

Chapter 1: Introduction

1.1 Overview

Leukocyte regulation, activation and function are directed by the interactions of the numerous receptors expressed on their cell surfaces. Insight into both the general nature of such interactions and the array of proteins involved is therefore crucial to understanding these processes. In addition to the obvious medical importance of studying these interactions in the immune system, they are the most characterised of all such cell surface interactions. Thus, their full characterisation may enable the modelling of the extent and nature of other important sets of proteins involved in cell-cell communication.

This thesis takes a two-pronged approach to investigating protein-protein interactions at the leukocyte cell surface. Firstly, the structural nature of such interactions is investigated using the CD2 subset of the immunoglobulin superfamily (IgSF) as a model sequence-related family. All the members of this family appear to interact with other members of the family and to function at the leukocyte cell surface (Tangye et al. 2000). Secondly, two approaches are used to establish the extent of current knowledge about the range of proteins involved and to extend it. These approaches are homology searching (again using the CD2 subset as a model family) and serial analysis of gene expression (SAGE (Velculescu et al. 1995)) in a cytotoxic CD8⁺ T cell clone.

This introductory chapter reviews current models of protein-protein interactions at the cell surface and functional properties of the molecules involved on leukocytes. It will also introduce the CD2 subset and describe the molecules investigated in this thesis.

1.2 The interface between cell surfaces: a unique environment

Virtually all processes that distinguish multi-cellular and single cell organisms require cell to cell communication. This communication can be achieved by the secretion of a soluble factor by one cell and its recognition by another, whether at the cell surface (e.g. for cytokines or growth hormones) or once the factor has penetrated the cell (e.g. steroid hormones).

However, in many cases cell-cell communication is achieved by direct interaction of proteins on one cell surface with those on another cell. These interactions can be long-lasting stable interactions, like those binding the cells of solid tissues together, or transient interactions where the cells bind, communicate via their cell surface receptors and then separate again.

Such transient interactions are best described in the immune system, where the function of leukocytes is often determined by their interaction with each other and with other cell types.

Transient cell surface interactions are also important in regulating processes such as development (reviewed extensively e.g. (Huh et al. 2000; Fercakova 2001; Grunwald et al. 2001)), neuronal plasticity (e.g. (Newman 1996; Haag et al. 1999; Wedlich 2002)) and wound healing (reviewed in (Shock and Laurent 1996; Mutsaers et al. 1997)).

Whether two cells interact stably or transiently, the contact zone between them is a unique environment with different mechanisms and rules for protein-protein interactions from those for interactions where at least one binding partner is in solution. The unique features of this environment derive from the fact that the proteins involved are locked into a membrane, which both restricts their movement and links them to many other proteins. This membrane is a bilayer consisting mainly of phospholipids, the properties of which have long been thought to fit a 'fluid-mosaic' model (Singer and Nicolson 1972). This model proposes that membrane components - both lipids and integral membrane proteins - whilst being held in the plane of the membrane by hydrophobic contacts, are free to diffuse laterally. The logical conclusion of this model is that the individual components would be randomly distributed

within the membrane. More recently, however, certain groups of lipids and protein have been found to form microdomains within the membrane, which diffuse together and have distinct characteristics from other regions of the membrane (Simons and Ikonen 1997). This means that certain components of the membrane do not have random distributions; rather they are linked within these microdomains (Somerharju et al. 1999).

The membrane localisation of the proteins involved leads to four distinctive properties of protein interactions in the contact zone between cells. Firstly, although the proteins involved can move freely in two dimensions, they are restricted in the third because they cannot leave the membrane. This restriction of diffusion means that standard affinities for their interactions (i.e. those calculated by measuring the binding of their extracellular domains in solution) are not valid, interaction instead being determined by their ‘2-dimensional affinities’ (Dustin et al. 1997).

Secondly, because both cells will express multiple copies of the proteins on their surfaces, multiple interactions will occur. This leads to an increased *avidity* of interactions between cells expressing proteins that bind with a given *affinity* – even proteins with modest affinities for each other can induce tight binding of two cells when expressed at reasonably high levels. For example, this effect is evident in the activation of cell adhesion mediated by the LFA-1 integrin, in which clustering of the protein on the cell surface has a larger effect on the avidity of the cell-cell interaction than changes in the affinity of individual molecules (reviewed in (van Kooyk and Figdor 2000)).

Thirdly, proteins in a cell membrane are able to function interactively in complex networks without such processes being limited by diffusion. Ordered assemblies and arrays of membrane components can form by the lateral interaction of proteins. For example, the T

cell receptor complex is composed of associated T cell receptor (TCR) and CD3 polypeptide chains (Malissen and Malissen 1996) and can associate with several signalling proteins (Wange and Samelson 1996) or with other membrane proteins (Beyers et al. 1992).

Alternatively, components can be ordered due to their preferential association with particular lipid microdomains (sometimes called rafts), as has been proposed for T cell activation (reviewed in (Alonso and Millan 2001)).

Finally, the membrane forms a barrier between extracellular factors and the intracellular processes that define cellular function. Thus, for protein-protein recognition between two cells to alter their function, it must occur in a manner that enables the transfer of this signal across the membrane. The possible mechanisms of such signal transduction are still being intensively investigated by many groups and are fundamental to understanding cell-cell communication events. Such mechanisms will be discussed later (section 1.5) in the context of T cell activation following recognition of antigen on the surface of a target cell.

1.3 Proteins expressed at the leukocyte cell surface

1.3.1 General functions

A wide variety of proteins are expressed on the cell surfaces of different types of leukocytes, however they can be grouped into a few classes based on structure or function. Broadly speaking, leukocyte cell surface proteins (other than the ion channels and transporters needed by all cells) function either as receptors - for soluble molecules, extracellular matrix (ECM) components or proteins or carbohydrates on the surface of other cells - or as ligands for such receptors on other cell types. There are also various extracellular enzymes on the cell surface such as the metalloproteases, which digest the extracellular matrix and allow leukocytes to pass into tissues, but these molecules fall outside the remit of this thesis.

Of the receptors and ligands on the leukocyte surface, 'antigen specific' receptors can be distinguished as those whose genes undergo random rearrangement in somatic cells so that the clonal populations of cells produced can each recognise a different antigen using its particular rearranged receptor. These antigen specific receptors are the immunoglobulins on B cells and T cell receptors (TCRs) on T cells. The major histocompatibility complex (MHC) molecules, which are expressed on antigen presenting cells, are not rearranged, but are polygenic and highly polymorphic resulting in a wide range of different MHC molecules being expressed in a population. MHC molecules bind short peptides derived from protein antigens and present them at the cell surface. T cells, via their TCRs, can then bind these peptide-MHC complexes on the antigen presenting cell's surface. The majority of other proteins at the leukocyte surface have evolved a tight specificity for either a single ligand or a group of ligands. Due to the avidity effects discussed in the previous section, affinities of interactions between receptors and ligands that are both membrane bound tend to be low. The affinities of receptors for soluble ligands are higher, allowing them to bind the ligand although it is at low concentration in the solution around the cell. On binding a ligand, most receptors transmit a signal into the cell they are expressed on in order to initiate the appropriate response to that ligand. However, some receptors with ligands on the ECM or on other cells function simply as adhesion molecules, holding the two cells together or holding the cell on the ECM for a limited time.

As for all biological recognition, the functions of leukocyte cell surface molecules require that they are capable of interacting with their ligands (usually other proteins) with great specificity. However, receptors whose ligands are also membrane bound must do this with only modest affinity - thus allowing them to interact transiently rather than permanently.

1.3.2 Common Domains

Most proteins consist of several domains i.e. sequence segments likely to form discrete structural units (Doolittle 1995). Domains with similar three dimensional structures occur in many different proteins and the set of domains with that structure is known as a 'superfamily'. Superfamilies are actually defined by similarities of primary amino acid sequence between domains, and are assumed to have arisen by duplication and later mutation of exons encoding ancestor domains. Proteins can consist of domains from only one such superfamily or of a combination of domains from several superfamilies, the latter being called mosaic proteins. The proteins at the leukocyte cell surface contain a limited number of domain types in their extracellular regions and these will be discussed briefly.

The most common of these is the immunoglobulin superfamily (IgSF) domain (Barclay et al. 1997) that is present in 34% of leukocyte surface antigens. IgSF domains consist of about 100 amino acids and have a similar structure to the domains of immunoglobulins, the receptors present on the B cell surface and secreted as antibodies by activated plasma cells. The topology of these domains (called the Ig-fold, (Amzel and J. 1979)) is a sandwich of two β sheets, each consisting of antiparallel β strands of 5-10 amino acids. The number of strands in each sheet divides the superfamily into families known as the V- (for variable), C1-, C2- (C for constant) and I- (intermediate) sets, these distinctions having first been seen at the level of primary sequence (Williams, A. F. and Barclay 1988). V- and C1-set domains are found in immunoglobulins and T cell receptors, whereas C2- and I-set domains are found in non-antigen receptors such as adhesion molecules. The longest domains are the V-set sequences which have four β strands in one sheet (called A, B, E and D) and five in the other (called C, C', C'', F and G), the strands being labelled alphabetically by their order of occurrence in the primary amino acid sequence. C1-set domains are missing the C' and C'' strands, whilst C2-set domains have the C' strand but not the C'' or D strands (Bork et al. 1994). I-set domains

were previously thought to fall within the C2-set but re-examination of primary sequences (Harpaz and Chothia 1994) and, subsequently, three dimensional structure (Thomsen et al. 1996) showed that they were a distinct 'hybrid' set with both the C' and D strands but no C" strand. IgSF domains represent a stable structure that is well suited for recognition of small molecules and other protein domains (often also members of the IgSF). It is likely that they are so common at the cell surface because of this stable fold and because each IgSF domain is usually encoded in a separate, single exon (Williams, A. F. and Barclay 1988).

Three other domain types have a similar structure, consisting of two β sheets, but do not share sequence similarities with the IgSF. This lack of discernable sequence homology suggests that these domain superfamilies may have arisen by convergent rather than divergent evolution, i.e. four separate ancestor sequences evolved a similar protein fold due to its exceptional stability. However, it remains possible that these domains diverged from a single ancestor too long ago for any residual sequence homology to be discernable. The other three domain types with this fold are the fibronectin type III (Fn3) (de Vos et al. 1992; Main et al. 1992), cytokine receptor (Cosman et al. 1990) and cadherin (Angst et al. 2001) superfamilies. The first two of these are also found at the leukocyte cell surface and the β strands of all three domain types are lettered as in IgSF domains. Cytokine receptor domains, as their name indicates, are found in cytokine receptors - usually in combination with IgSF and/or Fn3 domains (Cosman et al. 1990). Fn3 domains are also commonly found in conjunction with IgSF domains amongst membrane proteins in the nervous system (Brummendorf and Rathjen 1993) and in muscle proteins (Labeit and Kolmerer 1995).

Other common domains in the extracellular regions of leukocyte surface molecules are complement control protein (CCP), epidermal growth factor (EGF), fibronectin type II (Fn2), galectin, integrin, lectin C-type, link, Ly6, tumour necrosis factor (TNF) superfamily, TNF

receptor (TNFR) superfamily and scavenger receptor cysteine rich domains. The diversity of domain types found in these proteins is not surprising considering the diversity of functions that they perform. Their ligands can be soluble, or membrane bound, protein or carbohydrate and span a wide range of structures; their affinities may be weak (e.g. for transient cell-cell interactions) or strong (e.g. for binding soluble factors at low concentration); they may bind one, two or a set of different ligands; they may need to dimerise or form higher order oligomers and they may or may not be required to transmit signals into the cell. Rather, the limited number of domain types used is surprising and demonstrates that the evolution of new protein specificities by rearrangement and adaptation of existing domains is more common than the evolution of new domain types, as suggested in the initial analysis of the human genome (Lander et al. 2001).

In addition to these discrete domains, there are certain families of non-modular proteins that share overall sequence similarities and often transmembrane topologies. The commonest of these are the seven transmembrane or G protein-coupled receptors (Watson and Arkinstall 1994), which contain seven transmembrane alpha helices and activate signalling cascades beginning with intracellular guanine nucleotide binding proteins in response to binding their ligands extracellularly. The tetraspanin protein superfamily (Maecker et al. 1997) is a family of proteins thought to cross the lipid bilayer four times. Its members may function directly as receptors and also by linking clusters of membrane receptors and intracellular signalling molecules (Hemler 2001).

The majority of newly identified leukocyte cell surface molecules would be expected to contain one or more of these domain types in their extracellular domains, although it remains possible that novel domain types or sequence superfamilies remain to be discovered. The intracellular regions of these membrane proteins often also contain certain common sequence

motifs or even domains but these are less clearly defined than those of the extracellular regions. The importance of these motifs is in their ability to transmit signals when the extracellular receptor binds its ligand. The most common signalling motifs of receptors contain tyrosine residues that can be phosphorylated. Two of these motifs are the immunoreceptor tyrosine-based inhibition motif (ITIM) and the immunoreceptor tyrosine-based activation motif (ITAM) (Isakov 1997). These motifs are phosphorylated by various kinases and dephosphorylated by a range of phosphatases in response to events at the cell surface. Phosphorylation allows the binding of signalling molecules containing Src homology 2 (SH2) domains, a process that initiates intracellular signalling cascades. As their names suggest, signalling through ITIM motifs usually leads to termination of certain leukocyte functions whereas signalling through ITAMs leads to cellular activation. Other motifs also occur in the cytoplasmic regions of membrane bound receptors, such as proline-rich motifs that allow binding of SH3 domain-containing signalling molecules (Pawson and Gish 1992), but these will not be considered in detail in this thesis.

1.3.3 Molecules at the T cell surface

To investigate and extend current knowledge of the range of proteins expressed at the leukocyte surface, it has been advantageous to focus on a single model cell type. A full description of the molecules involved on the surface of this one cell would be a first step towards understanding how such a complex network of proteins and interactions can function in the cell's responses to its environment. Because of their importance in the immune responses, T lymphocytes or T cells have been one of the most studied lineages of leukocyte over recent years. These cells are therefore ideal as such a model cell type. In addition, T cells are crucial in initiating both cellular and humoral adaptive immune responses, and understanding how they interact with their environment is therefore central to understanding the immune response.

The molecules on the T cell surface can be divided into several groups. Firstly, those directly associated with the central antigen recognition event: the components of the T cell receptor (TCR) itself, the chains of CD3 that form the signalling complex associated with the TCR and the co-receptors CD4 and CD8. Secondly, other accessory molecules i.e. proteins that affect the T cell response to antigen in a non-antigen specific manner. Thirdly, molecules responsible for the migration of the cells into and out of lymphoid organs or damaged / infected tissues and for recognising recruitment signals such as chemokines. Fourthly, effector molecules expressed on the T cell surface once activated. Finally, other molecules expressed on the T cell surface that are not specific to T cell function such as general metabolic transporters and proteins involved in antigen presentation to other lymphocytes (e.g. MHC molecules). This thesis aims to investigate the full complement of T cell expressed molecules but will be especially concerned with antigen recognition and the accessory molecules that affect its outcome. This section therefore describes the well known molecules involved in this process.

T cells carry clonotypic, antigen-specific TCRs, which usually recognise short peptide antigens, bound to MHC molecules on the surface of antigen presenting cells (APCs). T-cell activation, however, is not only dependent on this antigen specific interaction but is also affected by several other cell surface proteins collectively called accessory molecules. The interactions of these additional proteins and the signals they transmit to the T-cell affect the outcome of TCR-peptide-MHC recognition (i.e. activation, ignorance or anergy) and regulate the termination of the response.

The TCR is produced by random rearrangement of the TCR genes during thymocyte development and T cells are then positively and negatively selected in the thymus so that

circulating T cells recognise peptides bound to self-MHC molecules but do not recognise self peptides bound to these MHC molecules. Each T cell expresses TCR from only a single rearranged gene, produced by the recombination of one allele each of the genes for either the α and β chains or the γ and δ chains. The remaining alleles and genes do not successfully recombine. Thus, T cells express either an $\alpha\beta$ TCR or a $\gamma\delta$ TCR capable of recognising one specific ligand. $\alpha\beta$ T cells are the better studied of these two sub-types and appear to be the main type of T cell involved in the classical adaptive immune response. The TCR is expressed at the cell surface with the various chains of CD3, forming a complex capable of transmitting a signal into the cell on binding to its ligand (see the discussion of T cell activation in section 1.5).

$\alpha\beta$ T cells can be further subdivided by their expression of either the CD4 or the CD8 co-receptor. These molecules can bind MHC molecules on APCs with CD4 recognising class II MHC molecules and CD8 recognising class I MHC molecules. The co-receptors are thought to enhance antigen recognition by T cells. This division of $\alpha\beta$ T cells therefore has functional significance. Expression of CD4 classically indicates cells whose TCR recognises class II MHC molecules and which function as a 'helper' T cell i.e. one that primarily functions to secrete cytokines and regulate other leukocytes. Expression of CD8, on the other hand, indicates cells whose TCR recognises class I MHC molecules and which function as a cytotoxic T cell (CTL) that, when activated, can trigger the lysis of cells expressing its MHC-peptide ligand on their surface. Neither CD4 nor CD8 are expressed on early thymocytes but as $\alpha\beta$ thymocytes develop, they become 'double positives' expressing both CD4 and CD8, before finally being positively selected and losing expression of one of these co-receptors. As stated above, other accessory molecules are not directly involved in this antigen recognition event but affect its outcome. Such molecules can generally be classified as adhesion

molecules (Springer 1990), co-stimulatory molecules (Lafferty et al. 1983) or inhibitory molecules (Veillette et al. 2002).

The adhesion molecules function to initiate and maintain contact between the T cell and the APC, allowing antigen recognition to occur. Larger adhesion molecules mediate the initial contact between the cells and hold them together long enough to establish better contact, the best example of this being the integrin LFA-1 on T cells (Springer et al. 1987) which binds to ICAM-1 (Marlin and Springer 1987; Makgoba et al. 1988), -2 (Staunton et al. 1989) or -3 (Fawcett et al. 1992) on APCs or target cells. Smaller molecules such as CD2 (on the T cell) and CD58 (on the target cell), may hold the cell membranes in close apposition, allowing the TCR to sample MHC molecules on the APC ((Davis, S. J. and van der Merwe 1996), section 1.5).

Co-stimulatory molecules transduce activatory signals to the T cell on binding their ligands but cannot activate the T cell on their own; rather they supplement the antigen specific signal received through the TCR. This 'two-signal hypothesis' was first proposed by Lafferty et al (Lafferty et al. 1983) before the molecules involved had been identified. Several co-stimulatory molecules have been described (Watts and DeBenedette 1999) including CD137 (4-1BB), CD150 (SLAM) and CD134 (OX-40) but none have as large an effect as CD28, which binds CD80 or CD86 (B7-1 and -2) on APCs (Yashiro et al. 1998). However, recently another costimulatory molecule related to CD28 has been identified, called ICOS for Inducible COStimulator (Hutloff et al. 1999).

Inhibitory molecules counteract T cell activation and function either to prevent inappropriate activation or to end activation and effector function after the appropriate period. The best described inhibitory molecule on T cells is CD152 (CTLA-4), which is related to CD28 and

ICOS, binds to CD80 and CD86, and is usually expressed only after T cell activation in order to bring the T cell response to an end (Oosterwegel et al. 1999).

No one model has yet been devised to explain the complex interactions of all these accessory molecules, and it is possible that several such molecules remain to be described.

1.3.4 The discovery of novel genes, transcripts and proteins

Several strategies have been used to identify new proteins expressed on leukocytes. These can be grouped into four broad classes: those starting with a novel phenotype, those based on finding novel protein antigens with antibodies, those looking for molecules related to others that have already been identified and those based on the analysis of gene expression profiles. Phenotype based methods involve targeted searches for the molecule responsible for a given effect e.g. the gene responsible for an immunological disorder or one that confers a novel function on a cell line. Antibodies have been used extensively to identify new surface antigens on leukocytes and this is the basis of the 'cluster of differentiation' or CD nomenclature that now includes over 200 such antigens (Mason et al. 2001).

Methods based on DNA sequence can be applied more systematically and universally and two main types of such analysis have been used. Firstly it is possible to search for homologues of genes with known immune function - either searching for a few close homologues of a single known gene or searching for all molecules containing domains or motifs often associated with immune function (such as IgSF domains or ITIM motifs). This latter technique can be readily automated but may result in several false positives because these motifs and domains are also seen outside the immune system, whereas the former approach gives few false positives but may not identify many new genes as such close homologues are rare.

Finally, new molecules can be identified by analysing the messenger RNA expressed in a cell of interest. This can be done by sequencing expressed sequence tags (ESTs) (Adams et al. 1993), or by hybridising cDNA (produced from cellular RNA) to a microarray of DNA sequences (Southern 2001). The use of EST sequencing does not require any previous knowledge of the gene sequence and thus is ideal for identifying novel genes expressed in a cell. However, it is highly inefficient because there are, on average, around 300,000-350,000 mRNA molecules in a cell and allowing for the statistical effects of sampling, known genes must therefore be sequenced many times over in order to find novel genes. Sequencing sufficient EST to identify all the genes in a cell is therefore impractical. The use of microarrays is more systematic but it requires some previous knowledge of the sequence to be found and is only quantitative when used to compare the expression of a gene between two cells. However, this technique has been applied extensively in leukocyte populations.

Several groups have used microarrays to study leukocytes, most notably using an array enriched in lymphocyte cDNAs (the 'Lymphochip' (Alizadeh, A. et al. 1999)). This array has been used to investigate the expression profiles of known molecules in different lymphocyte lineages and developmental stages, identify clusters of transcripts involved in particular processes (such as proliferation, B/T cell activation or cell cycle regulation), delineate particular signalling pathways and characterise the expression profiles of different disease states – notably identifying two separate classes of diffuse B cell lymphomas based on their gene expression (all reviewed by Alizadeh et al. (Alizadeh, A. A. and Staudt 2000)).

Microarray studies on T cells themselves have included identification of transcripts differentially expressed in cells with type 1 and type 2 cytokine secretion profiles (Rogge et al. 2000; Chtanova et al. 2001), investigation of the effect of calcium signalling on gene expression (Feske et al. 2001) and characterisation of transcripts whose expression is altered on activation (Ollila and Vihinen 1998; Teague et al. 1999; Ellisen et al. 2001), HIV infection

(Geiss et al. 2000) or heat shock (Schena et al. 1996). These studies have identified known genes important in particular processes and T cell states, but due to the limited numbers of genes represented on the arrays and the non-quantitative nature of absolute fluorescence measurements, current microarray technology has been unable to identify the complete set of transcripts whose expression characterises a particular T cell in a particular state.

An alternative approach is serial analysis of gene expression (SAGE, (Velculescu et al. 1995)), which allows a quantitative description of the transcriptome of a cell population without the need for prior knowledge of gene sequence. Because it gives a direct measure of transcript abundance (e.g. copies per cell), SAGE also provides a lasting reference library against which other populations can be compared in the future, without the need to reanalyse the original population. Among immune cells, SAGE has been used to identify transcripts differentially expressed in certain cell lines and conditions, for example between sub-types of CD4⁺ T cells (Nagai et al. 2001; Zelenika et al. 2002) and between monocytes, macrophages and dendritic cells (Hashimoto, S. et al. 1999; Hashimoto, S. et al. 1999; Hashimoto, S. I. et al. 2000). SAGE has also been used to investigate expression changes after HIV infection of CD4⁺ T cells (Ryo et al. 1999). Finally, one group has used SAGE to investigate the expression profile of a relatively poorly characterised lineage of cells, intra-epithelial lymphocytes (Shires et al. 2001). However, SAGE was initially proposed as a systematic and exhaustive technique both for comparisons of gene expression between cells or conditions and for identifying complete transcriptomes (Velculescu et al. 1995). To date it has not been used to its full potential in this second area – to characterise the complete transcriptome of a particular cell. The initial application of SAGE for this purpose to a clonal population of cytotoxic CTL and methods of analysis that can be used to identify novel transcripts of interest from this data are described in chapter 7.

1.4 The nature of protein-protein interactions at the cell surface

1.4.1 Interactions of soluble proteins

Early work on protein-protein interactions focussed on binding and recognition events between proteins in solution, especially antibody-antigen and proteinase-inhibitor binding and the recognition of soluble hormones by their receptors. These interactions are mono- or oligomeric and are characterised by both high affinity and high specificity. This allows these proteins to recognise their ligands at the low concentrations found in solution and to bind them for the relatively long periods of time (at least several seconds) needed for these proteins to perform their function. The binding sites of these proteins are largely non-planar and the surface-shape complementarity between the two proteins along the interface is thought to play a central role in generating both specificity and high affinity (Chothia and Janin 1975; Creighton 1993). The composition of the binding face is not usually significantly different from that of the average protein surface, although it tends towards increased hydrophobicity (Janin and Chothia 1990; Jones, S. and Thornton 1996). The entropic benefit of excluding water from the hydrophobic residues on the complementary binding faces is often the most important source of energy driving complex formation (Chothia and Janin 1975; Creighton 1993).

1.4.2 The need for weak interactions between cell surface molecules

As indicated in section 1.2, cell-cell interactions are generally multimeric, involving large numbers of binding events between arrays of proteins on the cell surface. Often however, these cell-cell interactions must be easily reversible, as the generation of the correct intracellular response may require relatively brief interactions (van der Merwe and Barclay 1994). The multimeric nature of these interactions and the need for regulated, rapid dissociation mean that the binding of individual cell surface proteins must be much weaker than that of proteins in solution, such as those involved in the better characterised interactions

discussed above (Davis, S. J. et al. 1998). This low binding affinity, however, must be achieved without any decrease in specificity for cognate recognition to result in correct cell function. In the immune system, such high specificity, low affinity recognition allows rapid dissociation after transient cellular contact so that lymphocytes, for example, can sample many targets or antigen presenting cells in succession.

1.4.3 Affinities and kinetics of interactions between leukocyte cell surface molecules

Despite the fact that the affinities for all transient protein-protein interactions at the cell surface must be weak compared to those involving soluble proteins, these affinities still span a thousand-fold range ($K_d \sim 0.2 - 200\mu\text{M}$), as indicated in *Table 1.1*. The precise affinity of each interaction is determined by the kinetics of that interaction ($K_d = k_{\text{off}}/k_{\text{on}}$ where k_{on} and k_{off} are the rate constants for association and dissociation respectively). The importance of these quantitative binding properties of cell surface molecules for their functions is clearly shown by studies on TCR-peptide-MHC interactions, for which these values can vary significantly depending on the receptors and antigen involved. The ability of peptide-MHC complexes both to influence selection of the TCR repertoire in the thymus (Alam et al. 1996; Williams, C. B. et al. 1999) and to induce T cell activation (Kersh et al. 1998) has been shown to correlate with the off rate (k_{off}) of their interactions with TCRs. In both cases, the outcome of the interactions is thought to be determined by kinetic thresholds. Given the range of affinities seen for other interactions at the cell surface, it is likely that quantitative differences in their binding properties also correlate with their function. In line with this, quantitative analysis of the interactions of CD80 and CD86 (B7.1 and B7.2, (Collins et al. 2002)) have suggested a mechanism by which they can exert differential effects on T cells despite binding the same ligands (CD28 and CD152/CTLA-4).

Interaction	Species	T	K_d (μM)	k_{on} ($\text{M}^{-1}\text{s}^{-1}$)	k_{off} (s^{-1})	References
CD4 / MHC class II	Mouse	25 °C	>100	N.D.	N.D.	(Weber and Karjalainen 1993)
KIRs / MHC I	Human		10	$>2 \times 10^5$	1.5	(Maenaka et al. 1999)
TCR / MHC II	Mouse		6	4×10^3	0.02	(Boniface et al. 1999; Wu et al. 2002)
CD8 α / MHC I	Human	37 °C	~200	$>1 \times 10^5$	>18	(Wyer et al. 1999)
CD2 / CD48	Mouse		90	$>1 \times 10^5$	>11	(Davis, S. J. et al. 1998)
CD2 / CD48	Rat		75	$>1 \times 10^5$	>6	(van der Merwe et al. 1993)
CD86 / CD28	Human		20	$>1 \times 10^6$	>28	(Collins et al. 2002)
CD2 / CD58	Human		15	$>4 \times 10^5$	>4	(van der Merwe et al. 1994)
TCR / MHC I	Human		15	7×10^4	1	(Willcox et al. 1999)
CD80 / CD28	Human		4	6×10^5	>1.6	(van der Merwe et al. 1997)
LICOS / ICOS	Human		4	N.D.	N.D.	(Brodie et al. 2000)
CD86 / CD152	Human		3*	2×10^6	5	(Collins et al. 2002)
OX2/OX2R	Human		2	4×10^5	0.8	(Wright et al. 2000)
CD80 / CD152	Human		0.2*	2×10^6	>0.4	(van der Merwe et al. 1997)

Table 1.1: Affinities and kinetics of interactions occurring at the leukocyte cell surface
* Affinities shown are for first ligand molecule to bind CD152, the affinity of the second site for its ligand is reduced, possibly by steric effects, to ~22 and ~2 for CD86 and CD80 respectively.

Although the precise binding properties of individual proteins may determine their particular function, it is also clear that all these molecules interact with relatively rapid kinetics. These kinetics, particularly the fast dissociation rates, coupled with the fluid nature of the cell membrane mean that any region of cell-cell contact must be highly dynamic, unless protein mobility is constrained by cytoskeletal association. Clusters of molecules might be expected to form at the cell-cell interface if their ligands are found on the opposing cell, as transient interactions would hold them there briefly with the result that the concentration of that particular molecule would be elevated at the interface relative to the rest of the cell surface. However, these clusters would be dynamic structures, with receptor-ligand pairs constantly associating and dissociating and molecules diffusing in and out of the cluster.

The formation of such clusters is the basis of a technique for assessing the two-dimensional (i.e. membrane bound) affinities of the interactions of cell surface molecules. As discussed in section 1.2, it is this two-dimensional affinity of an interaction, rather than its affinity in solution, that determines its functional effects on a cell. The two dimensional affinities of human and rat CD2 for their ligands were investigated (Dustin et al. 1997) and it was shown that these weak interactions were sufficient to induce membrane alignment and therefore generate a stable, cooperative contact area between a cell and an artificial lipid bilayer containing these molecules at physiological levels. However, as expected, the degree of membrane alignment decreases with the solution affinity of the interaction inducing it. The interaction between rat CD2 and CD48 (solution $K_d \sim 75\mu\text{M}$) appeared to be close to the lower limit at which membrane alignment could be expected to occur. Thus interactions of significantly lower affinity than this, such as those of CD4 and CD8 with MHC molecules, may not be physiologically relevant in isolation. In the case of these coreceptors, they may interact only with peptide-MHC complexes bound to a cognate TCR - the interaction being stabilised by association of the coreceptor with the TCR (Wyer et al. 1999).

1.4.4 The CD2-CD58 paradigm

CD2 was one of the first adhesion molecules to be implicated in T-cell recognition (Bierer et al. 1989) and its interactions with its ligands (CD58 in humans and CD48 in rodents) have become an important paradigm for protein interactions at the cell surface (Davis, S. J. and van der Merwe 1996; Davis, S. J. et al. 1998; Davis, S. J. et al. 1998). The NMR (Withka et al. 1993) and crystal (Jones, E. Y. et al. 1992; Bodian et al. 1994) structures of rat CD2 were the first such structures of an adhesion molecule to be solved and CD2 was the first adhesion molecule shown to have high specificity but low affinity for its ligand as suggested above (van der Merwe et al. 1993; van der Merwe et al. 1994). Extensive study of this system over recent years has yielded insights into the structural basis of these types of interaction.

The crystal structures of rat (Jones, E. Y. et al. 1992) and later human (Bodian et al. 1994) CD2 showed crystal contacts between adjacent CD2 molecules that were proposed to model the binding of CD2 to its ligands. The binding face in these crystal contacts was the GFCC'C" β -sheet of the variable (V-set) domain. Subsequent mutational (Arulanandam et al. 1993; Somoza et al. 1993; Davis, S. J. et al. 1998) and NMR (McAlister et al. 1996) analyses confirmed that this face of CD2 was responsible for binding and defined the extent of the binding site. Similarly, analysis confirmed that the rat CD48 binding site is located on the equivalent GFCC'C" face (van der Merwe et al. 1995).

The initial crystal structures of soluble CD2 (containing the complete extracellular domain) showed that this binding face was unusual, compared to those of proteins interacting in solution, in two ways. Firstly, it was remarkably flat and secondly, it contained a high proportion of charged and polar residues (Bodian et al. 1994). Thus a model for transient, low affinity interactions at the cell surface was proposed (Davis, S. J. and van der Merwe 1996; Davis, S. J. et al. 1998; Davis, S. J. et al. 1998).

This proposal suggested that the small amount of binding energy involved is generated by the few hydrophobic contacts between the interacting surfaces (in rat CD2 mediated by L38, F49 and Y81), and that high specificity is generated by contacts between charged and polar residues, which are energetically neutral but require an opposite charge on the opposite binding surface. The neutrality of this electrostatic binding is due to the unfavourable effect of disrupting interactions between the charged residues and surrounding water, which counteracts any favourable interactions generated on binding. This proposal is supported by thermodynamic data, which shows that favourable entropic effects account for a significant proportion of the total ligand-binding energy between rat CD2 and CD48 (O'Brien R., van der Merwe P., Davis S. and Ladbury J., unpublished data). The role of electrostatic interactions in conferring specificity is confirmed by charge-swap mutagenesis where the CD2 mutants E41R and K43E fail to bind wild type CD48 but bind the CD48 mutants R31E and E44K respectively (van der Merwe et al. 1995; Davis, S. J. et al. 1998).

More recently, a crystal structure has been described for the complex between the V-set domains of CD2 and CD58 (Wang et al. 1999). This structure confirms the earlier predictions, with binding occurring in the same orientation as the contacts seen in the CD2 crystals. A lack of shape complementarity is observed and there are many salt bridges formed between charged residues on the two protein surfaces.

1.4.5 The CD80-CD152 interaction: intermediate affinities

The CD2 paradigm presents a possible structural basis for the low affinity interactions seen between cell surface molecules. However, as discussed previously, the interactions of CD2 with its ligands are some of the weakest observed, with the affinity of the rat interaction being right on the threshold below which the two dimensional affinity is insufficient to induce membrane alignment (Dustin et al. 1997). An important interaction with a higher affinity is

that between CD80 (B7-1) and CD152 (CTLA-4): This interaction has a K_d of $\sim 0.3\mu\text{M}$ (van der Merwe et al. 1997) compared to $\sim 15\mu\text{M}$ for human CD2-CD58 (van der Merwe et al. 1994) and $\sim 75\mu\text{M}$ for rat CD2-CD48 (van der Merwe et al. 1993)). This affinity is still one hundred fold lower than that of interactions in solution such as antibody-antigen binding, as expected given the arguments in the section 1.4.2, but it is one of the strongest interactions seen at the leukocyte surface. The recent crystal structure of the complex between CD80 and CD152 (Stamper et al. 2001) reveals the structural basis of this intermediate affinity. The binding interface between the two proteins has many of the characteristics of classical protein-protein interfaces in solution, particularly high surface complementarity and mainly hydrophobic contacts. These features contribute high specificity and also high affinity to classical protein interactions, so other features have evolved to limit the affinity of this interaction. Firstly, the buried surface area on each protein is $\sim 600\text{\AA}^2$, which is at the low end of the range seen for protein-protein interactions ($600\text{-}900\text{\AA}^2$ (Janin and Chothia 1990)) and similar to that seen in the CD2-CD58 complex. Secondly, there is an entropic barrier to binding ($\Delta S = -2.7\text{kcal mol}^{-1}$), seen in thermodynamic analysis (R. O'Brien, personal communication) that may be explained by conformational changes between the free and bound proteins. A similar effect is seen in TCR binding to peptide-MHC molecules where the entropic barrier is larger (ΔS between -12 and -15 kcal mol^{-1} (Willcox et al. 1999)). The structure of the MYPPY loop in CD152, which forms most of the interactions with CD80, differs between the complex (Stamper et al. 2001) and the structure observed in solution by NMR (Metzler et al. 1997). The three proline residues in this loop adopt a *cis-trans-cis* conformation in the complex but a *trans-trans-cis* conformation is seen in the NMR structure. However, it is doubtful whether such a large conformational change could occur in solution and the fast binding kinetics of the interaction (van der Merwe et al. 1997) suggest that only small conformational changes could occur on binding.

Together with the TCR-peptide-MHC interaction, the structure of this complex extends our understanding of low affinity but high specificity interactions. Here the high surface complementarity and hydrophobic contacts characteristic of protein-protein interactions in solution are maintained and the reduced affinity is due to entropic effects. However, the affinity of the CD80-CD152 interaction is unusually high for interacting cell-surface molecules - probably due to its function in creating a rapid and potent inhibitory signal for the T cell - and the TCR-peptide-MHC interaction has unusually slow kinetics for cell surface interactions due to the entropic barrier to binding. Therefore, it seems likely that the CD2-CD58 paradigm of reduced surface complementarity and hydrophobic contacts to reduce affinity and electrostatic interactions to generate specificity will represent the pattern for many other protein-protein interactions at the cell surface. Overall, it seems that a variety of structural mechanisms ensure binding at the cell surface is both weak and specific, with the mechanisms used in each interaction being determined by the binding properties it requires for its function.

1.5 Organised assemblies of proteins at the leukocyte cell surface

1.5.1 The complexity of cell-cell communication

As is evident from the discussion above, many different interactions can occur between two cells. Cell-cell communication results from the complex interplay of all these interactions and the different signalling pathways they can initiate in each cell. Individual interactions are often investigated in isolation and ‘signalling pathways’ and functions are assigned to this interaction. However, in reality, the effect of any one interaction between a pair of proteins may depend on the other interactions occurring simultaneously or on the results of previous interactions that have altered the content or organisation of the proteome in the cell. Both

membrane proteins and cytoplasmic signalling molecules can interact with one another so that the number of possible configurations increases exponentially with the number of different proteins involved. The organisation of molecules at the cell surface and in the cytoplasm is therefore a key element in the determining the outcome of cell-cell communication.

1.5.2 The example of T cell activation

T cell activation is perhaps one of the best studied examples of cell-cell communication via a complex matrix of different cell surface molecules. The T cell has to be able to identify cells expressing antigen at very low levels on their surface and yet it has to be extremely specific so that it does not initiate an immune response against cells expressing only self antigens. As discussed previously, this process integrates signals from the TCR itself and from numerous accessory molecules, for example CD28, ICOS, CD152, CD150, 4-1BB and OX-40. This system is therefore an ideal initial model in which to try to understand how the array of cell surface molecules present and the nature of their interactions lead to correct responses to the similar array of molecules encountered on the surfaces of other cells.

The first level of organisational complexity to consider is the array of signalling molecules that can be activated by events at the cell surface. These molecules are often recruited into organised assemblies by adaptor molecules, which bind to the cell surface molecules initiating the signal. The signalling pathways resulting from engagement of the TCR with its ligand and from the interactions of the accessory molecules with their ligands have been investigated over many years and regularly reviewed (van Leeuwen and Samelson 1999; Kane et al. 2000; Leo and Schraven 2001). There is not space for a full description of the many signalling molecules involved in this introduction. However, it would be helpful to describe, briefly, some of the most fully characterised pathways resulting from TCR

engagement in order to understand the complexity of such pathways and identify some of the major cytoplasmic components associated with the cell surface molecules discussed in this thesis.

The earliest detectable biochemical events that occur on engagement of the TCR are the activation of Src family tyrosine kinases - most notably Lck, which is associated with the CD4/CD8 co-receptor cytoplasmic domain. These kinases phosphorylate the ITAM motifs in the cytoplasmic domains of the CD3 chains associated with the TCR, and importantly the ζ chain (CD247). This results in the recruitment of the Syk family tyrosine kinase ZAP70 (ζ associated protein, 70kDa) via its SH2 domain and its phosphorylation by the Src kinases. The activation of ZAP70 enables the phosphorylation of the important adaptor molecule LAT (linker for activation of T cells). This, in turn, enables recruitment of further adaptor molecules such as Gads, SLP-76 (SH2-containing leukocyte protein, 76kDa) and Grb2 and of the enzyme phospholipase C γ 1 (PLC γ 1). Both SLP-76 and PLC γ 1 are then phosphorylated by ZAP70. The phosphorylated trimeric complex of LAT, Gads and SLP-76 provides the binding site for the SH2 domains of the Tec family tyrosine kinases such as Itk, which also phosphorylate PLC γ 1. Active PLC γ 1 splits the phospholipid phosphatidylinositol 4,5-bisphosphate (PIP₂) into the second messengers diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP₃). IP₃ triggers calcium signalling, whilst DAG activates protein kinase C (PKC) and recruits Ras guanine releasing protein (RasGRP) to the membrane thus contributing to Ras activation. Grb2 recruits the SH3 binding protein Sos and therefore also activates Ras. Ras activates the MAP kinase signalling pathway. Calcium, PKC and MAP kinase signalling result in changes of gene expression in the activated T cell via activation of transcription factors such as NF-AT, AP1 and NF κ B. In addition to these well described signalling pathways, Src and Syk kinases also seem to phosphorylate other adaptor molecules such as TRIM, which can activate phosphatidylinositol 3 kinase to phosphorylate PIP₂ to

produce phosphatidylinositol 1,4,5-triphosphate, an activator of kinases such as Itk and of guanine exchange factors (GEFs) such as Vav. Vav is among a group of other adaptor molecules that have been shown to be recruited by SLP-76. These also include SLAP and Nck and together these molecules can signal cytoskeletal rearrangements, cap formation and/or cell migration via pathways involving the Rho/Rac G proteins and WASP.

This represents just some of the signalling pathways activated by TCR engagement, and does not consider the negative regulatory signals that are also instigated to counteract these activatory signals and maintain homeostasis (e.g. via PAG, SIT, Csk and Cbl). Each of these signalling pathways could be affected by additional signals from costimulatory or inhibitory molecules binding their ligands on the target cell or by the expression of inhibitors or other signalling molecules in response to previous encounters with other cells. Molecules such as LAT and SLP-76, which contain multiple tyrosine signalling motifs, can integrate multiple signals in their pattern of phosphorylation so that the relevant signalling cascades trigger the appropriate responses to the particular combination of interactions occurring at the cell surface.

1.5.3 Membrane microdomains or "lipid rafts"

As indicated in section 1.2, a further level of organisation exists within the plasma membrane itself. The large diversity of lipids in plasma membranes results in the formation of microdomains or 'rafts' with high sphingolipid and cholesterol content which are organised in a tightly packed, ordered formation distinct from the disordered arrangement of phospholipids in the bulk of the membrane (Simons and Ikonen 1997). These microdomains 'float' in the bulk membrane and have a different composition of proteins than the rest of the membrane. In particular GPI-linked and acetylated proteins accumulate in these rafts together with a few specific transmembrane proteins. Other integral membrane proteins are excluded from them.

These microdomains are also heterogeneous (Roper et al. 2000), most likely due to the cholesterol binding properties of different sphingolipids.

Membrane microdomains were identified in T cells as detergent resistant membrane complexes as long ago as 1983 (Hoessli and Rungger-Brandle 1983). More recently, it has been proposed that they form the platform for the signalling complexes formed during T cell activation (reviewed in (Alonso and Millan 2001)). Lipid rafts have also been implicated in signalling by other immune receptors such as the B cell receptor complex (Cheng et al. 2001) and the high-affinity IgE receptor of mast cells (Wilson et al. 2000; Wilson et al. 2001).

The exact role of lipid rafts in these immune processes is still unclear, but they do seem to be necessary at least for T cell activation (Stulnig et al. 1998; Xavier et al. 1998). Several important signalling molecules appear to be preferentially associated with rafts, including the Src kinases Lck and Fyn (Xavier et al. 1998) and the adapter molecule LAT (Zhang et al. 1998)). On T cell activation, these kinases phosphorylate several substrates (see above) and these newly phosphorylated substrates also then seem to be recruited to lipid rafts (Montixi et al. 1998). However, the presence of the TCR-CD3 complex itself in membrane microdomains before and after T cell activation remains controversial (Kosugi et al. 1999). Engagement of CD28 (Viola et al. 1999) and other costimulatory molecules (Yashiro-Ohtani et al. 2000) can also lead to reorganisation of membrane microdomains and this may be one of the mechanisms that result in their effects on T cell activation. Furthermore, some molecules implicated in negative signalling are excluded from these microdomains e.g. the large membrane bound molecule CD45 which has tyrosine phosphatase activity (Rodgers and Rose 1996).

1.5.4 The immunological synapse: form and function

The organised assembly most linked with immune function is the ‘immunological synapse’ - the highly organised area of contact between immune cells and in particular between a T cell and an APC, named for its similarity to neurological synapses (Paul and Seder 1994). This organisation was revealed by imaging studies of the mature synapse which showed a bull’s eye arrangement of groups of molecules termed supramolecular activation clusters (SMACs) (Monks et al. 1998; Wulfiging and Davis 1998; Grakoui et al. 1999). The central region (cSMAC) is enriched for TCR on the T cell and peptide-MHC complexes on the APC whilst the outer ring or periphery of the synapse (pSMAC) contains the interacting adhesion molecules LFA-1 and ICAM-1. The important signalling molecules of T cell activation are also seen in the cSMAC. The formation of this synapse is dependent on an intact T cell cytoskeleton (Grakoui et al. 1999), suggesting that synapse formation occurs by cytoskeletally-driven movement of molecules and possibly membrane microdomains. Supporting this idea, the synapse begins as an inverted structure with adhesion molecules clustered in the centre and TCR/peptide-MHC interactions occurring on the periphery - this then progresses to the mature synapse structure over a period of several minutes (Grakoui et al. 1999).

Several researchers have suggested that a principle function of the immunological synapse is to stabilise signal transduction through the TCR (Delon and Germain 2000; Dustin and Chan 2000). However, while synapse formation correlates well with full T cell activation (Monks et al. 1998) and inhibition of synapse formation with inhibitors of cytoskeletal movement prevents full activation (Grakoui et al. 1999), several lines of evidence suggests that synapse formation *results* from TCR triggering rather than enabling it. Doubts that the synapse was involved in TCR signal transduction first arose because CD45 was excluded from the synapse and yet this molecule is necessary to maintain Lck activity (Leupin et al. 2000; Davis, S. J.

and van der Merwe 2001). In addition, the CD4 co-receptor thought to function in TCR triggering vacates the synapse very soon after its formation (Krummel et al. 2000). Thirdly, the cytoskeletal organisation and protein redistributions shown to be important for synapse formation require TCR and costimulatory signals. Finally, and most significantly, TCR signalling peaks before synapse formation and only occurs at a low level in the mature synapse. The calcium signal peaks after only 30 seconds (Grakoui et al. 1999) whilst tyrosine phosphorylation in general, and specifically that of Lck and ZAP70, mainly occurs at the periphery of the synapse and occurs well before mature synapse formation (Lee et al. 2002). Some phosphorylation of tyrosine residues can be detected in the synapse but it is well below the peak seen before and during the initial stages of synapse formation (Leupin et al. 2000; Lee et al. 2002).

If the immunological synapse is not a scaffold for TCR signalling, what is its function?

Three other principle roles have been suggested (discussed in (van Der Merwe and Davis 2002)). Firstly, the immunological synapse may well be important for the targeted delivery of T cell effector molecules on activation; this is similar to the function of neurological synapses. The possibility of such a role for the synapse is supported by the observation that the secretory granules of CD8⁺ CTL accumulate at and empty into the central region of the synapse formed with their target cells (Stinchcombe et al. 2001; Stinchcombe et al. 2001). The outer ring of large adhesion molecules that surrounds this central region may help to limit diffusion of these effector molecules away from the targeted cell, reducing 'bystander' effects. The synapse may serve not only to focus secreted factors but also to direct cell surface molecules that are stored in internal vesicles to the interface with the target cell. This has been shown for the inhibitory molecule CD152 (CTLA-4), which is expressed after T cell activation and appears to be targeted selectively to the contact site with the APC (Linsley et al. 1996).

Secondly, synapse formation may be an important checkpoint in T cell activation after TCR triggering. It is ideally suited to the recruitment of accessory molecules whose signals could then be integrated to enable the T cell to respond correctly to the antigenic signal already received. CD28 ligation with its ligands has been shown to be enhanced by synapse formation (Bromley et al. 2001), suggesting that the synapse works to enhance costimulatory signals after TCR triggering. Many other accessory molecules are also of a similar size to the TCR and may be expected to collect in the synapse where larger cell surface molecules are excluded (Sperling et al. 1998; Johnson et al. 2000; Leupin et al. 2000). Thus, the synapse may integrate the signals that determine T cell fate on antigen stimulation, for example determining whether the cell is activated or becomes anergic (as expected for a naïve cell stimulated in the absence of costimulation), or whether an activated helper T cell takes on a T_h1 or a T_h2 phenotype.

Thirdly, Lee et al. suggest that TCR clustering in the synapse may also serve as a prelude to endocytosis of the receptor (Lee et al. 2002), which in turn may serve as a negative regulatory mechanism (see (Valitutti and Lanzavecchia 1997)). While the exact role of the synapse remains unclear, it is clear that TCR triggering involves earlier events, which probably do not depend on active cytoskeletal movement of molecules.

1.5.5 Models of T cell activation 1: conformational change

If cytoskeletal activity and active transport occur only after TCR triggering, what is the mechanism that initiates this event? As indicated above many of the signalling pathways involved are well understood, however the question of how the binding of the TCR to its antigen results in these pathways becoming activated remains to be answered. Historically, models of signal transduction across cell membranes have always been of two major types.

Firstly, conformational change models have been described in which the binding of a receptor to its ligand results in small structural changes in the receptor that alter its interactions with signalling molecules in the cytoplasm. Such mechanisms have been described for several receptors such as members of the 7TM or G-protein coupled receptor superfamily (Bockaert and Pin 1999), the bacterial aspartate receptor (Ottemann et al. 1999) and ion channels such as the nicotinic acetylcholine receptor (Unwin et al. 2002). However, the mechanisms of signal transduction in these receptors all share one key feature: the receptor contains multiple transmembrane helices and ligand binding alters the arrangement of these helices with respect to each other in order to transmit the conformational change to their cytoplasmic regions. The receptors involved in T cell activation (the TCR, CD4/CD8, CD28 and most other accessory molecules) all have only single transmembrane helices so it is hard to see how such a conformational change could be transmitted across the membrane. In addition the random nature of TCR diversity results in significant variation in the structure of the ligand binding regions of these receptors, again making it hard to conceive how a common conformational change could result in each case. In fact, structural studies have shown no evidence of conformational changes in TCRs on binding to peptide-MHC ligands (Garcia et al. 1999).

However, two more complex conformational change models have been proposed for the TCR. The first suggests that TCR/CD3 complexes exist as dimers in the cell membrane (Fernandez-Miguel et al. 1999) and that the binding of peptide-MHC ligands to the two receptors alters their relative orientation in the dimer (Ding et al. 1999). This is analogous to recently described mechanisms of erythropoietin (EPO, (Remy et al. 1999)) and epidermal growth factor (EGF, (Moriki et al. 2001)) receptor activation. However, evidence for dimerisation of the TCR/CD3 complexes is indirect (via labelled antibodies) and the mechanism is again difficult to imagine given that the TCR does not have its own signalling domain and binds a ligand with only a single binding site (in contrast to the EPO and EGF

receptors). Any change in orientation would have to be transmitted to the large multi-chain CD3 signalling complex associated with the TCR and no such mechanism has yet been proposed.

The second conformational change model suggested recently depends on a change in the structure of the CD3 ϵ cytoplasmic domain on TCR-MHC recognition, revealing a proline-rich region able to bind the SH3 domain of Nck (Gil et al. 2002). This effect apparently precedes tyrosine phosphorylation, so is unlikely to be caused by intracellular events occurring after TCR triggering. As with other models, it is unclear how binding of TCR to peptide-MHC could cause such a change in the CD3 ϵ structure considering that its own structure does not change as a result of binding. It has been suggested that ligand binding and pressure on the membrane due to cell contact could cause a piston like effect on the TCR/CD3 complex, moving CD3 ϵ away from the membrane and causing it to unfold sufficiently for Nck binding (Davis, M. 2002). This mechanism is suggested by comparison to the piston effect seen in the aspartate receptor (Ottemann et al. 1999). However, in that case multiple transmembrane helices in the same protein alter their orientation with respect to each other whereas in the mechanism suggested here, the TCR moves through the membrane due to force alone - thus any non-specific force on the T cell would trigger the TCR. In addition, experiments showing the importance of the SH3 domain of Nck for TCR triggering are not conclusive, as it is already implicated in TCR activation since Nck is recruited by SLP-76.

1.5.6 Models of T cell activation 2: aggregation

The second major group of models used to explain signal transduction across cell membranes involve the aggregation of receptor molecules. In these models, multiple receptors are brought together on ligand binding and this aggregation of receptors leads to their mutual

activation due to their proximity of their cytoplasmic domains or associated cytoplasmic molecules (reviewed in (Weiss and Schlessinger 1998)). Such mechanisms were thought to be common to many receptors such as those for the growth factors EGF (Yarden and Schlessinger 1987)), human growth hormone (hGH (Cunningham et al. 1991)), platelet-derived growth factor (PDGF, (Heldin et al. 1989)), stem cell factor (SCF (Philo et al. 1996)) and fibroblast growth factor (FGF (Spivak-Kroizman et al. 1994)), those for the cytokines interferon- γ (IFN γ (Walter et al. 1995)) and tumour necrosis factor β (TNF β (Banner et al. 1993)) and those for apoptosis-inducing ligands such as tumor necrosis factor-related apoptosis-inducing ligand (TRAIL (Mongkolsapaya et al. 1999)). In many cases, the ligand is a constitutive dimer (or trimer in the case of TNF β and TRAIL) and therefore can bind and join two receptors. In other cases, a monomeric ligand is bivalent e.g. hGH (de Vos et al. 1992). Oligomerisation of the FGF receptor occurs by a different mechanism: molecules of FGF are bound to heparin sulphate and thus multiple receptors binding to them are held close enough for mutual phosphorylation of their cytoplasmic kinase domains (Spivak-Kroizman et al. 1994). However, even well established ligand induced dimerisation models have become controversial under recent scrutiny. The EPO receptor is now thought to exist as a pre-formed dimer whose conformation is changed by ligand binding ((Remy et al. 1999), see earlier). Also, EGF was thought to associate monomerically with its receptor, with dimers forming between two receptor-ligand complexes (Lemmon et al. 1997), but more recently it has been suggested that activation occurs due to a conformational change in a pre-formed receptor dimer on EGF binding (Moriki et al. 2001). Recent structures of both EGF (Ogiso et al. 2002) and TNF α (Garrett et al. 2002) bound to the EGF receptor show that dimerisation appears to be receptor-mediated not ligand induced and that changes of orientation between the two receptors in the dimer are induced on ligand binding. The recent structure of the related ErbB3 receptor also supports the possibility of conformational changes on ligand binding (Cho and Leahy 2002).

That dimerisation or oligomerisation is crucial to TCR signalling is a widely held view (e.g. (Brown, J. H. et al. 1993; Bachmann and Ohashi 1999)). An alternative model suggests that TCR aggregation and mutual activation results not from direct interaction between TCR chains but by recruitment of multiple TCRs to lipid rafts on ligand engagement (Harder 2001). However, all aggregation models of TCR triggering fail to account for the fact that TCR triggering requires only very low doses of peptide-MHC on the APC (Kimachi et al. 1997) and is most efficient at these low doses (Lanzavecchia et al. 1999). This is the opposite of what would be expected if two TCRs needed to simultaneously engage their ligands and aggregate together. Such a situation could be conceivable if the TCR ligand was dimeric on the APC but despite some MHC-peptide dimers seen in crystals (e.g. (Brown, J. H. et al. 1993)), it seems unlikely that such a dimer could form on the APC surface where self-peptide-MHC complexes can outnumber antigenic peptide-MHC by up to a hundred fold. Finally, the structural diversity of different TCRs and their peptide-MHC ligands, the monovalent nature of the interactions, their low affinities and their rapid kinetics also all mitigate against oligomerisation. Given these considerations, it is not surprising that no dimers have been seen in any of the TCR-peptide-MHC complex crystals produced so far (Garboczi et al. 1996; Ding et al. 1998; Garcia et al. 1998; Teng et al. 1998; Ding et al. 1999; Reinherz et al. 1999; Degano et al. 2000; Hennecke et al. 2000; Reiser et al. 2000). Despite some initial reports (Reich et al. 1997; Alam et al. 1999), studies of TCR-peptide-MHC complexes in solution have also failed to detect any indication of oligomerisation (Willcox et al. 1999; Baker and Wiley 2001). Oligomers of soluble peptide-MHC ligands do activate T cells more effectively than monomers (Boniface et al. 1998); however, this result would be expected for any model of TCR triggering requiring simultaneous ligation of multiple TCRs (i.e. even if they do not need to be associated). Thus, there is no direct evidence for aggregation models of TCR triggering and there are numerous theoretical problems that make it hard to conceive of an effective mechanism for the production of such aggregates.

1.5.7 Kinetic segregation model of T cell activation

The inability of conformational change and aggregation models to explain TCR triggering has led to the proposition of alternative models of signal transduction. One such model is similar to aggregation models: engagement of peptide-MHC by the TCR allows coreceptor (CD4 or CD8) binding to this complex, in turn enabling Lck, associated with the coreceptor, to phosphorylate the chains of the CD3 complex and trigger the T cell (Delon et al. 1998). However, this simple model can be quickly dismissed as a general mechanism because the presence of coreceptors is not necessary for T cell activation (Locksley et al. 1993; Schilham et al. 1993). In addition to this, co-receptor interactions are extraordinarily weak (Table 1.1) and the half-life of this association may be insufficient to allow the series of phosphorylation reactions required for TCR triggering to occur.

Another model is the kinetic segregation model, which proposes that TCR triggering is driven by passive segregation of T cell surface molecules on contact with an APC. This creates a small contact zone between the two cells that excludes certain inhibitory molecules (mainly phosphatases) with large extracellular domains. This allows TCRs held in this zone by specific interactions with their ligands on the APC to be phosphorylated and so begin the signal cascade towards T cell activation. This model was proposed in 1996 (Davis, S. J. and van der Merwe 1996), and a similar model has been suggested subsequently (Shaw and Dustin 1997), to account for several observations about the nature of antigen recognition at the T cell surface:

1. The dimensions of the TCR-peptide-MHC complex are close to those of many of the important accessory molecules (e.g. CD2-CD58 or CD28-CD80) but much smaller than other abundant cell surface proteins such as CD43 and CD45 or the molecules involved in initial cell adhesion (LFA-1 and ICAM1).

2. Precise alignment of the T cell and APC membrane is required for optimal TCR-peptide-MHC interaction and this would therefore require that the large molecules mentioned above were excluded from the contact region.
3. Inhibitors of protein tyrosine phosphatases (PTPs) can activate T cells non-specifically but with signalling cascades similar to those seen on specific T cell activation - this suggested that the activatory kinases are constitutively active but that the steady state is held in balance by the PTPs.
4. CD45 is a large and highly abundant surface glycoprotein whose cytoplasmic domain has phosphatase activity. It plays a role in maintaining Lck in a 'primed' state by dephosphorylating an inhibitory tyrosine but could also constitutively inhibit TCR triggering by dephosphorylating molecules such as the ζ chain and ZAP70. CD45 may be associated with the TCR/CD3 complex on resting cells via CD45-associated protein.

In this model, after initial integrin-based adhesion between the T cell and an APC, small regions of close contact between the cells are formed by smaller adhesion molecules such as CD2 binding to CD58. These regions of close contact are still fluid, so that membrane components continue to move through them and encounter molecules on the other cell. However, molecules with large extracellular domains, such as the integrins, CD45 and CD43, are excluded. If a TCR encounters an MHC molecule in this close contact region and it is specific for the peptide displayed, it will bind that MHC molecule and be held in the contact zone for a finite length of time. However, if the incorrect peptide is displayed, both the TCR and MHC molecule will continue to diffuse and will soon leave the contact zone. Because phosphatases such as CD45 are excluded from the contact region, TCRs held in this region are phosphorylated by constitutively active kinases and signal transduction can be achieved. *Figure 1.1* shows the model diagrammatically.

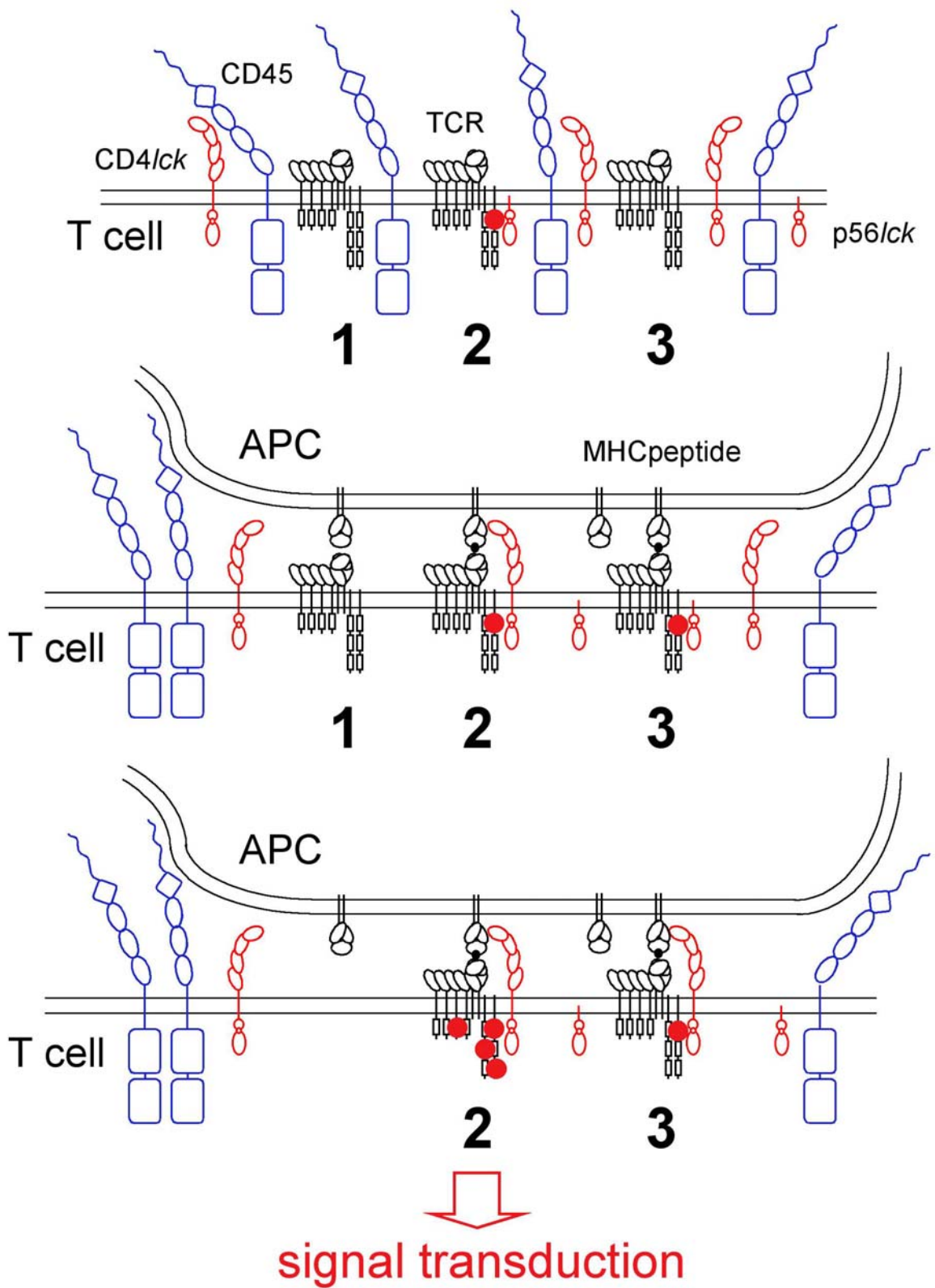


Figure 1.1: The kinetic segregation model of TCR triggering
 For legend see following page

Figure 1.1 (opposite page): The kinetic segregation model of TCR triggering

For simplicity, this diagram only shows a subset of the molecules involved. Red circles indicate tyrosine phosphorylation of receptor cytoplasmic regions.

On the resting T cell surface (top panel), molecules are distributed evenly so that both kinases and phosphatases have access to the TCR and other receptors. Thus, there is no *net* phosphorylation of these receptors. The Src family kinase Lck is maintained in its 'primed' state by the phosphatase CD45.

After initial adhesion with the APC via integrins (not shown), regions of close contact are formed between the T cell and the APC where smaller molecules (such as CD2 or the TCR itself) bind their ligands (e.g. CD58 or peptide-MHC). Molecules with large extracellular domains, including the phosphatase CD45, are excluded from this region (central panel).

TCRs recognising the correct ligand (2 and 3) are retained in this contact region long enough for their associated CD3 chains to be phosphorylated and begin signalling cascades (bottom panel). The other signalling molecules involved may be concentrated in the contact zone by the recruitment of lipid rafts, which contain many of these signalling molecules but not the large molecules excluded from the contact zone. TCRs encountering empty MHC molecules or molecules displaying the wrong peptide are not retained in the contact zone (1).

Putatively, TCR that bind ligands with a faster off rate (2) may require coreceptor binding to bring associated kinases into contact with CD3 quickly enough for signalling to begin before they leave the contact zone. Other TCR, which bind with slower off rates, may be retained in the contact zone long enough to be phosphorylated without kinases being targeted to them by coreceptor association (3).

While it has not been demonstrated that the kinetic-segregation model is a true reflection of TCR triggering, it fits well with several lines of evidence (in addition to the findings that promoted its proposal) that are difficult to incorporate into the other models discussed above. Firstly, because it depends on both TCR retention time in the contact zone and the net kinase activity in this zone, it leaves open possible roles for the many accessory molecules not yet accounted for in other models. Some costimulatory molecules, such as CD28, may initiate separate signalling pathways that affect the outcome of TCR signalling but others could affect the activities or presence of kinases or phosphatases in the contact zone and therefore alter the threshold of TCR signalling itself. Other molecules could affect the threshold of TCR signalling without initiating any signalling from their own cytoplasmic domain, by maintaining the contact zone (a possible role of CD2), retaining the TCR in the zone for longer or bringing kinases into contact with TCR/CD3 complexes more rapidly (a possible role of CD4/CD8). This model may also explain how accessory molecules can generate signals, as they would be held in the kinase enriched contact zone only if they encountered their ligand on the APC. Secondly, in common with others, this model fits with the observation that the half life or off rate of TCR-peptide-MHC interactions, rather than the affinity or on rate correlates with efficient T cell activation (Kersh et al. 1998; Kalergis et al. 2001). Thirdly, it allows for the role of membrane microdomains in the early events of T cell activation, as such domains contain many of the important signalling molecules and exclude large molecules such as CD43 and CD45. These microdomains would therefore aid the passive segregation of molecules in the contact zone and enable the correct signalling complexes to form rapidly by ensuring that the necessary molecules moved into the contact zone together. Fourthly, the formation of multiple, small, fluid, close contact sites across the T cell surface would allow the rapid detection of even small amounts of specific peptide-MHC ligand amongst the excess of self peptide-MHC molecules. This supports the idea of large molecules such as CD45 and the integrins being excluded from areas of close contact

with the APC. Finally, experiments where the dimensions of CD48, the ligand of the accessory molecule CD2, are altered by the insertion of additional domains show that these dimensions are critical for T cell activation - as would be predicted by this model (Wild et al. 1999).

1.5.8 New concepts for cell surface interactions and signalling

The findings and models discussed in this section have added an extra dimension to the consideration of cell surface interactions. They have demonstrated the organisation of molecules into signalling complexes, membrane domains and large cytoskeleton-dependent surface structures. The possibility of kinetically driven segregation of surface molecules has also been suggested. T cell activation is probably the best studied system involving these different forms of molecular interaction and yet it is still not fully understood. Again, knowledge of all the molecules involved and the nature of these interactions both with molecules on the T cell and with their ligands on the APC would aid understanding of this complex system. It is unlikely that T cell activation and regulation are unique in their complexity or in the mechanisms used to achieve the fine balance between failure to respond and over-sensitivity. Other systems, both within (e.g. B cell activation, dendritic cell maturation etc.) and outside (e.g. neuronal plasticity) the immune system may well be found to involve these types of molecular assemblies. Indeed, B cells have already been shown to form synapses when antigen is bound on the surface of cells (Batista et al. 2001) and neurological synapses have been shown to be fluid structures and to involve direct cell-cell contact, making them appear more similar to the immunological synapse (Donnadieu et al. 2001).

1.6 The CD2 subset of the immunoglobulin superfamily

1.6.1 Identification and characteristics of the subset

Interest in CD2 began when antibodies specific for this protein were shown to block erythrocyte rosetting, a phenomenon used to recognise T cells at the time (Springer et al. 1987). Further work on CD2 continued to show that it had important immune functions. When its ligands were identified, they were shown to share sequence similarity with CD2 itself and so this group of molecules were identified as a subset of the IgSF thought to have evolved from a common ancestor (Killeen et al. 1988).

The CD2 family is characterised by sequence-similarity in the extracellular regions of its members, which have a common structure consisting of a membrane proximal C2-set IgSF domain and a membrane distal V-set IgSF domain (Davis and van der Merwe, 1996). The exception is CD229 (Ly9) where this domain pair is duplicated to give a four domain extracellular region. CD48 and one splice variant of CD58 are GPI (glycosyl-phosphatidylinositol) anchored membrane proteins, whilst the remainder of the family contain transmembrane and cytoplasmic domains which can be alternatively spliced. All these proteins are involved in the type of interactions at the leukocyte cell surface discussed above and seem to have functions that influence the immune response.

The members that had been identified at the start of this work were CD2, CD58 (LFA-3), CD48, CD244 (2B4), CD150 (signalling lymphocytic activation molecule, SLAM), CD84 and CD229 (Ly9). The domain structure of these molecules is represented in **Figure 1.2**.

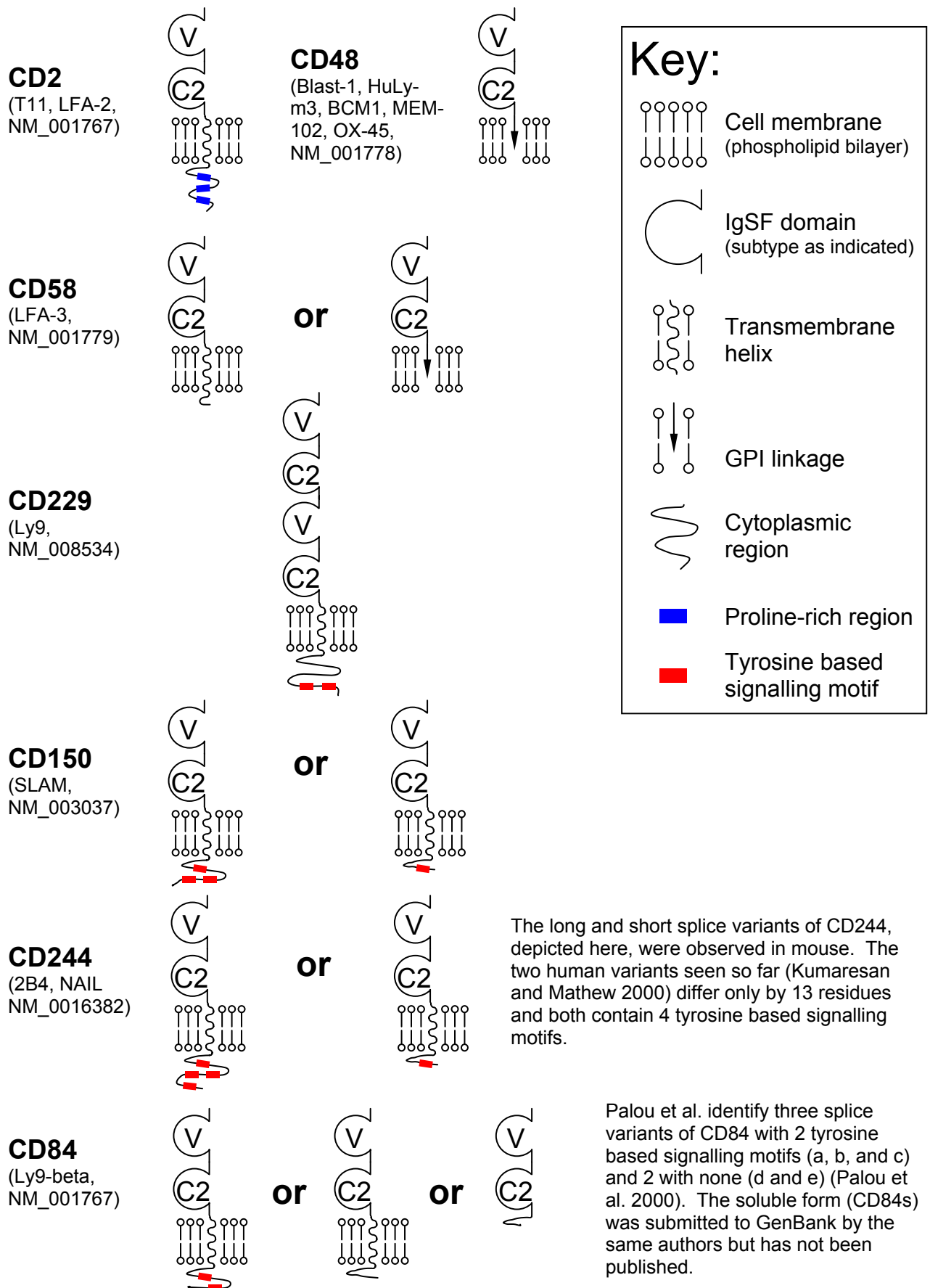


Figure 1.2: Representation of the known members of the CD2 family.

The domain structures of the family members and their alternatively spliced variants are shown according to the key. Alternate names and GenBank accession numbers for one human mRNA are given in brackets.

1.6.2 The CD2 subset as a model for the investigation of cell surface interactions

The CD2 subset is ideally suited to the investigation of the general properties of cell surface interactions for four reasons. (1) Its members contain IgSF domains, the commonest domain type at the leukocyte cell surface. (2) They interact with low affinity and high specificity, as do most cell surface proteins. (3) They interact monomerically, facilitating simple analysis of their interactions. (4) They appear to interact with other members of the subset, aiding ligand identification. The use of members of this subset to investigate the structural mechanisms behind the binding properties of cell surface molecules was discussed in section 1.4.3. This subset also forms a model of the evolution of new specificities at the leukocyte cell surface and of the network of interactions that can result.

The study of this family is therefore useful for three reasons: Firstly, its characterised members have important immunological functions in their own right (see below), and other members can therefore also be expected to have such functions. Secondly, the interactions of this family serve as a model for the nature of low affinity, high specificity interactions at the cell surface. Thirdly, this family is an example of the evolution of multiple receptor specificities from an ancestral protein. This form of duplication and modification has long been thought to be the most common mechanism for the evolution of new genes (see (Ohno 1999)).

1.6.3 Evolution of the CD2 family

The CD2 family is believed to have evolved via successive duplications of a single ancestral gene, encoding a homophilic receptor (Wong, Y. W. et al. 1990). This belief is supported by the fact that these proteins are encoded by genes clustered at two duplicated loci in humans (both on chromosome 1, either side of the centromere, 1q21-24 and 1p13 (Sewell et al. 1988; Kingsmore et al. 1989; Kingsmore et al. 1995; Tangye et al. 2000)) and at the syntenic

regions in mice (on chromosomes 1 and 3 (Wong, Y. W. et al. 1990; Tangye et al. 2000)). A common evolutionary origin is also suggested by the significant overlaps in their specificities and signalling properties: CD2 binds CD58 in humans (Selvaraj et al. 1987) and CD48 in rodents (Kato et al. 1992) but CD48 also binds 2B4 in mice and humans (Brown, M. H. et al. 1998). The cytoplasmic domains of CD244 (Tangye et al. 1999), CD150 (Sayos et al. 1998), CD84 and CD229 (Sayos et al. 2001) bind the signalling molecule SLAM-associated protein (SAP). Finally, antibody cross-linking of either CD2 (van der Merwe 1999) or SLAM (Cocks et al. 1995) can induce T cell proliferation and antibodies specific for 2B4 can induce natural killer (NK) cells to proliferate and kill target cells (Garni-Wagner et al. 1993). In addition to this evidence, it has been suggested that CD150 is a self-ligand (Punnonen et al. 1997), raising the possibility that it is more closely related to the putative ancestral homophilic molecule than any of the other proteins investigated so far.

As noted above, there are differences in the specificities of members of the CD2 family between rodents and humans. The study of these differences may help our understanding of how receptor specificities evolve. In mice, CD48 appears to be the dominant ligand of both CD2 (van der Merwe et al. 1993) and 2B4 (Brown, M. H. et al. 1998), whereas in humans the principle ligand of CD2 is CD58 (Selvaraj et al. 1987; van der Merwe et al. 1994) and that of 2B4 is still CD48 (Brown, M. H. et al. 1998).

1.6.4 The function of the members of the CD2 family

The various molecules of the CD2 subset have been shown to have a major role in at least NK and T cell functions (reviewed in (Tangye et al. 2000)). The functions of individual members of this subset will be considered in turn below, but first a feature of the signalling induced by several members of the family needs to be addressed. The cytoplasmic domains of CD84, CD150, CD229 and CD244 all contain one or more copies of the tyrosine-based motif

TxYxxV/I/A (Cocks et al. 1995; Sandrin et al. 1996; Palou et al. 2000). X-linked lymphoproliferative syndrome (XLP) is a genetic disease in which patients are unable to control B cell proliferation induced by Epstein-Barr virus infection. The gene whose disruption causes XLP is SAP, which has been shown to bind to one of these motifs in the cytoplasmic domain of all four proteins (Sayos et al. 1998; Tangye et al. 1999; Sayos et al. 2001). SAP regulates signal transduction through these receptors and appears to do so by controlling the binding of SHP-2 (SH2-containing protein tyrosine phosphatase) or SHIP (SH2-containing inositol phosphatase) to their cytoplasmic domains (Shlapatska et al. 2001). This motif has therefore been named an immunoreceptor tyrosine-based switch motif (ITSM), as it has a similar but distinct function to the ITIM and ITAM motifs.

The function of CD2 and its ligands

Different anti-CD2 monoclonal antibodies (mAb) have been shown to block human and mouse T cell function in vitro (Guckel et al. 1991) and to be able to activate $\gamma\delta$ T cells (Wesselborg et al. 1991). $\alpha\beta$ T cells cannot be activated by single anti-CD2 mAbs, but certain combinations of such mAbs, recognising separate epitopes on CD2, can induce their activation and proliferation (Meuer et al. 1984; Clark et al. 1988).

CD2 is expressed on T cells, thymocytes, NK cells and a small subset of thymic B cells (Krensky et al. 1983; Yagita et al. 1989; Yagita et al. 1989). As indicated above, it is now well established that the major CD2 ligand in humans is CD58 (LFA-3, (Selvaraj et al. 1987; van der Merwe et al. 1994)), however no homologue of CD58 has been found in rats or mice where the related molecule, CD48, is the major ligand (Kato et al. 1992; van der Merwe et al. 1993). Human CD48 binds CD2 with such a low affinity ($\sim 1\text{mM}$) (Arulanandam et al. 1993; van der Merwe et al. 1994) that the interaction is unlikely to be physiologically significant. Thus, humans and rodents appear to have evolved divergent functional CD2 ligands. CD48

and CD58 are widely expressed on haemopoietic cells and CD58 is also expressed on non-haemopoietic lineages, notably endothelium. These ligands are therefore expressed on nearly all the antigen presenting cells commonly interacting with T cells (Sanders et al. 1988; Yokoyama et al. 1991).

The binding of CD2 to CD58 has long been known to significantly enhance the efficiency of antigen recognition *in vitro* (Moingeon et al. 1989; Koyasu et al. 1990). This interaction is believed to promote adhesion between T-cells and APCs and in this way promote TCR binding to cognate peptide-MHC complexes, as suggested in the kinetic segregation models (Davis, S. J. and van der Merwe 1996; Shaw and Dustin 1997) described in section 1.5.7. Evidence from immunoprecipitation experiments suggests CD2 is weakly associated with the CD3/TCR complex in the membrane (Brown, M. H. et al. 1989; Beyers et al. 1992) and this association could increase the density of the TCR in the 'contact zone'. Confocal microscopy pictures show CD2 recruitment into the immunological synapse (Leupin et al. 2000). In addition to its role as an adhesion molecule, CD2 is believed to have a signalling function because of the effects on T-cell activation of combinations of mAbs specific for this molecule (Meuer et al. 1984; Clark et al. 1988). Such a role is supported by the identification of molecules associated with the cytoplasmic tail of CD2: CD2-associated protein (CD2AP (Dustin et al. 1998)) and CD2 binding proteins 1 (CD2BP1 (Li et al. 1998)) and 2 (CD2BP2 (Nishizawa et al. 1998)). These molecules are believed to signal cytoskeletal rearrangements in response to CD2 ligation in the immunological synapse.

The importance of CD2 in T-cell activation *in vivo* became unclear when CD2-deficient mice appeared to develop a normal immune system (Killeen et al. 1992). Since then, however, detailed investigation of CD2-null mice has revealed defects in positive selection and antigen-specific activation of CD8⁺ T-cells (Teh et al. 1997). Furthermore, CD2-deficient mice

expressing single recombinant TCR chains (Bachmann et al. 1999) have revealed that CD2 has a subtle role in T cell activation, possibly setting TCR affinity thresholds for activation or fine tuning the TCR response (reviewed in (van der Merwe 1999; Dumont et al. 2001)). More recently, similar studies in mice expressing another transgenic TCR β chain and lacking CD2 have shown defects in T cell development, α chain repertoire and T cell activation (Sasada and Reinherz 2001).

Thus, CD2 and its ligands are important to T cell function but their effect is subtle, affecting individual clonal populations of T cells differently. Such fine tuning of the T cell response may be necessary to maintain a sufficient T cell repertoire while avoiding activation of self-specific T cells.

The function of CD244 (2B4)

Using mAbs, CD244 was found to be expressed on all cells capable of non-MHC restricted killing i.e. NK cells and some T-cells, notably those that have been cultured in high concentrations of IL-2 and dendritic epidermal $\gamma\delta$ T-cells (DETCs) (Garni-Wagner et al. 1993). If spleen cells are sorted using CD244 as a marker, all non-MHC restricted cytotoxicity is contained within the CD244⁺ population (Garni-Wagner et al. 1993) and in DETCs, CD244 expression correlates with capacity to lyse transformed keratinocytes (Schuhmachers et al. 1995, 1995). mAbs specific for CD244 can greatly enhance the lytic activity of NK cells and non-MHC-restricted T-cells against a wide variety of targets (Garni-Wagner et al. 1993; Schuhmachers et al. 1995).

CD244 appears to be expressed in three alternatively spliced forms in mice, as seen by northern blot analysis (Mathew et al. 1993). Only two protein forms are known, which differ only in their cytoplasmic domains, with the shorter form lacking three of the tyrosine based

signalling motifs of the longer form (Schatzle et al. 1999; Stepp et al. 1999). Transfection of the two forms into a murine NK cell line suggests that the long form inhibits lysis and the short form activates lysis (Schatzle et al. 1999). The long form, unlike most NK cell inhibitory receptors, does not have a classical ITIM motif and cannot bind SHP-1 [Schatzle, 1999 #39]. However, as indicated above, it can bind both SHP-2 and SAP via its non-classical ITSM signalling motifs (Tangye et al. 1999).

2B4 has been shown to be a ligand for CD48 in both humans and mice with an affinity comparable to that of other cell surface protein-protein interactions ($K_d = 8\mu\text{M}$ or $15\mu\text{M}$ respectively (Brown, M. H. et al. 1998)). This affinity is slightly higher than that of CD2 binding with CD48 in rats and mice ($K_d \sim 75\mu\text{M}$ and $90\mu\text{M}$ respectively, see **Table 1.1**).

The function of CD150 (SLAM)

CD150 was initially identified with a mAb that activated T-cells and bound a previously unknown cell-surface antigen (Cocks et al. 1995). It is expressed by early immature thymocytes, CD45RO⁺ T-cells and some B-cells, but not on mature CD45RO⁻ CD3^{high}, single CD4⁺ or CD8⁺ thymocytes (Sidorenko and Clark 1993; Cocks et al. 1995; Punnonen et al. 1997). Its expression is rapidly induced on CD45RO⁻ T-cells and enhanced on CD45RO⁺ T-cells by their activation. In mice, it can also be upregulated on activation in B-cells and dendritic cells, but not in NK cells. Expression of CD150 appears to be maintained on highly polarized T_h1 but not T_h2 cells (Castro et al. 1999) after activation. Antibody cross-linking of CD150 influences both T-cell (Aversa et al. 1997; Aversa et al. 1997) and B-cell (Sidorenko and Clark 1993; Punnonen et al. 1997; Mikhalap et al. 1999) responses to antigen.

Three forms of CD150 have been identified from activated T- and B-cell mRNA: two encoding membrane-bound forms, one with a truncated cytoplasmic domain, and a soluble,

secreted form (Cocks et al. 1995; Punnonen et al. 1997). The longest form has three SAP binding ITSM motifs in its cytoplasmic domain.

The prevailing view is that SLAM is a self-ligand, involved in T- and B-cell interactions via bidirectional signalling (Sayos et al. 1998). However, preliminary reports suggested a homophilic affinity for SLAM that was extremely high compared to other cell-surface interactions ($K_d \sim 0.1\text{nM}$, (Punnonen et al. 1997)). This affinity would make the interactions of two cells expressing SLAM effectively irreversible, contradicting the models of cell surface interactions discussed so far. Further investigation of this interaction was therefore undertaken as part of this thesis.

The functions of the CD84 and CD229

The functions and ligands of CD84 and CD229 are not yet known. However, CD84 is expressed on lymphocytes and monocytes (de la Fuente et al. 1997) and CD229 on lymphocytes and thymocytes (de la Fuente et al. 2001), indicating roles in the immune system. The presence of ITSM motifs in their cytoplasmic tails and their ability to bind SAP (Sayos et al. 2001) may suggest roles as activatory or inhibitory receptors in a similar manner to CD150 and CD244.

1.7 Aims of thesis

1) *Investigation of the nature of low affinity protein-protein interactions at the leukocyte cell surface using members of the CD2 subset of the IgSF.*

In particular, this thesis aims to determine the structural basis of low affinity, high specificity interaction of rat CD48 for CD2, and to understand its implications for models of protein recognition at the cell surface. In addition, the dual specificity of rat CD48 for both CD2 and CD244 was studied by surface plasmon resonance. Reports of an unusually high affinity homophilic interaction of the related molecule CD150, which call into question the general nature of these types of interaction, were also investigated.

2) *Investigation of the evolution of different specificities among cell surface molecules.*

During these studies, two recently duplicated isoforms of rat 2B4 were isolated. The significance of the considerable variation focussed in the binding face of these isoforms, and its implications for the evolution of new specificities, was investigated. In addition, genomic methods were employed to identify the full set of CD2-related genes and pseudogenes.

3) *Investigating the full complement of proteins expressed at the leukocyte cell surface, using a cytotoxic T cell clone as a model cell type.*

In this work, we will use the recently developed SAGE method, along with the recently completed first draft of the human genome sequence, to characterise systematically gene expression patterns in leukocytes. The ultimate aim of such studies is to identify the complete set of molecules in the immune function of these cells. This thesis begins the identification of this complete set by describing the transcriptome of one particular cell - a cytotoxic T cell clone.