Functional Analysis of Myeloid Cell Antigens

by

Elma Zaven Tchilian

A thesis submitted for the degree of Doctor of Philosophy in the University of London

Department of Haematology, Faculty of Sciences, University College London

February 1994
ABSTRACT

Based on their structures and homologies to other adhesion receptors, the lymphomyeloid progenitor cell antigen, CD34, the myelomonocytic cell antigen, CD33, and the mature myeloid cell surface molecule, BGPC, are candidates to mediate specific cell-extracellular matrix and/or cell-cell adhesion. This thesis describes the analysis of adhesion of the CD34, CD33 and BGPC molecules to stromal layers formed by long-term bone marrow cultures, extracellular matrix components and various cell lines. BGPC mediated hemophilic aggregation of OHO cells transfected with BGPC cDNA as well as binding to carcinoembryonic antigen (CEA) and perhaps another cell surface receptor on HL-60 cells. The CD33 and CD34 defined antigens do not appear to be involved in adhesion to stromal layers; to the extracellular matrix components fibronectin, fibrinogen, laminin, collagen I, III, IV, IX and X and hyaluronic acid; or to other cell surface receptors expressed on Raji, Daudi, MOLT4, CEM, U937, HL-60, KG1 and peripheral blood mononuclear cells.

This work also describes the molecular cloning of a murine homologue of the human CD33 myeloid antigen. Two cDNA clones, differing by an 83 nucleotide insertion in the cytoplasmic region were isolated. The insertion generated a shift in the reading frame within the cytoplasmic tail, resulting in two mouse CD33 isoforms, m33-A and m33-B, with distinct cytoplasmic domains and with predicted protein core molecular weights of 37 kD and 45 kD, respectively. The cDNAs and deduced amino acid sequences show extensive similarity to the human CD33 sequence with the highest homology in the first and second Ig-like domains (61% amino acid identity). The most significant divergence between the human and murine proteins occurs in their cytoplasmic portions. Murine CD33 mRNAs were detected in bone marrow, spleen, thymus, brain, liver and in the multipotential progenitor cell line, A4, the myelomonocytic cell line, WEHI 3B, the myeloid cell line, M1 and the macrophage cell line, P388, by Northern analysis. The expression pattern of the murine CD33 homologue suggests that the function of CD33 antigen in haemopoiesis may be conserved between man and mouse.
To Mark Noble, whose initial support opened up opportunities for me to do a PhD and to Peter Beverley, without whom this work would never have been completed.
ACKNOWLEDGMENTS

I am indebted to Peter Beverley for advice, friendship and freedom of expression while working in his laboratory and to the Imperial Cancer Research Fund for supporting this project and my studentship. I would like to thank Sue Watt for her supervision during the studies described in chapter 3, which were carried out in the Medical Oncology Unit, St. Bartholomew's Hospital. I am grateful to Bryan Young, Hans Stauss, Alastair Reith and Barbara Pym for helpful discussions and technical advice about the work described in the rest of the thesis, which was done at the Tumour Immunology Unit, University College London Medical School. I am also grateful to all those mentioned in the text for their assistance and to Ray Hicks and Lorna Stewart for proofreading this manuscript. Many thanks are due to my colleagues for their help, friendship, encouragement and cheerfulness, especially to Maria Dahl, Elaine Dorey, Lindsey Goff, Sarah Murdoch, Lorna Stewart and Harry White.

Finally I thank my family and husband for their patience and support throughout the thesis.
CONTENTS

Title page 1
Abstract 2
Dedication 3
Acknowledgments 4
Contents 5
List of Tables 10
List of Figures 11
Abbreviations 13

Chapter 1: Introduction 15
1.1. General introduction 15
1.2. Bone marrow microenvironment 17
  1.2.1. Stromal cells 17
  1.2.2. Extracellular matrix 20
  1.2.3. Haemopoietic growth factors 23
    (a) functional diversity of haemopoietic growth factors 23
    (b) growth factor receptors and signal transduction 26
    (c) soluble growth factor receptors 28
1.3. Adhesion molecules 29
  1.3.1. Integrins 30
    (a) structure and function 30
    (b) signal transduction by integrins 32
  1.3.2. Adhesion molecules of the immunoglobulin superfamily 35
  1.3.3. Selectins 39
    (a) structure and function 39
    (b) rolling and tight adhesion: the cooperative roles of selectins and integrins 42
  1.3.4. Cadherins 43
1.5. CD34 defined antigen 44
1.6. CD33 defined antigen 47
1.7. CD66 defined antigens. BGPC splice variant 49
1.8. Aims of the thesis 53

Chapter 2: Materials and methods 55
2.1. Materials 55
2.1.1. Biochemical reagents and materials
2.1.2. Antibodies
2.1.3. Oligonucleotide primers
2.1.4. Cell preparation and tissues
   (a) cell lines
   (b) isolation of peripheral blood mononuclear cells and neutrophils
   (c) murine tissues
2.1.5. Suppliers addresses

2.2. Methods
   2.2.1. Preparation of plasmid DNA
      (a) small-scale preparation of plasmid DNA
      (b) large-scale preparation of plasmid DNA
      (c) restriction endonuclease digestion of plasmid DNA
   2.2.2. Production of CHO cell transfectants, stably expressing CD34, CD33 and BGPc proteins
   2.2.3. Flow cytometric analysis and cell sorting
   2.2.4. Western blot analysis of CHO cell transfectants
      (a) lysing CHO cells
      (b) SDS - polyacrylamide gel electrophoresis
      (c) western blot analysis
   2.2.5. Production of CD34-Fc and CD33-Fc fusion proteins
      (a) CD34-Fc and CD33-Fc constructs
      (b) transfection of Cos 7 cells
      (c) purification and characterisation of CD34-Fc and CD33-Fc fusion proteins
      (d) production of labelled CD33-Fc and CD34-Fc fusion proteins
   2.2.6. Binding to stromal layers
      (a) Long Term Bone Marrow Cultures (LTBMC)
      (b) binding of \(^{51}\text{Cr}\)-labelled CHO cells to stromal layers
      (c) binding of CD34-Fc and CD33-Fc fusion proteins to stromal layers
   2.2.7. Binding to ECM substrates
      (a) binding of \(^{51}\text{Cr}\)-labelled CHO cells to ECM substrates
      (b) binding of CD34-Fc and CD33-Fc fusion proteins to ECM substrates
   2.2.8. Binding to bone marrow ECM extract
2.2.9. Cell-cell binding studies
(a) aggregation assay
(b) immunofluorescence
(c) binding of $^{125}$I-labelled CEA to CHO monolayers
(d) binding of $^{51}$Cr-labelled cells to CHO monolayers
(e) binding of cells to Cos cell transfectants

2.2.10. Preparation of total RNA

2.2.11. RT-PCR analysis of mouse bone marrow and WEHI 3B cells

2.2.12. Cloning of PCR products

2.2.13. Sequence analysis

2.2.14. Computer analysis

2.2.15. Southern blot analysis
(a) preparation of radiolabelled DNA probe
(b) isolation of genomic DNA
(c) electrophoresis, transfer and hybridisation of restriction fragments of genomic DNA

2.2.16. Mouse bone marrow cDNA library screening
(a) plating of bacteriophage $\lambda$gt11
(b) hybridisation of bacteriophage plaques
(c) rapid PCR analysis of positive bacteriophages
(d) rescreening of the mouse bone marrow cDNA library

2.2.17. Cloning of bacteriophage DNA into pBluescript SK vector (p33-A and p33-B clones)
(a) preparation of plate lysate stock
(b) extraction of bacteriophage DNA
(c) cloning of bacteriophage DNA into pBluescript vector

2.2.18. 5' RACE

2.2.19. Northern blot analysis

2.2.20. In situ hybridisation

2.2.21. Immunohistochemical localisation of human CD33 antigen in thymus

2.2.22. PCR analysis of murine CD33 splice variants
Chapter 3: Binding studies of CD34 and CD33 defined antigens and of BGPC molecule

3.1. Introduction 101
3.2. Results 103
3.2.1. Production and characterisation of CHO cell transfectants and of soluble fusion proteins 103
(a) CHO cells, expressing CD34, CD33 and BGPC proteins 103
(b) production of CD34-Fc and CD33-Fc soluble fusion proteins 107
3.2.2. Binding to stromal layers 110
3.2.3. Binding to ECM components and bone marrow ECM extracts 112
(a) binding to ECM substrates 112
(b) binding to bone marrow ECM extract 114
3.2.4. Cell-cell interactions 116
(a) homophilic binding studies 116
(b) binding of $^{125}$I-labelled CEA to CHO cell transfectants 119
(c) binding of cells to CHO cell transfectants 121
(d) binding of cells to Cos cell transfectants 123
3.3. Discussion 125

Chapter 4: Molecular cloning of two isoforms of the murine homologue of myeloid CD33 antigen

4.1. Introduction 128
4.2. Results 130
4.2.1. Cloning approach 130
(a) PCR cloning 130
(b) library screening 132
(c) isolation of 5' end of murine CD33 gene 137
4.2.2. Nucleotide sequence 139
4.2.3. Predicted amino acid sequence 143
4.2.4. Homology to human CD33 antigen 146
4.3. Discussion 148

Chapter 5: Sequence comparisons of murine CD33

5.1. Introduction 151
5.2. Results 152
5.3. Discussion 155
<table>
<thead>
<tr>
<th>Chapter 6: Expression of murine CD33 antigen and analysis of splice variants</th>
<th>165</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1. Introduction</td>
<td>165</td>
</tr>
<tr>
<td>6.2. Results</td>
<td>166</td>
</tr>
<tr>
<td>6.2.1. Expression of murine CD33</td>
<td>166</td>
</tr>
<tr>
<td>(a) northern blot analysis</td>
<td>166</td>
</tr>
<tr>
<td>(b) in situ hybridisation analysis</td>
<td>169</td>
</tr>
<tr>
<td>6.2.2. Analysis of splice variants</td>
<td>172</td>
</tr>
<tr>
<td>6.3. Discussion</td>
<td>176</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 7: General discussion and future perspectives</th>
<th>179</th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>190</td>
</tr>
<tr>
<td>Publications</td>
<td>224</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Chapter 1
Table 1.1: Examples of association of growth factors with extracellular matrix 25
Table 2.1: Characteristics of selectins 41

Chapter 2
Table 2.1: Sequence of oligonucleotide primers 61
Table 2.2: Cell lines and culture medium 63

Chapter 5
Table 5.1: Protein homologies to murine CD33 153
LIST OF FIGURES

Chapter 1
Fig. 1.1: A model of haemopoietic stem cell differentiation 16
Fig. 1.2: Multifunctional interactions in the bone marrow microenvironment 18
Fig. 1.3: Schematic representation of an integrin molecule as an αβ heterodimer 31
Fig. 1.4: Schematic representation of adhesion molecules of the immunoglobulin superfamily 36
Fig. 1.5: Schematic representation of selectin molecules 40
Fig. 1.6: BGP splice variants 51

Chapter 2
Fig. 2.1: Schematic representation of the pIG vector 73

Chapter 3
Fig. 3.1: Flow-sorting of CHO cell transfectants 104
Fig. 3.2: FACScan analysis of CHO-CD33, CHO-CD34 and CHO-BGPc cell transfectants after three months in culture 105
Fig. 3.3: Western blot of CHO cell transfectants 106
Fig. 3.4: Analysis of CD34-Fc and CD33-Fc fusion proteins 108
Fig. 3.5: Antibody reactivity of CD34-Fc fusion protein 109
Fig. 3.6: Dot blot analyses of CD34-Fc and CD33-Fc fusion proteins 109
Fig. 3.7: Binding of CHO cells to stromal layers 111
Fig. 3.8: Binding of 35S-labelled CD33-Fc fusion protein to ECM substrates 113
Fig. 3.9: Binding of CD34-Fc fusion protein to ECM components 113
Fig. 3.10: Binding of CD34-Fc and CD33-Fc fusion proteins to rabbit bone marrow ECM extract 115
Fig. 3.11: Effect of EDTA and trypsin on the expressed antigens in CHO cell transfectants 117
Fig. 3.12: Aggregation of CHO-BGPc cell transfectants 118
Fig. 3.13: Binding of 125I-labelled CEA to CHO cell transfectants 120
Fig. 3.14: Binding of cell lines to CHO cell transfectants 122
Chapter 4

Fig. 4.1: RT-PCR analysis of mouse bone marrow and WEHI 3B cells

Fig. 4.2: Southern blot analysis

Fig. 4.3: Schematic representation of PCR analysis and PCR cloning and of cloning into pBluescript vector of positive λgt11 bacteriophages

Fig. 4.4: Mouse cDNA clones

Fig. 4.5: 5' RACE system

Fig. 4.6: p33-RACE/PCR clones

Fig. 4.7: Mouse CD33 clones

Fig. 4.8: The nucleotide and deduced amino acid sequences of murine CD33 cDNAs

Fig. 4.9: Protein sequences of murine CD33 isoforms

Fig. 4.10: Comparisons of the human and murine CD33 protein sequences

Chapter 5

Fig. 5.1: Comparisons between murine L-MAG and murine CD33 protein sequences

Fig. 5.2: Protein comparisons to the cytoplasmic domains of murine CD33 isoforms

Chapter 6

Fig. 6.1: Northern blot analysis

Fig. 6.2: In situ hybridisation analysis of mouse spleen

Fig. 6.3: In situ hybridisation analysis of lymph node

Fig. 6.4: Immunohistochemical analysis of human CD33 antigen in thymus

Fig. 6.5: RT-PCR analysis of murine CD33 splice variants
LIST OF ABBREVIATIONS

aa  amino acid(s)
ABTS  2,2' azino-di-[3-ethylbenzthiazoline sulphonate]
AML  acute myelogeneous leukaemia
bp  basepair(s)
BGP  biliary glycoprotein
BSA  bovine serum albumin
CD  cluster of differentiation
CEA  carcinoembryonic antigen
CFU-S  colony forming unit spleen
CNTF  ciliary neurotrophic factor
CK2  casein kinase 2
DAB  diaminobenzidine
DMEM  Dulbecco's modified Eagle's medium
ECM  extracellular matrix
EGF  epidermal growth factor
Epo  erythropoietin
FACS  fluorescence activated cell sorter
FAK  focal adhesion kinase
FCS  foetal calf serum
FGF  fibroblast growth factor
FITC  fluorescein isothiocyanate
GAG  glycosaminoglycans
GAPDH  glyceraldehyde phosphate dehydrogenase
G-CSF  granulocyte-colony stimulating factor
GM-CSF  granulocyte-macrophage colony stimulating factor
GPI  glycosyl phosphatidyl inositol
hCD33  human CD33
HRP  horseradish peroxidase
HSPG  heparan sulphate proteoglycan
ICAM  intercellular adhesion molecule
IFN  interferon
Ig  immunoglobulin
IL  interleukin
IMDM  Iscove's modified Dulbecco's medium
IPTG  isopropyl β-D-thiogalactopyranoside
kb kilobase(s)
kD kilodaltons
LFA leucocyte functional antigen
LIF leukaemia inhibitory factor
LTBMC long-term bone marrow culture
m33 murine CD33
MAG myelin-associated glycoprotein
MAP mitogen activated protein
M-CSF macrophage colony stimulating factor
MoAb monoclonal antibody
MW molecular weight
NCA non-specific cross-reacting antigen
NCAM neural cell adhesion molecule
OSM oncostatin M
PAGE polyacrylamide gel electrophoresis
PCR polymerase chain reaction
PDGF platelet-derived growth factor
PKC protein kinase C
PRL liver prolactin
PSG pregnancy specific glycoprotein
pfu plaque forming unit
RACE rapid amplification of cDNA ends
RT-PCR reverse transcriptase polymerase chain reaction
TGF transforming growth factor
TNF tumour necrosis factor
UAP universal amplification primer
UTR untranslated region
VLA very late antigen
β-gal β-galactosidase
Chapter 1
INTRODUCTION

1.1. General introduction

The hallmark of the haemopoietic system is a precisely controlled production of at least eight cell lineages, including neutrophils, eosinophils, basophils, monocytes and macrophages, erythrocytes, megakaryocytes, T and B lymphocytes (Fig. 1.1). At the centre of this process lies the haemopoietic stem cell, which possesses both the abilities to self-renew and to produce committed progenitors for all haemopoietic lineages. In the adult human, the renewal of stem cells and the commitment and differentiation to a particular lineage takes place in the bone marrow and these processes are regulated by critical interactions occurring between haemopoietic progenitor cells and the bone marrow microenvironment.

On their maturation most of the haemopoietic cells migrate from the bone marrow and enter the blood circulation, where they mediate a variety of functions related to immune recognition, combating infectious disease, O₂ transport and haemostasis. Leucocytes, unlike erythrocytes, must also make their way across the walls of small blood vessels and migrate into tissues to perform their functions, processes which require adherence to and migration over the surface of vascular and connective tissue cells and the extracellular matrix. Cell surface molecules, such as integrins, selectins and members of the immunoglobulin gene superfamily, are all broadly distributed and mediate adhesion during early haemopoiesis, inflammation and immune responses. Other molecules are relatively restricted to particular lineages or stages of differentiation and must therefore function in processes related to these.
Figure 1.1: A model of haemopoietic stem cell differentiation (After Zucker-Franklin et al., 1988)
This introduction will outline some of the interactions occurring between haemopoietic cells and the bone marrow microenvironment. It will also discuss the major groups of cell adhesion molecules mediating cell-cell and cell-extracellular matrix interactions as well as the growing understanding of the role of adhesion receptors in signal transduction. Knowledge of CD34, CD33 and CD66 defined antigens (in particular the BGPc molecule), which are expressed on different stages of myeloid development and whose significance in myeloid development is little understood will be reviewed.

1.2. Bone marrow microenvironment

The bone marrow microenvironment provides the supportive milieu within which haemopoietic cells proliferate and differentiate to give rise to the cellular elements of circulating blood and other tissues. This microenvironment consist of three interrelated components: stromal cells, extracellular matrix and soluble growth factors, actively interacting with each other and with haemopoietic cells (Fig. 1.2).

1.2.1. Stromal cells

The stromal cell network includes heterogeneous cell populations of endothelial cells, fibroblasts, adipocytes, macrophages and reticular cells (Dorshkind, 1990). The fundamental role of these cells was convincingly demonstrated by the long-term bone marrow culture (LTBMC) system, first described by Dexter et al. (1977). The important feature of these cultures is that sustained haemopoiesis occurs in the absence of added growth factors but is absolutely dependent on the formation of an adherent layer of bone-marrow derived stromal cells. Stromal cells support haemopoiesis by a combination of (a) cell adhesion molecules that allow the binding of haemopoietic cells, (b) the synthesis, secretion and
Figure 1.2: Multifunctional interactions in the bone marrow microenvironment
Growth factors and inhibitors, extracellular matrix molecules (ECM), progenitor and stromal cells with their respective cell adhesion molecules (CAM) interact to direct the coordinate regulation of blood cell development.
presentation of the appropriate range of growth factors, and (c) by the synthesis of extracellular matrix molecules. Numerous studies have demonstrated that a direct contact between stromal cells and haemopoietic stem cells is required for the proliferation and differentiation of haemopoietic progenitors (Bentley and Tralka, 1983; Gordon et al., 1990; Liesveld et al., 1991). However the mechanisms and the adhesion receptors mediating the interactions between haemopoietic progenitor cells and stromal cells remain poorly understood.

The presence of NCAM, an Ig-like cell adhesion molecule widely distributed throughout the central and peripheral nervous system and at the neuromuscular junction, has been identified in long-term bone marrow cultures adapted to promote lymphoid proliferation (Thomas et al., 1988). Several isoforms of NCAM resulting from alternative splicing have been described in the nervous system and several mRNA species have been found in stromal cells (Kincade, 1990). NCAM mediates cell-cell adhesion interactions via homophilic mechanisms as well as binding to heparan sulphate and heparin (Pigott and Power, 1993), but its role in the stromal environment is still unknown. Similarly the expression of another cell adhesion molecule, ICAM-1, also a member of the Ig superfamily, has been reported on marrow stromal cells in response of tumor necrosis factor alpha (TNF-α) and interferon gamma (IFN-γ) (Wawryk et al., 1989). ICAM-1 would appear to be a potential adhesive ligand on stromal cells, since ICAM-1 can bind specifically to the promonocytic cell line U937 (Makgoba et al., 1988) and a wide variety of lymphohaemopoietic cell types express LFA-1 (Miller et al., 1985; 1986), one of the ligands for ICAM-1. However so far there is no direct evidence for such a role of ICAM-1 in the context of the bone marrow microenvironment and blocking studies with CD18 antibodies suggest that the LFA-1 does not mediate
early progenitor cell attachment in the bone marrow microenvironment (Liesveld et al., 1991).

The cell adhesion receptors on haemopoietic progenitors mediating adhesion with stromal cells are also not well defined. The cell adhesion molecules expressed by blast-colony forming cells, which mediate specific interactions with stromal layers have been proposed to exist in transmembrane or phosphatidylinositol-linked forms and also to be trypsin sensitive (Gordon et al., 1990). However blocking experiments using antibodies against phosphatidylinositol-anchored structures as CD14, CD16, NCAM, and Thy-1, or antibodies against the stem cell marker CD34 do not block the binding of haemopoietic progenitor cells to stromal layers (Gordon et al., 1990).

In chronic myelogenous leukaemia early progenitors have been shown to be defective in their ability to bind to stromal monolayers (Gordon et al., 1987a). It will therefore be important to determine whether normal and leukaemic progenitors use similar mechanisms in their stromal associations and to identify the factors regulating their expression and function.

1.2.2. Extracellular matrix

An important component of the bone marrow microenvironment is the extracellular matrix (ECM), which is composed of glycoproteins, proteoglycans and glycosaminoglycans.

At least eighteen types of collagen have been described (Hay, 1991; Mayne and Brewton, 1993) and types I, III and IV have been detected in the adherent layers of LTBMC (Zuckerman and Wicha, 1983; Campbell and Wicha, 1988). The inhibition of collagen synthesis by cis-hydroxyproline inhibits stromal and haemopoietic cell production, however the mechanism underlying this effect is unclear (Zuckerman et
Laminin and fibronectin are also produced by the stroma (Campbell and Wicha, 1988), although there is little evidence that laminin is a major adhesion protein for haemopoietic cells. Fibronectin has been reported to be involved in the interactions with the haemopoietic stem cells as well as with various haemopoietic cell lines and erythroid progenitors. Thus day-12 spleen colony-forming-unit (CFU-S) attach to stromal cell-derived ECM via an interaction of the CS-1 segment of human plasma fibronectin with the VLA-4 integrin receptor (Williams et al., 1991). In addition a second adhesion mechanism involving membrane associated chondroitin sulphate proteoglycan on the side of the haemopoietic progenitor cells interacting with the heparin domain(s) of fibronectin molecules has been proposed (Minguell et al., 1992). PHA activated T cells and the haemopoietic cell lines Jurkat, Ramos and U937 all bind to CS-1 fragment of fibronectin through the VLA-4 receptor (Wayner and Kovach, 1992), while the adhesion of erythroid progenitors is inhibited by antibodies against VLA-5, which binds the RGDS peptide of fibronectin. Fibronectin adhesion is of particular importance in erythroid development, since after induction of differentiation, erythroid cells lose the ability to attach to fibronectin, suggesting that downregulation of the fibronectin receptor provides a release mechanism for erythroid cells from the bone marrow into the circulation (Patel and Lodish, 1984; 1986).

In contrast granulocytic progenitors (CFU-GM) do not bind well to fibronectin (Tsai et al., 1987), but bind to another adhesion molecule haemonectin (Campbell et al., 1987), thus implicating different adhesion pathways in the regulation of these two lineages. Developmentally regulated loss of ability to bind to haemonectin has also been observed in HL-60 cells when they are induced to differentiate with DMSO, suggesting that this could be part of the physiological release mechanism of granulocytic cells from bone marrow into circulating blood (Campbell,
1992). The receptor for haemonectin has yet to be defined or haemonectin cloned, and recently White et al. (1993) have shown that haemonectin is homologous to, and probably identical to the abundant plasma glycoprotein fetuin. Phosphorylated rat fetuin has an insulin receptor tyrosine kinase inhibitory activity (Auberger et al., 1989; Brown et al., 1992), but whether phosphorylated forms of fetuin would have cell adhesion or growth regulating activities on bone marrow cells is still unclear.

Aizawa and Tavassoli (1987) have implicated galactosyl- and mannosyl-binding sites in adherence of primitive murine stem cells to stroma, but it is not known on which ECM or cell surface glycoproteins or glycolipids these critical carbohydrate residues are expressed. Conversely, studies on human cultures have indicated that normal human CD34+ myeloid or leukaemic progenitors adhere to marrow stroma by mechanisms which do not involve galactosyl or mannosyl specificities nor fibronectin, laminin or RGDS-containing peptides (Liesveld et al., 1991).

The most complex of the ECM components are the proteoglycans. These consist of a core protein to which various glycosaminoglycans (GAG) are covalently attached. Wight et al. (1986) using [35S] sulphate- and [3H] glucosamine-labelled LTBMCs identified as the most abundant a chondroitin sulfate proteoglycan containing glycosaminoglycan chains of ~38 000 kD. Dermatan sulphate, hyaluronic acid and heparan sulphate are also present in the bone marrow microenvironment (Campbell and Wicha, 1988; Morris et al., 1991). Heparan sulphate has been shown to be essential for the binding of haemopoietic progenitors to stromal layers in vitro (Gordon et al., 1988). Hyaluronic acid could also contribute to the binding of immature haemopoietic cells to the microenvironment, presumably through CD44 antigen (Lewinsohn et al., 1990), and in murine systems the addition of CD44 antibodies to LTBMC prevented
output of myeloid cells and blocked lymphocyte growth in Whithlock-Witte type cultures (Miyake et al., 1990). Furthermore two splice variants of CD44 antigen have been identified in a various haemopoietic cell lines that may additionally regulate and contribute to the specificity of these interactions (Dougherty et al., 1991).

The adhesive interactions between cells and ECM components may play a vital role in the regulation of gene expression. Numerous examples of matrix- or adhesion induced changes in cell differentiation and/or gene expression have been observed in a variety of cell types, including monocytes (Sporn et al., 1990) and bone cells (Vukicevic et al., 1990). ECM can also regulate cell behaviour through synergistic interactions between growth factors and matrix molecules or the ECM molecules themselves may be mitogenic (Panayotou et al., 1989). Examples of such effects and the intracellular events that may transduce signals between ECM receptors and the nucleus are discussed below.

1.2.3. Haemopoietic growth factors

(a) functional diversity of haemopoietic growth factors

The haemopoietic growth factors are a diverse group of acidic glycoproteins, controlling the production, maturation and functional activity of the eight major lineages of blood cells. The regulatory mechanisms are those of redundancy and pleiotropy: more than one factor controls cells in one lineage and most factors are active on cells of more than one lineage (Metcalf, 1988; Nicola and Metcalf, 1991). For example neutrophil production can be stimulated, at least in vitro, by several colony-stimulating factors (CSF) - granulocyte CSF (G-CSF), granulocyte-macrophage CSF (GM-CSF) and multi-CSF (also called interleukin-3, IL-3) and also by interleukin-6 (IL-6) and the kit ligand (stem cell factor, Steel factor). Similarly eosinophil production can be stimulated by IL-3,
GM-CSF or IL-5. None of these factors is exclusively neutrophil or eosinophil specific, as each acts on at least some other cells. For instance GM-CSF acts to promote differentiation and survival of peripheral blood dendritic cells (Markowicz and Engelman, 1990), augments the primary antibody response by enhancing function of antigen-presenting cells (Morrisey et al., 1987), activates endothelial cells to proliferate and migrate (Bussolino et al., 1991) and together with erythropoietin (Epo) directly stimulates the proliferation and differentiation of intermediate and late erythroid progenitor cells (Emerson et al., 1985).

A further complexity is provided by the fact that some of the growth factors in addition to the soluble form can exist in membrane- and/or ECM-bound forms. For example macrophage CSF (M-CSF), GM-CSF, the kit ligand and TGF-α can exist in membrane-bound forms. The kit ligand is synthesized as either a 248 amino acid transmembrane polypeptide or a soluble form of 164 residues, which is released by the action of a still unknown protease. Two alternatively spliced kit ligand mRNAs encode transmembrane proteins that differ in the rate of processing into soluble form (Flanagan et al., 1991; Broxmeyer et al., 1992). The soluble form induces mast cell proliferation, while the cell surface form promotes cell adhesion as well as cell proliferation (Thiery and Boyer, 1992).

Gordon et al. (1987b) have demonstrated that GM-CSF can bind to marrow glycosaminoglycans and Roberts et al. (1988) confirmed this observation, indicating that heparan sulphate was the major ECM component to which growth factors bind. Subsequently many growth factors have been demonstrated to bind to ECM via heparan sulphate, chondroitin sulphate or the core proteins of specific matrix molecules (Table 1.1). Thus growth factors and ECM collaborate in creating distinct cellular environments or "niches" in the bone marrow that could regulate the proliferation and differentiation of haemopoietic cells.
Thrombospondin and kit ligand have been shown to promote synergistically the adhesion and growth of particular lineages from progenitor cell populations (Long et al., 1992). It is also possible that binding to heparan sulphate can protect growth factors from enzymatic degradation as has been found for fibroblast growth factor (FGF) (Klagsbrun et al., 1990).

Table 1.1: Examples of associations of growth factors with extracellular matrix

<table>
<thead>
<tr>
<th>Binding to heparin/heparan sulphate chains</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FGFs (aFGF, bFGF, KGF)</td>
<td>(Klagsbrun et al., 1990)</td>
</tr>
<tr>
<td>IL-3</td>
<td>(Roberts et al., 1988)</td>
</tr>
<tr>
<td>GM-CSF</td>
<td>(Gordon et al., 1987b)</td>
</tr>
<tr>
<td>Platelet factor 4</td>
<td>(Ruoslhti and Yamaguchi, 1991)</td>
</tr>
<tr>
<td>PDGF-B</td>
<td>(La Rochelle et al., 1991)</td>
</tr>
<tr>
<td>heparin binding-FGF</td>
<td>(Higashiyama et al., 1991)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Binding to chondroitin sulphate chains</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Platelet factor 4</td>
<td>(Perin et al., 1988)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Binding to proteoglycan core proteins</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TGF-β (betaglycan, decorin)</td>
<td>(Andres et al., 1989; Yamaguchi et al., 1990)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Binding to ECM glycoproteins</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TGF-β (fibronectin)</td>
<td>(Fava and McClure, 1987)</td>
</tr>
<tr>
<td>TGF-β (thrombospondin)</td>
<td>(Murphy-Ullrich et al., 1992)</td>
</tr>
<tr>
<td>PDGF-A, PDGF-B (SPARC)</td>
<td>(Raines et al., 1992)</td>
</tr>
<tr>
<td>β-endorphin (vitronectin)</td>
<td>(Hilderbrand et al., 1989)</td>
</tr>
</tbody>
</table>
growth factor receptors and signal transduction

The haemopoietic growth factors elicit their target cell responses by binding to specific receptors at the cell surface. These receptors fall into two broad categories: those with intrinsic kinase domains and those lacking such sequences. In the former family are the receptors for fibroblast growth factor (FGF), platelet-derived growth factor (PDGF), macrophage colony-stimulating factor (M-CSF/CSF-1) and kit ligand. Ligand binding to the extracellular domain of receptor tyrosine kinases usually results in receptor dimerisation, followed by activation of the catalytic function and autophosphorylation of tyrosine residues in the cytoplasmic domains. These tyrosine autophosphorylation sites in the growth factor receptors function as high affinity binding sites for SH2 (src homology) domains of signalling molecules (Fantl et al., 1992).

Several signalling pathways can then be initiated. One possibility is that Grb2 protein binds with its SH2 domain to the autophosphorylated receptor, which in turn recruits the Sos protein. Grb2-Sos complex then can activate the monomeric guanine nucleotide-binding protein Ras, leading to the activation of Raf (product of raf oncogene) and MAP (mitogen-activated protein) kinases. This cascade ultimately leads to the stimulation of transcription factors in the nucleus (Pelech, 1993). Activation of the Raf - MAP kinase pathway may also be achieved independently of Ras by a second route, following binding of phospholipase Cγ to the ligand occupied and tyrosine-autophosphorylated receptor (for example PDGF receptor). Phospholipase C may also initiate another signalling pathway involving inositol triphosphate and diacylglycerol. The latter activates protein kinase C, while inositol triphosphate mobilises Ca²⁺ from internal stores and promotes influx of external Ca²⁺ (Berridge, 1993).
The other class of growth factor receptors, lacking an intrinsic tyrosine kinase domain, includes the "haemopoietin" receptor superfamily. These receptors are characterised by four conserved cysteine residues in the extracellular domain distal to the membrane, whereas the domain proximal to the membrane contains in most cases the Trp-Ser-X-Trp-Ser motif (Bazan, 1991; Boulay and Paul, 1993). In this group are the receptors for G-CSF, GM-CSF, Epo, IL-2, IL-3, IL-4, IL-5, IL-6 and IL-7. Some of these receptors share common subunits. Thus IL-3, IL-5 and GM-CSF, bind to distinct specificity-determining α receptor components, but share a common signal-transducing β receptor component (Nicola and Metcalf, 1991). Similarly, leukaemia inhibitory factor (LIF) and IL-6 (together with ciliary neurotrophic factor (CNTF) and oncostatin M (OSM)) form another subgroup of cytokines, using a specificity-determining α receptor component in concert with a two different shared β receptor components - gp130 and LIFRβ. Subsequently it has been shown that the activated receptor complex for CNTF, LIF and OSM contain a heterodimer of LIFRβ-gp130, while the activated IL-6 receptor complexes contains gp130 homodimers (Davis et al., 1991; Murakami et al., 1993).

The molecular mechanism of cytokine receptor-mediated signal transduction is unclear, because the cytoplasmic portion of cytokine receptors and β chains do not have tyrosine kinase activity. However rapid intracellular tyrosine phosphorylation occurs in response to cytokine stimulation and the use of tyrosine kinase inhibitors has suggested that these phosphorylations are required for subsequent cellular responses (Ip et al., 1992; Campbell et al., 1993). In several cases tyrosine kinase activity has been found to copurify with cytokine receptor complexes, suggesting that non-receptor tyrosine kinases might associate with β-component cytoplasmic domains (Murakami et al., 1993; Wang et al., 1993). Recently it has been shown that the members of a newly described
family of cytoplasmic tyrosine kinases, Tyk2, JAK1 and JAK2, are involved. Epo, IL-3, GM-CSF and G-CSF can induce tyrosine phosphorylation and kinase activation of JAK2 and JAK2 can directly associate with β receptor components (Velazquez et al., 1992; Argetsinger et al., 1993; Wittuhn BA et al., 1993). This is followed by the induction of Ras-MAP kinase cascade and ultimately the activation of various transcription factors.

The potential convergence of so many different cytokine signalling pathways just inside the membrane, at the level of JAK, suggest that many different cytokines signal through fundamentally similar mechanisms and via overlapping pathways. Also cell responsive to one cytokine often respond similarly to other cytokine when receptors for these cytokines are introduced. Furthermore an elegant series of experiments has shown that the receptor tyrosine kinases and haemopoietic receptors may activate common nuclear signal transduction pathways. Thus IL-4 receptor signals by a mechanism similar to the insulin and insulin-like growth factor receptors, since IL-4 and insulin receptors induce the phosphorylation of closely related proteins (Wang et al., 1993). These studies reinforced the concept that nature has exploited only a limited number of ways to transduce signals across the membranes. However different cytokine receptors are also capable of generating qualitatively different signals in the same cells. This specificity may be determined by the use of different combinations of JAKs by different cytokine receptors or because distinct sets of potential substrates or additional yet unidentified kinases are involved.

(c) soluble growth factor receptors

Many of the haemopoietic growth factor receptors have been found to have soluble isoforms in addition to the membrane bound forms
including the receptors for G-CSF, GM-CSF, IL-1, IL-2, IL-4, IL-5, IL-6, LIF, M-CSF/CSF-1, TNF and Epo (Heaney and Golde, 1993). Their physiological role is incompletely understood but there are at least three possibilities: (1) that the soluble receptor stabilizes its ligand and functions as a serum binding protein, to prevent the degradation of the ligand until it is delivered to the membrane-associated receptor; (2) that the soluble receptor competes with its membrane bound counterparts for binding to the ligand and in this way the soluble receptor modulates the concentration of active ligand in the extracellular milieu and (3) that the soluble receptor may work at a distance as a mediator of hormonal responsiveness in cells that do not express the hormone specific receptor but do express the associated subunits. Thus a soluble IL-6 receptor could confer IL-6 sensitivity to cells that express only the β subunit of the IL-6 receptor complex (gp 130) (Taga et al., 1989).

1.3. Adhesion molecules

Although much is known about the individual components of the bone marrow microenvironment, little is known about the adhesion interactions among these various components and their counter-receptors. Better understood are the adhesion receptors mediating cell-cell and cell-extracellular matrix interactions within the immune and nervous systems. Since the same ligand/receptor pairs have been identified, similar mechanisms may regulate adhesion during inflammation, the immune response or early haemopoiesis. This section will describe the structure, function and signal transduction of the major families of adhesion receptors - the integrins, the selectins, the immunoglobulin superfamily and the cadherins.
1.3.1. Integrins

(a) **structure and function**

Integrins are a family of cell-surface glycoproteins. At least twenty different integrins have been characterised to date and they are formed by various noncovalent associations of 14 α and 8 β subunits to form heterodimers (Fig. 1.3) (Hynes, 1992; Juliano and Haskill, 1993). Each subunit has a large extracellular domain, a single transmembrane region and usually a short cytoplasmic domain ranging from 40 to 60 amino acids in length (with the exception of β4 cytoplasmic domain which is 1000 amino acids long) (Suzuki and Naitoh, 1990). The integrins were originally identified as receptors for extracellular matrix proteins such as collagens, fibronectin, laminin and vitronectin, but some integrins can also function as cell-cell adhesion molecules. There is considerable redundancy within the integrin family in that most of the integrins bind to more than one ligand and ligands can be recognised by more than one integrin. In some cases two integrins that share the same ligand will actually recognise different regions of the ligand molecule, as with the α5/β1 and α4/β1 fibronectin receptors (Guan and Hynes, 1990), or the α1/β1 and α6/β1 laminin receptors (Hall et al., 1990). In other cases such as α5/β1 and α3/β1, two integrins bind to the same region of the ECM protein (Elices et al., 1991).

One possible explanation for the very large number of different integrins, their overlapping specificities and redundancy, could be that they may serve different functions in the cell. This suggestion is based on the fact that the cytoplasmic domains of integrin subunits, particularly those of the various α subunits (Sastry and Horwitz, 1993) are highly heterogeneous, indicating that each subunit may contribute to discrete intracellular functions. In addition, it has been shown that several integrin α subunits undergo alternative splicing in a tissue and
Figure 1.3: Schematic representation of an integrin molecule as an αβ heterodimer. Ca, divalent cation binding sites, C, cystein-rich amino acid sequence. Some α chains are proteolitically cleaved but linked together over disulfide bridges. The right-hand figure shows the possible α,β combinations (After Dunon et al., 1992).
developmentally regulated manner, suggesting again that different cellular responses can be initiated from binding to a single ligand (Tamura et al., 1991; Hogervorst et al., 1991). Although integrins do not have the characteristics of a signal generating receptor, such as kinase or phosphatase domains, nonetheless is clear that integrins are capable of transducing signals from the ECM to the cell interior and that these signals may trigger changes in gene expression as described below.

(b) **signal transduction by integrins**

*Involvement of the cytoskeleton*

One of the ways by which integrins could transmit signals is by organising the cytoskeleton, thus regulating cell shape and internal cellular architecture. Integrin cytoplasmic domains do indeed interact with cytoskeletal components (Horwitz et al., 1986; Otey et al., 1990) and the β1 and β3 cytoplasmic domains are of particular significance for cytoskeletal association (Reszka et al., 1992; Solowska et al., 1991). Internal deletions and point mutations within the β1 cytoplasmic domain indicate that three regions contribute to focal contact localisation and cytoplasmic domain swapping between β1 and β3 showed functional interchangeability with respect to focal adhesion localisation. The exact mechanism by which the cytoskeleton plays a role in signal transduction in not very well understood. Certainly, many elements of the metabolic machinery of the cell, such as polyribosomes and mitochondria, are associated with the cytoskeleton and changes in their position and organisation could result in changes in function. In addition mechanical forces could be transmitted directly to the nucleus from the ECM since intermediate filaments physically link the plasma membrane to the nuclear envelope and thus external forces could result in changes in nuclear size and DNA packaging (Adams and Watt, 1993).
Although many developmental and differentiative processes involve changes in cell shape, integrin-mediated changes in gene expression can occur in the absence of changes in cell morphology (Werb et al., 1989). It has been shown that integrins could function as agents for the transduction of two types of signals (Hynes, 1992; Juliano and Haskill, 1993; Sastry and Horwitz, 1993). One is the transmission of signals from the matrix into the cell which may modulate gene expression. This type of signalling has been called outside-in signalling. Another type of signalling arises from effectors in the cytoplasm which modulate the affinity and/or specificity of integrins for its extracellular ligands and this cytoplasmic-mediated signalling has been called inside-out. The evidence suggest that the cytoplasmic domain is a mediator in both modes of signalling.

**Inside-out signalling**

Many examples of affinity modulation initiated in the cytoplasm have been described for different integrins in various cell types. The \( \alpha IIb\beta3 \) integrin on unstimulated resting platelets does not bind fibrinogen, however, activation by thrombin rapidly induces high-affinity fibrinogen binding to the \( \alpha IIb\beta3 \) receptor (Shattil and Brass, 1987). Another example is the leucocyte integrin \( \alpha L\beta2 \), which is inactive on resting leucocytes, but following activation by engagement of the T-cell receptor or addition of phorbol esters \( \alpha L\beta2 \) binds to its counter-receptors ICAM-1 and ICAM-2 (Dustin, 1990). A lot of evidence suggest that this signalling is mediated by the cytoplasmic domains. Deletions of the \( \beta1 \) and \( \beta2 \) cytoplasmic domains inhibit adhesion of NIH 3T3 cells to fibronectin and laminin substrates (Hayashi et al., 1990) or binding of Cos cells to ICAM-1 (Hibbs et al., 1991). However the exact mechanism by which the cytoplasmic domain modulates affinity is still unclear. Whether these changes result
from direct interactions between the α and β subunit cytoplasmic domains, or from binding of cytoplasmic (cytoskeletal) associated molecules remains to be determined. Protein phosphorylation is likely to regulate the affinity/specificity changes, since the tyrosine phosphorylation of the β1 cytoplasmic domain by pp60v-src inhibits ligand binding (Tapley et al., 1989). Furthermore there have been several reports of phosphorylations of α and β subunit cytoplasmic domains following activation of kinases (Hynes, 1992). However further evidence is required to confirm that phosphorylation is the major physiological mechanism for cytoplasmic induced affinity modulation.

**Outside-in signalling**

Various signalling events in response to integrin occupancy have been described including cytoplasmic alkalinisation due to activation of the sodium/hydrogen antiporter (Ingber et al., 1990; Schwartz et al., 1991), alterations in cAMP level (Nathan and Sanches, 1990) and increases or decreases in the intracellular concentration of Ca++ (Ng-Sikorski et al., 1991, Jaconi et al., 1991). It has also been reported that ligation of integrins can alter the phosphorylation state of intracellular proteins (Guan et al., 1991; Hanks et al., 1992). Thus in at least three cell systems (carcinoma cells, fibroblasts and platelets) ligation-induced phosphorylation of the 125 kD focal adhesion kinase (FAK), pp125fak, has been described (Juliano and Haskill, 1993). FAK is a prominent phosphoprotein in adhesion plaques and is also known to be a substrate for src family protein kinases, while as a tyrosine kinase it displays a limited substrate repertoire, which includes tensin and paxillin, two adhesion plaque proteins. However how FAK is activated by integrin ligation and the particular signalling pathways that it initiates are unknown.
An intriguing possibility would be that the activation of FAK or other proteins is related to control of gene expression. Thus one of the most interesting consequences of cell adhesion to ECM is the induction of specific gene expression. For example in monocytes specific genes are activated depending on the substrate to which they attach. IL-ra (receptor antagonist) and CSF-1 show specificity for fibronectin coated surfaces, while the genes for the cytokines IL-1 and IL-8 are induced on all surfaces (Eierman et al., 1989; Sporn et al., 1990). It has been noted that all these genes, induced by the adhesion of monocytes, share the NF-κB transcriptional motif. Since translocation of enhanced levels of NF-κB into the nucleus is known to occur after the adhesion of human monocytes, it has been proposed that this nuclear translocation of NF-κB is mediated by an integrin-dependent kinase (Juliano and Haskill, 1993). However further studies are required to determine whether phosphorylation of FAK or other proteins is a necessary step in this induction of transcription.

1.3.2. Adhesion molecules of the immunoglobulin superfamily

Other cell adhesion molecules include the immunoglobulin (Ig) superfamily members. The Ig superfamily includes a wide variety of proteins that all share the immunoglobulin homology unit, which consist of 70 to 110 amino acids organised in two parallel β-sheets made by up to 7-9 β-strands (Fig. 1.4) (Pigott and Power, 1993). With the exception of a few highly conserved residues, there is an enormous variety in the primary structures of the members of the Ig superfamily, but the tertiary structure is remarkably conserved. It is likely that this basic structure serves as a scaffold on which unique determinants can be displayed for recognition. Members of the Ig superfamily are involved in a variety of cell functions including lymphocyte recognition of antigen, signal transduction, cell adhesion and virus binding.
Figure 1.4: Schematic representation of adhesion molecules of the immunoglobulin superfamily. Filled circles represent N-glycosylation sites; C2, Ig C2-like domain; V, Ig V-like domain (After Dunon et al., 1992).
The earliest recognition events mediated by members of the Ig superfamily were probably simple homophilic adhesion, between identical single domains on opposing cells (Williams, 1987). As multidomain molecules have evolved this capacity for homophilic adhesion has been maintained and also extended to adhesion between non-identical family members such as CD2 and LFA-3. Furthermore other non-Ig family members have been identified as ligands. Amongst these are the integrins, including LFA-1 and Mac-1 which bind ICAM-1 and VLA-4 (α4/β1) and α4/β7 which bind VCAM-1. Components of the extracellular matrix can also serve as ligands and these include collagen and heparin, recognised by myelin associated glycoprotein (MAG) and heparan sulphate recognised by NCAM (Pigott and Power, 1993).

It has become obvious that molecules originally identified as adhesion molecules may have diverse functions. For example NCAM (Doherty and Walsh, 1992), L1 (Williams et al., 1992) and MAG (Johnson et al., 1989), expressed in fibroblasts by gene transfer have been found capable of inducing a neurite outgrowth response from neurons cultured on monolayers of transfected cells. It has been shown that these adhesion molecules promote neurite outgrowth not via adhesion per se but by transmitting signals to the cell interior. A common but specific second messenger pathway has been defined for NCAM and L1-mediated outgrowth. This involves a number of intracellular targets including increased tyrosine phosphorylation, pertussis toxin sensitive G protein activation and finally Ca++ influx into the growth cone through L- and N-type Ca++ channels.

Alternative splicing of the NCAM gene also generates isoforms with different capacities for cell signalling events. One of the major alternatively spliced forms of NCAM contains the product of the variable alternatively spliced exon (VASE) in the fourth Ig domain. Expression of
VASE containing NCAM results in a failure to activate neurite outgrowth (Doherty et al., 1992) and it has been suggested that the use of VASE may reflect an important developmental mechanism for modulating neuronal responsiveness to environmental cues (Walsh et al., 1992). Recent observation in chicken (Scholey et al., 1993) and Aplysia (Mayford et al., 1992) suggest that NCAM and related molecules might be involved in processes like learning and memory. Inactivation of the NCAM gene in mice results in size reduction of the olfactory bulb and also leads to impaired learning ability and changes in explorational behaviour (Cremer et al., 1994).

Finally the mechanisms regulating tissue specific patterns of adhesion molecule gene expression are just beginning to be explored. The NCAM promoter has been isolated from a number of species including human (Barton et al., 1990), rat (Chen et al., 1990), mouse (Hirsch et al., 1990) and chicken (Colwell et al., 1992). A number of highly conserved sequence motifs that may bind transcription factors, including homeobox (HOX) gene products have been identified (Hirsch et al., 1991; Colwell et al., 1992). The HOX2.5 protein binds to the NCAM promoter and enhances NCAM gene transcription in transfected cultured cells, showing that the NCAM gene is a direct target of HOX gene products (Jones et al., 1992). Similarly the homophilic cell adhesion molecule connectin in Drosophila is a target of the ultrabithorax homeotic gene (Gould and White, 1992). The regulation of cell adhesion molecules genes by genes that control embryonic pattern has very important implications for the spatial regulation of tissue morphogenesis.
1.3.3. Selectins

(a) **structure and function**

The selectin family consists of three membrane glycoproteins, including E-, L- and P-selectins, that mediate leukocyte-endothelial cell interactions by binding carbohydrate ligands on opposing cells (McEver, 1991; Varki et al., 1992). The extracellular region of each of them contains an amino-terminal motif that is characteristic of Ca++ dependent or C-type lectins, followed by an epidermal growth factor (EGF)-like domain and a series of consensus repeats like those in complement-regulatory proteins (CRP) (Fig. 1.5). Although both lectin and EGF domains are required to mediate neutrophil adhesion, the three dimensional structure of the ligand-binding region of human E-selectin revealed limited contact between the two domains (Graves et al., 1994). The transmembrane region of selectins is followed by a short cytoplasmic domain. The longest cytoplasmic tail is found in P-selectin (35 residues), one function of which is to direct P-selectin to the secretory granule membranes of endothelial cells and platelets (Disdier et al., 1992). When these cells are activated by thrombin or histamine, P-selectin is redistributed within minutes to the cell surface as granule membranes fuse with the plasma membrane. No function has yet been shown for the cytoplasmic regions of L- and E-selectins. E-selectin is expressed by cytokine activated endothelium. It appears on the endothelial cell surface 2-4 hours following induction of its synthesis by inflammatory cytokines such as IL-1 or TNF-α. L-selectin is constitutively expressed on the surface of monocytes, neutrophils and a subsets of lymphocytes until they are activated (Table 1.2).

P- and E-selectin bind to myeloid cells, eosinophils (Bochner et al., 1991), a subset of natural killer cells (Lobb et al., 1991) and memory T lymphocytes (Picker et al., 1991; Moore et al., 1992). Thus both molecules may mediate leucocyte extravasation during acute, chronic and allergic
Figure 1.5: Schematic representation of selectin molecule
(After Dunon et al., 1992)
inflammatory responses. L-selectin binds to the HEV of peripheral lymph nodes (Imai et al., 1991) and to cultured endothelial cells activated by cytokines (Spertini et al., 1991) (Table 1.2). Therefore, L-selectin has been implicated in both homing of lymphocytes to peripheral lymph nodes and the "rolling" of neutrophils on the endothelium near acute inflammatory sites (described further below).

Table 1.2: Characteristics of selectins

<table>
<thead>
<tr>
<th>Name</th>
<th>Expressed by</th>
<th>Target cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-selectin</td>
<td>rapidly-activated platelets and endothelial cells</td>
<td>neutrophils, monocytes, eosinophils, lymphocyte subsets</td>
</tr>
<tr>
<td>E-selectin</td>
<td>cytokine-activated endothelial cells</td>
<td>neutrophils, monocytes, eosinophils, lymphocyte subsets</td>
</tr>
<tr>
<td>L-selectin</td>
<td>monocytes, neutrophils, lymphocyte subsets</td>
<td>activated endothelial cells of peripheral lymph nodes, high endothelial venule</td>
</tr>
</tbody>
</table>

All selectins appear to recognise a sialylated carbohydrate determinants on their counter-receptors. The carbohydrate ligand for L-selectin is related to tetrasacharide sialyl Lewis\(^a\) (sLe\(^a\)) and Lewis\(^x\) (sLe\(^x\)) and contains sialic acid and sulphate (Rosen, 1993). E- and P-selectin recognise carbohydrate structures that are distinct but are closely related to the sLe\(^a\) and sLe\(^x\). The carbohydrate ligands for L- and P-selectin are O-linked to specific mucin-like molecules. Thus L-selectin recognizes two mucins in high endothelial venules, glycosylation dependent cell adhesion
molecule 1 (GlyCAM-1), which is secreted (Rosen, 1993) and CD34, which is on the cell surface (Baumhueter et al., 1993), while the mucin-like P-selectin glycoprotein ligand (PSGL-1) is a disulphide-linked dimer of 120 kD (Sako et al., 1993).

(b) rolling and tight adhesion: the cooperative roles of selectins and integrins

The current model for leucocyte movement from the blood circulation into tissues involves an early phase of rolling during which a brief and reversible attachment to the endothelium occurs, followed by a phase of tight adhesion and transendothelial migration. Selectins are involved in the earliest rolling interactions, while integrins are required for the tight adhesion (Lasky, 1993). The latter process needs an appropriate activation signal to be delivered to the leukocytes which will trigger an increase in the affinity of leukocyte integrins for endothelial ligands.

The mechanism underlying the ability of selectins to support rolling is not very well understood, but an involvement of P-selectin has been demonstrated in vitro (Lawrence and Springer, 1991). Other studies have suggested a role for L-selectin, since antibodies against L-selectin and recombinant soluble forms of L-selectin can inhibit rolling in vivo (Ley et al., 1991; von Adrian et al., 1991). The recent studies of Mayadas et al. (1993) in P-selectin deficient mice showed approximately 2.5 times more circulating neutrophils than in wild type mice. These findings suggest that defects in P-selectin expression lead to a deficiency in neutrophil-endothelial cell adhesion, preventing neutrophils from migrating efficiently into tissues. All these results support the view that P- and L-selectins act cooperatively at the early stages of neutrophil rolling and inflammation, while E-selectin, which is maximally expressed 4 hours
after induction of an inflammatory response, may be the major mediator of
leucocyte rolling during the later phases of inflammation (Lasky, 1993).

1.3.4. Cadherins

The last major family of cell adhesion molecules are cadherins. They are cell surface membrane proteins expressed in a wide variety of
tissues and species. Most members of this family have the same division
into sub-domains, including presequences that are cleaved upon
processing, a large extracellular domain, a single hydrophobic
transmembrane region and a cytoplasmic tail. The short cytoplasmic tail
is the most highly conserved region of homology between cadherins and is
particularly important for cadherin function (Takeichi et al., 1988; Geiger
and Ayalon, 1992).

Cadherins mediate homophilic Ca++ dependent cell-cell adhesion,
although heterophilic binding between different cadherin molecules is
possible. Cadherins are responsible for the selective cell-cell adhesion or
cell sorting which is necessary to allocate different cell types to their
proper positions during development and to maintain the integrity of the
multicellular organism. The complex spatio-temporal pattern of cadherin
expression during embryonic development further confirms the important
role of cadherins in morphogenesis (Geiger and Ayalon, 1992). Cadherins
have generally been associated with cell interactions among solid tissue,
however it has recently been shown that Langerhans cells express E-
cadherin, and adhere to keratinocytes by an E-cadherin mediated
adhesion (Tang et al., 1993). This is the first demonstration of cadherin
function in leucocytes and may indicate a much broader role of cadherins
in cell interactions than previously suspected.

In summary the response of cells to adhesion to extracellular
matrix and to neighbouring cells are diverse and complex and include
changes in cell proliferation, motility and cell differentiation. These responses are mediated by multiple adhesion receptors, including integrins, selectins, cadherins and members of the Ig superfamily, which are able to transduce signals, similar to those of traditional growth factor receptors. The regulation of expression of cell adhesion receptors and associated regulatory proteins also has important implications for the growth and spread of cancers. In particular changes in the expression of both integrins and cadherins seems to be involved in tumour growth and metastasis. However much more remains to be learned and the functions of many other potential adhesion molecules need to be resolved. The rest of this introductory chapter will focus on three such potential adhesion molecules, the CD34, CD33 and CD66 defined antigens, which are expressed on sequential stages of myeloid development.

1.5. **CD34 defined antigen**

The CD34 group of monoclonal antibodies recognise a 105-120 kD cell surface glycoprotein which is selectively expressed by 1% to 4% of normal human bone marrow cells, including myeloid and lymphoid progenitor cells and probably haemopoietic stem cells (Civin et al., 1989). CD34 antigen has also been shown to be expressed in blasts from approximately 30% of patients with acute lymphoid or myeloid leukemias. Less differentiated French-American-British (FAB) subtypes are more likely to be positive. CD34 antigen has been detected on a subset of bone marrow stromal cells (Simmons and Torok-Storb, 1991) and on vascular endothelium, where it is concentrated on the surface of interdigitating membrane processes (Fina et al., 1990). Among other tumours CD34 antibodies react strongly with angioblastomas and with some Kaposi's sarcomas (Fina et al., 1990; Sankey et al., 1990). Blood vessels in all types of tumours are CD34 positive and where angiogenesis is active there is an
extensive labelling of vascular sprouts (Schlingemann et al., 1990). A subset of human fibroblasts also bind CD34 antibodies (Greaves et al., 1992) in agreement with the detection of CD34 mRNA in murine fibroblast cell lines (Brown et al., 1991). The murine CD34 gene is expressed in a wider variety of tissues including brain and testis, the former at ten times the level observed in haemopoietic tissues or cell lines (Brown et al., 1991).

The human CD34 cDNA has been cloned and predicts a 40 kD type I integral membrane protein with nine potential N-linked and numerous potential O-linked glycosylation sites in the extracellular domain (Simmons et al., 1992), which is in agreement with previous biochemical studies, showing that CD34 antigen is heavily glycosylated (Sutherland et al., 1988). The murine CD34 cDNA sequence also has the characteristic features of heavily glycosylated molecules, suggesting that the carbohydrate moieties may have an important functional role. Two alternatively spliced murine CD34 cDNAs have been described, generating two murine CD34 isoforms with different cytoplasmic domains of 73 and 16 amino acid residues (Suda et al., 1992). There is a 90% amino acid identity between the cytoplasmic domains of human and (the longer) murine CD34 deduced sequences, with two of the potential phosphorylation sites conserved, suggesting that the cytoplasmic domain may be involved in signal transduction (Brown et al., 1991; Suda et al., 1992). In this respect it is interesting that the CD34 antigen can be stoichiometrically phosphorylated by activated protein kinase C both in CD34-positive cell lines and fresh leukaemic lymphoblasts (Fackler et al., 1990). Although the human and the mouse CD34 antigens have no significant sequence homologies to other proteins in the databases, they have some structural similarities to the surface molecule CD43, since both
are highly O-sialoglycosylated, mucin-like molecules (Brown et al., 1991; Simmons et al., 1992).

The human and mouse CD34 genomic DNA regions have also been characterised and have revealed the presence of myc, myb and multiple ets-like binding motifs, suggesting that these proteins may regulate CD34 expression (He et al., 1992; Burn et al., 1992). The CD34 gene has been mapped to human chromosome 1q32 (Molgaard et al., 1989; Tenen et al., 1990), a region that contains genes encoding a number of haemopoietic signalling and regulatory molecules including CD45, the protooncogene TRK (Miozzo et al., 1990) and the regulator of complement activation cluster (Weis et al., 1987). This genomic region is also of interest in view of its frequent involvement in deletions and rearrangements observed in acute leukaemias and myelodysplastic syndromes (Auerbach and Allen, 1991). Furthermore the KG1 cell line, which shows constitutive high expression of CD34 antigen, has chromosome alterations involving the 1q32 region (Furley et al., 1986).

The structural analogies to the sialomucin CD43 and the polar distribution of CD34 in endothelial cells has led to the hypothesis that CD34 antigen may function as an adhesion molecule, mediating the attachment of stem cells to the bone marrow extracellular matrix or directly to stromal cells. However using blocking experiment with CD34 MoAbs, Gordon et al. (1990) were unable to demonstrate a role for CD34 antigen in progenitor cell adhesion to cultured stromal layers. This does not eliminate the possibility that CD34 has a role in adhesion since the MoAbs used may not have been derived from CD34 epitopes involved in binding. Furthermore, it has recently been demonstrated that the vascular form of CD34 antigen binds specifically to L-selectin (Baumhueter et al., 1993). Because mucins could bear lineage specific carbohydrate determinants, it is possible that the adhesive properties of
CD34 in the bone marrow microenvironment and in the vascular endothelium are modulated by differential glycosylation or sialic acid content. CD34 antigen may also act as a signalling molecule, as is suggested by the chromosomal localisation and the phosphorylation of CD34 by PKC. Furthermore MoAbs against the structurally similar protein CD43 can induce lymphocyte and monocyte activation via the phospholipase C signalling pathway (Silverman et al., 1989).

1.6. CD33 defined antigen

In the human haemopoietic system, CD33 monoclonal antibodies recognise a 67 kD cell surface molecule, whose expression is largely restricted to the myelomonocytic lineage (Peiper et al., 1989). CD33 antigen is expressed both later than CD34 antigen and in concert with CD34, but is not detected on the earliest CD34+ haemopoietic stem cell. CD33 is detected on multi-potential colony-forming cells, progenitors of granulocytes and macrophages, early erythroid progenitors and mast cells (Andrews et al., 1983; Griffin et al., 1984; Sabbath and Griffin, 1986). In addition, the antigen is found on myelocytes, a proportion of metamyelocytes and on monocytes. It is expressed only weakly by terminally differentiated granulocytes and not at all by normal lymphoid cells. Leukaemic myeloblasts from approximately 85% of patients with acute myelogeneous leukaemia (AML) express CD33, but typically lymphoblasts from patients with acute lymphoblastic leukaemia and other lymphoid malignancies do not. Thus, CD33 antibodies are helpful diagnostic markers for distinguishing undifferentiated cases of AML from acute lymphoblastic leukaemia.

CD33 monoclonal antibodies have also been used therapeutically to purge residual leukaemic myeloblasts prior to autologous bone marrow transplantation for AML in remission, as the CD33 antigen is expressed by
leukaemic myeloblasts but not by the earliest normal haemopoietic stem cells (Ritz et al., 1987; Robertson et al., 1992). Furthermore human bone marrow depleted of CD33+ cells in vitro is able to grow normally in long-term bone marrow cultures (Andrews et al., 1986; 1990). In contrast to most other myeloid cell surface antigens, CD33 is strictly haemopoietic in its distribution. These properties make CD33 antibodies an ideal candidate for therapy of AML and clinical trials have demonstrated specific targeting of 131I-labelled, M195 antibody to leukaemic cells in bone marrow (Scheinberg et al., 1991). The construction of chimaeric and humanised CD33 antibodies with higher affinity, improved effector functions and lower immunogenicity would increase their potential for AML therapy (Co et al., 1992).

At least eight different monoclonal antibodies having CD33 specificity have been produced and characterised (Peiper et al., 1989). All of these antibodies bind to mouse cell transformants, expressing CD33 protein and competitive inhibition experiments suggest that all of them recognise the same or closely related epitopes. Immunoprecipitation analysis of KG1-a cells and mouse cell transformants expressing CD33 antigen, with CD33 monoclonal antibody (My9) under non-reducing conditions showed two polypeptide species with apparent molecular weights of 65 kD and 140 kD, suggesting that, in addition to the monomer, a homodimERIC form of the protein also exists at the cell membrane. Endoglycosidase digestion of CD33 immune complexes from surface iodinated mouse cell transformants and data from metabolical labelling of KG1-a cells with [3H]-galactose revealed that CD33 antigen is a glycoprotein, designated as gp 67 (Peiper et al., 1989).

The human CD33 genomic locus, spanning approximately 35 kb, has been cloned and the gene localised to the long arm of chromosome 19 by in situ hybridisation (Peiper et al., 1988). A cDNA for the CD33 antigen has also been isolated by expression cloning from a cDNA library of U937 cells.
(Simmons and Seed, 1989). The predicted protein sequence indicates that this molecule is a member of the immunoglobulin gene superfamily, having an N-terminal IgV-like domain, followed by one IgC2-type domain, a transmembrane region and a cytoplasmic tail of 82 residues. The CD33 sequence shows extensive similarity to two adhesion receptors, the myelin-associated glycoprotein (MAG) and the B cell antigen, CD22. Two CD33 transcripts, of 1.5 kb and 1.8 kb, have been identified in a panel of human myeloid leukaemia cell lines (Simmons and Seed, 1989). However, treatment of U937 cells with a variety of pharmacological agents, known to induce granulocytic or monocytic differentiation (PMA, DMSO, IFN-γ and IL-1β) does not alter the level of CD33 mRNA transcripts or protein expression. Independent studies confirmed that CD33 expression by U937 cells was unaltered by exposure to IFN-γ, lipopolysacharide or CSF-1. Treatment of U937 cells with an active metabolite of vitamin D3, 1, 25(OH)2D3 was found to enhance the expression of CD33 as determined by cellular radioimmunoassay (Peiper et al., 1989).

Despite interest in CD33 antibodies as diagnostic and therapeutic reagents, the function of the CD33 antigen is still unknown. Based on the observed homologies with the adhesion molecules CD22 and MAG it could be hypothesised that CD33 antigen might be an adhesion molecule.

1.7. CD66 defined antigens. BGPC splice variant.

Within the haemopoietic system CD66 defined antigens are expressed at a later stage than CD33 and are detected on cells of the granulocyte lineage, but not on a majority of haemopoietic progenitor cells from human bone marrow (Watt et al., 1991; Watt et al., 1994a). Colonic epithelium is also CD66 positive. Antigenic analysis revealed that CD66 (YTH71.3.2 and CE6/2D3.1) and CD66-like (YPC2/12.2.1) antibodies identify members of the carcinoembryonic antigen (CEA) family including
CEA itself, non-specific cross-reacting antigen (NCA), biliary glycoprotein (BGP), CGM1 and CGM6 (NCA-95). All these molecules except CEA are expressed by cells of the neutrophil lineage.

The CEA-related proteins are members of the Ig gene superfamily, with the extracellular region having a structure consisting of a leader or signal peptide, an IgV-like N-terminal domain and a variable number of (or no) IgC-like domains (Thompson et al., 1991). At least 29 different CEA-related cDNAs or genes have been identified and these are tandemly arranged on the long arm of chromosome 19 (19q13.1-3) (Thompson et al., 1991; 1992). The CEA superfamily has been arbitrarily divided into two subgroups: the membrane bound CEA subfamily and the secreted pregnancy specific glycoprotein (PSG) subfamily. The membrane linked CEA subfamily members may be anchored into the plasma membrane via glycosyl phosphatidyl inositol (GPI)-linkage of the carboxy terminal domain or may occur as transmembrane proteins, which possess cytoplasmic tails of variable length (Hefta et al., 1988; Kolbinger et al., 1989). The former GPI-linked proteins include CEA itself and the non-specific cross reacting antigen (NCA), while BGP, CGM1 and CGM7 are transmembrane molecules (Thomas et al., 1991).

The BGP molecules are thought to occur as a group of at least eight differentially spliced and phosphorylated transmembrane proteins (Fig. 1.6). These all carry the N-terminal IgV like domain and have none (BGPx and BGPx'), two (the A1 and B1 domains of BGPb and BGPd) or three (the A1, B1 and A2 domains of BGAa) IgC2 set domains. The third IgC domain may be replaced by a 31 amino acid non IgC2 set domain (BGPy and BGPz) (Barnett et al., 1989; 1993). These variants are also differentially spliced in their cytoplasmic region, having long (BGAa, b and x) or short (BGPc, d, x') cytoplasmic tails. Additional splicing may also generate three soluble BGP molecules (BGPg, h and l) as determined by
Figure 1.6: BGP splice variants
Different domains are indicated by letters above the blocks and different shadings. L, leader peptide; V, IgV-like domain; A1, A2 and B1, IgC-like domains; TM, transmembrane domain; Cyt, cytoplasmic domain.
cDNA cloning from a human leucocyte library (Kuroki et al., 1991). Unlike CGM1 and CGM6, it appears that BGP molecules are not restricted in their expression to haemopoietic cells (Kuroki et al., 1991; Skubitz et al., 1992; Khan et al., 1993), but are also expressed by a number of other cell types such as bile canaliculi, colonic and lung epithelium and hepatocytes (Barnett et al., 1989; 1993; Kim et al., 1992; Takahashi et al., 1993; Shupert and Chan, 1993). However the regulation of splicing and protein expression in different tissues and at different times in development is not yet understood.

The CEA related superfamily members are thought to act as both homophilic and heterophilic adhesion molecules in vitro. Thus CEA mediates Ca++ independent homophilic aggregation of rodent cells transfected with CEA cDNA (Benchimol et al., 1989) as well as interacting with NCA and BGP (Rojas et al., 1990). Similar interactions have been observed with NCA and the human, mouse and rat equivalents of BGP (Rojas et al., 1990; Oikawa et al., 1991; 1992; McCuaig et al., 1992; Tingstrom et al., 1990). Examples of other heterophilic adhesive interactions include (a) the presentation of sLe^x carbohydrate moieties by NCA and a phosphorylated BGP molecule on mature neutrophils (Stocks et al., 1990; Kuijpers et al., 1992), providing ligands for the E-selectin on activated endothelium, and (b) the presence of RGD sequence in the N-terminal domain of some PSG subgroup members (Thomas et al., 1991) possibly allowing these molecules to mimic matrix proteins and/or act as negative regulators of adhesion by interfering with receptor-ligand interactions.

Recently Watt et al. (1994b) have isolated by expression cloning, using CD66 MoAbs, another BGP cytoplasmic spliced variant, BGPa. This protein has an identical extracellular and transmembrane sequence to BGPa with one N-terminal IgV like domain, three IgC-like extracellular
domains (A1, B2, A2) plus a transmembrane domain, but the cytoplasmic domain is spliced out by 53 nucleotides (Fig. 1.6). RT-PCR analysis show that this splice variant can be detected in three primary colonic adenocarcinomas and in colonic carcinoma (Colo 201, HT29), in myeloid (KG1, KG1b and HL-60) and B (Ramos and Raji) cell lines, but not in the monocytic cell line, U937. However the function of this newly identified BGPC splice variant is not known. Based on the previously described adhesive functions of the members of BGP subfamily it could be expected that BGPC may mediate homophilic and heterophilic adhesion, although it is possible that the cytoplasmic splicing may modify these functions.

1.8. Aims of the thesis

Despite a considerable amount of work, the functions of CD34, CD33 and BGPC molecules remain unclear. The work described in this thesis was intended to clarify this. Taken together the structural information for these molecules including, (a) the sialomucin like structure of the CD34 antigen and its polar distribution on endothelial cells, (b) the Ig relatedness of CD33 antigen with highest homology to adhesion receptors, MAG and CD22, and (c) the relationship of the BGPC protein to the BGP subfamily of adhesion molecules, suggest that they all may act as adhesion molecules. Since CD34, CD33 and BGPC proteins are expressed on sequential stages of myeloid development, they are candidates to mediate cell-ECM and cell-cell adhesion within the bone marrow microenvironment, as well as at later stages of the maturation and function of myeloid cells. Furthermore, although BGP molecules are known to be adhesion molecules the significance of the large number of members of this family is ill understood as is the role of isoforms exhibiting alternative splicing in their cytoplasmic domains. To test the hypothesis that CD34, CD33 and BGPC molecule may mediate cell-ECM
or cell-cell adhesion and to examine whether the alternative splicing of BGPc may influence its binding properties, we analysed their binding to stromal layers from LTBMC, purified ECM substrates and bone marrow ECM extracts, and cell lines. The results of these studies are described in chapter 3.

Approaches for investigating the function of human proteins are limited by ethical constraints, while the mouse has obvious advantages for experimental manipulations. A first step in using the mouse model system is to identify the murine counterpart of a gene product of interest. Since during this study the murine homologue of CD34 antigen was cloned by others (Brown et al., 1991) we focused our attention on cloning the murine homologue of the CD33 myeloid antigen. Chapters 4-6 of this thesis describe the molecular cloning, sequence analysis and tissue and cell distribution of the murine homologue of CD33 antigen.
Chapter 2
MATERIALS AND METHODS

2.1. Materials

2.1.1. Biochemical reagents and materials

Biochemical reagents and materials are listed below. The addresses of suppliers are given at the end of this section.

Reagents

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABTS</td>
<td>Amersham</td>
</tr>
<tr>
<td>Agarose (electrophoresis grade)</td>
<td>Life Technologies</td>
</tr>
<tr>
<td>Agarose type VII (low melting temperature)</td>
<td>Sigma</td>
</tr>
<tr>
<td>Amido Black 10B</td>
<td>Sigma</td>
</tr>
<tr>
<td>Ammonium acetate</td>
<td>Fisons</td>
</tr>
<tr>
<td>Ammonium persulfate</td>
<td>Life Technologies</td>
</tr>
<tr>
<td>Ampicillin (sodium salt)</td>
<td>Sigma</td>
</tr>
<tr>
<td>Amplify fluorographic reagent</td>
<td>Amersham</td>
</tr>
<tr>
<td>Aprotinin</td>
<td>Sigma</td>
</tr>
<tr>
<td>BALB/c mouse bone marrow λgt11 cDNA library</td>
<td>Clontech</td>
</tr>
<tr>
<td>BamH I (20 000 U/ml)</td>
<td>Pharmacia</td>
</tr>
<tr>
<td>Boric acid</td>
<td>Fisons</td>
</tr>
<tr>
<td>Bromophenol Blue</td>
<td>Sigma</td>
</tr>
<tr>
<td>BSA, fraction V</td>
<td>Sigma</td>
</tr>
<tr>
<td>CaCl2</td>
<td>ICRF</td>
</tr>
<tr>
<td>Carcinoembryonic antigen (CEA)</td>
<td>Calbiochem</td>
</tr>
<tr>
<td>Calf intestinal alkaline phosphatase</td>
<td>Pharmacia</td>
</tr>
<tr>
<td>Caesium Chloride</td>
<td>Fisons</td>
</tr>
<tr>
<td>Chloroquine</td>
<td>Sigma</td>
</tr>
<tr>
<td>Collagen I (human)</td>
<td>Telios</td>
</tr>
<tr>
<td>Collagen III (human)</td>
<td>Telios</td>
</tr>
<tr>
<td>Collagen IV (human)</td>
<td>Sigma</td>
</tr>
<tr>
<td>Collagen IX (human)</td>
<td>Sigma</td>
</tr>
<tr>
<td>Collagen X (human)</td>
<td>Sigma</td>
</tr>
<tr>
<td>Coomassie Brilliant Blue R-250</td>
<td>Merck</td>
</tr>
<tr>
<td>DAB</td>
<td>Sigma</td>
</tr>
<tr>
<td>DEAE Dextran (Mw 500 000)</td>
<td>Sigma</td>
</tr>
</tbody>
</table>
Deoxycholic acid Sigma
Dextran sulphate (sodium salt) Pharmacia
Dithiothreitol Sigma
DMEM ICRF
DMSO Merck
DNA polymerase I Klenow fragment Pharmacia
DNase I (bovine pancreas) Pharmacia
Dried skimmed milk (Marvel) Sainsbury
EDTA Merck
ECL western blotting detection system Amersham
EcoR I (25 000U/ml) Pharmacia
Ecoscint A scintillation solution National Diagnostic
Ethanol (99.7%) Merck
Ethidium Bromide (10 mg/ml) Sigma
FCS Life Technologies
Ficoll-Hypaque Pharmacia
Fluorescein diacetate Sigma
Formaldehyde Merck
Formamide Fluka
Geneclean II kit Stratech Scientific
Glutamine Sigma
Glutaraldehyde Sigma
Guanidine HCl Sigma
Guanidine isothiocyanate (enzyme grade) Life Technologies
G418 sulphate (Geneticin) Life Technologies
Hae III DNA fragments Life Technologies
Ham's F10 Life Technologies
Heparin (injection 1000U/ml) CP Pharmaceuticals
HEPES Life Technologies
Hind III (18 000U/ml) Pharmacia
Horse Serum Life Technologies
Hydrocortisone Sigma
Hyaluronic acid Sigma
IMDM Life Technologies
Iodo-beads Sigma
Isopropanol Life Technologies
IPTG Sigma
Laminin (human) Life Technologies
Leupeptin Sigma
LB - medium
LB - agar
Lysozyme
Magic PCR preps DNA purification system
Maloney murine leukemia virus transcriptase
Maltose
Methylprednisolone
Methionine-free DMEM
MgCl₂
MgSO₄
Mineral oil
Nco I (5 000U/ml)
Nova Blue competent cells
Nu serum
Nonidet P-40
NZC broth
pBluescript SK
pT7Blue T-Vector Kit for PCR cloning
PBS
Penicillin
PEG 8 000
PEG 20 000
PMSF
Phenol
Potassium acetate
Propidium iodide
Protein A Sepharose
Proteinase K
Protogel acrylamide (37.5:1 ratio of acrylamide:bisacrylamide)
Pst I (15 000 U/ml)
Ralmount
Random Hexamer
Restriction digest buffers 1-7 (10x)
RNase I "A" (bovine pancreas)
RNase inhibitor
RPMI
Salmon Sperm DNA
Sequenase 2.0 kit

ICRF
ICRF
Sigma
Promega
Life Technologies
Sigma
Sigma
ICRF
ICRF
Merck
Sigma
Pharmacia
AMS Biotechnology
Serva
Sigma
Life Technologies
Stratagene
AMS Biotechnology
ICRF
Life Technologies
Fisons
Fisons
Sigma
Life Technologies
Sigma
Sigma
Pharmacia
Sigma
National Diagnostic
Pharmacia
BDH
Boehringer Mannheim
Life Technologies
Pharmacia
Promega
Life Technologies
Sigma
USB
SSC  
SDS  
SDS molecular weight markers  
Sephadex G 25  
Sfi I (10 000 U/ml)  
Silver Stain kit  
Sodium citrate  
Sodium N-lauroylsarcosinate  
Soybean trypsin inhibitor  
Streptomycin  
Taq polymerase (5 U/µl)  
Taq polymerase reaction buffer  
TEMED (electrophoresis grade)  
Trizma base  
Tween-20  
Triton X 100  
Trypsin  
T4 DNA ligase (FPLC pure)  
X-gal  
XL1-Blue rec A-competent cells  
ZnCl₂  
α-MEM  
2-mercaptoethanol  
2'-Deoxyadenosine-5'-Triphosphate  
2'-Deoxyctydine-5'-Triphosphate  
2'-Deoxyguanosine-5'-Triphosphate  
2'-Deoxythymidine-5'-Triphosphate  
5' RACE System kit  
[α-³⁵S] dATP (1000 Ci/mmol 10 µCi/µl)  
[α-³⁵P]dCTP (3000 Ci/mmol, 10 µCi/µl)  
³⁵S-Methionine (80 Ci/mmol, 1mCi/ml)  
⁵¹Chromium ( 250-500 mCi/mg, 1mCi/ml)  
¹²⁵I-labelled Protein A (30 mCi/mg)  
¹²⁵Iodine (17.4 mCi/µg iodine, 100 mCi/ml)  

Miscellaneous materials  

Autoradiography film  

ICRF  
Merck  
Sigma  
Pharmacia  
Pharmacia  
Bio-Rad  
Merck  
Fluka  
Sigma  
Life Technologies  
Promega  
Promega  
Life Technologies  
Sigma  
Merck  
Sigma  
Sigma  
Pharmacia  
Sigma  
Stratagene  
Sigma  
Life Technologies  
Sigma  
Pharmacia  
Pharmacia  
Pharmacia  
Pharmacia  
Life Technologies  
Amersham  
Amersham  
Amersham  
Amersham  
Amersham  
Kodak XAR  
Amersham
2.1.2. Antibodies

Monoclonal antibodies (MoAb):
- CD34 antibody: ICH3 (Watt et al., 1987) - mouse, IgG2a, kindly provided by Dr. R. Levinsky (ICH, London)
- CD33 antibodies: WM53 - supernatant, mouse, IgG1 (Seralab); Anti-Leu-M9 - mouse, purified IgG1 (Becton Dickinson)
- CD66 antibodies (Watt et al., 1991): YTH 71.3.2 - rat, IgG2a; YPC 2/12.1-rat IgG2a; CE6/2D3.1 - rat, IgG1, all kindly provided by Dr. S. Watt
- VCAM-1 - mouse, purified IgG1 (Immunotech)

The CD34 and CD66 monoclonal antibodies were provided as culture supernatants or ascites and were used purified on MAb Trap G kit (Pharmacia), following the suppliers instructions.

Horseradish Peroxidase (HRP) - conjugates (Amersham)
- sheep F(ab')2 anti-mouse Ig
- sheep F(ab')₂ anti-rat Ig
- sheep F(ab')₂ anti-human Ig
- Protein A

Streptavidin labelled (Dako)
- horseradish peroxidase

Biotinylated antibodies (Dako)
- rabbit anti-mouse Ig

Fluorescein isothiocyanate (FITC) - labelled antibodies (Dako)
- rabbit F(ab)₂ anti-mouse Ig
- rabbit F(ab)₂ anti-rat Ig

2.1.3. Oligonucleotide primers

Oligonucleotide primers were synthesized on an Applied Biosystems Model 380A automatic synthesizer at Clare Hall Laboratories, ICRF. Oligonucleotide primer sequences are listed in Table 2.1.

2.1.4. Cell preparation and tissues

(a) cell lines

The cell lines used are listed in Table 2.2. They were obtained from the Cell Production Unit, ICRF. Multipotential haemopoietic progenitor cell line, A4 (FDCP mix) (Spooncer et al., 1984) was a gift from Dr. Elaine Spooncer, Paterson Institute, Manchester.

The cells were cultured at 37°C in a humidified 5% CO₂ atmosphere in medium as indicated in Table 2.2, supplemented with 10% (v/v) heat inactivated fetal calf serum (FCS), L-glutamine (2 mmol/L), penicillin (100 U/ml) and streptomycin (100 µg/ml). Suspension cells were maintained in exponential growth at 2 x 10⁵ to 10⁶ cells/ml. The medium of monolayer cultures was changed every 2-3 days and the cells passaged when close to
### Table 2.1: Sequence of oligonucleotide primers

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>human CD33 primers</strong></td>
<td></td>
</tr>
<tr>
<td>D1 (43-63) sense</td>
<td>5'-TGGGC(A/G/T)GG(A/G/T)GC(A/G/T)(C/T)T(A/G/T)GC(A/G/T)ATG-3'</td>
</tr>
<tr>
<td>D2 (361-381) sense</td>
<td>5'-TT(C/T)(C/T)(A/G)ATGGA(A/G)(A/G)(A/G/T)GG(A/G/T)-3'</td>
</tr>
<tr>
<td>D3 (1012-1032) anti-sense</td>
<td>5'-(G/A)TA(G/A)TG(A/G/T)A(G/A)T(C/T)TC(T/C)TC(G/A)TCCAT-3'</td>
</tr>
<tr>
<td><strong>sense</strong></td>
<td></td>
</tr>
<tr>
<td>H1 (3-22)</td>
<td>5'-TTTTCCTCAGACATGCCGCTGC-3'</td>
</tr>
<tr>
<td>H2 (21-41)</td>
<td>5'-CTGCTACTGCTGCGCCCTGCT-3'</td>
</tr>
<tr>
<td>H3 (254-274)</td>
<td>5'-AAGTACAGGAGGAGACTCAGG-3'</td>
</tr>
<tr>
<td>H4 (732-753)</td>
<td>5'-TGGTATCTTTCCAGGAGATGGC-3'</td>
</tr>
<tr>
<td>H5 (853-873)</td>
<td>5'-ATAGTGAAGACCCACAGGAGG-3'</td>
</tr>
<tr>
<td><strong>anti-sense</strong></td>
<td></td>
</tr>
<tr>
<td>H6 (611-593)</td>
<td>5'-GTGATTATGACCCAGGAGG-3'</td>
</tr>
<tr>
<td>H7 (998-967)</td>
<td>5'-AACAGCCTGAGGTTTCAGTGGG-3'</td>
</tr>
<tr>
<td>H8 (1126-1108)</td>
<td>5'-TGATGCTTCTGGAGGGTC-3'</td>
</tr>
<tr>
<td>H9 (1142-1122)</td>
<td>5'-ATCTTCTAGCTGAGCTGATG-3'</td>
</tr>
<tr>
<td>H10 (1342-1329)</td>
<td>5'-TGGGACATGGAAGCTTTAAGG-3'</td>
</tr>
<tr>
<td><strong>GAPDH primers</strong></td>
<td></td>
</tr>
<tr>
<td>sense (831-853)</td>
<td>5'-ATCAAGAAGGTGGTGAAGCAGG-3'</td>
</tr>
<tr>
<td>anti-sense (1005-987)</td>
<td>5'-GTCATACAGGAAATGAGC-3'</td>
</tr>
<tr>
<td><strong>pT7BlueT-Vector primers</strong></td>
<td></td>
</tr>
<tr>
<td>T7 sense</td>
<td>5'-AATACGACTCACTATAG-3'</td>
</tr>
<tr>
<td>U19 anti-sense</td>
<td>5'-GTTCACCTGACGAGTC-3'</td>
</tr>
<tr>
<td><strong>Agg-t11-primers</strong></td>
<td></td>
</tr>
<tr>
<td>P7 sense</td>
<td>5'-GGTGGCAGCGACTCCTGGAGCCCG-3'</td>
</tr>
<tr>
<td>P8 anti-sense</td>
<td>5'-TTGACACCAGACCAGCTTGGGTAATG-3'</td>
</tr>
<tr>
<td><strong>pBluescript SK-primers</strong></td>
<td></td>
</tr>
<tr>
<td>SK sense</td>
<td>5'-CTCTGACACCTAGTGCC-3'</td>
</tr>
<tr>
<td>KS anti-sense</td>
<td>5'-CGAGGTGGTGATCG-3'</td>
</tr>
</tbody>
</table>
murine CD33 primers **
sense
M1 (2-22)  5’-CTGTGCTATGCTGTGCGAC-3’
M2 (24-43) 5’-CCGCTGTCTTGTGCTGTGG-3’
M3 (557-576) 5’-CTCACCTCCTGAGTTCC-3’
M4 (666-685) 5’-GGGTGTCACGTGGAAAGGAC-3’
M5 (867-884) 5’-AGGCAAGAGAGTCACA-3’
M6 (979-999) 5’-CAATCTGAGAAATCCAAGACC-3’
M7 (1110-1129) 5’-GCTCCTCCTGAGAAGACCT-3’
M8 (1251-1271) 5’-GAGAGAGCTGTCGTGCTCTTGA-3’
M9 (1463-1481) 5’-TCAGTCCTTCCTCTTCTTT-3’
M10 (1577-1594) 5’-CACAATTAGAAGGTGTGG-3’
M11 (1733-1750) 5’-TATCTATGACGACCAG-3’
M12 (1851-1871) 5’-CTGTTGTCACAGATATCTAC-3’

anti-sense
M13 (1770-1791) 5’-CACAGCAACATGCAATAGA-3’
M14 (1453-1474) 5’-CTAGCCTCTACATGCTTCT-3’
M15 (1280-1263) 5’-GGATATAAATGAAAGGAC-3’
M16 (1093-1074) 5’-GATCCTGAGATCTCCTCTGT-3’
M17 (931-914) 5’-GGAGTCACAGCATCAGAT-3’
M18 (826-808) 5’-GGTCTCCTTCTTCTGACAGA-3’
M19 (600-580) 5’-TCAGACGGAGAGCTGTGG-3’
M20 (386-380) 5’-TGTAAGAATATCCTCAACA-3’
M21 (330-313) 5’-GGTCTCCTTCTCAGATA-3’

5’ RACE primers
Anchor primer 5’-CUACCUACUACUAGGCGACCAGCTCGACTAGTACGGGIGGGIIGGIGIG-3’
Universal Amplification Primer 5’-CUACCUACUACUAGGCGACCAGCTCGACTAGTAC-3’

murine β-actin primers
sense (545-564) 5’-CGTGCTGACATCAAAAGAG-3’
anti-sense (973-954) 5’-TGGACAGTGAGCCAAGATG-3’

* The nucleotides in brackets for degenerate primers D1-D3 represent all potential coding combinations. A comprehensive pool of all possible combinations was synthesized.
** The numbering of the murine CD33 primers refers to the m33-B isoform.
Table 2.2: Cell lines and culture medium.

<table>
<thead>
<tr>
<th>Cell</th>
<th>Description</th>
<th>Culture medium</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Human</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CEM</td>
<td>T cell acute lymphoblastic leukemia</td>
<td>RPMI 1640</td>
</tr>
<tr>
<td>Daudi</td>
<td>Burkitt's lymphoma</td>
<td>RPMI 1640</td>
</tr>
<tr>
<td>HL-60</td>
<td>promyelocytic leukemia</td>
<td>IMDM</td>
</tr>
<tr>
<td>KG1</td>
<td>bone marrow acute</td>
<td>IMDM</td>
</tr>
<tr>
<td>K562</td>
<td>chronic myelogenous leukemia</td>
<td>RPMI 1640</td>
</tr>
<tr>
<td>MOLT4</td>
<td>T cell acute lymphoblastic leukemia</td>
<td>RPMI 1640</td>
</tr>
<tr>
<td>RAJI</td>
<td>Burkitt's lymphoma</td>
<td>RPMI 1640</td>
</tr>
<tr>
<td>U937</td>
<td>acute myelogenous leukemia</td>
<td>RPMI 1640</td>
</tr>
<tr>
<td><strong>Mouse</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>410.4</td>
<td>breast carcinoma</td>
<td>DMEM</td>
</tr>
<tr>
<td>A4 (FDCP mix)</td>
<td>multipotential hemopoietic progenitor cell line</td>
<td>Fisher's medium*</td>
</tr>
<tr>
<td>A-20</td>
<td>B lymphoma</td>
<td>DMEM</td>
</tr>
<tr>
<td>EL-4</td>
<td>T lymphoma</td>
<td>DMEM</td>
</tr>
<tr>
<td>M1</td>
<td>myeloblast</td>
<td>DMEM</td>
</tr>
<tr>
<td>NS-1</td>
<td>myeloma</td>
<td>DMEM</td>
</tr>
<tr>
<td>P388</td>
<td>monocyte-macrophage</td>
<td>DMEM</td>
</tr>
<tr>
<td>P815</td>
<td>mastocytoma</td>
<td>DMEM</td>
</tr>
<tr>
<td>RMA</td>
<td>T lymphoma</td>
<td>RPMI 1640</td>
</tr>
<tr>
<td>WEHI 3B</td>
<td>myelomonocytic leukemia</td>
<td>DMEM</td>
</tr>
<tr>
<td><strong>Others</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHO</td>
<td>Chinese Hamster Ovary cells</td>
<td>Ham's F-10</td>
</tr>
<tr>
<td>COS-7</td>
<td>SV 40 transformed African Green Monkey kidney, derived from CV1cells</td>
<td>DMEM</td>
</tr>
</tbody>
</table>

*A4 cell line was a gift from Dr. Elaine Spooncer, Paterson Institute, Manchester and was cultured as described in Spooncer et al. (1984).*
confluency. For cell adhesion assays cells were diluted to $2 \times 10^5$ cell/ml with fresh medium on the day before analysis. Single cell suspensions of the monolayered cultures were obtained by trypsinisation with 0.12% ($w/v$) trypsin or with 2 mM EDTA in phosphate buffered saline (PBS: 20 mM phosphate buffer containing 145 mM sodium chloride pH 7.4). Cells were stored in liquid nitrogen in a freezing mixture consisting of 10% ($v/v$) DMSO, 40% ($v/v$) medium and 50% ($v/v$) FCS.

(b) **isolation of peripheral blood mononuclear cells and neutrophils**

Heparinised human peripheral blood, freshly drawn from consenting adults, was diluted 1:2 in DMEM, containing 10% ($v/v$) FCS, overlaid over Ficoll-Hypaque (density of 1.077 g/ml) and centrifuged at 400 x g for 25 minutes at 15°C. The mononuclear cells were collected from the interface and washed several times in DMEM or PBS. Peripheral blood neutrophils were enriched by centrifugation on a discontinuous gradient of Ficoll-hypaque with densities 1.077 and 1.114 g/ml, collected from the 1.114 g/ml interface and washed with DMEM or PBS.

(c) **murine tissues**

Liver, thymus, spleen, kidney, brain and bone marrow were obtained from 6-8 week-old BALB/c mice.

2.1.5. **Suppliers addresses**

The addresses of suppliers for reagents, materials and instruments are listed below.

- Amersham International plc - Amersham, Bucks, UK
- AMS Biotechnology (UK) Ltd - Witney, Oxon, UK
- Amicon Ltd - Upper Mill, Stonehouse, Glos, UK
- Beckman Instrument (UK) Ltd - High Wycombe, Bucks, UK
- Becton Dickinson UK Ltd - Oxford, UK
- Bio-Rad Laboratories Ltd - Hemel Hempsted, Herts., UK
- Boehringer Mannheim GmbH - Mannheim, Germany
- BPL, Bio Products Laboratory - Dagger Lane, Elstree, Herts., UK
- BRL Bethesda Research Laboratories - Gaithersburg, MD, USA
- Calbiochem Novabiochem - Nottingham, UK
- CP Pharmaceuticals Ltd - Wrexham, UK
- Clontech Laboratories Inc - Palo Alto, CA, USA
- Dako Ltd - High Wycombe, Bucks, UK
- Falcon - Becton Dickinson U.K. Ltd, Cowley, Oxon., UK
- Fisons Scientific Equipment - Loughborough, Leics., UK
- Fluka - Gillingham, Dorset, UK.
- Hoefer Scientific Instruments - Newcastle, Staffordshire, UK
- Hybaid Ltd - Teddington, Middlesex, UK
- Immunotech S.A. - Luminy, Marseille Cedex, France
- ICRF - Media Department, Clare Hall, South Mimms, Herts., UK
- Merck Ltd - Poole, Dorset, UK
- Kodak - Amersham, Bucks, UK
- Life Technologies Ltd - Pailsey, Refrewshire, UK.
- National Diagnostics - Itling Lane, Hessle, Hull, UK
- NEN Research Products - Du Pont (UK) Ltd, Stevenage, Herts, UK
- New Brunswick Scientific Co Inc - Edison, New Jersey, USA
- Nunc - Kamstrup, Roskdile, Denmark
- Perkin-Elmer Ltd - Beaconsfield, Bucks, UK
- Pharmacia Biotech Ltd - St Albans, Herts, UK
- Pierce & Warriner (UK) Ltd - Upper Northgate St Chester, UK
- Promega - Madison, WI, USA
- Schleicher & Schuell GmbH - Posfach 4.D-3354 Dassel Germany
- Sera-lab - Crawley-Down, Sussex, UK
- Serva - Heidelberg, Germany
- Sigma Chemical Co Ltd - Fancy Road, Poole Dorset, UK.
- Sterilin - Stone, Staffs, UK
- Stratagene Cloning Systems - La Jolla, Ca, USA
- Stratech Scientific Ltd - Luton, Beds, UK
- Telios Pharmaceuticals Inc - Science Park Road, San Diego, CA, USA
- U. S. B. United States Biochemical Corporation -Cleveland, Ohio, USA
2.2. Methods

2.2.1. Preparation of plasmid DNA

a) small-scale preparation of plasmid DNA

Bacteria transformed with the relevant plasmid were initially streaked out onto LB-agar plates (1% (w/v) NaCl, 0.5% (w/v) bacto-yeast extract, 1% (w/v) bacto-tryptone, 1.5% (w/v) bacto-agar) with the appropriate antibiotic present (15 μg/ml tetracycline and 100 μg/ml ampicillin). The plates were incubated overnight at 37°C to allow the growth of bacterial colonies. 5 ml of LB medium (1% (w/v) NaCl, 0.5% (w/v) bacto-yeast extract, 1% (w/v) bacto-tryptone) with antibiotics were inoculated with a single bacterial colony and incubated in a shaking incubator overnight at 37°C (Model G25 Incubator shaker, New Brunswick). 850 μl of transformed bacteria was added to 150 μl glycerol and stored at -70°C. 1.5 ml of bacterial culture was centrifuged at 12 000 x g for 1 minute at 4°C. The bacterial pellet was resuspended in 200 μl of Solution I (25 mM Tris-HCl pH 8.0, 50 mM glucose, 10 mM EDTA) with vigorous shaking and 400 μl of Solution II (0.2 N NaOH, 1% (w/v) SDS) and 200 μl of Solution III (3/5 M potassium acetate) were added. The mixture was shaken briefly, left on ice for 5 minutes and centrifuged at 12 000 x g for 10 minutes. The supernatant was collected and the DNA extracted by adding an equal volume of phenol:chloroform (1:1 mix). After centrifugation at 12 000 x g for 2 minutes, the supernatant was transferred to a fresh tube, an equal volume of isopropanol was added and the mixture left at room temperature for 5 minutes. The pellet of nucleic acids was recovered by spinning for 5 minute at 12 000 x g. The DNA pellet was washed in 70% ethanol, air dried and dissolved in 50 μl TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0), containing 20 μg/ml RNAase. The solution was heated for 10 min at 70°C and DNA stored at -20°C.
b) **large-scale preparation of plasmid DNA**

100 µl of transformed bacteria was inoculated into 500 ml of LB medium with antibiotic and the culture grown overnight at 37°C with shaking. The culture was treated as described above for small-scale preparation (up to the phenol/chloroform extraction stage) except that the volumes were scaled up - 20 ml of Solution I, 40 ml of Solution II and 20 ml Solution III were used. The bacterial lysate was centrifuged at 3 000 x g for 20 minutes at 4°C and the supernatant filtered through four layers of cheesecloth. An equal volume of isopropanol was added and the mixture left at room temperature for 30 minutes. The nucleic acids were recovered by centrifugation at 3 000 x g for 30 minutes at 20°C. The pellet was dried and redissolved in 10 ml of high salt TE buffer (20 mM Tris pH 8.0, 10 mM EDTA). For each ml of solution 1 g of caesium chloride and 80 µl of 10 mg/ml ethidium bromide were added. The gradient was run at 55 000 rpm for 24 hours at 20°C in a Beckman rotor Ti 70.1 (Beckman L8-M Ultracentrifuge). The closed circular plasmid DNA was drawn off with a syringe avoiding the RNA and debris at the bottom. Ethidium bromide was removed by (5-6) butanol extraction. The plasmid DNA was precipitated by adding 1 volume of 1 M ammonium acetate and 2 volumes of 95% ethanol, spun immediately at 3 000 x g for 30 minutes at room temperature, washed in 75% ethanol, dried and resuspended in TE buffer. The concentration and purity of DNA were determined by measuring the optical density at 260 and 280 nm (1 OD<sub>260</sub>=50 µg/ml DNA; pure preparation OD<sub>260</sub>/OD<sub>280</sub> > 1.8)

c) **restriction endonuclease digestion of plasmid DNA**

Approximately 1-2 µg of plasmid DNA in 10 µl final volume was digested with 10-15 Units of the appropriate restriction enzyme for 1 hour at 37°C in the recommended restriction digest salt buffer. The digestion
mix was separated by agarose gel electrophoresis in 0.5 x TBE buffer (0.5 x TBE: 45 mM Tris-borate, 1 mM EDTA) at 80 V and visualised by ethidium bromide staining (10 μg/ml final concentration).

2.2.2. Production of CHO cell transfectants, stably expressing CD34, CD33 and BGPC proteins

cDNAs for CD34 (Simmons et al., 1992), in the pCDM8 vector, and CD33 (Simmons and Seed, 1989), in the pH3M vector, antigens were kindly provided by Dr. D. Simmons, ICRF, and for BGPC (Watt et al., 1994b), in pH3M vector, by Dr. S. Watt. These were used to produce stable transfectants in CHO cells.

CHO cells were grown in Ham's F10 with 10% (v/v) FCS to 70% confluency. The cells were removed from one 75 cm² flask (approximately 3 x 10⁶ cells) for each gene, with PBS, containing 2 mM EDTA to give a single cell suspension and washed twice in DMEM without serum. The cells were made up to 1.5 ml in DMEM without serum and 0.75 ml placed into an electroporation cell (Gene Pulser Transfection Apparatus, Bio Rad). The plasmids containing the CD34 and CD33 defined cDNAs were linearised with Hind III. The plasmid containing the BGPC cDNA was linearised with SfiI. The pSV2neo construct was linearised with BamHI. Linearised plasmids containing the cell surface protein genes were added with linearized pSV2neo at 10:1 concentration by weight (i.e. 15 μg of the appropriate linearised gene to 1.5 μg linearised pSV2neo) and electroporated into CHO cells using 400V and 500μF. The cells were plated into one 25 cm² flask in Ham's F10 medium with 10% (v/v) FCS. After 48 hours the transfectants were selected in 400 μg/ml G418 and the medium changed every 2 to 3 days. The stable CHO-CD34, CHO-CD33 and CHO-BGPC transfectants were selected by three cycles of fluorescence
activated cell sorting with the respective CD34, CD33 and CD66 antibodies as described below.

2.2.3. Flow cytometric analysis and cell sorting

Cells were harvested, washed in PBS, 0.2% (w/v) BSA and resuspended in the same medium at a concentration of 2 x 10^7 cells/ml. Fc receptor binding was blocked by adding 1% (v/v) human serum (heat inactivated at 56°C for 30 minutes) to the cell suspension for 30 minutes at 4°C. Cells were washed once in PBS with BSA and suspended to 2 x 10^7 cells/ml in the same medium. Saturating levels of primary antibodies (CD33, WM53 supernatant - 1:10 dilution; CD34, purified ICH3 antibody - 40 μg/ml and CD66, purified YPC2/12.1 and CE6/2D3.1 antibodies - 20 μg/ml of each) were added to the cells for 30 minutes at 4°C. Cells were washed twice and made up to 2 x 10^7 cells/ml in medium, containing the appropriate FITC-conjugated rabbit F(ab)_2 anti-mouse Ig or anti-rat Ig (1:40 dilution). The cells were incubated as before, washed twice and resuspended to 2 x 10^6 cell/ml in PBS, 0.2% (w/v) BSA for cell sorting. For the negative control cells were incubated with PBS, 0.2% (w/v) BSA without the relevant first layer antibody followed by the respective FITC conjugate.

Cells were analysed on the FACScan (Becton Dickinson) and sorted on a FACStar Plus (Becton Dickinson) on the basis of forward and orthogonal light scatter and fluorescein fluorescence using a single tunable Argon ion laser set at an output power of 0.3W and an emission wavelength of 488 nm. Cells were sorted under aseptic conditions at a rate of 500 - 1000 cells per second into Ham's F10 medium containing 15% (w/v) FCS. Only the highest 10-30% of fluorescent cells were sorted. To ensure collection of viable cells, dead cells were excluded both by scatter and propidium iodide staining (4 μg/ml final concentration). For each cell
type at least three sorts were performed at intervals of one week to generate a population of strongly positive cells (see Fig. 3.1). The CHO transfectants were periodically checked for expression by FACScan analysis (see Fig. 3.2) and analysed by western blotting as described below.

2.2.4. Western blot analyses of CHO cell transfectants

(a) Lysing CHO cells

The confluent CHO monolayer of one 75 cm² flask (approximately 3 x 10⁶ cells) was washed three times with ice cold PBS and overlaid with 3 ml of lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1% (v/v) Nonidet P-40, 0.5% (w/v) deoxycholic acid, 0.1% (w/v) SDS, 1% (v/v) aprotonin, 200 mM EDTA, 1 mM PMSF) for 45 minutes on ice. The cells were scraped from the flask and the lysate was centrifuged at 15 000 x g for 5 minutes in a microfuge. If the mixture appeared viscous the DNA was sheared by passing several times through a 23-gauge needle. After spinning at 15 000 x g, 1 ml of cell lysate was precleared for 1 hour at 4°C by end-over-end rotation with 50 μl of Protein A Sepharose beads prewashed in the lysis buffer. The cell lysate was centrifuged at 15 000 x g and separated by SDS-polyacrylamide gel electrophoresis as described below.

(b) SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

The samples were separated by 7.5% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) in discontinuous buffer system according to Laemmli (1970). They were heated for 5 minutes at 100°C in sample buffer (0.0625 M Tris-HCl pH 6.8, 10% (v/v) glycerol, 1% (w/v) SDS, 0.01% (w/v) Bromophenol Blue and for reducing conditions 5% (v/v) β-mercaptoethanol). The SDS molecular weight markers were treated similarly to the samples and were myosin, 205 kD; β-galactosidase, 116
kD; phosphorylase b, 97 kD; bovine serum albumin, 66 kD; ovalbumin, 45 kD; carbonic anhydrase, 29 kD. Samples were electrophoresed at 60 V in the stacking (3% acrylamide) and at 160 V in the separating (7.5% acrylamide) gel in a Tris-glycine running buffer (0.05 M Tris-HCl pH 8.3, 0.384 M glycine, 0.1% (w/v) SDS) (gel plate apparatus 16 cm x 12 cm x 0.75 mm, Bio-Rad). For the miniature slab gel electrophoresis unit (8.3 cm x 6.3 cm x 0.75 mm; Mighty II, Hoeffer Scientific Instruments) the samples were run at 60 V and 90 V in the stacking and resolving gels, respectively. After completion of the electrophoresis the gels were processed for immunoblotting as described below or stained for 2 hours with 0.1% (w/v) Coomassie Brilliant Blue R-250 in 25% (v/v) methanol, 10% (v/v) acetic acid at room temperature with gentle agitation. Gels were destained in 25% (v/v) methanol, 10% (v/v) acetic acid and were vacuum dried at 80°C for 1-2 hours. When more sensitive staining was required (1-10 ng protein per band), the gels were stained with silver salts (Merril et al., 1981), using Silver Stain kit (Bio-Rad).

(c) western blot analysis

For western blot analysis, after SDS-PAGE the gel was soaked in transfer buffer (50 mM Tris-HCl pH 8.3, 380 mM Glycine, 0.1% (w/v) SDS, 20% (v/v) methanol) for 15 minutes. The separated proteins were transferred to nitrocellulose Hybond-C Extra or Hybond-C Super membranes (Amersham) in a Trans-Blot Electrophoretic Transfer Cell (Bio-Rad), according to the manufacturers instructions, at 20 V overnight and for another 90 minutes at 60 V with cooling. Proteins from the mini-gels (8.3 x 6.3 cm x 0.75 mm, Mighty II, Hoeffer) were transferred for 1.5 hours at 90 V with cooling. To check the successful completion of protein transfer, the gels were stained with Coomassie brilliant blue or silver salts as described above (2.2.4b) or the nitrocellulose membrane was
stained with 0.1% (w/v) Amido Black 10B, 45% (v/v) methanol, 10% (v/v) acetic acid for 5 minutes and destained with 90% (v/v) methanol, containing 2% (v/v) acetic acid.

Excess binding sites on nitrocellulose membranes were blocked with 5% (w/v) dried skimmed milk in PBS, 0.15% (v/v) Tween-20 at 4°C overnight. The blocked membrane was washed in PBS, 0.5% (v/v) Tween-20 once for 15 minutes and twice for 5 minutes at room temperature. The membrane was cut into strips (12 x 0.5 cm) for incubation with 1-2 ml of appropriate antibodies, diluted in PBS, 0.15% (v/v) Tween-20, 1.5% (w/v) dried skimmed milk for 1 hour at room temperature. The optimal concentration for each antibody was determined by titration and for both CD66, YTH71.3.2., and for CD34, ICH3, antibodies concentration was determined as 20 µg/ml. The unbound antibody was removed with three washes as above. Nitrocellulose strips were incubated with the appropriate HRP-labelled sheep F(ab')2 anti-rat Ig or anti-mouse Ig (1:300 dilution). The excess second antibody was removed with five washes (1 x 15 minutes and 4 x for 5 minutes) with PBS, 0.5% (v/v) Tween-20. The blots were developed with the enhanced chemiluminescence western blotting system (ECL) (Amersham) according to the suppliers instructions. They were incubated in equal volumes of ECL detection reagents 1 and 2 using 0.125 ml cm⁻² for 1 minute at room temperature, the excess detection reagent was drained and the blots were wrapped in Saran Wrap and exposed to autoradiography film (Kodak-XAR) for 30 seconds to 10 minutes depending on the amount of target protein.
2.2.5. Production of CD34-Fc and CD33-Fc fusion proteins

(a) CD34-Fc and CD33-Fc constructs

The constructs for production of soluble CD33-Fc and CD34-Fc fusion proteins were kindly provided by Dr. D. Simmons. They were created by the genetic fusion of the Fc genomic fragment of human IgG1 molecule and the respective extracellular domains of CD33 and CD34 proteins. The pIG vector, containing the Fc genomic fragment of IgG1 is shown in Fig. 2.1. In brief the IgG1 gene segment containing hinge, CH2 and CH3 domains was amplified from human genomic DNA from K562 cells by PCR. After subcloning into pBluescript, the Fc genomic fragment was moved as a Hind III/Not I fragment from Bluescript and ligated into Hind III/Not I digested BamH I-minus pCDM8, bringing in the additional cloning sites, Hind III/EcoR V - EcoR I - Pst I - Sma I and BamH I, at the 5' extracellular domain cloning site.

![Figure 2.1. Schematic representation of the pIG vector.](image-url)
The cDNA sequences encoding the whole extracellular domains of CD34 and CD33 were amplified by PCR, using primers incorporating additional digestion sites (Watt et al., 1994b). The digests of pIG vector and of CD33 and CD34 extracellular domains respectively were purified on 0.8% (w/v) agarose gels prior to ligation with T4 ligase. After MC1061 bacterial transformation with these constructs, plasmid preparations of the CD34-Fc and CD33-Fc were purified on caesium chloride gradient as described in 2.2.1(b).

(b) transfection of Cos 7 cells

The Fc plasmid constructs were transfected into Cos 7 cells using the DEAE Dextran method (Aruffo and Seed, 1987). The confluent Cos 7 cells from one 75 cm\(^2\) flask were trypsinised and replated into three 15-cm petri dishes. The cells in the dishes were grown for an additional two to three days in DMEM, containing 10% (v/v) FCS to 50 - 70% confluency. The plates were washed with DMEM medium without serum and to each plate was added 15 ml of DMEM containing, 10% (v/v) Nu serum, 400 µg/ml DEAE Dextran, 2-5 µg/ml of DNA for transfection and 100 µM chloroquine (chloroquine was added last just before adding the 15 ml of the final mixture to each plate). The transfection mix was added dropwise over each portion of the plate. The cells were left for two to four hours at 37°C. The medium was then aspirated and the cells were shocked by adding 15 ml of DMEM, containing 10% (v/v) DMSO for one to two minutes at room temperature. The cells were washed three times with 10 ml of DMEM and incubated overnight with DMEM, containing 10% FCS. On the next day the medium was changed with DMEM containing 1% FCS and the cells allowed to grow for 7 days. Typically 10 to 20 15-cm petri dishes of Cos 7 cells were transfected with each plasmid Fc construct.
(c) **purification and characterisation of CD34-Fc and CD33-Fc proteins**

The fusion proteins secreted in the culture supernatants were collected at days 4 and 7. The supernatants were centrifuged to remove nonadherent cells and debris, pooled and the fusion proteins purified by Protein A-Sepharose chromatography using 0.1 M glycine HCl pH 3.0 as the elution buffer. The eluted proteins were desalted through a PD-10 desalting column, concentrated if necessary by Amicon Diaflo ultrafiltration (PM30 > 30 000 MW) according to the manufacturer's instructions and stored at 4°C. The purified fusion proteins were separated by 7.5% SDS-PAGE and analysed by western blotting with 30 μg/ml of CD34 antibody, ICH3, followed by HRP-labelled sheep F(ab')2 anti-mouse Ig (1:300 dilution) as described in 2.2.4(b) and (c).

Alternatively the reactivity of CD33-Fc and CD34-Fc soluble proteins with the respective antibodies was tested by dot blot analysis. CD33 (20 µl of WM53 supernatant) and CD34 (20 µl of 20 μg/ml ICH3 antibody) antibodies were applied directly to the Hybond-C Extra or Hybond-C Super nitrocellulose membrane as dots. The antibodies were added in aliquots of 5-10 µl and allowed to bind to the membrane for 1 hour at room temperature. The nitrocellulose membrane was washed three times for 5 minutes in PBS, blocked overnight with 5% (w/v) dried skimmed milk in PBS, 0.15% (v/v) Tween-20 and incubated with 30 μg/ml of purified CD33-Fc or CD34-Fc fusion proteins for 1 hour at room temperature. The membranes were processed for immunodetection, using HRP-conjugated sheep F(ab')2 anti-human Ig and the ECL detection system, as described in 2.2.4(c).
production of labelled CD33-Fc and CD34-Fc fusion proteins

Twelve hours after transfection, a fraction of Cos cells was seeded into a 75 cm² flask. Thirty six hours post transfection the cells were washed with DMEM and equilibrated with 10 ml of Methionine-free DMEM for 30 minutes at 37°C. ³⁵S-Methionine and dialysed FCS were added to a final concentration of 100 µCi/ml and 0.5% (v/v) respectively. The cells were allowed to incorporate the label overnight. The supernatant was harvested, centrifuged and incubated with 200 µl Protein A-Sepharose beads overnight at 4°C with end-over-end rotation. The beads were collected by centrifugation, washed twice with PBS and eluted with 200 µl of 0.1 M glycine-HCl buffer, pH 3.0. Aliquots of the samples were analysed by 7.5% SDS-PAGE. Before drying, the gels were incubated with Amplify fluorographic solution for 30 minutes at room temperature with occasional agitation. The dried gels were exposed to autoradiography film (Kodak XAR) and developed after 1-7 days.

2.2.6. Binding to stromal layers

(a) Long Term Bone Marrow Cultures (LTBMC)

These cultures were grown by Mrs. E. Dorey, ICRF, using the Dexter type long term culture system (Dexter et al., 1977). Human sternal bone marrow from patients undergoing cardiac surgery was used. In some of the experiments bone marrow was obtained from lymphoma patients in haematological complete remission undergoing autologous BM harvest prior to high dose therapy. All bone marrow samples were collected in IMDM, containing 200 U/ml heparin. The bone marrow aspirate was teased out and passed through a 16 gauge needle attached to a syringe to obtain a single cell suspension. The cells were spun for 8-10 minutes at 500 x g and resuspended at 1 x 10⁶ cells/ml maximum in long term culture medium (IMDM, 10% (v/v) FCS, 10% (v/v) horse serum, 10⁻⁸ β -
mercaptoethanol, 10^{-6} M methylprednisolone. They were seeded in 24 well tissue culture plates and maintained at 37°C in a humidified atmosphere of 4.5% CO_2. Once per week half of the supernatant medium was removed and replaced with fresh medium. After 4 to 8 weeks confluent homogeneous stromal layers free of haemopoietic activity were formed. They were washed twice with DMEM, blocked with 1% (w/v) heat-denatured BSA in DMEM for 1 hour at 37°C, washed and used for adhesion studies described below.

(b) binding of ^{51}Chromium-labelled CHO cells to stromal layers

The harvested CHO cells were centrifuged at 400 x g for 10 minutes and resuspended in 200-300 µl of DMEM with 10% (v/v) FCS. The cells were labelled with ^{51}Chromium (Cr) at 10-20 µCi/10^{6} cells at 37°C for 45-60 minutes with occasional mixing. The labelled cells were washed twice in 20 ml of fresh DMEM without serum and resuspended at 1 x 10^{6} cells/ml in serum free DMEM for the adhesion assays.

300 µl (300 000 cells) of labelled cell suspension was added to each well. The density of the cells was checked microscopically to ensure they were enough to form a confluent monolayer in the well. The cells were incubated for 70 minutes (10 minutes to settle and 60 minutes adherence) at 37°C. The samples were washed three times by gently exchanging the medium and the adherent cells were solubilised with 300 µl of 2 M NaOH and counted in a gamma counter. The results of adhesion assays are presented as the percentage of total counts recovered in the adherent layer and the values are the mean ± SD from four replicate wells.

For antibody blocking experiments, the labelled CHO cells (1 x 10^{7} cells/ml) were incubated with saturating amounts (100 µg/ml) of the appropriate antibodies for 30 minutes at 4°C, washed twice and then added to the stromal layers for adherence assays as described above.
(c) binding of CD34-Fc and CD33-Fc fusion proteins to stromal layers

0.5 ml (10 000 cpm) of $^{35}$S-labelled CD33-Fc or CD34-Fc fusion proteins (see 2.2.5d) were added to the stromal layers for 2 hours at 37°C. After three washes with DMEM, protein binding was determined by lysing the layers with 1 ml of 3% (v/v) Triton X-100, which was aspirated and mixed with 5 ml of ecscint scintillant and counted in a $\beta$-counter. The results are presented as described above.

2.2.7. Binding to ECM substrates

100 µl of the ECM components, fibronectin, fibrinogen, collagen I, III, IV, IX, X and hyaluronic acid, diluted in PBS at 20 µg/ml were adsorbed onto PVC 96-well plates overnight at 4°C. The plates were rinsed three times in PBS and uncoated plastic blocked by incubation for 1 hour with PBS, containing 1% (w/v) heat-denatured BSA. The plates were rinsed again and used for adhesion assays with $^{51}$Cr-labelled cells or CD33-Fc and CD34-Fc fusion proteins as described below.

(a) binding of $^{51}$Cr-labelled CHO cells to ECM substrates

100 µl (100 000 cells) of $^{51}$Cr-labelled CHO cells in DMEM (see 2.2.6b) were added to each well and incubated for 70 minutes at 37°C. The samples were washed three times by gentle immersion in PBS, containing 1 mM CaCl$_2$ and 1 mM MgCl$_2$. The adherent cells were lysed with 100 µl of 2 M NaOH and samples were quantitated as above (2.2.6b).

(b) binding of CD34-Fc and CD33-Fc fusion proteins to ECM substrates

100 µl (30 µg/ml) of the fusion proteins (see 2.2.5) in PBS, 0.1% (w/v) heat-denatured BSA was incubated for 2 hours at 37°C and then the
plates were washed three times with PBS, 0.1% (v/v) Tween-20. The binding was monitored by adding HRP-labelled sheep F(ab')\textsubscript{2} anti-human Ig or Protein A (1:1000) for 1 hour at 37°C. The plates were washed as above and 100 μl of substrate solution added - 1mM ABTS. After the desired colour intensity was achieved, the reaction was stopped by adding 100 μl of 0.1 M citric acid. The absorbance at 410 nm was measured using a Titertek Multiscan ELISA plate reader.

As an alternative way to monitor the binding of fusion proteins to ECM substrates, the soluble chimaeric molecules were preincubated for 1 hour at room temperature with \textsuperscript{125}I-labelled Protein A (1.5 μCi \textsuperscript{125}I-labelled Protein A with 200 μg fusion protein in 2 ml PBS, containing 1mM CaCl\textsubscript{2} and 1 mM MgCl\textsubscript{2}). 100 μl of this mixture was then added to the plates, incubated for 2 hour, washed, lysed with 100 μl of 2M NaOH and bound \textsuperscript{125}I-Protein A measured in a gamma counter.

When the binding of \textsuperscript{35}S-labelled fusion proteins (2.2.5d) to ECM substrates was determined 100 μl of \textsuperscript{35}S-labelled fusion proteins (2500cpm) were added per well and the adhesion assay was performed as above. The bound protein was dissolved with 100 μl 1M NaOH, added to 5 ml of ecoscintillant liquid and counted in a β- counter.

2.2.8. Binding to bone marrow ECM matrix

(a) extraction of bone marrow ECM

Bone marrow ECM was extracted according to the method of Campbell et al. (1985). The bone marrow was obtained from femurs and tibias of 4-week old New Zealand white rabbits and was placed immediately in a large volume of 50 mM Tris-HCl pH 7.4, containing 3.4 M NaCl, 1 μg/ml leupeptin and 10 μg/ml soybean trypsin inhibitor. All steps were performed at 4°C. The sample was homogenised in a polytron homogeniser for 30 seconds at high speed, homogenate was centrifuged at
1 500 x g for 10 minutes and the supernatant and fatty layer removed.
The pellet was washed exhaustively (5 x) in the 3.4 M NaCl solution,
digested with RNase (100 µg/ml) plus DNase (25 µg/ml) in DMEM for 1
hour at 37°C and washed overnight in 3.4 M NaCl, by end-over-end
rotation. The ECM was extracted from the pellet with 0.05 M Tris-HCl
buffer pH 7.4, containing 4 M guanidine-HCl and 2 mM dithiothreitol
with constant mixing for 16 hours at 4°C. The extract was centrifuged at 1
500 x g for 10 minutes and the supernatant dialysed exhaustively against
double distilled water, containing 1 µg/ml leupeptin and 10 µg/ml soybean
trypsin inhibitor. The sample was concentrated if necessary by 25% PEG
20 000. At this stage, samples from the dialysis including the insoluble
precipitates were lyophilised for storage or dissolved in Laemmli sample
buffer for SDS-PAGE. After SDS-PAGE and transfer of separated ECM
proteins to nitrocellulose membrane (see 2.2.4b and c), the membrane was
washed in 2.5% Triton X-100 for 30 minutes at room temperature to
promote renaturation of blotted proteins and blocked with 5% (w/v) dried
skimmed milk in PBS.

(b) binding of 51Cr-labelled cells to bone marrow ECM blots

51Cr-labelled CHO, KG1 and murine bone marrow cells (see 2.2.6b)
were incubated over the nitrocellulose blot using 1 x 10^6 cells cm^-2 of
nitrocellulose paper for 1 hour at 37°C or at room temperature. The blot
was rinsed very gently 5 times with DMEM. After fixation with 4% (v/v)
glutaraldehyde in PBS for 3 minutes the blots were dried and visualised
by autoradiography (Kodak XAR) at -70°C.

(c) binding of fusion proteins to bone marrow ECM blots

The CD34-Fc and CD33-Fc fusion proteins at a concentration of 20
µg/ml in PBS, 1mM CaCl2, 1 mM MgCl2 and 1.5% (w/v) dried skimmed
milk were incubated at 37°C for 1 hour, followed by HRP-labelled sheep F(\text{ab'})\text{2} anti-human Ig (1:300 dilution) and visualised as described in section 2.2.4(c).

### 2.2.9. Cell-cell binding studies

(a) **aggregation assay**

An aggregation assay for CHO cells was carried out according to Benchimol et al. (1989). Cells from cultures at the point of confluency were removed from plastic surfaces by a 3 minute incubation at 37°C with 0.12% (w/v) trypsin in PBS, containing 15 mM sodium citrate. The cells were washed twice with DMEM. Suspensions of 1 x 10^6 cell/ml in α-MEM, 0.8% (v/v) FCS, 10 μg/ml DNase I and 10 mM HEPES were prepared by passing several times through a pipette. 3 ml of the cell suspension were incubated at 37°C for 1 hour with agitation. Single cells remaining were counted and subtracted from the total number of cells in order to provide an estimate of the number of aggregated cells. Every experiment was repeated 4 times (in 6 replicates each). The results are expressed as % cells as aggregates ± SD from four representative experiments (n=6).

For the antibody blocking experiments the CHO cell transfectants (1 x 10^7 cells/ml) were preincubated in α-MEM containing 100 μg/ml of YTH71.3.2, CE6 or YPC2/12.1 antibodies for 30 minutes at 4°C. The cell suspension was washed, then diluted with α-MEM, 0.8% (v/v) FCS, 10 μg/ml DNase I to 1 x 10^6 cell/ml and the aggregation was determined as above.

In experiments to determine whether the parental CHO cell line could participate in aggregate formation, CHO-BGPc cells (1 x 10^6 cells/ml) were stained with 10 ng/ml of fluorescein diacetate for 30 minutes at 4°C. The cells were washed and mixed in a ratio 1:7 of CHO-BGPc : CHO cells and assayed for aggregation as described above. After 1
hour in suspension samples were applied to microscope slides and examined by light and fluorescence microscopy.

(b) **immunofluorescence**

Confluent CHO cell transfectants on plastic two chamber slide flasks were rinsed in PBS, fixed in 70% (v/v) acetone for 10 minutes and Fc binding blocked with 0.5% (w/v) human immunoglobulin in PBS for 30 min at room temperature in a humidified atmosphere. Cells were washed and stained with 20 μg/ml of YTH 71.3.2 or ICH3 antibodies in PBS, 0.2% (w/v) BSA and after extensive washing followed by appropriate FITC-conjugated rabbit F(ab)\(_2\) anti-rat or anti-mouse Ig (1:40 dilution). Cells were washed, mounted in a drop of 80% (v/v) glycerol in PBS, sealed with transparent nail varnish and examined by fluorescence microscopy.

(c) **binding of \(^{125}\)I-labelled CEA to CHO monolayers**

The CEA was iodinated using Iodo-beads (Pierce), following the manufacturer's instructions. Two Iodo-beads were washed twice with 1 ml of 50 mM phosphate buffer, pH 7.4. The beads were dried and incubated with 10 μl of \(^{125}\)Iodine solution (1 mCi) diluted with 200 μl phosphate buffer for 5 minutes. 10 μg of CEA in 100 μl of 0.1 M Tris-HCl buffer, pH 7.5 were added to the two preloaded Iodo-beads and the reaction was allowed to continue for 10 minutes at room temperature. The radiolabelled CEA was recovered by chromatography through Sephadex G-25. The CEA was labelled to a specific activity of 6 μCi/μg. 20 000 cpm of \(^{125}\)I-labelled CEA in PBS, 0.1% BSA was added to the confluent monolayers of CHO cells in 24 well tissue culture plates and incubated for 1 hour at 37°C. After three washes with PBS, the cells were lysed with 200 μl of 2M NaOH and counted in a gamma counter. The results are
presented as a percentage of the total counts recovered in the adherent layer.

(d) binding of $^{51}$Cr labelled cells to CHO monolayers

Confluent CHO cells on 24 well tissue culture plates were washed twice with DMEM and blocked for 1 hour with 1% heat denatured BSA in DMEM. The plates were rinsed again and 300 μl (300,000 cells) of $^{51}$Cr-labelled KG1, Raji, Daudi, HL-60, U937, CEM, MOLT-4 and peripheral blood mononuclear cells (see 2.1.4) in DMEM with 20 mM HEPES were added to each well and allowed to adhere for 70 minutes at 37°C. Unbound cells were removed by gently rinsing the wells three times with DMEM (200 μl). To each well 300 μl of 2 M NaOH was added and counted in a gamma counter.

(e) binding of cells to Cos cell transfectants

Cos 7 cells were transfected with cDNAs, encoding CD34, CD33, BGPC or VCAM-1 (kindly provided by Dr. D. Simmons) or were mock transfected using DEAE-Dextran as described in 2.2.5(b). They were trypsinised 12 hours after transfection, replated in 6-well tissue culture plates at 25% confluency to facilitate rosette scoring and cultured for 1-2 additional days before performing the adhesion assays. Each Cos cell transfectant was also seeded in two chamber slides and expression of desired molecule was monitored 48 to 72 hours post transfection, before the adhesion assay, by immunofluorescence as described in 2.2.9(b).

KG1, Raji, Daudi, HL-60, U937, CEM, MOLT-4 cell lines, diluted to $2 \times 10^5$ cells/ml on the day before analysis, and peripheral blood mononuclear cells (see 2.1.4) were washed several times in PBS and were resuspended in DMEM in the presence of 500 U/ml heparin and 20 mM HEPES. Forty eight hours post transfection Cos cell transfectants in the 6
well plates were overlaid with the above cells in 2 ml DMEM, 500 U/ml heparin and 20 mM HEPES and incubated for 70 minutes at 37°C or 4°C. Nonadhering cells were removed by gentle washing with DMEM, and the remaining cells were fixed in 4% (v/v) formaldehyde and examined by light microscopy.

2.2.10. Preparation of total RNA

RNA was isolated by the modified guanidinium thiocyanate method (Chirguin et al., 1979; Sambrook et al., 1989). To a fragment of tissue or a pellet of cells 5 volumes of guanidinium thiocyanate homogenization buffer (4 M guanidinium thiocyanate, 0.1 M Tris-HCl pH 7.5, 1% (v/v) 2-mercaptoethanol) was added. The sample was mixed until the cells were lysed and DNA sheared by passing several times through a 23-gauge needle. Sodium N-lauroylsarcosinate was added to the cell lysate to a final concentration of 0.5% (w/v) and the suspension centrifuged at 5 000 x g for 10 min. The supernatant was overlaid onto a cushion of 5.7 M caesium chloride, 10 mM EDTA pH 7.5 and centrifuged at 32 000 rpm for 24 hours at 20°C in Beckman SW41 rotor. The pellet of RNA was rinsed twice with 70% ethanol and suspended in TES buffer (10 mM Tris-HCl pH 7.6, 1 mM EDTA, 0.1% (w/v) SDS). 0.1 volume of 3 M sodium acetate pH 5.2 and 3 volumes of ice-cold ethanol were added and the mixture incubated for at least 30 minutes at 0°C. RNA was recovered by centrifugation at 12 000 x g for 10 minutes at 4°C, redissolved in a small volume of water and stored with 3 volumes of ethanol at -70°C. (To recover the RNA, an aliquot was removed, 3 M sodium acetate pH 5.2 added to a final concentration of 0.3 M and centrifuged at 12 000 x g at 4°C). The quality of RNA was analysed by 1% (w/v) agarose gel electrophoresis in TBE buffer. The amount and purity of RNA were determined spectrophotometrically (1 OD$_{260}$ = 40 μg/ml; pure preparation
The extracted RNA was used in RT-PCR analysis, 5'RACE and northern blot analysis.

2.2.11. RT-PCR analysis of mouse bone marrow and WEHI 3B cells

2 μg of total RNA (see above) from mouse bone marrow and WEHI 3B cells were reverse transcribed in 20 μl of buffer comprising 50 mM Tris-HCl pH 8.3, 3 mM MgCl\(_2\), 75 mM KCl, 10 mM dithiothreitol, 10 μg BSA, 20 U RNase Inhibitor, 1 μM random hexamer and 200 U of Moloney murine leukemia virus transcriptase. 1 to 5 μl of the resultant cDNA was used in each PCR reaction containing 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl\(_2\), 200 μM of each dATP, dGTP, dCTP and dTTP and 1 μM of each oligonucleotide primer. The pairs of primers used were D1/D2 and D2/D3 and H1 to H10 in all possible combinations. The positions of these primers are shown in Fig.4.1 and the primer sequences are listed in Table 2.1. After initial denaturation at 94°C for 10 minutes, 1 unit of Taq polymerase was added and PCR amplification was performed for 30 cycles in an automated Perkin Elmer Cetus DNA thermal cycler. Each cycle consisted of denaturing at 94°C for 1 minute, annealing at 57°C for 1 minute and elongation at 72°C for 2 minutes. Immediately after the last cycle, the sample was maintained at 72°C for 10 minutes. 20% of the amplification product was separated by 1.5% (w/v) agarose gel electrophoresis and visualised by ethidium bromide staining. HAE III DNA fragments of ΦX174 were used as size markers. To detect possible genomic contamination, for each sample, a control without reverse transcriptase was amplified as well.

A 358 bp PCR product amplified with H3/H6 primers from mouse bone marrow cDNA was further characterised by cloning and sequencing as described below.
2.2.11. Cloning of PCR product

The 358 bp PCR amplified product (see above) was cloned into pT7BlueT-Vector, designed for cloning PCR products (AMS Biotechnology). The cloning was based on the 3' A-nucleotide overhangs left by Taq polymerase in the PCR product and the compatible T-nucleotide overhangs provided within the pT7BlueT-Vector.

The 358 bp PCR amplified product was separated by 1.5% (w/v) agarose gel electrophoresis and gel purified with Magic PCR Preps DNA Purification System (Promega), according to the manufacturer's instruction. DNA recovered from the gel slice was resuspended in a total volume of 10 µl TE (about 20 ng/µl). Ligation reactions were performed for 16 hours at 16°C at a molar ratio of 0.15:1 vector to insert in 10 µl of buffer comprising 20 mM Tris-HCl pH 7.6, 5 mM MgCl₂, 5 mM DTT, 50 µg/ml BSA, 0.5 mM ATP and 8 Weiss units T4 DNA ligase. 1-2 µl of the ligation reaction were used to transform NovaBlue competent cells following the supplier's recommendations (AMS Biotechnology). The presence of appropriate insert in the recombinant clones (white colonies) was determined by direct colony PCR screening. The bacteria were transferred with a sterile toothpick to a 1.5 ml tube containing 50 µl sterile distilled water. The pellet was dispersed by vortexing and tubes placed in boiling water for 5 minutes to lyse the cells and denature DNases. The cell debris were removed by centrifugation for 1 minute at 12,000 x g and 10 µl of the supernatant amplified by PCR, as described in 2.2.11, using pT7Blue T-vector specific primers T7 and U-19 (Table 2.1). A positive clone was isolated and sequenced as described below. This clone will be referred to as H3/H6-PCR.
2.2.13. Sequence analysis

Double-stranded DNA of H3/H6-PCR clone was sequenced by the dideoxy termination method (Sanger et al., 1977) using Sequenase 2.0 kit (U.S.B.). Supercoiled plasmid DNA was obtained and purified on a caesium chloride gradient using the standard maxiprep procedure (see 2.2.1b). The sequence of the H3/H6-PCR clone was also confirmed by sequencing three additional clones, generated in three separate PCR amplifications of mouse bone marrow cDNA with primers H3/H6. The PCR products were purified and cloned into pT7Blue T-Vector as described for the H3/H6-PCR clone (2.2.12). The DNA for sequencing from these three clones was obtained by the following "mini-prep boiling" method, by which sufficient quality and quantity of DNA was produced. The bacterial pellet (obtained in the first step in 2.2.1a) was suspended in 250 μl TES (50 mM Tris-HCl pH 8.0, 50 mM EDTA, 15% (w/v) sucrose) and 20 μl of freshly prepared lysozyme solution (10 mg/ml) was added. The tube was boiled for 1 minute and the bacterial lysate centrifuged at 12 000 x g for 10 minutes at room temperature. The supernatant was precipitated by addition of 0.5 volume of 7.5 M ammonium acetate and one volume of isopropanol for 10 minutes at -70°C. The DNA pellet was recovered by centrifugation at 12 000 x g for 15 minutes at 4°C, washed with 70 % ethanol and redissolved in 18 μl TE pH 8.0, containing DNase free pancreatic RNase (20 μg/ml). The typical yield of high-copy-number plasmids such as pT7Blue T-vector (pUC-based origin of replication) is about 3-5 μg per ml of original bacterial culture.

3-5 μg of template DNA was used per sequencing reaction. Both strands of each insert were sequenced entirely using specific oligonucleotide primers T7 and U19 (Table 2.1) in a 1:1 (primer:template) molar stoichiometry. Annealing reaction was performed at 72°C for 5 minutes, the mixture was then allowed to cool at room temperature for 10
minutes and stored on ice. 0.5 μl of [α-\(^{35}\)S]dATP (>1000Ci/mmol,10 μCi/μl) was used per labelling reaction. Labelling and termination reactions were carried out according to the manufacturer's instructions. Sequence analysis was performed on a 5% (w/v) polyacrylamide gel and the sequencing gel was run at 50°C at a constant voltage of 1800 V (Sequi-Gen Nucleic Acid Sequencing Cell, Bio-Rad). On completion of the run, the gel was fixed in 10% (v/v) acetic acid, 10% (v/v) methanol, blotted onto 3MM paper and dried at 80°C in a vacuum dryer. Autoradiography was carried out using high speed X-ray films (Kodak XAR or Amersham Hyperfilm-MP) without intensifying screens. The sequences were analysed as described below.

2.2.14. Computer analysis

Computer analyses of DNA and amino acid sequences were performed on a DAP 610 computer using Bio Search Software, developed by J. F. Collins and A.F.W. Coulson (1987). For assessing amino acid homologies, the program searches the compiled Swiss-Prot, PIR, Brookhaven and GenBank databases (containing 58 701 entries) and for nucleic acid sequences the EMBL database.

The ALIGN program was also used for direct comparison of the sequences. Pattern detection and searching was performed with PROSITE program. ALIGN and PROSITE are programmes from Intelligenetics Software and computer analysis with them was performed on a VAX computer.

2.2.15. Southern blot analysis

Southern blots of human and mouse genomic DNA were probed with the H3/H6-PCR clone. The \(^{32}\)P-labelled H3/H6-PCR probe was prepared as described below.
(a) **preparation of radiolabelled DNA probe**

The H3/H6-PCR insert was amplified with specific T7 and U19 primers, annealing to the sequences either side of the cloning site of pT7BlueT-vector (see 2.2.11, Table 2.1). The whole amplification product was separated by 1.5% (w/v) agarose gel electrophoresis and gel purified using a Magic PCR preps DNA purification kit, according to the manufacturer’s instructions.

The DNA probe was labelled by random priming (Feinberg et al., 1983). 29.5 µl of DNA (containing approximately 20-30 ng) were boiled for 5 minutes and labelled by adding sequentially 10 µl OLB* buffer (see below), 2 µl BSA (10 µg/ml), 5 µl [α-32P]dCTP (>3000Ci/mmole, 10 µCi/µl) and 2 µl (2 Units) DNA polymerase I Klenow fragment. The labelling was performed for 2.5 hours at 37°C. The reaction was stopped by adding 1 µl of 0.5 M EDTA and labelled DNA was purified by chromatography through a Nick Column. The probe was denatured by heating to 100°C for 5 minutes prior to hybridisation. Probes were labelled to a specific activity > 1 x 10^8 cpm/µg and used in Southern hybridisation as below.

*OLB buffer contained 100 µl of solution A (1.2 M Tris-HCl pH 8.0, 0.12 M MgCl2, 1.7% (v/v) 2-mercaptoethanol and 500 µM of each dATP, dGTP and dTTP), 250 µl of solution B (2 M Hepes pH 6.6) and 150 µl of solution C (90 Unit/ml random hexamers in 3 mM Tris, 0.2 mM EDTA pH 7.5).

(b) **isolation of genomic DNA**

Genomic DNA was isolated from human peripheral blood mononuclear cells and BALB/c mouse spleen cells. 1 x 10^7 cells were washed in PBS and lysed in 1 ml of lysis buffer (50 mM Tris HCl pH 7.5, 100 mM EDTA, 0.5% (w/v) SDS, 0.4 mg/ml Proteinase K) overnight at
55°C. The solution was extracted twice with phenol, followed by two chloroform extractions. DNA was precipitated with 0.1 volume of 3 M sodium acetate and 2 volumes of 100% ethanol. The DNA was recovered by centrifugation, washed in 70% ethanol, air dried and dissolved in 50 μl of TE, containing 20 μg/ml RNase. DNA concentration and quality were analysed as described for plasmid DNA (2.2.1).

(c) electrophoresis, transfer and hybridisation of restriction fragments of genomic DNA

A total of 10 μg of genomic DNA was digested with 5 μl (75 Units) Pst I and fractionated in 0.9% (w/v) agarose gel electrophoresis in TBE overnight at constant voltage of 30 V. The gel was denatured in 1.5 M NaCl, 0.5 M NaOH for 1 hour at room temperature by gentle shaking and neutralised in 0.5 M Tris pH 7.4, 1.5 M NaCl twice for 30 minutes. DNA fragments were transferred from agarose gel to Hybond N nylon membrane (Amersham) by the capillary method (Southern, 1975) in 20 x SSC (SSC: 150 mM NaCl, 15 mM trisodium citrate). DNA was fixed to the membrane by 3 minutes exposure to UV light. The blot was prehybridised with 10 ml of 5 x SSC, 5 x Denhardt's solution, 0.5% (w/v) SDS and 100 μg/ml denatured sonicated salmon sperm DNA in a hybridization bottle for 1 hour at 65°C (Hybaid Maxi Hybridization Oven System). 32P-labelled H3/H6-PCR probe (see 2.2.15a) was added to give a final 1-2 x 10^6 cpm/ml and allowed to hybridise overnight at 65°C. The blot was washed at 65°C in solutions with increasing stringency, starting with 2 x SSC, 0.1% (w/v) SDS, followed by 1 x SSC, 0.1% (w/v) SDS to 0.1 x SSC, 0.1% (w/v) SDS for 30-40 minutes each, depending on the monitored radioactivity background. After washing the filter was sealed in a plastic bag with 1 ml of final wash solution to prevent drying and exposed to fast X-ray film (Kodak XAR) with intensifying screens at -70°C.
2.2.16. Mouse bone marrow cDNA library screening

A BALB/c mouse bone marrow cDNA library in λgt11 (Clontech) was screened.

a) plating of bacteriophage λgt11

The host bacteria, Y1090-, was initially streaked onto LB agar plate and a single colony was inoculated into LB medium, containing 0.2% (w/v) maltose. The titer of the library was determined by plating serial dilutions (10⁻², 10⁻⁴, 10⁻⁶) of bacteriophages. Dilutions were prepared in 300 μl SM buffer (35 mM Tris-HCl pH 7.5, 10 mM MgSO₄, 0.1 M NaCl, 0.01% (w/v) gelatin) and to each tube 0.6 ml of the host bacteria were added. The bacteriophage particles were allowed to adsorb to the bacteria by shaking for 20 minutes at 37°C. 3 ml of molten (47°C) 0.7% (w/v) LB-agarose were added to each dilution, mixed gently and poured onto a 90-mm plate containing 30-35 ml of hardened bottom LB-agar. After 8 to 11 hours incubation at 37°C, plaques appeared and were counted. The estimated titer of the library was 10⁷ pfu/ml. A working aliquot of library was stored at 4°C. For long term storage an aliquot was stored in 7% (v/v) DMSO at -70°C.

b) hybridisation of bacteriophage plaques

Initially 150 000 bacteriophage plaques were screened with the H3/H6-PCR clone. They were plated at a density of 30 000 pfu per 150 mm petri dish, as described above by scaling up the volumes of top 0.7% LB-agarose to 10 ml and of the bottom LB-agar to 80 ml. The plates containing the infected bacteria were incubated for 8 to 10 hours at 37°C until the plaques were just beginning to make contact with one another but not showing confluent lysis. An imprint of the plaques was obtained.
by gently layering a nitrocellulose filter (Schleicher and Schuell) onto the surface of the top agarose. Two lifts per plate were performed. The filters were marked by stabbing through it and into the agar beneath with a needle containing waterproof black drawing ink. The DNA on the filters was denatured by floating for 30 seconds and immersing for 60 seconds in 1.5 M NaCl, 0.5 M NaOH and neutralized for 5 minutes in 1.5 M NaCl, 0.5 M Tris-HCl pH 7.5. After washing in 3 x SSC, DNA was fixed to the filter by baking for 2 hours at 80°C. Ten filters were hybridised together in a heat-sealed bag with 30 ml of prehybridisation solution. The hybridisation conditions were the same as described for Southern blotting in 2.2.15(a) and (c). Filters were washed at 65°C with 1 x SSC, containing 0.1% (w/v) SDS and then exposed to autoradiographic film.

Hybridising plaques, identified by aligning the film with the original agar plate, were picked with a sterile tip and placed in 1 ml of SM buffer, containing a drop of chloroform. If the alignment of the filters with the plate did not permit identification of an individual hybridising plaque, an agar plug containing several plaques was isolated. An aliquot of the bacteriophage stock (usually 50 µl of a 10^-2 dilution) was replated so as to obtain 500-1000 plaques on 90-mm plate. These plaques were then screened a second time by hybridisation. A single well-isolated positive plaque was picked from the secondary screen and was further analysed and cloned as described below. Four such positive plaques were isolated from this screening.

c) rapid PCR analysis of positive bacteriophages

Following identification of the four positive clones by hybridisation, 1 µl of the eluted agar phage stock was used per PCR reaction to determine the insert sizes of the clones. PCR was carried out (see 2.2.11) with λgt11 specific oligonucleotide primers, P7 and P8, annealing to the
flanking EcoR I cloning site sequences of the phage (Table 2.1, Fig. 4.3A). It was determined that the four clones contained one type of insert of 1330 bp and two of these PCR amplified inserts were cloned into pT7Blue T-Vector (see 2.2.12 and Fig. 4.3A). Partial sequencing showed that they are identical. The insert from one of the initial phage clones was subcloned into Bluescript SK plasmid (see below 2.2.17) and will be referred as p33-A clone. Most of the 5' end of the p33-A clone was used for generating a probe for rescreening the mouse bone marrow cDNA library as described below.

(d) **rescreening of the mouse bone marrow cDNA library**

300 000 pfu of the mouse bone marrow cDNA library were rescreened using as a probe, a 461 bp fragment 5' to the Bgl II digestion site in clone p33-A. This probe was referred to as Probe A (see Fig. 4.4A) and was produced by digesting 10-20 μg of p33-A clone DNA with EcoR I and Bgl II enzymes (75 units of each) for 4 hours at 37°C. The 461 bp fragment was gel purified with a Geneclean II purification kit (Stratech Scientific), following the supplier's instructions and labelled as described in 2.2.15(a). The hybridisation conditions for library rescreening were the same as described in 2.2.16(b). Eleven positive clones were identified and characterised by PCR using λgt11 specific primers, P7 and P8, as described above. Eight of the positive clones were identical to the initial p33-A clone containing an insert of 1330 bp. The other three clones have an insert of 1213 bp. One of these inserts was isolated from the λ vector and subcloned into Bluescript SK plasmid as described below. The 1213 bp clone will be referred to as p33-B.
2.2.17. Cloning of bacteriophage DNA into pBluescript SK vector (p33-A and p33-B clones)

The 1330 bp and 1213 bp inserts were isolated from λ vectors and subcloned into the Bluescript SK plasmid (see Fig. 4.3B), according to standard methods (Sambrook et al., 1989).

(a) **preparation of plate lysate stock**

10⁵ pfu were plated on a 90-mm plate (2.2.16a) and incubated at 37°C until semiconfluent lysis occurred. 5 ml of SM buffer were added to the plate and incubated for 5 hours at 4°C with occasional shaking. After harvesting as much of the SM buffer as possible, another 1 ml of fresh SM was added. After 15 minutes this was combined with the first harvest, 0.1 ml of chloroform added, the mixture vortexed briefly and centrifuged at 3000 x g for 10 minutes at 4°C. A drop of chloroform was added to the supernatant and the titer of the plaque stock determined (see 2.2.16a). The stock usually contained approximately 10¹⁰ to 10¹¹ pfu/ml.

(b) **extraction of bacteriophage DNA**

The bacterial host was grown overnight at 37°C in 50 ml NZC Broth medium, supplemented with 0.2% (w/v) maltose. 500 ml of prewarmed NZC Broth was inoculated with 1 ml of overnight culture and incubated at 37°C with vigorous shaking until the OD₆₀₀ of the culture reached 0.5. Approximately 5 x 10⁹ - 10¹⁰ phages (500 μl of 90-mm plate lysate stock, see above) were added to the culture and incubation was continued until lysis occurred. 2.5 ml of chloroform and solid NaCl to a final concentration of 1 M were added to the lysed culture and incubated for a further 15 minutes. Bacterial debris was removed by centrifugation at 11 000 x g for 10 minutes at 4°C. The bacteriophage particles were precipitated with PEG 8 000 (final concentration 10% (w/v)). PEG was dissolved by stirring.
at room temperature and the precipitate was allowed to form by leaving at 4°C for a minimum of six hours. The pellet of precipitated bacteriophage particles was recovered by centrifugation at 11 000 x g for 10 minutes at 4°C and resuspended in 13 ml caesium chloride solution (1g caesium chloride was added per ml of SM). The gradient was centrifuged at 32 000 rpm at 4°C for 24 hours in a SW41 Beckman rotor. The band of phage particles was collected by side puncture. The caesium chloride was removed from the purified bacteriophage preparation by dialysis against a 1000-fold volume of 10 mM NaCl, 50 mM Tris-HCl pH 8.0, 10 mM MgCl₂ for 2 hours at 4°C. The phage suspension was adjusted to 20 mM EDTA, 0.5% (v/v) SDS, 50 µg/ml proteinase K and incubated for one hour at 65°C. Following two phenol extractions, DNA was precipitated with two volumes of 100% ethanol, centrifuged at 12 000 x g for 15 minutes at 4°C, washed with 70% ethanol, dried and resuspended in TE.

(c) cloning of bacteriophage DNA into pBluescript vector

The inserts were isolated from λgt11 vectors by EcoR I digestion (Fig. 4.3B). Bacteriophage DNA (25-50 µg) was digested in a final volume of 200 µl with an excess of EcoR I (75-100 units) for 1 hour at 37°C. Digestion was analysed by agarose gel electrophoresis and the inserts were gel purified with a GeneClean II purification kit.

The vector used for cloning of the phage inserts was pBluescript SK digested with EcoR I restriction enzyme. Since the vector and the inserts were carrying identical protruding termini the 5'-phosphate groups were removed with alkaline phosphatase to suppress self-ligation and circularisation of the plasmid vector DNA. Digested vector was extracted with phenol:chloroform and precipitated with two volumes of ethanol for 15 minutes at 0°C. DNA was recovered by centrifugation and redissolved in 100 µl of dephosphorylation CIP buffer (10 mM Tris-HCl pH 8.3, 1 mM
ZnCl₂, 1 mM MgCl₂). 20 µg of linearised plasmid DNA was dephosphorylated with 1 unit calf intestinal alkaline phosphatase (CIP) for 30 minutes at 37°C. The CIP was removed by digestion with proteinase K (final concentration of 100 µg/ml) in the presence of 0.5 % (w/v) SDS and 5 mM EDTA for 30 minutes at 56°C. The dephosphorylated DNA was purified by extraction with phenol:chloroform, ethanol precipitated and resuspended in TE at a concentration of 100 µg/ml.

Ligations were set up at a molar ratio of 0.125:1 vector:insert in 10 µl ligation buffer (50 mM Tris-HCl pH 7.6, 10 mM MgCl₂, 10 mM DTT, 50 µg/ml BSA), containing 1 mM ATP and 8 Weiss units T4 DNA ligase. Ligations were performed overnight at 16°C.

100 µl of XL1-Blue rec A- competent cells were transformed with 1-2 µl of the ligation mixture. Cells were incubated on ice for 40 minutes, heat shocked for 90 seconds at 42°C, transferred onto ice for 1 minute and diluted in 0.4 ml of LB medium. After incubation for 60 minutes at 37°C cells were plated onto LB agar plates with 100 µg/ml ampicillin and 15 µg/ml tetracycline. 30 minutes prior to plating of transformants 100 µl of 100 mM IPTG and 40 µl of 2% (w/v) X-gal were spreaded on the plates. White colonies were analysed by direct colony PCR analysis using SK and KS primers (2.2.11, Table 2.1) or by Eco RI restriction digestion of small-scale preparations of plasmid DNA (2.2.1). Positive clones were grown for plasmid isolation and used for sequencing as described in 2.2.13. The sequencing strategy using specific oligonucleotide primers for p33-A and p33-B clones is shown in Fig. 4.4.

2.2.18. 5'RACE

The 5'RACE System was used to isolate the 5' end of the mouse CD33 cDNA according to the manufacturer's instructions. In brief total RNA was extracted from BALB/c mouse bone marrow, as described in
2.2.10., and the first strand cDNA synthesis was carried out, using M16 anti-sense mouse CD33 primer (Table 2.1, Fig. 4.5B). After cDNA synthesis, the original template was destroyed with RNase H and the cDNA purified by a GlassMAX Spin Cartridge. An oligo-dC anchor sequence was added to the 3' end of the cDNA using TdT and dCTP. The poly(C) tailed DNA was amplified by PCR using a 5'RACE anchor primer and M19, an anti-sense mouse CD33 primer. Additional rounds of PCR using 5'RACE Universal Amplification Primer and M19 mouse CD33 primer were performed to ensure an adequate level of specificity. PCR conditions were as described in 2.2.11. PCR products were cloned into pT7Blue T-Vector (2.2.12). The size of the inserts of recombinant clones were analysed by PCR using specific primers, T7 and U19, flanking the pT7BlueT-Vector cloning site. PCR products were size fractionated by agarose gel electrophoresis. These were further analysed by Southern blot hybridisation with probe A as described in 2.2.15 and 2.2.16(d) and the positive clones were sequenced using pT7Blue T-Vector specific primers, T7 and U19 and gene specific primers (see Fig.4.6 for sequencing strategy). The clone, containing the 5' end of the murine CD33 gene will be referred as p33-RACE.

The RACE derived 5' sequence was confirmed by sequencing three additional independent clones, amplified from mouse bone marrow RNA using M3 and M2 primers (Table 2.1, Fig. 4.6). These clones will be referred to as PCR a, b and c.

2.2.19. Northern blot analysis

20 µg of total RNA (isolated as in 2.2.10), was separated by 1% (v/v) formaldehyde gel electrophoresis (Sambrook et al., 1989) overnight at 40 V. The gel was incubated with gentle shaking successively in 50 mM NaOH, followed by 0.1 M Tris-HCl pH 7.5 for 30 minutes each. RNA from
the gel was transferred to Gene Screen Plus membrane (NEN) by the capillary method in 10 x SSC. DNA was fixed to the membrane by baking for 2 hours at 80°C. The blot was hybridised to the p33-B probe, isolated by EcoR I digestion, gel purified with Geneclean II kit and labelled as in 2.2.15(a). The nitrocellulose membrane was incubated with $^{32}$P-labelled p33-B probe (final 4 x $10^5$ cpm/ml) in 50% (v/v) formamide, 1% (w/v) SDS, 1 M NaCl, 10% (w/v) dextran sulphate and 100 µg/ml denatured salmon sperm DNA in a heat-sealed bag overnight at 42°C. The blot was washed at 60°C in 0.2 x SSC, 0.1% (w/v) SDS. The filter was not allowed to dry and after the appropriate exposure to X-ray film (Kodak XAR) was stripped by (4-5 times) pouring on the membrane a boiling solution of 0.01% (w/v) SDS in 0.01 x SSC. The blot was rehybridised to a 1.8 kb mouse β-actin probe, kindly provided by Dr. Alastair Reith (Ludwig Institute for Cancer Research, London Middlesex Branch).

2.2.20. **In situ** hybridisation

**In situ** hybridisation was performed by Dr. R. Poulsom, ICRF. In brief the p33-B clone was linearised with Nco I to serve as a template for T7 polymerase in the presence of $^{35}$S-UTP giving rise to an anti-sense RNA probe. The DNA template was removed by digestion with RNase-free DNase and the $^{35}$S-labelled riboprobe was separated from the unincorporated nucleotides by Chromaspin-30 column.

Wax was removed from the paraffin sections of fixed murine spleen, lymph nodes, thymus, brain, sternum, muscle, kidney and epithelial tumour W26, by incubation in fresh xylene. The sections were rehydrated through an ethanol series, permeabilised with Proteinase K and postfixed in 4% paraformaldehyde. The slides were treated with triethanolamine and acetic anhydride and then dehydrated through an ethanol series and air dried. For the **in situ** hybridisation analysis of bone marrow cells and
A4, WEHI3B, P388 and RMA cell lines, cytospins were prepared. These were fixed in 4% paraformaldehyde, dehydrated through an ethanol series and air dried.

Hybridisation was carried out with $^{35}$S-labelled p33-B anti-sense riboprobe overnight at 55°C in a humidified chamber. The slides were subjected to high stringency washing, followed by incubation with RNase A. The high stringency washing was repeated and the slides were dehydrated through an ethanol series and air dried. The slides were dipped in Ilford K5 emulsion at 45°C and dried at room temperature before exposure at 4°C. Slides were exposed for 6 days or more and then developed in Kodak D19 developer, rinsed in 1% acetic acid and fixed in 30% sodium thiosulphate. Following rinsing with water the slides were stained with Giemsa solution and examined microscopically under dark and bright field illuminations.

2.2.21. Immunohistochemical analysis of human CD33 antigen in thymus

The immunohistochemical staining of human thymus was performed by Ms L. Happerfield, ICRF. In brief the frozen sections of human thymus were thawed at room temperature, fixed in acetone for 30 minutes and washed in TBS (25 mM Tris-buffered saline, pH 7.6) for 5 minutes, followed by 30 minutes incubation with 100 μl of CD33 MoAb, Anti-Leu-M9 (1:10 dilution) in a humidified chamber. The slides were washed for 10 minutes with TBS and incubated for an additional 30 minutes with the second layer of biotinylated rabbit anti-mouse Ig (1:200 dilution). The slides were washed as above and the streptavidin-conjugated horseradish peroxidase was applied for 30 minutes. The slides were washed as above. DAB enzyme substrate solution was added (1.3 mM DAB, 0.02% (v/v) H$_2$O$_2$ in 5 mM Tris-HCl pH 7.6) and after the
desired coloured was achieved, the enzyme reaction was stopped by extensive washing with water. The sections were counterstained with haematoxylin for 5 minutes, exposed to ascending ethanol gradient, followed by 3 minutes incubation with xylene. The slides were mounted with Ralmount and examined by light microscopy.

2.2.22. PCR analysis of murine CD33 splice variants

Primers M1, M4 and M5, encompassing the spliced region were designed for detection of alternatively spliced murine CD33 variants (see Fig. 6.8, Table 2.1). PCR was performed on cDNAs from mouse bone marrow, spleen, thymus, kidney, liver, brain and the cell lines, WEHI 3B, A4, M1, P388, EL-4 and A-20 as described in 2.2.11. As a PCR control, a 428 bp murine β-actin fragment was amplified with primers shown in Table 2.1. To detect possible genomic DNA contamination, for each sample a control without reverse transcriptase was amplified as well.
3.1. Introduction

Cell-ECM and cell-cell interactions are crucial for normal haemopoiesis, for direction and control of leucocyte traffic and migration through tissues, in the development of immune and non-immune inflammatory responses and for tumour growth and metastasis. Receptors specific for these processes have been described and rapid progress in this field in the last decade is in part due to the identification and cloning of the receptor/ligand pairs involved. The transient expression cloning procedure of Aruffo and Seed (1987; Seed, 1987) in particular, has been central to this and has found extensive application for the isolation of functional cDNA clones, encoding various cell surface proteins. This technique is based on the transient expression of cDNA libraries in mammalian cells and rescue of specific cDNA clones by antibody capture and panning. One of the main advantages of this system is that the cloned cDNAs are in an efficient expression vector, πH3M or pCDM8, and can be used immediately for functional experiments such as cell adhesion assays. Thus the cloned cDNAs could be expressed transiently in Cos cells or stably in other cell types such as L- or CHO- cells and their binding to various cell lines or ECM components examined. Using this approach the adhesive function of CD44 (St John et al., 1990), ICAM-1 (Berendt et al., 1992), CD22 (Stamenkovic et al., 1991) or ELAM-1 (Lowe et al., 1990) have been established.

Purified adhesion molecules have a valuable role for defining adhesive interactions. They can be incorporated in liposomes to study
their interactions with cell surfaces (Sadoul et al., 1983; 1990), coupled to microspheres to monitor aggregation (Kuhn et al., 1991), or coated onto plastic or glass surfaces as substrates for cultured cells to measure cell adhesion and spreading (Lochter et al., 1991, Frei et al., 1992). However this requires the purification of surface molecules in a functionally competent state. An extension of this approach has been the production of soluble forms of the adhesion molecules by gene fusion between the predicted extracellular domain of the cell surface protein and the Fc portion of the IgG1 molecule. Thus an easy to detect probe for binding to various ECM substrates, tissue sections and cell lines, is created. Examples of this strategy include the binding of CD44-Fc protein to hyaluronate in tissue sections and high endothelial cells in primary cultures (Aruffo et al., 1990), CD62-Fc binding to granulocyte and tumour cell sulphatides as well as to sulphatides absorbed on plastic (Aruffo et al., 1991) and CD22-Fc adherence to CD45RO antigen on MOLT 4 cells (Stamenkovic et al., 1990).

To test the proposed role in adhesion of CD34, CD33 and BGPC molecules we analysed their binding to purified ECM substrates or various cell lines. Since the CD33 and CD34 antigens are expressed on early haemopoietic progenitors, they are also candidates for mediating adhesion with bone marrow microenvironmental ECM and cells and therefore we investigated their interactions with stromal layers, formed by LTBMC. Because the CD34, CD33 and BGPC molecules have been cloned using the expression cloning system of Aruffo and Seed, as an experimental design for studying these adhesive interactions we applied their strategy of expressing CD34, CD33 and BGPC molecules stably in CHO and transiently in Cos cells. Furthermore we have produced soluble forms of CD33-Fc and CD34-Fc antigens. In this chapter the binding of CHO and Cos cell transfectants, expressing CD33, CD34 and BGPC
molecules and of CD33-Fc and CD34-Fc soluble proteins to stromal layers, ECM substrates and various cells is analysed.

3.2. Results

3.2.1. Production and characterization of CHO cell transfectants and of soluble fusion proteins

(a) CHO cells, expressing CD33, CD34 and BGPC proteins

The relevant linearised cDNAs for CD34, CD33 and BGPC were co-transfected with the pSV2neo plasmid, containing the neomycin resistance gene, into CHO cells and resistant cells were selected by the addition of G0418. The transfected CHO cells were stained with the respective CD34, CD33 and CD66 antibodies and sorted by flow cytometry to maximize the expression of each surface antigen. The sorting was repeated, taking the highest 10-30% of fluorescence until a purity of 99% positive cells was obtained. Three or four sorts at intervals of one week were sufficient to achieve this. Figure 3.1 shows the results from the sequential sorts for each of the CHO cell transfectants. After establishing stable CHO cell lines expressing the desired molecules, the cell populations were regularly FACScan analysed in order to monitor expression and to check for overgrowth of non-expressing variants (Fig. 3.2).

The CHO cell transfectants were further characterised by western blotting. The cells were lysed and the lysate electrophoresed, transferred to nitrocellulose membrane and analysed with the appropriate antibodies. The CD66 antibody, YTH 71.3.2, recognizes a 120 kD molecule in CHO-BGPC cells (lane 4 in Fig. 3.3), while the CD34 antibody, ICH3, binds to a protein of MW 116 kD in CHO-CD34 transfectants (lane 6 in Fig. 3.3), which is in agreement with the MW determined for CD34 antigen (Civin et al., 1989). The CD33 antibody, WM53 did not immunoblot with CHO-
Figure 3.1: Flow-sorting of CHO cell transfectants
CHO-CD33, CHO-CD34 and CHO-BGPc cell transfectants were labelled with saturating amounts of the CD33 (WM53), CD34 (ICH3) and CD66 (YPC2/12.1 and CE6/2D3.1) antibodies, followed by FITC-labelled rabbit F(ab)2 anti-mouse or anti-rat Ig. The sort windows (indicated by dash lines) for the sequential sorts of each of the transfectants are shown. Only the highest percent of fluorescent cells were selected and grown. The dotted profiles in the bottom panels represent the negative control of cells stained only with the second layer antibody.
Figure 3.2: FACScan analysis of CHO-CD33, CHO-CD34 and CHO-BGpc cell transfectants after three months in culture
CHO-CD33 (A), CHO-CD34 (B) and CHO-BGpc (C) cell transfectants were labelled with saturating amounts of CD33 (WM53), CD34 (ICH3) and CD66 (YPC2/12.1 and CE6/2D3.1) antibodies, followed by FITC-labelled rabbit F(ab)_2 anti-mouse or anti-rat Ig. Negative controls were cells stained with the second layer antibody only.
Figure 3.3: Western blot of CHO cell transfectants
The CHO-BGPc (tracks 3 and 4) and CHO-CD34 (tracks 5 and 6) cell transfectants and CHO cells (1 and 2) were lysed, precleared with Protein A Sepharose, separated in 7.5% SDS-PAGE (non-reducing conditions) and immunoblotted with CD66 (20 µg/ml of YTH71.3.2; tracks 1 and 4), with CD34 (20 µg/ml of ICH3; tracks 2 and 6) or without the first layer of antibody (BSA; tracks 3 and 5), followed by HRP-labelled Protein A and ECL detection system. Molecular weight markers were myosin, 250 kD; β-galactosidase, 116 kD; phosphorylase b, 97 kD; bovine serum albumin, 66 kD; ovalbumin, 45 kD; carbonic anhydrase, 29 kD.
CD33 cell lysate nor with the lysate of other CD33 positive cell lines, such as the monocytic cell line, U937, indicating that the WM53 antibody does not recognise the denatured form of CD33 antigen.

These results suggest that the CHO transfectants stably express the CD34 and CD33 defined antigens and BGPc protein.

(b) production of CD34-Fc and CD33-Fc soluble fusion proteins

The soluble fusion proteins were created by genetic fusion of the predicted extracellular domains of CD34 and CD33 antigens to the Fc portion of IgG1 molecule, encoding the hinge, CH1 and CH2 regions (Fig. 3.4A). The CD34-Fc and CD33-Fc constructs were transfected into Cos 7 cells and the resulting soluble proteins were secreted and accumulated in Cos cell supernatant to a concentration of 0.5 µg/ml on the seventh day post transfection. 35S-labelled CD33-Fc or CD34-Fc proteins were also produced by metabolic labelling of Cos cells with 35S-Methionine 12 hours post transfection. The fusion proteins were purified by Protein A-Sepharose chromatography and were analysed by western blotting or by autoradiography when produced as radiolabelled molecules. Under non-reducing conditions 35S-labelled CD33-Fc shows a MW of 60 kD (lanes 1 and 2 in Fig. 3.4B), while 35S-labelled CD34-Fc migrates at 116 kD (lanes 3 and 4 in Fig. 3.4B). CD34-Fc soluble protein was recognised by the CD34 antibody, ICH3, in a western blot analysis. The ICH3 antibody binds to 116 kD protein (lane C in Fig. 3.5), with the same MW as shown by autoradiography for 35S-labelled CD34-Fc (lane D in Fig. 3.5).

Since the CD33 specific antibody, WM53, does not immunoblot, as was shown from the analysis of CHO-CD33 transfectants (see above), CD33-Fc protein was tested in an alternative way. The WM53 antibody was immobilised on the nitrocellulose membrane and was probed with CD33-Fc protein, followed by HRP-labelled sheep F(ab')2 anti-human Ig.
Figure 3.4: Analysis of CD34-Fc and CD33-Fc fusion proteins

(A) Schematic representation of immunoglobulin fusion gene.
(B) Cos cells were transfected with CD33-Fc and CD34-Fc constructs and
36 hours post transfection the cells were labelled with $^{35}$S-Methionine.
The secreted in the supernatant fusion proteins were purified with Protein
A Sepharose and eluted in 200 µl 0.1 M glycine-HCl buffer, pH 3.0. The
autoradiography of $^{35}$S-labelled CD33-Fc (1µl and 10 µl of the eluate,
tracks 1 and 2 respectively) and CD34-Fc (1 µl and 10 µl of the eluate,
tracks 3 and 4) fusion proteins separated in 7.5% SDS-PAGE (non-
reducing conditions) is shown. Molecular weight markers were myosin,
250 kD; β-galactosidase, 116 kD; phosphohorylase b, 97 kD; bovine serum
albumin, 66 kD; ovalbumin, 45 kD; carbonic anhydrase, 29 kD.
Figure 3.5: Antibody reactivity of CD34-Fc fusion protein
CD34-Fc protein (lanes 1, 2, 3) was separated by 7.5% SDS-PAGE and immunoblotted with CD34 (30 μg/ml of ICH3; lane C), VCAM-1 (30 μg/ml; lane B) and without first layer of antibody (BSA; lane A). The bound antibodies were detected by HRP-conjugated rabbit F(ab')2 anti-mouse Ig and ECL system. Lane D is an autoradiograph of 35S-labelled CD34-Fc, separated in 7.5% SDS-PAGE. Molecular weight markers were myosin, 250 kD; β-galactosidase, 116 kD; phosphorylase b, 97 kD; bovine serum albumin, 66 kD; ovalbumin, 45 kD; carbonic anhydrase, 29 kD.

Figure 3.6: Dot blot analyses of CD34-Fc and CD33-Fc proteins.
The immobilised CD33 (20 μl of WM53 suprenatant; A, B, C) and CD34 (20 μl of 20 μg/ml of ICH3; D, E, F) antibodies were incubated with 30 μg/ml of purified CD33-Fc (B, E), CD34-Fc (C, F) proteins and BSA (A, D) for 1 hour at room temperature. The bound proteins were visualised with HRP-conjugated sheep F(ab')2 anti-human Ig and ECL detection system.
Specific recognition of WM53 antibody was detected with CD33-Fc protein (Fig. 3.6B), but not with CD34-Fc (Fig. 3.6C). No signal was observed between CD33-Fc protein and the immobilised CD34 antibody (Fig 3.6C).

These data show that chimaeric CD34-Fc and CD33-Fc proteins were efficiently secreted and were immunologically reactive with the appropriate antibodies. At the time of this study the BGP-Fc construct was not available and the binding of BGP-Fc protein was not tested in the following adhesion assays.

3.2.2. Binding to stromal layers

As a starting point for studying the possible role in adhesion of CD33 and CD34 antigens and BGPc protein, we decided to analyse their interactions within the bone marrow microenvironment and to do this we established an adhesion system using the stromal layers, formed by LTBMC. These layers comprise of stromal cells and their attendant ECM. In the first series of experiments ⁵¹Cr-labelled CHO transfectants, expressing CD34, CD33 and BGPc molecules, were added to the confluent stromal layers for 70 minutes at 37°C. No difference in binding between the transfectants and the CHO cells was observed (Fig. 3.7), suggesting that CHO cells, due to their high background attachment (nearly 60% binding) may not be suitable for studying adhesion interactions to stromal layers. No changes in the adhesion to stromal layers were detected when CHO cells were preincubated with CD33, CD34 or CD66 antibodies (data not shown).

Then the binding of CD33-Fc and CD34-Fc chimaeric proteins to the stromal layers was tested. The assay was performed by adding ³⁵S-labelled CD34-Fc and CD33-Fc fusion proteins for 2 hours at 37°C to the stromal layers. Negligible binding of less than 0.1% with both ³⁵S-labelled CD34-Fc and CD33-Fc fusion proteins to the stromal layers was observed,
Figure 3.7: Binding of CHO cells to stromal layers

$^{51}$Cr labelled CHO transfectants (0.3 ml/300 000 cells) were incubated with stromal layers in 24 well TC plates for 70 min at 37°C. The percentage of total counts recovered in the adherent layer is shown. Data are presented as the mean ± SD from 4 replicate wells. The results are representative of three such experiments.
which was comparable to their binding in the negative controls wells, containing bovine serum albumin (data not shown). These results suggest that CD33-Fc and CD34-Fc do not mediate specific interactions with stromal layers under the conditions of these experiments.

3.2.3. Binding to ECM components and bone marrow ECM extracts

(a) binding to ECM substrates

To provide more direct evidence for the possible interactions of CD33 and CD34 antigens and BGPc protein with ECM, we next studied the adherence to purified ECM components. Attachment was examined at 37°C, and the ECM substrates used in the adhesion assay were fibronectin, fibrinogen, laminin, collagens I, III, IV, IX, X and hyaluronic acid. Preliminary adhesion assays with CHO transfectants to the ECM substrates showed, as expected from the above results with stromal layers, high background binding and no differences in adherence between the CHO, CHO-CD33, CHO-CD34 and CHO-BGPc cells to the ECM substrates (data not shown).

We next examined the binding of CD34-Fc and CD33-Fc fusion proteins. No specific adherence was observed when the binding of ^{35}S-labelled CD33-Fc (Fig. 3.8) or CD34-Fc (data not shown) fusion proteins to the ECM substrates was measured. Similar results were obtained when the binding of CD34-Fc and CD33-Fc fusion proteins to ECM components was detected by ELISA using HRP-conjugated Protein A, followed by ABTS substrate (data not shown) or when the binding was monitored by ^{125}I-labelled Protein A. The last is shown in Fig. 3.9, where the binding of CD34-Fc fusion protein does not exceed that of ^{125}I-labelled Protein A alone to any of the ECM components (data for CD33-Fc fusion protein and ^{125}I-labelled Protein A not shown). These results suggest that CD34-Fc
Figure 3.8: Binding of \(^{35}\)S-labelled CD33-Fc fusion protein to ECM substrates
100 µl of \(^{35}\)S-labelled CD33-Fc fusion proteins (2,500 cpm) was added to the wells precoated with the indicated ECM substrates (Fn-fibronectin; Fg-fibrinogen; Ln-laminin; Col I- collagens I, III, IV, IX and X, HA-hyaluronic acid and BSA as a negative control). Binding was performed for 2 hours at 37°C. The counts recovered in the adherent layer in comparison to the total counts added (total) are shown.

Figure 3.9: Binding of CD34-Fc fusion protein to ECM components
200 µg of CD34-Fc protein was preincubated with 1.5 µCi of \(^{125}\)I-labelled Protein A for 1 hour at room temperature and 100 µl of this mixture was added to the wells precoated with the indicated ECM substrates (Fn-fibronectin; Fg-fibrinogen; Ln-laminin; Col I- collagens I, III, IV, IX, X, HA-hyaluronic acid and BSA as a negative control). As a negative control, the binding of Protein A only was monitored. Binding was performed for 2 hours at 37°C. The percentage of the total radioactivity recovered in the adherent layer is shown.
and CD33-Fc fusion proteins do not adhere specifically to fibronectin, fibrinogen, laminin, collagens I, III, IV, IX, X and hyaluronic acid.

(b) binding to bone marrow ECM extract

Campbell et al. (1987) have shown that bone marrow ECM, isolated by high salt precipitation, followed by guanidine extraction, preferentially promotes the adhesion of granulocytic bone marrow precursors. To identify the putative lineage-specific adhesion protein they have separated this ECM extract by SDS-PAGE, transferred the proteins to nitrocellulose and probed using $^{51}$Cr-labelled unfractionated murine marrow cells. Using this assay they have attributed the strong granulocytic cytoadhesion to a 60 kD haemonectin. In order to test whether CD34 and CD33 antigens interact with such bone marrow ECM extract or with haemonectin we prepared bone marrow ECM following Campbell's procedure (Campbell et al., 1985), and probed the ECM blots with $^{51}$Cr-labelled CHO-CD33, CHO-CD34, CHO-BGPc cell transfectants and KG1 cells, which as shown by FACscan analysis express CD33 and CD34 antigens (data not shown). No specific binding was observed with any of the CHO cell transfectants or KG1 cells to the ECM blots. Since at that time we did not have anti-haemonectin sera, we used as a positive control the binding of $^{51}$Cr-labelled murine bone marrow cells, as described in the original identification of haemonectin (Campbell et al., 1987). However no specific binding of $^{51}$Cr-labelled murine bone marrow cells to the ECM blots was observed.

The ECM blots were then probed with CD33-Fc and CD34-Fc fusion proteins, followed by HRP-conjugated sheep F(ab')$_2$ anti-human Ig. Specific binding to a 110 kD protein and a weaker band at 100 kD was detected with CD34-Fc fusion protein (Fig. 3.10). This result was obtained on three separate occasions using one batch of CD34-Fc proteins,
Figure 3.10: Binding of CD34-Fc and CD33-Fc fusion proteins to rabbit bone marrow ECM extract
The rabbit bone marrow extract was separated in 7.5% SDS-PAGE (non-reducing conditions) and transferred to nitrocellulose membrane. The binding of CD34-Fc (20 µg/ml; lane C), CD33-Fc (20 µg/ml, lane B) fusion proteins and PBS, 0.1% (w/v) BSA (lane A) was performed for 1 hour at room temperature, followed by HRP-labelled sheep F(ab')2 anti-human Ig and ECL detection system.
generously provided by Dr. David Simmons, which preparations unfortunately we have not characterised by western blotting. However no specific binding was observed with the CD34-Fc fusion proteins produced in our laboratory using the Fc-constructs, provided also by Dr. D. Simmons.

3.2.4. Cell-cell interactions

To determine whether CD34 and CD33 defined antigens and BGpC protein interact with other cell surface receptors we performed homophilic and heterophilic adhesion assays using the stable CHO and transient Cos cell transfectants.

(a) homophilic binding studies

To test whether CD33, CD34 and BGpC proteins mediate a homophilic adhesion, an aggregation assay was performed, measuring the ability of single cells to form aggregates in suspension. CHO cells were removed from plastic surfaces by 3 minutes incubation with 0.12% Trypsin in PBS, containing 15 mM sodium citrate (Benchimol et al., 1989). This appears to be a crucial step in performing the assay, since if the cells are dislodged only with EDTA, spontaneous self aggregation occurs. Trypsin probably digests ECM components, present in the pericellular matrix of CHO cells which are involved in adhesion. FACS analysis shows that the trypsin treatment protocol used does not affect the surface expression of transfected CD34, CD33 and BGpC molecules (Fig. 3.11). Using this assay it was shown that CHO-BGpC cells aggregate in DMEM with Ca++ and Mg++ (Fig. 3.12). No aggregation was observed with CHO, CHO-CD33 and CHO-CD34 cells or between CHO-CD33 and CHO-CD34 transfectants. The clumping observed in suspensions of control cells was confined to clusters of 2-3 cells (Fig.3.12C), whereas CHO-BGpC
Figure 3.11: Effect of EDTA and trypsin on the expressed antigens in CHO cell transfectants

Overlapping fluorescence histograms revealed that detachment of CHO transfectants with 2 mM EDTA in PBS or by mild trypsinization (0.12% (w/v) trypsin in PBS, 15 mM sodium citrate for 3 minutes at 37°C) did not affect the positive staining with the respective antibodies. No staining of non-transfected CHO cells was observed with either CD33, CD34 or CD66 antibodies. Negative controls were cells stained with the second layer reagent only.
Figure 3.12: Aggregation of CHO-BGPc cell transfectants

CHO-CD34 (C) and CHO-BGPc (B) transfectants (10^6 cells/ml) were incubated for 1 hour at 37°C with agitation in DMEM medium. Single cells remaining were counted and subtracted from the total number of cells in order to provide an estimate of the number of aggregated cells. The percentage of cells forming aggregates were determined from 6 replicate samples (A). Values are means ± SD.
transfectants developed massive aggregates of 50 and more cells (Fig. 3.12B). CD66 (YTH71.3.2, CE6/2D3.1 and YPC2/12.1) antibodies did not affect the aggregation. The effect of Fab fragments of CD66 MoAbs or of CD66 polyclonal antibodies have not been tested.

To test whether the clumping is the result of an exclusive homophilic interaction between BGPc molecules on different cells we examined whether non-transfected CHO cells were included in aggregates. Fluorescein diacetate-labelled CHO-BGPc transfectants were mixed with unlabelled parental CHO cells in a ratio of 1:7 and allowed to aggregate. This ratio limited the chance that non-transfected cells would be non-specifically entrapped in aggregates. The majority of aggregates formed under these conditions contained only labelled cells whereas about 90% of the single cells were unlabelled (data not shown).

Confluent CHO-BGPc cells were stained with CD66 MoAb, YTH 71.3.2 and examined by confocal microscopy. The cell surface distribution of BGPc protein is localised at the sites of cell-cell contacts (data not shown). This was also confirmed by electron microscopy using an immunogold labelling procedure (Watt et al., 1994b), thus further supporting the homophilic adhesion interactions mediated by BGPc molecule.

(b) binding of $^{125}$I-labelled CEA to CHO cell transfectants

BGPc is a member of the CEA family and since some members of this family, like CEA, NCA and BGP, have been shown to form heterophilic associations with each other (Oikawa et al., 1989; Zhou et al., 1990; Rojas et al., 1990), we tested whether BGPc interacts with CEA. Incubation of CHO transfectants with $^{125}$I-labelled CEA showed increased binding to CHO-BGPc, but not CHO, CHO-CD33 and CHO-CD34 cells (Fig. 3.13). Although dose response experiments or inhibition by cold CEA
Figure 3.13: Binding of $^{125}$I-labelled CEA to CHO cell transfectants

20 000 cpm of $^{125}$I-labelled CEA were incubated with the CHO monolayers in 24 well plates for 1 hour at 37°C. The percentage of total counts recovered in the adherent layer is shown. Data are presented as the mean ±SD from 4 replicate wells.
were not performed, these preliminary results suggest that BGPC may interact with CEA.

(c) **binding of cells to CHO cell transfectants**

To determine whether CD33, CD34 and BGPC proteins react with other cell surface receptors, CHO transfectants were tested for adhesion to a variety of cell lines representing T (CEM, MOLT4), B (Raji, Daudi), myeloid (HL-60, U937), lymphomyeloid (KG1) and erythroid (K562) cells. Confluent monolayers of CHO cells and the CHO transfectants were overlayered with \( ^{51} \text{Cr} \)-labelled cells in DMEM for 70 minutes at 37\(^\circ\)C. The non adherent cells were removed by gentle washing and the remaining cells were lysed with 2 M NaOH and counted. No significant differences in the binding of Raji, Daudi, MOLT-4, CEM, KG1, K562 and U937 cell lines to the CHO transfectants were observed, suggesting the lack of specific adherence to the CD33, CD34 and BGPC proteins (Fig. 3.14A, data for Daudi cells not shown). The only exception was the myeloid cell line HL-60, which showed an increase in binding to CHO-BGPC transfectants (Fig. 3.14B). As MoAbs specific to the BGPC splice variant are not available, the presence of BGPC-related molecules on HL-60 cells was determined with CD66 MoAb, YTH71.3.2. Only HL-60 of all cells tested reacted with YTH 71.3.2. The binding of HL-60 cells to CHO-BGPC transfectants was not blocked by CD66 antibody, YTH 71.3.2 (Fig. 3.14B).

The binding of peripheral blood mononuclear cells to CHO transfectants was also studied. These failed to adhere specifically to any of the CHO transfecants (data not shown). Further studies with separated populations of blood cells need to be performed.
Figure 3.14: Binding of cell lines to CHO cell transfectants
The indicated $^{51}$Cr-labelled cell lines (A) and HL-60 (B) were incubated with CHO transfectants for 70 minutes at 37°C. The percentage binding is determined from four replicates. Values are the mean ± SD. Increased binding of HL-60 cells to CHO-BGPc transfectants was observed and is not blocked by preincubation of HL-60 cells with CD66 antibody, YTH71.3.2 (B).
(d) binding of cells to Cos cell transfectants

In order to determine whether the specific interactions of the cell lines with CHO cell transfectants were masked by the presence of competitive adhesion molecules present on the CHO cells, the cell lines were tested for adhesion to Cos cells transiently expressing CD33, CD34 and BGPe proteins. Cos cells were transfected with cDNAs encoding CD33, CD34, BGPe or were mock transfected using DEAE-dextran. Transient expression of the introduced cDNA was monitored 48 and 72 hours post transfection by immunofluorescence. Optimal expression was observed at 72 hours after transfection. On average from 20% to 40% of the transfected Cos cells expressed the specific molecules.

Binding of cells to Cos transfectants was assessed morphologically because of the variable percentage of Cos cells with expressed transfected genes and because there were difficulties in washing Cos monolayers after radiolabelled cells were allowed to adhere. Many Cos cells detached when washing was attempted. For the morphological assay, cells were allowed to adhere to subconfluent monolayers of Cos transfectants for 70 minutes. Non-specific adhesion of peripheral blood cells or cell lines was reduced by performing the assays at 4°C and by the presence of 500 U/ml of heparin as previously described (Stamenkovic et al., 1991). After gentle washing the monolayers were fixed in 4% formaldehyde and scored for the presence of rosettes. Characteristic rosettes of adherent cells are shown by the binding of HL-60 cells to Cos-BGPe transfectants (Fig.3.15). Rosettes were also observed with the positive control for the adhesion assays, representing Cos cells transfected with VCAM-1, since all of the cell lines tested express VLA-4, one of the ligands for VCAM-1. No binding was detected for Raji, Daudi, MOLT-4, CEM, KG1, K562, U937 or peripheral blood cells to Cos-CD34, Cos-CD33 or Cos-BGPe cell transfectants. No
Figure 3.15: HL-60 adhesion to Cos-BGPc cell transfectants

HL-60 cells were incubated for 70 minutes at 4°C with Cos-CD34 (A) and Cos-BGPc (B) cell transfectants. After fixing with 4% formaldehyde, HL-60 rosettes are observed with Cos-BGPc (B) cells under transmitted illumination.

However, it is noted that the antibody used for CD34 and CD94, myeloid lineage markers, does not recognize multiple independent ligand-receptor interactions, thus making difficult to analyze the contribution of a single expressed protein. The sensitivity of neutralizing antibodies against non-adhesive molecules on the surface of the fusion protein further supports this view. It is established that CD34 is not only essential for hematopoietic stem cells but also for mature hematopoietic cells. In this respect the recognition of 110 kD and 100 kD proteins by CD34-Fc fusion protein in the bone marrow ECM extracts is interesting. Further
specific binding was detected when the adherence assay was performed at 4°C or 37°C.

3.3. Discussion

In this chapter the binding of lymphomyeloid antigen CD34, myeloid antigen CD33 and the mature myeloid molecule BGPc to stromal layers, ECM components and various cells is described. In the bone marrow microenvironment defining these interactions is particularly difficult because of its heterogeneity. Various cell populations and ECM molecules are present in the stromal layers, formed by LTBMC. Experimentally, further obstacles were presented by the nonspecific binding of CHO cells to stromal layers, which is a sum of the action of multiple independent ligand-receptor interactions, thus making difficult to analyse the contribution of a single expressed protein. The sensitivity of similar studies could be increased by the use for transfection of non-adherent cells or by testing the binding of purified soluble forms of the surface molecules, as described here for CD33-Fc and CD34-Fc fusion proteins.

It is difficult to interpret negative results but from the binding studies of CD34-Fc and CD33-Fc chimaeric molecules it appears that CD33 and CD34 antigens do not play a major role in cell attachment to stromal layers or to purified ECM substrates such as fibronectin, fibrinogen, laminin, collagens I, III, IV, IX, X and hyaluronic acid. However it is still possible that the hybrid soluble forms of CD34 and CD33 antigens may have reduced or altered affinities for their cognate ligands or that CD34 and CD33 antigens could bind to some other untested ECM ligands, typical of the bone marrow microenvironment. In this respect the recognition of 110 kD and 100 kD proteins by CD34-Fc fusion protein in the bone marrow ECM extracts is interesting. Further
studies with purified CD34 proteins and ECM extracts, as well as blocking experiments with CD34 antibodies, may give a clearer answer as to whether such interaction occurs. Positive controls for the binding of an Fc-fusion protein of a receptor for a known ECM component must also be included in similar experiments. The affinity and binding of the chimaeras may be further improved if CD34-IgM or CD33-IgM fusion proteins are used.

Recently it has been shown that the CD34 antigen expressed by high endothelial venules can function as a ligand for L-selectin (Baumhueter et al., 1993). It is therefore surprising that no specific binding was observed of CHO or Cos cell transfectants expressing CD34 antigen with T cell lines or peripheral blood cells, since the majority of them are known to express L-selectin (Pigott and Power, 1993). A possible explanation may be that the transfected molecules have an altered glycosylation pattern or that another glycoform of CD34 on haemopoietic cells serves as a ligand for a bone marrow stromal lectin.

The results presented here also show that BGPC protein mediates the homophilic aggregation of CHO cell transfectants expressing BGPC. These observations confirm and extend previous studies demonstrating that BGP, like CEA and NCA, can function in vitro as a homophilic adhesion molecule (Rojas et al., 1990). BGPC is a splice variant of BGP, containing the same extracellular domain, but is spliced out in the cytoplasmic tail by 53 bp (Watt et al., 1994b). Thus it appears that the splicing in the cytoplasmic domain of BGP does not affect this initial adhesion event, although ideally these experiments should be done with both splice variants simultaneously. Furthermore the preliminary binding studies of $^{125}$I-labelled CEA to CHO-BGPC cells showed that like BGP, NCA and CEA, which could form heterophilic associations with each other, BGPC may react with CEA. More detailed experiments with cold
CEA or inhibition studies with CD66 antibodies are needed to analyse the binding of BGPC with CEA. The significance of such interactions between closely related molecules is unclear, but could reside in the strength of adhesion or in secondary functions. The following order of adhesive strength has been proposed: CEA>NCA>BGP (Rojas et al., 1990). Whether the mechanism of binding of HL-60 cells to Cos or CHO transfectants, expressing BGPC molecule is homophilic or whether it is mediated by other members of the BGP or CEA family on HL-60 cells, recognised by CD66 MoAbs, remains to be determined. Further studies are required to elucidate the significance of BGPC adhesive interactions in myeloid development.

In summary CHO-BGPC cell transfectants could aggregate in a homophilic manner as well as bind to CEA and a cell surface receptor on HL-60 cells. The CD33 and CD34 antigens do not appear to be involved in adhesion to stromal layers, formed by LTBMC; with the ECM substrates fibronectin, fibrinogen, laminin, collagen I, III, IV, IX, X, hyaluronic acid; or with other cell surface receptors expressed on Raji, Daudi, MOLT4, CEM, U937, HL-60, KG1 and peripheral blood mononuclear cells. Alternative ways for determining the function of CD34 and CD33 antigens need to be explored and the use of such an approach for CD33 antigen is described in the next chapter.
Chapter 4

MOLECULAR CLONING OF TWO ISOFORMS OF THE MURINE HOMOLOGUE OF CD33 MYELOID ANTIGEN

4.1. Introduction

The adhesion studies, described in the previous chapter, did not show CD33 and CD34 antigens to mediate ECM or cell-cell adhesion, under the conditions of binding systems used. As an alternative approach to help elucidate the function of CD33 antigen we decided to clone the murine homologue of CD33 molecule. Having the murine counterpart would make possible a variety of new approaches to identifying the role of CD33 antigen.

The most popular approach to cDNA cloning involves screening a cDNA library. The screening of recombinant clones can be carried out with two types of reagents: antibodies and nucleic acid probes. Antibodies against the mouse CD33 antigen are not available. Furthermore the existing anti-human CD33 antibodies were raised in mouse, suggesting that even if it is assumed that human CD33 is highly homologous to its mouse counterpart, these antibodies will not be reactive with the murine CD33 molecule. On the other hand nucleic acid probes may be preferable for library screening, because they can be used under different stringencies, minimising the chances of undesirable cross-hybridisation. They allow extremely large number of clones to be analysed simultaneously and do not require that the cDNA clones be full-length.

When attempting to clone genes thought likely to be closely related to a known gene sequence, PCR can simplify the cloning procedure. PCR can be used to generate a specific probe for an uncloned gene with primers derived from a related gene. By synthesising a perfectly matched cDNA
probe, stringent hybridisation conditions can be used for subsequent screening of a cDNA library, thus eliminating false positive signals. To generate such authentic probes "degenerate" pools of oligonucleotides can be used. The "degenerate" pool is a mixture of oligonucleotide primers containing all possible sequences that can code for a given tract of amino acids. Examples of this approach were the detection of the novel integrin β6 subunit using degenerate primers based on conserved regions of the known integrin β subunits (Sheppard et al., 1990) or the identification of a new multigene family for odorant receptors with degenerate primers against the conserved regions of the superfamily of G proteins (Buck and Axel, 1991). Alternatively new members of a known gene family can be detected using PCR with specific (not degenerate) oligonucleotide primers representing consensus sequences within the family. Following the latter strategy the murine homologue of CD22 antigen has been cloned using specific human CD22 oligonucleotide primers encompassing the conserved regions between human CD22 and MAG sequences (Torres et al., 1992).

In this chapter the molecular cloning of the murine homologue of the human CD33 antigen is described. This study was initially designed to explore a PCR based approaches to generating a specific mouse CD33 cDNA fragment, to be used as a probe to screen a mouse cDNA library. Subsequently, two types of cDNA clones were isolated from a mouse bone marrow cDNA library. Sequence analysis showed that these encode two murine CD33 isoforms with distinct cytoplasmic domains.
4.2. Results

4.2.1. Cloning approach

(a) **PCR cloning**

Initial attempts to amplify a specific murine CD33 fragment by PCR using degenerate primers for the human CD33 sequence were unsuccessful. Two sense degenerate oligonucleotides, D1 and D2, matching sequences within the beginning and the end of the first Ig-like domain of human CD33 cDNA and an anti-sense, degenerate oligonucleotide, D3, corresponding to a sequence in the cytoplasmic region were designed (Table 2.1, Fig 4.1A). The sequences of these oligonucleotides were deduced from the respective amino acid sequences, using the degeneracy of the genetic code. D1, D2 and D3 were used in combinations - D1/D3 and D2/D3. PCR amplification was performed on cDNAs, from the mouse myelomonocytic cell line, WEHI 3B and from mouse bone marrow, since human CD33 is known to be highly expressed in this tissue. No detectable bands were observed under the conditions of the experiment (data not shown).

As an alternative approach specific oligonucleotide primers corresponding to different regions of the human CD33 sequence - the first and second Ig domains, transmembrane (TM), cytoplasmic domains and 3' untranslated region (UTR) were synthesized. These are shown in Table 2.1 and Fig. 4.1A and were used in all possible combinations for PCR amplification of RNA from WEHI 3B cells and mouse bone marrow. No bands were observed with any of the primer combinations except with the pair H3/H6, where a single band of the expected size of 358 bp was detected in both WEHI 3B cells and mouse bone marrow (Fig.4.1B). This fragment corresponds to most of the second Ig domain of the human CD33 sequence. As a positive control for the amplification the human lymphomyeloid CD33 positive cell line, KG1, was used and all bands of the
Figure 4.1: RT-PCR analysis of mouse bone marrow and WEHI 3B cells

(A) Schematic representation of hCD33 protein and cDNA. The location of the degenerate (D1-D3) and specific (H1-H10) human CD33 primers are indicated by arrows. The position of H3/H6-PCR amplified fragment in relation to the final hCD33 cDNA is shown below. (B) RT-PCR analysis with primers H2/H8, H3/H6 and H4/H8 on cDNA from the indicated cell lines and tissues. Two different amounts (1 µl and 5 µl) of the respective cDNAs, were used in the amplifications with H3/H6 and H4/H8 primers. Hae III DNA fragments of φX174 was used as a size marker. The presence of RNA in the samples was confirmed by amplification of a 174 bp GAPDH fragment. No bands are detected in the absence of DNA (negative control).
expected sizes were amplified there. The presence of RNA in the samples was confirmed by amplification of a 174 bp glyceraldehyde phosphate dehydrogenase (GAPDH) cDNA fragment (Fig. 4.1B).

The 358 bp PCR product, amplified from mouse bone marrow with H3/H6 primers, was cloned into the pT7BlueT-Vector and sequenced, using vector specific primers T7 and U19 (Table 2.1). The sequence of this clone, referred as H3/H6-PCR, was also confirmed by sequencing three additional clones, generated by three separate PCR amplifications of mouse bone marrow cDNA with primers H3/H6. Sequence analysis revealed that the H3/H6-PCR sequence is identical to the corresponding human CD33 sequence, with only one difference occurring. "T" at position 495 in the human sequence is replaced by "G" in the H3/H6-PCR clone, giving a silent change (Leu). This indicated that the H3/H6-PCR clone almost certainly represented the human CD33 sequence and was a result of contamination in the PCR. The silent change between the published human CD33 sequence and that of H3/H6-PCR clone is probably due to a polymorphism.

(b) library screening

We next asked whether, nevertheless, the 358 bp PCR amplified human CD33 fragment, H3/H6-PCR, could be used for screening a mouse cDNA library. To this end a Southern blot of mouse genomic DNA was probed with the H3/H6-PCR clone (Fig. 4.2). The probe cross-hybridised under high stringency indicating some degree of conservation of the gene and thereby demonstrating that H3/H6-PCR probe could be used.

Approximately 150 000 plaques of unamplified BALB/c bone marrow cDNA library were screened with the H3/H6-PCR probe and four positive clones were isolated. These clones contained one type of insert, of 1330 bp, as determined by PCR analysis using λgt11 specific primers P7
Figure 4.2: Southern blot analysis

10 μg of mouse (A) and human (B) genomic DNA were digested with Pst I and hybridized with the H3/H6-PCR probe.
and P8, corresponding to sequences on either side of the EcoR I cloning site of the λgt11 phage (Fig. 4.3A). Two of the PCR amplified inserts were cloned into pT7BlueT-Vector and sequenced. Partial sequencing of both clones showed that they are identical. The close homology of their sequences to the human CD33 sequence (an open reading frame demonstrating 52% identity with the predicted human CD33) indicates that they probably represent the murine homologue for human CD33. To allow for any misincorporations introduced by Taq polymerase and to eliminate the necessity of sequencing at least two additional independent PCR clones, the insert of one of the initial positive phages was isolated by extraction and EcoR I digestion of phage DNA and subcloned into pBluescript SK vector (Fig. 4.3B). This clone, referred as p33-A, was used for sequencing analysis. p33-A is a 3'-clone stretching from the 3' UTR to a position, corresponding to 554 bp from the 5' end of the human CD33 cDNA sequence.

In order to obtain the remaining 5' end of the murine gene another 300,000 pfu of the primary cDNA library were rescreened with a 461 bp probe, derived from the 5' end of clone p33-A, by Bgl II digestion (probe A in Fig. 4.4A). The resulting 11 positives clones were analysed by PCR using λgt11 specific primers, P7 and P8, and mouse CD33 specific primers deduced from p33-A clone. Eight of the positives were identical to the initial p33-A clone, containing an insert of 1330 bp. The other three clones had an insert of 1213 bp. The 1213 bp insert was isolated from the λ vector, subcloned into Bluescript plasmid and sequenced. Sequence analysis showed that this clone contained an additional 201 nucleotides of the 5' end of clone p33-A. This 1213 bp clone is referred to as p33-B (Fig. 4.4A). The sequencing strategy using specific oligonucleotide primers is shown in Fig. 4.4B.
Figure 4.3: Schematic representation of PCR analysis and PCR cloning (A) and of cloning into Bluescript vector (B) of positive λgt11 bacteriophages
Figure 4.4: Mouse CD33 cDNA clones

(A) Clones p33-A and p33-B were isolated from a mouse bone marrow cDNA library. Their relative position in relation to the final murine CD33 cDNAs is shown. Probe A, is a 461 bp fragment 5' to the Bgl II site in clone p33-A and was used for library rescreening. (B) Sequencing strategy of the p33-A and p33-B clones. The arrows represent the direction and extent of sequencing with specific primers, indicated next to the arrows (the primers sequences are listed in Table 2.1). Both strands of each insert were sequenced completely.
p33-A and p33-B clones were identical in their overlapping regions, but in p33-B an additional 83 bp are inserted (Fig. 4.4A). Together cDNA clones covered a total of 1614 bp. Both clones were missing 5' sequences of mouse CD33 and these were determined by the 5'RACE system described below.

(c) isolation of the 5' end of the murine CD33 gene

The 5'RACE system was used to obtain the remaining part of the 5' end of the gene (Fig. 4.5). First strand cDNA synthesis was carried out on mouse bone marrow RNA using anti-sense M16 mouse CD33 primer. The cDNA was purified, poly(C) tailed and amplified by PCR using an anchor primer annealing to the homopolymeric tail and nested anti-sense M19 mouse CD33 primers which annealed 493 bp 3' to the M16 primer. This allowed amplification of unknown sequences between the M19 primer and the 5'end of the mRNA. The sequences of the M16 and M19 primers were deduced from clones p33-A and p33-B (see above).

Following PCR, amplification products were analysed by agarose gel electrophoresis. No bands were detected, which may be due to a very low concentration of the PCR product. Often a single PCR of 25-30 cycles is not sufficient to generate specific product bands detectable by ethidium bromide staining. An additional round of 30 cycles PCR amplification was performed on 0.05% of the initial PCR reaction, using the M19 primer and nested 5' universal amplification primer (UAP). After agarose gel fractionation of PCR products a broad diffuse smear with multiple discrete bands were detected. This would suggest the presence of numerous non specific products, dependent both on reverse transcription and dC tailing.

The whole PCR reaction mixture was purified and subcloned into pT7BlueT-Vector. 36 recombinant clones were isolated. The size of the inserts was determined by PCR analysis using pT7BlueT-Vector specific
Figure. 4.5: 5’RACE system
(A) Positions of clone p33-B, M16 and M19 primers and probe A.
(B) Schematic representation of 5’ RACE procedure.
primers T7 and U19. Based on the size of human CD33, the expected length for the clones containing the 5' end of the murine gene (and amplified with M19 primer) was predicted to be at least 600 bp or bigger. In order to identify the specific clones, PCR amplified inserts were also analysed by Southern blotting. Hybridisation was performed with probe A (the 461 bp Bgl II fragment from clone p33-A in Fig.4.4A) representing an internal probe for these putative 5' clones. Of the 36 recombinant clones screened, only four satisfied both criteria - with sizes of 600 bp or more and a strong hybridisation signal in Southern blot analysis. The sequence analyses revealed that only one of them represented the 5' end of the murine homologue of human CD33. This clone will be referred as p33-RACE (Fig.4.6). The 3' end sequence of p33-RACE is identical to the sequence of clones p33-B (and p33-A), while the 5' end sequence shows high homology to human CD33, covering the first ATG codon.

The RACE derived sequence was confirmed by sequencing three additional independent PCR clones, generated from the 5' end region. These PCR clones were amplified from mouse bone marrow RNA, using primers M19 and M1, deduced from the p33-RACE sequence. These clones will be referred to as PCRa, PCRb and PC Rc (Fig.4.6).

4.2.2. Nucleotide sequence

The relative positions of p33-RACE and the p33-A and p33-B cDNA clones is shown in Fig. 4.7. These clones were identical in their overlapping regions, but in p33-B an additional 83 bp are inserted after nucleotide 866. These clones together provided the longest cDNA (1962 nucleotides), designated m33-B, while the compiled sequence without the 83 bp insertion, designated m33-A, consisted of 1879 nucleotides (Fig. 4.7).

The composite cDNA sequences are shown in Fig. 4.8. There is a single open reading frame in each of the cDNAs starting with ATG at
Figure 4.6: p33-RACE/PCR clones
(A) Clone p33-RACE was obtained by 5'RACE system. The sequence of p33-RACE was confirmed by sequencing three additional independent PCR clones, PCRa, PCRb and PCRc. These were amplified from mouse bone marrow RNA, using M19 and M1 primers, deduced from the 5'RACE sequence and indicated by arrows. (B) Sequencing strategy of p33-RACE and PCRa clones. The arrows represent the direction and extent of sequencing with specific primers, indicated next to the arrows.
Figure 4.7: Mouse CD33 clones
The position of the clones in relation to the final combined m33-A and m33-B cDNA sequences is shown. Clones p33-A and p33-B were isolated from a mouse bone marrow cDNA library. An additional 83 bp, indicated by shaded box, is inserted in clone p33-B. Clone p33-RACE was obtained by 5'RACE system.
Figure 4.8: The nucleotide and deduced amino acid sequences of murine CD33 cDNAs

The DNA sequence was assembled from the sequences of the clones shown in Fig. 4.7. Nucleotides and amino acids are numbered on the right and refer to the longer m33-B isoform. In m33-B, 83 extra nucleotides were inserted after nucleotide 866 (286 aa). The additional segment is boxed and shaded. The differences in amino acids, resulting from the frameshift are shown under the m33-B sequence. Dots up to the inserted segment indicate that all sequences to that point are identical. The hydrophobic transmembrane domain is double underlined. The polyadenylation signal sequences are shown in lower case and underlined. The overlined T at position 1315 is replaced by C in clone p33-A.
position 9 with 866 bp of 3' UTR in m33-A and 742 bp 3'UTR in m33-B. Both m33-A and m33-B cDNAs contain the canonical AATAAA polyadenylation signal and 261 bp further downstream, the less common polyadenylation ATTAAA variant (Wilusz et al., 1989). Neither of the clones terminates in a poly(A) tail (probably truncated in vitro during cDNA synthesis) suggesting that one or both of the clones contain additional 3'UTR. The presence of two sizes of mRNA transcript, of 2.0 kb and 3.9 kb (see chapter 6.2.1a), also suggests longer 3'UTR for one or both of the clones, which may result from the use of an alternative polyadenylation signal further downstream. Additional studies are required to identify the complete 3'UTR.

No unusual features, such as capacity to form significant hairpin structures (Pandey and Marzluff, 1987) were noticed in the 3'UTR. There is one ATTTA motif at position 1271 in the 3'UTR. Similar motifs are commonly found in the 3'UTR of lymphokines and protooncogenes and are thought to be mRNA instability determinants (Shaw and Kamen, 1986; Schuler and Cole, 1988). One ATTTA motif is also found in the human CD33 3'UTR at position 1186.

5.2.3. Predicted amino acid sequence

The deduced protein sequence of the mouse CD33 has the typical features of a transmembrane protein, with an extracellular domain of 240 residues, hydrophobic transmembrane domain of 27 residues and a cytoplasmic tail of variable length. Deduced protein sequences together with domain organization and hydrophobicity profile (Kyte and Doolittle, 1982) are shown in Fig. 4.9. The 83 bp insertion in m33-B results in a shift of the reading frame generating a new cytoplasmic portion (after residue 286). The predicted mature forms of m33-A and m33-B consist of 334 and
403 residues, with cytoplasmic tails of 67 and 136 amino acids respectively.

The sequence begins with a putative hydrophobic signal peptide which is likely to be cleaved between Ala-16 and Gln-17, according to the predictive method of von Heijne (1986). The extracellular domain consists of two Ig-like domains between residues 18-130 and 152-232. The amino terminal Ig-like domain is of the V-SET family, whereas the second Ig-like domain is of the C2-SET type (Williams and Barclay, 1988). This region contains three potential N-glycosylation sites (Asn-X-Ser/Thr) and a possible glycosaminoglycan attachment site (SerGlyAlaGly) at residue 217 (Fig. 4.9A). The predicted mass for the polypeptide backbone of 37 kD and 45 kD for m33-A and m33-B respectively is expected to differ significantly from that observed due to the presence of carbohydrates and probably glycans.

Interestingly a GPI anchor can be predicted. A GPI signal sequence is found (RKS), that could be cleaved at position 235, to which the lipid could be attached (Barclay et al., 1993). Such signal sequences are usually followed by hydrophobic regions, starting 7-10 residues after the attachment point. The hydrophobic region can be 10-20 residues long and may be indistinguishable from a sequence that might form a transmembrane sequence as in the murine CD33 sequence.

Multiple putative phosphorylation targets are found in the cytoplasmic domains of m33-A and m33-B. The identical overlapping intracellular portion of both isoforms contains two cAMP and cGMP-dependent protein kinase phosphorylation sites at positions 269 and 270 (Fremisco et al., 1980) and one Protein Kinase C Serine/Threonine phosphorylation site (PKC) (Woodgett et al., 1986) is present at residue 272 (Fig. 4.9A). The cytoplasmic region of m33-A has two sequences corresponding to Casein Kinase 2 (CK 2) phosphorylation sites (Pinna,
Figure 4.9: Protein sequences of murine CD33 isoforms

(A) Amino acid sequences are deduced from m33-A and m33-B cDNAs. The single letter amino acid code is used. The 83 bp insertion in m33-B generates a new reading frame after residue 286, resulting in two mouse CD33 isoforms, m33-A and m33-B, with different cytoplasmic domains. The different cytoplasmic regions are shown below the identical sequence. The three potential sites for N-glycosylation are underlined and a possible glycosaminoglycan attachment site is circled. The hydrophobic transmembrane domain is doubled underlined. Sites for protein kinase C phosphorylation are in shaded boxes. (B) A schematic representation of the two murine CD33 isoforms, m33-A and m33-B. The coding regions are shown as boxes and the proposed structural domains are L, leader peptide; V, IgV-like domain; C, IgC2-like domain; TM, transmembrane domain; Cyt, cytoplasmic domain. The identical amino acids are indicated by similar patterns in the boxes. (C) Hydrophobicity plot of the predicted amino acid sequences of m33-A isoform by Kyte and Doolittle (1982).
1990) at positions 296 and 308. In m33-B, the 83 bp insertion, leading to the frame shift, introduces four additional PKC consensus recognition sequences at residues 319, 344, 355 and 359 (Fig. 4.9A) and changes the CK 2 sites to one site at position 301. There is one tyrosine residue in the cytoplasmic domain of m33-A (position 331) and there are two in m33-B (positions 294 and 332), not lying in "YLYL" motifs, which are known to be associated with signal transduction (Keegan and Paul, 1992).

4.2.4. Homology to human CD33 antigen

The cDNAs and their derived amino acid sequences described above are highly homologous to human CD33. The similarity is highest at the 5' end in the extracellular Ig domains, where 66% of the nucleotides and 61% of the amino acids are identical. Fig. 4.10 depicts the alignment between the two amino acid sequences. As expected for Ig-like proteins all of the extracellular cysteine residues are conserved between the murine and human proteins. The N-linked glycosylation sites are also maintained, while the possible GAG (Ser-Gly-Ala-Gly) attachment site in the murine sequence is not present in human CD33. However, another type of putative GAG attachment site, Ser-Gly, is found in the human CD33 sequence in a similar position, near to the transmembrane domain (residue 252). The homology between the human and murine CD33 sequences is interrupted just before the transmembrane domain (the dashed lines in Fig. 4.10), in the region containing the putative GPI signal sequence in murine CD33 (RKS, position 235). A putative GPI anchor signal sequence is also found in the human sequences (DGS, position 248).

Alignment of the transmembrane regions reveals a 55% amino-acid identity. The homology with the human sequence extends within the cytoplasmic region of the shorter murine m33-A isoform, although at a
Figure 4.10: Comparison of the human and murine CD33 protein sequences
The human CD33 sequence is shown above the murine m33-A sequence. The single amino acid letter code is used. Identity is represented by asterisks and conservative substitutions by dots. Gaps introduced in the sequence to maximize the alignment are designated by dashed lines. The conserved cysteine residues are underlined. The putative transmembrane domain is double underlined. The sequences were aligned with Bio Search Software.
much lower level (29% amino acid identity). Significant divergence, however, occurs between the 3' cytoplasmic domain of the human CD33 and murine m33-B isoform. The multiple phosphorylation sites identified in both murine isoforms are not conserved in the human CD33 sequence. In the human form one CK2 phosphorylation site is identified at position 358 and two serine/threonine PKC consensus phosphorylation sequence at positions 288 and 311 are predicted.

The overall amino acid identity between the human and murine mCD33-A isoform is 53% and considering the conservative replacements the similarity reaches 71%. Comparisons of murine CD33 to other proteins in the databases are discussed in the next chapter.

4.3. Discussion

In this chapter the molecular cloning of the murine homologue of the human CD33 myeloid antigen is described. The experimental design employed to isolate this was based on the assumptions, that they are likely to be homologous and that the expression of murine CD33, like that of human, should be restricted principally to haemopoietic cells of myelomonocytic lineage. The cloning strategy used included PCR analysis and screening of a mouse bone marrow cDNA library, from which two cDNA clones differing by an 83 bp insertion were isolated.

PCR amplification has found extensive application in molecular cloning and it has been a commonly used technique for the generation of specific probes for uncloned genes (Lee et al., 1988; Girgis et al., 1988; Buck and Axel, 1991; Torres et al., 1992). The poor PCR results, with both degenerate and specific human CD33 primers, described here were probably due to the inability to select the right primer combinations and/or too stringent PCR conditions. The degenerate and specific human CD33 oligonucleotide primers, used in this study, were randomly
designed, matching sequences within the beginning and the end of the human CD33 cDNA, with the expectation of amplifying larger fragments of the mouse gene. Some of the primers, such as D1, D2, H2, H3 and H6, lie in reasonably conserved regions as was shown later by the comparison between the human and murine CD33 sequences. The lack of PCR products in the case of degenerate primers, D1 and D2 (in both D1 and D2 pools there were primers showing more than 90% identity at the nucleic acid level to the murine sequence), was probably due to insufficiently good homology of the anti-sense D3 primer to the murine sequence. D3 sequence corresponds to a poorly conserved region in the cytoplasmic domain, where only one of the six amino acids is identical between the human and murine sequence (residues 313-318 of m33-A and 338-344 of human CD33, with the degenerate nucleotide primers having less than 30% homology to the murine sequence).

The absence of specific mouse PCR products with the relatively homologous H2 (60% identity on nucleic acid level to the respective murine sequence), H3 (75% identity) and H6 (77% identity) primers, probably can be attributed to too stringent PCR conditions - 57°C annealing temperature. The PCR amplification and cloning of the human CD33 fragment also confirmed the high sensitivity of PCR and the extreme precautions needed against contamination. Such precautions are essential when there are concentrated solutions of target DNA nearby (the cDNA from CD33 positive human KG1 cells).

On the other hand PCR very much simplified the cloning procedures. For example amplification of inserts in the bacteriophage λ vectors using oligonucleotide primers annealing to the flanking vector sequences or the isolation of the 5' end of the murine gene by 5'RACE. However, a limitation of the current method is its relatively high rate of misincorporation. The Taq polymerase lacks editing functions and
incorporates an incorrect nucleotide at a rate of $2 \times 10^{-4}$ nucleotides per cycle in polymerase chain reactions. This rate of misincorporation translates into an overall error frequency of 0.25% in 30 cycle amplification (Saiki et al., 1988), making the sequence of an individual DNA molecule cloned from an amplified pool unreliable. To eliminate the possibility of misincorporations, the H3/H6 and p33-RACE sequences were confirmed by sequencing three additional independent PCR clones, while for p33-A and p33-B, because of their size, conventional cloning procedures were preferred instead of PCR.

The cDNAs and derived amino acid sequences reported here show extensive similarity to human CD33 (71% overall amino acid similarity). The two cDNA clones encode two mouse CD33 isoforms, m33-A and m33-B, with distinct cytoplasmic regions of 67 and 136 aa, respectively, in contrast to a single CD33 form identified in human (Simmons and Seed, 1988). The regions with the most significant similarities between the human and mouse species include the first and second Ig-like domains. Less conserved are the cytoplasmic domains, with the most significant divergence in their C-termini occurring, between the human CD33 and murine m33-B isoform. Multiple phosphorylation sites were found in the cytoplasmic regions of murine CD33 isoforms, suggesting that they may be phosphoproteins, but whether they have a role in intracellular signalling remains to be determined. Additional insights into the biological function of murine CD33 homologue and its distinctive cytoplasmic domains may be provided by further sequence analysis, which is discussed in the next chapter.
5.1. Introduction

It is now common for the sequence of a part or the whole of a protein to be determined before its function is known. Prediction of function from the analysis of secondary or tertiary structures cannot yet be achieved, and the most profitable approach has been to find similarities with the sequences of known proteins. However it is not easy to assess the significance of similarities between sequences. For example the significance of homology in coding regions may be very different from that in untranscribed or untranslated regions. While homologies in coding sequences are likely to indicate structural similarities of proteins, those in regulatory regions may indicate similarities in tissue specific expression or turnover. In addition results that have strong statistical significance are not necessarily very helpful, since a proportion of the database for proteins is inferred from gene sequences, when the physical and chemical properties and biological role of the protein may have not been observed. Nevertheless valuable information may be gained from the presence of a groups of related alignments or by the detection of patterns, which may help define regions of functional importance. Another aspect of sequence analysis is that it may give insights into the origin and evolution of the proteins and genes, since sequences may be similar not only because they have similar function but because they have evolved from a common ancestral sequence.

In order to obtain further insights into the biological function of murine CD33 isoforms, database searches were performed with murine
CD33 sequences, using Bio Search Software and the ALIGN program. The significance of the homologies found are discussed in this chapter.

5.2 Results

Database searches were performed using Bio Search Software (Coulson et al., 1987), which searches the compiled Swiss-Prot, PIR, Brookhaven and GenBank databases. The databases were searched with the whole murine m33-A and m33-B sequences or with different domains or parts of the murine CD33 sequences. Database searches were also run under different stringency conditions with regard to the gap penalties. Usually between 50 and 100 of the alignments were collected for display. Some of the most statistically significant and biologically interesting alignments are shown in Table 5.1.

After the homology to human CD33 sequence, the next best alignment of murine CD33 is to the myelin-associated glycoprotein (MAG). A very good alignment (27% identity, 55% similarity over 240 amino acids) was observed between the first and second Ig domains of mCD33 and MAG (Fig. 5.1, Table 5.1), although the RGD (Arg-Gly-Asp) cell attachment site in MAG is not present in the murine CD33 sequence.

In its extracellular region, murine CD33 shows similarity to other members of the Ig gene superfamily such as CD22, murine transmembrane carcinoembryonic antigen 2 (TM CEA 2), murine BGP-1 precursor (or murine hepatitis virus receptor) and the T cell surface glycoprotein, CD4. As with mouse CD33, all contain the typical pattern of an amino terminal V-SET Ig domain followed by a variable number of C2-SET Ig-like domains, with an additional V-like domain in CD4. Other Ig-like proteins related to the extracytoplasmic portion of murine CD33 include transforming protein axl, poliovirus receptor, α-PDGF receptor, IL-1 receptor, CD48, bFGF receptor and CD19. Finally, in its extracellular
Table 5.1: Protein homologies to murine CD33

<table>
<thead>
<tr>
<th>Protein</th>
<th>Alignment Scores</th>
<th>% Amino Acid Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human CD33*</td>
<td>733</td>
<td>53% over 334 aa</td>
</tr>
</tbody>
</table>

**Extracellular domain + TM domain (267 aa)**
- MAG (mouse) 301 27% over 240 aa
- CD22 (human) 184 25% over 259 aa
- Transforming protein (axl) (human) 157 19% over 267 aa
- TM-CEA-2 (mouse) 147 20% over 260 aa
- CD4 (human) 142 21% over 176 aa
- Poliovirus receptor precursor (human) 137 23% over 138 aa
- BGP-1 precursor (murine hepatitis virus receptor) 137 20% over 220 aa
- α-PDGF receptor precursor (mouse) 132 20% over 267 aa
- Interleukin-1 receptor precursor (chicken) 130 23% over 236 aa
- CD48 (human) 129 23% over 207 aa
- Basic FGF receptor-1 (human) 127 22% over 267 aa
- CD19 (human) 126 15% over 262 aa
- Aggrecan (chicken) 124 16% over 200 aa
- Heparan sulfate proteoglycan core protein (human) 119 23% over 192 aa

**Cytoplasmic domain m33-A (67 aa)**
- CD33 (human) 95 27% over 67 aa
- Tropomyosin (drosophila) 77 25% over 59 aa
- TNF receptor 1 precursor (rat) 74 26% over 41 aa
- CD2 precursor (human) 73 26% over 52 aa
- α-PDGF receptor precursor (mouse)**
- BGP-1 precursor (murine hepatitis virus receptor)**
- CD4 (human)**
- Cysteine-rich FGF receptor (human)**

**Cytoplasmic domain m33-B (136 aa)**
- Hepatitis B virus surface antigen 95 21% over 125 aa
- putative RNA dependent RNA polymerase (yeast) 80 25% over 100 aa
- Homeobox protein Hox-3.1 (mouse) 76 10% over 50 aa
- cGMP-dependent protein kinase (Drosophila) 65 22% over 36 aa
- CD6 (human)**
- MAG-L (mouse)**
- CD4 (human)**
- Cysteine-rich FGF receptor (human)**

The compiled PIR, Swiss-Prot, Brookhaven and GenBank databases were searched for
similarities to the murine CD33 protein sequence using Bio Search Software. The databases were searched for distantly related proteins with less stringent treatment of gaps. The alignment scores represent the highest score between alignment of various domains of murine CD33 molecules and those within the data banks within one search.

*The overall comparison with human CD33 is with m33-A isoform.

**These comparisons are derived from separate searches with the whole m33-A or m33-B sequences and the alignment scores are not comparable to the others presented, which are derived from searches with only the respective cytoplasmic domain. Some of these comparisons are presented on Fig. 5.2.

***The comparison with MAG-L is derived using ALIGN program (see Fig. 5.2)

![Comparison between murine MAG-L and murine CD33 protein sequences](image)

**Figure 5.1: Comparison between murine MAG-L and murine CD33 protein sequences**

Alignment between the first two extracellular Ig domains of murine L-MAG and murine CD33 protein. The MAG sequence is shown above murine CD33 sequence. Identity is represented by asterisks and conservative substitutions by dots. The sequences were aligned with Bio Search Software.
domain murine CD33 is also homologous to the core proteins of aggrecan and heparan sulfate proteoglycan.

The inserted segment of 83 bp in m33-B and its derived amino acid sequence do not show statistically significant homologies to any other sequence in the database.

The cytoplasmic domains of neither murine CD33 variant bears any significant resemblance to other Ig superfamily members or other proteins currently in the databanks. Other weaker homologies however exist. The identical overlapping cytoplasmic regions of m33-A and m33-B (from residues 267 to 287) are similar to the entire cytoplasmic regions of CD4 antigen and cysteine-rich FGF receptor (Table 5.1, Fig. 5.2). The cytoplasmic domain of the short m33-A isoform is homologous to tropomyosin and to the cytoplasmic domains of human CD33, rat TNF-1 receptor, CD2, α-PDGF receptor and murine BGP-1 (Table 5.1, Fig. 5.2). In fact the homologies of m33-A with human CD33, α-PDGF receptor and BGP-1 are an extension of the similarities from their extracytoplasmic and transmembrane region.

The cytoplasmic domain of m33-B isoform demonstrates some homology with Hepatitis B Virus surface antigen, yeast DNA polymerase, murine homeobox protein Hox-3.1 and cGMP-dependent protein kinase. The cytoplasmic domain of m33-B is also homologous to the cytoplasmic domain of CD6 and by direct alignment a segment of similarity was demonstrated with the long isoform of MAG (L-MAG) (Table 5.1, Fig. 5.2).

5.3. Discussion

Examination of Table 5.1 shows that in its extracellular portion mCD33 is homologous to three major categories of proteins - cell adhesion molecules, growth factor receptors and proteoglycans, all of which contain
**HOMOLOGIES WITHIN THE IDENTICAL CYTOPLASMIC DOMAINS**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Homology</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD4 m33-A/B</td>
<td>GLCLVFILVMCRRKKTLVHM</td>
<td>280</td>
</tr>
<tr>
<td>Cysteine rich FGF-receptor m33A/B</td>
<td>GLCLVFILVMCRRKKTLVHM</td>
<td>279</td>
</tr>
</tbody>
</table>

**HOMOLOGIES WITH THE CYTOPLASMIC DOMAIN OF m33-A ISOFORM**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Homology</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>human CD33 m33-A</td>
<td>CCLLTFIVKTHRRKAAKNLAVGNTDPPTTGSSASPKHOKSKNKLHGPTETSSCGAAPTVMDEELHYA-SLNFTMQMG</td>
<td>350 (368)</td>
</tr>
<tr>
<td>Tropomyosin isoform A m33-A</td>
<td>KCLLTFIVKTHRRKAAKNLAVGNTDPPTTGSSASPKHOKSKNKLHGPTETSSCGAAPTVMDEELHYA-SLNFTMQMG</td>
<td>328</td>
</tr>
<tr>
<td>TNF receptor-1 precursor m33-A</td>
<td>CCLLTFIVKTHRRKAAKNLAVGNTDPPTTGSSASPKHOKSKNKLHGPTETSSCGAAPTVMDEELHYA-SLNFTMQMG</td>
<td>310</td>
</tr>
<tr>
<td>CD2 precursor m33-A</td>
<td>CCLLTFIVKTHRRKAAKNLAVGNTDPPTTGSSASPKHOKSKNKLHGPTETSSCGAAPTVMDEELHYA-SLNFTMQMG</td>
<td>318</td>
</tr>
<tr>
<td>α-PDGF receptor m33-A</td>
<td>CCLLTFIVKTHRRKAAKNLAVGNTDPPTTGSSASPKHOKSKNKLHGPTETSSCGAAPTVMDEELHYA-SLNFTMQMG</td>
<td>298</td>
</tr>
<tr>
<td>BGP-1 precursor m33-A</td>
<td>CCLLTFIVKTHRRKAAKNLAVGNTDPPTTGSSASPKHOKSKNKLHGPTETSSCGAAPTVMDEELHYA-SLNFTMQMG</td>
<td>482 (521)</td>
</tr>
</tbody>
</table>
Figure 5.2: Protein comparisons to the cytoplasmic domains of murine CD33 isoforms

The sequences were aligned using Bio Search Software. Identity is represented by asterisks and conservative substitutions by dots. For better orientation in terms of the relation to the whole amino acid sequence a part of the transmembrane domain is shown and underlined. The numbers of the amino acids for the complete sequences are shown in brackets. The cytoplasmic regions of MAG-L and m33-B isoform were directly aligned using the ALIGN programme. Identities are represented by solid lines (I) and conservative substitutions by dots (:).
Ig like domains (except the aggrecan) and are involved in various types of adhesion and recognition mechanisms.

The group of cell adhesion molecules homologous to murine CD33 includes MAG, CD22, CD4 and the members of the CEA family - BGPs. For all of these molecules homophilic and heterophilic binding mechanisms have been suggested. MAG is a cell surface molecule expressed by oligodendrocytes and Schwann cells (Quarles, 1983), which has been shown to mediate cell-cell adhesion (neuron-oligodendrocyte and oligodendrocyte-oligodendrocyte) (Poltorak et al., 1987; Sadoul et al., 1990) as well as cell-extracellular matrix interactions with heparin and different collagen types (Fahrig et al., 1987). Ligands for the B cell adhesion molecule CD22 have been identified on human T cells as the low molecular mass isoform of the leucocyte common antigen, CD45RO and as the CD75 antigen on B cells (Stamenkovic et al., 1991). Members of the CEA family such as BGP can mediate homophilic and/or heterophilic cell adhesion in vitro (Benchimol et al., 1989; Thompson et al., 1991; Watt et al., 1994a,b), while CD4 functions as an accessory molecule in the recognition of foreign antigens in association with MHC Class II antigens by T cells (Parnes, 1989).

It is interesting to note that some of these adhesion molecules can function also as virus receptors. The CD4 antigen has been shown to be a receptor for HIV (Dalgeish et al., 1984). In mice nine isoforms of BGP have been identified, four of which can serve as functional receptors for mouse hepatitis virus-A59 (Dveksler et al., 1991; 1993). The binding of the gp 120 glycoprotein of HIV to MAG has been demonstrated by ELISA (van den Berg et al., 1992). Interestingly murine CD33 shows homology to poliovirus receptor, which is also a member of the Ig gene superfamily (Mendelsohn et al., 1989). So far only four members of the Ig gene superfamily are known to function as a virus receptors and three of them
are related to the murine CD33 sequence. The fourth is ICAM-1, which is widely expressed in human tissues and is a major rhinovirus receptor (Greeve et al., 1989). An interesting question is whether the domain structure common to molecules of the Ig gene superfamily is a common feature of proteins that mediate the entry of certain viruses into cells, or simply reflects the fact that many cell surface molecules are Ig like. (It should be noted that known receptors for several other viruses such as influenza virus and Epstein-Barr virus are not Ig family members (Weiss et al., 1988; Fingeroth et al., 1984))

The second class of proteins, related within the extracellular Ig like domains to the murine CD33 homologue are the receptors for growth factors and cytokines - transforming protein axl, α-PDGF receptor, bFGF receptor and IL-1 receptor. These proteins (except for IL-1 receptor) are receptor tyrosine kinases which on activation induce phosphorylation of their own cytoplasmic domains and/or other proteins in the cytoplasm and initiate a complex signal transduction cascade. The axl protein (from Greek anexelekto, or uncontrolled) represents a novel subclass of receptor kinases, without as yet an identified ligand. The overexpression of axl cDNA in NIH 3T3 cells induces neoplastic transformation (Janssen et al., 1991; O'Bryan et al., 1991). The axl gene is localised to human chromosome 19q13.2 (O'Bryan et al., 1991), where another oncogene, bcl3, is also found (19q13.1-19q13.2) (McKeithan et al., 1987; Korneluk et al., 1989), which is close to the region containing the gene for CD33 (19q13.3) (Peiper et al., 1988). The α-PDGF receptor transduces the mitogenic activities of PDGF in fibroblasts, while the multple biological responses of bFGF in angiogenesis, mitogenesis and cellular differentiation are mediated by the bFGF receptor (Barclay et al., 1993). The binding of IL-1 to its receptor mediates thymocyte and T cell activation, fibroblast

There are also homologies between the Ig like domains of murine CD33 and the human CD19 and CD48 antigens. Although neither a known adhesion molecule or growth factor receptor, CD19 antigen is involved in the regulation of B cell proliferation and associates with the complement receptor 2 (CD21), TAPA-1 and the Leu13 antigen forming a signal transduction complex on the membrane of B cells (Barclay et al., 1993). CD48 consists of an Ig V-like domain and an Ig C2 set domain which is linked to the lipid bilayer by a GPI anchor (Barclay et al., 1993) and has been found to be associated with p56^ck (Stefanova et al., 1991).

In fact the homology between CD48 and murine CD33 encompasses the GPI signal sequence and the following hydrophobic region typical of consensus GPI cleavage sites. Similar GPI sequences have been predicted in the murine CD33 sequence with a hydrophobic region identical to the transmembrane domain. An example, where the same sequence forms the transmembrane-spanning region and the GPI-signal sequence is CD58 (LFA-3) antigen, in which two alternatively spliced forms, resulting in a GPI anchor or a transmembrane form of CD58 are known (Seed, 1987). A similar GPI signal sequence has been predicted in the human CD33 isoform, although only one transmembrane human CD33 form has been identified so far.

In its extracellular domain murine CD33 resembles the core proteins of two proteoglycans - aggrecan and heparan sulphate proteoglycan (HSPG). Aggrecan, named because it forms aggregates, is the major structural macromolecule of cartilage (Hay, 1991). Aggrecan has extensive substitutions of chondroitin sulphate, keratan sulphate and hyaluronic acid and the homology with murine CD33 lies in the second globular domain G2 of chicken aggrecan, which has been shown to bind
hyaluronic acid (Chandrasekaran and Tanzer, 1992). HSPG is an integral component of basement membranes and is implicated in a number of biological activities such as cell binding, low density lipoprotein metabolism, basement membrane assembly and interactions with other membrane macromolecules (Kallunki et al., 1992). HSPG contains a large core protein of 467 kD with three heparan sulphate chains (Kallunki et al., 1992). The sequence similarity of murine CD33 is with domain IV of HSPG core protein, which contains the Ig repeats and one of the consensus sequences for heparan sulphate attachment sites (Ser-Gly-X-Gly). A similar consensus (Ser-Gly-Ala-Gly) sequence is present in murine CD33, while in humans the putative GAG attachment site is a Ser-Gly doublet, followed by the charged residue Lys and precede by a neutral residue Gly. Thus the presence of GAG sites in both human and murine CD33 sequences and the observed similarities of murine CD33 homologue with proteoglycans may suggest interactions involving GAGs.

GAG has been described in two leucocyte membrane glycoproteins CD44 and syndecan (Barclay et al., 1993). In CD44 the GAG has been identified as chondroitin sulphate and CD44 has been shown to be the principal cell surface receptor for hyaluronate (Aruffo et al., 1990), an interaction involved in the binding of lymphocytes to high endothelial venules. In addition homophilic binding has been suggested and in the rat, one of the variants of CD44 confers metastatic potential (Gunthert et al., 1991). The heparan sulphate chains of syndecan interact with extracellular matrix proteins such as collagens type I, III and V, thrombospondin and fibronectin and the cytoplasmic domain appears to be involved either directly or indirectly in interactions with intracellular actin cables (Hay, 1991). Syndecan is also required for the binding of bFGF to its high affinity receptor (Yayon et al., 1991; Olwin and Repraeger, 1992). When complexed with the heparan sulphate
proteoglycans of the cell surface, basement membranes or other extracellular matrices bFGF retains its biological activities in cell proliferation and differentiation and is protected from inactivation by proteases.

The multiple phosphorylation sites predicted in the cytoplasmic domains of both isoforms (see 4.2.3), suggest that the intracytoplasmic regions of murine CD33 variants may be phosphorylated. In this respect it may be of interest that the intracytoplasmic regions of murine CD33 isoforms, reveal some homologies to the intracytoplasmic regions of known phosphoproteins or signal transduction molecules, described below.

The first residues of the intracytoplasmic region, identical in both murine CD33 variants, are homologous to the entire cytoplasmic domains of CD4 and the cysteine-rich FGF receptor. The cytoplasmic domain of CD4 is phosphorylated at Ser residues when T cells are activated by antigen or phorbol esters (Shin et al., 1990). Although short (38 aa) the cytoplasmic domain of CD4 transduces cell signals by interaction with a lymphocyte-specific tyrosine kinase p56^{lck} through a motif, which however is not present in the murine CD33 sequence (Turner et al., 1990). The cysteine rich FGF receptor is unusual since it exhibits specific binding of FGF-1 and FGF-2 in the absence of heparin or heparan sulphate and also contains an extremely short cytoplasmic domain (13 residue), suggesting association with additional intracellular proteins (Burrus et al., 1992).

The cytoplasmic domain of murine m33-A isoform shares some similarities to the intracytoplasmic regions of human CD33, TNF-1 receptor, CD2, PDGF receptor and murine BGP-1. All of these molecules are known to be involved in signal transduction or to be substrates for phosphorylation, except human CD33, the phosphorylation of which has not been studied yet. TNF receptors have been suggested to co-modulate with the Fas antigen, the cytoplasmic region of which they share some
similarities with (Barclay et al., 1993). CD2-mediated signal transduction mimics that through the TcR-CD3 complex (Beyers et al., 1989; Bierer, 1989), which is believed to involve tyrosine kinase and phospholipase C activation followed by phospho-inositide turnover and activation of second messenger pathways (Klausner and Samelson, 1991). Binding of PDGF to its receptor leads to autophosphorylation and phosphorylation and activation of various intracellular substrates including phospholipase Cγ, phosphatidylinositol 3-kinase and the Raf kinase (Ullrich and Schlessinger, 1991). However the homology between the cytoplasmic domains of m33-A and α-PDGF receptor is not within the consensus tyrosine kinase motifs. The intracellular domains of human and rat BGP isoforms are also known to be phosphorylated (Afar et al., 1992; Culic et al., 1992) and the intracellular domain of the rat homologue has been shown to be a substrate for the insulin receptor tyrosine kinase (Margoulis et al., 1990).

On the other hand the cytoplasmic portion of m33-B variant has limited similarity to the long form of murine MAG (L-MAG) and CD6 antigen. Two cytoplasmic variants of MAG have been described (short MAG, S-MAG, and large MAG, L-MAG) (Lai et al., 1987) and although both MAG isoforms have been demonstrated to be phosphorylated in a heterologous cell system (Afar et al., 1990), recently L-MAG has been shown to be associated with the Fyn tyrosine kinase (Umemori et al., 1994). The Fyn-L-MAG association requires amino-terminal domains of Fyn that include SH2 and SH3 and crosslinking of L-MAG with antibodies induces an increase in the specific activity of Fyn kinase. CD6 is another protein, which undergoes hyperphosphorylation on activation. CD6 is expressed on peripheral blood T cells and most thymocytes. (Barclay et al., 1993).
The parallels among the different systems outlined above are undoubtedly very diverse and speculative. However, these comparisons may serve to propose some testable hypotheses concerning the function of CD33 antigen. Based on the overall Ig-like structure of murine CD33 molecules and the homologies discussed it can be speculated that CD33 molecules are involved in some type of cell adhesion or recognition. The presence of a GAG attachment site in the second Ig-like domain could support the suggestion that mouse CD33 may have role in cell-cell or cell-matrix interactions or that it may bind growth factors. It would be also possible that murine CD33 molecules function as a virus receptor. The distinct cytoplasmic domains of murine CD33 isoforms are probably phosphorylated as suggested by the presence of multiple phosphorylation sites and may serve to transduce different signals from extracellular ligands to the cell interior.
6.1. Introduction

Previous biochemical and immunocytochemical studies have demonstrated the expression of human CD33 antigen to be restricted to the myelomonocytic lineage. CD33 antigen is detected on myeloid progenitors, some monocytes and on leukaemic myeloblasts from most patients with AML (see general introduction). The strictly haemopoietic expression of the CD33 molecule was further confirmed by the immunohistochemical analysis of a panel of human tissues (Peiper et al., 1989). Binding of CD33 mAb to mononuclear phagocytes was detected in sections of the spleen, lung (alveolar macrophages), testis, placenta, and liver (Kupffer cells). CD33 antibodies bound to Langerhans cells and to perivascular macrophages in frozen sections of skin. Muscle (skeletal, cardiac and smooth), peripheral nerve and brain were negative. Hepatic, pulmonary, renal and pancreatic parenchyma were also negative.

Thus, unlike other cluster groups of myeloid MoAbs, CD33 reactivity appears to be limited to cells within the haemopoietic compartment. It is therefore presumed that the function of the CD33 antigen relates specifically to haemopoietic cells. In this respect there is a particular interest in analysing the expression of the murine CD33 homologue, which if it is conserved would support the haemopoietic role of CD33 antigen. Furthermore the isolation of two murine cDNAs differing by 83 bp and encoding two murine CD33 isoforms with distinct cytoplasmic regions strongly implies that they represent products of alternative splicing. Therefore analysis of the expression pattern of
murine CD33 isoforms may provide useful information regarding their functional significance.

The aim of this chapter is to analyse the expression of murine CD33 isoforms in various haemopoietic and non-haemopoietic tissues and cell lines. Because no anti-murine CD33 antibodies are yet available, the experimental design for studying the tissue and cellular distribution of murine CD33 molecules was based on northern blot analysis, in situ hybridisation and PCR analysis. By northern analysis the amount and size of the specific mRNA molecules can be determined, while by in situ hybridisation identification of the type and localisation of the cells expressing the gene (in a complex tissue) is possible. PCR analysis was used for detection of different murine CD33 isoforms.

6.2. Results
6.2.1. Expression of murine CD33

(a) northern blot analysis

The p33-B clone (Fig. 4.7) was used to analyse the expression of the murine gene in various haemopoietic and non-haemopoietic tissues and cell lines. Northern blot analysis (Fig. 6.1) shows two species of approximately 3.9 kb and 2.0 kb. Under the conditions used, both sizes of transcript are observed in bone marrow only, with higher levels of the larger (3.9 kb) mRNA. The 3.9 kb transcript is expressed highly in the thymus and in the macrophage cell line, P388. After a longer exposure a very faint 3.9 kb band is found in the brain and liver, perhaps due to the presence of microglia and Kupffer cells, respectively (data not shown). The smaller 2.0 kb mRNA, in addition to being present in the bone marrow, is observed at a lower level in spleen and the myeloid cell line, M1, the myelomonocytic cell line, WEHI 3B, and in the multipotential progenitor cell line, A4. A band of 4.4 kb appears in all cell lines and is probably an
Figure 6.1: Northern blot analysis
20 μg of total RNA extracted from the indicated tissues (A) and cell lines (B, on the next page) was electrophoresed in formaldehyde-agarose gel, transferred to nylon and hybridised with ³²P-labeled p33-B DNA probe or mouse β-actin. The arrow indicates a band that is probably an incompletely spliced mRNA.
artefact due to non-specific binding to the 23s ribosomal RNA (Fig. 6.1B). When the filter was washed under low stringency conditions, this band was also present in the murine tissues. A minor 6-8 kb RNA species was evident in WEHI 3B, A4 and M1 cells (arrow in Fig. 6.1B). This may represent an incompletely spliced mRNA. CD33 message could not be detected by northern blot analysis in the cell lines EL4 and RMA (T cell), NS-1 and A20 (B cell), 410.4 (breast carcinoma cell line) or P815 (unidentified cell, probably microglia, data not shown). Thus murine CD33 mRNAs are detected in these regions, where resident macrophages can be identified (Russel and Gordon, 1996).

The preliminary in situ hybridisation of bone marrow cells suggests that the most prominent expression is in early myeloid cells, but not in lymphoid, erythroid, granulocytic or megakaryocytic lineages in significant amounts (data not shown). The results of in situ hybridisation

**Figure 6.1 cont.**
artefact due to nonspecific binding to the 28s ribosomal RNA (Fig. 6.1B). When the filter was washed under low stringency conditions, this band was also present in the murine tissues. A minor 6 kb RNA species was evident in WEHI 3B, A4 and M1 cells (arrow in Fig. 6.1B). This may represent an incompletely spliced mRNA. CD33 message could not be detected by northern blot analysis in the cell lines EL4 and RMA (T cell), NS-1 and A20 (B cell), 410.4 (breast carcinoma cell line) or P815 (mastocytoma cell line). These data suggest that CD33 expression is haemopoietic, limited to the myelomonocytic lineage.

The length of the smaller mRNA (2.0 kb) corresponds to the size of the cDNAs (1962 bp and 1879 bp) isolated and characterised in chapter 4. The presence of the larger CD33 transcript (3.9 kb) suggests that one or both of the murine cDNAs, contain longer 3'UTR, which may result from the use of an alternative polyadenylation signal further downstream.

(b) in situ hybridisation

The expression pattern of murine CD33, observed by northern analysis was confirmed and further extended by in situ hybridisation. In spleen the p33-B riboprobe hybridises to the red pulp (Fig. 6.2B), while in lymph node the signal is confined to the subcapsular regions (Fig. 6.3B). There is a clearly abundant mRNA for β-actin in the samples (Fig. 6.2 C and Fig. 6.3C). In brain strong expression was observed on scattered (unidentified) cells, probably microglia (data not shown). Thus murine CD33 mRNAs are detected in those regions, where resident macrophages can be identified (Russel and Gordon, 1992).

The preliminary in situ hybridisation of bone marrow cells suggests that the most prominent expression is on early myeloid cells, but not in lymphoid, erythroid, granulocytic or megakaryocytic lineages in significant amounts (data not shown). The results of in situ hybridisation
Figure 6.2: In situ hybridisation analysis of mouse spleen

*In situ* hybridisation was carried out with $^{35}$S-labelled mouse anti-sense p33-B probe. Top is bright-field (A) and below are dark field (B,C) illuminations respectively, showing white grains in reflected light revealing murine CD33 transcripts in the spleen red pulp (B). The presence of mRNA was confirmed by the abundant signals of β-actin transcript (C).
A.

B.

C.

Figure 6.3: In situ hybridisation of lymph node
Bright-field (A) and dark field (B, C) illuminations. In situ hybridisation was carried out with $^{35}$S-labelled mouse antisense p33-B probe. Murine CD33 transcripts are detected in the subcapsular regions of lymph node (B). The presence of mRNA was confirmed by the abundant hybridisation signal with $\beta$-actin riboprobe (C).
of cell lines are also consistent with those obtained by northern analysis. Strong expression is observed in A4, WEHI 3B and M1 cells (data not shown), while no murine CD33 transcripts were detected in RMA cells. Murine muscle, kidney and mouse epithelial tumour W26 were also negative.

The *in situ* analysis of thymus reveals strong signal on cells, possibly macrophages, throughout the whole thymus (data not shown) which is consistent with the high CD33 expression in thymus detected by northern hybridisation. Since no data is available on the expression of human CD33 in thymus, we also decided to test the expression of human CD33 antigen in thymus. The CD33 antibody, Anti-Leu-M9, was used for the immunohistochemical analysis of human thymus and the resulting staining is shown on Fig. 6.4. The antibody appears to stain macrophages and possibly dendritic cells. Slightly more staining is found to the corticomedullary region of thymus, where more dendritic cells are known to be localised.

Although incomplete, these preliminary data together with the northern blot analysis suggest that CD33 expression in the mouse, like that in humans is confined to the haemopoietic system and is limited to the myelomonocytic lineage.

### 6.2.2. Analysis of splice variants

The two murine CD33 cDNA clones, described in chapter 4, differed by an 83 bp insertion in the cytoplasmic region. The site of the 83 bp insertion in m33-B cDNA is flanked by consensus splice exon donor-acceptor sites. The 5' ends of exons are usually G or A and the 3' ends of exons C/AAG (Mount, 1982). This suggests alternative splicing between two exons (Fig. 6.5A). The 83 bp inserted fragment begins with A and ends with CAG, while the first nucleotide, after the insertion (the begining of
Figure 6.4: Immunohistochemical analysis of human CD33 antigen in thymus

Frozen thymus sections were stained with CD33 antibody, Anti-Leu-M9 (1:10 dilution), followed by biotinylated rabbit anti-mouse Ig, streptavidin-labelled HRP and developed with DAB. The positive signal appears as brown staining and is localised on cells concentrated at the corticomedullary junction, but is also found on cells scattered throughout the cortex and medulla.
the next putative exon, position 950) is G. The last three nucleotides of the previous exon, positions 863-865 are AAG.

To clarify whether the 83 bp insertion in m33-B is a cloning artefact or represents a differentially spliced variant, we looked for the presence of this region in cells and tissues by RT-PCR. Primers, allowing the specific detection of both spliced variants were designed. These are shown schematically in Fig. 6.5A. M4 and M1 are primers adjacent to the site of the insertion and so would give two bands differing by 83 bp, i.e. products of 344 bp for m33-A isoform and of 427 bp for m33-B. Bands of both expected sizes, corresponding to the two isoforms, are amplified in bone marrow, spleen, thymus, liver, brain and in the multipotential progenitor cell line, A4, the myelomonocytic cell line, WEHI3B and the myeloid cell line, M1 (Fig. 6.5B). The m33-A product is the most abundant form, suggesting that m33-B may be a minor transcript. Alternatively PCR may preferentially amplify the smaller m33-A form. Only one PCR product, corresponding to the m33-A isoform, is detected in the macrophage cell line, P388. No m33-A or m33-B transcripts were detected in EL-4, A-20 or the remaining cell lines which were also negative in Northern blot analysis (Fig. 6.1B and Fig. 6.5B).

The absence of m33-B isoform in the P388 cell line is also confirmed by amplification with primers M4 and M5 (Fig. 6.5C), M5 being an internal primer from the inserted segment (Fig. 6.5A). This pair of primers was designed to generate a band of 265 bp only if m33-B is expressed and such a band is detected in all the expected RT-PCR positive samples shown in Fig. 6.5B, namely, bone marrow, spleen, thymus, liver, brain and the cell lines, WEHI 3B, A4 and M1 (Fig. 6.5C). The presence of RNA in the samples is confirmed using β-actin oligonucleotide primers in RT-PCR analysis (Fig. 6.5D). The PCR analyses were performed with and without reverse transcription, in order to detect any genomic
Figure 6.5: RT-PCR analysis of murine CD33 splicing variants
(A) The 83 bp inserted fragment is boxed and shaded and the consensus exon splice donor/acceptor sites are in lower case and bold. The location of the primers, used for detecting alternatively spliced isoforms, is indicated by arrows. RT-PCR analysis with primers M4/M1 (B) and M4/M5 (C) was performed on RNA from the indicated tissues and cell lines. 20% of the amplification product were separated on 1.5% agarose gel and visualised by ethidium bromide staining. Hae III DNA fragments of ΦX174 was used as size marker. (D) The presence of RNA in the samples was confirmed by amplification of a 428 bp β-actin mRNA fragment after reverse transcription. Control experiments in the absence of reverse transcriptase produced no definitive bands (data not shown).
contamination. No definitive signals were found when reverse transcriptase was omitted (data not shown).

The consensus 3′/5′ splice exon boundaries together with RT-PCR results suggest that the two cDNA clones represent the products of alternative splicing. The pattern of expression of the murine CD33 splice variants, obtained by RT-PCR does not correlate with the northern blot distribution of the two size transcripts of 3.9 kb and 2.0 kb. Thus, both m33-A and m33-B isoforms are present in thymus as shown by PCR (Fig.6.5B) and only a 3.9 kb size transcript is detected by northern hybridisation (Fig.6.1A). Similarly in spleen, both murine isoforms, m33-A and m33-B are detected by PCR (Fig.6.5B), while only the 2.0 kb transcript is observed by northern hybridisation (Fig.6.1A). This suggests that both murine CD33 isoforms could be derived from either of the two mRNAs species of 2.0 kb and 3.9 kb. (Furthermore, an 83 bp difference in the size of the transcripts could not be detected by northern analysis). However, final proof of which mRNA transcript represents which splice variant could be obtained by reprobing the northern blot with a specific 83 bp oligonucleotide. A detailed study of the gene structure and, ultimately, analysis of the proteins should resolve this question.

6.3. Discussion

This chapter described the tissue and cell distribution of the murine homologue of the CD33 myeloid antigen. The mRNA expression pattern of the murine CD33 homologue is similar to the human, since both are strictly haemopoietic, being largely restricted to the myelomonocytic lineage. This suggests that the function of CD33 antigen in haemopoiesis may be conserved between man and mouse. Certainly more detailed in situ hybridisation analysis is required to determine the precise localisation of murine CD33 mRNAs in different tissues and their
association with various subpopulations of macrophages. These experiments should ideally include a negative control with sense probe. Immunocytochemical staining with myeloid specific antibodies of adjacent sections would allow more accurate identification of CD33 mRNA positive cells. Description of a detailed tissue distribution of the murine CD33 antigens must await the isolation of antibodies to the murine protein.

Two murine CD33 mRNA species have been identified by northern analysis and these appear to be differentially expressed in the adult tissues of mice. The larger (3.9 kb) transcript is expressed abundantly in bone marrow and thymus and is found at low levels in liver and brain, but not in spleen. On the other hand, the 2.0 kb mRNA is expressed in bone marrow and spleen only. The restricted expression of the 3.9 kb transcript in the mature macrophage cell line, P388, and of the 2.0 kb transcript in the multipotential progenitor cell line, A4, myelomonocytic cell line WEHI 3B and myeloid precursor cell line, M1 suggests that differential polyadenylation of one or both CD33 isoforms may occur in myeloid precursors and in mature macrophage subpopulations. There are other examples of 3' end heterogeneity in mRNAs that arise from the use of different polyadenylation sites. Two sizes of mRNA transcripts (of 1.5 kb and 1.8 kb) have been shown for human CD33 (Simmons and Seed, 1989) and two mRNA species, of 5.2kb and 2.9 kb, generated by the use of different polyadenylation signals, have been identified for NCAM-120 (Barbas et al., 1988). However the biological significance of this remains unclear.

The murine CD33 molecules are alternatively spliced in the cytoplasmic region as suggested by the presence of consensus exon splice donor/acceptor sequences and the RT-PCR analysis. It also appears that both murine splicing variants, m33-A and m33-B, could be derived from either of the two mRNA species, of 2.0 kb and 3.9 kb, observed by northern
hybridisation. Alternative forms of proteins, derived from the differentially processed gene, are usually generated by appropriate splice selection of inframe exons. An interesting feature of the alternatively spliced murine CD33 isoforms, described here, is splicing that results in out of frame translation. Splicing in the cytoplasmic tail of m33-B isoform by 83 bp results in a production of m33-A, in which the 3' sequence is out of frame with the m33-A sequence. In addition to viruses, such splicing has been observed for the members of the CEA family, BGP (Barnett et al., 1989; Watt et al., 1993; 1994) and CGM1 (Nagel et al., 1993) resulting in a frame shift and generating unique cytoplasmic regions.

The biological significance of different murine CD33 isoforms is unknown. Alternative splicing is the major strategy for diversification of gene products in eukaryotic cell. In many cases differential processing of primary RNA transcripts is regulated in a developmental or tissue-specific fashion. Examples for such regulated splicing are NCAM (Small et al., 1988; Santoni et al., 1989), MAG (Ishiguro et al., 1991), CD44 (Jackson et al., 1992) and CD45 (Ralph et al., 1987). Splicing of the exons in the extracellular site could modulate the ligand binding properties of different variants as is with CD44 or NCAM, while the isoforms with distinct cytoplasmic domains may transduce different signals from extracellular ligands to the cell interior such as MAG. RT-PCR analyses indicate that the mCD33-B isoform occurs in myeloid precursors but not in mature macrophages, which may suggest that these molecules are differentially expressed during myeloid development. However the biological consequences of this remain to be determined.
The aim of this study was to understand the function of the lymphomyeloid progenitor antigen CD34, the myeloid antigen CD33 and the mature myeloid molecule BGPC. In particular, attention has been focused on the possible role of these proteins in cell-extracellular matrix and cell-cell adhesive interactions. The first result chapter describes various binding assays, while the second part of the thesis (chapters 4-6) describes the molecular cloning of the murine homologue of CD33 myeloid antigen and its tissue and cell distribution.

The binding studies described here, showed that BGPC mediates homophilic aggregation of CHO cells transfected with BGPC cDNA as well as binding to CEA and perhaps another cell surface receptor on HL-60 cells. BGPC is a member of the BG family of adhesion molecules, which encompasses at least eight variants differentially spliced in the extracellular and/or intracytoplasmic domains (Barnett et al., 1989; 1993). Only BGPA and BGPB have been analysed to any extent and it has been shown that these can mediate both homophilic and heterophilic interactions, being able to bind to BGPA and BGPB as well as to CEA and NCA (Rojas et al., 1990; Oikawa et al., 1991; 1992). BGPC is a splice variant of BGPA, with the same extracellular region but is differentially spliced out in the cytoplasmic domain by 53 bp (Watt et al., 1994b). Thus the splicing in the cytoplasmic tail does not appear to modify the adhesive properties of BGPC. The biological significance of different splice variants and the relationships between the extracellular domains of BGPC molecules and long and short cytoplasmic tails are unclear. It is interesting to note that different phosphorylation sites have been found in the cytoplasmic domains of members of the CEA family, like BGP, CGM1 and CGM7.
Analysis of phosphorylation motifs shows three protein kinase C phosphorylation sites in the cytoplasmic tail of BGPa and b, which are spliced out in the BGPa splice variant (Watt et al., 1994b). Thus, while BGPa and c may both be able to mediate similar adhesion, they may transduce different intracellular signals.

These experiments and the homophilic binding described here, do not explain the function of BGP and CEA-related molecules in myeloid development. It has been suggested that BGP, together with NCA, on mature neutrophils presents sialyl Le\(^\alpha\) oligosaccharide moieties, thereby mediating neutrophil adhesion to E-selectin molecules on activated endothelium (Bevilacqua et al., 1989; Stocks et al., 1990; Kuijpers et al., 1992). Preliminary studies by Watt et al. (1994a) have failed to identify sLe\(^\alpha\) determinants attached to BGPa, although these results may be due to incomplete glycosylation of BGPa in the experimental system of CHO-BGPa cell transfectants and soluble BGPa molecules. Furthermore BGP molecules are not restricted in their expression to haemopoietic cells, but are expressed by a number of other cell types such as bile caniculi, colonic and lung epithelium and hepatocytes (Kim et al., 1992; Takahashi et al., 1993) and additional functions have been suggested for these BGP molecules. It has been shown that BGP has the ability to bind bacteria in a mannose dependent manner and appears to regulate the normal bacterial colonisation of the gut (Watt et al., 1994a). C-CAM-105, the rat equivalent of human BGP may act as an ecto-ATPase (Lin and Guidotti, 1989). Further studies indicate that the murine equivalent of BGP acts as a receptor for coronaviruses such as the mouse hepatitis virus (Dveksler et al., 1993). Thus it seems likely that the type of BGP splice variant expressed in a particular cell lineage or in a particular species may determine its specific functions. Additional studies are needed to
determine the expression patterns and in vivo functions of BGP splice variants and their role in myeloid lineage development.

As far as CD34 and CD33 antigens are concerned binding studies with CD33-Fc and CD34-Fc fusion proteins do not provide evidence for specific interactions with stromal layers, formed by LTBMC, or with fibronectin, fibrinogen, laminin, collagen I, III, IV, IX, X or hyaluronic acid. However it still could be possible that changes in the affinity of CD33 and CD34 receptors occur in the soluble fusion proteins due to conformational changes. On the other hand multiple forms of individual ECM molecule could exist. Although it has not been demonstrated what variant forms are expressed in the bone marrow, alternative splicing may confer the ability to bind to a specific receptor as has been shown in the case of stem cells, which bind only to the CS-1 fragment of fibronectin (Minguell et al., 1991). Posttranslational modifications in the ECM molecules may also affects the ways in which ECM components interact with each other or with cells (Jones et al., 1986; Dean et al., 1990; Lawler et al., 1988). It will be important to determine the stromal cell expression of ECM splicing variants and ECM remodelling by proteolysis or posttranslational modifications, which may be important for the specific adhesion and observed nonrandom distribution of progenitor cells in the bone marrow microenvironment.

The adhesion assays of various cell lines to CHO and Cos cells transfectants, expressing CD33 and CD34 antigens, also suggest that CD34 and CD33 molecules do not interact with other cell surface receptors presented on lymphomyeloid (KG1), T (MOLT4, CEM), B (Raji, Daudi), erythroid (K562), myeloid (HL-60, U937) or peripheral blood mononuclear cells. However, altered glycosylation, sialylation, phosphorylation or other posttranslational modifications in the transfected molecules may modulate receptor function. Whether such changes can explain the lack of
binding of CD34 cell transfectants to the L-selectin on peripheral blood cells or T cell lines in our experimental system remains to be determined. Alternatively other regulatory pathways may be involved in the regulation of the adhesive properties of vascular and haemopoietic stem cell CD34 forms.

As an alternative approach to study the function of CD33 antigen we decided to clone its murine homologue, which would allow further functional and in vivo studies to be performed. Two cDNA clones, differing by an 83 bp insertion in the cytoplasmic portion, were isolated from a mouse bone marrow cDNA library and were predicted to encode two murine CD33 isoforms. The two murine CD33 isoforms, m33-A and m33-B, share identical extracellular (240 aa) and transmembrane (27 aa) domains, as well as the first 19 residues in their cytoplasmic regions, but have alternate COOH-terminal sequences of 48 and 117 residues respectively. The predicted protein core molecular masses are 37 kD for the m33-A and 45 kD for the m33-B isoforms. The two murine CD33 cDNA clones are probably derived from alternatively spliced transcripts of the same gene, as suggested by the presence of exon splice donor/acceptor sequences flanking the 83 bp insertion. The cDNAs and derived amino acid sequences share the general structural characteristics of members of the Ig gene superfamily and show extensive similarity to human CD33 (71% overall amino acid similarity). The regions with the most significant similarities include the first and second Ig-like domains. This suggests that these regions are important for CD33 function. Less conserved are the cytoplasmic domains, with the most significant divergence in their C-termini occurring, between the human CD33 and murine m33-B isoform.

The similarity between the murine and human CD33 molecules is also paralleled by their mRNA expression pattern. Northern blot analysis showed expression of murine CD33 isoforms in bone marrow, spleen,
thymus, brain, liver and in the multipotential progenitor cell line A4, myelomonocytic cell line WEHI 3B, myeloid cell line M1 and macrophage cell line P388D1. These data, together with the preliminary in situ hybridisation analysis, indicate that CD33 expression in the mouse, like that in humans, is confined to the haemopoietic system and probably limited to the myelomonocytic lineage. This suggests that the function of CD33 antigen in haemopoiesis may be conserved between man and mouse.

What could be the function of CD33 antigen? Analysis of the predicted protein sequences of murine CD33 isoforms reveals some interesting structural features and homologies with previously characterised proteins, and gives some insight into the probable function of this protein.

Although the binding studies, described in the first results chapter, did not indicate that CD33 mediated adhesion in the binding systems tested, in the light of the high degree of similarity found between the extracellular domains of murine CD33 isoforms and a number of adhesion molecules, it still can be hypothesised that the murine CD33 homologue may function as an adhesion molecule. Thus the extracellular domain of murine CD33 antigen is homologous to MAG, CD22, CD4 and members of the CEA family, TM CEA-2 and BGP-1, for all of which molecules homophilic and heterophilic cell-extracellular matrix and/or cell-cell adhesion interactions have been demonstrated (Poltorak et al., 1987; Stamenkovic et al., 1991; Benchimol et al., 1989; Thompson et al., 1991; Barclay et al., 1993). Furthermore the human CD33 locus has been mapped to chromosome 19q13.3 (Peiper et al., 1988), very close to the region containing the MAG and CEA family of genes 19q13.1 and 19q13.3 (Barclay et al., 1993), suggesting that this region of chromosome 19 contains a cluster of adhesion molecules. Similarly the mouse
counterparts of the MAG and CEA genes are linked on mouse chromosome 7 (Thompson et al., 1991). The chromosome localisation of the murine CD33 gene should be helpful in clarifying the relation with the MAG and CEA cluster of genes.

The possible role of murine CD33 in adhesion is further supported by the presence of a GAG attachment site in the extracellular region of the murine CD33 sequence. Glycosaminoglycans (GAG) have been described in two leucocyte adhesion molecules CD44 and syndecan and have been shown to be important regulators of cell-extracellular matrix interactions. Thus the interaction of syndecan with extracellular matrix proteins collagens type I, III and V, thrombospondin and fibronectin may be regulated by differences in the types and numbers of glycosaminoglycans attached to syndecan (Marynen et al., 1989). Whether similar changes in the attached GAG and/or changes in glycosylation of the murine CD33 molecule could regulate the adhesion of myeloid progenitors in the bone marrow stroma and/or of monocytes within the tissues remains to be determined. However the identification of putative GAG attachment sites near the transmembrane domain in both murine and human CD33 sequences and the conservation of N-glycosylation sites between the human and mouse CD33 sequences, suggest that they may be of functional importance.

The GAG, heparan sulphate, of syndecan in addition to binding extracellular matrix components is also required for the binding of basic fibroblast growth factor to its high affinity receptor (Yayon et al., 1991; Olwin et al., 1992). Furthermore CD44 was shown in a model system to bind to MIP-1β, thereby immobilising it for presentation to its receptor on lymphocyte (Tanaka et al., 1993). Now many growth factors have been demonstrated to bind to heparan sulphate and chondroitin sulphate. Whether a dual requirement for growth factor binding to protein receptors
and glycosaminoglycan-type receptor is a general phenomenon, has yet to be demonstrated. It is however conceivable that a heparin-like glycosaminoglycan attached to CD33 molecules could sequester and concentrate growth factors, and/or serve as a low affinity receptor for growth factors and deliver them to one or several high affinity receptors, regulating cellular processes such as proliferation or differentiation within the bone marrow or in other tissues. Alternatively, the sequence similarity in the extracellular domains between murine CD33 antigen and bFGF receptor (and of PDGF receptor and IL-1 receptor) may suggest that murine CD33 may bind bFGF or other cytokines through a mechanism distinct from GAG recognition and similar to those of growth factor and cytokine receptors. Finally CD33 might recognise simultaneously an ECM component and a growth factor. This would have interesting functional implications so that for example, the response of myeloid progenitors or monocytes to the growth factor would depend on their specific adherence to sites where the growth factor is secreted or concentrated locally.

The relationship between the two alternatively spliced murine CD33 isoforms and the single CD33 form identified in humans, is not clear. It remains to be established whether CD33 splice variants exist in man and whether the murine splice variants have a functional significance, which would be expected to be conserved between evolutionarily distant species. The biological significance of the different cytoplasmic domains of murine CD33 needs to be resolved. Differential splicing, resulting in the tissue-specific expression of transcripts encoding alternate cytoplasmic domains has been identified for rat liver prolactin (PRL) receptor (Edery et al., 1989) and MAG (Ishiguro et al., 1991). The tissue specific expression of murine CD33 isoforms remains to be determined at both the mRNA and protein level, but the preliminary PCR analysis suggests that the m33-B isoform occurs in myeloid precursors.
(A4, M1 and WEHI 3B cell lines), but not in a subset of mature macrophages (P388 cell line). This raises the possibility that perhaps the two murine CD33 isoforms are associated with different stages of myeloid differentiation and suggests that the alternate cytoplasmic domains may confer functional differences.

In this respect it is interesting that the 83 bp insertion generates distinct potential phosphorylation sites in the cytoplasmic tails of the m33-B isoform. Five Protein Kinase C (PKC) consensus sequences were found in m33-B, compared to one PKC site in m33-A. Casein kinase 2 and cAMP and cGMP dependent protein kinase phosphorylation sites are also predicted. In addition tyrosine residues were found in both isoforms. These multiple phosphorylation sites, together with the homologies of the cytoplasmic domains of murine CD33 variants to the cytoplasmic regions of CD4, CD2, MAG-L, BGP-1 and to a number of molecules known to be phosphorylated or involved in intracellular signalling, suggest that the intracytoplasmic region of murine CD33 variants may be phosphorylated, although the functional consequences of this are unknown. Some speculations however can be advanced.

One possibility is as for several integrins, where phosphorylation of the cytoplasmic domain of the β subunit is commonly involved in the regulation of affinity for their ligands (inside-out signalling) (Hibbs et al., 1991; Hynes, 1992). The adhesive properties of murine CD33 isoforms, might be modulated by a similar mechanism. For example the signal for such phosphorylation may be provided by some inflammatory mediator, produced by the endothelium where the attachment of monocytes is needed. However so far no examples for inside-out signalling exist within the members of the Ig gene superfamily.

Another alternative could be, as in the case of BGPa and c variants, that the splicing does not affect the adhesive properties of the
murine CD33 isoforms but the biological significance of the distinct cytoplasmic domains is to transduce different signals from extracellular ligands to the cell interior (outside-in signalling). An example for such a versatility can be given with integrins, where has been shown that different cytoplasmic domains trigger different functions (collagen gel contraction or migration) when transfected into cells (Chan et al., 1992). It has also been proposed that the smaller form of PRL receptor functions in ligand transport across epithelial barriers in liver rather than in signal transduction (Edery et al., 1989). The developmentally regulated expression of MAG isoforms suggests an important role of L-MAG in myelin formation (Ishiguro et al., 1991) and recently an association between Fyn tyrosine kinase and the cytoplasmic domain of L-MAG has been demonstrated (Umemori et al., 1994). The impaired myelination in Fyn-deficient mice further confirmed their interaction and role in myelination.

The cytoplasmic domains of neither murine CD33 isoform is large enough to posses intrinsic enzyme activity and does not show sequence homologies to the catalytic domains of tyrosine kinases or phosphatases, or signal transduction sequence motifs. Nevertheless, the cytoplasmic domains of murine CD33 isoforms may be implicated in signal transduction, for example by directly interacting with tyrosine kinases, similarly to CD4 and L-MAG (the cytoplasmic domains of which share some degree of similarity with CD33 cytoplasmic region), or by interaction with other membrane proteins or cytoskeletal components, which could in turn activate the intracellular machinery that controls cell growth, behaviour and differentiation.

Finally there is the possibility that more murine CD33 alternatively spliced forms may exist. These could be soluble forms (by analogy with CEA, NCAM, growth factor receptors), GPI anchored forms or other
transmembrane forms containing additional sequences in the extracellular domain. In the latter case it may be that additional sequences in the extracellular domain would mediate different binding functions such as those observed with inserted sequences in NCAM splice variants (Santoni et al., 1989) and in the FGF receptor (Miki et al., 1992).

The biological significance of the two murine CD33 mRNA species (of 2.0 kb and 3.9 kb), identified by northern blot analysis, also remains unclear. A further level of complexity is present in that both murine CD33 transcripts appear to encode for either of the spliced murine CD33 isoforms. The transcripts are differentially expressed in murine tissues and in mature macrophage subpopulations (P388 cell line) and myeloid precursors (A4, WEHI 3B and M1 cell lines). The presence of two mRNA transcripts with different 3'UTR may confer heterogeneity in signals controlling the length and stage of polyadenylation, in motifs affecting translation or in determinants associated with mRNA stability. For example, an instability ATTTA motif (Shaw and Kamen, 1986; Schuler and Cole, 1988) was found in the 3'UTR of the murine sequence and such a motif is present at a similar position in the human CD33 sequence, for which two mRNA species have been identified as well. Thus the existence of two transcripts, each able to encode either of the splice variants might suggest that the expression of the two murine CD33 isoforms is closely regulated at different stages of myeloid maturation and in different tissues or during neoplastic development.

The function of CD33 myeloid antigen is unknown. The availability of murine CD33 cDNA clones will undoubtedly help in determining the function of CD33 antigen and its distinctive murine cytoplasmic domains. One of the next steps surely should be the expression of the gene products in homologous or heterologous cell systems for functional studies. This will allow antibodies to be made against the murine CD33 variants,
adhesion of both isoforms to various ECM and cell surface ligands as well as growth factors to be tested, the putative GAG and carbohydrates attached to the murine CD33 molecules to be analysed and phosphorylation and possible signalling mechanisms to be studied. Full-length genes and associated flanking sequences can be isolated. Eventually cis and trans acting regulatory elements can be identified. Monoclonal antibodies will enable the precise immunohistochemical localisation of both murine CD33 isoforms. In vivo experiments with monoclonal antibodies may perturb the consequences of CD33 mediated interactions and may be revealing in regard to the function of CD33 antigen. Ultimately gene knockout through homologous recombination may help in defining its function.
REFERENCES


Afar DEH, Stanners CP, Bell JC: Tyrosine phosphorylation of biliary glycoprotein, a cell adhesion molecule related to carcinoembryonic antigen. Biochem Biophys Acta 1134:46, 1992

Aizawa S, Tavassoli M: In vitro homing of hematopoietic stem cells is mediated by a recognition system with galactosyl and mannosyl specificities. Proc Natl Acad Sci USA 84:4485, 1987


Aruffo A, Seed B: Molecular cloning of a CD28 cDNA by a high efficiency Cos cell expression system. Proc Natl Acad Sci USA 84:8573, 1987


Barbas JA, Chaix JC, Steinmetz M, Goridis C: Differential splicing and alternative polyadenylation generates distinct NCAM transcripts and proteins in mouse. EMBO J 7:625, 1988


Berendt A, McDowall A, Craig AG, Bates PA, Sternberg MJE, Marsh K, Newbold CI, Hogg N: The binding site on ICAM-1 for plasmodium falciparum-infected erythrocytes overlaps, but is distinct from the LFA-1 binding site. Cell 68:71, 1992


Beyers AD, Barclay AN, Law DA, He Q, Williams AF: Activation of T lymphocytes via monoclonal antibodies against rat cell surface


Brown J, Greaves MF, Molgaard HV: The gene encoding the stem cell antigen, CD34, is conserved in mouse and expressed in haemopoietic progenitor cell lines, brain and embryonic fibroblasts. Int Immunol 3:175, 1991

Brown WM, Christie DL, Dzegielewska KM, Saunders NR, Yang F: The rat protein encoded by clone pp63 is a fetuin/α2-HS glycoprotein-like molecule, but is it the tyrosine kinase inhibitor pp63? Cell 68:7, 1992


Burn TC, Satterthwaite AB, Tenen DG: The human CD34 hematopoietic cell antigen promoter and a 3' enhancer direct hematopoietic expression in tissue culture. Blood 80:3051, 1992


Campbell AD: The role of haemonectin in the cell adhesion mechanisms of bone marrow. Hematol Pathology 6:51, 1992


Co MS, Avdalovic NM, Caron PC, Avdalovic MV, Scheinberg DA, Queen C: Chimeric and humanized antibodies with specificity for the CD33 antigen. J Immunol 148:1149, 1992


Dalgleish AG, Beverley PCL, Clapham PR, Crawford DH, Greaves MF, Weiss RA: The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature 312:763, 1984


Dougherty GJ, Landorp PM, Cooper DL, Humphries RK: Molecular cloning of CD44R1 and CD44R2, two novel isoforms of the human


Dustin ML: Two-way signalling through the LFA-1 lymphocyte adhesion receptor. Bioessays 12:421, 1990


Flanagan JG, Chan DC, Leder P: Transmembrane form of the kit ligand growth factor is determined by alternative splicing and is missing in the Sl^d mutant. Cell 64:1025, 1991


Gordon MY, Riley GP, Clarke D: Heparan sulphate is necessary for the adhesive interactions between human early haemopoietic progenitor cells and the extracellular matrix of the marrow microenvironment. Leukaemia 2:804, 1988


Hayashi Y, Haimovich B, Reszka A, Boettiger D, Horwitz A: Expression and function of chicken integrin β1 subunit and its


Heaney ML, Golde DW: Soluble hormone receptors. Blood 82:1945, 1993

Hefta SA, Hefta LJF, Lee TD, Paxton RJ, Shively JE: Carcinoembryonic antigen is anchored to membranes by covalent attachment to a glycosylphosphatidylinositol moiety: Identification of the ethanolamine linkage site. Proc Natl Acad Sci USA 85:4648, 1988


Higashiyama S, Abraham JA, Miller, Fiddles JC, Klagsbrun M: A heparin-binding growth factor secreted by macrophage-like cells that is related to EGF. Science 251:936, 1991


Hunter DD, Murphy MD, Olsson CV, Brunken WJ: S-laminin expression in adult and developing retinae: a potential cue for photoreceptor morphogenesis. Neuron 8:399, 1992


Khan WN, Hammarstrom S, Ramos T: Expression of antigens of the carcinoembryonic antigen family on B cell lymphomas and Epstein-Barr immortalised B cell lines. Int Immunol 5:265, 1993


Kolbinger F, Schwartz K, Brombacher F, von Kleist S, Grunert F: Expression of an NCA cDNA in NIH/3T3 cells yields a 110 kD glycoprotein, which is anchored into the membrane via glycosyl phosphatidyl inositol. Biochem Biophys Res Commun 161:1126, 1989

Korneluk RG, McLeod HL, McKeithan TW, Brooks JD, MacKenzie AE: A chromosome 19 clone from a translocation breakpoint shows close linkage and linkage disequilibrium with myotonic dystrophy. Genomics 4:146, 1989


Lai C, Brow, MA, Nave KA, Noronha AB, Quarles RH, Bloom FE, Milner RJ, Sutcliffe JG: Two forms of 1B236/myelin associated glycoprotein, a cell adhesion molecule for postnatal neural development, are produced by alternative splicing. Proc Natl Acad Sci USA


Lowe J, Stoolman LM, Nair RP, Larsen RD, Berhend TL, Marks RM: ELAM-1-dependent cell adhesion to vascular endothelium
determined by a transfected human fucosyltransferase cDNA. Cell 63:475, 1990


McKeithan TW, Rowley JD, Shows TB, Diaz MO: Cloning of the chromosome translocation breakpoint junction of the t(14;19) in chronic lymphocytic leukemia. Proc Natl Acad Sci USA 84:9257, 1987


Metcalf D: The molecular control of blood cells. Cambridge, Massachusetts, Harvard University Press, 1988


Molgaard HV, Spurr NK, Greaves MF: The hemopoietic stem cell antigen, CD34, is encoded by a gene located on chromosome 1. Nucleic Acids Res 10:459, 1989


Murakami M, Hibi M, Nakagawa N, Nakagawa T, Yasukawa K, Yamanishi K, Taga T, Kishimoto T: IL-6-induced homodimerization of gp130 and associate activation of tyrosine kinase. Science 260:1808, 1993


210


Pandey NB, Marzluff WF: The stem-loop structure at the 3' end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability. Mol Cell Biol 7:4557, 1987


Raines EW, Lane TF, Iruela-Arispe ML, Ross R, Sage EH: The extracellular glycoprotein SPARC interacts with platelet derived growth factor (PDGF)-AB and -BB and inhibits the binding of PDGF to its receptors. Proc Natl Acad Sci USA 89:1281, 1992


Rojas M, Fuks A, Stanners CP: Biliary glycoprotein, a member of the immunoglobulin supergene family, functions in vitro as a Ca++ dependent intercellular adhesion molecule. Cell Growth Diff 1:527, 1990


Russel SW, Gordon S: Macrophage Biology and Activation. Current Topics in Microbiology and Immunology. Springer-Verlag, Berlin, 1992


Scholey AB, Rose SP, Zamani MR, Bock E, Schachner M: A role of the neural cell adhesion molecule in a late, consolidating phase of glycoprotein synthesis six hours following passive avoidance training of the young chick. Neuroscience 55:499, 1993


217


Walsh FS, Furness J, Moore SE, Ashton SV, Doherty P: Use of the neural cell adhesion molecule VASE exon by neurons is associated with a specific down regulation of NCAM dependent neurite outgrowth in the developing cerebellum and hippocampus. J Neurochem 59:1959, 1992


(HPCA-1) associated with human hemopoietic progenitor cells. Leukemia 1:417, 1987


Williams AF: A year in the life of the immunoglobulin superfamily. Immunol Today 8, 298, 1987


Wittuhn BA, Quelle FW, Silvennoinen O, Yi T, Tang B, Miura O, Ihle JN: JAK2 associates with the erythropoietin receptor and its
tyrosine phosphorylated and activated following stimulation with erythropoietin. Cell 74:227, 1993


Yayon A, Klagsburn M, Esko JD, Leder P and Ornitz DM: Cell surface heparin-like molecules are required for binding of basic fibroblast growth factor to its high-affinity receptor. Cell 64:841, 1991


The work described in this thesis has been published in part in:

