LEUKOSIALIN AND CD2:
CONTRASTING STRUCTURES AT THE LEUKOCYTE SURFACE

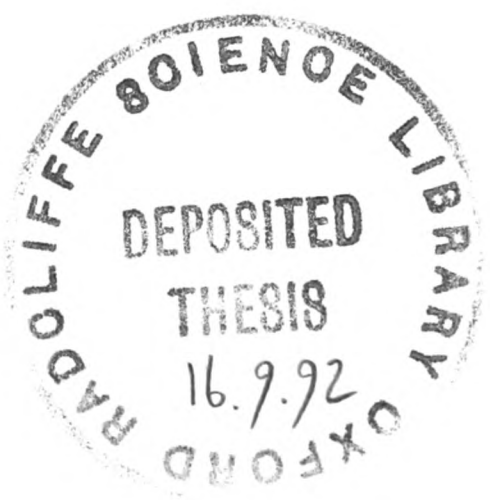
Jason George Cyster

Wolfson College

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Trinity Term 1992



To Chamorro and to my family

As the field of light expands,
so too does the surrounding field
of darkness.

Albert Einstein

CONTENTS

	Page
Abstract	i
Acknowledgements	ii
Publications	iii
Abbreviations	iv
CHAPTER 1 General Introduction	1
CHAPTER 2 Materials and Methods	43
CHAPTER 3 Protein sequence and gene structure for mouse leukosialin	72
CHAPTER 4 Expression of leukosialin on B cells in transgenic mice	83
CHAPTER 5 The dimensions of leukosialin and identification of linear protein epitopes that can be modified by glycosylation	93
CHAPTER 6 Characterization of the cytoplasmic domain of leukosialin	106
CHAPTER 7 The importance of cross-linking and epitope location in the homotypic aggregation of lymphocytes induced by anti-leukosialin antibodies	118
CHAPTER 8 Structure of domain 1 of rat CD2 antigen and expression of human CD2 and LFA-3 domain 1 in *E.coli*	127
CHAPTER 9 General Discussion	141
REFERENCES	158

ABSTRACT

Leukosialin and CD2: Contrasting structures at the leukocyte surface

Jason George Cyster

A thesis submitted for the degree of Doctor of Philosophy

Trinity Term 1992 Wolfson College Oxford University

Leukosialin is a major cell-surface sialoglycoprotein of unknown function expressed on all leukocytes except B cells.

Two clones for mouse leukosialin were isolated from a genomic library and sequencing revealed one of these to be a non-functional pseudogene, whereas the other contained an uninterrupted open reading frame that encoded leukosialin. The possibility that leukosialin might be involved in cell migration pathways was tested by expressing the antigen on B cells in transgenic mice but these cells did not show any aberrant migratory behaviour.

The heavy glycosylation of the leukosialin extracellular domain, with one O-linked glycan per 3 amino acids, suggests an extended structure and this has now been confirmed by electron microscopy. The antigenicity of this domain was studied using nine monoclonal antibodies (mAbs) whose binding is differentially affected by the cell type on which leukosialin is expressed and by the removal of sialic acid. Seven of these mAbs reacted clearly with the extracellular domain expressed in an unglycosylated form in *E.coli* and epitope mapping demonstrated that these epitopes correspond to short protein sequences. It thus appears that linear protein epitopes are recognized and these can be modified in the native structure by glycosylation.

MAbs to human leukosialin are potent at inducing cell clustering and this was shown to hold also for mAbs to rat leukosialin. Monomeric Fab fragments were inactive, demonstrating that cross-linking of leukosialin molecules is required for cell clustering.

The cytoplasmic domain of rat leukosialin was expressed in *E.coli* and shown to have an elongated structure and to be amenable to phosphorylation by protein kinase C.

CD2 is an important T cell adhesion molecule, interacting with LFA-3 on other cells through its NH_2-terminal domain. This domain is also of interest as it has been predicted to have an immunoglobulin type fold despite the absence of the normally conserved disulphide bond. The expression of rat CD2 domain 1 in *E.coli* and the three-dimensional structure of the domain, determined by NMR spectroscopy, are reported.

ACKNOWLEDGEMENTS

I sincerely thank Alan Williams for his support and encouragement, and for inspiring a rigorous and clear minded approach to scientific problems. I regard myself as being very privileged to have had the chance to conduct a D.Phil under Alan's supervision and I hope that I can carry forth some of the enthusiasm he expressed for science and the generosity he showed to the scientific community.

Neil Barclay, Don Mason, Simon Davis and Albert Beyers must be thanked for their helpful advice and stimulating discussions at many times over the three years I have spent in the unit. I thank Dimitris Kioussis, David Shotton and Paul Driscoll for their willingness to collaborate and for their important contributions in what became some exciting experiments. I express my gratitude to all members of the MRC cellular immunology unit as each and every one provided advice, reagents, and encouragement at one time or another.

The perpetual supply of monoclonal antibodies from Mike Puklavec and his advice on cell culture was much appreciated and will be sorely missed as will the unrelenting enthusiasm of Mike Coates in looking after so many mice. Keith Gould's willingness to synthesize oligonucleotides at almost any hour was greatly appreciated. I also thank Tony Willis for protein sequence and compositional analysis and Steve Simmonds for help and advice.

Thanks to Katie Lee and Jill Lively for washing so many flasks and bottles without ever a complaint.

Simon Davis, Neil Barclay, Don Mason, Marion Brown, Anton van der Merwe, Albert Beyers, Chamorro Somoza and Mark Wright I thank for proof reading Chapters of this thesis.

I gratefully acknowledge the financial support of the British Council.

Finalmente muchas gracias a Chamorro por estar alli y por intentar hacerme mas tolerante.

PUBLICATIONS

Cyster, J.G., Somoza, C., Killeen, N. and Williams, A.F. (1990) Protein sequence and gene structure for mouse leukosialin (CD43), a T lymphocyte mucin without introns in the coding sequence. *Eur. J. Immunol* **20**, 875-81.

Cyster, J.G., Shotton, D.M. and Williams, A.F. (1991) The dimensions of the T lymphocyte glycoprotein leukosialin and identification of linear protein epitopes that can be modified by glycosylation. *EMBO J.* **10**, 893-902.

Driscoll, P.C., Cyster, J.G., Campbell, I.D. and Williams, A.F. (1991) Structure of domain 1 of rat T lymphocyte CD2 antigen. *Nature* **353**, 762-65.*

*A copy of this publication is included at the end of this thesis.

ABBREVIATIONS

A_{280}	absorbance at 280 nm over a 1cm path length
β2m	β-2 microglobulin
bp	base pair
BM	bone marrow
BSA	bovine serum albumin
CCP	complement control protein
CD	cluster of differentiation
CHO	chinese hamster ovary
Con A	concanavalin A
cpm	counts per minute
D_2O	deuterated water
DAG	diacylglycerol
DEAE	diethylaminoethyl
DFP	diisopropylfluorophosphate
dH_2O	double deionized water
DMSO	dimethyl sulphoxide
DNP-BGG	dinitrophenylated bovine γ-globulin
DOC	deoxycholate
DxS	dextran sulphate
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
EGTA	ethyleneglycol-bis-(β-aminoethylether)-N,N'-tetraacetic acid
ER	endoplasmic reticulum
FACS	fluorescein activated cell sorter
FcR	Fc receptor
FCS	foetal calf serum
FITC	fluorescein isothiocyanate
Fuc	fucose
Gal	galactose
GalNAc	N-acetyl-galactosamine
Glc	glucose
GlcNAc	N-acetyl-glucosamine
G418	geneticin
GPI	glycosyl-phosphatidylinositol
GM-CSF	granulocyte-macrophage colony stimulating factor
GST	glutathione-S-transferase
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
HEV	high endothelial venule
hr	hour(s)
IAA	iodoacetic acid
ICAM	intercellular adhesion molecule
IFNγ	gamma interferon
Ig	immunoglobulin

IgSF	immunoglobulin Superfamily
IL	interleukin
i.p.	intraperitoneal
IP_3	inositol (1,4,5) triphosphate
IPTG	isopropyl β-D-thiogalactopyranoside
i.v.	intravenous
kb	kilobase pairs
kd or kD	kilodalton
Le	Lewis
LFA	leukocyte function associated antigen
LPS	lipopolysaccharide
mAb	monoclonal antibody
Man	mannose
MBP	myelin basic protein
MHC	major histocompatability complex
min	minute(s)
Mr	apparent molecular weight
NCAM	neural cell adhesion molecule
NeuNAc	N-acetyl-neuraminic acid
NK	natural killer
NMR	nuclear magnetic resonance
NP40	nonidet P 40
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDGF	platelet derived growth factor
PHA	phytohaemagglutinin
PKC	protein kinase C
PLC	phospholipase C
PMA	phorbol 12-myristate 13-acetate
PMSF	phenylmethylsulphonyl fluoride
PNA	peanut agglutinin
RAM	rabbit anti-mouse Ig $F(ab')_2$ antibodies
RT	room temperature
SDS	sodium dodecyl sulphate
SPF	specific pathogen free
SRBC	sheep red blood cell
TcR	T cell receptor
TDL	thoracic duct lymphocytes
TEMED	N,N,N',N''-tetramethylethylenediamine
TNF	tumour necrosis factor
WAS	Wiskott Aldrich Syndrome
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

CHAPTER 1 GENERAL INTRODUCTION

1.1 Introduction

1.2 Antigen specific recognition by T lymphocytes

- 1.2.1 The TcR/CD3 complex
- 1.2.2 Other TcR associated molecules
- 1.2.3 Signaling via the TcR complex

1.3 Membrane-membrane interactions in the immune system

- 1.3.1 CD4 positive T cell- antigen presenting cell
- 1.3.2 Leukocyte-endothelial cell
- 1.3.3 Cell surface charge and membrane-membrane interactions

1.4 Structural aspects of cell surface proteins

- 1.4.1 Superfamilies of surface molecules: a) Immunoglobulin, b) Fibronectin type III, c) cytokine receptor, d) Complement control protein, e) Epidermal growth factor
- 1.4.2 Multi-domain proteins

1.5 Glycosylation

- 1.5.1 N-glycosylation
- 1.5.2 O-glycosylation
- 1.5.3 Types of O-glycosylated cell surface molecules
- 1.5.4 Molecules with alternatively spliced exons that encode O-glycosylated regions

1.6 Mucins

1.7 Integrins

- 1.7.1 LFA-1 and its ligands ICAM-1, ICAM-2 and ICAM-3
- 1.7.2 VLA-4 and VCAM-1

1.8 Selectins and their ligands

1.9 Signal transduction and cytoplasmic domains

1.9.1 Signal transduction and single transmembrane regions

1.9.2 Cytoplasmic domains

1.9.3 Protein kinase C

1.10 CD2, LFA-3 and related molecules

1.10.1 The CD2 molecules

1.10.2 The epitopes of CD2

1.10.3 CD2 and cell activation

1.10.4 LFA-3 (CD58)

1.10.5 CD2 - LFA-3 mediated adhesion

1.10.6 GPI vs transmembrane linkage of LFA-3 and adhesion

1.10.7 Further ligands for CD2: a) CD48 (OX45, Blast 1), b) CD59

1.10.8 A second rosette receptor on T cells (E2)

1.11 Leukosialin

1.11.1 The leukosialin molecule

1.11.2 Leukosialin tissue distribution

1.11.3 Leukosialin heterogeneity

1.11.4 Biochemical properties

1.11.5 Leukosialin as an activation molecule

1.11.6 The O-linked glycans of leukosialin

1.11.7 Wiskott Aldrich Syndrome (WAS)

1.11.8 Soluble leukosialin

1.11.9 Leukosialin-ICAM-1 interaction

Aims of the work described in this thesis

1.1 General introduction

Effective protection against pathogenic agents entering the body through almost any surface requires continual surveillance of these surfaces and to achieve this the immune system has evolved as a complex system of migratory cells that circulate in the body ready to be recruited to sites of antigen entry in the tissues. Lymphocytes emerging as naive cells from thymus (T cells) and bone marrow (B cells) travel passively in the blood stream until they reach venules which pass through lymphoid organs. The cells then adhere to the endothelium, become polarized and motile and enter the lymphoid organ where they migrate along specific paths that bring them into contact with large numbers of antigen presenting cells. Some of the antigen presenting cells will have entered the lymphoid organ after migrating from peripheral locations where they may have been exposed to foreign antigens. Any lymphocytes that recognize the antigen complex carried by these cells will be activated. Lymphocytes which are not stimulated exit the lymph node via the efferent lymphatics and are returned to the blood stream to repeat the cycle. Cells that are never stimulated by specific recognition of antigen, which will be the majority of lymphocytes in the normal individual, become aged over a period of days and are removed by phagocytic cells. Lymphocytes that have been stimulated undergo proliferation and differentiation into longer lived cells, and they express an altered repertoire of surface proteins. These cells, in particular the T lymphocytes, upon being returned to the blood, circulate until they encounter a receptive endothelium in a peripheral tissue. The cells cross the endothelium and migrate actively over the extracellular matrix until encountering a target cell, and then carry out their effector responses which will include release of cytokines and cell killing. A schematic illustration of some of these events is presented in Fig 1.1.

It is clear that in addition to the specific antigen receptors which these cells must carry, a variety of other surface molecules will be crucial for effective immune function. This must include molecules which support adhesion to endothelial cells, to antigen presenting cells and to target cells, molecules which support adhesion to and migration over the extracellular matrix, and molecules which prevent unwanted interactions with non-receptive endothelial cells, non-antigen presenting cells and non-target cells and to inappropriate components of the extracellular matrix. This thesis is concerned with two leukocyte surface molecules, leukosialin and CD2, which may have roles in adhesion to endothelial, antigen presenting and target cells, in sending co-stimulatory signals to T cells during antigen recognition, and possibly in reducing unwanted adhesion events.

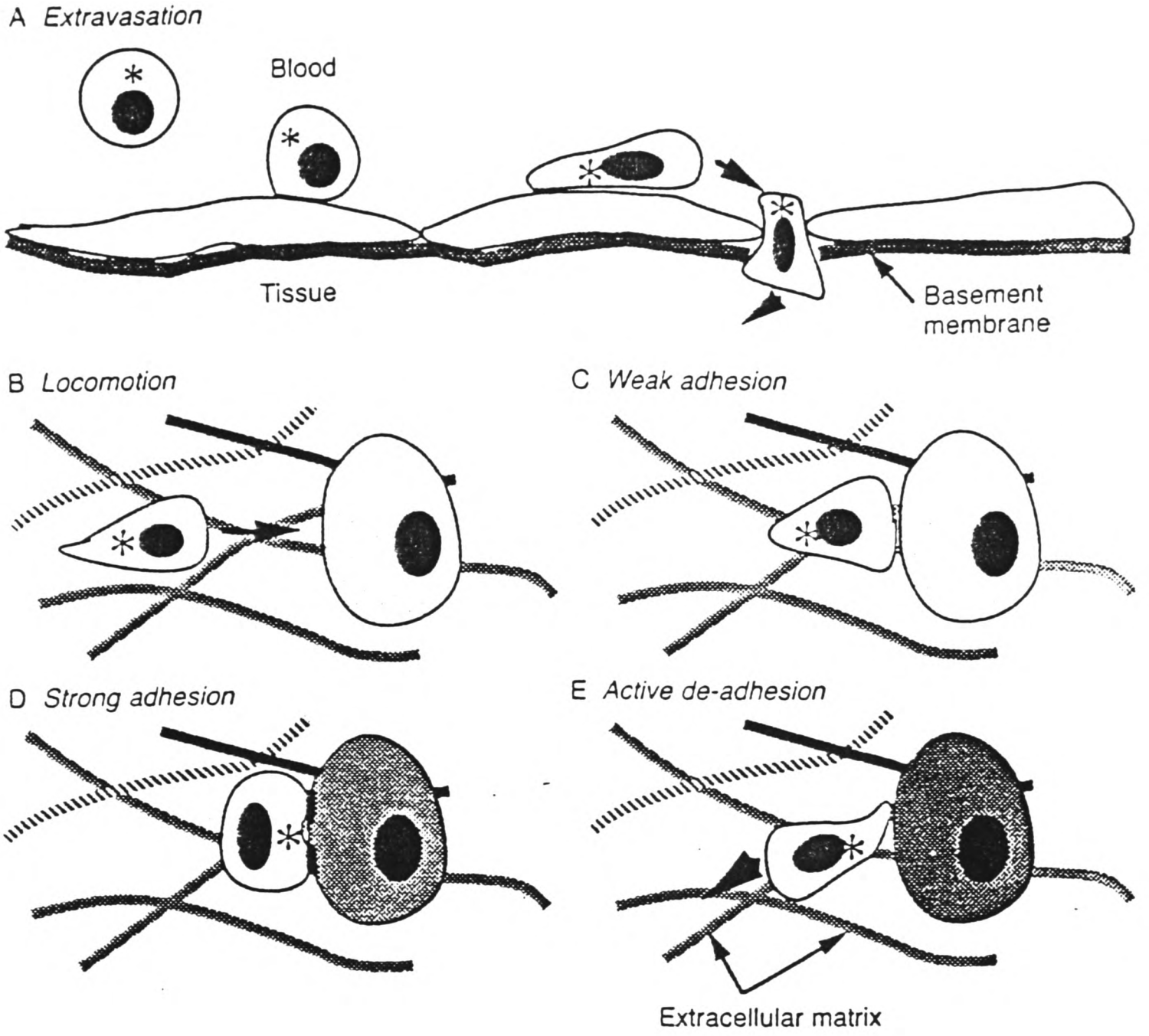


Figure 1.1 Some of the interactions leukocytes undergo in performing their function. The darkly coloured cell is a target cell and the * indicates the location of the microtubule organizing centre in the lymphocyte. From Dustin and Springer (1991).

To put the involvement of CD2 and leukosialin in T cell adhesion and costimulation into perspective, a brief summary of T cell antigen recognition and activation follows. The leukocyte endothelial cell interaction is then introduced, and cell surface charge, to which leukosialin makes a major contribution, and its influence on cell-cell interactions is addressed.

1.2 Antigen-specific recognition by T lymphocytes

The acquired branch of the immune response involves recognition of foreign antigen by clonotypic surface receptors on lymphocytes and the ensuing activation of these cells. Most T lymphocytes with receptors that have high affinity for self components are eliminated or inactivated in the thymus during maturation. T cell receptors (TcR) are predicted to be immunoglobulin superfamily molecules (IgSF) and are specific for processed antigen in the form of peptide presented in the groove between α-helices on the MHC molecules (Bjorkman et al., 1987). Peptides of 8-9 amino acids, derived largely from endogenously synthesized proteins, are presented by class I molecules (Townsend et al., 1984; van Bleek and Nathenson, 1990; Falk et al., 1991) whereas class II molecules generally present larger peptides (13-17 amino acids) generated from endocytosed proteins (Harding et al., 1988; Rudensky et al., 1991).

Recognition by the TcR involves the complementarity determining regions interacting with both outpointing side chains of the bound peptide and with the α-helical regions of the MHC molecule (Jorgensen et al., 1992). T cells expressing the accessory molecule CD4 recognize antigenic peptides in the context of MHC class II molecules on specialised antigen presenting cells, whereas CD8 positive cells are restricted to peptide-MHC class I complexes which can be expressed on almost any cell of the body (Brodsky and Guagliardi, 1991). The CD4 cells act as helper or inducer cells for both CD8 cell and B cell responses and they also produce cytokines which stimulate proliferation and differentiation of cells involved in other aspects of the immune response. The CD8 cells are responsible for the cell mediated response, killing cells that present foreign peptides by the directed release of cytotoxic mediators. Activation of B cells by T cell dependent antigens requires cross-linking of surface immunoglobulin molecules on the B cells by multivalent antigen, uptake and processing of the antigen and presentation, via class II MHC molecules, to CD4 positive T cells (Lanzavecchia, 1990). A subset of CD4 cells produce cytokines that, combined with signals from the cell-cell contact, lead to proliferation and differentiation of the B cells into Ig secreting cells and memory cells.

1.2.1 The TcR/CD3 complex

The antigen recognition complex of the T cell is made up of the clonotypic α and β (or γ and δ) chains of the T cell receptor, and the invariant CD3 chains, γ, δ, ε, and ζ or η. The γ, δ and ε chains are similar in having a single extracellular IgSF domain and relatively long cytoplasmic domains. The ζ chain, and the alternatively spliced form, η, are structurally different from the other CD3 chains and are located at a different site in the genome. They have a 9 amino acid extracellular domain and long cytoplasmic domains and can be associated with the TcR as ζζ or ζη dimers (Klausner and Samelson, 1991). The cytoplasmic domain of the ζ chain alone is able to mediate signaling if attached to an extracellular domain and cross-linked by mAb (Irving and Weiss, 1991; Romeo and Seed, 1991). More recently is was shown that other components of the TcR complex are able to mediate signaling independent of ζ (Wegener et al., 1992). The finding that a motif present 3 times in the cytoplasmic domain of ζ, which is essential for signaling, is also present in the cytoplasmic domains of CD3 γ, δ and ε supports the possiblity that all the CD3 chains have signaling roles. This motif has also be seen in the Ig associated molecules mb-1 and B29, and in the FcεRI γ and β chains (Reth, 1989; Wegener et al., 1992).

1.2.2 Other TcR associated molecules

Less tightly associated but coprecipitable with the TcR complex are the accessory molecules CD4 or CD8, CD2, CD5 and CD45 (Brown et al., 1989; Schraven et al., 1990; Volarevic et al., 1990; Burgess et al., 1991; Beyers et al., 1992). The CD4 and CD8 molecules associate through a motif in their cytoplasmic domains with the src family tyrosine kinase, $p56^{lck}$ (Turner et al., 1990) and this molecule coprecipitates with the complex. A second kinase of the src family, $p59^{fyn}$, is also copreciptitated although less efficiently (Samelson et al., 1990; Beyers et al., 1992). The CD4 molecule has an extracellular region composed of four IgSF domains (Clark et al., 1987; Maddon et al., 1987) the first two of which are important in binding to MHC class II and there is some evidence that the TcR and CD4 must interact simultaneously with the same class II molecule for efficient signal transmission (Dianzani et al., 1992). CD8, unlike CD4, exists on membranes as a heterodimer of α and β chains or as an αα homodimer. Both chains are composed of a single IgSF domain followed by an O-glycosylated region but only the α chain cytoplasmic domain is able to associate with $p56^{lck}$. Like the CD4 interaction, CD8 interacts with class I simultaneously with TcR recognition of the class I-peptide complex and provides a costimulatory signal for T cell activation (Connolly et al., 1990; Salter et al., 1990). Direct interactions between purified CD4 or CD8 and MHC molecules have not been

demonstrated but overexpression in transfected cells has provided indirect evidence for binding by CD4 to class II and CD8 to class I MHC molecules (Doyle and Strominger, 1987; Norment et al., 1988). The interaction of CD2 with glycoproteins on the antigen presenting cell contributes to adhesion and enhances signaling through the TcR complex. CD2 and its interaction with LFA-3 and related molecules is described in detail in section 1.10. More recent work has shown that the T cell molecule CD5 is able to interact directly with CD72, a C-type lectin expressed on B cells (Van de Velde et al., 1991). MAb cross-linking of CD5 can enhance responses mediated through the TcR/CD3 complex in a similar manner to CD2 (Spertini et al., 1991a; Altman et al., 1990) and all these molecules may be described as T cell co-receptors rather than just adhesion molecules (Springer, 1990).

CD45 is a large and heavily glycosylated protein that has protein tyrosine phosphatase activity in its cytoplasmic domain (Thomas, 1989; Trowbridge, 1991). The presence of CD45 in T cells is essential for effective signal transmission (Koretzky et al., 1991) and this may in part relate to its ability to dephosphorylate and activate $p56^{lck}$ (Trowbridge, 1991). In addition to coprecipitation and co-capping data (Volarevic et al., 1990; Schraven et al., 1990; Dianzani et al., 1992) the possibility that CD45 is also a co-receptor in T cell interactions is the evidence that CD45 binds the IgSF molecule CD22 (Stamenkovic et al., 1991b). These associations are depicted in Fig 1.2, which also illustrates the molecules LFA-1 and leukosialin binding to ICAM-1. The latter interactions will be described in more detail in sections 1.7 and 1.11. A further association recently described is that between the 4 transmembrane protein OX44 (CD53) and CD2 on rat NK cells and splenic T cells (Bell et al., 1992). Certain mAb specific for OX44 are able to stimulate a pattern of responses similar to anti-CD2 mAb. Whether the molecule is associated with CD2 as part of in the TcR/CD3 complex is not yet clear.

1.2.3 Signaling via the TcR complex

Triggering of the T cell through the TcR, by recognition of peptide- MHC complexes, super antigen- MHC complexes, or by mAb mediated cross-linking, activates at least two signal transduction pathways. The earliest detectable events involve the activity of protein tyrosine kinases leading to phosphorylation of several substrates including phospholipase Cγ-1 (PLCγ-1) and the CD3 ζ chain (Klausner and Samelson, 1991). The second pathway, which is probably driven by protein tyrosine kinase activation, involves PLC-γ1 mediated hydrolysis of phosphatidylinositols to generate diacylglycerol and a number of different phosphoinositides, in particular IP_3 (Alexander and Cantrell, 1989). IP_3 binds to receptors and induces opening of

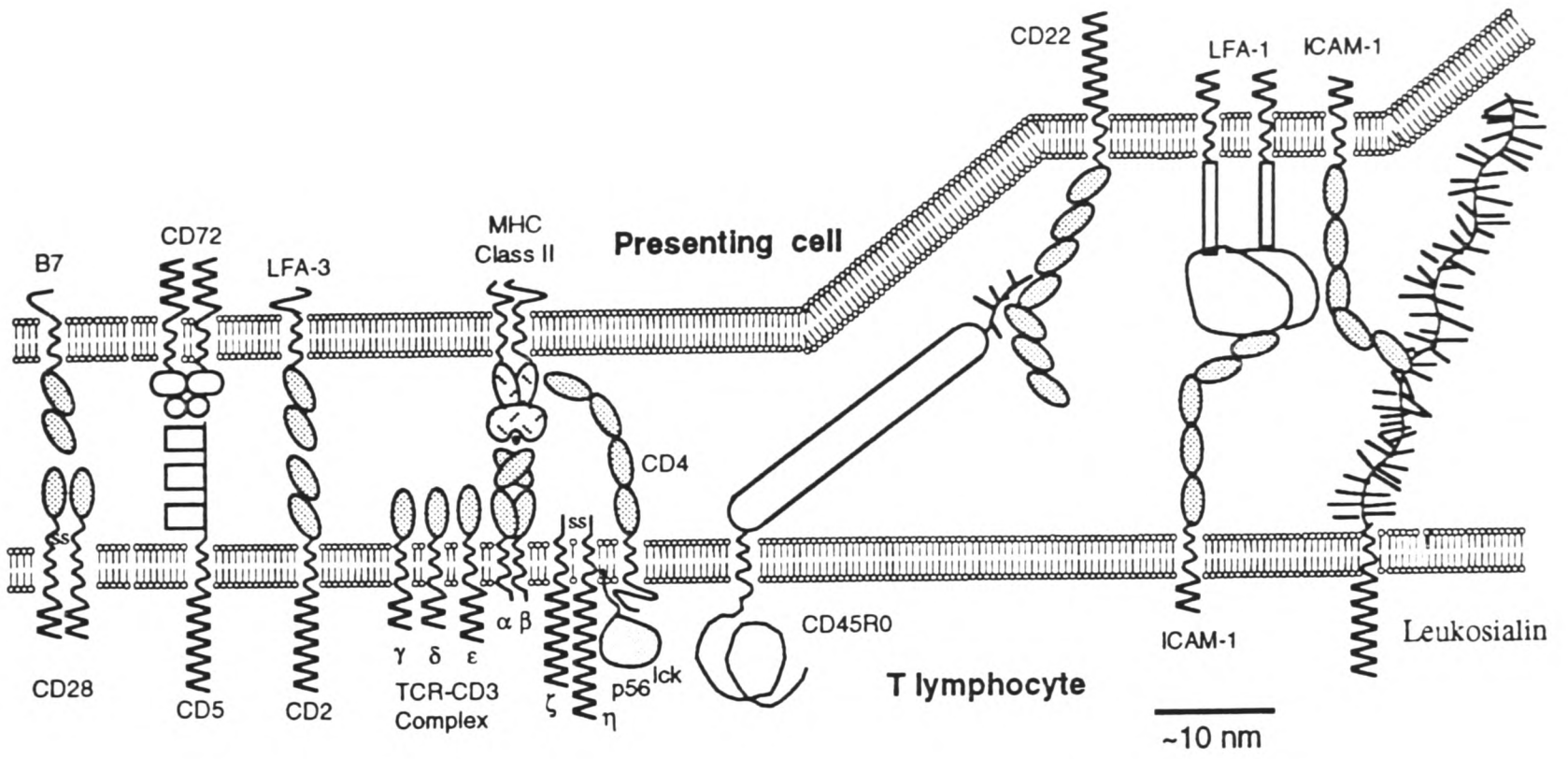


Figure 1.2 Schematic representation of molecules involved in CD4 positive T cell adhesion and activation. The domain structures are described in the text. O-linked glycans on CD45 and leukosialin are illustrated but N-linked glycans are not shown. The molecules are drawn rougly to scale from known dimensions of some of the molecules or their constituent domains.

calcium channels leading to an increase in cytosolic free calcium. This is important for the activity of a number of calcium dependent kinases and phosphatases. In particular, protein kinase C (PKC) is activated by association with diacylglycerol and (for some isoforms) by binding of calcium and this leads to the Ser and Thr phosphorylation of multiple substrates (Alexander and Cantrell, 1989), and causes the rapid activation of secondary signal transducers such as $p21^{ras}$, and c-raf kinase. There is also evidence for involvement of G proteins in some aspects of signaling and a novel GTP binding-protein has been identified associated with CD4/CD8 (Telfer and Rudd, 1991). Ultimately the phosphorylation state of a number of transcription factors, and their relative distribution between cytosol and nucleus, is changed leading to an altered pattern of expression of over 100 genes (Ullman et al., 1990). Some of the earliest genes expressed are transcription factors and the genes for IL-2 and the IL-2Rβ. This generates an autocrine loop and the binding of IL-2 to the high affinity IL-2R is essential for the activated cell to progress from the G1 stage of the cell cycle and undergo cell division. Further genes whose expression is elevated include other cytokines and, some days after the initial stimulation, various adhesion molecules including CD2, LFA-1, LFA-3, CD44 and the VLA antigens (Springer, 1990).

Signaling via the TcR/CD3 complex alone in resting T cells is an inadequate switch for cell proliferation, and may in some cases lead to inactivation or anergy (Mueller et al., 1989). This is seen, for example, when antibodies to the TcR or CD3 are used alone, and may occur *in vivo* when T cells interact with B cells that have taken up monomeric antigen, processed it and presented it without themselves being activated (Eynon and Parker, 1992). It is generally argued that, in the normal situation of a resting T cell interacting with a specialized antigen presenting cell which has taken up and processed antigen, at least two signals are transmitted to the T cell, one via the TcR/CD3 complex (and possibly enhanced by the other associated molecules), and one via a different accessory molecule such as CD28. CD28 is a member of the IgSF and exists on the membrane of T cells as a disulphide linked dimer (June et al., 1990). Signaling through the CD28 molecule is currently the only known method of enhancing IL-2 levels that appears to be independent of the TcR/CD3 complex (June et al., 1990; Altman et al., 1990). A natural ligand for CD28, the IgSF molecule B7, is expressed on B cells and the CD28-B7 interaction is able to induce signaling through CD28 (Damle et al., 1991; Linsley et al., 1991). The signaling mechanism is not understood but it leads to an increase in stability of IL-2 mRNA and increased transcription of IL-2 (June et al., 1989; Fraser et al., 1991).

1.3 Membrane-membrane interactions in the immune system

1.3.1 CD4 positive T cell - antigen presenting cell

As apparent from the above discussion, the activation of a T cell by recognition of antigen on an antigen presenting cell involves a number of molecular interactions in addition to those of the TcR itself. It is still unclear which interactions occur first when a migrating T cell meets a presenting cell or target cell. Non-specific interactions with cells not presenting the antigen generally last for less than an hour and do not involve any major changes in the T cell. Specific interactions in contrast can last for more than 24 hr and involve redistribution of surface molecules, a change in morphology of the T cell such that there is a large area of contact with the antigen presenting cell, and the re-orientation of the microtubule organizing centre towards the site of interaction. Distribution of only a small number of surface molecules have been examined in these systems but, in addition to the TcR and CD4 or CD8, LFA-1 and CD2 have been observed to focus in the region of conjugate formation (Kupfer and Singer, 1989a, b; Andre et al., 1990; Moingeon et al., 1991). The molecular dimensions of the TcR-MHC molecule interaction and the CD4/CD8 - MHC interaction are similar to the CD2-LFA-3, CD28-B7 and CD5-CD72 interactions, involving a membrane separation of approximately 15 nm and this distance is in agreement with the closest appostion of membranes in cell conjugates seen by electron microcopy (Biberfeld and Johansson, 1975). The LFA-1-ICAM interactions will be of larger dimensions and these molecules are likely to be excluded to areas of less tight association (Fig 1.2). The location or orientation of large and highly glycosylated molecules, such as leukosialin and CD45, at sites of cell contact are not well characterized although, as indicated above, there is evidence that such molecules are important at the site of interaction.

1.3.2 Leukocyte - endothelial cell interactions

The interaction of lymphocytes, polymorphonuclear cells, monocytes and dendritic cells with the endothelium of blood vessels is essential in allowing the cells to enter tissues in search of antigen. Interactions with endothelium must differ from those which take place in the tissues because the cells in the blood stream are not in a polarized or motile state (Parrott and Wilkinson, 1981). This interaction must be of adequate kinetics to occur during very short periods of contact between leukocyte and endothelium and be of adequate affinity to secure the leukocyte against the shear forces of the blood flow. The initial interaction of neutrophils with activated enothelium involves repetitive adhesion and deadhesion leading to a rolling effect which is followed by strong adhesion and migration through the endothelium. It appears that the rolling interaction is supported by selectin molecules on the endothelium interacting with

carbohydrate ligands on the neutrophil (Butcher, 1991; Lawrence and Springer, 1991). The subsequent, more rigid adhesion that follows is necessary for neutrophil migration through the endothelial layer, and involves LFA-1-ICAM adhesion. For the integrin mediated adhesion and thus migration to occur, the integrin needs to be in the high avidity state and this may result from exposure to activating/ chemoattracting molecules (such as IL-8 and platelet activation factor) which can be bound to the activated endothelium or released locally by the cells (Huber et al., 1991; Lorant et al., 1991). These stimuli also provoke shedding of the selectin molecules which may be important in allowing diapedesis (Huber et al., 1991).

Homing of activated T lymphocytes to sites of infection in peripheral tissues is probably by interacting with endothelium in the tissues in a similar way to neutrophils (Berg et al., 1991; Picker et al., 1991a; Shimizu et al., 1991). Naive cells must first be exposed to antigen in lymphoid tissues, however, which are specialized for efficient antigen presentation and for extensive cell-cell interactions. Entry into lymph nodes is through "high endothelial venules" (HEV) so called because of the cuboid nature of the endothelial cells which line them. The mechanism of lymphocyte attachment to this endothelium is not fully understood but involves the interaction of selectin molecules on the lymphocyte with molecules carrying carbohydrate ligands on the HEV (Butcher, 1991). The interaction of CD44 with hyaluronic acid maybe important in some stages of lymphocyte attachment and diapedesis (Haynes et al., 1989; Aruffo et al., 1990). Integrins are also involved and in addition to the LFA-1-ICAM interactions, VLA-4-VCAM-1 interactions are important at least in lymphocyte binding to peyers patch HEV (Butcher, 1991). It is likely that further molecular interactions which have not yet been defined and which lead to high tissue specificity and also to a level of redundancy, will be involved.

1.3.3 Cell surface charge and membrane-membrane interactions

Close cell-cell contact between circulating cells is opposed by charge repulsion and by the decrease in entropy required for interdigitation of the surface glycocalyces (Bell, 1984; Springer, 1990). The majority of the charge on the cell surface is due to sialic acid (Wigzell and Hayry, 1974) and a similar amount of sialic acid is carried by glycolipids and glycoproteins (Snary et al., 1976), with the majority of the latter being present on a small number of heavily glycosylated and highly expressed proteins, in particular leukosialin and CD45 (Williams and Barclay, 1986). The large amount of carbohydrate bound to the surface of cells is often termed the glycocalyx, from the Greek word meaning "sweet husk" (Bennett, 1963), and the lipid associated component of the glycocalyx extends about 10 nm from the cell membrane (Jentoft,

1990). In addition to the overall charge of the cell surface, the location of the charge groups and the distance of the groups from the membrane may be important in cellular interactions. Early studies of the electrophoretic mobility of lymphocytes showed that T cells were more mobile than B cells. Treatment of the T cells with neuraminidase reduced their rate to that of the B cells and it was argued that T cells carried more cell surface sialic acid (Nordling et al., 1972; Shortman et al., 1975) but direct measurements showed T and B cells to carry similar amounts (Despont et al., 1975). It was concluded that a subpopulation of sialic residues on lymphocytes was important in determing their electrophoretic mobility.

Cell-cell interactions which are directly influenced by sialylation include NK cell and CTL killing of targets (Kaufman et al., 1981; Rooney et al., 1984; Jiang et al., 1990), responder-stimulator interactions in the mixed lymphocyte reaction (Taira et al., 1986; Powell et al., 1987), T cell-antigen presenting cell interactions (Kearse et al., 1988; Boog et al., 1989; Neefjes et al., 1990), and cancer cell interactions where increased metastatic potential correlates with increased sialylation (Pearlstein et al., 1980; Yogeeswaran and Salk, 1981; Altevogt et al., 1983; Fogel et al., 1983; Collard et al., 1986). In a number of these studies it appeared that only a subset of the cell surface sialic acid was significant (Altevogt et al., 1983; Powell et al., 1987; Laferte and Dennis, 1988; Boog et al., 1989), supporting the idea that the disposition of the charge on the membrane is more important than its total. Thus the type of sialylated molecules a cell expresses is likely to be important in determining general cell-cell interactions. Evidence that sialylation can influence interactions at the molecular level comes from work with NCAM, a homophilic adhesion molecule of the IgSF family (Rutishauser et al., 1988). NCAM has two forms, an embryonic form which carries a poly-sialylated N-glycan, and an adult form which lacks the poly-sialic acid moiety. *In vitro* kinetic studies of the binding interaction shows a reduced on rate for the embryonic form (Moran and Bock, 1988; Acheson et al., 1991). While NCAM is somewhat unusual in carrying a poly-sialylated glycan, the results indicate that interactions between molecules that are heavily glycosylated can be directly influenced by the extent of sialylation of the glycans. The extent of sialylation of MHC molecule glycans, for example, has been argued to be directly reflected in the efficiency of the antigen presenting cell (Cullen et al., 1981; Cowing and Chapdelaine, 1983; Boog et al., 1989).

Activated T and B cells show increased binding of peanut agglutinin (PNA, specific for terminal Gal) (Butcher et al., 1982; Jung et al., 1988) and are less negatively charged than resting cells (Springer, 1990) and this may in part be responsible for the increased adhesion seen between

activated cells and other cells. The interaction between CD2 and LFA-3, for example, appears very sensitive to net charge and this will be described in more detail in section 1.10. Immature cells in the hematopoietic system also tend to have lower levels of sialylation than mature cells (Pink, 1983). It could be imagined with immature T cells, for example, that a large surface of interaction of a thymocyte with an epithelial cell may allow more effective scanning of MHC molecules on the cell membrane which must be important in the process of tolerance induction. Adhesion between red cells and T cells is not only enhanced by removal of sialic acid from either cell but also the type of contact changes from small points to large regions of close association (Bentwich et al., 1973).

1.4 Structural aspects of cell surface proteins

1.4.1 Superfamilies of surface molecules.

The two molecules studied in this thesis are shown to be of very contrasting structure. Both structures were accurately predicted, however, before this work was done based largely on primary sequence data and comparison with other known sequences, some of which had known structures. This highlights the information that can be gained from grouping molecules based on sequence similarity. As more cell surface molecules have been defined it has become apparent that many of them are made up from common building blocks or structural domains. The term Superfamily is used to group segments of protein with sequence similarity of less than 50% whereas molecules or domains with a higher level of identity are said to be in a family (Dayhoff et al., 1983). In protein Superfamilies there is often only 15-25% identity between sequences and the molecules are grouped together largely because they show a number of classical sequence patterns that typify the superfamily (Williams, 1987) although statistical tests are usually applied to support the superfamily argument. The following list includes well established Superfamily domains that are present in a variety of surface molecules. Particular emphasis is placed on the IgSF due to its relevance to work presented in Chapter 8.

a) Immunoglobulin (Ig)

Based initially on sequence similarity to domains within immunoglobulins themselves (Williams and Barclay, 1988), IgSF domains have now been identified in over 100 cell surface proteins including adhesion molecules (CD2, NCAM), signaling molecules (CD3 chains) and growth factor receptors (PDGF receptor). The Ig-fold consists of a sandwich of two beta-sheets, each consisting of anti-parrallel beta-strands of 5-10 amino acids usually with a conserved disulphide bond joining the two sheets (Amzel and Poljak, 1979; Williams and Barclay, 1988). The

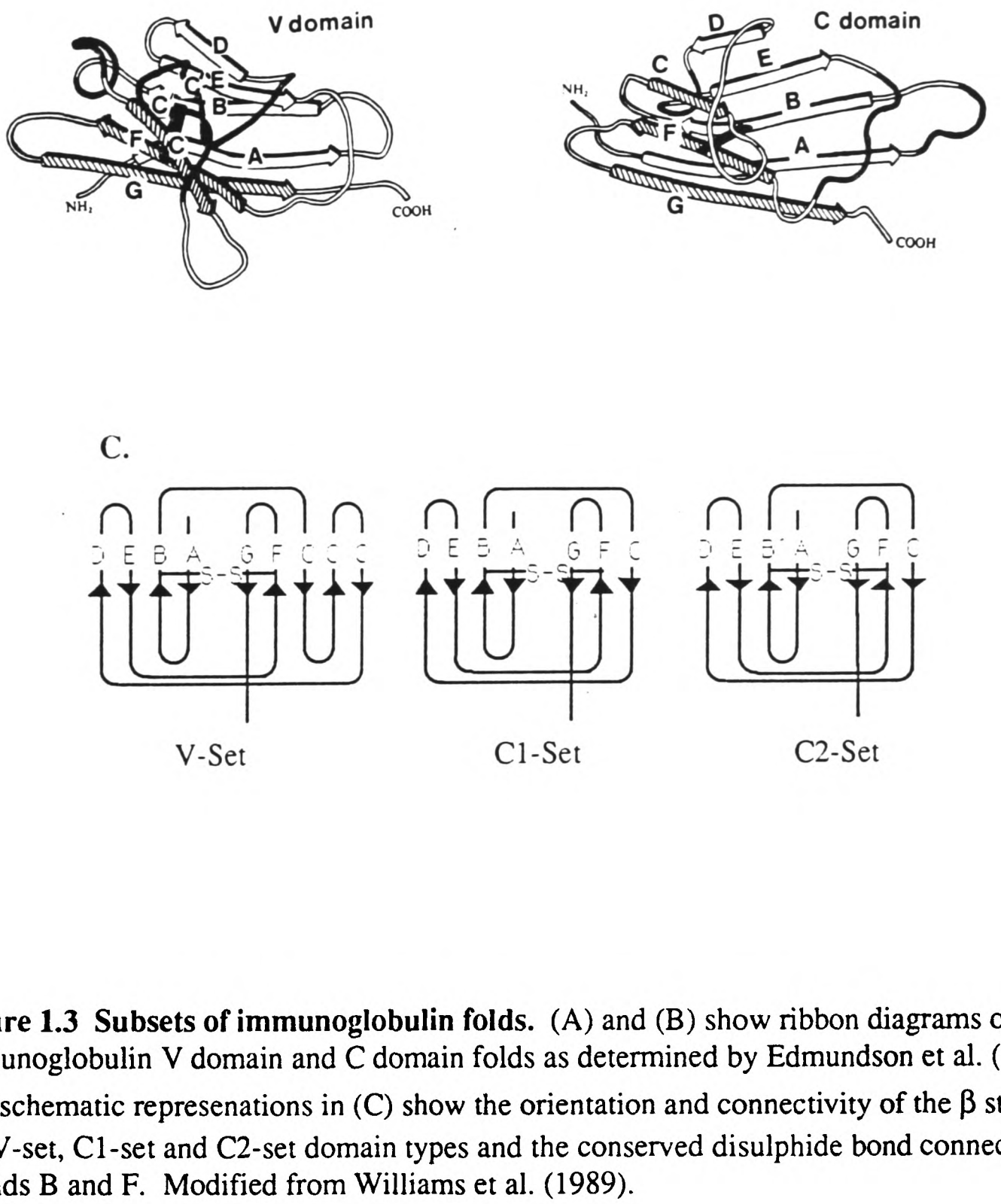


Figure 1.3 Subsets of immunoglobulin folds. (A) and (B) show ribbon diagrams of immunoglobulin V domain and C domain folds as determined by Edmundson et al. (1975). The schematic represenations in (C) show the orientation and connectivity of the β strands in the V-set, C1-set and C2-set domain types and the conserved disulphide bond connecting strands B and F. Modified from Williams et al. (1989).

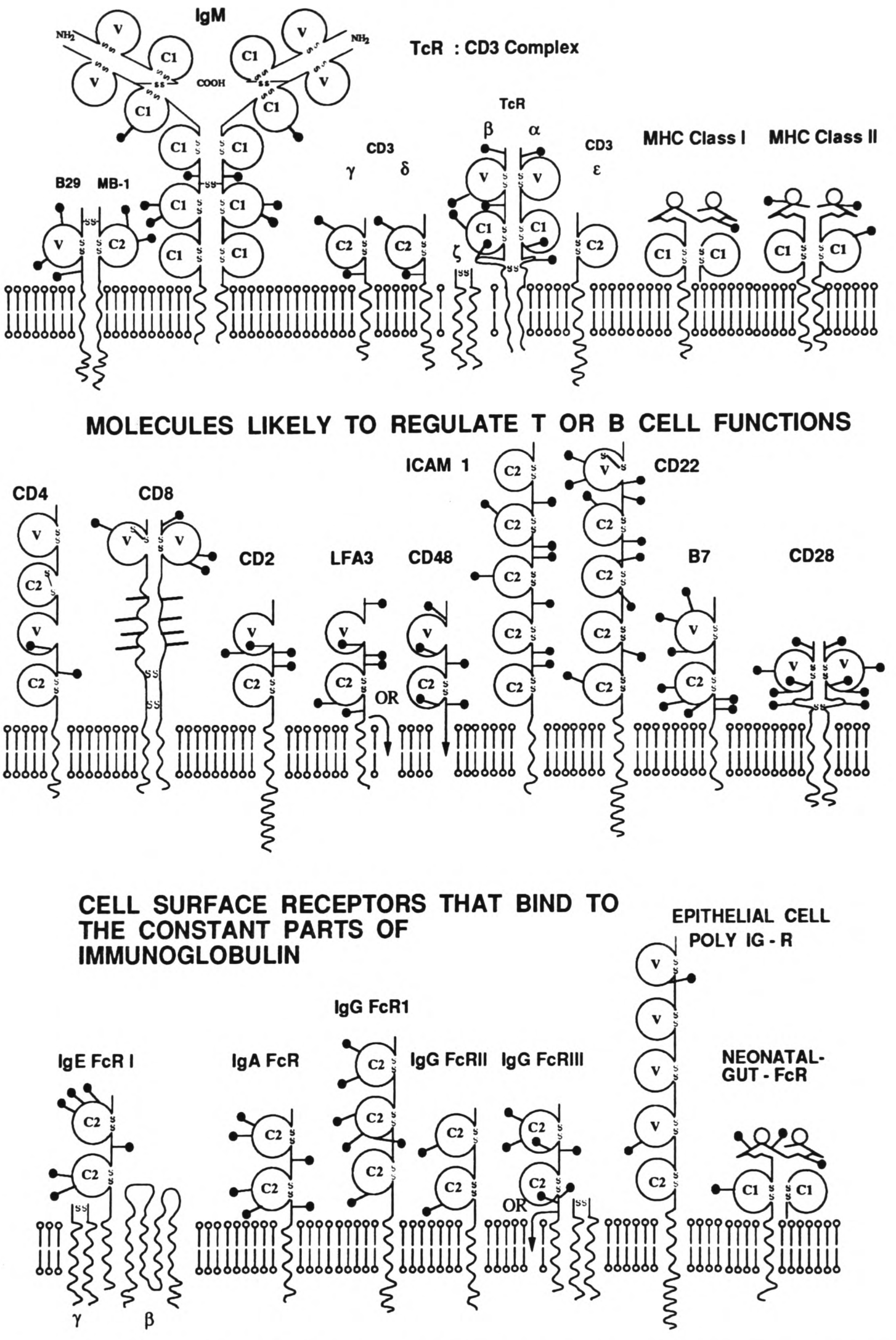


Figure 1.4 A subset of the immunoglobulin superfamily. Each circle corresponds to a predicted IgSF fold of the indicated set. N-linked glycans are shown as knobs and S-S indicates predicted disulphide bonds. This figure was kindly provided by Neil Barclay.

sequence similarities are found mainly in the beta strands at the ends of the domains with considerable differences in the loops between the strands and also in the length of sequence in the centre of the domain. These differences have lead to the definition of three subgroups of IgSF domains (Williams, 1987) termed V-set, C1-set and C2-set. The V-set domains, so called because of their similarity to Ig variable domains, typically have 65-75 amino acids betwen the conserved Cys positions whereas the C1- and C2- set folds, which are based on the Ig constant domains, have only 55-60 amino acids. The C1- and C2-set folds are made up of a 4 stranded (A, B, D and E) and a 3 stranded beta sheet (C, F and G, see Fig 1.3). The additional sequence in the V-set domains form two further beta strands (C' and C") with the same sheet as strand C. In some cases the A strand is hydrogen bonded to strand G in the CFG sheet and the D strand can be absent or replaced by a strand equivalent to C' (Ryu et al., 1990; Wang et al., 1990), but the location of the core strands, B, C, E and F is invariant (Williams, 1987). C2-set domains are distinguished from C1-set by the presence of sequence patterns more similar to the V-set fold, while lacking the extra sequence between the disulphide positions. The Ig fold has dimensions of 4 nm parallel to the β-strands and approximately 2.5 nm in the two perpendicular dimensions. Confirmation that proteins argued to be in the IgSF based on primary sequence patterns have similar three-dimensional folds comes from the structures of β2 microglobulin (Becker and Reeke, 1985), the MHC class I α3 domain (Bjorkman et al., 1987) and domains 1 and 2 of CD4 (Ryu et al., 1990; Wang et al., 1990). Some members of the IgSF family are depicted in Fig 1.4.

b) Fibronectin type III

First seen in fibronectin, this domain type is also found in a large number of surface molecules (Bazan, 1990), often in structures which also have IgSF domains. The three dimensional structures of two fibronectin type III domains have been determined (Baron et al., 1992; de Vos, 1992) and they reveal an Ig-type fold of two anti-parallel beta sheets forming a sandwich. The domain involves 100-130 amino acids and lacks a pattern of conserved cysteines. In the crystal structure of the growth hormone receptor- growth hormone complex the domain served a role both in contacting the growth hormone molecule and in self associating, providing contacts which supported receptor dimerization (de Vos et al., 1992).

c) Cytokine receptor

Found in cytokine and growth hormone receptors these domains contain about 100 amino acids and have a characteristic Cys X Trp sequence together with three other conserved Cys residues

(Bazan, 1990; Alan Williams, pers. comm.). The three dimensional structure of the cytokine receptor domain in the growth hormone receptor indicates again an Ig-type fold of anti-parallel beta sheets (de Vos et al., 1992). In the growth hormone receptor- growth hormone complex, the domain made many of the direct contacts necessary for binding the growth hormone molecule (de Vos et al., 1992).

d) Complement Control Protein (CCP)

Also called the short consensus repeat , this domain of 60-65 amino acids is found in many proteins of the complement cascade. It is also present in a wide variety of cell surface molecules alone or in combination with other types of Superfamily domain. The 3-dimensional structure of the 16th CCP domain from human complement factor H reveals a β-sandwich arrangement with one face being made up of a 3 stranded β-sheet and the other of two strands (Norman et al., 1991). The longest dimension of the domain from NMR data is 3.8 nm. Many of the cell surface proteins containing CCP domains are involved in interactions with complement components (Farries and Atkinson, 1991).

e) Epidermal Growth Factor

First defined in EGF itself, this sequence pattern is also present in a variety of surface proteins, sometimes in very large numbers such as the 36 domains in TAN-1, notch and xotch (Ellisen et al., 1991). The EGF domain is relatively small (40 amino acids) and is composed largely of a triple stranded beta sheet which is 2.3 nm in the longest dimension (Cooke et al., 1987; Campbell et al., 1989). A role in cell adhesion has been shown for EGF domains in notch and delta (Fehon et al., 1990).

Recent determination of structures for two fibronectin type III domains and a cytokine receptor domain have shown them to have an Ig-type fold (de Vos et al., 1992; Baron et al., 1992) as predicted (Bazan, 1990). This indicates that clearly different sequences can give rise to the same 3-dimensional fold and serves to strengthen arguments that members of a superfamily defined on the basis of primary sequence patterns have derived from a common ancestor (Williams and Barclay, 1988). It seems unlikely that convergent evolution to yield the same stucture would give rise to any common sequence pattern. The fact that sequence patterns are actually conserved at all may relate to the requirement for resitance to proteolysis of cell surface molecules. Surface molecules are generally resistant to proteolytic enzymes and this is based on

the folded structure since the denatured molecules are easily degraded. It has been argued (Williams, 1987) that mutation to give new recognition epitopes would have to be within the context of preserving resistance to proteolysis and that in general this lead to preservation of certain sequence patterns that determine one particularly stable solution to the fold. Numerous alternate patterns may exist that could also give the stable fold but to reach these a number of simultaneous mutations would be required and hence a switch to a new pattern may be a rare event in evolution (Williams, 1987).

Additional types of domain which recurr in more than one membrane protein include the scavenger receptor domain (Freeman et al., 1990) which is also seen in CD5, the cartilage link domain also seen in CD44 (Stamenkovic et al., 1989), the nerve growth factor receptor domain (Mallett and Barclay, 1991), the C-type and S-type lectin domains (Drickamer, 1988) and the cadherin domain (Mahoney et al., 1991). Proteins with cadherin domains have not been seen on leukocytes but they serve important calcium depedendent homotypic adhesion functions in many other cell types (McNeill et al., 1990; Takeichi, 1991). A further type of structural motif present in many surface proteins is that involving sequences rich in Ser, Thr and Pro and carrying large numbers of O-linked glycans. These regions are not described as domains because they do not have a tightly folded structure which necessitates conservation of sequence patterns. They will be described further in the section on O-linked glycosylation.

1.4.2 Multi-domain proteins

The majority of surface proteins are made up of more than one domain in the extracellular region, be they of the same type or of different types, and some very large molecules have been identified. For example, CD35 contains 28 CCP domains (Kalli et al., 1991), P-selectin contains 9 CCP domains, one EGF domain and one C-type lectin domain (Larsen et al., 1989), TAN-1 36 EGF domains (Ellisen et al., 1991) and the *fat* protein 34 cadherin domains and 4 EGF domains (Mahoney et al., 1991). Where a function for such surface molecules has been identified it is often restricted to the NH_2-terminal domain(s) (Hynes and Lander, 1992). This raises the question as to whether the other domains have evolved as spacer elements to project the functional domain from the membrane and associated glycocalyx, or whether in most cases all the domains have specific interaction functions. Some evidence to support the latter possibility comes from the observation that domain 3 of ICAM-1 is involved in a different interaction to domain 1 (Diamond et al., 1991), the EGF domain and at least two of the CCP domains in the selectins are important for high affinity binding of the terminal lectin domain

(Pigott et al., 1991; Watson et al., 1991) and CCP domains 8-11 and 15-18 of CD35 are involved in binding complement components C3b and C4b (Kalli et al., 1991). The domains nearest the membrane may be important in determining associations on the same cell, but it is still unclear why molecules such as CD4 and CD8, which appear to have analogous functions, should differ so markedly in the membrane proximal sequences.

1.5 Glycosylation

The majority of cell surface proteins carry (1-20) N-linked carbohydrates while a smaller number carry O-linked glycans (Kornfeld and Kornfeld, 1985) N-linked glycans have important roles in the folding and stability of a variety of proteins. Mutations at particular glycosylation sites often lead to a failure of expression at the cell surface or secretion, aggregation in the ER or rapid degradation (Elbein, 1991; Grinnell et al., 1991). Particular forms of N-linked glycans are important in directing proteins to specific compartments (Neufeld and Ashwell, 1980). For example, proteins destined for lysosomes are marked by attachment of Man-6-phosphate and this signal is recognized by a specific receptor, and glycoprotein hormones produced with the sequence SO_4-4GalNAcβ1-4GlcNAcβ1-2Manα are marked for rapid clearance by a receptor on hepatic reticuloendothelial cells (Fiete et al., 1991).

O-glycosylation is initiated post- rather than co- translationally and thus it is not likely to have a role in protein folding. However, the attachment of a large number of O-glycans in a short region of sequence contributes to maintaining chain stiffness in an otherwise random coil sequence (Shogren et al., 1989; Jentoft, 1990) and this is likely to be important with large mucin glycoproteins as it contributes to a large radius of gyration (the statistical average distance of all the elements of the chain from its centre). As a result the solution domains of these molecules begin to overlap when concentrations reach 1-2 mg/ml. At physiological concentrations (10-50 mg/ml) overlap becomes quite extensive leading to the formation of stable, intertangled networks (Jentoft, 1990; Strous and Dekker, 1992).

With smaller membrane molecules the glycosylation may also be important as a form of molecular extension and because of its resistance to proteolysis (Jentoft, 1990; Reddy et al., 1989; Kozarsky et al., 1988). The O-glycans themselves may be involved in interactions with cell-surface lectins and a well characterized example is the murine sperm-egg cell interaction. This involves an O-linked carbohydrate on the egg ZP3 protein interacting with a binding protein on the sperm (Florman and Wassarman, 1985). Stromal macrophages express a lectin

that recognizes the commonly O-linked structure, NeuNAc5α2-3Galβ1-3GalNAc, and this is important in conjugate formation with myelomonocytic cells (Crocker et al., 1991).

1.5.1 N-glycosylation

N-linked glycosylation is initiated during translation of the protein into the ER by transfer en block of the structure $Glc_3Man_9(GlcNAc)_2$ from a lipid-linked intermediate, to an Asn in the protein sequence. The recognition motif for the transfer enzyme is Asn-X-Ser/Thr-X (where X is any amino acid except Pro) although there are also structural constraints on recognition such that not all the sites present in the lumenal domain of a protein are necessarily glycosylated (Kornfeld and Kornfeld, 1985). Rapidly following addition of the sugar structure the glucose residues are trimmed off and subsequently varying numbers of mannose residues are removed. The structure is then modified in the medial and trans Golgi by addition of further sugars to generate any of a large array of structures which fall into three main categories termed high mannose, hybrid and complex. They all share the common core structure: Manα1-6(Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAc-Asn. A common mature form is the biantennary complex chain where the core region is lengthened by trisaccharide sequences composed of sialic acid-Gal-GlcNAc. Fucose is frequently attached to the Gal or GlcNAc in place of the sialic acid (Kornfeld and Kornfeld, 1985; Rademacher et al., 1988).

1.5.2 O-glycosylation

O-glycosylation of cell surface and secreted proteins is initiated by the attachment of GalNAc to the hydroxyl group of either Ser or Thr in the protein sequence. The motif recognized by the GalNAc transferase(s) is not known, although glycosylated regions are generally enriched for Pro, and there is some evidence that the sequence XPXX (where X is any amino acid but at least one is Thr) is a minimal recognition motif for some, but definitely not all, glycosylation events (Gooley et al., 1991; O'Connell et al., 1991a). It may be the case that when the region is very rich in Ser and Thr (>30%) the requirement for Pro residues is reduced (Table 1.1). The precise point at which attachment of GalNAc takes place is not well defined, but recent data with leukosialin indicate a very rapid addition in a cis Golgi compartment (Piller et al., 1989a; Piller et al., 1990). The nature of the sugars subsequently attached depends largely on the glycosyltransferases expressed within the cell but all the O-linked sugars formed have three well-defined regions, a core, a backbone region containing Gal and GlcNac residues, and non-reducing termini usually derived from α-linked sialyl, Fuc, Gal and GalNAc residues.

A. Core structures of O-linked glycans

Core class 1: Galβ1→3GalNAc-Ser/Thr

Core class 2: GlcNAcβ1↘6 / Galβ1→3GalNAc-Ser/Thr

Core class 3: GlcNAcβ1→3GalNAc-Ser/Thr

Core class 4: GlcNAcβ1↘6 / GlcNAcβ1→3GalNAc-Ser/Thr

Core class 5: GalNAcα1→3GalNAc-Ser/Thr

B. Examples of terminal structures

Fucα1↘4 / Galβ1→3GlcNAc- Blood group Lewis^a (Le^a)

Fucα1↘4 / Fucα1→2Galβ1→3GlcNAc- Blood group Lewis^b (Le^b)

Galβ1↘4 / Fucα1→3GlcNAc Lewis^x (Le^x)

NeuNAcα2→3Galβ1↘4 / Fucα1→3GlcNAc Sialyl-Lewis^x (sLe^x)

Polylactosaminoglycan

Galβ1→4GlcNAcβ1→3Galβ1→4GlcNAcβ1→3-R Blood group i

Figure 1.5 The five main classes of core structures defined in O-linked oligosaccharides and some antigenic structures associated with O-linked glycans. Modified from Schachter and Brockhausen, 1989.

The O-glycans can be divided into groups based on 5 core structures, defined by the sugars directly attached to the initial GalNac residue (Fig 1.5) (Paulson, 1989; Paulson and Colley, 1989; Schachter and Brockhausen, 1989). Leukosialin has only been shown to carry core class 1 and 2 and thus only these will be further considered here. Synthesis of core 1 involves the core 1 β1-3Gal transferase which transfers Gal from UDP-Gal to the 3 position of the GalNAc. This core structure may then be used as a substrate for α2-3 and α2-6 sialyltransferases, to form mono- or di-sialylated forms of O-glycan. Alternatively core 1 can serve as a substrate for the core 2 β1-6GlcNAc transferase to form core 2 (Fig 1.5). This enzyme requires the GalNAc to have the Gal attached in the 3 position for it to be a substrate; it is totally inactive on GalNAc-Ser/Thr. If sialic acid is already attached in the 6 position of GalNAc in the core 1 structure then core 2 cannot be formed, so the relative levels of core 2 β1-6-GlcNAc transferase and α2-6 sialyltransferase have a major role in determining which of these sugar structures is formed. This is particularly significant because the core 2 structure is a substrate for a number of "extending enzymes". Elongation generally involves addition of Gal and GlcNAc residues in β1-3 , β1-4 and β1-6 linkages and extensions can be 40 residues long. The poly-N-acetyllactosamine structure (Galβ1-4GlcNAcβ1-3-)n often serves in O-glycans as an intermediary sequence between core and non-reducing terminus. GlcNAc residues can be inserted in β1-6 linkage to a Gal residue in this repeating sequence to create a branched structure. In general the structures have non-reducing terminal α-linked sugars, the most common being sialic acid (as already discussed for core 1), and L-fucose (Schachter and Brockhausen, 1989). The O-glycosylation of leukosialin has been studied in some detail and is described in the section on leukosialin.

1.5.3 Types of O-glycosylated cell surface molecules

Cell surface proteins carrying O-linked oligosaccharides can be divided into 3 groups based on the location of the O-glycosylated region in the molecule.

1) O-glycosylated membrane proximal region. A large number of surface proteins, as listed in Table 1.1 for molecules on cells of hematopoietic origin, have a region near the membrane domain (or point of GPI-anchor attachment) of 15-250 amino acids which is rich in Ser, Thr and Pro and which in many cases has been shown to carry O-linked glycans. For most of these proteins the membrane distal domain has a demonstrated ligand binding activity, either with a soluble ligand (eg, the IL2-R) or a molecule on another cell (eg, CD8). The lengths of the glycosylated regions in CD46 and CD44 can be modified by alternative splicing as described in more detail below for CD44. With this group the possibility that the glycosylated region serves

solely as a protease resistant hydrophilic spacer, a "stalk" to project the globular NH_2-terminal domain above the membrane and associated glycocalyx, seems reasonable. With the following groups a more direct functional role for the O-glycosylated region must be envisaged.

2) O-glycosylated region encompassing the whole extent of the extracellular domain with no clear regional specialization. Glycophorin was the first example of this type of glycoprotein (Tomita et al., 1978) although it is not as heavily glycosylated as other members of this family (Table 1.1), in particular leukosialin. These molecules are often related to the mucin proteins expressed on epithelial and adenocarcinoma cells, although the latter molecules are generally substantially larger (as described further below) than those found naturally on circulating cells. In addition a set of highly expressed GPI-linked surface proteins of only 12-34 amino acids in length is included in this group because they have no evidence for regional specialization and are all likely to be heavily glycosylated (Table 1.1).

3) O-glycosylated membrane distal region. The distinction of this type of glycosylation from that in (1) is that it can not have a role in projecting a folded domain from the cell surface. One member of this group, CD45, is amongst the most highly expressed glycoproteins on T and B cells and is described in more detail below. While proteins of types (2) and (3) may have an important role in contributing to cell surface charge this cannot be the full explanation for the variety of molecules seen, the differential expression on different cells, and the co-expression of different members on the same cell. Rather the glycosylated domains are almost certain to have direct roles in interacting with other proteins which may include masking of specific structures on the same cell membrane and adhering to structures on other cells or to soluble factors.

1.5.4 Two molecules with alternatively spliced exons that encode O-glycosylated regions

CD45 is expressed on all cells of hematopoietic origin except red cells and like leukosialin, it is expressed at a high level, there being $7x10^4$ molecules on rat thymocytes. The extracellular domain varies in length from 391 to 552 amino acids in human CD45, depending on whether any of exons 4(A), 5(B), or 6(C) are present. The NH_2-terminal 40 amino acids of the form lacking all 3 exons (CD45R0) contain 45% Ser+Thr and has been shown to contain O-glycans in the rat (Barclay et al., 1987). The following 350 amino acids include 16 Cys residues and are folded into a rod like structure (Woollett et al., 1985). The alternate forms of CD45 include one, two or all three of exons A, B or C. Exon A is 74 amino acids long and 34% Ser+Thr, exon B 47 amino acids and 32% and exon C 48 amino acids and 40%, and it has been shown

Table 1.1 Glycoproteins on cells of hematopoietic origin with O-glycosylated regions

Protein	Length	% S+T	% Pro		Comments	Reference
Type 1: O-glycosylated membrane proximal region						
CD6	30	31	7	+	Sulphated	Swack et al., 1991
CD7	36	17	34		4 repeats 9aa	Aruffo & Seed, 1987
CD8α	50	16	20	+	5 glycans, 2 Cys	Classon et al., 1992
CD8β	40	20	15		2 Cys, dimer with α	Norment & Littman, 1988
CD27	70	21	14	++	1 Cys, homodimer	Camerini et al., 1991
CD30	35	38	14			Durkop et al., 1992
CD44	100-288	35	9	++	Splice, CS	Jackson et al., 1992a
CD46	29-44	35	23	+	Splice	Post et al., 1991
GP1Bα	155	37	20	++	35nm, repeats	Lopez et al., 1987
DAF	70	41	11	++	GPI	Medof et al., 1987
TNFR-II	57	33	21		TNFα receptor II	Smith et al., 1990
OX40	46	33	15	+	rat	Mallett et al., 1990
IL-2Rα	25	36	0	+		Leonard et al., 1984
CMRF35	55	33	18		CD8 like	Jackson et al.,1992b
Type 2: O-glycosylated region encompassing whole extracellular domain						
Glycophorin	72	33	5	++	15 glycans	Tomita et al., 1978
CD43	230	40	12	++	~80 glycans	Pallant et al., 1989
CD24	33	48	6	+	GPI	Kay et al., 1991
HSA	27	26	15		GPI, mouse	Kay et al., 1990
CDw52	12	50	8		GPI	Xia et al., 1991
B7(2)	26	38	4		GPI, mouse	Kuboto et al., 1990
Type 3: O-glycosylated membrane distal region						
CD34	140	41	7	++	50kd O-glycans	Simmons et al., 1991
CD45	40-209	40	11	++	Splice variants	Thomas, 1989
CD68	145	46	12	++	Macrosialin	Holness et al.,1992
114/A10	216	55	12		8 repeats, 27aa	Dougherty et al., 1989

Molecules listed are human unless indicated. Length represents the number of amino acids in the extracellular domain which encompasses the region likely to carry the glycans. The percentage of Ser+Thr and Pro in this region is shown. Where O-glycans are known to be present they are indicated + (small number) or ++ (>10 glycans). CS indicates chondroitin sulphate also present. With the forms that have splice variants the average content of Ser+Thr and Pro for all the exons is shown.

that these exons do carry large amounts of O-linked oligosaccharide (Jackson and Barclay, 1989). Exons A and C also contain potential N-glycosylation sites. As typical of heavily O-glycosylated regions, the exons are not highly conserved between species, there being about 40% identity between rodent and man (Thomas, 1989). The most interesting feature of the alternate exon usage by CD45 is its correlation with different functional cell subsets. B cells express a form that contain all three exons (B220), naive T cells forms that have combinations of 1 or 2 exons and activated T cells and most memory T cells express the form lacking all three exons. These differences, which appear to be conserved across species, imply important functional roles for the heavily glycosylated exons. The observations that different forms of CD45 expressed on the same cell preferentially associate with different molecules in the membrane (Dianzani et al., 1990; Dianzani et al., 1992) argues that some exon specific interactions are occuring to support these associations. Most recently the form of CD45 lacking all three exons has been shown, by antibody blocking studies, to be a ligand for CD22 and forms with inserted exons did not bind (Stamenkovic et al., 1991b). Thus one role for the exons may be to disrupt this adhesion interaction but why three exons should be used differentially remains unanswered.

CD44 (Pgp-1) has an NH_2-terminal region with similarities to the cartilage link protein and it binds hyaluronic acid (Aruffo et al., 1990). This interaction may be important in mediating lymphocyte binding to HEV and for migration of the cells through the tissues. The segment following this, approximately 100 amino acids long (150-248, 35% Ser+Thr) in the shortest form, is O-glycosylated and contains sites for chondroitin sulphate attachment. A minor high molecular weight form (180-200kd) present on lymphocytes carries chondroitin sulphate. Further variation results because a number (at least 5) of different exons can be inserted after Thr202 (human numbering), within the glycosylated region, and most of these also appear good substrates for GalNAc transferase (28-43% Ser+Thr) (Gunthert et al., 1991; Stamenkovic et al., 1991a; Jackson et al., 1992b). An epithelial cell form which has a 135 amino acid insert at Thr 202 is unable to support binding to hyaluronate (Stamenkovic et al., 1991a). Earlier work showed that mAb binding to the glycoyslated region reduces hyaluronate binding (Goldstein et al., 1989) and these results indicate a direct influence of the region on the link protein domain. Gunther et al., (1991) observed that transfection of carcinoma cells with a form of CD44 containing a 162 amino acid insert conferred metastatic potential on the cells whereas the form without the insert did not. Treating the cells with a mAb specific for the insert prevented metastasis. Thus it would appear that this region of CD44 is acting as more than just a "stalk"

to project the functional domain above the glycocalyx.

1.6 Mucins

Mucus secretions form a viscous gel covering epithelial surfaces and are made up of large, heavily O-glycosylated proteins (2%), salt (1%) and water (95%) and small amounts of lipids and cell debris (Harding, 1989). The mucus acts to lubricate epithelial surfaces and to protect the epithelial cells from the lumenal contents. Mucin glycoproteins have been subdivided into two classes, those produced by specialized epithelial cells that make up a gelatinous layer which also covers neighbouring cells, and those produced by non-specialized cells and by many carcinomas which do not form a gel (Hilkens, 1988; Strous and Dekker, 1992). Both have large protein cores (generally >200 kd) rich in Ser and/or Thr, Pro, Ala, and Gly, low in aromatic and sulfur containing amino acids, and both are heavily O-glycosylated (50-80 % by weight) with few or no N-glycans. The gel forming mucins are linked together through disulphide bonds to form long polymers and the presence of the bonds is important for gel formation. The non-gel forming epithelial mucins are expressed on the apical cell membrane, and many are likely to be anchored via a transmembrane domain (Hilkens, 1988; Ligtenberg et al., 1990; Wreschner et al., 1990). These mucins are not linked by disulphide bonds. The O-glycans on both types of mucins can vary markedly between tissues and also within tissues although the majority carry the core 1 or core 2 structures as described above, with sialic acid, galactose and fucose as the most common terminal residues. Mucin carbohydrates are frequently altered during differentiation and carcinogenesis, with most changes occuring in the peripheral region of the glycans, frequently reverting them to a fetal state (Hilkens, 1988). A variety of mAb have been raised that detect epitopes present in mucins on cancer cells and some of these are used in differential diagnosis.

Most of the mucin core proteins examined contain large numbers of Ser and/or Thr and Pro rich repeats (Table 1.2). Frequently the number of repeats is found to vary between alleles, and many alleles can be present in the population. The MUC-1 mucin, for example, which is expressed on mammary gland and pancreatic epithelium, can have between 40-90 tandem repeats of a 20 amino acid segement (Gendler et al., 1990; Hareuveni et al., 1990; Lan et al., 1990; Ligtenberg et al., 1990). A number of splice variants of this mucin also exist and in particular transmembrane and soluble forms are predicted (Wreschner et al., 1990).

High level expression (>10^6 molecules/cell) of large mucins of this type on carcinoma cell lines

Table 1.2 Repeat elements detected in mucin glycoproteins

Mucin	Species	Repeat consensus sequence	No.	Tissue	Ref.
MUC-1[a]	Human	PGSTAPPAHGVTSAPDTRPA	40-90	Breast & Pancreas	1
MUC-2	Human	PTTTPITTTTTVTPTPTPTGTQT	?>14	Intestine & Trachea	2 3
MUC-3	Human	HSTPSFTSSITTTETTS	?>7	Intestine	4
MUC-4	Human	$TSS^{V}/ASTGH^{T}/AT^{S}/PLVTD$	?>39	Colon & Tracheo-bronchus	5
FIM-A.1	Frog	VPTTPETTT	~15	Skin	6
FIM-B.1	Frog	GESTPAPSETT	?>6	Skin	7
PSGP	Fish	AATGPSGDDATSE	25	Eggs	8
PSM	Pig	GAGPGTTASSVGVTETARPSV-AGSGTTGTVSGASGSTGSSSG-SPGATGASIGQPETSRIRVAGS-SGPAPAVSSGASQAAGTS(81)	~25	Sub-maxillary gland	9

[a] MUC-1 has also been called episialin, DF3, PEM and PUM. No. indicates number of repeats where ?> means that the number of repeats estimated to be present from RNA and protein analysis is greater than the number present in the cDNA clones isolated.

References
1. Ligtenberg et al., 1990
2. Gum et al., 1989
3. Gerard et al., 1990
4. Gum et al., 1990
5. Porchet et al., 1991
6. Hoffman, 1988
7. Probst et al., 1990
8. Kitajima et al., 1986
9. Eckhardt et al., 1991

correlates with allo- and xeno- transplantability and this is argued to result from masking of MHC molecules by the mucin (Codington and Frim, 1983; Carraway and Spielman, 1986). Masking of adhesion molecules may also be important.

1.7 Integrins

Integrins are a large family of heterodimeric molecules composed of transmembrane α and β chains. More than 8 β chains and 13 α chains have been defined and many different combinations of the chains exist (Humphries, 1990; Hynes and Lander, 1992). The integrins are usually classified into β chain families because each β chain generally associates with multiple α chains. Leukocytes express a variety of forms but in particular the β2 (CD18) family of integrins, LFA-1 (CD11a/CD18), Mac-1 (CD11b/CD18) and gp150,95 (CD11c/CD18), are restricted to cells of the immune system (Dustin and Springer, 1991). The structure of the chains has not been determined but the presence of three or four cation binding sites in the α chain has been observed from the sequence and cations are necessary for ligand binding (Dustin and Springer, 1991; Dransfield et al., 1992). The dimensions of the integrin VLA-5 have been estimated by electron microscopy as 20 x 15 nm allowing potential binding of ligands at a distance of 20 nm from the lipid bilayer (Nermut et al., 1988).

It has been shown for a number of integrins, and may be true for all, that the extracellular domain can exist in a low and high avidity state for ligand. Switching from the low to high avidity state can be triggered by stimuli that activate the cell, such as phorbol esters (Danilov and Juliano, 1989; Dustin and Springer, 1989), by mAb or Fab binding to the extracellular domain and fixing the molecule in the high affinity state (Dransfield and Hogg, 1989; Gulino et al., 1990; O'Toole et al., 1990; van Kooyk et al., 1991) and by ligand binding itself (Du et al., 1991; Phillips et al., 1991). The mechanism by which signaling into the cell can lead to an altered avidity of the integrin molecules on the outside of the cell is not understood, although certain regions of the cytoplasmic domain are necessary (Hibbs et al., 1991). A novel lipid factor may also have a role (Hermanowski-Vosatka et al., 1992).

The majority of integrins have been shown to bind one or more extracellular matrix molecules including fibronectin, fibrinogen, vitronectin, collagen and laminin (Humphries, 1990) and the migration of lymphocytes through the tissue matrix depends at least in part on the interaction of integrins with matrix components (Shimizu et al., 1990b). In addition, binding of some extracellular matrix components to lymphocytes provides a costimulatory signal for activation

(Dang et al., 1990; Shimizu et al., 1990a). The integrins LFA-1 and Mac-1 of the β2 family, and VLA-4 of the β1 family, also act by binding to other cell surface molecules.

1.7.2 LFA-1 and its ligands ICAM-1, ICAM-2 and ICAM-3

LFA-1 is expressed on lymphoid and myeloid cells whereas Mac-1 and gp150,95 are restricted to cells of myeloid origin. The importance of LFA-1 (CD11a/CD18) on T cells in cell-cell interactions was revealed by the ability of anti-LFA-1 mAb to block a variety of T cell mediated events, including target cell killing (Springer, 1990). Subsequently 3 cell surface glycoproteins (termed ICAMs) have been defined which are ligands for LFA-1. ICAM-1 is a member of the IgSF, having five Ig domains in the extracellular region which give the domain a total length of 18 nm with a point of flexibility between domains 2 and 3 or 3 and 4 (Staunton et al., 1990). ICAM-1 is expressed at low levels on resting endothelium, B cells and dendritic cells but expression is increased following activation of these cells for example by IL-1, TNFα or IFNγ. LFA-1 binds ICAM-1 through the first domain of ICAM-1(Staunton et al., 1990). The third domain is a binding site for Mac-1 (Diamond et al., 1991). ICAM-2 is composed of only two IgSF domains which are similar to the NH_2-terminal domains of ICAM-1 (Staunton et al., 1989a). It is expressed on endothelium at higher levels than ICAM-1 and expression is not increased by activation. The sequence of ICAM-3 has not been reported but it is a heavily glycosylated molecule and is not expressed on endothelial cells. Rather it is present on resting T and B cells and may be important for LFA-1 mediated T-B cell interactions (de Fougerolles and Springer, 1992). The avidity of LFA-1 for the ICAMs is increased when cells are activated by a variety of stimuli, in particular by triggering through the TcR/CD3 complex or CD2 on T cells (Dustin and Springer, 1989; van Kooyk et al., 1989), and the Ig receptor complex on B cells (Dang and Rock, 1991), and when neutrophils and monocytes are activated by such factors as IL-8 and fMLP (Altieri and Edgington, 1988; Huber et al., 1991; Lawrence and Springer, 1991). Mac-1 avidity on neutrophils can be increased by binding to E-selectin (Kuijpers et al., 1991; Lo et al., 1991). The binding of β2 integrins to their ligands may also transmit costimulatory signals to the lymphocyte (Mishra et al., 1986; van Noesel et al., 1988) and it appears that related integrins can transmit different signals (Berton et al., 1992), a result not unexpected given the lack of sequence similarity in the α chain cytoplasmic domains. The increased LFA-1 avidity induced by stimuli that activate the cell is spontaneously downregulated after a period of minutes and this may be important in allowing de-adhesion of leukocytes from target cells, antigen presenting cells, or endothelial cells (Dustin and Springer, 1989).

1.7.3 VLA-4 and VCAM-1

VCAM-1, another member of the IgSF family with 6 or 7 IgSF domains (depending on alternative splicing events) (Hession et al., 1991), is expressed on cytokine activated endothelium and is a ligand for the VLA-4 (CD49d/CD29) β1 family integrin and the interaction of these molecules is important in lymphocyte adherence to endothelial cells (Yednock et al., 1992). Expression of VCAM-1 has also been seen on follicular dendritic cells and binding of B cells can involve the LFA-1/ICAM-1 and VLA-4/VCAM-1 pairs (Koopman et al., 1991). As with the β2 family integrins, VCAM-1 binding to VLA-4 can promote CD3-dependent T cell proliferation (Burkly et al., 1991).

1.8 Selectins and their ligands

The selectins are a group of cell surface molecules which have extracellular domains composed of a C-type lectin domain, a single EGF domain, and a number of CCP domains. P-selectin (GMP-140, CD62) is expressed in secretory granules within platelets and endothelial cells and is rapidly mobilized to the cell surface following stimulation, in particular by products of the clotting cascade such as thrombin (Larsen et al., 1989; Geng et al., 1990). E-selectin (ELAM-1) is expressed on the surface of endothelial cells following activation by a variety of inflammatory mediators (Bevilacqua et al., 1989; Hession et al., 1990) whereas L-selectin (LECAM-1, LAM-1, Mel14, Ly22) is expressed on the surface of lymphocytes and neutrophils (Lasky et al., 1989; Siegelman et al., 1989). The lectin domains of the P- and E-selectins are specific for the sugar structure sialyl-Lewis^x (sLe^x, Fig 1.5), having only a low affinity for the non-sialylated form (Le^x or CD15) (Lowe et al., 1991; Polley et al., 1991; Tiemeyer et al., 1991; Zhou et al., 1991). Le^x and sLe^x are present on a variety of molecules on the neutrophil surface including lactosylceramides (Tiemeyer et al., 1991), L-selectin (Picker et al., 1991b), Mac-1, and NCA 160 (Stocks et al., 1990). There is also evidence that sLe^x and Le^x determinants on human neutrophils can be O-linked, possibly carried by leukosialin (Fukuda et al., 1986; Leeuwenberg et al., 1991). In addition to sLe^x, P-selectin is able to bind sulfatides, heterogenous 3-sulfated galtactosyl ceramides, which are present on the surface of granulocytes (Aruffo et al., 1991). Binding of E-selectin to a sialylated form of Le^x not seen by anti-sLe^x antibodies has been reported (Picker et al., 1991a). This is of added significance because sLe^x has not been detected on human lymphocytes whereas the modified form of Le^x appears to be present on a subpopulation of skin homing, memory T cells (Berg et al., 1991; Shimizu et al., 1991). The nature of the L-selectin ligand is not fully defined but binding is neuraminidase

sensitve and can be inhibited by sulphate or phosphate containing carbohydrate ligands and the target structure may be present on proteins of 50 and 90 kd that are restricted to peripheral lymph node HEV (Imai et al., 1990; True et al., 1990; Watson et al., 1990).

A variety of other carbohydrate-lectin interactions operate in the immune system (Sharon and Lis, 1989). For example, macrophages express a Gal/GalNAc specific lectin that can act in uptake and destruction of tumour cells (Sharon and Lis, 1989); IgD receptors on T cells may act partially by binding to carbohydrate (Amin et al., 1991); and stromal macrophages may bind to myelomonocytic cells through a sialic acid dependent lectin (Crocker et al., 1991).

1.9 Signal transduction and cytoplasmic domains

1.9.1 Signal transduction and single transmembrane regions

The mechanism of signal transduction through molecules that have only a single transmembrane domain is poorly understood. Generally it is assumed that signals cannot be transmitted through a single transmembrane sequence, and that signaling comes about by an induced association of, or change in an existing association of, two or more surface glycoproteins. The result of this change is to bring two or more cytoplasmic domains into close proximity and the interactions between them culminate in signal transmission. The strongest evidence for this scenario has come from the transmembrane tyrosine kinases, where dimerization of the extracellular domains is induced by binding of soluble ligand, leading to cross-phosphorylation of the cytoplasmic domains and an activation of their kinase activity (Williams, 1989; Ullrich and Schlessinger, 1990). The molecular associations induced when two cells come together are likely to be more complex with the local concentration and relative orientation of a large number of cytoplasmic domains being important in determining the nature of the signal transmitted.

1.9.2 Cytoplasmic domains

Cytoplasmic domains of transmembrane proteins have roles in cell signaling, intracellular routing, compartmentalization, cytoskeletal association, endocytosis and recycling.
In contrast with the majority of extracellular domains which can be classified into superfamilies on the basis of sequence similarities, the majority of cytoplasmic domains have not been found to show large regions of sequence similarity to each other. This is despite the fact that the cytoplasmic domains of species homologs are often conserved at the level of 70-90 %. A small number which have clearly defined functional activites, the kinases and phosphatases, do show sequence patterns within families over large areas of the cytoplasmic domain. Recent studies

have demonstrated the existence of motifs in some other types of molecules involved in signaling events (Reth, 1989; Turner et al., 1990; Ellisen et al., 1991) and in molecules that interact with clathrin coated pits (Bansal and Giersach, 1991; Eberie et al., 1991). It is likely that more motifs will be discovered as the interactions and structures of further cytoplasmic domains are studied.

Phosphorylation of cytoplasmic domains on Ser, Thr and/or Tyr residues is a very common form of reversible modification. In the domains with kinase or phosphatase activity, phosphorylation can alter enzymatic activity (Williams, 1989; Ullrich and Schlessinger, 1990). Phosphorylated Tyr residues can also act as targets for binding of proteins containing SH2 (src homology region 2) domains (Anderson et al., 1990; Moran et al., 1990) including a PI-3 Kinase associated protein (Escobedo et al., 1991), GTPase activating protein and PLC-γ1 (Anderson et al., 1990). The exact specificities of the various SH2 domains has not been defined but the ability to bind Tyr-phosphate is essential in the formation of multiprotein complexes which follow phosphorylation of such proteins as the EGF-R and PDGF-R (Escobedo et al., 1991) and probably also the CD3 complex. The formation of these complexes is likely in turn to be essential for signal transmission. A similar binding domain for Ser- or Thr-phosphate has not been identified but roles for such phosphorylation in down-regulation from the cell surface and intracellular routing have been demonstrated (Cantrell et al., 1988; Casanova et al., 1990; Lund et al., 1990; Beyers et al., 1991). Leukosialin is phosphorylated in its cytoplasmic domain by protein kinase C, and this enzyme will be described further below.

Evidence for the association of a large number of transmembrane proteins with the cytoskeleton of the cell has been reported but only in a few cases have associations of purified molecules been demonstrated. Erythrocyte band 3 (the anion transporter) binds to ankyrin (Nelson and Veshnock, 1987) which in turn binds to the cytoskeletal molecules spectrin and vimentin (Lux et al., 1990). This interaction protects red cells from circulatory shear stresses and helps maintain their biconcave shape. A second interaction in the red cell is that between glycophorin and protein 4.1 (Anderson and Lovrien, 1984). The β chain of certain integrins may interact with talin or α-actinin (Horwitz et al., 1986; Otey et al., 1990) and co-localization of these molecules at the cell surface with integrins has been frequently reported (Burridge et al., 1988). High affinity interactions occur between molecules which are endocytosed via clathrin coated pits and adaptor proteins, associated with the coated pits, and are important for endocytosis (Pearse, 1988; Beltzer and Spiess, 1991). The GPIBα-IX complex of platelets, which acts as a

von Willebrand factor receptor, associates with an actin binding protein and this is suggested to be important in holding platelets, bound through GPIBα-IX to the endothelium, against the shear stress of the blood flow (Andrews and Fox, 1991). More recently the IgG FcγRI receptor was shown to be associated with an actin binding protein in U937 cells and following IgG binding the receptor showed a rapid decrease in affinity for the binding protein (Ohta et al., 1991). Perhaps the most successful series of studies looking at adhesion protein-cytoskeletal protein interactions have been with uvomorulin, a member of the cadherin class of adhesion proteins (Takeichi, 1991). It associates with three proteins through its cytoplasmic domain and the formation of this complex links uvomorulin to the actin filament network which is important for the adhesion function of the extracellular domain (Nagafuchi and Takeichi, 1988; Jaffe et al., 1990). The associating proteins have been purified and the sequence of one of them, catenin α, shown to be homologous to vinculin, a previously characterized cytoskeletal protein (Herrenknecht et al., 1991; Nagafuchi et al., 1991). Thus cytoskeletal interactions can be important in determing cell-shape, in regulating modulation from the membrane and in regulating cell-surface adhesion events. Such interactions may also be important in determining diffusion rates in the membrane, localization of the molecule in particular regions of the membrane and in regulating cell motility.

1.9.3 Protein Kinase C (PKC)

Protein kinase C mediated phosphorylation of Ser or Thr residues is important in many activation events of leukocytes. For example, signaling through the TcR/CD3 complex, CD2, surface Ig, and the IL-1 receptor involve PKC mediated phosphorylation of many substrates and a number of the events triggered through these receptors can be mirrored by direct activation of PKC with phorbol esters. Nine isoforms of the enzyme have so far been identified, differing in tissue distribution and in requirements for allosteric activators (Nishizuka, 1988; Pears and Parker, 1991). In particular the α, β, and γ isoenzymes all contain a regulatory calcium binding region whereas the δ, ε, ζ and η forms lack this region and do not require calcium for acitivity. All forms of PKC (except possibly ζ) share in common the property of binding and being activated by diacylglycerol, and this acts as a link between external stimuli which lead to production of DAG as a second messenger, and cell activation events (Bell and Burns, 1991). Phosphatidyl serine (PS), a lipid present naturally at high levels in the cell membrane, is also an important cofactor for enzyme activity and the sequence of events leading to PKC activation are thought to be as follows (Bell and Burns, 1991; Pears and Parker, 1991). Stimulation by external stimuli leads to hydrolysis of phosphoinositides and other phospholipids by

phospholipase C or D with subsequent elevelation in levels of DAG in the inner leaflet of the cell membrane. The regulatory subunit of PKC interacts cooperatively with the DAG and PS in such a manner that the conformation of the enzyme is changed, leading to relocation of the pseudosubstrate sequence and exposure of the active site. Phorbol esters, such as PMA, substitute for DAG in these events with the difference that they are less rapidly degraded by the cell and thus result in more prolonged PKC activation. Isoforms with a calcium binding site in the regulatory subunit are also dependent, to varying extents, on elevations in intracellular calcium for full activity. T cells and neutrophils express predominantly the α and β isoforms of PKC although the relative levels of these isoforms can vary with the activation state of the cells (Alexander and Cantrell, 1989; Pontremoli et al., 1990). This is significant because the β isozyme is less dependent on Ca^{2+} than the α species and shows substantial activity in the absence of Ca^{2+}. It is unclear whether these forms differ in substrate specificity, and the identity of their pseudo-substrate sequences suggest they may not (Kemp and Pearson, 1990; Gschwendt et al., 1991).

As stated previously, a large number of molecules are phosphorylated by PKC including cell membrane and cytosolic proteins. Cytoplasmic domains that are targeted include those of CD3 γ and ε, CD4, CD5, CD8, CD45, MHC class 1, the α chain of the IL-2 receptor, the β chain of integrins, and leukosialin (Chatila and Geha, 1988; Alexander and Cantrell, 1989). In most cases the immediate effect of phosphorylation is unkown although with CD3 and CD4 there is a correlation between phosphorylation and down-modulation from the cell surface (Alexander and Cantrell, 1989). LFA-1 cytoskeletal interactions may also be regulated by PKC (Pardi et al., 1992) and this could explain the involvement of PKC in modulating integrin avidity seen in some systems (Hibbs et al., 1991).

1.10 CD2, LFA-3 and related molecules

1.10.1 The CD2 molecule

The sheep erythrocyte rosetting receptor on T cells, now called CD2, was the first cell surface protein shown to be involved in heterotypic adhesion. This followed from the observation that a fraction of human peripheral blood lymphocytes formed spontaneous rosettes when incubated with sheep red blood cells (SRBC). Thymus derived non-Ig bearing cells were found to be involved and the formation of rosettes between SRBC and lymphocytes served as the earliest marker of the human T cell lineage (Brain et al., 1970; Jondal et al., 1972; Moingeon et al., 1989). MAb specific for T cells that blocked this interaction were isolated (Howard et al.,

1981; Kamoun et al., 1981; Bernard et al., 1982) and immunoprecipitation of the rosetting receptor showed it to have an Mr of 50-55 kd, running as a heterogenous band on SDS PAGE. N-glycanase treatment reduced the Mr to 40 kd indicating the presence of 2-3 N-linked glycans (Sayre et al., 1987). Staining by the mAb was restricted to human T cells, thymocytes and NK cells. MAb OX-34 raised against rat T blasts (Jefferies et al., 1985) recognized a protein of similar Mr to the human protein and showed a similar tissue distribution with the additional staining of spleen macrophages. More recently mAb to mouse CD2 were isolated and the novel observation was made that mouse CD2 is also expressed on B cells (Yagita et al., 1989).

CD2 cDNA clones have been isolated for human (Sewell et al., 1986; Sayre et al., 1987; Seed and Aruffo, 1987), rat (Williams et al., 1987) and mouse (Clayton et al., 1987; Sewell et al., 1987; Yagita et al., 1989) molecules and they show extracellular domains of 185, 181, and 181 amino acids with 3, 4 and 5 N-glycosylation sites, respectively. The amino acid sequence conservation between the extracellular domains of human and rodent CD2 is low (48%) and the location of only one of the glycosylation sites (in domain 2) is conserved. The single transmembrane region and the 116 amino acid cytoplasmic domain is highly conserved, however, at 68% between the rat and human molecules. The human CD2 gene contains 5 exons with the signal peptide, each predicted extracellular domain, the transmembrane region and the cytoplasmic domain encoded by separate exons (Diamond et al., 1988; Lang et al., 1988).

The extracellular domain of CD2 contains conserved sequence motifs characteristic of IgSF molecules although the Cys residues which form the conserved disulphide bond between sheets are absent. Sewell (1986) and Williams et al., (1987) concluded that the CD2 extracellular domain was constituted by two IgSF domains, in particular a V-set domain and a C2-set domain (Williams et al., 1989). Reinherz and coworkers questioned this classification, however, and predicted the domain to have an α/β structure on the basis of computer modelling from the primary sequence (Clayton et al., 1987; Recny et al., 1990). This controversy is resolved in Chapter 8 where domain 1 of rat CD2 is shown to have a V-type fold by NMR spectroscopy (Driscoll et al., 1991).

1.10.2 The epitopes of CD2

A large array of mAbs against human CD2 have been generated (Meuer, 1989) and many of these react with one of three regions of the molecule, centred around (1) Lys 43, (2) Gly 90 and

(3) Tyr 135 (Peterson and Seed, 1987; Meuer, 1989). MAbs specific for region 1 or 2 are able to reduce or inhibit SRBC rosetting. A still poorly defined observation is that binding of some mAb to region 3 is very weak in the unperturbed CD2 molecule, but is increased after binding of mAb to other regions or binding of LFA-3. The induction of binding in region 3 is rapid and energy independent and has been argued to result from a change in the conformation of CD2 (Meuer et al., 1984; Meuer, 1989). The region 3 determinant(s) are also induced following T cell activation.

1.10.3 CD2 and cell activation

Of particular interest in understanding the function of CD2 is the ability of a combination of mAb specific for regions 1 and 3 (and certain other pairs whose epitopes were not mapped by Peterson and Seed), without further cross-linking antibody, to induce T cell activation (Meuer et al., 1984; Brottier et al., 1985; van Lier et al., 1989). This has also been observed with mAb OX-54 and OX-55 to rat CD2 (Clark et al., 1988). In most cases the activation is accessory cell dependent. A combination of aggregated LFA-3 and one anti-CD2 mAb was also able to induce calcium fluxes whereas monomeric LFA-3 is inactive (Dustin et al., 1989). In addition single mAbs, if further cross-linked with secondary antibody, can induce IL-2 production (He et al., 1988; Ledbetter et al., 1988; Beyers et al., 1989) and, in some cases, proliferation (Wesselborg et al., 1991).

Activation of T cells through CD2, referred to as the alternative activation pathway (Meuer et al., 1984), leads to a similar pattern of events to triggering through the CD3/TcR complex, including tyrosine phosphorylation of a number of proteins (Ley et al., 1991), inositol phosphate and diacylglycerol mobilization and a rise in cytoplasmic free calcium (Beyers et al., 1989; Moingeon et al., 1989), activation of $p21^{ras}$ (Graves et al., 1991), and expression of the IL-2R, IL-2 secretion and mitogenesis (Beyers et al., 1989; Moingeon et al., 1989). Signaling through CD2 on T cells is dependent on the presence of the TcR/CD3 complex at the cell surface (Breitmeyer et al., 1987; Alcover et al., 1988) and at least a fraction of CD2 is present on the membrane in association with this complex (Brown et al., 1989; Beyers et al., 1992). Signaling via CD2 reported to occur on NK cells (Seamen et al., 1987) appears to require the Fc portion of the triggering antibody to bring about cross-linking of CD2 and CD16 (Spruyt et al., 1991). The requirement for ligand binding to two sites on CD2 for signal transmission has been argued to mean CD2 has two natural ligands and some evidence for additional ligands is presented below. It is possible, however, as CD2 is most likely to have largely an enhancing

role in activation through the TcR complex, that the second site is normally involved in interactions with the TcR complex. The orientation of the CD2 molecule in the complex may be very important in the normal activation events that occur at cell-cell interfaces.

1.10.4 The CD2 cytoplasmic domain

The 116 amino acid cytoplasmic domain of CD2 is quite highly conserved across species (68% between rat and human) and is very rich in Pro and His (Williams et al., 1987). An autonomous role for the domain in CD2 signaling has been shown (Beyers et al., 1991) and through the generation of deletion mutants the COOH-terminal 70 amino acids were found to be required for signal transduction as measured by increases in $[Ca^{2+}]i$ or IL-2 secretion (He et al., 1988; Chang et al., 1989). Alteration of the His or the Arg in the sequence PPPGHR (residues 260-265) in human CD2 inhibited calcium mobilization and IL-2 production demonstrating a significant role for this region (Chang et al., 1990).

1.10.5 LFA-3 (CD58)

The dominant target structure on sheep erythrocytes involved in rosette formation with T cells was also identified by mAb which blocked rosette formation (Hunig, 1985). Independently it was observed that mAb to the human surface molecule, lymphocyte function associated antigen 3 (LFA-3), blocked rosetting and blocked binding of purified CD2 to LFA-3 positive cells (Selvaraj et al., 1987a, b). LFA-3 exists as both a transmembrane and GPI linked protein (Seed, 1987; Wallner et al., 1987) and the sequence indicates a molecule with an extracellular domain of similar structure to CD2, being constituted by a V-set fold lacking the conserved disulphide bond and a C2-set fold with 2 disulphide bonds. Six potential N-linked glycosylation sites are present, 3 in each domain, and differential usage of these sites could account for the observed range in apparent Mr from 40 -70 kd (Dustin et al., 1987). The LFA-3 gene is situated in the same region of chromosome 1 (1p13) as CD2 and it is possible that the two genes arose by duplication from a common ancestor (Brown et al., 1987; Sewell et al., 1988). It has been speculated that heterotypic adhesion interactions within the IgSF, such as that between CD2 and LFA-3, evolved from a more primitive homotypic type of recognition (Williams, 1987).

LFA-3 shows a broad tissue distribution in humans, being expressed on endothelial cells, B cells, macrophages and dendritic cells, activated T cells and red blood cells. The ability of SRBC but not untreated autologus red cells to form rosettes lies in the higher level of LFA-3

expression on SRBC (10 000 vs 4 000 molecules per cell) and their smaller surface area, which is only 35% that of human RBC (Selvaraj et al., 1987a).

LFA-3 has not yet been isolated in mouse or rat. The distribution of CD2 binding activity in the mouse has been studied using a decameric form of CD2 (an IgM-CD2 fusion protein) and was found to be more restricted than LFA-3 distribution in man, being expressed on immature and mature lymphoid cells but not on endothelial cells (Rutschmann and Karjalainen, 1991). This data should not be taken to represent murine LFA-3 distribution, however, as there is increasing evidence that CD2 can interact with more than one ligand and the construct may not have detected tissues expressing low levels of ligand.

1.10.6 CD2-LFA-3 mediated adhesion

All the binding activity of CD2 for LFA-3 resides in the first 115 amino acids of human CD2 (Richardson et al., 1988; Recny et al., 1990). This single domain form binds to LFA-3 on lymphoblastoid cells with a dissociation constant of 0.4 μM (Recny et al., 1990), a value identical to that obtained for the two domain form (Sayre et al., 1989) and similar to that for binding of monomeric LFA-3 to cell surface CD2 (Dustin et al., 1989). The adhesion interaction between CD2 and LFA-3 can be important in conjugate formation between cytotoxic T cells or NK cells and certain target cells (Shaw et al., 1986; Nakamura et al., 1990), T cells and B cells (Emilie et al., 1988), and T cells and other antigen presenting cells (Moingeon et al., 1989; Koyasu et al., 1990). A role for the interaction in thymocyte adhesion to thymic epithelium has been seen (Vollger et al., 1987; Nonoyama et al., 1989). However, the apparent lack of CD2 ligand on murine epithelium, the lack of CD2 expression on immature thymocytes in mouse and sheep, and the lack of effect of injecting anti-CD2 mAb into newborn mice or adding the mAb to thymic organ cultures, argue against a major role for CD2 in thymocyte maturation (Mackay et al., 1988; Kyewski et al., 1989; Yagita et al., 1989; Rutschmann and Karjalainen, 1991).

The general lack of spontaneous adhesion between cells expressing CD2 and LFA-3 is a result both of surface densities of these molecules, and surface charge of the cells. If autologous red cells are treated to reduce their negative charge they can form rosettes with resting human T cells (Bentwich et al., 1973; Plunkett et al., 1987). Similarly when T cells are activated they show both an increase in CD2 expression and a decrease in net negative charge and these cells will form rosettes with autologous red cells. Activated B cells also down regulate cell surface sialic

acid and CD2-LFA-3 mediated T-B cell conjugate formation is enhanced when either or both cell types are pre-activated (Kearse et al., 1988; Sen et al., 1992). The need for a balance between levels of CD2 and LFA-3 expression to control adhesion is seen in comparing human with sheep. In the sheep, where LFA-3 expression on red cells is high, CD2 expression on resting T cells is significantly lower than on human T cells (Mackay et al., 1988). As emphasized by Springer (1990), cellular interactions are an equilibrium between charge repulsion, and adhesive interactions.

1.10.7 GPI vs transmembrane linkage of LFA-3 and adhesion

LFA-3 can be expressed as both transmembrane and GPI-linked forms. The functional significance of this is unclear but there is evidence that GPI-linked proteins are more mobile in the cell membrane (Ferguson and Williams, 1988). Comparison of Jurkat cell binding to lipid bilayers containing LFA-3 showed that the GPI-linked form supported more rapid adhesion than the transmembrane form (Chan et al., 1991). The nature of the contact formed was also different, with more force being needed to detach a Jurkat cell from the GPI-linked substrate (Tozeren et al., 1992). These observations may not be a good representation of cell-cell interactions, however, as the transmembrane form in this system is totally immobile, but they do indicate that differences in adhesion can result depending on mobility of an adhesion receptor in the membrane.

1.10.8 Further ligands for CD2

a) CD48 (OX45, Blast 1)

The OX-45 antigen is a GPI linked protein predicted to have have a structure very similar to LFA-3 and CD2. The sequences are more similar to LFA-3 than CD2, and in particular the first domains of OX-45 and LFA-3 are most similar (Killeen et al., 1988; Staunton et al., 1989b; Wong et al., 1990b). The genomic location of OX-45 in mouse and human indicates it may have evolved from the same precursor as CD2 and LFA-3 (Staunton et al., 1989b; Wong et al., 1990b). The molecule is weakly expressed on all thymocytes and lymphocytes and in large amounts on activated T cells and macrophages in the rat. All endothelium examined were also positive (Arvieux et al., 1986). The corresponding human antigen, identified with the 17D6 mAb, was initially thought to be restricted to B blast cells and was called Blast 1 (Thorley-Lawson et al., 1982). More recently Blast-1 has been shown to be the CD48 antigen and is expressed on resting T and B cells and on monocytes (Korinek et al., 1991; Yokoyama et al., 1991). The restricted expression of the 17D6 epitope is not understood but the large number of

glycosylation sites (5) and the variability in the molecular weight of the molecule from different sources (Arvieux et al., 1986; Korinek et al., 1991) make it possible that the effect is due to differential glycosylation. There is some evidence that the CD48 molecule is able to bind both to itself and to CD2 (Yokoyama et al., 1991; Mouhtouris et al., 1991). One mAb specific for CD48 is able to block killing of an EBV transformed B cell line by a restricted set of γ/δ T cell clones (Del Porto et al., 1991). Most recently, during efforts to identify mouse LFA-3, it has been found that mouse CD48 is a major ligand for mouse CD2 (K. Okumura, pers. comm.).

b) CD59

CD59 is a member of the LY-6 family of GPI-linked surface glycoproteins (Philbrick et al., 1990) and is expressed on all peripheral blood lymphocytes, monocytes, granulocytes and erythrocytes. CD59 has been implicated in the protection of cells from complement lysis, inhibiting C5b-C8 catalysed insertion of C9 into lipid bilayers (Meri et al., 1990). More relevant to this discussion, mAbs to CD59, and purified CD59, are able to block formation of rosettes between T cells and red blood cells (Groux et al., 1989). CD59 transfected CHO cells support a low level of rosetting of human T cells but cells co-transfected with LFA-3 and CD59 show at least an additive, and possibly a synergistic, effect in rosetting. The CD59 transfected cells are also able to enhance proliferation of T cells stimulated by PHA and IL-1α, and co-transfected cells show a synergistic potentiating effect over CD59 or LFA-3 transfected cells alone (Deckert et al., 1992). Bernard's group have suggested that CD59 is able to bind directly to CD2 (Deckert et al., 1992).

1.10.8 A second E rosette receptor on T cells (E2)

The E2 antigen on human T cells was identified as a second target of rosette blocking antibodies (Bernard et al., 1988). E2 is also present on red cells, B cells and monocytes and certain non-hematopoietic cells. Biochemical analysis demonstrated that the molecule carries O-linked glycans but no N-linked sugars (Aubrit et al., 1989) and this was supported by sequence analysis (Gelin et al., 1989). The molecule has a 100 amino acid extracellular domain rich in Pro and Gly and lacking Cys, a single transmembrane domain and a 38 amino acid cytoplasmic domain. The sequence was found to be identical with MIC2, a pseudoautosomal gene, which is expressed on Xg(a+) red blood cells but is absent from the surface of Xg(a-) cells (Gelin et al., 1989). More recently it has been reported that E2 transfected CHO cells are unable to support rosette formation but if cells are co-transfected with CD2, rosette formation is observed and can be blocked with mAb to either CD2 or E2. The blocking was seen with rosettes of red cells or

of LFA-3 transfected CHO cells (Zoccola et al., 1991). While this data is suggestive it has not yet been formaly demonstrated that E2 binds to LFA-3.

1.11 LEUKOSIALIN (CD43, Sialophorin, Large sialoglycoprotein)

1.11.1 The Leukosialin molecule

Early studies of leukocyte surface glycoproteins involved direct labelling of surface carbohydrate, SDS PAGE and visualization by autoradiography (Gahmberg and Andersson, 1977). Such studies lead to the observation that leukocytes had only 1-3 major glycoproteins. Thymocytes had bands at 150, 95-115 and 25 kd, T cells (in the rat) two bands at 95-105 and 150-170 kd and B cells a single band at 200-220 kd (Gahmberg and Andersson, 1977; Standring et al., 1978; Brown et al., 1981). Human myeloid and erythroid cell lines had a single dominant band at between 105 and 130 kd depending on their state of differentiation (Fukuda et al., 1981). These bands on the different cell types have been shown to correspond to three major glycoproteins, Thy-1 (25 kd), leukosialin (95-135 kd) and alternatively spliced forms of the leukocyte common antigen or CD45 (150-220 kd) (Williams et al., 1977; Brown et al., 1981; Dalchau et al., 1980). Their dominance in this labelling system is a result of both their heavy glycosylation and their high level of expression with site numbers on rat thymocytes being estimated as 10^6, 10^5 and $7x10^4$ for Thy-1, leukosialin and CD45, respectively (Williams and Barclay, 1986).

1.11.2 Leukosialin tissue distribution

In short, leukosialin expression has been found on all cells of the immune system other than resting B cells and dendritic cells. In the rat expression has been demonstrated on all T cells and thymocytes, on polymorphonuclear cells, on 40% of fetal liver cells and 60-70% of bone marrow cells (40% brightly, 25% dimly) including all B and T stem cells, on brain tissue (Williams et al., 1977; Dyer and Hunt, 1981) and on activated B cells and plasma cells (Williams, unpublished). In humans a similar pattern of expression has been demonstrated (Dalchau et al., 1980; Kamps et al., 1989; Stross et al., 1989). In addition expression has been shown on basophils and mast cells (Valent et al., 1990), monocytes and macrophages, lymphokine activated killer (LAK) cells, large granular lymphocytes (LGL) and NK cells (Stross et al., 1989; Stoll et al., 1989). The large majority of cell lines of T or B cell origin are positive (Stoll et al., 1989). Red blood cells are negative but it is still unclear as to whether the molecule is (Gahmberg et al., 1988) or is not (Stross et al., 1989) expressed on platelets. Dendritic cells are negative (Prickett et al., 1992). Tissue distribution in the mouse, to the extent

that it has been examined, shows the same pattern as in rat and human (Baecher et al., 1988; Gulley et al., 1988). Hardy and colleagues have shown that murine B lineage cells at the pre-pro and pro stage express leukosialin and then downregulate it upon V-D-J rearrangement and transition to the pre-B cell stage (Hardy et al., 1991). Although it is generally stated that resting B cells do not express leukosialin, there is some evidence for low level expression with certain antibodies in human (Dalchau et al., 1980; Stross et al., 1989) and rat (Chapter 5) but the level of expression in these cases, detected by flow cytometry, is several orders of magnitude lower than that seen on T cells or activated B cells and thus resting B cells will continue to be considered negative in this thesis.

1.11.3 Leukosialin heterogeneity

a) Heterogeneity in Mr and lectin binding

Leukosialin molecules can show heterogeneity in apparent molecular mass as indicated above, and one prominent characteristic of the molecule is its reduced migration on SDS PAGE following neuraminidase treatment, an effect seen with other heavily O-glycosylated and sialylated proteins (Gahmberg and Andersson, 1982). Early studies in the mouse demonstrated the induction of a protein of 145 kd (the asialo form) on T cells 2-3 days after activation with Con A or in the MLR (Kimura and Wigzell, 1978). This protein had the selective property of binding *Vicia villosa* lectin (Kimura et al., 1979) and it was shown to be directly related to a lower M_r form on resting cells (130 kd in the asialo form) that was specifically bound by *Helix pomatia* lectin (Conzelmann et al., 1980; Pink, 1983). The *Helix pomatia* lectin binding protein in human cells was subsequently purified and shown to correspond to leukosialin (Axelsson et al., 1985; Fukuda and Carlsson, 1986). Human T cells were also observed to lose the 120 kd surface labelled band following 3 days of stimulation and acquire a 130 kd band (the asialo forms; Andersson et al., 1978). If the blasts were then left unstimulated for 6 days the 130 kd band was replaced by the 120 kd band. More recent studies have shown that human lymphocytes, monocytes and thymocytes have a form of 115 kd whereas granulocytes and activated lymphocytes have a 125 kd form (Carlsson and Fukuda, 1986; Remold-O'Donnell et al., 1987; Gahmberg et al., 1988). In the mouse the 3E8 antibody (Baecher et al., 1988) was shown to bind a 115 kd form of leukosialin on non-stimulated speen cells and thymocytes, 115 and 125 kd forms on bone marrow and stimulated spleen cells, and forms ranging in size from 115 to 145 kd on various T cell lines.

The existence of molecular weight heterogeneity in rat leukosialin has not been examined. The

thymocyte and T cell forms which have been analyzed in detail (Brown et al., 1981) are of similar apparent Mr (95-100 kd), 5-15 kd smaller than the corresponding forms in mice and humans.

b) Antigenic heterogeneity

A number of the mAb specific for human leukosialin only react with certain leukosialin expressing cell types. The GA3 mAb, raised against K562 cells, stains only 4% of CD4 cells and 9% of bone marrow cells and reacts with erythoid but not hematopoietic cell lines. The epitope is neuraminidase sensitive and present only on 105 kd forms of the molecule (Bettaieb et al., 1988). MAb T305, raised against a T-ALL line, reacts with only a fraction of peripheral T cells (22%), shows little reactivity with granulocytes, but binds monocytes (80%) and bone marrow cells (65%). Reactivity of T cells with the mAb increases in disease states and after *in vitro* activation by poke weed mitogen (Fox et al., 1983; Sportsman et al., 1985). The epitope is present on the high molecular weight form of leukosialin which carries mainly hexasaccarhides, and it requires the presence of sialic acid linked α2-3 to galactose (Sportsman et al., 1985; Saitoh et al., 1991). The mAbs B1B6 and E11B also recognize sialic acid dependent epitopes but they react with the majority of cells expressing leukosialin, exceptions being monocytes and the majority of thymocytes (Axelsson et al., 1988). Similarly mAb F10-44-2 (Dalchau et al., 1980) showed reactivity with only a fraction (29%) of thymocytes. These were shown to be the more mature cells in the medulla. Lack of reactivity with immature thymocytes may be due to their low level of sialylation (see below). Of the additional mAb reported at the 4th Leukocyte typing workshop (Stoll et al., 1989) eight out of eight were sensitive to removal of sialic acid from leukosialin transfected cells. The only mAb specific for human leukosialin reported to be neuraminidase insensitive is L10 (Remold-O'Donnell et al., 1984) a mAb which has been shown to activate T cells and monocytes, as described in more detail below.

A small number of mAb to mouse leukosialin have been isolated and they show similar staining profiles (Baecher et al., 1988; Gulley et al., 1988). Antigenic heterogeneity for rat leukosialin has been demonstrated by the finding that the W3/13 mAb does not bind all of the leukosialin on rat thymocytes (Brown et al., 1981).

1.11.4 Biochemical properties

Leukosialin purified from rat thymocytes was estimated to be 60% carbohydrate by weight and the compostion of the carbohydrate, and its sensitivity to release by β-elimination, indicated it was probably all O-linked (Brown et al., 1981). Similar observations were made for human leukosialin although some evidence for a low level of N-linked glycosylation was obtained (Remold-O'Donnell et al., 1986; Carlsson and Fukuda, 1986). Pulse chase experiments with human leukosialin indicated a single precursor protein (approx. 54 kd) in different cell types that could be reduced to 52 kd by endoglycosidase H treament, and this was modified to different M_r forms by attachment of O-glycans in the Golgi compartment (Carlsson and Fukuda, 1986; Remold-O'Donnell et al., 1987). The heavy O-glycosylation of the molecule apparent from early work lead to the prediction that leukosialin was a lymphocyte mucin, acting solely to contribute to the glycocalyx and negative charge of the cell.

Isolation of cDNA clones for rat (Killeen et al., 1987) and human (Pallant et al., 1989; Shelley et al., 1990) leukosialin has confirmed the biochemical data. The molecules have extracellular domains of 224 and 235 amino acids in rat and man, respectively, that are composed of approximately 40% Ser and Thr, allowing for the attachment of 80-90 O-linked glycans. A single N-linked site is present in human but not rat leukosialin. A finding not predicted from the biochemical data was the long cytoplasmic domain (124 amino acids in rat, 123 in man). This domain is 77% conserved between rat and human providing indirect evidence that leukosialin may act as more than a cell surface mucin. This observation is extended in Chapter 3 where the mouse leukosialin sequence is reported (Cyster et al., 1990) and the level of conservation between rat and mouse cytoplasmic domains is shown to be 90%. A single mRNA size has been detected in Northern blots of rat RNA (Killeen et al., 1987) whereas two mRNAs exist for human leukosialin (Pallant et al., 1989). Recent isolation of genomic clones for human leukosialin (Shelley et al., 1990; Kudo and Fukuda, 1991) has revealed that the molecule is encoded by an uninterrupted open reading frame and that the different mRNAs result from alternate usage of polyadenylation sites. The single gene present in the human genome is localized to chromosome 16 band p11.2, close to the location of the α-subunits of the glycoproteins LFA-1, Mac-1 and p150,95 (16p11-p13.1) (Shelley et al., 1989). Mouse leukosialin has recently been mapped to chromosome 7, which also carries the gene for the mouse LFA-1α chain (Baecher et al., 1990).

1.11.5 Leukosialin as an activation molecule

One mAb to human leukosialin, L10, was observed to stimulate proliferation of peripheral blood T lymphocytes, in an accessory cell dependent manner, to an extent similar to anti-CD3 mAb (Mentzer et al., 1987). Induction of phosphotidylinositide hydrolysis was observed in the absence of accessory cells whereas a significant calcium flux could only be observed in their presence. Of particular significance, it was argued that signalling was independent of the TcR/CD3 complex, for a normal calcium flux was seen in a cell line largely defective in TcR expression (Silverman et al., 1989). However, pretreament of peripheral blood T cells with PMA, which is known to down regulate TcR expression, was observed to inhibit the calcium flux induced by L10 mAb (Wong et al., 1990a). This was argued to result from hyperphosphorylation of the leukosialin cytoplasmic domain induced by the PMA treatment, but other possible explanations were not ruled out (Wong et al., 1990a). L10 induced activation of monocytes has also been shown involving a 5-7 fold increase in H_2O_2 production and homotypic aggregation of the cells (Nong et al., 1989). In no case have any other mAb specific for leukosialin, including 5 mAb to the rat molecule, been shown to have such potent stimulatory effects. Weak potentiating effects of mAb on the MLR (Webb et al., 1979) and on phorbol ester and anti-CD3 induced proliferation of T cells have been reported (Axelsson et al., 1988; Beyers et al., 1989). A potentiating effect on activation was also seen when human leukosialin was transfected into a murine T cell hybridoma specific for human Daudi cells, but not when a construct lacking the ctoplasmic domain was transfected (Park et al., 1991). In summary it is still unclear whether leukosialin has the capacity to act as a signaling molecule independent of the TcR. More extensive evidence indicates a role in enhancing TcR mediated responses.

1.11.6 The O-linked glycans of leukosialin

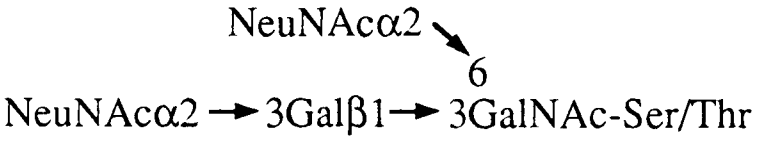
Initial studies on the O-glycosylation of leukosialin on K562 (erythroid), HL-60 (myeloid) and HSB-2 (T lymphoid) cell lines (Carlsson et al., 1986) have been followed by work on leukosialin from freshly isolated cells. Resting human T cells carry mainly the disialylated type 1 core (Fig 1.6). Following 3 days of activation by anti-CD3 mAb and IL-2, or PHA or in the MLR, the dominant sugar structure (70%) becomes the disialylated hexasaccacharide form of the type 2 core (Piller et al., 1988; Fig 1.6) which accounts for the shift to higher M_r seen following T cell activation and may also explain the differences in lectin binding of leukosialin on resting versus activated T cells in the mouse (Kimura and Wigzell, 1978). The change in structure results from an increase in core 2 β1-6GlcNAc transferase activity and reduced α2-6

sialyltransferase activity (Piller et al., 1988). The increase in Mr of leukosialin following activation is detectable at 48 hr (Baecher et al., 1990) and presumably reflects the time course over which leukosialin carrying the hexassaccharide replaces the existing form. In an early study where the cells where left for 6 days after stimulation the leukosialin had returned to the lower Mr form (Andersson et al., 1978) but it has not been shown whether the oligosaccharides attached are changed back to the core 1 structures.

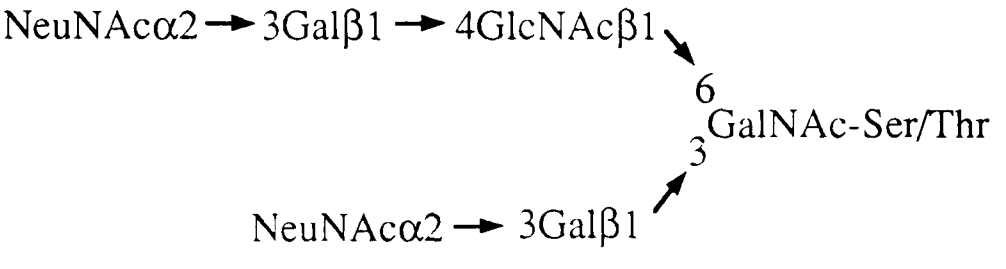
Granulocyte leukosialin carries mainly the hexasaccharide glycans with more extended forms of the type 2 core as minor components (Fukuda et al., 1986; Fig 1.6). In particular it was shown that the O-glycan structures of leukosialin could contain poly-N-acetyllactosamines (Fukuda et al., 1986) which are acceptors for α1-3 fucosyltransferases and as a result, Le^x and sLe^x determinants can be formed. Such O-linked structures have been shown to be present at low levels on human granulocyte glycoproteins, although it has not been directly demonstrated that they are on leukosialin (Fukuda et al., 1986). It should also be noted that poly-N-acetyllactosamine extension preferentially takes place on the side chain arising from the GlcNAc residue with β1-6 linkage to GalNAc, so branched (core 2) structures formed on activated T cell leukosialin may also carry some poly-N-acetyllactosamine.

The O-linked glycans attached to thymocyte leukosialin may differ between rat and human. Rat thymocyte leukosialin (purified with the W3/13 mAb) does not carry detectable amounts of glucosamine and carries mainly core 1 structures (Brown et al., 1981). In contrast, human thymocyte leukosialin reacts with a mAb specific for the Galβ1-4GlcNAcβ1-6 structure (Childs and Feizi, 1981), and the core 2 hexasaccharide sugar has been estimated to account for 40% of thymocyte leukosialin O-glycosylation (Pink, 1983; Saitoh et al., 1991). However, both rat and human are similar with regard to the low level of sialylation of leukosialin and other major glycoproteins on thymocytes (Despont et al., 1975; Brown et al., 1981; Pink et al., 1983). In the rat, untreated thymocyte leukosialin is the major PNA binding protein whereas peripheral T cell leukosialin will only bind following neuraminidase treatment (Brown and Williams, 1982), and similar results have been observed in mouse and man with *Helix pomatia* lectin binding (Axelsson et al., 1978; Pink, 1983). As the cells mature the level of sialylation increases, and thymocytes in the medulla show decreased PNA binding (Rose, 1982) and increased reactivity with mAbs that are dependent on the presence of sialic acid for binding (Dalchau et al., 1980). Whether the core 2 sugar structures present on human leukosialin are enriched on mature or immature thymocytes has not been examined.

1. **Resting T cells, thymocytes, lymphoid and erythroid cell lines**
   core 1 structures



2. **Activated T cells, thymocytes, lymphoid cell lines**
   core 2 structures



3. **Granulocytes, myeloid cell lines**
   elongated core 2 structures

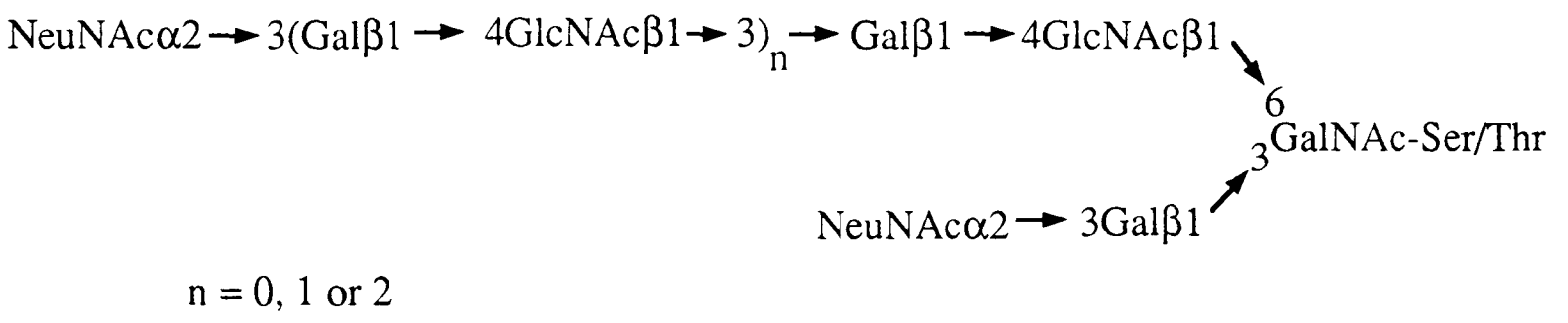


Figure 1.6 Major forms of O-linked oligosaccharides identified on leukosialin.
Fully sialylated forms are shown only.

Leukosialin from patients with acute T-lymphoblastic leukemia carries a large amount of the branched hexasaccharides characteristic of activated T cells whereas that from patients with chronic T lymphocytic leukemia mainly carries the tetrasaccharide (Saitoh et al., 1991). These observations are in agreement with the pattern of reactivity with the T305 mAb (Fox et al., 1983), which has been shown to react preferentially with leukosialin containing hexasaccharides (Saitoh et al., 1991). The O-linked glycans on Wiskott Aldrich syndrome patient lymphocytes are also unusual as discussed below.

1.11.7 Wiskott Aldrich Syndrome (WAS)

The WAS is an X-linked primary immunodeficiency affecting 1 in 250 000 male births. The disease is characterized by reduced T cell numbers, reduced response to carbohydrate antigens, eczema and thrombocytopenia and susceptibility to opportunistic infection. Serum IgA and IgE levels are elevated whereas IgM is decreased and IgG is normal. The T cells exhibit a progressive decline in number and function so that profound lymphopenia does not become apparent until 6 years of age (reviewed in Rosen et al., 1984). Scanning electron microscopy shows the cells have a reduction in surface microvilli (Kenney et al., 1986) and in some cases of cell volume (Parkman et al., 1981). Platelets are also half the normal size and turnover rapidly. All these defects can be corrected by bone-marrow transplantation (Ochs et al., 1980; Rosen et al., 1984). Parkman and colleagues observed that the 115 kd band present in normal lymphocyte membrane preparations, which corresponds to leukosialin, was absent in a number of WAS patients (Parkman et al., 1981; Parkman et al., 1983). Some abnormalities in platelet surface proteins were seen (Parkman et al., 1981) but this result has not been consistent (Pidard et al., 1988). Subsequent work has shown that leukosialin is not absent from lymphocytes of WAS patients, but the form present on resting T cells is of higher molecular weight than in controls (Remold-O'Donnell et al., 1984) resembling a form which has had sialic acid removed or the high M_r form which arises on normal T cells after activation. The gene for human leukosialin is not on the X chromosome but chromosome 16 (Pallant et al., 1989; Shelley et al., 1989), ruling out a direct defect of leukosialin in this disease.

Recent evidence indicates a defect in the expression of certain glycosyl-transferases in WAS lymphocytes and platelets (Higgins et al., 1991; Piller et al., 1991). In particular, resting T cells have an elevated level of β1-6GlcNAc transferase activity, being comparable to the level in activated cells of controls. As a result the leukosialin on resting T cells carries mainly O-linked

hexasaccharides rather than tetrasaccharides, accounting for both the increase in apparent Mr of WAS leukosialin (Remold-O'Donnell et al., 1984) and the increased reactivity of most WAS T cells with the T305 mAb (Piller et al., 1991). This transferase activity was also three times higher in WAS platelets than in normal platelets (Higgins et al., 1991). The pattern in B cells is not as clear but it would appear that the β1-6GlcNAc transferase activity is reduced compared to controls (Higgins et al., 1991). It has not yet been determined whether other non-glycosylation based changes occur in WAS lymphocytes, but these observations provide preliminary evidence that changes in cell-surface O-linked glycosylation can have major effects on the normal functioning of lymphocytes.

1.11.8 Soluble Leukosialin

Leukosialin has joined an increasing family of cell-surface proteins in having a soluble form (Schmid et al., 1992). The family includes: CD8, CD14, CD23, CD30, CD44, the selectins, decay accelerating factor, transferrin R, IL-1Rβ, IL-2Rα, IL-4R, IL-5Rα and IL-7R. Extensive biochemical studies on galactoglycoprotein, a protein isolated from human serum, revealed similarities to the extracellular domain of leukosialin (Schmid et al., 1980; Brown et al., 1981; Akiyama et al., 1984) and these have now been confirmed by protein sequencing (Schmid et al., 1992). The presence of only a single gene for human leukosialin and the lack of introns in the coding region means that the soluble form can only be generated by proteolysis of the cell membrane form. Cleavage is not at a precisely defined site but appears to be within 10-20 amino acids of the transmembrane domain (Schmid et al., 1992). The product generated is stable and of relatively uniform size (120 kd) and carries predominantly hexasaccharide forms of O-linked glycan. This rules out resting T cells or monocytes as the source of the majority of the soluble form, and indicates it is probably released from neutrophils or activated T cells (Schmid et al., 1992; Piller et al., 1988; Fukuda et al., 1986). The glycosylation state of plasma cell leukosialin has not yet been characterized and may also be a source of the soluble form.

1.11.9 Leukosialin - ICAM-1 interaction

An interaction between purified leukosialin and ICAM-1 has recently been demonstrated by Burakoff and colleagues (Rosenstein et al., 1991). The study showed both partial blocking of cell binding to purified leukosialin or ICAM-1 by anti-ICAM-1 and anti-leukosialin mAb, and clustering of ICAM-1 coated beads with leukosialin coated beads. It was not demonstrated in these experiments whether the interaction involved the leukosialin protein core, the carbohydrate or both. The physiological significance of this association remains to be assessed.

Aims of the work described in this thesis

The principle aim of this thesis was to examine the structure and function of leukosialin.

- The initial aim was to isolate a mouse genomic clone for leukosialin to complete the sequence of the mouse molecule and to address the question of the origin of heterogeneity of mouse leukosialin. Isolation of a genomic clone was also considered important to facilitate transgenic studies and ultimately to allow gene targeting experiments.
- The possibility that leukosialin might be involved in cell migration pathways was tested by expressing the antigen on B cells in transgenic mice.
- The structure of the molecule was examined by electron microscopy and the role of carbohydrate in the antigenic character of the extracellular domain assessed.
- The highly conserved nature of the cytoplasmic domain prompted its expression and structural analyses were performed combined with some initial steps towards isolation of proteins with which the domain interacts.
- The abilitiy of anti-leukosialin antibodies to induce cell aggregation was compared and the activity of monomeric Fab fragments was tested.

Arising from an interest in localizing epitopes in CD2, the first domain of rat CD2 was found to be very efficiently expressed in *E.coli*. As a result, and due to the interest in this domain as a predicted example of an IgSF fold lacking the conserved disulphide bond and because of its interaction with LFA-3, samples were prepared to allow NMR spectroscopy. Paul Driscoll's successful determination of the solution structure of the domain stimulated efforts to produce human CD2 and human LFA-3 domain 1 to allow a structural assessment of the adhesion interaction.

CHAPTER 2 MATERIALS AND METHODS

2.1 Materials

2.2 General buffers and media

2.3 General DNA methods

- 2.3.1 Phenol extraction and ethanol precipitation
- 2.3.2 Determination of DNA concentration
- 2.3.3 Restriction enzyme digestion of DNA
- 2.3.4 Agarose minigels
- 2.3.5 Isolation of DNA restriction fragments and PCR products
- 2.3.6 Ligation of inserts into vectors
- 2.3.7 Preparation of competent cells
- 2.3.8 Transformation of competent bacteria
- 2.3.9 Transfer of bacterial colonies to nitrocellulose filters
- 2.3.10 ^{32}P-labelling of cDNA and oligonucleotides
- 2.3.11 Screening of nitrocellulose filters with ^{32}P-probes
- 2.3.12 Southern blotting
- 2.3.13 Plasmid preparations

2.4 Nucleotide sequencing

- 2.4.1 Preparation of ssDNA templates for nucleotide sequencing
- 2.4.2 Sequencing reactions
- 2.4.3 Sequencing gels

2.5 *In vitro* mutagenesis

2.6 Polymerase chain reaction and Bal 31 deletion

- 2.6.1 PCR conditions
- 2.6.2 Isolation of DNA fragments by PCR for expression in the pGEX-2T vector
- 2.6.3 Generation of clones by Bal 31 deletion

2.7 Plating out EMBL-3 genomic library

- 2.7.1 Isolation and characterization of genomic clones
- 2.7.2 Isolation of lambda DNA
- 2.7.3 Computer programmes

2.8 Preparation of a full length rat leukosialin cDNA clone

2.9 Analysis of transgenic mouse DNA

- 2.9.1 Isolation of genomic DNA
- 2.9.2 PCR detection of transgene
- 2.9.3 Southern and Northern blotting of transgenic mouse DNA and RNA

2.10 Transfection, analysis and storage of eukaryotic cell lines

- 2.10.1 Electroporation

2.10.2 Calcium phosphate mediated transfection
2.10.3 Transient transfection mediated by DEAE-Dextran
2.10.4 Storage of cell lines in liquid nitrogen
2.10.5 Flow cytometry and FACS
2.10.6 Neuraminidase treatment
2.11 Immunoperoxidase staining of tissue sections
2.12 Expression in *E.coli*
2.12.1 Mini-protein inductions to screen clones for fusion protein expression
2.12.2 Bulk protein production
2.12.3 Cleavage of fusion proteins with thrombin
2.13 Analysis of proteins
2.13.1 SDS PAGE
2.13.2 Western blotting
2.13.3 Plate binding assay
2.13.4 Inhibition assay
2.13.5 Amino acid compositional analysis
2.13.6 Determination of protein concentration
2.14 Purification of antibodies and proteins
2.14.1 Purification of IgG and preparation of Fab fragments
2.14.2 Coupling of antibodies to Sepharose 4B
2.14.3 Antibody affinity columns
2.14.4 Gel filtration
2.14.5 Concentration of protein samples
2.14.6 Sucrose gradients
2.15 Labelling of antibodies and proteins for use as probes
2.15.1 FITC conjugation and biotinylation of proteins
2.15.2 Iodination of proteins
2.16 Phosphorylation by Protein kinase C
2.17 V8 protease cleavage and HPLC of leukosialin cytoplasmic domain
2.18 Rotary shadowing and electron microscopy
2.19 Detection of interactions with the leukosialin cytoplasmic domain
2.19.1 Preparation of eukaryotic cell lysates and cell membranes
2.19.2 Ligand blotting with cytoplasmic domain fusion protein
2.19.3 Affinity purification from cell lysates with fusion proteins
2.20 Aggregation assay

2.1 Materials

DNA reagents, kits and enzymes: The *in vitro* mutagenesis kit was from Amersham International. Sequenase Version 2 kits were from United States Biochemical Corporation. DNA restriction enzymes were from Amersham International, New England Biolabs, or Boehringer Mannheim. T4 polynucleotide kinase, T4 DNA ligase and the Klenow fragment of DNA polymerase were from United States Biochemical Corporation. T7 DNA polymerase was from Pharmacia. Ultrapure deoxynucleotides (dATP, dTTP, dCTP and dGTP) were from Pharmacia. DEAE-Dextran, and chloroquine diphosphate were from Sigma. Redistilled phenol was from Gibco, Paisley, UK. Yeast extract and agar were from Oxoid, Ltd. and bactotryptone (Difco 012301) was from Difco Ltd. Other reagents, if not specifically stated in the text, were from Aldrich Chemical Company, British Drug House (BDH), Boehringer Mannheim or Sigma, and were usually of analytical grade.

Oligonucleotides were synthesized by Keith Gould, Sir William Dunn School of Pathology, on an Applied Biosystems 380B DNA synthesizer.

Radiochemicals, α-[^{35}S]-dATP (10mCi/ml, >1000Ci/mmol), α-[^{32}P]-dATP (10mCi/ml, 3000Ci/mmol), γ-[^{32}P]-ATP(10mCi/ml, >5000Ci/mmol), methyl-[^{3}H]-thymidine (1mCi/ml, 2 Ci/mmol) and Na^{125}I (100mCi/ml) were purchased from Amersham International, U.K.

Vectors: M13 and pTZ vectors were from Amersham. The pKG5 expression vector was a gift from Keith Gould (Oxford). pGEX-2T (Fig 2.1) was provided by Duncan Campbell (Oxford) and pGM-T7 by Geoff Smith (Oxford). pHSE3' was from Dave Ferrick (Toronto) and the CDM8 vector was provided by Sue Mallet (MRC CIRU). pCM1-Neo A+ and A- were from Stratagene and the pTK(ES) plasmid was from Bill Colledge (Cambridge).

***E.coli* Strains**: The genotypes of the strains used are given in Sambrook et al., (1989). The aspects of the strains relevant to this work are: MC1061, high efficiency of transformation; Y1090, deletion in the *lon* gene which encodes a protease that is involved in degradation of aberrantly expressed proteins (Goff and Goldberg, 1985; Chin et al., 1988), the strain was cured of the pMC9 plasmid which carries the lacI^{q} gene and the ampicllin resistance gene by repetitive plating on non-selective plates; TG1, carries the F' pilus allowing infection by M13 phage, and is able to grow on minimal medium supplemented with thiamine; BL21(DE3), carries the bacteriophage T7 RNA polymerase on the bacteriophage lambda DE3 which is integrated into the chromosome (Studier and Moffatt, 1986) allowing high level expression of genes cloned into expression vectors containing the bacteriophage T7 promoter, and the strain is prototrophic and thus able to grow on minimal medium alone; K802, used to propogate bacteriophage lambda vectors; UT5600, a B strain which has a deletion encompassing the *omp T* gene encoding outer membrane protease T (Grodberg and Dunn, 1988).

Eukaryotic cell lines: The Jurkat (E6-1) line was from A.Weiss, UCSF. The mouse myeloma line NSO (Clark and Milstein, 1981) and rat myeloma line Y3 (Galfre et al., 1979) and the

simian fibroblast line COS (Gluzman, 1981) were provided by Mike Puklavec (MRC CIRU). The Madin Darby Canine Kidney (MDCK) cells, a polarized epithelial cell line, and the K562 erythroleukemia cell line was from the cell bank, Sir William Dunn School of Pathology.

Antibodies: A list of the mAbs used most frequently is given in table 2.1. Additional mAbs are stated in the text or listed in table 7.1. 5H4, 8B8 and 5G7 were provided by Sigjorn Fossum, Oslo, and had been raised against rat lymphokine ativated killer cells. The S7 and S11 mAbs, believed to be specific for mouse leukosialin, were kindly provided by David Kemp, Iowa. The L3T4, Ly2 and Ly5 mAb from Coulter were direclty conjugated to phycoerythrin and were used at 1/100. OX-21 (mouse IgG1), directed against human C3b inactivator, was from the MRC CIRU and was used as a control in experiments involving mouse antibodies. MT-463 (IgG_{2b}), mouse anti-human CD4, was used as an isotype matched control for some studies employing IgG2b antibodies. NDS58 (rat IgG2b), specific for a rat CD45 allotype (Newton et al., 1986), was used as a control in experiments involving rat antibodies. RAM $F(ab')_2$ conjugated to FITC was from Steve Simmonds. Sheep anti-mouse IgGγ1 $F(ab')_2$ was provided by Alan Williams. Peroxidase conjugated rabbit antiserum against mouse, rat and rabbit immunoglobulins was from Dakopatts, Denmark.

2.2 General Buffers and Media

Buffers:

Phosphate-Buffered Saline (PBS): 137 mM NaCl, 2.7 mM KCl, 8.1 mM Na_2HPO_4, 1.5 mM KH_2PO_4, pH 7.3, was prepared from tablets obtained from Oxoid Ltd.

PBS/BSA is PBS containing 0.25% (w/v) bovine serum albumin (BSA)

tris/sal: 140 mM NaCl, 25 mM Tris-HCl (pH 7.4), 0.02% sodium azide

tris/sal/EDTA: tris/sal with 5 mM EDTA

5x TBE: 445 mM Tris-HCl, 445 mM boric acid, 2 mM EDTA (pH 7.4)

20x SSC: 3 M NaCl, 0.3 M Na-citrate (pH 7.0)

SM: 100 mM NaCl, 8 mM $MgSO_4.7H_2O$, 50 mM Tris-HCl (pH 7.5), 0.01% Gelatin

TE : 10 mM Tris-HCl (pH 8), 1 mM EDTA (prepared from a 0.5 M stock at pH 8.0)

$T_{10}E_{0.1}$: 10 mM Tris-HCl (pH 8), 0.1 mM EDTA

6x glycerol dye: 30% glycerol (v/v), 10 mM EDTA (pH 8), 0.5% (w/v) bromophenol blue

Coomassie blue staining solution: 0.125% (w/v) Coomassie blue R250, 45.5% (v/v) methanol, 9% (v/v) glacial acetic acid

Destaining solution: 7.5% (v/v) glacial acetic acid, 5% (v/v) methanol

All other buffers are described in the text or are from Sambrook et al., (1989).

Bacterial growth media:

L broth: 1% (w/v) bactotryptone, 0.05% (w/v) yeast extract, 171 mM NaCl (autoclaved)

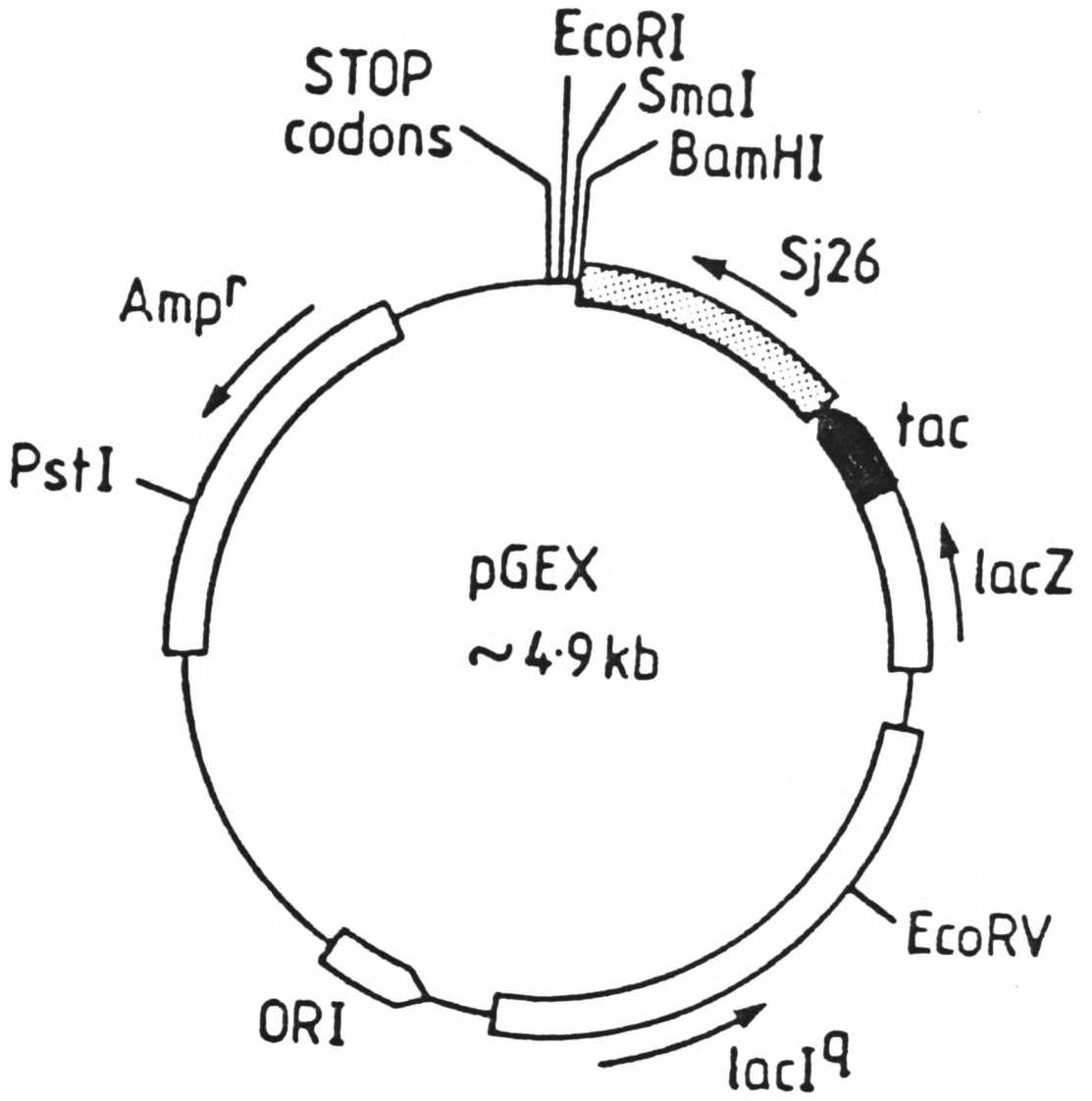
2x TY : 1.6% (w/v) bactotryptone, 1% (w/v) yeast extract, 86 mM NaCl (autoclaved)

TYM broth: 2% (w/v) bactotryptone, 0.5% (w/v) yeast extract, 0.1 M NaCl, 10 mM $MgS0_4$

Table 2.1 Monoclonal antibodies used in this thesis

MAb	Isotype	Specificity	Reference
Mouse mAb		Specific for rat antigens	
OX-1	IgG1	CD45, common	Sunderland et al., 1979
OX-2	IgG1	Thymocyte, brain, B cells, dendritic cells, endothelium	McMaster and Williams, 1979
OX-6	IgG1	MHC class II non-polymorphic	
OX-7	IgG1	Thy-1	Mason and Williams, 1980
OX-8	IgG1	CD8	Brideau et al., 1980
OX-34	IgG2a	CD2	Jefferies et al., 1985
OX-53	IgG1	CD2, OX34 competitive	
OX-54	IgG1	CD2, non-competitive	Clark et al., 1988
OX-55	IgG1	CD2, non-competitive	
W3/13	IgG1	Leukosialin	Williams et al., 1977
OX-56	IgG2b	Leukosialin	Killeen et al., 1987
OX-57	IgG2b	Leukosialin	Produced by William Brown and Mike Puklavec. Primary immunization with purified leukosialin and secondary boost with thymocyte membranes
OX-58	IgG2b	Leukosialin	
OX-74	IgG1	Leukosialin	
OX-75	IgG2b	Leukosialin	
5H4	IgG1	Leukosialin	Provided by Sigbjorn Fossum
8B8	IgM	Leukosialin	
5G7	IgG1	Leukosialin	
W3/25	IgG1	CD4	Williams et al., 1977
W3/15	IgG1	thymocytes, bone marrow	
Rat mAb		Specific for mouse antigens	
S7	IgG2a	Leukosialin	provided by David Kemp
S11	IgG2a	Leukosialin	
OX-20	IgG	Ig kappa chain	Mason et al., 1981
YA9	ND	mouse IgG1	J. Jarvis unpublished
YBM	ND	Thy 1.1 and Thy 1.2	S. Cobbold unpublished
L3T4	IgG2b	CD4	Coulter Immunology, Coulter corporation Hialeah, FL
Ly2	IgG2b	CD8	
Ly5		CD45R (B220)	

A.



B.

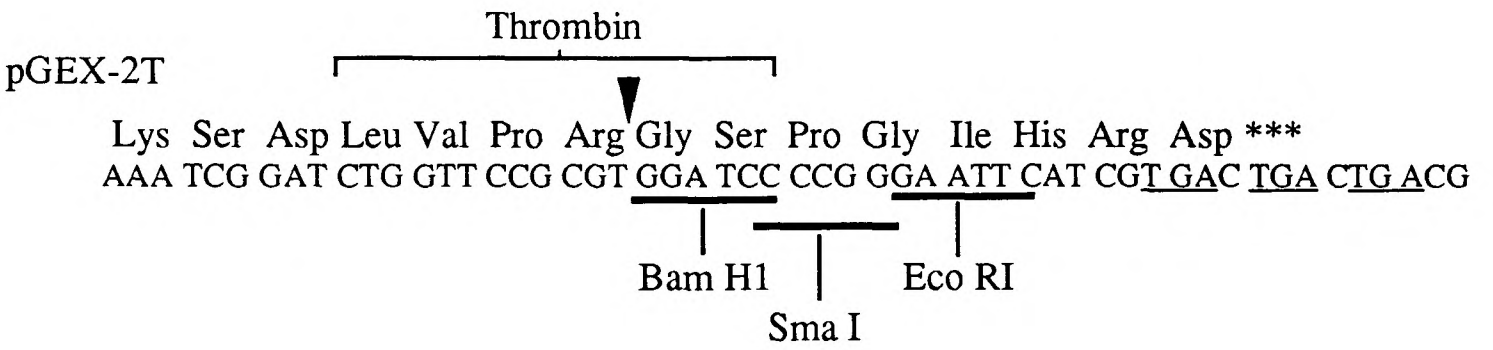


Figure 2.1 Structure of the pGEX vector and the pGEX-2T thrombin cleavage and cloning site. A) Schematic representation of a pGEX vector. Sj26 encodes the glutathione-S-transferase of *Schistosoma japonicum* and expression is driven by the tac promoter. Lac I^q encodes the lac repressor which binds the tac promoter and prevents transcription; lactose or lactose analogues (such as IPTG) block repressor binding and allow transcription. The plasmid also includes an ampicillin resistance gene and a bacterial origin of replication. B) The thrombin cleavage site present at the COOH-terminus of the Sj26 gene in the pGEX-2T vector is shown and the arrow indicates the site of cleavage. Stop codons are present in all three reading frames and are underlined. Three unique restriction sites present between the sequence encoding the thrombin cleavage site and the stop codons are shown.

Top agar: 1% (w/v) bactotryptone, 0.8% (w/v) NaCl, 0.8% (w/v) agar (autoclaved)
L plates consisted of L-broth containing 1.2% (w/v) agar and 10 mM $MgSO_4$ (autoclaved). LA plates also contained ampicillin at 100 µg/ml.
Minimal (M9) Medium: 48 mM Na_2HPO_4, 22 mM KH_2PO_4, 9 mM NaCl, 19 mM NH_4Cl, $MgSO_4$ 10 mM, $CaCl_2$ 1 mM, and glucose 0.4% (w/v). The $MgSO_4$ and $CaCl_2$ were prepared as 100x stocks and the glucose as a 50x stock (20%) and added just before culture to the autoclaved salt solution. For ^{15}N labelling, no NH_4Cl was included above and the $^{15}NH_4Cl$ was prepared as a filter sterilized 10% solution, stored at 4°C, and added just before culture.
Tissue culture media:
RPMI 1640 (Gibco, Paisley, UK.) medium contained 10% heat-inactivated foetal calf serum (FCS), 2 mM glutamine, 1 mM sodium pyruvate, 25 µM β-mercaptoethanol and antibiotics (50 µg/ml streptomycin sulphate, 100 µg/ml kanamycin sulphate and 50 µg penicillin G, termed SKP). RPMI medium was used for all eukaryotic cell culture except with the MDCK cell line where DMEM (Gibco) was substituted for RPMI.

2.3 General DNA methods

2.3.1 Phenol extraction and ethanol precipitation

Protein was removed from DNA solutions by extraction with an equal volume of phenol saturated with 0.1 M Tris-HCl (pH 8.0), followed by sequential extraction with phenol/chloroform (1:1) and chloroform/isoamyl alcohol (24:1, v/v). DNA was precipitated by the addition of 1/10 vol of 3 M NaOAc and 2.5 vol of ethanol, vortexing and incubating on dry ice for 30 min. The sample was centrifuged for 15 min at 4°C and the pellet was washed in 70% ethanol, dried at 37°C for 5 min and dissolved in an appropriate volume of TE.

2.3.2 Determination of DNA concentration

DNA was quantitated spectrophotometrically by measuring the Absorbance (A) of appropriately diluted samples at 260 nm. An A_{260} of 1 corresponds to 50 µg/ml for double-stranded DNA, 40 µg/ml for single-stranded DNA and 20 µg/ml for oligonucleotides (Sambrook et al., 1989). The ratio of readings at 260 and 280 nm (A_{260}/A_{280}) was used to estimate DNA purity. If the ratio was significantly less than 2 further phenol extractions were performed.

2.3.3 Restriction enzyme digestion of DNA

DNA was routinely digested with 2 units of enzyme per ug of DNA in 20 µl 1X restriction enzyme buffer. The appropriate E series buffer (Amersham International) or the appropriate low, medium, or high salt buffer (Sambrook et al., 1989), was used. Digestions were performed in 1.5 ml Eppendorf tubes at 37°C, usually for 2-4 hours, except with genomic DNA where digestion was o/n.

2.3.4 Agarose minigels

The concentration of agarose was chosen according to the size of the fragments that were to be separated (Sambrook et al., 1989) and was in the range 0.7-1.4 %. The appropriate amount of agarose was melted in 30 ml 1x TBE and allowed to cool to 50°C before the addition of ethidium bromide (final concentration 1µg/ml) and pouring of the gel. Samples (6-15 µl containing 1x glycerol dye) were run at 40-70 mA for 20-60 min.

2.3.5 Isolation of DNA restriction fragments and PCR products

The DNA was mixed with 1/6 volumes of 6x glycerol dye and loaded at approximately 1 µg per 20 mm well of a 30 ml low melting temperature agarose (Gibco, Life Technologies Ltd.) gel containing ethidium bromide and the gel was run at 50 mA until the desired band was well separated from possible contaminating fragments. The gel was viewed only under long wave ultraviolet light to minimize DNA damage. The band(s) was excised from the gel and transferred to 1.5 ml Eppendorf tubes at a maximum of 0.4 ml gel per tube and 40 µl 1 M NaCl and 200 µl TE was added and the tube incubated for 5 min at 65°C to melt the agarose before an equal volume of phenol/tris, was added. DNA was then extracted as usual.

2.3.6 Ligation of inserts into vectors

Vectors were cut with the appropriate restriction enzyme(s) and treated with calf intestinal phosphatase (Sambrook et al., 1989) to prevent vector self-ligation. Vectors were phenol/ chloroform and chloroform extracted, ethanol precipitated and resuspended at 200 ng/ml. Individual ligation reactions contained: 20 ng/ml cut vector, 1 mM ATP, 0.1 vol of 10x ligation buffer (700 mM Tris-HCl, pH 7.6, 70 mM $MgCl_2$, 80 mM DTT), 20-100 ng of insert and 1 unit of T4 DNA ligase in a total volume of 20 µl. Control ligations included vector without insert and insert without vector to allow estimation of how many colonies were derived from the ligation of insert into vector. Samples were incubated at 15°C for 3-6 hr (cohesive ends) or 6 hr to o/n (blunt ends) and 12 µl was used per transformation.

2.3.7 Preparation of competent cells

In the majority of transformations competent MC1061 cells were used due to their high level of competence and their stability on storage at -70°C. Strains other than MC1061 were prepared using a shorter technique as below. One MC1061 colony was inoculated into 20 ml TYM and grown in a 250 ml flask with shaking at 37°C until the A_{600} reached 0.2-0.8. This culture was transferred into a 2 litre flask containing 100 ml TYM, prewarmed to 37°C, and grown with vigorous shaking at 37°C to an A_{600} of 0.5-0.9. The cells were diluted to 500 ml with prewarmed TYM and grown until the A_{600} reached 0.6 and were then rapidly cooled by swirling the flask in ice water for 10 min. Cells were pelleted in pre-cooled 50 ml sterile tubes at

4°C in a Beckman benchtop centrifuge at 2000 rpm (600 g) for 10 min. Each pellet was gently resuspended in a minimal volume of supernatant, after which 10 ml ice cold transformation buffer 1 (30 mM KOAc, 50 mM $MnCl_2$, 100 mM KCl, 10 mM $CaCl_2$, 15% glycerol) was added. Cells were pooled in two 50 ml, sterile tubes, kept on ice for 30 min and then again pelleted as above. The pellets were gently resuspended in 10 ml ice cold transformation buffer 2 (10 mM MOPS-Na.H_2O, pH 7.0, 75 mM $CaCl_2$, 10 mM KCl, 15% v/v glycerol) and aliquoted into 1.5 Eppendorf tubes on dry ice. Aliquots were stored at -70°C.

Rapid protocol for production of competent *E.coli*. One colony was transferred to 10 ml of 2x TY medium and cultured o/n in a shaking incubator at 37°C. 0.5 ml of this culture was diluted to 50 ml with 2x TY and grown until A_{600}=0.5-0.6. The flask was cooled in iced water for 10 min, after which the bacteria were pelleted at 2000 rpm (600 g) at 4°C. Cells were resuspended in 12 ml ice cold 100 mM $MgCl_2$, immediately pelleted, resuspended in 12 ml ice cold 100 mM $CaCl_2$ and left on ice for 30 min before pelleting again at 4°C and gently resuspending in 2.5 ml ice cold 100 mM $CaCl_2$. The cells were kept at 4°C and used within two days.

2.3.8 Transformation of competent bacteria

For transformation with M13 phage DNA, aliquots (3 ml) of autoclaved top agar was dispensed into 15 ml sterile tubes and kept at 45°C. 35 µl IPTG (25 mg/ml in H_2O) and 35 µl X-gal (2% w/v in dimethylformamide) was added to each tube. The ligation mix was added to 0.2 ml of competent cells, mixed by flicking the tube and left on ice for 30 min. The cells were heat shocked at 42°C for 2-3 min. Freshly growing late log phase TG1 (0.2 ml) were added to the prewarmed top agar, following which the transformed cells were added. The agar mix was poured onto prewarmed (37°C) L plates. Plates were left to set and then incubated inverted o/n at 37°C.

For transformation of ampicillin resistant plasmids the ligation products were gently mixed with 100 µl competent cells. Cells were incubated on ice for 30 min and were then heat shocked at 37°C for 5 min with MC1061 or 42°C for 2-3 min with other strains. 200 µl 2X TY (prewarmed to 37°C) was added and the cells were then incubated for 30-60 min at 37°C. Transformed cells (100-200 µl) were plated on LA plates and grown o/n at 37°C. With the CDM8 vector, competent MC1061 cells containing the P3 plasmid (Seed and Aruffo, 1987) were used and selection was on plates containing 12 µg/ml tetracycline and 7.5 µg/ml ampicillin.

2.3.9 Transfer of bacterial colonies to nitrocellulose filters

Bacterial colonies were either toothpicked onto a pre-labelled nitrocellulose filter (Hybond C, Amersham) on an LA plate and grown on the filter until just visible (4-8 hr) at 37°C, or the nitrocellulose filter was lowered onto a plate of o/n colonies, the orientation marked with ink

and then it was carefully lifted off with colonies attached (and the master plate was returned to 37°C for 8 hr). Sheets of Whatman 3MM paper were placed into three trays and soaked with (A) 10% SDS, (B) 0.5 M NaOH, (C) 1.0 M Tris-HCl pH 7.4. Nitrocellulose filters were then placed, colony side up, for 4 min on A, B, and C successively. Filters were wiped once with tissues soaked in 2x SSC, 0.1% SDS and washed in 2x SSC with rocking for 15 min. Filters were allowed to dry on 3MM paper and were baked at 80°C for 2 hr.

2.3.10 ^{32}P labelling of cDNA and oligonucleotides

Random hexanucleotide labelling of cDNA probes and DNA fragments greater than 100 bp was with the Multi-Prime DNA labelling kit (Amersham) as recommended by the manufacturer. Unincorporated α-^{32}P-dATP was removed by adding 20 µg herring sperm DNA , 20 µl 3 M NaOAc and 220 µl isopropanol vortexing and spinning in a microfuge at RT for 10 min. Incorporation levels were calculated by comparing the counts in 2 µl prior to precipitation with 2 µl after resuspension. The pellet was resuspended in 200 µl TE, 1% SDS and stored at -20°C or used directly. For labelling oligonucleotides, 400 ng of oligonucleotide was mixed with 5µl 10x kinase buffer (1 M tris pH 8, 100 mM $MgCl_2$, 70 mM DTT), 1-3 µl γ-^{32}P-ATP and 10-20 units of T4 polynucleotide kinase in a total volume of 50 µl and incubated for 30-60 min at 37°C. The labelled oligonucleotide was separated from unincorporated γ-^{32}P-ATP by size exclusion chromatography using Sephadex G50 fine (Pharmacia). A 5 ml plastic pipette was blocked with glass wool and 2 ml of G50 added and washed with TE. The oligonucleotide mixture was made up to 200 µl and loaded onto the column and 5 drop fractions were collected. The fractions were counted and those corresponding to the early peak were pooled.

2.3.11 Screening of nitrocellulose filters with ^{32}P labelled probes

Nitrocellulose filters to be probed with oligonucleotides were prehybridized in a minimal volume of 6xSSC containing 0.1 mg/ml denatured herring sperm DNA, for approximately one hour. Prehybridization and hybridization was routinely performed at 42°C for 17 mer oligonucleotides. Labelled oligonucleotide was added to the prehybridization solution to approximately $5x10^5$ cpm/ml and allowed to hybridize for 2-4 hours. Filters were washed twice in 6xSSC at RT for 15 min, and the filter monitored with a hand held coulter counter. If the background was high, washing was continued with 6xSSC containing 0.1% SDS at RT. Filters were rinsed in 6xSSC, dried on Whatman paper and autoradiographed at -70°C for 1-24 hr.

Filters to be probed with cDNA probes were pre-hybridized in 50% formamide, 2x Denhardt's solution, 0.2% SDS, 5xSSC and 150 µg/ml herring sperm DNA at 42°C for 1-2 hr. The probe was denatured by boiling for 3 min and then placing on ice. Probe was added to $5x10^5$ cpm/ml and hybridization allowed to occur o/n. Washing was at 65°C with 2xSSC, 1xSSC and 0.2xSSC plus 0.1% SDS for approximately 20 min each time. Excess buffer was removed and

the filters wrapped in clingfilm (to prevent drying out) and exposed at -70°C for 1-24 hr. Hybridization and washing was performed in a Hybaid hybridization oven using the recommended volumes of solution in the vials. Wash solutions were pre-heated in water baths before use. When membranes were to be re-probed they were stripped of existing probe by repeated washing in boiling 0.1% SDS. Herring sperm DNA and Denhardt's solution were prepared as in Sambrook et al., (1989).

2.3.12 Southern blotting

Genomic, phage or plasmid DNA was digested to completion with the appropriate restriction enzymes, subjected to electrophoresis through horizontal slab gels and transferred to nitrocellulose filters (Hybond N, Amersham) by standard procedures. The transfer was either by capillary action o/n using stacks of paper towels (Sambrook et al., 1989) or using a Stratagene Pressure Blot apparatus. Pressure blotting was at 75 mmHg for 30 min with minigels and 60 min with large gels. The probes used with the mouse genomic Southern blot were the region of the partial cDNA clone corresponding to the coding region, and the 1.7kb Eco RI-Bam HI fragment of GL11a (Fig 4.4c). Those used with the rat genomic Southern blot were the 350 bp 5' Eco RI-Sca I fragment and the 730 bp 3' Acc I-Bam HI fragment of pLSGP 10. Probes used in analysis of transgenic mice are described in section 2.9.

2.3.13 Plasmid preparations

Minipreparations of plasmid DNA and M13 replicative forms were obtained by the alkline lysis method (Sambrook et al., 1989). Large scale preparations were by the CsCl-ethidium bromide equilibrium centrifugation method as follows. A 300ml o/n culture was pelleted at 7000 rpm (6000 g) in a Beckman JA14 rotor and the cells resuspended in the JA14 tube in 15 ml of freshly prepared 0.2 M NaOH, 1% SDS, mixed by vigorously inverting the capped tube and placed on ice for 5 min. 9 ml of ice cold 3 M potassium acetate (pH 4.8 prepared by mixing 70 ml 1M KOAc with 30 ml glacial acetic acid) was added with mixing and incubation continued for 20 min before centrifugtion at 10 000 rpm (8 600 g) at 4°C. The s/n was poured through two layers of muslin into a clean JA14 tube and 15 ml of isopropanol added followed by centrifugation at RT for 20 min at 10 000 rpm. The well drained pellet was resuspended in 4.5 ml of TE using a pasteur pipette, mixed with 4.5 g CsCl and incubated at 37°C for 10 min before addition of 0.3ml of ethidium bromide (10 mg/ml). After standing at room temperature for 10 min if significant preciptiate appeared the mixture was centrifuged at 3000 rpm for 5 min. The soluble CsCl-DNA mixture was transferred to quick seal ultracentruge tubes and centrifuged in a TL100 rotor at 85 000 rpm (400 000 g) for 15 hr or in a VTi80 rotor at 80 000 rpm (510 000 g) for 3 hr. The lower band was isolated from the tube using a 1 ml sterile syringe and the ethidium bromide removed by repeated extraction with water saturated butanol.

2.4 Nucleotide sequencing

2.4.1 Preparation of single-stranded DNA templates for nucleotide sequencing

An o/n culture of TG1 was diluted 1:100 in 2x TY and dispensed into 1.5 ml aliquots in sterile tubes. Each aliquot was inoculated with a single colourless M13 plaque using a sterile wide bore Pasteur pipette, and incubated for 5 hr at 37°C with vigorous agitation. Cultures were transferred to 1.5 ml Eppendorf tubes and bacteria were pelleted by centrifugation for 5 min. This step was repeated and the resulting cell-free supernatants (1ml) were transferred to Eppendorf tubes containing 0.2 ml 20% (w/v) PEG-6000 in 2.5 M NaCl, mixed well and incubated at RT for 15 min. After the incubation period, phage particles were pelleted by centrifugation for 5 min and the supernatant was aspirated with a drawn-out Pasteur pipette. Samples were centrifuged again and the remaining traces of PEG were aspirated. 100 µl TE and 75 µl Tris-saturated phenol was added to each tube. The suspension was vortexed (15 seconds), incubated at RT for 15 min, vortexed again and then centrifuged for 3 min. 75 µl of the aqueous phase was removed and transferred to a new tube and the DNA was ethanol precipitated, washed with 70% ethanol and resuspended in 35 µl 10 mM Tris-HCl pH 8, 0.1 mM EDTA. The cell pellet with remaining supernatant was kept at 4°C as a phage stock.

2.4.2 Sequencing reactions

Single-stranded M13 DNA templates were sequenced by the dideoxy chain-termination method of Sanger, using a Sequenase Version 2 kit (USB) and T7 DNA polymerase. Reactions were carried out according to the Sequenase Version 2 protocol. In cases where a sequence was unclear due to apparent GC compressions, dITP labelling mixture was substituted for the dGTP mixture. For sequencing of double stranded templates 5-10 µg of CsCl purified plasmid DNA was used. Denaturation was by incubating in 0.5 ml Eppendorf tubes the DNA in 8 µl with 2 µl of 2 M NaOH for 10 min at RT followed by addition of 3 µl 3 M NaOAc (pH 4.5) and 7 µl dH_20. 60 µl ethanol was added the mixture vortexed and placed on dry ice for 15 min before centrifugation and washing with 70% ethanol. The dry pellet was resuspended in 7 µl of dH_2O and used as template in reactions identical to those for single stranded DNA.
PCR products were sequenced by a modification of the protocol of Winship (1989). 20 pmol of oligonucleotide (approximately 2 µl of a 40 µg/ml stock) and 100-200 ng of template was used. The sequenase reagents were used and DMSO was included in the annealing and labelling reactions (but not the termination reactions) at a final concentration of 10 %. The annealing mixture was placed in a boiling water bath for 3 min and immediately transferred to dry ice where the tubes were left for upto 15 min. The annealed template was then thawed and used in the labelling step as described in the Sequenase protocol.

2.4.3 Sequencing gels

The products of the sequencing reactions were separated on 6% acrylamide/ bisacrylamide

(19:1), 8 M urea, 1x TBE gels. The electrophoresis buffer was 1x TBE and the bottom chamber contained 1x TBE, 1 M KOAc or simply 1x TBE when gels were run for longer time periods. Gels were run at a constant power of 37 Watts and were fixed in 10% ethanol/10% acetic acid before drying onto Whatman 3MM paper using a BIO-RAD gel dryer. Kodak-X-Omat films were exposed to the gels at RT for 16-48 hours.

2.5 *In vitro* mutagenesis

The oligonucleotide directed *in vitro* Mutagenesis System Version 2 of Amersham was used and the protocol of Amersham followed. Single stranded template for mutagenesis was prepared as in section 2.4.2 with the following differences. 5 ml 2x TY medium in a 50 ml tube was inoculated with 50 µl of a fresh o/n culture of TG1 cells. A phage plaque was transferred into the medium with a sterile Pasteur pipette and the mixture was cultured with shaking at 37°C for 5 hr. The culture was aliquoted into five 1.5 ml Eppendorf tubes and single-stranded DNA was purified from the phage. After the 70% ethanol wash, however, the dried DNA pellets of all five tubes were resuspended in a total volume of 100 µl TE. The DNA solution was centrifuged for 5 min to remove any remaining protein precipitate and the supernatant was transferred to a clean tube. Mutagenesis of the mouse leukosialin gene (fragment B1c) to introduce a 5' Sal I site and delete the internal Xho I site was performed with two oligonucleotides, gctccacccagtcgacccctgcat and catctcactggagggcc, simultaneously. The second oligonucleotide changed only a single nucleotide in the Xho I site without changing the encoded amino acid. The success of the mutagenesis was tested by treating purified replicative form DNA (Sambrook et al., 1989) with Sal I and Xho I to see gain and loss of cleavage products, respectively. Mutated inserts were then sequenced to confirm the presence of both changes and to ensure no mismutations were present elsewhere in the open reading frame.

2.6 Polymerase chain reaction and Bal 31 deletion

2.6.1 PCR conditions

Template DNA (50-1000ng) in 1-10 µl was mixed with 0.5 µg of each primer, 0.1 vol of 10 mM dNTPs, 0.5 vol 2x PCR buffer (50 mM KCl, 15 mM Tris-HCl pH 9.3, 2 mM $MgCl_2$, 0.05% gelatin) and 0.4 µl Taq polymerase per 100 µl reaction (Cetus Corporation). Reactions were prepared in 0.5 ml Eppendorf tubes and were overlayed with 1-2 drops of mineral oil. Reactions were carried out in a Perkin Elmer Cetus DNA thermal cycler and unless stated otherwise conditions were: denaturation (1 min, 94°C), annealing (2 min, 60°C) and elongation (1 min, 72°C), for a total of 20-30 cycles. Colony PCR involved the same conditions except the template was DNA spontaneously released from an *E. coli* colony, containing the plasmid of interest, that was transferred with a toothpick to a tube containing PCR mix. The toothpick was also spotted onto a master plate before transfer to the mix. 20 µl reactions were used for analytic purposes and 100 µl reactions for preparative samples.

2.6.2 Isolation of DNA fragments by PCR for expression in the pGEX-2T vector

DNA encoding domains or parts of domains was isolated by PCR using the oligonucleotides indicated in the results. The oligonucleotides were designed to yield a fusion protein in-frame with GST in the pGEX-2T vector and to have upto 6 nucleotides overhang on the 5' end and restriction enzyme sites plus 17-20 nucleotides of complementary sequence. The restriction sites were 5' Bam HI, 3' Eco RI except with LFA3.D1 where an internal Eco RI site necessitated the use of a 3' Sma I site. Single stranded DNA was used preferentially as template, and PCR products were phenol extracted, ethanol precipitated, treated with Bam HI and Eco RI (or Sma I) for 3-4 hr, the restriction enzymes were heat inactivated and the fragments ligated into the pGEX-2T expression vector which had been treated with the same enzymes.

For expression of the octapeptide PPVTITNP as a fusion protein two complementary oligonucleotides were made which included 5' and 3' overhangs that would generate open Bam HI and Eco RI sites following annealing: sense, gatcccctcctgtcactataactaatccttg; anti-sense, aattcaaggattagttatagtgacaggaggg. The oligonucleotides were treated with T4 polynucleotide kinase as described by Mallet et al.(1990). They were then ligated into the pGEX-2T vector which had been treated with Bam HI and Eco RI, annealing occuring during the 4 hour ligation reaction at 16°C.

Oligonucleotides specific for pGEX-2T 5' and 3' of the multiple cloning site were designed and used for analytical PCR and sequencing, 5'side: gcaagccacgtttggtg, 3' side: ttcaccgtcatcaccga. pGEX-2T clones were screened for fusion protein expression as described in section 2.12 and if a product of the expected size was apparent plasmid DNA was isolated by the CsCl method and the sequence of the insert confirmed. In some cases, where the fusion proteins were being screened for solubility, the inserts were not sequenced but a minimum of 2 independent clones and usually 3 or 4 were tested.

2.6.3 Generation of clones by Bal 31 deletion

The pGEX-2T plasmid containing the rat leukosialin extracellular domain sequence was linearized with Eco RI. After an initial trial to access the activity of the Bal 31 preparation, 30 µg was treated with 12 µl of Bal 31 in a total of 600 µl (600 mM NaCl, 12.5 mM $CaCl_2$, 20 mM Tris pH 8, 1 mM EDTA) at 30°C with 100 µl aliquots taken at 1, 2, 3.5, 5, 7.5 and 10 minutes and mixed with 10 µl 0.5 M EDTA to stop the reaction. The samples were phenol extracted, ethanol precipitated, and resuspended in 200 µl of 10 mM Tris-HCl pH 8, 100 mM

NaCl, 7 mM $MgCl_2$, 7 mM 2-mercaptoethanol buffer with 0.1 mM dNTPs and 10 units of Klenow. After 30 min at 37°C the DNA was re-extracted and precipitated before being taken up in 100 µl of the above buffer and treated with Bam HI at 37°C for 2 hours. 7 µl samples were analysed for size by agarose gel electrophoresis, and the fragments were then purified by gel fractionation. The purified fragments were ligated into the pGEX-2T vector that had been cut with Bam HI and Sma I and treated with phosphatase. Resultant clones were grown and analysed by Western blotting and PCR. The pGEX-2T vector has stop codons present in all 3 reading frames 4, 6, or 7 codons past the Sma I restriction enzyme site. Thus, depending on the digestion end point, expressed fusion proteins could have the residues XGIHRD, GNSS, or XEFIVTD at the C-terminus where X is variable. To identify constructs with inserts of particular size, colonies corresponding to the appropriate digestion time point were picked and analysed by PCR using oligonucleotides specific for vector sequences 5' and 3' of the cloning site. The insert size was determined by agarose gel electophoresis. The end points of the clones of interest were determined by sequencing the PCR products (section 2.4.2).

2.7 Plating out EMBL-3 genomic library

2.7.1 Isolation and characterization of genomic clones

A DBA/2J mouse genomic library was obtained from CLONTECH laboratories, Inc. (California). The library had been generated by partial Mbo I digestion of adult liver DNA and cloning into the EMBL-3 vector system, such that the average insert size was 15 kb, and the complexity was $1.33x10^6$. The EMBL-3 genomic library was screened at a plaque density of $4x10^5$ pfu/plate using the protocol supplied by CLONTECH laboratories which was briefly as follows. K802 cells competent for infection by phage were prepared by growing o/n in L broth supplemented with 0.2% maltose and 10 mM $MgSO_4$. 100 µl of the appropriate library dilution was mixed with 4.5 ml of competent bacteria in a 50 ml tube and incubated at 37°C for 30 min. 30 ml of molten (50°C) top agarose (0.7% agarose, 0.2% maltose and 10mM $MgSO_4$ in L broth) was added and the mixture poured over a pre-dried large square L plate (220x220mm). The agarose was allowed to set and the plates incubated at 37°C for 12-16 hr. The plates were placed at 4°C for 15-60 min and then nitrocellulose cut to size was carefully lowered onto the plates and orientation marked with a needle dabbed in drawing ink. After 1 min the filters were peeled off and floated, plaque side up, on 0.5 M NaOH, 1.5 M NaCl for 30 sec and then immersed and left 3 min. The filters were transferred to 0.5 M Tris-HCl pH 7.5, 1.5 M NaCl for 5 min and then rinsed in 2x SSC and dried. A second lift of each plate was made in the same manner except the filter was left for 3 min before lifting off the plate. Dried filters were sandwiched between 3MM paper and baked for 2 hr at 80°C before hybridization.

The probe was a 460bp Sty I - Bam HI fragment of the murine cDNA corresponding to the coding region. After screening a single complexity, one clone (GL1p) was isolated. DNA was

isolated as described below and digested with restriction enzymes. Fragments of the 18 kb insert which were shown to be reactive by Southern blotting were subcloned into pTZ19R and smaller fragments were subsequently cloned into M13 and sequenced. This analysis indicated GL1p was not a functional gene. Further screening of the library at the level of two complexities lead to the isolation of 11 more clones. These were then screened with an oligonucleotide (17mer) specific for the murine cDNA (coding region). Only one of the 11 clones (termed GL11a) reacted. Fragments of the insert which were shown to bind to the oligonucleotide or cDNA probe by Southern blotting were subcloned and sequenced. The GL11a clone was sequenced in both directions from a position 580 nucleotides 5' to the coding region to 150 nucleotides on the 3' side. Sequencing of GL1p was over a similar region, but only one strand was sequenced completely.

2.7.2 Isolation of lambda DNA

A plaque was picked into 1 ml SM buffer using a flamed Pasteur pipette and the tube rotated at 4°C for 4 hr. 10 ml competent K803 cells were pelleted, resuspended in 700 µl 10 mM Tris-HCl pH 7.5, 10 mM $MgSO_4$ and 750 µl of resuspended phage added. After 30 min at 37°C, 7.5 ml of top agarose (50°C) was added and the mixture poured over a large round L plate and incubated o/n at 37°C. 10 ml SM was poured on the plate, rocked at 4°C for 4 hr and poured into a JA20 tube with 10 µl DNase I (1 mg/ml) and 10 µl RNase (10 mg/ml) with incubation at 37°C for 30 min. Addition of 10 ml 20% PEG, 2 M NaCl in SM was followed by 1 hr on ice and centrigugation at 12 000 rpm for 20 min. The pellet was resuspended in 0.5 ml SM in an Eppendorf tube, spun for 5 min to remove debris, incubated with 5 µl 10% SDS and 5 µl 0.5 M EDTA (pH 8) at 68°C for 15 min and then phenol extracted as usual. Precipitation was with 50 µl 3 M NaOAc (pH 5.2) and 550 µl isopropanol.

2.7.3 Computer programmes

Sequences were analyzed by the computer programmes of Staden (Staden, 1987). Alignments of sequences were with the IALIGN program (Dayhoff et al., 1983). The NBRF database (Protein Identification Resource, National Biomedical Research Foundation, Georgetown, USA) was searched using the FASTP program (Lipman and Pearson, 1985).

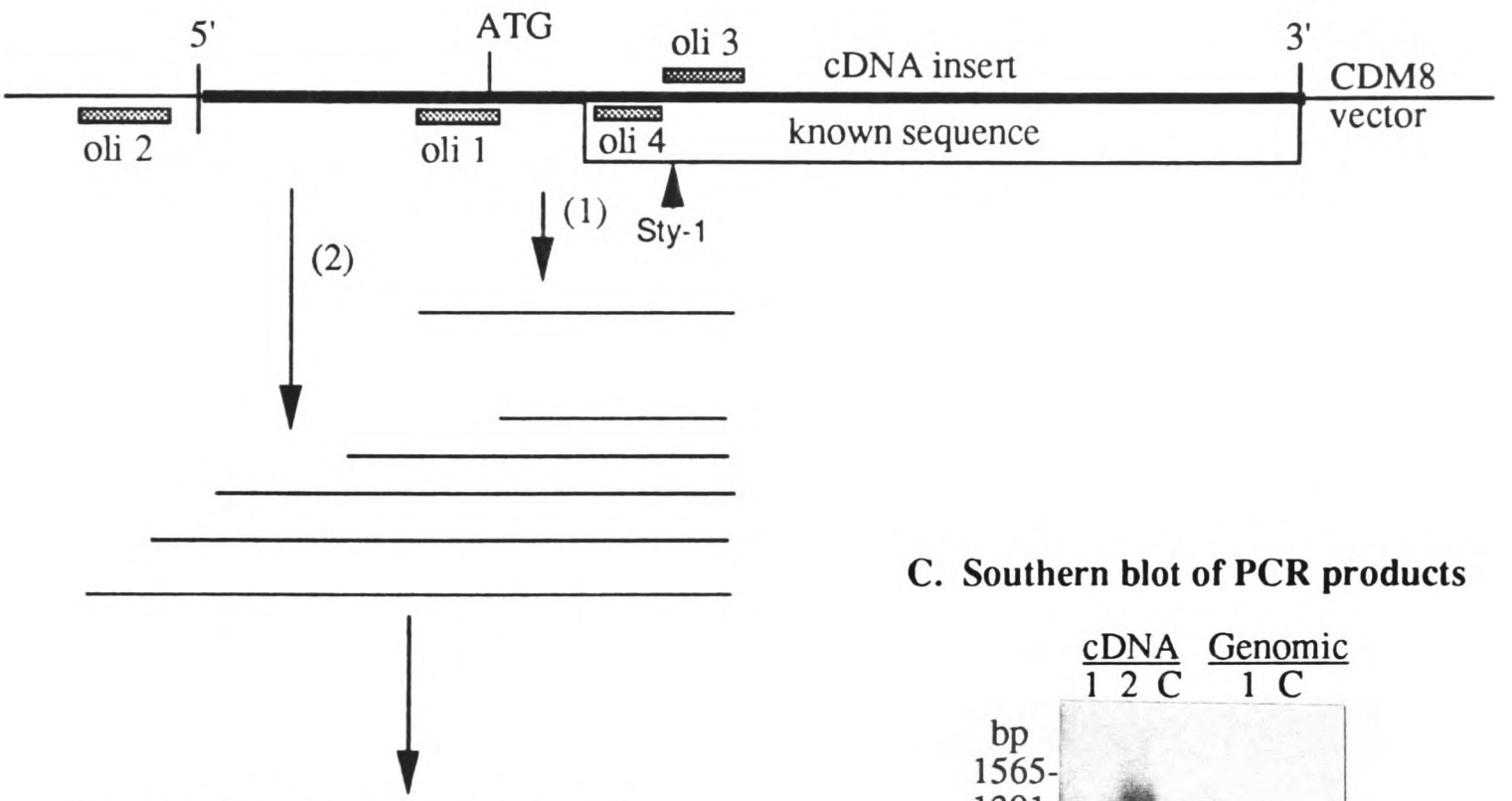
2.8 Preparation of a full length rat leukosialin cDNA clone

The cDNA for rat leukosialin lacked the complete leader sequence (Killeen et al., 1987). Alignment of mouse and human leukosialin nucleotide sequences in a region immediately 5' to the initiation ATG codon revealed a patch with 16/21 identities. Thus an oligonucleotide was designed with a Sal I site plus 17 nucleotides on the 5' side of the ATG (oligo 1, Fig 2.2). On the 3' side an oligonucleotide was made including rat sequence from a region immediately 3' to the unique Sty I site in the coding region (oligo 3, Fig 2.2). These oligonucleotides were used

A. Oligonucleotides

oli 1	5'	c gtcgac ctgtccctctggagatgg Sal I
oli 2	5'	c gaattc gactcactatagggag Eco RI
oli 3	3'	agaagacaattctggagttccc
oli 4	internal	tggcaacatcgtcatcg

B. Strategies



C. Southern blot of PCR products

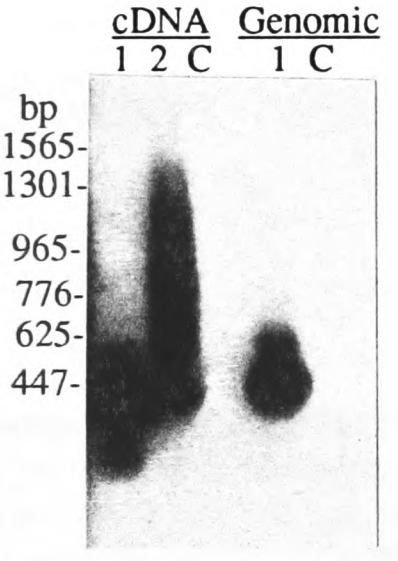


Figure 2.2 Isolation of rat leukosialin leader sequence and 5' untranslated region. A) The oligonucleotides were designed as described in the text and used either in strategy (1) or (2) as illustrated in (B). The PCR conditions were: denaturation 94°C, 60 sec, annealing 60°C 60 sec, and extension 72°C 2 min, for 30 cycles with 1 μg of rat genomic DNA or Con A T blast library in CDM8 plasmid DNA (kindly provided by Susan Mallett) as template and 0.5 μg of oligonucleotides 1 and 3 or 2 and 3 in 100 μl of reaction mixture. C) 5 μl of PCR products from strategy (1) or (2) using the cDNA library (cDNA) or genomic DNA as template, or from control (C) reactions to which no template was added, were run on a 1% agarose gel, transferred to nitrocellulose and probed with 32P-end labelled oligonucleotide 4. Exposure was for 2 hr at -70°C.

A.

```
ggaggaagag cctctggtcg gcagtgagga cgctgcccgg agccatccgt ctccggcgct -130

ccggagctct tgttgaccct ctgcttcctg gctctggatt caaggtgagg cagtggtctc -70

aagctccagc tgtcaagact cccaccacag tggacctagc tgtcccaact cctgctcccg -10
                                            g--------- --------t-

gctgtcact  -1
c-----c--

ATGgagATGgccttgcatcttctcctcctctttgggggcttctgggcccaggtggtgagccaa
(M)(E) M  A  L  H  L  L  L  L  F  G  G  F  W  A  Q  V  V  S  Q
```

B.

```
Oligonucleotide 1  cGTCGACctgtccctctggagatgg
rat leukosialin    ctcccggctgtcactatggagatgg
```

C.

```
R (M)(E)MALHLLLLFGGFWAQVVSQ
M          ----------AC-V--A-P
H          ---T-----L-   -VL---PDALG
```

Figure 2.3 Rat leukosialin leader sequence and putative sequence of rat leukosialin 5' untranslated region. A) The nucleotide sequence for the leader peptide and the encoded amino acids preceded by the 5' untranslated sequence. Differences between rat and mouse leukosialin over the first 30 nucleotides 5' of the ATG are shown; further 5' there is no significant homology with the mouse or human sequences. M and E are shown in parentheses because it is not clear whether translation would begin at this or the succeeding ATG in the rat sequence. B) Alignment of the oligonucleotide used in strategy 1 (Fig 2.2) and the corresponding rat leukosialin sequence determined by strategy 2. Capital letters indicate the Sal I site in the oligonucleotide. Nucleotide differences outside this site are underlined. C) Alignment of the rat leukosialin leader sequence with mouse and human leukosialin leader sequences. The end of the leader sequence, as determined from the mature NH_2-terminus of the protein, is known for rat and human leukosialin (Killeen et al., 1987; Schmid et al., 1992).

in PCR with 1 μg rat genomic DNA or a T blast cDNA library (Mallett et al., 1990) as template in 100 μl reactions to yield products of the expected size that reacted with an oligonucleotide derived from known rat sequence 5' of the Sty I site (oligo 4, Fig 2.2). The fragment was gel purified and treated with Sal I and Sty I. This fragment was used to replace the incomplete region in a derivative of the original clone for rat leukosialin (pLSGP-2 (Killeen et al., 1987)) which extended to the Hind III site 200 bp 3' of the stop codon (pLSGP-H, it was necessary to remove the rest of the 3' untranslated sequence present in pLSGP-2 due to the presence of endogenous Sty I sites) and this product was used for expression purposes.

A second strategy was employed where an oligonucleotide was designed specific for vector (CDM8) sequence 5' of the cloning site (oligo 2, Fig 2.2) and this was used in PCR with the same 3' oligonucleotide as above, and with the rat T blast cDNA library in CDM8 as template. Probing the PCR products showed that a series of products were produced (Fig 2.2) their length probably depending on the point at which reverse transcription during copy DNA synthesis ceased. Regions corresonding to intermediate and high molecular weight DNA fragments were gel fractionated and sublconed into M13. Two clones with 500 bp inserts were sequenced and both were identical (Fig 2.3). The differences between the rat nucleotide sequence 5' of the ATG and the oligonucleotide used for stategy 1 above are shown (Fig 2.3) and it is seen that 16/18 3' nucleotides are identical, explaining the efficiency of the PCR with this oligonucleotide. One of the differences is such that an additional ATG is present in the rat sequence. The context of this ATG should make it as strong an initiation site as the conserved ATG (Kozak, 1987) and thus translation is likely to initiate two residues earlier in rat leukosialin. The 30 nucleotides 5' of the inititation site are strongly conserved between mouse, human and rat (Fig 2.3). 5' of where the splice junction exists in human leukosialin (and probably exists in mouse leukosialin, Chapter 3) the homology of rat leukosialin is lost and the sequence is not similar to any of the available mouse or human 5' untranslated sequence . The most 5' region of the rat untranslated sequence is also unusual because the first 30 nucleotides are identical to 30 nucleotides in the rat leukosialin coding region (nucleotides 1080-1110 with respect to the initiation ATG) and the 45 nucleotides that follow these are identical to the anti-sense of the region 983-939. The significance of this observation and whether it is due to a PCR or cloning artefact will require further analysis.

2.9 Analysis of transgenic mouse DNA

2.9.1 Isolation of genomic DNA

Tail snips (approx. 1 cm) were taken from 10 day old mice, placed in 0.7 ml tail buffer (50 mM Tris-HCl pH 8, 100 mM EDTA, 100 mM NaCl and 1% SDS) with 25 μl proteinase K (10 mg/ml) and incubated o/n at 55°C. 10 μl RNAse A (10 mg/ml) was added and incubation continued at 37°C for 1 hr followed by tris/phenol, phenol/chloroform and chloroform

extraction and the purified DNA was precipitated from the aqueous phase by addition of an equal volume of isopropanol. The visible DNA precipitate was immediately removed with a flame-sealed pasteur pipette, shaken briefly in 1 ml 70% ethanol and then the tip with attached DNA was broken off inside an eppendorf tube containing 100 µl dH_2O.

2.9.2 PCR dectection of transgene

1 µl of the tail DNA preparation was used in a 20 µl reaction. The oligonucleotides were: 5' oligonuceotide specific for mouse leukosialin cytoplasmic domain, acgaaatggggtggtgg; 5' oligonucleotide specific for rat leukosialin cytoplasmic domain, cgccagaggcagaagcgc; and a common 3' oligonucleotide specific for human β-globin exon 2 sequence derived from the pHSE3' vector, tagactcaccctgaagttctcag. PCR conditions were as in section 2.6. for 30 cycles and 10 µl was analyzed on a 1% agarose gel as shown in Fig 2.4a. Product sizes expected were 400 bp for mouse transgenics and 580 bp for rat transgenics. In most cases the result was definitive and in no case was a clear PCR positive found to be negative in Southern blotting. If the negative control showed a distinct product the PCR was repeated. Screening for homozygosity required Southern blotting.

2.9.3 Southern and Northern blotting of transgenic mouse DNA and RNA

7 µg of tail DNA was digested with 20-30 units of Pst I in a 100 µl volume o/n. Digestion efficiency was tested with a minigel and if complete the DNA was ethanol precipitated and resuspended in 15 µl TE/1x glycerol dye and loaded onto a 0.8% 60 ml agarose gel (135x108 mm) and run at 50-60 mA for 4 hr. Transfer to nitrocellulose was usually with the pressure blotting apparatus (Stratagene) for 1 hr. PCR products containing most of the open reading frame of rat and mouse leukosialin were used as probes. When detecting the rat transgene both rat and mouse probes were included to expose the endogenous leukosialin gene and so allow loading and digestion efficiency to be compared between samples. Some examples are shown in Fig 2.4b. Expected product sizes from Pst I digestion of the mouse transgene were 2.4, 1.3 and 0.3 kb, and a single product of 3.7 kb from the rat transgene. The endogenous mouse gene produced fragments of 7, 1.5 and 0.3 kb. The 0.3 kb fragments are not evident in Fig 2.4b. The M7 line showed an additional band at 2.8 kb which may have resulted from aberrant insertion of one transgene copy. Homozygosity was confirmed by backcrossing to a non-transgenic and performing PCR analysis on the progeny.

The Northern blot presented in Fig 4.9 was performed as described in Sambrook et al. (1989) with the details as indicated in the figure legend. RNA was prepared from freshly isolated tissues by grinding them in a minimal volume (0.5-1 ml) of lysis buffer (18.92 g of guanidinium thiocyanate and 1 ml of 1 M sodium citrate in 40 ml). 0.5 ml samples were transferred to Eppendorf tubes and 50 µl 2 M NaOAc (pH 4) and 500 µl Tris/Phenol added,

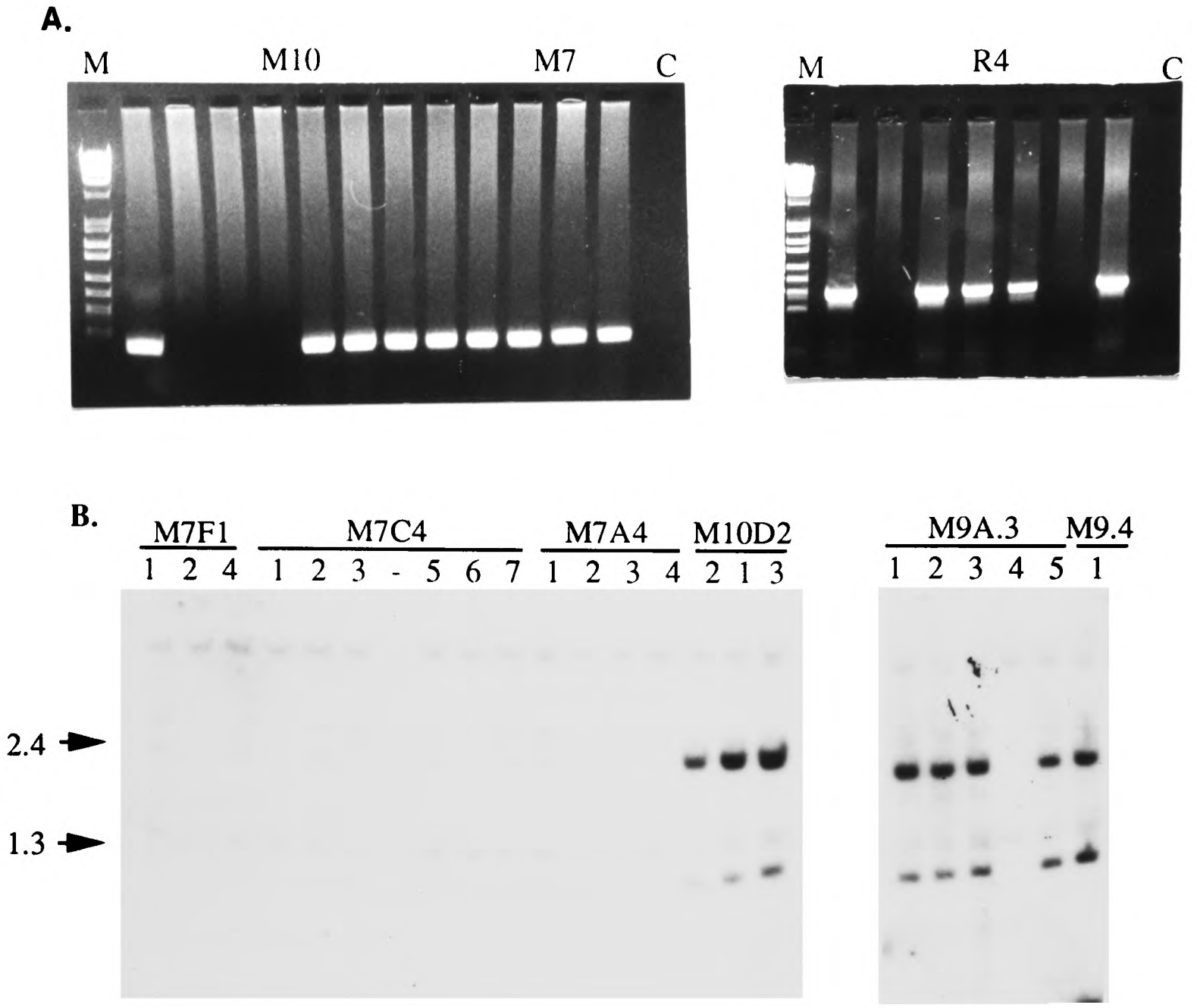


Figure 2.4 PCR and Southern blot detection of transgene transmission. A) PCR was as described in the text with 0.5-1 μg of tail DNA from litters of mice of the indicated lines. 10 μl samples were loaded on 1% agarose gels and each lane corresponds to a different mouse. C indicates control PCR reactions to which no template was added and M marker fragments. The size of the 4 smallest markers is 965, 776, 625 and 447 base pairs. B) Tail DNA samples from litters of the indicated lines of trangenic mice (pairs of heterozygotes) were digested with Pst I, separated on 0.8 % agarose gels, transferred to nitrocellulose and probed with a labelled PCR product corresponding to most of the coding region. The blot was exposed at -70°C for 30 hr and the position of two expected products at 2.4 and 1.3 kb is indicated. M7A4.2 is a non-transgenic and lacks these bands. The origin of the band at approximately 2.8 kb in the M7 line is not understood. Potential homozygotes include M7F1.4, M7C4.7 and M7A4.1.

followed by 100 µl chloroform/isoamyl alcohol. Tubes were shaken vigorously and placed on ice 10 min. The preparation was microfuged 15 min at 4°C and the RNA precipitated from the aqueous phase with 500 µl isopropanol, and incubation at -20°C for 1 hr prior to pelleting. The pellet was dissolved in 100 µl 0.5% SDS at 65°C for 10 min.

2.10 Transfection, analysis and storage of eukaryotic cell lines

The selective agent used in all transfections was neomycin (G418 -phosphate, Gibco, Life Technologies Ltd.) which acts by inhibiting protein synthesis. Resistance to this antibiotic was conferred by the neomycin resistance gene, which encodes an aminoglycoside phosphotransferase, driven by the thymidine kinase promoter on the pKG5 plasmid. Before transfection the cells were tested for their level of natural resistance to G418 and a concentration was chosen which lead to complete death of untransfected cells by day 4 or 5. This was 0.8 mg/ml in all cases except with Jurkat cells were 1.0 mg/ml was previously determined to be optimal (He et al., 1988).

2.10.1 Electroporation

NSO, Jurkat E6-1 and K562 cells were transfected by electroporation (He et al., 1988). Cells in the exponential phase of growth were pelleted and resuspended in PBS at a concentration of 1.0x 10^7/ml and kept on ice for 10 min. The electroporation cuvette was precooled on ice and 0.5 ml of the cell suspension was transferred to it with the plasmid DNA, which had been sterilized by extraction with an equal volume of ether, and the mixture was mixed by gently pipetting and kept on ice for 5 min. The cuvette was then pulsed at 0.52 kV and 25 uF using the Gene Pulser apparatus (Biorad, Richmond, CA) and placed on ice for 10 min. The cell suspension was transferred into a 25 ml tissue culture flask containing 1 ml RPMI/10 % FCS at RT, and incubated for 10 min. 7.5 ml of complete medium was then added to the flask and it was gassed with 5% CO_2 and cultured at 37°C. After 48 hr of growth, cells were selected for neo expression by culturing in medium containing G418 at 0.8 mg/ml. The medium was changed daily for four or five days and then on alternate days. In the case of the NSO cells, thymocyte feeders prepared by exposure of freshly isolated rat thymocytes to 2000 rad of γ-irradiation from a ^{137}Cs source (Gamma Cell 40, Atomic Energy of Canada, Ltd.), were added to rescue small numbers of drug resistant cells. Cells were analyzed for expression of the transfected gene by FACSCANing as detailed in Chapter 4.

2.10.2 Calcium phosphate mediated transfection

Y3 and MDCK cells were transfected by the calcium phosphate technique (Sambrook et al., 1989). 1.5x 10^7 Y3 cells were placed in 80 cm^2 flasks in 50 ml of fresh medium, gassed with 5% CO_2 and placed at 37°C o/n. 2x HBS buffer (280 mM NaCl, 10 mM KCl, 1.5 mM $Na_2HPO_4.2H_2O$ and 50 mM Hepes) was prepared and the pH adjusted carefully to 7.05. DNA

sterilized by ether extraction was mixed with 186 µl 1M $CaCl_2$ and 540 µl sterile H_2O was added. 750 µl 2x HBS was put in a 15 ml tube and a 1 ml disposable pipette was inserted and air bubbled through slowly while the DNA mixture was added dropwise using a pasteur pipette over about 1 min. The bubbling was continued for a further 15 seconds and the mixture left at RT for 30-40 min before transferring into the medium of the o/n Y3 cultures, gassed and incubated at 37°C. Some cultures were glycerol shocked after 4 hr. The cells were washed in the flasks by gently pouring on and off 12 ml fresh RPMI. 2 ml of 15% glycerol in 1x HBS was added and the flasks incubated at 37°C for 2 min. The cells were washed 2 further times and 40 ml of fresh medium was added, the cells gassed and returned to the incubator. After 24 hr the cells were removed from the flask by vigorous knocking, washed and resuspended in 50 ml medium containing G418 at 0.8 mg/ml. The glycerol shock step was not observed to increase expression. Cells were analyzed for expression of the transfected gene at various time points as detailed in Chapter 4.

Polarized epithelial cells do not efficiently take up material by endocytosis through the apical plasma membrane so a modified version of the calcium phosphate technique was employed for transfection of MDCK cells (Sambrook et al., 1989). The DNA was prepared as above and while the calcium phosphate precipitate was forming the MDCK cells were freed from tissue culture flasks by treatment with trypsin (0.25%) in PBS for 15 min, washed and resuspended at 2x 10^6/ml in 2 ml of medium in 10 ml tubes. The cells were pelleted and resuspended in the calcium phosphate/DNA mixture, left at room temperature for 15 min and then mixed with 5 ml of medium and transferred to 25 cm^2 tissue culture flasks, gassed with 5% CO_2 and placed at 37°C o/n. After 24 hr the medium was poured off and replaced with medium containing 0.8 mg/ml G418 and this medium was changed every two days. Flow cytometry was performed after 2 weeks and 15-30% of cells were positive. The fraction of positive cells was found to decrease with continued culture. The transfected cells were labelled with W3/13 and the 10% brightest cells isolated. A second sort was required to increase the percent of positive cells to more than 80%.

Jurkat E6-1 cells and NSO cells were co-transfected with pHSE3'-rat leukosialin linearized with Xho I and uncut pKG5 at a ratio of 5 : 1 (30-60 µg total DNA). Y3 cells were transfected with pHSE3'-mouse leukosialin and pKG5 in a similar manner. K562 cells and MDCK cells were transfected with uncut CDM8-RL and pKG5 at a similar ratio (30 µg DNA).

2.10.3 Transient transfection by DEAE-Dextran

COS-7 cells were plated out the night before transfection at 2 x 10^6 cells per 75 cm^2 flask containing RPMI/10% FCS/SKP. The next day, cells in each flask were washed twice with 25 ml RPMI, before adding 5 ml of DEAE-Dextran transfection mix made up in a 15 ml sterile tube. The transfection mix was made up as follows: 20 µl DEAE Dextran (Sigma, molecular weight 500 000, 100 mg/ml in PBS, stored at 20°C), 50 µl chloroquine diphosphate (Sigma, 10

mM in PBS, made up fresh each time), 5 ml RPMI and 15 μg of DNA, made by CsCl plasmid purification. Cells were gassed with 5% CO_2 and left for 3-4 hours at 37°C. The DEAE-Dextran transfection mix was poured off, the cells washed once with 25 ml PBS, and 5 ml of 10% DMSO in PBS added. Cells were left for 2 min in the 37°C incubator and the DMSO/PBS solution poured off. Cells were washed twice with 25 ml PBS and 30 ml RPMI/10% FCS/SKP was added. Cells were gassed with 5% CO_2 and left in the 37°C incubator o/n. The next day, cells were trypsinized and transferred to new flasks. Cells were grown one more days and then detached with 50 ml PBS/EDTA and labelled with mAbs for flow cytometry.

2.10.4 Storage of cell lines in liquid nitrogen

Cells in the exponential phase of growth were pelleted and resuspended at 10^7cells/ml in complete medium/10% DMSO, snap frozen on dry ice for 1-2 hr and then stored in liquid nitrogen. Cells were thawed at 37°C, washed twice in complete medium and were then cultured in 10 ml complete medium at 37°C.

2.10.5 Flow cytometry and FACS

Lymphoid cells from rat or mouse spleen, lymph nodes, thymus or bone marrow were prepared by conventional procedures (Williams et al., 1977) and washed in PBS/BSA. Cell lines were isolated directly from suspension culture (NSO, Jurkat and K562), by tapping the flask 3-4 times to release loosely bound cells (Y3) or by incubation in PBS/EDTA for 15-30 min to release plastic bound cells (MDCK, COS). Cells, 1-2x 10^6 per sample, were pelleted in 3.5 ml tubes and resuspended in 50 μl of mAb in the form of tissue culture supernatant, ascites fluid diluted 100 fold in PBS/BSA or purified Ig at 20 μg/ml in PBS/BSA. After incubation on ice for 45-60 min, cells were washed twice in 1 ml cold PBS/BSA and then resuspended in 50 μl RAM-FITC at approximately 15 μg/ml (pre-titred for optimal staining). After 30-60 min on ice 1 ml PBS/BSA was added and the cells analyzed on a FACSCAN (Becton Dickinson). Dead cells were gated out according to the forward scatter. For cell sorting, cells were labelled similarly but with sterile, dialysed antibody at a ratio of 10^8 cells/ml of tissue culture supernatant or secondary antibody. Cells were washed before sorting, resuspended at 3x 10^7/ml in complete medium and usually the 5-10% brightest cells were collected into a petri dish containing complete medium.

2.10.6 Neuraminidase treatment

For treatment with neuraminidase, cells were washed and resuspended in PBS titrated to pH 6 and containing 2 mM $CaCl_2$ at 10^7 cells per ml, and incubated with or without 0.05 U of *Vibrio cholera* or *Arthrobacter ureafaciens* neuraminidase (Calbiochem Corporation, La Jolla, CA) at 37°C for 1 hour.

2.11 Immunoperoxidase staining of tissue sections

Tissues were isolated into PBS/BSA, excess fat removed and cut into appropriate size pieces, and placed in a base mould containing OCT embedding medium (Miles Inc. USA). The tissue was covered with OCT and the base mould placed in a dry ice/isopentane bath until the block was frozen. Blocks were wrapped in aluminium foil and stored at -20°C. Cryostat sections (5-6 μM) were cut onto clean glass slides by Elizabeth Bailey, Sir William Dunn School of Pathology. The slides were air dried and sections circled with a DAKO pen prior to fixing in absolute ethanol at 4°C for 10 min. Following 3 washes with cold PBS, excess PBS was removed with Whatman 3MM paper and antibody (50-70 μl) was added on top of the section. Mab tissue culture supernatant was used diluted 1/3 in 10% BSA + 10% normal CBA/ca mouse serum (except with OX20 staining where mouse serum was not used). After 1 hr at 4°C in a moist atmosphere, slides were washed 3x with PBS and peroxidase conjugated antibody, diluted as recommended by the producer (usually 1/50) in BSA/mouse serum, was placed on the section for 1 hr at 4°C. During this incubation the enzyme substrate was prepared. 150 mg of diaminobenzidine was dissolved in 300 ml 0.14 M NaCl, 25 mM Tris-HCl (pH 7.6) and filtered through No. 1 Whatman filter paper to give a clear solution. Immediately prior to use 0.5 ml of 6% w/v H_2O_2 was added. Slides were washed 3x in PBS and transferred to enzyme substrate at RT and incubated for 5 min. Slides were washed again and counterstained for 10 sec with Harris' haematoxylin with immediate washing in water. Slides were then dehydrated by serial incubation (1-2 min) in 70%, 90% and 100% ethanol followed by 3 incubations in separate baths of xylene. Each slide was removed individually from the xylene with forceps and mounted in DPX. Staining for leukosialin expression in mouse tissues was with the S7 mAb; the S11 mAb gave higher background staining on FACS and was not used.

2.12 Expression in *E.coli*

Expression in *E.coli* using the pGEX-2T vector (Fig 2.1) was based on the original protocol of Smith and Johnsson (1988) with the specific details as follows.

2.12.1 Mini-protein inductions to screen clones for fusion protein expression.

A single colony was inoculated into 2 ml of broth and grown in a shaking incubator o/n at 37°C. 0.5 ml of the o/n culture was added to 4.5 ml of broth in a 50 ml (falcon) tube and shaken vigorously at 37°C for 1 hr. 10 μl of IPTG was added from a 100 mM stock and the cultures returned immediately to the incubator and culture continued for 3 hr. To screen for protein expression 1ml of culture was transferred to an Eppendorf tube and pulsed for 30 sec. The pelleted cells were resuspended in 100 μl sample buffer and 5-8 μl loaded per lane of an SDS PAGE gel. The presence of a fusion protein was visualized by coomassie blue staining of the gel. To test for fusion protein solubility the whole 5 ml culture was pelleted and resuspended in tris/sal/EDTA, PMSF (1 mM) and iodoacetamide (2 mM) were added, and the

cells lysed. Initally lysis was by addition of lysozyme to 0.2 mg/ml and incubation at 37°C for 30 min followed by addition of $MgCl_2$ to 5 mM and DNAse I to 20 µg/ml. The cells were then freeze/thawed 3 times using a dry ice/ethanol bath and a 37°C water bath, and then sonicated for 3x 10 seconds on an amplitude setting of 8 using an MSE sonicator. Subsequently it was observed that effective lysis was achieved by sonication alone and the first steps were omitted. After sonication the cells were microfuged for 10 min at 4°C and the s/n transferred to a new tube and mixed with 50-200 µl glutathione agarose (Sigma, G-4510, 50% in tris/sal) for 20 min, rotating at 4°C. The agarose beads were pelleted in a microfuge and washed 3-4 times with 1 ml tris/sal and either boiled directly in sample buffer or the protein was eluted as stated in the results.

2.12.2 Bulk Protein production

Batches of 3 l were usually grown. An o/n culture of 300 ml was diluted 1/10 in fresh medium with no more than 500 ml per 2 l flask to ensure good aeration. Induction was as described for the small cultures and cells were pelleted in a JA14 rotor at 7 K for 15 min. The cells from 3 l were resuspended in a total of 40-50 ml of tris/sal/EDTA using a 10 ml plastic disposable pipette and either frozen on dry ice or sujected directly to sonication. The sample was divided into 3x 15 ml vials and each was sonicated on ice using a probe of 10 mm diameter for 3x 30 seconds. The lysed cells were centrifuged in a SW35 rotor at 30 K for 30 min. Fusion protein was isolated from the lysate s/n by batch absorption with glutathione agarose (using approximately 5 ml per 30 ml lysate s/n) or by passage over a glutathione agarose column. Glutathione agarose columns were prepared as in section 2.14.3 and the lysate passed over the column at 1 column volume per hour. Often the column rate slowed significantly with time, especially when the lysate s/n was turbid and viscous, and it was necessary to resuspend the agarose in the column at intervals using a pasteur pipette. The column was washed with tris/sal and then eluted with glutathione elution buffer (50 mM Tris-HCl (pH 8), 5 mM reduced glutathione, Sigma) and 3-5 ml samples were collected. Separate columns were used to purify different fusion proteins. The columns could be re-used by washing with tris/sal. If the column was to be used for purification of a different protein it was first washed with 5 column volumes of 3 M NaCl.

2.12.3 Cleavage of fusion proteins with thrombin

The fusion protein, in tris/sal or glutathione elution buffer, was mixed with 0.05 vols of 3 M NaCl, 50 mM CaCl2, and bovine thrombin (Sigma T-7513) was added to 1 NIH unit per 2 mg protein from a 100 units/ml stock (in dH_2O). The sample was incubated at 37°C for the appropriate time, usually for 1 hr, and the efficiency of cleavage assessed by SDS PAGE. To remove the GST the cleaved sample, in tris/sal, was passed over a glutathione agarose column equilibrated with tris/sal and the fractions passing directly through collected.

2.13 Analysis of proteins

2.13.1 SDS PAGE

SDS PAGE analysis of proteins, base on the method of Laemmli, was performed using BIO-RAD mini-Protein slab gel apparatus. Resolving gels were between 7 and 18 % (w/w) acrylamide/ bisacrylamide (30:0.8), 0.375 M Tris-HCl (pH 8.8), 0.1% (w/v) SDS, 0.05% (w/v) ammonium persulphate and 0.1 % (v/v) TEMED. Stacking gels were 3% (w/w) acrylamide:bisacrylamide, 0.125 M Tris-HCl (pH 6.8), 0.1 % (w/v) SDS, 0.05% (w/v) ammonium persulphate and 0.2% (v/v) TEMED. 2x non-reducing sample buffer consisted of 0.2 M Tris-HCl (ph 6.8), 6% (w/v) SDS, 10 mM EDTA, 20% (v/v) glycerol and 0.005% (w/v) bromophenol blue. β-mercaptoethanol was added to 10% (v/v) to obtain 2x reducing sample buffer. Protein samples were prepared for electrophoresis by boiling in an equal volume of sample buffer, or as stated in the results, for 5 min. The running buffer for electrophoresis was 25 mM Tris-HCl, 192 mM glycine, 0.1% (w/v) SDS and gels were run at constant voltage (200-220 V) for 1 hr.

2.13.2 Western Blotting

Protein was transferred from SDS PAGE gels to nitrocellulose (Schleicher & Schuell, Dassell FRG) for 1 hr at 100 V (4°C) using the BioRad blotting apparatus. 3 µl rainbow molecular weight markers (Amersham) were run on gels to be used in Western blotting. The membrane was placed in 5% milk powder in PBS for 30-60 min at RT. Then 1/5 diluted tissue culture supernatant or 1/500 diluted ascites fluid (in 5% milk powder) containing the appropriate mAb was added for 2 h at 4°C with gentle rocking. The membranes were washed twice for 20 min in PBS/0.05% Tween 20 before addition of ^{125}I-RAM (10^6 cpm/ml) in PBS/0.05% Tween 20/1% BSA. After 1 hour at 4°C the membranes were washed twice for 10 min each with PBS/0.05% Tween 20, and then exposed to X-OMAT S film (Kodak, France) at -70°C, usually for 16 hr.

2.13.3 Plate binding assay

Purified target proteins at 50 µg/ml were used to coat flexible assay plates (Falcon 3911, Microtest III) by incubation at RT for 30-60 min. Free sites on the plates were blocked by incubation with 2% BSA for 10 min. The plates were washed and used in assays with the first binding step involving incubation with 50 µl mAbs for 60 min at 4°C and the second step a similar incubation with ^{125}I-RAM at 1.5 x 10^5 cpm/well. The wells were counted by cutting up the dried soft plate into individual wells and trasferral to an LKB 1261 Multigammer counter. Counts were collected over 15 seconds for each sample.

2.13.4 Inhibition assay

Before use mAbs were titrated in a direct binding assay. Usually 12 three fold serial dilutions in

PBS/BSA were made starting at 1/100 of ascites or neat tissue culture s/n, by mixing 70 µl with 140 µl of PBS/BSA in a microtitre tray. Counts were plotted against antibody concentration and the dilution chosen was where the curve was steepest and had just begun to fall. The protein to be tested was diluted usually to a starting concentration of 100 µg/ml and serial 3 fold dilutions were made. 60 µl of each dilution was mixed with 60 µl of diluted mAb and incubated at 4°C for 1-2 hr. During this time a target plate was set up as for the direct binding assay. 50 µl of the mAb/target protein mixture was transferred in duplicate to wells of the target plate and incubated shaking at 4°C for 1 hr. The plates were washed and incubated with ^{125}I-RAM as in the direct binding assay and then the wells counted. The inhibition assay performed with cells rather than soluble protein in the inhibition step was performed in the same way except the cells were pelleted at each washing step by centrifugation at 600g for 3 min.

2.13.5 Amino acid compositional analysis

Samples were freeze dried in Waters PICO-TAG sample tubes using a Savant Speedvac concentrator. The tubes were placed in purpose-made reaction vials which contained 200 µl of 5.7 M constant boiling HCl and 10 µl of 0.1 M phenol. The vials were sealed, purged with nitrogen, evacuated and left at 110°C in a PICO-TAG work station for 16-24 hr. The sample tubes were removed and dried under vacuum before derivatization and analysis which was performed by Tony Willis of the MRC Immunochemistry Unit, Department of Biochemistry, using an LKB amino acid analyzer and a Waters PICO-TAG system. The system was calibrated using a known mixture of amino acids and saccharide residues. An internal standard of nor-leucine was included in all samples.

2.13.6 Determination of protein concentration

The absorbance at 280 nm (A_{280}) over a path length of 1 cm was measured and an A_{280}=1 was taken as 1 mg/ml except with antibodies where A_{280}=1.4, and with Fab where A_{280}=1.5, was taken as 1 mg/ml. In addition, the concentration of a sample of the leukosialin cytoplasmic domain (CY-1) purified by gel filtration was determined by compositional analysis, and from this it was calculated that A_{280}=1 corresponded to a 1.56 mg/ml solution. The aborbance at 260 nm was also routinely measured to gauge the purity of the protein solution.

2.14 Purification of antibodies and proteins

2.14.1 Purification of IgG and preparation of Fab fragments

Monoclonal IgG was prepared from ascites fluid as described (Mason and Williams, 1980). Briefly, the ascites was spun at 10 000g for 20 min to remove clots and the antibody precipitated by adding Na_2SO_4 to 18% w/v. The sample was incubated at 37°C for 30 min and the precipitate pelleted by centrifugation at 25°C, 8 K (7 000 g) for 15 min in a SW35 rotor. The pellet was resuspended to 33% of the initial volume and Na_2SO_4 precipitation repeated.

The pellet was resuspended in 20% of the initial volume and in some cases further purified by ion-exchange chromatography using DEAE Biogel A (Pharmacia). The column was equilibrated with and the sample loaded in 25 mM Tris-HCl (pH 7.4), 50 mM NaCl and then washed further with this buffer. Elution was with 25 mM Tris-HCl, 125 mM NaCl. Further purification was by gel filtration using a Sephacryl S200 column. Fabs were prepared from the OX56, OX57 and OX75 Igs by treatment with papain (Sigma Chemical Co.) as in (Davis et al., 1990a). The cleaved Fab fragments were purified by chromatography on DEAE-cellulose equilibrated with 15mM sodium phosphate (pH 8), eluting with 100 mM sodium phosphate. The Fabs were further purified by gel filtration on a Sephacryl S-200 (Pharmacia Ltd. Milton Keynes) column in tris/sal. The W3/13 Fab was provided by Alan Williams and had been prepared by treating with pepsin to generate F(ab')2 and this was purified in the same manner as the Fab and then reduced and alkylated (Parham, 1986). The OX75 Fab was further purified using a Sephadex G75 column as described in the results.

2.14.2 Coupling of antibodies to Sepharose 4B

IgG at 10 mg/ml was dialysed o/n against 50 mM disodium tetraborate HCl pH8. Sepharose CL (Pharmacia) at 1 ml per 10 mg of IgG was washed on a filter funnel with dH_2O to remove fines. The sepharose was placed in a beaker in the fume hood, stirring with 0.4 vols of ice-cold dH_2O and 1 vol of ice-cold 4 M K_2HPO_4, KOH pH 11.5 was added. CNBr was dissolved at 300 mg/ml and added to the stirred sepharose 4B over 1 min to give 40 mg CNBr/ml. After stirring on ice for 10 min, cold dH_2O was added and the beads collected on a sintered glass funnel. The beads were quickly washed with water until the pH of the eluate was neutral (about 1 l). The gel was then added to the protein and incubated at 4°C o/n on a rotating wheel. The beads were washed with 50 mM ethanolamine HCl pH 8 for 10 min and then with 10 vols of elution buffer (0.5 M propionic acid pH 2.5 or 50 mM diethylamine pH 11.5), neutralized and stored at 4°C in tris/sal.

2.14.3 Antibody affinity columns

Affinity columns were set up in 10 or 20 ml glass syringes with scintered glass as a base. All columns were run at 4°C. During the binding step flow rate was kept below a maximum of 1-2 column volumes an hr. Flow rate was increased during washing which was usually with tris/sal and was continued until the A_{280} of the wash buffer was less than 0.03 (approximately 10 column volumes). Elution of the OX-34 column was with 0.5 M propionic acid (pH 2.5) and fractions were immediately neutralized with solid Tris-HCl. With the W3/13 and OX-56 columns elution was with 50 mM diethylamine (pH 11.5) and the fractions were neutralized with solid glycine. The fractions were monitored as they emerged and as soon as the A_{280} peaked and began to fall the column was returned to tris/sal. The peak fractions were pooled, concentrated with a centricon apparatus and dialysed against tris/sal for 24-48 hr.

2.14.4 Gel filtration

Sephacryl S200 or Sephadex G75 (Pharmacia Ltd. Milton Keynes) columns were poured in columns of dimensions 100x1 cm. Columns were equilibrated either with tris/sal or 0.1 M ammonium bicarbonate. S200 columns were run at a flow rate of 20 ml/hr and G75 columns at 10 ml/hr and 1.5-5 ml fractions were collected using LKB fraction collectors. Samples were loaded in a maximum volume of 1-2 ml. When samples were being collected for NMR spectroscopy acid washed tubes were used. Fractions were assayed for protein by measuring the A_{280} and peak fractions were run on SDS PAGE to determine purity before fractions were pooled.

2.14.5 Concentration of protein samples

For most proteins, centricon 10 concentrators (Amicon) were used. Samples were centrifuged in the concentrators at 4°C and 3 K (1 300 g) for the time required. Smaller aliquots were concentrated in 3 ml concentrators (Amicon) spinning at 8 K (6 500 g) in a JA20 rotor. With the leukosialin cytoplasmic domain concentration was by vacuum drying in a Savant Speedvac concentrator. Samples were prepared in 0.1 M ammonium bicarbonate and centrifuged under vacuum until the desired concentration or until dry. The lyophilised protein was then resuspended as desired.

2.14.6 Sucrose Gradients

Sucrose density gradients (1.6ml) for hydrodynamic studies were prepared in 2 ml polyalomer ultracentrifuge tubes (Beckman Ltd., High Wycombe, UK) by diffusion for greater than 4 hours at 4°C of 0.36 ml aliquots of 30, 25, 20, 15 and 10% sucrose 140 mM NaCl, 10 mM Tris-HCl, pH 8. The protein solution in a total volume of 100 µl was loaded immediately prior to centrifugation at 55 000 rpm for 24 hr at 4°C (accel 9, deaccel 0) in an SW50 swing out rotor. Fractions were collected by insertion of a hollow needle (diameter 2 mm) attached to tubing prefilled with 30% sucrose through the gradient to its base and then holding the other end of the tubing below the base of the gradient such that the solution was syphoned out. 7 drop fractions (approx. 18 per gradient) were collected in 0.5 ml Eppendorf tubes. Step gradients for separation of membrane preparations (Standring and Williams, 1978) were poured in Beckman L8-M tubes (5 ml) using 1.35 ml of cold 40%, 28% and 10% sucrose in 10 mM Tris-HCl (pH 7.4) per step. The lysate sample was loaded immediately and the gradients spun at 29 000 rpm (80 000 g) at 4°C for 20 hr. Fractions (15 drop) were isolated as above.

2.15 Labelling antibodies and proteins for use as probes

2.15.1 FITC conjugation and biotinylation of proteins

For FITC conjugation, 5.3% Na_2CO_3 and 4.2% $NaHCO_3$ were prepared and 5.8 ml of

Na_2CO_3 was mixed with 10 ml $NaHCO_3$ and then diluted 1:10 with 0.15 M NaCl to make the cojugation buffer (pH 9.5). The buffer was used to equilibrate a 5 ml G50 column and the protein desalted over the column. FITC was dissolved in conjugation buffer to 1 mg/ml and 0.3 ml was added per ml of protein solution (at approximately 5 mg/ml). The sample was incubated for 2 hr in the dark and then separated on a 10 ml G50 column equilibrated with PBS/10 mM azide. The ratio A_{495}/A_{280} was usually greater than 1.
For biotinylation the protein was used at 2-5 mg/ml in PBS (no azide). The biotin-N-hydroxysuccinimide ester (NHS-LC-Biotin; Pierce, USA) was dissolved in DMSO at 20-40 mg/ml and added to the protein at 0.2 mg ester per mg protein. The reaction was allowed to continue for 1-2 hr at RT and the free ester was removed using a 10 ml G50 column or by dialysis o/n at 4°C into PBS/azide.

2.15.2 Iodination of proteins

A 10 ml G50 column (PD10, Pharmacia) was washed with PBS/10 mM azide and pre-blocked with 0.5 ml of 10% BSA followed by 2 volumes of PBS/azide. In a fume hood, 10 µl freshly prepared chloramine T (2 mg/ml in 0.3 M phosphate buffer), 5µl (0.5 mCi) ^{125}I and 20 µg of protein were mixed and after reacting for 2 min, 25 µl of tyrosine (0.3 mg/ml in 0.3M phosphate buffer, stored at 4°C) was added and left for 30 sec. 50 µl 10% BSA was then added and the sample loaded onto the 'dry' PD10 column. 2.4 ml PBS/azide was washed through and then a further 2 ml was applied and collected as labelled protein. Free iodine was washed from the column and it was then discarded. An aliquot of labelled protein was counted on the gamma counter and usually around 200 000 cpm per ul were incorporated, corresponding to a specific activity of approximately 200 cpm/ pmol for antibodies. Antibody probes were stored at 4°C. Labelled cytoplasmic domain fusion protein was divided into 0.6 ml aliquots and stored at -20°C. 0.3 M phosphate buffer pH 7.4 is made up as: 9.8g $NaH_2PO_4.2H_2O$, 37.9g Na_2HPO_4 anhydrous, 10 ml 1 M NaN_3 upto 1 l with dH_2O and stored at -20°C.

2.16 Phosphorylation by Protein Kinase C

The following protocol and the protein kinase C, which was a mixture of bovine brain α, β and γ isoforms (~500 units/ml), was kindly provided by Peter Parker (ICRF London). The lipid preparation was prepared by mixing 100 µl of phosphatidyl serine (Lipid Products, Surrey) at 10 mg/ml in chloroform/methanol with 1 µl of PMA (2 mg/ml in DMSO) in an acid washed glass tube, dryed down under N_2 for 10 min and then redissolving in 200 µl of 1% Triton X-100, 20 mM Hepes pH 7.5. The preparation was vortexed, incubated at 30°C for 5-10 min, further vortexed and stored on ice until used. The suspension usually became clear after a few minutes on ice. The enzyme mixture was diluted 1/12.5 in 20 mM Tris-HCl, pH 7.5, 2mM EDTA, 5 mM DTT, 0.02% Triton X-100. The reaction was then set up using the following

volumes or multiples thereof: 10 µl substrate, 5 µl $MgCl_2$ 0.1 M, 5 µl $CaCl_2$ 6 mM, 10 µl lipid preparation, 5 µl diluted enzyme and 5 µl ATP 1mM containing γ-[^{32}P]-ATP to a specific activity of 200-500 cpm/pmol. To assess incorporation, samples were run on SDS PAGE and then gel dried down onto Whatmann 3MM paper using a gel drier. After exposure to X-OMAT film, the protein bands were cut from the gel, placed in Scintillation fluid and counted for 1 min using a LKB 1211 Rackbeta Liquid Scintillation Counter. A sample of reaction mixture containing a total amount of ATP equivalent to that run in each lane, was spotted directly onto a segment of dried gel lacking protein and allowed time to absorb. This was then counted in the same way. The activity present per mole of phosphate was determined and from this, and the known number of moles of protein loaded in each lane, the number of moles of phosphate incorporated per mole of protein were calculated. To detect phosphorylated residues by amino terminal sequencing, Tony Willis employed a modified protocol using an ABI 473A sequencer. In particular, samples were applied to a Sequelon-DITC membrane (Millipore, UK Ltd.) so that they would be covalently bound. This is a polyvinylidene difluoride membrane which has been derivatised with 1,4-phenylene diisothiocyanate (DITC) to produce a surface that reacts witth amino groups of both the N-terminal amine and the ε-amino groups of lysines within the protein to form thioureas. Modified washing conditions were employed to prevent precleavage of the DITC amino acid, and a fraction collector was connected to the sequencer such that 80% of the residue released at each cycle was injected into the HPLC system and 20% was collected and counted for ^{32}P.

2.17 V8 Protease cleavage and HPLC of leukosialin cytoplasmic domain

For V8 protease cleavage 100 µg of leukosialin cytoplasmic domain, with or without pre-phosphorylation, was desalted using a 10 ml G50 column pre-equilibrated with 25 mM NH_4HCO_3 (pH 7.8) and V8 protease (endoproteinase glu-C, Boehringer Mannheim) was added at 2% w/w and the sample placed at 25°C o/n. The sample was then used directly in HPLC. This was on an Applied Biosystems HPLC setup (1480A Injector/mixer, 1400A Solvent delivery system, 1783A Absorbance detector-controller and a dual pen recorder) using an RP18 analytical column. Linear gradients were run over 30 min at 1 ml/min from 5 to 70% acetonitrile in 0.1 % (v/v) trifluoroacetic acid. Effluent with absorbance at 206 nm was collected by hand in Eppendorf tubes with separate tubes for each peak. With phosphorylated samples the tubes were counted for ^{32}P in a Rackbeta counter. Compositional analysis was performed on a sample from each peak.

2.18 Rotary shadowing and electron microscopy

Rat thymocyte leukosialin was from existing preparations isolated by William Brown (Brown et al., 1981) and Nigel Killeen (Killeen et al., 1987). Sample concentration was estimated by compositional analysis and purified leukosialin, at 100 µg/ml in 0.1 M NH_4HCO_3 with or

without 0.5% sodium deoxycholate, was diluted 10-fold in 77% (v/v) glycerol : 0.1 M NH_4HCO_3 prior to rotary shadowing. Leukosialin : Fab complexes were formed by incubating equal weights of the two proteins o/n at 4°C and then treating as above in the absence of deoxycholate. The following steps were carried out by David Shotton. 50 µl of the samples in glycerol were sprayed without delay as a mist onto the surfaces of freshly cleaved mica sheets, using a laboratory aerosol spray canister, and dried at RT by evacuation to better than 2×10^{-5} mbar for 10 min on the uncooled specimen stage of a Balzers BAF 301 freeze-fracture apparatus. They were then rotary shadowed with platinum-carbon at an incident angle of 6° while rotating at 2 Hz as previously described (Shotton et al. 1979; Woollett et al. 1985). Replicas were floated onto distilled water, picked up on bare 360 hexagonal mesh copper grids and photographed on Ilford EM film using a Philips EM 400T electron microscope operated conventionally at 80 kv.

2.19 Detection of interactions with the leukosialin cytoplasmic domain

2.19.1 Preparation of eukaryotic cell lysates and cell membranes

Lysates were routinely prepared in the following buffer: 10 mM Tris-HCl pH 7.4, 140 mM NaCl, 1 mM EDTA with NP40 detergent (Sigma) added to 1%. Protease inhibitors added immediately before use included PMSF to 200 µM, iodoacetamide to 2 mM, leupeptin to 2 µM and aprotinin at 1/100. Freshly isolated cells were washed and resuspended in cold lysis buffer at 1-5x 10^8/ml, incubated on ice with occassional vortexing for 1 hr and then microfuged at 4°C for 20 min. The supernatant was stored in aliquots at -20°C or used directly (freeze thawing of samples lead to precipitation of large amounts of material and was avoided). The pellet was discarded. For preparation of membranes, cells were resuspended in lysis buffer without detergent or with tween 40 at 2.5% , incubated on ice for 1 hr with vortexing and the cells in the absence of detergent were lysed by sonication for 12 seconds. The cells were spun at 600 g for 10 min and the s/n transferred to eppendorfs which were then spun at 100 000 g in a TL100 rotor for 30 min. The pellets were resuspended in 0.5 ml tris/sal at $6x10^8$ cells/ml.

2.19.2 Ligand blotting with cytoplasmic domain fusion protein

Proteins run on SDS PAGE were transferred to nitrocellulose as for Western blotting and incubated in 5% milk powder in PBS (Blotto) for 1 hr ar RT. The membranes were transferred to appropirate size dishes with a minimum volume of fresh blotto and purified GST was added to approximately 0.25 mg/ml. Iodinated probe, cytoplasmic domain fusion protein or a control fusion protein or GST, was added to 10^6 cpm per ml and incubation was for 2 hr at 4°C. The membranes were rinsed in cold PBS/0.05% Tween 20 and washed in this for 2x 10 min. If background was still high a further wash was performed. Exposure to autoradiograph film was usually at -70°C o/n.

2.19.3 Affinity purification from cell lysates with fusion proteins

Fusion protein or GST was added to samples of lysate in Eppendorf tubes to a final concentration of approximately 0.5 mg/ml (from a stock at 3-10 mg/ml in tris/sal) and incubated at 4°C for 1-2 hr. Glutathione agarose was then added in excess (50-100 µl per 200 µl of fusion protein/lysate mixture) and incubated for 20 min with rotation at 4°C. The absorbed lysate was isolated and the agarose with bound fusion protein and any attached proteins was washed rapidly at 4°C 3-4 times by addition of 1 ml cold lysis buffer to the Eppendorf tube, briefly vortexing, pelleting in a microfuge for 30 seconds and aspirating the s/n. In some instances the agarose pellet was resuspended directly in sample buffer. In other cases the fusion protein and attached proteins were eluted with glutathione elution buffer and samples of the elution analyzed. Elution of the bound 110 kd protein alone by increased salt concentration was also performed as described in the results.

2.20 Aggregation Assay

Thymocytes, spleen and lymph node cells were isolated from organs of 6-10 week old AO or PVG rats. Aggregation assays were performed in 96 well flat bottomed microtitre plates (Falcon 3072, Becton Dickinson, New Jersey). In most cases wells were pre-coated with 50% heat inactivated FCS in PBS for 1 hr at RT. The FCS was removed and mAbs or chemicals were added in a maximum volume of 30 µl, followed by 200 µl of freshly isolated thymocytes at $3x10^6$/ml in RPMI,10mM Hepes, and antibiotics (streptomycin, kanamycin and penicillin). With lymph node and spleen cells, $1.5x10^6$/ml were used. The plates were placed at 37°C or 4°C and aggregation scored at a number of time points. When testing cation requirements cells were washed and plated in PBS. Aggregation was measured in a semi-quantitative manner based on a modification of the system used by Rothlein and Springer (1986). Scores ranged from 0 to 5, where 0 indicates that less than 1% of cells were clumped; 1 that 1-5% were in small aggregates (5-20 cells); 2 that 5-30% were in small clumps (10-50 cells); 3 that 30-60% were in medium sized clumps (20-100 cells); 4 that 60-90% were in large clumps (>50 cells) and 5 that >90% of cells were aggregated in large balls of cells (>100 and often >1000 cells). Only integer values were assigned to a given well in any experiment; fractional values are derived from the means of multiple experiments.

Anti-rat LFA-1 (WT1) and anti-ICAM-1 (IA29) mAbs used in this study were the kind gift of M. Miyaska (Tamatani and Miyasaka, 1990; Tamatani et al., 1991). R73 was provided by T. Hunig and HIS41 by R.Aspinall. Phorbol 12-myristate 13-acetate (PMA), cycloheximide, cytochalasin B, and 2-deoxyglucose were obtained from Sigma Chemicals Co., actinomycin D from P-L Biochemicals, Inc. Milwaukee Wis. PMA and cytochalasin B were prepared in DMSO at 2 mg/ml and 10 mM, respectively. DMSO at 1/200 was not observed to have any effect on aggregation.

CHAPTER 3

PROTEIN SEQUENCE AND GENE STRUCTURE FOR MOUSE LEUKOSIALIN

3.1 Introduction

3.2 Isolation of partial cDNA clones and full-length genomic clone for mouse leukosialin

3.3 Sequence analysis demonstrates that clone GL11a contains a single open reading frame likely to encode leukosialin

3.4 Identification of a mouse leukosialin pseudogene

3.5 The BALB/c mouse genome contains two genes homologous to leukosialin compared with the possibility of five genes in the rat

3.6 Structural features of leukosialin proteins

3.6.1 O-linked glycosylation

3.6.2 N-linked glycosylation

3.6.3 Extracellular regions of sequence similarity

3.7 Phosphorylation sites

3.8 Homology

3.9 Discussion

3.1 Introduction

Isolation of a genomic clone for mouse leukosialin is of interest in relation to the origin of heterogeneity in leukosialin molecules (see Chapter 1). The biosynthesis and genomic data for human leukosialin are in accord with only one gene for all forms of the molecule. However, in the rat, Southern analysis on genomic DNA suggested the possibility of up to 5 leukosialin genes (Nigel Killeen, D.Phil Thesis, 1988). Furthermore, alternative splicing of small exons as is found for CD45 (Thomas, 1989) might have gone undetected in previous work. Characterization of the mouse leukosialin gene is also important in allowing *in vivo* studies of the molecule and the possibility of interrupting the gene in the genome by homologous recombination.

In this chapter I describe the isolation of a genomic clone for mouse leukosialin and demonstrate that the open reading frame is without introns. The second gene present in the murine genome is shown to be a pseudogene. These findings demonstrate that the variations in antigenicity and molecular weight of mouse leukosialin that have been reported must be due to differences in post-translational modifications.

3.2 Isolation of partial cDNA clones and full-length genomic clone for mouse leukosialin

The isolation and sequence analysis of the partial cDNA clone was performed by Chamorro Somoza and Nigel Killeen (prior to my arrival in the laboratory). A segment of the rat leukosialin cDNA from the C-terminal coding region was used as a probe to screen a BALB/c thymocyte cDNA library in lambda gt10. Sequence analysis of a clone isolated from this library demonstrated that its 5' sequence began about 310 nucleotides after the beginning of the coding region based on a comparison with the rat sequence (Killeen et al., 1987) and it had an approximately 2.8 kb 3' untranslated sequence. In order to complete the sequence I used this clone to probe a genomic library for the mouse leukosialin gene. This strategy was chosen as the human leukosialin gene had recently been shown not to have any introns in the coding sequence (Shelley et al., 1990).

A genomic library prepared from DBA/2J liver DNA (obtained from Clontech Laboratories Inc.) was screened with a cDNA probe corresponding to the coding section of the partial

murine cDNA clone. Screening $1.2x10^6$ plaques produced a single clone, GL1p. This clone did not hybridize to an oligonuclueotide specific for the 5' region of the sequenced murine cDNA clone and its restriction map did not correspond to that of the cDNA. Preliminary sequence analysis indicated it had only 80% identity with the murine leukosialin cDNA. Further screening of $2.5x10^6$ plaques produced a clone, GL11a, which was found to bind the oligonucleotide and to have a restriction map compatible with that of the cDNA.

3.3 Sequence analysis demonstrates that Clone GL11a contains a single open reading frame likely to encode leukosialin.

The region of the GL11a clone which bound to the cDNA probes was subcloned and sequenced in both directions (Fig. 3.1). An open reading frame which corresponds to a predicted protein of 395 amino acids, with a 19 amino acid leader was identified (Fig. 3.2). The sequence of the partial cDNA clone exactly matched sequence in the proposed open reading frame of the genomic clone (data not shown). The open reading frame is suggested to start at the ATG indicated in Fig. 1 because of the high level of identity with both the rat and human cDNA sequences generated by such an initiation, and because the consensus rules for initiation codons are satisfied (Kozak, 1987). In addition the predicted size of the protein encoded by this open reading frame (38.0 kd) is in good agreement with that of the protein backbone of the human and rat leukosialin molecules (38.4 and 37.6 kd, respectively). The mature mouse leukosialin glycoprotein has a similar apparent Mr to rat and human leukosialin although both the mouse and human forms have been reported to be slightly larger than rat leukosialin (Pink, 1983), probably because of the presence of an N-linked carbohydrate structure in the mouse and human forms.

A 19 amino acid leader sequence ending at the indicated Asp residue (Fig. 3.2) is suggested from a comparison with the rat sequence for which the NH_2-terminus has been determined by protein sequencing (Killeen et al., 1987). The Pro at -1 in mouse leukosialin is not expected from the -3, -1 rule, but its occurence is not unprecedented (von Heijne, 1986) and Ala is very common at the -3 position.

It would appear that the coding sequence lacks introns since there is only one open reading

frame and this corresponds to the rat sequence over its entire length. In addition no consensus sequences for acceptor sites (Padgett et al., 1986) in intron splicing could be found in the open reading frame. The point at which transcription begins and ends in the mouse gene is not known and thus there is the possibility of introns in the 5' or 3' noncoding regions. Shelley and coworkers (Shelley et al., 1990) have demonstrated the presence of an intron in the 5' untranslated region of the human gene and when the first 554 nucleotide 5' of the initiation ATG are aligned between mouse and human leukosialin the splice acceptor and donor sites for this intron are seen to be conserved (Fig. 3.3). Thus it is likely that an intron of 379 bp exists in the 5' untranslated region of mouse leukosialin. The sequence conservation of this intron, 66% identity between mouse and human (compared to 75% in the flanking 5' untranslated regions) is suprisingly high and implies that the intron has some functional activity. There is recent evidence that it may have enhancer activity (Kudo and Fukuda, 1991).

3.4 Identification of a mouse leukosialin pseudogene

Analysis of clone GL1p revealed a sequence with significant homology to GL11a at the nucleotide level (Fig. 3.1). At the amino acid level the sequence corresponding to the coding region of GL11a is interrupted by a number of stop codons in GL1p, and there have been three frame-shifts due to base deletion or insertion. This clone also lacks an ATG in the expected position and the 5' region in the area of the leader in GL1p shows no homology to GL11a or to leader sequences in general. Based on the homology with the GL11a clone and with the rat and human leukosialin sequences, and the apparent absence of consensus splice sites, this pseudogene also appears to be intronless in the region corresponding to the coding region. The 5' regions of the two murine genes demonstrate little homology (33% over 540 residues, data not shown).

3.5 The BALB/c mouse genome contains two genes homologous to leukosialin compared with the possibility of five genes in the rat

Southern blot analysis was performed with BALB/c genomic DNA using a number of enzymes, and probing with the partial cDNA corresponding to the region from nucleotide 310 to the second Bam HI site in GL11a (Fig. 3.4c, probe 1). The fragments which

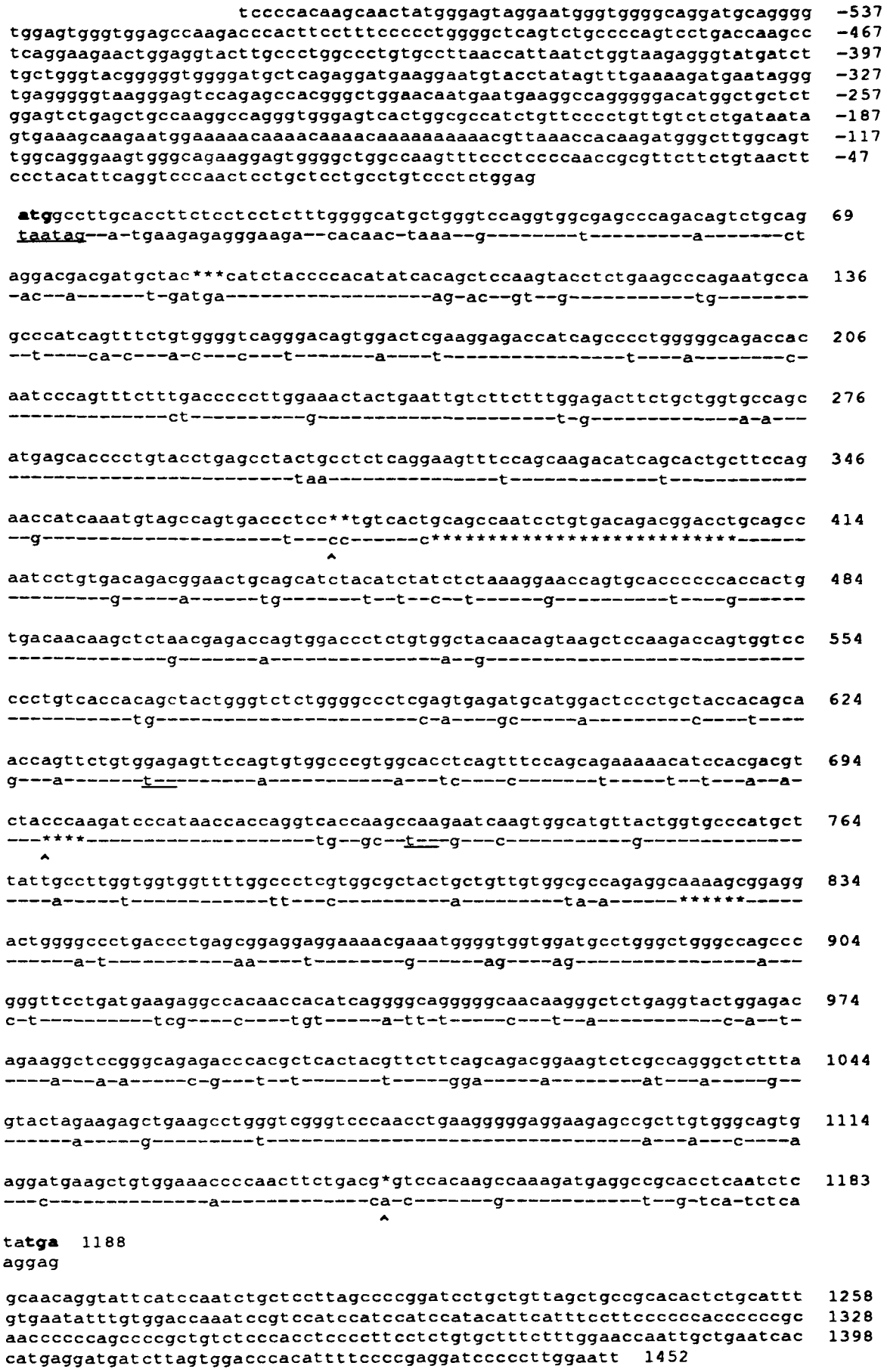


Figure 3.1 Nucleotide sequence of mouse genomic leukosialin gene (GL11a) and related pseudogene (GL1p). The nucleotide sequence of the mouse leukosialin gene (GL11a) is numbered from the initiator ATG and includes the 5' non-coding (to residue -587) and 3' noncoding (residues 1189 to 1452) nucleotides. Differences in the sequence of the pseudogene (GL1p) are shown for the region corresponding to the open reading frame of GL11a, where dashes indicate nucleotide identities and * indicate gaps inserted to maximize homology. ^ indicates insertions or deletions in the pseudogene that lead to frameshifts in the coding sequence as argued by homology. Stop codons which are present in the pseudogene after correcting for frameshifts are underlined. The sequences shown, and a further 147 nucleotides of 5' and 240 nucleotides of 3' sequence of clone GL1p, are in the EMBL/GenBank/DDBJ Nucleotide Sequence Databases under the accession numbers X17018 (leukosialin gene) and X17017 (leukosialin pseudogene).

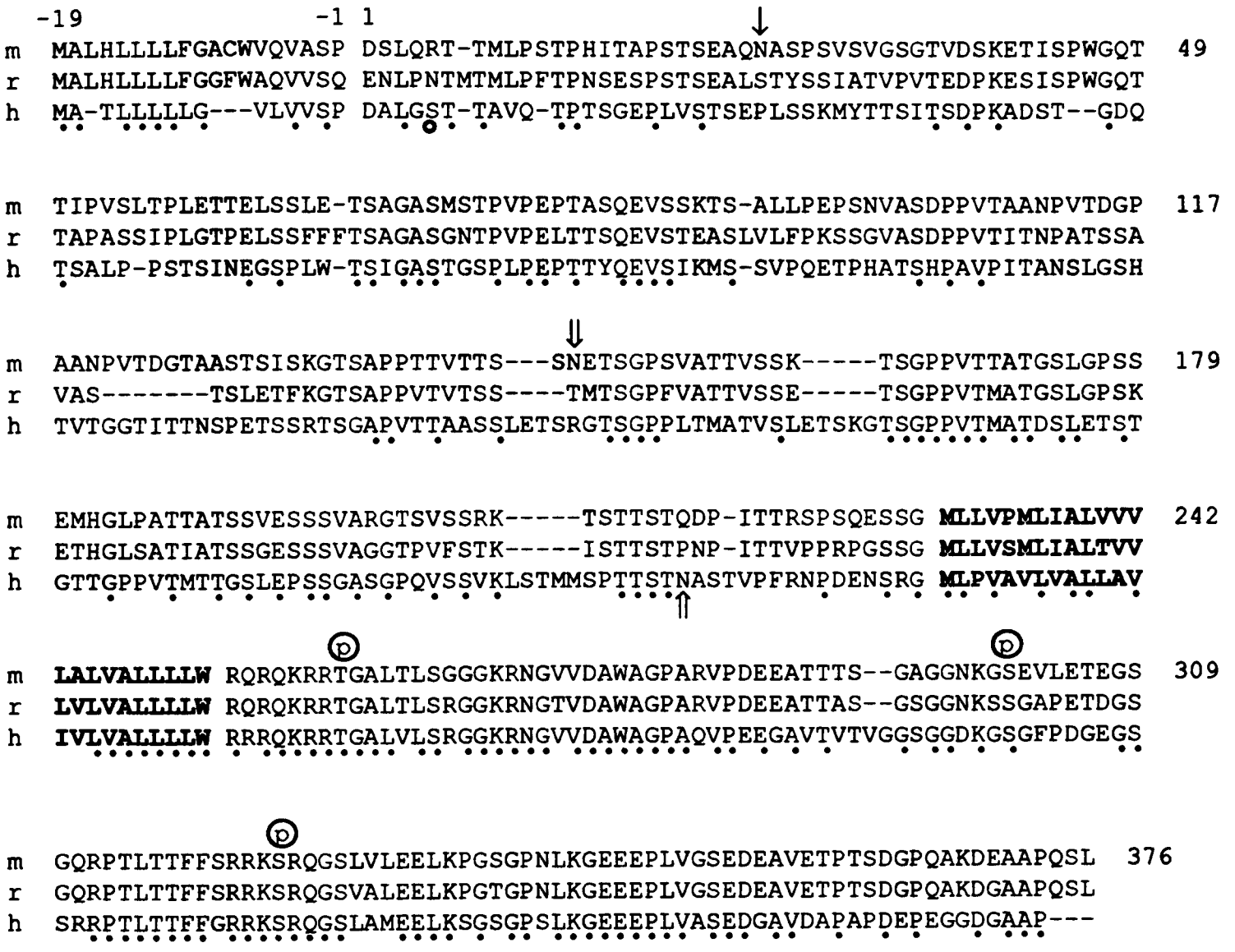


Figure 3.2 Mouse (m) leukosialin protein sequence and its alignment with rat (r) and human (h) leukosialin sequences. Amino acids are shown in single letter code. The numbering refers to the mouse sequence with residue 1 being the Asp that is proposed as the NH_2-terminal residue of the mature protein by alignment with the known NH_2-terminus of rat leukosialin. The mature NH_2-terminus of human leukosialin is likely to be the ser indicated by the open circle (Schmid et al., 1992). The region in bold corresponds to the proposed transmembrane domain. Spaces introduced to maximize homology are shown as dashes. Dots under the sequence alignments indicate amino acid residues that are identical in the three species. Double arrows indicate potential sites of N-linked glycosylation in the mouse and human sequences, and the single arrow indicates a site in the mouse sequence which is unlikely to be used. Conserved potential sites for phosphorylation by protein kinase C are marked Ⓟ. Data for human leukosialin is from (Pallant et al., 1989; Shelley et al., 1989). The rat sequence is from (Killeen et al., 1987) with the leader sequence being determined as described in Chapter 2.

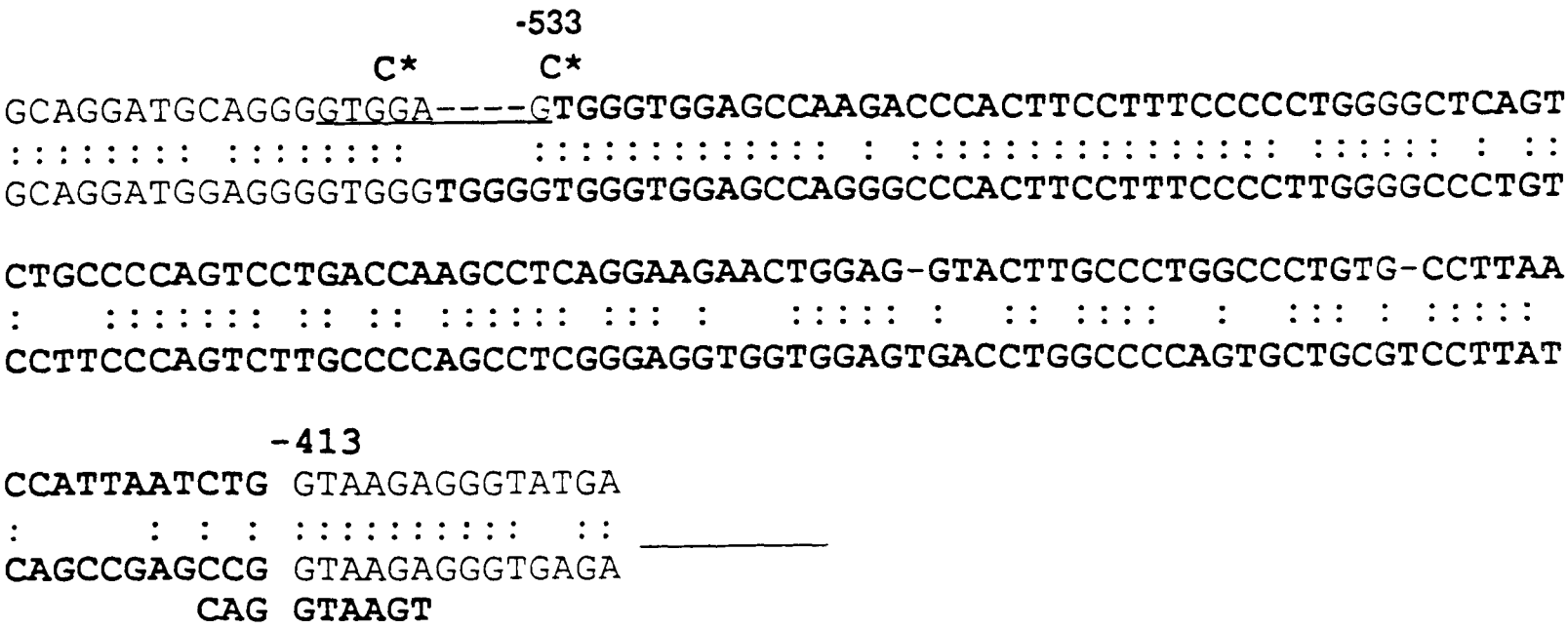


Figure 3.3 Alignment of mouse and human leukosialin gene sequences 5' of the initiation ATG showing the presence of an intron and the location of a Sal I site introduced to allow expression. The mouse sequence is shown in the upper line. The regions in bold correspond to 5' untranslated sequence shown to be present in processed human leukosialin transcripts and the location of the intron is based on the human gene sequence (Shelley et al., 1990). The consensus splice acceptor and donor sites are shown beneath the human sequence where Y indicates pyrimidine and N any nucleotide (Padgett et al., 1986). Dashes indicate spaces introduced to maximise homology. Numbering is with respect to the mouse sequence where the A of the initiation ATG is taken as +1. The C* indicates the two nucleotides changed by mutagenesis to introduce a Sal I site (gtcgac).

hybridize following Bam HI digestion (Fig. 3.4a) correspond exactly to those predicted from the restriction maps of the two isolated genomic clones. That is, a single fragment of 2280bp for GL11a, and two fragments of 1232 and 717 bp for GL1p. The greater intensity of the 2280bp band (Fig. 3.4a) is in agreement with its having 100% homology to the cDNA probe, whereas the two smaller bands, corresponding to the pseudogene, have only 80% homology and so bind the probe less strongly. The same pattern is obtained when the whole gene, including 600bp of 5' noncoding sequence (Fig. 3.4c, probe 2) is used as a probe (data not shown). Digestion with Sty I produces a number of low molecular weight fragments. The restriction map for GL11a indicates 2 fragments should have been generated by Sty I treatment (543 and 673bp) reactive with the probe and these are seen in Figure 3.4a in addition to a larger band corresponding to the 879bp fragment expected from the map of GL1p. The patterns of hybridization after Acc I and Bgl II digestion (Fig. 3.4a) are also commensurate with there being only two genes. This contrasts with the situation in the rat, where Southern blot analysis (performed by Nigel Killeen) using rat leukosialin cDNA probes generates a complex pattern. Figure 3.4b shows an example involving Eco RI digestion of rat genomic DNA. Five or six bands clearly hybridize with probes from the 5' and 3' regions of the coding sequence implying that each fragment includes the equivalent of a whole coding region. Thus there would appear to be at least 5 genes homologous to leukosialin in the rat genome. The different reactivities of some of the bands may imply some are pseudo genes with divergent sequences.

3.6 Structural features of leukosialin proteins

3.6.1 O-linked glycosylation.

The extracellular domain of leukosialin can be regarded as a mucin structure (Hilkens, 1988). Both leukosialin and mucins have a high molar percentage of Ser and Thr, significant amounts of Ala, Gly and Pro, and low levels of aromatic and sulphur containing amino acids. The extracellular domain of mouse leukosialin is composed of 40% Ser and Thr (compared to about 12% in proteins in general, Wilson et al., 1991), 11% Pro, 9% Ala, and 6% Gly. It does not contain any Phe (in contrast with the rat protein which has a Phe triplet), Tyr, or Cys. A single Trp and 3 Met residues are present. It seems likely that the majority of the 93 Ser and Thr residues in mouse leukosialin will carry O-linked

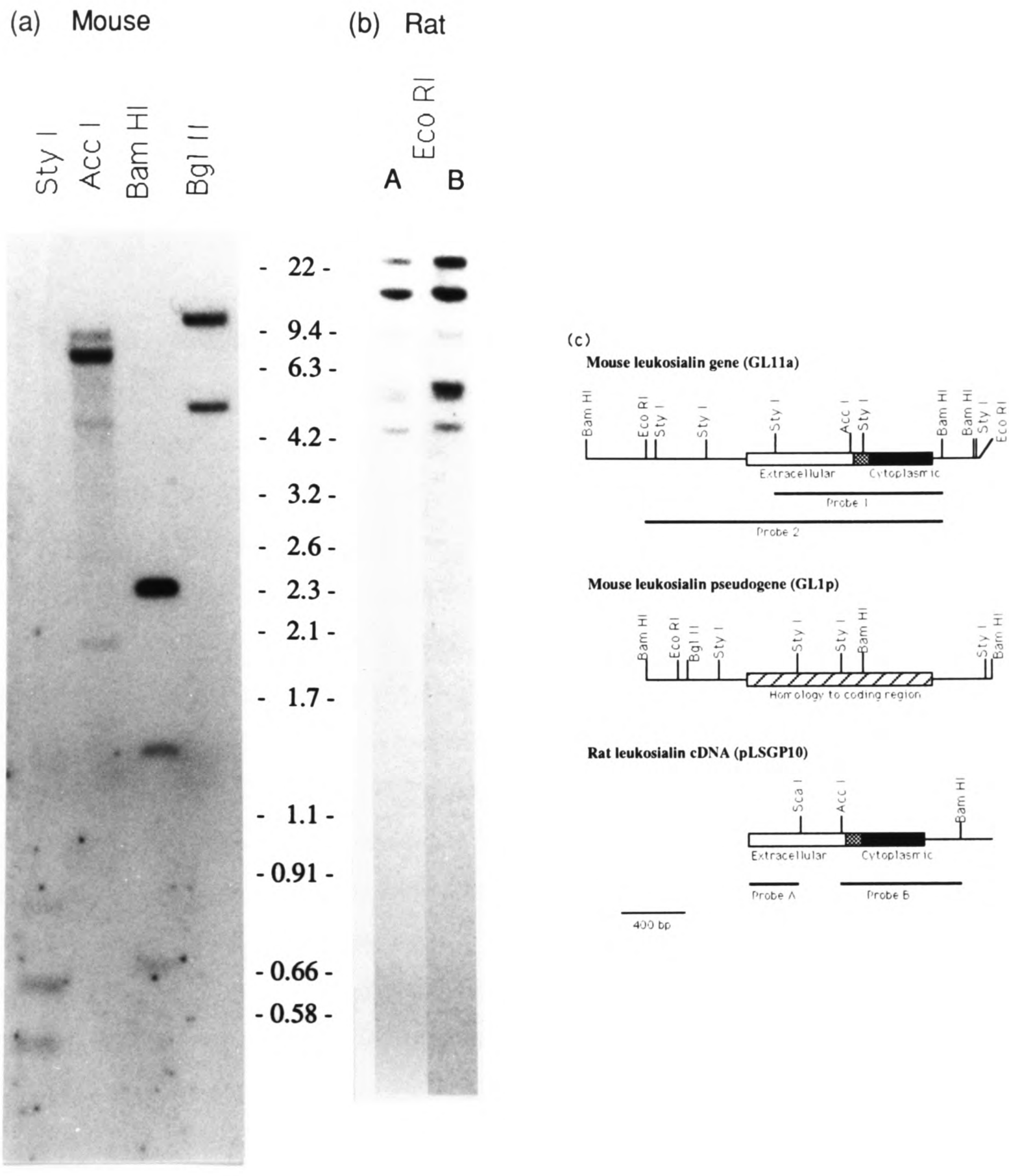


Figure 3.4 Southern blot of mouse and rat genomic DNA to show number of leukosialin genes. (a) Balb/c genomic DNA digested with the indicated restriction enzymes and probed with the cDNA probe (probe 1, section c). The final wash after hybridization was with 0.2x SSC, 0.1 % SDS at 65°C for 30 min. Exposure was at -70°C for 48 hr. (b) Genomic rat DNA digested with Eco RI and probed with cDNA probes specific for 5' (probe A) and 3' (probe B, section c) regions of the gene (this blot was performed by Nigel Killeen). The position of molecular weight markers is indicated in kilobases. (c) Partial restriction enzyme maps of mouse genomic leukosialin clones GL11a and GL1p, and of rat cDNA clone pLSGP-10 (Killeen et al., 1987) showing relevant restriction sites, the region corresponding to the coding sequence, and the fragments used as probes for library screening (probe 1) and Southern blot analysis of mouse DNA (probes 1 and 2) and rat DNA (probes A and B).

carbohydrate structures. This is known to be the case for rat (Brown et al., 1981) and human (Remold-O'Donnell et al., 1987) leukosialin. The mouse molecule is similar to the rat and human forms in analysis on SDS PAGE (Pink, 1983) and in its binding characteristics for peanut lectin (Brown and Williams, 1982). Recent work by Schmid et al., (1992) indicates that 82 of the 92 Ser and Thr in human leukosialin will carry O-glycans.

3.6.2 N-linked glycosylation.

An N-linked glycosylation site that seems likely to be utilised exists in the extracellular domain of mouse leukosialin at residue 148 (Fig. 3.2). A second potential site exists at residue 25 but this is unlikely to be glycosylated due to the presence of a proline at residue 28. An analogous sequence (NCTP) is seen in the rat CD45 antigen and was shown to be unglycosylated (Barclay et al., 1987). Rat leukosialin does not contain any potential N-linked glycosylation sites (Killeen et al., 1987) and it is free of N-linked sugar on the basis that there is no glucosamine found in the composition of the molecule isolated from thymocytes (Brown et al., 1981). The human protein contains a single N-linked structure (Pallant et al., 1989; Shelley et al., 1989) but the position of the mouse and human sites differ (residues 148 and 220, respectively). Although no compositional data is available for mouse leukosialin, the fact that the apparent Mr of this molecule has usually been reported to be higher than the rat molecule (Pink, 1983) is consistent with the presence of an N -linked structure in the mouse protein. Contrasting with this is the finding that mouse leukosialin does not bind to lentil lectin and that its apparent Mr is not affected by N-glycanase treament (6). However, these results do not conclusively argue against the presence of an N-linked site since human leukosialin is also negative for binding to lentil lectin (Remold et al., 1984) and it seems possible that the action of N-glycanase could be inhibited by adjacent O-linked structures. Further work is needed to establish the state of N-glycosylation of the mouse molecule.

3.6.3 Extracellular regions of sequence similarity

Patches of sequence similarity have been seen in the extracellular sequences of rat and human leukosialin. Initially 4 repeat sequences were reported in human leukosialin (Pallant et al., 1989; Shelley et al., 1989) but more recently 5 repeats have been observed (Schmid et al.,

1992). If the stringency of alignment is reduced to 30% identity (without introduction of gaps or insertions) between a given repeat and 2 or more others then 7 repeat elements can be identified in human leukosialin (Fig 3.5a and b). Only the last 5 of these are in a tandem array. As previously reported (Killeen et al., 1987; Cyster et al., 1990) the repeat elements are not as well conserved in mouse or rat leukosialin and only 4 elements can be reasonably extended to 18 amino acids while 3 further segments can be aligned with part of the repeat (Fig 3.5, regions shorter than 7 amino acids were not considered). The most NH_2-terminal repeat in human leukosialin does not align with a repeat like sequence in mouse or rat, whereas there is a sequence in mouse and rat leukosialin starting at amino acid 67 that does not align with a conserved human repeat. An additional finding with mouse leukosialin is the presence of two segments of 8 identical amino acids separated by one amino acid (Fig. 3.5c). The rat sequence and the mouse pseudogene sequence do not show this repeating element. The repeat has apparently resulted from a duplication of 27 nucleotides and this can be seen in the comparison of the functional and pseudogene mouse sequences in Fig 3.1 (nucleotides 382 to 408). This repeat would appear to have arisen after the duplication to produce the two genes now seen in the mouse. Whether or not one of the other genes seen in the rat contains the repeated segment remains to be established.

3.7 Phosphorylation sites.

The cytoplasmic domains of the leukosialin sequences are free of Tyr residues, and thus cannot be targets for tyrosine kinases. However, it has been shown that leukosialin of human peripheral blood mononuclear cells is phosphorylated and that the level of phosphorylation can be upregulated by phorbol esters (Chatila and Geha, 1988). This suggests phosphorylation at Ser or Thr residues by a protein kinase C. Target sites for protein kinase C include Ser or Thr residues in the vicinity of basic amino acids (Turner et al., 1985; Woodgett et al., 1986; House et al., 1987; Kennelly and Krebs, 1991) and in the leukosialin sequences possible sites that are conserved across the species are; KRRT(260), KGS(301) and RRKSR(324) (Fig.3.2). A recent study of human leukosialin demonstrated that only Ser residues were phosphorylated, both in untreated cells, and in cells treated with phorbol esters. In particular, a site at 332 (corresponding to potential site 324 in mouse leukosialin) was shown to be phosphorylated (Piller et al., 1989b). An analysis of protein

A.

```
  -19                -1 1
m  MALHLLLLFGACWVQVASP DSLQRT-TMLPSTPHITAPSTSEAQNASPSVSVGSGTVDSKETISPWGQT  49
r  MALHLLLLFGGFWAQVVSQ ENLPNTMTMLPFTPNSESPSTSEALSTYSSIATVPVTEDPKESISPWGQT
h  MA-TLLLLLG---VLVVSPDALG ST-TAVQ-TPTSGEPLVSTSEPLSSKMYTTSITSDPKADST--GDQ

m  TIPVSLTPLETTELSSLE-TSAGASMSTPVPEPTASQEVSSKTS-ALLPEPSNVASDPPVTAANPVTDGP  117
r  TAPASSIPLGTPELSSFFFTSAGASGNTPVPELTTSQEVSTEASLVLFPKSSGVASDPPVTITNPATSSA
h  TSALP-PSTSINEGSPLW-TSIGASTGSPLPEPTTYQEVSIKMS-SVPQETPHATSHPAVPITANSLGSH

m  AANPVTDGTAASTSISKGTSAPPTTVTTS---SNETSGPSVATTVSSK-----TSGPPVTTATGSLGPSS  179
r  VAS-------TSLETFKGTSAPPVTVTSS----TMTSGPFVATTVSSE-----TSGPPVTMATGSLGPSK
h  TVTGGTITTNSPETSSRTSGAPVTTAASSLETSRGTSGPPLTMATVSLETSKGTSGPPVTMATDSLETST

m  EMHGLPATTATSSVESSSVARGTSVSSRK-----TSTTSTQDP-ITTRSPSQESSG   229
r  ETHGLSATIATSSGESSSVAGGTPVFSTK-----ISTTSTPNP-ITTVPPRPGSSG   224
h  GTTGPPVTMTTGSLEPSSGASGPQVSSVKLSTMMSPTTSTNASTVPFRNPDENSRG   230
```

B.

Human

```
6    QTPTSGEPLVSTSEPLSS
83   MSSVPQETPHATSHPAVP
116  ITTNSPETSSRTSGAPVT
134  TAASSLETSRGTSGPPLT
152  MATVSLETSKGTSGPPVT
170  MATDSLETSTGTTGPPVT
188  MTTGSLEPSSGASGPQVS
```

C.

```
Mouse
     109 AANPVTDG
     118 AANPVTDG
Pseudo
     100 AANPVADR
Rat
     112 ITNPATSS
```

Figure 3.5 Location and sequence of repeats in the extracellular domain of leukosialin and the amino acid sequence of the novel 8 amino acid repeat in mouse leukosialin. A) The locations of seven 18 amino acid repeats in the human leukosialin extracellular domain are shown by thick underlining. The regions of mouse leukosialin which show a similar repeat motif or partial motif are indicated by a thin overline. The motifs in rat leukosialin are similar to those in the mouse. B) Alignment of the 18 amino acid motifs of human leukosialin. Residue numbers indicate the position in human leukosialin with respect to the mature NH_2-terminus (Schmid et al., 1992). C) The novel eight amino acid repeat present in the mouse gene and the corresponding sequence in the mouse pseudogene (Pseudo) and the rat gene. The repeat is shown in bold in (A).

kinase C mediated phosphorylation of the cytoplasmic domain of rat leukosialin is presented in Chapter 6.

3.8 Homology

At the nucleotide level, the murine gene was found to have 82% identity with the rat and 53% identity with the human leukosialin coding region. The level of identity at the amino acid level is similar (Table 3.1). When the extracellular, transmembrane, and cytoplasmic domains are compared separately it is seen that the transmembrane and cytoplasmic domains are much more similar than are the extracellular domains (Table 3.1, Fig 3.2). However, this phenomenon is not seen when the pseudogene is compared to the other sequences. This indicates that strong selection is occurring at the protein level in the cytoplasmic sequences of the coding genes and is suggestive of an important function for this region. In the extracellular domain of leukosialin, the over-riding feature is the high level of O-glycosylation and the sequence is conserved in the sense of providing O-linked sites even though the level of amino acid identity is not high. The extracellular and cytoplasmic domains of the leukosialins were not found to have significant homology with any other sequences in the NBRF database.

Table 3.1. Percentage identities of mouse, rat, and human leukosialin genes and the mouse pseudogene at the nucleotide and amino acid levels[a)].

Alignment	Total coding region	Extra-cellular	Trans-membrane	Cyto-plasmic
Mouse/Rat	82 (74)	76 (64)	87 (87)	91 (90)
Mouse/Human	53 (53)	40 (42)	71 (65)	75 (72)
Rat/Human	59 (54)	48 (37)	69 (70)	77 (73)
Mouse/Pseudo[b)]	81 (66)	82 (68)	89 (83)	76 (57)
Rat/Pseudo	77 (61)	76 (61)	84 (83)	74 (55)
Human/Pseudo	52 (43)	41 (38)	70 (65)	67 (48)

a) Alignments of nucleotides and amino acids were determined using the IALIGN program (Dayhoff et al., 1983). Percentages of nucleotide identity are given together with percentage identity of amino acids in brackets, for the various pairs of sequences indicated. b) Pseudo refers to mouse pseudogene (clone GL1p). The rat and human sequences were taken from references (Killeen et al., 1987) and (Pallant et al., 1989; Shelley et al., 1989).

3.9 Discussion

The lack of introns in the human (Shelley et al., 1990) and mouse leukosialin genes is surprising given the clear demarcation of different domains of structure in this molecule. The result is significant regarding possible genetic heterogeneity in leukosialin, since the lack of introns in the coding sequence rules out the possibility of different forms being produced by alternative splicing of exons. From this data and the evidence for only one functional gene in the mouse, it seems clear that all heterogeneity in mouse leukosialin must be produced by post-translational mechanisms. Direct evidence that heterogeneity of rat leukosialin is generated by post-translational mechanisms is presented in Chapter 5.

The lack of introns in the coding sequence is uncommon for genes of cell surface molecules. Three membrane proteins which have single trans-membrane regions and which are encoded by genes without introns in the coding region are human thrombomodulin (Jackman et al., 1987), the GA733 protein isolated from human gastrointestinal tumours (Linnenbach et al., 1989), and the human blood platelet glycoprotein Ibα (Wenger et al., 1988). CD14 is also similar in having only a single intron located immediately following the initiation ATG (Ferrero et al., 1990). Other integral membrane proteins whose genes lack introns include a number of rhodopsin-related proteins (eg, the beta-2- adrenergic receptor of the hamster and human (Kobilka et al., 1987); and the M2 muscarinic receptor (Peralta et al., 1987)). These molecules all have 7 transmembrane domains and are approximately 450 amino acids in length. Some members of this family, including rhodopsin itself (Nathans and Hogness, 1984), do contain introns and it is possible that the intronless members arose from these interrupted genes by reverse transposition (Weiner et al., 1986).

The origins of genes without introns is unknown. If primordial genes evolved with introns (Gilbert et al., 1986) then the lack of introns can either be due to intron loss or to reinsertion into the genome of genetic material copied from mRNA by reverse transcription (Weiner et al., 1986). The finding of an intron in the 5' non-coding sequence of the human leukosialin gene and the probability that this is also found in the mouse gene may argue for intron loss rather than a mechanism involving reverse transcription. Alternatively leukosialin evolved from an intronless precursor (Dibb and Newman, 1989) and an intron became inserted in the

5' untranslated region by a transposition event. The intron may have been maintained because it conferred some advantageous enhancer activity. It will be instructive to look at the leukosialin gene structure in early vertebrates, particularly as processed genes are largely restricted to mammalian genomes (Wagner, 1986), to address more directly the question of the origin of introns in transmembrane proteins.

The 5' untranslated region and sequence immediately 5' of the transicription initiation site in human leukosialin is highly conserved between human and mouse leukosialin. Recent work by Kudo and Fukuda, (1991) revealed a 14 nucleotide sequence, GGGTGGGTGGAGCC, critically involved in expression from the leukosialin promoter in Jurkat cells and this motif is conserved in the mouse sequence except for the exchange of A for G at the second position. In their study of human leukosialin (Kudo and Fukuda, 1991) they mapped the transcription start site to a position 70 nucleotides upstream of the 3' end of exon 1, in contrast with Shelley et al., (1990) who mapped the major site a further 55 nucleotides upstream (the site indicated in Fig 3.3). Further work is needed to determine if the transcription start site is the same in all cells that express leukosialin and if the 14 nucleotide element (110-123 nucleotides upstream of the 3' end of exon 1, Kudo and Fukuda, 1991) is critical for leukosialin expression in all cells. The high level of sequence conservation between mouse and human leukosialin around this region suggests that other parts of the sequence will be important, at least in some cells. The region immediately upstream of either transcription start site in mouse and human (Shelley et al., 1990; Kudo and Fukuda, 1991) leukosialin is rich in guanine and lacks TATA and CAAT boxes. These are features typical of house keeping genes (Shelley et al., 1990) and may be significant in relation to the very early expression of leukosialin in hematopoietic stem cells (Dyer and Hunt 1981; Hardy et al., 1991).

Internal repeats of Thr and/or Ser and Pro rich sequence occur in most mucin type glycoproteins (Strous and Dekker, 1992). The repeat sequence is often present at very high copy number, the repeats are generally in a tandem array, and there are frequent variations in the number of repeats between alleles (Gum et al., 1990; Ligtenberg et al., 1990; Eckhardt et al., 1991). The leukosialin sequences also show internal repeats, although they are not as

highly conserved as those in the larger mucin glycoproteins, and it may be that the whole extracellular domain has derived from a short precursor fragment. It appears that some repeat events have occurred in human leukosialin and rodent leukosialin after the separation of the species. In particular, one repeat of 8 amino acids present in mouse leukosialin is clearly absent from rat and human leukosialin. Despite the low sequence identity and the apparent occurrence of repeat events at different times in evolution, the extracellular domains are all of similar length, being 235 amino acids in human, 224 in rat and 229 in mouse leukosialin. This may indicate a need for conservation of molecular length not seen in the large mucins expressed at epithelial surfaces. There is a strong correlation in heavily O-glycosylated molecules between molecular extension and number of amino acids (described in detail in Chapter 5) and further accumulation of repeat sequences may be selected against because this would increase the molecular extension and add significantly to the amount of sialic acid at the periphery of the leukocyte. This in turn would have implications for the ability of the leukocyte to undergo adhesion interactions. It will be interesting to study the leukosialin molecule in more divergent species to see if the length of the extracellular domain has been conserved and to address the wider issue of the evolution of molecular extension at the cell surface.

The high level of sequence conservation over the entire length of the leukosialin cytoplasmic domain is strong evidence that the domain interacts with one or probably multiple intracellular proteins. The domain is more conserved than the CD2 cytoplasmic domain (68% over 116 amino acids) for example, which is essential for signal transduction through CD2 (He et al., 1988; Beyers et al., 1991). An expression study of the leukosialin cytoplasmic domain is reported in Chapter 6.

It is demonstrated in the next chapter that clone GL11a is a functional gene. Transfection of the leukosialin gene into cells leads to expression of a membrane protein that reacts with mAb previously predicted to react with mouse leukosialin.

CHAPTER 4

EXPRESSION OF LEUKOSIALIN ON B CELLS IN TRANSGENIC MICE

4.1 Introduction

4.2 Construction of modified mouse and rat leukosialin genes to target expression to B cells of transgenic mice

4.3 Expression of mouse and rat leukosialin in cell lines

4.4 Generation and screening of leukosialin transgenic mice

4.5 Analysis of transgene expression by flow cytometry

4.6 Expression of mouse leukosialin transgene in different tissues

4.7 Increased ratio of B cells to T cells in the spleen and lymph nodes of mouse leukosialin transgenic mice

4.8 Mouse leukosialin transgenic mice show normal antibody responses to complex antigens and *in vitro* proliferative responses to LPS

4.9 Construction of a targeting vector for interruption of the mouse leukosialin gene by homologous recombination in embryonic stem cells

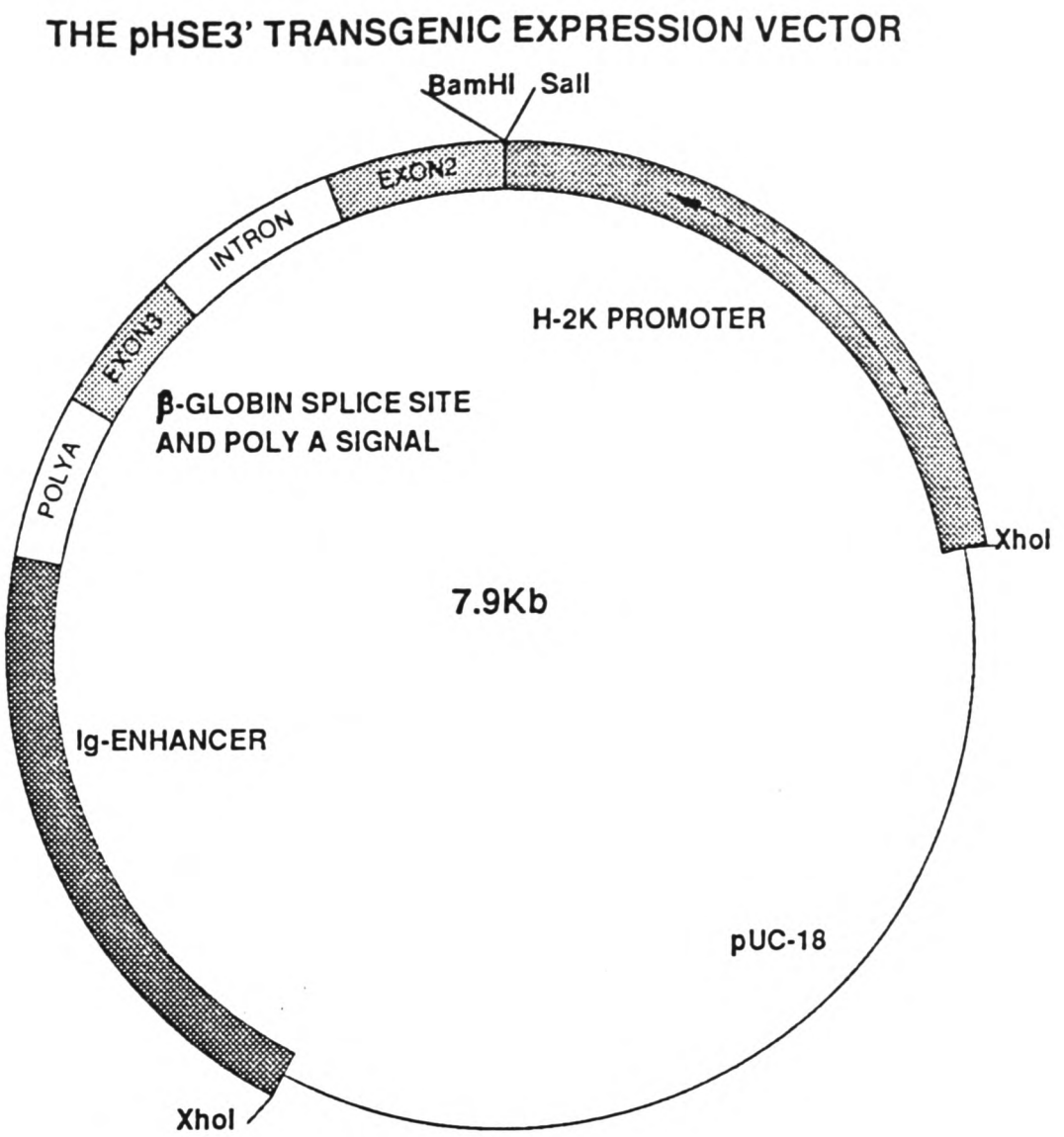
4.10 Discussion

4.1 Introduction

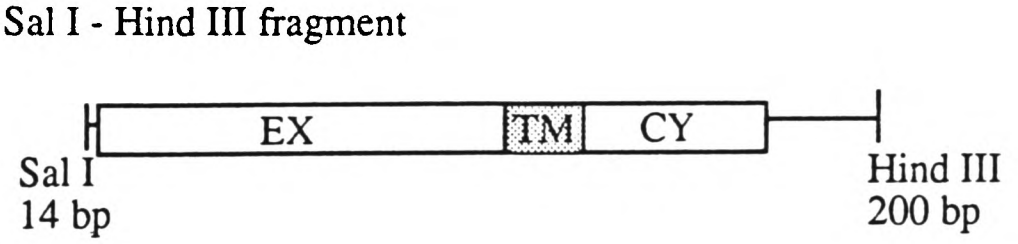
The observation that B cells have a lower electrophoretic mobility than T cells (Nordling and Anderson, 1972; Shortman et al., 1975) and that this difference may be due to differences in the disposition of cell surface sialic acid (Despont et al., 1975) may in turn be a result of the lack of leukosialin expression on resting B cells (Williams et al., 1977; Remold-O'Donnell et al., 1987). Leukosialin has been estimated to carry approximately half of the sialic acid attached to glycoproteins on T cells (Brown et al., 1981; Williams and Barclay, 1986). B cells also migrate more slowly through lymph nodes than T cells (Howard, 1972), and show a greater stickiness for nylon wool (Julius et al., 1973). The highly glycosylated and charged nature of leukosialin and its very extended structure suggest that the molecule may have major effects on cell-cell interactions in general and on cell migratory behaviour in particular. The hypothesis was made that expression of leukosialin on resting B cells may alter their migratory behaviour, leading to a migration pattern more similar to T cells. To test this hypothesis leukosialin was expressed on B cells in transgenic mice. The first steps towards a second approach of examining function *in vivo*, interruption of the leukosialin gene in the genome by homologous recombination, are described in the last section of this chapter.

4.2 Construction of modified mouse and rat leukosialin genes to target expression to B cells of transgenic mice

The generation and use of transgenic mice has been reviewed extensively (Palmiter and Brinster, 1986; Storb, 1987; Hanahan, 1989). The pHSE3' vector (Ferrick et al., 1989; Pircher et al., 1989) which contains the enhancer from the JH-CH intron of the Ig heavy chain (Fig 4.1a), was chosen as the means of targeting leukosialin expression to B cells in transgenic mice. Immunoglobulin transgenes show a high level of expression in B cells, a lower level in T cells and occassional expression in other tissues (Gerlinger et al., 1986; Chen et al., 1987; Reik et al., 1987; Storb, 1987) and expression is largely position and copy number independent (Storb, 1987). The pHSE3' vector does not contain all the enhancer activity associated with Ig heavy chains (Pettersson et al., 1990) and this may be a cause for its apparent copy number dependence (Pircher et al., 1989). The enhancer fragment used does contain negative regulators that reduce expression in non-lymphoid tissue (Imler et al., 1987; Libermann et al., 1990; Wang et al., 1991), however, and expression from this vector, which also includes the H-2K



B. Rat leukosialin Sal I - Hind III fragment



Mouse leukosialin Sal I - Bam HI fragment

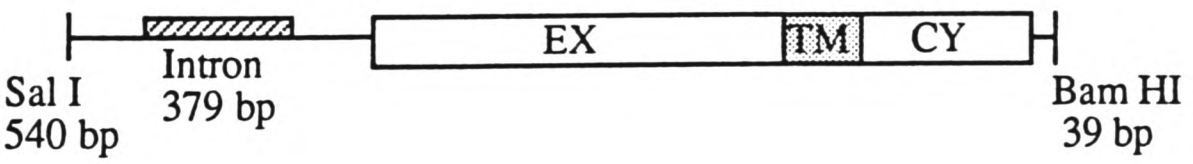


Figure 4.1 The pHSE3' transgenic expression vector and the coding fragments of mouse and rat leukosialin used in the transgenic constructs. A) The pHSE3' vector (Ferric et al., 1989) has a H-2K promoter 5' of the cloning site and human β-globin exons 2 and 3 with the intervening intron and splice sites and the poly A signal 3'. Further distal is the Ig heavy chain enhancer. These components are present within the pUC-18 plasmid which contains an origin of replication and the ampicllin resistance gene. B) The fragments of the rat leukosialin cDNA (Killeen et al., 1987) and mouse leukosialin gene inserted into the pHSE3' vector are illustrated and the amount of 5' and 3' untranslated sequence present in each construct are shown. EX, TM and CY correspond to the extracellular, transmembrane and cytoplasmic sequences, respectively.

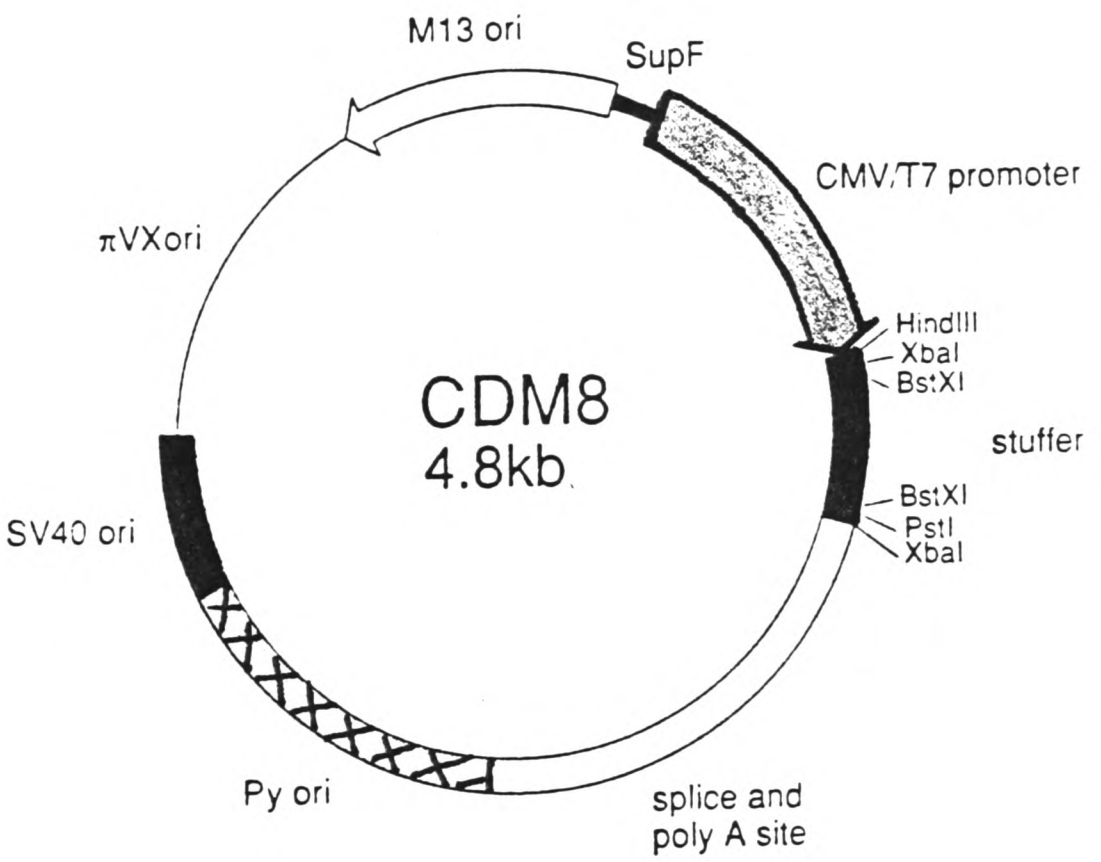


Figure 4.2 CDM8 vector. Transcription of DNA inserted in place of the stuffer DNA is driven by the cytomegalovirus (CMV) promoter and enhancer. A splice site and poly A signal is present 3' of the stuffer. The SV40 virus origin of replication (SV40 ori) allows the vector to replicate as episomal DNA to a high copy number in cells which express the SV40 T antigen. The polyoma virus ori (Py ori) can by used to replicate the plasmid in cells expressing the polyoma large T antigen. Bacterial plasmid replication uses the pVT ori of pBR322. Selection in bacteria relies on the presence of the Sup F tRNA gene, to read nonsense mutations in the tetracycline and ampicillin resistance genes encoded in the p3 plasmid, which is stably transformed into the bacteria MC1061/p3.

promoter, has a similar pattern to that of Ig transgenes (Pircher et al., 1989; Ferrick et al., 1989).

To allow subcloning of the mouse leukosialin gene, including the intron in the 5' untranslated region, a Sal I site was introduced by site directed mutagenesis at the position corresponding to the start of transcription in the human leukosialin gene (Chapter 3, Fig 3.3 and Shelley et al., 1990). In the same mutagenesis step the internal Xho I site (Fig 3.4c) was removed by a single base change (Chapter 2, section 2.5). The gene fragment from the Sal I site to the first Bam HI site 3' of the open reading frame was isolated and inserted into the pHSE3' vector (Fig 4.1). The rat leukosialin cDNA clone originally isolated was truncated in the leader sequence (Killeen et al., 1987) and was extended using PCR with a 5' oligonucleotide based on the mouse sequence immediately upstream of the initiation ATG (Chapter 2, section 2.8, Fig 2.2) that also contained a Sal I site. This created a product that incorporated 14 nucleotides of mouse 5' untranslated sequence followed by the rat cDNA upto the Hind III site 200 bp 3' of the stop codon. The Sal I- Hind III fragment was isolated and blunt end treated and this was inserted into the pHSE3' vector which had been cut with Bam HI and blunt end treated (Fig 4.1). The mouse leukosialin Sal I-Bam HI gene fragment was also blunt end treated and both mouse and rat leukosialin were inserted into the CDM8 expression vector as blunt end fragments (Fig 4.2).

4.3 Expression of mouse and rat leukosialin in cell lines

Expression of the mouse and rat genes was initially tested in COS cells using a transient expression protocol (Seed and Aruffo, 1987; Mallett et al., 1990). Three days after transfection with the leukosialin genes in the CDM8 vector the cells were stained with mAb specific for mouse or rat leukosialin and high levels of expression were evident (Fig 4.3). Subsequently the mouse leukosialin-pHSE3' construct was cotransfected with the pKG5 vector, which contains the neomycin resistance gene, into Y3 rat myeloma cells and the cells selected in medium containing 0.8 mg/ml neomycin. After 7 days 8% of cells were stained with the S7 mAb and this increased spontaneously to 62% after 5 weeks (Fig 4.4a). The rat leukosialin-pHSE3' construct was cotransfected with pKG5 into the mouse myeloma line NSO. Following 2 weeks of selection very few cells remained viable so the cells were plated with thymocyte feeders to rescue the small numbers of drug resistant cells. 80% of cells were labelled by the W3/13 mAb

Transient expression of rat and mouse leukosialin in COS cells

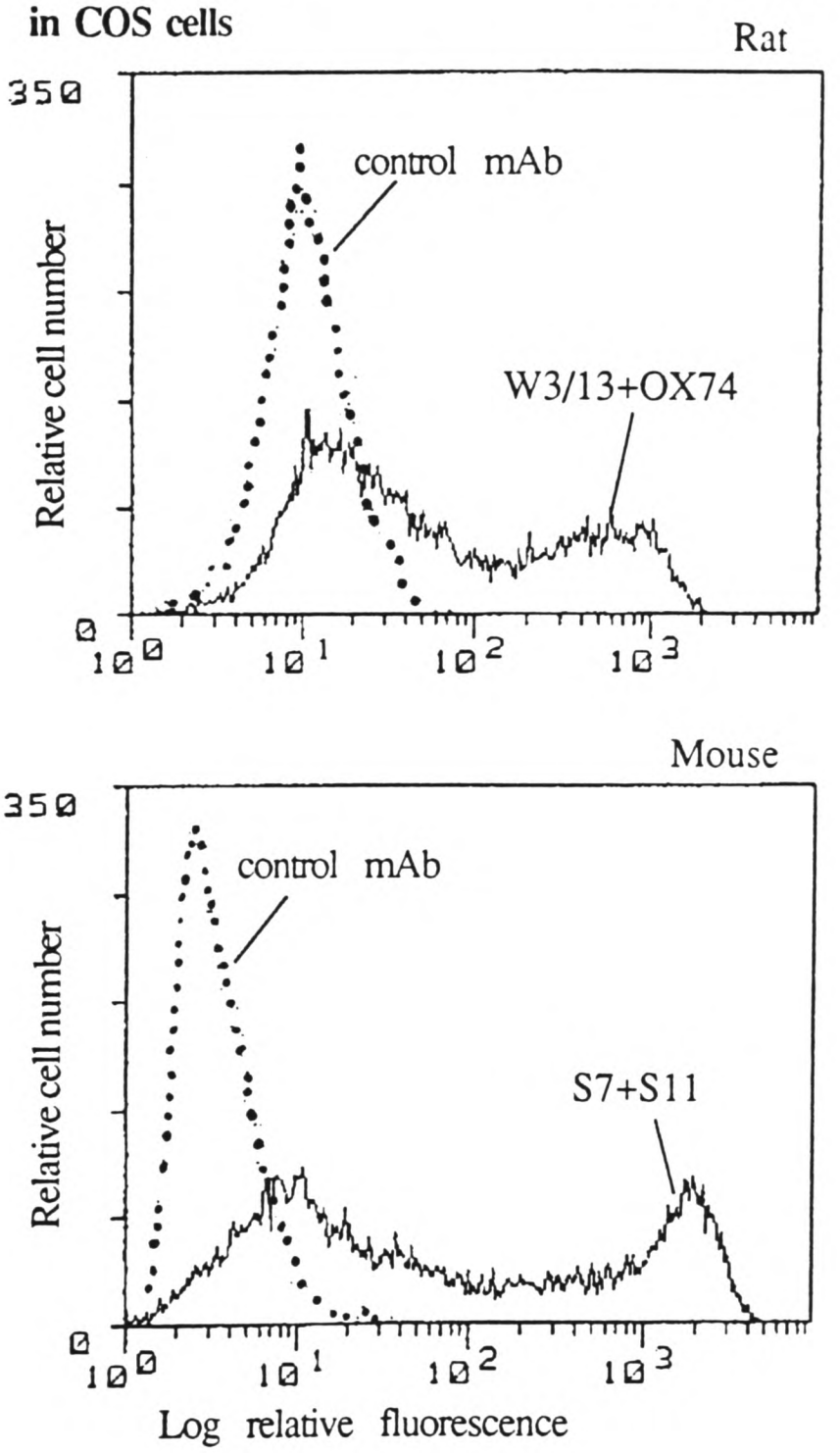


Figure 4.3 Transient expression of rat and mouse leukosialin in COS cells. COS cells were transfected with rat or mouse leukosialin in the CDM8 vector by the DEAE dextran method and after 3 days of culture the transfected cells were released from the flask with PBS/EDTA and 10^6 were labelled with saturating levels of OX21 or NDS58 (control mAb) or the indicated mAbs. Following incubation with FITC-labelled $F(ab')_2$ anti-mouse or anti-rat Ig antibody, the bound antibody was measured by flow cytometry. Mock transfected COS cells stained with any of the indicated mAb gave a profile similar to that shown for the control mAb.

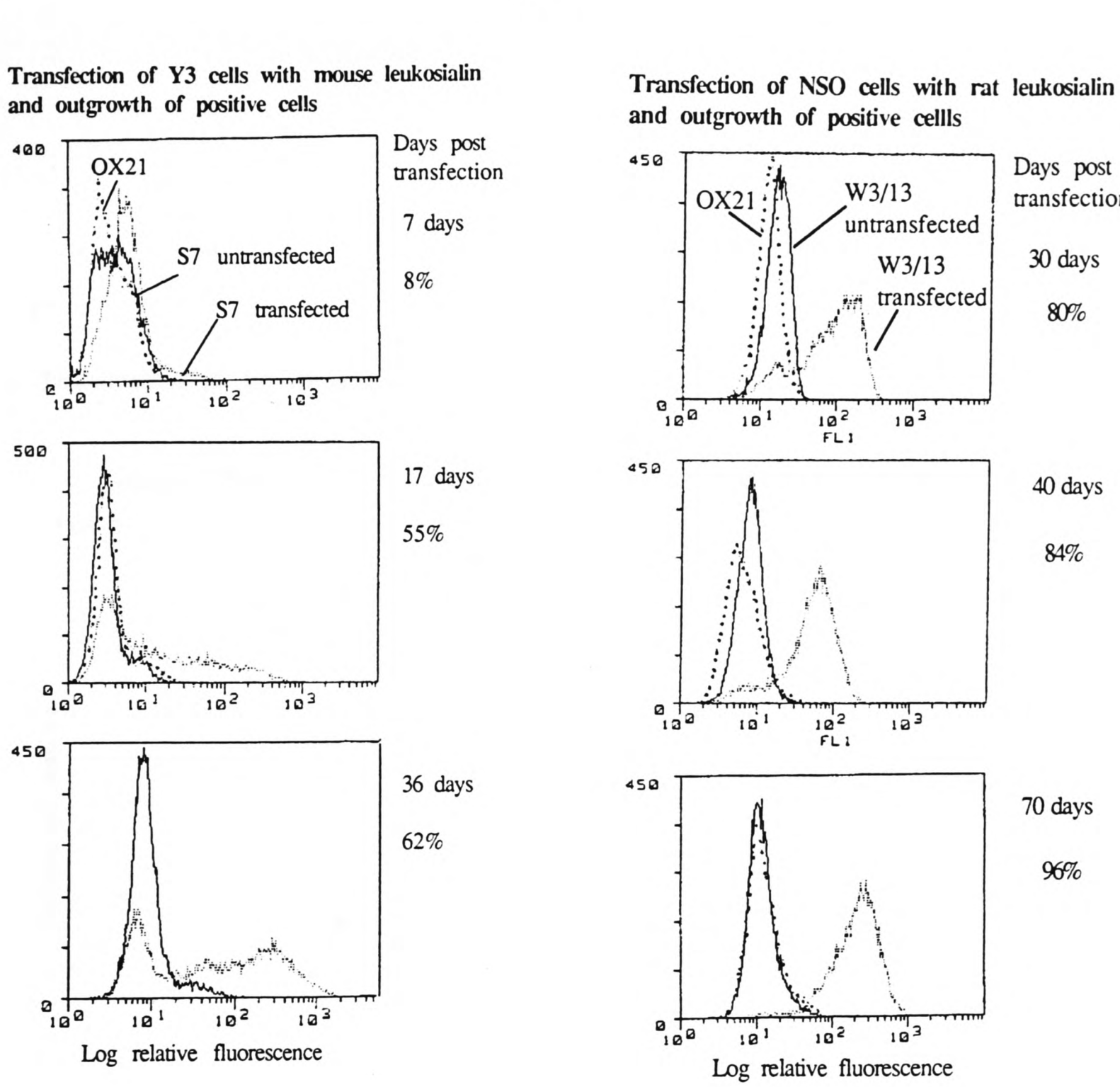


Figure 4.4 Expression of the mouse leukosialin-pHSE3' construct in Y3 cells (A) and the rat leukosialin-pHSE3' construct in NSO cells (B) and Jurkat cells (C). Cells were taken at the indicated time points after the transfection and labelled as described for Fig 4.3. Two or three profiles have been superimposed at each time point as indicated. The percent of cells which were strongly labelled by the mAb is shown. In (C) the 10% of Jurkats which were the most strongly labelled with W3/13 were isolated by FACS. The intensity of W3/13 staining at 8 and 30 days following the sorting is shown with the approximate number of positive cells.

C.

Transfection of Jurkat cells with rat leukosialin and outgrowth of negative cells

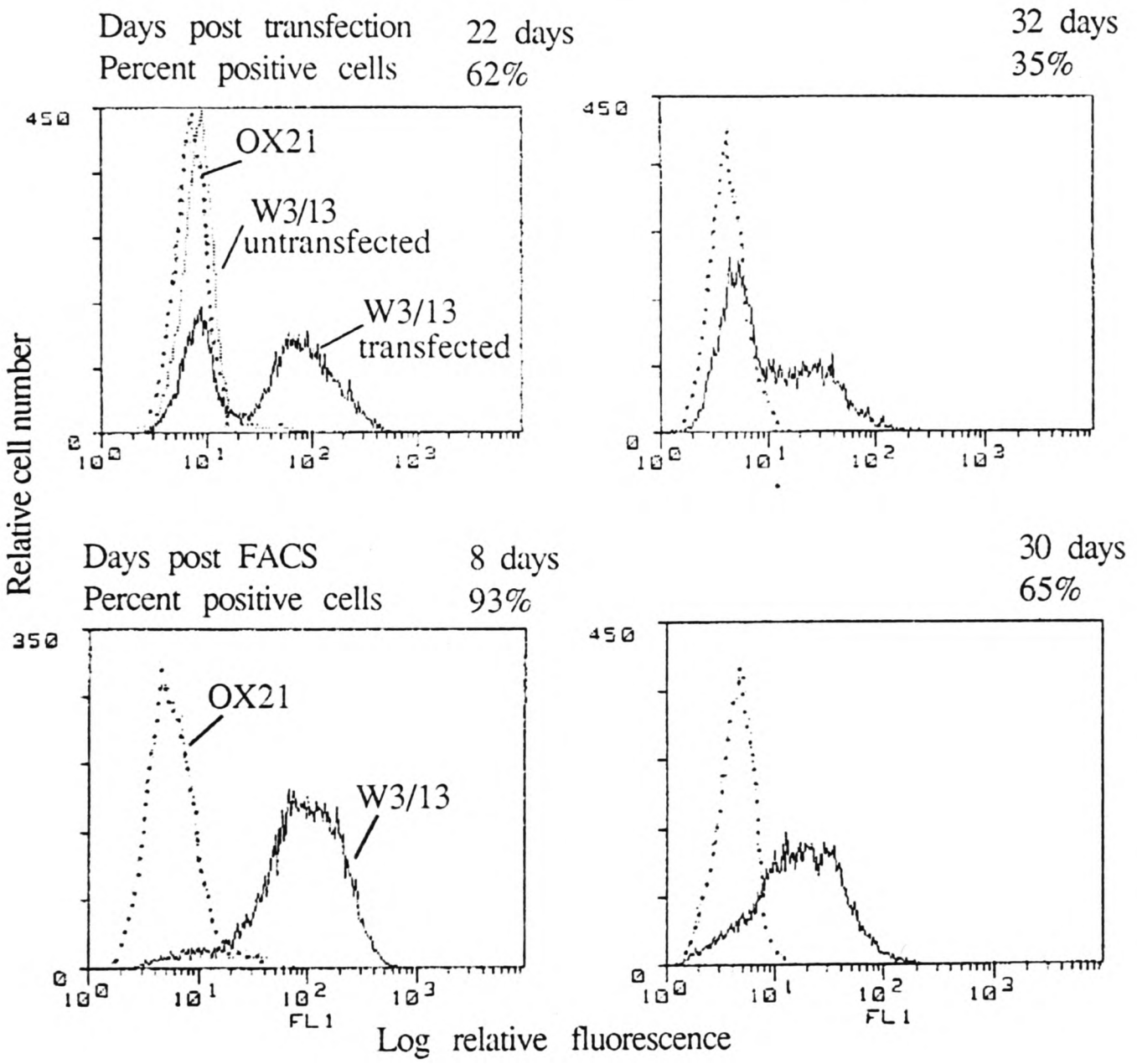


Figure 4.4 continued

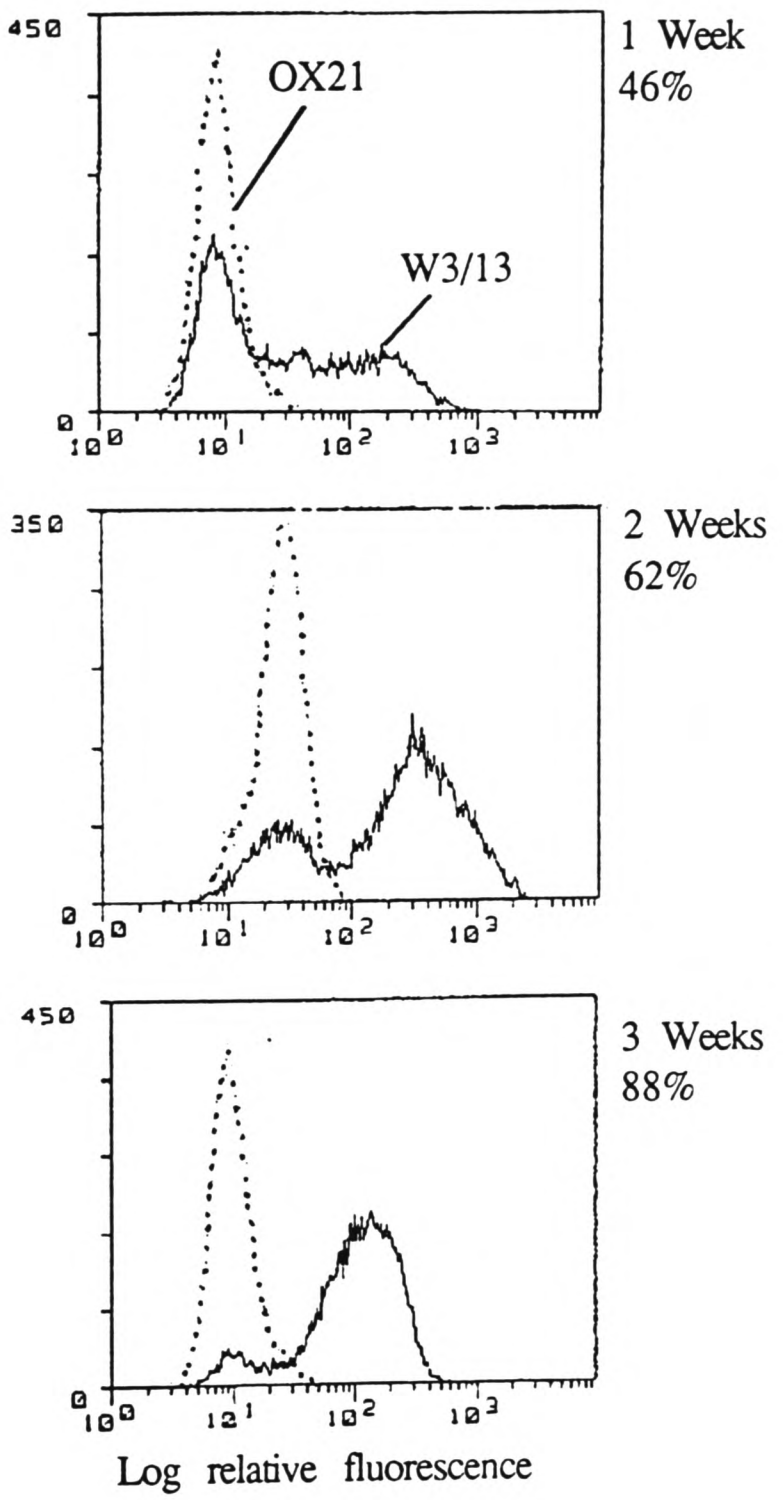


Figure 4.5 Outgrowth of NSO cells by NSO cells expressing rat leukosialin from the pHSE3' vector. 1.7x 10^6 transfected NSO cells were mixed with 1.7x 10^6 untransfected NSO cells and the cells were cultured in 10 ml non-selective medium in a 25cm^2 flask. The medium was replaced every 3-4 days as it began to yellow and the cells were split to keep numbers below 2x 10^6/ml. Samples were taken at the indicated time points and analyzed by flow cytometry as in Fig 4.3.

after a further 2 weeks of culture and the level of expression continued to increase with 96% of cells being strongly positive at 10 weeks (Fig 4.4b). The construct was also transfected into the human T lymphoma line, Jurkat, and 62% of cells stained with the mAb after 3 weeks of selection (Fig 4.4c) but expression decreased with continued culture. Transfected Jurkat cells sorted for W3/13 expression by FACS were also observed to lose expression with extended culture (Fig 4.4c).

The ability of leukosialin transfected NSO cells to outgrow their non-transfected counterparts was examined in non-selective medium. Equal numbers of untransfected and transfected NSO cells were mixed in non-selective medium and after 3 weeks the fraction of cells staining with the W3/13 mAb had increased from 46% to 88% (Fig 4.5). It is not clear if this effect is due to the expression of leukosialin in these cells or to the introduction of the pHSE3' and pKG5 vectors. An experiment in which a different cell surface molecule is introduced in the pHSE3' vector will be needed to control for this possibility.

4.4 Generation of leukosialin transgenic mice

The rat and mouse leukosialin genes with the 5' H-2K promoter sequences and the 3' β-globin sequence containing the intron and poly A signal and the Ig-enhancer were isolated from the prokaryotic elements of the pHSE3' vector by Xho I digestion (Fig 4.1) and gel fractionation. The purified constructs were sent to Dimitris Kioussis (Mill Hill, London) who microinjected the DNA fragments into the pronuclei of one-cell embryos from CBA/ca mice. These were introduced to the oviducts of pseudo-pregnant foster mothers to complete development. Ten days following the birth of these mice tail DNA was isolated and Southern blot analysis performed (Chapter 2). Of the mice born to the foster mothers 4 of 13 carried the rat transgene and 7 of 52 the mouse transgene (Table 4.1). The low frequency of mice carrying the mouse leukosialin transgene may have been because the foster mothers were pregnant rather than pseudo-pregnant. Table 4.1 lists the approximate transgene copy number in the mice estimated from the Southern blots and shows that two lines of rat leukosialin transgenic mice were established (R3 and R4) and 4 lines of mouse leukosialin transgenics (M5, M7, M9 and M10).

Table 4.1 Transgenic founder mice and transgene transmission

Founder	Copy No.*	Transmission	Expressed
R1	5-10	No	?
R2	1	No progeny	?
R3	10-20	Yes	+
R4	10-20	Yes	+
M5	1	Yes	-
M6	1	No progeny	?
M7	2	Yes	++
M8	1	No progeny	?
M9	5-10	Yes	+
M10	10-20	Yes	++
M11	5-10	Died	?
M12	20-40	Died	?

R lines carry the rat leukosialin transgene and M lines the mouse leukosialin transgene. * copy number of the transgene was roughly estimated by Southern blotting and comparison of band intensities with those derived for the endogenous gene (as in Fig 2.4).

4.5 Analysis of transgene expression by flow cytometry

Lymphoid cells from heterozygous transgenic mice and non-transgenic littermates were labelled with anti-leukosialin mAbs (Figs 4.6 and 4.7). The lymph node and spleen cells from the rat leukosialin transgenic lines, R3 and R4, showed expression of rat leukosialin on a subpopulation of cells. After double staining with W3/13 and OX-20, the transgene expression was found to be restricted to the non-B cell population (Fig 4.6). Staining with W3/13 and anti-thy 1 confirmed that the W3/13 positive cells were all T cells. W3/13 also labelled a subpopulation of thymocytes in both transgenic lines (30% in R3 and 60% in R4) while no labelling of bone marrow cells was evident. Breeding of the R4 line was continued and this pattern of expression in the lymphoid system remained unchanged over successive generations.

Three of the four mouse leukosialin transgenic lines showed increased anti-mouse leukosialin mAb (S7) staining of their spleen and lymph node cell populations compared to non-transgenic controls. The M7 and M10 lines showed significant expression of leukosialin on their B cell populations (Fig 4.7a and b) whereas the M9 line showed increased expression on only a fraction of the B220 positive cells. The forward scatter profiles of the B cell populations in mice of the M7 and M10 lines did not differ detectably from controls. From Fig 4.7b it is seen that the T cell populations also show increased leukosialin expression. Thymocytes of M7 mice did not show a significant increase in expression whereas in the M10 line the level of expression was increased to a similar extent to the increase on mature T cells in M7 mice (Fig 4.7). Bone marrow cells of M7 mice showed an increase in staining of the intermediate population but no increase in the brightest population (bone marrow cells of M10 mice have not been examined). This pattern of expression in the three lines was stably transmitted to the progeny.

4.6 Expression of the mouse leukosialin transgene in different tissues

Tissue sections from organs of M7 and M10 transgenic mice were stained with the S7 mAb and some examples are shown in Fig 4.8. A generalized pattern of expression was seen in mice of the M10 line with cells in all the non-lymphoid tissues tested (liver, heart, kidney, small intestine, stomach and brain) staining with S7 in addition to lymphoid organs. In the heart, brain and gut sections the cells stained were very elongated and may be endothelial cells. In the kidney sections staining was heaviest in the glomeruli and in the small intestine staining was

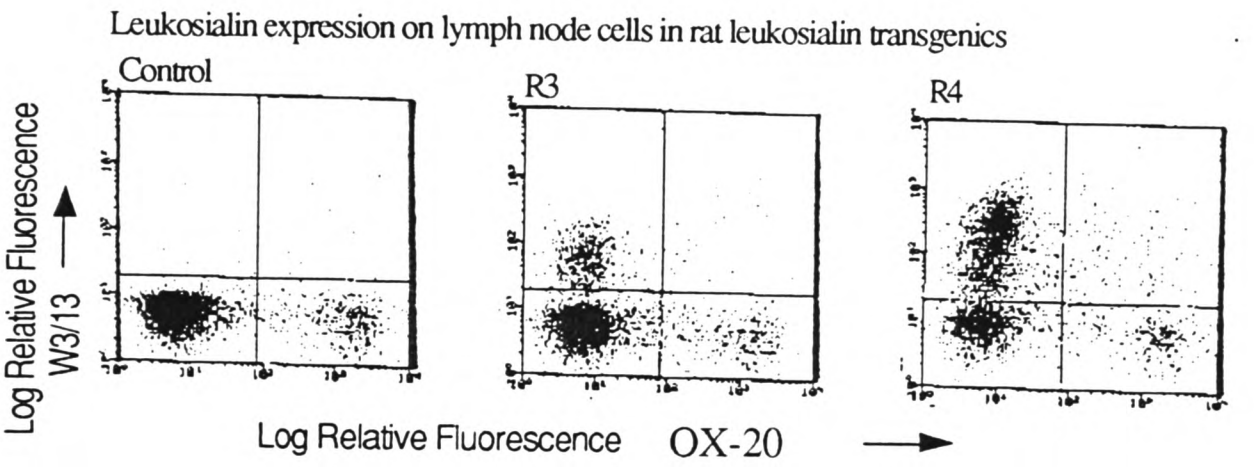
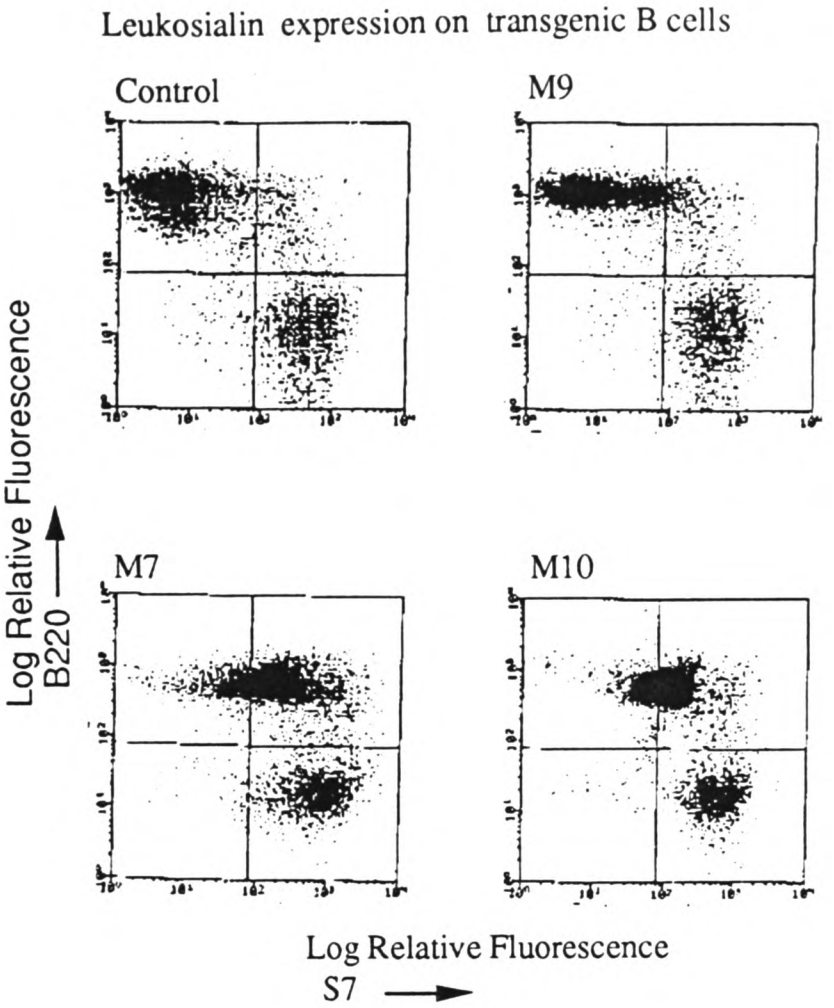


Figure 4.6 Expression of rat leukosialin on spleen and lymph node cells of transgenic mice analyzed by flow cytometry. Spleen and lymph node cells were isolated from transgenic mice of the R3 and R4 lines or from littermate controls and 10^6 were incubated with FITC-conjugated OX-20 (anti-mouse κ chain) followed by 10% normal mouse serum and then W3/13. Biotinylated YA9 (anti-rat IgGγ1) and PE-labelled streptavidin were used to detect the bound W3/13. For single colour staining cells were incubated with W3/13 followed by FITC-labelled YA9.



B. Leukosialin expression in mouse leukosialin transgenics

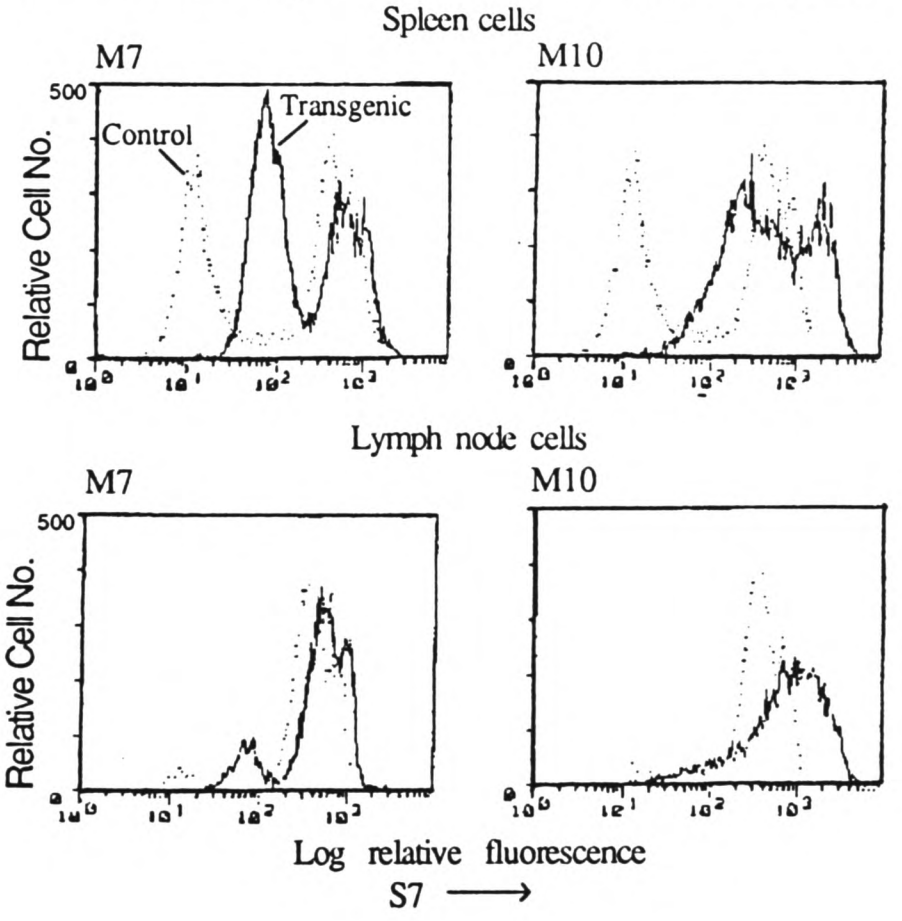


Figure 4.7 Expression of mouse leukosialin on B cells in transgenic mice analyzed by flow cytometry. A) Spleen cells were isolated from transgenic mice of the M7, M9 and M10 lines or from littermate controls and 10^6 were incubated with anti-leukosialin mAb S7 and then with FITC-labelled anti-rat IgGγ2 followed by PE-conjugated anti-B220 mAb Ly5. B) Spleen and lymph node cells from mice of the M7 and M10 lines were incubated with S7 and then FITC-labelled anti-rat IgGγ2.

evident on the epithelium of the lamina propria but not on the microvilli. Expression in the M7 line, in contrast, was detected only in the lymphoid tissues, including the lymphocytes in the sections of small intestine. Northern blotting of RNA from thymus, spleen and liver of mice of the M7 line showed that transcription of the transgene was also restricted to the lymphoid tissues (Fig 4.9).

4.7 Increased ratio of B cells to T cells in the spleen and lymph nodes of mouse leukosialin transgenics

The relative number of cells staining for the B cell markers Ig or B220 was slightly elevated (5-10%) in the spleen and lymph nodes of M7 and M10 transgenic mice when compared to T cell numbers (Fig 4.10; this is also seen with S7 staining alone if Fig 4.7b is studied). A similar relative decrease as a fraction of total cells was seen in both the CD4 and CD8 compartments of the T cell population (Fig 4.10). The difference was less obvious in the lymph node populations but this may have been due to the preparations being a mixture of cervical and mesenteric lymph nodes, with the relative contribution of each type of node varying with each isolation. The size of spleens or lymph nodes from the transgenic mice did not differ detectably from the organs of their non-transgenic counterparts and the number of cells isolated from the spleens were equivalent. The small size of the change in the relative number of B cells vs T cells means it would be difficult to determine if the lymphocyte numbers of transgenic spleens were this much larger than their littermates or if the total number of lymphocytes were unchanged and the B cells had displaced a small number of T cells. Analysis of spleen and lymph node sections of transgenic and control mice did not show any notable increase in the size of B cell areas or increased numbers of B cells in the T cell areas or red pulp.

4.8 Mouse leukosialin transgenic mice show normal antibody responses to complex antigens and *in vitro* proliferative responses to LPS

The primary antibody response of M7 transgenic mice to DNP-coupled bovine gamma globulin (BGG) was similar to that of littermate controls (Fig 4.11a). The antibody response of transgenic mice to sheep red blood cells 7 days following a primary injection and 4 days following a secondary injection did not differ significantly from controls (Fig 4.11b) with all mice showing increased antibody production following the secondary injection.

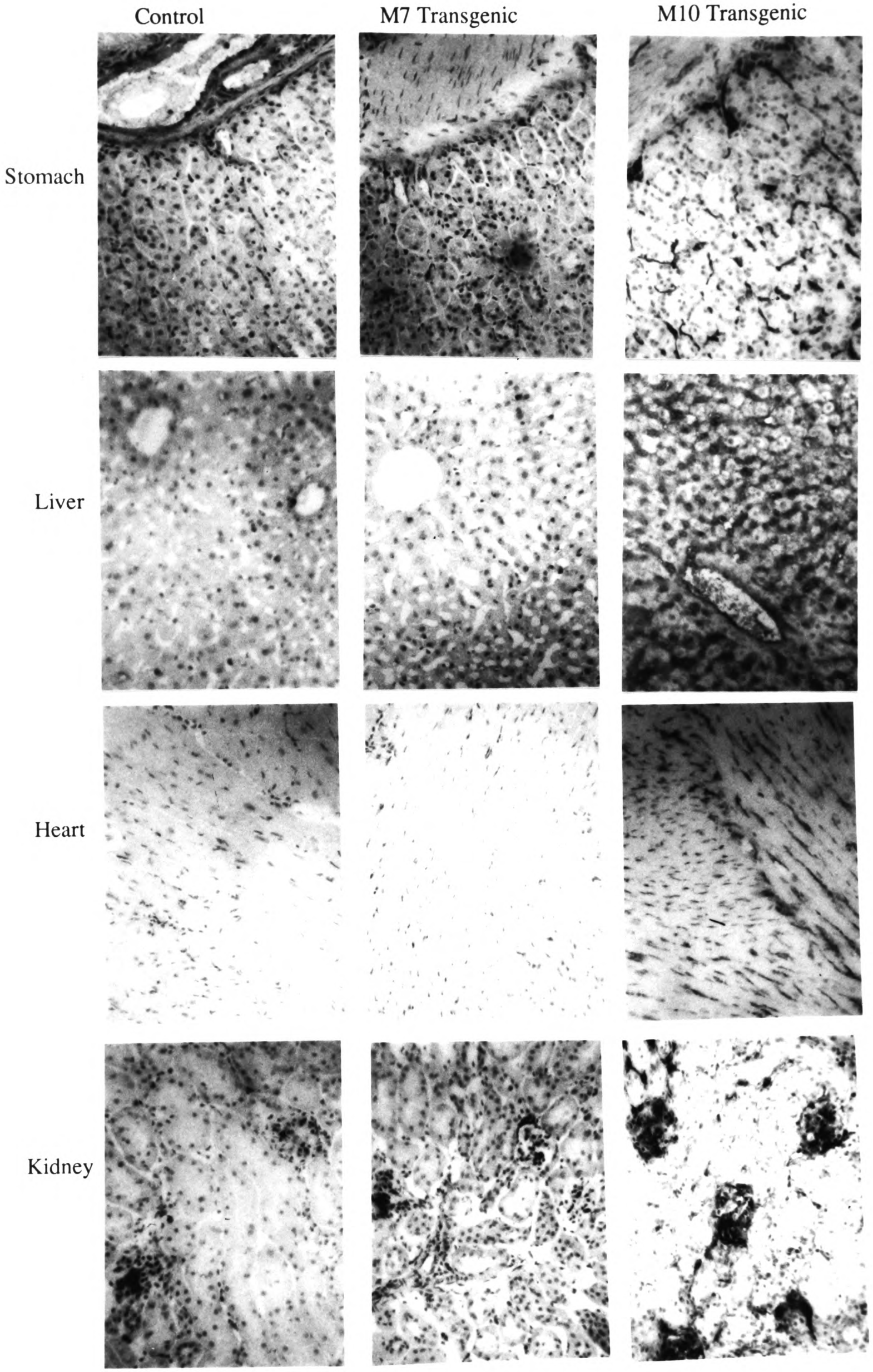


Figure 4.8 Detection of leukosialin expression on cryostat sections of organs from transgenic and control mice by immunoperoxidase staining. Sections were cut and stained as described in Chapter 2. All sections were counterstained with hematoxylin.

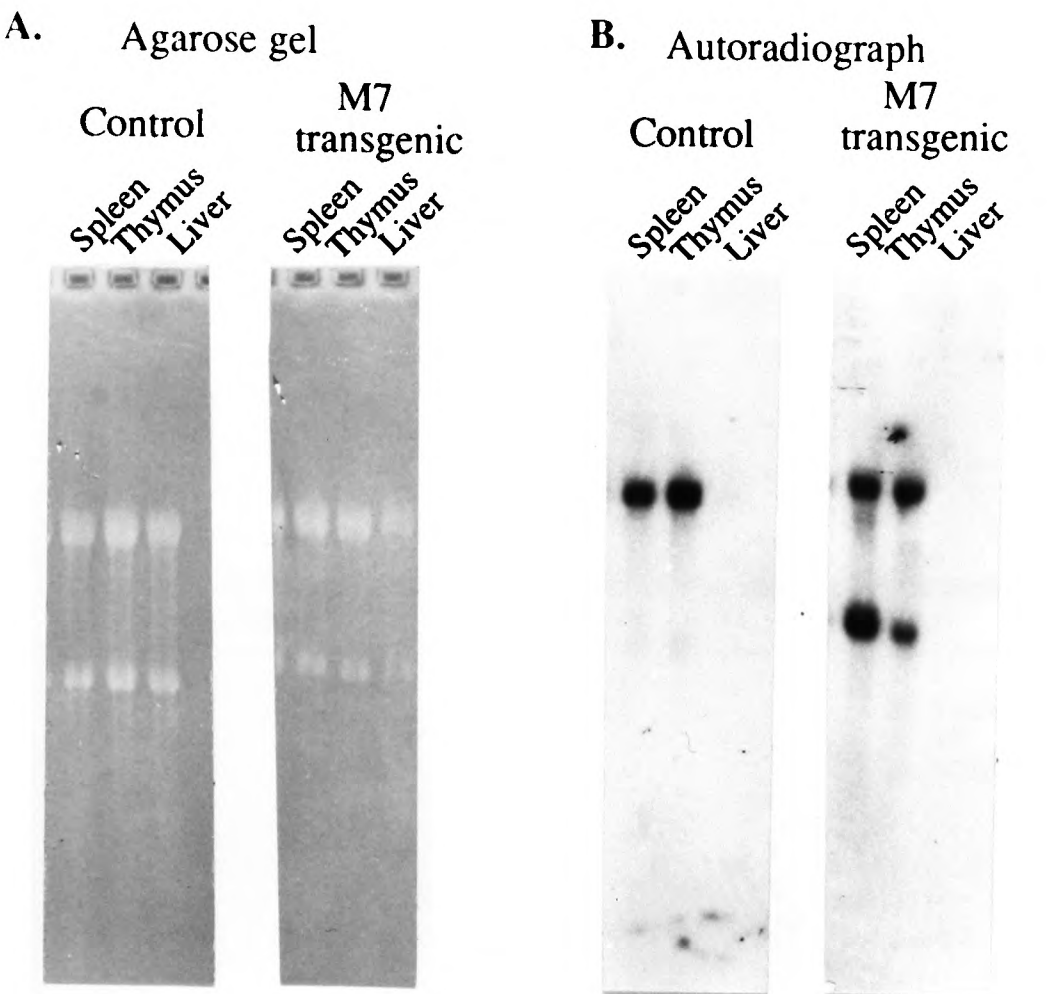
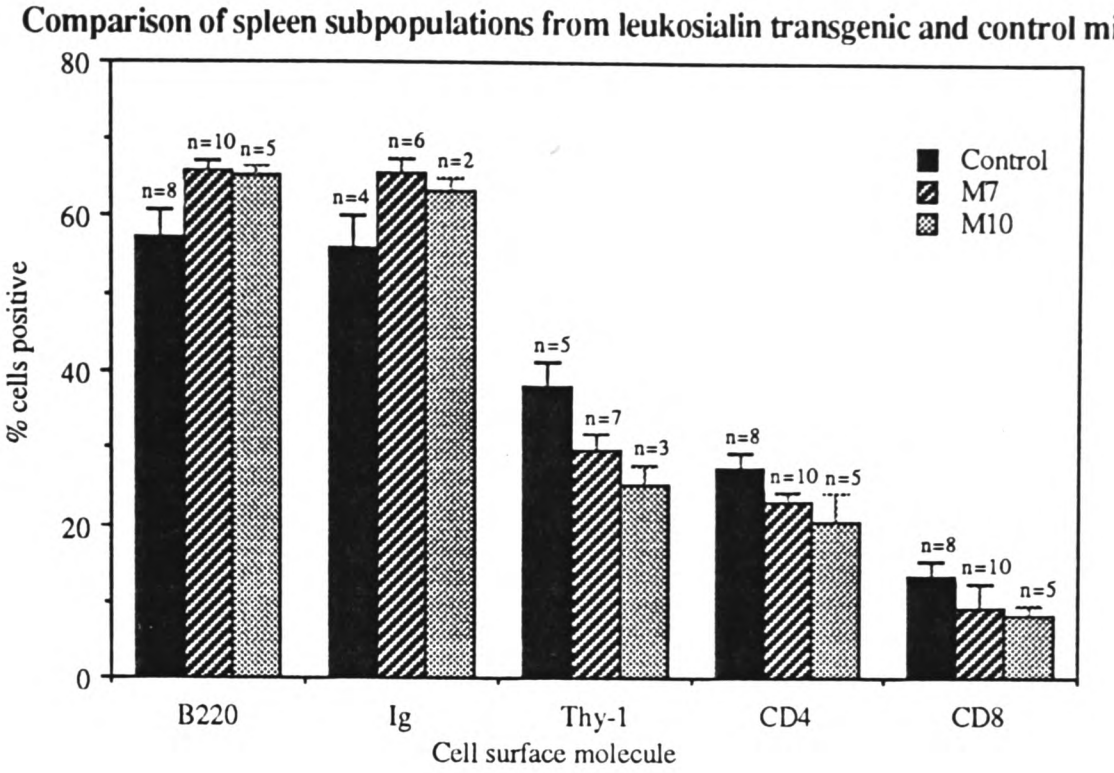


Figure 4.9 Northern blot detection of transgene RNA in spleen and thymus. RNA was prepared by the acid guanidine thiocyanate phenol chloroform method from homogenized spleen, thymus or liver of 8 week old control or M7 transgenic mice. A) RNA samples (10 μg) were electrophoresed on a 1.2% agarose gel in formaldehyde and the 28 S and 18 S RNA is evident. B) The RNA was transferred to nitrocellulose using a pressure blotter and the filter was baked at 80°C for 2 hr and hybridized o/n at 42°C with a probe spanning the extracelluar region of mouse leukosialin and the filters were washed at a stringency of 0.1 x SSC at 60°C for 30 min and exposed at -70°C for 24 hr. The expected size of the transgene transcript is approximately 1.9 kb and the size of the endogenous transcript is 4.2 kb (Baecher et al., 1990).

A.



B.

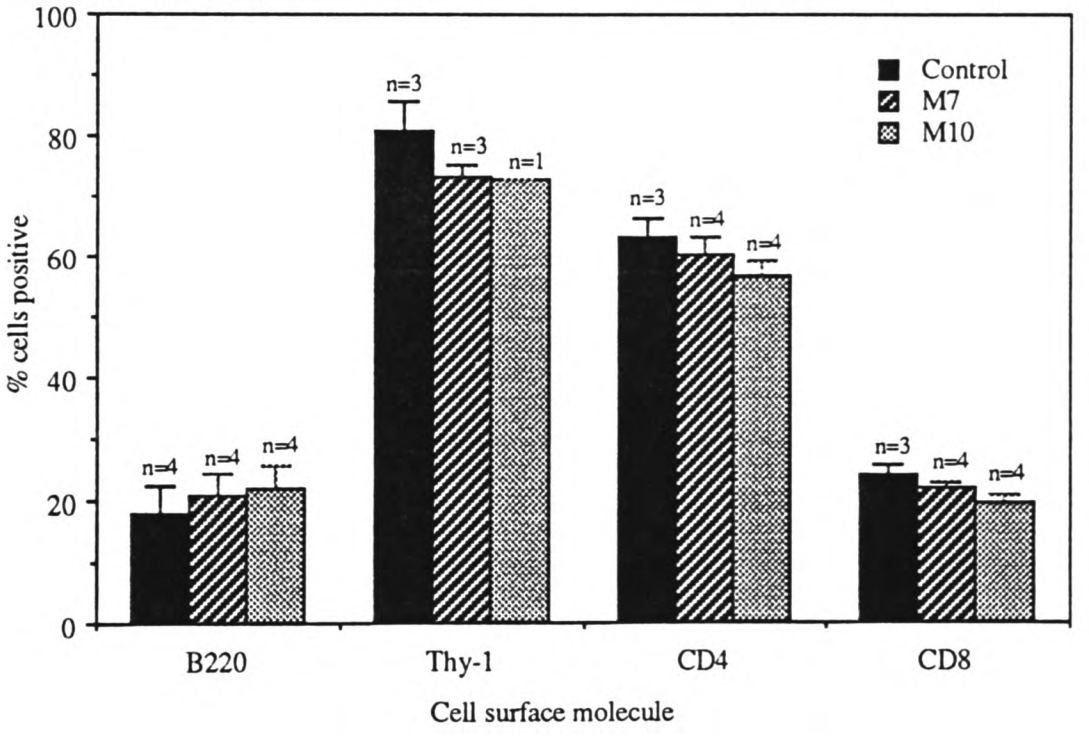
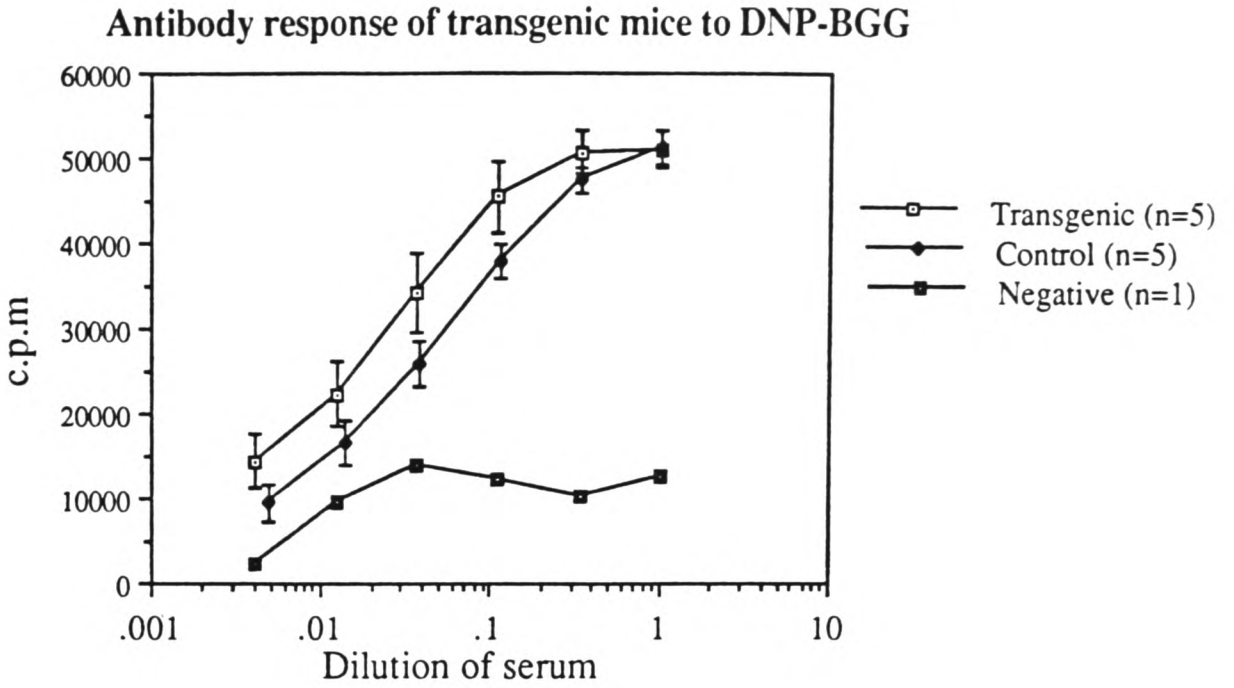


Figure 4.10 Increase in the ratio of B cells to T cells in the spleen (A) and lymph nodes (B) of mice from the M7 and M10 transgenic lines. Cells (1-2 x 10^6) from the spleen or lymph nodes (mesenteric and cervical) of individual mice were labelled with one of the PE conjugated mAbs (specific for): L3T4 (CD4), Ly2 (CD8), Ly5 (B220) or the FITC conjugated mAbs YBM (Thy-1) or OX-20 (κ chain) and analyzed by flow cytometry. A single marker was set to separate positive and negative cells with each mAb and the percent of positive cells labelled in each population was determined. Data have only been taken from labelling experiments where there was a clear separation of positive cells from negative cells. The data shown are the average and standard deviation of samples from the indicated number of mice of each type. All mice were between 7 and 35 weeks of age.

A.



B.

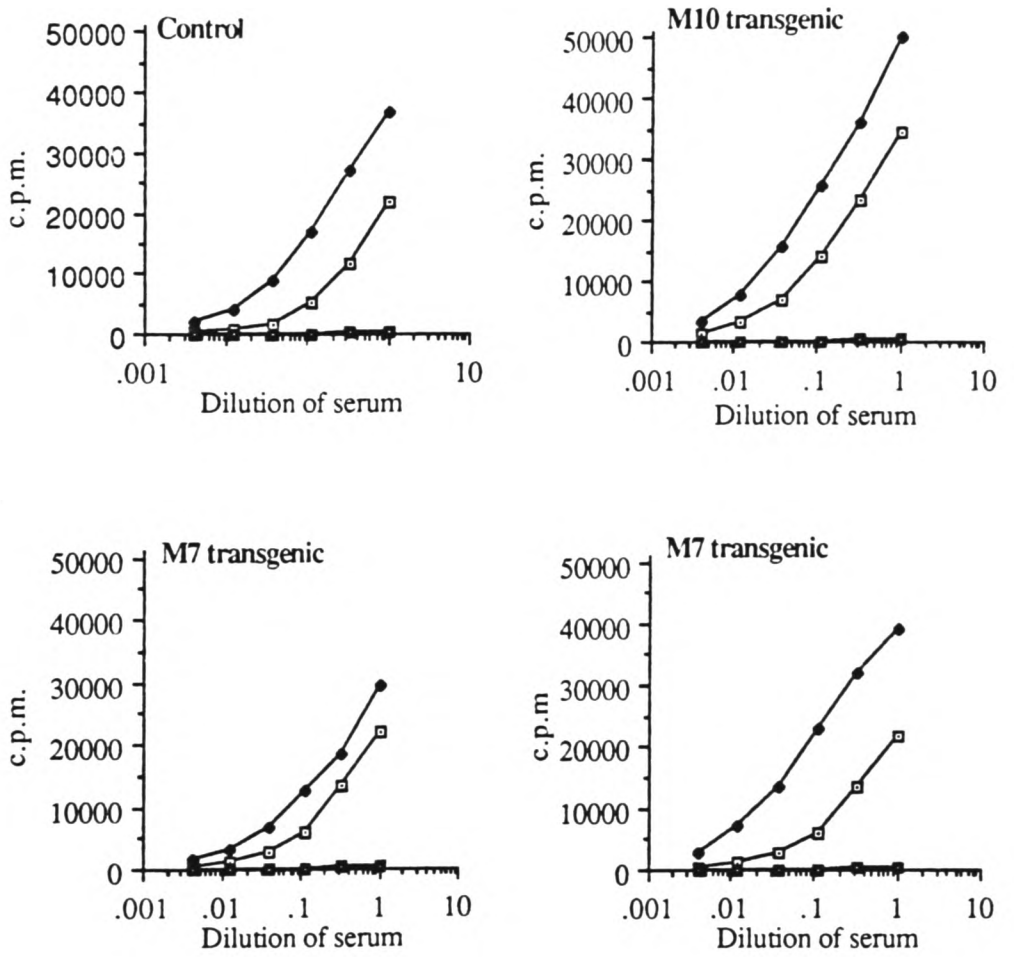


Figure 4.11 Antibody responses of M7 and M10 transgenic mice to bovine gamma globulin and sheep red blood cells. A) Five M7 and 5 control mice were injected i.p. with 0.5 ml of 20 µg/ml DNP-BGG as a potassium aluminate precipitate and 7 days later tail bleeds were performed. The sera were diluted 1/100 before use and then serial 1/3 dilutions were made and 50 µl samples were transferred in duplicate to wells pre-coated with DNP-BGG. ^{125}I-RAM was used in the second step and the counts shown are the average and standard deviation for each group of mice. B) Mice were injected with 0.3 ml of 1% SRBC in saline i.v. and 0.1 ml i.p. and tail bled at 7 days followed by a second injection at 10 days and a second bleed 4 days later. Sera were diluted as in (A) and mixed with 5 µl SRBC (10^9/ml) for 60 min at 4°C. The cells were washed 3x and ^{125}I-RAM added. After washing the bound counts were measued and the average of duplicate wells are shown. Open squares show the response after 7 days and filled diamonds the response 4 days after boosting. Also shown is an unimmunized control (filled squares).

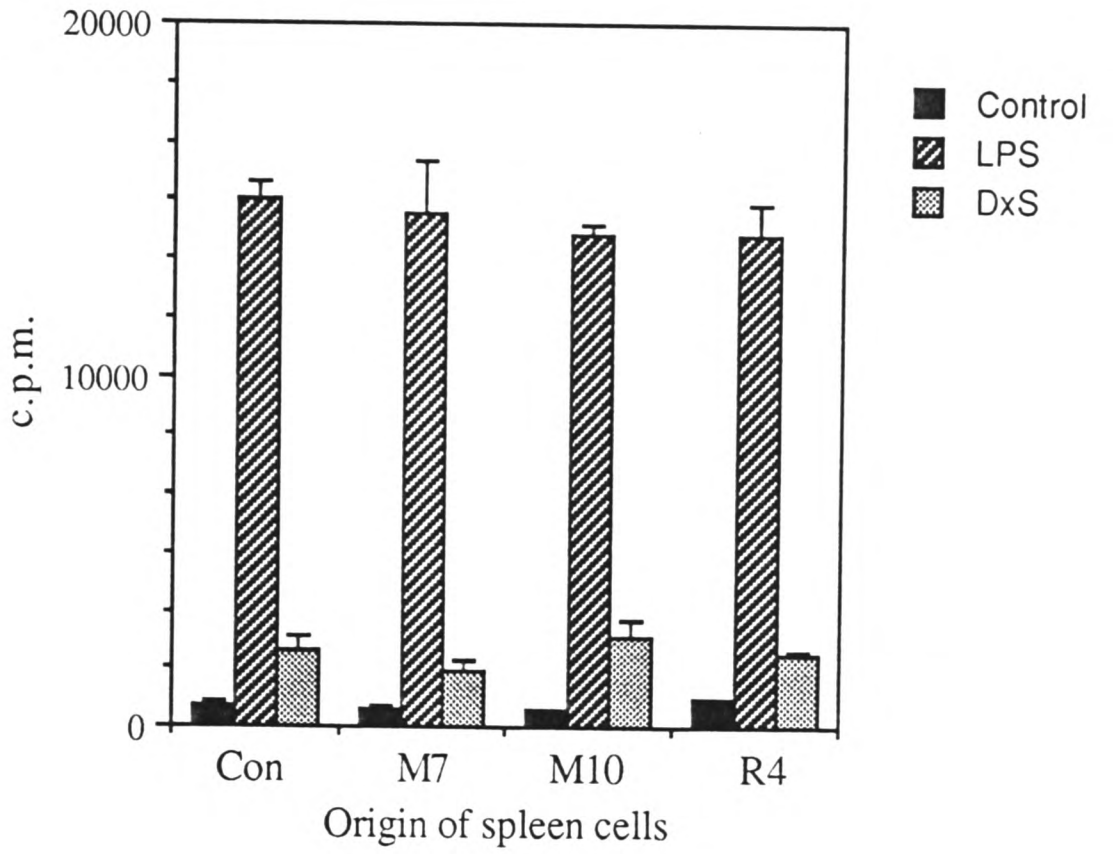


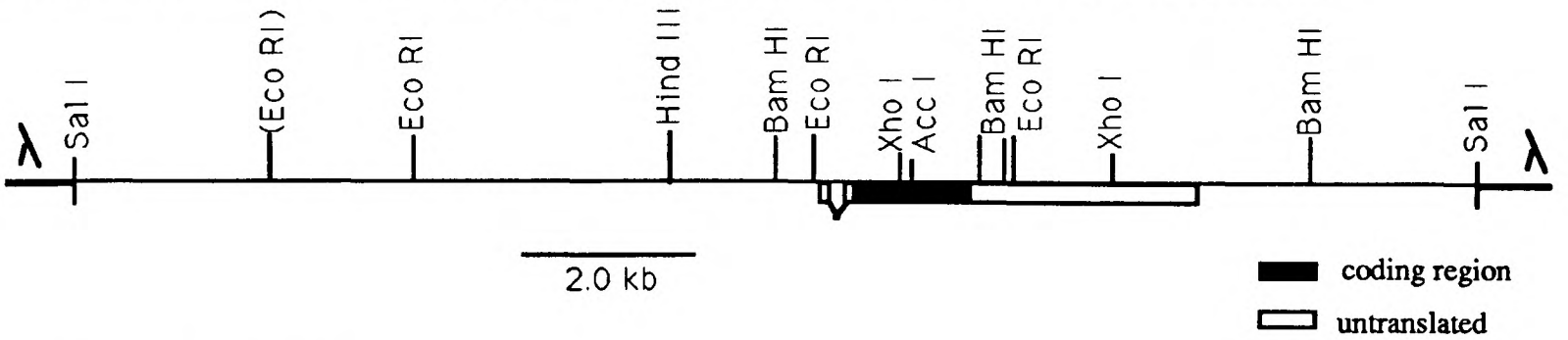
Figure 4.12 Proliferative response of spleen cells from M7 and M10 transgenic mice to LPS and dextran sulphate. Spleen cells were isolated from two spleens of each mouse type indicated and incubated in medium at $2x10^5$ cells/ml in round bottomed microtitre trays without addition (control) or with LPS at 10 μg/ml or dextran sulphate at 10 μg/ml. After 3 days 0.5 μCi of ^{3}H-thymidine was added and incubation continued for 14 hr. The cells were harvested on to filter mats using a Skatron Combi cell harvester and the filters counted in the presence of scintillant. Shown are the averages and standard deviations of triplicate cultures.

Spleen cells from transgenic and control mice showed very similar proliferative responses following *in vitro* stimulation with LPS and dextran sulphate (Fig 4.12) and no differences in ^{3}H-thymidine uptake in the absence of stimulatory agents was apparent.

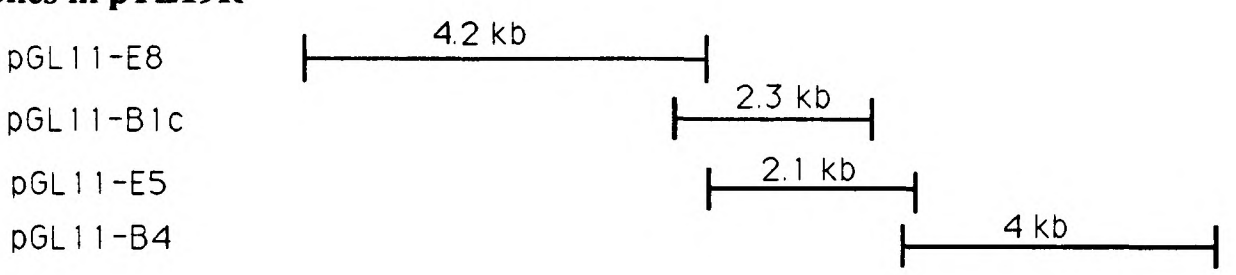
4.9 Construction of a targeting vector for interruption of the mouse leukosialin gene by homologous recombination in embryonic stem cells

The original EMBL-3 genomic mouse leukosialin clone isolated (GL11a, Chapter 3) contained an insert of approximately 16 kb. Bam HI and Eco RI fragments of the insert were subcloned into pTZ-19R and their orientation and location with respect to each other was determined by Southern blotting of restriction enzyme digests and comparing these to the original EMBL clone (Fig 4.13a). To generate a larger fragment containing the coding region the pGL11-E8 plasmid, which contains 4.2 kb of 5' untranslated sequence, was cleaved with Bam HI and the B1c fragment containing the open reading frame was inserted to generate a 6 kb product of identical sequence to the corresponding region of the endogenous gene. This plasmid was then treated with Xho I which cleaves at a unique site in the coding sequence, near the transmembrane region (Fig 4.13a), and the Neo resistance gene with or without a polyadenylation signal (poly A+ or -), and with the thymidine kinase (tk) promoter, was inserted. This serves both to interrupt the gene and to provide a selectable marker. Clones carrying the poly A+ form of the gene in both orientations were isolated. To allow negative selection of non-homologous recombinants, the herpes simplex virus thymidine kinase gene was inserted at the 3' end of the construct in the Sal I site which was present in the vector multiple cloning site immediately 3' of the Bam HI site. The final constructs contained a total of 6 kb of sequence homologous to the endogenous gene, 5.4 kb 5' of the Neo gene and 0.6 kb 3' (Fig 4.12b). The amount of sequence introduced on the 3' side was kept to a minimum to allow PCR for homologous recombinants using a sense oligonucleotide specific for the 3' end of the Neo gene and an anti-sense oligonucleotide specific for mouse leukosialin sequence immediately 3' of the sequence present in the construct. This construct is being used by Mike Carroll (Boston, USA) to interrupt the endogenous gene in embryonic stem cells.

A. EMBL-3 mouse leukosialin genomic clone (GL11a) restriction map and subclones



Sublcones in pTZ19R



B. Targeting vector for disruption of mouse leukosialin gene by homologous recombination

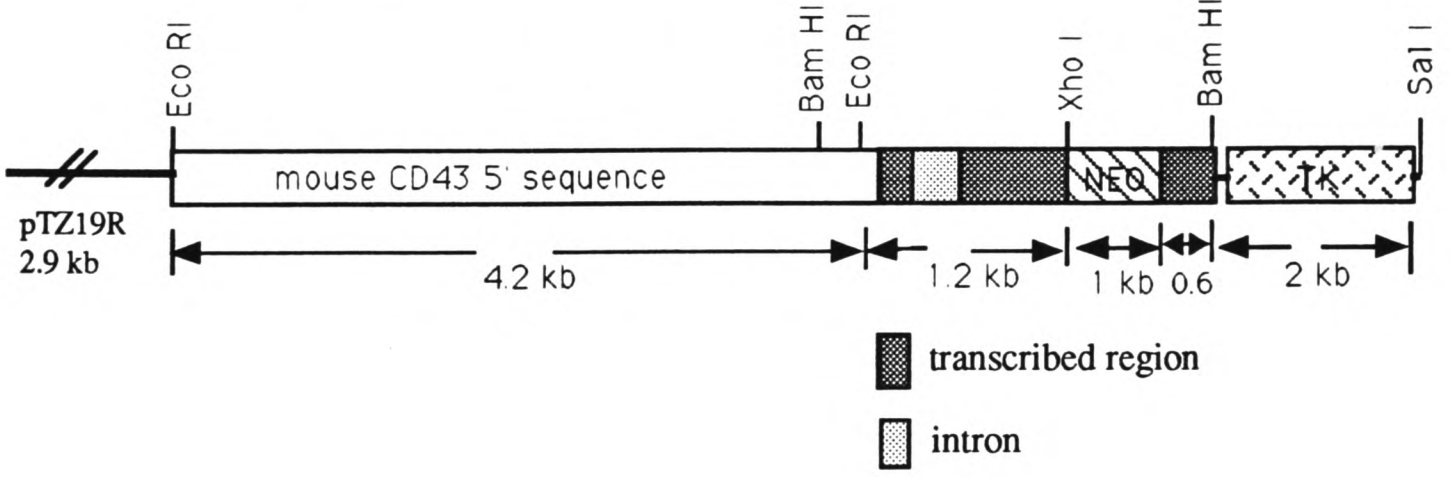


Figure 4.13 Restriction map and subcloned fragments of mouse genomic clone GL11a (A) and targeting vector for disruption of mouse leukosialin gene by homologous recombination (B). The restriction map was determined and the targeting vector constructed as described in the text.
The Eco RI in parentheses is only approximately located.
λ indicates DNA derived from the lambda arms of the EMBL-3 vector.
Neo is the neomycin resistance gene and TK the thymidine kinase gene.

4.10 Discussion

Mouse leukosialin was expressed on B lymphocytes of two lines of transgenic mice (M7 and M10). In one of these lines (M10) leukosialin was also expressed on cells in a variety of non-lymphoid tissues. Both lines grew normally and they showed no difference in breeding capacity or morbidity compared to non-transgenic CBA/ca mice kept under identical conditions. The conditions were clean but not SPF and thus the animals are likely to have been exposed to various bacteria and possibly viruses. Some problems were encountered in breeding initially but these were common to all the CBA/ca mice being bred. Inclusion of shredded tissues as bedding in addition to the standard floor of wood shavings was found to improve breeding significantly. The normal life span of these animals and their general good health makes it unlikely that they were suffering from any severe autoimmune diseases.

The immunological status of the transgenic mice appeared normal. The architecture of their cervical and mesenteric lymph nodes and their spleens was normal and the size of the B cell areas appeared the same as the controls and there was no detectable increase of B cells in the T cell areas. No B cells were detected in the thymus. Flow cytometry of labelled spleen and lymph node cells of M7 and M10 transgenic mice showed that they had a slight increase in the ratio of B cells to T cells in these tissues. The small size of this change makes it difficult to determine if it is due to an increased number of total B cells compared to controls, or a displacement of T cells. The antibody response of both lines of mice to a protein antigen and sheep red blood cells was similar to controls with a slight enhancement apparent in some cases. The level of expression of the major surface glycoprotein of B cells, B220 (or CD45R) which is likely to carry a large fraction of the B cell sialic acid, was not altered in the transgenic mice. This and the observation, in preliminary experiments where transgenic B cells were purified and lysates Western blotted with the S7 mAb, that the leukosialin on the B cells is of a similar molecular weight to that on T cells, argues that the transgenic B cells will carry more sialic acid at their periphery than control B cells. However, the leukosialin expression is still an order of magnitude lower on the transgenic B cells than on normal T cells so the possibility that leukosialin expression at the level of T cells has a greater effect on B cell behaviour cannot be ruled out.

A number of parameters have not been tested in the transgenic mice including: their ability to mount an antibody response to type I and type II T independent antigens; the composition of the Ly-1 B cell compartment; the production of different antibody isotypes; the efficiency of germinal centre formation; the actual rate at which the B cells pass through the lymphoid organs; the life span of the B cells. The relative increase in B cell numbers in spleen and lymph nodes may be a consequence of increased migration to these organs, reduced rates of migration through them, or increased B cell numbers in general. A preliminary analysis of the ratio of B cells to T cells in the blood did not reveal any significant difference but it will be valuable to perform a more comprehensive analysis of the blood lymphocyte populations as this should give an indication as to whether B cell numbers in the whole animal are elevated. The increased expression of leukosialin on T cells means the possibility that the alteration is in the T cell compartment cannot be ruled out at present. It will be necessary to perform adoptive transfer experiments, where labelled transgenic B cells are transferred to controls, to conclude that the effect is within the B cell compartment. Such a study will be important in testing the migration rate of leukosialin expressing B cells through lymph nodes and spleen.

The restriction in expression of the rat leukosialin transgene to the T cell compartment is not understood. It is possible that the presence of an intron in the 5' untranslated region will enhance expression of the trangene generally (Brinster et al., 1988; Palmiter et al., 1991) and that this accounts for expression of the mouse gene in both B and T cells of two transgenic lines. However, it will be necessary to analyse a larger number of lines to determine if the restricted expression was related to the site of insertion in the genome or to some inherent property of the construct. It is interesting that only a subpopulation of the T cells and thymocytes in the R3 and R4 lines show expression of rat leukosialin. Why this should be so and whether the expression pattern corresponds to a defined T cell subset is not known.

M10 mice showed much broader transgene expression than mice of the M7 line. This may be a consequence of the M10 line carrying multiple copies of the transgene (M9 mice also had multiple copies and, although expression on B cells was low, they too showed expression in non-lymphoid tissues). When multiple copies of the transgene are present they are almost invariably in a tandem array (which is why they are inherited in a Mendelian fashion) (Palmiter

and Brinster, 1986). Placing promoter elements in a tandem array can increase their potency (Verweij et al., 1990) and it may be the case that the combined activity of the Ig enhancer and H-2K promoter becomes less tissue specific when present as multiple copies in a tandem array.

The M7 transgenic mice have been bred to homozygosity. This will allow the line to be maintained without continual assessment of transgene transmission and should facilitate future studies of the type described above. Leukosialin from the transgenic B cells of these mice will also be useful for a study of the O-linked glycans attached to glycoproteins on resting B cells.

In conclusion, expression of leukosialin on B cells in transgenic mice at a level approximately 10 fold lower than that on T cells does not compromise antibody responses to T cell dependent antigens and does not lead to a major change in B cell localization in lymphoid organs.

CHAPTER 5

THE DIMENSIONS OF LEUKOSIALIN AND IDENTIFICATION OF LINEAR PROTEIN EPITOPES THAT CAN BE MODIFIED BY GLYCOSYLATION

5.1 Introduction

5.2 Electron microscopy of leukosialin

5.3 Antigenic heterogeneity as seen with leukosialin on transfected cells

5.4 Labelling of lymphoid cell populations and cell lines with the anti-rat leukosialin mAbs

5.5 Effect of neuraminidase digestion on the antigenic determinants

5.6 Binding of mAbs to the rat leukosialin extracellular domain expressed in *E.coli*

5.7 Epitope localization by fusion protein expression

5.8 Further characterization of the OX-56 and OX-74 epitopes

5.9 Localization of epitopes by electron microscopy

5.10 Discussion

5.1 Introduciton

Leukosialin consists of an extracellular portion of 224, 229, or 235 amino acids in rat (Killeen et al., 1987), mouse (Cyster et al., 1990) and man (Pallant et al., 1989; Shelley et al., 1989; Schmid et al., 1992) respectively, and attached to this there are 70-85 O-linked carbohydrate chains but few or no N-linked structures. On the basis of the high level of glycosylation of leukosialin it has been argued that the extracellular part of the molecule would have an extended structure (Brown et al.,1981).

There is considerable evidence for heterogeneity in leukosialin in terms of apparent M_r (Fukuda and Carlsson, 1986; Remold-O'Donnell et al., 1987), antigenicity (Standring et al., 1978; Carlsson and Fukuda, 1986; Bettaieb et al., 1988) and glycosylation (Carlsson et al., 1986; Fukuda et al., 1986; Piller et al., 1988). This heterogeneity must be ascribed to post-translational modifications for mouse and human leukosialin because in these species there is only one functional leukosialin gene which lacks introns in the coding sequence (Cyster et al., 1990; Shelley et al., 1990) precluding diversity introduced by alternative splicing of exons.

Antigenic heterogeneity has been seen in human leukosialin in terms of determinants that are specific either for T lymphocytes, or for a larger form on activated lymphocytes and granulocytes (Carlsson and Fukuda, 1986). For many epitopes, antigenicity is reduced if cells are treated with neuraminidase (Sportsman et al., 1985; Axelsson et al., 1988; Stoll et al., 1989). These results might be taken to suggest that leukosialin determinants involve both a protein and a carbohydrate component. The antigenicity of leukosialin is given added interest by the finding that anti-leukosialin antibodies are commonly found in patients infected with HIV-1 (Ardman et al., 1990).

In this chapter the extracellular part of rat leukosialin is shown to have an extended structure and the relative roles of protein and carbohydrate in the antigenicity of the molecule are studied. Epitope positions in the molecule are established by the complementary techniques of molecular biology and electron microscopy.

5.2 Electron microscopy of leukosialin

The molecular structure of the leukosialin molecule was clearly revealed by electron microscopy after low angle rotary shadowing (Fig 5.1). In the absence of detergent, rosette-like structures were seen, with a variable number of arms of equal length and a central density of size roughly proportional to the number of arms (Fig 5.1a). The addition of deoxycholate almost completely dissociated these rosettes into individual flexible rods, without discernable polarity or regional specialization (Fig 5.1b). It is concluded that the rosettes consist of leukosialin molecules that have formed micelles by association of their hydophobic transmembrane domains, while the structures seen in the presence of deoxycholate are single extended leukosialin molecules.

The length of the structures as seen in Fig 5.1 can be measured with reasonable accuracy and a correction of -1 nm for shadow thickness is applied at each free end of the rods (Shotton et al., 1979). In contrast, the widths were too narrow for meaningful measurement. Measurements of length were made for 70 arms of the rosettes illustrated in Fig 5.1a, using the criteria that the longest straight arms were chosen and measured from the edge of the central density. The mean length $\pm$ SD was 41 $\pm$ 3.3 nm after the -1 nm correction was applied. Measurements were also made of single molecules shown in Fig 5.1b with straight forms that were clearly isolated in the micrograph being chosen. The mean length $\pm$ SD after corrections of -2 nm was 49.3 $\pm$ 7.7 nm from 47 values.

The extra length of single molecules compared with the arms of the rosettes could be due to shadowing of the transmembrane and cytoplasmic segments or to part of the extracellular segment that has collapsed onto the centre of the protein micelle. It seems that the latter possibility will be part of the explanation since a globular cytoplasmic domain of 124 amino acids will be expected to be less that 4 nm in length and if the structure were extended it may not be visualized by this technique. In the case of the CD45 leukocyte molecule the extracellular part was found to overlay the cytoplasmic domain in shadowing studies (Woollett et al., 1985). Thus in subsequent discussion we take a value of 45 nm, which is the mean between the measurements from Fig 5.1a and 5.1b, as the length of the extracellular segment of leukosialin.

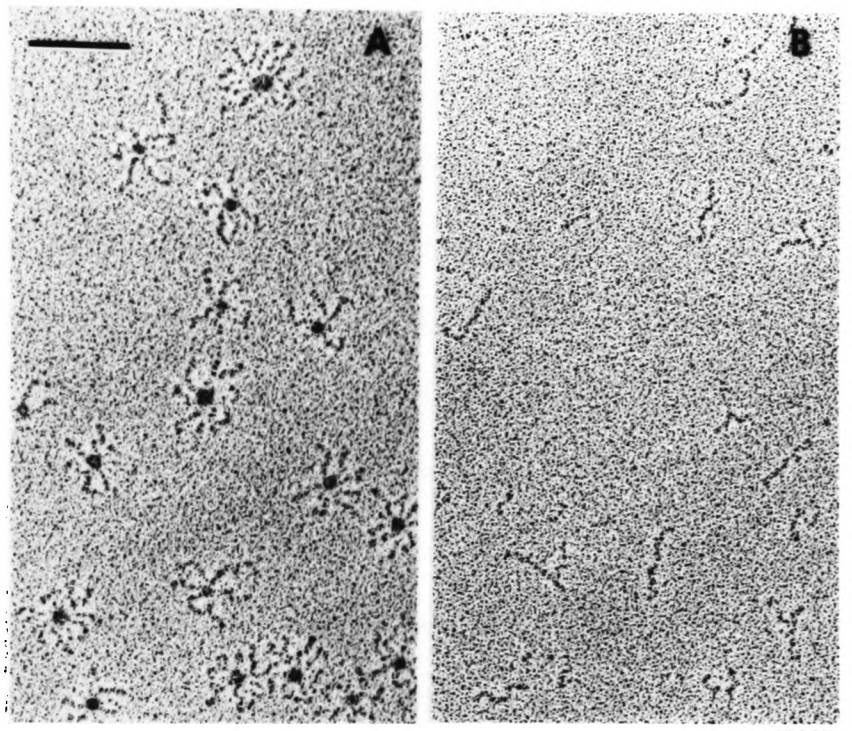


Figure 5.1. Electron micrographs of rat thymocyte leukosialin following low-angle rotary shadowing. (A) Leukosialin micelles in the absence of detergent. (B) Leukosialin monomers sprayed from a solution containing 0.05% sodium deoxycholate. Magnification x 100,000; Bar = 100 nm.

5.3 Antigenic heterogeneity as seen with leukosialin on transfected cells

Full length cDNA for rat leukosialin in either the pHSE3' or CDM8 eukaryotic expression vector, was transfected into mouse NSO myeloma cells (as described in chapter 4), human Jurkat T lymphoma cells, Madin-Darby canine kidney (MDCK) epithelial cells and K562 erythroleukemia cells. Leukosialin on the transfected cells was studied by labelling with nine different monoclonal antibodies, of which eight were of the IgG class. All the mAbs gave strong binding to leukosialin on NSO cells (Fig 5.2) proving that they recognized the product of one leukosialin gene. In contrast, with the OX-58 and OX-75 mAbs, leukosialin was not labelled at all on Jurkat, MDCK and K562 cells, and some of the other mAbs gave weaker binding on these cells in comparison with that of the OX-56 and OX-74 mAbs. On K562 cells W3/13 and 5H4 gave the strongest staining. It is presumed that these antigenic differences are due to the effects of differential glycosylation.

5.4 Labelling of lymphoid cell populations and cell lines with the anti-rat leukosialin mAbs

All of the mAbs reacted with most thymocytes and T lymphocytes in the lymph node cell population (Fig 5.3). In contrast the labelling of B lymphocytes in the lymph node population was not distinguishable from that of the negative control, except with the OX-56 and 5H4 mAbs which gave weak binding. Thus it seems that OX-56 and 5H4 detect leukosialin at low levels on B cells. The alternative that they both cross-react with another molecule is unlikely, since these mAbs recognize different epitopes (see below) and it is improbable that different cross-reactions would yield a similar pattern of binding.

With bone marrow cells clear differences in labelling were seen. The W3/13, OX-57, OX-75 and 5G7 mAbs labelled various subsets of the total cell population while the OX-56 and 5H4 mAbs labelled all the cells. W3/13 labels polymorphs among bone marrow cells but does not bind to the lymphoid cells which are mostly of pre-B lineage (Williams et al., 1977). The OX-56 and 5H4 mAbs must be labelling B cell progenitors as well as polymorphs and this correlates with the finding that these mAbs weakly label mature B cells. Expression of leukosialin on mouse pro- and pre-pro- B cells (Hardy et al., 1991) has

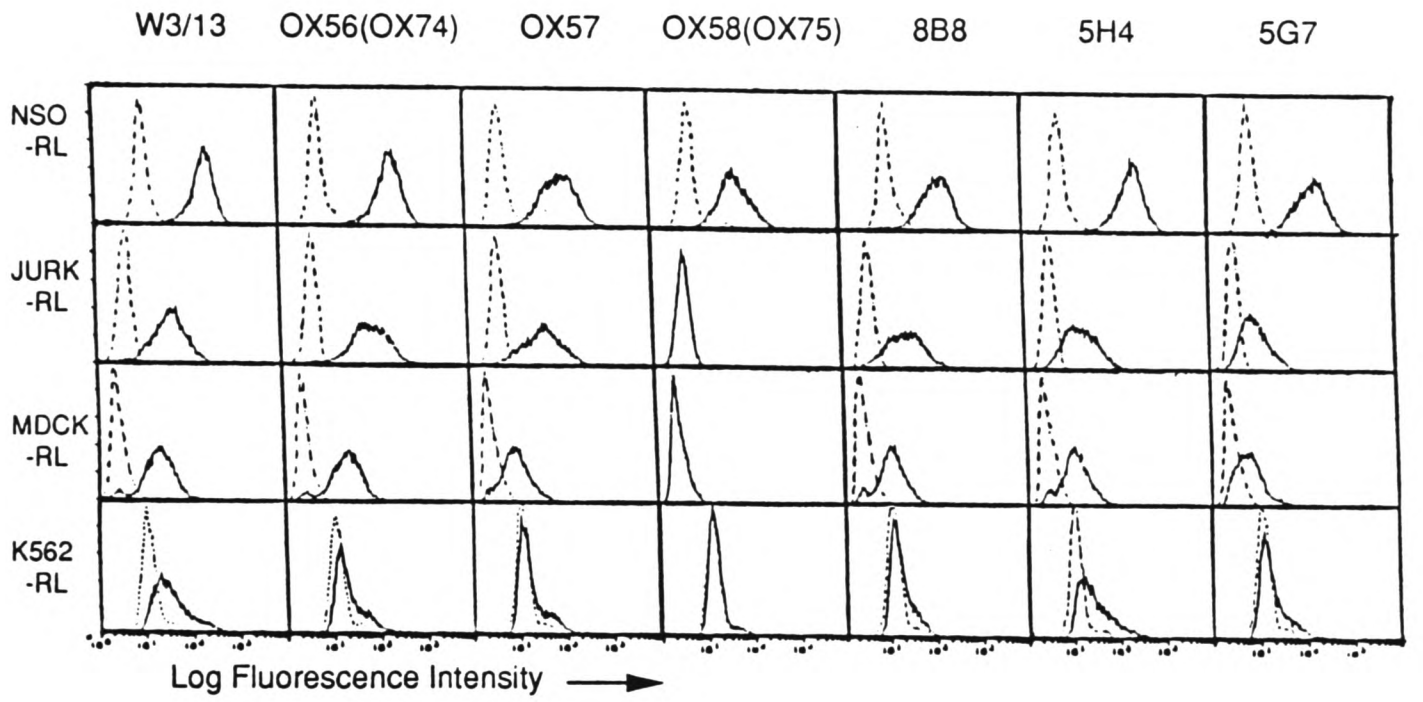


Figure 5.2. Binding of mAbs to rat leukosialin expressed on NSO, Jurkat, MDCK or K562 cells transfected with rat leukosialin cDNA. Transfected cells (10^6) were incubated with saturating levels of the mAbs followed by a second incubation with fluorescein-labelled $F(ab')_2$ anti-mouse IgG antibody, and bound antibody was measured by flow cytometry. The solid lines show labelling of the transfected cells and dotted lines non-transfected control cells. The horizontal axis shows log fluorescence intensity and the vertical axis relative cell number. Each horizontal row is for one cell type and each vertical row for one mAb. The profiles for OX-74 and OX-75 were indistinguishable from those for OX-56 and OX-58, respectively, and thus the data for OX-74 and OX-75 are not shown separately.

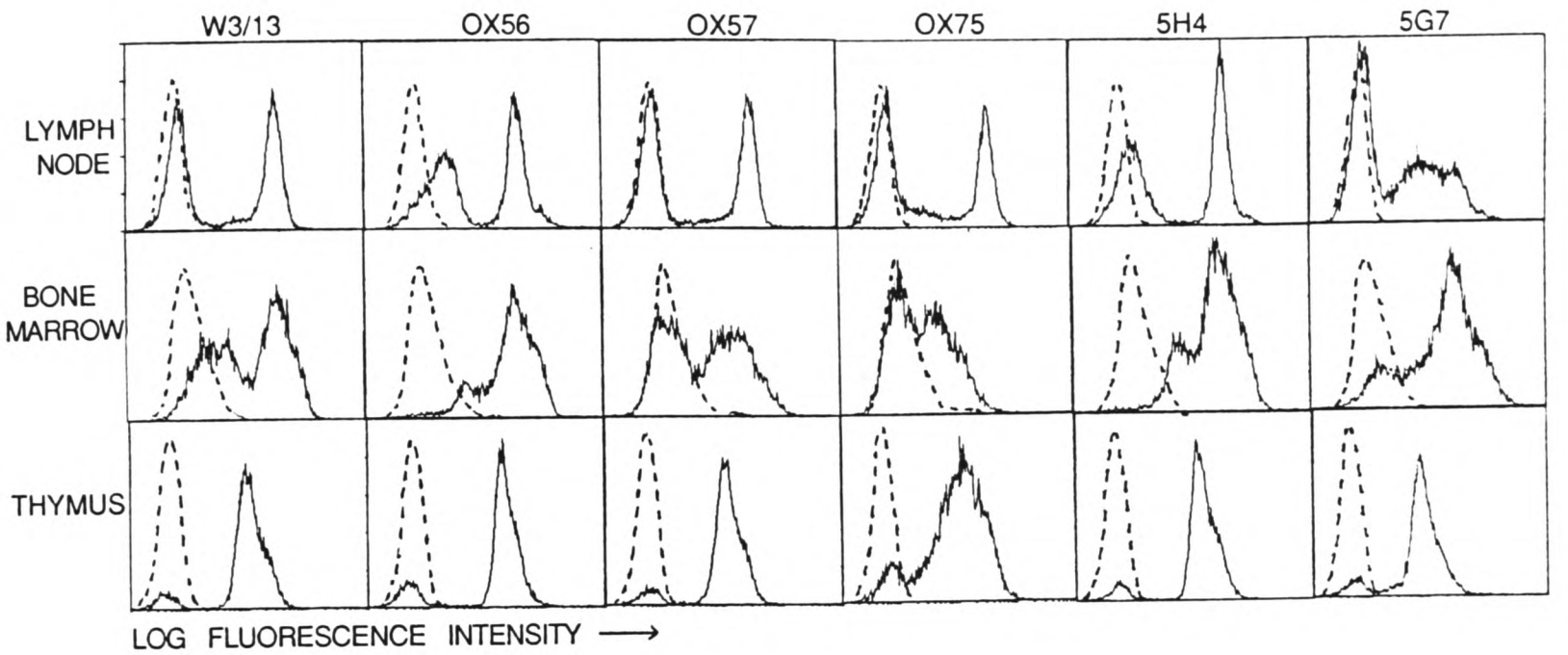
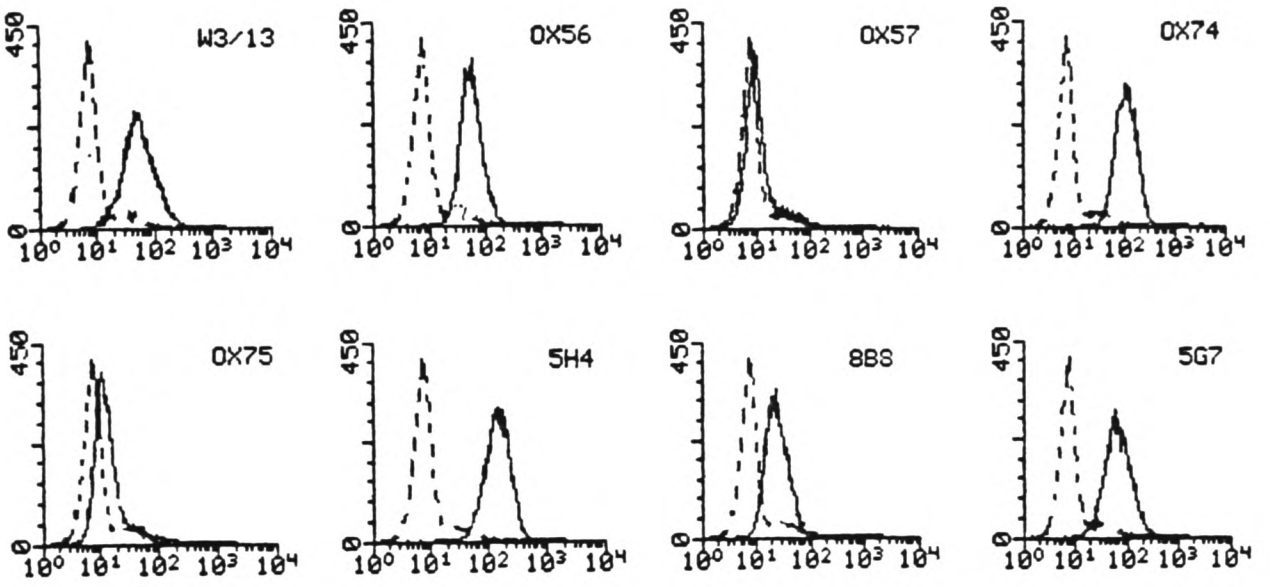


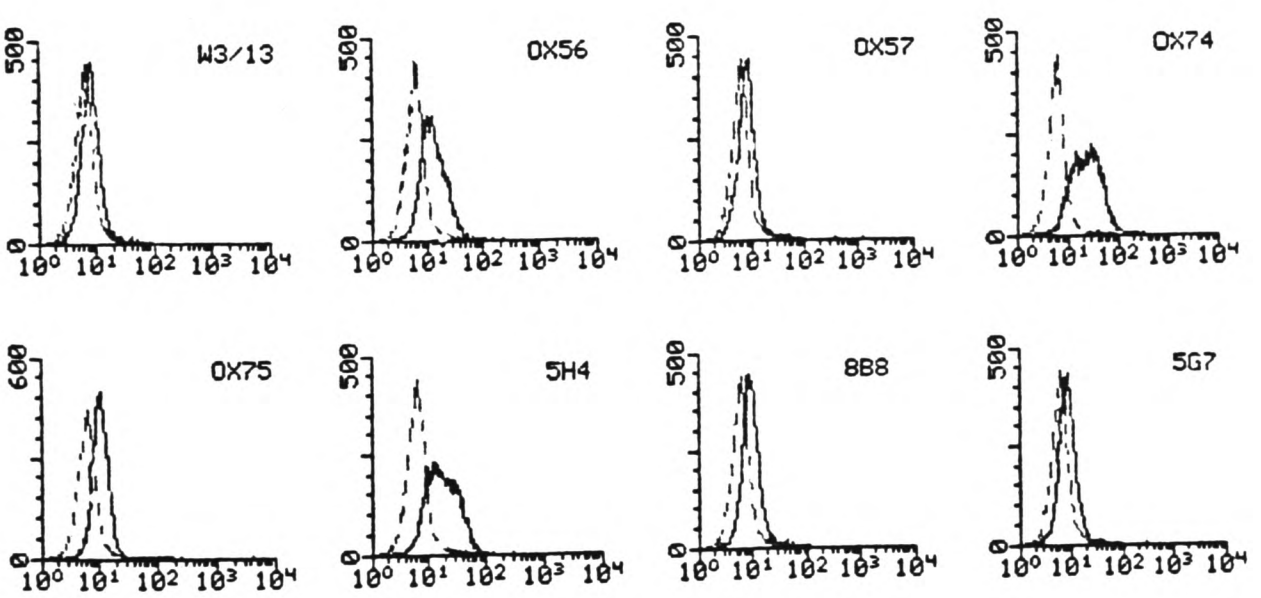
Figure 5.3. Binding of anti-rat leukosialin mAbs to rat lymphoid cells analysed by flow cytometry. Cells were isolated from the lymph nodes, bone marrow and thymus of a PVG rat and analysed with anti-rat leukosialin antibodies as described for Fig 5.2. Profiles for 6 MAbs are shown. In each case the solid line represents the indicated mAb and the dotted line an isotype matched control. Other mAbs matched one of the patterns shown as follows: OX-58 = OX-75; OX-74 = OX-56; 8B8 = 5H4 and thus the data are not shown separately. Other experiments on PVG and AO rats gave similar profiles.

(A)

A181



RNK-16



Log Fluorescence Intensity

Figure 5.4. Binding of anti-rat leukosialin mAbs to rat NK cell lines A181 and RNK-16 (A) and to the rat T cell line, Line 3, following antigenic stimulation (B). A) The NK cell lines were kindly provided by Louise Spruyt and were stained as in Fig 5.2. B) The Line 3 cells were provided by Albert Beyers 4 days after stimulation with antigen (MBP). Live cells were isolated on a ficoll gradient and the cells were stained directly (—) or placed in fresh medium and left for 3 days (···) or 10 days (----) before staining with the indicated mAb. (····) indicates staining of cells isolated on the gradient with control mAb OX21. The horizontal axis shows log fluorescence intensity and the veritcal axis relative cell number.

(B)

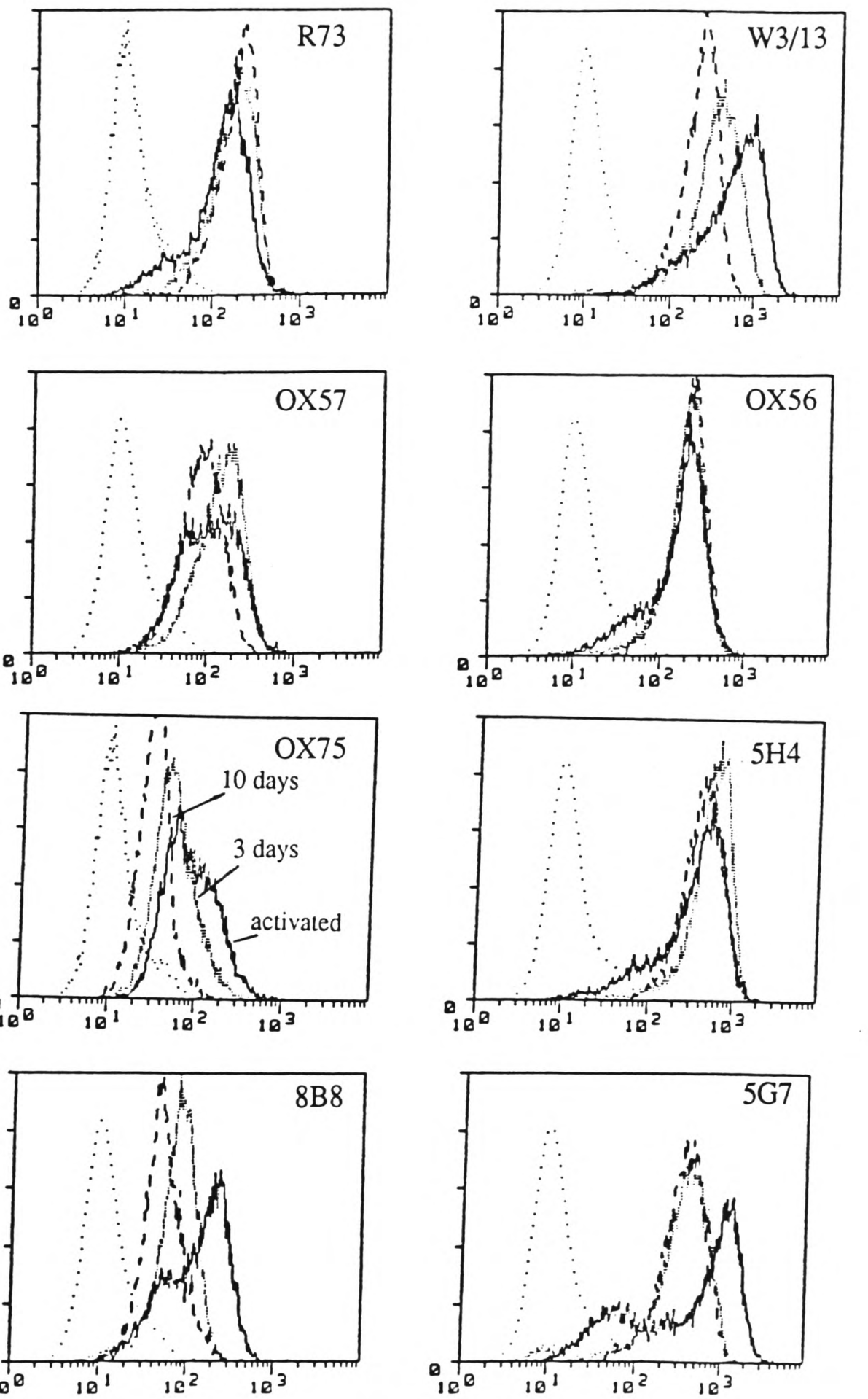


Figure 5.4 continued

recently been demonstrated. Thus in the rat bone marrow population there are clear differences in the recognition of determinants of the leukosialin molecule by different mAbs.

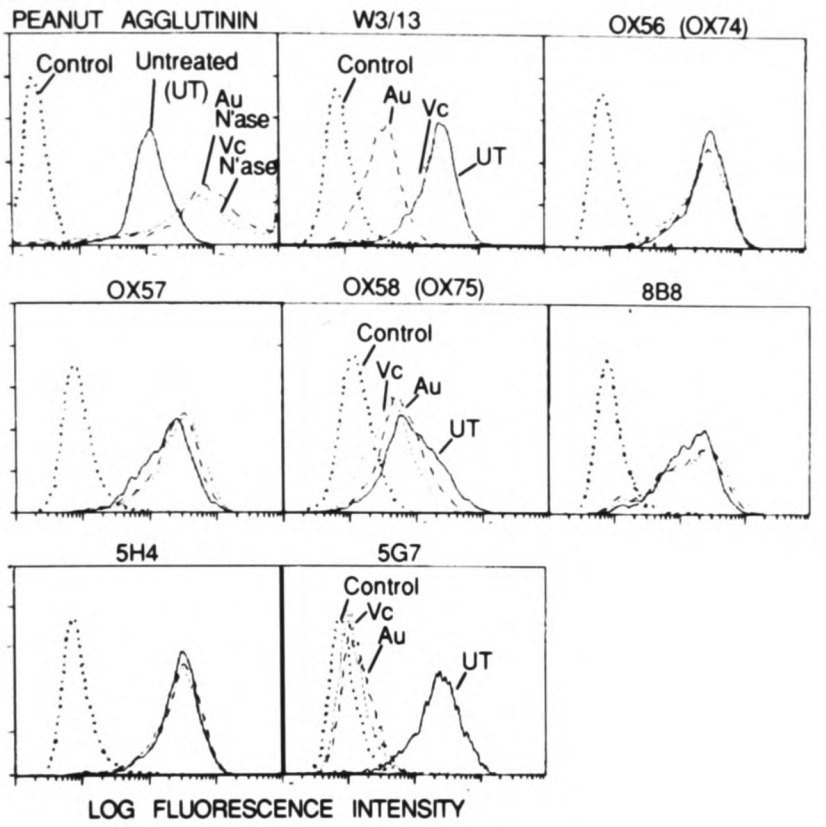
The anti-leukosialin mAb staining profiles of two rat NK cell lines were examined and both showed different patterns to the transfected cells and to general lymphoid populations (Fig 5.4a). The A181 line was unusual in being OX57 negative. The RNK-16 line showed stronger staining with the OX-75 mAb than with W3/13, a reciprocal pattern to that seen with the transfected cell lines. To examine the possibility that the antigenic structure of leukosialin on T cells changes with activation, the MBP specific CD4+ T cell line, Line 3 (Sedgwick et al., 1989), was labelled at an early stage after antigenic stimulation and at subsequent time points as the cells returned to the resting state (Fig 5.4b). At the first time point (4 days) after antigen stimulation, when TCR expression was reduced, W3/13 labelling was high and the cells showed a marked fall in expression of this epitope after a further 10 days of culture. OX-75 and 8B8 showed the opposite pattern with mAb binding increasing as the cells returned to the resting state. OX-56 labelling was unchanged over the whole period whereas OX-57 and 5H4 showed more complex patterns (Fig 5.4b).

5.5 Effect of neuraminidase digestion on the antigenic determinants

Treatment with neuraminidase of NSO cells expressing rat leukosialin was effective in revealing ligands for peanut agglutinin and in producing a significant reduction of binding of 4 of the anti-leukosialin mAbs (Fig 5.5a). Furthermore, the results obtained differed depending on the type of neuraminidase used. *Arthrobacter ureafaciens* neuraminidase is selective for the removal of α 2-6 linked sialic acid, and this enzyme had more effect on the binding of W3/13 mAb than digestion with *Vibrio cholera* neuraminidase which has a higher activity on α 2-3 linkages (Uchida et al., 1979). In contrast the opposite result was seen for the binding of OX-58 and OX-75 mAbs, whilst digestion with either enzyme resulted in loss of 5G7 binding. Figure 5.5a also shows that neuraminidase digestion gave a slight enhancement of OX-57 mAb binding.

The effects of neuraminidase treatment were also analysed by Western blotting after SDS PAGE (Fig 5.5b), giving results that qualitatively supported the cell binding data. In

(A)



(B)

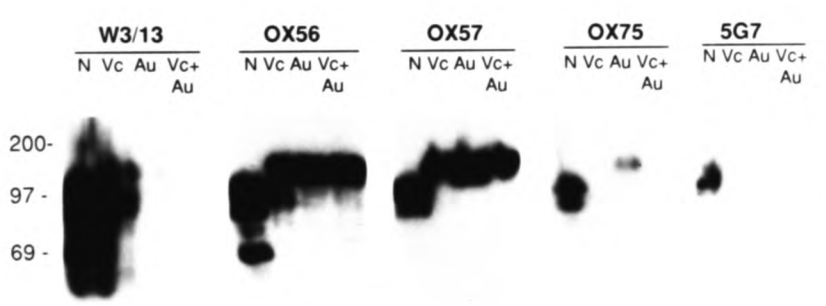


Figure 5.5. Effects of neuraminidase treatment on the antigenicity of rat leukosialin assayed by flow cytometry and Western blotting. A) Transfected NSO cells expressing rat leukosialin were treated with either *V. cholera* (····) or *A. ureafaciens* (- - -) neuraminidase for 60 min or left untreated (—), washed in PBS/BSA and analysed directly by flow cytometry as in Fig 5.2. The first profile shows binding of FITC-conjugated peanut agglutinin before and after the neuraminidase treatments as compared to the background where galactose was present in excess (50mM, control (· · ·)). The remaining profiles show the binding of anti-rat leukosialin mAb before and after the neuraminidinase treatment as compared to a negative control mAb OX-21 (· · ·). As in Fig 5.2, the profiles for OX-74 and OX-75 were indistinguishable from those for OX-56 and OX-58 respectively, and are not shown separately. B) Transfected NSO cells were left untreated (Native, N), treated with *V. cholera* neuraminidase (Vc) or *A. ureafaciens* neuraminidase (Au) or a combination of the two (Vc+Au). 10^6 cells were then boiled in 100 µl of reducing sample buffer and 10 µl was run on 7.5% SDS-PAGE and electroblotted to nitrocellulose. The nitrocellulose membranes were probed with the indicated antibodies and ^{125}I-RAM as the second antibody prior to autoradiography. Numbering on the left indicates the apparent Mr determined from marker proteins (kD).

addition the results in Fig 5.5b established that both the *A. ureafaciens* and the *V. cholera* neuraminidase had removed substantial amounts of sialic acid from the leukosialin molecule, since the characteristic shift of leukosialin to a higher apparent M_r on SDS PAGE (Standring et al., 1978; Brown et al., 1981) was seen in both cases.

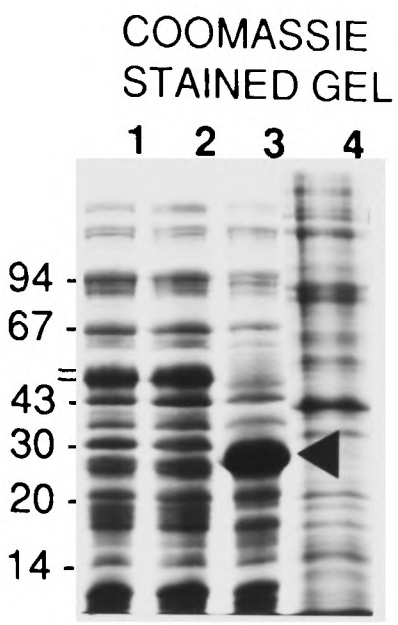
5.6 Binding of mAbs to the rat leukosialin extracellular domain expressed in *E. coli*

The extracellular domain of rat leukosialin was isolated by PCR using the oligonucleotides in Table 5.1 and expressed in *E. coli* as a fusion protein with glutathione-S-transferase (GST) in the pGEX vector system (Smith and Johnson, 1988). The binding of mAbs to the fusion protein was assessed by Western blotting. Of the 9 mAbs available, 7 gave positive binding reactions. Figure 5.6 shows the binding obtained with W3/13, OX-56 and OX-58 MAbs and the lack of reaction with the OX-57 antibody. W3/13 mAb bound to a form of the fusion protein having the apparent M_r expected for the full length protein. OX-56 and OX-58 bound both to this form and to lower M_r forms believed to be degradation products. These differences in binding of the mAbs to degradation products are consistent with the epitope mapping described below.

The antigenicity of the affinity-purified fusion protein was also tested in a plate binding assay with the results shown in Table 5.2. Binding to the fusion protein in this assay was clearly significant for 6 of the 7 mAbs that showed positive binding after Western blotting. The exception was the W3/13 mAb, for which binding to the fusion protein was low, although above that of the GST control. It was also of interest that the OX-57 and 5G7 mAbs gave measureable binding above the control even though these mAbs were not active in Western blotting. It is perhaps not surprising that the results from the plate assay and from Western blotting were not totally concordant, given that the conformation of the protein might differ in these two assay systems.

5.7 Epitope localization by fusion protein expression

To localize epitopes, the plasmid containing the GST-rat leukosialin construct was cut with Eco RI to generate linear DNA ending on the 3' side of the cDNA encoding the extracellular segment. This was further digested with Bal 31 exonuclease for various periods to yield



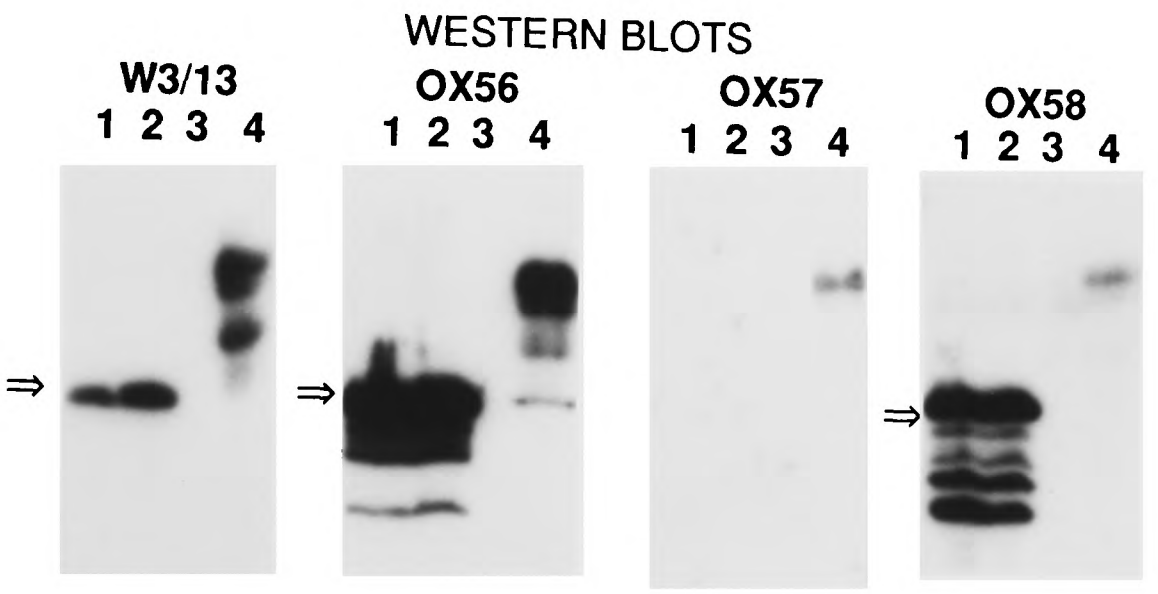


Figure 5.6. Reactivity of mAbs with recombinant (non-glycosylated) extracellular domain of rat leukosialin. Coomassie stained gel: Samples of *E.coli* clones expressing the full length extracellular domain as a fusion protein with glutathione-S-transferase (lanes 1 and 2, fusion protein indicated by ⇒) or the glutathione-S-transferase alone (lane 3, GST indicated by arrowhead) were boiled in sample buffer and run on a 12% SDS-PAGE gel. A sample of NSO cells expressing rat leukosialin in glycosylated form was included in lane 4. Western blots: Identical gels were Western blotted and probed with the indicated mAb and ^{125}I-RAM second antibody prior to autoradiography. Numbering on the left indicates apparent M_r (kD).

Table 5.1 Oligonucleotides used in the PCR to isolate the rat leukosialin extracellular domain and fragments of the domain for insertion into the pGEX-2T vector

No.	5'/3'	Amino acid		Sequence 5'-3'
1079	5'	+1	rat	tagtagggatccgaaaatctgccgaatacg
1080	3'	+224	rat	ctgaattcagccactcgatcctgg
1155	3'	+28	rat	ctagaattcattcgaggtacttggg
1156	3'	+39	rat	ctagaattcagtcctctgtcactggcac
1157	3'	+55	rat	ctagaattcaggaagctggggcagtg
1076	5'	+1	mouse	tagtagggatccctgcagaggacgacgatg
1077	3'	+229	mouse	ctagaattcagccacttgattcttggct
1161	3'	+114	mouse	ctagaattcatgtcacaggattggctgcagtgacagga
1162	3'	+114	m, rat codons bold	ctagaattcatgt**cgc**aggatt**ggttat**agtgacagga
1260	5'	+108	rat, octapeptide	gatcccctcctgtcactataactaatccttg
1261	3'	+115	rat, octapeptide	aattcaaggattagttatagtgacaggaggg

All 5' oligonucleotides include a Bam H1 site and all 3' oligonucleotides a stop codon followed by an Eco RI site. Amino acid indicates the position in the rat or mouse template sequence at which the coding region of the oligonucleotide begins (5') or ends (3'). In 1162 the codons changed from mouse to rat sequence are shown in bold. 1260 and 1261 were not used in PCR but were directly annealed and ligated in to the vector as detailed in Chapter 2.

Table 5.2 Summary of the data for anti-rat leukosialin mAbs including results for the plate binding assay with fusion protein

MAb	Anti-body Class	Cell Type Variation[a]	Neuraminidase Sensitivity	Western Blot Reactivity		Plate Binding Assay		
				Native Leukosialin	Fusion Protein	Qualitative Score[b]	Quantitative Data (cpm) GST-fusion	GST
OX21	IgG1						154	242
W3/13	IgG1	+	+	+	+	?	498	87
OX56	IgG2b	+	-	+	+	+	8 324	160
OX57	IgG2b	+	-	+	-	?	929	368
OX58	IgG2b	++	+	+	+	+	9 421	469
OX74	IgG1	+	-	+	+	+	12 045	1188
OX75	IgG2b	++	+	+	+	+	8 369	803
8B8	IgM	+	-	+	+	+	9 738	241
5H4	IgG1	+	-	+	+	+	8 210	302
5G7	IgG1	++	+	+	-	?	680	222

a) the pattern of staining with a given mAb on at least one cell type was less than the maximum staining seen with one or more of the other anti-leukosialin mAbs on that cell type. b) ? indicates that the result from the binding assay was neither clearly positive nor clearly negative.

fragments with different lengths of 3' DNA eliminated. The digested material was then cleaved with Bam HI to release fragments of the leukosialin cDNA which were recloned into the pGEX vector and expressed in *E. coli* for Western blotting with mAbs. A succession of clones were detected that reacted with different mAbs as the size of the construct increased. For finer mapping, the size of inserts was determined by PCR on bacteria picked from colonies, and further clones were selected with sizes intermediate between those that had already been studied by Western blotting.

Results for the Western blotting of the polypeptides expressed from 13 constructs are shown in Figure 5.7. It should be noted that the stability of the fusion proteins varied considerably and in a number of cases a fusion protein band was not visible amongst the other *E. coli* proteins after Coomassie blue staining. However, for all constructs with leukosialin fragments of 65 residues or greater, expression could be confirmed by Western blotting with OX-58 and OX-75 mAbs (Figs 5.7 and 5.8). To ensure the production of fragments shorter than 65 residues, constructs were made by PCR techniques containing the 23, 39 and 55 NH_2-terminal amino acids of leukosialin (Table 5.1). Fusion proteins expressed from these constructs were clearly visible with Coomassie blue staining after SDS PAGE. Thus the negative antibody binding reactions recorded for these fragments (Figs 5.7 and 5.8) were not due to a lack of expression. The epitopes were mapped for all the mAbs that react on Western blotting (Fig 5.8) and the localization was clear-cut in all cases except that of W3/13, whose epitope was difficult to localise exactly because variable reactivity was seen with clones smaller than residues 1-129. Weak reactivity, not evident in Fig 5.7, was seen with clones 1-122, 1-120 and 1-118, but no binding at all was seen to clone 1-115 despite repeated blots with this construct. Thus the W3/13 epitope is assigned to the region 116-129. All the epitopes were found to be localized to one or other of two general regions of the extracellular domain, namely between residues 55 to 79 and 107 to 130 (Fig 5.8). The data in Figs 5.7 and 5.8 firmly establishes a physical limit to the extent of epitopes on the COOH-terminal side, but leaves open the possibility that some residues other than those assigned might be necessary on the NH_2-terminal side.

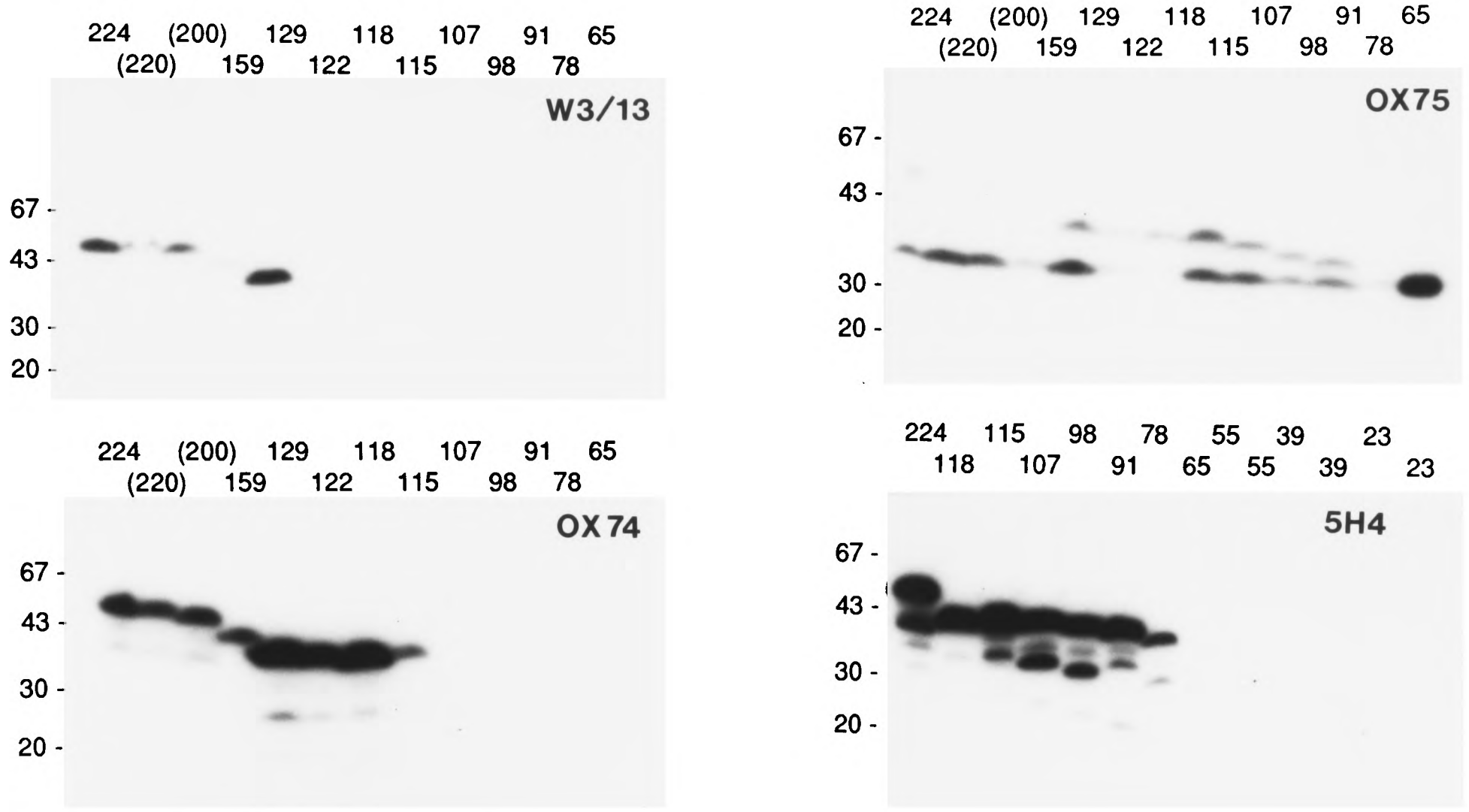


Figure 5.7. Epitope mapping of rat leukosialin using deletion clones expressed in *E.coli*. Plasmids containing coding sequences for the extracellular part of leukosialin with deletions at the 3' end were expressed as fusion proteins in *E.coli*. The cells were then boiled in sample buffer and run on 15% SDS-PAGE for Western blotting. The membranes were probed with the indicated mAb and ^{125}I-RAM as second antibody prior to autoradiography. Numbering above the lanes refers to the leukosialin sequence number at the COOH-terminus of the fusion protein Western blotted in the underlying lane. Values in parentheses are approximate end points, estimated from insert size only, whereas all other end points were determined by sequence analysis. Numbers on the left of the blots indicate apparent Mr (kD). Reactive degradation products of the fusion proteins are evident in many of the lanes in which positive clones had been run. These were not seen in cases where the fusion protein was non-reactive.

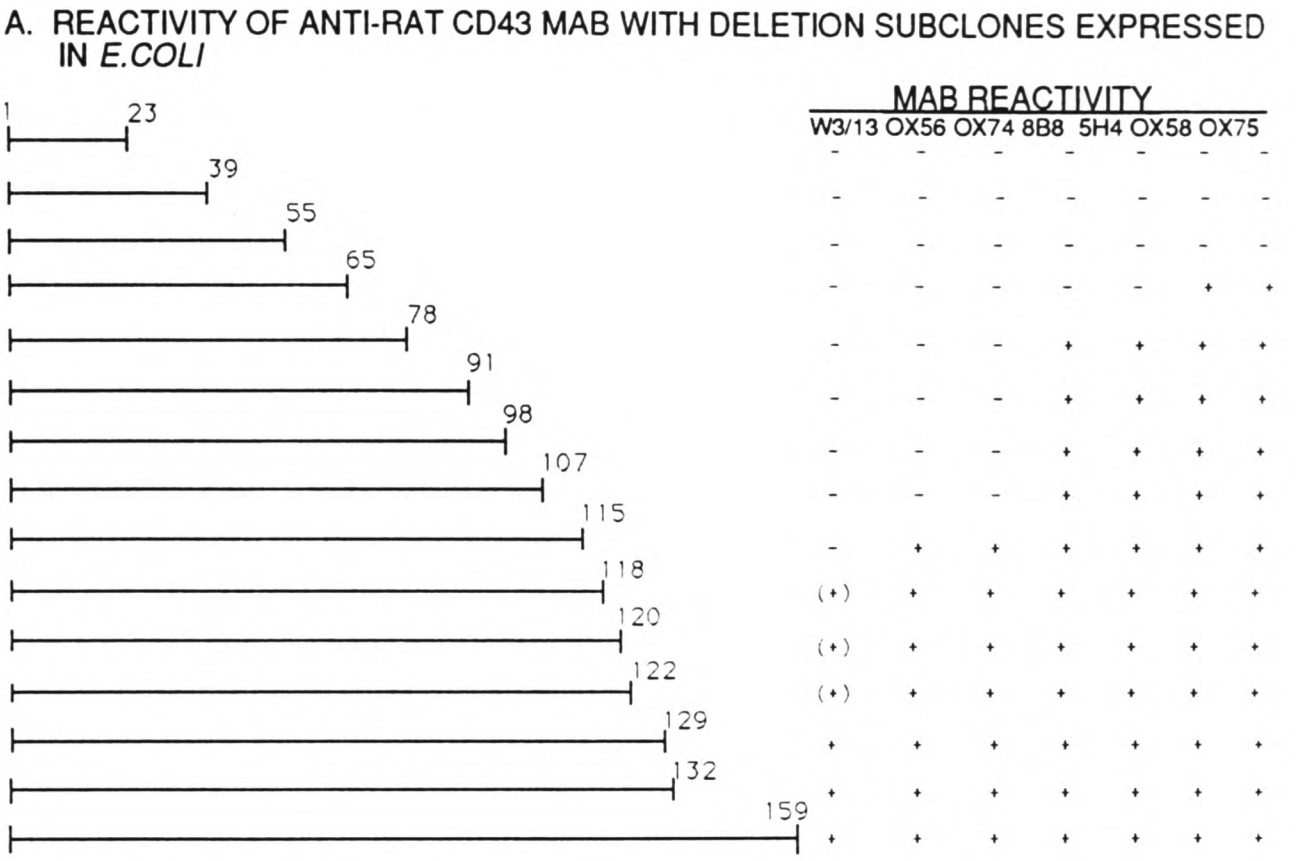


Figure 5.8. Summary of the reactivity of anti-leukosialin mAbs with deletion clones of rat leukosialin expressed in *E.coli*. (A) The clones were analysed for reactivity by Western blotting with all the antibodies showing clear differences between positive and negative clones with the exception of mAb W3/13, which showed varied reactivity with clones 1-118, 1-120, and 1-122 on repeated blotting. The horizontal lines represent the clone inserts where residue 1 is the NH_2-terminal amino acid of mature leukosialin and the number on the right is the COOH-terminal residue of each deletion clone. A further 4-7 amino acids derived from the vector may be present at the COOH-terminus (see Chapter 2). (B) The location of the epitopes in the extracellular domain is shown to scale. The shaded areas correspond to the region between the longest negative and shortest positive clone reactive with the indicated antibody, except with the W3/13 epitope where the shortest "positive" subclone was taken as 1-129. The epitopes for OX-56 and OX-74 were found to be totally overlapping by this analysis, as were the OX-58 and OX-75 epitopes and the 8B8 and 5H4 epitopes, confirming the similar flow cytometry results obtained for these pairs of antibodies (Figs 5.2-5.4).

5.8 Further characterization of the OX-56 and OX-74 epitopes

To confirm the assignment of epitopes in the leukosialin extracellular region in one case, the epitopes recognized by the OX-56 and OX-74 mAbs were investigated. These epitopes had been mapped by Western blotting of the deletion proteins to the sequence P (108) PVTITNP (Fig 5.8). An alignment of the rat and mouse leukosialin sequences around this region is shown below, where the dots correspond to sites of O-glycosylation in rat leukosialin.

```
                  •       •       •       •
Rat      98 F [P] KS [S] G [VASDPPVT] IT [NP] A [T] 117
Mouse    95 L [P] EP [S] N [VASDPPVT] AA [NP] V [T] 114
```

From this alignment and the observation that OX-56 and OX-74 do not react with mouse leukosialin, the residues Ile Thr were candidates to determine the epitope. In transferring the rat epitope to mouse leukosialin, two extra COOH-terminal residues of rat sequence were included (see chapter 2). These extra residues were thought unlikely to be contact amino acids but might influence the local conformation of the peptide. Thus changes were made in mouse residues 109, 110 and 113 and a fusion protein created with a stop codon after Thr 114 (mouse numbering). In a separate experiment, the octapeptide PPVTITNP alone was expressed as a fusion protein with GST. Results from Western blots of these constructs are shown in Fig 5.9 where it can be seen that OX-74 and OX-56 reacted with the mutated mouse molecule but not the wild-type form. Furthermore the octapeptide fusion protein gave a clear reaction albeit at a somewhat lower level than might have been expected from the control rat sequence on the same blot. The quantitative aspects of these experiments are difficult to assess because of possible variations including, degradation, transfer to the blots, or decrease in antibody affinity due to an incompletely constituted epitope. The clear conclusion, however, is that the OX-56 and OX-74 antibodies can recognise the octapeptide.

To assess the binding further a synthetic peptide with the sequence VASDPPVTITNPAT, corresponding to residues 104-117, was obtained. The peptide was active in blocking OX-56 and OX-74 binding to the recombinant extracellular domain, but not OX-75 which is specific for a different epitope (Fig 5.10). The molar potency of the peptide was equivalent to that of the fusion protein indicating that the whole epitope is contained within the peptide. The high concentrations of peptide and fusion protein needed for 50% inhibition of mAb binding may

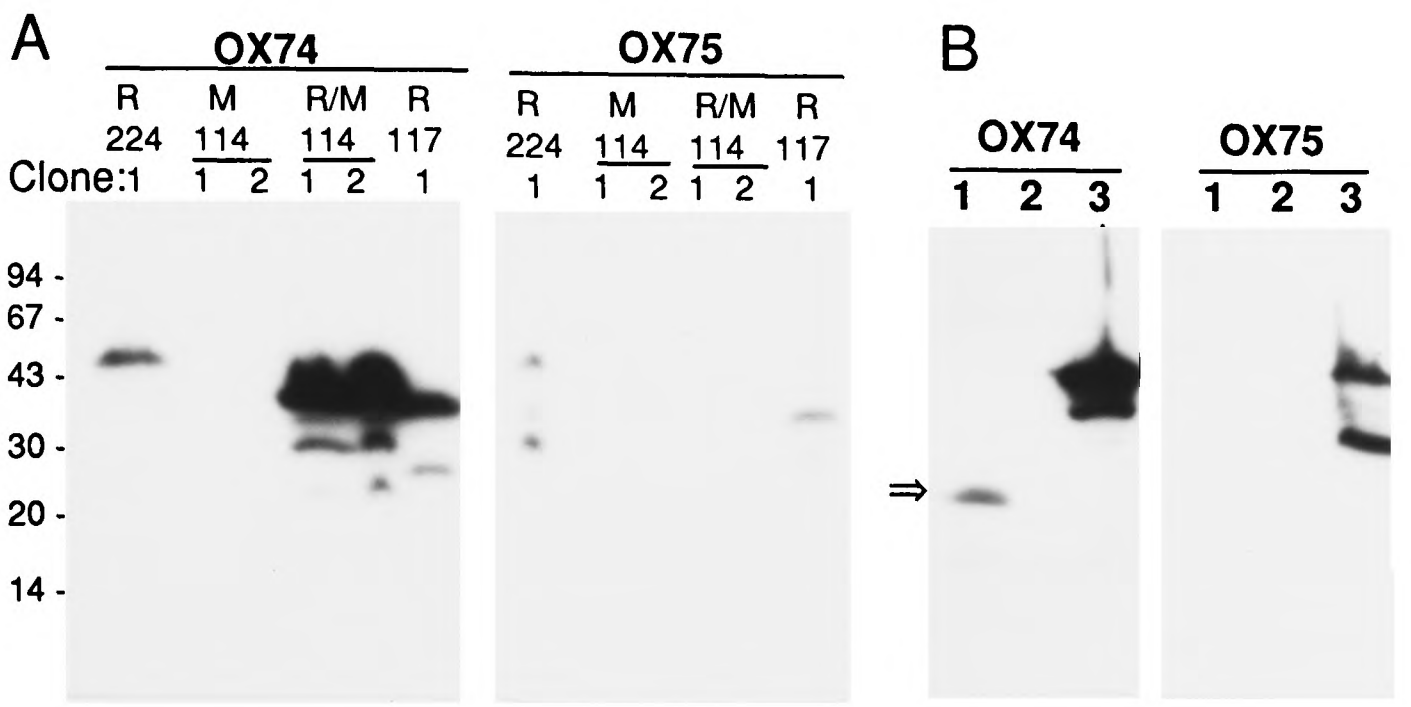
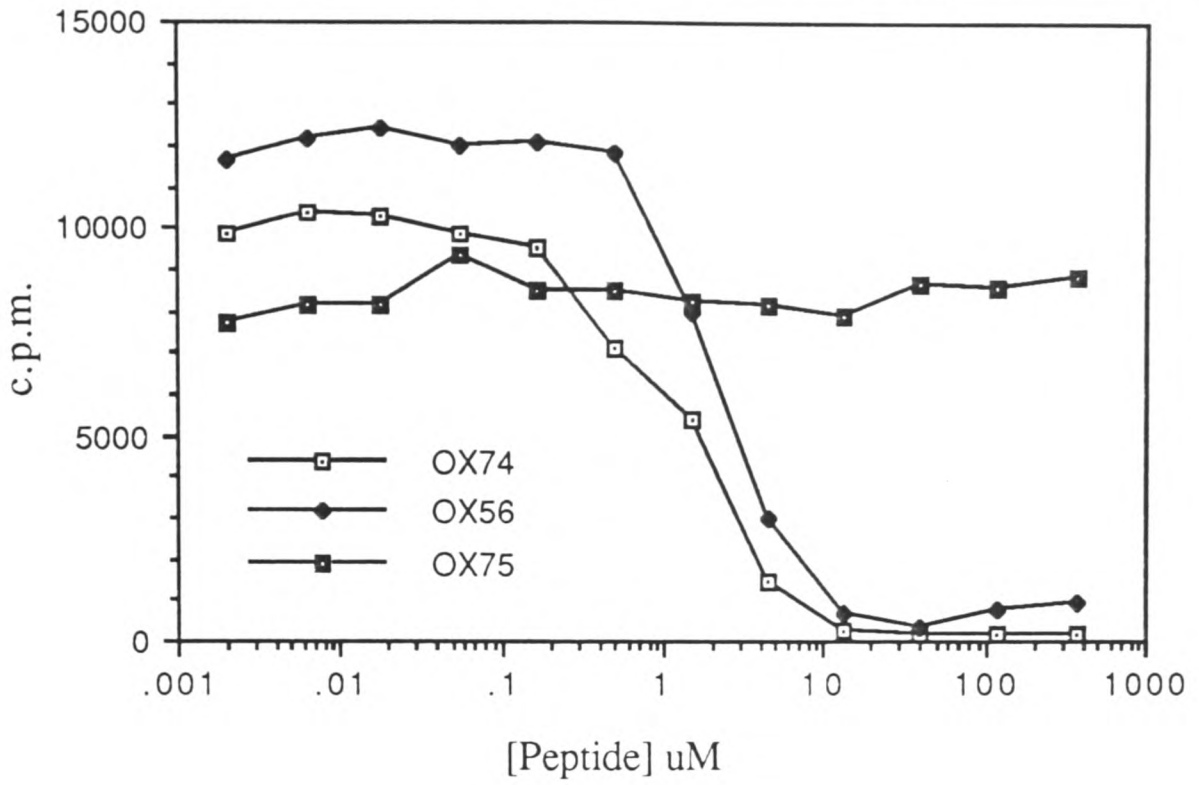


Figure 5.9. Reactivity of OX-74 mAb with modified mouse leukosialin and with the octapeptide PPVTITNP-fusion protein. *E.coli* clones expressing fragments of normal or mutated leukosialin, or the octapeptide, as fusion proteins with GST were boiled in sample buffer, run on 15% SDS-PAGE and Western blotted with OX-74 or with OX-75 as a control. In (A) the tracks are: R 224 clone 1, rat leukosialin residues 1-224; M 114 clones 1 and 2, two separate clones containing mouse leukosialin residues 1-114; R/M 114 clones 1 and 2, two clones containing mouse leukosialin residues 1-114 with amino acids 109, 110 and 113 mutated to rat sequence; R 117 clone 1, rat leukosialin residues 1-117. In (B) the tracks are: 1. octapeptide (rat leukosialin residues 108-115) indicated ⇒, 2. GST alone, and 3. rat leukosialin residues 1-224. Numbering on the left indicates the apparent M_r determined from marker proteins. The OX-56 mAb showed similar reactivity with the modified mouse construct and detectable but weaker reactivity with the octapeptide-fusion protein.



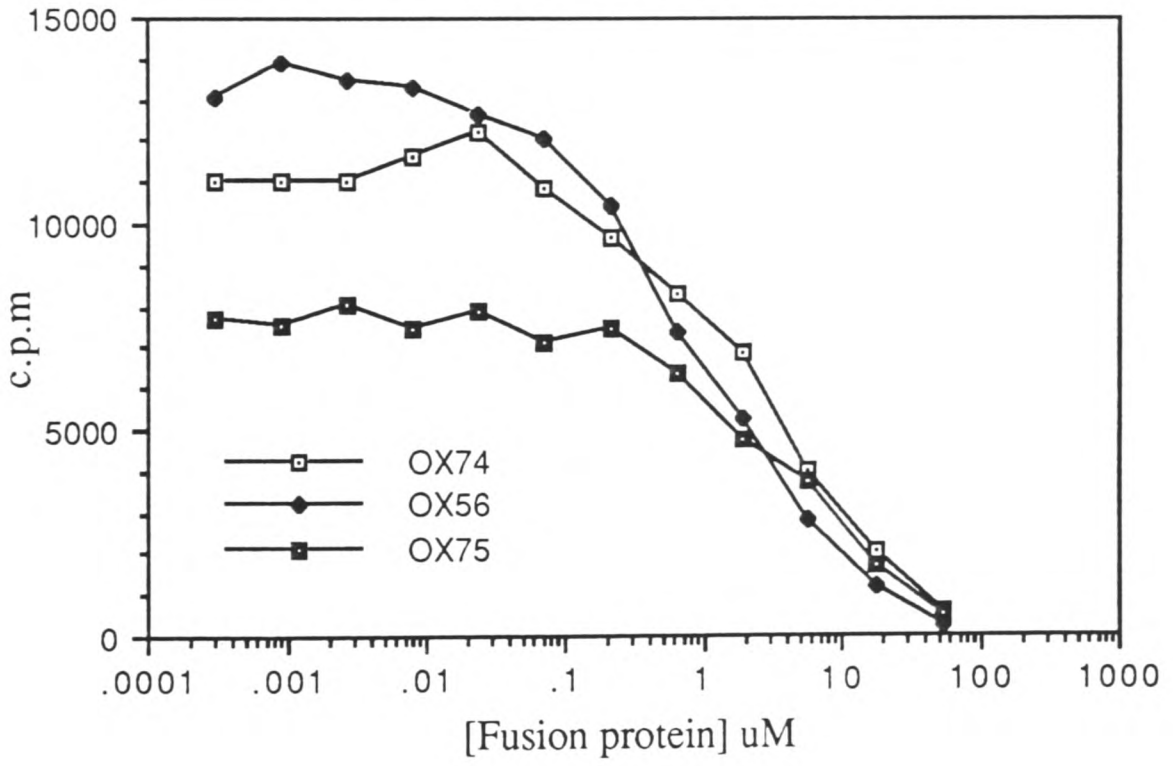


Figure 5.10. Inhibition of OX-56 and OX-74 mAb binding by the synthetic peptide VASDPPVTITNPAT. Three fold serial dilutions of peptide or extracellular domain-GST fusion protein were incubated with OX-56, OX-74 or OX-75 mAbs for 1 hour at 4°C. 50µl of each mixture was transferred to duplicate wells of a plate coated with the extracellular domain-GST fusion protein and incubated for 1hr at 4°C. ^{125}I-labelled RAM binding was then measured. The counts shown are the average of duplicate assays and are from representative experiments. The starting fusion protein concentration was estimated from the absorbance at 280nm, taking A_{280}=1 to be 1mg/ml. The peptide Mr is 1.38kD and the fusion protein Mr approximately 50kD.

indicate that only a fraction of the molecules in either preparation are in the correct conformation. It will be necessary to test glycosylated forms of the extracellular domain to define the affinity of the mAbs.

The complete epitope is thus contained within residues 104-117, the residues present in the synthetic peptide, and possibly within residues 108-115 since the original deletion clone finished at rat residue 115 and the corresponding octapeptide is reactive with both mAb.

5.9 Localization of epitopes determined by electron microscopy

To visualize epitope positions in glycosylated leukosialin, complexes were made between pure leukosialin and Fab fragments of the W3/13, OX-57 and OX-75 mAbs. According to Figure 5.8, OX-75 and W3/13 would be expected to bind near the distal end and middle, respectively, of the arms of the rosette form of leukosialin. This result was clearly seen in the complexes visualized by electron microscopy (Fig 5.11 A, B) confirming the predicted orientation of the leukosialin molecule in the rosettes. The epitope for the OX-57 mAb could not be mapped with the fusion proteins, because no reaction was seen with the *E. coli* material in Western blotting. However, it seemed possible that this mAb reacted near the NH_2-terminus because it was very potent in causing aggregation of rat thymocytes (see Chapter 7). This possibility was confirmed by low angle shadowing which showed binding of OX-57 at or near the ends of the rosette arms (Fig 5.11C).

To quantitate the differences in position of the W3/13 Fab compared with the OX-75 or OX-57 Fabs distances were measured between the outer face of the central density of the rosette and the inner side of the Fab. More than 100 measurements were made in each case and the mean values $\pm$ SE of mean were: W3/13, 11.6 $\pm$ 0.3 nm; OX-57, 21.6 $\pm$ 0.4 nm; and OX-75, 21.3 $\pm$ 0.5 nm. These values clearly support the finding that W3/13 Fab binds closer to the transmembrane sequence than the other Fabs. It is unclear why the the leukosialin extracellular domain is not seen extending passed the bound Fab fragments, even with the W3/13 Fabs which bound approximately 30 nm from the tip. This effect has also been seen with binding of Fab fragments within 30 nm of the end of the *Acanthamoeba* myosin II tail (Rimm et al., 1990).

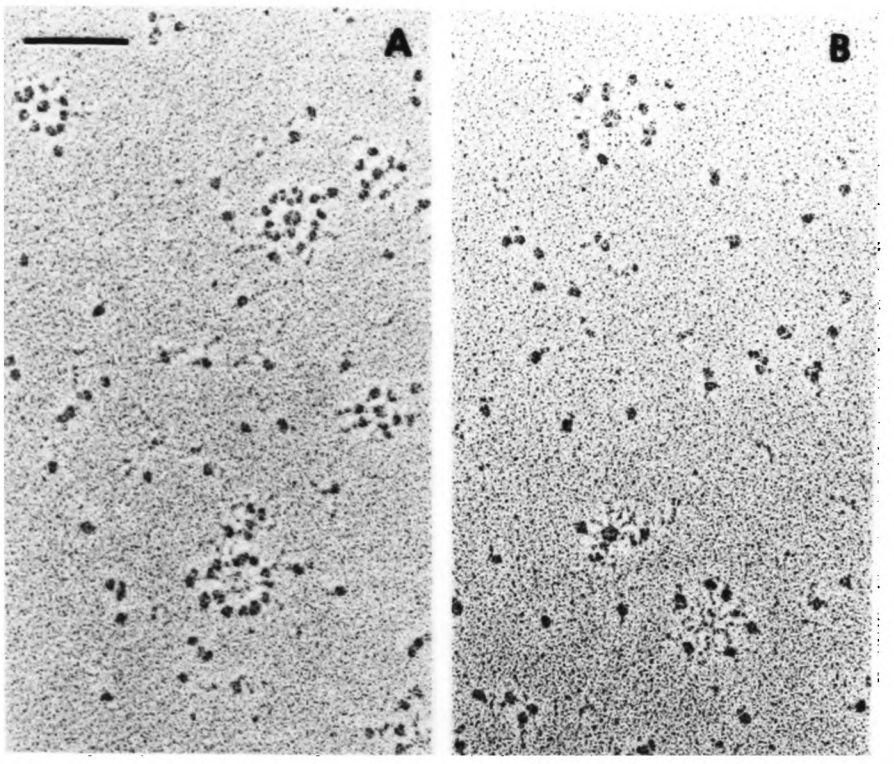


Figure 5.11. Electron microscopy of complexes of Fabs and rat leukosialin following low-angle rotary shadowing. The complexes were for leukosialin plus Fabs from the anti-leukosialin MAbs: (A) W3/13 (B) OX-75 (C) OX-57. (D) shows leukosialin plus unbound Fab from OX-35, a control mAb that is specific for CD4. All complexes were prepared in the absence of detergent. Magnification x 100,000; Bar = 100 nm.

5.10 Discussion

The visualisation of leukosialin by electron microscopy shows an extended conformation for the glycosylated 224 amino acid extracellular segment of about 45 nm in length. This yields an extension per amino acid residue of about 0.2 nm which is less than the theoretical value of 0.34 nm per residue for a fully extended peptide backbone and greater than the 0.15 nm per residue seen in an alpha helix. A value of 0.25 nm per amino acid residue has been determined for porcine and ovine submaxillary gland mucins (Shogren et al., 1989; Jentoft, 1990). Maintenance of the extended conformation in the porcine and ovine mucins required the O-linked sugars to be present, although the major chain-stiffening effects could be attributed to steric interactions between the peptide-linked GalNAc and adjacent amino acids in the peptide core. Residues attached directly to the GalNAc had some effect on chain stiffness while more peripheral residues were not involved (Rose et al., 1984; Shogren et al., 1986; Gerken et al., 1989). Similarly in the case of the cell surface mucin called episialin (MUC-1), removal of sialic acid did not affect the extended conformation (Bramwell et al., 1986). These data indicate that changes in the outer regions of the carbohydrate structures attached to the leukosialin backbone, as occurs for example in T cell activation (Piller et al., 1988), should not alter the dimensions of the extracellular domain. Evidence that this will be the case comes from the recent work on galactoglycoprotein where it is shown using both electron microscopy and circular dichroism that removal of sialic acid or sialic acid and galactose does not significantly alter the extended conformation of the molecule (Watzlawick et al., 1991). Galactoglycoprotein is now known to be the extracellular domain of human leukosialin (Schmid et al., 1992).

Leukosialin antigenic determinants are of interest as examples of linear epitopes and an understanding of their structural basis provides an indication as to whether the protein part of the extracellular domain is obscured by carbohydrate or free for interaction with other molecules. The surprising finding was that epitopes for 7 of the 9 mAbs studied here were expressed in the unglycosylated protein chain made in *E. coli*. This included the W3/13, OX-58 and OX-75 mAbs for which there is strong evidence that the epitopes recognized are

affected by glycosylation. For all these mAbs, neuraminidase treatment of leukosialin affects binding, and for OX-58 and OX-75 the expression of epitopes on transfected leukosialin is dramatically affected by the host cell type. It could be argued that epitopes for all the other mAbs are also affected by glycosylation, on the basis that they show different patterns of labelling on normal lymphoid cells and cell lines.

The position of epitopes in rat leukosailin was mapped for 7 of the mAbs by expressing fusion proteins of different lengths, and the assignments were confirmed in two cases by electron microscopy on Fab : leukosialin complexes. In addition the epitopes recognized by the OX-56 and OX-74 mAbs were shown to be present within a single 14 residue synthetic peptide. The overall conclusion was that the epitope seen by each of the 7 mAb was determined by a single stretch of protein sequence. Laver and colleagues have recently queried the existence of linear epitopes in native protein structures (Laver et al., 1990) but the epitopes of leukosialin would appear to be unambiguous examples in this category.

It is concluded that the specificity of leukosialin epitopes is largely determined by linear protein sequence, but that their conformation or accessibility can be modified by glycosylation. The M/N antigenic determinants of glycophorin A provide the classical example in which antigenicity appears to be affected by both protein sequence and carbohydrate, but in this case the antigenicity of the unglycosylated form has not, to my knowledge, been characterized (Lisowska and Wasniowska, 1978; Sadler et al., 1979; Prohaska et al., 1981; Lisowska et al., 1987). In another study, a mAb specific for human fibronectin was shown to react with the peptide VTHPGY, but only when the Thr residue had an O-linked sugar attached (Matsuura et al., 1988). It was further demonstrated that the minimal carbohydrate requirement for antigenicity was the presence of GalNAc. Results more similar to the leukosialin data have recently been reported for the cell surface mucin, episialin. In this case mAbs were found that reacted with native, but not asialo episialin, yet these mAbs reacted with β-galactosidase-episialin fusion proteins isolated from lambda gt11 expression cloning systems (Siddiqui et al., 1988; Gendler et al., 1990; Ligtenberg et al., 1990). In addition an immunodominant region of the episialin molecule has been studied by synthesizing peptides corresponding to the mucin repeat segment (Price et al., 1990). A

number of antibodies whose binding had earlier been suggested to involve carbohydrate (Burchell et al., 1983; Sekine et al., 1985) were shown to bind the unglycosylated peptides.

Recent work on CD13 (O'Connell et al., 1991b) showed that variable O-glycosylation was responsible for mAbs being able to precipitate only a fraction of the CD13 from a given cell population. Different CD13 mAbs precipitated a different fraction of the pool. In addition to the possibility of subtle differences in oligosaccharide composition causing variable masking of epitopes, they suggested there could be differential utilization of glycosylation sites. Differential usage of glycosylation sites is unable to explain all the observations with leukosialin because treatment with neuraminidase alone was able to cause the masking of protein epitopes. Furthermore, the observations with human galacto-glycoprotein that 82 of the 92 ser+thr residues are likely to be glycosylated (Schmid et al., 1992) leaves little room for differential utilization of sites. The possibility could be a factor in some cases, however, and it will be important to assess the frequency with which a given residue is glycosylated and whether this differs between cell types. This should be possible using a modified protocol of Edman degradation where the glycosylated form of the amino acid is detected (Gooley et al., 1991).

The O-linked carbohydrate synthesized by Jurkat and K562 cells have been characterized. Jurkat cells generate mainly GalNacα1-Ser/Thr with only minor amounts of more complex O-glycans (Piller et al., 1990). K562 cells synthesize the type 1 core and carry predominantly mono-sialylated forms (Carlsson et al., 1986). The lack of reactivity of OX-58 and OX-75 with both cell lines indicates that the epitope is disrupted by the attachment of GalNac alone and extension to core 1 structures carrying a single sialic acid is inadequate to restore the epitope. Thus aspects of the O-linked sugar structure both near to and distant from the peptide backbone can affect mAb binding to the peptide core.

Similar observations to those described in this chapter have been made for a heavily N-glycosylated glycoprotein, the herpes simplex virus glycoprotein C. A number of mAb specific for a region of this molecule rich in N-glycans fail to bind following removal of sialic acid or sialic acid and terminal galactose. Both sets of mAbs were able to bind the

glycoprotein isolated from virus grown in the presence of tunicamycin and thus lacking N-linked glycan structures, although the level of binding was somewhat reduced (Sjoblom et al., 1987). Expression of the virus in a cell line which attached fucose in place of terminal galactose supported mAb binding and in this case fucosidase treatment, but not galactosidase treatment, abolished activity of the carbohydrate-dependent epitopes (Olofsson et al., 1990).

It may thus be the general case with antibodies raised against mammalian cell surface proteins that carbohydrate can modify protein epitopes rather than being part of the contact sites for antibody binding.

In this chapter the dimensions of the extracellular part of the leukosialin molecule are determined and it is established that the antigenic determinants are largely based on linear protein sequences, with glycosylation having a modulating influence on epitope conformation or accessibility. The implications of these findings for cell-cell interactions are discussed in Chapter 9.

CHAPTER 6

CHARACTERIZATION OF THE CYTOPLASMIC DOMAIN OF LEUKOSIALIN

6.1 Introduction

6.2 Expression of the leukosialin cytoplasmic domain in *E.coli* and its sensitivity to degradation following cell lysis

6.3 Purification of the cytoplasmic domain from the Omp T negative *E.coli* strain UT5600 and cleavage with thrombin

6.4 Hydrodynamic properties of the leukosialin cytoplasmic domain

6.5 NMR spectroscopy analysis of the full length cytoplasmic domain

6.6 *In vitro* phosphorylation by protein kinase C

6.7 Identification of a cytoplasmic domain binding protein

- 6.7.1 Ligand Blotting Analysis indicates the presence of a dominant cytoplasmic domain binding protein in cell lysates
- 6.7.2 Purification of the 110kd protein
- 6.7.3 The 110kd band corresponds to nucleolin

6.8 Discussion

6.1 Introduction

The 124 amino acid cytoplasmic domain of leukosialin is highly conserved between species arguing that the domain plays an important role in the function of leukosialin. It is constitutively phosphorylated in resting peripheral blood lymphocytes and the level of phosphorylation can be rapidly increased by stimulation with PMA, Con A or anti-CD3 (Chatila and Geha, 1988; Axelsson and Perlmann, 1989). The phosphorylation of human leukosialin in K562 and Jurkat cells is restricted to Ser residues (Piller et al., 1989b). Indirect evidence for the association of leukosialin with the cytoskeleton comes from the observation that leukosialin caps spontaneously to the uropod of thymocytes and T cells whereas CD45 remains evenly distributed over the cell surface (De Petris and Baumgartner, 1982; de Petris, 1984). Some mAbs to human leukosialin (L10) have a co-stimulatory effect on T lymphocyte activation and most recently the cytoplasmic domain was found to be essential for enhancing signal transduction in a T cell hybridoma following interaction with an antigen presenting cell (Park et al., 1991). The following studies were undertaken to examine the structure of the leukosialin cytoplasmic domain, its ability to be phosphorylated by protein kinase C, and to begin an assessment of its interactions with other proteins.

6.2 Expression of the leukosialin cytoplasmic domain in *E.coli* and its sensitivity to degradation following cell lysis

The cytoplasmic domain of rat leukosialin was considered to span amino acids 248 to the COOH-terminal residue 371 (Fig 6.1) and this region (CY-1) was isolated by PCR and inserted into the pGEX-2T expression vector. The construct was introduced into *E.coli* strain MC1061 and following induction, high level expression was observed when whole *E.coli* cells were directly lysed by boiling in SDS sample buffer and run on SDS PAGE (Fig 6.2). When the cells were lysed in the absence of denaturant, however, the fusion protein was degraded and only a product of similar size to the GST could be purified. This degradation was not reduced by the standard protease inhibitor cocktail of EDTA, PMSF, aprotinin, iodoacetamide and leupeptin. Recognizing that poly-basic regions of proteins are targets for a number of trypsin type proteases, a second construct (CY-2) was designed in which the NH_2-terminal poly-basic region was deleted (Fig 6.1). This construct was also efficiently expressed but degradation was

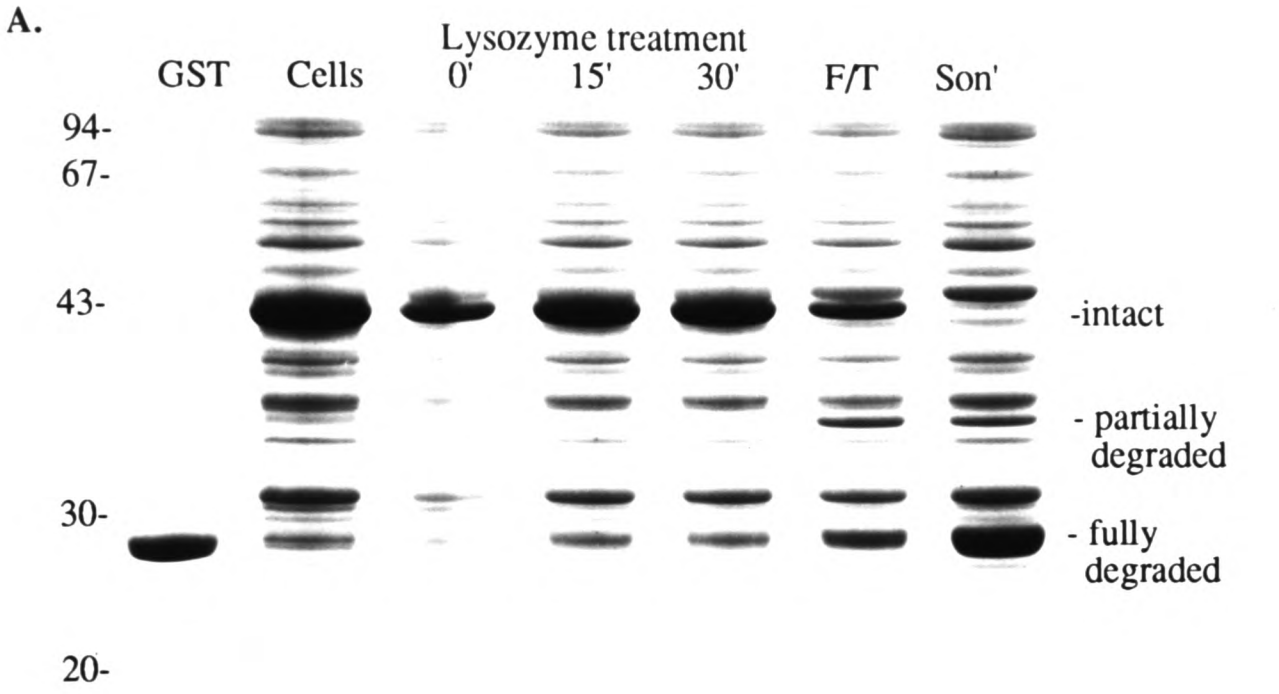


Figure 6.2 Leukosialin cytoplasmic domain fusion protein degradation following cell lysis (A) is inhibited by benzamidine (B). A) A 5ml induced culture of the CY-1 construct in *E.coli* MC1061 was pelleted, resuspended in 800µl tris/sal/EDTA, lysozyme was added to 0.2mg/ml and the cells incubated for 30 min followed by freeze-thawing (F/T) x3 and sonication (Son') for 3x10 seconds. 10µl aliquots were removed at the indicated points, mixed with 30µl of sample buffer, boiled and 8µl was loaded per lane. B) A 300ml induced culture of the CY-2 construct in *E.coli* MC1061 was pelleted, resuspended in 7 ml tris/sal/EDTA and 0.7ml aliquots were transferred to each of 9 eppendorf tubes and the following protease inhibitors added to the final concentrations: 1. nil, 2. PMSF 1mM, 3. Iodoacetamide 2mM, 4. Soy-bean trypsin inhibitor 40µg/ml, 5. 1,10-phenanthroline 80µM, 6. benzamidine 50mM, 7. aprotinin 1/100, 8. leupeptin+pepstatin A 20µM, 9. DFP 10mM. Following lysozyme and DNAse treatments more protease inhibitor was added and the cells lysed by freeze/thawing and sonication. C indicates cells and C+L cells treated with lysozyme, boiled directly in sample buffer. Samples were analyzed as in (A). Numbering on the left indicates apparent Mr (kd) of marker proteins.

again observed following cell lysis. In this case the degraded product was larger than previously, indicating that removal of the NH_2-terminal basic region had protected the junction from proteolysis and suggesting that at least one of the proteases involved was specific for basic residues. Expression of the CY-1 fusion protein in a strain (Y1090) defective in endoprotease La did not lead to a reduction in degradation. A wider panel of proteolytic inhibitors was tested and it was oberved that benzamidine, an arginine analogue, was very effective at reducing degradation (Fig 6.2). When cells were lysed in the presence of 50mM benzamidine, approximately 60% of the protein purified was of the expected M_r for the intact fusion protein with the second major product corresponding to a form that had been cleaved near the centre of the cytoplasmic domain. The existence of an outer-membrane protease of *E.coli* that is specific for poly-basic regions and sensitive to inhibition by benzamidine has been previously reported (Grodberg and Dunn, 1988; Sugimura and Higashi, 1988). The protease, designated Omp T (outer membrane protease T), was reported also to be partially sensitive to inhibition by Cu^{2+}, Zn^{2+} and Fe^{2+} (Sugimura and Nishihara, 1988). $CuCl_2$, $ZnCl_2$ and $FeSO_4$ were thus tested (2 mM) and found to reduce degradation of the cytoplasmic domain fusion proteins (Fig 6.3) with a similar relative potency to that reported (Zn was most effective and Fe least), providing strong evidence that Omp T was the dominant protease involved.

6.3 Purification of the cytoplasmic domain from the Omp T negative *E.coli* strain UT5600 and cleavage with thrombin

E.coli strains that do not express the OmpT protease have been identified (Grodberg and Dunn, 1988) and one of these was obtained from B. Bachmann (*E.coli* Genetics Stock Centre, Yale), termed UT5600. Both constructs were transformed into this strain and the level of degradation following lysis was markedly reduced (Fig 6.3). Greater than 90% of the material purified from the UT5600 strain is intact fusion protein compared to less than 60% from MC1061 in the presence of benzamidine (Figs 6.3 and 6.4).

The purified fusion proteins were efficiently cleaved by thrombin after 1 hour at 37°C. Following release from the GST the intact cytoplasmic domain showed unusual migration characteristics on SDS-PAGE, running at approximately twice its predicted molecular weight (approximately 22 kd compared to 12.8 kd, Fig 6.4). Aberrant migration of proteins rich in

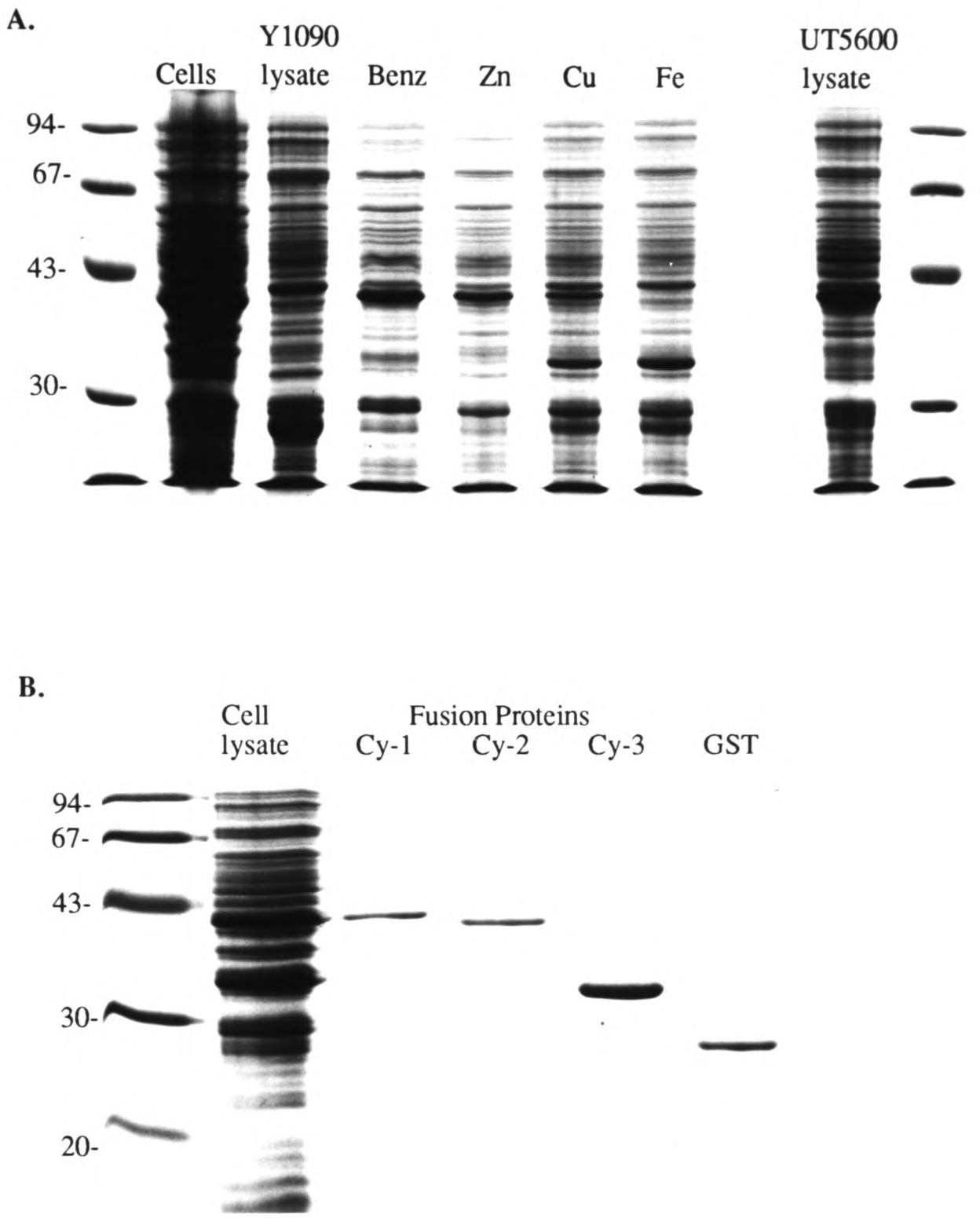
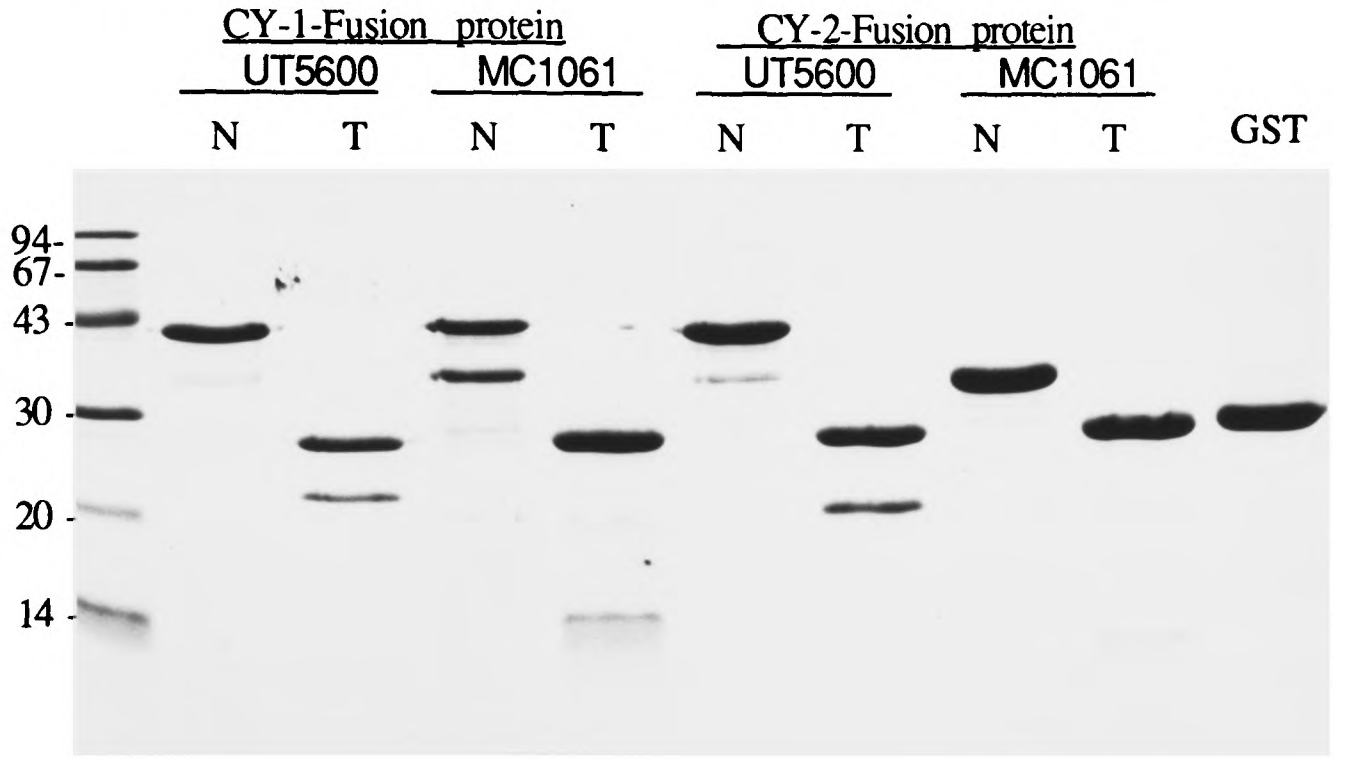


Figure 6.3 Characterization of proteolytic activity and purification of intact fusion proteins. A) Sensitivity of proteolytic activity to inhibition by Zn^{2+}, Cu^{2+}, and Fe^{2+}. 5ml induced cultures of the CY-1 construct in *E.coli* Y1090 were boiled directly in sample buffer (Cells) or lysed either in the absence of protease inhibitors (Y1090 lysate) or in the presence of 2mM $ZnCl_2$ (Zn), 2mM $CuCl_2$ (Cu) or 2mM $FeSO_4$ (Fe) before mixing with sample buffer and running on 15% SDS PAGE as indicated. In the lane UT5600 lysate, a 5ml induced culture of the CY-1 construct in *E.coli* UT5600 cells was lysed in the absence of protease inhibitors and an identical sample loaded as for Y1090 lysate. Partially degraded fusion protein is present in lanes Cu and Fe and no intact fusion protein can be seen in the Y1090 lysate lane. Other lanes contain intact fusion protein. B) CY-1 and CY-2 fusion protein was purified from Omp T negative UT5600 cells and 2 μg loaded per lane. CY-3 (see section 6.6) was isolated by growing CY-2 in an Omp T positive strain MC1061 in the absence of benzamidine and 4 μg was loaded. The cell lysate sample corresponds to MC1061 cells expressing the CY-1 construct lysed in the presence of benzamidine. The leftmost lanes contain marker proteins of the indicated Mr (kd).

A. Thrombin cleavage of cytoplasmic domain fusion proteins



B. Purification of leukosialin cytoplasmic domain by gel filtration

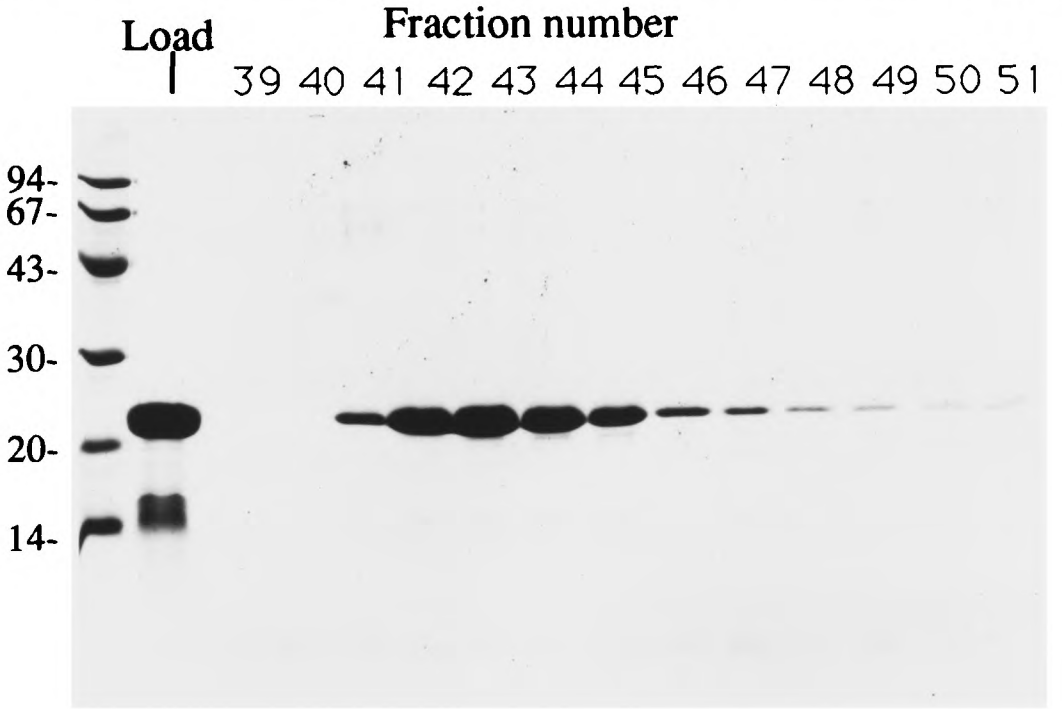


Figure 6.4 Thrombin cleavage of cytoplasmic domain fusion proteins produced in Omp T positive and negative *E.coli* strains (A) and purification of the full length domain by gel filtration. A) Approximately 4 μg of the indicated fusion proteins not treated (N) or treated (T) with thrombin for 60 min at 37°C were run on 18% SDS PAGE. The fusion proteins were purified from the *E.coli* strains indicated. CY-1 purification from MC1061 cells was in the presence of 50mM benzamidine whereas CY-2 was purified in its absence. GST is included in the rightmost lane for comparison. B) Thrombin cleaved CY-1 fusion protein was passed over glutathione agarose to remove the GST and 5mg of cytoplasmic domain was loaded onto a Sephadex G75 column and 2ml fractions collected. 5μl of the column load sample and 10μl of the indicated fractions were analyzed. The leftmost lanes in both gels contain marker proteins of the indicated Mr (kd). The leukosialin cytoplasmic domain runs at approximately twice its molecular weight on SDS PAGE.

basic residues has been previously reported (See and Jackowski, 1989). To confirm the nature of the material released by thrombin treatment, the GST was removed and a sample subjected to Edman degradation. A single sequence was obtained with the domain from UT5600 cells (Table 6.1), that was the expected NH_2-terminal sequence with the two additional residues Gly Ser derived from the thrombin cleavage site. The material from the MC1061 cells generated two sequences at similar molar ratios (Table 6.1) corresponding to the NH_2-terminal sequence of the intact domain and sequence starting at Arg 317.

Omp T is specific for exposed di- or poly-basic regions of proteins (Sugimura and Nishihara, 1988) and such a specificity is in agreement with the results obtained in that there are two dominant sites of proteolysis in the cytoplasmic domain, corresponding to the location of two of the three di- or poly-basic regions (Fig 6.5).

6.4 Hydrodynamic properties of the leukosialin cytoplasmic domain

The hydrodynamic properties of the leukosialin cytoplasmic domain were compared with those of three molecules that have compact globular structures (Schulz and Schirmer, 1979), carbonic anhydrase (29 kd), soybean trypsin inhibitor (20.1 kd), and cytochrome c (12.4 kd). Size exclusion analysis using a Sephadex G75 gel filtration column demonstrated that the cytoplasmic domain has a larger Stokes radius than all of these proteins (Fig 6.6a), including carbonic anhydrase, a protein of more than twice its molecular weight. To test the possibility that the domain was present in solution as a dimer, sucrose gradients were run. Following sedimentation through linear 10-30% sucrose gradients the leukosialin cytoplasmic domain migrated at a similar location to cytochrome c (Fig 6.6b) which has a similar Mr (12.4 kd), that is, it migrated in approximate accordance with the Mr deduced from the sequence. Sedimentation velocity is less dependent on asymmetry than gel filtration (Siegel and Monty, 1966) and therefore it is unlikely that the cytoplasmic domain exists as a dimer.

The potential effect of phosphorylation on the structure of the domain was tested by performing sucrose gradient analysis of material which had been phosphorylated to the level of approximately 2 molecules of phosphate per molecule of cytoplasmic domain (see below). No change in the sedimentation rate could be detected when both phosphorylated and

Table 6.1 NH_2-terminal sequence analysis of recombinant leukosialin cytoplasmic domain

	CY-1 UT5600		CY-1 MC1061*			
No.	AA	pmol	AA	pmol	AA	pmol
1	G	105	G	82	R	51
2	S	17	S	53	K	68
3	R	48	R	42	S	46
4	Q	69	Q	55	R	23
5	R	56	R	27	Q	64
6	Q	69	Q	78	G	68
7	K	57	K	27	S	9
8	R	48	R	12	V	22
9	R	52	R	14	A	24
10	T	27	T	19	L	27
11	G	59				
12	A	56				
13	L	63				
14	T	25				
15	L	60				
16	S	11				
17	R	32				
18	G	43				
19	G	46				
20	K	26				

CY-1was isolated from the indicated *E.coli* strain in the presence of benzamidine.
*Two residues were released per cycle with the MC1061 material at similar molar ratios.

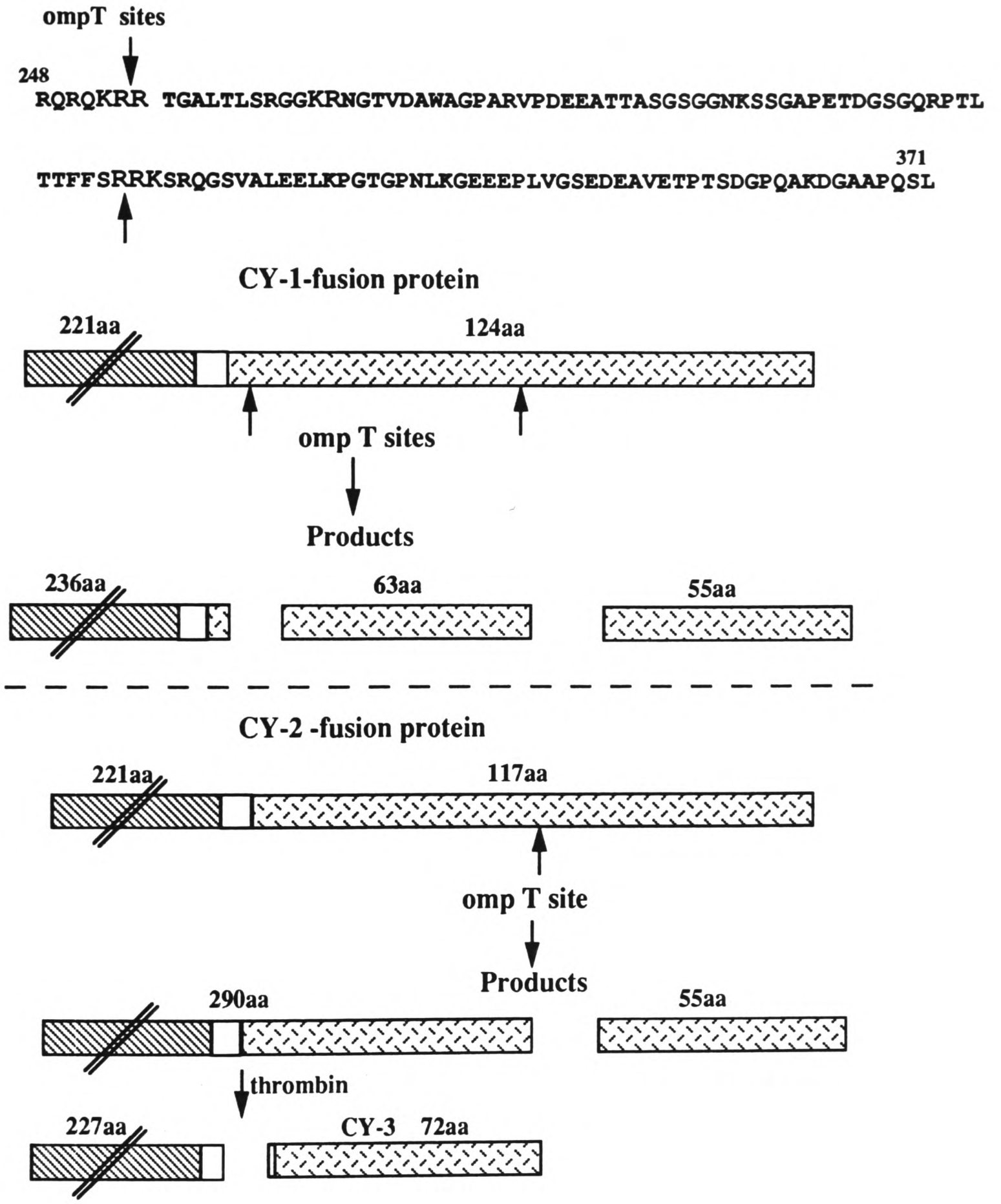
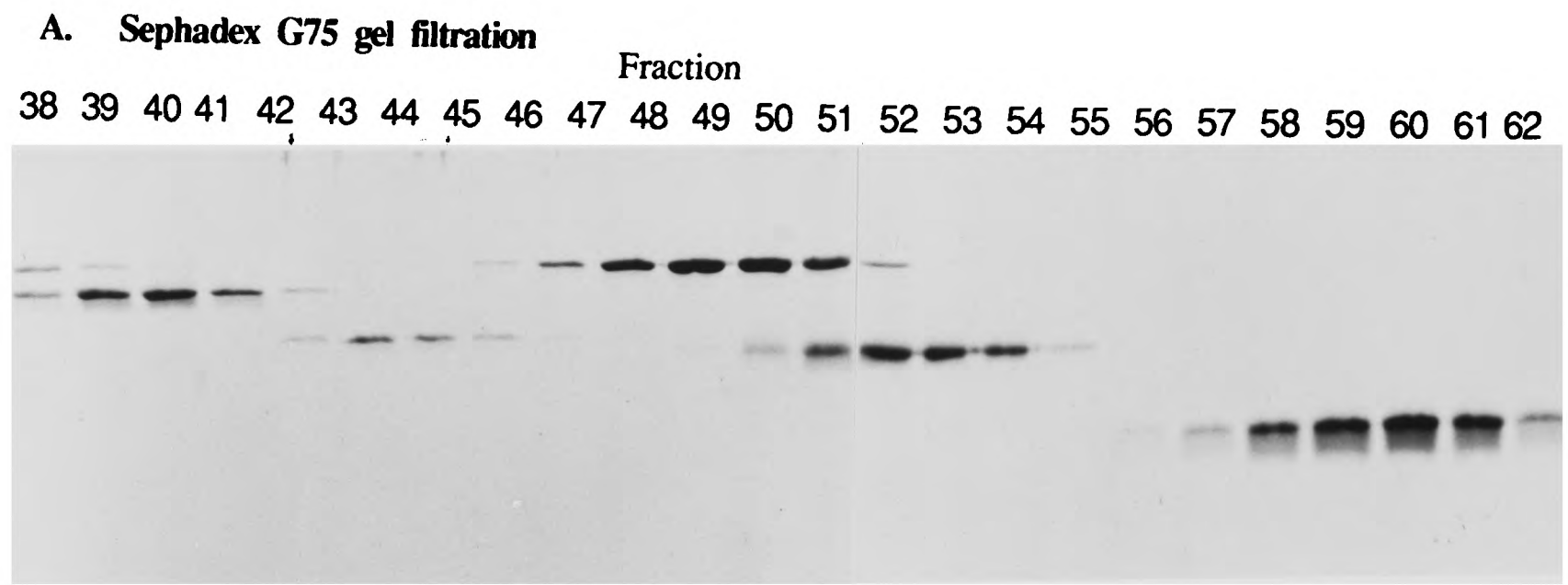


Figure 6.5 Cleavage of leukosialin cytoplasmic domain fusion proteins by outer membrane protease T. The location of the cleavage sites in the amino acid sequence is shown and the products generated by omp T cleavage of CY-1 and CY-2 fusion proteins illustrated. The size of the cleavage products generated in amino acids (aa) is shown. Also shown is the generation of fragment CY-3 by cleavage of the truncated protein remaining after omp T proteolysis of the CY-2 fusion protein with thrombin.

Figure 6.6 Hydrodynamic properties of the leukosialin cytoplasmic domain. A) A protein mixture containing 10µg of carbonic anhydrase, soy-bean trypsin inhibitor, cytochrome c and thrombin cleaved CY-1 fusion protein was loaded onto a Sephadex G75 gel filtration column (100x1cm) at a flow rate of 10 ml/hr and 2.1 ml fractions were collected. 10µl samples from the indicated fractions were analyzed. B) An identical protein mixture to (A) was loaded onto a linear 10-30% sucrose gradient and centrifuged at 260 000g for 20 hr. Fractions were collected from the base of the tube and 10µl samples were analyzed. Lettering on the right indicates the position of each protein in the mixture on SDS PAGE; a. carbonic anhydrase (29 kd), b. glutathione-S-transferase (27 kd), c. leukosialin cytoplasmic domain (12.8 kd), d. soy-bean trypsin inhibitor (20.1 kd) and e. cytochrome c (12.4 kd). The leukosialin cytoplasmic domain runs at approximately twice its molecular weight on SDS PAGE.



ıphosphorylated material was run on the same gradient.

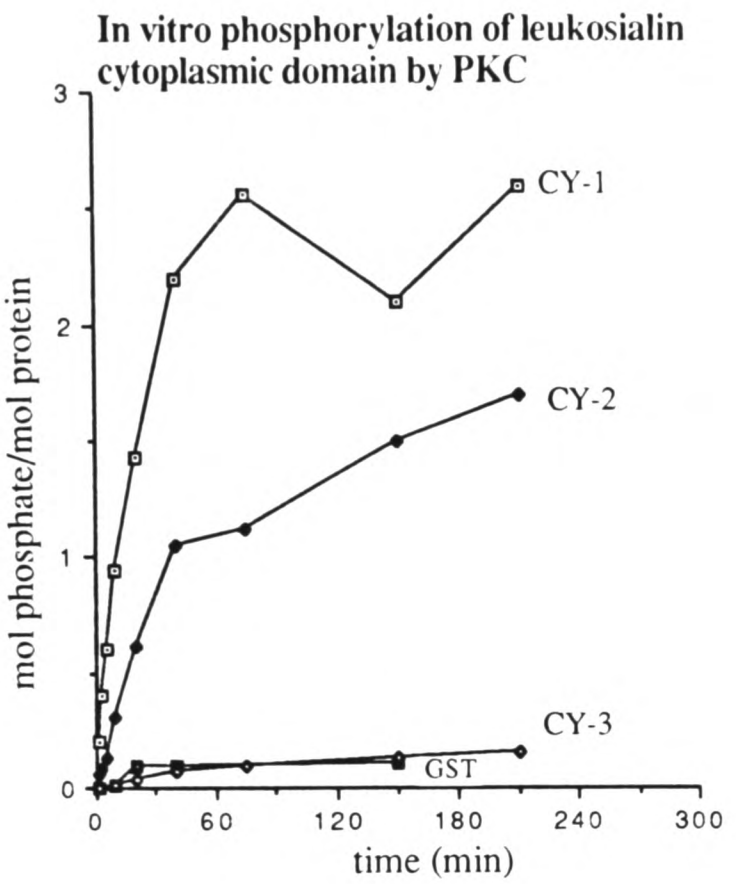
5 NMR spectroscopy analysis of the full length cytoplasmic domain

) mg of the cytoplasmic domain was prepared as a lyophilised powder from material purified y gel filtration in ammonium bicarbonate (Fig 6.4b). The protein was resuspended in 90% $_2O$/10% D_2O and a ^2H spectrum was recorded on a 600 MHz NMR machine by Paul Driscoll. he spectrum exhibited very limited chemical shift dispersion and there were no resolved amide H signals that would be expected in the presence of a compact folded (H-bonded) structure. his spectrum is characteristic of a random coil (Paul Driscoll, pers. comm.) although it does ot rule out the possibility that there are local regions of secondary structure.

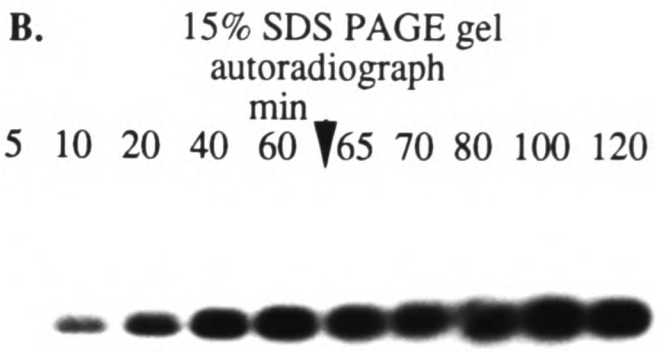
.6 *In vitro* phosphorylation by protein kinase C

hosphorylation of the full length cytoplasmic domain (CY-1) by PKC lead to the incorporation pproximately 1 mole phosphate per mole of protein by 10 min and 2 moles by 30 min (Fig .7). The rate of addition of further phosphate was significantly reduced. The form of the ytoplasmic domain lacking the NH_2-terminal basic residues (CY-2) showed a similar rate of ıcorporation of the first phosphate group but addition of the second group was slower, howing a rate more similar to that of the third group in CY-1 (Fig 6.7). This indicates that emoval of the NH2-terminal basic residues destroys dominant PKC site(s). By producing the :Y-2 construct in an Omp T positive *E.coli* strain it was possible to purify a fragment, CY-3, panning amino acids 255 to 316 (Fig 6.5 and Fig 6.3). This fragment did not show any ignificant incorporation of phosphate over a 4 hr period (Fig 6.7) making it unlikely that Ser 95 or 296 is a natural target for phosphorylation. Residues near the NH_2- or COOH-termini of he CY-3 fragment cannot be excluded as potential sites as these residues would be adjacent to egions rich in basic amino acids in their normal context (Fig 6.5). Direct NH_2-terminal equencing of phosphorylated CY-1 revealed that the Ser at position 2, derived from the hrombin cleavage site, and the Thr at position 24 (amino acid 269; the large number of counts luted at cycle 25 is consistent with the low yield of Thr at cycle 24) was phosphorylated (Fig .8). As a fraction of the total phosphate bound to the domain the two residues accounted for pproximately 10 and 14 %, respectively. To map internal phosphorylation sites, hosphorylated CY-1 was cleaved with V8 protease and the peptides purified by HPLC. The

A.



B.



C.

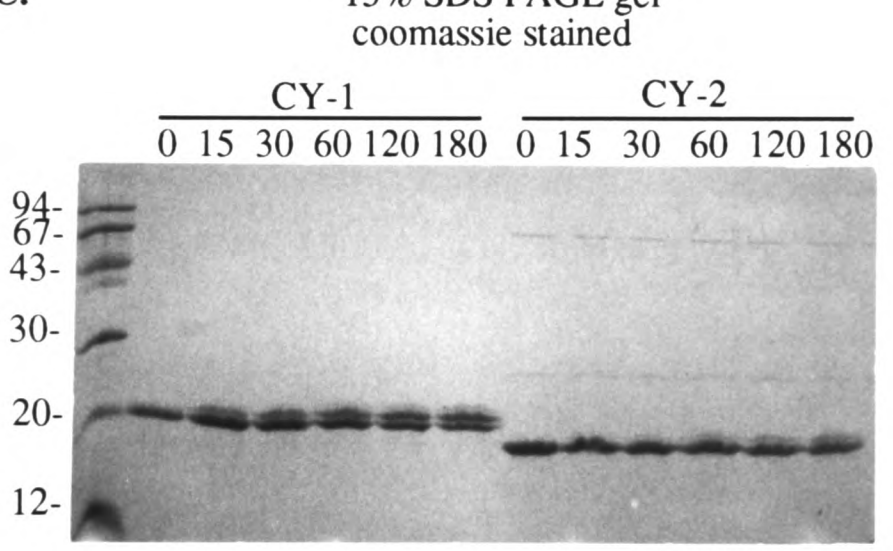
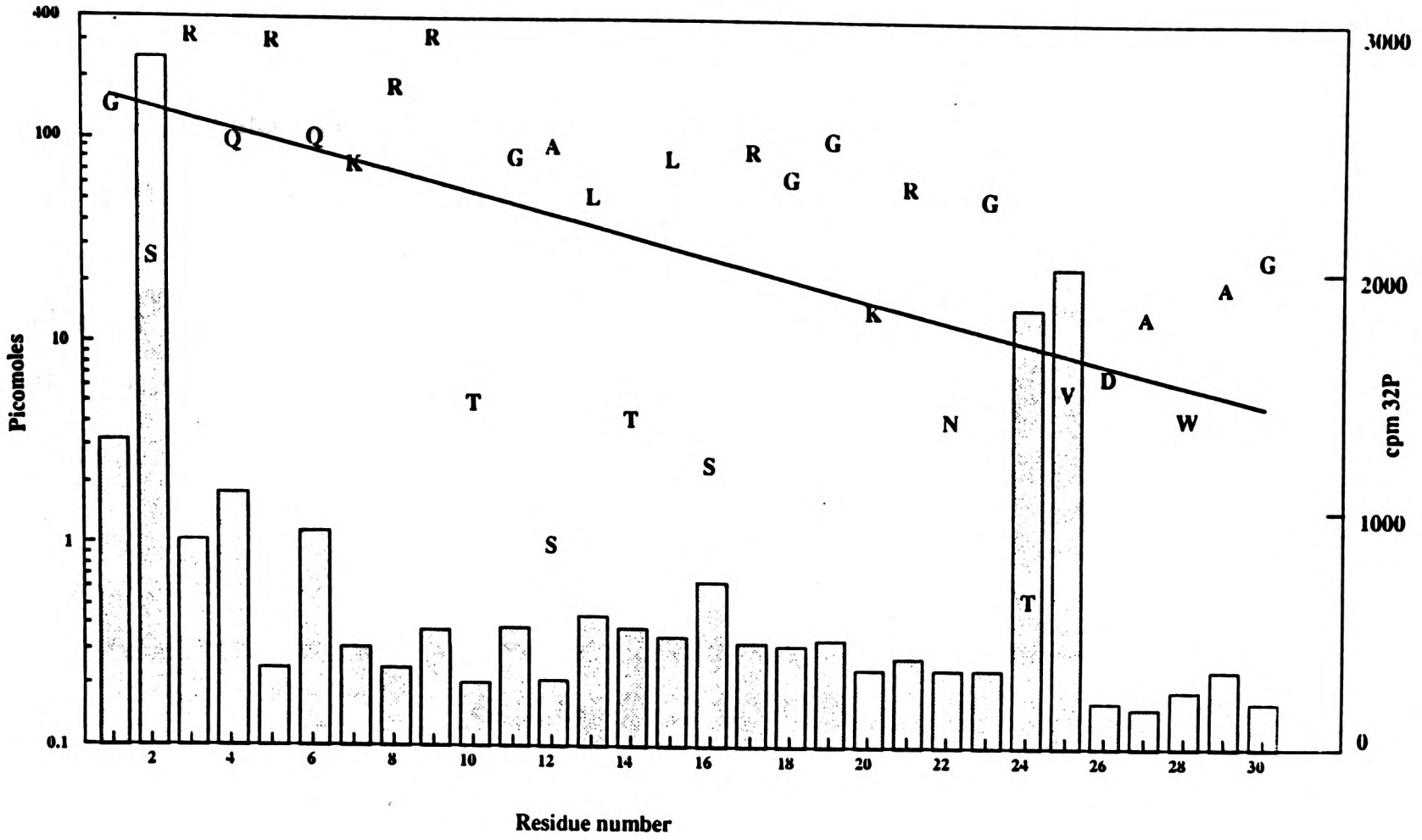


Figure 6.7 *In vitro* phosphorylation of leukosialin cytoplasmic domain by protein kinase C. A) 3 μg of the CY-1, CY-2 or CY-3 cytoplasmic domain constructs or of GST in a volume of 20 μl were incubated with PKC and ATP as described in Chapter 2 and 8 μl aliquots were isolated at successive time points and mixed with 10 μl reducing sample buffer. The samples were run on SDS PAGE and the bands cut from the dried gel and counted and the moles of phosphate present per mole of protein were determined. The CY-1 construct was purified by gel filtration to greater than 98% purity; the CY-2 and CY-3 constructs were only about 90% pure by SDS PAGE (see section C) and thus the moles of phosphate incorporated per mole of CY-2 protein is not exact. B) CY-1 was treated as in (A) except after 60 min fresh PKC was added and the incubation continued. The SDS PAGE gel was dried down and exposed to autoradiograph film for 30 min at room temperature. C) CY-1 and CY-2 were treated with PKC and samples extracted at the indicated times, run on SDS PAGE and the gel was stained with coomassie blue.

A.



B.

248 V1 V2

GSRQRQKRRTGALTLSRGGKRNGTVDAWAGPARVPDEEATTASGSGGNKSSGAPETDGSGQRPTL

V3 V4 V5 V6 371

TTFFSRRKSRQGSVALEELKPGTGPNLKGEEEPLVGSEDEAVETPTSDGPQAKDGAAPQSL

Peak	Peptide	cpm
1	V6	254
2	V4+V5	88
3	u.d.	453
4	V1+V2+V3	34 517
5	V1+V2+V3	51 921

Figure 6.8 NH_2-terminal sequence analysis of CY-1 to detect phosphorylated amino acids (A) and phosphorylated peptides isolated by V8 protease digestion and HPLC (B).
A) CY-1 was phosphorylated with PKC for 15 min, the sample desalted using a G50 column, covalently bound to Sequelon-DITC membrane and subjected to NH_2-terminal sequencing on a ABI 473A sequencer. The total activity of the sample bound to the membrane was 140 090 cpm. 20% of the residue released at each cycle was counted. B) Peptides generated by V8 protease digestion of CY-1 phosphorylated as in (A) were isolated by HPLC and counted. u.d. indicates an undefined peptide whose amino acid composition did not correspond to sequence in CY-1. Peptides V1, V2 and V3 did not elute from the column as isolated peptides but as two broad peaks composed of mixtures of all three. In an unphosphorylated sample V2 was isolated in a separate peak. V4 and V5 did not elute as independent peptides in either the phosphorylated or unphosphorylated samples, possibly due to the adjacent Pro residue.

peptide(s) isolated in each peak were defined by compositional analysis and Fig 6.8b shows the counts isolated in each peak and the corresponding peptide(s) present. The three NH_2 terminal peptides expected from V8 proteolysis were not effectively separated in this analysis but it is clear the the COOH-terminal region of the domain (amino acids 329-371) is not phosphorylated. Further mapping will be necessary to define all the residues phosphorylated in the NH_2-terminal 81 amino acids.

SDS PAGE analysis of phosphorylated CY-1 or CY-2 revealed that the single band separated into a doublet, consisting of one band at the same Mr as the unphosphorylated material and one band at a higher Mr, as phosphorylation proceeded (Fig 6.7c). Autoradiography showed that both bands carried phosphate (Fig 6.7b). With extended phosphorylation of CY-1 the band often became a triplet (Fig 6.7c) although in no case was the lower band observed to be converted totally into higher M_r bands. Incorporation of phosphate into proteins is known to increase their apparent M_r on SDS PAGE (Casnellie and Lamberts, 1986) but it is unclear why a significant amount of phosphate could be incorporated into the lower band without causing a shift. The shift in Mr may only result from incorporation at some sites.

6.7 Identification of a cytoplasmic domain binding protein

6.7.1 Ligand Blotting Analysis indicates the presence of a dominant cytoplasmic domain binding protein in cell lysates

The CY-1 cytoplasmic domain fusion protein was iodinated by the chloramine T technique and used to probe nitrocellulose blots of cell lysates run on SDS PAGE, in a manner similar to Western blotting. Lysates from Y3 cells showed a dominant reactive band at approx. 110 kd (Fig 6.9a). To confirm that the interaction of the iodinated fusion protein preparation was through the leukosialin cytoplasmic domain, the iodinated material was treated with thrombin. The cytoplasmic domain does not contain tyrosine and thus should not contain label and this was confirmed by SDS PAGE of iodinated CY-1 fusion protein before and after thrombin treatment. When the thrombin cleaved probe was used in a blotting assay (Fig 6.9a) no binding was detected. Competitive inhibition of probe binding was performed using unlabelled CY-1 fusion protein and a 1000 fold excess was needed for significant inhibition (Fig 6.9c). The requirement for such a large excess is not understood but may indicate that the iodinated material

had a higher affinity for the 110 kd protein than did the unlabelled fusion protein. Iodinated GST or GST-CD2 domain 1 did not show binding. Reactivity of the iodinated fusion protein but not GST alone with a 110 kd band was also seen with lysates of rat TDL, mouse spleen cells and thymocytes and Jurkat cells (Fig 6.9b).

When higher percentage SDS PAGE gels were used to separate lysate proteins before the blotting step, two further bands were evident in some lysate preparations at approximately 20 and 25 kd. As an initial test of the compartmentalization of these proteins, membrane preparations were made using tween 40 or by shearing (Standring and Williams, 1978) and it was found that the 110 kd protein was present in similar amounts in both membrane and nuclear fractions and lesser amounts in the cytosol whereas the lower Mr bands were present only in the nuclear fractions and most likely corresponded to histones (Fig 6.10a). The distribution of the 110 kd protein is consistent with a cytoskeletal component as a major part of the cytoskeleton is insoluble in non-ionic detergent and would pellet with the nucleus (Bourguignon and Bourguignon, 1984).

Sucrose gradient sedimentation analysis of cell membrane preparations is a means of separating membrane vesicles derived from different cellular compartments (Standring and Williams, 1978). Samples of tween membrane preparations from Y3 cells were run on sucrose gradients and it was found that the 110 kd protein present in the membrane fraction migrated at a similar location to the major peak of leukosialin (Fig 6.10b and c). This indicated the two proteins could be present in the same compartment and thus a physiologically relevent interaction was possible.

6.7.2 Purification of the 110 kd protein

To purify the 110 kd protein, soluble CY-1 fusion protein was mixed with cell lysates and then extracted using glutathione-agarose. In initial experiments, the stability of the fusion protein in the lysate was assessed and the effect of detergent on the efficiency of purification tested. It was found that incubation for extended periods (up to 24 hr at 4°C) in NP40 lysates did not lead to significant degradation of the CY-1 fusion protein and the presence of up to 2% NP40 or 1% tween did not reduce binding of the fusion protein to glutathione agarose. Ligand blotting with

the iodinated CY-1 fusion protein was performed following the affinity purification step to assess whether any 110 kd protein was isolated. Initial experiments were unsuccessful but when the amount of CY-1 fusion protein used in the purificaion step was increased to approximately 0.5 mg per ml of lysate, some 110 kd protein was isolated (Fig 6.11a). Direct SDS PAGE analysis of the purified material demonstrated a band at 110 kd that was not present in the control preparation (Fig 6.11b,c). This band corresponded to a dominant band in the tween membrane preparation (Fig 6.11c).

Elution of the 110 kd protein from the fusion protein on the gluthione agarose by increasing ionic strength was examined and complete elution was achieved at salt concentrations of 300 mM or higher (Fig 6.11b). When the CY-2 rather than the CY-1 fusion protein was used as the affinity ligand, no 110 kd material was purified from the lysate dectable by visual inspection of the gel (Fig 6.11c) or by blotting with the iodinated CY-1 fusion protein. This implies that ionic interactions are important and that the NH_2-terminal basic amino acids are involved in the interaction.

6.7.3 The 110 kd band corresponds to nucleolin

Attempts to purify large quantities of the 110 kd protein using a glutathione agarose column were unsuccessful possibly because the interaction with the CY-1 fusion protein is of low affinity or because the 110 kd protein is unstable in the lysate and becomes degraded during the time involved in running and washing the column. Due to the paucity of material which could be directly purified and noting the dominant nature of a 110 kd band in the tween membrane preparations, it was decided to isolate the band directly from the membrane preparation and analyze it by microsequencing. Samples of tween membrane preparation were run on 7% SDS PAGE and transferred to problot membrane, stained with coomassie blue and the material corresponding to 110 kd was excised as a strip and subject to Edman degradation by Tony Willis. The sequence obtained (Table 6.2) is identical to the NH_2-terminus of rat nucleolin (Bourbon and Amalric, 1990).

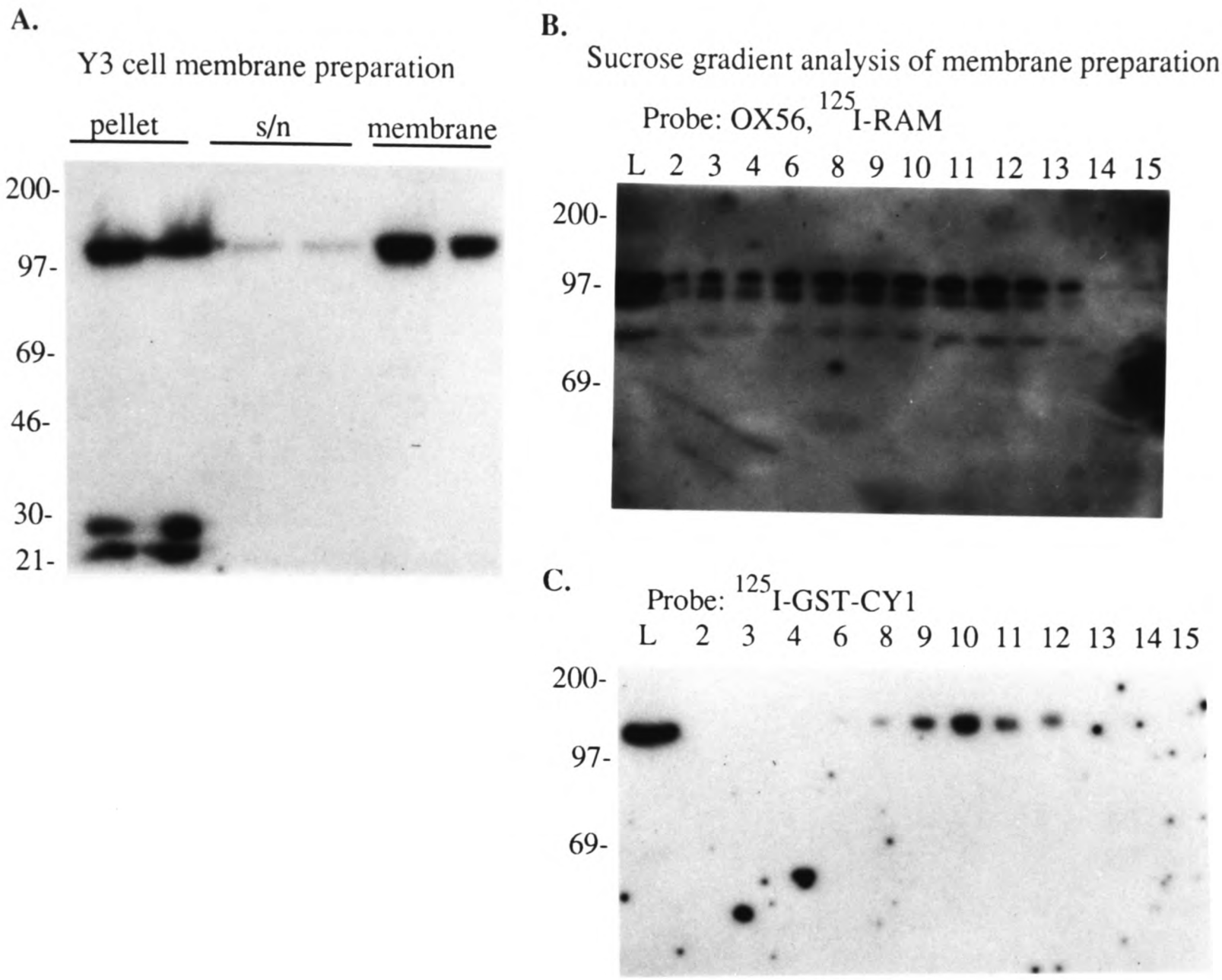
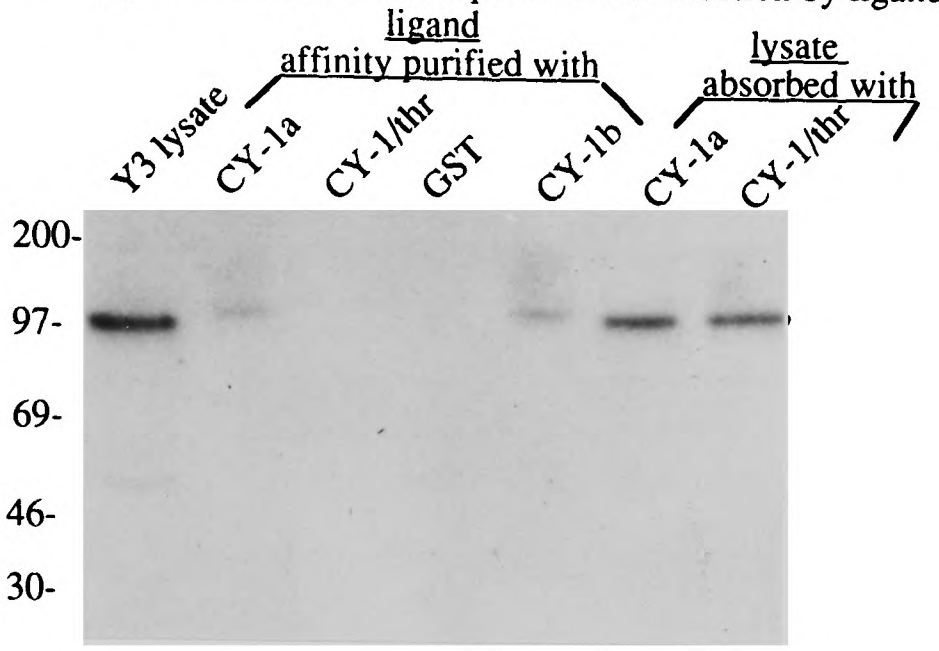


Figure 6.10 The 110kd protein is present in Y3 cell membrane preparations and comigrates with leukosialin on sucrose gradients. A) Y3 cell membranes were prepared using tween 40 (left lanes) or by shearing (right lanes). Samples of the low speed spin (pellet, equivalent to $5x10^6$ cells), the s/n of the high speed spin (s/n, 10^6 cells) and the pellet of the high speed spin (membrane, $5x10^6$ cells) were mixed with equal volumes of non-reducing sample buffer, run on 12 % SDS PAGE and probed with ^{125}I-GST-CY-1. In (B) and (C) 180 µl of the tween membrane sample was loaded onto a stepwise (10, 28, 40%) sucrose gradient and spun at 80 000g for 20 hr. Samples were collected from the base of the tube and 12 µl of the indicated fractions were loaded onto each of two 10 % SDS PAGE gels. After running the protein was transferred to nitrocellulose and probed with (B) OX-56 mAb followed by ^{125}I-RAM or (C) ^{125}I-GST-CY-1. Films were exposed o/n at -70°C. Numbering on the left of the blots indicates apparent Mr (kd) determined from marker proteins.

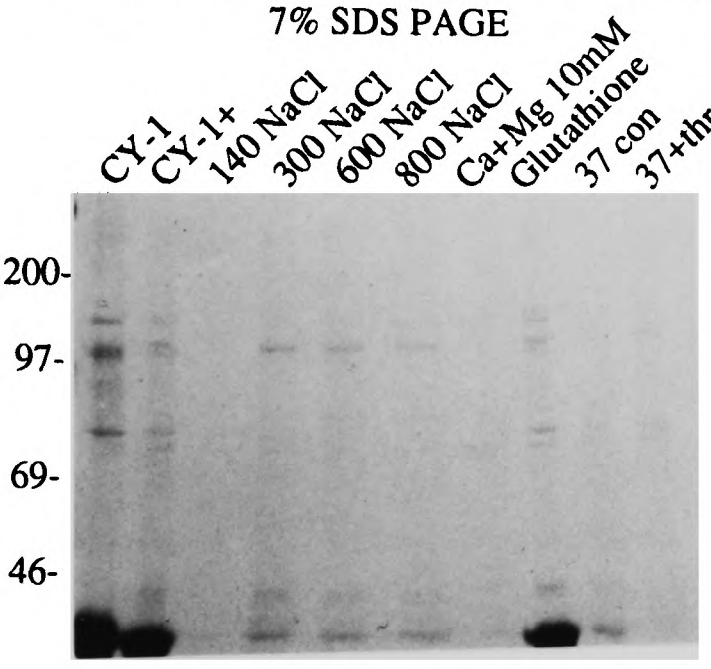
Figure 6.11 Affinity purification of the 110 kd protein from Y3 cell lysates. A) Samples were run on 10 % SDS PAGE, transferred to nitrocellulose and probed with ^{125}I-GST-CY-1. The samples loaded in each lane were: lane 1, Y3 lysate; lanes 2-5, material purified from the lysate with the indicated fusion protein; lanes 6 and 7, samples of lysate after absorbtion with the indicated fusion protein. CY-1a and CY-1b are separate samples that were treated identically. CY-1/thr indicates the CY-1 fusion protein had been cleaved with thrombin before use as an affinity ligand.
In (B) and (C) samples of CY-1 fusion protein that had been incubated with Y3 lysate and absorbed on glutathione agarose (Chapter 2, section 2.19) were transferred to 10 Eppendorf tubes and incubated with lysis buffer containing the indicated amount of NaCl (mM) or $CaCl_2$+$MgCl_2$, or with glutathione elution buffer, for 30 min at 4°C. Some samples were also incubated at 37°C in lysis buffer with or without 2 units of thrombin. In (B) samples of the eluted material were mixed with equal volumes of non-reducing sample buffer and run on SDS PAGE. In (C) the agarose beads were washed, resuspended in non-reducing sample buffer and run on SDS PAGE. CY-1 is a sample of untreated fusion protein and CY-1+, material bound to the glutathione agarose before incubation with any of the indicated elution buffers. Gels were stained with coomassie blue. It is seen that a 110 kd band is eluted following incubation in 300-800 mM NaCl and by incubation with thrombin whereas all the bands are eluted by glutathione. The dominant band at 40 kd is the fusion protein. The location of the uneluted 110 kd band in (C) is indicated by an arrow. D) A Y3 cell tween membrane preparation was absorbed with CY-1 or CY-2 fusion protein and after washing the complex on the glutathione agarose, bound protein was eluted with 300 mM NaCl. The samples loaded were: Y3 Tween, 4 µl of the membrane prepartion after absorption with CY-1; CY-1a, the elution of the CY-1 bound sample; CY-2, the elution of the CY-2 sample; CY-1b, identical to CY-1a except 2% NP40 was added to the membrane preparation prior to incubation with the fusion protein; CY-1c, identical to CY-1a except 1% NP40 and 1% DOC were added to the membrane preparation. Numbering on the left of the gels indicates apparent Mr (kd) determined from marker proteins.

A. Affinity purification of 110 kd protein and detection by ligand blotting



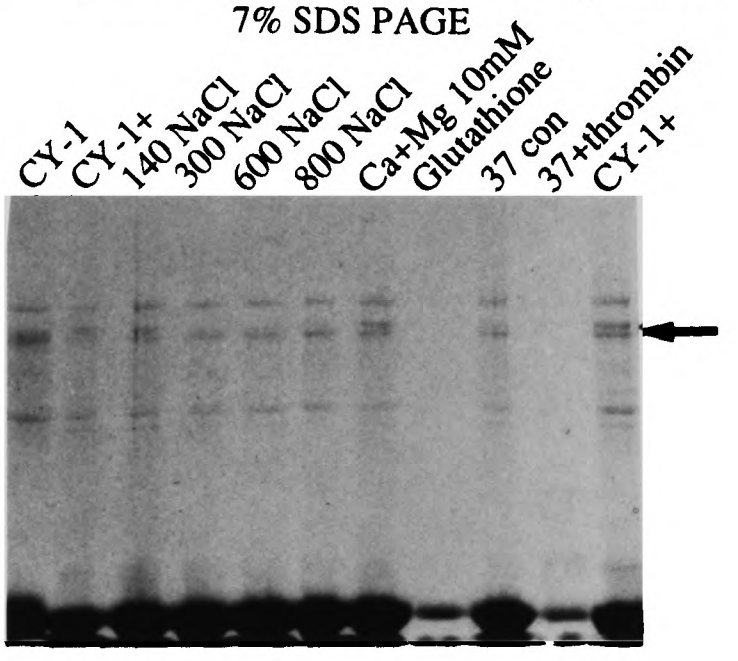
B.

Elution of 110 kd protein from CY-1 fusion protein
7% SDS PAGE



C.

Protein remaining bound following elution
7% SDS PAGE



D. CY-2 does not bind the 110 kd protein
7% SDS PAGE

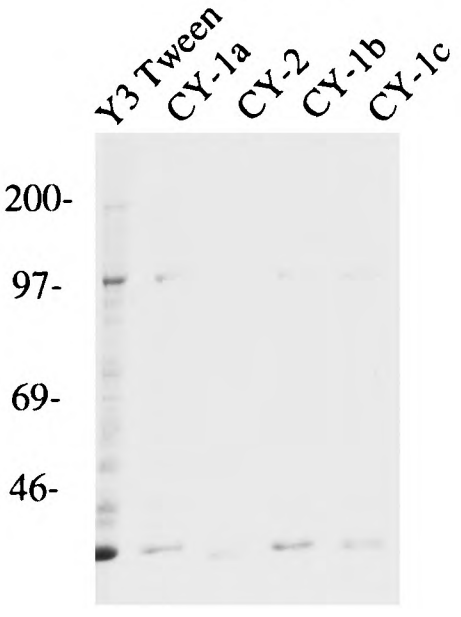


Table 6.2 NH_2-terminal sequence analysis of 110 kd band in Tween membrane preparation

	110 kd Band	
No.	amino acid	picomol
1	V	5.6
2	K	4.4
3	L	6.3
4	A	5.6
5	K	2.3
6	A	5.4
7	G	4.6
8	K	2.2
9	T	3.8
10	H	1.1
11	G	4.0
12	E	4.1
13	S	0.9
14	K	1.2
15	K	1.3
16	M	2.6

6.8 Discussion

The high level of sequence conservation in the leukosialin cytoplasmic domain and the evidence that this is a result of selective pressure argues that the domain will be involved in interactions with other proteins. *In vitro* expression and purification of large quantities of the domain was chosen as one approach to allow such interactions to be studied. Cytoplasmic domains are generally rich in hydrophilic amino acids which should help reduce the tendency for aggregation and the formation of insoluble inclusion bodies. This may make them more exposed for proteolysis, however, and this was observed for the leukosialin cytoplasmic domain. The specificity of the dominant protease involved, Omp T, for poly-basic regions means that this protease is likely to be a problem for expression of cytoplasmic domains in general as they are frequently rich in basic amino acids, especially near the transmembrane domain where they act as stop-transfer signals. Omp T deficient strains of *E.coli* are available, however, and as shown here their use can significantly reduce degradation problems of this type. Other proteases (Swamy and Goldberg, 1981) are also likely to present problems with other cytoplasmic domains. The rat CD2 cytoplasmic domain was expressed in the pGEX system but it did not accumulate significantly inside the cells and use of *lon*$^-$ or OmpT$^-$ strains did not improve yields.

The hydrodynamic properties of the leukosialin cytoplasmic domain demonstrate that it is very elongated and the lack of slowly exchanging protons in the NMR spectroscopy analysis is in agreement with this observation. This may also explain why the polybasic regions in the domain are susceptible to proteolysis whereas the di- and tri-basic regions in the glutathione-S-transferase (Smith et al., 1986) are not cleaved. The lack of proteolysis between Lys 265 and Arg 266 in the cytoplasmic domain may be due to some aspect of primary sequence specificty of the Omp T not yet defined or to local conformation in this region that obscures the site. The possibility that regions of secondary structure exist which may be of functional importance, such as the reverse-turn that acts as a coated pit internalization signal in a number of cytoplasmic domains (Bansal and Giersach, 1991; Eberie et al., 1991), has not been eliminated.

The generally open conformation of the leukosialin cytoplasmic domain may be important in allowing its phosphorylation by PKC. The substrate specificty of PKC is not totally defined

but the presence of basic residues on one, and preferrably both sides of the Ser or Thr residue, are essential (Kennelly and Krebs, 1991). It is common that there is a spacer residue between the basic amino acids and the Ser or Thr. The PKC preparation used in these experiments was a mixture of bovine α, β and γ isozymes (provided by Peter Parker). It is not clear whether different PKC isozymes have different substrate specificities but the similarity of the pseudosubstrate sequences in the α,β and γ isozymes argues that these members of the PKC family at least will have very similar specificities (Pears and Parker, 1991; Gschwendt et al., 1991). α and β are the dominant species expressed in human T cells (Alexander and Cantrell, 1989). Two molecules of phosphate were rapidly incorporated into the cytoplasmic domain by *in vitro* phosphorylation with this PKC preparation and sequence analysis indicates that the Ser from the thrombin cleavage site at +2 is phosphorylated and the Thr at +24 (amino acid 269). Phosphorylation of the Thr was unexpected because human leukosialin is not phosphorylated on Thr residues (Piller et al., 1989b). However, the site is not conserved in human leukosialin leaving open the possibility that phosphorylation of Thr is important in rat leukosialin. The result was also unexpected because the CY-3 fragment, which begins at residue 255 and encompasses Thr 269 and the adjacent Lys (265) and Arg (266), is not phosphorylated. This indicates some long range effect of amino acids present in CY-1 but absent from CY-3 that is essential for recognition of Thr 269 by PKC. It is possible, although unlikely given the stability of the CY-3 fusion protein isolated in the presence of Omp T, that CY-3 was cleaved at Arg 266 prior to the phosphorylation step. NH_2-terminal sequencing of a sample of CY-3 after incubation with PKC will be necessary to eliminate this possibility

One site of phosphorylation of human leukosialin has been mapped to Ser 332, 75 amino acids from the membrane domain (Piller et al., 1989b). This Ser and the flanking basic amino acids are conserved in rat and mouse leukosialin. Evidence that the residue (Ser 319 in rat leukosialin) is phosphorylated to high levels following *in vitro* phosphorylation of the rat cytoplasmic domain was the observation that the peptide encompassing amino acids 319-338, generated by endoproteinase lys C cleavage, carried the largest amount of phosphate (data not shown due to an ambiguity in the assignment of a number of peaks after HPLC). A number of efforts to sequence the peptide were unsuccessful due to the first residue being blocked (Tony Willis, personal communication) which could result from its being phosphorylated to high

stoichiometry. In summary, two sites that are likely to be phosphorylated in rat leukosialin include Thr 269 and Ser 319. Ser 315 and Ser 323 are also possible candidate sites and have not been excluded by this analysis. Residues further to the COOH-terminus are ruled out by the V8 protease analysis and residues between Thr 269 and Ser 315 are unlikly to be phosphorylated given the lack of phosphorylation of CY-3. Further peptide mapping and sequencing will be necessary to precisely define all the sites and the preference with which they are phosphorylated.

The cytoplasmic domain bound to a prominent band in lysates of Y3 cells and a number of other cell types including mouse and human cells. The sensitivity of the association to ionic strength argues that ionic interactions are important and the observation that, even after an effective affinity purification step, the majority of the reactivity remained in the absorbed lysate argues that the protein is present at high concentration. For these reasons I argue that nucleolin, the protein shown to be present in the tween membrane preparations at the same apparent Mr as the band which is both labelled and affinity purified, is the protein with which the cytoplasmic domain is interacting. Nucleolin is a highly charged protein having a region near the NH_2-terminus containing three stretches of 25, 28 and 33 Glu or Asp residues (Bourbon and Amalric, 1990). The sensitivity of the interaction to ionic strength and the requirement for the NH_2-terminal basic amino acids of the leukosialin cytoplasmic domain argues that the cytoplasmic domain may interact with this acidic region. Nucleolin is expressed at high levels in rapidly dividing cells, such as cell lines, accounting for 10% of total nucleolar proteins (Bugler et al., 1982).

Nucleolin, however, is located largely in the nucleolus of the cell where there is strong evidence that it takes part in ribosomal assembly processes (Olson et al., 1981; Bugler et al., 1982; Pfeifle and Anderer, 1983; Peter et al., 1990). There is some evidence that the molecule shuttles to and from the cytosol (Borer et al., 1989) and that it can be present at the cell membrane (Pfeifle et al., 1981; Pfeifle and Anderer, 1983; Semenkovich et al., 1990). The great majority of evidence supports the role of nucleolin in RNA binding and ribosomal assembly, however, and it is most likely that the interaction observed here is not physiologically relevant. Pfeifle and Anderer (1983) observed that maintenance of the nucleolar location of

nucleolin during disruption of cells required the presence of calcium; cell disruption in the absence of calcium and at low magnesium concentrations rendered most of the nucleolin into a soluble form. The lysis buffer used in this work was free of Ca and Mg and contained EDTA. Indeed, in a recent experiment where cells were lysed in tris/sal alone the reactivity of membrane preparations with the iodinated cytoplasmic domain was markedly reduced.

The iodinated cytoplasmic domain fusion protein was used to screen 3x 10^6 plaques of a mouse spleen lambda gt11 expression library and a similar number of plaques of a rat thymus lambda ZAP expression library but no positive clones were identified. The technique has been used successfully by others both with purified cytoplasmic domains (Skolnik et al., 1991) and with GST-fusion proteins (Blackwood and Eisenman, 1991). Screening with the phosphorylated cytoplasmic domain may be worth considering. Further affinity purification efforts using the CY-1 fusion protein and metabolically labelled cells as a source of cell lysates could also be considered.

In conclusion, the 124 amino acid cytoplasmic domain of rat leukosialin can be expressed at high levels in *E.coli*. The isolated domain is very elongated and lacks significant secondary or tertiary structure. PKC is able to phosphorylate the domain at a number of sites to a level of 2 molecules of phosphate per molecule of cytoplasmic domain after 30 minutes.

CHAPTER 7

THE IMPORTANCE OF CROSS-LINKING AND EPITOPE LOCATION IN THE HOMOTYPIC AGGREGATION OF LYMPHOCYTES INDUCED BY ANTI-LEUKOSIALIN ANTIBODIES

7.1 Introduction

7.2 Lymphocyte aggregation by anti-leukosialin mAbs

7.3 Sensitivity of aggregation induced by anti-leukosialin mAbs to metabolic and cytoskeletal inhibitors

7.4 Importance of divalent cations and lack of effect of mAbs against LFA-1 and ICAM-1 in anti-leukosialin induced aggregation

7.5 Effect of epitope position on aggregation induced by anti-leukosialin mAbs

7.6 Fab fragments are inactive in thymocyte aggregation

7.7 The Fc portion of anti-leukosialin mAb is not required for aggregation

7.8 Discussion

7.1 Introduction

Leukosialin is notable as a target antigen that can strongly mediate the mAb induced clustering of lymphocytes, monocytes and certain cell lines. Axelsson and co-workers (1988) reported that antibodies B1B6 and E11B were able to induce aggregation of human T cells. The aggregation was found to be LFA-1 dependent. Induction of LFA-1 dependent aggregation has also been seen with antibodies to the T cell receptor/CD3 complex (Dustin and Springer, 1989), CD2 (van Kooyk et al., 1989), Ig (Dang and Rock, 1991), CD14 (Lauener et al., 1990), CD19 (Kansas and Tedder, 1991; Smith et al., 1991), CD20 and CD39 (Kansas and Tedder, 1991), CD40 (Barrett et al., 1991), CD44 (Belitsos et al., 1990; Koopman et al., 1990), MHC class II (Mourad et al., 1990; Kansas and Tedder, 1991) and LFA-1 itself (Keizer et al., 1988; Dransfield and Hogg, 1989). In the case of CD43 (Axelsson et al., 1988), CD44 (Belitsos et al., 1990), CD14 (Lauener et al., 1990) and MHC class II (Mourad et al., 1990) it was reported that Fab fragments of the antibodies were capable of inducing clustering. If correct, this would rule out the possibility that molecular cross-linking either on the cell, or between cells was necessary for the aggregation. Another mAb to human leukosialin (L10) has been found to induce clustering of PMA treated monocytes, detectable at 1-2 hours and reaching a maximum at 20 hours. This aggregation event was also LFA-1 dependent and argued to be independent of cell-cell crosslinking by the mAb (Nong et al., 1989).

Antibodies against other antigens have been shown to induce aggregation that is not inhibited by antibodies against LFA-1 and ICAM-1. The target antigens include CD9 (Masellis et al., 1990), Thy-1 (Isobe and Nakashima, 1991), VLA-4 (Bednarczyk and McIntyre, 1990; Campanero et al., 1990), Leu 13 (Evans et al., 1990) and TAPA-1 (Takahashi et al., 1990). Also in those cases where an LFA-1 dependent mechanism was argued, a component of the aggregation was sometimes not inhibited by mAb to LFA-1 or ICAM-1. The physiological significance of antibody induced cell-cell aggregation, particularly when this is homotypic and independent of known adhesion molecules, is unclear. When the aggregation is seen only with the bivalent IgG it may be due to direct cross-linking of cells. Testing the activity of monomeric Fab fragments in the assay is an important step in defining the nature of the adhesion event.

The aim of this study was to compare the ability of anti-rat leukosialin antibodies which bound at different distances from the transmembrane region to induce homotypic aggregation, and to examine in detail the requirement for molecular cross-linking in this event.

7.2 Lymphocyte aggregation by anti-rat leukocyte mAbs

All 9 mAbs specific for rat leukosialin were able to induce rapid and marked aggregation of rat thymocytes. Significant clustering was detectable after 2 hours and increased to a maximum by 48 hours at 37°C (Fig 7.1 and Table 7.1). The antibodies were also tested for clustering of lymph node and spleen cells and all nine were active, some examples being shown in Fig 7.1. The aggregation was not as extensive as with thymocytes, which is probably due in part to the presence of resting B cells that do not express leukosialin. The background aggregation was also higher (Fig 7.1) than with thymocytes where, in the absence of the antibody, there were rarely more than 2% of cells in clumps after 48 hours. All the following studies were performed with freshly isolated thymocytes. A panel of mAbs that bind to rat thymocyte surface molecules was studied (Table 7.1) and it was found that in addition to anti-leukosialin antibodies, the mAbs W3/15, OX2, OX7, and OX10 were active in cell clustering while mAbs specific for 11 other surface antigens had no effect. Of these four active mAb, only OX7 (anti-Thy1) has been shown previously to induce clustering (Isobe and Nakashima, 1991), in that case of murine thymocytes. Significantly, antibodies which were unable to cause aggregation included those specific for highly expressed and elongated molecules such as CD45 and CD44. The ability to induce clustering did not necessarily correlate with a high level of antigen expression since OX2 and W3/15 are found at low levels on the thymocyte surface (Williams et al., 1977; McMaster and Williams, 1979).

7.3 Sensitivity of aggregation induced by anti-leukosialin mAbs to metabolic and cytoskeletal inhibitors

The activity of 4 of the anti-leukosialin mAb was further analyzed using purified IgG (Fig 7.2). The location of the epitopes recognized by these antibodies, mapped in Chapter 5, is illustrated in Fig 7.2. Titration analysis showed OX75 to be the most potent and W3/13 the least (Fig 7.3a) with OX56 and OX57 being of intermediate activity. The maximum level of aggregation induced by the antibodies showed a different pattern with OX57 being the most potent. W3/13

Table 7.1 MAbs specific for a restricted set of thymocyte surface molecules induce aggregation [a)]

mAb	Isotype	^%	Molecule	Aggregation	
				3 hr	24 hr
OX21	IgG1	0	hC3bIna	0	0
W3/13	**IgG1**	**98**	**CD43**	**0**	**4**
OX56	**IgG2b**	**98**	**CD43**	**1**	**4**
OX57	**IgG2b**	**98**	**CD43**	**2**	**5**
OX58	**IgG2b**	**97**	**CD43**	**2**	**5**
OX74	**IgG1**	**98**	**CD43**	**1**	**4**
OX75	**IgG2b**	**97**	**CD43**	**2**	**5**
5H4	**IgG1**	**98**	**CD43**	**2**	**5**
8B8	**IgM**	**98**	**CD43**	**2**	**5**
5G7	**IgG1**	**97**	**CD43**	**1**	**4**
OX1	IgG1	98	CD45	0	0
OX30	IgG2a	97	CD45	0	0
OX2	**IgG1**	**90**	**OX2**	**1**	**3**
OX7	**IgG1**	**97**	**THY-1**	**1**	**3**
OX8	IgG1	86	CD8	0	0
OX10	**IgG1**	**82**	**CD8**	**2**	**3**
OX18	IgG1	35	RT1A	0	0
OX19	IgG1	95	CD5	0	0
OX27	IgG2a	35	MHC I	0	0
OX34	IgG2a	95	CD2	0	0
OX55	IgG1	95	CD2	0	0
OX44	IgG1	15	CD53	0	0
OX45	IgG1	95	CD48	0	0
OX47	IgG1	95	OX47	0	0
OX49	IgG2a	80	CD44	0	0
OX50	IgG1	80	CD44	0	0
OX52	IgG2a	80	?	0	0
R73	IgG1	50	TCR	0	0
HIS41	IgG1	95	CD45	0	0
W3/15	**IgG1**	**95**	**?**	**1**	**3**
W3/25 [b)]	IgG1	90	CD4	0	0
WT.1	IgG1	98	CD11a	0	0
IA29	IgG1	0	ICAM-1	0	0

a) 10µl of mAb tissue culture supernatant was used per well (final 1/20) and thus sodium azide was present at a final concentration of 0.001%. ^% indicates the approximate % of rat thymocytes stained by the antibody on FACScan analysis. Aggregation was scored as described in Chapter 2 on a scale of 1 (1-5% cells aggregated) to 5 (>90% cells aggregated). ? Indicates molecular structures not yet defined. b) In addition to W3/25, 9 other mAb were tested (kindly provided by A.N. Barclay, unpublished mAb) specific for different epitopes on CD4, and all were found to be inactive in aggregation.

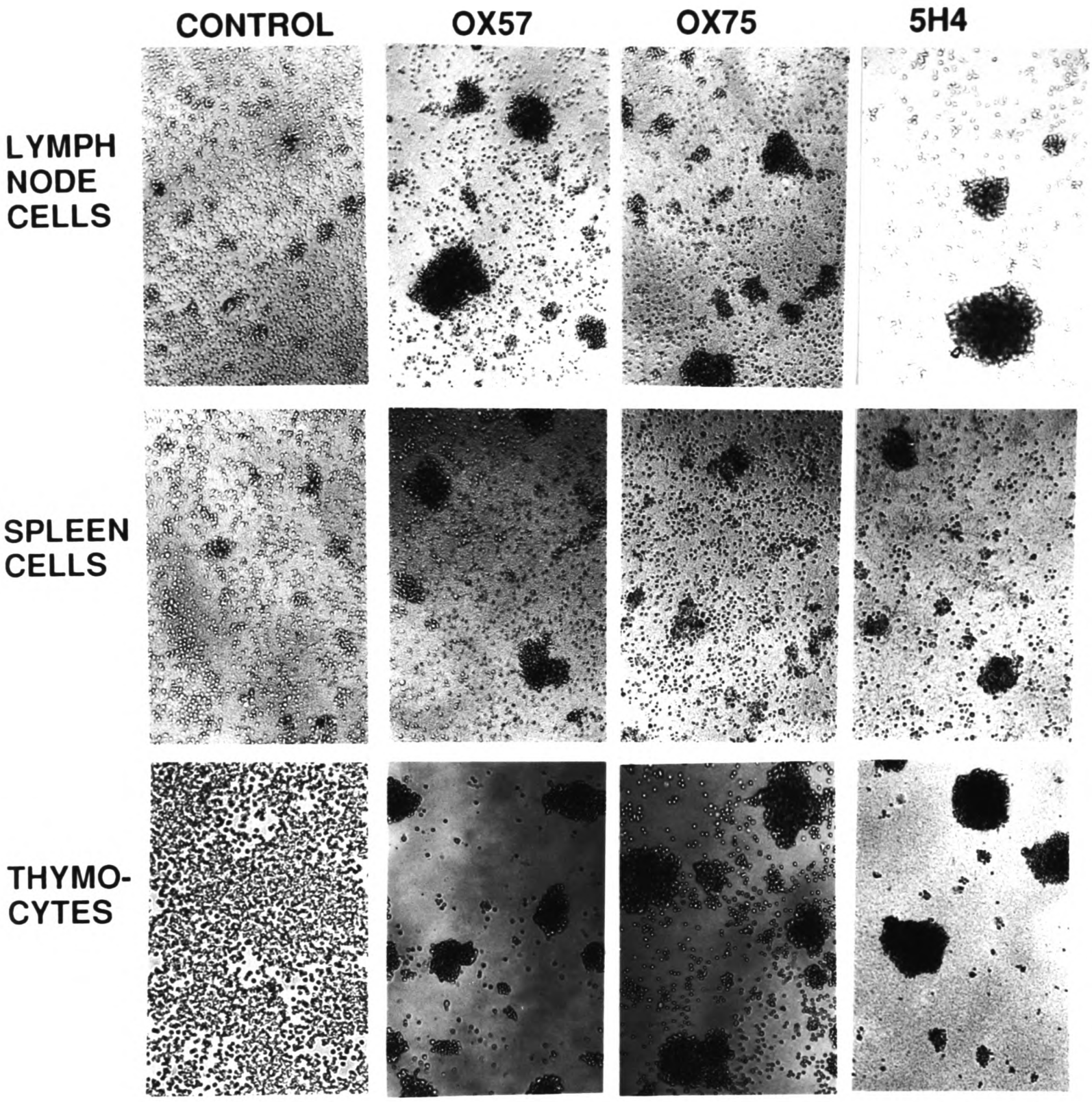


Figure 7.1 Induction of lymphocyte aggregation by anti-rat leukosialin antibodies. The assays were performed at 37°C for 24 hrs with mAb at 1μg/ml (OX55, OX57, OX75) or tissue culture supernatant at 1/20 (5H4). Lymph node and spleen cell control wells were without mAb, thymocyte control wells contained OX55. Magnification, 44 times.

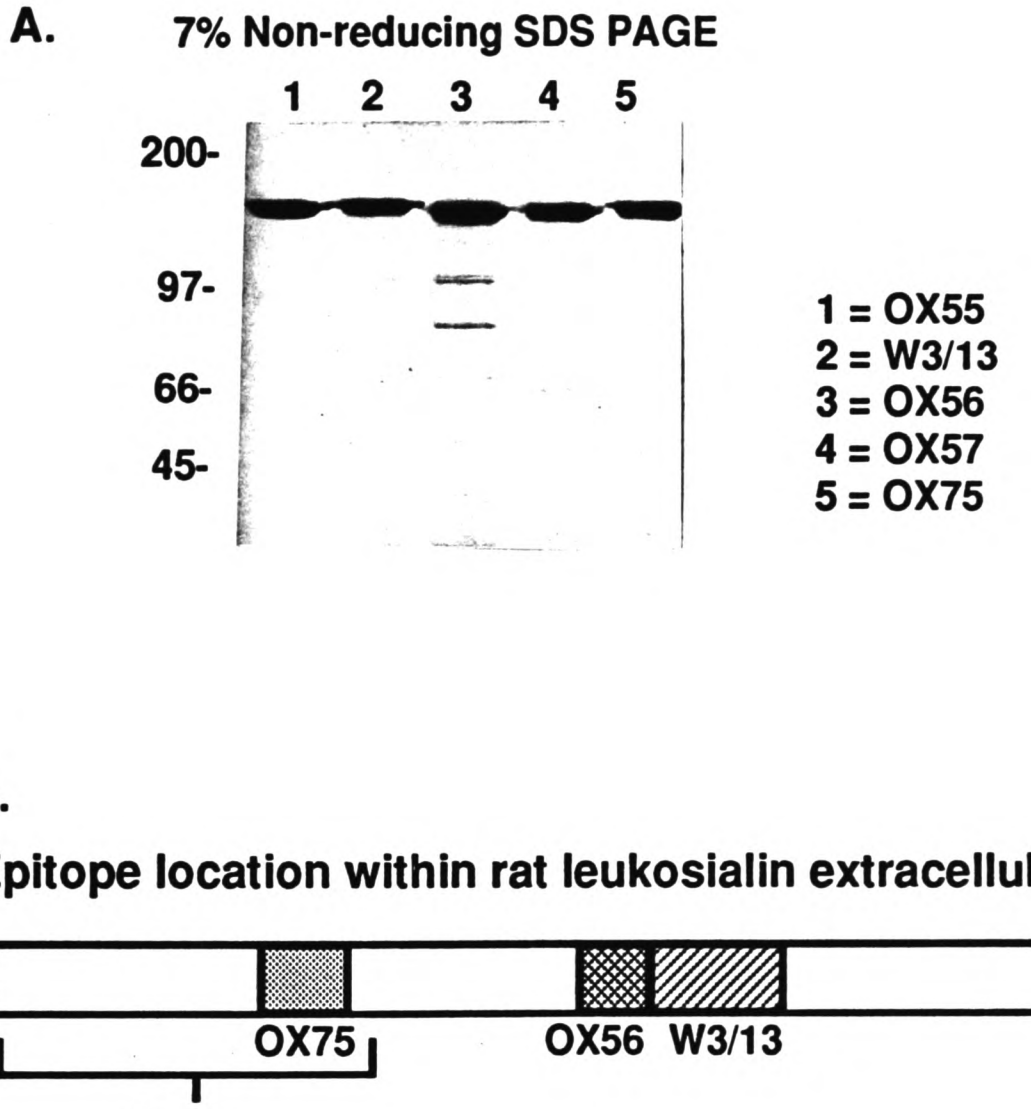


Figure 7.2 SDS PAGE analysis of purified anti-leukosialin IgG (A) and location of the W3/13, OX56, OX57 and OX75 epitopes within the rat leukosialin extracellular domain (B). A) 2μg of each antibody preparation was boiled in 15μl of non-reducing sample buffer and run on a 7% gel. The gel was stained with coomassie blue. The position of molecular weight standards are shown on the left (kD). B) The W3/13, OX56 and OX75 epitopes were mapped by deletion and the approximate location of the OX57 epitope by electron microscopy (Chapter 5). The location of the epitopes is shown roughly to scale assuming each amino acid contributes equally to the extension of the molecule. TM= transmembrane domain.

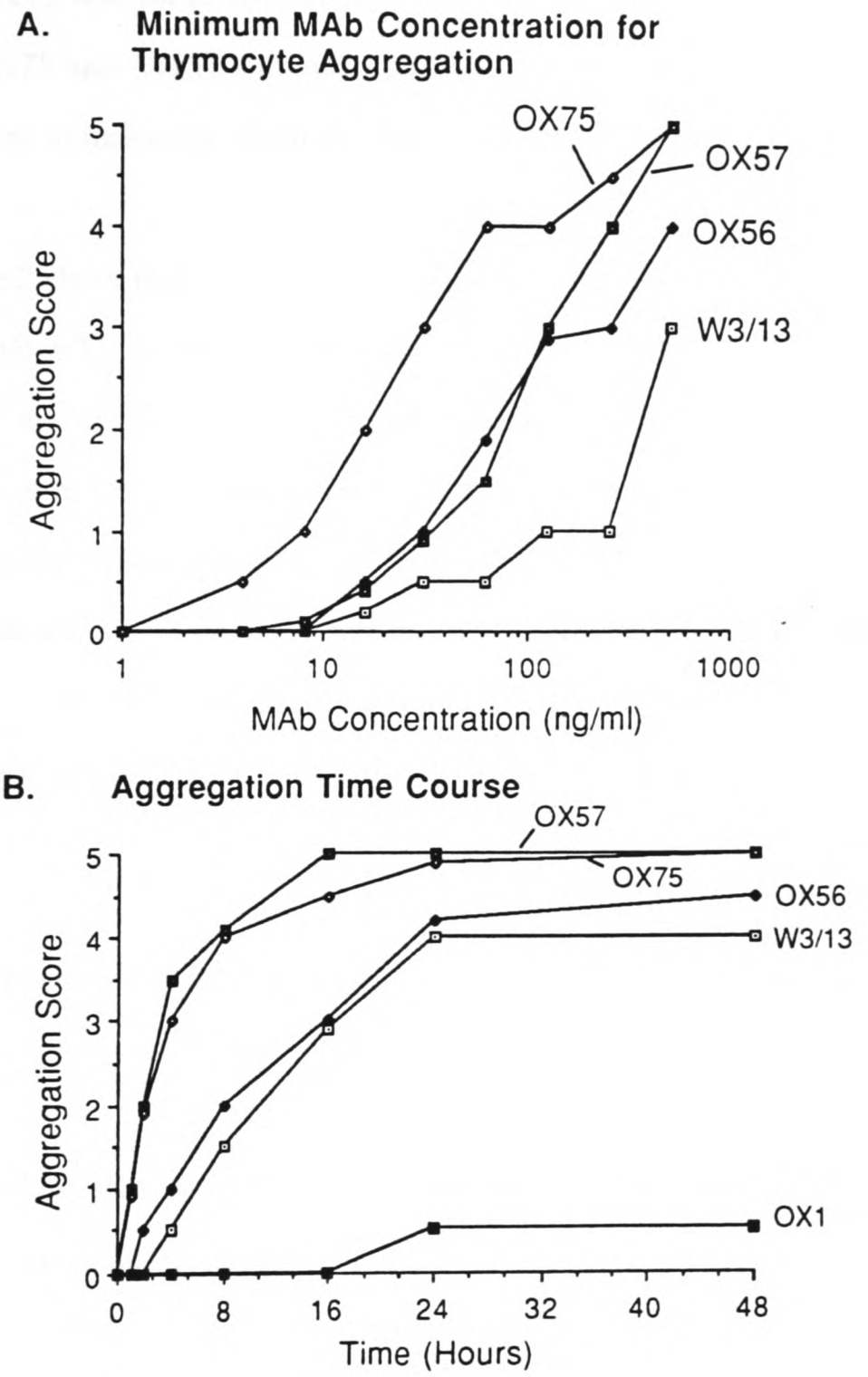


Figure 7.3 Minimum concentration of anti-leukosialin mAb required to induce thymocyte aggregation (A) and time course of aggregation induced by anti-leukosialin antibodies (B). A) Two fold serial dilutions of the mAb were made in PBS, 0.05% BSA starting at 10μg/ml. 10μl aliquots of the dilutions were added to wells in duplicate to which 200μl of thymocytes were then added. Aggregation was scored after 24 hr. B) MAb were used at a final concentration of 1μg/ml and aggregation was scored at successive time points up to 48 hrs. In some experiments a background level of aggregation (with or without control mAb) was seen. The scores shown are the average of 4 independent assays.

and OX56 were generally unable to induce more than an aggregation score of 4 within 24 hours, whereas OX75 was more similar to OX57. Following the time course of aggregation it was found that OX75 and OX57 showed a similar profile whereas W3/13 and OX56 were considerably slower at inducing maximal aggregation (Fig 7.3b).

Figure 4 and table 2 show that the antibody induced clustering was reduced in rate and extent by depletion of cellular ATP by incubation with sodium azide and 2-deoxyglucose, by reduced temperature (4°C) and by the cytoskeletal inhibitor, cytochalasin B. In Fig. 4 the inhibitors were added at the same time as the mAb. Incubation of the cells with the inhibitors for 1 hour prior to addition of the mAb reduced the extent of aggregation further but did not change the overall pattern (Table 7.2). Where the aggregation was not totally blocked, the inhibitors were found to alter the nature of the aggregates, preventing the formation of large tight clusters and allowing more typically formation of a loose matrix (Fig 7.4).

The inhibition due to metabolic inhibitors was not due to effects on protein or mRNA synthesis since neither cycloheximide (5-50 μg/ml) nor actinomycin D (0.1-1.0 μg/ml) had any effect on aggregate formation induced by the anti-leukosialin antibodies over a 48 hour period.

7.4 Importance of divalent cations and lack of effect of mAbs against LFA-1 and ICAM-1 in anti-leukosialin induced aggregation

Depletion of Ca^{2+} and Mg^{2+} from the medium using EGTA and EDTA gave an inhibition of aggregation that was similar to that seen with the metabolic inhibitors (Fig 7.4, Table 7.2). The requirement for cations was also assessed by plating cells in PBS or PBS supplemented with $CaCl_2$ or $MgCl_2$ (Table 7.2). In PBS alone no significant aggregation was observed with W3/13 or OX56 whereas a limited amount of cell association was seen with OX57 and OX75. The extent of aggregation was very similar to that in RPMI in the presence of EDTA or EGTA (Fig 7.4). Addition of $MgCl_2$ (0.1 to 1.0mM) enhanced aggregation but did not support the same level as RPMI. Addition of $CaCl_2$ (0.1 to 0.4mM) did not increase clustering significantly. At concentrations of 1mM or higher the thymocytes did not settle to the base of the well, possibly due to formation of a $Ca(PO_4)_2$ precipitate. A combination of $MgCl_2$ and

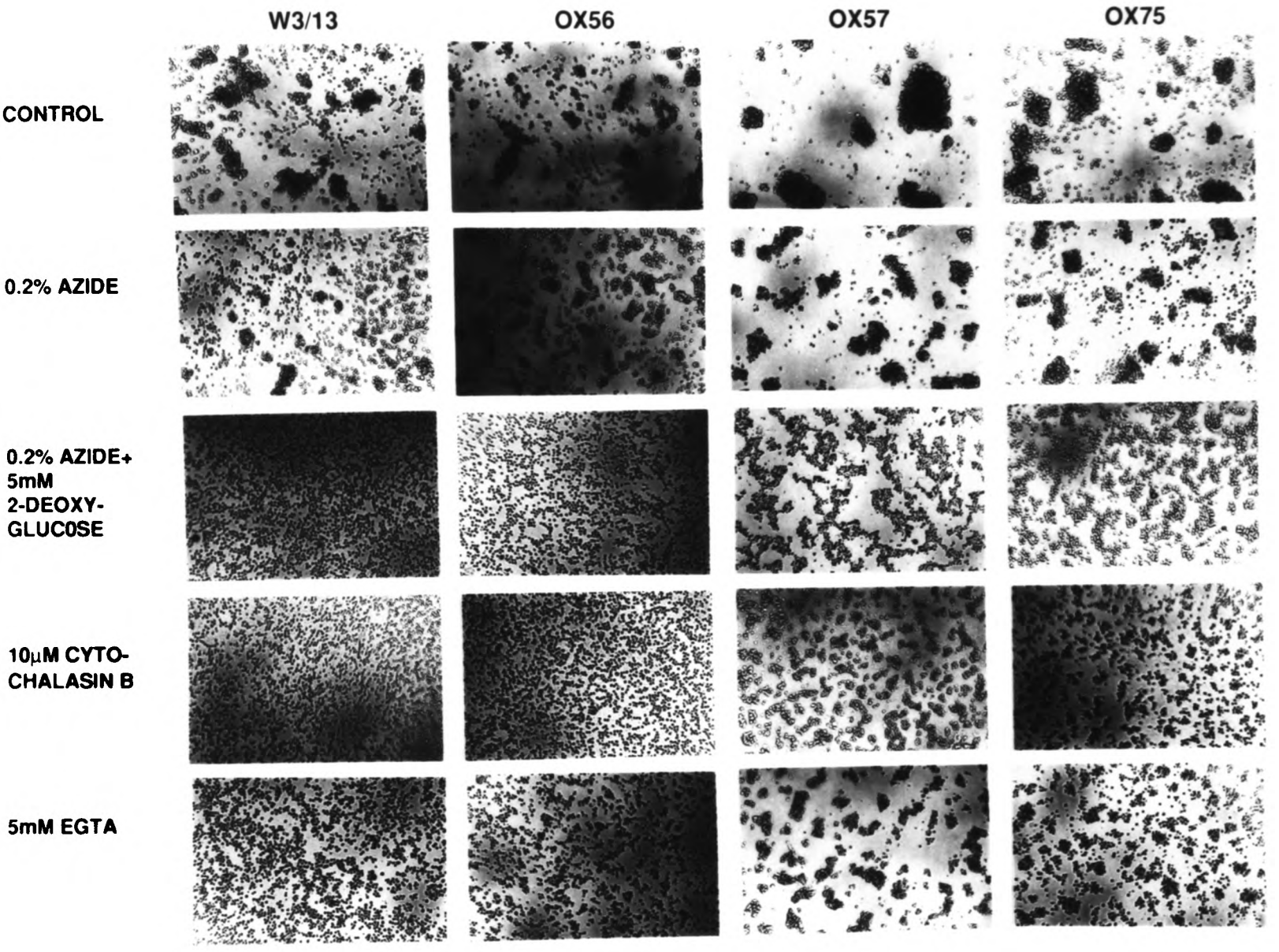


Figure 7.4 Effect of metabolic and cytoskeletal inhibitors on anti-leukosialin mAb induced thymocyte aggregation and requirement for divalent cations. Thymocytes (200μl) were plated at $3x10^6$/ml, the indicated mAb was added to 1μg/ml without (Control) or with inhibitors as indicated and the cells incubated at 37°C for 24 hr. Magnification, 56 times. The level of aggregation with the negative control mAb OX55 was identical to Fig 7.1. Results with 5mM EDTA were identical to EGTA. Incubation at 4°C lead to a pattern similar to that with Cytochalasin B.

Table 7.2. Effect of metabolic and cytoskeletal inhibitors on aggregation induced by anti-leukosialin mAb and requirement for divalent cations [a)]

Conditions	W3/13		OX56		OX57		OX75	
	4 hr	24 hr	4 hr	24 hr	4 hr	24 hr	4 hr	24hr
Control	1	4	1	4	3	5	3	5
4°C	0	1/2	0	1	1	3	1	3
Azide 0.2%	0	1	0	1	1	3	1	2
Azide+2DG [b)]	0	0	0	0	0	2	0	1
Cyto B 10μM	0	0	0	1/2	0	2	0	1
EDTA 5mM	0	1/2	0	1	2	3	2	3
EGTA 5mM	0	1/2	0	1	2	3	2	3
PBS	0	0	0	1/2	1	2	1	2
+$MgCl_2$	0	0	0	1/2	2	3	2	3
+$CaCl_2$	0	0	0	1/2	1	2	1	2
+$MgCl_2$+$CaCl_2$	0	1/2	0	1	2	3	2	3

a) Thymocytes were incubated with the inhibitors for 1 hr prior to addition of mAb to 3 μg/ml. b) Sodium azide was used at 0.2% with 2DG, 2-deoxyglucose at 5mM. Cyto B, cytochalasin B. $MgCl_2$ and $CaCl_2$ were at 0.4mM alone or together.

$CaCl_2$ showed a slight enhancement over $MgCl_2$ alone although not to the same level as in RPMI.

Anti-leukosialin induced aggregation was not inhibited by antibodies to LFA-1α (WT1) or ICAM1 (IA29), at a dilution of tissue culture supernatant of 1/10 whereas aggregation induced by PMA at 50ng/ml was totally inhibited by WT1 but unaffected by IA29, as previously reported (Tamatani et al., 1991). In addition to WT1 and IA29, all the antibodies listed in Table 7.1 were tested at 1/20 of tissue culture supernatant for their ability to inhibit OX75 induced aggregation. In no case was the level of aggregation reduced.

7.5 Effect of epitope position on aggregation induced by anti-leukosialin mAbs

The inhibition of aggregation obtained with the procedures above were greater with the anti-leukosialin mAbs which bound nearer the transmembrane region (W3/13 and OX56) than with mAbs that bound at a more distal site (OX57 and OX75; Fig 7.2). This correlation was also seen with the kinetics and magnitude of aggregation (Fig 7.3). Furthermore, using tissue culture supernatants at 1/20 dilution the aggregation induced by mAbs OX58, 5H4, and 8B8 showed similar kinetics to that by OX57 and OX75 whereas OX74 induced aggregation that was more similar to OX56 and W3/13 induced clustering (Table 7.1 and data not shown). This pattern also fits the epitope map (Chapter 5).

7.6 Fab fragments are inactive in thymocyte aggregation

Fab fragments were generated from the W3/13, OX56, OX57 and OX75 mAb. This was either by direct papain digestion of IgG to create Fab (with OX56, OX57 and OX75) or by pepsin digestion to generate $F(ab')_2$ followed by reduction and alkylation (W3/13). The Fab fragments were all purified by ion exchange chromatography followed by size exclusion chromatography to separate them from contaminating IgG and $F(ab')_2$. It can be seen from the non-reducing SDS PAGE analysis of the pooled Fab preparations that the samples are contaminated with small amounts of disulfide-linked $F(ab')_2$ (Fig 7.5a). The level of contamination was estimated from coomassie stained gels to be less than 2% in all cases. The Fab preparations all showed a similar level of binding to thymocytes on FACScan analysis (Fig 7.5b). The broad staining profile with OX75 Fab is also seen with the intact IgG (Chapter 5, Fig 5.2) and may be due to

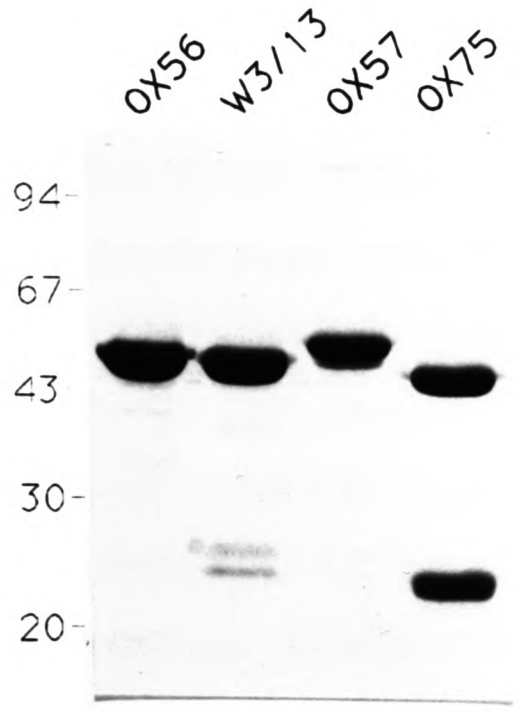
the dependency of the epitope on sialic acid. When these preparations were tested for clustering activity, the W3/13, OX56 and OX57 Fabs were totally inactive, even at 10 ug/ml, a concentration more than 50 times that necessary for aggregation with the intact IgGs. The OX75 Fab preparation was active in aggregating thymocytes, however, at only 5 times the concentration required with the IgG (50 ng/ml vs 10 ng/ml). To test that this effect was due to the monomeric interaction of the Fab with leukosialin, the OX75 Fab preparation was further purified by passage over a sephadex G75 gel filtration column. The material eluted as two peaks and was analyzed by SDS PAGE without reduction (Fig 7.5c). The first peak consisted of a small amount of $F(ab')_2$ but also material that had the properties of Fab on SDS PAGE. This material presumably consists of Fab dimers that are non-covalently associated. The second peak had the properties of unaggregated Fab since its elution from the G75 column was as expected for Fab. Material from fractions 37 and 47 was dialysed against PBS and tested immediately and without concentration for its ability to promote thymocyte clustering. It was found that fraction 37 (dimeric) was effective at approximately 40 ng/ml whereas fraction 47 (monomeric) was ineffective even at 5000 ng/ml (Fig 7.5d). Aggregation could be induced by subsequent addition of rabbit anti-mouse Ig antiserum confirming that the Fab had bound to the cells. The FACScan profiles using Fab from both peaks were equivalent, being similar to that for the original preparation (Fig 7.5b). Thus in no case was a monomeric Fab preparation able to induce aggregation.

7.7 The Fc portion of anti-leukosialin mAb is not required for aggregation

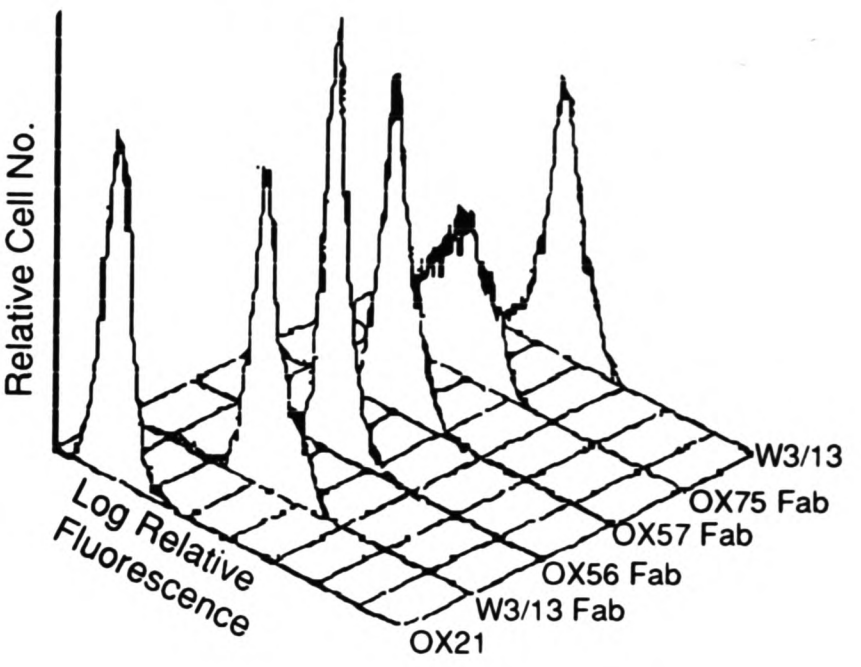
The activity of homo-dimeric OX75 Fab in thymocyte aggregation (Fig 7.5) excludes any involvement of the Fc region in aggregation by this mAb. After storage at 4°C for 3 months (at 0.1 mg/ml) the OX57 Fab preparation also demonstrated clustering activity (a score of 4 after 16 hrs at 37°C). This reveals again the propensity of Fab fragments to self-associate and it also demonstrates that OX57 is active in thymocyte aggregation in the absence of the Fc region.

Figure 7.5 Isolation of monomeric OX75 Fab and lack of thymocyte aggregating activity. A) 7% non-reducing SDS PAGE analysis of indicated Fabs (7μg/track). The lower M_r band in OX75 is a mixture of Fd fragment of the heavy chain and light chain and is presumably derived from Fab, and termed Fab* in section C, that lacks the inter-chain disulphide. B) Binding of Fab fragments to rat thymocytes analyzed by flow cytometry. Cells were incubated with saturating amounts of the mAb or Fab shown followed by a second incubation with fluorescein-labelled $F(ab')_2$ anti-mouse IgG antibody, and bound mAb was measured by flow cytometry. C) 12% non-reducing SDS PAGE analysis of OX75 Fab fractions collected from a G75 column. 10 mg of the OX75 Fab preparation (lane L) was loaded onto a Sephadex G75 column at a flow rate of 10 ml/hr in 25 mM tris-HCl (pH 7.4), 140 mM NaCl. 2ml fractions were collected and 10μl samples analyzed. D) Monomeric OX75 Fab does not induce aggregation. Thymocytes were incubated with OX75 IgG, Fab material from fraction 37 or Fab from fraction 47, at 1μg/ml for 16 hr at 37°C. Magnification, 80 times.

A. Fab Preparations

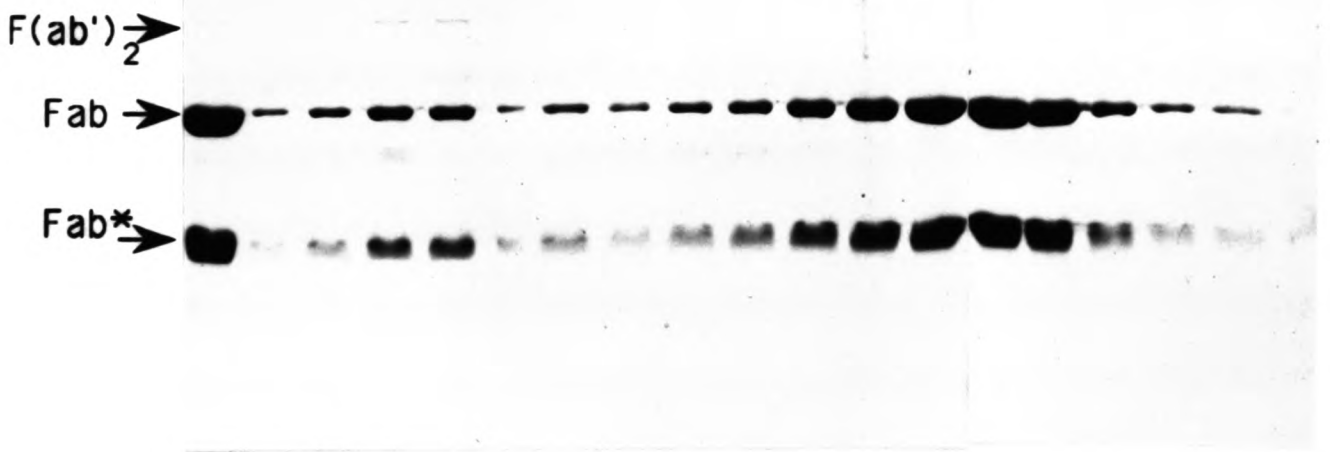


B. Binding of Fab Fragments to Rat Thymocytes



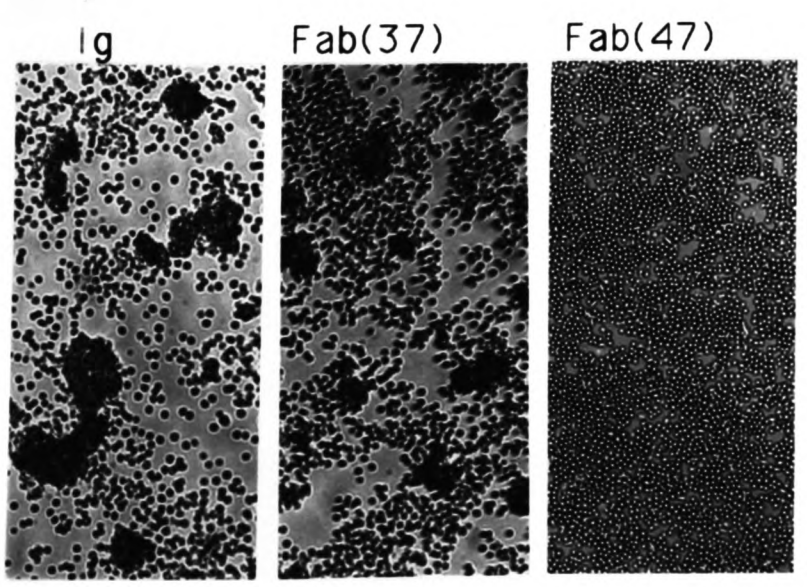
C.

Isolation of Monomeric OX75 Fab



Fab*- non-disulphide linked Fab

D. Monomeric OX75 Fab Does Not Induce Aggregation



7.9 Discussion

Antibodies to a variety of surface molecules have been found to induce homotypic aggregation of leukocytes. To understand the significance of this activity it is important to distinguish between a number of potential mechanisms as follows. (1) The clustering could be due to direct cross-linking of molecules between cells that is adequate to maintain cell adhesion without other interactions. (2) The antibody provides a cross-linking effect between cells that overcomes repulsion and this then allows other adhesion molecules to come into play to achieve pronounced clustering. (3) The antibody induces or stabilizes a conformational state in an adhesion molecule that potentiates adhesion and cell clustering. (4) The antibody cross-links a species of cell surface molecule that can mediate signal transduction and activate the cell to a state in which adhesion reactions are favoured by changes in the disposition or conformation of a number of cell surface molecules.

The features of the clustering via leukosialin reported in this Chapter are that: monovalent Fab is nonfunctional; divalent cations are required for the full effect; the potency of antibodies is greater in rough proportion to the distance of the epitopes they recognize from the cell surface; and clustering is reduced by metabolic inhibitors and cytochalasin B. The relationship to epitope position can be argued to favour hypotheses (1) or (2) and bivalent antibody is also essential for these mechanisms. Inhibition by metabolic inhibitors and cytochalasin B could be relevent to either case since membrane mobility may be crucial for effective clustering. The requirement for cations appears to favour hypothesis (2) since the integrins require divalent cations and these are potential molecules in secondary adhesion events. Inhibition was not seen by the anti-CD11a mAb WT1 but the epitopes involved in the clustering of resting thymocytes may differ from those mediating the clustering of PMA activated T cells that is inhibited by the anti-CD11a antibody and other integrin molecules may be involved. Aggregation of human T cells by the anti-leukosialin mAb B1B6 was inhibited by anti-CD18 but not anti-CD11a (Axelsson et al., 1988). Hypothesis (3) seems unlikely since Fab was ineffective and it is difficult to imagine how numerous epitopes on a molecule with an extended structure could all mediate conformational changes that lead to adhesion. This type of mAb mediated adhesion has

been demonstrated with certain mAb specific for integrin molecules (Keizer et al., 1988; Dransfield and Hogg, 1989; Gulino et al., 1990). The possibility that mAb binding is masking sialic acid and so reducing surface charge to an extent that promotes aggregation (King and Katz, 1990) is also unlikely given the lack of activity of the Fab fragments. Hypothesis (4) is compatible with most of the data except that it is not obvious why epitopes that are more distant from the cell surface should favour signal transduction events in isolated cells. Furthermore, the mAb to leukosialin used in this study are not able to activate cells in isolation although some of them do enhance proliferation stimulated by PMA (Beyers et al., 1989).

It could be argued that differences in antibody affinity, rather than epitope location, account for the differences in rate and extent of aggregation seen in this study. This is unlikely, however, because all the mAbs were used in excess in experiments testing rate of aggregation and senstivity to inhibition, and the relationship to epitope position was seen with all 8 mAbs whose epitope location has been mapped. That is, the 5 mAb that bound near the NH_2-terminus induced more rapid and inhibitor resistant aggregation than 3 mAb that bound near the centre of the molecule. In addition, equimolar amounts of recombinant extracellular domain were needed to inhibit binding of the OX-56 and OX-75 mAbs (Chapter 5, Fig 5.10), indicating a similar affinity for the protein core despite the different location of their epitopes.

Taking all the data together, hypothesis (2) in which the antibody overcomes repulsion between cells leading to adhesion via the established adhesion molecules, seems most likely. This possibility is also consistent with the fact that cell adhesion reactions involving T cells can be considerably enhanced by removal of sialic acid with neuraminidase (Bentwich et al., 1973; Boog et al., 1989; Kapsenberg et al., 1989). Much of the T lymphocyte sialic acid is carried by leukosialin (Brown et al., 1981) a fact consistent with the view that this molecule may be a factor in natural repulsion between cells.

There have been many studies on cell clustering induced by mAbs and in most cases a key role for mAb cross-linking between cells is not proposed. In some cases Fab fragments have been reported to function but these experiments must be considered with caution since it was not stated that conditions to minimise the possibility of aggregation were followed. Most

commonly it is argued that adhesion is due to cell activation by cross-linking of surface molecules on individual cells leading to an activation state (hypothesis (4)). Inhibition by metabolic and cytoskeletal inhibitors is often proposed as supporting this mechanism but complex events and membrane rearrangement may equally be essential in hypotheses (1) and (2).

Hypothesis (2) can be argued to be in accord with current views on some cell interactions in physiological conditions. For example in the case of neutrophil adherence to endothelial cells it is now established that cells first bind via selectins that presumably react with carbohydrate at the outer surface of the membrane and that this is followed by integrin-mediated adhesion (Lawrence and Springer, 1991). Hypothesis (2) similarly proposes a two stage mechanism with the antibody providing the initial cross-linking contacts between cells.

The inability of mAbs to the majority of surface molecules to induce aggregation by any mechanism, in particular by mechanisms (1) or (2), may be due to a variety of factors including the cell surface density of the protein, its mobility and distribution in the membrane as a whole and in local environments such as microvilli, the structure of the molecule and the location of the epitope with respect to the membrane. Further study of features which distinguish molecules that support antibody mediated aggregation from those that do not should improve our understanding of factors regulating normal adhesion events.

CHAPTER 8

STRUCTURE OF DOMAIN 1 OF RAT CD2 ANTIGEN AND EXPRESSION OF HUMAN CD2 AND LFA-3 DOMAIN 1 IN E.COLI

8.1 Introduction

8.2 Expression of rat CD2 domain 1 in *E.coli*

8.3 Growth in minimal medium for ^{15}N-labelling

8.4 Three-dimensional structure of rat CD2 domain 1

8.5 Expression of two domain form of rat CD2

8.6 Human CD2 domain 1 expression in *E.coli* and lack of conformational stability

8.7 Human LFA-3 domain 1 is expressed as an insoluble fusion protein

8.8 Effect of reduced temperature on fusion protein solubility

8.9 Expression of the isolated LFA-3 domain 1 in *E.coli*

8.10 Expression of CD2.D1 containing Cys residues in *E.coli*

8.11 Discussion

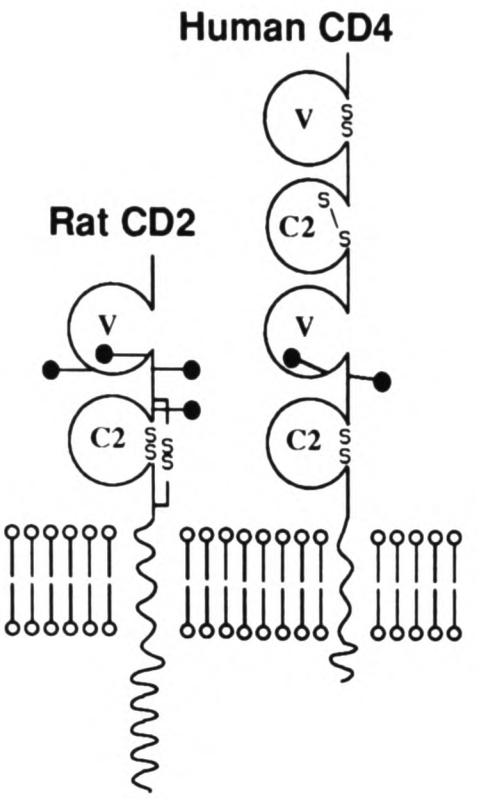
8.1 Introduction

A schematic model proposed for the structure of CD2 is shown in Fig 8.1a along with CD4 whose extracellular domain has been argued to be derived from a CD2-like structure by gene duplication of a two-domain segment (Williams et al., 1987; Williams et al., 1989). The structure of CD2 domain 1 is of interest as an example of a putative IgSF domain without the conserved disulphide bond (Williams et al., 1987) and because of its interaction with LFA-3 (Selvaraj et al., 1987b; Recny et al., 1990). Furthermore, predictions on the basis of primary sequence data lead to a proposed α/β type structure containing 22% α-helix and circular dichroism spectroscopy on the first domain was said to support these predictions (Clayton et al., 1987; Recny et al., 1990). Thus CD2 domain 1 is an important test case for the validity of IgSF assignments based on sequence patterns.

8.2 Expression of rat CD2 domain 1 in *E.coli*

Domain 1 of rat CD2 (CD2.D1) is considered to encompass amino acid residues 1-99 and the nucleotide sequence encoding this region was isolated by PCR using the oligonucleotides shown in Table 8.1. This sequence was expressed as a fusion protein with glutathione-S-transferase (GST) in the pGEX-2T vector in *E.coli* strain MC1061. The fusion protein was expressed at 40 mg/l and was soluble in the lysate of *E. coli* cells, allowing purification on a glutathione agarose column (Fig 8.1b, tracks 1,2, 3). Cleavage by thrombin was complete after 60 minutes and the CD2 domain was purified by affinity chromatography using an OX-34 affinity column followed by gel filtration on a Sephadex G75 column (Fig 8.1b, tracks 4, 5 and 6). The capacity of the OX-34 column was limited to 5 mg of CD2.D1, however, and in later purifications the cleaved GST/CD2 preparation was dialysed against tris/saline and passed back over the glutathione agarose column to remove the GST and then applied directly to the gel filtration column. This yielded a product of identical purity to that which had been OX-34 affinity purified. CD2.D1 emerged as a single peak from the gel filtration column at an elution volume of 140 ml compared with 120 ml for cytochrome c, a monomeric protein of similar Mr (Schulz and Schirmer, 1979), indicating that CD2.D1 was behaving as a monomer in solution. The CD2 domain inhibited the binding of two non-competitive monoclonal antibodies against rat CD2 (OX-34 and OX-55) (Beyers et al., 1989) with a molar potency equal to that of the extracellular domain of CD2 expressed in CHO cells (Fiona Gray, pers.com.).

A.



B.

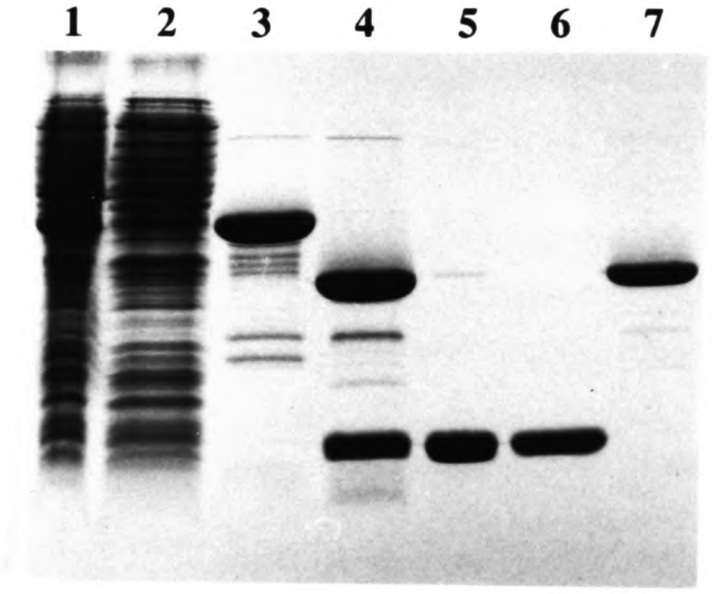


Figure 8.1 Schematic representations for CD2 and CD4 at a cell surface (A) and the expression and purification of rat CD2 domain 1 (B). A) The circles indicate IgSF domains and the symbols, N-linked carbohydrate structures. B) Coomassie stained 16% SDS PAGE gel. Lanes: 1, Soluble lysate from *E.coli* expressing the fusion protein of GST-CD2 domain 1; 2, lysate passed through a glutathione agarose affinity column; 3, purified fusion protein; 4, fusion protein cleaved with thrombin; 5, CD2 domain 1 after purification by OX-34 affinity chromatography; 6, CD2 domain 1 after purification by gel filtration on a Sephadex G75 column; 7, GST.

Table 8.1 Oligonucleotides used in the PCR to isolate IgSF domains for insertion into the pGEX-2T expression vector

No.	Domain		AA	Sequence 5' - 3'
1053	rCD2.D1	5'	+1	tagtggatccagagacagtgggaccgtc
1054	rCD2.D1	3'	+99	ttgaattctactctagaatcctcaagtc
1039	rCD2ex	3'	+177	atgagatagagggatcctcatttctctggaca
1247	hCD2.D1	5'	+1	tagtagggatccaaagagattacgaatgcc
1620	hCD2.D1	5'	+4	tagtagggatccacgaatgccttggaaacc
1248	hCD2.D1	3'	+104	ctagaattcactcttgaatcttcaaatcaa
1278	hLFA-3	5'	+2	tagtagggatcccaacaaatatatggtg
1309	hLFA-3	5'	+1	tagtagggatccttttcccaacaaatat
1470	hLFA-3	5'	RD+2	tagtagggatccagagactcccaacaaatatatggtg
1471	hLFA-3	5'	GSGSG+1	tagtagggatccggtagcggtagcggtttttcccaacaaatatatgg
1279	hLFA-3	3'	+95	gtacccgggctaagactcaagcacataaag
1310	hLFA-3	3'	+91	ctacccgggtcaataaagaaagaacttc
1311	hLFA-3	3'	+88	ctacccgggtcagaacttcatggtatcag
1069	rCD4.D3	5'	+179	tagtagggatccggctttgcgagcacatc
1070	rCD4. D3	5'	+181	tagtagggatccgcgagcacatccatcac
1071	rCD4.D3	3'	+290	ctagaattcatttcatcaccaccaggt
1072	rCD4.D3	3'	+287	ctagaattcacaccaggttcacttcct

All 5' oligonucleotides include a Bam HI site and all 3' oligonucleotides except those used with LFA-3 include a stop codon followed by an Eco RI site; with LFA-3 the stop codon was followed by a Sma I site due to the presence of an internal Eco RI site. AA indicates the amino acid number in the template sequence at which the coding region of the oligonucleotide begins (5') or ends (3') and any additional coding information present in the oligonucleotide.

8.3 Growth in minimal medium for ^{15}N-labelling

To allow growth in minimal (M9) medium, the construct was initially introduced into the prototrophic *E. coli* strain, TG1, but no induction of fusion protein could be demonstrated in these cells upon repeated attempts. The reason for this is unclear. A second prototrophic strain, BL21, was tested and induction was observed. The growth of these cells in minimal medium was compared with that in 2xTY by measuring the absorbance at 550nm (A_{550}). 18-20 hrs were needed for the minimal medium culture to reach an A_{550}=1 which was considered to be stationary phase, the point at which the 1 in 10 dilution could be made for fusion protein induction. Following the dilution 2-2.5 hrs were required to reach the same density of growth as a 1 hr 2x TY culture (A_{550}=0.5-0.6). IPTG was then added and culture continued for a further 3 hr. Upon scaling up to larger cultures it was observed that more IPTG was needed than with MC1061 or other strains used (where maximal induction could be achieved with 10 μM). A comparison of the induction by different IPTG concentrations was made (Fig 8.2) and the yield was found to be maximal at 0.12 mM. Higher concentrations caused a slight reduction in growth rate and protein yields whereas at lower concentrations growth rate was not affected but yield was significantly reduced (Fig 8.2).

It was suggested (Paul Driscoll, personal communication) that the minimum NH_4Cl concentration that could be used without significantly reducing protein yield was 0.52 g/l. To confirm this for fusion protein production in the BL21 strain, NH_4Cl concentrations of 1.0, 0.52 and 0.26 g/l were compared. Following 1 in 10 dilution of a 20 hr culture (which had been supplemented with 0.7 g/l NH_4Cl) the growth rate over the first two hours was similar with each culture reaching A_{550}=0.6 by 2 hr at which point IPTG was added and culture continued for 3 hr. The final absorbances (550nm) of the cultures at 1.0, 0.52 and 0.26 g/l NH_4Cl were 1.0, 1.0 and 0.9 and the A_{280} of the fusion protein preparations (purified as desribed in the legend to Fig. 8.2) were 2.22, 1.83, and 1.12, respectively. 0.52 g/l was chosen as an optimum for minimizing the amount of $^{15}NH_4Cl$ needed without significantly compromising protein yields.

The yield of fusion protein with NH_4Cl at 0.52 g/l reached a maximum after 3 hr induction,

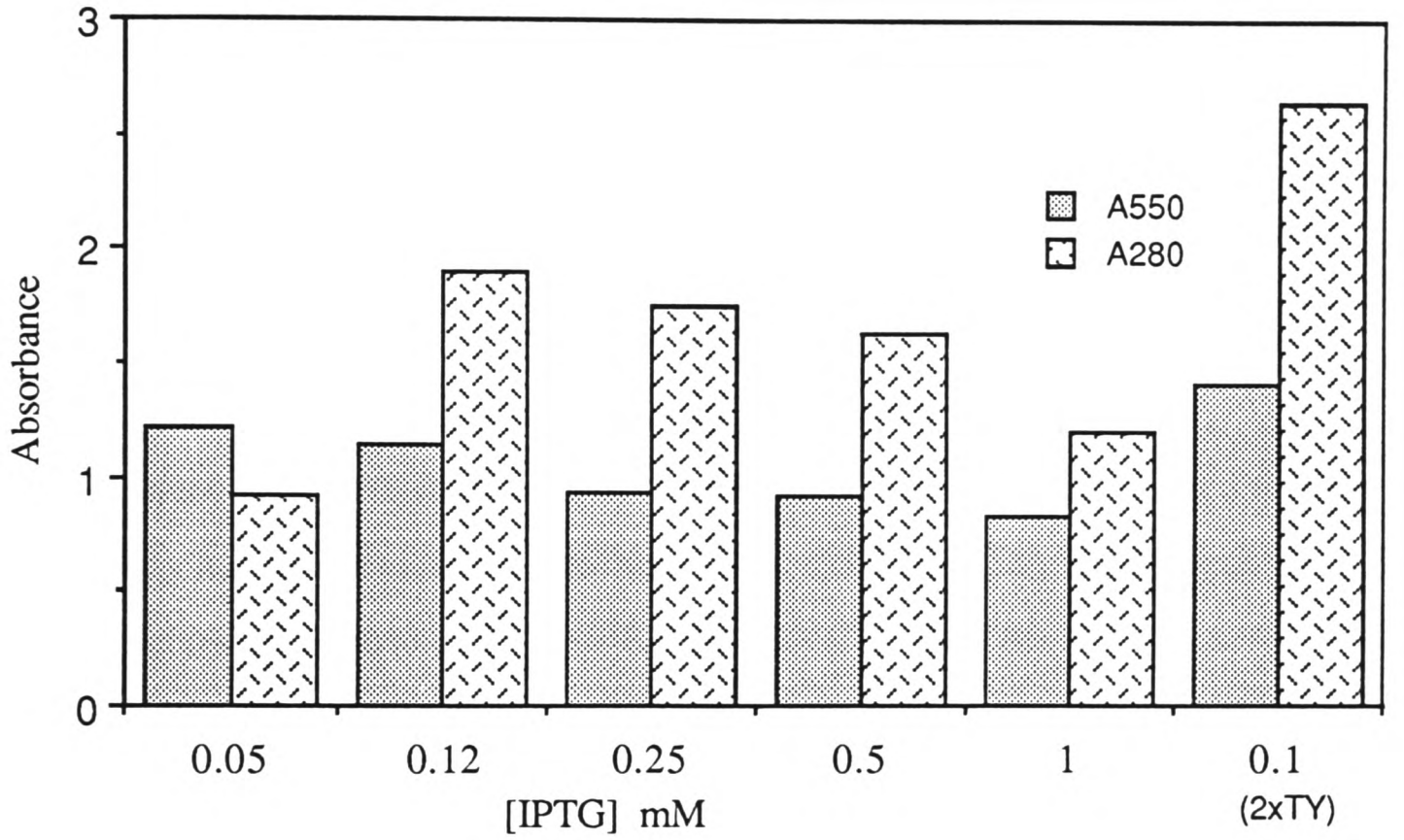


Figure 8.2 Effect of IPTG concentration on fusion protein induction in BL21 cells grown in minimal medium. 5ml of an 18 hr minimal medium culture (A_{550}=1.1) was added to each of 5 250ml-flasks containing 45ml minimal medium. A 6th flask containing 45ml 2xTY was inoculated with 5ml of an o/n 2xTY culture. Flask 6 was incubated for 1 hr and flasks 1-5 for 2 hr before IPTG was added to the indicated concentration. All cultures were incubated for a further 3 hr, the absorbance at 550nm was measured, and the cells were harvested, resuspended in 800µl tris/sal/EDTA and lysed by sonication. The lysate supernatant was mixed with 200µl 50% glutathione agarose at 4°C for 15 min followed by 3 washes with tris/sal and then elution with 400µl of glutathione elution buffer. The A_{280} of the elutions is shown.

decreasing slightly as induction was extended to 4 or 5 hr. The reduced yield with increased time may be due to exhaustion of the NH_4Cl, the culture reaching saturation and the cells undergoing lysis (Studier et al., 1990). In summary, for ^{15}N labelling BL21cells containing the pGEX-CD2.D1 plasmid were grown in minimal medium supplemented with $^{15}NH_4Cl$ (MSD Isotopes, Merck Frost, Canada) at 0.52 g/l (9.7mM) for 18-20 hr until the absorbance at 550nm was >1.0. The cells were then diluted 1:10 and grown until A_{550} was 0.5-0.6 when IPTG was added to 0.12 mM and growth continued for a further 3 hr. The ^{15}N labelled CD2 was purified as described for the unlabelled material and both forms were dialysed against double deionized water and concentrated using centricon apparatus. Occasionally small amounts of preciptiate appeared and these were removed by centrifugation. Approximately 5 mM samples were used for NMR spectroscopy.

8.4 Three-dimensional solution structure of rat CD2 domain 1

The determination of the 3-D solution structures of small proteins by 2-D and 3-D NMR spectroscopy has been reviewed extensively (Wuthrich, 1989; Clore and Gronenborn, 1991). Figure 8.3 shows representative ^1H NMR spectra of CD2.D1 obtained from a 600 MHz spectrometer by Paul Driscoll. Figure 8.3a shows part of the 2-D homonuclear Hartmann-Hahn (HOHAHA) spectrum and Fig 8.3b the corresponding region of the 2-D nuclear Overhauser effect (NOESY) spectrum. The off diagonal cross-peaks of the HOHAHA spectrum indicate proton resonances which are correlated 'through bonds', that is within individual amino acids. The cross-peaks in the NOESY spectrum yield information about protons which are separated by less than 5 Å in space. These peaks are the information source for the 3-D structure determination. The spectra indicate a compact fold with mainly β-sheet secondary structure. The complexity of the spectra is sufficient to warrant the application of ^{15}N-labelling and 3-D heteronuclear NMR spectroscopy (Clore and Gronenborn, 1991). All the NMR spectra were obtained and assignments made by Paul Driscoll (Driscoll et al., 1991).

The CD2 structures calculated clearly show a fold like an Ig V-domain with the core strands having a very similar orientation to those in CD4 domain 1, and the Ig V domains NEW V_H and REI V_L (Driscoll et al., (1991) and Fig 8.4). The distance between the Cα atoms of the residues corresponding to the conserved disulphide bond of IgSF domains (Ile 18 and Val 78)

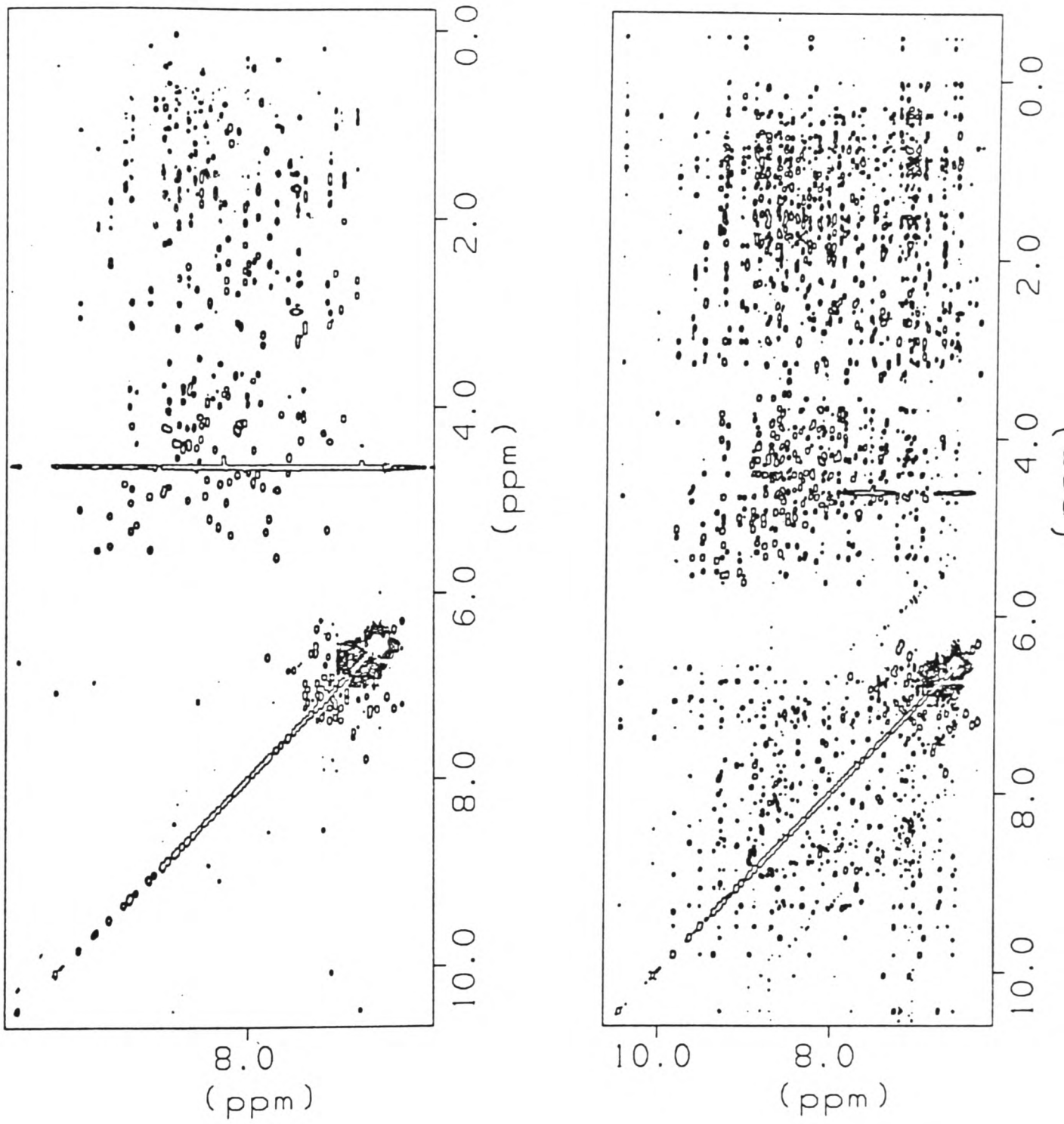


Figure 8.3 2-Dimensional homonuclear Hartmann-Hahn spectrum (A) and the corresponding region of the 2-Dimensional nuclear Overhauser effect spectrum (B) of rat CD2 domain 1.

was within the range seen for the Cα atoms of the Cys residues forming the bond in other determined structures (Driscoll et al., 1991). Thus the structure is commensurate with the formation of a disulphide bond at this location as had been previously predicted (Williams et al., 1989) and demonstrated (Gray et al., 1992, and section 8.10).

When the full CD2 domain is compared with CD4 domain 1 and the V-domains, differences in the regions connecting the β-strands are evident. Fig 8.4 shows comparisons of the folds for the four domains and the sequence comparisons for CD2 and CD4 are shown in Fig 8.5. After the Ile-18 residue of β-strand B in CD2, there is a Pro residue and the CD2 chain folds immediately into a long, kinked loop connecting to β-strand C. In the V domains and CD4, β-strand B extends further and the loop joining strands B and C is not distorted. In CD2, β-strands D and E are truncated in comparison with the same regions in the other domains. This appears to be required to accomodate the kink in the loop between strands B and C since there are many internal side chain contacts between the loops connecting strands B to C and D to E (Paul Driscoll, personal communication). These contacts plus the unusual structure for the B to C loop may explain the fact that sequence in this region is highly conserved between human and rat CD2. In contrast, in the same region of CD4 there is little conservation of sequence between the species (Fig 8.5). In CD2 the loop between β-strand F and G is longer than in CD4 and is more similar to the same region in the V-domains where this loop constitutes hypervariable region 3.

The structure for CD4 (Wang et al., 1990; Ryu et al., 1990) is unusual in comparison with Ig chains in that domain 2 follows on directly from domain 1 without any hinge residues such as are found between Ig V domains and the following C_L or C_{H1} domains. From Fig 8.5 it is probable that the juxtaposition of domains in CD2 will be like that in CD4 since the distance of the β-strand G of domain 1 from β-strand B of domain 2 is in fact one amino acid less in the CD2 sequence than in CD4.

β-strand assignments for human CD2 domain 1 can be made by alignment with the rat CD2 domain 1 structure (Fig 8.5). All the LFA-3 binding activity has been shown to reside in domain 1 of CD2 (Recny et al., 1990) and mutagenesis of human CD2 has indicated regions of

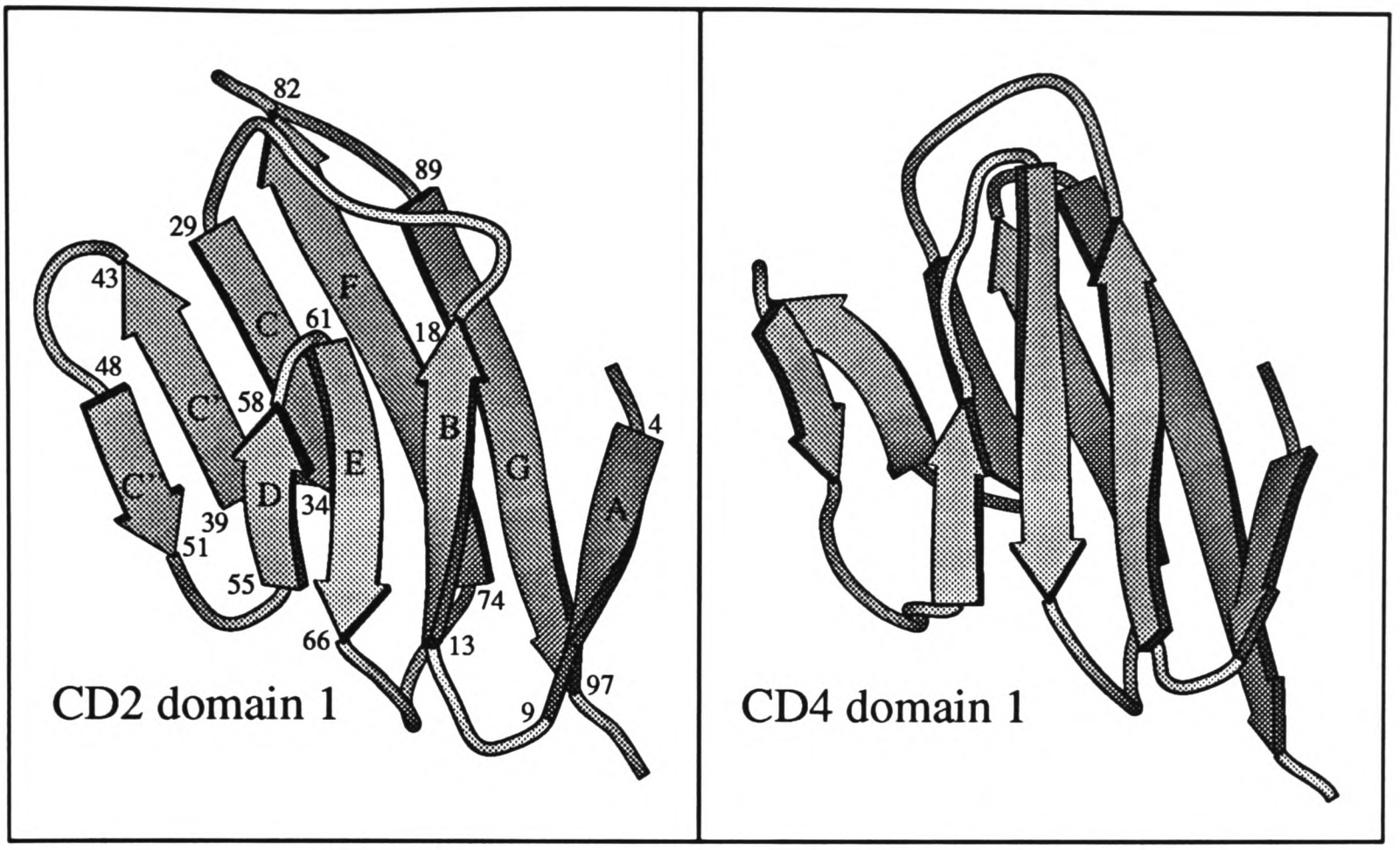


Figure 8.4 Illustrations of the folds for CD2 domain 1, CD4 domain 1, Ig NEW V_H and Ig REI V_L. The coordinates for CD4 domain 1 were kindly provided by Stephen Harrison and those for the V domains were from the Brookhaven database. The diagrams were computed by Paul Driscoll using the MOLSCRIPT program (Driscoll et al., 1991).

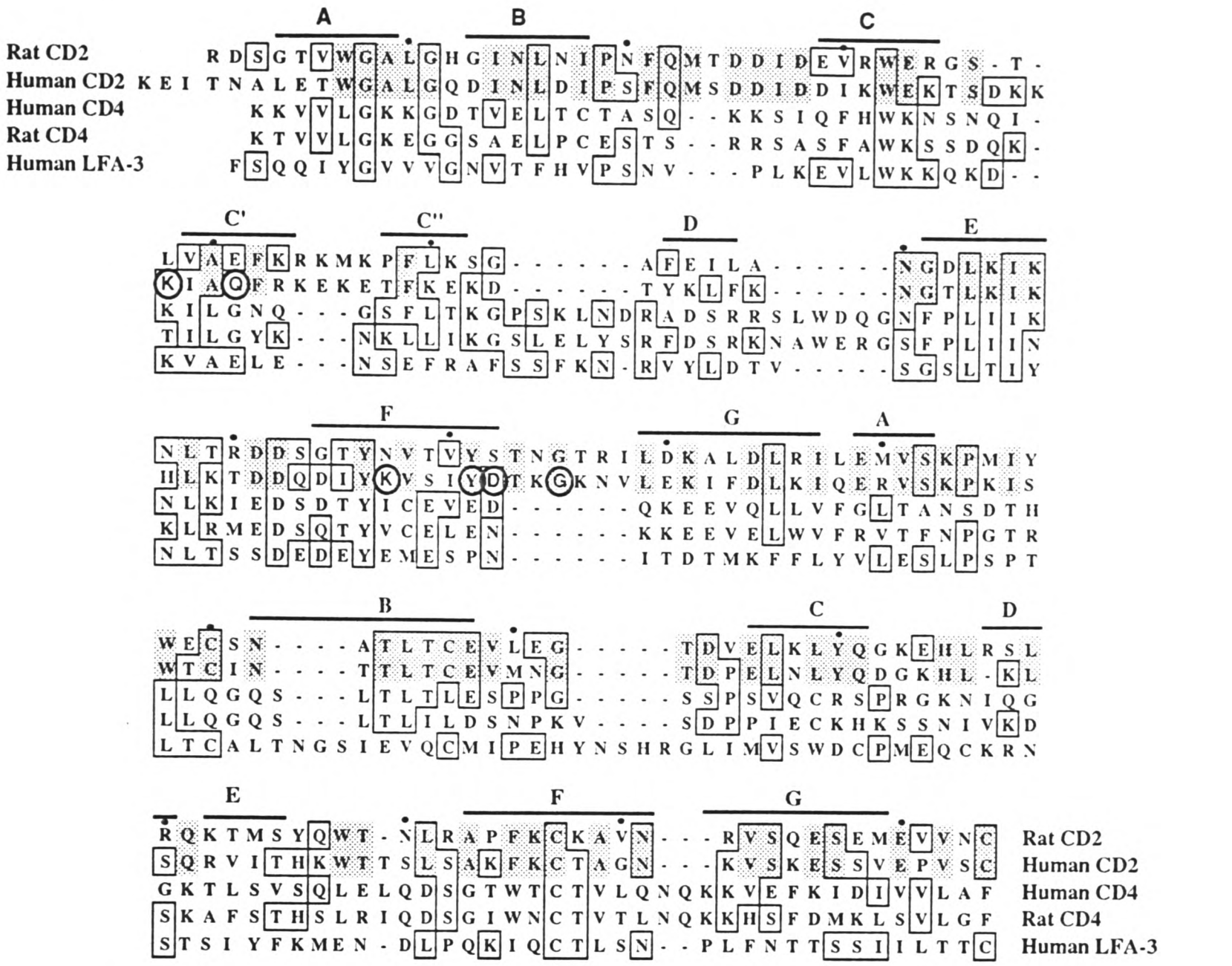


Figure 8.5 Sequence alignments of CD2, CD4 and LFA-3 domains 1 and 2. Human and rat sequences of CD2 and CD4 and the human sequence of LFA-3 are shown derived from the Swissprot database. The black dots above the rat CD2 sequence indicate each tenth residue of that sequence. The β-strand assigments have been made on the basis of the rat CD2 structure for the first domains and on the basis of the CD4 structure for the second domains. In the first domain sequences, the positions of CD4 β-strands are also correctly identified by the bars but their extent is not exactly shown in all cases. Residues which are identical between CD2 and CD4, CD2 and LFA-3, CD4 and LFA-3 or all 3 sequences are boxed. Identities between rat and human CD2 are coloured grey. Residues in human CD2 which are candidates for involvement in LFA-3 binding are circled and include human residues 43,46,82,86,87,&90. These data are from mutational analysis (Peterson and Seed, 1987; note residue 43 is equivalent to 48 in this reference where the amino terminus was assigned 5 residues before the correct position) and the criteria for considering a residue as a candidate for binding is that there should be a single mutation, not involving a Pro residue, that leads to the loss of LFA-3 binding and that the equivalent residue in the rat structure (38,41,77,81,82,&85) is accessible to solvent.. All of these residues except rat Ser 82 are exposed on the same β-sheet face.

contact with LFA-3 (Peterson and Seed, 1987). Residues which appear to be involved in this interaction are marked in Fig 8.5. These are positioned in β-strand C' and in the region of ß-strands F to G. This suggests that the face of the β-sheet C,C',C'',F,G may be involved in the CD2 : LFA-3 interaction.

8.5 Expression of two domain form of rat CD2

A two domain form of rat CD2 (CD2ex) encompassing amino acids 1-177 was expressed using the pGEX system but in contrast to the first domain alone, this form was largely insoluble (Fig 8.6). The antigenicity of the soluble material was tested in a direct binding assay (Table 8.2) and mAbs OX-34, OX-53 and OX-55 were found to react with both the single domain and two domain forms whereas OX-54 did not react with either form but did show reactivity with the CHO cell expressed two domain form. This indicates that either the soluble fraction of the two domain form expressed in *E.coli* is not correctly folded or that carbohydrate is an essential part of the OX-54 epitope.

8.6 Human CD2 domain 1 expression in *E.coli* and lack of conformational stability

The sequence corresponding to the residues from the mature NH_2-terminus to amino acid 104 of human CD2 (equivalent to amino acid 99 in rat CD2, Fig 8.5), was isolated by PCR (Table 8.1) and inserted into the pGEX-2T vector. The fusion protein was efficiently expressed (Fig 8.6) although at levels only approximately 50% of those seen with rat CD2.D1. The hCD2.D1 fusion protein was soluble and was readily cleaved by thrombin. However, when the hCD2.D1 as either the fusion protein or the isolated domain was concentrated in a centricon apparatus, significant amounts of precipitate accumulated. 150 mg of fusion protein could be isolated from 9 litres of culture which corresponded to 50 mg of hCD2.D1. After passing the cleaved material over the glutathione agarose column to remove the GST and then concentrating the domain for gel filtration, only 5 mg of domain remained in solution. It was also found that when the soluble material was passed over a Sephadex G75 gel filtration column, the domain did not emerge as one sharp peak, but as an early broad peak followed by a peak at the expected position for the monomeric hCD2.D1 material. The ratio of total protein in each peak varied with the preparation and Fig 8.7 shows an example where the majority of material emerged in the high M_r peak (this preparation had only been partially purified of GST). After 3 runs of this

type, pooling the peak corresponding to the monomeric material and concentrating (which saw further precipitation) 0.5 ml of a 0.5 mM solution was obtained and this was analyzed by NMR spectroscopy by Paul Driscoll and Arthur Crawford. The signal obtained indicated that a large fraction of the material was unfolded or at least not in a compact, stable fold. Further hCD2.D1 was isolated by Arthur Crawford, with similar results. It was also observed that significant quantities of the hCD2.D1 fusion protein were not being eluted from the glutathione-agarose column by the glutathione elution buffer. Elution was achieved using 0.5 M propionic acid (pH 2.5) followed by immediate titration to pH 7 with solid Tris-HCl. The material isolated in this way showed similar properties to the fusion protein isolated with glutathione elution.

The reactivity of the recombinant CD2 with 10 anti-CD2 mAb was tested and all reacted with the fusion protein in a direct binding assay (Table 8.3) except OCH-217 which had previously been shown to be specific for domain 2 (Peterson and Seed, 1987). The antigenicity of the isolated domain was also assessed by inhibition of mAb binding to the immobilized fusion protein. Figure 8.8 shows that the isolated domain was able to inhibit binding of the 7 mAb specific for domain 1 that were tested. The inhibition curve for mAb 453 (95-5-49) was unusual, however, and this mAb also gave the lowest counts in the direct binding assay indicating that the epitope seen by this mAb may not be totally in the correct conformation. The amount of hCD2.D1 required to block the other mAb was an order of magnitude higher than the amount of rat CD2.D1 needed to block binding of OX-34 or OX-55 (see Fig 8.12, for example). Thus as indicated by the NMR results it appears that only a fraction of the CD2 is folded correctly. In contrast to the results with the inhibition assay, no mAb binding to the isolated domain when it was bound to nitrocellulose or plastic could be detected. This was also in contrast with the isolated rat CD2.D1 which was able to bind both OX-34 and OX-55 when bound to plastic or nitrocellulose as efficiently as the fusion protein. The explanation for this is unclear but implies that the conformation of hCD2.D1 was totally lost following binding to nitrocellulose or plastic. A second hCD2.D1 construct was examined in which the first 3 amino acids were deleted (Table 8.1) to make the NH_2-terminus more similar to rat CD2 (Fig 8.5). Expression and purification of this construct was performed by Arthur Crawford and myself but it was found to have similar properties to the original form.

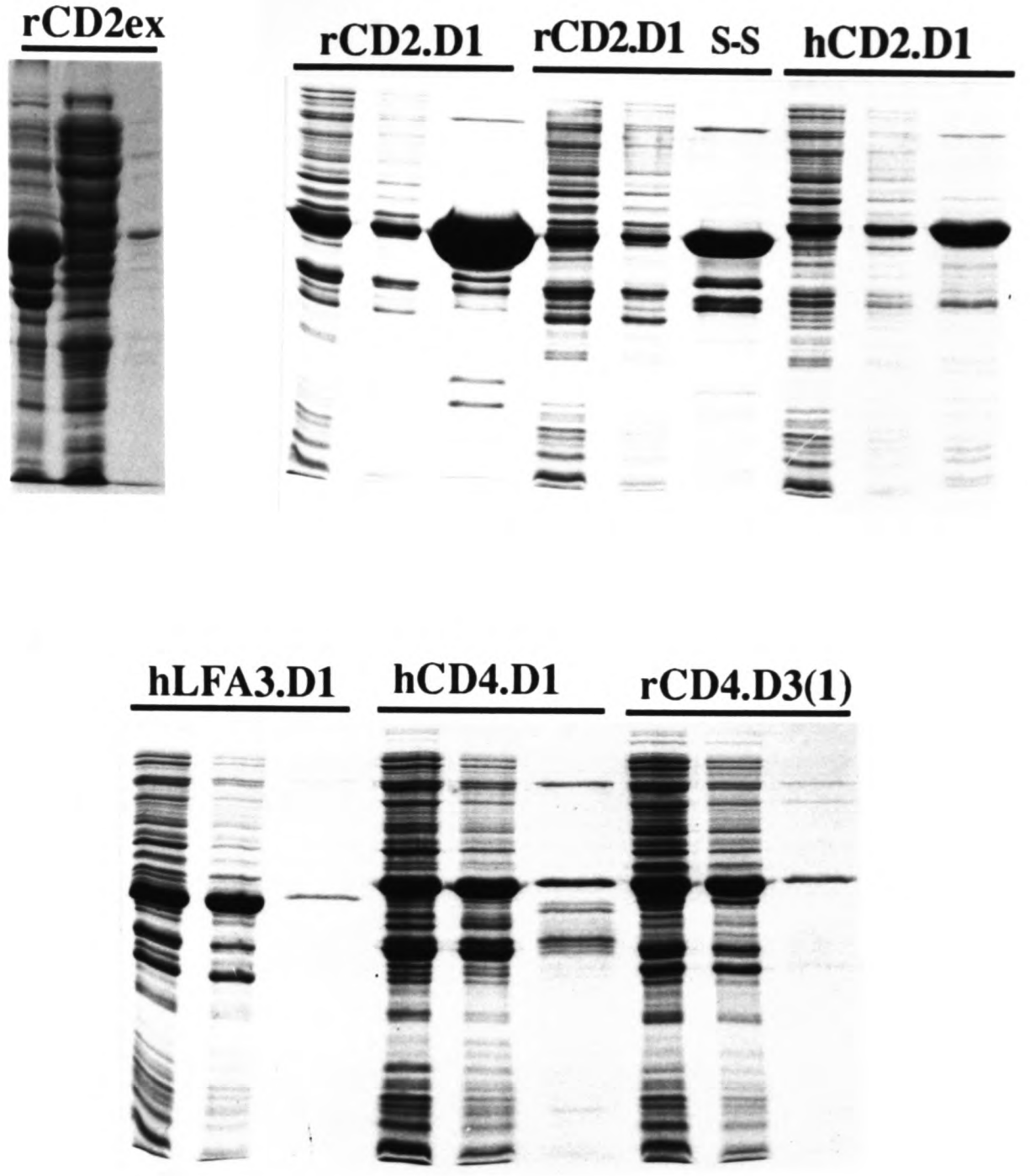


Figure 8.6 Comparison of the expression and solubility of CD2, LFA-3 and CD4 domains in *E.coli* using the pGEX system. Each sample was prepared as follows. 0.5 ml o/n culture was diluted to 5ml using 2xTY in a 50 ml tube and incubated at 37°C for 1hr. IPTG was added and culture continued for 3 hr when the cells were pelleted, resuspended in 800 µl tris/sal/EDTA and transferred to an eppendorf tube, 50 µl was extracted and mixed with 50 µl sample buffer, protease inhibitors were added to the remaining 750 µl and the cells lysed by sonication. The lysate was microfuged for 10 min and the supernatant mixed with 200 µl 50% gluathione agarose for 15 min. The agarose beads were washed 4x with 1 ml tris/sal and then resuspended in 100 µl sample buffer. The lysate pellet was resuspended in 800 µl tris/sal and 50 µl mixed with 50 µl sample buffer. The samples were boiled for 5 min and then run on 15% SDS PAGE in the lane order, from left to right: 8µl cells, 8µl lysate pellet and 8µl agarose/fusion protein, corresponding to 20, 20 and 200 µl of the initital culture, respectively. The fusion proteins involved the following amino acids from the indicated proteins: rCD2ex, 1-177; rCD2.D1, 1-99; rCD2.D1s-s, 1-99 with residues 18 and 78 mutated to cys; hCD2, 1-104; hLFA-3, 2-95; hCD4.D1, 1-97; rCD4.D3(1), 179-290. r indicates rat origin and h human origin.

Table 8.2 Direct Binding of anti-rat CD2 mAb to 1 and 2 domain forms of rat CD2[a]

MAb	GST	GST-CD2.D1	GST-CD2ex	CHO CD2ex
Control(OX21)	256	644	288	1 909
OX34	548	36 822	36 792	27 024
OX53	185	18 473	19 854	nd
OX54	142	1 486	1 073	13 445
OX55	175	28 832	21 859	12 422

[a]Data are shown as cpm and are the average of duplicates. CHO CD2ex is the two domain form of CD2 expressed in CHO cells and was provided by F. Grey. nd indicates not determined

Table 8.3 Direct binding of anti-CD2 mAb to hCD2.D1 fusion protein[a]

MAb	GST	GST-hCD2.D1
F92-3A11	1 841	22 211
95-5-49	1 790	6 512
T11/3T4	3 464	25 478
TS1/8.1.1	1 294	16 746
TS2/18.1.1	1 522	14 154
CLB-T11/1	1 415	17 312
NU-TER	1 992	19 764
RFT11	649	14 140
OKT11	4 124	16 220
OCH-217	1 366	1 816
Control (OX74)	3 284	3 630

[a]Data are shown as cpm and are the average of duplicates. Anti-CD2 mAb RFT11 and OKT11 were kindly provided by M.Brown and all the others by F.Gotch. OX74 is an anti-CD43 mAb.

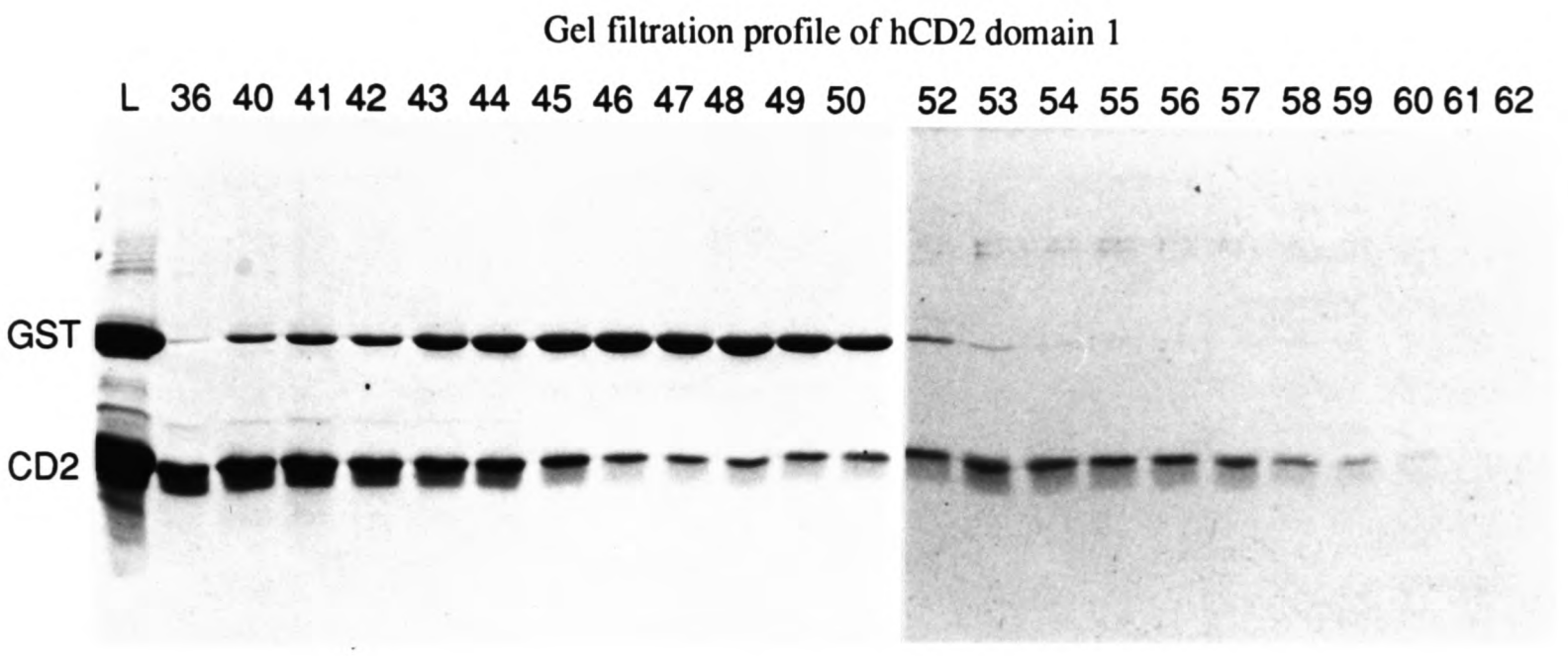


Figure 8.7 Gel filtration profile of human CD2 domain 1. hCD2.D1 from which the GST had been partially removed by batch absorption with glutathione agarose, was loaded onto a Sephadex G75 column (100x1 cm) at a flow rate of 10ml/hr in tris/sal. 2ml fractions were collected and 10µl of the indicated fractions was analyzed on 18% SDS PAGE.

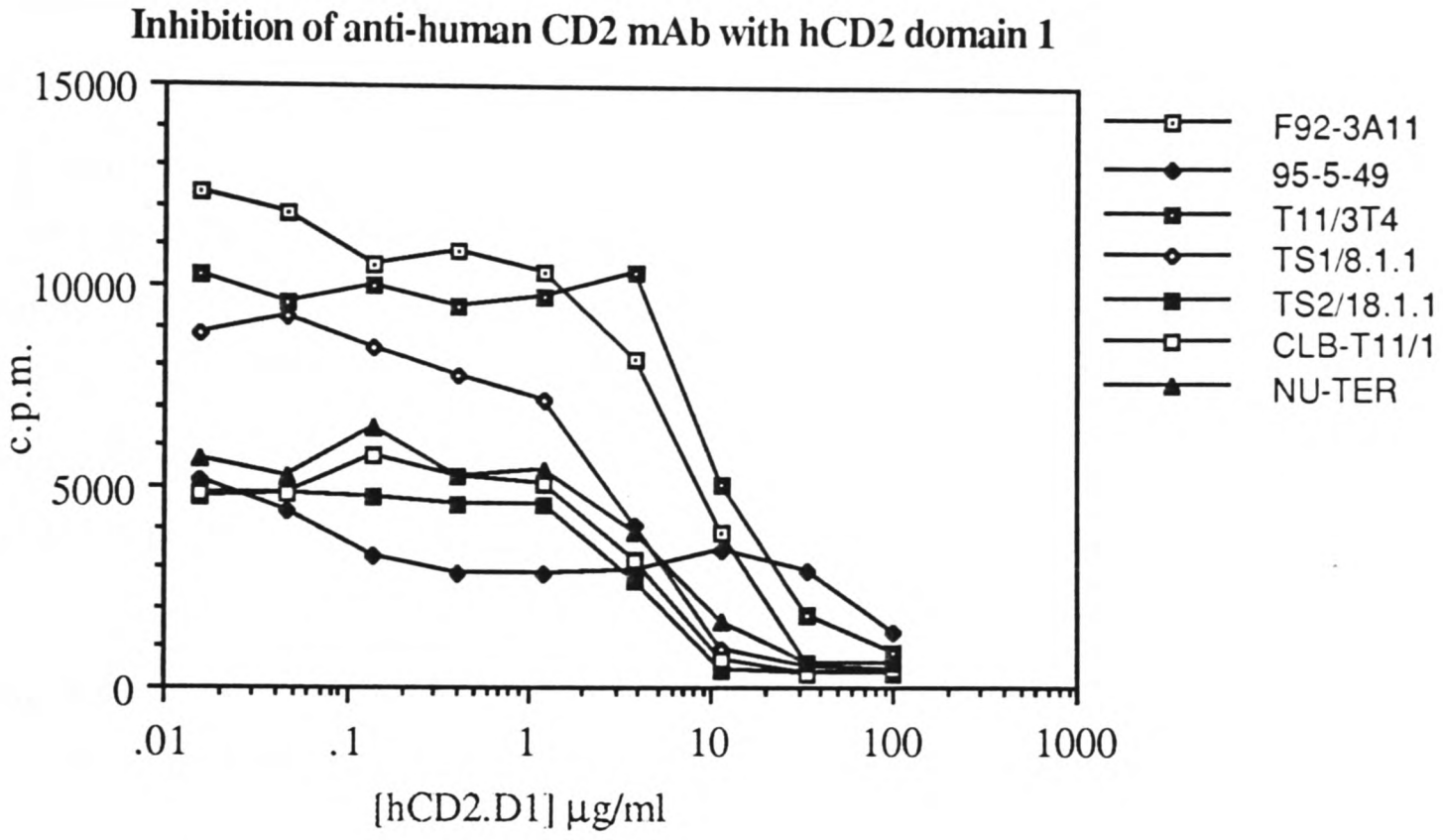


Figure 8.8 Inhibition of anti-human CD2 mAb binding to GST-hCD2.D1 with soluble human CD2 domain 1. Dilutions of soluble CD2 domain 1 were made in PBS/BSA and equal volumes incubated with the indicated mAb for 60 min. 50 µl of each mixture was transferred in duplicate to wells pre-coated with GST-hCD2.D1 and incubated for a further 60 min prior to washing and addition of ^{125}I-labelled RAM. Bound counts were measured and the average of duplicates are shown. The concentration of hCD2.D1 was estimated assuming A_{280}=1 corresponded to approximately 1 mg/ml.

8.7 Human LFA-3 domain 1 is expressed as an insoluble fusion protein

The first domain of human LFA-3 was considered to span amino acids 1 to 95 (Fig 8.5). Because the thrombin cleavage site involves a Gly Ser sequence and the second residue of LFA-3 was Ser, the 5' oligonucleotide was designed leaving out the NH_2-terminal Phe (Table 8.1). This form of LFA-3 was efficiently expressed, using the pGEX-2T vector, in terms of total fusion protein production (Fig 8.9) which was estimated from the gel as 40 mg/l and it reacted with anti-LFA-3 mAB TS2/9 (Sanchez-Madrid et al., 1982) in Western blotting. Only a small amount of material could be purified from the lysate supernatant (0.2 mg/l), however, the majority being present in an insoluble form in the pellet (Fig 8.9). The soluble material showed reactivity with mAb TS2/9 in a direct binding assay (24 300 cpm bound compared with 3 600 cpm bound by the GST control) and this is the first demonstration that this mAb, which can block SRBC rosetting (van Seventer et al., 1989), is specific for domain 1 of LFA-3. To try and improve solubility a number of modifications on the initial construct were examined. In particular, due to the large number of hydrophobic residues near the COOH-terminus (Fig 8.5), different 3' end points were tested. The NH_2-terminus was modified in some cases by introducing one of two types of spacer or linker to allow good separation from the GST partner (Table 8.1). None of the modifications lead to any significant improvement in solubility (Fig 8.9). The inclusion of the GSGSG linker or the sequence RD (the NH_2-terminal amino acids of rat CD2) between the GST and the LFA-3 sequence was found to improve the efficiency of thrombin cleavage of the fusion protein that could be affinity purified from approximately 60% to 95% after 3 hr.

Domain 3 of rat CD4, based on either the intron-exon boundaries (amino acids 179-290) or on IgSF domain alignments (181-287), was expressed in the pGEX system and found to be largely insoluble (Fig 8.6). Human CD4 domain 1 expressed in this system by Chamorro Somoza (Fig 8.6) also gave very low yields of soluble fusion protein.

8.8 Effect of reduced temperature on fusion protein solubility

Growth at reduced temperature has been reported to improve yields of otherwise insoluble fusion proteins (Cabilly, 1989). The various LFA-3 domain 1 fusion proteins and the CD4.D3

Modified Forms of LFA-3 Domain 1 Expressed in *E.coli*

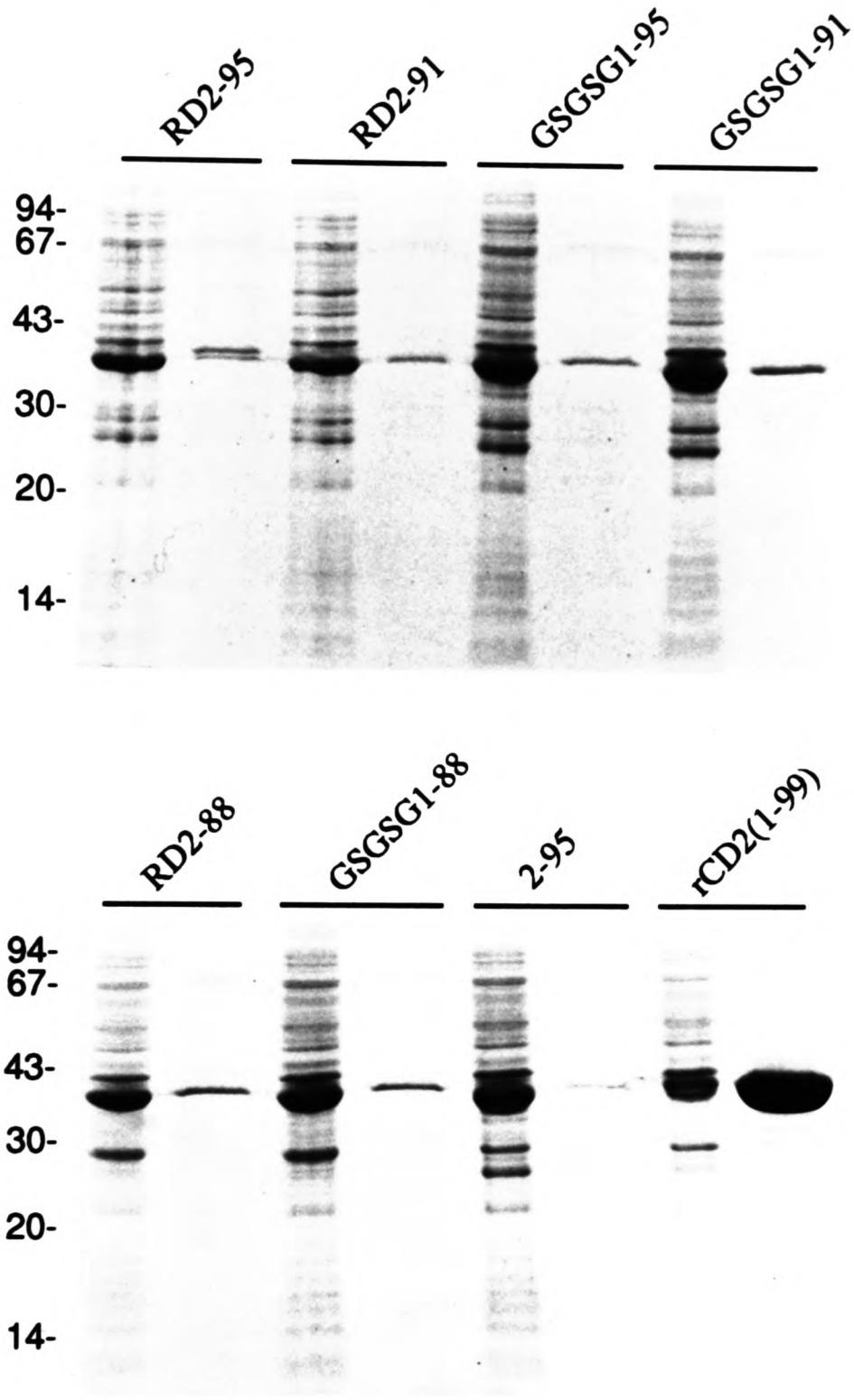
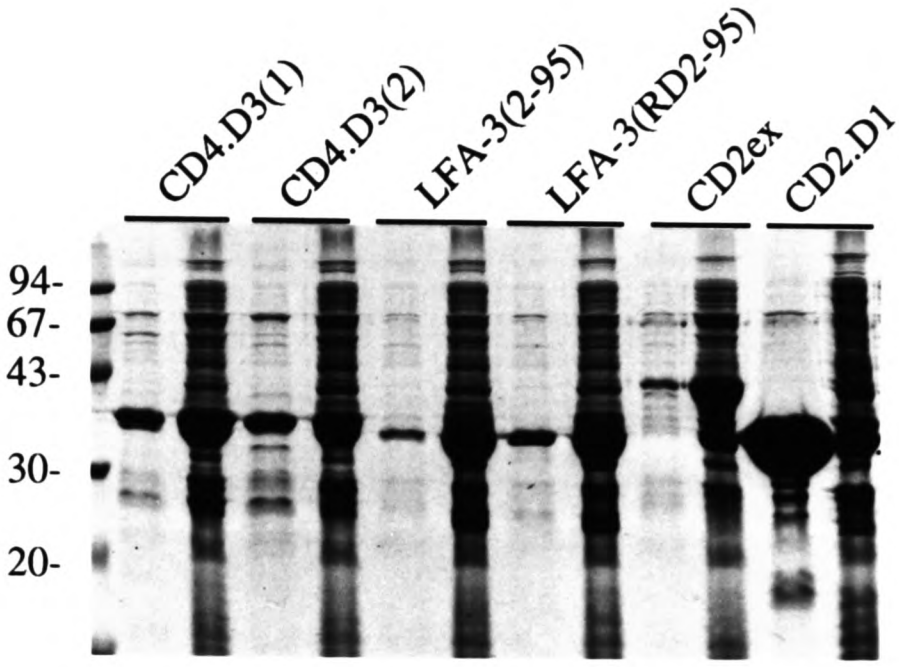


Figure 8.9 Expression of modified forms of human LFA-3 domain 1 and lack of fusion protein solubility. Fusion protein was produced as described for Fig 8.6 except 100µl of glutathione agarose was used in the purification step. The pairs of samples loaded for each construct were: left lane 8 µl lysate pellet, right lane 16 µl agarose/ soluble fusion protein, corresponding to 25 and 500 µl of the original culture. Numbering on the left indicates apparent Mr (kd) determined from marker proteins. Each LFA-3 domain 1 construct is designated by any additional amino acids introduced at the NH_2-terminus followed by the numbers of the most NH_2-terminal and COOH-terminal amino acids present.

A.



B.

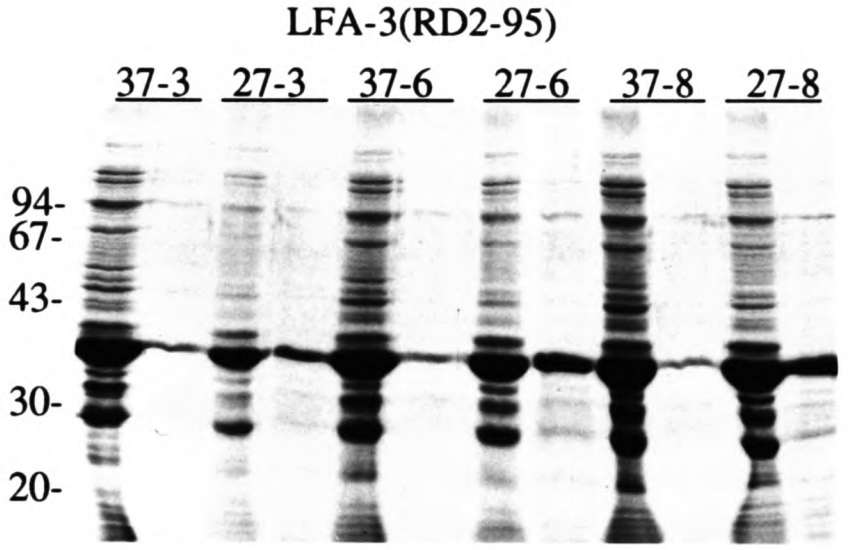


Figure 8.10 Induction of fusion protein expression at reduced temperature and effect on solubility. A) Following o/n growth of the culture all incubations were at 27°C and the fusion proteins were produced as in Fig 8.6 except the 1/10 dilution was grown for 1.5 hr before addition of IPTG and induction was for 6 hr. 16µl of the glutathione agarose / fusion protein sample and 8µl of the lysate pellet was loaded in the left and right lanes of each pair respectively. CD4.D3(1) corresponds to amino acids 179-290 and (2) to 181-287. B) Time course of induction of LFA-3(RD2-95) at 37°C and 27°C. Following 1/10 dilution of an o/n culture the cells were grown for 1.5 hr, IPTG was added and incubation continued for 3 hr, 6 hr, or 8 hr at 37°C or 27°C as indicated. Samples loaded were as in (A) except the pellet sample is on the left and the agarose/fusion protein sample on the right. Numbering on the left indicates apparent Mr (kd) determined from marker proteins.

and CD2ex fusion proteins were grown at 25-27°C during the induction phase and yields of soluble protein compared to those at 37°C (compare Fig 8.10a and Figs 8.6 and 8.9). Improved solubility was seen with some of the fusion proteins, in particular LFA-3(RD2-95). An induction time course of LFA-3(RD2-95) expression at 27°C was performed and yield was found to be maximal after 6 hr (Fig 8.10b) with approximately 2 mg of soluble fusion protein being produced per litre. Cleavage with thrombin was 95% complete after 3 hr and after removal of the GST the LFA-3 was concentrated and 2mg was passed over a Sephadex G75 column. No dominant peak of absorbance was detected but SDS PAGE analysis indicated some material emerged with an elution volume of 60-70 ml and lesser amounts at 140-150 ml, the elution volume for rat CD2.D1. Thus the majority of the LFA-3 produced in this way was aggregated.

8.9 Expression of the isolated LFA-3 domain 1 in *E.coli*

The LFA-3 domain 1 sequence isolated by PCR using the oligonucleotide combination 1278+1279 or 1309+1311 (Table 8.1), was inserted into the Bam HI- Eco RI cut pGM-T7 vector such that the initiation Met in the vector was immediately followed by the Gly Ser of the Bam HI site and then the LFA-3 domain 1 sequence. Transcription in this vector system is driven by the T7 promoter and the constructs were transformed into BL21 (DE3) cells which have the bacteriophage T7 polymerase gene integrated into the genome and under the control of the *lac* repressor (Rosenberg et al., 1987; Studier et al., 1990). rCD2.D1 was also introduced into this vector as a control. Both forms of LFA-3 domain 1 and the CD2 domain 1 were expressed at similar levels in the induced BL21 cultures (Fig 8.11). However, following cell lysis the majority of the LFA-3 remained in the insoluble fraction. The rCD2.D1 appeared totally soluble as in the pGEX system (Fig 8.11). Thus the insolubility of the LFA-3 domain expressed in *E.coli* is not a direct consequence of being linked to a fusion partner.

8.10 Expression of rat CD2 domain 1 containing Cys residues in *E.coli*

Mutated forms of rat CD2 were constructed by He Qi and Fiona Gray in the laboratory with cysteines inserted either at what were predicted to be the conserved positions (replacing Ile 18 and Val 78, construct VW, Fig 8.12a) or with one at the conserved position (Val 78) and one displaced 4 residues further towards the NH_2-terminus along β-strand B (Ile 14, construct VY,

Fig 8.12a). These constructs were designed prior to the determination of the 3-dimensional structure to test the IgSF arguments about the domain 1 fold. The constructs were expressed as soluble forms in CHO cells and a disulphide was found only in the construct where the cysteines were at the conserved positions (Gray et al., 1992). To test the effect of including Cys residues on expression in *E.coli*, these two forms were introduced into the pGEX-2T system (the mutated constructs were used as templates for PCR with the 5' and 3' oligonucleotides used previously to isolate the wild-type domain, Table 8.1). Both were efficiently expressed in *E.coli* as soluble forms although at levels only approximately 50% those seen with the wild-type construct (Fig 8.6; VW and VY expression was similar). Inhibition of OX-34 and OX-55 binding to CD2 by the VW and VY mutated single domains was similar to the wild type (Fig 8.12b). Samples of both the VW and VY fusion proteins were purified and the domains isolated and purified by gel filtration in 0.1M ammonium bicarbonate. Each domain was then denatured (using 8 M guanidinium chloride, 0.1 M Tris) and alkylated (with 4-vinylpyridine) without reduction and subjected to 24 cycles of Edman degradation by Tony Willis. The yield of free Cys (as pyridylethyl-cysteine) in both VW and VY was only 10-20 % of the expected yield had all the Cys been in the reduced form. The lack of free Cys with the VY construct was in contrast with the results obtained for the two domain form expressed in CHO cells (Gray et al., 1992) where the yield was as expected for fully reduced residues. This may, at least in part, be due to formation of intermolecular disulphide bonds and non-reducing SDS PAGE (Fig 8.12c) showed that a fraction of both VW and VY was present as disulphide linked aggregates. The amount of protein in aggregates would not appear adequate to account for the low level of free Cys, however, and it may be the case that VY was aberrantly folded allowing an intramolecular bond to form. The reactivity with mAbs (Fig 8.12b) argues against this and a third possibility is that the domain was not fully denatured prior to alkylation, and some Cys present in the core of the structure remained reduced and thus was not detected on sequencing. It will be important to repeat these experiments, with the inclusion of a control that is both denatured and reduced, to distinguish these possibilities and to confirm that the majority of the VW preparation had the expected intramolecular disulphide bond.

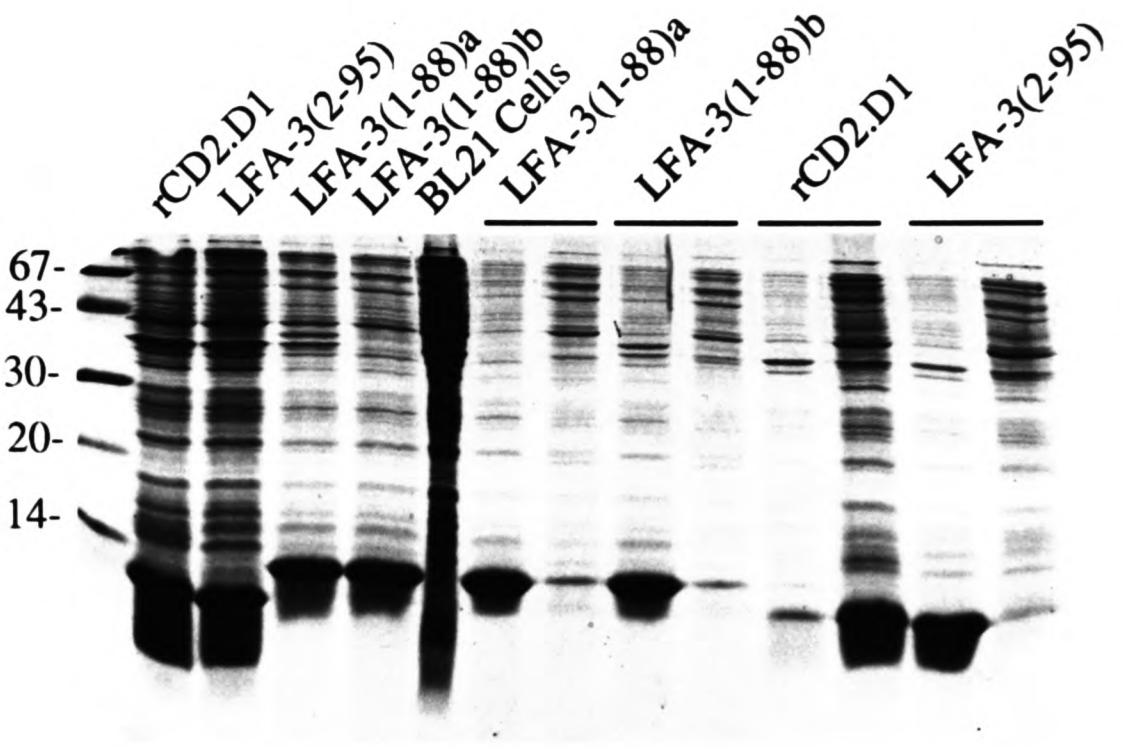
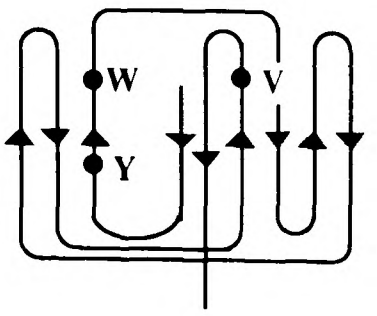
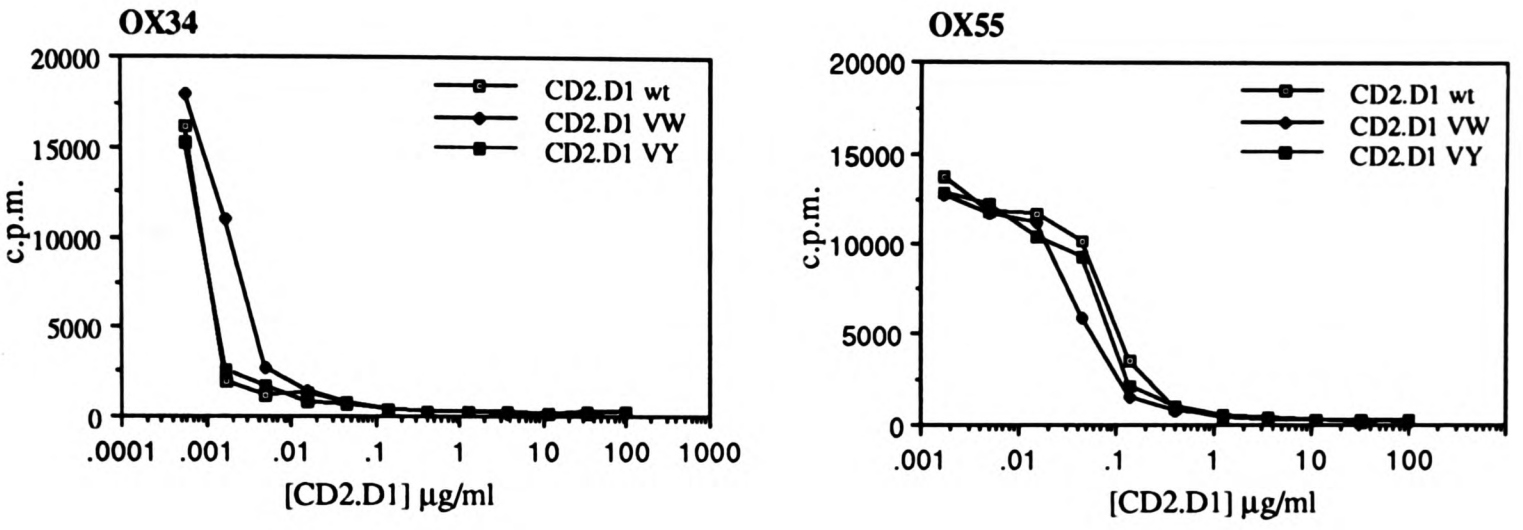


Figure 8.11 Expression of LFA-3 domain 1 as an isolated domain in *E.coli*. BL21 (DE3) cells containing the indicated pGM-T7 constructs were grown o/n in 2xTY, diluted 1/10 and grown for 1 hr prior to addition of IPTG to 0.12mM. After a further 3 hr growth the cells were harvested. For the first 5 lanes 1ml of culture was pelleted and resuspended in 100 µl sample buffer and 8 µl was loaded. For the remaining lanes, 5 ml cultures were pelleted, resuspended in 800 µl tris/sal/EDTA and sonicated 3x 10 sec. The lysate was microfuged for 10 min and the pellet resuspended in 800 µl tris/sal. Samples of lysate supernatant and lysate pellet were mixed with equal volumes of sample buffer and 15 µl loaded per lane on an 18% SDS PAGE gel. For each pair the pellet sample was loaded in the left lane and the lysate supernatant in the right. a and b indicate independent clones. The blurred nature of the bands corresponding to the expressed domains is not due to proteolysis but to incomplete resolution in the gel. This effect was also seen with the rat and human CD2 domains isolated from pGEX fusion proteins and was prevented only by very precise adjustment of the upper and lower gel buffer and the running buffer pH. Numbering on the left indicates apparent Mr (kd) determined from marker proteins.

A



B.



C.

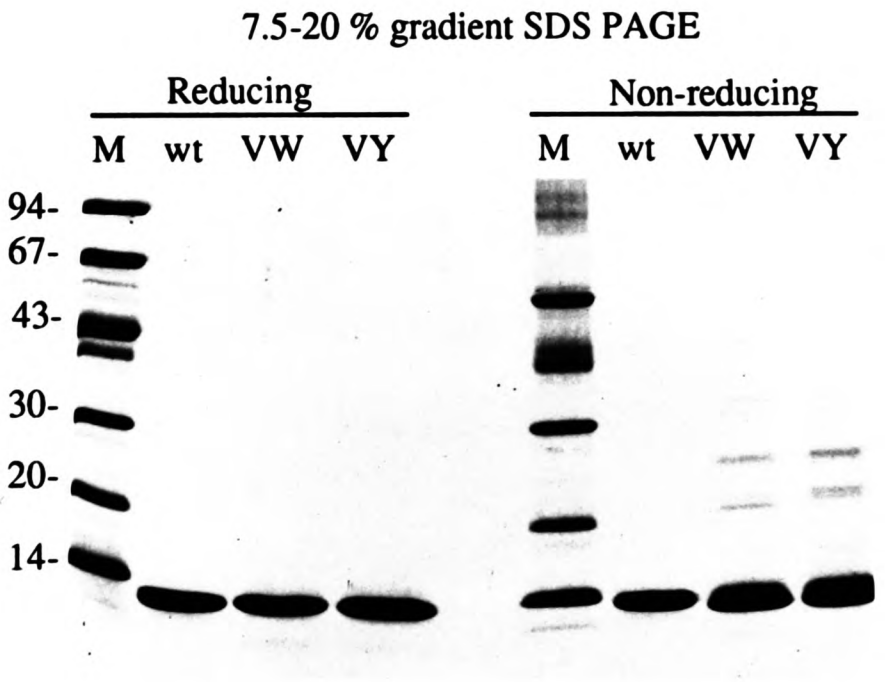


Figure 8.12 Inhibition of anti-rat CD2 mAb by Cys modified rat CD2 domain 1 and SDS PAGE analysis of the mutated domain. A) Schematic illustration of the CD2 fold with the position of Cys reisdues inserted into rat CD2 domain 1 indicated (see text). B) Dilutions of VW, VY or wild type (wt) rat CD2 domain 1 were made in PBS/BSA and equal volumes incubated with OX-34 or OX-55 for 60 min. 50 µl of each mixture was transferred in duplicate to wells pre-coated with GST-rCD2 domain 1 fusion protein and incubated for a further 60 min prior to washing and addition of ^{125}I-labelled RAM. Bound counts were measured and the average of duplicates are shown. C) 5 µg of the wild type CD2 domain 1 or the VW or VY forms, which had been purified by gel filtration and frozen once, were boiled in reducing or non-reducing sample buffer and loaded on a 7.5-20 % gradient gel as indicated. M indicates molecular weight markers and the numbering on the left indicates the Mr (kd).

8.11 Discussion

The expression and purification of rat CD2 domain 1 in *E.coli* and its solubility and stability in deionized water at high concentration (50 mg/ml) has allowed the determination of a low resolution 3-D solution structure of the domain. The complexity of the NMR spectra necessitated the use of ^{15}N-labelled material to allow 3-D NMR spectroscopy and this was produced by expressing the fusion protein in a prototrophic *E.coli* strain in minimal medium supplemented with $^{15}NH_4Cl$. Once the conditions for growth had been optimized the yields of ^{15}N-CD2.D1 were similar to those of the ^{14}N form produced in rich medium.

The structure of CD2 domain 1 fits closely with that predicted from sequence on the basis of superfamily considerations with β-strands B,C,E and F that make up the core of the fold being exactly predicted (Killeen et al., 1988; Williams et al., 1987). The proposal of an α/β type fold from secondary structure predictions (Recny et al., 1990; Clayton et al., 1987) is incorrect. CD2 domain 1 and CD4 domain 2 (Wang et al., 1990; Ryu et al., 1990) were test cases for IgSF arguments since both sequences exhibited unusual features which lead to contradictions between predictions from superfamily sequence patterns (Williams et al.,1989) and *ab initio* computer calculations based on the primary sequences. The superfamily arguments were correct in both cases and this suggests that other IgSF domains predicted on the basis of conserved sequence patterns will also be correct (Williams and Barclay, 1988).

The structure for CD2 domain 1 adds significantly to known variants of the Ig-fold since it provides a prototype for IgSF domains that lack the conserved disulphide bond. Structures determined previously include Ig V and C domains (Amzel and Poljak, 1979), MHC Class 1 α3 domain (Bjorkman et al., 1987) and β-2 microglobulin (Becker and Reeke, 1985; Bjorkman et al., 1987) and CD4 domains 1 and 2 (Ryu et al., 1990; Wang et al., 1990). Ig-like folds have been seen in domains 1 and 2 of the Pap D bacterial chaperone protein (Holmgren and Branden, 1989) and, since the determination of this structure, in the growth hormone receptor domains 1 and 2 (de Vos et al., 1992) and in the fibronectin type III repeat from fibronectin (Baron et al., 1992). These folds should not be categorised as IgSF domains, however, since they show no detectable sequence similarity to IgSF domains (Freeman et al., 1990, and Alan Williams, personal communication). CD2 domain 1 typifies IgSF domains that lack the conserved

disulphide bond. Such domains are found at a moderate frequency amongst domains of cell surface molecules (e.g., LFA-3 domain 1, CEA domain 1, CD4 domain 3, PDGF receptor domain 4 (Williams and Barclay, 1988)) but are the major type in the myosin-binding muscle proteins that are in the IgSF (Labeit et al., 1990).

To facilitate a study of the CD2-LFA-3 interaction in the absence of a rodent LFA-3 sequence the first domains of human CD2 and human LFA-3 were expressed in the pGEX-system. Human CD2 domain 1 was expressed as a soluble fusion protein that could be easily purified with no evidence for the accumulation of insoluble inclusion bodies. The lower yields of human CD2 than rat CD2 may reflect a reduced rate of transcription or translation as eukaryotic codon preferences are not the same as prokaryotic ones and the two sequences may differ in the number of non-prokaryotic type codons (Ikemura, 1981) or some degradation of the domain may have occurred inside the cells. Unexpectedly, the purified domain showed a tendency to aggregate in solution, and a major fraction of the protein precipitated as the concentration was increased above 2-5 mg/ml. Most of the material that remained in solution emerged as a multimer from a gel filtration column. Although at least a fraction of the domain isolated from *E.coli* and present in solution at low concentration was able to bind conformationaly dependent antibodies, NMR spectroscopy indicated that the majority of the protein present at high concentration was not tightly folded. MAb 453 (95-5-49) showed weaker binding characteristics than the other mAb studied and it binds a different region of the molecule (Peterson and Seed, 1987), with mutations around Lys 43 affecting binding. From alignment with rat CD2 this region corresponds to the C-C' loop and C' strand, a region of the protein well separated from the core that may be more susceptible to unfolding. It will be important to test further mAb specific for this region (Peterson and Seed, 1987) to see if it can be incorrectly folded even when the core of the domain is folded.

The instability of the human domain 1 fold may be due to the absence of domain 2 with the possibility that human CD2 has a larger area of contact between domains than does the rat molecule. In addition the presence of the single N-linked glycan on CD2 domain 1 (Moingeon et al., 1989) may be necessary for maintaining a stable fold or for covering an area of hydrophobicity on the surface of the structure. The attachment of the N-glycan core occurs co-

translationally and there is evidence for a variety of proteins that the N-linked sugars are important either in folding or in maintaining the normal conformation of the protein (Elbein, 1991). Some evidence for the importance of the N-glycan on human CD2 domain 1 comes from the work of Luther et al., (1991) showing that N-glycanase treatment of CD2 disrupts epitopes critical for CD2-LFA-3 binding. However, these workers also argued that CD2 domain 1 cannot be expressed in a soluble form in *E.coli* which clearly is not the case. It will be important to test whether the soluble CD2.D1 described here is able to bind LFA-3. The possibility that glycosylation may improve stability is currently being investigated by expression of the domain in yeast. Production of the domain in an environment containing eukaryotic chaperones may also enhance correct folding.

The rules governing solubility and insolubility of proteins expressed in *E.coli* are not understood (Marston, 1986; Hartley and Kane, 1988). Thus the insolubility of LFA-3 domain 1 (the two domain form of CD2 and CD4 domains 1 and 3) could be due to a variety of factors including an inability to fold correctly in the absence of domain 2, an instability of the fold in the absence of domain 2, a requirement for eukaryotic chaperones, the exposure of hydrophobic surfaces during one or more folding steps or in the final product, or an instability of the fold in the absence of N-glycans. N-glycans should not be important for LFA-3 folding as LFA-3 isolated from cells grown in the presence of tunicamycin is able to bind CD2 (Chan et al., 1991, and Po-Ying Chan, personal communication). LFA-3 domain 1 contains more hydrophobic residues in its sequence than CD2 particularily in the region of strand G. The rapid rate of synthesis and accumulation of proteins expressed in *E.coli* from expression vectors means that any tendency to aggregate will be pushed forward.

The introduction of Cys residues into CD2 domain 1 did not lead to fusion protein insolubility and the purified Cys containing domain inhibited antibody binding with equal potency to the wild type domain. This result was not unexpected since the environment inside *E.coli* is reducing and disulphides generally do not form (Schoemaker et al., 1985; Tuggle and Fuchs, 1985; Marston, 1986) leading to the prediction that the presence of Cys *per se* should not lead to insolubility. That is, if the structure can form a stable fold in the absence of disulphide bonds, and the presence of Cys residues does not interrupt this fold, then there is no reason to

expect aggregation to occur. The general problem with expressing Cys rich domains that normally have disulphide bonds inside *E.coli* may lie in the requirement for the bond(s) to maintain some elements of the fold. If the fold is not stable, associations through abnormally exposed surfaces may occur resulting in insolubility and evidence that some Ig V domains require the disulphide bond for fold stability has recently been reported (Glockshuber et al., 1992). It is unlikely, therefore, that deleting Cys residues will improve solubility of Ig domains expressed inside *E.coli*. The advantage of the periplasmic expression systems, which have been used successfully for the soluble expression of many antibody fragements (Better et al., 1988; Skerra and Pluckthun, 1988; Wetzel, 1988; Ward et al., 1989; Glockshuber et al., 1990; Winter and Milstein, 1991), may relate both to the oxidizing conditions in the periplasm and the presence of disulphide isomerases (Bardwell et al., 1991; Kamitani et al., 1992). The slower rate of accumulation of recombinant protein may also be important in reducing aggregation of partially folded intermediates.

Expression of single IgSF domains in soluble form has been problematic not only in *E.coli* (Kuwana et al., 1987; Udaka et al., 1990) but also in eukaryotic systems (Traunecker et al., 1989; Davis et al., 1990b; Classon et al., 1992). The high yields of soluble rat CD2 domain 1 and its stability as an isolated fold make further studies with this domain worthwhile. The construction of LFA-3-rCD2 chimeras is presently being pursued in the laboratory by Doug McPherson where strands A and G of LFA-3 are exchanged for those of rat CD2. Preliminary results indicate an improved solubility. The possibility of taking loops from IgSF members of unkown structure and replacing the corresponding region in rat CD2 (especially in the C'-C" region) is also being considered as a tool in the analysis of unkown IgSF domain structures.

The production of ^{13}C labelled rat CD2 domain 1 will allow an increased number of assignments to be made and should lead to a high resolution structure for the domain. The structural information available at this point will allow an extension of the work by Peterson and Seed (1987) in a directed mutagenesis approach, targeting outpointing residues on the face of the sheet that appears to be involved in the interaction with LFA-3. In future work the site of interaction with other potential ligands, especially CD48 (OX45) and CD59, will be studied.

CHAPTER 9

GENERAL DISCUSSION

9.1 General discussion on leukosialin

9.1.1 Summary

9.1.2 Molecules at the cell periphery

9.1.3 Glycosylation and protein-protein interactions

9.1.4 Carbohydrate-lectin interactions: a role for leukosialin?

9.1.5 Leukosialin as a leukocyte mucin

9.1.6 Conclusions

9.2 General discussion on CD2

9.2.1 Summary

9.2.2 Ig type folds and their surfaces of interaction

9.2.3 Modulation of adhesion molecule binding sites by conformational changes

9.2.4 Glycosylation and the LFA-3 binding site on CD2

9.1 General discussion on leukosialin

9.1.1 Summary

The studies described in Chapters 3-7 lead to the following observations. A single functional gene for leukosialin lacking introns in the coding region, and a non-functional pseudogene, are present in the mouse genome. Thus all the heterogeneity reported for mouse leukosialin must be at the post-translational level. Expression of leukosialin on B cells in transgenic mice did not lead to any major alteration in the migratory behaviour of these cells or in their antibody response to complex antigens. Electron microscopy of purified leukosialin showed the extracellular domain can extend to a length of 45 nm. The majority of mAb specific for this domain were able to bind a non-glycosylated form expressed in *E.coli* and this was despite the sensitivity of binding of some of these mAbs to the nature of the O-linked glycans attached to forms on eukaryotic cells. These epitopes mapped to short linear protein sequences. Thus the exposure of many protein determinants in the leukosialin backbone is modulated by the nature of the O-linked glycans. Homotypic aggregation of thymocytes induced by anti-leukosialin mAbs required bivalent antibodies, energy and an intact cytoskeleton. A correlation was observed between the potency of mAb in inducing cell aggregation and the location of the epitope in the extracellular domain with mAb that bound at the greatest distance from the transmembrane region being most potent. The 124 amino acid cytoplasmic domain of leukosialin, produced in *E.coli,* had an elongated structure and was amenable to phosphorylation at two or more sites by protein kinase C.

9.1.2 Molecules at the cell periphery

Immunochemical studies indicate that the major surface glycoproteins of rodent T lineage cells include leukosialin, CD45, and Thy-1 for thymocytes and leukosialin, CD45 and MHC Class I antigen for T lymphocytes (Williams and Barclay, 1986). Other glycoproteins, including the TcR and the CD2, CD4, CD5, CD8 and LFA-1 antigens, are more minor constituents on T cells. The topology of the three major glycoproteins of thymocytes can be envisaged as illustrated in Fig 9.1. Leukosialin and CD45 are likely to extend much further from the cell surface than Thy-1, and the TcR, MHC antigens and CD2 have dimensions only marginally larger than Thy-1. In the case of CD45, the extra exons shown as A, B, or C (Fig 9.1) have a high content of O-linked sugars and are therefore likely to have an extended conformation similar to leukosialin. Thus the B lymphocyte form of CD45 which includes the A, B, and C segments, would be expected to have a total extracellular length of about 55 nm with 28 nm

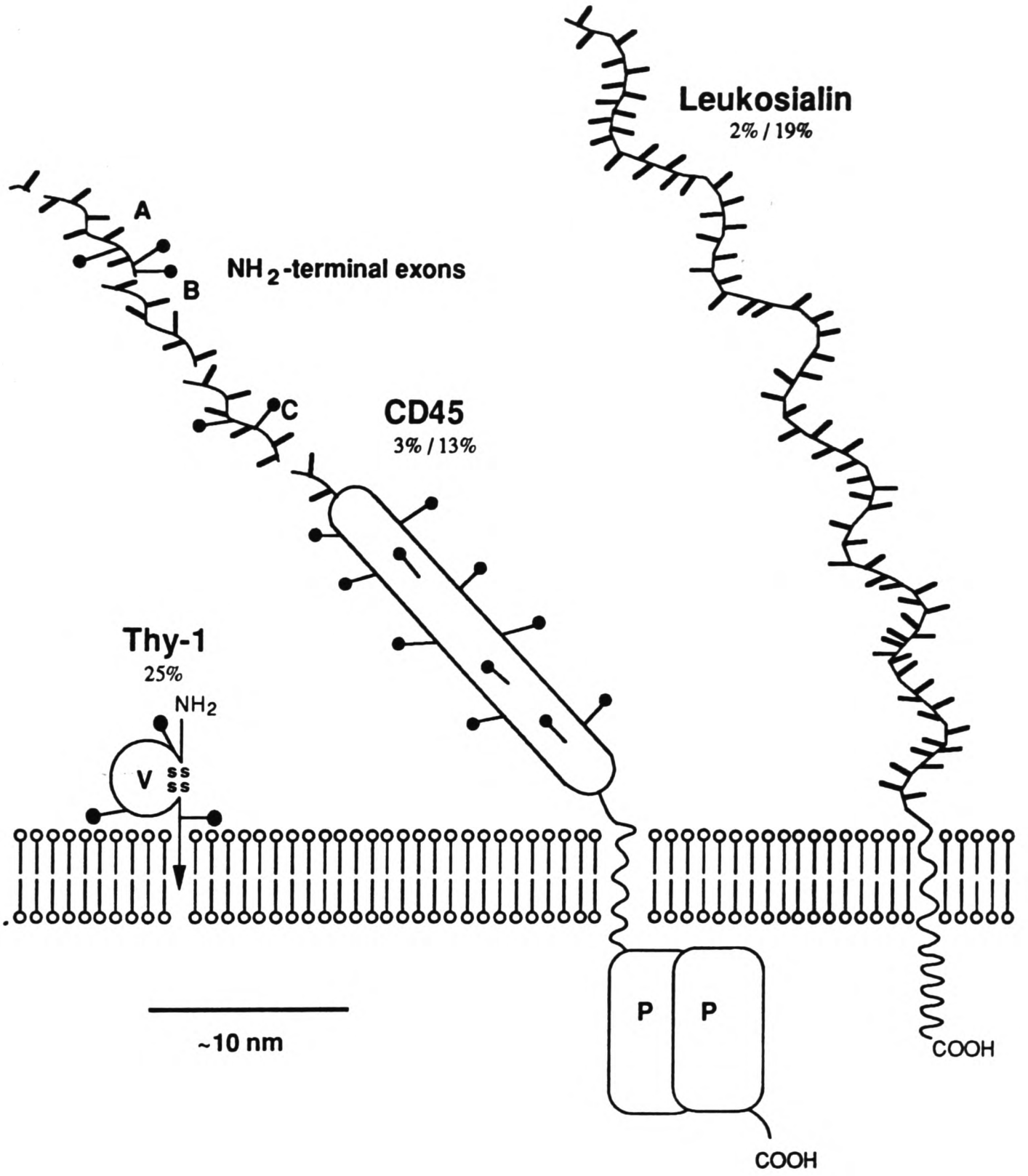


Figure 9.1 A model of the major glycoproteins of rat thymocytes drawn roughly to scale. The dimensions for Thy-1 are based on an Ig fold and the Stokes' radius of the molecule (Williams and Barclay, 1986) while those for leukosialin and CD45 are based on electron microscopy from Chapter 5 and Woollett et al., 1985. The symbols ♀ and | indicate N- and O- linked carbohydrate respectively. The figures beside each molecule represent the percentage of the thymocyte surface that is calculated to be covered by each of these molecules. For these calculations the thymocyte was taken to have a smooth surface of area 120 μm^2, Thy-1 as a sphere of radius 3 nm, and leukosialin and CD45 as rods of 45 x 5 and 28 x 8 nm, respectively. The site numbers was taken to be 10^6, 10^5 and $7x10^4$, respectively (Williams and Barclay, 1986). With leukosialin and CD45 the area covered was calculated either from the cross-sectional area assuming the molecules project perpendicular to the membrane (value on the left), or from the maximal area the molecules would cover if lying across the lipid bilayer in a fixed position. The values for all three molecules are likely to be overestimates as the cell surface generally contains numerous microvilli and these can increase the surface area at least 3 fold (Yang et al., 1992).

derived from the membrane proximal region (Woollett et al., 1985) and 27 nm from the 140 amino acids of the A, B, and C segments. This has recently been confirmed by electron microscopy on a soluble form of CD45R(ABC) (McCall et al., 1992). On the basis of the models in Fig 9.1, leukosialin and CD45 are likely to be the leukocyte membrane glycoproteins that most readily make contact with other surfaces.

The possiblity that leukosialin generally does not extend from the cell, as depicted in Fig 9.1, but rather is bent back towards the membrane is unlikely for a number of reasons. 1) Heavily O-glycosylated molecules have less flexibility than unglycosylated proteins of similar sequence and most of the leukosialin molecules viewed by electron microscopy are extended in one direction. 2) Epitope mapping indicates that mAb which are raised against leukosialin on the cell membrane tend to bind epitopes nearer the NH_2-terminus of the molecule than the membrane proximal region. 3) The number of amino acids in the extracellular domain of mouse, rat and human leukosialin is similar indicating that the molecular extension of the molecules is likely to be similar, a prediction confirmed for rat and human leukosialin by electron microscopy. This conservation of length is despite a very low level of sequence identity, and the apparent occurrence of repeat events at different times in evolution. If the molecules were not projecting in an extended form from the membrane, at least part of the time, the selective pressure for conservation of length would not be expected to be high. 4) The greater potency of mAb specific for epitopes nearer the NH_2-terminus than in the middle of the domain at inducing cell aggregation, which is most likely a result of cell-cell cross-linking, argues that the NH_2-terminus of the molecule is generally displaced further from the membrane than the middle of the molecule. 5) All forms of leukosialin studied, including the thymocyte form, contain sialic acid and thus charge repulsion between leukosialin molecules and between leukosialin and other sialylated glycoproteins and glycolipids could act to keep the molecule in an extended form. 6) If such structures did not project from the surface then molecules like CD44 and GP1bα, which have very extended stalk regions of similar composition to leukosialin (Table 1.1), would gain little obvious advantage from having these regions between the ligand binding NH_2-terminal domain and the transmembrane region.

9.1.3 Glycosylation and protein-protein interactions

The possibility that the peptide backbone of the leukosialin molecule is open for interaction with other proteins is shown by the experiments described in Chapter 5. Thus despite the attachment of one O-linked glycan approximately every three residues (Fig 9.2), 7 out of 9 mAbs raised

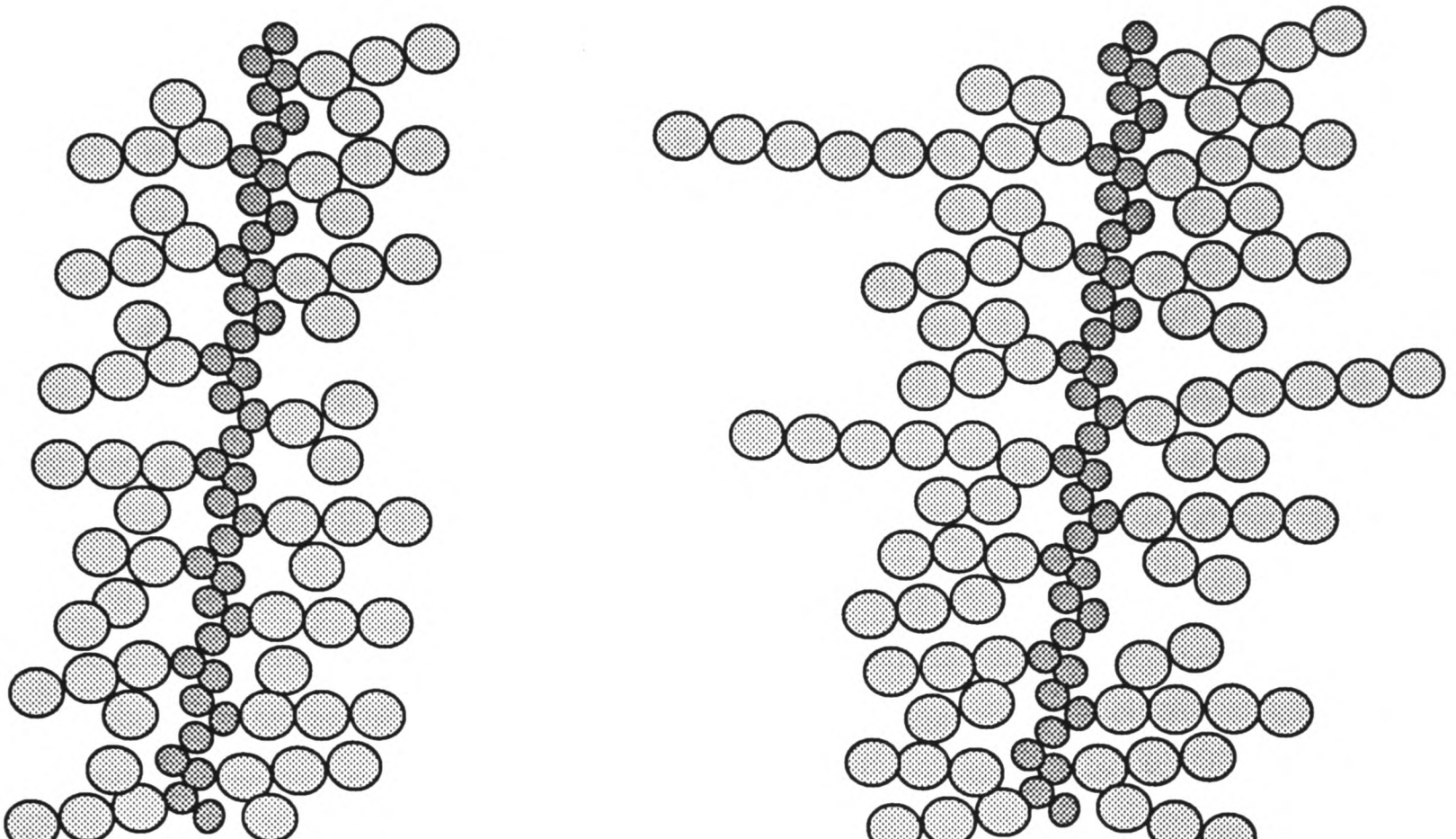


Figure 9.2 Models for segments of resting T cell and granulocyte leukosialin. Each amino acid is shown as a small circle and each monosaccharide as a large circle. The components are drawn roughly to scale (1 mm = 0.1 nm) based on an extension per amino acid of 0.38 nm and an average extension per 10 amino acids of 2 nm (as the length of the 224 amino acid extracellular domain is 45 nm). Monosaccharides were taken to correspond to spheres of 0.6 nm diameter (Rademacher et al., 1988). One O-linked glycan is attached per three amino acids. The carbohydrate structures correspond to those shown in Fig 1.6.

against a glycosylated form of leukosialin are able to bind the peptide core. The possibility that all these mAb bind regions of the core free of O-glycans is excluded by the ability to obscure some of the peptide epitopes by removing sialic acid from the carbohydrate chains. The sensitivity of mAb binding to expression on different cell types, some of which are known to attach different O-glycans to leukosialin, also supports the argument that the mAbs do bind regions which are readily glycosylated. As described in Chapter 5, a number of reports have shown that mAbs specific for epithelial mucins bind the peptide core. Many of these antibodies are also sensitive to removal of sialic acid from the glycosylated forms of the protein. This ability of antibodies to bind the protein core alone in these heavily O-glycosylated regions was not generally expected and a number of groups attempting to clone mucin glycoproteins raised anti-sera against the deglycosylated mucin to allow screening of lambda gt11 expression libraries (Gendler et al., 1987; Timpte et al., 1988; Gum et al., 1989; Bhargava et al., 1990; Gum et al., 1990; Lan et al., 1990; Porchet et al., 1991) only to find subsequently that pre-existing mAbs, which had been raised against the glycosylated mucin, also bound the recombinant proteins they isolated. Indeed, there are still frequent reports which conclude that the epitope seen by a mAb is carbohydrate in nature solely because reactivity is lost following treatment of the antigen with neuraminidase (Waibel et al., 1987; Stahli et al., 1988; Kim et al., 1989). This points to our lack in understanding of the way proteins interact with their attached carbohydrates, and how the attached carbohydrates interact with each other. The stability of such interactions is of importance in considering how changes in glycosylation could modulate protein-protein interactions.

Changes in the nature of N-linked glycans can alter protein-protein interactions. Possibly the first example was the reduced binding of fibronectin to gelatin resulting from the presence of poly-lactosamine chains on the N-glycans of fibronectin (Zhu and Laine, 1985). The specific bioactivity of a variety of hormones, such as GM-CSF, follicle stimulating hormone and thyroid stimulating hormone can be influenced by their glycosylation state and in some cases this is altered by physiological stimuli (Rademacher et al., 1988). In relation to cell-surface adhesion molecule interactions, in addition to the effect of the polysialyl N-glycan on NCAM mediated adhesion described earlier, the binding of MAC-1 to ICAM-1 domain 3 is also affected by glycosylation. This domain of ICAM-1 has two N-linked glycosylation sites and disruption of either site by mutagenesis enhances the interaction with Mac-1 and there is also a correlation between glycan size and the strength of the interaction (Diamond et al., 1991). ICAM-1, like many other cell membrane glycoproteins (Rademacher et al., 1988), shows differences in

apparent M_r depending on the tissue from which it is purified and these have been attributed to differences in N-glycosylation. A regulatory role for these differences in ICAM interactions is apparent and it is possible that this is also true for many other membrane glycoprotein interactions.

Protein-protein interactions involving adhesion molecules where the protein core is in a heavily O-glycosylated protein may include interactions between leukosialin and ICAM-1, the NH_2-terminal region of CD45 and CD22 and possibly between E2 and LFA-3. It remains to be shown whether the protein backbone, the carbohydrate or both are involved in each case. This is an important point because CD22 also binds to CD75, which has recently been shown to be a carbohydrate determinant (Bast et al., 1992), or at least CD22 binding to B cells is blocked by antibodies that are specific for CD75. ICAM-1, CD22 and LFA-3 have extracellular domains made up of IgSF domains. Excluding carbohydrate specific mAbs, there are no reports of IgSF domains interacting specifically with N- or O-linked carbohydrate although there is some evidence for binding to glycosaminoglycans (Nybroe et al., 1989). It will be important to determine if the above interactions have a protein-protein component, as the results in Chapter 5 imply they could, and to see how they are modulated by the nature of the O-glycans.

9.1.4 Carbohydrate-lectin interactions: a role for leukosialin?

Carbohydrate-lectin interactions have a role in most, if not all, circulating cell - endothelial cell interactions (Aizawa and Tavassoli, 1988; Cornil et al., 1990; Butcher, 1991). A likely explanation for this (Williams, 1991) is that the high density of carbohydrate ligands on the cell surface supports the rapid kinetics of interaction needed to retain or capture a cell from the blood stream. In addition to the density of the carbohydrate ligands it is likely that their distance from the cell membrane will be significant. Although the majority of work indicates that selectin ligands are carried on N-linked carbohydrates, the O-linked glycans on human neutrophils have been shown to include sLe^x determinants (Fukuda et al., 1986) and selectin binding to the U937 cell line is not dependent on N-linked glycosylation (Leeuwenberg et al., 1991). Thus the possibility that leukosialin carries ligands for selectins in a favourable location for binding has not been ruled out. The occurrence of soluble leukosialin in serum that most likely derived from granulocytes, and the data indicating that partial downregulation of leukosialin following TNFα treatment of granulocytes can be prevented by protease inhibitors (Campanero et al., 1991), suggests that leukosialin has a rapid release mechanism similar to the selectin molecules (Huber et al., 1991). This may be important in allowing adherent cells to migrate across the

endothelium. It was intended to examine the possibility that leukosialin on granulocytes could be a selectin ligand in the rat but some preliminary experiments where bone marrow and spleen cells were probed with anti-Le^x and anti-sLe^x antibodies revealed minimal expression of these determinants. Western blotting of partially purified leukosialin from rat bone marrow (which is a rich source of mature granulocytes, Williams et al. (1977)) did not reveal mAb binding. Sialyl-Lewis x is expressed as a stage specific antigen in the mouse embryo (termed SSEA-1; Solter and Knowles, 1978) but, to my knowledge, expression in adult mice has not been reported. This may explain why antibodies to this determinant could be raised in rodents. Thus the selectin ligands in rodents may differ from those in man and the possiblity that some forms of leukosialin carry these determinants remains to be examined.

A role for cell-type specific O-linked carbohydrates in lymphocyte-target cell interactions in the tissues has been suggested by work with two mAb raised against CTL that block the killing of target cells (Lefrancois and Bevan, 1985). These mAbs, CT1 and CT2, react with related O-linked carbohydrate structures present on CD45, a protein of Mr 140 kd that may be leukosialin, and a smaller protein (Lefrancois et al., 1985; Conzelmann and Lefrancois, 1988). Expression of these determinants is restricted to neonatal thymocytes, intestinal intra-epithelial lymphocytes and activated CD8 T cells (Conzelmann and Lefrancois, 1988). The nature of the natural ligand for these carbohydrate determinants and how widely it is expressed remains to be determined. This work is also interesting because it shows that specific modification of O-linked glycans can be restricted to a subset of O-glycosylated molecules. Related studies have shown that NK cell killing of targets can be inhibited by poly-lactosamines and there is evidence for a direct interaction between such structures on NK cell CD45 and lectins on tumour target cells (Gilbert et al., 1988). Leukosialin may also have a role in this type of interaction as its glycosylation state on NK cells has not been characterized.

Another case for a leukosialin-lectin interaction may be the binding of thymocytes to a galactose specific lectin in the thymic cortex. Leukosialin and other molecules on thymocytes are less sialylated than on mature T cells (Pink, 1983) and have more carbohydrate with terminal galactose residues. The existence of a Gal specific lectin in the thymic cortex (Levi and Teichberg, 1983) suggests that a Gal-lectin type interaction may serve to retain thymocytes in the cortex until they pass through a developmental stage that leads to increased sialylation of surface proteins (Sharon and Lis, 1989). Leukosialin is the dominant PNA binding protein on thymocytes and the reduced PNA binding of the more mature thymocytes in the medulla (Rose,

1982) indicates that leukosialin in particular undergoes changes in the number of exposed Gal residues as the cells mature, and thus may be an important ligand for such a lectin.

The distribution of lectins that bind leukosialin may best be studied by purification of leukosialin from specific cell populations (for example, thymocytes or bone marrow derived granulocytes) and using this as a first step reagent in staining tissue sections. If the purified material is used in the absence of detergent it should exist as multivalent complexes (seen in Fig 5.1 and 5.11, some complexes have more than 20 leukosialin molecules) allowing avidity to compensate for a potentially low affinity. The high level of glycosylation of leukosialin makes it a particularly good reagent to assess lectin distribution.

9.1.5 Leukosialin as a leukocyte mucin

Leukosialin is frequently classified as a mucin (Hilkens, 1988; Strous and Dekker, 1992). A major function of epithelial mucins is to provide lubrication of the epithelial surface and protection of the epithelial cells from the lumenal contents. The intertangled network they form reduces penetration of large molecules while allowing relatively free diffusion of small molecules and ions. Leukosialin may serve analogous roles, such as protection of circulating lymphocytes from unwanted interactions with other cells or plasma components. Large mucin molecules present on carcinoma cells appear active in masking small surface glycoproteins, such as class I MHC, both from interactions with other cells and from recognition by antibodies (Codington and Frim, 1983). The mucins may also act to mask adhesion molecules and it has been found that cancer cells expressing high levels of mucin glycoproteins down regulate expression before they adhere together to grow as a solid tumour (Carraway and Spielman, 1986). Leukosialin expression, however, is not downregulated when lymphocytes enter the tissues where they must partake in numerous adhesion interactions. Indeed, following activation, leukosialin expression may be upregulated (Baecher et al., 1990) and the size of the sugar chains increased (Piller et al., 1988). Thus if leukosialin is a masking molecule it must be argued that either (1) the size of the molecule and/or its level of expression is not so great as to prevent small molecules on the membrane interacting with similar molecules on other cells, (2) the molecule has a polarized distribution on motile cells, with the majority of leukosialin being moved off the face of primary interaction or (3) leukosialin is directly involved in the first steps of the adhesion process and its masking function then becomes irrelevant. The situation is unclear in relation to points 1 and 2 whereas there is evidence to support point 3.

Leukosialin expression has not been found to correlate with a reduced ability of mAb to bind to other cell surface molecules. It is possible that leukosialin and similar molecules reduce the kinetics of antibody interaction with small molecules on the membrane compared to binding to the isolated protein but this has not been analyzed. Indirect evidence that leukosialin can negatively influence the efficiency of cell-cell interactions is the lack of leukosialin expression on the professional antigen presenting cells, dendritic cells (Prickett et al., 1992) and B cells (Williams et al., 1977). The picture is complicated, however, at least with B cells, as they express a heavily O-glycosylated form of CD45 at similar levels to leukosialin expression on T cells. Furthermore, the work reported in Chapter 4 indicates that expression of leukosialin on B cells (admittedly at a level 10 fold down from T cells) does not affect their ability to recieve T cell help and mount an antibody response. In addition, leukosialin expression is up-regulated on B cells within 72 hours of activation (by LPS; Gulley et al., 1988) and thus the cells express leukosialin during a period in which they are likely to be involved in many interactions.

Support for the possiblity that leukosialin is a good defence against close apposition between membranes of non-polarized cells is that the dimensions of known molecular interactions involved in primary adhesion events of leukocytes are similar to the maximal length of leukosialin. P- and E- selectins have dimensions of approximately 40 and 30 nm, respectively, and CD44 has a heavily glycosylated stalk region of greater than 100 amino acids, that may place the hyaluronate binding domain more than 20 nm from the membrane. The ligands for the selectins and CD44 may themselves be displaced from the lipid bilayer as already suggested for the selectins. The LFA-1 - ICAM-1 and VLA-4 - VCAM-1 interactions are likely to be active at a membrane separation of approximately 40 nm. Although this argument is very indirect, Yang et al. (1992) recently reported that the presence of poly-sialic acid on NCAM increases the intracellular space between NCAM expressing cells brought together by centrifugation. They suggest that interactions between cells carrying polysialylated NCAM would be through adhesion molecules of greater length than would need to be the case with cells carrying unmodified NCAM.

Evidence for the second possibility, that leukosialin is redistributed away from the face of interaction, comes from the work of de Petris and Baumgartner (de Petris and Baumgartner, 1982; de Petris, 1984). They found that the majority of leukosialin on polarized thymocytes, T cells and granulocytes was concentrated on the uropod. Cells without clear polarity showed an even distribution of label over the cell surface. The uropod is the trailing edge of the cell (Parrot

and Wilkinson, 1981) indicating that the leading edge is preferentially cleared of leukosialin. This picture is complicated however because, while there is evidence that the first contact between a migrating lymphocyte and a second cell occurs at the leading edge, the lymphocyte reorientates such that the uropod makes greater and more extended contact with the target or antigen presenting cell (Parrot and Wilkinson, 1981). Furthermore, ICAM-1 also shows an increased density on the uropod of polarized cells and Rothlein and Springer observed that cells coming together in homophilic aggregates after PMA treatment remained in the aggregates through their uropods (Rothlein and Springer, 1986). These observations suggest that leukosialin can actually be focussed on a region of the cell membrane involved in interactions. Thus it is unclear whether the first adhesion processes do occur in a region of reduced leukosialin concentration or if the converse situation holds. The distribution of leukosialin on cells at various stages of conjugate formation needs to be studied to address this issue.

The third possibility, that leukosialin can take a direct part in the interaction process, is indicated by three lines of work. Firstly, as already described, leukosialin may be involved in binding to cell surface lectins. Secondly, the data presented in Chapter 7 and some of the data previously presented by others (Axelsson et al., 1988; Nong et al., 1989), indicate that bivalent mAb to leukosialin are very potent at inducing homotypic cell aggregation. These mAbs are unable to activate cells on their own and it appears most likely (although it remains to be directly proven) that cells can be aggregated by cross-linking leukosialin molecules on different cells. The third more direct line of evidence comes from the ability of leukosialin to interact with the cell adhesion molecule, ICAM-1 (Rosenstein et al., 1991). Although this interaction has not yet been shown to support direct cell-cell adhesion, earlier work from the same group indicated that an interaction was occuring between leukosialin on a T hybridoma and a ligand on the target cell and this interaction enhanced IL-2 production by the hybridoma (Park et al., 1991). The ability of mAb binding to leukosialin to augment T cell proliferation also supports a role of the molecule at sites of interaction. These data imply that leukosialin can be involved at sites of cell-cell contact as is also argued for CD45 (see Fig 1.2). How molecules of this type are able to interact at the interface at the same time as small molecules such as CD2 and the TcR remains unclear. Uropods of motile cells have numerous microvilli and it may be the case that, at the same time as weak interactions between these large molecules and their ligands are occurring, small areas of membrane come close enough together for short molecules to interact while the longer molecules are excluded to areas of greater membrane separation. How complete this exclusion from a common area would be is important in interpreting the types of interactions

that can occur between cytoplasmic domains. The large molecules may also serve a role at the interface in preventing the membranes being brought close together, by high affinity interactions between small adhesion molecules, over such a large area as to generate an association of very high avidity that it is not readily reversed. Thus the data presently available support hypothesis (3), with leukosialin acting to reduce the general interactions of (unpolarized) cells but being involved in certain specific, and probably early, adhesion processes to cells expressing ligands (through lectin-carbohydrate or protein-protein interactions). If the cells do not express ligands for other receptors, the interaction is readily reversed as the membranes are still far apart, whereas if other ligands are expressed and they support interactions of higher affinity then these will occur and leukosialin may be excluded depending on the dimensions of these interactions.

At the level of the individual cell, ICAM-1 and leukosialin are co-expressed on a variety of activated leukocytes. In this situation, the leukosialin-ICAM-1 adhesion interaction (Rosenstein et al., 1991) could take place on the same cell and be important in regulating the distribution of one or other of these molecules on the membrane. For example, leukosialin may be linked to the cytoskeleton and be actively transported to the uropod whereas ICAM-1 is not so linked but is drawn to the uropod by the association with leukosialin in its extracellular domain. Whether such an interaction could regulate the availability of ICAM-1 for LFA-1 binding remains to be assessed.

Recent work implicates leukosialin in binding to the collagen-like region of complement component C1q and thus possibly playing a role in C1q-mediated enhancement of phagocytosis by monocytes. A variety of molecules have been argued to have C1q receptor activity, however, and the case for leukosialin does not seem strong as the interaction is disrupted by 100 mM NaCl (Guan et al., 1991).

9.1.6 Conclusions

A clarification of leukosialin's role on the cell membrane should come from the production of leukosialin deficient mice. The construct described in Chapter 4 has at this stage been used successfully for homologous recombination in embryonic stem cells by Mike Carroll. The dominance of leukosialin on T cells and neutrophils as the most heavily glycosylated and one of the more abundant glycoproteins suggests that its deletion will not be readily compensated by other molecules. I predict that the T cells and granulocytes in these animals will show an increased tendency to aggregate and to become involved in non-specific interactions. This may

lead to a reduction in the efficiency of the immune response and an increased frequency of autoimmune diseases.

9.2 General discussion on CD2

9.2.1 Summary

The studies described in Chapter 8 and the work of Paul Driscoll (Driscoll et al., 1991) establish that the first domain of rat CD2 has a V-set IgSF domain fold. The orientation of the core β-strands is highly conserved in comparison to other V domains and most of the structural differences reside in the loops connecting the β-strands. Mutations that affect binding of CD2 to LFA-3 (Peterson and Seed, 1987) are in residues on the face of the domain including strands C, C', C'', F and G. This structure was an important test case for IgSF arguments and it adds significantly to known variants of the Ig-fold since it provides a prototype for IgSF domains that lack the conserved disulphide bond.

9.2.2 Ig type folds and their surfaces of interaction

A large fraction (about 40%) of the molecules identified on the surface of leukocytes and other cell types contain IgSF domains. In most cases it is predicted that these domains will be involved in protein recognition or adhesion events, as reviewed in Williams and Barclay (1988) and Springer (1990). A very similar fold to that of IgSF domains is seen in the cytokine receptor domain and the fibronectin type III domain. The wide deployment of this fold type may reflect its resistance to proteolysis and simultaneous capacity to display a large number of unique residues for specialized interactions. This ability is most clearly demonstrated with antibodies where differences, largely restricted to loops at one end of the domain (connecting strands B-C, C'-C'' and F-G), generate a potential repertoire of more than 10^{10} different specificities. Crystallographic analysis of antibody-antigen complexes has confirmed that contacts are largely restricted to these loops (Davies et al., 1990). TcR recognition of antigen MHC complexes also appears to be through these loops, making contact with both the α-helices of the MHC molecule and outpointing side chains of the bound peptide (Jorgensen et al., 1992).

Structural understanding of adhesion molecule interactions has not progressed to the extent of that for antibody-antigen interactions. Combined structural and mutagenesis analysis of MHC

class I indicate a restricted site of interaction with CD8, with residues in the loop connecting β strands C and D and some residues in strand D being important (Salter et al., 1991). Mutagenesis of CD8 indicates involvement of residues on loops connecting strands B-C and C-C' (Sanders et al., 1991). The structure of a CD8α2 homodimer, determined very recently (Leahy et al., 1992), suggests the interation may also involve residues on the face of the domain made up of strands A, B, E and D. The CD4 binding site for class II, although still controversial, may lie on one face of the molecule including residues in loops connecting strands B-C and F-G of domain 1 and possibly some residues in domain 2 (Fleury et al., 1991).

The binding site on CD2 for LFA-3 may also reside on one face of the molecule, covering a large part of the β sheet including strands C, C', C'', F and G. As suggested by Springer (1991) this would make the interaction crudely resemble that between the V domains in the light and heavy chains of antibodies. Structural features important for the dimerization of IgV domains include conserved β-bulges present in strands C' and G (Colman, 1988). These bulges are present in CD8 and are involved in dimer formation (Leahy et al., 1992) but are absent from CD2 indicating that the contacts made between CD2 and LFA-3 will not be identical to those between immunoglobulin V domains. This is not unexpected as the orientation of the domains in the adhesion interaction is almost certain to differ from that in an antibody and because the CD2 surface, unlike that in an IgV domain, must be stable in a form exposed to water. In addition, the V_H-V_L interaction is unusual compared to other intermolecular interactions between proteins with β-sheets in that most of the contacts are between strands at the edge of the interface (C' and G) rather than involving residues in the central strands (C and F). This type of packing may accommodate more variability which is likely to be important as the complementarity determining regions contribute about one-quarter of the buried surface between Ig V domains (Colman, 1988). This will not be a factor in CD2-LFA-3 packing and it is most likely that this association will be the more classical type with residues over the whole surface contributing to packing interactions. If it is confirmed that CD2 and LFA-3 bind to each other through one face of each V-set domain, then it is likely that cell-membranes would need to come closer together for CD2-LFA-3 adhesion to occur (roughly 12 nm) than would be the case if both molecules interacted through the terminal loops (like an Ig molecule). The result also

implies that one face of CD2 would remain exposed for other possible interactions.

Further demonstration of the diversity of surfaces that can be involved in interactions of Ig type folds is seen in the growth hormone receptor - growth hormone complex. Contacts with growth hormone involve residues in the loops connecting β-strands A to B and E to F of domain 1 of the receptor, some residues on strand G immediately preceding the linker connecting to domain 2, a Glu in the linker and residues in the loop between strands B and C in domain 2 (de Vos et al., 1992). The complex involves two receptor molecules binding a single hormone molecule with the binding site on the hormone for the second receptor being smaller than that for the first. Binding of the second receptor molecule is suggested to occur following binding of the first, and its interaction is stabilized by a dimerization of surfaces in the second domains of the receptors.

A detailed characterization of the CD2-LFA-3 interaction, by mutagenesis and if possible co-crystallization, will be informative with regard to the types of interactions that can occur between adhesion molecules of the IgSF family and also the cytokine receptor and fibronectin type III domain superfamilies. The orientation of molecules in this association is important in considering other interactions that may simultaneously take place such as between CD2 and members of the TcR/CD3 complex. An extension of the observations with the growth hormone receptor- growth hormone complex, may be that concentrating molecules such as CD2 and the TcR/CD3 complex in a small region of membrane stabilizes low affinity associations on the same membrane which place cytoplasmic domains in appropriate orientations for cross-phosphorylation (or de-phosphorylation) events leading to signal transduction. A structural analysis of the orientation of molecules in adhesion pairs should enhance understanding of how groups of molecules can come together. This is illustrated in the structure of the CD8α_2 homodimer, where the site of interaction with class I MHC may be through a surface of one CD8 molecule leaving the other CD8 molecule to interact with a second class I molecule. In the physiological situation where most of the CD8 exists as an $\alpha\beta$ heterodimer it may be the case that β interacts with a molecule other than class I, possibly an adhesion molecule, bringing the two together. This type of information should also give an improved understanding of which associations can occur at a given membrane separation. For example, recent structural

predictions on the membrane proximal region of CD72, the ligand for CD5, indicate the presence of a collagenous stalk that could project the lectin domain more than 14 nm from the lipid bilayer (Beavil et al., 1992). This would place the CD5-CD72 interaction into a different category from that of CD2-LFA-3 in terms of the membrane separation that can be tolerated while allowing adhesion/signaling to occur.

The evidence that CD59 and CD48 are additional ligands for CD2 (Decker et al., 1992; Yokoyama et al., 1991; Mouhtouris et al., 1991) raises questions regarding the possiblility of simultaneous interaction with more than one ligand. The observation that effective antibody mediated signaling through CD2 requires simultaneous binding of two antibodies specific for different sites has lead to speculation that CD2 has two physiological ligands and their simultaneous binding is important for signal transmission. An alternative possibility, as stated earlier, is that a second site on CD2 is involved in binding to molecules on the same cell membrane, such as CD3 molecules (Yang et al., 1986; Ley et al., 1991). It will be important to distinguish between these possibilities particularly because if the former mechanism operates then CD2 signaling may occur in the absence of TcR recognition of peptide in cases where a cell expresses adequate amounts of two CD2 ligands. Simultaneous binding of two ligands also has implications for signaling into the target or antigen presenting cell. There is evidence that signals can be transmitted through all three potential ligands of CD2 by antibody cross-linking (Webb et al., 1990; Hideshima et al., 1990; Yee-Wah Wong pers. comm.). The effect of cross-linking pairs of ligands, for example CD59 with LFA-3, on signaling has not yet been examined.

9.2.3 Modulation of adhesion molecule binding sites by conformational changes

A further point to be addressed in relation to the CD2 structure is the possibility of conformational changes occurring in CD2 and related molecules that can affect interactions with other proteins. This has been suggested for CD2 because of the requirement of pre-activation of the cell, or pre-binding of a mAb to one site on CD2 (for example, region 1; Peterson and Seed, 1987), to allow a number of mAbs specific for another region (region 3, located in the second domain in the case of the 9-1 mAb) to bind. This type of binding site modulation is much more rapid than upregulation of surface expression or down-regulation of surface charge and thus

could be of great significance in relation to early events in T cell and NK cell interactions. Integrins show avidity modulation in the sense that stimulating a cell by various means, or binding of certain mAbs to the extracellular domain of the integrin, enhance cell binding to ligands. Evidence that this modulation is at the level of individual molecules, and thus is a true affinity modulation, comes from the ability of monomeric Fab fragments to induce the change, and from the ability to show increased ligand binding of purified molecules (of GPIIb/IIIa) in solution (Campanero et al., 1990; Du et al., 1991; Kovach et al., 1992). Integrins are heterodimers and this modulation can be imagined as a shift in the orientation of one subunit with respect to the other, exposing the high affinity ligand binding site (Sims et al., 1991). Modulation of binding site conformation is difficult to imagine for monomeric proteins with single transmembrane domains. In addition to CD2, this type of modulation has been argued for L-selectin, and has been suggested for CD44. The studies with L-selectin demonstrated that activation of the cell with cytokines or mAb elevated, within minutes, binding of a polysaccharide (PPME) and this was attributed to an increased affinity (Spertini et al., 1991b). The ligand in this case was polyvalent and the possiblity that the increased binding was the result of an altered distribution of the molecule on the membrane, leading to an increased avidity without invoking the need for conformational changes necessary to explain an increase in affinity, was not eliminated. The study with CD44 showed elevated hyaluronic acid binding following addition of a mAb specific for CD44. Again the ligand was polyvalent and these workers suggest that the increased binding may indeed result from a multimerization effect increasing avidity rather than individual molecules showing elevated affinity (Lesley et al., 1992).

Modulation of conformation following antibody binding has also been suggested for MHC class II molecules, CD14, CD43 and CD44 by the ability of Fab fragments of mAb to these molecules to induce LFA-1 dependent aggregation, as discussed in Chapter 7, and for glycophorin A by the ability of Fab fragments to induce increased membrane rigidity (Chasis et al., 1985). Fab binding is speculated to transmit a signal that activates the cell or that alters associations with the cytoskeleton. However, it was not shown in these studies that the Fab preparations were free of non-covalent dimers, and in most cases the amount of Fab needed to induce an effect was more than an order of magnitude greater than needed with the bivalent

antibody. Thus the possibility that all these observations are a result of molecular cross-linking has not been ruled out.

There are a number of possible explanations for the observed induction of region 3 on CD2 following mAb or SRBC binding to CD2. The simplest explanation is that mAb specific for site 3 are of low affinity and significant binding is only detectable when the antibody binds two sites simultaneously. This may be facilitated by bringing two molecules together with a first antibody or by elevating expression as occurs after activation. Testing the ability of monovalent Fab fragments specific for region 1 to induce region 3 determinants should address this issue. The inability of some mAb to region 1 or 2 to induce region 3 (Peterson and Seed, 1987) suggests the explanation may be more complicated. From a structural view point, the tightly folded nature of CD2 domain 1, and the likelihood that domains 1 and 2 are closely packed, raises questions as to what aspect of conformation could be altered. It appears generally the case that Fab binding to a protein antigen does not induce major conformational changes in the Fab (Davies et al., 1990). Where conformational changes have been seen they are largely restricted to the antigen binding site (Rini et al., 1992). Thus it is unlikely that antibody binding to a site in domain 1 will alter the conformation of a region in domain 2. A more plausible explanation is that the orientation of carbohydrate molecule(s) on the protein surface is shifted by binding of one ligand, exposing a second site of interaction. As described in the studies on leukosialin, a small change in the carbohydrate structure can significantly alter the exposure of a protein surface. Binding of one ligand to CD2 may create a new surface with which the carbohydate interacts with higher affinity, exposing the surface previously covered. The ability of cell activation to elevate region 3 exposure is also difficult to explain from a structural view point as it is unlikely that a signal can be transmitted through a single transmembrane domain leading to altered conformation. A more probable explanation, other than the possible effect of upregulated expression, may be an altered association with a second molecule (such as CD3, OX44 or CD45). This could also be an explanation for the antibody mediated modulation. To my knowledge there is at present no convincing evidence for an altered conformation in the extracellular domain of an adhesion molecule, outside the integrin superfamily, as a result of signals coming from inside the cell or ligand binding to the molecule on the cell surface, that is physiologically relevant. It will be important to assess the concept for the IgSF more directly

when the 2 domain human CD2 structure is available. Cocrystals of CD2 and LFA-3 would be particularly informative.

9.2.4 Glycoyslation and the LFA-3 binding site on CD2

One aspect of the CD2 binding site for LFA-3 which has to this point been ignored is the difference in N-glycosylation sites between rat and human CD2. Rat CD2 has four glycosylation sites and it is likely that all are used to some extent on the basis of the apparent Mr of a soluble form produced in CHO cells (Simon Davis, pers. comm.). Three of the four sites lie in the first domain of rat CD2 and two of these sites fall on the face equivalent to that suggested to be involved in interaction of human CD2 with LFA-3. The single site in domain 1 of human CD2 is on the opposite face, at the start of strand E. In addition the sequence conservation over the region from the start of strand C to the end of C" and from the start of F to the end of G is low (31%). Rat CD2 does not support binding of sheep red blood cells (Chamorro Somoza, pers.comm.). Mouse CD2 also shows a low level of sequence conservation and has 3 potential glycosylation sites in domain 1. These are in different locations to rat CD2, however, one site being near the NH_2-terminus, a second in the same location as the site in human CD2 domain 1, and a third in the loop between strands E and F. This should leave the predicted LFA-3 binding site more exposed than in rat CD2 and there is evidence that mouse CD2 can bind human LFA-3 (Jones et al., 1990). A rodent LFA-3 cDNA clone has not yet been isolated. Okumura and colleagues, in efforts to isolate mouse LFA-3, found CD48 to be a dominant CD2 ligand. The possibility thus exists that a gene duplication event giving rise to LFA-3 in man has not occurred in the rodent lineage and these species have only the CD48 molecule. It will be interesting to see if the site of CD48 interaction with CD2 is conserved between human and rodent CD2.

REFERENCES

Acheson, A., Sunshine, J.L. and Rutishauser, U. (1991) NCAM polysialic acid can regulate both cell-cell and cell-substrate interactions. *J Cell Biol* **114**, 143-53.

Aizawa, S. and Tavassoli, M. (1988) Molecular basis of the recognition of intravenously transplanted hemopoietic cells by bone marrow. *Proc Natl Acad Sci U S A* **85**, 3180-3.

Akiyama, K., Simons, E.R., Bernasconi, P., Schmid, K., van, H.H., Vliegenthart, J.F., Haupt, H. and Schwick, H.G. (1984) The structure of the carbohydrate units of human plasma galactoglycoprotein determined by 500-megahertz 1H NMR spectroscopy. *J. Biol. Chem.* **259**, 7151-4.

Alcover, A., Alberini, C., Acuto, O., Clayton, L.K., Transy, C., Spagnoli, G.C., Moingeon, P., Lopez, P. and Reinherz, E.L. (1988) Interdependence of CD3-Ti and CD2 activation pathways in human T lymphocytes. *EMBO J.* **7**, 1973-7.

Alexander, D.R. and Cantrell, D.A. (1989) Kinases and phosphatases in T-cell activation. *Immunol. Today* **10**, 200-5.

Altevogt, P., Fogel, M., Cheingsong, P.R., Dennis, J., Robinson, P. and Schirrmacher, V. (1983) Different patterns of lectin binding and cell surface sialylation detected on related high- and low-metastatic tumor lines. *Cancer Res* **43**, 5138-44.

Altieri, D.C. and Edgington, T.S. (1988) A monoclonal antibody reacting with distinct adhesion molecules defines a transition in the functional state of the receptor CD11b/CD18 (Mac-1). *J. Immunol.* **141**, 2656-60.

Altman, A., Coggeshall, K.M. and Mustelin, T. (1990) Molecular events mediating T cell activation. *Adv Immunol* **48**, 227-360

Amin, A.R., Tamma, S.M., Oppenheim, J.D., Finkelman, F.D., Kieda, C., Coico, R.F. and Thorbecke, G.J. (1991) Specificity of the murine IgD receptor on T cells is for N-linked glycans on IgD molecules. *Proc Natl Acad Sci U S A* **88**, 9238-42.

Amzel, L.M. and Poljak, R.J. (1979) Three-dimensional structure of immunoglobulins. *Annu Rev Biochem* **48**, 961-97.

Anderson, D., Koch, C.A., Grey, L., Ellis, C., Moran, M.F. and Pawson, T. (1990) Binding of SH2 domains of phospholipase C gamma 1, GAP, and Src to activated growth factor receptors. *Science* **250**, 979-82.

Anderson, R.A. and Lovrien, R.E. (1984) Glycophorin is linked by band 4.1 protein to the human erythrocyte membrane skeleton. *Nature* **307**, 655-8.

Andersson, L.C., Gahmberg, C.G., Kimura, A.K. and Wigzell, H. (1978) Activated human T lymphocytes display new surface glycoproteins. *Proc Natl Acad Sci U S A* **75**, 3455-8.

Andre, P., Benoliel, A.M., Capo, C., Foa, C., Buferne, M., Boyer, C., Schmitt, V.A. and Bongrand, P. (1990) Use of conjugates made between a cytolytic T cell clone and target cells to study the redistribution of membrane molecules in cell contact areas. *J Cell Sci* **97**, 335-47.

Andrews, R.K. and Fox, J.E. (1991) Interaction of purified actin-binding protein with the platelet membrane glycoprotein Ib-IX complex. *J. Biol. Chem.* **266**, 7144-7.

Ardman, B., Sikorski, M.A., Settles, M. and Staunton, D.E. (1990) Human immunodeficiency virus type 1-infected individuals make autoantibodies that bind to CD43 on normal thymic lymphocytes. *J. Exp. Med.* **172**, 1151-8.

Aruffo, A. and Seed, B. (1987) Molecular cloning of two CD7 (T-cell leukemia antigen) cDNAs by a COS cell expression system. *EMBO J.* **6**, 3313-6.

Aruffo, A., Kolanus, W., Walz, G., Fredman, P. and Seed, B. (1991) CD62/P-selection recognition of myeloid and tumor cell sulfatides. *Cell* **67**, 35-44.

Aruffo, A., Stamenkovic, I., Melnick, M., Underhill, C.B. and Seed, B. (1990) CD44 is the principal cell surface receptor for hyaluronate. *Cell* **61**, 1303-13.

Arvieux, J., Willis, A.C. and Williams, A.F. (1986) MRC OX-45 antigen: a leucocyte/endothelium rat membrane glycoprotein of 45,000 molecular weight. *Mol. Immunol.* **23**, 983-90.

Aubrit, F., Gelin, C., Pham, D., Raynal, B. and Bernard, A. (1989) The biochemical characterization of E2, a T cell surface molecule involved in rosettes. *Eur. J. Immunol.* **19**, 1431-6.

Axelsson, B., Hammarstrom, S., Finne, J. and Perlmann, P. (1985) The large sialoglycoprotein of human lymphocytes. II. Biochemical features. *Eur. J. Immunol.* **15**, 427-33.

Axelsson, B. and Perlmann, P. (1989) Persistent superphosphorylation of leukosialin (CD43) in activated T cells and in tumour cell lines. *Scand. J. Immunol.* **30**, 539-47.

Axelsson, B., Youseffi, E.R., Hammarstrom, S. and Perlmann, P. (1988) Induction of aggregation and enhancement of proliferation and IL-2 secretion in human T cells by antibodies to CD43. *J. Immunol.* **141**, 2912-7.

Baecher, C.M., Dorfman, K.S., Mattei, M.G. and Frelinger, J.G. (1990) cDNA cloning and localization of the mouse leukosialin gene (Ly48) to chromosome 7. *Immunogen* **31**, 307-14.

Baecher, C.M., Infante, A.J., Semcheski, K.L. and Frelinger, J.G. (1988) Identification and characterization of a mouse cell surface antigen with alternative molecular forms. *Immunogenetics* **28**, 295-302.

Bansal, A. and Giersach, L.M. (1991) The NPXY internalization signal of the LDL receptor adopts a reverse-turn conformation. *Cell* **67**, 1195-1201.

Barclay, A.N., Jackson, D.I., Willis, A.C. and Williams, A.F. (1987) Lymphocyte specific heterogeneity in the rat leucocyte common antigen (T200) is due to differences in polypeptide sequences near the NH2-terminus. *EMBO J.* **6**, 1259-64.

Bardwell, J.C., McGovern, K. and Beckwith, J. (1991) Identification of a protein required for disulfide bond formation in vivo. *Cell* **67**, 581-9.

Baron, M., Main, A.L., Driscoll, P.C., Mardon, H.J., Boyd, J. and Campbell, I.D. (1992) ^{1}H NMR assignment and secondary structure of the cell adhesion type III module of fibronectin. *Biochemistry* **31**, 2068-2073.

Barrett, T.B., Shu, G. and Clark, E.A. (1991) CD40 signaling activates CD11a/CD18 (LFA-1)-mediated adhesion in B cells. *J. Immunol.* **146**, 1722-9.

Bast, J.E.G, Zhou, L.J., Freeman, G.J., Colley, K.J., Ernst, T.J., Munro, J.M. and Tedder, T.F. (1992) The HB-6, CDw75 and CD76 differentiation antigens are unique cell-surface carbohydrate determinants generated by the β-galactoside α2,6-sialylatransferase. *J. Cell Biol.* **116**, 423-435.

Bazan, J.F. (1990) Structural design and molecular evolution of a cytokine receptor superfamily. *Proc Natl Acad Sci U S A* **87**, 6934-8.

Beavil, A., Edemeades, R.L., Gould, H.J. and Sutton, B.J. (1992) α-Helical coiled-coil stalks in the low-affinity receptor for IgE (FceRII/CD23) and related C-type lectins. *Proc. Natl. Acad. Sci. USA* **89**, 753-757.

Becker, J.W. and Reeke, G.J. (1985) Three-dimensional structure of beta 2-microglobulin. *Proc Natl Acad Sci U S A* **82**, 4225-9.

Bednarczyk, J.L. and McIntyre, B.W. (1990) A monoclonal antibody to VLA-4 alpha-chain (CDw49d) induces homotypic lymphocyte aggregation. *J. Immunol.* **144**, 777-84.

Belitsos, P.C., Hildreth, J.E. and August, J.T. (1990) Homotypic cell aggregation induced by anti-CD44(Pgp-1) monoclonal antibodies and related to CD44(Pgp-1) expression. *J. Immunol.* **144**, 1661-70.

Bell, G.I., Dembo, M. and Bongrand, P. (1984) Cell adhesion: competition between non-specific repulsion and specific binding. *Biophys. J.* **45**, 1051-1064.

Bell, G.M., Seamen, W.E., Niemi, E.C. and Imboden, J.B. (1992) The OX-44 molecule couples to signaling pathways and is associated with CD2 on rat T lymphocytes and a natural killer cell line. *J. Exp. Med.* **175**, 527-536.

Bell, R.M. and Burns, D.J. (1991) Lipid activation of protein kinase C. *J. Biol. Chem.***266**, 4661-4.

Beltzer, J.P. and Spiess, M. (1991) In vitro binding of the asialoglycoprotein receptor to the β adaptin of plasma membrane coated vesicles. *EMBO J.* **10**, 3735-3742.

Bennett, H.S. (1963) Morphological aspects of extracellular polysaccharide. *J. Histochem. Cytochem.* **11**, 14-23.

Bentwich, Z., Douglas, S.D., Skutelsky, E. and Kunkel, H.G. (1973) Sheep red cell binding to human lymphocytes treated with neuraminidase; enhancement of T cell binding and identification of a subpopulation of B cells. *J. Exp. Med.* **137**, 1532-7.

Berg, E.L., Robinson, M.K., Warnock, R.A. and Butcher, E.C. (1991) The human peripheral lymph node vascular addressin is a ligand for LECAM-1, the peripheral lymph node homing receptor. *J Cell Biol* **114**, 343-9.

Bernard, A., Aubrit, F., Raynal, B., Pham, D. and Boumsell, L. (1988) A T cell surface molecule different from CD2 is involved in spontaneous rosette formation with erythrocytes. *J. Immunol.* **140**, 1802-7.

Bernard, A., Gelin, C., Raynal, B., Pham, D., Gosse, C. and Boumsell, L. (1982) Phenomenon of human T cells rosetting with sheep erythrocytes analyzed with monoclonal antibodies. *J. Exp. Med.* **155**, 1317-1333.

Berton, G., Laudanna, C., Sorio, C. and Rossi, F. (1992) Generation of signals activating neutrophil functions by leukocyte integrins: LFA-1 and gp150,95, but not CR3, are able to stimulate the respiratory burst. *J. Cell Biol.* **116**, 1007-1017.

Bettaieb, A., Farace, F., Mitjavila, M.T., Mishal, Z., Dokhelar, M.C., Tursz, T., Breton, G.J., Vainchenker, W. and Kieffer, N. (1988) Use of a monoclonal antibody (GA3) to demonstrate lineage restricted O-glycosylation on leukosialin during terminal erythroid differentiation. *Blood* **71**, 1226-33.

Better, M., Chang, C.P., Robinson, R.R. and Horwitz, A.H. (1988) *Escherichia coli* secretion of an active chimeric antibody fragment. *Science* **240**, 1041-3.

Bevilacqua, M.P., Stengelin, S., Gimbrone, M.J. and Seed, B. (1989) Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils related to complement regulatory proteins and lectins. *Science* **243**, 1160-5.

Beyers, A., Spruyt, L.L. and Williams, A.F. (1992) Molecular associations between the T lymphocyte antigen receptor complex and the surface antigens CD2, CD4, or CD8 and CD5. *Proc. Natl. Acad. Sci. USA* **89**, In press.

Beyers, A.D., Barclay, A.N., Law, D.A., He, Q. and Williams, A.F. (1989) Activation of T lymphocytes via monoclonal antibodies against rat cell surface antigens with particular reference to CD2 antigen. *Immunol. Rev.* **111**, 59-77.

Beyers, A.D., Davis, S.J., Cantrell, D.A., Izquierdo, M. and Williams, A.F. (1991) Autonomous roles for the cytoplasmic domains of the CD2 and CD4 T cell surface antigens. *EMBO J.* **10**, 377-85.

Bhargava, A.K., Woitach, J.T., Davidson, E.A. and Bhavanandan, V.P. (1990) Cloning and cDNA sequence of a bovine submaxillary gland mucin-like protein containing two distinct domains. *Proc Natl Acad Sci U S A* **87**, 6798-802.

Biberfeld, P. and Johansson, A. (1975) Contact areas of cytotoxic lymphocytes and target cells. An electron microscopic study. *Exp Cell Res* **94**, 79-87.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L. and Wiley, D.C. (1987) Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* **329**, 506-12.

Blackwood, E.M. and Eisenman, R.N. (1991) Max: A helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* **251**, 1211-1217.

Boog, C.J., Neefjes, J.J., Boes, J., Ploegh, H.L. and Melief, C.J. (1989) Specific immune responses restored by alteration in carbohydrate chains of surface molecules on antigen-presenting cells. *Eur. J. Immunol.* **19**, 537-42.

Borer, R.A., Lehner, C.F., Eppenberger, H.M. and Nigg, E.A. (1989) Major nucleolar proteins shuttle between nucleus and cytoplasm. *Cell* **56**, 379-90.

Bourbon, H.M. and Amalric, F. (1990) Nucleolin gene organization in rodents: highly conserved sequences within three of the 13 introns. *Gene* **88**, 187-96.

Bourguignon, L.Y.W. and Bourguignon, G.J. (1984) Capping and the cytoskeleton. *Int. Rev. Cytology* **87**, 195-224.

Brain, P., Gordon, J. and Willetts, W.A. (1970) Rosette formation by peripheral lymphocytes. *Clin Exp Immunol* **6**, 681-8.

Bramwell, M.E., Wiseman, G. and Shotton, D.M. (1986) Electron-microscopic studies of the CA antigen, epitectin. *J Cell Sci* **86**, 249-61.

Breitmeyer, J.B., Daley, J.F., Levine, H.B. and Schlossman, S.F. (1987) The T11 (CD2) molecule is functionally linked to the T3/Ti T cell receptor in the majority of T cells. *J. Immunol.* **139**, 2899-905.

Brinster, R.L., Allen, J.M., Behringer, R.R., Gelinas, R.E. and Palmiter, R.D. (1988) Introns increase transcriptional efficiency in transgenic mice. *Proc Natl Acad Sci U S A* **85**, 836-40.

Brodsky, F.M. and Guagliardi, L.E. (1991) The cell biology of antigen processing and presentation. *Annu Rev Immunol* **9**, 707-44.

Brottier, P., Boumsell, L., Gelin, C. and Bernard, A. (1985) T cell activation via CD2 [T, gp50] molecules: accessory cells are required to trigger T cell activation via CD2-D66 plus CD2-9.6/T11(1) epitopes. *J. Immunol.* **135**, 1624-31.

Brown, M.H., Gorman, P.A., Sewell, W.A., Spurr, N.K., Sheer, D. and Crumpton, M.J. (1987) The gene coding for the human T-lymphocyte CD2 antigen is located on chromosome 1p. *Hum. Genet.* **76**, 191-5.

Brown, M.H., Cantrell, D.A., Brattsand, G., Crumpton, M.J. and Gullberg, M. (1989) The CD2 antigen associates with the T-cell antigen receptor CD3 antigen complex on the surface of human T lymphocytes. *Nature* **339**, 551-3.

Brown, W.R., Barclay, A.N., Sunderland, C.A. and Williams, A.F. (1981) Identification of a glycophorin-like molecule at the cell surface of rat thymocytes. *Nature* **289**, 456-60.

Brown, W.R.A. and Williams, A.F. (1982) Lymphocyte cell surface glycoproteins which bind to soybean and peanut lectins. *Immunology* **46**, 713-726.

Bugler, B., Caizergues-Ferrer, M., Bouche, G., Bourbon, H. and Amalric, F. (1982) Detection and localization of a class of proteins immunologicaly related to a 100kDa nucleolar protein. *Eur. J. Biochem.* **128**, 475-480.

Burchell, J., Durbin, H. and Taylor-Papdimitriou, J. (1983) Complexity of expression of antigenic determinants, recognized by monoclonal antibodies HMFG-1 and HMFG-2, in normal and malignant human mammary epithelial cells. *J. Immunol.* **131**, 508-513.

Burgess, K.E., Odysseos, A.D., Zalvan, C., Druker, B.J., Anderson, P., Schlossman, S.F. and Rudd, C.E. (1991) Biochemical identification of a direct physical interaction between the CD4:p56lck and Ti(TcR)/CD3 complexes. *Eur. J. Immunol.* **21**, 1663-8.

Burkly, L.C., Jakubowski, A., Newman, B.M., Rosa, M.D., Chi, R.G. and Lobb, R.R. (1991) Signaling by vascular cell adhesion molecule-1 (VCAM-1) through VLA-4 promotes CD3-dependent T cell proliferation. *Eur. J. Immunol.* **21**, 2871-5.

Burridge, K., Fath, K., Kelly, T., Nuckolls, G. and Turner, C. (1988) Focal adhesions: transmembrane junctions between the extracellular matrix and the cytoskeleton. *Annu Rev Cell Biol* **4**, 487-525.

Butcher, E.C. (1991) Leukocyte-endothelial cell recognition: three (or more) steps to speciticity and diversity. *Cell* **67**, 1033-1036.

Butcher, E.C., Rouse, R.V., Coffman, R.L., Nottenburg, C.N., Hardy, R.R. and Weissman, I.L. (1982) Surface phenotype of peyer's patch germinal centre cells: implications for the role of germinal centres in B cell differentiation. *J. Immunol.* **129**, 2698-2708.

Cabilly, S. (1989) Growth at sub-optimal temperatures allows the production of functional, antigen-binding Fab fragments in Escherichia coli. *Gene* **85**, 553-7.

Camerini, D., Walz, G., Loenen, W.A., Borst, J. and Seed, B. (1991) The T cell activation antigen CD27 is a member of the nerve growth factor/tumor necrosis factor receptor gene family. *J. Immunol.* **147**, 3165-9.

Campanero, M.R., Pulido, R., Alonso, J.L., Pivel, J.P., Pimentel, M.F., Fresno, M. and Sanchez, M.F. (1991) Down-regulation by tumor necrosis factor-alpha of neutrophil cell surface expression of the sialophorin CD43 and the hyaluronate receptor CD44 through a proteolytic mechanism. *Eur. J. Immunol.* **21**, 3045-8.

Campanero, M.R., Pulido, R., Ursa, M.A., Rodriguez-Moya, M., de Landazuri, M. and Sanchez-Madrid, F. (1990) An alternative leukocyte homotypic adhesion mechanism, LFA-1/ICAM-1-independent, triggered through the human VLA-4 integrin. *J Cell Biol* **110**, 2157-65.

Campbell, I.D., Cooke, R.M., Baron, M., Harvey, T.S. and Tappin, M.J. (1989) The solution structures of epidermal growth factor and transforming growth factor alpha. *Prog Growth Factor Res* **1**, 13-22.

Cantrell, D.A., Verbi, W., Davies, A., Parker, P. and Crumpton, M.J. (1988) Evidence that protein kinase C differentially regulates the human T lymphocyte CD2 and CD3 surface antigens. *Eur. J. Immunol.* **18**, 1391-6.

Carlsson, S.R. and Fukuda, M. (1986) Isolation and characterization of leukosialin, a major sialoglycoprotein on human leukocytes. *J. Biol. Chem.* **261**, 12779-86.

Carlsson, S.R., Sasaki, H. and Fukuda, M. (1986) Structural variations of O-linked oligosaccharides present in leukosialin isolated from erythroid, myeloid, and T-lymphoid cell lines. *J. Biol. Chem.* **261**, 12787-95.

Carraway, K.L. and Spielman, J. (1986) Structural and functional aspects of tumor cell sialomucins. *Mol Cell Biochem* **72**, 109-20.

Casanova, J.E., Breitfeld, P.P., Ross, S.A. and Mostov, K.E. (1990) Phosphorylation of the polymeric immunoglobulin receptor required for its efficient transcytosis. *Science* **248**, 742-5.

Casnellie, J.E. and Lamberts, R.J. (1986) Tumor promoters cause changes in the state of phosphorylation and apparent molecular weight of a tyrosine protein kinase in T lymphocytes. *J. Biol. Chem.* **261**, 4921-5.

Chan, P.Y., Lawrence, M.B., Dustin, M.L., Ferguson, L.M., Golan, D.E. and Springer, T.A. (1991) Influence of receptor lateral mobility on adhesion strengthening between membranes containing LFA-3 and CD2. *J Cell Biol* **115**, 245-55.

Chang, H.C., Moingeon, P., Lopez, P., Krasnow, H., Stebbins, C. and Reinherz, E.L. (1989) Dissection of the human CD2 intracellular domain. Identification of a segment required for signal transduction and interleukin 2 production. *J. Exp. Med.* **169**, 2073-83.

Chang, H.C., Moingeon, P., Pedersen, R., Lucich, J., Stebbins, C. and Reinherz, E.L. (1990) Involvement of the PPPGHR motif in T cell activation via CD2. *J. Exp. Med.* **172**, 351-5.

Chasis, J.A., Mohandas, N. and Shohet, S.B. (1985) Erythrocyte membrane rigidity induced by glycophorin A-ligand interaction. *J. Clin Invest* **75**, 1919-26.

Chatila, T.A. and Geha, R.S. (1988) Phosphorylation of T cell membrane proteins by activators of protein kinase C. *J. Immunol.* **140**, 4308-14.

Chen, S., Botteri, F., van der Putten, H., Landel, C.P. and Evans, G.A. (1987) A lymphoproliferative abnormality associated with inappropriate expression of the Thy-1 antigen in transgenic mice. *Cell* **51**, 7-19.

Childs, R.A. and Feizi, T. (1981) Differences in carbohydrate moieties of high molecular weight glycoproteins of human lymphocytes of T and B origins revealed by monoclonal autoantibodies with anti-I and anti-i specificities. *Biochem Biophys Res Commun* **102**, 1158-64.

Chin, D.T., Goff, S.A., Webster, T., Smith, T. and Goldberg, A.L. (1988) Sequence of the *lon* gene in *Escherichia coli*. A heat-shock gene which encodes the ATP-dependent protease La. *J. Biol. Chem.* **263**, 11718-28.

Clark, M.R. and Milstein, C. (1981) Expression of spleen cell immunoglobulin phenotype in hybrids with myeloma cell lines. *Somatic Cell Genet* **7**, 657-66.

Clark, S.J., Jefferies, W.A., Barclay, A.N., Gagnon, J. and Williams, A.F. (1987) Peptide and nucleotide sequences of rat CD4 (W3/25) antigen: evidence for derivation from a structure with four immunoglobulin-related domains. *Proc Natl Acad Sci U S A* **84**, 1649-53.

Clark, S.J., Law, D.A., Paterson, D.J., Puklavec, M. and Williams, A.F. (1988) Activation of rat T lymphocytes by anti-CD2 monoclonal antibodies. *J. Exp. Med.* **167**, 1861-72.

Classon, B.J., Brown, M.H., Garnett, D., Somoza, C., Barclay, A.N., Willis, A.C. and Williams, A.F. (1992) The hinge region of the CD8α chain: structure, antigenicity, and utility in expression of immunoglobulin superfamily domains. *International Immunology* **4**, 215-225.

Clayton, L.K., Sayre, P.H., Novotny, J. and Reinherz, E.L. (1987) Murine and human T11 (CD2) cDNA sequences suggest a common signal transduction mechanism. *Eur. J. Immunol.* **17**, 1367-70.

Clore, G.M. and Gronenborn, A.M. (1991) Structures of larger proteins in solution: three- and four-dimensional heteronuclear NMR spectroscopy. *Science* **252**, 1390-9.

Codington, J.F. and Frim, D.M. (1983) Cell-surface macromolecular and morphological changes related to allotransplantability in the TA3 tumor. *Biomembranes* **11**, 207-58.

Colman, P.M. (1988) Structure of antibody-antigen complexes: implications for immune recognition. *Adv. Immunol.* **43**, 99-132.

Collard, J.G., Schijven, J.F., Bikker, A., La Riviere, G., Bolscher, J.G. and Roos, E. (1986) Cell surface sialic acid and the invasive and metastatic potential of T-cell hybridomas. *Cancer Res* **46**, 3521-7.

Connolly, J.M., Hansen, T.H., Ingold, A.L. and Potter, T.A. (1990) Recognition by CD8 on cytotoxic T lymphocytes is ablated by several substitutions in the class I alpha 3 domain: CD8 and the T-cell receptor recognize the same class I molecule. *Proc Natl Acad Sci U S A* **87**, 2137-41.

Conzelmann, A. and Lefrancois, L. (1988) Monoclonal antibodies specific for T cell-associated carbohydrate determinants react with human blood group antigens CAD and SDA. *J. Exp. Med.* **167**, 119-31.

Conzelmann, A., Pink, R., Acuto, O., Mach, J.P., Dolivo, S. and Nabholz, M. (1980) Presence of T 145 on cytolytic T cell lines and their lectin-resistant mutants. *Eur. J. Immunol.* **10**, 860-8.

Cooke, R.M., Wilkinson, A.J., Baron, M., Pastore, A., Tappin, M.J., Campbell, I.D., Gregory, H. and Sheard, B. (1987) The solution structure of human epidermal growth factor. *Nature* **327**, 339-41.

Cornil, I., Kerbel, R.S. and Dennis, J.W. (1990) Tumor cell surface beta 1-4-linked galactose binds to lectin(s) on microvascular endothelial cells and contributes to organ colonization. *J Cell Biol* **111**, 773-81.

Cowing, C. and Chapdelaine, J.M. (1983) T cells discriminate between Ia antigens expressed on allogeneic accessory cells and B cells: a potential function for carbohydrate side chains on Ia molecules. *Proc Natl Acad Sci U S A* **80**, 6000-4.

Crocker, P.R., Kelm, S., Dubois, C., Martin, B., McWilliam, A.S., Shotton, D.M., Paulson, J.C. and Gordon, S. (1991) Purification and properties of sialoadhesin, a sialic acid-binding receptor of murine tissue macrophages. *EMBO J.* **10**, 1661-9.

Cullen, S.E., Kindle, C.S., Shreffler, D.C. and Cowing, C. (1981) Differential glycosylation of murine B cell and spleen adherent cell Ia antigens. *J. Immunol.* **127**, 1478-84.

Cyster, J., Somoza, C., Killeen, N. and Williams, A.F. (1990) Protein sequence and gene structure for mouse leukosialin (CD43), a T lymphocyte mucin without introns in the coding sequence. *Eur. J. Immunol.* **20**, 875-81.

Cyster, J.G., Shotton, D.M. and Williams, A.F. (1991) The dimensions of the T lymphocyte glycoprotein leukosialin and identification of linear protein epitopes that can be modified by glycosylation. *EMBO J.* **10**, 893-902.

Dalchau, R., Kirkley, J. and Fabre, J.W. (1980) Monoclonal antibody to a human brain-granulocyte-T lymphocyte antigen probably homolgous to the W3/13 antigen of the rat. *Eur. J. Immunol.* **10**, 745-749.

Damle, N.K., Linsley, P.S. and Ledbetter, J.A. (1991) Direct helper T cell-induced B cell differentiation involves interaction between T cell antigen CD28 and B cell activation antigen B7. *Eur. J. Immunol.* **21**, 1277-82.

Dang, L.H. and Rock, K.L. (1991) Stimulation of B lymphocytes through surface Ig receptors induces LFA-1 and ICAM-1-dependent adhesion. *J. Immunol.* **146**, 3273-9.

Dang, N.H., Torimoto, Y., Schlossman, S.F. and Morimoto, C. (1990) Human CD4 helper T cell activation: functional involvement of two distinct collagen receptors, 1F7 and VLA integrin family. *J. Exp. Med.* **172**, 649-52.

Danilov, Y.N. and Juliano, R.L. (1989) Phorbol ester modulation of integrin-mediated cell adhesion: a postreceptor event. *J Cell Biol* **108**, 1925-33.

Davies, D.R., Padlan, E.A. and Sheriff, S. (1990) Antibody-antigen complexes. *Ann. Rev. Biochem.* **59**, 439-473.

Davis, S.J., Brady, R.L., Barclay, A.N., Harlos, K., Dodson, G.G. and Williams, A.F. (1990a) Crystallization of a soluble form of the rat T-cell surface glycoprotein CD4 complexed with Fab from the W3/25 monoclonal antibody. *J Mol Biol* **213**, 7-10.

Davis, S.J., Ward, H.A., Puklavec, M.J., Willis, A.C., Williams, A.F. and Barclay, A.N. (1990b) High level expression in Chinese hamster ovary cells of soluble forms of CD4 T lymphocyte glycoprotein including glycosylation variants. *J. Biol. Chem.* **265**, 10410-8.

Dayhoff, M.O., Barker, W.C. and Hunt, L.T. (1983) Establishing homologies in protein sequences. *Methods Enzymol* **91**, 524-45.

de Fougerolles, A.R. and Springer, T.A. (1992) Intercellular adhesion molecule 3, a third adhesion counter-receptor for lymphocyte function-associated molecule 1 on resting lymphocytes. *J. Exp. Med.* **175**, 185-190.

de Petris, S. (1984) Lectin-binding and spontaneous capping characteristics of the thymocyte glycophorin-like glycoprotein. *Exp Cell Res* **152**, 510-9.

de Petris, S. and Baumgartner, U. (1982) Spontaneous segregation of receptors for Peanut and *Helix pomatia* agglutinins to the uropod region of polarized lymphocytes. *J. Ultrastructure Res.* **80**, 323-338.

de Vos, A.M., Ultsch, M. and Kossiakoff, A.A. (1992) Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* **255**, 306-312.

Deckert, M., Kubar, J. and Bernard, A. (1992) CD58 and CD59 molecules exhibit potentializing effects in T cell

adhesion and activation. *J. Immunol.* **148**, 672-677.

Del Porto, P., Mami, C.F., Bruneau, J.M., Jitsukawa, S., Dumas, J., Harnois, M. and Hercend, T. (1991) TCT.1, a target molecule for gamma/delta T cells, is encoded by an immunoglobulin superfamily gene (Blast-1) located in the CD1 region of human chromosome 1. *J. Exp. Med.* **173**, 1339-44.

Despont, J.P., Abel, C.A. and Grey, H.M. (1975) Sialic acids and sialyltransferases in murine lymphoid cells: indicators of T cell maturation. *Cell. Immunol.* **17**, 487-94.

Diamond, D.J., Clayton, L.K., Sayre, P.H. and Reinherz, E.L. (1988) Exon-intron organization and sequence comparison of human and murine T11 (CD2) genes. *Proc Natl Acad Sci U S A* **85**, 1615-9.

Diamond, M.S., Staunton, D.E., Marlin, S.D. and Springer, T.A. (1991) Binding of the integrin Mac-1 (CD11b/CD18) to the third immunoglobulin-like domain of ICAM-1 (CD54) and its regulation by glycosylation. *Cell* **65**, 961-71.

Dianzani, U., Luqman, M., Rojo, J., Yagi, J., Baron, J.L., Woods, A., Janeway, C.J. and Bottomly, K. (1990) Molecular associations on the T cell surface correlate with immunological memory. *Eur. J. Immunol.* **20**, 2249-57.

Dianzani, U., Shaw, A., Al-ramadi, B.K., Kubo, R.T. and Janeway Jr, C.A. (1992) Physical association of CD4 with the T cell receptor. *J. Immunol.* **148**, 678-688.

Dibb, N.J. and Newman, A.J. (1989) Evidence that introns arose at proto-splice sites. *EMBO J.* **8**, 2015-21.

Dougherty, G.J., Kay, R.J. and Humphries, R.K. (1989) Molecular cloning of 114/A10, a cell surface antigen containing highly conserved repeated elements, which is expressed by murine hemopoietic progenitor cells and interleukin-3-dependent cell lines. *J Biol Chem* **264**, 6509-14

Doyle, C. and Strominger, J.L. (1987) Interaction between CD4 and class II MHC molecules mediates cell adhesion. *Nature* **330**, 256-259.

Dransfield, I., Cabanas, C., Craig, A. and Hogg, N. (1992) Divalent cation regulation of the function of the leukocyte integrin LFA-1. *J. Cell Biol.* **116**, 219-226.

Dransfield, I. and Hogg, N. (1989) Regulated expression of Mg2+ binding epitope on leukocyte integrin alpha subunits. *EMBO J.* **8**, 3759-65.

Drickamer, K. (1988) Two distinct classes of carbohydrate-recognition domains in animal lectins. *J. Biol. Chem.* **263**, 9557-9560.

Driscoll, P.C., Cyster, J.G., Campbell, I.D. and Williams, A.F. (1991) Structure of domain 1 of rat T lymphocyte CD2 antigen. *Nature* **353**, 762-5.

Du, X.P., Plow, E.F., Frelinger, A.3., O'Toole, T.E., Loftus, J.C. and Ginsberg, M.H. (1991) Ligands "activate" integrin alpha IIb beta 3 (platelet GPIIb-IIIa). *Cell* **65**, 409-16.

Durkop, H., Hummel, M., Eitelbach, F., Seed, B. and Stein, H. (1992) Molecular cloning and expression of a new member of the nerve growth factor receptor family that is characteristic for Hodgkin's disease. *Cell* **68**, 421-427.

Dustin, M.L. and Springer, T.A. (1991) Role of lymphocyte adhesion receptors in transient interactions and cell locomotion. *Annu Rev Immunol* **9**, 27-66.

Dustin, M.L. and Springer, T.A. (1989) T-cell receptor cross-linking transiently stimulates adhesiveness through LFA-1. *Nature* **341**, 619-24.

Dustin, M.L., Olive, D. and Springer, T.A. (1989) Correlation of CD2 binding and functional properties of multimeric and monomeric lymphocyte function-associated antigen 3. *J. Exp. Med.* **169**, 503-17.

Dustin, M.L., Sanders, M.E., Shaw, S. and Springer, T.A. (1987) Purified lymphocyte function-associated antigen 3 binds to CD2 and mediates T lymphocyte adhesion. *J. Exp. Med.* **165**, 677-92.

Dyer, M.J. and Hunt, S.V. (1981) Committed T lymphocyte stem cells of rats. Characterization by surface W3/13 antigen and radiosensitivity. *J. Exp. Med.* **154**, 1164-77.

Eberie, W., Sander, C., Klaus, W., Schmidt, B., von Figura, K. and Peters, C. (1991) The essential tyrosine of the internalization signal in lysosomal acid phosphatase is part of a β turn. *Cell*

Eckhardt, A.E., Timpte, C.S., Abernethy, J.L., Zhao, Y. and Hill, R.L. (1991) Porcine submaxillary mucin contains a cystine-rich, carboxyl-terminal domain in addition to a highly repetitive, glycosylated domain. *J. Biol. Chem.* **266**, 9678-86.

Edmundson, A.B., Ely, K.R., Abola, E.E., Schiffer, M.. and Panagiotopoulos, N. (1975) Rotational allomerism and divergent evolution in immunoglobulin light chains. *Biochem.* **14**, 3953-61.

Elbein, A.D. (1991) The role of N-linked oligosaccharides in glycoprotein function. *Trends in Biotechnology* **9**, 346-352.

Ellisen, L.W., Bird, J., West, D.C., Soreng, A.L., Reynolds, T.C., Smith, S.D. and Sklar, J. (1991) TAN-1, the human homolog of the Drosophila notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. *Cell* **66**, 649-61.

Emilie, D., Wallon, C., Galanaud, P., Fischer, A., Olive, D. and Delfraissy, J.F. (1988) Role of the LFA3-CD2 interaction in human specific B cell differentiation. *J. Immunol.***141**, 1912-8.

Escobedo, J.A., Navankasattusas, S., Kavanaugh, W.M., Milfay, D., Fried, V.A. and Williams, L.T. (1991) cDNA cloning of a novel 85 kd protein that has SH2 domains and regulates binding of PI3-kinase to the PDGF beta-receptor. *Cell* **65**, 75-82.

Evans, S.S., Lee, D.B., Han, T., Tomasi, T.B. and Evans, R.L. (1990) Monoclonal antibody to the interferon-inducible protein Leu-13 triggers aggregation and inhibits proliferation of leukemic B cells. *Blood* **76**, 2583-2593.

Eynon, E.E. and Parker, D.C. (1992) Small B cells as antigen-presenting cells in the induction of tolerance to soluble protein antigens. *J. Exp. Med.* **175**, 131-138.

Falk, K., Rotzschke, O., Deres, K., Metzger, J., Jung, G. and Rammensee, H.G. (1991) Identification of naturally processed viral nonapeptides allows their quantification in infected cells and suggests an allele-specific T cell epitope forecast. *J. Exp. Med.* **174**, 425-34.

Farries, T.C. and Atkinson, J.P. (1991) Evolution of the complement system. *Immunol. Today* **12**, 295-300.

Fehon, R.G., Kooh, P.J., Rebay, I., Regan, C.L., Xu, T., Muskavitch, M.A. and Artavanis, T.S. (1990) Molecular interactions between the protein products of the neurogenic loci Notch and Delta, two EGF-homologous genes in Drosophila. *Cell* **61**, 523-34.

Ferguson, M.A. and Williams, A.F. (1988) Cell-surface anchoring of proteins via glycosyl-phosphatidylinositol structures. *Annu Rev Biochem* **57**, 285-320.

Ferrero, E., Hsieh, C.L., Francke, U. and Goyert, S.M. (1990) CD14 is a member of the family of leucine-rich proteins and is encoded by a gene syntenic with multiple receptor genes. *J. Immunol.* **145**, 331-6.

Ferrick, D.A., Sambhara, S.R., Ballhausen, W., Iwamoto, A., Pircher, H., Walker, C.L., Yokoyama, W.M., Miller, R.G. and Mak, T.W. (1989) T cell function and expression are dramatically altered in T cell receptor V gamma 1.1J gamma 4C gamma 4 transgenic mice. *Cell* **57**, 483-92.

Fiete, D., Srivastava, V., Hindsgaul, O. and Baenziger, J.U. (1991) A hepatic reticuloendothelial cell receptor specific for SO4-4GalNAcβ1,4GlcNAcβ1,2Manα that mediates rapid clearance of lutropin. *Cell* **67**, 1103-1110.

Fleury, S., Lamarre, D., Meloche, S., Ryu, S.E., Cantin, C., Hendrickson, W.A. and Sekaly, R.P. (1991) Mutational analysis of the interaction between CD4 and class II MHC: class II antigens contact CD4 on a

surface opposite the gp120-binding site. *Cell* **66**, 1037-49.

Florman, H.M. and Wassarman, P.M. (1985) O-linked oligosaccharides of mouse egg ZP3 account for its sperm receptor activity. *Cell* **41**, 313-24.

Fogel, M., Altevogt, P. and Schirrmacher, V. (1983) Metastatic potential severely altered by changes in tumor cell adhesiveness and cell-surface sialylation. *J. Exp. Med.* **157**, 371-6.

Fox, R.I., Hueniken, M., Fong, S., Behar, S., Royston, I., Singhal, S.K. and Thompson, L. (1983) A novel cell surface antigen (T305) found in increased frequency on acute leukemia cells and in autoimmune disease states. *J. Immunol.* **131**, 762-7.

Fraser, J.D., Irving, B.A., Crabtree, G.R. and Weiss, A. (1991) Regulation of interleukin-2 gene enhancer activity by the T cell accessory molecule CD28. *Science* **251**, 313-6.

Freeman, M., Ashkenas, J., Rees, D.J., Kingsley, D.M., Copeland, N.G., Jenkins, N.A. and Krieger, M. (1990) An ancient, highly conserved family of cysteine-rich protein domains revealed by cloning type I and type II murine macrophage scavenger receptors. *Proc Natl Acad Sci U S A* **87**, 8810-4.

Fukuda, M. and Carlsson, S.R. (1986) Leukosialin, a major sialoglycoprotein on human leukocytes as differentiation antigens. *Med Biol* **64**, 335-43.

Fukuda, M., Carlsson, S.R., Klock, J.C. and Dell, A. (1986) Structures of O-linked oligosaccharides isolated from normal granulocytes, chronic myelogenous leukemia cells, and acute myelogenous leukemia cells. *J. Biol. Chem.* **261**, 12796-806.

Fukuda, M., Koeffler, H.P. and Minowada, J. (1981) Membrane differentiation in human myeloid cells: expression of unique profiles of cell surface glycoproteins in myeloid leukemic cell lines blocked at different stages of differentiation and maturation. *Proc Natl Acad Sci U S A* **78**, 6299-303.

Gahmberg, C.G. and Andersson, L.C. (1982) Role of sialic acid in the mobility of membrane proteins containing O-linked oligosaccharides on polyacrylamide gel electrophoresis in sodium dodecyl sulfate. *Eur. J. Biochem* **122**, 581-6.

Gahmberg, C.G. and Andersson, L.C. (1977) Selective radioactive labeling of cell surface sialoglycoproteins by periodate-tritiated borohydride. *J. Biol. Chem.* **252**, 5888-94.

Gahmberg, C.G., Autero, M. and Hermonen, J. (1988) Major O-glycosylated sialoglycoproteins of human hematopoietic cells: differentiation antigens with poorly understood functions. *J Cell Biochem* **37**, 91-105.

Galfre, G., Milstein, C. and Wright, B. (1979) Rat x rat hybrid myelomas and a monoclonal anti-Fd portion of mouse IgG. *Nature* **277**, 131-3.

Gelin, C., Aubrit, F., Phalipon, A., Raynal, B., Cole, S., Kaczorek, M. and Bernard, A. (1989) The E2 antigen, a 32 kd glycoprotein involved in T-cell adhesion processes, is the MIC2 gene product. *EMBO J.* **8**, 3253-9.

Gendler, S.J., Burchell, J.M., Duhig, T., Lamport, D., White, R., Parker, M. and Taylor-Papadimitriou, J. (1987) Cloning of partial cDNA encoding differentiation and tumor-associated mucin glycoproteins expressed by human mammary epithelium. *Proc Natl Acad Sci U S A* **84**, 6060-4.

Gendler, S.J., Lancaster, C.A., Taylor, P.J., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E.N. and Wilson, D. (1990) Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. *J. Biol. Chem.* **265**, 15286-93.

Geng, J.G., Bevilacqua, M.P., Moore, K.L., McIntyre, T.M., Prescott, S.M., Kim, J.M., Bliss, G.A., Zimmerman, G.A. and McEver, R.P. (1990) Rapid neutrophil adhesion to activated endothelium mediated by GMP-140. *Nature* **343**, 757-60.

Gerken, T.A., Butenhof, K.J. and Shogren, R. (1989) Effects of glycosylation on the conformation and dynamics of O-linked glycoproteins: carbon-13 NMR studies of ovine submaxillary mucin. *Biochemistry* **28**, 5536-43.

Gerlinger, P., LeMeur, M., Irrmann, C., Renard, P., Wasylyk, C. and Wasylyk, B. (1986) B-lymphocyte

targeting of gene expression in transgenic mice with the immunoglobulin heavy-chain enhancer. *Nucleic Acids Res.* **14**, 6565-77.

Gerard, C., Eddy, R.J. and Shows, T.B. (1990) The core polypeptide of cystic fibrosis tracheal mucin contains a tandem repeat structure. Evidence for a common mucin in airway and gastrointestinal tissue. *J. Cli. Invest.* **86**, 1921-7.

Gilbert, C.W., Zaroukian, M.H. and Esselman, W.J. (1988) Poly-N-acetyllactosamine structures on murine cell surface T200 glycoprotein participate in natural killer cell binding to YAC-1 targets. *J. Immunol.* **140**, 2821-8.

Gilbert, W., Marchionni, M. and McKnight, G. (1986) On the antiquity of introns. *Cell* **46**, 151-3.

Glockshuber, R., Malia, M., Pfitzinger, I. and Pluckthun, A. (1990) A comparison of strategies to stabilize immunoglobulin Fv-fragments. *Biochemistry* **29**, 1362-7.

Gluzman, Y. (1981) SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23**, 175-182.

Goff, S.A. and Goldberg, A.L. (1985) Production of abnormal proteins in E. coli stimulates transcription of lon and other heat shock genes. *Cell* **41**, 587-95.

Goldstein, L.A., Zhou, D.F., Picker, L.J., Minty, C.N., Bargatze, R.F., Ding, J.F. and Butcher, E.C. (1989) A human lymphocyte homing receptor, the hermes antigen, is related to cartilage proteoglycan core and link proteins. *Cell* **56**, 1063-72.

Gooley, A.A., Classon, B.J., Marschalek, R. and Williams, K.L. (1991) Glycosylation sites identified by detection of glycosylated amino acids released from Edman degradation: the identification of Xaa-Pro-Xaa-Xaa as a motif for Thr-O-glycosylation. *Biochem Biophys Res Commun* **178**, 1194-201.

Graves, J.D., Downward, J., Rayter, S., Warne, P., Tutt, A.L., Glennie, M. and Cantrell, D.A. (1991) CD2 antigen mediated activation of the guanine nucleotide binding proteins p21ras in human T lymphocytes. *J. Immunol.* **146**, 3709-12.

Gray, F., Cyster, J.G., Willis, A.C. and Williams, A.F. (1992) Analysis of the structure of domain I of the CD2 T lymphocyte antigen by site-directed mutagenesis to form a disulphide bond. In preparation.

Grinnell, B.W., Walls, J.D. and Gerlitz, B. (1991) Glycosylation of human protein C affects its secretion, processing, functional activities, and activation by thrombin. *J. Biol. Chem.* **266**, 9778-85.

Grodberg, J. and Dunn, J.J. (1988) ompT encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *J Bacteriol* **170**, 1245-53.

Groux, H., Huet, S., Aubrit, F., Tran, H.C., Boumsell, L. and Bernard, A. (1989) A 19-kDa human erythrocyte molecule H19 is involved in rosettes, present on nucleated cells, and required for T cell activation. Comparison of the roles of H19 and LFA-3 molecules in T cell activation. *J. Immunol.* **142**, 3013-20.

Gschwendt, M., Kittstein, W. and Marks, F. (1991) Protein kinase C activation by phorbol esters: do cysteine-rich regions and pseudosubstrate motifs play a role? *Trends Biochem. Sci.* **16**, 167-9.

Guan, E., Burgess, W.H., Robinson, S.L., Goodman, E.B., McTigue, K.J. and Tenner, A.J. (1991) Phagocytic cell molecules that bind the collagen-like region of C1q. *J. Biol. Chem.* **266**, 20345-20355.

Gulino, D., Ryckewaert, J.J., Andrieux, A., Rabiet, M.J. and Marguerie, G. (1990) Identification of a monoclonal antibody against platelet GPIIb that interacts with a calcium-binding site and induces aggregation. *J. Biol. Chem.* **265**, 9575-81.

Gulley, M.L., Ogata, L.C., Thorson, J.A., Dailey, M.O. and Kemp, J.D. (1988) Identification of a murine pan-T cell antigen which is also expressed during the terminal phases of B cell differentiation. *J. Immunol.* **140**, 3751-7.

Gum, J.R., Byrd, J.C., Hicks, J.W., Toribara, N.W., Lamport, D.T. and Kim, Y.S. (1989) Molecular cloning of human intestinal mucin cDNAs. Sequence analysis and evidence for genetic polymorphism. *J. Biol. Chem.* **264**, 6480-7.

Gum, J.R., Hicks, J.W., Swallow, D.M., Lagace, R.L., Byrd, J.C., Lamport, D.T., Siddiki, B. and Kim, Y.S. (1990) Molecular cloning of cDNAs derived from a novel human intestinal mucin gene. *Biochem Biophys Res Commun* **171**, 407-15.

Gunthert, U., Hofmann, M., Rudy, W., Reber, S., Zoller, M., Haussmann, I., Matzku, S., Wenzel, A., Ponta, H. and Herrlich, P. (1991) A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma cells. *Cell* **65**, 13-24.

Hanahan, D. (1989) Transgenic mice as probes into complex systems. *Science* **246**, 1265-75.

Harding, C.V., Leyva, C.F. and Unanue, E.R. (1988) Mechanisms of antigen processing. *Immunol. Rev.* **106**, 77-92.

Harding, S.E. (1989) The macrostructure of mucus glycoproteins in solution. *Adv Carbohydr Chem Biochem* **47**, 345-81.

Hardy, R.R., Carmack, C.E., Shinton, S.A., Kemp, J.D. and Hayakawa, K. (1991) Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *J. Exp. Med.* **173**, 1213-25.

Hareuveni, M., Tsarfaty, I., Zaretsky, J., Kotkes, P., Horev, J., Zrihan, S., Weiss, M., Green, S., Lathe, R., Keydar, I. and Wreschner, D.H. (1990) A transcribed gene containing a variable number of tandem repeats codes for a human epithelial tumor antigen. *Eur. J. Biochem.* **189**, 475-486.

Hartley, D.L. and Kane, J.F. (1988) Properties of inclusion bodies from recombinant Escherichia coli. *Biochem. Soc. Trans.* **16**, 101-2.

Haynes, B.F., Telen, M.J., Hale, L.P. and Denning, S.M. (1989) CD44--a molecule involved in leukocyte adherence and T-cell activation [published erratum appears in Immunol Today 1990 Mar;11(3):80]. *Immunol. Today* **10**, 423-8.

He, Q., Beyers, A.D., Barclay, A.N. and Williams, A.F. (1988) A role in transmembrane signaling for the cytoplasmic domain of the CD2 T lymphocyte surface antigen. *Cell* **54**, 979-84.

Hermanowski-Vosatka, A., Van Strijp, J.A.G., Swiggard, W.J. and Wright, S.D. (1992) Integrin modulating factor-1: A lipid that alters the function of leukocyte integrins. *Cell* **68**, 341-352.

Herrenknecht, K., Ozawa, M., Eckerskorn, C., Lottspeich, F., Lenter, M. and Kemler, R. (1991) The uvomorulin-anchorage protein alpha catenin is a vinculin homologue. *Proc Natl Acad Sci U S A* **88**, 9156-60.

Hession, C., Osborn, L., Goff, D., Chi, R.G., Vassallo, C., Pasek, M., Pittack, C., Tizard, R., Goelz, S., McCarthy, K. and et, a.l. (1990) Endothelial leukocyte adhesion molecule 1: direct expression cloning and functional interactions. *Proc Natl Acad Sci U S A* **87**, 1673-7.

Hession, C., Tizard, R., Vassallo, C., Schiffer, S.B., Goff, D., Moy, P., Chi, R.G., Luhowskyj, S., Lobb, R. and Osborn, L. (1991) Cloning of an alternate form of vascular cell adhesion molecule-1 (VCAM1). *J. Biol. Chem.* **266**, 6682-5.

Hibbs, M.L., Jakes, S., Stacker, S.A., Wallace, R.W. and Springer, T.A. (1991) The cytoplasmic domain of the integrin lymphocyte function-associated antigen 1 beta subunit: sites required for binding to intercellular adhesion molecule 1 and the phorbol ester-stimulated phosphorylation site. *J. Exp. Med.* **174**, 1227-38.

Hideshima, T., Okada, T.H. and Okada, H. (1990) Expression of HRF20, a regulatory molecule of complement on peripheral blood mononuclear cells. *Immunology* **69**, 396-401.

Higgins, E.A., Siminovitch, K.A., Zhuang, D.L., Brockhausen, I. and Dennis, J.W. (1991) Aberrant O-linked oligosaccharide biosynthesis in lymphocytes and platelets from patients with the Wiskott-Aldrich syndrome. *J. Biol. Chem.* **266**, 6280-90.

Hilkens, J. (1988) Biochemistry and function of mucins in malignant disease. *Cancer Rev.* **11-12**, 25-54.

Hoffman, W. (1988) A new repetitive protein from Xenopus laevis skin highly homologous to pancreatic spasmolytic polypeptide. *J Biol Chem* **263**, 7686-90.

Holmgren, A. and Branden, C.I. (1989) Crystal structure of chaperone protein PapD reveals an immunoglobulin fold. *Nature* **342**, 248-51.

Holness, C.L., DaSilva, R. and Simmons, D.L. (1992) In preparation

Horwitz, A., Duggan, K., Buck, C., Beckerle, M.C. and Burridge, K. (1986) Interaction of plasma membrane fibronectin receptor with talin-a transmembrane linkage. *Nature* **320**, 531-3.

House, C., Wettenhall, R.E. and Kemp, B.E. (1987) The influence of basic residues on the substrate specificity of protein kinase C. *J. Biol. Chem.* **262**, 772-7.

Howard, F.D., Ledbetter, J.A., Wong, J., Bieber, C.P., Stinson, E.B. and Herzenberg, L.A. (1981) A human T lymphocyte differentiation marker defined by monoclonal antibodies that block E-rosette formation. *J. Immunol.* **126**, 2117-22.

Huber, A.R., Kunkel, S.L., Todd, R.3. and Weiss, S.J. (1991) Regulation of transendothelial neutrophil migration by endogenous interleukin-8. *Science* **254**, 99-102.

Humphries, M.J. (1990) The molecular basis and specificity of integrin-ligand interactions. *J. Cell Science* **97**, 585-592.

Hunig, T. (1985) The cell surface molecule recognized by the erythrocyte receptor of T lymphocytes. Identification and partial characterization using a monoclonal antibody. *J. Exp. Med.* **162**, 890-901.

Hynes, R.O. and Lander, A.D. (1992) Contact and adhesive specificities in the associations, migrations and targeting of cells and axons. *Cell* **66**, 303-322.

Imai, Y., True, D.D., Singer, M.S. and Rosen, S.D. (1990) Direct demonstration of the lectin activity of gp90MEL, a lymphocyte homing receptor. *J Cell Biol* **111**, 1225-32.

Imler, J.L., Lemaire, C., Wasylyk, C. and Wasylyk, B. (1987) Negative regulation contributes to tissue specificity of the immunoglobulin heavy-chain enhancer. *Mol. Cell. Biol.* **7**, 2558-67.

Irving, B.A. and Weiss, A. (1991) The cytoplasmic domain of the T cell receptor zeta chain is sufficient to couple to receptor-associated signal transduction pathways. *Cell* **64**, 891-901.

Isobe, K. and Nakashima, I. (1991) Homotypic aggregation of murine T lymphocytes induced by anti-Thy-1 monoclonal antibodies. *Immunology* **73**, 159-64.

Jackman, R.W., Beeler, D.L., Fritze, L., Soff, G. and Rosenberg, R.D. (1987) Human thrombomodulin gene is intron depleted: nucleic acid sequences of the cDNA and gene predict protein structure and suggest sites of regulatory control. *Proc Natl Acad Sci U S A* **84**, 6425-9.

Jackson, D.G., Buckley, J. and Bell, J.I. (1992a) Multiple variants of the human lymphocyte homing receptor CD44 generated by insertions at a single site in the extracellular domain. *J. Biol. Chem.* In press.

Jackson, D.G., Hart, D.N.J., Starling, G. and Bell, J.I. (1992b) Molecular cloning of a novel member of the immunoglobulin gene superfamily homologous to the polymeric immunoglobulin receptor. *Eur. J. Immunol* **267**, 4732-39.

Jackson, D.I. and Barclay, A.N. (1989) The extra segments of sequence in rat leucocyte common antigen (L-CA) are derived by alternative splicing of only three exons and show extensive O-linked glycosylation. *Immunogenetics* **29**, 281-7.

Jaffe, S.H., Friedlander, D.R., Matsuzaki, F., Crossin, K.L., Cunningham, B.A. and Edelman, G.M. (1990) Differential effects of the cytoplasmic domains of cell adhesion molecules on cell aggregation and sorting-out.

Proc Natl Acad Sci U S A **87**, 3589-93.

Jefferies, W.A., Green, J.R. and Williams, A.F. (1985) Authentic T helper CD4 (W3/25) antigen on rat peritoneal macrophages. *J. Exp. Med.* **162**, 117-27.

Jentoft, N. (1990) Why are proteins O-glycosylated? *Trends Biochem. Sci.* **15**, 291-4.

Jiang, S.B., Ojcius, D.M., Persechini, P.M. and Young, J.D. (1990) Resistance of cytolytic lymphocytes to perforin-mediated killing. Inhibition of perforin binding activity by surface membrane proteins. *J. Immunol.* **144**, 998-1003.

Jondal, M., Holm, G. and Wigzell, H. (1972) Surface markers on human T and B lymphocytes. I. A large population of lymphocytes forming nonimmune rosettes with sheep red blood cells. *J. Exp. Med.* **136**, 207-15.

Jones, W.J., Sen, J. and Burakoff, S.J. (1990) *Year Immunol* **6**, 95-111.

Jorgensen, J.L., Esser, U., de St. Goth, B.F., Reay, P.A. and Davis, M.M. (1992) Mapping T-cell receptor-peptide contacts by variant peptide immunization of single chain transgenics. *Nature* **355**, 224-230.

Julius, M.H., Simpson, E. and Herzenberg, L.A. (1973) A rapid method for the isolation of functional thymus-derived murine lymphocytes. *Eur. J. Immunol.* **3**, 645-9.

June, C.H., Ledbetter, J.A., Lindsten, T. and Thompson, C.B. (1989) Evidence for the involvement of three distinct signals in the induction of IL-2 gene expression in human T lymphocytes. *J. Immunol.* **143**, 153-61.

June, C.H., Ledbetter, J.A., Linsley, P.S. and Thompson, C.B. (1990) Role of the CD28 receptor in T-cell activation. *Immunol. Today* **11**, 211-6.

Jung, T.M., Gallatin, W.M., Weissman, I.L. and Dailey, M.O. (1988) Down-regulation of homing receptors after T cell activation. *J. Immunol.* **141**, 4110-7.

Kalli, K.R., Hsu, P.H., Bartow, T.J., Ahearn, J.M., Matsumoto, A.K., Klickstein, L.B. and Fearon, D.T. (1991) Mapping of the C3b-binding site of CR1 and construction of a (CR1)2-F(ab')2 chimeric complement inhibitor. *J. Exp. Med.* **174**, 1451-60.

Kamitani, S., Akiyama, Y. and Ito, K. (1992) Identification and characterization of an Escherichia coli gene required for the formation of correctly folded alkaline phosphatase, a periplasmic enzyme. *EMBO J.* **11**, 57-62.

Kamoun, M., Martin, P.J., Hansen, J.A., Brown, M.A., Siadak, A.W. and Nowinski, R.C. (1981) Identification of a human T lymphocyte surface protein associated with the E-rosette receptor. *J. Exp. Med.* **153**, 207-12.

Kamps, W.A., Timens, W., De, B.G., Spanjer, H.H. and Poppema, S. (1989) In situ study of haemopoiesis in human fetal liver. *Scand. J. Immunol.* **30**, 399-408.

Kansas, G.S. and Tedder, T.F. (1991) Transmembrane signals generated through MHC class II, CD19, CD20, CD39 and CD40 antigens induce LFA-1 dependent and independent adhesion in human B cells through a tyrosine kinase-dependent pathway. *J. Immunol.* **147**, 4094-4102.

Kapsenberg, M.L., Stiekema, F.E., Kallan, A., Bos, J.D. and Roozemond, R.C. (1989) The restrictive role of sialic acid in antigen presentation to a subset of human peripheral CD4+ T lymphocytes that requires antigen-presenting dendritic cells. *Eur. J. Immunol.* **19**, 1829-34.

Kaufman, Y., Berke, G. and Eshhar, Z. (1981) Cytotoxic T lymphocyte hybridomas that mediate specific tumor-cell lysis in vitro. *Proc Natl Acad Sci U S A* **78**, 2502-506.

Kay, R., Takei, F. and Humphries, R.K. (1990) Expression cloning of a cDNA encoding M1/69-J11d heat-stable antigens. *J. Immunol.* **145**, 1952-9.

Kay, R., Rosten, P.M. and Humphries, R.K. (1991) CD24, a signal transducer modulating B cell activation responses, is a very short peptide with a glycosyl phosphatidylinositol membrane anchor. *J. Immunol.* **147**, 1412-6.

Kearse, K.P., Cassatt, D.R., Kaplan, A.M. and Cohen, D.A. (1988) The requirement for surface Ig signaling as a prerequisite for T cell:B cell interactions. A possible role for desialylation. *J. Immunol.* **140**, 1770-8.

Keizer, G.D., Visser, W., Vliem, M. and Figdor, C.G. (1988) A monoclonal antibody (NKI-L16) directed against a unique epitope on the alpha-chain of human leukocyte function-associated antigen 1 induces homotypic cell-cell interactions. *J. Immunol.* **140**, 1393-400.

Kemp, B.E. and Pearson, R.B. (1990) Protein kinase recognition sequence motifs. *Trends Biochem. Sci.* **15**, 342-6.

Kennelly, P.J. and Krebs, E.G. (1991) Consensus sequences as substrate specificity determinants for protein kinases and protein phosphatases. *J. Biol. Chem.* **266**, 15555-8.

Kenney, D., Cairns, L., Remold, O.E., Peterson, J., Rosen, F.S. and Parkman, R. (1986) Morphological abnormalities in the lymphocytes of patients with the Wiskott-Aldrich syndrome. *Blood* **68**, 1329-32.

Killeen, N., Barclay, A.N., Willis, A.C. and Williams, A.F. (1987) The sequence of rat leukosialin (W3/13 antigen) reveals a molecule with O-linked glycosylation of one third of its extracellular amino acids. *EMBO J.* **6**, 4029-34.

Killeen, N., Moessner, R., Arvieux, J., Willis, A. and Williams, A.F. (1988) The MRC OX-45 antigen of rat leukocytes and endothelium is in a subset of the immunoglobulin superfamily with CD2, LFA-3 and carcinoembryonic antigens. *EMBO J.* **7**, 3087-91.

Kim, Y.D., Robinson, D.Y., Manderino, G.L., Tribby, I.I. and Tomita, J.T. (1989) Molecular characterization of the epitope in prostate and breast tumor-associated PR92 antigen. *Cancer Res* **49**, 2379-82.

Kimura, A., Wigzell, H., Holmquist, G., Ersson, B. and Carlsson, P. (1979) Selective affinity fractionation of murine cytotoxic T lymphocytes (CTL). *J. Exp. Med.* **149**, 473-484.

Kimura, A.K. and Wigzell, H. (1978) Cell surface glycoproteins of murine cytotoxic T lymphocytes. I. T 145, a new cell surface glycoprotein selectively expressed on Ly 1-2+ cytotoxic T lymphocytes. *J. Exp. Med.* **147**, 1418-34.

King, P.D. and Katz, D.R. (1990) Mechanisms of dendritic cell function. *Immunol. Today* **11**, 206-11.

Kitajima, K., Inoue, Y. and Inoue, S. (1986) Polysialoglycoproteins of Salmonidae fish eggs. Complete structure of 200-kDa polysialoglycoprotein from the unfertilized eggs of rainbow trout (*Salmo gairdneri*). *J. Biol. Chem.* **261**, 5262-9.

Klausner, R.D. and Samelson, L.E. (1991) T cell antigen receptor activation pathways: the tyrosine kinase connection. *Cell* **64**, 875-8.

Kobilka, B.K., Frielle, T., Collins, S., Yang, F.T., Kobilka, T.S., Francke, U., Lefkowitz, R.J. and Caron, M.G. (1987) An intronless gene encoding a potential member of the family of receptors coupled to guanine nucleotide regulatory proteins. *Nature* **329**, 75-9.

Koopman, G., Parmentier, H.K., Schuurman, H.J., Newman, W., Meijer, C.J. and Pals, S.T. (1991) Adhesion of human B cells to follicular dendritic cells involves both the lymphocyte function-associated antigen 1/intercellular adhesion molecule 1 and very late antigen 4/vascular cell adhesion molecule 1 pathways. *J. Exp. Med.* **173**, 1297-304.

Koopman, G., van Kooyk, Y., de Graaff, M., Meyer, C.J., Figdor, C.G. and Pals, S.T. (1990) Triggering of the CD44 antigen on T lymphocytes promotes T cell adhesion through the LFA-1 pathway. *J. Immunol.* **145**, 3589-93.

Koretzky, G.A., Picus, J., Schultz, T. and Weiss, A. (1991) Tyrosine phosphatase CD45 is required for T-cell antigen receptor and CD2-mediated activation of a protein tyrosine kinase and interleukin 2 production. *Proc Natl Acad Sci U S A* **88**, 2037-41.

Korinek, V., Stefanova, I., Angelisova, P., Hilgert, I. and Horejsi, V. (1991) The human leucocyte antigen

CD48 (MEM-102) is closely related to the activation marker Blast-1. *Immunogenetics* **33**, 108-12.

Kornfeld, R. and Kornfeld, S. (1985) Assembly of asparagine-linked oligosaccharides. ***Annu Rev Biochem*** **54**, 631-64.

Kovach, N.L. Carlos, T.M., Yee, E. and Harlan, J.M. (1992) A monoclonal antibody to b1 integrin (CD29) stimulates VLA-dependent adherence of leukocytes to human umbilical vein endothelial cells and matrix components. *J. Cell Biol* **116**, 499-509

Koyasu, S., Lawton, T., Novick, D., Recny, M.A., Siliciano, R.F., Wallner, B.P. and Reinherz, E.L. (1990) Role of interaction of CD2 molecules with lymphocyte function-associated antigen 3 in T-cell recognition of nominal antigen. *Proc Natl Acad Sci U S A* **87**, 2603-7.

Kozak, M. (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15**, 8125-48.

Kozarsky, K., Kingsley, D. and Krieger, M. (1988) Use of a mutant cell line to study the kinetics and function of O-linked glycosylation of low density lipoprotein receptors. *Proc Natl Acad U S A* **85**, 4335-39

Kubota, H., Okazaki, H., Onuma, M., Kano, S., Hattori, M. and Minato, N. (1990) Identification and gene cloning of a new phosphatidylinositol-linked antigen expressed on mature lymphocytes. Down-regulation by lymphocyte activation. *J. Immunol.* **145**, 3924-31.

Kudo, S. and Fukuda, M. (1991) A short, novel promoter sequence confers the expression of human leukosialin, a major sialoglycoprotein on leukocytes. *J. Biol. Chem.* **266**, 8483-9.

Kuijpers, T.W., Hakkert, B.C., Hoogerwerf, M., Leeuwenberg, J.F. and Roos, D. (1991) Role of endothelial leukocyte adhesion molecule-1 and platelet-activating factor in neutrophil adherence to IL-1-prestimulated endothelial cells. Endothelial leukocyte adhesion molecule-1-mediated CD18 activation. *J. Immunol.* **147**, 1369-76.

Kupfer, A. and Singer, S.J. (1989a) Cell biology of cytotoxic and helper T cell functions: immunofluorescence microscopic studies of single cells and cell couples. *Annu Rev Immunol* **7**, 309-37.

Kupfer, A. and Singer, S.J. (1989b) The specific interaction of helper T cells and antigen presenting B cells. *J Exp Med* **170**, 1697-1713.

Kuwana, Y., Itoh, S., Nagase, F., Nakashima, I. and Kurosawa, Y. (1987) Production of the constant domain of murine T-cell receptor beta-chain in Escherichia coli. *Febs Lett* **219**, 360-4.

Kyewski, B.A., Jenkinson, E.J., Kingston, R., Altevogt, P., Owen, M.J. and Owen, J.J. (1989) The effects of anti-CD2 antibodies on the differentiation of mouse thymocytes. *Eur. J. Immunol.* **19**, 951-4.

Labeit, S., Barlow, D.P., Gautel, M., Gibson, T., Holt, J., Hsieh, C.L., Francke, U., Leonard, K., Wardale, J., Whiting, A. and et, a.l. (1990) A regular pattern of two types of 100-residue motif in the sequence of titin. *Nature* **345**, 273-6.

Laferte, S. and Dennis, J.W. (1988) Glycosylation-dependent collagen-binding activities of two membrane glycoproteins in MDAY-D2 tumor cells. *Cancer Res* **48**, 4743-8.

Lan, M.S., Batra, S.K., Qi, W.N., Metzgar, R.S. and Hollingsworth, M.A. (1990) Cloning and sequencing of a human pancreatic tumor mucin cDNA. *J. Biol. Chem.* **265**, 15294-9.

Lang, G., Wotton, D., Owen, M.J., Sewell, W.A., Brown, M.H., Mason, D.Y., Crumpton, M.J. and Kioussis, D. (1988) The structure of the human CD2 gene and its expression in transgenic mice. *EMBO J.* **7**, 1675-82.

Lanzavecchia, A. (1990) Receptor-mediated antigen uptake and its effect on antigen presentation to class II-restricted T lymphocytes. *Annu Rev Immunol* **8**, 773-93.

Larsen, E., Celi, A., Gilbert, G.E., Furie, B.C., Erban, J.K., Bonfanti, R., Wagner, D.D. and Furie, B. (1989) PADGEM protein: a receptor that mediates the interaction of activated platelets with neutrophils and monocytes. *Cell* **59**, 305-12.

Lasky, L.A., Singer, M.S., Yednock, T.A., Dowbenko, D., Fennie, C., Rodriguez, H., Nguyen, T., Stachel, S. and Rosen, S.D. (1989) Cloning of a lymphocyte homing receptor reveals a lectin domain. *Cell* **56**, 1045-55.

Lauener, R.P., Geha, R.S. and Vercelli, D. (1990) Engagement of the monocyte surface antigen CD14 induces lymphocyte function-associated antigen-1/intercellular adhesion molecule-1-dependent homotypic adhesion. *J. Immunol.* **145**, 1390-4.

Laver, W.G., Air, G.M., Webster, R.G. and Smith, G.S. (1990) Epitopes on protein antigens: misconceptions and realities [published erratum appears in Cell 1990 Aug 10;62(3):following 608]. *Cell* **61**, 553-6.

Lawrence, M.B. and Springer, T.A. (1991) Leukocytes roll on a selectin at physiologic flow rates: distinction from and prerequisite for adhesion through integrins. *Cell* **65**, 859-73.

Leahy, D.J., Axel, R. and Hendrickson, W.A. (1992) Crystal structure of a soluble form of the human T cell coreceptor CD8 at 2.6Å resolution. *Cell* **68**, 1145-62.

Ledbetter, J.A., Tonks, N.K., Fischer, E.H. and Clark, E.A. (1988) CD45 regulates signal transduction and lymphocyte activation by specific association with receptor molecules on T or B cells. *Proc Natl Acad Sci U S A* **85**, 8628-32.

Leeuwenberg, J.F., Tan, A., Jeunhomme, T.M., Ploegh, H.L. and Buurman, W.A. (1991) The ligand recognized by ELAM-1 on HL60 cells is not carried by N-linked oligosaccharides. *Eur. J. Immunol.* **21**, 3057-9.

Lefrancois, L. and Bevan, M.J. (1985) Novel antigenic determinants of the T200 glycoprotein expressed preferentially by activated cytotoxic T lymphocytes. *J. Immunol.* **135**, 374-383.

Lefrancois, L., Puddington, L., Machamer, C.E. and Bevan, M.J. (1985) Acquisition of cytotoxic T lymphocyte-specific carbohydrate differentiation antigens. *J. Exp. Med.* **162**, 1275-93.

Leonard, W.J., Depper, J.M., Crabtree, G.R., Rudikoff, S., Pumphrey, J., Robb, R.J., Kronke, M., Svetlik, P.B., Peffer, N.J., Waldmann, T.A. and et, a.l. (1984) Molecular cloning and expression of cDNAs for the human interleukin-2 receptor. *Nature* **311**, 626-31.

Lesley, J., He, Q., Miyake, K., Hamann, A., Hyman, R. and Kincade, P.W. (1992) Requirements for hyaluronic acid binding to CD44: a role for the cytoplasmic domain and activation by antibody. *J. Exp. Med.* **175**, 257-266.

Levi, G. and Teichberg, V.I. (1983) Selective interactions of electrolytes from eel electric organ and mouse thymus with mouse immature thymocytes. *Immunol. Lett.* **7**, 35-39.

Levy, S.C., Davies, A.A., Druker, B. and Crumpton, M.J. (1991) The T cell receptor/CD3 complex and CD2 stimulate the tyrosine phosphorylation of indistinguishable patterns of polypeptides in the human T leukemic cell line Jurkat. *Eur. J. Immunol.* **21**, 2203-209.

Libermann, T.A., Lenardo, M. and Baltimore, D. (1990) Involvement of a second lymphoid-specific enhancer element in the regulation of immunoglobulin heavy-chain gene expression. *Mol. Cell. Biol.* **10**, 3155-62.

Ligtenberg, M.J., Vos, H.L., Gennissen, A.M. and Hilkens, J. (1990) Episialin, a carcinoma-associated mucin, is generated by a polymorphic gene encoding splice variants with alternative amino termini. *J. Biol. Chem.* **265**, 5573-8.

Linnenbach, A.J., Wojcierowski, J., Wu, S.A., Pyrc, J.J., Ross, A.H., Dietzschold, B., Speicher, D. and Koprowski, H. (1989) Sequence investigation of the major gastrointestinal tumor-associated antigen gene family, GA733. *Proc Natl Acad Sci U S A* **86**, 27-31.

Linsley, P.S., Brady, W., Grosmaire, L., Aruffo, A., Damle, N.K. and Ledbetter, J.A. (1991) Binding of the B cell activation antigen B7 to CD28 costimulates T cell proliferation and interleukin 2 mRNA accumulation. *J. Exp. Med.* **173**, 721-30.

Lipman, D.J. and Pearson, W.R. (1985) Rapid and sensitive protein similarity searches. *Science* **227**, 1435-41.

Lisowska, E., Messeter, L., Duk, M., Czerwinski, M. and Lundblad, A. (1987) A monoclonal anti-glycophorin A antibody recognizing the blood group M determinant: studies on the subspecificity. *Mol. Immunol.* **24**, 605-13.

Lisowska, E. and Wasniowska, K. (1978) Immunochemical characterization of cyanogen bromide degradation products of M and N blood-group glycopeptides. *Eur. J. Biochem* **88**, 247-52.

Lo, S.K., Lee, S., Ramos, R.A., Lobb, R., Rosa, M., Chi, R.G. and Wright, S.D. (1991) Endothelial-leukocyte adhesion molecule 1 stimulates the adhesive activity of leukocyte integrin CR3 (CD11b/CD18, Mac-1, alpha m beta 2) on human neutrophils. *J. Exp. Med.* **173**, 1493-500.

Lopez, J.A., Chung, D.W., Fujikawa, K., Hagen, F.S., Papayannopoulou, T. and Roth, G.J. (1987) Cloning of the alpha chain of human platelet glycoprotein Ib: a transmembrane protein with homology to leucine-rich alpha 2-glycoprotein. *Proc Natl Acad Sci U S A* **84**, 5615-9.

Lorant, D.E., Patel, K.D., McIntyre, T.M., McEver, R.P., Prescott, S.M. and Zimmerman, G.A. (1991) Coexpression of GMP-140 and PAF by endothelium stimulated by histamine or thrombin: a juxtacrine system for adhesion and activation of neutrophils. *J Cell Biol* **115**, 223-34.

Lowe, J.B., Kukowska, L.J., Nair, R.P., Larsen, R.D., Marks, R.M., Macher, B.A., Kelly, R.J. and Ernst, L.K. (1991) Molecular cloning of a human fucosyltransferase gene that determines expression of the Lewis x and VIM-2 epitopes but not ELAM-1-dependent cell adhesion. *J. Biol. Chem.* **266**, 17467-77.

Lund, K.A., Lazar, C.S., Chen, W.S., Walsh, B.J., Welsh, J.B., Herbst, J.J., Walton, G.M., Rosenfeld, M.G., Gill, G.N. and Wiley, H.S. (1990) Phosphorylation of the epidermal growth factor receptor at threonine 654 inhibits ligand-induced internalization and down-regulation. *J. Biol. Chem.* **265**, 20517-23.

Luther, M., Knoppers, M., Neidhardt, E., Kandeekar, S., Gordon, K., Concino, M., Schmike, P., Francis, M., Rabin, E., Reinherz, E. and Recny, M. (1991) Role of N-linked carbohydrate chains in mediating the immunoadhesion properties of the human CD2 receptor. *J. Cell Biochem. Suppl.15G* Abstract R115

Lux, S.E., John, K.M. and Bennett, V. (1990) Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. *Nature* **344**, 36-42.

Mackay, C.R., Hein, W.R., Brown, M.H. and Matzinger, P. (1988) Unusual expression of CD2 in sheep: implications for T cell interactions. *Eur. J. Immunol.* **18**, 1681-8.

Maddon, P.J., Molineaux, S.M., Maddon, D.E., Zimmerman, K.A., Godfrey, M., Alt, F.W., Chess, L. and Axel, R. (1987) Structure and expression of the human and mouse T4 genes. *Proc Natl Acad Sci U S A* **84**, 9155-9.

Mahoney, P.A., Weber, U., Onofrechuk, P., Biessmann, H., Bryant, P.J. and Goodman, C.S. (1991) The fat tumor suppressor gene in Drosophila encodes a novel member of the cadherin gene superfamily. *Cell* **67**, 853-68.

Mallett, S. and Barclay, A.N. (1991) A new superfamily of cell surface proteins related to the nerve growth factor receptor. *Immunol. Today* **12**, 220-3.

Mallett, S., Fossum, S. and Barclay, A.N. (1990) Characterization of the MRC OX40 antigen of activated CD4 positive T lymphocytes--a molecule related to nerve growth factor receptor. *EMBO J.* **9**, 1063-8.

Marston, F.A. (1986) The purification of eukaryotic polypeptides synthesized in Escherichia coli. *Biochem. J.* **240**, 1-12.

Masellis, S.A., Jensen, G.S., Seehafer, J.G., Slupsky, J.R. and Shaw, A.R. (1990) Anti-CD9 monoclonal antibodies induce homotypic adhesion of pre-B cell lines by a novel mechanism. *J. Immunol.* **144**, 1607-13.

Mason, D.W. and Williams, A.F. (1980) The kinetics of antibody binding to membrane antigens in solution and at the cell surface. *Biochem. J.* **187**, 1-20.

Mason, D.W., Pugh, C.W. and Webb, M. (1981) The rat mixed lymphocyte reaction: roles of a dendritic cell in

intestinal lymph and T-cell subsets defined by monoclonal antibodies. *Immunology* **44**, 445-64.

Matsuura, H., Takio, K., Titani, K., Greene, T., Levery, S.B., Salyan, M.E. and Hakomori, S. (1988) The oncofetal structure of human fibronectin defined by monoclonal antibody FDC-6. Unique structural requirement for the antigenic specificity provided by a glycosylhexapeptide. *J. Biol. Chem.* **263**, 3314-22.

McCall, M.N., Shotton, D.M. and Barclay, A.N. (1992) Expression of soluble isoforms of rat CD45. Analysis by electron microscopy and use in epitope mapping of anti-CD45R monoclonal antibodies. *Immunology* In press

McMaster, W.R. and Williams, A.F. (1979) Identification of Ia glycoproteins in rat thymus and purification from rat spleen. *Eur. J. Immunol.* **9**, 426-433.

McNeill, H., Ozawa, M., Kemler, R. and Nelson, W.J. (1990) Novel function of the cell adhesion molecule uvomorulin as an inducer of cell surface polarity. *Cell* **62**, 309-16.

Medof, M.E., Lublin, D.M., Holers, V.M., Ayers, D.J., Getty, R.R., Leykam, J.F., Atkinson, J.P. and Tykocinski, M.L. (1987) Cloning and characterization of cDNAs encoding the complete sequence of decay-accelerating factor of human complement. *Proc Natl Acad Sci U S A* **84**, 2007-11.

Mentzer, S.J., Remold-O'Donnell, E., Crimmins, M.A., Bierer, B.E., Rosen, F.S. and Burakoff, S.J. (1987) Sialophorin, a surface sialoglycoprotein defective in the Wiskott-Aldrich syndrome, is involved in human T lymphocyte proliferation. *J. Exp. Med.* **165**, 1383-92.

Meri, S., Morgan, B.P., Davies, A., Daniels, R.H., Olavesen, M.G., Waldmann, H. and Lachmann, P.J. (1990) Human protectin (CD59), an 18,000-20,000 MW complement lysis restricting factor, inhibits C5b-8 catalysed insertion of C9 into lipid bilayers. *Immunology* **71**, 1-9.

Meuer, S.C. (1989) Cluster report: CD2. In Knapp, W., Dorken, B., Gilks, W.R., Rieber, E.P., Schmidt, R.E., Stein, H. and Von dem Borne, A.E.G. (eds.), *Leukocyte Typing IV*. Oxford Unviersity Press, Oxford, pp.270-72

Meuer, S.C., Hussey, R.E., Fabbi, M., Fox, D., Acuto, O., Fitzgerald, K.A., Hodgdon, J.C., Protentis, J.P., Schlossman, S.F. and Reinherz, E.L. (1984) An alternative pathway of T-cell activation: a functional role for the 50 kd T11 sheep erythrocyte receptor protein. *Cell* **36**, 897-906.

Mishra, G.C., Berton, M.T., Oliver, K.G., Krammer, P.H., Uhr, J.W. and Vitetta, E.S. (1986) A monoclonal anti-mouse LFA-1 alpha antibody mimics the biological effects of B cell stimulatory factor-1 (BSF-1). *J. Immunol.* **137**, 1590-8.

Moingeon, P., Chang, H.C., Sayre, P.H., Clayton, L.K., Alcover, A., Gardner, P. and Reinherz, E.L. (1989) The structural biology of CD2. *Immunol. Rev.* **111**, 111-44.

Moingeon, P.E., Lucich, J.L., Stebbins, C.C., Recny, M.A., Wallner, B.P., Koyasu, S. and Reinherz, E.L. (1991) Complementary roles for CD2 and LFA-1 adhesion pathways during T cell activation. *Eur. J. Immunol.* **21**, 605-10.

Moran, M.F., Koch, C.A., Anderson, D., Ellis, C., England, L., Martin, G.S. and Pawson, T. (1990) Src homology region 2 domains direct protein-protein interactions in signal transduction. *Proc Natl Acad Sci U S A* **87**, 8622-6.

Moran, N. and Bock, E. (1988) Characterization of the kinetics of neural cell adhesion molecule homophilic binding. *Febs Lett* **242**, 121-4.

Mouhtouris, E., Vaughan, H.A. and Sandrin, M.S. (1991) Characterization of CD48 as an adhesion molecule. *Aust. Soc. Immunol.* Meeting Abstract 242

Mourad, W., Geha, R.S. and Chatila, T. (1990) Engagement of major histocompatibility complex class II molecules induces sustained, lymphocyte function-associated molecule 1-dependent cell adhesion. *J. Exp. Med.* **172**, 1513-6.

Mueller, D.L., Jenkins, M.K. and Schwartz, R.H. (1989) Clonal expansion versus functional clonal inactivation:

a costimulatory signalling pathway determines the outcome of T cell antigen receptor occupancy. *Annu Rev Immunol* **7**, 445-80.

Nagafuchi, A. and Takeichi, M. (1988) Cell binding function of E-cadherin is regulated by the cytoplasmic domain. *EMBO J.* **7**, 3679-84.

Nagafuchi, A., Takeichi, M. and Tsukita, S. (1991) The 102 kd cadherin-associated protein: similarity to vinculin and posttranscriptional regulation of expression. *Cell* **65**, 849-57.

Nakamura, T., Takahashi, K., Fukazawa, T., Koyanagi, M., Yokoyama, A., Kato, H., Yagita, H. and Okumura, K. (1990) Relative contribution of CD2 and LFA-1 to murine T and natural killer cell functions. *J. Immunol.* **145**, 3628-34.

Nathans, J. and Hogness, D.S. (1984) Isolation and nucleotide sequence of the gene encoding human rhodopsin. *Proc Natl Acad Sci U S A* **81**, 4851-5.

Neefjes, J.J., De Bruijn, M.L., Boog, C.J., Nieland, J.D., Boes, J., Melief, C.J. and Ploegh, H.L. (1990) N-linked glycan modification on antigen-presenting cells restores an allospecific cytotoxic T cell response. *J. Exp. Med.* **171**, 583-8.

Nelson, W.J. and Veshnock, P.J. (1987) Ankyrin binding to (Na+ + K+)ATPase and implications for the organization of membrane domains in polarized cells. *Nature* **328**, 533-6.

Nermut, M.V., Green, N.M., Eason, P., Yamada, S.S. and Yamada, K.M. (1988) Electron microscopy and structural model of human fibronectin receptor. *EMBO J.* **7**, 4093-9.

Neufeld, E.F. and Ashwell, G. (1980) In Lennarz, W.J. (eds.), *Carbohydrate recognition systems for receptor-mediated pinocytosis*. Plenum Press, New York, pp. 241-262.

Newton, M.R., Wood, K.J. and Fabre, J.W. (1986) A monoclonal alloantibody detecting a polymorphism of the rat leucocyte common (LC) antigen. *J Immunogenet* **13**, 41-50.

Nishizuka, Y. (1988) The molecular heterogeneity of protein kinase C and its implications for cellular regulation. *Nature* **334**, 661-5.

Nong, Y.H., Remold-O'Donnell, E., LeBien, T.W. and Remold, H.G. (1989) A monoclonal antibody to sialophorin (CD43) induces homotypic adhesion and activation of human monocytes. *J. Exp. Med.* **170**, 259-67.

Nonoyama, S., Nakayama, M., Shiohara, T. and Yata, J. (1989) Only dull CD3+ thymocytes bind to thymic epithelial cells. The binding is elicited by both CD2/LFA-3 and LFA-1/ICAM-1 interactions. *Eur. J. Immunol.* **19**, 1631-5.

Nordling, S., Andersson, L.C. and Hayry, P. (1972) Thymus-dependent and thymus-independent lymphocyte separation: relation to exposed sialic acid on cell surface. *Science* **178**, 1001-2.

Norman, D.G., Barlow, P.N., Baron, M., Day, A.J., Sim, R.B. and Campbell, I.D. (1991) Three-dimensional structure of a complement control protein module in solution. *J Mol Biol* **219**, 717-25.

Norment, A.M., Salter, R.D., Parham, P., Engelhard, V.H. and Littman, D.R. (1988) Cell-cell adhesion mediated by CD8 and MHC class 1 molecules. *Nature* **336**, 79-81.

Nybroe, O., Moran, N. and Bock, E. (1989) Equilibrium binding analysis of neural cell adhesion molecule binding to heparin. *J Neurochem* **52**, 1947-49.

O'Connell, B., Tabak, L.A. and Ramasubbu, N. (1991a) The influence of flanking sequences on O-glycosylation. *Biochem Biophys Res Comm* **180**, 1024-1030.

O'Connell, P.J., Gerkis, V. and d'Apice, A.J. (1991b) Variable O-glycosylation of CD13 (aminopeptidase N). *J. Biol. Chem.* **266**, 4593-7.

O'Toole, T.E., Loftus, J.C., Du, X.P., Glass, A.A., Ruggeri, Z.M., Shattil, S.J., Plow, E.F. and Ginsberg,

M.H. (1990) Affinity modulation of the alpha IIb beta 3 integrin (platelet GPIIb-IIIa) is an intrinsic property of the receptor. *Cell Regul* **1**, 883-93.

Ochs, H.D., Slichter, S.J., Harker, L.A., Von, B.W., Clark, R.A. and Wedgwood, R.J. (1980) The Wiskott-Aldrich syndrome: studies of lymphocytes, granulocytes, and platelets. *Blood* **55**, 243-52.

Ohta, Y., Stossel, T.P. and Hartwig, J.H. (1991) Ligand-sensitive binding of actin-binding protein to immunoglobulin G Fc receptor I (Fc gamma RI). *Cell* **67**, 275-82.

Olofsson, S., Sjoblom, I. and Jeansson, S. (1990) Activity of herpes simplex virus type 1-specified glycoprotein C antigenic site II epitopes reversibly modulated by peripheral fucose or galactose units of glycoprotein oligosaccharides. *J Gen Virol*

Olson, M.O., Guetzow, K. and Busch, H. (1981) Localization of phosphoprotein C23 in nucleoli by immunological methods. *Exp Cell Res* **135**, 259-65.

Otey, C.A., Pavalko, F.M. and Burridge, K. (1990) An interaction between alpha-actinin and the beta 1 integrin subunit in vitro. *J Cell Biol* **111**, 721-9.

Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) Splicing of messenger RNA precursors. *Annu Rev Biochem* **55**, 1119-50.

Pallant, A., Eskenazi, A., Mattei, M.G., Fournier, R.E., Carlsson, S.R., Fukuda, M. and Frelinger, J.G. (1989) Characterization of cDNAs encoding human leukosialin and localization of the leukosialin gene to chromosome 16. *Proc Natl Acad Sci U S A* **86**, 1328-32.

Palmiter, R.D. and Brinster, R.L. (1986) Germ-line transformation of mice. *Annu Rev Genet* **20**, 465-99.

Palmiter, R.D., Sandgren, E.P., Avarbock, M.R., Allen, D.D. and Brinster, R.L. (1991) Heterologous introns can enhance expression of transgenes in mice. *Proc Natl Acad Sci U S A* **88**, 478-82.

Pardi, R., Inveradi, L., Rugarli, C. and Bender, J.R. (1992) Antigen-receptor complex stimulation triggers protein kinase C-dependent CD11a/CD18-cytoskeleton association in T lymphocytes. *J Cell Biol.* **116**, 1211-1220.

Parham, P. (1986) Preparation and purification of active fragments from mouse monoclonal antibodies. In Weir, D.M. (ed.), *Handbook of Experimental Immunnology*. Blackwell Scientific Publications, Oxford, Vol. 1, pp. 14.1-14.25.

Park, J.K., Rosenstein, Y.J., Remold-O'Donnell, E., Bierer, B.E., Rosen, F.S. and Burakoff, S.J. (1991) Enhancement of T-cell activation by the CD43 molecule whose expression is defective in Wiskott-Aldrich syndrome. *Nature* **350**, 706-9.

Parkman, R., Remold-O'Donnell, E., Kenney, D.M., Perrine, S. and Rosen, F.S. (1981) Surface protein abnormalities in lymphocytes and platelets from patients with Wiskott-Aldrich syndrome. *Lancet* **2**, 1387-9.

Parkman, R., Remold, O.E., Kenney, D.M. and Rosen, F.S. (1983) GPL-115 deficiency: a new class of immunodeficiencies. *Birth Defects* **19**, 97-9.

Paulson, J.C. (1989) Glycoproteins:what are the sugar chains for? *TIBS* **14**, 272-276.

Paulson, J.C. and Colley, K.J. (1989) Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. *J. Biol. Chem.* **264**, 17615-8.

Pearlstein, E., Salk, P.L., Yogeeswaran, G. and Karpatkin, S. (1980) Correlation between spontaneous metastatic potential, platelet-aggregating activity of cell surface extracts, and cell surface sialylation in 10 metastatic-variant derivatives of a rat renal sarcoma cell line . *Proc. Natl. Acad. Sci. USA* **77**, 4336-4339.

Pears, C.J. and Parker, P.J. (1991) Domain interactions in protein kinase C. *J. Cell Science* **100**, 683-686.

Pearse, B.M. (1988) Receptors compete for adaptors found in plasma membrane coated pits. *EMBO J.* **7**, 3331-6.

Peralta, E.G., Winslow, J.W., Peterson, G.L., Smith, D.H., Ashkenazi, A., Ramachandran, J., Schimerlik, M.I. and Capon, D.J. (1987) Primary structure and biochemical properties of an M2 muscarinic receptor. *Science* **236**, 600-5.

Peter, M., Nakagawa, J., Doree, M., Labbe, J.C. and Nigg, E.A. (1990) Identification of major nucleolar proteins as candidate mitotic substrates of cdc2 kinase. *Cell* **60**, 791-801.

Peterson, A. and Seed, B. (1987) Monoclonal antibody and ligand binding sites of the T cell erythrocyte receptor (CD2). *Nature* **329**, 842-6.

Pettersson, S., Cook, G.P., Bruggemann, M., Williams, G.T. and Neuberger, M.S. (1990) A second B cell-specific enhancer 3' of the immunoglobulin heavy-chain locus. *Nature* **344**, 165-8.

Pfeifle, J. and Anderer, F.A. (1983) Localization of phosphoprotein PP 105 in cell lines of various species. *Biochem Biophys Res Commun* **116**, 106-12.

Pfeifle, J., Hagmann, W. and Anderer, F.A. (1981) Cell adhesion-dependent differences in endogenous protein phosphorylation on the surface of various cell lines. *Biochim Biophys Acta* **670**, 274-84.

Philbrick, W.M., Palfree, R.G., Maher, S.E., Bridgett, M.M., Sirlin, S. and Bothwell, A.L. (1990) The CD59 antigen is a structural homologue of murine Ly-6 antigens but lacks interferon inducibility. *Eur. J. Immunol.* **20**, 87-92.

Phillips, D.R., Charo, I.F. and Scarborough, R.M. (1991) GPIIb-IIIa: the responsive integrin. *Cell* **65**, 359-62.

Picker, L.J., Kishimoto, T.K., Smith, C.W., Warnock, R.A. and Butcher, E.C. (1991a) ELAM-1 is an adhesion molecule for skin-homing T cells. *Nature* **349**, 796-9.

Picker, L.J., Warnock, R.A., Burns, A.R., Doerschuk, C.M., Berg, E.L. and Butcher, E.C. (1991b) The neutrophil selectin LECAM-1 presents carbohydrate ligands to the vascular selectins ELAM-1 and GMP-140. *Cell* **66**, 921-33.

Pidard, D., Didry, D., Le Deist, F., Durandy, A., Griscelli, C., Bellucci, S. and Nurden, A.T. (1988) Analysis of the membrane glycoproteins of platelets in the Wiskott-Aldrich syndrome. *Br J Haematol* **69**, 529-35.

Pigott, R., Needham, L.A., Edwards, R.M., Walker, C. and Power, C. (1991) Structural and functional studies of the endothelial activation antigen endothelial leucocyte adhesion molecule-1 using a panel of monoclonal antibodies. *J. Immunol.* **147**, 130-5.

Piller, F., Le Deist, F., Weinberg, K.I., Parkman, R. and Fukuda, M. (1991) Altered O-glycan synthesis in lymphocytes from patients with Wiskott-Aldrich syndrome. *J. Exp. Med.* **173**, 1501-10.

Piller, F., Piller, V., Fox, R.I. and Fukuda, M. (1988) Human T-lymphocyte activation is associated with changes in O-glycan biosynthesis. *J. Biol. Chem.* **263**, 15146-50.

Piller, V., Piller, F. and Fukuda, M. (1990) Biosynthesis of truncated O-glycans in the T cell line Jurkat. Localization of O-glycan initiation. *J. Biol. Chem.* **265**, 9264-71.

Piller, V., Piller, F., Klier, F.G. and Fukuda, M. (1989a) O-glycosylation of leukosialin in K562 cells. Evidence for initiation and elongation in early Golgi compartments. *Eur. J. Biochem* **183**, 123-35.

Piller, V., Piller, F. and Fukuda, M. (1989b) Phosphorylation of the major leukocyte surface sialoglycoprotein, leukosialin, is increased by phorbol 12-myristate 13-acetate. *J. Biol. Chem.* **264**, 18824-31.

Pink, J.R. (1983) Changes in T-lymphocyte glycoprotein structures associated with differentiation. *Contemp Top Mol Immunol* **9**, 89-113.

Pircher, H., Mak, T.W., Lang, R., Ballhausen, W., Ruedi, E., Hengartner, H., Zinkernagel, R.M. and Burki, K. (1989) T cell tolerance to Mlsa encoded antigens in T cell receptor V beta 8.1 chain transgenic mice. *EMBO J.* **8**, 719-27.

Plunkett, M.L., Sanders, M.E., Selvaraj, P., Dustin, M.L. and Springer, T.A. (1987) Rosetting of activated human T lymphocytes with autologous erythrocytes. Definition of the receptor and ligand molecules as CD2 and lymphocyte function-associated antigen 3 (LFA-3). *J. Exp. Med.* **165**, 664-76.

Polley, M.J., Phillips, M.L., Wayner, E., Nudelman, E., Singhal, A.K., Hakomori, S. and Paulson, J.C. (1991) CD62 and endothelial cell-leukocyte adhesion molecule 1 (ELAM-1) recognize the same carbohydrate ligand, sialyl-Lewis x. *Proc Natl Acad Sci U S A* **88**, 6224-8.

Pontremoli, S., Melloni, E., Sparatore, B., Michetti, M., Salamino, F. and Horecker, B.L. (1990) Isozymes of protein kinase C in human neutrophils and their modification by two endogenous proteinases. *J. Biol. Chem.* **265**, 706-12.

Porchet, N., Nguyen, V.C., Dufosse, J., Audie, J.P., Guyonnet, D.V., Gross, M.S., Denis, C., Degand, P., Bernheim, A. and Aubert, J.P. (1991) Molecular cloning and chromosomal localization of a novel human tracheo-bronchial mucin cDNA containing tandemly repeated sequences of 48 base pairs. *Biochem Biophys Res Commun* **175**, 414-22.

Post, T.W., Liszewski, M.K., Adams, E.M., Tedja, I., Miller, E.A. and Atkinson, J.P. (1991) Membrane cofactor protein of the complement system: alternative splicing of serine/threonine/proline-rich exons and cytoplasmic tails produces multiple isoforms that correlate with protein phenotype. *J. Exp. Med.* **174**, 93-102.

Powell, L.D., Whitehcart, S.W. and Hart, G.W. (1987) Cell surface sialic acid influences tumor cell recognition in the mixed lymphocyte reaction. *J. Immunol.* **139**, 262-70.

Price, M.R., Hudecz, F., O'Sullivan, C., Baldwin, R.W., Edwards, P.M. and Tendler, S.J. (1990) Immunological and structural features of the protein core of human polymorphic epithelial mucin. *Mol. Immunol.* **27**, 795-802.

Prickett, T.C.R., McKenzie, J.L. and Hart, D.N.J. (1992) Adhesion molecules on human tonsil dendritic cells. *Transplantation* **53**, 483-490.

Probst, J.C., Gertzen, E.M. and Hoffmann, W. (1990) An integumentary mucin (FIM-B.1) from Xenopus laevis homologous with von Willebrand factor. *Biochemistry* **29**, 6240-4.

Prohaska, R., Koerner, T.J., Armitage, I.M. and Furthmayr, H. (1981) Chemical and carbon-13 nuclear magnetic resonance studies of the blood group M and N active sialoglycopeptides from human glycophorin A. *J. Biol. Chem.* **256**, 5781-91.

Rademacher, T.W., Parekh, R.B. and Dwek, R.A. (1988) Glycobiology. *Annu Rev Biochem* **57**, 785-838.

Recny, M.A., Neidhardt, E.A., Sayre, P.H., Ciardelli, T.L. and Reinherz, E.L. (1990) Structural and functional characterization of the CD2 immunoadhesion domain. Evidence for inclusion of CD2 in an alpha-beta protein folding class. *J. Biol. Chem.* **265**, 8542-9.

Reddy, P., Caras, I. and Krieger, M. (1989) Effects of O-linked glycosylation on the cell surface expression and stability of decay accelerating factor, a glycophospholipid-anchored membrane protein. *J. Biol. Chem.* **264**, 17329-36

Reik, W., Williams, G., Barton, S., Norris, M., Neuberger, M. and Surani, M.A. (1987) Provision of the immunoglobulin heavy chain enhancer downstream of a test gene is sufficient to confer lymphoid-specific expression in transgenic mice. *Eur. J. Immunol.* **17**, 465-9.

Remold-O'Donnell, E., Kenney, D. and Rosen, F.S. (1987) Biosynthesis of human sialophorins and analysis of the polypeptide core. *Biochemistry* **26**, 3908-13.

Remold-O'Donnell, E., Kenney, D.M., Parkman, R., Cairns, L., Savage, B. and Rosen, F.S. (1984) Characterization of a human lymphocyte surface sialoglycoprotein that is defective in Wiskott-Aldrich syndrome. *J. Exp. Med.* **159**, 1705-23.

Remold-O'Donnell, E., Zimmerman, C., Kenney, D. and Rosen, F.S. (1987) Expression on blood cells of

sialophorin, the surface glycoprotein that is defective in Wiskott-Aldrich syndrome. *Blood* **70**, 104-9.

Reth, M. (1989) Antigen receptor tail clue [letter]. *Nature* **338**, 383-4.

Richardson, N.E., Chang, H.C., Brown, N.R., Hussey, R.E., Sayre, P.H. and Reinherz, E.L. (1988) Adhesion domain of human T11 (CD2) is encoded by a single exon. *Proc Natl Acad Sci U S A* **85**, 5176-80.

Rimm, D.L., Kaiser, D.A., Bhandari, D., Maupin, P., Kiehart, D.P. and Pollard, T.D. (1990) Identification of functional regions on the tail of Acanthamoeba myosin-II using recombinant fusion proteins. I. High resolution epitope mapping and characterization of monoclonal antibody binding sites. *J Cell Biol*

Rini, J.M., Schulze-Gahmen, U. and Wilson, I.A. (1992) Structural evidence for induced fit as a mechanism for antibody-antigen recognition. *Science* **255**, 959-965.

Romeo, C. and Seed, B. (1991) Cellular immunity to HIV activated by CD4 fused to T cell or Fc receptor polypeptides. *Cell* **64**, 1037-46.

Rooney, C.M. and Munro, A.J. (1984) NK cells can recognize asialylated autologous lymphocytes and ABO-mismatched lymphocytes. *Immunology* **51**, 193-9.

Rose, M.C., Voter, W.A., Brown, C.F. and Kaufman, B. (1984) Structural features of human tracheobronchial mucus glycoprotein. *Biochem. J.* **222**, 371-7.

Rose, M.L. (1982) Binding of peanut lectin to lymphocytes: a marker of immaturity? *Immunology Todat* **3**, 6-8.

Rosen, F.S., Cooper, M.D. and Wedgwood, R.J. (1984) The primary immunodeficiencies. (2). *N Engl J Med* **311**, 300-10.

Rosenberg, A.H., Lade, B.N., Chui, D.S., Lin, S.W., Dunn, J.J. and Studier, F.W. (1987) Vectors for selective expression of cloned DNAs by T7 RNA polymerase. *Gene* **56**, 125-35.

Rosenstein, Y., Park, J.K., Hahn, W.C., Rosen, F.S., Bierer, B.E. and Burakoff, S.J. (1991) CD43, a molecule defective in Wiskott-Aldrich syndrome, binds ICAM-1. *Nature* **354**, 233-235.

Rudensky, A., Preston, H.P., Hong, S.C., Barlow, A. and Janeway, C.J. (1991) Sequence analysis of peptides bound to MHC class II molecules. *Nature* **353**, 622-7.

Rutishauser, U., Acheson, A., Hall, A.K., Mann, D.M. and Sunshine, J. (1988) The neural cell adhesion molecule (NCAM) as a regulator of cell-cell interactions. *Science* **240**, 53-7.

Rutschmann, R. and Karjalainen, K. (1991) Mouse LFA-3 studied with chimeric soluble CD2 shows preferential expression on lymphoid cells. *Eur. J. Immunol.* **21**, 1379-84.

Ryu, S.E., Kwong, P.D., Truneh, A., Porter, T.G., Arthos, J., Rosenberg, M., Dai, X.P., Xuong, N.H., Axel, R., Sweet, R.W. and Hendrickson, W.A. (1990) Crystal structure of an HIV-binding recombinant fragment of human CD4. *Nature* **348**, 419-26.

Sadler, J.E., Paulson, J.C. and Hill, R.L. (1979) The role of sialic acid in the expression of human MN blood group antigens. *J. Biol. Chem.* **254**, 2112-9.

Saitoh, O., Piller, F., Fox, R.I. and Fukuda, M. (1991) T-lymphocytic leukemia expresses complex, branched O-linked oligosaccharides on a major sialoglycoprotein, leukosialin. *Blood* **77**, 1491-9.

Salter, R.D., Benjamin, R.J., Wesley, P.K., Buxton, S.E., Garrett, T.P., Clayberger, C., Krensky, A.M., Norment, A.M., Littman, D.R. and Parham, P. (1990) A binding site for the T-cell co-receptor CD8 on the alpha 3 domain of HLA-A2. *Nature* **345**, 41-6.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning. A Laboratory Manual. 2nd edition, Cold Spring Harbour Laboratory Press,

Samelson, L.E., Phillips, A.F., Luong, E.T. and Klausner, R.D. (1990) Association of the fyn protein-tyrosine

kinase with the T-cell antigen receptor. *Proc Natl Acad Sci U S A* **87**, 4358-62.

Sanchez-Madrid, F., Krensky, A.M., Ware, C.F., Robbins, E., Strominger, J.L., Burakoff, S.J. and Springer, T.A. (1982) Three distinct antigens associated with human T-lymphocyte-mediated cytolysis: LFA-1, LFA-2, and LFA-3. *Proc Natl Acad Sci U S A* **79**, 7489-93.

Sanders, S.K., Fox, R.O. and Kavathas, P. (1991) Mutations in CD8 that affect interactions with HLA class I and monoclonal anti-CD8 antibodies. *J. Exp. Med.* **174**, 371-9.

Sayre, P.H., Chang, H.C., Hussey, R.E., Brown, N.R., Richardson, N.E., Spagnoli, G., Clayton, L.K. and Reinherz, E.L. (1987) Molecular cloning and expression of T11 cDNAs reveal a receptor-like structure on human T lymphocytes. *Proc Natl Acad Sci USA* **84**, 2941-5.

Sayre, P.H., Hussey, R.E., Chang, H.C., Ciardelli, T.L. and Reinherz, E.L. (1989) Structural and binding analysis of a two domain extracellular CD2 molecule. *J. Exp. Med.***169**, 995-1009.

Schachter, H. and Brockhausen, I. (1989) The biosynthesis of branched O-glycans. *Symp Soc Exp Biol* **43**, 1-26.

Schmid, K., Hediger, M.A., Brossmer, R., Collins, J.H., Haupt, H., Marti, T., Offner, G.D., Schaller, J., Takagaki, K., Walsh, M.T., Schwick, G.H., Rosen, F.S. and Remold-O'Donnell, E. (1992) Amino acid sequence of human plasma galactoglycoprotein: Identity with the extracellular region of CD43 (sialophorin). *Proc. Natl. Acad. Sci. USA* **89**, 663-667.

Schmid, K., Mao, S.K.Y., Kimura, A., Hayashi, S. and Binette, J.P. (1980) Isolation and characterization of a Serine-Threonin-rich galactoglycoprotein from normal human plasma. *J. Biol. Chem.* **255**, 3221-3226.

Schoemaker, J.M., Brasnett, A.H. and Marston, F.A. (1985) Examination of calf prochymosin accumulation in Escherichia coli: disulphide linkages are a structural component of prochymosin-containing inclusion bodies. *EMBO J.* **4**, 775-80.

Schraven, B., Samstag, Y., Altevogt, P. and Meuer, S.C. (1990) Association of CD2 and CD45 on human T lymphocytes. *Nature* **345**, 71-4.

Schulz, G.E. and Schirmer, R.H. (1979) *Principles of protein structure.* Springer-Verlag, New York.

Seamen, W.E., Eriksson, E., Dobrow, R. and Imboden, J.B. (1987) Inositol triphosphate is generated by a rat natural killer cell tumor in response to target cells or to crosslinked monoclonal antibody OX-34: possible signaling role for the OX-34 determinant during activation by target cells. *Proc. Natl. Acad. Sci. USA* **84**, 4239-4245.

Sedgwick, J.D., MacPhee, I.A. and Puklavec, M. (1989) Isolation of encephalitogenic CD4+ T cell clones in the rat. Cloning methodology and interferon-gamma secretion. *J Immunol Methods* **121**, 185-96.

See, Y.P. and Jaskowski, G. (1989) In Creighton, T.E. (eds.), Estimating molecular weights of polypeptides by SDS electrophoresis. IRL Press

Seed, B. (1987) An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. *Nature* **329**, 840-2.

Seed, B. and Aruffo, A. (1987) Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. *Proc Natl Acad Sci U S A* **84**, 3365-9.

Sekine, H., Ohno, T. and Kufe, D.W. (1985) Purification and characterization of a high molecular weight glycoprotein detectable in human milk and breast carcinomas. *J. Immunol.* **135**, 3610-5.

Selvaraj, P., Dustin, M.L., Mitnacht, R., Hunig, T., Springer, T.A. and Plunkett, M.L. (1987a) Rosetting of human T lymphocytes with sheep and human erythrocytes. Comparison of human and sheep ligand binding using purified E receptor. *J. Immunol.* **139**, 2690-5.

Selvaraj, P., Plunkett, M.L., Dustin, M., Sanders, M.E., Shaw, S. and Springer, T.A. (1987b) The T lymphocyte glycoprotein CD2 binds the cell surface ligand LFA-3. *Nature* **326**, 400-3.

Semenkovich, C.F., Ostlund, R.J., Olson, M.O. and Yang, J.W. (1990) A protein partially expressed on the surface of HepG2 cells that binds lipoproteins specifically is nucleolin. *Biochemistry* **29**, 9708-13.

Sen, J., Bossu, P., Burakoff, S.J. and Abbas, A.K. (1992) T cell surface molecules regulating noncognate B lymphocyte activation: role of CD2 and LFA-1. *J. Immunol.* **148**, 1037-42

Sewell, W.A., Brown, M.H., Dunne, J., Owen, M.J. and Crumpton, M.J. (1986) Molecular cloning of the human T-lymphocyte surface CD2 (T11) antigen. *Proc Natl Acad Sci U S A* **83**, 8718-22.

Sewell, W.A., Brown, M.H., Owen, M.J., Fink, P.J., Kozak, C.A. and Crumpton, M.J. (1987) The murine homologue of the T lymphocyte CD2 antigen: molecular cloning, chromosome assignment and cell surface expression. *Eur. J. Immunol.* **17**, 1015-20.

Sewell, W.A., Palmer, R.W., Spurr, N.K., Sheer, D., Brown, M.H., Bell, Y. and Crumpton, M.J. (1988) The human LFA-3 gene is located at the same chromosome band as the gene for its receptor CD2. *Immunogenetics* **28**, 278-82.

Sharon, N. and Lis, H. (1989) Lectins as cell recognition molecules. *Science* **246**, 227-34.

Shaw, S., Luce, G.E., Quinones, R., Gress, R.E., Springer, T.A. and Sanders, M.E. (1986) Two antigen-independent adhesion pathways used by human cytotoxic T-cell clones. *Nature* **323**, 262-4.

Shelley, C.S., Remold-O'Donnell, E., Davis, A., Bruns, G.A., Rosen, F.S., Carroll, M.C. and Whitehead, A.S. (1989) Molecular characterization of sialophorin (CD43), the lymphocyte surface sialoglycoprotein defective in Wiskott-Aldrich syndrome. *Proc Natl Acad Sci U S A* **86**, 2819-23.

Shelley, C.S., Remold-O'Donnell, E., Rosen, F.S. and Whitehead, A.S. (1990) Structure of the human sialophorin (CD43) gene. Identification of features atypical of genes encoding integral membrane proteins. *Biochem. J.* **270**, 569-76.

Shimizu, Y., Shaw, S., Graber, N., Gopal, T.V., Horgan, K.J., Van Seventer, G.A. and Newman, W. (1991) Activation-independent binding of human memory T cells to adhesion molecule ELAM-1. *Nature* **349**, 799-802.

Shimizu, Y., van Seventer, G.A., Horgan, K.J. and Shaw, S. (1990a) Costimulation of proliferative responses of resting CD4+ T cells by the interaction of VLA-4 and VLA-5 with fibronectin or VLA-6 with laminin. *J. Immunol.* **145**, 59-67.

Shimizu, Y., van Seventer, G.A., Horgan, K.J. and Shaw, S. (1990b) Roles of adhesion molecules in T-cell recognition: fundamental similarities between four integrins on resting human T cells (LFA-1, VLA-4, VLA-5, VLA-6) in expression, binding, and costimulation. *Immunol. Rev.* **114**, 109-43.

Shogren, R., Gerken, T.A. and Jentoft, N. (1989) Role of glycosylation on the conformation and chain dimensions of O-linked glycoproteins: light-scattering studies of ovine submaxillary mucin. *Biochemistry* **28**, 5525-36.

Shogren, R.L., Jamieson, A.M., Blackwell, J. and Jentoft, N. (1986) Conformation of mucous glycoproteins in aqueous solvents. *Biopolymers* **25**, 1505-17.

Shortman, K., Von, B.H., Lipp, J. and Hopper, K. (1975) Subpopulations of T-lymphocytes. Physical separation, functional specialisation and differentiation pathways of sub-sets of thymocytes and thymus-dependent peripheral lymphocytes. *Transplant Rev* **25**, 163-210.

Shotton, D.M., Burke, B.E. and Branton, D. (1979) The molecular structure of human erythrocyte spectrin. Biophysical and electron microscopic studies. *J Mol Biol* **131**, 303-29.

Siddiqui, J., Abe, M., Hayes, D., Shani, E., Yunis, E. and Kufe, D. (1988) Isolation and sequencing of a cDNA coding for the human DF3 breast carcinoma-associated antigen. *Proc Natl Acad Sci U S A* **85**, 2320-3.

Siegelman, M.H., van de Rijn, M. and Weissman, I.L. (1989) Mouse lymph node homing receptor cDNA clone encodes a glycoprotein revealing tandem interaction domains. *Science* **243**, 1165-72.

Silverman, L.B., Wong, R.C., Remold-O'Donnell, E., Vercelli, D., Sancho, J., Terhorst, C., Rosen, F., Geha, R. and Chatila, T. (1989) Mechanism of mononuclear cell activation by an anti-CD43 (sialophorin) agonistic antibody. *J. Immunol.* **142**, 4194-200.

Simmons, D.L., Satterthwaite, A.B., Tenen, D.G. and Seed, B. (1992) Molecular cloning of a cDNA encoding CD34, a sialomucin of human hematopoietic stem cells. *J. Immunol.* **148**, 267-271.

Sims, P.J., Ginsberg, M.H., Plow, E.F. and Shattil, S.J. (1991) Effect of platelet activation on the conformation of the plasma membrane glycoprotein IIb-IIIa complex. *J. Biol. Chem.* **266**, 7345-52.

Sjoblom, I., Lundstrom, M., Sjogren -Jansson, E., Glorioso, J.C., Jeansson, S. and Olofsson, S. (1987) Demonstration and mapping of highly carbohydrate-dependent epitopes in the herpes simplex virus type 1-specified glycoprotein C. *J Gen Virol*

Skerra, A. and Pluckthun, A. (1988) Assembly of a functional immunoglobulin Fv fragment in Escherichia coli. *Science* **240**, 1038-41.

Skolnik, E.Y., Margolis, B., Mohammadi, M., Lowenstein, E., Fischer, R., Drepps, A., Ullrich, A. and Schlessinger, J. (1991) Cloning of PI3 kinase-associated p85 utilizing a novel method for expression/cloning of target proteins for receptor tyrosine kinases. *Cell* **65**, 83-90.

Smith, C.A., Davis, T., Anderson, D., Solam, L., Beckmann, M.P., Jerzy, R., Dower, S.K., Cosman, D. and Goodwin, R.G. (1990) A receptor for tumor necrosis factor defines an unusual family of cellular and viral proteins. *Science* **248**, 1019-23.

Smith, D.B., Davern, K.M., Board, P.G., Tiu, W.U., Garcia, E.G. and Mitchell, G.F. (1986) Mr 26,000 antigen of Schistosoma japonicum recognized by resistant WEHI 129/J mice is a parasite glutathione S-transferase [published erratum appears in Proc Natl Acad Sci U S A 1987 Sep;84(18):6541]. *Proc Natl Acad Sci U S A* **83**, 8703-7.

Smith, D.B. and Johnson, K.S. (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione transferase. *Gene* **67**, 31-40.

Smith, S.H., Rigley, K.P. and Callard, R.E. (1991) Activation of human B cells through the CD19 surface antigen results in homotypic adhesion by LFA-1-dependent and -independent mechanisms. *Immunology* **73**, 293-7.

Snary, D., Allen, A.K., Faulkes, R.A. and Crumpton, M.J. (1976) Carbohydrate composition of lymphocyte plasma membrane from pig mesenteric lymph node. *Biochem. J.* **153**, 75-8.

Solter, D. and Knowles, B.B. (1978) Monoclonal antibody defining a stage-specific embryonic antigen (SSEA-1). *Proc Natl Acad Sci U S A* **75**, 5565-69.

Spertini, F., Stohl, W., Ramesh, N., Moody, C. and Geha, R.S. (1991a) Induction of human T cell proliferation by a monoclonal antibody to CD5. *J. Immunol.* **146**, 47-52.

Spertini, O., Kansas, G.S., Munro, J.M., Griffin, J.D. and Tedder, T.F. (1991b) Regulation of leukocyte migration by activation of the leukocyte adhesion molecule-1 (LAM-1) selectin. *Nature* **349**, 691-4.

Sportsman, J.R., Park, M.M., Cheresh, D.A., Fukuda, M., Elder, J.H. and Fox, R.I. (1985) Characterization of a membrane surface glycoprotein associated with T-cell activation. *J. Immunol.* **135**, 158-64.

Springer, T.A. (1990) The sensation and regulation of interactions with the extracellular environment: the cell biology of lymphocyte adhesion receptors. *Annu Rev Cell Biol* **6**, 359-402.

Spruyt, L.L., Glennie, M.J., Beyers, A.D. and Williams, A.F. (1991) Signal transduction by the CD2 antigen in T cells and natural killer cells: requirement for expression of a functional T cell receptor or binding of antibody Fc to the Fc receptor, Fc gamma RIIIA (CD16). *J. Exp. Med.* **174**, 1407-15.

Staden, R. (1986) The current status and portability of our sequence handling software. *Nuc Acid Res* **14**, 217-21

Stahli, C., Caravatti, M., Aeschbacher, M., Kocyba, C., Takacs, B. and Carmann, H. (1988) Mucin-like carcinoma-associated antigen defined by three monoclonal antibodies against different epitopes. *Cancer Res* **48**, 6799-802.

Stamenkovic, I., Amiot, M., Pesando, J.M. and Seed, B. (1989) A lymphocyte molecule implicated in lymph node homing is a member of the cartilage link protein family. *Cell* **56**, 1057-62.

Stamenkovic, I., Aruffo, A., Amiot, M. and Seed, B. (1991a) The hematopoietic and epithelial forms of CD44 are distinct polypeptides with different adhesion potentials for hyaluronate-bearing cells. *EMBO J.* **10**, 343-8.

Stamenkovic, I., Sgroi, D., Aruffo, A., Sy, M.S. and Anderson, T. (1991b) The B lymphocyte adhesion molecule CD22 interacts with leukocyte common antigen CD45RO on T cells and alpha 2-6 sialyltransferase, CD75, on B cells. *Cell* **66**, 1133-44.

Standring, R. and Williams, A.F. (1978) Glycoproteins and antigens of membranes prepared from rat thymocytes after lysis by shearing or with the detergent tween-40. *Biochim. Biophys. Acta* **508**, 85-96.

Standring, R., McMaster, W.R., Sunderland, C.A. and Williams, A.F. (1978) The predominant heavily glycosylated glycoproteins at the surface of rat lymphoid cells are differentiation antigens. *Eur. J. Immunol.* **8**, 832-9.

Staunton, D.E., Dustin, M.L., Erickson, H.P. and Springer, T.A. (1990) The arrangement of the immunoglobulin-like domains of ICAM-1 and the binding sites for LFA-1 and rhinovirus [published erratum appears in Cell 1990 Jun 15;61(2):1157]. *Cell* **61**, 243-54.

Staunton, D.E., Dustin, M.L. and Springer, T.A. (1989a) Functional cloning of ICAM-2, a cell adhesion ligand for LFA-1 homologous to ICAM-1. *Nature* **339**, 61-4.

Staunton, D.E., Fisher, R.C., LeBeau, M.M., Lawrence, J.B., Barton, D.E., Francke, U., Dustin, M. and Thorley-Lawson.D. (1989b) Blast-1 possesses a glycosyl-phosphatidylinositol (GPI) membrane anchor, is related to LFA-3 and OX-45, and maps to chromosome 1q21-23. *J. Exp. Med.* **169**, 1087-99.

Stocks, S.C., Albrechtsen, M. and Kerr, M.A. (1990) Expression of the CD15 differentiation antigen (3-fucosyl-N-acetyl-lactosamine, LeX) on putative neutrophil adhesion molecules CR3 and NCA-160. *Biochem. J.* **268**, 275-80.

Stoll, M., Dalchau, R. and Schmidt, R.E. (1989) Cluster Report: CD43. In Knapp, W., Dorken, B., Gilks, W.R., Rieber, E.P., Schmidt, R.E., Stein, H. and Von dem Borne, A.E.G. (eds.), *Leukocyte Typing IV.* Oxford Unviersity Press, Oxford, pp. 604-608.

Storb, U. (1987) Transgenic mice with immunoglobulin genes. *Annu Rev Immunol* **5**, 151-74.

Stross, W.P., Warnke, R.A., Flavell, D.J., Flavell, S.U., Simmons, D., Gatter, K.C. and Mason, D.Y. (1989) Molecule detected in formalin fixed tissue by antibodies MT1, DF-T1, and L60 (Leu-22) corresponds to CD43 antigen. *J Clin Pathol* **42**, 953-61.

Strous, G.J. and Dekker, J. (1992) Mucin-Type Glycoproteins. *Crit. Rev. Bioch. Mol. Biol.* **27**, 57-92.

Studier, F.W. and Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113-30.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**, 60-89.

Sugimura, K. and Higashi, N. (1988) A novel outer-membrane-associated protease in Escherichia coli. *J Bacteriol* **170**, 3650-4.

Sugimura, K. and Nishihara, T. (1988) Purification, characterization, and primary structure of Escherichia coli protease VII with specificity for paired basic residues: identity of protease VII and OmpT. *J Bacteriol* **170**, 5625-32.

Sunderland, C.A., McMaster, W.R. and Williams, A.F. (1979) Purification with monoclonal antibody of a predominant leukocyte-common antigen and glycoprotein from rat thymocytes. Eur. J. Immunol. 9, 155-59

Swack, J.A., Mier, J.W., Romain, P.L., Hull, S.R. and Rudd, C.E. (1991) Biosynthesis and post-translational modificaiton of CD6, a T cell signal transducing molecule. *J. Biol. Chem.* **266**, 7137-7143.

Swamy, K.H. and Goldberg, A.L. (1981) E. coli contains eight soluble proteolytic activities, one being ATP dependent. *Nature* **292**, 652-4.

Taira, S., Kakiuchi, T., Minami, M., Nariuchi, H. and Taiara, S.[.t.T.S. (1986) The regulatory role of sialic acids in the response of class II reactive T cell hybridomas to allogeneic B cells [published erratum appears in J Immunol 1986 Dec 15;137(12):4021]. *J. Immunol.* **137**, 2448-54.

Takahashi, S., Doss, C., Levy, S. and Levy, R. (1990) TAPA-1, the target of an antiproliferative antibody, is associated on the cell surface with the Leu-13 antigen. *J. Immunol.* **145**, 2207-13.

Takeichi, M. (1991) Cadherin cell adhesion receptors as a morphogenetic regulator. *Science* **251**, 1451-1455.

Tamatani, T., Kotani, M. and Miyasaka, M. (1991) Characterization of the rat leukocyte integrin, CD11/CD18, by the use of LFA-1 subunit-specific monoclonal antibodies. *Eur. J. Immunol.* **21**, 627-33.

Tamatani, T. and Miyasaka, M. (1990) Identification of monoclonal antibodies reactive with the rat homolog of ICAM-1, and evidence for a differential involvement of ICAM-1 in the adherence of resting versus activated lymphocytes to high endothelial cells. *Int Immunol* **2**, 165-71.

Telfer, J.C. and Rudd, C.E. (1991) A 32-kD GTP-binding protein associated with the CD4-p56lck and CD8-p56lck T cell receptor complexes. *Science* **254**, 439-41.

Thomas, M.L. (1989) The leukocyte common antigen family. *Annu Rev Immunol* **7**, 339-69.

Thorley-Lawson, D.T., Schooley, A.K., Bhan, A.K. and Nadler, L.M. (1982) Epstein-Barr virus superinduces a new human B cell differentiation antigen (Blast-1) expressed on transformed lymphoblasts. *Cell* **30**, 415-422.

Tiemeyer, M., Swiedler, S.J., Ishihara, M., Moreland, M., Schweingruber, H., Hirtzer, P. and Brandley, B.K. (1991) Carbohydrate ligands for endothelial-leukocyte adhesion molecule 1. *Proc Natl Acad Sci U S A* **88**, 1138-42.

Timpte, C.S., Eckhardt, A.E., Abernethy, J.L. and Hill, R.L. (1988) Porcine submaxillary gland apomucin contains tandemly repeated, identical sequences of 81 residues. *J. Biol. Chem.* **263**, 1081-8.

Tomita, M., Furthmayr, H. and Marchesi, V.T. (1978) Primary structure of human erythrocyte glycophorin A. Isolation and characterization of peptides and complete amino acid sequence. *Biochemistry* **17**, 4756-70.

Townsend, A.R., McMichael, A.J., Carter, N.P., Huddleston, J.A. and Brownlee, G.G. (1984) Cytotoxic T cell recognition of the influenza nucleoprotein and hemagglutinin expressed in transfected mouse L cells. *Cell* **39**, 13-25.

Tozeren, A., Sung, K.-L.P., Sung, L.A., Dustin, M.L., Chan, P.-Y., Springer, T.A. and Chien, S. (1992) Micromanipulation of adhesion of a Jurkat cell to a planar bilayer membrane containing lymphocyte function-associated antigen 3 molecules. *J. Cell Biol.* **116**, 997-1006.

Traunecker, A., Dolder, B., Oliveri, F. and Karjalainen, K. (1989) Solubilizing the T-cell receptor--problems in solution. *Immunol. Today* **10**, 29-32.

Trowbridge, I.S. (1991) CD45. A prototype for transmembrane protein tyrosine phosphatases. *J. Biol. Chem.* **266**, 23517-20.

True, D.D., Singer, M.S., Lasky, L.A. and Rosen, S.D. (1990) Requirement for sialic acid on the endothelial ligand of a lymphocyte homing receptor. *J Cell Biol*

Tuggle, C.K. and Fuchs, J.A. (1985) Glutathione reductase is not required for maintenance of reduced glutathione

in *Escherichia coli* K-12. *J Bacteriol* **162**, 448-50.

Turner, J.M., Brodsky, M.H., Irving, B.A., Levin, S.D., Perlmutter, R.M. and Littman, D.R. (1990) Interaction of the unique N-terminal region of tyrosine kinase p56lck with cytoplasmic domains of CD4 and CD8 is mediated by cysteine motifs. *Cell* **60**, 755-65.

Turner, R.S., Kemp, B.E., Su, H.D. and Kuo, J.F. (1985) Substrate specificity of phospholipid/Ca2+-dependent protein kinase as probed with synthetic peptide fragments of the bovine myelin basic protein. *J. Biol. Chem.* **260**, 11503-7.

Uchida, Y., Tsukada, Y. and Sugimori, T. (1979) Enzymatic properties of neuraminidases from Arthrobacter ureafaciens. *J Biochem Tokyo* **86**, 1573-85.

Udaka, K., Chua, M.M., Tong, L.H., Karush, F. and Goodgal, S.H. (1990) Bacterial expression of immunoglobulin VH proteins. *Mol. Immunol.* **27**, 25-35.

Ullman, K.S., Northrop, J.P., Verweij, C.L. and Crabtree, G.R. (1990) Transmission of signals from the T lymphocyte antigen receptor to the genes responsible for cell proliferation and immune function: the missing link. *Annu Rev Immunol* **8**, 421-52.

Ullrich, A. and Schlessinger, J. (1990) Signal transduction by receptors with tyrosine kinase activity. *Cell* **61**, 203-12.

Valent, P., Majdic, O., Maurer, D., Bodger, M., Muhm, M. and Bettelheim, P. (1990) Further characterization of surface membrane structures expressed on human basophils and mast cells. *Int Arch Allergy Appl Immunol* **91**, 198-203.

van Bleek, G.M. and Nathenson, S.G. (1990) Isolation of an endogenously processed immunodominant viral peptide from the class I H-2K^b molecule. *Nature* **348**, 213-6.

van de Velde, H., von Hoegen, I., Luo, W., Parnes, J.R. and Thielemans, K. (1991) The B-cell surface protein CD72/Lyb-2 is the ligand for CD5. *Nature* **351**, 662-5.

van Kooyk, Y., van de Wiel van Kemenade, P., Weder, P., Kuijpers, T.W. and Figdor, C.G. (1989) Enhancement of LFA-1-mediated cell adhesion by triggering through CD2 or CD3 on T lymphocytes. *Nature* **342**, 811-3.

van Kooyk, Y., Weder, P., Hogervorst, F., Verhoeven, A.J., van Seventer, G., de Velde, A., Borst, J., Keizer, G.D. and Figdor, C.G. (1991) Activation of LFA-1 through a Ca2(+)-dependent epitope stimulates lymphocyte adhesion. *J Cell Biol* **112**, 345-54.

van Lier, R., Bloemena, E., Brouwer, M., Van Heijst, J., Weinreich, S. and Aarden, L. (1989) Studies on the monocyte dependence of T-cell proliferation induced by monoclonal antibodies directed against regions I and II of the CD2 antigen. *Immunology* **67**, 333-8.

van Noesel, C., Miedema, F., Brouwer, M., de Rie, M.A., Aarden, L.A. and van Lier, R.A. (1988) Regulatory properties of LFA-1 alpha and beta chains in human T-lymphocyte activation. *Nature* **333**, 850-2.

van Seventer, G.A., Shimizu, Y., Horgan, K.J., Wallner, B.P. and Shaw, S. (1989) The three LFA-3 specific mAb in the non-lineage panel of workshop mAb all inhibit rosetting. In Knapp, W., Dorken, B., Gilks, W.R., Rieber, E.P., Schmidt, R.E., Stein, H. and Von dem Borne, A.E.G. (eds.), *Leukocyte Typing IV*. Oxford Unviersity Press, Oxford, pp. 717.

Verweij, C.L., Guidos, C. and Crabtree, G.R. (1990) Cell type specificity and activation requirements for NFAT-1 (nuclear factor of activated T-cells) transcriptional activity determined by a new method using transgenic mice to assay transcriptional activity of an individual nuclear factor. *J. Biol. Chem.* **265**, 15788-95.

Volarevic, S., Burns, C.M., Sussman, J.J. and Ashwell, J.D. (1990) Intimate association of Thy-1 and the T-cell antigen receptor with the CD45 tyrosine phosphatase. *Proc Natl Acad Sci U S A* **87**, 7085-9.

Vollger, L.W., Tuck, D.T., Springer, T.A., Haynes, B.F. and Singer, K.H. (1987) Thymocyte binding to human

thymic epithelial cells is inhibited by monoclonal antibodies to CD-2 and LFA-3 antigens. *J. Immunol.* **138**, 358-63.

von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* **14**, 4683-90.

Wagner, M. (1986) A consideration of the origin of processed pseudogenes. *Trends in Genetics*

Waibel, R., O'Hara, C.J. and Stahel, R.A. (1987) Characterization of an epithelial and a tumor-associated human small cell lung carcinoma glycoprotein antigen. *Cancer Res* **47**, 3766-70.

Wallner, B.P., Frey, A.Z., Tizard, R., Mattaliano, R.J., Hession, C., Sanders, M.E., Dustin, M.L. and Springer, T.A. (1987) Primary structure of lymphocyte function-associated antigen 3 (LFA-3). The ligand of the T lymphocyte CD2 glycoprotein. *J. Exp. Med.* **166**, 923-32.

Wang, J., Oketani, M. and Watanabe, T. (1991) Positive and negative regulation of immunoglobulin gene expression by a novel B-cell-specific enhancer element. *Mol. Cell. Biol.* **11**, 75-83.

Wang, J.H., Yan, Y.W., Garrett, T.P., Liu, J.H., Rodgers, D.W., Garlick, R.L., Tarr, G.E., Husain, Y., Reinherz, E.L. and Harrison, S.C. (1990) Atomic structure of a fragment of human CD4 containing two immunoglobulin-like domains. *Nature* **348**, 411-8.

Ward, E.S., Gussow, D., Griffiths, A.D., Jones, P.T. and Winter, G. (1989) Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. *Nature* **341**, 544-6.

Watson, S.R., Imai, Y., Fennie, C., Geoffrey, J., Singer, M., Rosen, S.D. and Lasky, L.A. (1991) The complement binding-like domains of the murine homing receptor facilitate lectin activity. *J Cell Biol* **115**, 235-43.

Watson, S.R., Imai, Y., Fennie, C., Geoffroy, J.S., Rosen, S.D. and Lasky, L.A. (1990) A homing receptor-IgG chimera as a probe for adhesive ligands of lymph node high endothelial venules. *J Cell Biol* **110**, 2221-9.

Watzlawick, H., Walsh, M.T., Ehrhard, I., Slayter, H.S., Haupt, H., Schwick, H.G., Jourdian, G.W., Hase, S., Schmid, K. and Brossmer, R. (1991) The effect of the carbohydrate moiety upon the size and conformation of human plasma galactoglycoprotein as judged by electron microscopy and circular dichroism. Structural studies of a glycoprotein after stepwise enzymic carbohydrate removal. *Biochem. J.*

Webb, D.S.A., Shimizu, Y., Van Seventer, G.A., Shaw, S. and Gerrard, T.L. (1990) LFA-3, CD44, and CD45: Physiologic triggers of human monocyte TNF and IL-1 release. *Science* **249**, 1295-97.

Webb, M., Mason, D.W. and Williams, A.F. (1979) Inhibition of mixed lymphocyte response by monoclonal antibody specific for a rat T lymphocyte subset. *Nature* **282**, 841-3.

Wegener, A.-M.K., Letourneur, F., Hoeveler, A., Brocker, T., Luton, F. and Malissen, B. (1992) The T cell receptor/CD3 complex is composed of at least two autonomous transduction modules. *Cell* **68**, 83-95.

Weiner, A.M., Deininger, P.L. and Efstratiadis, A. (1986) Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu Rev Biochem* **55**, 631-61.

Wenger, R.H., Kieffer, N., Wicki, A.N. and Clemetson, K.J. (1988) Structure of the human blood platelet membrane glycoprotein Ib alpha gene. *Biochem Biophys Res Commun* **156**, 389-95.

Wesselborg, S., Janssen, O., Pechhold, K. and Kabelitz, D. (1991) Selective activation of gamma/delta + T cell clones by single anti-CD2 antibodies. *J. Exp. Med.* **173**, 297-304.

Wetzel, R. (1988) Active immunoglobulin fragments synthesized in E. coli--from Fab to scantibodies. *Protein Eng* **2**, 169-70.

Wigzell, H. and Hayry, P. (1974) Specific fractionation of immunocompetent cells. Application in the analysis of effector cells involved in cell mediated lysis. *Curr Top Microbiol Immunol* **67**, I-42.

Williams, A.F. (1991) Cellular interactions. Out of equilibrium. *Nature* **352**, 473-4.

Williams, A.F. (1987) A year in the life of the immunoglobulin superfamily. *Immunol. Today* **8**, 298-303.

Williams, A.F. and Barclay, A.N. (1988) The immunoglobulin superfamily--domains for cell surface recognition. *Annu Rev Immunol* **6**, 381-405.

Williams, A.F. and Barclay, A.N. (1986) Glycoprotein antigens of the lymphocyte surface and their purification by antibody affinity chromatography. In Weir, D.M. (ed.), *Handbook of Experimental Immunology.* Blackwell Scientific Publications, Oxford, Vol. 1, pp. 22.1-22.24.

Williams, A.F., Barclay, A.N., Clark, S.J., Paterson, D.J. and Willis, A.C. (1987) Similarities in sequences and cellular expression between rat CD2 and CD4 antigens. *J. Exp. Med.* **165**, 368-80.

Williams, A.F., Davis, S.J., He, Q. and Barclay, A.N. (1989) Structural diversity in domains of the immunoglobulin superfamily. *Cold Spring Harb Symp Quant Biol* **2**, 637-47.

Williams, A.F., Galfre, G. and Milstein, C. (1977) Analysis of cell surfaces by xenogeneic myeloma-hybrid antibodies: differentiation antigens of rat lymphocytes. *Cell* **12**, 663-73.

Williams, L.T. (1989) Signal transduction by the platelet-derived growth factor receptor. *Science* **243**, 1564-70.

Winship, P.R. (1989) An improved method for directly sequencing PCR amplified material using dimethylsulphodie. *Nuc. Acid Res.* **17**, 1266.

Winter, G. and Milstein, C. (1991) Man-made antibodies. *Nature* **349**, 293-9.

Wong, R.C., Remold, O.E., Vercelli, D., Sancho, J., Terhorst, C., Rosen, F., Geha, R. and Chatila, T. (1990a) Signal transduction via leukocyte antigen CD43 (sialophorin). Feedback regulation by protein kinase C. *J. Immunol.* **144**, 1455-60.

Wong, Y.W., Williams, A.F., Kingsmore, S.F. and Seldin, M.F. (1990b) Structure, expression, and genetic linkage of the mouse BCM1 (OX45 or Blast-1) antigen. Evidence for genetic duplication giving rise to the BCM1 region on mouse chromosome 1 and the CD2/LFA3 region on mouse chromosome 3. *J. Exp. Med.* **171**, 2115-30.

Woodgett, J.R., Gould, K.L. and Hunter, T. (1986) Substrate specificity of protein kinase C. Use of synthetic peptides corresponding to physiological sites as probes for substrate recognition requirements. *Eur. J. Biochem* **161**, 177-84.

Woollett, G.R., Williams, A.F. and Shotton, D.M. (1985) Visualisation by low-angle shadowing of the leucocyte-common antigen. A major cell surface glycoprotein of lymphocytes. *EMBO J.* **4**, 2827-30.

Wreschner, D.H., Hareuveni, M., Tsarfaty, I., Smorodinsky, N., Horev, J., Zaretsky, J., Kotkes, P., Weiss, M., Lathe, R., Dion, A. and et, a.l. (1990) Human epithelial tumor antigen cDNA sequences. Differential splicing may generate multiple protein forms. *Eur. J. Biochem.* **189**, 463-73.

Wuthrich, K. (1989) Protein structure determination in solution by Nuclear Magnetic Resonance spectroscopy. *Science* **243**, 45-50.

Xia, M.Q., Tone, M., Packman, L., Hale, G. and Waldmann, H. (1991) Characterization of the CAMPATH-1 (CDw52) antigen: biochemical analysis and cDNA cloning reveal an unusually small peptide backbone. *Eur. J. Immunol.* **21**, 1677-84.

Yagita, H., Nakamura, T., Asakawa, J., Matsuda, H., Tansyo, S., Iigo, Y. and Okumura, K. (1989) CD2 expression in murine B cell lineage. *Int Immunol* **1**, 94-8.

Yang, P., Yin, X. and Rutishauser, U. (1992) Intercellular space is affected by the polysialic acid content of NCAM. *J. Cell Biol.* **116**, 1487-1496.

Yang, S.Y., Chouaib, S. and Dupont, B. (1986) Common pathway for T lymphocyte activation involving both the CD3-Ti complex and CD2 sheep erythrocyte receptor determinants. *J. Immunol.* **137**, 1097-100.

Yednock, T.A., Cannon, C., Fritz, L.C., Sanchez-Madrid, F., Steinman, L. and Karin, N. (1992) Prevention of experimental autoimmune encephalomyelitis by antibodies against α4β1 integrin. *Nature* **356**, 63-66.

Yogeeswaran, G. and Salk, P.L. (1981) Metastatic potential is positively correlated with cell surface sialylation of culture murine tumour cell lines. *Science* **212**, 1514-1516.

Yokoyama, S., Staunton, D., Fisher, R., Amiot, M., Fortin, J.J. and Thorley, L.D. (1991) Expression of the Blast-1 activation/adhesion molecule and its identification as CD48. *J. Immunol.* **146**, 2192-200.

Zhou, Q., Moore, K.L., Smith, D.F., Varki, A., McEver, R.P. and Cummings, R.D. (1991) The selectin GMP-140 binds to sialylated, fucosylated lactosaminoglycans on both myeloid and nonmyeloid cells. *J Cell Biol* **115**, 557-64.

Zhu, B.C. and Laine, R.A. (1985) Polylactosamine glycosylation on human fetal placental fibronectin weakens the binding affinity of fibronectin to gelatin. *J. Biol. Chem.* **260**, 4041-5.

Zoccola, D., Kubar, J., Deckert, M. and Bernard, A. (1991) The participation of E2 in T-cell rosettes. *Faseb J* **5**, A614 Abs 1414.

Structure of domain 1 of rat T lymphocyte CD2 antigen

Paul C. Driscoll*, Jason G. Cyster†, Iain D. Campbell*, & Alan F. Williams†

* Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK
† MRC Cellular Immunology Research Unit, Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK

THE CD2 antigen is largely restricted to cells of the T-lymphocyte lineage and has been established as an important adhesion molecule in interactions between human T lymphocytes and accessory cells[1]. In the adhesion reaction, CD2 on T cells binds to LFA-3 on other cells, with binding through domain 1 of CD2 (ref. 2). CD2 can also be a target for the delivery of mitogenic signals to T lymphocytes cultured with combinations of anti-CD2 antibodies[3,4]. Two predictions that are contradictory have been made for the structure of CD2 domain 1. One suggests an immunoglobulin (Ig) fold, on the basis of sequence patterns conserved in the Ig-superfamily (IgSF)[5], whilst the other proposes a pattern of alternating α-helices and β-strands, on the basis of secondary structure predictions[6]. Thus CD2 domain 1 is an important test case for the validity of IgSF assignments based on sequence patterns. We report here the expression of domain 1 of rat CD2 in an *Escherichia coli* expression system and have determined a low-resolution solution structure by NMR spectroscopy.

A schematic model proposed for the domain structure of CD2 is shown in Fig. 1*a* together with that for CD4 whose extracellular region may have been derived in evolution from a CD2-like structure by gene duplication of a two-domain segment[7,8]. Domain 1 of rat CD2 was considered to encompass amino-acid residues 1–99 and this sequence was expressed as a fusion protein with glutathione-S-transferase (GST) of the blood fluke *Schistosoma mansoni*[9] in *E. coli* with the thrombin cleavage sequence, LVPRGS, between GST and CD2 domain 1. Thrombin cleaves after the Arg(R) residue in this sequence and thus the cleaved product contains the extra residues Gly Ser at the amino terminus of the CD2 domain 1 sequence. The fusion protein was expressed at 40 mg l^{-1} and was soluble in the lysate of *E. coli* cells, allowing purification on a glutathione agarose column (Fig. 1*b*, lanes 1–3). Cleavage by thrombin was effective and the CD2 domain was purified by affinity chromatography followed by gel filtration (Fig. 1*b*, lanes 4, 5 and 6). The CD2 domain was judged to be folded correctly because it bound to two noncompetitive monoclonal antibodies against rat CD2 (OX34 and OX55) (ref. 4) with an affinity equal to that of the extracellular domain of CD2 expressed in Chinese hamster ovary cells (data not shown). The fusion protein was also expressed in *E. coli* cells grown with $^{15}NH_4Cl$ as the sole nitrogen source to yield uniformly ^{15}N-labelled CD2 domain 1, suitable for multidimensional heteronuclear (NMR) studies.

The CD2 domain 1 preparations yielded excellent NMR spectra and the assignment of the ^1H resonances to individual amino acids was done using a combination of three-dimensional ^1H nuclear Overhauser enhancement ^{15}N-^1H heteronuclear multiple quantum coherence (NOESY-HMQC) and ^1H homonuclear Hartmann-Hahn-^{15}N-^1H-HMQC (HOHAHA-HMQC) NMR spectroscopy[10,11]. Side-chain assignments were obtained using two-dimensional ^1H NMR of an unlabelled sample of the protein in D_2O solution.

Three-dimensional structures for CD2 domain 1 were calculated on the basis of the NOE data, slowly exchanging amide NH groups and $^3J_{NH\alpha}$ spin-spin coupling constants. The calculations were done using the X-PLOR program based on dynamical simulated annealing[12,13]. In total, 16 independent structures were obtained that satisfied the NOE distance restraints with no violation in any individual structure greater than 0.5 Å.

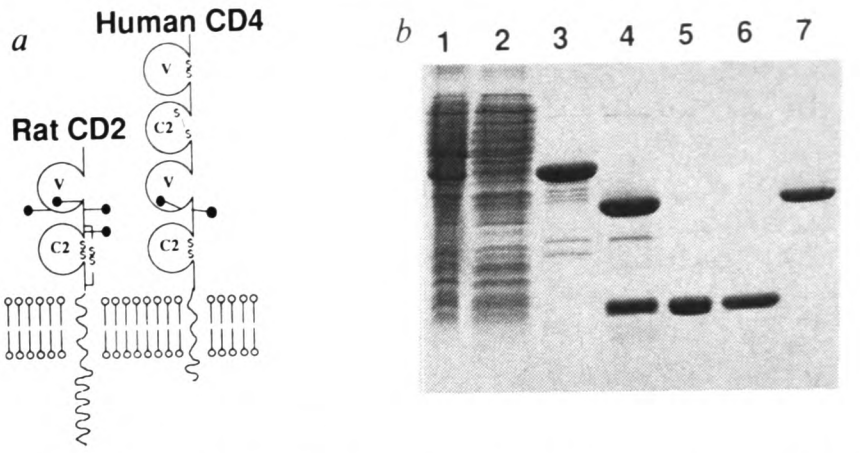


FIG. 1 *a*, Schematic representations for CD2 and CD4 at a cell surface. The circles indicate IgSF domains and the ♀ symbols, N-linked carbohydrate structures. *b*, Expression and purification of rat CD2 domain 1. Lanes: 1, soluble lysate from *E. coli* expressing the fusion protein of GST–CD2 domain 1; 2, lysate passed through a glutathione agarose affinity column; 3, purified fusion protein; 4, fusion protein cleaved with thrombin; 5 and 6, CD2 domain 1 after purification by affinity chromatography followed by gel filtration; 7, GST.

METHODS. Nucleotide sequence encoding rat CD2 protein residues 1–99 was isolated from the complementary DNA[7] by polymerase chain reaction (PCR) with the introduction of a *Bam*H1 restriction site immediately preceding the codon for protein residue 1 and a stop codon after the codon for amino-acid 99 followed by an *Eco*R1 site. The PCR product was inserted into the pGEX-2T vector and the sequence was confirmed using oligonucleotides specific for pGEX. The construct was transformed into *E. coli* BL21, a prototrophic strain. For fusion protein production, cells were grown overnight, diluted 1:10 in fresh medium and grown for 1 h at which point IPTG (isopropylthiogalactoside) was added to 0.12 mM, and the cells grown for a further 3 h. The cells were pelleted and resuspended in 1/60 volume 25 mM Tris–HCl, pH7.4, 150 mM NaCl, 5 mM EDTA, 0.2 mM PMSF, 2 mM iodoacetamide acid and sonicated for 3 × 30 s using an MSE sonicator on amplitude setting 8. Triton X-100 was added to a final concentration of 1% and the lysate spun at 30,000*g* for 30 min. The supernatant was passed over a glutathione-agarose column which was then washed with 25 mM Tris–HCl pH7.4, 140 mM NaCl and eluted with 5 mM glutathione, 50 mM Tris–HCl, pH8. The eluted material was mixed with 0.05 vol 3 M NaCl, 50 mM $CaCl_2$ and 1 NIH u of bovine thrombin per 2 mg protein. Cleavage was complete after 60 min and the CD2 domain was then purified on an OX34 monoclonal antibody affinity column[7] followed by gel filtration using a 200 ml Sephadex G75 column. For ^{15}N-labelling, cells were grown in M9 medium (10 mM $MgSO_4$, 1 mM $CaCl_2$, 0.4% glucose, 48 mM Na_2HPO_4, 22 mM KH_2PO_4, 9 mM NaCl) and $^{15}NH_4Cl$ (MSD Isotopes, Merck Frost, Canada) at 9.7 mM for ~18 h until the absorbance at 550 nm (A_{550}) was >1.0. The cells were then diluted 1:10 and grown for 2 h until A_{550} = ~0.5 when IPTG was added and growth continued for a further 3–4 h.

The structures were based on 713 interproton NOE distance restraints, 64 ϕ dihedral angle restraints and 68 H-bond distance restraints. The 16 structures are shown in Fig. 2*a*, best-fitted on the backbone atoms of residues 3–98. A restrained minimized average structure was obtained by computing the mean coordinate positions of the 16 best-fitted structures followed by restrained minimization of the same target function as that used in the final stages of the structure calculation (Fig. 2*b*). The r.m.s. deviation of the 16 structures from the restrained minimized average structure is 0.98 ± 0.11 Å for backbone atoms and 1.59 ± 0.19 Å for all atoms in the well-defined portion of the structure spanning residues 3–98. This level of definition is sufficient to determine accurately the polypeptide fold of the protein and the conformation of some of the internal side chains.

The calculated CD2 structures clearly show a fold like an immunoglobulin V-domain and in Fig. 2*c* the β strands for CD2 are overlayed with equivalent strands from CD4 (refs 14, 15) and the immunoglobulin V domains NEW V_H (ref. 16) and REI V_L (ref. 17). It can be seen that the core strands are very similar with r.m.s. differences for the three-dimensional alignments being 0.92, 1.43 and 0.95 Å for CD2 versus CD4 and the V_H and V_L domains, respectively (calculations for 28 Cα positions from β strands B, C, D, E and F).

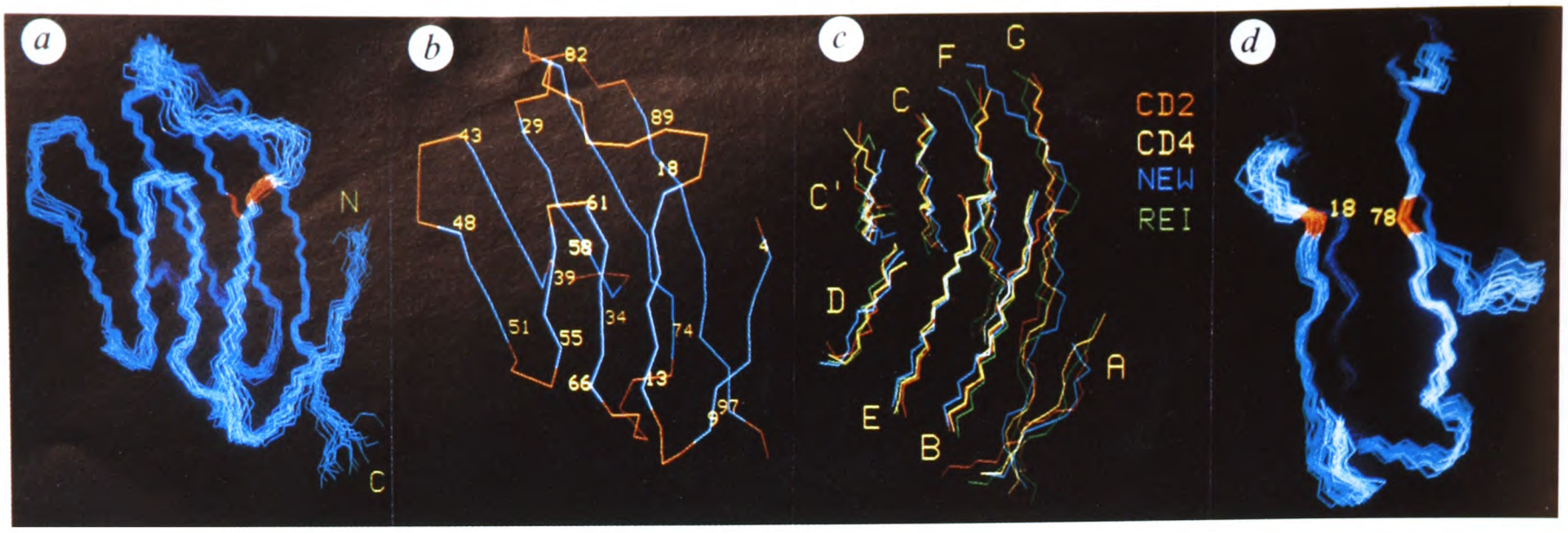


FIG. 2 *a*, Best-fit superposition of the 16 structures computed for CD2 domain 1 based on 713 NOE distance restraints (20 intraresidue, 346 sequential (residue *i*, residue *j*; $j=i+1$), 102 short range (*i*, *j*; $j \leq i+5$), 60 medium range (*i*, *j*; $j \leq i+10$) and 185 long range (*i*, *j*; $j>i+10$), 64 ϕ dihedral angle restraints and 68 H-bond distance restraints. Backbone N, C and Cα atoms of residues 5–98 are shown. NMR spectra were recorded using a 3 mM pH4.2 CD2 domain 1 sample, at 500 or 600 MHz ^{1}H frequency and at 23 °C or 30 °C. Distance restraint upper limits were categorized according to the estimated intensity of the NOE cross peak in the spectrum: strong 2.7 Å; medium 3.5 Å; weak 5.0 Å. Appropriate corrections were made to the upper limits in the cases of NOEs connecting degenerate proton resonances. Sequential distance restraints were estimated on a 10-point sliding scale from the integrated cross-peak intensities of the three-dimensional ^{15}N-^{1}H NOESY-HMQC NMR experiment (mixing time = 150 ms). In all cases the lower limit of the distance restraint was given by the van der Waals' contact distance. $^3J_{NH\alpha}$ spin–spin coupling constants were measured by lineshape fitting the traces from a two-dimensional ^{15}N-^{1}H HMQC-*J* experiment[27]. For $^3J_{NH\alpha}$ smaller than 6 Hz the corresponding ϕ torsion angle was restrained in the range −20° to −105°; for $^3J_{NH\alpha}$ larger than 7.5 Hz the ϕ angle was restrained to the range −70° to −120°. H-bonded amide NH groups were identified on the basis of the presence of an NH resonance 2 h after dissolving the protein in D_2O solution. The distance restraints used for each H-bond identified were r(HN-O) ≤ 2.3 Å, r(N-O) = 2.5–3.3 Å. Initially 30 structures were computed starting from randomized coil structures using a dynamical simulated annealing protocol[12,13] using the program X-PLOR. After two rounds of refinement in which the NOE restraints list was revised and added to, 16 structures had converged with good covalent geometry and van der Waals' contacts and no single violation greater than 0.5 Å. During the calculations the protein moves in a force field consisting of the sum of a number of terms. Bond (F_{bond}) and angle terms (F_{angle}) are used to maintain idealized covalent geometry. An improper dihedral term ($F_{improper}$) is incorporated to maintain the planarity and chirality of certain covalent groups including the peptide bond which is assumed to be planar and universally *trans* in CD2 domain 1. Square well terms are used for experimental distance (F_{NOE}) and dihedral angle restraints (F_{cdihed}). A simple quartic repulsion term (F_{vdW}) is used to represent the van der Waals' component of the target function. The N-terminal region of the protein, comprising the Gly-Ser N-leader dipeptide and residues 1 and 2, was omitted from the calculations because of the absence of structurally significant NMR data. The final values of each term of the target function for the 16 structures were: $F_{bond} = 72 \pm 4$ kcal mol^{-1}; $F_{angle} = 1842 \pm 14$ kcal mol^{-1}; $F_{improper} = 100 \pm 4$ kcal mol^{-1}; $F_{NOE} = 168 \pm 18$ kcal mol^{-1}; $F_{cdihed} = 4.5 \pm 2$ kcal mol^{-1}; $F_{vdW} = 59 \pm 8$ kcal mol^{-1}; $F_{total} = 2{,}247 \pm 38$ kcal mol^{-1}. For each of these terms the final force constant used was identical to that given in ref. 12. The hard sphere van der Waals' radii were set to 0.8 of the standard values. Using the CHARMM[28] empirical energy function to measure a Lennard–Jones van der Waals' energy for each structure gave a value $E_{L-J} = -243 \pm 12$ kcal mol^{-1}, indicating good nonbonded contacts. Note that this Lennard–Jones term is not used as part of the target function in the structure calculations. The rms deviations from the experimental restraints and idealized covalent geometry are small: interproton distances (781) 0.085 ± 0.004 Å; experimental dihedral constraints (64) 1.047 ± 0.266 degree; bonds (1552) 0.010 ± 0.0003 Å; angles (2979) 2.078 ± 0.008 degree; improper dihedrals (634) 1.023 ± 0.018 degree. *b*, The backbone Cα trace of the mean structure for CD2 domain 1 computed from the 16 best-fitted structures by coordinate averaging and energy minimization. The β strands are shown in blue, loops in red. *c*, The β sheets of CD2 superimposed with those of CD4 and REI V_L and NEW V_H. The structures were aligned with each other by matching 28 Cα positions of the core residues of β strands B, C, D, E and F. *d*, A cutaway view of the superimposed CD2 structures showing the Ile-18 and Val-78 residues (with Cα atoms coloured red) that are at the positions of the conserved IgSF disulphide bond in antibody V-domains. The separation of the Cα atoms of these residues is the range observed for disulphide links in IgSF domains.

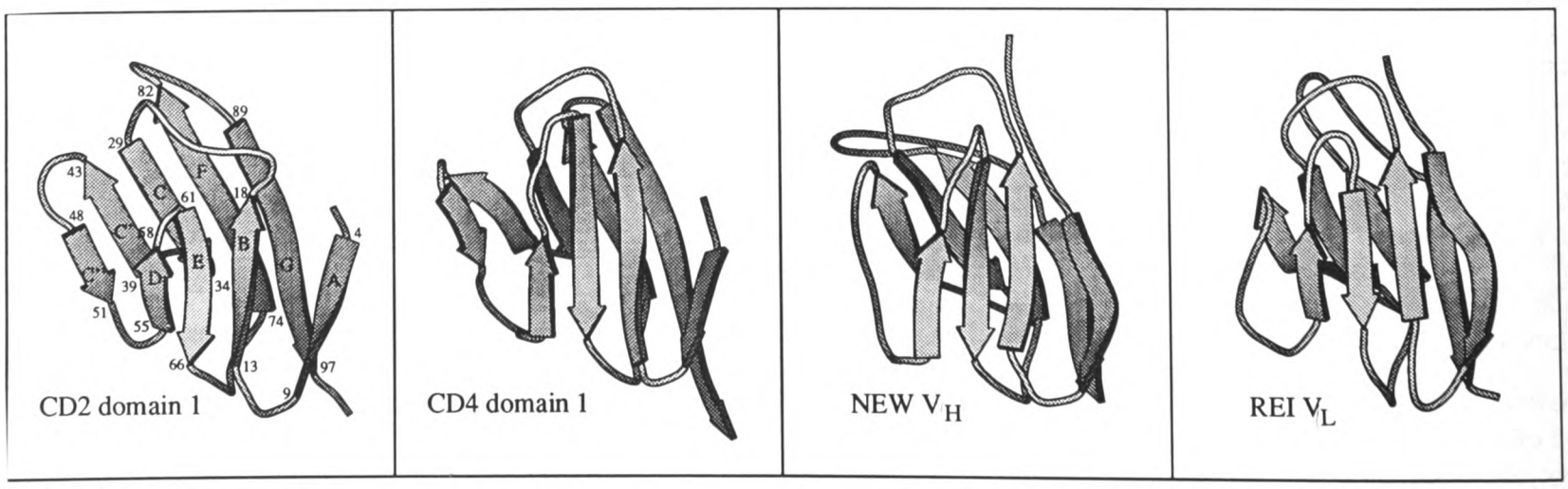


FIG. 3 Illustrations of the folds for CD2 domain 1, CD4 domain 1, immunoglobulin NEW V_H and immunoglobulin REI V_L. The coordinates for CD4 domain 1 were kindly provided by S. Harrison and those for the V domains were from the Brookhaven database. The diagrams were computed using the MOLSCRIPT program[29].

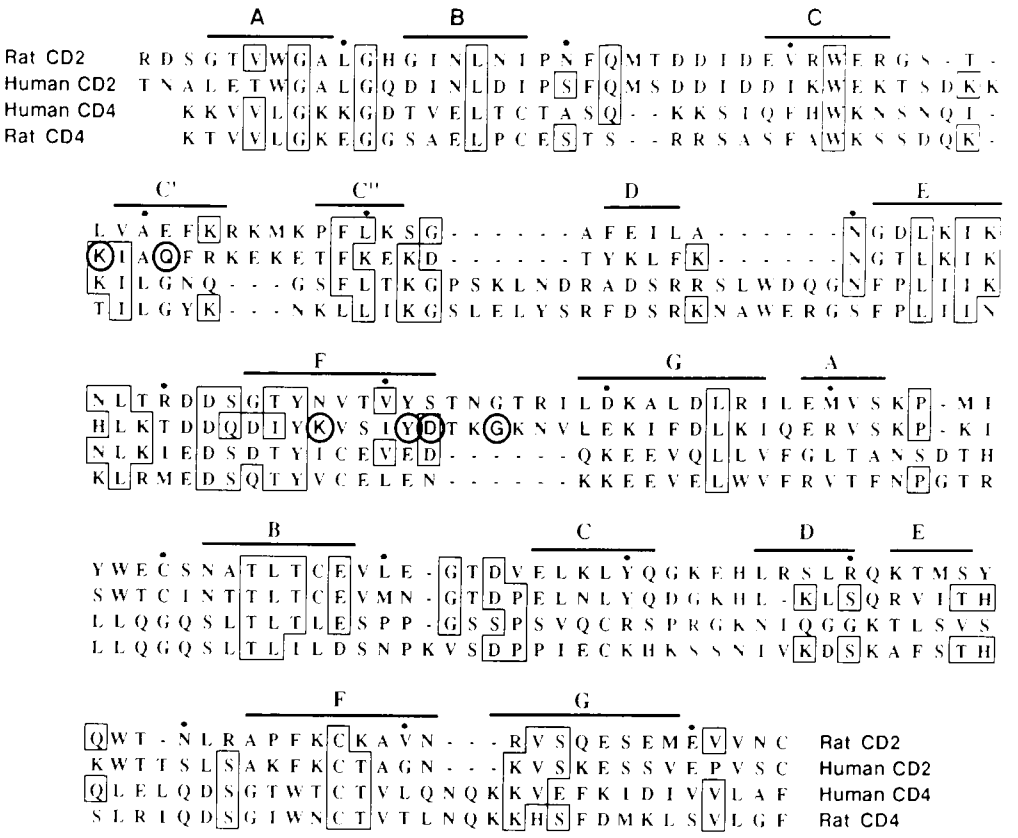


FIG. 4 Sequence alignments of CD2 and CD4 domains 1 and 2. Human and rat sequences are shown derived from the Swissprot database. The first three amino acids of human CD2 are not shown and the black dots above the rat CD2 sequence indicate each tenth residue of that sequence. The β-strand assignments have been made on the basis of the rat CD2 structure for the first domains and on the basis of the CD4 structure for the second domains. In the domain 1 sequences, the positions of CD4 β strands are also correctly identified by the bars but their extent is not exactly shown in all cases. Residues in human CD2 that are candidates for involvements in LFA-3 binding are circled and include human residues 43, 46, 82, 86, 87 and 90. These data are from mutational analysis[19] (note residue 43 is equivalent to 48 in ref. 19 where the amino terminus was assigned five residues before the correct position). The criteria for considering a residue as a candidate for binding is that there should be a single mutation, not involving a Pro residue, that leads to the loss of LFA-3 binding and that the equivalent residue in the rat structure (38, 41, 77, 81, 82 and 85) is accessible to solvent. All of these residues except rat Ser 82 are exposed on the same β-sheet face.

Ile 18 and Val 78 are the residues of CD2 domain 1 corresponding to the conserved disulphide bond of IgSF domains[5] and these residues are highlighted in Fig. 2*d*. In the 16 structures the Cα atoms of these residues are separated by 7.0 ± 0.2 Å. In IgSF domains the Cα–Cα distance of the Cys residues forming the conserved disulphide bond is in the range 5.6–7.4 Å (ref. 18). In studies before the determination of the CD2 stucture, cysteine residues were substituted by site-directed mutagenesis at Ile 18 and Val 78 and the mutant CD2 was expressed in a secreted two-domain form in Chinese hamster ovary cells. In this CD2 mutant a disulphide bond was formed (F. Gray, J.G.C., T. Willis, A.F.W. manuscript in preparation).

When the full CD2 domain is compared with CD4 domain 1 and the V-domains, differences in the regions connecting the β-strands are evident. Figure 3 shows comparisons of the folds for the four domains and the sequence comparisons for CD2 and CD4 are shown in Fig. 4. After the Ile-18 residue of β-strand B in CD2, there is a Pro residue and the CD2 chain folds immediately into a long, kinked loop connecting to β-strand C. In the V domains and CD4, β-strand B extends further and the loop joining strands B and C is not distorted. In CD2, βstrands D and E are truncated in comparison with the same regions in the other domains. This seems to be required to accommodate the kink in the loop between strands B and C since there are many internal side-chain contacts between the loops connecting strands B to C and D to E. These contacts plus the unusual structure for the B to C loop may explain the fact that sequence in this region is highly conserved between human and rat CD2. By contrast, in the same region of CD4 there is little conservation of sequence between the species (Fig. 4). In CD2 the loop between β-strand F and G is longer than in CD4 and is more similar to the same region in the V-domains where this loop constitutes hypervariable region 3.

The structure for CD4 (refs 14, 15) is unusual in comparison with immunoglobulin chains in that domain 2 follows on directly from domain 1 without any hinge residues such as are found between immunoglobulin V domains and the following C_L or C_{H1} domains. From Fig. 4 it is probable that the juxtaposition of domains in CD2 will be like that in CD4 as the distance of the β-strand G of domain 1 from β-strand B of domain 2 is one amino acid less in the CD2 sequence than in CD4.

The β-strand assignments for human CD2 domain 1 can be made by alignment with the newly solved rat CD2 domain 1 structure (Fig. 4). All the LFA-3 binding activity resides in domain 1 of CD2 (ref. 2) and mutagenesis of human CD2 has indicated regions of contact with LFA-3 (ref. 19). Residues which appear to be involved in this interaction are marked in Fig. 4. These are positioned in β-strand C' and in the region of β-strands F to G. This suggests that the face of the β-sheet C, C', C'', F, G may be involved in the CD2:LFA-3 interaction. This can be further investigated by mutation analysis on the basis of the known three-dimensional structure.

The structure of CD2 domain 1 fits closely with that predicted from sequence on the basis of superfamily considerations with β-strands B, C, E and F that make up the core of the fold being exactly predicted[5,7]. The proposal from secondary structure predictions[2,6] is incorrect. CD2 domain 1 and CD4 domain 2 (refs 14, 15) were test cases for IgSF arguments because both sequences had unusual features which lead to contradictions between predictions from superfamily sequence patterns[8] and *ab initio* computer calculations based on the primary sequences. The superfamily arguments were correct in both cases and this suggests that other IgSF domains predicted on the basis of conserved sequence patterns will also be correct[20].

The structure for CD2 domain 1 adds significantly to known variants of the immunoglobulin-fold as it is the first structure for a domain involved in cell adhesion and provides a prototype for IgSF domains that lack the conserved disulphide bond. Structures determined previously include immunoglobulin V and C domains[21], MHC class 1 α3 domain[22] and β-2 microglobulin[22,23] and CD4 domains 1 and 2 (refs 14, 15). Also immunoglobulin-like folds have been seen in domains 1 and 2 of the PapD bacterial chaperone protein[24], but these should not be categorized as IgSF domains as they show no detectable sequence similarity to IgSF domains (ref. 25 and A.F.W., unpublished data). CD2 domain 1 typifies IgSF domains that lack the conserved disulphide bond. Such domains are found at a moderate frequency among domains of cell-surface molecules (for example LFA-3 domain 1, CEA domain 1, CD4 domain 3, PDGF receptor domain 4 (ref. 20)) but are the main type in the myosin-binding muscle proteins that are in the IgSF (ref. 26).

Received 4 July; accepted 23 August 1991.

1. Springer, T. A. *Nature* **346,** 425–434 (1990).
2. Recny, M. A., Neidhardt, E. A., Sayre, P. H., Ciardelli, T. L. & Reinherz, E. L. *J. biol.. Chem.* **265,** 8542–8549 (1990).
3. Meuer, S. C. *et al. Cell* **36,** 897–906 (1984).
4. Beyers, A. D., Barclay, A. N., Law, D. A., He, Q. & Williams, A. F. *Immun. Rev.* **111,** 59–77 (1989).
5. Killeen, N., Moessner, R., Arvieux, J., Willis, A. & Williams, A. F. *EMBO J.* **7,** 3087–3091 (1988).
6. Clayton, L. K., Sayre, P. H., Novotny, J. & Reinherz, E. L. *Eur. J. Immun.* **17,** 1367–1370 (1987).
7. Williams, A. F., Barclay, A. N., Clark, S. J., Paterson, D. J. & Willis, A. C., *J. exp. Med.* **165,** 368–380 (1987).
8. Williams, A. F.,Davis, S. J., He, Q. & Barclay, A. N. *Cold Spring Harb. Symp. quant. Biol.* **54,** 637–647 (1989).
9. Smith, D. B. & Johnson, K. S. *Gene* **67,** 31–40 (1988).
10. Marion, D. *et al. Biochemistry* **28,** 6150–6156 (1989).
11. Driscoll, P. C., Clore, M., Marion, D., Wingfield, P. T. & Gronenborn, A. M. *Biochemistry* **29,** 3542–3556 (1991).
12. Clore, G. M. & Gronenborn, A. M. *CRC crit. Rev. Biochem. molec. Biol.* **24,** 479–564 (1989).
13. Brünger, A. T. *X-PLOR Manual, Version 2.1* (Yale University Press, 1990).
14. Wang, J. *et al. Nature* **348,** 411–418 (1990).
15. Ryu, S-E. *et al. Nature* **348,** 419–426 (1990).
16. Saul, F. A., Amzel, M. & Poljak, R. J. *J. biol. Chem.* **253,** 585–597 (1978).
17. Epp, O., Lattman, E. E., Schiffer, M., Huber, R. & Palm, W. *Biochemistry* **14,** 4943–4952 (1975).
18. Richardson, J. S. *Adv. Protein Chem.* **34,** 167–339 (1981).
19. Peterson, A. & Seed, B. *Nature* **329,** 842–846 (1987).
20. Williams, A. F. & Barclay, A. N. *A. Rev. Immun.* **6,** 381–405 (1988).
21. Amzel, L. M. & Poljak, R. J. *A. Rev. Biochem.* **48,** 961–967 (1979).
22. Bjorkman, P. J. *et. al. Nature* **329,** 506–512 (1987).
23. Becker, J. W. & Reeke, G. N. *Proc. natn. Acad. Sci. U.S.A.* **82,** 4225–4229 (1985).
24. Holmgren, A. & Branden, C.-I. *Nature* **342,** 248–251 (1989).
25. Freeman, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **87,** 8810–8814 (1990).
26. Labiet, S. *et al. Nature* **345,** 273–276 (1990).
27. Kay, L. E. & Bax, A. *J. Magn. Reson.* **86,** 110–126 (1990).
28. Brooks, B. R. *J. comput. Chem.* **4,** 187–217 (1983).
29. Kraulis, P. J. *J. appl. Crystallogr.* (in the press).

ACKNOWLEDGEMENTS. We thank J. Boyd, T. Harvey and N. Soffe from the Oxford Department of Biochemistry for helpful discussions and S. Harrison, Harvard University for supplying the CD4 coordinates. I.D.C. is a member of the Oxford Centre for Molecular Sciences supported by the SERC and MRC. P.C.D. is supported by a Royal Society University Research Fellowship and J.G.C. by a Commonwealth Overseas Studentship.