Comparisons of Solution Structures of Free and IL-1β Bound Fab Suggests a Model for B Cell Receptor Signalling

Catherine Jane Hall

MChem (Leicester)

Thesis submitted for the degree of Doctor of Philosophy

Department of Biochemistry

University of Leicester

October 2009

Abstract

Comparisons of Solution Structures of Free and IL-1β bound Fab Suggests a Model for B Cell Receptor Signalling

Catherine Jane Hall

Therapeutic antibodies are an important and growing class of biotherapeutics with the potential to treat a range of major human diseases. Structural knowledge of the complexes formed between antibodies and target proteins would provide valuable information about the mode of action of the therapeutic and potentially reveal a structural mechanism underlying antigen induced B cell receptor signalling. Often the Fab fragment of the antibody is used in therapeutics, however, there are currently no NMR-based structures available for Fab or Fab/antigen complexes, presumably due to the difficulties of producing suitable samples and obtaining NMR data for large proteins and complexes. The work described here illustrates the development of approaches that have made it possible to obtain high quality NMR-based structures for the gIC8 Fab alone and in complex with its antigen, IL-1β, including the development of approaches that allow the measurement of reliable RDC data. Expression and purification of a $^{15}N/^{13}C/^{2}H$-labelled gIC8 Fab fragment allowed for collection of triple resonance data for the free Fab and over 90% of the visible backbone resonances to be assigned. A refined homology model of the free gIC8 Fab and a docked and refined structure for the gIC8 Fab/IL-1β complex were obtained using experimental NMR restraints. Comparisons made between the free and bound Fab structures highlight small antigen induced changes in domain orientation, which suggest a mechanism for B cell receptor signalling. The work reported here shows that NMR spectroscopy can be used as a tool to obtain detailed structural data for large Fab and Fab-target protein complexes and provide an improved understanding of how antibodies interact with their target proteins.

Acknowledgements

First and foremost I would like to offer my sincerest gratitude to my supervisor Professor Mark Carr for giving me the opportunity to work on this project and for his guidance and encouragement throughout. I extend my thanks to Drs Richard Taylor and Alistair Henry for their advice and support during my PhD, and to their colleagues at UCB-Celltech who kindly provided me with some of the data presented in this thesis. In addition I would like to thank Dr Mark Pfuhl, for his helpful advice during my committee meetings. I am also grateful to the Biotechnology and Biological Sciences Research Council and UCB-Celltech for funding this studentship and to Bruker Biospin for allowing us to collect some data on their 950 MHz spectrometer.

I offer special thanks to Dr Vaclav Veverka as without his guidance, patience and friendship I would not have been able to complete this research. I also gratefully acknowledge Ian Wilkinson, both for his important scientific input to this project and for his friendship. It is a pleasure to thank all of the other people who made this thesis possible; in particular I would like to acknowledge Drs Lorna Waters, Fred Muskett, Phil Renshaw and Kirsty Lightbody for their instruction and support. I would also like to thank all my others friends and colleagues, especially Ojay, Dariush, Ali, Sami, Sarah, Phil, Alice, Sid, Stacey, Kiran, Helen, Jo and Alan for helping me to enjoy my time at Leicester.

I offer a heartfelt thank you to my Mum, Dad, Will and to my grandparents for the love, support and encouragement they have given me throughout my education. Finally, I thank Jon for his love and support, and for keeping me smiling through the good and bad times.

Catherine Jane Hall, Leicester, October 2009

Table of Contents

Abbreviations

A_a	Axial component of the alignment tensor
AIR	Ambiguous interaction restraint
A_r	Rhombic component of the alignment tensor
A_{xx}, A_{yy}, A_{zz}	Principal axes of the alignment tensor
B cell	B lymphocyte
B_0	Static magnetic field
BCR	B cell receptor
CDR	Complementarity determining regions
C_H	Constant heavy domain
C_L	Constant light domain
Cryo-ET	Cryo-electron tomography
CSA	Chemical shift anisotropy
DD	Dipole-dipole interactions
DHPC	Dihexanoyl-phosphatidylcholine
DMPC	Dimyristoyl-phosphatidylcholine
DNA	Deoxyribonucleic acid
ELISA	Enzyme-linked immunosorbent assay
EM	Electron microscopy
Fab	Antigen binding fragment
Fc	Crystallisable fragment
FPLC	Fast protein liquid chromatography
FR	Frame work
FRET	Förster resonance energy transfer

Fv	Variable fragment
HADDOCK	High ambiguity driven docking approach
HSQC	Heteronuclear single quantum correlation
Ig	Immunoglobulin
IL-1β	Interleukin-1 beta
INEPT	Insensitive nuclei enhanced by polarization transfer
IPAP	In-phase and anti-phase
IPTG	Isopropyl-1-thio-β-D-galactopyranoside
ITAM	Immunoreceptor tyrosine-based activation motif
mAb	Monoclonal antibody
mIg	Membrane-bound immunoglobulin
NEM	N-Ethylmaleimide
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser Effect Spectroscopy
OD_{600}	Optical density at 600 nm
PALES	Prediction of alignment from structure
PBS	Phosphate buffered saline
PDB	Protein data bank
PEG	Poly(ethylene glycol)
Q factor	Quality factor
R	Rhombicity
RDC	Residual dipolar coupling
RDC_{calc}	Theoretical RDC

RDC_{obs}	Measured RDC
RMSD	Root mean square deviation
rmsdp	Root mean squared displacement
scFv	Single chain variable fragment
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
T_2	Spin-spin relaxation time
TALOS	Torsion angle likelihood obtained from shift and sequence similarity
TEV	Tobacco etch virus
TROSY	Transverse relaxation-optimised spectroscopy
V(D)J	Variable, diversity and joining immunoglobulin gene segments
V_H	Variable heavy domain
V_L	Variable light domain
WATERGATE	Water suppression through gradient tailored excitation
γ	Gyromagnetic ratio
$\Delta\delta$	Combined chemical shift
τ_c	Rotational correlation time

Amino acid abbreviations and one letter codes

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 Structure and function of antibodies

All multicellular organisms need to protect themselves against infection by harmful pathogens. Along with an innate immune response, which relies on protective barriers, toxic molecules and phagocytic cells to destroy invading microorganisms, vertebrates have a more sophisticated immune response to invading pathogens, called the adaptive response (Janeway, 2005). This immune response can respond to millions of different foreign antigens in a specific way, by the production of a vast range of antigen receptors (immunoglobulins; Igs), which are presented on the surface of B lymphocytes (B cells). Antigen binding activates the B cells to produce soluble immunoglobulins (antibodies) which bind antigen with high specificity, resulting in neutralisation or destruction of the pathogen (Roitt and Delves, 2001).

The first atomic resolution structure of an antibody fragment was the New IgG1 (λ) Fab (antigen binding fragment), solved in 1973 by X-ray crystallography (Poljak et al., 1973). Even before this early structure, chemical methods had been used to establish that the basic structure of antibodies consisted of two identical light (L) chains and two identical heavy chains (H), linked by noncovalent interactions and disulphide bonds. The heavy chains are covalently linked together at a hinge region while the light chains are linked to the heavy chains with disulphide bonds (Figure 1.1). The heavy and light chains are folded into repeated immunoglobulin folds (Poljak et al., 1973), compact globular structures consisting of two layers of antiparallel β sheets, stabilised with a disulphide bond. Each IgG antibody molecule has two identical Fab fragments, which

consist of the complete light chains associated with the variable and one constant domain from the heavy chains. A hinge region links the Fabs to the Fc (crystallisable fragment) which consists of constant domains from the heavy chain (Figure 1.1) (Ramsland and Farrugia, 2002). The Fc interacts with effector cells and complement to initiate immune responses. Antigens bind to the antibody variable domains (Fv) where the binding sites are primarily formed by six hypervariable loops, three from the V_L domain (L1, L2, L3) and three from the V_H domain (H1, H2, H3) known as complementarity determining regions (CDRs) shown in Figure 1.1. The regions of the variable domains outside of these loops are called the framework (FR) (Davies et al., 1975).

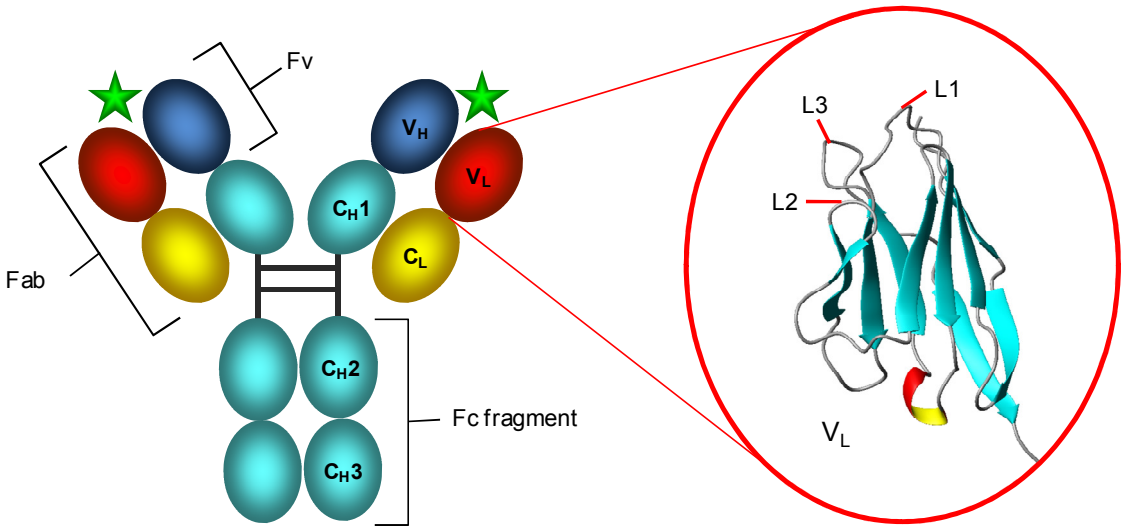


Figure 1.1

Schematic diagram of an IgG antibody. Each immunoglobulin domain is shown as an oval. The heavy chain has three constant domains (C_H1, C_H2 and C_H3) shown in light blue and an N-terminal variable domain (V_H) coloured in dark blue. The shorter light chains consist of one constant domain (C_L, yellow) and one variable domain (V_L, red). The Fc, Fab and Fv regions are indicated and the antigen binding sites are highlighted with green stars. The enlargement shows a ribbon representation of the variable light domain showing the β sheet framework and the CDR loops.

There are 5 human antibody isotypes, each of which contains a different heavy chain and this determines the antibody structure and effector function. The isotypes are known as IgM, IgD, IgG, IgE and IgA and the corresponding heavy chains are called μ, δ, γ, ε and α respectively (Figure 1.2). IgG is the most abundant antibody in the serum, accounting for about 75% of the total circulating immunoglobulin. It is monomeric with

two antigen binding sites and a molecular mass of ~150 kDa. IgG is important in complement activation and is also able to bind to cell surface Fc receptors. The IgG isotype is further divided into four classes in humans called IgG_1, IgG_2, IgG_3 and IgG_4, which is based on their relative concentrations in the serum. The classes are structurally very similar, differing only in the hinge region. IgG_1, IgG_2, IgG_3 and IgG_4 have hinges consisting of 15, 12, 62 and 12 residues respectively, meaning that the flexibility of the Fab fragments with respect to the Fc varies between the subclasses. IgM circulates in the blood and lymph and accounts for 5 to 10% of the total antibody. In an initial immune response to an unknown antigen, IgM is the first antibody to be produced. Before affinity maturation has occurred (see later), most antibodies will bind antigen with a low affinity. IgM has a pentameric structure, which allows it to bind to polyvalent antigens with high avidity and trigger reactions with complement, which can kill microorganisms directly or facilitate their phagocytosis. IgA accounts for about 15% of the antibody population and can exist as a monomer or a dimer. The monomeric form is secreted into the bloodstream, whilst the dimeric form is secreted into the lumen of the gut and is also the main antibody in secretions such as milk, saliva, sweat and tears. IgD accounts for less than 1% of the total immunoglobulins. It has a monomeric structure and most of the population is present on the surface of lymphocytes. IgE is the least abundant immunoglobulin isotype, accounting for only 0.002% of the total immunoglobulins. Its heavy chain is structurally different to the other classes in that the hinge region is replaced by an extra constant domain. IgE plays a significant role in allergy and is associated with type 1 hypersensitivity. Two types of light chain, λ and κ, are found in antibodies of all isotypes. Unlike the heavy chains, no functional difference has been found between antibodies containing λ or κ light chains (Roitt and Delves, 2001).

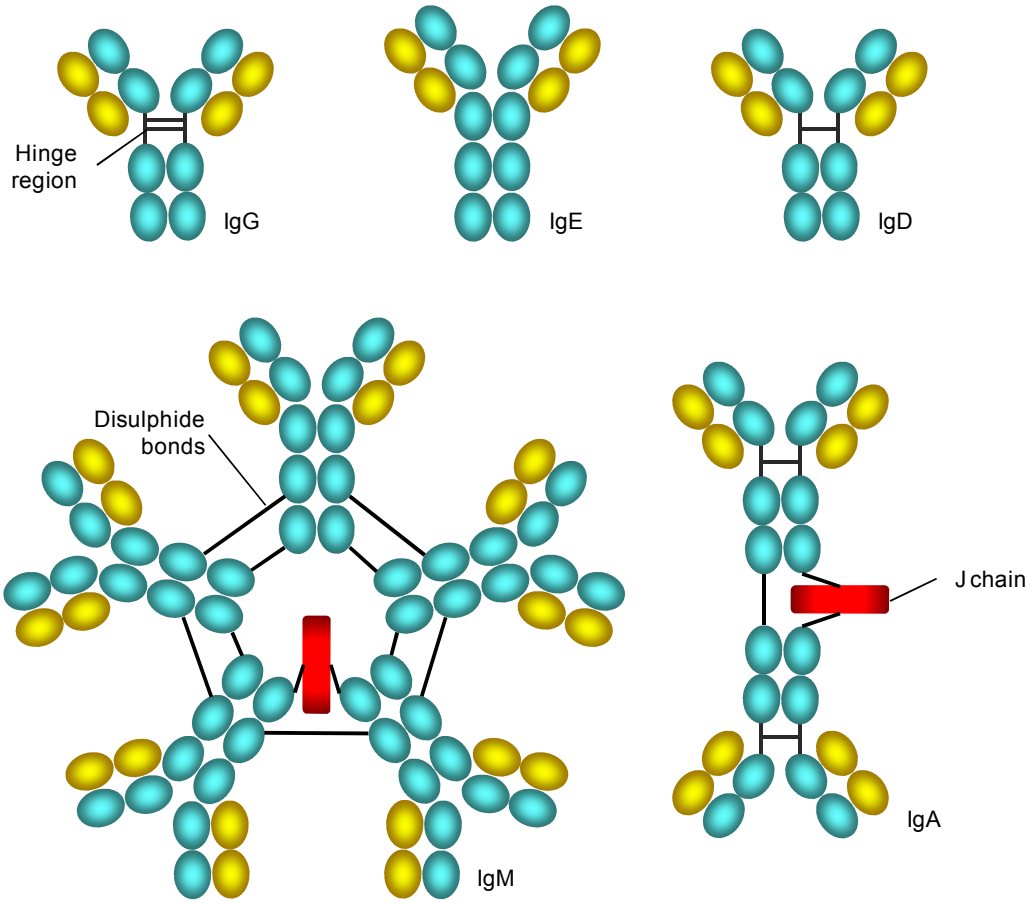


Figure 1.2

Schematic representation of the structures of the antibody classes. Each of the classes has two light chains shown in yellow and two heavy chains shown in blue. J chains, shown in red, are 15 kDa polypeptide chains which promote polymerisation of IgM and IgA which can exist as multimers, as shown here.

Membrane-bound immunoglobulin (mIg) receptors on the surface of B cells form part of the B cell receptors (BCRs). All immunoglobulin isotypes can occur in a membrane bound form although IgM and IgD are the most abundant as they occur on circulating B cells. IgG, IgA and IgE are expressed on the surface of memory B cells and plasma

cells. Each B cell only presents antibodies with the same variable regions and therefore the same antigen specificity and affinity. BCRs are activated by antigen binding, which initiates multiple cellular responses ultimately resulting in B cell proliferation and differentiation (Janeway, 2005). Firstly, BCR activation triggers a signalling pathway which results in transcription of genes important for proliferation and differentiation. Antigen binding to the BCR also activates the internalisation of the antibody-bound antigen, which leads to its degradation into peptide fragments. The fragments are then presented on the surface of the B cell, bound to a membrane glycoprotein called a major histocompatibility complex II (MHC II). Helper T cells specifically recognise the MHC II-bound antigen fragments and secrete cytokines which induces the B cells to proliferate and differentiate into soluble antibody secreting plasma cells and memory B cells (Janeway, 2005).

B cells are able to initiate a specific immune response to antigens due to the enormous diversity of antigen binding sites expressed across the B cell population. This diversity partially arises in the initial B cell development stage in the lymphoid organs where gene segments encoding the variable domains are rearranged. Further diversity is introduced by the process of somatic hypermutation which occurs in mature, activated B cells. The variable regions of antibody light chains are encoded by variable (V) and joining (J) gene segments. In addition to the V and J gene segments, variable regions of heavy chains are also encoded by a third segment called the diversity (D) gene segment. These gene segments are recombined in a B and T lymphocyte-specific process known as V(D)J recombination to create a continuous exon which codes for a specific variable domain (Jung et al., 2006). Diversity in the variable domain occurs as the germline DNA contains multiple different copies of each of the gene segments of which only one

of each is randomly selected for recombination. The endonuclease RAG (encoded by the recombination activating genes) introduces DNA double-strand breaks specifically at the border between the coding sequences and their flanking recombination sequences (RS). The sequences encoding CDRs 1 and 2 is contained within the V segment, whist the region encoding CDR3 lies at the interface between the V and J and is also partially encoded by the D segment in the heavy chain (Bassing et al., 2002). The multiple copies of the V, D and J gene segments combined with different pairing of the heavy and light chains, could potentially give over 10^6 different antibody specificities and further diversity is introduced as a result of the imprecise repair of the DNA double strand breaks. Addition or subtraction of nucleotides at the joining points between the gene segments translates to increased diversity in CDR3 in both the heavy and light chains leading to an estimated repertoire of 10^{11} antibodies of different specificities expressed on the surfaces of a population of B cells (Janeway, 2005).

Additional diversity is introduced to the heavy and light V genes in the peripheral lymphoid organs, after functional immunoglobulin genes have been assembled. This process, known as somatic hypermutation, introduces point mutations into the V region as a response to B cell activation by antigen binding, resulting in variant B cell receptors being expressed on the B cell surface (Diaz and Casali, 2002). Somatic hypermutation drives a process known as affinity maturation. Mutations that increase BCR affinity for antigen are selected for as they compete for a limited amount of antigen. Those B cells that are selected proliferate and expand, whilst B cells expressing BCRs with lower antigen affinity are removed by apoptosis. This process results in the expression of soluble antibodies with very high antigen affinity. IgM is the first antibody type secreted in the immune response, although later in the response the same

V region is expressed in IgG, IgA or IgE antibodies. This is due to a process called isotype, or class, switching and occurs as a result of irreversible DNA recombination between the V(J)D region and regions encoding the heavy chains, which determines the isotype of antibody secreted by the plasma cell. Soluble antibodies can then participate in host defences in three main ways (Roitt and Delves, 2001). The first important role of antibodies is neutralisation of bacteria, toxins and viruses by binding to them and preventing access to cells which they may invade or destroy. Secondly, the Fc region of antibodies is recognized by phagocytes, so antibodies can coat pathogens, targeting them for phagocytosis. Finally, antibodies are able to activate the complement system (Janeway, 2005).

1.2 B cell receptor signalling

B cells are activated by specific antigen binding to the B cell receptor (BCR), which consists of a membrane bound antibody, termed the antigen binding subunit, associated with the signalling subunits, Igα and Igβ (Figure 1.3). Membrane bound immunoglobulins (mIg) have identical antigen binding properties to the secreted antibodies and only differ structurally in that they contain a 25 amino acid transmembrane domain and short cytoplasmic tail C-terminal to the last constant domain. The cytoplasmic portion varies in length between isotypes, with the tails of IgM and IgD consisting of 3 amino acids whereas the other Ig classes show more extended cytoplasmic tails of either 28 (IgG, IgE) or 14 (IgA) amino acids (Reth, 1992).

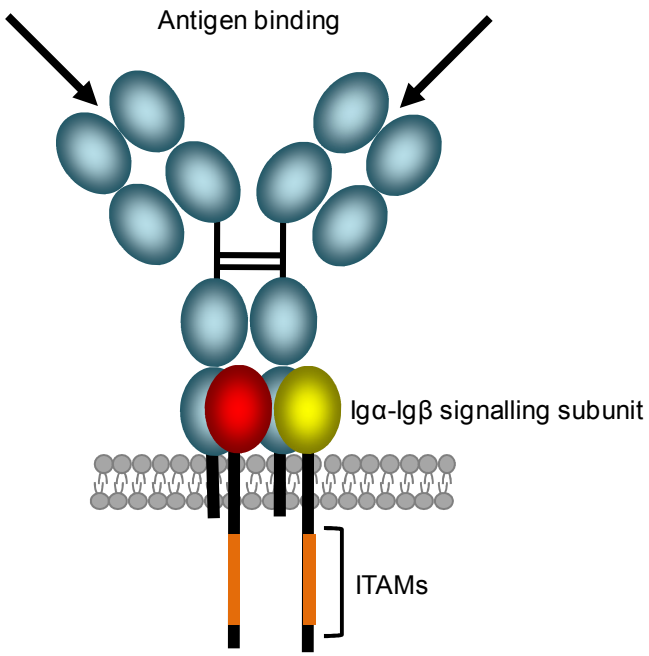


Figure 1.3

Schematic representation of a B cell receptor. The immunoglobulin antigen binding subunit is shown in blue with the antigen binding sites indicated with arrows. The associated Igα-Igβ heterodimers are shown in red and yellow and the approximate position of the ITAM motifs are shown.

The associated signalling subunits are transmembrane proteins with an extracellular Ig-like domain, a transmembrane domain and cytoplasmic domains of 61 and 48 amino acids for Igα and Igβ, respectively (Cambier et al., 1994). The cytoplasmic regions of both Igα and Igβ carry a signalling motif termed the immunoreceptor tyrosine-based activation motif (ITAM) with the consensus sequence D/ExxxxxxxD/ExxYxxL/IxxxxxxxYxxL/I (where x denotes any amino acid) (Reth,

1989; Sigalov, 2005). As explained below, phosphorylation of tyrosine residues within the Igα/Igβ cytoplasmic tails plays an important role in B cell signalling.

BCRs are activated by antigen binding to the mIg, triggering phosphorylation of tyrosine residues in the ITAM motifs of the Igα/Igβ cytosolic tails. The mechanism by which this occurs is unclear, but most evidence suggests that the initial phosphorylation is mediated by the Src-family tyrosine kinases Lyn, Fyn, Blk and Lck, which are anchored in the membrane in close proximity to the BCR (Figure 1.4) (Cambier et al., 1994).

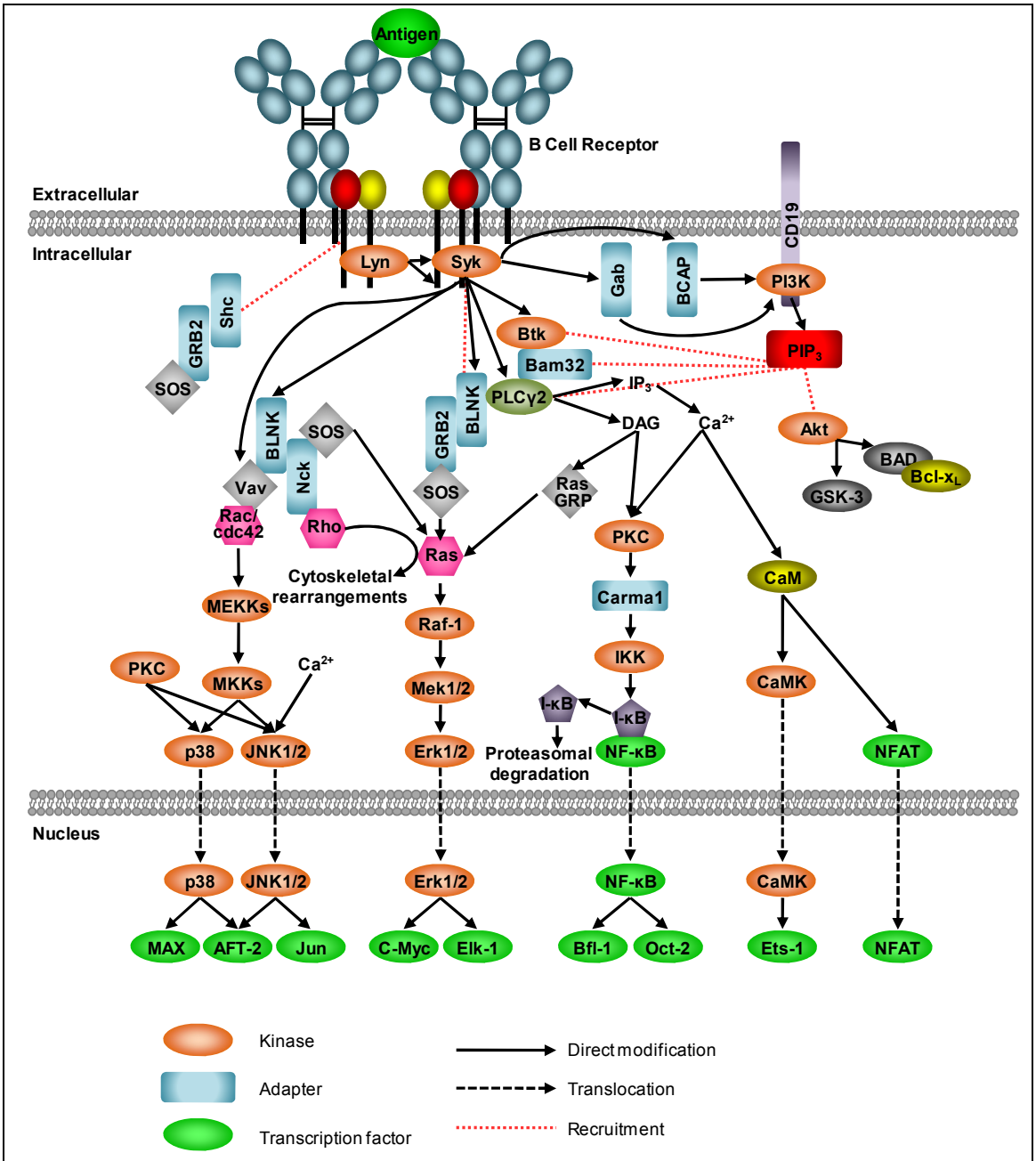


Figure 1.4

B cell receptor signal transduction cascade. BCR signalling is initiated by antigen binding to the BCR at the cell surface, triggering phosphoylation of the ITAM domains in the cytoplasmic tails of the signalling subunits Igα-Igβ. Signals are then propagated by protein phosphorylation, interactions and modification. The pathway ultimately regulates transcription factors and expression of genes involved in B cell proliferation and differentiation. This figure is modified from (Dal Porto et al., 2004).

After initial phosphorylation of the ITAM tyrosines, Lyn can bind Igα/Igβ via its Src homology 2 (SH2) domain, which increases its kinase activity thereby resulting in amplified ITAM phosphorylation and the subsequent recruitment and activation of the cytosolic tyrosine kinase Syk. Activated Syk then recruits and phosphorylates the scaffold molecule BLNK (B cell linker protein), which plays an important role in several different signalling pathways. BCR activation also triggers Lyn-mediated phosphorylation of tyrosines within the cytoplasmic tail of the integral transmembrane protein CD19, which create binding sites for PI-3K. Once localised to the plasma membrane, PI-3K phosphorylates its substrate, phosphatidylinosidol 4,5-bisphosphate [$PI(4,5)P_2$], to form phosphatidylinosidol 3,4,5-trisphosphate [$PI(3,4,5)P_3$]. This molecule is essential for the recruitment and activation of pleckstrin homology (PH) domain-containing proteins including Btk (Bruton's tyrosine kinase), which is crucial for proper BCR signal transduction.

Once the key kinases Lyn, Syk and Btk are activated, the next stage of BCR signalling is the activation of phospholipase C (PLC)-γ2 and the elevation of Ca^{2+} levels. The adaptor protein BLNK plays an important role as a platform for effector molecule assembly in several different pathways. Firstly, BLNK recruits PLCγ2 to the plasma membrane where it is activated by Syk and Btk and can cleave the membrane associated phosphoinositide $PI(4,5)P_2$ into the second messengers inositol 1,4,5-trisphosphate (IP_3) and diacyl glycerol (DAG). Production of IP_3 immobilises Ca^{2+} from intra- and extra-cellular stores, which is required for the activation of the transcription factors NF-κB (nuclear factor kappa-light-chain-enhancer of activated B cells) by protein kinase C (PKC) and nuclear factor of activated T-cells (NFAT) by Ca^{2+}-calmodulin. DAG activates PKC isotypes which regulate the mitogen activated protein kinase (MAPK)

family. Secondly, phosphorylated BLNK associates with the guanine nucleotide exchange factor Vav, which can activate the Rac pathway, and the growth factor receptor-bound protein 2 (Grb2)-son of sevenless homolog (Sos) complex, which can activate the Ras-MAPK pathway. Finally, BLNK recruits other adaptors, such as Nck, which is associated with cytoskeletal remodelling.

The mechanisms by which signalling pathways, induced by the activation of BCRs, are connected to the activation of transcription and gene regulation are not well characterised. MAPK-family members, extracellular signal-regulated kinase (ERK), c-Jun NH2-terminal kinase (JNK) and p38 MAPK, are activated and regulated by a system of cytoplasmic kinases which are activated by BCR aggregation. Upon activation, each MAPK family member phosphorylates a different set of transcription factors including Elk-1 and c-Myc (ERK), c-Jun and AFT-2 (JNK) and MAX (p38 MAPK). Another molecule which connects events occurring after receptor activation to processes occurring in the nucleus is the serine/threonine kinase, Akt, which has an important role in promoting cell survival. Akt is recruited to the cell membrane as a result of PIP_3 production, and is activated. Akt can then interact with a number of cytosolic proteins including the pro-apoptotic protein Bad, causing it to dissociate from the anti-apoptotic protein Bcl-x_L. In addition, Akt can inhibit the expression of several pro-apoptotic proteins by directly regulating transcription factors such as E2F (E2 transcription factor), NF-κB and CREB (cAMP response element binding). Akt also inhibits the cell cycle regulatory protein, glycogen synthase kinase-3 (GSK-3). BCR signalling also causes the translocation of the NF-κB family of transcription factors to the nucleus by targeting the inhibitory protein, I-κB for ubiquitiation.

The strength of the signal from Igα/Igβ can be influenced by several other molecules including CD19, CD22 and a low affinity receptor for IgG known as FcγRIIB. CD19 positively contributes to the signal by increasing the recruitment and activation of Lyn, PI-3K, Btk and Vav. CD22 and FcγRIIB both contain immunoreceptor tyrosine based inhibitory motifs (ITIMs) in their cytoplasmic domains, which are phosphorylated by Lyn and suppress the BCR response.

However, arguably the greatest unknown factor of this signalling pathway is the mechanism by which antigen binding to an extracellular domain can initiate cell signalling. Several models have been put forward, which postulate an allosteric change in the cytoplasmic tail region of the mIg upon antigen binding, which initiates phosphorylation of the Igα/Igβ ITAMs and triggers downstream signalling events. One model suggests that an allosteric change occurs in the tail region of the receptor on antigen binding, which activates a pre-associated or proximal tyrosine kinase to phosphorylate its own ITAM tyrosines (Cambier et al., 1994). Alternatively, conformational changes in the mIg could lead to hydrogen peroxide formation, which could act as a second messenger in lymphocyte activation (Reth, 2002). It has also been suggested that an antigen-induced structural change may translocate the BCR into distinct membrane signalling domains (rafts) that are enriched in protein tyrosine kinases (PTKs) such as Lyn, which could increase the phosphorylation of the ITAMs (Geisberger and Achatz, 2003).

Another general model suggests that clustering of BCRs activates cytoplasmic kinases to phosphorylate the ITAM tyrosines of a neighbouring BCR. This could potentially occur by BCR interaction with a multivalent antigen or as a result of random

interactions of BCRs diffusing in the plasma membrane. The latter model assumes that the BCR is stabilised on antigen binding, increasing the chances of an encounter with an unbound BCR which is diffusing freely in the membrane. Lyn could be pre-associated with the BCR, either bound to the Igα/Igβ, or the mIg tail of the inactive BCR. It is therefore unclear if Lyn binds to Igα/Igβ through its SH2 domain and whether a background level of ITAM phosphorylation is necessary, or if the interaction is independent of other phosphotyrosines. There is also evidence for the presence of a transducer complex which binds to the unphosphorylated Igα/Igβ ITAMs and is dissociated on BCR activation making ITAM tyrosines available for phosphorylation by Src family kinases (Reth and Wienands, 1997). Another model suggests that BCRs exist in preformed oligomers that are disrupted upon antigen binding, thus initiating cell signalling (Reth, 2001).

The mechanism of BCR activation by membrane bound antigen was recently investigated using confocal microscopy and Förster resonance energy transfer (FRET) in live cells (Tolar et al., 2005; Tolar et al., 2008). Cell lines were engineered to express fluorescent probes in both the extracellular domains and cytoplasmic tails of mIg, Igα and Igβ, and these were used to analyse the change in BCR conformation upon antigen binding. Their results indicated firstly, that the inactive BCR is monomeric in live B cells contradicting the suggestion that BCRs exist as preformed oligomers. Secondly, the results showed an increase in FRET between extracellular domains on antigen binding, providing evidence to support antigen-induced BCR clustering. Based on FRET data obtained from the intracellular domains, a model for an antigen binding-induced conformational change of the BCR was proposed, described as opening of the cytoplasmic domains, which was likened to the opening of an umbrella. FRET was also

observed between the labelled BCRs and fluorescently labelled 'raft lipids' supporting the idea of BCR translocation into lipid rafts on activation by antigen. Further work by the same group analysed the dynamics of individual BCRs interacting with membrane-bound monovalent and multivalent antigen using single-molecule total internal reflection fluorescence (TIRF) microscopy (Tolar et al., 2009). Their results suggested that following antigen binding, the BCRs cluster and are then immobilised on engagement with several BCRs in close proximity. They propose that it is this immobilisation which initiates BCR signalling pathways. When monovalent antigen was used, the immobilisation appeared to be dependent on the presence of the membrane proximal domain and the N-terminal part of the transmembrane region of the mIg. However, the deletion of these regions had no effect on BCR immobilisation in the presence of multivalent antigen, suggesting that crosslinking of the BCRs by multivalent antigen may immobilise the BCR using a different mechanism to monovalent antigen binding. The results also suggested that the membrane proximal mIg domain clusters spontaneously and activates cell signalling when in BCR-like complexes containing Igα/Igβ. Based on this evidence, they proposed that there may be an interaction site in this domain which is obscured in the complete, unbound mIg and is exposed by a conformational change induced by antigen binding. However, it is also possible that the domain aggregation is a result of many other new surfaces exposed by the deletion of surrounding domains, not due to the exposure of one specific BCR clustering site. Current opinion supports a model where antigen binding induces a structural change that is propagated to the signalling subunit and induces B cell clustering followed by signalling, although there is no evidence to support this (Tolar and Pierce, 2009).

1.3 Therapeutic antibodies

Antibodies have been selected through evolution for their ability to recognise and eliminate pathogenic and disease antigens with high specificity and affinity, making them ideal molecules for use as targeted therapies. Generation of mouse monoclonal antibodies (mAbs) by Köhler and Miller in 1975 (Kohler and Milstein, 1975) paved the way for the release of the first approved therapeutic antibody, muromomab (Orthoclone) in 1986, a murine antibody that is used to treat organ rejection. There are currently 22 molecules in clinical use (summarised in Table 1.1) and with a predicted market value of $26 billion by 2010, therapeutic antibodies are clearly developing into a beneficial and profitable group of products (Chames et al., 2009).

Two other examples of murine antibodies raised against human antigens have been approved, however, although this class of antibody have excellent specificity for the human targets, in many cases their use was limited as they were recognized as foreign and quickly eliminated by the immune system when injected into patients. In addition, the mouse antibodies were unable to interact with components of the human immune system, further limiting their therapeutic potential (Chames et al., 2009).

Generic name	Trade name	Type of molecule	Antibody format	Antigen	Approved indication	Year approved
Muromomab	Orthoclone	Murine	IgG2a	CD3	Allograft rejection in allogeneic renal	1986
Abciximab	ReoPro	Chimeric	IgG1 (Fab)	GPIIb/IIIa R	Maintenance of coronary patency	1994
Rituximab	Mabthera	Chimeric	IgG1	CD20	CD20-positive B-cell non-Hodgkin's lymphoma	1997
Daclizumab	Zenapax	Humanized	IgG1	CD25 (IL-2R)	Allograft rejection	1997
Basiliximab	Simulect	Chimeric	IgG1	CD25 (IL-2R)	Allograft rejection	1998
Palivizumab	Synagis	Humanized	IgG1	Protein F	Respiratory syncytial virus (RSV inhibitor) in children	1998
Infliximab	Remicade	Chimeric	IgG1	TNFα	Crohn's disease and rheumatoid arthritis	1998
Trastuzumab	Herceptin	Humanized	IgG1	HER2/Neu	Metastatic breast cancer	1998
Gemtuzumab	Mylotarg	Humanized	IgG4 (Coupled to calicheamicin)	CD33	CD33-positive acute myeloid leukemia	2000
Alemtuzumab	Mabcampath	Humanized	IgG1	CD52	B-cell chronic lymphocytic leukemia	2001
Ibritomomab	Zevalin ^{90}Y	Mouse	IgG1 (Coupled to radioisotope)	CD20	B-cell non-Hodgkin's lymphoma	2002
Adalimumab	Trudexa	Human	IgG1 (PD)	TNFα	Crohn's disease and rheumatoid arthritis	2002
Omalizumab	Xolair	Humanized	IgG1	IgE	Treatment of asthma	2003
Tositumomab	Bexxar ^{131}I	Murine	IgG2a (Coupled to radioisotope)	CD20	CD20-positive B-cell non-Hodgkin's lymphoma	2003
Efalizumab	Raptiva	Humanized	IgG1	CD11a	Moderate to severe plaque psoriasis	2003
Cetuximab	Erbitux	Chimeric	IgG1	EGFR	Metastatic colorectal and head and neck carcinoma	2004
Bevacizumab	Avastin	Humanized	IgG1	VEGF-A	Metastatic colorectal and non-small cell lung	2004
Natalizumab	Tysabri	Humanized	IgG4	Integrin-α4	Multiple sclerosis	2004
Ranibizumab	Lucentis	Humanized	IgG1	VEGF-A	Wet-type age-related macular degeneration	2006
Panitumumab	Vectibis	Human	IgG2	EGFR	Metastatic colorectal carcinoma	2006
Eculizumab	Soliris	Humanized	IgG2/4	C5	Paroxysmal nocturnal haemoglobinuria	2007
Certolizumab	Cimzia	Humanized	IgG1 (PEGylated Fab)	TNFα	Crohn's disease and rheumatoid arthritis	2008

Table 1.1

Monoclonal antibodies approved for therapeutic use. A summary of the 22 monoclonal antibodies approved by the US Food and Drug Administration (FDA) in the years indicated. This data highlights the progress made in the field since the first

approved therapeutic in 1986 and the range of diseases that can be treated by monoclonal antibodies, including cancers and autoimmune diseases. Abbreviations shown in the table are: EGFR, epidermal growth factor receptor; HER, human epidermal growth factor receptor; TNF, tumour necrosis factor; VEGF, vascular endothelial growth factor. The table was adapted and updated from (Chames et al., 2009).

Advances in molecular biology in the 1990s made it possible to clone the genes of IgG molecules into eukaryotic expression vectors (Winter and Milstein, 1991). This technology was used to create chimeric antibodies, containing mouse variable domains grafted onto human constant domains (Chadd and Chamow, 2001). These chimeras are about 70% human and contain a full human Fc region, resulting in considerably less immunogenic molecules that can interact with components of the human immune system. Further developments in antibody engineering technology allowed for the mouse contribution to the therapeutic to be reduced further to create 'humanised' antibodies, where only mouse hypervariable regions are grafted onto a human framework. However, it was often seen that the humanised antibody had a lower affinity for the antigen than that of the parent murine antibody, implying that framework regions played an important role in the interaction with antigen (Jain et al., 2007). Indeed, as discussed later, of eight antibody/antigen complex structures selected for detailed interface analysis three showed direct interactions between the antigen and framework regions in the variable domains (Wilkinson et al., 2009). This highlights the necessity for rational design of therapeutic antibodies, where residues involved in the interaction surface are specifically identified and conserved. Another major advance in antibody

technology was the development of *in vitro* selection methods, including phage display (McCafferty et al., 1990). In this technique, the variable (V) domains are expressed on the surface of bacteriophage and are selected by binding to the antigen of interest. Once the phage displaying the correct variable domains is selected, the encoding DNA can be amplified, allowing for a large library of potential targets to be screened (Hoogenboom, 2005). Fully human antibodies can also be produced from transgenic mice, in which the mouse immunoglobulin loci are replaced with human immunoglobulin genes (Filpula 2007)(Green, 1999).

The simplest mode of action of a therapeutic antibody is binding to an antigen and disrupting its activity by blocking interactions with binding partners. The antigen can be a soluble ligand, as is the case for the anti-TNFα antibodies infliximab (Remicade), adalimumab (Trudexa) and certolizumab (Cimzea), and the anti-vascular endothelial growth factor antibody bevacizumab (Avastin). Antibodies can also target cell surface receptors, blocking the interaction with its ligand as is the case with the anti-epidermal growth factor receptor antibody cetuximab (Erbitux) and the anti-HER2 antibody trastuzumab (Herceptin) (Chames et al., 2009). Therapeutic antibodies can also function through destruction of target cells by activation of components of the human immune system. After antigen binding the antibody can recruit effector cells such as natural killer cells, macrophages or neutrophils. The antibody could also activate complement to destroy the target either through antibody-dependant cell cytotoxicity (ADCC) or complement-dependant cytotoxicity (CDC), both of which are mediated through the Fc region of the antibody (Brekke and Sandlie, 2003). An example of a drug which functions in this way is the anti-CD20 antibody rituximab (Rituxan). Monoclonal antibodies can also be fused to therapeutic agents such as toxins, cytokines or drug-

activating enzymes, to specifically target antigens that are overexpressed in cancer cells, directing the therapeutic to the tumour microenvironment (Jain et al., 2007; Wu, 2005). An example of an immunotoxin is the anti CD33-antibody gemtuzumab (Mylotarg), a humanised mAb conjugated to the anti-tumour antibiotic calicheamicin (Reichert and Valge-Archer, 2007). Most of the therapeutic mAbs available on the market are chimeric or humanised and are almost exclusively of the IgG class, as shown in Table 1.1. IgG antibodies have two Fab fragments, making them bivalent, and a single Fc domain. This can interact with the neonatal Fc receptor (FcRn) resulting in the long (>10 days) serum half-life of IgG, and can also recruit cytotoxic effector functions.

There are a range therapeutic applications where the mAb Fc domain is unnecessary or undesirable, for example, *in vivo* imaging which requires short serum half-lives (Nelson and Reichert, 2009). IgGs have been dissected into smaller fragments using proteolytic techniques and later genetically engineered into a range of mono- and multi-valent fragments (Figure 1.5). Antibody Fab and scFv (V_H and V_L domains joined by a flexible linker, in a single polypeptide) fragments usually retain the specific antigen binding affinity of the full length IgG. These smaller fragments have different properties to full length mAbs, resulting in differences in their potential applications. Firstly, they are smaller than full size mAbs and are therefore able to penetrate into tissues and tumours more rapidly and deeply. However, their small size also means that they have short circulating half-lives, which although useful for some applications for example, delivery of cytotoxic isotopes, can result in insufficient accumulation of the therapy at the target site. The circulating half-life of a fragment can be increased by conjugation to poly(ethylene glycol) (PEG) in a process known as PEGylation, which is discussed in more detail later. Also, the smaller, unglycosylated fragments can be produced in

microbial expression systems, which is less costly than the mammalian expression systems required for full length mAbs. Finally, whilst therapeutic fragments can function by blocking the interaction between two binding partners, they are unable to function by activation of components of the human immune system, as they lack the Fc region (Nelson and Reichert, 2009).

More recently, antibody fragments have been engineered into multivalent designs in order to increase the functional affinity (avidity) and to provide slower dissociation rates from cell surface or multimeric antigens (Holliger and Hudson, 2005). Figure 1.5 shows Fab fragments that have been chemically cross linked into divalent and trivalent multimers. These di- and trivalent molecules show improved retention and cell internalisation compared with the parent IgG molecule. Single chain Fvs have been engineered into successful multimeric structures by reducing the scFv linker to less than five amino acids which promotes self assembly into bivalent dimers (diabodies), trivalent trimers (triabodies) or tetravalent tetrabodies (Holliger and Hudson, 2005).

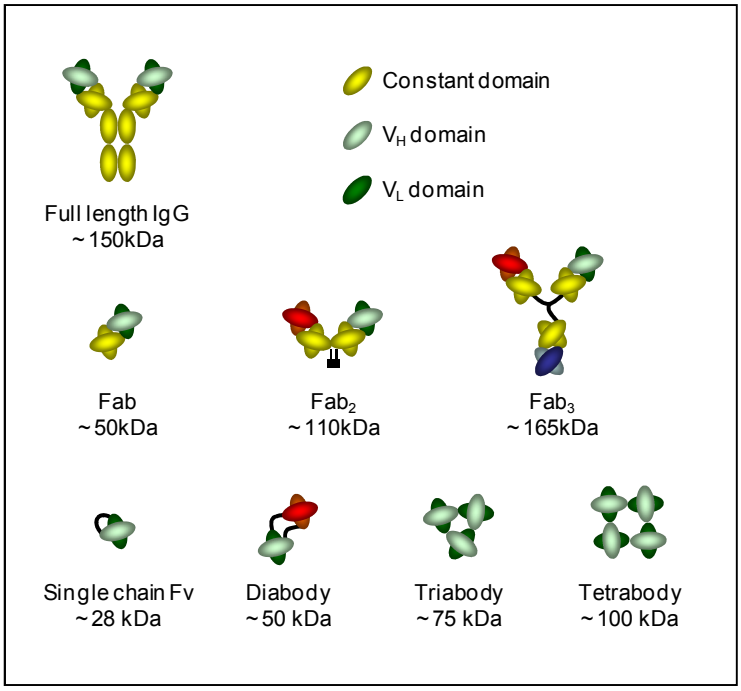


Figure 1.5

Schematic representation of different antibody fragments. The full length IgG molecule is shown alongside a variety of antibody fragments. The Fab fragment consists of two constant and two variable antigen binding domains. Fab fragments can be modified to form divalent (Fab_2) or trivalent (Fab_3) fragments. Single chain Fv (scFv) fragments consist of only the variable domains linked by a 15-25 amino acid linker. These fragments can also be modified to form multivalent fragments (diabodies, triabodies and tetrabodies are shown). This figure was adapted from (Holliger and Hudson, 2005)

1.4 IL-1β

The cytokine interleukin-1 (IL-1) is a member of a group of proteins which mediate immunologic and inflammatory responses to infection and tissue damage (Dinarello, 2002). Interleukin-1β (IL-1β) was the first in the family of 11 IL-1 cytokines to be studied and it was found to be involved in stimulating expression of the early response genes *c-fos* and *c-jun* and also other cytokines and inflammatory factors. Inhibition of IL-1β has been shown to have a positive effect in several inflammatory diseases including rheumatoid arthritis, adult onset Still's disease (effecting the joints and several visceral organs) and systemic-onset juvenile rheumatoid arthritis which causes severe inflammation in very young children (Dinarello, 2009). IL-1β acts by binding to a membrane-bound receptor protein, the interleukin-1 receptor (IL-1R), on the surface of primary blood monocytes or tissue macrophages (Dinarello, 1988). The activity of IL-1β could therefore be inhibited by interrupting the interaction between the cytokine and the receptor and there are several neutralising antibodies both to IL-1β and IL-1R under development (McCulloch et al., 2006). In addition to being an important therapeutic target, IL-1β has been shown to be easily expressed in *E. coli* and several high resolution crystal structures (Priestle et al., 1989) and an NMR structure (Clore et al., 1991) have been published for the free form. This makes IL-1β an ideal model protein for the development of improved NMR approaches.

1.5 Determination of the structures of large proteins using NMR spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy, X-ray crystallography and electron microscopy (EM) are currently the only techniques able to determine the three-dimensional structure of proteins at atomic resolution. Since the first experiments performed with biological macromolecules in 1957 (Saunders et al., 1957), continuing

advances in NMR technology, most notably superconducting magnets, Fourier transform spectroscopy and computer control of instruments, combined with development in techniques for collection of multidimensional NMR, led to the first 3D structure determined solely using NMR data in 1985 (Williamson et al., 1985). However, using early methodology it was difficult to obtain NMR structures for proteins with a molecular weight exceeding 20-25 kDa. This size limit is attributable to increased spectral overlap from increased NMR signals in the spectra of large proteins and also because larger proteins have short transverse relaxation rates which results in increased decay of the signal and poor spectral sensitivity. Significant advances have been made in techniques to address these issues, however, NMR solution structures available for proteins greater than about 40-50 kDa are still limited. This is demonstrated in Figure 1.6 which shows a histogram summarising the molecular weight distribution of all the solution state NMR structures in the Protein Data Bank (PDB) (data collected on 15/09/09).

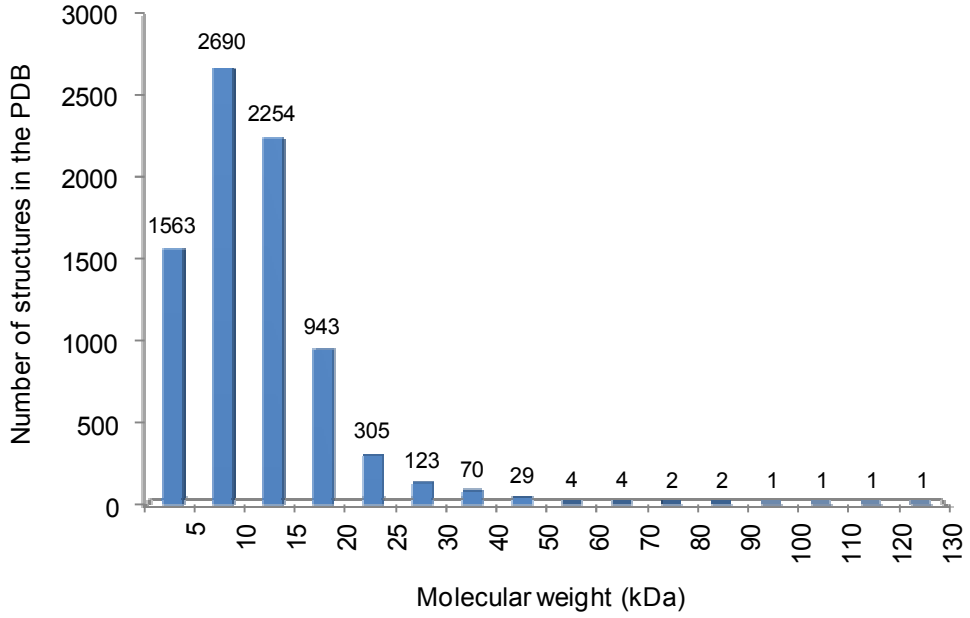


Figure 1.6

Molecular weight distribution of the solution state NMR structures deposited in the PDB. The histogram shows a current (15/09/09) summary of all the NMR solution structures that have been deposited in the PDB and their molecular weights. The number of structures in each group is indicated above the bar. This histogram shows clearly that there are very few structures with molecular weights greater than ~40-50 kDa due to the difficulties of obtaining NMR data for large proteins.

NMR is unique from other spectroscopic techniques in that each spin ½ nuclei or group of equivalent nuclei gives rise to a specific resonance in the spectrum and the position of the signals in the spectra are dependent on the chemical environment of the nuclei. The initial stage of data analysis is to associate each of the resonances in the NMR spectrum with a specific nucleus from a specific residue in the protein sequence.

Spectra acquired from larger proteins have more resonances in the spectra from the increased number of residues in the protein backbone resulting in spectral overlap (Wuthrich, 1990). To reduce the level of overlap, three dimensional experiments are acquired for larger proteins which spread the NMR peaks over three frequency axes. To collect 3D experiments, the protein sample must be uniformly labelled with ^{15}N and/or ^{13}C. This is readily achieved by expressing the protein in bacterial over-expression systems in minimal media containing labelled nitrogen and carbon compounds (e.g. $(^{15}NH_4)_2SO_4$ and ^{13}C glucose) as the sole nitrogen and carbon sources (Clore and Gronenborn, 1991).

Triple resonance spectra can be used to assign ^{1}HN, ^{15}N, ^{13}CO, $^{1}H\alpha$, $^{13}C\alpha$, $^{13}C\beta$ and $^{1}H\beta$ backbone atoms using sequential and intra-residue correlations through heteronuclear one and two bond couplings. Triple resonance experiments are primarily built from a series of INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) building blocks. INEPT transfer is an effective method of transferring magnetisation from sensitive nuclei with a high gyromagnetic ratio (usually proton) to less sensitive nuclei with lower gyromagnetic ratios (for example nitrogen or carbon) by means of the scalar coupling interaction (Cavanagh et al., 1996). This is illustrated in Figure 1.7, which shows an example of an HNCA pulse sequence (Sattler et al., 1999).

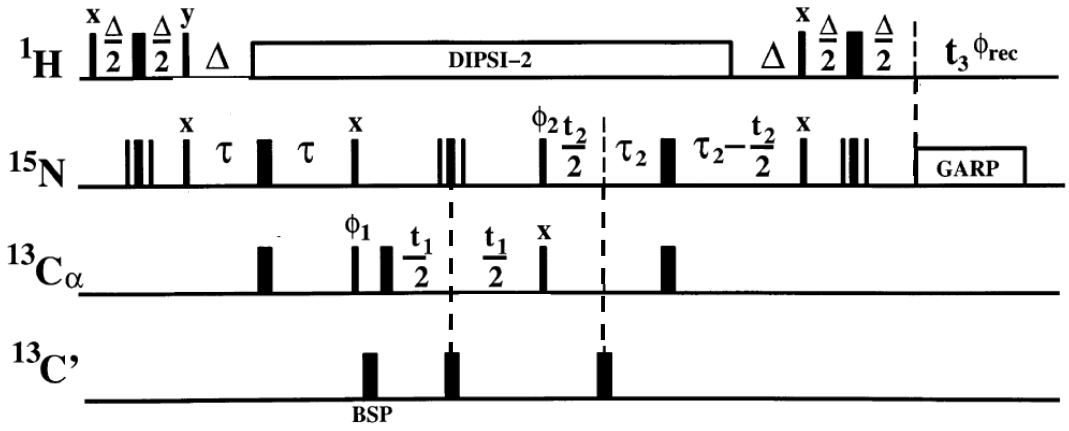


Figure 1.7

Pulse scheme of an out and back 3D HNCA experiment. The magnetisation transfer follows a HN→N→Cα(t_1)→N(t_2) →HN(t_3) pathway, where t_1 and t_2 are evolution periods and t_3 is the detection period. INEPT type transfer sequences are used to transfer magnetism between nuclei. This figure was adapted from (Sattler et al., 1999).

A strategy for the assignment of ^{1}HN, ^{15}N, and ^{13}Cα backbone resonances combines the use of HNCA, which provides intraresidue and sequential correlations between ^{1}HN, ^{15}N and ^{13}Cα resonances, and HN(CO)CA spectra which provide correlations between ^{1}HN and ^{15}N of one residue and ^{13}Cα of the preceding residue. Ambiguities in backbone assignments can be resolved by the collection of HNCACB and HN(CO)CACB spectra which, in addition to the information provided by the HNCA and HN(CO)CA spectra, also provide intra and/or inter-residue Cβ correlations. The chemical shifts of Cα and Cβ are characteristic of the different types of amino acids and can therefore be used to position a sequentially connected stretch of amino acids in a known sequence (Cavanagh et al., 1996).

One of the most important NMR parameters used for structure determination are NOEs (nuclear Overhauser effect). The cross peaks seen in the 3D NOESY-HSQC spectra are generated from through space magnetisation transfer and, for non-deuterated samples, can be seen between two hydrogen atoms that are separated by a maximum distance of ~5-6 Å. NOE connections can be seen between sequential residues in which case they are used to assist in backbone assignments. However, the most important role of NOEs is determination of protein structure as they can provide through space connections between residues that are far apart in the sequence but close together in space. Also, the NOE peak intensity is dependent on the distance between the two interacting ^{1}H spins and so for structure calculation this information is used to define approximate distances between protons. NOE restraints provide direct information about protein conformation and are, therefore, very important in the determination of protein global folds. Residual dipolar couplings (RDCs) are restraints that provide information about the global orientation of the whole protein. RDCs are through space interactions between NMR active nuclei which are observed as a contribution to scalar coupling values when the protein is in an anisotropic environment and can be used to refine the orientation of individual bond vectors in structural calculations. Further details on measurement, applications and the theory of RDCs will be discussed later.

Phi and psi backbone torsion angles can also be used as restraints in protein structure determination. Phi and psi angles can be predicted from a combination of five chemical shift assignments (Hα, Cα, Cβ, CO and N) and the protein sequence, using the program TALOS (Torsion Angle Likelihood Obtained from Shift and sequence similarity) (Cornilescu et al., 1999). The prediction is based on knowledge that the secondary shift

values (which are defined as the observed value of the chemical shift minus the value expected for the same residue in a random coil) are highly correlated with aspects of protein secondary structure. TALOS compares fragments of the observed secondary shift values to a database containing fragments of previously determined secondary shifts which correspond to known high resolution structures. Phi and psi angles of the unknown sequence can then be predicted based on secondary shift similarities.

Another issue for NMR data collection from large proteins is that spectral linewidth increases with the size of the protein. Linewidth is proportional to the overall rotational correlation time of the protein (τ_c). For an approximately spherical, globular protein, τ_c can be determined using the equation:

$$\tau_c = \frac{4\pi\eta_w r_H^3}{(3k_B T)}$$

Equation 1.1

where η_w is the viscosity of the solvent, r_H is the effective hydrodynamic ratio of the protein, k_B is the Boltzmann constant and T is the temperature. As the rotational correlation time increases, spin-spin relaxation times (T_2) decrease, resulting in spectral linewidth broadening and decreased sensitivity. This contributes to poor spectral resolution and reduction in the signal to noise ratio of the NMR spectra (Cavanagh et al., 1996).

In order to collect high quality triple resonance data from large proteins, techniques have been developed to reduce or avoid unfavourable relaxation pathways. The

transverse relaxation-optimised spectroscopy (TROSY) technique can significantly reduces transverse relaxation rates thus reducing spectral linewidths significantly (Riek et al., 1999). The high magnetic fields used for NMR data acquisition of proteins means that chemical shift anisotropy interactions (CSA) between ^{15}N, ^{13}C and ^{1}H contribute heavily to spin relaxation along with dipole-dipole (DD) interactions. This results in an increase in overall transverse relaxation rates which are amplified with increase in magnetic field strength. Cross-correlation between DD and CSA interference results in different relaxation rates, and consequently different linewidths, for the multiplet components in a system of two coupled spins, for example ^{15}N-^{1}H of an amide backbone bond. For one component of the multiplet the interference causes cancellation of the T_2 relaxation, meaning the linewidth is significantly narrower and TROSY observes exclusively this narrow component (Figure 1.8) (Pervushin et al., 1997). The TROSY effect is dependent on field strength and the effect is most pronounced for larger proteins as the line broadening caused by T_2 relaxation becomes more severe as molecular mass increases. Since only one component of the ^{15}N-^{1}H multiplet is retained there is intrinsic loss in signal sensitivity, however, for measurements of large proteins at ^{1}H field strengths of 500 MHz or greater, TROSY provides an increase in both sensitivity and resolution compared to the decoupled HSQC spectrum (Pervushin et al., 1997).

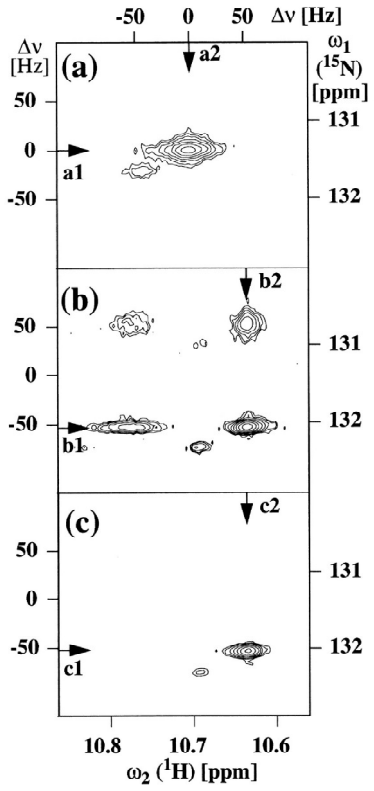


Figure 1.8

Contour plots of $^{15}N/^{1}H$ correlation spectra of one residue from a 40 kDa protein at 750 MHz showing the linewidth difference between standard HSQC and TROSY spectra. Panel A shows a signal from a decoupled HQSC spectrum. Panel B shows the same section as A recorded without decoupling highlighting the difference in linewidths between the four multiplet components. Panel C shows a TROSY type spectrum where only the sharpest peak is selected. Figure is taken from (Pervushin, 2000).

Sensitivity in 3D NMR spectra can also be improved by reducing the transverse relaxation rate constants for ^{13}C and ^{1}HN spins. This is achieved by producing a protein sample in which most, if not all, of the carbon-bound hydrogen atoms are replaced with

deuterium. Deuterons have an approximately six-fold lower gyromagnetic ratio relative to protons, resulting in a considerable increase in the relaxation times of ^{13}C and ^{1}HN spins in proximity of a substituted deuteron. Deuterated proteins show spectra with significantly improved resolution and sensitivity (Clore and Gronenborn, 1998). Multiple deuteration strategies have been developed and can be selected as appropriate depending on the protein and the data required. Deuterated proteins can be produced in bacterial expression systems designed for unlabelled proteins which contain the relevant labelled substrate in the minimal growth media with water replaced by D_2O. Perdeuterated proteins, where all the non-exchangeable protons are replaced with deuterium, are over-expressed in bacteria grown in media containing 100% D_2O and a perdeuterated carbon source (Venters et al., 1995). Although protein perdeuteration greatly assists assignment of backbone atoms, only ^{1}HN-HN NOEs can be collected from the samples. A very effective method of retaining the sensitivity provided by perdeuteration and obtaining ^{1}H-^{1}H NOEs to provide side chain distance restraints in large proteins, is to reintroduce ^{1}H spins at selected sites in otherwise perdeuterated proteins. Most strategies for selective protonation focus on the labelling of side chain methyl groups (Ile, Leu, Val, Ala) as they occur frequently in the hydrophobic core and because their spins relax slowly, therefore, providing sensitive reporters for molecular structure and dynamics (Tugarinov et al., 2006). Aromatic side chain rings (Phe, Tyr, Trp) can also be selectively protonated to provide useful long distance restraints. Selective proton labelling is achieved by expressing the protein in fully deuterated minimal media containing a selectively protonated carbon source. For example, $^{1}H/^{13}C$ pyruvate was used as the primary carbon source to produce deuterated protein with selected protonation at the methyl groups of Ala, Val, Leu and Ile-γ2 (Rosen et al., 1996), and 3-$^{2}H/^{13}C$ α-ketoisovalerate and ^{13}C α-ketobutyrate were used with $^{13}C/^{2}H$

glucose to selectively protonate the Leu-δ, Val-γ and Ile-δ methyl groups (Goto et al., 1999). Selective labelling of aromatic residues has been described in proteins expressed in minimal media containing D_2O and perdeuterated glucose, supplemented with fully protonated shikemic acid (Rajesh et al., 2003).

For NMR data collection of large proteins the most successful results are achieved if the TROSY spectra are obtained from deuterated proteins. For example, assignments of the backbone atoms of $^{15}N/^{13}C/^{2}H$ malate synthase G (MSG), a 723 residue (82 kDa) enzyme were determined using a suite of 4D TROSY experiments (Kay, 2005). Selective protonation of MSG methyl groups then allowed for the global fold to be determined using long range HN-HN (99), CH_3-CH_3 (386) and HN-CH_3 (357) NOEs augmented with ^{1}H-^{15}N residual dipolar couplings (RDCs), along with phi and psi angle restraints from the chemical shifts (Tugarinov and Kay, 2003).

1.6 Thesis overview

The work in this thesis will describe the development of approaches that have made it possible to obtain NMR-based structures of a free gIC8 Fab and gIC8 Fab in complex with its target protein, IL-1β. Chapter 2 will discuss the development of approaches for the determination of backbone amide RDC data, investigating the properties of Pf1 phage and phospholipid bicelle alignment media using IL-1β (17 kDa) as a model protein. The transfer of this technology to larger (~65 kDa) IL-1β/Fab complexes is also described in this chapter.

Chapter 3 describes expression of a $^{15}N/^{13}C/^{2}H$-labelled gIC8 Fab fragment for NMR spectroscopy in *E. coli*. The initial expression trials were carried out on unlabelled

protein to select an optimal *E. coli* strain which was used for the expression of ^{15}N-labelled material. An HSQC spectrum collected from this sample was well dispersed with many resolved peaks suggesting that it would be possible to obtain good quality triple resonance data from the Fab. The second part of Chapter 3 describes the optimisation of a protocol for expression and purification of $^{15}N/^{13}C/^{2}H$ and $^{15}N/^{2}H$-labelled gIC8 Fab which was used to collect high quality 2D and 3D NMR data.

Chapter 4 describes the collection of NMR data from the Fab which was used as restraints to refine a model of free gIC8 Fab and to obtain a docked structure of the gIC8 Fab/IL-1β complex. In this chapter the high quality of the triple resonance data is highlighted, which allowed for a high percentage of the HN, N, Cα, Cβ and CO backbone resonances to be assigned and the residues at the interaction interface to be determined using chemical shift perturbation techniques. The sucessful refinement of a gIC8 homology model using backbone amide RDCs, dihedral angles obtained from the chemical shift data and long range HN-HN NOEs restraints is described. The refinement and docking of a gIC8 homology model and a crystal structure of IL-1β using the same NMR derived restraints in addition to a knowledge of the protein interaction surfaces derived from chemical shift perturbation data. Detailed comparisons made between the free and bound Fab structures show significant differences in the orientation of the V_L domain relative to the other domains and based on this observation a mechanism for antigen-induced B cell signalling is suggested.

Chapter 2 - Development of Approaches for the Collection of RDC Data from Large Complexes

2.1 Introduction

2.1.1 Partial alignment of proteins in solution

As discussed previously, traditional methods of structure determination of macromolecules by NMR relied on the measurement of semiquantitative short range restraints, most significantly NOEs but also dihedral angles. More recently, dipole-dipole couplings between spin ½ nuclei (^{1}H, ^{13}C, ^{15}N) have been used to obtain structural information about the global orientation of a protein with respect to the magnetic field.

The dipolar coupling, D_{ij}, between a pair of ½ spin nuclei is shown in Equation 2.1.

$$D_{ij} = -\frac{\mu_0 \gamma_i \gamma_j h}{(2\pi r)^3}\left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle$$

Equation 2.1

where r is the distance between a specific pair of nuclei, γ_{ij} are the gyromagnetic ratios for the nuclei, μ_0 is the permittivity of vacuum and h is Planck's constant (Prestegard, 2004). The angle θ relates the inter-nuclear vector between the coupled spins to the static external magnetic field (B_0). The measured dipolar coupling is an average of every value of θ at all orientations and, therefore, in isotropic solution the dipolar

coupling averages to zero. A non-zero value of θ can be observed when there is an anisotropic distribution of orientations relative to the magnetic field (Blackledge, 2005). In solid state NMR the orientation of the molecules are fixed and so effectively all the inter-nuclear vectors have fixed orientations with respect to the magnetic field. This means that all possible dipolar couplings are seen to up to their maximum value and the resolution achievable in solution state NMR is lost. Since residual dipolar couplings are usually measured between a pair of bonded nuclei, the value of *r* is fixed, so dipolar couplings contain information on the orientation of internuclear vectors with respect to the external magnetic field, regardless of where the vector is situated in the protein. Dipolar coupling can be used to provide information on protein orientation, whilst preserving the resolution provided by solution state NMR, by introducing a weak degree of alignment into the sample. If only a fraction of the protein molecules are aligned residual dipolar couplings (RDCs) are seen as a positive or negative contribution to the scalar *J* coupling in NMR spectra (Tolman, 2001).

Early experiments using paramagnetic myoglobin uniformly enriched with ^{15}N showed a clear field dependence of the one bond ^{1}H-^{15}N splitting (Tolman et al., 1995). This finding was also observed in the model diamagnetic protein ubquitin, although much smaller dipolar coupling values were seen in this case (Tjandra et al., 1996). However, for most diamagnetic proteins the degree of alignment caused by magnetic susceptibility anisotropy is too weak for accurate determination of RDC values. It is possible in some circumstances to label proteins with paramagnetic centres which have anisotropic electron spin distributions, for example members of the lanthanide family. Lanthanides can displace a different metal ion (e.g. Ca^{2+}, Mg^{2+} or Mn^{2+}) but these binding sites are quite rare. Recently, the availability of lanthanide tags that can be site specifically

positioned in proteins and oligonucleotides have made this method of protein alignment more viable (Otting, 2008).

Partial alignment of a protein sample in a magnetic field can be conveniently achieved by the addition of an alignment media to the sample. These act as anisotropic barriers, resulting in a non uniform distribution of protein orientation within the sample. A common liquid crystalline media used for partial alignment of proteins is bicelles. These are flat, disk-like structures typically formed from regular saturated phospholipids, most commonly dimyristoyl-phosphatidylcholine (DMPC) and dihexanoyl-phosphatidylcholine (DHPC). In solution, a mixture of these lipids form disk-like structures of 30-40 nm in diameter (Prestegard, 2004) with the DMPC making the flat surfaces and DHPC covering the edges. When they are placed in a magnetic field the bicelles align with their long edge parallel to the magnetic field (Ottiger and Bax, 1998). Bicelles carry no surface charge and therefore induce partial alignment by acting as steric barriers (Figure 2.1A). The properties of bicelles can be altered by varying the DHPC:DMPC ratio which changes the bicelle size or by using other uncharged lipid. Bicelle stability can be increased by the addition of charged amphiphiles, such as CTAB (hexadecyl(cetyl)trimethylammonium bromide) which is positively charged and SDS (Sodium dodecyl sulfate) which is negatively charged, or by using charged lipids to form the bicelles. In addition to improving stability, bicelles bearing a charged surface will change the average partial alignment of proteins in the media as alignment would be induced by both steric barriers and electrostatic repulsion or attraction (Bax et al., 2001). Although bicelles have proved to be a useful system for measuring dipolar couplings, their long term chemical stability is sensitive to pH and the liquid crystalline phase is very dependent on ionic strength and temperature (Ottiger and Bax, 1998).

Liquid crystal phases consisting of filamentous bacteriophage provide a more robust alternative (Figure 2.1B), the most widely used being Pf1 and Fd (Bax, 2003). Pf1 are ~2 μm in length (Hansen et al., 2000), almost twice as long as Fd, however both have a diameter of 6.5 nm (Bax, 2003). As a result of these high aspect ratios, Pf1 can remain in a liquid crystalline phase to very low concentrations (Bax et al., 2001), however, higher concentrations of Fd are required due to its lower diameter to length ratio (Barrientos et al., 2001). Phage form a magnetic field induced, ordered liquid crystalline phase where the long axis of the phage orientates parallel to the magnetic field (Bax et al., 2001). Unlike DHPC/DMPC bicelles, both strains of phage have a negative surface charge at physiological pH and so induce protein alignment through both steric and electrostatic interactions, therefore, phage and bicelle media will generally align the same protein in different orientations. The negative surface charge also makes phage an ideal media for determining RDCs for DNA (MacDonald et al., 2001) and RNA (McCallum and Pardi, 2003). Pf1 phage is favoured as it is relatively inexpensive and the degree of alignment can be easily tuned by varying the phage concentration (Zweckstetter and Bax, 2001).

Another liquid crystalline phase alignment media that has been used for partial alignment of proteins is cetylpyridium halide (Cl^- or Br^-) and hexanol. Cetylpyridium carries a positive charge and so would partially align a protein in a different orientation to negatively charged Pf1 (Bax et al., 2001). Mixtures of n-hexanol and alkylethylene glycol also form liquid crystals in a magnetic field. This carries little net charge, however, unlike bicelles, its components are chemically inert (Bax, 2003). The integral membrane protein, bacteriorhodopsin, surrounded by seven or eight lipids form disk like, two dimensional crystals known as purple membrane (PM) fragments (Koenig et

al., 1999). These biocrystals align in a magnetic field and are stable over a wide range of pH values, temperatures and ionic strengths (Sass et al., 1999) making PM another alternative media for partial protein alignment.

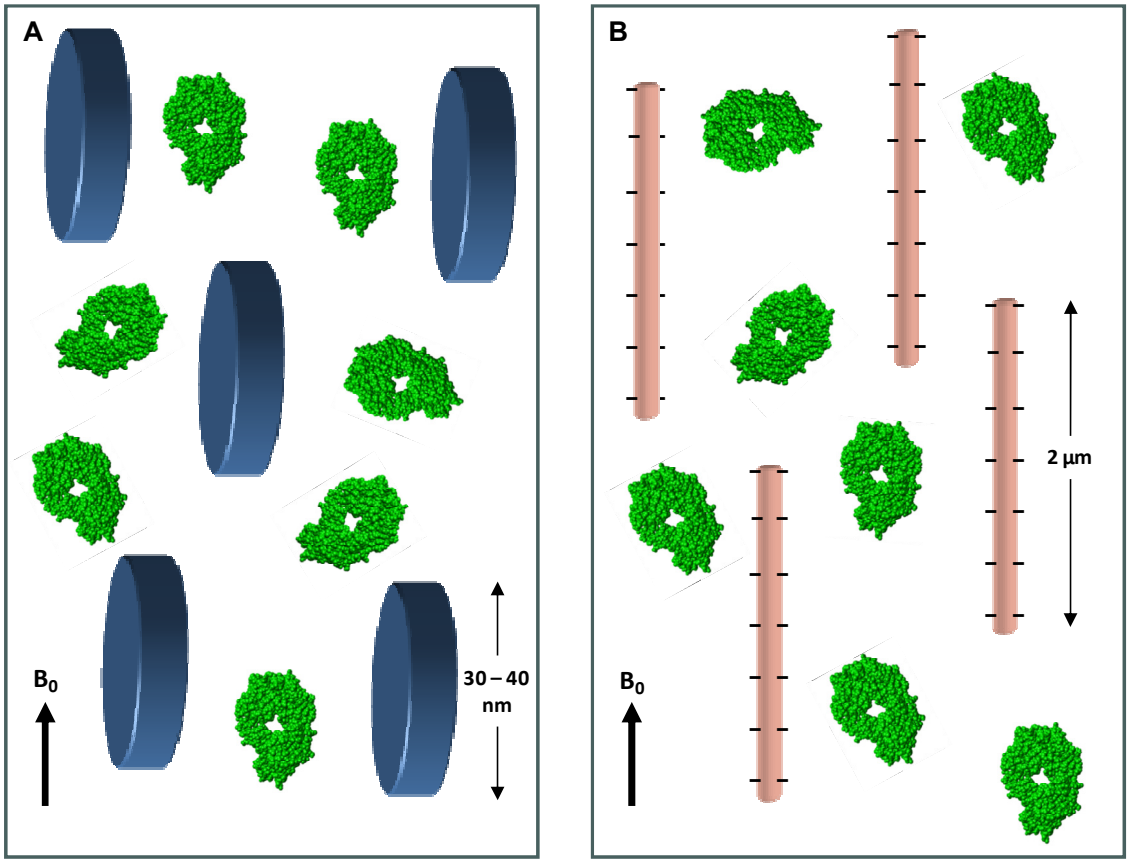


Figure 2.1

Partial alignment of a protein by liquid crystal alignment media. A schematic representation of partial protein alignment in bicelles (A) and phage (B) both of which form a liquid crystalline phase in an external magnetic field (B_0). Phage and bicelles partially align proteins in different orientations as phage particles have a negative surface charge resulting in alignment through both steric and electrostatic modes.

Partial alignment of proteins can also be achieved independently of the magnetic field using strained polyacrylamide gels (Ishii et al., 2001) where anisotropy is induced by compressing or stretching a gel away from its natural isotropic state (Tycko et al., 2000). As a result of the strain, cavities in the gel develop a slightly oblate shape, the orientation of which depends on the direction of the strain. Protein is allowed to diffuse into these cavities and will preferentially partially align in an orientation that is dependent on the shape of the stretched or compressed cavity.

An optimal degree of protein alignment in any media gives RDCs that are scaled down by about 10^3 relative to their static values, which is achieved by adjusting the concentration of alignment media in the sample. Stronger alignment can result in broadening of ^{1}H multiplets resulting from unresolvable homonuclear dipolar couplings. If there is an intention to use the RDC values to determine the orientation between two protein domains, it is useful to optimise protein alignment in more than one media. This is due to the non-unique nature of the angular dependence of the residual dipolar coupling function:

$$\left(\frac{3\cos\theta - 1}{2}\right)$$

Equation 2.2

where θ is the angle between the internuclear vector and the magnetic field. Therefore, a single dipolar coupling measurement can only be confined to two cones of orientation, meaning that relative fragment orientation cannot be uniquely determined from a single set of RDCs (Al-Hashimi et al., 2000). A large RDC data set can restrict the possible

solutions, however, to determine a unique orientation it is necessary to collect data in more than one alignment media (Prestegard, 2004).

The average orientation of a partially aligned protein with respect to the magnetic field is described by the alignment tensor. The magnitude and direction of the alignment tensor is described within the eigenframe of the alignment tensor by three eigenvalues A_{xx}, A_{yy} and A_{zz} where conventionally $A_{xx} \leq A_{yy} \leq A_{zz}$. The axial component of the alignment tensor (A_a) = $A_{zz}/2$ and the rhombic component (A_r) = (1/3)(A_{xx} - A_{yy}). The orientation of the alignment tensor with respect to the coordinate frame of the molecule is defined as a three-dimensional rotation around three Euler angles (α, β, γ).

The measured coupling (D_{ij}) is related to the alignment tensor as described in Equation 2.3.

$$D_{ij} = \frac{\gamma_i \gamma_j \mu_0 h}{16\pi^2 r^3_{ij,eff}} \left[A_a (3\cos^2\theta - 1) + \frac{3}{2} A_r \sin^2\theta \cos 2\phi \right]$$

Equation 2.3

where γ_i and γ_j are the gyromagnetic ratios, μ_0 is the permeability of a vacuum, h is Plank's constant, $r^3_{if,eff}$ is the effective internuclear distance, A_a is the axial component of the alignment tensor, A_r is the rhombic component of the alignment tensor, θ and ϕ are polar angles describing the internuclear vector with respect to the frame of the alignment tensor (Figure 2.2) (Blackledge, 2005).

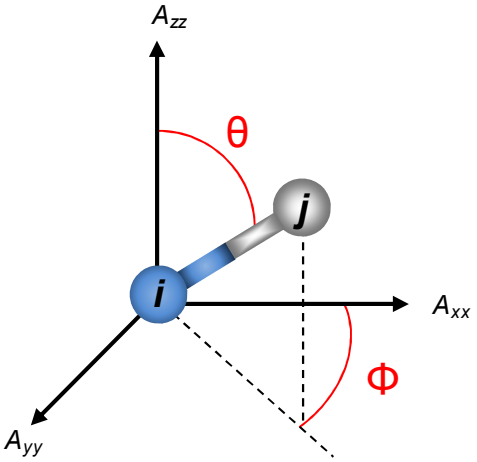


Figure 2.2

Representation of the internuclear vector within the frame of the alignment tensor. The principal axes of the alignment tensor are labelled as A_{xx}, A_{yy}, A_{zz}. The internuclear vector between *i* and *j* is described within the principal axis frame by the polar angles θ and ϕ.

2.1.1.1 Applications of RDCs

RDCs can be used as an additive set of restraints in structural refinement as they provide important information about the global orientation of proteins which complement the short-range information provided by NOEs, *J* coupling constants and chemical shifts. Significant improvements in accuracy were seen when sets of one bond ^{15}N-^{1}H and ^{13}C-^{1}H dipolar coupling were used in addition to conventional NMR restraints for structural refinement of a complex of the DNA-binding domain of the transcription factor GATA-1 with a 16 base pair oligodeoxyribonuclotide (Tjandra et al., 1997). RDCs can also be used to solve the backbone structure of a protein directly,

using a molecular fragment replacement technique. This was demonstrated for ubquitin for which four backbone couplings (N-H, CO-N, CO-HN, Cα-Hα) have been determined in two different alignment media. The best fit between the RDCs measured from 7 residue fragments from ubiqitin and 7 residue fragments from the PDB were determined and the protein structure was calculated by assembling the fragments and superimposing overlapping fragments. Further refinement using chemical shift data gave a 3D structure with a backbone rmsd of 0.88 Å when compared to a previously determined crystal structure (Delaglio et al., 2000). RDCs have also proved to be a useful tool to validate structures of protein (Cornilescu et al., 1998) and protein-protein complexes (Bewley, 2001) that have been previously solved by other means. The accuracy of a solved structure can be determined by observing the agreement between experimentally determined RDCs and RDCs calculated from the known structure. The relative ease of RDC data collection and analysis provides a role for them in rapid classification of protein fold families. Backbone N-H RDCs can be used to determine structural similarity between unknown proteins and structures from a database, without the need for backbone assignment (Valafar and Prestegard, 2003). RDCs can be used to differentiate between monomeric and homodimeric states. For example, crystal structures show that cyanovirin-N exists as a domain swapped dimer, however, RDCs calculated from the alignment tensor of the monomeric protein correlate well with the experimentally determined RDCs whilst those calculated from the dimeric show poorer correlation. Hence the RDC data provides strong evidence to suggest that protein is monomeric in solution (Zweckstetter and Bax, 2000). NMR is an increasingly important technique in screening for ligands that bind to a specific potential therapeutic target. RDCs can provide information to rapidly determine the average ligand orientation with

respect to the protein, which could be used to design modified ligands with improved binding properties (Lipsitz and Tjandra, 2004).

Structural determination of proteins that consist of several well folded domains is often complicated due to the intrinsic flexibility of the systems. Crystal packing forces often result in crystal structures that show only a snapshot of one possible conformation of the flexible molecule and NMR structures are limited by the availability of cross domain ^{1}H-^{1}H NOEs (Fushman et al., 2004). RDCs can be used to determine interdomain orientations due to the unique, long-range orientational information they provide. When a multidomain protein is partially aligned, both domains are aligned to the same extent and therefore have a common alignment tensor. RDC data collected from each domain is used to obtain an alignment tensor for each of the domains individually which are then rotated until the tensors coincide. This approach was used for a two domain fragment of the barley lectin protein, as previous attempts were unable to conclusively determine the relative orientations between the B and C domains as very few interdomain NOEs were observed. Backbone N-H dipolar couplings were collected from the fragments using two types of media that aligned the protein in different orientations. Calculations of the alignment tensor from each domain showed that the relative domain orientations were different from the domain orientation seen in a crystal structure of a homologous protein and proposed a possible solution where alignment tensors determined for the domains in each alignment media coincided (Fischer et al., 1999). Similarly, RDCs can be used alongside other restraints to determine the structures of protein-protein complexes. The structures of protein-protein complexes can be determined using only RDC restraints and chemical shift perturbation data providing that high resolution structures of the unbound proteins are available and that

they do not undergo significant backbone rearrangements on complex formation. This was demonstrated by the determination of the EIN-HPr complex which was determined from crystal structures of unbound EIN and HPr using only RDCs and chemical shift perturbations as restraints in docking. The resulting complex structure was comparable with a NOE-based structure previously determined by NMR (McCoy and Wyss, 2002).

2.1.1.2 RDC data acquisition

Measurement of dipolar couplings has predominantly focused on one bond ^{1}H-^{15}N ($^{1}D_{HN\text{-}N}$) RDCs as the HN-N internuclear distances are quite short and invariant. This means that sizeable dipolar interactions are observed for directly bonded spins and the values can, therefore, be interpreted simply terms of angular constraints with no contribution from variations in internuclear distance (Prestegard, 2004). Various techniques have been developed for the measurement of dipolar couplings. The *J* coupling observed for data measured in the partially aligned sample represents the sum of the scalar and dipolar contributions; however, since variation can exist in the scalar coupling it is also necessary to collect data for the protein in both the isotropic and aligned environments (Bax et al., 2001). Initially one-bond $^{1}D_{HN\text{-}N}$ dipolar couplings were obtained from ^{1}H decoupled HSQC spectra, however, this method of data collection resulted in crowded spectra and resonance overlap, limiting the applicability of this method for larger proteins. The IPAP (in-phase and anti-phase) method reduces spectral overlap by separating the F_1 coupled HSQC into two separate spectra. This is achieved by acquiring two separate, interleaved experiments, one which yields spectra with in-phase doublets and the other with anti-phase doublets. Addition/subtraction of the spectra gives two individual spectra each containing one component of the doublet (Ottiger et al., 1998). Assuming that the *J* coupling does not change between the

isotropic and the partially aligned states, RDCs can be determined from the difference in splitting between the two states. The linewidth of the upfield component relative to the downfield component is broadened at higher magnetic field strengths due to CSA/dipole relaxation interference. This linewidth broadening and loss of sensitivity limits the precision of RDC measurement. Linewidth broadening also increases with molecular mass due to the increased rotational correlation times of larger proteins which further contributes to the imprecision of RDC determination using this method (Cavanagh et al., 1996). An alternative approach for determining RDCs for large proteins and complexes can be obtained from HSQC and TROSY spectra acquired for both aligned and unaligned samples. The separation between the HSQC and TROSY peaks is equal to half the *J* coupling plus half the RDC (($^1J_{NH}$ +D_{NH}))/2). The value of the RDC is the difference in splitting between the corresponding peaks in the aligned and unaligned sample. The linewidths in these spectra are significantly narrower than in the IPAP spectra and so RDCs can be measured with higher precision (Kontaxis et al., 2000). To further reduce spectral crowding when collecting data for very large proteins and protein-protein complexes it is also possible to measure $^1J_{NH}$ from a 3D HNCO spectrum without ^{1}H decoupling in the ^{15}N dimension (Cavanagh et al., 1996). In addition to $^1D_{HN-N}$, a variety of experimental methods have been developed to collect $^1D_{C-H}$, $^1D_{N-C'}$, $^1D_{N-C\alpha}$, $^1D_{C\alpha-C'}$, $^1D_{C\alpha-C\beta}$, $^2D_{N-C\alpha}$, $^2D_{HN-C'}$, $^2D_{HN-C\alpha}$, and $^3D_{HN-C\alpha}$ RDCs in ^{15}N/^{13}C or ^{15}N/^{13}C/^{2}H-labelled proteins.

RDCs provide valuable information about the global orientation of proteins in solution and have clear applications as restraints in obtaining the structures of protein-protein complexes using docking methods. The work in this chapter describes an optimised method for determining RDCs from IL-1β/Fab complexes. Initially the properties of Pf1

phage and bicelle alignment media were investigated and the determination of RDCs for the model protein IL-1β was optimised in both media. The technology was then applied to the IL-1β/Fab complex and RDCs were determined for two complexes in phage alignment media. The affect of protein backbone mobility on the agreement between experimentally determined RDCs and RDCs calculated from a known structure was also considered.

2.2 Materials and Methods

2.2.1 NMR sample preparation

2.2.1.1 Preparation of bicelle samples

A blend of synthetic phospholipids for bicelle preparation was purchased from Avanti Polar Lipids, Inc. The molecular composition was a 3:1 molar ratio of dimyristoyl-phosphatidylcholine (DMPC) and dihexanoyl-phosphatidylcholine (DHPC), premeasured into 50 mg aliquots. A 15% stock solution was made by the addition of 280 µl of 100 mM sodium chloride, 25 mM sodium phosphate, 0.02% sodium azide, 10% D_2O buffer at pH 6.5 to the 50 mg DMPC/DHPC mixture and this was left to hydrate at 18 ^{o}C for 3 hours. To ensure sample homogeneity, the mixture was cooled on ice for 2 minutes until a clear, free flowing liquid consistency was seen and then immediately placed into a preheated heat block at 40 °C for 2 minutes, during which time the solution became very viscous. This cycle was repeated three times to ensure sample homogeneity (Ottiger and Bax, 1998). The bicelle solution was stored on ice to use immediately for NMR sample preparation or stored at -20 °C. The ^{2}H quadrupolar splitting of the HDO signal was used to monitor bicelle temperature stability were acquired from a 0.35 ml sample containing 5% bicelles in 100 mM sodium chloride, 25 mM sodium phosphate, 0.02% sodium azide, buffer at pH 6.5. Backbone amide RDCs were determined from samples containing 0.50 mM ^{15}N/^{2}H IL-1β containing 5% bicelles in the same buffer as above. NMR samples containing bicelles were kept on ice until being placed directly into a probe preheated to 35 °C. The samples required at least 30 minutes to equilibrate before data could be recorded.

2.2.1.2 Pf1 phage NMR sample preparation

Pf1 phage was purchased from ASLA Ltd., Riga, Latvia. The 50 mg/ml phage stock solution was dialysed into 100 mM sodium chloride, 25 mM sodium phosphate, 0.02% sodium azide, buffer at pH 6.5 and the phage concentration was confirmed to be approximately 50 mg/ml by UV absorbance at 270 nm using an extinction coefficient of 2.25 cm^{-1} mg^{-1} (Hansen et al., 2000). Positive displacement pipettes were used to dispense the phage samples, which have a high viscosity. NMR experiments used to trial phage alignment properties were acquired from 0.35 ml samples of 0.50 mM ^{15}N IL-1β containing either 2.5, 5 or 7.5 mg/ml Pf1 phage in 100 mM sodium chloride, 25 mM sodium phosphate, 0.02% sodium azide, buffer at pH 6.5. RDC data was acquired for complexes using 0.35 ml samples of 0.4 mM $^{15}N/^{2}H$ IL-1β and 0.45 mM g168/gIC8 Fab containing 4 mg/ml Pf1 phage in the same buffer as above. NMR samples containing phage required 30 minutes to equilibrate in the spectrometer.

2.2.1.3 Preparation of unaligned samples

NMR data were acquired from 0.35 ml samples of 0.5 mM $^{15}N/^{2}H$ in the same buffer as above to obtain appropriate spectra of non aligned samples. Similarly, NMR data were acquired for unaligned, 0.35 ml samples of 0.4 mM $^{15}N/^{2}H$ IL-1β bound to gIC8 and g168 Fabs in the same buffer.

2.2.2 NMR spectroscopy

NMR spectra were acquired on a 500 MHz Bruker DRX, 600 MHz Bruker DRX, 600 MHz Bruker Avance, an 800 MHz Bruker Avance II and a 950 MHz Bruker Avance III spectrometer as appropriate. All experiments were collected at 35 ^{o}C. ^{15}N IPAP-HSQC (Ottiger et al., 1998) experiments were acquired to determine the ^{1}H-^{15}N backbone

RDCs of free IL-1β. Spectra were acquired for both isotropic and partially aligned IL-1β samples. Typical acquisition times were 80 ms in F_2 (^{1}H) and 200 ms in F_1 (^{15}N) and acquisitions were over 16 hours. Backbone amide RDC values for the IL-1β/Fab complex samples were derived ^{15}N/^{1}H HSQC and TROSY spectra (Kontaxis et al., 2000) collected from both isotropic and aligned samples. Acquisition times were 60 ms in F_2 (^{1}H) and 50 ms in F_1 (^{15}N) and experiments were collected for between 5 and 12 hours. The WATERGATE method was used to suppress the water signal when required (Piotto et al., 1992). All NMR data was processed using Topspin (Bruker Biospin Ltd.). Spectra were zero filled to 2048 points in the proton dimension and 4096 points in the nitrogen dimension.

2.2.3 Analysis of RDC data

NMR spectra were analysed using Sparky (Kontaxis et al., 2000). A complete set of backbone resonance assignments had been previously made for free IL-1β and IL-1β in complex with gIC8 by a colleague (Dr Vaclav Veverka) and the assignments for the IL-1β/g168 complex were made using the spectra of free IL-1β and only assigning residues that were not significantly perturbed by g168 binding. RDCs were only derived from well resolved peaks which were not excessively broadened. Backbone amide RDCs were derived from free IL-1β by calculating the difference in the $^1J_{NH}$ splittings measured in IPAP-HSQC spectra in isotropic and aligned samples. Backbone amide RDC values for the IL-1β/Fab complex samples were derived from the differences between the ^{15}N/^{1}H scalar couplings between the isotropic and partially aligned samples using ^{15}N/^{1}H HSQC and TROSY spectra (Kontaxis et al., 2000). The derived RDCs were analysed using PALES (Prediction of AlignmEnt from Structure) (Zweckstetter and Bax, 2000). Previously reported IL-1β structures used for alignment tensor

determination were the X-ray structure 2I1B (Priestle et al., 1989) and the NMR structure 6I1B (Clore et al., 1991).

2.3 Results and Discussion

2.3.1 Properties of DMPC:DHPC bicelles

The temperature stability of samples from a 5% solution of bicelles containing a 1:3 molar ratio of DHPC to DMPC was monitored by observing the ^{2}H quadrupolar splitting of the HDO resonance, which arises from the exchange between isotropic HDO and aligned HDO molecules associated with the bicelles (Lipsitz and Tjandra, 2004). Samples were kept in ice until being inserted into a probe preheated to 38 °C. The sample was allowed to equilibrate for at least 30 minutes before a ^{2}H spectrum was acquired, giving a doublet splitting of 9.4 Hz. The temperature of the spectrometer was then gradually increased or decreased and experiments were acquired over a range of temperatures. The ^{2}H splitting observed is shown in Figure 2.3 which illustrates that the phase was stable between 26 and 48 °C. The ^{2}H splitting increases with temperature until the liquid crystalline phase collapses at ~50 °C, indicated by the presence ^{2}H singlet in the D_2O spectrum and white flakes in the sample.

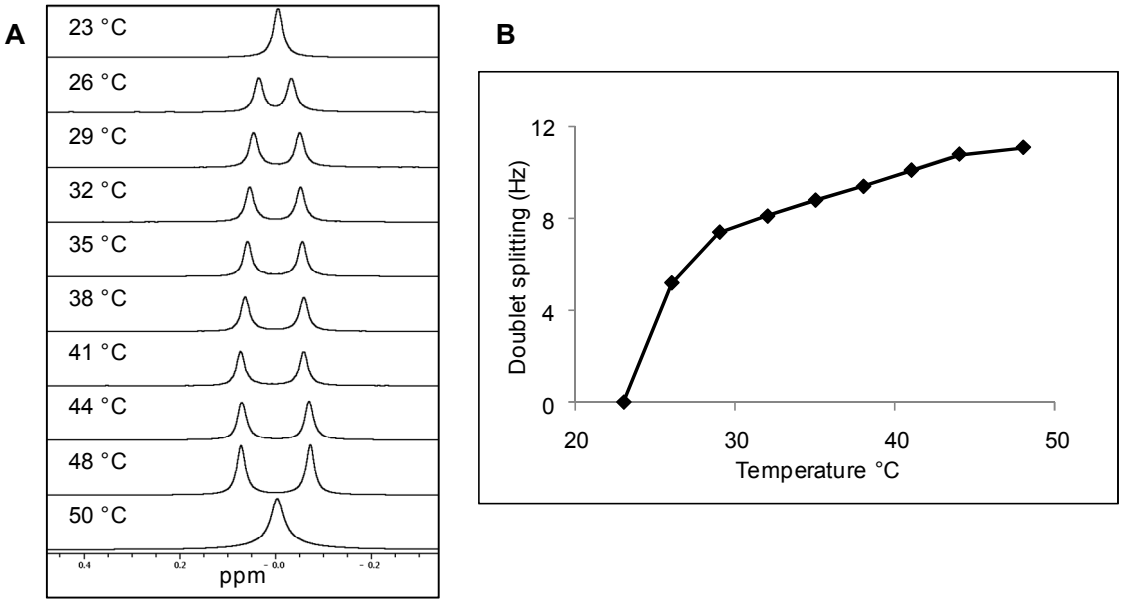


Figure 2.3

Stability of phospholipid bicelles over increasing temperatures. The stability of a 1:3 mixture of DHPC to DMPC was monitored by NMR at 500 MHz using the splitting of the deuterium signal. NMR spectra for samples containing 5% phospholipid bicelles were initially recorded at 38 °C The temperature was gradually increased or decreased and a spectra was recorded at each temperature increment. The spectra from the 1D deuterium observe experiments are shown in A with the values of the doublet splitting shown in B.

In a magnetic field, bicelles make a transition from being randomly ordered into a nematic phase liquid crystal, which has orientational ordering. This means that the bicelles are disordered with respect to one another but their long molecular axes statistically point in one direction, in this case parallel with the external magnetic field (Figure 2.1). Previously reported ^{31}P spectra collected from bicelles with the same

DHPC:DMPC ratio show two resonances when the liquid crystalline phase is observed, one corresponding to DHPC and the other to DMPC. The changes in the ^{31}P spectra were also observed as a result of gradual temperature increase. At the point of liquid crystalline phase collapse, a narrow peak corresponding to isotropic DHPC is seen and the DMPC resonance strongly upfield shifted which is indicative of highly ordered DMPC. Since no splitting is observed in the ^{2}H solvent signal it has been suggested that the DMPC forms a hexagonal phase, which is characterised as a hexagonal arrangement of long cylindrical rods of amphiphilic molecules and this is visible as the white particles in the sample (Ottiger and Bax, 1998). As the temperature was decreased the ^{2}H splitting reduced, until an isotropic phase was induced at ~23 °C, characterised by a ^{2}H singlet peak. The induction of an isotropic phase is confirmed by the ^{31}P spectrum which shows one narrow resonance at the isotropic shift position showing that the DHPC and DMPC have no order and are randomly tumbling in solution (Ottiger and Bax, 1998).

NMR data for ^{15}N-labelled IL-1β were collected at 35 °C from a sample containing 5% 1:3 DHPC to DMPC bicelle solution. The difference between the upfield and downfield peaks of IPAP-HSQC spectra of aligned and unaligned IL-1β were used to determine backbone amide RDCs at 500, 600 and 800 MHz (Figure 2.4). The values of the measured RDCs range from about -10 Hz to 12 Hz and were unaffected by field strength of between 500 and 800 MHz confirming that bicelle alignment is independent of magnetic field strength. All dipolar couplings were measured in the indirect ^{15}N dimension, as transverse relaxation is considerably slower in the ^{15}N dimension compared to the ^{1}H dimension, resulting in sharper signals in the spectra, improving the accuracy by which the RDC values can be determined.

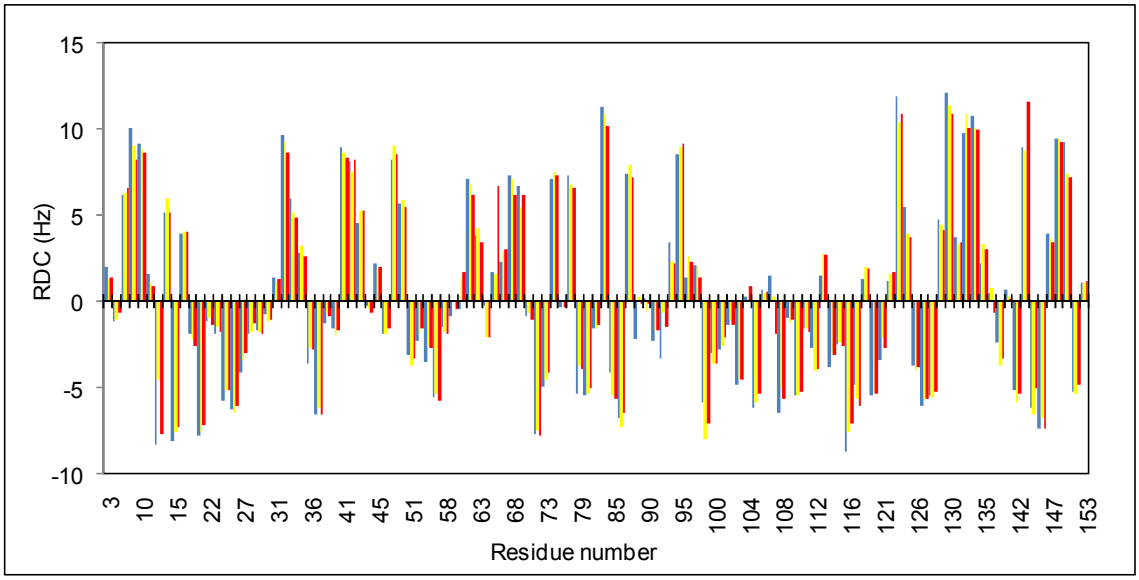


Figure 2.4

Alignment of phospholipid bicelles is independent of magnetic field strength. The histogram shows the magnitude of ^{1}H-^{15}N dipolar couplings collected from a sample of ^{15}N-labelled IL-1β containing a 5% mixture of 1:3 DHPC to DMPC. RDCs were collected at 500 MHz (blue), 600 MHz (yellow) and 800 MHz (red). The errors on RDCs determined from IL-1β were estimated to be 1.0 Hz, based on a conservative estimate of the peak position in the processed spectra.

Bicelles form a nematic phase when they are orientated by the external magnetic field. Alignment can be induced at magnetic fields of ~2 T (Hare et al., 1995), so in the magnetic fields described here the bicelles are fully nematic and therefore an increase in magnetic field strength would not be expected to induce an increase in alignment. The range of values of the measured RDCs are consistent with values measured for other proteins at the same DHPC/DMPC ratio and lipid concentration of 5%. RDCs values of about -15 to +12 Hz were measured for ubquitin (8.6 kDa) at 38 °C (Tjandra, 1997),

values of about -20 to +15 Hz were measured for putidaredoxin (11 kDa) at 39 °C (Jain et al., 2005) and values of about -15 to +12 Hz were seen for a PBX-DNA complex (20 kDa) at 37 °C (Sprules et al., 2003). The literature values are all slightly larger than the RDCs observed for IL-1β, most probably because they were recorded at higher temperatures and, as shown previously, the degree of bicelle alignment is increased at higher temperatures.

2.3.2 Properties of Pf1 filamentous phage alignment media

Partial alignment of IL-1β was also achieved using Pf1 filamentous phage purchased from Asla Biotech. The degree of protein alignment in phage solution can be optimised and is dependent on phage concentration, salt concentration and magnetic field strength (Hansen et al., 2000; Zweckstetter and Bax, 2001). Pf1 phage samples have been reported to show only a modest change in splitting of the ^{2}H D_2O signal between 10 and 45 °C and NMR data can be collected at temperatures between 5 and 45 °C (Hansen et al., 2000). IPAP-HSQC spectra were collected at 600 MHz from samples of ^{15}N-labelled IL-1β containing phage at concentrations of 2.5 mg/ml, 5.0 mg/ml and 7.5 mg/ml and also an unaligned reference sample. Figure 2.5 shows a small section from the IPAP spectra of IL-1β at increasing phage concentrations, focusing on residue H30. At 2.5 mg/ml phage small RDCs ranging between -2.6 and 2.7 Hz could be measured. Increasing the phage concentration to 5 mg/ml in the sample gave dipolar couplings in the range of -7.4 to 11.1 Hz and at phage concentrations of 7.5 mg/ml RDCs of between -25.1 and 36.5 Hz were measured (Figure 2.6).

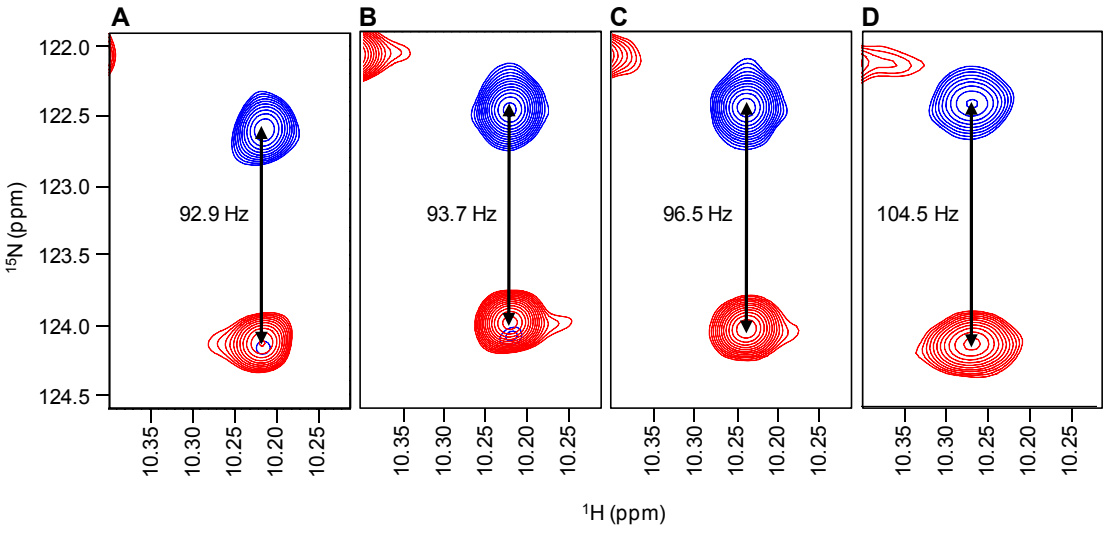


Figure 2.5

The magnitude of ^{1}H-^{15}N RDCs increase with Pf1 phage concentration. Sections of overlaid IPAP-HSQC spectra showing the separation between the upfield (blue) and downfield (red) components of the peak corresponding to histidine 30 of IL-1β. Panel A shows a section from the isotropic spectrum and spectra collected from samples containing 2.5, 5 and 7 mg/ml Pf1 are shown in B-D respectively. The measured RDC is the difference between the coupling seen in aligned and isotropic samples. In this case the measured RDC for the sample containing 2.5 mg/ml phage is only 0.8 Hz, for the 5 mg/ml sample is it 3.6 Hz and 11.6 Hz for the sample containing 7.5 mg/ml phage.

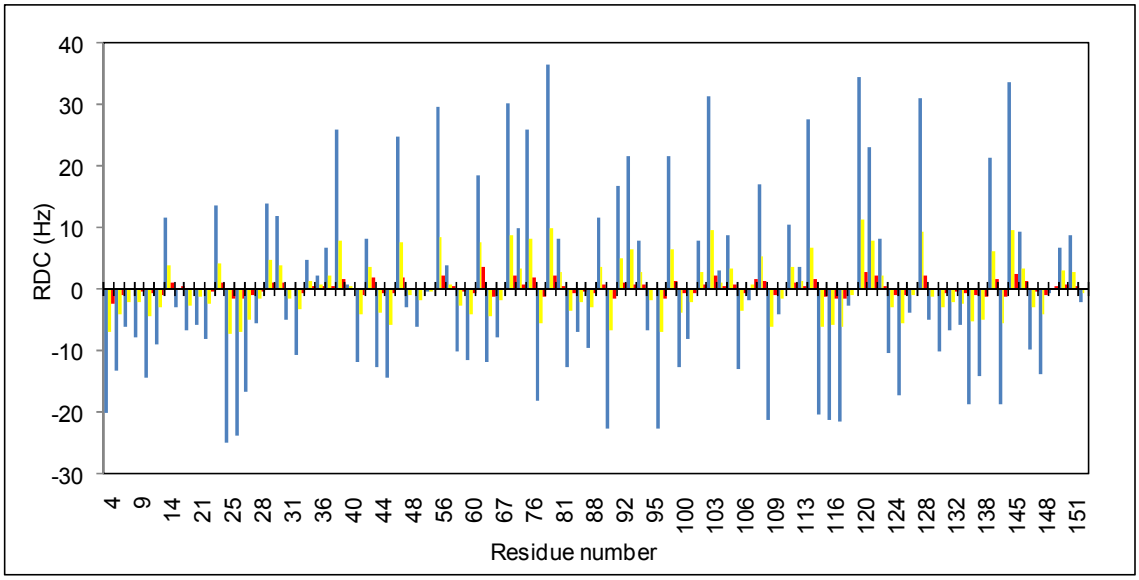


Figure 2.6

Magnitude of ^{1}H-^{15}N RDCs measured from IL-1β increase with Pf1 phage concentration. A histogram showing ^{1}H-^{15}N RDCs measured at 600 MHz from IL-1β in samples containing 7.5, 5 and 2.5 mg/ml phage in blue, yellow and red respectively. The errors on RDCs determined from IL-1β were estimated to be 1.0 Hz, based on a conservative estimate of the peak position in the processed spectra.

A representative set of well resolved peaks were used to calculate the average backbone amide ^{1}H linewidths in the IPAP-HSQC spectra of IL-1β at each phage concentration. Addition of 5 mg/ml phage to an isotropic sample increased the average line widths by approximately 2 Hz, however, the proton linewidths were 1.8 times greater in the sample containing 7.5 mg/ml phage compared to the 5 mg/ml phage sample. The magnitude of alignment of Pf1 phage at low concentrations is reported to be dependent on the strength of the magnetic field (Hansen et al., 2000). RDCs determined from a ^{15}N/^{2}H IL-1β sample containing 5 mg/ml phage, collected at 600 and 950 MHz are

shown in Figure 2.7. On average, RDCs measured at 950 MHz were nearly two fold greater than collected for the same residue at 600 MHz.

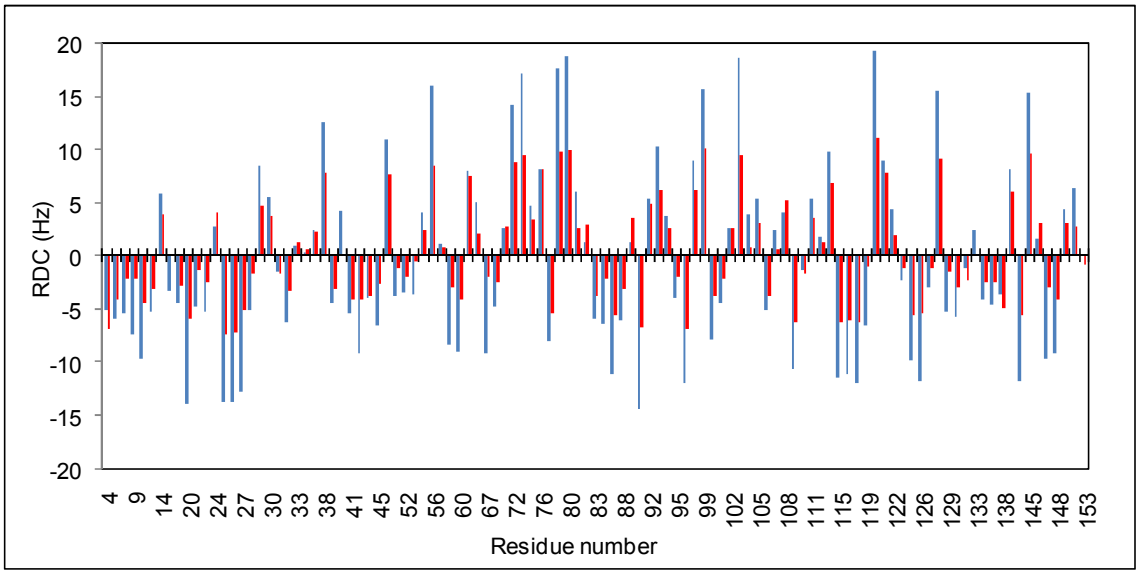


Figure 2.7

Alignment of filamentous Pf1 phage is dependent on magnetic field strength. The histogram shows the magnitude of ^{1}H-^{15}N dipolar couplings obtained from a sample of ^{15}N IL-1β containing 4 mg/ml pf1 phage. RDCs were collected at 950 MHz (blue) and 600 MHz (red). The errors on RDCs determined from IL-1β were estimated to be 1.0 Hz, based on a conservative estimate of the peak position in the processed spectra.

Phage is predicted to form a fully nematic liquid crystalline phase at concentrations of 10-20 mg/ml, however the nematic threshold is dependent on the salt concentration in the sample. At lower phage concentrations or in samples containing higher salt concentrations the phage exist in a so called paranematic phase, which is transition period between isotropic and nematic phases. Below the nematic phase threshold, Pf1

alignment is dependent on the strength of the magnetic field, however, above the nematic threshold phage alignment becomes independent of magnetic field strength (Zweckstetter and Bax, 2001). The results shown here suggest that in samples containing 5 mg/ml Pf1 and 100 mM NaCl, the phage exist in a paranematic phase as the degree of alignment is dependent on the magnetic field strength.

Pf1 phage have a length of about 2 μm and a diameter of 6.5 nm. This high aspect ratio allows for these particles to remain in a liquid crystalline phase to very low concentrations as demonstrated here (Bax et al., 2001). As mentioned previously (Equation 1.1), increase in solvent viscosity increases the rotational correlation time of the protein and consequently decreases the spin-spin relaxation time (T_2), resulting in broadening of spectral linewidths. The optimal phage concentration is that which results in measurable RDCs with minimal increase in solvent viscosity, and hence line broadening, which might limit the accuracy with which the RDCs can be determined.

A comparison of RDCs determined for the same residues in IL-1β in bicelle and phage alignment media is shown in Figure 2.8. There is clearly no correlation between the values obtained, showing that bicelles and phage partially align IL-1β in different orientations. Alignment tensors calculated for IL-1β using the PALES software (Prediction of AlignmEnt from Structure) gave alignment tensor values of D_a = 5.40 Hz and R = 0.45 when bicelle alignment media was used and D_a = 5.27 Hz and R = 0.28 when aligned in phage media.

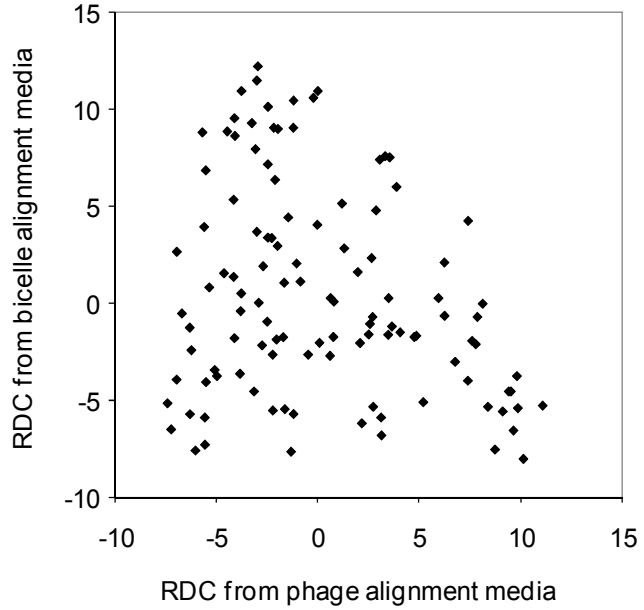


Figure 2.8

Bicelles and phage align IL-1β in different orientations. The ^{1}H-^{15}N RDC values collected from a ^{15}N-labelled IL-1β sample containing 5% phospholipid bicelles are plotted against RDCs for the corresponding residues measured from a sample containing 5 mg/ml Pf1 phage. Both data sets were collected at 800 MHz. The lack of correlation between the RDC values shows that IL-1β has different orientations when partially aligned in the two media.

The difference in direction and magnitude of alignment of IL-1β between phage and bicelle is expected, due to the different in properties of the two media. Bicelles have no net surface charge and so act only as steric barriers, preventing the random protein rotation. In contrast, phage have a negative surface charge and so induce partial alignment sterically and also by electrostatic interactions (Bax et al., 2001).

2.3.3 Determination of backbone amide RDCs for large IL-1β/Fab complexes (~ 65 kDa)

The IPAP-HSQC experiment is restricted to determination of RDCs for proteins and complexes with a molecular weight of less than 30-40 kDa (Tate et al., 2004). In the IPAP-HSQC spectra for larger proteins, interference between the ^{15}N chemical shift anisotropy and the ^{15}N-H dipolar coupling relaxation mechanism result in a significant difference in linewidth between the doublet components. Whilst the downfield doublet component remains relatively narrow, extreme line broadening is seen in the upfield component, which increases with greater magnetic field strengths. The precision with which $^{1}D_{N-H}$ RDCs can be measured is dependent on the accuracy with which the frequency of the broad upfield component can be determined if indeed it can be resolved at all (Kontaxis et al., 2000).

Two unlabelled Fab proteins with high affinity for IL-1β, known as gIC8 (M.W. 48 kDa) and g168 (M.W. 47 kDa), were provided though a collaboration with UCB-Celltech and used to form complexes with $^{15}N/^{2}H$ IL-1β for NMR analysis. IPAP-HSQC experiments collected for the IL-1β/Fab complexes showed very severe line broadening of the upfield component and consequently the dipolar couplings for the complexes were obtained from the difference in backbone amide signal positions in TROSY spectrum and ^{1}H decoupled HSQC spectrum. HSQC and TROSY spectra were collected for both the isotropic and partially aligned complexes. Backbone amide RDC values for IL-1β in the complexes were derived from differences between the $^{15}N/^{1}H$ scalar couplings in isotropic and partially aligned samples (Kontaxis et al., 2000) as shown in Figure 2.9.

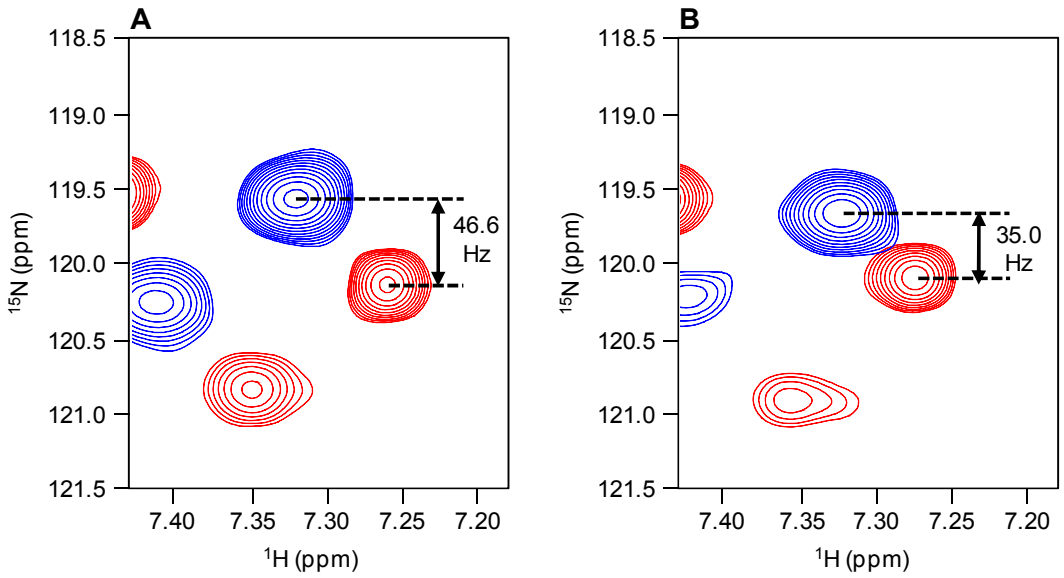


Figure 2.9

Measurement of backbone amide RDCs from ~65 kDa complexes. For the large IL-1β/Fab complexes, backbone amide RDCs were obtained by collecting HSQC and TROSY spectra for both the aligned and unaligned samples. An overlay of a section from the HSQC (blue) and TROSY (red) spectra collected from an unaligned sample of $^{15}N/^{2}H$ IL-1β/gIC8 is shown in panel A. The same section from the spectra collected from a sample of the complex containing 5 mg/ml Pf1 phage is shown in panel B. Backbone amide RDC values were derived from the differences in $^{15}N/^{1}H$ scalar couplings between the isotropic and partially aligned samples. The separation between the HSQC and TROSY peaks is equal to half the *J* coupling plus half the RDC (($^{1}J_{NH}$ + $^{1}D_{NH}$))/2). The value of the RDC is therefore the difference in separation between the corresponding peaks in the aligned and unaligned sample. In this example for residue E25 RDC = (2x35.0) – (2x46.6) = -23.2 Hz.

Optimised partial alignment of the complex samples was achieved using 4 mg/ml Pf1 phage and NMR data were collected on an 800 MHz spectrometer. The RDC values obtained for the complexes are summarised in Figure 2.10. A complete set of backbone resonance assignments had been previously made for IL-1β in complex with gIC8 by a colleague (Dr Vaclav Veverka) and backbone amide RDCs were determined for 102 of the 145 non-proline residues in the protein with values ranging from -37.4 to 38.9 Hz (Figure 2.10A). Backbone assignments were not available for IL-1β in complex with g168, so HSQC and TROSY spectra from the isotropic and partially aligned spectra were assigned using the spectra obtained for free IL-1β. A proportion of the resonances in the spectra were perturbed on antigen binding which, combined with spectral overlap and line broadening as a result of addition of phage, meant that RDCs could only be determined for 50 of the backbone amide proton peaks with a high level of confidence. The values of the determined RDCs range between -40.9 and 32.6 Hz and are shown in Figure 2.10B.

Examples of published values of RDCs determined from phage alignment media demonstrate how much the degree alignment can vary between different proteins. Hemoglobin gives RDCs of about -20 to +15 Hz when aligned in 10 mg/ml Pf1 in 100 mM sodium phosphate buffer (Lukin et al., 2003), while RDCs determined from calmodulin are about -20 to +20 Hz in 15 mg/ml Pf1 alignment media in 100 mM sodium phosphate buffer (Chou et al., 2001) and ubquitin has RDCs of about -22 to +22 Hz when aligned in 5 mg/ml Pf1 in 50 mM NaCl buffer (Clore and Schwieters, 2004). That approximately the same alignment is achieved for these different proteins despite large difference in Pf1 concentration demonstrates that, unlike bicelles, the negative surface charge on phage causes alignment to vary significantly between proteins. The

effect of salt concentration in the sample buffer also contributes to the difference in alignment.

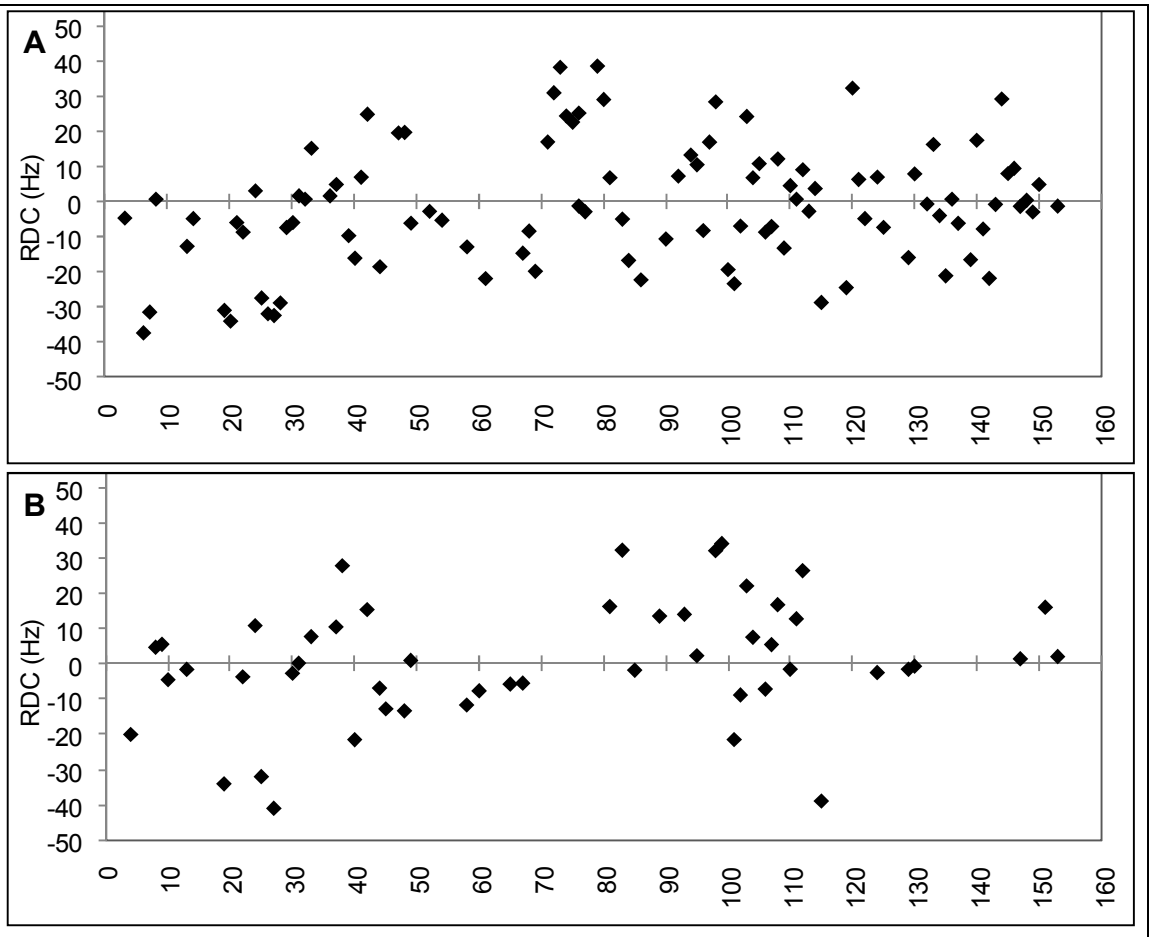


Figure 2.10

A summary of backbone amide RDCs obtained from IL-1β/Fab complexes. The ^{1}H-^{15}N RDCs determined for ^{15}N/^{2}H IL-1β in complex with unlabelled g168 Fab are shown in panel A with the RDCs collected from the ^{15}N/^{1}H IL-1β/gIC8 complex shown in panel B. For both complexes the RDCs were determined from spectra acquired at 800 MHz and from samples containing 4 mg/ml Pf1 phage alignment media. The errors on RDCs were estimated to be 1.4 Hz, based on a conservative estimate of the peak position in the processed spectra.

2.3.4 Agreement between the RDC data and reported high resolution structures for IL-1β

One high resolution NMR structure (6I1B (Clore et al., 1991)) and multiple crystal structures of IL-1β are available in the PDB. The crystal structure 2I1B (Priestle et al., 1989) was found to show the best agreement with the RDC data. The agreement between RDCs predicted from the reported structures and RDCs determined from both alignment media was compared using the PALES software (Zweckstetter and Bax, 2000). Theoretical RDCs from the PALES output can be used to compare the agreement between experimentally determined RDCs and theoretical RDCs derived from the known structure. PALES also includes a statistical analysis of the comparison that evaluates the agreement between the measured RDCs (RDC_{obs}) and the theoretical RDCs (RDC_{calc}). This quality factor (Q) is defined as:

$$Q = \text{rms}\,(RDC_{calc} - RDC_{obs}) \,/\, \text{rms}(RDC_{obs})$$

Equation 2.4

where rms is root mean square. The Q factor gives a good indication of the agreement between the two values and an ideal correlation between experimentally determined and theoretical RDCs would give a Q factor of zero. Plots showing the theoretical RDCs versus the measured RDCs for both structures, together with the Q factor values, are shown in Figure 2.11A-D i.

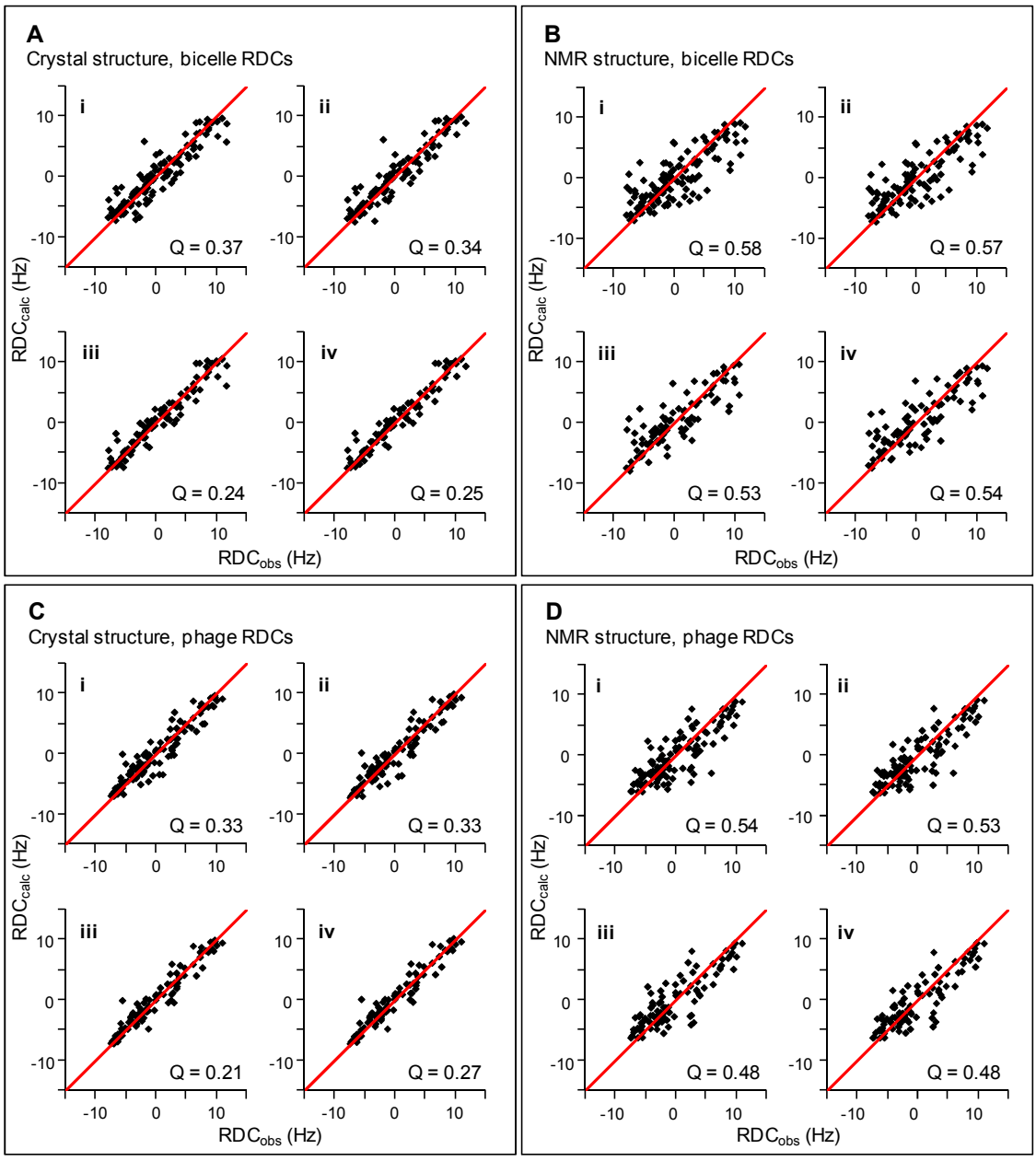


Figure 2.11

Agreement between RDCs determined for IL-1β and a set of RDCs provided by high resolution solution and crystal structures reported for IL-1β. All plots show the agreement between experimental RDCs measured for IL-1β (RDC_{obs}) against RDCs calculated with PALES using a reported structure of IL-1β from the PDB (RDC_{calc}). All NMR data were collected at 600 MHz and samples were aligned using either 5%

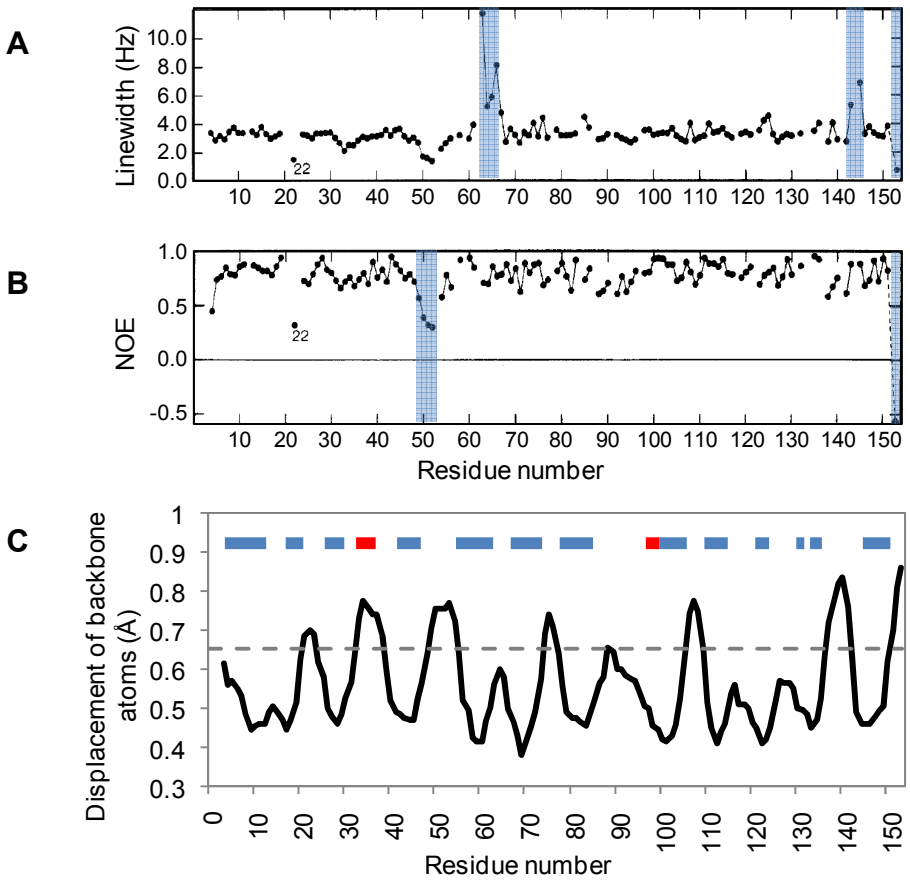
phospholipid bicelles or 5 mg/ml phage. The IL-1β structures used in the analysis were a crystal structure with the PDB code 2I1B and an NMR structure with the PDB code 7I1B. The Q factors are shown on each plot to indicate the degree of agreement and the red line shown on the plot indicates perfect agreement. The plots labelled as (i) in each panel show the agreement when all experimentally determined RDCs were involved in the calculation. Panel A shows the agreement between RDCs collected from a sample aligned using bicelles and the crystal structure. Panel B shows the agreement between RDCs collected using bicelle alignment media and the NMR structure. Panel C shows the correlation between RDCs collected using phage alignment media and the crystal structure and panel D shows the agreement between RDCs collected using phage alignment media and the NMR structure of IL-1β. In panels (ii), (iii) and (iv), RDCs obtained for regions identified to be flexible or disordered were excluded from the comparisons. The following definitions apply: i) All measured RDCs are included in the calculations; ii) RDCs for residues identified as being flexible from NMR data were excluded (Clore et al., 1990); iii) RDCs from residues with a high B-factor (atomic rmsd of 0.65Å or greater) were excluded (Priestle et al., 1989); iv) RDCs with residues identified as being flexible from NMR data and/or with a high B-factor were excluded. The errors on RDCs determined from IL-1β were estimated to be 1.0 Hz, based on a conservative estimate of the peak position in the processed spectra.

The analysis clearly shows that the measured RDCs from both alignment media are in better agreement with the crystal structure than the NMR structure. Also slightly better agreement is seen with RDCs collected in phage alignment media compared to bicelle media. Reasons for the deviation from the perfect agreement are experimental

uncertainties, protein dynamic and exchange effects and variability in the 3D structures, for example the effect of crystal packing in an X-ray structure (Meiler et al., 2001). The same experimentally determined RDCs were compared to theoretical RDCs from both the crystal and NMR structures, so the experimental uncertainties and effects from dynamics and exchange remain constant. The errors associated with the RDC values are more than an order of magnitude smaller than the range of dipolar couplings measured, implying that non-perfect agreement is a more likely result of the uncertainty in the coordinates of the structure (Zweckstetter and Bax, 2002). It would be expected that the effect of structural differences would be more severe for the crystal structure as the amide protons were added to the original coordinates by model building. The fact that the experimental RDC data correlate better with the crystal structure demonstrated by the higher Q factor for the NMR structure strongly suggests that the that the latter it is a poorer model for the experimental RDC data.

In the presence of mobility within the protein structure, the experimental dipolar coupling corresponds to a conformational average. This means that dipolar coupling measured from regions of the backbone with internal motion (for example loops and termini) may show poor agreement with theoretical RDCs calculated from a known structure. To assess the effect of potential mobility and disorder in regions of IL-1β, backbone atoms that were identified to be dynamic from NMR data, together with residues that exhibit a large atomic root mean square displacement (rmsdp) on the basis of crystallographic B-factor were excluded from the RDC data set. The ^{15}N linewidth and ^{1}H-^{15}N heteronuclear NOE data reported for IL-1β are shown in Figure 2.12 (Clore et al., 1990), with regions involved in chemical exchange marked on the plot. Residues that show ^{15}N T_2 exchange line broadening (Figure 2.12A) indicate motion on the 30 ns

to 10 ms time scale, involving at least two species with different chemical shifts. Figure 2.12B highlights regions where there is a slow conformational equilibrium between a major and minor species (Clore et al., 1990).



D

```
APVRSLNCGL RDSQQKSLVM SGPYELKALH LQGQDMEQQV VFSMSFVQGE ESNDKIPVAL
APVRSLNCGL RDSQQKSLVM SGPYELKALH LQGQDMEQQV VFSMSFVQGE ESNDKIPVAL

GLKEKNLYLS CVLKDDKPTL QLESVDPKNY PKKKMEKRFV FNKIEINNKL EFESAQFPNW
GLKEKNLYLS CVLKDDKPTL QLESVDPKNY PKKKMEKRFV FNKIEINNKL EFESAQFPNW

YISTSQAENM PVFLGGTKGG QDITDFTMQF VSS
YISTSQAENM PVFLGGTKGG QDITDFTMQF VSS
```

Figure 2.12

NMR and X-ray crystallography data that identifies flexible regions in the structure of IL-1β. Panel A shows previously reported backbone amide ^{15}N linewidth data (Clore et al., 1990), with residues that have motion in the 30 ns to 10 ms range highlighted in blue. Panel B shows previously reported ^{1}H-^{15}N NOE data (Clore et al., 1990), with residues that are in slow conformational exchange between major and minor species indicated. Panel C shows a plot of the rms displacement (rmsdp) of the backbone nitrogen atoms of IL-1β derived from reported crystallographic B-factors (Priestle et al., 1989). Residues with an rmsdp of greater than 0.65 Å (indicated by a dashed line) are considered to be flexible and/or disordered. The secondary structure of IL-1β is highlighted in panel C, with sheets shown as blue bars and helices shown as red bars. Panel D shows the sequence of IL-1β, with residues that have an rmsdp of 0.65Å or greater highlighted in yellow and residues shown to be flexible by NMR in green.

The crystallographic B-factor is a measure of the mean square displacement of the atoms in the crystal. It describes the uncertainty of atomic position and combines uncertainty from two sources. The B-factor represents uncertainty that comes from thermal motion which is the vibration of the atoms around its rest position. The B-factor also includes a disorder parameter, which is uncertainty due to atoms that do not occupy the same space in every unit cell. A plot of the atomic root mean squared displacement (rmsdp) of the backbone nitrogen atoms in the IL-1β structure is shown in Figure 2.12C (Priestle et al., 1989). Surface exposed loops would be expected to show more mobility and the data indicates these regions have a rmsdp of 0.65 Å or greater. Therefore, residues with a rmsdp of 0.65 Å based on crystallographic B-factor were defined as regions of IL-1β structure with high uncertainty in atomic position.

RDCs determined for residues defined as having high uncertainty in atomic position were excluded from the PALES analysis. Figure 2.11A-D ii shows that the agreement between the measured and theoretical RDCs did not significantly change when residues affected by high mobility from the NMR data were excluded. Figure 2.11A-D iii shows the agreement when residues affected by high mobility, based on the crystallographic data, are excluded. This improves the agreement with both structures, and with RDCs collected in both alignment media. Figure 2.11A-D iv shows no significant change in the agreement between measured and back-calculated RDCs when both the residues identified as being uncertain from NMR data and those identified as being uncertain from crystallography data are excluded.

The agreement between RDCs determined for IL-1β/gIC8 and RDCs values predicted from the IL-1β crystal structure, 2I1B (Priestle et al., 1989), are shown in Figure 2.13A. Excluding RDCs affected by high mobility based on the crystallographic data improves the Q factor from 0.37 to 0.28. The agreement of RDCs determined for IL-1β/g168 with the complete data set and with residues identified as being uncertain based on the crystallographic data excluded is shown in Figure 2.13B. The agreement between the experimentally determined RDCs and those calculated from the known crystal structure does not improve significantly in this case. It is highly possible the poor agreement that is seen is due to changes in the structure induced on complex formation, since the calculated RDCs were determined from a crystal structure of free IL-1β. It may be also be explained by the relatively small RDC data set that was used as it has been shown that the error in the orientation of the alignment tensor decreases if a larger RDC data set is used (Zweckstetter and Bax, 2002).

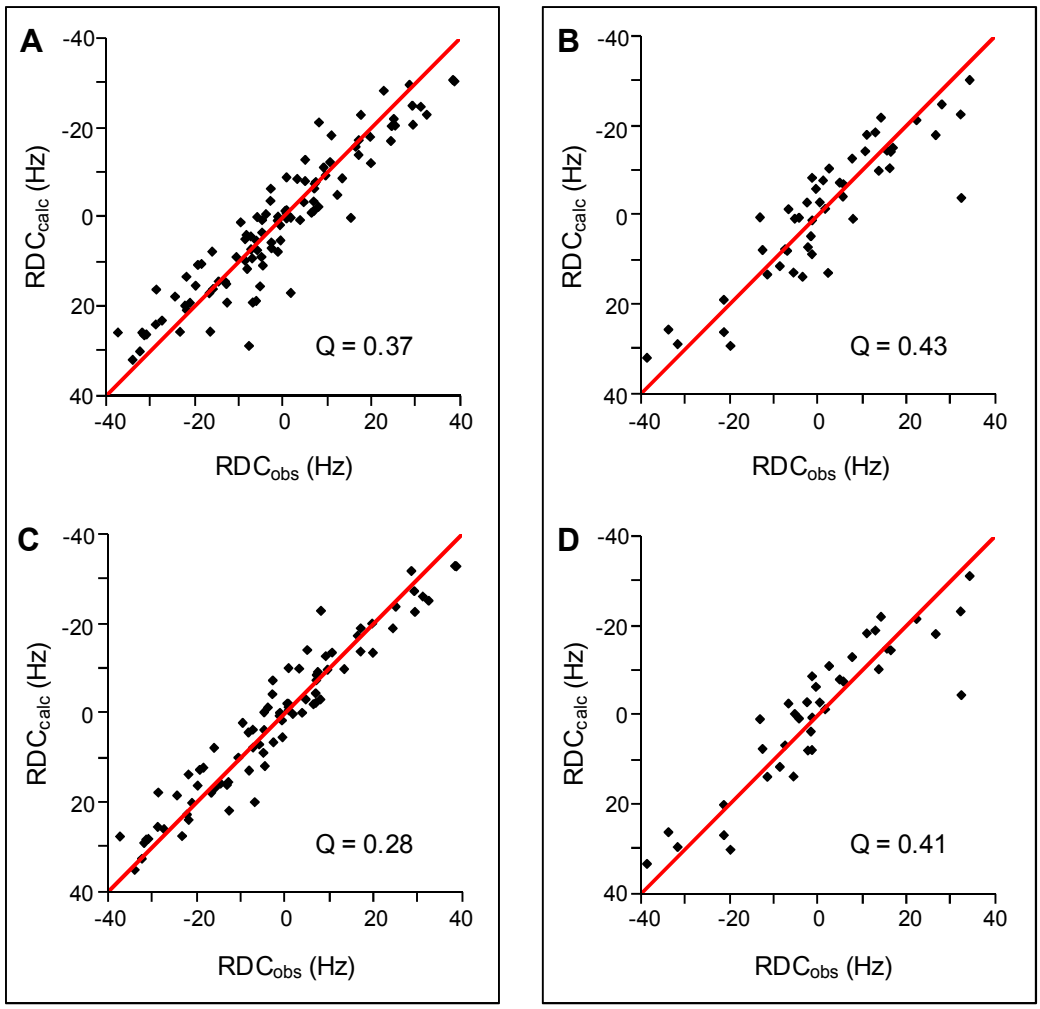


Figure 2.13

Agreement between backbone amide RDCs determined for IL-1β bound to two Fabs and RDCs predicted from a reported structure of free IL-1β. Panel A shows the agreement between experimental RDCs for the IL-1β/gIC8 (RDC_{obs}) against RDCs calculated with PALES using a known crystal structure of IL-1β, 2I1B (RDC_{calc}). The agreement when RDCs from residues with a high B-factor (equivalent to an atomic rmsd of 0.65Å or greater) are excluded from the data set is shown below. Panel B shows the agreement between RDCs obtained from IL-1β/g168 and calculated from the crystal structure, 2I1B. Below is shown the correlation when residues with a high B-factor (atomic rmsd of 0.65Å or greater) are excluded from the data set. All NMR data were

collected at 800 MHz and samples were aligned using 4 mg/ml phage. The Q factors are shown on each plot to indicate the degree of agreement and the red line on the plot indicates perfect agreement.

2.4 Conclusions

The work reported in this chapter describes the assessment of two alignment media, phospholipid bicelles and filamentous Pf1 phage, to determine RDCs from the model protein, IL-1β. The dependence of both alignment media on temperature, concentration, salt concentrations and magnetic field was investigated. Preparation of both bicelles and phage was optimised and RDCs for IL-1β were collected using IPAP-HSQC methods. RDCs were also determined for the large (~65 kDa) IL-1β/Fab complexes. The optimal concentration of Pf1 alignment media was established and RDCs were determined from the difference in the backbone amide signal positions in TROSY spectrum and decoupled HSQC spectrum.

To improve the agreement between the experimentally determined RDCs and RDCs predicted from a reported structure, RDCs from regions of the protein affected by mobility were excluded from the data set. These were identified using previously reported NMR dynamics and crystallographic B-factor data. The results showed that agreement could be improved if residues affected by high mobility based on the crystallographic data are excluded. The agreement between experimentally determined RDCs and RDCs calculated from a structure is significantly improved if the structure is refined against RDCs as shown in Chapter 4 for an antibody Fab protein.

Chapter 3 - Expression and Purification of gIC8 Fab

3.1 Introduction

The increasing realisation of the desirable therapeutic properties of antibodies has driven the development of a diverse range of prokaryotic and eukaryotic expression systems for the expression of antibodies and antibody fragments. These include bacterial, yeast, filamentous fungus, eukaryotic algae, insect cell, plant cell, mammalian cells and transgenic animal systems (Arbabi-Ghahroudi et al., 2005). The system selected is dependent on the structural requirement of the expressed protein, for example certain IgG therapeutic antibodies require glycosylation which requires a mammalian expression system. However, bacterial expression systems are usually favoured for the expression of antibody fragments (scFv, Fab) which have a less complex structure and do not require glycosylation. *E. coli* expression systems are most commonly used as they have the advantage of being straightforward to scale up for large scale therapeutic production and relatively inexpensive for fermentation. In the *E. coli* expression systems, antibody product can be directed either to the cytoplasm, the periplasmic space between the cytoplasmic and cell membranes or to the culture medium (Arbabi-Ghahroudi et al., 2005). The periplasmic space is a more oxidising environment than the cytoplasm and chaperones which exist there assist the folding of the Ig domains and form the correct disulphide bonds to stabilise the structure (Fernandez, 2004).

Fab fragments present an attractive alternative to full-length antibodies for therapeutic use as they have identical binding sites to the complete IgG but are simpler and cheaper

to produce (Holliger and Hudson, 2005). However, a disadvantage of using Fabs is that their small size makes them subject to renal filtration and as a result they have a very short circulating half-life (Covell et al., 1986). This has been addressed by conjugation of the Fab to poly(ethylene glycol) (PEG) in a process known as PEGylation. This modification improves the *in vivo* circulating half-life of the Fab by increasing the size of the molecule to above the glomerular filtration limit whilst also improving protein solubility, reducing immunogenicity and enhancing proteolytic resistance of the therapeutic (Chapman, 2002). The PEG molecule can be site specifically conjugated to cysteine residues of the Fab heavy chain that would be involved in inter-heavy chain disulphide bridges in a full length antibody but are not needed for this purpose in a Fab (Humphreys et al., 2007). PEGylation at this site ensures that the PEG molecule is situated away from the CDRs and, therefore, does not interfere with antigen binding (Chapman, 2002). Cimzia (certolizumab) is a recombinant humanized Fab fragment against TNF-α approved for the treatment of Crohn's disease and rheumatoid arthritis which acts by selectively binding and neutralising human TNF-α. The Fab is conjugated to a 40-kDa PEG increasing the circulating half-life of Cimzia to two weeks, reducing the frequency with which the therapeutic needs to be administered (Choy et al., 2002). Unlike other anti-TNF-α agents (Remicade (infliximab), Humira (adalimumab) and Enbrel (etanercept)) which contain intact human Fc domains, Cimzia does not suffer from adverse effect associated with Fc mediated complement activation and cellular toxicity (Rivkin, 2009).

There has been very little structural data obtained from antibodies and antibody fragments using NMR techniques. One notable example showed that it was possible to obtain ^{15}N and ^{1}H backbone assignments from the scFv and Fv of the phosphorylcholine

antibody, MCPC603 (Freund et al., 1994). Backbone assignments have also been reported for a human IgG1 CH_3 domain (Liu et al., 2007b) and complete human IgG1 Fc region (Liu et al., 2007a). Backbone and sidechain assignments have been made for a human kappa light chain V region (Mukherjee et al., 2009) and there are also structures reported for IgE Cε2 domain (McDonnell et al., 2001), the V_HH domain, the camelid single V_H-like binding domain (Renisio et al., 2002) and a single chain T cell receptor (Hare et al., 1999).

There are currently no NMR structures available for any Fab fragment, presumably due the difficulties of producing suitable samples and obtaining NMR data for large proteins. The work in this chapter describes the expression of a labelled gIC8 Fab fragment suitable for NMR spectroscopy. Initially, the expression of two Fab fragments selected for their affinity to IL-1β was optimised in minimal media. The results were used to successfully express and purify a ^{15}N sample of a Fab fragment. Further trials in minimal media containing 100% D_2O allowed for the expression and purification of high quantities of $^{15}N/^{2}H$ and $^{15}N/^{13}C/^{2}H$-labelled Fab fragments which were used to collect high quality 2D and 3D NMR data.

3.2 Materials and Methods

3.2.1 Protein Expression and Purification

3.2.1.1 Preparation of Competent Cells

Competent cells were made from wildtype W3110 E.coli (K12 strain; genotype F- mcrA mcrB IN[rrnD-rrnE] lambda) and also the mutant strains MXE001, MXE002, MXE003, MXE004, MXE005 MXE006 and DPH5. Cells were grown in 50 ml LB broth to an optical density at 600 nm (OD_{600}) of 0.5 to 0.8, followed by incubation on ice for 15 minutes and centrifugation (4000 rpm, 10 minutes). The cell pellet was resuspended in cold TSS (85% LB, 10% PEG_{8000}, 5% DMSO, 50 mM $MgCl_2$), frozen on dry ice and stored at -80 °C. Fab IC8 and 8516 vectors were transformed into the competent cells using a heat shock protocol consisting of 5 minutes on ice, 30 seconds at 42 °C followed by a further 2 minutes on ice. The cells were used to induce cultures in minimal media containing the SM6 blend of trace elements from which glycerol stocks were made and these were used to seed subsequent cultures.

3.2.1.2 Expression Trials

Trials for expression of IC8 and 8516 Fabs were performed in 200ml of minimal media containing the SM6 blend of trace elements (media optimised for use in fermentation tanks at UCB) with the addition of 5 mg/ml tetracycline. Cultures were seeded to an OD_{600} of 0.1 from overnight cultures and then incubated at 37 °C until an OD_{600} of 2.0 was reached, at which point the expression of the proteins was induced by the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 200 μM. At 1, 2, 4, 6, 8 and 24 hours post-induction a 10 ml aliquot was extracted from the main cultures, the cells were harvested and the pellets resuspended in 1 ml of lysis buffer

containing 30 mM tris pH 8.0, 10 mM EDTA, 20% sucrose and 100 μg/ml lysozyme and then incubated on ice for 30 minutes. The lysed cells were separated by centrifugation and the supernatant containing the periplasmic fraction was retained for Enzyme-Linked Immunosorbent Assay (ELISA) analysis.

3.2.1.3 Enzyme-Linked Immunosorbent Assay (ELISA)

Maxisorp ELISA plates (Nunc) were coated with 2.0μg/ml mouse anti-human CH1 antibody in phosphate buffered saline (PBS) and incubated at room temperature for 1 hour. The plates were washed three times using PBST (PBS with 0.01% (v/v) Tween 20; Sigma) and all subsequent washes were performed in the same way. The plates were blocked using 3% BSA, 1% Tween 20 in PBS which was applied for 1 hour at room temperature and then washed. Samples were diluted in PBS and the standard antibody (gIC8 or g8516 Fab) was added at 1.0μg/ml (diluted in PBS). The plates were incubated at room temperature for 1 hour and washed. Polyclonal mouse anti-human kappa constant peroxidase-conjugate (Southern Biotech) in a 1:5000 dilution in PBS was applied to the plates for 1 hour at room temperature. The plates were washed and developed using the horse radish peroxidase substrate, TMB. Quantification was achieved by reading the absorbance at 630nm and 490nm by spectrophotometry (Tecan GENios).

3.2.1.4 Expression and purification of unlabelled and uniformly ^{15}N-labelled gIC8 Fab

To prepare unlabelled and uniformly ^{15}N-labelled gIC8 samples, cells were grown in SM6 blend minimal media containing 10 μg/ml tetracycline and 5.2 g/l ^{15}N ammonium sulphate when required. Transformants were grown at 37 °C to an OD_{600} of approximately 1.0 when they were induced by the addition of IPTG to a concentration

of 200 μM. Cells were harvested 4 hours post-induction and resuspended in sucrose/EDTA lysis buffer containing 30 mM Tris pH 8.0, 10 mM EDTA, 20% sucrose and 100 μg/ml lysozyme. Cells were lysed using a French press and the insoluble debris was removed by centrifugation (~50 000 g, 60 minutes). The Fab containing supernatant was adjusted to pH 6.0 using 0.1 M citric acid and loaded onto a 10 ml protein G column preequilibrated with 75 mM citrate, 75 mM sodium phosphate buffer at pH 6.0 allowing 20 minutes of contact time per ml of supernatant. The Fab was eluted using 0.1 M glycine-HCl buffer at pH 2.7. Purified Fab was dialysed into 100 mM sodium phosphate, 2 mM EDTA buffer at pH 6.0 and the light chain C terminal cysteine was reduced using 2 mM THP and capped using 50 mM NEM (N-ethylmaleimide) to prevent Fab dimerisation. The capped Fab was then dialysed into 100 mM sodium chloride, 25 mM sodium phosphate, 0.02% sodium azide buffer at pH 6.5 and was finally purified on a Superdex 75 16/60 gel filtration column (GE Healthcare).

3.2.1.5 Protein expression and purification of deuterated gIC8 Fab

Cell growth and expression of ^{2}H-labelled gIC8 Fab in fully deuterated media was achieved by conditioning 50 ml overnight cultures of w3110 cells to grow with increasing percentages of deuterium in the media (0, 30, 70 and 100). Glycerol stocks were made from the fully conditioned cells and these were used to seed subsequent cultures.

To prepare ^{15}N/^{2}H and ^{15}N/^{2}H/^{13}C labelled gIC8 samples, cells were grown in minimal media containing 10 μg/ml tetracycline, 100% D_2O, 1 g/l ^{15}N ammonium sulphate and 3 g/l ^{13}C glucose when required. Transformants were grown at 37 °C to an OD_{600} of approximately 1.0 when they were induced by the addition of IPTG to a concentration

of 200 μM. Cells were harvested 28 hours post-induction then the Fab containing supernatant passed through a 22 μm filter and concentrated using an Amicon concentration cell to a final Fab concentration of >0.2 mg/ml. The concentrated supernatant was adjusted to pH 6.0 using 0.1 M citric acid and purification was performed as for the unlabelled Fab sample.

3.2.1.6 Expression and purification of IL-1β

Unlabelled and uniformly ^{15}N and $^{15}N/^{2}H$-labelled samples of IL-1β were prepared from a pET-21 based *E.coli* expression vector, which included an N-terminal hexahistidine tag with a TEV protease cleavage site in the linker. The vector was transformed into Tuner (pLysS) DE3 competent cells (Novagen). For the expression of ^{2}H-labelled IL-1β, *E.coli* cultures were grown in minimal media containing 100% D_2O. Cells were conditioned to grow in this media by successively increasing the percentage of D_2O (0, 30, 70, and 100%) in the media of 50 ml overnight cultures. Glycerol stocks were made from the fully conditioned cells and these were used to inoculate subsequent cultures.

Inoculated cultures were incubated at 30 °C in minimal media containing 100 μg/ml carbenicillin, 1 g/l > 99% ^{15}N ammonium sulphate as the sole nitrogen source when required and >99.8% D_2O where appropriate. After reaching an OD_{600} of 0.7 to 1.0, the temperature was then reduced to 25 °C and the overnight IL-1β expression induced by the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG) to a concentration of 0.01 mM. Cells were harvested by centrifugation (6000 g for 20 minutes), the cell pellet resuspended in PBS containing protease inhibitor (Roche), lysed using a French press and the insoluble debris removed by centrifugation at 48 000 g for 30 minutes. The clarified lysate was dialysed into running buffer for the Ni-NTA column (100 mM

sodium chloride, 25 mM Tris, 15 mM imidazole buffer at pH 7.5). IL-1β was purified on a 6 ml Ni-NTA affinity column (Qiagen) and eluted with a linear imidazole gradient (15-500 mM). Fractions containing IL-1β were pooled and the His-tag was removed by incubation with TEV protease (TEV:IL-1β 1:15 w:w) at room temperature overnight. This was then dialysed into running buffer for the Ni-NTA column (as above) and IL-1β was separated from the cleaved His-tag and the TEV protease by passing through a Ni-NTA column. IL-1β containing fractions were then dialysed into running buffer for gel filtration (100 mM sodium chloride, 25 mM sodium phosphate, 0.02% sodium azide buffer at pH 6.5) and final purification was performed on a Superdex 75 16/60 gel filtration column (GE Healthcare) at a flow rate of 1 ml/min.

3.2.2 Thermal stability of gIC8 Fab by Thermofluor

Sypro Orange dye was used to monitor the Fab unfolding process. The assay was run in quadruplicate in 384 well plates (96 samples/plate). To each well was added 1 µl of sample at a concentration of 1 mg/ml, 1 µl of 30x Sypro Orange dye and 8 µl of buffer. A 7900HT fast real-time PCR system which contained a heating device was used for accurate temperature control. Fluorescence changes induced by Sypro Orange binding to exposed hydrophobic regions were monitored between 20 ^{o}C and 99 ^{o}C (ramp rate = 0.9 ^{o}C /min).

3.2.3 NMR spectroscopy

NMR spectra were acquired from a 0.35 ml sample of 0.45 mM ^{15}N-labelled gIC8 Fab in 100 mM sodium chloride, 25 mM sodium phosphate, 0.02% sodium azide buffer at pH 6.5. Spectra were collected between 35 and 50 °C on either a 600 MHz Bruker DRX or an 800 MHz Bruker Avance system. ^{15}N/^{1}H HSQC and TROSY spectra were

acquired for approximately 1 hour with acquisition times of 60 ms in ^{1}H (F_2) and 40 ms in ^{15}N (F_1). The WATERGATE method was used to suppress the water signal when required (Piotto et al., 1992) and all NMR data was processed using Topspin (Bruker Biospin Ltd.) and analysed using the Sparky package (Goddard and Kneller).

3.3 Results and Discussion

3.3.1 Fab expression trials

Fab expression vectors were obtained from UCB-Celltech for two antibody Fab fragments with high affinity for IL-1β, known as gIC8 and g8516, and assessed for their suitability to produce samples of both proteins. It had been shown previously at UCB-Celltech that high yields of Fab could be obtained in a mutant strain of the *E. coli* cell line w3110 containing a deletion of the periplasmic protease Tsp and an inactivating mutation in the active site of another periplamic protease, DegP. Expression trials were carried out for both gIC8 and g8516 Fab vectors transformed into the wildtype w3110 cells and mutant strains called MXE001, which had deleted Tsp, and MXE005, which had deleted Tsp and inactive DegP. A cell line called DPH5 was also used which has a repair mutation that aids the secretion pathway. Protein expression levels in 200 ml minimal media containing the SM6 blend of elements were determined by ELISA at 1, 2, 4, 6, 8 and 24 hour time points after induction. Figure 3.1 shows that the greatest Fab yields were obtained for gIC8 expressed in wildtype w3110 and DPH5 cells and g8516 yields were similar in all the cell lines. In further selected expression trials, carried out in 800 ml cultures, both gIC8 and g8516 were expressed in wildtype w3110 and DPH5 cells, harvested at the optimal time as determined by the expression trials and fully lysed using a French press to extract all protein from the bacterial periplasm.

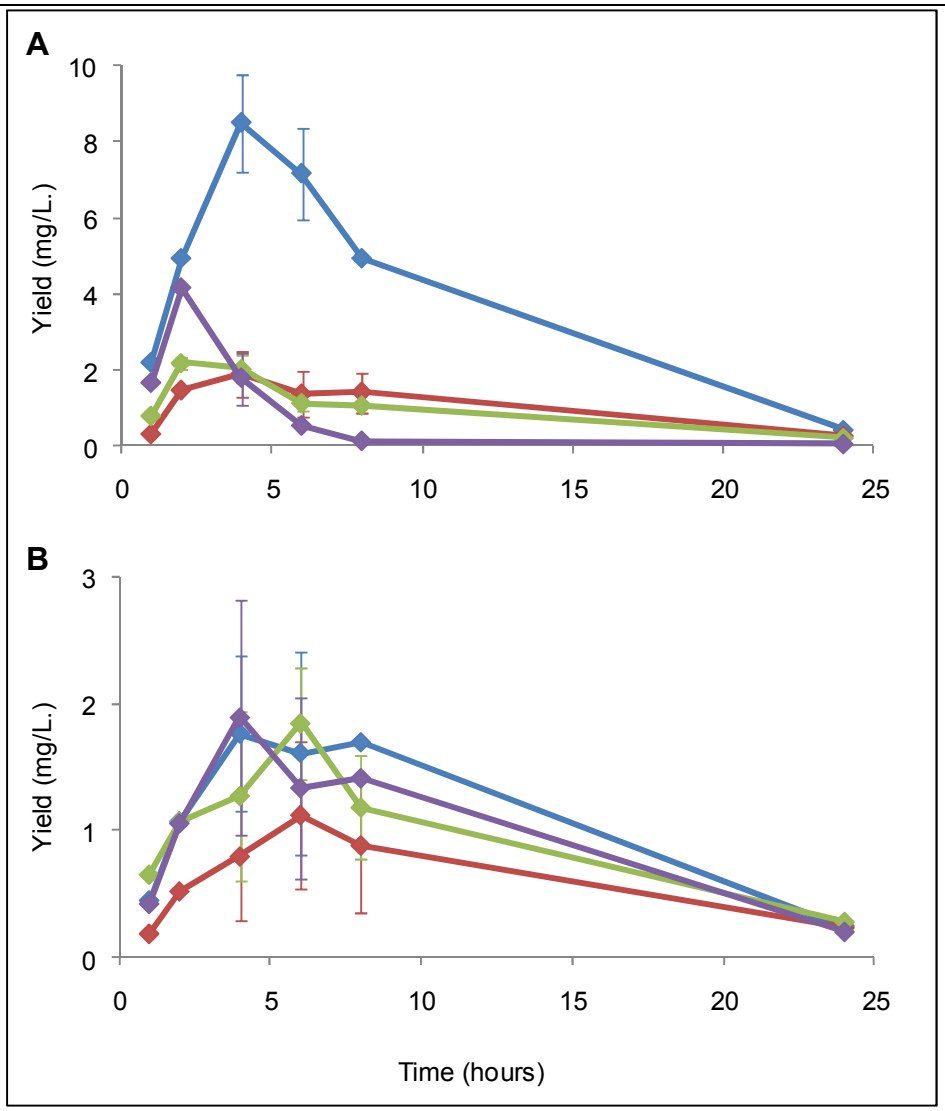


Figure 3.1

gIC8 and g8516 expression trials. The expression of gIC8 at 1, 2, 4, 6, 8 and 24 hour time points post-induction in the wildtype w3110 (blue), MXE001 (red), MXE005 (green) and DPH5 (purple) cells lines was determined by ELISA (Panel A). Duplicate readings were taken for time points where there was high levels of expression and the range of these values are indicated on the appropriate points. Panel B shows the expression of g8516 at 1, 2, 4, 6, 8 and 24 hour time points post-induction in the wildtype w3110, MXE001, MXE005 and DPH5 cell lines with the same colour scheme as described above.

Analysis by ELISA showed that for g8516 expressed in DPH5 and w3110 the yields were 1.1 mg/l and 2.3 mg/l respectively. For gIC8, ELISA analysis showed final protein concentrations to be 3.1 mg/l when expressed in DPH5 cells and 4.7 mg/l in w3110. A protein G column was used to purify the gIC8 protein expressed in w3110 cells to give a purified yield of 2.6 mg.

3.3.2 Expression and purification of ^{15}N-labelled gIC8 Fab

Uniformly ^{15}N-labelled gIC8 for NMR trials was expressed and purified from 4 l of w3110 cell culture in SM6 minimal media containing 5.2 g/l ^{15}N-labelled ammonium sulphate as the sole nitrogen source. Protein was extracted from the periplasm and purified using a protein G column to give a yield of 8.6 mg. Uniform ^{15}N labelling was confirmed using mass spectrometry (Figure 3.2) and the mean melting temperature of gIC8 Fab was determined by Thermafluor to be 76.5 +/- 0.6 °C (Figure 3.3).

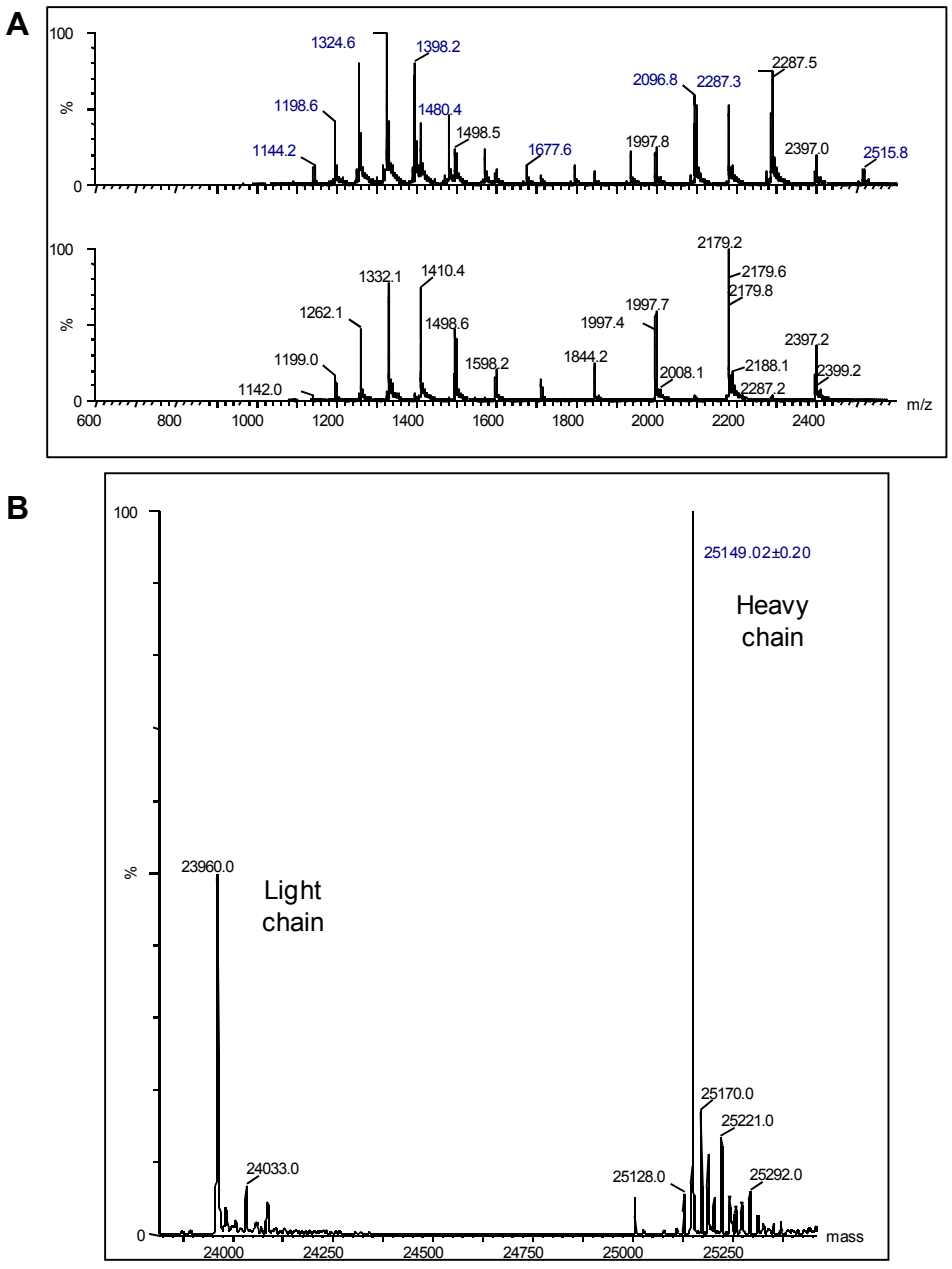


Figure 3.2

Mass spectrometry analysis of ^{15}N-labelled gIC8 Fab. Panel A shows the mass to charge envelope obtained by electrospray mass spectrometry of the heavy chain (top) and light chain (bottom). Panel B shows the relevant regions of a deconvoluted mass spectrum for the light chain (left) and the heavy chain (right). The expected mass for the ^{15}N-labelled heavy chain was 25149.4 kDa and for the ^{15}N light chain the expected mass

was 23961.1 kDa. The observed masses for the chains are indicated on the spectra and show that the Fab has been successfully ^{15}N-labelled. Mass spectrometry data was obtained by collaborators at UCB-Celltech.

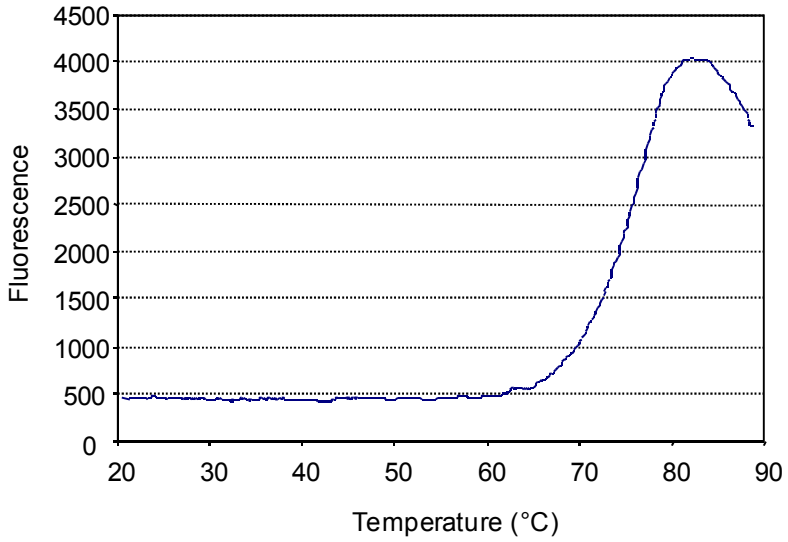


Figure 3.3

Determination of the thermal stability of gIC8 Fab by Thermofluor. The unfolding of the Fab was monitored by raising the temperature from 20 °C to 99 °C in the presence of Sypro Orange fluorescent dye. This dye binds to hydrophobic regions of the protein that become exposed upon unfolding, resulting in change in the emission spectrum which is shown plotted against temperature. The midpoint of the slope indicates that the melting temperature is 76.5 °C for the Fab. Thermofluor data was obtained by collaborators at UCB-Celltech.

3.3.3 NMR spectroscopy

A $^{15}N/^{1}H$ HSQC spectrum collected from a 0.35 ml sample of ^{15}N gIC8 Fab at 0.45 mM at 45 °C and 800 MHz (Figure 3.4A) is very well dispersed with many well resolved peaks. The average linewidth of a selection of resolved peaks in the spectrum is 30.0 Hz +/- 3.1 Hz. Experience from previous structural biology projects in the laboratory (summarised in Table 3.1) has shown that it is difficult to obtain high quality triple resonance data from proteins with backbone amide proton linewidths in the $^{15}N/^{1}H$ HSQC spectra of greater than approximately 25 Hz.

Table 3.1 shows that as protein size increases, the average spectral linewidth also increases. This is due to short ^{1}HN T_2 relaxation, as a direct result of the increased τ_c of larger proteins. A significant contribution to the ^{1}HN T_2 relaxation is from dipolar coupling to surrounding aliphatic protons. Replacing non-exchangeable protons in the structure with deuterons can reduce the ^{1}HN T_2 relaxation time by approximately 1.7-fold and can therefore have a considerable impact on the spectral sensitivity and resolution (Venters et al., 1995). Deuteration also reduces the ^{13}C T_2 relaxation times by approximately 12.5-fold as the most significant contribution to ^{13}C T_2 relaxation is strong dipolar coupling between a ^{13}C spin and its directly bonded proton (Cavanagh et al., 1996). Therefore, $^{15}N/^{13}C/^{2}H$ gIC8 Fab protein was expressed and purified for the collection of triple resonance data.

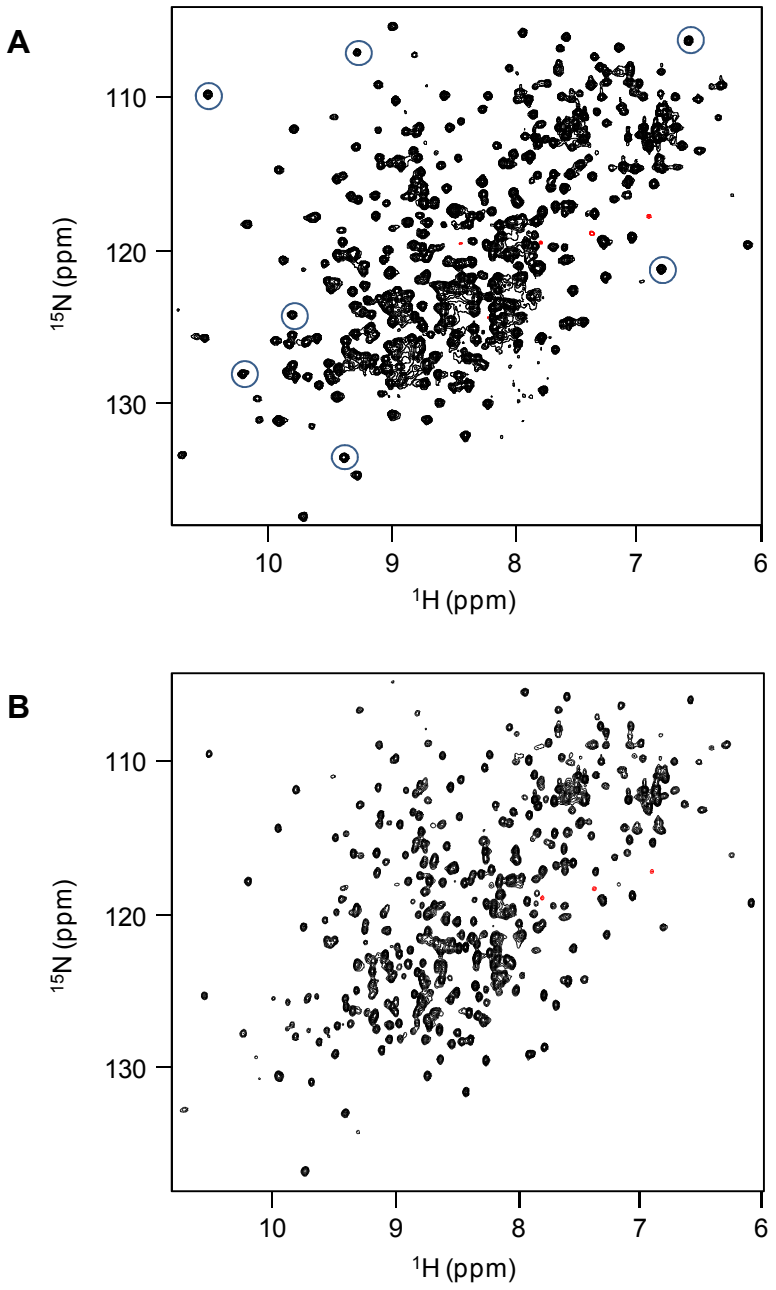


Figure 3.4

^{15}N/^{1}H HSQC spectra of gIC8 Fab. Panel A shows a ^{15}N/^{1}H HSQC spectrum of ^{15}N-labelled gIC8 Fab at a concentration of 0.45 mM and Panel B shows an HSQC spectrum of ^{15}N/^{2}H-labelled gIC8 Fab at 0.30 mM. Both spectra were collected at 45 °C on a 800 MHz spectrometer and demonstrate the reduction in backbone proton linewidth when proteins are deuterated. The peaks highlighted with a blue circle were selected to

measure the linewidths which are shown in Table 3.1. Negative peaks are shown in red. These peaks correspond to folded residues arising from arginine sidechains. The spectra have been adjusted so that peaks are shown just above the noise level. Peaks that are missing in the spectrum of $^{15}N/^{2}H$ gIC8 compared to the $^{15}N/^{1}H$-labelled sample are due to incomplete backbone amide proton back-exchange. This is discussed further in Chapter 4.

Protein	Pdcd4 MA-3C	ESAT-6/ CFP-10	Rv0288/ Rv0287	Pdcd4 2MA-3	IC8 (scFv)/ IL-1β	IC8 (scFv)/ IL-1β	gIC8 Fab	gIC8 Fab
Mass (Da)	15487	20564	24495	33002	44753	48137	49109	51989
No. of residues	136	195	233	298	407	407	444	444
Labelling	^{15}N	^{15}N	^{15}N/^{13}C	^{15}N	^{15}N	^{15}N/^{2}H	^{15}N	^{15}N/^{2}H
Temp	25°C	35°C	35°C	25°C	40°C	40°C	45°C	45°C
	17.2	18.9	22.1	28.4	22.6	19.0	32.6	16.6
	15.0	19.7	22.5	28.7	24.4	16.6	26.2	22.2
	21.9	15.4	21.4	26.4	28.0	16.0	32.6	23.0
	15.6	19.6	29.8	25.1	26.8	16.0	30.2	23.8
	14.1	17.4	36.6	25.4	31.0	16.0	34.2	16.6
	15.0	16.2	29.2	27.1	30.4	17.2	25.4	19.8
	15.7	15.7	28.8	29.5	28.6	22.0	29.4	15.8
	17.5	18.4	16.2	22.7			29.4	22.2
	18.2	15.0	22.4	21.1				
	18.1	15.4	23.7	24.0				
Mean (Hz)	**16.8**	**17.2**	**25.3**	**25.8**	**27.4**	**17.5**	**30.0**	**20.0**
S.D.	2.3	1.8	5.8	2.7	3.1	2.2	3.1	3.2

Table 3.1

Table of backbone amide proton linewidths measured from ^{15}N/^{1}H HSQC spectra from a selection of proteins. The table shows amide proton linewidths from 9 proteins and protein complexes, the structures of which have been solved within the laboratory. The peaks selected for linewidth measurement in the Fab HSQC are highlighted in Figure 3.4. The linewidth reduction on protein perdeuteration is illustrated by the comparisons between the ^{15}N and ^{15}N/^{13}C/^{2}H scFv IC8/IL-1β complex and the ^{15}N and ^{15}N/^{2}H gIC8 Fab, shown in the last four columns.

3.3.4 Expression and purification of $^{15}N/^{13}C/^{2}H$ gIC8 Fab

The SM6 minimal media used in the Fab expression trials was unsuitable for expression of ^{13}C-labelled protein as it contained high concentrations of the unlabelled carbon sources glucose and citrate. Expression trials in an alternative minimal media containing 3 g/l glucose showed reasonable cell growth and expression (data not shown) and so this was used for expression of both $^{15}N/^{2}H$ and $^{15}N/^{13}C/^{2}H$ gIC8, using ^{13}C-labelled glucose when required. Expression trials in media containing 100% deuterium over 28 hours were monitored by ELISA analysis. This showed an increase in the levels of periplasmic protein that reached a plateau between 4 and 12 hours after which point the protein levels reduced gradually (Figure 3.5) however, the protein levels in the periplasm were significantly lower than seen for the cells cultured in H_2O. To investigate potential causes for this reduced yield, protein concentration in the growth media was analysed. The protein levels in the growth media steadily increased after 2 hours to 11.3 mg/l after 28 hours (Figure 3.5). This suggested that the most of the protein was leaking from the periplasm into the growth media when cells were grown in deuterated media. Therefore, for the expression of $^{15}N/^{2}H$ and $^{15}N/^{13}C/^{2}H$ gIC8, cells were harvested from the media 28 hours post-induction.

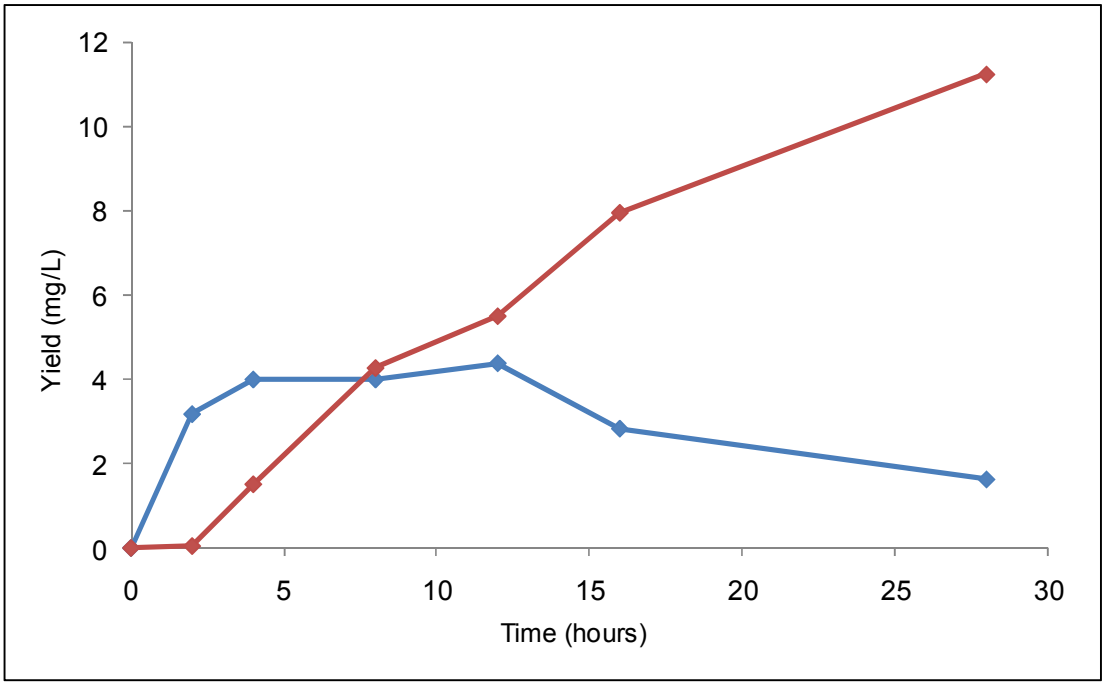


Figure 3.5

gIC8 expression in fully deuterated minimal media. ELISA was used to determine the concentration of Fab in samples extracted from the periplasm of w3110 cells (blue) or isolated from the growth media (red) between 1 and 28 hours after induction. The growth media contained 100% D_2O.

Lane L in the SDS-PAGE gel shown in Figure 3.6 Panel B contains the growth media after concentration. A band corresponding to the Fab heavy and light chains can clearly be seen at about 25 kDa in addition to many other protein bands, showing that other proteins were leaking into the media and gIC8 was not selectively transported across the membrane.

Protein leakage from the periplasm into the growth media has previously been reported. In some cases leakiness has been induced by selective mechanical and enzymatic disruption of the cells. Others have identified protein leakage from the *E. coli* periplasm into the growth medium as a result of the addition of specific components to the growth medium (for example chromate or glycine) or changes in pH, aeration or temperature. However, the mechanisms by which these changes in conditions affect leakage are poorly characterised (Shokri et al., 2002). The cause of the leakage of gIC8 from the periplasm to the growth medium when the bacteria were cultured in media containing D_2O is unclear. It is possible that a similar process was occurring when the cells were cultured in H_2O. However, this was not investigated as the protein yields obtained from the periplasm was adequate. To the best of my knowledge there are no reported cases where *E. coli* cells become leakier when grown in media containing D_2O and the W3110 *E. coli* strain has not been specifically shown to have leaky cell membranes.

The Fab protein was purified from the growth media, which was initially concentrated in an Amicon cell to reduce the total volume and ensure the Fab concentration was high enough to bind to the protein G column (< 0.2 mg/ml). The concentrated growth media was then purified on a protein G column (Figure 3.6). A non-reduced SDS-PAGE gel showed that a significant proportion of the Fab was forming dimers in solution due to an exposed thiol group designed to facilitate PEGylation if the Fab was to be used as a therapeutic. Therefore, complete reduction and capping of cysteines with NEM (N-ethylmaleimide) was performed to prevent further dimerisation was prior to purification of the Fab by gel filtration (Figure 3.7). Typical final yields were 13.1 mg for $^{15}N/^{2}H$ gIC8 and 9.4 mg for $^{15}N/^{13}C/^{2}H$ gIC8 from a starting volume of 1.6 l in both cases.

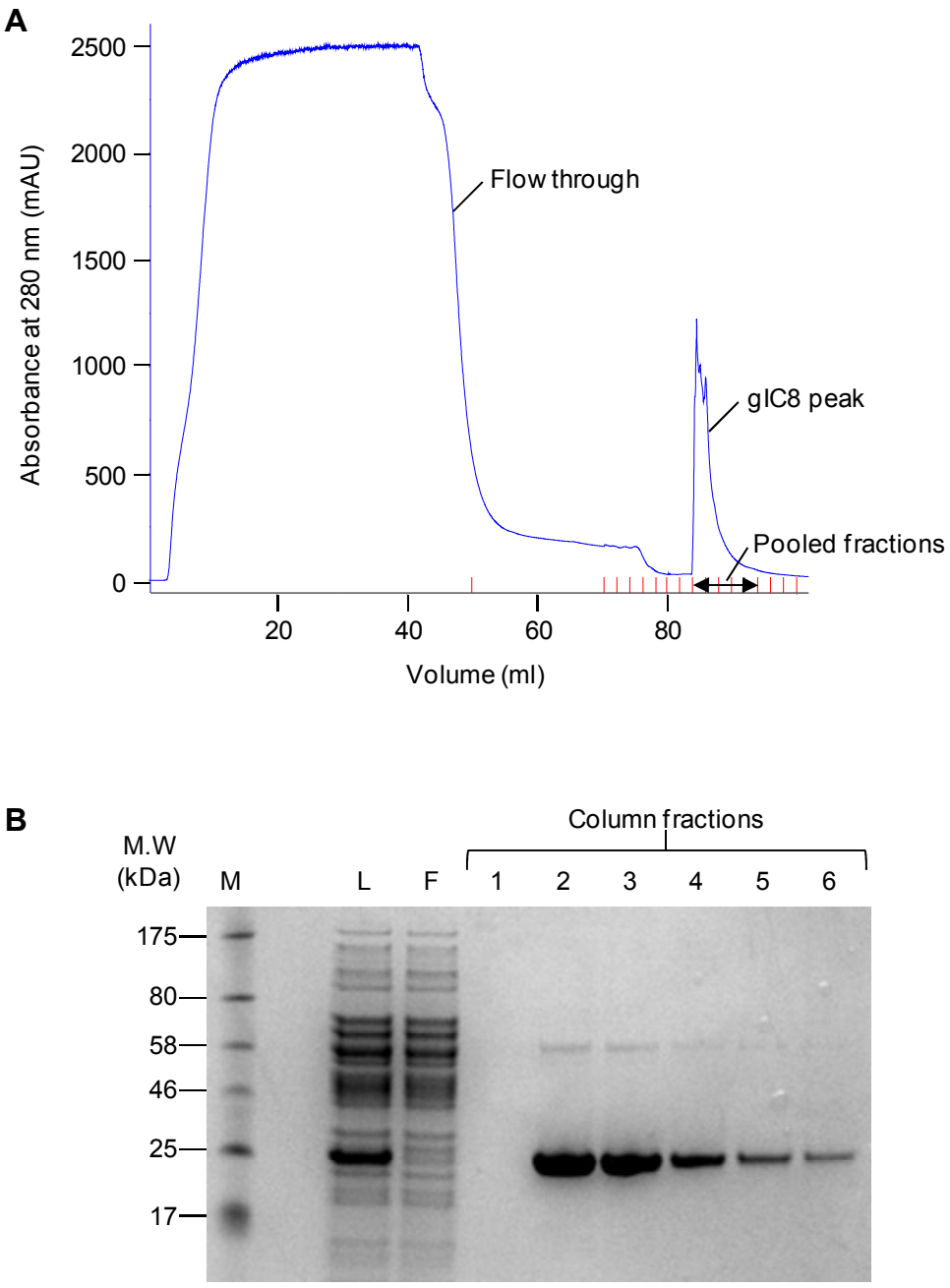


Figure 3.6

Purification of gIC8 Fab. Panel A shows a typical FPLC profile attained for the Protein G column based purification of gIC8 Fab. The protein was eluted from the column with a pH gradient from pH 6.0 to pH 2.7. The flow through and elution peaks are labelled and the pooled fractions are indicated. Panel B shows an SDS-PAGE gel

displaying FPLC elution fractions. Lane M contains molecular markers as indicated, lane L contains the original column load and lane F contains the column flow through. Lanes 1 to 6 show column fractions. The fractions in lanes 2 to 6 were pooled for further purification.

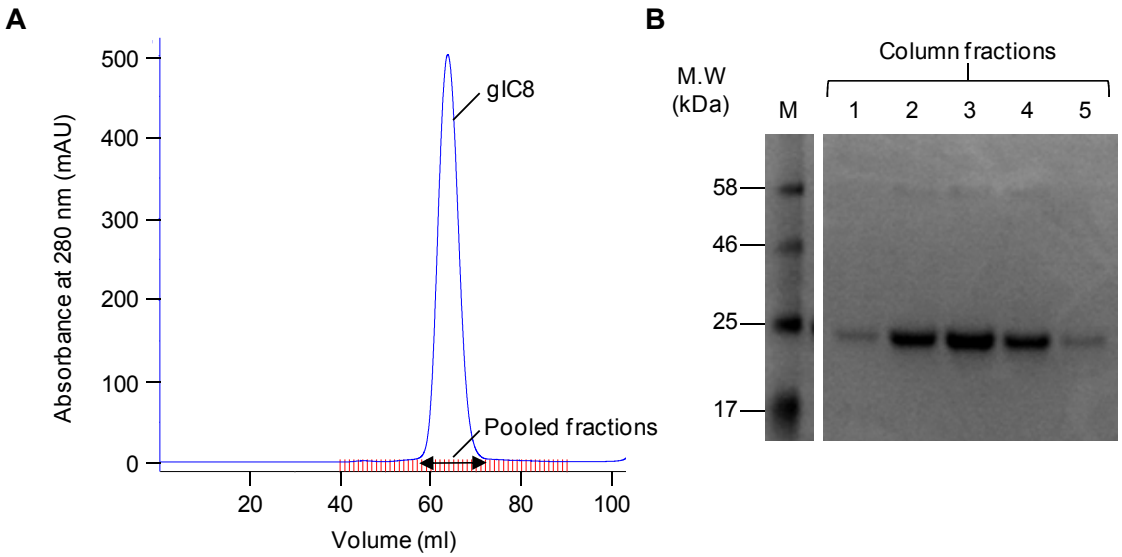


Figure 3.7

Purification of gIC8 Fab by gel filtration. Panel A shows a typical FPLC profile attained by gel filtration of gIC8 Fab on a Superdex 75 16/60 column with the elution peak and pooled fractions indicated. Panel B shows an SDS-PAGE gel displaying some of the pooled fractions in Panel A. Lane M contains molecular markers as indicated. Lanes 1-5 are samples from across the gIC8 peak as shown in Panel A.

Figure 3.4B shows a $^{15}N/^{1}H$ HSQC spectrum of $^{15}N/^{12}H$-labelled Fab and the amide proton linewidths for the set of indicated peaks is shown in Table 3.1. The average backbone proton linewidth in this spectrum has reduced to 20.0 +/- 3.3 Hz (based on the same set of peaks as measured for the $^{15}N/^{1}H$ protein), allowing for the collection of high quality triple resonance spectra.

3.3.5 Expression and purification of IL-1β

Previous work carried out at UCB-Celltech had established optimal conditions for the expression of IL-1β in Tuner (pLysS) DE3 competent *E. coli* cells (Novagen). The presence of a His-tag allowed for initial purification of IL-1β on a Ni-NTA column (Figure 3.8). The IL-1β containing fractions were pooled and the His-tag was then cleaved using TEV protease. The IL-1β was passed through a second Ni-NTA column to remove the cleaved His-tag and the His-tagged TEV protease, which bound to the column, and the purified IL-1β was collected in the flow through (Figure 3.9). The IL-1β was subjected to final purification on a gel filtration column (Figure 3.10).

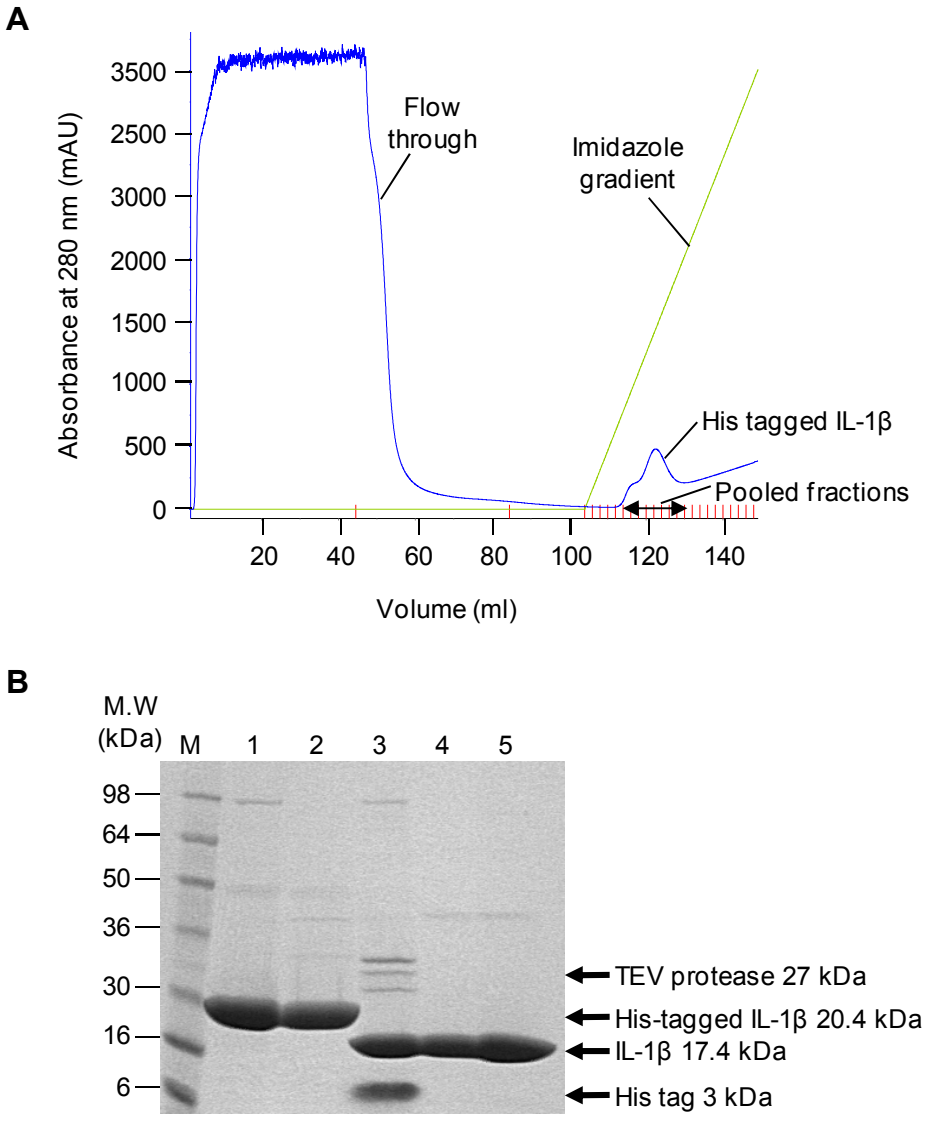


Figure 3.8

Purification of IL-1β. Panel A shows a typical FPLC profile attained for the Ni-NTA column based purification of His-tagged IL-1β. The protein was eluted from the column over a linear gradient from 15-500 mM imidazole (indicated in green). The flow through and elution peaks are labelled and the pooled fractions are also indicated. Panel B shows an SDS-PAGE gel displaying FPLC elution fractions and the subsequent TEV cleavage. Lane M contains molecular markers as indicated. Lanes 1 and 2 contain purified His-tagged IL-1β from the pooled fractions. Lane 3 shows the successful cleavage of the His-tag with TEV protease. Lane 4 shows the purified untagged IL-1β

following the purification step shown in Figure 3.9 and lane 5 shows IL-1β after final purification by gel filtration (Figure 3.10).

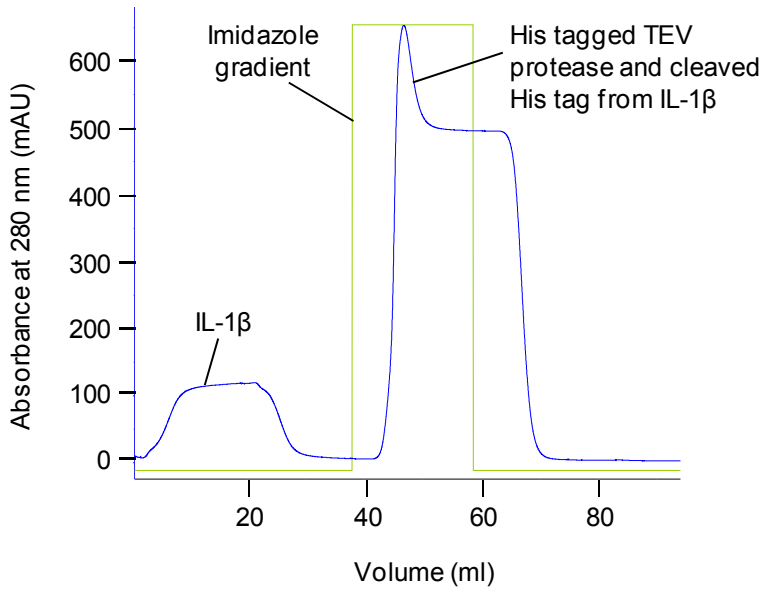


Figure 3.9

Separation of cleaved IL-1β on a Ni-NTA column. A typical FPLC profile attained for the Ni-NTA column based purification of IL-1β. After TEV cleavage the untagged IL-1β has no affinity for the column so flows through, whilst the cleaved His-tag and the His-tagged TEV protease both bind the column. Once the IL-1β has been collected, the bound proteins are eluted over an imidazole step gradient from 15 to 500 mM (indicated in green). An SDS-PAGE gel showing TEV cleavage and purification of untagged IL-1β is shown in Figure 3.8B in lanes 3 and 4.

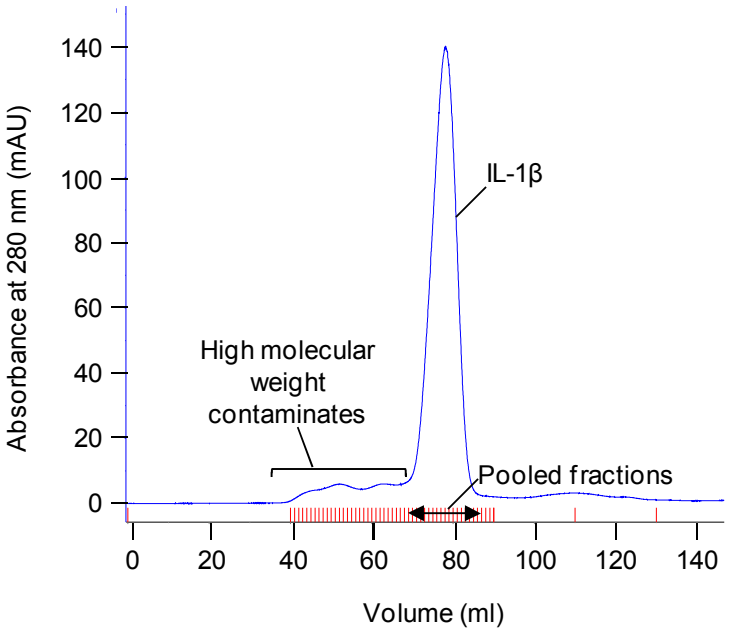


Figure 3.10

Final purification of IL-1β by gel filtration. A typical FPLC profile attained for gel filtration based purification of untagged IL-1β on a Superdex 75 16/60 column. A small amount of high molecular weight contaminates are highlighted, and the elution peak of purified IL-1β and fractions the pooled are indicated. A gel showing the purified IL-1β is shown in Figure 3.8B in lane 5.

3.4 Conclusions

The work reported in this chapter has described the expression trials of two Fab fragments in four cell lines to obtain an optimal expression protocol for the Fab fragment gIC8, allowing for the expression and purification of the ^{15}N-labelled Fab. Initial $^{15}N/^{1}H$ HSQC experiments collected from this sample gave well dispersed spectra and the protocol was further optimised for expression of $^{15}N/^{2}H$ and $^{15}N/^{13}C/^{2}H$ Fab.

Expression trials in minimal media containing 100% D_2O showed that 24 hours post-induction the periplasmic protein had leaked into the growth media. Therefore, the deuterated Fab could be purified from the supernatant and despite the decreased cell growth in D_2O it was possible to obtain at least a two-fold increase of the yield compared to Fab expressed in media containing 100% H_2O.

$^{15}N/^{1}H$ HSQC spectra collected from the $^{15}N/^{1}H$ Fab showed significantly decreased amide proton linewidths making the collection of high resolution NMR data possible. The successful $^{15}N/^{13}C/^{2}H$ isotope labelling of a Fab fragment shown here has overcome a considerable barrier that has so far prevented structural studies of Fab proteins by NMR.

Chapter 4 - Determination of a NMR Based Model for the Structure of gIC8/IL-1β Complex

4.1 Introduction

4.1.1 Protein-protein docking

Most proteins achieve their function by interacting with other proteins to form a complex. Interactions between proteins are critical in many mechanisms of cellular control, including competitive inhibition, allosteric regulation, gene regulation and signal transduction. Obtaining three dimensional structures of these complexes is essential to fully understanding the mechanisms of their cellular function. Protein structures are obtained by two main sources, X-ray crystallography and NMR and while both are effective tools for determining the structures of single molecules, obtaining the structures of protein-protein complexes is a greater challenge. X-ray crystallography depends on protein crystallization and the dynamic nature of complex formation results in difficulties in obtaining crystals. Use of NMR techniques is limited by the size of the protein that can be observed, as discussed previously. Apart from the problems related to data collection from large proteins, the traditional NMR approach to solving protein-protein complexes requires the collection of intermolecular NOEs. Often intermolecular NOEs involve side chains protons, necessitating complete assignment of NMR signals, making determination of complexes by NMR a lengthy and difficult process. Prediction of protein-protein interactions is, therefore, a very important tool in providing information about the function of biological complexes (Russell et al., 2004; Smith and Sternberg, 2002).

When the atomic structures of the individual interacting structures are known, either by experimental means or homology modelling, computational methods can be used to predict the structure of the interaction. Protein docking predictions are based on the premise that the structure of the complex is the lowest energy state available to the system, and function by sampling many states and selecting the conformations with the lowest free energy. It is important to establish an effective scoring system that can discriminate between correctly or nearly correctly docked structures and those which are incorrectly docked.

The earliest example of computerised protein-protein docking (Wodak and Janin, 1978) used low resolution structures with a rigid body structure to perform comprehensive sampling over the surface of one docking partner. Advances in technology led to the development of improved rigid body docking techniques, which matched the shapes of protein interfaces to provide low energy docked structures. These techniques are rapid and fairly successful but fail to account for conformational changes that may occur in the protein structure on binding, for example sidechain and backbone movements (Gray, 2006). Not allowing protein flexibility during docking algorithms could prevent the native complex structure being calculated. The accuracy of the determined complex structure also relies on some protein flexibility during docking and this becomes especially important if the docked structures are homology models. Allowing flexibility in a docking algorithm increases the number of degrees of freedom in the protein, which means that flexible protein-protein docking requires longer running times than rigid body docking and also produces a greater number of false positive solutions. These structures must be scored properly to identify near native results (Andrusier et al., 2008).

Biochemical and/or biophysical data obtained from the protein complexes can also be combined with docking. Several sources of experimental data can be used to define protein interfaces. Mutagenesis can be used to identify interface residues as those that influence complex formation when mutated. Mass spectrometry can also be used to obtain information on biomolecular complexes based on the changes of H/D exchange rate between the free and bound protein, which identifies interface residues that become less solvent exposed on complex formation (van Dijk et al., 2005a). NMR techniques can provide information about the interaction surface by chemical shift perturbation techniques. Signals in an NMR spectrum from residues involved in the protein-protein interaction surface will show a chemical shift change on binding to the partner protein, highlighting residues that are perturbed on complex formation. RDCs and ^{15}N relaxation data can also be introduced as restraints in docking protocols and can provide information about the relative orientations of the binding partners (van Dijk et al., 2005b).

The high ambiguity driven docking approach (HADDOCK) uses experimental biochemical and/or biophysical data to provide information about interacting residues which are introduced as ambiguous interaction restraints (AIRs) to drive docking (Dominguez et al., 2003). The ambiguity refers to the fact that residues at the protein-protein interface of the components are identified but there is no knowledge of specific pairwise interactions across the interface. AIRs are defined as two types of residues; active residues are experimentally identified as being at the protein-protein interface and passive residues are neighbours of active residues. This allows some 'fuzziness' at the interface to allow for incomplete experimental identification of interface residues. Experimental data that can be used to define the interaction site include chemical shift

perturbation data from NMR titrations and mutagenesis data, which can be combined with RDCs and/or ^{15}N relaxation data to provide information on the relative orientation of the complex components. RDCs can be used as direct restraints (called SANI) where the relative orientations of each component are optimised with respect to a single, specified alignment tensor. In addition, RDCs can be introduced as intervector (NH-NH) projection angle restraints (VEAN), obtained by combining RDCs in a pairwise manner. These restraints do not require knowledge of the orientation of the molecule with respect to the alignment tensor (Meiler et al., 2000). The HADDOCK protocol follows three consecutive stages: (i) randomisation of the starting orientation and rigid body energy minimisation (ii) semi-flexible simulated annealing and (iii) final refinement with an explicit solvent. During the simulated annealing and water refinement stages, amino acids located at the interface (backbone and side chain) are allowed to move to optimise interface packing (Dominguez et al., 2003).

There are currently no NMR-based antibody Fab fragment or Fab complex structures available, presumably to the large size of the protein and difficulties in producing labelled protein. The work in this chapter describes the collection of 2D and 3D NMR spectra for gIC8 Fab and backbone resonance assignment (HN, N, Cα, Cβ, CO) of the spectra. In addition, backbone amide RDCs and long range HN-HN NOEs were also obtained and this data was used to refine a homology model for the free Fab. Analysis of the chemical shift changes induced on Fab binding to its antigen IL-1β allowed for the interaction surface on the Fab to be determined. Previous knowledge about the interaction surface on IL-1β was combined with backbone amide RDC and long range HN-HN NOEs to dock a homology model of the Fab and a reported crystal structure of IL-1β using HADDOCK. Comparisons made between the free and bound Fab structures

show a significant difference in the orientation of the V_L domain relative to the other domains, suggesting a mechanism for antigen induced B cell signalling.

4.2 Materials and Methods

4.2.1 NMR spectroscopy

NMR spectra were acquired from 0.35 ml samples of 0.3 mM free Fab and Fab/IL-1β complex in 25 mM sodium phosphate, 100 mM sodium chloride, 0.02% w/v sodium azide buffer at pH 6.5. All data were collected at 45 °C on an 800 MHz Bruker Avance System. The TROSY based 3D spectra recorded to obtain backbone assignments of the free Fab were: $^{15}N/^{13}C/^{1}H$ HNCACB (Wittekind and Mueller, 1993), HN(CO)CACB (Grzesiek and Bax, 1992), HNCA, HN(CO)CA and HNCO (Kay et al., 1990). Typically 3D experiments were collected for with acquisition times of 60-90 ms in F_3 (^{1}H), 22-25 ms in F_2 (^{15}N) and 8-9 ms in F_1 (^{13}C) apart from the HNCO which was 22 ms for free and 41 ms for Fab/IL-1β. Long range NOE data were obtained using $^{15}N/^{1}H$ NOESY-HSQC (Marion et al., 1989) experiments collected over 87 hrs with a mixing time of 500 ms. RDCs were collected at 800 MHz from 0.35 ml samples of free $^{15}N/^{2}H$ Fab at a concentration of 0.3 mM with the addition of 4 mg/ml Pf1 phage (Asla Biotech) and for $^{15}N/^{2}H$ Fab/IL-1β complex (0.2 mM) and $^{15}N/^{2}H$ IL-1β/Fab complex (0.3 mM) by the addition of 3 mg/ml Pf1 phage (Asla Biotech) in the NMR buffer described previously. Backbone amide RDC values were derived from the differences between the $^{15}N/^{1}H$ scalar couplings in the isotropic and partially aligned samples using $^{15}N/^{1}H$ HSQC and TROSY spectra (Kontaxis et al., 2000) which were measured in triplicate for the free Fab. Acquisition times were 60 ms in F_2 (^{1}H) and 50 ms in F_1 (^{15}N) and experiments were collected for 5 to 12 hours. The WATERGATE pulse scheme was used to suppress the water signal when required (Piotto et al., 1992). All NMR data was processed using Topspin (Bruker Biospin Ltd.) and the spectra were analysed using Sparky (Goddard and Kneller).

4.2.2 Chemical shift mapping

The minimal shift approach (Veverka et al., 2008; Waters et al., 2007; Williamson et al., 1997) was used to identify residues in the Fab which are perturbed on IL-1β binding. The minimal combined chemical shift change (N, HN, CO) between all the assigned peaks in the HNCO spectrum of the free Fab and peaks in the HNCO spectrum of the Fab bound to IL-1β was calculated using Microsoft Excel. The combined amide proton, nitrogen and carbonyl minimal shift (Δδ) was calculated using the following equation:

$$\Delta\delta = \sqrt{(\Delta\delta_{NH})^2 + (\Delta\delta_N\alpha_N)^2 + (\Delta\delta_C\alpha_C)^2}$$

Equation 4.1

where $\Delta\delta_{NH}$, $\Delta\delta_N$ and $\Delta\delta_C$ are the amide proton, nitrogen and carbonyl differences in chemical shift between peaks in the HNCO spectra. Scaling factors of 0.20 (α_N) and 0.33 (α_C) were used to account for the differences in spectral range between the amide proton, nitrogen and carbonyl chemical shifts. The minimal shift induced on IL-1β binding was taken to be the smallest calculated combined shift. Assignments had been made previously for both free and Fab bound IL-1β so it was possible to calculate actual shifts for each residue. The combined shift (Δδ) was calculated in the same way as above for each assigned residue in HNCO spectra of free and bound IL-1β.

4.2.3 Sequence specific assignments

Sequence specific assignments (N, HN, Cα and Cβ) were obtained for the Fab from identification of intra- and inter-residue connectivities in the $^{15}N/^{13}C/^{1}H$ HNCACB,

HN(CO)CACB, HNCA, HN(CO)CA and HNCO spectra with additional data provided by sequential NOEs observed in the $^{15}N/^{1}H$ NOESY-HSQC spectrum.

4.2.4 Homology modelling

The homology modelled Fab was produced by Jiye Shi at UCB. Antibody Modeling Tool (Riley, 2005) was used to select potential templates from known antibody structures for the homology modeling of the Fab. This tool performed two BLAST (Altschul et al., 1997) searches against the Protein Data Bank (Berman et al., 2000), one with the heavy chain sequence of the Fab as the query and the other with the light chain sequence. Antibody structures that have E-values lower than 1E-10 in both searches were aligned to the Fab sequence using FUGUE (Shi et al., 2001). Each of the pairwise alignment was subsequently scored based on the sequence similarity outside of CDR, the conservation of solvent inaccessible residues, the similarity in CDR length, as well as the resolution of the experimental structure. Ten best-scored alignments were retained and the corresponding antibody structures designated as potential templates for homology modelling. Each of the ten potential templates was structurally aligned, using FATCAT (Ye and Godzik, 2003), to the NMR structure of an scFv (Wilkinson et al., 2009) that shares the same variable domain sequences as the Fab. Structures having a Cα RMSD greater than 1Å to the scFv (excluding the flexible linker between the two variable domains) were discarded. Sequence conservation at the variable-constant domain interface, defined as residues within 8Å from the opposing domain, was evaluated for each of the remaining structures. The structure with fewest interface residues non-identical to the Fab was selected as the final template. Both the final template and the structure of the scFv were supplied to MODELER (Sali and Blundell, 1993) to build a homology model of the Fab. Instructions were given so that the

variable domains were modelled based on the scFv, while the constant domains were modelled based on the final template.

4.2.5 Docking of the Fab/IL-1β complex

The structure of the Fab/IL-1β complex was determined by NMR restraint driven docking using HADDOCK (Dominguez et al., 2003). A homology model of gIC8 Fab and a free crystal structure of IL-1β (PDB code 2I1B (Priestle et al., 1989)) were used as starting structures. The docking was driven by information about residues in the interacting surfaces from chemical shift perturbation data in addition to backbone amide RDCs, HN-HN NOEs and dihedral angles restraints determined from CA, CB, CO and N chemical shifts. The interaction surfaces were defined as active or passive ambiguous interaction restraints (AIRs). Active AIRs are solvent exposed residues that have been shown experimentally to be involved in the interaction whilst passive AIRs are defined as solvent exposed neighbours of active residues. Residues in IL-1β perturbed by more than 0.15 ppm on binding to gIC8 Fab and with a solvent accessibility of greater than 10% calculated using MolMol (Koradi et al., 1996) were defined as active. Due to extensive chemical shift changes throughout the Fab on IL-1β binding, active Fab residues were defined as CDR residues that were perturbed by more than 0.05 ppm on IL-1β binding with solvent accessibility of greater than 10%. For the Fab 23 active and 14 passive residues were selected and for IL-1β 26 active and 15 passive were defined as shown below.

gIC8 active: 27, 30, 31, 92, 93, 94, 95, 245, 246, 247, 267, 268, 269, 270, 271, 272, 314, 315, 316, 317, 318, 321, 322

gIC8 passive: 1, 26, 28, 32, 50, 216, 217, 241, 242, 244, 260, 273, 274, 286

IL-1β active: 5, 6, 15, 39, 46, 48, 49, 50, 52, 54, 55, 56, 92, 93, 105, 106, 107, 108, 109, 147, 148, 149, 150, 151, 152, 153

IL-1β passive: 1, 2, 4, 11, 13, 14, 36, 37, 41, 51, 53, 91, 94, 96, 145

Residues at the interaction sites and at the interfaces between Fab domains may experience conformational changes on complex formation. In the docking calculation these residues were defined as semi-flexible allowing them to move during simulated annealing step. Active and passive residues +/- 2 sequential residues were defined as semi-flexible to allow for conformational changes in the antigen binding region. Residues involved in the Fab domain-domain interfaces were identified as residues that gain more than 5% solvent accessibility on the removal of a proximal domain in MolMol (Koradi et al., 1996). These residues +/- 2 sequential residues were also defined as semi-flexible. All the semi-flexible residues are shown below.

gIC8 semi-flexible regions: 1-9, 18-51, 63-74, 78-111, 214-219, 222-227, 239-289, 292-298, 307-335, 108-113, 138-143, 163-175, 112-140, 158-169, 172-182, 205-216, 332-338, 364-371, 420-424, 339-365, 382-405, 427-438

IL-1β semi-flexible regions: 1-17, 34-58, 89-98, 103-111, 143-155

Residual dipolar couplings from Fab/IL-1β complex were assigned by pattern matching from the assigned spectra of the free Fab and used as direct restraints (SANI). The axial (A_a) and rhombic (R) alignment tensor components were obtained from RDCs collected from the IL-1β crystal structure 2I1B (Priestle et al., 1989) using PALES (Zweckstetter

and Bax, 2000). Intervector projection angle restraints (VEAN) (Meiler et al., 2000) were also derived from the RDC data and used in the docking and refinement. Long range HN-HN NOEs from the Fab were used as distance restraints in the docking calculations. Distances between residues were determined based on peak height and NOEs were divided into groups corresponding to ^{1}H-^{1}H distances of <5 Å, 5 to 6.5 Å and 6.5 – 8 Å. The average peak heights for sets of *i*+1 and *i*+2 peaks were taken to be the thresholds for the <5 Å and 6.5 Å respectively.

Phi and psi dihedral angles were determined from CA, CB, CO and N chemical shifts for both IL-1β and the Fab were used as additional restraints in the docking. The Cα and Cβ chemical shifts were corrected for deuteration effects and the dihedral angles were calculated in TALOS (Cornilescu et al., 1999).

The docking protocol follows three stages: (i) randomisation of the starting orientation and rigid body energy minimisation (ii) semi-flexible simulated annealing and (iii) final refinement with an explicit solvent. In the first step the molecules were separated by at least 25 Å and randomly rotated around their centre of mass. Each protein was then allowed to rotate in turn to optimise relative orientation followed by a period where translations and rotations were allowed. The proteins were docked by rigid body energy minimisation generating 1000 docked structures of which the best 200 structures were further refined. The simulated annealing step followed three stages. Firstly the relative orientations of the proteins were optimised, followed by a stage where the side chains were allowed to move, finishing with a stage where both sidechains and the backbone were allowed to move. In the final docking step, the structures were refined in an 8 Å water layer. Intervector projection angle restraints were used in the rigid body energy

minimisation and simulated annealing steps and the RDCs were used as direct restraints in the final water refinement step. 132 satisfactorily converged structures were obtained from the 200 docked results.

4.2.6 Refinement of a structure for the free IC8 Fab

A refined structure of the Fab was obtained using HADDOCK. Residues in domain interfaces and CDR residues +/- 2 sequential residues and were defined as being semi-flexible (as listed previously). Phi and psi dihedral angles were determined from CA, CB, CO and N chemical shifts, long range HN-HN NOEs NOEs and backbone amide RDCs were used as restraints in the structural refinement. RDCs were used as direct restraints and as intervector projection angle restraints (Meiler et al., 2000). A set of 200 docked structures were produced and 115 satisfactorily converged complex structures were obtained from these. Analysis of the structures was carried out in MolMol (Koradi et al., 1996) and the APBS tool in PyMol was used to calculate the electrostatic surface potential of the Fab and a structure of the IgG Fc (1FC1 (Deisenhofer, 1981)).

4.2.7 Structural analysis

Backbone atom (N, Cα, CO) root mean square deviation (RMSD) values were calculated to the mean structure and were obtained by superposition of backbone atoms of residues in elements of regular secondary structure. The residues in regular secondary structure are as follows:

IL-1β: 5-12,16-22,25-30,34-38,42-46,55-62,67-74,77-84,97-99,101-105,110-114,121-123,130-131,133-135,145-150

gIC8 Fab: 5-12,16-22,25-30,34-38,42-46,55-62,67-74,77-84,97-99,101-105,110-114,121-123,130-131,133-135,145-150,157-160,163-166,173-178,186-191,197-201,215-220,223-228,238-243,256-259,267-271,275-278,283-190,297-303,306-307,312-316,327-334,336-340,344-351,358-363,370-374,378-379,385-392,401-406,412-417,435-440,445-450,459-465,475-476,480-484,493-497,501-505,508-516,524-527,536-538,542-543,549-561,568-576,578-585.

4.2.8 Statistical analysis of the structural difference between free and bound Fab

An unpaired Student's t-test (Van Emden, 2008) was performed to test the significance of the difference between backbone Cα atom coordinates from 30 representative low energy structures of free Fab and 30 representative low energy structures of bound Fab. Differences in the Cα atom coordinates between the free and bound structures were deemed to be statistically significant if the resulting p values were equal to or less than 0.05.

4.3 Results and Discussion

4.3.1 Sequence specific assignments for gIC8 Fab

Deuteration of the gIC8 Fab, as described in Chapter 3, gave rise to well resolved NMR spectra, as illustrated in the TROSY spectrum of $^{15}N/^{2}H$ gIC8 shown in Figure 4.1. Triple resonance spectra (HNCO, HNCA, HNCACB, HNCOCA and HNCOCACB) were obtained for the Fab which allowed for the assignment of 91% of the visible (84% of the total) backbone atoms (HN, N, Cα, Cβ) of $^{15}N/^{13}C/^{2}H$ gIC8. The representative set of strips from the HNCACB spectrum (Figure 4.2) and the HNCO spectrum is shown in Figure 4.3 both underline the high quality of the triple resonance data obtained. The excellent signal to noise ratio in the spectra (highlighted in Figure 4.2) allowed for all connections between Cα and Cβ resonances for stretches of typically between 6 to 40 residues as shown in A.1. The 3D spectra also only contain minimal spectral overlap which is shown very clearly in Figure 4.3. A set of ^{15}N strips from the $^{15}N/^{1}H$ NOESY-HSQC spectrum, used to assist in the backbone atom assignments, are shown in Figure 4.4B. ^{1}HN-^{1}HN NOEs are clearly visible for *i* to *i*+2 and *i* to *i*+3 indicating NOE distances of 7 to 8 Å compared to the maximum of 5-6 Å that would be visible if the protein was protonated. Distances like these can be observed due to the reduced spin diffusion and T_1 relaxation of the highly deuterated protein making it possible to use long mixing times of 500 ms (Koharudin et al., 2003).

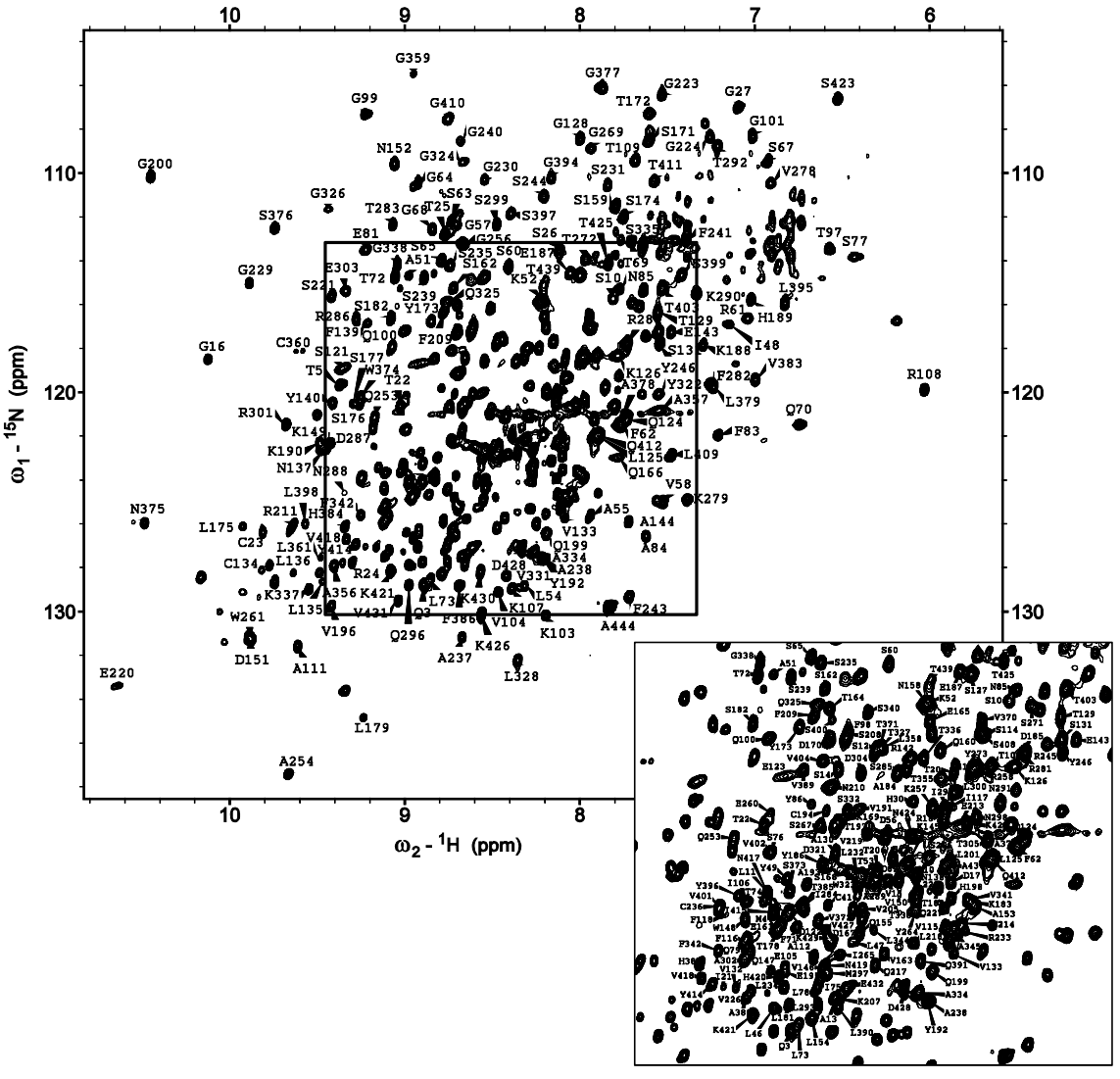


Figure 4.1

^{15}N/^{1}H TROSY spectrum of ^{15}N/^{2}H-labelled gIC8 Fab. The assignment of the signals from the backbone amide groups are indicated by residue type and number. The insert shows an enlargement of the highly populated central region of the spectrum.

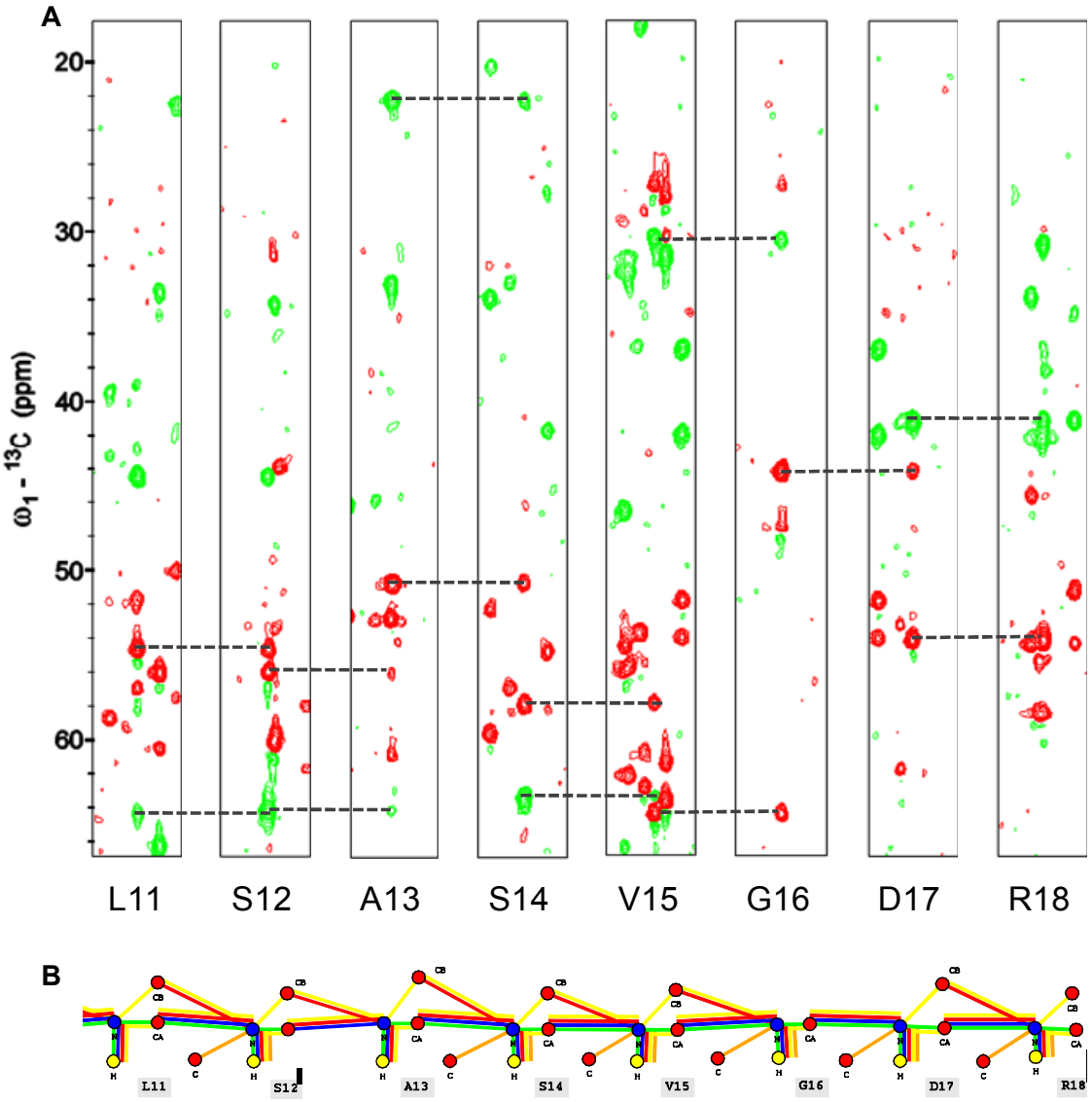


Figure 4.2

^{15}N strips from the HNCACB spectrum of ^{15}N/^{13}C/^{2}H gIC8 Fab. Panel A illustrates the intra- and inter-residue cross peaks linking residues L11 to R18. Cα signals are shown in red and Cβ signals are shown in green. Dashed lines indicate the Cα and Cβ connections. Panel B shows a schematic representation of the connectivities that were used to make backbone assignments for this section of the sequence. Coloured circles represent backbone atoms (red for carbon, blue for nitrogen and yellow for hydrogen) and atom types are labelled. Lines connecting the atoms represent triple resonance

spectra used to make the sequence specific backbone assignment with HNCA represented using a green line, HNCOCA with blue, HNCOCACB with red, HNCACB with yellow and HNCO with orange. A summary of the connectivities in the complete sequence is shown in A.1.

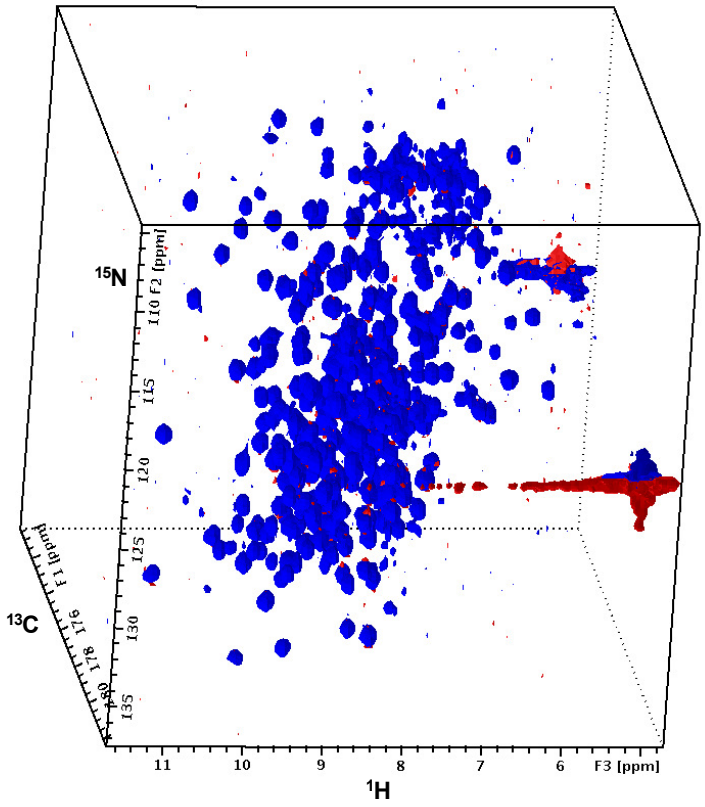


Figure 4.3

View of the complete HNCO spectrum of $^{15}N/^{13}C/^{2}H$ gIC8 Fab. The cube view of the HNCO spectrum illustrates the good spectral resolution that was seen in all the triple resonance spectra that were obtained.

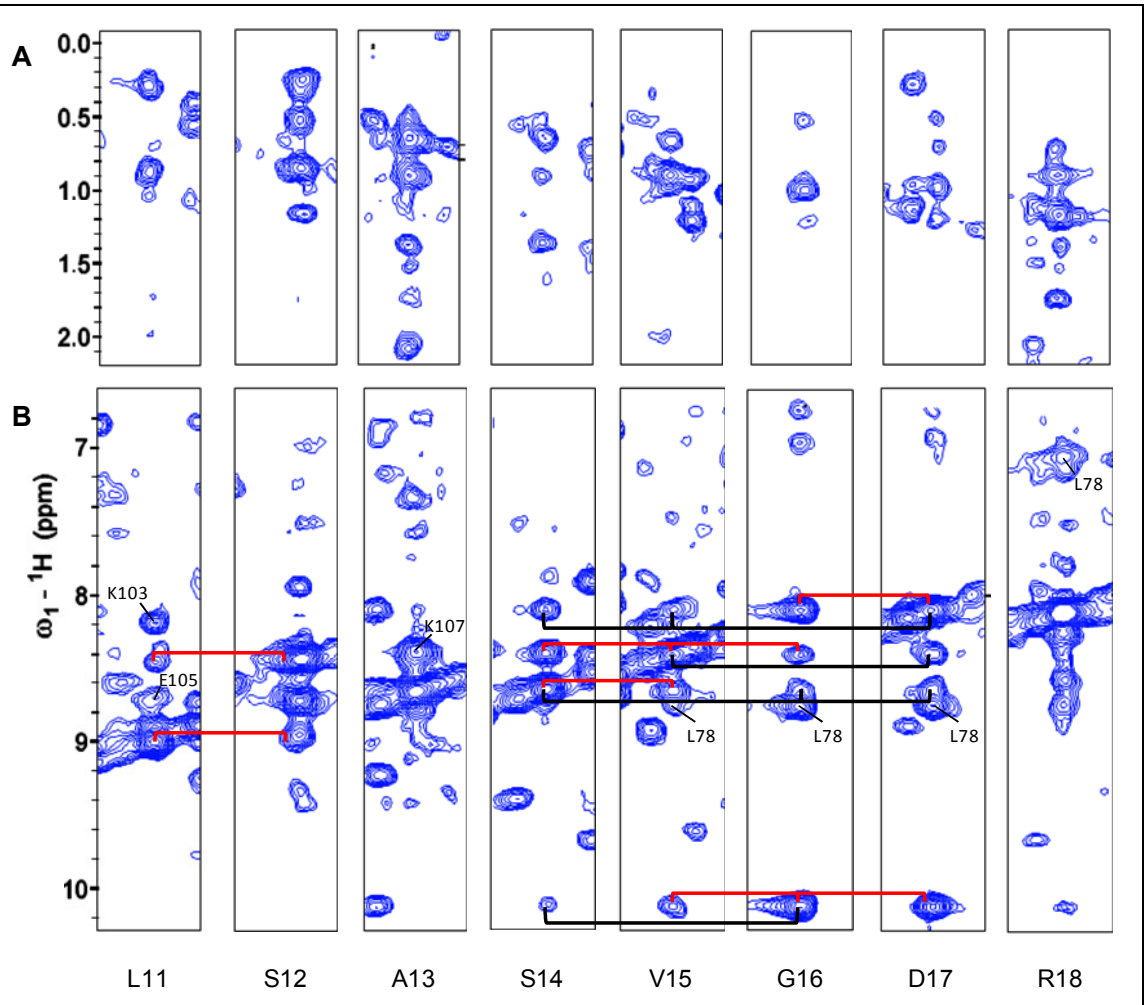


Figure 4.4

^{15}N strips from the NOESY-HSQC spectrum of ^{15}N/^{2}H gIC8 Fab. Panel B shows ^{15}N strips from the NOESY-HSQC spectrum of ^{15}N/^{2}H gIC8 Fab illustrating NOE connections between residues L11 to R18. Sequential HN-HN NOEs are indicated in red and black lines indicate *i* to *i*+2 and *i* to *i*+3 HN-HN NOEs. Residues L11 to S14 are within a β strand and so the fact that *i* to *i*+3 HN-HN NOEs are visible means that NOE distances of up to 8 Å are being seen in the spectrum. Also indicated are long range NOEs to residues further away in the primary sequence. No sequential NOEs are seen for A13 as the crosspeaks are obscured by the diagonal. Panel A shows amide proton to methyl group NOEs in the residues as a result of selective protonation of methyl groups.

The gIC8 Fab protein was expressed in minimal media containing 100% D_2O and protonated glucose was used as a carbon source. Using this labelling strategy ~60-80% of the non-exchangeable proton positions are replaced with deuterium (Cavanagh et al., 1996), however, the isotopic enrichment is not random. Studies on five proteins expressed using the same minimal media composition showed that protonation was sequence independent, side chain specific and selective. Analysis of $^{13}C/^{1}H$ HSQC spectra acquired for the proteins showed the methyl groups of Ala, Val, Leu, Ile, Met and Thr, and a methine group of Thr had peak intensities corresponding to 55% to 64% protonation. In addition, methylene groups of Arg (δ), Ile (γ), Lys (γ, δ, ε), Pro (δ), Ser (β), Asp (β), Glu (β, γ), Asn (β) and Gln (β, γ) showed peak intensities correlating to 27% to 36% protonation. The methyl groups existed in two isotopomers, CHD_2 (73%) and CH_2D (22%), and for the methylene groups CHD (87%) constituted the major isotopmer and less than 10% of the methyl and methelene groups were fully protonated (Leiting et al., 1998). The result of this labelling strategy when applied to $^{15}N/^{2}H$ gIC8 Fab can be seen in Figure 4.4A, which shows strips from the NOESY-HSQC spectrum with connections to the selectively labelled methyl and methylene groups. Figure 4.5 shows a cube view of the NOESY-HSQC spectrum which again highlights NOEs to the protonated methyl groups. This selective labelling was straightforward to achieve and has great potential for use in identifying residue types, which will be very valuable in assisting backbone assignment of large proteins and determination of global folds. For gIC8 Fab, the extra information was not used due to time restraints, however, with further optimisation this technique could be used on future projects.

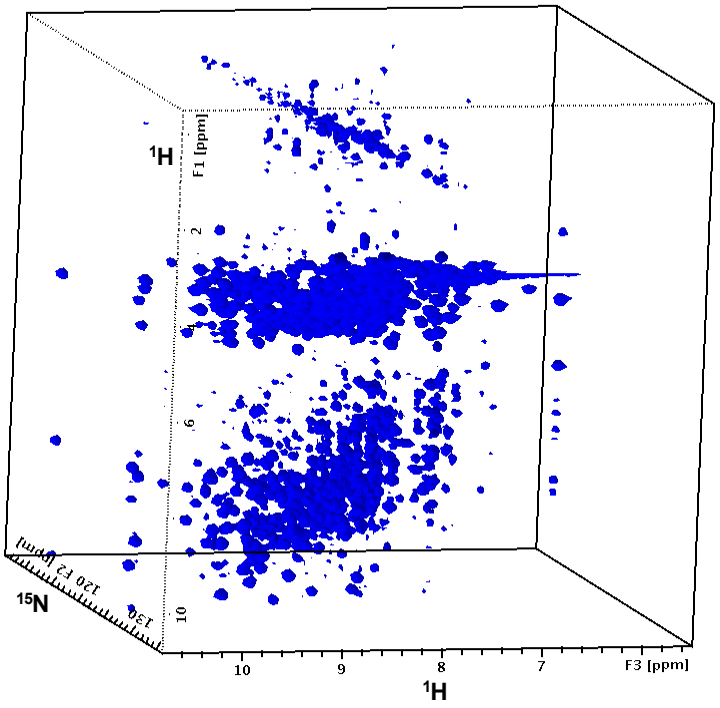


Figure 4.5

View of the complete NOESY-HSQC spectrum of ^{15}N/^{2}H gIC8 Fab. A cube view of the NOESY-HSQC spectrum highlights HN-HN NOEs and also shows HN-CH_3 connections to methyl groups.

As mentioned previously, some resonances that can be been in the spectra of the ^{15}N/^{1}H gIC8 are not visible in the spectra obtained from the ^{15}N/^{2}H sample. In the HNCO spectrum of ^{15}N/^{2}H gIC8, 347 signals of a potential 419 from non-proline signals are visible. After expression of the gIC8 in minimal media containing 100% D_2O, the protein was purified in H_2O containing buffers to back exchange amide protons. However, an issue for the expression of deuterated large proteins is that amide protons involved in strong hydrogen bonds or those buried inside the protein and not accessible to the solvent can remain deuterated (Cavanagh et al., 1996). Protocols are described

where the protein is unfolded using chemical denaturants and refolded in the presence of H_2O to ensure complete back exchange at the amide proton positions (Venters et al., 1995). In this case we were concerned that chemical denaturation may not be reversible and incubation of the protein at 65 °C showed no increase in D/H exchange. A $^{15}N/^{1}H$ HSQC spectrum of $^{15}N/^{1}H$ gIC8 allowed to incubate in buffer containing 100% D_2O for 12 hours (Figure 4.6), showed 57 peaks from backbone amide protons with a slow H/D exchange rate. A proportion of these peaks are not seen in the $^{15}N/\ ^{1}H$ HSQC spectrum of $^{15}N/\ ^{2}H$ gIC8 in buffer containing 100% H_2O, even though this sample had been allowed several months for back-exchange during the collection of triple resonance experiments.

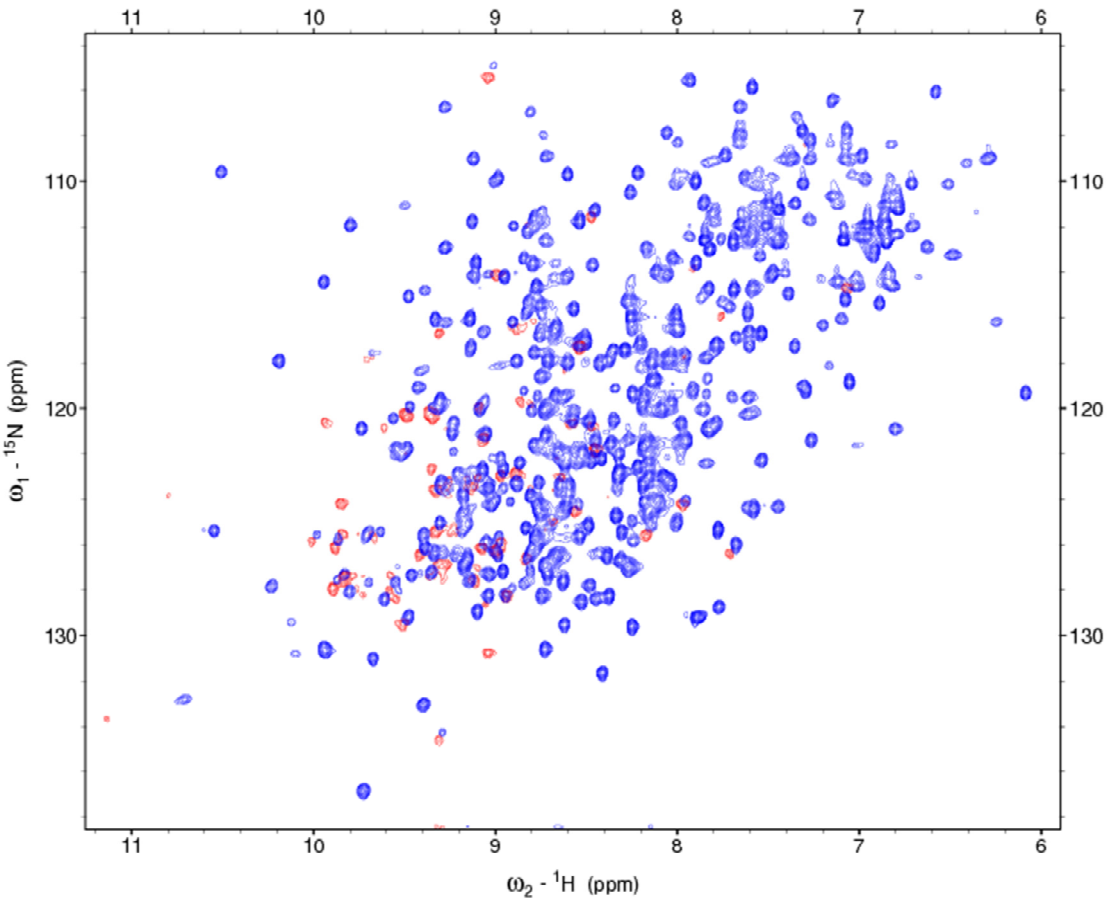


Figure 4.6

Amide protons in the backbone of deuterated Fab that do not back-exchange. Overlaid $^{15}N/^{1}H$ HSQC spectra of $^{15}N/^{1}H$ gIC8 recorded in D_2O in red and $^{15}N/^{2}H$ gIC8 recorded in H_2O in blue.

The structures of the Fab variable and constant domains are different. The two β pleated sheets in the constant domain consist of four (A-B-E-D) and three (G-F-C) antiparallel β strands. The variable domains have the same four stranded sheet as the constant domain (A-B-E-D) but there is an additional two strands on the second sheet (G-F-C-C'-C'') (Figure 4.7B) (Ramsland and Farrugia, 2002). Figure 4.7A shows the sequence of gIC8

with the assigned residues highlighted. The secondary structure is indicated above the sequence with core β strands, defined as those positioned between two other strands, highlighted with dark blue arrows. This shows that there are stretches of unassigned residues in strand C in both the variable domains, in addition to strand F in the V_H domain. It is highly possible that the amide protons in these core β strands are involved in tight hydrogen bonds and therefore did not back exchange when the protein was purified in H_2O.

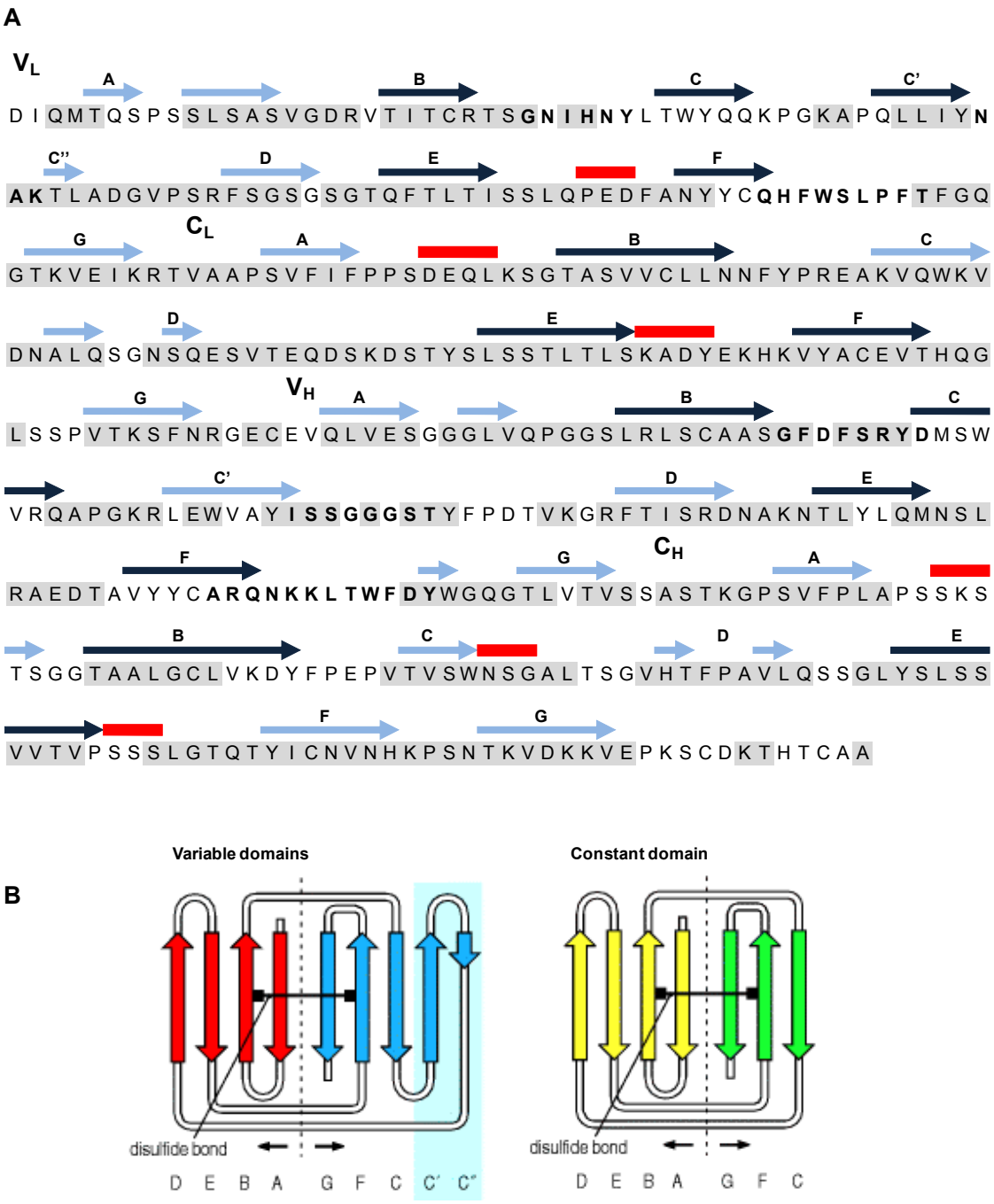


Figure 4.7

Summary of backbone (NH, H, Cα and Cβ) assignments made for gIC8 Fab. Panel A shows the amino acid sequence of gIC8 Fab with assigned residues highlighted in grey. The secondary structure of the Fab is shown above the sequence with β-sheets represented as blue arrows, core β strands, defined as those sandwiched between two

other strands, are shown in dark blue and helices are shown as red blocks. The β strands are labelled as described in Panel B. The Fab domains are indicated and residues in the CDRs are highlighted with bold print (defined using the IMGT numbering system (Giudicelli et al., 1997)). Panel B shows a general, schematic structure of the immunoglobulin constant and variable domains to demonstrate the positioning of core β strands. The β strands are lettered sequentially with respect to their order of occurrence in the amino acid sequence. The β strands C' and C'', which are seen in the variable but not constant domains, are highlighted with a light blue background. Panel B is from (Janeway, 2005).

Figure 4.7A also highlights that a proportion of the unassigned residues occur in CDRs specifically CDR3 in both the V_H and V_L. NMR is extremely sensitive to the exchange of a nucleus between conformational states. Nuclei in slow exchange show two peaks in a spectrum corresponding to the two separate states, while nuclei in fast exchange show a single resonance at the population weighted average of the chemical shifts of the nuclei in the two sites. As the exchange rate decreases, the linewidths broaden so that nuclei in intermediate exchange show very broad peaks that become indistinguishable from the baseline and disappear (Cavanagh et al., 1996). It is therefore possible that resonances from CDR resonances are not visible in the spectra as the residues in these loops are in intermediate exchange between 2 or more conformations. This could be validated by recording the NMR experiment at a higher temperature which would increase the exchange rate showing a single peak in the spectrum. Alternatively, the spectrum could be obtained at a lower temperature which would result in a decrease in

the exchange rate and peaks corresponding to the separate states would be visible in the spectrum.

The immune system is able to generate antibodies which are highly specific for virtually any antigenic target due to the diversity in the CDRs generated by V(J)D recombination, junctional diversity and somatic mutations introduced by antigen binding. Affinity maturation may require major conformational changes of the antigen binding surface to produce antigen specific, high affinity antibodies and can involve significant changes in conformational flexibility of the CDRs (Jimenez et al., 2003). Several models of molecular recognition have been proposed which are dependent on molecular flexibility. The "lock and key" model assumes that no structural change is induced in the binding partners on complex formation and so both are rigid and exist preformed in the binding conformation. In the "induced fit" model, binding partners form an unoptimised complex and then structural flexibility allows molecular reorganisation to optimise the binding interaction. A related model, "conformational selection" assumes that due to molecular flexibility, molecules transiently exist in the correct conformational state for an optimal interaction and these are selected for complex formation (Berger et al., 1999). It has been suggested that the somatic mutations involved in affinity maturation increase affinity for the antigen by reducing flexibility in the antigen binding site (Jimenez et al., 2003). A sequence of mutations which increase the binding affinity of the antibody 4-4-20 with its antigen, fluorescein, have been deduced. These show that the flexibility of the binding pocket decreases as antigen affinity increase (Jimenez et al., 2004). A mechanism is proposed by which initial germline antibodies have flexible antigen binding sites which can adopt many different combinations and interact with a range of epitopes in an induced fit type

mechanism. Affinity maturation somatic mutations rigidify the CDRs into an epitope specific binding site so that interactions can occur via a lock and key type mechanism. This is supported by the observation that somatic mutations occur in residues that are not directly involved in the interaction with fluorescein, suggesting that their role is to stabilise the binding site (Thorpe and Brooks, 2007). Crystal structures obtained for a germline Fab fragment, both free and bound to its antigen, nitrophenyl phosphonate, were compared to crystal structures of free and bound forms of the affinity matured Fab, 48G7 (Wedemayer et al., 1997). The structures showed that significant reorganisation of the binding site residues occurred on hapten binding to the germline antibody, while very few structural changes were induced on 48G7 binding to the antigen, again supporting the idea that antigen binding in affinity matured antigen occurs by an induced fit mechanism. The gIC8 Fab is an affinity matured antibody, however, flexibility can be seen in the binding site of the free protein suggesting that some degree of induced fit is employed on gIC8 binding to IL-1β. Evidence of the mutations that drove affinity maturation of gIC8 is unavailable, and so it is not possible to determine if gIC8 flexibility was reduced during maturation, however, this result shows clearly that the CDR3 loops in an affinity matured antibody are in intermediate exchange in the free Fab structure. In an HNCO spectrum collected for $^{15}N/^{13}C/^{2}H$ gIC8 in complex with unlabelled IL-1β a total of 379 signals were visible, 32 more than then HNCO spectra obtained from free Fab. This suggested that CDRs which are in intermediate exchange in the free Fab, are stabilised in one discrete conformation in the Fab complex and resonances for these residues are therefore visible in the spectra.

This was supported by NMR data collected for the equivalent scFv/IL-1β complex where it was possible to make essentially complete backbone assignments for the CDRs

(Wilkinson et al., 2009). Comparisons of the scFv/IL-1β, free Fab and Fab/IL-1β HNCO spectra revealed 17 peaks (N31, F91,W92, S93, L94, F96 in the V_L and D247, G268, G270, Q313, N314, K315, K316, L317, T318, W319, F320 in the V_H) that had been assigned as CDR residues in the scFv complex and had corresponding peaks in the Fab/IL-1β spectrum. However, these residues had no visible signals in the HNCO spectrum of the free Fab (Figure 4.8). In addition, residues L33 and M248, both of which are in positions adjacent to CDRs, and 4 framework residues (F274, D276, T277 and G280) showed signals in the complex spectra that were absent from the spectra of the free Fab.

All of the framework residues shown to be stabilised on complex formation with the Fab are in a region of the variable domain with nonregular secondary structure. In the structure of the corresponding scFv/IL-1β complex, three framework residues were seen at the protein interface, D1, D276 and K279. There is no evidence that residue D1 in the Fab interacts with IL-1β as no chemical shift perturbation data is available for this residue. It is however, possible to speculate that residues D276 and K279 could interact with IL-1β in the Fab as shown in the structure of the scFv/IL-1β complex. Interactions made with these residues could act to stabilise a region of the backbone that was in intermediate exchange between two or more forms in the free state, consequently making the Fab framework residues F274, D276, T277 and G280 visible in the Fab HNCO spectrum.

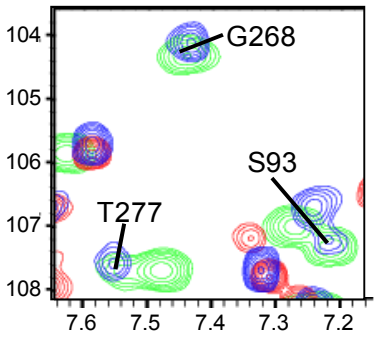


Figure 4.8

Identification of residues that are in intermediate exchange in the structure of the free Fab and are stabilised into one position on binding to IL-1β. A small region from the overlaid $^{15}N/^{1}H$ HSQC spectra of free $^{15}N/^{2}H$-labelled Fab in red, $^{15}N/^{2}H$ Fab-unlabelled IL-1β complex in blue and $^{15}N/^{1}H$-labelled scFv/unlabelled IL-1β complex in green. Labelled assignments indicate CDR residues which are visible in the scFv/IL-1β and Fab/IL-1β spectra but are missing in the free Fab spectra. This suggests that these CDR residues are in intermediate exchange in the free Fab so no peaks are visible in the NMR spectrum. On antigen binding the CDRs are stabilised into one conformation in the complex. Figure 4.11B shows a summary of the residues that are in intermediate exchange in the free Fab and are stabilised into one conformation on IL-1β binding.

It is not uncommon for framework regions to be involved in the interaction surface. Indeed, analysis of the interface properties of eight scFv or Fab complex structures showed that three (2GHW, 1LK3, 1DZB) had interactions with the target protein involving framework residues in addition to residues in the CDR loops (Wilkinson et al., 2009). It is clearly important in the rational design of therapeutic antibodies to determine

whether framework residues play a role in antigen binding, as changes in these residues would presumably have a negative effect on antibody affinity.

4.3.2 Identification of residues with signals perturbed by binding.

Backbone assignments of IL-1β bound to gIC8 Fab and free IL-1β had been previously made in the group. These were used to determine actual combined backbone HN, N and CO chemical shift perturbations from HNCO spectra. Figure 4.9 Panel A shows a histogram of the combined chemical shift changes and indicates a specific interaction site involving 25 residues with a chemical shift change of greater than 0.3 ppm (S5, L6, N7, C8, G9, L10, D12, S17, V19, V40, S45, F46, S52, D54, I56, E105, I106, N107, N108, K109, Q149, F150, V151, S152 and S153).

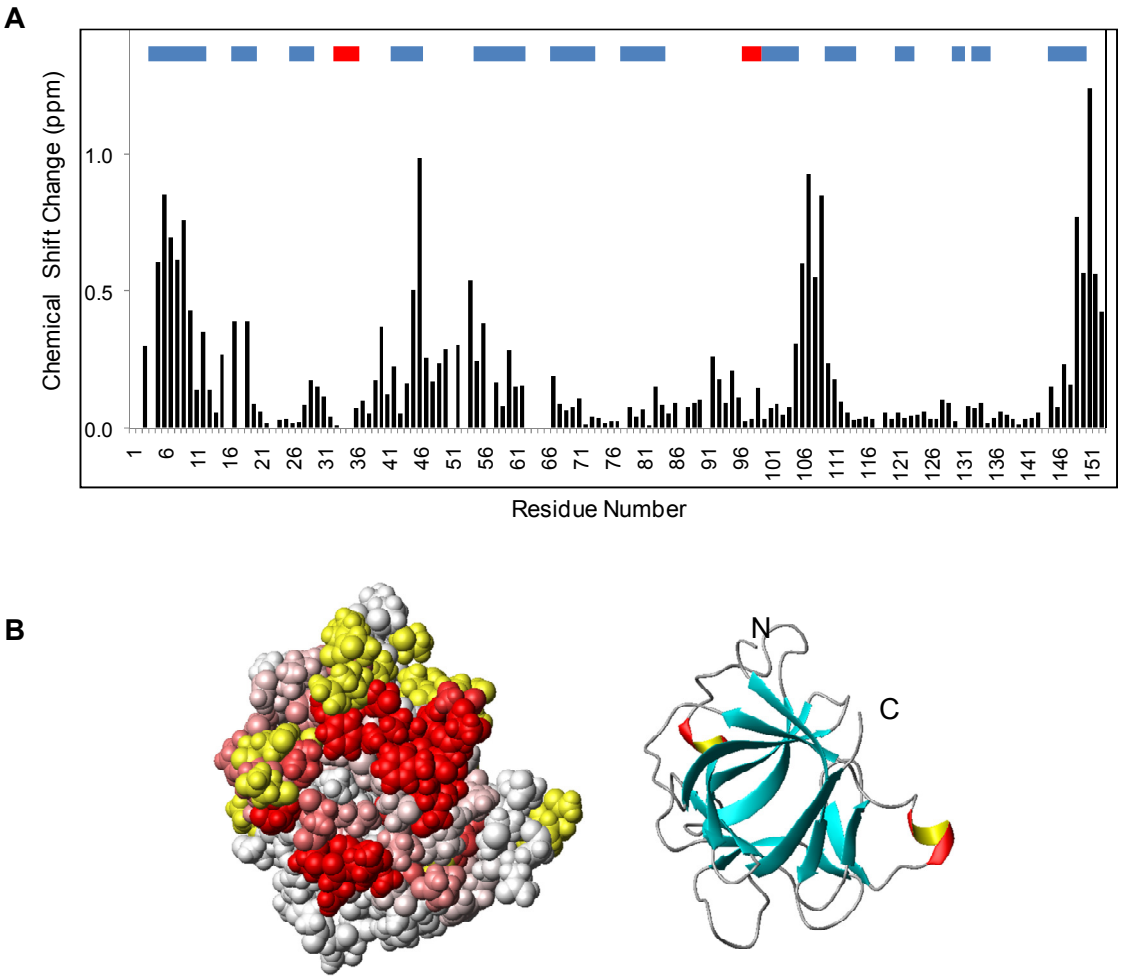


Figure 4.9

Perturbation of HN, N and CO signals from IL-1β on binding to IC8 Fab. The histogram in Panel A summarises the combined backbone HN, N and CO chemical shifts changes seen in HNCO spectra of IL-1β on binding to IC8 Fab. The secondary structure of IL-1β is indicated above the plot with red bars for helices and blue bars for β-sheets. Panel B highlights residues that show significant chemical shift differences between free and bound IL-1β on a space-filled view of the protein. Residues which show a minimal shift change of less than 0.1 ppm are coloured white, residues perturbed by more than 0.5 ppm are shown in red and residues showing chemical shift changes between 0.1 ppm and 0.5 ppm are coloured on a linear gradient between white and red.

Residues with no available chemical shift perturbation data are shown in yellow. A ribbon representation of IL-1β in the same orientation as the space filled model is shown to the right.

Residues involved in the interaction between gIC8 Fab and IL-1β were identified from chemical shift changes seen for peaks in HNCO spectra of free and bound $^{15}N/^{13}C/^{2}H$ protein. As no backbone assignments were available for the gIC8 in the complex, combined minimal shift changes in HN, N and CO were used to identify perturbed residues (Veverka et al., 2008; Waters et al., 2007; Williamson et al., 1997). The HNCO spectrum shown in Figure 4.3 emphasizes how little spectral overlap is present in the 3D data, which resulted in accurate chemical shift perturbation data that could be interpreted with confidence. The results of the minimal combined shift analysis are shown in the histogram in Figure 4.10. Significant chemical shift changes on IL-1β binding can be seen in CDR loops, where NMR data is available. In addition to chemical shift differences in CDR residues, significant shifts in signals are also seen for residues at the V_L-C_L interface. Noticeably, there are no chemical shift changes in either constant domain, apart from residues Q166, D170 and S171, which are located at the V_L-C_L interface.

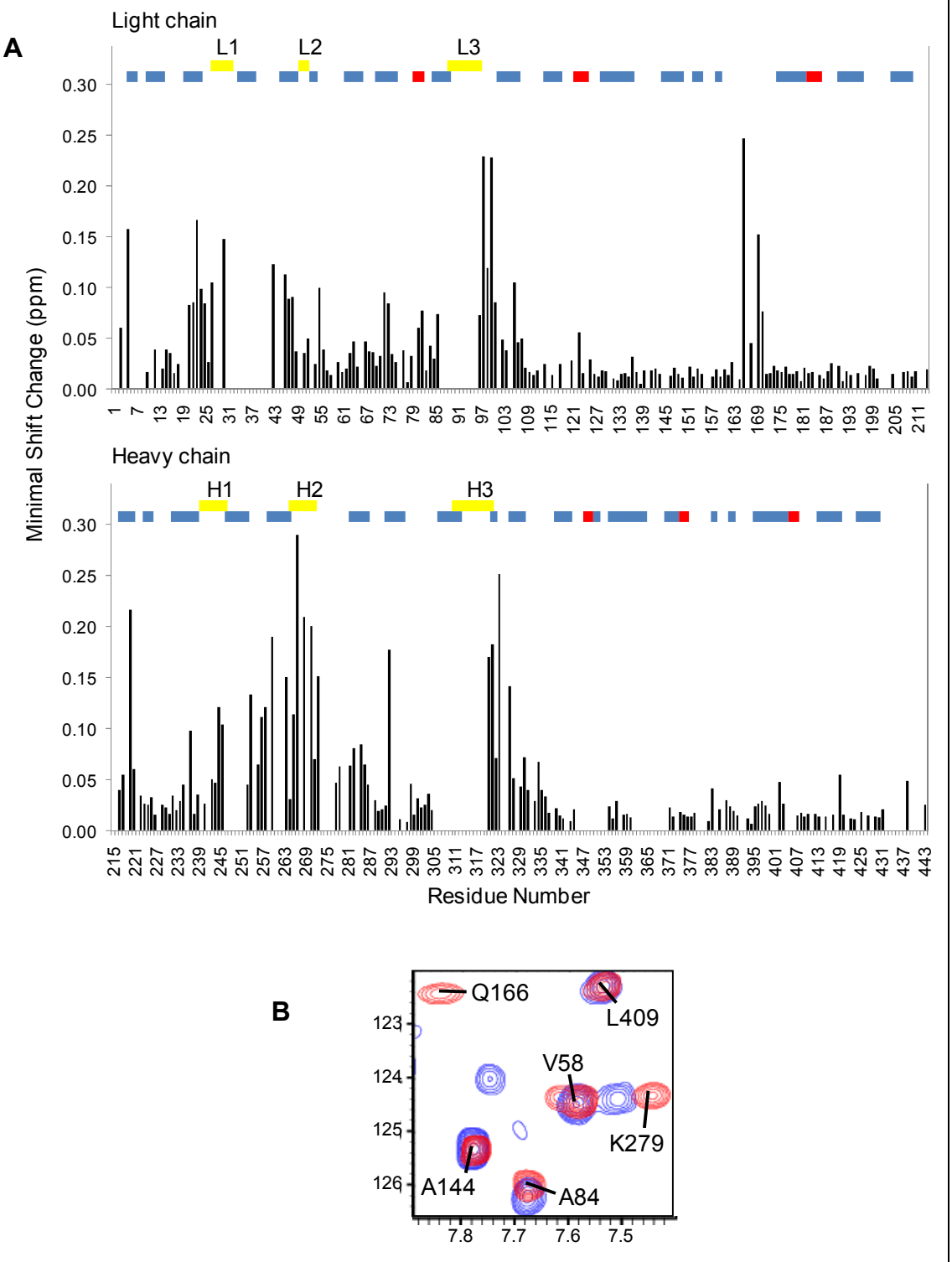


Figure 4.10

Minimal perturbation of HN, N and CO signals from gIC8 Fab on binding to IL-1β. The histograms in Panel A summarize the combined backbone HN, N and CO

minimal shifts seen for the Fab light chain (top) and heavy chain (bottom) on binding to IL-1β. The secondary structure of the Fab is indicated above the plot with red bars for helices and blue bars for β-sheets. The positions of the CDRs in the light (L1, L2 and L3) and heavy chain (H1, H2 and H3) are indicated with yellow bars (defined using the IMGT numbering system (Giudicelli et al., 1997)). Panel B shows a small region from a TROSY spectrum of free $^{15}N/^{2}H$-labelled Fab in red and $^{15}N/^{2}H$ Fab-unlabelled IL-1β complex in blue with the assignments for the free protein indicated. Signals from some residues, for example L409 and A144, show no significant shifts on antigen binding, whilst the backbone amide peaks for Q166 and K279 show significant antigen binding induced changes.

Figure 4.11A shows the chemical shift perturbation data mapped onto the surface of gIC8 Fab, highlighting the chemical shift changes seen at the V_L-V_H and V_L-C_L interfaces. Unlike IL-1β, a discrete binding surface cannot be defined for the Fab structure using chemical shift perturbations alone, as significant chemical shift changes are seen throughout the variable domains. It is, therefore, unclear whether chemical shifts are seen as a result of a direct interaction with IL-1β or due to a structural change that was induced by binding. Figure 4.11B shows the chemical shift perturbation data mapped onto the surface of gIC8 and also highlighting residues which were shown to be in intermediate conformational exchange in the free Fab but stabilised into one conformation on binding to IL-1β.

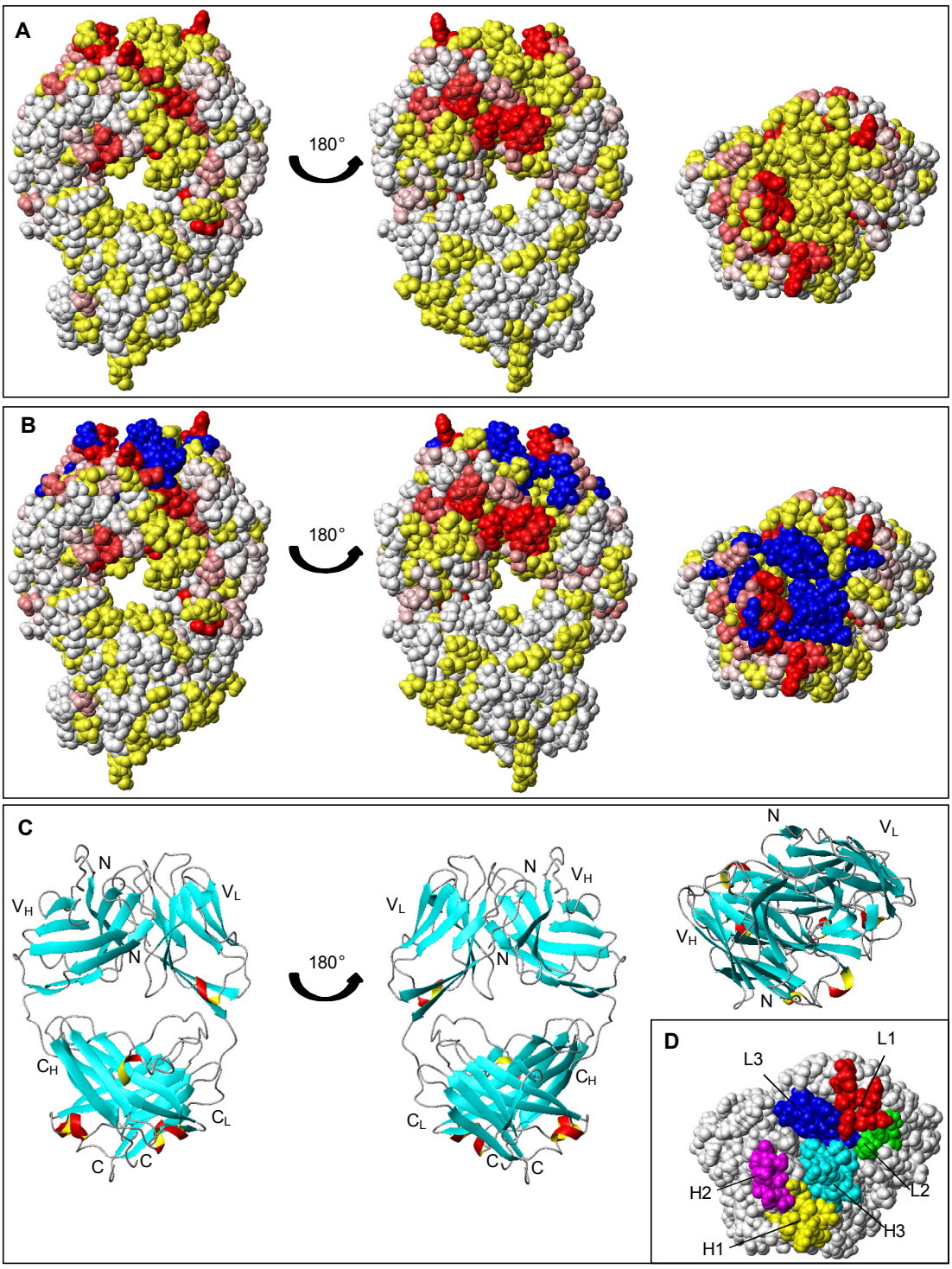


Figure 4.11

Localisation of gIC8 Fab residues that are significantly perturbed on IL-1β binding. Panel A highlights residues that show chemical shift differences between free

and bound gIC8 Fab on a space-filled view of the Fab in three orientations. Residues that showed a minimal shift change of less than 0.03 ppm are shown in white, minimal shifts of greater than 0.12 ppm are in red and between 0.03 and 0.12 ppm are coloured according to the magnitude of the shift on a linear gradient between white and red. Residues with no available chemical shift perturbation data are highlighted in yellow. Panel B shows the same space-filled views of the Fab as in A but also highlights in blue the CDR residues that are in intermediate exchange in the free Fab but are stabilised into one conformation on IL-1β binding. Ribbon representations of the Fab in the same orientation as in Panels A and B are shown in Panel C. Panel D shows a space filled view of the IL-1β binding site with the CDRs highlighted and in the same orientation as the far right view in Panels A, B and C. The light chain CDRs 1, 2 and 3 are coloured red, green and blue and the heavy chain CDRs 1, 2 and 3 are coloured yellow, magenta and cyan respectively.

The mechanism by which antigen binding to the immunoglobulin domain of a B cell receptor induces a cell signalling pathway in B cells is not understood. Current opinion is that structural changes induced by antigen binding triggers intracellular signalling, but it is unclear how this structural change is propagated through the cell membrane. One possible explanation is that structural changes induced by antigen binding send a structural signal though the antibody domains, like a ripple through the structure.

The chemical shift perturbation shows very clearly that no structural signal is propagated to the Fab constant domains, as only a select few residues show significant chemical shift perturbations on antigen binding and these are specifically positioned at

the V_L-C_L interface. It is, therefore, unlikely that structural changes can be passed through the Fc constant domains as suggested above, since no signal was passed from the variable domains though to the Fab constant domains. The minimal shift data does, however, show that significant structural changes are induced by antigen binding in the variable domains. Chemical shift changes were expected in the CDRs as these are in the established antigen binding site, however, if the membrane bound antibody in the BCR acted purely as a antigen acceptor, significant structural changes would not necessary be expected at the V_L-V_H or the V_L-C_L interface. It is also interesting that no significant chemical shift changes were observed at the V_H-C_H interface. An explanation for this pattern of chemical shift changes is that on IL-1β binding to gIC8 Fab the entire V_L domain changes in orientation relative to the constant and V_H domains, which remain fixed in position. This would result in chemical shift changes at the V_L-V_H interface as the nuclei here would be experiencing a change in environment. In this model, the V_L-C_L interface would act as a pivot point for the V_L domain which would explain the chemical shift changes seen at this interface, both in the variable and constant domains. Since the constant and V_H domains would remain in fixed positions relative to each other, no chemical shift changes would be expected at the V_H-C_H domain interface.

4.3.3 Analysis of NMR restraints

4.3.3.1 Residual dipolar couplings

Backbone amide residual dipolar couplings (RDCs) were determined for free $^{15}N/^{2}H$ gIC8 Fab, $^{15}N/^{2}H$ gIC8 Fab bound to unlabelled IL-1β and $^{15}N/^{2}H$ IL-1β in a complex with unlabelled gIC8 Fab (Figure 4.12). Partial alignment of the free Fab was achieved using 4 mg/ml Pf1 filamentous phage and 3 mg/ml Pf1 was used for the complex

samples. These phage concentrations were selected based on the work done on optimisation of RDC data collection for large complexes in Chapter 2.

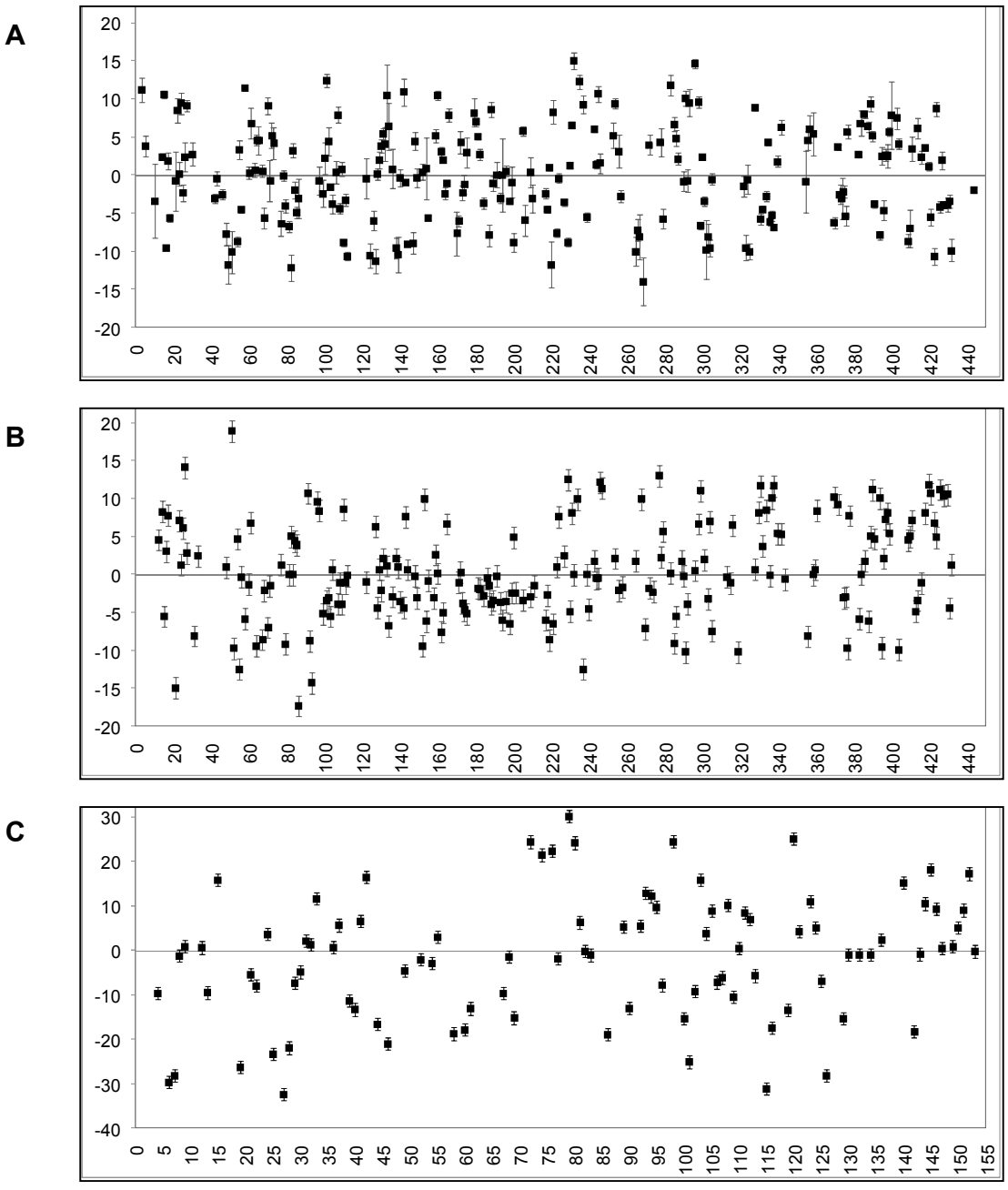


Figure 4.12

A summary of the backbone amide RDC values determined for the free and IL-1β bound Fab. Panel A contains RDCs values obtained from the free Fab collected in 4

mg/ml Pf1 phage alignment media. The standard deviations in peak position obtained from spectra recorded in triplicate are indicated by vertical error bars. Panel B shows RDC values determined for $^{15}N/^{2}H$ Fab bound to IL-1β and for the $^{15}N/^{2}H$ IL-1β-unlabeled Fab complex in C. RDCs for the Fab complexes were collected from samples containing 3 mg/ml Pf1 alignment media. The errors on RDCs determined from the protein complexes was estimated to be 1.4 Hz, based on a conservative estimate of the peak position in the processed spectra.

Backbone amide RDCs were derived from the difference in the $^{15}N/^{1}H$ scalar couplings between isotropic and partially aligned samples which were determined from $^{15}N/^{1}H$ HSQC and TROSY spectra (Kontaxis et al., 2000), acquired at 800 MHz. For the free Fab, 219 RDCs (109 from the constant domains and 110 from the variable domains), covering 49% of the protein, were determined with values ranging from -14 Hz to +15 Hz. The remaining residues that were visible in the spectra could not be assigned due to overlap in the $^{15}N/^{1}H$ HSQC and TROSY spectra. Each HSQC and TROSY spectrum from both aligned and unaligned free Fab samples were recorded in triplicate, giving and an average error in peak position of 1.25 +/- 0.88 Hz. The agreement between the measured RDCs and values calculated from the unrefined model using PALES are shown in Figure 4.13A-C. RDCs determined from the constant domain (Figure 4.13A) show a better agreement with the unrefined model than RDCs determined from the variable domains (Figure 4.13B), with Q factors of 0.37 and 0.47 respectively.

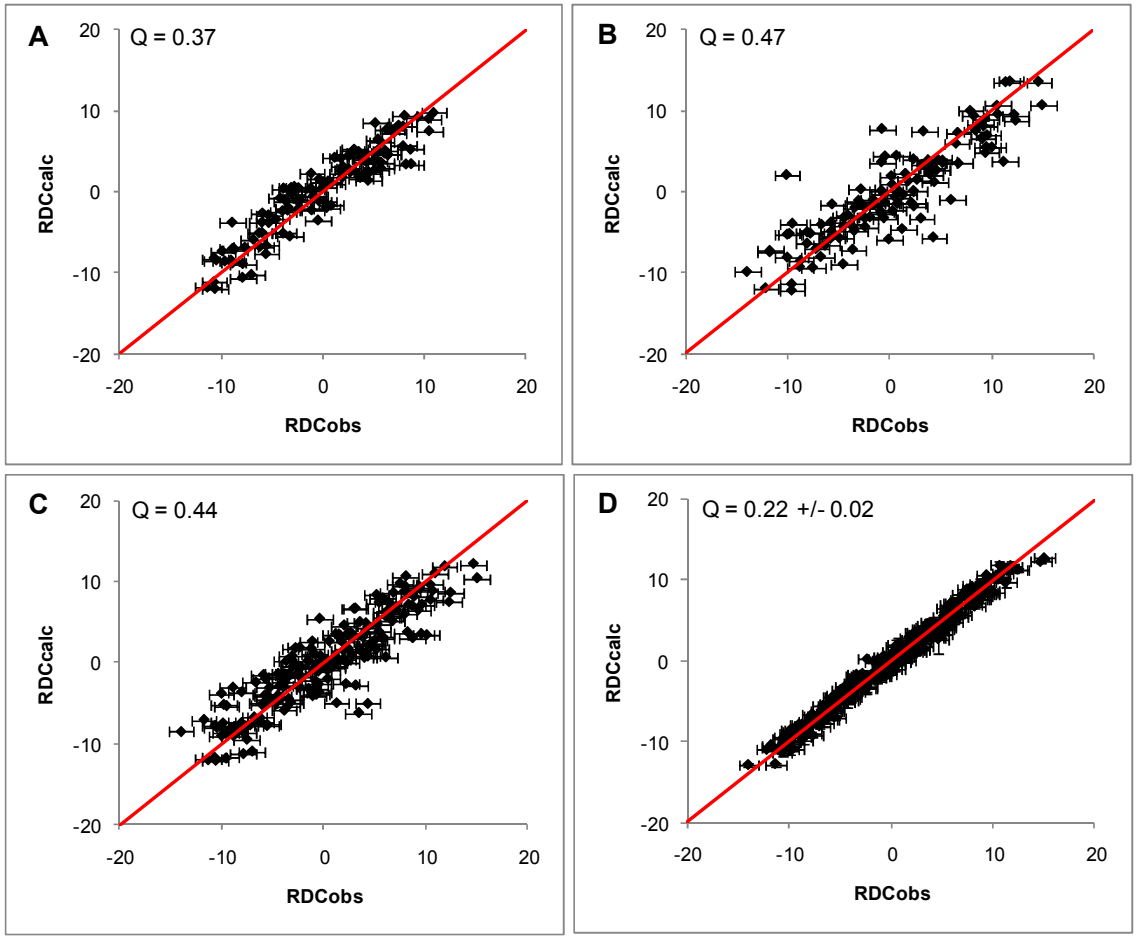


Figure 4.13

Comparison of the agreement between measured and calculated backbone amide RDCs for the Fab before and after refinement. Panels A, B and C show the agreement between RDCs measured for the free Fab (RDC_{obs}) and RDCs calculated from the constant domains (A), variable domains (B) and complete Fab structures (C) of the model which was used as a starting point for Fab refinement (RDC_{calc}). Panel D shows the agreement between the measured Fab RDCs and the RDCs calculated from the refined Fab structures. The mean RDC value for each residue is shown with the standard deviation of the values over the 115 Fab structures indicated. The red line on each graph shows perfect agreement and the Q factors are also indicated.

The gIC8 homology model was based on the PDB structure 1L7I (Vajdos et al., 2002) as it provided the best template. 1L7I and gIC8 Fab shared exactly the same sequence in their constant domains meaning that these domains could be modelled with high confidence. The gIC8 variable domains shared less sequence homology with 1L7I and so these were based on the structure of the corresponding scFv 2KH2 (Wilkinson et al., 2009) as it shares the variable domain sequences with the Fab and so the variable domains are expected to be very similar. The scFv structure, however, was obtained as part of a complex. The orientation of the variable domains may, therefore, be influenced by complex formation and there may also be structural differences arising from the fact that variable domains in Fab fragments are associated with constant domains. The method by which the Fab was modelled may explain why RDCs determined from the constant domain of the free Fab agree better with the homology model than those determined from the variable domains.

For the $^{15}N/^{2}H$ Fab bound to unlabelled IL-1β 201 RDCs (100 from the constant domains and 101 from the variable domains) covering 46% of the protein were obtained with values ranging from -17 Hz to +19 Hz. The agreement between RDCs determined for gIC8 Fab in complex with IL-1β and RDCs calculated for the unrefined model of the free Fab are shown in Figure 4.14A. Again, overlap in the $^{15}N/^{1}H$ HSQC and TROSY spectra prevented assignment of a full set of backbone amide RDCs. For $^{15}N/^{2}H$ IL-1β bound to unlabelled Fab, 96 RDCs were measured, covering 63% of the protein, with values ranging from -32 Hz to +30 Hz. The agreement between experimentally determined RDCs for IL-1β and RDCs calculated from a crystal structure of free IL-1β (2I1B) is shown in Figure 4.14C. Errors on RDC values obtained from both components

in the complex were estimated to be 1.4 Hz based on a conservative estimate of the peak position in the processed spectra.

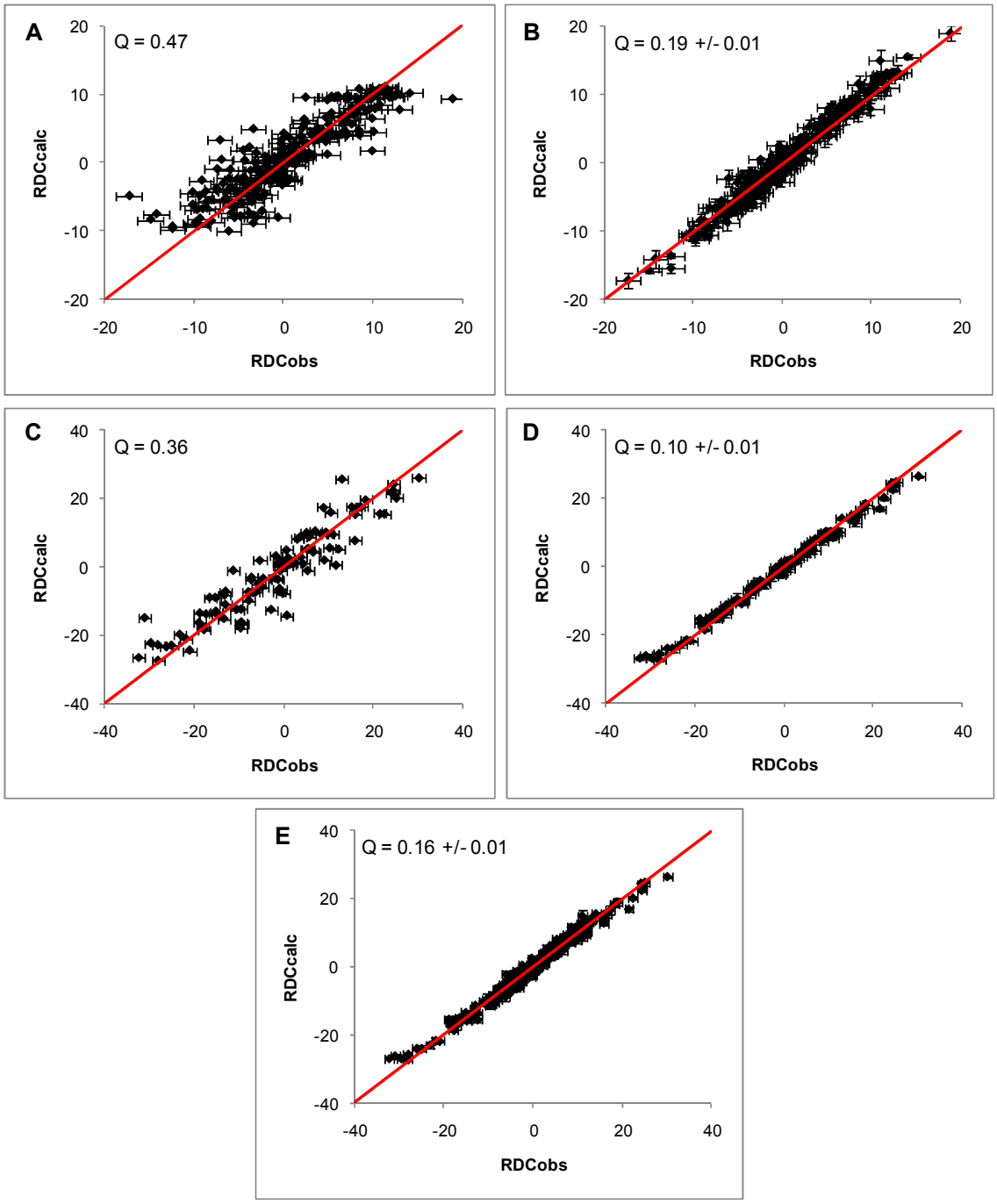


Figure 4.14

Comparison of the agreement between measured and back-calculated backbone amide RDCs for IL-1β and gIC8 Fab before and after docking to form a complex. The panels show graphs of the experimentally determined RDCs (RDC_{obs}) compared to values calculated with PALES (RDC_{calc}). Panel A corresponds to RDCs measured for IL-1β in complex with Fab compared to values calculated for the structure of free IL-1β (2I1B). Panel B compares the measured RDCs for bound IL-1β to those calculated from each of the converged docked and refined structures of the IL-1β/Fab complex. The mean calculated RDC value for each residue is shown together with error bars indicating the standard deviation over the 132 converged complex structures. Panel C compares RDCs measured for the Fab in complex with IL-1β to the values calculated from the unrefined model, which was used as a starting point for docking. In contrast, Panel D compares the measured RDCs for the bound Fab to those predicted by the refined IL-1β/Fab complex. The mean calculated RDC value for each residue is shown together with error bars indicating the standard deviation over the 132 converged complex structures. These plots clearly show a significantly improved agreement between the measured and calculated RDC values after restraint driven docking to produce the complex. Panel E shows agreement between measured and calculated RDCs for the complex as a whole. The red line on the graphs is an indication of perfect agreement and the Q factor for each plot is shown.

4.3.3.2 NOEs

A total of 559 HN-HN NOE-derived distance restraints were obtained from the NOESY-HSQC spectrum of free gIC8 Fab. Distances between residues were determined based on peak height and NOEs were divided into groups corresponding to

^{1}H-^{1}H distances of <5 Å, 5 to 6.5 Å and 6.5 – 8 Å. The average peak heights for sets of *i*+1 and *i*+2 peaks were taken to be the thresholds for the <5 Å and 6.5 Å respectively. This gave groups of 167 sequential NOEs (i, i + 1) corresponding to distances of <5 Å, 196 medium range NOEs (i, i + ≤ 4) corresponding to distances of 5 to 6.5 Å and 196 long range NOEs (i, i + ≥ 5) corresponding to distances of 6.5 to 8 Å.

4.3.3.3 Dihedral angles

Phi and psi angles were determined from CA, CB, CO and N chemical shifts using TALOS for both the free Fab and bound IL-1β (Cornilescu et al., 1999). For the free Fab 432 phi and psi angles were determined and 178 phi and psi angle restraints were determined for IL-1β in complex with the Fab.

4.3.4 Refinement of the homology model produced for free gIC8 Fab

Refinement of a homology model of free gIC8 Fab was performed in HADDOCK (Dominguez et al., 2003) with the docking element disabled. Refinement was driven by 219 experimentally determined backbone amide RDCs, which were used as direct restraints and also as intervector projection angle restraints (Meiler et al., 2000). Other experimental restraints used were 559 HN-HN NOE derived distance restraints and 432 phi and psi dihedral angles determined from experimental CA, CB, CO and N chemical shift values as described previously. Residues in the Fab domain interface and CDRs were defined as semi-flexible during the refinement (Figure 4.15A).

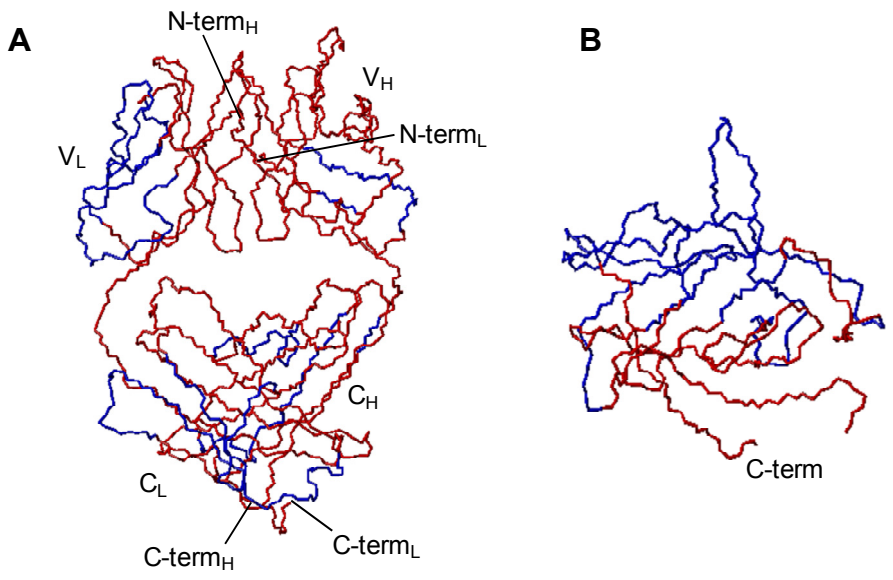


Figure 4.15

Localisation of semi-flexible and fixed residues in docking calculations. Residues at the protein-protein interfaces (active and passive residues +/- 2 sequential residues) for Fab (A) and IL-1β (B) were defined as semi-flexible. For the Fab, residues in the domain interfaces were also defined as semi-flexible. In the backbone views of Fab (A) and the IL-1β (B) the semi-flexible residues are shown in red.

In the final iteration of the HADDOCK refinement protocol, 115 satisfactorily converged complex structures were obtained from 200 starting conformations (Figure 4.16). A best fit superposition of the converged structures on the backbone atoms of residues in elements of regular secondary structure in gIC8 gives a backbone atom (N, Cα, CO) RMSD to the mean of 0.36 +/- 0.07 Å. Figure 4.13C shows the agreement between backbone amide RDCs experimentally determined for gIC8 Fab and RDCs calculated for the unrefined homology model of gIC8 Fab. The agreement between

RDCs determined for gIC8 and RDCs calculated for the refined Fab homology model are shown in Figure 4.13D, illustrating an essentially perfect agreement (within error) as a result of refinement.

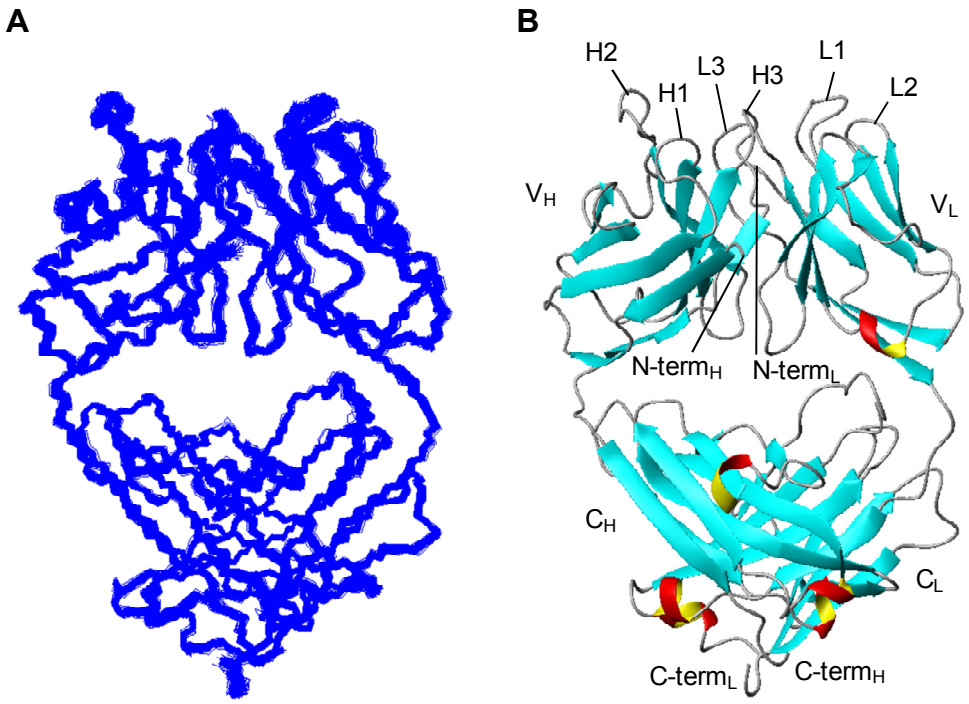


Figure 4.16

Refined model of free gIC8 Fab structure. Panel A shows a best fit superimposition of the 115 converged structures obtained after refinement of the gIC8 Fab homology model. A best fit superposition of the converged structures on the backbone atoms of residues in elements of regular secondary structure in gIC8 gives a backbone atom (N, Cα, CO) RMSD to the mean of 0.36 +/- 0.07 Å. Panel B shows a ribbon representation of the structure closest to the mean in the same orientation as in Panel A.

4.3.5 Determination of the interaction surface on gIC8 and IL-1β

For protein-protein docking the HADDOCK software requires the interaction surface on each protein to be defined from experimentally determined parameters. For IL-1β, these so called ambiguous interaction restraints (AIRs), were determined from combined chemical shift changes in HN, N and CO induced on complex formation from HNCO spectra. A defined binding footprint can be identified when chemical shift changes between free and bound IL-1β are shown on the surface of a space-filled view of IL-1β (Figure 4.17A). The thresholds used to define the AIRs for IL-1β were, therefore, selected to ensure that the entire footprint was defined as being in the interaction surface. Active AIRs were therefore defined as residues with a chemical shift change of greater than 0.15 ppm on binding to gIC8. Residues in the interaction surface also have to be surface exposed and so as an additional restraint active residues also had a solvent accessibility of greater than 10% (calculated in MolMol (Koradi et al., 1996)). This gives 26 residues being actively involved in the IL-1β interaction surface. Passive AIRs were defined as neighbours of the active residues with a solvent accessibility of greater than 10% which gave 15 passive residues for IL-1β, shown in Figure 4.17A.

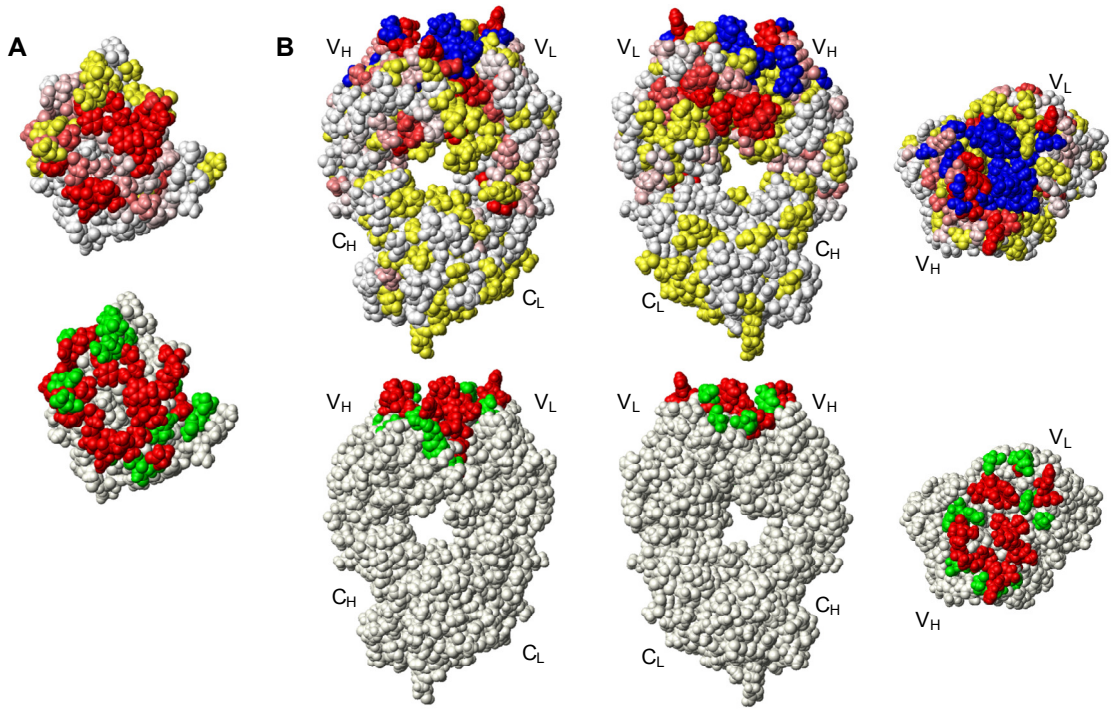


Figure 4.17

Mapping of interaction surfaces and definition of ambiguous interaction restraints (AIRs). Residues with signals significantly perturbed on complex formation are shown on space filled structures of IL-1β in Panel A and for the Fab in Panel B (top). Residues which show no significant chemical shift change on complex formation are coloured white, whilst substantially perturbed residues are shown in red. Other residues with available minimal shift data are coloured on a linear gradient between white and red, depending on the magnitude of the chemical shift change as described previously. Residues with no available chemical shift data are coloured yellow and CDR residues that are stabilised on antigen binding are shown in blue. For docking using HADDOCK, the residues experimentally identified to be at the interaction surface and solvent exposed are defined as active. Solvent exposed neighbours of the active residues are defined as passive. These are also used in docking to help define the interaction surface but carry less weight than active residues. Active residues are shown in red on the space

filled models of gIC8 and IL-1β (bottom) while passive residues are shown in green. For IL-1β, active residues were defined as solvent exposed residues which were perturbed my more than 0.15 ppm. For gIC8 Fab, active residues were defined as solvent exposed CDR residues which were perturbed on complex formation. This includes residues in the CDRs which show a minimal shift change of more than 0.05 ppm on binding to IL-1β and also residues in the CDRs that are stabilised on complex formation. Passive residues in gIC8 Fab are any solvent exposed neighbours of the active residues.

When the combined minimal shift changes in HN, N and CO induced on complex formation were shown on the surface of a space filled Fab homology model (Figure 4.17B), there was no discrete interaction surface, as chemical shift perturbations could be seen throughout the V_H-V_L interface. It was, therefore, difficult to establish whether framework residues in the vicinity of the CDRs were perturbed through a direct interaction with IL-1β or due to another conformational change induced by binding. In addition, there was no minimal shift perturbation data available for a significant proportion of the CDRs, notably CDR3 in the heavy and light chains. A reasonable compromise to address this issue was to define the interaction surface in the Fab as residues in the CDRs which demonstrate evidence of direct interaction with IL-1β. For residues where minimal shift data was available, perturbed residues were defined as those with a chemical shift change of greater than 0.05 ppm. CDR residues with no available chemical shift data were also assumed to directly interact with IL-1β if they were shown to be in intermediate exchange in the free Fab but stabilised into one position in the Fab complex. The residues at the interaction surface must also have a

solvent accessibility of greater than 10% and passive residues are defined as the solvent exposed neighbours of the active residues. This gives 23 active and 14 passive residues for gIC8 Fab (Figure 4.17B). Although all the residues that have been defined as active almost certainly interact directly with IL-1β in the complex, it is possible that some framework residues that also interact with IL-1β have been excluded.

4.3.6 Docking of IC8 Fab/IL-1β complex

The structure of the gIC8 Fab bound to IL-1β was obtained by NMR restraint driven docking using HADDOCK, which is software specifically designed for protein-protein docking (Dominguez et al., 2003), from a homology model of gIC8 and a crystal structure of IL-1β (PDB code 2I1B (Priestle et al., 1989)). This crystal structure was selected as it showed good agreement with the experimental RDCs collected for free IL-1β, as discussed in Chapter 2. The docking calculation was driven by information about residues in the interacting surfaces from chemical shift perturbation data for gIC8 Fab (23 active, 14 passive) and IL-1β (26 active, 15 passive). Backbone amide RDCs (201 for gIC8 and 96 for IL-β) were used to define the relative orientations of the proteins within the complex. RDCs were used both as direct restraints and also as intervector projection angle restraints (Meiler et al., 2000). For IL-1β 178 phi and psi angles, derived from CA, CB, CO and N chemical shifts were also used in the docking. The minimal shift perturbation data between free and bound Fab shows that there are many residues which are essentially structurally identical in the free and bound form, especially in the constant domains. This allowed for selected phi and psi dihedral angles derived from triple resonance spectra obtained for the free Fab to be used to aid in structural refinement of bound Fab during docking. Only phi and psi angles obtained for residues with chemical shift differences of less than 0.05 ppm between the free and

bound form were used in the docking giving a total of 178 phi and psi dihedral angle restraints. In a similar way, HN-HN NOEs determined between residues in the free Fab that showed no significant minimal shift on IL-1β binding (>0.05 ppm) were used in structural refinement of bound Fab during docking. A total of 195 HN-HN NOE derived distance restraints were used in the docking calculation, 69 sequential NOEs (i, i + 1) corresponding to distances of <5 Å, 54 medium range NOEs (i, i + ≤ 4) corresponding to distances of 5 to 6.5 Å and 72 long range NOEs (i, i + ≥ 5) corresponding to distances of 6.5 to 8 Å.

To allow for potential structural changes at the protein-protein interface, residues forming the interaction surface of both proteins were defined as semi-flexible during the docking calculation. The semi-flexible residues were defined as all residues in the Fab interaction surface (both active and passive), +/- 2 sequential residues. To allow for domain-domain movement on antigen binding, the residues at all the domain interfaces (V_H-V_L, V_H-C_H, V_L-C_L and C_H-C_L) were also defined as semi-flexible during the docking. Domain interface residues were identified in MolMol and defined as those that gain solvent accessibility when the adjoining domain was removed. These residues +/- 2 neighbouring residues were defined as semi-flexible as shown in Figure 4.15. In the final iteration of HADDOCK docking, 132 satisfactorily converged complex structures were obtained from 200 starting conformations as shown in Figure 4.18.

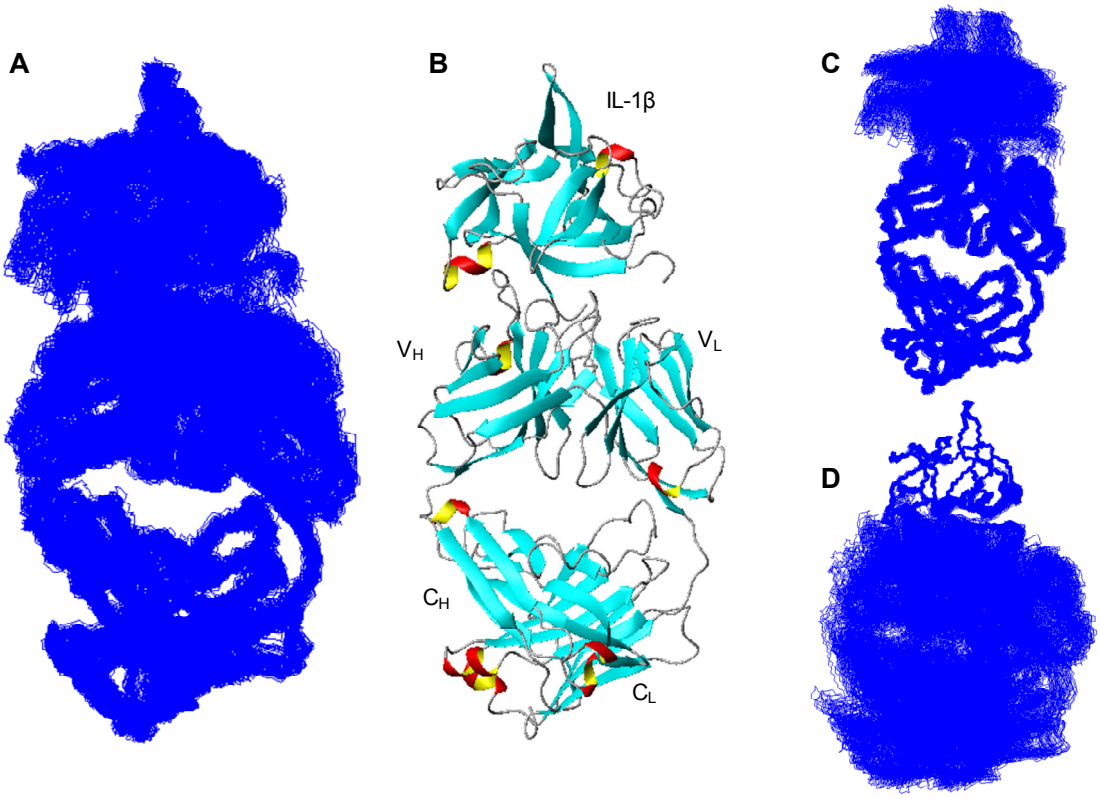


Figure 4.18

Structure of the gIC8 Fab/IL-1β complex. Panel A shows a best fit superposition of the 132 converged structures on the backbone atoms of residues in elements of regular secondary structure in the IL-1β/gIC8 complex, which gives a backbone atom (N, Cα, CO) RMSD to the mean of 2.39 +/- 0.70 Å. Panel B shows a ribbon representation of the closest to mean structure in the same orientation as Panel A. Panel C shows the converged structures when only the elements of regular secondary structure in the backbone atoms of the Fab are superimposed giving a backbone RMSD to the mean of 0.88 +/- 0.2 Å for the Fab. Panel D shows the converged structures when only elements of secondary structure in the backbone atoms of IL-1β are superimposed, giving a backbone RMSD of 0.34 +/- 0.07 Å for IL-1β.

A best fit superposition of the converged structures on the backbone atoms of residues in elements of regular secondary structure in the IL-1β/gIC8 complex gives a backbone atom (N, Cα, CO) RMSD to the mean of 2.39 +/- 0.70 Å. When only superimposed on elements of secondary structure in IL-1β the backbone atom RMSD to the mean of IL-1β was 0.34 +/- 0.07 Å and when superimposed on the elements of secondary structure in the Fab the backbone atom RMSD to the mean was 0.88 +/- 0.2 Å. Individually, IL-1β and gIC8 Fab were refined very well during the docking, as illustrated by the improved agreement between the experimental RDCs obtained for both components and RDCs calculated in PALES before and after docking (Figure 4.14B,D,E). The backbone RMSD of the converged complex structures was significantly higher than the backbone RMSD values of the individual complex components, which suggested that there was variation in the relative positioning of gIC8 and IL-1β within the set of structures.

The structure for a complex of the corresponding scFv bound to IL-1β was determined using a very similar docking protocol (Wilkinson et al., 2009). The resulting scFv/IL-1β complex structures produced one main cluster containing 77 of the 200 calculated structures and the remaining structures were grouped into eight sparsely populated clusters. Clustering of the converged gIC8 Fab/IL-1β structures into groups with a clustering cut off of 3.5 Å did not produce one highly populated cluster and several less populated clusters, as was seen for the scFv/IL-1β complex, instead a group of 5 equally populated groups were produced.

To investigate further the structural variation that was occurring within the set of converged gIC8 Fab/IL-1β complex structures, the relative positioning between the Fab and IL-1β in a representative structure from each of the five clusters was examined. In

Figure 4.19, space filled views of each of the five representative structures are shown in an orientation where the IL-1β binding site is viewed from above. Only the Fab variable domains are shown and these are in the same orientation in each of the structures, highlighting the changes in IL-1β position of approximately 10 Å between the structures. Figure 4.19 also shows the representative structures, in the same orientation, with a ribbon representation of IL-1β, highlighting that there is very little change in the orientation of IL-1β between the representative structures from each of the five clusters. This shows that the results of the docking provided a series of structures in which the relative orientations of the two proteins were consistent, however, there was variation in the interaction surface, resulting in a set of structures where IL-1β is 'sliding' around the Fab binding site. The structure closest to the mean of the complete set of converged structures is shown in Figure 4.19F. Interestingly, in this structure, the IL-1β is positioned fairly centrally on the Fab variable domains. There is clearly some variability in the contact surface of IL-1β on the Fab which presumably reflects uncertainty in the docking restraints. In contrast, the relative orientation of IL-1β with respect to the Fab remains constant and appears well defined by the RDCs.

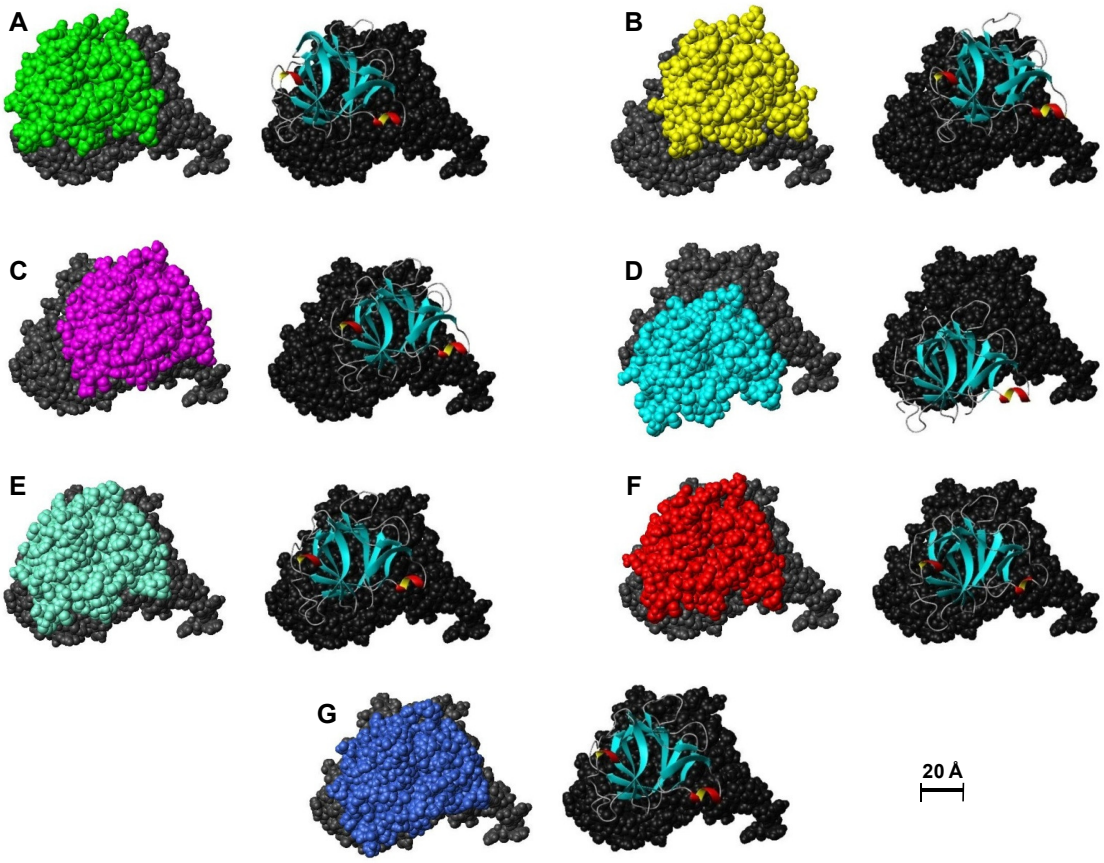


Figure 4.19

Variability in the IL-1β/gIC8 Fab contact surface in converged docked structures. After docking calculations to produce the IL-1β/Fab complex, five almost equally populated, converged clusters were obtained, with the closest to mean structure from each cluster shown in Panels A-E. Each panel shows the variable (scFv) portion of the Fab in the same orientation in grey with a coloured space filled view of IL-1β on the left, and a corresponding ribbon representation of IL-1β on the right. Panel F shows the position of IL-1β in the closest to overall mean structure and Panel G shows the position of IL-1β in the corresponding scFv complex (PDB code 2KH2) (Wilkinson et al., 2009).

RDCs determined from IL-1β/gIC8 complex were used both as restraints during the docking protocol and for the initial determination of the alignment tensor of the partially aligned IL-1β/gIC8 complex from the crystal structure of IL-1β (2I1B). The values of the backbone amide RDC data contain experimental error, however, difficulties in determining the effect of these errors in the final docked gIC8/IL-1β complex structure arise from the combination of experimental parameters used in the docking, each of which contain a certain level of experimental error. The very few available examples from the literature show that, assuming the RDC data set samples the protein sufficiently, increasing the errors on RDCs does not notably effect the outcome of structural refinement protocols.

A protocol using 217 simulated backbone amide RDCs calculated from a crystal structure and small angle scattering was used to refine the structure of the 31 kDa nuclear export protein TAP (TIP-associated protein) (Gabel et al., 2008). The effect of RDC error on the refined structures was assessed by evaluating the difference between structures that had been derived from RDC data sets after the introduction of experimental noise (1 Hz and 2 Hz). The results showed that the introduction of errors to the RDCs did not notably change the main structural parameters and similar ensembles of structures were obtained.

The alignment tensor of ubiquitin was determined from a set of simulated RDCs with 1.8 Hz error introduced to the values (Zweckstetter and Bax, 2002). The data showed that errors in the value of the alignment tensor increase substantially if less than 20 RDCs are used in the calculation, however inclusion of more than 20 RDCs in the calculation provided little improvement in the accuracy of the alignment tensor. These

results indicate that increasing the errors on RDCs does not introduce substantial errors in the orientation of the alignment tensor, as long as a sufficiently large RDC data set is used. Experiments evaluating the dependence of the HADDOCK docking procedure of a scFv and IL-1β on the completeness of backbone amide RDC coverage showed that if RDC data was available for 46% of the scFv and 48% of the IL-1β a reliable docked complex structure could be obtained (Wilkinson et al., 2009). According to this data, the number of RDC used in the Fab/IL-1β docking protocol (46% coverage of gIC8, 63% coverage of IL-1β) was sufficient to provide a reliable complex structure.

4.3.7 Comparison between Fab and scFv complexes with IL-1β

We have previously established that the gIC8 Fab binds to IL-1β in essentially identical manner to the corresponding scFv fragment by comparing a $^{15}N/^{1}H$ HSQC spectrum obtained for ^{15}N-labelled IL-1β bound to gIC8 Fab to a HSQC spectrum obtained for ^{15}N-labelled IL-1β bound to the corresponding scFv (Wilkinson et al., 2009). The positions of the peaks in the spectra were nearly identical, meaning that the binding of IL-1β by Fab or scFv is very similar. It is, therefore, expected that the scFv/IL-1β complex will have a very similar structure to IL-1β bound to the Fab variable domains. The interaction surface on the scFv was very well defined in the docking calculation, as chemical shift perturbation data was available for essentially all the residues in the CDRs in their bound form. Combined with backbone amide RDCs and long range, intramolecular HN-HN NOE derived distance restraints, the docking produced a single converged solution which was consistent with other published antibody-antigen structures. Figure 4.19G shows the scFv/IL-1β complex with the scFv in the same orientation as the variable domains of the previously discussed Fab complexes. This figure highlights that in the scFv/IL-1β complex, IL-1β occupies a very central position

in the scFv binding site, similar to the positioning of IL-1β in the closest to mean IL-1β/Fab complex structure (Figure 4.19F). A best fit superposition of the backbone atoms of residues in elements of regular secondary structure in the IL-1β/scFv complex and IL-1β bound to the Fab variable domains from the closest to mean structure, shown in Figure 4.20, gives a backbone atom (N, Cα, CO) RMSD to the mean of 1.65 Å.

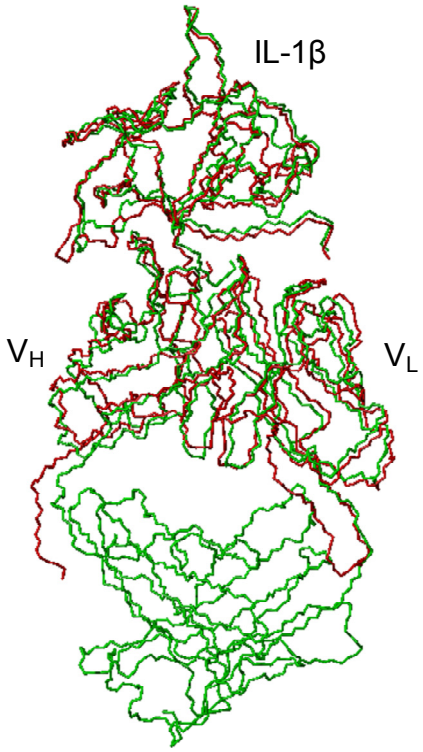


Figure 4.20

Comparison of the complexes formed between IL-1β and a scFv and Fab, with identical variable domains. Panel A shows an overlay of the structures closest to the mean for Fab/IL-1β complex in green with the scFv/IL-1β complex (PDB code 2KH2) in red. A best fit superposition of the backbone atoms of residues in elements of regular secondary structure in the IL-1β/scFv complex and IL-1β bound to the Fab variable domains gives a backbone atom (N, Cα, CO) RMSD to the mean of 1.65 Å.

4.3.8 Analysis of gIC8 Fab/IL-1β complex interface

Figure 4.18 showed that the Fab CDR loops make most contact with IL-1β, as would be expected for an antibody-antigen complex. The interaction surface is fairly large, with a buried surface area of 1937 +/- 237 $Å^2$. This is consistent with the scFv/IL-1β complex which has a buried surface area of 1930 +/- 130 $Å^2$. Analysis of the protein-protein interfaces was performed on the structure closest to the mean from the set of 132 converged gIC8 Fab/IL-1β complexes. Residues at the protein-protein interfaces were defined as those that lost more than 5% solvent accessibility on complex formation using MolMol (Koradi et al., 1996). On this basis, 17 interface residues were identified on IL-1β (A1, R4, L6, Q15, F46, N53, E105, N107, N108, K109, T147, M148, Q149, F150, V151, S152, S153). The residues identified in the IL-1β interface were essentially the same in both the scFv and Fab complexes.

Seventeen interface residues were also identified on the Fab, (I2, N28, H30, N31, Y32, W92, S93, L94, S224, D247, G268, G269, G270, S271, T272, Y273, K316). This shows that residues from CDRs L1, L3, H1, H2 and H3 were involved in the interaction along with framework residues from light and heavy chain (Table 4.1).

		Fab	scFv
V_L	FR L1	I2	D1
	CDR L1	N28, H30, N31, Y32	N28, H30, Y32
	CDR L3	W92, S93, L94	W92, S93, L94, F96
V_H	CDR H1	S224, D247	
	CDR H2	G268, G269, G270, S271, T272	G268, G270, S271, T272
	FR H2	Y273	Y273, F274, D276, K279
	CDR H3	K316	K316

Table 4.1

Summary of the interface residues identified in the Fab when bound to IL-1β. Residues at the Fab interface with IL-1β were identified as those which lost more than 5% solvent accessibility on complex formation (calculated in MolMol). The contact residues identified for the Fab are shown here, compared to those identified for the corresponding scFv in complex with IL-1β, with their position in the structure of the variable domains indicated.

Despite having the same number of residues in the interaction surface, five of the interface residues determined for the Fab complex differed from the scFv interface residues. Apart from I2, N31, S224, D247 and G269 which were identified in the Fab interface and D1, F96, F274, D276 and K279 which were identified as interface residues in the scFv complex the other interface residues were the same (Table 4.1). The scFv interface residues F274, D276 and K279 are all framework residues within the region of nonregular secondary structure which was shown to be stabilised into one conformation on Fab binding to IL-1β. It is, therefore highly probable that these

residues are also involved in the Fab interaction with IL-1β but not recognised because of the restricted interaction surface that was used in the docking calculation. The other discrepancies between the interaction residues in the scFv and Fab complexes are most likely due to the slight difference in IL-1β positioning between the two complexes, again due to slight differences in the interaction surfaces that were used as restraints in docking as shown in (Figure 4.21). A summary of the amino acid composition of both interfaces in the Fab/IL-1β and scFv/IL-1β complexes are compared in Table 4.2.

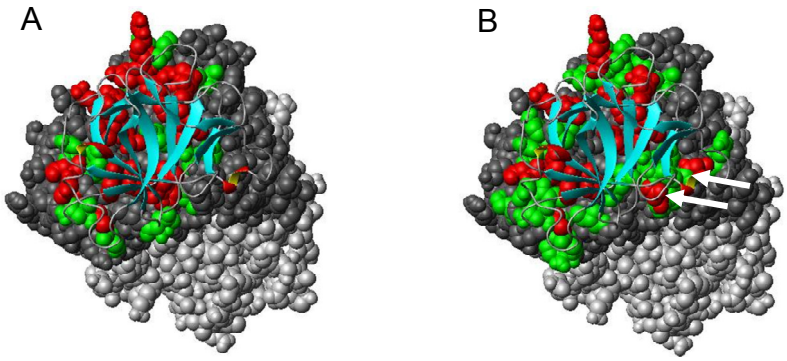


Figure 4.21

Comparison of the interaction surfaces defined in the docking calculation for gIC8 Fab and the corresponding scFv. The active (red) and passive (green) residues defined in the docking calculation for the Fab complex (left) and scFv complex (right) plotted on a space filled view of the Fab. The relative orientation of IL-1β is shown as a ribbon structure when the variable domains are superimposed with the scFv. Framework residues defined as active in the scFv docking are indicated by arrows.

PDB code	Antibody type	Target protein	BSA (Å^2)	Hydrogen bonds	Interface (%)			
					Aromatic	Non-polar	Polar	Charged
1NMC	scFv	Neuraminidase	1416	6	31.6	45.0	35.0	20.0
1DZB	scFv	Lysozyme	1856	6	32.1	46.4	14.3	39.3
2GHW	scFv	SARS spike	2528	5	13.5	27.0	43.2	29.7
1LK3	Fab	IL-10	1732	14	26.9	50.0	34.6	15.4
1FSK	Fab	Pollen antigen	1761	5	41.7	54.2	29.2	16.7
1ZTX	Fab	West Nile protein	1691	10	26.9	42.3	34.6	23.1
1JPS	Fab	Tissue factor	2103	18	25.8	45.2	22.6	26.1
3B2V	Fab	EGF receptor	1869	8	30.8	61.5	30.8	7.7
Mean			**1870**	**9**	**28.7**	**46.5**	**30.5**	**22.3**
S.D.			**329**	**5**	**8**	**10**	**9**	**10**
2KH2	scFv	IL-1β	1930	11	27.6	48.3	31.0	20.7
	Fab	**IL-1β**	**1937**	**10**	**23.5**	**47.1**	**35.3**	**17.6**

Table 4.2

A comparison of the interface features of a representative selection of antibody-target protein complexes. The interface properties of the antibody-target protein complexes ((Ay et al., 2000), (Faelber et al., 2001), (Josephson et al., 2002), (Hwang et al., 2006), (Li et al., 2008), (Malby et al., 1998), (Mirza et al., 2000), (Nybakken et al., 2005)) are compared with those of the gIC8 Fab and corresponding scFv complexes. The buried surface area (BSA), interface residues and hydrogen bonds were determined in MolMol. Interface residues are defined as those that loose more than 5% solvent accessibility on complex formation. Table was adapted from (Wilkinson et al., 2009).

Ten intermolecular hydrogen bonds between gIC8 Fab and Il-1β were calculated in the closest to mean Fab/IL-1β complex using MolMol and these are summarised in Table 4.3. The intermolecular hydrogen bonds were also calculated for every docked gIC8/IL-1β structure to give a more complete picture of hydrogen bonding across all the converged structures. Four of the hydrogen bonds seen in the closest to mean gIC8 Fab/IL-1β complex occur in 10% or more of the converged structures and these are highlighted in Table 4.3.

A comparison of the interface properties of a representative selection of antibody fragment (Fab and scFv)/target protein complexes is shown in Table 4.2 with the interface properties of the IL-1β/Fab complex added to the bottom of the table. The table highlights that the interface properties of gIC8 Fab/IL-1β are consistent with the interface properties of other antibody-protein antigen complexes.

Fab	IL-1β
H30	A1
S93	**S153**
L94	**S152**
D247	**N108**
G270	K109
S271	K109
S271	N108
T272	**Q15**
Y273	M148
K316	E105

Table 4.3

Summary of the intermolecular hydrogen bonds formed between gIC8 Fab and IL-1β. Intermolecular hydrogen bonds determined in MolMol for the structures closest to the mean from the converged gIC8/IL-1β structures are shown in the table. Intermolecular hydrogen bonds were also determined for each of the other converged structures. The hydrogen bonds which occur in the closest to mean structure and also in more than 10% of the other converged structures are shown in red.

4.3.9 Comparison between free and bound Fab

Comparisons were made between the structures of the free and bound Fab to characterise any changes that occur on IL-1β binding. It was not possible to draw any conclusions on structural changes within the CDR loops on antigen binding because too little experimental NMR data was available for these regions. However, the RDC

restraints used in docking and refinement, which provide information on the global orientation of the protein, were obtained for backbone amides distributed throughout the structure. Therefore, changes in domain orientations induced by antigen binding should be well characterised.

NMR is very sensitive to changes in the chemical environment of nuclei which makes the chemical shift perturbation data a highly reliable tool for the identification of residues that occupy a different environment in the free and bound Fab. The chemical shift perturbation data shown in Figure 4.11 provides very clear evidence of structural changes occurring in the Fab on complex formation. Residues in the variable domains that are perturbed on IL-1β binding fall predominantly in the CDR regions, at the V_L-V_H interface and at the V_L-C_L interface. Notably, the only residues that show a substantial chemical shift perturbation in the constant domains are those at the interface between the V_L and C_L domains which strongly indicates that no structural changes occur in the constant domains on IL-1β binding to gIC8 Fab.

Backbone atom (N, Cα, CO) RMSD to the mean values for each domain in the 115 converged, refined free Fab structures, when superimposed on elements of regular secondary structure, are shown in Table 4.4. The same information is also shown for the 132 converged Fab structures from the docked IL-1β/gIC8 Fab complex. The backbone RMSD of the closest to mean structures of each individual domain (V_H, V_L, C_H, C_L) in the free and bound Fab, when superimposed on elements of secondary structure, were determined to establish if any structural changes occurred within the domains. The RMSD values ranged from 0.38 to 0.50 Å indicating that no substantial structural changes occurred within the domains on antigen binding and that any structural changes

induced by antigen binding involved reorientation of a whole domain rather than structural rearrangements within the domains. The pattern of the chemical shift perturbations in the Fab indicates that the constant domains remain structurally unchanged on antigen binding. Since no substantial structural rearrangements are seen within the regular secondary structure of the domains, it is possible that the complete V_L domain rotates with the V_L-V_H interface as a pivot point, explaining the chemical shift changes seen in these regions.

Figure 4.22A shows the closest to mean free structure and closest to mean bound structure overlaid on the elements of regular secondary structure in their constant domains. In this orientation the constant domains overlay with a backbone atom (N, Cα, CO) RMSD of 0.55 Å. Figure 4.22A shows clearly that when the closest to mean free and bound structures are overlaid on the constant domains, the V_H domain occupies a very similar orientation in the free and bound structures, however a difference in orientation of the V_L domain can be seen. When the closest to mean structures of the free and bound Fab were superimposed on elements of secondary structure in the constant and V_H domains (Figure 4.22B), the backbone RMSD of these regions is 0.56 Å which implies that there is little change in orientation of V_H between the free and bound state. However, when the closest to mean free and bound structures are superimposed in this way, backbone RMSD of the V_L domains is 1.16 Å implying a change in orientation of the V_L with respect to the other three domains.

	Backbone atom (N, Cα, CO) RMSD to the mean when superimposed on elements of regular secondary structure (Å)	
	Bound Fab	**Free Fab**
V_L	0.29 +/- 0.05	0.24 +/- 0.03
V_H	0.30 +/- 0.05	0.17 +/- 0.02
C_L	0.30 +/- 0.05	0.21 +/- 0.02
C_H	0.36 +/- 0.07	0.20 +/- 0.03
Fab	0.88 +/- 0.20	0.36 +/- 0.07

Table 4.4

Summary of RMSD values for the converged refined free Fab structures and the converged structures of gIC8 bound to IL-1β. The backbone atom (N, Cα, CO) RMSDs to the mean are shown for the converged structures when superimposed on backbone atoms of residues in elements of secondary structure. RMSD values for the complete Fab, both in the free and bound structures, were calculated as were the RMSD values for the individual Fab domains. This data was used to confirm that no significant antigen induced structural changes occurred within the structures of the individual domains.

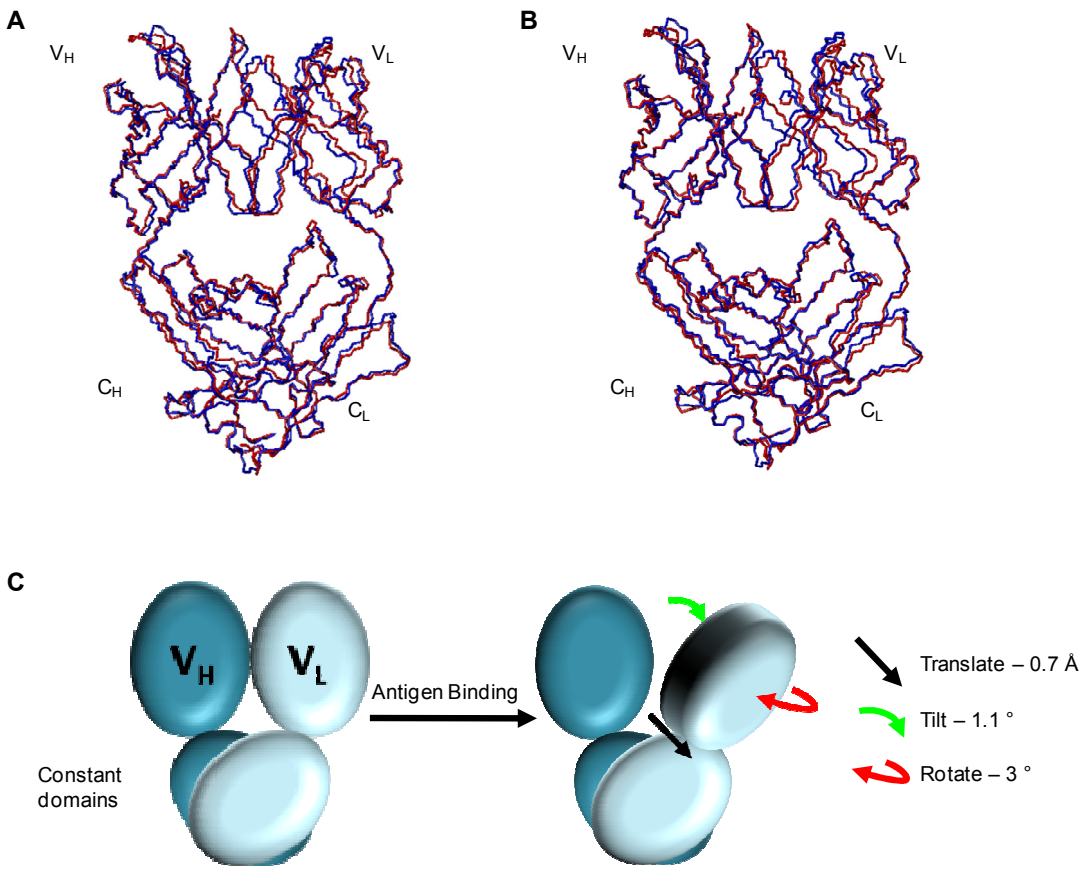


Figure 4.22

Change of domain orientation in Fab on IL-1β binding to gIC8 Fab. Panel A shows a best fit of the free (blue) and IL-1β bound (red) Fab, superimposed on the backbone atoms of elements in regular secondary structure in the constant domains only. Panel B shows a best fit of the free and bound Fab when superimposed on the backbone atoms of residues in elements of regular secondary structure in the constant and V_H domains. Comparisons of Panels A and B show that the V_H domain does not change in orientation relative to the constant domains on IL-1β binding, while the V_L domain undergoes a small change in orientation. Panel C shows a schematic representation of the change of orientation in V_L on antigen binding. The domain movement is described as a forwards and downwards slide or translation (black arrow), a tilt (green arrow) and a rotation (red arrow) around the tilt axis. The values shown are based on the difference between the

closest to mean structures and can be used to provide an idea of the scale of the structural changes induced by antigen binding.

To further explore the differences between the free and bound structures the coordinates of the backbone Cα atoms were obtained from 30 low energy structures from each of the sets of free and bound Fab structures when they were superimposed on elements of secondary structure in the free and bound Fab. This allowed for a mean backbone position to be obtained from each set of structures and for the spread within the sets of structures to be analysed. The mean difference in Cα position for the free and bound structures was calculated using the equation:

$$\Delta C\alpha = \sqrt{\left(x_f - x_b\right)^2 + \left(y_f - y_b\right)^2 + (z_f - z_b)^2}$$

Equation 4.2

where $x/y/z_f$ is the mean atomic position of a Cα coordinate in the set of free structures and $x/y/z_b$ is the mean atomic position of a Cα coordinate in the set of bound structures. The mean Cα positions were shown on the surface of a space filled model of the Fab (Figure 4.23B).

Structural differences of greater than 0.5 Å can be seen in 84% of the backbone Cα positions between the free and bound structures in the V_L domain with large differences at the V_L-V_H interface (residues H90-G101 and Q38-P40) and residues in CDRs L1, L2

and L3 also showing greater structural differences. Large changes can also be seen at the V_L-C_L interface which is consistent with the hypothesis that the whole V_L domain undergoes a structural change with the pivot point being the V_L-C_L interface.

Large differences between the structure of the free and bound Fab are also seen in the V_H domain at the V_L-V_H interface (G324 and Q325) and also in CDRs H1 and H2. Large differences are also seen in the stretch of residues between G256 and D287. This corresponds to CDR3 and also includes residues which are stabilised on IL-1β binding and are shown to interact with IL-1β in the scFv/IL-1β structure which suggests that differences between the free and bound structures in this region is due to direct interactions with the antigen. Large differences between the free and bound structures are also seen for residues G222 to L225. These residues lie in a region where no NMR data was available meaning that no restraints were driving the structural refinement in HADDOCK and could result in high uncertainties in atomic structure of the backbone of the free, bound or both structures in this region.

Large differences between the free and bound structures are also seen in the C_L. Differences seen in residues Y173 to K169 correlate with chemical shift perturbations that are induced by IL-1β binding and so are consistent with the existing data. Figure 4.23 highlights that the differences seen in residues V146 to N137 and E195 to V205 correspond to regions where only sparse NMR data is available and so these differences may be a result of poor model refinement in these regions rather than genuine structural differences. This is supported by the limited chemical shift perturbation data that is available (Figure 4.23) which does not show large chemical shift differences in these regions. Isolated residues in the C_H domain display structural differences between the

free and bound Fab structures, however in all cases these correspond to residues for which there is no NMR data and are therefore likely to be a result of poor model refinement.

An unpaired t-test was performed comparing the coordinates of backbone Cα atoms for the set of free and bound Fab structures to determine whether the changes seen in the Fab structure on binding to IL-1β were statistically significant. Since changes in chemical shift provide a much more sensitive indication of structural change than comparison of the position of backbone atoms between two structures, a residue was considered to be significantly different between the free and bound Fab if the chemical shift perturbation data showed a difference and the backbone Cα coordinates were statistically significantly different. Residues which were shown to be significantly different in the V_L domain were CDR residues (G27, A51, K52, T97), residues at the V_L-V_H interface (F98–G101, K103), residues at the V_L-C_L interface (I106-R108) in addition to residues Q3, T5, R24, I48, Y49, G68 and T69. Residues shown to the significantly different between the free and bound Fab in the V_H were the CDR residues, (S244, R245, S267, G269, S271, T272), framework residues that are thought to interact directly with IL-1β (Y273, V278, K279), residues at the V_L-V_H interface (A254, G324, T327) residues that are adjacent to this region (F282, T283, S285, R286, D287) in addition to residues 220, 221, 223, 328 and 330. Residues shown to be significantly different between the free and bound structures in the C_L were Q166 and S171 both of which are at the V_L-C_L interface and residues E123 and N137 which only show small chemical shifts. Residues in the C_H which were shown to be significantly different (S335, T336, H384 and H420) only just crossed the significance threshold.

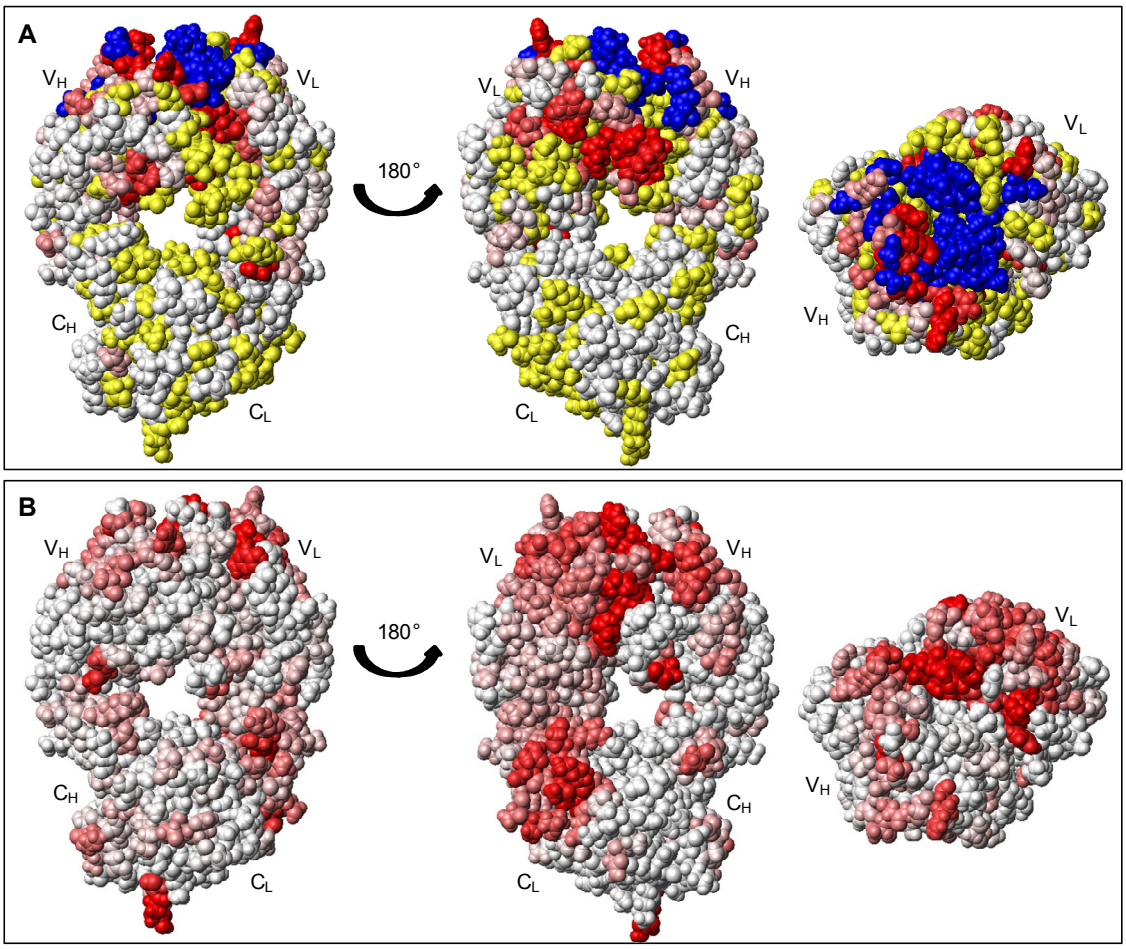


Figure 4.23

Localisation of gIC8 Fab residues showing structural changes on complex formation with IL-1β. Panel A highlights residues that show chemical shift differences between free and bound gIC8 Fab on a space filled view of the Fab in three orientations. Residues that showed a minimal shift change of less than 0.03 ppm are shown in white, minimal shifts of greater than 0.12 ppm are in red and between 0.03 and 0.12 ppm are coloured according to the magnitude of the shift on a linear gradient between white and red. Residues with no available chemical shift perturbation data are highlighted in yellow and residues highlighted in blue show CDR residues that are in intermediate exchange in the free Fab but are stabilised into one conformation on IL-1β binding.

Panel B highlights residues that show differences in the mean backbone Cα position between free and bound structures on space filled views of the Fab in the same orientation as Panel A. Residues whose backbone Cα position showed a mean difference of less than 0.5 Å between the free and bound structure are shown in white, residues with a mean difference in Cα position of greater than 1.5 Å are shown in red and between 0.5 and 1.5 Å are coloured according to the difference on a linear gradient between white and red.

This data strongly suggests that significant differences exist between the structures of the free and IL-1β bound gIC8 Fab, based on the chemical shift perturbation data and statistical comparisons between the refined homology model of the free Fab and the structure of the docked Fab. Based on this data we propose that the constant and V_H domains remain in a fixed orientation on binding to IL-1β whilst the whole V_L domain undergoes a small change in orientation. A schematic representation of the movement of the V_L domain is shown in Figure 4.22C. To provide an idea of the direction and magnitude of the change in the V_L orientation the closest to mean structures of the free and bound Fab were compared. It was found that between these two structures the V_L domain slides 0.7 Å in a forward and downward direction, tilts by 1.1° relative to the fixed domains, and rotates in a clockwise direction by 3° around the tilt axis.

Although there are over 100 unique structures of protein-antibody complexes (both scFv and Fab) currently accessible in the PDB, there are limited examples where both the structure of the antibody in the free and bound state is known. For the five examples of antibody fragments where structures have been obtained in the free and protein bound

form, three are anti-lysozyme antibodies (D44.1 Fab (Braden et al., 1994), HyHEL-63 Fab (Li et al., 2000) and D1.3 scFv (Bhat et al., 1990)), a tissue factor binding Fab (Huang et al., 1998) and a cytochrome c binding Fab (Mylvaganam et al., 1998). For each of these examples detailed comparisons have been made to identify structural changes that occur on antigen binding. Small antigen induced structural changes are reported in each, predominately relating to rearrangements of sidechains, rearrangements in the CDRs and changes in the relative disposition of the V_L and V_H domains.

Figure 4.24 shows examples of the free (blue) and bound (red) Fabs complexes formed with cytochrome c (A), tissue factor (B) and lysozyme (C) superimposed on elements of secondary structure in the constant domains, to highlight any changes in domain orientation comparable to those seen for gIC8 Fab. The effect of antigen binding to the Fabs varies quite significantly between the three examples. Cytochrome c binding seems to shift both variable domains in the same direction relative to the constant domains which is a different movement to that seen in the gIC8 Fab. The tissue factor binding Fab shows no real differences in variable domain orientation between the free and bound Fabs. The resolution of these structures is quite low however (2.5 Å for the free and 3.0 Å for the bound), so it is possible that subtle changes in domain orientation may have been lost in structural noise. The free and bound structures for the lysozyme binding Fab show that on complex formation the V_L domain remains in a fixed orientation with respect to the constant domains and there is a small change in the relative orientation of V_H. This is comparable to the structural changes that are seen in gIC8 Fab except that the relative change in orientation is seen in the V_H as opposed to the V_L. It is possible that this difference arises due to the nature of the antigen contact

with the CDRs. Of the 17 residues involved in the gIC8 Fab interaction surface, 8 are contributed by the V_L domain and 9 from the V_H domain. There are 12 residues identified as being in contact with lysozyme in the anti-lysozyme Fab, with only 3 of these being in CDRs in the V_L. Notably, there are no structural changes seen in the constant domains in any of the structures for which free and bound forms are available, suggesting that there is no structural signal propagated through the to the constant domains on antigen binding, confirming that antigen binding does not initiate B cell via a structural signal being passed though the constant domains.

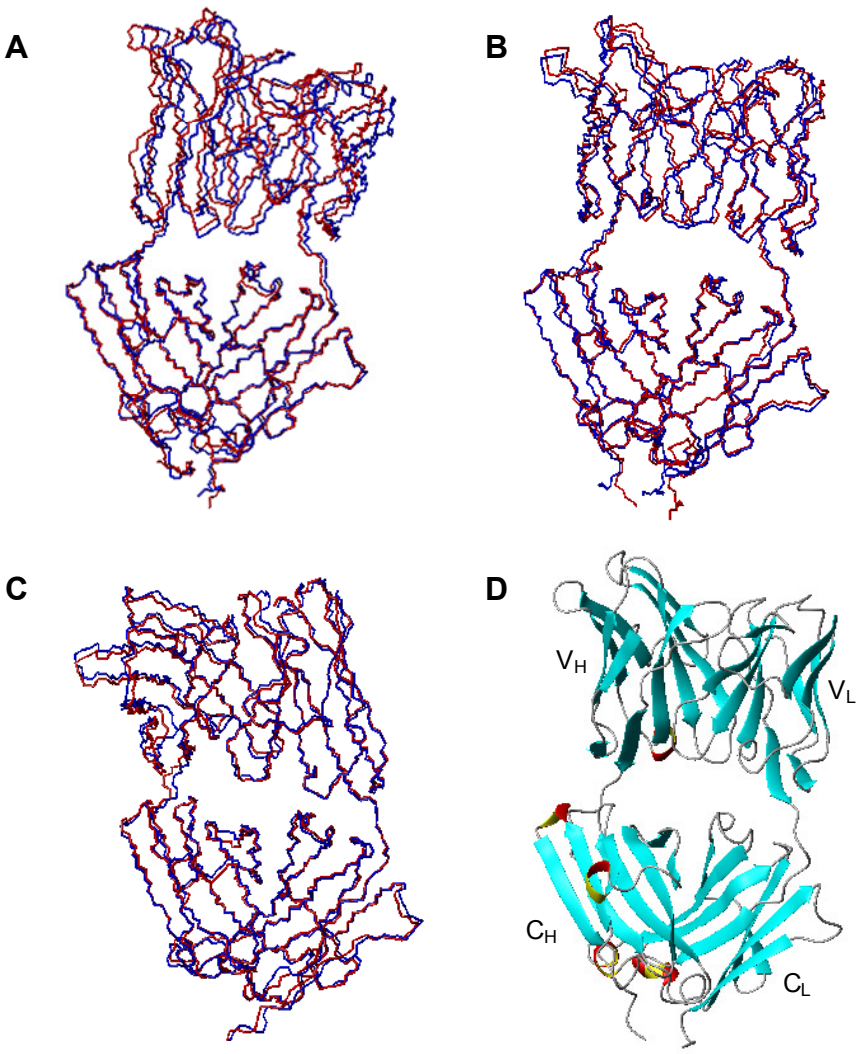


Figure 4.24

Structural comparisons between free and protein bound Fab fragments. The free (blue) and protein antigen bound (red) structures of Fabs against; A - cytochrome c (Mylvaganam et al., 1998), B - tissue factor (Huang et al., 1998) and C - lysozyme (Braden et al., 1994), are shown in the same orientation as the ribbon representation shown in panel D. The free and bound Fab structures are superimposed on elements of regular secondary structure in the constant domains only.

^{15}N, ^{13}C and ^{1}H signals in the NMR spectra of proteins are very sensitive to changes in chemical environment, for example changes induced by formation of protein-protein complexes (Cavanagh et al., 1996). For this reason, changes in signal position seen on binding as shown in Figure 4.11 are much more sensitive to small structural changes induced by antigen binding that comparing the backbones of crystal structures, as there is always some uncertainty in the atomic coordinates of structures. Although only a small set of structures are available for the free and protein antigen bound Fabs, the evidence suggests a mechanism where a structural change is seen in the variable domains in regions other than the CDRs and that no changes are propagated to the constant domains.

A fundamental unanswered question in immunology is the mechanism by which antigen binding to the B cell receptor can activate cytoplasmic B cell signalling. The current opinion is that antigen binding induces a structural change in the BCR which triggers signalling, however, the work presented here shows that structural changes at the interaction site are not propagated to the constant domains. A promising explanation is that the observed relative change in V_L orientation may form a new interaction site which could increase the affinity for the Fab with an Fc fragment, thus providing a direct method of transmitting the signal to membrane proximal Fc domains.

To explore our theory further, electrostatic potential surfaces were calculated for the free and bound Fab structures (Figure 4.25A-F). Changes in charge distribution occurred at the antigen binding site (Figure 4.25A,B), caused by rearrangement of surface loops on Fab contact with IL-1β. The other significant difference in surface potential between the free and bound form occurs in the V_L domain (Figure 4.25C-F)

where the relative change of orientation in this domain brings a negatively charged patch to a more central position in the face of the domain. This negative patch is brought closer to an existing negative patch across the V_L and C_L producing an extended potential interaction site.

An electrostatic surface was also calculated for the surface of a human IgG1 Fc fragment from the PDB (PDB code 1FC1 (Deisenhofer, 1981)) to ascertain if an antigen induced acidic binding patch on the Fab could indeed interact with an Fc domain (Figure 4.25G,H). The surface features on the Fc clearly show a basic portion on C_H2 and an extended basic patch from all four domains which lines the centre of the horseshoe shaped fragment. This shows that the Fab could form electrostatic interactions with the Fc domain.

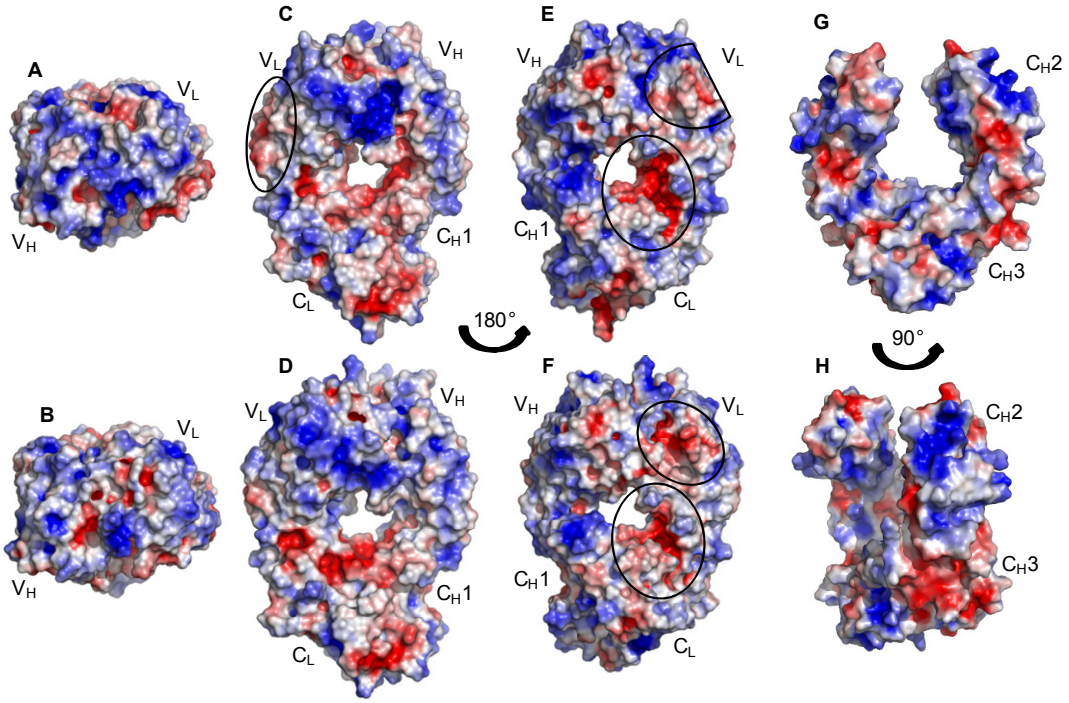


Figure 4.25

Electrostatic surface potential of free and bound Fab. Panel A shows a surface view of the IL-1β binding site in the free Fab, coloured according to electrostatic potential with areas of significant negative charge in red, positive charge in blue and neutral in white. A surface view of the bound Fab in the same orientation as A is shown in Panel B. Panel C shows a surface view of the free Fab showing V_L, V_H, C_L, and C_H1 with the same colours as used previously and the bound Fab shown in the same orientation in Panel D. Panel E shows the structure of the free Fab rotated 180° about the Y axis from the view in C with the bound Fab shown in the same orientation in Panel F. Regions of the structure that show significant changes in electrostatic potential on IL-1β binding are highlighted with black rings. Panel G shows a surface view of a human IgG1 Fc fragment (PDB code 1FC1) and a 90° rotation of the same structure shown in H, coloured according to electrostatic potential with the colours as above.

Although seven crystal structures of complete intact antibody molecules have been reported in the PDB, only three of these have full electron density and complete hinge regions (Saphire et al., 2002). These are murine IgG1 and IgG2 and a human IgG2. Structural comparisons show that the individual subunits are very similar but the structures differ significantly in the relative orientations of the subunits, indeed the Fab arm is rotated by 158°, the Fab-Fc angles ranges from 66° to 123° and the Fab-Fab angle between 115° and 172° (see Figure 4.26) (Sandin et al., 2004). The differences in relative subunit orientation is most likely due to the flexibility of antibodies with each of these structures providing a snapshot of a possible antibody conformation in solution. However, since the structures are from three different subclasses of IgG and are crystallised in three different space groups, it is difficult to draw conclusions about antibody flexibility from this data (Saphire et al., 2002).

Fluorescence resonance energy transfer (FRET) (Oi et al., 1984) has proved to be powerful tool for providing an insight into antibody segmental flexibility. Energy transfer measured between the C-terminal and N-terminal of an IgG1 molecule showed an average end to end distance of 75 Å, with a range of roughly 57 to 143 Å (Zheng et al., 1992). This value is significantly less than the 120 to 170 Å expected for a planar T or Y shaped IgG1 (Dangl et al., 1988) and suggests that the Fc portion can bend in both directions relative to the Fabs.

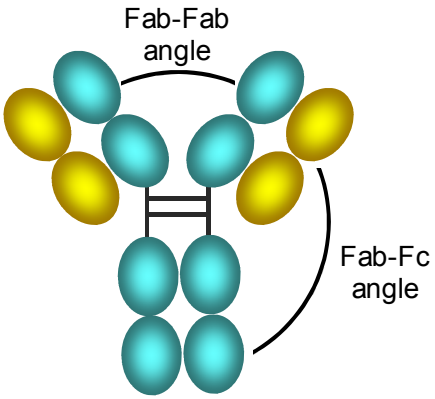


Figure 4.26

Schematic representation of an IgG antibody with the Fab-Fab and Fab-Fc angles indicated. Experimental techniques including X-ray crystallography, FRET and cryo-EM have highlighted the flexibility of antibody molecules, indicating that they do not necessarily exist in the classic 'Y' shape and in fact there is considerable variation the in the magnitude of the Fab-Fab and Fab-Fc angles.

Unlike crystallographic techniques, which show an average structure of a population of proteins, cryo-electron tomography (cryo-ET) provides visualisation of individual molecules in solution state conformations. Cryo-ET data collected from IgG2a showed good agreement with the individual subunits of a published crystal structure. It was shown that the Fc domains of the structures overlaid very well but the Fabs exhibited different relative orientations arising from hinge bending and rotational displacement of the Fab arms and, interestingly, the arrangements of the two Fab arms in relation to the Fc are not symmetrical (Sandin et al., 2004). The Fab-Fab and two Fab-Fc angles were measured for a set of 42 reliable IgG2a cryo-ET structures showing that the Fab arms are separated by approximately 110° with a distribution of about 30° and the Fab-Fc

angle ranges between about 165° and 52° (Bongini et al., 2004). This data agrees well with Fab-Fab and Fab-Fc angles calculated from electron microscopy which again visualises individual molecules but provides no 3D information. Angles measured for at least 200 IgG2 molecules gave a mean Fab-Fab angle of 127° with a standard deviation of 32° and an average Fab-Fc angle of 99° with a standard deviation of 32° (Roux et al., 1997). This data all supports a very flexible antibody structure where Fab fragments have a substantial degree of rotational and hinge bending movement relative to an Fc portion, which is also wagging and possibly rotating. It must be noted that none of these experiments have been performed on monovalent antigen bound antibody and so there is no evidence of the affect this will have on relative flexibility.

We have proposed a mechanism by which an antigen induced structural change in the variable domain can propagate a signal to initiate a B cell signalling pathway (
Figure 4.27). On antigen binding, the V_L domain changes in orientation relative to the V_H and constant domains, revealing a site which increases the affinity of the Fab for the Fc fragment. It is possible, due to the proven flexibility of the antibody hinge region, that a Fab could interact with its own Fc (
Figure 4.27A) providing the structural signal to initiate B cell signalling or B cell clustering. Perhaps a preferable model is that antigen binding increases the affinity of a Fab to an Fc fragment on neighbouring BCR (
Figure 4.27B). This would bring the ITAMs in the cytoplasmic regions of the signalling subunits into close proximity and as a consequence initiate B cell signalling in the currently accepted way. This model is entirely consistent with published data and provides an explanation of how small levels of antigen can initiate a rapid immune response as the BCRs do not need to be saturated with antigen to cluster.

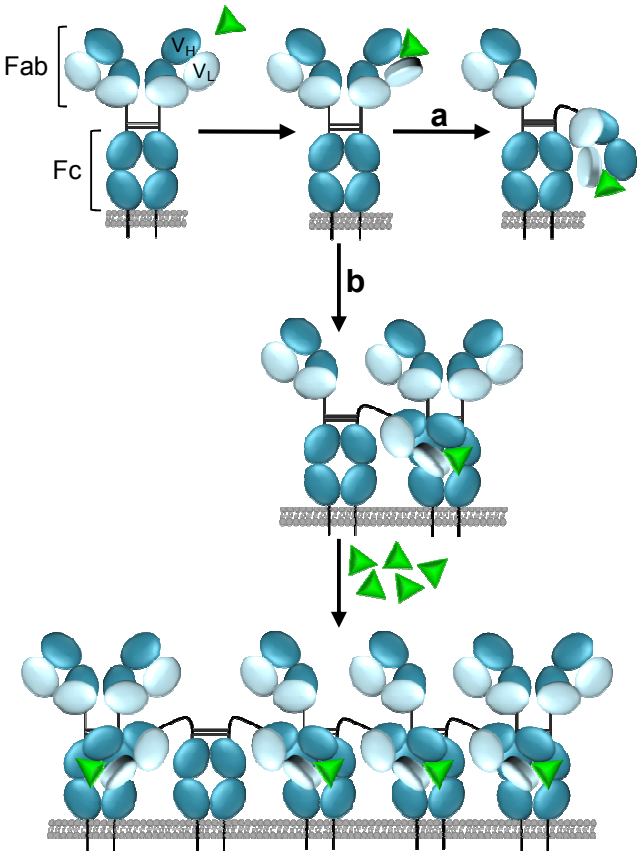


Figure 4.27

Proposed model of antigen induced B cell receptor signalling. Schematic of a membrane bound IgG BCR with the heavy chain shown as dark blue ovals and the light chain shown in light blue. Monovalent antigen is represented as green triangles. On antigen binding to the Fab, the V_L changes orientation with respect to the V_H and constant domains revealing an interaction site which has increased affinity for a region in the Fc fragment. Two possible pathways of initiation of BCR signalling are shown here: (**a**) Due to the flexibility of the hinge region the Fab fragment can interact with its own Fc thereby transferring the antigen induced structural change in the Fab to the

membrane proximal Fc portion which could induce structural changes in the associated signalling subunits and initiate cell signalling. (**b**) The Fab has increased affinity for the Fc domain of a neighbouring BCR. This induces BCR clustering, bringing together the ITAM motifs on the cytoplasmic tails of the signalling subunits allowing for phosphorylation of signalling proteins.

The data reported in this thesis focuses on a specific Fab fragment and the presented hypotheses are based on this single example. To determine whether the changes seen in gIC8 Fab on IL-1β binding occur in all Fabs on protein antigen binding, similar work will have to be carried out on other examples of Fab fragments. Evidence to support the hypothesis of the exposure of a potential Fc binding site on the Fab on antigen binding could be obtained by proving an interaction between an Fc domain and an antigen bound Fab fragment. Chemical shift perturbation mapping provides a sensitive method for validation of protein-protein interactions (Zuiderweg, 2002). TROSY experiments obtained for the $^{15}N/^{2}H$ labelled sample of the Fab in complex with IL-1β could be monitored whilst unlabelled Fc was titrated into the sample. Chemical shift perturbations would be seen for any residues in the Fab that were involved in an interaction with the Fc. This technique would not only confirm that Fab and Fc interact but also identify whether the interaction surface on the Fab is in the V_L domain as predicted. Chemical shift perturbation techniques are effective for probing weak interactions however, this approach may experience problems as the size of the combined Fab-Fc protein would be approximately 120 kDa, which could result in severe line broadening. Surface plasmon resonance is a versatile technique for characterising macromolecular interactions that permits the analysis of protein-protein

interactions over a wide range of molecular weights, affinities and binding rates (McDonnell, 2001), and could also be used to verify an interaction between the antigen bound Fab and an Fc domain. Other techniques commonly used to verify protein-protein interactions which also may be useful in verifying an interaction between IL-1β bound Fab and an Fc domain include size exclusion chromatography, fluorescence spectroscopy (Ward, 1985) and isothermal titration calorimetry (Pierce and Nall, 1999).

In this chapter it is proposed that on binding to antigen a conformational change occurs in the Fab revealing a negative patch which could potentially interact with positively charged regions in the Fc, thereby initiating BCR signalling. This hypothesis could be tested using site directed mutagenesis techniques where charged residues implicated in this interaction are mutated to Alanine or an oppositely charged residue. B cells lines stably expressing this mutant or wild type BCR could then be generated, and the antigen initiated downstream signalling monitored by Western blot analysis with a phospho-specific antibody directed against Erk1/2 (Carsetti et al., 2009).

4.4 Conclusions

The work reported in this chapter has shown that high quality triple resonance NMR data can be obtained for $^{15}N/^{13}C/^{2}H$-labelled gIC8 Fab fragment allowing for assignment of a high proportion of the backbone (HN, N, Cα, Cβ, CO) resonances and collection of long range HN-HN NOEs. Backbone amide RDCs determined from the Fab were combined with HN-HN NOE derived distance restraints and dihedral angle restraints determined from Cα, Cβ, CO and N chemical shifts to refine a homology model of the free Fab.

An HNCO spectrum was also obtained for $^{15}N/^{13}C/^{2}H$-labelled gIC8 Fab bound to unlabelled IL-1β and the interaction surface with IL-1β was indentified using combined minimal shift changes in HN, N, CO between the spectra of the free and bound Fab. The interaction surface on IL-1β had been previously determined within the group. Backbone amide RDCs were successfully determined for $^{15}N/^{1}H$ Fab bound to IL-1β and $^{15}N/^{1}H$ IL-1β bound to unlabelled Fab. An NMR driven docking approach driven by knowledge of the protein interaction surfaces, dihedral angle restraints determined from Cα, Cβ, CO and N chemical shift data, backbone amide RDCs and long range HN-HN NOE restraints was used to obtain a structure of gIC8 Fab/IL-1β complex from a homology model of gIC8 Fab and a crystal structure of IL-1β.

The NMR data has provided some interesting structural insights into the mechanisms involved in antigen binding to an antibody fragment. Firstly, it was shown that some residues in the CDR loops, notably CDR3 in both the heavy and light chain, are in intermediate exchange between two or more conformations but are stabilised on

complex formation. This implies a degree of induced fit antigen binding, contrary to the lock and key model that has been proposed for mature antibody-antigen complex formation (Jimenez et al., 2003). Secondly, the chemical shift perturbation data between the free and bound Fab shows no chemical shift changes occur in the constant domains on antigen binding. This is a strong indication that B cell receptors do not propagate a structural signal though their constant domains on antigen binding and that antigen induced signalling must occur via another route. Chemical shift differences were also seen at the V_L-V_H and V_L-C_L interfaces, however no significant changes were seen at the V_H-C_H interface, hinting that the constant domains and the V_H remain in a fixed position while the V_L domain changes orientation using the V_L-C_L interface as a pivot point. Comparisons made between the structures of free and bound Fab confirmed this hypothesis showing that the constant and V_H domain did indeed remain fixed on complex formation, while the V_L domain undergoes a small but significant relative change in orientation. This change in orientation reveals a negatively charged patch on the Fab surface which, due to the proven flexibility of antibodies, we postulate could interact with the Fc domain of a neighbouring BCR, initiating BCR clustering and triggering the signalling pathway.

Chapter 5 - Final Conclusions

Detailed structural information on the complexes formed between therapeutic antibodies and their target proteins would provide insights into the nature of the interaction between the proteins and provide valuable information to improve the rational design and humanisation of therapeutics. An improved understanding of the structural changes that occur on antigen binding would also help determine a mechanism for antigen induced B cell signalling. Intracellular B cell signalling is believed to be the result of an antigen induced structural change in the antibody, however the mechanism by which this occurs is currently unknown. Crystal structures of antibody fragment-target protein complexes are available, however, to date the only NMR-based structure of an antibody-target protein complex is that of a scFv/IL-1β complex, which was very recently solved within our laboratory (Wilkinson et al., 2009). NMR is an ideal tool for mapping precise interaction interfaces on both the antibody and target protein and can also provide information on protein dynamics, although, until recently, obtaining structural information for large proteins and protein complexes was difficult and time consuming. Recent advances in technology have made it possible to obtain detailed NMR structures for proteins and complexes of over 80 kDa (Kanelis et al., 2001; Tugarinov and Kay, 2003).

Obtaining the structures of complexes is challenging by X-ray crystallography, as the dynamics of complex formation result can lead to difficulties in obtaining crystals, and NMR techniques are limited by the molecular mass of the complexes that can be observed. The prediction and modelling of protein-protein interactions is therefore a very important tool in providing information about the function of biological complexes

(Russell et al., 2004; Smith and Sternberg, 2002). The results by de novo docking of two proteins known to interact are dramatically improved if experimentally derived data is included in docking (van Dijk et al., 2005a). Residual dipolar couplings provide information about the global orientation of a protein and therefore can play an important role in defining the relative orientations of partner proteins when used as restraints in docking. The assessment of alignment media for RDC measurement and optimisation of approaches to make it possible to determine backbone amide RDCs for large IL-1β/Fab complexes is described in Chapter 2. The methodology developed was used to obtain reliable backbone RDC data for the gIC8 Fab/IL-1β complex, which played a key role in obtaining a reliable structure of the complex by restraint driven docking.

Increasingly antibody Fab fragments are being developed as therapeutics in applications where Fc mediated effects are not required, as Fabs contain the complete, highly specific antigen binding site, are very stable and can be efficiently produced in *E.coli*. Isotopic labelling of proteins (^{15}N and/or ^{13}C) is essential for the collection of triple resonance NMR spectra that greatly simplify obtaining backbone resonance assignments for proteins. For proteins greater than about 30 kDa, it is also highly advantageous to replace non-exchangeable protons in the structure with deuterium, as this reduces the ^{1}HN and ^{13}C T_2 relaxation rate, resulting in improved spectral sensitivity and resolution (Venters et al., 1995). Successful optimisation of a protocol for the expression and purification of ^{15}N/^{13}C/^{2}H gIC8 Fab is described in Chapter 3. To the best of my knowledge, this is the first time that uniformly ^{15}N/^{13}C/^{2}H-labelled Fab has been produced and used for detailed NMR-based structural studies.

Examples from the high quality triple resonance NMR data obtained for the $^{15}N/^{13}C/^{2}H$-labelled gIC8 Fab fragment are shown in Chapter 4. This allowed for assignment of a high proportion of the backbone (HN, N, Cα, Cβ, CO) resonances. An HNCO spectrum obtained for $^{15}N/^{13}C/^{2}H$-labelled gIC8 Fab bound to unlabelled IL-1β highlighted that resonances corresponding to some residues in the CDRs and also some framework residues could be seen in the spectrum of the Fab complex, but were not visible in the spectrum of the free Fab. This strongly suggests that these residues were in intermediate exchange between two or more conformations in the free structure but stabilised on complex formation, implying an induced fit mechanism for at least some CDRs on complex formation. This data also suggests that some framework residues may be involved in the formation of a stable antigen binding site, which is an important consideration when deciding which residues to conserve in antibody humanisation. Combined minimal shift (HN, N, CO) perturbation data obtained from the spectra of free and bound gIC8 indicated that significant chemical shift changes could be seen in the CDR regions and at the V_L-V_H and V_L-C_L interfaces. Significantly, no chemical shift changes were seen at the V_H-C_H or in the constant domains. Based on this chemical shift perturbation data, it appears that on binding of IL-1β to gIC8 Fab the constant and V_H domains remain unchanged, whilst the V_L domain changes in orientation relative to the fixed domains, with the V_L-C_L interface acting as a pivot point (see Figure 4.11).

A reliable model of the gIC8/IL-1β complex was obtained using a NMR data driven docking approach. Knowledge of the protein interaction surfaces from chemical shift perturbation, dihedral angle restraints determined from Cα, Cβ, CO and N chemical shift data, backbone amide RDCs and long range HN-HN NOE restraints were used to obtain the complex structure by restraint driven docking of a homology model of gIC8

Fab and a crystal structure of IL-1β. The structure of the free gIC8 was obtained by NMR restraint driven refinement of the same gIC8 homology model, using dihedral angle restraints, backbone amide RDCs and long range HN-HN NOE restraints.

As suggested by chemical shift perturbation, comparisons of the structures of the free and bound Fab showed that the relative orientations of the V_H and constant domains remained fixed and that V_L undergoes a small but significant relative change in orientation, which appeared to reveal a negatively charged patch on the Fab surface. Fluorescence resonance energy transfer (FRET) (Dangl et al., 1988; Zheng et al., 1992) and cryo-electron tomography (cryo-ET) (Sandin et al., 2004) have both shown that antibody molecules are very flexible and that is it possible for Fab regions to be in close proximity to the Fc domains. A mechanism for antigen induced B cell signalling is proposed in Chapter 4, where antigen binding to the Fab causes a change in orientation of one of the variable domains with respect to the constant and V_H domains. This reveals a negatively charged patch on the Fab which could interact with a positively charged patch on the surface of its own Fc or the Fc of a neighbouring receptor. This could cause BCR clustering and the initiation of cell signalling pathways (see Figure 4.27).

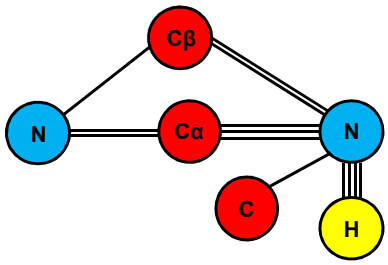
The work described in this thesis shows the potential for NMR techniques in the structural determination of antibody-target protein complexes and the impact that this could have on rational design and humanisation of future therapeutics. In addition, interesting observations have been made on the structural changes induced by antigen binding which has resulted in a proposed mechanism for BCR signalling, which appears

to explain the known key features of activation of the B cell receptor and provides testable predictions for future investigation.

Appendix

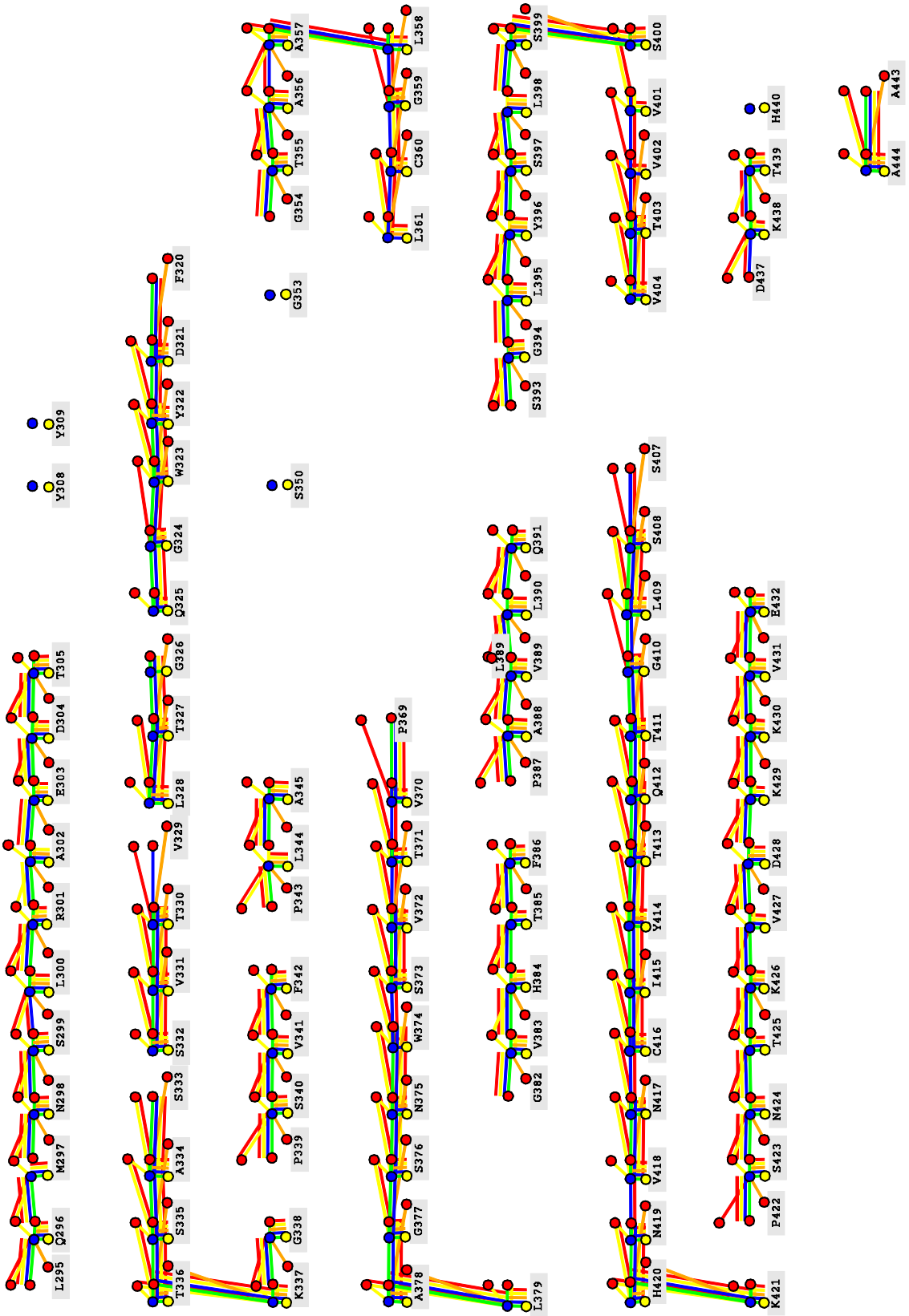
A.1 Summary of backbone assignments

Schematic representation of the connectivities that were used to make backbone assignments for gIC8 Fab. Coloured circles represent backbone atoms (red for carbon, blue for nitrogen and yellow for hydrogen) and atom types are shown in the schematic below. Lines connecting the atoms represent triple resonance spectra used to make the sequence specific backbone assignment with HNCA represented using a green line, HNCOCA with blue, HNCOCOCB with red, HNCACB with yellow and HNCO with orange.



I21
T20
R18
D17
G16
V15
S14
A13
S12
L11
S10
S9
T22
C23
R24
T25
S26
G27
N28
I29
H30
S63
F62
R61
S60
P59
V58
G57
D56
A55
L54
T53
K52
A51
N50
G64
S65
G66
S67
G68
T69
Q70
F71
T72
L73
T74
I75
S76
S77
L78
Q79
P80
E81
D82
F83
A84
E105
V104
K103
T102
G101
Q100
G99
F98
T97
F96
I106
K107
R108
T109
V110
A111
A112
P113
S114
V115
F116
I117
F118
P120
S121
D122
E123
Q124
L125
K126
Q147
V146
K145
A144
E143
R142
P141
Y140
F139
N138
N137
L136
L135
C134
V133
V132
S131
A130
T129
G128
S127
T5
M4
Q3
I2
G41
K42
A43
Y49
I48
L47
L46
Q45
Y86
N85

W148
K149
V150
D151
N152
A153
L154
Q155
G157
N158
S159
Q160
E161
S162
V163
T164
E165
Q166
D167
S168
H189
K188
E187
Y186
D185
A184
K183
S182
L181
T180
L179
T178
S177
S176
L175
Y173
T172
S171
D170
K169
K190
V191
Y192
A193
C194
E195
V196
T197
H198
Q199
G200
L201
P204
V205
T206
K207
S208
F209
N210
R211
G212
E213
C214
S231
G230
G229
P228
Q227
V226
L225
G224
G223
G222
S221
E220
V219
L218
Q217
V216
L232
R233
L234
S235
C236
A237
A238
S239
G240
F241
D242
F243
S244
R245
Y246
R252
Q253
A254
P255
G256
K257
R258
L259
E260
W261
A263
Y264
I265
S266
S267
G268
G269
G270
S271
T272
Y273
T277
Y278
K279
G280
R281
F282
T283
I284
S285
R286
D287
N288
A289
K290
N291
T292
L293


L295 Q296 M297 N298 S299 L300 R301 A302 E303 D304 T305 Y308 Y309
T336 S335 A334 S333 S332 V331 T330 V329 L328 G326 T327 Q325 G324 W323 Y322 D321 T320
K337 G338 P339 S340 V341 F342 P343 L344 A345 S350 G353 G354 T355 A356 A357 L358 G359 C360 L361
A378 G377 S376 N375 W374 S373 V372 T371 V370 P369
L379 G382 V383 H384 T385 F386 P387 A388 V389 L389 L390 Q391 S393 G394 L395 Y396 S397 I398 S399 S400 V401 V402 T403 V404
P422 S423 N424 T425 K426 D428 V427 K429 K430 V431 E432 D437 K438 T439 H440 A443 A444
H420 N419 V418 N417 C416 I415 Y414 T413 Q412 T411 G410 L409 S408 S407

A.2 Chemical shift index data

A.2.1 Sequence specific assignments obtained for gIC8 Fab

Residue	C	CA	CB	H	HE21	HE22	N	NE2
I2	177.1	59.84	36.29	-	-	-	-	-
Q3	-	54.36	29.89	8.9	-	-	128.7	-
M4	176.4	52.98	33.07	8.978	-	-	124	-
T5	-	61.78	70.62	9.381	-	-	119.5	-
S9	177.3	60.98	-	-	-	-	-	-
S10	-	57.04	64.4	7.818	-	-	115.6	-
L11	178.5	54.68	44.47	9.003	-	-	123.2	-
S12	-	56.05	64.15	8.479	-	-	117.6	-
A13	177.5	50.8	22.26	8.67	-	-	127.4	-
S14	178.8	57.9	63.41	8.66	-	-	118.3	-
V15	174.3	64.44	30.44	8.412	-	-	122.7	-
G16	178	44.16	-	10.13	-	-	118.4	-
D17	179.8	54.18	41.15	8.104	-	-	122.8	-
R18	-	58.54	30.74	8.133	-	-	120	-
T20	179.5	60.93	71.13	8.082	-	-	118.5	-
I21	178.1	60.61	40.34	9.227	-	-	126.9	-
T22	178.7	61.91	70.54	9.032	-	-	120.5	-
C23	180.7	55.82	47.51	9.811	-	-	126.3	-
R24	176.8	53.88	32.76	9.308	-	-	127.7	-
T25	177.2	59.48	70.05	8.772	-	-	112.6	-
S26	178.8	59.22	62.64	8.121	-	-	113.4	-
G27	-	43.29	-	7.096	-	-	106.9	-
N28	-	53.57	37.37	-	-	-	-	-
I29	177.2	59.65	38.23	8.115	-	-	119.9	-
H30	-	57.35	27.04	8.314	-	-	119.5	-
G41	-	45.38	-	-	-	-	-	-
K42	177.1	53.36	34.49	8.025	-	-	120.5	-
A43	-	50.11	18.22	8.137	-	-	122.1	-
Q45	177	53.22	31.77	-	-	-	-	-
L46	177.8	56.64	41.77	8.984	-	-	127.7	-
L47	177.5	54.89	45.49	8.594	-	-	125.2	-
I48	178.8	57.77	43.42	7.149	-	-	116.8	-
Y49	-	54.63	39.23	8.915	-	-	122.6	-
N50	174.5	55.28	36.39	-	-	-	-	-
A51	174.5	57.28	18.96	8.984	-	-	114.6	-
K52	-	55.48	35.05	8.212	-	-	115.7	-
T53	180.6	62.92	68.19	8.46	-	-	122.7	-
L54	175.5	55.38	42.19	8.321	-	-	128.7	-
A55	176	51.35	18.22	7.939	-	-	125.5	-
D56	175	55.9	40.4	8.425	-	-	121	-
Residue	C	CA	CB	H	HE21	HE22	N	NE2

G57	178.5	45.23	-	8.719	-	-	112.1	-
V58	-	60.26	32.27	7.528	-	-	124.9	-
P59	175.7	62.95	33	-	-	-	-	-
S60	178.4	59.94	62.41	8.411	-	-	114.1	-
R61	177	56.45	28.68	7.049	-	-	116.5	-
F62	178.5	58.08	39.89	7.763	-	-	121.4	-
S63	180.1	57.14	65.11	8.739	-	-	112.1	-
G64	180.2	43.74	-	8.932	-	-	110.3	-
S65	-	56.88	65.31	8.799	-	-	113.8	-
G66	180.8	44.84	-	-	-	-	-	-
S67	178.3	58.04	64.03	6.936	-	-	109.3	-
G68	180.3	47.14	-	8.834	-	-	112.4	-
T69	177.9	59.97	70.6	7.714	-	-	113	-
Q70	179.9	55.8	30.9	6.751	7.056	6.939	121.4	112.6
F71	177.2	56.57	42.3	8.863	-	-	124.6	-
T72	180	60.11	72.06	9.053	-	-	114.7	-
L73	178.1	53.11	40.97	8.85	-	-	128.4	-
T74	179.4	60.67	70.31	9.029	-	-	123.6	-
I75	-	59.8	37.66	8.751	-	-	126.7	-
S76	179.4	60.14	62.41	8.987	-	-	121.6	-
S77	179.1	56.15	61.72	6.431	-	-	113.7	-
L78	176.4	56.5	41.97	8.764	-	-	126.9	-
Q79	-	52.3	29.09	9.103	-	-	125.5	-
P80	174.4	65.83	30.92	-	-	-	-	-
E81	176.1	57.28	27.66	9.226	-	-	113.4	-
D82	175.5	54.56	40.39	8.39	-	-	121.8	-
F83	178.7	59	37.8	7.209	-	-	121.8	-
A84	178	51.29	19.91	7.624	-	-	126.5	-
N85	178	51.89	40.82	7.772	-	-	115.2	-
Y86	-	56.86	41.13	8.789	-	-	119.7	-
F96	177.2	55.64	37.18	-	-	-	-	-
T97	181.3	60.05	72.02	6.581	-	-	113.3	-
F98	174.2	56.24	43.08	8.629	-	-	117.1	-
G99	180.1	44.65	-	9.236	-	-	107.2	-
Q100	176.7	57.5	27.14	9.016	-	-	117.1	-
G101	-	44.95	-	7.012	-	-	108.2	-
T102	179.1	61.62	72.69	7.897	-	-	118.2	-
K103	177.7	55.75	33.37	8.197	-	-	130	-
V104	-	61.72	32.82	8.469	-	-	129	-
E105	178.3	54.41	32.68	8.918	-	-	126.2	-
I106	175.9	59.13	37.18	9.114	-	-	123.5	-
K107	178.1	55.22	32.48	8.39	-	-	128.8	-
R108	178.6	54.24	29.51	6.034	-	-	119.8	-
T109	177.1	61.5	69.82	7.689	-	-	109.3	-
V110	177.4	63.74	31.41	8.362	-	-	122.7	-
Residue	C	CA	CB	H	HE21	HE22	N	NE2

A111	176.9	50.85	20.14	9.619	-	-	131.4	-
A112	-	49.59	17.04	8.784	-	-	125.7	-
P113	175.6	61.9	30.43	-	-	-	-	-
S114	-	57.85	63.52	7.936	-	-	117	-
V115	178.5	60.85	32.7	8.109	-	-	125.1	-
F116	-	55.26	42.52	9.129	-	-	125.2	-
I117	180.4	58.51	42.09	8.127	-	-	120	-
F118	-	53.89	41.07	9.215	-	-	124.1	-
P120	176.1	62.04	30.6	-	-	-	-	-
S121	-	57.31	65.33	9.37	-	-	118.8	-
D122	173.6	57.31	40.05	8.902	-	-	124	-
E123	173.9	54.4	29.17	8.918	-	-	118.6	-
Q124	-	58.59	25.7	7.737	-	-	121.1	-
L125	172.3	57.9	40.69	7.902	-	-	121.9	-
K126	174.8	58.57	31.61	7.777	-	-	118.2	-
S127	177.3	59.1	64.24	8.007	-	-	114.5	-
G128	178.2	45.71	-	8.006	-	-	108.3	-
T129	170.8	60.94	72.14	7.561	-	-	116.2	-
A130	-	50.41	21.53	8.677	-	-	121.6	-
S131	180.6	57.31	62.17	7.549	-	-	117.2	-
V132	177.8	60.16	31.58	9.168	-	-	127	-
V133	178.3	61.25	34.68	8.09	-	-	125.7	-
C134	180.8	54.13	43.99	9.822	-	-	128	-
L135	179.6	52.92	45.97	9.472	-	-	128.6	-
L136	177.4	53.93	43.22	9.642	-	-	128.2	-
N137	177.7	51.89	41.98	9.474	-	-	122.5	-
N138	180.7	54.03	36.91	8.274	-	-	122.6	-
F139	178.6	54.47	42.74	9.22	-	-	116.8	-
Y140	-	58.93	41.8	9.424	-	-	120.3	-
P141	177.3	61.74	34.78	-	-	-	-	-
R142	176.9	57.95	29.3	8.3	-	-	117.8	-
E143	177.8	57.83	27.99	7.482	-	-	117.1	-
A144	-	50.98	20.76	7.722	-	-	125.8	-
K145	-	54.84	33.84	8.305	-	-	122.1	-
V146	177.3	60.68	32.77	8.728	-	-	126.2	-
Q147	178.8	53.65	31.82	8.994	-	-	126.2	-
W148	176.4	56.48	31.15	9.124	-	-	124.2	-
K149	177.4	54.49	35.72	9.427	-	-	122.2	-
V150	-	59.85	33.16	8.274	-	-	123.4	-
D151	177.2	56.35	39.13	9.879	-	-	131.2	-
N152	179.4	54.72	38.05	9.067	-	-	109.4	-
A153	174.6	50.71	18.94	7.975	-	-	123.7	-
L154	175.6	55.8	41.38	8.795	-	-	128.3	-
Q155	-	53.48	29.99	8.557	-	-	124.8	-
G157	178.3	45.85	-	-	-	-	-	-
Residue	C	CA	CB	H	HE21	HE22	N	NE2

N158	179.3	52.32	37.04	8.232	-	-	115.7	-
S159	178.6	56.2	66.52	7.803	-	-	111.4	-
Q160	177.5	55	33.11	8.156	-	-	117.6	-
E161	177.6	55.13	32.89	8.95	-	-	124.4	-
S162	180.2	56.65	64.68	8.726	-	-	115.1	-
V163	-	59.93	35.07	8.433	-	-	125.6	-
T164	176.6	61.81	70.26	8.698	-	-	115.9	-
E165	176.2	55.54	29.26	8.206	-	-	116.4	-
Q166	-	57	27.89	7.799	-	-	122.9	-
D167	174	55.16	43.52	8.697	-	-	125.2	-
S168	-	61.16	63.27	8.663	-	-	122.2	-
K169	179.6	57.19	32.34	8.664	-	-	120.5	-
D170	176.5	52.87	41.68	8.689	-	-	117.2	-
S171	179.9	58.84	59.61	7.615	-	-	108.4	-
T172	179.6	59.99	71.47	7.623	-	-	107.2	-
Y173	178.3	57.09	41.08	8.857	-	-	116.6	-
S174	181.4	58.55	65.96	7.744	-	-	111.8	-
L175	179	55.15	46.94	9.93	-	-	126.1	-
S176	179.3	55.16	64.55	9.262	-	-	120.5	-
S177	178.3	56.28	63.45	9.29	-	-	120.4	-
T178	179.2	61.25	69.89	8.968	-	-	124.7	-
L179	177.6	53.31	41.25	9.24	-	-	134.8	-
T180	179.5	61.74	69.46	8.168	-	-	123	-
L181	176.6	52.79	46.2	8.9	-	-	127.6	-
S182	177.9	57.82	63.74	9.08	-	-	116.5	-
K183	175.7	60.08	32	8.01	-	-	123.6	-
A184	-	54.45	17.55	8.322	-	-	118.4	-
D185	174.2	56.87	40.25	7.728	-	-	117.7	-
Y186	173.2	58.59	38.52	8.735	-	-	122.1	-
E187	176.6	57.24	28.63	8.061	-	-	114.5	-
K188	176.8	56.49	32.41	7.298	-	-	117.7	-
H189	-	53.68	30.73	7.036	-	-	115.5	-
K190	177.6	58.19	34.54	9.477	-	-	122.3	-
V191	178.2	61.59	32.75	8.56	-	-	119.9	-
Y192	179	57.57	41.44	8.209	-	-	127.5	-
A193	178.6	50.13	22.48	8.814	-	-	122.9	-
C194	-	52.39	43.96	8.718	-	-	120	-
E195	177	54.16	32.15	8.915	-	-	126.5	-
V196	-	61.17	34.04	9.422	-	-	129.6	-
T197	179.3	60.86	70.03	8.675	-	-	120.6	-
H198	177.7	57.72	34.6	8.102	-	-	123.4	-
Q199	176.2	58.24	28.76	8.201	-	-	126.3	-
G200	179.7	45.1	-	10.46	-	-	110	-
L201	-	52.75	42.15	8.113	-	-	121.9	-
P204	175.3	63.32	31.17	-	-	-	-	-
Residue	C	CA	CB	H	HE21	HE22	N	NE2

V205	-	60.5	34.15	8.542	-	-	123.8	-
T206	-	60.88	69.63	8.464	-	-	122.4	-
K207	178.3	52.9	33.07	8.674	-	-	127.4	-
S208	179.4	56.43	67.49	8.614	-	-	116.9	-
F209	181.3	56.67	41.39	8.782	-	-	116.1	-
N210	176.6	50.46	39.61	8.704	-	-	119	-
R211	-	57.83	29.82	9.653	-	-	126	-
G212	-	45.12	-	-	-	-	-	-
E213	177	55.97	30.23	8.057	-	-	120.6	-
C214	-	56.75	35.91	8.052	-	-	124.5	-
V216	177.4	63.29	31.16	-	-	-	-	-
Q217	178.9	55.18	32.21	8.481	-	-	126	-
L218	177	53.48	43.25	8.284	-	-	125.2	-
V219	176.4	61.8	34.43	8.521	-	-	120.9	-
E220	175.8	56.82	30.14	10.65	-	-	133.3	-
S221	-	57.64	65.13	9.423	-	-	115.5	-
G222	178.3	44.52	-	-	-	-	-	-
G223	179.9	45.3	-	7.538	-	-	106.3	-
G224	181.7	44.36	-	7.261	-	-	108.2	-
L225	176.7	53.8	43.63	8.264	-	-	123.2	-
V226	180.2	59.54	34.37	9.095	-	-	127.1	-
Q227	-	53.84	27.1	8.133	-	-	123.8	-
P228	174.9	63.21	30.34	-	-	-	-	-
G229	177.5	44.72	-	9.893	-	-	114.9	-
G230	182	44.23	-	8.549	-	-	110.2	-
S231	179.8	56.12	66.4	7.841	-	-	110.4	-
L232	179.5	54.49	46.51	8.564	-	-	122.9	-
R233	177.1	54.01	31.85	8.068	-	-	124.7	-
L234	176	52.68	43.27	9.008	-	-	127.2	-
S235	-	56.46	66.07	8.748	-	-	114.1	-
C236	181.3	53.99	42.73	9.237	-	-	123.7	-
A237	176.4	50.71	19.52	8.68	-	-	131.1	-
A238	176.3	50	23.11	8.225	-	-	127.4	-
S239	-	57.43	65.04	8.895	-	-	114.6	-
G240	178.8	45.32	-	8.689	-	-	108.5	-
F241	-	54.36	39.18	7.398	-	-	113	-
D242	177.4	52.73	38.38	-	-	-	-	-
F243	177	60.18	39.54	7.722	-	-	129.2	-
S244	179.2	60.4	62.62	8.211	-	-	110.9	-
R245	176.1	54.92	30.02	7.624	-	-	117.3	-
Y246	-	58.25	40.57	7.551	-	-	117.6	-
R252	178.9	52.29	33.08	-	-	-	-	-
Q253	177.6	54.72	31.87	9.175	-	-	121.1	-
A254	-	50.02	17.4	9.676	-	-	137.3	-
P255	174.6	64.07	-	-	-	-	-	-
Residue	C	CA	CB	H	HE21	HE22	N	NE2

G256	178.2	45.61	-	8.68	-	-	113.1	-
K257	176.7	54.45	33.78	8.195	-	-	119.8	-
R258	-	55.32	31.34	7.981	-	-	118.3	-
L259	177.4	55.6	40.74	-	-	-	-	-
E260	-	54.51	33.06	8.993	-	-	120.1	-
W261	-	58.47	28.57	9.887	-	-	131.1	-
A263	177.6	52.46	-	-	-	-	-	-
Y264	179.6	56.52	44.17	8.274	-	-	124	-
I265	181	58.51	40.33	8.651	-	-	125.6	-
S266	174.4	57.46	64.46	8.212	-	-	121.8	-
S267	-	60.96	64.45	8.751	-	-	120.6	-
G268	176.6	44.99	-	-	-	-	-	-
G269	-	44.13	-	7.947	-	-	108.7	-
G270	178.8	45.76	-	-	-	-	-	-
S271	180.3	57.17	64.19	7.691	-	-	115.8	-
T272	179.3	58.52	72.53	7.978	-	-	113.8	-
Y273	-	57.54	42.57	7.966	-	-	118.3	-
T277	176.4	63.35	68.29	-	-	-	-	-
V278	177.6	60.35	31.62	6.91	-	-	110.4	-
K279	-	57.91	31.59	7.394	-	-	124.8	-
G280	-	45.02	-	-	-	-	-	-
R281	177.4	56.9	30.01	7.725	-	-	117.7	-
F282	178.2	52.54	39.18	7.248	-	-	119.5	-
T283	-	60.94	71.27	9.074	-	-	112.2	-
I284	180.8	58.46	42.5	8.824	-	-	123.8	-
S285	180.8	57.31	64.92	8.368	-	-	118.4	-
R286	180	54.23	34.83	9.282	-	-	116.6	-
D287	-	51.84	41.74	9.189	-	-	121.5	-
N288	175.6	55.14	36.82	9.344	-	-	124.5	-
A289	173.7	53.79	17.92	8.492	-	-	122.7	-
K290	176.3	54.56	32.42	7.322	-	-	115.4	-
N291	180.5	53.92	36.45	7.855	-	-	119.6	-
T292	181.9	60.9	72.24	7.214	-	-	108.6	-
L293	-	52.91	45.89	8.765	-	-	127.6	-
L295	178.7	53.52	42.2	-	-	-	-	-
Q296	-	54.3	28.36	8.983	-	-	128.7	-
M297	179.1	54.32	34.76	8.704	-	-	126.4	-
N298	179.5	51.24	41.1	7.976	-	-	120.2	-
S299	176.8	57.32	61.26	8.481	-	-	112.2	-
L300	175.6	56.25	42.17	8.073	-	-	119.2	-
R301	175.7	53.42	33.38	9.686	-	-	121.3	-
A302	172	55.81	17.37	9.134	-	-	125.8	-
E303	176.7	58.25	27.69	9.343	-	-	115.2	-
D304	175.9	54.84	41.7	8.548	-	-	118.4	-
T305	-	64.25	69.57	7.922	-	-	121.1	-
Residue	C	CA	CB	H	HE21	HE22	N	NE2

F320	176.4	55.22	-	-	-	-	-	-
D321	177	54.07	40.55	8.707	-	-	122.3	-
Y322	179.8	58.94	42.43	7.559	-	-	119.9	-
W323	175.1	55.62	32.04	8.607	-	-	122.4	-
G324	-	45.35	-	8.68	-	-	109.4	-
Q325	-	56.47	29.66	8.757	-	-	115.8	-
G326	179.1	43.86	-	9.435	-	-	111.6	-
T327	180.8	59.67	70.22	8.438	-	-	117.5	-
L328	-	55.53	41.93	8.36	-	-	132.1	-
V329	178.8	61.69	33.16	-	-	-	-	-
T330	180.5	62.09	69.69	8.282	-	-	123.4	-
V331	176.6	59.59	30.49	8.423	-	-	128.2	-
S332	-	56.63	63.79	8.618	-	-	119.9	-
S333	178.8	58.52	62.9	-	-	-	-	-
A334	177.1	51.76	18.11	8.263	-	-	127.1	-
S335	178.5	57.18	64.43	7.641	-	-	113.1	-
T336	177.9	63.64	68.78	8.241	-	-	117.9	-
K337	177.2	55.76	35.56	9.554	-	-	128.9	-
G338	-	44.1	-	9.05	-	-	114	-
P339	175.4	61.85	32.51	-	-	-	-	-
S340	180.7	57.9	64.45	8.514	-	-	116	-
V341	178.9	60.3	32.61	8.021	-	-	123.5	-
F342	-	52.59	40.48	9.258	-	-	125.5	-
P343	176.6	62.58	32.97	-	-	-	-	-
L344	178.7	52.67	38.29	8.476	-	-	124.7	-
A345	-	49.77	18.54	8.091	-	-	124.6	-
G354	177.8	45.41	-	-	-	-	-	-
T355	180.5	61.63	70.93	8.157	-	-	118.6	-
A356	178.5	49.86	22.1	9.405	-	-	127.8	-
A357	-	49.42	20.29	7.553	-	-	120.7	-
L358	177.7	53.49	43.67	8.439	-	-	117.5	-
G359	180.9	48.48	-	8.951	-	-	105.4	-
C360	181.4	54.3	46.12	9.582	-	-	118	-
L361	-	54.12	43.12	9.481	-	-	127.5	-
P369	-	61.73	34.18	-	-	-	-	-
V370	-	59.93	34.33	7.947	-	-	116.4	-
T371	179.2	60.3	70.34	8.443	-	-	117.5	-
V372	177.5	60.58	33.63	8.759	-	-	124.3	-
S373	-	56.09	66.27	8.903	-	-	123.1	-
W374	175.2	56.16	31.06	9.252	-	-	120.1	-
N375	175.7	53.85	35.65	10.49	-	-	125.9	-
S376	178	58.74	60.91	9.748	-	-	112.4	-
G377	178.3	44.09	-	7.878	-	-	106	-
A378	175.3	53.05	18.56	7.779	-	-	121.4	-
L379	-	54.05	41.79	7.239	-	-	119.7	-
Residue	C	CA	CB	H	HE21	HE22	N	NE2

G382	179.4	45.52	-	-	-	-	-	-
V383	175.8	61.3	32.44	7.013	-	-	119.4	-
H384	-	54.28	30.18	9.35	-	-	125.7	-
T385	178.9	61.95	68.16	8.849	-	-	123.9	-
F386	-	58.35	38.19	8.707	-	-	128.7	-
P387	175.3	62.49	31.38	-	-	-	-	-
A388	174.5	52.32	20.22	9.119	-	-	127.3	-
L389	-	-	44.45	-	-	-	-	-
V389	177.8	59.7	33.97	8.831	-	-	118.3	-
L390	175.9	54.22	41.2	8.659	-	-	127.8	-
Q391	-	54.96	32.12	8.252	-	-	125.9	-
S393	177.2	57.42	63.74	-	-	-	-	-
G394	180.9	44.72	-	8.169	-	-	110.1	-
L395	175.3	51.93	43.23	6.833	-	-	115.8	-
Y396	176	58.77	39.5	9.148	-	-	123.4	-
S397	180.1	57.37	65.64	8.396	-	-	111.7	-
L398	178.1	55.17	-	9.57	-	-	125.9	-
S399	180.1	57.6	65.72	7.423	-	-	114.5	-
S400	-	56.6	63.9	8.691	-	-	116.4	-
V401	-	58.72	35.35	9.238	-	-	123.8	-
V402	178.4	58.64	33.6	8.986	-	-	121.7	-
T403	178.3	58.51	69.34	7.507	-	-	115.1	-
V404	-	57.04	33	8.736	-	-	118	-
S407	177	59.74	62.3	-	-	-	-	-
S408	177.3	59.09	64.04	7.947	-	-	116.9	-
L409	174.4	56.77	40.51	7.481	-	-	122.7	-
G410	177.8	45.15	-	8.765	-	-	107.4	-
T411	-	62.02	69.76	7.584	-	-	110.2	-
Q412	178.8	54.11	30.3	7.887	-	-	122	-
T413	179.4	61.62	69.8	8	-	-	118.3	-
Y414	-	58.19	40.3	9.287	-	-	126.8	-
I415	178.1	58.79	41.49	8.973	-	-	124	-
C416	-	52.23	42.72	8.714	-	-	123.7	-
N417	178.5	51.67	39.06	9.009	-	-	123.3	-
V418	-	61.18	33	9.341	-	-	126.6	-
N419	178.9	51.78	40.95	8.72	-	-	126.1	-
H420	179.3	54.33	32.62	8.954	-	-	126.5	-
K421	-	60.2	29.01	9.088	-	-	128	-
P422	176.8	65.76	31.37	-	-	-	-	-
S423	178.6	57.08	63.68	6.526	-	-	106.5	-
N424	-	54.3	37.38	8.295	-	-	121	-
T425	179.2	62.93	69.21	7.83	-	-	114.1	-
K426	-	55.4	33.7	8.57	-	-	130	-
V427	179.8	60.47	35.84	8.722	-	-	124.7	-
D428	-	52.42	42.48	8.337	-	-	126.9	-
Residue	C	CA	CB	H	HE21	HE22	N	NE2

K429	179.3	53.02	33.56	8.705	-	-	125	-
K430	176.5	56.26	32.3	8.573	-	-	128.1	-
V431	177.7	62.46	31.42	9.043	-	-	129.4	-
E432	-	53.02	30.91	8.615	-	-	126.9	-
D437	-	54.28	41.24	-	-	-	-	-
K438	175.8	56.1	32.05	8.311	-	-	122.2	-
T439	-	62.29	69.32	8.201	-	-	114.7	-
A443	176.7	52.17	18.71	-	-	-	-	-
A444	-	53.43	19.65	7.845	-	-	129.6	-

A.2.2. Sequence specific assignments obtained for IL1β

Residue	C	CA	CB	H	N
P2	175.6	60.45	28.79	-	-
V3	-	58.19	31.03	8.055	118.5
R4	177	54.01	28.18	-	-
S5	172.2	55.23	61.73	8.083	118.3
L6	175.5	50.86	40.27	9.052	126.6
N7	176	50.61	36.54	9.17	123.6
C8	171.6	53.33	29.5	9.346	114.2
G9	172.9	41.48	-	8.555	105.9
L10	176.1	51.61	44.17	9.152	121.7
R11	176.2	51.97	31.84	8.575	119.5
D12	179.7	50.9	38.4	8.84	121
S13	175.4	58.04	59.62	7.568	114.5
Q14	175.2	52.73	25.54	8.095	120.5
Q15	-	55.1	22.25	8.434	111
K16	176	56.43	27.19	8.483	119.9
S17	-	57.02	62.66	8.383	119.7
L18	176	52.62	38.81	8.484	122.2
V19	174.6	57.08	33.27	8.917	116.5
M20	176.3	53.24	29.38	8.736	121
S21	174.4	54.65	60.55	8.632	120
G22	-	42.23	-	8.136	111.6
P23	176.4	61.89	28.62	-	-
Y24	175.1	53.72	36.39	8.05	110.5
E25	173.2	53.65	31.65	7.238	119.5
L26	176.5	50.76	42.26	8.607	122.9
K27	174	52.73	34.72	9.215	119.6
A28	177.6	47.37	18.65	7.576	118.7
L29	175.3	53.08	41.84	9.384	125.7
H30	174.4	57.07	25.42	10.19	122.1
L31	176.7	51.5	41.01	8.496	126
Q32	176.8	52.65	27.67	8.933	120.4
Residue	C	CA	CB	H	N

G33	-	44.33	-	8.733	109.6
Q34	176.6	54.65	25.49	8.814	121.7
D35	177.1	52.94	38.33	8.031	118.9
M36	177.7	55.14	29.06	7.793	118.3
E37	176.8	55.37	26.37	8.134	118
Q38	176	53.86	26.48	8.021	116.8
Q39	175.8	54.09	26.36	7.591	118.1
V40	174.8	59.97	29.68	8.461	123.9
V41	176	59.55	29.49	7.959	125.6
F42	175.9	55.06	38.4	9.459	126.7
S43	174	54.58	61.27	9.16	115.4
M44	175.2	52.09	33.28	9.62	131.4
S45	172.5	54.53	63.65	8.57	119.3
F46	176.5	56.52	35.52	9.44	124.2
V47	175.5	57.44	31.02	7.442	116.9
Q48	176.1	53.95	26.67	8.508	119.6
G49	172	42.74	-	8.135	111.2
E50	-	54.57	26.74	8.496	122.3
E51	175.4	53.18	28.7	8.457	124.5
S52	-	54.82	62.4	8.462	117.9
N53	175	53.12	-	-	-
D54	174.9	52.7	38.44	8.415	113.3
K55	175	53.68	31.93	7.431	118.9
I56	-	55.59	39.05	8.179	121.2
P57	177.5	60.51	28.57	-	-
V58	173.6	56.73	34.39	10.1	118.3
A59	176.2	48.35	20.6	8.728	120.5
L60	173.3	51.67	-	10.45	123
G61	173.1	40.46	-	8.271	110.3
L62	-	52.03	38.58	8.208	120.7
K63	-	53.51	35.66	8.077	120.5
N66	173.8	52.08	-	-	-
L67	175.2	50.65	44.69	7.007	116.1
Y68	175.8	55.1	40.3	9.166	120.8
L69	176.5	54.51	39.2	8.597	121.8
S70	173.3	53.73	64.52	9.165	115.7
C71	174.3	54.4	27.6	8.456	119.7
V72	174.2	56.39	33.25	8.817	116.9
L73	176.7	52.21	39.76	8.567	123.6
K74	176	53.42	31.13	8.259	126.2
D75	174.9	53.73	36.85	9.255	127.5
D76	174.7	53.59	37.64	8.498	109.9
K77	-	50.79	31.75	7.734	119.9
P78	175	61.03	28.35	-	-
T79	171.8	59.89	70.23	8.8	119.6
Residue	C	CA	CB	H	N

L80	175.5	51.69	41.53	8.874	126.7
Q81	173.7	51.3	29.04	9.772	126.2
L82	175	52.01	38.98	8.512	122
E83	175.5	52.5	30.37	9.184	124.5
S84	174.7	56.35	60.97	8.722	121.5
V85	174.7	56.01	32.17	8.021	117.1
D86	-	48.73	39.62	8.047	121.5
P87	177.7	61.68	-	-	-
K88	177.7	55.59	29.12	8.498	117.2
N89	172.6	50.61	39.3	7.815	114
Y90	-	54.09	38.75	7.363	116
P91	173.8	60.42	-	-	-
K92	174	52.04	31.87	7.11	115.8
K93	176.5	56.91	29.59	8.376	119.5
K94	176.1	51.78	27.86	7.903	117.8
M95	176.1	52.03	31.63	7.538	122.2
E96	177.3	57.34	25.43	9.241	122.3
K97	178.2	57.52	29.29	8.238	121.5
R98	176.4	56	25.04	7.921	113.9
F99	175.3	55.58	-	7.992	116.7
V100	175.4	61.14	28.99	7.519	117.8
F101	174.2	53.75	39.39	9.825	126.8
N102	173.8	49.72	36.75	10.49	120.8
K103	176.2	54.14	28.57	9.376	127.6
I104	175.1	58.48	38.25	9.318	133.4
E105	175.3	53.47	26.26	8.668	126.7
I106	174.6	57.43	36.95	8.371	122.9
N107	174.3	52.39	34.81	8.879	122.3
N108	173.5	52.58	35.04	8.69	112.8
K109	174.4	52.17	34.22	7.725	116.1
L110	175.9	51.19	42.76	9.579	121.1
E111	174.4	51.44	30.6	8.673	117.6
F112	173.9	54.14	38.99	10.33	119
E113	177.4	51.02	30.42	8.807	125.8
S114	174.6	56.72	60.64	9.153	121.7
A115	179.2	51.28	15.46	8.153	128.1
Q116	174.7	54.36	26.45	7.862	116.9
F117	-	50.95	36.08	7.427	116.9
P118	178.2	61.09	28.82	-	-
N119	173.2	53.31	35.4	10.3	114
W120	175.4	54.93	26.28	8.143	119.6
Y121	179	54.62	39.09	9.438	119
I122	175.3	63.66	33.84	8.751	124.7
S123	171.4	55.19	64.76	8.987	122
T124	174.6	56.48	68.41	9.116	110
Residue	C	CA	CB	H	N

S125	176	54.64	61.7	9.603	113.6
Q126	177.6	56.03	25.88	8.966	122
A127	176.7	49.05	16.88	8.208	120.8
E128	177.3	54.03	28.17	8.491	117.1
N129	173.5	52.34	35.45	8.077	113.8
M130	-	49.64	-	8.906	117.7
P131	178.6	60.61	-	-	-
V132	175.7	60.46	29.19	8.04	125.3
F133	171.8	53.36	39.1	8.776	124.5
L134	176.4	50.94	40.87	8.72	120
G135	173.2	42.73	-	9.372	114.7
G136	174.7	41.9	-	8.592	109.9
T137	174	59.03	67.88	7.432	112.4
K138	177.2	53.57	30.05	8.47	124.6
G139	174.9	43.18	-	8.78	111.4
G140	174.3	42.45	-	8.218	109.4
Q141	176.2	54.35	26.64	8.492	120
D142	175.4	52.16	39.27	7.802	119.2
I143	175.7	59.34	33.89	9.271	125.6
T144	172.9	57.97	-	8.072	112
D145	174.2	50.25	40.56	7.315	120.2
F146	176.1	54.58	42.04	9.354	117.3
T147	173.4	58.82	69.22	9.54	108.5
M148	175	51.63	35.78	8.75	120.5
Q149	177	52.04	27.82	8.374	126.8
F150	174.9	57.43	37.18	8.891	124.8
V151	174.6	57.43	32.64	8.502	116.5
S152	173.8	57.53	61.42	8.664	117.3
S153	-	59.13	61.63	8.351	124

References

Al-Hashimi, H. M., Valafar, H., Terrell, M., Zartler, E. R., Eidsness, M. K., and Prestegard, J. H. (2000). Variation of molecular alignment as a means of resolving orientational ambiguities in protein structures from dipolar couplings. Journal Of Magnetic Resonance *143*, 402-406.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research *25*, 3389-3402.

Andrusier, N., Mashiach, E., Nussinov, R., and Wolfson, H. J. (2008). Principles of flexible protein-protein docking. Proteins-Structure Function And Bioinformatics *73*, 271-289.

Arbabi-Ghahroudi, M., Tanha, J., and MacKenzie, R. (2005). Prokaryotic expression of antibodies. Cancer and Metastasis Reviews *24*, 501-519.

Ay, J., Keitel, T., Kuttner, G., Wessner, H., Scholz, C., Hahn, M., and Hohne, W. (2000). Crystal structure of a phage library-derived single-chain Fv fragment complexed with turkey egg-white lysozyme at 2.0 angstrom resolution. Journal Of Molecular Biology *301*, 239-246.

Barrientos, L. G., Louis, J. M., and Gronenborn, A. M. (2001). Characterization of the cholesteric phase of filamentous bacteriophage fd for molecular alignment. Journal of Magnetic Resonance *149*, 154-158.

Bassing, C. H., Swat, W., and Alt, F. W. (2002). The mechanism and regulation of chromosomal V(D)J recombination. Cell *109*, S45-S55.

Bax, A. (2003). Weak alignment offers new NMR opportunities to study protein structure and dynamics. Protein Science *12*, 1-16.

Bax, A., Kontaxis, G., and Tjandra, N. (2001). Dipolar couplings in macromolecular structure determination. Nuclear Magnetic Resonance of Biological Macromolecules, Pt B *339*, 127-174.

Berger, C., Weber-Bornhauser, S., Eggenberger, J., Hanes, J., Pluckthun, A., and Bosshard, H. R. (1999). Antigen recognition by conformational selection. Febs Letters *450*, 149-153.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000). The Protein Data Bank. Nucleic Acids Research *28*, 235-242.

Bewley, C. A. (2001). Rapid validation of the overall structure of an internal domain-swapped mutant of the anti-HIV protein cyanovirin-N using residual dipolar couplings. Journal Of The American Chemical Society *123*, 1014-1015.

Bhat, T. N., Bentley, G. A., Fischmann, T. O., Boulot, G., and Poljak, R. J. (1990). Small Rearrangements In Structures Of Fv And Fab Fragments Of Antibody D1.3 On Antigen-Binding. Nature *347*, 483-485.

Blackledge, M. (2005). Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. Progress in Nuclear Magnetic Resonance Spectroscopy *46*, 23-61.

Bongini, L., Fanelli, D., Piazza, F., De los Rios, P., Sandin, S., and 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.

Braden, B. C., Souchon, H., Eisele, J. L., Bentley, G. A., Bhat, T. N., Navaza, J., and Poljak, R. J. (1994). 3-Dimensional Structures Of The Free And The Antigen-Complexed Fab From Monoclonal Antilysozyme Antibody-D44.1. Journal Of Molecular Biology *243*, 767-781.

Brekke, O. H., and 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. C., Pleiman, C. M., and Clark, M. R. (1994). Signal-Transduction By The B-Cell Antigen Receptor And Its Coreceptors. Annual Review Of Immunology *12*, 457-486.

Carsetti, L., Laurenti, L., Gobessi, S., Longo, P. G., Leone, G., and Efremov, D. G. (2009). Phosphorylation of the activation loop tyrosines is required for sustained Syk signaling and growth factor-independent B-cell proliferation. Cellular Signalling *21*, 1187.

Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, and Skelton, N. J. (1996). Protein NMR spectroscopy: Principles and practice, In Protein NMR spectroscopy: Principles and practice (Academic Press, Inc.; Academic Press Ltd.).

Chadd, H. E., and Chamow, S. M. (2001). Therapeutic antibody expression technology. Current Opinion in Biotechnology *12*, 188-194.

Chames, P., Van Regenmortel, M., Weiss, E., and 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.

Chou, J. J., Li, S. P., Klee, C. B., and Bax, A. (2001). Solution structure of Ca2+-calmodulin reveals flexible hand-like properties of its domains. Nature Structural Biology *8*, 990-997.

Choy, E. H. S., Hazleman, B., Smith, M., Moss, K., Lisi, L., Scott, D. G. I., Patel, J., Sopwith, M., and Isenberg, D. A. (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., Driscoll, P. C., Wingfield, P. T., and Gronenborn, A. M. (1990). Analysis of the Backbone Dynamics of Interleukin-1-Beta Using 2-Dimensional Inverse Detected Heteronuclear N-15-H-1 Nmr-Spectroscopy. Biochemistry *29*, 7387-7401.

Clore, G. M., and Gronenborn, A. M. (1991). Structures of larger proteins in solution: three- and four-dimensional heteronuclear NMR spectroscopy. Science *252*, 1390-1399.

Clore, G. M., and Gronenborn, A. M. (1998). NMR structure determination of proteins and protein complexes larger than 20 kDa. Current Opinion In Chemical Biology *2*, 564-570.

Clore, G. M., and Schwieters, C. D. (2004). How much backbone motion in ubiquitin is required to account for dipolar coupling data measured in multiple alignment media as assessed by independent cross-validation? Journal of the American Chemical Society *126*, 2923-2938.

Clore, G. M., Wingfield, P. T., and Gronenborn, A. M. (1991). High-Resolution 3-Dimensional Structure of Interleukin-1-Beta in Solution by 3-Dimensional and 4-Dimensional Nuclear-Magnetic-Resonance Spectroscopy. Biochemistry *30*, 2315-2323.

Cornilescu, G., Delaglio, F., and Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. Journal of Biomolecular Nmr *13*, 289-302.

Cornilescu, G., Marquardt, J. L., Ottiger, M., and Bax, A. (1998). Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. Journal of the American Chemical Society *120*, 6836-6837.

Covell, D. G., Barbet, J., Holton, O. D., Black, C. D. V., Parker, R. J., and Weinstein, J. N. (1986). Pharmacokinetics Of Monoclonal Immunoglobulin G1, F(Ab')2, And Fab' In Mice. Cancer Research *46*, 3969-3978.

Dal Porto, J. M., Gauld, S. B., Merrell, K. T., Mills, D., Pugh-Bernard, A. E., and Cambier, J. (2004). B cell antigen receptor signaling 101. Molecular Immunology *41*, 599-613.

Dangl, J. L., Wensel, T. G., Morrison, S. L., Stryer, L., Herzenberg, L. A., and Oi, V. T. (1988). Segmental Flexibility and Complement-Fixation of Genetically Engineered Chimeric Human, Rabbit and Mouse Antibodies. Embo Journal *7*, 1989-1994.

Davies, D. R., Padlan, E. A., and Segal, D. M. (1975). 3-Dimensional Structure of Immunoglobulins. Annual Review of Biochemistry *44*, 639-667.

Deisenhofer, J. (1981). Crystallographic Refinement And Atomic Models Of A Human Fc Fragment And Its Complex With Fragment-B Of Protein-A From Staphylococcus-Aureus At 2.9-A And 2.8-A Resolution. Biochemistry *20*, 2361-2370.

Delaglio, F., Kontaxis, G., and Bax, A. (2000). Protein structure determination using molecular fragment replacement and NMR dipolar couplings. Journal Of The American Chemical Society *122*, 2142-2143.

Diaz, M., and Casali, P. (2002). Somatic immunoglobulin hypermutation. Current Opinion In Immunology *14*, 235-240.

Dinarello, C. A. (1988). BIOLOGY OF INTERLEUKIN-1. FASEB JOURNAL *2*, 108-115.

Dinarello, C. A. (2002). The IL-1 family and inflammatory diseases. Clinical and Experimental Rheumatology *20*, S1-S13.

Dinarello, C. A. (2009). Immunological and Inflammatory Functions of the Interleukin-1 Family. Annual Review of Immunology *27*, 519-550.

Dominguez, C., Boelens, R., and Bonvin, A. M. J. J. (2003). HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. Journal of the American Chemical Society *125*, 1731-1737.

Faelber, K., Kirchhofer, D., Presta, L., Kelley, R. F., and Muller, Y. A. (2001). The 1.85 angstrom resolution crystal structures of tissue factor in complex with humanized Fab

D3h44 and of free humanized Fab D3h44: Revisiting the solvation of antigen combining sites. Journal Of Molecular Biology *313*, 83-97.

Fernandez, L. A. (2004). Prokaryotic expression of antibodies and affibodies. Current Opinion In Biotechnology *15*, 364-373.

Fischer, M. W. F., Losonczi, J. A., Weaver, J. L., and Prestegard, J. H. (1999). Domain orientation and dynamics in multidomain proteins from residual dipolar couplings. Biochemistry *38*, 9013-9022.

Freund, C., Ross, A., Pluckthun, A., and Holak, T. A. (1994). Structural And Dynamic Properties Of The Fv Fragment And The Single-Chain Fv Fragment Of An Antibody In Solution Investigated By Heteronuclear 3-Dimensional Nmr-Spectroscopy. Biochemistry *33*, 3296-3303.

Fushman, D., Varadan, R., Assfalg, M., and Walker, O. (2004). Determining domain orientation in macromolecules by using spin-relaxation and residual dipolar coupling measurements. Progress in Nuclear Magnetic Resonance Spectroscopy *44*, 189-214.

Gabel, F., Simon, B., Nilges, M., Petoukhov, M., Svergun, D., and Sattler, M. (2008). A structure refinement protocol combining NMR residual dipolar couplings and small angle scattering restraints. Journal of Biomolecular Nmr *41*, 199-208.

Geisberger, R., and Achatz, G. (2003). Models of signal transduction through the B-cell antigen receptor. IMMUNOLOGY *110*, 401.

Giudicelli, V., Chaume, D., Bodmer, J., Muller, W., Busin, C., Marsh, S., Bontrop, R., Marc, L., Malik, A., and Lefranc, M. P. (1997). IMGT, the international ImMunoGeneTics database. Nucleic Acids Research *25*, 206-211.

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

Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C., and Kay, L. E. (1999). A robust and cost-effective method for the production of Val, Leu, Ile (delta 1) methyl-protonated N-15-, C-13-, H-2-labeled proteins. Journal of Biomolecular Nmr *13*, 369-374.

Gray, J. J. (2006). High-resolution protein-protein docking. Current Opinion In Structural Biology *16*, 183-193.

Green, L. L. (1999). Antibody engineering via genetic engineering of the mouse: XenoMouse strains are a vehicle for the facile generation of therapeutic human monoclonal antibodies. Journal of Immunological Methods *231*, 11-23.

Grzesiek, S., and 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.

Hansen, M. R., Hanson, P., and Pardi, A. (2000). Filamentous bacteriophage for aligning RNA, DNA, and proteins for measurement of nuclear magnetic resonance dipolar coupling interactions. Rna-Ligand Interactions Pt A *317*, 220-240.

Hare, B. J., Prestegard, J. H., and Engelman, D. M. (1995). Small angle x-ray scattering studies of magnetically oriented lipid bilayers. Biophysical Journal *69*, 1891-1896.

Hare, B. J., Wyss, D. F., Osburne, M. S., Kern, P. S., Reinherz, E. L., and Wagner, G. (1999). Structure, specificity and CDR mobility of a class II restricted single-chain T-cell receptor. Nature Structural Biology *6*, 574-581.

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

Hoogenboom, H. R. (2005). Selecting and screening recombinant antibody libraries. Nature Biotechnology *23*, 1105-1116.

Huang, M. D., Syed, R., Stura, E. A., Stone, M. J., Stefanko, R. S., Ruf, W., Edgington, T. S., and Wilson, I. A. (1998). The mechanism of an inhibitory antibody on TF-initiated blood coagulation revealed by the crystal structures of human tissue factor, Fab5G9 and TF center dot 5G9 complex. Journal Of Molecular Biology *275*, 873-894.

Humphreys, D. P., Heywood, S. P., Henry, A., Ait-Lhadj, L., Antoniw, P., Palframan, R., Greenslade, K. J., Carrington, B., Reeks, D. G., Bowering, L. C.*, et al.* (2007). Alternative antibody Fab' fragment PEGylation strategies: combination of strong reducing agents, disruption of the interchain disulphide bond and disulphide engineering. Protein Engineering Design & Selection *20*, 227-234.

Hwang, W. C., Lin, Y. Q., Santelli, E., Sui, J. H., Jaroszewski, L., Stec, B., Farzan, M., Marasco, W. A., and Liddington, R. C. (2006). Structural basis of neutralization by a human anti-severe acute respiratory syndrome spike protein antibody, 80R. Journal Of Biological Chemistry *281*, 34610-34616.

Ishii, Y., Markus, M. A., and Tycko, R. (2001). Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel. Journal of Biomolecular Nmr *21*, 141-151.

Jain, M., Kamal, N., and Batra, S. K. (2007). Engineering antibodies for clinical applications. Trends In Biotechnology *25*, 307-316.

Jain, N. U., Tjioe, E., Savidor, A., and Boulie, J. (2005). Redox-dependent structural differences in putidaredoxin derived from homologous structure refinement via residual dipolar couplings. Biochemistry *44*, 9067-9078.

Janeway, C. (2005). Immunobiology: the immune system in health and disease, 6th edn (New York: Garland Science).

Jimenez, R., Salazar, G., Baldridge, K. K., and Romesberg, F. E. (2003). Flexibility and molecular recognition in the immune system. Proceedings Of The National Academy Of Sciences Of The United States Of America *100*, 92-97.

Jimenez, R., Salazar, G., Yin, J., Joo, T., and Romesberg, F. E. (2004). Protein dynamics and the immunological evolution of molecular recognition. Proceedings Of The National Academy Of Sciences Of The United States Of America *101*, 3803-3808.

Josephson, K., Jones, B. C., Walter, L. J., DiGiacomo, R., Indelicato, S. R., and Walter, M. R. (2002). Noncompetitive antibody neutralization of IL-10 revealed by protein engineering and X-ray crystallography. Structure *10*, 981-987.

Jung, D., Giallourakis, C., Mostoslavsky, R., and Alt, F. W. (2006). Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. Annual Review Of Immunology *24*, 541-570.

Kanelis, V., Forman-Kay, J. D., and Kay, L. E. (2001). Multidimensional NMR methods for protein structure determination. Iubmb Life *52*, 291-302.

Kay, L. E. (2005). NMR studies of protein structure and dynamics. Journal of Magnetic Resonance *173*, 193-207.

Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990). 3-Dimensional Triple-Resonance Nmr-Spectroscopy of Isotopically Enriched Proteins. Journal of Magnetic Resonance *89*, 496-514.

Koenig, B. W., Hu, J. S., Ottiger, M., Bose, S., Hendler, R. W., and Bax, A. (1999). NMR measurement of dipolar couplings in proteins aligned by transient binding to

purple membrane fragments. Journal of the American Chemical Society *121*, 1385-1386.

Koharudin, L. M. I., Bonvin, A. M. J. J., Kaptein, R., and Boelens, R. (2003). Use of very long-distance NOEs in a fully deuterated protein: an approach for rapid protein fold determination. Journal of Magnetic Resonance *163*, 228-235.

Kohler, G., and Milstein, C. (1975). Continuous Cultures of Fused Cells Secreting Antibody of Predefined Specificity. Nature *256*, 495-497.

Kontaxis, G., Clore, G. M., and Bax, A. (2000). Evaluation of cross-correlation effects and measurement of one-bond couplings in proteins with short transverse relaxation times. Journal of Magnetic Resonance *143*, 184-196.

Koradi, R., Billeter, M., and Wuthrich, K. (1996). MOLMOL: A program for display and analysis of macromolecular structures. Journal of Molecular Graphics *14*, 51-&.

Leiting, B., Marsilio, F., and O'Connell, J. F. (1998). Predictable deuteration of recombinant proteins expressed in Escherichia coli. Analytical Biochemistry *265*, 351-355.

Li, S., Kussie, P., and Ferguson, K. M. (2008). Structural basis for EGF receptor inhibition by the therapeutic antibody IMC-11F8. Structure *16*, 216-227.

Li, Y. L., Li, H. M., Smith-Gill, S. J., and Mariuzza, R. A. (2000). Three-dimensional structures of the free and antigen-bound Fab from monoclonal antilysozyme antibody HyHEL-63. Biochemistry *39*, 6296-6309.

Lipsitz, R. S., and Tjandra, N. (2004). Residual dipolar couplings in NMR structure analysis. Annual Review of Biophysics and Biomolecular Structure *33*, 387-413.

Liu, D., Cocco, M., Rosenfied, R., Lewis, J., Ren, D., Li, L., Remmele, R., and Brems, D. (2007a). Assignment of backbone 1H, 13C and 15N resonances of human IgG1 Fc (51.4 kDa). Biomolecular NMR Assignments *1*, 233.

Liu, D. J., Cocco, M., Matsumura, M., Ren, D., Becker, B., Remmele, R. L., and Brems, D. N. (2007b). Assignment of H-1, C-13 and N-15 resonances of the reduced human IgG1 C(H)3 domain. Biomolecular Nmr Assignments *1*, 93-94.

Lukin, J. A., Kontaxis, G., Simplaceanu, V., Yuan, Y., Bax, A., and Ho, C. (2003). Quaternary structure of hemoglobin in solution. Proceedings of the National Academy of Sciences of the United States of America *100*, 517-520.

MacDonald, D., Herbert, K., Zhang, X. L., Polgruto, T., and Lu, P. (2001). Solution structure of an A-tract DNA bend. Journal Of Molecular Biology *306*, 1081-1098.

Malby, R. L., McCoy, A. J., Kortt, A. A., Hudson, P. J., and Colman, P. M. (1998). Three-dimensional structures of single-chain Fv-neuraminidase complexes. Journal Of Molecular Biology *279*, 901-910.

Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989). 3-Dimensional Heteronuclear Nmr of N-15-Labeled Proteins. Journal of the American Chemical Society *111*, 1515-1517.

McCafferty, J., Griffiths, A. D., Winter, G., and Chiswell, D. J. (1990). Phage Antibodies - Filamentous Phage Displaying Antibody Variable Domains. Nature *348*, 552-554.

McCallum, S. A., and Pardi, A. (2003). Refined solution structure of the iron-responsive element RNA using residual dipolar couplings. Journal Of Molecular Biology *326*, 1037-1050.

McCoy, M. A., and Wyss, D. F. (2002). Structures of protein-protein complexes are docked using only NMR restraints from residual dipolar coupling and chemical shift perturbations. Journal of the American Chemical Society *124*, 2104-2105.

McCulloch, C. A., Downey, G. P., and El-Gabalawy, H. (2006). Signalling platforms that modulate the inflammatory response: new targets for drug development. Nature Reviews Drug Discovery *5*, 864-876.

McDonnell, J. (2001). Surface plasmon resonance: towards an understanding of the mechanisms of biological molecular recognition. CURRENT OPINION IN CHEMICAL BIOLOGY *5*, 572.

McDonnell, J. M., Calvert, R., Beavil, R. L., Beavil, A. J., Henry, A. J., Sutton, B. J., Gould, H. J., and Cowburn, D. (2001). The structure of the IgE C epsilon 2 domain and its role in stabilizing the complex with its high-affinity receptor Fc epsilon Rl alpha. Nature Structural Biology *8*, 437-441.

Meiler, J., Blomberg, N., Nilges, M., and Griesinger, C. (2000). A new approach for applying residual dipolar couplings as restraints in structure elucidation (vol 16, pg 245, 2000). Journal of Biomolecular Nmr *17*, 185-185.

Meiler, J., Prompers, J. J., Peti, W., Griesinger, C., and Bruschweiler, R. (2001). Model-free approach to the dynamic interpretation of residual dipolar couplings in globular proteins. Journal Of The American Chemical Society *123*, 6098-6107.

Mirza, O., Henriksen, A., Ipsen, H., Larsen, J. N., Wissenbach, M., Spangfort, M. D., and Gajhede, M. (2000). Dominant epitopes and allergic cross-reactivity: Complex formation between a Fab fragment of a monoclonal murine IgG antibody and the major allergen from birch pollen Bet v 1. Journal Of Immunology *165*, 331-338.

Mukherjee, S., Pondaven, S. P., Hofer, N., and Jaroniec, C. P. (2009). Backbone and side-chain H-1, C-13 and N-15 resonance assignments of LEN, a human immunoglobulin kappa IV light-chain variable domain. Biomolecular Nmr Assignments *3*, 255-259.

Mylvaganam, S. E., Paterson, Y., and Getzoff, E. D. (1998). Structural basis for the binding of an anti-cytochrome c antibody to its antigen: Crystal structures of FabE8 cytochrome c complex to 1.8 angstrom resolution and FabE8 to 2.26 angstrom resolution. Journal Of Molecular Biology *281*, 301-322.

Nelson, A. L., and Reichert, J. M. (2009). Development trends for therapeutic antibody fragments. Nature Biotechnology *27*, 331-337.

Nybakken, G. E., Oliphant, T., Johnson, S., Burke, S., Diamond, M. S., and Fremont, D. H. (2005). Structural basis of West Nile virus neutralization by a therapeutic antibody. Nature *437*, 764-768.

Oi, V. T., Vuong, T. M., Hardy, R., Reidler, J., Dangl, J., Herzenberg, L. A., and Stryer, L. (1984). Correlation Between Segmental Flexibility And Effector Function Of Antibodies. Nature *307*, 136-140.

Ottiger, M., and Bax, A. (1998). Characterization of magnetically oriented phospholipid micelles for measurement of dipolar couplings in macromolecules. Journal of Biomolecular Nmr *12*, 361-372.

Ottiger, M., Delaglio, F., and Bax, A. (1998). Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. Journal of Magnetic Resonance *131*, 373-378.

Otting, G. (2008). Prospects for lanthanides in structural biology by NMR. Journal Of Biomolecular Nmr *42*, 1-9.

Pervushin, K. (2000). Impact of Transverse Relaxation Optimized Spectroscopy (TROSY) on NMR as a technique in structural biology. Quarterly Reviews Of Biophysics *33*, 161-197.

Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997). Attenuated T-2 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 of the United States of America *94*, 12366-12371.

Pierce, M., and Nall, B. (1999). Isothermal titration calorimetry of protein-protein interactions. METHODS-A COMPANION TO METHODS IN ENZYMOLOGY *19*, 213.

Piotto, M., Saudek, V., and Sklenar, 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., Chen, B. L., Phizacke.Rp, and Saul, F. (1973). 3-Dimensional Structure Of 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*, 3305-3310.

Prestegard, J. (2004). Residual dipolar couplings in structure determination of biomolecules. CHEMICAL REVIEWS *104*, 3519-3540.

Priestle, J. P., Schar, H. P., and Grutter, M. G. (1989). Crystallographic Refinement of Interleukin-1-Beta at 2.0 a-Resolution. Proceedings of the National Academy of Sciences of the United States of America *86*, 9667-9671.

Rajesh, S., Nietlispach, D., Nakayama, H., Takio, K., Laue, E. D., Shibata, T., and Ito, Y. (2003). A novel method for the biosynthesis of deuterated proteins with selective protonation at the aromatic rings of Phe, Tyr and Trp. Journal Of Biomolecular Nmr *27*, 81-86.

Ramsland, P. A., and Farrugia, W. (2002). Crystal structures of human antibodies: a detailed and unfinished tapestry of immunoglobulin gene products. Journal of Molecular Recognition *15*, 248-259.

Reichert, J. M., and Valge-Archer, V. E. (2007). Outlook - Development trends for monoclonal antibody cancer therapeutics. Nature Reviews Drug Discovery *6*, 349-356.

Renisio, J. G., Perez, J., Czisch, M., Guenneugues, M., Bornet, O., Frenken, L., Cambillau, C., and Darbon, H. (2002). Solution structure and backbone dynamics of an antigen-free heavy chain variable domain (VHH) from llama. Proteins-Structure Function And Genetics *47*, 546-555.

Reth, M. (1989). Antigen Receptor Tail Clue. Nature *338*, 383-384.

Reth, M. (1992). Antigen Receptors On B Lymphocytes, In Paul, W. E. (Ed.). Annual Review Of Immunology, Vol. 10. X+910p. Annual Reviews Inc.: Palo Alto, California, Usa. Illus, pp. 97-121.

Reth, M. (2001). Oligomeric antigen receptors: a new view on signaling for the selection of lymphocytes. Trends In Immunology *22*, 356-360.

Reth, M. (2002). Hydrogen peroxide as second messenger in lymphocyte activation. Nature Immunology *3*, 1129-1134.

Reth, M., and Wienands, J. (1997). Initiation and processing of signals from the B cell antigen receptor. Annual Review of Immunology *15*, 453-479.

Riek, R., Wider, G., Pervushin, K., and Wuthrich, K. (1999). Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. Proceedings Of The National Academy Of Sciences Of The United States Of America *96*, 4918-4923.

Riley, D. (2005) Development of an Automated Template Selection and Alignment Tool for Protein Structure Homology Modeling., M.Sc., Rochester, NY, USA.

Rivkin, A. (2009). Certolizumab Pegol for the Management of Crohn's Disease in Adults. Clinical Therapeutics *31*, 1158-1176.

Roitt, I. M., and Delves, P. J. (2001). Roitt's essential immunology, 10th edn (Oxford, UK; Malden, MA: Blackwell Science).

Rosen, M. K., Gardner, K. H., Willis, R. C., Parris, W. E., Pawson, T., and Kay, L. E. (1996). Selective methyl group protonation of perdeuterated proteins. Journal of Molecular Biology *263*, 627-636.

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

Russell, R. B., Alber, F., Aloy, P., Davis, F. P., Korkin, D., Pichaud, M., Topf, M., and Sali, A. (2004). A structural perspective on protein-protein interactions. Current Opinion In Structural Biology *14*, 313-324.

Sali, A., and Blundell, T. L. (1993). Comparative Protein Modeling By Satisfaction Of Spatial Restraints. Journal Of Molecular Biology *234*, 779-815.

Sandin, S., Ofverstedt, L. G., Wikstrom, A. C., Wrange, O., and Skoglund, U. (2004). Structure and flexibility of individual immunoglobulin G molecules in solution. Structure *12*, 409-415.

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

Sass, J., Cordier, F., Hoffmann, A., Cousin, A., Omichinski, J. G., Lowen, H., and Grzesiek, S. (1999). Purple membrane induced alignment of biological macromolecules in the magnetic field. Journal of the American Chemical Society *121*, 2047-2055.

Sattler, M., Schleucher, J., and Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Progress In Nuclear Magnetic Resonance Spectroscopy *34*, 93-158.

Saunders, M., Wishnia, A., and Kirkwood, J. G. (1957). THE NUCLEAR MAGNETIC RESONANCE SPECTRUM OF RIBONUCLEASE1. Journal of the American Chemical Society *79*, 3289.

Shi, J. Y., Blundell, T. L., and Mizuguchi, K. (2001). FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. Journal Of Molecular Biology *310*, 243-257.

Shokri, A., Sanden, A. M., and Larsson, G. (2002). Growth rate-dependent changes in Escherichia coli membrane structure and protein leakage. Applied Microbiology And Biotechnology *58*, 386-392.

Sigalov, A. (2005). Multi-chain immune recognition receptors: spatial organization and signal transduction. SEMINARS IN IMMUNOLOGY *17*, 51.

Smith, G. R., and Sternberg, M. J. E. (2002). Prediction of protein-protein interactions by docking methods. Current Opinion In Structural Biology *12*, 28-35.

Sprules, T., Green, N., Featherstone, M., and Gehring, K. (2003). Lock and key binding of the HOX YPWM peptide to the PBX homeodomain. Journal Of Biological Chemistry *278*, 1053-1058.

Tate, S.-i., Shimahara, H., and Utsunomiya-Tate, N. (2004). Molecular-orientation analysis based on alignment-induced TROSY chemical shift changes. Journal of Magnetic Resonance *171*, 284-292.

Thorpe, I. F., and Brooks, C. L. (2007). Molecular evolution of affinity and flexibility in the immune system. Proceedings of the National Academy of Sciences of the United States of America *104*, 8821-8826.

Tjandra, N. (1997). Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. SCIENCE *278*, 1111-1114.

Tjandra, N., Grzesiek, S., and Bax, A. (1996). Magnetic field dependence of nitrogen-proton J splittings in N-15-enriched human ubiquitin resulting from relaxation interference and residual dipolar coupling. Journal of the American Chemical Society *118*, 6264-6272.

Tjandra, N., Omichinski, J. G., Gronenborn, A. M., Clore, G. M., and Bax, A. (1997). Use of dipolar H-1-N-15 and H-1-C-13 couplings in the structure determination of magnetically oriented macromolecules in solution. Nature Structural Biology *4*, 732-738.

Tolar, P., Hanna, J., Krueger, P. D., and 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.

Tolar, P., and Pierce, S. K. (2009). Change we can believe in-of the conformational type Workshop on the Initiation of Antigen Receptor Signaling. Embo Reports *10*, 331-336.

Tolar, P., Sohn, H. W., and Pierce, S. K. (2005). The initiation of antigen-induced B cell antigen receptor signaling viewed in living cells by fluorescence resonance energy transfer. Nature Immunology *6*, 1168-1176.

Tolar, P., Sohn, H. W., and Pierce, S. K. (2008). Viewing the antigen-induced initiation of B-cell activation in living cells. Immunological Reviews *221*, 64-76.

Tolman, J. R. (2001). Dipolar couplings as a probe of molecular dynamics and structure in solution. Current Opinion in Structural Biology *11*, 532-539.

Tolman, J. R., Flanagan, J. M., Kennedy, M. A., and Prestegard, J. H. (1995). Nuclear Magnetic Dipole Interactions In Field-Oriented Proteins - Information For Structure Determination In Solution. Proceedings Of The National Academy Of Sciences Of The United States Of America *92*, 9279-9283.

Tugarinov, V., Kanelis, V., and Kay, L. E. (2006). Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. Nature Protocols *1*, 749-754.

Tugarinov, V., and Kay, L. E. (2003). Quantitative NMR studies of high molecular weight proteins: Application to domain orientation and ligand binding in the 723 residue enzyme malate synthase G. Journal of Molecular Biology *327*, 1121-1133.

Tycko, R., Blanco, F. J., and Ishii, Y. (2000). Alignment of biopolymers in strained gels: A new way to create detectable dipole-dipole couplings in high-resolution biomolecular NMR. Journal of the American Chemical Society *122*, 9340-9341.

Vajdos, F. F., Adams, C. W., Breece, T. N., Presta, L. G., de Vos, A. M., and Sidhu, S. S. (2002). Comprehensive functional maps of the antigen-binding site of an anti-ErbB2 antibody obtained with shotgun scanning mutagenesis. Journal Of Molecular Biology *320*, 415-428.

Valafar, H., and Prestegard, J. H. (2003). Rapid classification of a protein fold family using a statistical analysis of dipolar couplings. Bioinformatics *19*, 1549-1555.

van Dijk, A. D. J., Boelens, R., and Bonvin, A. (2005a). Data-driven docking for the study of biomolecular complexes. Febs Journal *272*, 293-312.

van Dijk, A. D. J., Fushman, D., and Bonvin, A. M. J. J. (2005b). Various strategies of using residual dipolar couplings in NMR-driven protein docking: Application to Lys48-linked Di-ubiquitin and validation against N-15-relaxation data. Proteins-Structure Function and Bioinformatics *60*, 367-381.

Van Emden, H. F. (2008). Statistics for terrified biologists (Malden, MA; Oxford: Blackwell Pub.).

Venters, R. A., Huang, C. C., Farmer, B. T., Trolard, R., Spicer, L. D., and Fierke, C. A. (1995). High-Level H-2/C-13/N-15 Labeling of Proteins for Nmr-Studies. Journal of Biomolecular Nmr *5*, 339-344.

Veverka, V., Crabbe, T., Bird, I., Lennie, G., Muskett, F. W., Taylor, R. J., and Carr, M. D. (2008). Structural characterization of the interaction of mTOR with phosphatidic acid and a novel class of inhibitor: compelling evidence for a central role of the FRB domain in small molecule-mediated regulation of mTOR. Oncogene *27*, 585-595.

Ward, L. (1985). MEASUREMENT OF LIGAND-BINDING TO PROTEINS BY FLUORESCENCE SPECTROSCOPY. METHODS IN ENZYMOLOGY *117*, 400.

Waters, L. C., Veverka, V., Bohm, M., Schmedt, T., Choong, P. T., Muskett, F. W., Klempnauer, K. H., and Carr, M. D. (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.

Wedemayer, G. J., Patten, P. A., Wang, L. H., Schultz, P. G., and Stevens, R. C. (1997). Structural insights into the evolution of an antibody combining site. Science *276*, 1665-1669.

Wilkinson, I. C., Hall, C. J., Veverka, V., Shi, J. Y., Muskett, F. W., Stephens, P. E., Taylor, R. J., Henry, A. J., and Carr, M. D. (2009). A high resolution NMR-based model for the structure of a ScFv-IL-1{beta} Complex: Potential for NMR as a Key tool in therapeutic antibody design and development. J Biol Chem.

Williamson, M. P., Havel, T. F., and Wuthrich, K. (1985). Solution Conformation Of Proteinase Inhibitor-Iia From Bull Seminal Plasma By H-1 Nuclear Magnetic-Resonance And Distance Geometry. Journal Of Molecular Biology *182*, 295-315.

Williamson, R. A., Carr, M. D., Frenkiel, T. A., Feeney, J., and 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.

Winter, G., and Milstein, C. (1991). Man-Made Antibodies. Nature *349*, 293-299.

Wittekind, M., and Mueller, L. (1993). Hncacb, a High-Sensitivity 3D Nmr Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha-Carbon and Beta-Carbon Resonances in Proteins. Journal of Magnetic Resonance Series B *101*, 201-205.

Wodak, S. J., and Janin, J. (1978). Computer-Analysis Of Protein-Protein Interaction. Journal Of Molecular Biology *124*, 323-342.

Wu, A. (2005). Arming antibodies: prospects and challenges for immunoconjugates. NATURE BIOTECHNOLOGY *23*, 1137-1146.

Wuthrich, K. (1990). Protein structure determination in solution by NMR spectroscopy. The Journal of Biological Chemistry *265*, 22059-22062.

Ye, Y., and Godzik, A. (2003). Flexible structure alignment by chaining aligned fragment pairs allowing twists. Bioinformatics *19 Suppl 2*, ii246-255.

Zheng, Y., Shopes, B., Holowka, D., and Baird, B. (1992). Dynamic Conformations Compared For Ige And Igg1 In Solution And Bound To Receptors. Biochemistry *31*, 7446-7456.

Zuiderweg, E. R. P. (2002). Mapping protein-protein interactions in solution by NMR Spectroscopy. Biochemistry *41*, 1-7.

Zweckstetter, M., and Bax, A. (2000). Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. Journal of the American Chemical Society *122*, 3791-3792.

Zweckstetter, M., and Bax, A. (2001). Characterization of molecular alignment in aqueous suspensions of Pf1 bacteriophage, In Journal of Biomolecular Nmr, pp. 365-377.

Zweckstetter, M., and Bax, A. (2002). Evaluation of uncertainty in alignment tensors obtained from dipolar couplings. Journal of Biomolecular Nmr *23*, 127-137.