Structural Features of the Extracellular Region of Programmed Cell Death Ligand-1 and the Implications for Therapeutic Targeting

Thesis submitted for the degree of

Doctor of Philosophy

University of Leicester

Kayleigh Walker BSc (Hons), University of Manchester

Department of Molecular and Cell Biology

University of Leicester

June 2019

Acknowledgements

Structural Features of the Extracellular Region of PD-L1 and the Implications for Therapeutic Targeting

Abstract

The use of antibody immune checkpoint inhibitors of the PD-1/PD-L1 interaction has recently emerged as a great success in the treatment of cancer. The development of small-molecule inhibitors targeting the PD-1/PD-L1 interaction lags behind. There are however, a series of bi-phenyl derived compounds patented by Bristol-Myers Squibb, that have been shown to block the PD-1/PD-L1 interaction *in vitro* but have no cellular activity.

The work reported here uses NMR chemical shift perturbation analysis to characterise the binding of one of these small-molecules to PD-L1. Currently the only published structural information regarding these molecules binding to PD-L1 focuses on only the IgV-like domain of PD-L1. The work reported here shows that the BMS compound appears to bind to the full extracellular region of PD-L1 with a much-reduced affinity than what has been reported for the single IgV-like domain, shedding light on why these compounds have not been shown to have cellular activity and suggesting that development of small-molecules targeting PD-L1 should be performed with the full extracellular region of PD-L1.

Backbone resonance assignment of the full extracellular region of PD-L1 and the two Ig-like domains in isolation has allowed the identification of an interface between the domains of PD-L1. Chemical shift perturbation analysis suggests that this interface may be augmented when PD-L1 binds to PD-1. Since PD-L1 has been shown to bind to binding partners on the same membrane, as well as across cell membranes, we propose that an interface between the domains of the extracellular region of PD-L1 is a mechanism through which PD-L1 can switch the orientation of the IgV-like domain depending on whether it is binding to a partner on the same membrane or across membranes. This could have important implications for the design of therapeutics that seek to modulate the binding of PD-L1 to its binding partners.

Kayleigh Walker

Contents

Abbreviations

aa	Amino acid
Bcl-XL	B-cell lymphoma-extra large
BMS	Bristol-Myers Squibb
BPI	Bi-phenyl inhibitor
CAR	Chimeric Antigen receptor
CD	Circular Dichroism
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
DC	Dendritic cell
DTT	1,4 Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
EMA	European Medicines Agency
Fab	Antigen-binding fragment
FDA	Federal Drug Administration
HPV	Human Papilloma Virus
HSQC	Heteronuclear single quantum coherence
HTRF	Homogenous Time-Resolved Fluorescence
Hz	Hertz
IFN-γ	Interferon gamma
Ig	Immunoglobin
IgC	Immunoglobin constant domain
IgV	Immunoglobin variable domain

IL	Interleukin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
ITIM	Immunoreceptor Tyrosine-based Inhibitory Motif
ITSM	Immunoreceptor Tyrosine-based Switch Motif
LB	Luria-Bertani
LILRB	Leukocyte immunoglobulin-like receptor
MCC	Merkel Cell Carcinoma
MHC	Major Histocompatibility Complex
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect spectroscopy
NSCLC	Non-small cell lung cancer
OD_{600}	Optical Density at 600nm
PBS	Phosphate-buffered saline
PD-1	Programmed cell death protein 1
PD-L1	Programmed cell death ligand 1
PEG	Polyethylene glycol
PI3K	Phosphoinositide 3-kinase
RI	Refractive Index
RT-PCR	Reverse transcription polymerase chain reaction
SDS-PAGE	Sodium dodecyl sulphate–polyacrylamide gel electrophoresis
SEC-MALS	Size Exclusion Chromatography-Multi-angle Light Scattering

SHP	Src Homology region 2 domain-containing Phosphatase
TCR	T-cell receptor
TROSY	Transverse relaxation optimised spectroscopy
UV	Ultraviolet
WCL	Whole cell lysate
ZAP70	Zeta-chain-associated protein kinase 70

Chapter 1

Introduction

1.1 Cancer Immunotherapy

Cancer immunotherapy can be defined as therapeutic treatments that use and enhance the immune system to recognise and target cancer cells. Due to the innate ability of the immune system to provide immunosurveillance throughout the entire body, this exciting therapy area has the potential to treat all cancer types. In 1891, Dr William B. Coley first attempted immune-modulating cancer treatment by injecting a neck cancer patient with live streptococcal cultures. Encouraged by the success of the treatment of this initial patient, Coley expanded the number of patients he treated and his research culminated in the creation of "Coley's toxin", which was a combined injection of heat-killed *Streptococcus* and *Serratia marcescens* (Coley 1898). A lack of knowledge of the mechanism of action of "Coley's toxin" at the time meant surgery and radiotherapy remained the standard treatments of the day. It wasn't until 1957 that the theory of cancer immunosurveillance was proposed by Burnet, who suggested that lymphocytes could recognise somatic cells transformed by spontaneous mutations (Burnet 1957). However, a lack of data supporting this theory and the inability to culture lymphocytes *in vitro* at the time hindered progress in this area. Finally, in 1976, the T-cell growth factor cytokine interleukin-2 (IL-2) was identified and T-cell cultures could be grown *in vitro* for the first time (Morgan *et al.* 1976). In 1985 it was shown that administration of recombinant IL-2 to patients with metastatic cancers caused significant tumour regression (Rosenberg *et al.* 1985), and IL-2 was approved by the FDA for treatment of metastatic kidney cancer in 1991, and metastatic melanoma in 1998.

Since then cancer immunotherapy treatment has grown to include other cytokines, vaccines, adoptive T-cell transfer and monoclonal antibodies (Farkona *et al.* 2016). Cancer vaccines involve introducing antigens to the immune system such that immune cells can recognise antigens expressed

on cancer cells and eliminate them. The most widely used cancer vaccine today is the human papilloma virus (HPV) vaccine which is a preventative vaccine that protects against the HPV types that cause the greatest risk of cervical cancer (Kash *et al.* 2015). Adoptive T-cell transfer, or chimeric antigen receptor (CAR) T-cell therapy, is an emerging area of personalised cancer immunotherapy that involves the genetic engineering of patient-derived T-cells such that they express a CAR that recognises a specific tumour antigen on cancer cells and these T-cells are then transferred back to the patient. Clinical trials of an anti-CD19 CAR T-cell therapy showed very promising results with 90% remission rates for patients with relapsed/refractory acute lymphoblastic leukaemia (ALL) (Singh *et al.* 2016). Following positive clinical trials, in 2017 CAR T-cell therapies were approved by the FDA to treat children with ALL, and adults with advanced lymphomas.

A current trend in cancer immunotherapy is the use of monoclonal antibody checkpoint inhibitors targeting co-inhibitory molecules on the surface of T-cells, and their binding partners (Ayyar *et al.* 2016). The immune checkpoint refers to a series of activating co-stimulatory, and co-inhibitory molecules important for the commencement of productive immune responses to pathogens, as well as the termination of immune responses to prevent tissue damage, and maintenance of peripheral tolerance (Bour-Jordan *et al.* 2011). For a T-cell to become activated two signals are required. Firstly the signal produced when the T-cell receptor (TCR) binds to antigen-MHC complexes on the surface of antigen-presenting cells. The second signal is an activating signal provided by the interaction of co-stimulatory interactions such as CD28/B7-1/2 (Bretscher 1999). There are, however, co-inhibitory interactions such as CTLA-4/B7-1/2 and PD-1/PD-L1. The interaction of these molecules causes intracellular signalling in the T-cell that serves to dampen a T-cell response and plays an important role in the maintenance of immune tolerance. It is well established that over-expression of both CTLA-4 (Contardi *et al.* 2005) and PD-L1 is a mechanism through which tumours escape an immune response (Nakanishi *et al.* 2007; Thompson *et al.* 2004). These immune checkpoint interactions are a well-characterised therapeutic target with the first monoclonal antibody checkpoint inhibitor, Ipilimumab

targeting CTLA-4, FDA-approved for the treatment of metastatic melanoma in 2011 (Fellner 2012). Since then five monoclonal antibodies targeting either PD-1 or PD-L1 have been approved by the FDA for the treatment of a wide variety of different cancers with the list growing rapidly as results from clinical trials are published (Gong J *et al.* 2018).

1.2 The PD-1/PD-L1 interaction

1.2.1 PD-1 and PD-L1 expression

Programmed cell death-1 (PD-1) is a 288 amino acid (aa) transmembrane protein consisting of a 20 aa leader sequence, an extracellular immunoglobin (Ig) family domain, a single pass transmembrane domain and a 95 aa intracellular domain containing an immunoreceptor tyrosine-based inhibitory motif (ITIM) and an immunoreceptor tyrosine-based switch motif (ITSM). PD-1 is expressed at low basal levels on naïve T-cells and this is associated with immune tolerance (Agata *et al.* 1996). PD-1 is transiently expressed on T-cells, B-cells, natural killer T-cells, activated monocytes and dendritic cells (DCs) following an immune stimulus before being down regulated in the case of acute infection (Bally *et al.* 2016). Two extracellular ligands for PD-1 have been identified, PD-L1 and PD-L2 (Dong *et al.* 1999; Latchman *et al.* 2001). PD-L1 and PD-L2 differ in their expression patterns with PD-L1 being more widely expressed than PD-L2. PD-L1 has been shown to be constitutively expressed on mouse T-cells, B-cells, DCs, macrophages, and bone marrow-derived mast cells (Yamazaki *et al*. 2002). PD-L1 expression has also been identified on a wide range of non-hematopoietic cells in heart, skeletal muscle, placenta, and lung tissue (Dong *et al.* 1999). PD-L2 expression has been found to be more restricted, being expressed on activated macrophages and DCs (Yamazaki *et al*. 2002). Since it is PD-L1 over-expression that is associated with cancer, this work focuses on the interaction of PD-1 with PD-L1 (Pardoll DM, 2012).

1.2.2 PD-1/PD-L1 signalling

As discussed above, the PD-1/PD-L1 interaction serves to inhibit T-cell activation. PD-1 and PD-L1 interact at the immunological synapse via their immunoglobin-like domains and the resulting signalling via PD-1 counteracts T-cell receptor signalling and co-stimulatory signalling via CD28 (Figure 1.1). The cytoplasmic domain of PD-1 contains both an ITIM and ITSM motif. PD-1 has been shown to recruit both Src homology region 2 domain-containing phosphatase-1 (SHP-1) and SHP-2 to its ITIM and ITSM motifs (Chemnitz *et al.* 2004) leading to the inhibition of ZAP70 kinase and PI3K-mediated phosphorylation, which is involved in downstream T-cell receptor signalling (Sheppard *et al.* 2004) (Figure 1.2). RT-PCR measurements of human T-cells have shown that PD-1 engagement results in the inhibition of CD28-mediated up-regulation of cytokines involved in T-cell activation and effector function, such as IL-2, IL-10 and IFN-γ, and decreases transcription of the pro-survival factor Bcl-X_L (Chemnitz *et al.* 2004).

(Antigen-presenting) cell

MHC I

B7

Co-stimulation

CD28

PD-L1

Co-inhibition

PD-1

TCR

Cytotoxic T-cell

T-cell effector function

Inhibition of T-cell signaling
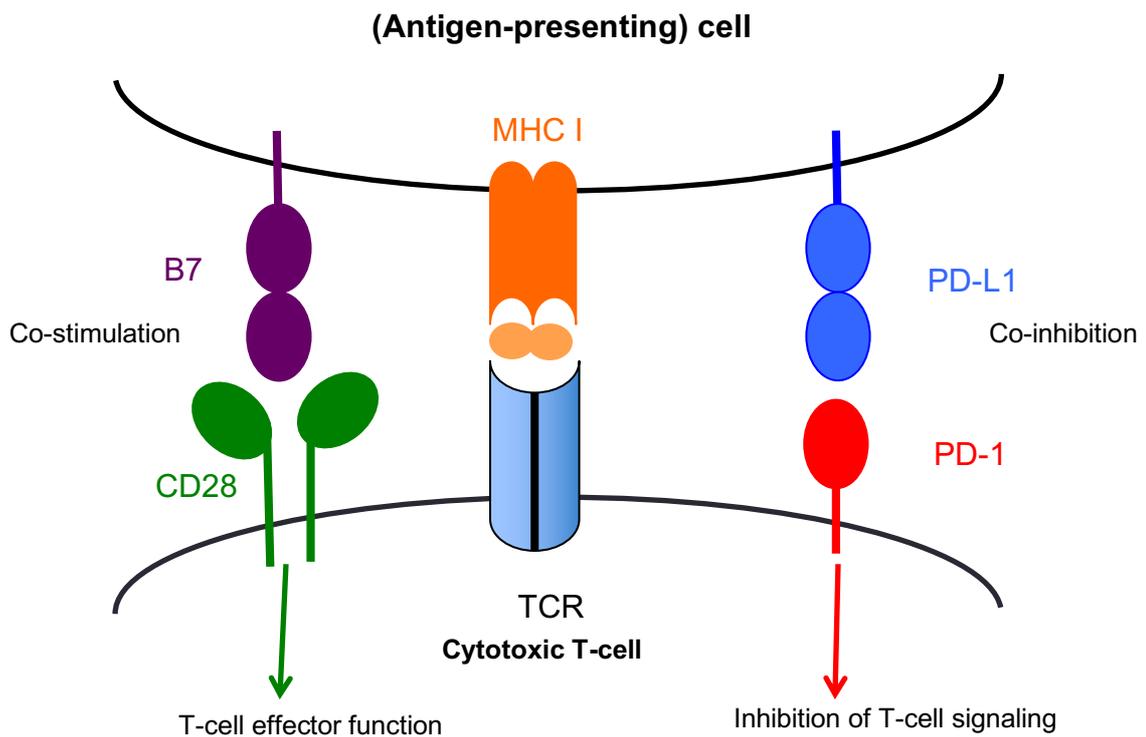
Figure 1.1 Two-signal T-cell activation.

A figure representing the immunological synapse between a cytotoxic T-cell and an MHC I-antigen presenting cell. T-cell activation and effector function requires co-stimulatory signals as well as T-cell receptor (TCR) signalling. The PD-1/PD-L1 interaction is inhibitory of TCR signalling. The role of this interaction is to promote immune tolerance and protect from tissue damage caused by chronic inflammation.

There are also questions remaining as to the roles of SHP-1 and SHP-2 in PD-1 signalling since it has been shown that, in T-cells, SHP-1 can bind to both the ITIM and ITSM motifs of PD-1, whereas SHP-2 binds only to the ITSM (Chemnitz *et al*. 2004). However, PD-1 recruitment of SHP-1 in B-cells (Okazaki *et al*. 2001) and Jurkat T-cells (Latchman *et al*. 2001) has not been reported, raising the question of the importance of SHP-1 to PD-1 signalling. Also, point mutation of the ITSM tyrosine disrupts PD-1 signalling whereas mutation of the ITIM has no effect on PD-1 signalling (Chemnitz *et al.* 2004) suggesting that SHP-2 recruitment to the ITSM is the main driver of PD-1 signalling.
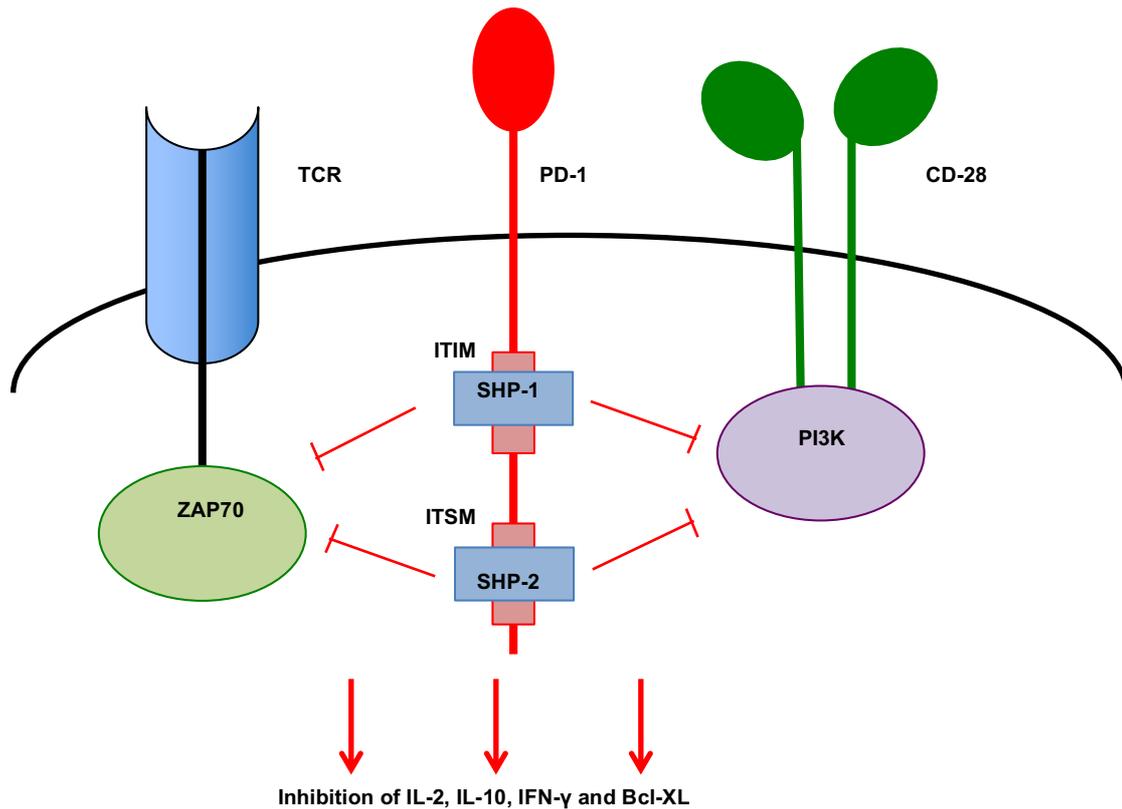
Figure 1.2. Modulation of T-cell signalling by PD-1.

A figure representing the recruitment of SHP phosphatases to the cytoplasmic ITIM and ITSM motifs of PD-1. This inhibits TCR and co-stimulatory signalling by either inhibition of Zap70 and PI3K phosphorylation or direct dephosphorylation, or both. This leads to abrogation of ZAP70 and PI3K-mediated downstream signalling events.

The specific mode of action of how ligation of PD-1 by PD-L1 at the cell membrane leads to the transduction of intracellular signalling remains unknown. PD-1 has been shown to be monomeric at the cell surface (Zhang *et al.* 2004), in contrast to CD28 and CTLA-4, which are disulphide-linked homodimers, and B7-1, which is predominantly a non-covalently-linked dimer at the cell surface (Bhatia *et al.* 2005). The oligomeric state of PD-L1 on the cell membrane has not been determined, however cross-linking experiments have previously shown that PD-L1 has a tendency to form a dimer in solution (Chen *et al.* 2008). It could be possible that a PD-L1 dimer binds to PD-1 and the resulting proximity of the PD-1 cytoplasmic domains initiates intracellular signalling. The stoichiometry of the PD-L1/PD-1 complex in solution will be explored in this thesis.

1.2.3 Structure of PD-1 and PD-L1

There are many crystal structures of the extracellular regions of PD-1, PD-L1 and the complex of the two, as well as a solution NMR structure of PD-1 (Figure 1.3). Structures of PD-1 confirm that the extracellular region of PD-1 is an Ig-like domain with V-type topology, made up of two anti-parallel β-sheets (Zhang *et al.* 2004; Cheng *et al.* 2013). There are crystal structures of both the unbound full extracellular region of PD-L1 as well as the N-terminal domain alone. These structures confirm that the domains of PD-L1 are Ig-like domains with two anti- parallel β-sheets. The C-terminal membrane-proximal Ig-like domain of PD-L1 has C-type topology and the N-terminal domain has V-type topology (Chen *et al.* 2010; Zak *et al*. 2015).
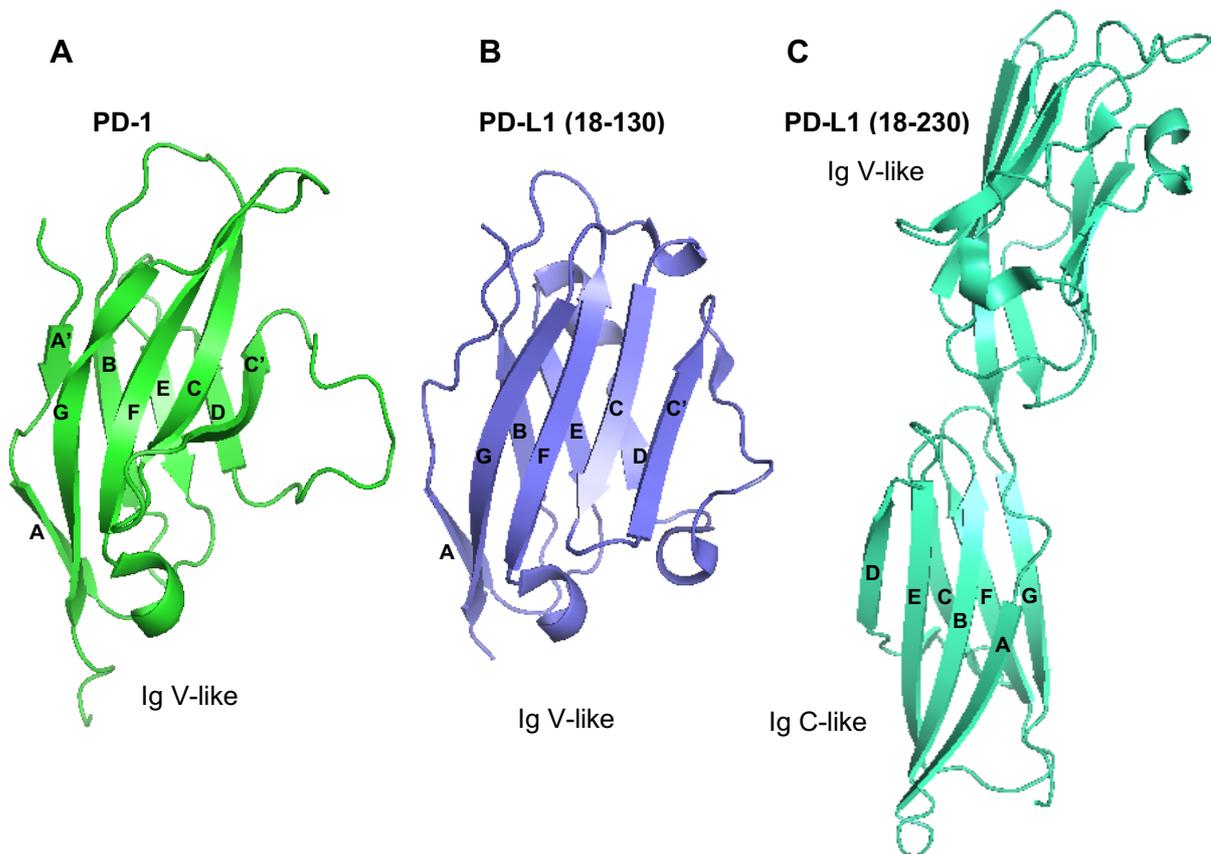
Figure 1.3. Structures of PD-1 and PD-L1.

A) Ribbon representation of the extracellular domain of human PD-1. The crystal structure shows Immunoglobulin V-like topology with the β-strands labelled as such. (PDB: 5GGR) **B)** Ribbon representation of the N-terminal domain of the extracellular region of human PD-L1 (18-132). β-strands are labelled with the Ig V-like topology that the crystal shows. (PDB: 5C3T) **C)** Ribbon representation of the crystal structure of the full extracellular region of human PD-L1 (18-230). The membrane proximal domain of PD-L1 shows Ig C-like topology and the β-strands are labelled as such. (PDB: 3FN3)

At the time of writing there are no structures of the full extracellular domains of wild-type human PD-1 and PD-L1 in complex. Murine PD-1 and human PD-1 share 64% sequence identity and the affinity of murine PD-1 for human PD-L1 has been shown to be similar to that of murine PD-1 for murine PD-L1 (Zhang *et al.* 2004). The structure of murine PD-1 in complex with the full extracellular region of PD-L1 was solved by x-ray crystallography in 2008. This crystal structure showed that PD-1 and PD-L1 interact via their GFCC' β-sheets, and only the N-terminal IgV-like domain of PD-L1 is involved in the binding of PD-1 (Lin *et al.* 2008). There was some ambiguity as to the stoichiometry of the interaction since these crystals contained two PD-1 molecules and one PD-L1 molecule in asymmetric unit. The presence of the second PD-1 molecule was suggested to be a crystallisation artefact (Lin *et al*. 2008). The structure of the human PD-1 and the single PD-1 interacting N-terminal domain of human PD-L1, in complex, was solved using x-ray crystallography and this confirmed that murine PD-1 and human PD-1 bound to human PD-L1 via the same GFCC' β-sheet. Table 1.1 summarises the macromolecular structures of PD-1 and PD-L1 alone and in complex, at the time of writing. The binding interface is made up of a mixture of polar and non-polar contacts with a hydrophobic core surrounded by polar interactions that bury a total surface area of 1970$Å^2$ (Figure 1.4). A clear 1:1 stoichiometry was present in these crystals (Zak et al. 2015). The affinity of human PD-1 for human PD-L1 has been determined by isothermal titration calorimetry (ITC) to be 8 µM (Cheng *et al.* 2013).

Table 1.1 Summary of PD-1/PD-L1 structures.

Protein	PDB Code	Method	Year
Murine PD-1 (34-150; C60S)	1NPU	X-ray Crystallography	2004
Human PD-L1 (18-239) Murine PD-1 (25-157; C60S)	3BIK	X-ray Crystallography	2008
Murine PD-1 (34-150; C60S) Murine PD-L2 (20-220)	3BP5	X-ray Crystallography	2008
Human PD-1 (33-150; C93S)	3RRQ	X-ray Crystallography	2011
Human PD-1 (34-150; C93S)	2M2D	NMR	2013
Human PD-1 (33-150; C93S) Human PD-L1 (18-134)	4ZQK	X-ray Crystallography	2015
Human PD-L1 (18-134)	5C3T	X-ray Crystallography	2015
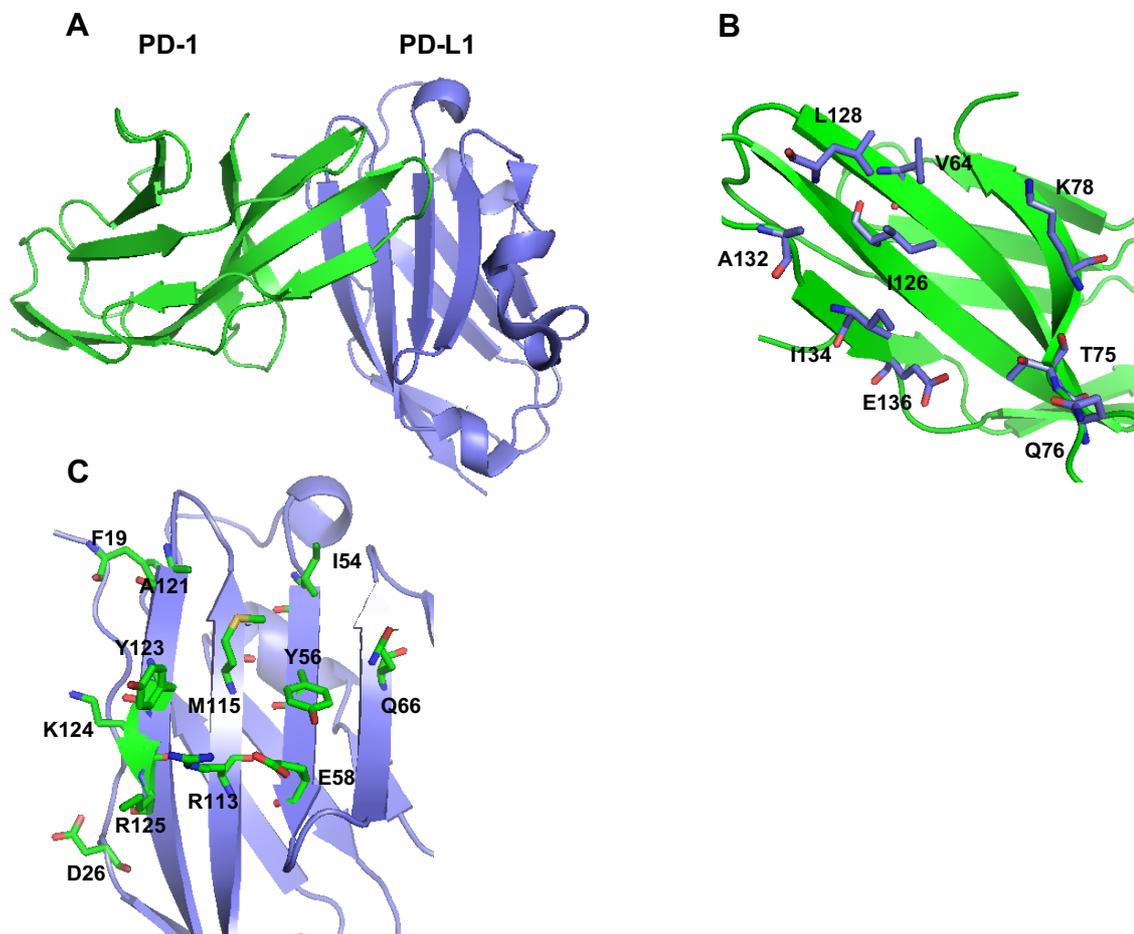
Figure 1.4. The human PD-1/PD-L1 interaction interface.
A) Ribbon representation of the PD-1 (green) PD-L1 (18-130) (purple) complex. PDB: 4ZQK. **B)** Ribbon representation of PD-1 with residues involved in the binding of PD-L1 shown in purple. **C)** Ribbon representation of PD-L1 with residues involved in the binding of PD-1 shown in green.

1.3 Therapeutic targeting of the PD-1/PD-L1 interaction

1.3.1 Antibodies targeting PD-1

There are currently two licensed therapeutics targeting PD-1 in the treatment of cancer, Merck's pembrolizumab and Bristol-Myers Squibb's nivolumab. These were first FDA-approved in 2014 for the treatment of metastatic melanoma in patients where prior treatment has failed. In 2015 FDA-approval for both pembrolizumab and nivolumab was extended to the second-line treatment of non-small cell lung cancer (NSCLC). Since then these drugs have been approved by the FDA and European Medicines Agency (EMA) for the treatment of kidney cancer, Hodgkin lymphoma, head and neck cancer, bladder cancer, colorectal cancer, liver cancer, stomach and gastroesophageal cancer, and cervical cancer. The structures of PD-1 in complex with Fab-fragments of both pembrolizumab and nivolumab have been solved by x-ray crystallography. These structures show that pembrolizumab and nivolumab block the binding of PD-L1 to PD-1 by directly competing for the PD-L1 binding site of PD-1 (Lee *et al.* 2016). Figure 1.5 shows the binding interfaces of the antibodies. It is clear that these antibodies bind to different, slightly overlapping epitopes on PD-1, directly competing for the PD-L1 binding site. The total buried surface areas of the nivolumab/PD-1 and pembrolizumab/PD-1 complexes are 1487$Å^2$ and 2126$Å^2$ respectively (Lee *et al.* 2016). Despite the significant difference in the size of the interaction sites between the two antibodies they both bind with similar affinities in the low picomolar range and showed similar response rates in preclinical assessments (Fessas *et al.* 2017).
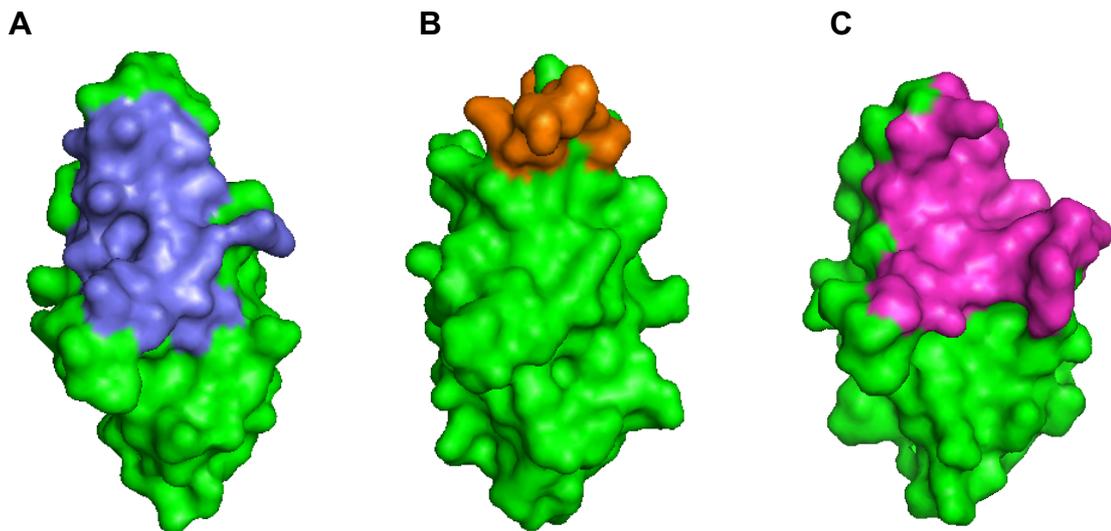
Figure 1.5. Antibodies binding to PD-1.
Surface representations of PD-1 from crystal structures in complex with **A** PD-L1 (4ZQK), **B** nivolumab (5GGR), **C** pembrolizumab (5GGS). The binding interfaces are shown in purple, orange and pink respectively. PD-1 is shown in the same orientation in each figure.

1.3.2 Antibodies targeting PD-L1

There are currently three licensed therapeutics targeting PD-L1 in the treatment of cancer. Atezolizumab, developed by Roche Genentech, is FDA and EMA-approved to treat urothelial cancer and chemotherapy-resistant NSCLC. Avelumab, developed by Merck and Pfizer, is FDA-approved for the second-line treatment of bladder cancer and is approved by the FDA as a first-line treatment for metastatic Merkel Cell Carcinoma (MCC). This is the first treatment of any kind to be FDA-approved for this rare type of skin cancer. Due to how rare MCC is, and the aggressive nature of the cancer, Avelumab has conditional approval from the EMA for the treatment of metastatic MCC so patients can be treated before full regulatory approval. Durvalumab, developed by AstraZeneca, is approved by the FDA for the treatment of advanced bladder cancer and NSCLC. Crystal structures of PD-L1 in complex with all three licensed antibodies have been solved and all three work by blocking the PD-1 binding site (Lee *et al.* 2017; Liu *et al.* 2017). The antibodies bind to different epitopes of the PD-L1 CC'FG β-sheet but have a shared binding region made up of the PD-L1 residues Y56, E58, R113, M115, Y123 (Figure 1.6).
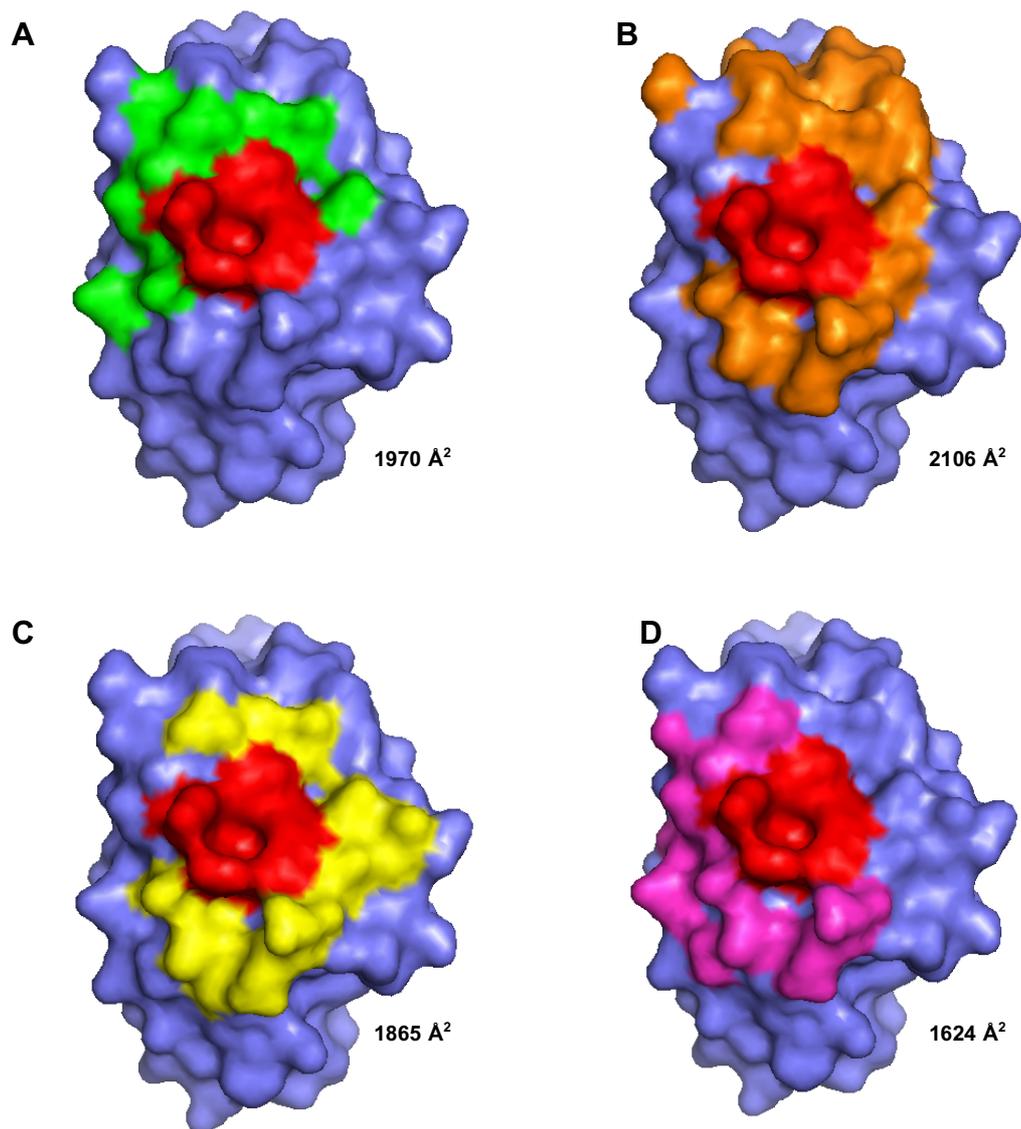
Figure 1.6. Antibodies binding to PD-L1.

Surface representation of PD-L1 with the binding interfaces of **A** PD-1, **B** atezolizumab, **C** avelumab, **D** durvalumab shown in green, orange, yellow and pink respectively. The binding region shared by PD-1 and all three antibodies, made up of PD-L1 residues Y56, E58, R113, M115, Y123, is shown in red. Labels indicate the total buried surface area of the antibody/PD-L1 interactions.

With the development of small-molecules targeting the PD-1/PD-L1 interaction lagging behind that of monoclonal antibodies, there are currently no licensed small-molecules targeting this interaction. There are however small-molecule peptides, cyclic peptides and traditional chemically derived compounds that have been shown to inhibit the PD-1/PD-L1 interaction *in vitro* (Zarganes-Tzitzikas *et al.* 2016).

1.4 Small-molecule targeting of protein-protein interactions

The number of protein-protein interactions (PPIs), the interactome, in humans has been estimated to include up to ~650,000 interactions (Stumpf *et al.* 2008). This represents a large number of potential therapeutic targets. However, PPIs often involve large flat surfaces, without the deep binding pockets found in traditional drug targets of high-affinity small-molecules. PPIs can be classified into 6 main types of interaction (Scott *et al.* 2016). These are summarised in Table 1.2. Critical to the understanding of PPIs was the discovery that interactions driving the affinity of two proteins for one another are not evenly distributed across the interaction surface. In fact, there are so-called "hotspots" that contribute disproportionately more to the binding energy (Clackson and Wells 1995). Alanine-scanning hotspot analysis has shown that Tryptophan, Tyrosine and Arginine, and to a lesser extent Aspartate and Histidine residues, are enriched at PPI hotspots (Bogan and Thorn 1998). This is consistent with PPIs often being predominantly hydrophobic with a surrounding ring of polar residues (Miller 1989).

The presence of hotspots on PPIs means that they can be targeted by small-molecules. However, due to these hotspots being smaller in size than the deep pockets typically targeted by small-molecules, small-molecules targeting PPIs may need to bind multiple hotspots to achieve a similar affinity (Buchwald 2008). For this reason, small-molecule inhibitors of PPIs tend to have larger molecular sizes than traditional small-molecule drugs (Neugebauer *et al.* 2007).

Table 1.2 Classification of protein-protein interfaces (Adapted from Scott *et al.* 2016)

PPI	Description	Example Structure
Globular protein-helical peptide	Helix with a discontinuous epitope binding into a groove	PDB: 2XA0 BCL-X$_L$ (grey) bound to BAX (green)
Globular protein-peptide	Continuous epitope on β-sheet or β-strand binding onto surface with pockets	PDB:1G73 XIAP (grey) bound to SMAC (green)
Globular protein-peptide	Continuous epitope binding into pocket in a β-propeller	PDB: 2DYH KEAP1 (grey) bound to NRF2 peptide (green)
Globular protein-peptide, anchor residue	Peptide with an anchor residue with post-translational modification binding into a pocket	PDB:3UVW Bromodomain of BRD4 (grey) bound to diacetylated histone 4 peptide (green)
Globular protein-globular protein	Two proteins binding via discontinuous epitopes	PDB: 1Z92 IL-2R (grey) bound to IL-2 (green)
Peptide-peptide	Pair of helices with an elongated binding interaction	PDB:1NKP MYC (grey) bound to MAX (green)
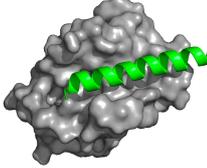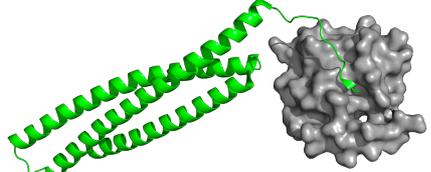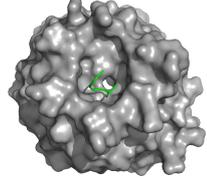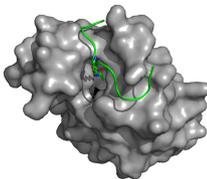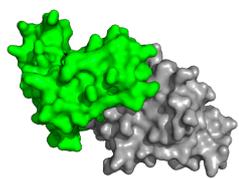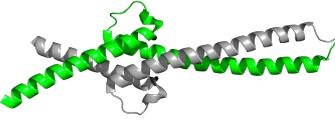
The PD-1/PD-L1 interaction is that of a globular protein-globular protein PPI with both proteins binding via β-sheets with discontinuous epitopes (Figure 1.4). The crystal structure of human PD-1 (33-150; C93S) and human PD-L1 (18-134) shows that the PPI consists of a hydrophobic core surrounded by a network of polar residues providing hydrogen bonds between PD-1 and PD-L1 (Zak *et al.* 2015). Based on the crystal structure Zak *et al*. identified 3 potential hotspots for small molecule targeting of PD-L1 (Figure 1.7). Firstly, a hydrophobic pocket comprised of PD-L1Tyr56, PD-L1Glu58, PD-L1Arg113, PD-L1Met115, and PD-L1Tyr123, that accommodates the side-chain of PD-1Ile134 in the PD-1/PD-L1 complex but could accommodate a six-membered ring of a small-molecule drug. Secondly, a small hydrophobic groove formed by PD-L1Met115, PD-L1Tyr123 and PD-L1Ala121 that accommodates PD-1Ile126 in the complex but could be filled with a branched aliphatic component of a small-molecule drug with the carbonyl oxygen of PD-L1Ala121 available to accept a hydrogen bond. Thirdly, a groove made up of residues from PD-L1Asp122 to PD-L1Arg125, and PD-L1Asp26 containing multiple hydrogen bond donors and acceptors. This groove however is shallow and therefore mat not be suitable for targeting with a small-molecule.
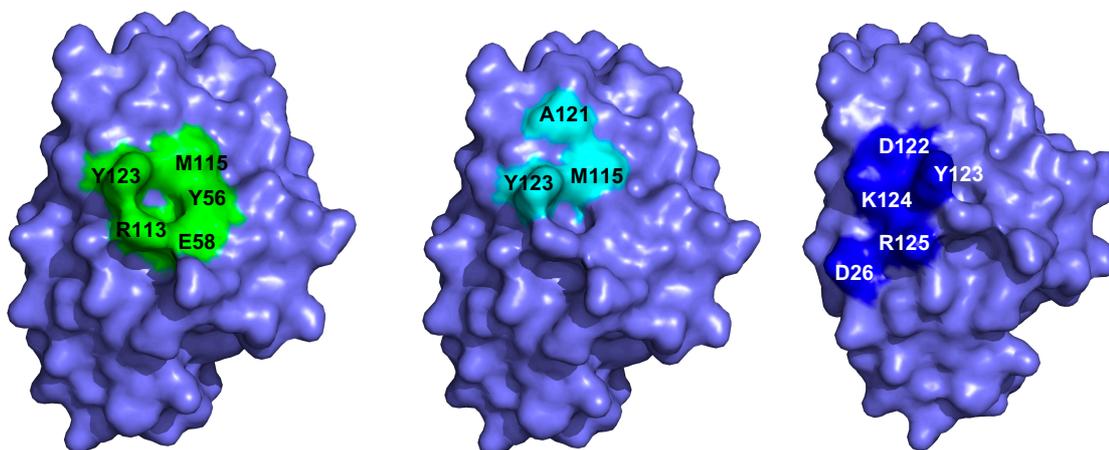
Figure 1.7. Hotspots on PD-L1.
Surface representations of PD-L1 (PDB:4ZQK) showing the first (green), second (cyan), and third (blue) hotspots identified in Zak *et al*. 2015.

1.5 Aims and objectives

The work described in this thesis aims to characterise the binding site of a PD-1/PD-L1 inhibitor, synthesised by collaborators from The University of Victoria (Canada), from a series of compounds patented by Bristol-Myers Squibb in 2015 (WO2015/034820), using NMR spectroscopy methods. The BMS compounds are (2-methyl-3-biphenyl) methanol derivatives and share the same core and the compound used in this work is shown below (Figure 1.8). From this point the compound will be referred to as bi-phenyl inhibitor (BPI).
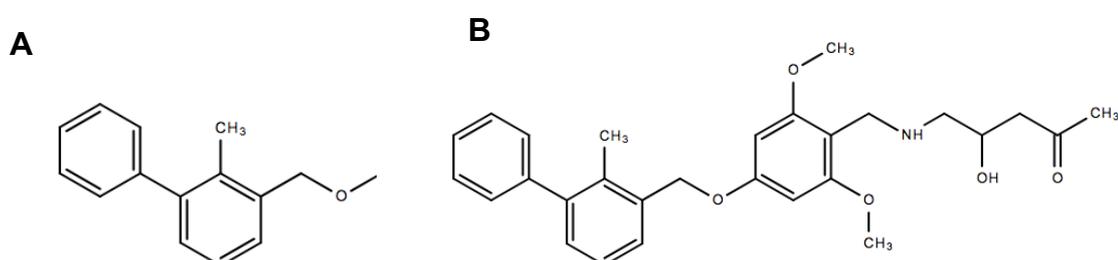
Figure 1.8. Chemical structure of bi-phenyl inhibitors of PD-1/PD-L1.

A) Core structure shared by compounds in Bristol-Myers Squibb patent WO2015/034820.

B) Structure of BPI used in this work.

At the time of writing no BPIs have been shown to have cellular activity but have been shown to block the interaction of PD-1 and PD-L1 *in vitro* (Zak *et al.* 2016). It is hoped that by characterising the binding of these BPIs to PD-L1 that it will guide the development of small-molecules that block PD-L1 binding to PD-1 and also show cellular activity.

This work also aims to better understand the extracellular domains of PD-L1. Specifically, investigating if there is an interaction between the two extracellular domains of PD-L1 and the orientation of the domains in relation to each other, with the objective of identifying sites on PD-L1 that could be targeted to achieve a therapeutic outcome.

Chapter 2

The production and characterisation of PD-L1 and PD-1 proteins

2.1 Introduction

This chapter describes the expression, purification and characterisation of PD-L1 and PD-1 proteins. In all cases an *E. coli* bacterial expression system was used to most easily allow isotopic labelling for NMR studies. Three human PD-L1 constructs were produced, pertaining to different domains of the extracellular region of hPD-L1. These are the full extracellular region (A18-T239, PD-L1_D1D2), the membrane proximal domain (N130-T239, PD-L1_D2), and the PD-1-interacting domain (A18-Y134, PD-L1_D1) (Figure 2.1). In addition to this an expression construct containing the extracellular domain of hPD-1 (P33-E150) was produced.
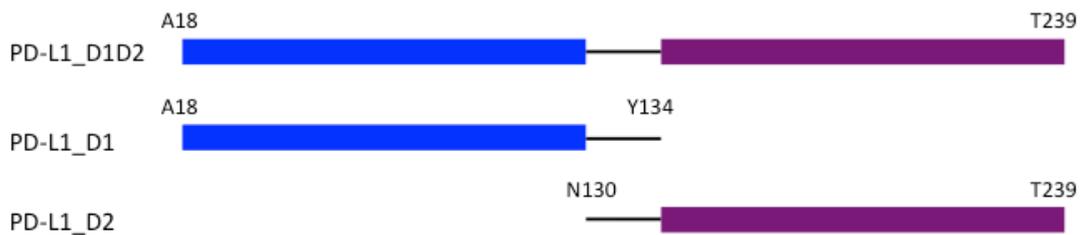
PD-L1_D1D2:

```
    20        30        40        50        60        70        80
AFTVTVPKDLYVVEYGSNMTIECKFPVEKQLDLAALIVYWEMEDKNIIQFVHGEEDLKVQHSSYRQ

      90       100       110       120       130       140       150
RARLLKDQLSLGNAALQITDVKLQDAGVYRCMISYGGADYKRITVKVNAPYNKINQRILVVDPVTSE

          160       170       180       190       200       210
HELTCQAEGYPKAEVIWTSSDHQVLSGKTTTTNSKREEKLFNVTSTLRINTTTNEIFYCTFRRLD

   220       230       239
PEENHTAELVIPELPLAHPPNERT
```

PD-L1_D1:

```
    20        30        40        50        60        70        80
AFTVTVPKDLYVVEYGSNMTIECKFPVEKQLDLAALIVYWEMEDKNIIQFVHGEEDLKVQHSSYRQ

      90       100       110       120       130
RARLLKDQLSLGNAALQITDVKLQDAGVYRCMISYGGADYKRITVKVNAPY
```

PD-L1_D2:

```
          140        150        160        170        180        190
NAPYNKINQRILVVDPVTSEHELTCQAEGYPKAEVIWTSSDHQVLSGKTTTTNSKREEKLFNVT

          200        210        220        230        239
STLRINTTTNEIFYCTFRRLDPEENHTAELVIPELPLAHPPNERT
```

PD-1:

```
          40         50         60         70         80         90
PPTFSPALLVVTEGDNATFTCSFSNTSESFVLNWYRMSPSNQTDKLAAFPEDRSQPGQDCRF

          100        110        120        130        140        150
RVTQLPNGRDFHMSVVRARRNDSGTYLCGAISLAPKAQIKESLRAELRVTERRAE
```

Figure 2.1. PD-L1 and PD-1 constructs.

This shows the domain boundaries of PD-L1 and sequences for al constructs with cysteines involved in disulphide bonds highlighted.

2.1.1 Protein Refolding

All four of the expressed proteins contain disulphide bonds and since *E. coli* do not contain a specialised intracellular oxidising compartment, like the endoplasmic reticulum in eukaryotic cells, but can form disulphide bonds in the periplasm, these proteins were expressed as insoluble inclusion bodies. The proteins therefore needed to be resolubilised and refolded *in vitro*. The production of native proteins from inclusion bodies involves 4 major steps: isolation and washing of the inclusion bodies; solubilisation of the inclusion bodies; protein refolding, and finally, protein purification. Inclusion bodies are easily isolated from the cell lysate by centrifugation and washing with buffers containing detergents can produce inclusion bodies with very high purity of the desired protein (Khan *et al.* 1998). Solubilisation of inclusion bodies is generally achieved by using high concentrations (5-8 M) of chaotropic agents such as urea and guanidine hydrochloride. Reducing agents are used when resolubilising cysteine-containing proteins to ensure the cysteines are in a reduced state so that the correct disulphide bonds can be formed (Fischer *et al.* 1993). Refolding of denatured proteins is contingent on the removal of the chaotropic agent so that the polypeptide chain can start to form secondary and tertiary structure. This is done

primarily by three methods: dilution, dialysis, or solid-phase refolding. In the case of disulphide containing proteins the refolding buffer often contains a redox shuffling system to allow the formation of native disulphide bonds. This requires the addition of reduced and oxidised thiol reagents, such as glutathione, cysteine or cysteamine, the ratio of which needs to be optimised for each protein (Wetlaufer *et al.* 1991). In 1991, Keifhaber *et al.* proposed a quantitative kinetic model whereby the yield of folded protein is determined by the competing rates of folding and aggregation. This model shows that at high protein concentrations, aggregation is favoured over folding, leading to insoluble protein. For this reason, *in vitro* protein refolding is done at low protein concentrations generally between 10-100µg/ml (Lilie *et al*. 1998). The method of refolding used here is drop-wise dilution. This involves the slow drop-wise addition of denatured protein into the refolding buffer, such that the final dilution will be 100-fold. Drop-wise addition of the resolubilised inclusion bodies removes the denaturant whilst keeping protein concentrations low to reduce aggregation and promote correct folding. Low molecular weight additives are often added to refolding mixtures to improve the folding yields of specific proteins. Common additives include non-denaturing amounts of chaotropic agents, detergents such as Triton X-100, metal ions, polyethylene glycol (PEG) and arginine (Rudolph and Lilie 1996). The refolding buffers used for the work reported here contain arginine at a concentration of 1M for PD-L1 and 0.4M for PD-1. Arginine contains a guanidine group but it does not destabilise the native structure of proteins as strongly as guanidine. The mechanism of how arginine improves refolding yields is not known but it is most probably due to solubilising effect of arginine on folding intermediates, therefore reducing aggregation (Lilie *et al.* 1998).

2.1.2 Circular Dichroism

For use in structural and functional studies the PD-L1 and PD-1 proteins needed to be correctly folded from inclusion bodies, and the stability of the proteins determined. Circular Dichroism (CD) is a spectroscopy method based on asymmetric molecules, in this case a polypeptide chain made up of chiral amino acids, differentially absorbing left- and right-handed circularly

polarised light. In CD spectroscopy the far-UV region, 190-250nm, is used, meaning that the peptide bond is the chromophore (Greenfield 2006). Because α-helix and β-sheet secondary structure have the peptide bonds arranged differently, this gives rise to distinct CD spectra (Greenfield and Fasman 1969). Since the absorption and therefore the CD spectra is dependent on the secondary structure of a protein, CD can be used to analyse changes in the secondary structure of a protein due to a range of different stimuli, for example, mutations, ligand binding and temperature. The unfolding of a protein can be studied by CD by measuring a loss of secondary structure as temperature increases. CD is therefore a good tool in determining the folding and stability of recombinant proteins.

2.1.3 Thermo Fluorescence Assay

In order to inform the optimal conditions for NMR studies of PD-L1 a thermal shift assay was performed using the dye SYPRO Orange. As the temperature is gradually increased, proteins begin to unfold exposing hydrophobic patches, to which SYPRO Orange binds. The SYPRO Orange fluorescence is strongly quenched by water and so when this is expelled upon binding to the hydrophobic patches there is an increase in fluorescence (Huynh and Partch 2015). This assay allows the melting point, and therefore stability, of a protein to be determined in a number of different buffer conditions.

2.2 Methods

2.2.1 Expression and purification of PD-L1 and PD-1

The genes encoding human PD-L1_D1 (18-134) and PD-L1_D1D2 (18-237) were synthesised and cloned into pET28a by GenScript, whilst the gene encoding human PD-1 (33-150) was synthesised and cloned into pET9a, also by GenScript. A construct for human PD-L1_D2 (130-239) was cloned into the pLEICS-05 vector by PROTEX (University of Leicester). All constructs had codon usage optimised for expression in *E. coli.* All constructs were expressed as untagged proteins. All proteins were expressed as insoluble inclusion bodies in *E. coli* BL21 (DE3) cells grown in

LB broth. Cells were cultured at 37°C and protein production induced with 1mM IPTG at OD_{600} 0.8-1.0. The cells were then cultured for a further 5 hours before being collected by centrifugation.

Cell pellets were resuspended in PBS, pH 7.2, and lysed by sonication before the inclusion bodies were collected by centrifugation. Inclusion bodies were washed twice with 50mM Tris-HCl pH 8.0, 200mM NaCl, 0.5% Triton-X100, 10mM EDTA, 10mM DTT, and once in the same buffer without Triton-X100. Washed inclusion bodies were resolubilised in 50mM Tris-HCl pH 8.0, 5M Guanidine-HCl, 200mM NaCl, 20mM DTT. Resolubilised inclusion bodies containing PD-L1_D1 were refolded by drop-wise dilution into 0.1M Tris-HCl pH 8.0, 1M Arginine, 0.25mM oxidised glutathione, and 0.25mM reduced glutathione (Zak *et al.* 2015). PD-L1_D1D2 and PD-L1_D2 were refolded by drop-wise dilution into the same buffer but with 0.5mM oxidised glutathione and 2mM reduced glutathione (adapted from Zak *et al.* 2016), and 0.25mM oxidised glutathione and 0.5mM reduced glutathione respectively. Resolubilised inclusion bodies containing PD-1 were refolded by drop-wise dilution into 0.1M Tris-HCl pH 8.0, 0.4M Arginine, 2mM EDTA, 0.5mM oxidised glutathione, 2mM reduced glutathione (Zak *et al.* 2015).

The refolding mixtures were then concentrated by tangential flow filtration before being dialysed into 10mM Tris pH 7.5, 20mM NaCl for size exclusion chromatography. The correctly folded PD-L1 and PD-1 proteins were separated from misfolded aggregates and contaminants by size exclusion chromatography using a 16/60 Seperdex 75 gel filtration column (GE Healthcare) equilibrated with 10mM Tris pH 7.5, 20mM NaCl. The purity of proteins and presence of disulphide bonds was checked by SDS-PAGE.

2.2.2 Circular Dichroism Spectroscopy

Circular Dichroism (CD) thermal denaturation studies were performed for the PD-L1 proteins to investigate the effect of increasing temperature on their structure as well as determine their melting temperature. Spectra were recorded on a Chirascan plus spectrometer between 190 nm – 260 nm at a protein concentration of 9µM, 14µM, and 24µM for PD-L1_D1D2, PD-L1_D1 and PD-L1_D2 respectively, in a 10mM Tris, 20mM NaCl buffer at pH 7.5 in a cell with path length 0.5 mm. The thermal denaturation analysis was performed between 20°C and 90°C with increments of 1°C for PD-L1_D1D2 and 5°C for PD-L1_D1 and PD-L1_D2.

2.2.3 Thermo Fluorescence Buffer Screen

Thermal fluorescence assays were performed over a range of salt concentrations from 0-150mM and with a pH range of 6.5-8.0 to determine the conditions in which PD-L1 is most stable for NMR studies. The assays were performed using a concentration of 5µM PD-L1_D1D2 and 10µM PD-L1_D1 and SYPRO Orange (Thermo Fischer) as per manufacturer's instructions. The assay was performed with a Bio Rad iCycler using a 96-well plate allowing each buffer condition to be screened in triplicate. The stability of PD-L1 was tested in 25mM sodium phosphate at pH 6.5, 7.0 and 7.5, as well as 25mM Tris at pH 7.5 and 8.0 with NaCl concentrations of 0, 50, 100 and 150mM tested at each pH.

2.3 Results

2.3.1 Expression and purification of PD-L1 and PD-1

All of the PD-L1 and PD-1 constructs expressed in BL21 (DE3) *E. coli* cells showed good post-induction expression and washed inclusion bodies showed high levels of desired protein, as analysed by SDS-PAGE. After refolding and concentration by tangential flow all correctly folded PD-L1 and PD-1 constructs were separated from misfolded disulphide linked aggregates by size exclusion chromatography producing pure protein as analysed by SDS-PAGE.

Figure 2.2 shows the expression and purification of PD-L1_D1D2. Panel C shows the size exclusion chromatogram after refolding and concentration of PD-L1_D1D2, where three major peaks elute from the column. Comparison of reduced and non-reduced SDS-PAGE gels shows the first peak to elute at 45ml contains higher order disulphide-linked aggregates of PD-L1_D1D2 (Figure 2.2). SDS-PAGE analysis of the peak to elute at 72ml shows that it contains pure monomeric PD-L1_D1D2. Comparison of reduced and non-reduced PD-L1_D1D2 SDS-PAGE gels shows a mobility shift with non-reduced PD-L1_D1D2 traveling further through the gel suggesting that disulphide bonds had formed during refolding. PD-L1_D1D2 is a 25kDa protein that elutes from the size exclusion column where expected as determined by calibration with known proteins. SDS-PAGE analysis of the peak eluted at 103ml shows that no protein is contained in this peak. The absorbance seen here is likely due to residual arginine and/or glutathione eluting from the column. The final yield of PD-L1_D1D2 was 20 mg per litre of original culture with refolding efficiency estimated at 20% based on 100 mg of PD-L1_D1D2 being present in the inclusion bodies, estimated from absorbance at 280nm of inclusion bodies resolubilised in 5M guanidine.
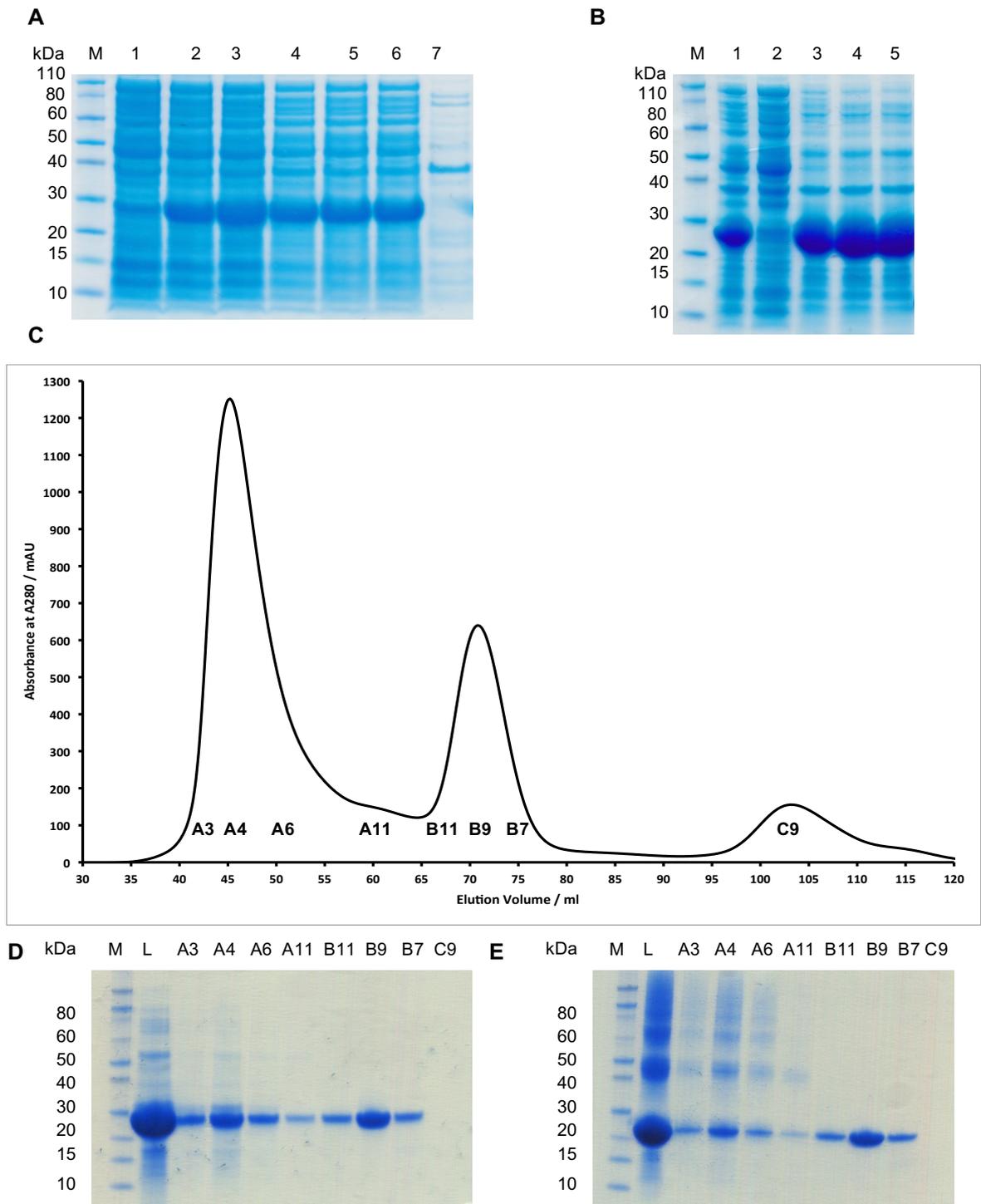
Figure 2.2. Expression and purification of PD-L1_D1D2.

Panel A shows SDS-PAGE analysis of the expression of PD-L1_D1D2. Lane 1 contains whole cell lysate (WCL) pre-induction. Lanes 2-6 contain WCL at hourly intervals post-induction. Lane 7 contains the soluble fraction 5 hours post-induction showing that PD-L1_D1D2 is expressed insolubly. Panel B shows the WCL (1) and soluble fraction (2) of PD-L1_D1D2 expression. Lanes 3-5 show the inclusion body washes. Panel C shows a size exclusion chromatogram of refolded PD-L1_D1D2. Panels D and E show SDS-PAGE analysis of the load and eluted fractions from the size exclusion column with and without the reducing agent DTT respectively. Monomeric PD-L1_D1D2 has been separated from disulphide-linked misfolded aggregates.

The size exclusion chromatogram after refolding and concentration of PD-L1_D1 is shown in Figure 2.3. Two major peaks elute from the column with smaller minor peaks eluting first between 40 and 75 ml. SDS-PAGE analysis of the minor peaks shows that they do contain PD-L1_D1 in small quantities but there are faint bands of higher order contaminants also present. SDS-PAGE analysis of the peak to elute at 82ml shows that this contains pure monomeric PD-L1_D1 (Figure 2.3). Again, there is a mobility shift between reduced and non-reduced samples suggesting that disulphide bonds have formed during refolding. The peak to elute at 103ml contains no protein and the absorbance seen here is likely due to residual arginine and/or glutathione eluting from the column. The final yield of PD-L1_D1 was 15mg per litre of culture with an estimated refolding efficiency of 20%.
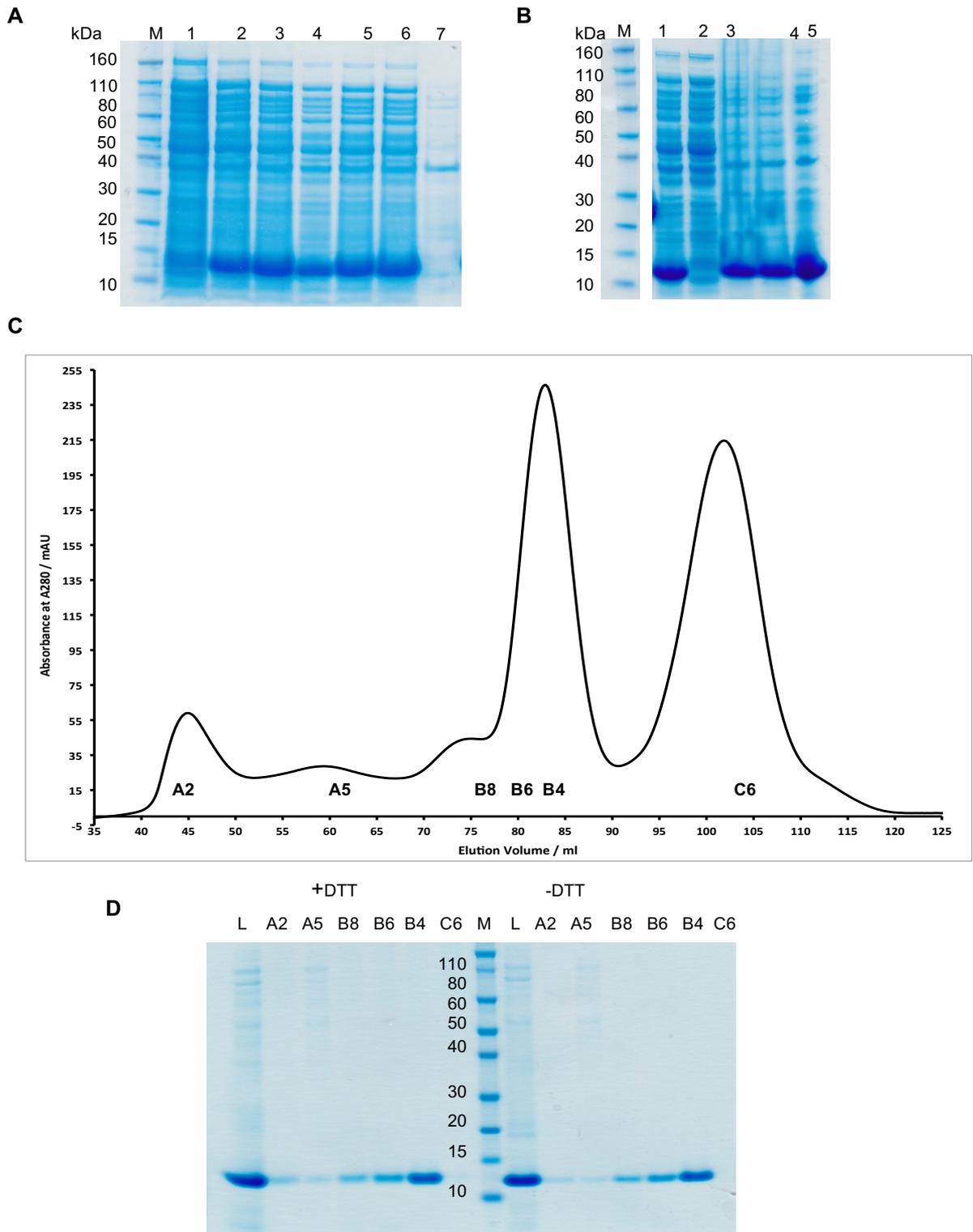
Figure 2.3. Expression and purification of PD-L1_D1.

Panel A shows SDS-PAGE analysis of the expression of PD-L1_D1. Lane 1 contains whole cell lysate (WCL) pre-induction. Lanes 2-6 contain WCL at hourly intervals post-induction. Lane 7 contains the soluble fraction 5 hours post-induction showing that PD-L1_D1 is expressed insolubly. Panel B shows the WCL (1) and soluble fraction (2) of PD-L1_D1 expression. Lanes 3-5 show the inclusion body washes. Panel C shows a size exclusion chromatogram of refolded PD-L1_D1. Panel D shows SDS-PAGE analysis of the load and eluted fractions from the size exclusion column with and without the reducing agent DTT. Monomeric PD-L1_D1 has been separated from higher molecular weight contaminants.

34

Figure 2.4 shows the expression and purification of PD-L1_D2. Panel C shows the size exclusion chromatogram of PD-L1_D2 after refolding and concentration. There are three peaks that elute from the column. The first in the void volume at 40ml contains higher molecular weight contaminants and small amounts of PD-L1_D2 as visualised by SDS-PAGE analysis. The second peak to elute at 76ml contains pure monomeric PD-L1_D2. Comparison of reduced and non-reduced SDS-PAGE analysis shows that there is a leading shoulder on this peak that contains disulphide-linked dimers of PD-L1_D2 but by fraction B9 all of this mis-folded contaminant had eluted (Figure 2.4). Comparison of reduced and non-reduced SDS-PAGE gels again shows a mobility shift, suggesting that disulphide bonds have formed during refolding. SDS-PAGE analysis of the peak to elute at 100ml shows that this peak contains no protein. The absorbance seen here is likely due to residual arginine and/or glutathione eluting from the column. The final yield of PD-L1_D2 was 22mg per litre of culture with an estimated refolding efficiency of 20%.

Finally, SDS-PAGE analysis of the expression and purification of PD-1 is shown in Figure 2.5. The major peak seen in the size exclusion chromatogram eluting at 76 ml contains pure monomeric PD-1 as determined by SDS-PAGE analysis. The final yield of PD-1 was 20mg per litre of culture with an estimated protein folding efficiency of 20%.
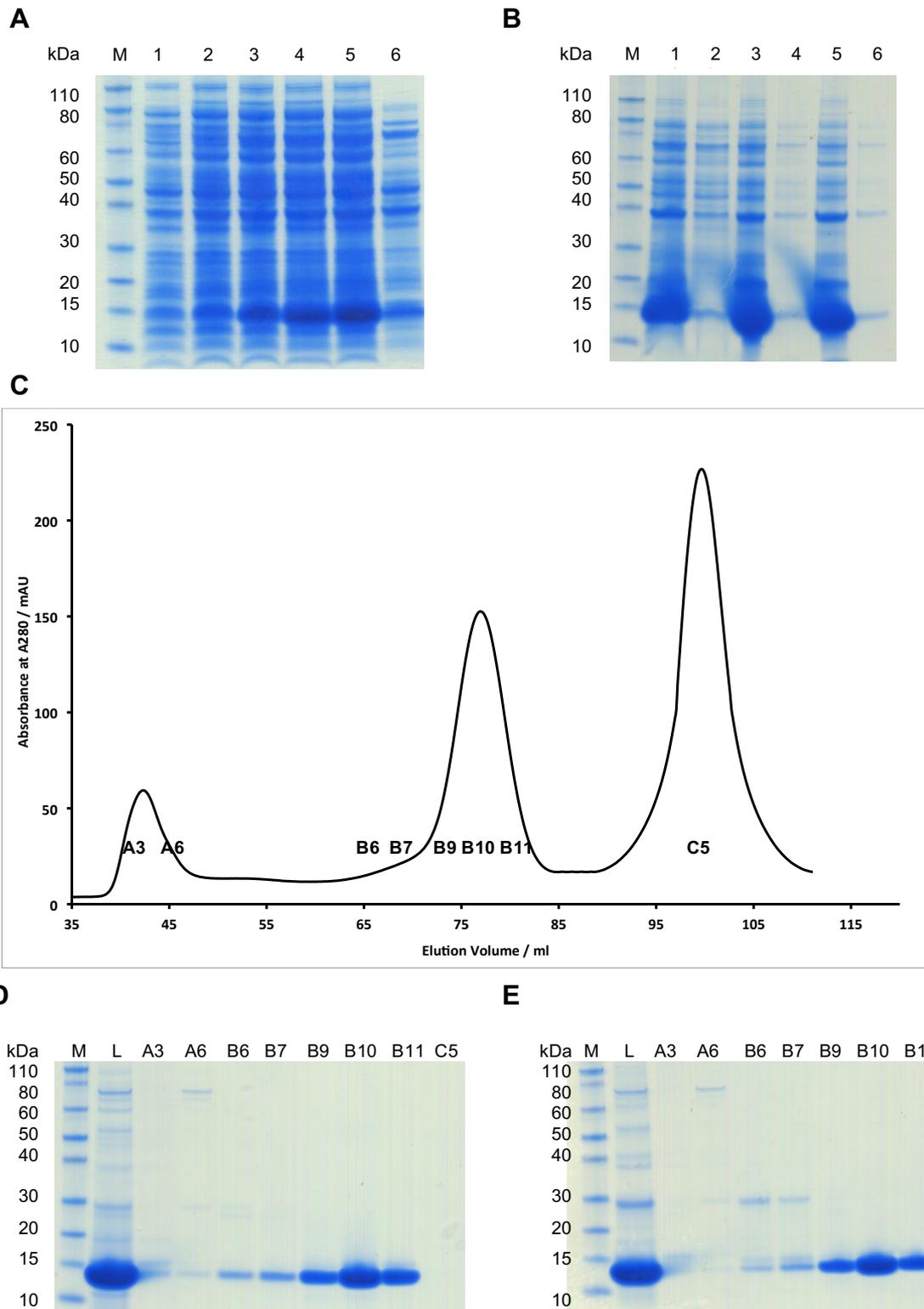
Figure 2.4. Expression and purification of PD-L1_D2.

Panel A shows SDS-PAGE analysis of the expression of PD-L1_D2. Lane 1 contains whole cell lysate (WCL) pre-induction. Lanes 2-5 contain WCL at hourly intervals post-induction. Lane 6 contains the soluble fraction 5 hours post-induction showing that PD-L1_D2 is expressed insolubly. Panel B shows SDS-PAGE of the inclusion body washes. Lanes 2, 4, 6 show what is solubilised in each wash the Panel C shows a size exclusion chromatogram of refolded PD-L1_D2. Panels D and E show SDS-PAGE analysis of the load and eluted fractions from the size exclusion column with and without the reducing agent DTT respectively. Monomeric PD-L1_D2 has been separated from disulphide-linked misfolded aggregates and higher molecular weight contaminants.
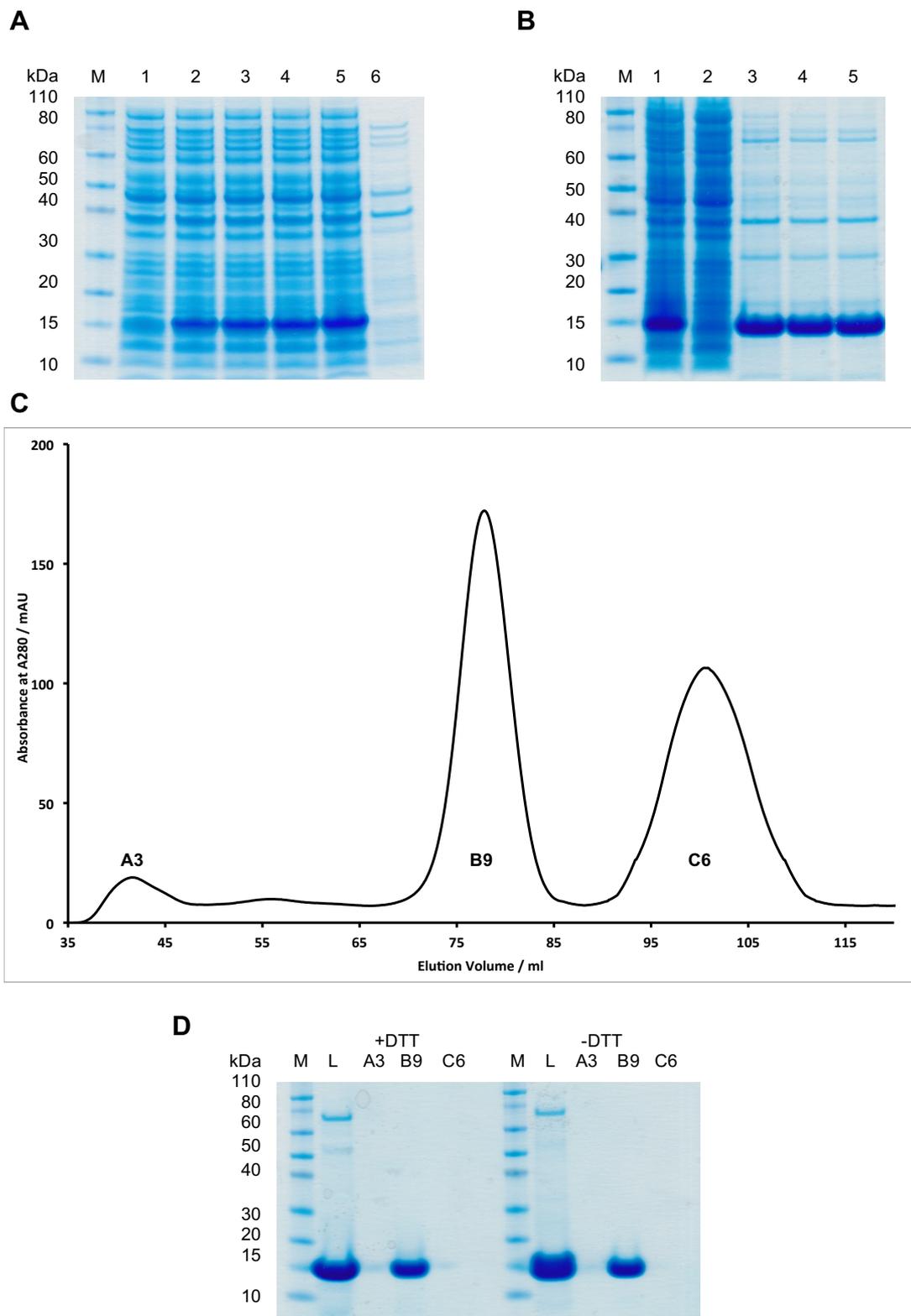
Figure 2.5. Expression and purification of PD-1.

Panel A shows SDS-PAGE analysis of the expression of PD-1. Lane 1 contains whole cell lysate (WCL) pre-induction. Lanes 2-5 contain WCL at hourly intervals post-induction. Lane 6 contains the soluble fraction 5 hours post-induction showing that PD-1 is expressed insolubly. Panel B shows the WCL (1) and soluble fraction (2) of PD-1 expression. Lanes 3-5 show the inclusion body washes. Panel C shows a size exclusion chromatogram of refolded PD-1. Panel D shows SDS-PAGE analysis of the load and eluted fractions from the size exclusion column with and without the reducing agent DTT. Monomeric PD-1 has been separated from higher molecular weight contaminants.

2.3.2 Circular Dichroism Thermal Stability Assay

CD spectroscopy was used to assess whether the PD-L1 proteins were folded. The far-UV spectra for PD-L1_D1D2 showed typical β-sheet folded structure with a negative peak at 215nm (Figure 2.6). A thermal melt was performed where the far-UV absorbance of PD-L1_D1D2 was measured at 1°C intervals between 10°C and 90°C (Figure 2.7a). A thermal denaturation curve was produced by plotting the intensity of the negative peak at 215nm against increasing temperature (Figure 2.7b). The midpoint of this curve is the temperature at which 50% of PD-L1_D1D2 is denatured, and this is the melting temperature. This showed a melting temperature of 52°C for PD-L1_D1D2, with the sample being 25% denatured at 49°C and unfolding beginning at approximately 40°C (Figure 2.7b). NMR experiments were performed at 30°C and it is shown in Figure 2.7b that PD-L1_D1D2 is essentially 100% folded at this temperature.
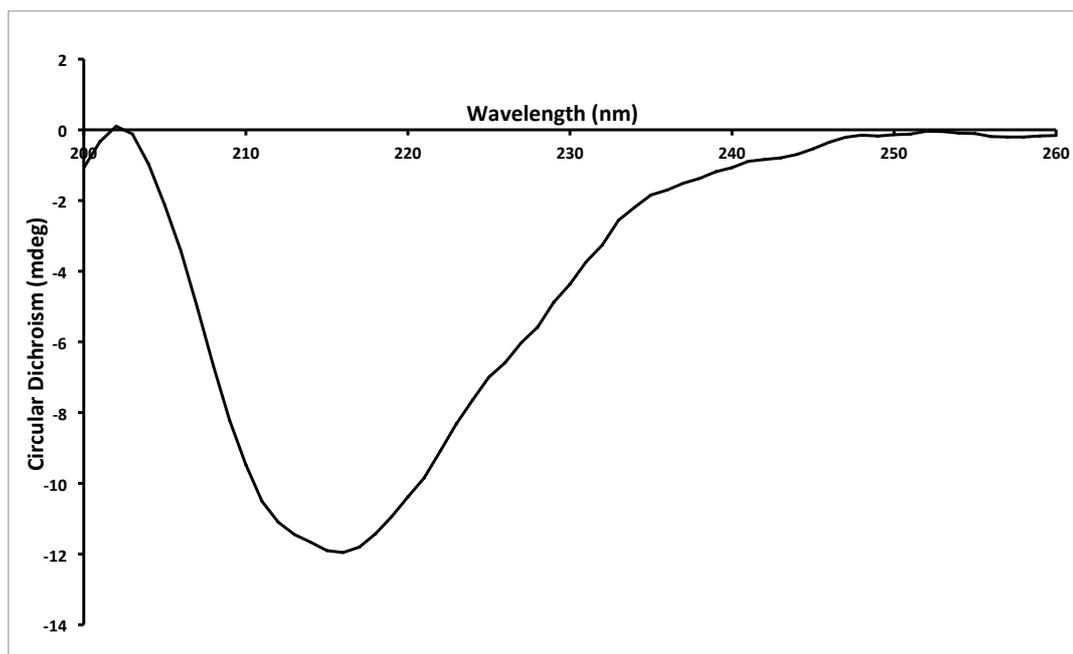
Figure 2.6. Circular Dichroism absorption spectrum of PD-L1_D1D2.
This shows the far-UV CD spectrum of 9 µM PD-L1_D1D2 in 10 mM Tris pH8.0 at 20°C.

A
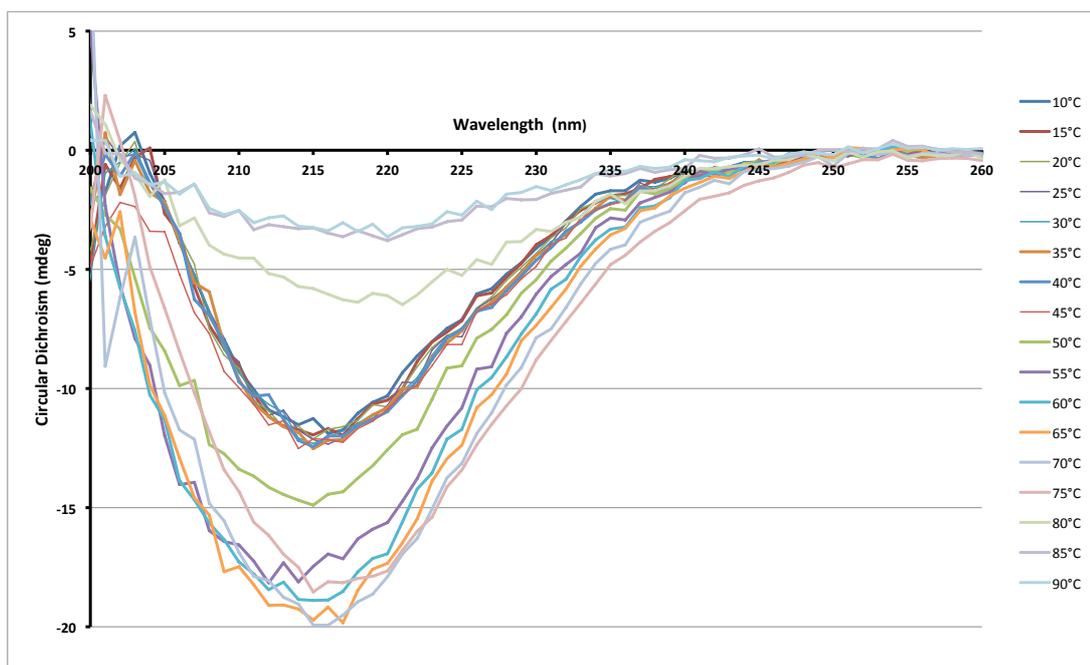
B
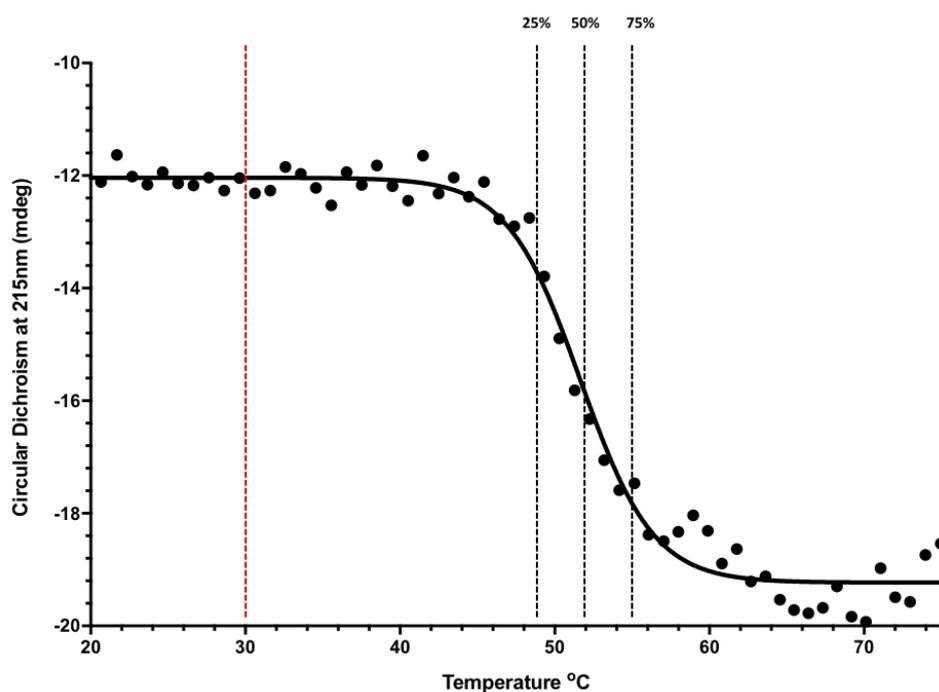
Figure 2.7. Thermal denaturation studies of PD-L1_D1D2.

A) Overlay of the far-UV CD spectra of PD-L1_D1D2 at increasing temperatures of 5°C increments. **B)** The thermal denaturation of PD-L1_D1D2 as measured by CD spectroscopy at a wavelength of 215nm as the temperature is increased in 1°C increments. Black dashed lines show the temperature at which PD-L1_D1D2 is 25%, 50% and 75% unfolded. The red dashed line is the temperature at which NMR experiments were performed.

The far-UV spectra for PD-L1_D1 also showed typical β-sheet folded structure with a negative peak at 217nm (Figure 2.8). A thermal melt was performed where the far-UV absorbance of PD-L1_D1 was measured at 1°C intervals between 20°C and 90°C (Figure 2.9a). A thermal denaturation curve was produced by plotting the intensity of the positive peak at 200 nm as temperature is increased showed a melting temperature of 44°C but with unfolding starting at as low temperature as 35°C (Figure 2.9b). 200nm was chosen because this is where there is the greatest difference in absorbance between folded and denatured PD-L1_D1. NMR experiments were performed at 25°C and it can be seen from Figure 2.9b that PD-L1_D1 is essentially 100% folded at this temperature.
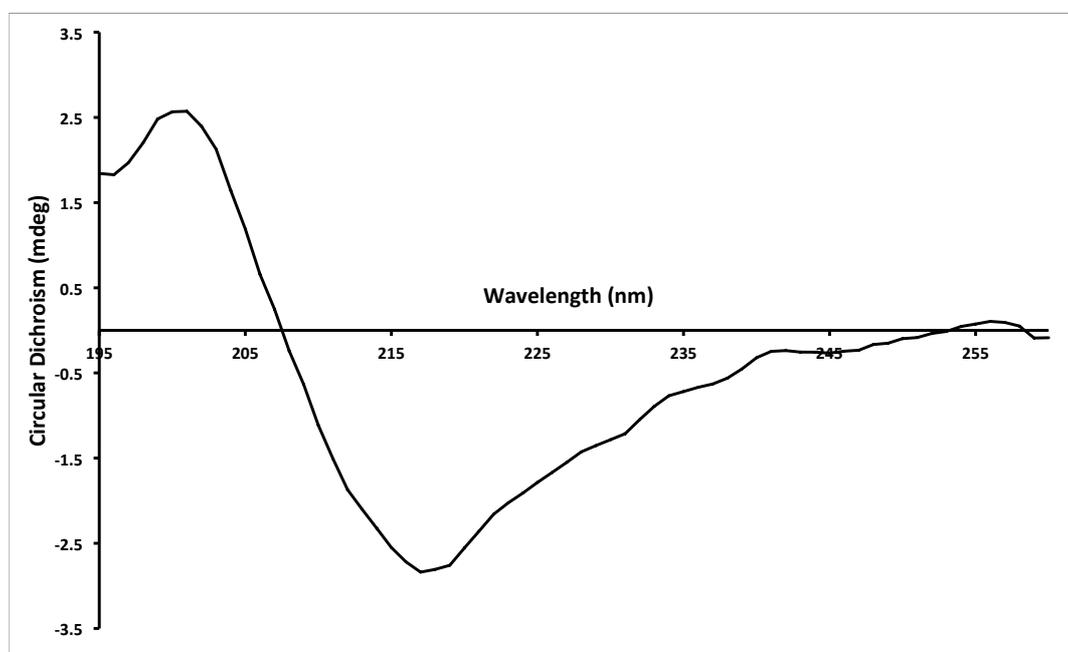
Figure 2.8. Circular Dichroism absorption spectrum of PD-L1_D1.
This shows the far-UV CD spectra of 14 µM PD-L1_D1 in 10 mM Tris pH 8.0 at 20°C.
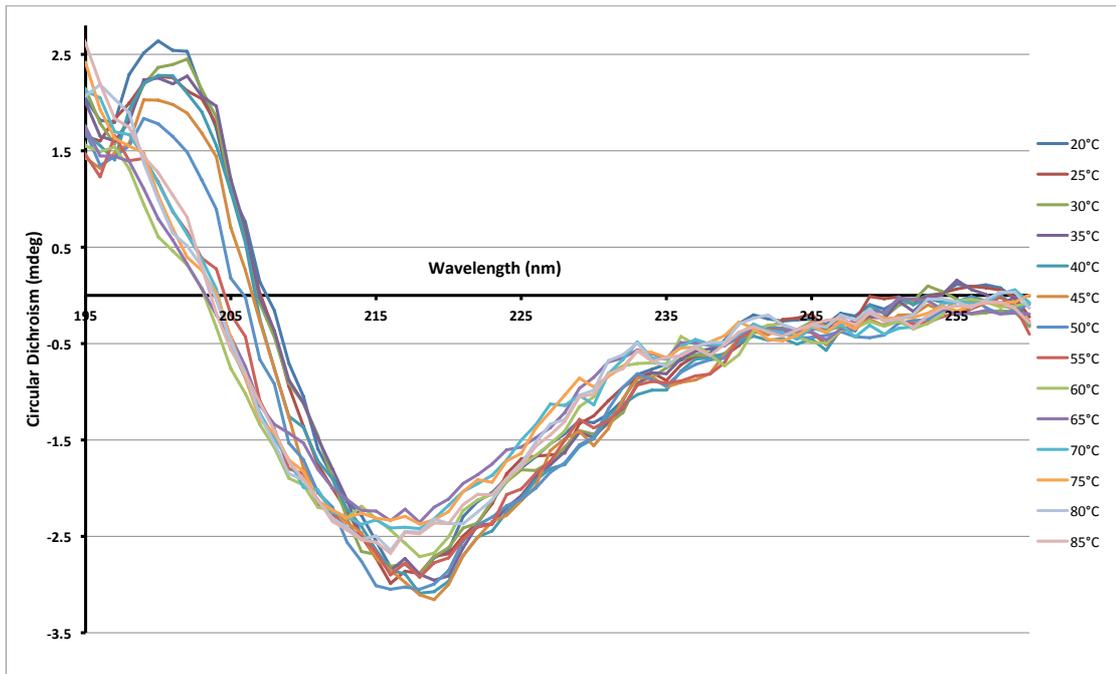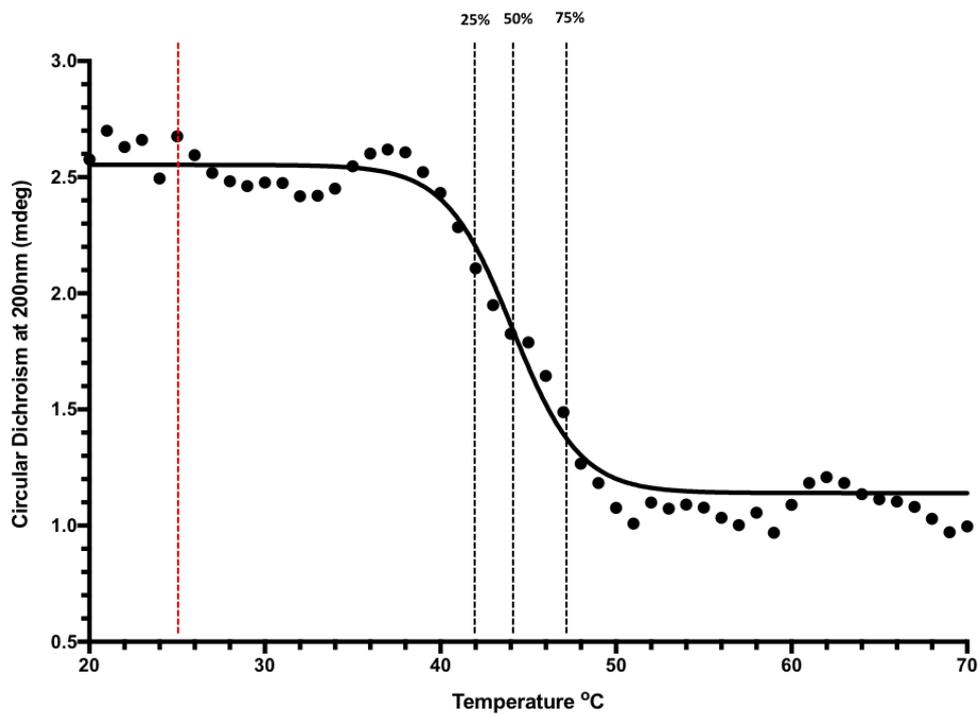
Figure 2.9. Thermal denaturation studies of PD-L1_D1.

A) Overlay of the far-UV CD spectra of PD-L1_D1D2 at increasing temperatures of 5°C increments. **B)** Thermal denaturation of PD-L1_D1 as measured by CD spectroscopy at a wavelength of 200nm as the temperature is increased in 1°C increments. Black dashed lines show the temperature at which PD-L1_D1 is 25%, 50% and 75% unfolded. The red dashed line is the temperature at which NMR experiments were performed.

The far-UV spectrum for PD-L1_D2 shows characteristics of β-sheet secondary structure, with a negative peak at 212nm, and random coil structure, with a second negative peak at 198nm (Figure 2.10). This can be explained by the loops between the β-strands of the IgC-like PD-L1_D2 significantly contributing to the CD absorption, as well as the β-sheet core. Thermal denaturation studies measuring the intensity at 204 nm as temperature is increased showed a melting temperature of 46°C for PD-L1_D1D2, with unfolding beginning below 30°C (Figure 2.11b). At 41°C 25% of PD-L1_D2 is unfolded and at 25°C, the temperature at which NMR experiments were performed, there is an insignificant amount of unfolded PD-L1_D2, estimated at 1% (Figure 2.11b).
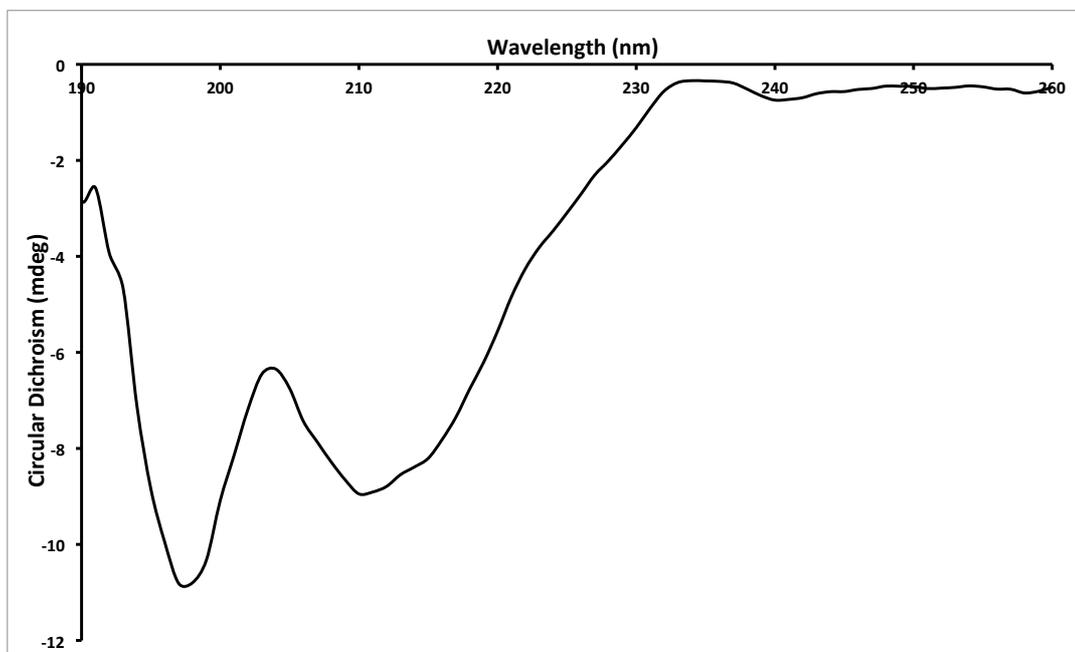
Figure 2.10. Circular Dichroism absorption spectrum of PD-L1_D2.
This shows the far-UV CD spectra of 24 μM PD-L1_D2 in 10 mM Tris pH 8.0 at 20°C.

A
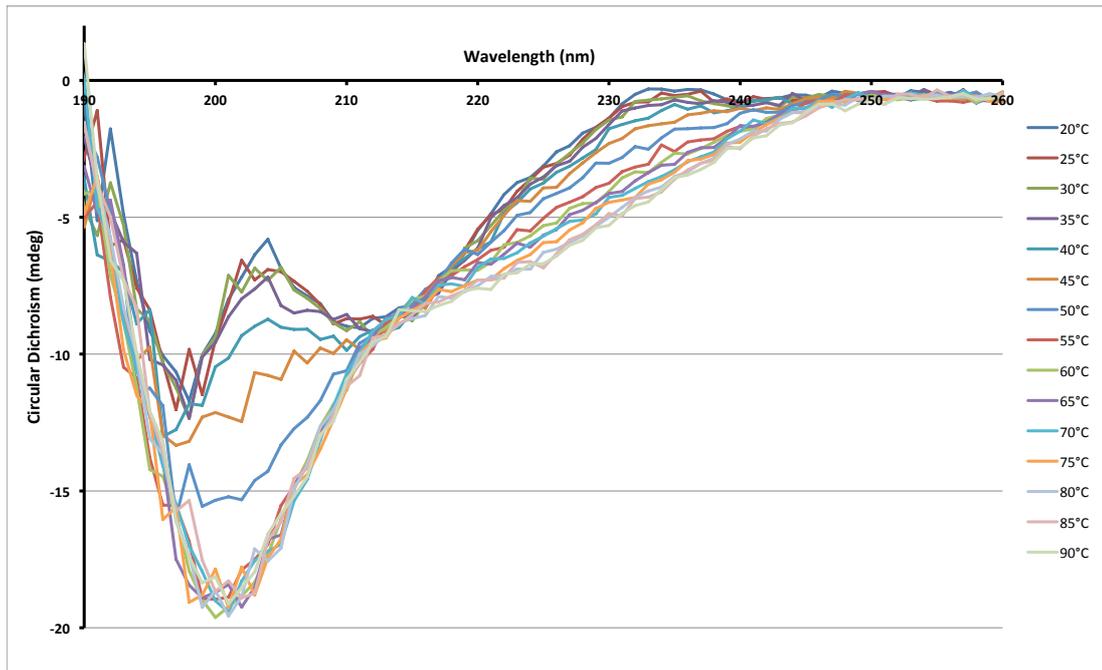
B
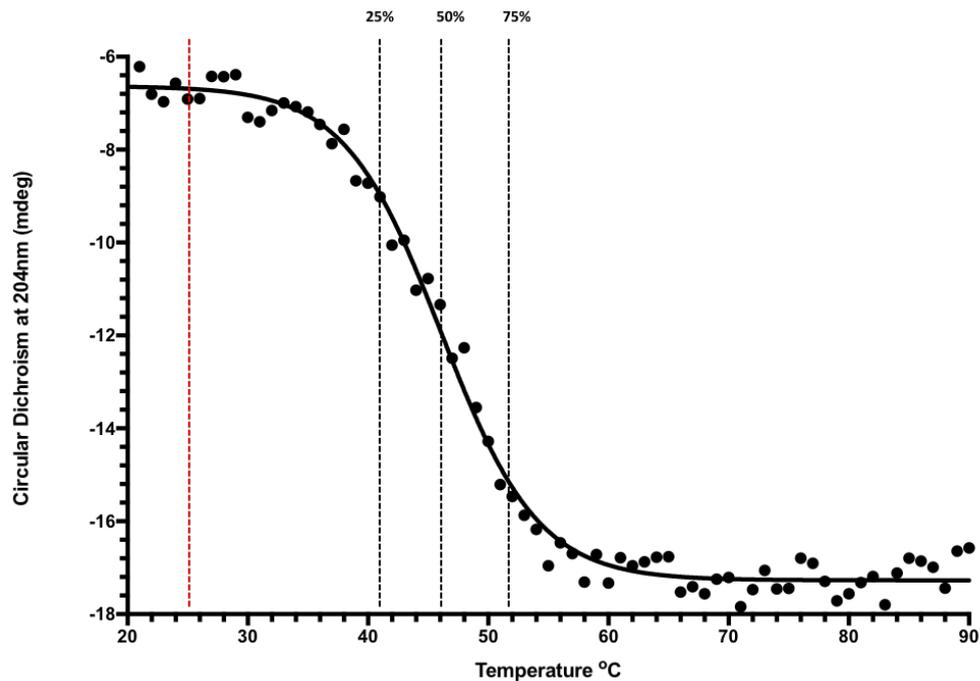
Figure 2.11. Thermal denaturation studies of PD-L1_D2.
A) Overlay of the far-UV CD spectra of PD-L1_D2 at increasing temperatures of 5°C increments. **B)** Thermal denaturation of PD-L1_D2 as measured by CD spectroscopy at a wavelength of 204nm as the temperature is increased in 1°C increments. Black dashed lines show the temperature at which PD-L1_D2 is 25%, 50% and 75% unfolded. The red dashed line is the temperature at which NMR experiments were performed.

2.3.3 1D ^{1}H NMR Spectroscopy

NMR spectroscopy is a good method to determine the folded state of a protein sample since the spectra of unfolded and folded proteins have easily visible differences. Since the spectrum of an unfolded protein is effectively the sum of the random coil chemical shifts of the amino acids in the sequence, the signals in a 1D ^{1}H NMR spectrum of an unfolded protein have a narrow dispersion (Bundi & Wuthrich 1979). In contrast, for a folded protein signals in the 1D ^{1}H NMR spectrum are more widely dispersed due to the protons sampling a more diverse range of chemical microenvironments (Bundi & Wuthrich 1977). Figure 2.12 shows the 1D ^{1}H NMR spectra for PD-L1_D1D2, PD-L1_D1 and PD-L1_D2. All spectra show wide dispersion in the amide region between 6-9 ppm as well as visible peaks in the ~10 ppm region, where tryptophan indole signals are expected. All spectra also show low field shifted methyl signals below 0 ppm, indicative of a folded protein.
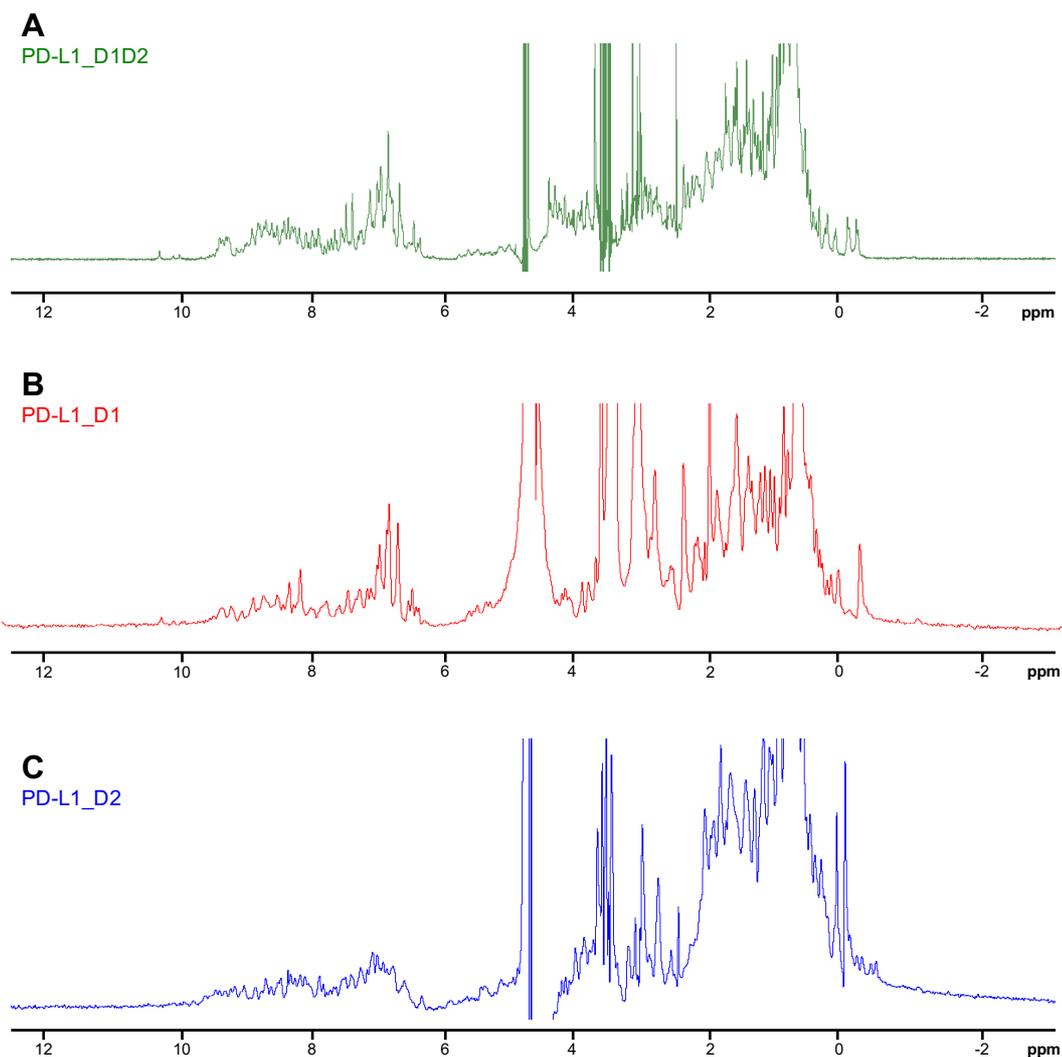
Figure 2.12. 1D ^{1}H NMR Spectroscopy of PD-L1 constructs.

A) 1D ^{1}H NMR spectrum for 200μM PD-L1_D1D2 with ^{15}N,^{13}C and non-exchangable proton ^{2}H labelling. The experiment (32 scans) was collected on a 950 MHz spectrometer at 30°C in 25 mM KH2PO4 (pH 7.5), 25 mM NaCl buffer conditions. Water suppression was achieved using excitation sculpting and pulse gradients. **B)** 1D ^{1}H NMR spectrum for 100μM PD-L1_D1. The experiment (4 scans) was collected on a 600 MHz spectrometer at 25°C in 25 mM KH2PO4 (pH 7.5), 25 mM NaCl buffer conditions. Water suppression was achieved by pre-saturation of the water signal. **C)** 1D ^{1}H NMR spectrum for 170μM PD-L1_D2 with ^{15}N,^{13}C labelling. The experiment (4 scans) was collected on an 800 MHz spectrometer at 25°C in 25 mM KH2PO4 (pH 7.5), 25 mM NaCl buffer conditions. Water suppression was achieved by pre-saturation of the water signal.

2.3.4 Thermo Fluorescence Buffer Screen

A thermal fluorescence buffer screen was performed with PD-L1_D1D2 and PD-L1_D1 to determine the optimum conditions for stability during NMR experiments. The pH of the buffer over the range tested (6.5-8.0) does not appear to affect the stability of PD-L1_D1D2 or PD-L1_D1. The stability of PD-L1_D1 was also not affected by the concentration of NaCl in the buffer, over the 0-150mM range tested (Figure 2.13).
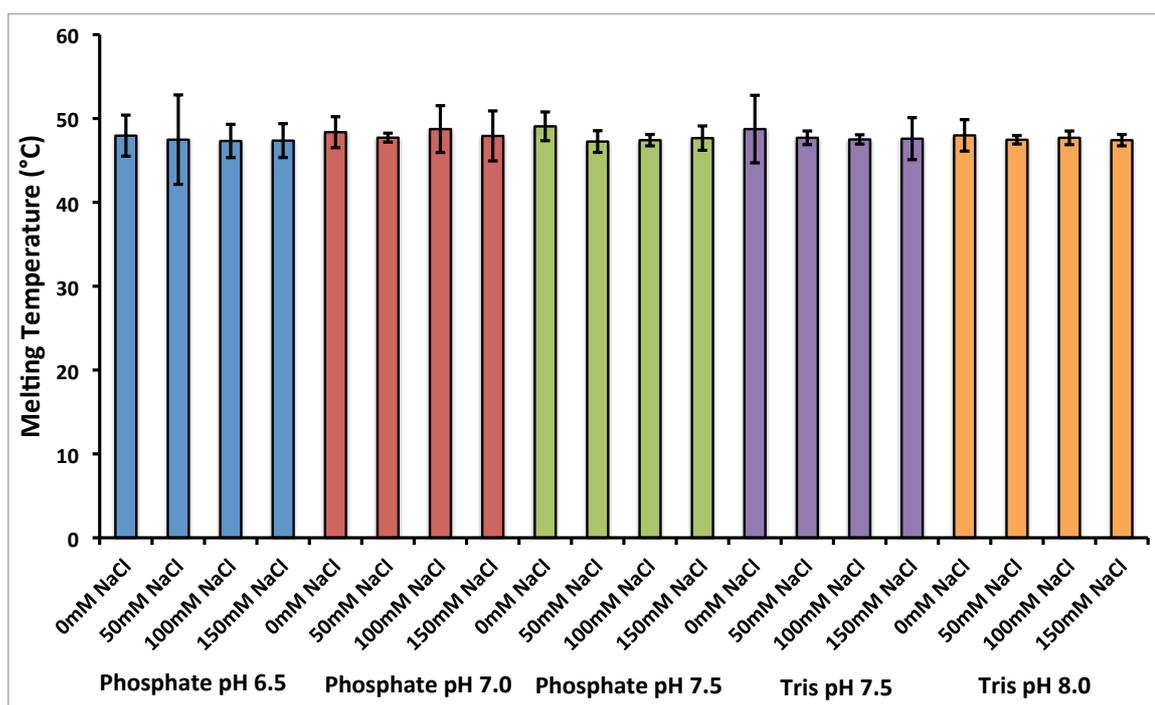
Figure 2.13. PD-L1_D1 buffer screen.

A bar chart showing the melting temperature of PD-L1_D1 in different buffers with varying pH and NaCl content. This shows that varying pH and/or NaCl has no significant difference on the melting temperature of PD-L1_D1. Error bars show standard deviation across three repeats.

The stability of PD-L1_D1D2, however, improved with low NaCl, as shown by a reduction in the melting temperature as the concentration of NaCl was increased (Figure 2.14). This pattern was seen across all the buffers tested and there was a drop in average melting temperature from 46.5°C at 0 mM Nacl to 44 °C at 150 mM NaCl.
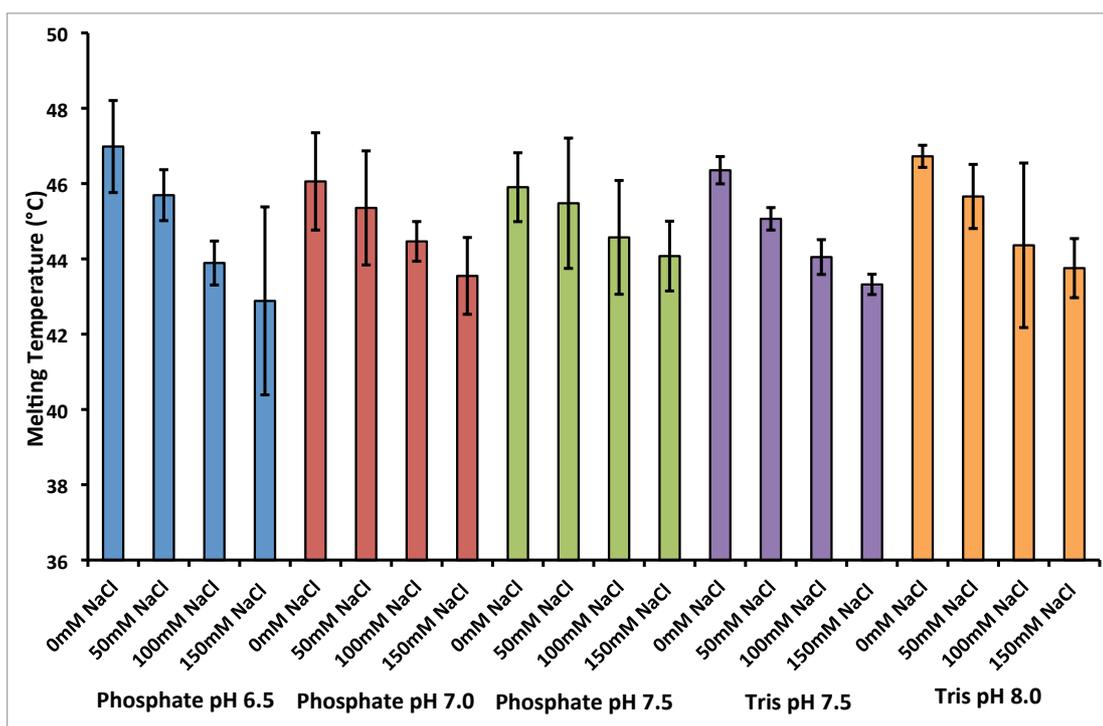
Figure 2.14. PD-L1_D1D2 buffer screen.

A bar chart showing the melting temperature of PD-L1_D1D2 in different buffers with varying pH and NaCl content. This shows that pH does not affect the stability PD-L1_D1D2 over the pH range tested. Increasing the NaCl concentration reduces the melting temperature of PD-L1_D1D2. Error bars show standard deviation across three repeats.

2.4 Discussion

The results presented here show that the selected PD-L1 and PD-1 constructs can be expressed in *E. coli* and refolded from inclusion bodies with yields of up to 20 mg per litre of original culture. This is achieved by having high initial expression and relatively good refolding efficiency, estimated at 20%, for all proteins. The high yield of unlabelled proteins from cultures grown in LB demonstrates that enough labelled protein should be produced from *E. coli* grown in minimal media for NMR studies, where there is often a reduced protein yield, especially when cultures are grown using D_2O.

CD analysis of the different PD-L1 constructs show that all proteins have the expected β-sheet secondary structure and are stable to high enough temperatures to perform NMR experiments at 25 or 30°C. Inspection of 1D ^{1}H NMR spectra of the PD-L1 constructs showed a wide dispersion of signals indicating that the proteins are folded.

The buffer screen shows that, over the range of salt concentration tested, low salt slightly improves the melting temperature for PD-L1_D1D2 and makes no difference to the melting temperature of PD-L1_D1. Using a buffer with low salt is advantageous for NMR studies when using a cryogenic probe as the signal-to-noise ratio, and therefore the sensitivity, is reduced by samples that are electrically conductive (Kelly *et al.* 2002). Kelly *et al.* have shown that greater sensitivity can be achieved during NMR experiments by using buffers with lower conductivity than phosphate, however due to phosphate being non-protonated it makes an ideal buffer for NMR and is the most commonly used buffer for NMR (Kelly *et al.* 2002). Based on this the NMR studies discussed in Chapters 3 and 4 were conducted in 25 mM KH_2PO_4 (pH 7.5), 25 mM NaCl buffer conditions.

Chapter 3

The binding of a bi-phenyl inhibitor to PD-L1

3.1 Introduction

As described in the introductory chapter of this thesis, the PD-1/PD-L1 interaction is a well-established therapeutic target with multiple biologic licensed therapeutics targeting both PD-1 and PD-L1. The discovery of small molecules targeting this interaction however, lags behind with no licensed drugs currently available. The PD-1/PD-L1 interface is large, flat and hydrophobic and therefore considered difficult to target with small molecules. Nonetheless, there are peptides, cyclic peptides and small molecules that have been shown to target this interaction (Zarganes-Tzitzikas *et al.* 2016). Some of these are summarised in Table 3.1.

This chapter describes the investigation of the binding of a bi-phenyl (BPI) inhibitor of PD-L1, using NMR. The BPI, shown in Table 3.1, was chosen from a set of molecules patented by Bristol-Myers Squibb (WO2015/034820A1) described as "compounds useful as immunomodulators" based on its ability to be synthesised by collaborators at the University of Victoria, Canada. The patent discloses that the compounds block the PD-1/PD-L1 interaction in an Homogenous Time-Resolved Fluorescence (HTRF) assay. Collaborators from the University of Victoria tested their synthesised compound in a Biacore assay (unpublished). In this assay the extracellular domain of PD-1 was immobilised on the chip and PD-L1, alone and in the presence of BPI, was in solution. This gave an IC_{50} value of 20 nM for the BPI used in this work. This IC_{50} was used as a guide, assuming affinities in the nanomolar range, when determining concentrations for use in NMR experiments.

Table 3.1 Summary of small molecules targeting PD-L1

Molecule	Description	Company	Activity
	A series of molecules with an oxadiazole flanked by a serine/threonine and a glutamine/ asparagine/ glutamate/aspartate	Aurigene Ltd. WO2015/033301A1	Rescue assay of mouse splenocytes in the presence of recombinant mouse PD-1/PD-L1. Rescued cells to 92% at 100 nM
SNTSESF-NH SNTSESFK FRVTQLAPKAQIKE-NH$_2$	A peptide antagonist of PD-L1 that was designed using human and murine PD-1 sequences involved in the binding of PD-L1	Aurigene Ltd. US2011/0318373A3	Highly effective in antagonizing PD-1 signalling, with in vivo exposure upon subcutaneous dosing
	Macrocyclic peptides	Bristol-Myers Squibb US2014/0294898A1	Nanomolar potency in HTRF assays. Were shown to block binding of PD-L1 to Jurkat-PD-1 cells with nanomolar IC$_{50}$ values
 Used in this work	Series of molecules with a biphenyl scaffold	Bristol-Myers Squibb WO2015/034820A1	Nanomolar potency in HTRF assays
	Optimised molecules from biphenyl scaffold series	Bristol-Myers Squibb WO2015/160641 A2	EC$_{50}$ values in triple digit nanomolar range in T-cell activation assay
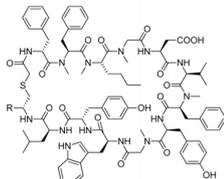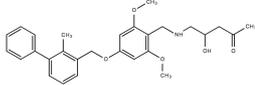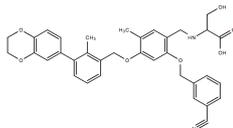
Up until recently no small molecules of this kind had been shown to have an effect in cellular assays. However, Bristol-Myers Squibb have recently patented a series of optimised molecules that have been shown to have some activity in cellular assays (WO2015/160641A2) (Skalniak *et al*. 2017). However, there were notable differences in the potency of the molecules in these cellular assays depending on whether PD-L1 (18-134) or PD-L1 (18-239) was used. This will be commented on further in Chapter 5.

3.1.1 Chemical shift mapping

NMR is a useful tool for the investigation of ligand binding to a protein due to chemical shift perturbation. This is where the chemical shift changes that occur when a ligand is added to a protein are used to determine the location of the ligand-binding site. The standard methodology of this experiment uses ^{15}N-labelled protein, and 2D HSQC spectra are acquired with and without unlabelled ligand, whether that is a small molecule or protein (Zuiderweg 2002). The ^{15}N-HSQC experiment shows H-N correlations and therefore it can be expected to see a signal corresponding to the backbone amide for each amino acid of the polypeptide chain excluding prolines. One can determine which signal belongs to which residue by using triple resonance NMR experiments using ^{13}C/^{15}N-labelled protein. This allows the assignment of the protein backbone by determining the sequential neighbours of a given spin system (Kay *et al.* 1990). Any changes in the chemical environment of protein backbone amides caused by ligand binding are represented by changes in the chemical shift of the signal corresponding to that particular amino acid. This can then be used to map the ligand-binding site onto the protein in question, as well as determine any conformational changes upon binding. This chapter describes the protein backbone assignment of PD-L1_D1 and PD-L1_D1D2 as well as the chemical shift perturbation analysis of a BPI binding to PD-L1. Here, what is known as a minimal shift approach is used (Williamson *et al.* 1997). The rationale for using a minimal shift approach is that the bound form of the protein does not need to be assigned. In this instance the geometrical distance between each signal in the free spectrum and the nearest signal in the bound spectrum is taken as the minimal shift.

3.2 Methods

3.2.1 Isotopic labelling of PD-L1

PD-L1_D1 and PD-L1_D1D2 was expressed as described in Chapter 2 with the exception that BL21(DE3) cells were grown in minimal media with ^{15}N-NH$_4$Cl as the sole nitrogen source to produce ^{15}N-labelled protein. To produce ^{13}C/^{15}N-labelled protein ^{15}N-NH$_4$Cl was used as the sole nitrogen source and ^{13}C-glucose as the sole carbon source. To produce ^{13}C/^{15}N/^{2}H-PD-L1_D1D2 100% D$_2$O was used in preparation of the media. Isotopically labelled proteins were refolded and purified as described in Chapter 2.

Table 3.2 Minimal media ingredients per litre of culture

NH$_4$Cl	1.0 g	ZnCl$_2$	0.5 mg
Na$_2$PO$_4$	6.8 g	CuCl$_2$	0.1 mg
KH$_2$PO$_4$	3.0 g	CoCl$_2$	0.1 mg
NaCl	0.5 g	H$_3$BO$_3$	0.1 mg
Na$_2$SO$_4$	0.3 mM	MgSO$_4$	1.0 mM
EDTA	50 mg	CaCl$_2$	0.3 mM
MnCl$_2$	16 mg	d-Biotin	1.0 mg
FeCl$_3$	5.0 mg	Thiamine	1.0 mg
		Glucose	4.0 g

3.2.2 NMR backbone assignment of PD-L1_D1 and PD-L1_D1D2

Sequence specific backbone resonance assignments (N, NH, Cα, Cβ, C') were obtained for PD-L1_D1 and PD-L1_D1D2 using intra- and inter-residue connectivities where amide NH to Cα and Cβ peaks were identified in triple resonance spectra and used to search for i-1 or i+1 residues in the sequence. Assignments for PD-L1_D1 were determined using ^{15}N/^{13}C/^{1}H HNCACB (Wittekind & Mueller 1993), CBCA(CO)NH (Grzesiek & Bax 1992) and HNCO (Kay *et al.* 1990) spectra whereas assignments for PD-L1_D1D2 were determined using ^{15}N/^{13}C/^{2}H HNCACB, HN(CO)CACB spectra.

NMR spectra were collected from 190µM and 240µM samples of PD-L1_D1 and PD-L1_D1D2 respectively in a 25mM potassium phosphate, 20mM sodium chloride, 10µM EDTA, 0.02% sodium azide (w/v) buffer at pH 7.5,

containing 5% D_2O. NMR experiments for PD-L1_D1 were collected at 25°C whereas NMR experiments of PD-L1_D1D2 were collected at 30°C on either an 800 MHz Bruker system or a 950 MHz Bruker system. Typical acquisition times for 3D experiments of PD-L1_D1 were 8-9ms in ^{13}C (25ms for HNCO) 18-25ms in ^{15}N and 80ms in ^{1}H. Typical acquisition times for 3D experiments of PD-L1_D1D2 were 6ms in ^{13}C, 22ms in ^{15}N and 70ms in ^{1}H. Non-uniform sampling of 25% was conducted for the HNCACB experiment of PD-L1_D1 and HNCACB and HN(CO)CACB experiments for PD-L1_D1D2. The $^{15}N/^{1}H$ HSQC spectrum of PD-L1_D1 was recorded with acquisition times of 40ms in ^{15}N and 9ms in ^{1}H. The $^{15}N/^{1}H$ TROSY-HSQC spectrum of PD-L1_D1D2 was recorded with acquisition times of 40ms in ^{15}N and 80ms in ^{1}H (Pervushin *et al.* 1997). All NMR data was processed using NMRPipe software with linear prediction used to extend the effective acquisition time in ^{15}N by 2-fold in the CBCA(CO)NH and HNCO experiments for PD-L1_D1 and the HNCO for PD-L1_D1D2 (Delaglio *et al.* 1995). Spectra were analysed using the Sparky software package (Lee *et al.* 2015).

3.2.3 Chemical shift mapping of BPI binding to PD-L1_D1

To determine which residues of PD-L1_D1 were affected by the binding of the BPI ligand the minimal shift approach was used (Williamson *et al.* 1997). ^{15}N-HSQC spectra of free PD-L1_D1 and PD-L1_D1 in the presence of a 10% molar excess of the BPI were collected at 25°C on a 600 MHz Bruker spectrometer. NMR samples contained 100µM PD-L1_D1 in a 25mM potassium phosphate, 20mM sodium chloride, 10µM EDTA, 0.02% sodium azide (w/v), 5% D_2O buffer at pH 7.5 containing 1% DMSO (v/v) or 110µM BPI, 1% DMSO (v/v). The acquisition times were 50ms in ^{15}N and 90ms in ^{1}H.

The minimal combined amide proton and amide nitrogen chemical shift difference (Δδ) was determined for each backbone amide peak in Microsoft Excel using the following equation:

$\Delta\delta = \sqrt{(\Delta\delta_{HN})^2 + (0.2\,\Delta\delta_N)^2}$, where a scaling factor of 0.2 for the amide nitrogen chemical shift change accounts for the differences in the range of amide proton and amide nitrogen chemical shifts. The minimal shift upon BPI binding for each PD-L1_D1 backbone peak was taken as the lowest combined shift ($\Delta\delta$).

3.2.4 Investigation of the BPI ligand binding to PD-L1_D1D2

To investigate the binding of the BPI to PD-L1_D1D2 ^{15}N-TROSY-HSQC NMR experiments were performed in the presence of varying concentrations of BPI. All NMR samples contained 100µM PD-L1_D1D2 in 25mM potassium phosphate, 20mM sodium chloride, 10µM EDTA, 0.02% sodium azide (w/v), 5% D_2O buffer at pH 7.5. Titration experiments were carried out in the presence of 0, 20, 40, 60, 80, 100, 150 and 200µM of BPI. NMR acquisition times were 40ms in ^{15}N and 60ms in [1]H. All NMR data was processed in Topspin with the signal intensity for each peak also measured in Topspin. The decrease in signal intensity as the concentration of BPI increased was plotted in Prism, and non-linear regression curves for each signal were fitted using a one-site specific binding model.

3.3 Results
3.3.1 NMR backbone assignment of PD-L1_D1 and PD-L1_D1D2

Sequence specific backbone resonance assignment of PD-L1_D1 was completed to 90% using triple resonance HNCACB and CBCA(CO)NH spectra. Figure 3.1 shows the strip plot for a representative section of the amino acid sequence. The quality of the spectra is good with little overlap of the signals and inter-residue connectivities were visible for most residues. Those residues that are not assigned are due to the signals being absent from the spectra most likely due to them being flexible causing broadening of signals beyond detection. For example, the unassigned residues between C40 and K46 are a section of the BC loop, a long loop connecting β-strands across the sheets in the β-sandwich. This section is therefore likely flexible and exchanging with the solvent. Figure 3.2 shows a summary of the assigned and unassigned residues of PD-L1_D1.
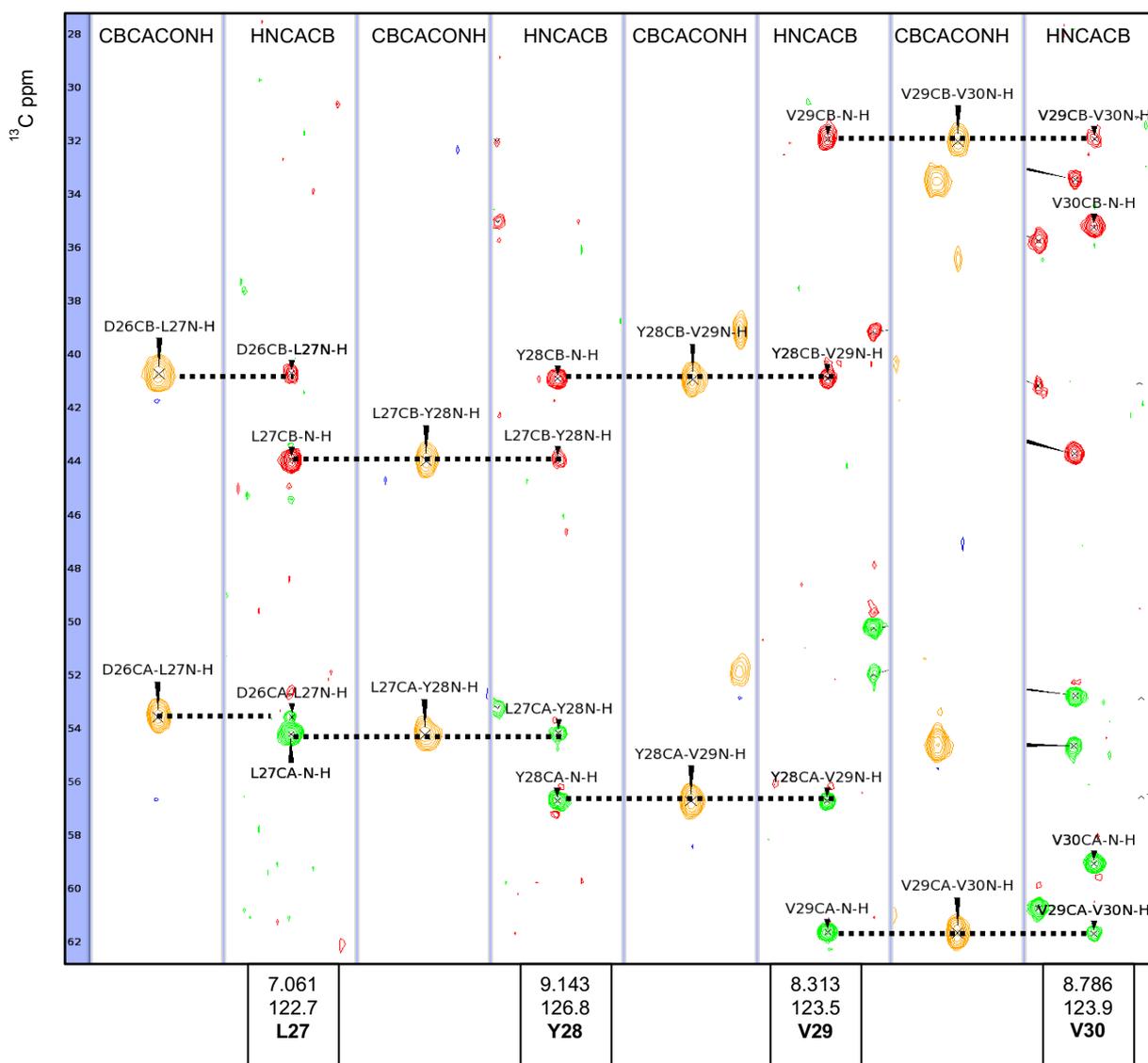
Figure 3.1. PD-L1_D1 sequential backbone assignments.
Strips taken at the ^{1}H,[15]N frequencies of L27 to V30 in the HNCACB and CBCACONH spectra of ^{15}N/^{13}C labelled PD-L1D_D1, showing the inter-residue connections linking the residues. The HNCACB spectrum shows Cα signals in green and Cβ signals are shown in red. Dashed black lines indicate sequential Cα and Cβ connections.

Figure 3.2. Backbone amide assignments of PD-L1_D1.

A) ^{15}N-HSQC spectrum of 100µM PD-L1_D1 at 25ºC collected on an 800 MHz Bruker spectrometer. Backbone amide assignments are labelled by one letter code and number. B) Cartoon representation of PD-L1_D1 with unassigned residues coloured in yellow.

Sequence specific backbone assignment of PD-L1_D1D2 was completed to 75% using triple resonance HNCACB and HN(CO)CACB spectra. Figure 3.3 shows the strip plot for a representative section of the amino acid sequence. The completeness of the assignments is much reduced for the membrane proximal domain with only 60% of the residues assigned. These missing residues consist of the BC and DE loops as well as the D β-strand (Figure 3.4). Given that these residues are unable to be assigned due to lack of signals in the 3D spectra it is likely that this section of the protein is dynamic. Examination of NOESY-HSQC data further commented on in Chapter 4 did not provide any additional assignments for those unassigned peaks in Figure 3.4 either due to there being no NOE cross-peaks or cross-peaks to other unassigned residues.
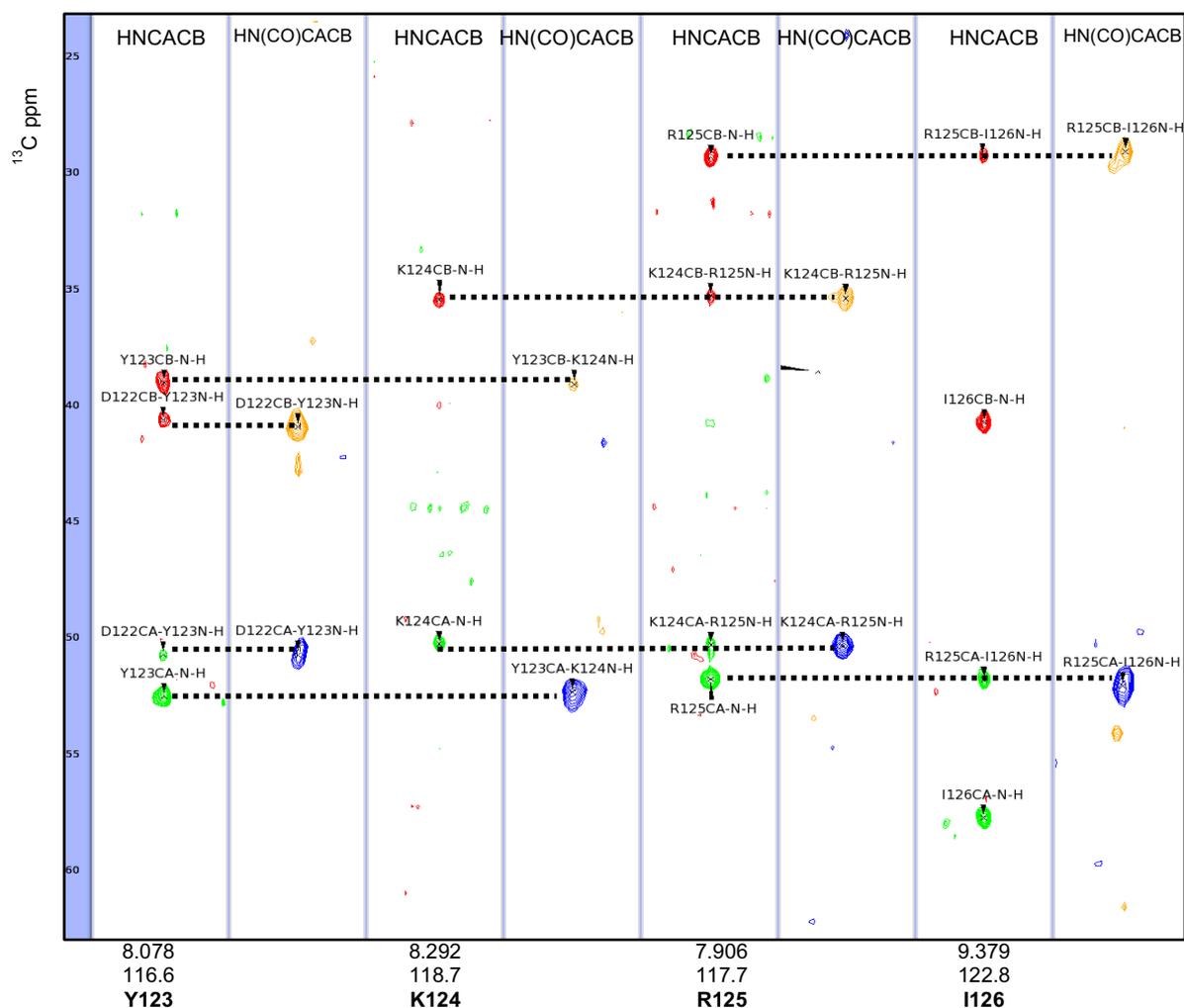
Figure 3.3 PD-L1_D1D2 sequential backbone assignments.

Strips taken at the ^{1}H,^{15}N frequencies of Y123 to I126 in the HNCACB and HN(CO)CACB spectra of ^{15}N/^{13}C/^{2}H labelled PD-L1D_D1D2, showing the inter-residue connections linking residues Y123 to I126. Cα signals are shown in green/blue and Cβ signals are shown in red/yellow. Dashed black lines indicate sequential Cα and Cβ connections.
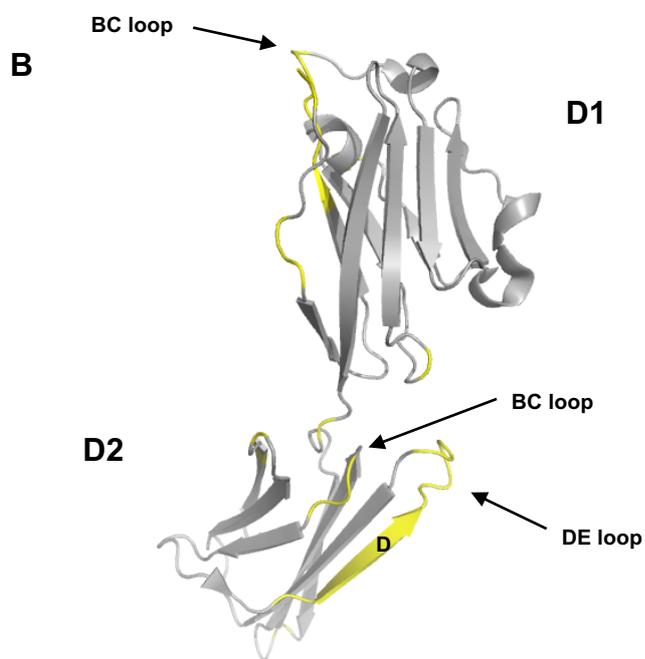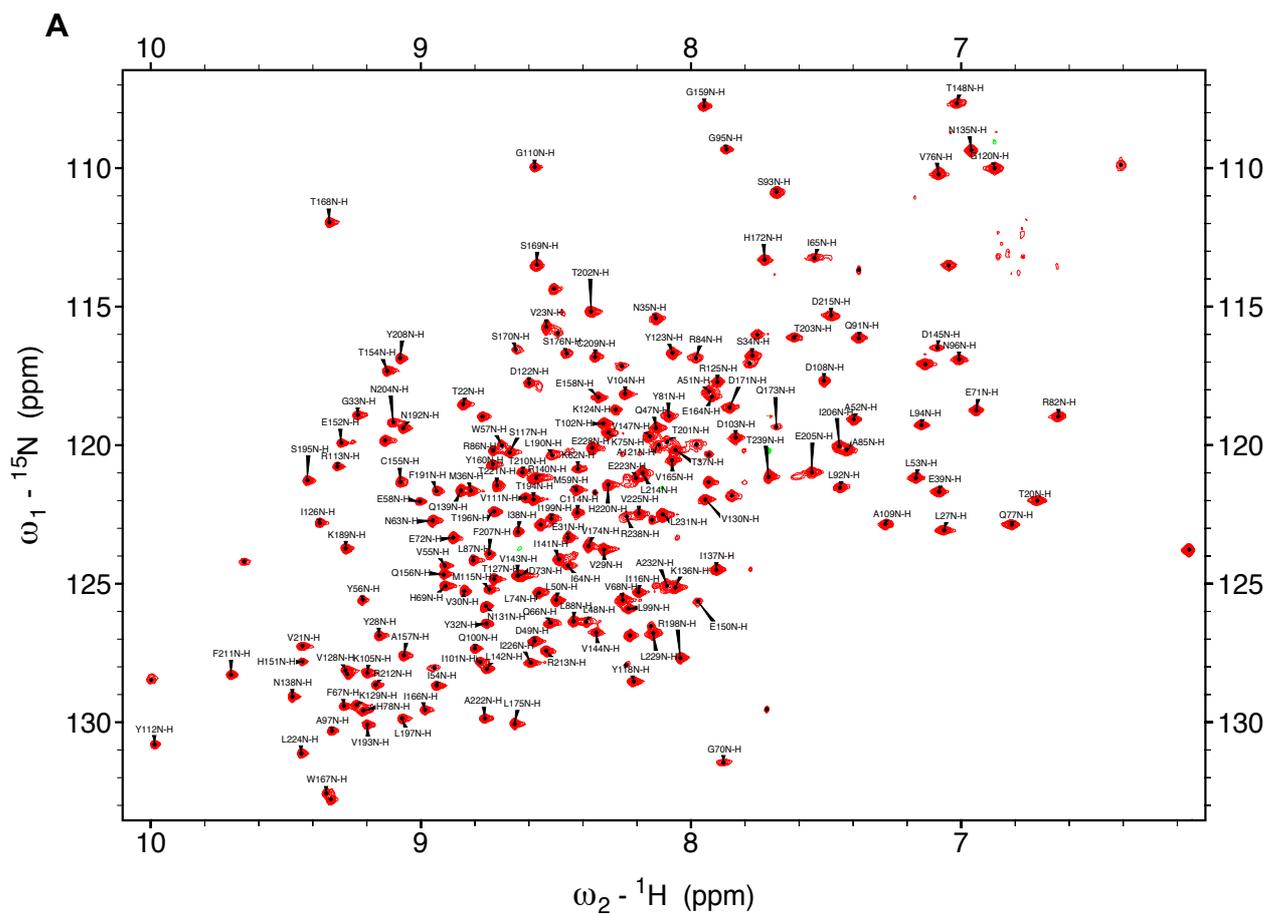
Figure 3.4. Backbone amide assignments of PD-L1_D1D2.

A) ^{15}N-TROSY-HSQC spectrum of 240µM PD-L1_D1D2 at 30ºC collected on an 800 MHz Bruker spectrometer. Backbone amide assignments are labelled by one letter code and number. **B)** Cartoon representation of PD-L1_D1D2 with unassigned residues coloured in yellow.

3.3.3 Investigation of BPI binding to PD-L1_D1

Investigation of the binding of the BPI to PD-L1_D1 was performed using backbone amide chemical shift perturbation analysis. Figure 3.5 shows an overlay of the ^{15}N-HSQC spectrum of PD-L1_D1 with the corresponding ^{15}N-HSQC spectrum of PD-L1_D1 in the presence of a 10% molar excess of BPI. In the presence of the BPI there are some peaks that stay in the same position and also a large number of peaks that show a change in chemical shift. This is indicative of specific binding of the BPI to PD-L1_D1 and, given the large number of shifts, a potential conformational change. Interestingly, there are also a subset of peaks that as well as shifting, appear to split into two distinct signals (Figure 3.5). Overlay of the ^{15}N-HSQC spectrum of PD-L1_D1 with the corresponding ^{15}N-HSQC spectrum of PD-L1_D1 in the presence of a 1:2 molar ratio of BPI:PD-L1_D1 shows that at this ratio there is almost complete saturation of PD-L1_D1, suggesting a binding event where one BPI molecule is binding to a PD-L1_D1 dimer.

Slow exchange creates a situation where as a ligand is titrated into a protein NMR sample, the intensity of the free state decreases and the intensity of the bound state increases representative of the concentration of free and bound protein. This is due to the off-rate being slower than the difference in Hz between the chemical shift of the free and bound state of the protein. In contrast, during fast exchange the ligand is coming on and off the protein of interest many times over the NMR experiment so that the chemical shift of the signal is the weighted average of free and bound shifts (Williamson 2013). This results in individual signals moving gradually from the free chemical shift to the bound chemical shift as the ligand is titrated in. Overlaying of free and BPI bound spectra of PD-L1_D1 shows slow exchange behaviour which is in fitting with an IC_{50} value of 20nM determined by collaborators using Biacore. Due to the system being in slow exchange this means peaks that are shifting are not easily followed as the ligand concentration is increased, as is the case during fast exchange. For this reason, minimal shift analysis was used for chemical shift mapping. This is where the chemical shift change from peaks in the free spectrum is determined from the peak in the bound spectrum that is nearest.
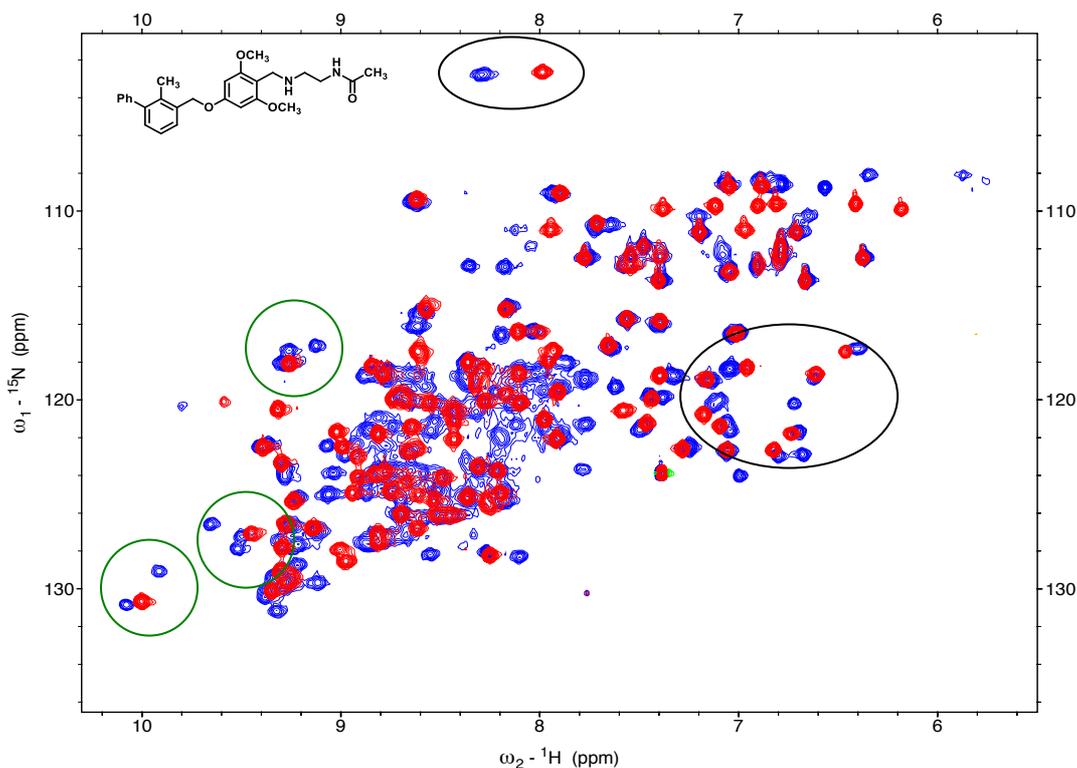
Figure 3.5. BPI binding to PD-L1_D1.

An overlay of the ^{15}N-HSQC spectra for free PD-L1_D1 (red) and PD-L1_D1 in the presence of a 10% molar excess of BPI (blue). The structure of the BPI is shown in the top left corner. Examples of signals that split into two are circled in green. Examples of shifts are circled.

Figure 3.6 shows a bar chart representation of the minimal chemical shift changes observed for PD-L1_D1 upon binding of the BPI ligand. These shift changes have been mapped onto a crystal structure of PD-L1_D1 with white representing no significant shift, minimal shifts between 0.02 and 0.06 represented on a scale of white to red and large shifts over 0.06 represented in red (Figure 3.6). The mapping of chemical shift changes onto PD-L1_D1 shows a patch of large chemical shift changes at the PD-1 binding face involving the CC'FG β-sheet. This patch of chemical shift changes covers 4120 Å^2 and is significantly larger than would be expected for the binding of a 500 Dalton molecule such as is binding to PD-L1_D1 in this instance and more consistent with a protein-protein interface. Figure 3.6 shows the structure of PD-L1_D1 with the peaks that have split into two highlighted. This shows a smaller patch of residues more consistent with the direct binding site of the BPI. Since one molecule of the BPI binds to a PD-L1_D1 dimer it is therefore likely that the larger patch of chemical shift changes represents a PD-L1_D1 dimer interface.

A

B

Size of shift perturbation
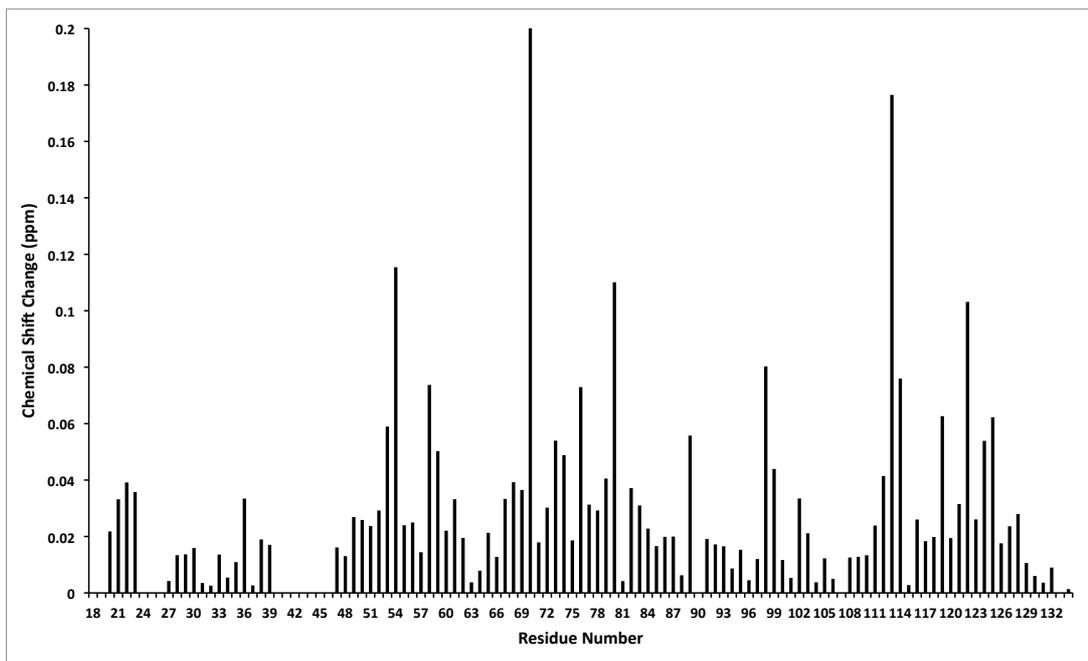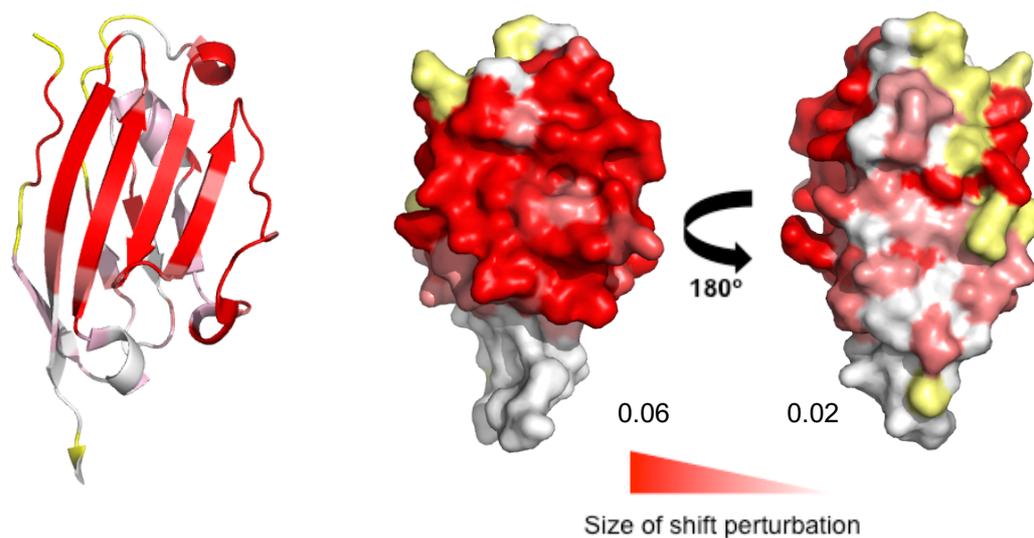
C
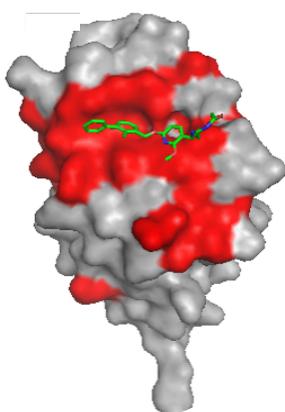
Figure 3.6. Chemical shift mapping of the binding of the BPI to PD-L1_D1.

A) A bar chart representation of the combined ^{15}N and ^{1}H chemical shift changes of PD-L1_D1 residues upon binding of the BPI. **B)** The mapping of chemical shift changes onto the structure of PD-L1_D1. Shifts over 0.06 are shown in red and shifts less than 0.02 are shown in white. Residues for which there is no backbone amide assignment are shown in yellow. **C)** Peaks that split into two signas upon ligand binding are shown in red on the structure of PD-L1_D1.

There are also a number of smaller chemical shift changes involving residues on the opposite face to PD-1 binding on PD-L1_D1 (Figure 3.6). The crystal structure of the full-length extracellular region of PD-L1 shows a PD-L1 dimer interface involving some of the same residues that show these smaller chemical shift changes. This dimer interface is made up of a number of polar and ionic interactions as shown in Figure 3.7, but has not been shown to have any physiological significance (Chen *et al*. 2010, PDB: 3FN3). Analytical size exclusion chromatography was used to determine whether PD-L1_D1 was a dimer in solution, over a range of concentrations, under NMR buffer conditions. Figure 3.8 shows that two peaks elute from the column under these conditions that correspond to a 26kDa dimer and a 13kDa monomer after comparison with calibration standards. SDS-PAGE analysis confirms that the presence of dimers is not due to disulphide-linked dimers. It appears that under NMR conditions PD-L1_D1 is a mixture of both monomer and dimer. It could be that at the low NaCl concentrations used for NMR that PD-L1_D1 has a propensity to form the ionic interactions seen in Figure 3.7.
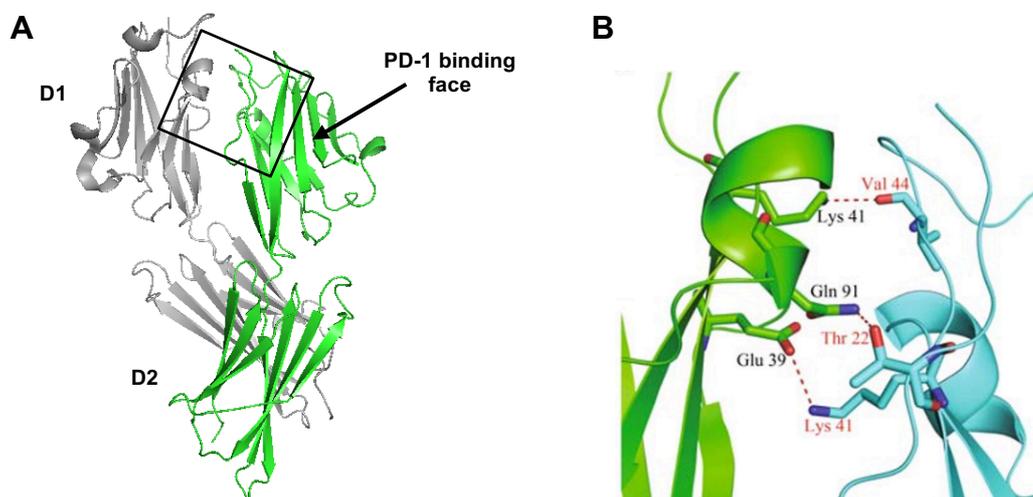
Figure 3.7. Dimeric PD-L1.

A) PDB:3FN3. Crystal structure of dimeric PD-L1_D1D2. **B)** Zoom in on the dimeric interface between D1 shown in the black box. Taken from Chen, 2010.

A
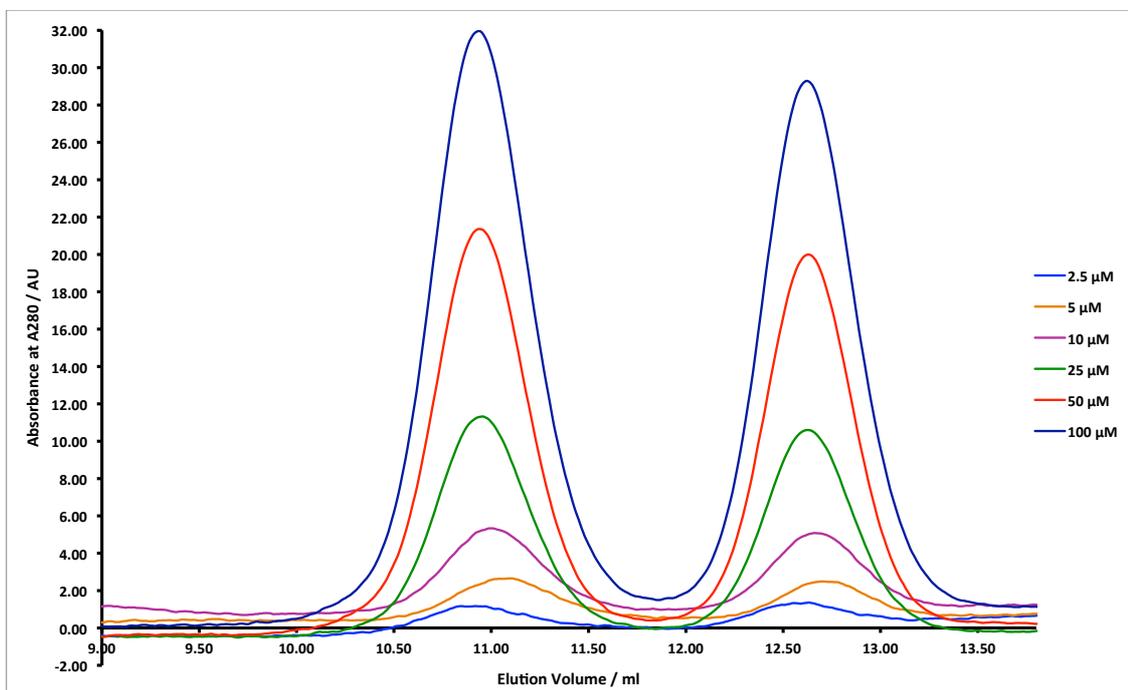
B
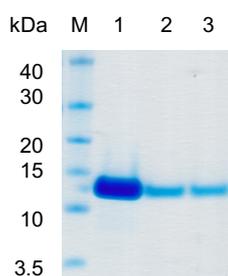
Figure 3.8. Investigation of dimerization of PD-L1_D1 in solution.

A) Size exclusion chromatogram of PD-L1_D1 at different concentrations. **B)** Non-reduced SDS-PAGE analysis of the eluted fractions from size exclusion chromatography of a 50µM sample of PD-L1_D1. Lane 1 contains the column load. Lane 2 contains a sample from the first peak to elute at 10.9ml. Lane 3 contains a sample from the second peak to elute at 12.6ml.

3.3.4 Investigation of the BPI binding to PD-L1_D1D2

Since at the cell membrane PD-L1 consists of two domains, the binding of the BPI to the full extracellular region of PD-L1 (PD-L1_D1D2) was also investigated using NMR. PD-L1_D1D2 was titrated with BPI at 1:0, 1:0.2, 1:0.4, 1:0.6, 1:0.8, 1:1 and 1:2 ratios of protein to ligand. Figure 3.9 shows the ^{15}N-TROSY-HSQC of free PD-L1_D1D2 overlaid with the ^{15}N-TROSY-HSQC of PD-L1_D1D2 in the presence of BPI at a 1:1 molar ratio. When the BPI is added to PD-L1_D1D2 there are no chemical shift changes in the NMR spectra but there is significant reduction in signal intensity of all signals. Figure 3.9 shows a subset of signals that depict how as the concentration of ligand is increased the intensity of signals decreases. It became apparent that PD-L1_D1D2 precipitated upon addition of the BPI, therefore the decrease in signal intensity in the NMR is due to reduced amounts of free PD-L1_D1D2 in solution. The precipitation of PD-L1_D1D2 upon addition of the ligand is most likely due to binding of the ligand since precipitation does not occur when DMSO is added alone and addition of the ligand to NMR samples does not significantly change the pH of the sample. The drop in free PD-L1_D1D2 signal intensity was plotted against concentration of the BPI and curves were fitted for each signal in Prism using a one-site specific binding model (Figure 3.10). The mean Kd value determined from these curves was 58 +/-6µM. The binding affinity of this BPI to PD-L1_D1 as determined using SPR by collaborators is 20nm. The data presented in this chapter clearly shows that the affinity of the BPI for PD-L1_D1D2 is an order of magnitude less than the reported binding affinity to PD-L1_D1.
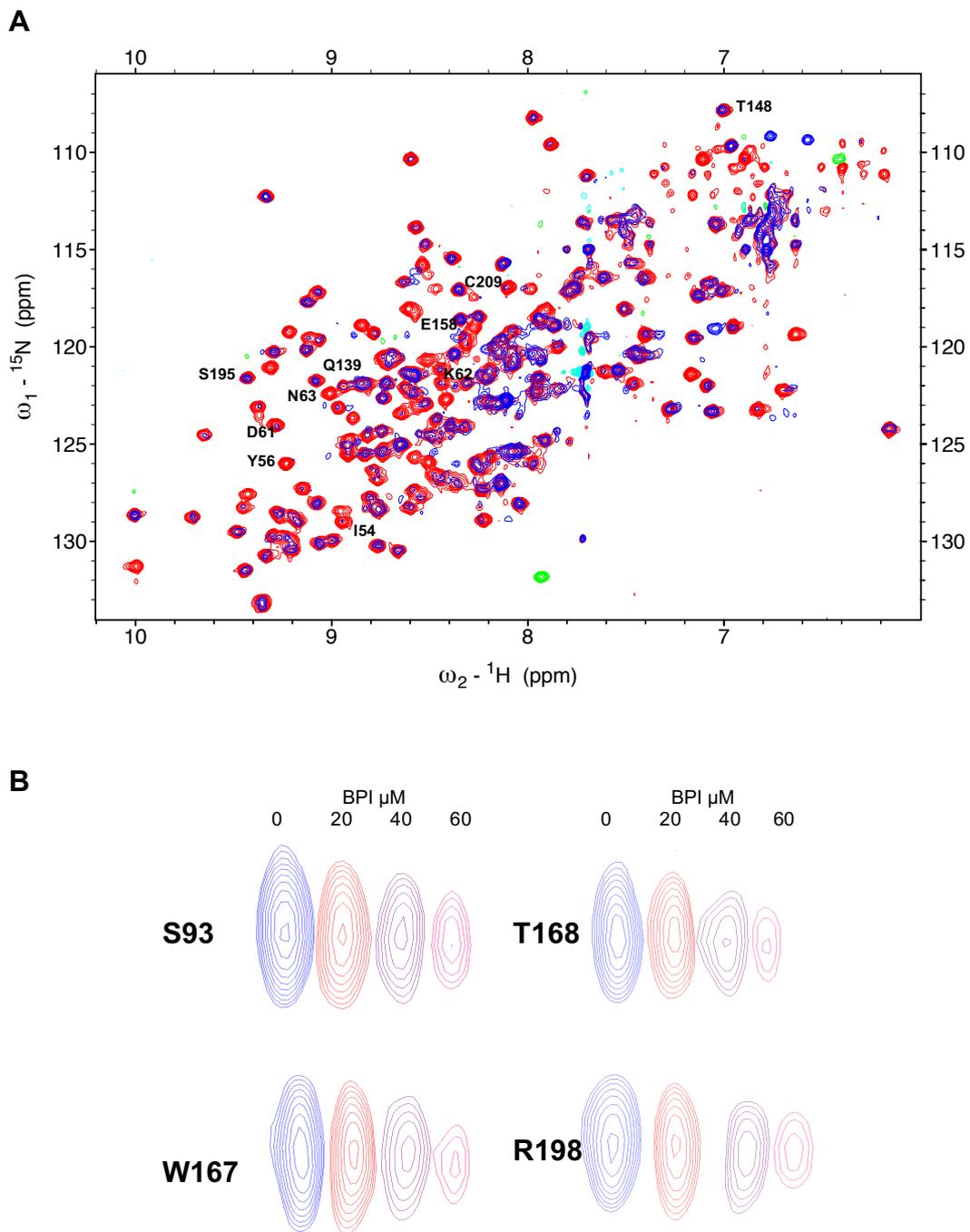
Figure 3.9. BPI binding to PD-L1_D1D2.

A) An overlay of the ^{15}N-TROSY-HSQC of 100μM free PD-L1_D1D2 (shown in red) and the ^{15}N-TROSY-HSQC of PD-L1_D1D2 in the presence of a 1:1 ratio of BPI (shown in blue) collected on an 800MHz spectrometer. **B)** A subset of signals showing the decrease in intensity as the concentration of BPI is increased.
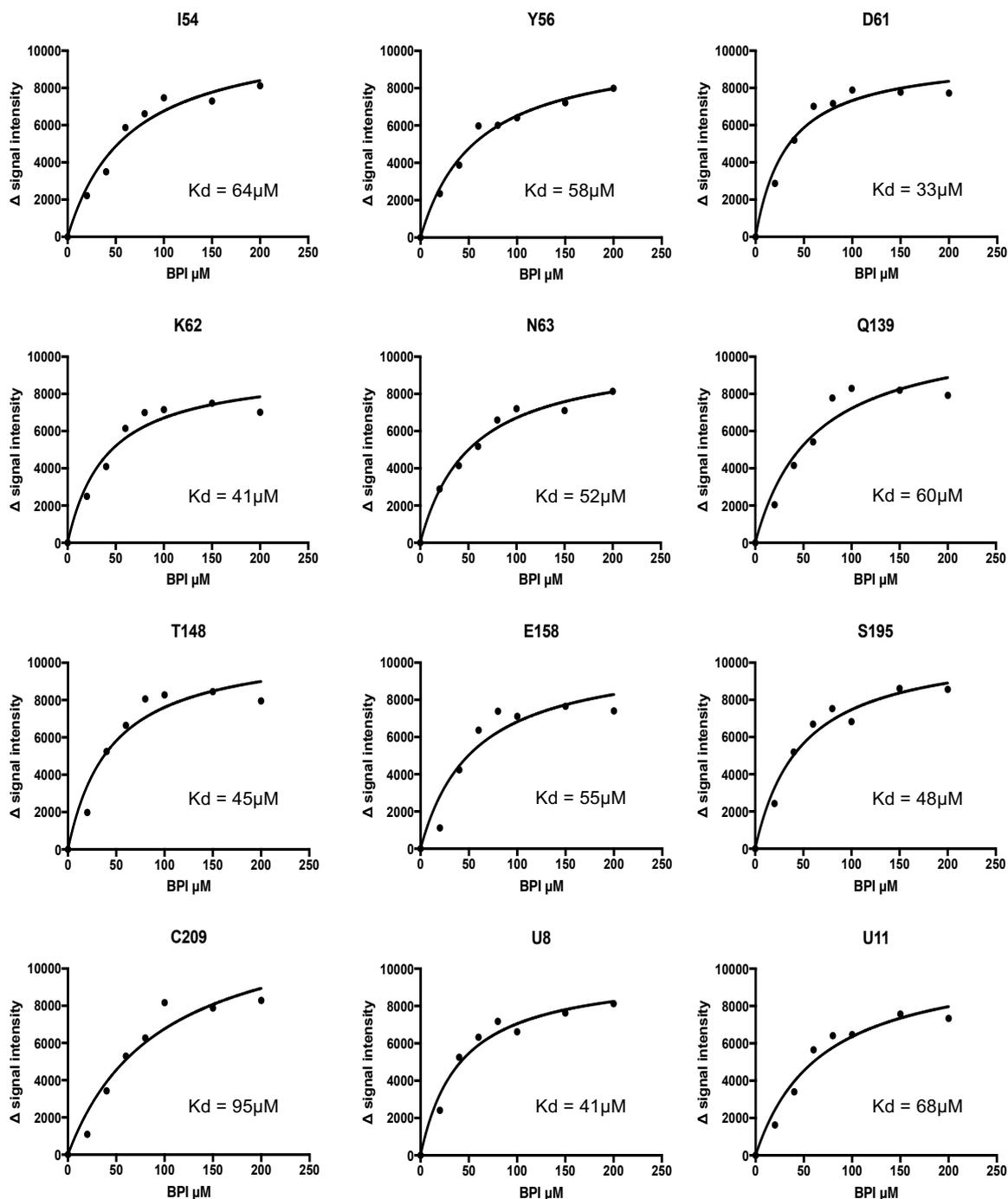
Figure 3.10. Binding curves of BPI binding to PD-L1_D1D2.

A representative sample of the binding curves produced from the plotting the decrease in NMR signal intensity of PD-L1_D1D2 signals upon the titration of BPI (Figure 3.9). Unassigned resonances are labelled U. Curves were fitted in Prism using a one-site specific binding model. $Y = Bmax*X/(Kd + X)$.

3.4 Discussion

Sequence specific backbone resonance assignments for PD-L1_D1 and PD-L1_D1D2 were obtained with 90% completion for PD-L1_D1 and 75% completion for PD-L1_D1D2. The backbone assignment of PD-L1_D1 allowed the chemical shift mapping of the binding of a BPI to PD-L1_D1. NMR chemical shift perturbations upon addition of the BPI to PD-L1_D1 show that the BPI binds to the PD-1 binding face of PD-L1. Almost complete saturation of PD-L1_D1, as determined by NMR, with half the molar amount of BPI suggests the BPI binds to PD-L1_D1 with a 1:2 stoichiometry. This finding was later confirmed by the publication of a crystal structure of PD-L1_D1 bound to another similar BPI from the same series of Bristol-Myers Squibb patented molecules (Zak *et al.* 2016, PDB: 5J89). The latter showed the BPI sandwiched between two PD-L1_D1 molecules via their PD-1 binding faces with residues T20, Y56, M115, A121, D122, Y123, K124 making important hydrophobic and electrostatic interactions (Figure 3.11). These residues identified in the crystal structure align well with the signals that split into two peaks upon binding in the NMR studies presented here. The peak splitting seen is therefore due to the individual PD-L1_D1 monomers experiencing different chemical environments as they bind to different sides of the asymmetric BPI.
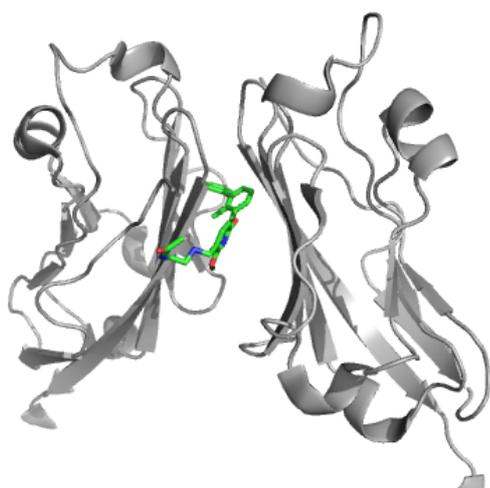
Figure 3.11. BMS compound cause PD-L1_D1 dimerisation.

This shows the crystal structure of a BMS compound (green) bound to PD-L1_D1 (grey). PDB: 5J89. One BMS compound is bound to a dimer of PD-L1_D1 formed via the GFCC' β-sheet face.

There are also a number of chemical shift changes induced by the binding of the BPI on the opposite face to where it binds (Figure 3.6).

It is possible that BPI-induced chemical shift perturbations on the opposite face to binding are due to changes in the polar dimer interface shown in Figure 3.7. The two dimer forms are shown in Figure 3.12. In this scenario free PD-L1_D1 exists as a dimer in solution, evidenced by analytical gel filtration (Figure 3.7), via the interface shown in Figure 3.7. Ligand binding at the PD-1 binding face would cause disruption of this ionic interface and PD-L1_D1 + BPI forms a dimer via the PD-1 binding face (Figure 3.12).

A PD-L1 dimer via ABED β-sheet

B BPI-induced PD-L1 dimer via PD-1 binding face
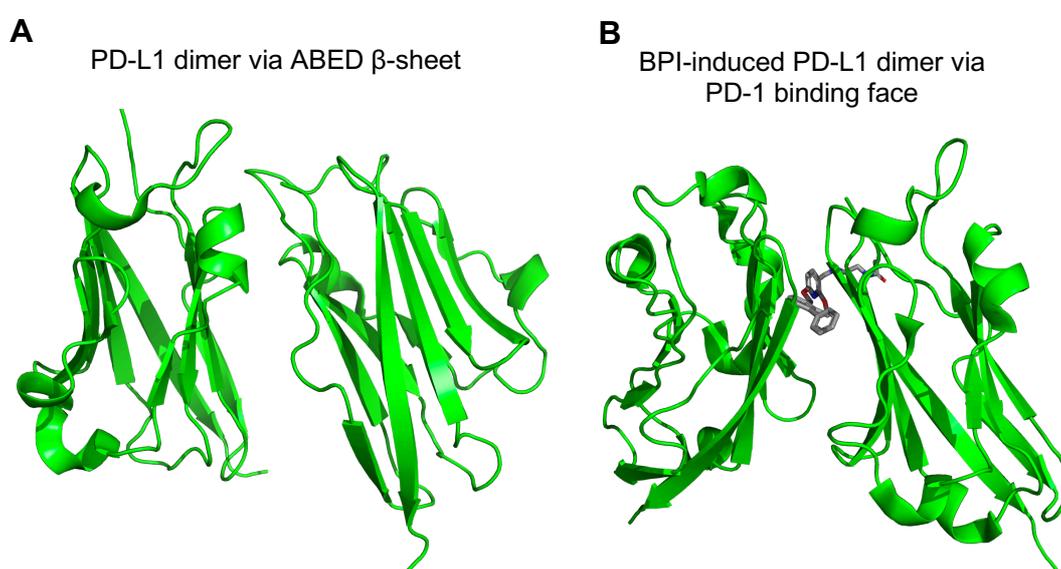
Figure 3.12. PDL1_D1 dimeric forms.
A) Dimeric crystal structure of free PD-L1 (PDB: 3FN3) via polar interactions of the ABED β-sheet leaving the PD-1 binding face exposed. **B)** Crystal structure of PD-L1_D1 bound to a BPI ligand (PDB: 5J89). Ligand binding causes dimerization of PD-L1 occluding the PD-1 binding site.

The NMR studies presented here characterising the binding of a BPI to PD-L1_D1D2 show that BPI binding to PD-L1_D1D2 does not occur in a similar way as binding to PD-L1_D1. Precipitation of PD-L1_D1D2 upon binding of the ligand, under the NMR conditions used, has allowed an estimate of the binding affinity of a BPI to PDL1_D1D2 via the measurement of loss of NMR signal intensity. The data yield a binding affinity of 58+/-6µM for the BPI binding to PD-L1_D1D2. This is an order of magnitude less than the affinity

of the BPI to PD-L1_D1 determined as 20nM by collaborators. The binding affinity of PD-L1_D1D2 to PD-1 is 8μM (Cheng *et al.* 2013). Taken together it is likely that whilst these compounds perform well in *in vitro* assays using PD-L1_D1, they do not work in cell-based assays because they have a much-reduced affinity for PD-L1_D1D2. The affinity of the compounds for PD-L1_D1D2 is most likely not tight enough, when compared to the binding affinity of PD-1 for PD-L1_D1D2, to compete with PD-1 for binding to PD-L1_D1D2 on the cell surface. It is likely that the presence of the membrane-proximal Ig-like domain of PD-L1_D1D2 prevents the BPI-induced dimer formation seen with PD-L1_D1. The following chapter explores this further and aims to better understand the structural features of PD-L1_D1D2 in solution.

Chapter 4

Investigation of structural features of the extracellular region of PD-L1

4.1 Introduction

As detailed in Chapter 1, interaction of PD-L1 and PD-1 at the cell surface, triggers intracellular signalling within T-cells via recruitment of SHP1 and SHP2 phosphatases to tyrosine-phosphorylated ITIM and ITSM motifs (Sheppard *et al.* 2004). The specific mechanism of how signalling is transduced from the cell surface remains unknown. However, other B7-family members signal via dimerization (Schwartz *et al.* 2001). This, taken together with the data presented in Chapter 3 that shows how the presence of the membrane-proximal Ig-like domain (D2) of PD-L1 reduces the affinity of the BPI compounds to PD-L1, led to the investigation of structural features of PD-L1 in solution. PD-L1 has previously been shown to interact with binding partners other than PD-1 and PD-L1 has also been shown to participate in *cis*-binding with molecules on the surface of the same cell (Zhao *et al*. 2018; Chaudri *et al*. 2018). The hope is that better understanding of the structural features of the extracellular region of PD-L1 in solution could shed light on the PD-1 signalling mechanism, how PD-L1 interacts with other binding partners, as well as why the BPI compounds do not bind to PD-L1_D1D2 with a similar affinity as for PD-L1_D1. This chapter details the investigation of PD-1 binding to PD-L1 as well as how NMR was used to investigate the structural features of the extracellular region of PD-L1. In particular the position of the two PD-L1 domains in relation to each other, and the possibility of PD-L1 sampling different conformations is explored.

4.1.1 PD-1/PD-L1 in cis and in trans binding

PD-L1 has been shown to interact with multiple binding partners including B7-1 (CD80), suggesting there is a degree of plasticity within the PD-L1 molecule. The binding of B7-1 to PD-L1 has been shown to be *in cis*, between molecules on the surface of the same cell (Chaudri *et al.* 2018). Also, in addition to PD-L1

binding PD-1 across membranes, PD-L1 has been found to bind PD-1 *in cis*. PD-L1 binding PD-1 in *cis* has been shown to be inhibitory of PD-L1 binding PD-1 in *trans*, having a threshold modulation effect on T-cell activation (Zhao *et al.* 2018).

Cis/trans ligand-receptor interactions, as described above, have been described for a number of different systems including MHC-I/Ly49. Specifically, mouse Ly49A is an inhibitory receptor for a number of different MHC-I molecules (Karlhofer *et al.* 1992). This interaction is involved in Natural Killer (NK) cell "missing-self recognition" where when MHC-I is low, inhibitory signalling through Ly49 does not occur, leading to NK cell activation (Ljunggren and Karre 1990). Ly49A has also been found to interact *in cis* with MHC-I molecules expressed on the NK cell via the same binding site as binding across membranes (Doucey *et al.* 2004). This has been shown to serve a threshold modulation effect by reducing the availability of Ly49A to bind MHC-I *in trans*, effectively lowering the point at which NK cell activating signalling causes activation (Back *et al.* 2007). The structural basis for how Ly49 binds MHC-I both *in cis* and *in trans* has been determined using x-ray crystallography. Ly49 receptors are homodimeric and made up of a ligand binding domain (NKD) and a long stalk region consisting of approximately 70 amino acids. A crystal structure of Ly49L with the stalk region showed that the NKD was folded back onto the α^3S helix of the stalk (Back *et al.* 2009). Comparison of the crystal structure of Ly49L NKD homodimer alone showed that without the stalk region the NKD subunits are in a different relative orientation with the NKD subunits interacting via their α2 helices (Back *et al*. 2009). In this situation the NKD subunits would not fold back onto the stalk and would be in an extended conformation. In addition to this cell-cell adhesion assays using Ly49A with stalk region deletions showed that deletion of the α^3S helix significantly reduced *trans* binding whereas deletion of α^1S or α^2S helices did not. It has also been determined by co-immunoprecipitation that association of Ly49A, with α^3S deletions, and MHC-I occurs *in cis*, thus α^3S is required for *trans* interactions of Ly49 but not *cis* interactions (Back *et al.* 2009).
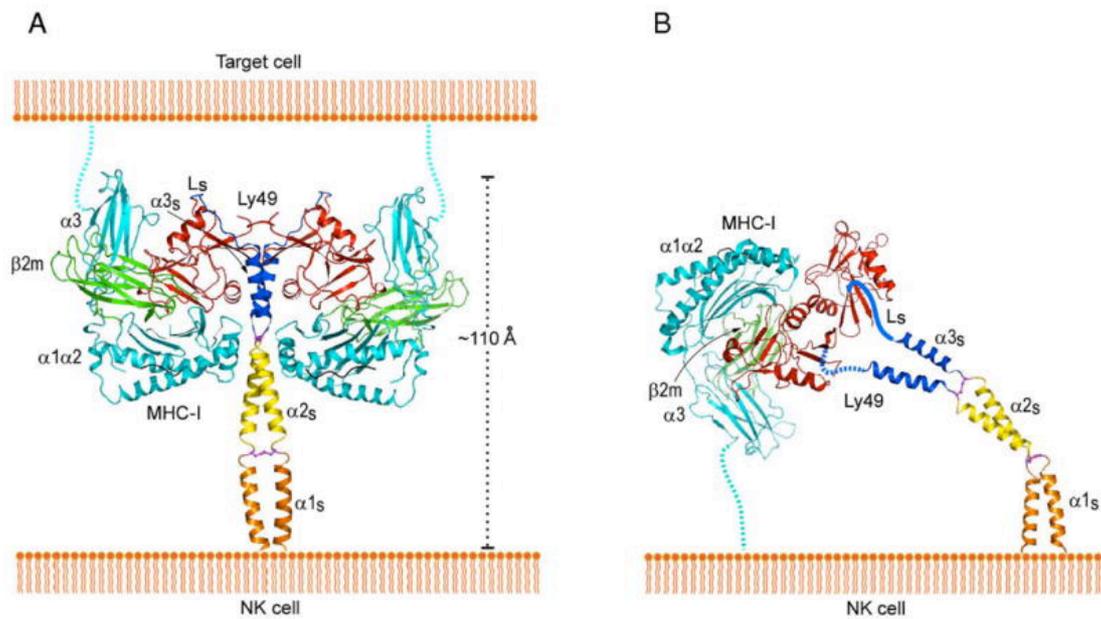
Figure 4.1. A model of the *cis* and *trans* interactions of Ly49. Taken from Back, 2009.
A) shows that proposed model of Ly49 interacting with MHC-I *in trans*. Here the NKD domains are coloured in red and are bent back onto the stalk region α^3S helices (blue). **B)** shows the model of Ly49 binding MHC-I *in cis*. The Ly49 receptor is in an extended conformation in order to orientate the NKD domains such that they can bind to MHC-I on the same membrane. MHC-I is coloured in cyan.

Figure 4.1 shows a proposed model for *cis* and *trans* interactions of Ly49. *Trans* interaction involves the bending back of the NKD domains onto the α^3S helix in the stalk such that a Ly49 homodimer on an NK cell can symmetrically bind two MHC-I molecules on the target cell. In order for a Ly49 receptor to bind an MHC-I molecule on the same NK cell the orientation of the NKD domains would need to be reversed. In this case an extended conformation of Ly49 where the NKDs are disengaged from the stalk allows the *cis* interaction of a Ly49 homodimer with a single MHC-I molecule (Figure 4.1). This example shows a precedent for interactions between domains of immune molecules serving to regulate function. The threshold modulation mechanism seen in this example could also serve as a regulatory mechanism for the PD-1/PD-L1 interaction where PD-1 binding to PD-L1 *in cis* reduces the inhibitory threshold for PD-1/PD-L1 signalling.

73

PD-1 and PD-L1 are immunoglobin-like molecules. There are also immunoglobin-like molecules that have been shown to interact *in cis*. Human LILRB proteins are also receptors for MHC-I that have been shown to bind to MHC-I on the same cell (Masuda *et al.* 2007). In this case the extracellular domains of LILRB receptors are composed of four or six Ig-like domains with the two most membrane-distal domains involved in MHC-I binding. It has been proposed that to reverse the orientation of these domains, such that they can bind in *cis*, the Ig-like domains are required to bend back on one another via flexible hinges between the domains (Held and Mariuzza 2011). There is a precedent for this with the four Ig-like domains of the *Drosophilia* protein Dscam taking on a horseshoe shape via a flexible hinge between Domains 2 and 3 (Meijers *et al.* 2007) (Figure 4.2). The flexible hinge has the amino acid sequence TEPISS. The linker between D1 and D2 of PD-L1 is also made from 6 amino acids containing 1 proline residue with the sequence APYNKI. It is possible then that this PD-L1 linker could also be a flexible hinge.
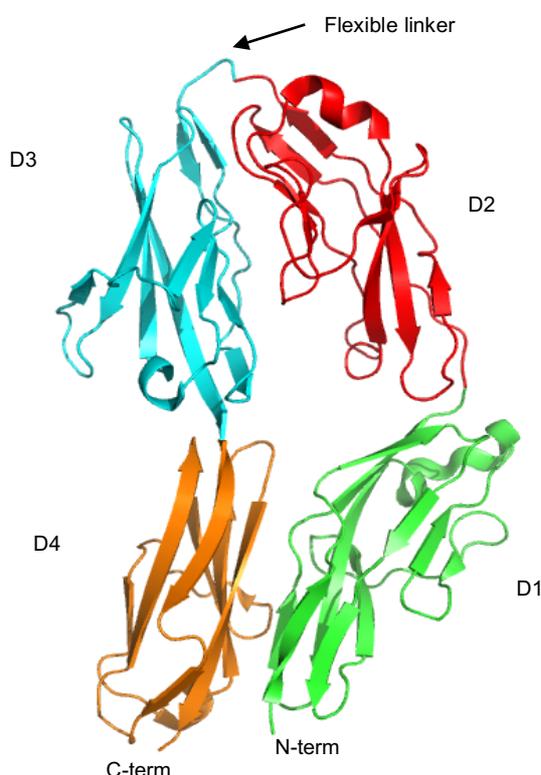
Figure 4.2. Crystal structure of *Drosophilia* Dscam. PDB: 2V5M

Cartoon representation of the crystal structure of *Drosophilia* protein Dscam taking on a horseshoe shape via a flexible hinge between Domains 2 and 3.

4.2 Methods

4.2.1 Size Exclusion Chromatography-Multi-angle Light Scattering

Size exclusion chromatography followed by multi-angle light scattering was performed with 130 µM PD-L1_D1D2 and 120 µM PD-1 samples in 25 mM potassium phosphate pH 7.5, 100 mM sodium chloride buffer conditions. SEC was performed using a Superdex S75 analytical grade column equilibrated with 25 mM potassium phosphate pH 7.5, 100 mM sodium chloride. Light scattering analysis coupled to Refractive Index (RI) detection of the eluent was performed using a Dawn 8+ light scattering instrument and Optilab T-rEX RI detector. Light scattering data was analysed using the ASTRA software package.

4.2.2 NMR spectroscopy: PD-1 binding to PD-L1

To determine which residues of PD-L1 were affected by the binding of PD-1, the minimal shift approach was used (Williamson *et al*. 1997). ^{15}N/^{1}H TROSY spectra of free PD-L1_D1D2 at 110µM and PD-L1_D1D2 in the presence of a 20% molar excess of PD-1 at a final concentration of 135µM were collected at 30°C on an 800 MHz Bruker spectrometer equipped with a cryoprobe. The acquisition time was 40 ms in ^{15}N and 60 ms in ^{1}H, with 32 scans. In addition, ^{15}N/^{1}H HSQC spectra of free PD-L1_D1 at 120µM and PD-L1_D1 in the presence of a 50% molar excess of PD-1 at a final concentration of 180µM were collected at 25°C on a 600 MHZ Bruker spectrometer. The acquisition time was 50 ms in ^{15}N and 90 ms in ^{1}H, with 32 scans.

The minimal combined amide proton and amide nitrogen chemical shift difference (Δδ) was determined for each backbone amide peak in Microsoft Excel using the following equation:

$\Delta\delta = \sqrt{(\Delta\delta_{HN})^2 + (0.2\,\Delta\delta_N)^2}$ where a scaling factor of 0.2 for the amide nitrogen chemical shift change accounts for the differences in range of amide proton and amide nitrogen chemical shifts. The minimal shift upon PD-1 binding for each PD-L1_D1D2 and PD-L1_D1 backbone amide peak was taken as the lowest combined shift (Δδ).

4.2.3 NMR spectroscopy: Sequence specific assignments of PD-L1_D2

^{13}C, ^{15}N-labelled PD-L1_D2 was expressed in BL21(DE3) E. coli grown in minimal media with ^{15}N-NH$_4$Cl as the sole nitrogen source and ^{13}C-glucose as the sole carbon source. ^{13}C, ^{15}N-labelled PD-L1_D2 was refolded and purified as described in Chapter 2.

Sequence specific backbone resonance assignments (N, NH, Cα, Cβ, C') were obtained for PD-L1_D2 using intra- and inter-residue connectivities where amide NH to Cα and Cβ peaks were identified in triple resonance spectra and used to search for i-1 or i+1 residues in the sequence. Assignments for PD-L1_D2 were determined using ^{15}N, ^{13}C, ^{1}H HNCACB, CBCA(CO)NH, HNCA, HN(CO)CA (Wittekind and Mueller 1993; Kay *et al.* 1990). NMR spectra were collected from 220µM samples of PD-L1_D2 in a 25mM potassium phosphate, 20mM sodium chloride, 10µM EDTA, 0.02% sodium azide (w/v) buffer at pH 7.5, containing 5% D$_2$O. NMR experiments of PD-L1_D1 were collected at 25°C on an 800 MHz Bruker system. Typical acquisition times for 3D experiments of PD-L1_D2 were 6-9 ms in ^{13}C, 21-24 ms in ^{15}N and 80 ms in ^{1}H. Non-uniform sampling was used to 25% for the HNCA collection. The ^{15}N/^{1}H HSQC spectrum of PD-L1_D1 was recorded with acquisition times of 50 ms in ^{15}N and 90 ms in ^{1}H. All NMR data was processed using NMRPipe software (Delaglio F *et al.* 1995). Linear prediction was used to extend the effective acquisition time in ^{15}N by 2-fold in the CBCA(CO)NH experiment. Spectra were analysed using the Sparky software package (Lee *et al.* 2015).

4.2.4 NMR spectroscopy: Chemical shift analysis of PD-L1 domains

To investigate whether there was a difference between the individual domains of the extracellular region of PD-L1 depending on whether they were stand-alone domains or as one polypeptide chain, the 2D ^{15}N/^{1}H spectra were compared. ^{15}N/^{1}H HSQC spectra for PD-L1_D1, PD-L1_D2, and PD-L1_D1D2 were collected as described previously. Using Sparky, the spectra for PD-L1_D1 and PD-L1_D2 were referenced to PD-L1_D1D2 to account for chemical shift differences caused by the use of different temperatures during collection. Where residues were assigned in both the individual domain constructs and PD-L1_D1D2, the actual chemical shift change was measured. For all other

residues a minimal shift approach was used, as described above, to determine residues that were shifted in PD-L1_D1D2 when compared to the single domain spectra of PD-L1_D1 and PD-L1_D2.

4.2.5 NMR spectroscopy: Long range backbone NOEs of PD-L1_D1D2

Long range backbone amide to backbone amide NOEs in PD-L1_D1D2 were searched for with the aim of orientating the two domains in relation to each other. ^{15}N/ ^{13}C/ ^{2}H-labelled PD-L1_D1D2 was expressed and purified as described previously. A ^{15}N/^{1}H NOESY-HSQC (Zuiderweg & Fesik 1989; Marion et al. 1989) experiment for PD-L1_D1D2 was obtained from a 200μM sample in a 25mM potassium phosphate, 20mM sodium chloride, 10μM EDTA, 0.02% sodium azide (w/v) buffer at pH 7.5, containing 5% D_2O. The ^{15}N/^{1}H NOESY-HSQC experiment was carried out at 30°C on a 950 MHz Bruker spectrometer. Acquisition times were 17 ms in ^{15}N, 90 ms in direct ^{1}H and 17 ms indirect ^{1}H. A mixing time of 500 ms was used to search for ^{1}H-^{1}H NOEs corresponding to distances of up to 8Å. Non-uniform sampling was used to 50%. All NMR data was processed using NMRPipe software. Linear prediction was used to extend the effective acquisition time in ^{15}N by 2-fold. Spectra were analysed using the Sparky software package.

4.2.6 HADDOCK docking of PD-L1 domain interface

Modelling of a potential D1/D2 interface was performed using HADDOCK 2.2 based on chemical shift perturbation data (de Vries *et al*. 2010; van Zundert *et al*. 2015). HADDOCK requires the definition of active and passive residues which are then used as Ambiguous Interaction Restraints to drive docking. Residues in PD-L1_D1 and PD-L1_D2 that had a chemical shift changes over 0.02 ppm were chosen as active residues and residues contacting the active residues at the surface of the protein were chosen as passive residues. These are summarised in Table 4.1 and Figure 4.3.

Table 4.1. Active and Passive residues for HADDOCK

	Active Residues	Passive Residues
PD-L1_D1	Y32, G33, S34, M36, S79, S80, R82, K105, L106, D108, G110, K129, V130, N131, A132	L27, V29, V30, E31, N35, T37, H78, Y81, Q83, R84, D103, V104, Q107, A109, V111, T127, V128
PD-L1_D2	Y134, K136, I137, Q156, E158, G159, Y160, K189, L190, F191, N192, R212, L214, D215, E217	P133, N135, N138, A157, P161, K162, A163, V165, T182, N183, K185, R186, E187, E188, T194, F211, R213, P216, E218, N219
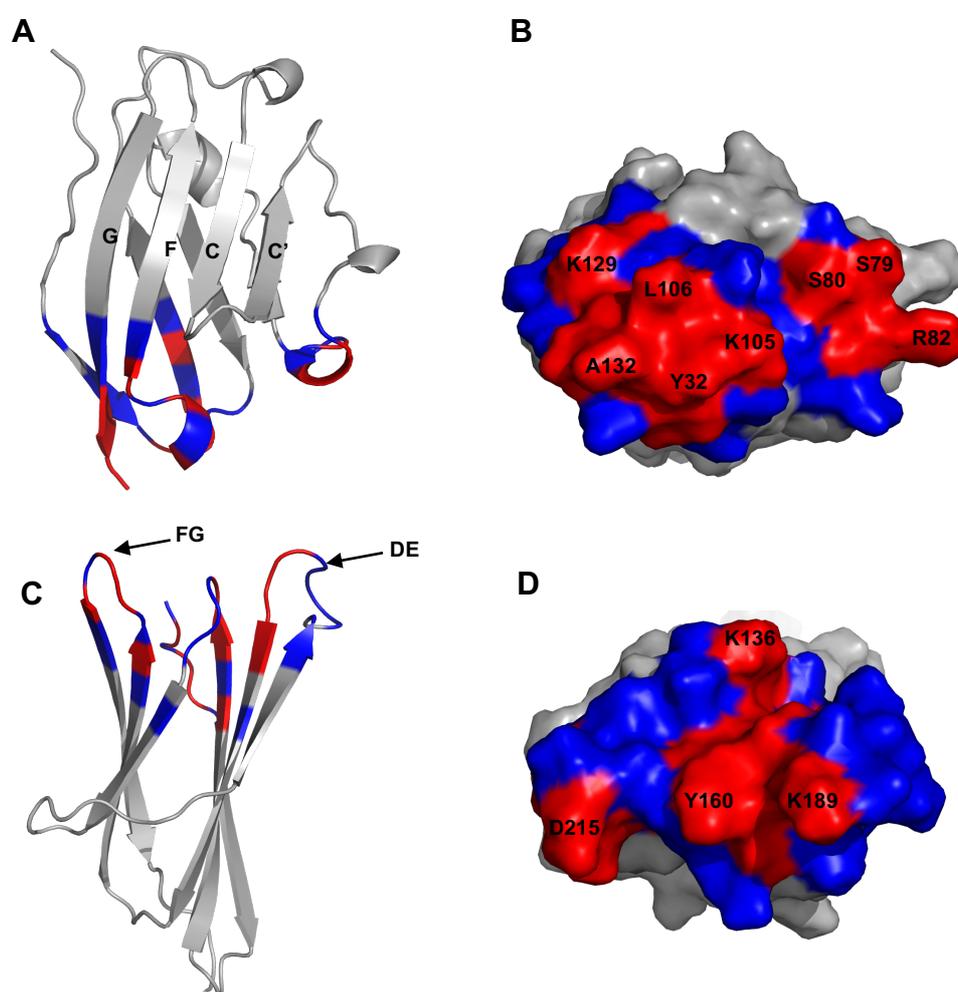
Figure 4.3. Active and Passive residues for HADDOCK.
A) Ribbon representation of PD-L1_D1 with active residues coloured in red and passive residues coloured in blue. **B)** Surface view from the bottom of PD-L1_D1. **C)** Ribbon representation of PD-L1_D2 with active and passive residues coloured as in A). **D)** Surface view from the top of PD-L1_D2.

For the docking PD-L1_D1 and PD-L1_D2 were held together as if one polypeptide chain through the use of Unambiguous Interaction Restraints (Box 4.1). These restraints serve to mimic a planar peptide bond between the C-terminus of PD-L1_D1 (Alanine 132) and the N-terminus of PD-L1_D2 (Proline 133).

Box 4.1. Unambiguous Restraint File

```
!
assign  (segid A and resid 132 and name C)
        (segid B and resid 133 and name N) 1.500 0.100 0.100
!
assign  (segid A and resid 132 and name O)
        (segid B and resid 133 and name NH) 3.200 0.100 0.100
```

The first distance restraint aims to have a distance of 1.5 Å between the terminal carboxyl carbon of Alanine 132 and the backbone nitrogen of Proline 133, with distance margins of 0.1 Å. This distance was chosen based on the distance of the peptide bond measured from the crystal structure (PDB: 3FN3). The second distance restraint aims to have a distance of 3.2 Å between the carbonyl oxygen of Alanine 132 and the protons of the N-terminus of Proline 133, with distance margins of 0.1 Å. This serves to keep the mimicked peptide bond planar as rotation about this bond would result in shorter distances between the atoms.

4.3 Results

4.3.1 Size Exclusion Chromatography-Multi-angle Light Scattering

SEC-MALS was used to determine the oligomeric state of both PD-1 and PD-L1 in solution, as well as the stoichiometry of the complex formed when the two bind. Size exclusion chromatography showed that PD-L1 (25kDa) eluted at 10.7ml and PD-1 (13kDa) eluted at 12.2 ml (Figure 4.4). Previous calibration of the column with known protein standards estimates that proteins eluting at 10.7 ml and 12.2 ml are 40kDa and 24kDa respectively, which are both larger than expected and could be indicative of dimers of PD-L1 and PD-1. Size exclusion chromatography of the PD-L1_D1D2/PD-1 mixture showed the elution of two peaks. The first peak elutes at 10 ml and a second peak at 10.8ml, likely representing the PD-L1/PD-1 complex and free PD-L1 respectively (Figure 4.5). MALS analysis, as material is eluted from the size exclusion column, determines the molecular weight of free PD-L1_D1D2 to be 23.1 +/- 1.1 kDa and PD-1 is determined as 13.4 +/- 0.9 kDa which is consistent with the expected molecular weights (Figure 4.4). The first peak to elute from the column when the PD-L1/PD-1 mixture is loaded is shown by MALS analysis to have a molecular weight of 31.2 +/- 0.7 kDa. This is consistent with a PD-L1/PD-1 complex of 1:1 stoichiometry. The second peak to elute has a molecular weight of 23.6 +/- 1.0 kDa which is consistent with free PD-L1 (Figure 4.5). Due to incomplete separation of PD-L1/PD-1 complex and free PD-L1, as shown by the SEC trace, the eluted fractions containing the complex were loaded back onto the column an hour after their elution. This is shown in Figure 4.5B. It is apparent that the PD-L1/PD-1 complex is eluting as a broader peak and has shifted to an elution volume of 10.5 ml at the apex. MALS analysis shows that across the peak the average molecular weight stays relatively level at 29.4 +/- 0.3 kDa which is still consistent with a 1:1 complex. SDS-PAGE analysis shows the presence of both PD-L1 and PD-1 in the eluted fractions (Figure 4.6). This is suggesting that the PD-L1 and PD-1 complex is dissociating as the proteins move through the column which is to be expected considering a relatively weak binding affinity of 8µM has been reported previously (Cheng *et al*. 2013).
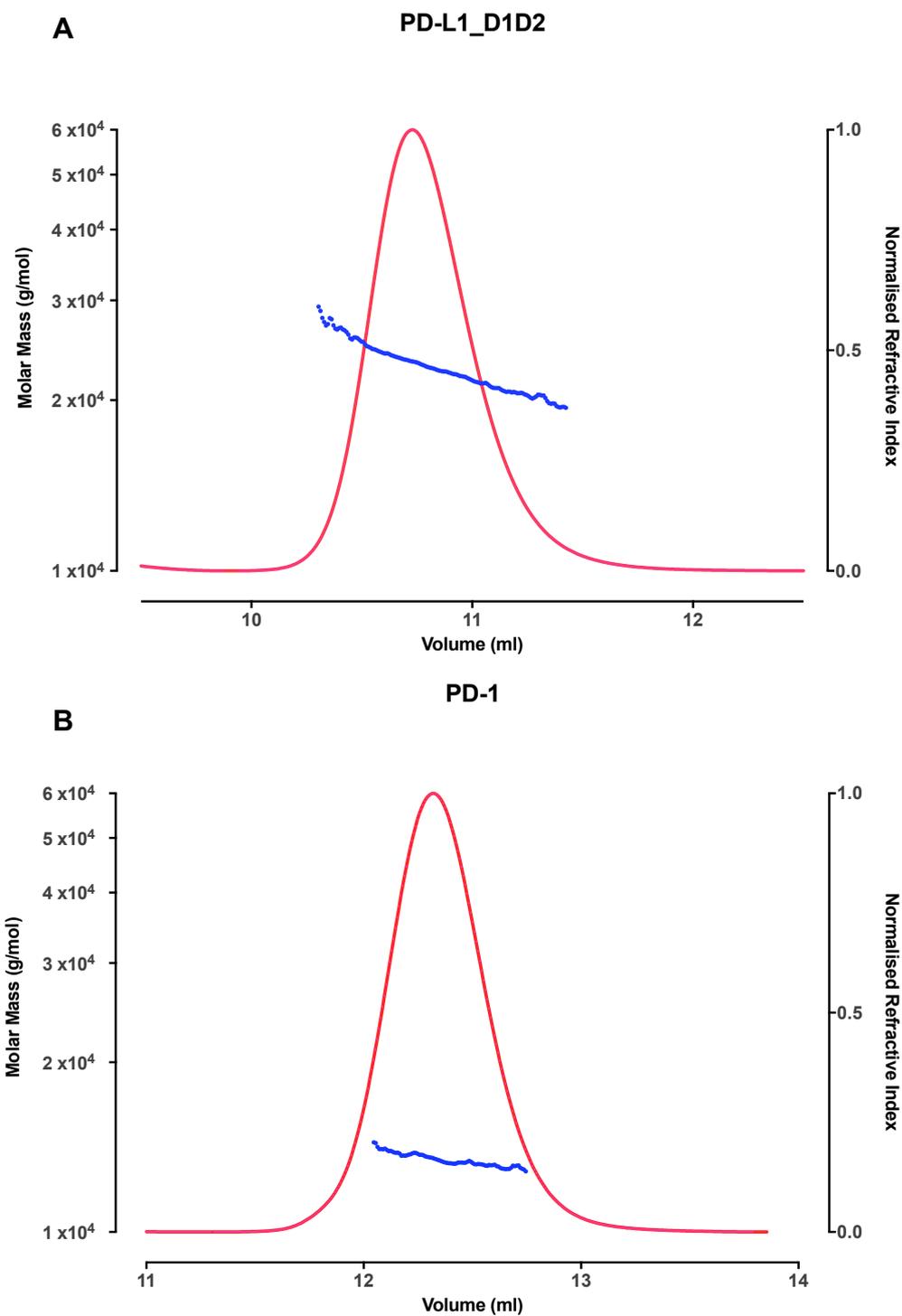
Figure 4.4. SEC-MALS analysis of PD-L1_D1D2 (A) and PD-1 (B).

The normalised Refractive Index (RI) of the eluted material from size exclusion chromatography is plotted as a function of elution volume (red) with the molar mass of the peak region, as calculated by multi-angle light scattering, shown in blue.
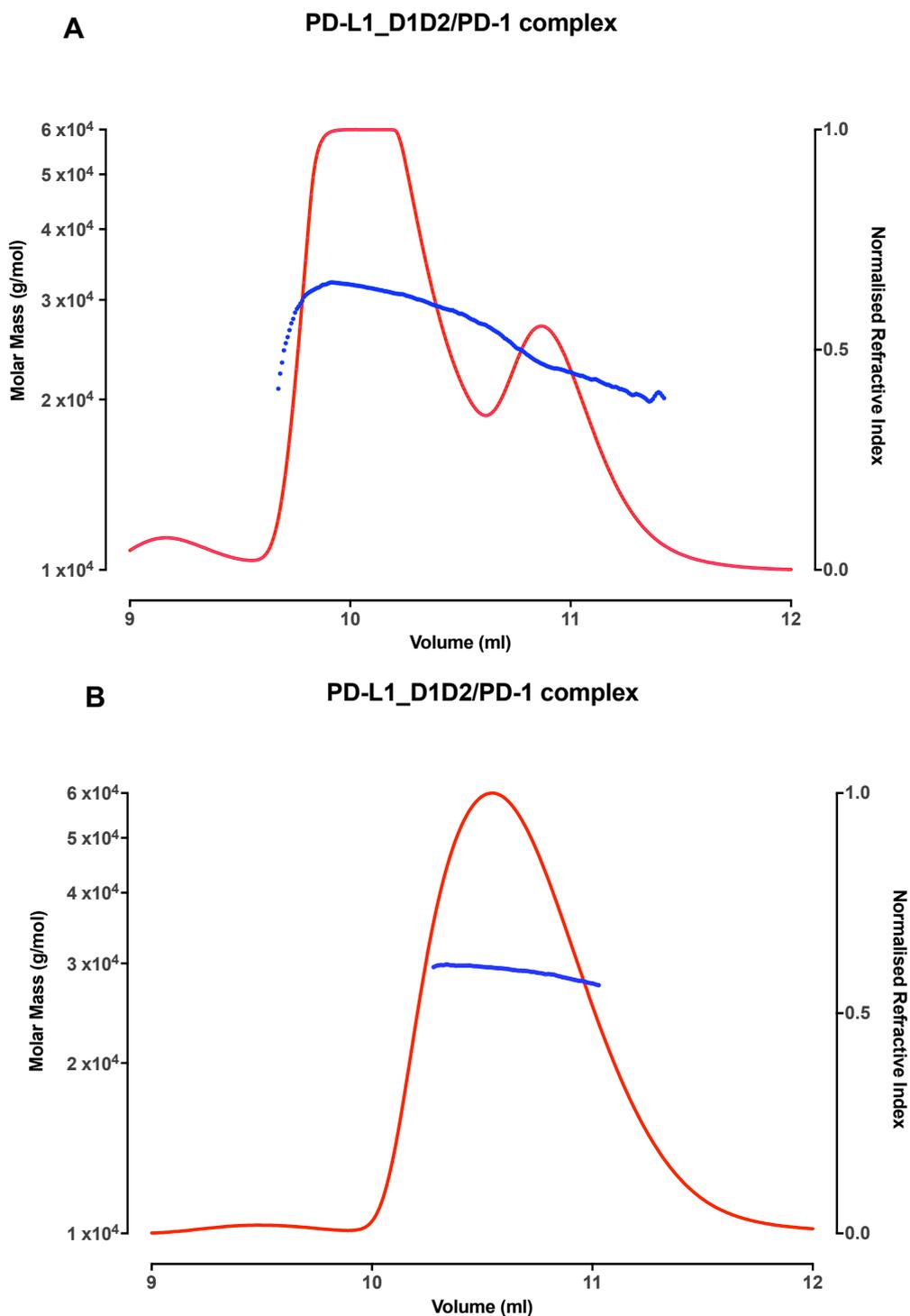
Figure 4.5. SEC-MALS analysis of the PD-L1_D1D2/PD-1 complex.

A) Analysis of a 130μM PD-L1_D1D2 and 120μM PD-1 mixture. The normalised RI of eluted peaks is shown in red with the molar mass of the peak region, as calculated by multi-angle light scattering, shown in blue. **B)** The peak eluting at 10ml was collected and re-loaded onto the column an hour later. Colours as in A).
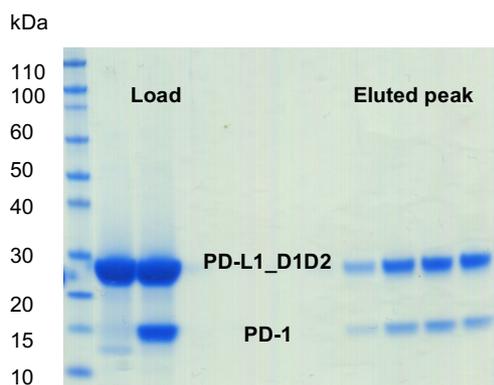
Figure 4.6. SDS-PAGE analysis of SEC-MALS eluted fractions. Shown is the non-reduced SDS-PAGE gel of samples of the PD-L1_D1D2/PD-1 complex loaded onto the column and fractions across the eluted peak. PD-L1 and PD-1 is eluting across the peak.

	Expected MW (kDa)	MW (kDa) Peak 1	MW (kDa) Peak 2
PD-L1_D1D2	25.3	23.1 +/- 1.1	
PD-1	13.3	13.4 +/- 0.9	
PD-L1/PD-1	38.6	31.2 +/- 0.7	23.6 +/- 1.0
PD-L1/PD-1 re-load		29.4 +/- 0.3	

Table 4.2. Expected and measured molecular weights of PD-L1_D1D2, PD-1 and the complex formed when these two proteins bind as determined by SEC-MALS

4.3.2 NMR spectroscopy: PD-L1 binding to PD-1

Investigation of PD-1 binding to ^{15}N-labelled PD-L1_D1D2 was performed using backbone amide chemical shift perturbation analysis. Figure 4.7 shows an overlay of the ^{15}N-TROSY spectrum of PD-L1_D1D2 and the ^{15}N-TROSY spectrum of PD-L1_D1D2 in the presence of unlabelled PD-1. In the presence of PD-1 there are a number of PD-L1_D1D2 signals that show chemical shift changes as well as some signals that do not, indicative of specific binding of PD-1 to PD-L1_D1D2. Figure 4.8 shows a bar chart representation of the minimal shift analysis of PD-1 binding to PD-L1. Unassigned backbone amide signals were included in the analysis and, where possible, it was determined whether these signals were from D1 or D2. In the case of signals that were unassigned and shifted between single domain PD-L1_D1/2 spectra and full extracellular PD-L1_D1D2, these signals were labelled as Sh (shifted).
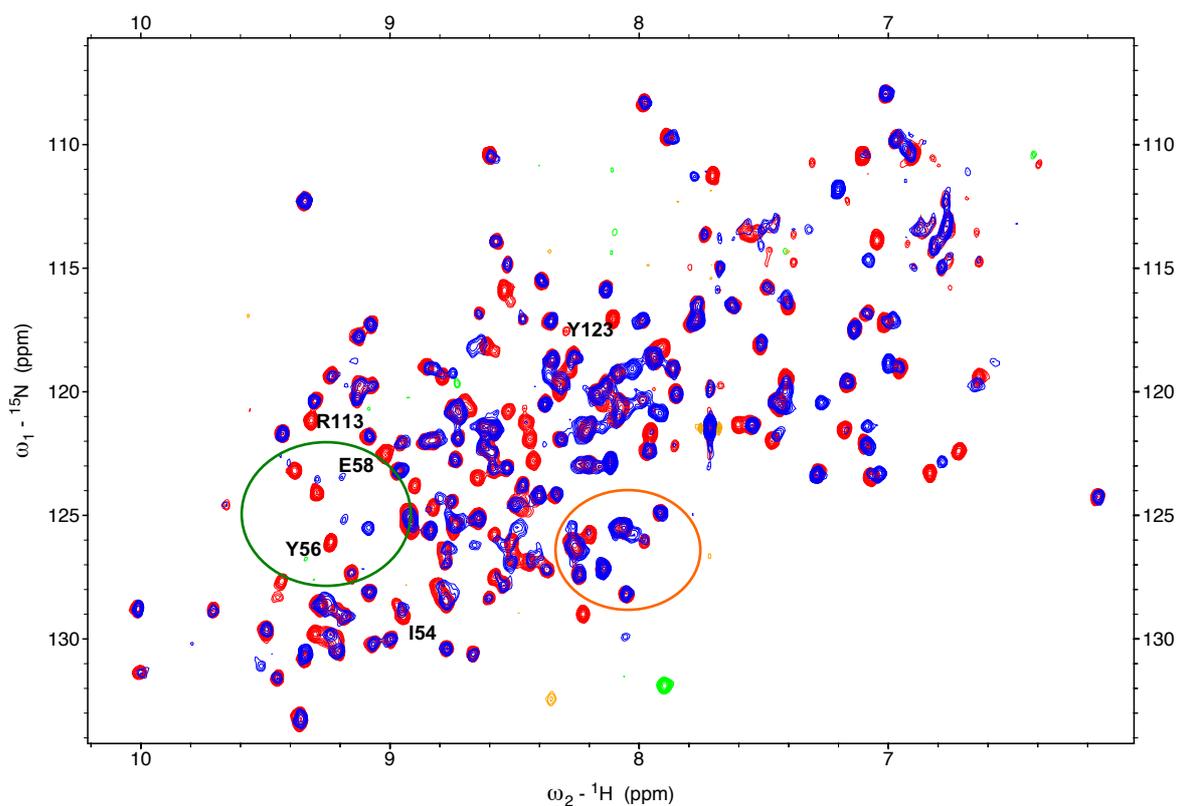
Figure 4.7. PD-1 binding to PD-L1_D1D2.

An overlay of the ^{15}N/^{1}H TROSY spectra of free ^{15}N-PD-L1_D1D2 at 110µM (red) and ^{15}N-PD-L1_D1D2 in the presence of a 20% molar excess of PD-1 at a final concentration of 135µM (blue). Spectra were collected at 30°C on an 800 MHz Bruker spectrometer. Examples of signals that remain in the same position are circled in orange. Examples of signals that have shifted in the presence of PD-1 are circled in green.
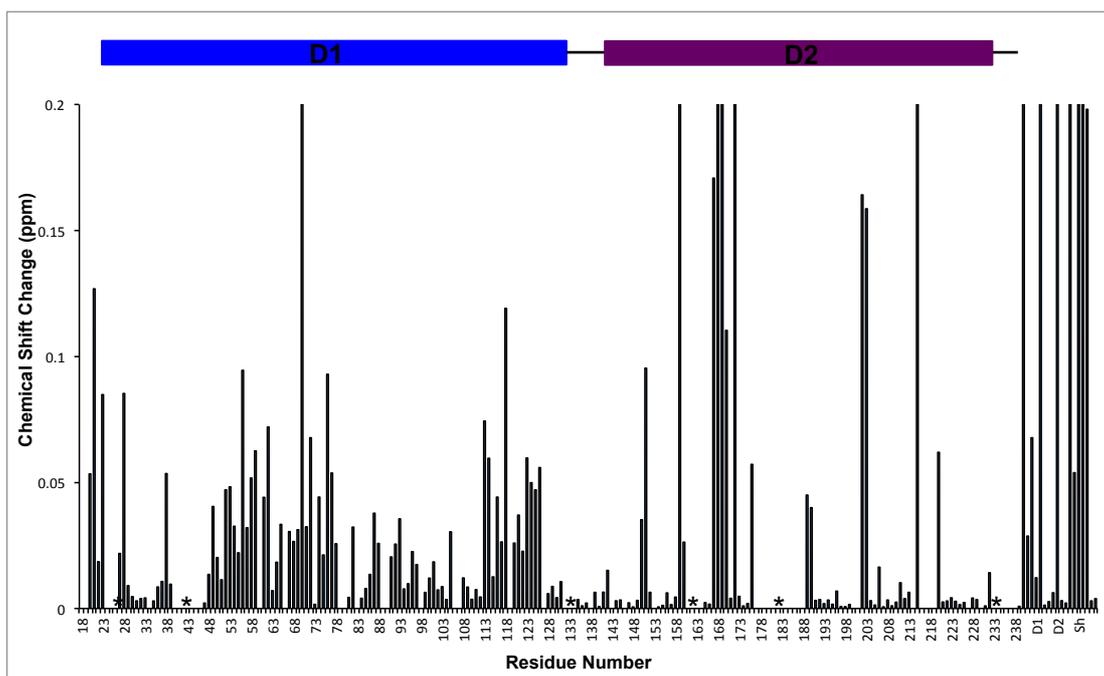
Figure 4.8. Chemical shift mapping of the binding of PD-1 to PD-L1_D1D2.
A histogram representation of the combined ^{15}N and ^{1}H backbone amide chemical shift changes of PD-L1_D1D2 residues upon binding of PD-1. Unassigned backbone amide signals are included to the right and are labelled as D1, D2 or Sh (shifted) as determined by overlaying of spectra. An * indicates unassigned residues.

The bar chart shows that the N-terminal Ig-like domain (D1) of PD-L1 is more affected by the binding of PD-1 with widespread chemical shift changes. Interestingly there are also some chemical shift changes in both assigned and unassigned residues of D2 as well as unassigned signals labelled shifted (Sh) (Figure 4.8). Figure 4.9A shows the mapping of the chemical shift changes upon PD-1 binding onto the structure of PD-L1_D1D2 with white representing no significant shift; shifts between 0.01 and 0.04 are represented in pink on a scale from white to red and shifts over 0.04 are shown in red.

85
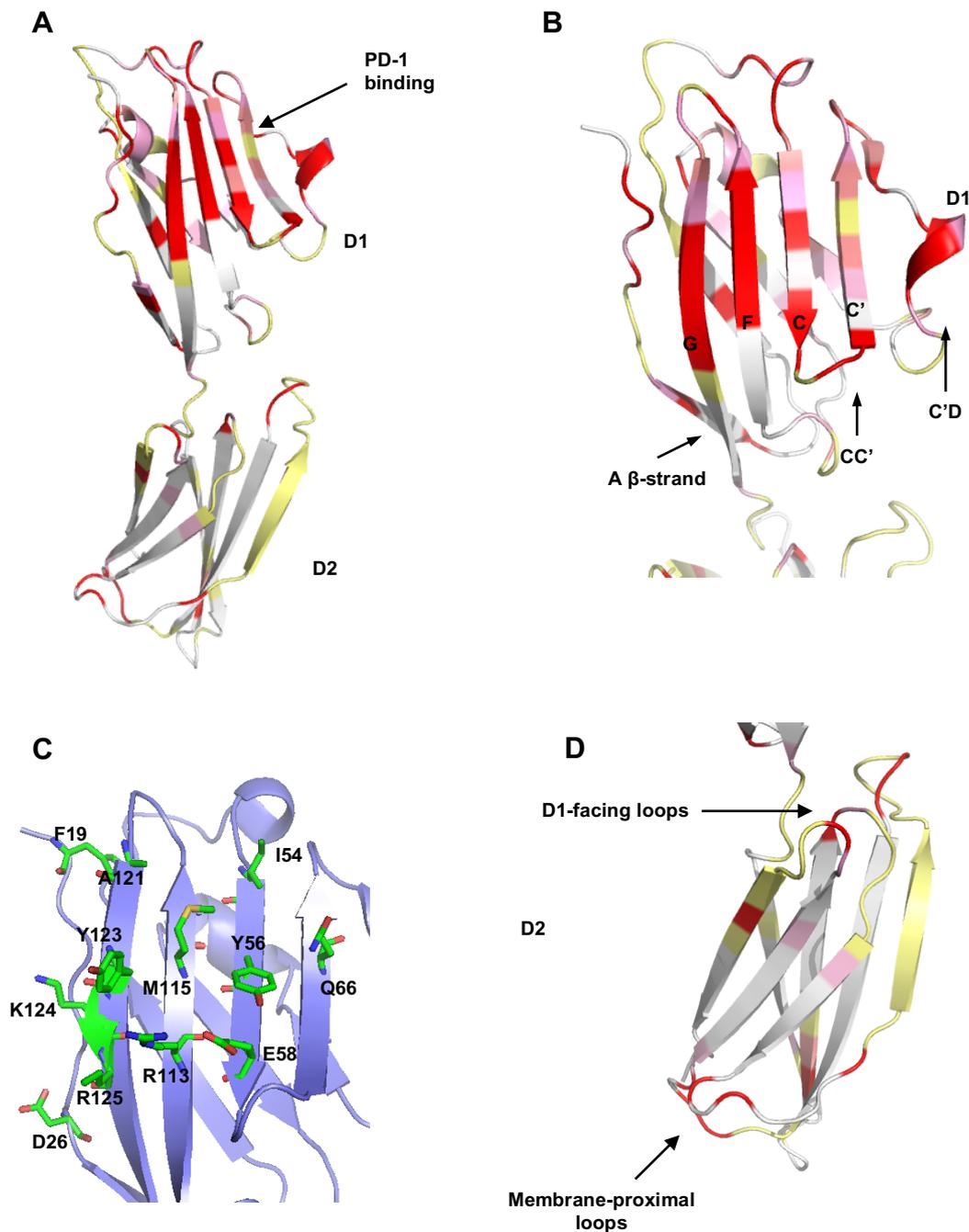
Figure 4.9. Chemical shift mapping of PD-1 binding to PD-L1_D1D2.
A) The mapping of chemical shift changes onto the structure of PD-L1_D1D2. Shifts over 0.04 are shown in red and shifts less than 0.01 are shown in white. Shifts in between 0.01 and 0.04 are shown in pink on a linear scale from white to red. Residues for which there are no backbone amide assignments are shown in yellow. **B)** A view of the chemical shift changes at the PD-1 binding face of domain 1 (D1). **C)** PD-L1 residues that have been previously shown, using x-ray crystallography, to be involved in PD-1 binding are shown in green. **D)** A view of the chemical shift changes on domain 2 (D2).

The mapping of chemical shift changes upon binding of PD-1 onto the structure of PD-L1_D1D2 shows a patch of large chemical shift changes covering the G, F and C β-strands of D1 of PD-L1_D1D2 (Figure 4.9B). This includes resides I54, Y56, E58, R113, M115, Y123, K124, and R125, that have been shown previously by crystallography to be important contacts in the binding interface with PD-1 (Figure 4.9C) (Zak *et al.* 2015). The presence of widespread chemical shift changes on D1 suggests there could be a conformational change occurring upon the binding of PD-1. Surprisingly there appears to be a number of residues at the membrane-proximal loops (H151, S169, H171, and T201-I206) and D1-facing loops (G159, Y160, K189, L190, L215, D215) of D2 (Figure 4.9D) that show chemical shift changes when PD-1 binds. It could be possible that binding of PD-1 induces a conformational change through D1 and into D2, perhaps changing the orientation of the domains in relation to each other. This would explain the chemical shift changes seen away from the PD-1 binding site on D1 on the A β-strand, CC' and C'D loops (Figure 4.9B).

The binding of PD-1 to PD-L1_D1 was also investigated using NMR chemical shift perturbation to determine if there were differences between PD-L1_D1 and PD-L1_D1D2 chemical shift changes. Figure 4.10 shows a histogram showing the chemical shift changes for residues A18-Y134 of PD-L1_D1 and PD-L1_D1D2 upon PD-1 binding. There are a number of differences between the chemical shift changes of PD-L1_D1 and PD-L1_D1D2 when binding to PD-1, summarised in Table 4.3. Since a minimal shift approach was used to investigate chemical shift perturbations when PD-1 binds, it is reasonable to expect that isolated peaks in the PD-L1_D1 spectrum will have larger shifts than in the PD-L1_D1D2 spectra due to there being less crowding of signals. This could explain the difference in the case of V21, Y28, I54, Y56, F67, A97, K105, Y112, and V128, however this does not explain the incidences of there being a substantially larger shift change in the case of PD-L1_D1D2. Figure 4.10 shows these residues highlighted in red on the structure of PD-L1. A number of these residues are grouped together on the CC' β-strands and the long C'D loop, however, and this could be suggestive a situation where this section of PD-L1 is in a different conformation depending on whether the

construct is PD-L1_D1 or PD-L1_D1D2. There is also the possibility that these D1 residues represent a contact surface with D2.

A
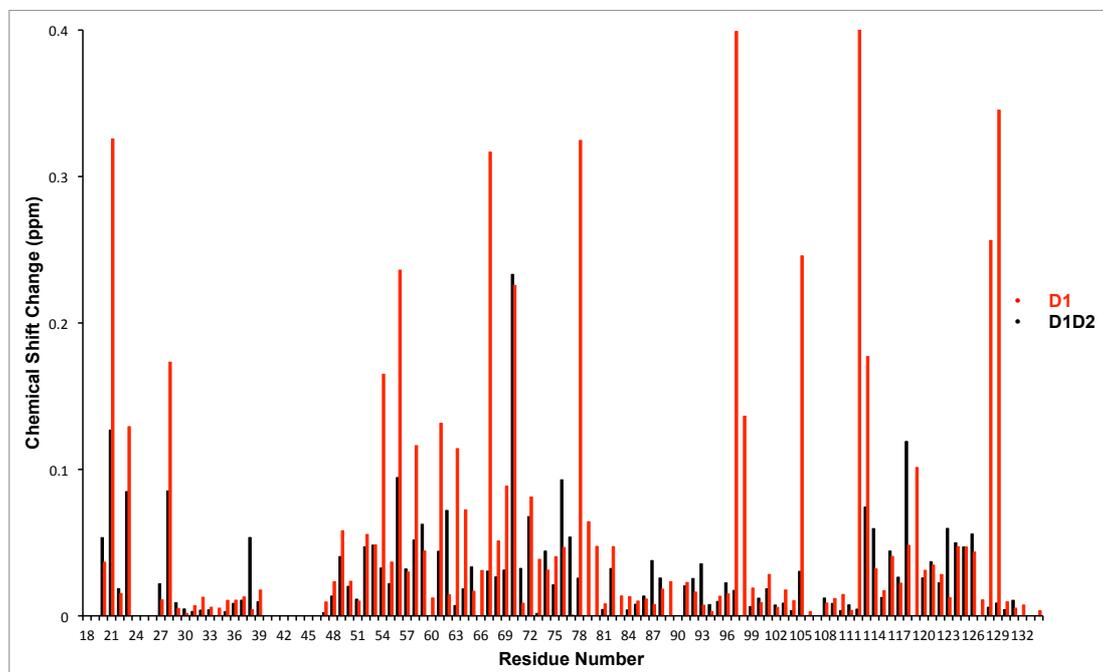
B
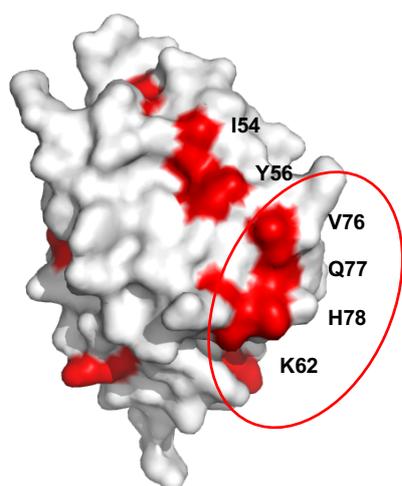
C
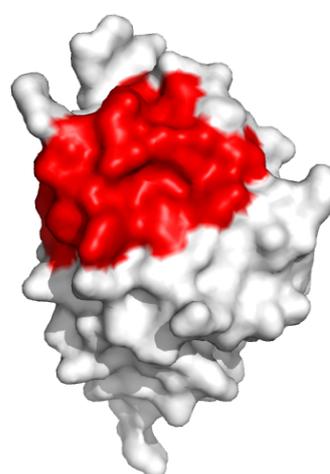
Figure 4.10. Comparison of the chemical shift changes of PD-L1_D1 and PD-L1_D1D2 when binding PD-1.

A) Histogram representation of the chemical shift changes for residues A18-Y134 of PD-L1_D1 (red) and PD-L1_D1D2 (black). **B)** Residues with significant differences in PD-1-induced shifts between D1 and D1D2 are shown in red. Circled residues represent a possible D2 interface. **C)** Surface view of PD-L1_D1 with PD-1 binding site coloured in red.

Residue	Difference between D1 and D1D2
V21	+ 0.20
Y28	+ 0.09
I38	- 0.05
I54	+ 0.13
Y56	+ 0.14
K62	- 0.06
W67	+ 0.29
V76	- 0.05
Q77	- 0.05
H78	+ 0.30
A97	+ 0.38
K105	+ 0.22
Y112	+ 0.69
Y118	- 0.07
V128	+ 0.25
K129	+ 0.34

Table 4.3. A table summarising the major chemical shift differences between PD-L1_D1 and PD-L1_D1D2 when bound to PD-1. Positive values have larger shifts in D1 and conversely negative values have larger shifts in D1D2.

4.3.3 NMR spectroscopy: Sequence specific assignments of PD-L1_D2

Sequence specific backbone assignment of PD-L1_D2 was completed to 60% using triple resonance HNCACB, HN(CO)CACB, HNCA and HN(CO)CA spectra. Figure 4.11 shows the strip plot for a representative section of the amino acid sequence. The incompleteness of the backbone assignments is due mainly to a single region of D2, consisting of the BC and DE loops, as well as the entirety of the D β-strand remaining unassigned, due to lack of signals in the 3D spectra (Figure 4.11). It is likely that these loops are occupying multiple conformational states that are exchanging on a relatively low timescale. This would lead to broadening of the signals to such an extent that they cannot be detected.
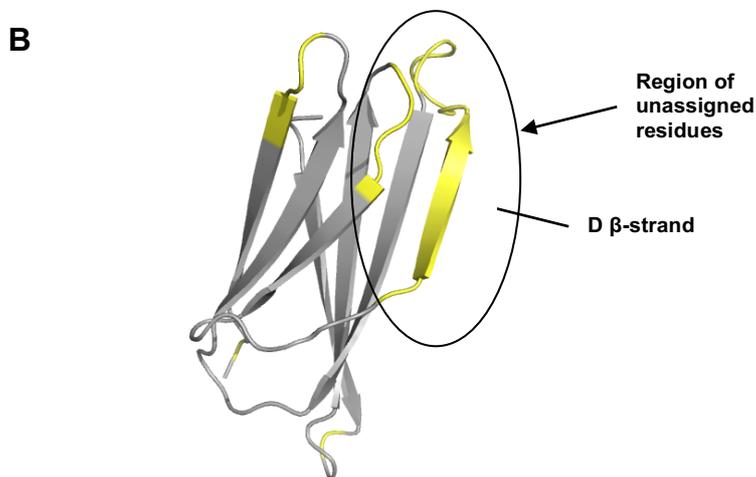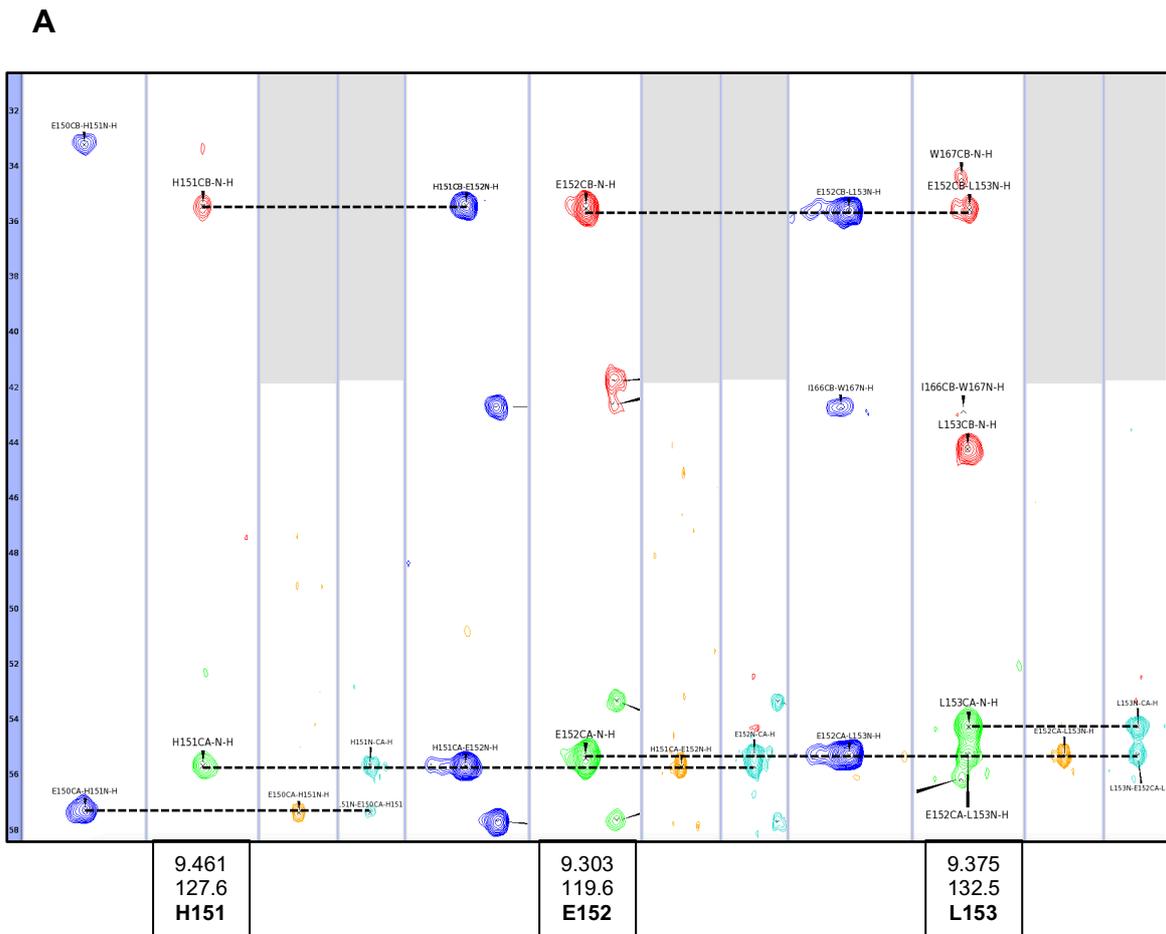
Figure 4.11. PD-L1_D2 sequential backbone assignments.

A) Protein backbone strips from the HNCACB, CBCA(CO)NH (blue), HNCA (cyan) and HN(CO)CA (orange) spectra of ^{15}N/^{13}C labelled PD-L1_D2, showing the inter- and intra-residue connections linking signals from residues H151 to L153. The HNCACB spectrum shows Cα signals in green and Cβ signals are shown in red. Dashed black lines indicate sequential Cα and Cβ connections. **B)** Cartoon representation of PD-L1_D2 with unassigned residues coloured in yellow.

4.3.4 NMR spectroscopy: Chemical shift analysis of PD-L1 domains

Since binding of PD-1 to PD-L1_D1D2 causes chemical shift changes in D2 and the affinity of PD-L1_D1D2 for the BPI was an order of magnitude less than the affinity of PD-L1_D1, I set out to investigate whether the two domains of the extracellular region of PD-L1 interact with each other. Overlay of the ^{15}N-HSQC spectra of the individual domains, PD-L1_D1 and PD-L1_D2, and the full extracellular region of PD-L1, PD-L1_D1D2 shows that a subset of signals in both D1 and D2 are shifted (Figure 4.12). This suggests that as opposed to the domains acting independently of each other with a flexible linker there is some interaction between the domains. Figure 4.13 shows bar chart representations of the chemical shift changes between the individual domains and the full extracellular region of PD-L1. Those residues with a chemical shift change greater than 0.02 were mapped onto a crystal structure of PD-L1_D1D2 (PDB: 3FN3; Chen *et al.* 2010) (Figure 4.14). There is a patch of chemical shift changes on the bottom of PD-L1_D1 involving the CC', C'D and DF loops that appear to show a potential protein interface (Figure 4.14). The incompleteness of D2 assignments makes the identification of a binding interface on D2 difficult but the presence of shifts in the FG loop and either side of BC loop suggests a situation where the loops of D1 and D2 form a binding interface. A schematic model of this is shown in figure 4.14. Interestingly, some of these residues (G159, Y160, K189, L190, F211, L214, D215, H220) also shift upon PD-1 binding, suggesting that a conformational change of PD-L1_D1D2 upon binding of PD-1 could alter a possible PD_L1 D1/D2 interface. This could go some way to explaining the mechanisms by which PD-L1 binds with different partners both *in cis*, on the same membrane, and *in trans*, across membranes.
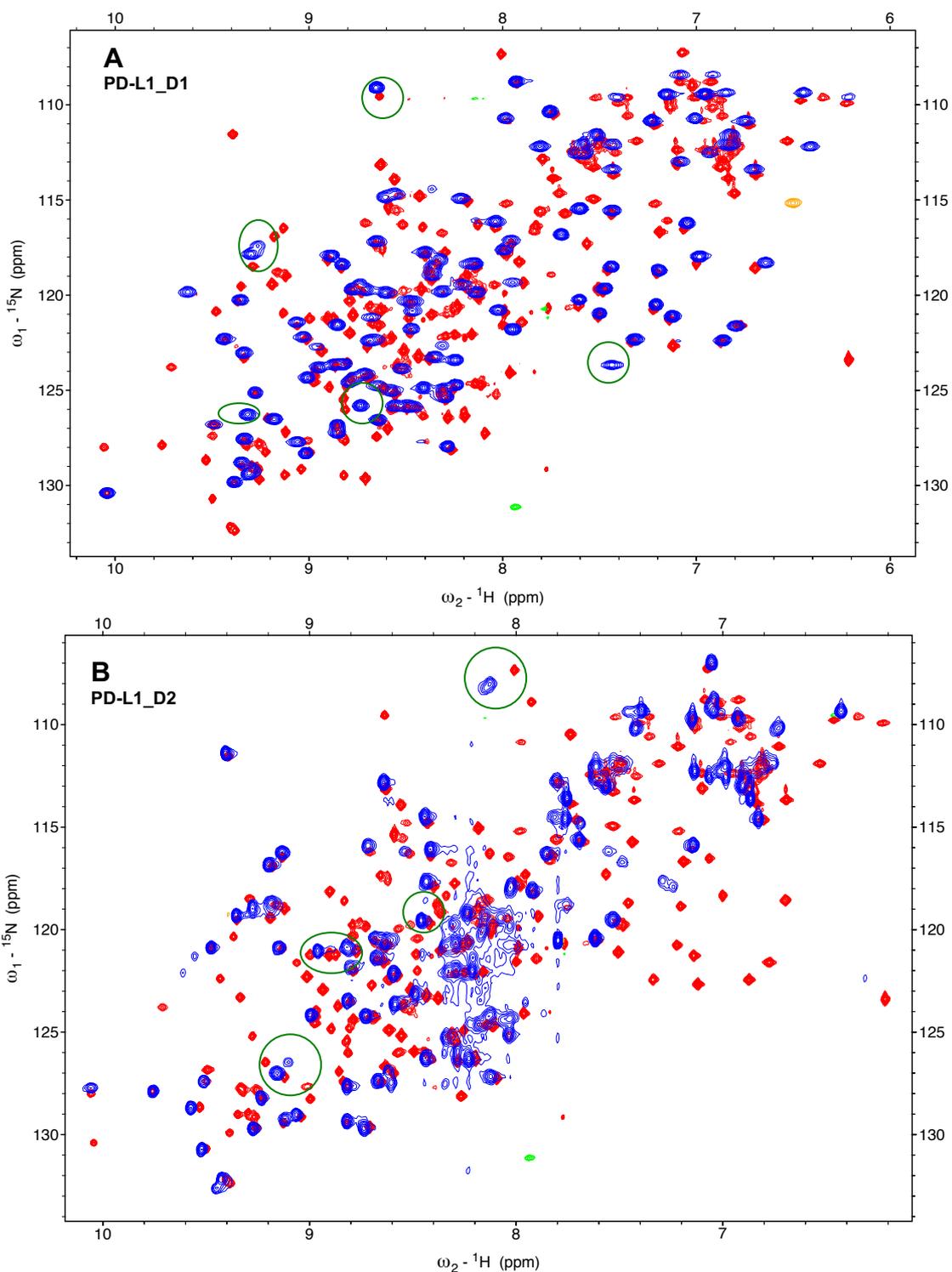
Figure 4.12. Overlay of PD-L1_D1, PD-L1_D2, and PD-L1_D1D2 spectra.
A) An overlay of the ¹⁵N-HSQC spectra of PD-L1_D1D2 (red) collected on a 950 MHz spectrometer at 30°C and the ¹⁵N-HSQC spectra of PD-L1_D1 (blue) collected on a 600 MHz spectrometer at 25°C. **B)** An overlay of the same ¹⁵N-HSQC spectra of PD-L1_D1D2 (red) and the ¹⁵N-HSQC spectra of PD-L1_D2 (blue) collected on a 800 MHz spectrometer at 30°C. Examples of residues that shift are circled in green.
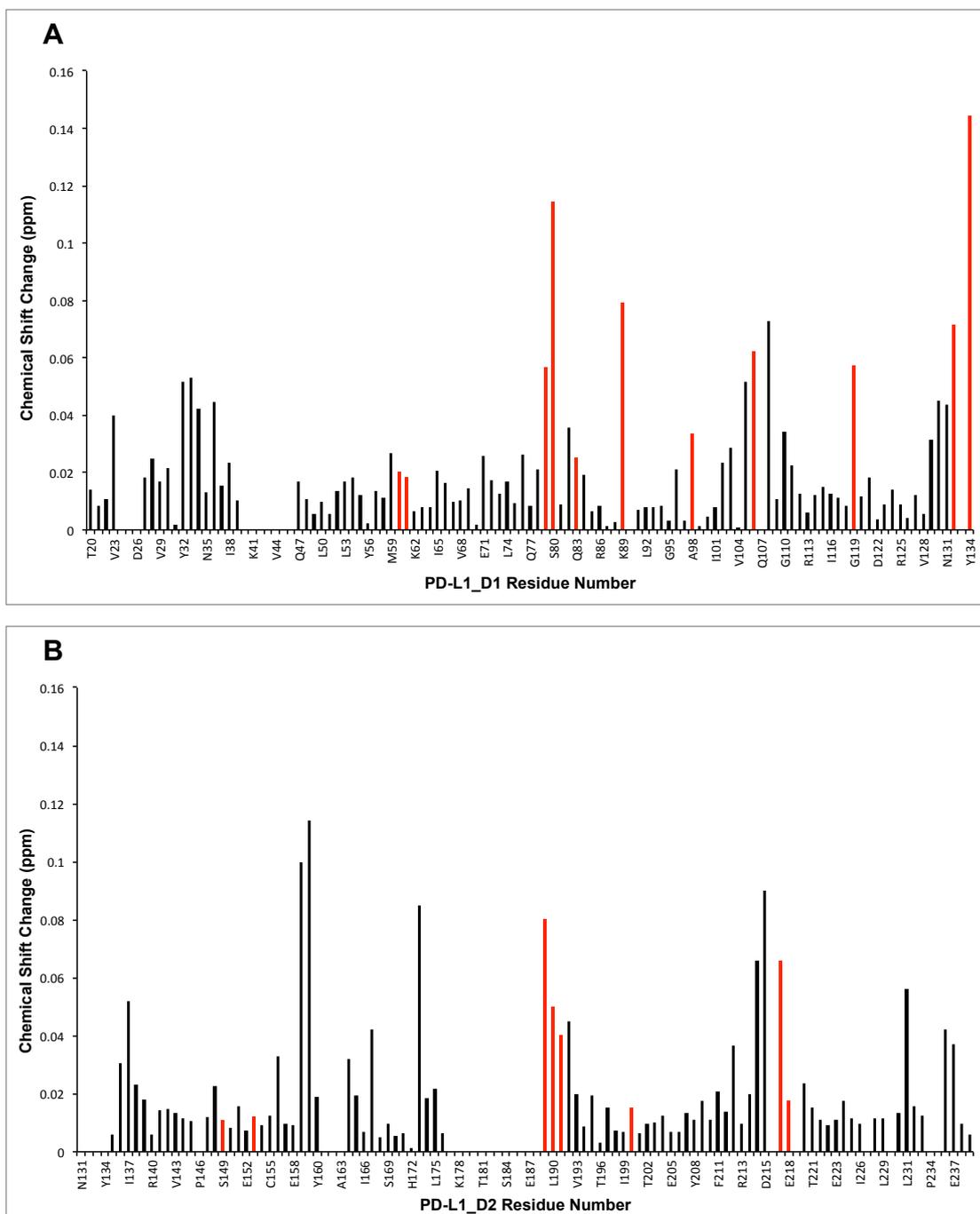
Figure 4.13. Chemical shift changes between individual PD-L1 domains and the full extracellular region of PD-L1.

Histogram representation of the chemical shift changes between PD-L1_D1 **(A)** and PD-L1_D1D2 and PD-L1_D2 **(B)** and PD-L1_D1D2. Shifts highlighted in red are minimal shift values due to only being assigned in one protein. Values of zero indicate unassigned residues.
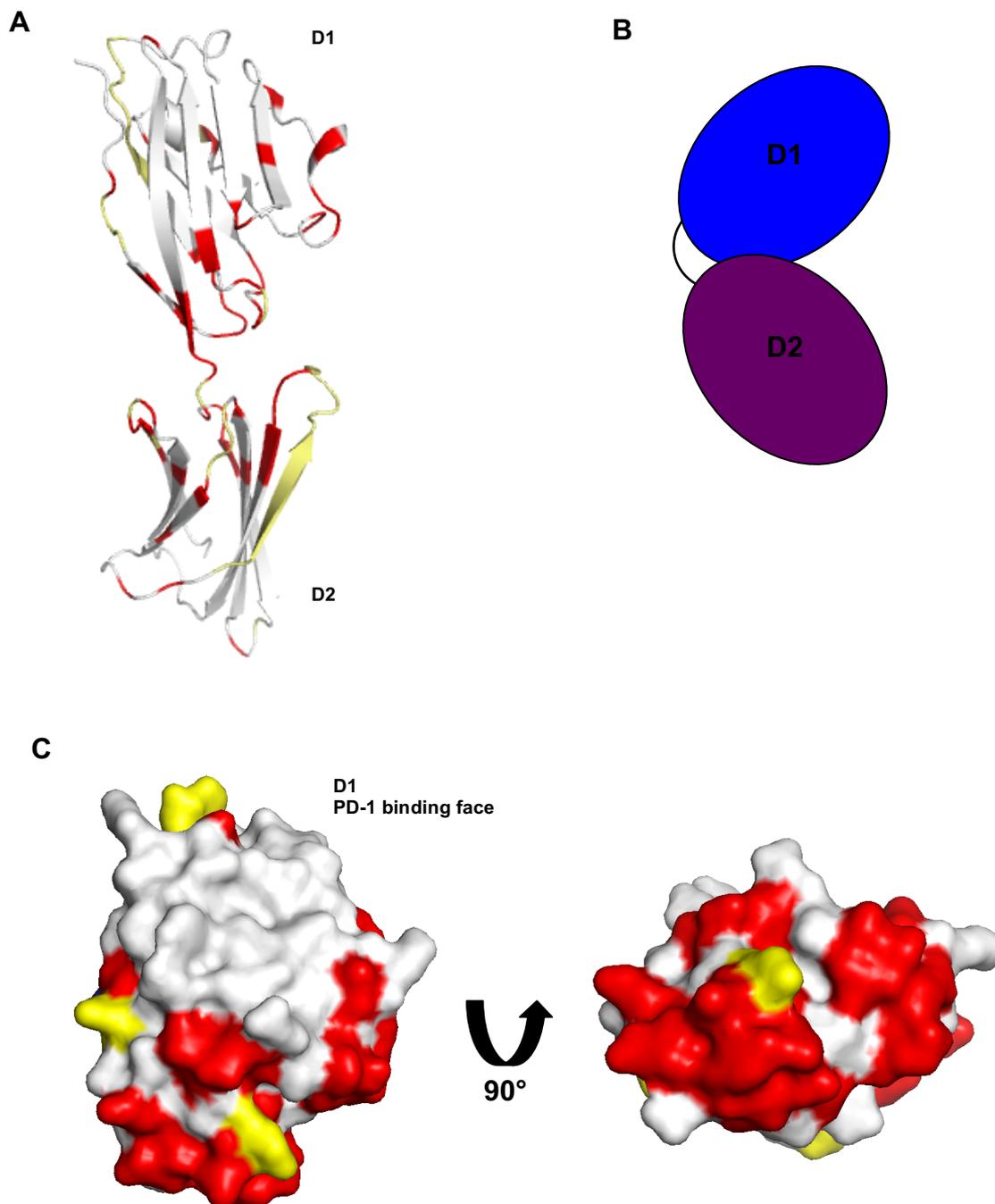
Figure 4.14. Mapping of chemical shift changes between individual PD-L1 domains and the full extracellular region of PD-L1.

A) Ribbon representation of a crystal structure of PD-L1_D1D2 (PDB: 3FN3) with chemical shift changes greater than 0.02 ppm shown in red. Unassigned residues are shown in yellow. **B)** Schematic of D1/D2 interface. **C)** Surface representation of PD-L1_D1 showing the PD-1 binding face and a 90° rotation showing a patch of chemical shift changes made up of the CC', C'D and DF loops of PD-L1_D1.

4.3.5 Long-range backbone NOE cross-peaks in PD-L1_D1D2

To further explore whether there is an interface between D1 and D2 of PD-L1_D1D2, a ^{15}N-NOESY-HSQC experiment was performed to search for long-range NOEs between backbone amides of PD-L1 residues in D1 and D2. Figure 4.15 shows NOE cross-peaks observed for V104, which is representative of the quality of the data.
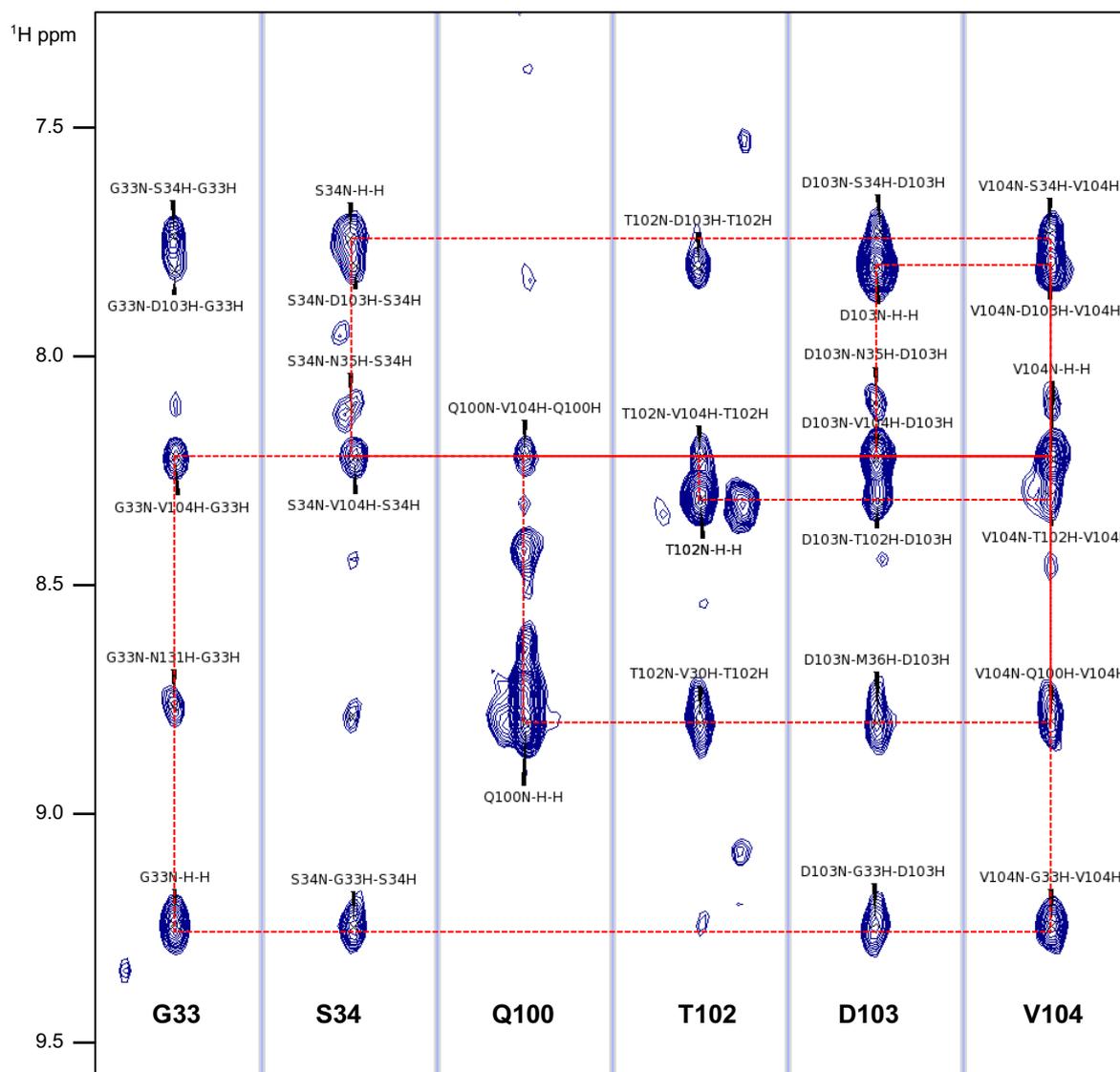
Figure 4.15. NOE cross-peaks for V104.

Strips from the ^{15}N-NOESY-HSQC spectra of ^{15}N/^{13}C/^{2}H labelled PD-L1_D1D2. NOE cross-peaks between the backbone amide of V104 and G33, S34, Q100, T102 and D103 are shown by red rectangles.

NOE cross-peaks for residues in D1 that were shown to shift significantly between PD-L1_D1 and PD-L1_D1D2 were investigated to determine if there were any cross-peaks to D2. These are summarised in Table 4.4. All NOE cross-peaks identified for the residues that show chemical shift changes between PD-L1_D1 and PD-L1_D1D2 are between residues in D1, meaning that NOE cross-peaks across a D1/D2 interface are not visible. There are a number of reasons why this could be. Firstly, due to time restrictions, the mixing time used when collecting the spectra was decided based on previous studies in the group on antibody Fab fragments, instead of being experimentally optimised by collecting a series of spectra with different mixing times. This could have limited the distance of NOE cross-peaks that were observed. Using distances between backbone amides in a crystal structure of PD-L1_D1D2 as a guide for NOE cross-peak distances, long-range (5-6Å) NOEs appear to be observable. There are NOE cross-peaks between residues that span over 8Å as measured in the crystal structure (Table 4.4).

Table 4.4. A table summarising the NOE cross-peaks for PD-L1_D1 residues that show chemical shift changes in PD-L1_D1D2 spectra. Intramolecular distances were measured in PyMOL from a crystal structure of PD-L1_D1D2. PDB: 3FN3. All NOE cross-peaks observed were in D1.

Residue	NOE cross-peaks	Maximum distance in crystal (Å)
Y28	V30, T127, K129	6.4 (V30)
V30	Y28, E31, Y32, K129, V130	6.4 (Y28, Y32)
Y32	V30, E31, K129	9.6 (K129)
G33	S34, V104, N131	8.1 (N131)
S34	G33, N35, D103, V104	4.7 (D103)
V76	K75, Q77	2.9
Q77	K75, V76	4.8 (K75)
T102	D103, V104	4.9 (V104)
D103	G33, S34, N35, M36, T102, V104	6.3 (M36)
V104	G33, S34, Q100, T102, D103	10.8 (Q100)
K105	D108	4.9 (D108)
D108	K105, A109	4.9 (K105)
A109	D108, V130	5.7 (V130)
K129	Y28, V30, Y32, V111, V130	9.6 (Y32)
V130	V30, A109, K129, N131	5.7 (A109)
N131	G33, V130	8.1 (G33)
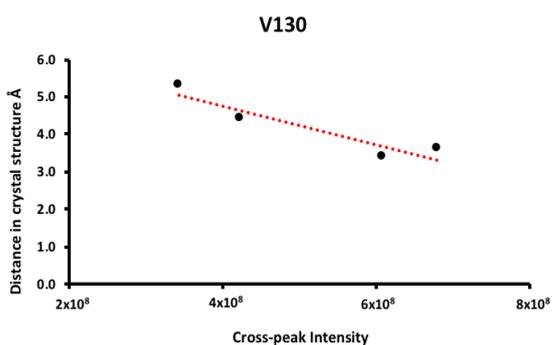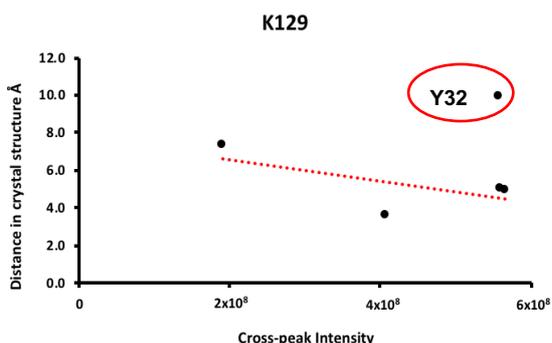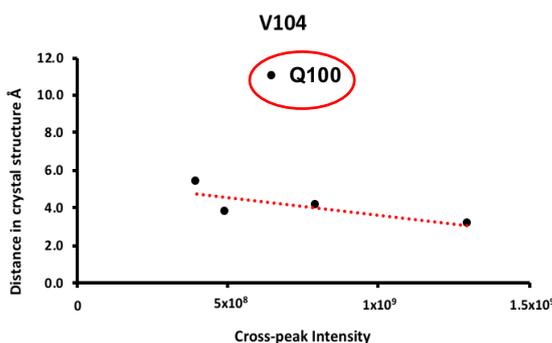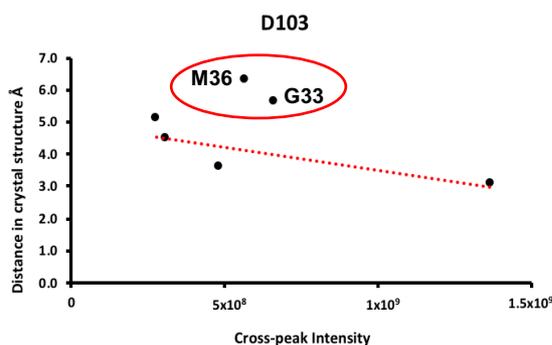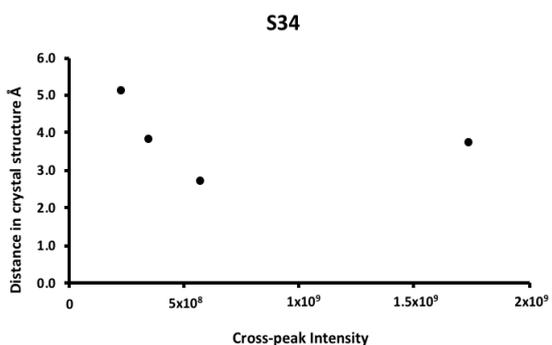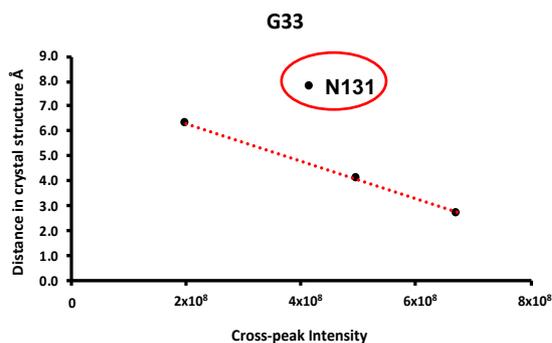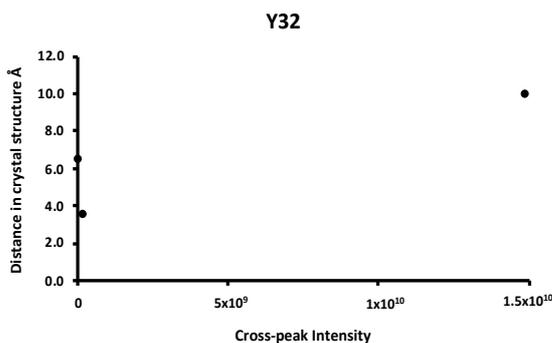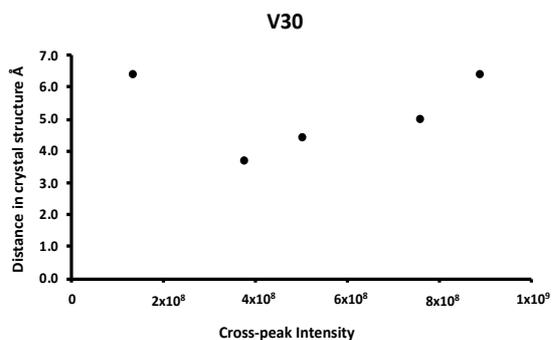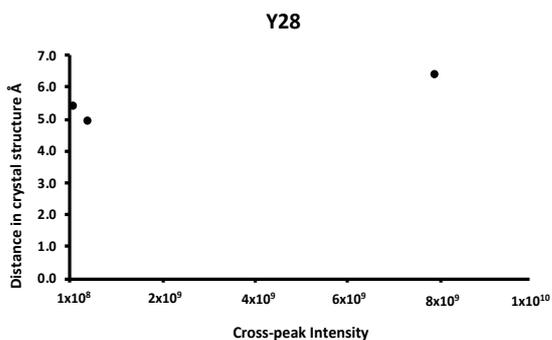
Figure 4.16. NOE distance plots.

The intensity of backbone NH to backbone NH NOE cross-peaks was plotted against distance in the PD-L1 crystal structure PDB: 3FN3. Trendlines are shown in red and outliers are circled.

However, it is unlikely that NOEs are observable across a distance over 8Å suggesting that the residues at the potential interface are in a different conformation to that in the crystal structure (PDB: 3FN3). To further investigate this NOE distance plots were produced for PD-L1 residues that make up a potential interface (Figure 4.16). Since the intensity of an NOE cross-peak is proportional to the distance between the protons involved you would expect to see cross-peaks with the largest intensity having smaller measured distances. Figure 4.16 shows that for the residues at the potential interface there is not a good correlation between NOE cross-peak intensity and distance in the crystal structure overall. For residues G33, D103, V104, K129, and V130 there is a trend of NOE cross-peaks with the largest intensity having the smallest distance as measured in the crystal structure, however, there are outliers (Figure 4.16). For example the distance between the backbone amides of V104 and Q100 is 11Å in the crystal structure but the cross-peak has an intensity that you would expect for a distance of 4Å. It is likely therefore that, in solution, the backbone amides of V104 and Q100 are closer together than in the crystal structure. Taken together this suggests that the potential D1/D2 interface is in a different conformation in solution to that seen in the crystal structure. It could be then, that NOE cross-peaks across a D1/D2 interface have not been observed because the data collection was not fully optimised to observe NOEs across a very large (8-10Å) distance. In addition to this, if the interface is dynamic then the time that the residues spend in the interaction is a factor in the NOE cross-peaks being observed. For example if the interaction is only present for 20% of the time, over the NMR experiment, then the effective concentration of the residues when taking part in the D1/D2 interaction would be only 40µM. This would significantly reduce the signal to noise ratio for the NOE cross-peaks across the interface. Given more time it would be possible to determine distance restraints for the NOE cross-peaks based on their intensities, and therefore determine whether backbone amide NOE cross-peaks across an interface should be observable. It would also be possible to optimise the NOE mixing time such that cross-peaks for very long range NOEs are being observed at their maximum intensity.

4.3.6 HADDOCK docking of interface between the domains of PD-L1

The HADDOCK docking program was used to model a potential interface between the D1 and D2 domains of PD-L1 using NMR chemical shift perturbation data. HADDOCK produced 185 docked structures which were grouped into seven clusters. These are summarised in Table 4.5, along with the HADDOCK scores and subsequent ranking. HADDOCK scoring is a weighted sum of intermolecular van der Waals and electrostatic energies as well as distance restraint energies and desolvation energy, such that the smallest weighted sum is ranked highest (Dominguez *et al.* 2003).

Table 4.5. Summary of HADDOCK structures.

Cluster	Number of structures	HADDOCK score	Rank
1	71	-22.8	1
2	35	-17.8	2
3	30	-14.9	4
4	28	-3.1	6
5	10	-15.2	3
6	6	-2.8	7
7	5	-6.0	5

Figure 4.17 shows each cluster of structures aligned on D1 as well as all 185 structures produced by HADDOCK aligned on D1. After alignment on D1 it can be seen from Figure 4.17A that all 185 solutions have D2 within an approximately 60° cone. This suggests that D1 and D2 cannot significantly fold back on one another via the APYNKI linker. Inspection of the clusters of the docked structures reveals that the cluster can be grouped into two types. Clusters 1, 2, 4 and 5 show the domains twisting in relation to each other whereas clusters 3, 6 and 7 show significant kinking of up to 90° of D2 in relation to D1. Some of these solutions would allow for a D1/D2 interface that orientates D1 in such a way that the PD-1 binding face of PD-L1 is most exposed at the membrane although it is not likely this has any biological significance. It is worth noting that within clusters 1, 2, 4 and 5 (the clusters that show twisting), the structures are better converged than in the remaining clusters, that show a bend between the domains, perhaps indicating that they are more likely to be the true docking solution.
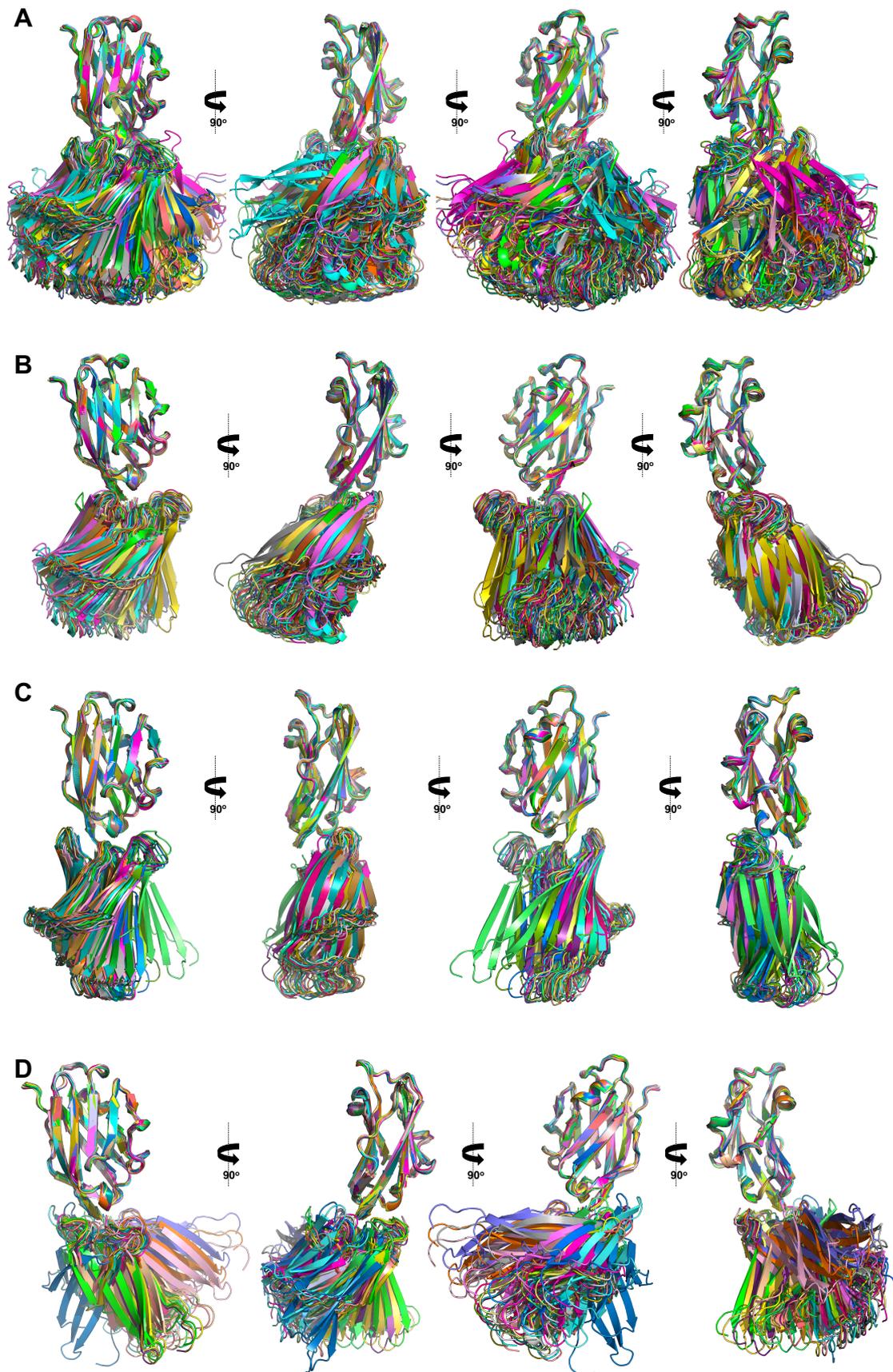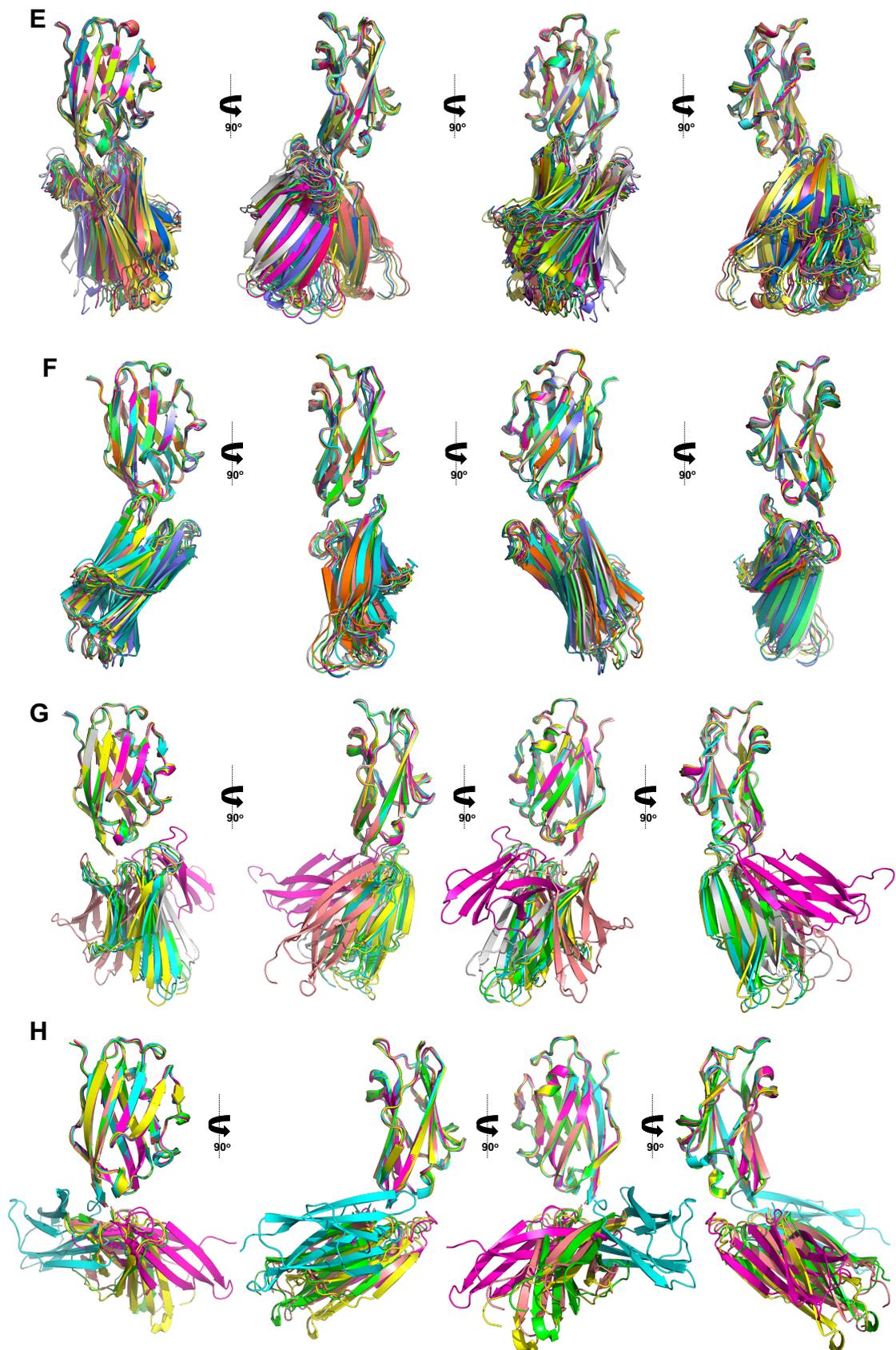
Figure 4.17 Docked structures produced by HADDOCK.

Ribbon representations showing 360° view of HADDOCK structures aligned on D1. **A)** all 185 structures **B)** Cluster 1 **C)** Cluster 2 **D)** Cluster 3 **E)** Cluster 4 **F)** Cluster 5 **G)** Cluster 6 **H)** Cluster 7.

NMR data shown here suggests that the BPI ligand binds to PD-L1_D1D2 with a lower affinity than PD-L1_D1. An interface between D1 and D2 that holds D2 in an orientation that it would sterically hinder the ligand-induced dimer forming could be a reason for this. Figure 4.18 shows all 185 D1/D2 HADDOCK docking solutions aligned via D1 onto the crystal structure of the PD-L1_D1 dimer bound to a BPI. There are only 5 structures, all from Cluster 5, that show a small steric clash between D2 of the two PD-L1 molecules forming a BPI-induced dimer (Figure 4.18). This suggests that it is not likely that the presence of D2 sterically hinders the formation of the dimer seen in Figure 4.18A.
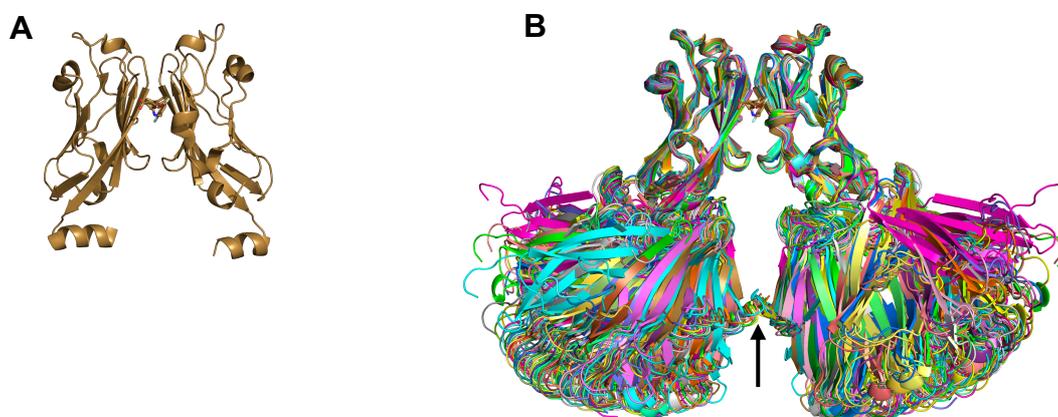
Figure 4.18 Alignment of docking solutions onto ligand-induced PD-L1 dimer.
A) Crystal structure of PD-L1_D1 bound to a BPI ligand. The ligand induces a PD-L1 dimer occluding the PD-1 binding face. PDB: 5J89 **B)** 185 HADDOCK structures aligned via D1 onto the ligand-induced PD-L1_D1 dimer seen in A). Steric clashes of D2 are indicated by an arrow.

4.4 Discussion

As discussed previously, it is not currently known how PD-L1 binding to PD-1 at the cell surface triggers an intracellular signal. The SEC-MALS analysis shown here confirms that the full extracellular region of PD-L1, PD-L1_D1D2, binds to PD-1 forming a 1:1 complex under the conditions tested, suggesting that intracellular signalling is not through dimerization. NMR chemical shift perturbation analysis of PD-L1 when bound to PD-1 revealed widespread chemical shift changes in D1, as well as chemical shift changes in D2, suggesting a conformational switch in the orientation of the domains of PD-L1_D1D2 when binding to PD-1. Overlay of the 2D NMR spectra of PD-L1_D1D2 and PD-L1_D1 or PD-L1_D2, and subsequent chemical shift analysis, reveals an interface between the two domains, which appears to be altered by the binding of PD-1. Since PD-L1 has been shown to bind B7-1 both in *trans* and in *cis*, it is possible that the interface between D1 and D2, suggested here by NMR chemical shift perturbation data, is a mechanism through which PD-L1 is able to switch the orientation of its binding domain.

NOEs between backbone amides of the two domains of PD-L1 were not observed, likely due to the large distances. Given more time other methods could have been used to try and define the interface between the two domains and/or establish their relative orientations. Using Paramagnetic Relaxation Enhancements (PREs) is one such method. This method can be used to obtain distance restraints of approximately 20 Å using spin labels such as MTSL (1-Oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) (Methanethiosulfonate) or IASL (3-(2-iodoacetamido-2,2,5,5,tetramethyl-1-pyrrolidinyloxy radical (iodoacetamido-PROXYL) (Battiste & Wagner 2000; Huang *et al*. 2016). These labels contain radicals and the unpaired electron causes line broadening of surrounding NMR signals in a distance dependant manner. Therefore, modifying a protein with a spin-label at a specific site and measuring changes in peak heights allows the determination of distances from this site (Battiste & Wagner 2000). Labelling of PD-L1 with spin-labels at a number of sites on both D1 and D2 would allow the determination of longer distance restraints which could provide more information for improving the HADDOCK docking. There are a number of considerations, however. Spin labels such as MTSL and IASL

require a free surface-exposed cysteine residue to be attached to the protein (Berliner *et al*. 1982) (Figure 4.19). PD-L1 does not contain such a cysteine so a mutation would have to be made. Since this work uses *E. coli* to express PD-L1 and it is subsequently refolded *in vitro* mutation of a residue to cysteine could lead to a scenario where incorrect disulphide bond formation would cause incorrect folding of the protein.

A   **B**
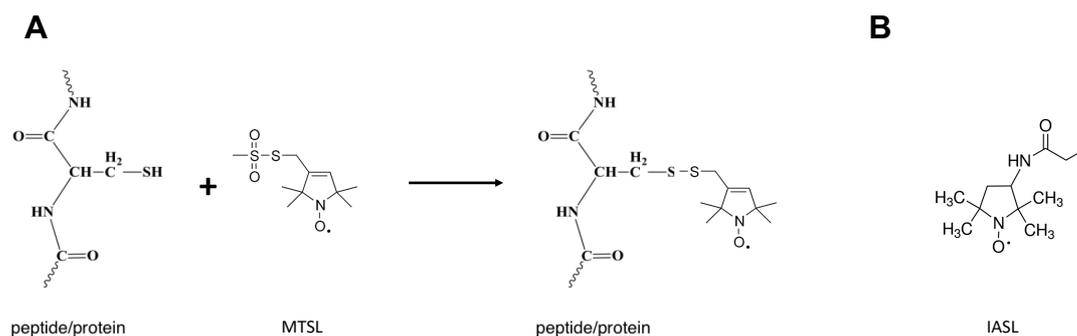
Figure 4.19. Sulfhydryl specific spin labels.

A) shows the attachment of MTSL to a cysteine residue. **B)** shows the chemical structure of IASL.

Alternatively, determination of residual dipolar couplings (RDCs) of PD-L1_D1D2 would provide information on the orientation of the two domains in relation to each other (Al-Hashimi *et al*. 2000). These could be used as restraints to improve docking using HADDOCK. Residual dipolar coupling between two nuclei spins occurs when there is alignment between molecules, leading to incomplete averaging of anisotropic interactions (Bax 2003). In solution, rotational Brownian motion of molecules averages dipolar couplings to zero, so to measure RDCs of a protein, alignment of the molecules must be induced (Bax 2003). A number of methods have been used to align protein molecules for the measurement of RDCs. These include the use of lipid bicelles, filamentous bacteriophages and stressed polyacrylamide gels (Tjandra and Bax 1997; Clore *et al*. 1998; Sass *et al*. 2000). RDCs allow the orientation of the bond between two nuclei to be determined, relative to the external magnetic field thus providing orientational restraints throughout the entire molecule.

Chapter 5

Final Conclusions

The use of antibody therapeutics targeting immune checkpoint molecules has been a huge success in the treatment of cancer in recent years, providing patients with improved responses and increased overall survival (Fessas *et al.* 2017). Antibodies targeting the PD-1/PD-L1 immune checkpoint in particular have achieved great success with five licensed therapeutics targeting this interaction at the time of writing. Anti-PD-1 and anti-PD-L1 drugs are used in the treatment of an ever-growing number of cancers including melanoma and NSCLC. In contrast, the development of small-molecular weight inhibitors targeting this interaction remains in the early stages. There are several advantages of developing small molecule drugs over therapeutic antibodies including lower production cost, the ability to be administered orally as opposed to injection, improved solid tumour penetration, and reduced immune-related adverse effects (Michot *et al.* 2016). Adverse side effects of antibody therapeutics targeting the PD-1/PD-L1 interaction mainly affect the gut, skin, endocrine glands and lungs with patients experiencing fatigue, nausea, decreased appetite, itching, diarrhoea and pneumonitis (Michot *et al.* 2016). For example, during a phase I trial of the anti-PD-1 antibody nivolumab in patients with multiple cancers, 22% of patients experienced Grade 3 or Grade 4 adverse side effects with pneumonitis being the most common (Topalian *et al*. 2012). Immune-related adverse effects of anti-PD-1 and anti-PD-L1 antibodies are associated with their high affinity and long half-life resulting in prolonged target occupancy. The development of a small-molecule inhibitor of the PD-1/PD-L1 interaction could improve outcomes for patients by reducing severe immune-related adverse side effects by modulating the half-life of the drug.

The series of bi-phenyl inhibitor (BPI) compounds patented by Bristol-Myers Squibb were previously shown to inhibit the PD-1/PD-L1 interaction *in vitro* with IC50s in the nM range, but had no cellular activity. The work contained in Chapter 3 of this thesis uses NMR chemical shift perturbation analysis to

show that these compounds bind to the GFCC' β-sheet face of PD-L1, with one BPI binding to a PD-L1_D1 dimer. This is consistent with a crystal structure of the IgV-like domain of PD-L1 bound to a BMS BPI compound that was published at the same time (Figure 5.1).
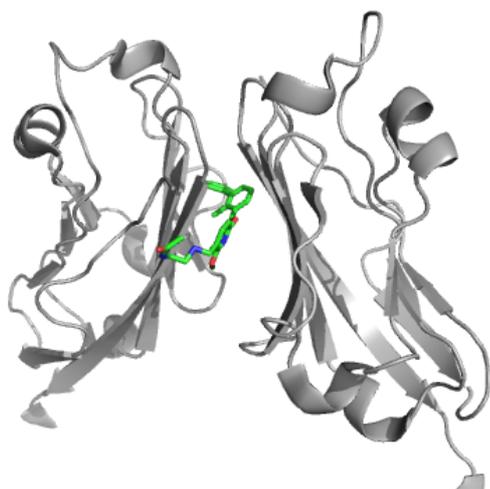
Figure 5.1. BMS compound cause PD-L1_D1 dimerisation.

This shows the crystal structure of a BMS compound (green) bound to PD-L1_D1 (grey). PDB: 5J89. One BMS compound is bound to a dimer of PD-L1_D1 formed via the GFCC' β-sheet face.

This crystal structure showed that binding of the BPI induced dimerization of PD-L1_D1 via the GFCC' β-sheet, blocking PD-1 from binding (Zak *et al*. 2016). Since these compounds have no cellular activity, and on the cell membrane PD-L1 consists of two Ig-like domains, the binding of a BPI to the full extracellular region of PD-L1, PD-L1_D1D2, was investigated. NMR was used to measure the drop in signal intensity as titration of BPI ligand to PD-L1_D1D2 caused PD-L1 to precipitate. This work has determined that the BPI ligand appears to bind to the full extracellular region of PD-L1 with an approximate Kd of 60µM in solution. This is much reduced from the Kd values in the low nM range determined previously for the single IgV-like domain of PD-L1, PD-L1_D1. It was thought that the presence of the second Ig-like domain could sterically hinder the formation of the ligand-induced PD-L1 dimer (Figure 5.1). Presented in Chapter 4 of this thesis is the identification of a potential interface between the two domains of PD-L1. HADDOCK docking based on chemical shift perturbations presented here shows that a D1/D2 interface would not sterically hinder the formation of the ligand-induced PD-L1 dimer. Bristol-Myers Squibb have since released a second patent of improved PD-1/PD-L1 modulating compounds which are larger and an example of which is shown in Figure 5.2.

A
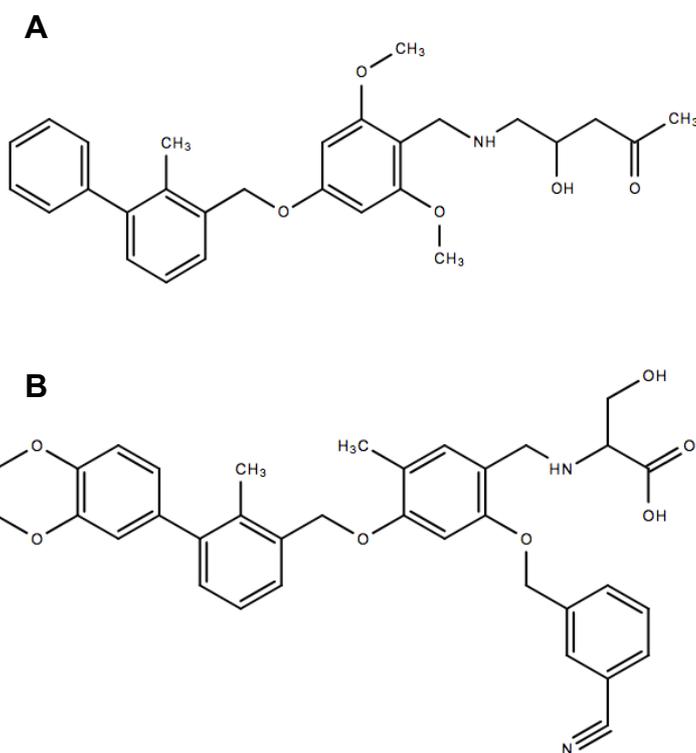
Figure 5.2. Structure of BMS compounds.
A) Structure of the bi-phenyl derived compound used in this work. **B)** Structure of a modified BMS compound that has been shown to have cellular activity.

B

These new compounds, based on the series of compounds used in this work, have been shown to have cellular activity. These compounds were tested in T-cell activation assays where the blocking of PD-L1, expressed on antigen-presenting cells, binding to PD-1, expressed on T-cell-like Jurkat cells, results in the expression of a luciferase reporter gene (Skalniak *et al.* 2017). The modified compounds tested were found to have EC50 values of around 300nM compared to the old compounds showing no activity, and anti-PD-1 and anti-PD-L1 antibodies having EC50 values ranging from 0.3 to 1.2 nM. However, in assays where the ability of the compounds to block soluble PD-L1_D1 from binding to PD-1 expressing Jurkat cells the new compounds performed as well as the anti-PD-1 antibody nivolumab, and the old compounds also showed a good level of inhibitory activity (Skalniak *et al.* 2017). The conflicting nature of these results was not commented on in the paper. The results presented in Chapter 3 show that the BMS compounds have a much lower affinity for PD-L1_D1D2 than they do for PD-L1_D1. These new compounds do show cellular activity and this is probably due to the modified compounds having a higher affinity for the PD-

L1_D1D2 monomer. It would be interesting to see further study of these compounds as structural information about the way they bind to the full extracellular region of PD-L1 would shed light on this. Any future determination of the affinity and activity of small-molecule inhibitors of PD-L1 should be performed using the full extracellular region of PD-L1.

The work set out in Chapter 4 of this thesis sought to investigate the structural features of the extracellular region of PD-L1 in a bid to better understand the impact of the membrane-proximal IgC-like domain of PD-L1 on ligand binding, and also the mechanism through which PD-1 signalling occurs. SEC-MALS analysis determined that, unlike other B7-family molecules, PD-1 and PD-L1 are monomers. They bind in a 1:1 complex in solution therefore it is unlikely that PD-1 signalling is initiated by a PD-L1 dimer bringing the intracellular domains of PD-1 together. NMR investigation of the two extracellular domains of PD-L1 in isolation and the full extracellular region of PD-L1 revealed an interface between D1 and D2 of PD-L1. NMR chemical shift perturbation analysis of the binding of PD-1 to PD-L1_D1D2 shows that as well as chemical shift changes at the PD-1 binding site in D1, there are more widespread chemical shift changes in D1 and also in D2, including potential augmentation of a D1/D2 interface. Recently it has been shown that PD-L1 can interact with both PD-1 and B7-1 (CD80) on the same membrane (Zhao *et al.* 2018; Chaudri *et al.* 2018). Here I propose that the D1/D2 interface may be a mechanism through which PD-L1 switches the orientation of its ligand-binding domain in relation to the cell membrane in order to interact with binding partners both *in cis* and *in trans.* Further investigation into the mechanisms through which PD-L1 interacts with its binding partners may lead to the identification of allosteric sites for the development of drugs targeting PD-L1, as well as different ways of modulating the PD-1/PD-L1 interaction for therapeutic use.

Appendix

A.1 Sequence specific resonance assignments for PD-L1_D1D2

Resonance Assignment

Residue	CA	CB	H	N
F19	57.35	36.27	-	-
T20	57.27	68.73	6.72	122
V21	57.42	29.88	9.445	127.2
T22	56.97	68.63	8.849	118.5
V23	53.48	30.01	8.54	115.8
P24	-	-	-	-
K25	-	-	-	-
D26	50.91	37.7	-	-
L27	51.48	40.4	7.068	123.1
Y28	53.97	37.83	9.161	126.8
V29	58.72	28.39	8.326	123.7
V30	56.5	31.78	8.844	125.2
E31	52.68	27.45	8.46	123.3
Y32	57.08	36.56	8.762	126.4
G33	41.97	-	9.238	118.9
S34	54.97	61.56	7.776	116.8
N35	48.56	36.26	8.132	115.4
M36	51.29	33	8.819	121.6
T37	58.24	67.13	8.062	120.2
I38	56.65	36.36	8.646	123.1
E39	55.17	26.21	7.085	121.7
C40	-	-	-	-
K41	-	-	-	-
F42	-	-	-	-
P43	-	-	-	-
V44	-	-	-	-
E45	-	28.39	-	-
K46	59.72	30.73	7.939	121.3

Residue	CA	CB	H	N
Q47	53.64	25.85	8.131	119.4
L48	52.34	39.36	8.389	126.4
D49	49.29	37.13	8.584	127
L50	54.75	38.64	8.504	125.6
A51	51.01	14.65	7.937	118.1
A52	48.39	16.62	7.402	119.1
L53	52.31	40.84	7.17	121.1
I54	58.03	37.77	8.947	128.6
V55	57.87	32.31	8.919	124.3
Y56	51.75	37.92	9.224	125.5
W57	53.29	30.46	8.701	120
E58	51.47	31.37	9.01	122
M59	53.25	30.67	8.433	121.6
E60	-	-	-	-
D61	51.76	36.74	-	-
K62	52.52	30.01	8.426	120.9
N63	52.16	36.42	8.961	122.7
I64	60.21	35.25	8.464	124.3
I65	55.5	38.12	7.546	113.3
Q66	52.84	27.3	8.521	126.3
F67	54.14	38.5	9.293	129.3
V68	57.84	30.87	8.258	125.6
H69	53.2	24.77	8.913	125
G70	42.5	-	7.88	131.4
E71	51.26	29.72	6.947	118.8
E72	53.65	27.85	8.884	123.3
D73	49.99	37.96	8.641	124.6
L74	52.53	37.61	8.567	125.3
K75	56.32	28.2	8.31	119.5
V76	57.67	28.27	7.086	110.3
Q77	53.53	27.29	6.817	122.8
H78	56.95	28.42	9.218	129.5
S80	58.01	60	-	-

Residue	CA	CB	H	N
Y81	56.67	36.44	8.087	119
R82	53.29	30.76	6.645	119
Q83	-	-	-	-
R84	53.64	28.82	7.982	116.9
A85	47.02	19.83	7.428	120.2
R86	51.79	29.99	8.737	120.2
L87	49.96	40.16	8.81	124.1
L88	50.96	34.59	8.44	126.3
D90	53.4	36.29	-	-
Q91	53.41	25.16	7.383	116.2
L92	55.35	38.05	7.451	121.5
S93	58.05	59.82	7.687	110.9
L94	50.66	37.9	7.151	119.3
G95	42.3	-	7.874	109.4
N96	48.49	37.33	7.01	116.9
A97	46.89	15.36	9.336	130.2
A98	47.86	16.7	7.726	129.6
L99	49.58	38.57	8.236	125.8
Q100	50.42	29.24	8.805	127.3
I101	56.54	36.31	8.784	127.7
T102	57.36	67.53	8.328	119.2
D103	51.32	35.68	7.839	119.7
V104	61.19	28.57	8.247	118.2
K105	50.75	31.47	9.204	128.2
L106	-	-	-	-
Q107	54.18	24.53	-	-
D108	51.78	37.57	7.511	117.7
A109	49.46	16.23	7.283	122.8
G110	41.8	-	8.584	110
V111	59.53	28.69	8.617	121.9
Y112	54.16	37.9	9.986	130.8
R113	51.71	30.98	9.316	120.8
C114	49.16	40.17	8.426	122.5

Residue	CA	CB	H	N
M115	51.7	30.63	8.751	125.2
I116	56.58	36.43	8.202	125.3
S117	52.39	61.52	8.68	120.3
Y118	54.72	33.9	8.221	128.5
G119	43.5	-	-	-
G120	40.34	-	6.879	110.1
A121	48.56	18.63	8.12	120
D122	50.78	40.9	8.61	117.8
Y123	52.55	39.18	8.073	116.7
K124	50.44	35.51	8.286	118.7
R125	52.02	29.38	7.906	117.7
I126	57.9	40.91	9.381	122.8
T127	60.81	66.54	8.735	124.8
V128	58.37	31.05	9.284	128.1
K129	52.09	31.53	9.244	129.3
V130	57.66	30.42	7.95	121.9
N131	48.59	37.77	8.765	125.8
A132	-	-	-	-
P133	-	-	-	-
Y134	53.73	32.18	-	-
N135	49.58	34.6	6.966	109.4
K136	51.88	27.03	8.061	125.1
I137	58	36.23	7.909	124.5
N138	49.62	36.73	9.481	129
Q139	50.41	30.84	8.855	121.6
R140	53.02	30.61	8.577	121.2
I141	58.44	37.38	8.494	124.1
L142	50.24	42.31	8.76	128
V143	59.53	27.81	8.637	124.7
V144	60	28.89	8.355	126.7
D145	-	-	7.088	116.5
P146	61.01	29	-	-

Residue	CA	CB	H	N
V147	62.69	27.63	8.155	119.7
T148	58.17	66.01	7.018	107.8
S149				
E150	-	-	7.975	125.6
H151	-	-	9.439	127.8
E152	52.26	29.57	9.298	119.9
L153	51.12	39.12	-	-
T154	58.39	69.03	9.129	117.3
C155	51.93	44.32	9.082	121.3
Q156	51.57	28.35	8.92	124.6
A157	49.19	20.01	9.067	127.5
E158	50.56	29.84	8.347	118.3
G159	42.1	-	7.955	108.1
Y160	55.23	40.08	8.738	120.7
P161	-	-	-	-
K162	-	-	-	-
A163	47.16	27.6	-	-
E164	51.05	28.7	7.926	118.3
V165	57.38	28.47	8.073	120.5
I166	56.64	37.54	8.99	129.5
W167	53.25	28.77	9.34	132.6
T168	55.89	70.23	9.344	112
S169	53.76	63.04	8.577	113.6
S170	-	-	8.649	116.5
D171	50.25	36.75	7.857	118.6
H172	54.63	24.1	7.733	113.4
Q173	52.71	26.05	7.684	119.3
V174	59.94	28.52	8.383	123.6
L175	50.19	40.49	8.658	130
S176	-	-	8.461	116.7
G177	-	-	-	-
K178	-	-	-	-
T179	-	-	-	-

Residue	CA	CB	H	N
T180	-	-	-	-
T181	-	-	-	-
T182	-	-	-	-
N183	-	-	-	-
S184	-	-	-	-
K185	-	-	-	-
R186	-	-	-	-
E187	-	-	-	-
E188	-	-	-	-
K189	52.59	30.76	9.285	123.7
L190	52.38	43.03	8.516	120.3
F191	55.73	37.35	8.946	121.7
N192	50.05	36.76	9.069	119.4
V193	59.33	29.99	9.204	130
T194	57.94	67.63	8.588	122
S195	53.31	62.16	9.427	121.3
T196	57.54	69.41	8.734	122.4
L197	50.21	39.81	9.075	129.8
R198	51.24	27.82	8.043	127.6
I199	56.19	38.97	8.522	122.6
N200	-	-	-	-
T201	-	-	8.09	119.9
T202	57.06	67.82	8.372	115.2
T203	61.29	66.8	7.611	116.2
N204	41.44	-	9.113	119.2
E205	55.12	26.16	7.557	120.9
I206	57.44	37.52	7.452	120.1
F207	53.22	41.17	8.75	123.9
Y208	53.35	36.67	9.081	116.9
C209	49.84	39.16	8.359	116.8
T210	58.12	67.7	8.63	121
F211	53.68	36.03	9.709	128.2
R212	51.8	29.37	9.172	128.6

Residue	CA	CB	H	N
R213	51.79	29.79	8.541	127.4
L214	52.01	39.47	8.183	121
D215	50.13	38.07	7.485	115.4
P216	-	-	-	-
E217	-	-	-	-
E218	-	-	-	-
N219	49.32	39.04	-	-
H220	53.66	30.61	8.311	121.4
T221	59.16	69.06	8.722	121.4
A222	47.68	18.24	8.768	129.8
E223	51.92	29.23	8.205	121.2
L224	52.98	39.78	9.448	131
V225	57.05	29.59	8.196	122.4
I226	54.17	31.58	8.598	127.8
P227	59.36	29.39	-	-
E228	53.97	26.57	8.368	120.1
L229	50.44	38.13	8.144	126.7
P230	59.75	28.27	-	-
L231	51.93	38.76	8.108	122.5
A232	-	-	8.088	125.1
H233	-	-	-	-
P234	-	-	-	-
P235	-	-	-	-
N236	-	-	-	-
E237	-	-	-	-
R238	53.03	27.13	8.239	122.6
T239	60.19	67.59	7.715	121.1

A.2 Sequence specific resonance assignments for PD-L1_D1

Resonance Assignment

Residue	C'	CA	CB	H	N
A18	-	53.69	19.98	7.77	130.3
F19	174.1	60.26	39.58	-	-
T20	172.3	60.26	71.9	6.748	121.9
V21	175.3	60.38	33.32	9.376	126
T22	172.9	59.75	71.64	8.846	118.2
V23	-	56.3	33.52	8.68	105.7
P24	-	-	-	-	-
K25	-	-	-	-	-
D26	175.1	53.61	40.75	-	-
L27	174.9	54.24	43.99	7.059	122.7
Y28	173.5	56.75	40.95	9.141	126.8
V29	176.2	61.69	32	8.308	123.4
V30	174.3	59.13	35.26	8.783	123.9
E31	177	55.07	30.04	8.432	120.6
Y32	176.8	59.96	38.26	8.369	125.2
G33	174.1	44.91	-	9.259	118.1
S34	171.4	57.82	64.51	7.561	115.7
N35	174.7	51.36	39.27	8.174	115.2
M36	173.1	54.21	36.35	8.813	121.8
T37	173.4	61.05	70.05	8.096	120.2
I38	174.8	59.49	39.64	8.65	122.7
E39	-	58.1	29.7	7.085	121.4
C40	-	-	-	-	-
K41	-	-	-	-	-
F42	-	-	-	-	-
P43	-	-	-	-	-
V44	-	-	-	-	-
E45	-	-	-	-	-
K46	175.6	57.77	32.27	-	-

Residue	C'	CA	CB	H	N
Q47	175.1	55.44	29.6	8.134	118.7
L48	175.8	55.27	42.96	8.424	126.2
D49	176.7	52.12	40.23	8.606	126.8
L50	178.3	57.62	42.21	8.535	125.3
A51	177.9	53.9	17.94	7.952	117.9
A52	176.6	51.28	19.97	7.397	118.8
L53	175.9	55.14	44.44	7.171	120.8
I54	174.8	60.91	41.29	8.975	128.6
V55	174.3	60.78	35.78	8.907	124.1
Y56	175.2	54.61	41.15	9.237	125.4
W57	175.6	56.05	33.51	8.701	119.7
E58	174.1	54.31	34.82	9.024	121.7
M59	175.3	56.13	34.24	8.429	121.2
E60	171.9	55.24	46.1	8.568	120.2
D61	174.8	54.77	34.07	9.301	123.3
K62	175.2	54.8	35.66	8.792	118.7
N63	173.9	55.03	39.51	8.99	122.4
I64	174.6	63.19	38.78	8.487	124.2
I65	174.9	58.45	41.62	7.539	112.8
Q66	172.7	55.66	30.74	8.512	126
F67	173.1	57.01	41.7	9.311	129.1
V68	174.7	60.78	34.4	8.263	125.3
H69	176	56.05	27.92	8.953	124.8
G70	172.6	45.57	-	8.004	102.7
E71	174.8	54.19	33.18	6.948	118.3
E72	175.9	56.58	31.19	8.916	122.9
D73	176.8	52.86	41.09	8.678	124.4
L74	179.2	55.34	41.15	8.611	125
K75	178.3	59.34	31.8	8.32	119.2
V76	175.3	60.55	31.64	7.115	109.7
Q77	176.2	56.36	30.82	6.82	122.7
H78	177.8	59.82	31.76	9.259	129.5
S79	176.6	61.88	62.26	8.278	122.9

Residue	C'	CA	CB	H	N
S80	175.2	60.99	62.94	9.585	120.2
Y81	174.7	59.47	39.66	8.106	118.6
R82	174.9	56.15	34.18	6.601	118.6
Q83	176	58.07	27.2	8.621	122.6
R84	174.2	56.43	32.54	8	116.4
A85	176.4	49.93	23.31	7.432	119.9
R86	173.2	54.67	33.5	8.737	120
L87	177	52.79	43.74	8.825	123.9
L88	178.2	53.86	38.09	8.451	126.1
K89	-	59.8	31.98	8.379	128
D90	177.5	56.13	39.21	-	-
Q91	177.9	56.32	28.55	7.392	115.9
L92	179.4	58.23	41.61	7.458	121.2
S93	174.1	61.01	62.54	7.717	110.6
L94	177.6	53.67	41.45	7.155	119
G95	172.2	45.27	-	7.894	109.1
N96	174.1	51.3	40.51	7.008	116.5
A97	173.3	49.86	18.46	9.349	130.1
A98	174.1	50.64	20.06	9.021	128
L99	173.4	52.39	42.18	8.251	125.7
Q100	175.2	53.21	32.71	8.813	127.1
I101	175.2	59.35	39.88	8.81	127.5
T102	173.5	60.21	70.51	8.313	118.8
D103	174.8	54.16	38.86	7.908	119.6
V104	176.2	64.03	32.08	8.362	118
K105	177.8	53.29	35.15	9.278	126.5
L106	-	58.98	40.19	8.269	120.1
Q107	175.3	57.72	27.2	-	-
D108	176.1	54.64	40.88	7.658	117.2
A109	175.8	52.55	19.87	7.279	122.6
G110	121.5	44.76	-	8.616	109.4
V111	175.8	62.49	32.14	8.636	129.7
Y112	175.2	56.85	41.04	10.01	117.3

Residue	C'	CA	CB	H	N
R113	174.5	54.52	34.56	9.318	120.5
C114	172	51.87	42.81	8.434	122.1
M115	174.1	54.59	33.99	8.748	124.9
I116	173.9	59.45	39.89	8.2	125
S117	173.4	55.13	64.33	8.673	120
Y118	173.4	57.58	37.08	8.244	128.2
G119	174.7	46.44	-	8.519	115
G120	170	43.34	-	6.917	109.7
A121	176.5	51.36	22.05	8.171	119.7
D122	172.9	53.47	43.98	8.617	117.5
Y123	174	55.34	42.4	8.098	116.3
K124	174.4	53.19	39.16	8.285	118.4
R125	175.9	54.78	32.95	7.918	117.4
I126	175	60.77	44.37	9.402	122.6
T127	172.8	63.66	69.48	8.727	124.7
V128	175.9	61.25	34.48	9.299	127.8
K129	173.8	54.93	35	9.274	129.7
V130	175.5	60.06	34.05	7.91	122.1
N131	173.8	52	39.17	8.694	126.1
A132	-	50.27	17.84	8.202	123.7
P133	175.3	62.99	31.34	-	-
Y134	-	58.61	38.85	7.393	124

A.3 Sequence specific resonance assignments for PD-L1_D2

Resonance Assignment

Residue	C'	CA	CB	H	N
Y134	-	56.81	35.39	-	-
N135	172.8	52.79	38	6.989	109.1
K136	176.1	54.82	30.98	8.075	124.5
I137	174	60.89	39.89	7.982	124.6
N138	176.4	52.46	40.02	9.52	129
Q139	174.3	53.3	34.49	8.909	121.3
R140	174.9	56.04	34.3	8.583	120.9
I141	173.7	61.38	41.06	8.531	123.9
L142	174.2	53.08	46.06	8.767	127.9
V143	175.7	62.63	31.4	8.673	124.5
V144	175.5	63.09	32.43	8.386	126.5
D145	176.3	51.08	43.31	7.087	116.2
P146	-	63.95	32.38	-	-
V147	175.8	65.77	31.26	8.18	119.5
T148	178	61	69.1	6.999	107.3
S149	175.8	60.05	62.9	7.797	116.6
E150	175.6	57.36	30.44	7.979	125.3
H151	175.8	55.57	32.97	9.462	127.7
E152	172.9	55.17	33.13	9.305	119.6
L153	174.4	54	42.82	9.377	132.5
T154	174.9	61.33	72.02	9.144	117.1
C155	174.2	54.84	47.28	9.094	121.1
Q156	172.5	54.45	31.72	8.943	124.5
A157	172.6	52.19	23.63	9.104	127.3
E158	174.5	53.46	33.46	8.382	117.9
G159	175.6	45.02	-	8.074	108.3
Y160	170.7	58.59	-	8.84	121.2
P161	-	-	-	-	-
K162	-	-	-	-	-
A163	-	50.16	21.52	-	-

Residue	C'	CA	CB	H	N
E164	176.1	53.98	32.28	7.972	118.2
V165	174.7	60.37	32.01	8.109	120.3
I166	175.5	59.65	41.17	9.01	129.3
W167	176.2	56.11	31.83	9.396	132.9
T168	176.3	58.66	73.36	9.355	111.7
S169	174.6	56.62	66	8.589	113.1
S170	176.7	61.07	62.37	8.663	116.2
D171	174.4	53.06	39.84	7.868	118.3
H172	175.3	57.58	27.39	7.752	113
Q173	174.6	55.59	29.45	7.702	119
V174	175.5	63.12	32.07	8.434	123.4
L175	176.4	53.21	44.08	8.682	130
S176	176.2	-	-	8.484	116.5
G177	-	-	-	-	-
K178	-	-	-	-	-
T179	-	-	-	-	-
T180	-	-	-	-	-
T181	-	-	-	-	-
T182	-	-	-	-	-
N183	-	-	-	-	-
S184	-	-	-	-	-
K185	-	-	-	-	-
R186	-	-	-	-	-
E187	-	-	-	-	-
E188	-	-	-	-	-
K189	-	-	-	-	-
L190	-	-	-	-	-
F191	-	57.85	41.05	-	-
N192	174.3	52.99	40.01	9.225	119.2
V193	174.7	62.32	33.67	9.224	130
T194	174.1	60.71	70.67	8.619	121.7
S195	172.9	56.15	65.17	9.425	121.2
T196	173.5	60.49	72.51	8.748	122.1

Residue	C'	CA	CB	H	N
L197	172.4	53.08	43.45	9.065	129.5
R198	174	54.2	31.49	8.067	127.4
I199	174.8	59.15	42.58	8.54	122.4
N200	172.6	53.62	38.05	8.284	125.6
T201	173.8	59.8	67.69	8.097	119.7
T202	171.8	60	70.76	8.388	114.8
T203	171.8	64.4	69.92	7.64	115.9
N204	174.3	55.43	37.19	9.13	119
E205	175	58.07	29.68	7.562	120.7
I206	174	60.42	41.13	7.478	119.8
F207	174.3	56.09	44.53	8.766	123.7
Y208	172.7	56.18	40	9.078	116.5
C209	175.1	52.61	42.22	8.361	116.3
T210	172.5	61.11	70.82	8.629	120.7
F211	172.8	56.66	39.25	9.707	128.2
R212	174.5	54.78	33.17	9.183	128.5
R213	174.9	54.78	33.39	8.552	127.7
L214	174.5	55	43.35	8.173	120.9
D215	176.9	53.19	41.24	7.632	115.1
P216	-	62.72	34.15	-	-
E217	176	56.75	29.54	8.253	120.1
E218	177	56.34	33.62	8.229	126.3
N219	-	53.33	38.85	-	-
H220	-	56.62	34.26	8.274	121.1
T221	174.1	62.07	72.12	8.764	121.1
A222	173.5	50.66	21.9	8.768	129.6
E223	173.8	54.89	32.8	8.238	120.9
L224	174.8	55.99	43.51	9.469	131
V225	173.8	60.02	33.21	8.236	122.2
I226	176.1	57.14	35.02	8.613	127.7
P227	-	62.37	33.01	-	-
E228	176.6	56.88	30.13	8.406	119.8
L229	176.2	53.35	41.73	8.175	126.6

Residue	C'	CA	CB	H	N
P230	-	62.75	31.87	-	-
L231	176.6	54.9	42.44	8.157	122.3
A232	176.9	52.02	19.52	8.128	124.9
H233	175.9	54.24	33.4	7.977	120
P234	-	-	-	-	-
P235	-	-	-	-	-
N236	-	-	-	-	-
E237	-	-	-	-	-
R238	-	56.01	30.73	-	-
T239	175.6	63.13	70.58	7.741	120.8

References

Agata Y, Kawasaki A, Nishimura H, Ishida Y, Tsubata T, Yagita H, Honjo T (1996). **Expression of the PD-1 antigen on the surface of stimulated mouse T and B lymphocytes.** *Int. Immunol.* 5: 765-772.

Al-Hashimi HM, Valafar H, Terrell M, Zartler ER, Eidsness MK, Prestegard JH (2000). **Variation of Molecular Alignment as a Means of Resolving Orientational Ambiguities in Protein Structures from Dipolar Couplings.** *J. Mag. Res.* 143: 402-406.

Ayyar BV, Arora S, O'Kennedy R (2016). **Coming-of-age of antibodies in cancer therapeutics.** *Trends in Pharm. Sci.* 37(12): 1009-1028.

Back J, Chalifour A, Scarpellino L, Held W (2007). **Stable masking by H-2Dd cis ligand limits Ly49A relocalization to the site of NK cell/target cell contact.** *PNAS* 104: 3978-3983.

Back J, Malchiodi EL, Cho S, Scarpellino L, Schneider P, Kerzic MC, Mariuzza RA, Held W (2009). **Distinct conformations of Ly49 natural killer cell receptors mediate MHC class I recognition in trans and cis.** *Immunity* 31: 598-608.

Bally APR, Austin JW, Boss JM (2016). **Genetic and epigenetic regulation of PD-1 expression.** *J. Immunol.* 196: 2431-2437.

Battiste JL & Wagner G (2000). **Utilization of Site-Directed Spin Labeling and High-Resolution Heteronuclear Nuclear Magnetic Resonance for Global Fold Determination of Large Proteins with Limited Nuclear Overhauser Effect Data.** *Biochemistry* 39(18): 5355-5365.

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

Berliner LJ, Grunwald J, Hankovszky HO, Hideg K (1982). **A novel reversible thiol-specific spin label: papain active site labeling and inhibition.** *Anal. Biochem.* 119: 450-455.

Bhatia S, Edidin M, Almo SC, Nathenson SG (2005). **Different cell surface oligomeric states of B7-1 and B7-2: Implications for signalling.** *PNAS* 43: 15569-15574.

Bogan AA & Thorn KS (1998). **Anatomy of hotspots in protein interfaces.** *J. Mol. Biol.* 280: 1-9.

Bour-Jordan H, Esensten JH, Martinez-Llordella M, Penaranda C, Stumpf M, Bluestone JA. (2011). **Intrinsic and extrinsic control of peripheral T-cell tolerance by costimulatory molecules of the CD28/B7 family**. *Immunol. Rev.* 241: 180-205.

Bretscher P (1999). **A two-step, two-signal model fro the primary activation of precursor T-cells.** *PNAS* 96: 185-190.

Buchwald P (2008). **Glucocorticoid receptor binding: a biphasic dependence on molecular size as revealed by the bilinear LinBiExp model.** *Steroids* 73: 193-208.

Bundi A & Wuthrich K (1977). **H-1-NMR titration shifts of amide proton resonances in polypeptide chains.** *Febs Letters* 77: 11-14.
Bundi A & Wuthrich K (1979). **H-1-NMR parameters of the common amino-acid residues measured in aqueous-solution of the liner tetrapeptides H-Gly-Gly-X-L-Ala-OH**. *Biopolymers* 18: 285-297.

Burnet M (1957). **Cancer: a biological approach.** *BMJ* 1: 841-847.

Chaudri A, Xiao Y, Klee AN, Wang , Zhu B, Freeman GJ (2018). **PD-L1 Binds to B7-1 Only In Cis on the Same Cell Surface.** *Cancer Immunol. Res.* 6 (8): 921-929.

Chemnitz JM, Parry RV, Nichols KE, June CH, Riley JL (2004). **SHP-1 and SHP-2 associate with immunoreceptor tyrosine-based switch motif of programmed death 1 upon primary human T cell stimulation, but only receptor ligation prevents T cell activation.** *J. Immunol.* 2: 945-954.

Chen Y, Liu P, Gao F, Cheng H, Qi J, Gao GF (2010). **A dimeric structure of PD-L1: functional units or evolutionary relics?** *Protein Cell* 1(2): 153-160.

Cheng X, Veverka V, Radhakrishnan A, Waters LC, Muskett FW, Morgan SH, Huo J, Yu C, Evans EJ, Leslie AJ, Griffiths M, Stubberfield C, Griffin R, Henry AJ, Jansson A, Ladbury JE, Ikemizu S, Carr MD, Davis SJ (2013). **Structure and interactions of the human programmed cell death 1 receptor.** *J. Biol. Chem.* 288 (17): 11771-11785.

Clackson T & Wells JA (1995). **A hotspot of binding energy in a hormone-receptor interface.** *Science* 267: 383-386.

Clore GM, Starich MR, Gronenborn AM (1998). **Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses**. *J. Am. Chem. Soc.* 120: 10571-10572.

Coley WB (1898). **The treatment of inoperable sarcoma with the mixed toxins of erysipelas and bacillus prodigiosus – immediate and final results in one hundred and forty cases.** *JAMA* 31: 389-421.

Contardi E, Palmisano GL, Tazzari PL, Martelli AM, Falà F, Fabbi M, Kato T, Lucarelli E, Donati D, Polito L, Bolognesi A, Ricci F, Salvi S, Gargaglione V, Mantero S, Alberghini M, Ferrara GB, Pistillo MP (2005). **CTLA-4 is constitutively expressed on tumor cells and can trigger apoptosis upon ligand interaction.** *Int. J. Cancer* 117: 538-550.

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995). **NMRPipe: A multidimensional spectral processing system based on UNIX pipes**. *Journal of Biomolecular NMR* 6(3): 277-293.

Dominguez C, Boelens R, Bonvin AMJJ (2003). **HADDOCK: a protein-protein docking approach based on biochemical and/or biophysical information.** *JACS* 125: 1731-1737.

Dong H, Zhu G, Tamada K, Chen L (1999). **B7-H1, a third member of the B7 family, co-stimulates T-cell proliferation and interleukin-10 secretion.** *Nature Med.* 5(12): 1365-1369.

Doucey MA, Scarpellino L, Zimmer J, Guillaume P, Luescher IF, Bron C, Held W (2004). **Cis-association of Ly49A with MHC class I restricts natural killer cell inhibition.** *Nat. Immunol.* 5: 328-336.

Farkona S, Diamandis EP, Blasutig IM (2016). **Cancer immunotherapy: the beginning of the end for cancer?** *BMC Medicine* 14:73.

Fellner C (2012). **Ipilimumab (Yervoy) Prolongs Survival In Advanced Melanoma.** *Pharmacy & Therapeutics* 9: 503-511.

Fessas P, Lee H, Ikemizu S, Janowitzd T (2017). **A molecular and preclinical comparison of the PD-1–targeted T-cell checkpoint inhibitors nivolumab and pembrolizumab.** *Semin. Oncol.* 44(2): 136-140.

Fischer B, Sumner I, Goodenough P (1993). **Isolation, renaturation and formation of disulfide bonds of eukaryotic proteins expressed in *E. coli* as inclusion bodies.** *Biotech. Bioeng.* 41: 3-13.

Gong J, Chehrazi-Raffle A, Reddi S, Salgia R (2018). **Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: a comprehensive review of registration trials and considerations.** *J. Immuno. Cancer* 6:8.

Greenfield NJ (2006). **Using circular dichroism spectra to estimate protein secondary structure.** *Nat. Protoc.* 1(6): 2876-2890.

Greenfield NJ and Fasman GD (1969). **Computed Circular Dichroism Spectra for the Evaluation of Protein Conformation.** *Biochemistry* 8: 4108-4116.

Grzesiek S, Bax A. (1992). **Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR.** *JACS* 114(16): 6291-6293.

Held W & Mariuzza RA (2011). ***Cis-trans* interactions of cell surface receptors: biological roles and structural basis.** *Cell Mol. Life Sci.* 68: 3469-3478.

Huang S, Umemoto R, Tamura Y, Kofuku Y, Uyeda TQP, Nishida N, Shimada I (2016). **Utilization of paramagnetic relaxation enhancemnts for structural analysis of actin-binding proteins in complex with actin.** *Sci. Reports* 6: 1-9.

Huynh K & Partch CL (2015). **Analysis of protein stability and ligand interactions by thermal shift assay.** *Curr. Protoc. Protein Sci.* 79: 1-14.

Karlhofer FM, Ribaudo RK, Yokoyama WM (1992). **MHC class I alloantigen specificity of Ly-49+ IL-2 activated natural killer cells.** *Nature* 358: 66-70.

Kash N, Lee MA, Kollipara R, Downing C, Guidry J, Tyring SK (2015). **Safety and Efficacy Data on Vaccines and Immunization to Human Papillomavirus**. *Journal of Clinical Medicine* 4(4): 614–33.

Kay LE, Ikura M, Tschudin R, Bax A (1990). **Three-Dimensional Triple-Resonance NMR Spectroscopy of Isotopically Enriched Proteins**. *Journal of Magnetic Resonance.* 89: 496-514.

Kelly AE, Ou HD, Withers R, Dötsch (2002). **Low- conductivity buffers for high-sensitivity NMR measurements.** *J. Am. Chem. Soc.* 124: 12013-12019.

Khan RH, AppaRao KBC, Eshwari ANS, Totey SM, Panda AK (1998). **Solubilization of recombinant ovine growth hormone with retention of native-like secondary structure and its refolding from the inclusion bodies of Escherichia coli**. *Biotech. Progress* 14: 722-728.

Kiefhaber T, Rudolph R, Kohler HH, Buchner J (1991). **Protein aggregation in vitro and in vivo: a quantitative model of the kinetic competition between folding and aggregation.** *Biotechnology* 9: 825-829.

Latchman Y, Wood CR, Chernova T, Chaudhary D, Borde M, Chernova I, Iwai I, Long AJ, Brown JA, Nunes R, Greenfield EA, Bourque K, Boussiotis VA, Carter LL, Carreno BM, Malenkovich N, Nishimura H, Okazaki T, Honjo T, Sharpe AH, Freeman GJ (2001). **PD-L2 is a second ligand for PD-1 and inhibits T cell activation.** *Nat. Immunol.* 2: 261-268.

Lee HT, Lee JY, Lim H, Lee SH, Moon YJ, Pyo HJ, Ryu SE, Shin W, Heo YS (2017). **Molecular mechanism of PD-1/PD-L1 blockade via anti-PD-L1 antibodies atezolizumab and durvalumab.** *Sci. Reports* 7: 5532

Lee JY, Lee HT, Shin W, Chae J, Choi J, Kim SH, Lim H, Won Heo T, Park KY, Lee YJ, Ryu SE, Son JY, Lee JU, Heo YS (2016). **Structural basis of checkpoint blockade by monoclonal antibodies in cancer immunotherapy.** *Nat. Comms.* 7: 13354.

Lee W, Tonelli M, Markley JL (2015). **NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy.** *Bioinformatics* 31(8): 1325-1327.

Lilie H, Schwarz E, Rudolph R (1998). **Advances in refolding of proteins produced in *E. coli.*** *Curr. Opin. Biotech*. 9: 497-501.

Lin DY, Tanaka Y, Iwasaki M, Gittis AG, Su HP, Mikami B, Okazaki T, Honjo T, Minato N, Garboczi DN (2008). **The PD-1/PD-L1 complex resembles the antigen-binding Fv domains of antibodies and T cell receptors.** *PNAS* 105: 2011-3016.

Liu K, Tan S, Chai Y, Chen D, Song H, Zhang CW, Shi Y, Liu J, Tan W, Lyu J, Gao S, Yan J, Qi J, Gao GF (2017). **Structural basis of anti-PD-L1 monoclonal antibody avelumab for tumor therapy**. *Cell Research* 27: 151-153.

Ljunggren HG, Kärre K (1990). **In search of the 'missing self': MHC molecules and NK cell recognition.** *Immunol. Today* 11: 237-244.

Marion D, Kay LE, Sparks SW, Torchia DA, Bax A (1989). **Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins.** *J. Am. Chem. Soc.* 4: 1515-1517.

Masuda A, Nakamura A, Maeda T, Sakamoto Y, Takai T (2007). **Cis binding between inhibitory receptors and MHC class I can regulate mast cell activation.** *J. Exp. Med.* 204: 907-920.

Meijers R, Puettmann-Holgado R, Skiniotis G, Liu JH, Walz T, Wang JH, Schmucker D (2007). **Structural basis of Dscam isoform specificity.** *Nature* 449: 487-491.

Michot JM, Bigenwald C, Champiat S, Collins M, Carbonnel F, Postel-Vinay S, Berdelou A, Varga A, Bahleda R, Hollebecque A, Massard C, Fuerea A, Ribrag V, Gazzah A, Armand JP, Amellal N, Angevin E, Noel N, Boutros C, Mateus C, Robert C, Soria JC, Marabelle A, Lambotte O (2016). **Immune-related adverse events with immune checkpoint blockade: a comprehensive review.** *Eur. J. Cancer* 54: 139-148.

Miller S (1989). **The structure of interfaces between subunits of dimeric and tetrameric proteins.** *Protein Eng.* 3: 77-83.

Morgan DA, Ruscetti FW, Gallo R (1976). **Selective in vitro growth of T lymphocytes from normal human bone marrows.** *Science* 193: 1007-8.

Nakanishi J, Wada Y, Matsumoto K, Azuma M, Kikuchi K, Ueda S (2007). **Overexpression of B7-H1 (PD-L1) significantly associates with tumor grade and postoperative prognosis in human urothelial cancers**. *Cancer Immunol Immunother.* 56: 1173-1182.

Neugebauer A, Hartmann RW, Klein CD (2007). **Prediction of protein-protein interaction inhibitors by chemoinformatics and machine learning methods**. *J. Med. Chem.* 50: 4665-4668.

Okazaki T, Maeda A, Nishimura H, Kurosaki T, Honjo T (2001). **PD-1 immunoreceptor inhibits B cell receptor-mediated signalling by recruiting src homology 2-domain-containing tyrosine phosphatase 2 to phosphotyrosine.** *PNAS* 98: 13866-13871.

Pardoll DM (2012). **The blockade of immune checkpoints in cancer immunotherapy.** *Nat. Rev. Cancer* 12(4): 252-264.

Pervushin K, Riek R, Wider G, Wüthrich K (1997). **Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution.** *PNAS.* 94(23): 12366-12371.

Rosenberg SA, Lotze MT, Muul LM, Leitman S, Chang AE, Ettinghausen SE, Matory YL, Skibber JM, Shiloni E, Vetto JT, Seipp CA, Simpson C, Reichert CM (1985). **Observations on the systemic administration of autologous lymphokine-activated killer cells and recombinant interleukin-2 to patients with metastatic cancer**. *N Engl J Med.* 313(23): 1485-1492.

Sass HJ, Musco G, Stahl SJ, Wingfield PT, Grzesiek S (2000). **Solution NMR of proteins within polyacrylamide gels: Diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes**. *J. Biomo.l NMR.* 18: 303-309.

Rudolph R, Lilie H (1996). ***In vitro* refolding of inclusion body proteins.** *FASEB J* 10: 49-56.

Scott DE, Bayly AR, Abell C, Skidmore J (2016). **Small molecules, big targets: drug discovery faces the protein-protein interaction challenge.** *Nat. Rev. Drug Discovery.* 15: 533-550.

Schwartz JC, Zhang X Fedorov AA, Nathenson SG, Almo SC (2001). **Structural basis for co-stimulation by the human CTLA-4/B7-2 complex.** *Nature* 410 (6828): 604-608.

Sheppard KA, Fitz LJ, Lee JM, Benander C, George JA, Wooters J, Qiu Y, Jussif JM, Carter LL, Wood CR, Chaudhary D (2004). **PD-1 inhibits T-cell receptor induced phosphorylation of the ZAP70/CD3zeta signalosome and downstream signaling to PKCtheta.** *FEBS Lett.* 574:37-41.

Singh N, Frey NV, Grupp SA, Maude SL (2016). **CAR T cell therapy in acute lymphoblastic leukemia and potential for chronic lymphocytic leukemia**. *Curr. Treat. Options Oncol.* 17(6): 28.

Skalniak L, Zak KM, Guzik K *et al.* (2017). **Small molecule inhibitors of PD-1/PD-L1 immune checkpoint alleviate the PD-L1-induced exhaustion of T-cells.** *Oncotarget.* 8: 72167-72181.

Skalniak L, Zak KM, Guzik K, Magiera K, Musielak B, Pachota M, Szelazek B, Kocik J, Grudnik P, Tomala M, Krzanik S, Pyrc K, Dömling A, Dubin G, Holak TA (2017). **Small-molecule inhibitors of PD-1/PD-L1 immune checkpoint alleviate the PD-L1-induced exhaustion of T-cells.** *Oncotarget* 8: 72167-72181.

Thompson RH, Gillett MD, Cheville JC, Lohse CM, Dong H, Webster WS, Krejci KG, Lobo JR, Sengupta S, Chen L, Zincke H, Blute ML, Strome SE, Leibovich BC, Kwon ED (2004). **Costimulatory B7-H1 in renal cell carcinoma patients: Indicator of tumor aggressiveness and potential therapeutic target**. *PNAS* 49: 17174-17179.

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

Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, Powderly JD, Carvajal RD, Sosman JA, Atkins MB, Leming PD, Spigel DR, Antonia SJ, Horn L, Drake CG, Pardoll DM, Chen L, Sharfman WH, Anders RA, Taube JM, McMiller TL, Xu H, Korman AJ, Jure-Kunkel M, Agrawal S, McDonald D, Kollia GD, Gupta A, Wigginton JM, Sznol M (2012). **Safety, activity, and immune correlates of anti-PD-1 antibody in cancer.** *N. Engl. J. Med.* 26: 2443-2454.

van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastritis PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ, Bovin AMJJ (2016). **The HADDOCK2.2 webserver: User-friendly integrative modelling of biomolecular complexes.** *J. Mol. Biol.* 428: 720-725.

Wan T, Beavil RL, Fabiane SM, Beavil AJ, Sohi MK, Keown M, Young RJ, Henry AJ, Owens RJ, Gould HJ, Sutton BJ (2002). **The crystal structure of IgE Fc reveals an asymmetrically bent conformation.** *Nat. Immunol.* 3: 681-686.

Wetlaufer DB, Branca PA, Chen GX (1987). **The oxidative folding of proteins by disulfide plus thiol does not correlate with redox potential.** *Protein Eng.* 2: 141-146.

Williamson MP (2013). **Using chemical shift perturbation to characterise ligand binding.** *Progress in NMR Spec.* 73: 1-16.

Williamson RA, Carr MD, Frenkiel TA, Feeney J, Freedman RB (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.

Wittekind M, Mueller L (1993). **HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins.** *Journal of Magnetic Resonance.* 101(2): 201-205.

Yamazaki T, Akiba H, Iwai H, Matsuda H, Aoki M, Tanno Y, Shin T, Tsuchiya H, Pardoll DM, Okumura K, Azuma M, Yagita H (2002). **Expression of programmed death 1 ligands by murine T cells and APC.** *J. Immunol.* 169(10): 5538-5545.

Zak KM, Grudnik P, Guzik K, Zieba BJ, Musielak B, Dömling A, Dubin G, Holak TA (2016). **Structural basis for small molecule targeting of the programmed death ligand 1 (PD-L1).** *Oncotarget* 7: 30323-30335.

Zak KM, Kitel R, Przetocka S, Golik P, Guzik K, Musielak B, Dömling A, Dubin G, Holak TA (2015). **Structure of the Complex of Human Programmed Death 1, PD-1, and Its Ligand PD-L1.** *Structure* 23 (12): 2341-2348.

Zarganes-Tzitzikas T, Konsatinidou M, Gao Y, Krzemien D, Zak K, Dubin G, Holak TA, Dömling A (2016). **Inhibitors of programmed cell death 1 (PD-1): a patent review (2010-2015).** *Expert Opin. Ther. Pat.* 26(9): 973-977.

Zhang X, Schwartz J-CD, Guo X, Bhatia S, Cao E, Lorenz M, Cammer M, Chen L, Zhang ZY, Edidin MA, Nathenson SG, Almo SC (2004). **Structural and Functional Analysis of the Costimulatory Receptor Programmed Death-1.** *Immunity* 20: 337-347.

Zhao Y, Harrison DL, Song Y, Ji J, Huang J, Hui E (2018). **Antigen-Presenting Cell-Intrinsic PD-1 Neutralizes PD-L1 in cis to Attenuate PD-1 Signaling in T Cells.** *Cell Reports* 24: 379-390.

Zuiderweg ERP (2002). **Mapping protein-protein interactions in solution by NMR spectroscopy.** *Biochemistry* 41: 1-7.

Zuiderweg ERP & Fesik SW (1989). **Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a.** *Biochemistry* 6: 2387-2391.