UNIVERSITY OF BIRMINGHAM

Immune presentation and recognition of class I MHC phosphopeptide antigens

by

Daniel Henry Stones

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Supervisor: Professor Benjamin E. Willcox

Abstract

Alterations to metabolic pathways, in particular post-translational modification, are a recognised hallmark of diseases such as autoimmunity, inflammation and cancer, and potentially provide a source of altered self antigens that can stimulate immune responses. Most notably, phosphorylated peptides have emerged as a group of tumour associated antigens which can be presented by MHC molecules and recognised by T-cells, and therefore represent promising candidates for future cancer immunotherapy strategies. However, how antigen phosphorylation impacts upon antigen presentation and recognition remains unclear. During this study I demonstrated that the phosphate moiety of phosphopeptides bearing the canonical P4 phosphorylation is more structurally diverse in its binding mode than previously thought. Strikingly, two epitopes exhibited a major conformational change upon addition of the phosphate moiety, effectively creating “conformational neoantigens”. This occurred through a similar mechanism for each epitope, whereby the presence of the phosphate moiety raised the position of the P4 Serine, allowing phosphate-mediated contacts with MHC residues and distorting the conformation of the central epitope region most critical for T- cell receptor recognition. Finally, I found that recognition of phosphopeptides can be both phosphate-dependent and epitope-specific at the level of the T-cell receptor. Therefore, this study shows that phosphorylation can have a profound and diverse effect on antigen binding, epitope identity and T-cell receptor recognition. In summary, my studies suggest that phosphopeptides are not only tumour associated but also highly antigenically distinct, establishing them as attractive candidates for cancer immunotherapy strategies.

For Leanne and Ella

Acknowledgements

I would firstly like to thank my supervisor Ben Willcox for giving me the opportunity to undertake a PhD as part of his group and for all his support both in and out of the Lab. Also, thanks to Fiyaz Mohammed for mentoring in all things crystallography related and for all the collaborative efforts and late nights/early mornings spent shooting X-rays at crystals. A big thanks also to the past and present members of the Willcox lab; especially Carrie Willcox, Mahboob Salim and Sarah Nicholls, for all their help and advice in the lab and for mutual support during those much needed coffee breaks.

Thanks to Mark Cobbold for discussions regarding HLA-B7 phosphoantigens and the cellular aspects of phosphoantigen presentation and recognition. I would like to thank Angela Zarling for providing the TCR gene constructs used for cloning into expression vectors and also Victor Engelhard and Kara Cummings for providing peptide affinity data on all of the phospho/nonphospho-peptides.

Finally I would like to thank all my friends and family who have supported me over the course of this thesis, even when I was at my most stressed and irritable, this applies especially to Leanne, Ella and my parents, without your constant love and support this thesis would not have been possible.

Contents

List of Figures

List of Tables

Table of Abbreviations

Å	Angstrom
µg	Microgram
µl	Microlitre
µM	Micromolar
µm	Micrometer
β2M	β-2-Microglobulin
Ala	Alanine
AIRE	Autoimmune regulator
APC	Antigen presenting cell
Arg	Arginine
Asn	Asparagine
Asp	Aspartic Acid
bp	Base pair
$CaCl_2$	Calcium chloride
cDNA	Complementary DNA
CO_2	Carbondioxide
Cys	Cysteine
DC	Dendritic cell
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium Bromide
FPLC	Fast protein liquid chromatography
g	Gram
g	Gravity
Glu	Glutamic acid
Gln	Glutamine
Gly	Glycine
HCl	Hydrochloric acid
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
Hz	Hertz
Ile	Isoleucine
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kb	Kilobase
KCl	Potasium chloride
KDa	Kilodalton
L	Litre
LB	Luria Bertani

Leu	Leucine
Lys	Lysine
M	Molar
MΦ	Macrophage
MES	2-(N-morpholino)ethanesulfonic acid
Met	Methionine
mg	Milligram
MHC	Major histocompatability complex
min	Minute
ml	Millilitre
mM	Millimolar
Mr	Molecular mass
mTEC	Medullary thymic epithelial cells
NaCl	Sodium chloride
ng	Nanogram
$NiCl_2$	Nickle chloride
Ni-NTA	Nickel-nitriloacetic acid
nm	Nanometer
nM	Nanomolar
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
Phe	Phenylalanine
Pro	Proline
rpm	Revolutions per minute
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	Seconds
Ser	Serine
pSer	Phosphorylated-Serine
TBE	Tris/Borate/EDTA
TBS-T	TRIS Buffered Saline Tween 20
Thr	Threonine
TRIS	tris(hydroxymethyl)aminomethane
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine
w/v	Wieght/Volume
v/v	Volume/Volume

1 Introduction

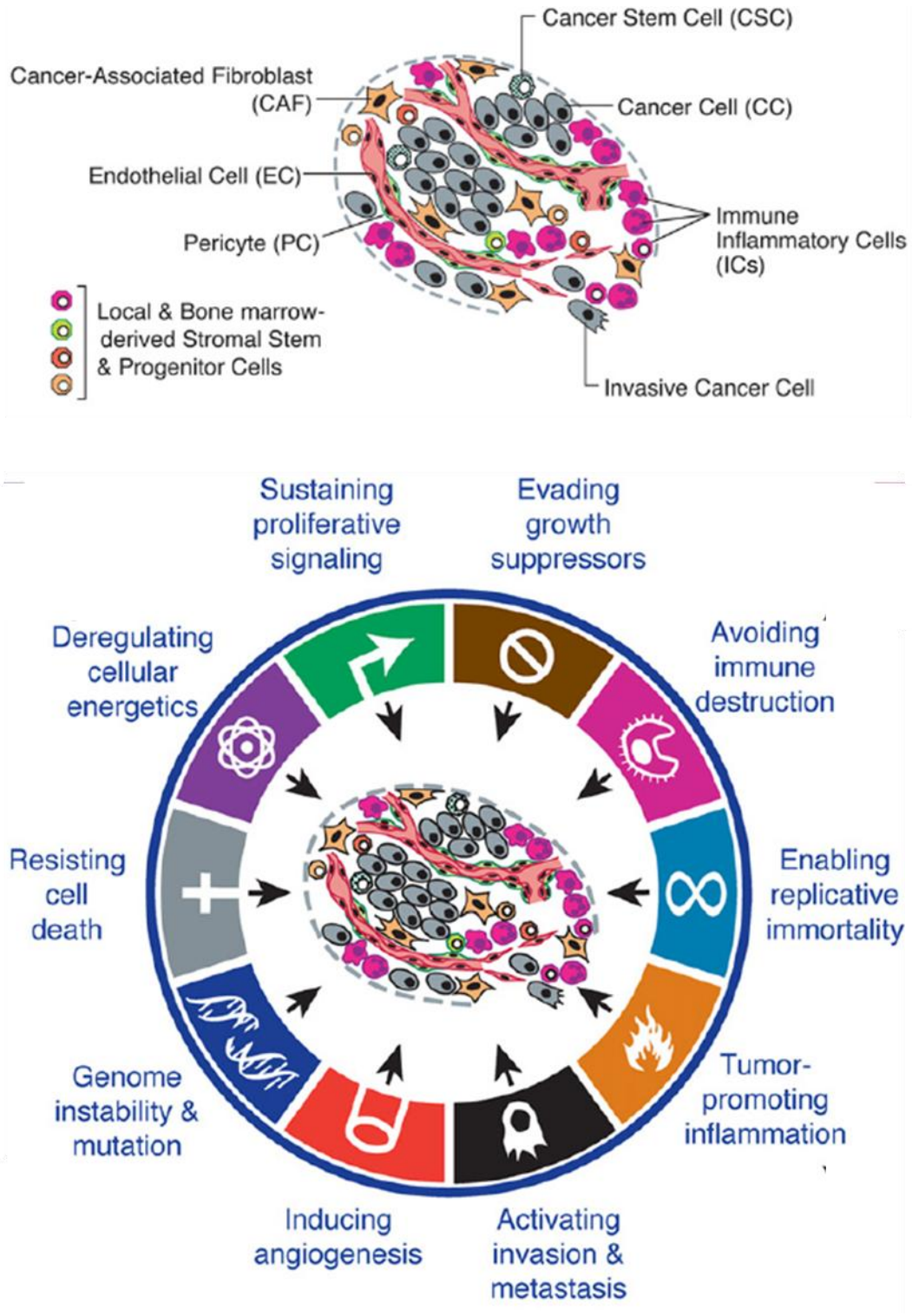
1.1 Cancer and the immune system

1.1.1 Tumour biology

Although death rates for some cancers have fallen in recent years, cancer remains a leading cause of death worldwide and is projected to increase by 45% between 2007-2030 (WHO, 2010).

Tumour development in humans is thought to involve a multistep process of genetic alterations originating from defects in DNA repair mechanisms (Kinzler and Vogelstein 1996; Jiricny and Marra 2003; Peltomaki 2003), thus helping to drive the progression of normal cells into highly malignant cancers. These multiple genetic changes range from single nucleotide point deletions to chromosome complement disruption (Kinzler and Vogelstein 1996) and can be due to hereditary defects or induced through chemical carcinogens, UV exposure or other external factors. Due to the decreased functionality of DNA repair mechanisms these cells often exhibit an increased mutation rate in pathways which are thought to be crucial to cancer formation, survival and metastasis. Hanahan & Weinberg (Hanahan and Weinberg 2000) describe in their seminal paper six pathways cancers exploit in order to survive; self sufficiency in growth signals, insensitivity to anti-growth signals, tissue invasion and metastasis, sustained angiogenesis, limitless replicative potential and evasion of apoptotic signals. Since the publication of this article in 2000, other attributes have been added to the list (Figure 1.1) (Hanahan and Weinberg 2011), most notably, the ability of tumour cells to evade host immune responses (Dunn, Old et al. 2004).

Figure 1.1 Hallmarks of cancer



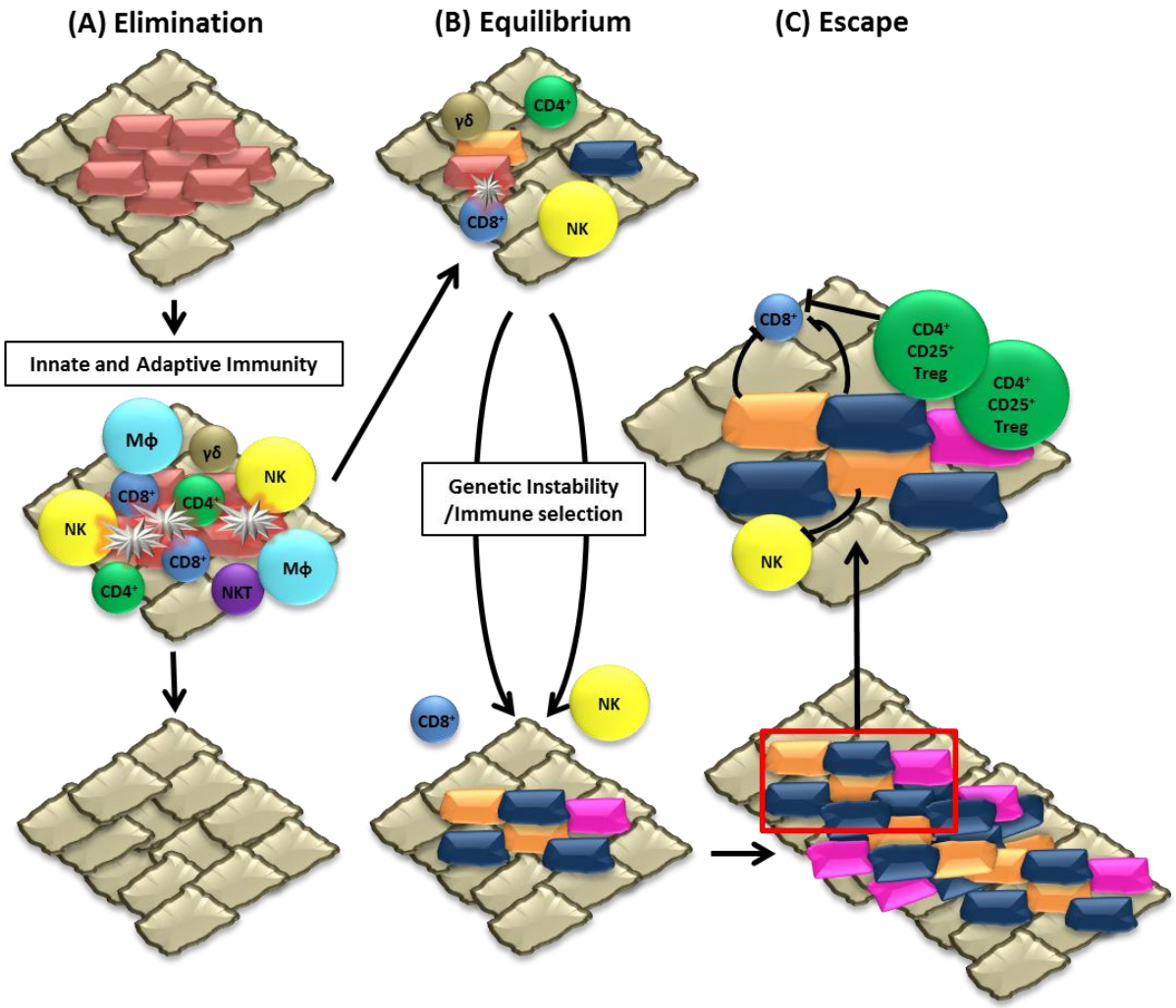
In order for tumours to grow and survive, tumour cells must first exploit several of these pathways. Figure adapted from (Hanahan and Weinberg 2011)

1.1.2 Evidence for tumour suppression by the immune system

It was first hypothesised by Ehrlich in 1909 that the immune system plays a role in recognising and eliminating tumour formation (Ehrlich 1909). However it was not until the 1950's that Burnet and Thomas proposed the theory of immunosurveillance in which the small, continual accumulation of nascent tumour cells possessing detrimental genetic alterations are recognised and destroyed by the host immune system (Burnet 1957; Thomas 1959). However, for the tumour to grow to an observable clinical grade the tumour must evade the host cancer immunosurveillance mechanisms through a process termed cancer immunoediting (Figure 1.2) (Dunn, Bruce et al. 2002; Dunn, Old et al. 2004). Dunn *et al* proposed that during immunoediting the cancer cell population undergoes Darwinian evolution in order to survive. The first stage of immunoediting consists of the "elimination" phase mentioned above, in which immunosurveillance mechanisms destroy the majority of arising pre-cancerous cells. This however applies selective pressures upon the arising tumour cells and the second stage, "equilibrium", is reached where the selective pressures of the immune response have preferentially selected arising tumour cells with decreased immunogenicity and enhanced survival factors. These surviving cells, with increased genetic instability and survival advantages lead to the final stage of "tumour escape".

Numerous studies in both mouse and human models provide evidence for immunosurveillance mechanisms acting to control tumour cell growth. Although early mouse models failed to prove the existence of tumour immunosurveillance (Gorer and Kaliss 1959; Stutman 1974), which lead to the immunosurveillance hypothesis falling out of favour, it was not until the development of knockout mouse models that the theory of immunosurveillance could be formally proved.

Figure 1.2 Cancer Immunoediting



Cancer Immunoediting is thought to involve three phases. Firstly, nascent tumour cells (red rectangular cells) arising from normal cells (large grey rectangles) are recognised and eliminated by cells of the innate and adaptive immune response (circles). This is followed by a period of equilibrium in which some tumour cells, through genetic instability, become resistant to immune recognition and elimination (orange, blue and pink rectangles). This is followed by the third phase of tumour escape in which the cells able to persist during the equilibrium phase outgrow to form a tumour mass capable of suppressing the immune response directly and through recruitment of suppressor cells such as $CD4^+CD25^+$ regulatory T-cells. This figure was modified from (Strausberg 2005).

In particular the Interferon-γ receptor (IFN-γR) knockout mouse was found to be 10-20 times more susceptible to tumour induction by chemical carcinogenesis (Kaplan, Shankaran et al. 1998). Also, over-expression of a dominant-negative mutant IFN-γR α subunit in sarcomas completely ablated tumour sensitivity to IFN-γ and following transplantation into naive syngeneic hosts displayed increased tumourigenisis and reduced immunogenicity (Dighe, Richards et al. 1994). These results demonstrate that IFN-γ, a key product of many immune responses, is a critical component of tumour control by the host. Further studies using other mouse models, such as perforin knockout and RAG-2 knockout mice (van den Broek, Kägi et al. 1996; Shankaran, Ikeda et al. 2001), provided additional evidence to support the immunosurveillance theory. The most striking of these showed that chemically induced tumours generated in RAG-2 immune-deficient mice are strongly immunogenic and were able to be killed upon transplantation to a WT immune-competent host. However, tumours generated in the same manner in WT immune-competent hosts are weakly immunogenic and were not able to be controlled upon transplantation (Shankaran, Ikeda et al. 2001). This provided direct evidence for immunoediting, where highly immunogenic tumour cells arising in the immune-competent host are eliminated by the immune response early in the tumourigenisis process, thus leaving a subset of weakly immunogenic tumour cells to persist.

The evidence for immune control of tumour growth in human models comes firstly from studies in immunocompromised hosts; including organ transplant patients who are receiving immunosuppressive drugs and human immunodeficiency virus (HIV) positive patients. Following solid organ transplantation, patients have a 4-fold increased risk of developing cancer (Adami, Gabel et al. 2003). Although many of these tumours are virally associated and may arise due to the lack of anti-viral immune responses rather than anti-tumour

responses, increased relative risk ratios have been observed for a range of tumours with no apparent link to viral infection (Sheil 1986; Penn 1996). Secondly there is a large body of evidence which demonstrates that in immunocompetent patients the presence of tumour infiltrating lymphocytes (TILs) can correlate with increased survival in many tumour settings (Clemente, Mihm et al. 1996; Ropponen, Eskelinen et al. 1997; Schumacher, Haensch et al. 2001; Zhang, Conejo-Garcia et al. 2003; Galon, Costes et al. 2006). Although these data clearly show that the immune system can act to curtail tumour growth, there is also evidence in some tumour settings which shows the presence of TIL infiltration can result in poor prognosis (Kataki, Scheid et al. 2002). Potentially this highlights another aspect of the tumour immune-editing hypothesis, in that cancer cell survival may in some cases involve not simply overcoming but even utilising the immune response and this will be discussed in greater detail in section 1.4.2.

1.2 Vertebrate Immunity

The vertebrate immune system consists of complex cellular networks that have evolved to defend the host against a wide array of pathogens and has been implicated in host protection from tumourogenesis (Dunn, Bruce et al. 2002; Kim, Emi et al. 2007; Swann and Smyth 2007). These processes are commonly separated into two categories; the innate immune response and the adaptive immune response. The innate response to an antigen occurs early and does not involve processing of antigen. Cells of the innate immune system express pattern recognition receptors (PRRs), which allow them to recognise pathogen-encoded and damage associated molecules (pathogen-associated molecular patterns, PAMPs and damage-associated molecular paterns, DAMPs) directly (Bianchi 2007; Sancho,

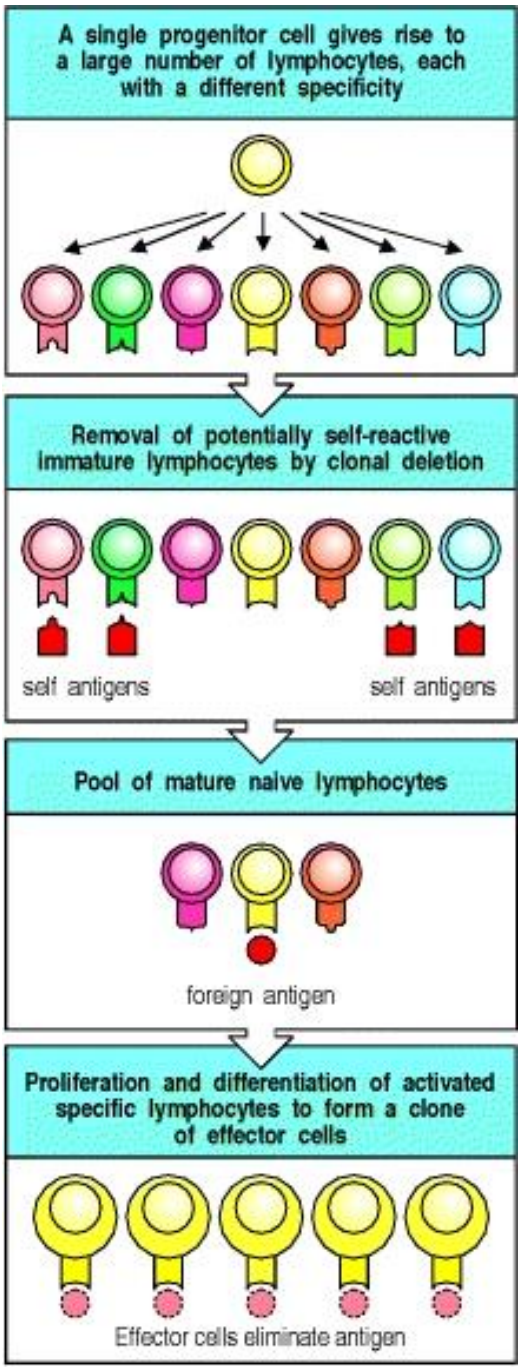
Joffre et al. 2009). The innate response therefore does not require prior contact with the pathogen and subsequent encounters with the same antigen do not lead to increased activation of the innate response. In contrast, the adaptive response requires processing and presentation of antigen by antigen presenting cells (APC), which initiate the clonal expansion of antigen specific lymphocytes in response to pathogen challenge. As a consequence, the adaptive immune response is slower to initiate than the innate response, but allows the generation of "immunological memory" through survival of clonally expanded antigen specific lymphocyte populations (Gowans and Uhr 1966; Zinkernagel, Bachmann et al. 1996; Zinkernagel 2002). Together the innate and adaptive immune responses allow the immune system to respond quickly to an antigen but also establish long-lasting pathogen specific memory.

Unlike cells of the innate immune system, which carry multiple germline encoded receptors that can recognise a wide array of pathogen determinants, the lymphocytes of the adaptive immune system, B-cells and T-cell, bare only a single specificity of antigen receptor, namely the B-cell receptor (BCR) and the T-cell receptor respectively, whose structure belongs to the immunoglobulin super family. The specificity of lymphocyte antigen receptors is diverse and is reliant upon a unique ability of lymphocyte cells to rearrange the genes encoding its antigen receptor. The antigen receptor consists of multiple inherited gene segments which are able to undergo somatic gene rearrangement to produce a distinct combination of segments for each antigen receptor (Nemazee 2006). The diversity in the receptor repertoire is further enhanced due to events that introduce variations during the rearrangement of each segment such as N-nucleotide addition (Cabaniols, Fazilleau et al. 2001) and in B-cells somatic hypermutation (Teng and Papavasiliou 2007). This diversity is most apparent in the

complimentary determinant regions 1-3 (CDR1-3) (Wu and Kabat 1970) produced during VDJ and VJ recombination. However this quasi-random generation of lymphocyte antigen receptors, which collectively express a wide array of antigen specificities, could inevitably lead to the generation of an auto-reactive lymphocyte population. This problem is addressed to a large extent during lymphocyte development, where clones which are potentially self-reactive are eliminated from the repertoire by clonal deletion (Blackman, Kappler et al. 1990; Rajewsky 1996), thus allowing the expansion of self-tolerant lymphocyte populations (Figure 1.3). Therefore the diversity of lymphocyte receptors generated is extensive, although only a small proportion of naïve lymphocytes finally progress past the selection process and develop into matured lymphocytes.

B-cells and the antibodies they produce are able to recognise and bind to free antigen whereas in contrast, T-cells are known to recognise antigen in the context of MHC molecules (Zinkernagel and Doherty 1974; Zinkernagel and Doherty 1974) and therefore require antigen processing. The selection process that determines further T-cell development involves a fine balance between the initial positive selection of T-cells which recognise presented self-molecules, followed by negative selection of those that recognise self with a high affinity (Strasser, Puthalakath et al. 2007). Thus each receptor surviving into the mature repertoire is functional but at the same time potentially auto-reactive T-cell populations are eliminated. The survival of lymphocytes is thought to be mediated by signalling through the antigen receptor and co-receptors (Blackman, Kappler et al. 1990; Cantrell 2002; Pobezinsky, Angelov et al. 2012), and that these signals are provided by other cells in the lymphoid organs (Anderson and Takahama 2012). Lymphocytes that fail to receive signals undergo a form of programmed cell death called apoptosis.

Figure 1.3 Clonal Selection Theory



A multitude of lymphocytes with different antigen specificities arise from a single precursor. Self-reactive lymphocytes are deleted in the thymus during negative selection leaving a pool of mature naïve lymphocytes. Upon antigen recognition and in the presence of secondary signalling on activated antigen presenting cells, the antigen specific lymphocytes proliferate and differentiate to form a clonal pool of effector and memory cells. Figure taken from Immunobiology (Janeway, Travers et al. 2005).

Alternatively, those that do receive appropriate signals are then able to mature into naïve lymphocytes. In the case of T-cells, the proliferation of an antigen specific populations of naive lymphocytes relies upon interaction with presented molecules, co-stimulation and cytokine stimulation which is facilitated by APC present in the draining lymphoid tissue (Curtsinger, Johnson et al. 2003). Following antigen removal most of the activated antigen specific cells undergo apoptosis. However, a small proportion remains after antigen removal and there is strong evidence to suggest these form the basis of immunological memory (Gray 2002; Zinkernagel 2002; Fazilleau, McHeyzer-Williams et al. 2007).

The generation of effective immunity to pathogens requires coordination between the innate and adaptive immune responses. One major link between the two responses is provided by a group of cells that are collectively termed 'professional antigen presenting cells' (APC). An important class of APC is the dendritic cell (Liu 2001). Dendritic cells (DC) are thought to exist in peripheral tissues in what is termed an immature state, where they continuously "sample" the surrounding environment by endocytosis and pinocytosis (Guermonprez, Valladeau et al. 2002; Lutz and Schuler 2002). In this immature state they show low expression of MHC molecules at their surface and lack co-stimulatory molecule expression, but have an increased ability to phagocytose pathogens (Banchereau and Steinman 1998; Lutz and Schuler 2002). Immature DC also express high levels of the Toll-like receptor family and other PRRs on their surface, which have evolved to detect PAMPs present on a wide array of pathogens (Akira and Hemmi 2003) and DAMPs associated with cellular stress and necrosis (Bianchi 2007; Sancho, Joffre et al. 2009). During infection, DCs become activated and migrate into the lymphatics and eventually to the local lymph tissue. During this process the DC receives maturation signals that enable the up-regulation of MHC

class I and II molecules, allowing presentation of phagocytosed antigen (Lutz and Schuler 2002). In addition, they also show up-regulation of co-stimulatory molecules such as CD86, CD40, and adhesion molecules and chemokines, which collectively explain their potent ability to stimulate naïve T-cell responses (Banchereau and Steinman 1998; Lutz and Schuler 2002).

Once in the draining lymph node, DC are highly efficient at presenting antigen to naïve T-cells circulating through the lymphatic system, namely CD4+ T-cells via MHC class II complexes and/or CD8+ T-cells via MHC class I complexes. Successful priming of a response involves binding of the TCR to cognate antigen displayed by MHC molecules on the DC surface, along with co-stimulatory signals from DC to T-cell (Curtsinger, Johnson et al. 2003). Chief amongst these are the costimulatory B7 family molecules expressed on the DC surface, which are recognised by CD28 on the T cell surface (Carreno and Collins 2002). Subsequently, increases in adhesion molecule expression, cytokine production, and up regulation of CD40 ligand, which binds to CD40 on APC and transmits activating signals to the T-cell, help sustain T-cell signalling and drive clonal expansion and differentiation (Guermonprez, Valladeau et al. 2002).

CD4+ T-cells play key immune-regulatory roles and can profoundly influence any subsequent responses to a given pathogen. Their differentiation is relatively complex but subsets of functionally distinct CD4+ T-cells can be separated on the basis of the cytokines they produce into Th1, Th2 or the more recently recognized regulatory-T-cell (Treg) and Th17 cell lineages. The Th1 lineage produces IL-2 and IFN-γ, and favours the cell-mediated immune response by activating macrophages and stimulating the CD8+ T cell response, thus enabling

elimination of intracellular pathogens (Szabo, Sullivan et al. 2003). In contrast, Th2 CD4+ T-cells produce IL-4, IL-13 and IL-10, and are involved in co-ordination of the B-cell humoral response (Paul and Zhu 2010). Since Th1 cytokines can inhibit Th2 differentiation and vice-versa, CD4+ T cell responses have the potential to become polarised (Neurath, Finotto et al. 2002; Romagnani 2004). The factors affecting Th1/Th2 polarisation are not entirely clear, but include the nature of the cytokines induced by a pathogen, the costimulatory signals from APC and the nature of the peptide-MHC ligand (Akira, Takeda et al. 2001). Since Th1/Th2 bias is thought to be evident relatively early in the immune response (Rothenberg, Moore et al. 2008), production of particular cytokines by cells of the innate immune system is thought to be relevant (Romagnani 2004). A more recent identification of a novel subset of IL-17 producing CD4+ T-cells has been defined as Th17 cells (Stockinger and Veldhoen 2007). The Th17 subset has been shown to directly mediate inflammation and play an important role in many autoimmune diseases (Furuzawa-Carballeda, Vargas-Rojas et al. 2007).

Understanding how adaptive immune responses are generated is of key importance in establishing how durable immunity to specific antigens is elicited, and how these processes might be harnessed therapeutically. In particular, stimulation of class I MHC-restricted T cell responses forms the basis of many vaccination and cancer immunotherapy strategies, and understanding the key interactions that determine the ability of a given antigen to be processed and presented to $CD8^+$ T-cells is likely to have a major impact on the design of such therapies.

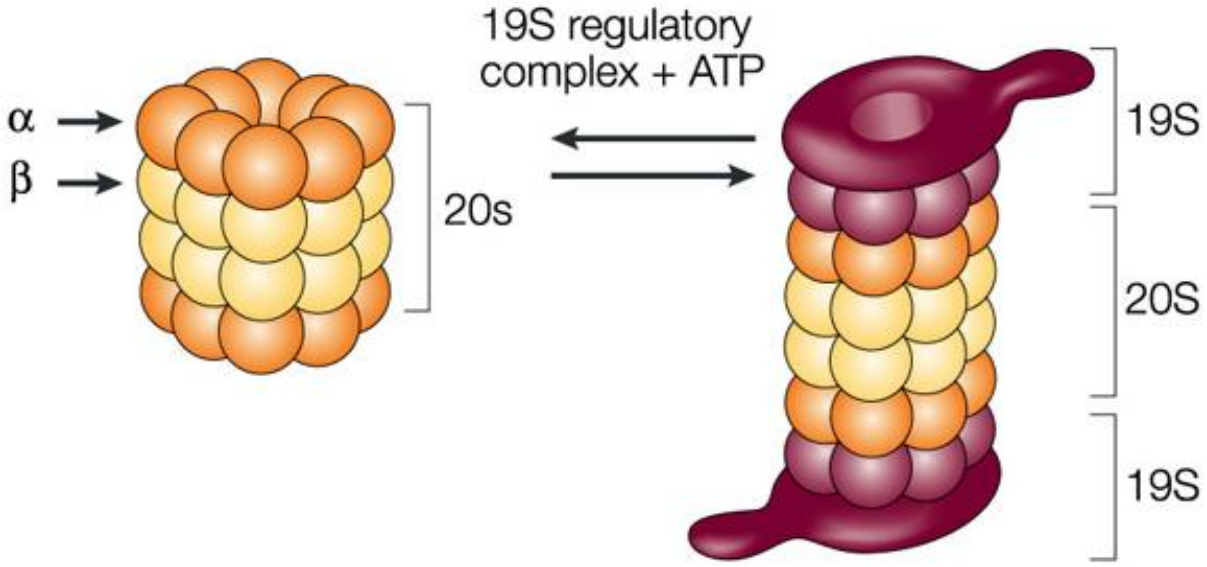
1.3 Antigen Presentation

1.3.1 Proteasomal degradation of cytosolic antigens

MHC class I molecules bind short peptide fragments, that are from pathogen-derived or self-molecules, enabling stable expression on the cell surface and presentation of such antigen to $CD8^+$ T-cells (Bjorkman, Saper et al. 1987). Class I molecule presentation can occur via two routes; processing of intra cellular antigens (Townsend and Bodmer 1989), either self or pathogen derived, or secondly, through cross presentation of exogenous antigens. For the moment I will focus on the endogenous antigen processing pathway and will focus on cross presentation in section 1.3.6. For intracellular antigens to be presented they must first be processed into small fragments, since class I molecules typically bind peptide fragments of 8-10 amino acids in length. This consists of two processes; ATP- dependent ubiqutination and unfolding, and ATP-independent degradation of the polypeptide chain.

The proteasome, an abundantly expressed ATP-dependent protease present in the cytosol and nucleus, plays a major role in degrading intracellular proteins (Peters, Franke et al. 1994), including those destined for presentation by class I MHC molecules. Each proteasome consists of a 20S core proteosome formed from multiple α and β protein subunits that stack in four heptameric rings to form a hollow cylinder (Figure 1.4). At each end of the 20S cylinder, the PA700 (19S) complex acts as a cap, and contains subunits that have ATPase activity and help to unfold the proteins to be digested, as well as those involved in recognition of proteins targeted for destruction (Peters, Franke et al. 1994). The 19S caps therefore act as regulated “gates” controlling entrance of the polypeptide chain to the proteolytic core of the proteosome.

Figure 1.4 The proteasome



The proteasome consists of a central 20S subunit comprised of α and β proteins stacked in a heptameric ring structure. The entry of protein chains into the 20S core is regulated by 19S which forms a cap structure at either end of the S20 proteasome. Figure taken from (Adams 2004).

Once the polypeptide chain is within the proteolytic core of 20S, the chain is degraded by different endopeptidases, aminopeptidases and carboxypeptidases in an ATP-independent fashion. The 3-15 amino acid peptide chains then diffuse out of the proteosome into the cytosol. Unlike conventional proteases, the function of this mechanism allows the protein substrate to remain bound until all of it is converted into short fragments.

The proteasome can exist in two forms, the 'constitutive' proteasome is expressed in all mammalian cells, and the 'immunoproteosome', which contains three distinct IFN-γ-inducible subunits, LMP2, LMP7 and PA28, and is mainly expressed during infection in cells stimulated with inflammatory cytokines such as IFN-γ (Griffin, Nandi et al. 1998). Replacement of constitutive subunits with LMP2 and LMP7 alters the activity of the proteasome, such that increased cleavage is observed after hydrophobic residues, favouring production of peptides suitable for both TAP transport and MHC binding (Rivett and Hearn 2004). In addition, incorporation of the PA28 ring structure at either end of the 20S proteasome is thought to increase efflux of peptides from the proteasome, potentially allowing antigenic peptides that might otherwise be cleaved to escape further processing (Wang and Maldonado 2006). Due to this it has also been found that the immunoproteosome is highly important for generating a variety of immune responses (Van den Eynde and Morel 2001; Strehl, Seifert et al. 2005).

In most cases, the proteins that proteasomes act on are targeted for destruction by covalent attachment of the small protein ubiquitin (Hershko and Ciechanover 1998). Ubiquitin can exist free in the cell or covalently linked to a large number of intracellular proteins, the majority of which are ultimately destined for proteasomal destruction. Ubiquitin is prepared

for conjugation by the ATP-dependent ubiquitin-activating enzyme E1, which activates the ubiquitin and transfers it to one of a number of E2 ubiquitin-conjugating enzymes. The E2 enzymes act in conjunction with E3 accessory proteins to form E2-E3 ubiquitin ligases, whereby the E3 moiety recognises specific degradation signals in protein substrates, and the E2 protein ultimately attaches a polyubiquitin chain to specific Lysine residues in the target protein (Hershko and Ciechanover 1998). This polyubiquitin attachment is then recognised by subunits of the PA700 proteasome complex which targets the protein for proteolytic degradation (Pickart 1997). The complexity of the ubiquitin system is emphasised by the existence of multiple (>30) distinct E2 enzymes and hundreds of different E3 proteins that complex with specific E2 enzymes. The system therefore consists of multiple distinct proteolytic pathways, which share the E1 enzyme and the proteasome but differ in the composition of their ubiquitin ligases, allowing recognition of different degradation signals (Pickart 1997; Pickart 2001) and consequently targeting of distinct subsets of intracellular proteins for destruction.

The nature of the degradation signals that influence which proteins are targeted for proteasomal destruction and how these are regulated is highly relevant when considering which proteins serve as substrates for the class I MHC processing pathway. Newly synthesised proteins can be substrates for class I MHC peptide production due to the production of defective ribosomal products (DRiPs) which can contain large patches of abnormally exposed hydrophobic residues that act as signals for destruction of the protein (Yewdell, Antón et al. 1996; Minami, Hayakawa et al. 2010). These can usually be rectified by chaperone proteins that aid in the correct refolding of the protein, however, if these are unable to repair the misfolded protein then the protein is ultimately eliminated, avoiding

formation of protein aggregates. Regulated destruction of protein can also occur through targeted phosphorylation events at specific sites on a protein which unmask a normally hidden degradation signal (Hao, Oehlmann et al. 2007). The main mechanism by which cells recognise these abnormal degradation signals are through the multiple E3 molecules that facilitate protein entry into the ubiquitin mediated-proteasomal degradation pathway. Once the poly-peptide has been cleaved the peptide fragment is able to be transported into the endoplasmic reticulum (ER) by the transporter associated with antigen processing (TAP).

1.3.2 TAP transport

Class I MHC molecules typically present peptides derived from cytosolic proteins. However, the peptide-binding domains of nascent class I MHC proteins are never exposed to the cytosol, and instead face the lumen of the ER. Therefore, in order for peptide antigens to enter the ER and be loaded onto class I MHC molecules, they must first be transported across the ER membrane. This is facilitated by TAP, which carries out ATP-dependent peptide transport (Abele and Tampe 2004). TAP is a member of the ABC transporter family of proteins and is a heterodimer formed from two MHC-encoded proteins, TAP-1 and TAP-2 (Reits, Griekspoor et al. 2000). The heterodimer consists of 3 structural domains; the pore domain, the peptide binding domain and finally the ATP-binding domain. The pore domain is ~50Å in diameter and therefore restricts the length of peptides able to enter and be transported across the ER membrane (8-16 amino acids). Binding of peptide to the peptide binding domain is somewhat sequence-dependent, favouring hydrophobic residues at the COOH terminal end of the peptide antigen (van Endert, Riganelli et al. 1995). These features favour transport of peptides suitable for binding to class I MHC molecules. Binding of peptide to TAP results in a conformational change in the TAP molecule and may act as the

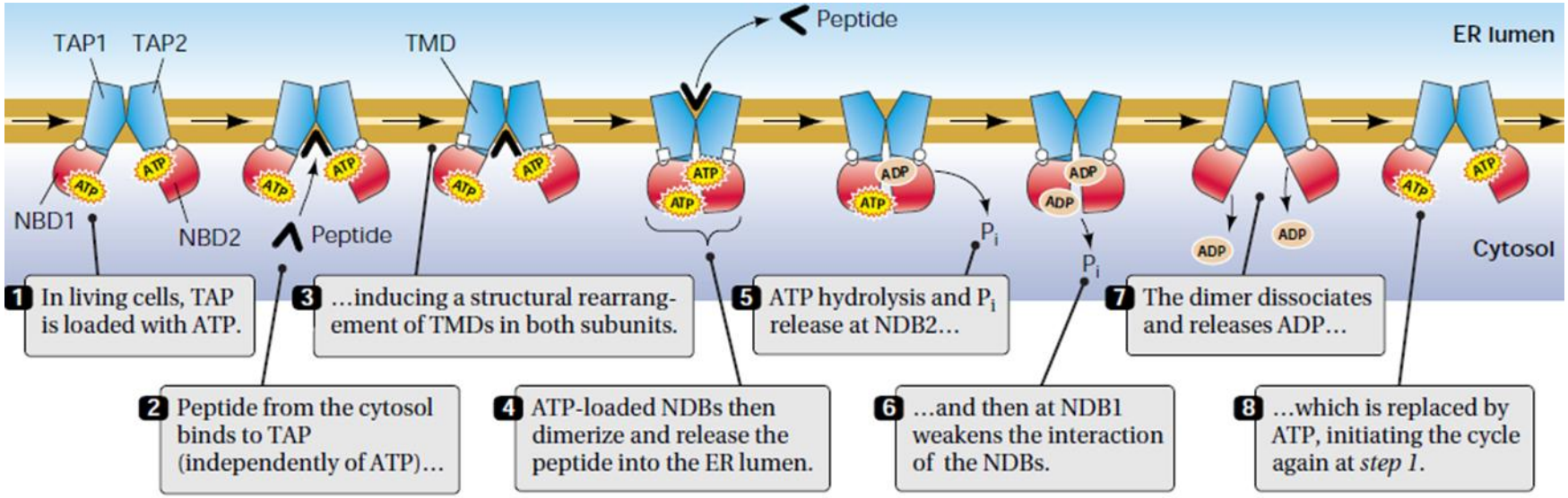
switch for ATP hydrolysis. ATP is bound to TAP at the third domain; the Nucleotide Binding Domain. The exact mechanism of peptide transport across TAP and the ER membrane is still not fully understood; however a mechanism has been proposed but not yet proven experimentally (Figure 1.5) (Abele and Tampe 2004; Procko, O'Mara et al. 2009).

1.3.3 Antigen processing in the endoplasmic reticulum

Assembly of class I MHC molecules in the endoplasmic reticulum (ER) depends on association of the class I MHC heavy chain with β2-microglobulin and peptide (Purcell and Elliott 2008; Lee, Park et al. 2009). This in turn is dependent on the further proteolytic processing of the ER peptide pool, and on a series of chaperone-mediated events on the lumenal face of the ER membrane that, assuming appropriate peptides are present, ultimately result in stable assembly of the class I MHC molecule (Elliott and Williams 2005). Stable association with peptide results in release of class I MHC to the cell surface, whereas unstable association with class I MHC proteins, results in retention in the ER.

Subsequent to transport of proteasomally generated peptides to the ER via TAP, various aminopeptidases associated with the ER are involved in additional cleavage and trimming of the peptide antigens present. A key aminopeptidase implicated in the final proteolytic processing of peptides presented on MHC class I molecules is the ER-associated aminopeptidase 1 (ERAP1) (Chang, Momburg et al. 2005). ERAP1 has been found to trim N-extended precursor peptides but spares peptides ~ 9 amino acids, the length that optimally binds to MHC class I. In this sense, ERAP1 acts as a 'molecular ruler', binding to peptide substrates of 9-16 amino acids in length and accurately trimming them to a length of 9 amino acids but no less (Chang, Momburg et al. 2005).

Figure 1.5 TAP transport of peptides



TAP consists of two transmembrane domains (TMD) and two nucleotide binding domains (NBD). TAP resides in a peptide receptive conformation bound to two molecules of ATP. Upon peptide binding the structure of TAP alters to allow the peptide into the ER. During the alteration of TAP's conformation the ATP molecules are hydrolysed to ADP which subsequently is lost from the TAP molecule. This causes the TAP molecule to take up new ATP molecules and revert to the peptide receptive conformation and thus starting the cycle again. Figure taken from (Abele and Tampe 2004).

In addition, the ERAP1 molecule possesses substrate specificity in that peptides with hydrophobic C-termini are able to bind to ERAP1 whereas peptides with charged C-termini do not bind, therefore favouring the appropriate trimming of peptides with C-terminal residues suitable for binding class I MHC molecules (Kochan, Krojer et al. 2011). Aminopeptidases such as ERAP-1 therefore have profound effects on the repertoire of peptides generated for class I MHC loading.

Newly expressed class I MHC heavy chains initially associate with the ER chaperone calnexin, which catalyses association with β2-microglobulin. Subsequently, the heavy chain:β2-microglobulin complexes associate with a number of other proteins that are involved class I MHC stabilisation and loading, and are collectively termed the peptide loading complex (PLC), which contains the proteins tapasin, ERp57 and calreticulin (Elliott and Williams 2005). The function of this complex is to maintain the class I MHC molecule in a peptide-receptive conformation and retain it in close proximity to peptides transported into the ER lumen by TAP. A key member of the complex is Tapasin, which is required for class I antigen presentation and aids the stability of class I molecules. To date, three functions of Tapasin have been discovered, emphasising its vital role in the folding of class I MHC molecules. Firstly, the tapasin molecule stabilises the TAP heterodimer and so promotes peptide transport into the ER. Secondly, it links the TAP molecules to the MHC molecules and so is structurally integral to the PLC. Thirdly, it plays a role in the loading of peptide antigens onto the MHC (Sadasivan, Lehner et al. 1996; Elliott and Williams 2005). ERp57 is a thiol-oxidoreductase involved in ensuring correct folding of newly formed MHC heavy chain along with calreticulin. The ERp57 protein is also thought to be involved in facilitating conformational flexibility of the class I binding cleft to assist in peptide binding (Purcell and

Elliott 2008). However, cells defective in either tapasin or calreticulin show defects in class I MHC assembly and express class I MHC molecules at the surface that contain suboptimal low affinity peptides. Consequently, the PLC complex is thought to facilitate a key peptide-editing function, allowing the exchange of low affinity peptides bound to the class I MHC molecule for those of higher affinity.

1.3.4 Class I MHC structure

Class I MHC molecules consist of a type I transmembrane heavy chain comprising a α1-α2 peptide-binding platform and a membrane-proximal α3 domain, which associates non-covalently with the β2-microglobulin light chain and bound peptide (Bjorkman, Saper et al. 1987; Madden 1995). The α3 and β2-microglobulin domains, located proximal to the cell membrane, share (relative amino-acid sequence and intra-chain disulphide bridge) homology to immunoglobulin constant regions (Peterson, Cunningham et al. 1972; Bjorkman, Saper et al. 1987). The highly polymorphic α1 and α2 domains, located distal to the cell membrane, associate together as an "intra-chain dimer" to form a distinctive peptide-binding platform oriented away from the presenting cell at the membrane-distal end of the molecule, and supported by the membrane-proximal α3 and β2-microglobulin domains.

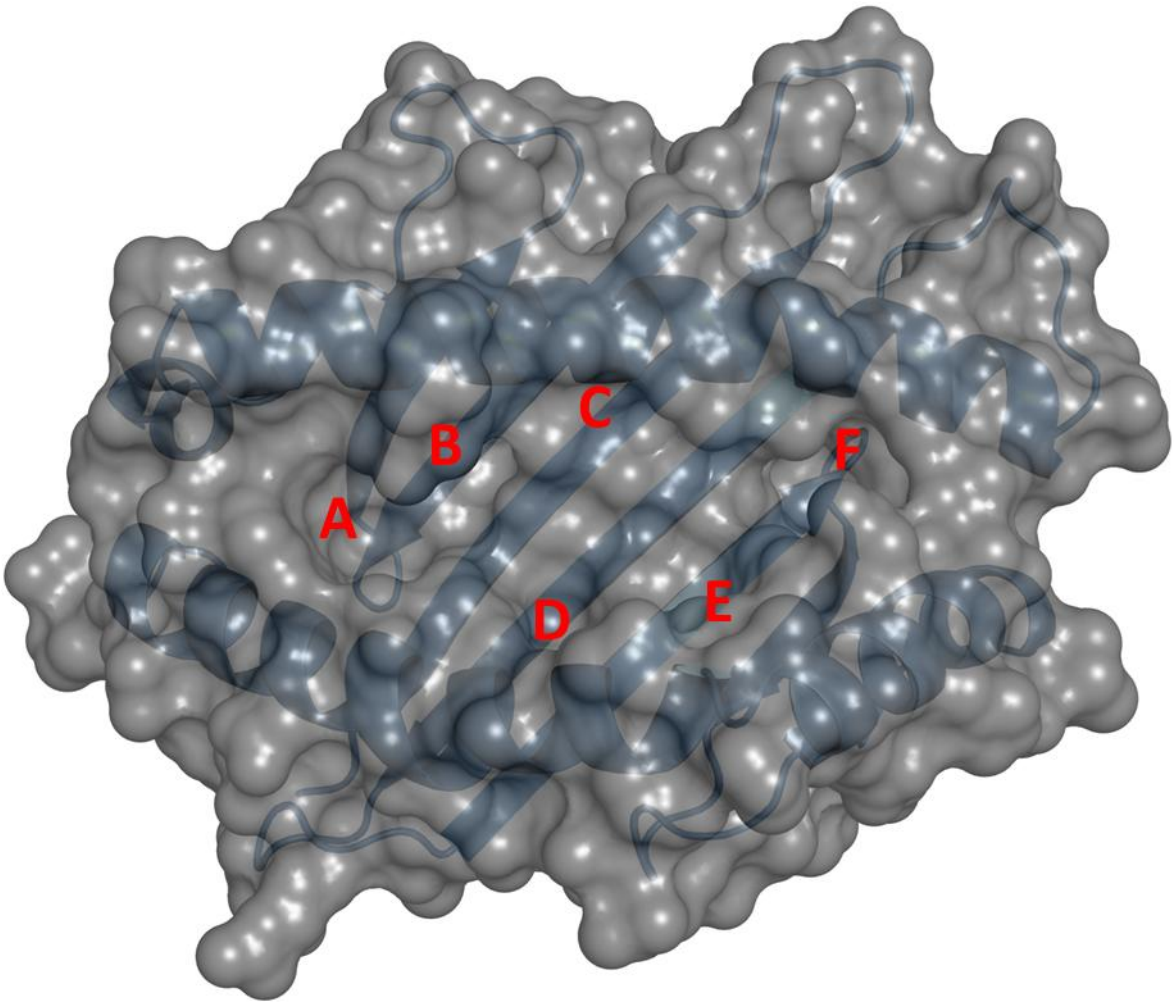
Within the α1-α2 platform, the peptide binding site is located in a groove situated between two α-helices positioned on top of an eight-stranded antiparallel β sheet. The antigen binding site of HLA-A2 is approx 25Å long and 10Å wide (Bjorkman, Saper et al. 1987), enclosed at either end by bulky amino acid side chains that are conserved in almost all class I MHC molecules. Early structures of class I MHC molecules noted electron density un-

attributable to HLA-A2 heavy chain along this groove and it was hypothesised that this represented a mixture of bound peptides (Madden, Garboczi et al. 1993).

Subsequent structures of class I MHC molecules in the presence of individual peptide species has clearly defined the molecular mechanisms governing class I MHC antigen presentation. The peptide occupies an extended conformation along the groove, with variations in length accommodated over the central region of the antigen. The enclosed nature of the groove explains why the length of peptides bound to MHC class I molecules generally falls into a limited range, from ~8-14 amino acids long. The interactions that bind the peptide within the groove of the HLA-A2 molecule occur in defined pocket regions termed A-F (Saper, Bjorkman et al. 1991) (Figure 1.6). Interactions to the amino and carboxyl termini involve conserved amino acid residues and involve Pocket A and Pocket F, respectively. Pockets B and F play critical roles in accommodating primary anchor residues. Pocket B is located adjacent to pocket A and protrudes under the α1 domain helix, reaching the backbone of the α1 domain β-sheet. The opening of pocket B is composed of entirely polar side chains, however the inner surface is completely hydrophobic, explaining the presence of Leu as the primary anchor at position 2. In addition, the hydrophobic Pocket F permits interaction with the hydrophobic C-terminal primary anchor residue, which is oriented downwards to the β-sheet floor of the groove (Saper, Bjorkman et al. 1991).

These studies have defined the key structural requirements for peptide-MHC binding. Importantly, the structures emphasise that peptide is critical to the structural integrity of the complex (Madden 1995), forming interactions with the α1 and α2 helix and the β-sheet floor forming the base of the peptide-binding groove.

Figure 1.6 Pockets of the MHC-class I binding groove



Interactions to the amino and carboxyl termini involve conserved amino acid residues and involve Pocket A and Pocket F, respectively. Pocket B is involved in P2 anchor residue binding as is the F pocket for PC anchor residues. Pockets D, C and E are not thought to play as prominent a role in peptide binding as pockets A, B and F.

In addition, the strikingly high number of polymorphic residues in class I MHC genes map to residues oriented into the groove, or protruding from the surface of the alpha helices (Bjorkman, Saper et al. 1987). Consequently, allelic differences between MHC molecules give rise to distinct peptide-binding and TCR-binding properties.

1.3.5 MHC-peptide binding and complex stability

Since class I MHC molecules are highly unstable in the absence of antigen, the affinity of peptides for the class I MHC molecule plays an important role in determining presentation and stability at the cell surface, and can profoundly affect immunogenicity (Micheletti, Bazzaro et al. 1999; Skokos, Shakhar et al. 2007). Below a certain threshold affinity, peptides are likely to be out-competed by others that possess a higher binding affinity for the MHC molecule in question, which will be presented at the cell surface at a greater level than their low affinity counterparts. Peptide binding affinity, typically in the nanomolar range for immunogenic peptides, relies upon several factors. The optimal length of peptides bound to class I MHC molecules is between 8-10 amino-acids in length, however larger peptides have been shown to bind (Parker, Bednarek et al. 1992). For a peptide to bind to class I MHC, it must have a free N-terminal amino group and a free C-terminal carboxylic acid group, suggesting interactions with the N and C termini form important contacts to the MHC molecule. Analysis of peptides eluted from class I MHC molecules indicated thousands of peptides are able to bind each allele, and highlighted this promiscuous binding is dependent on the presence of key primary anchor residues (Engelhard 1994). For HLA-A2, primary anchor residues are located at positions 2 and 9, where Leucine and Valine are preferred respectively. The nature of the primary anchor residues has been shown to have a dramatic impact upon the affinity of a peptide for MHC (Parker, Bednarek et al. 1992; Madden,

Garboczi et al. 1993). Not surprisingly, anchor modification is commonly employed in peptide vaccine design to improve the efficacy of the particular peptide antigen in generating an immune response (Irvine, Parkhurst et al. 1999; Borbulevych, Baxter et al. 2005). In addition, secondary anchor residues, which can be located at various points along the peptide chain, can also modify the affinity of a given peptide for MHC.

1.3.6 Cross Presentation

Cross-presentation is defined as the mechanisms by which exogenous antigens gain access to the MHC class I processing pathway for presentation to cytotoxic $CD8^+$ T-cells. Professional APCs are the principle cells capable of cross-presenting extracellular antigen on MHC class I. During this process the APC acquires antigen from the surrounding environment through endocytic mechanisms, most notably phagocytosis and macropinocytosis. The internalised antigen can then be processed by at least two different mechanisms (Cresswell, Ackerman et al. 2005; Rock and Shen 2005). One method is by antigen escape from the phagosome, thereby entering the usual cytosol-proteasome degradation pathway for MHC class I antigens, however, at present the exact mechanisms of antigen escape are unclear. Secondly the phagosome can fuse with the endosome, thus allowing phagocytosed antigen to be degraded by endosomal proteases (particularly cathepsin S (Shen, Sigal et al. 2004)) and binding to MHC class I molecules within the endocytic compartment itself. Depending upon the nature of the antigen, one or both of these pathways can contribute to cross-presentation *in vivo* (Rock and Shen 2005). The result of cross-presentation can also have a profound effect on the activation or tolerance of an immune response towards a particular antigen, depending upon other factors such as “danger signals” and the presence of co-stimulatory molecules (Kurts, Kosaka et al. 1997). The concept of cross-presentation is

important when formulating potential vaccination strategies because the ability of an antigen to be cross-presented can determine the success of a vaccination in generating $CD8^+$ T-cell responses.

1.4 Immunotherapy

1.4.1 Immunotherapy strategies

Standard treatments for many malignancies involve surgical resection, chemotherapy and/or radiotherapy, however these are subject to numerous side effects, are frequently unable to provide a lasting protection against cancer and for some malignancies these methods prove ineffective. To solve this, many different immunotherapy stratergies have been devised exploiting different aspects of the immune system, all with varying degrees of success. So far Stem Cell Transplantation (SCT) has proved to have powerful anti-tumour effects through allo-reactive responses, as evidenced by the graft versus leukaemia effect (Riddell, Berger et al. 2003). Antibody therapy has also proved very successful in the clinic, with several antibody treatments available for the treatment of different malignancies (Scott 1998; Los, Roodhart et al. 2007). However, these treatments can often have drastic side effects (e.g. graft versus host disease) and in the case of antibody therapy are very expensive and do not form a lasting memory response, therefore relying on a prolonged treatment period. It is therefore desirable to further investigate the possibility of a cellular based treatment for malignancy by harnessing the patient's own immune system. To date, adoptive transfer of MHC matched donor cells into patients treated with radiotherapy has been the most successful treatment for blood borne malignancies but has little effect upon solid malignancies (Dudley, Wunderlich et al. 2002; Dudley and Rosenberg 2003). Other methods

involving *in vitro* priming of autologous T-cells with various tumour antigens have had varying degrees of success depending upon the tumour antigen targeted, method of T-cell priming and method of T-cell administration to the patient. Studies by Rosenberg *et al* have shown promising results of *ex vivo* antigen primed T-cell reactivity against tumours when re-administered back to the patients (Dudley, Wunderlich et al. 2002). In spite of this, the cell numbers used, man hours and the cost of this method would be unsuitable for most clinics around the world. TCR gene transfer of TCRs specific for tumour associated antigens is also the focus of much interest as the transfer of these TCRs into a recipients T-cells can result in circumventing the problems of central tolerance against tumour antigen (Stanislawski, Voss et al. 2001).

As well as T-cell vaccines, DC vaccines are also of key interest (Figdor, de Vries et al. 2004; Melief and Van der Burg 2008) as a successful vaccine would be able to prime a host response in vivo, therefore reducing time and cost in generating antigen specific T-cells, and could potentially lead to a strong lasting memory response involving multiple arms of the immune system. There are many methods with which to prime tumour specific dendritic cell-induced responses used in the literature; such as DNA delivery using virus or gene gun approaches (Scheerlinck, Casey et al. 2001), tumour cell lysate (Gatza and Okada 2002), recombinant protein and exosome delivery (Siguo Hao 2007). Another important variable in DC vaccination is the maturation state of the dendritic cell and the method by which it is matured (Gilboa 2007). Finally, the site and method of administration can also play a key role in the induction of immune responses. For example the site of administration either intra lymphatic, intra dermal or intravenous, and the addition of cytokines all have a key role in the migration and type of immune response generated (Morse, Coleman et al. 1999;

Morel and Turner 2010). To date, Sipuleucel-T (Provenge) is the only cellular based treatment for solid malignancy which has FDA approval for use in the clinic. It is currently used for the treatment of patients with prostate cancer and has been shown to increase the median survival of patients with metastatic prostate cancer by 4.1 months (Kantoff, Higano et al. 2010). The personalised treatment consists of harvesting PBMC from a patient through leukapheresis, followed by incubation of the dendritic cells with a fusion peptide, consisting of the prostatic acid phosphatase (PAP) antigen and granulocyte-macrophage colony stimulating factor (GM-CSF). The activated dendritic cells are then infused back into the patient to facilitate an immune response against PAP expressing tumour cells. Finally, BCG vaccination for the treatment of bladder cancer has been shown to be clinically effective (Biot, Rentsch et al. 2012), however the precise immunological mechanisms of this effect are still largely unknown.

1.4.2 Challenges facing immunotherapy

Clearly due to the immunosuppressive environment and poor immunogenicity exhibited by many tumours (Foss 2002; Zou 2005) there remain many challenges facing future immunotherapy strategies.

Although various immunotherapies have shown some benefit to patients enrolled on clinical trials (Dudley, Wunderlich et al. 2002; Kantoff, Higano et al. 2010), these responses have typically been modest and short lived, highlighting the fact there is still no cell based vaccination that provides complete regression in the majority of cases. This is not only due to each individual's tumour being slightly different to that of another, at various stages of tumour progression and altered antigen display, but also because the survival and progression of malignancy involves complex "evolutionary" processes which tailor the

microenvironment and immune response to evade destruction and aid in tumour formation and metastasis. For solid tumours to grow past a certain size of around $2mm^3$ the tumour must form blood vessels to supply oxygen and nutrient-rich blood to the tumour. This process involves many cytokines and growth factors such as vascular endothelial growth factor (VEGF) produced by the tumour (Gabrilovich, Chen et al. 1996). These molecules have been shown to have a potentially negative impact upon immune responses by down regulating the activation of dendritic cells, thus inhibiting the activation of antigen specific T-cells (Gabrilovich, Chen et al. 1996). Loss of MHC molecules is also a common T-cell subversion mechanism of malignant cells (Bubeník 2004) as cells that do not possess MHC are unable to be targeted by T-cells. However, this would usually result in the targeted destruction of the cell by NK-cells, which rely upon a balance between activatory and inhibitory signals for their activation (Di Santo 2006). The tumour may exploit this by down-regulation of activatory ligand such as MICA and MICB on the cell surface of the tumour and also through production of inhibitory cytokines (Khong and Restifo 2002). This therefore provides immunity to NK-cell attack due to MHC loss and protection against T-cell targeting. Subversion of the immune response can also occur through the tumour's expression of so called "death receptors" such as FasL (Hahne, Rimoldi et al. 1996) and TRAIL (Radoja and Frey 2000) which upon binding counterpart receptors upon T-cells, activate programmed cell death pathways within the T-cells leading to T-cell elimination.

Also the role of the tumour microenvironment and surrounding cell populations can also negatively impact upon immunotherapy strategies. Key cellular components of this immunosuppressive environment include T-reg (Beyer and Schultze 2006) and Monocyte

derived suppressor cells (MDSC) (Lesokhin, Hohl et al. 2012), both of which are able to suppress anti-tumour responses and also aid the growth of the tumour.

Hence there are multiple challenges that need to be considered when designing any potential immunotherapy. However, one key aspect of immunotherapy approaches frequently relies upon the targeting of specific antigens expressed by the tumour. However these antigens are often self-antigens and so the majority of potential tumour reactive T-cells are either deleted from the repertoire, have very low affinity or have become anergic due to peripheral tolerance mechanisms (Zheng, Gao et al. 2002; Mapara and Sykes 2004). One potential approach to a successful immunotherapy strategy therefore involves targeting tumour specific antigen by moderate/high affinity T-cells present within the repertoire. Potential tumour specific antigen targets for immunotherapy include; tumour associated viral antigens (Restifo, Dudley et al. 2012), cancer testis antigens (CTAgs), which are usually expressed only in immuneprivileged sites (such as the testes) but commonly overexpressed by certain tumours (Scanlan, Gure et al. 2002) and antigens which are altered during tumourogenesis, either through mutation or over-expression (Wang, Zhou et al. 2002). Most importantly these antigens must not be presented within the thymus, thus allowing reactive T-cells to evade central tolerance. However, given that some of the methods used to target these antigens, such as TCR gene transfer, can generate potent anti-tumour effects but can also elicit strong auto-immune effects if the TCR reactivity were not specific for tumour antigen (Bendle, Linnemann et al. 2010). A second consideration for the selection of optimal tumour associated antigens for immunotherapy strategies would be the functional importance of these antigens in maintaining the tumour phenotype and survival, which in the case of CTAgs is poorly understood.

1.5 Targeting post-translationally modified antigens as novel tumour antigens for immunotherapy

1.5.1 Types of post-translational modification

One potential approach to address the issues outlined above is to focus immunotherapeutic strategies on tumour associated antigens inherently linked to tumour phenotype in some way. Targeting metabolic changes that occur upon tumourigenisis would be one possible axis for such approaches. Clearly, signalling pathways are well recognised to be altered in cancer cells, with alterations in kinase pathways highlighted as being particularly relevant (Greenman, Stephens et al. 2007). This raises the question of whether post-translational modification, and especially phosphorylation, might provide a means to specifically target cancer cells. Clearly a wide range of post-translational modifications occurs in cells including; glycosylation, citrinulation, acetylation and phosphorylation. However, phosphorylation of cellular proteins is one of the most wide spread, in addition, kinase pathways are also altered/upregulated in cancer and are a hallmark of malignant transformation (Vogelstein and Kinzler 2004; Greenman, Stephens et al. 2007). Targeting kinases themselves has proved promising for some cancers (Baselga 2011; Gherardi, Birchmeier et al. 2012) and raises the question of whether the altered phosphoproteome in cancer cells gives rise to an immunological signature of tumourigenisis, and one that may be targeted therapeutically.

1.5.2 Phosphorylation

The pool of phosphoproteins that could potentially provide a source of class I MHC epitopes is large. Firstly, up to a third of cellular proteins are phosphorylated (Alberts, Johnson et al. 2002). Secondly, the majority of these take place in the cytosol, the key site where proteins are degraded for class I MHC presentation by the proteasome. Finally, phosphorylation is

known to play an important role in the regulated destruction of certain cellular proteins (Harper 2002; Hao, Oehlmann et al. 2007). The source of phosphorylated antigen for class I MHC presentation can arise from a number of different cellular processes and can be from functional proteins that have been targeted for degradation in order to limit their activity or from a large pool of miss-folded/abnormalities that may occur during post translational modifications, such as aberrant phosphorylation. This highlights the potential for phosphorylation to provide antigens for presentation by MHC class I molecules to the immune system. Since many viral proteins are also phosphorylated during an infection, it is conceivable that these may also be presented by class I MHC molecules. Therefore, mechanistically how phosphorylated proteins might be processed and presented on MHC molecules is an important area of study in order to fully understand the repertoire of antigens that are available for immune recognition. Phosphorylated proteins therefore provide a potential source of antigens for class I and II MHC presentation that are functionally linked to the transformation process and might prove tumour selective.

1.5.3 Presentation of phosphopeptide antigens

Within the last decade it has become clear that post translationally modified proteins, including phosphoproteins, can enter the class I presentation pathway, ultimately giving rise to phosphopeptides that can be presented on class I molecules at the surface of cells, for recognition by the immune system. The first study to demonstrate natural phosphopeptide presentation on class I MHC (Hogan, Eisinger et al. 1998) involved sequencing of peptides eluted from class I molecules by mass spectrometry, and revealed presentation of a single phosphopeptide by HLA-A68.2 on the surface of a lung tumour cell line. However this study did not study the further effects of phosphorylation in antigen binding and T-cell recognition.

In 2000 Zarling *et al* used mass spectrometry to identify the presence of phosphopeptides on a range of MHC class I alleles (Zarling, Ficarro et al. 2000). By using various EBV transformed LCL cell lines, each expressing a different class I MHC allele, it was demonstrated that each class I MHC allele could present a varying degree of phosphopeptides and interestingly found that HLA-B alleles present a wider range of phosphopeptides than HLA-A alleles. The authors proposed that this is likely to reflect the common structural features shared by MHC binding motifs and phosphokinase motifs. Presentation of phosphopeptides was also shown to be completely TAP dependent and therefore many of these peptides were likely to originate from processed proteins in the cytosol. The selective recognition of phosphopeptide by mouse CTL was also demonstrated, which I shall go on to discuss further in section 1.5.4. This study therefore provided the first comprehensive insight into how posttranslational effects such as phosphorylation might play an important role in antigen presentation and immune recognition. However the study did not fully resolve the structural aspects of how these phosphopeptides might bind to class I MHC molecules.

To further characterise the binding mode and sequence repertoire of phosphopeptide antigens presented by class I MHC molecules in a cancer background, Zarling *et al* in 2006 published results that identified a common motif of phosphopeptides presented by HLA-A*0201 in various cancer cell lines (Zarling, Polefrone et al. 2006). 36 phosphopeptides were found to be presented by HLA-A*02001 (HLA-A2) on one or more of the 4 cell lines tested, each containing a single phosphoserine (pSer) or phosphothreonine (pThr). Of these 9-13 amino acid phosphopeptides, 68% contained a pSer or pThr at P4. In addition, 62% of the phosphopeptides contained a charged amino acid residue at the P1 position (Arg or Lys). However, of the peptides that contained pSer at P4, 50% also contained a positively charged

amino acid at P1. Finally, 41% of the phosphopeptides contained subdominant anchor residues (Met, Ile, Val or Gln rather than the optimal Leu) at P2 compared to only 23-24% of non-phosphorylated peptides. Therefore a distinct canonical HLA-A2 binding phosphopeptide motif had been determined, which contained a phosphorylation at P4, a positively charged residue at P1 and could tolerate a higher frequency of subdominant anchor residues compared to non-phosphorylated peptides.

The molecular basis of canonical phosphopeptide presentation by HLA-A2 was first investigated by Mohammed *et al*, who compared the structure of four canonical phosphopeptide/HLA-A2 complexes determined by x-ray crystallography (Mohammed, Cobbold et al. 2008). The four peptides were canonical phosphopeptides in that they all contained a P4 phosphorylation and a charged P1 residue. The group included one nonamer and three decamers. The overall mode in which phosphopeptides bind to the HLA-A2 molecule was found to be similar to that of non-phosphorylated peptides in that the stabilising interactions usually observed between the peptide termini and the A and F pockets of the HLA-A2 groove are conserved for all peptides. The anchor residues at P2 and the carboxy-terminus were also orientated similarly to most previous HLA-A2 structures. In all four phosphopeptides, the phosphate at P4 was structurally conserved in that the phosphate head group was orientated upwards and towards the N-termini P1 position. The highly conserved orientation of the phosphate head group allows for a variety of stabilising interactions between the phosphate, P1 residue and residues of the HLA-A2 peptide binding groove. HLA-A2/phosphopeptide binding assays confirmed that the phosphate can have a profound effect upon MHC binding affinity. Finally, the study suggested that phosphorylation might affect overall peptide conformation itself. These data therefore not

only suggest that phoshopeptides can bind with altered affinities compared to their non-phosphorylated counterparts, due to the ability to tolerate poor anchor residues, but also that the phosphate could potentially play a key role in antigen presentation to TCR; not only through direct interaction with the solvent exposed phosphate but also through possible conformational change upon addition of the phosphate moiety.

A second study by Petersen *et al* was the first to directly compare the structures of HLA-A2 associated phosphorylated epitopes and non-phosphorylated counterparts (Petersen, Wurzbacher et al. 2009). The study compared one canonical epitope, RVASPTSGV (RVAS), and two non-canonical epitopes, GLLGSPVRA (GLLGS) and YLDpSGIHSGA (YLDS), in their phosphorylated and non-phosphorylated forms. Most strikingly the study demonstrated that for these epitopes phosphorylation has very little effect upon overall antigen conformation. This highlights that discrimination of these antigens by phosphopeptide specific T-cells is likely to rely solely on direct binding of the TCR with the phosphate moiety.

1.5.4 T-cell recognition of post-translational modified antigens

In addition to how phosphopeptides are presented by MHC molecules, which is still relatively poorly understood, our understanding of how they are recognised by the TCR is another area worthy of investigation. Of relevance, mice immunised with phosphopeptide pulsed DC generated from bone marrow stem cells are able to prime phosphopeptide specific T-cells *in vivo*, and expand CD8+ T-cell populations able to recognise peptide pulsed cells in a phosphate-dependent and epitope-specific manner *in vitro* (Zarling, Ficarro et al. 2000). More recently a study reported that phosphopeptide-specific T cells generated in a similar manner kill human tumour cell lines presenting phosphopeptide antigen (Zarling, Polefrone et al. 2006).

These studies provide somewhat limited insight into how T-cells are likely to recognise class I MHC-restricted phosphopeptides. Although they indicate that phosphopeptides can survive *in vitro* cross-presentation, confirmed by pulsing exogenous antigen onto DC and using these cells to prime T-cell responses, both studies use phosphopeptides from a human source to generate T-cells from mice. Thus the lack of central and peripheral tolerance in this scenario suggests this is in effect similar to immunising with foreign antigen. Therefore, the experimental system used in these studies is likely to be very different to priming of natural responses to phosphopeptide antigens, or of therapeutic vaccination strategies in a human setting. Despite these caveats, these two studies provide clear evidence that tumour cells presenting phosphopeptides can in principle be recognised by T-cells in a manner that is both phosphate-dependent and epitope-specific.

Collectively, these results along with structural studies of phosphopeptide-HLA-A2 complexes have generated some hypotheses as to how phosphate-specific recognition might take place. In the case of antigens such as RVAS, discrimination of phosphopeptide epitopes by phosphopeptide specific T-cells may rely solely upon direct TCR interaction with the phosphate moiety, due to similar conformations. In summary, the full impact of phosphorylation on antigen structure and discrimination by the TCR remains unclear.

1.5.5 Phosphopeptides as novel tumour antigens for immunotherapy

As mentioned previously, many of the abnormalities in cancer also occur in posttranslational modification events and targeted activation/deactivation of cell signalling pathways (Greenman, Stephens et al. 2007). Of these, phosphorylation is of key importance. These modifications are able to be presented by MHC class I molecules on the surface of tumour cell lines and have the potential to be recognised by T-cells in a phosphate as well as epitope

specific manner (Zarling, Ficarro et al. 2000; Zarling, Polefrone et al. 2006). These therefore represent a novel and qualitatively distinct set of tumour associated antigens that are linked to the transformation process. Although the presence of phosphopeptide/MHC complexes has been shown, the immunological relevance to cancer therapy in a human setting has not and questions in this area still remain.

Despite these issues, phosphopeptides are arguably promising targets for future immunotherapy approaches. However, to fully comprehend the mechanisms of presentation and the responses generated by phosphoantigens, a molecular understanding of their presentation and recognition is important. By understanding the structural basis of phosphopeptide presentation, such as the effects of phosphate position on MHC stabilisation, peptide conformation and TCR recognition, may in the future provide the ability to predict how a phosphorylated antigen might function as a potential target for cancer therapy. Just as the determination of primary anchor residues for class I MHC molecules helped improve the design of new vaccines for both viral and cancer therapy, so might the knowledge generated by studying the structure of phosphoantigen presentation and recognition help improve future vaccine design.

1.6 Aims

Clearly there remain many unresolved questions regarding phosphopeptide presentation and recognition. This study aims to directly address some of these issues by determining the effect of phosphorylation on epitope binding and overall conformation, and also, whether epitope specific recognition, which is clearly highly dependent upon the phosphate at a

cellular level, is innate to the TCR. These issues were examined through X-ray crystallographic and biophysical analysis methods.

2 Methods

2.1 Molecular Cloning Techniques

2.1.1 Plasmids

pET23a vectors (Novagen) were used for transformation and inducible expression of proteins to high-levels in the form of *E. coli* inclusion bodies. HLA-A2 and HLA-B7 heavy chains (encompassing α1, α2 and α3 domains), β2-microglobulin (β_2M) and LILRB1 encoding pET23a plasmids had previously been generated in the Willcox laboratory prior to the commencement of my PhD. Constructs enabling biotinylation of HLA-A2 heavy chain were also previously produced by cloning the HLA-A2 heavy chain region into the JMB002 biotinylation construct (generated by Jonathan M. Boulter and based on the pGMT7 vector). Human CD8α constructs (Geneservices 100005953 ORF plasmid) were sub-cloned into pET23a vectors as described (Gao, Gerth et al. 1998). For characterisation of the phosphopeptide specific T-cell receptor (TCR) binding, constructs of the IRS2, IRS2BK and CDC25B αβ TCR clones were kindly donated by Angela L. Zarling (University of Virginia, USA) and utilised as template for sub-cloning into TOPO-TA, pMT/BiP/V5-His-A (*Drosophila* expression vector) and pMT/Bip/V5-BirA-His (Novagen). The hygromycin-encoding co-transfection plasmid, pCoHygro, was obtained from Novagen.

2.1.2 Bacterial cell culture

DH5α (Invitrogen), TOP10 (Invitrogen) or BL-21 (Novagen) *E. coli* cells were cultured in Luria Bertani (LB) media (10g Bactotryptophan, 10g NaCl and 5g Yeast extract per litre of 18 Ohm milli-Q water) or low salt LB (10g Bactotryptophan and 5g Yeast extract per litre). LB agar plates were prepared by adding 15g of agar powder to a litre of LB. All media was autoclaved prior to use. Ampicillin was added to the culture media (following autoclaving and cooling to 50°C) at a final concentration of 100µg/ml. LB agar plates were stored at 4°C.

2.1.3 Preparation of chemically competent bacteria

Stocks of chemically competent *E. coli* DH5α (Invitrogen) or BL-21 (Novagen) cells were produced by streaking for single colonies onto LB agar plates containing no antibiotic and incubating at 37°C for 16 hours. A single colony was then selected and inoculated into a 5ml starter culture of LB media with no antibiotic and grown at 37°C, 180rpm until slightly cloudy. This was then diluted into 100ml of fresh LB media and incubated at 37°C as above. Once the cells had reached an OD_{600} of 0.4-0.6 the culture was spun at 4000rpm in a Sorval RC 5C plus, using the SS-34 rotor for 15 minutes at 4°C. The resulting supernatant was removed and the pellet re-suspended in 20ml of pre-chilled TFB1 (30mM Potassium acetate, 50mM Manganese chloride, 100mM Potassium chloride, 10mM Calcium chloride, 15% v/v glycerol). A second round of centrifugation was then performed as described above. Following this the supernatant was removed before re-suspending the pellet in 4ml of pre-chilled TFB2 (10mM Sodium-MOPS pH 7.0, 75mM Calcium chloride, 10mM Potassium chloride, 15% v/v glycerol). The cell suspension was rapidly divided into 200µl aliquots and flash frozen in liquid nitrogen. Chemically competent bacterial cells were stored at -80°C prior to use.

2.1.4 Bacterial transformation with plasmid DNA

50-100µl of chemically competent E. *coli* cells was thawed on ice. Typically, 0.5-1.0µg of plasmid DNA was then added and the cells and incubated for 30 minutes on ice. The cells were then heat shocked at 42°C for 90 seconds, 500µl SOC media (20g Bactotryptophan, 0.5g NaCl, 5g Yeast extract, 2.5Mm KCl, 10mM $MgCl_2$ and 20mM glucose per litre of 18 Ohm milli-Q water) added and incubated at 37°C with shaking at 180rpm for 1 hour. Following incubation, 150µl of the cells were subsequently plated out onto LB/Amp plates and incubated at 37°C for no more than 16 hours.

2.1.5 Insect S2 cell culture and transfection with plasmid DNA

Non-transfected Parental S2 cells (Invitrogen) were cultured in complete growth media consisting of; Schneiders Drosophila medium (Lonza), 1.8g/L L-Glutamine and 10% FCS (Geneflow). For transfection of parental S2 cells with plasmid DNA, 3ml of cells were plated out into a single well of a 6 well plate (Iwaki) at a density of $1x10^6$ cells/ml. These were then grown until the cell density reached $2\text{-}4x10^6$ cells/ml, which was typically after 36 hours. The Invitrogen Calcium Phosphate transfection kit was used for all transfections. Firstly Solution A (240mM $CaCl_2$, 9.5µg plasmid DNA, 0.5µg pCoHygro plasmid DNA, sterile H_2O to final volume of 150µl) was slowly mixed with 150µl of Solution B (2X HBS) and incubated at room temperature for 30-40 minutes. The resultant mixture was then added drop wise to the cells and incubated for 24 hours at 27°C. Media was then removed and the wells washed twice with fresh media. Each wash was pooled and the cells pelleted at 100xg for 10 minutes. Cells were re-suspended in 3ml fresh complete growth media, returned to the same well and incubated at 27°C for 2 days. Finally the cells were harvested and transferred to selective media (complete growth media containing 300µg/ml hygromycin B (Roche)) which was

replaced every 4-5 days. After 12-14 days cells were transferred to T25 flasks (Iwaki) in 5-6ml selective media.

2.1.6 Isolation of plasmid DNA

Plasmids were transformed into either DH5α or TOP10 (Invitrogen) E. *coli* cells as mentioned above. After 16 hours of incubation at 37°C a single colony was picked from the plate and transferred to 5ml of LB media containing 100µg/ml Ampicilin. The cultures were incubated overnight at 37°C, with vigorous shaking at 180 rmp. The cells were then pelleted using a 1.5ml tube in consecutive 5 minute spins at 17,000 g using a bench top centrifuge. The supernatant was discarded after each spin. Once all cells were pelleted, plasmid DNA was then purified from the cell pellet using QIAprep Spin Miniprep Kit (QIAGEN) using the manufacturer's instructions and reagents. Briefly; this involved re-suspending the cell pellet, lysing the cells with an alkaline buffer containing sodium hydroxide, neutralisation in high salt conditions and finally centrifugation (13,000 g for 10 minutes) through silica matrix causing adsorption and separation of DNA from lysate. Following this the matrix was washed using an ethanol-based solution, and DNA was then eluted with 50µl PCR clean dH_2O. DNA stocks were stored at -20°C.

2.1.7 Polymerase chain reaction

Typically, 50µl reactions were performed containing; <0.1µg DNA (0.5-1µl from a QIAprep spin Miniprep sample), 20pmol of each forward and reverse primer, 10mM dNTPs, 5µl of relevant 10X stock enzyme buffer, 1µl PFU Ultra or Expand DNA polymerase enzyme and made to a final volume of 50µl with PCR clean dH_2O. Reactions were carried out using a 2720 thermocycler (Applied Biosystems). The program used for individual reactions varied,

however a typical example consisted of; 95°C for 2 minutes, 95°C for 45 seconds, 65°C for 1 minute, 72°C for 5 minutes followed by 24 cycles of 95°C for 45 seconds, 65°C for 1 minute, 72°C for 5 minutes. PCR products were then analysed by agarose gel electrophoresis.

2.1.8 Agarose gel electrophoresis

Agarose gels were made by adding 0.8-1.0% (w/v) agarose powder to TBE buffer (89mM Tris base, 89mM boric acid, 2mM EDTA) and heating to dissolve the agarose. Before cooling, ethidium bromide solution was mixed to a final concentration of ~0.5µg/ml with the agarose before pouring. A 10X loading buffer consisting of 16g sucrose, 0.25% (w/v) Bromophenol Blue, 40ml dH_2O was used in the preparation of all samples. Agarose gels were immersed in TBE buffer before samples were loaded and gels ran at 80V. The size of DNA fragments were assessed by comparing their migration to a 100 base pair marker ladder (Roche). The ladder bands are sized 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500 and 2642 base pairs. Gels were examined and recorded under UV light using a Syngene InGenius image capture system.

2.1.9 Gel purification of DNA

Digested plasmid and insert DNA were visualised under UV light. Bands were excised from the agarose gel and subsequently purified using the QIAquick gel extraction kit following the manufacturer's guidelines. Briefly, agarose bands were dissolved at 50°C in buffer QG (this buffer provides the appropriate pH for optimal DNA binding to the silica membrane), DNA is then bound to a silica matrix and washed with further buffer QG to remove residual impurities followed by a final wash with buffer PE containing ethanol to remove salts. DNA was then eluted from the matrix using 50µl PCR clean dH_2O.

2.1.10 Restriction enzyme DNA digests and DNA ligation

The restriction enzymes XbaI, BglII and XhoI (Roche) and the appropriate buffers were used to digest insert and vectors in preparation for cloning and also for verification of successful cloning. Restriction digests were performed for 2 hours at 37°C. To prevent re-circularisation of cut vectors following restriction enzyme digest, 1µl of alkaline phosphatase (NEB) was added and incubated for 1 hour at 37°C. Following agarose gel purification of the insert and vector, DNA ligation was conducted overnight using T4 DNA ligase and the appropriate DNA ligase buffer, as per manufacturer's instructions (NEB).

2.1.11 DNA sequencing

DNA sequencing reactions were performed by the Functional Genomics service, University of Birmingham, utilising the ABI BigDye Terminator v3.0 reaction and an ABI 3730 Sequencer.

2.2 General Protein Techniques

2.2.1 S2 insect cell protein expression

Transfected S2 cells were sub-cultured into 500ml Erlenmyer flasks (Corning) at a cell density of $1.2x10^6$ cells/ml in 125ml of selective media supplemented with 0.05% Pluronic F-68 surfactant (Gibco) and incubated for 5 days at 27°C with constant shaking at 80rpm. Following this cells were transferred into 250ml complete growth media plus 0.05% Pluronic F-68 surfactant in 1L Erlenmyer flasks at a cell density of $1.2x10^6$ cells/ml. Upon cells reaching a cell density of $3\text{-}5x10^6$ cells/ml (typically 3 days), the cells were induced by addition of 500µM copper sulphate ($CuSO_4$). The cells were then incubated for a further 5 days before harvesting of the supernatant by centrifugation at 100g for 20 minutes followed

by 5000 rpm for 20 minutes in a Beckman Coulter Avanti 5-26xp centrifuge using the JLA8.1 rotor.

2.2.2 Ni-NTA column purification of His tagged proteins

Proteins generated in the S2 insect cell expression systems contained a C-terminal His tag allowing purification of expressed proteins supernatant. One litre of supernatant was first dialysed overnight in 10L of phosphate buffered saline (PBS) (Oxoid) before addition to a Ni-NTA column overnight at 4°C. The column was then moved to room temperature and washed with 300-500ml of 10mM Imidazole in PBS to remove unbound proteins. Elution of His-tagged proteins was performed with 250mM Imidazole in PBS. Eluted protein was buffer exchanged using a PD-10 column into 5mM Tris pH 8, 150mM NaCl before storing at -20°C.

2.2.3 E.*coli* protein expression

BL-21 E. *coli* cells (Novagen) were transformed with plasmid DNA as described above. After 16 hour incubation on antibiotic plates a single colony was selected and transferred to 10ml of LB containing 100µg/ml ampicillin. The cells were then incubated at 37°C with vigorous shaking at 180 rpm for ~3 hours or until the media turned cloudy. The starter culture was then used to inoculate 2 litres of LB-ampicillin media (ampicillin at final concentration of 100µg/ml) using 5ml of the starter culture per litre. This was further incubated as above. As the optical density increased, 1ml samples were removed and the OD_{600} was measured against an LB media blank. Upon reaching the desired OD range of 0.4-0.6 the cultures were induced with 1M IPTG to a final concentration of 0.5mM and incubated under the same conditions for 4 hours. After this period the cultures were incubated at 6°C, with reduced shaking of 100 rpm.

2.2.4 Inclusion body purification

Inclusion bodies were harvested from post induction protein expression cultures. The cells were pelleted at 6000 rpm for 20 minutes at 4°C (Beckman Coulter Avanti 5-26xp centrifuge using the JLA8.1 rotor), supernatant discarded and the cells re-suspended in 50ml ice-cold PBS. The cells were then sonicated for five 45 second bursts at 50% duty cycle, using a Misonix Sonicator and a standard 12mm probe. The cells were then centrifuged at 13500 rpm in a Sorval RC5C plus using the SS-34 rotor at 4°C for 15 minutes. The supernatant was discarded and the inclusion body homogenised in Triton wash buffer (50mM Tris-HCl, 0.5% Triton-X100, 200Mm NaCl, 10Mm EDTA, 0.1% (w/v) Sodium Azide, 2mM DTT, pH 8.0) so that cell debris and membrane components were removed, before being pelleted again as above. This wash step was repeated two additional times. Following the third wash with Triton wash buffer, the pellet was re-suspended and homogenised in wash buffer containing no Triton-X100 (50mM Tris-HCl, 1mM EDTA, 0.1% (w/v) Sodium Azide, 2 mM DTT, pH 8.0) so as to remove detergent and salt from the inclusion bodies. The purified inclusion bodies were then finally pelleted as above and solubilised in a denaturant urea solution consisting of; 8M Urea, 50mM MES pH 6.5, 10mM EDTA and 2mM DTT or for CD8αα inclusion bodies; adjustment of pH to 8.0 along with addition of 500mM NaCl and 10% Glycerol. Inclusion bodies were then incubated overnight at 4°C on a rocking platform. Insoluble material was separated from the soluble inclusion bodies by centrifugation at 15000 rpm for 30 minutes at 4°C (Sorval). Supernatant was harvested into 1ml aliquots and stored at -80°C. Purity and concentration was assessed by SDS-PAGE and Bradford-Dye binding assay respectively.

2.2.5 SDS-PAGE

SDS-PAGE separating gels of 15% acrylamide or 12.5% acrylamide consisted of; 4.9ml SDS solution (115ml deionised water, 125ml 1.5M Tris pH8.8 and 5ml 10% SDS), 5ml 30% acrylamide, 100µl 10% APS and 4µl TEMED for the 15% solution and 5.74ml SDS solution (78.5ml deionised water, 62.5ml 1.5M Tris pH 8.8 and 2.5ml 10% SDS), 4.16ml 30% acrylamide, 100µl 10% APS and 4µl TEMED for the 12.5% gels. This was then transferred to a 1mm BIORAD gel cassette. 1ml of Isopropanol was added in order to level the gel. Stacking gels consisted of; 3.2 ml SDS solution (136ml deionised water, 25ml 1M Tris pH6.8 and 2ml 10% SDS), 0.67ml 30% acrylamide, 40µl 10% APS and 4µl TEMED. The stacking gel was then added to the top of the dried separating gel and a 1.0mm comb inserted.

Samples were prepared by mixing ~1-20µl of protein with 2-4µl of 5x loading buffer (made by mixing 5ml 1M Tris pH 6.8, 3g SDS, 1g bromophenol blue, 100µl 2M DTT in a final volume of 25ml with deionised water) and making up to a final volume of 10-20µl with deionised water. Samples were then boiled at 90°C for 10 minutes before loading onto the SDS-PAGE gel.

SDS-PAGE running buffer was made up of 75g Tris base, 470g glycine and 25g SDS and made up to 2.5L with deionised water. SDS-PAGE was carried out using a Bio-Rad SDS-PAGE apparatus that was set to 100 volts for 10 minutes through the stacking gel and then increased to 180 volts for electrophoresis in the separating gel. SDS-PAGE gels were stained for 30 minutes with Coomassie Blue (1.25g Coomassie Brilliant Blue R250, 250ml methanol and 50ml acetic acid) and destained overnight with destain solution (300ml water, 150ml methanol and 50ml acetic acid).

2.2.6 Quantification of protein

For solubilised inclusion body preparations an approximation of protein concentration in each sample was determined by a Bradford dye binding assay. 1ml of Bio-Rad Protein Assay Dye Reagent Concentrate was added to 4ml milli-Q water to make a working stock solution. An aliquot of the protein sample was taken and diluted 1:10 in 8M urea. This was then mixed with the working stock dye solution in a 1:200 mix. The absorption was then measured at 595nm using an 8M urea control sample as a blank. The concentration of protein was calculated on the basis that 1µg of Bovine Serum Albumin has an absorbance of ~0.05 at 595nm.

For refolded complexes, accurate determination of the protein concentration was performed by measuring the absorbance of a 1:50 protein:buffer solution at 280nm. The extinction coefficients for each complex were determined using the following equation:

$$\frac{(W\times 5500)+(Y\times 1490)+(S\times 125)}{Mr}$$

Where W is the number of tryptophan residues, Y is the number of tyrosine residues, S is the number of disulphide bonds and Mr is the molecular weight of the protein. Absorbance readings at 280nm were then divided by the extinction coefficient to provide protein concentrations in mg/ml. The extinction coefficient for HLA-A2 and HLA-B7 complexes was determined to be 1.5. The TCRs, LILRB1 and CD8αα were determined to have an extinction coefficient of 1.5, 2.2 and 1.3 respectively.

2.2.7 Protein refolding

HLA-A2 complexes containing single phosphopeptides were produced for crystallisation trials using dilution refolding protocols described by Garboczi et al (Garboczi, Hung et al. 1992), allowing renaturation of β_2M and HLA-A2 heavy chain with incorporation of the desired synthetic phosphopeptide. Refolding buffers (100mM Tris pH 8.0, 2mM EDTA, 0.4M L-arginine, 0.5mM oxidised glutathione, 0.5mM reduced glutathione and 0.1mM PMSF) were maintained at 4°C with constant stirring. 10mg of denatured β_2M was added to the 500ml refold solution by taking a third of the β_2M solution and mixing with ~20ml of the refold buffer in a 25ml universal. The diluted β_2M was then added back to the main refold buffer solution slowly in a drop wise fashion and this procedure repeated for the other two thirds of the β_2M. Adding the protein in this pulsed fashion is thought to aid in the correct folding of the protein by decreasing the total concentration of denatured protein present at one time, and is likely to increase the efficiency of renaturation. The refold was then left for at least 30 minutes. 2mg of the relevant synthetic phosphopeptides/peptides (Sigma, EZBiolabs and Genescript) were then added to the refold mix. At the same time 10mg of HLA-A2 heavy chain was added to the refold in the same fashion as the β_2M. The refold was then left for more than 4 hours before subsequent additions of 10mg HLA-A2 heavy chain. The total amount of HLA-A2 heavy chain added to each refold was 50mg. The refold mixes were concentrated down to ~10ml using Amicon stirred cells that were pressurised using gaseous nitrogen. Samples were then centrifuged at 670 g for 5 minutes and filtered using a 0.45µm syringe filter (Sartorius) to remove any precipitated/aggregated protein before loading onto a Pharmacia 26/60 Superdex 200 size exclusion chromatography column (GE Healthcare).

Refolding of β_2M alone was carried out using the same refold conditions as HLA-A2 complexes. To the 500ml refold 5 x 10mg pulses of β_2M were added over 3 days before loading onto a Pharmacia 26/60 Superdex 200 size exclusion chromatography column

Refolding of HLA-B7 complexes was similar to HLA-A2 complexes, however different amounts of 12mg, 15mg and 4mg o β_2M, MHC heavy chain and peptide respectively were used. The β_2M used in these refolds was also pre refolded and purified by size exclusion then concentrated to 15-20mg/ml before adding the 12mg directly to the refold mix. The heavy chain was initially diluted to 1mg/ml in 1M urea before addition to the refold and was added in 3 x 5mg amounts 24 hours apart over three days. The refold mix was incubated at 4°C for 2 weeks with constant stirring before purification by size exclusion chromatography as described above followed by anion exchange using a Resource Q anion exchange column (GE Healthcare).

The renaturation of LILRB1 was performed in a similar manner to β_2M. However before refolding commenced the refold buffer was pH adjusted to pH 8.3.

CD8αα refolding was performed in a 500ml refold buffer containing; 200mM Tris pH 8, 10mM EDTA pH 8, 1M L-Arginine, 6.5mM cysteamine, 3.7mM cystamine and 0.1mM PMSF. Prior to addition to the refold buffer 20mg of inclusion body was incubated for 30 minutes in 15ml of a reducing and denaturing buffer (6M Guanidine-HCl, 50mM Tris-HCL pH 8, 10mM EDTA pH 8, 100mM NaCl, 10mM DTT). A total of 60mg CD8α was added to the refold over 3 days then left for 48 hours before purification by size exclusion (26/60 S200) and cation exchange (Resource S; GE Healthcare) chromatography.

2.2.8 Size exclusion chromatography

Size exclusion chromatography was performed with a Pharmacia 26/60 Superdex 200 column in 20mM Tris pH 8, 50 mM NaCl (5mM NaCl if loading onto Resource Q or Resource S columns following size exclusion), using the Pharmacia AKTA FPLC system. Samples were loaded in volumes of 5-10ml with a pre-washed super-loop (10ml), and the flow rate was set to 3ml/minute. Fractions were collected and concentrated by centrifugation using an Amicon Ultra (Millipore) at 670 g until the volume reached ~250µl. The concentrate was then transferred to a microcon centrifugal filter units with 10 KDa cutoff membranes (Millipore) and centrifuged at 4,100 g in a bench top centrifuge until the volume reached ~100µl.

2.2.9 Anion and Cation exchange chromatography

Anion and cation exchange chromatography was performed using 6ml Resource Q and Resource S columns, respectively with a flow rate of 4ml/minute. Columns were pre-equilibrated with low salt buffer (Buffer A), typically 5mM NaCl. Once protein was loaded a gradient of either 50% or 100% of a high salt buffer containing 1M NaCl (Buffer B) was run over 10-15 column volumes. Fractions were collected every 0.5ml and concentrated as above.

2.2.10 Western blotting

Western blotting was conducted by wet transfer electroblotting of proteins (including pre-stained markers, NEB) from 12.5% SDS-PAGE gels to PDVF (Polyvinylidene Fluoride) membranes in the presence of Transfer Buffer (25mM Tris-HCL pH 8, 192mM glycine, 10% methanol), followed by a wash in TBS-T (100mM Tris-HCL pH7.5, 150mM NaCl, 0.1% Tween

20) and overnight blocking at 4°C in TBS-T plus 5% powdered milk. Probing with anti-His antibody (Invitrogen), made up at a dilution of 1:1000 in TBS-T plus 5% milk, was performed for 1 hour at room temperature with agitation. This was followed by three 10 minute washes with TBS-T and incubation with goat anti-mouse-HRP-conjugated secondary antibody at a dilution of 1:10,000 in TBS-T for 1 hour at room temperature. Blots were then washed three times in TBS-T before enhanced Chemiluminescence detection was performed using the Geneflow ECL Kit. Exposure of Kodak Scientific Imaging film (Kodak) was performed in a darkroom using an intensifier cassette and a developer machine (Kodak).

2.2.11 3C cleavage of T-cell receptors

Cleavage of purified αβ T-cell receptors, containing a 3C cleavage site following the C-terminal cysteine of each chain, was performed using HRV 3C protease enzyme (Novagen) at a concentration of 1U/100µg protein. The cleavage reaction was incubated overnight at 4°C and then either analysed directly by 12.5% SDS-PAGE or cleaved T-cell receptors purified by addition of 200µl Ni-NTA beads, overnight incubation at 4°C with constant mixing and the supernatant harvested.

2.2.12 Biotinylation of protein complexes

Protein complexes were buffer exchanged using PD-10 columns into Biotinylation buffer (100mM TRIS pH 8, 20mM NaCl, 5mM $MgCl_2$ and 0.1mM PMSF). Biotinylation reactions were carried out by addition of 0.5mM d-Biotin, 5mM ATP, BirA enzyme and a 1X protease inhibitor cocktail (Sigma), then incubated in the dark at room temperature overnight.

2.3 Surface Plasmon Resonance

SPR binding studies were performed at 25°C using a BIAcore™ 3000 (BIAcore AB) using HBS-EP as running buffer (10mM HEPES pH 7.4, 150mM NaCl, 3.4mM EDTA, 0.005% surfactant P20). For immobilisation of His-tagged proteins, Ni-NTA chip surfaces were first regenerated to remove metal ions that may have been present in previously used reagents by addition of Regeneration solution (10mM HEPES pH 8.4, 150mM NaCl, 350mM EDTA, 0.005% surfactant P20) followed by a one minute injection of Nickel solution (500µM $NiCl_2$ in running buffer). Streptavidin (Sigma) was covalently coupled to CM5 sensor chips (BIAcore AB) via primary amines using the Amine Coupling Kit (BIAcore AB). Streptavidin was dissolved in 10mM sodium acetate pH 5.5 and injected over the sensor chip at 0.5mg/ml during coupling. Analytes were kept on ice and centrifuged at 17,000 g in a bench top centrifuge prior to use in order to remove aggregated proteins. Analytes were serially injected over all four flow cells at a flow rate of 10µl/minute. For test injections, 10µl was injected (60 second contact time). Data were collected at 10Hz. Data analysis was performed using BIAeval software (BIAcore AB) and Origin Pro 8 graphing and statistical analysis software (OriginLab Corporation).

2.4 Overview of X-ray crystallography

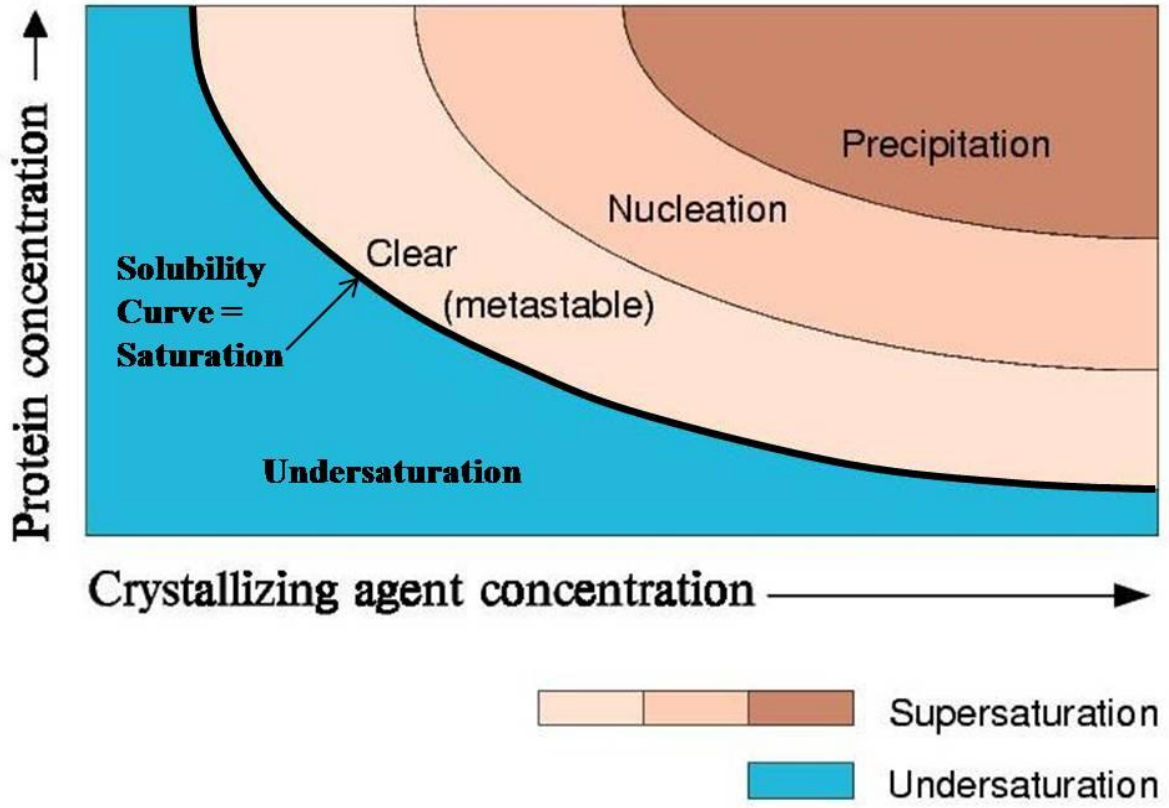
2.4.1 Crystal generation

Crystallisation is achieved by slowly increasing the protein concentration to the point at which the protein molecules are forced to interact with one another in an ordered manner. This is termed the supersaturated state and may further develop into a crystalline or an

amorphous state. Crystal growth can be separated into three phases termed nucleation, growth and cessation. During the nucleation phase the formation of well ordered aggregates occurs. This is followed by the growth phase whereby protein molecules are recruited to the growing face of the crystal to form large well ordered crystals. The final stage is cessation in which the growth of the crystal halts due to reduced levels of free protein molecules in the solution or through incorporation of impurities within the crystal itself. The conditions affecting the stages of crystal growth are often represented in a phase diagram (Figure 2.1). This can be separated into two zones; the unsaturated zone, where crystal growth does not occur, and the saturated zone, which facilitates crystal growth and is located above the solubility curve. The saturation zone can be further divided into three regions. The clear or metastable region is where crystal growth can occur but not the initial crystal formation. The nucleation zone is the phase in which molecules interact with one another and precipitate into ordered crystalline structures. The third phase is the precipitation stage in which protein molecules interact with one another to form an amorphous precipitant.

The overall aim in the generation of crystals for X-ray crystallographic analysis is to identify a crystallisation condition in which the saturation point is just within the nucleation zone and above the metastable zone. This facilitates a slow and reduced crystal formation with less crystal defects in comparison to a more saturated solution, which typically results in rapid formation of many nucleation sites resulting in the formation of small and often defective crystals. As the crystals grow the saturation levels drop to the metastable phase, in which nucleus initiation will cease and the stable growth of the few nucleations present will increase the size of existing crystals.

Figure 2.1 Two dimensional crystallisation phase diagram

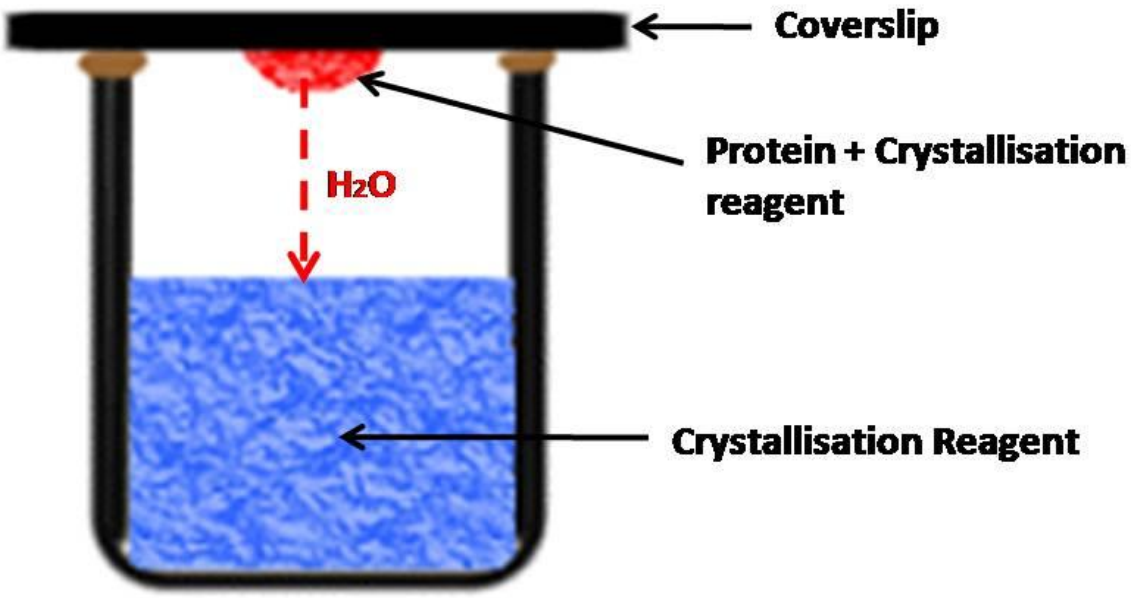


Conditions below the solubility curve are under-saturated and will be unable to support crystal growth. Saturation levels above the solubility curve can be separated into three distinct zones; metastable, nucleation and precipitation. The nucleation phase can support nucleation and crystal growth. In the metastable zone crystal growth is only maintained for existing nucleation sites. The precipitation zone is the point at which the protein forms amorphous aggregates.

Several factors, including pH, salt concentration, protein concentration and precipitant concentration can have a profound effect upon nucleation and crystal growth by reducing the solubility of a protein and causing the solution to enter a supersaturated state. Another crucial factor in determining the ability of a protein to nucleate and crystal growth to occur is the method by which the protein crystallisation takes place. Many methods are used to facilitate the crystallisation of proteins (vapour diffusion, batch methods, and equilibrium dialysis), all of which aim to bring the protein solution to a supersaturated state. However, easily the most common method is the hanging drop vapour diffusion method which is also the method employed in this study.

The vapour diffusion method involves making “sitting” or “hanging” drops that usually contain equal amounts of protein solution (0.1-3µl), equilibrated over a reservoir solution that contains the crystallisation agents (pH buffer, salt and precipitant). In this study the hanging drop method was used and therefore will be discussed here in greater detail. The two solutions (protein solution and crystallisation solution) are mixed on a coverslip (1µl+1µl drops and above) or a viewdrop 2 96-well plate seal (100nl+100nl drops). The coverslip or plate seal is then inverted so that the drop hangs downwards over a corresponding reservoir solution (Figure 2.2). Differences in precipitant concentration between the drop and reservoir cause a gradient resulting in water vapour diffusing out of the drop, such that the concentration of crystallisation reagents in the drop gradually approaches that of the reservoir solution and the protein concentration also increases. Ultimately, these changes in the drop commonly lead to supersaturation of the protein and possible crystal formation.

Figure 2.2 Diagram of the hanging drop vapour diffusion method



Equal volumes of protein and crystallisation reagent are suspended on a coverslip above a well containing crystallisation reagent. The well is sealed to form an enclosed system. Differences in precipitant concentration between the drop and the reservoir solution allows water to diffuse out of the drop by a process called vapour diffusion until equilibrium is achieved. During this process the protein concentration within the drop slowly increases leading to enhanced protein-protein interactions that can either be specific (crystal formation) or non-specific (aggregate formation). The formation of ordered protein-protein interactions and crystal nucleation can be dependent upon several factors including protein purity, crystallisation reagent conditions and protein concentration.

2.4.2 X-ray diffraction

Once optimal crystal formation has been achieved, peptide–HLA-A2 complex crystals were soaked in reservoir buffer containing increasing concentrations (5%, 10% and 15% (v/v)) of ethylene glycol or glycerol as cryoprotectants before being 'flash-cooled' at 100K in a stream of nitrogen gas (Oxford Cryosystems). X-ray data were collected to a resolution of 1.6–2.7Å on an in house MicroMax 007HF microfocus rotating anode X-ray generator (Rigaku) with a Saturn charge-coupled device detector. This results in the typical X-ray diffraction pattern, which represents the reflections of each atom within an asymmetric unit of the crystal.

2.4.3 Data processing, model building and refinement

Following X-ray diffraction, each data set was integrated, scaled and merged with programs of the XDS suite (Kabsch 1993). Peptide-MHC complex structures were determined by molecular replacement with MOLREP (Collaborative Computational Project No 4 1994) using as the search model a previously determined HLA-A2 structure with peptide residues omitted to eliminate model bias. The LILRB1/HLA-A2-RQA_Vnp complex structure was solved by molecular replacement with crystallography and nuclear magnetic resonance system (CNS) (Brunger, Adams et al. 1998) using LILRB1/HLA-A2 as the search model (PDB code; Willcox et al 2003).

In each case, the molecular-replacement calculations yielded unambiguous rotation and translation function solutions. The molecular models were refined with CNS (Brunger, Adams et al. 1998) and REFMAC5 (Murshudov, Vagin et al. 1997). The progress of refinement was verified by monitoring of the variation of the R_{free} value (Brunger 1992) calculated from an independent set of reflections (5%), which were set aside for cross-validation purposes. The

models were subjected to several rounds of alternating simulated annealing and positional refinement followed by isotropic B factor refinement. Examination of the resulting electron-density maps showed unbiased features in the electron density (full sequence of each peptide), which confirmed the validity of the molecular replacement solution. All model manipulations were performed with COOT (Emsley and Cowtan 2004). Once the R factor values were below 30%, water molecules were included in the models if they appeared in Fo – Fc (observed structure factor – calculated structure factor) maps contoured at over 3σ and were within hydrogen-bonding distance to chemically acceptable groups. These water molecules were added in successive steps and were included in the subsequent refinement cycles.

2.4.4 Structural Verification, Analysis and Figure creation

The quality of the final refined models was verified with PROCHECK (Collaborative Computational Project No 4 1994) and WHATCHECK (Hooft, Vriend et al. 1996). Structural figures were produced with Pymol (http://www.pymol.org), with molecular surfaces generated using DELPhI (Rocchia, Alexov et al. 2001; Rocchia, Sridharan et al. 2002). Structural analysis was performed using Contact (Collaborative Computational Project No 4 1994).

3 Structural plasticity of phosphopeptide binding to HLA-A2

3.1 Background

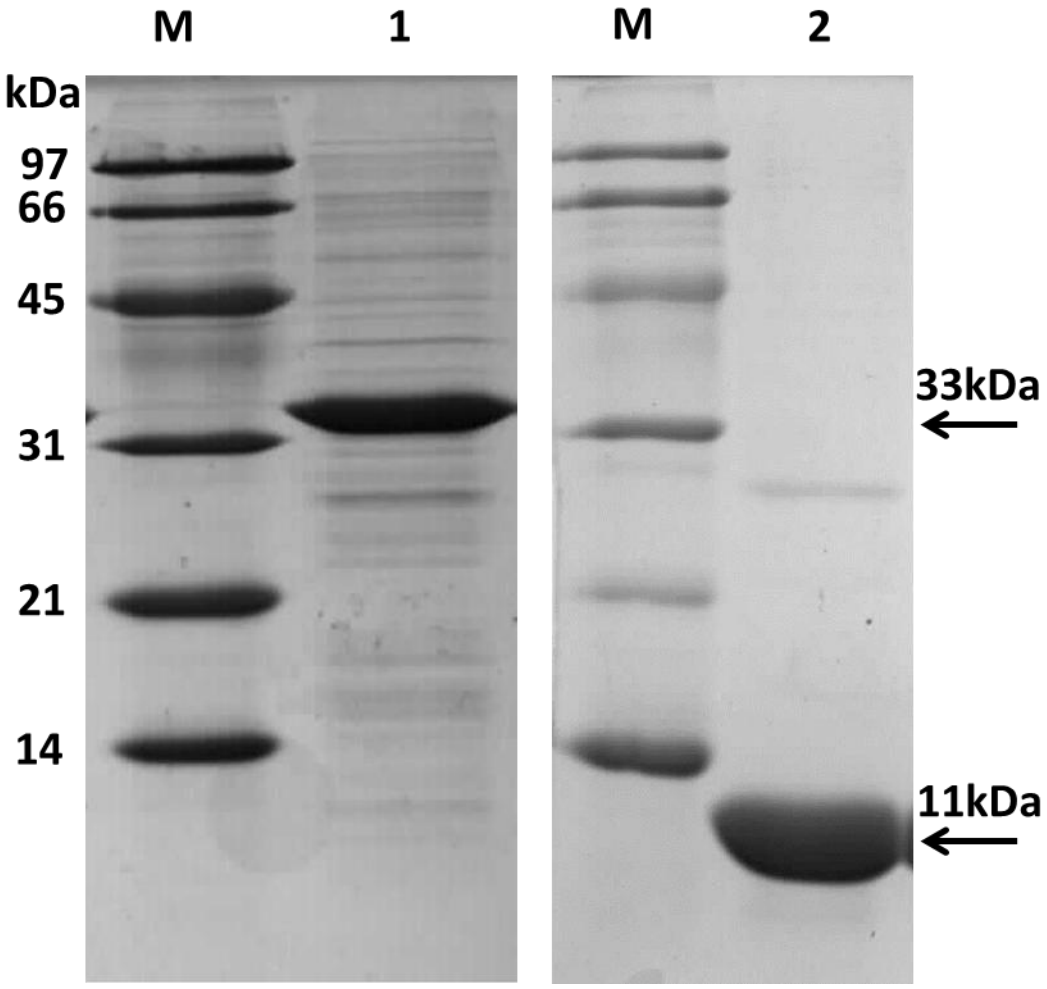
It has been previously shown that phosphorylation can alter the binding of naturally occurring HLA-A2 restricted phosphopeptide epitopes (Mohammed, Cobbold et al. 2008; Petersen, Wurzbacher et al. 2009) through the introduction of additional contacts between the phosphate moiety, the peptide and residues of the surrounding MHC. Mohammed *et al* also demonstrated that of the 37 HLA-A2-restricted phosphopeptides identified directly from cell lines by mass spectrometry (Zarling, Polefrone et al. 2006), 62% consisted of epitopes that possessed a positively charged residue (Arg or Lys) at position one (P1) and 68% a single phosphorylation at position four (P4). Collectively, these epitopes were described as possessing a canonical phosphopeptide motif. This was also commonly associated with sub-optimal anchor residues. The remaining peptides which do not possess a positively charged residue and a P4 phosphorylation were defined as having a non-canonical phosphopeptide motif. Within the canonical phosphopeptide repertoire the phosphate moiety was shown to form key stabilising interactions with the Arg residue at P1 of the peptide, Arg65 and Lys66 of the MHC and also numerous indirect contacts with surrounding residues of the peptide main chain and MHC. The study also highlighted that this binding mode seemed to be highly conserved between the four canonical HLA-A2-phosphopeptides complex structures studied (Mohammed, Cobbold et al. 2008).

These studies left several questions unresolved. Firstly, of the 37 HLA-A2 restricted phosphopeptide epitopes described by Zarling *et al* only a very small number of epitopes have been extensively studied to date and it is therefore unclear how representative these structural features are of the entire group. Secondly, the degree of structural plasticity exhibited by the phosphate moiety between different epitopes and the exact impact this would have on the binding energetics, which vary from quite extreme phosphate dependent binding to no dependence on phosphate, is undefined. In this chapter these questions were specifically addressed by determining the binding affinities of a range of canonical phosphopeptide/non-phosphopeptide epitope pairs using established peptide-MHC binding assays together with structural studies of each phosphopeptide/HLA-A2 complex with X-ray crystallography. Conceivably these issues could have major effects upon antigen presentation under different metabolic conditions and on T-cell recognition.

3.2 HLA-A2-phosphopeptide complex production

In order to generate soluble HLA-A2-phosphopeptide complexes for use in structural studies two separate pET23a vectors incorporating either the HLA-A2 heavy chain or the β_2M -domain were generated prior to this study (Mohammed, Cobbold et al. 2008). Full expression was performed in *E. coli* BL21 (DE3) cells and inclusion bodies purified then solubilised in 8M urea (Figure 3.1). The purity of HLA-A2 heavy chain and β_2M were deemed sufficient for further *in vitro* dilution refolding of HLA-A2-phosphoeptide complexes. For this study canonical phosphopeptide epitopes were selected from the Zarling *et al* list of HLA-A2 presented phosphopeptides (Zarling, Polefrone et al. 2006) whose HLA-A2-peptide complex structure are yet to be determined. HLA-A2-phosphopeptide complexes were

Figure 3.1 SDS-PAGE analysis of purified HLA-A2 and β_2M inclusion bodies



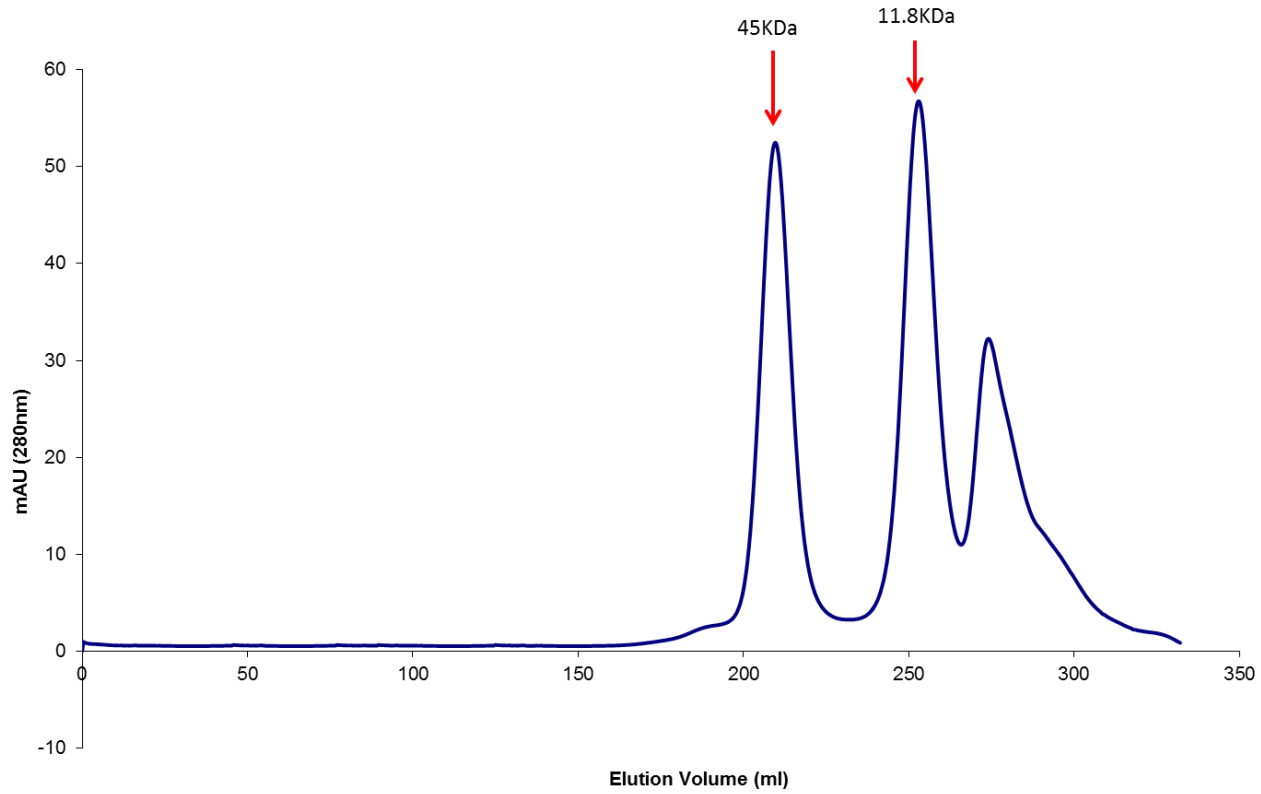
Purified HLA-A2 and β_2M- inclusion bodies were solubilised in 8M Urea and 1µl analysed by SDS-PAGE. Molecular weight markers (M) were analysed beside HLA-A2 (Lane 1) and β_2M- (Lane 2) to assess protein size and purity. Yields of HLA-A2 and β_2M typically reached 200mg/L.

generated by standard *in vitro* dilution refolding conditions (Garboczi, Hung et al. 1992) and subsequently purified by size exclusion chromatography. The yield and purity of each purified HLA-A2-phosphopeptide complex was sufficient for crystallisation screening trials as demonstrated by a large single peak eluting at 215ml (corresponding to a molecular weight of 45kDa) on the S200 size exclusion elution profile (Figure 3.2) and by subsequent analysis with SDS-PAGE (Figure 3.3).

3.3 Crystallisation of HLA-A2-phosphopeptide complexes

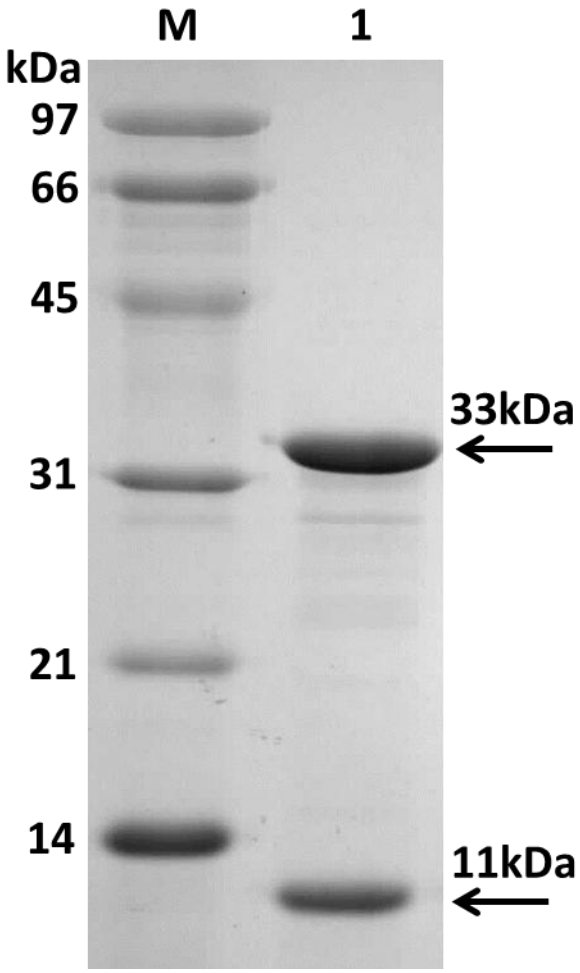
Purified HLA-A2-phosphopeptide complexes were concentrated to 9-15mg/ml before use in various sparse matrix based commercial crystallisation screens. Each screen contained 96 unique conditions and was set up on a small scale (100nl+100nl drops) via the hanging drop vapour diffusion method, using a Mosquito nanolitre crystallisation robot (Mohammed, Cobbold et al. 2008). Each well, containing 100µl of crystallisation reagent reservoir, was then sealed and allowed to equilibrate at 23°C. After 3 days, crystallisation trials were inspected manually using a Leica MZI6 microscope in order to identify conditions that yielded microcrystals. The success of crystallisation for each HLA-A2-phos complex varied (Table 3.1). Interestingly, the most preliminary hits were identified for the HLA-A2-RQIpS complex with 48/96 (50%) conditions tested yielding crystals. In contrast, HLA-A2-RLQpS and HLA-A2-RLSpS both yielded far fewer conditions. The conditions that yielded the most promising crystals, in terms of morphology and size, were further optimised on a large scale to grow diffraction-grade crystals (Table 3.1 and Figure 3.4).

Figure 3.2 Size exclusion chromatography purification of a typical HLA-A2-Phosphopeptide complex



Refolded HLA-A2-phosphopeptide complexes were purified by S200 size exclusion chromatography. The molecular weight of HLA-A2-phosphoeptide complex is ~45kDa and thus elutes from the column at around 215ml. Free β_2M, which has a molecular weight of ~11.8kDa, elutes from the column at around 255ml. The third peak at around 270ml consists of low molecular weight refold buffer components.

Figure 3.3 SDS-PAGE analysis of refolded HLA-A2-phosphopeptide complex following purification by size exclusion chromatography



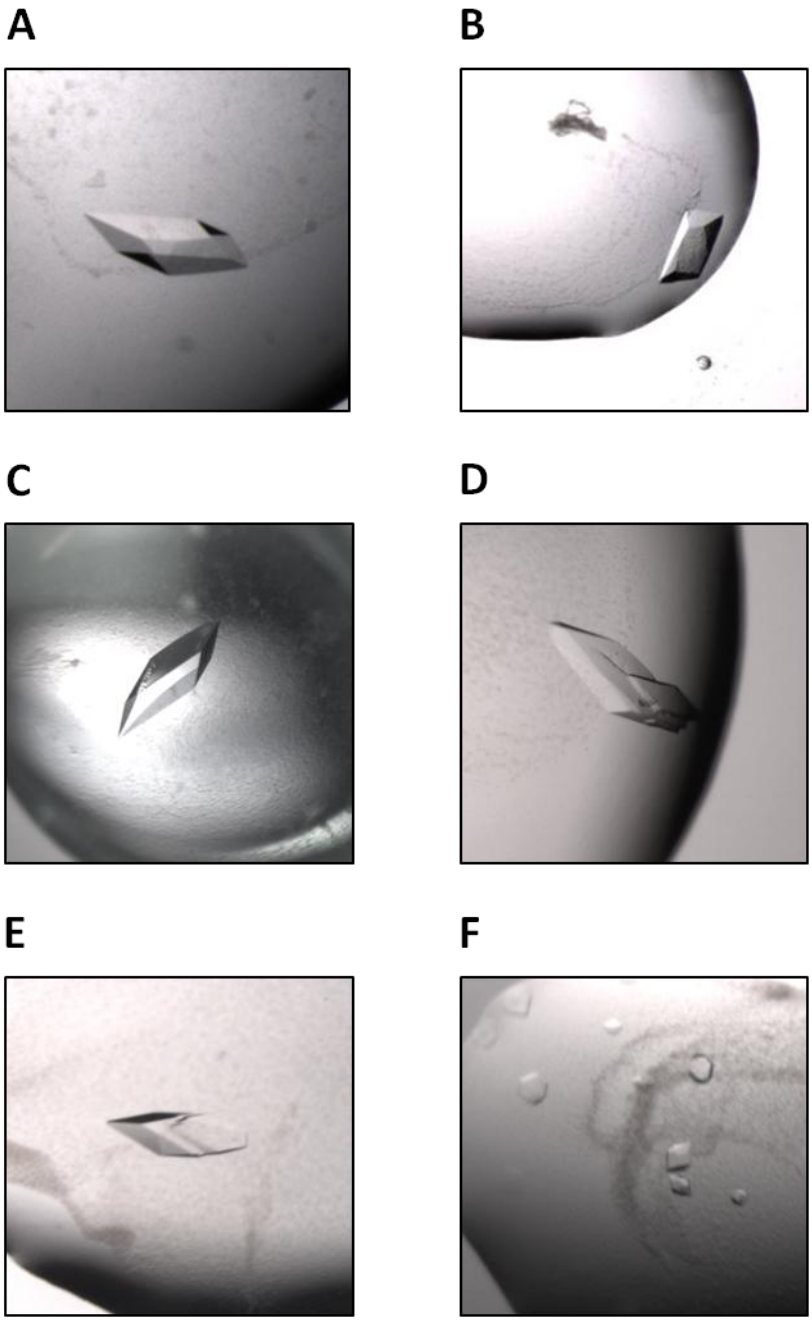
The purity of renatured and gel filtrated HLA-A2-phosphopeptide complex was assessed by SDS-PAGE. The analysis shown above is a representative of all HLA-A2-phosphopeptide complexes. Molecular Weight markers (M) were analysed alongside HLA-A2-phosphopeptide complexes (Lane 1). The HLA-A2 heavy chain and β_2M are indicated by the 33kDa and 11kDa arrows respectively.

Table 3.1 Crystallisation conditions for HLA-A2-phosphopeptide complexes

Epitope Sequence	Source Protein	Protein Concentration (mg/ml)	Total Number of 'Hits'	Final Optimised Condition
RLQpSTSERL	Mitochondrial escape 1-like 1	11	1/96	0.2M Sodium Formate, 20% (w/v) PEG3350
RLSpSPLHFV	ORF17, chromosome 2	11	1/96	0.2M Sodium Formate, 20% (w/v) PEG3350
RQIpSQDVKL	Adenosine monophosphate deaminase 2	11	48/96	8% (v/v) Tacsimate pH 8, 20% (w/v) PEG3350
RTFpSPTYGL	β-synemin/Desmuslin	7.2	4/576	0.1M BisTris Propane pH 7.5, 0.02M Sodium/Potassium Phosphate, 20% (w/v) PEG 3350
RTLpSHISEA	FLJ13725	12.2	6/288	0.1M Bicine pH 9, 20% (w/v) PEG 6000

The number of conditions which produced initial crystals for each complex is documented along with the number of separate conditions tested. The final condition used to grow crystals for X-ray diffraction experiments is also shown for each complex as well as the protein concentration at which this was achieved.

Figure 3.4 HLA-A2-phosphopeptide complex crystals used for X-ray diffraction



Crystals of HLA-A2-RTLpS (A), HLA-A2-RTFpS (B), HLA-A2-RQIpS (C), HLA-A2-RLSpS (D), HLA-A2-RLQpS (E) and HLA-A2-PKD2p (F) were grown to sufficient size and quality for X-ray diffraction experiments using the hanging drop vapour diffusion technique. Crystals were typically produced in 1µl+1µl hanging drops and appeared after 48 hours at 23°C. The size of crystals generated typically reached 100µmx100µm.

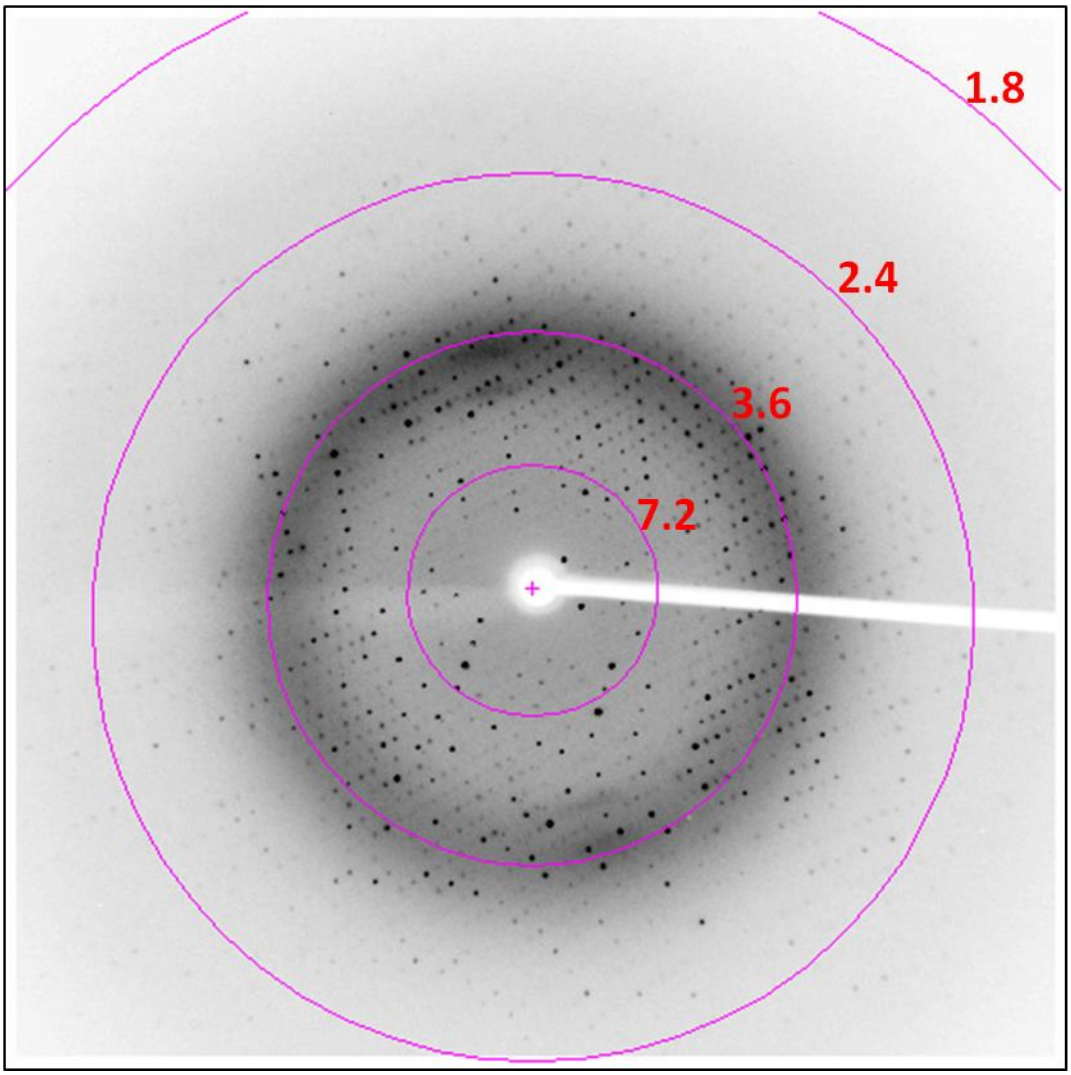
3.4 Data collection and processing

Prior to data collection HLA-A2-phosphopeptide complex crystals were soaked in reservoir solution supplemented with increasing concentrations of ethylene glycol (5, 10, 15 and 20%) prior to flash cooling in liquid nitrogen (Mohammed, Cobbold et al. 2008). Data collection was performed under the supervision of Dr Fiyaz Mohammed using the 'in house' Rigaku Micro-Max 007HF rotating anode X-ray generator with a Saturn CCD detector (University of Birmingham Macromolecular X-ray Facility). X-ray data were collected (Figure 3.5) to 1.7Å for HLA-A2-RQIpS, 1.8Å for HLA-A2-RTFpS, 1.9Å for HLA-A2-RTLpS, 1.8Å for HLA-A2-RLSpS and 2.1Å for HLA-A2-RLQpS under a constant nitrogen gas steam at 100K. The crystals typically belonged to the orthorhombic space group C2. Diffraction data sets were integrated, scaled and merged using programs of the XDS suite (Kabsch 1993). The relevant unit cell parameters and data processing statistics are listed in Table 3.2. All phosphopeptide crystals provided high resolution data which was easily processed, as demonstrated by the favourable R_{merge} values for each complex, thus allowing each complex to be taken forward for further structure solution.

3.5 HLA-A2-phosphopeptide complex structure solution

Initial phase information for each HLA-A2-phosphopeptide complex was determined by molecular replacement using MOLREP (Murshudov, Vagin et al. 1997) a component of the CCP4 suite (CCP4 1994). The phasing model consisted of an HLA-A2 complex (PDB ID code 1HHG) (Madden, Garboczi et al. 1993), but with the antigenic peptide residues and water molecules omitted so as to eliminate model bias. In each case, the molecular-replacement calculations yielded unambiguous rotation (rotation angles θ_1, θ_2 and θ_3) and translation

Figure 3.5 Typical example of HLA-A2-phosphopeptide X-ray diffraction pattern



A 0.5° oscillation image of the HLA-A2-RTLpS crystal complex, which diffracted X-rays to 1.9Å, was collected using an in house Micro-Max 007HF rotating anode X-ray generator (Rigaku) using a Saturn CCD detector (University of Birmingham Macromolecular X-ray Facility). The pink rings represent resolution shells (Å). This image was typical of diffraction data collected for each HLA-A2-phosphopeptide complex crystal.

Table 3.2 Data processing statistics for HLA-A2-phosphopeptide complexes

	HLA-A2-RLQpS	HLA-A2-RLSpS	HLA-A2-RQIpS	HLA-A2-RTFpS	HLA-A2-RTLpS
Space Group	C2	C2	C2	C2	C2
Unit Cell Length (Å)	a=110.8,b=54.1,c=75.6	a=118.7,b=53.5,c=74.9	a=118.0,b=54.2,c=75.1	a=120.1,b=54.3,c=74.8	a=117.4,b=54.8,c=75.4
Unit Cell Angles (°)	α=90, β=120.4, γ=90	α=90, β=104.8, γ=90	α=90, β=104.9, γ=90	α=90, β=104.6, γ=90	α=90, β=104.5, γ=90
Resolution (Å)	20-2.1	20-1.8	20-1.7	20-1.8	20-1.9
Observed Reflections	606386 (184123)	687316 (43473)	334999 (48808)	424744 (35684)	327869 (27616)
Unique Reflections	25449 (8229)	40570 (5048)	46573 (10750)	43093 (6172)	36088 (4747)
Multiplicity	23.8(22.4)	16.9 (8.6)	7.2 (4.5)	9.9 (5.8)	9.1 (5.8)
Completeness (%)	98.9 (97.8)	95.8 (80.2)	91.9 (75.4)	99.1 (95.7)	98.1 (92.2)
R_{merge} (%)	11.6 (42.2)	5.2 (68.7)	3.4 (32.8)	5.7 (76.9)	4.2 (45.2)
I/σ(I)	27.4 (9.3)	40.76 (3.06)	34.0 (4.7)	26.5 (2.4)	34.6 (4.3)

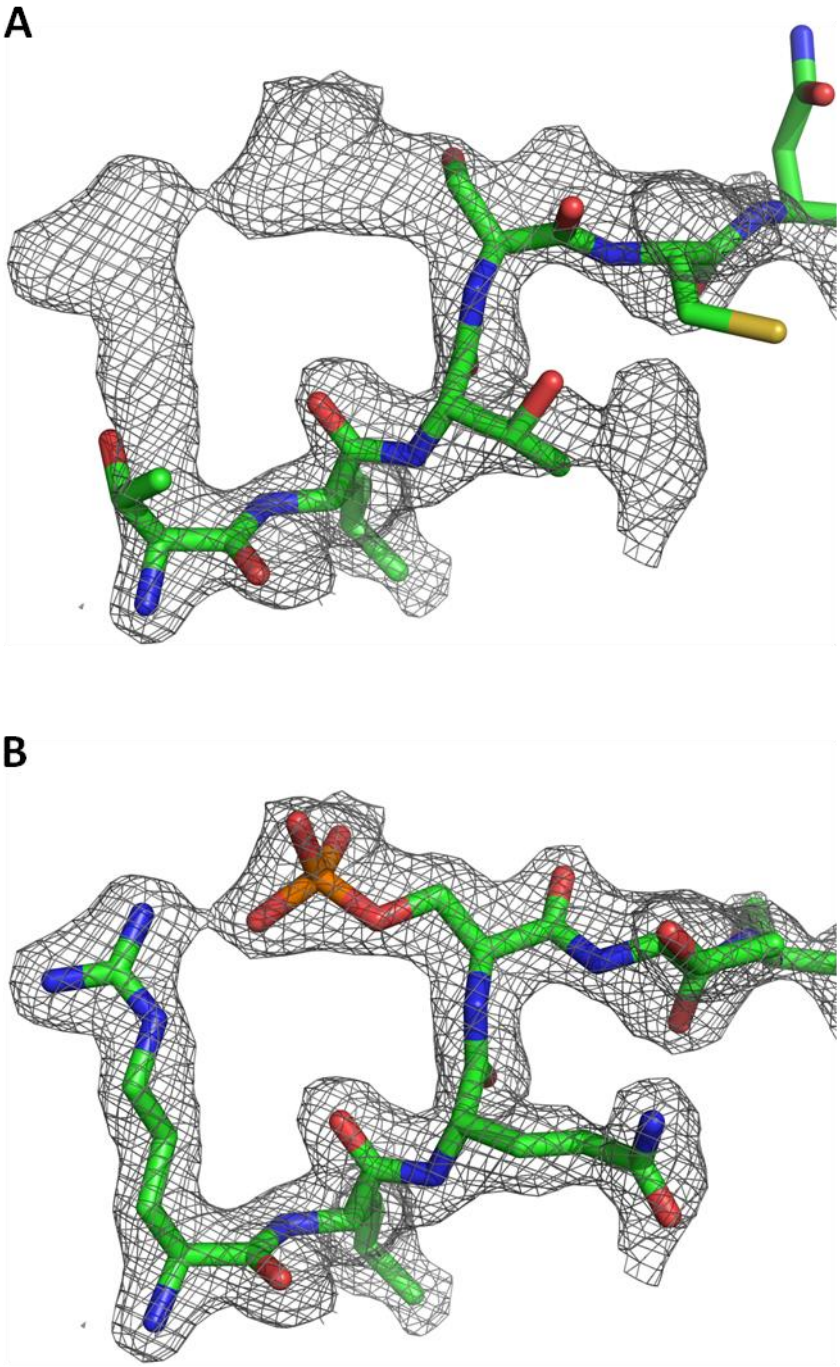
Data processing was performed for each HLA-A2-phosphopeptide complex using programs of the XDS suite. The values in parentheses represent data processing statistics in the highest resolution shell.

(translation vectors x, y and z) function solutions as demonstrated by peak RF-function and monitor scores respectively (Values for each complex are documented in Appendix A-E). The molecular replacement solutions for each HLA-A2-phosphopeptide complex were easily generated due to the fact that the phasing model used was highly similar (100% identity, excluding the peptide epitope) to the HLA-A2-phoshopeptide complexes.

3.6 HLA-A2-phosphopeptide structure refinement

The molecular models were refined with CNS (Brunger 1992) and REFMAC5 (Murshudov, Vagin et al. 1997). The progress of refinement was verified by monitoring of the variation of the R_{free} value (Brunger 1992) calculated from an independent set of reflections, which were set aside for cross-validation purposes. The models were subjected to several rounds of alternating simulated annealing and positional refinement followed by isotropic B factor refinement. Examination of the resulting electron-density maps showed unbiased features in the electron density (full sequence of each peptide), which confirmed the validity of the molecular replacement solution (Figure 3.6). All model manipulations were performed with COOT (Emsley and Cowtan 2004). Once the R factor values dropped below 30%, water molecules were included in the models if they appeared in *Fo–Fc* maps (observed structure factor – calculated structure factor) contoured at over 3σ and were within hydrogen-bonding distance to chemically acceptable groups. These water molecules were added in successive steps and were included in the subsequent refinement cycles. The final refinement statistics for each HLA-A2-phosphopeptide complex structure are summarised in Table 3.3. Due to the crystals diffracting to a high resolution and low R_{merge} values following data processing, the resulting *Fo-Fc* and 2*Fo-Fc* maps yielded well defined electron density thus aiding in the

Figure 3.6 Unbiased features of electron density maps



The electron density map (grey mesh contoured to 1σ) of HLA-A2-RLQpS is shown following molecular replacement and an initial round of refinement. (A) Sticks represent the peptide from the phase model 1HHG (Madden, Garboczi et al. 1993) used in the molecular replacement calculations. The 1HHG derived peptide does not fit into the electron density map, indicating that there is no model bias in the calculated electron density maps. (B) The RLQpS peptide fits snugly into the electron density further emphasising the lack of model bias.

Table 3.3 Refinement Statistics for HLA-A2-phosphopeptide complexes

	HLA-A2-RLQpS	HLA-A2-RLSpS	HLA-A2-RQIpS	HLA-A2-RTFpS	HLA-A2-RTLpS
Resolution (Å)	20-2.1	20-1.8	20-1.7	20-1.8	20-1.9
Ref. in working set	24223	38539	44256	40913	34292
Ref. in test set	1226	2031	2317	2159	1792
R_{Cryst} (%)	19.05	19.53	21.70	21.43	22.13
R_{Free} (%)	23.00	22.12	23.52	23.34	24.25
Num. of protein atoms	3155	3179	3156	3145	3103
Average B-Factor	Chain A – 20.14 Chain B – 23.11	Chain A – 20.34 Chain B – 21.25	Chain A – 20.74 Chain B – 20.50	Chain A – 20.19 Chain B – 22.97	Chain A – 25.71 Chain B – 23.64
Number of water mol.	343	440	395	303	224
RMS Bond Length (Å)	0.014	0.015	0.007	0.006	0.038
RMS Bond Angle (°)	1.384	1.533	0.987	0.985	2.590
Ramachandran Plot					
Allowed (%)	99.7	99.7	99.7	99.4	99.4
Generously Allowed (%)	0.3	0.3	0.3	0.3	0.3
Disallowed (%)	0	0	0	0.3	0.3

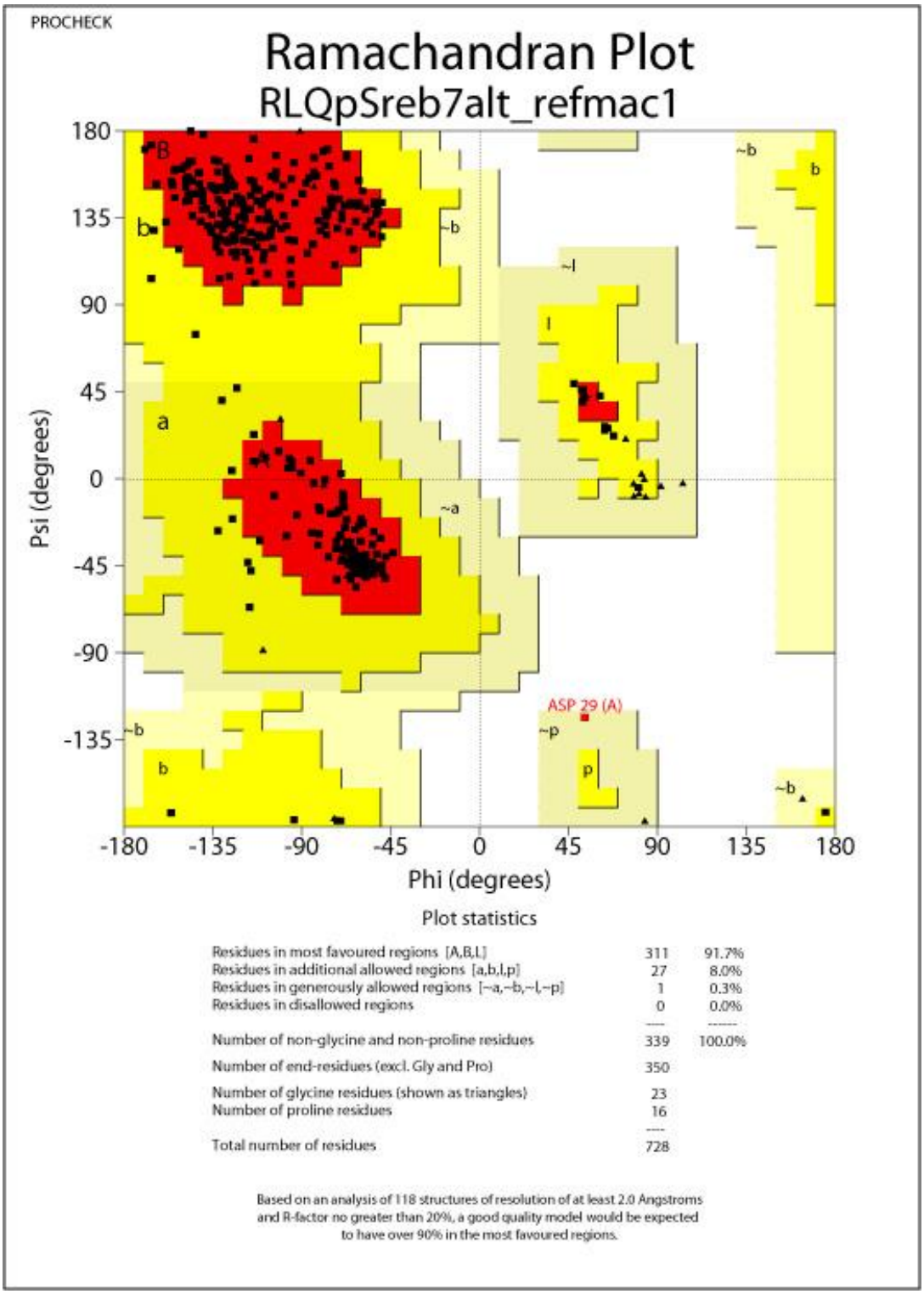
Final refinement statistics for each HLA-A2-phosphopeptide complex were generated following several rounds of alternating simulated annealing and positional refinement followed by isotropic B factor refinement and model building.

process of model building and refinement of each of the complexes. The quality of the final refined models was verified with PROCHECK (CCP4 1994) and WHATCHECK (Hooft, Vriend et al. 1996), which demonstrated in most cases that non-glycine residues were absent from the disallowed regions of the Ramachandran plot, although one residue (Ser195) was present in the disallowed region in HLA-A2-RTLpS and HLA-A2-RTFpS complexes, the residue was in well-defined electron density and was situated within a solvent exposed loop region therefore this residue was deemed to be in the correct position for these two structures (Figure 3.7). Most residues were well defined in all structures, except for a few solvent-exposed side chains indicating the flexibility of these residues. Hydrogen bonding, hydrophobic and van der Waals interactions were analyzed with CONTACT (CCP4 1994). The structural analysis of each HLA-A2-phosphopeptide complex was performed with COOT and the CCP4 software suite (CCP4 1994). All structural figures were produced using PYMOL (Schrodinger 2010).

3.7 Overall structure of HLA-A2-phosphopeptide complexes

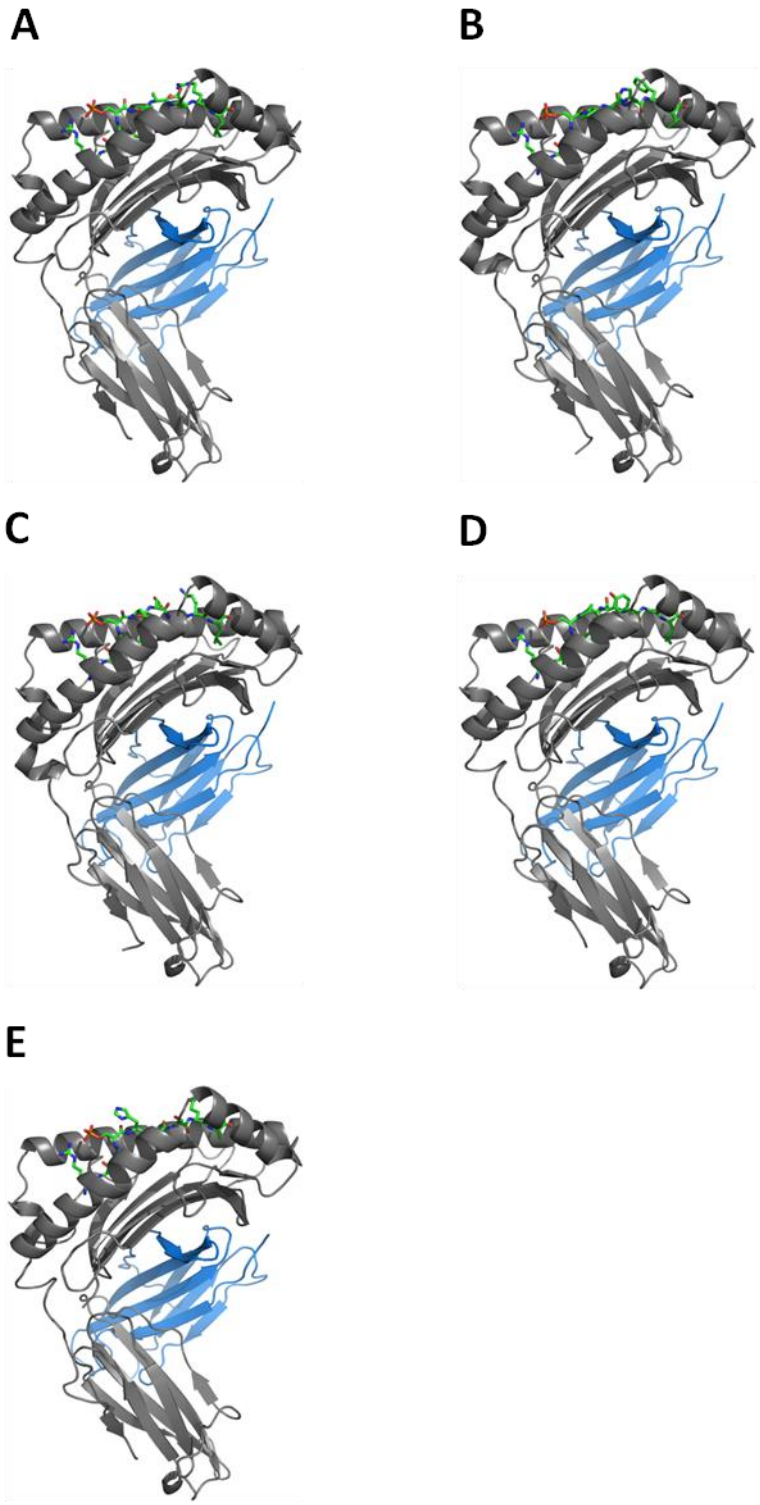
The overall structure of each HLA-A2-phosphopeptide complex is similar, both to each other and to previously solved structures of non-phosphorylated HLA-A2 complexes, with a typical overall RMSD of 0.8 (Figure 3.8 A-E). As expected, each complex is comprised of an α1-α2 peptide binding platform and an α3 membrane-proximal domain non-covalently associated with the β_2M-light chain and the bound peptide.

Figure 3.7 Ramachandran plot for HLA-A2-RLQpS complex



A Ramachandran plot for HLA-A2-RLQpS complex which demonstrates that non-glycine residues were absent from the disallowed regions indicating that the model is of high quality.

Figure 3.8 Overall structures of HLA-A2-phosphopeptide complexes



The overall complex structure for HLA-A2-RLQpS (A), HLA-A2-RLSpS (B), HLA-A2-RQIpS (C), HLA-A2-RTFpS (D) and HLA-A2-RTLpS (E) are shown in ribbon representation. The A2 heavy chain is depicted in grey whereas β2M is shown in light blue. Each peptide is shown in stick representation, with carbon (green), oxygen (red), nitrogen (blue) and phosphate (orange).

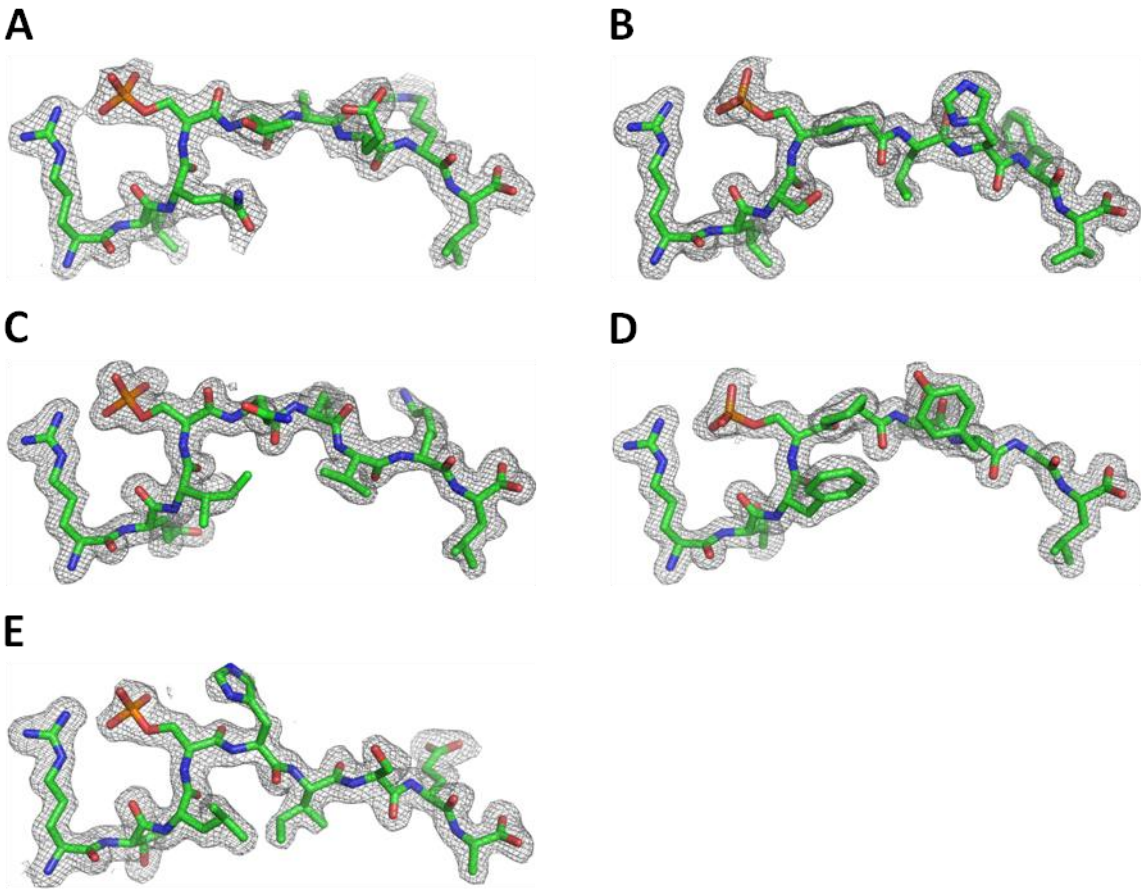
Analysis of each phosphopeptide epitope revealed unambiguous electron density for each epitope (Figure 3.9). The phosphate in each complex is orientated in a broadly similar way to other canonical phosphopeptides in that the phosphate is upward facing, solvent exposed and is orientated towards the P1 side chain position and is therefore in an extremely prominent position potentially available for recognition by TCRs (Figures 3.8, 3.9 and 3.10).

The basic mode of peptide binding for each epitope is also similar to that of non-phosphorylated peptides and previously determined phosphopeptides in that each peptide adopts an extended structure in the antigen binding groove. Each peptide maintains stabilising interactions between their N and C terminus and HLA-A2 residues as for non-phosphorylated peptides. In addition to this, the anchor residues at position 2 (P2) and the C-terminus (PC) are orientated in a broadly similar orientation to previous HLA-A2-phosphopeptide structures (Mohammed, Cobbold et al. 2008), in that they point downward into the B and F pockets, respectively (Figure 3.10).

3.8 Analysis of canonical phosphopeptide binding mode

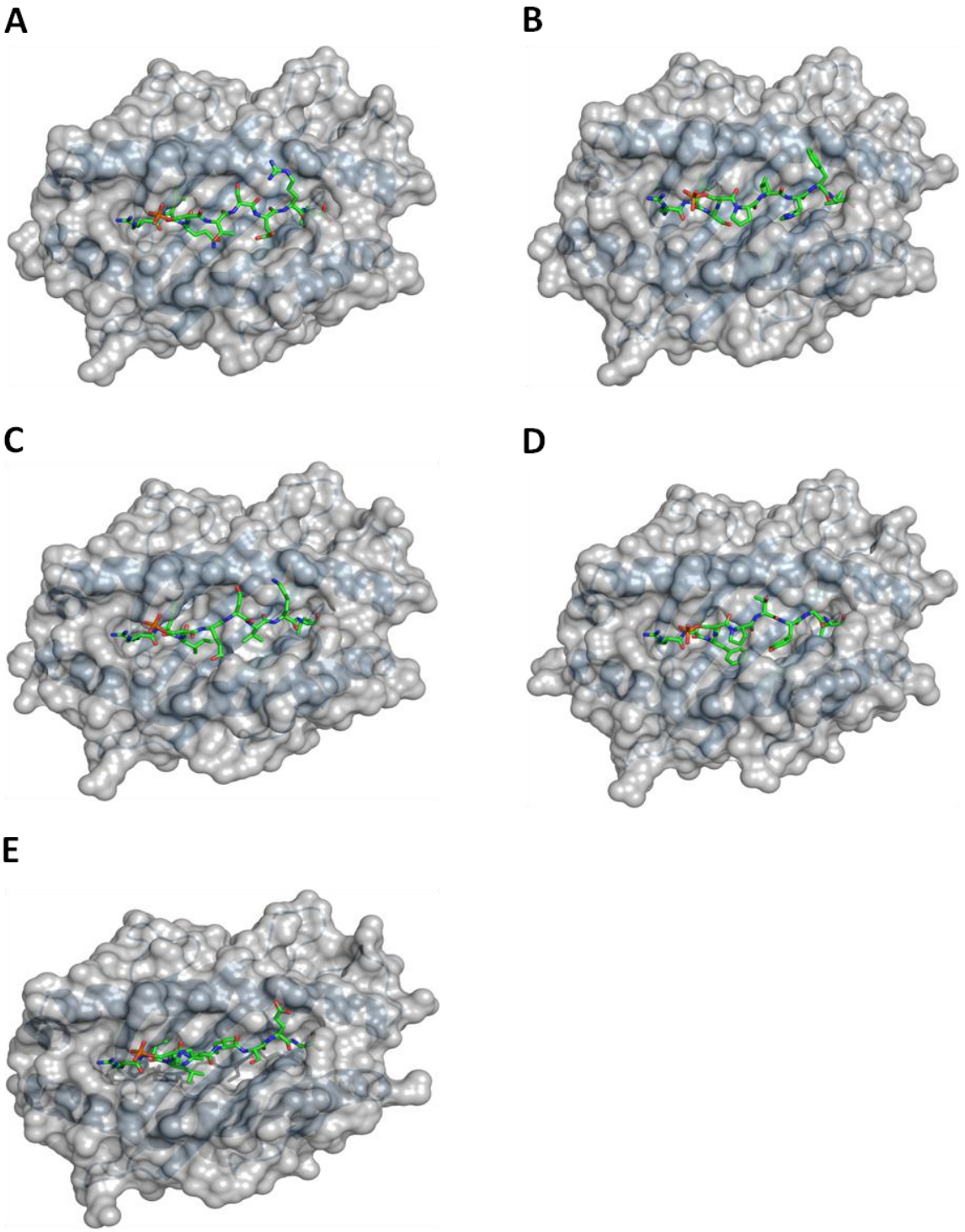
Having shown that the overall structure of the HLA-A2-phosphopeptide complexes in this study matched previously determined phosphorylated and non-phosphorylated epitopes bound to HLA-A2, the role of the phosphate group in binding to MHC was examined for each epitope. Previous results have demonstrated that for each epitope the phosphate moiety adopts a conserved orientation and mediates intra-molecular peptide contacts to the P1 Arg residue as well as hydrogen-bonding interactions with the α1 helical residues Arg65 and Lys66 (Mohammed, Cobbold et al. 2008).

Figure 3.9 Individual phosphopeptide epitope structures



Individual phosphopeptide epitope structures of RLQpS (A), RLSpS (B), RQIpS (C), RTFpS (D) and RTLpS (E). Each epitope is superimposed onto a 2Fo-Fc electron density map contoured to 1σ (grey wire).

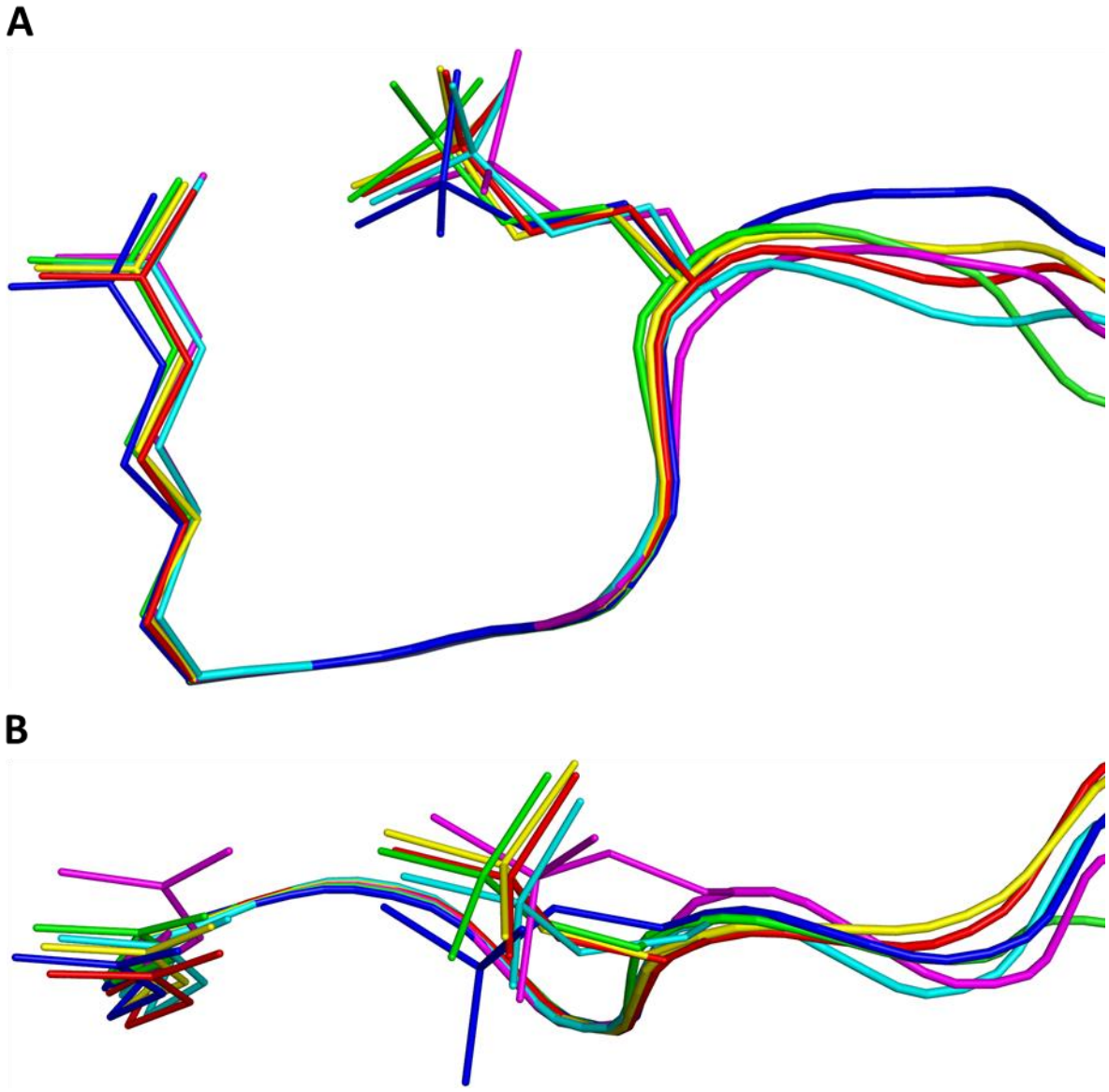
Figure 3.10 Orthogonal view of peptide binding groove



View of α1α2 peptide binding platform of; HLA-A2-RLQpS (A), HLA-A2-RLSpS (B), HLA-A2-RQIpS (C), HLA-A2-RTFpS (D) and HLA-A2-RTLpS (E). Molecular surfaces of α1α2 peptide binding platforms (grey) overlay ribbon representation of the α1α2 regions (dark blue). Phosphopeptides are shown in stick representation

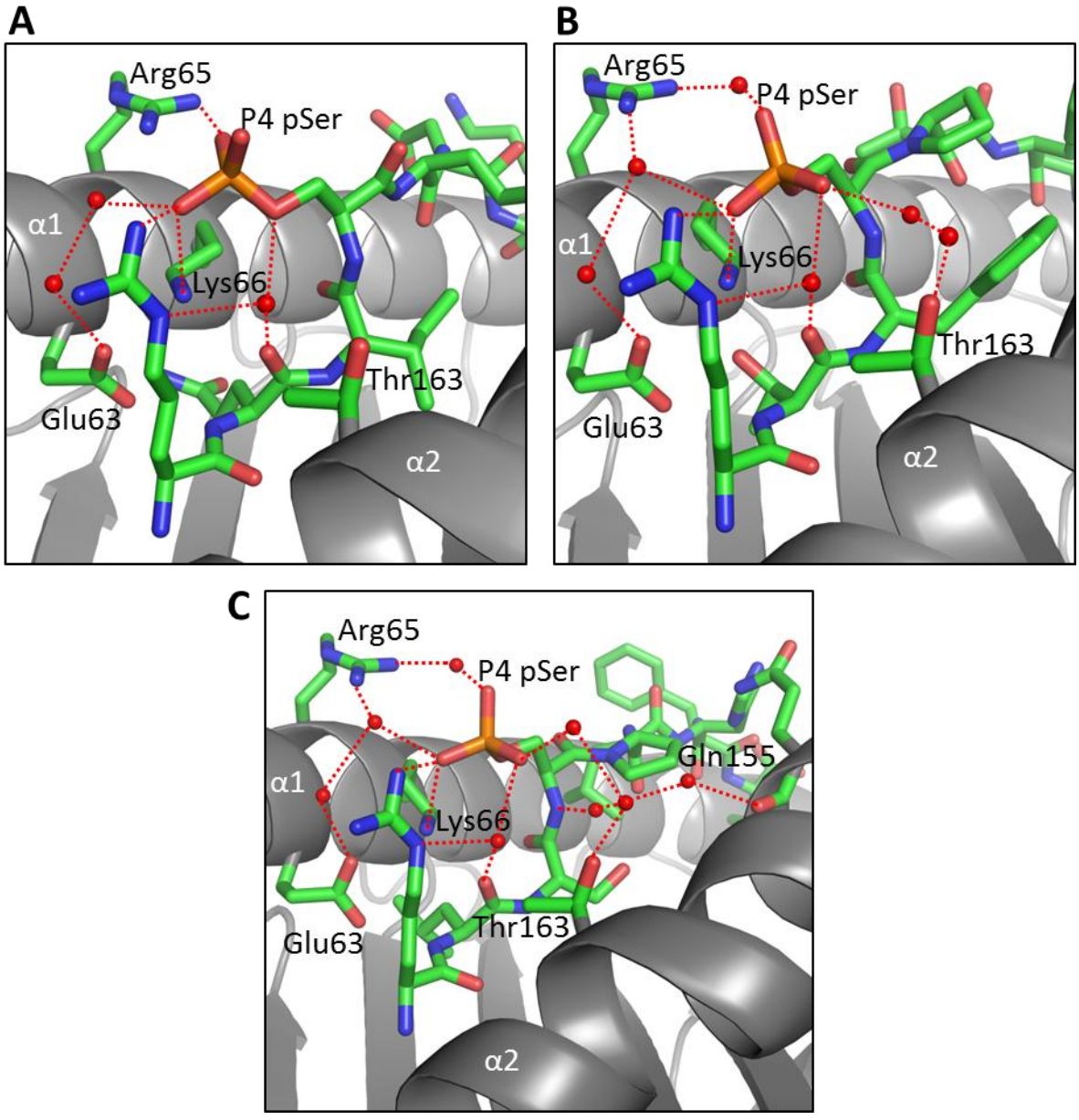
However, due to this study focusing on only a small number of canonical phosphopeptides, it was still unclear whether this is representative of the whole group or to what extent there is structural plasticity within the binding mode of the canonical phosphopeptide repertoire. Upon expanding the number of phosphopeptide epitope structures determined it was revealed that while some epitopes fall into the category previously described by Mohammed *et al,* others showed differences in this phosphate mediated binding implying there is a degree of plasticity in the canonical phosphopeptide repertoire as a whole (Figure 3.11). The RQIpS epitope is representative of an epitope in which the phosphate adopts a highly similar binding mode to those described by Mohammed *et al* (Figure 3.12 A). However, the most striking plasticity in phosphate moiety binding was observed for RTFpS in which the phosphate group is orientated away from Arg65 and positioned more towards the α2 helix of the MHC peptide binding groove compared to other phosphopeptide epitopes, resulting in the loss of a direct hydrogen-bonding interaction with Arg65 (Figure 3.12 B). Despite this re-orientation away from Arg65, the phosphate moiety mediates compensatory interactions with the hydroxyl group of Thr163, located on the α2 helix, through a network of water mediated hydrogen bonds. A second epitope, RLSpS, also demonstrated plasticity features with regards to the phosphate binding mode, albeit not as pronounced as RTFpS. In this case the reorientation of the phosphate moiety resulted in the loss of a direct interaction with Arg65 and like RTFpS, an indirect interaction with Arg65 was observed through water mediated hydrogen bonds (Figure 3.12 C). Also similarly to RTFpS, the subtle reorientation of the phosphate moiety of RLSpS allowed compensatory interactions with the hydroxyl moiety of Thr163 via a network of water mediated hydrogen bonds (Figure 3.12 C).

Figure 3.11 Plasticity of the canonical phosphopeptide binding mode



Phosphopeptide structures were superimposed onto one another to determine any difference in the phosphate binding mode. The side (A) and orthogonal (B) representations of RLQpS (cyan), RLSpS (pink), RQIpS (yellow), RTFpS (blue), RTLpS (green) and PKD2 (red) are shown. PKD2 represents an epitope which displays the same structural conformation as other epitopes previously described (Mohammed, Cobbold et al. 2008).

Figure 3.12 Phosphate mediated contacts

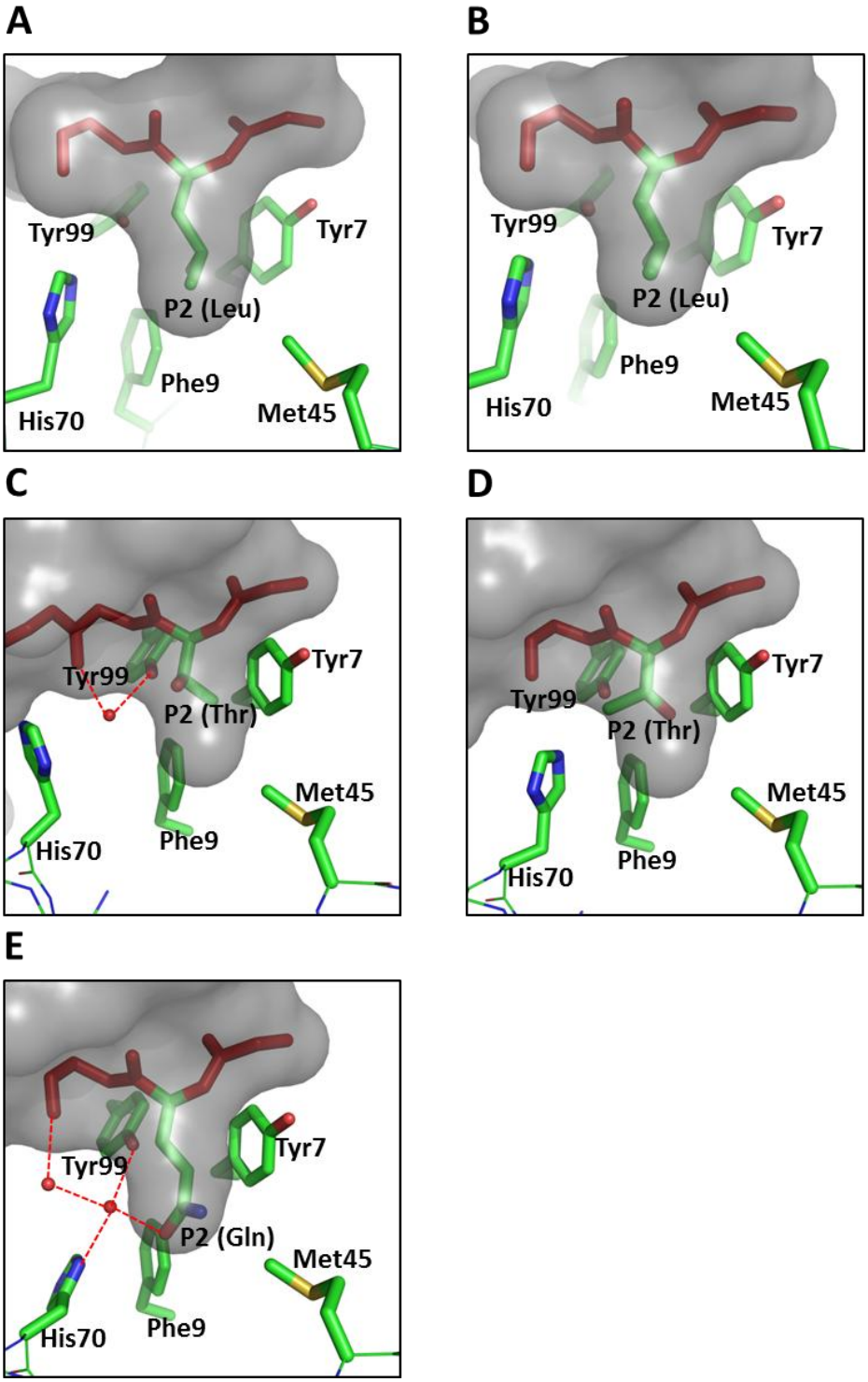


Phosphate mediated contacts of HLA-A2-RQIpS (A), HLA-A2-RTFpS (B) and HLA-A2-RLSpS (C). The α1 and α2 helices (grey) along with water molecules (red spheres) and interacting side chains of the MHC are shown. Red dash lines indicate hydrogen bond interactions.

A key observation of the HLA-A2 phosphopeptide antigen repertoire is that the presence of subdominant anchor residues is overrepresented within the group and this supports the hypothesis that contacts between the phosphate and MHC are able to compensate for suboptimal interactions of P2 anchor residues with surrounding residues of the B and F binding pockets of the MHC. Previous work by Mohammed *et al* highlighted the binding interactions of four subdominant anchor residue containing phosphopeptides to HLA-A2; namely KMDpS, RTYpS, RQApS_M and RQApS (PKD2) (Mohammed, Cobbold et al. 2008). This study showed that in the presence of phosphate mediated contacts, the toleration of subdominant anchor residues leads to structurally sub-optimal interactions with the MHC, including fewer or unusual binding interactions which may result in compensatory conformational rearrangement of B-pocket residues. The interactions between each P2 anchor residue and amino acids that line the B pocket of MHC was determined for RLQpS, RLSpS, RTFpS, RTLpS and RQIpS to evaluate how this correlates with peptide-MHC binding affinity. The dominant P2 anchor leucine of RLQpS and RLSpS fills the B-pocket optimally, thus facilitating numerous hydrophobic and van der Waals interactions with surrounding residues Phe 9, Met 45 and Val 67 (Figure 3.13 A-B and Table 3.4) similar to those described by previous structures (Madden, Garboczi et al. 1993; Madden 1995; Mohammed, Cobbold et al. 2008). In contrast, the threonine of RTFpS and RTLpS although orientated in a similar direction to P2 leucine forms considerably fewer interactions with the residues Phe 9, Met 45 and Val 67 of the B-pocket due to its smaller side chain (Figure 3.13 C-D and Table 3.4).

Finally, the P2 glutamine of RQIpS adopted a similar binding mode to PKD2 and RQA_M described by Mohammed *et al* (Mohammed, Cobbold et al. 2008) in that the accommodation of the polar glutamine residue in the mainly hydrophobic B-pocket is

Figure 3.13 The accommodation of anchor residues in HLA-A2 phosphopeptides



The accommodation of the P2 anchor residue in the B-pocket of the peptide binding groove is depicted for HLA-A2-RLQpS (A), HLA-A2-RLSpS (B), HLA-A2-RTFpS (C), HLA-A2-RTLpS (D) and HLA-A2-RQIpS (E). The main chain of each peptide is represented by red cylinder, with water molecules shown as red spheres. The surrounding MHC residues are represented in stick format.

Table 3.4 Hydrophobic and van der Waals contacts between peptide P2 anchor residues and HLA-A2

	HLA-A2-RLQpS	HLA-A2-RLSpS	HLA-A2-RQIpS	HLA-A2-RTFpS	HLA-A2-RTLpS
P2 Anchor	Leu	Leu	Gln	Thr	Thr
	C^{α} (7)	C^{α} (8)	C^{α} (7)	C^{α} (7)	C^{α} (7)
	C^{β} (8)	C^{β} (7)	C^{β} (7)	C^{β} (4)	C^{β} (6)
	C^{γ} (7)	C^{γ} (8)	C^{γ} (7)	$O^{\gamma 1}$ (7)	$O^{\gamma 1}$ (8)
	$C^{\delta 1}$ (12)	$C^{\delta 1}$ (12)	C^{δ} (9)	$C^{\gamma 2}$ (10)	$C^{\gamma 2}$ (9)
	$C^{\delta 2}$ (12)	$C^{\delta 2}$ (12)	$O^{\varepsilon 1}$ (12)		
			$N^{\varepsilon 2}$ (13)		
Total Contacts	46	47	55	28	30

The number of hydrophobic and van der Waals contacts between the individual atoms of the P2 side chain and the surrounding residues of HLA-A2 are shown in parenthesis.

compensated by an unusual set of direct and water mediated interactions with surrounding residues (Figure 3.13 E). Moreover, this is accompanied by the re-orientation of His70 (a residue which has a highly conserved position in other HLA-A2 structures) due to the introduction of ordered water molecules. This conformational adjustment of His70 results in a dramatic reshaping of the B-pocket.

3.9 Phosphate moiety effects on epitope binding affinity

Having established there is considerable structural plasticity in the binding mode of canonical phosphopeptides, involving altered phosphate mediated contacts to the surrounding residues of the MHC, an interesting question that arose is what implications do these altered phosphate binding modes have upon antigen binding energetics. Previous work has highlighted that the phosphate moiety can radically alter peptide-MHC binding affinity (Mohammed, Cobbold et al. 2008). However, the degree to which the phosphate affects binding affinity was also shown to be epitope dependent. Therefore the HLA-A2 binding affinities of the phosphopeptides used for this study (RQIpS, RLSpS, RTLpS, RTFpS, RLQpS and RQApS_V) were determined in collaboration with Prof. Victor Englehard using previously described methods (Mohammed, Cobbold et al. 2008) along with their non-phosphorylated counterparts. These results were collated with the previously determined canonical phosphopeptide epitope affinities (Mohammed, Cobbold et al. 2008) and documented in Table 3.5.

Comparison of the binding affinities of each canonical phosphopeptide with their non-phosphorylated counterpart for MHC reveals three distinct groups; firstly those with sub-dominant anchors (RQA_M) which display the highest phosphate induced increase in

Table 3.5 Affinity analysis of HLA-A2 phosphopeptides and their non-phosphorylated counterparts

Sequence	Source Protein	Peptide length (amino acids)	IC_{50} (nM) Nonphosphorylated peptide	IC_{50} (nM) Phosphorylated peptide	Fold increase
RQApSIELPSMAV	Lymphocyte-specific protein 1	12	26	60	-2.3
R**L**SpSPLHF**V**	ORF 17, chromosome 2	9	27.0	16.0	1.7
R**L**QpSTSER**L**	Mitochondrial escape 1-like 1	9	325.0	122.5	2.7
K**M**DpSFLDMQ**L**	Nedd4 binding protein 2	10	11.4	3.4	3.4
R**V**ApSPTSG**V**	Insulin receptor substrate 2	9	731.3	178.5	4.1
R**T**YpSGPMNK**V**	Premature ovarian failure, 1B	10	792.8	116.9	6.8
R**Q**ApSLSIS**V**	Protein kinase D2	9	284.5	38.5	7.4
R**T**LpSHISE**A**	FLJ13725	9	165.0	16.5	10.0
R**T**FpSPTYG**L**	β-synemin/Desmuslin	9	408.2	32.5	12.6
R**Q**IpSQDKV**L**	Adenosine monophosphate deaminase 2	9	1,925.0	25.5	75.5
R**Q**ApSIELPS**M**	Lymphocyte-specific protein 1	10	1,769.0	11.2	158.6

Competitive binding assays were performed as previously described (Mohammed, Cobbold et al. 2008). The peptide concentrations covered a 100,000 fold range. The concentration at which test peptide displaced 50% of the radiolabeled peptide (IC_{50}) was calculated allowing a reasonable approximation of the dissociation constant (Mohammed, Cobbold et al. 2008).

binding affinity, secondly those with a sub-dominant P2 anchor (T and Q) but with a dominant PC anchor (V, L or I) (RVApS, RTYpS, RQApS, RTFpS, RQIpS) and thirdly those that possess dominant P2 (L or M) and PC anchor residues (RLSpS, RLQpS, KMDpS). These results indicate that minimal phosphate enhanced binding effects are observed with epitopes that encompass dominant P2 and PC anchor residues. However, in cases where peptides contain sub-dominant anchor residues, the phosphate can dramatically enhance epitope binding compared to its non-phosphorylated counterpart. These new epitopes (RTFS, RQIS, RLSS and RLQS) confirm and extend the results shown by Mohammed *et al* (Mohammed, Cobbold et al. 2008) in that the phosphate can have a dramatic affect on binding affinity in the presence of subdominant anchor residues.

However, two epitopes (RQA_V and RTLpS) do not conform to this pattern. For RTLS this is most likely due to interactions mediated by residues other than anchor residues which serve to increase epitope binding for non-phosphorylated RTLS (which contains no dominant anchor residues). In the case of RQA_V a small decrease in binding affinity upon phosphate addition is observed. This may possibly result from a loss of stabilising interactions present in the non-phosphorylated version which may be lost upon phosphate addition.

3.10 Conclusions

Prior to this study canonical HLA-A2 phosphopeptide antigens were shown to adopt a conserved binding mode and that the phosphate moiety could dramatically increase antigen binding affinity for some epitopes. However, upon expanding the number of canonical phosphopeptide structures solved, this study has demonstrated that the binding mode of phosphopeptides is considerably more diverse than previously thought. The plasticity of the

phosphate moiety in certain epitopes allows the phosphate to form new or additional contacts to surrounding residues of the MHC peptide binding groove. Upon correlation of the phosphopeptide structures with their binding affinities it is clear that the plasticity of the phosphate seems to have little effect upon the degree of phosphate dependant binding as the two epitopes that display phosphate moiety plasticity, RTFpS and RLSpS, show a high and low degree of phosphate dependant binding respectively. This demonstrates that the plasticity of the phosphate moiety observed in these epitopes is possibly a result of the phosphate adopting an altered conformation with optimal energetic potential for the given epitope and that factors such as the presence or absence of canonical anchor residues play a much more prominent role in determining the overall antigen binding energetics. However, although observations in the binding affinities of canonical phosphopeptides can be made, such as how different anchor residues or how the plasticity of the phosphate may affect affinity, it is extremely difficult to generate predictions of binding affinities based on epitope sequence alone (as demonstrated by RQA_Vp and RTLpS). This may be due to the affinity being dependent upon not only P2 and PC anchor residues but also upon other interactions including intra-peptide and peptide-MHC interactions along the entire epitope which may change upon addition of the phosphate moiety.

Despite this plasticity in phosphate moiety positioning, the conformation of the main chain for each of the canonical phosphopeptide epitopes in this study show a relatively conserved raised P4 position consistent with those of Mohammed *et al* (Mohammed, Cobbold et al. 2008). However, it is still unclear to what extent the presence of the phosphate moiety has upon determining this raised conformation. These effects could conceivably impact upon the

overall structure of presented antigen but also upon antigen recognition. The molecular mechanisms that govern this will be explored in the following chapters.

4 Phosphate induced effects on HLA-A2 antigen identity

4.1 Background

Due to the alteration/dysregulation of the phophoproteome during infection and tumourigenesis (Greenman, Stephens et al. 2007), phosphopeptide antigens are emerging as potential targets for use in immunotherapeutic approaches. Studies have shown that phosphorylated antigens can be presented by MHC class I and II molecules at the cell surface and be recognised by phosphopeptide specific T-cells in a phosphate dependent and epitope specific manner (Zarling, Ficarro et al. 2000; Zarling, Polefrone et al. 2006; Depontieu, Qian et al. 2009). However the molecular basis of how phosphopeptide antigens are presented and distinguished from their non-phosphorylated counterpart is poorly understood. To date, only a single study has characterised the structures of phosphorylated antigen with their non-phosphorylated counterparts (Petersen, Wurzbacher et al. 2009). This study examined three epitopes, only one of which conformed to the canonical phosphopeptide motif which demonstrated no change in overall conformation. However, in a previous study four separate canonical phosphopeptide structures were compared with the structures of several unrelated non-phosphorylated epitopes obtained from the protein data bank (Mohammed et al., 2008). Although this was not a direct comparison of phosphorylated versus non-phosphorylated antigen pairs, the study demonstrated a strikingly conserved raised conformation of the P4 pSer compared to the other non-related non-phosphopeptide antigens, suggesting that the phosphate may have an effect in restricting antigen main chain

conformation between P1-P4. These studies highlighted two possible mechanisms by which phosphorylated antigens may be distinguished from their non-phosphorylated counterparts. Firstly, in the case of the IRS2 antigen described in the Petersen *et al* study, the phosphate may have no effect on overall antigenic structure and in such cases discrimination is likely to involve direct TCR-mediated contact to the phosphate moiety, which is likely going to depend on its position within the epitope. Alternatively, the Mohammed *et al* study demonstrated a conserved conformation of canonical phosphopeptides between P1-P4 thus highlighting the possibility that for some epitopes phosphorylation may induce a conformational change along the length of the epitope due to this apparent restriction in the binding mode of the antigen.

Based on this data it is apparent that the current understanding of how the phosphate affects antigen identity remains poorly defined. Having shown that the phosphate can dramatically influence antigen binding and that there is considerable plasticity in the binding mode of the phosphate moiety within canonical phosphopeptides, this chapter will focus on the effects the phosphate moiety has on antigen conformation and its impact on antigen recognition will be explored.

4.2 HLA-A2-nonphosphopeptide complex production

HLA-A2-non-phosphopeptide complexes were produced in the same manner as HLA-A2-phosphopeptide complexes as described in Chapter 3. Initially, epitope selection for the non-phosphopeptides was focused on those with available structural data for their phosphorylated equivalents. However, subsequent structural studies led to the inclusion of other epitopes that lacked structural data for their phosphorylated counterparts with attempts to obtain structural data in both phosphorylated and non-phosphorylated forms

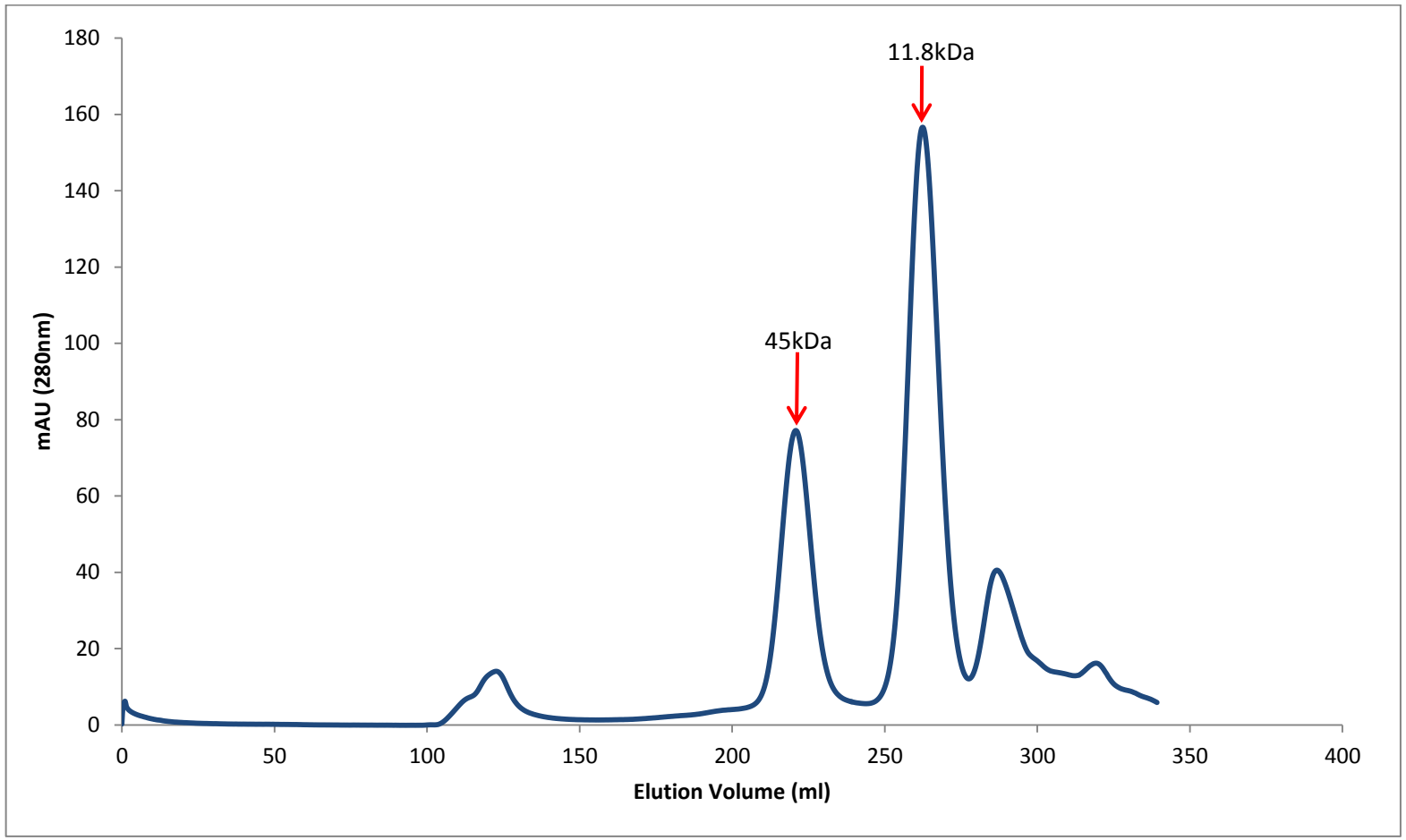
conducted in parallel. As with the production of previous HLA-A2-phosphopeptide complexes, renaturation of HLA-A2-non-phosphopeptide complexes with *in vitro* dilution refolding strategies provided high yields of protein to a sufficient purity for subsequent crystallisation screening studies (Figure 4.1 and 4.2).

4.3 Crystallisation

4.3.1 Crystallisation of HLA-A2-non-phosphopeptide complexes using commercial screens

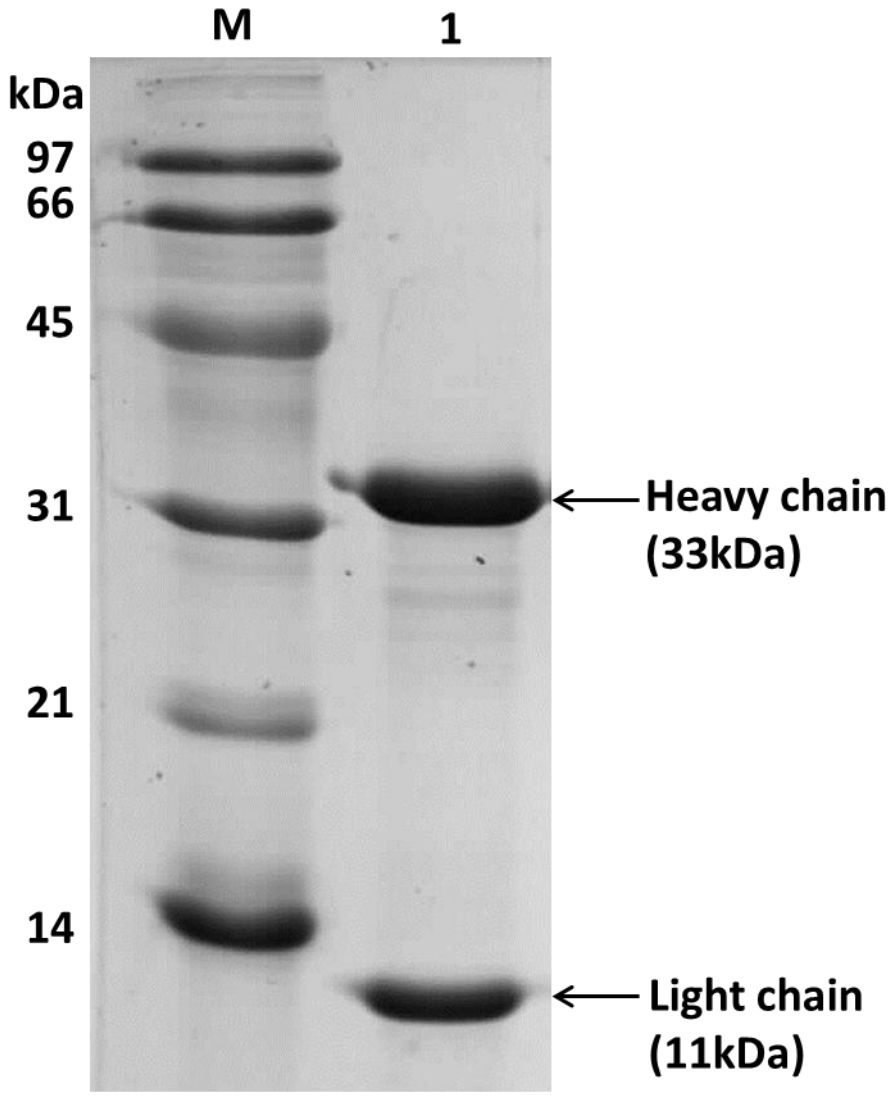
Initial crystallisation screening trials were performed using purified HLA-A2-non-phosphopeptide at 10-15 mg/ml with a range of sparse matrix based commercial screens (including JCSG-plus, PACT, PEG Ion I+II, PEGRx, Wizard I+II, Wizard III+IV, Index and Structure screens I+II). Generally, the non-phosphorylated complexes were intransigent to crystallisation in these standard trials with the exception of two cases (Table4.1). Firstly, the HLA-A2-RQISQDVKL (HLA-A2-RQISnp) complex crystallised in 3 unique conditions obtained from the PEG Ion I+II screen and secondly crystals of the HLA-A2-RTLSHISEA (HLA-A2-RTLSnp) complex were identified with the Wizard I+II and Proplex screens. In order to obtain crystals of sufficient size and quality for X-ray diffraction studies, multiple rounds of optimisation were performed. Large crystals of HLA-A2-RQISnp, which were sufficient for X-ray diffraction experiments, were manually grown in 0.1M Sodium Tartrate dibasic dehydrate, 0.1M Bis Tris propane pH6.5, 19% PEG 3350 using hanging drop vapour diffusion with a drop size of 1µl+1µl (Table 4.1 and Figure 4.3 A). Unfortunately, despite extensive efforts, it was not possible to generate diffraction-grade crystals for the HLA-A2-RTLSnp complex (Table 4.1 and Figure 4.3 B).

Figure 4.1 Size exclusion chromatography purification of a typical HLA-A2-nonphosphopeptide complex



Blue line shows a size exclusion elution profile of the HLA-A2-RLSSnp complex using a Superdex 200 column. Red arrows indicate the elution peaks of the HLA-A2-RLSSnp complex (45kDa), and free β_2M (11.8kDa) which has not been incorporated into the HLA-A2 complex.

Figure 4.2 SDS-PAGE analysis of purified HLA-A2-nonphosphopeptide complex used for crystallisation



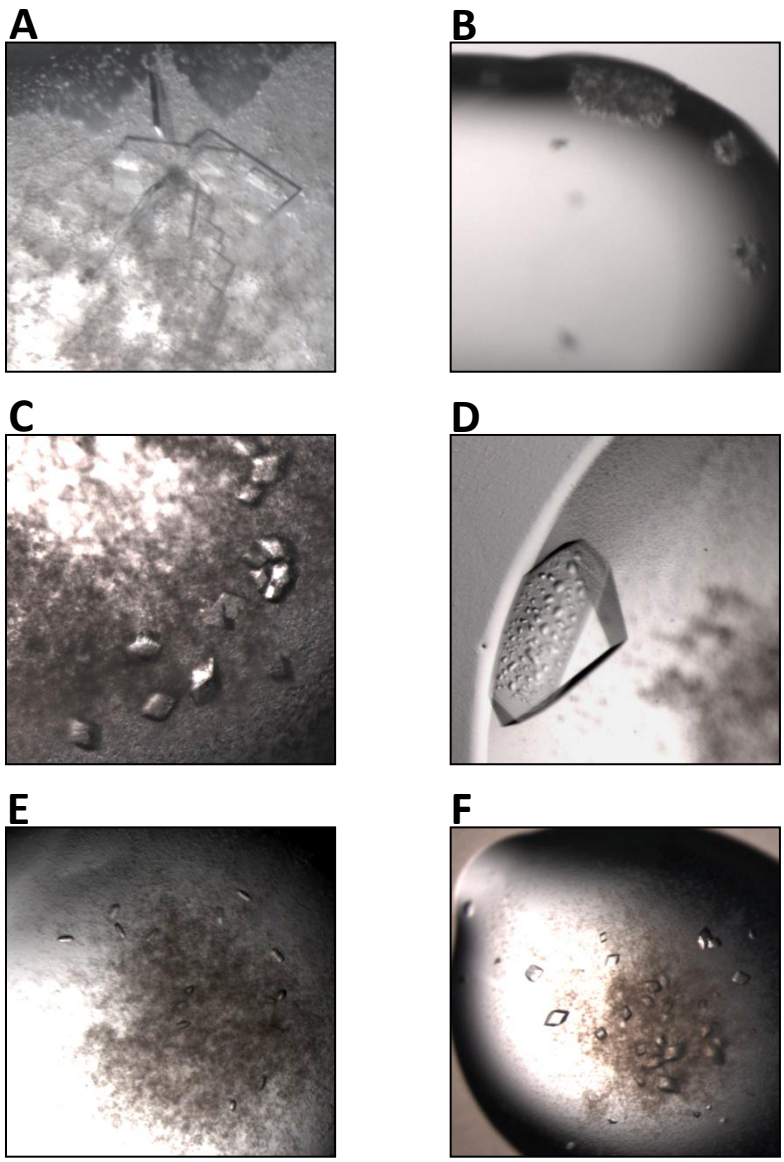
Typical SDS-PAGE analysis of gel filtration purified HLA-A2-nonphosphopeptide complexes. Molecular weight standards are shown in the first lane with masses indicated in kDa. Lane 1 shows HLA-A2-nonphosphopeptide complex analysed under reducing conditions. The samples were of sufficient purity for subsequent crystallisation trials.

Table 4.1 Conditions tested for crystallisation of HLA-A2-nonphosphopeptide complexes

Epitope	Source Protein	Protein conc. (mg/ml)	Screens Tested	Hits
RQLSSGVSEI	Heat shock protein 27	12.1	PACT	1
		12.1	Index	2
		12.1	Peg Ion I+II	5
RLSSPLHFV	ORF 17, chromosome 2	12, 11	Peg Ion I+II	0
		12	Index	0
		12	PACT	0
		10	Structure Screen I+II	0
		12	Petersen	0
RQISQDVKL	Adenosine monophosphate deaminase 2	12	**Peg Ion I+II**	3
		12	Index	0
		12	PACT	1
		12	Petersen	0
RLQSTSERL	Mitochondrial escape 1-like 1	12.46	Wizard III+IV	0
		12.46	JCSG-Plus	0
		12.46	Wizard I+II	1
		12.46	PACT	0
		12	Petersen	0
RTLSHISEA	FLJ13725	10	JCSG-Plus	0
		14, 10	Structure Screen I+II	0
		14, 10	Index	0
		10	Proplex	1
		14	Wizard I+II	1
		13.5	Peg Ion I+II	0
		14	PACT	0
		14	Wizard III+IV	0
		11	Petersen	0
RTFSPTYGL	β-synemin/Desmuslin	13.2, 11	PACT	0
		10	JCSG-Plus	0
		13.2, 10	Structure Screen I+II	0
		13.2, 10	Index	0
		13.2, 7.73	Wizard III+IV	0
		7.73	Peg Rx	0
		13.2	Wizard I+II	0
		10	Peg Ion I+II	0
		11	Petersen	0
RLDSYVRSL	Trafficking protein particle complex 1	11.16	Index	0
		11.16	JCSG-Plus	0
		11.16	Peg Ion I+II	0
		11.16	PACT	0
		11.16	Peg Rx	0
		11.16	Wizard I+II	0
		11	Petersen	0
RLFSKELRC	TFIID transcription initiation factor subunit 13	12	Peg Ion I+II	0
		12	Wizard I+II	0
		12	PACT	0
KMDSFLDMQL	Nedd4 binding protein 2	8, 9, 10, 11, 12	Petersen	0
RTYSGPMNKV	Premature ovarian failure, 1B	7.7	Petersen	0

Initial crystallisation conditions for each complex were tested with sparse-matrix based commercial screens using the hanging drop vapour diffusion method. All hanging drop screens were set up using the Mosquito nanolitre crystallisation robot. The success of crystal generation in the initial screens and further optimisation to produce diffraction grade crystals was very poor. Screens highlighted in red indicate conditions which yielded initial crystals but were unable to produce diffraction grade crystals. The screen shown in green yielded preliminary hits that enabled further optimisation and ultimately produced crystals suitable for X-ray diffraction experiments.

Figure 4.3 Crystals of HLA-A2-nonphosphopeptide complexes

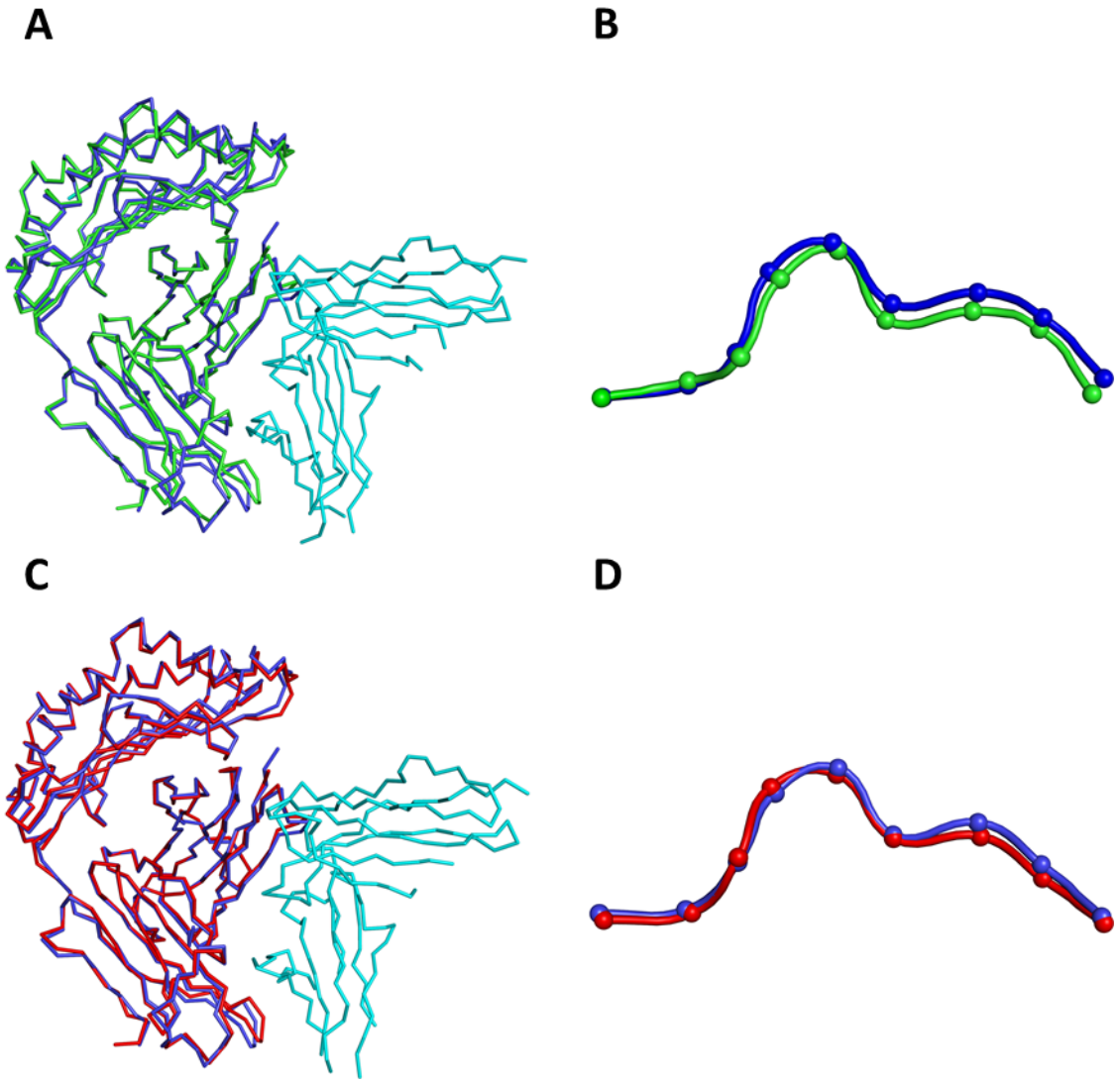


HLA-A2-nonphosphopeptide crystals of the HLA-A2-RQISnp complex which were crystallised in 0.1M Sodium Tartrate dibasic dehydrate, 0.1M Bis Tris propane pH6.5, 19% PEG 3350 using hanging drop vapour diffusion in a drop size of 1µ+1µl (A). These crystals typically took 2 days to appear at 23°C. The HLA-A2-RTLSnp complex, which was crystallised in 1M Sodium citrate, 0.01M Imidazole pH8, 0.1M Sodium Iodide as above, produced microcrystals intransigent to optimisation (B). (C) LILRB1-HLA-A2-RLSSnp crystals crystallised in 0.1M Tris pH 8.5, 0.2M Ammonium Acetate, 18% PEG 3350 using hanging drop vapour diffusion with a drop size of 2µl+2µl. Crystals in this condition typically took 2-3 weeks to appear at 23°C. LILRB1-HLA-A2-RTFSnp (D), LILRB1-HLA-A2-RLQSnp (E) and LILRB1-HLA-A2-RLSSnp (F) following optimisation with the Additive Screen (Molecular Dimensions). These crystals were produced using hanging drop vapour diffusion with a drop size of 1µl+1µl. Crystals typically took 2 days to appear. All crystals apart from HLA-A2-RTLSnp crystals were approximately 100µm x 100µm in size.

4.3.2 Crystallisation of HLA-A2-non-phosphopeptide complexes with LILRB1

As mentioned above the majority of HLA-A2-non-phosphopeptide complexes were intransigent to crystallisation using conventional approaches, including conditions which have previously led to crystallisation of a wide range of HLA-A2-peptide complexes (Table 4.1) (Petersen, Wurzbacher et al. 2009). In order to circumvent intransigent crystallisation, a novel method was explored in which the HLA-A2-non-phosphopeptide complex was co-crystallised with a previously identified MHC-binding protein. An obvious candidate was LILRB1 (Anderson and Allen 2009), a broadly expressed inhibitory receptor that binds MHC class I with low affinity and which has previously been co-crystallised in complex with HLA-A2 to 3.4Å resolution (Willcox, Thomas et al. 2003). Crucially, the LILRB1 binding site on HLA-A2 involves the non-polymorphic α3 and β_2M domains and does not interfere with the α1-α2 peptide-binding platform (Willcox, Thomas et al. 2003). In order to verify that binding of LILRB1 to MHC-class I does not alter the peptide binding platform or epitope conformation, two identical HLA-A2 structures previously crystallised with and without LILRB1 (Madden, Garboczi et al. 1993; Willcox, Thomas et al. 2003) were compared. The α1α2 domains of each complex were found to be highly conserved with a root mean square deviation (RMSD) of 0.6Å and the epitope also adopted essentially identical conformations within the antigen binding groove (Figure 4.4 A/B) highlighted by a low rmsd RMSD of 0.35Å. This was confirmed with a higher resolution structure of the LILRB1-HLA-A2-ILK complex determined to 2.4Å resolution, which enabled a more accurate structure of the peptide moiety. Notably, the final optimised crystallisation condition that yielded the higher-resolution LILRB1-HLA-A2-ILK complex structure consisted of 20% PEG 3350, 0.2 M ammonium acetate and 0.1 M Tris-HCL pH 8.5 (Index screen; Hampton Research), which is subtly different from the

Figure 4.4 Comparison of HLA-A2 structure with and without LILRB1



Comparison of the overall structure of the Garboczi *et al* HLA-A2-ILK complex (Blue; HLA-A2) with the LILRB1-HLA-A2-ILK complex from the Willcox *et al* study (Green; HLA-A2, Cyan; LILRB1) reveals very little difference in α1α2 domain conformation demonstrated by a low RMSD of 0.6Å (A). Analysis of each peptide (Garboczi - Green, Willcox - Blue) also shows that binding of LILRB1 does not affect the conformation of the bound epitope highlighted by a low RMSD of 0.35Å (B). Following the generation of a higher resolution structure of LILRB1-HLA-A2-ILK (Red; HLA-A2, Cyan; LILRB1), comparison of this structure with the HLA-A2-ILK complex from the Garboczi *et al* study (Blue; HLA-A2) conclusively demonstrated that binding of LILRB1 has no effect upon α1α2 domain and antigen conformation as shown by low RMSD values of 0.6Å and 0.24Å respectively.

chemical compounds used to determine the corresponding low resolution complex structure (Willcox et al., 2003); the precipitant 20% PEG 3350 is substituted with 30% PEG 4K. Despite these subtle differences in the LILRB1-HLA-A2-ILK crystallisation condition, the high resolution complex crystals adopted an alternate space group of $P3_221$ (rather than $P3_121$ in the low resolution structure), with a larger unit cell and twice the number of LILRB1-HLA-A2 complexes per asymmetric unit. To eliminate the possibility of altered crystal contacts affecting epitope conformation, we compared the high resolution ILK epitope with that of the same HLA-A2 bound peptide structure determined in the absence of LILRB1. Structural overlay comparisons demonstrated that the peptide binding platform in both complexes was very similar with an RMSD value of 0.6Å. Most crucially, the peptide itself showed no changes in epitope structure upon LILRB1 binding to HLA-A2 (Figure 4.4 C/D) with an RMSD of 0.24Å. Hence, the HLA-A2-LILRB1 co-crystallisation strategy was potentially an attractive strategy for overcoming intransigent crystallisation of HLA-A2-non-phosphopeptide complexes and could be used to explore phosphate mediated effects on antigen identity.

Initially, co-crystallisation trials for each non-phosphopeptide epitope utilised the same condition described above for crystallisation of the high resolution LILRB1-HLA-A2-ILK complex which consisted of 20% PEG 3350, 0.2M ammonium acetate and 0.1M Tris-HCL pH 8.5. For these experiments purified HLA-A2-non-phosphopeptide complexes were mixed with LILRB1 in a 1:1 ratio at 13.5mg/ml (Table 4.2). There was limited success in producing diffraction grade crystals for specific epitopes such as LILRB1-HLA-A2-RTFSPTYGL (LILRB1-HLA-A2-RTFSnp), LILRB1-HLA-A2-KMDSFLDMQL (LILRB1-HLA-A2-KMDSnp) and LILRB1-HLA-A2-RLQSTSERL (LILRB1-HLA-A2-RLQSnp) (Table 4.2).

Table 4.2 Crystallisation and optimisation of LILRB1-HLA-A2-nonphosphopeptide complex crystals

HLA-A2 Epitope	Protein conc. (mg/ml)	Screen	Hits	Condition/Crystallisation information
RTFSPTYGL	12, 11, 10, 9	Index Screen, Mohammed *et al* 2012	n/a	0.1M HEPES pH 7.4, 0.2M Ammonium Acetate, 18-20 % PEG 3350. Only small microcrystals
		PEG/Ion I+II		0.2M Potassium Sodium Tartrate, 20% PEG 3350. 1µl+1µl produced a single crystal which diffracted to 2.3Å
KMDSFLDMQL	12, 11, 10, 9	Index Screen, Mohammed *et al* 2012	n/a	0.1M HEPES pH 7.4, 0.2M Ammonium Acetate, 18-20 % PEG 3350. Only small microcrystals
RLSSPLHFV	12, 11, 10, 9	Index Screen, Mohammed *et al* 2012	n/a	0.1M Tris pH 8.5, 0.2M Ammonium Acetate, 18% PEG 3350. 2µl+2µl yielded 1 crystal that produced x-ray diffraction data to 3.4Å
	13.5	PEG/Ion I+II	20	Hits used for further optimisation F1 - 4% Tacsimate pH 6, 12% PEG 3350 G10 – 2% Tacsimate pH 7, 0.1M HEPES pH 7.5, 20% PEG 3350
	13.5	Molecular Dimensions Additive Screen	6	Condition B4 (10% DMSO) hit taken for further optimisation/x-ray diffraction. Final optimised condition 3% Tacsimate pH 6, 12.8% PEG 3350, 10% DMSO
	13.5	Homemade metal ion screen	0	Metal Ion screen based on 3% Tacsimate pH 6, 12.8% PEG 3350, 10% DMSO
RLQSTSERL	13.5	PEG/Ion I+II	13	0.2M Potassium Acetate, 20% PEG 3350 taken forward to 1µl+1µl trials which produced small crystals. This was taken forward for further optimisation.
	13.5	Molecular Dimensions Additive Screen	0	Additive screen based on 0.2M Potassium Acetate

Initial screening results of LILRB1-HLA-A2-nonphosphopeptide complexes and subsequent optimisation.

One of the complexes however, HLA-A2-RLSSnp, despite the poor reproducibility of crystal generation, did yield a single drop that provided diffraction quality crystals (Figure 4.3 C). Despite this, based on the lower levels of crystal reproducibility coupled with higher nucleation sites, the previously identified LILRB1-HLA-A2 crystallisation condition was deemed insufficient for generating diffraction-grade crystals (data not shown).

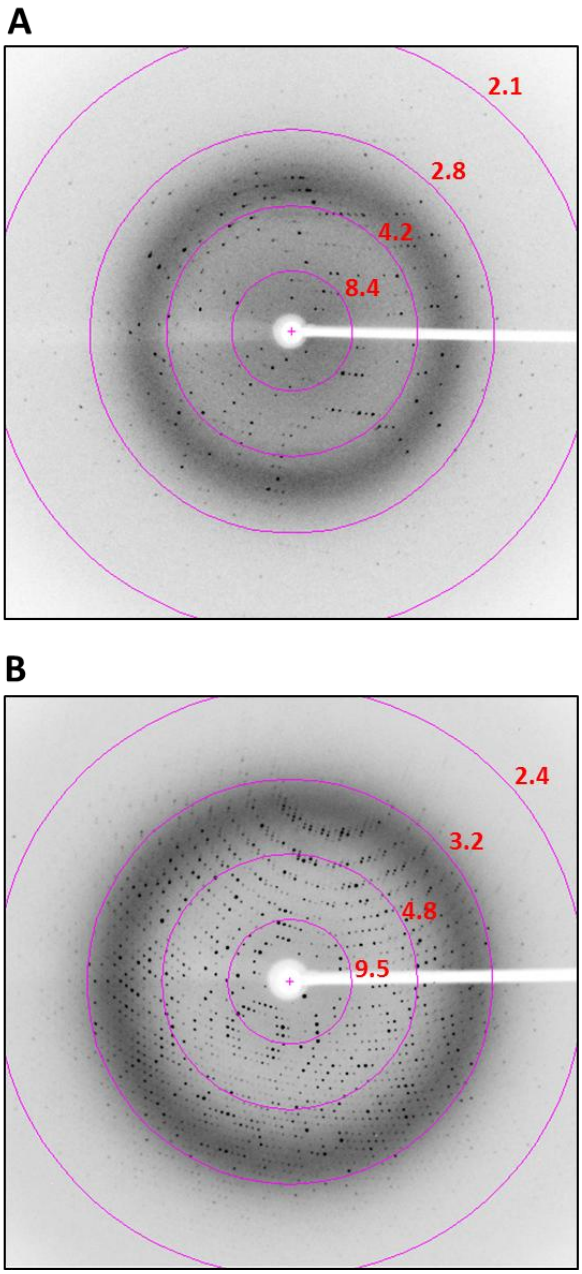
Due to the unreliable nature of this generic HLA-A2-LILRB1 crystallisation condition, a range of commercial screens were used in order to identify novel reagents that consistently produced crystals of sufficient size and quality for X-ray diffraction experiments. Of the several commercial screens tested (PACT, Index, PEG Ion I+II, PEGRx, Structure screen I+II, JCSG-Plus, Proplex, Wizard I+II and Wizard III+IV) the PEG Ion screen yielded promising initial hits for each LILRB1-HLA-A2-epitope complex (Table 4.2). Following optimization of the preliminary conditions (Table 4.2), diffraction-grade crystals for the LILRB1-HLA-A2-RLSSPLHFV (RLSSnp) (Figure 4.3C/F), LILRB1-HLA-A2-RTFSnp (Figure 4.3D) and LILRB1-HLA-A2-RLQSnp (Figure 4.3E) complexes were readily generated for X-ray diffraction studies.

4.4 X-ray Data collection and Processing

The HLA-A2-RQISnp crystals were first soaked in reservoir solution supplemented with increasing amounts (5, 10, 15 and 20%) of ethylene glycol (as cryoprotectant) before being flash-cooled in liquid nitrogen. X-ray data collection was performed as previously described in Chapter 3 and the data collected to a resolution of 2.1Å (Figure 4.5).

The LILRB1-HLA-A2-RLSSnp crystal grown in the previously optimised condition used for the crystallisation of the high resolution LILRB1-HLA-A2-ILK complex was also soaked in increasing amounts of ethylene glycol as above before being flash-cooled in liquid nitrogen.

Figure 4.5 X-ray diffraction patterns of HLA-A2-RQISnp and LILRB1-HLA-A2-RTFSnp complexes



A 0.5° oscillation image collected from (A) HLA-A2-RQISnp and (B) LILRB1-HLA-A2-RTFSnp crystals. Diffraction data for each complex crystal was collected on an in house Micro-Max 007HF rotating anode X-ray generator (Rigaku) using a Saturn CCD detector (University of Birmingham Macromolecular X-ray Facility). The HLA-A2-RQISnp and LILRB1-HLA-A2-RTFSnp crystals diffracted X-rays to 2.1Å and 2.3Å respectively.

However this crystal only diffracted to a resolution of 3.2Å and the resulting data proved extremely difficult to process. Therefore, LILRB1 complex crystals of HLA-A2-RTFSnp, HLA-A2-RLQSnp and HLA-A2-RLSSnp were initially soaked in reservoir solution containing increasing concentrations of ethylene glycol or glycerol (5, 10, 15 and 20 %) before being flash-cooled and stored in liquid nitrogen. Data collection was performed as above and revealed that the glycerol cryoprotectant was sub-optimal due to the presence of ice-rings in the diffraction pattern (data not shown). However, crystals vitrified in ethylene glycol cryoprotectant allowed data to be collected to 2.3Å, 2.8Å and 2.9Å for LILRB1-HLA-A2-RTFSnp, LILRB1-HLA-A2-RLQSnp and LILRB1-HLA-A2-RLSSnp complexes respectively.

Diffraction data sets for both LILRB1-HLA-A2-peptide and HLA-A2-peptide complexes were integrated, scaled and merged using programs of the XDS suite (Kabsch 1993). While X-ray data for HLA-A2-RQISnp and LILRB1-HLA-A2-RTFSnp complexes produced spherical spots with high quality profiles (Figure 4.5) allowing data to be processed to an acceptable standard, the equivalent data for LILRB1-HLA-A2-RLQSnp and LILRB1-HLA-A2-RLSSnp complexes generated diffraction images with streaky spots (data not shown). Hence, it was extremely difficult to process the latter two data sets with acceptable statistics, which in turn prevented subsequent structure solution; these were therefore not included in the analysis. The relevant data processing statistics are listed in Table 4.3.

4.5 HLA-A2-non-phosphopeptide complex structure solution

Initial phase information for HLA-A2-RQISnp and the LILRB1-HLA-A2-RTFSnp complexes were determined by molecular replacement using MOLREP within the CCP4 suite (Murshudov, Vagin et al. 1997). The phasing model used for the HLA-A2-RQISnp complex was as described in Chapter 3 (PDB ID code 1HHG)(Madden, Garboczi et al. 1993).

Table 4.3 Data processing statistics for HLA-A2-RQISnp and LILRB1-HLA-A2-RTFnp complexes

	HLA-A2-RQISnp	LILRB1-HLA-A2-RTFSnp
Space Group	$P2_1$	$P3_221$
Unit Cell Length (Å)	a=56.4,b=79.3,c=56.8	a=116.3,b=116.3,c=192.6
Unit Cell Angles (°)	α=90, β=115.5, γ=90	α=90, β=90, γ=120
Resolution (Å)	20-2.1	20-2.3
Observed Reflections	299127 (94287)	758810(41698)
Unique Reflections	26307 (8581)	66969(7529)
Multiplicity	11.4 (11.0)	11.3 (5.5)
Completeness	99.3 (98.8)	99.1(94.0)
R_{merge} (%)	6.5 (14.5)	11.3(67.3)
I/σ(I)	25.2 (11.4)	23.6(2.8)

X-ray data for all complexes were integrated, scaled and merged using programs of the XDS suite. Numbers in parentheses represent data processing statistics in the highest resolution shell.

In contrast, for the LILRB1-HLA-A2-RTFSnp complex, the phasing model consisted of the LILRB1-HLA-A2 complex (PDB ID code 1P7Q) (Willcox, Thomas et al. 2003), with the peptide co-ordinated omitted to minimise model bias.

Molecular replacement calculations for both the HLA-A2-RQISnp and LILRB1-HLA-A2-RTFSnp complexes yielded unambiguous rotation (θ_1, θ_2 and θ_3) and translation (translation vectors x, y and z) function solutions as demonstrated by peak RF-function and monitor scores, respectively (Values for each complex are documented in Appendix F and G)

4.6 HLA-A2-non-phosphopeptide structure refinement

The HLA-A2-non-phosphopeptide and LILRB1-HLA-A2-non-phosphopeptide structures were refined in the same manner as the HLA-A2-phosphopeptide structures discussed in Chapter 3. The final refinement statistics for each complex is listed in Table 4.4. As with some of the HLA-A2-phosphopeptide complexes both of these complexes display some residues within disallowed regions of the Ramachandran plot. For the HLA-A2-RQISnp complex the single residue in a disallowed region is within a loop region away from the peptide binding platform and is also within clear electron density therefore we can assume that the position of this residue is correct. For the LILRB1-HLA-A2-RTFSnp complex the several disallowed residues reside in flexible loop regions of the MHC and LILRB1 but also of note is that the refinement statistics show that further refinement is required in order to reduce the statistics to an acceptable level, however as the peptide and the surrounding residues of the peptide binding groove are all within electron density, structural comparisons between the non-phosphopeptide and phosphopeptide are able to be performed as the structure of these residues will alter very little during further refinement.

Table 4.4 Refinement Statistics for HLA-A2-RQISnp and LILRB1-HLA-A2-RTFnp complexes

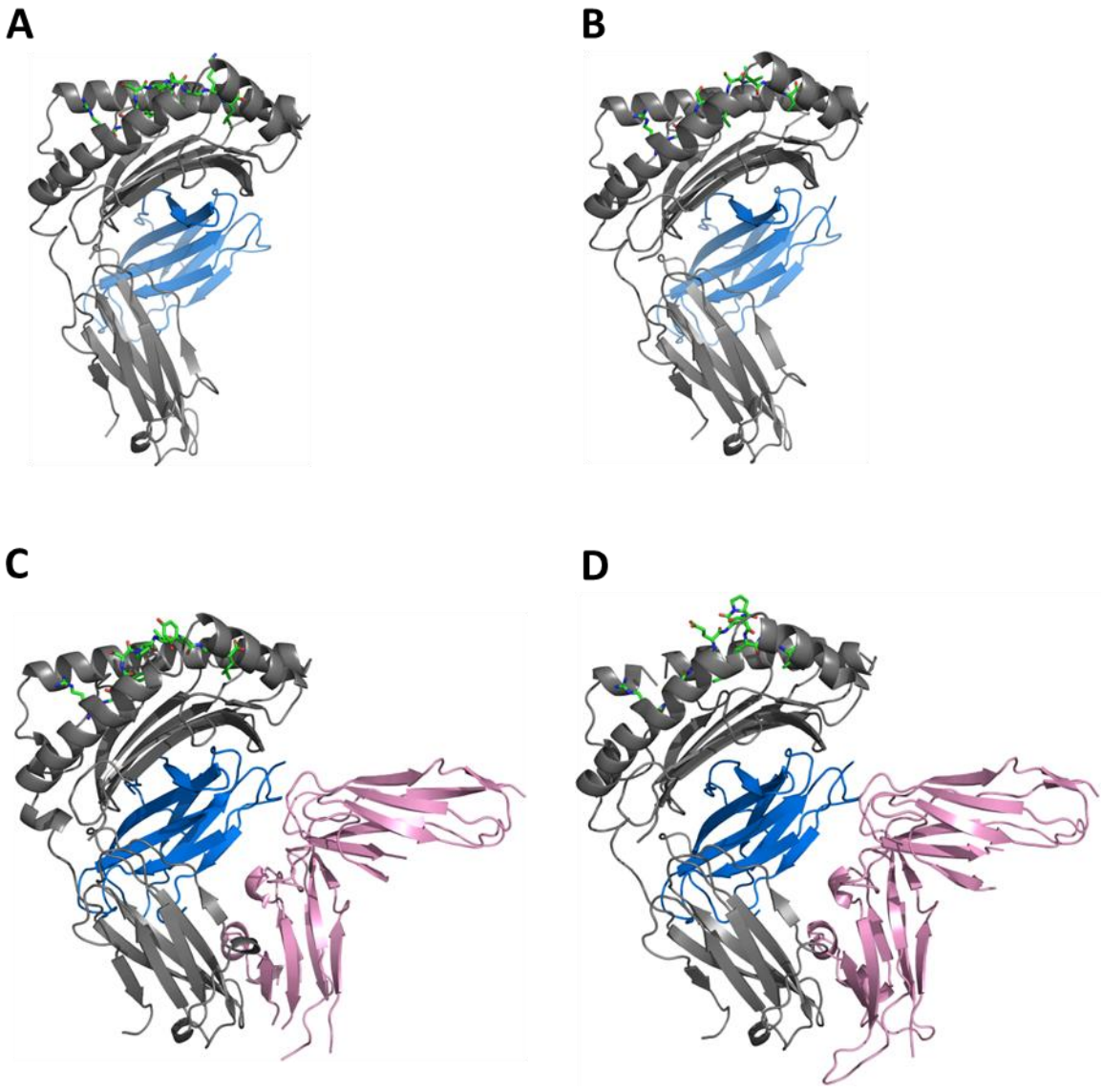
	HLA-A2-RQISnp	LILRB1-HLA-A2-RTFSnp
Resolution Range (Å)	20 - 2.1	20 - 2.3
Ref. in working set	25033	63557
Ref. in test set	1781	3401
R_{cryst} (%)	19.02	26.13
R_{free} (%)	23.67	29.58
Number of protein atoms	3151	9133
Average B-factor (Å^2)	Heavy Chain A – 12.55 Light Chain B – 16.19	Heavy Chain A – 33.80 Heavy Chain E – 35.62 Light Chain B – 26.95 Light Chain F – 26.71 LILRB1 D – 36.11 LILRB1 H – 42.10
Number of water mol.	265	202
RMS Bond Length (Å)	0.016	0.008
RMS Bond Angle (°)	1.46	1.48
Ramachandran Plot		
Allowed (%)	99.4	98.5
Generously Allowed (%)	0.3	0.9
Disallowed (%)	0.3	0.5

Final refinement statistics for each HLA-A2/phosphopeptide complex were generated following several rounds of alternating simulated annealing and positional refinement followed by isotropic B factor refinement and model building. Although both complexes demonstrate some residues in disallowed regions of the ramachandran plot, the disallowed residue in the HLA-A2-RQISnp structure does not form part of the peptide binding groove and is well defined in electron density. The LILRB1-HLA-A2-RTFSnp structure demonstrates a higher degree of disallowed residues due to their position in flexible loops and requires further refinement of these regions.

4.7 Overall Structure of HLA-A2-non-phosphopeptide complexes

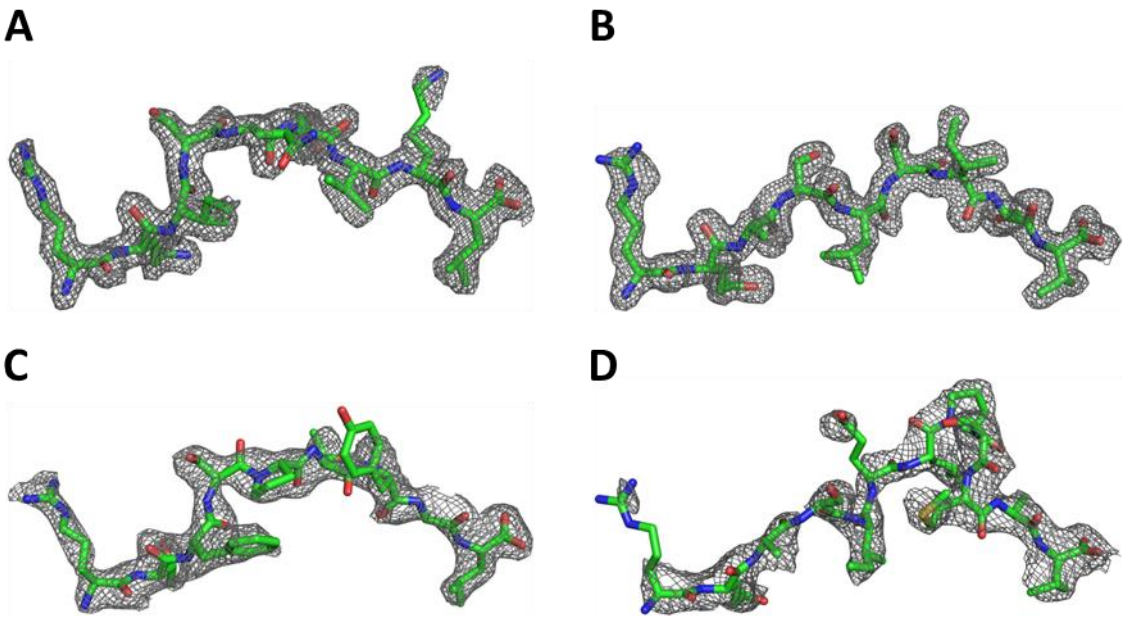
The structural analysis of each HLA-A2-phosphopeptide complex was performed with COOT and the CCP4 software suite (Murshudov, Vagin et al. 1997; Emsley and Cowtan 2004). All structural figures were produced using Pymol (Schrodinger 2010). In addition to the two non-phosphopeptide structures mentioned above, two additional HLA-A2-peptide complexes, HL-A2-RQA_V and HLA-A2-PKD2 (RQAS), were previously solved in their phosphorylated and non-phosphorylated forms by Dr Fiyaz Mohammed and will be included for the remainder of this chapter. As with the phosphorylated versions of each epitope described in Chapter 3 the overall structure of each HLA-A2-non-phosphopeptide complex is similar, both to each other and to previously solved structures of HLA-A2 complexes, including HLA-A2-RTFS and HLA-A2-RQA_V which were crystallised with LILRB1 co-crystallisation strategy (Figure 4.6). Analysis of each complex revealed unambiguous electron density for each epitope (Figure 4.7). Similar to previous observations, the overall mode of binding for each non-phosphopeptide complex is similar to their phosphorylated counterpart in that each peptide adopts an extended structure in the antigen binding groove formed by the α1-α2 platform. Furthermore, each peptide maintains a conserved network of stabilising hydrogen bonding interactions between their N and C terminus and HLA-A2 residues. In addition to this, the anchor residues at position 2 (P2) and the C-terminus (PC) are orientated in a broadly similar orientation to related HLA-A2-phosphopeptide structures (Mohammed, Cobbold et al. 2008), in that they protrude downwards into the B and F pockets, respectively (Figure 4.8).

Figure 4.6 Overall structure of HLA-A2-nonphosphopeptide and LILRB1-HLA-A2-nonphosphopeptide complexes



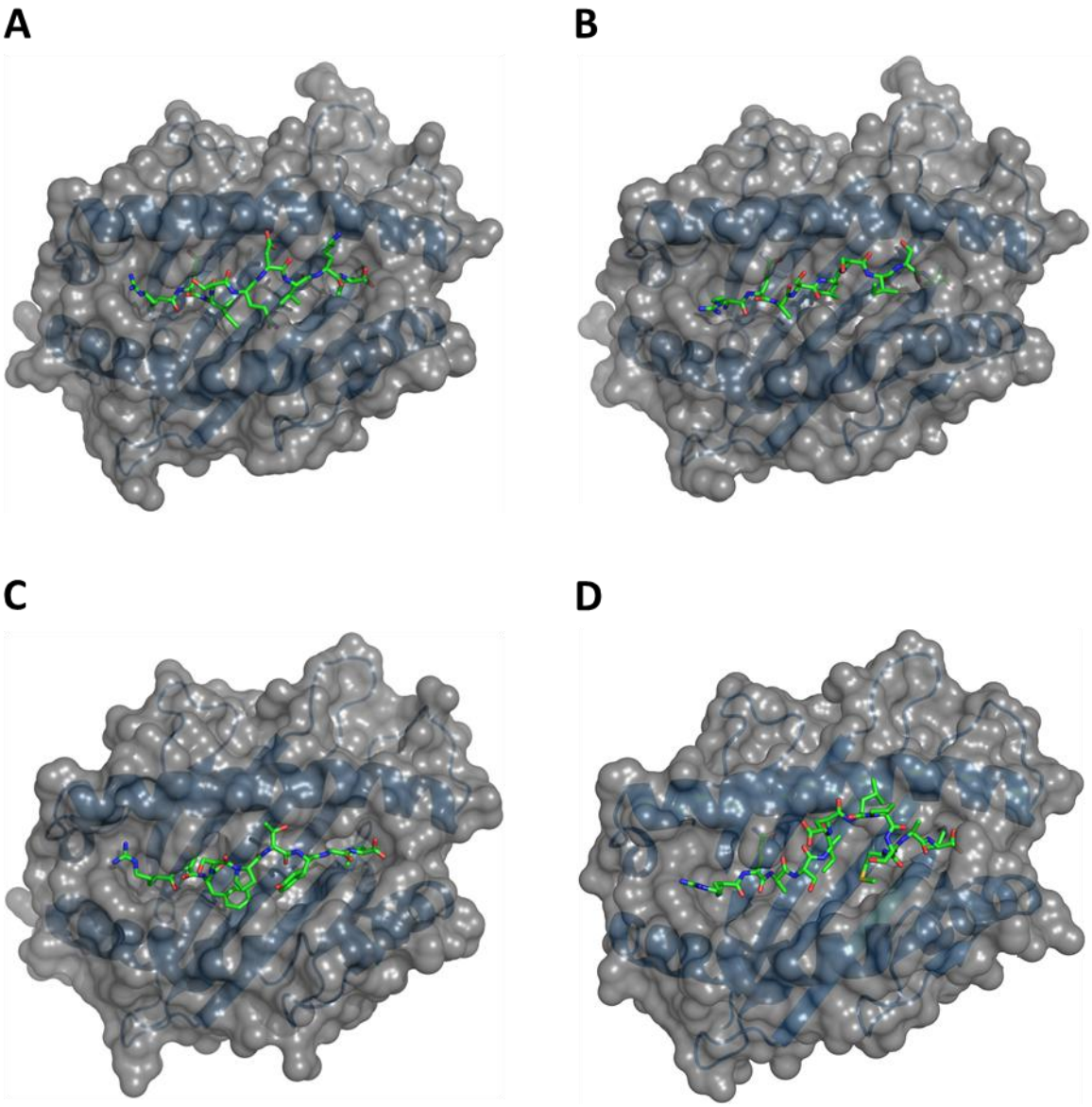
The overall structures of HLA-A2-RQInp (A), HLA-A2-PKD2np (B), LILRB1-HLA-A2-RTFnp (C) and LILRB1-HLA-A2-RQA_Vnp (D) are shown as cartoon representations. The heavy chain of HLA-A2 consisting of the α1, α2 binding platform and membrane proximal α3 domain are shown in grey with the β_2M domain shown in blue. LILRB1 molecules bound to HLA-A2 complexes are highlighted in pink (C and D). The peptide in each complex is shown within the peptide binding groove of the HLA-A2 molecule as stick representation.

Figure 4.7 Electron density for phosphopeptides epitopes in their non-phosphorylated state



Individual phosphopeptide epitope structures of RQISnp (A), PKD2np (B), RTFSnp (C) and RQA_Vnp (D). The 2Fo-Fc electron density map is contoured to 1σ (grey wire).

Figure 4.8 Orthogonal view of non-phosphopeptides in the antigen binding groove of HLA-A2



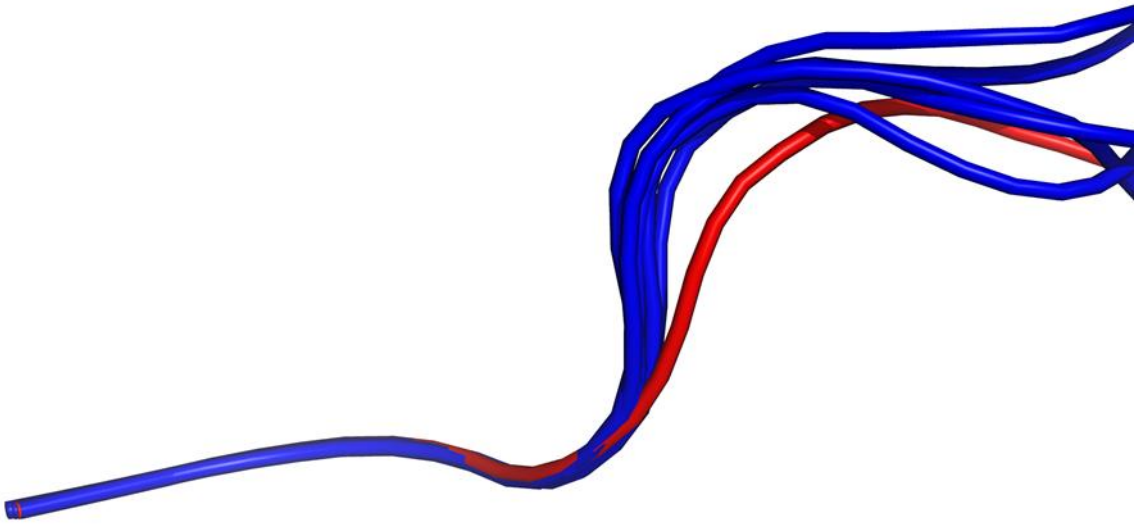
Orthogonal view of HLA-A2-RQISnp (A), HLA-A2-PKD2np (B), HLA-A2-RTFSnp (C) and HLA-A2-RQA_Vnp (D) complexes. Molecular surfaces of α1α2 peptide binding platforms (grey) overlay ribbon representation of the α1α2 regions (dark blue). Phosphopeptides are shown in stick representation.

4.8 Phosphate induced effects on epitope identity

To date there has only been a direct structural comparison of a single canonical phosphopeptide in its phosphorylated and non-phosphorylated state bound to HLA-A2. This study highlighted extremely limited changes in epitope structure upon addition of the phosphate moiety. However, this epitope adopted an unusual conformation relative to previous phosphopeptide structures suggesting it is not typical of the canonical phosphopeptide repertoire (Figure 4.9). This unusual conformation was hypothesised by Peterson *et* al to be due to the limited flexibility of the P5 Pro restricting the main chain from adopting a raised conformation upon phosphorylation of the antigen. Therefore the effect of the phosphate moiety on phosphopeptide that adopt a typical main-chain conformation has remained unclear. To address this issue the HLA-A2-phosphopeptide complexes of HLA-A2-PKD2p, HLA-A2-RQApS_V, HLA-A2-RQIpS and HLA-A2-RTFpS were compared to the structure of their non-phosphorylated counterparts.

In each case the phosphorylated epitopes maintain the conserved raised positioning of the main chain at position 4 of the peptide (see Chapter 3 for details). The PKD2 epitope exhibits phosphate dependent binding for HLA-A2 with a binding affinity of Kd 38.5nm versus Kd 284.5nm for phosphorylated and non-phosphorylated epitopes respectively (See Chapter 3, table 3.1). Although the phosphorylated form of PKD2 had been previously crystallised in complex with HLA-A2 (Mohammed, Cobbold et al. 2008), both the HLA-A2 bound phosphorylated and non-phosphorylated forms were re-crystallised for this study in a previously described condition (Petersen, Wurzbacher et al. 2009). This yielded crystals that in each case belong to the orthorhombic space group $P2_12_12_1$ with very similar unit cell parameters for both phosphorylated and non-phosphorylated complexes and allowed a

Figure 4.9 Comparison of RVApS (IRS2) peptide conformation with other phosphopeptide epitopes

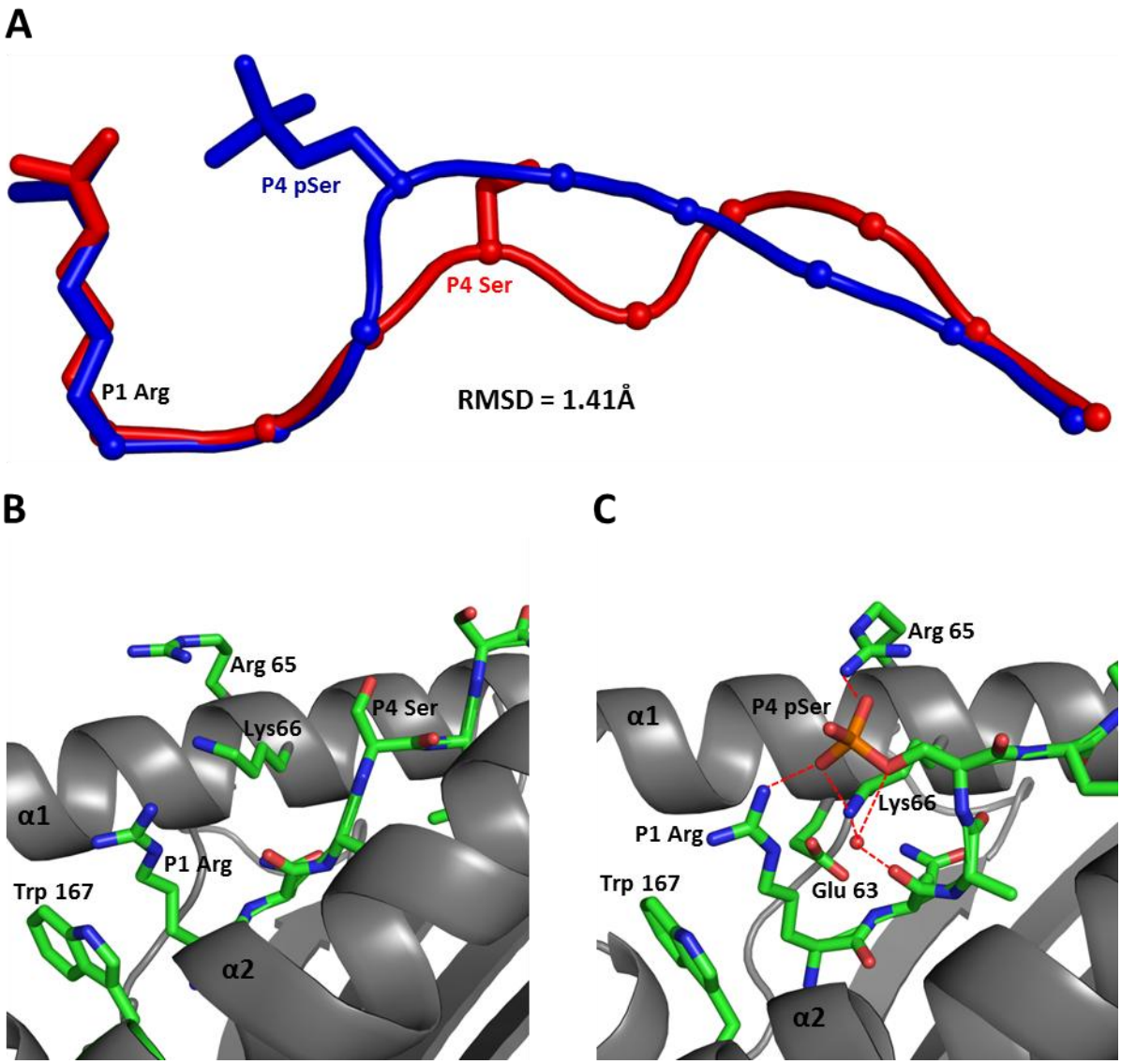


Comparison of the IRS2 epitope RVApS (red) conformation compared to previously solved phosphopeptides and phosphopeptides solved in this study (blue). The IRS2 epitope adopts an unusual conformation compared to other canonical phosphopeptide epitopes.

precise comparison of both structures. The overall MHC main chain structure of the phosphorylated and non-phosphorylated PKD2 structures was found to be very similar with an rmsd of 0.21Å. However the same comparison revealed a striking difference in the epitope main chain conformation as demonstrated by the high rmsd value of 1.41Å which is much greater than the previous comparisons of phosphorylated/non-phosphorylated structures solved by Petersen *etal* (range 0.09-0.45Å) (Petersen, Wurzbacher et al. 2009) (Figure 4.10 A). In the unmodified version of PKD2, the P4 Ser adopts a much lower position within the peptide binding groove and forms no contacts with surrounding residues of the MHC (Figure 4.10 B). Upon phosphorylation, the P4 pSer adopts a raised conformation, enabling the formation of multiple hydrogen-bonding interactions typical of the canonical phosphopeptide group (Figure 4.10 C). These additional interactions introduced upon addition of the phosphate are likely to contribute to the increased affinity of the phosphorylated PKD2p antigen binding to the MHC. The raising of the main chain at position 4 in PKD2p results in an altered main chain conformation within the central region of the epitope and re-orientation of the side chains at position 5, 6 and 7 (Figure 4.11 A). These residues are frequently involved in TCR binding and this alteration in antigen structure within the central region of the epitope causes a dramatic change in the molecular surface (Figure 4.11 B/C) and may therefore significantly impact upon antigen recognition and discrimination by phosphopeptide specific TCRs.

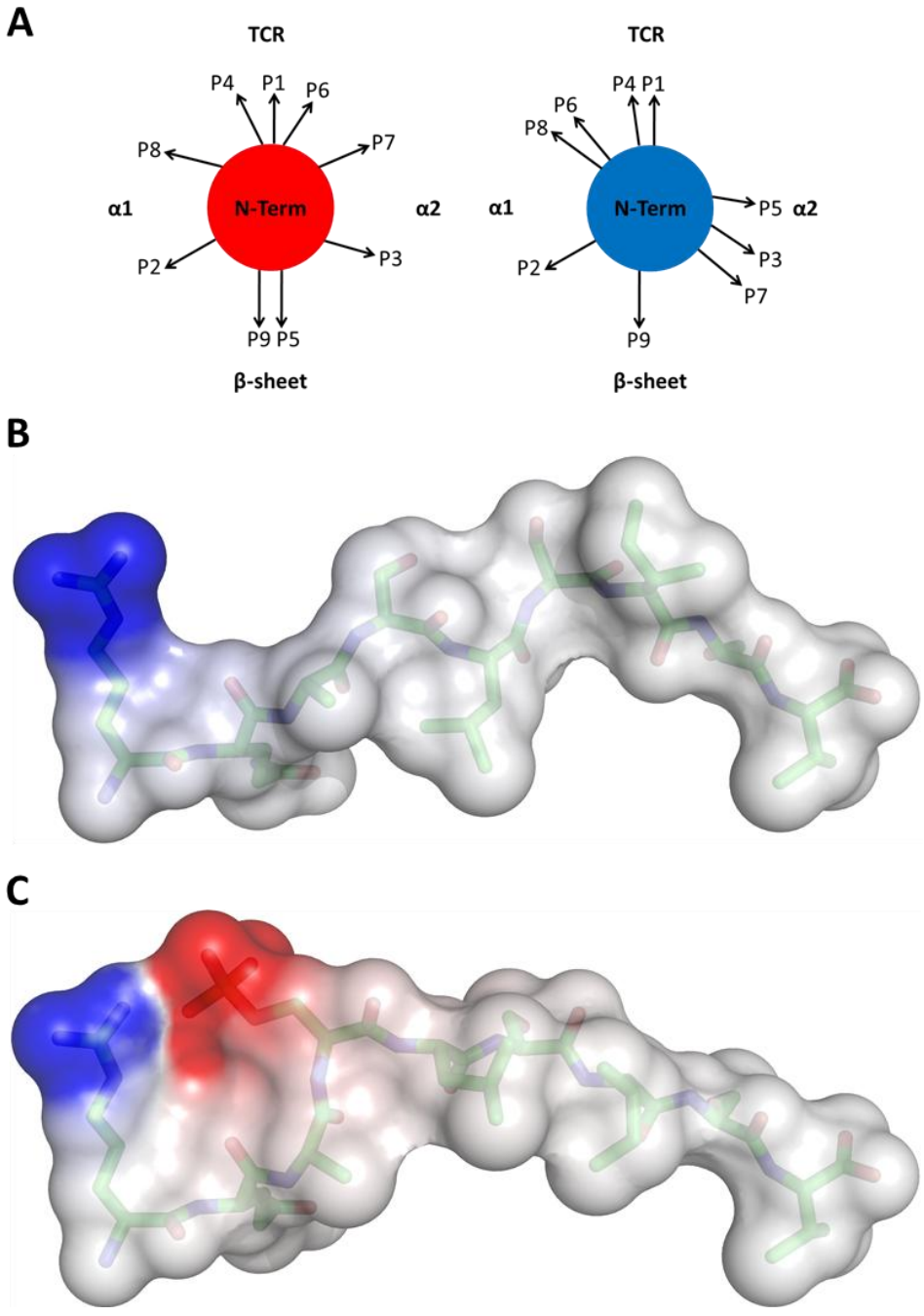
The RQApS_V peptide is a canonical phosphopeptide epitope composed of 12 residues, an unusual length for HLA-A2 binding. Also, peptide-MHC binding studies have highlighted that the binding affinities are highly similar with Kds of 60.0nM and 26nM for phosphorylated and non-phosphorylated epitopes, respectively (See Chapter 3 Table 3.1).

Figure 4.10 Comparison of phosphorylated and non-phosphorylated PKD2 epitope conformation



Phosphorylation of the PKD2 epitope causes substantial conformational change in the antigen main-chain (A), with the phosphopeptide depicted in blue and the non-phosphorylated epitope in red. The interactions mediated by P4 Ser (B)/pSer (C) to HLA-A2 are compared. Hydrogen bonding interactions are shown in red dashed lines. Solvent molecule is represented as a red sphere.

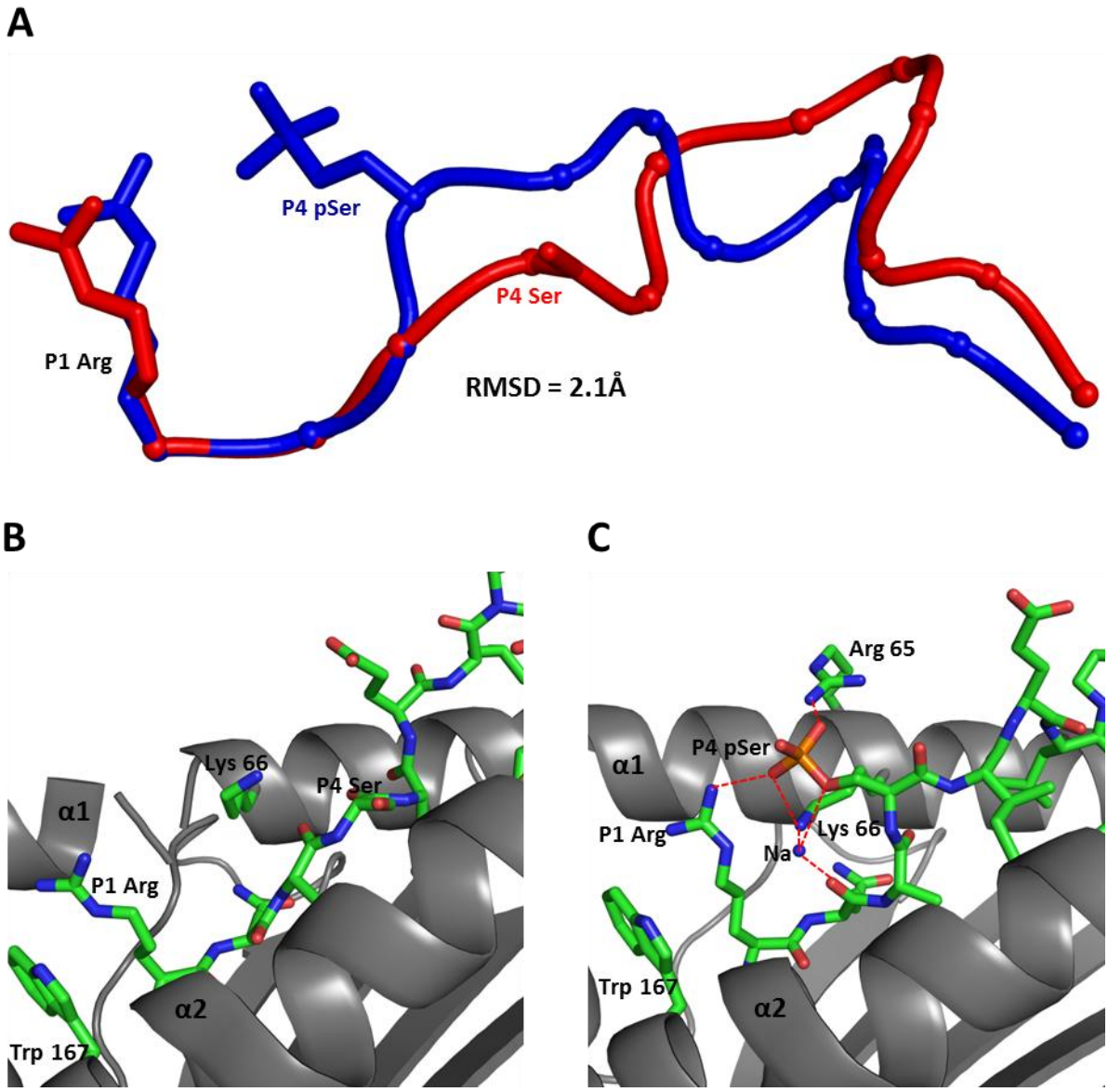
Figure 4.11 Comparison of phosphorylated and non-phosphorylated PKD2 epitope side chain orientation and surface electrostatics



The side chains of P5, P6 and P7 are re-orientated upon phosphorylation of PKD2 (A). The phosphopeptide is shown in blue whereas the non-phosphopeptide counterpart is highlighted in red (A). The surface representation of PKD2np (B) and PKD2p (C) epitopes highlight not only significant charge related differences on the epitope surface(blue, positive; grey, neutral; red, negative), but also substantially different conformation over the central region of the epitope.

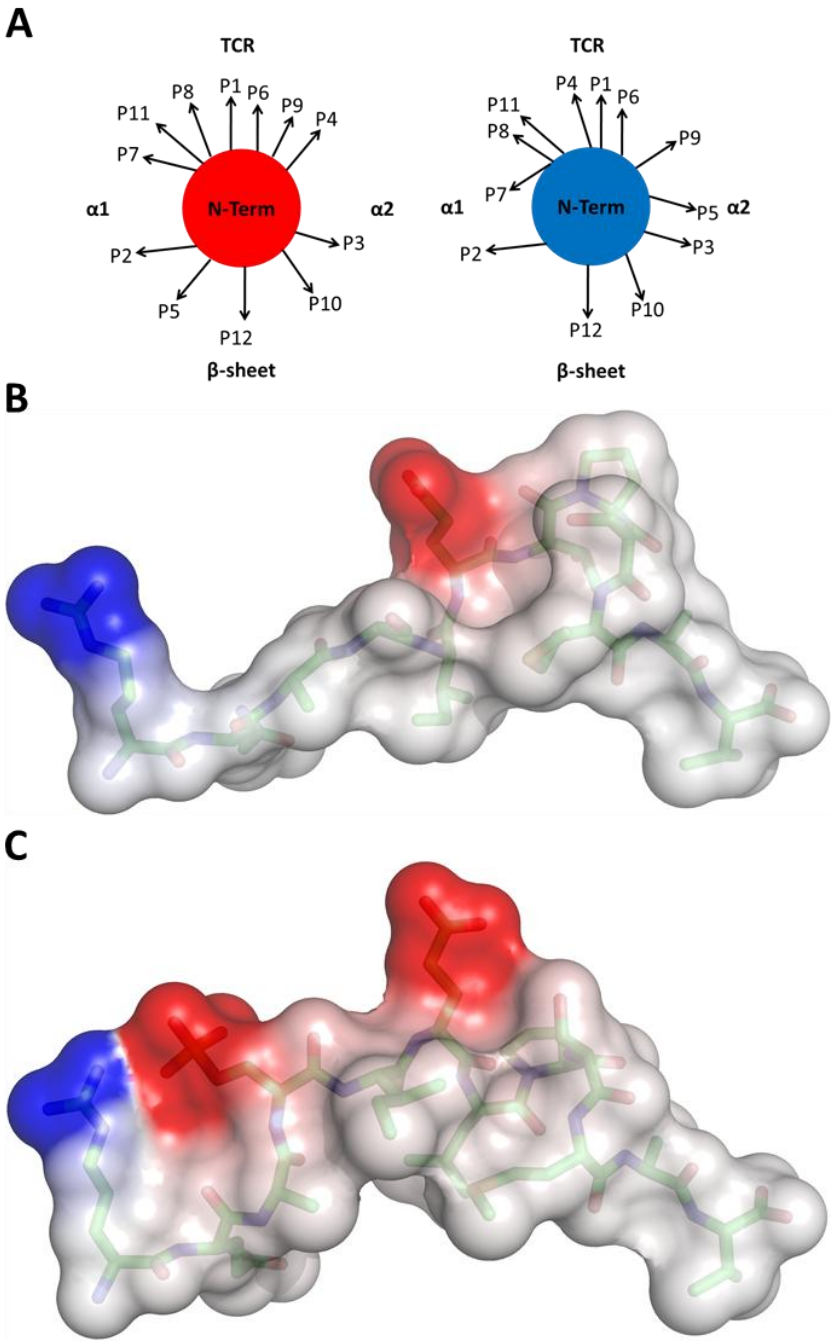
An initial hypothesis that suggested the unusual length of the epitope may disrupt the canonical phosphate binding mode observed for other phosphopeptides was proposed. However upon solving the structure of HLA-A2-RQApS_V to 2.1Å, despite exhibiting a novel helical segment towards the C-terminus of the peptide, the phosphate adopted a position and interactions similar to other canonical phosphopeptides which have been shown to contribute to increased antigen affinity (Chapter 3) (Mohammed, Cobbold et al. 2008). A second hypothesis envisaged that conformational rearrangement of the epitope upon addition of the phosphate moiety may potentially disrupt stabilising interactions which are present in the non-phosphorylated epitope. Therefore the structure of HLA-A2-RQAS_Vnp was solved to 2.7Å in complex with LILRB1. Once again following comparison of the two MHC structures very little difference in overall MHC main chain structure was observed, with an rmsd of 0.75Å. However comparison of the peptide main chains of phosphorylated and non-phosphorylated RQAS_V revealed a dramatic alteration in the conformation of the antigen highlighted by an rmsd of 2.1Å (Figure 4.12 A). In the non-phosphorylated epitope the P4 serine adopts a low position within the antigen binding groove and forms no contacts with surrounding residues (Figure 4.12 B). As with PKD2, when phosphorylated, the P4 pSer raises and permits interactions with surrounding residues (Figure 4.12 C). Also like PKD2 the conformational change is focused on the central region of the epitope with substantial differences in main chain position and side chain orientations at Positions 4, 5, and 7 (Figure 4.13 A), thereby altering the molecular surface of the epitope (Figure 4.13 B/C). These results may also explain the relatively similar binding affinities for the phosphorylated and non-phosphorylated epitopes (See Chapter 3, table 3.1) in that, when in the non-phosphorylated form, the low main chain position adopted by P4 serine and surrounding peptide main-chain allows additional side chain and intra-peptide contacts along the length

Figure 4.12 Comparison of RQA_Vp and RQA_Vnp epitope structure and P4 binding



Phosphorylation of the RQA_V epitope causes substantial conformational change in the antigen main-chain (A) with the phosphopeptide depicted in blue and the non-phosphorylated epitope in red. The interactions mediated by P4 Ser (B)/pSer (C) to HLA-A2 are compared. Hydrogen bonding interactions are shown in red dashed lines. Solvent molecule is represented as a red sphere.

Figure 4.13 Comparison of RQA_Vp and RQA_Vnp epitope side chain orientation and surface electrostatics



Repositioning of the side chains of P4, P5 and P7 upon phosphorylation of RQA_V with the phosphopeptide highlighted in blue and non-phosphopeptide in red. The surface representation of RQA_Vnp (B) and RQA_Vp (C) epitopes highlight not only significant charge related differences on the epitope surface (blue, positive; grey, neutral; red, negative), but also substantial differences in conformation over the central region of the epitope.

of the epitope which help stabilise the complex. These are lost upon phosphorylation due to elevation of P4 serine which results in a more extended conformation compared to the non-phosphorylated epitope. Most striking is the dramatic difference in the orientation and position of Ile5 in RQA_Vnp relative to RQApS_V, allowing numerous hydrophobic interactions with the side-chains of Arg97, His70 and Tyr99 (Table 4.5 and Figure 4.14 A). Also, Tyr159 is positioned closer to the epitope in RQA_Vnp and makes an additional hydrogen-bonding interaction to the backbone carbonyl group of Gln2 (Figure 4.14 B). Finally, although the 3/10 helix is maintained in both phosphorylated and non-phosphorylated epitopes, the intra-peptide hydrophobic interactions between Ile5 and Met10 are present in RQAS_Vnp but absent upon phosphorylation (Table 4.5 and Figure 4.14 C). Therefore, whilst phosphorylation introduces novel contacts between the phosphate moiety and surrounding residues the resulting conformational change leads to a loss of interactions present in the non-phosphorylated form, the net effect of which results in similar overall binding affinities.

Comparison of the RQIS epitope in phosphorylated and non-phosphorylated forms as expected revealed very little difference in overall MHC structure with an rmsd of 0.62Å. However, unlike PKD2 and RQAS_V, phosphorylation of the epitope resulted in very little change in epitope conformation highlighted by a low rmsd of 0.39Å (Figure 4.15 A) despite the phosphopeptide epitope conforming to the canonical binding motif and forming phosphate mediated contacts similar to those described previously (Mohammed, Cobbold et al. 2008). This may explain why the RQIS epitope has a dramatic 75-fold increase (Chapter 3, Table 3.5) in binding affinity upon addition of a phosphate moiety; due to retention of contacts present in the non-phosphorylated epitope coupled with the additional contacts

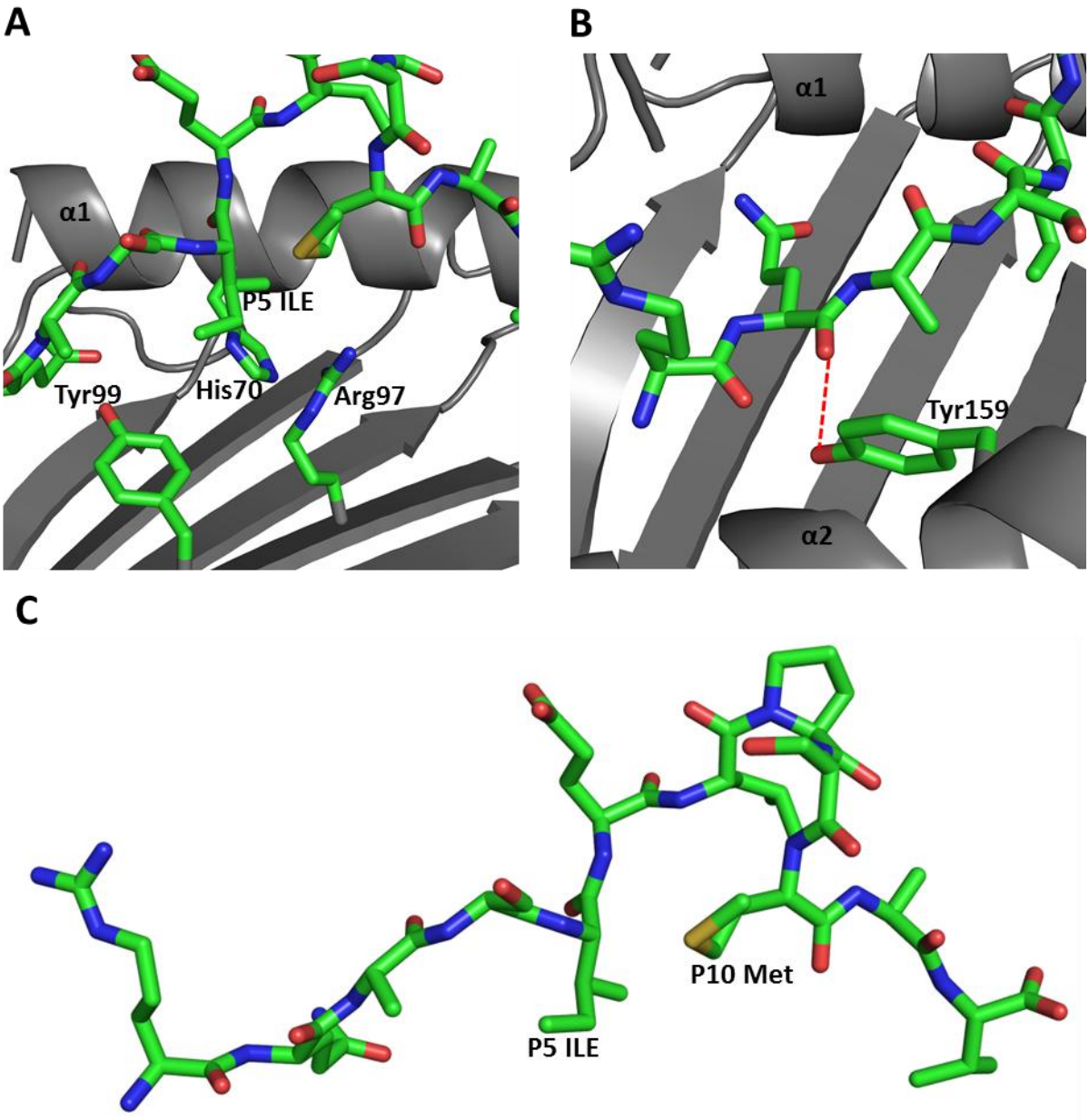
Table 4.5 Comparison of Hydrogen and Hydrophobic contacts in RQA_Vnp and RQA_Vp binding

Residue	Atom	RQA_V phos	RQA_V np
P1 Arg	C	7 (6)	7 (6)
	O	8 (5)	7 (5)
	N	11 (1)	13 (1)
	C^{α}	9 (3)	7 (3)
	C^{β}	9 (2)	7 (1)
	C^{γ}	5 (2)	4 (2)
	C^{δ}	10 (1)	5 (0)
	N^{ϵ}	2 (1)	10 (0)
	C^{ζ}	1 (1)	4 (0)
	NH^{1}	4 (0)	1 (0)
	NH^{2}	0 (3)	6 (0)
P2 Gln	C	6 (7)	6 (6)
	O	5 (12)	9 (7)
	N	8 (5)	7 (5)
	C^{α}	7 (6)	6 (6)
	C^{β}	6 (5)	3 (5)
	C^{γ}	7 (2)	6 (2)
	C^{δ}	8 (0)	8 (0)
	$O^{\epsilon 1}$	11 (0)	9 (0)
	$N^{\epsilon 2}$	12 (0)	11 (0)
P3 Ala	C	1 (8)	3 (9)
	O	3 (9)	5 (7)
	N	7 (7)	3 (7)
	C^{α}	8 (7)	5 (6)
	C^{β}	10 (3)	10 (5)
P4 Ser/pSer	C	0 (8)	0 (10)
	O	0 (6)	1 (11)
	N	0 (7)	1 (5)
	C^{α}	1 (5)	1 (8)
	C^{β}	0 (3)	0 (5)
	O^{γ}	3 (3)	0 (2)
	P	3 (1)	N/A
	O^{1}	4 (3)	N/A
	O^{2}	4 (0)	N/A
	O^{3}	0 (1)	N/A
P5 Ile	C	0 (9)	6 (9)
	0	0 (12)	10 (5)
	N	0 (7)	0 (9)
	C^{α}	0 (7)	0 (8)
	C^{β}	1 (3)	5 (3)
	$C^{\gamma 1}$	4 (2)	8 (3)
	$C^{\gamma 2}$	3 (5)	15 (1)
	$C^{\delta 1}$	9 (0)	2 (6)
P6 Glu	C	0 (13)	2 (7)
	O	0 (18)	3 (7)
	N	0 (7)	2 (8)
	C^{α}	0 (8)	5 (10)
	C^{β}	0 (5)	1 (5)
	C^{γ}	0 (1)	0 (5)
	C^{δ}	0 (0)	0 (1)
	$O^{\epsilon 1}$	0 (1)	0 (0)
	$O^{\epsilon 2}$	0 (0)	0 (0)

Residue	Atom	RQA_V phos	RQA_V np
P7 Leu	C	2 (12)	0 (10)
	O	3 (19)	0 (12)
	N	0 (10)	1 (8)
	C^{α}	1 (11)	1 (13)
	C^{β}	4 (5)	1 (8)
	C^{γ}	5 (1)	2 (6)
	$C^{\delta 1}$	11 (1)	5 (7)
	$C^{\delta 2}$	15 (0)	7 (1)
P8 Pro	C	0 (11)	0 (15)
	O	0 (10)	0 (13)
	N	1 (9)	0 (8)
	C^{α}	1 (10)	0 (10)
	C^{β}	0 (3)	0 (6)
	C^{γ}	0 (5)	0 (3)
	C^{δ}	0 (10)	0 (5)
P9 Ser	C	2 (12)	0 (9)
	O	4 (8)	3 (9)
	N	0 (14)	0 (11)
	C^{α}	0 (10)	0 (7)
	C^{β}	1 (6)	1 (3)
	O^{γ}	1 (5)	0 (4)
P10 Met	C	4 (8)	4 (8)
	O	8 (7)	6 (8)
	N	1 (13)	0 (10)
	C^{α}	4 (10)	1 (10)
	C^{β}	5 (4)	2 (5)
	C^{γ}	1 (6)	1 (3)
	S^{δ}	1 (1)	1 (0)
	C^{ϵ}	8 (0)	3 (3)
P11 Ala	C	3 (10)	4 (10)
	O	6 (9)	4 (10)
	N	2 (12)	2 (11)
	C^{α}	4 (11)	2 (10)
	C^{β}	2 (6)	1 (7)
P12 Val	C	9 (2)	11 (2)
	O	11 (2)	11 (2)
	N	4 (5)	5 (5)
	C^{α}	5 (3)	6 (3)
	C^{β}	7 (2)	6 (2)
	$C^{\gamma 1}$	13 (1)	11 (3)
	$C^{\gamma 2}$	12 (1)	10 (1)
	OXT	7 (4)	9 (4)
Total		**365 (260)**	**345 (248)**

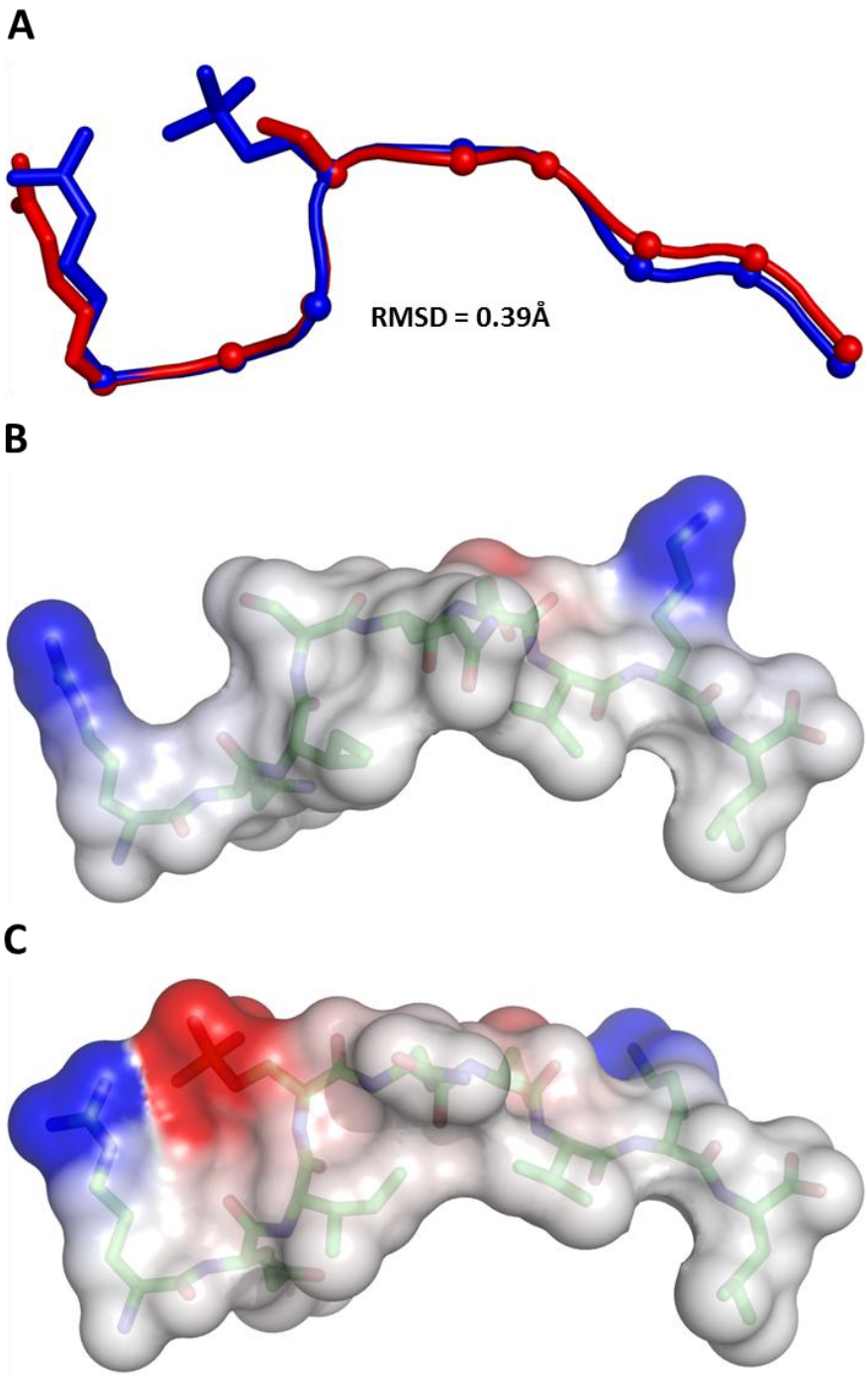
HLA-A2-peptide contacts and intra-peptide contacts (in brackets) between 2.3Å and 4.5Å in distance were calculated using CONTACT.

Figure 4.14 Differences in RQA_Vnp epitope interactions compared to RQA_Vp



The altered conformation of the RQA_Vnp epitope permits novel contacts to HLA-A2 that are absent in the phosphorylated counterpart. A, the P5 Ile side chain of RQA_Vnp forms multiple hydrophobic contacts with residues His70, Arg97 and Tyr99 of HLA-A2. B, The side chain of Tyr159 is positioned closer to the RQA_Vnp epitope and mediates a hydrogen bonding interaction with the backbone carbonyl group of P2 Gln. C, demonstrates the close proximity of P5 Ile and P10 Met therefore allowing intra-peptide hydrophobic interactions between the two residues.

Figure 4.15 Comparison of RQIpS and RQISnp epitope structure and surface electrostatics

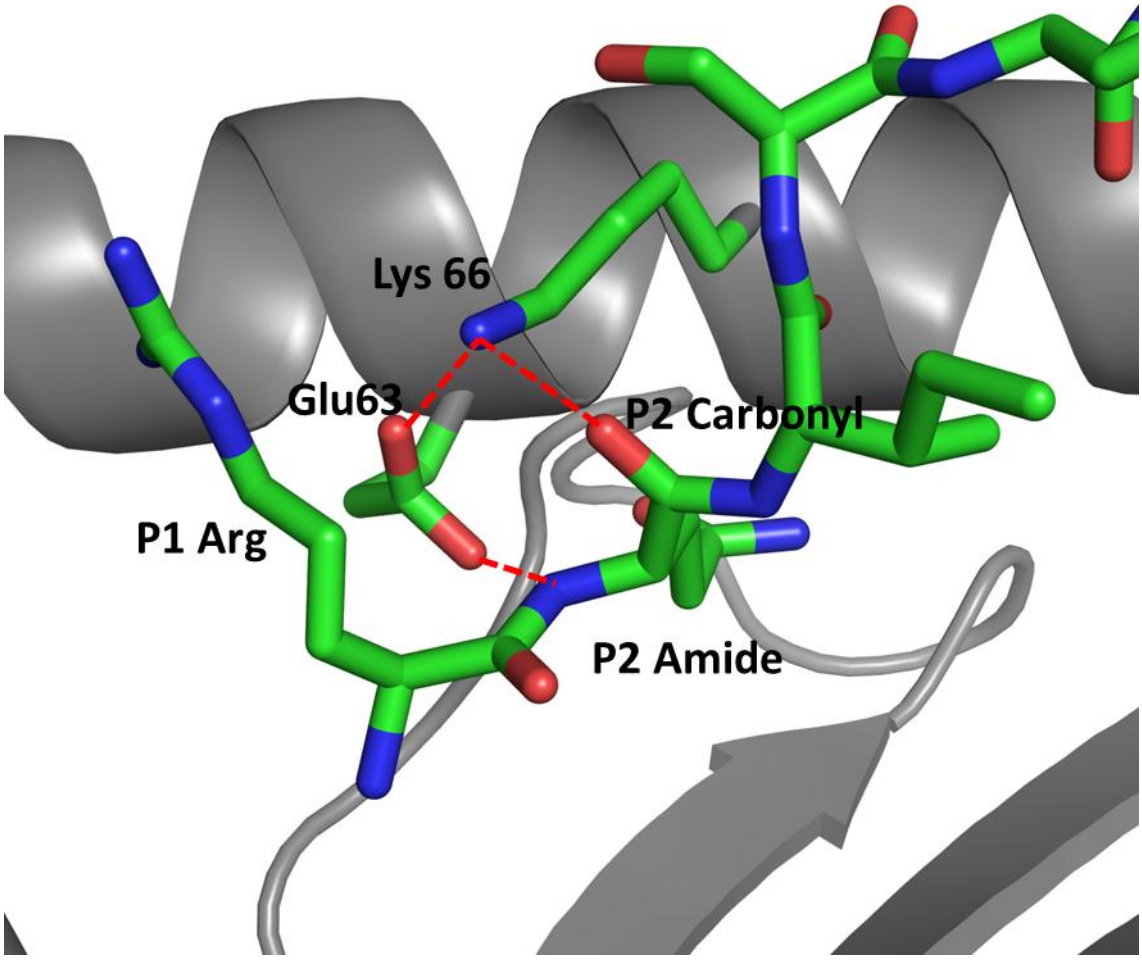


Comparison of the RQISnp (red) and RQIpS (blue) epitope structure reveals very little difference in antigen conformation (A). This results in the RQISnp epitope (B) possessing a similar molecular surface structure to the RQIpS epitope (C), the only difference being the presence of the phosphate, therefore altering surface electrostatics (blue, positive; grey, neutral; red, negative).

provided by the phosphate moiety. The molecular surfaces available for TCR binding of both phosphorylated and non-phosphorylated epitopes are therefore very similar (Figure 4.15 B/C), with the only difference being the presence or absence of the phosphate moiety. Two reasons underlie these observations: Firstly, in the RQISnp epitope, Glu63 and Lys66 interact with one another to form a hydrogen bond network with the backbone amide and carbonyl of Gln2 (Figure 4.16), thus forming a 'molecular clamp' which may act to stabilise and restrict the conformation within this region of the epitope. These interactions are conserved in canonical phosphopeptide structures but unlike in RQISnp they are disrupted in the non-phosphorylated versions of PKD2 and RQA_V which undergo altered antigen conformation. Secondly, in the majority of canonical phosphopeptides the phosphate contacts restrain the main chain conformation between position 4 and 5 (as shown in Chapter 3). In the non-phosphorylated forms of PKD2 and RQA_V these contacts are lost and the position of P4 serine lowers into the antigen binding groove, however for RQISnp the elevation in the main chain is maintained by His70, which can adopt two conformations one of which protrudes from the base of the α1 helix binding platform and re-orientates relative to its position adopted in the phosphorylated structure (Figure 4.17). This enables stabilising interactions with the backbone carbonyl of P3 and P5 (Figure 4.17) and allows RQISnp to naturally mimic the main chain conformation of the phosphorylated peptide.

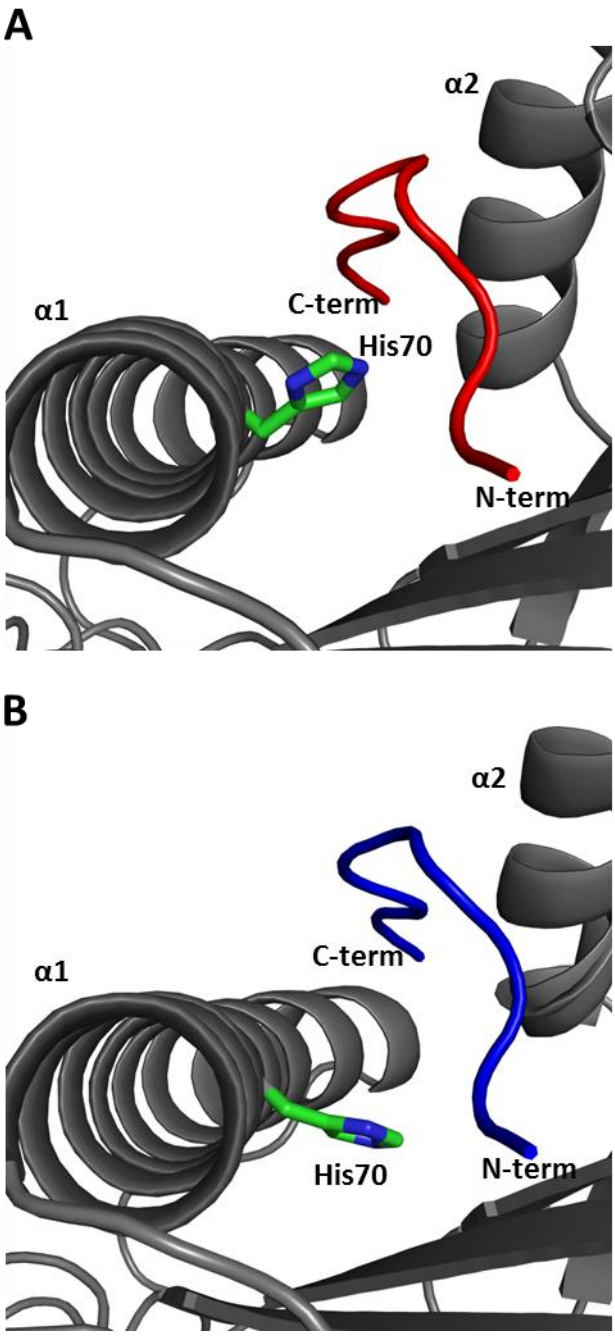
The RTFS epitope also revealed overall similarity of the MHC main chain upon comparison of the phosphorylated and non-phosphorylated structures and also showed very little difference in antigen main chain conformation, as demonstrated by a very low rmsd value of 0.07Å (Figure 4.18 A). As with RQIS the molecular surface of both the phosphorylated and non-phosphorylated epitopes are very similar (Figure 4.18 B/C).

Figure 4.16 RQISnp interactions with Glu63 and Lys66



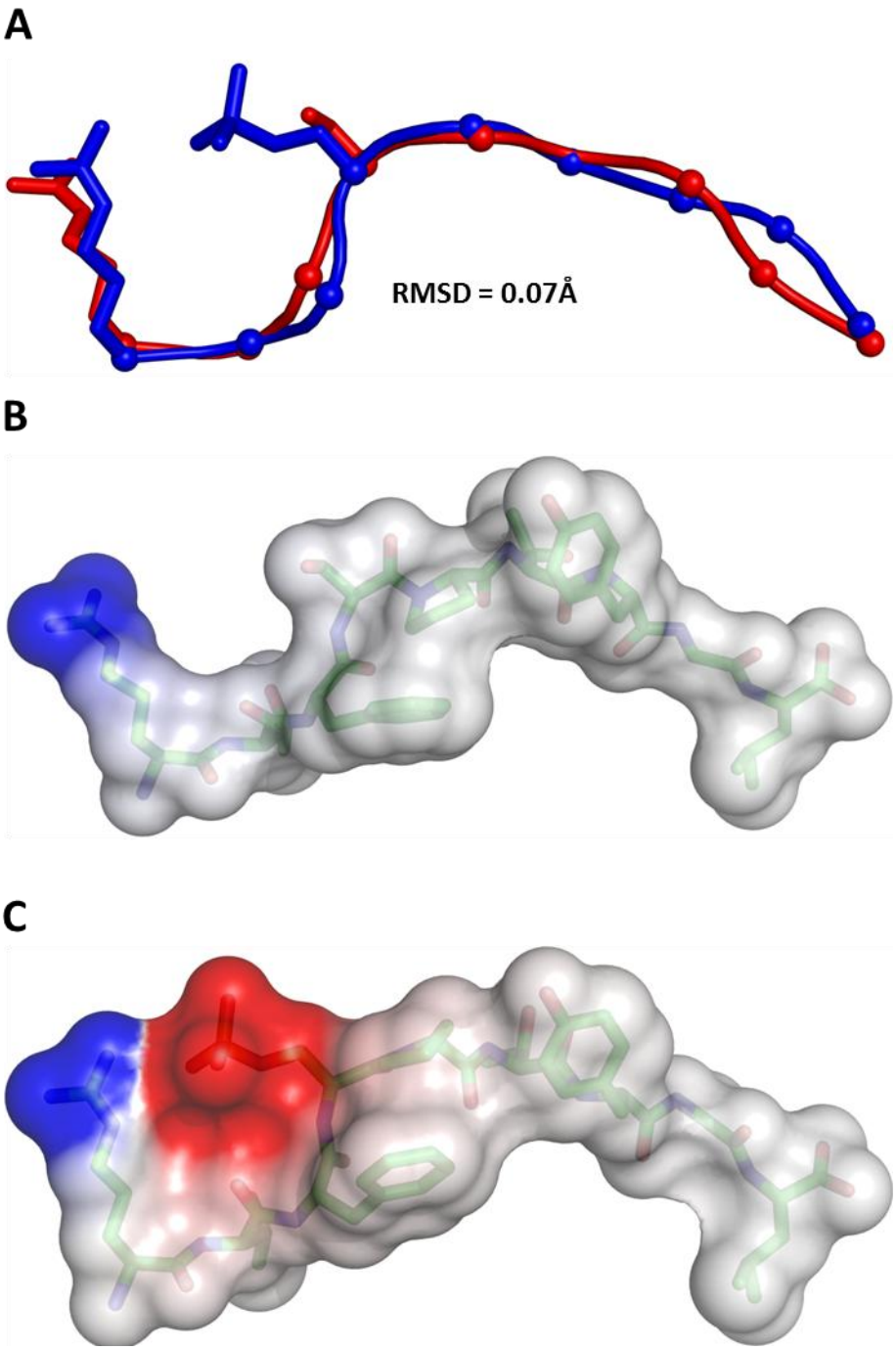
The P2 Carbonyl and amide groups of the RQISnp epitope form a network of hydrogen bonding interactions with the HLA-A2 side chains of Lys66 and Glu63. These interactions are commonly found in canonical phosphopeptide antigens and are absent in the non-phosphorylated forms of the PKD2 epitope and RQA_V epitope. In the RQISnp epitope these interactions act as a “mechanical clamp”, restricting the conformation of this region of the peptide.

Figure 4.17 RQISnp elevation of main chain by His70



The side chain of His70 in the RQISnp complex (A) is able to adopt an additional raised orientation in comparison to the RQIpS complex (B) thereby contributing to the maintenance of a raised conformation in the non-phosphorylated RQIS epitope. HLA-A2 heavy chain shown in grey with the side chain of His 70 depicted in stick representation. The main chain of RQISnp and RQIpS epitopes are shown in red and blue cartoon representation respectively.

Figure 4.18 Comparison of RTFSnp and RTFpS antigen structure and surface electrostatics

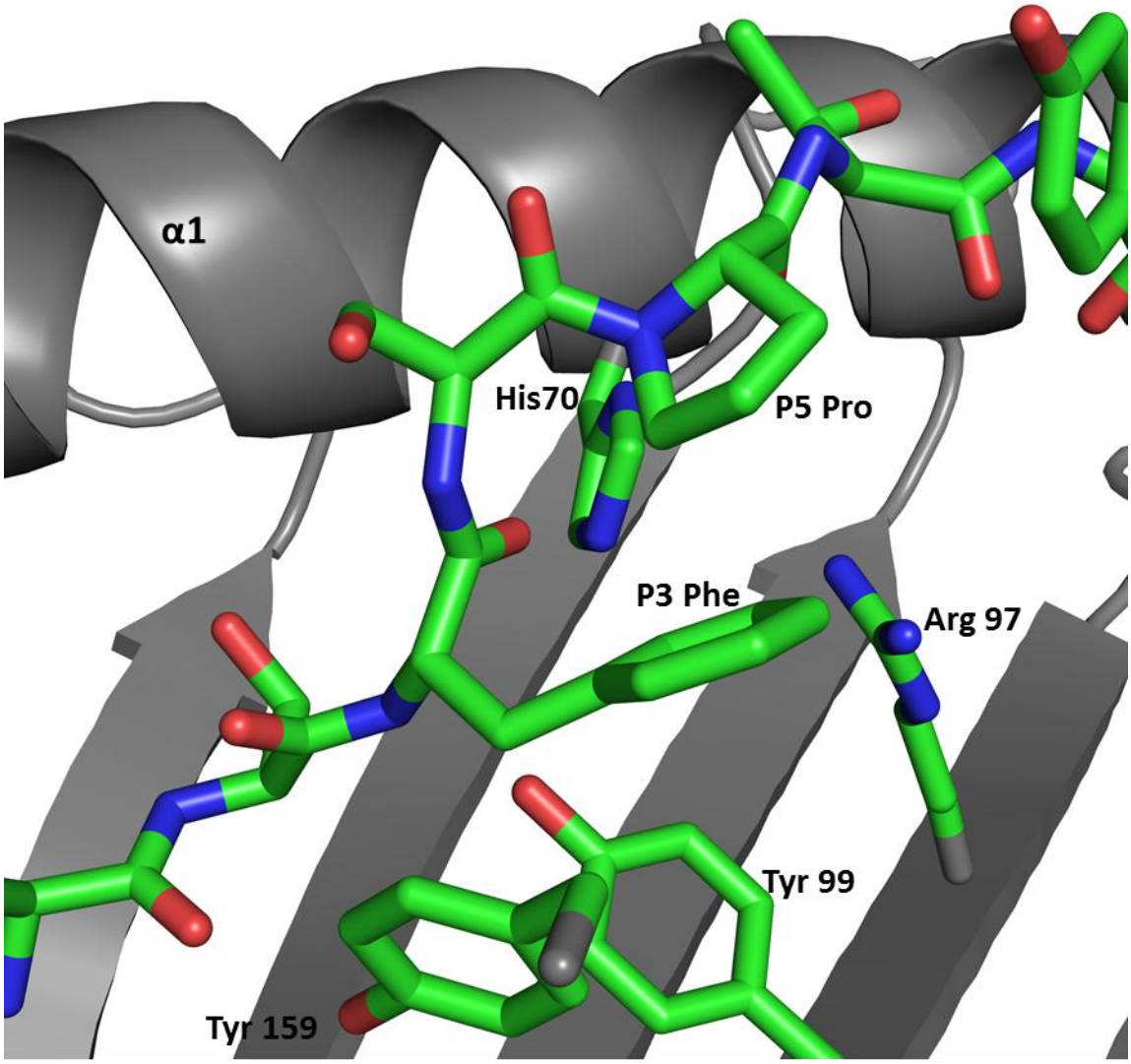


Comparison of RTFpS (blue) and RTFSnp (red) epitope conformation (A) reveals very little difference. This results in very similar surface structures for the RTFSnp (B) and RTFpS (C) epitopes with the only difference being the presence of the phosphate moiety which introduces differences in surface electrostatics (blue, positive; grey, neutral; red, negative).

The maintenance of main chain elevation in RTFSnp may be due to several factors. Firstly, RTFS contains a proline residue at position 5 of the epitope, which has been shown in other epitopes to restrict epitope conformation in this region (Petersen, Wurzbacher et al. 2009) due to its limited flexibility. However comparison of the RTFpS epitope structure with that of the P5 Pro containing epitope, RVApS demonstrated by Petersen *etal* to have limited flexibility, shows that while the RVApS epitope adopts an altered lower conformation compared to other canonical phosphopeptides the RTFpS epitope does not and adopts a position similar to other canonical phosphopeptides (as demonstrated in Figure 4.9), including those without P5 proline. Secondly, similar to RQISnp, the RTFSnp epitope maintains interactions with Glu63 and Lys66 through a hydrogen bond network with the backbone amide and carbonyl of Gln2 (as seen in Figure 4.16 for RQISnp) and also demonstrates a raised positioning of His70 which may contribute to the maintained raised position of the non-phosphorylated epitope (Figure 4.19). Finally, the presence of a relatively large Phe side chain at position 3 of the RTF epitope causes a restriction in the ability of the P3-P5 main chain region to adopt a lower conformation upon removal of the phosphate but also most notably forms hydrophobic contacts with P5 Pro further restraining the conformation around the P3-P5 region of the epitope (Figure 4.19).

These results may also explain the 12.6-fold increase in binding affinity upon phosphorylation of the RTFS epitope (See Chapter 3, Table 3.5) as unlike RQApS_V, which loses several interactions upon antigen phosphorylation due to conformational change, RTFpS maintains interactions observed in RTFSnp but also gains additional phosphate mediated interactions with surrounding residues (See Chapter 3, Figure 3.12B). In addition, an alteration in P1 Arg side chain orientation upon phosphorylation allows direct hydrogen

Figure 4.19 Maintenance of the raised conformation in RTFnp by P3 Phe



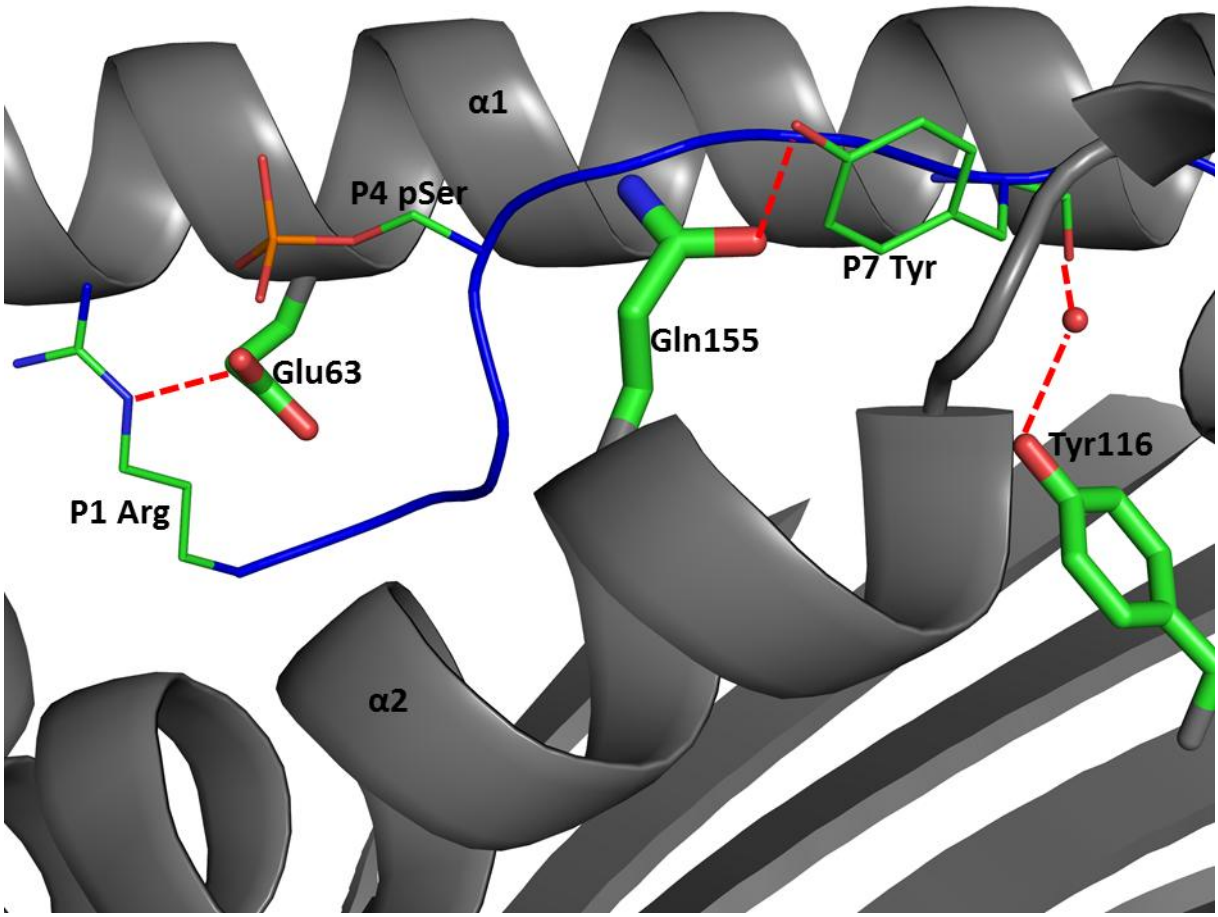
His70 of the MHC α1 helix is able to adopt a raised conformation which contributes to the elevation of the RTFSnp epitope main chain. The presence of the bulky side chain of Phe at P3 inhibits the lowering of the main chain and can adopt a conformation that allows hydrophobic interaction with the P5 Pro residue which acts to further restrict the conformation of the peptide in the P3-P5 region of the RTFnp epitope, thus enabling the RTFnp epitope to maintain the raised conformation observed in the RTFpS epitope. The α1α2 domains of HLA-A2 are shown in grey. The RTFnp epitope and His70, Arg97, Tyr99 and Tyr159 side chains of the HLA-A2 peptide binding groove are shown in stick representation.

bonding between the P1 Arg side chain and the carboxyl group of Glu63 (Figure 4.20). RTFpS also forms novel direct interactions between the hydroxyl group of P7 Tyr and the carboxyl group of Gln155, along with indirect water mediated contacts between the hydroxyl group of Tyr116 and the main chain carbonyl of P7 Tyr (Figure 4.20). These additional peptide-MHC contacts observed in RTFpS account for the substantial increase in epitope affinity compared to RTFSnp.

4.9 Conclusions

These results show that antigen phosphorylation can radically alter antigen conformation and epitope identity and demonstrate the molecular mechanisms which govern this. The epitopes PKD2 and RQAS_V which both undergo conformational rearrangement upon phosphorylation do so by a similar structural mechanism, whereby the P4 pSer induces the elevation of the surrounding main chain relative to the non-phosphorylated state, thereby allowing the phosphate moiety to satiate its bonding requirements and participate in stabilising interactions with surrounding residues of the peptide and MHC. Despite this similarity in phosphate binding and phosphate induced conformational rearrangement, the consequences of antigen phosphorylation on peptide-MHC affinity is clearly highly epitope dependent, and highlights the benefit of parallel structural and peptide-MHC affinity analysis to fully understand the effects of phosphorylation for individual epitopes. For PKD2 the conformational rearrangement resulted in the formation of additional phosphate mediated contacts without significant loss of contacts observed in the non-phosphorylated epitope and thus resulted in a net increase in epitope binding affinity upon phosphorylation. Conversely the conformational rearrangement of RQApS_V upon phosphorylation causes a loss of interactions in the central region of the epitope which act to balance out the

Figure 4.20 RTFpS peptide MHC contacts



Additional contacts between RTFpS epitope and HLA-A2, other than phosphate mediated contacts, which are not found in RTFnp. The main chain of the RTFpS epitope is shown in blue ribbon representation with side chains of P1 Arg, P4 pSer and P7 Tyr shown in line representation with Carbon (green), Oxygen (red) and Nitrogen (blue). The HLA-A2 heavy chain residues Glu63, Tyr116 and Gln155 are shown in stick representation with Carbon (green), Oxygen (red), Nitrogen (blue) and the main chain in grey. Red sphere represents a water molecule and relevant hydrogen bonds are indicated by red dashed line.

additional contacts introduced by the phosphate which results in no net gain in antigen binding. The effect of the conformational change both these epitopes undergo upon phosphorylation also highlights that the most drastic conformational rearrangement occurs within the central region of the epitope, which has been highlighted as a key region for antigen recognition and discrimination by TCRs.

However, the study also highlights two antigens (RQIS and RTFS) which do not alter in conformation upon phosphorylation and are consistent with a previously described canonical phosphopeptide RVAS which did not undergo conformational rearrangement upon phosphorylation (Petersen, Wurzbacher et al. 2009). This study has demonstrated that this can occur through a combination of mechanisms on an epitope specific basis and allow the non-phosphorylated epitopes to naturally mimic the conformation of the phosphorylated forms.

It can therefore be concluded that there exists two types of phosphopeptide antigen; those which undergo conformational change upon phosphate addition and therefore represent completely novel antigens for T-cell discrimination and those which do not undergo conformational rearrangement and consequently discrimination between phosphopeptide/non-phosphopeptide counterparts would be to a large extent based upon direct TCR engagement of the phosphate moiety. This knowledge of how particular epitopes are presented upon phosphorylation may lead to informed decisions on optimal therapeutic strategies targeting phosphopeptide antigens, with the antigens exhibiting a more dramatically altered structure upon phosphorylation being the less likely to display cross-reactivity upon use in vaccination strategies. However, the molecular mechanisms which govern phosphopeptide recognition by phosphopeptide specific T-cell receptors have never

been directly studied. Therefore, knowledge of how these T-cell receptors discriminate between phosphorylated and non-phosphorylated epitopes is lacking and so these questions will form the basis of the following chapter.

5 Phosphopeptide antigen recognition by phosphopeptide specific T-cell receptors

5.1 Background

Having demonstrated in Chapter 4 that phosphorylation can affect antigen identity through two distinct mechanisms; either conformational change or phosphate presence alone, this chapter will focus on defining the molecular mechanisms that govern phosphopeptide recognition and discrimination by phosphopeptide specific T-cells. Interestingly, although epitopes which undergo conformational change may be more antigenically distinct than those which do not, previous studies have highlighted that antigens that exhibit minimal structural rearrangement (Petersen, Wurzbacher et al. 2009) can also be discriminated and targeted by phosphopeptide specific T-cell clones (Zarling, Ficarro et al. 2000; Zarling, Polefrone et al. 2006). These T-cell clones, which have been highlighted to target and eliminate phosphopeptide-expressing tumour cells in a phosphopeptide-specific manner, were obtained from Dr Angela Zarling (University of Viriginia) and used to study the molecular basis of TCR phosphopeptide recognition and discrimination.

5.2 Production of soluble T-cell receptors

5.2.1 Expression vector cloning

Plasmids containing either the α or β chain genes from the three separate phosphopeptide specific T-cell clones; IRS2 (specific for the IRS2 epitope RVApS), IRS2BK which is also specific

for the IRS2 epitope RVApS but showed higher avidity than the previous clone in T-cell killing assays (Zarling *et al* unpublished data) and CDC25B (specific for the CDC25B epitope GLLGpS) were kindly donated by Dr Angela Zarling. Each chain was amplified from the plasmids using a pair of chain specific primers (Table 5.1 and Figure 5.1) which incorporated the restriction enzyme sites *BglII* and *XbaI* for subsequent cloning into pMTBipV5-His expression vectors, enabling expression in the Drosophila expression system. The sequence for each cloned T-cell receptor (TCR) chain was verified by plasmid sequencing performed by the Functional Genomics facility (University of Birmingham) and verification of correct sequence was performed using Sequencher® (Sequencher v.4.5, Genecodes) (See Appendix H-N for individual sequences).

5.2.2 Soluble T-cell receptor expression

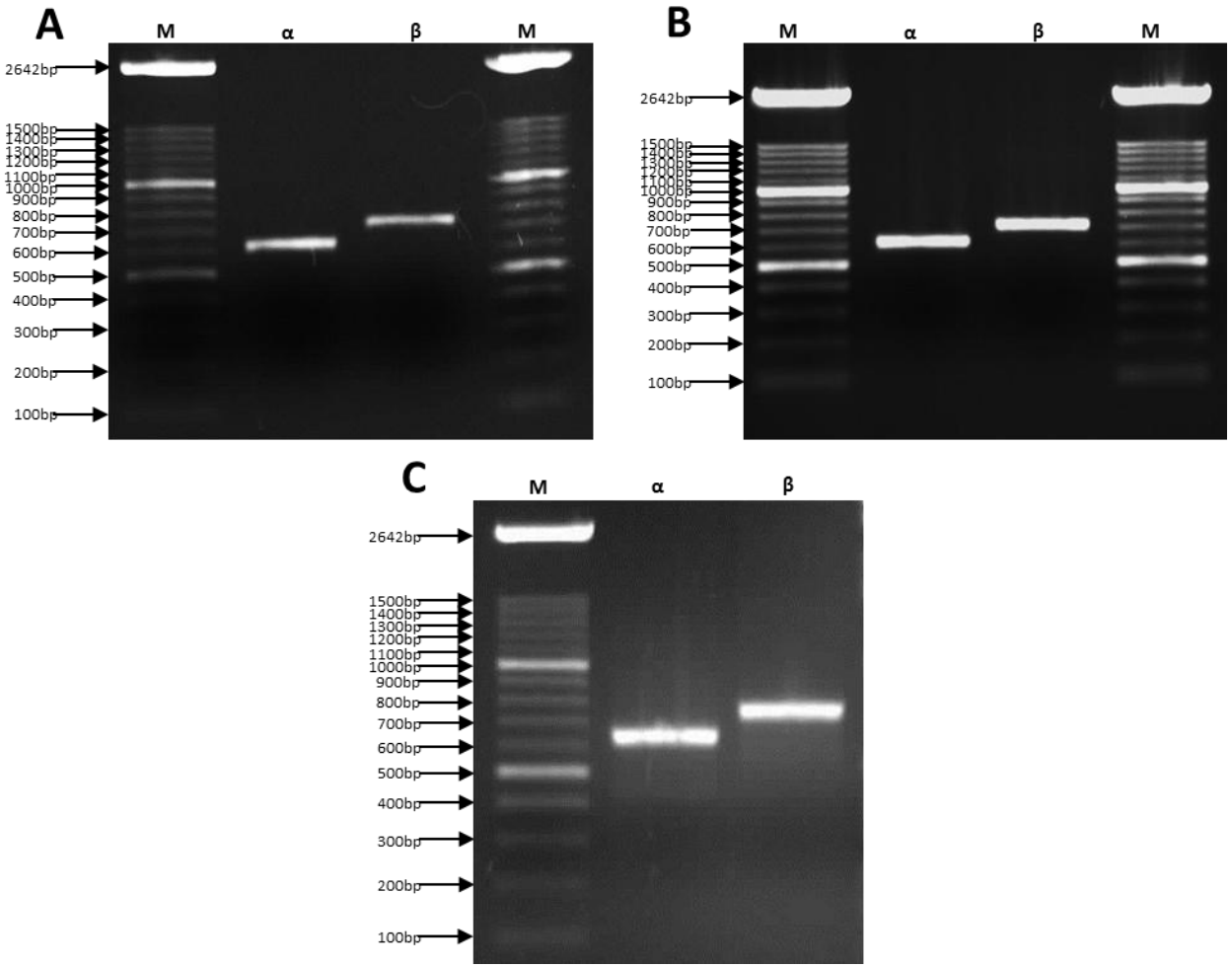
Following successful cloning of each TCR α and β chain into pMTBipV5-His expression vectors containing acid and base zipper constructs respectively (Chang, Bao et al. 1994; Adams, Chien et al. 2005; Willcox, Pitard et al. 2012), these constructs were used for co-transfection into insect S2 cells along with a plasmid conferring Hygromycin resistance. Transfected S2 cells were cultured in media containing Hygromycin as a selection agent. Following several weeks of selection the transfected S2 cells were induced with 500μM $CuSO_4$ and the expression of soluble TCR assessed by Western blot (Figure 5.2) of the S2 supernatant. Each TCR showed good levels of expression by day 5 induction of both the α-acid zipper chain and β-base zipper chain, however each seems to exhibit a higher molecular weight than the predicted 31KDa and 33KDa respectively, this is most likely due to glycosylation of the TCR chains in the Drosophilla expression system. Correct hetrodimer conformation was verified for each clone by dot blot analysis with an acid/base zipper

Table 5.1 TCR Specific primers

TCR	Chain	Primer	Sequence	T_m (°C)
IRS2	α	Forward	5'-GGA *AGA TCT* GCT CAG TCA GTG ACA CAG CC-3'	64
		Reverse	5'-GGA *TCT AGA* ACA GGG AAC GTC TGA ACT GGG-3'	66
	β	Forward	5'-GGA *AGA TCT* GAA GCT GGA GTC ACC CAG TC-3'	64
		Reverse	5'-GGA *TCT AGA* ACA GTC TGC TCG GCC CCA G-3'	64
IRS2BK	α	Forward	5'-GGA *AGA TCT* GCT CAG TCA GTG ACA CAG CCC GAT G-3'	78
		Reverse	5'-GGA *TCT AGA* ACA GGG AAC GTC TGA ACT GGG G-3'	70
	β	Forward	5'-GGA *AGA TCT* GAG GCT GCA GTC ACC CAA AGT C-3'	66
		Reverse	5'-GGA *TCT AGA* ACA GTC TGC TCG GCC CCA GGC-3'	72
CDC25B	α	Forward	5'-GGA *AGA TCT* GGC GAG CAG GTG GAG CAG-3'	62
		Reverse	5'-GGA *TCT AGA* ACA GGG AAC GTC TGA ACT GG-3'	62
	β	Forward	5'-GGA *AGA TCT* GCA GTC ACC CAA AGC CCA AG-3'	64
		Reverse	5'-GGA *TCT AGA* ACA GTC TGC TCG GCC CCA G-3'	64

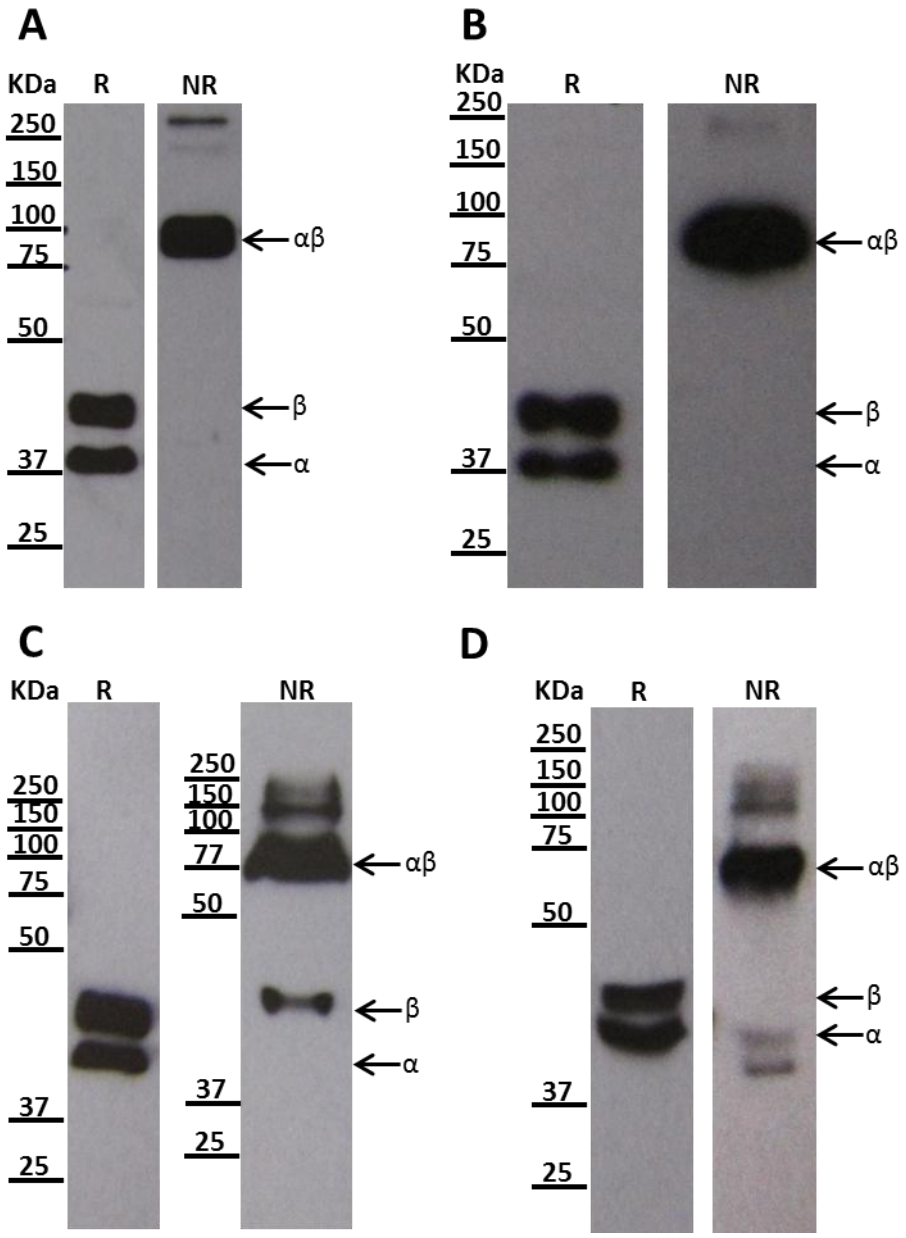
This table documents the primer sequences used for the cloning of phosphopeptide specific TCR α/β chain sequences into pMTBipV5-His expression vectors. Letters in italics represent positions within the primer which introduce *BglII* and *XbaI* restriction sites into the PCR product to enable ligation into pMTBipV5-His expression vectors.

Figure 5.1 PCR amplification of α and β TCR chains



TCR chains (α and β) for IRS2 (A) CDC25B (B) and IRS2BK (C) phosphopeptide-specific T-cells were amplified using chain-specific primers which introduced a *BglII* and *XbaI* restriction site, enabling ligation into pMTBipV5-His expression vectors.

Figure 5.2 Western Blot analysis of soluble TCR expression in transfected S2 insect cell supernatant following induction



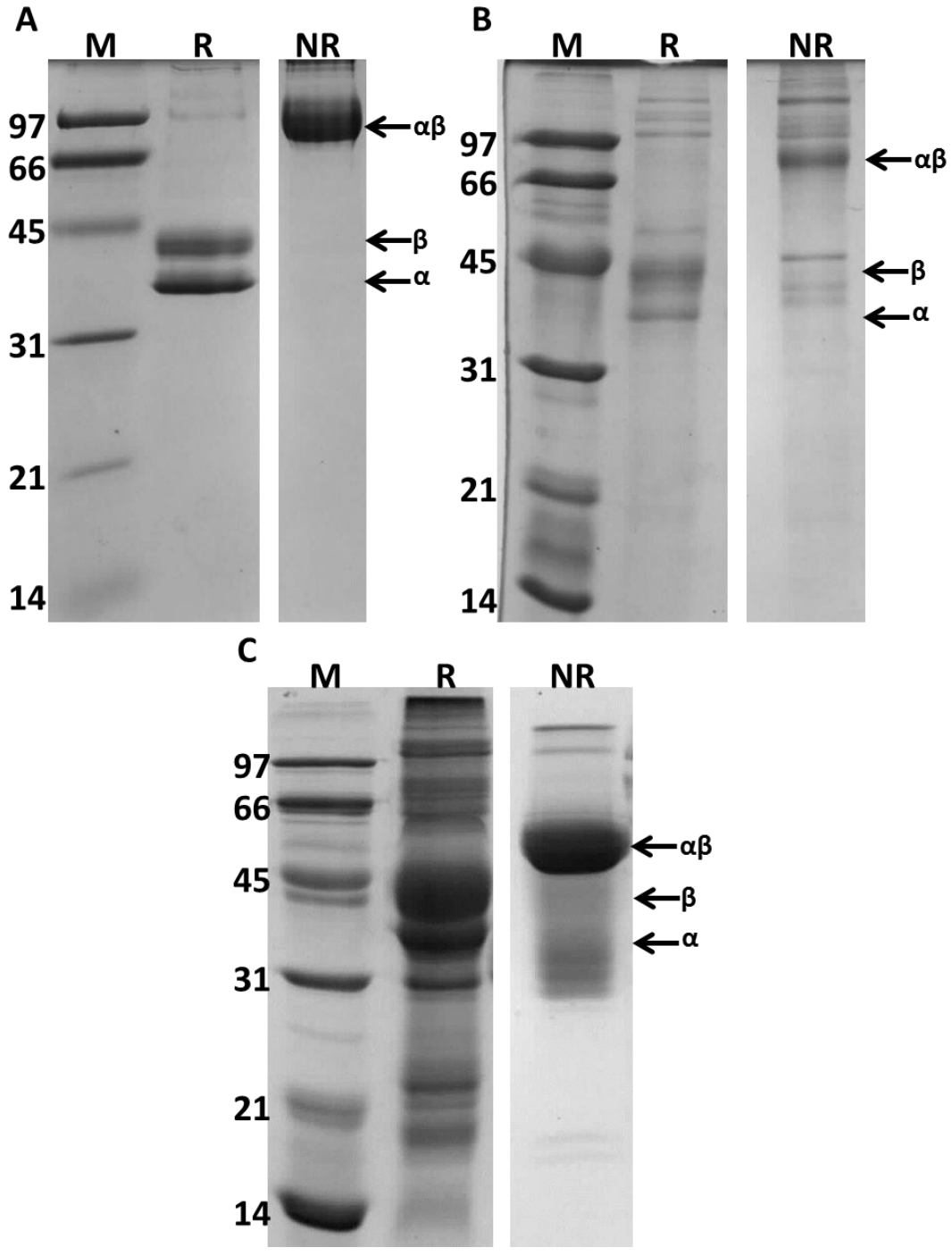
Western blot analysis of transfected S2 cell IRS2 TCR (A), IRS2BK TCR (B), CDC25B TCR (C) and CDC25BBirA (D) production following 5 days of expression with an anti-His antibody, reveals suitable production of both α (lower band) and β (upper band) chains for each TCR when assessed in reducing conditions (Lane R) and that these form a dimer when assessed under non-reducing conditions (Lane NR).

specific antibody (2H11) (Chang, Bao et al. 1994) (Figure 5.3). Following confirmation of TCR expression, large scale inductions of 1-2L were performed. His-tagged TCR-zippers were isolated from the resulting supernatant by binding to a Ni-NTA bead column followed by washing in PBS 10mM imidazole and elution of the bound TCR with PBS 250mM imidazole. The yield of each TCR varied somewhat between preparations but typically reached levels of 4-6mg/Litre of S2 supernatant. The purity of each TCR following Ni-NTA purification was assessed by SDS-PAGE (Figure 5.4 A-C). Both α and β chains were visualised for each TCR clone under reducing conditions. However, as noted in western blot analysis, the bands for each TCR chain were slightly higher than the sequence derived molecular weight of the α (31 kDa including the acid zipper) and β chains (33 kDa including the base zipper). As mentioned above the most likely explanation for the mobility shift on the SDS-PAGE is that the Drosophila derived TCRs are glycosylated to a significant extent (Figure 5.4 A-C). This was also the case for the heterodimer TCR observed under non-reducing conditions, which migrated higher than the expected sequenced-derived molecular weight (64 kDa with acid/base zipper moieties) for each clone (Figure 5.4 A-C). SDS-PAGE analysis also reveals that the different TCR clones obtained different levels of purity following Ni-NTA bead purification, with IRS2 exhibiting the highest purity (Figure 5.4A) and IRS2BK the lowest (Figure 5.4C). Following Ni-NTA bead purification each TCR was used in further studies which will now be discussed on a TCR by TCR basis.

Figure 5.3 Heterodimer dot blot analysis of TCR complexes

Heterodimer dot blot analysis using an acid/base zipper specific antibody (2H11) confirms the presence of TCR heterodimer in IRS2 TCR (3), IRS2BK TCR (4), CDC25B TCR (5) and CDC25B BirA TCR (6). Lane 1 and 2 are negative controls consisting of pure IRS2 α (1) and β (2) chains which are not detected by the heterodimer-specific antibody.

Figure 5.4 SDS-PAGE analysis of Ni-NTA bead purified TCRs

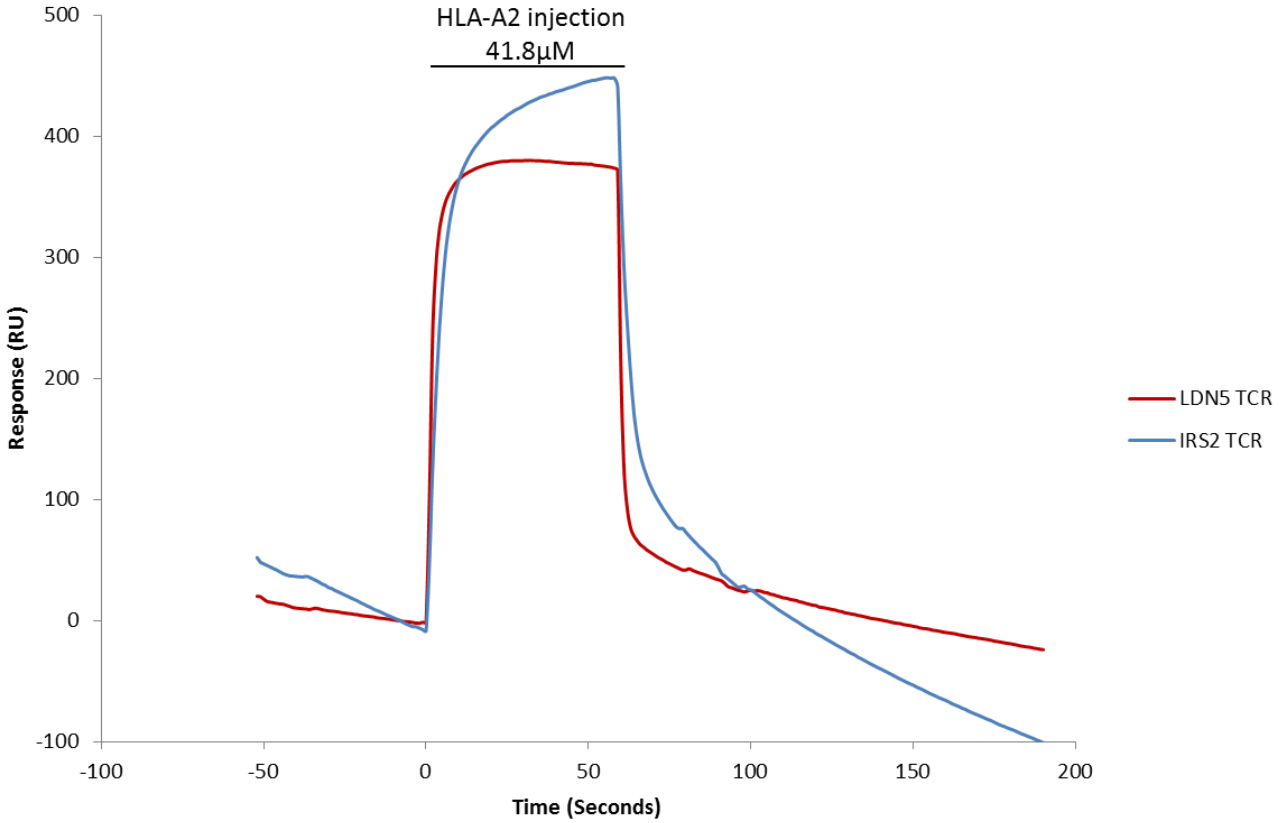


Following Ni-NTA bead purification soluble IRS2 (A), CDC25B (B) and IRS2BK (C) were assessed by 12.5% SDS-PAGE analysis under reducing (Lane R) or non-reducing (Lane NR) conditions. All samples were compared to a molecular weight marker (Lane M) so as to assess purity of each TCR preparation. The α and β bands were assigned based on single chain transfectants.

5.3 IRS2 TCR binding studies

The IRS2 TCR clone which recognises the RVApS epitope of insulin receptor substrate 2 was produced in high quantity with reasonable purity following Ni-NTA bead purification (Figure 5.4 A). This was used for a preliminary surface plasmon resonance (SPR) binding study in which 5118 Response units (RU) IRS2 TCR was immobilised onto a flow cell of an Ni-NTA chip with 2668RU of a negative control αβ TCR (LDN5) immobilised onto a separate flow cell. To assess specific binding purified soluble HLA-A2-RVApS complexes were injected over each flow cell. The resulting sensogram showed very similar increases in response for IRS2 TCR relative to the negative control LDN5 TCR even at very high concentrations of HLA-A2-phosphopeptide complex (Figure 5.5). This result was extremely surprising given that T-cells which express this TCR are known to target and kill cells expressing HLA-A2-RVApS complexes. There are several explanations as to why the TCR did not display any binding. Firstly, the strength of TCR binding to the Ni-NTA chip is poor as demonstrated by the significant reduction in levels of protein immobilised over-time, suggesting on-going TCR dissociation from the surface. Secondly, the purity of the TCR sample used for immobilisation is likely to play a significant role in the quality of binding observed. This may also have a greater impact in the Ni-NTA chip setting as the impurities eluted from the Ni-NTA purification step are proteins which are also likely to exhibit a degree of (non-specific) binding to Ni-NTA chip and therefore the observed levels of protein binding to the Ni-NTA chip may result in a gross overestimation of the levels of bound TCR. To address these issues, a second series of SPR experiments were performed using a CM5 chip with either

Figure 5.5 Surface Plasmon resonance sensogram of Ni-NTA purified IRS2 TCR using a Ni-NTA chip



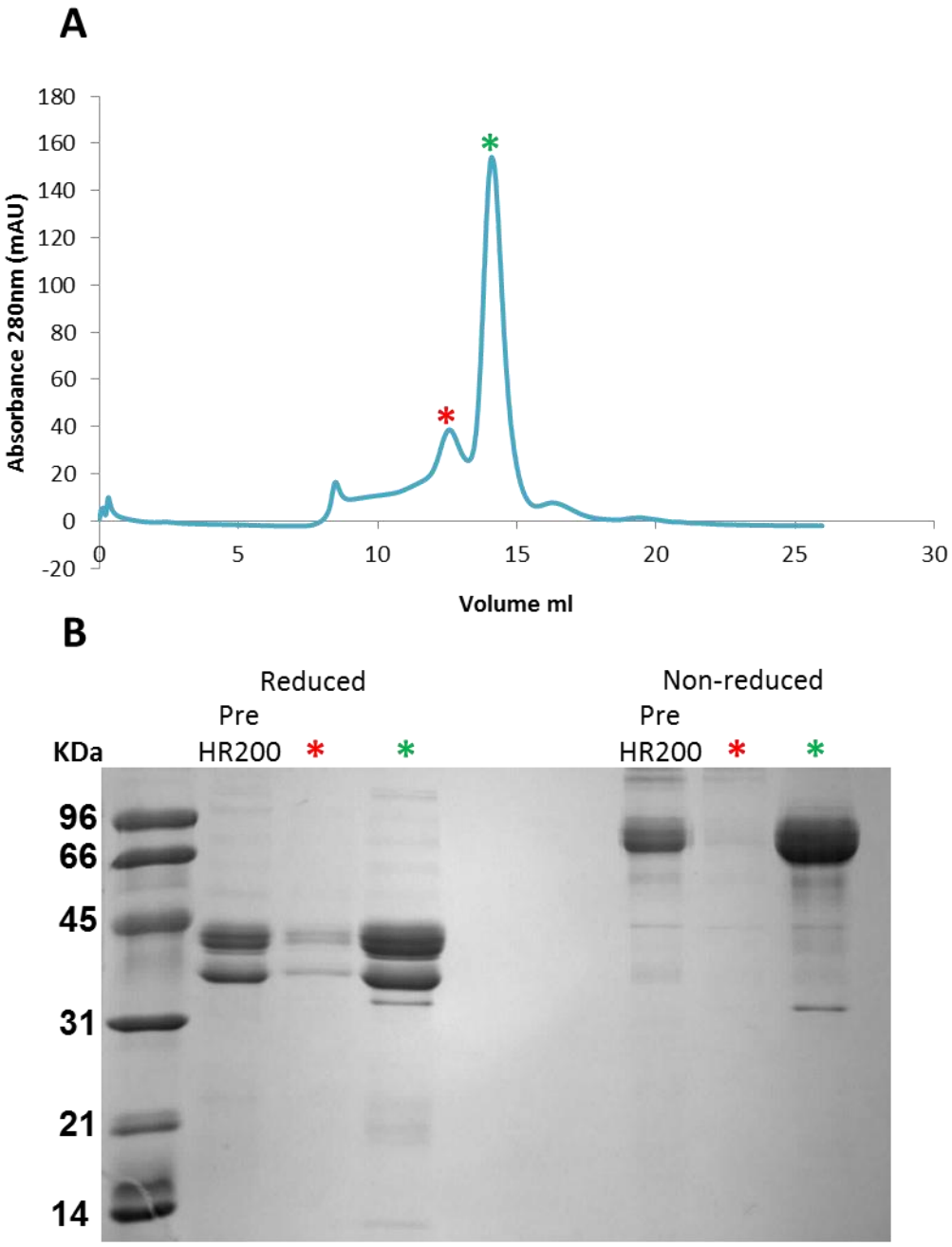
Following immobilisation of 5118RU NI-NTA purified IRS2 TCR (blue) or 2668RU LDN5 TCR (red) to separate flow cells of an Ni-NTA sensor chip via His tag-Ni coupling, 10µl of HLA-A2-RVApS, in solution at a concentration of 41.8µM, was injected over each flow cell using a flow rate of 10µl/min. The sensogram suggests a high degree of aggregate sticking in both flow cells with no credible increase in specific binding to the IRS2 TCR relative to the LDN5 negative control.

biotinylated HLA-A2-RVApS (5436RU) or an unrelated HLA-A2-WT1 (Willms tumour 1 antigen) epitope complex (5352RU) coupled with streptavidin to individual flow cells. Prior to injection, the IRS2 TCR was further purified by size exclusion chromatography using a Sephaadex HR-200 column (Figure 5.6). Despite improved purity of the IRS2-TCR, there was still no detectable specific binding of the IRS2 TCR above negative controls (Figure 5.7). These results indicate that the IRS2 TCR clone may recognise HLA-A2-RVApS with an affinity that is below the limit of detection by SPR. Furthermore, due to this low affinity, attempts to generate TCR/HLA-A2 complexes for X-ray crystallography would be extremely difficult.

5.4 IRS2BK T-cell binding studies

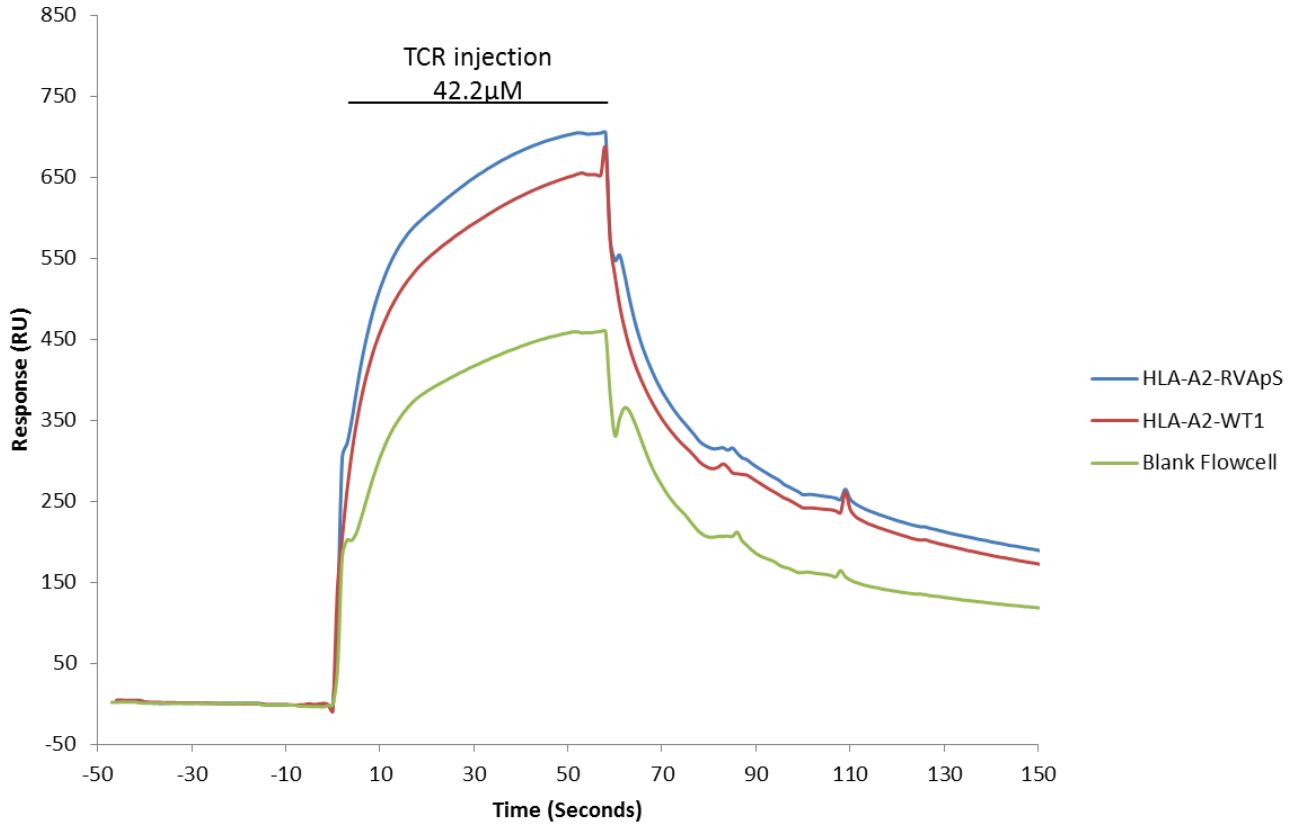
A second clone which recognises HLA-A2-RVApS with greater affinity in cellular assays than the previous IRS2 clone (Personal communication from Angela Zarling) was generated as described above. A preliminary SPR binding study was performed in which either biotinylated HLA-A2-RVApS (8780RU), HLA-A2-RVASnp (5539RU) or an unrelated HLA-A2-WT1 epitope control (6325RU) were coupled to individual flow cells of a CM5 sensor chip using streptavidin. Following MHC immobilisation, soluble IRS2BK TCR was concentrated to 7mg/ml, diluted to 4mg/ml in HBS-EP buffer and injected over the surface of each flow cell. The responses measured were marginally higher over HLA-A2-RVApS/HLA-A2-RVASnp than over the HLA-A2-WT1 negative control, however the signal for the non-phosphorylated version of RVAS was higher than that of the phosphopeptide, albeit by only a small amount (Figure 5.8). This is despite the fact that there was 8000 RU of HLA-A2-RVApS immobilised. These results were surprising as at the high concentration of TCR (62.5µM) specific binding to the phosphopeptide would be expected.

Figure 5.6 HR-200 gel filtration purification of IRS2 TCR



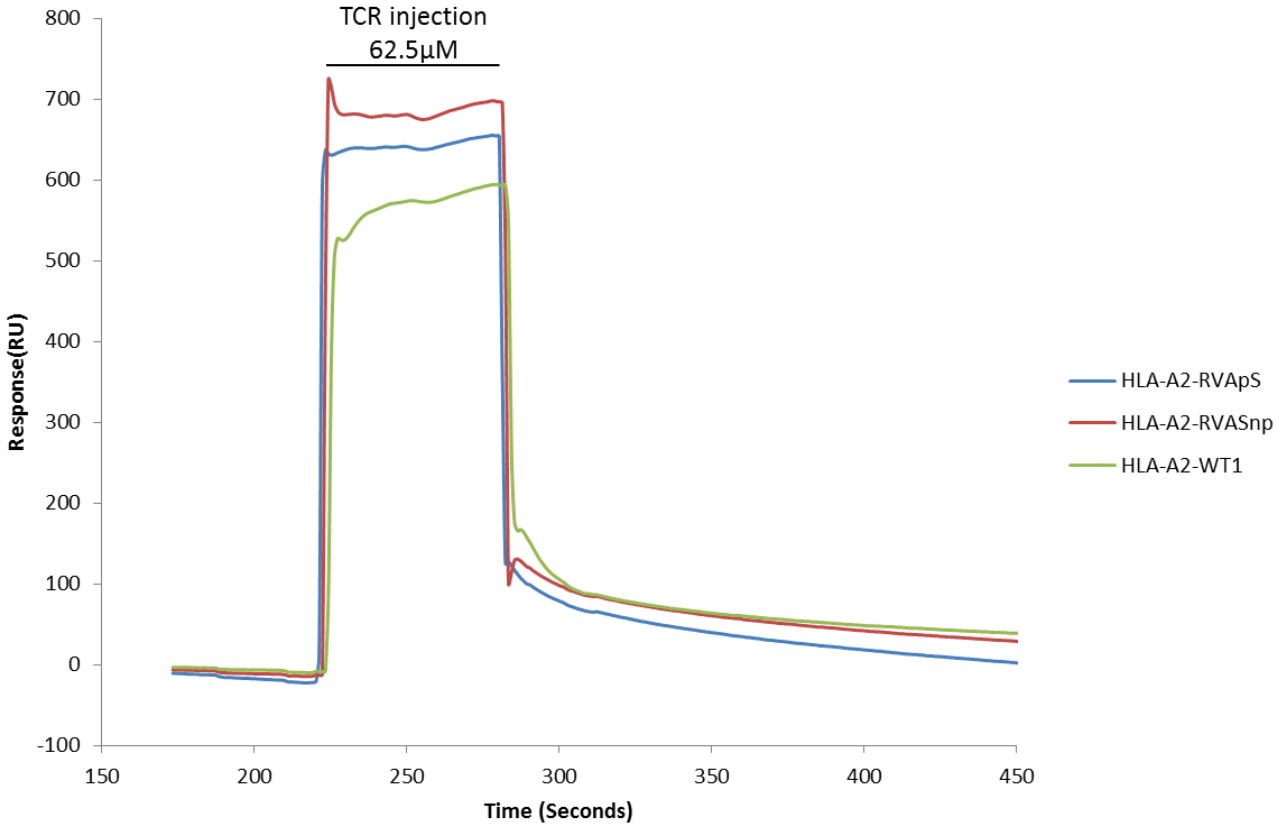
Purification of IRS2 TCR by HR200 gel filtration results in an elution profile with multiple peaks, including those indicated by red and green asterisk (A). Further analysis of these peaks by SDS-PAGE (B) reveals that although the first peak (red asterisk) contains some TCR α and β chains, the non-reduced analysis indicates the second peak (green asterisk) contains the majority of the TCR. This SDS-PAGE analysis also reveals that there are still small amounts of residual contaminants following HR-200 gel filtration.

Figure 5.7 Surface Plasmon resonance sensogram of HR-200 purified IRS2TCR using a CM5 chip



Following immobilisation of 5436RU HLA-A2-RVApS (blue) and 5352RU HLA-A2-WT1 negative control (red), to separate flow cells of a CM5 sensor chip via amine coupling, 10µl of HR200 purified IRS2 TCR, in solution at a concentration of 42.2µM, was injected over each flow cell (including one which contained no MHC complex, sown in green) using a flow rate of 10µl/min. The sensogram demonstrates essentially no difference in response between HLA-A2-RVApS and HLA-A2-WT negative control, consistent with no binding under these conditions.

Figure 5.8 Surface Plasmon resonance sensogram of HR-200 purified IRS2BK TCR using a CM5 chip



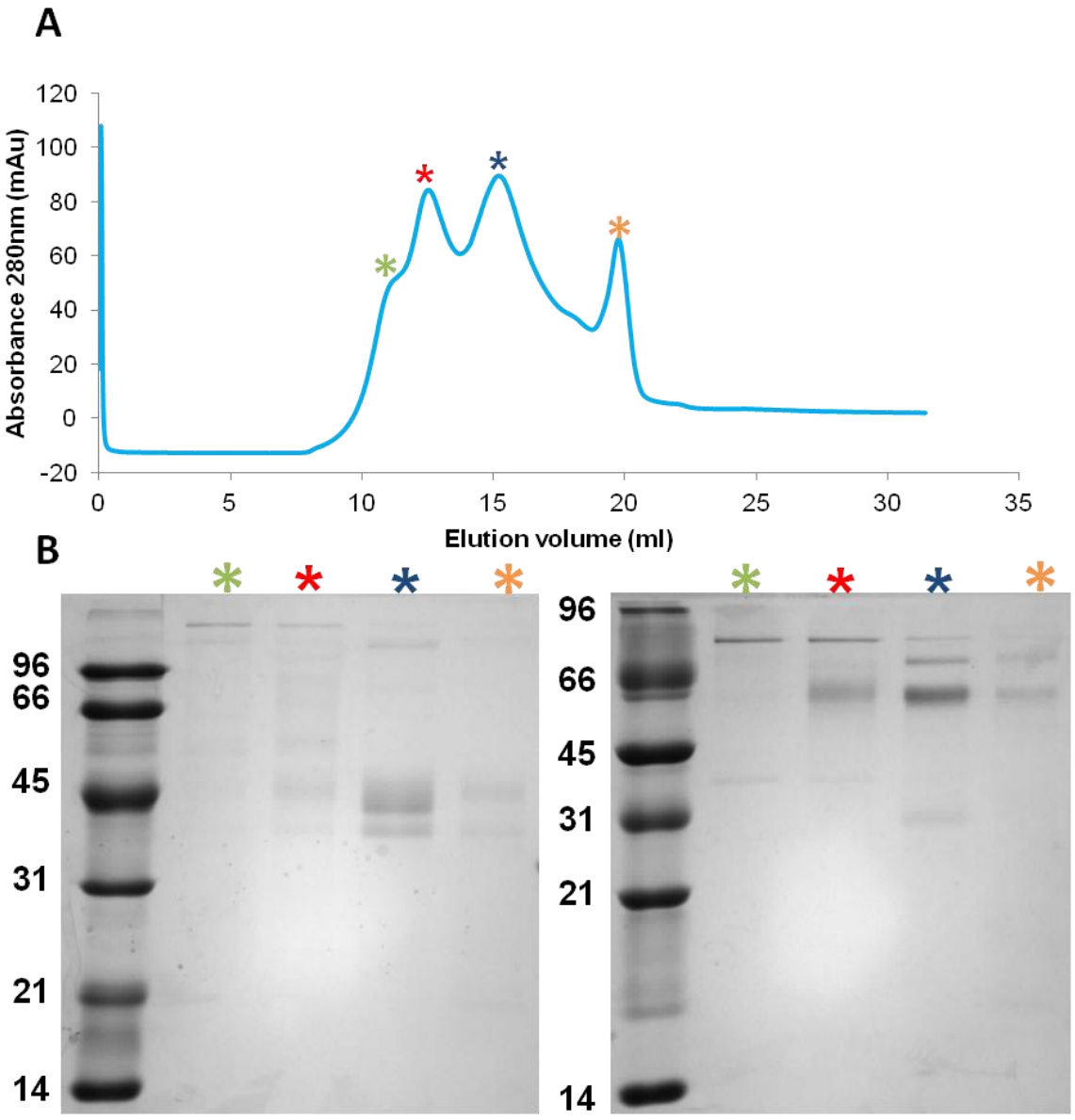
Following immobilisation of 8780RU HLA-A2-RVApS (blue), 5500RU HLA-A2-RVASnp (red) or 6300RU HLA-A2-WT1 negative control (green) to separate flow cells of a CM5 sensor chip via streptavidin coupling, 10µl of HR200 purified IRSBK TCR at a concentration of 62.5µM was injected over each flow cell at a flow rate of 10µl/min. Although there is an observable increase in binding over HLA-A2-WT1 the IRS2BK TCR does not have increased binding above the HLA-A2-RVASnp control.

This raises the question of whether these are genuine binding responses and alternatively of whether this receptor also exhibits a low binding affinity.

5.5 CDC25B T-cell clone binding studies

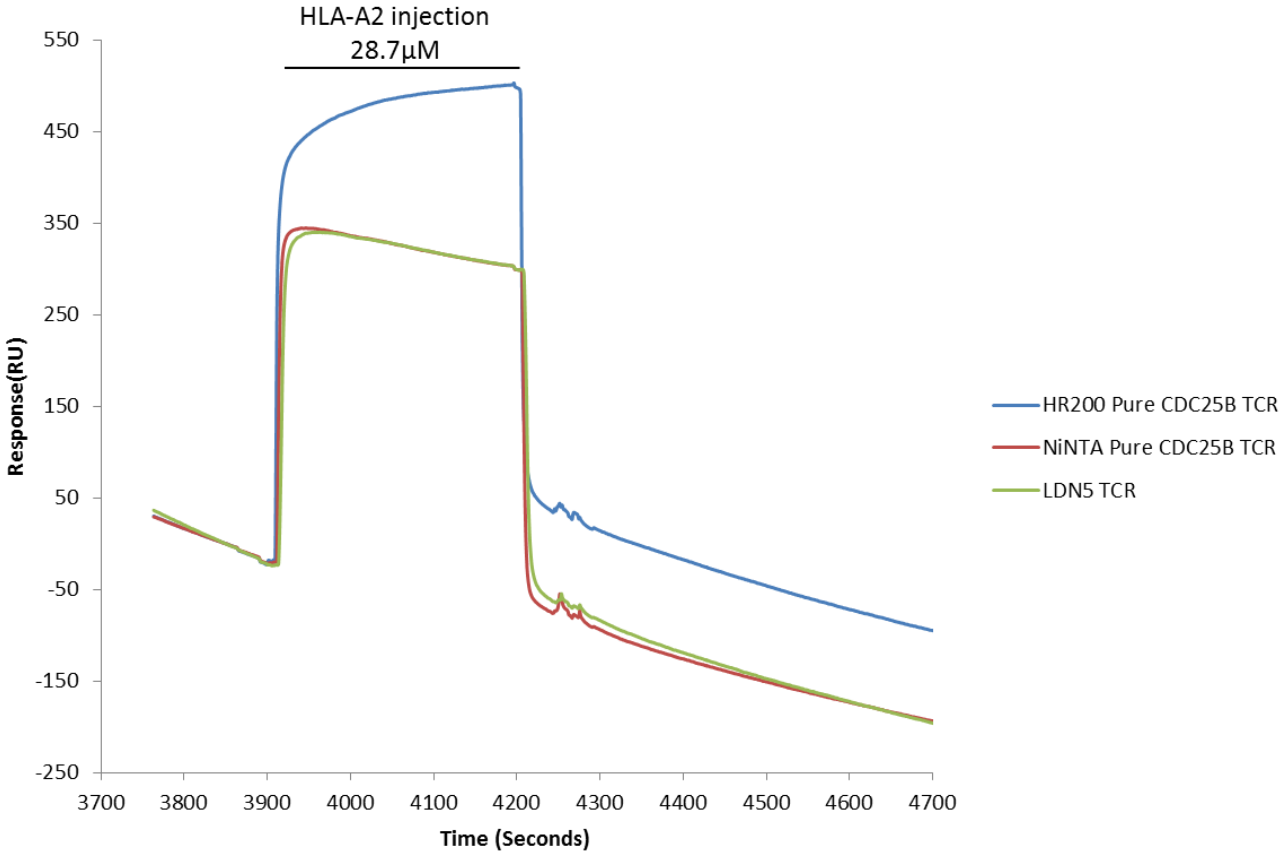
A third TCR which recognises HLA-A2-GLLGpS, an epitope of CDC25B, was produced as previously described and purified by Ni-NTA bead purification alone or by Ni-NTA bead purification followed by HR200 gel filtration (Figure 5.9). A preliminary SPR binding study was performed by coupling either 5898RU HR-200 gel filtrated CDC25B TCR, 7889RU Ni-NTA bead purified CDC25B TCR or 3969RU of an Ni-NTA bead purified negative control αβ TCR (LDN5) to separate flow cells using His-Ni binding. Following immobilisation of TCR to each flow cell S200 gel filtrated HLA-A2-GLLGpS complex in HBS-EP buffer was injected over each flow cell at a concentration of 1.3mg/ml. The resulting sensogram (Figure 5.10) showed increased levels of binding to the HR-200 purified CDC25B TCR relative to the negative control and also to the non-HR-200 purified CDC25B TCR. This result provided the first evidence that there was detectable specific binding of the HR200 purified CDC25B TCR to the HLA-A2-GLLGpS complex. However there are a number of caveats about this result which must be noted. Firstly, the sensogram shows that there was a high degree of sticking over the HR-200 CDC25B flow cell surface compared to control as indicated by the bi-phasic dissociation phase of the HLA-A2-GLLGpS ligand. As a result, the sensogram for this flow cell fails to return to pre-injection levels, which is typical of samples containing protein aggregates (van der Merwe, Barclay et al. 1994). This may be accounted for by the poor resolution of the TCR by HR-200 gel filtration prior to injection (Figure 5.9), or by an innate propensity to aggregate. The gel filtration profile also suggests plenty of opportunity for

Figure 5.9 Purification of CDC25B TCR by HR-200 gel filtration



Purification of the CDC25B TCR by HR-200 gel filtration shows a poorly refined elution profile (A) which contains multiple unresolved peaks (asterisks). Fractions corresponding to these peaks were analysed by SDS-PAGE (B). Each peak is denoted by green, red, blue and orange asterisks. The left hand panel represents samples loaded under reducing conditions and the right hand panel under non-reducing conditions. The analysis indicates that the majority of CDC25B TCR is contained in the peak denoted by the blue asterisk although multiple contaminants still exist.

Figure 5.10 Surface plasmon resonance sensogram of HLA-A2-GLLGpS binding to CDC25B TCR using an Ni-NTA chip

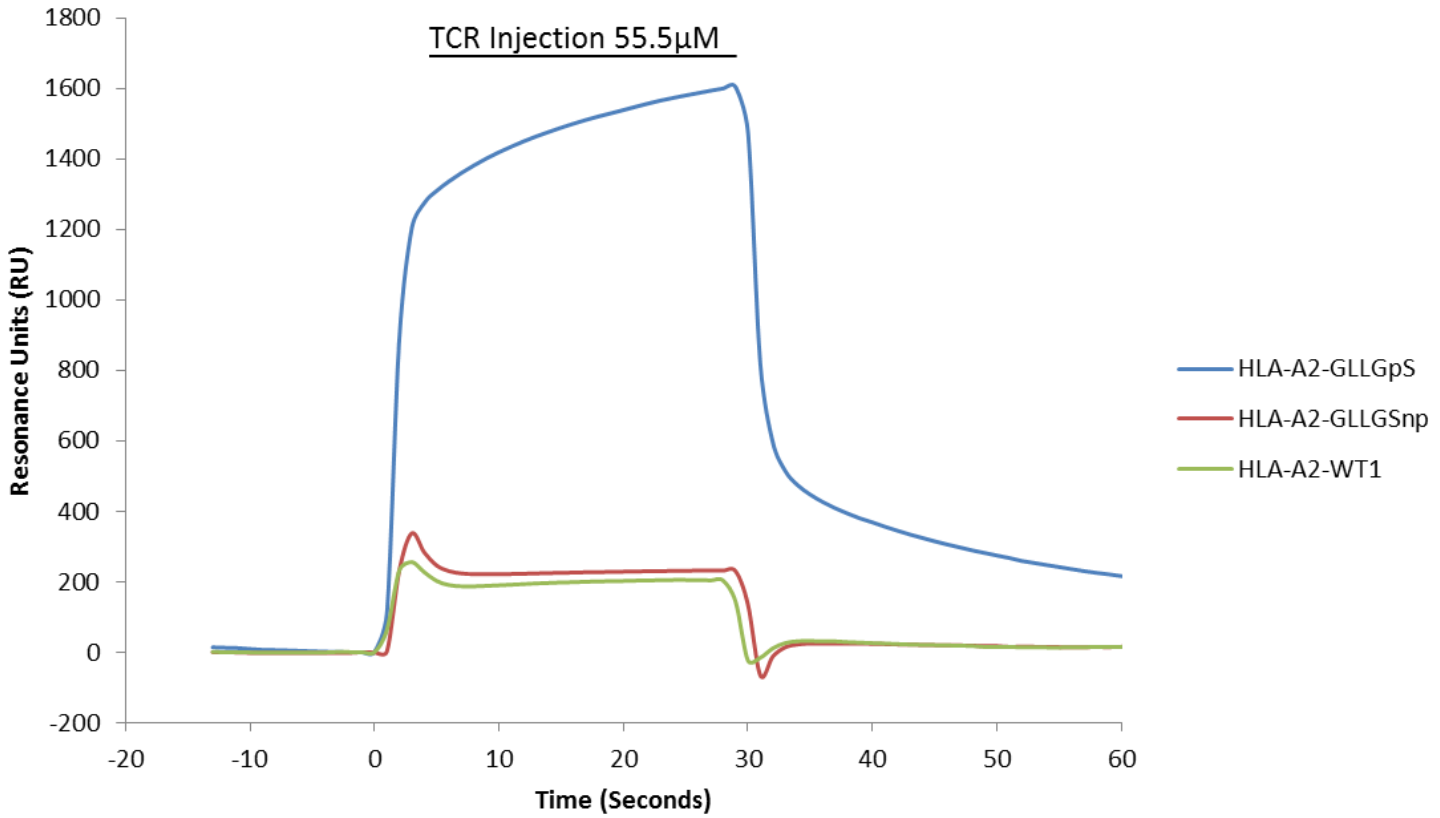


Following immobilisation of 5898RU HR200 purified CDC25B TCR (Blue), 7889RU Ni-NTA purified CDC25B TCR (Red) or 3969RU negative control LDN5 TCR (Green) to separate flow cells of an Ni-NTA via His tag-Ni coupling, HLA-A2-GLLGpS in solution at a concentration of 28.7µM was injected over each flow cell. The trace indicates a clear specific binding of HLA-A2-GLLGpS to the HR200 purified CDC25B TCR.

inactive (but His-containing) protein in pre-HR200 post-NiNTA samples which because of the poor resolution of gel filtration may not be fully removed from the HR-200 purified sample. This lack of purity may also account for the lack of binding seen in the non-HR200-purified sample as the presence of inactive His-containing proteins within the sample would also bind to the flow cell resulting in a gross overestimation of the levels of bound TCR. Secondly, the base level on all flow cells seems to be continually falling at a steady rate prior to and following the injection. This is indicative of on-going dissociation of immobilised TCR from the flow cell surface during the experiment, making accurate determination of relative amounts bound to each flow cell surface difficult. Finally, these results did not ascertain whether the CDC25B phosphopeptide specific TCR is phosphate and epitope specific.

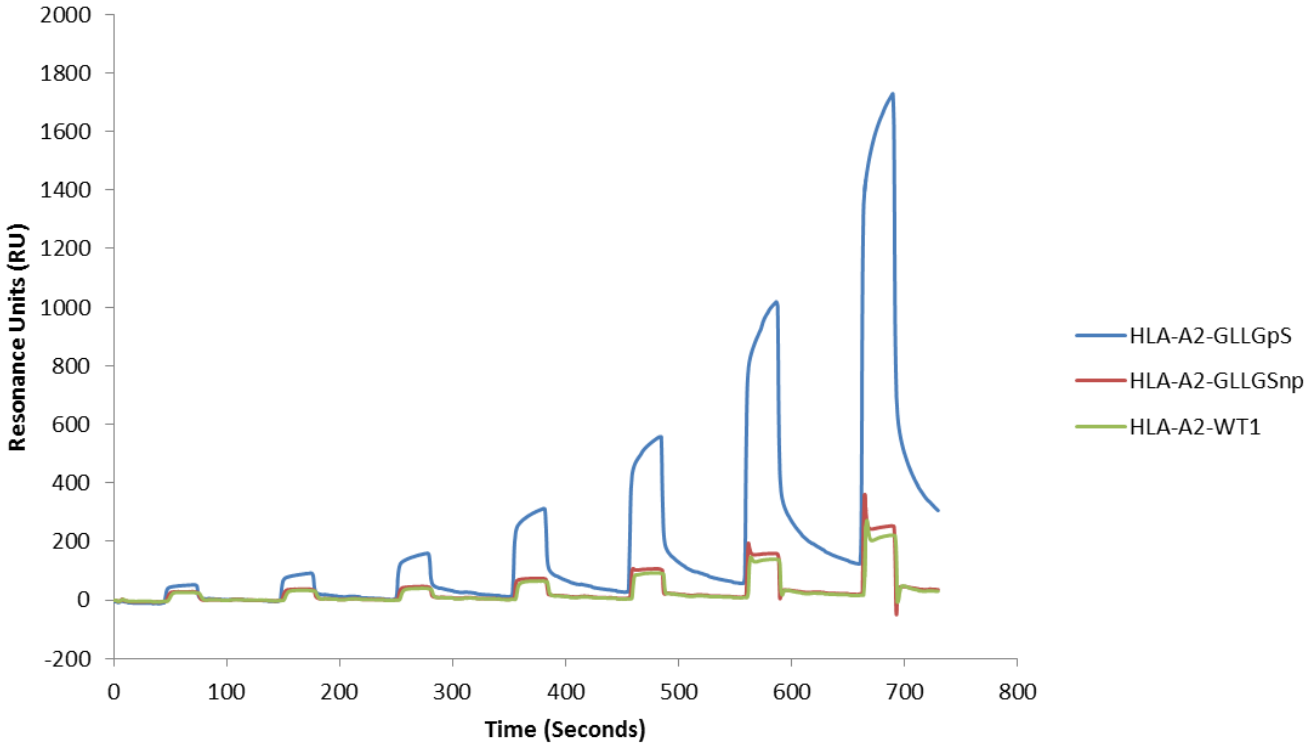
In order to address some of these issues a second SPR binding study was performed. Rather than coupling TCR to the surface of the sensor chip using His-Ni binding which proved to be suboptimal in the experiments above. A CM5 chip coated with streptavidin via amine coupling was used to immobilise biotinylated MHC molecules consisting of 2982RU HLA-A2-GLLGpS, 2768RU HLA-A2-GLLGnp and 2821RU of a negative control HLA-A2-WT1. CDC25B TCR which had been previously purified by HR-200 gel filtration was thawed and concentrated to 7.1mg/ml before dilution to 3.5mg/ml (55µM) in HBS-EP buffer. Initial test injections at this concentration revealed increases in signal which were much greater over the flow cell presenting HLA-A2-GLLGpS than over either of the two control flow cells, thus indicating specific binding (Figure 5.11). Therefore a series of TCR dilutions (ranging from 4 - 0.06mg/ml) were injected over each flow cell (Figure 5.12). The dilution series indicated that specific binding of CDC25B TCR to the phosphorylated epitope could be seen over the entire range of injected TCR concentrations. However, although specific binding to the

Figure 5.11 Surface plasmon resonance sensorgram of CDC25B TCR binding to HLA-A2-GLLGpS using a CM5 sensor chip



Following immobilisation of 2982RU HLA-A2-GLLGpS (blue), 2768RU HLA-A2-GLLGSnp (red) and 2821RU HLA-A2-WT1 negative control via streptavidin coupling to separate flow cells of a CM5 sensor chip, 5µl of CDC25B TCR was injected at a concentration of 55.5µM with a flow rate of 10µl/min. These results demonstrate clear specific binding of the CDC25B TCR to the HLA-A2-GLLGpS complex with no binding observed for the other HLA-A2 complexes.

Figure 5.12 Surface plasmon resonance sensorgram of CDC25B TCR dilution series binding to immobilised HLA-A2 complexes



Following immobilisation of 2982RU HLA-A2-GLLGpS (blue), 2768RU HLA-A2-GLLGSnp (red) and 2821RU HLA-A2-WT1 negative control via streptavidin coupling to separate flow cells of a CM5 sensor chip, 10µl of CDC25B TCR was injected in a dilution series of 0.9µM, 1.7µM, 3.5µM, 6.9µM, 13.9µM, 27.7µM and 55.5µM at a flow rate of 10µl/min. The sensogram indicates specific binding of CDC25B TCR to HLA-A2-GLLGpS at all concentrations.

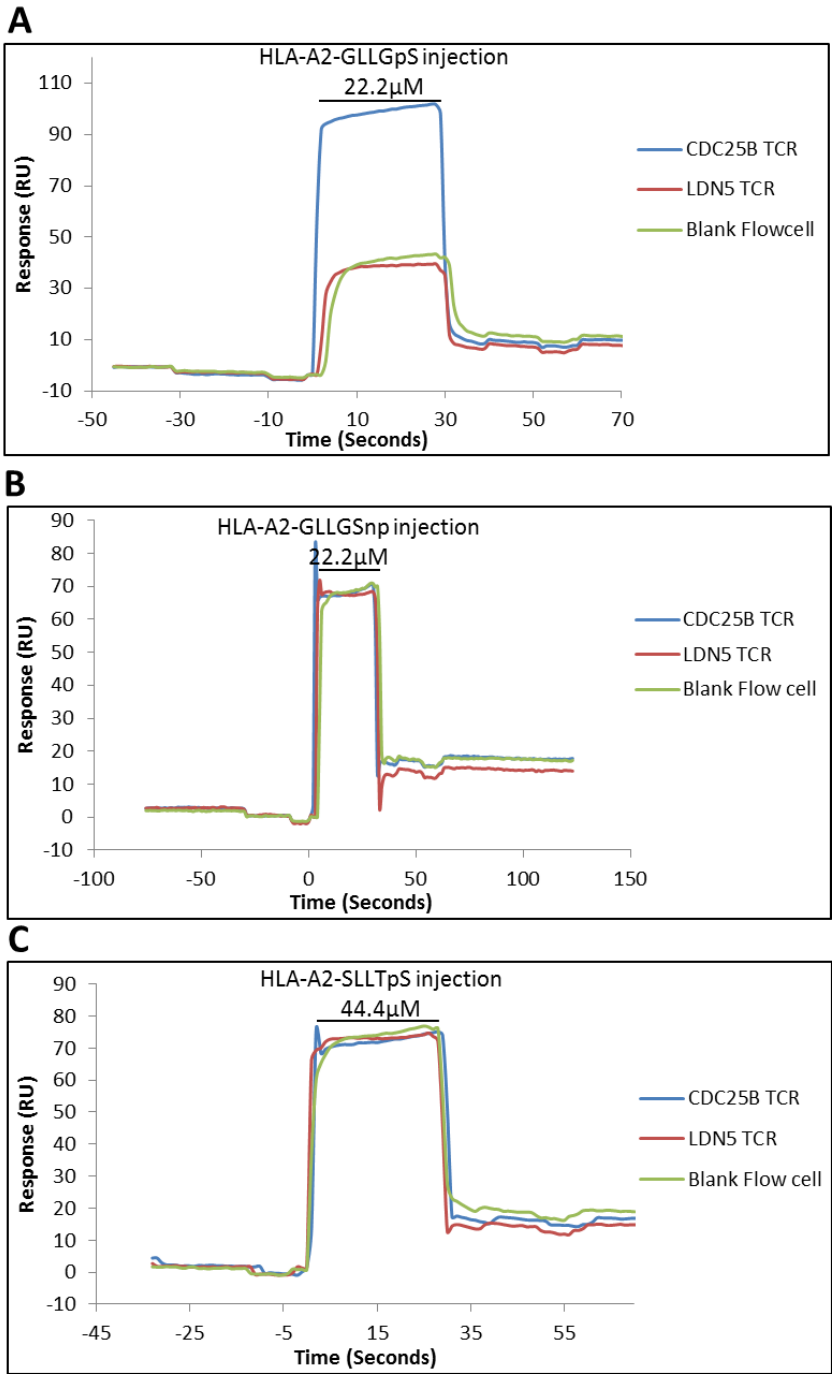
phosphorylated epitope was observed, there was still significant levels of TCR sticking to the HLA-A2-GLLGpS flow cell compared to the non-phosphorylated-MHC complex and negative control flow cells as evident by the bi-phasic dissociation phase of the profile. This problem is exacerbated at the higher protein concentration injections consistent with protein aggregation effects (van der Merwe, Barclay et al. 1994), thus limiting the accuracy of any specific binding readings. Due to these effects it is problematic to calculate an accurate affinity for the TCR-MHC interaction based on this data. The sticking effect observed in these experiments is likely due to protein aggregate formation within the injected solution, possibly due to freeze/thawing of the sample or following concentration.

Due to the inability to obtain pure CDC25B TCR samples at high concentrations without aggregation a third SPR binding study was designed in which biotinylated CDC25B TCRs were immobilised to the surface of a streptavidin-coated CM5 chip and MHC-peptide complexes injected over in solution. It was hypothesised that this method would reduce the sticking phenomenon observed in previous experiments because the HR-200 gel filtrated MHC complexes were highly purified and can be concentrated to high concentrations with minimal aggregate formation. However in order to immobilise the CDC25B TCR to the surface of a CM5 chip, the α chain of the CDC25B TCR needed to be cloned into the pMTBipV5-BirA/His vector to incorporate a C-terminal BirA tag, enabling biotinylation of the TCR protein. The CDC25B TCR containing a BirA biotinylation site was expressed as described above. 2995RU of Biotinylated CDC25B TCR was then immobilised on to streptavidin-coated flow cells of a CM5 sensor chip with 2985RU of LDN5 αβ TCR immobilised to a separate flow cell as a negative control. As well as reducing the amount of sticking observed in previous attempts this experiment also sought to determine the specificity of CDC25B TCR binding,

not only between phosphorylated and non-phosphorylated GLLGS epitope, but in comparison to other phosphorylated epitopes with a P5 pSer. Therefore complexes of HLA-A2-GLLGpS, HLA-A2-GLLGSnp and HLA-A2-SLLTpS were purified by HR-200 gel filtration immediately prior to concentration and injection over each flow cell. The resulting binding curves demonstrated unequivocal binding of HLA-A2-GLLGpS complex to the immobilised CDC25B TCR but not to the LDN5 TCR or blank flow cell (Figure 5.13A) and displayed minimal levels of sticking. Crucially, upon injection of HLA-A2-GLLGnp and HLA-A2-SLLTpS, no binding was observed over any of the flow cells indicating the CDC25B TCR binds in a phosphate-dependent and epitope specific manner (Figure 5.13 B/C). Subsequent affinity analysis using an HLA-A2-GLLGpS dilution range of 88.9µM – 0.7µM was performed by subtracting the control response (LDN5 flow cell) from the response measured in the CDC25B TCR flow cell at equilibrium for each dilution of injected HLA-A2-GLLGpS. The specific responses were then plotted vs free analyte concentration (HLA-A2-GLLGpS), or as a Scatchard plot (ie. specific binding vs specific binding/free ligand concentration) (Figure 5.14). This revealed the CDC25B TCR/HLA-A2-GLLGpS interaction to have an affinity of 41.2µM and 40.1µM by direct non-linear curve fitting and Scatchard analysis respectively (Figure 5.14).

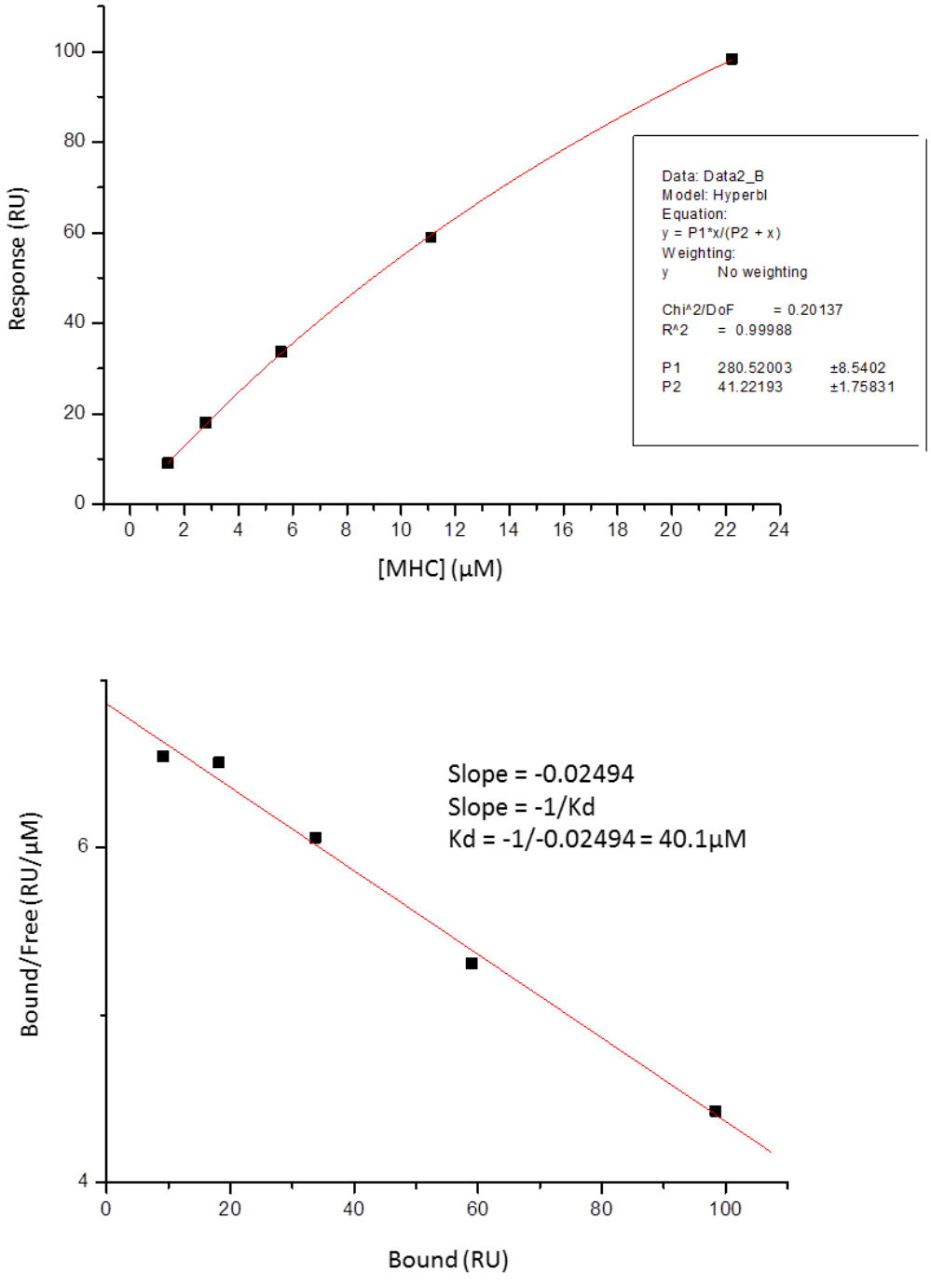
To verify this interaction and obtain accurate binding data in both orientations the CDC25B TCR was purified by HR200 gel filtration immediately prior to concentration and injection over 3024RU HLA-A2-GLLGpS, 3100RU HLA-A2-GLLGSnp and 3165RU HLA-A2-SLLTpS complex immobilised to separate flow cells of a CM5 chip. The top concentration was also reduced (to 2mg/ml rather than 7mg/ml) so as to minimise the level of aggregate formation during this step. A second alteration to previous experiments in this orientation was to increase the flow rate and decrease the volume of each injection, therefore reducing the

Figure 5.13 Surface plasmon resonance binding of CDC25B TCR with MHC in solution



Following immobilisation of CDC25B TCR (blue) or the negative control LDN5 TCR (red) onto separate flow cells of a CM5 sensor chip, 5µl of 22.2µM HLA-A2-GLLGpS (A), 22.2µM HLA-A2-GLLGSnp or 44.4µM HLA-A2-SLLTpS was injected over each flow cell using a flow rate of 10µl/min.

Figure 5.14 Equilibrium affinity measurement of CDC25B TCR-HLA-A2-GLGGpS interaction with MHC in solution

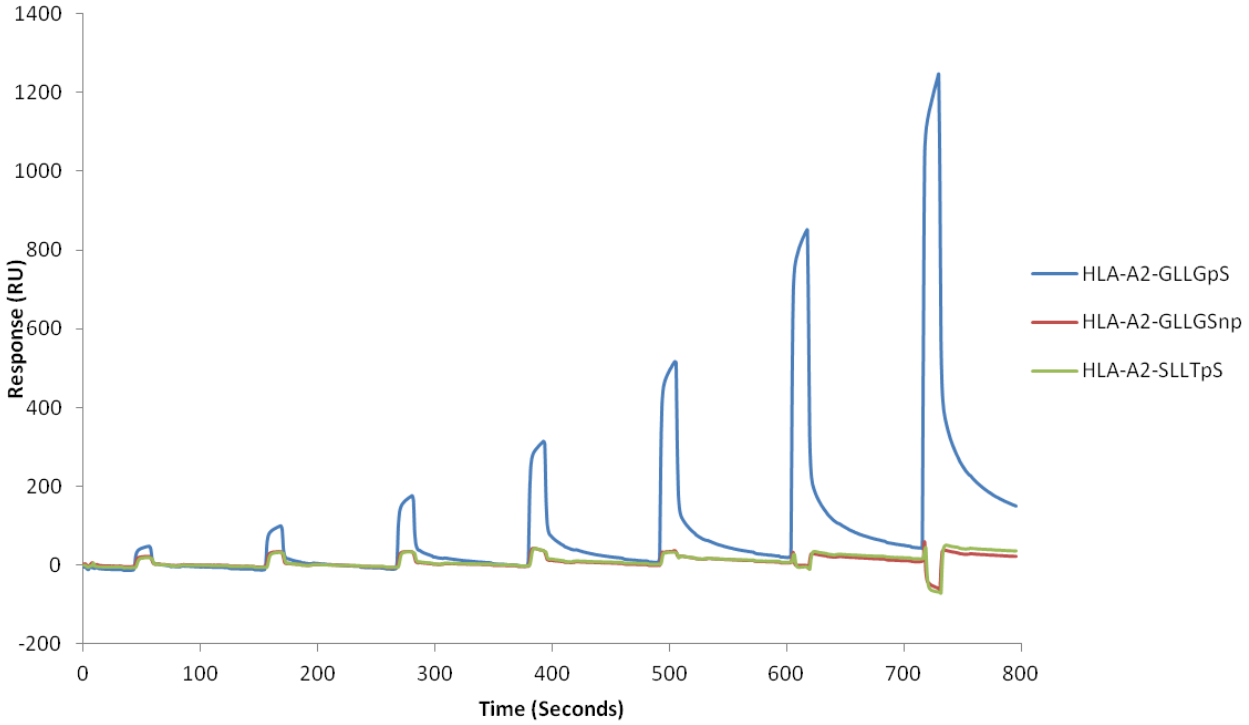


Equilibrium affinity analysis of CDC25B TCR with HLA-A2-GLLGpS complex using a CM5 sensor chip with TCR bound to the surface and HLA-A2 in solution at 22.2µM, 11.1µM, 5.6µM, 2.8µM and 1.3µM, demonstrates a specific binding affinity of 41.2µM and 40.1µM by direct non-linear curve and Scatchard plot analysis respectively.

time injected protein was in contact with the flow cell surface and the ability of aggregates to stick. Although there was still evidence of some sticking at the higher concentrations, the resulting injected dilution series (31.3µM – 0.5µM) revealed a considerably improved binding profile compared to previous experiments in this orientation and more importantly confirmed specific binding of the CDC25B TCR to HLA-A2-GLLGpS (Figure 5.15). Due to the improved quality of these results further affinity analysis in this alternate orientation was carried out. The measurement of specific binding response at each concentration revealed a similar affinity to that of the previous experiment in the opposite orientation (35µM and 38µM) as demonstrated by direct non-linear curve and Scatchard analysis respectively (Figure 5.16).

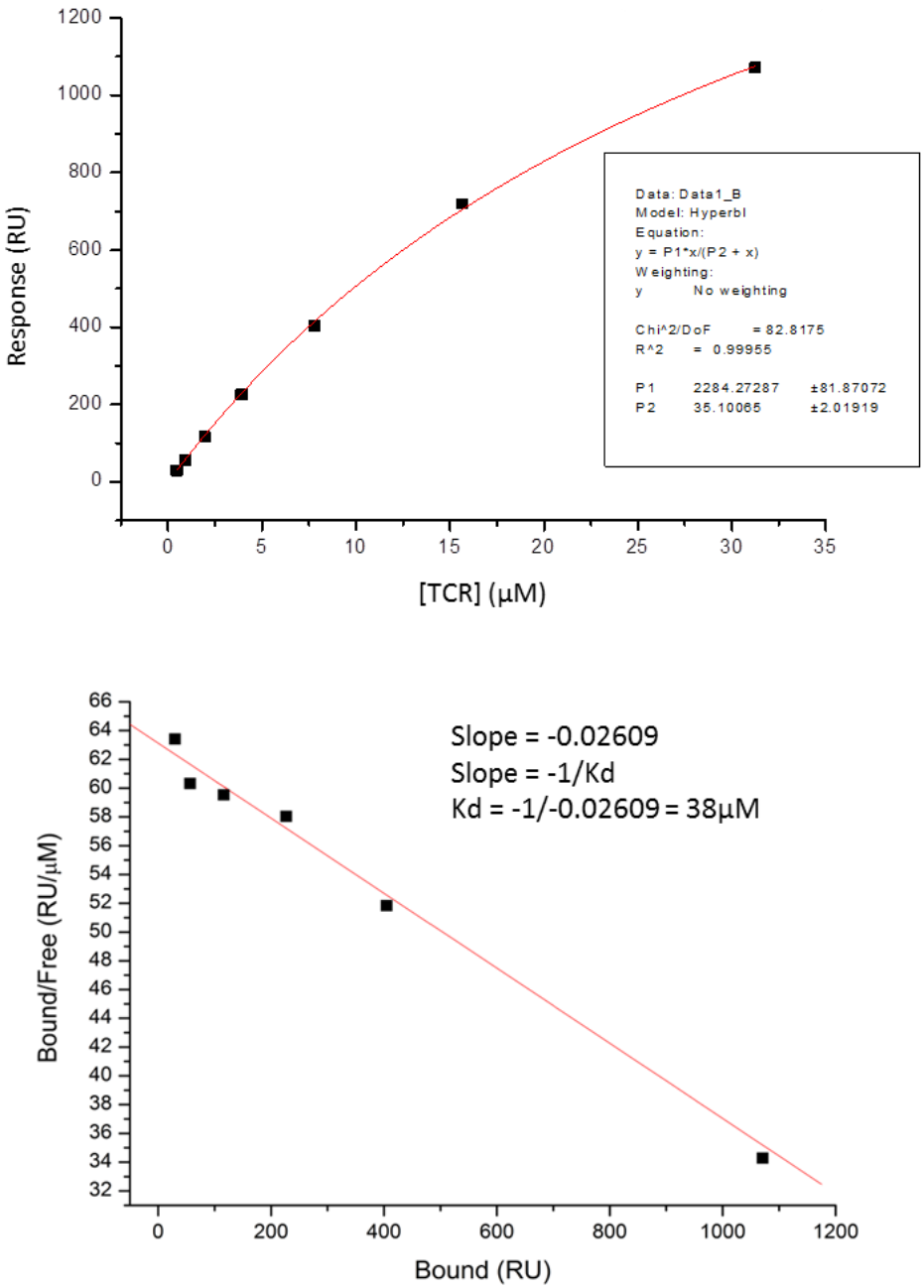
In order to determine how the phosphate moiety may influence TCR discrimination of GLLGpS, three unrelated TCR-HLA-A2 structures; 1AO7, 1BD2 and 1OGA were used to model the CDC25B TCR interaction with HLA-A2-GLLGpS complex. The HLA-A2 regions of the three TCR-HLA-A2 complex structures were superimposed onto the HLA-A2-GLLGpS complex, 3FQU, from the Petersen *et al* study using lsqman (Figure 5.17). The three resulting TCR-HLA-A2-GLLGpS models confirmed that the raised solvent exposed position of the phosphate would allow direct contacts to the TCR. In particular it highlighted the possibility of multiple direct contacts to α and β chain CDR3 loop regions. While direct TCR-phosphopeptide-MHC crystal structures will be required to validate these findings, they strongly suggest that even when phosphopeptide and non-phosphopeptide antigens are closely matched, the phosphate moiety is likely to be able to mediate direct contacts with appropriate phosphopeptide specific TCRs (Figure 5.17) thus ensuring TCR discrimination of the phosphopeptide from its non-phosphorylated counterpart.

Figure 5.15 Optimised surface plasmon resonance sensogram of CDC25B TCR dilution series binding to immobilised HLA-A2 complexes



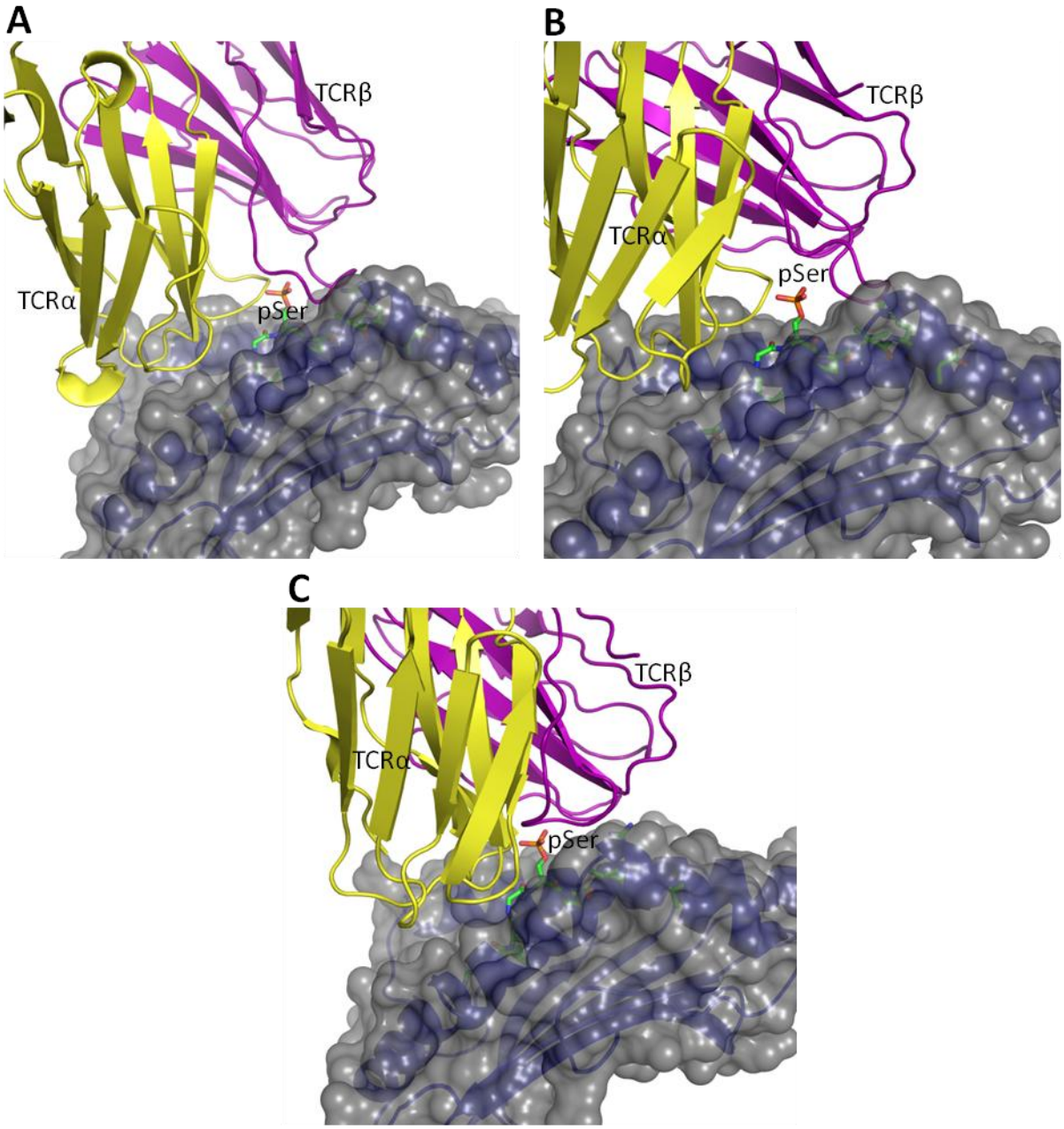
Following immobilisation of 3024RU HLA-A2-GLLGpS (blue), 3100RU HLA-A2-GLLGSnp (red) and 3165RU HLA-A2-WT1 negative control via streptavidin coupling to separate flow cells of a CM5 sensor chip, 5µl of CDC25B TCR was injected in a dilution series of 0.5µM, 0.9µM, 1.9µM, 3.9µM, 7.8µM, 15.6µM and 31.3µM at a flow rate of 20µl/min. The trace indicates specific binding of CDC25B TCR to HLA-A2-GLLGpS at all concentrations.

Figure 5.16 Equilibrium affinity measurement of CDC25B TCR, TCR in solution



Equilibrium affinity analysis of CDC25B TCR with HLA-A2-GLLGpS complex using a CM5 sensor chip with HLA-A2 bound to the surface and CDC25B TCR in solution demonstrates a specific binding affinity of 35.1µM and 38µM by direct non-linear curve and Scatchard plot analysis respectively.

Figure 5.17 Models of potential TCR interactions with GLLGpS phosphate moiety



TCR-HLA-A2 complex structures; 1AO7 (A), 1BD2 (B) and 1OGA (C) were used to model potential interactions between TCRs and the phosphate moiety of GLLGpS. This was performed by superimposing the HLA-A2 onto the HLA-A2-GLLGpS complex structure (PDB code 3FQU; (Petersen, Wurzbacher et al. 2009). Visualisation of the interactions and figure production was performed using Pymol. TCR α and β chains are highlighted in yellow and purple cartoon, respectively. HLA-A2 α1α2 regions are shown in blue cartoon with the surface depicted in grey. The GLLGpS epitope is shown in stick representation with carbons (green), oxygen (red), nitrogen (blue) and phosphate (orange).

5.6 Conclusions

Previous studies have shown that phosphopeptide specific T-cells are able to recognise and kill cells which present HLA-A2-phosphopeptide complexes and this can occur in a phosphate-dependent and epitope-specific manner, however the mechanism in which this occurs and the extent of any cross reactivity in TCR binding between matched phosphorylated and non-phosphorylated epitopes is poorly understood. These issues could potentially have profound implications in selection of phosphopeptide epitopes as targets in immunotherapy strategies. Conceivably if the non-phosphorylated epitope could induce cross-reactivity of phosphopeptide specific T-cells and was presented within the thymus or upon normal tissue this may result in impaired phosphopeptide specific T-cell responses due to central tolerance mechanisms or result in autoimmunity respectively. This study therefore sought to determine the evidence for direct phosphopeptide-specific TCR recognition of HLA-A2-phophopeptide complexes by examining the binding properties of several TCRs with their phosphorylated and non-phosphorylated epitope as well as to different epitopes with similar phosphate positioning. Overall the results revealed that the analysis of the TCR-HLA-A2-GLLGpS complex affinity by SPR was highly dependent upon the purity of TCR used. In each experiment, sticking was observed as a bi-phasic dissociation phase along with other features such as not reaching equilibrium which is indicative of aggregate binding (van der Merwe, Barclay et al. 1994). In the case of CDC25B TCR, the aggregate binding still demonstrated specificity for the HLA-A2-GLLGpS complex indicating that aggregated TCR still retains specific binding properties. Even following purification by HR-200 gel filtration, the subsequent concentration of the TCR preparation often resulted in aggregate formation which could potentially be exacerbated by residual impurities due to the poor resolution of

the HR-200 gel filtration purification (observed as multiple merged peaks on the elution profile). Although limiting the amount of TCR concentration prior to injection and decreasing the length of injection slightly alleviated some of the observed sticking, the employment of additional purification steps such as increased imidazole concentration during washing of the Ni-NTA bead column or anion/cation exchange may help to improve the initial TCR purity and reduce aggregate formation. However, the fact that the aggregate effects are only observed on specific flow cells suggests that they are aggregated TCR rather than non-specific impurities. This raises the question as to whether further buffer optimisation is required to alleviate this propensity for the TCR to aggregate following gel filtration purification.

The IRS2 and IRS2BK TCRs, whilst known to enable T-cell recognition and killing of HLA-A2-RVApS presenting cells (Zarling, Polefrone et al. 2006), demonstrated no binding by SPR analysis. One conclusion could again be the purity of these receptors used in SPR analysis affecting the quality of binding observed, however this is unlikely as the purity of the IRS2 TCR was much greater than that of the CDC25B TCR which displayed specific binding in SPR experiments. One possible explanation for these findings is that the IRS2 and IRS2BK TCRs may interact with HLA-A2-RVApS with a very low affinity that is below the detection limit for SPR methods. This has been noted before for immune-receptor complexes (such as CD4 binding to MHC-class II) which are known to bind but due to their low affinity are not observable by SPR. In these situations other methodologies, such as Alpha-screen, which relies on chemiluminecence-based detection of multimeric bead-bead interactions, could be used to detect if these TCRs can recognise HLA-A2-RVApS complexes.

However, despite no observable binding of the two HLA-A2-RVApS-specific TCRs the CDC25B TCR demonstrated specific binding to the HLA-A2-GLLGpS complex upon analysis using an Ni-NTA chip and CM5 in ether orientation (bound to surface or in solution). Further analysis revealed that the TCR has an affinity within the region of other TCRs specific for non-phosphorylated epitopes such as certain viral epitopes (1-100μM) (Davis, Boniface et al. 1998). The CDC25B TCR was also shown to be able to discriminate between epitopes in a phosphate-dependent but also an epitope-specific manner. This is further supported by data from Petersen *et al* (Petersen, Wurzbacher et al. 2009) which shows that the phosphate moiety of the GLLGpS epitope adopts a raised solvent exposed position within the central region of the epitope. Further to this, modelling of 3 unrelated TCR receptors onto the HLA-A2-GLLGpS complex structure revealed that the phosphate is situated in an optimal position close to the CDR loop regions of the TCR α and β chains. This may result in direct contact of the TCR with the phosphate moiety and other residues of the GLLGpS epitope.

As mentioned above, the ability and mechanism of phospho-antigen discrimination by phosphopeptide specific T-cells are important considerations in any potential vaccine design. Clearly phosphate induced conformational change is not a prerequisite for productive phosphopeptide-TCR interactions as demonstrated by CDC25B TCR-HLA-A2-GLLGpS. However, while some epitopes which do not undergo conformational change may be specifically recognised by phosphopeptide specific TCRs with an affinity which allows measurement by SPR, such as GLLGpS, others, such as RVApS may not. This difference in recognition of the two antigens by each of the TCR clones studied hear, which have been previously reported to exhibit activation of T-cells (both native $CD8^+$ T-cells and TCR transfected Jurkat cells) upon binding to specific HLA-A2-phosphopeptide complex (Zarling,

Polefrone et al. 2006), may be due to the fact that in the GLLGpS epitope, the pSer is in the P5 position rather than the P4 position in RVApS. This may result in the P5 phosphate mediating more contacts with the TCR as it is positioned centrally relative to the TCR binding footprint compared to the P4 positioning in RVApS, which adopts a more remote position closer to the N-terminus of the peptide, and may therefore only allow minimal contacts to the TCR. This conceivably provides an explanation for the lower affinity observed for the HLA-A2-RVApS specific TCRs. Also the HLA-A2-RVApS complex reported by Petersen *et al* (Petersen, Wurzbacher et al. 2009) demonstrated that the RVApS epitope adopts a low conformation in the binding groove resulting in flexibility of the phosphate moiety compared to that of other canonical phosphopeptides reported by Mohammed *et al* (Mohammed, Cobbold et al. 2008) and those from this study. This may therefore contribute to reducing the availability of the phosphate moiety to form stable contacts with RVApS specific TCRs thus reducing the overall stability and affinity of the HLA-A2-RVApS-TCR interaction. Unfortunately, due to the lack of detectable interaction observed for the RVApS specific TCRs used in this study, examining the structure of the HLA-A2-RVApS-TCR complex by X-ray crystallography methods proved unfeasible, although future analysis of the HLA-A2-GLLGpS in complex with CDC25B TCR may help to shed light on these issues.

Due to the discrepancies in TCR recognition of phosphopeptides which undergo little conformational change, it is therefore conceivable that the most optimal approach for future vaccination strategies would be to target antigens such as PKD2 and RQAS_V which were shown in chapter 4 to undergo dramatic conformational change upon phosphorylation. Although responses to the RQApS_V epitope have been raised in human T-cells, obtaining a clonal population has proved elusive (personal communication from Dr Mark Cobbold).

Further work to isolate a clonal population of HLA-A2-RQApS_V specific T-cells and isolate the TCR into expression systems for further structural and biophysical analysis would greatly aid in determining how conformational change as well as phosphate presence affects TCR binding.

These results show for the first time that certain phosphopeptide specific T-cells are able to discriminate in a phosphate-dependent and epitope-specific manner at the level of TCR binding. Due to the fact that phosphorylation is highly altered in tumour tissue, it is conceivable that if the phosphopeptide were inherently linked to the tumourigenesis process and, like GLLGpS, was able to be discriminated by T-cells from its non-phosphorylated counterpart, this would provide a highly suitable target for tumour immunotherapy strategies.

6 Discussion

This thesis has focused on the molecular basis of phosphopeptide antigen presentation and recognition of phosphopeptide antigens by specific TCR's. In chapter 3 I demonstrated that the phosphate moiety of canonical phosphopeptides displays a previously unidentified epitope specific plasticity in its orientation and the interactions formed with surrounding residues, with implications for the role the phosphate moiety plays in contributing to the binding energetics for certain epitopes. This was followed in chapter 4 by analysis of the overall structure of several epitopes in both their phosphorylated and non-phosphorylated states. This comparison revealed that for some epitopes, addition of a phosphate moiety had very little impact upon the overall main chain structure of the epitope. However for other epitopes, introduction of a phosphate moiety had a profound effect upon the overall antigen main chain conformation which, when compared to their non-phosphorylated counterparts, represent structurally distinct antigens and may therefore be more likely to be discriminated from their non-phosphorylated counterparts by phosphopeptide specific T-cells. Finally, in chapter 5, the molecular basis of phosphopeptide discrimination by phosphopeptide specific TCR's was studied. This demonstrated that one of these TCRs was able to specifically bind to HLA-A2-GLLGpS in a phosphate-dependent and epitope-specific manner, with an affinity comparable to that of previously documented TCR-MHC-antigen interactions (Davis, Boniface et al. 1998).

Previously, Mohammed *et al* demonstrated that inclusion of a phosphate moiety can increase the antigen affinity for HLA-A2 (Mohammed, Cobbold et al. 2008). By studying the

binding affinity of a greater number of antigens, this study has shown that the majority of HLA-A2 phosphopeptides possessing a canonical motif can fit into three categories depending upon the type of P2 and PC anchor residues they possess. Phosphopeptide antigens with canonical P2 and PC anchors display the least increase in binding upon phsophorylation, followed by antigens with non-canonical P2 but canonical PC, and finally those with non-canonical anchors at P2 and PC which demonstrate the greatest increase in binding upon addition of the phosphate moiety. However, two antigens in this study (RTLpS and RQApS_V) did not fit this classification. RQA_V showed a similar binding affinity upon phosphate addition compared to its non-phosphorylated form. This was surprising, as the epitope possess the canonical motif and also displays a conserved canonical binding mode as demonstrated in chapter 3. However, the structure of both the phosphorylated and non-phosphorylated forms of RQA_V may provide an explanation for the lack of phosphate induced increase in epitope binding by the fact that this epitope undergoes significant conformational rearrangement upon addition of the phosphate. Although addition of the phosphate results in additional phosphate mediated contacts the change in epitope conformation leads to the loss of several interactions between the peptide and the MHC leading to a net reduction in peptide MHC contacts. Therefore one hypothesis generated by this study is that this lack of phosphate-induced increase in binding affinity due to conformational rearrangement may also be true for other epitopes, such as RTLpS, however further comparison of the phosphorylated and non-phosphorylated forms of these epitopes would need to be performed to fully understand these effects for individual epitopes. Comparison of the canonical phosphopeptide structures have also revealed plasticity in the phosphate moiety orientation and the stabilising contacts the phosphate moiety forms for

each individual epitope. However, further correlation with antigen affinity data revealed no significant impact upon overall antigen affinity when the phosphate moiety adopted altered orientations to those seen in the majority of canonical phosphopeptides studied to date. This subtle re-orientation of the phosphate moiety for certain epitopes (RTFpS, RLSpS) may therefore highlight that the phosphate adopts altered confirmations in order to obtain the most optimal energetic potential for the given epitope. These results may also provide a model for the effects seen in mutational studies of RQA_M (Mohammed, Cobbold et al. 2008), which demonstrated that loss of both P1 Arg and Arg65 was required to significantly reduce phosphopeptide affinity, upon loss of only one of these interactions it is highly likely that the phosphate may adopt an alternate conformation in order to obtain the highest binding potential and complex stability.

Another point that this study has highlighted is that it is problematic to predict the overall binding properties of a phosphopeptide epitope based upon its sequence alone. Epitopes such as RTLpS and RQA_V not only demonstrate this point, but also the extent to which the phosphate moiety can influence the overall epitope structure. In some cases addition of a phosphate moiety can result in substantial rearrangement of the epitope main chain (RQA_V and PKD2); however, other epitopes (RQIpS and RTFpS) are able to naturally adopt the raised P4 positioning commonly observed in the canonical phosphopeptide group but without the presence of the phosphate moiety. This maintenance of the main chain conformation seems to be due to several factors which are epitope specific such as the presence of a proline at P5, which has been suggested to restrict the antigen main chain in this region due to its limited flexibility (Petersen, Wurzbacher et al. 2009). The positioning of His 70 may also adopt a raised position, thus contributing to the elevation of the main chain

around the P4-P5 region of the epitope. Also, the presence of interactions between P2 and residues of the MHC to form a “molecular clamp”, which are present in the canonical phosphopeptides but are lost upon removal of the phosphate in antigens that undergo main chain rearrangement, can help to restrict the conformation within this region of the epitope main chain. Finally, it is also conceivable that residues at P3 and P5 may also impact upon the ability of the antigen to maintain a raised position through interactions with surrounding residues of the MHC. This may be true for RQIpS in that although the side chain of Glu155 displayed only partial electron density, indicating a degree of flexibility, the residue is located in a position which may allow interactions with the P5 Glu side chain of RQIpS, thus potentially contributing to the maintained elevation of the main chain within this region. It is therefore clear that detailed structural analysis of individual epitopes in their phosphorylated and non-phosphorylated forms is required for determining the impact of phosphorylation upon not only antigen binding but also overall antigen conformation.

The ability of the phosphate moiety to influence epitope binding and epitope conformation are both important considerations for the design of any future vaccination strategies targeting phosphopeptide antigens. One major factor influencing the success of any potential vaccination strategy is the relative expression and presentation of the target antigen on various tissue types, including thymus and normal tissue.

As this study has shown, phosphopeptide epitopes can either undergo conformational change or maintain the overall epitope structure upon addition of the phosphate moiety. If the epitope undergoes substantial conformational change upon addition of the phosphate, these epitopes represent structurally distinct antigens compared to their non-

phosphorylated counterparts. It is therefore conceivable that even in the event of the non-phosphorylated form being presented within the thymus, due to the unique structural identity of these phosphorylated epitopes, phosphopeptide specific T-cells may not be deleted. This putative escape of T-cells specific for phosphopeptides which undergo conformational change is further supported by the fact that in order for the phosphopeptide epitope to be present within the thymus requires not only the expression of the parent protein but also the activity of the relevant kinase to form the specific phosphorylation event. Due to these factors, epitopes which undergo conformational change upon phosphorylation represent prime candidates for vaccination strategies using multiple approaches such as peptide vaccination, TCR gene transfer, antigen primed DC vaccination.

The second group of phosphopeptide antigens, those which do not undergo conformational change, may be more susceptible to central tolerance mechanisms due to the structural similarities between the phosphorylated and non-phosphorylated forms, with the only structural difference between the two being the presence or absence of the phosphate moiety. Therefore the position of the phosphate within these epitopes may play an important role in determining the degree to which they are able to induce epitope cross-reactivity of phosphopeptide and non-phosphopeptide specific T-cells. For example, the non-canonical epitope GLLGpS (which exhibits no conformational change upon addition of the phosphate), the p5 positioning of the phosphate lies within a region of the epitope that would allow direct phosphate mediated contacts to both the α and β chains of the TCR, most notably to the CDR3 regions. The presence or absence of the phosphate would therefore be likely to represent a significant structural difference in terms of TCR discrimination, meaning that if the non-phosphorylated epitope was present within the thymus it may not induce

cross-reactivity of phosphopeptide specific T-cells. I was able to address discrimination of phosphorylated and non-phosphorylated forms of GLLG by a phosphopeptide specific TCR in chapter 5 and demonstrated that this TCR was able to recognise HLA-A2-GLLGpS complex in a phosphate-dependent and epitope-specific manner with an affinity comparable to that of anti-viral responses.

However in the event of no conformational change to an epitope upon phosphate moiety addition to a residue, such as P1, may result in increased cross reactivity between the phosphorylated and non-phosphorylated epitopes due to the fact that P1 is a residue commonly ignored by the TCR. This could potentially lead to increased central tolerance of the T-cell repertoire not only in the event of the phosphorylated form being present within the thymus but also if the non-phosphorylated form were to be present. This hypothesis may explain why the two RVApS specific TCRs (IRS2 and IRS2BK) demonstrated no observable phosphopeptide specific binding by SPR. The fact that the RVApS epitope demonstrates no conformational change upon addition of the phosphate moiety (Petersen, Wurzbacher et al. 2009), and that the P4 positioning of the phosphate lies within a region which potentially would allow only minimal contacts to the TCR, may lead to an increased potential for TCR cross reactivity between the phosphorylated and non-phosphorylated epitopes. Therefore, if the non-phosphorylated RVAS epitope is expressed within the thymus this potential for cross reactivity could potentially lead to functional deletion or tolerisation of the T-cell repertoire with only the T-cells possessing low affinity phosphopeptide specific TCRs escaping negative selection. These factors would therefore impact upon the success of any vaccination strategy.

This also raises the therapeutic importance of some fundamental aspects of phosphopeptide presentation and recognition. Firstly, a major question is whether phosphopeptide antigens are presented within the thymus. For an immunotherapy approach to be effective there must be a pool of T-cells that can bind the antigen/MHC complex with high enough avidity to invoke a positive proliferative and activatory response. If the antigen is presented within the thymus upon medullary thymic epithelial cells (mTEC) then negative selection of antigen specific high avidity T-cells will occur and result in a pool of low avidity T-cells. However certain phosphopeptides may not be presented in the thymus due to lack of other components involved in the initial phosphorylation of the antigen. This means that although T-cells reactive with non-phosphorylated antigens could well be deleted in the thymus due to promiscuous expression of tissue specific antigen (TSA) by AIRE (Cheng, Shum et al. 2007; Mathis and Benoist 2009), T-cells specifically recognising the phosphorylation in the context of a certain antigen might not be deleted and so form a pool of high/moderate avidity T-cells able to target and kill cells presenting this antigen. This would be an attractive therapy in a cancer setting as the up-regulation of phosphorylation events coupled with the over-expression of many proteins provides an extra level of selectivity in targeting tumours. In relation to these points, on-going studies have shown that a broad range of HLA-A and HLA-B phosphopeptide specific T-cells are able to be obtained from healthy donor as well as cancer patient PBMC and that these are functionally active, producing high levels of IFN-γ upon phosphopeptide recognition (personal communication from Dr Mark Cobbold). This clearly demonstrates that certain phosphopeptide epitopes are able to escape central tolerance mechanisms. In addition, these phosphopeptide specific T-cells also seem to be of clinical relevance as cancer patients who do not posses phosphopeptide reactive T-cells have

a worse prognosis than patients who do mount a broad phosphopeptide specific response (personal communication from Dr Mark Cobbold).

A second question is how efficiently phosphopeptides are cross-presented, since this could significantly affect natural phosphopeptide-specific responses in cancer patients, as well as DC vaccination approaches. For a vaccine to elicit a CTL response it must be able to undergo cross-presentation. Early evidence in mouse models by Zarling *et al* (Zarling, Ficarro et al. 2000; Zarling, Polefrone et al. 2006) have shown generation of CTL by phosphopeptide pulsing of bone marrow derived DC and T-cell stimulation in vivo suggesting that the phosphate can survive cross-presentation. This however is still yet to be proven in a human setting.

In addition it is currently unclear how specific phosphopeptides are for expression on tumours verses normal tissue. Clearly HLA-B class I molecules, such as HLA-B7 and HLA-B27, represent a promising avenue for future vaccination strategies due to the HLA-B molecules presenting a far greater amount of phosphopeptide antigens than HLA-A molecules, and the fact that associated phosphopeptides are from source proteins that have been strongly linked to tumourigenisis, such as altered kinases (Michelotti, Price et al. 2000; Krivtsov and Armstrong 2007). The molecular basis for this increase in epitope presentation by HLA-B molecules is currently unknown, however they do possess a similar canonical motif to that seen in HLA-A2 binding phosphopeptides in that P1 is typically a charged residue and the phosphate is a P4 pSer. Therefore it is conceivable that the binding mode may be similar to that of HLA-A2 canonical phosphopeptides. During this study attempts were made to try to elucidate the molecular mechanisms governing the high degree of phosphopeptide

presentation in the context of HLA-B7 through analysis of the binding affinities for several epitopes in their phosphorylated and non-phosphorylated forms and also by determining the crystal structure of these epitopes. Although methods I developed for refolding and purification of HLA-B7-phosphopeptide complexes was successful (documented in Appendix section 6) the low yield of complex obtained rendered crystallisation challenging. Surprisingly, analysis of the binding affinities of HLA-B7 epitopes (performed by Clara Cummings and Victor Engelhard) revealed that upon addition of a phosphate moiety the affinity of the epitope was actually reduced compared to its non-phosphorylated counterpart rather than increasing like many of the HLA-A2 epitopes, suggesting the molecular ground rules governing HLA-B7-phosphopeptide binding may be different. Therefore if phosphorylation actually reduces epitope affinity we are left with the question as to why there is such a high prevalence of phosphopeptide antigens associated with HLA-B alleles. This could be due to factors other than antigen-MHC binding interactions. One possibility for the increased prevalence of phosphopeptides associated with HLAB molecules is that the non-phosphorylated HLA-B-restricted epitopes naturally possess a higher degree of similarity to Kinase motifs and therefore the pool of phosphopeptides able to bind to HLA-B molecules is also increased. A second possibility could be that these epitopes are preferentially selected for because the phosphorylation results in altered antigen processing such as increased or altered proteasome cleavage. Alternatively or in addition, phosphorylation may play an important role in HLA-B7 epitope presentation through alteration of TAP transport, possibly resulting in increased epitope transfer into the ER upon epitope phosphorylation. Finally, it is still unclear what effect phosphorylation of an epitope has upon antigen processing within the ER and how phosphorylation may affect antigen

processing by the peptide loading complex. Due to the fact that the effects of phosphorylation upon antigen processing pathways have to date not been investigated, it is currently difficult to determine the reasons behind the increased prevalence of phosphopeptide antigens in the context of HLA-B. Future work should therefore aim to study not only the structural aspects of HLA-B-phosphopeptide presentation but should also focus upon the impact of phosphorylation on events during antigen processing.

It is clear from this and other studies that phosphopeptide antigens are emerging as a novel set of targets for use in tumour immunotherapy strategies. This study has demonstrated that phosphorylation has a profound effect on the antigenic identity of class I MHC peptides and how they are recognised by the TCR. This information suggests that phosphopeptides are highly distinct antigens and raises hope that they are suitable candidates for cancer immunotherapy.

Appendix A: HLA-A2-RLQpS complex structure solution rotation and translation

Rotation Function						
Peak Number	θ (°)	ϕ(°)	X (°)	RF-function		
1	173.50	-111.60	173.73	20.37		
2	168.98	-138.61	177.19	5.48		
3	6.18	27.04	176.73	5.41		
4	14.00	68.42	134.34	5.18		
5	104.82	-86.78	80.70	4.83		
Translation Function						
Peak Number	TX	TY	TZ	TFcnt	R-fac	Score
1	0.404	0.000	0.227	49.83	0.509	0.558
2	0.899	0.000	0.200	1.97	0.615	0.308
3	0.258	0.000	0.349	2.81	0.618	0.304
4	0.542	0.000	0.222	1.86	0.620	0.302
5	0.300	0.000	0.321	1.96	0.620	0.300

Initial phase information was determined by molecular replacement using MOLREP (Murshudov, Vagin et al. 1997) with the HLA-A2 structure (PDB code; 1HHG) as the phasing model. The top five results for the rotation and translation functions are shown and the correct solution for each is highlighted in red. This is determined by a peak in the RF-function and TFcnt/Score values for the rotation and translation functions respectively.

Appendix B: HLA-A2-RLSpS complex structure solution rotation and translation values

Rotation Function						
Peak Number	θ (°)	ϕ(°)	X (°)	RF-function		
1	173.29	-116.98	171.40	20.02		
2	7.70	25.23	176.89	6.26		
3	168.88	-142.20	176.28	6.05		
4	105.18	-86.78	79.13	5.27		
5	114.52	-87.65	78.44	4.97		
Translation Function						
Peak Number	TX	TY	TZ	TFcnt	R-fac	Score
1	0.401	0.000	0.227	48.26	0.539	0.541
2	0.481	0.000	0.212	2.14	0.623	0.340
3	0.093	0.000	0.265	3.55	0.617	0.336
4	0.093	0.000	0.264	3.64	0.620	0.334
5	0.375	0.000	0.232	2.80	0.623	0.332

Initial phase information was determined by molecular replacement using MOLREP (Murshudov, Vagin et al. 1997) with the HLA-A2 structure (PDB code; 1HHG) as the phasing model. The top five results for the rotation and translation functions are shown and the correct solution for each is highlighted in red. This is determined by a peak in the RF-function and TFcnt/Score values for the rotation and translation functions respectively.

Appendix C: HLA-A2-RTFpS complex structure solution rotation and translation values

Rotation Function						
Peak Number	θ (°)	ϕ(°)	X (°)	RF-function		
1	40.28	-159.21	135.23	15.16		
2	47.28	-159.95	134.81	4.92		
3	47.31	-164.67	132.66	4.81		
4	63.09	-166.16	49.92	4.67		
5	140.18	-61.48	169.38	4.25		
Translation Function						
Peak Number	TX	TY	TZ	TFcnt	R-fac	Score
1	0.336	0.045	0.224	30.98	0.535	0.493
2	0.850	0.060	0.091	3.22	0.624	0.278
3	0.235	0.311	0.234	1.88	0.627	0.277
4	0.155	0.398	0.083	2.96	0.624	0.277
5	0.240	0.268	0.362	1.92	0.624	0.276

Initial phase information was determined by molecular replacement using MOLREP (Murshudov, Vagin et al. 1997) with the HLA-A2 structure (PDB code; 1HHG) as the phasing model. The top five results for the rotation and translation functions are shown and the correct solution for each is highlighted in red. This is determined by a peak in the RF-function and TFcnt/Score values for the rotation and translation functions respectively.

Appendix D: HLA-A2-RQIpS complex structure solution rotation and translation values

Rotation Function						
Peak Number	θ (°)	ϕ(°)	X (°)	RF-function		
1	173.25	-115.21	172.24	20.76		
2	7.18	23.40	175.93	5.90		
3	168.83	-140.41	176.00	5.50		
4	11.27	54.05	122.06	4.87		
5	104.71	-86.67	79.73	4.82		
Translation Function						
Peak Number	TX	TY	TZ	TFcnt	R-fac	Score
1	0.401	0.000	0.227	54.16	0.529	0.552
2	0.390	0.000	0.296	2.81	0.621	0.335
3	0.278	0.000	0.274	2.57	0.622	0.331
4	0.472	0.000	0.208	2.88	0.626	0.329
5	0.026	0.000	0.240	1.99	0.623	0.329

Initial phase information was determined by molecular replacement using MOLREP (Murshudov, Vagin et al. 1997) with the HLA-A2 structure (PDB code; 1HHG) as the phasing model. The top five results for the rotation and translation functions are shown and the correct solution for each is highlighted in red. This is determined by a peak in the RF-function and TFcnt/Score values for the rotation and translation functions respectively.

Appendix E: HLA-A2-RTLpS complex structure solution rotation and translation values

Rotation Function					
Peak Number	θ (°)	ϕ(°)	X (°)	RF-function	
1	173.58	-114.32	173.03	19.63	
2	14.30	66.67	134.54	5.68	
3	103.99	-86.76	81.08	5.45	
4	169.39	-140.42	175.90	5.43	
5	0.94	5.41	179.99	5.12	
Translation Function					
Peak Number	TX	TY	TZ	TFcnt	R-fac	Score
1	0.892	0.000	0.223	52.44	0.477	0.617
2	0.020	0.000	0.365	3.02	0.580	0.388
3	0.063	0.000	0.338	3.14	0.581	0.386
4	0.216	0.000	0.298	1.72	0.584	0.380
5	0.549	0.000	0.321	1.68	0.583	0.378

Initial phase information was determined by molecular replacement using MOLREP (Murshudov, Vagin et al. 1997) with the HLA-A2 structure (PDB code; 1HHG) as the phasing model. The top five results for the rotation and translation functions are shown and the correct solution for each is highlighted in red. This is determined by a peak in the RF-function and TFcnt/Score values for the rotation and translation functions respectively.

Appendix F: HLA-A2-RQInp complex structure solution rotation and translation values

Rotation Function					
Peak Number	θ (°)	ϕ(°)	X (°)	RF-function	
1	162.51	90.88	55.07	21.06	
2	149.41	-167.57	69.20	4.57	
3	36.04	155.98	140.37	4.48	
4	86.41	52.14	33.96	4.47	
5	143.11	53.20	45.36	4.43	
Translation Function					
Peak Number	TX	TY	TZ	TFcnt	R-fac	Score
1	0.189	0.000	0.137	31.57	0.499	0.533
2	0.272	0.000	0.420	2.27	0.603	0.277
3	0.799	0.000	0.486	2.32	0.604	0.275
4	0.054	0.000	0.167	1.66	0.604	0.275
5	0.339	0.000	0.411	1.70	0.605	0.274

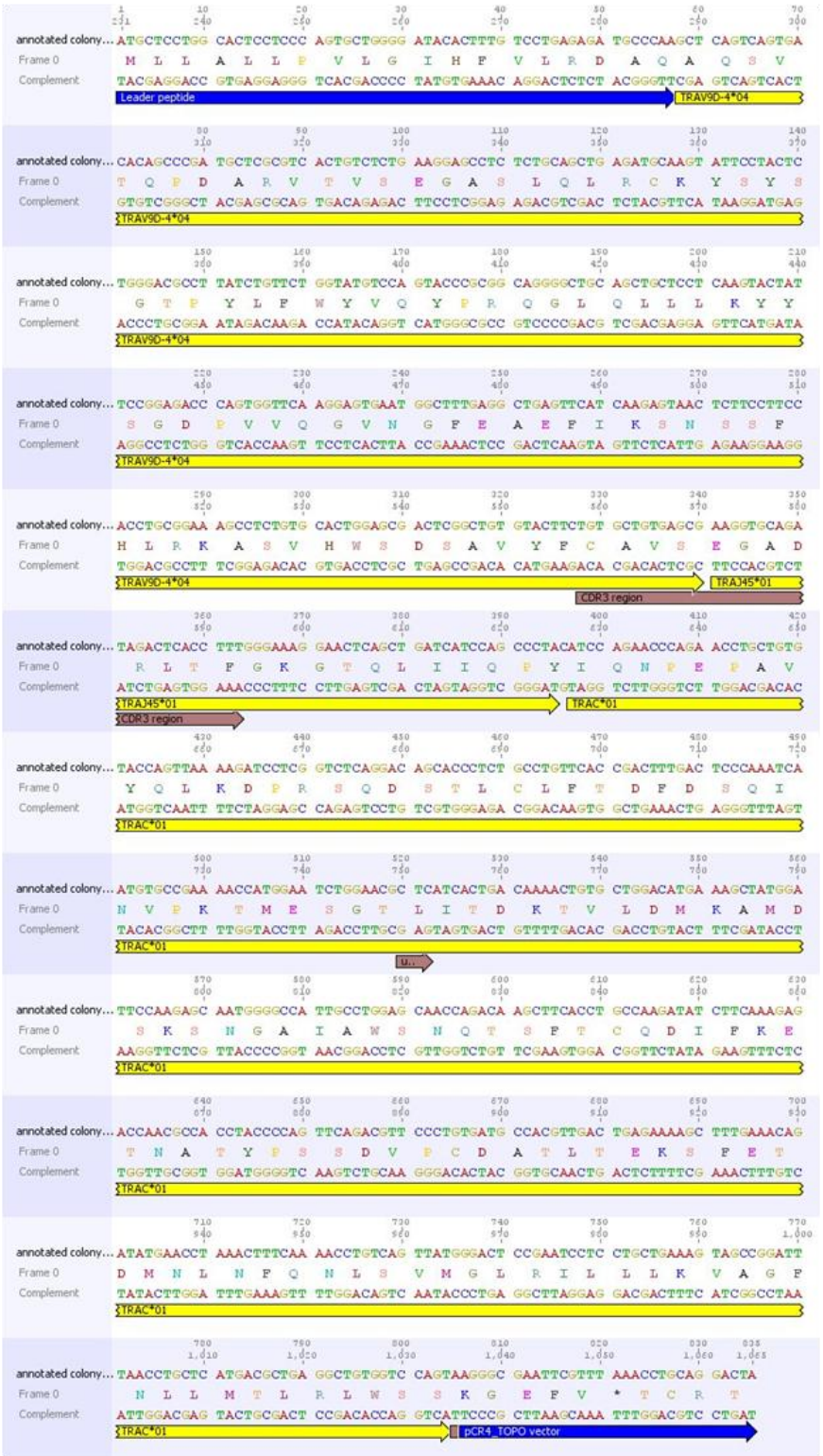
Initial phase information was determined by molecular replacement with MOLREP using the HLA-A2 structure (PDB ID code; 1HHG) as the phasing model. The top five results for the rotation and translation functions are shown and the correct solution for each is highlighted in red. This is determined by a peak in the RF-function and TFcnt/Score values for the rotation and translation functions respectively.

Appendix G: LILRB1-HLA-A2-RTFnp structure solution rotation and translation values

Rotation Function						
Peak Number	θ (°)	ϕ(°)	X (°)	RF-function		
1	90.00	-89.97	180.00	14.13		
2	95.63	-84.96	175.98	3.35		
3	94.44	-86.77	176.28	3.34		
4	45.40	54.84	136.08	3.34		
5	14.32	-179.15	49.67	3.26		
Translation Function 1						
Peak Number	TX	TY	TZ	TFcnt	R-fac	Score
1	0.734	0.228	0.019	10.84	0.507	0.360
2	0.489	0.109	0.432	3.55	0.607	0.087
3	0.833	0.406	0.057	2.91	0.605	0.087
4	0.878	0.361	0.158	4.51	0.606	0.087
5	0.761	0.649	0.214	2.53	0.609	0.083
Translation Function 2						
Peak Number	TX	TY	TZ	TFcnt	R-fac	Score
1	0.785	0.533	0.175	8.31	0.540	0.279
2	0.733	0.238	0.524	4.49	0.561	0.222
3	0.740	0.244	0.518	3.58	0.566	0.220
4	0.261	0.480	0.143	2.32	0.564	0.219
5	0.238	0.459	0.151	2.13	0.563	0.212

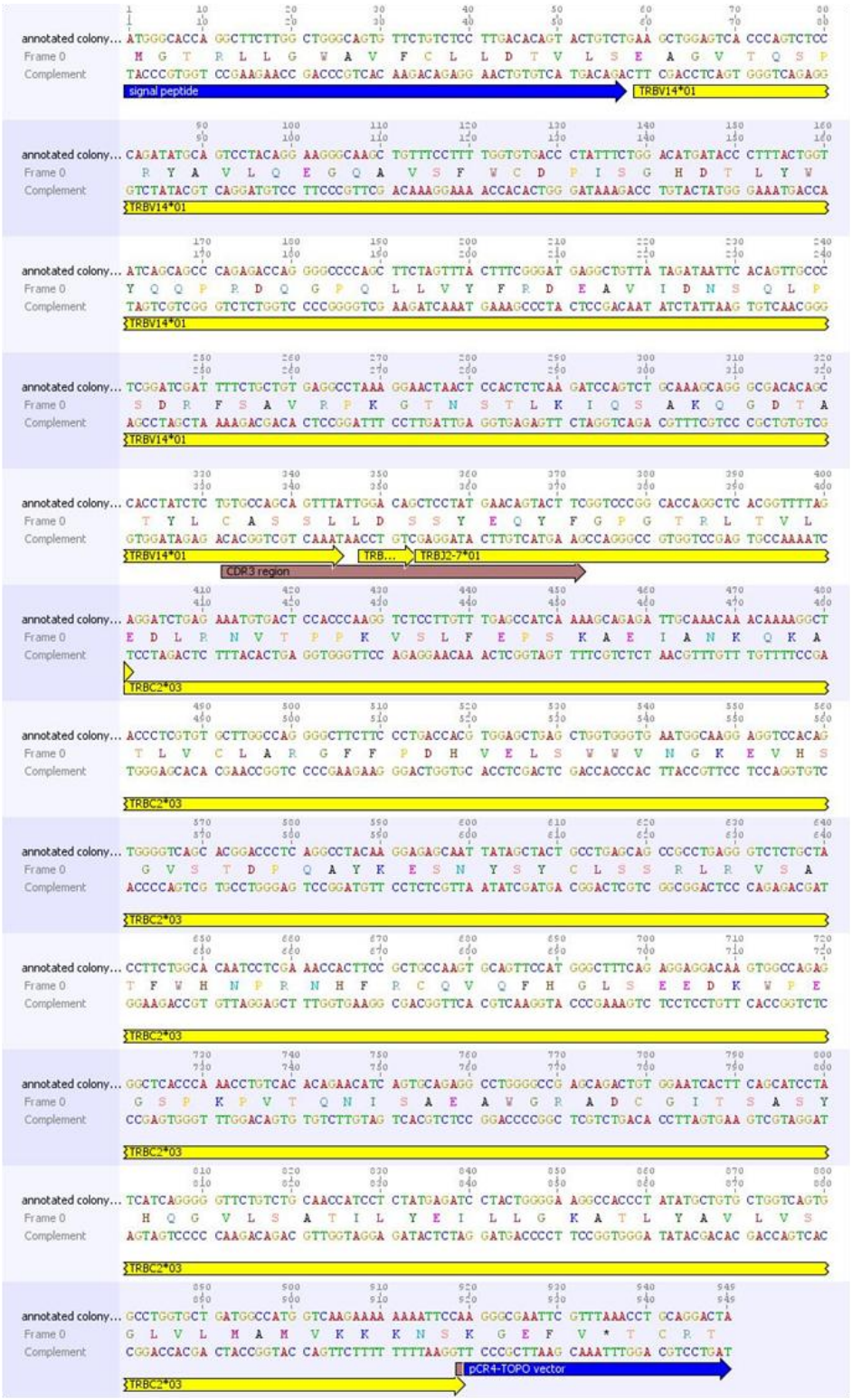
Initial phase information was determined by molecular replacement with MOLREP using the LILRB1-HLA-A2 structure (PDB ID code; 1P7Q) as the phasing model. The top five results for the rotation and translation functions are shown and the correct solution for each is highlighted in red. This is determined by a peak in the RF-function and TFcnt/Score values for the rotation and translation functions respectively. For the LILRB1-HLA-A2-RTFnp complex two translation functions were determined due to the presence of two molecules occupying the unit cell.

Appendix H: IRS2 TCR α chain gene sequence



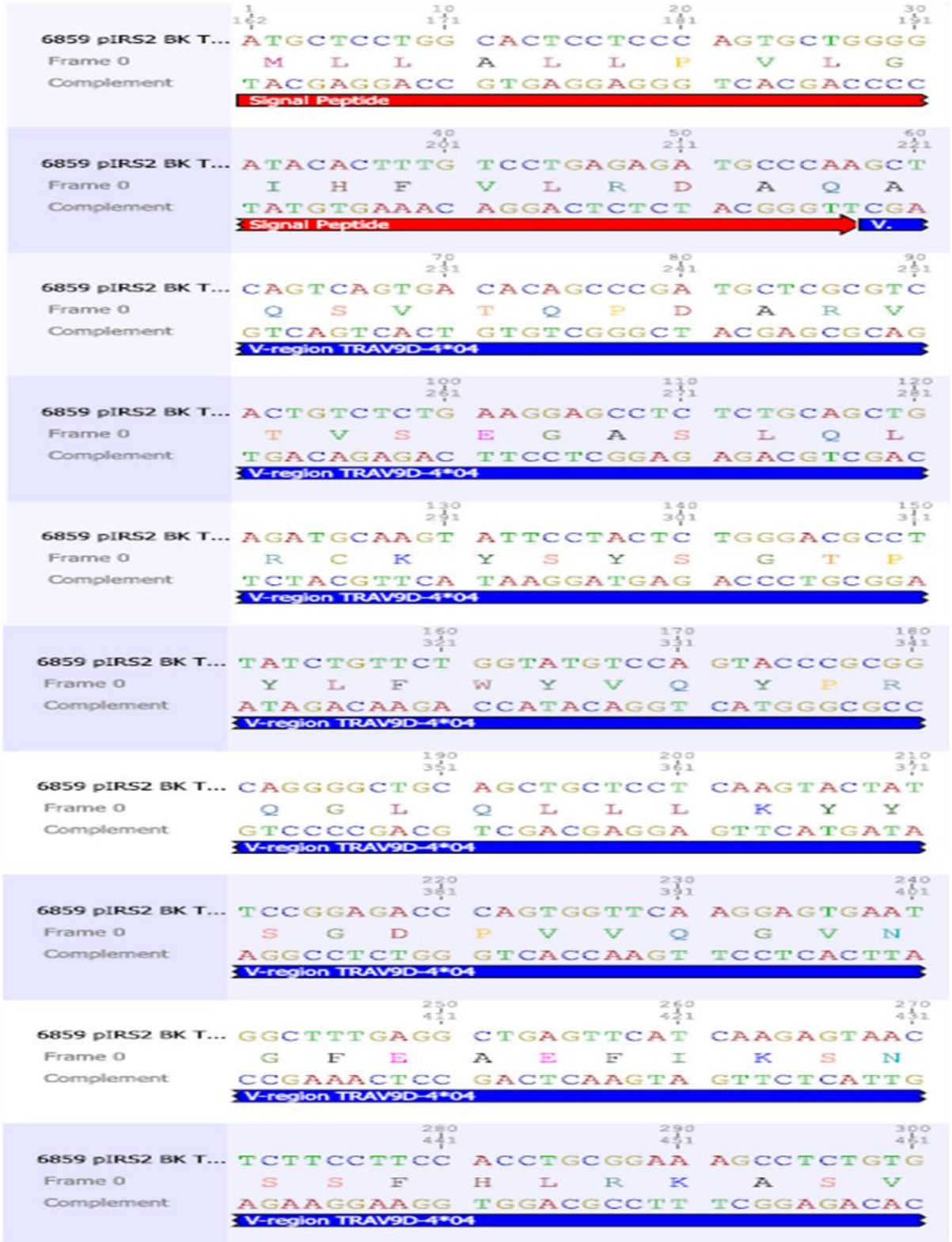
Gene sequence of IRS2α chain with Variable, Joining and Constant domains highlighted (kindly provided by Dr Angela Zarling).

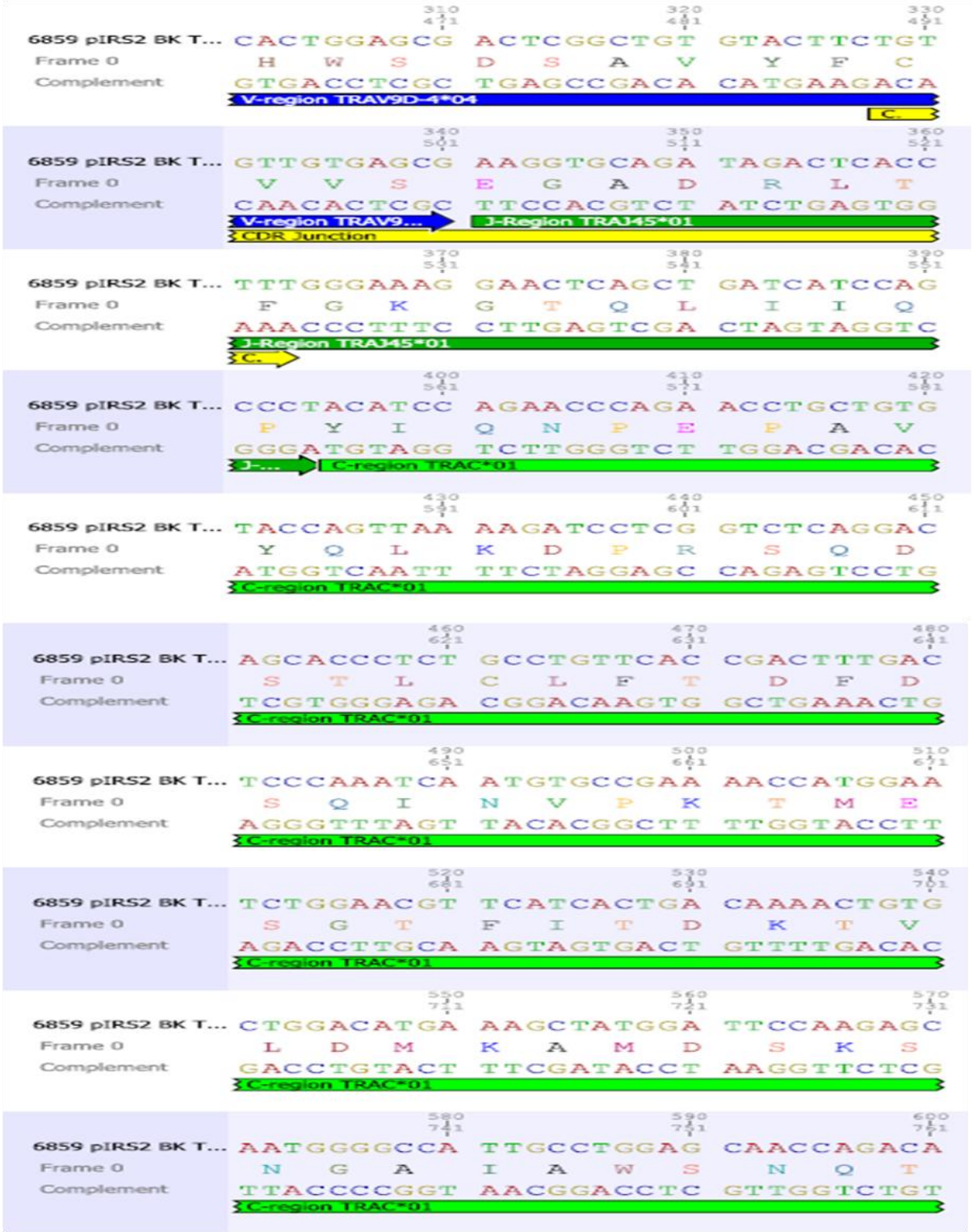
Appendix I: IRS2 TCR β chain gene sequence



Gene sequence of IRS2β chain with Variable, Joining and Constant regions highlighted (kindly provided by Dr Angela Zarling).

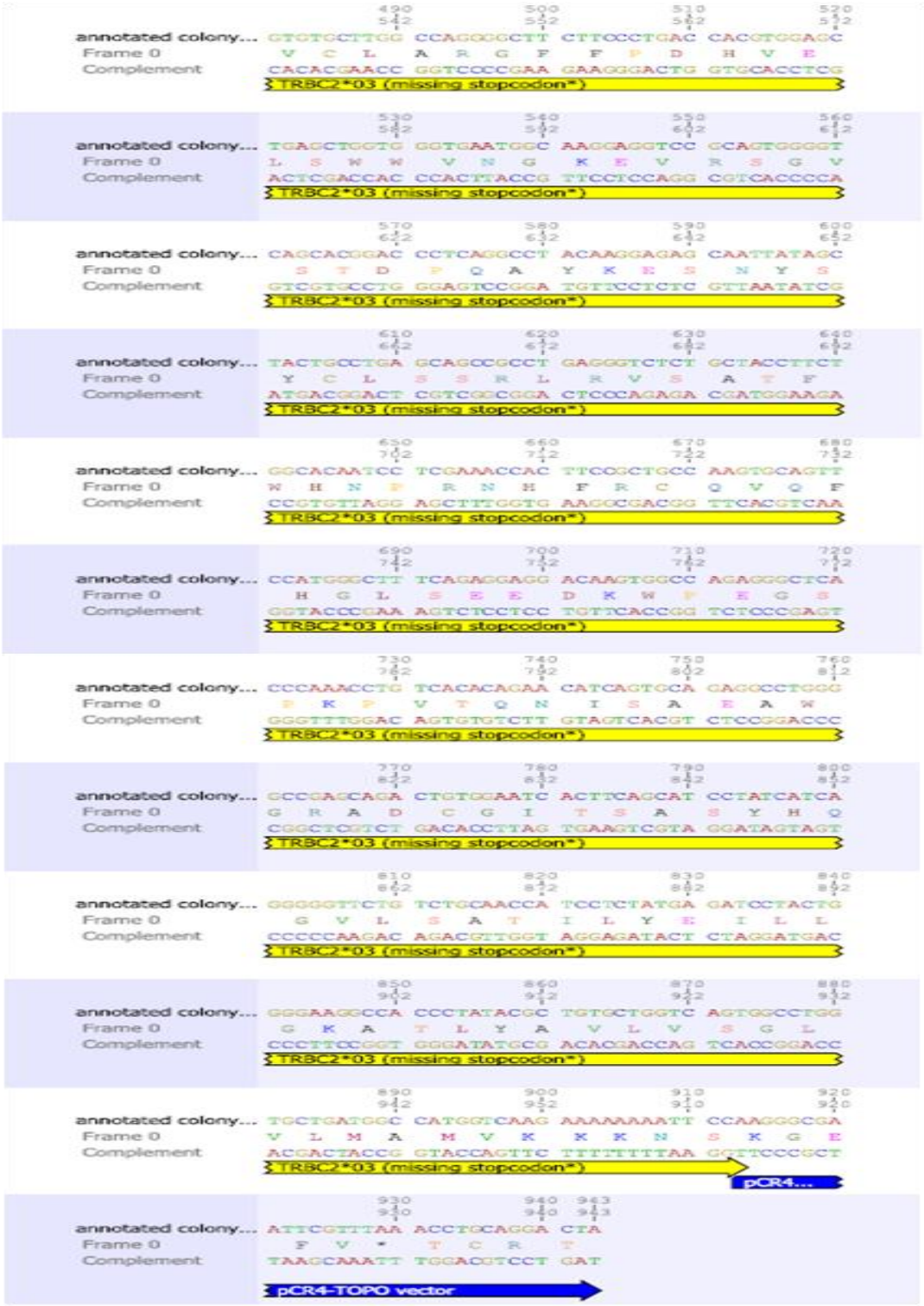
Appendix J: IRS2BK TCR α chain sequence




6859 pIRS2 BK T...
Frame 0
Complement
V-region TRAV9D-4*04
C...
V-region TRAV9...
J-Region TRAJ45*01
CDR Junction
J-Region TRAJ45*01
C...
J-...
C-region TRAC*01

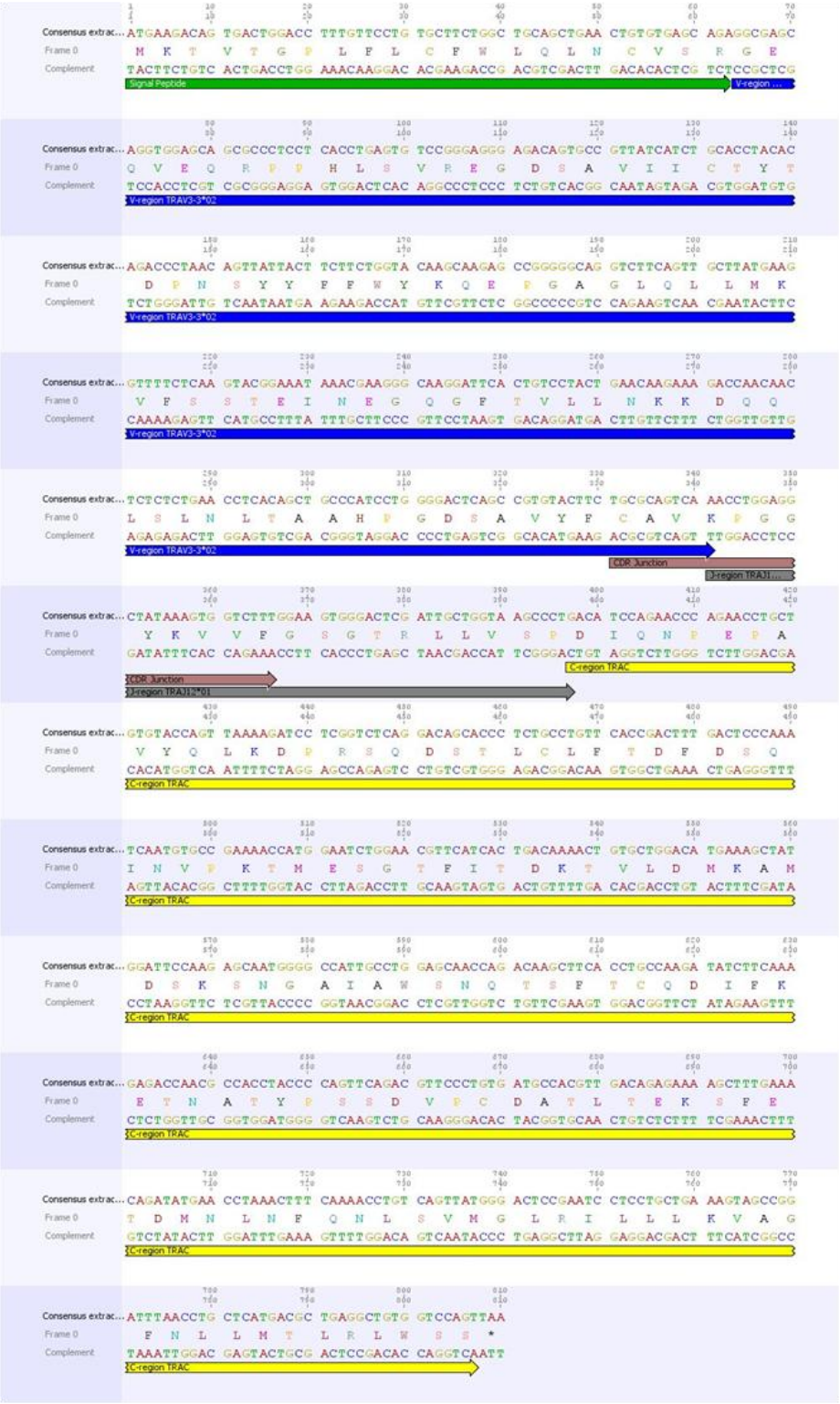
Gene sequence of IRS2BK α chain with Variable (Blue), Joining (Dark Green) and Constant (Light Green) regions highlighted (kindly provided by Dr Angela Zarling).

Appendix K: IRSBK TCR β chain sequence



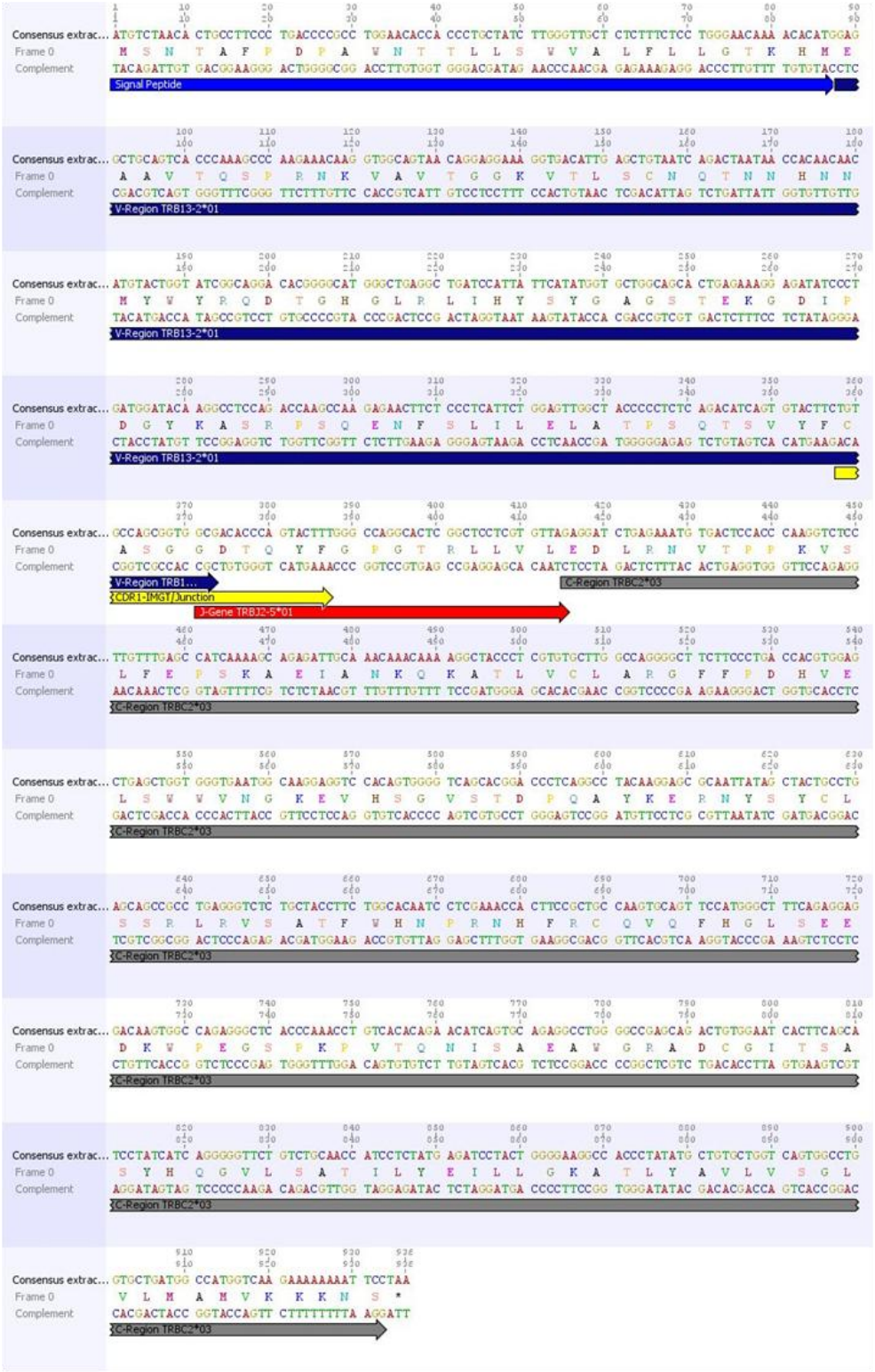
Gene sequence of IRS2BK β chain with Variable, Joining and Constant regions highlighted (kindly provided by Dr Angela Zarling).

Appendix L: CDC25B TCR α chain sequence



Gene sequence of CDC25B α chain with Variable (Blue), Joining (Grey) and Constant (Yellow) regions highlighted (kindly provided by Dr Angela Zarling).

Appendix N: CDC25B TCR β chain sequence



Gene sequence of CDC25B β chain with Variable (Dark Blue), Joining (Red) and Constant (Grey) regions highlighted (kindly provided by Dr Angela Zarling).

Appendix O

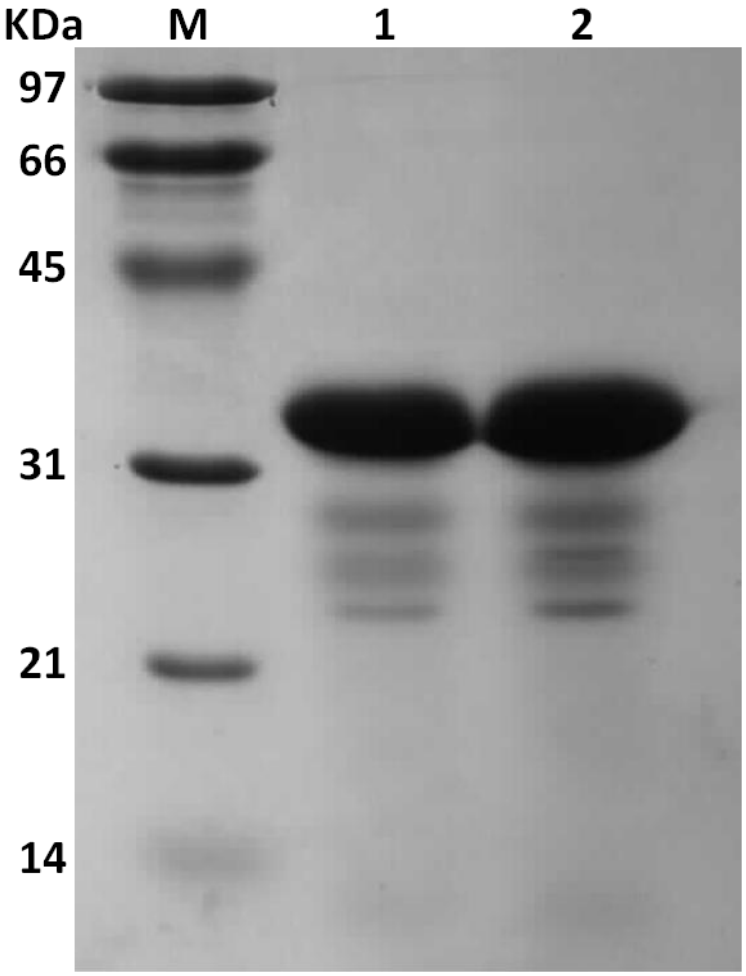
HLA-B7-phosphoeptide complex production

In order to generate soluble HLA-B7-phosphoeptide complexes use in this study two separate pet23a vectors incorporating either the HLA-B7 heavy chain or the β_2M - domain were generated prior to this study. Full expression was performed in *E. coli* BL21 (DE3) cells and inclusion bodies purified then solubilised in 8M urea as documented in chapter 3. For this study epitopes which were able to generate immune responses in ELISPOT assays of healthy human PBMC (personal communication from Dr Mark Cobbold) were selected. Initial attempts utilised standard dilution refolding (Garboczi, Hung et al. 1992) as used for the generation of HLA-A2 molecules documented in chapter 2, 3 and 4. However, purification by S200 gel filtration yielded a peak which, upon SDS-PAGE analysis, contained multiple contaminant bands and lacked stoichiometric levels of β_2M, indicating the presence of HLA-B7 heavy chain which was not in complex with β_2M (Apendix O Figure 1).

Previous studies of non-phosphorylated epitope presentation by HLA-B molecules have used various alterations to the levels of heavy-chain: β_2M: peptide ratio added to the refold mix as well as the length of time the refold mix was left before purification, indicating that the efficiency of complex refolding for HLA-B molecules may not be as high as HLA-A molecules (Macdonald, Williams et al. 2002). Attempts were therefore made to improve upon the refold efficiency but without increasing the levels of phosphopeptide too much in order to keep the cost of refolding low. Initial tests seemed to improve the refold efficiency, however when further purified by Resource-Q anion exchange this revealed multiple peaks due to

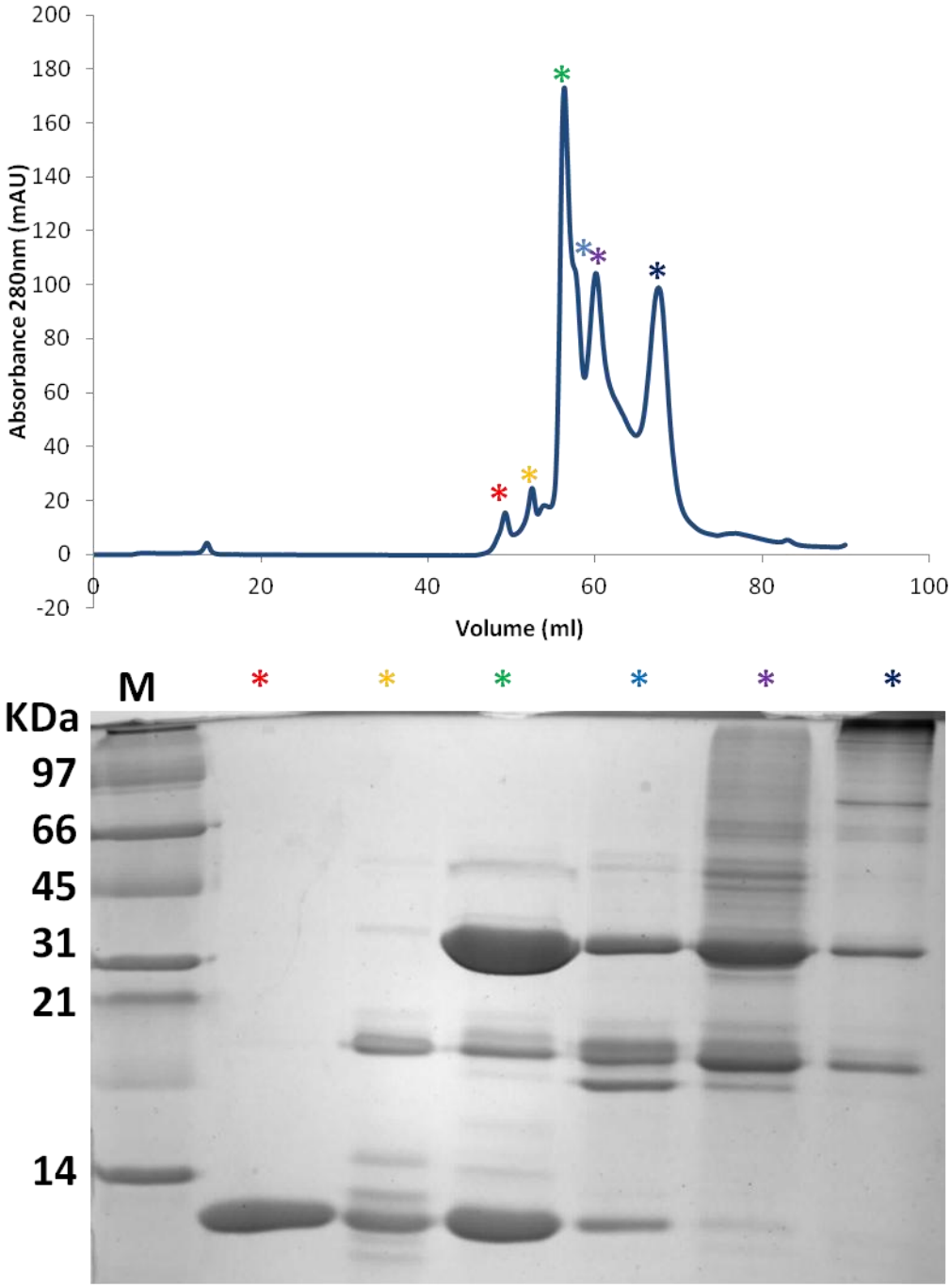
contaminating proteins such as free β2M and free HLA-B7 heavy chain (Appendix Figure 6.2). Following several rounds of refinement the most optimal refold protocol for generating HLA-B7-phosphopeptide complexes is documented in chapter 2. However although this method resulted in increased yields of stoichiometric HLA-B7-phosphoeptide complex (Appendix O Figure 3) due to the continued need for Resource-Q purification the yields were still too low for efficient use in crystallisation trials. Therefore further work is needed to improve upon this refold technique in order to provide enough soluble HLA-B7-phosphoeptide complex for use in crystallisation studies.

Appendix O Figure 1



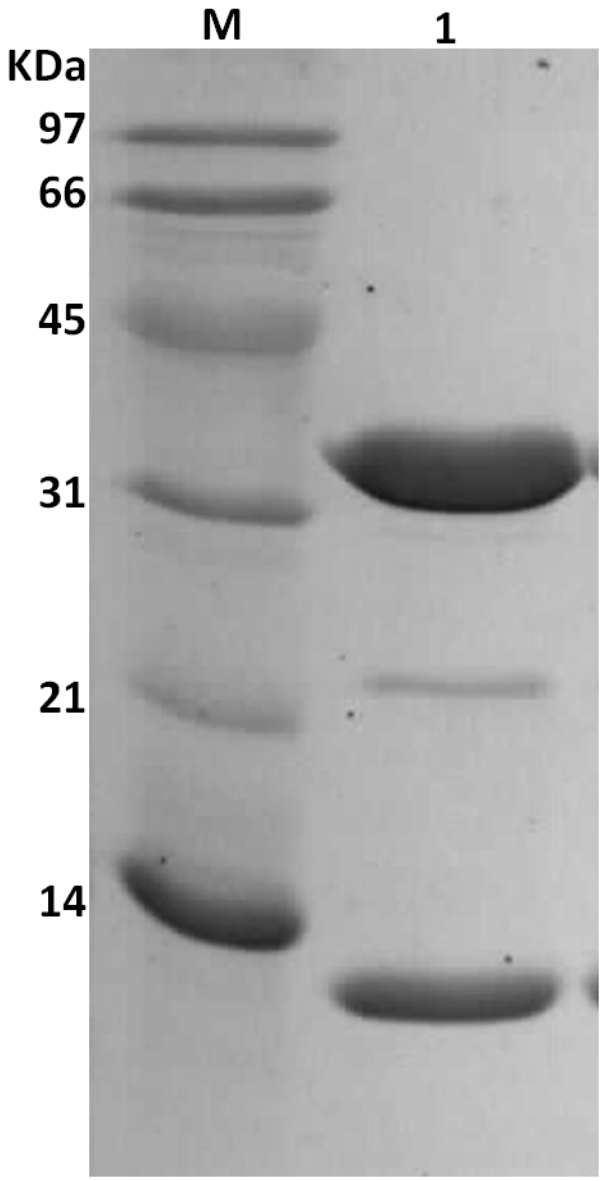
SDS-PAGE analysis of 2 HLA-B7-phosphoeptide complexes (Lane 1 HLA-B7-RPFpSPREAL, Lane 2 HLA-B7-RPApSAGAML) following standard dilution refolding and S200 gel filtration purification revealed that the complexes lacked stiochemetric β2m.

Appendix O Figure 2



Purification by Resource-Q anion exchange reveals that the single peak obtained from previous S200 gel filtration does not contain single HLA-B7 complex but also contains free β2M and HLA-B7 heavy chain only complex when assessed by SDS-PAGE under reducing conditions.

Appendix O Figure 3 Improved purity of HLA-B7-phosphoeptide complex



Optimisation of HLA-B7-phosphoeptide refolding yields significantly improved purity of HLA-B7 complex (Lane1) following Resource-Q anion exchange purification when assessed by SDS-PAGE under reducing conditions.

Appendix O Table 1 Affinity analysis of HLA-B7 phosphopeptides and their non-phosphorylated counterparts

Sequence	Source Protein	Peptide length (amino acids)	IC_{50} (nM) Nonphosphorylated peptide	IC_{50} (nM) Phosphorylated peptide	Fold increase
R**P**TpSRLNR**L**	Nuclear receptor co-activator 1	9	1.15	5.55	4.8
R**P**FpSPREA**L**	Leucine zipper protein 1	9	1.9	4.5	2.4
R**P**HpSPEKA**F**	ELG protein	9	1.75	17.5	10
K**P**RpSPVVE**L**	Beta adrenergic receptor kinase 1 variant	9	2.25	20	8.9
R**P**ApSAGAM**L**	Myocyte enhancer factor 2D	9	1.05	10	9.5
E**P**RpSPSHS**M**	Mixed-lineage leukemia	9	2.05	8.5	4.1
R**V**RpSPTRS**P**	Mixed-lineage leukemia	9	3.2	21	6.7

Competitive binding assays were performed as previously described (Mohammed, Cobbold et al. 2008). The peptide concentrations covered a 100,000 fold range. The concentration at which test peptide displaced 50% of the radiolabeled peptide (IC_{50}) was calculated allowing a reasonable approximation of the dissociation constant (Mohammed, Cobbold et al. 2008).

References

Abele, R. and R. Tampe (2004). "The ABCs of Immunology: Structure and Function of TAP, the Transporter Associated with Antigen Processing." Physiology **19**(4): 216-224.

Adami, J., H. Gabel, et al. (2003). "Cancer risk following organ transplantation: a nationwide cohort study in Sweden." Br J Cancer **89**(7): 1221-1227.

Adams, E. J., Y.-H. Chien, et al. (2005). "Structure of a gd T Cell Receptor in Complex with the Nonclassical MHC T22." Science 308(5719): 227-231.

Akira, S. and H. Hemmi (2003). "Recognition of pathogen-associated molecular patterns by TLR family." Immunology Letters **85**(2): 85-95.

Akira, S., K. Takeda, et al. (2001). "Toll-like receptors: critical proteins linking innate and acquired immunity." Nat Immunol **2**(8): 675-680.

Alberts, B., A. Johnson, et al. (2002). Molecular Biology of the cell. New York, Garland Science.

Anderson, G. and Y. Takahama (2012). "Thymic epithelial cells: working class heroes for T cell development and repertoire selection." Trends in Immunology **33**(6): 256-263.

Banchereau, J. and R. M. Steinman (1998). "Dendritic cells and the control of immunity." Nature **392**(6673): 245-252.

Baselga, J. (2011). "Targeting the Phosphoinositide-3 (PI3) Kinase Pathway in Breast Cancer." The Oncologist **16**(suppl 1): 12-19.

Bendle, G. M., C. Linnemann, et al. (2010). "Lethal graft-versus-host disease in mouse models of T cell receptor gene therapy." Nat Med **16**(5): 565-570.

Beyer, M. and J. L. Schultze (2006). "Regulatory T cells in cancer." Blood **108**(3): 804-811.

Bianchi, M. E. (2007). "DAMPs, PAMPs and alarmins: all we need to know about danger." J Leukoc Biol **81**(1): 1-5.

Biot, C., C. A. Rentsch, et al. (2012). "Preexisting BCG-Specific T Cells Improve Intravesical Immunotherapy for Bladder Cancer." Science Translational Medicine **4**(137): 137ra172.

Bjorkman, P. J., M. A. Saper, et al. (1987). "Structure of the Human Class-I Histocompatibility Antigen, Hla-A2." Nature **329**(6139): 506-512.

Bjorkman, P. J., M. A. Saper, et al. (1987). "The Foreign Antigen-Binding Site and T-Cell Recognition Regions of Class-I Histocompatibility Antigens." Nature **329**(6139): 512-518.

Blackman, M., J. Kappler, et al. (1990). "The role of the T cell receptor in positive and negative selection of developing T cells." Science **248**(4961): 1335-1341.

Borbulevych, O. Y., T. K. Baxter, et al. (2005). "Increased Immunogenicity of an Anchor-Modified Tumor-Associated Antigen Is Due to the Enhanced Stability of the Peptide/MHC Complex: Implications for Vaccine Design." The Journal of Immunology **174**(8): 4812-4820.

Brunger, A. T. (1992). "Free R value: a novel statistical quantity for assessing the accuracy of crystal structures." Nature **355**(6359): 472-475.

Brunger, A. T., P. D. Adams, et al. (1998). "Crystallography & NMR system: A new software suite for macromolecular structure determination." Acta Crystallogr D Biol Crystallogr **54**(Pt 5): 905-21.

Bubeník, J. (2004). "MHC class I down-regulation: tumour escape from immune surveillance?" International Journal of Oncology **25**(2): 487-491.

Burnet, M. (1957). "Cancer—A Biological Approach." BMJ **1**(5023): 841-847.

Cabaniols, J.-P., N. Fazilleau, et al. (2001). "Most α/β T Cell Receptor Diversity Is Due to Terminal Deoxynucleotidyl Transferase." The Journal of Experimental Medicine **194**(9): 1385-1390.

Cantrell, D. A. (2002). "T-cell antigen receptor signal transduction." Immunology **105**(4): 369-374.

Carreno, B. M. and M. Collins (2002). "THE B7 FAMILY OF LIGANDS AND ITS RECEPTORS: New Pathways for Costimulation and Inhibition of Immune Responses." Annual Review of Immunology **20**(1): 29-53.

Chang, H. C., Z. Bao, et al. (1994). "A general method for facilitating heterodimeric pairing between two proteins: application to expression of alpha and beta T-cell receptor extracellular segments." Proceedings of the National Academy of Sciences 91(24): 11408-11412.

Chang, S., F. Momburg, et al. (2005). "The ER aminopeptidase, ERAP1, trims precursors to lengths of MHC class I peptides by a ‘‘molecular ruler’’ mechanism." PNAS **102**(47): 17107–17112.

Cheng, M. H., A. K. Shum, et al. (2007). "What's new in the Aire?" Trends in Immunology **28**(7): 321-327.

Clemente, C. G., M. C. Mihm, et al. (1996). "Prognostic value of tumor infiltrating lymphocytes in the vertical growth phase of primary cutaneous melanoma." Cancer **77**(7): 1303-1310.

Collaborative Computational Project No 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50:** 760-763.

Cresswell, P., A. L. Ackerman, et al. (2005). "Mechanisms of MHC class I-restricted antigen processing and cross-presentation." Immunological Reviews **207**(1): 145-157.

Curtsinger, J. M., C. M. Johnson, et al. (2003). "CD8 T Cell Clonal Expansion and Development of Effector Function Require Prolonged Exposure to Antigen, Costimulation, and Signal 3 Cytokine." The Journal of Immunology **171**(10): 5165-5171.

Davis, M., J. Boniface, et al. (1998). "Ligand recognition by alpha beta T cell receptors " Annual Review of Immunology **16**: 523-544.

Di Santo, J. P. (2006). Natural killer cell developmental pathways: A question of balance. Annual Review of Immunology. Palo Alto, Annual Reviews. **24:** 257-286.

Dighe, A. S., E. Richards, et al. (1994). "Enhanced in vivo growth and resistance to rejection of tumor cells expressing dominant negative IFNγ receptors." Immunity **1**(6): 447-456.

Dudley, M. E. and S. A. Rosenberg (2003). "Adoptive-cell-transfer therapy for the treatment of patients with cancer." Nat Rev Cancer **3**(9): 666-675.

Dudley, M. E., J. R. Wunderlich, et al. (2002). "Cancer Regression and Autoimmunity in Patients After Clonal Repopulation with Antitumor Lymphocytes." Science **298**(5594): 850-854.

Dunn, G. P., A. T. Bruce, et al. (2002). "Cancer immunoediting: from immunosurveillance to tumor escape." Nat Immunol **3**(11): 991-998.

Dunn, G. P., L. J. Old, et al. (2004). "The Three Es of Cancer Immunoediting." Annual Review of Immunology **22**(1): 329-360.

Ehrlich, P. (1909). "Ueber den jetzigen stand der Karzinomforschung." Ned. Tijdschr. Geneeskd. **5**: 273-290.

Elliott, T. and A. Williams (2005). "The optimization of peptide cargo bound to MHC class I molecules by the peptide-loading complex." Immunological Reviews **207**(1): 89-99.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2126-32.

Engelhard, V. H. (1994). "Structure of peptides associated with MHC class I molecules." Current Opinion in Immunology **6**(1): 13-23.

Fazilleau, N., L. J. McHeyzer-Williams, et al. (2007). "Local development of effector and memory T helper cells." Current Opinion in Immunology **19**(3): 259-267.

Fenn, T. D., D. Ringe, et al. (2003). "POVScript+: a program for model and data visualization using persistence of vision ray-tracing." Journal of Applied Crystallography **36**(3 Part 2): 944-947.

Figdor, C. G., I. J. M. de Vries, et al. (2004). "Dendritic cell immunotherapy: mapping the way." Nat Med **10**(5): 475-480.

Foss, F. M. (2002). "Immunologic mechanisms of antitumor activity." Seminars in Oncology **29**(3, Supplement 7): 5-11.

Furuzawa-Carballeda, J., M. I. Vargas-Rojas, et al. (2007). "Autoimmune inflammation from the Th17 perspective." Autoimmunity Reviews **6**(3): 169-175.

Gabrilovich, D. I., H. L. Chen, et al. (1996). "Production of vascular endothelial growth factor by human tumors inhibits the functional maturation of dendritic cells." Nat Med **2**(10): 1096-1103.

Galon, J., A. Costes, et al. (2006). "Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome." Science **313**(5795): 1960-1964.

Gao, G. F., U. C. Gerth, et al. (1998). "Assembly and crystallization of the complex between the human T cell coreceptor CD8a homodimer and HLA-A2." Protein Science **7**: 1245-1249.

Garboczi, D. N., D. T. Hung, et al. (1992). "HLA-A2-Peptide Complexes: Refolding and Crystallization of Molecules Expressed in Escherichia coli and Complexed with Single Antigenic Peptides." Proceedings of the National Academy of Sciences **89**(8): 3429-3433.

Gatza, E. and C. Y. Okada (2002). "Tumor Cell Lysate-Pulsed Dendritic Cells Are More Effective Than TCR Id Protein Vaccines for Active Immunotherapy of T Cell Lymphoma." J Immunol **169**(9): 5227-5235.

Gherardi, E., W. Birchmeier, et al. (2012). "Targeting MET in cancer: rationale and progress." Nat Rev Cancer **12**(2): 89-103.

Gilboa, E. (2007). "DC based cancer vaccines." The Journal of Clinical Investigation **117**(5): 1195–1203.

Gorer, P. A. and N. Kaliss (1959). "The Effect of Isoantibodies in Vivo on Three Different Transplantable Neoplasms in Mice." Cancer Res **19**: 824-830.

Gowans, J. L. and J. W. Uhr (1966). "THE CARRIAGE OF IMMUNOLOGICAL MEMORY BY SMALL LYMPHOCYTES IN THE RAT." J. Exp. Med. **124**(5): 1017-1030.

Gray, D. (2002). "A role for antigen in the maintenance of immunological memory." Nat Rev Immunol **2**(1): 60-65.

Greenman, C., P. Stephens, et al. (2007). "Patterns of somatic mutation in human cancer genomes." Nature **446**(7132): 153-158.

Griffin, T. A., D. Nandi, et al. (1998). "Immunoproteasome Assembly: Cooperative Incorporation of Interferon gamma (IFN-gamma)-inducible Subunits." J. Exp. Med. **187**(1): 97-104.

Guermonprez, P., J. Valladeau, et al. (2002). "ANTIGEN PRESENTATION AND T CELL STIMULATION BY DENDRITIC CELLS." Annual Review of Immunology **20**: 621-667.

Hahne, M., D. Rimoldi, et al. (1996). "Melanoma Cell Expression of Fas(Apo-1/CD95) Ligand: Implications for Tumor Immune Escape." Science **274**(5291): 1363-1366.

Hanahan, D. and R. A. Weinberg (2000). "The Hallmarks of Cancer." Cell **100**(1): 57-70.

Hanahan, D. and Robert A. Weinberg (2011). "Hallmarks of Cancer: The Next Generation." Cell **144**(5): 646-674.

Hao, B., S. Oehlmann, et al. (2007). "Structure of a Fbw7-Skp1-Cyclin E Complex: Multisite-Phosphorylated Substrate Recognition by SCF Ubiquitin Ligases." Molecular Cell **26**(1): 131-143.

Harper, J. W. (2002). "A phosphorylation-driven ubiquitination switch for cell-cycle control." Trends in Cell Biology **12**(3): 104-107.

Hershko, A. and A. Ciechanover (1998). "The Ubiquitin System." Annual Review of Biochemistry **67**(1): 425-479.

Hogan, K. T., D. P. Eisinger, et al. (1998). "The Peptide Recognized by HLA-A68.2-restricted, Squamous Cell Carcinoma of the Lung-specific Cytotoxic T Lymphocytes Is Derived from a Mutated Elongation Factor 2 Gene." Cancer Res **58**(22): 5144-5150.

Hooft, R. W., G. Vriend, et al. (1996). "Errors in protein structures." Nature **381**: 272.

Irvine, K. R., M. R. Parkhurst, et al. (1999). "Recombinant Virus Vaccination against “Self” Antigens UsingAnchor-fixed Immunogens." Cancer Research **59**(11): 2536-2540.

Jiricny, J. and G. Marra (2003). "DNA repair defects in colon cancer." Current Opinion in Genetics & Development **13**(1): 61-69.

Kabsch, W. (1993). "Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants." Journal of Applied Crystallography **26**(6): 795-800.

Kantoff, P. W., C. S. Higano, et al. (2010). "Sipuleucel-T Immunotherapy for Castration-Resistant Prostate Cancer." N Engl J Med **363**: 411-422.

Kaplan, D. H., V. Shankaran, et al. (1998). "Demonstration of an interferon γ-dependent tumor surveillance system in immunocompetent mice." Proceedings of the National Academy of Sciences **95**(13): 7556-7561.

Kataki, A., P. Scheid, et al. (2002). "Tumor infiltrating lymphocytes and macrophages have a potential dual role in lung cancer by supporting both host-defense and tumor progression." Journal of Laboratory and Clinical Medicine **140**(5): 320-328.

Khong, H. T. and N. P. Restifo (2002). "Natural selection of tumor variants in the generation of "tumor escape" phenotypes." Nat Immunol **3**(11): 999-1005.

Kim, R., M. Emi, et al. (2007). "Cancer immunoediting from immune surveillance to immune escape." Immunology **121**(1): 1-14.
Kinzler, K. W. and B. Vogelstein (1996). "Lessons from Hereditary Colorectal Cancer." Cell **87**(2): 159-170.
Kochan, G., T. Krojer, et al. (2011). "Crystal structures of the endoplasmic reticulum aminopeptidase-1 (ERAP1) reveal the molecular basis for N-terminal peptide trimming." Proceedings of the National Academy of Sciences.
Krivtsov, A. V. and S. A. Armstrong (2007). "MLL translocations, histone modifications and leukaemia stem-cell development." Nat Rev Cancer **7**(11): 823-833.
Kurts, C., H. Kosaka, et al. (1997). "Class I–restricted Cross-Presentation of Exogenous Self-Antigens Leads to Deletion of Autoreactive CD8+ T Cells." JEM **186**(2): 239-245.
Lee, S., B. Park, et al. (2009). "Redox-Regulated Export of the MHC Class I-Peptide Complexes from the Endoplasmic Reticulum." Mol. Biol. Cell: E09-03-0238.
Lesokhin, A. M., T. M. Hohl, et al. (2012). "Monocytic CCR2+ Myeloid-Derived Suppressor Cells Promote Immune Escape by Limiting Activated CD8 T-cell Infiltration into the Tumor Microenvironment." Cancer Research **72**(4): 876-886.
Liu, Y.-J. (2001). "Dendritic Cell Subsets and Lineages, and Their Functions in Innate and Adaptive Immunity." Cell **106**(3): 259-262.
Los, M., J. M. L. Roodhart, et al. (2007). "Target Practice: Lessons from Phase III Trials with Bevacizumab and Vatalanib in the Treatment of Advanced Colorectal Cancer." The Oncologist **12**(4): 443-450.
Lutz, M. B. and G. Schuler (2002). "Immature, semi-mature and fully mature dendritic cells: which signals induce tolerance or immunity?" Trends in Immunology **23**(9): 445-449.
Madden, D. R. (1995). "The Three-Dimensional Structure of Peptide-MHC Complexes." Annual Review of Immunology **13**(1): 587-622.
Madden, D. R., D. N. Garboczi, et al. (1993). "The antigenic identity of peptide-MHC complexes: A comparison of the conformations of five viral peptides presented by HLA-A2." Cell **75**(4): 693-708.
Macdonald, W., D. S. Williams, et al. (2002). "Identifcation of a dominant self-ligand bound to three HLA B44 alleles and the preliminary crystallographic analysis of recombinant forms of each complex." FEBS Lett. **527**: 27-32.
Mapara, M. Y. and M. Sykes (2004). "Tolerance and Cancer: Mechanisms of Tumor Evasion and Strategies for Breaking Tolerance." Journal of Clinical Oncology **22**(6): 1136-1151.
Mathis, D. and C. Benoist (2009). "Aire." Annual Review of Immunology **27**(1): 287-312.
Melief, C., J.M. and S. Van der Burg, H. (2008). "Immunotherapy of established (pre)malignant disease by synthetic long peptide vaccines." Nature Reviews Cancer **8**: 351-360.
Micheletti, Bazzaro, et al. (1999). "The lifespan of major histocompatibility complex class I/peptide complexes determines the efficiency of cytotoxic T-lymphocyte responses." Immunology **96**(3): 411-415.
Michelotti, G. A., D. T. Price, et al. (2000). "α1-Adrenergic receptor regulation: basic science and clinical implications." Pharmacology & Therapeutics **88**(3): 281-309.
Minami, R., A. Hayakawa, et al. (2010). "BAG-6 is essential for selective elimination of defective proteasomal substrates." The Journal of Cell Biology.

Mohammed, F., M. Cobbold, et al. (2008). "Phosphorylation-dependant interaction of antigenic peptides with the class I MHC molecule HLA-A2." Nature Immunology **9**(11): 1236-1243.

Morel, P. A. and M. S. Turner (2010). "Designing the Optimal Vaccine: the Importance of Cytokines and Dendritic Cells." The Open Vaccine Journal **3**: 7-17.

Morse, M. A., R. E. Coleman, et al. (1999). "Migration of Human Dendritic Cells after Injection in Patients with Metastatic Malignancies." Cancer Research **59**(1): 56-58.

Murshudov, G. N., A. A. Vagin, et al. (1997). "Refinement of Macromolecular Structures by the Maximum-Likelihood method." Acta Cryst **D53**: 240-255.

Nemazee, D. (2006). "Receptor editing in lymphocyte development and central tolerance." Nat Rev Immunol **6**(10): 728-740.

Neurath, M. F., S. Finotto, et al. (2002). "The role of Th1/Th2 polarization in mucosal immunity." Nat Med **8**(6): 567-573.

Nicholls, A., K. A. Sharp, et al. (1991). "Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons." Proteins: Structure, Function, and Genetics **11**(4): 281-296.

Parker, K. C., M. A. Bednarek, et al. (1992). "Sequence motifs important for peptide binding to the human MHC class I molecule, HLA-A2." J Immunol **149**(11): 3580-3587.

Paul, W. E. and J. Zhu (2010). "How are TH2-type immune responses initiated and amplified?" Nat Rev Immunol **10**(4): 225-235.

Peltomaki, P. (2003). "Role of DNA Mismatch Repair Defects in the Pathogenesis of Human Cancer." Journal of Clinical Oncology **21**(6): 1174-1179.

Penn, I. (1996). "Malignant melanoma in organ alograft recipients." Transplantation **61**(2): 274-278.

Peters, J. M., W. W. Franke, et al. (1994). "Distinct 19 S and 20 S subcomplexes of the 26 S proteasome and their distribution in the nucleus and the cytoplasm." J. Biol. Chem. **269**(10): 7709-7718.

Petersen, J., S. J. Wurzbacher, et al. (2009). "Phosphorylated self-peptides alter human leukocyte antigen class I-restricted antigen presentation and generate tumor-specific epitopes." Proceedings of the National Academy of Sciences **106**(8): 2776-2781.

Peterson, P. A., B. A. Cunningham, et al. (1972). "{beta} 2-microglobulin--A Free Immunoglobulin Domain." Proceedings of the National Academy of Sciences **69**(7): 1697-1701.

Pickart, C. M. (1997). "Targeting of substrates to the 26S proteasome." FASEB J. **11**(13): 1055-1066.

Pickart, C. M. (2001). "MECHANISMS UNDERLYING UBIQUITINATION." Annual Review of Biochemistry **70**(1): 503-533.

Pobezinsky, L. A., G. S. Angelov, et al. (2012). "Clonal deletion and the fate of autoreactive thymocytes that survive negative selection." Nat Immunol **13**(6): 569-578.

Procko, E., M. L. O'Mara, et al. (2009). "The mechanism of ABC transporters: general lessons from structural and functional studies of an antigenic peptide transporter." The FASEB Journal **23**(5): 1287-1302.

Purcell, A. W. and T. Elliott (2008). "Molecular machinations of the MHC-I peptide loading complex." Current Opinion in Immunology **20**(1): 75-81.

Radoja, S. and A. B. Frey (2000). "Cancer-induced Defective Cytotoxic T Lymphocyte Effector Function: Another Mechanism How Antigenic Tumors Escape Immune-mediated Killing." Molecular Medicine **6**(6): 465-479.

Rajewsky, K. (1996). "Clonal selection and learning in the antibody system." Nature **381**(6585): 751-758.

Reits, E. A. J., A. C. Griekspoor, et al. (2000). "How does TAP pump peptides? Insights from DNA repair and traffic ATPases." Trends in Immunology **21**(12): 598-600.

Restifo, N. P., M. E. Dudley, et al. (2012). "Adoptive immunotherapy for cancer: harnessing the T cell response." Nat Rev Immunol **12**(4): 269-281.

Riddell, S. R., C. Berger, et al. (2003). "The graft versus leukemia response after allogeneic hematopoietic stem cell transplantation." Blood Reviews **17**(3): 153-162.

Rivett, A. J. and A. R. Hearn (2004). "Proteasome Function in Antigen Presentation: Immunoproteasome Complexes, Peptide Production, and Interactions with Viral Proteins." Current Protein and Peptide Science **5**: 153-161.

Rocchia, W., A. Sridharan, et al. (2002). "Rapid grid-based construction of the molecular surface for both molecules and geometric objects: applications to the finite difference Poisson-Boltzmann method." J Comp Chem **23**: 128-137.

Rocchia, W., E. Alexov, et al. (2001). "Extending the applicability of the non-linear Poisson-Boltzmann equation: multiple dielectric constants and multivalent ions." J Phys Chem B **105**: 6507-14.

Rock, K. L. and L. Shen (2005). "Cross-presentation: underlying mechanisms and role in immune surveillance." Immunological Reviews **207**(1): 166-183.

Romagnani, S. (2004). "Immunologic influences on allergy and the TH1/TH2 balance." The Journal of allergy and clinical immunology **113**(3): 395-400.

Ropponen, K. M., M. J. Eskelinen, et al. (1997). "Prognostic value of tumour-infiltrating lymphocytes (TILs) in colorectal cancer." The Journal of Pathology **182**(3): 318-324.

Rothenberg, E. V., J. E. Moore, et al. (2008). "Launching the T-cell-lineage developmental programme." Nat Rev Immunol **8**(1): 9-21.

Sadasivan, B., P. J. Lehner, et al. (1996). "Roles for Calreticulin and a Novel Glycoprotein, Tapasin, in the Interaction of MHC Class I Molecules with TAP." Immunity **5**(2): 103-114.

Sancho, D., O. P. Joffre, et al. (2009). "Identification of a dendritic cell receptor that couples sensing of necrosis to immunity." Nature **458**(7240): 899-903.

Saper, M. A., P. J. Bjorkman, et al. (1991). "Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution." Journal of Molecular Biology **219**(2): 277-319.

Scanlan, M. J., A. O. Gure, et al. (2002). "Cancer/testis antigens: an expanding family of targets for cancer immunotherapy." Immunological Reviews **188**(1): 22-32.

Scheerlinck, J.-P. Y., G. Casey, et al. (2001). "The immune response to a DNA vaccine can be modulated by co-delivery of cytokine genes using a DNA prime-protein boost strategy." Vaccine **19**(28-29): 4053-4060.

Schrodinger, LLC (2010). The PyMOL Molecular Graphics System, Version 1.3r1.

Schumacher, K., W. Haensch, et al. (2001). "Prognostic Significance of Activated CD8+ T Cell Infiltrations within Esophageal Carcinomas." Cancer Research **61**(10): 3932-3936.

Scott, S. D. (1998). "Rituximab: A New Therapeutic Monoclonal Antibody for Non-Hodgkin's Lymphoma." Cancer Practice **6**(3): 195-197.

Shankaran, V., H. Ikeda, et al. (2001). "IFN[gamma] and lymphocytes prevent primary tumour development and shape tumour immunogenicity." Nature **410**(6832): 1107-1111.

Sheil, R., A. G. (1986). "Cancer after transplantation." World Journal of Surgery **10**(3): 389-396.

Shen, L., L. J. Sigal, et al. (2004). "Important Role of Cathepsin S in Generating Peptides for TAP-Independent MHC Class I Crosspresentation In Vivo." Immunity **21**(2): 155-165.

Siguo Hao, O. B. F. L. J. Y. S. L. J. X. (2007). "Mature dendritic cells pulsed with exosomes stimulate efficient cytotoxic T-lymphocyte responses and antitumour immunity." Immunology **120**(1): 90-102.

Skokos, D., G. Shakhar, et al. (2007). "Peptide-MHC potency governs dynamic interactions between T cells and dendritic cells in lymph nodes." Nat Immunol **8**(8): 835-844.

Stanislawski, T., R.-H. Voss, et al. (2001). "Circumventing tolerance to a human MDM2-derived tumor antigen by TCR gene transfer." Nat Immunol **2**(10): 962-970.

Stockinger, B. and M. Veldhoen (2007). "Differentiation and function of Th17 T cells." Current Opinion in Immunology **19**(3): 281-286.

Strasser, A., H. Puthalakath, et al. (2007). "What do we know about the mechanisms of elimination of autoreactive T and B cells and what challenges remain." Immunol Cell Biol **86**(1): 57-66.

Strehl, B., U. Seifert, et al. (2005). "Interferon-gamma: The functional plasticity of the ubiquitin-proteasome system, and MHC class I antigen processing." Immunological Reviews **207**(1): 19-30.

Stutman, O. (1974). "Tumor Development after 3-Methylcholanthrene in Immunologically Deficient Athymic-Nude Mice." Science **183**(4124): 534-536.

Swann, J. B. and M. J. Smyth (2007). "Immune surveillance of tumors." Journal of Clinical Investigation **117**(5): 1137-1146.

Szabo, S. J., B. M. Sullivan, et al. (2003). "MOLECULAR MECHANISMS REGULATING TH1 IMMUNE RESPONSES." Annual Review of Immunology **21**: 713-758.

Teng, G. and F. N. Papavasiliou (2007). "Immunoglobulin Somatic Hypermutation." Annual Review of Genetics **41**: 107-120.

Thomas, L. (1959). Cellular and Humoral Aspects of the Hypersensitive States. New York, Hoeber-Harper.

Townsend, A. and H. Bodmer (1989). "Antigen Recognition by Class I-Restricted T Lymphocytes." Annual Review of Immunology **7**(1): 601-624.

van den Broek, M. E., D. Kägi, et al. (1996). "Decreased tumor surveillance in perforin-deficient mice." The Journal of Experimental Medicine **184**(5): 1781-1790.

Van den Eynde, B. J. and S. Morel (2001). "Differential processing of class-I-restricted epitopes by the standard proteasome and the immunoproteasome." Current Opinion in Immunology **13**(2): 147-153.

van der Merwe, P. A., A. N. Barclay, et al. (1994). "Human Cell-Adhesion Molecule CD2 Binds CD58 (LFA-3) with a Very Low Affinity and an Extremely Fast Dissociation Rate but Does Not Bind CD48 or CD59." Biochemistry 33(33): 10149-10160.

van Endert, P. M., D. Riganelli, et al. (1995). "The peptide-binding motif for the human transporter associated with antigen processing." The Journal of Experimental Medicine **182**(6): 1883-1895.

Vogelstein, B. and K. W. Kinzler (2004). "Cancer genes and the pathways they control." Nat Med **10**(8): 789-799.

Wang, H. Y., J. Zhou, et al. (2002). "Identification of a Mutated Fibronectin As a Tumor Antigen Recognized by CD4+T Cells." The Journal of Experimental Medicine **195**(11): 1397-1406.

Wang, J. and M. A. Maldonado (2006). "The Ubiquitin-Proteasome System and Its Role in Inflammatory and Autoimmune Diseases." Cellular Molecular Immunology **3**(4): 255-261.

Willcox, C. R., V. Pitard, et al. (2012). "Cytomegalovirus and tumor stress surveillance by binding of a human [gamma][delta] T cell antigen receptor to endothelial protein C receptor." Nat Immunol 13(9): 872-879.

Wu, T. T. and E. A. Kabat (1970). "An analysis of the sequences of the variable regions of bence jones proteins and myeloma light chains and their implications for antibody complementarity." J. Exp. Med. **132**(2): 211-250.

Yewdell, J. W., L. C. Antón, et al. (1996). "Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules?" The Journal of Immunology **157**(5): 1823-1826.

Zarling, A. L., J. M. Polefrone, et al. (2006). "Identification of class I MHC associated phosphopeptides as targets for cancer immunotherapy." PNAS **103**(40): 14889-14894.

Zarling, A. L., S. B. Ficarro, et al. (2000). "Phosphorylated Peptides Are Naturally Processed and Presented by Major Histocompatability Complex Class I Molecules In Vivo." J. Exp. Med. **192**(12): 1755-1762.

Zhang, L., J. R. Conejo-Garcia, et al. (2003). "Intratumoral T Cells, Recurrence, and Survival in Epithelial Ovarian Cancer." New England Journal of Medicine **348**(3): 203-213.

Zheng, X., J.-X. Gao, et al. (2002). "Clonal Deletion of Simian Virus 40 Large T Antigen-Specific T Cells in the Transgenic Adenocarcinoma of Mouse Prostate Mice: An Important Role for Clonal Deletion in Shaping the Repertoire of T Cells Specific for Antigens Overexpressed in Solid Tumors." The Journal of Immunology **169**(9): 4761-4769.

Zinkernagel, R. M. (2002). "On differences between immunity and immunological memory." Current Opinion in Immunology **14**(4): 523-536.

Zinkernagel, R. M. and P. C. Doherty (1974). "Immunological surveillance against altered self components by sensitised T lymphocytes in lymphocytes choriomeningitis." Nature **251**(5475): 547-548.

Zinkernagel, R. M. and P. C. Doherty (1974). "Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system." Nature **248**(5450): 701-702.

Zinkernagel, R. M., M. F. Bachmann, et al. (1996). "On Immunological Memory." Annu. Rev. Immunol. **14**: 333-367.

Zou, W. (2005). "Immunosuppressive networks in the tumour environment and their therapeutic relevance." Nat Rev Cancer **5**(4): 263-274. Brunger, A. T. (1992). "Free R value: a novel statistical quantity for assessing the accuracy of crystal structures." Nature **355**(6359): 472-475.