum (Narazaki *et al*, 1993). This pathway plays an important part in the regulation of the immune system. A breakdown in the control of this pathway is, however, commonly associated with a large portion of IL-6 related diseases, many of which can be alleviated by specific blocking with gp130 (Rose-John *et al*, 2006; Scheller *et al*, 2006; Scheller *et al*, 2011).

Figure 1.4.2 – The signalling complexes of IL-6. Two signalling active receptor complexes have been proposed for IL-6/IL-6R/gp130 signalling. The tetrameric complex, as suggested by Scheller *et al* (2006), consists of two gp130 molecules (with immunoglobulin like domain (Ig) and cytokine-binding module (CBM)) associated with IL-6 and IL-6R. Wang *et al* (2009), however, suggest that a hexameric complex is formed with an additional IL-6/IL-6R dimer. The only crystal structure produced is of the hexameric complex and is also shown (Boulanger *et al* 2003).

Figure 1.4.3 – IL-6 classic and trans-signalling. Cells expressing IL-6R are able to directly bind to IL-6 and initiate classical IL-6 signalling. Cells which do not express IL-6R are also susceptible to IL-6 signalling through the soluble IL-6R initiating trans-signalling.

IL-6 signalling, whether through classic- or trans-signalling, functions through the use of two intracellular signalling pathways (Figure 1.4.4). These two pathways are the Janus Kinase (JAK)/signal transducer and activator of transcription (STAT) pathway and the SH2-containing tyrosine-specific protein phosphatase (SHP-2)/mitogen activated protein kinase (MAPK) pathway. Both of these signalling events bring about gene expression resulting in the increased survival and proliferation of cells, the formation of the acute phase response and the negative regulation of cytokine signalling (Scheller *et al*, 2006).
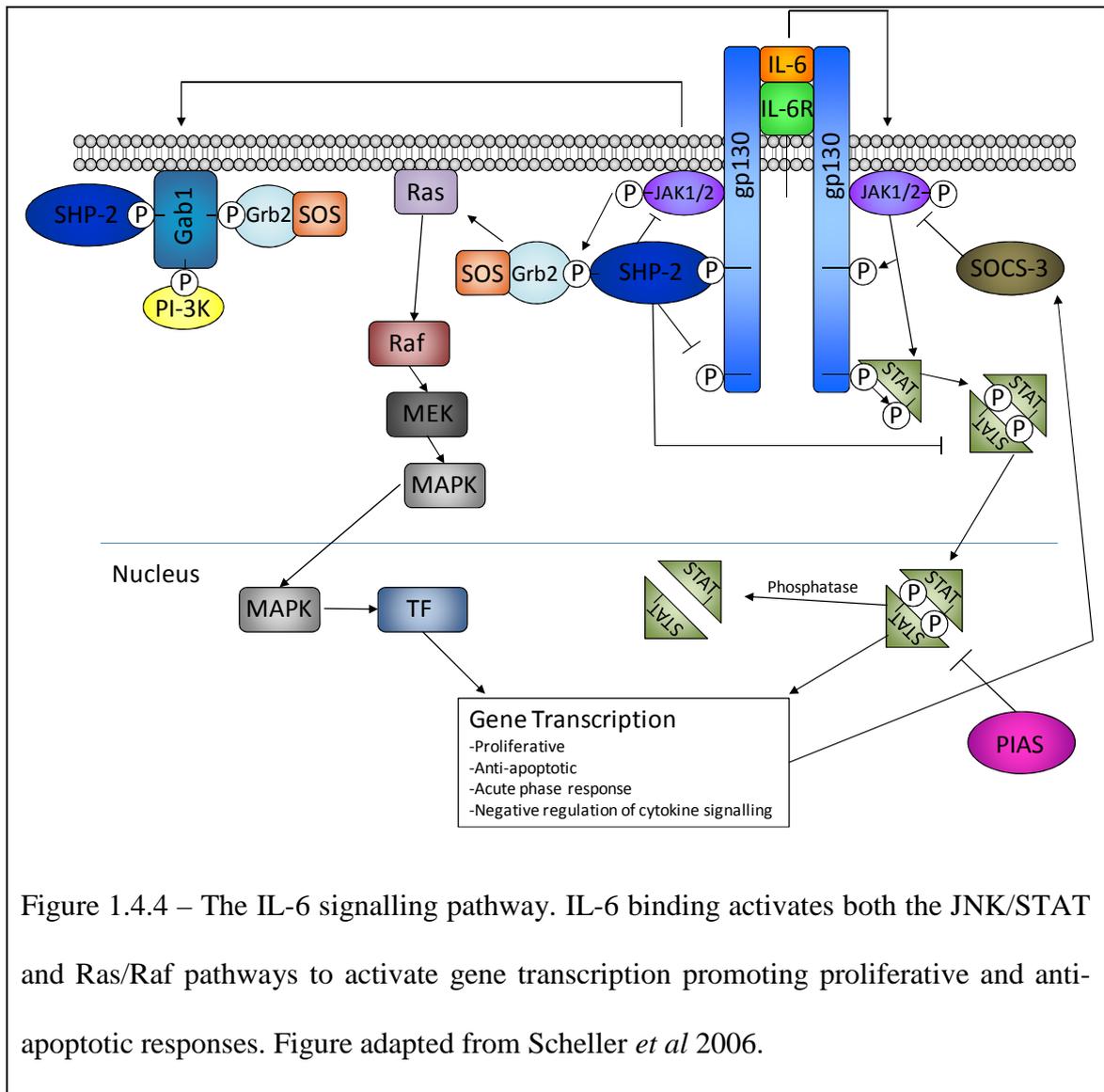
Figure 1.4.4 – The IL-6 signalling pathway. IL-6 binding activates both the JNK/STAT and Ras/Raf pathways to activate gene transcription promoting proliferative and anti-apoptotic responses. Figure adapted from Scheller *et al* 2006.

It is clear that the role of IL-6 within the body is both complex and highly important, with a breakdown in regulation leading to a wide variety of diseases. This importance, and wide pathological role, makes IL-6 and its signalling pathways an important target for therapeutic intervention. By blocking either the classical or trans IL-6 signalling pathways, it may be possible to treat a wide range of diseases by reducing inflammation related symptoms and degeneration and improve the outlook for a number of cancers. Therapeutic antibodies are an ideal candidate for specifically targeting extracellular

proteins such as these and present an interesting and highly valuable tool for the development of new treatments.

1.5 – Thesis Overview

The work presented in this thesis describes the detailed investigation of the structural properties of antibody/antigen interactions, focussing in particular on the features and properties of the CDR loops in both the bound and unbound states and the characterisation of structural changes associated with antigen binding. Further understanding of the structural properties of these interactions is critical for the continued development of antibody-derived therapeutics. Previous work using an anti-IL-1β scFv and Fab suggests that the CDR loops exist in an exchange between two or more states and are stabilised upon binding (Hall, 2009; Wilkinson, 2009). These studies also identified a subtle but significant change in orientation between the V_L domain and the rest of the Fab structure which may be important in the binding and further function of antibodies. The work presented here aims to build upon this information and determine general features of antibody binding by providing additional detail on the nature of interactions of this type. Therapeutic candidate antibodies targeted at IL-6 were used to generate an independent and structurally unique experimental system. Chapter two describes the production of isotopically labelled scFv proteins in *E. coli* that were used to determine the validity and characteristics of these proteins as amenable and reliable models for antibody/antigen interaction. This chapter also describes the optimisation of a small volume, high cell density *E. coli* expression protocol for the production of deuterated scFv proteins and its development to include

unlabelled amino acids, which may help in the sequence specific assignment of large or difficult proteins.

Chapter three describes the collection and analysis of NMR spectroscopy data for an anti-IL-6 scFv derived from a candidate therapeutic antibody to provide information on the structural properties of the scFv/antigen interaction. Of particular interest was the structural nature of the CDR loops when in an antigen bound and unbound state and the identification of areas that demonstrate significant structural changes upon binding outside of these loops. High quality triple resonance data was obtained for the scFv/IL-6 complex and, as far as can be determined from the literature, for the first unbound scFv in solution. H_N, N, Cα, Cβ and C$'$ assignments for both states were determined to a high level of completeness, allowing the location of structural changes within the proteins to be determined by chemical shift perturbation data and mapped onto a homology model of the scFv. Additionally, $^{15}N/^{13}C/^{2}H$ labelled IL-6 was produced and used to confirm that the scFv protein is a good model for studying antibody antigen interactions by NMR.

Previous work within the group using anti-IL-1β antibody fragments had led to the discovery of a V_L domain movement that may be important for antibody function and the development of a model for the initiation of BCR signalling. On the basis of this information, and the work described in chapters two and three, various aspects of this model and the domain movements seen were investigated in chapter four. This more detailed investigation into the potential function of the V_H/V_L domain re-orientation focussed on the potential for an antigen modulated interaction between the Fab and Fc parts of the antibody molecule. Isotopically labelled IC8Fab was produced and used, in conjunction with unlabelled IL-1β and IgG/IgE Fc, to collect high quality $^{15}N/^{1}H$

TROSY spectra for the Fab in the presence or absence of Fc and/or IL-1β. The NMR data obtained suggests a weak, antigen modulated, interaction between the Fab and Fc. This has led us to form a revised model for BCR signalling that provides both a basis for future studies to build upon and key insights into the as yet poorly understood mechanism of activation via the B-cell receptor.

Chapter 2 – Cloning, Expression and Purification

2.1 – Introduction

2.1.1 – The scFv and Fab constructs

The successful development of antibodies as therapeutics has driven the pursuit of more easily obtainable derivative molecules that contain the desirable antigen binding activity, without the potential complications involved with Fc mediated immune responses. The Fab fragment was one of the initial portions of the antibody produced by enzymatic digest (Porter, 1950; Porter, 1959), and has subsequently been found to offer these characteristics whilst being easy to produce and develop. The smallest fragment of the antibody that maintains full antigen binding capacity is the Fv (Bird *et al* 1988), consisting of only the V_L and V_H domains (Figure 1.1.1). Both the Fab and Fv fragment hold a number of advantages over the use of full length antibodies when considered as therapeutics. They are easier and cheaper to produce in a much wider range of expression systems, and have improved tissue penetration properties whilst avoiding any Fc mediated responses (Carter, 2006). They do, however, also hold a number of potential disadvantages. The fragments are significantly smaller in size than the full antibody and more readily removed from the blood stream by glomerular filtration. Affinity for the neonatal Fc receptor has also been shown to be critical in regulating serum half life for antibody molecules (Suzuki *et al* 2010). As these fragments do not contain the Fc portion of the antibody they have a circulation half life of as little as 90 minutes (Milenic *et al*, 1991), compared to as much as 21 days for whole IgG antibody

(Morell *et al* 1970). The Fv fragments also have the additional problem of a tendency for the two variable domains to dissociate from each other and form V_L dimers and V_H monomers which, despite some V_H domains retaining antigen binding activity, are prone to aggregation due to the exposed hydrophobic patches that form the interface between the two domains (Chothia *et* al, 1985, Ward *et al*, 1989). The serum half life can be improved by attaching conjugates to increase the total molecular weight of the constructs above the threshold for retention within the blood stream (Chapman, 2002). The dissociation properties of the Fv can be partially alleviated by introducing a flexible linker between the V_L and V_H domain keeping the protein tethered together without disrupting the interface or the CDR residues. This construct is known as the single chain Fv (scFv) (Bird *et al*, 1988 and Huston *et al*, 1988). This single chain solution, however, does not resolve the problem entirely as the domains can still dissociate and interact with the domains from one or more other scFv molecules (Wörn and Plückthun 2001), resulting in a slow exchange between the monomeric form and other domain swapped multimers, most often dimers (Figure 2.1.1). This is a well studied property of the scFv which can vary significantly between different scFv constructs with different linker lengths (Desplancq *et al*, 1994; Essig *et al*, 1993; Griffiths *et al*, 1993; Holliger *et al*, 1993; Whitlow *et al*, 1993), and under different conditions with pH, ionic strength and scFv concentration among the contributing factors (Arndt *et al*, 1998; Lee *et al*, 2002). The nature of this exchange provides the potential for the development of multivalent proteins that may bind to multiple antigens. These constructs, however, will retain some instability unless modified.

Figure 2.1.1 – A schematic representation of the process underlying the domain swap dimerisation of the scFv. The V_L and V_H domain of the monomeric scFv (A) are able to dissociate from one another and enter an open monomer state (B). This open monomer may either re-associate with itself, or may associate with another scFv molecule undergoing the same process to form a domain swap dimer (C). This interaction is in exchange, and under some conditions may allow the formation of higher order multimers (not shown). The scFv will be locked into whichever oligomeric state it is in upon binding to antigen.

2.1.2 – Therapeutic potential of scFv and Fab fragments

The push for the development of therapeutic monoclonal antibodies has produced 28 FDA licensed therapeutics, with more undergoing various stages of clinical trials and licensing (Reichert, 2010). A significant number of the candidate therapeutics still in clinical trials are derived from antibody fragments. However, currently the only licensed therapeutics are based on Fab fragments. Fragments such as these are often conjugated to other large molecules such as poly ethylene glycol (PEG) (Certolizumab pegol (Cimzia) UCB, Choy *et al*, 2002) to increase the serum half life. Conjugates such as PEG also provide other desirable properties such as reduced immunogenicity and toxicity, improved resistance to proteolytic cleavage, improved solubility and improved bioavailability (Chapman, 2002). The scFv has, so far, not been successfully developed as a therapeutic, mainly due to its instability and small size. It may, however, be useful for medical imaging and diagnostics or the targeted delivery of therapeutic agents, such as drugs, radionuclides or other proteins, where multiple specificities and/or its ability to penetrate tissues and tumours and then be rapidly cleared from the system may be advantageous (Ahmad *et al*, 2012; Milenic *et al*, 1991; Weisser & Hall, 2009). The formation of multivalent or multi-specific forms, or the addition of a variety of conjugates to these molecules, may again provide a number of interesting options for future development.

2.1.3 – Expression of scFv and Fab Fragments

Both the scFv and Fab fragments are readily producible in *E. coli* based expression systems as they are much less complex than full length antibodies and do not require

glycosylation. This method of expression is preferable over other eukaryotic expression systems as it is cheaper and easier to scale up to large volume fermenter cultures. The expressed protein can be targeted to the periplasm of the bacteria using leader sequences included in the expression construct that are removed during the transport process (Skerra & Plückthun, 1988; Skerra & Plückthun, 1991). The less reducing environment of the bacterial periplasm, when compared to the cytoplasm, allows the formation of disulphide bonds, which can aid the folding of the immunoglobulin domains and the linking of the heavy and light chains of the Fab fragment. Bacterial expression is also advantageous for structural studies by NMR as it provides the potential for isotopic labelling in a relatively easy and affordable manner. The use of NMR active isotopes of carbon and nitrogen, and the replacement of the majority of hydrogen atoms for deuterium, is essential for the study of proteins of this size to allow the collection of good quality, interpretable data. Isotopic labelling is most often carried out in minimal media containing sources of these isotopes (Gardner & Kay, 1998). This approach can produce mixed results and in many cases can severely reduce protein yields when compared to rich media, particularly when using deuterated minimal media. Recent studies into the optimisation of *E. coli* expression systems has resulted in the introduction of small volume, high density cultures that produce up to 680 mg l^{-1} of unlabelled or 500 mg l^{-1} of deuterated protein (Sivashanmugam *et al*, 2009). This method relies on the proper selection of high expressing colonies, and the use of small cultures to provide careful tuning of aeration, pH, nutrient concentrations and trace element levels, to provide very high levels of expressed protein (Sivashanmugam *et al*, 2009, Studier, 2005). This approach significantly reduces the cost of producing deuterated protein for NMR on a laboratory scale, and is also easily adaptable to include other forms of labelling that may have previously been cost prohibitive.

2.1.4 – The scFv as a model for antibody binding

Despite the occurrence of the domain swap multimerisation, which has traditionally limited the quality of data provided by structural studies due to the problems associated with samples containing a mixture of multimeric species, the scFv provides the potential for a good model system for the study of antigen/antibody interactions by NMR. It is advantageous in that it contains a full antigen binding component of an antibody in a relatively small and easily expressible construct, which can be locked into the monomeric state upon antigen binding. The work contained in this chapter describes the cloning and expression of four scFv constructs that target Interleukin-6 (IL-6), along with an associated Fab fragment which has been used to validate the scFv as a model for antibody binding. Buffer composition and the extent of the multimerisation under different conditions have been explored to produce samples that are optimised for the monomeric form of the scFv.

2.2 – Cloning of the scFv Constructs

2.2.1 – Materials and Methods

A set of four tetracycline resistant, IPTG inducible pTTod vectors (Figure 2.2.1) containing affinity matured IgG1 Fab constructs, 1189, 1160, 488 and 271, were supplied by UCB which had been generated in the process of selecting an anti-IL-6 therapeutic candidate. The vectors are designed for the tandem expression of the heavy and light chain of the Fabs. An OmpA periplasmic export sequence is included for the targeting of the expressed protein to the periplasm to aid folding and the formation of disulphide bonds which is cleaved off following transport. These vectors were transformed into InvαF *E. coli* cells (Invitrogen) for the production of glycerol stocks and purified plasmid DNA (Mini Prep Kit, Qiagen).

Figure 2.2.1 – Vector map for the Fab pTTod vectors provided by UCB. These vectors are directly inducible by IPTG and confer tetracycline resistance (Tet^R) for selection. The tandem expression of the Fab heavy and light chains is targeted to the periplasm by the OmpA periplasmic export sequence.
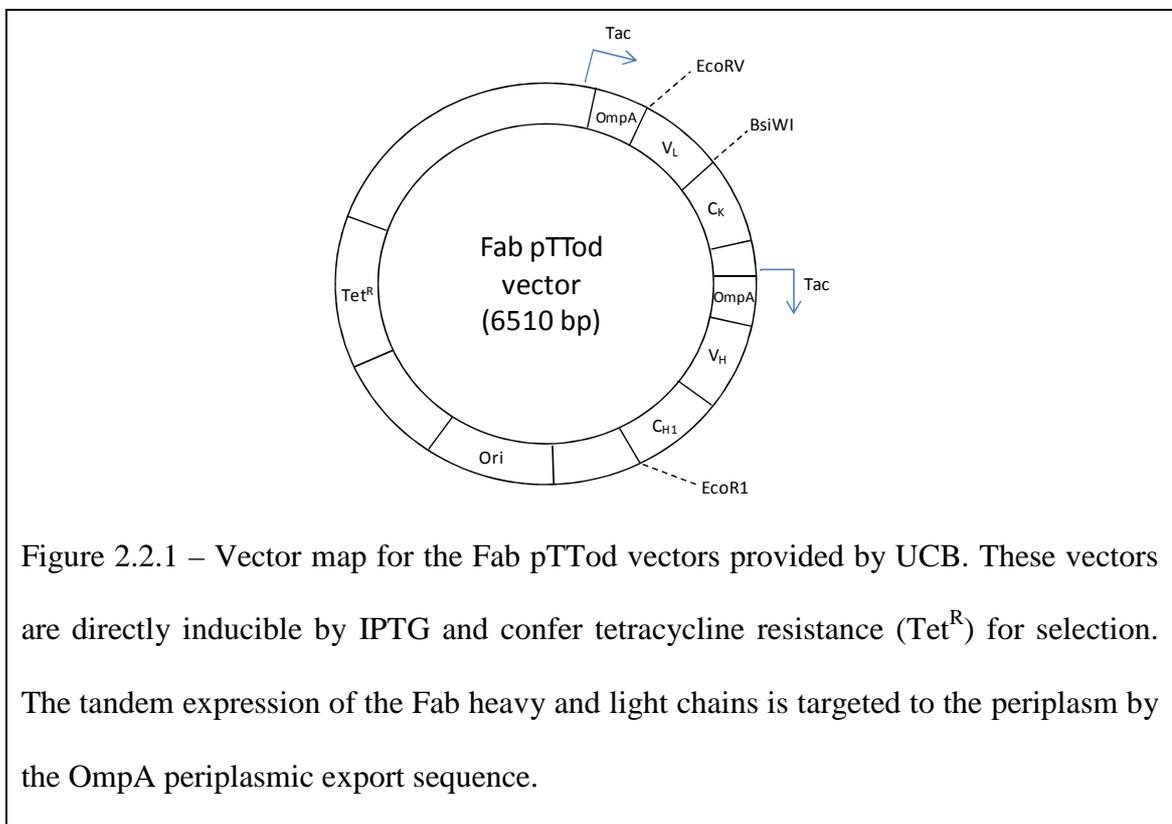
Two methods were employed to produce the scFv constructs. The first method produced scFv vectors by modifying the set that contained the Fab fragment constructs (Figure 2.2.2). The V_H domain was amplified from the plasmid using PCR (Taq Plus Precision polymerase, Stratagene), adding in coding for a BsiWI restriction site, and $(Gly_4Ser)_4$ linker onto the 5' end of the amplified fragment, and a His_6 tag and EcoRI restriction site onto the 3' end, using the PCR primers described in Figure 2.2.3. The amplified fragments were analysed by agarose gel electrophoresis (5 µl sample, 5 µl dH_2O, 2µl 6x Gel Loading Dye (NEB)) and run on a 1 % w/v agarose gel (containing a 1:20000 dilution of SYBR Safe (Invitrogen)), made with and run in TAE buffer (40 mM Tris, 1 mM ethylenediaminetetraacetic acid (EDTA), 20 mM acetic acid at 200 V against 100 b.p (Promega), Lambda/EcoRI + HindIII (Promega) or 2-log (NEB) ladders as appropriate. Gels were visualised under Ultra Violet (UV) light. Larger volumes of PCR product (mixed 5:1 with 6x gel loading dye) were gel purified (Gel Extraction Kit, Qiagen) following agarose gel electrophoresis and stored at -20 ˚C.
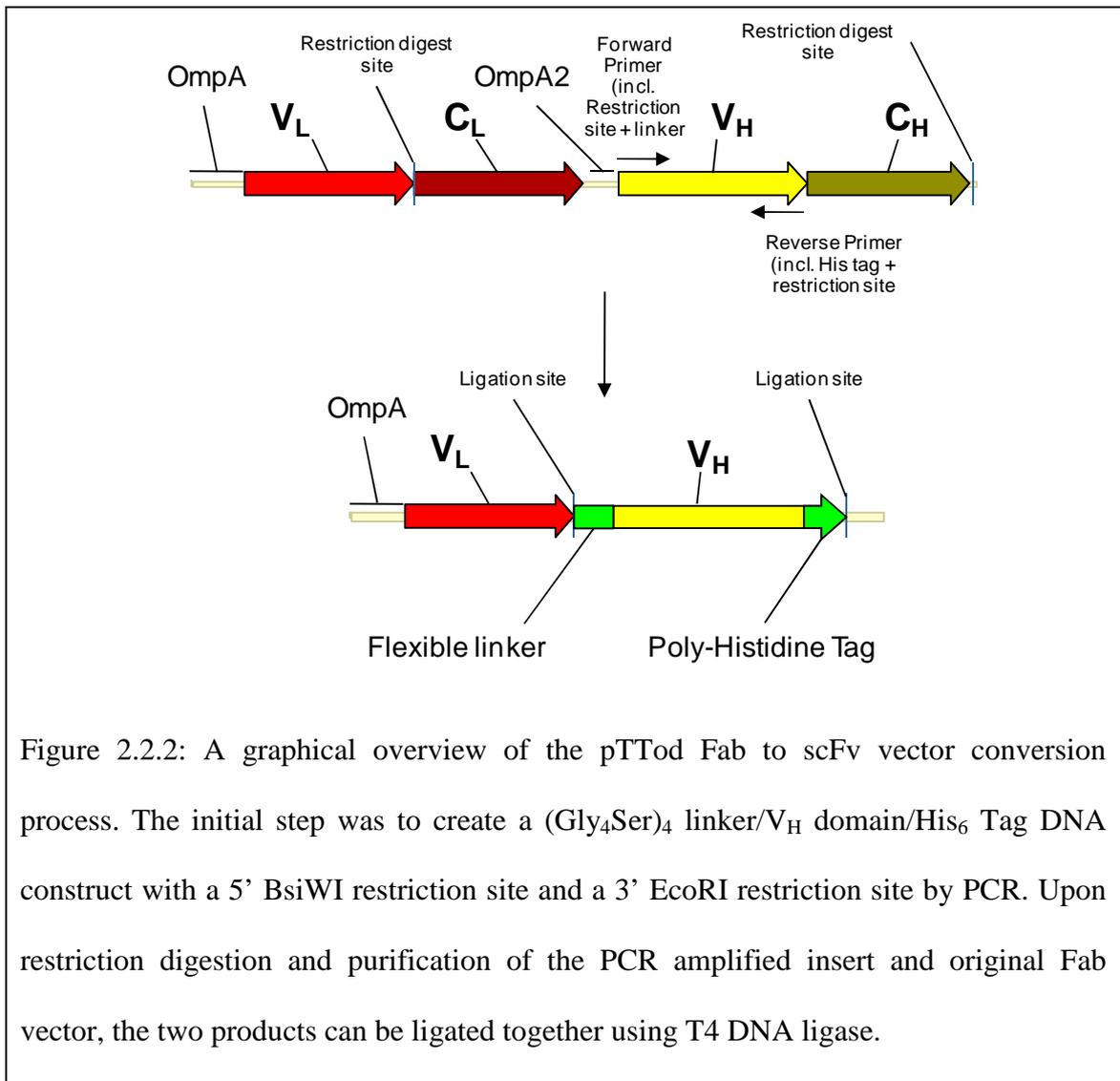
Figure 2.2.2: A graphical overview of the pTTod Fab to scFv vector conversion process. The initial step was to create a $(Gly_4Ser)_4$ linker/V_H domain/His_6 Tag DNA construct with a 5' BsiWI restriction site and a 3' EcoRI restriction site by PCR. Upon restriction digestion and purification of the PCR amplified insert and original Fab vector, the two products can be ligated together using T4 DNA ligase.

Forward:

1189, 271 and 488

GGACTG`CGTACG`GGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGAGGTTCAGCTGGTCGAGTCTGGAGGC
(99bp)

1160

GGACTG`CGTACG`GGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGAGGTTCAGCTGTTGGAGTCTGGAGGC
(99bp)

Reverse:
1160 and 271

CGCGTT`GAATTC`TTATCAGTGATGGTGATGGTGATGGCTCGAGACTGTCACCATGGTCCCCTG
(63bp)

1189
CGCGTT`GAATTC`TTATCAGTGATGGTGATGGTGATGGCTCGAGACAGTCACCATGGTCCCCTG
(63bp)

488
CGCGTT`GAATTC`TTATCAGTGATGGTGATGGTGATGACTCGAGACTGTCACAAGGGTCCCCTG
(63bp)

Figure 2.2.3: An outline of the primers produced for the amplification of the Fab V_H domain with 5' BsiWI restriction site and $(Gly_4Ser)_4$ linker and 3' EcoRI restriction site and His_6 tag. The regions coding for the linker are shown in green, hexa-histidine tag in yellow, stop codons in red and complementary regions in blue. The restriction sites (BsiWI for forward and EcoRI for reverse primers) are shown in grey.

To produce the completed scFv vector both the Fab plasmid and corresponding PCR product were restriction digested with BsiWI and EcoRI (NEB) and gel purified to remove any unwanted DNA. The digested PCR products were ligated into the appropriate digested vector at a 2:1 molar ratio of insert to vector using T4 DNA ligase

(NEB) at 16 °C overnight. This ligation reaction was used to transform InvαF *E. coli* cells, which were plated onto 10 μg ml^{-1} tetracycline LB Agar plates to select for colonies containing the plasmids.

The second cloning method employed a two step PCR method and empty pTTod vector that was unable to transform cells on its own (Figure 2.2.4). A set of four primers (Figure 2.2.5) were used to amplify the V_L domain, with EcoRV site and $(Gly_4Ser)_4$ linker, and the V_H domain, with linker, His_6 tag and EcoRI restriction site, in the first step. The second step then used the overlapping linker region and the primers at the 5' and 3' ends of the whole DNA construct to produce the entire scFv coding region as one PCR amplified DNA fragment. The products of these steps were visualised by agarose gel electrophoresis to ensure that the DNA produced was of the correct size. The resulting DNA fragment was double restriction digested with EcoRV and EcoRI (NEB), gel purified and ligated into empty, double digested, alkaline phosphatase treated (Antarctic phosphatase (NEB)), pTTod plasmid using T4 DNA ligase at room temperature for two hours and used to transform InvαF *E. Coli* as described above.
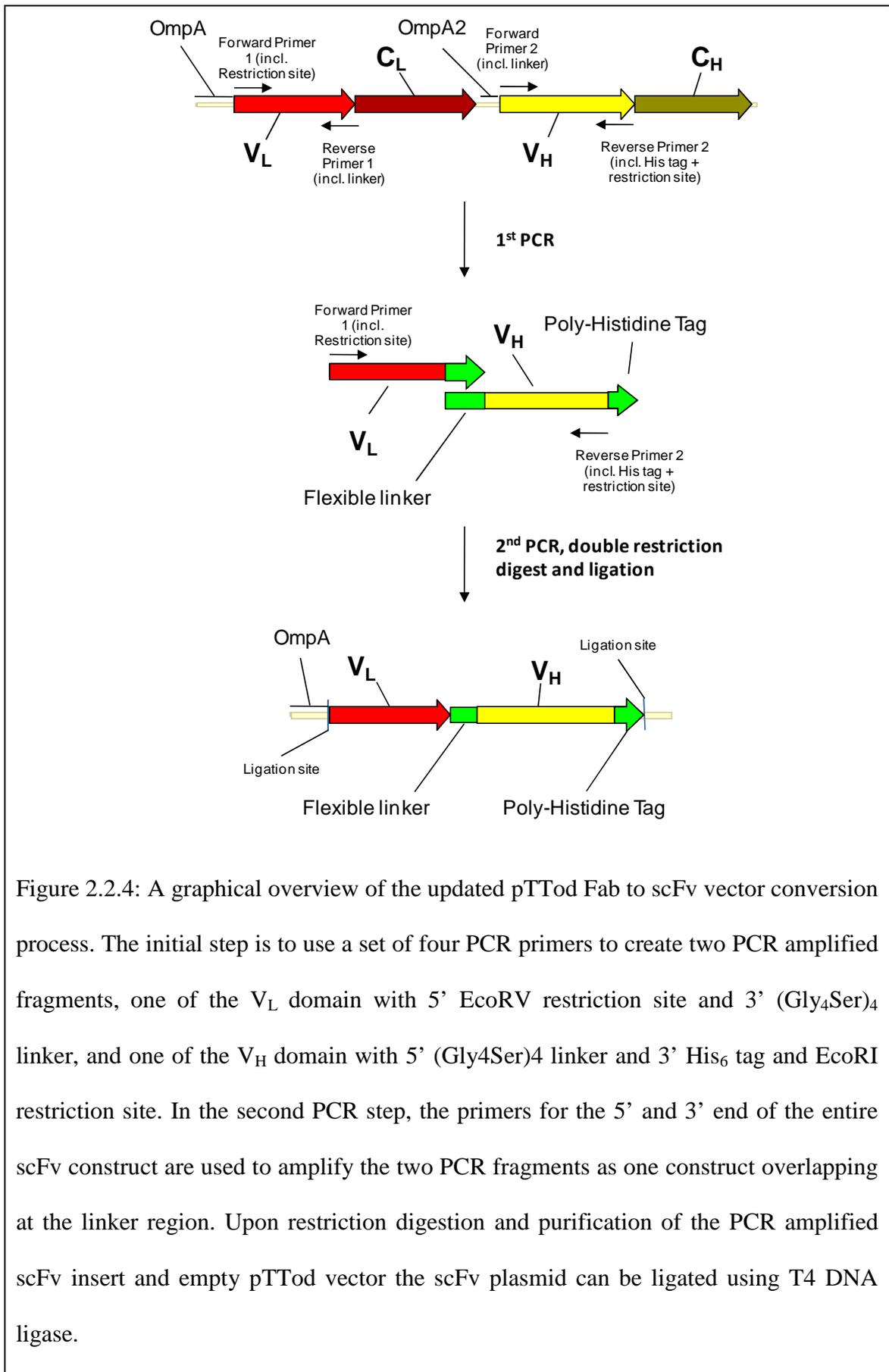
Figure 2.2.4: A graphical overview of the updated pTTod Fab to scFv vector conversion process. The initial step is to use a set of four PCR primers to create two PCR amplified fragments, one of the V_L domain with 5' EcoRV restriction site and 3' $(Gly_4Ser)_4$ linker, and one of the V_H domain with 5' (Gly4Ser)4 linker and 3' His_6 tag and EcoRI restriction site. In the second PCR step, the primers for the 5' and 3' end of the entire scFv construct are used to amplify the two PCR fragments as one construct overlapping at the linker region. Upon restriction digestion and purification of the PCR amplified scFv insert and empty pTTod vector the scFv plasmid can be ligated using T4 DNA ligase.

L488-EcoRv-Forward
GGGGGGATATCCAGATGACCCAGAGTCCAAG

L488-4G4S-Reverse
TGACCCGCCTCCACCTGACCCGCCTCCACCTGACCCGCCTCCTGACCCGCCTCCACCCGTACGTTTGATTTCTACTTTAGTG

H488-4G4S-Forward
GGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGAGGTTCAGCTGGTCGAGTCTGGAGGC

H488-HIS-EcoRI-Reverse
GGGGGAATTCTCAGTGATGGTGATGGTGATGACTCGAGACTGTCACAAGGGTCCCCTG

L1160-EcoRv-Forward and L1160-4G4S-Reverse are the same sequence as the 488 versions

H1160-4G4S-Forward
GGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGAGGTTCAGCTGTTGGAGTCTGGAGGC

H1160-HIS-EcoRI-Reverse
GGGGGAATTCTCAGTGATGGTGATGGTGATGGCTCGAGACTGTCACCATGGTCCCCTG

L271-EcoRv-Forward and H271-4G4S-Forward are the same as the 488 versions
H271-HIS-EcoRI-Reverse is the same as the 1160 version

H271-4G4S-Reverse
GGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGGTGGAGGCGGGTCAGGTGGAGGCGGGTCACGTACGTTTGATTTCCAGTTTAGTT

Figure 2.2.5: An outline of the primers produced for the amplification of the scFv construct containing 5' EcoRV restriction site, $(Gly_4Ser)_4$ linker, His_6 tag and 3' EcoRI restriction site. The regions coding for the linker are shown in green, His_6 tag in yellow, stop codons in red and complementary regions in blue. The restriction sites are shown in grey as specified in the primer name.

60

In addition to the cloning method described, an additional EcoRI site within the DNA coding for the V_H domain of the 488 construct was removed by introducing a point mutation into the Fab encoding vector (QuikChange Lighting Site Directed Mutagenesis Kit, Stratagene) prior to cloning. This mutation altered the codon for Asn 204 from AAT to AAC, which removed the restriction site whilst still coding for an asparagine residue.

The colonies produced from the cloning process were screened for inserts of the correct size using colony PCR reactions (GoTaq, Promega) utilising sequencing primers that bind to the plasmid outside the protein coding regions. The products of the colony PCR screen were visualised by agarose gel electrophoresis and, once inserts of the correct size were found, samples of the plasmid were generated by mini prep (Qiagen, as per kit instructions) and sent to PNACL or UCB for sequencing to confirm that the ligation had completed properly and that no mutations were present in the sequence.

Once vectors containing the correct DNA sequence were confirmed, W3110 *E. coli* cells (UCB) were transformed. Glycerol stocks were produced for each vector in both InvαF and W3110 cells.

2.2.2 – Results

The 488 Fab vector was successfully mutated prior to cloning and scFv vectors were successfully produced from all four Fab vectors using the two PCR methods described. The inserts produced by the one step PCR method (Figure 2.2.6-A) were ligated into the digested Fab pTTod vectors with mixed success. Incomplete digestion of the initial Fab containing plasmid limited the effectiveness of this method as the original Fab vector

was carried though into the transformation process. The screening of a large number of different transformed colonies, examples of which can be seen in Figure 2.2.6-B, was necessary to identify colonies containing the scFv vector. DNA sequencing of identified colonies confirmed that a construct had been made for a scFv based upon the 1189 Fab fragment (1189scFv) that was free of mutations or alterations.

The remaining three scFv constructs (1160scFv, 488scFv and 271scFv) were cloned using the multi-step PCR process. The DNA products of this method were visualised by agarose gel electrophoresis (Figure 2.2.7), gel purified, ligated and transformed into InvαF cells. Colony screens were again used to identify colonies containing the correct plasmids. Once found, plasmid DNA was generated by mini prep and sequenced to confirm the quality of the constructs.
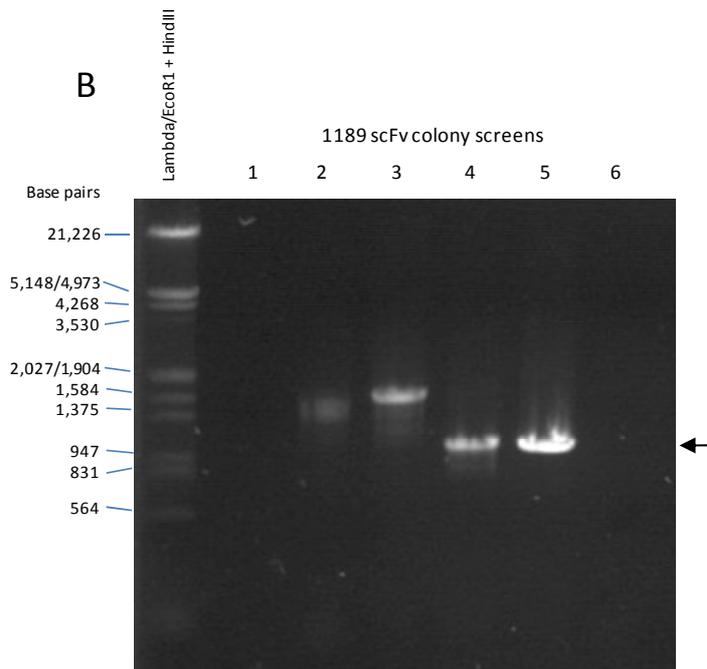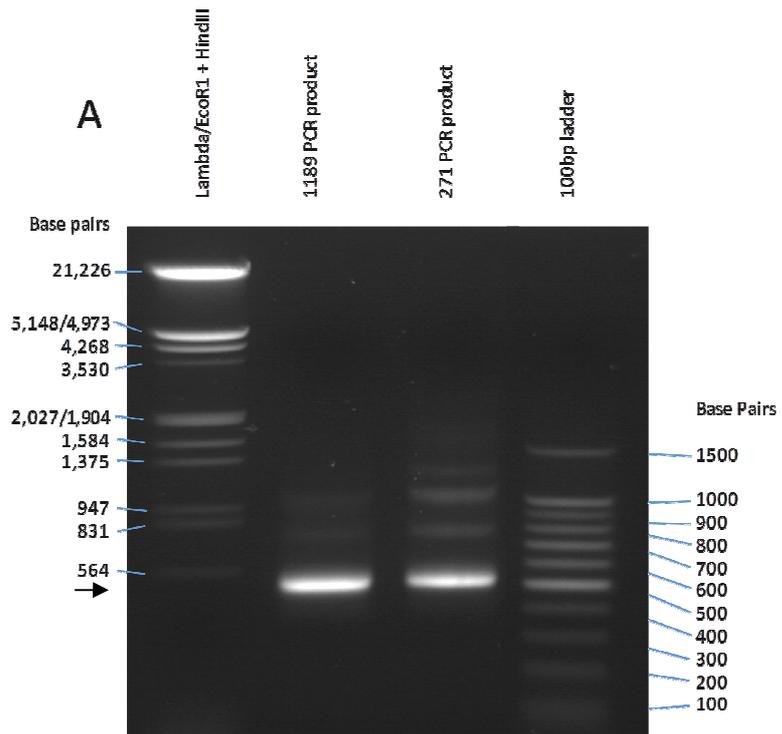
Figure 2.2.6 – Ultra-violet fluorescence visualisation of SYBR Safe (Invitrogen) stained 1% w/v agarose gels. Examples typical of the single step PCR reaction are shown displaying the successful amplification of the linker-V$_H$-His PCR inserts against a

Lambda/EcoRI + HindIII marker (Promega) (A). Examples of the colony PCR screen are also shown, demonstrating 1189scFv constructs of the correct size (lanes 4 and 5, approximately 900 base pairs), the incorrect size (lane 2, approximately 1400 base pairs) and full Fab constructs (lane 3, approximately 1700 base pairs).
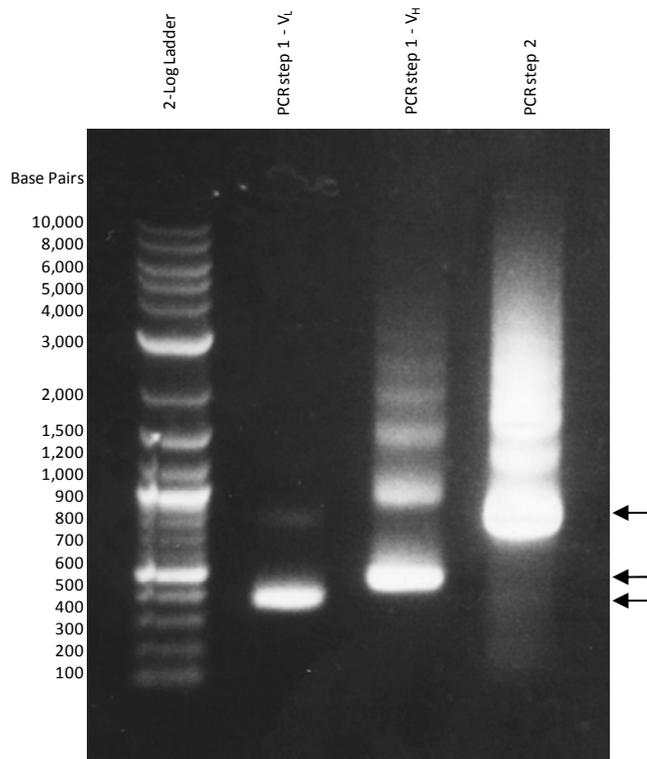
Figure 2.2.7 – An ultra-violet fluorescence visualisation of a SYBR Safe stained 1% w/v agarose gel. PCR products are shown for both steps of the multi-step cloning method a 2-log ladder (NEB). The required products (as marked) were separated from concatenated multimers by gel purification.

2.3 – Standard Cell Density Protein Expression

2.3.1 – Materials and Methods

2.3.1.1 – ScFv Expression trials

The initial expression trials for each of the four scFv constructs were conducted in two 50ml LB cultures, containing 10 μg ml^{-1} tetracycline, in 250 ml baffled flasks. Cultures were inoculated to an absorbance at 600 nm (A_{600}) of 0.1, with plasmid containing W3110 cells from a 10 ml overnight culture grown in LB (containing 10 μg ml^{-1} tetracycline) at 37 °C in a shaking incubator set at 200 rpm. Once seeded, the 50ml cultures were grown to an A_{600} of 0.8 at 37 °C/200 rpm. At this point, a 1 ml pre-induction sample (T_0) was taken and stored at -20 °C and the remaining culture was induced with IPTG to a final concentration of 200 μM. One flask for each construct remained at 37°C while the other was transferred to an incubator set to 15 °C. From this point onwards, 1 ml samples were taken and the A_{600} recorded at 1, 2 ,3 and 4 hour time points (T_1, T_2, T_3, and T_4 respectively) and after approximately 16 hours (overnight, T_{16}). All samples were stored at -20 °C until processing.

The samples taken from the expression trials were thawed at room temperature and normalised to an A_{600} of 0.8 by dilution with water. Once normalised, 200 μl of each sample was centrifuged at 20,000 x g for 10 minutes in a bench top centrifuge to pellet the bacterial cells. These cells were resuspended in 20 μl water, 10 μl 4x LDS sample buffer (Invitrogen) and 10 μl 200 mM Dithiothreitol (DTT) and heated at 70 °C for 10 minutes. The heated samples were centrifuged at 20,000 x g for 10 minutes to pellet any

DNA present in the sample, which was removed prior to analysis by coomassie stained Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE, section 2.3.1.2). In addition to these normalised samples, the remainder of the expression culture was pelleted at 6000 x g for 15 minutes in a benchtop centrifuge, resuspended in 2 ml lysis buffer containing 50 mM Tris pH 8.0, 2 mM EDTA, 50 μM phenylmethylsulphonyl fluoride (PMSF), 0.1 % Triton X-100 and 100 μg ml^{-1} lysozyme and incubated at room temperature for 30 minutes. Once lysed, bovine pancreas deoxyribonuclease (Sigma-Aldrich) was added with magnesium chloride to a final concentration of 5 mM. The sample was incubated for 30-60 minutes at room temperature to fully degrade the DNA. 20 μl samples were taken both prior to and following centrifugation at 50,000 x g for analysis by SDS-PAGE alongside the normalised time point samples.

2.3.1.2 – SDS-Polyacrylamide Gel Electrophoresis

20 μl samples for analysis by SDS-PAGE were combined with 10 μl 200 mM dithiothreitol (DTT) and 10 μl 4X LDS sample buffer (Invitrogen) and heated to 70˚C for 10 minutes. 20 μl of the heated samples were applied to wells on a NuPAGE Novex 4-12 % Bis-Tris pre-cast 1.0 mm polyacrylamide gel (Invitrogen) and run in MES-SDS buffer (Invitrogen) for 35 minutes at 200 V with a current setting of 125 mA per gel. The SDS-PAGE gels were analysed by staining in coomassie brilliant blue (2.5 g l^{-1} in 45 % methanol/10 % acetic acid) or by western blot (2.3.1.3). All gels were run with one lane containing either a 14-66 kDa molecular weight marker (Sigma-Aldrich) or Novex Sharp pre-stained protein standards (Invitrogen)

2.3.1.3 – Western Blot

Western blotting of histidine tagged proteins was used in cases where coomassie staining was not sensitive enough to detect low levels of expression. SDS-PAGE gels were run as in the previous section and, instead of staining, transferred onto a polyvinylidene difluoride (PVDF, 0.2 μM pore size, Invitrogen) membrane. The PVDF membrane was activated in methanol for 10 seconds and soaked for 30 minutes in transfer buffer (1.45 g Tris and 7.2 g glycine in 20 % methanol) together with any blotting pads and paper to be used in the transfer process. The transfer of protein bands from gel to membrane was done in an XCell II blotting module (Invitrogen) at 30 V/175 mA for 1 hour. Following transfer, the membrane was blocked in a 5 % (w/v) solution of ECL blocking agent (GE Healthcare) dissolved in TBST (135 mM sodium chloride, 15 mM Tris, 0.1% (v/v) tween-20, pH 7.6) for 1 hour at room temperature. Once blocked, the membrane was incubated for 1 hour with a 1:4000 dilution of mouse anti-His antibody (Merck) in 1 % blocking solution. The membrane was washed for 5 minutes in TBST and incubated with a 1:5000 dilution of HRP conjugated goat-anti-mouse secondary antibody in 1 % blocking solution for 1 hour. Following incubation, the membrane was washed five times in TBST for 5 minutes per wash and treated with ECL plus reagents (GE healthcare), as per the manufacturer's instructions, and exposed on X-ray film (Fujifilm).

2.3.1.4 – Expression of unlabelled scFvs

Starter cultures (50 ml LB, 10 μg ml^{-1} tetracycline in 250 ml baffled flasks) of W3110 *E. coli* containing the appropriate scFv plasmid were grown overnight at 37 ˚C/200 rpm

and used to inoculate expression cultures (500 ml LB, 10 μg ml^{-1} tetracycline in 2.2 l baffled flasks) to an A_{600} of 0.1. The main expression cultures were grown at 37 ˚C/ 150 rpm until an A_{600} of 0.6. Once the correct A_{600} was achieved, the cultures were transferred to an incubator of the correct temperature (37 ˚C for 1189scFv and 15 ˚C for 488scFv as determined in section 2.3.1.1/2.3.2.1) to equilibrate until an A_{600} of 0.8. At this point, IPTG was added to a final concentration of 200 μM. In all cases, the cultures were left to express overnight. Samples were taken at 1, 2, 3 and 4 hours and after overnight expression (16 hours) and analysed by SDS-PAGE as described previously with samples taken from the cell lysis procedure (section 2.3.1.8).

Once the expression had been completed, the bacterial cells were pelleted at 6000 x g for 30 mins and stored at -80 ˚C until processing.

2.3.1.5 – Expression of ^{15}N and ^{15}N/^{13}C labelled scFvs

Expression of the ^{15}N and ^{15}N/^{13}C labelled scFvs was carried out in 500 ml minimal medium (Appendix A.1.1) cultures containing 10 μg ml^{-1} tetracycline with 1 g l^{-1} ^{15}N ammonium sulphate, and/or 2 g l^{-1} ^{13}C$_6$ D-glucose where appropriate, at 37 ˚C for 16 hours following the protocol described in the previous section. Starter cultures were grown initially in 10 ml LB (10 μg ml^{-1} tetracycline) at 37 ˚C/200 rpm over the course of 8 hours and then in two 100 ml minimal medium cultures overnight at 37 ˚C/200 rpm to improve growth rates. Samples were again taken at 1, 2, 3 and 4 hours and after overnight expression and analysed by SDS-PAGE as described previously.

Once the expression had been completed, the bacterial cells were pelleted at 6000 x g for 30 mins and stored at -80 ˚C until processing.

2.3.1.6 – Expression of ^{15}N/^{2}H and ^{15}N/^{13}C/^{2}H labelled scFvs

W3110 *E. coli* containing the scFv vectors were conditioned on medium containing increasing concentrations of deuterium oxide (^{2}H$_2$O or D$_2$O). Initially, the bacteria were grown in 10 ml LB (10 μg ml^{-1} tetracycline) and then 50 ml cultures of minimal medium containing 10 μg ml^{-1} tetracycline and 0 %, 30 %, 70 % and 100 % D$_2$O. Glycerol stocks were made of these cells to use in inoculating future starter cultures for deuterated expression. All cultures were grown at 37 ˚C/200 rpm in 50 ml flasks for the 10 ml cultures and 250 ml baffled flasks for the 50 ml cultures.

Four 100 ml minimal medium starter cultures containing 10 μg ml^{-1} tetracycline, 1 g l^{-1} ^{15}N ammonium sulphate, 2 g l^{-1} ^{13}C$_6$ D-glucose (where appropriate) and 100 % D$_2$O were grown in 500 ml baffled flasks at 37 ˚C/200 rpm overnight (to an A$_{600}$ of 0.6) from four 10 ml 100 % D$_2$O LB cultures containing 10 μg ml^{-1} tetracycline that had been grown for 8 hours. These starter cultures were used to inoculate eight 500 ml expression cultures as described previously. Samples were taken as in section 2.3.1.4 and analysed by SDS-PAGE and western blot as described in section 2.3.1.3.

Once the expression had been completed, the bacterial cells were pelleted at 6000 x g for 30 mins and stored at -80 ˚C until processing.

2.3.1.7 – Expression of unlabelled 1189 Fab

Expression of the 1189 Fab construct supplied by UCB was carried out in LB cultures containing 10 μg ml^{-1} tetracycline. Four 100 ml cultures were used as starter cultures and inoculated from glycerol stocks of W3110 *E. coli* containing the pTTod 1189 vector. These were grown at 37 ˚C/200 rpm overnight in 500 ml baffled flasks and used

to inoculate eight 500 ml cultures in 2.2 l baffled flasks to an A_{600} of 0.1. The main cultures were grown at 37 ˚C/200 rpm and were induced with IPTG to a final concentration of 200 μM when they reached an A_{600} of 0.8, and left to express overnight. Cells were pelleted and samples were taken and analysed as in section 2.3.1.4.

2.3.1.8 – Purification of the scFvs

The bacterial pellets from the scFv expression cultures were thawed and resuspended fully by vortex in 20 ml of lysis buffer (50 mM Tris pH 8.0, 2 mM EDTA, 0.1 % Triton-X100) per litre of culture for the large scale expression method, or 10 ml per 100 ml of culture for the small scale expression method. Once resuspended, the cells were lysed by five passes through a French pressure cell. The resulting lysate was centrifuged at 50,000 x g for 30 minutes and the supernatant passed through a 0.2 μm filter to remove any insoluble material. The remaining soluble material was dialysed against 100 mM sodium chloride, 25 mM Tris pH 7.5 and 30 mM imidazole at 4 ˚C.

Once dialysed, the scFv containing sample was loaded onto a 6 ml nickel affinity column containing a nickel nitrilotriacetic acid (NTA) resin (Ni-NTA His-Bind Superflow, Merck) pre-equilibrated with 100 mM sodium chloride, 25 mM Tris pH 7.5, 30 mM imidazole) on an Akta Fast Protein Liquid Chromatography (FPLC) system (GE healthcare). Once the protein containing solution had been loaded, the column was washed with five column volumes (CV) of binding buffer to remove any unbound material. The protein could then be eluted with a linear gradient of 30-500 mM

imidazole over 5 CV and collected in 2 ml fractions using the Frac-950 attached to the FPLC system.

The scFv containing fractions were identified by SDS-PAGE, pooled and dialysed into a size exclusion chromatography (SEC)/NMR suitable buffer (100 mM sodium chloride, 25 mM sodium acetate pH 5.5, 0.02 % sodium azide (NaN_3)) at 4 °C. Dialysed samples were concentrated down to 5 ml by ultrafiltration (Vivaspin 20, Sartorius) and loaded at 1 ml min^{-1} onto a 120 ml size exclusion column (HiLoad 16/60 Superdex 75 Prep Grade, GE Healthcare) pre equilibrated with two CV of SEC buffer. Fractions were collected throughout the runs, with any scFv containing fractions (as determined by SDS-PAGE) being pooled and stored at 4°C with 10 μM EDTA and 200 μM 4-(2-Aminoethyl) benzenesulphonyl fluoride hydrochloride (AEBSF) added. The final amount of purified protein was quantified by absorbance at 280 nm using extinction coefficients calculated from the protein sequence (1189scFv ε = 47890 M^{-1} cm^{-1}, 488scFv ε = 46040 M^{-1} cm^{-1}).

Deuterated scFvs were denatured and refolded prior to dialysis into the SEC/NMR buffer to exchange buried amide deuterons for protons. The pooled fractions from the nickel affinity column were denatured by adding guanidine hydrochloride to 5 M (as determined in section 2.5.1.2/2.5.2.2). The guanidine hydrochloride could then be removed by four steps of dialysis, initially in two 2 l steps of a 100 mM sodium chloride, 25 mM Tris pH 7.5 buffer, and then in two 2 l steps of the SEC/NMR buffer prior to concentration and gel filtration.

2.3.1.9 – Purification of 1189 Fab

1189Fab containing pellets were thoroughly resuspended by vortex into lysis buffer, lysed by French press, centrifuged and filtered as described above. The resulting 1189Fab containing solution was adjusted to pH 6.0 using 0.1 M citric acid and loaded onto a 10 ml protein G column (recombinant protein G sepharose FF resin, Generon), allowing 20 minutes of contact time per ml of sample (0.5 ml min^{-1} in this case). The column had been pre equilibrated with protein G binding buffer (75 mM sodium phosphate, 75 mM sodium citrate, pH 6.0) on an Akta FPLC system. The column was washed with 4 CV of the binding buffer before the protein was eluted with 100 mM Glycine HCl, pH 2.7 and collected in 1 ml fractions.

The Fab containing fractions were pooled and immediately dialysed into SEC/NMR buffer in two steps at 4°C. The dialysed 1189 Fab sample could then be concentrated by ultrafiltration and stored as described previously.

2.3.2 – Results

2.3.2.1 – scFv Expression Trials

Previous expression trials conducted in the group for a different set of scFv constructs indicated that optimal protein production was achieved by inducing expression using 200 μM IPTG over a period of approximately 16 hours (Wilkinson *et al*, 2009). The initial expression trials of the anti-IL-6 scFv constructs produced soluble protein under these conditions at 37 °C for 1189scFv and at 15 °C for 488scFv. The expressed proteins were detectable by Coomassie stained SDS-PAGE (Figure 2.3.1) and western blot (Figure 2.3.2), showing increasing quantities over time in both cases. An estimation of the amounts of protein produced could be determined from the SDS-PAGE gel using reference samples of known lysozyme quantities (Figure 2.3.3, 1.25 μg of 1189scFv and 0.75 μg of 488scFv). Using this information, and the approximate amount of culture used to make the samples (100 μl), it was estimated that 25 mg l^{-1} of 1189scFv and 15 mg l^{-1} 488scFv were being produced, indicating that sufficient quantities of the scFvs could be produced for analysis by NMR spectroscopy. Both the 1160scFv and 271scFv constructs showed no signs of expression and were not carried forward past this stage.
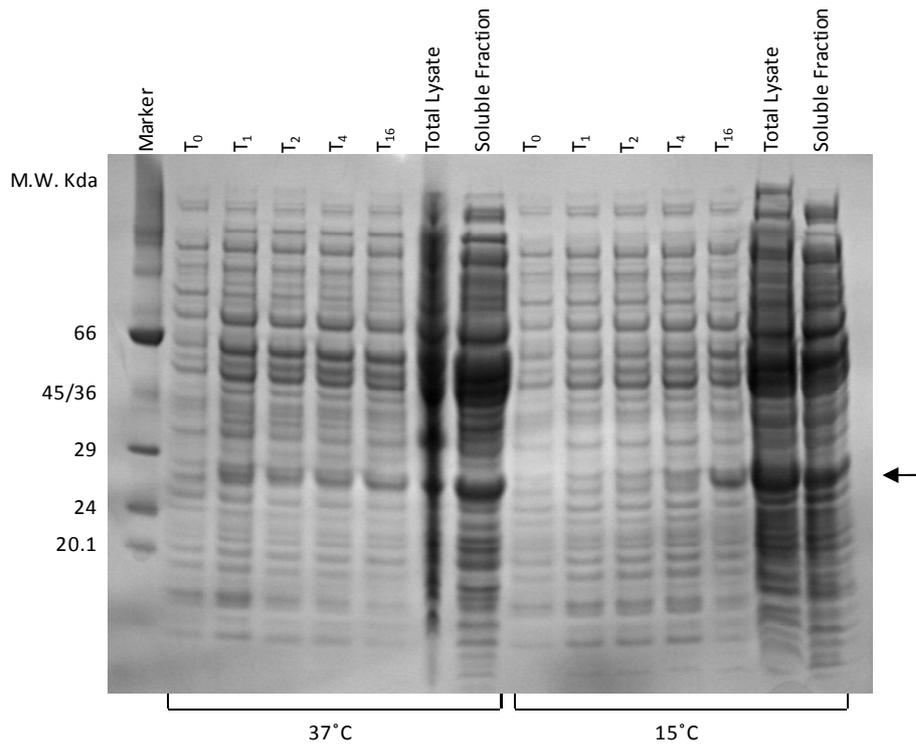
Figure 2.3.1 – Samples taken from the expression trials for 1189scFv analysed by SDS-PAGE and visualised by staining with Coomassie brilliant blue. Samples taken prior to induction with IPTG (T_0) and samples taken at 1, 2, 3, 4 and 16 hours following induction (T_1, T_2, T_3, T_4 and T_{16} respectively) are shown together with the total lysate and soluble fraction of the lysed cell pellet from the full 50 ml culture following the 16 hour expression. Samples were run against a 14-66 kDa molecular weight marker (Sigma-Aldrich) to assess the size of the protein produced. Soluble protein can clearly be seen at approximately 27 kDa in both trial experiments.
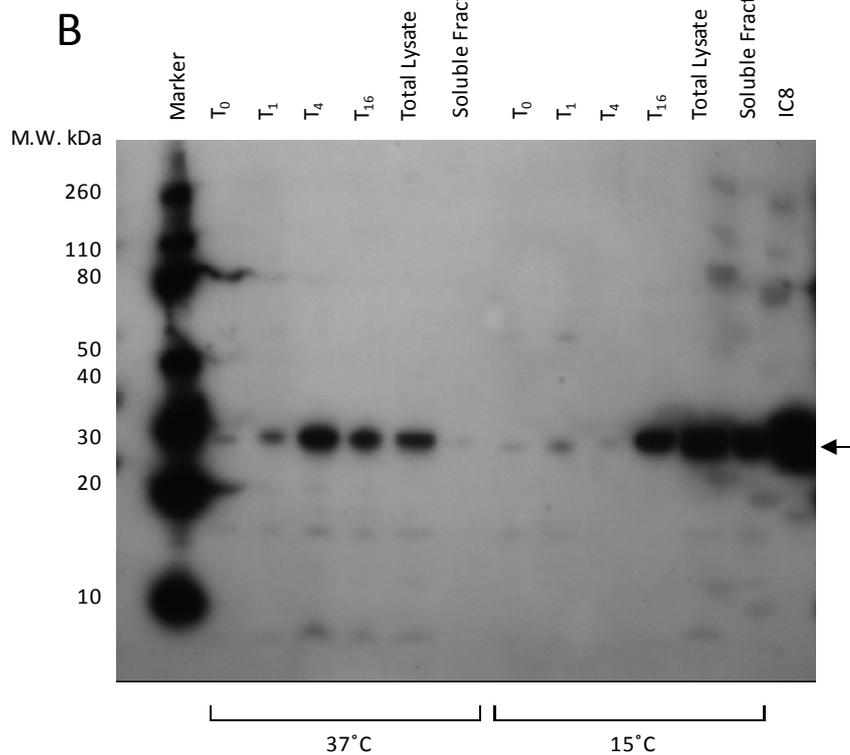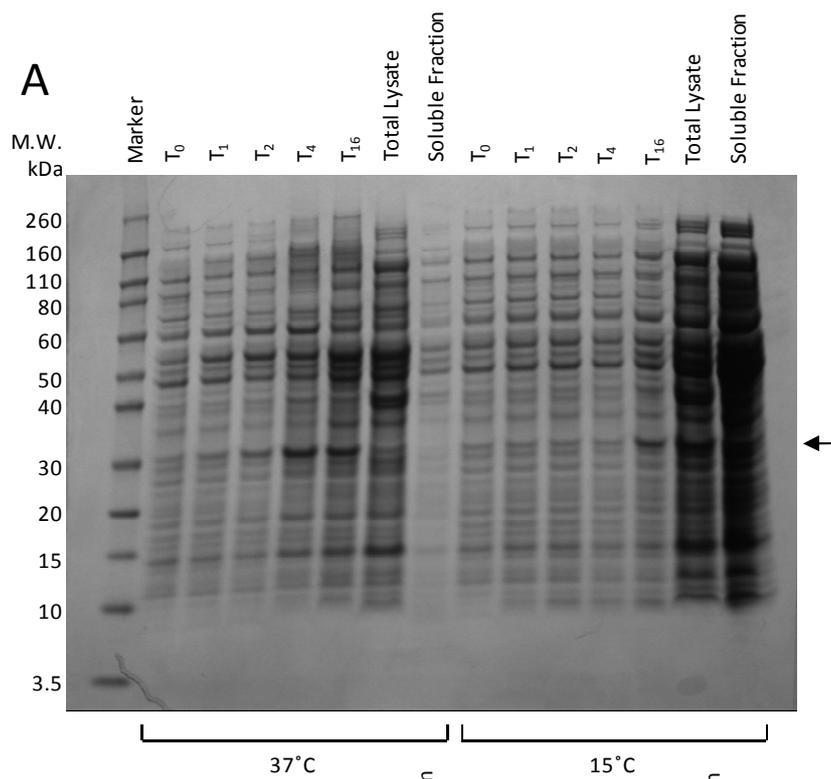
Figure 2.3.2 – Samples taken from the expression trials of 488scFv analysed using SDS-PAGE visualised both by staining with Coomassie brilliant blue (A) and by

western blot of the histidine tag (B). Samples taken prior to induction with IPTG (T_0) and samples taken at 1, 2, 4 and 16 hours following induction (T_1, T_2, T_4 and T_{16}) (A) or samples taken at 1, 4 and 16 hours following induction (T_1, T_3 and T_{16}) (B) are shown together with the total lysate and soluble fractions of the lysed cell pellets following 16 hours of expression. Samples were run against a Novex Sharp molecular weight marker (Invitrogen) to assess the size of the protein produced (A and B) and a previously expressed and purified histidine tagged scFv as a positive control (B only). The scFv is clearly visible at a molecular weight of approximately 27 kDa at both trial temperatures. Soluble expression for this 488scFv construct was only seen at 15 ˚C.
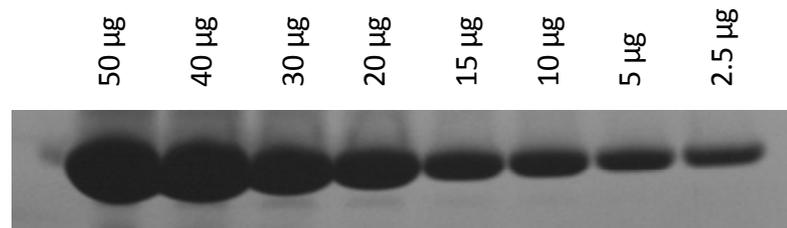
Figure 2.3.3 – Coomassie stained SDS-PAGE of known quantities of lysozyme. The intensity of the bands seen allows the estimation of protein quantities from other Coomassie stained SDS-PAGE gels. (Image obtained from Alice Barkell).

2.3.2.2 – Standard Cell Density Expression and Purification of unlabelled, ^{15}N labelled and ^{15}N/^{13}C labelled scFvs

The cultures grown to produce the scFvs for further analysis grew to an A_{600} of 3-4 in LB and 2-2.5 in minimal medium, showing levels of protein production similar to the expression trials when assessed by SDS-PAGE. The expressed scFvs were purified following the protocol described in section 2.3.1.8. The protein eluted from the nickel affinity chromatography column contained mainly scFv along with some *E. coli* contaminant proteins that had also bound to the column (Figure 2.3.4). Further purification, using a pre-calibrated (Gel filtration calibration kit, low molecular weight, GE Healthcare) SEC column, removed contaminants to a level below that which could be seen on a Coomassie stained SDS-PAGE gel and highlighted the formation of the domain swap multimers by providing an estimation of the molecular weight of the proteins eluted (Figure 2.3.5). The collected fractions corresponding to the scFv containing peaks were pooled and stored at 4˚C. The amount of purified scFv produced was approximately 25 mg l^{-1} from LB and approximately 15 mg l^{-1} from minimal media for 1189scFv and 12 mg l^{-1} from LB and 5 mg l^{-1} from minimal media for 488scFv (as determined by absorbance at 280nm).

Figure 2.3.4 – A typical example of the nickel affinity purification of scFvs. This purification resulted in two overlapped elution peaks absorbing at 280 nm as seen on an FPLC system (A, 1 and 2). When analysed by coomassie stained SDS-PAGE against a
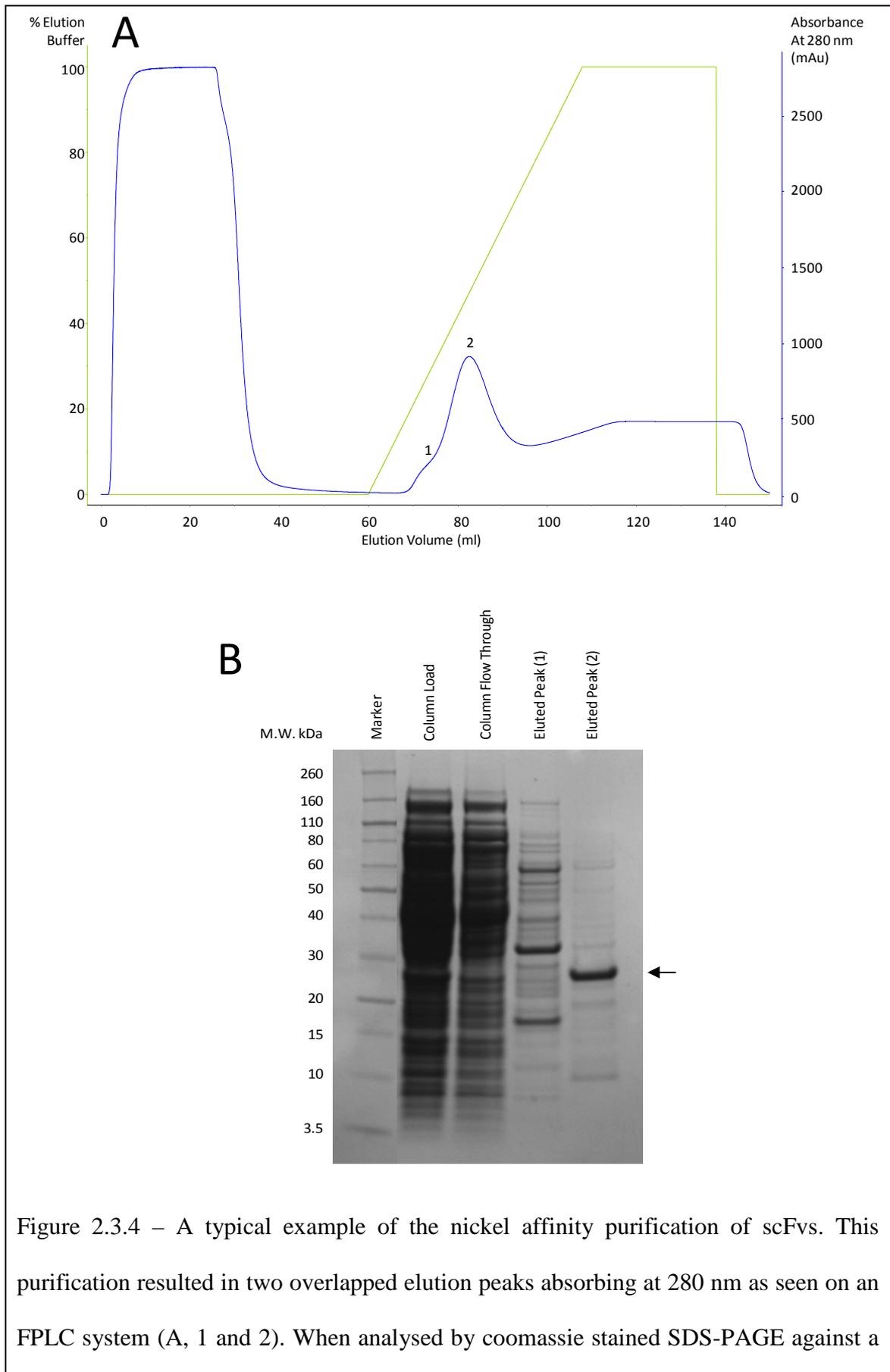
Novex Sharp marker it was seen that peak 1 contained a number of contaminant *E. coli* proteins that had bound to the nickel affinity column whereas peak 2 contained the expressed scFv with a low level of other contaminant *E. coli* proteins. The column flow through sample indicates that all soluble scFv was recovered during this process.

Figure 2.3.5 – A typical example of the SEC purification of the scFvs. This process resulted in the detection of three elution peaks absorbing at 280 nm as seen on an FPLC system (A, 1, 2 and 3). When analysed by coomassie stained SDS-PAGE against Novex
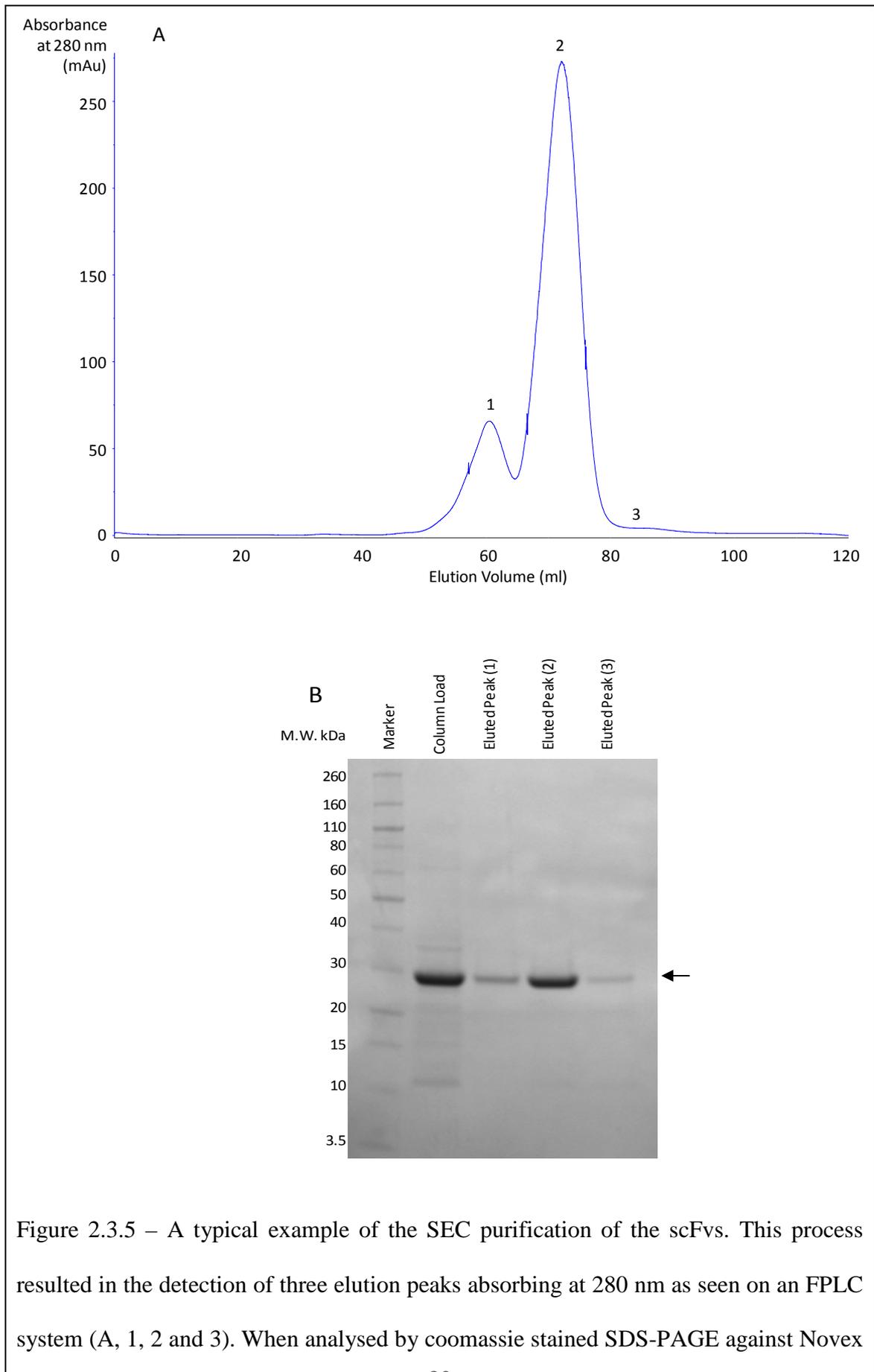
Sharp molecular weight standards (B) it was seen that peak 1 (approximately 45 kDa) contained scFv in a form roughly corresponding to the molecular weight expected for domain swapped dimer, peak 2 (approximately 26 kDa) contained the apparently monomeric scFv and peak 3 (approximately 11 kDa) contained no visible protein other than the remainder of peak 2. The fractions collected for peak 1 and 2 were pooled and stored at 4˚C.

2.3.2.3 – Standard Cell Density Expression and Purification of Deuterated scFvs

The expression levels seen in the production of triply labelled scFv were severely reduced when compared to non-deuterated minimal media. Growth rates and expression levels were impaired, with cultures failing to exceed a final A_{600} of 1.8-2.0 after overnight expression and expressed scFv only being visible on western blots. The resulting protein was purified as described for the non-deuterated scFv samples and partially denatured using 5 M guanidine hydrochloride. This denaturation allowed the back exchange of the backbone amide deuterons for protons in the core of the scFv structure that are normally protected from solvent exchange. The total protein yield was approximately 0.5 mg l^{-1} of deuterated culture as quantified by absorbance at 280 nm. Despite these low expression levels, sufficient protein was generated to form an NMR sample due to the large culture volumes used (approximately 10 l).

2.3.2.4 – Standard Cell Density Expression and Purification of 1189Fab

Unlabelled 1189Fab was expressed and purified by protein G (Figure 2.3.6) as described in sections 2.3.1.7 and 2.3.1.9. A final yield of approximately 5 mg l^{-1} was produced as quantified by absorbance at 280 nm ($\varepsilon = 47890$ M^{-1} cm^{-1}).

Figure 2.3.6 – Protein G purification of 1189Fab using an FPLC system. The separation

and elution of the protein resulted in an elution peak absorbing at 280 nm (A, 2). When
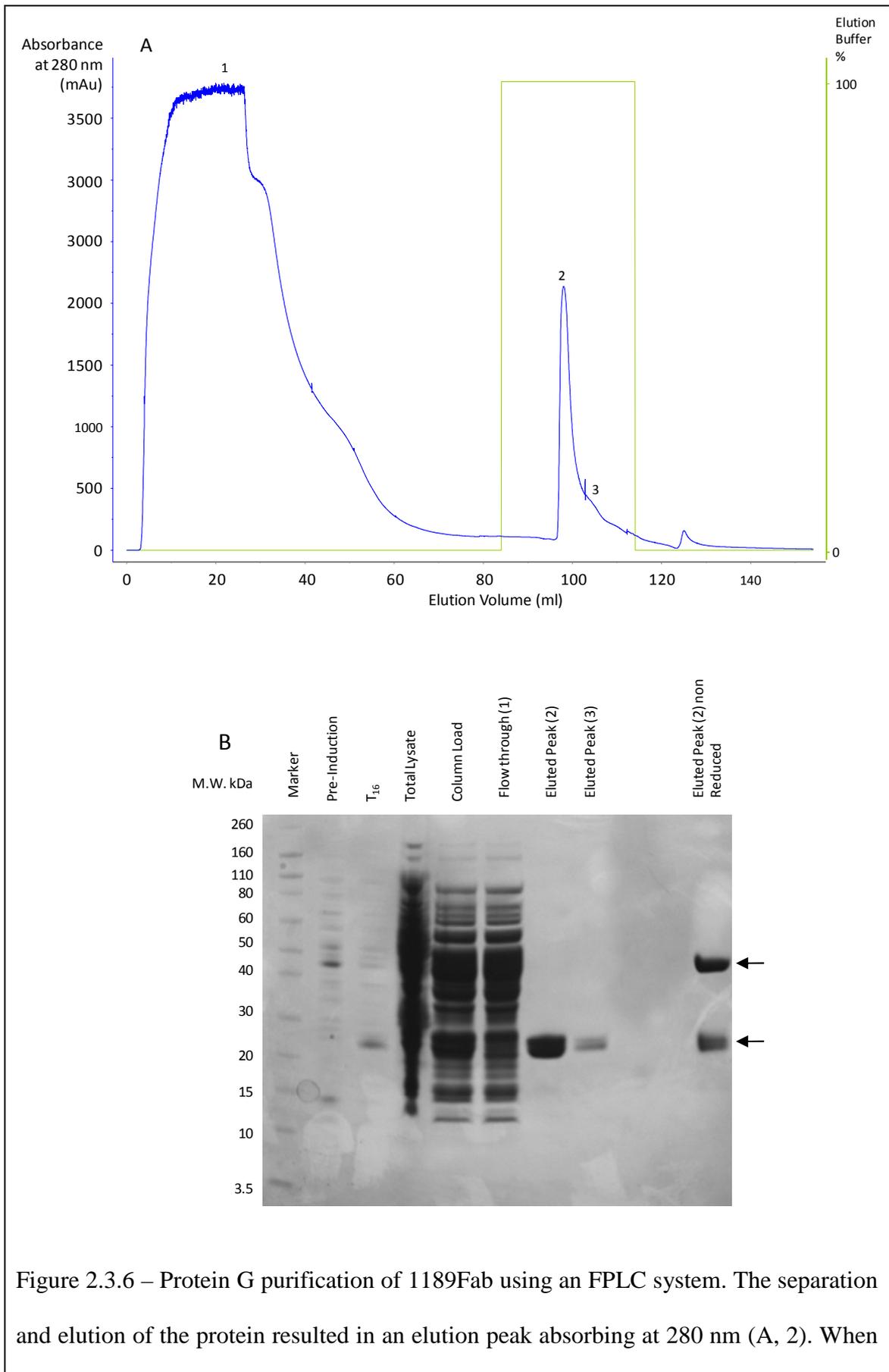
analysed by Coomassie stained SDS-PAGE against Novex Sharp molecular weight marker a characteristic double band representing the heavy and light chain of the Fab could be seen. In addition a sample was run without the addition of the DTT to retain the inter chain disulphide bond. The band shift seen to the molecular weight expected for 1189Fab (47 kDa) further confirms successful purification.

2.4 – Small Volume, High Cell Density Protein Expression

2.4.1 – Materials and Methods

2.4.1.1 – Selection of High Expressing colonies

The initial step in the high density method is to doubly select transformed colonies that show the highest level of protein expression. Glycerol stocks of these selected cells will create a reliable starting point for the production of *E. coli* cultures expressing large quantities of protein (Sivashanmugam *et al.* 2009). W3110 *E. coli* were transformed with either the 1189scFv or the 488scFv plasmid and plated out onto 70 % D_2O LB agar plates containing 10 μg tetracycline and grown at 37 ˚C overnight. Six small expression trials of separate colonies were set up with 2 ml starter cultures (100 % D_2O LB with tetracycline) and 10 ml trial cultures (standard 100 % D_2O minimal medium used in large scale expression, with tetracycline). Time point samples were taken and treated as described in section 2.3.1.1 and analysed using SDS-PAGE/western blot. Once completed, the highest expressing trial culture was streaked out onto a fresh 70 % D_2O plate and grown overnight. Six of these new colonies were analysed with the highest expressing culture being used to prepare glycerol stocks.

2.4.1.2 – Protein Expression Using the High Cell Density Method

The double selected glycerol stocks were used to inoculate 50 ml 100 % D_2O LB cultures containing 10 μg l^{-1} tetracycline. These were grown overnight to an A_{600} of 3-5

at 37 ˚C in D$_2$O (lower temperatures were required if the cultures were over growing). Once at an appropriate A$_{600}$, the cultures were centrifuged at 5000 x g for 15 minutes without refrigeration and the cell pellet resuspended in the high density labelling medium as specified by Sivashmanmugam *et al* (2009) (Appendix A.1.2), containing ^{15}N ammonium chloride and/or ^{13}C$_6$ D-glucose (where appropriate) in D$_2$O containing 10 μg l^{-1} tetracycline. The resuspended culture was returned to the incubator for 1-3 hours to allow the A$_{600}$ to increase by one unit. Once achieved, IPTG was added and the cultures placed in an incubator according to the concentration and temperature determined for each construct in 2.3.1.1. The cultures were allowed to express until the A$_{600}$ began to decrease or the protein appeared to degrade (determined by SDS-PAGE analysis of culture time points). When using antibiotics, such as ampicillin, that are degraded by resistance mechanisms or the pH of the culture medium, the addition of extra antibiotic is required regularly throughout the expression run to prevent plasmid loss (Dever & Dermody, 1991; Sivashanmugam *et al*, 2009). The resistance mechanism used here for the scFv vectors employs an efflux pump located in the cell membrane to keep tetracycline concentrations low within the bacterial cells and, as it is not degraded, no additional antibiotic is required (Schnappinger & Hillen, 1996). Initially this protocol tested protein expression in 50 and 100 ml cultures, in 250 and 500 ml baffled flasks respectively, to optimise aeration before moving to 50 ml cultures once the protocol was optimised.

2.4.1.3 – Specific Non-Labelling of Amino Acids

The method outlined in 2.4.1.2 was adapted to specifically non-label certain methyl and aromatic side chain amino acid types. This process was trialled on 1189scFv with

leucine, isoleucine, valine and tryptophan/tyrosine but may also be done with phenylalanine. Amino acid, of the type to be non-labelled, was added at a concentration of 200 mg l^{-1} in the high density labelling medium. This concentration was based on previous standard cell density expression protocols of a similar nature (Vuister *et al*, 1994), and compensates for an up to ten-fold higher final A_{600} in the high cell density protocol, with a ten-fold higher concentration of the unlabelled amino acids compared to the 20 mg l^{-1} used in the original protocol. Apart from the additional amino acid, the same method and stocks were used as described in the previous section.

2.4.2 – Results

2.4.2.1 – Selection of High Expressing Colonies

W3310 *E. coli* colonies freshly transformed with pTTod scFv vector for 1189scFv and 488scFv underwent the two step small scale expression trial process described in section 2.4.1.1 and were analysed by western blotted SDS-PAGE (section 2.3.1.3). The variability of expression levels was more apparent in the first step of colony selection than the second (Figure 2.4.1), indicating that consistently high expression levels were being successfully selected for. The improvement in expression levels was seen for both 1189scFv and 488scFv.
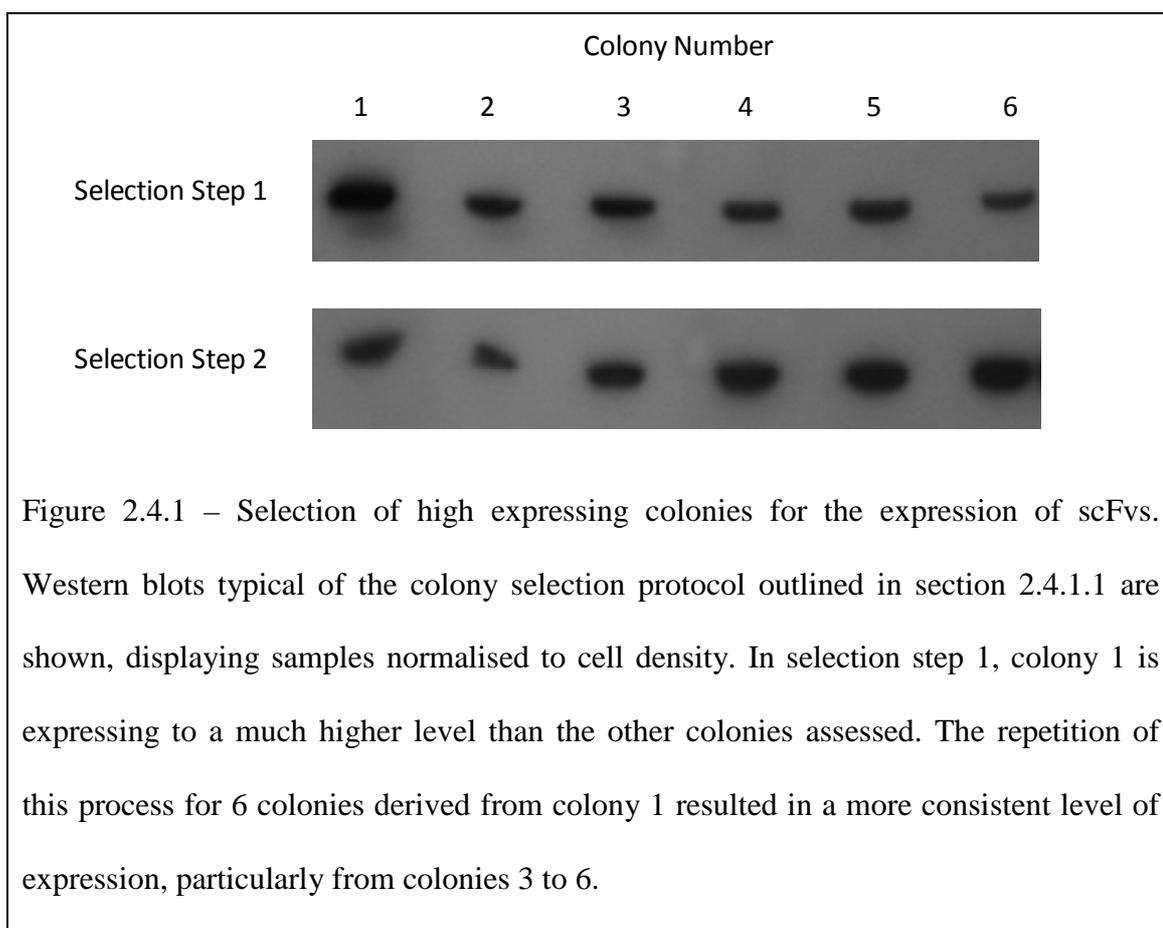
Figure 2.4.1 – Selection of high expressing colonies for the expression of scFvs. Western blots typical of the colony selection protocol outlined in section 2.4.1.1 are shown, displaying samples normalised to cell density. In selection step 1, colony 1 is expressing to a much higher level than the other colonies assessed. The repetition of this process for 6 colonies derived from colony 1 resulted in a more consistent level of expression, particularly from colonies 3 to 6.

2.4.2.2 – Protein Expression Using the High Cell Density Method

High cell density protein expression was initially trialled using 50 and 100 ml cultures, in 250 and 500 ml baffled flasks, with the aim of achieving optimal aeration. Both sets of trials were successful in producing high yields of 1189scFv and 488scFv (estimated to be 60 mg l^{-1} and 30 mg l^{-1} respectively, see section 2.4.2.1), which could clearly be seen by Coomassie stained SDS-PAGE (Figure 2.4.2). These trials also indicated that expressing the protein in 50 ml cultures in 250 ml baffled flasks produced the optimal amount of protein. The duration of expression was also optimised during this trial process by monitoring the A_{600} of the cultures at 0, 2, 4, 16 and 24 hours after induction with IPTG for 1189scFv, and 0, 4, 16, 24, 40 and 48 hours following induction for 488scFv. As no evidence of protein degradation was seen over the course of the expression, the optimal amount of protein would be obtained at the highest cell density, and so at the time point before the A_{600} begins to drop. For 1189scFv and 488scFv this was 16 and 40 hours respectively.

The scFvs were purified using the same protocols described for the standard cell density expression procedure described in section 2.3.1.8. The increase in deuterated protein production was particularly significant, with final purified protein yields of 50 mg l^{-1} for 1189scFv and 25 mg l^{-1} for 488scFv. These yields were significantly below those reported by Sivashanmugam *et al* (2009), but represent a 100 fold increase when compared to the standard cell density minimal medium expression method. This reduced both the quantities of expensive, isotopically labelled media components required, and the time taken to get the amounts of protein needed for NMR spectroscopy.
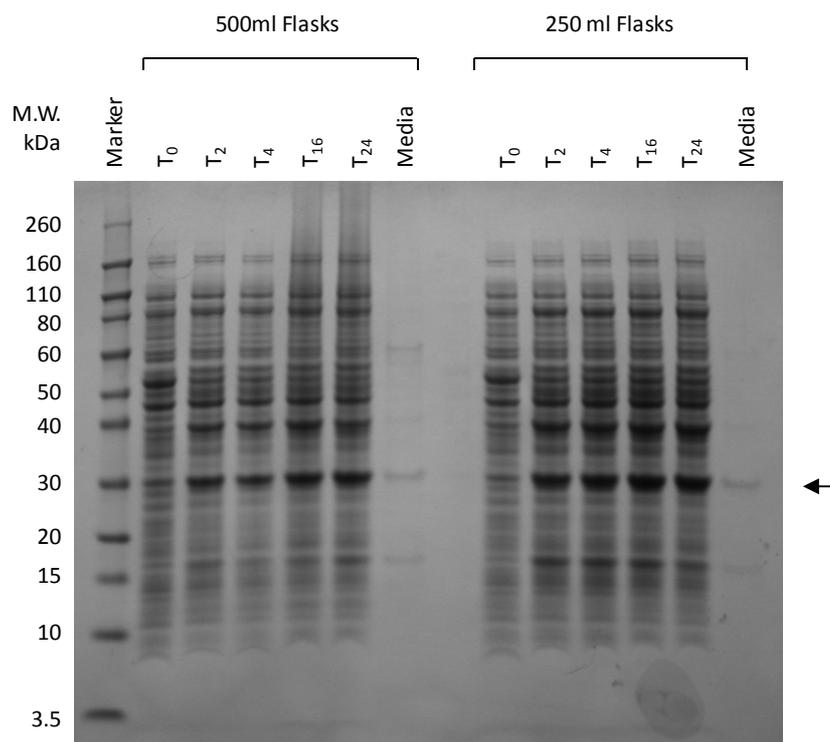
Figure 2.4.2 – Small scale, high cell density expression trials of 1189scFv in 250 and 500 ml flasks. Analysis by coomassie stained SDS-PAGE of expression trial samples taken at 0, 2, 4, 16 and 24 hours following IPTG induction (T_0, T_2, T_4, T_{16} and T_{24}) clearly show protein bands at the appropriate molecular weight (27 kDa). All samples were normalised to an A_{600} of 0.8 to assess protein degradation in the bacterial cells. At all time points a significant amount of protein was produced even up until the A_{600} of the culture began to drop at 24 hours. A sample of the centrifuged culture medium was also run to determine if any protein had leaked out of the cells during expression (Media). The low levels of protein visualised in this sample most likely represent protein released as the cells began to lyse. From this analysis it was determined that expression for between 16 and 24 hours following the addition of IPTG in 50 ml cultures in 250 ml baffled flasks produced the highest protein yield. All samples were run against Novex Sharp molecular weight standards (Marker).

2.4.2.3 – Specific Non-Labelling of Amino Acids in ^{15}N/^{13}C/^{2}H labelled 1189scFv samples

The final yields of protein obtained when introducing unlabelled amino acids into the labelling media of the high cell density protocol (as described in section 2.4.1.1) were slightly lower than those seen with the un-modified method. Protein yields for the 1189scFv construct were between 30 and 40 mg l^{-1}, depending on the amino acid type used. A set of four ^{15}N/^{13}C/^{2}H samples were produced with valines, leucines, isoleucines or tryptophans/tyrosines specifically not labelled.

2.5 – Protein characterisation

2.5.1 – Materials and Methods

2.5.1.1 – Circular Dichroism Spectroscopy

Far UV Circular Dichroism (CD) spectroscopy was performed using a Jasco J-715 spectropolarimeter at wavelengths between 190 and 250 nm, collecting 10 accumulations per spectrum at a resolution of 1 nm and a scan rate of 50 nm minute^{-1}. All spectra were corrected for buffer contribution, by the subtraction of a buffer only spectrum, and converted into molar CD per residue based upon the protein concentration. The corrected CD spectra were analysed using the CD Pro software package (Sreerama & Woody, 2000) to determine the secondary structure composition of the sample. All scFv samples were analysed in a buffer comprising of 100 mM sodium fluoride, 25 mM sodium acetate, pH 5.5. Sodium fluoride was used to maintain the ionic strength of the buffer, whilst removing chlorine from the sample to improve signal to noise at low wavelengths. Thermal denaturation curves were also produced monitoring the change in Circular Dichroism between 20 and 95 °C at a number of wavelengths.

2.5.1.2 – Fluorescence Spectroscopy

Intrinsic tryptophan fluorescence was used to monitor the unfolding of the 1189scFv in increasing concentrations of guanidine hydrochloride (0 – 6 M in 0.5 M increments). Fluorescence spectra were collected using a Perkin Elmer LS50B luminescence spectrometer, with excitation at 280 nm and collection between 300 and 450 nm at a scan rate of 150 nm min^{-1}, for a total of five accumulations. Samples were 2 μM scFv in 100 mM sodium chloride, 25 mM sodium acetate pH 5.5 and were allowed to equilibrate following the addition of guanidine hydrochloride for 30 minutes prior to data collection.

2.5.1.3 – NMR Spectroscopy of the scFvs

Preliminary ^{15}N/^{1}H HSQC (Bodenhausen & Ruben, 1980) spectra were collected from ^{15}N labelled 1189scFv and 488scFv at 40˚C on an 800 MHz Bruker Avance or a 600 MHz Bruker DRX spectrometer. Samples were prepared in SEC buffer and concentrated to 230 μM by ultrafiltration (Vivaspin 6, Sartorius) and analysed in 5 mm Shigemi NMR tubes. Typical acquisition times were 35 ms in F_1 (^{15}N) and 60 ms in F_2 (^{1}H) with a total experiment time of 40-60 minutes.

Solvent suppression was carried out using the WATERGATE pulse scheme (Piotto *et al*, 1992) and data was processed using the Topspin 2.1 software (Bruker Biospin Ltd.).

2.5.1.4 – Analytical Size Exclusion Chromatography

Analysis of the effects of protein concentration on the scFv domain swap equilibrium was performed for 1189scFv on a pre-calibrated (gel filtration calibration kit, low molecular weight, GE Healthcare) Superdex 75 10/300 GL size exclusion column (GE Healthcare) operating at a flow rate of 1 ml minute^{-1}. Samples were run at 7.5, 15, 30, 140 and 530 µM in a 100 mM sodium chloride, 25 mM sodium acetate buffer at pH 5.5 containing 0.02 % sodium azide. Additional 7.5 µM samples to study the effect of pH on the multimerisation process were also run in the same buffer at pH 5.0, 4.5, and 4.0. An additional sample was run in buffer containing 25 mM Tris pH 7.5 instead of sodium acetate. In all cases, the column was equilibrated with two column volumes of the appropriate buffer before use. Estimates of the proportions of monomer, dimer and trimer were determined from the area of the elution peaks using the peak integration function contained within the UNICORN 5.11 software on the Akta FPLC system (GE Healthcare).

2.5.2 – Results

2.5.2.1 – Circular Dichroism Spectroscopy of scFvs

The CD spectra acquired for the scFvs displayed the characteristics of a predominantly β-sheet protein (Figure 2.5.1). Spectral analysis using the CD pro package (Sreerama & Woody, 2000) indicated that the secondary structure composition of the scFvs was approximately 1 ±2 % α-helix, 49 ±3 % β-sheet, 22 ±2 % turn and 28 ±2 % unstructured. This data is consistent with the expected secondary structure of 3 % α-helix, 50 % β-sheet and 47 % turn/unstructured (taken from the structure of IC8scFv, Wilkinson *et al*, 2009), indicating that the scFvs were correctly folded.

Figure 2.5.1 – Far UV CD spectrum of 1189scFv. This data was collected from a 3 μM sample of the scFv at 25 ˚C. The spectrum displays the characteristic features of a β-sheet protein showing a negative peak at 218 nm (Kelly *et al*, 2005). The small positive peak at 235 nm is most likely a contribution the intra-domain disulphide bonds. The data has been buffer corrected, converted to molar CD per residue and smoothed.

To further characterise the properties of the scFvs, CD thermal denaturation curves were collected monitoring at 220 nm between 20 and 95 ˚C. This data was used to determine if the scFvs showed co-operative unfolding, a characteristic property of correctly folded proteins. The sigmoidal unfolding curves observed for both 1189scFv and 488scFv represent co-operative unfolding events, further confirming that the scFvs are correctly folded. The mid points of the denaturation curves were 80 ˚C for 1189scFv and 69 ˚C for 488scFv. A typical temperature denaturation curve for each scFv is shown in Figure 2.5.2.

96
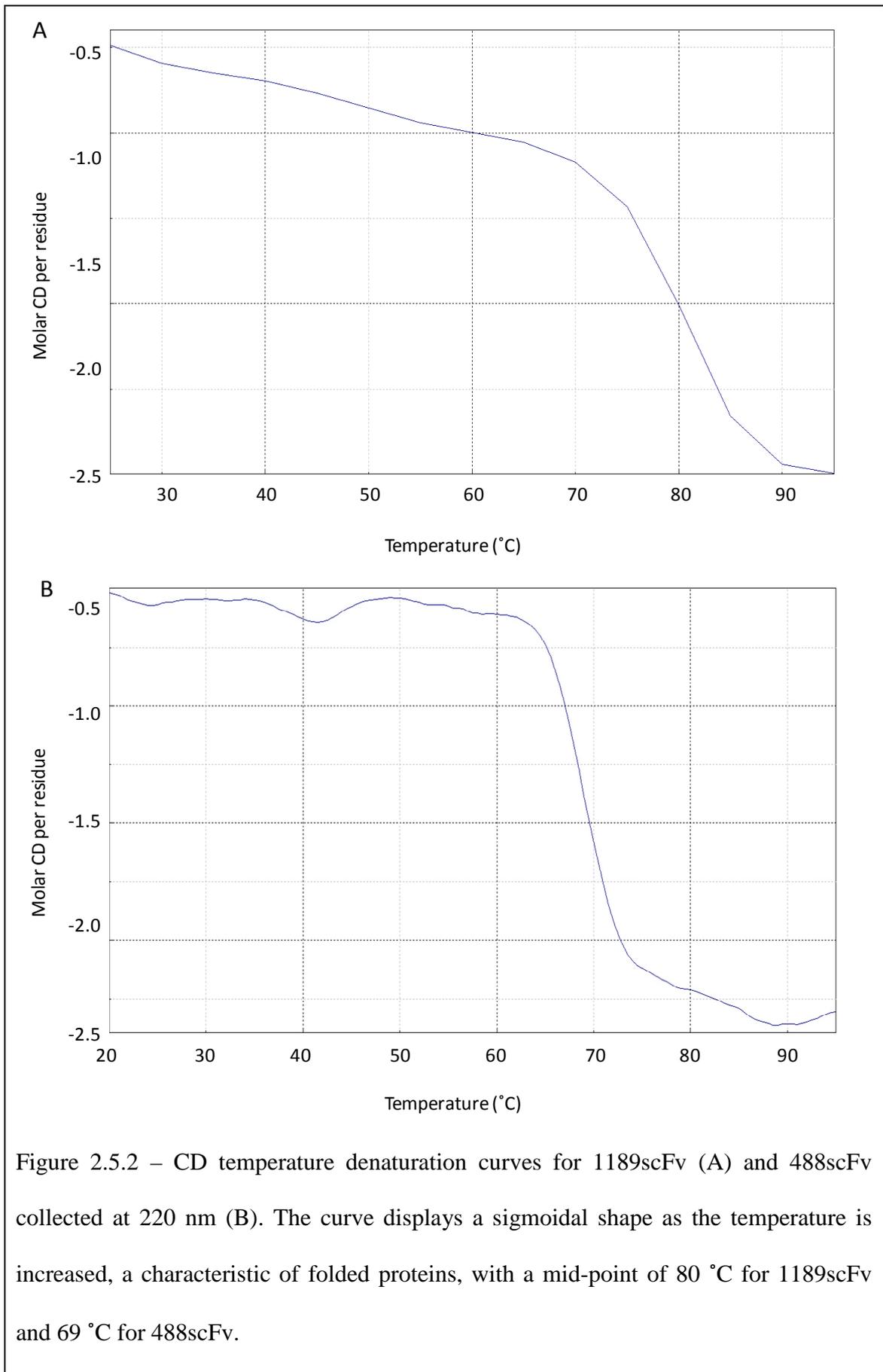
Figure 2.5.2 – CD temperature denaturation curves for 1189scFv (A) and 488scFv collected at 220 nm (B). The curve displays a sigmoidal shape as the temperature is increased, a characteristic of folded proteins, with a mid-point of 80 ˚C for 1189scFv and 69 ˚C for 488scFv.

2.5.2.2 – Fluorescence Spectroscopy of 1189scFv

Intrinsic tryptophan fluorescence spectra of proteins are often sensitive to changes in the chemical environment of tryptophan residues, particularly with respect to the solvent accessibility of the side chains (Royer, 2006). Florescence spectra were collected for 1189scFv, and the change wavelength of the fluorescence maxima (F_{max}) was monitored as the concentration of guanidine was increased (Figure 2.5.3). The changes seen in the F_{max} of the samples were used to follow the unfolding of the protein from 327 nm to 348 nm as the buried tryptophan residues became more solvent exposed. The sigmoidal shape of the curve is characteristic of stable, folded protein with a mid-point at approximately 4.2 M guanidine hydrochloride and the protein becoming fully denatured between 5.0 and 5.5 M. The guanidine hydrochloride induced denaturation of the scFv was found to be fully reversible. The removal of the denaturant by dialysis resulted in spectra identical to those collected for the original native protein sample. This information indicates that partial denaturation of the scFv should allow the solvent exchange of buried deuterons for protons (as described in section 2.3.1.8) to recover backbone amide NMR signals for the core regions of deuterated scFv samples.

Figure 2.5.3 – Chemical denaturation of 1189scFv. The F_{max} wavelength of the intrinsic tryptophan fluorescence emission spectra collected from 1189scFv is shown against increasing concentrations of guanidine. The sigmoidal curve, characteristic of co-operative unfolding events, further confirms the correctly folded nature of the scFv.
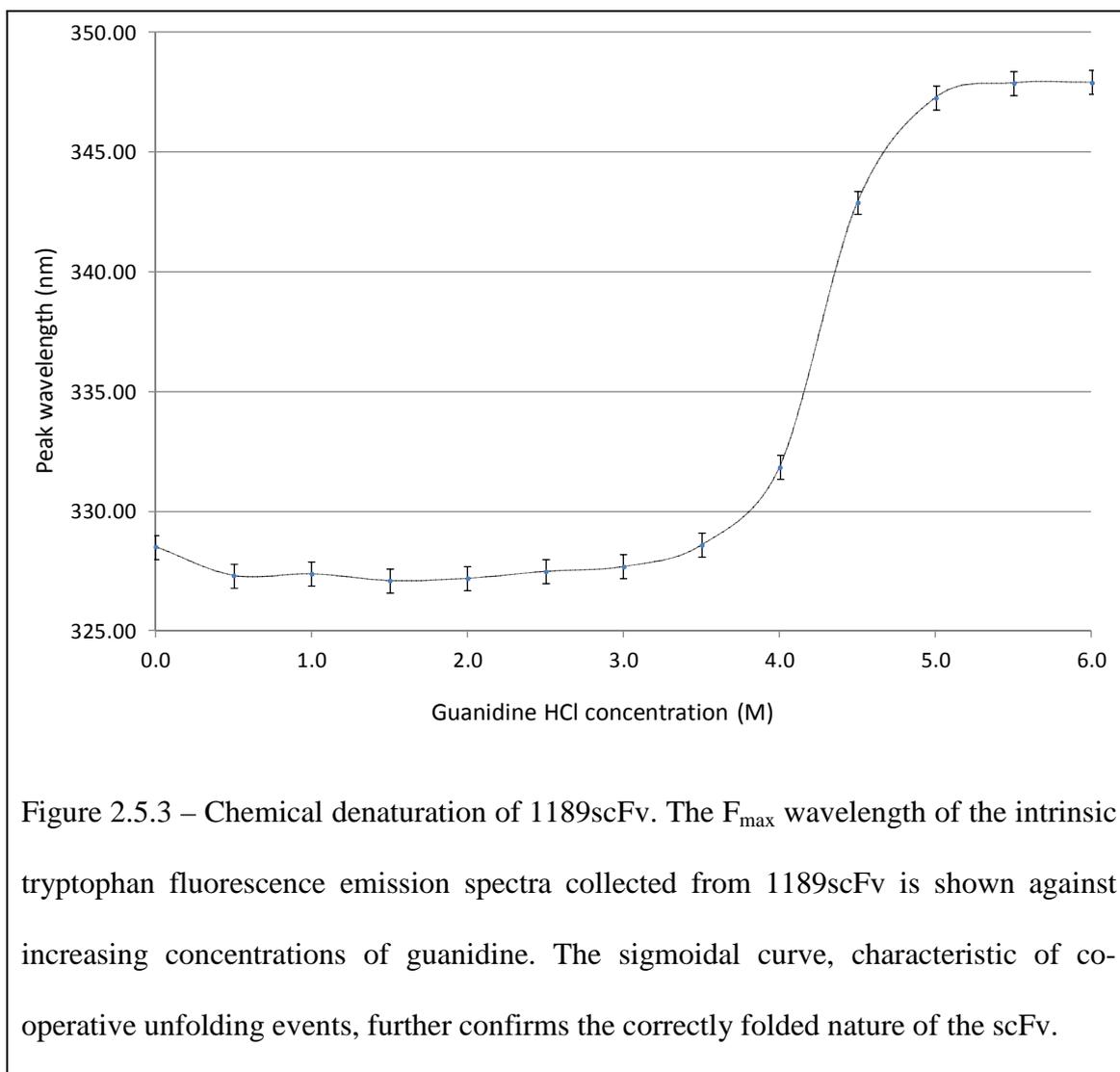
2.5.2.3 – NMR Spectroscopy of scFvs

Preliminary $^{15}N/^{1}H$ HSQC spectra of ^{15}N labelled 1189scFv and 488scFv produced well dispersed spectra, with high levels of signal to noise and very little signal overlap (Figure 2.5.4). The high quality of the spectra and dispersion of the signals seen indicate that the scFvs are folded and are suitable for further study by NMR.

Figure 2.5.4 – ^{15}N/^{1}H HSQC spectra of 1189scFv (A) and 488scFv (B). The spectra, collected at 40 ˚C on an 800 MHz (A) or 600 MHz (B) spectrometer, are of a high quality, showing well dispersed signals and good signal to noise. These spectra are characteristic of folded β-sheet proteins and indicate that further NMR data may be gathered to investigate structural properties of the scFvs.

2.5.2.4 – Analysis of the oligomerisation state of 1189scFv

Figure 2.5.5 shows results typical of the SEC analysis of the oligomerisation state of 1189scFv over a range of protein concentrations, with results of the proportions of the oligomeric states of 1189scFv at varying concentrations and pH values, summarised in Table 2.5.1 and Table 2.5.2.
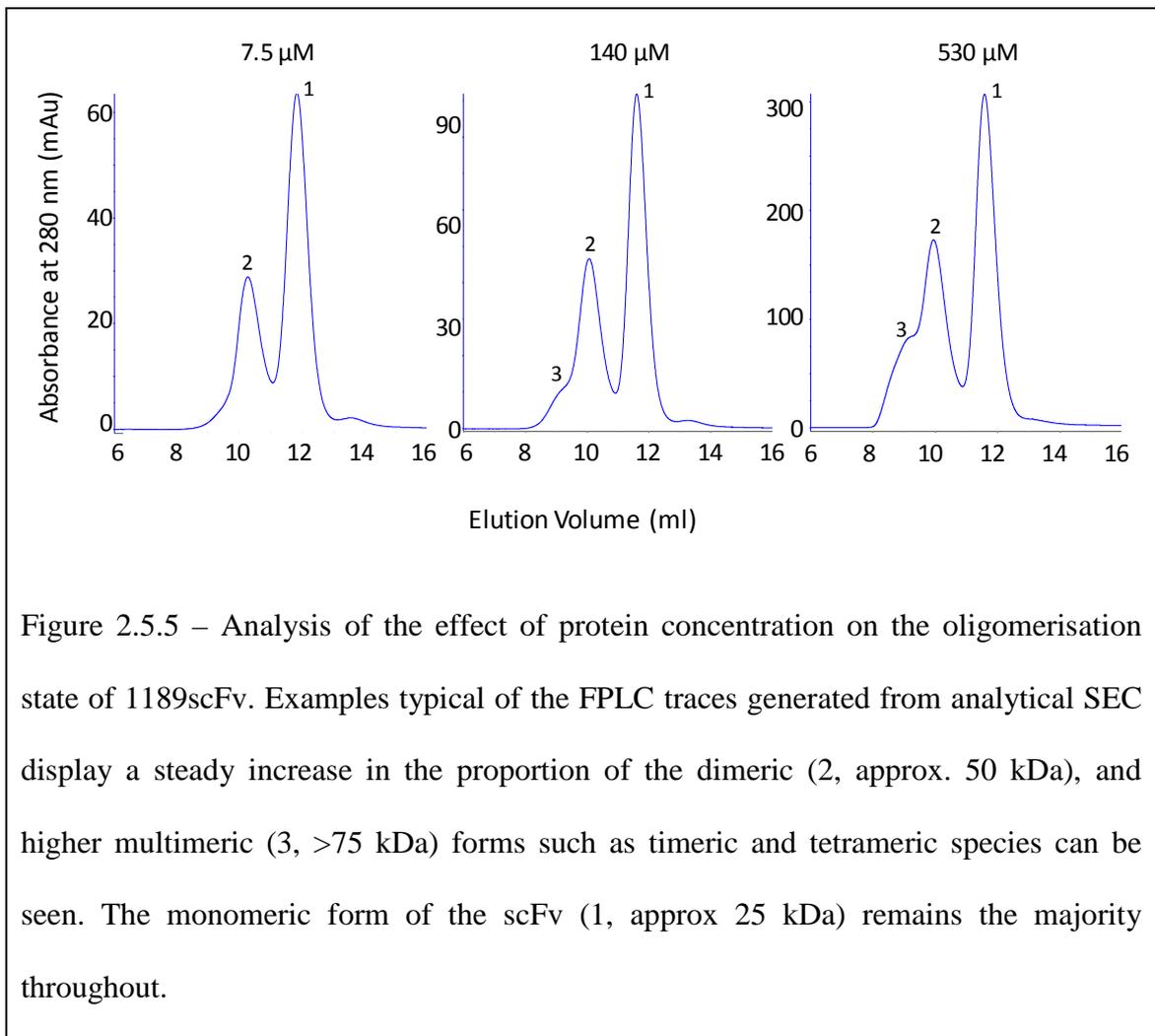
Figure 2.5.5 – Analysis of the effect of protein concentration on the oligomerisation state of 1189scFv. Examples typical of the FPLC traces generated from analytical SEC display a steady increase in the proportion of the dimeric (2, approx. 50 kDa), and higher multimeric (3, >75 kDa) forms such as timeric and tetrameric species can be seen. The monomeric form of the scFv (1, approx 25 kDa) remains the majority throughout.

Concentration at pH5.5 (μM)	Monomer	Dimer	Trimer/Tetramer
7.5	65%	35%	-
15	63%	37%	-
30	63%	37%	-
140	58%	42%	-
265	57%	43%	-
530	52%	33%	15%

Table 2.5.1 – The effect of increasing protein concentration on the oligomerisation state of 1189scFv. This table reports the proportion of each form of 1189scFv as a percentage of the total scFv eluted from the SEC column. Even at very high concentrations the monomeric species remains in the majority.

pH at 7.5μM	Monomer	Dimer	Trimer/Tetramer
7.5	45%	30%	25%
5.5	65%	35%	-
5.0	58%	35%	7%
4.5	59%	33%	8%
4.0	60%	33%	7%

Table 2.5.2 – The effect of pH on the oligomerisation state of 1189scFv. This table reports the proportion of each form of 1189scFv as a percentage of the total scFv eluted from the SEC column. At pH values below the calculated isoelectric point (6.6) there appears to be relatively little change in the proportion of multimeric species. However, above the isoelectric point there seems to be a dramatic increase in the proportion of the trimeric/tetrameric forms.

The scFv shows relatively little change in the equilibrium between oligomerisation states as the concentration is increased, even up to levels used in NMR samples. pH also seems to have a relatively minor effect on the equilibrium when below the calculated isoelectric point of 1189scFv (6.6). A significant change is seen, however, when the pH rises above this value with a 20 % reduction in monomer, compared to pH 5.5 and a shift towards the trimeric/tetrameric species.

2.6 – Discussion

The results described in this chapter report the successful cloning and expression of two scFv constructs based on high affinity anti-IL-6 Fab fragments. The unsuccessful expression of two of the constructs (1160scFv and 271scFv) is in line with previous scFv expression trials performed within the group, which show a 50 % success rate when attempting to express scFv constructs of previously expressed Fab fragments in *E. coli*. It is unclear exactly why this is the case, as the sequences of the scFv constructs are very similar outside of the CDR loops (Figure 2.4.1), and the Fab fragments from which they are derived, are all readily produced in the same *E. coli* expression system. The expression levels of antibody fragments reported in the literature vary significantly (Bedzyk *et al.*, 1990; Carter *et al.*, 1992; Denzin *et al.*, 1991; Wörn & Plückthun, 1998; Wörn & Plückthun, 1999), and proper comparisons are difficult to draw due to the wide range of expression systems and antibody fragment types used. In an investigation into the yields of periplasmic antibody fragment expression (Knappik & Plückthun, 1995), it was shown that, despite a 75 % sequence identity between two antibody sequences, the levels of protein produced were variable when using the same expression system. It is likely that a conserved number of sequence differences that are present within the framework or the CDR loops of the proteins may have an overall effect on the stability of the protein or its ability to fold properly, subsequently affecting the overall expression levels.

Despite the problems experienced with 1160scFv and 271scFv, 1189scFv and 488scFv showed very good soluble expression levels in both rich and minimal media that were above what has been seen previously in the lab and in line with other, high expressing, scFvs seen in the literature such as those expressed by Wörn & Plückthun (1999). The application of the high cell density expression protocol for the deuteration of these proteins also produced levels of soluble scFv far in excess of the amounts possible with the standard expression methods, at a fraction of the cost. This reduction in cost is mostly attributable to the vastly reduced quantities of D_2O required, providing a more widely accessible route to the production of deuterated proteins. Further development of this expression system to produce protein samples with specifically unlabelled methyl containing amino acids, such as leucine, isoleucine, valine, phenylalanine and tyrosine/tryptophan, was also successful for use in assisting the backbone assignment process as part of NMR data analysis. These samples, that would have previously been useful but cost prohibited, were produced and used to provide further data to assist in the backbone assignment process for 1189scFv.

Characterisation of the scFvs by CD and fluorescence indicated that soluble, stable protein had been produced and demonstrated the high thermal and chemical stability of these molecules. Contrary to previous studies of scFvs performed in the group and elsewhere (Lee *et al*, 2002), the extent of oligomerisation shows relatively little dependence on the total concentration of scFv in the sample with the monomeric species remaining the majority, even when at NMR sample concentrations. The effect of pH on the domain swap equilibrium also had little effect apart from at pH 7.5 where

the formation of the dimeric and trimeric/tetrameric species was clearly favoured. These results indicate that the optimal pH for the monomeric form of the scFvs is 5.5.

The characteristics of the domain swap equilibrium and the extent of oligomerisation seen for 1189scFv are noticeably different to scFvs studied previously. Analytical SEC results indicate that, at NMR sample concentrations, the IC8scFv was shown to be up to 50 % multimer. 1189scFv, however, was shown to be approximately 60 % monomer at equivalent concentrations. Preliminary NMR experiments for the scFvs also show high quality, well resolved data that do not display the line broadening and signal overlap observed for IC8 scFv when free in solution. This improvement in spectral quality is likely to be an effect of the differing characteristics of the domain swap equilibrium, where exchange processes or the contribution from multimer signals are reduced. The effect of pH on this equilibrium was not a factor determined for IC8scFv, but has been shown to have some significance for 1189scFv. This feature of 1189scFv may well be related to the isoelectric point of the protein (calculated to be 6.6) as moving the pH of the protein sample above this value appears to have significant effects on the domain swap equilibrium. IC8, however, has a much higher isoelectric point (calculated to be 9.3) and would not have encountered a buffer with a pH higher than 8.0. It is possible that similar characteristics may be observed if IC8scFv were to be analysed outside of the normal pH range of NMR or protein purification buffers. These differing characteristics of the proteins are apparent despite only subtle differences between the framework regions of the two scFvs. Interestingly, a significant number of the less conserved differences between the 1189scFv and IC8scFv that lie outside of the CDR loops reside within or around the interface of the two variable domains. A55-T55, K66-Q66 and T106-N106 (1189scFv compared to IC8scFv, see Appendix 4 for Kabat numbering) all lie on the V_L domain interface whilst G194-R194 lies on the V_H

interface (Figure 2.6.2). These differences appear to make the IC8 interface slightly more polar, and the presence of R194 in particular may be significant as it will give the interface between the two domains a more charged nature. This could potentially make the interface more hydrophilic, favouring solvent exposure and the open forms of the scFv and promoting the domain swap exchange. K66 of the 1189scFv may also be able to form an ionic interaction with E257 (Figure 2.6.2) on the V_H domain, further stabilising the interface of the scFv. The buried surface area and strength of interaction of the interface between the two domains can vary significantly from one antibody to another, allowing the re-orientation of the variable domains (Stanfield *et al*, 1993). It is this varying strength of interface interaction that gives rise to the range of characteristics observed for the domain swap equilibrium (Arndt *et al*, 1998, Lee *et al*, 2002). This inherent instability, and propensity to form domain swap multimers regardless of the antibody type or selection method used, appears to be a key character of the scFv even when optimised to favour the monomeric form. Despite a large number of attempts to improve the stability of the variable domain interface, it has been difficult to produce a method suitable for all scFv constructs (Wörn & Plückthun, 2001). From a structural and functional point of view, the inherent instability of almost all scFv constructs implies that this is a fundamental feature of the two domains and may be important for antigen binding. The structural changes that binding brings about within the molecule may also be important for additional functions of the antibody molecule, such as the role it plays in BCR signalling.
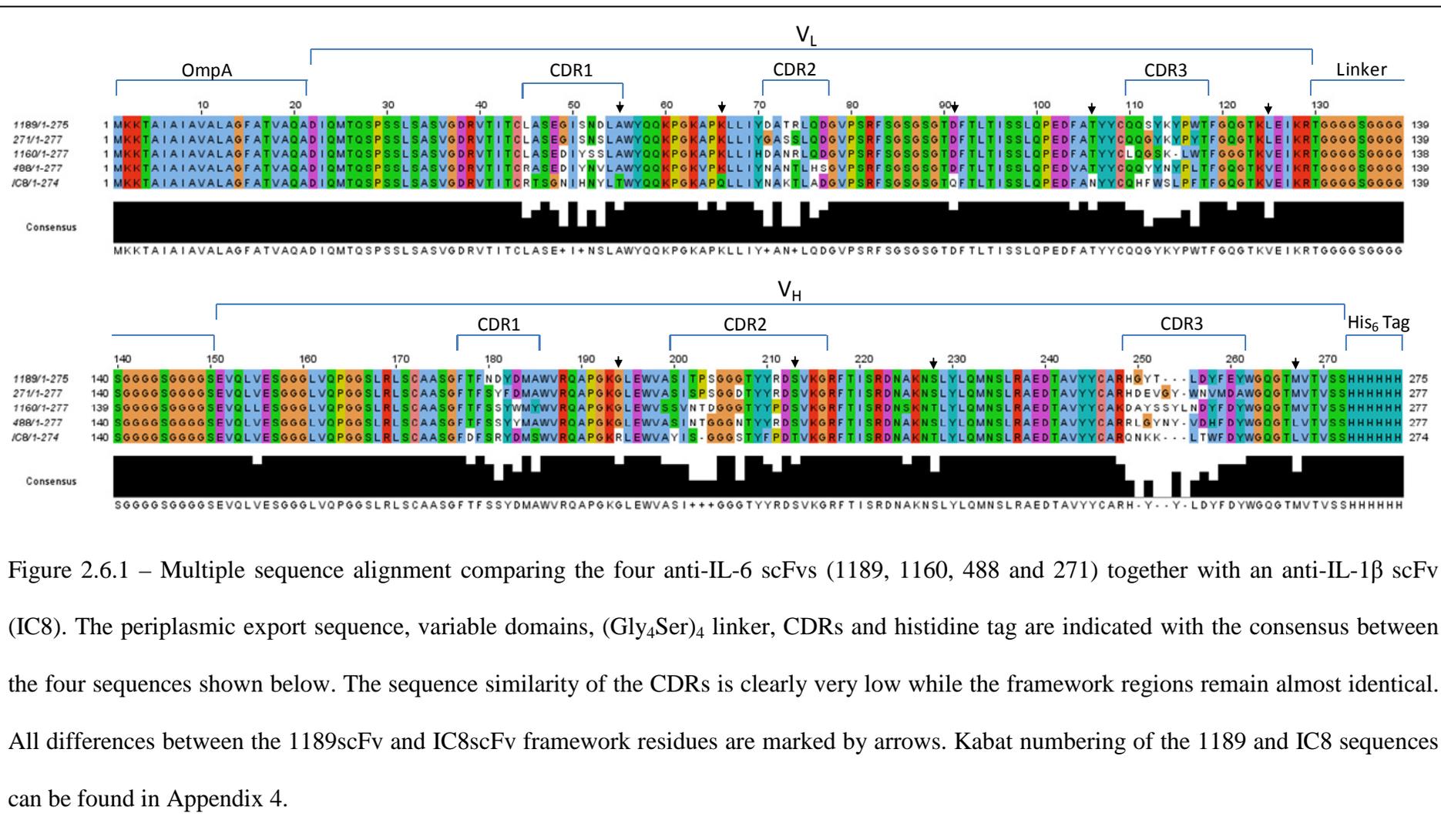
Figure 2.6.1 – Multiple sequence alignment comparing the four anti-IL-6 scFvs (1189, 1160, 488 and 271) together with an anti-IL-1β scFv (IC8). The periplasmic export sequence, variable domains, (Gly$_4$Ser)$_4$ linker, CDRs and histidine tag are indicated with the consensus between the four sequences shown below. The sequence similarity of the CDRs is clearly very low while the framework regions remain almost identical. All differences between the 1189scFv and IC8scFv framework residues are marked by arrows. Kabat numbering of the 1189 and IC8 sequences can be found in Appendix 4.
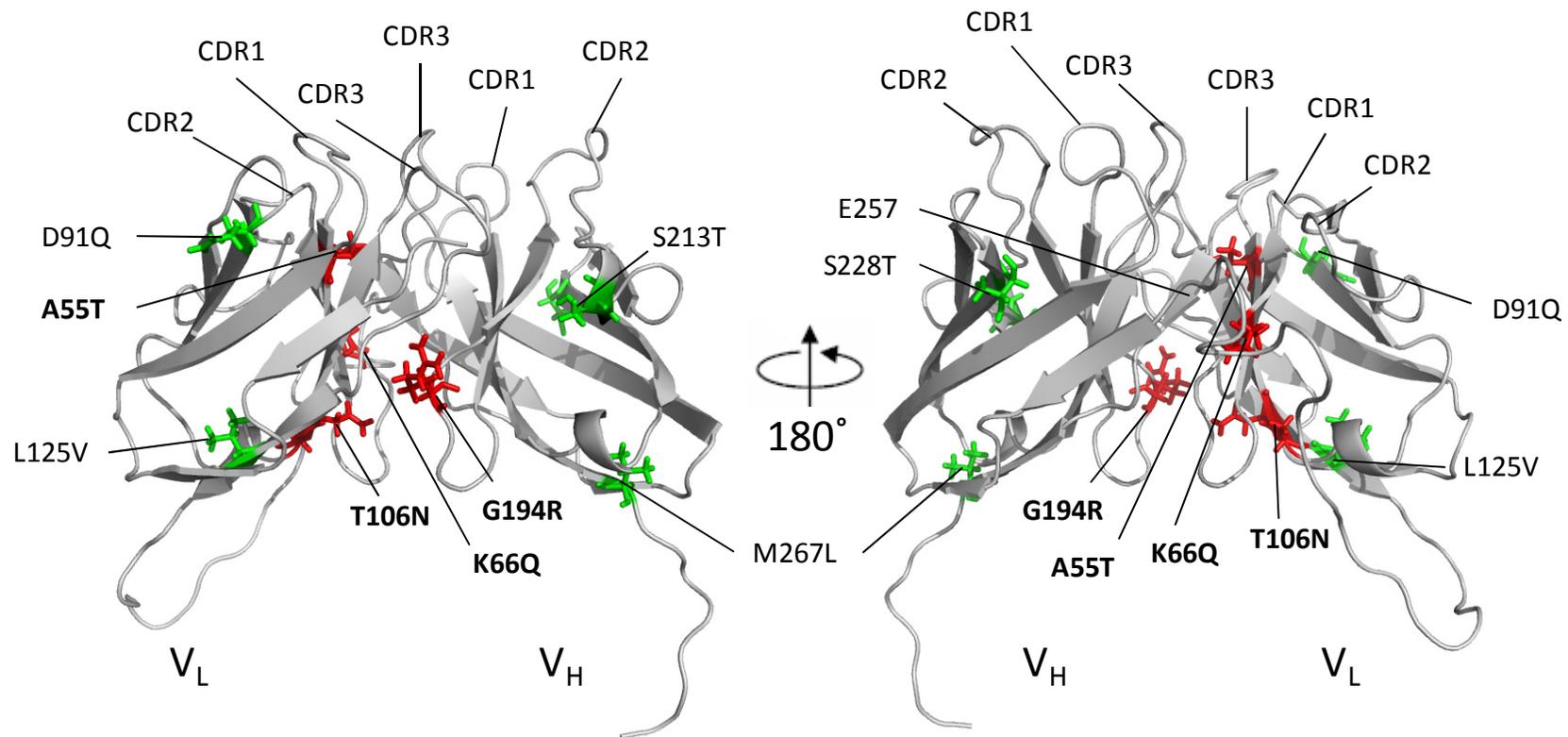
Figure 2.6.2 – The framework residue differences between 1189scFv and IC8scFv shown on the IC8scFv structure (PDB accession code 2KH2). The residues marked in red indicate differences located in the V_H/V_L interface that may affect the nature of the domain swap equilibrium. Other changes within the framework residues are shown in green. The Kabat numbering of these residues can be found in Appendix 4.1.

2.7 – Conclusions

The work reported in this chapter describes the cloning, expression and characterisation of a number of scFvs produced for the study of antibody/antigen interactions. Four scFv constructs were produced from high affinity, therapeutic candidate, anti IL-6 Fab fragment vectors with the soluble expression of two being confirmed and optimised in expression trials.

1189scFv and 488scFv were successfully expressed and purified with yields of up to 25 mg l^{-1} of unlabelled material in LB and up to 50 mg l^{-1} of ^{15}N/^{13}C/^{2}H labelled material in deuterated minimal media. The recently reported small volume, high cell density expression method was assessed, and proved essential for the production of deuterated NMR samples. This method was also extended to include the specific non-labelling of methyl containing amino acids that may aid the NMR data analysis process.

The expressed scFvs were characterised by CD and fluorescence spectroscopy to ensure that correctly folded protein had been produced, and to determine the chemical and thermal stability of the molecules for sample preparation and analysis by NMR.

The domain swap equilibrium was assessed by analytical SEC to determine the optimal buffer conditions to reduce the presence of dimer, both during sample generation and purification and also in the final NMR samples. A 50-60 % level of monomer was seen throughout a range of concentrations, increasing up to 530 μM, suggesting that the free scFv may be analyzable by NMR at high concentrations for the first time. This would allow comparisons of the antigen bound and unbound states further investigating the changes produced upon antigen recognition.

In addition to the scFv constructs, IL-6 (unlabelled and ^{15}N/^{13}C/^{2}H labelled, as discussed in chapter 3) and the 1189 Fab fragment (unlabelled) were successfully expressed in large quantities. The IC8/IL-1β Fab/antigen pair (^{15}N/^{2}H/unlabelled, chapter four) were also expressed to allow the exploration of further structural changes within the Fab molecule and what effect this may have beyond antigen recognition.

3.1 – Introduction

The 1980s saw the bringing together of a number of NMR spectroscopic techniques into a process that could determine the structure of proteins. The first protein structure produced by NMR was that of the bull seminal trypsin inhibitor (Williamson *et al*, 1985). This process relied on the sequential assignment method (Dubs *et al*, 1979; Wagner & Wüthrich, 1982) and the use of recently developed distance geometry calculations (Havel & Wüthrich, 1985) to determine the fold of the protein. This culmination of techniques itself relied upon significant technological advances leading from the initial development of NMR as a spectroscopic technique in the 1940s and 1950s. The development of Fourier transform spectroscopy, superconducting magnets, computational methods and the collection of two dimensional spectra were key factors in realising the potential of NMR as a tool for determining macromolecular structure. These initial steps into structure determination were limited by a number of factors, primarily relating to the size of the proteins studied. The limits of homonuclear NMR experiments, the magnetic fields of the time and quality of instrumentation restricted structure determination to proteins of less than 10 kDa. The proteins used also had to be stable at concentrations in the region of a few millimolar to provide sufficient signal to noise, further limiting the available candidates for study.

Since these first structures were produced, there have been significant advancements in technology, sample preparation and experimental design. The development of higher

magnetic fields, currently commercially available up to 22.3 T (or a 950 MHz proton Larmor Frequency), has allowed the collection of higher resolution spectra. The development of the cryoprobe, a key development in increasing the sensitivity of data collection, has also allowed the collection of data from lower concentration samples. Combining this with the advancement in electronics and signal processing over this time period, the current NMR spectrometers provide powerful tools for the investigation of ever more complex systems. The initial development of routine methods to generate ^{15}N and ^{13}C isotopically labelled protein was another critical step in developing NMR for the study larger proteins and protein complexes (LeMaster & Richards 1985). Advancements in labelling techniques have since led to the uniform isotopic labelling of samples in high yields from small culture volumes (Sivashanmugam *et al*, 2009). These labelling methods have allowed the collection of new and more complex multidimensional heteroneuclear spectra, both to improve the analysis of NOE peaks (Otting *et al,* 1990, Clore *et al* 1991) and to allow the unambiguous assignment of chemical shifts to the protein backbone using triple resonance spectra (as summarised by Gardner & Kay, 1998 and Ferentz & Wagner, 2000). The replacement of protons for deuterons to improve relaxation times, particularly by attenuating ^{13}C-^{1}H dipolar interactions, has also been fundamental in studying proteins larger than 20 kDa, where relaxation times would otherwise be too short for longer multidimensional experiments (Grzesiek *et al*, 1993). Combinations of these labelling techniques are now a fundamental part of NMR spectroscopy.

In addition to the development of heteronuclear and triple resonance experiments, methods for the collection of data from larger proteins have been created. The most notable of these being the Transverse Relaxation Optimised Spectroscopy (TROSY) method (Pervushin *et al*, 1997) which allows the collection of sharper, higher signal to

noise spectra, particularly at very high fields. TROSY modules have subsequently been added to the majority of the triple resonance and NOESY spectra used for sequential assignments (as summarised by Ferentz & Wagner, 2000).

The current potential of NMR as a tool for studying the macromolecular structure of proteins and other biomolecules is significantly greater than first seen in the 1980s when the first NMR structure was published. The molecular weight of a protein is no longer such a limiting factor, with structures produced for proteins approaching 100 kDa in size (Grishaev *et al*, 2008) and the collection of data investigating interactions and dynamics for protein complexes over 200 kDa (Gelis *et al,* 2007, Religa *et al*, 2010). Protein concentration and stability is also less of a limiting factor with the use of samples with concentrations as low as 100 μM. Non-uniform sampling is now also widely available for the reduction of experiment times to a third (or less) than the total time taken for uniformly sampled experiments, whilst maintaining spectral resolution (Rovnyak *et al*, 2004; Schmieder *et al*, 1994). Developments such as these have even led the way into developing in-cell NMR experiments that observe proteins, and generate structures, in a near native environment (Inomata *et al*, 2009; Sakakibara *et al*, 2009). Further development into new and existing experiments has been critical in improving the potential of the technique and reducing the time taken to get structural data. These include the collection of additional data such as RDCs (Tjandra & Bax, 1997), the collection and processing of non-uniform sampled data in four dimensional experiments (Mobli *et al*, 2010) and advancements in spectral processing and analysis. NMR spectroscopy is also a powerful tool in detecting subtle changes in protein structures. Movements of less than 1 Å are also detectable by chemical shift differences, surpassing the capability of other structural techniques.

114

3.2 – Materials and Methods

3.2.1 – NMR Sample Preparation

3.2.1.1 – ScFv

SEC purified ^{15}N/^{13}C/^{2}H labelled scFvs were concentrated to NMR concentrations by ultrafiltration (Vivaspin 20/Vivaspin 6/Vivaspin 2 10,000 MWCO, Sartorius). Final NMR samples were approximately 400 μl in volume, and between 200 and 400 μM in concentration. The buffer composition for the NMR samples was 100 mM sodium chloride, 25 mM sodium acetate, 0.02 % sodium azide, 10 μM EDTA, 200 μM AEBSF at pH 5.5. Additionally, ^{15}N/^{13}C/^{2}H 1189scFv samples with unlabelled leucines, isoleucines, valines or tyrosine/tryptophans and ^{15}N/^{2}H 1189scFv samples were prepared in the same way.

3.2.1.2 – Cloning and Expression of Interleukin-6

A human IL-6 construct, with Tobacco Etch Virus Protease (TEV) cleavable hexa-histidine tag, was provided by UCB in a modified pET21a vector containing an ampicillin resistance gene for selection. This plasmid was transformed into InvαF *E. coli* cells, Origami B DE3 *E. coli* cells (Merck) and Origami B DE3 pLysS *E. coli* cells (Merck).

Glycerol stocks of these cells were used to inoculate 100 ml LB cultures containing 12.5 μg ml^{-1} tetracycline, 15 μg ml^{-1} kanamycin and 100 μg ml^{-1} carbenicillin. These cultures were grown overnight at 37 ˚C/200 rpm in a shaking incubator, spun down, re-suspended in fresh antibiotic containing media and used to inoculate 500 ml antibiotic containing LB cultures to an A$_{600}$ of 0.1. The main expression cultures were grown at 37 ˚C/150 rpm until an A$_{600}$ of 0.6. At this point, the flasks were placed in an incubator, pre-cooled to 17 ˚C until an A$_{600}$ of 0.8 was achieved. IPTG was added to a final concentration of 100 μM and the cultures were left to express overnight (approx. 16 hours). Samples were taken and analysed as in section 2.2.2.4.

3.2.1.3 – Expression of ^{15}N/^{13}C/^{2}H labelled Interleukin-6

To form starter cultures for the expression of triple labelled IL-6, newly transformed Origami B DE3 pLysS *E. coli* colonies were selected from LB agar plates, containing 12.5 μg ml^{-1} tetracycline, 15 μg ml^{-1} kanamycin, 34 μg ml^{-1} chloramphenicol and 100 μg ml^{-1} carbenicillin, and used to inoculate 10 ml LB cultures grown at 37 ˚C/200 rpm overnight. These cultures were used to inoculate six 10 ml cultures of 30 % D$_2$O minimal medium grown for 8 hours at 37 ˚C/200 rpm, which were in turn used to begin another six 50 ml cultures of 70 % D$_2$O minimal medium grown overnight at 37 ˚C/200 rpm. The six overnight cultures were pooled, spun down (6000 x g, 15 mins), resuspended in fresh 100 % D$_2$O minimal medium, containing 1 g l^{-1} ^{15}N ammonium sulphate, 2 g l^{-1} ^{13}C$_6$ D-glucose and 100 μg ml^{-1} carbenicillin, and used to inoculate ten 500 ml expression cultures of the same media in 2.2 l baffled flasks to an A$_{600}$ of 0.1. The expression cultures were grown at 37 ˚C/200 rpm for 8-10 hours until an A$_{600}$ of 0.2-0.3 was achieved. At this point the flasks were transferred to an incubator pre-

cooled to 17 ˚C and left to equilibrate for 30 minutes. Following equilibration, the cultures were induced with IPTG to a concentration of 100 μM and left to express overnight (16 hours). Samples were taken and analysed as described in section 2.2.2.4.

3.2.1.3 – Purification of Interleukin-6 and the Formation of the scFv/Interleukin-6 Complex

Bacterial cell pellets containing IL-6 were lysed by French press, centrifuged, filtered and purified by nickel affinity chromatography as described in 2.2.4.1. The pooled fractions from this step were dialysed into a buffer containing 100 mM NaCl and 25 mM Tris at pH 7.5, before adding histidine tagged Tobacco Etch Virus (TEV) protease (Protex) and incubating overnight at 4˚C.

The TEV cleaved IL-6 was re-loaded onto the same, re-equilibrated, nickel affinity column, collecting the cleaved protein as flow through. The remaining histidine tag and TEV protease, which were bound to the column, could be eluted with elution buffer (see section 2.2.4.1). Samples of the collected proteins were taken and analysed by coomassie stained SDS-PAGE to assess the purity of the protein and the effectiveness of the cleavage step. The purified, cleaved protein was divided into aliquots, frozen in liquid nitrogen and stored at -80 ˚C.

The scFv/IL-6 complex was formed at a scFv concentration below 30 μM and at pH 5.5 by adding a 10 % molar excess of unlabelled IL-6 to labelled scFv or at a 10 % molar excess or an excess of unlabelled scFv to labelled IL-6. The complex could then be concentrated into a 5 ml sample to be purified by SEC. The fractions containing only

monomeric scFv/IL-6 complex were pooled and concentrated to NMR volumes and concentrations as described in section 3.2.1.1.

3.2.1.4 – The 1189 Fab/Interleukin-6 Complex

Unlabelled 1189 Fab was added to ^{15}N/^{13}C/^{2}H labelled IL-6 at a 10 % molar excess and purified and concentrated as described in 3.2.1.3 to 200-300 μM.

3.2.2 – NMR Spectroscopy

3.2.2.1 – Data Collection and Processing for Sequence Specific Assignments

All data was collected from 400 μl samples between 200 and 400 μM in concentration (as described section 3.2.1) in 5 mm Shigemi NMR tubes. All data was acquired at 40 ˚C on either a 600 MHz Bruker DRX or 800 MHz Bruker Avance spectrometer. The two and three dimensional TROSY based spectra (Pervushin et al, 1997) collected to allow the sequence specific assignment of NMR signals for 1189scFv in both the complexed and uncomplexed state were: ^{15}N/^{1}H TROSY (Pervushin et al, 1997), ^{15}N/^{13}C/^{1}H HNCO, HNCA, HN(CO)CA (Kay et $al.$, 1990), HNCACB (Wittekind & Mueller, 1993) and HN(CO)CACB (Grzesiek & Bax, 1992). Typical acquisition times for the 2D experiments were 30 ms in F_1 (^{15}N) and 60 ms in F_2 (^{1}H) with a total experiment time of 30 to 90 minutes. Typical acquisition times for the 3D experiments were 8 ms in F_1 (^{13}C), except for the HNCO experiment which was 18 ms, 20 ms in F_2 (^{15}N) and 60 ms in F_3 (^{1}H) with a total experimental time of 60 to 96 hours. In addition to this, long range NOE data was collected with a ^{1}H/^{15}N/^{1}H NOESY-TROSY (Zhu et al, 1998) experiment collecting 12 ms in F_1 (^{1}H), 10 ms in F_2 (^{15}N) and 60 ms in F_3 (^{1}H) with a NOESY mixing time of 450 ms for the 1189scFv/IL6 complex and 700 ms for the uncomplexed scFv over a total time of approximately 92 hours.

Solvent suppression was carried out using the WATERGATE pulse scheme (Piotto et al, 1992). All data was processed using Topspin 2.1 (Bruker Biospin Ltd.) using linear prediction to extend ^{15}N acquisition times up 30 ms in data sets with sufficient signal to noise. All data was analysed using SPARKY (Goddard & Kneller).

3.2.2.2 – Data Collection and Processing for the ^{15}N/^{13}C/^{2}H Interleukin-6 samples

All data was collected from the ^{15}N/^{13}C/^{2}H labelled IL-6 samples in 5 mm Shigemi NMR tubes. Three dimensional TROSY based ^{15}N/^{13}C/^{2}H HNCO spectra were collected for IL-6 in complex with 1189 Fab and 1189scFv and in an uncomplexed state. Typical acquisition times for these experiments were 18 ms in F_1 (^{13}C), 20 ms in F_2 (^{15}N), and 50 ms in F_3 (^{1}H), in a total experiment time of approximately 100 hours. All data was collected on a 600 MHz Bruker DRX system.

Solvent suppression was carried out using the WATERGATE method. All data was processed in Topspin 2.1 and analysed using SPARKY.

3.2.3 – NMR Data Analysis

3.2.3.1 – Sequence Specific Backbone Assignments

The sequence specific backbone resonance assignments for the amide proton (H_N),
amide nitrogen (N), alpha carbon (C_α), beta carbon (C_β) and carbonyl carbon (C′) were
obtained for residues within 1189 scFv both when in complex with IL-6 and in an
unbound state. These assignments were produced using the inter- and intra-residue
connectivities shown in the spectra produced by the $^{15}N/^{13}C/^1H$ HNCA, HN(CO)CA,
HNCACB, HN(CO)CACB and HNCO experiments (Ikura *et al*, 1990; Bax 1994).
Signals relating to the inter- and intra-residue amide NH to Cα and Cβ were identified
and used to search through the spectra for signals relating to neighbouring residues in
the protein sequence. Long range H_N to H_N NOE data collected in the $^{15}N/^1H$ NOESY-
TROSY spectra was used to help confirm sequential assignments. Patterns of H_N to
methyl proton NOEs from the same spectra were also used as an additional
confirmation in places where H_N to H_N NOEs were unobservable. Data from missing
peaks in $^{15}N/^{13}C/^1H$ HNCO spectra of the selectively unlabelled amino acid samples
(Leu, Ile, Val, Phe and Tyr/Trp) helped provide confirmation and location of the
assignments made, by introducing anchor points beyond what can traditionally be
determined from chemical shifts alone (i.e. Gly, Ala and Ser/Thr).

3.2.3.2 – Chemical Shift Mapping of the 1189 scFv

Both minimal (Veverka *et al*, 2009; Waters *et al*, 2007; Williamson *et al*, 1997) and actual chemical shift comparisons between the bound and unbound 1189scFv were used to investigate the location of structural changes within the scFv protein upon antigen binding. The minimal shift approach estimates the magnitude of chemical shift change between two spectra by comparing an assigned and an unassigned spectrum (Figure 3.2.1). The difference in position between an assigned peak and all peaks in the unassigned spectrum can be calculated and the lowest possible value selected. This process is repeated for all of the peaks in the assigned spectrum to produce the smallest possible chemical shift change, or minimal shift value, for each residue in the protein. The combined chemical shift change ($\Delta\delta$) between a pair of peaks can be calculated from ^{15}N/^{13}C/^{1}H TROSY-HNCO spectra using the formula:

$$\Delta\delta = \frac{\sqrt{(\Delta\delta_{HN}\alpha_{HN})^2 + (\Delta\delta_N\alpha_N)^2 + (\Delta\delta_{C'}\alpha_{C'})^2}}{n}$$

Where $\Delta\delta_{HN}$, $\Delta\delta_N$, and $\Delta\delta_{C'}$ are the ^{1}H, ^{15}N and ^{13}C chemical shift differences between pairs of compared ^{15}N/^{13}C/^{1}H HNCO peaks, α_{HN}, α_N and $\alpha_{C'}$ are scaling factors of 1.0, 0.2 and 0.35 respectively accounting for the differences in the spectral range of the amide proton, amide nitrogen and carbonyl chemical shifts and n is the number of dimensions of the experiment (Wishart *et al*, 1991; Wishart *et al*, 1995). When using ^{15}N/^{1}H TROSY data the C′ portion of the formula is omitted.

Figure 3.2.1- An example of the minimal shift analysis performed for 1189scFv. The minimal shift analysis calculates the nearest peak (black arrows) when comparing an unassigned 1189scFv/IL-6 ^{15}N/^{1}H TROSY spectrum (red) and an assigned 1189scFv ^{15}N/^{1}H TROSY spectrum (blue). This approach may under represent some shifts or may see a number of assigned peaks using the same unassigned peak to calculate minimal shift values in the cases where large shifts occur or peaks disappear (G173, S183, G185).

Minimal shift data (^{15}N/^{13}C/^{1}H HNCO) comparing assigned free 1189scFv spectra to unassigned bound 1189scFv spectra and assigned bound 1189scFv spectra to unassigned free 1189scFv spectra could be combined selecting the maximum value to produce a more complete minimal shift data set that benefits from differences in the assignments of the free and bound scFv.

The two sets of assignments were also directly compared by evaluating the resonances assigned to each residue to determine the actual chemical shift for each pair of assignments. The chemical shift difference ($\Delta\delta$) was calculated by subtracting the δ_{HN}, δ_N, and $\delta_{C'}$ chemical shifts for each residue in the bound scFv from the free, applying the weighting factor, removing the sign and dividing by three:

$$\Delta\delta = \frac{\sqrt{\left(\left(\Delta\delta_{HN(Free)} - \Delta\delta_{HN(Bound)}\right)\alpha_{HN}\right)^2 + \left(\left(\Delta\delta_{N(Free)} - \Delta\delta_{N(Bound)}\right)\alpha_N\right)^2 + \left(\left(\Delta\delta_{C'(Free)} - \Delta\delta_{C'(Bound)}\right)\alpha_{C'}\right)^2}}{3}$$

3.2.3.3 – Chemical Shift Mapping onto a Homology Model for 1189scFv

A homology model of 1189scFv was produced by Jiye Shi (UCB) based upon the IC8 scFv structure (Wilkinson *et al*, 2009). This model was used to produce images of the chemical shift analysis performed in section 3.2.3.2. Chemical shift differences were visualised by colour on the model of the scFv using the PyMOL molecular graphics system (Version 1.3, Schrödinger, LLC) with thresholds determined from the data sets.

3.2.3.4 – Minimal Shift Mapping of Antibody Binding on Interleukin-6

Assignments for human IL-6 (Veverka, unpublished) allowed the determination of chemical shift changes upon scFv and Fab binding by minimal shift analysis. The chemical shift differences observed between free and 1189Fab bound IL-6 and free and 1189scFv bound IL-6 were compared to determine if the binding site, and so properties of binding, were the same for both the Fab and scFv.

Minimal shift values were mapped onto an NMR derived structure of IL-6 (Xu *et al*, 1996) to identify the position of the binding site and any conformational changes that were induced by binding as described in 3.2.3.3.

3.3 – Results

3.3.1 – Expression and Purification of Interleukin-6

The expression cultures for unlabelled IL-6 grew to an A_{600} of approximately 3.5 following overnight expression at 17 ˚C showing a high level of expressed protein (Figure 2.3.8-B, T_{16}). The growth and expression levels were again reduced in deuterated media when compared to non-deuterated media with a final A_{600} of approximately 0.8. Despite this reduction in cell density, sufficient protein was produced from both the labelled and unlabelled expression for purification. The IL-6 produced was purified by nickel affinity chromatography in two steps. The initial purification separated the IL-6 from the bulk of other *E. coli* proteins using the histidine tag (Figure 3.3.1). The second purification step used histidine tagged TEV protease to cleave off the histidine tag and linker. This cleaved protein was re-applied to the nickel affinity column to remove the cleaved tags, TEV protease and any *E. coli* proteins that had bound to the column in the initial step (Figure 3.3.2). The amount of purified IL-6 produced was approximately 30 mg l^{-1} for the unlabelled protein and 2.75 mg l^{-1} for the triply labelled protein.
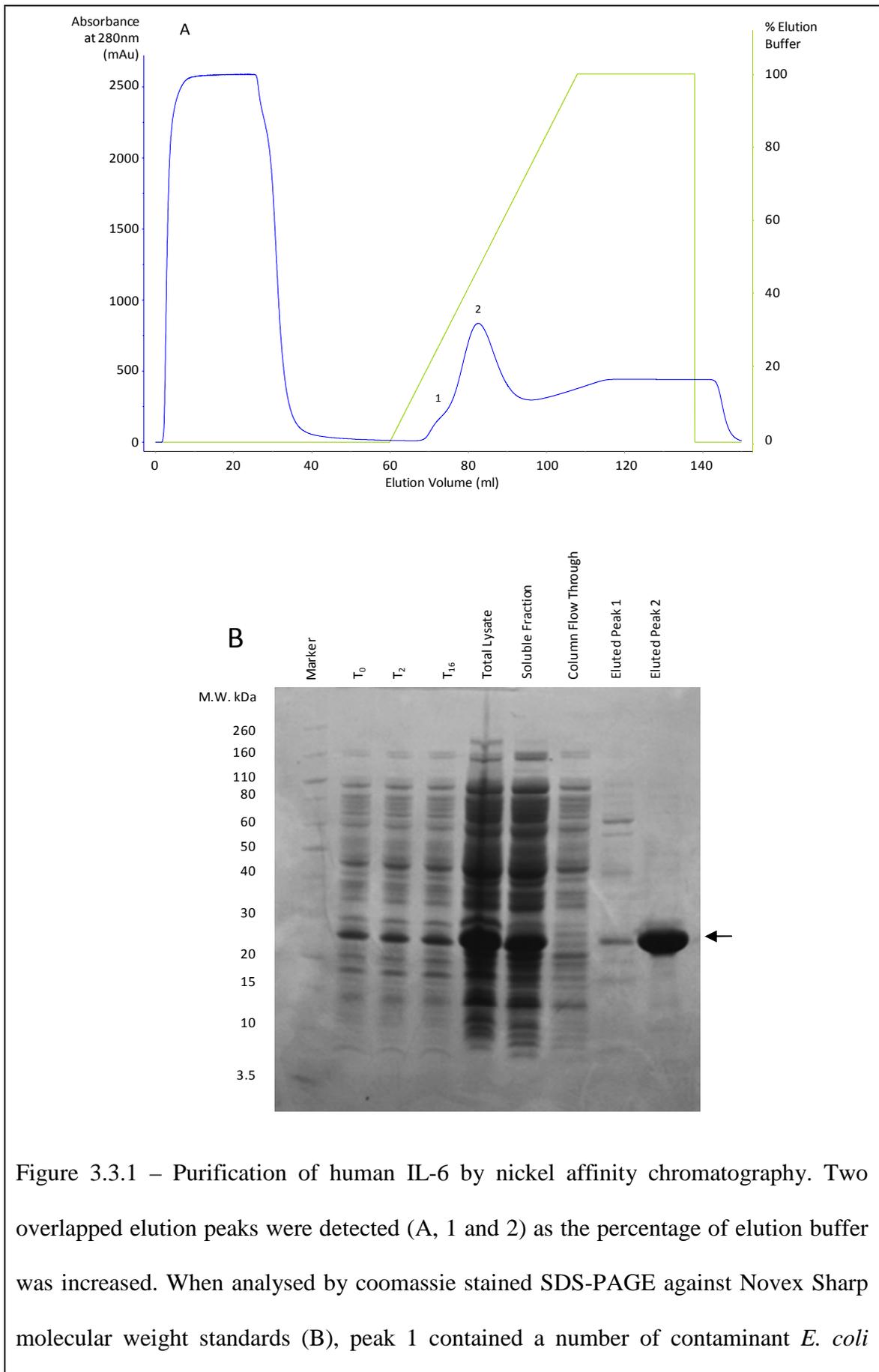
Figure 3.3.1 – Purification of human IL-6 by nickel affinity chromatography. Two overlapped elution peaks were detected (A, 1 and 2) as the percentage of elution buffer was increased. When analysed by coomassie stained SDS-PAGE against Novex Sharp molecular weight standards (B), peak 1 contained a number of contaminant *E. coli*

proteins that had bound to the nickel affinity column and peak 2 contained the expressed IL-6 with a low level of other contaminant *E. coli* proteins. Also shown are the samples for the 0, 2 and 16 hour time points taken during the IL-6 expression protocol (T_0, T_2 and T_{16}). It is clear that IL-6 is being highly expressed, both before and after the addition of IPTG. The total lysate corresponds to a sample of the cell lysate prior to centrifugation and the soluble fraction corresponds to the remainder of the lysate following both centrifugation and filtration, which was subsequently loaded onto the nickel affinity column. The column flow through is shown indicating that all soluble IL-6 was recovered using this procedure.

Figure 3.3.2 – Purification of TEV cleaved IL-6 by nickel affinity chromatography (A).

The cleaved IL-6 was passed through a nickel affinity column and separated from
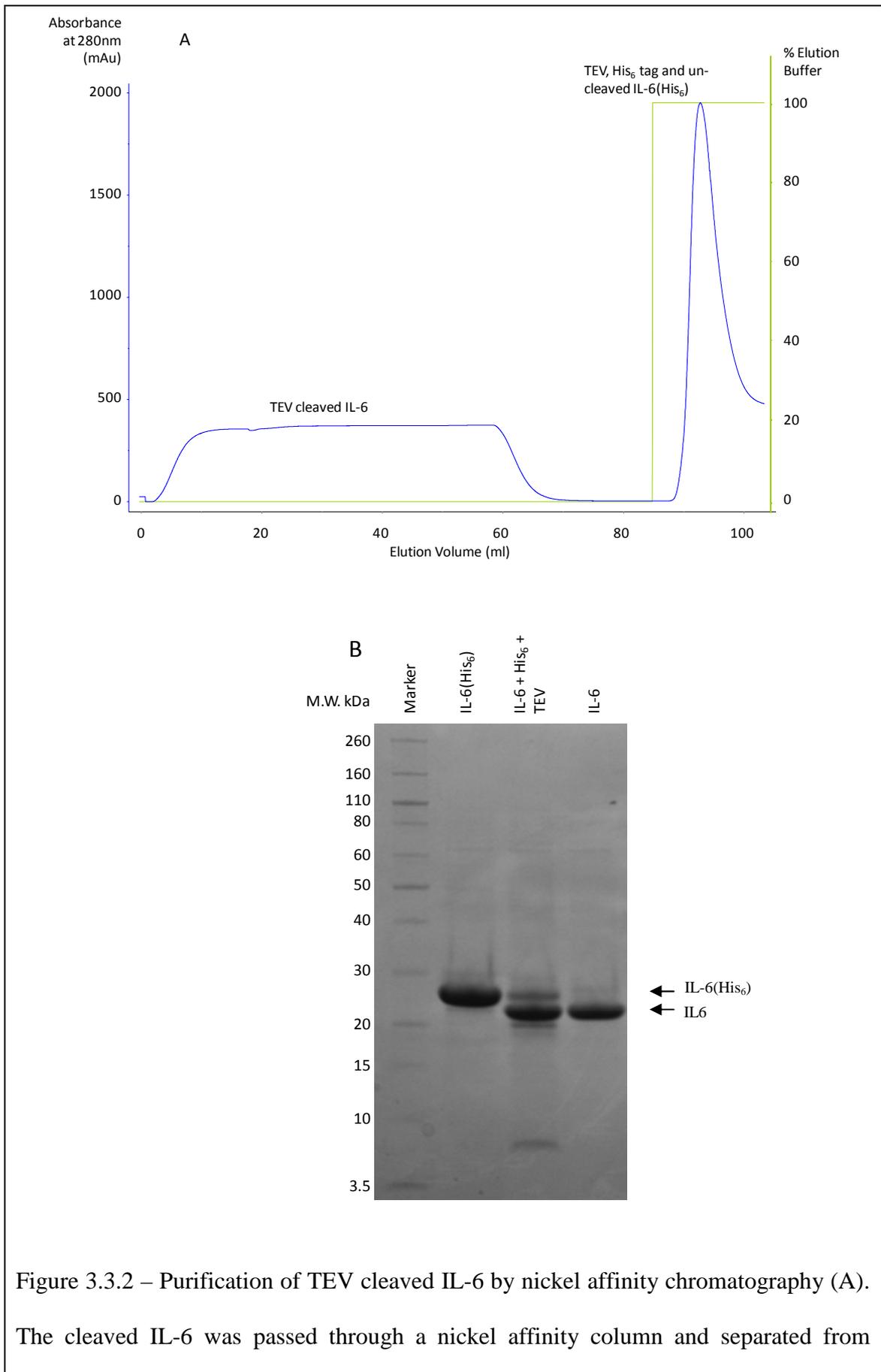
residual histidine tagged IL-6 (IL-6(His$_6$)), cleaved histidine tag (His$_6$) and TEV protease (TEV). The extent of cleavage and the purity of the final IL-6 protein sample were assessed by coomassie stained SDS-PAGE (B). The majority of the tagged IL-6 was clearly cleaved and the final IL-6 sample was free from contaminants.

3.3.2 – Sequence Specific Backbone Assignments

Despite the potential effects of the domain swap equilibrium discussed in chapter 2, full, high quality, backbone NMR data sets were collected for both 1189scFv and 488scFv. In contrast to the spectra previously seen for alternative scFv proteins (Wilkinson *et al,* 2009; Wilkinson, 2009), the HSQC and TROSY spectra of the unbound scFv were well resolved despite an increase in line width when compared to the bound (Table 3.3.1, Figure 3.3.3). The broader lines suggest a contribution from the dimeric form of the scFv in the NMR sample. This did not prevent the collection of well resolved, high quality two and three dimensional NMR spectra for 1189scFv, allowing comprehensive assignments to be obtained for the backbone when free in solution and when bound to 87 and 91 % complete respectively. Strip plots from the HNCACB and HN(CO)CACB spectra used in the backbone assignment process are shown in Figures 3.3.4 (1189scFv) and 3.3.5 (1189scFv/IL-6). HNCO data was collected from selectively amino acid unlabelled samples to help in the assignment process. Peaks missing from spectra collected for samples unlabelled for a specific amino acid type (for example isoleucine, Figure 3.3.6) could be used to locate assignments to specific points on the scFv sequence. Additionally, long range H$_N$ to H$_N$

and H_N to methyl NOE data collected from $^{15}N/^1H$ NOESY-TROSY experiments was used to help confirm the backbone assignments produced (Figure 3.3.7).

Protein		1189scFv	1189scFv	1189scFv/IL-6
Mass (Da)		27124	27124	48087
Labelling		$^{15}N/^{13}C$	$^{15}N/^{13}C/^2H$	$^{15}N/^{13}C/^2H$
Line width (Hz)	Q3	29.0	21.6	21.5
	G28	21.2	13.5	13.7
	K39	33.0	21.3	20.2
	F83	24.3	17.6	18.4
	S146	24.0	13.6	14.0
	T157	32.7	21.7	20.2
	W176	27.7	21.1	16.1
	K194	23.1	22.6	21.0
	R201	27.9	20.6	15.4
	G239	33.2	30.7	17.1
Mean (Hz)		27.6	20.4	17.8
Standard Deviation		4.4	4.9	2.9

Table 3.3.1 – Backbone amide proton line widths determined from $^{15}N/^1H$ TROSY spectra of $^{15}N/^{13}C$ labelled 1189scFv and $^{15}N/^{13}C/^2H$ labelled 1189scFv when free in solution and when bound to IL-6. The narrowest lines are observed for deuterated 1189scFv in complex with IL-6. The broader lines observed for 1189scFv without IL-6 are a result of the domain swap equilibrium taking place in the sample.
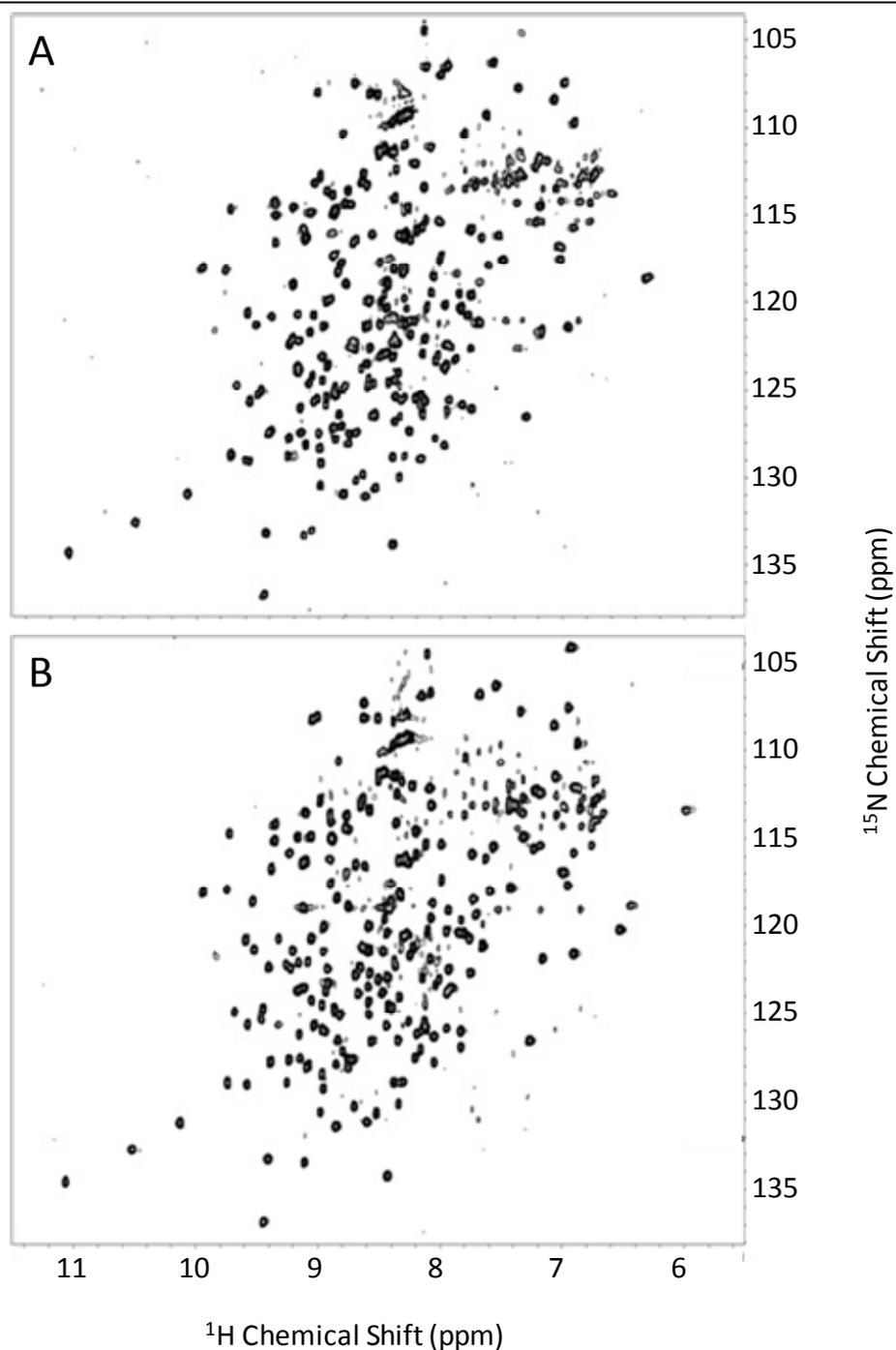
Figure 3.3.3 – ^{15}N/^{1}H TROSY spectra of ^{15}N/^{13}C/^{2}H labelled 1189scFv when free in solution (A) and complexed with IL-6 (B). Both spectra are of high quality showing resolved and sharp peaks with good signal to noise indicating the suitability of both samples for further multi-dimensional NMR experiments.

Figure 3.3.4 – ^{15}N strips from HN(CO)CACB and HNCACB spectra of ^{15}N/^{13}C/^{2}H labelled 1189scFv. These strips display the correlation of intra- (HNCACB) and inter-residue (HN(CO)CACB/HNCACB) connectivities for the Cα (red) and Cβ (green) atoms of residues A25 to E27 used in the process of sequential backbone assignment.
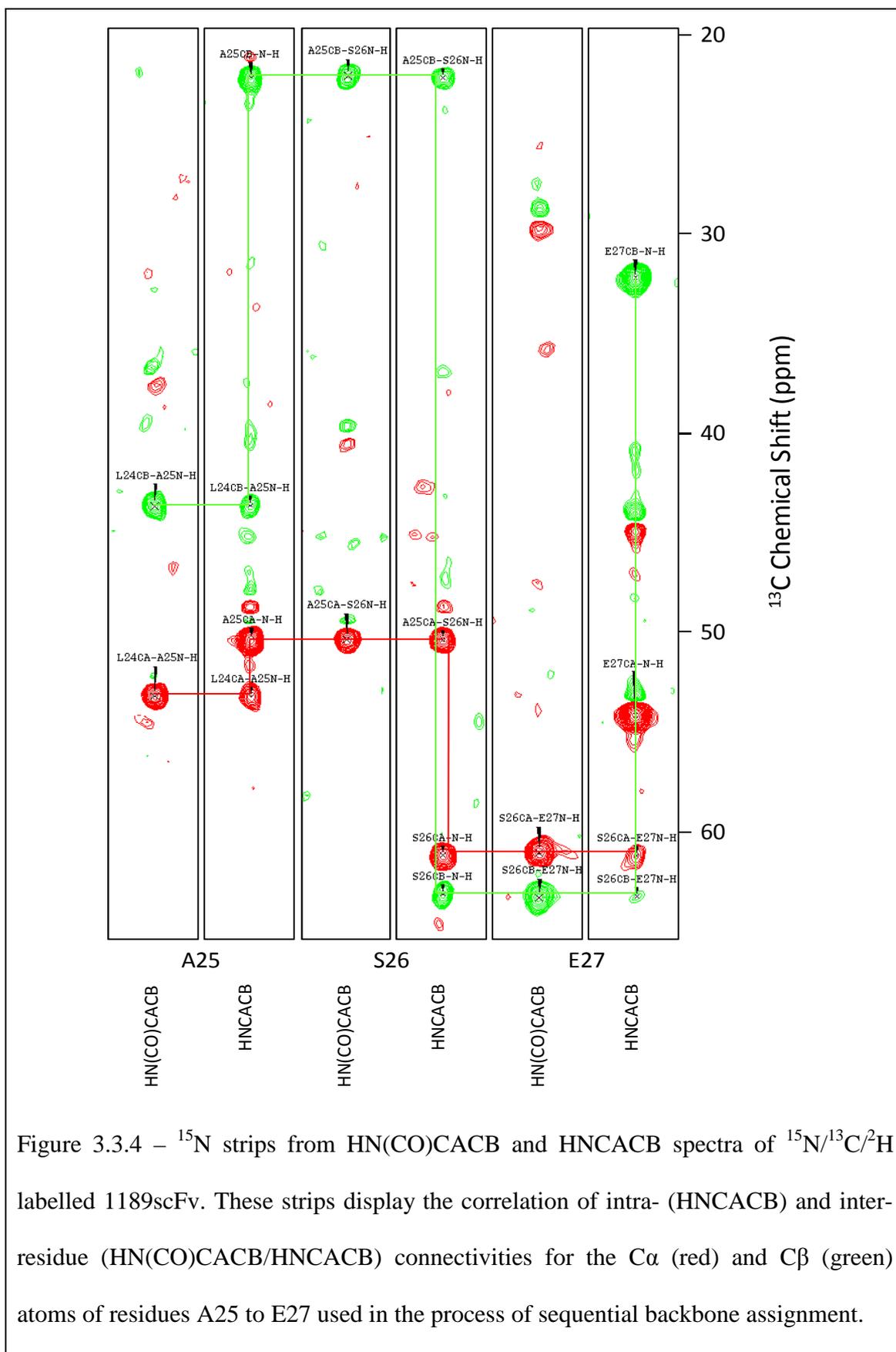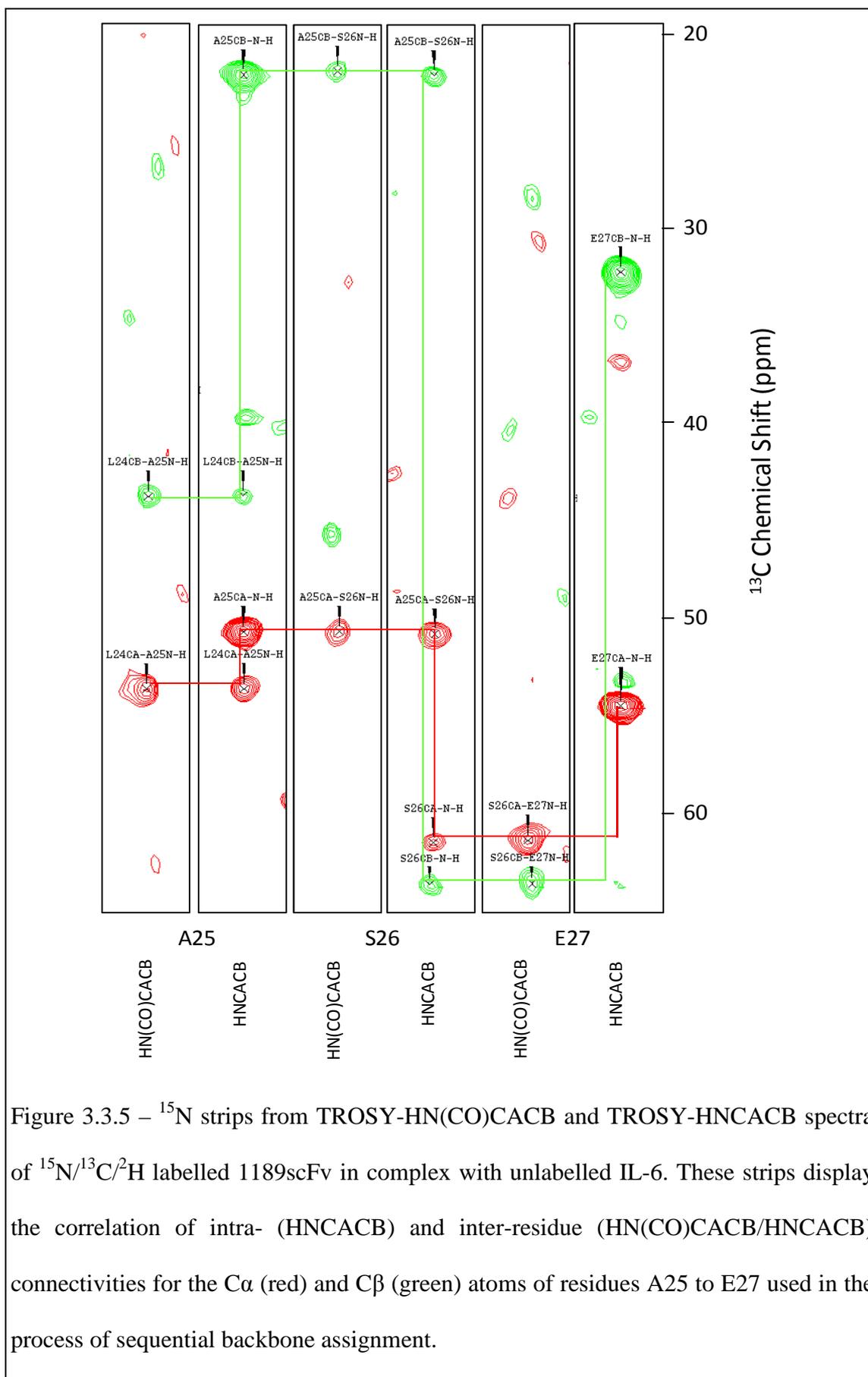
Figure 3.3.5 – [15]N strips from TROSY-HN(CO)CACB and TROSY-HNCACB spectra of ^{15}N/^{13}C/^{2}H labelled 1189scFv in complex with unlabelled IL-6. These strips display the correlation of intra- (HNCACB) and inter-residue (HN(CO)CACB/HNCACB) connectivities for the Cα (red) and Cβ (green) atoms of residues A25 to E27 used in the process of sequential backbone assignment.
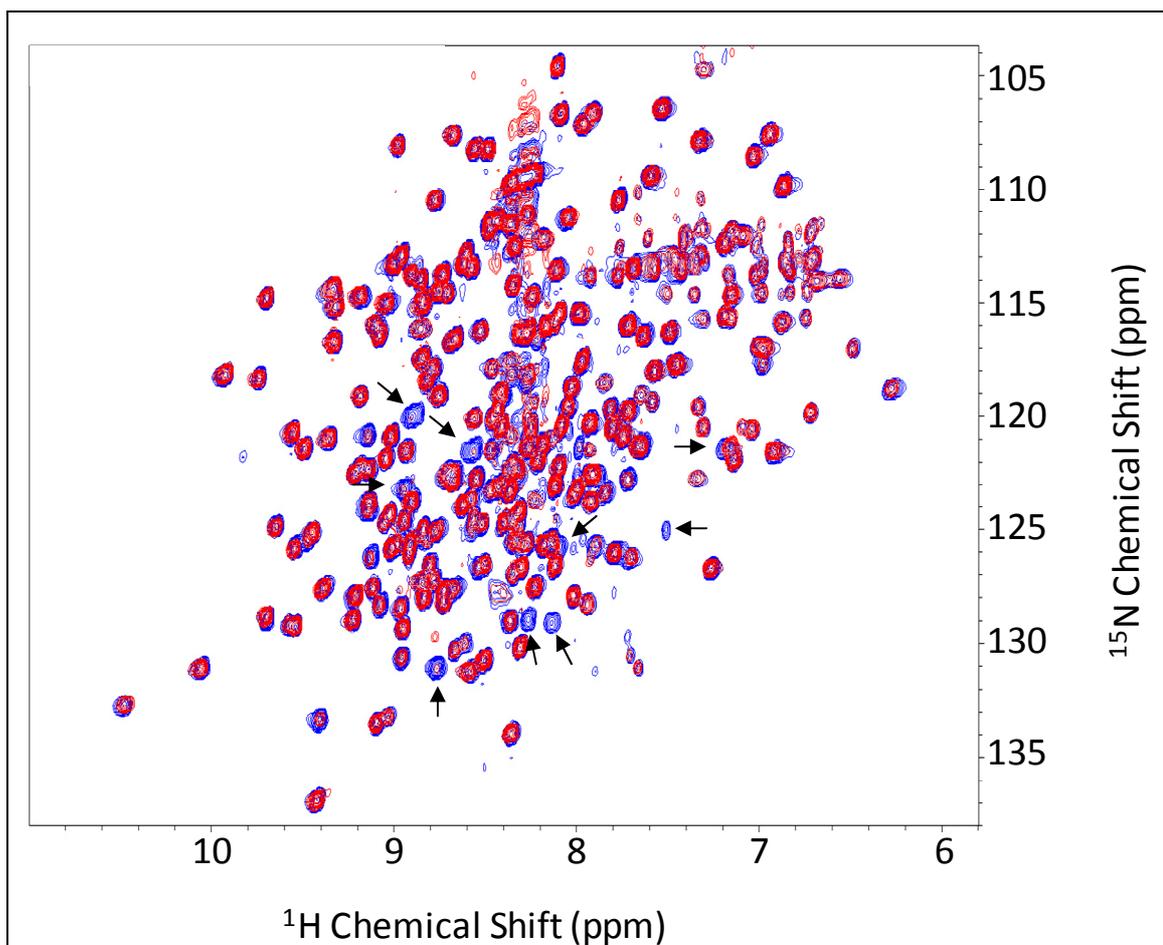
Figure 3.3.6 – A ^{15}N/^{1}H projection of a TROSY-HNCO spectrum representing a selectively isoleucine unlabelled ^{15}N/^{13}C/^{2}H labelled 1189scFv sample (Red) overlaid onto an equivalent projection for a uniformly ^{15}N/^{13}C/^{2}H labelled 1189scFv sample (Blue). When visualised in this way in both two and three dimensions, the peaks representing residues with isoleucines preceding them can be clearly identified (arrows) and used in the backbone assignment process.
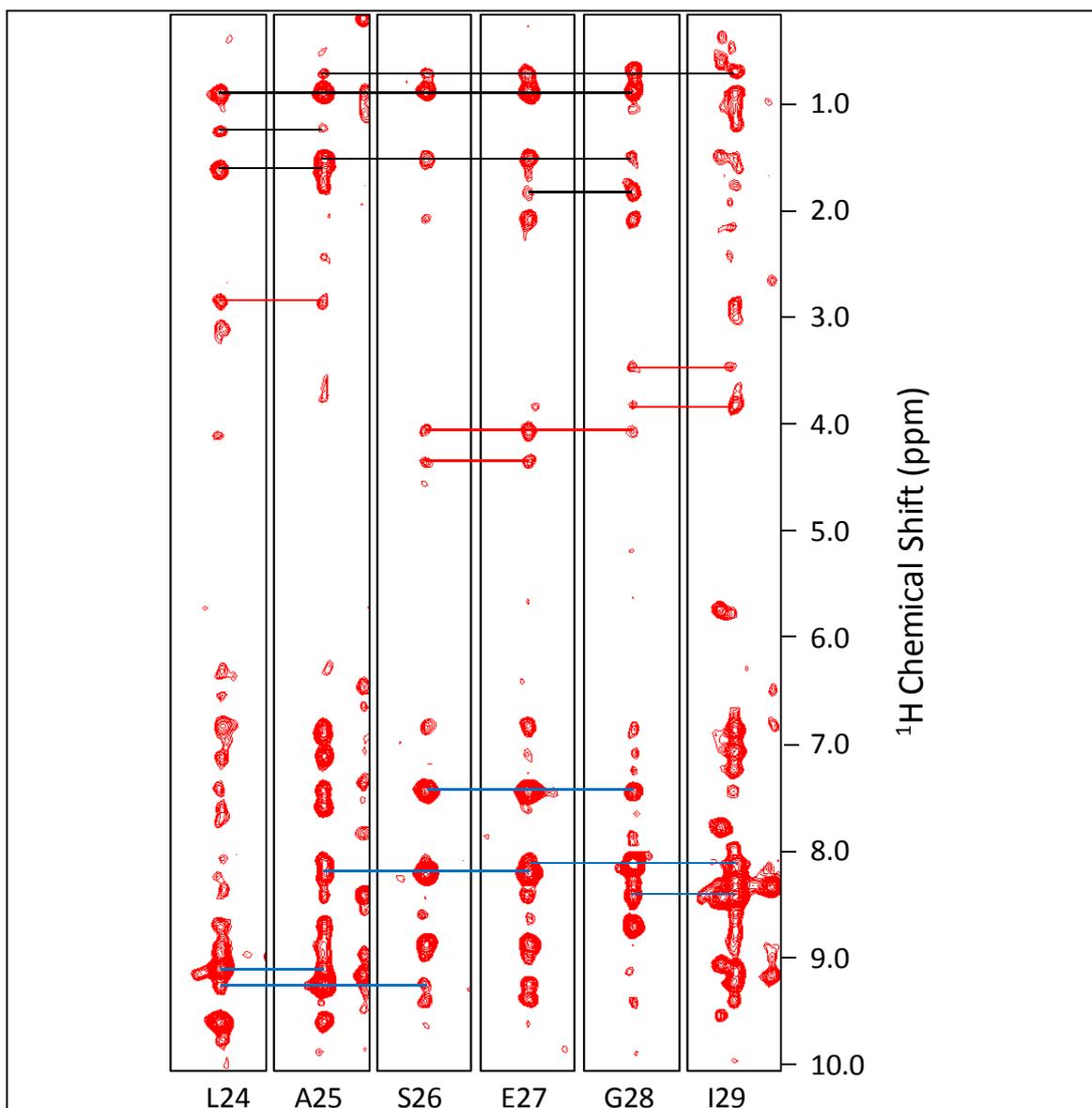
Figure 3.3.7 – ^{15}N strips from a ^{15}N/^{1}H NOESY-TROSY spectrum. The spectrum, optimised for the collection of long range NOE peaks using a mixing time of 400 ms, of ^{15}N/^{2}H labelled 1189scFv in complex with unlabelled IL-6. The spectrum was used primarily to confirm and extend assignments by linking the backbone amide NOE's of residues to the i-1 and i+1 peaks (blue). Connections to i±2 or i±3 may also be seen. In addition to this, patterns of αCH to NH and βCH to NH NOE's (red) and intra- and inter-residual CH$_3$ to NH NOE's (black) were used to confirm correct sequence specific assignment and provide further information to aid the assignment process.

Comprehensive backbone assignments (H_N, N, C_α, C_β, C′) were successfully obtained for 1189scFv with 190 of 218 assignable H_N resonances (excluding linker (20), His_6 tag (6), Prolines (9) and N-terminal residue (1)) assigned for the unbound scFv. An overview of the H_N assignments is highlighted in Figure 3.3.8-A and an assigned ^{15}N/^{1}H TROSY spectrum shown in Figure 3.3.9-A. The completeness of assignment for this set of data was 87 % with residues S9, L46, D50, S77, Y92, W96, E135, G145, M163, V177-S179, Y224, C225, R227-E236, Q240 and G241 remaining unassigned. The assignments for the scFv in complex with IL-6 showed a slightly increased level of completeness of 91 % with 198 of 218 residues assigned in generally longer fragments than for the unbound scFv (shown in Figure 3.3.8-B and Figure 3.3.9-B). Residues L47, D50, D56, S77, S91-Y94, W96, T97, G138, G145, Y224, C225, H228-T231, D233 Y234 and T242 remained unassigned. In both cases the majority of observed systems were assigned with 33 and 32 weak or incomplete systems remaining (HNCO, free and bound respectively from a total of 231 and 238) containing no inter-residual connectivities. A conversion table for the Kabat numbering of the antibody V_H domains can be found in Appendix 4.

1189scFv

A

V$_L$ CDR1 V$_L$ CDR2

DIQMTQSPSSLSASVGDRVTITCLASEGISNDLAWYQQKPGKAPKLLIYDATRLQDGVPS

V$_L$ CDR3 Linker

RFSGSGSGTDFTLTISSLQPEDFATYYCQQSYKYPWTFGQGTKLEIKRTGGGGSGGGGSG

V$_H$ CDR1

GGGSGGGGSEVQLVESGGGLVQPGGSLRLSCAASGFTFNDYDMAWVRQAPGKGLEWVASI

V$_H$ CDR2 V$_H$ CDR3

TPSGGGTYYRDSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARHGYTLDYFEYWGQ

His$_6$ Tag

GTMVTVSSHHHHHH

1189scFv/IL-6

B

V$_L$ CDR1 V$_L$ CDR2

DIQMTQSPSSLSASVGDRVTITCLASEGISNDLAWYQQKPGKAPKLLIYDATRLQDGVPS

V$_L$ CDR3 Linker

RFSGSGSGTDFTLTISSLQPEDFATYYCQQSYKYPWTFGQGTKLEIKRTGGGGSGGGGSG

V$_H$ CDR1

GGGSGGGGSEVQLVESGGGLVQPGGSLRLSCAASGFTFNDYDMAWVRQAPGKGLEWVASI

V$_H$ CDR2 V$_H$ CDR3

TPSGGGTYYRDSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCARHGYTLDYFEYWGQ

His$_6$ Tag

GTMVTVSSHHHHHH

Figure 3.3.8 – A graphical representation of the assignments obtained for 1189scFv in both the free (A) and IL-6 complexed state (B). Assigned residues (H$_N$) are highlighted in blue, with i-1 Cα and/or Cβ only assignments shown in yellow. The location of the CDR loops (raised line) and β-sheets (green) are shown for reference. The location of unassigned residues is similar for both samples located mainly in the CDR3 loops.

138

Figure 3.3.9 –Assigned ^{15}N/^{1}H TROSY spectra of ^{15}N/^{13}C/^{2}H labelled 1189scFv (A) and 1189scFv in complex with IL-6 (B). Assignments for the backbone amide signals are labelled by residue type and number. A good level of signal to noise and a low level of overlap in the spectra allowed a high level of assignment for the scFv.

3.3.3 – Chemical Shift Mapping of Interaction Sites

3.3.3.1 – Minimal Shift Analysis of 1189scFv Binding to Interleukin-6

Initially the assigned TROSY and HNCO spectra for the free 1189scFv were compared to the equivalent unassigned spectra for 1189scFv in complex with IL-6. Large chemical shift changes were seen between the two sets of spectra from which minimal shift analysis was easily conducted (Figure 3.3.10). The minimal shift changes observed are summarised on histograms with improved resolution from the three dimensional spectra (Figure 3.3.10-B). This provides improved clarity and definition of shifted regions when compared to the data produced from the two dimensional spectra (Figure 3.3.10-A), this is obtained primarily from a much reduced level of signal overlap and the additional chemical shift data from the C′ dimension. To further visualise the changes seen, and to properly assess their location on the scFv structure, the three dimensional minimal shift values were also mapped onto the homology model of 1189scFv as described in section 3.2.3.3 (Figure 3.3.11), using a noise threshold of 0.015 ppm and a highly shifted threshold of 0.040 ppm. This representation shows the largest shifts positioned mainly on and around the CDR loops and also in main framework of the variable domains, particularly at the interface between the V_L and V_H as apparent in Figures 3.3.10 to 3.3.16.

Figure 3.3.10 – Histograms summarising the minimal shift analysis of IL-6 binding to 1189scFv determined from assigned $^{15}N/^{1}H$ TROSY (A) and $^{15}N/^{13}C/^{1}H$ HNCO spectra of $^{15}N/^{13}C/^{2}H$ labelled 1189scFv. The increase in the clarity of shifted regions can be seen when using the three dimensional data, particularly around CDR 1 and 2 of the V_H

141
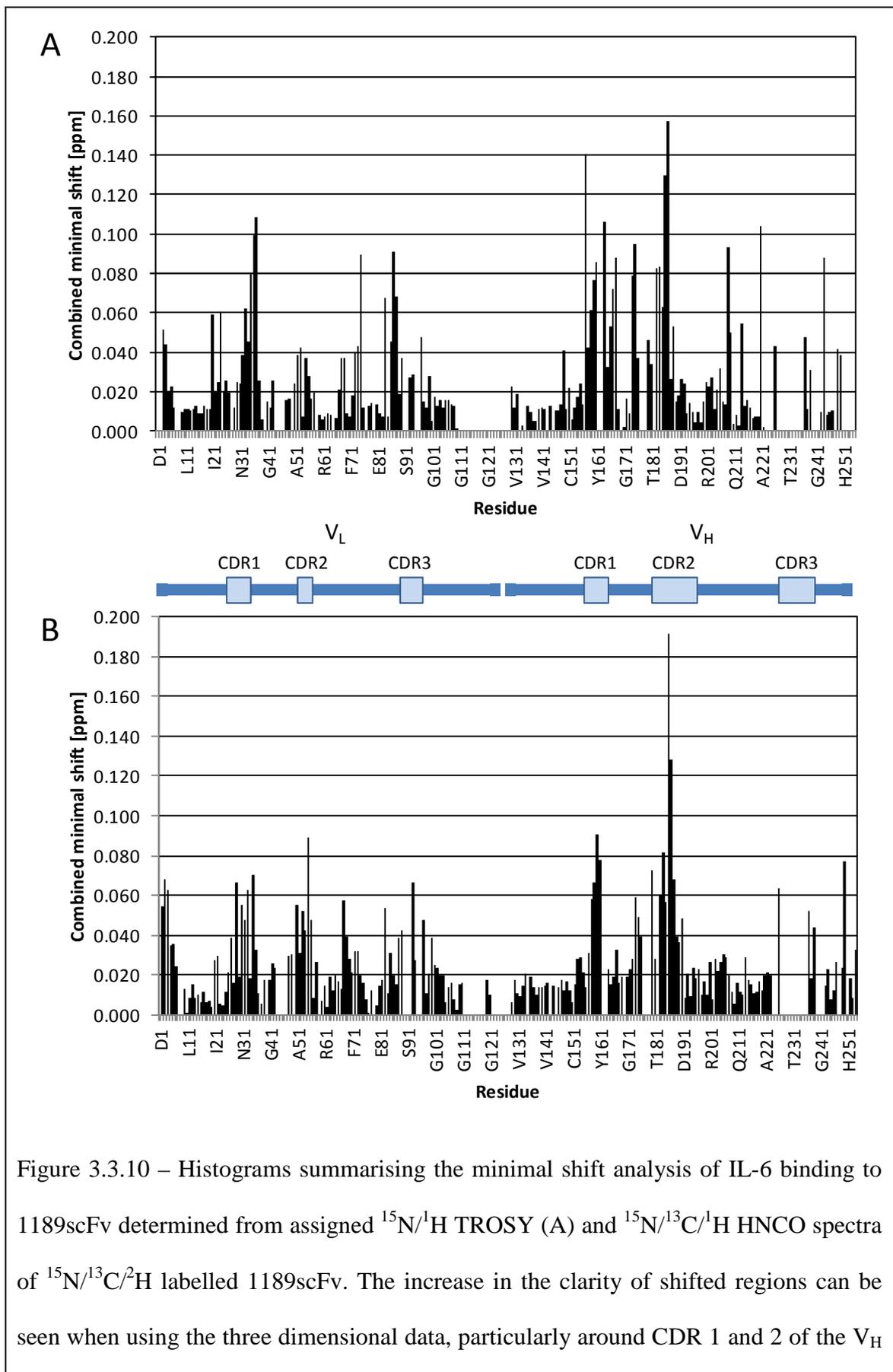
domain. No data was available for residues D1, I2, P8, S9, P40, P44-L47, D50, P59, S65, S77, P80, S91, Y92, P95, W96, G111-S129, L133, E135, P143, G145, D162, M163, P170, V177-S179, P182, Y223-K225, R227-E236, Q240, T242, S248 and H251-H254 in the 2D analysis and residues D1, P8, S9, P44-L47, D50, P59, Q90-Y92, P95, W96, G112-S119, G122-G128, E135, P143, G145, D162, M163, P170, V177-S179, P182, Y224, C225, R227-E236, G241, T242, S248, H251 and H254 in the 3D analysis.

Figure 3.3.11 – The locations of 1189scFv residues with NMR signals perturbed by the binding of IL-6. The two ribbon representations of the homology model produced for 1189scFv show the predicted locations of secondary structural features and enable the visualisation of parts of the structure not observable in a surface representation. All residues are coloured according to the combined minimal shift induced by IL-6 binding with minimal shift values below 0.015 ppm coloured white, above 0.040 ppm coloured red and those in between coloured on a linear, white to red gradient. Residues for which no chemical shift data could be obtained are coloured yellow. The majority of residues with perturbed chemical shifts sit on or around the CDR loops and at the interface between the V_H and V_L domains. In addition a surface representation looking down onto the CDR loops shows the binding surface of the scFv.

To complement this, the minimal shift analysis was repeated for HNCO spectra of the assigned 1189scFv in complex with IL-6 against unassigned spectra for the free 1189scFv (Figure 3.3.12). The results produced were highly similar for both minimal shift analyses validating the results seen.

Figure 3.3.12 - A histogram summarising the minimal shift changes observed in HNCO spectra of $^{15}N/^{13}C/^{2}H$ labelled 1189scFv in complex with unlabelled IL-6 when compared to unassigned spectra for the free scFv. Large chemical shift changes can again be seen around the CDR loops (as indicated above) and some framework residues. No data was obtained for residues D1, P8, P40, P44, L47, D50, D56, P59, S77, P80, S91-T97, G111, G113-124, G127, E135, G138, P143, G145, P170, P182, Y224, C225, H228-T231, D233, Y234, G241, T242, H249, H250 and H252-254.

As the assignment gaps for the bound and unbound scFv differ slightly, the fullest possible picture based on the minimal shift analysis was produced by selecting the highest minimal shift value from the two sets of data for the scFv ensuring that a value for the maximum possible number of residues was used. This helped to fill in some gaps that were present in the individual minimal shift analyses (Figures 3.3.13 and 3.3.14).
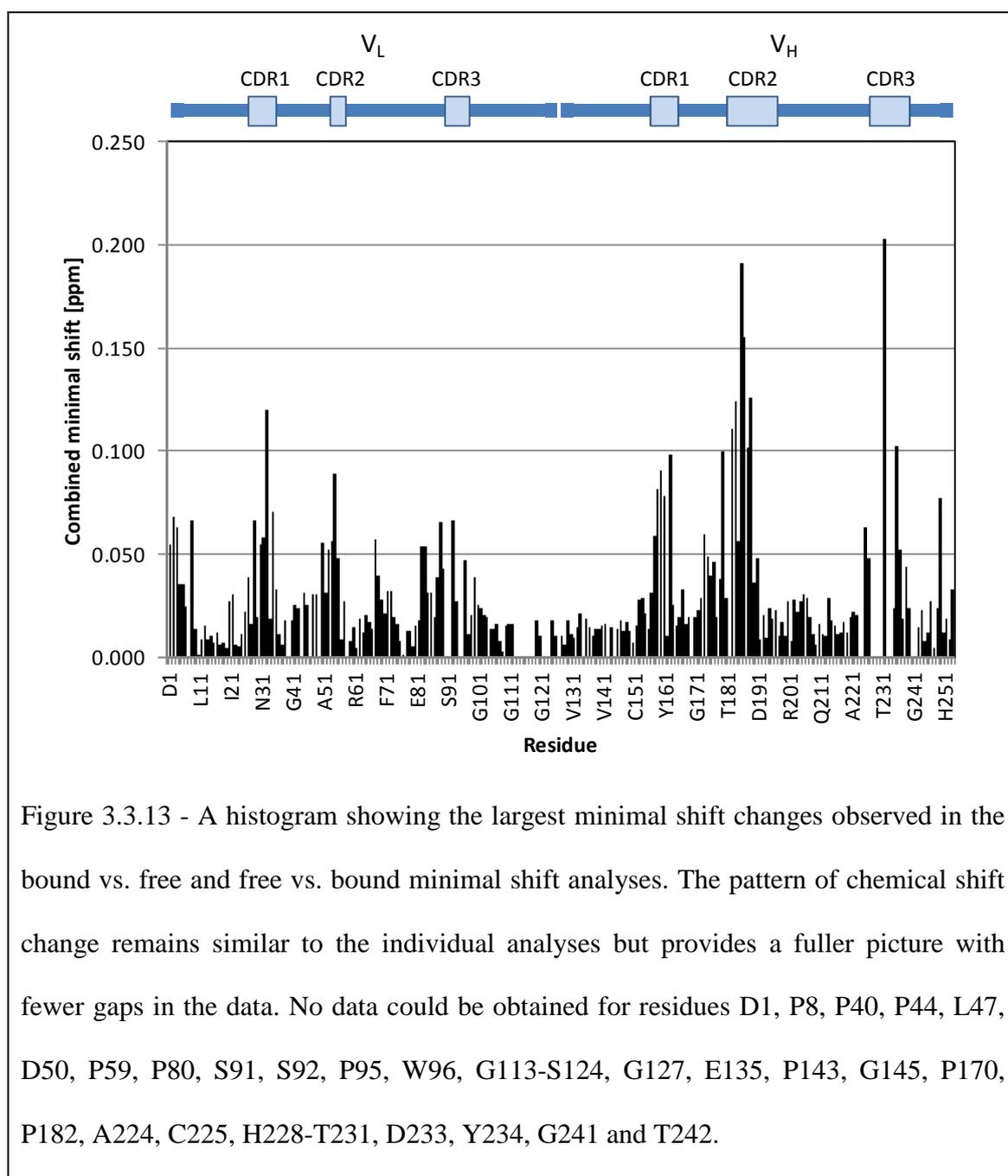
Figure 3.3.13 - A histogram showing the largest minimal shift changes observed in the bound vs. free and free vs. bound minimal shift analyses. The pattern of chemical shift change remains similar to the individual analyses but provides a fuller picture with fewer gaps in the data. No data could be obtained for residues D1, P8, P40, P44, L47, D50, P59, P80, S91, S92, P95, W96, G113-S124, G127, E135, P143, G145, P170, P182, A224, C225, H228-T231, D233, Y234, G241 and T242.

Figure 3.3.14 – Locations of the scFv residues with NMR signals perturbed by the binding of IL-6 as determined by combining free vs. bound and bound vs. free minimal shift approaches. The two ribbon representations of the homology model of 1189scfv show the predicted location of secondary structural features and enable the visualisation of parts of the structure not observable in a surface representation. All residues are coloured according to combined minimal shift induced by IL-6 binding with values below 0.015 ppm coloured white, above 0.040 ppm coloured red and those in between coloured on a linear, white to red gradient. Residues for which no chemical shift data could be obtained are coloured yellow. This figure gives a fuller picture of the locations of signal perturbation on 1189scFv. The majority of chemical shifts sit on or around the CDR loops and at the interface between the V_H and V_L domains.

The final comparison between the bound and unbound 1189scFv was of the two assigned sets of HNCO chemical shifts for each residue (Figures 3.3.15 and 3.3.16). This analysis provides a definite chemical shift change for each residue with an assigned peak in each spectrum. However, due to the assignment gaps and the differing position of these for the free and bound scFv, a number of residues have either only one set of chemical shifts or no shifts at all resulting in the picture produced having a larger proportion of missing data than the minimal shift analyses.

Figure 3.3.15 – Histograms displaying the actual chemical shift (combined HN, N, C′) for 1189scFv binding to IL-6. The left histogram displays the full range of the data set whilst the histogram on the right provides further clarity for smaller shifts. The largest shifts observed lie within the inter-domain interface on the V_L domain. There are, however, still significant shifts in the CDR loops and framework residues. There is no data for residues D1, S7-S9, K39, P40, A43-L47, Y49, D50, Q55, D56, V58, P59, S76-P80, Q89-T97, G110-128, V134, E135, G137, G138, Q142-G145, Y161-M163, A169, P170, W176, S179, T181, P182, Y223-E236, G239-T242 and S247-H254.

Figure 3.3.16 – Locations of scFv residues with NMR signals perturbed by IL-6 binding. The two ribbon representations of the homology model produced for 1189scFv show the predicted locations of secondary structural features and enable the visualisation of parts of the structure not observable on a surface representation. All residues are coloured according to combined actual shift induced by IL-6 binding with values below 0.002 ppm coloured white, above 0.030 ppm coloured red and those in between coloured on a linear, white to red gradient. Residues for which no chemical shift data could be obtained are coloured yellow. The majority of chemical shifts sit on or around the CDR loops. The largest shifts however sit on the interface on the V_L domain suggesting that very large changes are occurring in this area.

3.3.3.2 – Minimal Shift Analysis for scFv and Fab Binding to Interleukin-6

The minimal shift analysis for IL-6 to compare the binding of Fab and scFv was performed by comparing the Fab and scFv bound IL-6 HNCO spectra to spectra collected for IL-6 and by comparing the Fab and scFv bound IL-6 spectra directly to each other. The locations of the peaks in the IL-6 spectra when bound to the scFv and Fab were very similar with only a small number of subtle changes between the two sets of data (Figure 3.3.17), which were confirmed by the minimal shift comparison between the two samples (Figure 3.3.18). The chemical shift perturbation of signals from IL-6 was mapped onto a previously solved NMR structure of the protein (Xu *et al*, 1996, PDB accession code 2IL6) (Figure 3.3.19). The differences between the minimal shift data sets were also mapped on to the same structure in Figure 3.3.20.
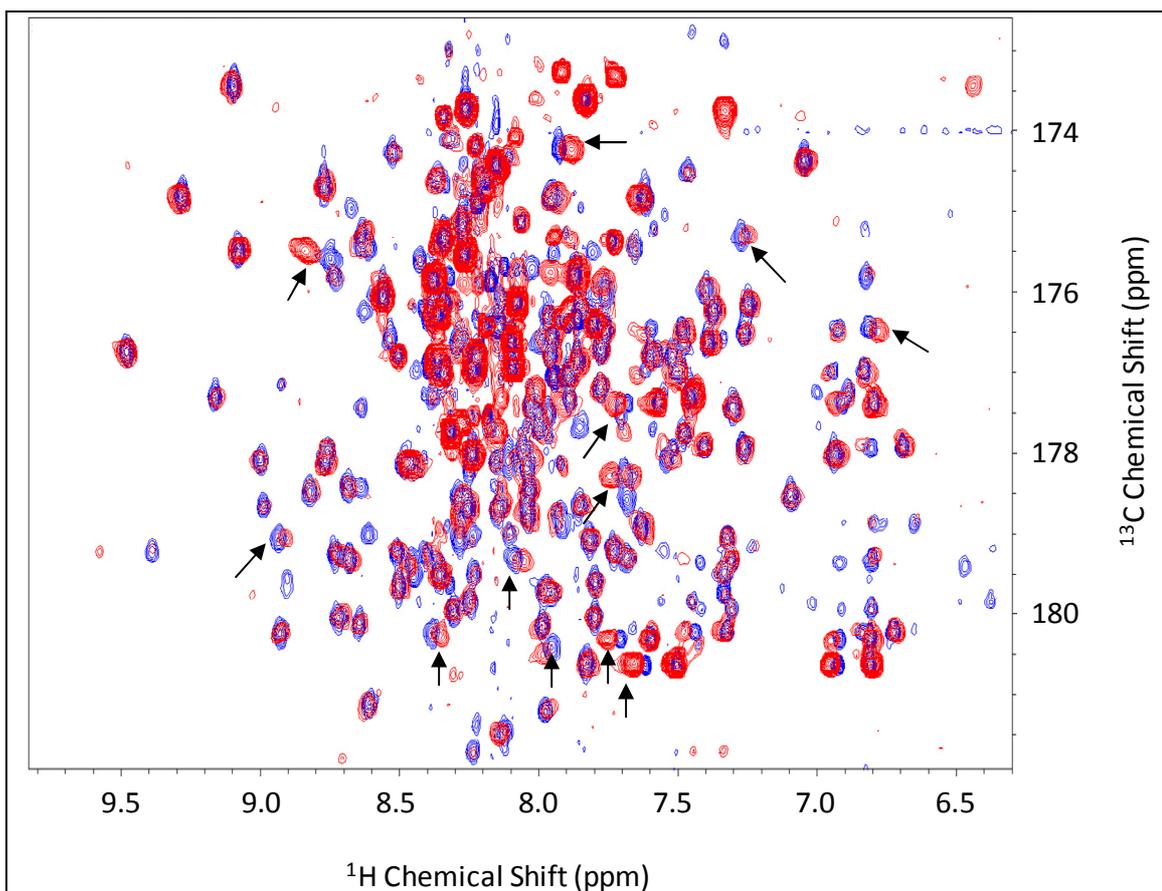
Figure 3.3.17 – Overlaid $^{13}C/^{1}H$ projections of $^{15}N/^{13}C/^{1}H$ HNCO spectra of $^{15}N/^{13}C/^{2}H$ labelled IL-6 in complex with 1189scFv (Blue) and 1189Fab (Red). The two spectra are highly similar, apart from a small number of shifted peaks (arrows) and a broader line width in the Fab bound spectrum due to the larger size of the complex.
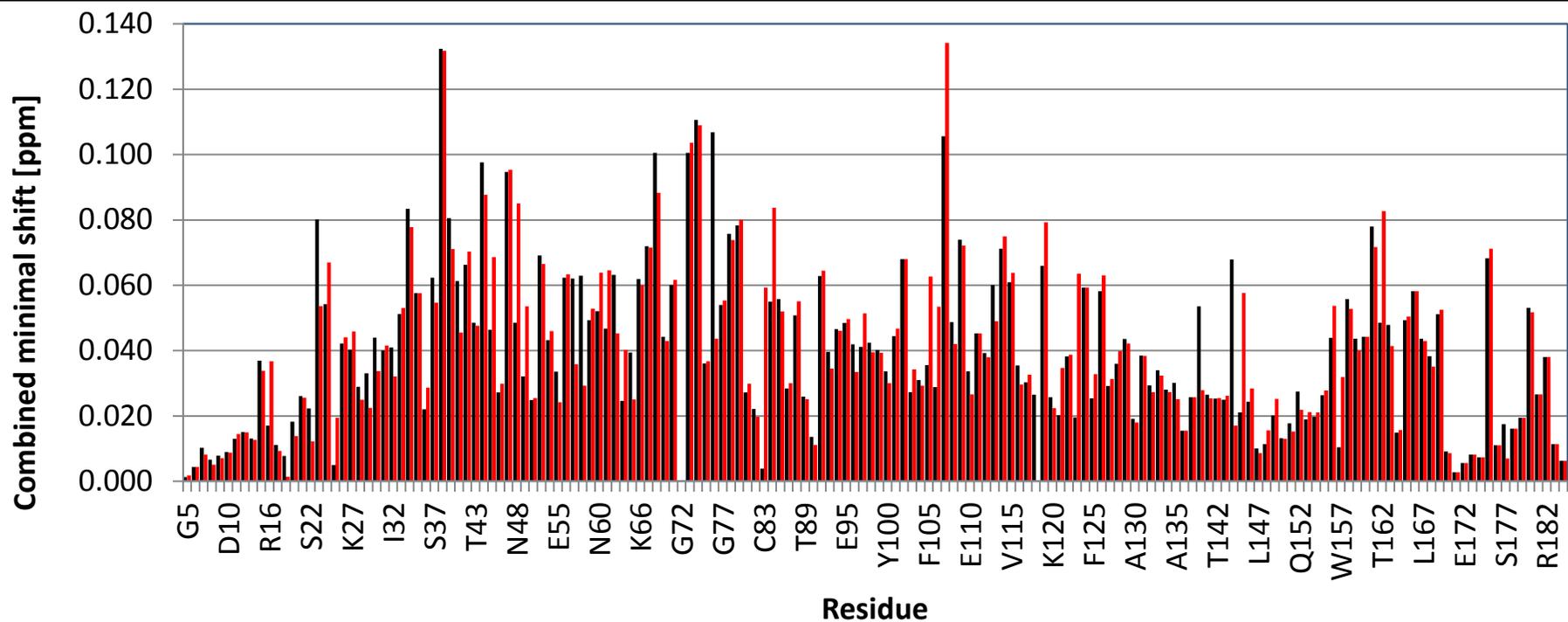
Figure 3.3.18 - A histogram showing the minimal shift changes observed in TROSY-HNCO spectra of ^{15}N/^{13}C/^{2}H labelled IL-6 in complex with unlabelled 1189scFv (Black) or 1189Fab (Red). Relatively large chemical shift changes can be seen throughout the protein, the pattern of which is highly similar when bound to both the scFv and the Fab. No data was obtained for residues A1-P5, P15, P19, K42, E52, K55, P66, N80, I89, G91, P140 and P142.
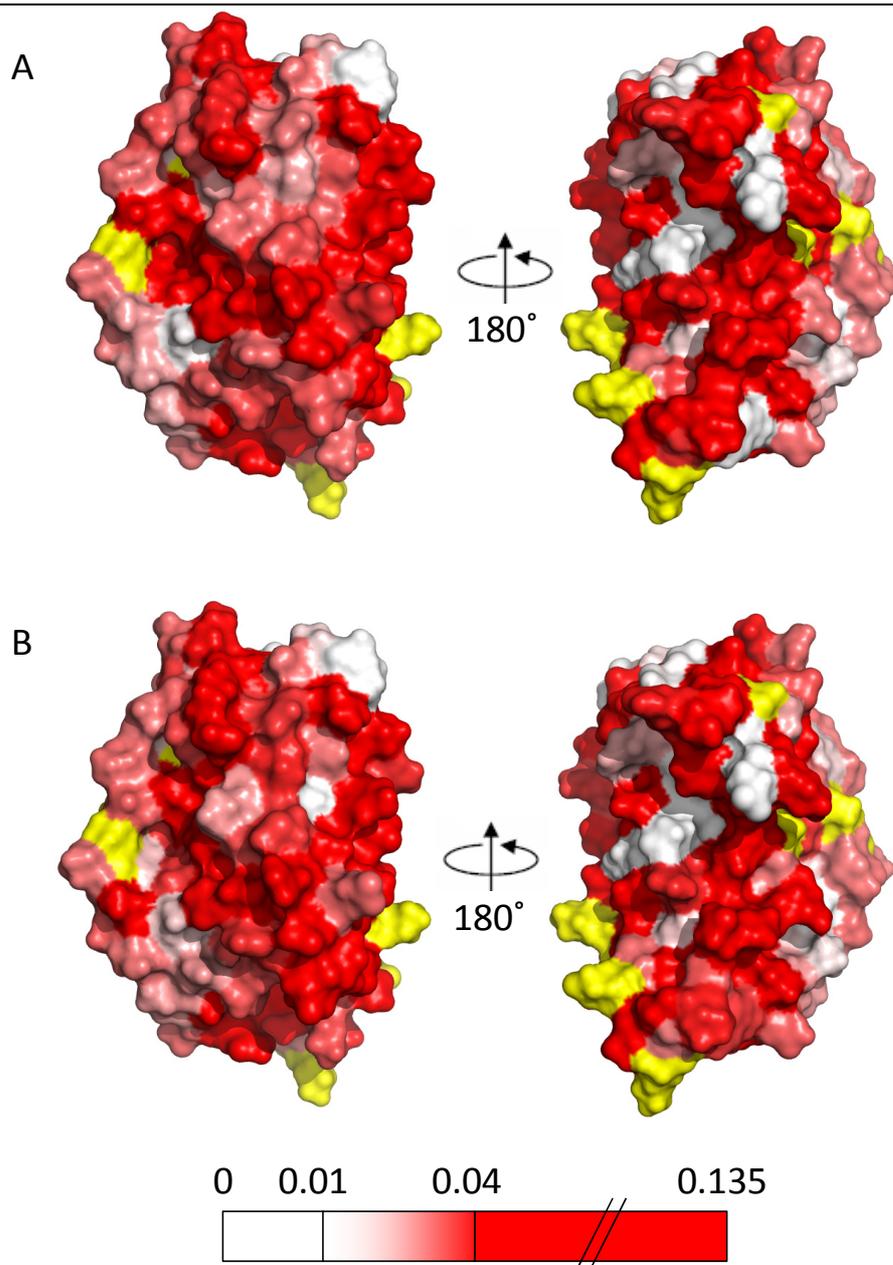
Figure 3.3.19 – Locations of IL-6 residues with NMR signals perturbed by the binding of 1189scFv (A) and 1189Fab (B). Residues corresponding to ^{15}N/^{13}C/^{1}H HNCO peaks with a scFv or Fab induced minimal shift value of less than 0.01 ppm are coloured white, more than 0.04 ppm red and those in between coloured on a linear gradient from white to red. Residues for which no data could be obtained are coloured yellow. The overall distribution of chemical shifts is similar with only a relatively small number of subtle differences.

Figure 3.3.20 – The location of chemical shift differences between the IL-6/Fab and IL-6/scFv data sets. Chemical shift differences between the data sets of less than 0.01 ppm are coloured white, more than 0.04 ppm red and those in between coloured on a linear gradient from white to red. Residues for which no data could be obtained are coloured yellow.
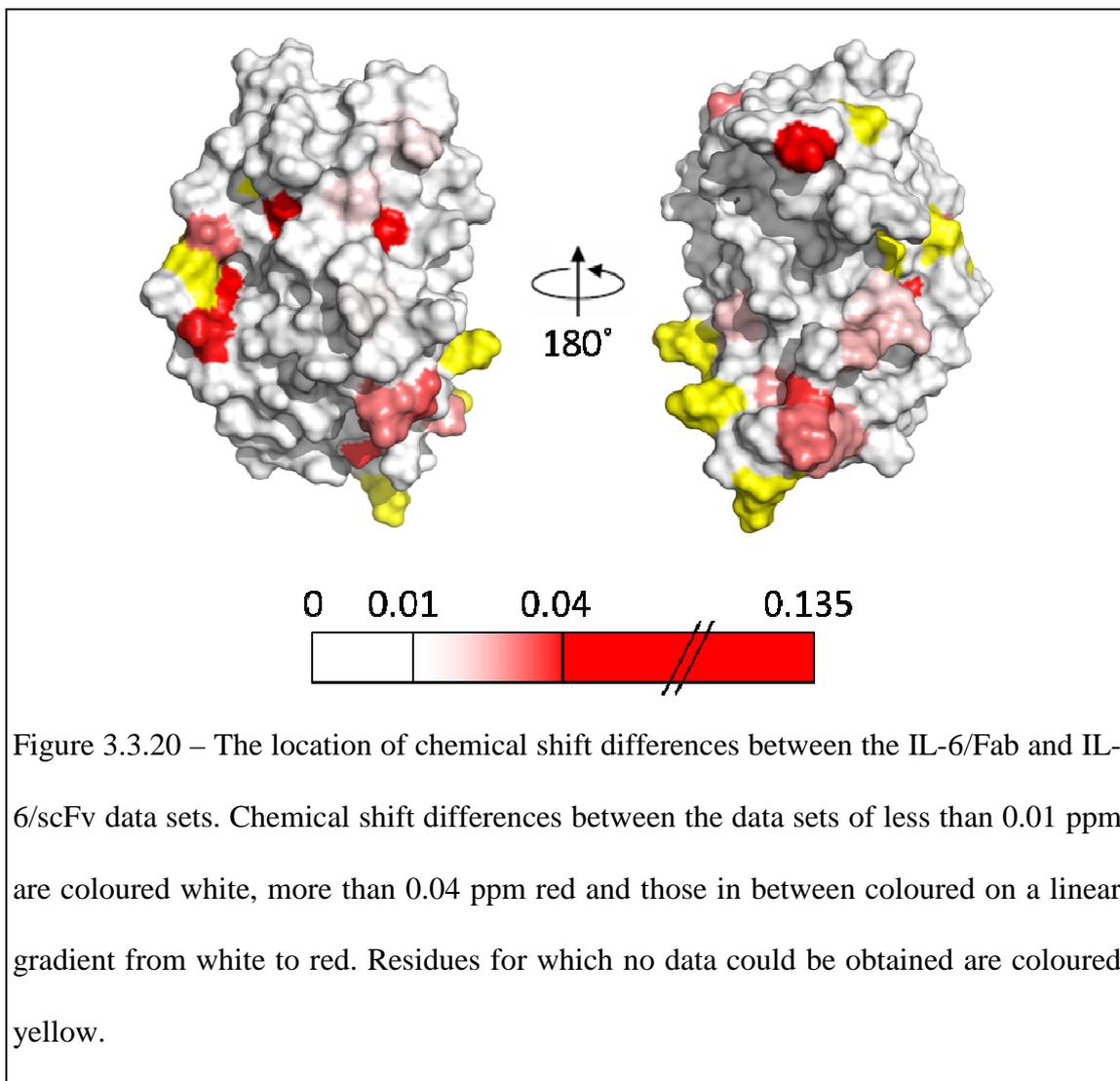
3.4 – Discussion

Despite the potential problem of the domain swap dimer formation on the quality and availability of NMR data for the scFv when unbound in solution, the results described in this chapter clearly show that a high level of backbone assignments can be achieved from good quality, high resolution spectra. The overall line widths of the spectra for the free scFv were broader than those observed for the scFv in complex with IL-6 (Table 3.3.1) suggesting that the domain swap equilibrium (Worn and Pluckthun, 2001) does still affect line shape and data quality. Stabilisation by antigen binding and purification can resolve this by locking the scFv into the monomeric form. The extent of line broadening by this equilibrium indicates that the signals may indeed be from a dimeric form of the scFv, as the dimer would have a similar molecular weight to the scFv/IL-6 complex and so a similar, if not slightly broader, average line width. The signals seen, however, should not be considered to be exclusively dimer due to the gel filtration analysis used in Chapter 2, which observed both monomer and dimer in samples of an NMR concentration. It is most likely that the signals seen are a composite of the monomer and dimer signals. Despite this line shape the data collected was of a sufficiently high quality to allow backbone resonance assignments to be made and comparisons to be drawn between the chemical shifts for 1189scFv, both on its own and in the presence of IL-6. As the longer triple resonance pulse sequences favour slower relaxation times, the smaller monomeric form of the scFv may be selected for over the larger dimeric species resulting in higher quality spectra with narrower lines than may have been expected from the two dimensional data.

The work reported here also validates the effectiveness of producing high quality NMR samples using the small volume, high density labelling method for use in multi-dimensional experiments. The specific non-labelling of amino acids by type has been successfully carried out using this method and incorporated into the backbone assignment process of challenging proteins, such as the unbound 1189scFv. This type of labelling is particularly useful where backbone assignments are fragmented and difficult to locate to specific portions of the sequence using the information on residue type provided by chemical shift alone. The chemical denaturation and refolding of the scFv was also critical in recovering signals lost during deuterated expression, giving a higher overall assignment level than would have been otherwise possible and shedding light on chemical shift changes within the core of the protein structure.

Backbone assignments for both the free and bound scFv were completed to a high level when the regions of the protein expected to give rise to no signals or heavily overlapped signals (histidine tag and linker) are excluded. A level of 87 % for the unbound scFv and 91 % for the bound scFv indicate a good level of assignment with 33 and 32 systems (as seen on HNCO spectra) remaining unassigned due to weak or incomplete inter- and intra-residual signals. This level of assignment was more than sufficient to provide a good overall picture of the chemical shift changes upon IL-6 binding and could offer the possibility of producing more structural information in the future for both forms of the protein. In both cases, specific regions of the scFv could not be assigned, the most obvious being CDR3 on the V_H domain, but also some residues within other CDRs, loops and the interface between the two variable domains. In the case of V_H CDR3, this is most likely to be due to the structural nature of the loop. Loops that are unstructured and more stable structural elements are both visible by NMR. The majority of the loops observed in the scFv give rise to clear, well resolved

signals that are similar in intensity to the majority of the signals from the protein. This suggests that they have a relatively stable conformation. V_H CDR3, however, shows no signals and is likely to be sampling a small number of distinct conformations on an intermediate (millisecond to second) NMR timescale, a characteristic that broadens NMR signals often to unobservable levels. This flexibility may help to reduce the entropic penalty of trying to dock a rigid, high affinity surface on to an antigen and may allow the initiation of binding at less than perfect orientations. Another interesting observation from this process is that whilst IL-6 locks the sample into the monomeric form and improves overall line width and spectral quality, it does not appear to stabilise any of the unobservable CDR residues into a single state that can be observed by NMR spectroscopy. This is in contrast to results previously seen for antibody fragments against IL-1β in which signals were restored for the unobservable CDR loops in both scFv and Fab proteins upon binding (Hall, 2009; Wilkinson, 2009). The most likely cause of this continued lack of signals is the inherent plasticity of IL-6 (compared to the more rigid IL-1β protein) which has been previously shown to broaden NMR signals (Xu *et al*, 1996) and may be a feature of IL-6 that is important for the formation of the IL-6/IL-6Rα/gp180 signalling complex (Boulanger *et al*, 2003). It could also be argued that the CDR3s do not bind and so continue to move between conformations. However, as these residues sit at the centre of the binding pocket, some interaction with IL-6 is highly likely.

The missing CDR loop signals for 1189scFv are primarily located on V_H CDR3. This CDR loop is known to show increased flexibility and so is not unexpected. There is data, however, that demonstrates that other CDR loops may also display some flexibility. For 1189scFv there are a small number of residues missing for V_L CDR3 and previous data from the group also demonstrates other CDRs missing a number

signals in a similar manner (Hall, 2009; Wilkinson, 2009). This lack of signals, and their recovery upon binding, is indicative of a structural motif in exchange on a millisecond to second timescale. As the other CDR loops appear to show these characteristics to a lesser extent than V_H CDR3 it suggests that they are more stable in structure. There may exist, however, an amount of flexibility in these loops, particularly towards the centre of the antigen binding surface and V_L CDR3, which may be a characteristic feature of antibodies. This has so far not been identified in crystal structures that characteristically may favour a low energy state, be affected by crystal packing or be influenced by the conditions used to form the crystals including buffer and temperature. Unfortunately, there is also a lack of supporting data from molecular dynamics simulations that show flexibility of this type. Simulations such as these, however, are often trained on crystal structures, which may not always show flexibility, and often focus on shorter timescales than the millisecond to second timescale that this type of exchange occurs on. The data produced in this group clearly demonstrates that V_H CDR3, often accompanied by other CDR loops, is in a structural exchange on an intermediate NMR timescale.

A large proportion of the work in this chapter focuses on the mapping of chemical shift changes observed for the scFv onto a homology model of the scFv. As discussed in chapter 2, the sequence similarity between 1189scFv and IC8scFv, of which a structure had previously been solved in the lab, is very high. The sequence similarity between the two proteins is so great (84 % identical) that the homology modelling process, which samples the entire Protein Data Bank database for proteins of a similar sequence, selected IC8scFv from which to generate a model. Due to the nature of the modelling process and the sequence similarity between the two proteins, the resulting model is very similar to the IC8 structure (Figure 3.4.1). As the model retains high similarity to

IC8scFv, which shares the same framework sequence, this model was used for all chemical shift mapping.

Figure 3.4.1 – Comparison of the structure of IC8scFv (yellow, PDB accession code 2KH2) and the homology model for 1189scFv. The two structures overlay well with only subtle differences in the core and loops of the proteins, showing that the model produced is sensible and is suitable for use as a template for chemical shift mapping.
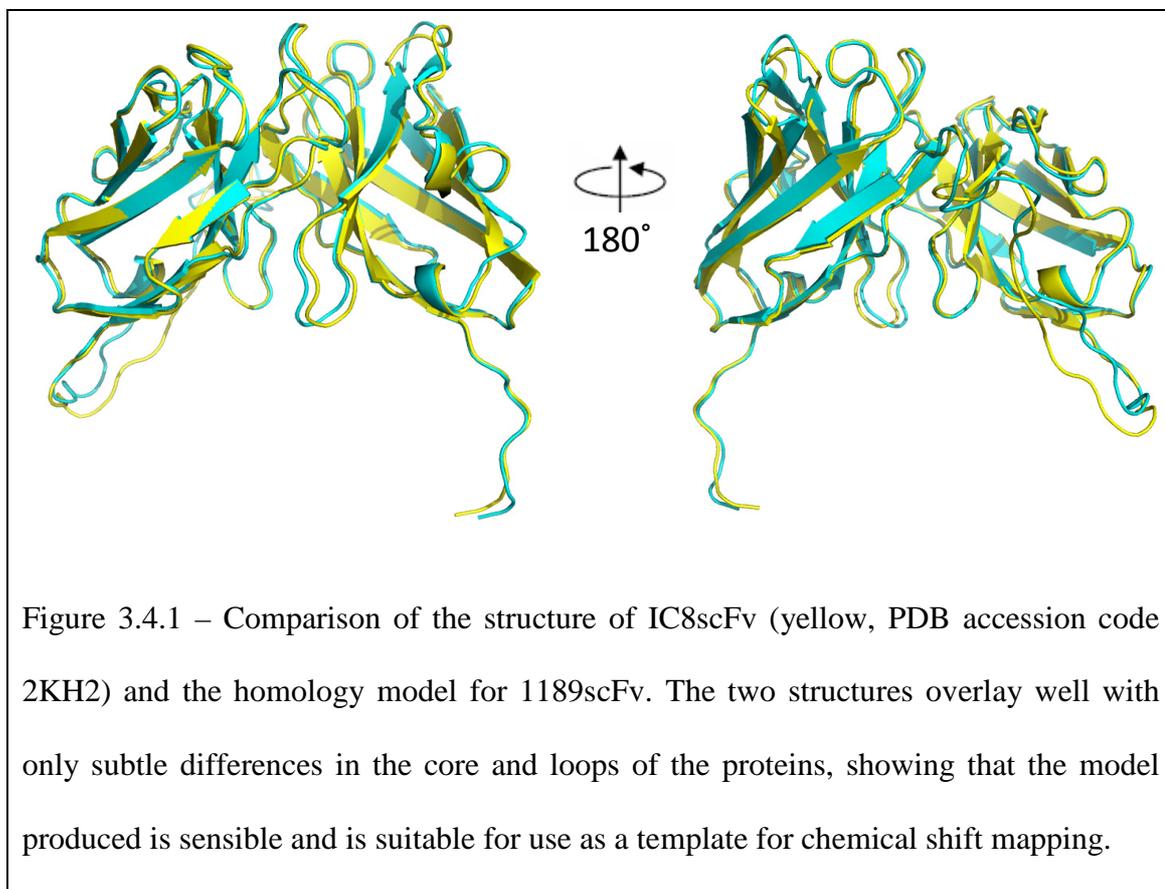
Once a suitable model structure had been produced, the chemical shift changes observed could be mapped to determine both where the binding site for IL-6 is positioned on the scFv and the location of any other interesting conformational changes that have occurred. The single minimal shift approaches and the combined approach produced an overview of 1189scFv binding to IL-6 that provided good quality information on the location of the binding site and regions, showing structural changes upon binding with low levels of insignificant shifts. As expected, the chemical shifts for all of the CDR loops show large changes when binding to IL-6 as the residues become less solvent exposed and form interactions with the antigen. The chemical shift changes

observed for the V_H CDR loops tend to be larger than those seen for the V_L CDR loops. This supports the proposal that the V_H domain may contribute more significantly to antigen binding affinity (Kabat and Wu, 1991) than the V_L, as the V_H is known to retain high affinity when observed as a single entity (Ward *et al*, 1989) and contribute more to the binding interface than the V_L (Davies *et al*, 1990). Also, in fitting with previous results seen for anti-IL-1β antibody fragments (Hall, 2009; Wilkinson 2009), large chemical shift changes are observed at the interface between the two variable domains, which have been shown to be caused by a movement of the V_L domain relative to the V_H upon binding (Hall, 2009). Movement in this region of the antibody structure has been reported previously (Bhat *et al*, 1990; Colman *et al,* 1987; Davies and Padlan, 1992; Stanfield *et al*, 1993) where the V_L is shown to rotate and alter its position relative to the V_H domain upon antigen binding. The data reported in this chapter further indicates that a conformational change of this type may be a fundamental characteristic of antibody binding and key for the proper interaction between an antibody and its antigen. This change in position may also reflect further functional properties of the antibody molecule, possibly relating to the initiation of B-cell signalling and the maturation and activation of antibody producing cells (as discussed in chapter 4). The chemical shift changes were not located exclusively to the CDR and interface residues, and a number of other changes were observed in the framework of the scFv. These changes are concentrated towards the antigen binding portion of the protein and may represent additional binding areas or movement within the molecule produced by the direct binding of the CDR loops to the antigen. These analyses further confirm the results observed for the anti-IL-1β scFvs used by Wilkinson (2009) that scFvs are a useful and easy to use model for antigen binding that can produce a high quality overview of the location and relative size of structural changes.

Further evidence of the changes occurring during antigen binding was produced by comparing the actual assignments obtained for the free and IL-6 bound 1189scFv. This method of analysis has the advantage of comparing residue to residue to directly assess the full scale of the chemical shift changes occurring. It does, however, suffer from lack of data, particularly when the two sets of assignments have residues missing in different places as was the case for this scFv. In contrast to the minimal shift approach, which revealed the largest chemical shift changes occurring at the CDR loops, this method found the largest shifts at the V_L interface which were almost 8 fold larger than the largest chemical shift seen for any other region of the scFv. These residues lie in a well assigned β-strand leading to CDR3 and are flanked by another β-strand also showing significant chemical shift changes and are unlikely to be suffering from assignment errors due to the completeness and confidence of the scFv assignments, the length of the assigned stretch and thorough re-checking of the data. The data missing in this region may hide further chemical shift changes in this area. These large changes clearly demonstrate a significant change at the interface, particularly on the V_L domain, which is in line with the results previously seen with IL-1β recognising Fabs and scFvs, where a movement of the V_L domain is observed relative to the other domains (Hall, 2009; Wilkinson, 2009). Similar results have also been seen elsewhere, such as in the study performed by Stanfield *et al* (2003). Further investigation into this change to determine its function may shed light into alternate functions of antibodies beyond antigen binding.

Further validation of the suitability of scFv proteins as models for whole antibodies was provided by monitoring the chemical shift changes observed for IL-6 when bound by both scFv and Fab. The Fab is known to be the antigen binding portion of the antibody molecule providing full functionality. The comparison of the spectra between the scFv

bound IL-6 and the Fab bound IL-6 suggest that the two proteins are binding in a very similar manner with only slight differences between the two spectra. There are a small number of missing peaks in the Fab bound IL-6 spectra that are likely to be due to the overall size of the complex broadening signals and producing lower quality spectra. There are also a small number of shifted peaks representing some difference between the binding properties of the scFv and Fab. This difference, however, will be subtle as so few peaks have moved. It is possible that the scFv shows an increased flexibility about the variable domain interface that may play a part in these observed shifts. Additionally, differences between the experimental conditions, such as pH or temperature, may also be a factor. These differences represent only a few peaks (13 of 168 assigned peaks: E23, N45, N48, L57, A58, F105, E106, S118, I123, T138, N144, A145 and T162) from a reasonably large data set that otherwise supports the proposition that the scFv and Fab bind in the same manner.

The chemical shift data obtained for IL-6 also highlights the inherent flexibility and plasticity of the protein as mentioned previously in this section. Chemical shift changes are observed for the majority of residues within IL-6 and are not restricted to the scFv and Fab binding area at site III (Boulanger *et al*, 2003. Figure 1.4.1), where the antibodies are targeted for the disruption of the interaction between the IL-6/IL-6R complex and domain 1 of gp130 to block signalling (section 1.4). Flexibility within the IL-6 structure has been noted previously in NMR experiments that observe shorter than expected T_2 relaxation times in some alpha-helical regions as a result of slow movements relative to other helices (Xu *et al*, 1996). This inherent flexibility may allow the continued movement of the CDRs or the V domains between structural conformations, resulting in the loss of NMR signals due to intermediate exchange

processes which explains why some of the CDR residues on 1189scFv are not recovered upon binding.

3.5 – Conclusions

The work reported in this chapter demonstrates that high quality NMR data can be collected from 1189scFv in a variety of labelling formats which can be used to produce backbone assignments for the protein in both the bound and unbound state. The chemical shifts of these assignments have been compared and the differences between them used to highlight both the binding site and other areas of significant structural change within the scFv upon antigen binding. Key areas that demonstrate these changes are the CDR loops that bind to IL-6 and the V_L-V_H interface. The CDR3 loops of both variable domains remained unassigned both when unbound and bound to IL-6. This suggests that a limited number of structural states may be sampled in both the free and bound states. This observation, along with the wide ranging chemical shift changes upon scFv/Fab binding, gives insight to the flexible and plastic nature of IL-6. The large chemical shifts seen at the variable domain interface give evidence for a change in the position of the domains relative to each other upon antigen binding. This re-orientation may be a key factor in antibody binding and may expose or protect additional binding sites that provide further functional roles.

Comparisons of the minimal shift spectra produced for IL-6 when bound to 1189scFv and 1189Fab demonstrate a very similar binding pattern for both proteins. This information confirms that the mechanism of binding for the scFv is very similar to the Fab, validating the scFv as a good, lower molecular weight, model system for antibody-antigen interactions.

Chapter 4 – Investigations into Fab/Fc Interactions and the Initiation of B-cell Receptor Signalling

4.1 – Introduction

The data reported in Chapter 3 supports work by Catherine Hall (2009) that demonstrated domain movements about the V_H/V_L interface. Movements of this type, which have been observed elsewhere (Bhat *et al*, 1990; Colman *et al*, 1987; Davies & Padlan, 1992; Stanfield *et al*, 1993), appear to be fundamental to the nature of antibody/antigen interactions. As this feature is common to a number of unrelated antibodies, and demonstrates similar features in both of the systems studied in this group, it is possible that this property may have functions beyond antigen binding which in itself would not necessarily require a repositioning of the domains. One of the key areas to which this change may be attributable is the initiation of B-cell maturation by signalling through the B-cell receptor (BCR), which currently lacks any significant detail concerning the transfer of an antigen binding event into a signalling event. The BCR itself is highly abundant (in the region of 120,000 receptors per B-cell cell (Yang & Reth, 2010a)) and consists of an antibody molecule of the appropriate class in a 1:1 complex (Schamel & Reth, 2000) with two disulphide linked accessory proteins, Igα and Igβ, which contain the immunoreceptor tyrosine activation motifs (ITAMs) necessary for signalling (Hombach *et al*, 1990) (Figure 4.1.1). Despite a significant amount of evidence gathered shortly after the identification of the BCR in 1970 (Raff *et*

al, 1970), and a general acceptance in the literature that the BCR may be composed of any of the five immunoglobulin classes (Abney *et al*, 1978; Coffman and Cohn, 1977; Mason, 1976; Okumura *et al,* 1976; Venkitaraman *et al*, 1991; Walters and Wigzell, 1970), there is still a surprising persistence of an inaccurate dogma that believes that only IgM and IgD may be expressed on the surface of B-cells. Whilst this is the case during much of the development of the B-cell, once class switching has occurred the nature of the BCR changes to reflect the class preference of the germinal centre, and subsequently memory, B-cells (Coffman and Cohn, 1977; Wienands *et al*, 1990). All heavy chains are able to be expressed along with a class specific transmembrane tail (Wang and Clark, 2003) and are able to associate with the Igα/Igβ signalling dimer on the cell surface. Igα may even show preference to heavy chain class depending on its glycosylation state (Venkitaraman *et al*, 1991). In the case of IgG and IgD, surface presentation is possible even without co-expression and association with the signalling heterodimer (Venkitaraman *et al*, 1991). More recently it has been shown that natural constructs of transmembrane IgG are able to cluster and signal in the same way as IgM (Tolar *et al*, 2009). This evidence demonstrates that IgG is a suitable and useful target for the study of BCR signalling as it is required for the maintenance and activation of both germinal centre B-cells that have undergone class switching and the long lived memory cells that help to provide the secondary adaptive immune response. Therefore any results gathered from this system are also relevant to other antibody classes due to the similar nature of antibody structures and the use of a common signal transducing protein dimer (Igα/Igβ).

Upon antigen binding the BCRs associate and form a structure known as an immune synapse (Fleire *et al,* 2006). In this structure, the ITAMs are phosphorylated by the Src family kinase Lyn and spleen tyrosine kinase (Syk) which initiate the formation of an

166

assembly of intracellular signalling components. Factors such as Syk, phospholipase-Cγ2 (PLCγ2), phospoinositide 3-kinase (PI3K), Bruton's tyrosine kinase (Btk), the Rho-family GTPase Vav and the adapter molecule B cell linker (BLNK) are involved, amongst others, (Kurosaki, 2011; Rolli *et al*, 2002; Schmitz, 1996) as described in Figure 4.1.2. This signalling event drives a number of cascades, leading to the activation of nuclear factor-κB (NF-κB), Jun N-terminal protein kinase (JNK), extracellular signal-related kinase (ERK) and protein kinase C (PKC), the mobilisation of calcium and the rearrangement of the cytoskeleton (Wang and Clark 2003). This ultimately leads to cell differentiation and maturation when paired with the appropriate CD4+ helper T-cell stimulation (Harwood & Batista 2009). The BCR is involved with a number of stages of this maturation process from the progression of the pro-B-cell to the pre-B-cell stage in an immature form using a surrogate light chain on the membrane immunoglobulin (mIg) to the differentiation of mature B-cells and the activation of plasma and memory B-cells of all classes upon antigen recognition (Wang & Clark, 2003). The mIg used in the BCR is developed in stages allowing the maturation and refinement of the antibody and the removal of self recognising cells. Despite the signalling pathways themselves and the downstream effects of antigen binding being well characterised (Cambier *et al,* 1994; Kurosaki, 2011; Kurosaki *et al,* 2009; Wang & Clark, 2003), the binding event itself, how the receptor is able to activate and regulate itself, and the method by which such an abundant receptor is able to remain silent is, as yet, poorly understood. The details of this initiation event, and the investigation of its molecular basis, are key in developing a full picture of the mechanisms of activation for BCR mediated signalling and B-cell development and an understanding of how conditions such as autoimmune diseases and B-cell lymphomas are developed, potentially leading the way into the generation of new treatments.

167
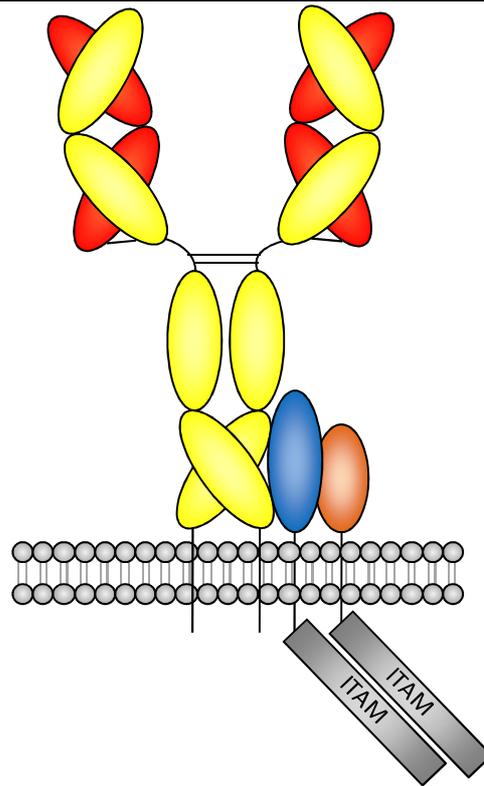
Figure 4.1.1 – A representation of the IgG BCR complex. The membrane bound antibody mIg (red and yellow) is associated in a 1:1 complex with a disulphide linked Igα (blue) and Igβ (orange) heterodimer containing the immunoreceptor tyrosine-based activation motifs necessary for signalling.
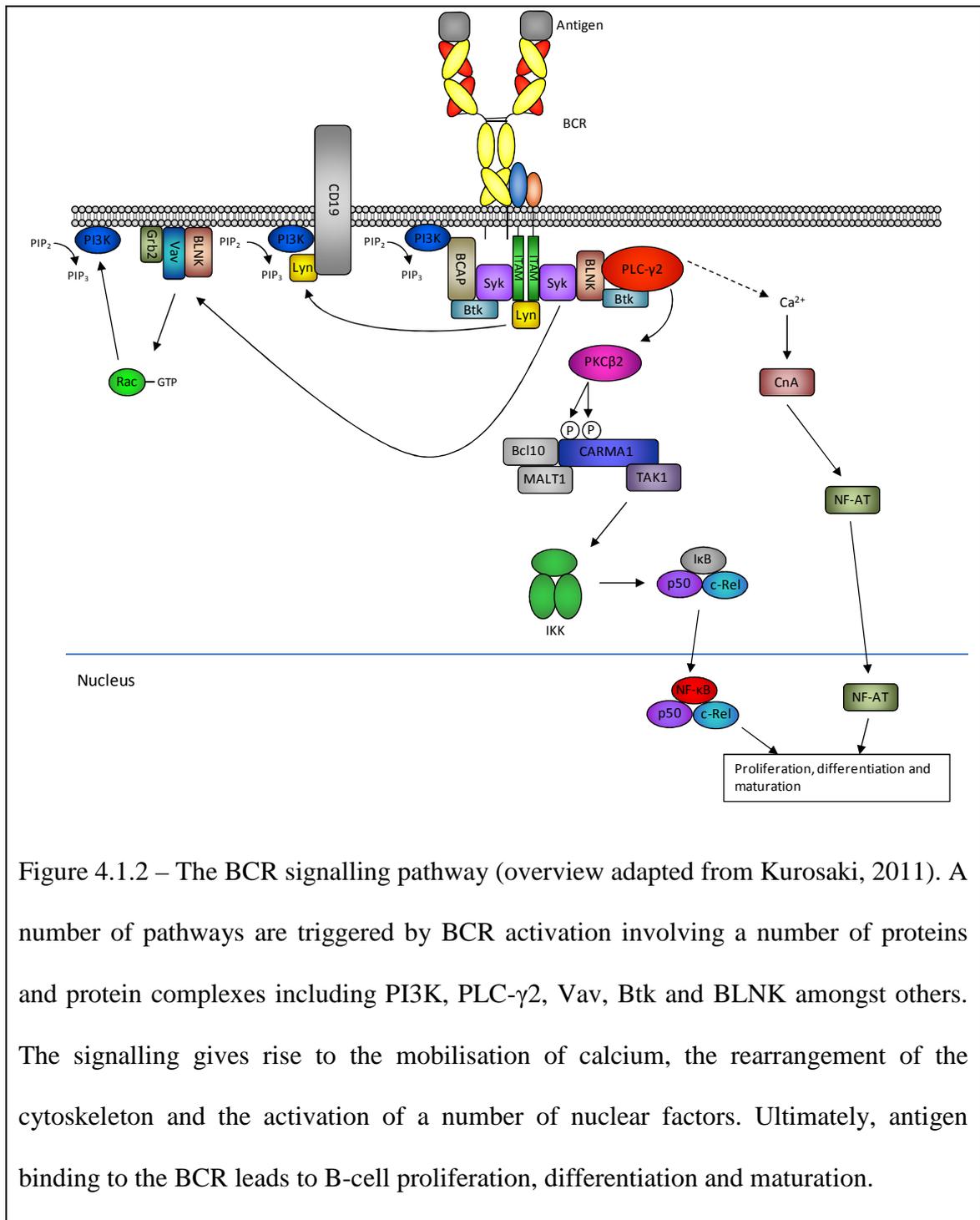
Figure 4.1.2 – The BCR signalling pathway (overview adapted from Kurosaki, 2011). A number of pathways are triggered by BCR activation involving a number of proteins and protein complexes including PI3K, PLC-γ2, Vav, Btk and BLNK amongst others. The signalling gives rise to the mobilisation of calcium, the rearrangement of the cytoskeleton and the activation of a number of nuclear factors. Ultimately, antigen binding to the BCR leads to B-cell proliferation, differentiation and maturation.

Traditionally it has been proposed that signalling is activated by a clustering of the BCR complexes by aggregation with polyvalent antigen. In many cases, this view gave a reasonable explanation of the data observed but has more recently been called into

question. It has been shown on a number of occasions that monovalent antigen is able to trigger a signalling event without itself being cross-linked when presented on a membrane (Tolar *et al*, 2009) and even when in solution eliciting a signalling event (Kim *et al*, 2006). In a study performed by UCB (unpublished) it was shown that soluble, exclusively monovalent, antigen was able to trigger a BCR signalling response using a cell based reporter assay. The interpretation of this data, however, is made less clear by the observation that soluble monovalent antigen may not fully activate a MHC class II mediated response (Kim *et al*, 2006). It may be that two or more separate and/or interacting mechanisms exist for polyvalent and monovalent antigens. Data collected by Tolar *et al* in 2009 demonstrates that, in contrast to polyvalent antigen, the strength of the signalling response produced by monovalent antigen is dependent upon the presence of both the membrane proximal constant domain and transmembrane section of the mIg. Others also argue that a conserved response is unlikely to be triggered by a very large set of ligands, which differ vastly in both size and shape, cross-linking a receptor (Yang & Reth, 2010). In the light of these discoveries it would appear that the traditional model of activation is, at least in part, no longer sufficient to explain the data and subsequently two different models have been proposed.

The first, and most closely related to the antigen cross-linking model, is the conformation-induced oligomerisation model. This model requires that for signalling to occur the BCRs must oligomerise, or form clusters, to bring the ITAMs of the complex into close proximity to allow phosphorylation and the initiation of downstream signalling (Pierce & Liu, 2010). This model is supported primarily by live cell imaging, using fluorescence resonance energy transfer (FRET), confocal laser scanning microscopy (CLSM), single particle tracking (SPT) and total internal reflection fluorescence microscopy (TIRFM), investigating the activation of the receptor. A

combination of these techniques has been used to detect and track BCR complexes and their interactions on live cells when in a resting state and when in the presence of antigen (Pierce & Liu, 2010). During one set of these experiments it was observed that both polyvalent antigen and monovalent antigen were able to induce signalling by clustering of BCRs when presented on a membrane such as the surface of an antigen presenting cell (APC) (Tolar *et al*, 2009). It was also noted during these investigations that the response observed for monovalent antigen was affected by alterations to the transmembrane and membrane proximal constant domains whilst that of polyvalent antigen was not, suggesting that two separate mechanisms may be present. This data was gathered in a study investigating which portions of the antibody molecule were essential for the clustering and signalling of the BCR. It was clear from the outcome that the membrane proximal constant domain (Cμ4 or Cγ3 in this case) was essential for the clustering and signalling of the BCR, and spontaneously signalled when expressed alone on the surface of the cell. The quality of the monovalent signalling response was also affected by mutations of conserved residues in the transmembrane section of the mIg (particularly a WTxxST motif in the N terminal part of the transmembrane helix, Tolar *et al*, 2009) which are known to differ between antibody classes (Schamel & Reth, 2000). This conformation-induced oligomerisation model (outlined in Figure 4.1.3-A) provides some further insight into the mechanism of initiation of B-cell signalling but ultimately still does not provide an explanation of how effects from antigen binding are transferred from the Fab to Cμ4 and into a signalling response.
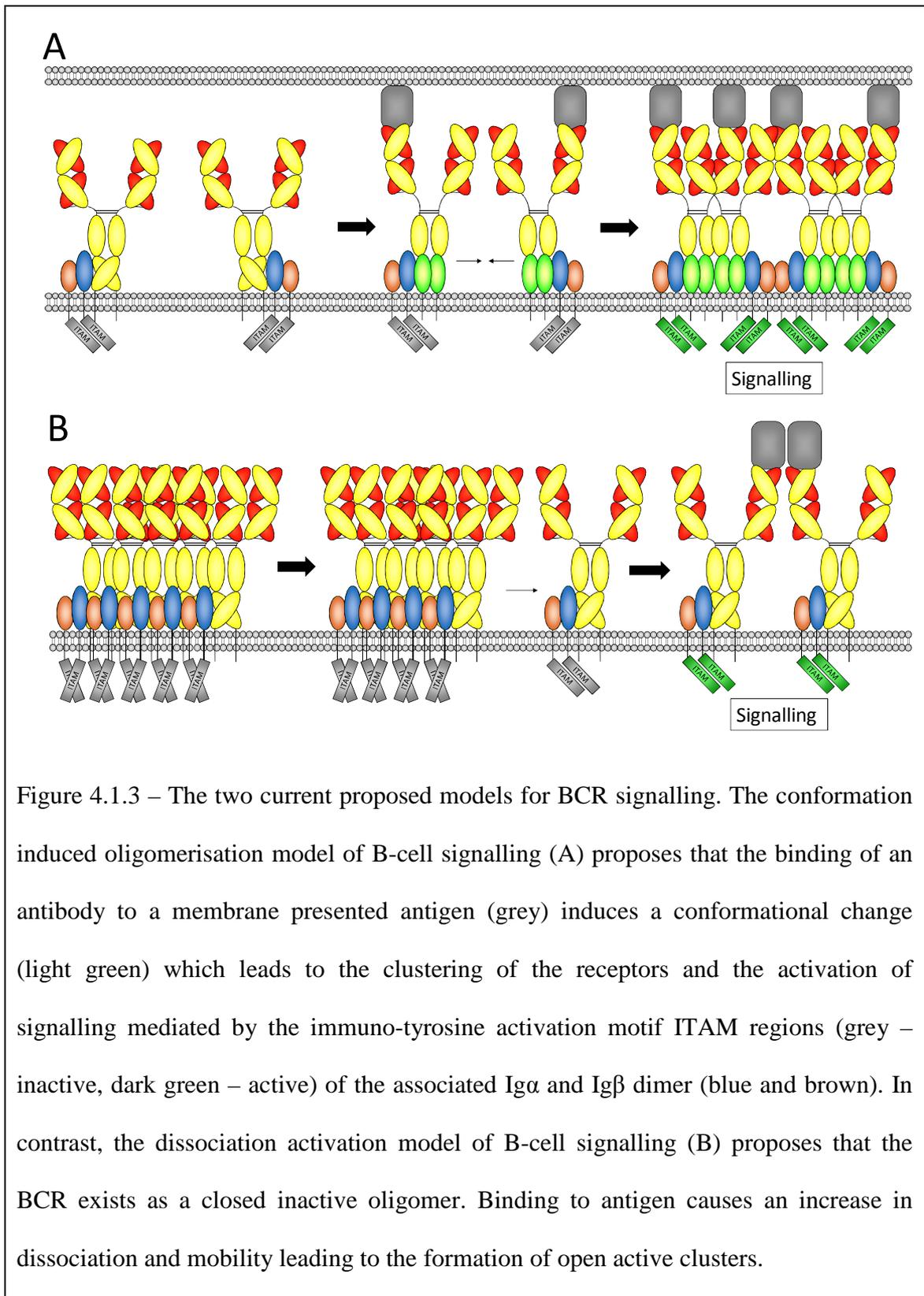
Figure 4.1.3 – The two current proposed models for BCR signalling. The conformation induced oligomerisation model of B-cell signalling (A) proposes that the binding of an antibody to a membrane presented antigen (grey) induces a conformational change (light green) which leads to the clustering of the receptors and the activation of signalling mediated by the immuno-tyrosine activation motif ITAM regions (grey – inactive, dark green – active) of the associated Igα and Igβ dimer (blue and brown). In contrast, the dissociation activation model of B-cell signalling (B) proposes that the BCR exists as a closed inactive oligomer. Binding to antigen causes an increase in dissociation and mobility leading to the formation of open active clusters.

The second proposal for BCR signalling, known as the dissociation activation model (Figure 4.3.1-B), opposes the conformation-induced oligomerisation model in its mechanism of action as it envisages BCRs in pre-associated, tight inhibitory clusters that prevent rather than initiate signalling (Yang & Reth, 2010). Upon antigen binding the cluster is disrupted to create signal active monomers or multimers by removing a mutual steric blocking of the ITAMs or a change in the conformation of the Igα/Igβ complex to allow phosphorylation. This mechanism differs somewhat from the oligomerisation model but may still require the receptor proteins to be in some degree of close proximity to allow phosphorylation (Yang & Reth, 2010). It has been suggested (as reviewed by Treanor, 2012 and Yang & Reth, 2010) that this model better explains how a large range of antigen sizes and shapes can induce the same signalling event. There is, however, significant debate as to the reliability and scope of the data and experiments used to come to this conclusion, with much of the initial data being disregarded as artefacts of the techniques used, particularly the potential of detergent lysis and blue native PAGE assays to detect non-native aggregates of proteins (Pierce, 2009; Peirce & Liu, 2010; Treanor, 2012; Yang & Reth, 2010). The initial work made use of this blue native PAGE technique to view the oligomerisation states of the BCR from detergent lysates of B-cells (Schamel & Reth, 2000). It was found that IgM and IgD BCRs ran as large macromolecular complexes. Further investigation into this model to provide complementary data and attempt to address the criticisms of the native PAGE data utilised bifluorescence complementation assays, reconstruction of the BCR on the surface of a Drosophila Schneider cell and the detection of immunoprecipitated multiprotein complexes by flow cytometry (Yang & Reth, 2010a). The data from these experiments again supported the conclusion that the BCR exists in a large pre-oligomerised state. Evidence is also provided for a difference in clustering propensity

and lipid interactions involving the transmembrane regions that may play a role in the organisation of the clusters and the signalling event. Again, whilst providing good evidence, the results have been questioned by some for focussing too heavily on biochemical assays rather than investigating live cells. The criticism in this debate, however, is not one sided as the majority of live cell imaging data has been gathered from FRET assays. Whilst FRET is good for defined heterodimers, it can be problematic, and potentially misleading, when used for large multiprotein complexes where donor and acceptor ratios may be variable and the orientation of the donor and acceptor proteins may be sub-optimal (Yang & Reth, 2010).

Currently the mechanism of B-cell signalling, and indeed silencing, remains unresolved. Recent work suggests a significant contribution to the regulation of this receptor by the cell membrane with a significant involvement of lipid rafts and membrane compartmentalisation and reorganisation (Sohn *et al*, 2006; Sohn *et al*. 2008; Treanor *et* al, 2010; Treanor *et al*, 2012). This reorganisation and disruption of the cell membrane and the "picket fence" model of membrane organisation (Kusumi *et al*, 1993, Kusumi *et al*, 2005) appears to be particularly mediated by Ezrin-Radixin-Moesin (ERM) proteins. These proteins may regulate the diffusion dynamics of the BCR (results for which have been interpreted for both models) and its association with lipid rafts which may contain the factors necessary for proper signalling. This process has been shown to be initiated by antigen binding which again highlights the poor understanding of this system. Interestingly, it has also been shown that the disruption of the membrane structure alone is enough to initiate a robust BCR signalling event (Treanor *et al*, 2010). This suggests that a combination of potentially weak protein-protein interactions may combine with protein-lipid interactions and membrane dynamics in order to regulate signalling events.

From the results of the work carried out to date it is clear that the overall picture of B-cell signalling is both complex and poorly understood in certain areas. The model initially put forward by Catherine Hall (Hall, 2009), which provided the impetus for the work in this chapter, provided a potential explanation for the activation of BCR signalling by monovalent antigen. This model provides a starting point based on reliable structural evidence for further work that may help to begin to fill in the gap that exists between the antigen binding event and the propagation of signals that allow the proper development of B-lymphocytes.

4.2 – Materials and methods

4.2.1 – Protein Expression and Purification

4.2.1.1 – Expression and Purification of ^{15}N/^{13}C/^{2}H IC8Fab′

IC8Fab′ vector (provided by UCB) was transformed into W3110 *E. coli* and conditioned for growth in D_2O as in section 2.2.2.6. These cells were used to start 100 ml, 100 % D_2O LB starter cultures containing 10 μg ml^{-1} tetracycline that were grown overnight at 37 ˚C/200 rpm. The starter cultures were used to inoculate 500 ml, 100 % D_2O minimal medium expression cultures containing 10 μg ml^{-1} tetracycline, 1 g l^{-1} ^{15}N ammonium sulphate and 3 g l^{-1} $^{13}C_6$ D-glucose to an A_{600} of 0.1. The expression cultures were grown to an A_{600} of approximately 1.0, induced with IPTG to 200 μM and left to express for approximately 28 hours, after which the bacterial cells were pelleted as described in section 2.2.2.4. In this case, the remaining clarified expression medium was also retained and concentrated by ultrafiltration using an Amicon 8400 stirred cell with Millipore polyethersulphone membrane (5,000 MWCO) to a final volume of approximately 50 ml. Cell pellets containing deuterated IC8Fab′ were lysed as described for 1189Fab in section 2.2.2.7 and added to the concentrated expression medium. This pool was adjusted to pH 6.0 using 0.1 M citric acid and filtered using a 0.2 μm filter, before being purified by protein G as in section 2.2.4.3.

Following purification the IC8Fab′ was dialysed into 100 mM sodium phosphate, 2 mM EDTA buffer at pH 6.0. The C terminal cysteine residue was reduced by adding DTT to 2 mM and incubating at room temperature for 30 minutes. The free cysteine could then

be capped by adding N-ethylmaleimide (NEM) to a final concentration of 50 mM and incubating at room temperature for one hour. The capped sample was dialysed into SEC/NMR buffer, concentrated and purified by SEC.

4.2.1.3 – Expression and Purification of Interleukin-1β

A pET21a vector containing a human IL-1β construct, with a T9G mutation to limit the activity of the expressed protein, was provided by UCB and expressed and purified according to the protocol outlined by Wilkinson *et al*, 2009.

4.2.2 – NMR Spectroscopy

4.2.2.1 – Preliminary IC8Fab′/Fcγ samples

Purified $^{15}N/^{13}C/^{2}H$ labelled IC8Fab′ was concentrated to form the basis of a set of preliminary NMR samples investigating the potential interactions between IC8Fab′ and IgG1 Fc.

Human IgG1 Fc (provided by UCB) was concentrated by dialysis into a 20 mM NaCl, 5 mM Sodium Phosphate, 0.004 % Sodium Azide buffer at pH 6.5 followed by lyophilisation and resuspension into one fifth of the original volume. This process was repeated until a concentrated stock of approximately 1 mM was produced.

Samples were generated by combining the two proteins, additional SEC buffer and IL-1β where necessary to generate samples containing 70 μM IC8Fab′ and a 0:1, 1:1, 2.5:1 or 6:1 ratio of Fc to Fab with or without 77 μM IL-1β.

4.2.2.2– The IC8Fab′/Interleukin-1β Complex and IgG1 Fc/IgE Fc Samples

A stock of purified $^{15}N/^{13}C/^{2}H$ labelled IC8Fab′ was divided into two. One portion was combined with a 10 % molar excess of unlabelled IL-1β in 100 mM sodium chloride, 25 mM sodium phosphate, 0.02 % sodium azide, 10 μM EDTA, 200 μM AEBSF at pH 6.5 and purified by SEC, whilst the other was dialysed into the same buffer. These two stocks of material were then concentrated to 300 μM by ultrafiltration to form concentrated stock solutions from which samples could be made.

Human IgG1 Fc (provided by UCB) was concentrated by dialysis into a 20 mM NaCl, 5 mM sodium phosphate, 0.004 % sodium azide buffer at pH 6.5 and lyophilised. The lyophilised protein was recovered in one fifth of the initial volume to increase concentration 5 fold. This process was repeated until a concentrated 1500 µM stock was produced from which samples could be made. EDTA and AEBSF were added to 10 µM and 200 µM respectively.

Human IgE Fc (provided by UCB) in 100 mM sodium chloride, 25 mM sodium phosphate, 0.02% sodium azide, 10 µM EDTA, 200 µM AEBSF pH 6.5 was concentrated by ultrafiltration (Sartorius Vivaspin 6, 10 kDa MWCO) to create a 1500 µM stock from which samples could be made.

In total a set of eight 400 µl NMR samples were made from these stocks. These samples were arranged into pairs (one with and one without Fc) that were dialysed into the same batch of SEC buffer. This would negate, as much as possible, any pH or ionic differences that may be introduced by using different stocks of the buffer made at different points in time. The pairs of samples were as follows:

- 80 µM IC8Fab′ vs. 80 µM IC8Fab′ with 480 µM IgG1 Fc

- 80 µM IC8Fab′/IL-1β complex vs. 80 µM IC8Fab′/IL-1β complex with 480 µM IgG1 Fc

- 80 µM IC8Fab′ vs. 80 µM IC8Fab′ with 480 µM IgE Fc

- 80 µM IC8Fab′/IL-1β complex vs. 80 µM IC8Fab′/IL-1β complex with 480 µM IgE Fc

4.2.2.3 – Data collection and Processing for the Fab/Fc Interaction Studies

All spectra were collected from samples in 5mm Shigemi NMR tubes at 40 ˚C on an 800 MHz Bruker Avance spectrometer. $^{15}N/^1H$ TROSY spectra were collected with acquisition times of 30 ms in F_1 (^{15}N) and 50 ms in F_2 (^1H) for both a 30 minute and 2 hour experiment for each sample.

Solvent suppression was carried out using the WATERGATE pulse sequence. All data was processed using Topspin 2.1 and analysed using SPARKY.

4.2.2.4 – NMR data analysis of the IC8Fab′/Interleukin 1-β interaction with IgG Fc

The data from these experiments was analysed by observing changes in position and signal to noise of all the peaks within the collected spectra. Shifted peaks were identified by eye from overlaid spectra and the changes in signal to noise were detected by comparing values between the two spectra calculated using the signal to noise function contained in SPARKY. Assignments from the IC8Fab′ (Hall, 2009) were transferred onto the spectra collected to identify any shifted peaks or regions of significant signal to noise reduction. Shifted peaks and variations in signal to noise were mapped onto the residual dipolar coupling (RDC) refined model of the IC8Fab′ (Hall, 2009).

4.3 – Results

4.3.1 – Expression and Purification of ^{15}N/^{13}C/^{2}H labelled IC8Fab′

^{15}N/^{13}C/^{2}H labelled IC8 Fab was expressed, purified, capped and concentrated as described in section 4.2.1.1 (Figures 4.3.1 and 4.3.2) with a final purified protein yield varying between 2 and 6 mg l^{-1}. SDS-PAGE analysis detected no other protein other than the IC8Fab′ and demonstrated that the free cysteine had been successfully capped.
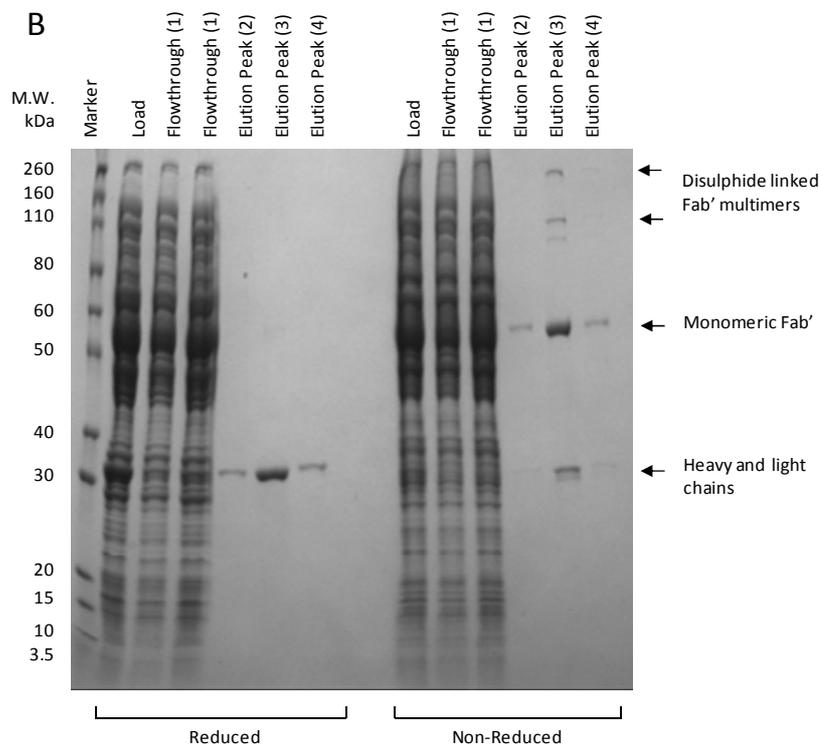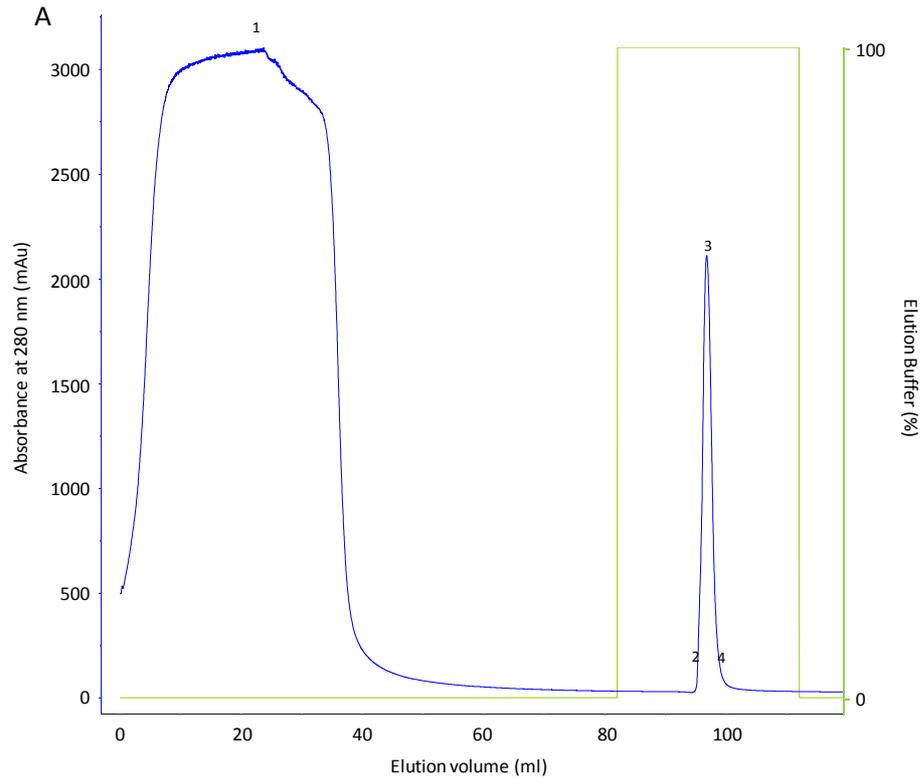
Figure 4.3.1 – The purification of IC8Fab′. IC8Fab′ was purified from the cell lysate
and concentrated expression medium (B - Load) using a protein G column attached to

an Akta FPLC system (A). Samples taken from the purification process were analysed by SDS-PAGE against Novex Sharp molecular weight markers (B) and show that all protein was removed from the cell lysate (1) and that IC8Fab′ was obtained (2, 3 and 4). The characteristic band shift of Fab proteins was seen on the gel when comparing reduced and non-reduced samples. This gel also shows additional disulphide linked multimers of the Fab′.
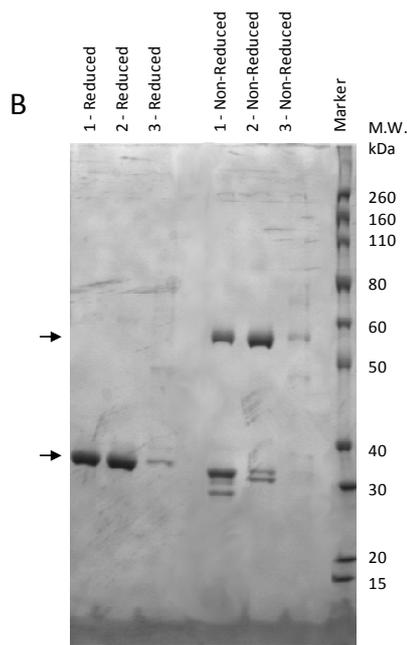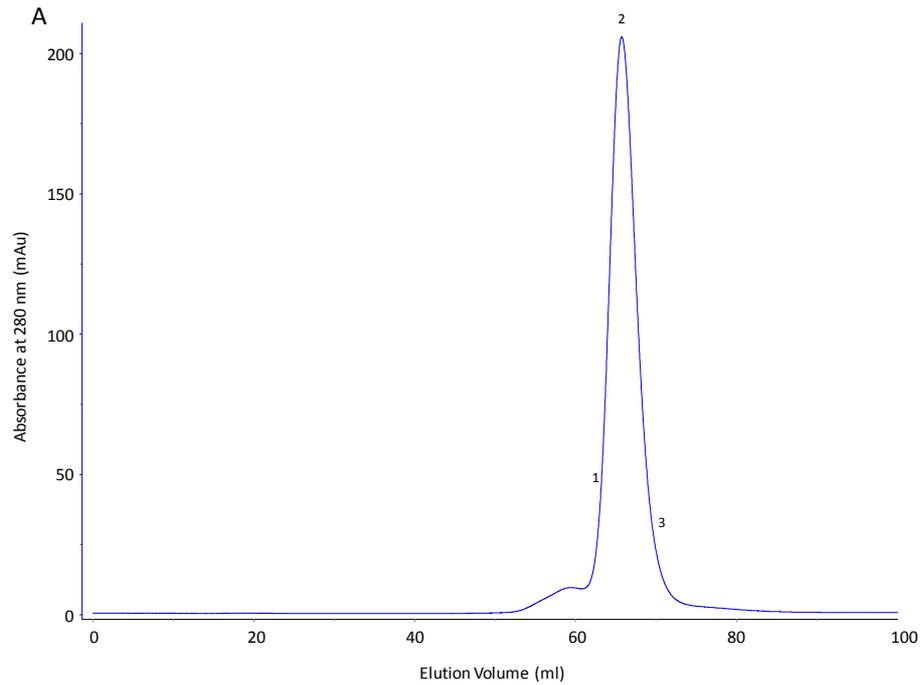
Figure 4.3.2 – The purification of capped IC8Fab′. Protein eluted from the SEC column

(A) was analysed by SDS-PAGE in both reduced and non-reduced samples (B) against

Novex Sharp molecular weight markers. Highly purified protein (2, 35 kDa observed,

48 kDa expected) was produced and the multimeric Fab′ species seen in Figure 4.3.1

were no longer present following the capping of the free cysteine.

4.3.2 – Purification of the IC8Fab′/IL-1β Complex

The IC8Fab′/IL-1β complex, formed as described in section 4.2.2.1, was purified by SEC (Figure 4.3.3-A). Analysis by SDS-PAGE (4.3.3-B) confirmed that the complex had formed and was of sufficient purity for analysis by NMR.
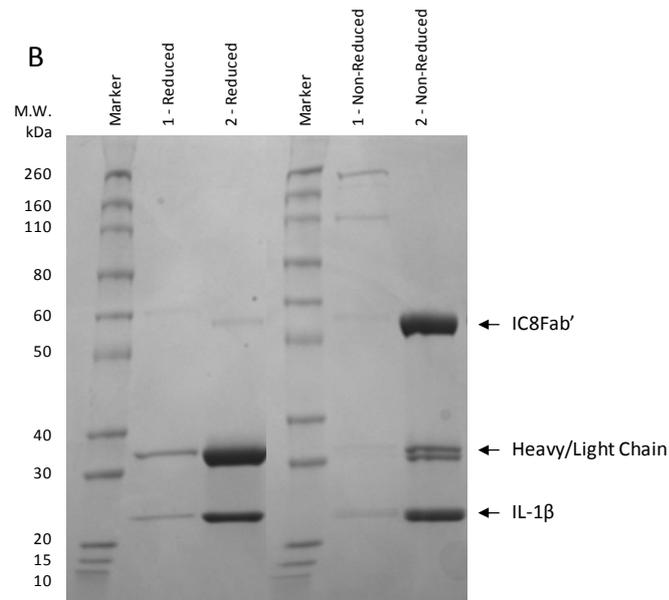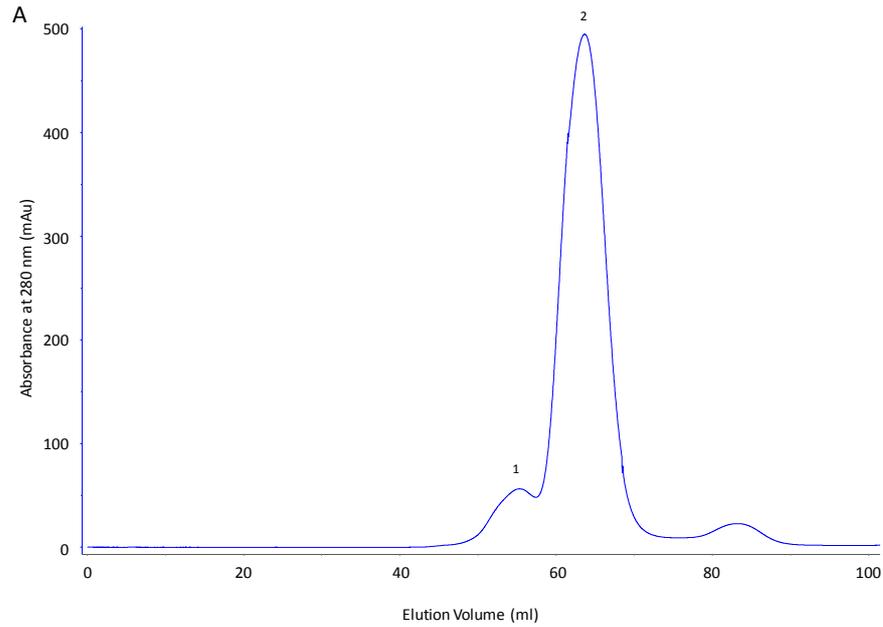
Figure 4.3.3 – Purification of the IC8Fab′/IL-1β complex. SEC purification of the complex resulted in two Fab′ containing elution peaks (A) as determined by SDS-PAGE (B). The major peak (2) was found to consist of monomeric IC8Fab′ bound to IL-1β (45 kDa observed, 65 kDa expected). The additional smaller peak (1) contained a small amount of un-capped IC8Fab′ as disulphide linked multimers bound to IL-1β (>75 kDa observed).

4.3.3 – NMR Spectroscopy

4.3.3.1 – Assignment of Spectra and the Identification of Shifted Peaks

The NMR spectra collected for the Fab/Fc interaction experiments were all of high quality with good signal to noise. The assignments produced by Catherine Hall (2009) were transferred onto the spectra from the reported chemical shift values and used for further analysis (Figure 4.3.4). The resulting assignments were 71 % complete (319 of 444 assigned, 76 % excluding Prolines) for IC8Fab′ and 71 % complete (315 of 444, 76 % excluding prolines) for IC8Fab′ bound to IL-1β. A summary of the assignments is shown in Figure 4.3.5.

Figure 4.3.4 – The assigned $^{15}N/^{1}H$ TROSY spectrum of $^{15}N/^{13}C/^{2}H$ labelled IC8Fab′. The spectrum shown is typical of the quality of spectra obtained for the Fc/Fab interaction experiments with a good level of signal to noise and a good dispersion of signals for a protein of this size. Assignments for the backbone amide signals are labelled by residue type and sequence number. The expanded region clarifies a crowded area of the spectrum.

A

B

Figure 4.3.5 – Summary of the backbone NH assignments used for IC8Fab′ (A) and IC8Fab′ in complex with IL-1β (B). The sequence is shown with the assignments highlighted in blue. Secondary structure is also shown with α-helices in red, β-sheets in green and CDRs as raised lines.

Preliminary ^{15}N/^{1}H TROSY experiments were overlaid to determine if any changes could be seen when increasing quantities of Fcγ were added to the IC8 sample. A number of small shifts could be seen with some evidence of the movement of signals as increasing quantities of Fcγ were added (Figure 4.3.6). A refined set of experiments were produced with carefully prepared samples to confirm this observation (Section 4.2.2.1). The assigned spectra were overlaid to compare non-Fc containing samples to Fc containing samples to determine if any peak shifts could be identified. Overall the spectra analysed showed relatively few minor shifts for IC8Fab with Fcγ (Figure 4.3.7) and IC8Fab/IL-1β with Fcγ (Figure 4.3.8). Fcε was initially chosen as a control with the intention of finding a protein of similar molecular weight and function that would not interact with the Igγ based IC8 Fab′. This, however, did not appear to be the case as IC8 Fab with Fcε (Figure 4.3.9) also produced a small number of shifts. While no discernible shifts were observed for IC8/IL-1β with Fcε, the data presented in the latter parts of this chapter also indicates that an interaction may be occurring and that Fcε was not performing as a good control. Of the shifts seen for all of the samples, only a small number were for peaks assigned to residues. G68, I265 and G326 were shown to be shifted for IC8Fab′ with Fcγ, C360, H384 and T403 for IC8Fab′/IL-1β with Fcγ and H30 and H384 for IC8Fab′ with Fcε, the locations of which are shown mapped onto the structure of IC8Fab′ in Figure 4.3.10. The sample displaying the largest number of shifts (as can be seen from Figures 4.3.7-9) was the IC8Fab′ with Fcγ in the absence of antigen.

Figure 4.3.6 – Overlaid preliminary ^{15}N/^{1}H TROSY spectra of samples containing increasing concentrations of Fcγ. As the Fc is added to the samples at a 1:1 Fc to Fab ratio (red), a 2.5:1 ratio (green) and a 6:1 ratio (purple), some peaks move away from the starting point (0:1, blue) in an Fc concentration dependent manner. Results such as these are the basis of the refined samples and repeated experiments.
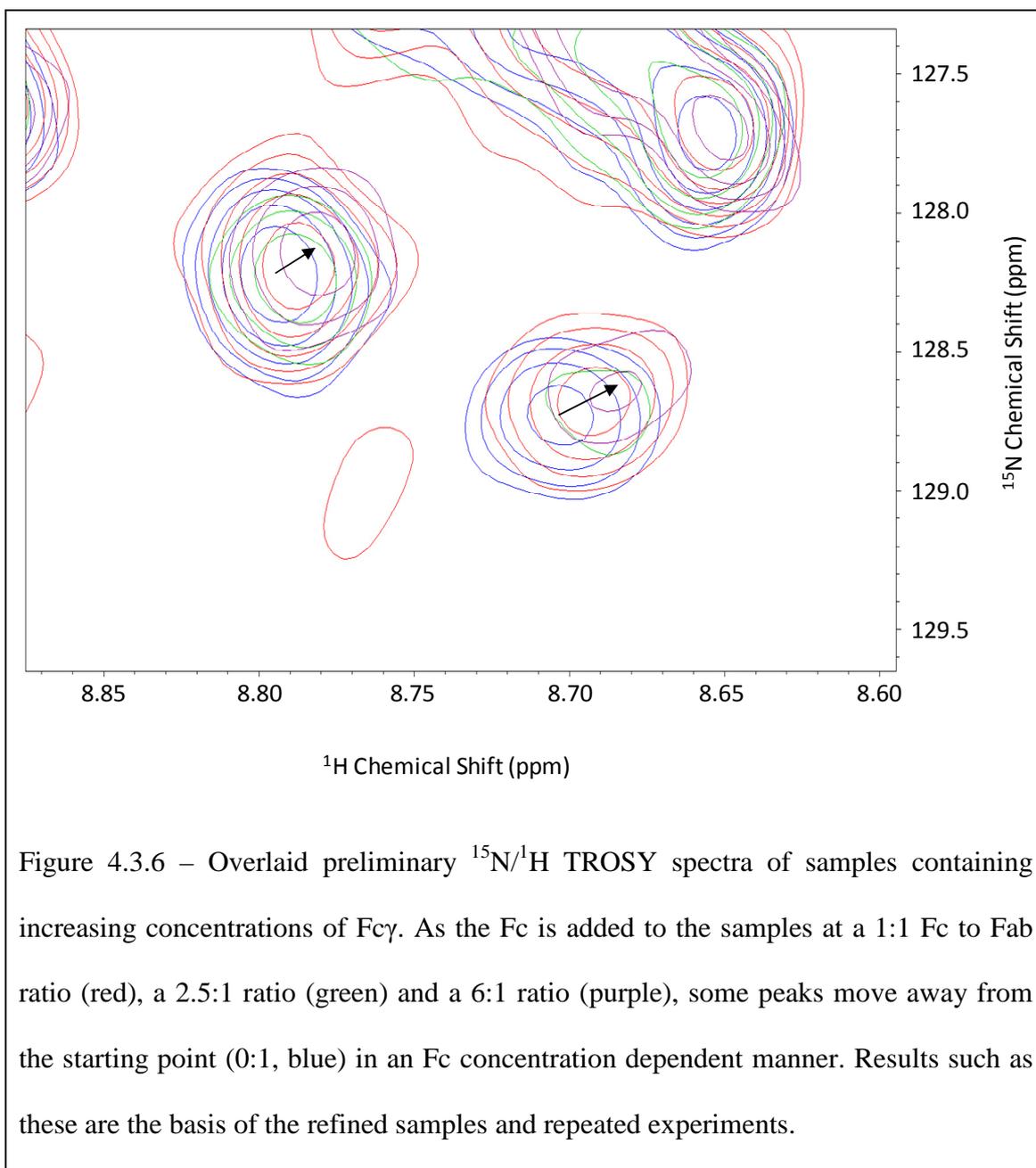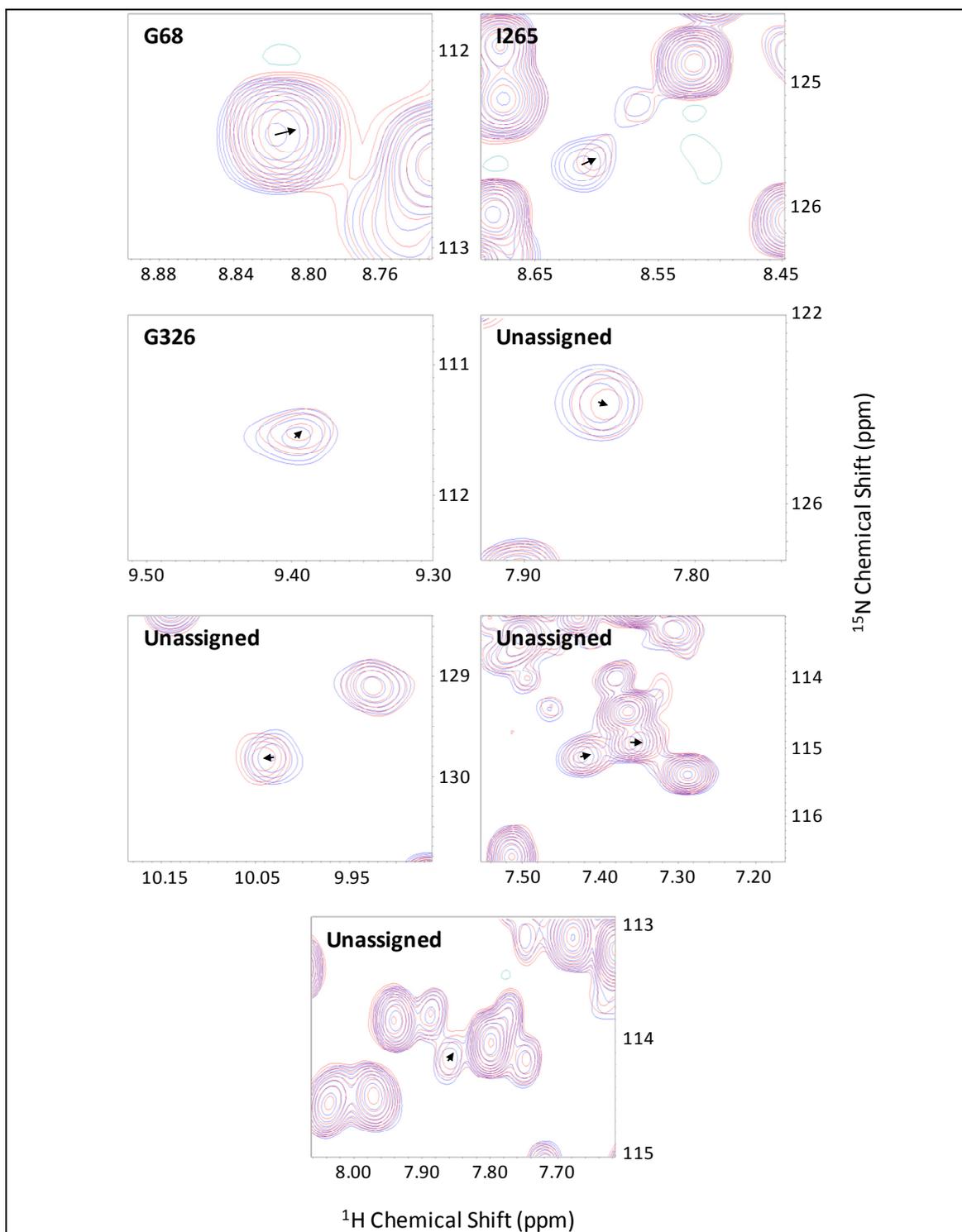
Figure 4.3.7 – Overlaid sections from ^{15}N/^{1}H TROSY spectra representing chemical shift changes observed between IC8Fab′ and IC8Fab′/Fcγ samples. The sample without Fcγ is shown in red and the sample with a 6:1 molar ratio of Fcγ to Fab shown in blue. Arrows highlighting the differences observed are also shown.

Figure 4.3.8 – Overlaid sections from $^{15}N/^{1}H$ TROSY spectra representing the chemical shift changes observed between IC8Fab′/IL-1β and IC8Fab′/IL-1β/Fcγ samples. The sample without Fcγ is shown in red and the sample with a 6:1 molar ratio of Fcγ to Fab is shown in blue. Arrows highlighting the differences observed are also shown.

Figure 4.3.9 – Overlaid sections from ^{15}N/^{1}H TROSY spectra representing the chemical shift changes observed between IC8Fab′ and IC8Fab′/Fcε samples. The sample without Fcε is shown in red and the sample with a 6:1 molar ratio of Fcε to Fab is shown in blue. Arrows highlighting the differences observed are also shown.

Figure 4.3.10 – Locations of shifted assigned peaks mapped onto the IC8Fab′ structure. Peaks shifted in the IC8Fab′/Fcγ experiments are shown in red, the IC8Fab′/IL-1β/Fcγ experiments shown in green and the IC8Fab′/Fcε in blue. H384 (as indicated) is shifted in both the IC8Fab′/IL-1β/Fcγ and IC8Fab′/Fcε spectra. The Kabat numbering of these residues can be found in Appendix 4.2.

4.3.3.2 – Analysis of Signal to Noise Differences in the Fab/Fc Interaction Spectra

It was noted whilst analysing the spectra collected for the investigation of potential Fab/Fc interactions that the signal to noise levels in the spectra observed were reduced in the presence of Fc (Figure 4.3.11). These reductions in signal to noise were not a result of line broadening which would indicate increased sample viscosity. Salt concentration may also affect the signal to noise without line broadening. However, samples were rigorously produced to maintain a stable ionic strength throughout. This

signal to noise reduction may be due to a proportion of the sample being too large to observe (i.e. when bound to the Fc). Further investigation into the nature of this phenomenon resulted in the discovery of a more significant signal to noise reduction in samples containing Fcγ (Figure 4.3.12) and Fcε (4.3.14) in the absence of antigen. The samples with Fcγ also showed specific areas of greater signal to noise reduction with residues C23, G57, F71, D82, A84, N85, T102, I106, K149, Y173, V191-A193, F209, V219, V226, L232, L234, Y246, E260, Y264, I265, T283, S285, T292, L293, Q296, M297, R301, G326, T330, S332, A357, T371, S373, S397, V401, V427 and V431 all being reduced by more than one standard deviation from the mean signal to noise level. This pattern of data loss was mapped on to the structure of the IC8 Fab′ to identify any localisations of data loss in Figure 4.3.16. The data collected in the presence of antigen demonstrated a loss of this specific pattern of signal to noise reduction for Fcγ or a signal to noise reduction of a less significant magnitude over the entirety of the Fab′ signals with Fcε (Figures 4.3.13 and 4.3.15).
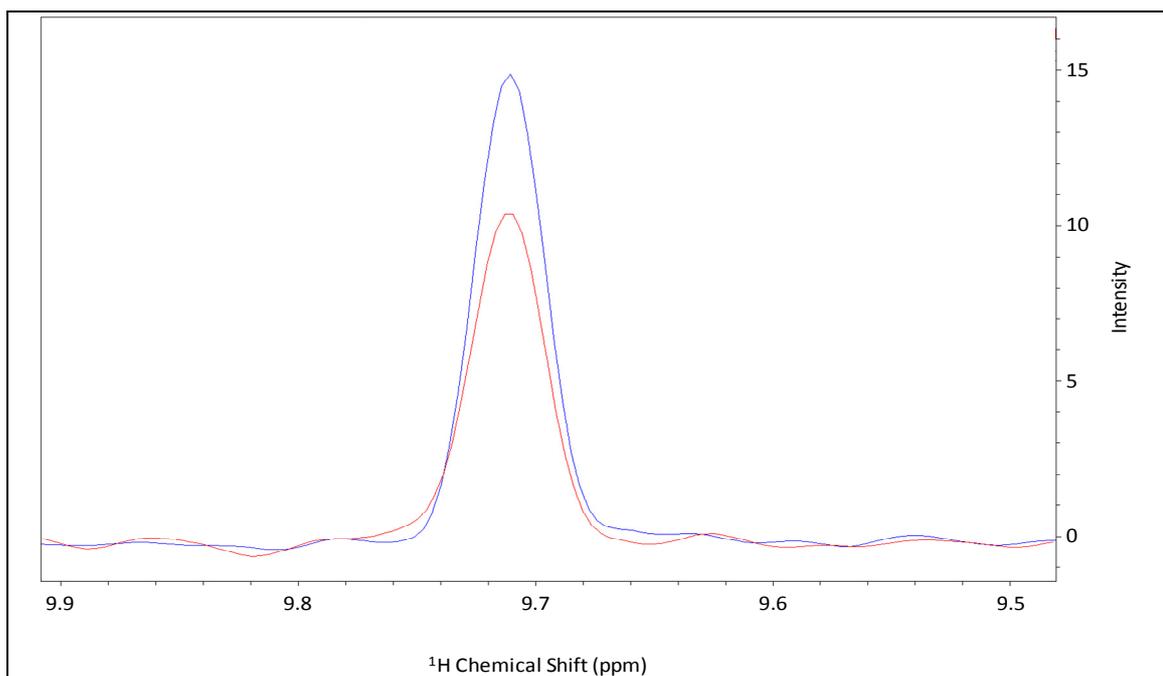
Figure 4.3.11 – Overlaid ^{1}H cross sections of ^{15}N/^{1}H TROSY peaks. These peaks demonstrate a typical example of the reduction in signal to noise seen for the Fc interaction experiments. The sample containing Fc (Red) shows a significant reduction in signal to noise compared to samples without Fc (Blue) that is not a result of line broadening.

Figure 4.3.12 – A histogram representing the signal to noise changes observed for each residue in IC8Fab′ upon the addition of Fcγ. The signals observed show an average of 81 ±12 % of the signal to noise values seen for the control sample (no Fc), with some residues showing a more significant reduction. Missing data is represented by asterisks (red).
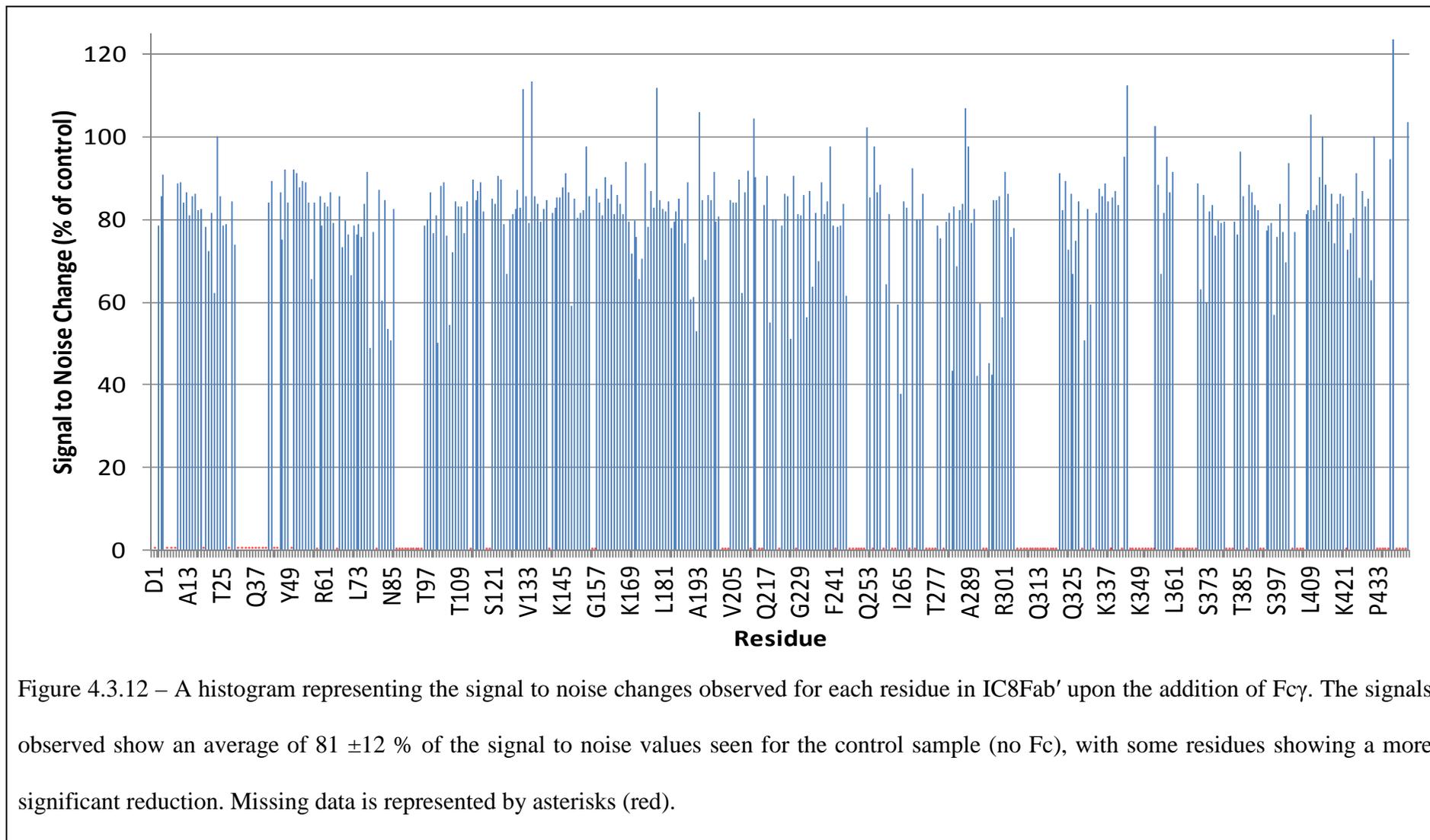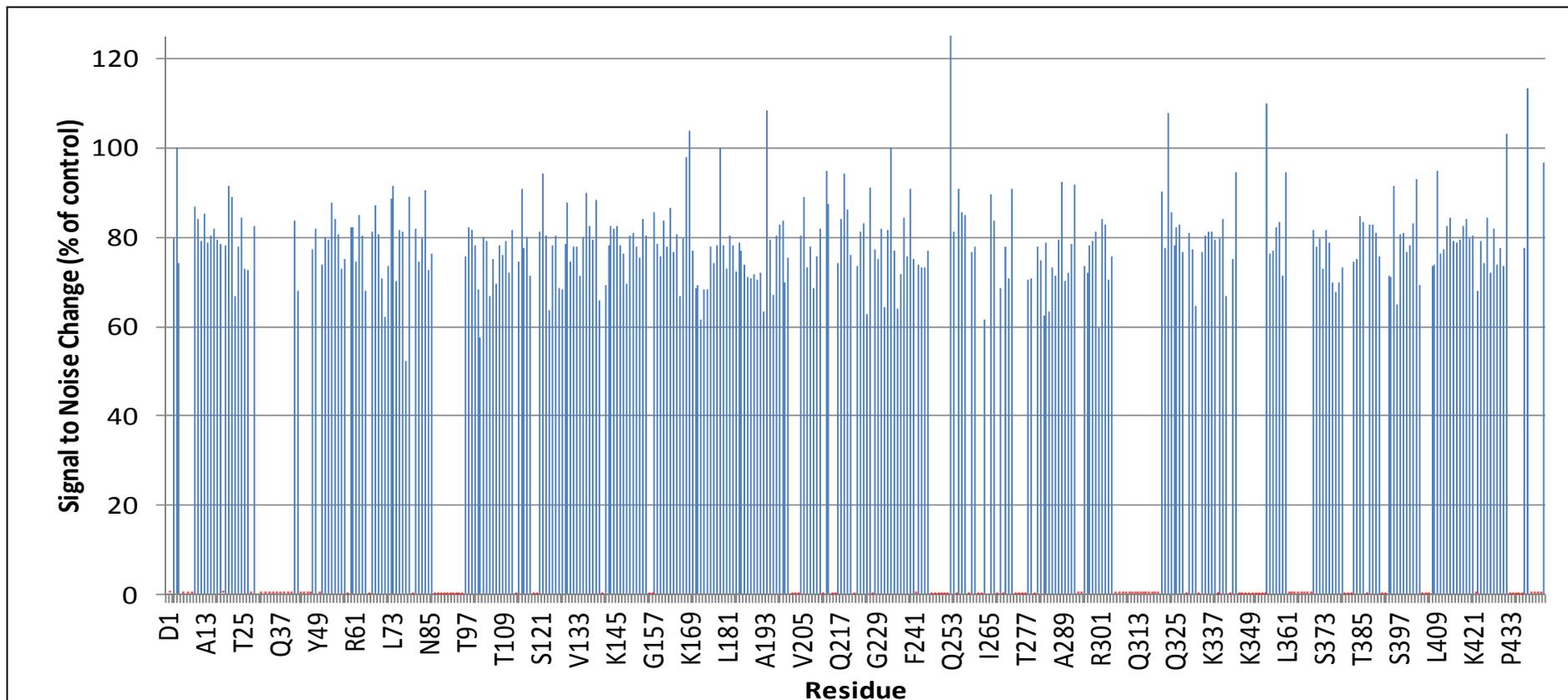
Figure 4.3.13 - A histogram representing the signal to noise changes observed for each residue in IC8Fab′ in complex with IL-1β upon the addition of Fcγ. The signals observed show an average of 79 ±9 % of the signal to noise values seen for the control sample (no Fc). However, the areas of further reduced signal to noise are no longer present. Missing data is represented by asterisks (red).

Figure 4.3.14 - A histogram representing the signal to noise changes observed for each residue in IC8Fab′ upon the addition of Fcε. The signals observed show an average of 58 ±12 % of the signal to noise values seen for the control sample (no Fc). The variation in the values of the signal to noise change makes it difficult to locate any specific areas of further signal reduction. Missing data is represented by asterisks (red).

Figure 4.3.15 - A histogram representing the signal to noise changes observed for each residue in IC8Fab′ in complex with IL-1β upon the addition of Fcε. The signals observed show an average of 73 ±9 % of the signal to noise values seen for the control sample (no Fc). Missing data is represented by asterisks (red).

Due to the significant and specific nature of the signal to noise losses observed for IC8Fab′ in the presence of Fcγ but absence of IL-1β areas of reduced signal to noise were mapped onto the model produced previously for IC8Fab′ (Hall, 2009. Figure 4.3.16).

Figure 4.3.16 – Areas of significant signal to noise reduction in the presence of Fcγ mapped onto the structure of IC8Fab′. Changes in signal to noise over one standard deviation away from the mean are coloured in light pink, over two standard deviations in pink and over three standard deviations in red. Missing data is coloured in yellow.
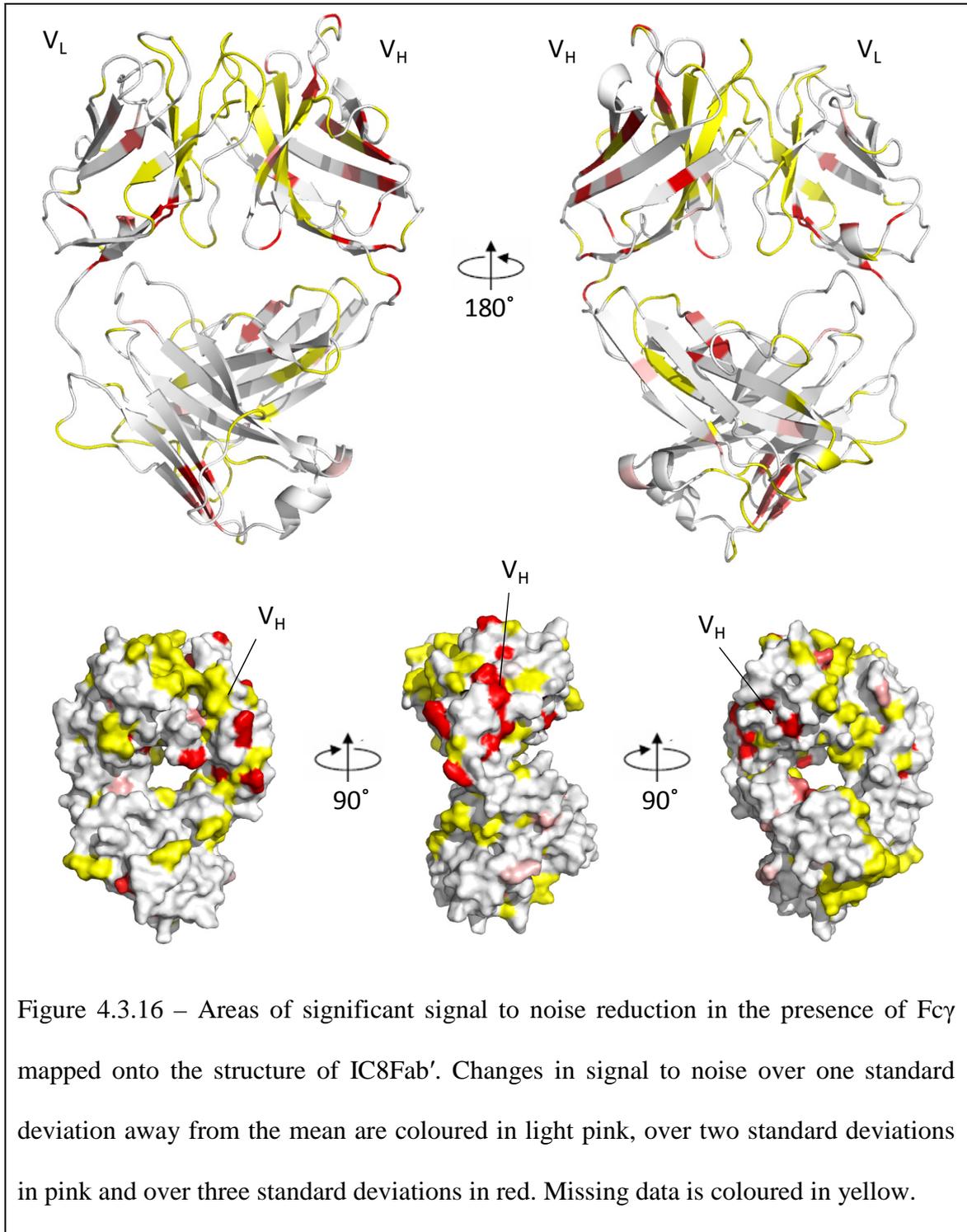
In addition to the effects seen, it was also noted that the initial preliminary experiments (as shown in Figure 4.3.6) also demonstrated an increasing decrease in signal to noise as increasing quantities of Fcγ were added to the sample (Figure 4.3.17).

Figure 4.3.17 – Overlaid cross sections of a typical peak from the preliminary ^{15}N/^{1}H TROSY spectra of samples containing increasing concentrations of Fcγ. As the Fc is added to the samples at a 1:1 Fc to Fab ratio (red), a 2.5:1 ratio (green) and a 6:1 ratio (purple), it can be demonstrated that peaks such as these show an increasing loss of signal to noise without any significant change in line width when compared to the starting point (0:1 ratio, blue).

4.4 – Discussion

The main aim of the work reported in this chapter was to provide further insight into the potential mechanisms of BCR mediated signalling, to assess the key features of a conformational change and interaction model developed within the group (Hall, 2009), and to determine how this model may fit in with the two main theories for the activation of BCR signalling. The investigation focussed on an NMR based method to detect interactions between the labelled IC8Fab′ and unlabelled Fc as described in section 4.1 and 4.2. NMR provides a useful tool that is able to observe weak or transient interactions that may not be detectable using other techniques, and may provide information on the location of interactions by mapping changes in signals from experiments onto a structure. The method used, however, will suffer from a number of factors, mainly relating to the way the samples are produced. As the Fc and Fab are no longer tethered, and the Fcs are no longer attached to a membrane, there will be a number of energetic penalties included in the system that may affect the dynamics and affinities of binding events. The limitations of NMR also mean that signals from any complex formed, which may be approaching 115 kDa, may be significantly broadened or unobservable.

The initial investigation into potential binding events between the Fc and Fab focussed on a minimal shift technique similar to that used in Chapter 3 for 1189scFv by monitoring changes in chemical shift values to determine potential binding sites and effects of interactions. The number and magnitude of the chemical shifts seen for these samples was low, with only a small number of shifted peaks moving a relatively short distance to the next peak. Often these peaks had no assignment and so were not useful

for mapping onto a structure. The lack of significant shifts rules out the possibility of any strong interactions with tight binding of the Fab to the Fc. This in turn leaves only two possibilities. The first is that there is no interaction at all and that the shifts observed are a result of pH or ionic strength changes within the sample. As the samples were dialysed into the same batch of buffer at the same time it is unlikely that this is the cause. In addition, the peaks that are shifting are also not exclusively of residues that are highly sensitive to changes such as these (i.e. acidic and basic residues and histidines). The second possibility is that the interaction is weak and transient in nature and that the proportion of sample in a complex is relatively low with very broad signals, potentially explaining the results observed. The peaks on any potential site of interaction would represent an average of both the interacting and non-interacting state. This would mean that for a relatively low proportion of a bound state only a relatively small change would be seen (for example 10 % bound would show only 10 % of the maximal chemical shift change). This may also be exacerbated if the peak representing the complex was significantly broader. This explanation may be proven by titrating in higher concentrations of the binding partner (in this case the Fc) and observing peaks moving away from the starting point. This feature was seen in some peaks during the preliminary experiments (Figure 4.3.5) and was confirmed in the refined experiments where a small number of peaks (examples in Figure 4.3.6-8) showed a change in chemical shift values. Throughout these experiments the peak shifts were small and observed only on a few peaks but may provide evidence for an interaction. Unfortunately these shifts are too few to provide any definite binding patch.

During the analysis of the NMR experiments collected for the investigation into Fab/Fc interactions, it was also noted that the $^{15}N/^{1}H$ signal intensity for the Fab′ in Fc containing samples was significantly reduced when compared to the controls. This

reduction was not a result of line broadening and therefore was unlikely to be due to viscosity. In addition, the 90 ˚ pulse observed for maximal signal in the NMR experiments did not change, suggesting the samples were of the same ionic strength, another factor which may have caused a loss in signal to noise. One remaining cause would be that a proportion of the sample was unobservable in some way, for example in a large complex that has significantly broader and potentially unobservable signals (i.e. Fab bound Fc). In this case it is possible to tentatively estimate a binding affinity based upon the average proportion of signal lost when Fc is added. If it is assumed that, for example, an average 19 % signal drop (as observed for IC8 with Fcγ) represents 19 % of the added Fab bound to Fc (19 % of 80 μM and so 15.2 μM bound). This would leave 64.8 μM unbound Fab and 464.8 μM unbound Fc. These numbers may then be used in the equation:

$$Kd = \frac{[Fab][Fc]}{[FabFc]}$$

From this equation, and the standard deviations observed for the data, we can estimate the potential Kd for each sample to be as follows:

IC8 + Fcγ = 1981 ± 701 μM

IC8/IL-1β + Fcγ = 1742 ± 465 μM

IC8 + Fcε = 616 ± 221 μM

IC8/ IL-1β + Fcε = 1239 ± 332 μM

These Kd values are estimates and should be taken as such. They do, however, provide additional insight into the effects of Fc binding in the presence and absence of antigen,

particularly when using Fcε instead of Fcγ. This data gives evidence to suggest that the binding affinity of IC8 and Fcε is reduced upon the addition of IL-1β.

In addition to the general signal to noise loss associated with the Fc containing samples, specific areas of further signal to noise loss were observed upon the addition of Fcγ to IC8 in the absence of IL-1β. These features that can be observed on the histogram in Figure 4.3.10 and may be as much as 3 standard deviations below the mean drop in signal to noise. These features were mapped onto the structure of IC8 to determine if any specific locations of binding could be observed (Figure 4.3.15). Once mapped it was observed that there was a concentration of these features on the V_H domain of the Fab of a more significant level than the rest of the protein. Additional features were located on the V_L domain, the C_H domain at the interface with V_H and at the base of the C_L domain. Unfortunately, due to a significant amount of missing data around the V_L/V_H interface, it is not clear if there is any pronounced effect in this region. What is significant from this data is the number of features located on the V_H and the possible importance of this domain for Fab/Fc interactions and potentially BCR signalling which appears to be modulated by the presence or absence of antigen.

From the data presented in this chapter there are strong indications of some form of interaction occurring between the Fab and Fc. What is less clear is the exact nature of this interaction and how this information impacts upon the models proposed for BCR signalling. The titration movement of a small set of signals and the signal to noise features observed imply a transient, low affinity interaction. The nature of this interaction, particularly the affinity of it, will be affected by the lack of a covalent link between the individual components of the antibody and the lack of a membrane attachment introducing entropic penalties on the K_D of the binding. The detachment of

207

the Fab from Fc may allow positions of interaction that otherwise would be impossible due to steric limitations or may disfavour a particular conformation due to lack of tethering. It is also well reported that membrane topology and rearrangement of the 'picket fence' model (Kusumi *et al*, 1993; Kusumi *et al*, 2005) is important for proper BCR signalling (Treanor, 2012) and the effect of removing this from the observed system is unknown. The evidence collected here suggests that a non-covalent binding interaction occurs between the Fab and the Fc. Furthermore, this interaction appears to be both transient and affected by the presence of antigen in the system to the extent that the Fc/Fab interaction is altered and/or reduced upon binding. This observation directly opposes the model developed initially (Hall, 2009) which proposed that antigen binding promotes an interaction. The observations made in this chapter, however, may indeed still fit with the two current proposals for the initiation of BCR signalling depending upon whether an intra- or inter-receptor interaction occurs.

The two current models of BCR signalling are the association activation model or the dissociation activation model (as described in section 4.1). Both of these models conclude that an interaction causing clustering of the BCR occurs but differ on what affect this has. The conformation-induced oligomerisation model states, as its name suggests, that the clustering of the BCR is triggered by a change in conformation of the receptor which in turn causes the association of the receptors which is required to produce a signalling event. This could be considered to be a more traditional view of the system, as it has long been known that polyvalent antigen is able to readily activate the BCR, although how such a diverse set of ligands is able to do this remains unclear. Recent studies have also shown that monovalent, membrane tethered antigen is capable of triggering clustering of the BCR despite being free to move within its membrane and not providing the cross-linking of polyvalent antigen (Tolar *et al*, 2009). It was this

study, in which the membrane proximal constant domain (e.g. Cμ4 or Cγ3) was found to be critical for the clustering of the receptor and the activation of signalling, that provides the basis for the first model proposed that potentially fits with our observations. It is the ability of the membrane proximal domain, in the absence of the other constant and variable domains, to spontaneously cluster and activate signalling that suggests that some form of occlusion of a binding surface may be provided by the rest of the antibody molecule (Pierce & Liu, 2010). Taking this information and the data gathered in this chapter, we suggest that one possible explanation for the Fab/Fc interaction is that it provides some form of block on the association of BCRs. This could either be by the protection of a binding site or by steric hindrance of a clustering event caused by the Fab arm rotating back to protect its own Fc. This protection may be disrupted upon antigen binding allowing the association of the Cμ4 (or equivalent) domains (Figure 4.4.1). For this theory to hold a number of additional criteria must also be met. The linker between Fab and Fc must be flexible enough to allow this interaction. In the case of IgG1 it has been reported that the angles and distances involved are possible with the antibody molecule even favouring a position that brings Fab and Fc closer together than the traditional 'Y' shape commonly depicted (Zheng *et al*, 1992). In the case of other IgG isotypes, and other antibody classes such as IgE, this may not be possible as linker lengths and flexibilities differ (Saphire *et* al, 2002; Roux *et al*, 1997; Zheng *et al* 1992). Even with the most flexible IgG subclasses such as IgG1 and IgG3, steric clashes and limits on flexibility around the linker may make this interaction impossible. Another issue this model must overcome is the fact that it is only Cμ4 that spontaneously clusters and not Cμ4/Cμ3. For this model to fit it requires an additional process promoting clustering, possibly suggesting why some evidence shows that monovalent antigen only produces a signalling event when presented on a

209

membrane, resulting in the generation of forces that expose a binding site (Pierce &

Liu, 2010).
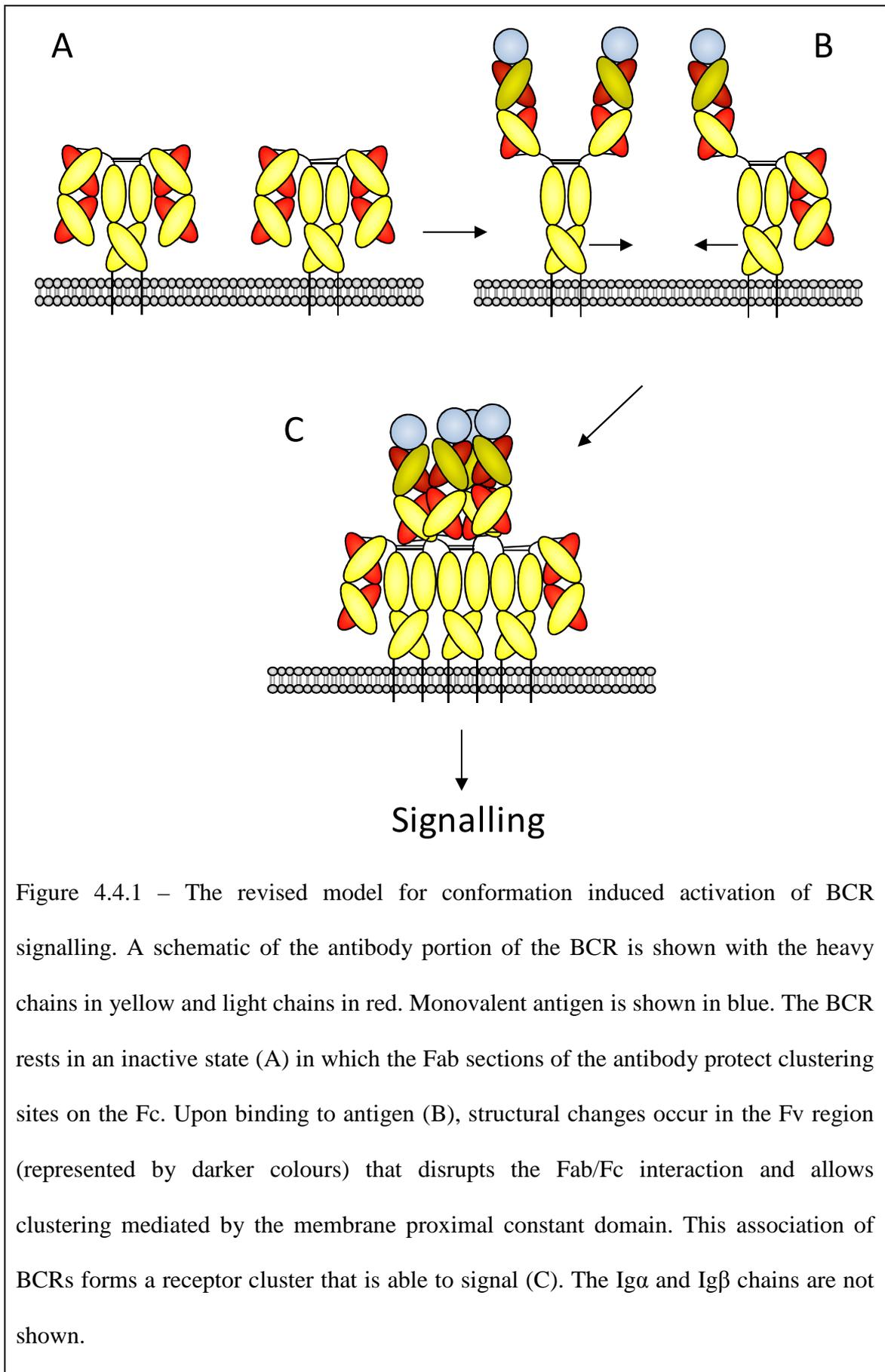
Figure 4.4.1 – The revised model for conformation induced activation of BCR signalling. A schematic of the antibody portion of the BCR is shown with the heavy chains in yellow and light chains in red. Monovalent antigen is shown in blue. The BCR rests in an inactive state (A) in which the Fab sections of the antibody protect clustering sites on the Fc. Upon binding to antigen (B), structural changes occur in the Fv region (represented by darker colours) that disrupts the Fab/Fc interaction and allows clustering mediated by the membrane proximal constant domain. This association of BCRs forms a receptor cluster that is able to signal (C). The Igα and Igβ chains are not shown.

The dissociation activation model proposes that rather than clustering acting to trigger BCR signalling it instead inhibits it by keeping the receptor in a tight, inactive state where the ITAMs are protected from phosphorylation by close proximity to other receptors (Yang & Reth, 2010). Upon binding to antigen, this interaction is disrupted and the BCRs are able to dissociate and signal. The data presented in this chapter may fit this model quite well. An interaction, possibly between a V_H and an Fc of an adjacent receptor, may help produce an inactive receptor cluster. Upon antigen binding this interaction may be altered by a movement of the variable domains to an extent where the inactive cluster begins to disperse and signalling is allowed to occur (Figure 4.4.2). This model may also help to describe the observations made at UCB where soluble monovalent antigen is able to initiate signalling by disrupting this interaction. It may also be able to account for the reduced flexibility of some antibody classes and subtypes, as the angle required in the linker region for the inter-molecular interaction of the Fab and Fc of a neighbouring BCR may be less severe than the angle required for an intra-molecular interaction.
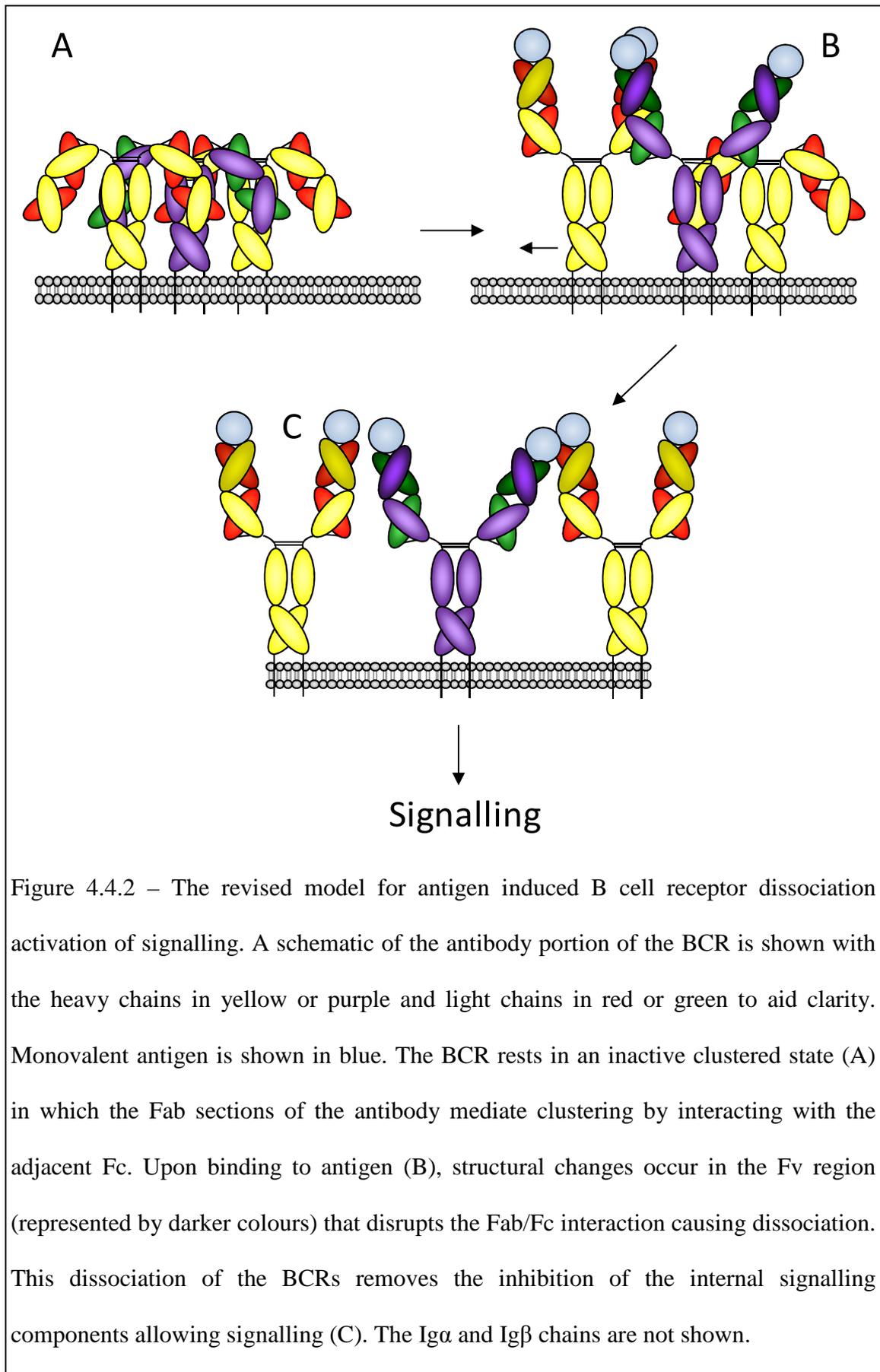
Figure 4.4.2 – The revised model for antigen induced B cell receptor dissociation activation of signalling. A schematic of the antibody portion of the BCR is shown with the heavy chains in yellow or purple and light chains in red or green to aid clarity. Monovalent antigen is shown in blue. The BCR rests in an inactive clustered state (A) in which the Fab sections of the antibody mediate clustering by interacting with the adjacent Fc. Upon binding to antigen (B), structural changes occur in the Fv region (represented by darker colours) that disrupts the Fab/Fc interaction causing dissociation. This dissociation of the BCRs removes the inhibition of the internal signalling components allowing signalling (C). The Igα and Igβ chains are not shown.

The antigen modulated Fab-Fc interaction described in this chapter may be accommodated within both of the currently proposed models. Despite not providing conclusive proof for either model, the results described in this chapter provide the first potential molecular basis for an antigen mediated interaction that may affect BCR signalling. The two key insights of an antigen induced change in the V_L/V_H domain interface and the resulting loss of interaction between the Fab and Fc have been added to the ideas surrounding BCR signalling. For either proposal to be definitively proven, additional work must be undertaken to further investigate the nature and potential positioning of binding events. Other potential effects that may regulate Fab and Fc interactions, such as steric hindrance of a binding site by an antigen, need to be ruled out. As the conformational change between the antigen free and antigen bound states are relatively small when compared to the change in orientation seen from one antibody to the next, a clear identification of a binding site would be essential for providing further evidence. This would help to determine whether movements in the V_L/V_H interface are able to influence this interaction, or whether other explanations may be more fitting. NMR experiments using paramagnetic relaxation enhancers are useful for identifying interactions of a transient nature on large protein complexes, and so may provide a useful tool for future work. Other techniques, such as introducing mutations to interrupt Fab/Fc binding, may also be useful but would require a large amount of work to identify a binding site on such a large protein, and may affect the affinity of the antibody for its antigen, potentially disrupting or altering the conformational changes that have been observed upon binding. Additional NMR control experiments must also be undertaken with the addition of similarly sized, completely unrelated proteins to samples rather than the Fc. Fcε was initially selected as a control as a similar protein that should not interact with a Fab of a different antibody class. This choice, however,

considering the structural similarities to Fcγ and the apparent presence of a higher affinity interaction, may have been ill advised. Further NMR experiments may be useful which look at higher concentrations of the Fc to see if further signal to noise decreases or chemical shifts may be seen. Chemical shifts may also be observed for the Fc itself, although isotopic labelling in the current mammalian expression systems may be a challenge. Other experiments such as cell based reporter assays that assess the BCR on the surface of a cell may also be useful in looking at potential ways of blocking signalling to attempt to provide evidence for or against the models proposed.

4.5 – Conclusions

The work reported in this chapter demonstrates the testing of a proposed model for BCR signalling by NMR spectroscopy and the subsequent updating of the model based upon the results observed for the interactions between Fab and Fc and the current opinions and data found in the literature. The results observed here have the potential to support both of the current theories of BCR activation and suggest new mechanisms of the initiation of signalling for both. High quality data has been collected and analysed to determine the nature of the proposed Fab/Fc interaction and demonstrates a definite antigen mediated effect. The observation of specific patterns of signal to noise reduction when observing IC8 and Fcγ in the absence of antigen implicates the V_H domain as being central in this interaction with an increased number of these features being located on the residues within this domain. This information, in combination with the work described in Chapter 3, helps to provide a fuller picture of the effects of antigen binding. It also helps determine the effects that these structural changes may have upon B-cell maturation and the initiation of BCR signalling, which is still, despite much research into the area, a significant gap in our understanding of this system.

Despite the preliminary nature of the experiments described, and the potential problems produced by studying the system using isolated components, significant progress has been made in understanding the potential signalling mechanisms. This work provides interesting and potentially exciting insights into the nature and mechanism of BCR activation and provides the basis for future research into this area.

Chapter 5 – Final Conclusions

Antibodies and their derived fragments are an important set of tools for the development of new therapeutics. Therapeutic antibodies form a large portion of the therapeutic market with global sales totalling many billions of US dollars (Reichert, 2010). Due to the highly specific nature of antibodies, and the relatively low toxicity of developed therapeutics, much effort is being invested into their development for new targets. New variants of the antibody molecule are also being investigated that look to further improve its potential as a therapeutic and tailor treatments to specific diseases and applications. A fuller understanding of the structural and functional properties of the antibody molecule is critical for this research. The work presented in this thesis aims to provide further information on a molecular level of the characteristics of antibody/antigen interactions from a structural perspective. Identification of the antigen binding site together with other functionally important areas within the molecule will provide a rationale for further modification of highly specific antibody derived therapeutics without affecting key properties such as affinity. NMR spectroscopy provides an important and highly sensitive tool for the study of antibody/antigen interactions and is able to reliably detect subtle structural changes which occur upon binding. Since the proteins are in solution they are able to move freely in a state relatively close to that which they would be *in vivo*, allowing a more realistic determination of features such as the dynamics of different parts of the protein and providing further insight into the properties of the system studied. The relative lack of antibody/antigen structures in the protein data bank (PDB) demonstrates the difficulty often experienced in crystallising these complexes and their perceived unsuitability for

217

study by NMR. The work presented in this thesis clearly demonstrates that antibody fragments, such as the scFvs, are viable and amenable tools for study by NMR.

Antibody fragments are important therapeutic tools in their own right with a number of Fab fragments already approved for use in the clinic (Yoon *et al*, 2010), but also represent important model systems for the study of the properties of antibody binding. The small size compared to the full antibody (~ 150 kDa), or even the Fab (~ 50 kDa), makes the scFv (~ 27 kDa) an attractive model system for study by NMR. The work reported in chapter two describes the cloning, expression and purification of isotopically labelled scFv proteins in a highly efficient small volume method allowing the generation of NMR samples from less than five percent of the D_2O required for the equivalent traditional methods for 1189scFv. This method was adapted to include the use of unlabelled amino acids in otherwise uniformly labelled $^{15}N/^{13}C/^{2}H$ samples to provide further information for the reliable determination of sequence specific backbone assignments in an efficient and cost effective manner. Chapter two also describes the equilibrium between monomeric and dimeric states of 1189scFv at a range of protein concentrations. This tendency to multimerise is a characteristic which has previously prevented the collection of triple resonance data from scFvs. In contrast to reports in the literature, 1189scFv shows a significantly reduced multimerisation dependence on concentration and, as far as can be determined from the available literature, has allowed for the first time the collection of high quality, assignable backbone triple resonance data for a scFv in the absence of its antigen. These differing characteristics may be due to subtle differences in the residue composition of the V_L/V_H interface that may be beneficial for the study of these proteins by NMR.

Chapter three describes the collection and analysis of high quality NMR spectroscopy data which allowed detailed structural comparisons of 1189scFv in both the presence and absence of its antigen, IL-6. Backbone assignments (H_N, N, Cα, Cβ and C$'$) were produced to a high level of completeness for the free and IL-6 bound scFv. Interestingly, missing assignments were primarily for regions on the CDR3 loops of both the V_L and V_H domain. This indicates, as seen previously for IC8scFv and Fab$'$ (Hall, 2009; Wilkinson, 2009), that the CDR3 loops are involved in intermediate exchange between a limited number of conformations and may indicate a stabilisation upon binding. In contrast to data gathered for the IC8 fragments and IL-1β, the signals for the unassigned CDR3 residues were still undetectable upon binding suggesting that an intermediate exchange may remain even in the bound state. This interesting observation may reflect the structural nature of IL-6 as an antigen which demonstrates a much greater plasticity within its structure than IL-1β. This is also demonstrated by the minimal shift analysis of 1189Fab and 1189scFv binding to triply labelled IL-6 in chapter 3, where chemical shift perturbations are transferred away from the binding site to a significant proportion of the protein. In addition to the large chemical shift changes observed for the CDR residues of 1189scFv, an further set of perturbations were found within the framework of the protein. A significant portion of these additional perturbations were located on the residues comprising the interface between the V_L and V_H domains. These chemical shift changes were shown to be much larger than those observed on the antigen binding site in the actual chemical shift analysis performed in chapter three, suggesting a significant change in chemical environment caused by a re-orientation of the domains. While this information does not provide the precise details on what structural changes are occurring, the patterns of chemical shift changes show many similarities to the data observed for the IC8/IL-1β system used by Hall (2009) and

Wilkinson (2009), indicating a similar conformational change. This similarity of structural change, despite the use of an antigen that is structurally very different, led to the hypothesis that this domain movement may be a conserved feature of antibody/antigen interactions, and may point to additional functionality within the molecule. These results may help to explain the gap in our understanding of the way in which antigen binding, particularly when in a monovalent form, leads to the initiation of signalling through the B-cell receptor.

Chapter four describes a series of NMR experiments designed to test essential elements of a model of B-cell signalling developed previously within the group (Hall, 2009), involving an interaction between the Fab and Fc which is regulated by antigen binding. The results observed were assessed, taking into account both the proposed model and results from other recent investigations into the initiation of BCR signalling, which have led to two potential models of their own (Tolar *et al*, 2009; Yang and Reth, 2010). The IC8Fab′/IL-1β/Fc system was used to investigate this as it contained all the domains present in the full antibody with comprehensive backbone resonance assignments previously determined for the Fab′ molecule (Hall, 2009). The experiments described in chapter 4 revealed a weak interaction between the Fab′ and Fc, represented by a reduction in the signal to noise of a specific subset of backbone amide NMR signals. Interestingly, the intensity of the affected signals was no longer affected when antigen was bound to the Fab′ indicating a weakening of the Fab′/Fc interaction. These experiments shed the first light onto a potential antigen regulated interaction that may be involved in BCR signalling and provide the first structural data specifically looking at an interaction between this combination of antibody fragments. There is also further evidence within the literature which indicates that structural changes occur within whole antibodies upon binding. Small angle x-ray scattering data observes

experimentally significant changes in the dynamics and volumes of the observed proteins upon antigen binding (Pilz *et al*, 1973; Pilz *et al*, 1977, Pilz *et al*, 1980). More recent evidence also observes changes in the manner in which antibodies interact with each other, highlighting the ability of antibodies to form dimers at high concentrations (Mosbaek *et al*, 2012). Whilst circulating antibodies will unlikely ever reach concentrations such as these, it is possible that localised high concentrations may be formed on the surface of a B-cell. Other research in to the flexibility of antibodies, particularly regarding the Fab-Fab and Fab-Fc angles, comes from a number of techniques. The crystallisation of whole antibodies is particularly difficult due to the nature of the linker regions between the Fab and Fc and subsequently only a few structures have been published. These highly flexible, unstructured regions are often missing from electron density maps and allow the Fabs to occupy multiple angles, disrupting the formation of high quality crystals. There are however a few examples of antibody crystal structures for antibodies such as human IgG1 (Harris *et al*, 1998; Saphire *et al*, 2001) and mouse IgG2A (Harris *et al*, 1997) that often contain modified linker regions to aid stability. These structures generally represent the antibody as a distorted T shape with a Fab-Fab angle approaching 180 ˚ in some cases, which is consistent with the small angle x-ray scattering data despite the modified linkers. A more recent study, using IgG1 with an intact hinge region, also confirms this observation but demonstrates the extreme flexibility and asymmetry that these molecules display (Saphire *et al*, 2002). FRET data has been gathered to deduce the distance between the ends of the Fab and the Fc to determine the flexibility and overall shape of antibodies (Zheng *et al*, 1992). It was noted in this study that the average end to end distance for the Fab and Fc was significantly less than that expected for a planar T or Y shape. Cryoelectron microscopy and tomography experiments have also been

used to attempt to determine the overall shape and flexibility of the antibody molecule. Whilst some experiments demonstrate that Fab fragments are able to bend quite far back towards the Fc (Roux *et al*, 1997), others favour the more classical Y shape (Bongini *et al*, 2004). Both the FRET and cryoelectron microscopy/tomography data both show a large variance in angle around the mean, further highlighting the flexibility of the linker region. Despite the data in this field being relatively sparse, and in some cases conflicting, there is significant evidence that suggests that the Fab is at least sufficiently flexible to interact with an Fc, even if not for the same antibody then a neighbouring one, potentially supporting the results outlined in this chapter. The results presented are the first steps into determining a mechanism for the initiation of B-cell signalling. The continued development of the models proposed hinges quite heavily on the ability to identify a binding site for the interaction that has been observed between the Fab and the Fc. This is particularly important for the determination of whether the structural changes observed upon binding are indeed required for signalling, or whether other processes may be involved, such as steric hindrance by antigen binding, that regulate the interaction. Further NMR experiments using paramagnetic relaxation enhancers may be useful for determining the exact location of a binding site by NMR. These agents are useful for the identification of binding sites for weak and transient interactions, particularly on large molecules where complex NMR data may be unavailable. Work using FRET fluorophores attached to the Fab and Fc that mirrors the NMR experiments already performed may also provide confirmation as to whether the distance between them is regulated by antigen binding. Other experiments, such as mutational or blocking studies using cell surface reporter assays, may also provide insight into the mechanics of this system.

The work presented in this thesis outlines investigations into the structural properties of antibody/antigen interactions, including the behaviour of the CDR loops, the re-orientation of the variable domains and the discovery of a potential, antigen regulated, interaction between the Fab and Fc, that may inform the current models of BCR signalling and antibody behaviour when expressed on the surface of B-cells. In addition, NMR is shown to be a very useful and highly sensitive tool for the observation of these proteins in an efficient and reliable manner, capable of identifying subtle changes within the protein structure. The results described in this work point to exciting future studies by both NMR and other techniques, that may lead to the further development of antibody molecules as therapeutics and the determination of a precise and structurally informed model for the activation of B-cell signalling.

A.1 – Isotopic Labelling Media

A.1.1 – Minimal Medium

The minimal medium for the expression of isotopically labelled proteins in the standard cell density expression protocols was made according to the following protocol per litre of media:

- 1.0g $(NH_4)_2SO_4$ (^{15}N if required)

- Add 100ml of 10x stock of PO_4/NaCl
    - 17g Na_2HPO_4
    - 7.5g KH_2PO_4
    - 1.25g NaCl
    - Make up to 250ml

- Add 1ml of 1000x Na_2SO_4 (0.3M)
    - 4.26g in 100ml water

- Add 10ml 110x stock solution of EDTA trace elements
    - 1g EDTA
    - Dissolve in 160ml water and adjust to pH7 with NaOH
    - 0.32g $MnCl_2$
    - 0.1g $FeCl_3$
    - 0.01g $ZnCl_2$

- 0.002g CoCl$_2$

- 0.002g CuCl$_2$

- 0.002g H$_3$BO$_3$

- Adjust to pH7 and make up to 200ml

- Add 864ml water, autoclave and allow to cool.

- Add 1ml of 1000x filter sterilised MgSO$_4$ (1M)

  - 24.6g MgSO$_4$.7H$_2$O in 100ml water

- Add 1ml of 1000x filter sterilised CaCl$_2$ (0.3M)

  - 4.4g CaCl$_2$.2H$_2$O in 100ml water

- Add 1ml of 1000x filter sterilised d-Biotin

  - 10mg of d-Biotin in 10ml 50% ethanol

- Add 1ml of 1000x filter sterilised Thiamine (Light sensitive)

  - 10mg of Thiamine in 10ml water

- Add 20ml of 50x (4 g l^{-1}) autoclaved glucose (^{13}C$_6$ D-glucose if required)

  - 20g of glucose in 100ml water

  - Autoclave

- Add appropriate antibiotic

For the expression of deuterated protein using this method replace all (including any used to make the individual reagents) water with deuterium oxide and filter using a 0.22 μm filter rather than autoclaving after the addition of EDTA trace elements.

A.1.2 – High Density Labelling Medium

The minimal medium for the expression of isotopically labelled proteins in the high cell density expression protocols was made according to the following protocol per 100 ml of media:

- 50mM Na_2HPO_4 (0.71g)
- 25mM KH_2PO_4 (0.34g)

- Add 1ml 100X NaCl (1M)
    - 5.84g in 100ml

- Add 1ml 100X $MgSO_4$ (0.5M)
    - 12.3g in 100ml

- Add 1 ml 100X $CaCl_2$ (0.02M)
    - 0.29g in 100ml

- Make up to 90ml with water (or D_2O) and adjust to pH 8.0-8.2

- Add 100ul Metals
    - 12.5mM $FeCl_3.7H_2O$ (0.338g)
    - 5mM $CaCl_2.2H_2O$ (0.073g)
    - 2.5mM $MnCl_2.4H_2O$ (0.050g)
    - 2.5mM $ZnSO_4.7H_2O$ (0.072g)
    - 0.5mM $CoCl_2.6H_2O$ (0.012g)
    - 0.5mM $CuCl_2.2H_2O$ (0.0085g)
    - 0.5mM $NiCl_2.6H_2O$ (0.012g)
    - 0.5mM Na_2MoO_4 (0.012g)

- 0.5mM Na_2SeO_3 (0.0076g)

- 0.5mM H_3BO_3 (0.0030g)

- Make up to 100ml

- Add 0.25ml 100x BME vitamins (Sigma)

- Add 0.1g NH_4Cl (or $^{15}NH_4Cl$ if required)

- Add 1g glucose (or $^{13}C_6$ glucose if required)

- Add appropriate antibiotic

- Top up to 100ml with water (or D_2O)

This expression medium cannot be autoclaved and will form a white precipitate prohibiting filtration. This appears to be normal and the precipitate will re-dissolve during expression. If deuteration is required, all stocks must be made in deuterium oxide in place of water.

A.2 – Chemical Shift Index Data

A.2.1 – Sequence Specific Assignments for 1189scFv

Residue	C′	Cα	Cβ	H	N
D1	172.4	53.1	38.87	-	-
I2	175.8	62.52	36.81	9.222	123.4
Q3	175.2	54.96	29.46	8.849	130.4
M4	176.5	51.54	29.58	8.996	124.9
T5	174.2	61.81	70.25	9.635	120
Q6	175.6	54.19	30.69	9.78	128.1
S7	-	55.41	64.81	8.895	117.7
S9	175.5	60.96	62.95	-	-
S10	172	57.02	64.85	7.799	115.3
L11	174.2	54.85	45.04	8.924	124.3
S12	173.6	56.13	63.89	8.515	119
A13	174.9	50.77	22.66	8.813	126.9
S14	173.9	57.1	64.39	8.746	115.9
V15	177.8	64.85	30.75	8.498	122.3
G16	174.8	44.28	-	10	117.5
D17	174.4	54.58	41.12	7.996	121.8
R18	176.7	54.19	30.87	8.163	119.7
V19	173.7	59.36	35.07	8.431	121.4
T20	174	60.89	71.26	8.052	116.8
I21	174.6	60.06	40.37	9.339	128.8
T22	174.3	60.9	70.61	8.991	119.4
C23	171.1	55.69	43.86	9.103	127.5
L24	176.3	52.75	43.51	9.161	127.6
A25	178.9	49.96	22.67	9.295	127.2
S26	173.5	60.41	62.56	8.326	114

Residue	C′	Cα	Cβ	H	N
E27	173.2	53.65	32.37	7.536	117
G28	174.8	46.47	-	8.189	103.9
I29	175.4	59.53	38.32	8.436	117.5
S30	173.7	58.3	61.1	8.303	115.9
N31	175.2	52.62	36.87	8.351	120.7
D32	173.6	53.14	37.78	8.186	120.9
L33	174.7	52.21	44.5	7.794	125.6
A34	176	49.08	22.8	8.811	127.5
W35	175.9	55.48	33.32	8.612	115.5
Y36	174.7	56.09	42.45	9.82	117.6
Q37	174.2	54.1	31.95	9.226	121.6
Q38	175	54.85	37.22	9.083	127
K39	-	58.66	31.67	8.328	128.2
P40	178.4	63.77	30.65	-	-
G41	174	45.5	-	8.804	113
K42	175.6	53.01	34.34	8.011	119.6
A43	-	50.05	18.39	8.18	121.5
P44	-	63.06	33.08	-	-
K45	-	59.89	33.01	8.425	113.5
L47	175.4	54.74	44.82	-	-
I48	173.4	57.56	43.36	7.064	117
Y49	-	54.44	39.82	9.031	122.5
D50	178.8	55.4	38.69	-	-
A51	179.8	57.42	19.29	9.266	114
T52	175.6	60.59	71.45	8.839	107.7
R53	174	55.41	28.09	8.556	125
L54	178.5	55	42.33	8.217	124.8
Q55	175.6	55.06	28.25	8.199	125.8
D56	177.3	56.25	40.4	8.26	124.9
G57	174.6	44.98	-	8.7	112.6
V58	-	60.31	32	7.877	125.3
P59	176.6	61.38	35.46	-	-

Residue	C′	Cα	Cβ	H	N
S60	174.4	59.88	62.94	8.543	117.4
R61	176	56.73	28.28	7.063	116.3
F62	174.2	58.05	39.55	7.739	120.5
S63	173	56.95	65.14	8.686	112.1
G64	172.7	43.55	-	9.043	112.3
S65	174.2	57.13	65.79	8.928	114.4
G66	171.8	44.84	-	8.419	111.8
S67	174.4	58.02	64.36	6.95	109.1
G68	172.4	46.99	-	9.08	112.6
T69	173.6	61.34	70.72	8.195	112.8
D70	173.8	53.52	41.56	7.004	120.8
F71	175.9	53.6	33.37	8.625	123.1
T72	172.6	60.64	71.43	8.531	113.9
L73	174.7	52.92	41.41	8.754	126.9
T74	173.4	59.8	69.3	8.908	118.4
I75	176.2	59.69	37.9	8.617	126
S76	-	59.61	62.75	8.647	120.8
S77	174.1	56.85	62.12	-	-
L78	177.3	56.28	42.37	8.648	123.8
Q79	-	52.48	28.72	9.107	123.6
P80	178.5	66.25	31.16	-	-
E81	176.5	57.23	27.98	9.414	113.7
D82	177.6	54.54	40.77	8.332	119.8
F83	174.2	60.06	37.21	7.225	121.2
A84	174.9	51.29	20.12	8.024	127.6
T85	173.2	61.98	68.68	8.075	114.8
Y86	176.8	56.57	41.34	8.855	124.4
Y87	175.4	57.4	43.7	9.438	120.3
C88	173.6	52.78	44.55	7.712	115.7
Q89	174.5	54.44	32.76	8.795	122
Q90	175.7	53.82	29.81	8.778	128
S91	-	-	-	8.886	117.6

Residue	C′	Cα	Cβ	H	N
Y92	175.4	54.06	43.04	-	-
K93	174.5	54.09	32.68	9.068	125
Y94	-	60.35	37.76	7.965	122
W96	175.9	-	-	-	-
T97	172.5	59.67	72.01	7.238	111.2
F98	178.5	55.88	43.15	8.837	118.4
G99	173.8	44.69	-	9.057	107.4
Q100	176.4	57.34	27.15	8.923	116.7
G101	172.4	44.71	-	7.027	106.9
T102	174	61.43	72.09	8.111	118
K103	173.7	56.22	32.68	8.395	129.4
L104	175.3	53.94	43.15	8.747	129.6
E105	174.5	54.19	32.18	8.999	125.3
I106	176.5	59.72	37.84	7.805	122
K107	176.3	56.31	32.15	8.349	128.2
R108	176.4	55.78	29.94	8.482	124
T109	175.2	61.29	69.48	8.179	114.8
G110	175.2	45.01	-	8.447	110.9
G111	-	45.03	-	8.445	109.1
S119	175.5	58.28	63.63	-	-
G120	174.6	45.06	-	8.519	110.7
G121	-	-	-	8.357	115.7
S124	173	56.88	65.28	-	-
G125	172.7	43.45	-	9.039	112.3
G126	-	-	-	8.926	114.3
G128	174.4	45.07	-	-	-
S129	174.7	57.92	63.84	8.253	115.3
E130	176	56.72	29.46	8.703	123.2
V131	175.7	62.72	31.3	7.728	120.6
Q132	173.9	54.91	31.61	8.39	126
L133	175.6	53.57	42.8	8.348	124.8
V134	-	61.86	34.1	8.55	120.8

Residue	C′	Cα	Cβ	H	N
E135	176.8	56.76	30.4	-	-
S136	173.6	57.72	65.4	9.404	114.5
G137	174.4	44.62	-	8.572	107.5
G138	173.1	45.14	-	7.619	105.7
G139	170.8	44.63	-	7.411	107.2
L140	176.8	54.07	42.4	8.097	122.8
V141	173.2	59.37	34.79	9.025	123.9
Q142	-	53.51	27.33	8.202	122.3
P143	177.8	63.49	30.3	-	-
G144	-	44.68	-	9.78	114.1
G145	170.6	44.29	-	8.166	-
S146	173.1	56.25	66.84	7.853	109.8
L147	173.6	54.35	46.71	8.629	122
R148	175.6	54.18	31.68	8.076	122.4
L149	177.1	52.83	42.75	8.864	126.5
S150	173.1	56.72	66.47	8.547	113.8
C151	171.6	53.61	42.19	9.048	122.9
A152	176.4	50.67	19.62	8.665	130.5
A153	176.4	50.21	23.21	8.315	126.8
S154	174.4	57.54	65.41	8.926	114.2
G155	173.4	45.36	-	8.643	107.5
F156	173.5	54.45	39.41	7.327	112.3
T157	174.6	62.04	67.39	9.268	118.4
F158	176.1	61.02	39.7	8.441	133.3
N159	173.5	53.91	36.93	8.352	110.5
D160	176.2	54.3	41.9	7.576	115.6
Y161	176.1	56.87	35.52	7.238	114
D162	-	-	-	8.312	118.7
M163	174.4	51.43	35.77	-	-
A164	175.8	53.1	23.32	9.27	121.5
W165	175.1	57.44	34.27	8.899	114
V166	173.1	60.58	35.08	9.363	122.9

Residue	C′	Cα	Cβ	H	N
R167	173.9	52.75	35.16	9.303	122.5
Q168	175.4	55.4	37.93	8.588	120.8
A169	-	52.54	19.25	8.865	130
P170	178.4	63.93	30.78	-	-
G171	174.3	45.26	-	8.783	112.7
K172	176.4	54.17	33.37	8.423	119.5
G173	173.4	44.34	-	8.036	106.1
L174	176.7	54.01	41.12	6.357	118.1
E175	175.4	55.19	32.94	9.13	121.2
W176	-	60.12	28.03	10.14	130.4
S179	174	58.45	61.83	-	-
I180	174.7	57.25	41.64	8.637	119.4
T181	-	60.5	69.62	8.647	115.5
P182	176.6	66.56	-	-	-
S183	176.7	58.33	64.08	7.996	105.9
G184	173.9	44.23	-	8.13	110.5
G185	173.3	45.61	-	8.045	106.4
G186	171.7	43.84	-	7.677	108.7
T187	172.5	58.31	71.53	7.769	112.8
Y188	174.1	55.86	40.57	8.299	121.3
Y189	176.4	56.98	42.86	8.868	117.2
R190	177.7	55.75	28.53	8.005	123
D191	179.1	58.3	39.9	9.472	126.9
S192	175.6	59.92	62.39	8.259	111.5
V193	175.4	59.69	30.32	7.503	112.6
K194	177.8	58.06	31.61	7.348	126
G195	174	45.1	-	9.166	115.8
R196	175.9	56.97	29.77	7.656	117.3
F197	175	52.27	38.65	7.879	119.7
T198	174.7	60.97	71.42	8.987	113.1
I199	171.6	58.52	40.81	9.49	132.7
S200	171.8	57.91	65.31	8.486	118.4

Residue	C′	Cα	Cβ	H	N
R201	173	54.14	34.99	9.411	116
D202	177.4	52.12	42.02	9.097	120.2
N203	177.4	55.32	36.86	9.528	124.4
A204	179	53.84	17.9	8.51	122.4
K205	176.2	54.54	32.52	7.27	114.9
N206	172.3	53.58	36.51	7.801	119
S207	170.9	57.12	67.49	7.115	107.8
L208	173.8	53.35	45.5	8.887	125.9
Y209	176.7	56.89	41.44	8.41	118.4
L210	174.4	53.95	41.45	8.662	123
Q211	174.4	55.75	28.78	8.728	125.1
M212	173	56.34	33.9	8.522	124.2
N213	174.5	51.03	40.51	7.887	119.9
S214	176.2	57.2	61.48	8.559	111.1
L215	177.1	56.53	41.92	8.13	118.9
R216	176.9	53.87	33.09	9.574	120.7
A217	175.7	54.34	17.3	9.254	129.5
E218	176	58.02	27.73	9.12	114.3
D219	177	54.76	41.97	8.493	118.3
T220	174.6	64.1	69.31	7.827	120.2
A221	175.5	52.43	20.35	9.309	128.2
V222	175.4	60.28	33.53	8.053	118.4
Y223	-	58.11	39.77	9.199	126.9
C225	172.5	52.96	42.62	-	-
A226	-	50.26	20	8.604	124.1
E236	176	58.55	31.34	-	-
Y237	173.3	57.59	42.98	7.175	111.4
W238	177.7	56.27	32.29	8.752	121.7
G239	-	45.51	-	8.856	109.8
T242	172.2	59.65	70.32	7.859	109.8
M243	175.3	55.67	32.44	8.6	128.4
V244	174.2	61.84	32.96	8.936	129.9

Residue	C′	Cα	Cβ	H	N
T245	172.1	61.62	69.46	8.401	123.4
V246	176	59.74	30.36	8.445	128.3
S247	173.9	57	64.52	8.493	119.7
S248	174.3	58.1	63.81	8.64	119.4
H249	176	55.71	29.82	8.45	122.5
H250	-	63.28	31.13	7.93	122.6
H251	176.5	60.21	-	-	-
H252	176.1	56.1	32.22	8.223	128.4
H253	176.4	55.5	29.9	8.591	125.9
H254	-	59.09	49.36	8.209	119.8

A.2.2 – Sequence Specific Assignments for 1189scFv in complex with IL-6

Residue	C′	Cα	Cβ	H	N
D1	172.4	53.13	38.95	-	-
I2	175.8	62.65	36.84	9.117	123.4
Q3	175.2	54.98	29.44	8.856	131.3
M4	176.5	51.53	29.71	8.965	125.8
T5	174.3	61.84	70.21	9.592	120.6
Q6	175.7	54.21	30.63	9.741	128.8
S7	-	55.44	64.71	8.841	118.2
P8	176.1	62.88	-	-	-
S9	175.5	60.96	62.88	9.105	116.2
S10	172	57.07	64.64	7.742	115.7
L11	174.2	54.93	45.11	8.861	124.7
S12	173.6	56.17	63.73	8.456	119.5
A13	174.9	50.81	22.76	8.752	127.5
S14	173.9	57.18	64.19	8.688	116.3

Residue	C′	Cα	Cβ	H	N
V15	177.8	64.87	30.79	8.438	122.8
G16	174.8	44.34	-	9.936	117.9
D17	174.3	54.68	41.2	7.936	122.3
R18	176.7	54.23	30.93	8.107	120.2
V19	173.7	59.4	35.14	8.378	121.9
T20	173.2	60.89	71.11	7.99	117.2
I21	174.6	60.51	40.24	9.153	126.1
T22	174.4	62.01	70.95	8.953	119.9
C23	171.1	55.82	47.77	9.576	125.4
L24	176.3	52.79	43.5	9.09	128
A25	178.9	50.01	22.76	9.237	127.5
S26	173.4	60.37	62.45	8.202	114.4
E27	173.2	53.67	32.49	7.426	117.7
G28	174.8	46.61	-	8.108	104.4
I29	175.1	59.49	38.46	8.405	118.4
S30	173.7	58.84	60.77	8.267	116.2
N31	174.7	51.61	38.03	8.244	121.5
D32	174.1	54.54	39.08	8.471	123.6
L33	174.5	52.43	44.21	7.835	126.8
A34	175.8	49.04	-	8.753	128
W35	175.9	55.62	33.68	8.618	116.4
Y36	174.7	56.27	42.79	9.745	117.8
Q37	174.2	53.43	34.27	9.162	121.9
Q38	175	54.7	31.05	9.579	128.9
K39	-	54.38	30.59	9.11	133.3
P40	178.4	63.82	30.68	-	-
G41	174	45.53	-	8.764	113.6
K42	175.5	53.08	34.39	7.953	120.2
A43	-	50.07	18.43	8.084	121.7
P44	172.8	61.91	-	-	-
K45	176.1	54.02	35.43	8.322	118.1
L46	-	55.98	41.25	8.965	128.3

Residue	C′	Cα	Cβ	H	N
L47	175.4	54.82	44.9	-	-
I48	173.4	57.63	43.61	6.958	117.5
Y49	-	54.45	40.34	8.917	123.1
D50	179.3	55.49	40.94	-	-
A51	179.8	57.51	19.13	9.164	114.8
T52	175.6	60.59	71.47	8.621	107.1
R53	175.5	55.45	28.13	8.444	125.6
L54	178.4	55.07	42.43	8.125	125.5
Q55	-	55.05	28.14	8.165	126.9
D56	177.3	56.3	40.33	-	-
G57	174.5	45.05	-	8.651	113
V58	-	60.38	32.11	7.837	125.9
P59	177	63.08	33.07	-	-
S60	174.4	59.93	62.6	8.362	113.9
R61	176	56.82	28.38	6.998	116.8
F62	174.1	58.05	39.56	7.654	121
S63	172.9	56.99	65.08	8.629	112.6
G64	172.6	43.57	-	8.978	112.7
S65	174.1	57.14	65.63	8.881	114.9
G66	171.7	44.85	-	8.354	112.3
S67	174.6	58.1	64.37	6.876	109.4
G68	172.3	47.11	-	9.103	113.4
T69	173.5	61.4	70.83	8.076	113
D70	173.8	53.54	41.62	6.911	121.4
F71	175.9	56.84	41.97	8.69	122.6
T72	172.6	60.34	72.39	8.894	113.9
L73	174.7	53.07	40.63	8.715	127.5
T74	173.4	60.85	70.13	8.935	123.6
I75	176.1	59.78	37.84	8.564	126.4
S76	-	59.66	62.8	8.588	121.2
S77	174.1	56.9	62.02	-	-
L78	177.3	56.28	42.5	8.588	124.2

Residue	C′	Cα	Cβ	H	N
Q79	-	52.46	28.82	9.055	124.1
P80	178.4	66.21	31.21	-	-
E81	176.5	57.3	28.03	9.349	114.1
D82	177.6	54.68	40.75	8.272	120.3
F83	174	60.08	37.15	7.163	121.7
A84	174.9	51.38	20.2	8.054	127.6
T85	173.2	61.92	68.66	7.993	115.2
Y86	176.9	56.61	41.41	8.82	124.9
Y87	175.4	57.52	43.84	9.324	120.6
C88	173.3	52.79	44.44	7.63	116
Q89	174.5	54.12	33.47	8.923	122.3
Q90	-	52.94	27.72	8.395	124.6
T97	172.3	59.8	71.88	-	-
F98	178.4	55.84	43.13	8.751	118.7
G99	173.8	44.74	-	9.002	107.9
Q100	176.4	57.43	27.22	8.897	117.4
G101	172.4	44.72	-	6.953	107.4
T102	174	61.43	72.26	8.062	118.6
K103	173.7	56.36	32.6	8.342	130
L104	175.4	54.03	43.2	8.704	130.1
E105	174.5	54.27	32.24	8.94	125.9
I106	176.5	59.77	37.94	7.756	122.5
K107	176.3	56.34	32.21	8.305	128.7
R108	176.4	55.95	29.99	8.428	124.4
T109	175.2	61.27	69.45	8.121	115.2
G110	175	45.11	-	8.373	111.4
G111	-	-	-	8.387	109.5
S124	175.3	58.26	63.55	8.304	116
G125	174.5	45.08	-	8.49	111.2
G126	-	-	-	8.303	116
G127	174.4	44.65	-	-	-
G128	174.3	45.13	-	7.546	106.2

Residue	C′	Cα	Cβ	H	N
S129	174.6	57.94	63.72	8.2	115.8
E130	176	56.82	29.55	8.67	123.7
V131	175.6	62.81	31.45	7.657	121
Q132	173.9	54.97	31.69	8.347	126.4
L133	175.5	53.58	42.85	8.258	125.3
V134	176.5	61.94	34.1	8.469	121.3
E135	176.8	56.79	30.49	10.52	132.6
S136	173.6	57.72	65.25	9.353	115
G137	-	44.66	-	8.511	108
G138	173.1	45.24	-	-	-
G139	170.8	44.68	-	7.343	107.7
L140	176.7	54.11	42.43	8.039	123.2
V141	173.2	59.43	34.86	8.968	124.4
Q142	-	53.56	27.43	8.145	122.8
P143	177.7	63.49	30.43	-	-
G144	-	44.76	-	9.722	114.6
G145	170.6	44.31	-	-	-
S146	173.1	56.29	66.66	7.795	110.2
L147	173.6	54.41	46.77	8.575	122.5
R148	175.6	54.27	31.79	8.015	122.9
L149	177	52.87	42.69	8.799	127
S150	173.1	56.71	66.32	8.773	114.2
C151	171.6	53.57	41.93	9.166	123.5
A152	176.4	50.75	19.54	8.598	131
A153	176.4	50.27	23.29	8.208	127.4
S154	174.3	57.62	65.28	8.891	114.8
G155	173.4	45.42	-	8.622	108
F156	173.5	54.55	39.59	7.222	112.1
T157	174.8	62.31	67.79	9.124	118.9
F158	176.2	61.08	39.7	8.433	134.1
N159	173.8	53.31	36.23	8.317	111.6
D160	176.8	54.91	43.24	7.567	115.3

239

Residue	C′	Cα	Cβ	H	N
Y161	174.7	58.21	41.5	7.322	114.8
D162	174.1	54.14	-	8.508	123.1
M163	174.3	51.36	35.59	7.935	119
A164	175.8	50.46	25.28	9.263	122
W165	175.5	55.54	32.44	8.746	114.3
V166	173	60.6	-	9.463	125.2
R167	173.9	52.49	32.78	9.684	124.8
Q168	175.4	54.23	32.25	8.955	121.3
A169	-	50.42	17.02	9.447	136.7
P170	178.3	64.01	30.85	-	-
G171	174.7	45.58	-	8.578	113.3
K172	176.4	54.2	33.33	8.224	121
G173	173.4	44.27	-	8.079	106.6
L174	176.8	53.97	41.24	6.442	118.7
E175	175.4	55.21	32.97	9.082	121.9
W176	174.3	60.11	-	10.13	131.1
V177	172.9	63.16	-	7.953	125.8
A178	176	51.48	22.28	6.91	115.7
S179	171.8	59.65	-	9.233	115.7
I180	173	57.79	41.2	9.209	121.3
T181	-	61.11	66.3	8.333	118
P182	176.7	66.87	-	-	-
S183	177	59.62	63.75	9.042	108.1
G184	173.7	43.83	-	8.091	112.1
G185	173.1	46.16	-	8.158	106.8
G186	171.4	42.96	-	7.679	106.7
T187	172.6	58.45	71.69	6.9	112
Y188	174.2	55.71	40.75	8.299	120.4
Y189	177	57.3	42.75	9.532	118.4
R190	177.5	55.9	28.54	7.915	123.4
D191	179.1	58.48	40.03	9.384	127.6
S192	175.6	59.94	62.22	8.236	111.9

Residue	C′	Cα	Cβ	H	N
V193	175.3	59.73	30.35	7.385	113
K194	177.7	57.98	31.63	7.269	126.4
G195	174	45.21	-	9.125	116.3
R196	175.8	56.99	29.88	7.598	117.8
F197	175	52.23	38.66	7.85	120.2
T198	174.7	60.94	71.41	8.9	113.5
I199	171.5	58.6	-	9.407	133.1
S200	171.7	57.92	65.12	8.413	118.9
R201	172.9	54.12	35.06	9.381	116.5
D202	177.4	52.15	42.04	9.055	120.6
N203	177.3	55.28	37.18	9.452	124.6
A204	178.9	53.86	18.02	8.51	122.9
K205	176.1	54.68	32.59	7.237	115.5
N206	172.3	53.62	36.42	7.714	119.2
S207	170.9	57.11	67.4	7.069	108.4
L208	173.8	52.75	46.83	8.835	126.3
Y209	176.7	56.41	41.63	9.226	122.3
L210	174.3	53.76	42.05	8.591	123.3
Q211	174.4	54.55	28.5	8.953	129.1
M212	173	54.54	34.52	8.85	127.7
N213	173.5	51.01	40.5	7.807	120.3
S214	176.2	57.38	61.44	8.504	111.5
L215	177.1	56.55	41.89	8.077	119.4
R216	176.9	53.61	33.48	9.522	121.2
A217	180.7	55.79	17.51	9.035	125.6
E218	176	58.01	27.75	9.049	114.8
D219	177	54.84	41.93	8.439	118.8
T220	174.7	64.19	69.43	7.763	120.6
A221	175.5	52.68	21.09	9.255	128.8
V222	175.4	62.61	31.89	7.831	119.5
Y223	-	58.23	-	9.146	127.4
C225	172.4	53.02	42.27	-	-

Residue	C′	Cα	Cβ	H	N
A226	174.4	50.17	19.95	8.584	124.9
R227	-	53.91	-	8.37	122.1
T231	173.3	60.49	69.35	-	-
L232	-	53.09	44.85	6.534	120.1
Y234	174.3	58.05	-	-	-
F235	174.4	55.83	37.56	8.414	123.2
E236	176	58.69	-	8.055	126.2
Y237	173.2	57.55	43.12	7.06	111.4
W238	177.8	56.25	32.19	8.65	122.2
G239	173.1	45.56	-	8.831	110.5
Q240	176.3	56.42	-	8.904	116
G241	-	44.94	-	6.932	104.1
T242	172.2	59.72	70.31	-	-
M243	175.3	55.73	32.31	8.525	130.5
V244	174.1	61.91	32.96	8.981	130.5
T245	172.1	61.66	69.51	8.337	123.9
V246	176	59.8	30.33	8.385	128.7
S247	173.8	57.04	64.34	8.435	120.2
S248	-	58.25	63.85	8.582	119.8
H250	176.5	56.48	-	-	-
H251	175.6	61.87	-	8.017	121.2
H252	173.4	59.71	-	7.179	112.4
H253	-	56.85	-	6.728	113.8
H254	-	-	-	-	-

A.3 – Protein Sequences

A.3.1 – 1189scFv

(M K K T A I A I A V A L A G F A T V A Q A) D I Q M T Q S P S S L S A S V G D
R V T I T C L A S E G I S N D L A W Y Q Q K P G K A P K L L I Y D A T R L Q D
G V P S R F S G S G S G T D F T L T I S S L Q P E D F A T Y Y C Q Q S Y K Y P
W T F G Q G T K L E I K R T G G G G S G G G G S G G G G S G G G G S E V Q L
V E S G G G L V Q P G G S L R L S C A A S G F T F N D Y D M A W V R Q A P G
K G L E W V A S I T P S G G G T Y Y R D S V K G R F T I S R D N A K N S L Y L
Q M N S L R A E D T A V Y Y C A R H G Y T L D Y F E Y W G Q G T M V T V S
S H H H H H H

A.3.2 – 1189Fab

>V$_L$C$_κ$

(M K K T A I A I A V A L A G F A T V A Q A) D I Q M T Q S P S S L S A S V G D
R V T I T C L A S E G I S N D L A W Y Q Q K P G K A P K L L I Y D A T R L Q D
G V P S R F S G S G S G T D F T L T I S S L Q P E D F A T Y Y C Q Q S Y K Y P
W T F G Q G T K L E I K R T V A A P S V F I F P P S D E Q L K S G T A S V V C L
L N N F Y P R E A K V Q W K V D N A L Q S G N S Q E S V T E Q D S K D S T Y
S L S S T L T L S K A D Y E K H K V Y A C E V T H Q G L S S P V T K S F N R G
E C

>$V_H C_{H1}$

(M K K T A I A I A V A L A G F A T V A Q A) E V Q L V E S G G G L V Q P G G S
L R L S C A A S G F T F N D Y D M A W V R Q A P G K G L E W V A S I T P S G
G G T Y Y R D S V K G R F T I S R D N A K N S L Y L Q M N S L R A E D T A V
Y Y C A R H G Y T L D Y F E Y W G Q G T M V T V S S A S T K G P S V F P L A
P S S K S T S G G T A A L G C L V K D Y F P E P V T V S W N S G A L T S G V H
T F P A V L Q S S G L Y S L S S V V T V P S S S L G T Q T Y I C N V N H K P S N
T K V D K K V E P K S C

A.3.3 – IC8Fab′

>$V_L C_\kappa$

(M K K T A I A I A V A L A G F A T V A Q A) D I Q M T Q S P S S L S A S V G D
R V T I T C R T S G N I H N Y L T W Y Q Q K P G K A P Q L L I Y N A K T L A D
G V P S R F S G S G S G T Q F T L T I S S L Q P E D F A N Y Y C Q H F W S L P F
T F G Q G T K V E I K R T V A A P S V F I F P P S D E Q L K S G T A S V V C L L
N N F Y P R E A K V Q W K V D N A L Q S G N S Q E S V T E Q D S K D S T Y S
L S S T L T L S K A D Y E K H K V Y A C E V T H Q G L S S P V T K S F N R G E
C

>V_HC_{H1}

(M K K T A I A I A V A L A G F A T V A Q A) E V Q L V E S G G G L V Q P G G S

L R L S C A A S G F D F S R Y D M S W V R Q A P G K R L E W V A Y I S S G G

G S T Y F P D T V K G R F T I S R D N A K N T L Y L Q M N S L R A E D T A V Y

Y C A R Q N K K L T W F D Y W G Q G T L V T V S S A S T K G P S V F P L A P

S S K S T S G G T A A L G C L V K D Y F P E P V T V S W N S G A L T S G V H T

F P A V L Q S S G L Y S L S S V V T V P S S S L G T Q T Y I C N V N H K P S N T

K V D K K V E P K S C D K T H T C A A

A.3.4 – IgG1 Fc

K T H T C P P C P P E L L G G P S V F L F P P K P K D T L M I S R T P E V T C V

V V D V S H E D P E V K F N W Y V D G V E V H N A K T K P R E E Q Y N S T Y

R V V S V L T V L H Q D W L N G K E Y K C K V S N K A L P A P I E K T I S K A

K Q P R E P Q V Y T L P P S R D E L T K N Q V S L T C L V K G F Y P S D I A V

E W E S N G Q P E N N Y K T T P P V L D S D G S F F L Y S K L T V D K S R W

Q Q G N V F S C S V M H E A L H N H Y T Q K S L S L S P G K

(Sequence for single chain, Fc exists as a dimer)

A.3.5 – IgE Fc

V A S R D F T P P T V K I L Q S S C D G G G H F P P T I Q L L C L V S G Y T P G

T I Q I T W L E D G Q V M D V D L S T A S T T Q E G E L A S T Q S E L T L S Q

K H W L S D R T Y T C Q V T Y Q G H T F E D S T K K C A D S N P R G V S A Y

L S R P S P F D L F I R K S P T I T C L V V D L A P S K G T V Q L T W S R A S G

K P V N H S T R K E E K Q R N G T L T V T S T L P V G T R D W I E G E T Y Q C

R V T H P H L P R A L M R S T T K T S G P R A A P E V Y A F A T P E W P G S R

D K R T L A C L I Q N F M P E D I S V Q W L H N E V Q L P D A R H S T T Q P R

K T K G S G F F V F S R L E V T R A E W E Q K D E F I C R A V H E A A S P S Q

T V Q R A V S V N P G K

(Sequence for single chain, Fc exists as a dimer)

A.4.1 – 1189

scFv	Res	Kabat	Fab	scFv	Res	Kabat	Fab	scFv	Res	Kabat	Fab
1	D	L1	1	37	Q	L37	37	74	T	L74	74
2	I	L2	2	38	Q	L38	38	75	I	L75	75
3	Q	L3	3	39	K	L39	39	76	S	L76	76
4	M	L4	4	40	P	L40	40	77	S	L77	77
5	T	L5	5	41	G	L41	41	78	L	L78	78
6	Q	L6	6	42	K	L42	42	79	Q	L79	79
7	S	L7	7	43	A	L43	43	80	P	L80	80
8	P	L8	8	44	P	L44	44	81	E	L81	81
9	S	L9	9	45	K	L45	45	82	D	L82	82
10	S	L10	10	46	L	L46	46	83	F	L83	83
11	L	L11	11	47	L	L47	47	84	A	L84	84
12	S	L12	12	48	I	L48	48	85	T	L85	85
13	A	L13	13	49	Y	L49	49	86	Y	L86	86
14	S	L14	14	50	D	L50	50	87	Y	L87	87
15	V	L15	15	51	A	L51	51	88	C	L88	88
16	G	L16	16	52	T	L52	52	89	Q	L89	89
17	D	L17	17	53	R	L53	53	90	Q	L90	90
18	R	L18	18	54	L	L54	54	91	S	L91	91
19	V	L19	19	55	Q	L55	55	92	Y	L92	92
20	T	L20	20	56	D	L56	56	93	K	L93	93
21	I	L21	21	57	G	L57	57	94	Y	L94	94
22	T	L22	22	58	V	L58	58	95	P	L95	95
23	C	L23	23	59	P	L59	59	96	W	L96	96
24	L	L24	24	60	S	L60	60	97	T	L97	97
25	A	L25	25	61	R	L61	61	98	F	L98	98
26	S	L26	26	62	F	L62	62	99	G	L99	99
27	E	L27	27	63	S	L63	63	100	Q	L100	100
28	G	L28	28	64	G	L64	64	101	G	L101	101
29	I	L29	29	65	S	L65	65	102	T	L102	102
30	S	L30	30	66	G	L66	66	103	K	L103	103
31	N	L31	31	67	S	L67	67	104	L	L104	104
32	D	L32	32	68	G	L68	68	105	E	L105	105
33	L	L33	33	69	T	L69	69	106	I	L106	106
34	A	L34	34	70	D	L70	70	107	K	L107	107
35	W	L35	35	71	F	L71	71	108	R	L108	108
36	Y	L36	36	72	T	L72	72	109	T	L109	109
				73	L	L73	73				

scFv	Res	Kabat	Fab
130	E	H1	215
131	V	H2	216
132	Q	H3	217
133	L	H4	218
134	V	H5	219
135	E	H6	220
136	S	H7	221
137	G	H8	222
138	G	H9	223
139	G	H10	224
140	L	H11	225
141	V	H12	226
142	Q	H13	227
143	P	H14	228
144	G	H15	229
145	G	H16	230
146	S	H17	231
147	L	H18	232
148	R	H19	233
149	L	H20	234
150	S	H21	235
151	C	H22	236
152	A	H23	237
153	A	H24	238
154	S	H25	239
155	G	H26	240
156	F	H27	241
157	T	H28	242
158	F	H29	243
159	N	H30	244
160	D	H31	245
161	Y	H32	246
162	D	H33	247
163	M	H34	248
164	A	H35	249
165	W	H36	250
166	V	H37	251
167	R	H38	252
168	Q	H39	253
169	A	H40	254
170	P	H41	255
171	G	H42	256
172	K	H43	257
173	G	H44	258
174	L	H45	259
175	E	H46	260
176	W	H47	261
177	V	H48	262
178	A	H49	263
179	S	H50	264
180	I	H51	265
181	T	H52	266
182	P	H52A	267
183	S	H53	268
184	G	H54	269
185	G	H55	270
186	G	H56	271
187	T	H57	272
188	Y	H58	273
189	Y	H59	274
190	R	H60	275
191	D	H61	276
192	S	H62	277
193	V	H63	278
194	K	H64	279
195	G	H65	280
196	R	H66	281
197	F	H67	282
198	T	H68	283
199	I	H69	284
200	S	H70	285
201	R	H71	286
202	D	H72	287
203	N	H73	288
204	A	H74	289
205	K	H75	290
206	N	H76	291
207	S	H77	292
208	L	H78	293
209	Y	H79	294
210	L	H80	295
211	Q	H81	296
212	M	H82	297
213	N	H82A	298
214	S	H82B	299
215	L	H82C	300
216	R	H83	301
217	A	H84	302
218	E	H85	303
219	D	H86	304
220	T	H87	305
221	A	H88	306
222	V	H89	307
223	Y	H90	308
224	Y	H91	309
225	C	H92	310
226	A	H93	311
227	R	H94	312
228	H	H95	313
229	G	H96	314
230	Y	H97	315
231	T	H98	316
232	L	H99	317
233	D	H100	318
234	Y	H100A	319
235	F	H100B	320
236	E	H101	321
237	Y	H102	322
238	W	H103	323
239	G	H104	324
240	Q	H105	325
241	G	H106	326
242	T	H107	327
243	M	H108	328
244	V	H109	329
245	T	H110	330
246	V	H111	331
247	S	H112	332
248	S	H113	333

248

scFv	Res	Kabat	Fab
1	D	L1	1
2	I	L2	2
3	Q	L3	3
4	M	L4	4
5	T	L5	5
6	Q	L6	6
7	S	L7	7
8	P	L8	8
9	S	L9	9
10	S	L10	10
11	L	L11	11
12	S	L12	12
13	A	L13	13
14	S	L14	14
15	V	L15	15
16	G	L16	16
17	D	L17	17
18	R	L18	18
19	V	L19	19
20	T	L20	20
21	I	L21	21
22	T	L22	22
23	C	L23	23
24	R	L24	24
25	T	L25	25
26	S	L26	26
27	G	L27	27
28	N	L28	28
29	I	L29	29
30	H	L30	30
31	N	L31	31
32	Y	L32	32
33	L	L33	33
34	T	L34	34
35	W	L35	35
36	Y	L36	36
37	Q	L37	37
38	Q	L38	38
39	K	L39	39
40	P	L40	40
41	G	L41	41
42	K	L42	42
43	A	L43	43
44	P	L44	44
45	Q	L45	45
46	L	L46	46
47	L	L47	47
48	I	L48	48
49	Y	L49	49
50	N	L50	50
51	A	L51	51
52	K	L52	52
53	T	L53	53
54	L	L54	54
55	A	L55	55
56	D	L56	56
57	G	L57	57
58	V	L58	58
59	P	L59	59
60	S	L60	60
61	R	L61	61
62	F	L62	62
63	S	L63	63
64	G	L64	64
65	S	L65	65
66	G	L66	66
67	S	L67	67
68	G	L68	68
69	T	L69	69
70	Q	L70	70
71	F	L71	71
72	T	L72	72
73	L	L73	73
74	T	L74	74
75	I	L75	75
76	S	L76	76
77	S	L77	77
78	L	L78	78
79	Q	L79	79
80	P	L80	80
81	E	L81	81
82	D	L82	82
83	F	L83	83
84	A	L84	84
85	N	L85	85
86	Y	L86	86
87	Y	L87	87
88	C	L88	88
89	Q	L89	89
90	H	L90	90
91	F	L91	91
92	W	L92	92
93	S	L93	93
94	L	L94	94
95	P	L95	95
96	F	L96	96
97	T	L97	97
98	F	L98	98
99	G	L99	99
100	Q	L100	100
101	G	L101	101
102	T	L102	102
103	K	L103	103
104	V	L104	104
105	E	L105	105
106	I	L106	106
107	K	L107	107
108	R	L108	108
109	T	L109	109

scFv	Res	Kabat	Fab								
130	E	H1	215	170	P	H41	255	211	Q	H81	296
131	V	H2	216	171	G	H42	256	212	M	H82	297
132	Q	H3	217	172	K	H43	257	213	N	H82A	298
133	L	H4	218	173	R	H44	258	214	S	H82B	299
134	V	H5	219	174	L	H45	259	215	L	H82C	300
135	E	H6	220	175	E	H46	260	216	R	H83	301
136	S	H7	221	176	W	H47	261	217	A	H84	302
137	G	H8	222	177	V	H48	262	218	E	H85	303
138	G	H9	223	178	A	H49	263	219	D	H86	304
139	G	H10	224	179	Y	H50	264	220	T	H87	305
140	L	H11	225	180	I	H51	265	221	A	H88	306
141	V	H12	226	181	S	H52	266	222	V	H89	307
142	Q	H13	227	182	S	H52A	267	223	Y	H90	308
143	P	H14	228	183	G	H53	268	224	Y	H91	309
144	G	H15	229	184	G	H54	269	225	C	H92	310
145	G	H16	230	185	G	H55	270	226	A	H93	311
146	S	H17	231	186	S	H56	271	227	R	H94	312
147	L	H18	232	187	T	H57	272	228	Q	H95	313
148	R	H19	233	188	Y	H58	273	229	N	H96	314
149	L	H20	234	189	F	H59	274	230	K	H97	315
150	S	H21	235	190	P	H60	275	231	K	H98	316
151	C	H22	236	191	D	H61	276	232	L	H99	317
152	A	H23	237	192	T	H62	277	233	T	H100	318
153	A	H24	238	193	V	H63	278	234	W	H100A	319
154	S	H25	239	194	K	H64	279	235	F	H100B	320
155	G	H26	240	195	G	H65	280	236	D	H101	321
156	F	H27	241	196	R	H66	281	237	Y	H102	322
157	D	H28	242	197	F	H67	282	238	W	H103	323
158	F	H29	243	198	T	H68	283	239	G	H104	324
159	S	H30	244	199	I	H69	284	240	Q	H105	325
160	R	H31	245	200	S	H70	285	241	G	H106	326
161	Y	H32	246	201	R	H71	286	242	T	H107	327
162	D	H33	247	202	D	H72	287	243	L	H108	328
163	M	H34	248	203	N	H73	288	244	V	H109	329
164	S	H35	249	204	A	H74	289	245	T	H110	330
165	W	H36	250	205	K	H75	290	246	V	H111	331
166	V	H37	251	206	N	H76	291	247	S	H112	332
167	R	H38	252	207	T	H77	292	248	S	H113	333
168	Q	H39	253	208	L	H78	293				
169	A	H40	254	209	Y	H79	294				
				210	L	H80	295				

References

Abney, E.R., Cooper, M.D., Kearney, J.F., Lawton, A.R., Parkhouse, R., 1978. Sequential expression of immunoglobulin on developing mouse B lymphocytes: a systematic survey that suggests a model for the generation of immunoglobulin isotype diversity. *The Journal of Immunology.* **120,** 2041-2049.

Ahmad, Z.A., Yeap, S.K., Ali, A.M., Ho, W.Y., Alitheen, N.B.M., Hamid, M., 2012. scFv Antibody: Principles and Clinical Application. *Clinical and Developmental Immunology.* **2012.**

Al-Lazikani, B., Lesk, A.M., Chothia, C., 1997. Standard conformations for the canonical structures of immunoglobulins. *Journal of Molecular Biology.* **273,** 927-948.

Arndt, K.M., Müller, K.M., Plückthun, A., 1998. Factors influencing the dimer to monomer transition of an antibody single-chain Fv fragment. *Biochemistry.* **37,** 12918-12926.

Bakema, J.E. & van Egmond, M., 2011. Immunoglobulin A: A next generation of therapeutic antibodies? *mAbs.* **3,** 352-361.

Bax, A., 1994. Multidimensional nuclear magnetic resonance methods for protein studies. *Current Opinion in Structural Biology.* **4,** 738-744.

Becker, C., Fantini, M.C., Schramm, C., Lehr, H.A., Wirtz, S., Nikolaev, A., Burg, J., Strand, S., Kiesslich, R., Huber, S., 2004. TGF-β Suppresses Tumor Progression in Colon Cancer by Inhibition of IL-6< i> trans-Signaling. *Immunity.* **21,** 491-501.

Bedzyk, W.D., Weidner, K., Denzin, L., Johnson, L.S., Hardman, K., Pantoliano, M., Asel, E.D., Voss Jr, E.W., 1990. Immunological and structural characterization of a high affinity anti-fluorescein single-chain antibody. *Journal of Biological Chemistry.* **265,** 18615-18620.

Bhat, T., Bentley, G., Fischmann, T., Boulot, G., Poljak, R., 1990. Small rearrangements in structures of Fv and Fab fragments of antibody D 1.3 on antigen binding. *Nature.* **347,** 483-485.

Bird, R.E., Hardman, K.D., Jacobson, J.W., Johnson, S., Kaufman, B.M., Lee, S.M., Lee, T., Pope, S.H., Riordan, G.S., Whitlow, M., 1988. Single-chain antigen-binding proteins. *Science.* **242,** 423-426.

Bodenhausen, G. & Ruben, D.J., 1980. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chemical Physics Letters.* **1,** 30.

Bongini, L., Fanelli, D., Piazza, F., De Los Rios, P., Sandin, S., Skoglund, U., 2004. Freezing immunoglobulins to see them move. *Proceedings of the National Academy of Sciences of the United States of America.* **101,** 6466-6471.

Boulanger, M.J., Chow, D., Brevnova, E.E., Garcia, K.C., 2003. Hexameric structure and assembly of the interleukin-6/IL-6 {alpha}-receptor/gp130 complex. *Science's STKE.* **300,** 2101.

Boulianne, G.L., Hozumi, N., Shulman, M.J., 1984. Production of functional chimaeric mouse/human antibody.

Brekke, O.H. & Sandlie, I., 2003. Therapeutic antibodies for human diseases at the dawn of the twenty-first century. *Nature Reviews Drug Discovery.* **2,** 52-62.

Cambier, J., Pleiman, C., Clark, M., 1994. Signal transduction by the B cell antigen receptor and its coreceptors. *Annual Review of Immunology.* **12,** 457-486.

Carter, P., Kelley, R.F., Rodrigues, M.L., Snedecor, B., Covarrubias, M., Velligan, M.D., Wong, W.L.T., Rowland, A.M., Kotts, C.E., Carver, M.E., 1992. High level Escherichia coli expression and production of a bivalent humanized antibody fragment. *Nature Biotechnology.* **10,** 163-167.

Carter, P.J., 2006. Potent antibody therapeutics by design. *Nature Reviews Immunology.* **6,** 343-357.

Chalaris, A., Rabe, B., Paliga, K., Lange, H., Laskay, T., Fielding, C.A., Jones, S.A., Rose-John, S., Scheller, J., 2007. Apoptosis is a natural stimulus of IL6R shedding and contributes to the proinflammatory trans-signaling function of neutrophils. *Blood.* **110,** 1748-1755.

Chames, P., Van Regenmortel, M., Weiss, E., Baty, D., 2009. Therapeutic antibodies: successes, limitations and hopes for the future. *British Journal of Pharmacology.* **157,** 220-233.

Chapman, A.P., 2002. PEGylated antibodies and antibody fragments for improved therapy: a review. *Advanced Drug Delivery Reviews.* **54,** 531-545.

Chen, Q., Fisher, D.T., Clancy, K.A., Gauguet, J.M.M., Wang, W.C., Unger, E., Rose-John, S., Von Andrian, U.H., Baumann, H., Evans, S.S., 2006. Fever-range thermal stress promotes lymphocyte trafficking across high endothelial venules via an interleukin 6 trans-signaling mechanism. *Nature Immunology.* **7,** 1299-1308.

Chothia, C. & Lesk, A.M., 1987. Canonical structures for the hypervariable regions of immunoglobulins. *Journal of Molecular Biology.* **196,** 901-917.

Chothia, C., Novotnı, J., Bruccoleri, R., Karplus, M., 1985. Domain association in immunoglobulin molecules:: The packing of variable domains. *Journal of Molecular Biology.* **186,** 651-663.

Choy, E., Hazleman, B., Smith, M., Moss, K., Lisi, L., Scott, D., Patel, J., Sopwith, M., Isenberg, D., 2002. Efficacy of a novel PEGylated humanized anti-TNF fragment (CDP870) in patients with rheumatoid arthritis: a phase II double-blinded, randomized, dose-escalating trial. *Rheumatology.* **41,** 1133-1137.

Clore, G.M., Wingfield, P.T., Gronenborn, A.M., 1991. High-resolution three-dimensional structure of interleukin 1. beta. in solution by three-and four-dimensional nuclear magnetic resonance spectroscopy. *Biochemistry.* **30,** 2315-2323.

Coffman, R.L. & Cohn, M., 1977. The class of surface immunoglobulin on virgin and memory B lymphocytes. *The Journal of Immunology.* **118,** 1806-1815.

Colman, P., Laver, W., Varghese, J., Baker, A., Tulloch, P., Air, G., Webster, R., 1987. Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. *Nature.* **326,** 358-363.

Conrath, K.E., Lauwereys, M., Wyns, L., Muyldermans, S., 2001. Camel single-domain antibodies as modular building units in bispecific and bivalent antibody constructs. *Journal of Biological Chemistry.* **276,** 7346-7350.

Coppieters, K., Dreier, T., Silence, K., Haard, H.D., Lauwereys, M., Casteels, P., Beirnaert, E., Jonckheere, H., Wiele, C.V.D., Staelens, L., 2006. Formatted anti–tumor necrosis factor α VHH proteins derived from camelids show superior potency and targeting to inflamed joints in a murine model of collagen-induced arthritis. *Arthritis & Rheumatism.* **54,** 1856-1866.

Davies, D.R. & Padlan, E.A., 1992. Twisting into shape. *Current Biology: CB.* **2,** 254.

Davies, D.R., Padlan, E.A., Sheriff, S., 1990. Antibody-antigen complexes. *Annual Review of Biochemistry.* **59,** 439-473.

Davies, D. & Chacko, S., 1993. Antibody Structure. *Accounts of Chemical Research.* **26,** 421-427.

Denzin, L., Whitlow, M., Voss Jr, E., 1991. Single-chain site-specific mutations of fluorescein-amino acid contact residues in high affinity monoclonal antibody 4-4-20. *Journal of Biological Chemistry.* **266,** 14095-14103.

Desplancq, D., King, D.J., Lawson, A.D.G., Mountain, A., 1994. Multimerization behaviour of single chain Fv variants for the tumour-binding antibody B72. 3. *Protein Engineering.* **7,** 1027-1033.

Dever, L.A. & Dermody, T.S., 1991. Mechanisms of bacterial resistance to antibiotics. *Archives of Internal Medicine.* **151,** 886.

Diaz, M. & Casali, P., 2002. Somatic immunoglobulin hypermutation. *Current Opinion in Immunology.* **14,** 235-240.

Dominitzki, S., Fantini, M.C., Neufert, C., Nikolaev, A., Galle, P.R., Scheller, J., Monteleone, G., Rose-John, S., Neurath, M.F., Becker, C., 2007. Cutting edge: trans-signaling via the soluble IL-6R abrogates the induction of FoxP3 in naive CD4 CD25–T cells. *The Journal of Immunology.* **179,** 2041-2045.

Dubs, A., Wagner, G., Wüthrich, K., 1979. Individual assignments of amide proton resonances in the proton NMR spectrum of the basic pancreatic trypsin inhibitor. *Biochimica Et Biophysica Acta (BBA)-Protein Structure.* **577,** 177-194.

Essig, N.Z., Wood, J.F., Howard, A.J., Raag, R., Whitlow, M., 1993. Crystallization of single-chain Fv proteins. *Journal of Molecular Biology.* **234,** 897-901.

Faulds, D. & Sorkin, E.M., 1994. Abciximab (c7E3 Fab). A review of its pharmacology and therapeutic potential in ischaemic heart disease. *Drugs.* **48,** 583-598.

Ferentz, A.E. & Wagner, G., 2000. NMR spectroscopy: a multifaceted approach to macromolecular structure. *Quarterly Reviews of Biophysics.* **33,** 29-65.

Fleire, S., Goldman, J., Carrasco, Y., Weber, M., Bray, D., Batista, F., 2006. B cell ligand discrimination through a spreading and contraction response. *Science's STKE.* **312,** 738.

Gardner, K.H. & Kay, L.E., 1998. The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annual Review of Biophysics and Biomolecular Structure.* **27,** 357-406.

Gauldie, J., Richards, C., Harnish, D., Lansdorp, P., Baumann, H., 1987. Interferon beta 2/B-cell stimulatory factor type 2 shares identity with monocyte-derived hepatocyte-stimulating factor and regulates the major acute phase protein response in liver cells. *Proceedings of the National Academy of Sciences.* **84,** 7251.

Gelis, I., Bonvin, A.M.J.J., Keramisanou, D., Koukaki, M., Gouridis, G., Karamanou, S., Economou, A., Kalodimos, C.G., 2007. Structural basis for signal sequence recognition by the 204-kDa translocase motor SecA determined by NMR. *Cell.* **131,** 756.

Goddard, T.D. and Kneller, D.G., SPARKY 3. University of California: San Fancisco.

Griffiths, A.D., Malmqvist, M., Marks, J., Bye, J., Embleton, M., McCafferty, J., Baier, M., Holliger, K.P., Gorick, B., Hughes-Jones, N., 1993. Human anti-self antibodies with high specificity from phage display libraries. *The EMBO Journal.* **12,** 725.

Grishaev, A., Tugarinov, V., Kay, L.E., Trewhella, J., Bax, A., 2008. Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints. *Journal of Biomolecular NMR.* **40,** 95-106.

Grzesiek, S., Anglister, J., Ren, H., Bax, A., 1993. 13C line narrowing by 2H decoupling in 2H/13C/15N-enriched proteins. Application to triple resonance 4D J

connectivity of sequential amides. *Journal of the American Chemical Society.* **115,** 4369-4370.

Grzesiek, S. & Bax, A., 1992. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *Journal of the American Chemical Society.* **114,** 6291-6293.

Haegeman, G., Volckaert, G., Derynck, R., Tavernier, J., FIERS, W., 1986. Structural analysis of the sequence coding for an inducible 26-kDa protein in human fibroblasts. *European Journal of Biochemistry.* **159,** 625-632.

Hall, C.J., 2009. *Comparisons of Solution Structures of Free and IL-1β bound Fab Suggests a Model for B Cell Receptor Signalling.* Ph. D. University of Leicester.

Hanes, J., Jermutus, L., Weber-Bornhauser, S., Bosshard, H.R., Plückthun, A., 1998. Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries. *Proceedings of the National Academy of Sciences.* **95,** 14130.

Harris, L.J., Larson, S.B., Hasel, K.W., McPherson, A., 1997. Refined structure of an intact IgG2a monoclonal antibody. *Biochemistry.* **36,** 1581-1597.

Harris, L.J., Skaletsky, E., McPherson, A., 1998. Crystallographic structure of an intact IgG1 monoclonal antibody. *Journal of Molecular Biology.* **275,** 861-872.

Harwood, N.E. & Batista, F.D., 2009. Early events in B cell activation. *Annual Review of Immunology.* **28,** 185-210.

Havel, T.F. & Wüthrich, K., 1985. An evaluation of the combined use of nuclear magnetic resonance and distance geometry for the determination of protein conformations in solution. *Journal of Molecular Biology.* **182,** 281-294.

Hirano, T., Yasukawa, K., Harada, H., Taga, T., Watanabe, Y., Matsuda, T., Kashiwamura, S., Nakajima, K., Koyama, K., Iwamatsu, A., 1986. Complementary DNA for a novel human interleukin (BSF-2) that induces B lymphocytes to produce immunoglobulin.

Hmila, I., Abdallah R, B.A.B., Saerens, D., Benlasfar, Z., Conrath, K., Ayeb, M.E., Muyldermans, S., Bouhaouala-Zahar, B., 2008. VHH, bivalent domains and chimeric Heavy chain-only antibodies with high neutralizing efficacy for scorpion toxin AahI′. *Molecular Immunology.* **45,** 3847-3856.

Holliger, P. & Hudson, P.J., 2005. Engineered antibody fragments and the rise of single domains. *Nature Biotechnology.* **23,** 1126-1136.

Holliger, P., Prospero, T., Winter, G., 1993. " Diabodies": Small bivalent and bispecific antibody fragments. *Proceedings of the National Academy of Sciences.* **90,** 6444.

Hombach, J., Tsubata, T., Leclercq, L., Stappert, H., Reth, M., 1990. Molecular components of the B-cell antigen receptor complex of the IgM class.

Honjo, T., 1983. Immunoglobulin genes. *Annual Review of Immunology*. **1,** 499-528.

Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.S., Novotný, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., Crea, R., 1988. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. *Proceedings of the National Academy of Sciences*. **85,** 5879.

Ikura, M., Kay, L.E., Bax, A., 1990. A novel approach for sequential assignment of 1H, 13C, and 15N spectra of larger proteins: heteronuclear tripleresonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry*. **29,** 659-654.

Inomata, K., Ohno, A., Tochio, H., Isogai, S., Tenno, T., Nakase, I., Takeuchi, T., Futaki, S., Ito, Y., Hiroaki, H., 2009. High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature*. **458,** 106-109.

Jones, P.T., Dear, P.H., Foote, J., Neuberger, M.S., Winter, G., 1986. Replacing the complementarity-determining regions in a human antibody with those from a mouse.

Jostock, T., Müllberg, J., Özbek, S., Atreya, R., Blinn, G., Voltz, N., Fischer, M., Neurath, M.F., Rose-John, S., 2001. Soluble gp130 is the natural inhibitor of soluble interleukin-6 receptor transsignaling responses. *European Journal of Biochemistry*. **268,** 160-167.

Kabat, E.A. & Wu, T., 1991. Identical V region amino acid sequences and segments of sequences in antibodies of different specificities. Relative contributions of VH and VL genes, minigenes, and complementarity-determining regions to binding of antibody-combining sites. *The Journal of Immunology*. **147,** 1709.

Kay, L.E., Ikura, M., Tschudin, R., Bax, A., 1990. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance*. **89,** 496-514.

Kelly, S.M., Jess, T.J., Price, N.C., 2005. How to study proteins by circular dichroism. *Biochimica Et Biophysica Acta (BBA)-Proteins & Proteomics*. **1751,** 119-139.

Kettleborough, C.A., Saldanha, J., Heath, V.J., Morrison, C.J., Bendig, M.M., 1991. Humanization of a mouse monoclonal antibody by CDR–grafting: the importance of framework residues on loop conformation. *Protein Engineering*. **4,** 773-783.

Kim, Y.M., Pan, J.Y.J., Korbel, G.A., Peperzak, V., Boes, M., Ploegh, H.L., 2006. Monovalent ligation of the B cell receptor induces receptor activation but fails to promote antigen presentation. *Proceedings of the National Academy of Sciences of the United States of America*. **103,** 3327-3332.

Knappik, A. & Plückthun, A., 1995. Engineered turns of a recombinant antibody improve its in vivo folding. *Protein Engineering*. **8,** 81-89.

Köhler, G. & Milstein, C., 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature.* **256,** 495-497.

Kurosaki, T., 2011. Regulation of BCR signaling. *Molecular Immunology.* **48,** 1287-1291.

Kurosaki, T., Shinohara, H., Baba, Y., 2009. B cell signaling and fate decision. *Annual Review of Immunology.* **28,** 21-55.

Kusumi, A., Nakada, C., Ritchie, K., Murase, K., Suzuki, K., Murakoshi, H., Kasai, R.S., Kondo, J., Fujiwara, T., 2005. Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: high-speed single-molecule tracking of membrane molecules. *Annu.Rev.Biophys.Biomol.Struct.* **34,** 351-378.

Kusumi, A., Sako, Y., Yamamoto, M., 1993. Confined lateral diffusion of membrane receptors as studied by single particle tracking (nanovid microscopy). Effects of calcium-induced differentiation in cultured epithelial cells. *Biophysical Journal.* **65,** 2021-2040.

Lee, Y.C., Boehm, M.K., Chester, K.A., Begent, R.H.J., Perkins, S.J., 2002. Reversible dimer formation and stability of the anti-tumour single-chain Fv antibody MFE-23 by neutron scattering, analytical ultracentrifugation, and NMR and FT-IR spectroscopy. *Journal of Molecular Biology.* **320,** 107-127.

LeMaster, D.M. & Richards, F.M., 1985. Proton-nitrogen-15 heteronuclear NMR studies of Escherichia coli thioredoxin in samples isotopically labeled by residue type. *Biochemistry.* **24,** 7263-7268.

Lonberg, N., 2005. Human antibodies from transgenic animals. *Nature Biotechnology.* **23,** 1117-1125.

Mason, D., 1976. The class of surface immunoglobulin on cells carrying IgG memory in rat thoracic duct lymph: the size of the subpopulation mediating IgG memory. *The Journal of Experimental Medicine.* **143,** 1122-1130.

McCafferty, J., Griffiths, A.D., Winter, G., Chiswell, D.J., 1990. Phage antibodies: filamentous phage displaying antibody variable domains. *Letters to Nature*, **348,** 552-554.

Mihara, M., Hashizume, M., Yoshida, H., Suzuki, M., Shiina, M., 2012. IL-6/IL-6 receptor system and its role in physiological and pathological conditions. *Clinical Science.* **122,** 143-159.

Mihara, M., Kasutani, K., Okazaki, M., Nakamura, A., Kawai, S., Sugimoto, M., Matsumoto, Y., Ohsugi, Y., 2005. Tocilizumab inhibits signal transduction mediated by both mIL-6R and sIL-6R, but not by the receptors of other members of IL-6 cytokine family. *International Immunopharmacology.* **5,** 1731-1740.

Milenic, D., Yokota, T., Filpula, D., Finkelman, M., Dodd, S., Wood, J., Whitlow, M., Snoy, P., Schlom, J., 1991. Construction, binding properties, metabolism, and tumor targeting of a single-chain Fv derived from the pancarcinoma monoclonal antibody CC49. *Cancer Research.* **51,** 6363.

Mobli, M., Stern, A.S., Bermel, W., King, G.F., Hoch, J.C., 2010. A non-uniformly sampled 4D HCC (CO) NH-TOCSY experiment processed using maximum entropy for rapid protein sidechain assignment. *Journal of Magnetic Resonance.* **204,** 160-164.

Morell, A., Terry, W.D., Waldmann, T.A., 1970. Metabolic properties of IgG subclasses in man. *Journal of Clinical Investigation.* **49,** 673.

Morrison, S.L., Johnson, M.J., Herzenberg, L.A., Oi, V.T., 1984. Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains. *Proceedings of the National Academy of Sciences.* **81,** 6851.

Mosbæk, C.R., Konarev, P.V., Svergun, D.I., Rischel, C., Vestergaard, B., 2012. High Concentration Formulation Studies of an IgG2 Antibody Using Small Angle X-ray Scattering. *Pharmaceutical Research.* 1-11.

Narazaki, M., Yasukawa, K., Saito, T., Ohsugi, Y., Fukui, H., Koishihara, Y., Yancopoulos, G.D., Taga, T., Kishimoto, T., 1993. Soluble forms of the interleukin-6 signal-transducing receptor component gp130 in human serum possessing a potential to inhibit signals through membrane-anchored gp130. *Blood.* **82,** 1120-1126.

Naugler, W.E. & Karin, M., 2008. The wolf in sheep's clothing: the role of interleukin-6 in immunity, inflammation and cancer. *Trends in Molecular Medicine.* **14,** 109-119.

Nelson, A.L., 2010. Antibody fragments: Hope and hype, *MAbs,* 2010, Landes Bioscience pp77.

Norman, D.*, et al*, 1993. Consensus statement regarding OKT3-induced cytokine-release syndrome and human antimouse antibodies, *Transplantation proceedings,* 1993, Elsevier pp89-92.

Oka, C. & Kawaichi, M., 1995. V (D) J recombination of immunoglobulin genes. *Advances in Biophysics.* **31,** 163-180.

Okumura, K., Julius, M., Tsu, T., Herzenberg, L.A., Herzenberg, L., 1976. Demonstration that IgG memory is carried by IgG-bearing cells. *European Journal of Immunology.* **6,** 467-472.

Otting, G., Qian, Y., Billeter, M., Müller, M., Affolter, M., Gehring, W., Wüthrich, K., 1990. Protein--DNA contacts in the structure of a homeodomain--DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *The EMBO Journal.* **9,** 3085.

Pant, N., Hultberg, A., Zhao, Y., Svensson, L., Pan-Hammarström, Q., Johansen, K., Pouwels, P.H., Ruggeri, F.M., Hermans, P., Frenken, L., 2006. Lactobacilli expressing

variable domain of llama heavy-chain antibody fragments (lactobodies) confer protection against rotavirus-induced diarrhea. *Journal of Infectious Diseases.* **194,** 1580-1588.

Pervushin, K., Riek, R., Wider, G., Wüthrich, K., 1997. Attenuated T2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences.* **94,** 12366.

Pierce, S.K., 2009. Understanding B cell activation: from single molecule tracking, through Tolls, to stalking memory in malaria. *Immunologic Research.* **43,** 85-97.

Pierce, S.K. & Liu, W., 2010. The tipping points in the initiation of B cell signalling: how small changes make big differences. *Nature Reviews Immunology.* **10,** 767-777.

Pilz, I., Kratky, O., Licht, A., Sela, M., 1973. Shape and volume of anti-poly-D-alanyl antibodies in the presence and absence of tetra-D-alanine as followed by small-angle x-ray scattering. *Biochemistry.* **12,** 4998-5005.

Pilz, I., Schwarz, E., Durchschein, W., Light, A., Sela, M., 1980. Effect of cleaving interchain disulfide bridges on the radius of gyration and maximum length of anti-poly (D-alanyl) antibodies before and after reaction with tetraalanine hapten. *Proceedings of the National Academy of Sciences.* **77,** 117-121.

Pilz, I., Schwarz, E., Palm, W., 1977. Small-Angle X-Ray Studies of the Human Immunoglobulin Molecule Kol. *European Journal of Biochemistry.* **75,** 195-199.

Piotto, M., Saudek, V., Sklenr, V., 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *Journal of Biomolecular NMR.* **2,** 661-665.

Poljak, R.J., Amzel, L.M., Avey, H.P., 1973. Three dimensional structure of the Fab′ fragment of a human immunoglobulin at 2.8 A resolution. *Proceedings of the National Academy of Sciences of the United States of America.* **70,** .

Porter, R., 1959. The hydrolysis of rabbit γ-globulin and antibodies with crystalline papain. *Biochemical Journal.* **73,** 119.

Porter, R., 1950. The formation of a specific inhibitor by hydrolysis of rabbit antiovalbumin. *Biochemical Journal.* **46,** 479.

Rabe, B., Chalaris, A., May, U., Waetzig, G.H., Seegert, D., Williams, A.S., Jones, S.A., Rose-John, S., Scheller, J., 2008. Transgenic blockade of interleukin 6 transsignaling abrogates inflammation. *Blood.* **111,** 1021-1028.

Raff, M.S. & Taylor, R., 1970. Immunoglobulin determinants on the surface of mouse lymphoid cells. *Nature.* **225,** 553

Reichert, J.M. and Reichert, J., 2010. Metrics for antibody therapeutics development, *MAbs,* 2010, Landes Bioscience pp695.

Riechmann, L., Clark, M., Waldmann, H., Winter, G., 1988. Reshaping human antibodies for therapy. *Nature.* **332,** 323-327.

Religa, T.L., Sprangers, R., Kay, L.E., 2010. Dynamic regulation of archaeal proteasome gate opening as studied by TROSY NMR. *Science.* **328,** 98-102.

Rolli, V., Gallwitz, M., Wossning, T., Flemming, A., Schamel, W.W.A., Zürn, C., Reth, M., 2002. Amplification of B cell antigen receptor signaling by a Syk/ITAM positive feedback loop. *Molecular Cell.* **10,** 1057-1069.

Rose-John, S., Scheller, J., Elson, G., Jones, S.A., 2006. Interleukin-6 biology is coordinated by membrane-bound and soluble receptors: role in inflammation and cancer. *Journal of Leukocyte Biology.* **80,** 227-236.

Roux, K.H., Strelets, L., Michaelsen, T.E., 1997. Flexibility of human IgG subclasses. *The Journal of Immunology.* **159,** 3372-3382.

Rovnyak, D., Frueh, D.P., Sastry, M., Sun, Z.Y.J., Stern, A.S., Hoch, J.C., Wagner, G., 2004. Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction. *Journal of Magnetic Resonance.* **170,** 15-21.

Royer, C.A., 2006. Probing protein folding and conformational transitions with fluorescence. *Chemical Reviews-Columbus.* **106,** 1769-1784.

Russell, N.D., Corvalan, J.R.F., Gallo, M.L., Davis, C.G., Pirofski, L., 2000. Production of protective human antipneumococcal antibodies by transgenic mice with human immunoglobulin loci. *Infection and Immunity.* **68,** 1820-1826.

Saerens, D., Ghassabeh, G.H., Muyldermans, S., 2008. Single-domain antibodies as building blocks for novel therapeutics. *Current Opinion in Pharmacology.* **8,** 600-608.

Sakakibara, D., Sasaki, A., Ikeya, T., Hamatsu, J., Hanashima, T., Mishima, M., Yoshimasu, M., Hayashi, N., Mikawa, T., Wälchli, M., 2009. Protein structure determination in living cells by in-cell NMR spectroscopy. *Nature.* **458,** 102-105.

Saphire, E.O., Parren, P.W.H.I., Pantophlet, R., Zwick, M.B., Morris, G.M., Rudd, P.M., Dwek, R.A., Stanfield, R.L., Burton, D.R., Wilson, I.A., 2001. Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. *Science.* **293,** 1155-1159.

Saphire, E.O., Stanfield, R.L., Max Crispin, M., Parren, P.W.H.I., Rudd, P.M., Dwek, R.A., Burton, D.R., Wilson, I.A., 2002. Contrasting IgG structures reveal extreme asymmetry and flexibility. *Journal of Molecular Biology.* **319,** 9-18.

Schamel, W.W.A. & Reth, M., 2000. Monomeric and oligomeric complexes of the B cell antigen receptor. *Immunity.* **13,** 5-14.

Scheller, J., Chalaris, A., Schmidt-Arras, D., Rose-John, S., 2011. The pro-and anti-inflammatory properties of the cytokine interleukin-6. *Biochimica Et Biophysica Acta (BBA)-Molecular Cell Research.*

Scheller, J., Grötzinger, J., Rose-John, S., 2006. Updating interleukin-6 classic-and trans-signaling. *Signal Transduction.* **6,** 240-259.

Schmieder, P., Stern, A.S., Wagner, G., Hoch, J.C., 1994. Improved resolution in triple-resonance spectra by nonlinear sampling in the constant-time domain. *Journal of Biomolecular NMR.* **4,** 483-490.

Schmitz, R., Götz, B., Gram, H., 1996. Catalytic specificity of phosphotyrosine kinases Blk, Lyn, c-Src and Syk as assessed by phage display. *Journal of Molecular Biology.* **260,** 664-677.

Schnappinger, D. & Hillen, W., 1996. Tetracyclines: antibiotic action, uptake, and resistance mechanisms. *Archives of Microbiology.* **165,** 359-369.

Schroeder, J., H.W, Cavacini, L., 2010. Structure and function of immunoglobulins. *Journal of Allergy and Clinical Immunology.* **125,** S41-S52.

Scott, A.M., Wolchok, J.D., Old, L.J., 2012. Antibody therapy of cancer. *Nature Reviews Cancer.* **12,** 278-287.

Sehgal, P.B., May, L.T., Tamm, I., Vilcek, J., 1987. Human beta 2 interferon and B-cell differentiation factor BSF-2 are identical. *Science.* **235,** 731-732.

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., Li, Q., 2009. Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Protein Science.* **18,** 936-948.

Skerra, A. & Pluckthun, A., 1988. Assembly of a functional immunoglobulin Fv fragment in Escherichia coli. *Science.* **240,** 1038.

Skerra, A. & Plückthun, A., 1991. Secretion and in vivo folding of the Fab fragment of the antibody McPC603 in Escherichia coli: influence of disulphides and cis-prolines. *Protein Engineering.* **4,** 971-979.

Skiniotis, G., Lupardus, P.J., Martick, M., Walz, T., Garcia, K.C., 2008. Structural organization of a full-length gp130/LIF-R cytokine receptor transmembrane complex. *Molecular Cell.* **31,** 737-748.

Sohn, H.W., Tolar, P., Jin, T., Pierce, S.K., 2006. Fluorescence resonance energy transfer in living cells reveals dynamic membrane changes in the initiation of B cell signaling. *Proceedings of the National Academy of Sciences.* **103,** 8143-8148.

Sohn, H.W., Tolar, P., Pierce, S.K., 2008. Membrane heterogeneities in the formation of B cell receptor–Lyn kinase microclusters and the immune synapse. *The Journal of Cell Biology.* **182,** 367-379.

Sreerama, N. & Woody, R.W., 2000. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Analytical Biochemistry.* **287,** 252-260.

Stanfield, R.L., Takimoto-Kamimura, M., Rini, J.M., Profy, A.T., Wilson, I.A., 1993. Major antigen-induced domain rearrangements in an antibody. *Structure.* **1,** 83-93.

Stangel, M. & Pul, R., 2006. Basic principles of intravenous immunoglobulin (IVIg) treatment. *Journal of Neurology.* **253,** 18-24.

Studier, F.W., 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification.* **41,** 207-234.

Suzuki, T., Ishii-Watabe, A., Tada, M., Kobayashi, T., Kanayasu-Toyoda, T., Kawanishi, T., Yamaguchi, T., 2010. Importance of neonatal FcR in regulating the serum half-life of therapeutic proteins containing the Fc domain of human IgG1: a comparative study of the affinity of monoclonal antibodies and Fc-fusion proteins to human neonatal FcR. *The Journal of Immunology.* **184,** 1968-1976.

Tjandra, N. & Bax, A., 1997. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science.* **278,** 1111-1114.

Tolar, P., Hanna, J., Krueger, P.D., Pierce, S.K., 2009. The constant region of the membrane immunoglobulin mediates B cell-receptor clustering and signaling in response to membrane antigens. *Immunity.* **30,** 44-55.

Tonegawa, S., 1983. Somatic generation of antibody diversity. *Nature.* **302,** 575-581.

Tramontano, A., Chothia, C., Lesk, A.M., 1990. Framework residue 71 is a major determinant of the position and conformation of the second hypervariable region in the VH domains of immunoglobulins. *Journal of Molecular Biology.* **215,** 175-182.

Treanor, B., 2012. B cell receptor: from resting state to activate. *Immunology.*

Treanor, B., Depoil, D., Gonzalez-Granja, A., Barral, P., Weber, M., Dushek, O., Bruckbauer, A., Batista, F.D., 2010. The membrane skeleton controls diffusion dynamics and signaling through the B cell receptor. *Immunity.* **32,** 187-199.

Vaughan, T.J., Williams, A.J., Pritchard, K., Osbourn, J.K., Pope, A.R., Earnshaw, J.C., McCafferty, J., Hodits, R.A., Wilton, J., Johnson, K.S., 1996. Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. *Nature Biotechnology.* **14,** 309-314.

Venkitaraman, A.R., Williams, G.T., Dariavach, P., Neuberger, M.S., 1991. The B-cell antigen receptor of the five immunoglobulin classes. *Nature.* **352,** 777-781

Verhoeyen, M., Saunders, J., Price, M., Marugg, J., Briggs, S., Broderick, E., Eida, S., Mooren, A., Badley, R., 1993. Construction of a reshaped HMFG1 antibody and comparison of its fine specificity with that of the parent mouse antibody. *Immunology.* **78,** 364.

Veverka, V., Henry, A.J., Slocombe, P.M., Ventom, A., Mulloy, B., Muskett, F.W., Muzylak, M., Greenslade, K., Moore, A., Zhang, L., 2009. Characterization of the Structural Features and Interactions of Sclerostin. *Journal of Biological Chemistry.* **284,** 10890-10900.

Vincenti, F., Kirkman, R., Light, S., Bumgardner, G., Pescovitz, M., Halloran, P., Neylan, J., Wilkinson, A., Ekberg, H., Gaston, R., 1998. Interleukin-2–receptor blockade with daclizumab to prevent acute rejection in renal transplantation. *New England Journal of Medicine.* **338,** 161-165.

Vuister, G.W., Kim, S.J., Orosz, A., Marquardt, J., Wu, C., Bax, A., 1994. Solution structure of the DNA-binding domain of Drosophila heat shock transcription factor. *Nature Structural Biology.* **1,** 605-614.

Wabl, M. & Steinberg, C., 1996. Affinity maturation and class switching. *Current Opinion in Immunology.* **8,** 89-92.

Wagner, G. & Wuthrich, K., 1982. Sequential resonance assignments in protein 1H nuclear magnetic resonance spectra. Basic pancreatic trypsin inhibitor. *Journal of Molecular Biology.* **155,** 347-366.

Walters, C.S. & Wigzell, H., 1970. Demonstration of heavy and light chain antigenic determinants on the cell-bound receptor for antigen similarities between membrane-attached and humoral antibodies produced by the same cell. *The Journal of Experimental Medicine.* **132,** 1233-1249.

Wang, L.D. & Clark, M.R., 2003. B-cell antigen-receptor signalling in lymphocyte development. *Immunology.* **110,** 411-420.

Wang, X., Lupardus, P., LaPorte, S.L., Garcia, K.C., 2009. Structural biology of shared cytokine receptors. *Annual Review of Immunology.* **27,** 29-60.

Ward, E.S., Güssow, D., Griffiths, A.D., Jones, P.T., Winter, G., 1989. Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli. *Nature.* **341,** 544-546.

Waters, L., Veverka, V., Böhm, M., Schmedt, T., Choong, P., Muskett, F., Klempnauer, K., Carr, M., 2007. Structure of the C-terminal MA-3 domain of the tumour suppressor protein Pdcd4 and characterization of its interaction with eIF4A. *Oncogene.* **26,** 4941-4950.

Wienands, J., Hombach, J., Radbruch, A., Riesterer, C., Reth, M., 1990. Molecular components of the B cell antigen receptor complex of class IgD differ partly from those of IgM. *The EMBO Journal.* **9,** 449.

Weinblatt, M.E., Keystone, E.C., Furst, D.E., Moreland, L.W., Weisman, M.H., Birbara, C.A., Teoh, L.A., Fischkoff, S.A., Chartash, E.K., 2003. Adalimumab, a fully human anti–tumor necrosis factor α monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. *Arthritis & Rheumatism.* **48,** 35-45.

Weisser, N.E. & Hall, J.C., 2009. Applications of single-chain variable fragment antibodies in therapeutics and diagnostics. *Biotechnology Advances.* **27,** 502-520.

Wesolowski, J., Alzogaray, V., Reyelt, J., Unger, M., Juarez, K., Urrutia, M., Cauerhff, A., Danquah, W., Rissiek, B., Scheuplein, F., 2009. Single domain antibodies: promising experimental and therapeutic tools in infection and immunity. *Medical Microbiology and Immunology.* **198,** 157-174.

Whitlow, M., Bell, B.A., Feng, S.L., Filpula, D., Hardman, K.D., Hubert, S.L., Rollence, M.L., Wood, J.F., Schott, M.E., Milenic, D.E., 1993. An improved linker for single-chain Fv with reduced aggregation and enhanced proteolytic stability. *Protein Engineering.* **6,** 989-995.

Wilkinson, I.C., 2009. *Solution structure of Interleukin-1β bound to a potential therapeutic antibody.* Ph. D. University of Leicester.

Wilkinson, I.C., Hall, C.J., Veverka, V., Shi, J.Y., Muskett, F.W., Stephens, P.E., Taylor, R.J., Henry, A.J., Carr, M.D., 2009. High Resolution NMR-based Model for the Structure of a scFv-IL-1β Complex. *Journal of Biological Chemistry.* **284,** 31928.

Williamson, M.P., Havel, T.F., Wüthrich, K., 1985. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by< sup> 1</sup> H nuclear magnetic resonance and distance geometry. *Journal of Molecular Biology.* **182,** 295-315.

Williamson, R.A., Carr, M.D., Frenkiel, T.A., Feeney, J., Freedman, R.B., 1997. Mapping the binding site for matrix metalloproteinase on the N-terminal domain of the tissue inhibitor of metalloproteinases-2 by NMR chemical shift perturbation. *Biochemistry.* **36,** 13882-13889.

Winau, F., Westphal, O., Winau, R., 2004. Paul Ehrlich—in search of the magic bullet. *Microbes and Infection.* **6,** 786-789.

Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S., Sykes, B.D., 1995. 1 H, 13 C and 15 N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *Journal of Biomolecular NMR.* **5,** 67-81.

Wishart, D., Sykes, B., Richards, F., 1991. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of Molecular Biology.* **222,** 311-333.

Wittekind, M. & Mueller, L., 1993. HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha-and beta-carbon resonances in proteins. *Journal of Magnetic Resonance.Series B.* **101,** 201-205.

Wörn, A. & Plückthun, A., 2001. Stability engineering of antibody single-chain Fv fragments1. *Journal of Molecular Biology.* **305,** 989-1010.

Wörn, A. & Plückthun, A., 1999. Different equilibrium stability behavior of ScFv fragments: identification, classification, and improvement by protein engineering. *Biochemistry.* **38,** 8739-8750.

Wörn, A. & Plückthun, A., 1998. Mutual stabilization of VL and VH in single-chain antibody fragments, investigated with mutants engineered for stability. *Biochemistry.* **37,** 13120-13127.

Xu, G.Y., Hong, J., McDonagh, T., Stahl, M., Kay, L.E., Seehra, J., Cumming, D.A., 1996. Complete 1 H, 15 N and 13 C assignments, secondary structure, and topology of recombinant human interleukin-6. *Journal of Biomolecular NMR.* **8,** 123-135.

Yang, J. & Reth, M., 2010. The dissociation activation model of B cell antigen receptor triggering. *FEBS Letters.* **584,** 4872-4877.

Yang, J. & Reth, M., 2010a. Oligomeric organization of the B-cell antigen receptor on resting cells. *Nature.* **467,** 465-469.

Yoon, S., Kim, Y.S., Shim, H., Chung, J., 2010. Current perspectives on therapeutic antibodies. *Biotechnology and Bioprocess Engineering.* **15,** 709-715.

Zheng, Y., Shopes, B., Holowka, D., Baird, B., 1992. Dynamic conformations compared for IgE and IgG1 in solution and bound to receptors. *Biochemistry.* **31,** 7446-7456.

Zhu, G., Kong, X.M., Sze, K.H., 1999. Gradient and sensitivity enhancement of 2D TROSY with water flip-back, 3D NOESY-TROSY and TOCSY-TROSY experiments. *Journal of Biomolecular NMR.* **13,** 77-81.