Analysis of the Tumour Microenvironment in CLL

Hamilton, Emma

Awarding institution:
King's College London

Title: Analysis of the Tumour Microenvironment in CLL

Author: Emma Hamilton

Analysis of the Tumour Microenvironment in CLL

Emma Jane Hamilton

A THESIS PRESENTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

KING'S COLLEGE LONDON

2013

Declaration

I hereby declare that I alone composed this thesis and that the work is my own, except where stated otherwise.

Emma Hamilton

September 2013

Acknowledgements

Many thanks to my primary supervisor, Professor Shaun Thomas, for giving me the opportunity to work in his laboratory and for his patience and guidance throughout my time at the Rayne Institute. I would also like to thank my secondary supervisor, Professor Stephen Devereux for his countless ideas and guidance of my work and Professor G J Mufti for giving me the opportunity to work in the Department of Haematological Medicine.

I would like to acknowledge our collaborators without whom the project would not have been possible. Thank you for being so generous with your time and knowledge: Professor Chris Pepper for giving me the opportunity to learn microarray analysis; Professor Edward Marcotte and Dr Daniel Boutz for their hospitality and training in proteomics and mass spectrometry; Dr Eric Blanc for his expertise in Bioinformatics.

To my colleagues in the Department of Haematological Medicine, thank you for your friendship, advice and support over the last four years particularly Dr Andrea Buggins in the CLL team for taking me under her wing and Dr Steve Orr and Dr Mina Lamadema in the Cell Cycle group. I am also indebted to the medical staff at King's College Hospital and especially the patients who kindly donated samples for research and without which these studies could not have been completed.

This PhD was generously funded by the MRC and my rotation in mass spectrometry at the University of Texas at Austin was funded by The British Society for Haematology.

List of abbreviations

7-AAD	7-Amino-Actinomycin D
ACN	Acetonitrile
APE1	Apurinic/apyrimidinic Endonuclease 1
APEX	Absolute Protein Expression
BCR	B Cell Receptor
BLAST	Basic Local Alignment Tool
BSA	Bovine Serum Albumin
BTK	Bruton's tyrosine kinase
CEBP	CCAAT/enhancer-binding protein
ChIP	Chromatin Immunoprecipitation
CID	collision induced dissociation
CLL	Chronic Lymphocytic Leukaemia
CTL	Cytotoxic T Lymphocytes
CXCL1	Chemokine ligand 1
DAVID	Database for Annotation, Visualisation and Integrated Discovery
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulphoxide
DTT	Dithiothreitol
ECGS	Endothelial Cell Growth Supplement
ESI	Electrospray Ionisation

FACS	Fluorescence activated cell sorting
F-actin	Filamentous Actin
FBS	Foetal Bovine Serum
FC	Fold Change
FDR	False Discovery Rate
FISH	Fluorescence in situ hybridisation
FITC	Fluorescein Isothiocyanate
FSC	Forward Scatter
GO	Gene Ontology
GOC	Gene Ontology Compartment
GOF	Gene Ontology Function
GOP	Gene Ontology Process
GSCA	Gene Set Control Analysis
hBM-MSCs	human bone marrow mesenchymal stromal cells
HCL	Hairy Cell Leukaemia
HDBEC	Human Dermal Blood Endothelial Cells
HDMEC	Human Dermal Microvascular Endothelial Cells
hEGF	Human Epidermal Growth Factor
HMEC-1	Human Microvascular Endothelial Cells
HMGB2	High Mobility Group Box 2
HPRD	Human Protein Reference Database
HRP	Horseradish Peroxidase
Hsp	Heat-Shock Protein

HTS	High-Throughput-Screening
IAM	Iodoacetamide
ID	identifier
IGF-1	Insulin Like Growth Factor 1
IL-6	Interleukin 6
iTRAQ	isobaric Tags for Relative and Absolute Quantification
IVT	*In Vitro* Transcription
LC	Liquid Chromatography
LC-MS/MS	Liquid Chromatography tandem Mass Spectrometry
LFA-1	Lymphocyte function-associated antigen-1
LTQ	LTQ-Orbitrap mass spectrometer
m/z	mass to charge ratio
mAb	Monoclonal Antibody
MALDI	Matrix Assisted Laser Desorption Ionisation
MEACS	Myosin Heavy Chain IIA Exposed apoptotic cell
MHC	Major Histocompatibility Complex
MM	Multiple Myeloma
MPRIP	Myosin Phosphatase Rho Interacting Protein
MS	Mass Spectrometry
MT	Metallothioneins
NCBI	National Center for Biotechnology Information
NFκB	Nuclear Factor kappa B
NHS	N-hydroxysuccinimide

OMIM	McKusick's Online Mendelian Inheritance in Man
PAGE	Polyacrylamide Gel Electrophoresis
PBMC	Peripheral Blood Mononuclear Cells
PBS	Phosphate Buffered Saline
PFA	Paraformaldehyde
PI	Propodium Iodide
PI3K1	p85α subunit of phosphatidylinositol 3-kinase
RIN	RNA Integrity Number
RMA	Robust Multichip Averaging
RNases	Ribonucleases
RPMI	Rosswell Park Memorial Institute medium
RT	Reverse Transcriptase
SDS	Sodium dodecyl sulphate
SILAC	Stable Isotope Labeling with Amino acids in cell Culture
SSC	Side Scatter
STAT1	Signal transducer and activator of transcription 1
Syk	Spleen tyrosine kinase
TAC	Tetrameric Antibody Complexes
TBZ	Thiabenazole
TCR	T cell receptor
TdT	Terminal deoxynucleotidyl Transferase
TF	Transcription Factor
TFE	Trifluorethanol

TBS	Tris Buffered Saline
TK	Thymidine kinase
TMHMM	Transmembrane hidden markov model
UDG	Uracil-DNA Glycosylase
VEGF	Vascular Endothelial Growth Factor
Velos	Velos-Orbitrap mass spectrometer
VLA-4	Vascular cell adhesion molecule-4
WBC	White blood cell
WHO	World Health Organisation
WT	Whole transcriptome
XPO1	Exportin 1

Abstract

This thesis applies systems biology approaches to obtain a better understanding of complex cellular systems in Chronic Lymphocytic Leukaemia (CLL).CLL is a cancer of the blood which affects B cells. CLL cells accumulate quickly in the body but die rapidly by apoptosis when cultured in the lab, suggesting that CLL cell survival can be modulated by tumour cell interactions with the microenvironment *in vivo*.

The first part of the thesis describes *in vitro* assays which demonstrated that CLL cell contact with endothelial cells in a co-culture system promotes CLL cell survival. The precise mechanisms involved remain unknown, however, they are thought to include both secreted factors and receptor:ligand interactions between CLL cells and the microenvironment.

The second part of the thesis is based on gene expression profiling and investigates the transcriptional effects induced in CLL cells by co-culture. We identified genes that were statistically significantly up regulated >1.5 fold compared with liquid culture ($q<0.001$) and were common to two endothelial systems.

The third part of the thesis focuses on cell surface proteins which allow CLL cells to communicate with other cells. In order to obtain a comprehensive analysis of the cell surface proteome of CLL cells, we developed a method utilising cell impermeable, cleavable sulfo-NHS-SS-biotin to enrich for cell surface proteins followed by high-content mass spectrometry (LC-MS/MS) for identification. Somewhat surprisingly, a number of proteins were identified which have traditionally been classified as intracellular. Some of these have been previously identified as specific binding partners for CLL B-cell receptor (BCR) stereotypes. It is possible that the BCR activation observed in CLL occurs as a consequence of binding to these aberrantly expressed autoantigens. This study provides insights into the mechanisms that underpin the cytoprotection afforded by CLL cell-endothelial cell co-culture and identifies potential new targets for disrupting these signals *in vivo*.

Supplementary Data

The CD provided with this thesis contains the following supplementary information:

Folder Chapter4:

Supplementary Table 4.2.6 Differential expression of mRNA by fold change

Supplementary Table 4.2.7 Enriched GO Term Annotations

Supplementary Table 4.2.8 Common up-regulated mRNA

Supplementary Table 4.2.10 DAVID Analysis of putative TF module controlled mRNA

Folder Chapter 5:

Supplementary Table 5.2.3 HeLa total lysate proteins

Supplementary Table 5.2.4a HeLa cell surface proteins

Supplementary Table 5.2.4b DAVID analyses of HeLa cell surface proteins

Supplementary Table 5.2.5 HeLa DNAse treatment

Supplementary Table 5.2.7 Preliminary CLL cell surface proteins

Supplementary Table 5.2.8a CLL cell surface proteins

Supplementary Table 5.2.8b DAVID analyses of CLL cell surface proteins

Supplementary Table 5.2.9 CLL total lysate proteins

Supplementary Table 5.2.10 Annexin V negative CLL cell surface proteins

Supplementary Table 5.3.1 TBZ treated HUVEC cell surface proteins

Table of Figures

Table of Tables

Chapter 1

General Introduction

1. General Introduction

1.1 Introduction to Chronic Lymphocytic Leukaemia (CLL)

Chronic Lymphocytic Leukaemia (CLL) is the most common form of adult leukaemia in the Western World, accounting for about 40% of all leukaemias over the median age of 65-70 years of age. The incidence of CLL in England and Wales has been reported as 6.5 per 100,000 per year (4). CLL is 20-30% more common in European, Australian and North American White and Black populations than in India, China or Japan (5) and is more common in males than females. Although the disease may present with systemic symptoms, such as night sweats, tiredness, unintentional weight loss and symptoms of anaemia or infection, 70-80% of patients are diagnosed due to an incidental finding of lymphocytosis upon a routine full blood count (6).

CLL is regarded as a clinically heterogeneous disease and is characterised by a neoplasm of morphologically mature clonal $CD5^+$ B lymphocytes (7), which are immunologically less mature. The World Health Organization (WHO) classification of haematopoieic neoplasias describes CLL as a stage of Small Lymphoytic Lymphoma (SLL), distinguishable only by its leukemic appearance (8). The cells found in a blood smear are small mature lymphocytes with a narrow border of cytoplasm and a dense nucleus, lacking distinctive nuclei and having aggregated chromatin. CLL cells arise from the bone marrow and progressively accumulate in the blood, bone marrow and lymphoid tissues (9). The major cause of morbidity and mortality in CLL are secondary infections, attributed to abnormal immune function either as a result of the primary disease or through the immunosuppressive effects of the management of CLL. 80% of CLL patients will suffer infections during the disease course (10), the most common immune defect is hypoglobulinemia rendering patients susceptible to bacterial infections. There are also various defects in the function of T cells that are induced *via* contact with CLL cells, including alterations in CD40L activation (11) and abnormalities in the expression of genes involved in the differentiation of $CD4^+$ cells, cytoskeleton formation and vesicular trafficking and cytotoxicity of $CD8^+$ cells (12). The secretion of soluble mediators has also been shown to alter T cell function (13). Impaired immune surveillance in the disease is characteristic of CLL, contributing to a high incidence of secondary malignancies (14).

1.2 Diagnosis

A diagnosis of CLL requires a count of greater than 5,000 B-lymphocytes per micorlitre in the peripheral blood for a period of at least three months. The clonality of these cells must be confirmed by flow cytometry. The monoclonal population of B cells in CLL typically express CD5, CD19 and CD23 and have reduced levels of IgM, IgD and CD79b present on the membrane surface compared

to those found on normal B cells (15). Other tests not required to establish diagnosis may be performed to give an insight into predicting prognosis or tumour burden. These include molecular genetics, fluorescence *in situ* hybridisation (FISH), mutational status of *IgVH,* serum marker and bone marrow examination.

Dohner *et al*. (16) showed that cytogenetic lesions can be identified in more than 80% of CLL cases, with deletions in the long arm of chromosome 13 (del(13q14.1)) being the most common abnormality. Evidence is growing to suggest that detection of chromosomal deletions has prognostic significance, in particular patients with disease with del (17p) have poor prognosis and appear to be resistant to standard chemotherapy (17, 18). Detection of cytogenetic abnormalities may influence therapeutic decisions and repetition of FISH analyses could be important prior to commencing second and third line treatment as additional defects may be acquired during the course of the disease (19).

The clinical course of patients with CLL varies dramatically with some patients having stable disease for a decade or more (20), whilst others have more aggressive disease. As CLL advances, the patient may experience swollen lymph nodes, spleen and liver, as well as anaemia, infections and ultimately bone marrow failure. There is currently no cure for the disease. Staging systems implemented by Rai and Binet are used to assess the extent of disease in a patient (**Table 1.1**). These systems are still used when assessing disease progression and treatment.

However, these staging systems have no value in predicting the clinical course of a patient diagnosed with early stage CLL. A growing number of prognostic factors have now been identified to predict the prognosis of such patients (21).

Rai Staging System	**Description**	**Binet Classification**	**Description**
Stage 0	Stage 0 CLL is characterized by absolute lymphocytosis (>15,000/mm^3) without adenopathy, hepatosplenomegaly, anaemia, or thrombocytopenia.	Clinical Stage A	Clinical stage A CLL is characterized by no anaemia or thrombocytopenia and fewer than three areas of lymphoid involvement (Rai stages 0, I, and II).
Stage I	Stage I CLL is characterized by absolute lymphocytosis with	Clinical Stage B	Clinical stage B CLL is characterized by no anaemia

	lymphadenopathy without hepatosplenomegaly, anaemia, or thrombocytopenia.		or thrombocytopenia with three or more areas of lymphoid involvement (Rai stages I and II).
Stage II	Stage II CLL is characterized by absolute lymphocytosis with either hepatomegaly or splenomegaly with or without lymphadenopathy.		
Stage III	Stage III CLL is characterized by absolute lymphocytosis and anaemia (haemoglobin <11 g/dL) with or without lymphadenopathy, hepatomegaly, or splenomegaly.	Clinical Stage C	Clinical stage C CLL is characterized by anaemia and/or thrombocytopenia regardless of the number of areas of lymphoid enlargement (Rai stages III and IV).
Stage IV	Stage IV CLL is characterized by absolute lymphocytosis and thrombocytopenia (<100,000/mm^3) with or without lymphadenopathy, hepatomegaly, splenomegaly, or anaemia.		

Table 1.1 CLL Staging System. Adapted from the National Cancer Institute (NCI) U.S. National Institutes of Health (http://www.cancer.gov/cancertopics/pdq/treatment/CLL/healthprofessional/10.cdr#Section_10).

1.2.1. Prognostic factors

CLL may progress slowly and patients with early, uncomplicated disease are not usually treated. Patients with advancing disease are treated with chemotherapy and monoclonal antibodies. CLL is considered incurable and treatments vary depending on diagnosis and the progression of the disease. As CLL progresses slowly in most cases, treatment is usually delayed until symptoms arise, indicating that the disease has progressed and may affect a patient's quality of life. This is known as the 'watch and wait' approach.

CLL is considered an indolent disease, with life expectancy of 7 or more years. The average age of patients at diagnosis is 65 years. CLL also has a highly variable clinical course with some patients progressing much more quickly and severely than others. It would therefore be extremely valuable to be able to predict which patients are likely to have progressive disease and which are not. There has been much research into potential prognostic markers associated with CLL (reviewed in (22)), including CD44, CD49d, CD38, ZAP-70 VEGF, thymidine kinase (TK), LPL, CLLU1 expression, genome wide gene expression profiling (23) and immunoglobulin heavy chain variable (*IgVH*) region mutational status (24).

CLL cells express immunoglobulin, which may or may not have undergone somatic mutations in the *IgVH* region. Patients with leukaemia cells which use an unmutated *IgVH* gene have a worse outcome than those patients using a mutated *IgVH* gene (25), whilst the use of the *VH3.21* gene is an unfavourable marker independent of the mutational status (26). *IgVH* status has also been used to give clues to the origin of the CLL cell. *IgVH* mutational status is one of the most powerful prognostic features. Within the germinal centre, normal B cells undergo somatic hypermutation of their *IgVH* genes after exposure to T-dependent antigens. Around half of all CLL patients show evidence of somatic hypermuation, which is arbitrarily defined as a 2% deviation from the germ line sequence. These patients have better prognosis than those with unmutated *IgVH* genes.

CD38 is another important prognostic feature in CLL. Its expression is dynamic and changes in response to contact with activated $CD4^+$ T cells (27). CD38 is a transmembrane protein which may function as an ecto-enzyme and receptor involved in adhesion processes. Interactions between CD38 and its cognate receptor, CD31 promote the proliferation and survival of CLL cells (28).

Patients with clones having few or no *IgVH*-gene mutations (29) or with many $CD38^+$ or ZAP-70^+ B cells (30) generally have an aggressive disease, whereas the course of the disease in patients with mutated *IgVH* clones or few $CD38^+$ or ZAP-70^+ B cells is more likely to be indolent. The doubling time of circulating CLL cells (proliferation index) may also be used as a prognostic indicator (31).

Whole-genome sequencing of four CLL patients followed by analysis of samples from a further 363 patients have recently identified four genes that are recurrently mutated: notch 1 (*NOTCH1*) and exportin 1 (*XPO1*), predominantly found in patients with unmutated *IgVH*, and myeloid differentiation primary response gene 88 (*MYD88*) and kelch-like 6 (*KLHL6*), which are predominantly found in cases with mutated *IgVH* (32). Constitutive activation of NOTCH1 signalling has been observed in CLL cells (33) and about 4% of CLL patients are believed to have *NOTCH1* mutations (32). Stromal cell-mediated Notch signalling has been shown to play a role in CLL resistance to chemotherapy. In one study, CLL cells were co-cultured with autologous and allogeneic human BM-mesenchymal stromal cells (hBM-MSCs). The co-culture protected CLL cells from apoptosis spontaneously and

following treatment of CLL cells with Fludarabine, Cyclophosphamide, Bendamustine, Prednisone and Hydrocortisone. However, treatment with a combination of anti-Notch-1, Notch-2 and Notch-4 antibodies reversed the protective effect of co-culture by day 3, even in presence of the drugs listed (34). Together these data suggest that blocking Notch signalling could be an additional tool to overcome drug resistance in CLL and show how understanding CLL: microenvironment interactions may help guide new therapeutic approaches.

1.2.2. CLL is a proliferative disease

Our understanding of CLL has greatly evolved over the past fifteen years. In the past, CLL was considered a disease of accumulation of mature B cells rather than proliferation. However, recent studies measuring the dynamic cellular kinetics of CLL cells have shown CLL clones to be more dynamic than previously assumed (35). Cellular birth rates were measured by calculating kinetic profiles after giving patients deuterated water. The birth rates varied from 0.1% to greater than 1% of the clone per day with a correlation between cellular birth rates over 0.35% and disease progression. If a patient has a clonal burden of 10^{12} cells, these birth rates suggest the production of 10^{9}-10^{10} leukaemic cells each day.

Recently, Lin *et al*., (36) performed single-molecule telomere length and telomere fusion analysis of CLL patients at different disease stages and identified the shortest telomeres ever recorded in primary human tissue. This supports the proposition that significant cell division occurs in CLL. They also show that critical telomere shortening, dysfunction and fusion contribute to disease progression. However, this remains a contentious issue as other groups debate which is causative, the short telomeres observed in CLL or chromosomal abnormalities (37).

1.2.3. The Microenvironment in CLL

The infiltration of the bone marrow and the lymph node by abnormal B cells is one of the main manifestations of the disease that leads to progression and ultimately organ failure (38). It is therefore important to understand how interactions with the tissue microenvironment contribute to the pathogenesis of the disease. Important parameters that govern the progression of the disease include the proliferative capacity of the cells (39), the potential to evade apoptosis (40) and cell migration (41). Some of the cell surface receptors and ligands which potentially play a role in the CLL microenvironment are shown in **Figure 1.1.**

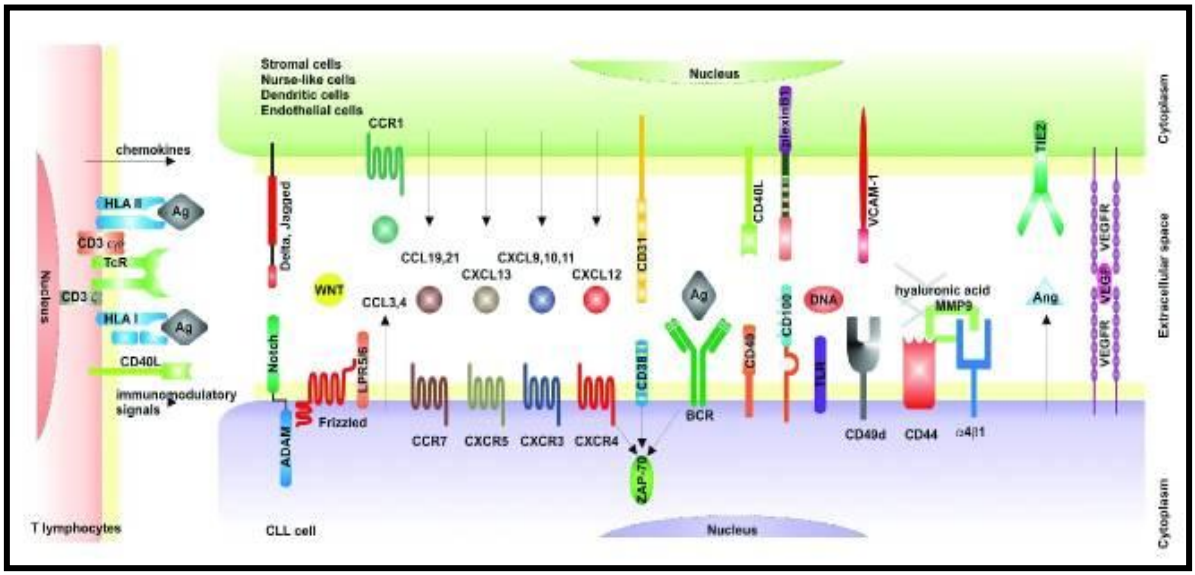

Figure 1.1 The CLL tissue microenvironment. The schema shows cell surface receptors on B CLL cells and their ligands on accessory cells. The CD38, BCR, CXCR4 receptors and adhesion molecules, CD49d, CD44 and CD18, together with MMP9, a cell surface docking molecule, appear to function in close proximity on the B cell membrane (from Deaglio and Malavasi (2)).

1.2.4. CLL cell adhesion and migration

Homing to the secondary lymphoid tissues and the bone marrow is an important feature of the pathophysiology of CLL, but the molecular signals which result in the accumulation of CLL cells in these organs are largely unknown. To enter these organs, circulating cells need to arrest on specific endothelial barriers, to locomote over the endothelial surface toward inter-endothelial junctions and to cross these junctions while resisting disruptive shear forces. These functions depend on the ability of circulating cells to establish dynamic adhesive interactions through their α4 integrins. The VLA-4 integrin dimer is composed of α4β1 (alpha 4 is also known as CD49d and beta 1 is also known as CD29). Other important integrins include α4β7 and the β2 integrins LFA-1 (αLβ2) and Mac-1 (αMβ2). During these interactions, the integrins undergo reversible activation by endothelial-presented chemokines. This is significant as VLA-4 has been shown to be activated by the bone marrow chemokine, CXCL12. It has been shown that CLL cells have impaired migration to lymph nodes and bone marrow (42). CLL cells express lower than expected levels of lymphocyte integrin LFA-1 (43). It is believed that low levels of LFA-1 result in failed transmigration across endothelium expressing ICAM-1, VCAM-1 and CXCL12. This suggests that CLL cells have a reduced capacity to adhere and transmigrate through multiple vascular endothelial beds and home poorly to lymphoid

organs other than spleen. Therefore, integrin blocking could thus be an efficient strategy to prevent circulating CLL cells from reaching pro-survival microenvironments.

Till *et al*. (44), have suggested that the failure of chemokines to induce clustering of adhesion molecules on the CLL cell surface results in defective trans-endothelial migration as CLL cells but not normal B cells are dependent on autocrine VEGF and alpha4beta1 integrin for chemokine-induced motility on and through endothelium. Integrins are responsible for mediating cell-cell or cell-matrix adhesion. The expression of VLA4 (45) adhesion molecule segregates CLL patients into high and low risk categories. The interaction of VLA-4 with VCAM might facilitate B cell activation, firstly by mediating B cell tethering to the endothelial membrane and then by facilitating BCR/ antigen engagement. VLA-4 interaction with VCAM1 appears to synergise with the B cell receptor (BCR) to enhance signalling *via* the BCR.

$CD38^+$ CLL cells have been shown to have greater chemotactic potential compared with $CD38^-$ CLL cells (41). Interactions between CD38 and its ligand CD31 define a genetic signature characterized by modulation of pathways involved in proliferation and migration of CLL cells (46, 47). The differential gene expression profile of $CD38^+$ compared to $CD38^-$ CLL cells highlights differential expression of CD44, CD49d, MMP9 and ZAP70, which are all known to be involved in cell adhesion and motility (41). The activity of CD38 can be blocked using domain-specific monoclonal antibodies (mAbs), which results in weakened CXCL12 responses. This may be partly due to the physical proximity on the cell membrane between CD38 and CXCR4 (the CXCL12 receptor). CD38 is associated with other cell surface molecules, such as ZAP70 amongst other adaptor membrane proteins and is in close proximity to the BCR (41). The BCR/CD38 complex is formed within lipid rafts upon binding of CD38 to CD31 (3). It is thought that stimuli from the microenvironment, such as the interaction between CD38 on CLL cells and CD31 found on endothelial cells (including HMEC-1, which is used in studies described in this Thesis), could activate CD38. This would result in downstream phosphorylation of Syk (spleen tyrosine kinase) and/ or ZAP70, leading to activation of intracellular cascades. These processes would result in turn in actin cytoskeleton changes, polarisation, chemotaxis and trans-endothelial cell migration. These mechanisms are illustrated in **Figure 1.2**.

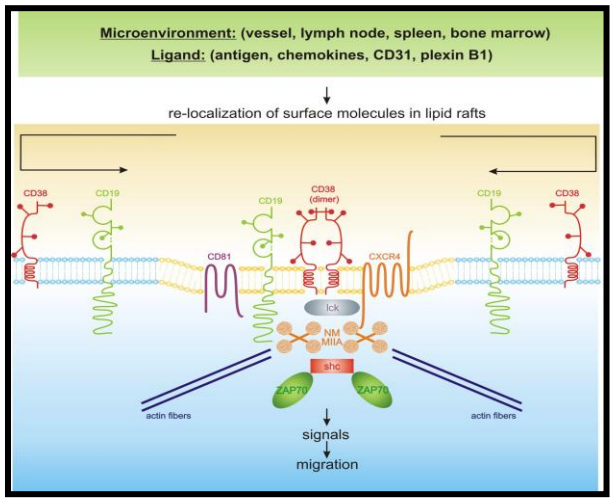

Figure 1.2 The role of CD38/BCR and CXCR4 in cell migration. CD38 engages CD31 ligand found on the endothelial membrane. An intracellular signalling cascade involving ZAP70 and or Syk may feed in to a BCR signalling pathway, leading to actin reorganization and cell motility (from Deaglio *et al*., (3)).

1.2.5. Accessory cells in the microenvironment

Factors that govern the progression of CLL include the proliferative capacity of the cells, the potential to evade apoptosis (48) and cell migration (49). Interactions between CLL cells and the tissue microenvironment contribute to these factors and ultimately to the pathogenesis of the disease (38).

In vivo, CLL cells are in close contact with accessory cells in the microenvironment. Dendritic cells, stromal cells, bone marrow derived endothelial cells, as well as umbilical vein endothelial cells can all confer a survival advantage to CLL cells (50). Data from our laboratory and others have shown that activated $CD4^+$ T cells and endothelial cells can be found in and around proliferation centres in the lymph nodes (51, 52). These data suggest that CLL cell survival is modulated by tumour cell interactions with the microenvironment *in vivo*.

In vitro assays have demonstrated that CLL cell contact with endothelial cells (HMEC-1) in a co-culture system promotes CLL cell survival, whilst CLL cells cultured alone undergo extensive apoptosis (**Figure 1.3**) (1). Therapeutic applications are clear as agents which block pro-survival CLL-endothelial interactions could be developed as new therapies to treat CLL patients.

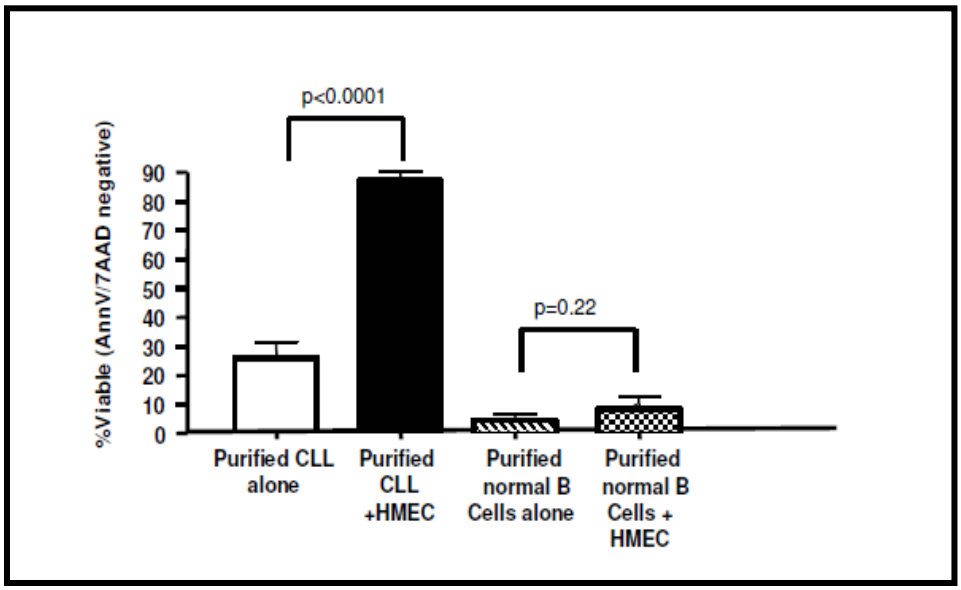

Figure 1.3 Co-culturing CLL cells with HMEC-1 cells protects CLL cells from apoptosis. Normal B cells do not receive a cytoprotective effect (from Buggins *et al*., (1)).

Our understanding of the mechanisms which operate in the CLL microenvironment is incomplete, although it is now appreciated that CLL is a dynamic disorder with significant tumour cell turnover every day with cellular birth rates between 0.1 and 1% of the entire leukaemic clone per day (53) as well as some cells undergoing apoptosis. These data suggest that the CLL clone is continually dying and replenishing itself, particularly in patients with stable WBC counts. Death rates approaching or exceeding 1% of the entire clone per day were reported in some patients, therefore in most patients a significant portion of the clone has the potential to undergo apoptosis. Primary CLL cells are notoriously difficult to culture *in vitro* (54) and establishing models for drug testing has been hindered by the poor survival of these cells. This raises the question as to what signals are provided *in vivo* that enable these cells to survive and proliferate and can it be recapitulated in the laboratory?

Recent work has highlighted the role of certain accessory cells within the tumour microenvironment in the survival and in inducing the proliferation of CLL cells (1, 55, 56). As a result of this, a variety of co-culture systems have been developed to mimic the tumour microenvironment. Tumour proliferation is believed to occur mainly in pseudofollicles, which are specialized structures that contain CLL-cells, T-lymphocytes and stromal cells (57). These structures develop in the lymph nodes, bone marrow and spleen (56, 58). Interactions with T-lymphocytes, the microvasculature, soluble factors and other stromal elements are all thought to play a major role in the survival and expansion of the tumour cells. In keeping with this, lymph node biopsies from CLL patients with aggressive disease contain activated T-lymphocytes. Work from our laboratory demonstrated that

proliferating CLL cells co-localize with activated $CD4^+$ T-cells in lymph node pseudofollicles (56) and ligation of CD40 on CLL cells by its ligand, CD40L (expressed by activated T lymphocytes) has recently been shown to induce responses, including up-regulation of surface markers and induction of chemokines (59). Lymph nodes of CLL patients with aggressive disease also contain large numbers of $CD31^+$ blood vessels (56) and $CD31^+$ nurse-like cells (28, 60). In a recent study, Buggins *et al.* (61) demonstrated that interactions with endothelial cells can promote the survival of CLL cells and induce the expression of CD38 and ITGA4 (CD49d) on the tumour cells (1). The expression of these molecules on CLL cells is associated with aggressive disease and clinical outcome (25, 45, 62). These studies indicate that interactions with accessory cells, such as activated T-lymphocytes and endothelial cells are likely to play a role in sustaining CLL cells *in vivo*. Therefore there is a need to model these interactions *in vitro* in order to define critical molecular interactions that may promote survival and proliferation in this disease.

1.2.6. Effects of cell:cell interactions *vs* effects of soluble factors in CLL culture systems

Various studies have investigated the contribution of cell:cell interactions and soluble factors in the cytoprotective effects of co-culture systems on CLL cells. Since CLL cells associate with one another at sites of tissue involvement in the lymph node, bone marrow and spleen, Pettitt *et al.* (63) postulated that homotypic interactions between the malignant cells might reduce CLL cell apoptosis. Highly pure CLL cell populations were cultured on a non-adherent surface and CLL cell viability was found to increase markedly with the level of crowding at the bottom of the culture vessel. The effect did not require direct cell:cell contact, indicating that cell survival was being regulated in an autocrine fashion by soluble products. Further experiments showed that conditioned media from crowded CLL cells enhanced the survival of autologous non-crowded cells, indicating that at least some of the autocrine survival factors produced by CLL cells could accumulate in the extracellular environment. Co-culture of CLL cells with an excess of autologous fixed cells also enhanced survival of CLL cells. Treatment of fixed cells with neuraminidase, which cleaves glycosidic linkages of neuraminic acids, abrogated the protective effect of cell:cell contact. These data indicate that cell surface specific post-translational modifications are necessary to mediate a cytoprotective effect during co-culture.

Other groups have also demonstrated the importance of the combination of cell:cell interactions and soluble factors for CLL cell survival in culture. In order to elucidate important survival signals acting on CLL cells, Burgess *et al.* (64) cultured primary CLL peripheral blood mononuclear cells (PBMCs) at high density and used antibody arrays to measure the level of 42 cytokines in tissue culture supernatants. These experiments showed that IL-6, IL-8, CXCL2 and CCL2 were highly up-regulated in culture. The addition of either CXCL2 or CCL2 to CLL cultures enhanced CLL cell

survival and antibodies blocking these chemokines reduced survival. Co-culture of CLL cells and PBMC accessory cells separated by transwells provided a similar degree of survival protection compared to high density culture, whereas CLL cells cultured alone died rapidly. It was shown that CCL2 and CXCL2 were produced by CLL cells only when co-cultured with accessory cells, leading the authors to speculate that accessory cells release soluble factors that promote the production of these pro-survival chemokines from CLL cells. Data from Burgess *et al.* (64) suggest that soluble factors may be more important than cell:cell interactions. However, in other co-culture systems it appears that direct cell:cell interactions are more important for CLL cell survival. Maffei *et al*. (65) utilise HUVEC (Human Umbilical Vein Endothelial Cells) in direct cell:cell contact with CLL cells in their co-culture system. CLL cells cultured in conditioned medium from HUVEC did not receive a cytoprotective effect and co-cultures separated by a microporous membrane, or transwell to prevent physical contact died by apoptosis. In agreement with data from Burgess *et al.* (64), Maffei *et al.* (65) tested the contribution of soluble factors present in the co-culture conditioned medium to the observed inhibition of apoptosis in CLL cells, by collecting media from CLL cells co-cultured on HUVEC for 48 hours (co-culture conditioned medium) and then added them to CLL cells cultured alone. An increase in CLL cell viability was observed with the co-culture conditioned medium, suggesting that CLL cells are able to shape their microenvironment.

1.2.7. Antigen stimulation, inflammation and CLL

In CLL and other cancers, a unique microenvironmental organisation is active in the development and survival of malignant cells: chronic inflammation exposes cells to growth factors, newly formed blood vessels provide nutrients and immune tolerance avoids immune-mediated elimination. The concept of the microenvironment being a regulator of CLL proliferation is linked to a role of antigen stimulation through the BCR on the surface of CLL cells. The microenvironment, along with dead or dying CLL cells may act as a source of antigen for CLL B cells. Observations indicating an important role of antigenic pressure in the pathogenesis of CLL include the following: at least half of CLL patients have somatically mutated *IgVH* region, which track the clonal history to *in vivo* BCR activation (25, 66); more than 20% of cases express closely homologous, stereotyped BCR which may recognise auto-antigens or bacterial components (67); and in the *TCL1* transgenic murine model of CLL (68) leukemic immunoglobulins are autoreactive and bind polysaccharides found in bacterial cell membranes. Autoantigens and molecular structures normally involved in scavenging debris, apoptotic cells and pathogenic bacteria appear relevant in triggering and/or facilitating the evolution of at least some CLL clones (69). Inflammatory receptors including Toll-like receptors (TLR) can be engaged concomitantly with the BCR, therefore TLR may also play a role in BCR co-stimulation of CLL cells. It was recently shown that bacterial lipopeptides protect CLL cells from spontaneous apoptosis mediated by TLR signalling (70). The relationship between antigen

stimulation/inflammation and CLL pathogenesis is currently highly contentious and is likely to be an area of intense further research.

Other evidence which supports the proposition that CLL cells are continuously exposed to antigen *in vivo* is the N-glycosylation pattern of surface IgM (sIgM). Krysov *et al.* (71) reported that sIgM exists in two distinct forms with different N-glycosylation patterns in the mu-constant region. One glycoform is similar to that found on normal B cells which bear mature complex glycans common to other cell-surface glycoproteins. The other glycoform is immature, mannosylated and more characteristic of mu chains usually found in the endoplasmic reticulum. Unmutated CLL (U-CLL) were found to expresses a higher proportion of mannosylated surface mu chains than mutated CLL, whilst normal B cells express only the mature glycoform. Persistent engagement of sIgM on normal B cells can induce the expression of the immature form and this suggests that glycan modification is a consequence of antigen exposure. CLL cells were also shown to revert to the mature form after incubation *in vitro*, suggesting the source of antigen had been removed. These findings support the concept that CLL cells are continuously exposed to antigen *in vivo*.

Evidence from numerous groups supports a role for antigen signalling and interactions with the microenvironment in the development and subsequent progression of CLL (72, 73). It has been shown that CLL cells resist apoptosis (74), but the mechanisms which connect BCR signalling to apoptotic resistance are not fully understood. Paterson *et al.* (75) investigated the downstream regulators of signalling pathways particularly the role of the pro-apoptotic molecule BIM and identified two major isoforms, BIM_{EL} and BIM_{L}, which undergo phosphorylation in CLL cells stimulated with anti-IgM. Co-culture with HK (a follicular dendritic cell line) promoted BIM_{EL} phosphorylation, suggesting that BIM may coordinate microenvironment and antigen-mediated survival signals. Studies such as this provide new insights into BCR signalling and its relevance to CLL providing therapeutic targets (76).

1.2.8. The use of endothelial cells in co-culture model systems

An inflammatory microenvironment will have effects not only on CLL cells but also on other accessory cells, including endothelial cells. This should be considered when using a CLL, endothelial co-culture system in order that a suitable endothelial cell is chosen. Inflammatory responses are determined by the response of the vascular endothelium to extracellular injury. Stimuli trigger gene expression programs, which result in transcription of genes encoding pro- and anti-inflammatory proteins. These in turn guide attraction and interaction with leukocytes, affect vascular permeability and determine the composition of infiltrating leukocytes. For example, the

stimulation of endothelium markedly enhances CLL interaction with endothelium (77) *via* integrin alpha 4 beta 1 and VCAM-1.

The diversity of the vascular bed is determined by the endothelial cells lining the inner vessel surface. The most frequently used models are micro-vascular (*e.g.* HMEC-1) and macro-vascular (*e.g.* HUVEC) cells. Results of studies based on one endothelial cell type are often assumed uncritically to apply to other endothelial cell types (78-80). However, E-selectin is frequently reported to be expressed differentially on HUVEC and HMEC-1 (81) and this highlights the potential for functional differences in models using particular endothelial cell types.

It is generally assumed that the responses of different endothelial cells to distinct inflammatory stimuli are comparable and this may not be a reasonable assumption. One study compared the effect of TNF-α on blood brain barrier *vs* macro-vascular endothelial cells (HUVEC). Studies have shown differences in basal gene expression profiles of micro- and macro-vascular endothelial cells. TNF-α stimulation (2ng/ml for 5 hours) induces distinct gene expression programs in micro-vascular and macro-vascular human endothelial cells (82). However it is unclear whether differences in expression after TNF-α treatment were due to differences in basal gene expression or a different response to the stimulus. In this study, the responses of a micro-vascular endothelial cell HMEC-1 and a macro-vascular endothelial cell type HUVEC to TNF-α were analysed by microarray analysis and compared. Many genes were comparably induced in HMEC-1 and HUVEC by TNF-α treatment but around half of the 86 genes induced by TNF-α treatment were specific for HMEC-1 or HUVEC. The genes which differed encode chemokines, cytokines and cell surface molecules. These specific effects were restricted to subtypes of endothelial cells rather than in cell type dependent regulation and included mRNA encoding proteins involved in cell differentiation or cell cycle control. A confounding factor in the design of this experiment was the use of one cell line and one primary cell when the vital comparison was between gene expressions of micro- and macro-vascular endothelial cells. In addition, the TNF-α stimulation used in this study (2ng/ml) is of a higher concentration than cytokine concentrations produced by CLL cells in culture (117pg/ml), and higher still than levels found in sera (39.6pg/ml) (83).

HMEC-1 is the first immortalized human micro-vascular endothelial cell line that retains the morphologic, phenotypic and functional characteristics of normal human micro-vascular endothelial cells. HMEC-1 cells express cell surface molecules typically associated with endothelial cells, including CD31 (PECAM-1) and CD36. CD31 is the only known ligand for CD38 (84), which is expressed by CLL cells and is a negative prognostic marker. The HMEC-1 cells also express the cell adhesion molecules ICAM-1 and CD44 and following stimulation with IFN-gamma, express major histocompatibility complex class II antigens. HMEC-1 cells also specifically bind lymphocytes in cell

adhesion assays (85). The studies described above demonstrate that because of the diversity of endothelial cells that are used in *in vitro* models, it may be more important to determine which effects occur in common when different endothelial cell lines are used in co-culture systems. Endothelial models have been used to study mechanisms in other haematological malignancies, such as the importance of VCAM-1 in Hairy Cell Leukaemia (HCL) (86) and cell adhesion-mediated immune resistance against cytotoxic T cell lysis in Multiple Myeloma (MM) (87).

1.2.9. The bi-directional effect of co-culture

Co-culture has effects on the accessory cells as well as on the CLL cells. For example, Plander *et al*. (88) showed that CLL cells have an anti-apoptotic effect on the bone marrow stroma. Briefly, CLL cells were purified by flow cytometric sorting, which removed cells with glycophorin A, CD14, CD56 and CD3 on their cell surface. Purified CLL cells were co-cultured with allogeneic, normal bone marrow stromal cells (BMSC) with or without CD40L, both resulted in CLL cells being rescued from apoptosis. The CLL cells up-regulated the expression of CD18 and CD49d, which are ligands for adhesion molecules found on BMSC. This may reflect what occurs *in vivo* since CLL cells from bone marrow aspirates express higher levels of CD49d compared with CLL cells from peripheral blood samples. The BMSC themselves increased secretion of IL-6 and IL-8 and up-regulated expression of ICAM-1 and CD40L mRNA. These studies demonstrate that CLL cells produce cytokines which can shape the microenvironment, resulting in bi-directional signalling between the different cell types in the co-culture system.

1.2.10. Effect of co-culture on different CLL subtypes

Co-cultures using various model systems provide the CLL cells with a cytoprotective effect when compared with liquid culture, regardless of the clinical features of the disease (16). However, there are some differences between CLL cells from patients with different clinical features when these cells are cultured alone *in vitro.* The percentage of CLL cells undergoing spontaneous apoptosis *in vitro* appears to be higher for samples from patients with unmutated (UM) *IGHV* compared with mutated (M) *IGHV*, and co-culture achieves a 2.2-fold increase in relative viability in M-CLL compared with a 6.1-fold increase in UM-CLL (65) with co-culture. This suggests that cells from UM-CLL patients are more dependent on the microenvironment for support. However, it is unknown whether the same pathways are up-regulated as a result of co-culture or if there is a completely different mechanism in each of these CLL subtypes. Other studies have highlighted the different functions of clinically relevant proteins in CLL. For example, CD49d is thought to have important roles in homing (89). Expression of CD49d and CD38 are higher on CLL cells from the bone marrow

compared to those in the peripheral blood and die quicker in liquid culture. This may suggest that these cells are more dependent on the tumour microenvironment for survival than those cells found in the peripheral blood.

1.2.11. Targeted therapies for CLL

Targeted therapies are medications which preferentially block the growth of cancer cells by interfering with specific molecules needed for carcinogenesis and tumour growth (90), rather than affecting all dividing cells as with traditional chemotherapy. The paradigms for such therapies are the development and use of all-trans retinoic acid (ATRA) for treating patients with acute promyelocytic leukaemia (APML) (91), Gleevec/Imatinib Mesylate for treating patients with Chronic Myeloid Leukaemia (CML) (92) and monoclonal antibody therapies for HER2 positive breast cancers (93). Traditionally, CLL was treated with chemotherapy including purine analogues such as fludarabine (F) and alkylating agents such as cyclophosphamide (C). Recent advances in the treatment of CLL have seen the introduction of monoclonal antibody therapies, including alemtuzumab (directed against CD52), rituximab (R) and ofatumumab (directed against CD20). Modern regimens combine chemotherapy with monoclonal antibodies such as FC, FR, FCR and CHOP (cyclophosphamide, doxorubicin- anthracycline, vincristine- mitotic inhibitor and prednisolone-corticosteroid), which produce a synergistic effect on CLL cells (94).

CLL is currently incurable even with combination therapies which have improved response rates but not overall survival, because the disease becomes resistant to the therapy being used. Therefore, novel therapies. Targeted therapies are particularly attractive prospects for the treatment of patients with complicated disease, including 17p deletion, *TP53* mutation, fludarabine-refractory CLL and those with suboptimal response to treatment. A personalized approach to treatment is likely to be important for CLL patients given the heterogeneity in both clinical manifestations and prognosis of the disease. For example, a Bruton's Tyrosine Kinase (BTK) inhibitor has been used to treat patients with relapsed disease, leading to a high frequency of durable remissions (95). It is hoped that a better understanding of novel aspects of CLL biology can lead to the rational design of other agents and combinations of agents which specifically target the pathophysiology of the disease, resulting in more effective and less toxic therapies. Some targeted therapies currently under investigation are described below in **Table 1.2**.

Treatment	Class	Rationale	Mechanism	Reference
Navitoclax (ABT-263)	Bcl-2 inhibitor	The anti-apoptotic protein Bcl-2 is over expressed in CLL and results in chemo-resistance (96).	BH3 mimetic, small molecule binds with high affinity to Bcl-2, BclXL and BclW, promoting apoptosis.	(97, 98)
Flavopiridol	Cyclin-dependent kinase inhibitor	Cyclin D1 is over expressed in a subset of CLL (99), cyclin D2 is over expressed in proliferation centres in CLL (100) reducing threshold for cell cycle checkpoints.	Caspase-3 activation, broad acting inhibitor including off target effects. Action is p53 independent, therefore could be used in high-risk cytogenetic features such as del(17p13).	(101) (102)
Ibrutinib	Bruton's tyrosine kinase (Btk) inhibitor	BCR signalling provides growth signals to CLL cells (72, 103).	Inhibition of Btk results in transient lymphoctosis, usually associated with a nodal response (104)	(105)
CAL-101 (GS-1101)	PI3Kδ inhibitor	In CLL, the PI3K pathway is constitutively activated and dependent on PI3Kδ (106).	Isoform-selective inhibitor of PI3Kδ that inhibits PI3K signalling and induces apoptosis of CLL cells.	(107)
Fostamatinib	Syk inhibitor	BCR signalling activates Syk, which leads to downstream signalling promoting cell survival and growth (108).	Blocks BCR signalling, inhibiting tumour growth (109).	(110)

Panobinostat (LBH589), Suberoylanilide hydroxamic acid (SAHA)	Histone deacetylase (HDAC) inhibitor	Elevated HDAC enzyme levels have been reported in CLL (111).	Altering histone modifications can restore apoptotic pathways, allowing CLL cells to undergo apoptosis (112, 113).	(114)
NVP-AUY922-AG	HSP90 inhibitor	Hsp90 expression data in CLL cells is equivocal (115, 116). However, CLL cells over express several Hsp90 client proteins making CLL potentially susceptible to Hsp90 inhibition.	Inhibits NF–κB signalling, overcomes microenvironmental cytoprotection and is highly synergistic with fludarabine (117, 118).	(117)
Carfilzomib	Proteasome inhibitor	Proteasomes mediate degradation of regulatory proteins that are aberrantly active in CLL *e.g.* p53, Bcl-2, and NF-κB families.	Cytotoxicity is caspase-dependent and p53 independent. CFZ promotes atypical activation of NF-κB.	(119)

Table 1.2 Targeted therapies currently under investigation in CLL.

Other potential therapies target accessory cells rather than the CLL cells directly, for example, the immunomodulatory drug lenalidomide. CLL cells like malignant cells in other cancers modify the immune microenvironment to block effective host anti-tumour responses. CD4 and CD8 T cells from patients with CLL exhibit exhaustion (120) and globally impaired LFA-1-mediated migration (43) and this defect is mediated by direct tumour cell contact. Treatment with lenalidomide was shown to reverse this T-cell defect, rescuing adhesion and motility function (43).

1.3 Systems Biology and Bioinformatics

Denis Noble, one of the pioneers of Systems Biology has said that “Systems biology...is about putting together rather than taking apart, integration rather than reduction...” and “ It means changing our philosophy, in the full sense of the term" (121).

Historically, scientific research has focussed on individual proteins or protein complexes and how they interact in a static time-frame in a particular cellular localisation. This is known as a reductionist approach and provides little information about the dynamic and temporal context in which these protein:protein interactions may occur. Systems Biology combines a series of overlapping concepts to integrate complex data from a number of diverse data sources and disciplines in a holistic approach (122) and shown in **Figure 1.4**. For the cellular proteome, analyses are carried out at various levels from individual protein:protein interactions, to networks of potential protein:protein interactions right through to the dynamics of these interactions and regulation of their encoding mRNA and gene expression. These studies may incorporate wet-laboratory experiments, including proteomics using high-content mass spectrometry, microarrays to monitor the expression levels of mRNA transcripts or RNA deep sequencing to analyse transcriptomes (123). These datasets are interrogated using ever more sophisticated bioinformatics methods. The application of these techniques results in a greatly increased volume of complex, interconnected data and an understanding of the cell as a complex system (124)(125). Once a biological system can be modelled, the ultimate aim of Systems Biology is to *predict* how the molecular behaviour of a cell will change in a certain environment, in response to a particular stimulus or when a component protein is removed or altered, a situation that occurs in cancer (125).

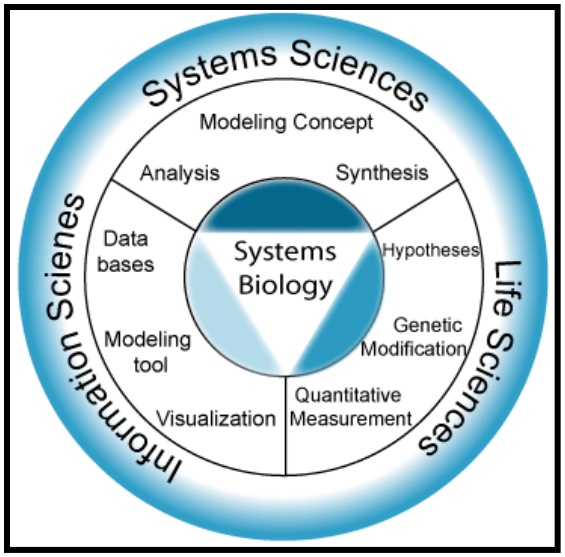

Figure 1.4 Schematic of Systems Biology. The interface between life sciences, physical sciences, computer science and mathematics (from http://www.sysbio.de/).

The concept of Systems Biology is being utilised to analyse and make intelligent use of high throughput, data-rich, 'omics' experiments in order to better understand the complexity of cell functions, which may potentially lead to improvements in our knowledge of health and disease. One aim of my Thesis is to apply Systems Biology approaches to obtain a better understanding of the behaviour of CLL cells as a complex biological system (126, 127).

Bioinformatics, the application of statistics and computer science to the study of biology, is at the centre of Systems Biology. Bioinformatics requires databases, algorithms, computational and statistical techniques to solve problems arising from the management and analysis of biological data, such as analyses of gene expression datasets and of high throughput mass spectrometry experiments. Large scale post genomic experiments require powerful tools in order to analyse and make sense of the high throughput data. Bioconductor (www.bioconductor.org) (128) is open source, open development software which uses the R statistical programming language to handle and analyse large data sets. It may be used for microarray analysis, high throughput datasets (such as mass spectrometry), sequence data and annotation of data (combining and converting multiple formats).

1.3.1. Protein:protein interaction networks

There is currently a great deal of interest in protein:protein interaction networks. Their appeal lies in their ability to integrate multiple data types, often from high throughput screens, which may be interrogated to identify novel targets for research and to provide a visual display of such data. Networks may be built based on high throughput screens such as yeast two-hybrid assays (which are able to detect a binary interaction between proteins). This technique has come under much criticism due to its high false positive rate (129). Other predictive protein: protein interaction networks are built on the knowledge of interactions of proteins in other organisms, which is then used to predict data on orthologous human proteins. Additional data on domain-domain interactions, co-occurrence and co-expression of genes can also be used. OPHID (130) and HomoMINT (131) both integrate data from protein:protein interactions in other organisms with predicted data on orthologous human proteins. A third method for creating networks employs data-mining, whereby protein interaction networks are based on literature searches. These databases may be created using language processing algorithms or curated manually, for example the GeneGO platform.

A combination of data are used by HumanNet (132), which is a probabilistic functional gene network created by our collaborators, Professor Marcotte's laboratory at the University of Texas at Austin which uses modified Bayesian integration of literature mining along with many different data types from different organisms. Each data type is weighted according to how well it is able to link genes

known to function together in *H.Sapiens.* The weighting of different types of data reduces the effect of bias upon a network and the different types of evidence used by HumanNet are shown in **Table 1.3.**

Networks generated by such algorithms may be used to create predictions of important protein:protein interactions in order to provide a guide or focus for a research project. It is now possible to begin the search for novel potential regulators of an important pathway or process before beginning any wet-laboratory experiments. It is widely regarded that the value of such networks lies in their selective use and validation of components predicted to be of functional importance by wet-laboratory experiments. A good example of this is the application of the Phenolog method (133). This method is based on the fact that many functionally important genes are conserved in different species. Phenologs are identified by mapping orthologous genes between species in which deletion of the orthologues leads to characteristic phenotypes in each organism. Orthologues are genes in different species which are descended from a single gene in an ancestral organism (134). However, the phenotype produced by gene deletion may be different in each organism and while proteins encoded by conserved groups of genes may still work together, they lead to different phenotypes in different species. A breakthrough study using the Phenolog approach was reported in the New York Times, where the method predicted a yeast model for mammalian angiogenesis (133) and recently, this method was used in a project that re-positioned/re-purposed thiabendazole, an anti-fungal drug to inhibit neo-vascularisation (135).

Co-citation of worm gene
Co-expression among worm genes
Worm genetic interactions
Literature curated worm protein physical interactions
High-throughput yeast 2-hybrid assays among worm genes
Fly protein physical interactions
Co-citation of human genes
Co-expression among human genes
Co-occurrence of domains among human proteins
Gene neighbourhoods of bacterial and archaeal orthologues of human genes

Literature curated human protein physical interactions
Human protein complexes from affinity purification/mass spectrometry
Co-inheritance of bacterial and archaeal orthologues of human genes
High-throughput yeast 2-hybrid assays among human genes
Co-citation of yeast genes
Co-expression among yeast genes
Yeast genetic interactions
Literature curated yeast protein physical interactions
Yeast protein complexes from affinity purification/mass spectrometry
Yeast protein interactions inferred from tertiary structures of complexes
High-throughput yeast 2-hybrid assays among yeast.

Table 1.3 Evidence used to construct HumanNet. A probabilistic functional gene network (http://www.functionalnet.org/humannet/.

1.3.2. Gene expression profiling

The principle of gene expression profiling is based on complimentary base pairing or hybridisation (see: http://www.ncbi.nlm.nih.gov/About/primer/microarrays.html). DNA probes are manufactured so that in theory each probe sequence only occurs once in the genome. Thousands of different probes are arrayed or synthesised on a surface and typically each array (chip) contains several probes corresponding to each transcript of every protein encoding gene. RNA is extracted from the cell of interest, copied to cDNA or cRNA and washed over the probe sets. If the sequence of RNA bases matches a probe, the sample will bind to the probe. A fluorescent dye is then washed over the probe and laser light is used to visualise the fluorescent stain. Fluorescence is therefore correlated with gene expression. Fluorescence is then quantified against a background fluorescence reading from mismatched probes using a statistical package, such as Bioconductor.

Important points to consider when analysing gene expression data include the promiscuity of the probe set being used. In practice, a specific probe will not always bind to a unique sequence within the genome. Promiscuous probes bind multiple sequences within the genome and therefore these

data cannot be used in further analyses. It is good practice to check the specificity of each probe used by running a Basic Local Alignment Tool (BLAST) search.

A bioinformatic package such as Biomart (www.biomart.org, a data integration system) may then be used in further analyses of the data created. Biomart is used to convert between proprietary probes (such as Affymetrix®) and their corresponding gene identifiers (IDs), such as Ensembl IDs or Entrez IDs. Depending on the downstream use of these data, multiple identifiers may be required. This time saving application is one of the strengths of using a high throughput programme. Care must be taken when converting between identifiers as these do not map uniquely. One Affymetrix® probe may map to several Ensembl identifiers, which may map to several Entrez identifiers. Therefore it is good practice to verify ID conversions independently of automated versions, for example using http://www.genecards.org/. Genecards is a searchable, integrated database of human genes that provides concise genomic related information on all known and predicted human genes using standard nomenclature and approved gene symbols.

1.3.3. Gene ontology annotation

The Gene Ontology (GO) project is a bioinformatics initiative which aims to 'standardise the representation of gene and gene product attributes across species and databases' (136). A controlled vocabulary is used to describe gene products and gene characteristics. GO annotations include the gene product identifier, a GO term, the reference used to make the GO annotation such as a journal article, an evidence code which describes the type of evidence upon which the annotation was based and the date and creator of the annotation. GO terms may refer to a process (GOP), a function (GOF) or a cellular component (GOC). For example, using the appropriate GO terms, a search may be carried out for gene products with a particular cellular localisation (through GOC) and which are involved in a certain function (GOF). GO is structured as a 'tree' with parent and child terms (see **Figure 1.5** for an illustration).

GO terms should be treated with caution. Incomplete annotation can be misleading and some gene products remain un-annotated. Some GO terms have extremely broad definitions, rendering them of little use when performing a very specific search.

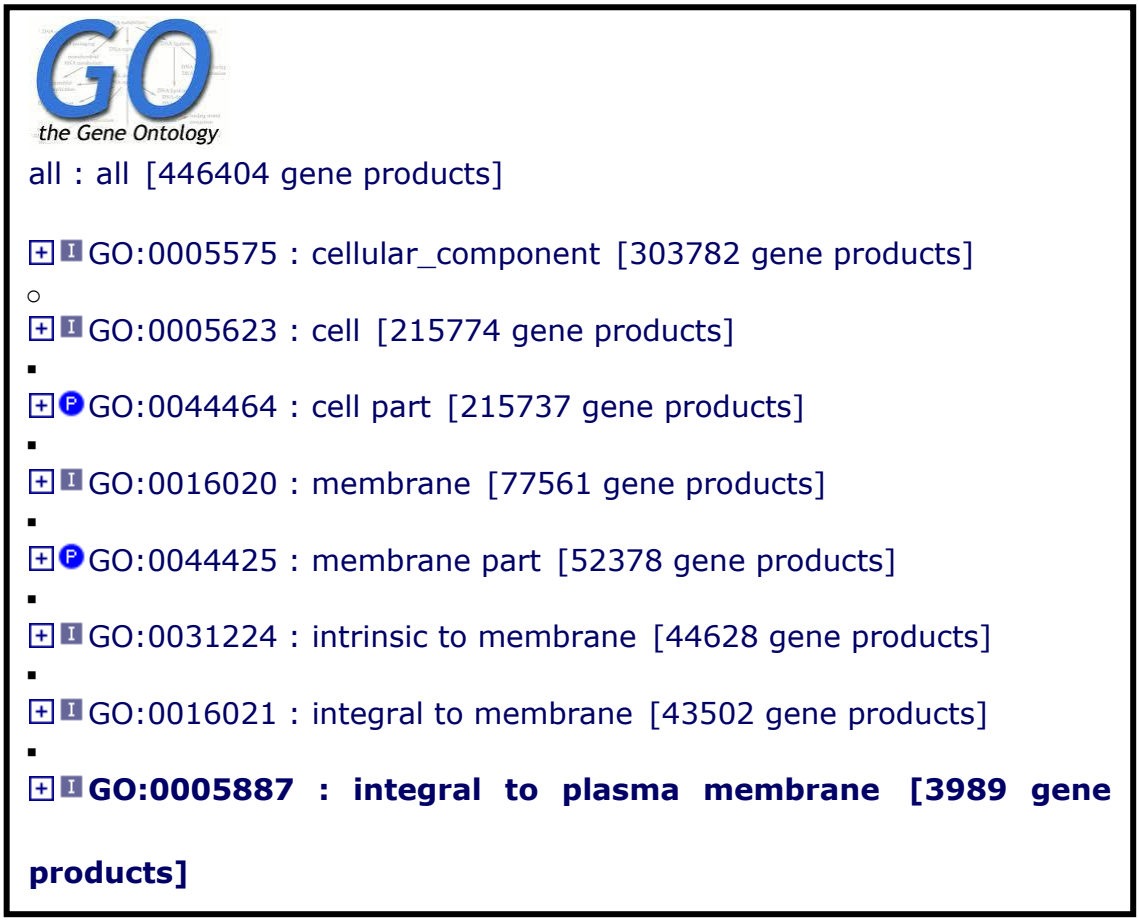

Figure 1.5 GO ontology hierarchy. An example of a search for gene products 'integral to plasma membrane'. Each layer of description results in a smaller number of gene products **(**http://www.geneontology.org/).

1.3.4. Cellular localisation prediction algorithms

Other bioinformatic tools use statistical analyses in order to make predictions about gene products. Examples of this include the prediction of transmembrane and secreted proteins. Statistical methods for predicting transmembrane helices use hydrophobicity analysis. A stretch of about twenty amino acids may indicate that these amino acids are part of an alpha helix spanning a lipid bilayer, which is composed of hydrophobic fatty acids. Hydrophilic amino acids following this sequence are likely to be in contact with aqueous environments and are therefore likely to be present either on the outer surface of the cell or as part of an intracellular tail. A second common analysis used in the prediction of transmembrane helices is the abundance of positively charged amino acids on the cytoplasmic side of the membrane. This is known as the 'postive inside rule' (137).

Transmembrane Hidden Markov Model (TMHMM http://www.cbs.dtu.dk/services/TMHMM/) is perhaps the most well-known transmembrane prediction algorithm. It combines hydrophobicity, charge bias, helix length and grammatical constraints to correctly predict 97-98% of transmembrane helices (138). It can also discriminate between soluble and membrane proteins at 99% accuracy, although this accuracy drops when a signal peptide is present. TMHMM predicts that 20-30% of all genes in most genomes encode membrane proteins. However, membrane proteins are notoriously difficult to characterise in wet–laboratory experiments due to their structure, requiring solubilisation and stabilisation. This means that transmembrane proteins are grossly underrepresented in databases of known proteins, especially in databases with structural information.

Other subcellular localisations of proteins can be predicted on the basis of amino acid sequence. Tools include:

TaretP (http://www.cbs.dtu.dk/services/TargetP) for secretory peptides, mitochondrial targeting peptides and chloroplast transit peptides,

SignalP3.0 (http://www.cbs.dtu.dk/services/SignalP) for secretory signal peptides in eukaryotes, Gram-negative and Gram-positive bacteria,

big-Pi (http://mendel.imp.ac.at/sat/gpi/gpi_server.html) for GPI membrane anchors

PredictNLS (http://www.predictprotein.org) and **NucPred** (http://www.sbc.su.se/~maccallr/nucpred) to predict nuclear localisation signals in eukaryotes.

When screening for unknown transmembrane or signal peptides, a combination of different algorithms should be used to maximise the number of proteins identified.

1.3.5. Membrane protein enrichment

Low abundance membrane proteins must be enriched from cell extracts before they can be analysed using proteomic tools. Methods available for enriching membrane proteins from cell extracts include subcellular fractionation, de-lipidation and affinity purification.

The stability of the non-covalent interaction between avidin and biotin (*K*d 10^{-15} M) has been exploited in many applications and is commonly used in chemistry and biology. Methods for modification of molecules with biotin have been used to allow protein recovery, immobilization and detection using avidin-based reagents. The study of cell surface proteins is one major area which has greatly benefited from this application. A reactive ester such as an N-hydroxysuccinamide (NHS) group is used to covalently link biotin to the molecule of interest. NHS undergoes a nucleophilic substitution reaction in the presence of primary amines, such as the amino group in

exposed lysine residues in proteins. The presence of a charged group in sulfo-NHS-biotin renders the reagent membrane impermeable, allowing the labelling of cell surface proteins only. This type of reagent has been used in a number of studies, including the isolation of viruses to increase effective titre (139) and analyses of individual cell surface proteins and identification of components of the cell surface proteome (140). The use of cleavable spacers has been introduced to some reagents with the aim of facilitating the release of biotinylated proteins after capture on immobilized avidin. The most common cleavable group is a disulfide bridge that can be broken by reducing agents such as dithiothreitol (DTT). **Figure 1.6** schematically summarises one method of concentrating membrane proteins that utilises the biotin-avidin interaction.

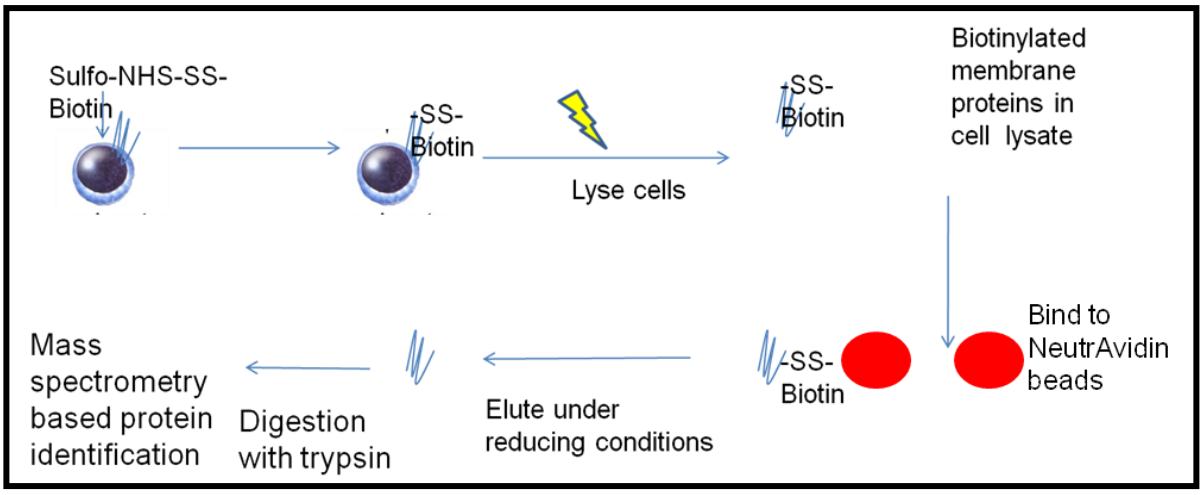

Figure 1.6 Schematic of biotinylation and extraction of cell surface proteins.

1.3.6. Protein identification by mass spectrometry

Mass spectrometry (MS) is a powerful tool in proteomics and can be used to identify proteins present in a complex mixture, such as those produced using a biotin-avidin capture system (see **Figure 1.6**). Two experimental approaches are commonly used in MS experiments. In the first method, biochemical fractionation methods are used to separate individual proteins present in cell lysates prior to MS analysis. Complex protein mixtures of a cell lysate may be separated by gel electrophoresis, such as SDS-PAGE or 2-dimensional (2-D) gels, which separate proteins based on size (SDS-PAGE) and pI (isoelectric focussing, IEF). Bands or spots at the size and pI of interest or bands/spots that are different in experimental compared with control samples are cut out form the gel. The protein(s) contained in the gel slice are digested by a proteolytic enzyme such as trypsin using an 'in gel' digestion method. The tryptic peptides are then separated by High Performance Liquid Chromatography (HPLC) (usually referred to as liquid chromatography (LC)) to reduce sample complexity by separating the peptides prior to MS analysis. The tryptic peptides are

reconstituted in a mobile phase, which is forced at high pressure through a column that contains the stationary phase particles with a particular surface chemistry (this varies depending on the type of HPLC being performed). The motion of the peptides in the mobile phase is retarded by interactions with the stationary phase. The time taken for an analyte to elute from the column is the 'retention time'. Analytes with different properties have different retention times, allowing unknown analytes to be separated.

The second common method used in MS experiments involves the digestion of all proteins in a cell lysate with trypsin, in a process known as 'in solution' digestion, releasing a complex mixture of thousands of peptides. The peptides can then be separated by one-dimensional LC, such as a C18 reverse phase column, or by two-dimensional LC, for instance employing a cation exchange column followed by a C18 column. The separated peptide mixtures are then analysed by MS.

In the next stage of MS analysis, the tryptic peptides are ionised in the MS machine, causing fragmentation of peptides into charged ions (141). Two common ionization techniques are Electrospray Ionisation (ESI) or Matrix Assisted Laser Desorption Ionisation (MALDI). These peptide ions are then introduced into the mass analyser using an electric field and the ions are sorted according to their mass to charge ratio (m/z). An identification of the amino acid composition of peptides analysed helps increase the number of identifications made. In effect, this is peptide sequencing and can be achieved by selecting peptides individually in the ion trap and fragmenting each one by collision-induced dissociation (CID). In CID, the molecular ions are accelerated by an electrical potential and allowed to collide with molecules of either helium, nitrogen or argon. The collision results in bond breakage and the fragment ions can then be analysed to give an accurate amino acid composition.

Two Orbitrap mass spectrometers were used in the study, the LTQ-Orbitrap (142) and the Velos Pro-Orbitrap (143). The LTQ quadrupole linear ion trap instrument (Thermo Finnegan) collects and traps ions in a plane between electrodes. When a specified number of ions have been collected (or after a defined time) the ions are ejected from the trap. In the LTQ instrument, the rate of increase in the ejection voltage allows different m/z ions to be ejected at different times and analysed, which creates an MS spectrum. The ions ejected from the LTQ are then directed into the Orbitrap, which increases the accuracy of the m/z measurements. The second mass spectrometer used in this study was a Velos Pro-Orbitrap machine. This uses an improved Velos ion trap as the front end rather than the LTQ described above.

Ions trapped by the LTQ or the Velos are subsequently injected into the Orbitrap. The Orbitrap uses an oscillating electrostatic field between two non-linear electrodes, ions oscillate around the central electrode and the frequency of oscillation is inversely proportional to the square root of the m/z. This

machine has a high mass accuracy (1ppm for the Velos Pro-Orbitrap), improved resolution and dynamic range over previous ion trap technologies and is easy to use for routine analyses and a wide dynamic range (144, 145).

Trypsin digestion of proteins results in cleavage at specific amino acids, namely at the carboxyl side of lysine or arginine, except when these residues are followed by proline. Using a process known as peptide mass fingerprinting (146), the spectra produced are compared against an online database of peptides generated from a predicted digest of all known proteins for a given organism (*e.g.* using MASCOT or SEQUEST (147)).

1.3.7. Quantification of protein expression in mass spectrometry

Quantification of differences in protein abundance by mass spectrometry is technically challenging. The MS signal intensity or peak height of a peptide does not correlate directly with the abundance of the protein. Different peptides have a different tendency to ionise due to their different chemical structures and properties. The chemical environment also affects ioinzation efficiency (148), such as the buffers used to dissolve the peptides before MS analysis. Therefore it is possible that a low abundance protein may be interpreted as high abundance because some of its peptides are more easily ionisable and so detected more frequently than those derived from other proteins. For these reasons, until recently only a fraction of the thousands of MS studies published provided a comprehensive quantification by MS and reports quantifying proteins in mixtures reliably are even more limited (149).

Advances are being made towards quantitative proteomics including relative and absolute quantification. One technique used for relative quantification requires *in vivo* stable isotope labelling prior to MS analysis. For example, 'Stable Isotope Labelling with Amino acids in cell Culture' (SILAC) (150). In this approach, two different samples are cultured in media containing amino acids labelled with different isotopes, such as 'heavy' ^{13}C-arginine and 'light' ^{12}C-arginine respectively. The cells in each culture incorporate the ^{13}C- or ^{12}C-arginine into all newly-synthesised proteins. Protein lysates from these cultures can be pooled and analysed together by MS. MS analysis can distinguish the same peptides from the two different cell cultures, because they will have different m/z values and the different characteristic MS spectra can be used for relative quantification (151).

Another *in vitro* labelling technique is iTRAQ (isobaric Tags for Relative and Absolute Quantification) (152). This method uses isotope-coded covalent tags, each of which fragments in the mass spectrometer to produce a distinct fragmentation m/z pattern. The ratio of peak intensities for a given protein can then be used for relative quantification between experimental conditions. The labelling

allows a 4-plex or 8-plex format, enabling 4 or 8 experimental samples to be compared in the same MS run. As an example, the iTRAQ method has been used to investigate the effects of Imatinib treatment on CML cells and identify potential protein markers for response to Imatinib treatment (153).

Label-free quantification methods are also advancing protein quantification in mass spectrometry. This approach harnesses statistical methods such as Absolute Protein Expression (APEX), which was developed by Professor Marcotte, our collaborator at the University of Texas at Austin (154). Previous label-free quantification approaches had concentrated on measuring the peak height of the peptide and did not take into account other information such as the peptide count. However, the peptide count can be used to provide further information as larger proteins will contribute more peptides than smaller proteins to an analysis (154). Therefore, the probability of observing a peptide from a larger protein is higher than that of the smaller protein and this may lead to an overestimation of the abundance of large proteins without normalisation. The APEX method includes a correction factor for each protein, which accounts for variable peptide detection by MS techniques. It corrects the observed peptide count (sampling depth) which has been biased by the MS technique by a pre-calculated estimate of the number of peptides which would be generated if a bias did not occur. This APEX tool provides a Z score for the identification of statistically significant differences in protein abundances between samples and has been successfully applied in our recent study, which quantifies changes in chromatin and nuclear matrix-bound proteins in human T cells during entry into the first cell cycle from quiescence (155).

The aim of my PhD project was to apply Systems Biology methods to answer biological questions about interactions between CLL cells and the tumour microenvironment, including: (i) the effects of co-culture model systems on primary CLL cell viability and phenotype, (ii) investigating the transcriptional effects of endothelial cell co-culture on primary CLL cells, particularly control mechanisms responsible for co-ordinated gene expression and (iii) a systematic identification of primary CLL cell surface proteins. Increasing our knowledge of CLL biology provides more potential targets for rational therapeutic intervention.

The study in Chapter 3 and recently published in the *British Journal of Haematology*, focuses on the effects of co-culture on primary CLL cells. CLL cells undergo apoptosis when cultured in the laboratory away from the supportive tumour microenvironment found in the body. In Chapter 3 a direct comparison of the effects of different co-culture systems on primary CLL cells is made. The effects of interactions between primary CLL cells and endothelial cells (HMEC-1), and with mouse fibroblasts expressing the human proteins CD40L (expressed on activated T cells) or CD31 (expressed on endothelial cells), are compared using cell based assays.

The study in Chapter 4 examines the mRNA changes which occur when cells from CLL patients are co-cultured with an endothelial cell line, HMEC-1 and the primary endothelial cells, HDBEC. Based on the observation that both co-culture systems improved CLL cell viability and induced a similar phenotype in CLL cells, I sought to determine whether there were mRNA changes which occurred in the cells from all CLL patients as a result of co-culture. The rationale behind these biofinformatic analyses was that any mRNA or pathways up-regulated in both systems may represent cellular mechanisms which the CLL cells have become addicted to and rely upon for survival and therefore present targets for intervention.

The final study in Chapter 5 focuses on identifying cell surface proteins on primary CLL cells which may be required for cell:cell contacts made in the tumour microenvironment and therefore may be important to receive signals which can result in the transcriptional changes observed in Chapter 4. I developed methods to identify cell surface proteins by MS using HeLa cells, as a test cell. I then applied these methods to analyse the cell surface proteins isolated from primary CLL cells.

Aims of thesis:

- To investigate the effects of different co-culture models on the viability of primary CLL cells.
- To investigate common transcriptional effects induced in CLL cells by endothelial cell co-culture.
- To use mass spectrometry to identify proteins enriched in CLL cell surface protein samples.

Chapter 2

Materials and Methods

2. Materials and Methods

2.1 Reagents

2.1.1. General chemicals, consumables and kits

7-amino-actinomycin (7AAD)	BD Bioscience
Acetonitrile (ACN)	Sigma
Annexin V-FITC	BD Bioscience
Anti-CD3/CD28 immunomagnetic beads (Dynabeads)	Invitrogen, Life Technologies
Anti-mouse IgG immunomagnetic beads (Depletion Dynabeads)	Invitrogen, Life Technologies
Bovine serum Albumin (BSA)	Sigma
Bromophenol Blue	Sigma
Cell Surface Protein Isolation Kit, comprising: EZ-Link Sulfo-NHS-SS-Biotin, Quenching Solution, Lysis Buffer, NeutrAvidin Agarose, Wash Buffer Dithiothreitol (DTT), PBS, TBS, Spin Columns, Collection Tubes	Pierce (Thermo Scientific)
Coomassie brilliant blue R250	Pierce (Thermo Scientific)
Dimethyl sulfoxide (DMSO) tested for tissue culture, endotoxin free	Sigma

Dithiothreitol (DTT)	Sigma
Dulbecco's Modified Eagle Medium (DMEM)	Sigma
ECL plus Chemiluminescent Western Blot Detection Kit	GE Healthcare
1.5ml microcentrifuge tubes 0.5ml micocentrifuge tubes	Starlabs
EasySep Human B Cell Enrichment Kit without CD43 depletion	STEMCELL Technologies
Endothelial Cell Growth Medium	Promocell
Ethanol	Fisher Scientific
15ml and 50ml Falcon tubes	VWR international
5ml FACS tubes	BD Biosciences
0.2µm filter	TPP, Helena Bioscience
Fixation/Permeabilisation concentrate	eBioscience
Fluorescein isothiocyanate (free FITC)	Sigma
Foetal bovine serum (FBS)	Sigma
GeneChip Human Gene 1.0 ST Array	Affymetrix
GeneChip WT Expression Kit (Buffers proprietary composition)	Ambion
GeneChip WT Terminal Labelling Kit (Buffers proprietary composition)	Affymetrix
GeneChip Expression Wash, Stain and Scan Kit (Buffers proprietary composition)	Affymetrix
Glycogen	Ambion

Histopaque 1077	Sigma
Hybond- C Extra membrane	GE Healthcare
Hyperfilm ECL	GE Healthcare
Hypersep C18 Columns	Thermo Scientific
IGEPAL-CA-630 (Nonidet P40)	Sigma
Iodoacetamide	Sigma
M199 Medium	Gibco
Methanol	Fisher Scientific
Mini-Protean TGX 4-15% (w/v) polyacrylamide gel	Bio Rad
Neubauer Improved Haemocytometer	VWR international LTD
Nitrocellulose membrane (Hybond C Extra)	GE Healthcare
Non-fat dried milk (Marvel)	Sainsbury's
Novex sharp protein standards	Invitrogen, Life Technologies
NuPAGE gels, running and transfer buffer	Invitrogen, Life Technologies
Paraformaldehyde (PFA)	Sigma
Penicillin	Sigma
Phosphate buffered saline (PBS) tablets	Oxoid
Pre-separation filter, 30µm nylon mesh	Miltenyi
Propodium Iodide (PI)	Sigma
RNeasy kit	Qiagen
RNAse A	Sigma
Rosswell Park Memorial Institute medium (RPMI-1640)	Invitrogen, Life Technologies

RNase Out	Invitrogen, Life Technologies
Single donor buffy coats/ cones	National Blood Transfusion Service
Streptomycin	Sigma
Trifluorethanol (TFE)	Sigma
T cell negative isolation kit	Invitrogen, Life Technologies
Trizol	Invitrogen, Life Technologies
Trypsin	Sigma
Trypsin (proteomics grade)	Sigma
Tween-20	Sigma
Whatman 3MM paper	VWR international LTD
X-ray film (Hyperfilm–ECL)	GE healthcare
X-Vivo 15 cell culture media	Lonza

Table 2.1 General chemicals, consumables and kits

2.1.2. Buffers and solutions

Cell cycle stain	40µg/ml Propidium Iodide, 5µg/ml Fluorescein isothiocyanate and 1µg/ml of RNase A in PBS
Mini Complete Protease inhibitor cocktail tables EDTA free (serine and cysteine protease inhibitor)	Roche proprietary composition
Fixing solution	1% (w/v) PFA
Fluorescein Isothiocyanate (FITC)	1µg/ml in PBS
10 x Iodoacetamide solution	550mM Iodoacetamide (IAM)

stock	
1 x MES SDS running buffer	50mM MES, 0.1% (w/v) SDS, 1mM EDTA, 50mM Tris-HCl pH 7.3
MS buffer A	100% (v/v) H_20, 0.1% (v/v) formic acid
MS buffer B	100% (v/v) Acetonitrile, 0.1% (v/v) formic acid
MS buffer C	95% (v/v) H_20, 5% (v/v) Acetonitrile, 0.1% (v/v) formic acid
1 x NuPage transfer buffer	25mM Bicine, 1mM EDTA, 25mM Bis-Tris pH 7.2, 20% (v/v) Methanol
Phosphate Buffered Saline (PBS)	1 tablet (phosphate buffer, 0.02% (w/v) potassium chloride, 0.8%(w/v) sodium chloride) dissolved per 100ml of H_2O and autoclaved
Phosphatase inhibitors	2mM β-glycerophosphate, 5mM NaF, 1mM Na_3VO_4, 0.1µM okadaic acid
Propodium Iodine (PI)	1mg/ml in dH_2O
RNAse A	10mg/ml in dH_2O
10% (w/v) sodium dodecyl sulphate (SDS)	50g SDS dissolved in 500ml of dH_2O
2x SDS Lysis Buffer	125mM Tris-HCl pH 6.8, 4% (w/v) SDS, 40% (v/v) Glycerol, 200mM DTT and 0.002% (w/v) Bromphenol blue
Secondary antibody incubation buffer	10% (w/v) Non-fat dried milk, PBS, 0.05% (v/v) Tween-20
Tris Buffered Saline (TBS)	25 mM Tris-HCl pH 7.2, 150 mM NaCl
Western blot wash buffer	PBS, 0.05% (v/v) Tween-20
Western blot blocking solution	10% (w/v) Non-fat dried milk, PBS, 0.05% (v/v) Tween-20

(PBST)	
Endothelial Cell Growth Medium-2	2% (v/v) FBS, (5ng/ml) Epidermal Growth Factor, (10ng/ml) Basic Fibroblast Growth Factor, (22.5µg/ml) Heparin, (0.5ng/ml) vascular endothelial growth factor 165, (20ng/ml) insulin-like growth factor, (1µg/ml) ascorbic acid, (0.2µg/ml) hydrocortisone
HMEC-1 Medium	M199 medium, 10% (v/v) FBS, 2mM L-glutamine, 10µg/ml Endothelial cell growth supplement (ECGS), 1µg/ml Hydrocortisone, 5µM 2-mercaptoethanol, 10ng/ml human epidermal growth factor (hEGF), 1µg/ml Ascorbic acid in M199, 0.5ng/ml vascular endothelial growth factor (VEGF), 10ng/ml insulin like growth factor 1 (IGF-1)

Table 2.2 Buffers and solutions

2.1.3. Antibodies for western blotting

Antibody	Clone or identifier	Supplier
BCL2	100/D5	Abcam
CD5	MEM-32	Abcam
CD44	KZ-1	Institute for Transfusion Sciences and International Blood Reference Laboratory, *National Blood Service*, *Bristol*, UK
CD79B	CD79B	Santa Cruz Biotechnology

Cdk6	C19	Santa Cruz Biotechnology
Histone H3	Ab1791	Abcam
HRP-conjugated rabbit anti goat antibody	P 0449	DAKO
HRP-conjugated goat anti rabbit antibody	P 0448	DAKO
HRP-conjugated goat anti mouse	P 0447	DAKO
MHC1	HC10	From Dr Linda Barber
ZAP70	D1C10E	Cell signalling technologies

Table 2.3 Antibodies for western blotting

2.1.4. Antibodies for flow cytomety

Antibody	**Clone or identifier**	**Supplier**
CD3-FITC	Okt3	eBiosciences
CD5-PECy7	UCHT2	eBioscience
CD11c-FITC	KB90	Dako
CD19-Pacific blue	48-0199	eBioscience
CD38-PE	HB7	eBioscience
CD44-PE	IM7	eBioscience

CD49d-FITC	Bu49	Serotec
CD69-APC	FN50	eBioscience
CD69-PE	FN50	eBbiosciences
CD103-PE	Ber-ACT8	Dako
CD138-APC	MI 15	Dako
Ki-67-FITC	51-36524	BD Pharmingen
Mouse isotype control PE	IgG1 kappa	Pharmingen
Mouse isotype control PE	IgG2a	eBioscience
ZAP70-FITC	IE7.2	eBioscience

Table 2.4 Antibodies for flow cytometry

2.2 Cell culture

2.2.1. Culturing HeLa cells

HeLa, a human cervical cancer cell line (162) were cultured as an adherent monolayer in DMEM supplemented with 10% (v/v) FBS, 1% glutamine, 100U/ml penicillin, 100µg/ml streptomycin at 37^{o}C in a humidified 5% (v/v) CO_2 atmosphere. HeLa were passaged when they reached about 80% confluency. The cells were washed in 1x PBS, followed by the addition of trypsin to the adherent monolayer of cells. The cells were incubated at 37^{o}C for 2-5 minutes until the cells began to detach. Approximately 5ml PBS containing 10% (v/v) FBS was added to the cells to quench the trypsin and cells were then centrifuged at 150 x G_{max} for 5 minutes and the supernatant containing trypsin was removed.

2.2.2. Culturing HMEC-1 cells

Adherent HMEC-1, Human microvascular endothelial cells (85) were passaged when they reached about 80% confluency. The cells were washed in 1x PBS, followed by the addition of trypsin to the adherent monolayer of cells. The cells were incubated at 37^{o}C for 2-5 minutes until the cells began to detach. Approximately 5ml PBS containing 10% (v/v) FBS was added to the cells to quench the trypsin and cells were then centrifuged at 150 x G_{max} for 5 minutes and the supernatant containing trypsin was removed. For assays in Chapter 3, the cells were then resuspended in HMEC-1 media and seeded at a lower density and returned to the 37^{o}C, 5% CO_2 incubator in HMEC-1 media without antibiotics. For assays in Chapter 4, the cells were resuspended in Endothelial Cell Growth Medium 2. When co-culturing HMEC-1 and CLL cells, 1% (w/v) BSA was added to the prepared HMEC-1 media and filter sterilised using a 0.2 µM filter.

2.2.3. Culturing primary endothelial cells

HDBEC (from Professor Mark Peakman) and HDMEC (Promocell) primary endothelial cells were cultured in Endothelial Cell Growth Medium 2 at 37^{o}C in a humidified 5% (v/v) CO_2 atmosphere and trypsinised as described in section 2.2.2.

2.2.4. Cryopreservation of cells

Cells were frozen in liquid nitrogen and stored long term for use at a later date. The appropriate quantity of cells were pelleted by centrifugation at 200 x G_{max} for 5 minutes and resuspended in a solution containing 50% (v/v) of the appropriate storage medium (either RPMI-1640 or DMEM, depending on cell type), 40% (v/v) FBS and 10% (v/v) DMSO. The cells were aliquotted into cryovials and frozen overnight at -80^{o}C in a Nalgene Mr Frosty freezer container. The cryovials containing the frozen cells were then transferred to a liquid nitrogen storage vessel for long term storage. The liquid nitrogen storage is in the vapour phase, maintained and monitored as part of our accredited departmental Tissue Bank.

2.2.5. Reviving cryopreserved cells

Frozen cells are revived by immediate thawing under warm water after retrieval from liquid nitrogen. Immediately after thawing, the cells are washed in PBS to remove the freezing mix, which contains DMSO. The cells are then transferred to the appropriate pre-warmed medium.

2.2.6. Isolation of quiescent T cells

Normal human T Cells were isolated from single donor Buffy Coats, recently superseded by Leukocyte Cones. Peripheral blood mononucleocytes (PBMCs) were obtained by density-gradient

separation using Histopaque 1077, the cells were centrifuged for 30 minutes at 560 x G_{max} with no brake. The PBMCs were then removed with a Pasteur pipette and washed in 50ml of PBS and centrifuged at 400 x G_{max} for 10 minutes. Platelets were removed by centrifuging twice in 50ml of PBS with 2% (v/v) FBS at 200 x G_{max}. Non-activated T cells were isolated from the PBMCs by negative selection using the T Cell Negative Isolation Kit, which contains antibodies against monocytes (CD14 and HLA Class II DR/DP), granulocytes (HLA Class II DR/DP), B Cells (HLA Class II DR/DP), NK Cells (CD16 a and b, CD56), erythroid cells (CD235a) and activated T Cells (HLA Class II DR/DP). The PBMCs were suspended in PBS with 2% (v/v) FBS at 1 x 10^8/ml and incubated at 4^{o}C with rotation for 20 minutes with 20µl of antibody mix per 1 x 10^7 cells. Unbound antibody was removed by washing with PBS/2% (v/v) FBS and centrifugation at 500 x G_{max} for 8 minutes. Cells were resuspended at 1 x 10^7/ml in PBS/2% (v/v) FBS with 100µl anti-immunoglobulin conjugated magnetic beads per 1 x 10^7 cells and incubated at room temperature, with gentle rolling, for 15 minutes. Bead clumps are dispersed with gentle pipetting and labelled cells were removed using a magnet (Dynal Magnetic Particle Concentrator). The supernatant containing the non-labelled quiescent T cells was transferred to a clean tube and collected by centrifugation. T cells were cultured at 4 x 10^6/ml in X-VIVO 15 with 10% (v/v) FBS.

2.2.7. Culturing T Cells

The quiescent (G_0) T cells were generally cultured at 4 x 10^6/ml in X-Vivo 15 with 10% (v/v) FBS, L-glutamine (2mM final), penicillin (2,000 units per ml), streptomycin (2mg/ml) at 37^{o}C in a fully humidified atmosphere of 5% CO_2.

2.2.8. Stimulation of quiescent (G_0) T cells

Where required, quiescent T cells were stimulated by the addition of anti-CD3/CD28 magnetic beads at a ratio of 0.5 beads/cell.

2.2.9. CLL cell isolation

An EasySep negative selection kit was used to enrich human B cells from patients with CLL without depletion of CD43. Bi-specific Tetrameric Antibody complexes (TAC) label unwanted cells by recognising antigens (CD2, CD3, CD14, CD16, CD56 and glycophorin A) expressed on T and NK cells. These complexes are then removed by dextran coated magnetic nanoparticles. PBMCs were obtained by density-gradient separation using Histopaque 1077, as described in 2.2.6 above for the isolation of T cells. The cells were centrifuged for 30 minutes at 560 x G_{max} with no brake. PBMCs were then removed using a Pasteur pipette and cells were re-suspended at a concentration of 5 x 10^7 cells/ml in PBS containing 2% (v/v) FBS. EasySep Negative Selection Human B cell enrichment

cocktail without CD43 depletion was then added at 50 µl/5 x 10^7 cells, mixed and incubated at room temperature for 10 minutes, after which the cell suspension was brought up to a total volume of 2.5 ml with PBS containing 2% (v/v) FBS to remove unbound antibody. The magnetically labelled unwanted cells are removed using a magnet. The negatively selected, enriched B cells remain in the supernatant.

2.2.10. CLL Patient Samples

PBMCs were collected and cryopreserved as previously described (52) from randomly selected patients with confirmed CLL. PBMCs were isolated by density gradient separation, as described in 2.2.6 (Histopaque-1077). Cells were either processed fresh or cryopreserved in RPMI 1640, 40% (v/v) FBS and 10% (v/v) DMSO as previously described (52). Ethical approval was obtained from the local institutional review board of King's College Hospital NRES Reference 08/H0906/94+5 (see appendix for Participant Information Sheet) and in every case, informed written consent was obtained.

LSL number	Binet Stage	CD38%	Mutational Status	Serotype	FISH	Cytogenetics
LSL/004448	A					
LSL/004467	C	52	unmutated	IGHV: 1-2 Identity: 100% IGHD: 6-19 IGHJ: 4	Deletion of one copy of 13q14.3	43~46,XX,-14,-14,add(15)(p10),add(16)(q24),add(17)(q25),+1~3mar [cp10]
LSL/004513	B	57	unmutated	IGHV: 1-69 Identity: 100% IGHJ: 3 IGHD: 3-16	normal (p53 mut)	45,XY,del(1)(q42q44),add(17)(p11),-18 [3]46,XY [27]
LSL/004539	A	35			normal	normal
LSL/004589	A	48	mutated	IGHV: 4-34 identity: 95.89% IGHJ: 6 IGHD: 6-13	normal	46,XX,add(19)(q13)[3]46,XX [7]
LSL/004631	B	96	mutated	IGHV 3-21	Del one copy 13q14.3	normal
LSL/004635	A	1.5				
LSL/004654	A	1				
LSL/004749	B	3	mutated	IGHV: 3-23 Identity: 86.5% IGHD: 6-19	del 13q14.3	normal
LSL/005030		85	unmutated	IGHV: 5-51 Identity: 98.7	Trisomy 12	46,XY [15]
LSL/005033	C	22	unmutated	IGHV: 4-4 identity: 100% IGHJ: 6 IGHD: 1-7		47,XY,+12,del(14)(q24q32) [15]
LSL/005195		0	mutated			
LSL/005417	A	49				
LSL/005502	A	1			del 13q14.3	46,XX,del(13)(q14.1q14.3)[12]46,XX [18]
LSL/005539	A	2	unmutated	IGHV: 2-70 Identitiy: 100% IGHD: 3-3 IGHJ:4	Del 13q14.3	46,XY,del(13)(q14.1q14.3)[6]46,XY [9]
LSL/005750	B	44			normal	normal
LSL/006022	B	98	mutated	IGHV: 1-8 Identity: 95.4% IGHJ: 5 IGHD: 3-9	Del one copy of p53	normal
LSL/006171	A	49	unmutated	IGHV:4-34 Identity: 100% IGHD:6-19 IGHJ:5	Del 13q14.3 del TP53	45,XY,add(5)(q13),-17 [12]
LSL/006811	B	98	unmutated	IGHV:1-69 Identity: 100% IGHD:2-21 IGHJ:3	Trisomy 12 detected	
LSL/006917		49	unmutated			
LSL/007739		43	unmutated	VH4-39		
LSL/007870		43	mutated	IGHV:1-3 Identity: 88.1% IGHJ:4 IGHD:6-19		

Table 2.5 Patient samples used in this study (Non routine diagnostic data provided by Dr Najeem Folarin)

2.2.11. CLL liquid culture conditions

CLL PBMCs were cultured at 2 x 10^6/ml in HMEC-1 recommended medium in Chapter 2 and Endothelial Cell Growth Medium 2 in Chapter 3, each with 2mM L-glutamine, 2,000 units per ml penicillin, 2mg/ml streptomycin, 1% (w/v) BSA. Cells were incubated at 37^{o}C in a fully humidified atmosphere of 5% CO_2.

2.2.12. CLL co-culture conditions

HMEC-1, HDMEC and HDBEC were seeded at 1 x 10^5/ml in 24 well plates in the recommended medium and incubated overnight to allow cells to adhere. CLL PBMCs were cultured alone and on endothelial cells at 2 x 10^6 cells/ml, as previously described (1) and harvested at the time points indicated.

2.3 Flow cytometry

2.3.1. Analysis of the cell cycle by flow cytometry

The percentage of cells in each phase of the cell cycle was determined by analysing the DNA and total protein content, as used for previous studies in our laboratory *e.g.* (163). 2 x 10^5 cells per sample were collected by centrifugation at 200 x G_{max} for 5 minutes. 2 x 10^5 T cells were taken at the indicated time points and were fixed in 500µl of 70% (v/v) ethanol. The fixed cells were then centrifuged in FACS tubes at 350 x G_{max} for 8 minutes and the supernatant was removed. The cell pellet was resuspended in 400µl FITC/PI cell cycle stain, which contains DNAase-free RNAase. RNA in the sample was digested for 30 minutes at 37ºC. Flow cytometric analysis was then performed using a Becton Dickinson FACS Calibur machine. The WinMDI2.9 or FlowJo program was used to analyse the data. Cells which pass the flow cytometer laser as doublets or aggregates were excluded from the analyses. This was done as, for example doublets of cells in G_0/G_1 will have 2 x 2n DNA content and will therefore be quantified as being in G_2/M and having 1 x 4n DNA content. To avoid the artefact that the doublets will be scored as being in G_2/M rather than in G_0/G_1, a doublet discriminator gate was applied in all flow cytometry analyses. The principle is that total fluorescence detection, which is equivalent to the area under the curve (FL2-A Area), is proportional to the time that a cell needs to pass a detector (FL2-W Width). Doublets will have a much larger FL2-W, due to their bigger size. Therefore plotting the area (FL2-A) against width (FL2-W) and applying a gate around cells with a correct pulse width corresponding to single cells, excludes the doublet events. PI (FL-2A) was then plotted against FITC (FL-1H), which quantifies total protein content and is a surrogate measure of cell size (see Figure S1C of Orr *et al.* (155)). The

percentages of cells in each cell cycle phase were calculated by manually applying gates around populations of cells in G_0, G_1, S phase, G_2/M and apoptosis. Cells undergoing apoptosis typically have a DNA content <2n (also known as sub-G_1).

2.3.2. Determination of the activation state of T cells

Quiescent T cells were stimulated with CD3/CD28 beads overnight and harvested by centrifugation at 200 x G_{max} for 5 minutes. Cells were then resuspended in 100 µl PBS at a concentration of 1- 2 x 10^6 cells/ml and stained with anti-human CD69-PE or PE-labelled isotype-matched control at 4^{o}C for 30 minutes. Expression of CD69 was determined by flow cytometric analysis using a Becton Dickinson FacsCalibur instrument.

2.3.3. Flow cytometric analysis of cell surface antigens

Phenotypic analyses were performed on cells isolated from CLL patient samples by flow cytometry. 2 x 10^5 cells were incubated with the appropriate antibodies at 4^{o}C for 20 minutes. The samples were then washed with PBS and centrifuged at 150 x G_{max} for 5 minutes. The samples were then fixed in 1% (w/v) PFA. The phenotype of CLL PBMCs at time zero and following co-culture for 24 hours with HMEC-1 cells or control medium was analysed by 5-colour flow cytometry using a BD FACS Canto II instrument. CD38, CD44, CD49d, ZAP-70 and CD69 expression were determined using CD19-PB, CD5-PeCy7, CD38-PE, CD44-PE, CD69-APC, ZAP-70-FITC, CD49d-FITC. Apoptosis of cells was assessed after 7 days co-culture using flow cytometry following labelling with CD19-PB, CD5-PeCy7, Annexin V–FITC, and 7-amino-actinomycin D (7-AAD) according to the manufacturer's instructions. Proliferation was analysed at various time points by measuring Ki-67 expression using flow cytometry (FACS Canto II). Cells were labelled with CD19-PB, CD5-PeCy7 before treatment with Fix and Perm reagent supplemented with 5% (v/v) NP-40 and labelling with Ki-67-FITC or matched isotype control.

2.3.4. Flow cytometric analysis of intracellular antigens

Antibodies to cell surface antigens were added using the above protocol. After washing with PBS, the cell pellet was resuspended by vortexing as 100µl fixation buffer was added. After thorough mixing, the sample was incubated in the dark at room temperature for 20 minutes. 1ml of permeabilisation buffer was then added to each sample, mixed and centrifuged at 150 x $G_{max.}$ An antibody to the intracellular protein, ZAP70 was then added to appropriate tubes and the samples were incubated at room temperature for 20 minutes. The samples were then washed once in permeabilisation buffer and once in PBS, with centrifugation performed at 150 x G_{max}. The samples

were then be fixed in 1% (w/v) PFA and stored in the dark at 4^{o}C until analysed using the FACS Canto II.

2.3.5. Flow cytometric viability assay

The viability of cells was assessed using flow cytometry. Annexin V binds to phosphatidyl serine, which is normally on the intracellular leaflet of the plasma membrane, but is moved to the extracellular leaflet when apoptosis is triggered. Detection of extracellular phosphatidyl serine can therefore be used as an early marker of apoptosis. 7AAD is a dye which stains DNA. Cells which are negative for both Annexin V and 7AAD are viable cells. This viability protocol was used after following the cell-surface antigen labelling protocol described above, or independently. Briefly, 100µl of Annexin V binding buffer was added to each tube and 2.5µl Annexin V and 1µl 7AAD was added as required. The samples were then incubated in the dark for 15 minutes before a further 400µl of Annexin V binding buffer was added to each tube. The samples were analysed immediately and the percentage of cells staining positive for Annexin V (FITC) and 7AAD (PE Cy5.5) was determined by flow cytometry (FACS Canto II).

2.4 Protein analysis

2.4.1. Total Protein Lysates

Cells were harvested by centrifugation at 200 x G_{max} for 5 minutes, washed with PBS and centrifuged again at 200 x G_{max} for 5 minutes. The cell pellet was disrupted by flicking the tube and an appropriate volume of 2X SDS Lysis Buffer was added. The tube was flicked vigorously to resuspend the pellet before heating at 100^{o}C for 10 minutes. Samples were briefly centrifuged (30 seconds, 1,000 x G_{max}) and stored at –20^{o}C if not used immediately.

2.4.2. Western Blotting

Protein samples were separated by Polyacrylamide Gel Electrophoresis (PAGE) using the NuPAGE system and 4-12% (w/v) polyacrylamide bis-Tris gels. Novex Sharp pre-stained ladders were used as protein standards. Gels were assembled in a NuPAGE gel apparatus and electrophoresis was carried out for 1 hour at 200V in 1X MES SDS Running Buffer. After electrophoresis, the separated proteins were transferred to a nitrocellulose-coated nylon membrane (Hybond C-Extra) using the NuPAGE gel blotting module at 25V and NuPAGE Transfer Buffer with 20% (v/v) methanol, according to the manufacturer's instructions. The membrane was then removed from the apparatus

and blocked for 30 minutes in 10% (w/v) non-fat dried milk in PBS, 0.05% (v/v) Tween-20. After three brief washes in PBS, Tween-20 the blot was placed in primary antibody.

Primary antibodies for Western blotting were typically used at dilutions of 1:1000 in PBS, 3% (w/v) BSA, 0.01% (v/v) NaN_3. The dilution corresponds typically to a final immunoglobulin concentration of 0.1-0.2 μg/ml. Blots were incubated in primary antibodies overnight at 4°C or for 1 hour at room temperature. They were then washed three times in PBS, 0.05% (v/v) Tween-20 and incubated in secondary antibody for 45 minutes. Horseradish peroxidase (HRP) conjugated secondary antibodies were used at a dilution of 1:2000 in PBS, 10% (V/V) dried milk. Blots were washed three times in PBS, 0.05% (v/v) Tween-20 before the addition of 1ml ECL-Plus for 10 minutes. To visualise protein bands, the blots were exposed to Hyperfilm-ECL X ray film and developed using a Compact X4 X-Ray developer.

2.5 RNA analysis

RNA isolation, sample processing and microarray analyses were carried out under the supervision of Mrs Megan Musson at the Central Biotechnology Service Laboratory, University of Cardiff as part of a collaboration with Professor Chris Pepper.

2.5.1. Isolation of total RNA

4×10^6 cells were centrifuged at 300 x G_{max} for 5 minutes and the supernatant was discarded. The pellet was resuspended in 1ml of Trizol by pipetting and incubated at room temperature for 5 minutes. The sample was stored at -80°C until processed further when the sample was transferred to a 1.5ml micro-centrifuge tube and 150μl of Chloroform was added to the sample for phase separation. After mixing vigorously, the sample was incubated at room temperature for 3 minutes, followed by centrifugation at 13,000 x G_{max} for 2 minutes. The separated upper aqueous phase was transferred to a fresh micro-centrifuge tube and RNA was co-precipitated with 1μl glycogen using 500μl of isopropanol. The addition of glycogen helps to visualise the RNA pellet. Samples were incubated at room temperature for 10 minutes before centrifugation at 12,000 x $G_{max,}$ 4°C for 10 minutes. The supernatant was removed and the pellet was washed in 750 μl of 75% (v/v) ethanol. The sample was then centrifuged at 7,500 x G_{max} for 5 minutes at 4°C. The supernatant was removed until approximately 10μl remained, then the pellet was air dried at room temperature. The pellet was dissolved in 100μl of RNase-free dH_2O at 65°C for 5 minutes. The samples were then cooled on ice for 15 minutes. The quantity and quality of the RNA was determined using a Nanodrop appatatus (see section 2.5.3 and the quality was determined by a second method using the 2100 Bioanalyser (Agilent) as described in section 2.5.4.

2.5.2. RNA further purification

Total RNA 'clean up' was performed using RNeasy mini kit. 350 µl buffer RLT was added to the sample of 100 µl total RNA and mixed well. 250 µl 100% (v/v) ethanol was then added to the diluted RNA and mixed by pipetting. The sample was immediately transferred to an RNeasy mini spin column, which was placed in a 2ml collection tube. After closing the lid gently, the samples were centrifuged at room temperature at 8,000 G_{max} for 15 seconds. The column was carefully removed from the collection tube to prevent contact with the flow through and the flow through was discarded. Using the same collection tube, 500 µl RPE buffer was then added to the RNeasy spin column. The lid was closed gently and samples were centrifuged at room temperature at 8,000 G_{max} for 15 seconds to wash the spin column membrane. The flow through was discarded. A further 500 µl RPE buffer was added to the RNeasy spin column and samples were centrifuged at room temperature 8,000 x G_{max} for 2 minutes to wash the spin column membrane. The long centrifugation step dries the spin column membrane, ensuring that no ethanol is carried over during RNA elution as residual ethanol may interfere with downstream reactions and measurements of the RNA. The RNeasy spin column was removed from the collection tube and placed in a new collection tube. Samples were centrifuged at room temperature at 13,000 x G_{max} for 1 minute to eliminate any possible carryover of RPE buffer or to remove any residual flow through from the outside of the RNeasy spin column. The RNeasy spin column was then placed in a new 1.5ml collection tube and 50 µl of RNase-free water was added directly to the spin column membrane. The lid was closed gently and samples were centrifuged at room temperature at 8,000 x G_{max} for 1 minute to elute the RNA. The quantity and quality of the RNA was determined using the Nanodrop apparatus and 2100 Bioanalyser (see sections 2.5.3 and 2.5.4). Samples were then stored at -80^{o}C.

2.5.3. Determination of RNA concentration

RNA concentration was determined by Nanodrop spectrophotometry. This method measures the absorbance of DNA or RNA in solution at 260nm and the quality is assessed by determining the 260nm/280nm ratio. The Beer-Lambert equation A= E x b x c is used to calculate the DNA/RNA concentration. A is the absorbance value, E is the wavelength-dependant molar absorption coefficient or extinction coefficient with units of $lmol^{-1}cm^{-1}$, b is the path length in cm, c is the analyte concentration in mol/l. Typically 1µl of DNA or RNA in dH_2O was placed on the Nanodrop pedestal and the absorbance of the material in the sample column was measured against a blank of dH_20. The 260nm/280nm ratio is considered to be a good indicator of protein contamination since proteins, particularly those containing aromatic amino acids which absorb light at 280nm (164) therefore the purity of the sample can be assessed The A260/A280 ratio should be close to 2.0 for pure RNA.

2.5.4. Determination of RNA quality

RNA quality was determined using an RNA 6000 Nano Assay and 2100 Bioanalyser. Briefly this micro fluidics-based platform method separates and quantifies RNA by electrophoresis in micro fluidic networks of channels and wells etched into glass chips. The RNA Integrity Number (RIN) algorithm allows the reproducible comparison of RNA quality between different samples. Sample integrity is not determined by just the ratio of the ribosomal RNA, but by the entire electrophoretic trace of the RNA sample, including the presence or absence of degradation products. A RIN score of greater than 8 out of 10 is usually preferred for microarray applications.

2.5.5. Sample preparation for microarray analysis

RNA samples were prepared for Affymetrix whole transcriptome microarray analysis using the Ambion Whole Transcript (WT) Expression Kit. First, sense strand cDNA was generated from total RNA. The WT Expression Kit uses a reverse transcription priming method that specifically primes non-ribosomal RNA from the sample, including both poly(A) and non-poly(A) mRNA. Primers that avoid rRNA binding are designed by using a proprietary-oligodeoxynucleotide matching algorithm. These primer sequences provide complete and unbiased coverage of the transcriptome while significantly reducing the priming of rRNA. cDNA was then fragmented and labelling was carried out using the Affymetrix GeneChip WT Terminal Labelling Kit. Hybridisation of labelled, fragmented cDNA to the GeneChips occurs overnight (16 hours) in a hybridisation oven at 45^{o}C.

100ng RNA was used for microarray analysis. A set of poly(A) RNA controls (Affymetrix GeneChip Poly(A) RNA Control Kit) were added to the RNA, which have been designed specifically to provide exogenous positive controls to monitor the entire target labelling process. Each eukaryotic GeneChip probe array contains probe sets for several *B. Subtilis* genes that are absent in eukaryotic samples. The polyadenylated transcripts for the *B. subtilis* genes are added directly into RNA samples ('Poly(A) spike').

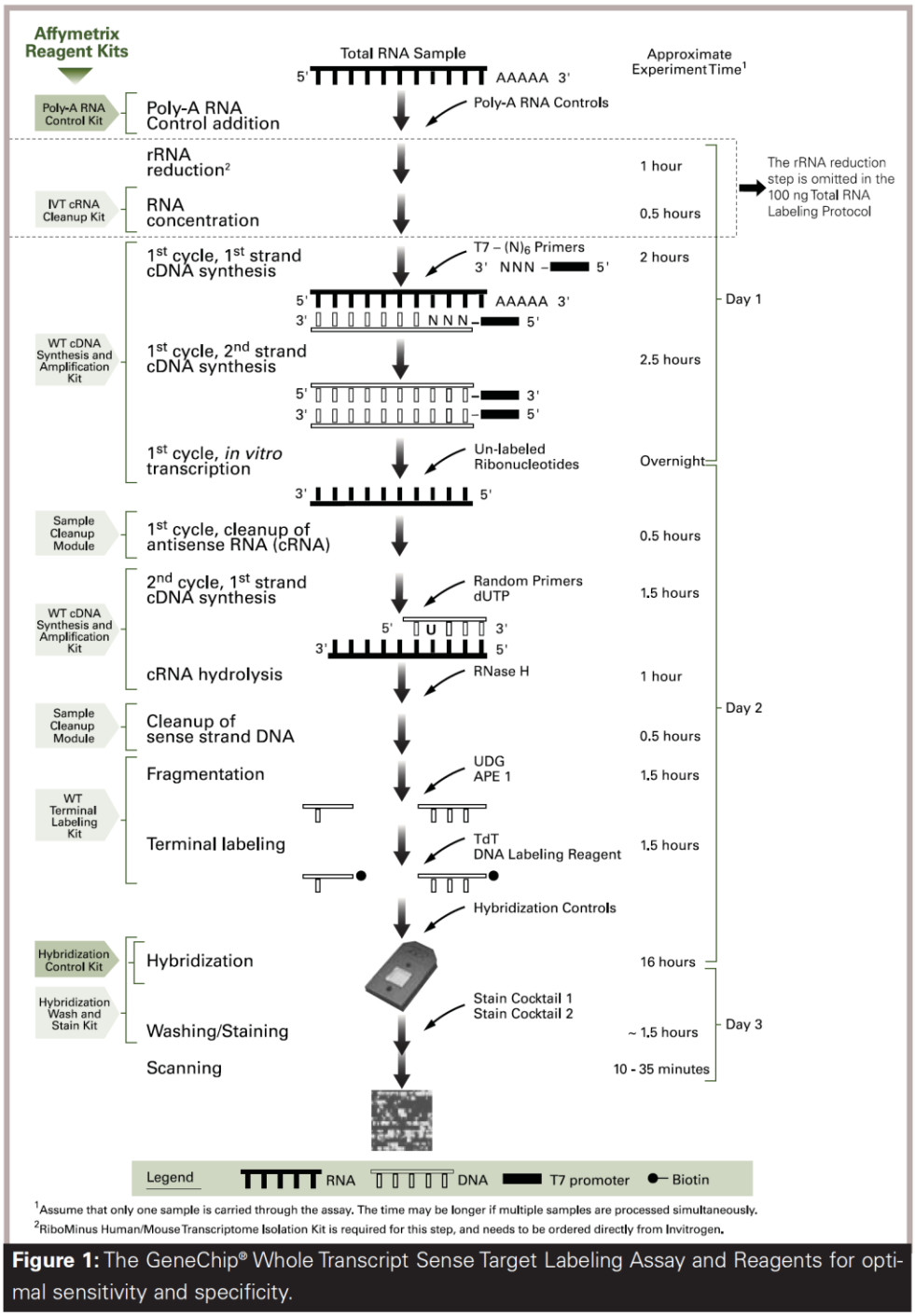

Figure 1: The GeneChip® Whole Transcript Sense Target Labeling Assay and Reagents for optimal sensitivity and specificity.

Figure 2.1 Sample work up for microarray analysis (from the Whole Transcript (WT) Expression Kit insert).

2.5.6. First-strand cDNA synthesis

In this reverse transcription procedure, total RNA was primed with engineered primers containing a T7 promoter sequence (165). The reaction then synthesizes single stranded cDNA containing a T7 promoter sequence. Briefly, first-Strand Master Mix was added to nuclease free PCR tubes, and then 5 µl of RNA (with poly(A) spike) was added to each tube. Samples were mixed by vortexing and centrifuged briefly to collect the reaction at the bottom of the tube. Samples were then incubated for 1 hour at 25^{o}C, then for 1 hour at 42^{o}C, then for 2 minutes at 4^{o}C in a thermal cycler. Immediately after the incubation the samples were centrifuged briefly (approximately 5 seconds) to collect the first-strand cDNA at the bottom of the tube. The samples were then placed on ice for 2 minutes to cool.

2.5.7. Second-strand cDNA synthesis

Single-stranded cDNA was converted to double-stranded cDNA, which acts as a template for transcription. The reaction uses DNA polymerase and RNase H to simultaneously degrade the RNA and synthesize second-strand cDNA.

Second-Strand Master Mix is placed in a nuclease-free tube and mixed by vortexing. 50µl of the second-strand master mix was then added to each tube containing the first-strand synthesis cDNA sample and mixed gently by flicking the tube. Samples were centrifuged briefly to collect the reaction at the bottom of the tube and then incubated for 1 hour at 16^{o}C, then for 10 minutes at 65^{o}C, then for at least 2 minutes at 4^{o}C in a thermal cycler. Immediately after the incubation, samples were centrifuged briefly (approximately 5 seconds) to collect the double-stranded cDNA at the bottom of the tube and samples were placed on ice to cool.

2.5.8. *In vitro* transcription cRNA synthesis

Antisense cRNA was synthesized and amplified by *in vitro* transcription (IVT) of the second-strand cDNA template using T7 RNA polymerase (165). Briefly, an IVT Master Mix was prepared in a nuclease-free tube, mixed by vortexing and centrifuging briefly before adding 30 µl of the IVT Master Mix to each tube containing a Second-Strand cDNA sample. Samples were mixed thoroughly by gently vortexing, then centrifuged briefly to collect the reaction at the bottom of the tube and then incubated for 16 hours at 40^{o}C, then overnight at 4^{o}C in a thermal cycler. After incubation, the reaction was placed on ice.

2.5.9. Purification of cRNA

Enzymes, salts, inorganic phosphates and unincorporated nucleotides were removed to improve the stability of the cRNA. Briefly, cRNA Binding Mix was prepared in a nuclease-free tube and 60µl of cRNA Binding Mix was added to each sample and mixed by pipetting. Samples were then transferred to a "U" bottomed plate and 60µl of isopropanol was added to each tube and mixed by pipetting. The samples were mixed for 2 minutes by shaking. The cRNA in the sample binds to the Nucleic Acid Binding Beads during this incubation. The plate was then transferred to a magnetic stand to capture the magnetic beads. After approximately 5 minutes the supernatant was carefully aspirated without disturbing the magnetic beads and then the plate was removed from the magnetic stand. Samples were then washed twice with 100µl of Nucleic Acid Wash Solution for 1 minute on a plate shaker. After the last wash, the pellets were allowed to air dry. Then, 40µl of preheated (55 to 58^{o}C) Elution Solution was added to each sample to elute the purified cRNA from the Nucleic Acid Binding Beads and incubated for 2 minutes. After incubation, the plate was shaken vigorously for 3 minutes before removing the supernatant containing the eluted cRNA to a fresh plate, which was cooled on ice.

2.5.10. Second-cycle cDNA synthesis

Sense-strand cDNA was synthesized by the reverse transcription of cRNA using random primers. The sense-strand cDNA contains dUTP at a fixed ratio relative to dTTP. 10µg of cRNA was required for second-cycle cDNA synthesis. Briefly, 10µg of cRNA was prepared in a volume of 22µl and 2µl of random primers were added. Samples were mixed by vortexing and centrifuged briefly to collect the reaction at the bottom of the tube. Samples were incubated for 5 minutes at 70^{o}C, then 5 minutes at 25^{o}C, then 2 minutes at 4^{o}C in a thermal cycler. After the incubation, samples were cooled on ice and centrifuged briefly to collect the second-cycle cDNA at the bottom of the tube. The second-Cycle Master Mix was prepared on ice and then 16µl was added to each cRNA/Random Primer sample. Samples were mixed by vortexing and centrifuged briefly. The samples were incubated for 10 minutes at 25^{o}C, then 90 minutes at 42^{o}C, then 10 minutes at 70^{o}C, then for at least 2 minutes at 4^{o}C in a thermal cycler. Immediately after incubation, samples were centrifuged briefly and cooled on ice.

2.5.11. Digestion using RNase H

RNase H (2µl) was added to the second-cycle cDNA. Samples were vortexed and centrifuged briefly before incubation for 45 minutes at 37^{o}C, then 5 minutes at 95^{o}C, then for at least 2 minutes at 4^{o}C

in a thermal cycler. Following incubation, samples were centrifuged briefly and cooled on ice. RNase H degrades the cRNA template leaving single-stranded cDNA.

2.5.12. Second-cycle cDNA purification

The second-strand cDNA is purified to remove enzymes, salts and unincorporated dNTPs, preparing the cDNA for fragmentation and labelling. Briefly, 18µl of nuclease free water and 60µl of cDNA Binding Mix was added to each sample and mixed by pipetting before being transferred to a "U" bottomed plate. 120µl of ethanol was added to each sample before shaking the plate gently for 2 minutes. The cDNA in the sample binds to the Nucleic Acid Binding Beads during this incubation. After incubation, the plate was placed on a magnetic stand to capture the Nucleic Acid Binding Beads and the supernatant was discarded, each sample was washed twice with 100µl Nucleic acid wash solution for 1 minute. The cDNA was eluted with 30µl of preheated (55-58^{o}C) Elution Solution. Samples were incubated with Elution Solution for 2 minutes without shaking then a further 3 minutes with vigorous shaking. The plate was placed on a magnetic stand and the supernatant containing the eluted cDNA was removed to a new nuclease free plate on ice.

2.5.13. Fragmentation and labelling the single-stranded cDNA

The Affymetrix GeneChip WT Terminal Labelling Kit is used for the fragmentation and labelling of the cDNA. The second-cycle (sense-strand) cDNA contains dUTP and the kit uses uracil-DNA glycosylase (UDG) and apurinic/apyrimidinic endonuclease 1 (APE1) to recognize and fragment the cDNA at the unnatural dUTP residues. The DNA is then labelled by terminal deoxynucleotidyl transferase (TdT) using the Affymetrix proprietary DNA Labelling Reagent.

16.8µl of the Fragmentation Master Mix containing UDG and APE was added to each ssDNA sample and mixed by vortexing and centrifuged briefly. Samples were incubated at 37^{o}C for 60 minutes then 93^{o}C for 2 minutes then 4^{o}C for at least 2 minutes. After incubation, samples were mixed by flicking the tube and centrifuged briefly. Size analysis was then performed using the Bioanalyzer RNA 6000 Nano LabChip. The range in peak size of the fragmented samples was approximately 40 to 70 nucleotides.

2.5.14. Labelling of fragmented single-stranded DNA

Component	Volume in one reaction
Fragmented single-stranded DNA	45µl
5X TdT Buffer	12µl
TdT	2µl
DNA labelling reagent, 5mM	1µl
Total Volume	**60µl**

Table 2.6 Labelling of fragmented single-stranded DNA

A master mix was made containing TdT and DNA labelling reagent was made and 15µl of the labelling master mix was added to each tube containing fragmented ssDNA. Samples were mixed by flicking the tubes and centrifuged briefly to collect the reaction at the bottom of the tube before incubating at 37^{o}C for 60 minutes then 70^{o}C for 10 minutes then 4^{o}C for 2 minutes.

2.5.15. Hybridisation

Samples were processed using the GeneChip Hybridization, Wash and Stain Kit.

Component	Volume for one 49/64 Format Array	Final Concentration
Fragmented and labelled DNA target	60µl	25ng/µl
Control oligodeoxynucleotide B2 (3nM)	3.7µl	50pM
20 x Eukaryotic hybridisation controls (*bioB, bioC, bioD, cre*)	11µl	1.5, 5, 25 and 100pM respectively
2 x Hybridisation mix	110µl	1X
DMSO	15.4µl	7%
Nucelase-free water	19.9µl	
Total Volume	**220.0µl**	

Table 2.7 Hybridization cocktail

The hybridization cocktail was mixed according to **Table 2.7** and heated at 99^{o}C for 5 minutes, cooled at 45^{o}C for 5 minutes and then centrifuged at 13,000 x G_{max} for 1 minute. Gene Chip ST Arrays were equilibrated to room temperature before use and then the sample was injected onto the array through the septa on the back of the array plastic casing. The array was placed in a hybridization oven at 45^{o}C overnight (17 hours). After hybridization, the arrays were vented by inserting a clean pipette tip into one of the septa and the probe arrays were re-filled with Wash Buffer A. Probe arrays were then washed and stained according to the GeneChip Expression Wash, Stain and Scan Kit (Affymetrix) before scanning.

2.6 Proteomic methods

2.6.1. Isolation of cell surface proteins

Cell surface proteins were isolated using the Cell Surface Protein Isolation kit as per the manufacturer's instructions. Briefly, approximately 3 x10^{7} cells were washed in ice-cold PBS. The cell pellet was resuspended in 30 ml of freshly prepared Sulfo-NHS-SS-Biotin solution and incubated at 4^{o}C for 30 minutes on a roller. The reaction was terminated by the addition of 1.5 ml of Quenching Solution and the cells were washed twice with TBS. The cell pellet was resuspended in 1.5 ml of Pierce Lysis Buffer containing protease inhibitors and phosphatase inhibitors, sonicated 6 x 10 seconds on medium power in a Bioruptor water bath (Diagenode), incubated on ice for 30 minutes and centrifuged at 10,000 x G_{max} for 2 minutes at 4^{o}C. The clarified supernatant containing the solubilised proteins was incubated with immobilized NeutrAvidin Gel for 1 hour at room temperature on a rolling platform. Unbound proteins were removed from the tube by three washes with TBS containing protease inhibitors. Biotinylated proteins bound to the gel were eluted by incubation with SDS-PAGE Sample Buffer containing 50 mM DTT for 1 hour at room temperature and subsequent centrifugation at 1,000 x G_{max} for 2 minutes. The eluted proteins were separated by electrophoresis for western blot using a NuPAGE Bis-Tris 4% to 12% (w/v) polyacrylamide gel, as described in 2.4.2 or stored at -80^{o}C for mass spectrometry analysis.

2.6.2. In solution trypsin digestion protocol for mass spectrometry (MS)

Protein samples were firstly eluted from the NeutrAvidin matrix by incubation in 40µl elution buffer containing 50mM DTT in 30mM Tris-HCl pH8.0 at 95^{o}C for 5 minutes. Protein samples were then solubilised by the addition of Trifuoroethanol (TFE) to a protein/TFE of 1-2mg/ml. Next, DTT was added to a final concentration of 15mM in order to break disulfide bonds, which stabilise tertiary protein structure. The samples were then heated to 55^{o}C for 45 minutes to fully denature the proteins. The samples are then allowed to cool to room temperature before iodoacetamide (IAM)

was added to a final concentration of 55mM. IAM is an alkylating agent, which is used to prevent peptides from forming disulfide bonds. It covalently modifies to the thiol group of cysteine residues. Samples were then incubated in the dark at room temperature for 30 minutes before diluting in 50mM Tris-HCl, 2mM $CaCl_2$ pH8.0 to reduce the TFE concentration to 5% (v/v). Sequencing grade trypsin was resuspended in 1mM HCl at a concentration of 100µg/ml. 1µg of the trypsin solution was used per mg of protein to be digested. Trypsin is a serine protease and cleaves peptides at the carboxyl side of the amino acids lysine or arginine, except when either is followed by proline. The samples were incubated for 5 hours at 37^{o}C and then formic acid was added to 1% (v/v) to stop trypsin digestion. The sample was then stored at 80^{o}C or processed further.

2.6.3. Detergent removal by electrophoresis for MS analysis

Samples were eluted from the NeutrAvidin matrix by incubating at 95^{o}C for 5 minutes in 40µl elution buffer containing 50mM DTT in 30mM Tris-HCl, pH8.0. The eluate was removed by centrifugation at 14000 x G_{max} for 5 minutes. 4 x SDS sample buffer was added to the sample and loaded into a BioRad pre-cast Mini-Protean TGX 4-15% (w/v) polyacrylamide gel. Once all the sample had entered the gel (typically after about 10 minutes), electrophoresis was stopped and the gel was stained with coomassie blue. To speed up staining, the gel was heated gently in the microwave in 20 second bursts, without boiling. The gel was then placed on a shaker for at least an hour. The gel was then destained overnight in 40% (v/v) Methanol, 10% (v/v) Glacial Acetic Acid on a shaking platform. The gel was then placed in dH_20 to fully rehydrate before cutting gel slices.

2.6.4. In-gel reduction, alkylation and digestion for MS

The protein in the gel appears as a blue band. The gel containing this protein band was excised using a scalpel blade and transferred to a micro centrifuge tube. The gel slice was completely rehydrated with dH_20 and then cut into small pieces of approximately 3x3mm. Destaining solution (200mM NH_4HCO_3, 50% (v/v) acetonitrile) was added to cover the gel pieces and the sample was incubated at 30^{o}C for 20 minutes with intermittent vortexing. The supernatant was discarded and the gel was rehydrated with 200mM NH_4HCO_3, before replacing with fresh destain solution until the gel pieces were fully destained. Once fully destained, the gel sample was dried in a vacuum centrifuge apparatus (SpeedVac). The dried gel pieces were then submerged in reducing buffer (10mM DTT, 100mM NH_4HCO_3) and incubated at 56^{o}C for 1 hour. Excess reducing buffer was removed and the gel pieces were resuspended in alkylating solution (100mM iodoacetamide in H_20) and incubated in the dark at room temperature for 30 minutes. The supernatant was then removed and the gel pieces were washed twice with 200mM NH_4HCO_3. The gel pieces were again washed twice with 100% (v/v) acetonitrile, which caused them to shrink and then rehydrated with 200mM NH_4HCO_3. After drying

the gel pieces in the SpeedVac, the gel pieces were rehydrated in enough digestion buffer (50ng/µl trypsin in 1mM HCl) to just cover the gel pieces. After allowing the gel pieces to re-swell for 5 minutes, 20µl 200mM NH_4HCO_3 was added. After a further 5 minutes, excess digestion buffer was removed. The gel pieces were covered with 200mM NH_4HCO_3 and incubated at 30^oC overnight.

The digestion reaction was stopped by adding 1% (v/v) formic acid. The supernatant was removed to a new micro centrifuge tube. Peptides were then extracted from the gel pieces using an extraction buffer containing 0.1% (v/v) formic acid, 60% (v/v) acetonitrile. Gel pieces were incubated in the extraction buffer at 30^oC for 40 minutes with intermittent vortexing. The extraction buffer containing liberated peptides was removed and combined with the supernatant from the digestion reaction. The extraction procedure was repeated and the supernatants were combined. The volume of the extraction supernatant was reduced to about 10-20µl in the SpeedVac.

2.6.5. Sample preparation for MS analysis

After trypsin digestion, protein samples were purified of contaminants (such as small molecules and salts) by solid phase extraction using Hypersep C18 tips. The stationary phase in the tip is formed using silica bound to C18 hydrocarbon chains and peptides in solution can be separated from other compounds in the mixture according to their physical and chemical properties based on how they interact with the C18 chains. The C18 resin was first equilibrated by washing three times with 50µl of 40% (v/v) Acetonitrile solution (40% (v/v) MS Buffer B, 60 % (v/v) MS Buffer A), followed by three washes with 50% (v/v) MS Buffer A. The sample was loaded onto the resin and the resin was then washed three times with 50µl of MS Buffer A. The sample was eluted with 50µl of 40% (v/v) Acetonitrile solution. All samples were dried to 10-20µl with a SpeedVac and resuspended in 60µl of MS Buffer C.

2.6.6. Mass spectrometry analysis

I spent three months in Prof Marcotte's laboratory at the University of Texas at Austin, USA during the Summer of 2012 where I carried out the initial mass spectrometry analyses, in collaboration with Dr Daniel Boutz, a Post-Doctoral Research Fellow. Further analyses of my samples were carried out subsequently, after my return to the UK. Mass spectrometry analyses of the processed peptide samples were carried out using a Surveyor Plus HPLC system connected to an LTQ-Orbitrap or Velos-Orbitrap mass spectrometer (Thermo Scientific).

High pressure liquid chromatography (HPLC) is used to separate a mixture of peptides before analysis by Mass Spectrometry (MS). The peptides were separated based on their partition coefficients between a polar mobile phase (ACN or MeOH) and a hydrophobic stationary phase

(C18 column). HPLC column conditions can be altered so that peptides can be separated better. The composition of the mobile phase is altered to change the retention time of analytes. Altering the time taken to complete the run can also adjust the separation of peptides.

The tryptic peptides were separated by reverse phase chromatography on a Zorbax C18 column in-line with the LTQ-Orbitrap mass spectrometer. The separation of peptides was accomplished by running a 5-38% (v/v) acetonitrile gradient over 230 minutes. Eluted peptides were directly injected by electrospray ionization into the LTQ-Orbitrap, which is equipped with a nano-spray ion source for analysis. Full parent spectra (MS1) were collected at 60,000 resolution and the dominant ion was chosen from these spectra for further analysis. This becomes the parent ion; this precursor ion is activated to undergo fragmentation by collision with inert gas (Collision-Induced Dissociation (CID)), which is known as tandem mass spectrometry. Therefore, in further analyses, a data-dependent scan can be performed where the mass spectrometer identifies products with these characteristic parent ion M/Z values. These parent ions can then be fragmented in order to confirm their composition. The subsequent ion fragmentation spectra (MS2) were collected in a data-dependent manner, with ions required to carry +2 or greater charge for MS2 selection. The top 12 most intense qualifying peaks were selected per round, with peaks selected twice within 30 seconds excluded from selection for 45 seconds to avoid the detection of only the most abundant peptide in the time period. Data were analysed using the Sequest search algorithm within the Proteome Discoverer 1.3 software package (Thermo). Sequest is a tandem mass spectrometry data analysis program, which identifies tryptic peptide sequences from the tandem mass spectra. The tryptic peptide sequences are generated computationally from protein sequences predicted from the latest build of the human genome. The spectra were searched against the non-redundant Ensembl v64 Homo sapiens (NCBI36) protein-coding data set. Results were filtered at 1% false discovery rate (FDR) using Percolator.

Chapter 3

Co-culture of CLL cells

3. Co-culture of CLL cells

3.1 Introduction

Previous work from our laboratory using confocal immunofluoresence microscopy of CLL lymph nodes showed the presence of microvessels (endothelium) at centres of CLL cell proliferation (52). Furthermore*, in vitro* assays have demonstrated that CLL cell contact with endothelial cells (HMEC-1) in a co-culture system promotes CLL cell survival, whilst CLL cells cultured alone undergo extensive apoptosis (1). Therapeutic applications are clear as agents which block pro-survival interactions could be developed as new therapies to treat CLL patients. The use of endothelial cell lines or primary endothelial cells have several advantages over other co-culture systems described in the literature, not least their ready availability and comparative ease of culture. Co-culture systems utilising mesenchymal stem cells (MSC) from bone marrow aspirates of CLL patients suffer from poor availability, low cell recovery and poor growth in ambient oxygen (166).

A variety of co-culture systems have been developed to mimic the tumour microenvironment by imitating different cell:cell interactions and the effects of soluble factors. The components of some of these systems are described in **Table 3.1**.

Culture	**Details**	**Reference**
Murine fibroblasts transfected with human CD40L	Drug testing model for Hsp90 inhibitor in combination with Fludarabine inhibits proliferative/activated CLL cell phenotype.	(118)
Soluble CD40L and IL4	Drug testing model for Fludarabine in combination with soluble CD40L, IL-4 or a combination.	(167)
High density CLL	Conditioned media from crowded cells increased survival of non-crowded cells. Co-culture of fixed cells with viable CLL cells was also protective, suggesting homotypic interactions are important.	(63)
Bone marrow stromal cells derived from MSC	Human and murine marrow stromal cells (murine BM-derived cell line M2-10B4) were shown to protect CLL cells from spontaneous and drug-induced apoptosis: development of a reliable and reproducible system to assess stromal cell adhesion-mediated drug resistance.	(168)

Blood derived nurse like cells	Blood-derived nurse-like cells protect chronic lymphocytic leukemia B cells from spontaneous apoptosis through stromal cell-derived factor-1.	(169)
Human microvascular endothelial cell line (HMEC-1)	Co-culture of CLL cells with a microvascular endothelial cell line protects cells from apoptosis and upregulates NFκB.	(1)
Human umbilical vein endothelial cells (HUVEC)	Physical contact with endothelial cells through β1- and β2- integrins rescues chronic lymphocytic leukemia cells from spontaneous and drug-induced apoptosis.	(170)
CCL2, CXCL2	Cytokine screen paper highlights the need for accessory cells. The addition of either CXCL2 or CCL2 enhanced CLL cell survival, while antibodies blocking these chemokines reduced survival.	(64)
T cells	Co-culture of pre activated T cells with CLL cells at different T:B cell ratios	(52, 171)
Bone marrow stromal cells (BMSC)	BMSC promote cell survival and drug resistance by modulating the redox status of CLL cells.	(172)

Table 3.1 CLL co-culture model systems and their uses

3.1.1. Comparison of co-culture systems

To date no study has systematically characterised the effects of the different co-culture systems or provided a direct comparison of CLL cell survival, proliferation and phenotype. In collaboration with Professor Chris Pepper (Cardiff University) we sought to compare three co-culture systems designed to mimic the lymph node and vascular microenvironments. The aim was to determine key effects on CLL cells which occur in common in each of the systems tested as well as differences between the systems. Two model systems utilised mouse embryonic fibroblasts transfected with human CD40L (expressed by activated T-lymphocytes) or human CD31 (CD38 ligand) and were carried out by Cardiff University. The third model used human micro vascular endothelial cell line, HMEC-1, described in Chapter 1.

These studies provide an insight into the different signals being provided by different co-culture systems and the cellular components required for an *in vitro* model of the CLL tumour microenvironment. Results of this study were published in the *British Journal of Haematology* (173) and according to the regulations governing PhD Theses at King's College London, I have chosen to present this work in the form published.

3.2 Results

The paper presented below was produced as part of our collaboration with Professor Chris Pepper at Cardiff University. I performed experiments and analysed data from assays using HMEC-1 cells at King's College London and experiments utilising transfected fibroblasts were carried out by the joint first author, Dr Laurence Pearce at the Cardiff University. I contributed to the writing of the Materials and Methods section of the paper and edited the manuscript. Dr Andrea Buggins and Professor Chris Pepper designed the study, analysed data and wrote the rest of the manuscript.

bjh research paper

Mimicking the tumour microenvironment: three different co-culture systems induce a similar phenotype but distinct proliferative signals in primary chronic lymphocytic leukaemia cells

Emma Hamilton,[1†] Laurence Pearce,[2†] Liam Morgan,[2] Sophie Robinson,[2] Vicki Ware,[2] Paul Brennan,[2] N. Shaun B. Thomas,[1] Deborah Yallop,[1] Stephen Devereux,[1] Chris Fegan,[2] Andrea G. S. Buggins[1†] and Chris Pepper[2†]

[1]*Department of Haematology, King's College London, London, and* [2]*Department of Medical Genetics, Haematology and Pathology, School of Medicine, Cardiff University, Cardiff, UK*

Received 16 January 2012; accepted for publication 04 May 2012

Correspondence: Dr Chris Pepper, Department of Haematology, Cardiff University, School of Medicine, Heath Park, Cardiff CF14 4XN, UK. E-mail: peppercj@cf.ac.uk

[†]These authors contributed equally to this work.

Summary

Interactions in the tumour microenvironment can promote chronic lymphocytic leukaemia (CLL) cell survival, proliferation and drug resistance. A detailed comparison of three co-culture systems designed to mimic the CLL lymph node and vascular microenvironments were performed; two were mouse fibroblast cell lines transfected with human CD40LG or CD31 and the third was a human microvascular endothelial cell line, HMEC-1. All three co-culture systems markedly enhanced CLL cell survival and induced a consistent change in CLL cell phenotype, characterized by increased expression of CD38, CD69, CD44 and ITGA4 (CD49d); this phenotype was absent following co-culture on untransfected mouse fibroblasts. In contrast to HMEC-1 cells, the CD40LG and CD31-expressing fibroblasts also induced ZAP70 expression and marked CLL cell proliferation as evidenced by carboxyfluorescein succinimidyl ester labelling and increased Ki-67 expression. Taken together, our data show that co-culture on different stroma induced a remarkably similar activation phenotype in CLL cells but only the CD40LG and CD31-expressing fibroblasts increased ZAP70 expression and CLL cell proliferation, indicating that ZAP70 may play a critical role in this process. This comparative study reveals a number of striking similarities between the co-culture systems tested but also highlights important differences that should be considered when selecting which system to use for *in-vitro* investigations.

Keywords: chronic lymphocytic leukaemia, microenvironment, proliferation, survival, co-culture systems.

Chronic lymphocytic leukaemia (CLL) is a highly heterogenous disease with a very variable clinical outcome. It is now appreciated that it is a highly proliferative disorder with significant tumour cell turnover every day (Messmer *et al*, 2005) yet despite this, primary CLL cells are notoriously difficult to culture *in-vitro* and drug testing models are hindered by the poor survival of these cells. This raises the question as to what signals are provided *in-vivo* that enables these cells to survive and proliferate. Recent work has highlighted the role of accessory cells within the tumour microenvironment in the survival and induction of proliferation in these malignant cells (Patten *et al*, 2008; Buggins *et al*, 2010; Ferretti *et al*, 2011; Herishanu *et al*, 2011). As a result of this, a variety of co-culture systems have been developed to mimic the tumour microenvironment. However, to date no study has systematically characterized the effects of these co-culture systems or provided a direct head to head comparison in terms of CLL cell survival, proliferation and phenotype (Patten *et al*, 2008; Plander *et al*, 2009; Buggins *et al*, 2010; Coscia *et al*, 2011; Ferretti *et al*, 2011; Pepper *et al*, 2011).

Tumour proliferation is believed to mainly occur in pseudofollicles, which develop in the lymph nodes, bone marrow and spleen (Schmid & Isaacson, 1994; Patten *et al*, 2008). Interactions with T-lymphocytes, the microvasculature, soluble factors and other stromal elements are all thought to play a major role in the survival and expansion of the tumour cells. In keeping with this concept, lymph node biopsies from CLL patients with aggressive disease contain activated T-lymphocytes. We have previously demonstrated that proliferating CLL cells co-localize with activated CD4+

T-cells in the lymph node (Patten *et al*, 2008) and ligation of CD40 on CLL cells by its ligand CD40LG (expressed by activated T-lymphocytes) has recently been shown to induce differential responses in terms of up-regulation of surface markers and induction of chemokines (Scielzo *et al*, 2011). Lymph nodes of CLL patients with aggressive disease also contain large numbers of CD31+ vessels (Patten *et al*, 2008) and CD31+ nurse-like cells (Deaglio *et al*, 2005). In a recent study, we demonstrated that interactions with endothelial cells can promote the survival of CLL cells (Buggins *et al*, 2010) and induce the expression of CD38 and ITGA4 (CD49d) on the tumour cells; both of these molecules are associated with aggressive disease and inferior clinical outcome (Damle *et al*, 1999; Shanafelt *et al*, 2008; Majid *et al*, 2011).

These studies indicate that it is interactions with accessory cells, such as activated T-lymphocytes and endothelial cells, that play a role in sustaining CLL cells *in-vivo*. Therefore there is a need to model these interactions *in-vitro* in order to define the critical molecular interactions that promote survival and proliferation in this disease. This study compared three different co-culture systems designed to mimic the lymph node and vascular microenvironments, and compared and contrasted their effects on CLL cells. Two of the model systems utilized were mouse embryonic fibroblasts transfected with human CD40LG or human CD31 (CD40L-TF and CD31-TF) and the third was a microvascular human endothelial cell line, HMEC-1. The aim of these experiments was to compare the effects of these different co-culture systems in order to identify the most appropriate *in-vitro* model system for mimicking the tumour microenvironment.

Materials and methods

Patient samples

Peripheral blood mononuclear cells (PBMCs) from 42 patients with confirmed CLL (Table I) were isolated by density gradient separation (Histopaque-1077; Sigma, Poole, UK). Cells were either cultured fresh or cryopreserved in RPMI 1640 medium, 40% fetal bovine serum (FBS) and 10% dimethylsulfoxide (Sigma) as previously described (Patten *et al*, 2008). Ethical approval was obtained from the local institutional review board of King's College Hospital and South East Wales ethics committee, respectively. In every case, informed written consent was obtained according to the Declaration of Helsinki.

Liquid culture conditions

Chronic lymphocytic leukaemia PBMCs were cultured at 2×10^6/ml in either HMEC-1 recommended medium [M199, 10% (v/v) FBS, L-glutamine (2 mmol/l final), penicillin (2000 units per ml), streptomycin (2 mg/ml), endothelial cell growth supplement (ECGS) (10 μg/ml), hydrocortisone (1 μg/ml), 2-mercaptoethanol (5 μmol/l), human epidermal growth factor (hEGF) (10 ng/ml), ascorbic acid in M199, (1 μg/ml), vascular endothelial growth factor (VEGF) (0·5 ng/ml), insulin-like growth factor 1 (IGF1) (10 ng/ml) supplemented with 1% bovine serum albumin (Sigma)] or Dulbecco's Modified Eagle's Media (DMEM) complete media (DMEM containing 10% fetal calf serum, 2% penicillin plus streptomycin, 1% L-glutamine and 5 ng/ml interleukin 4 [Miltenyi Biotech, Bisley, UK]). Cells were incubated at 37°C in a fully humidified atmosphere of 5% CO_2.

Co-culture conditions

Human microvascular endothelial cells (HMEC-1, Centres for Disease Control and Prevention, Atlanta, GA) were seeded at 10^5/ml in 24-well plates in the recommended medium and incubated overnight to allow cells to adhere. CLL PBMCs were cultured alone and on HMEC-1 cells at 2×10^6/ml as previously described (Buggins *et al*, 2010) and harvested at the time points indicated. Untransfected mouse fibroblasts (NTL) and genetically modified fibroblasts expressing CD31 (CD31-TF a kind gift from Dr Silvia Deaglio, University of Torino & Human Genetics Foundation, Turin, Italy) or CD40LG (CD40L-TF from Dr Aneela Majid, Medical Research Council Toxicology Unit, Leicester University, Leicester, UK) were seeded at 2×10^6/ml in DMEM complete media. The cells were left overnight to adhere to the plates. 2×10^6 CLL cells were added and the co-cultures were left at 37°C, 5% CO_2 for up to 21 d. The supplemented DMEM media was changed every 3–4 d and cells harvested at the time points indicated. All co-cultures were seeded at confluence to ensure optimal contact with primary CLL cells.

Flow cytometry

The phenotype of CLL PBMCs at time 0 and at 24 h following co-culture with HMEC-1 cells, CD40L-TF, CD31-TF, NTL fibroblasts or control medium was analysed by 5-colour flow cytometry. CD38, CD44, ITGA4, ZAP70, CD69, CD11c, CD103 and CD138 expression were determined using CD19-Pacific Blue (PB), CD5-phycoerythrin-cyanin 7 (PECy7), CD38-pycoerythrin (PE), CD44-PE, CD69-allophycocyanin (APC), ZAP70-fluorescein isothiocyanate (FITC) (all eBioscience, Hatfield, UK), ITGA4-FITC (Serotec, Kidlington, UK) and CD11c-FITC, CD103-PE and CD138-APC (Dako, Ely, UK). Apoptosis of cells was assessed after 7 d co-culture using flow cytometry following labelling with CD19-PB, CD5-PeCy7, Annexin V-FITC (Becton Dickinson, Oxford, UK), and 7-amino-actinomycin D (7-AAD; Becton Dickinson) according to the manufacturer's instructions. Proliferation was analysed at various time points by measuring Ki-67 expression using flow cytometry. Cells were labelled with CD19-PB, CD5-PeCy7 before treatment with Fix and Perm reagent (eBioscience) supplemented with 5% Nonidet P-40

Table I. Characteristics of CLL patients in the study.

ID	Binet stage	CD38 (%)	ZAP70 (%)	*IGHV* mutation status	*IGHV* gene usage	FISH
1	B	57·0	ND	Unmutated	*IGHV1-69*	Normal (*TP53* mutation)
2		85·0	ND	Unmutated	*IGHV5-51*	Trisomy 12
3	C	22·0	ND	Unmutated	*IGHV4-4*	ND
4	B	98·0	ND	Unmutated	*IGHV1-69*	Trisomy 12
5	B	3·0	ND	Mutated	*IGHV3-23*	del 13q14·3
6	A	1·0	ND	ND	ND	ND
7	A	48·0	ND	Mutated	*IGHV4-34*	Normal
8	A	ND	ND	ND	ND	ND
9	C	52·0	ND	Unmutated	*IGHV1-2*	del 13q14·3
10	A	49·0	ND	ND	ND	ND
11	A	1·0	ND	ND	ND	del 13q14·3
12	A	2·0	ND	Unmutated	*IGHV2-70*	del 13q14·3
13	B	44·0	ND	ND	ND	Normal
14	B	98·0	ND	Mutated	*IGHV1-8*	del 17p
15	A	49·0	ND	Unmutated	*IGHV4-34*	del 13q14·3/del 17p
16	B	43·0	ND	Mutated	*IGHV1-3*	ND
17	A	43·0	ND	Unmutated	*IGHV4-39*	ND
18	A	1·5	ND	ND	ND	ND
19	A	35·0	ND	ND	ND	Normal
20	B	96·0	ND	Mutated	*IGHV3-21*	del 13q14·3
21	A	3·9	22·2	Mutated	*IGHV2-5*	Normal
22	A	9·8	10·4	Mutated	*IGHV1-2*	del 13q14·3
23	C	5·7	3·4	Mutated	*IGHV4-59*	Normal
24	A	84·8	13·2	Mutated	*IGHV4-34*	Normal
25	A	84·2	1·0	Unmutated	*IGHV3-9*	del 11q
26	A	16·4	1·0	Mutated	*IGHV3-64*	del 13q14·3
27	A	46·0	31·0	Mutated	*IGHV3-74*	Normal
28	A	6·1	1·4	Mutated	*IGHV3-74*	Normal
29	A	100·0	87·0	Mutated	*IGHV3-7*	del 13q14·3
30	A	10·3	92·0	Mutated	*IGHV3-33*	ND
31	A	2·3	2·0	Mutated	*IGHV3-7*	del 13q14·3
32	A	99·7	80·4	Mutated	*IGHV1-3*	ND
33	A	13·0	11·0	Mutated	*IGHV3-9*	del 13q14·3
34	A	8·9	4·6	Mutated	*IGHV3-21*	del 13q14·3
35	C	28·7	58·6	Unmutated	*IGHV3-30*	del 13q14·3/del 17p
36	B	70·4	33·1	Unmutated	*IGHV3-74*	ND
37	A	12·0	20·0	Mutated	*IGHV3-21*	Trisomy 12
38	B	99·5	80·0	Mutated	*IGHV3-21*	del 11q
39	A	5·0	34·0	Mutated	*IGHV4-34*	ND
40	A	0·9	10·3	Mutated	*IGHV3-23*	Normal
41	A	36·0	34·0	Unmutated	*IGHV3-53*	del 13q14·3
42	C	47·0	25·0	Unmutated	*IGHV1-69*	del 11q

and labelling with Ki-67-FITC (Becton Dickinson) or matched isotype control. All antibodies and clones used are shown in Table II.

CFSE labelling of primary CLL cells

Primary CLL cells were labelled with 10 mmol/l carboxyfluorescein succinimidyl ester (CFSE) (Invitrogen, Carlsbad, CA, USA) in 1 ml of DMEM supplemented with 1% FBS. Cells were incubated at 37°C for 10 min before being washed twice in DMEM supplemented with 10% FBS. CLL cells were then seeded into the various co-culture systems at 2×10^6 cells/ml. Cell proliferation was assessed after 7 and 14 d in culture by the decrease in CFSE labelling (compared to day 0 cells) and was quantified using FlowJo 9.3.3 software (Tree-Star Inc., Ashland, OR, USA).

Statistical analysis

All statistical analyses were performed using GraphPad Prism 5.0 (GraphPad Software, La Jolla, CA, USA). All of the paired data were tested for normality and considered Gaussian, so data sets were compared using the paired *t*-test. *P*-values < 0·05 were considered to be statistically significant.

Table II. Antibodies used for flow cytometry.

Antibody	Clone or identifier	Supplier
CD38-PE	HB7	eBioscience
ITGA4-FITC	Bu49	Serotec
CD19-Pacific blue	HIB19	eBioscience
CD5-PE Cy7	UCHT2	eBioscience
CD44-PE	IM7	eBioscience
CD69-APC	FN50	eBioscience
ZAP70-FITC	IE7·2	eBioscience
Ki-67-FITC	51-36524	Becton Dickinson
CD11c -FITC	KB90	Dako
CD103-PE	Ber-ACT8	Dako
CD138-APC	MI 15	Dako

Results

We have previously reported that HMEC-1 cells promote the survival of CLL cells *in-vitro* compared with liquid culture (Buggins *et al*, 2010). We therefore initially compared the survival index of this co-culture system with NTL mouse fibroblasts and that of CD31-TF or CD40L-TF. Primary CLL cells from 42 different CLL patients were analysed for viability following co-culture alone or with NTL mouse fibroblasts, CD40L-TF or CD31-TF or HMEC-1 cells for 7 d. All four co-culture systems promoted survival of the CLL cells when compared with liquid culture (NTL $P = 0{\cdot}01$, CD40L-TF $P = 0{\cdot}0018$, CD31-TF $P = 0{\cdot}0009$ and HMEC-1 $P < 0{\cdot}0001$) with HMEC-1 cells providing the greatest cytoprotection (Fig 1).

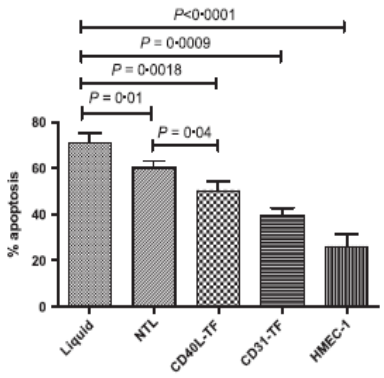

Fig 1. Human microvascular endothelial cell line-1 cells and both CD31-TF and CD40L-TF prevent apoptosis of CLL cells. Primary CLL cells from 21 different CLL patients were cultured alone or co-cultured with untransfected mouse fibroblasts (NTL), CD40L-TF, CD31-TF or HMEC-1 cells for 7 d. CD19+CD5+CLL cells were gated on and apoptosis measured by Annexin V/7-amino-actinomycin D positivity. All the co-culture systems prevented CLL cell apoptosis with HMEC-1 cells giving the best survival.

We then determined whether the different co-culture systems induced similar phenotypic changes in CLL cells by analysing them at an earlier time point of 24 h, chosen because time course studies had shown that the maximal changes in the phenotypic markers included in this study were observed this time point (data not shown). The antibody panels were chosen to quantify the effect of co-culture on the expression of markers associated with adhesion (CD38, CD44 and ITGA4), activation (CD38 and CD69), migration (ITGA4) and hairy cell leukaemia/plasmacytoid differentiation (CD103, CD11c and CD138). We also analysed the expression of ZAP70 due to its association with progressive disease (Rassenti *et al*, 2004). We have already shown that CD38 and ITGA4 are up-regulated following co-culture with HMEC-1 cells (Buggins *et al*, 2010). Here, we showed that both the transfected fibroblast cell lines did the same, with the CD31-TF cells being the most potent stimulator (CD38: CD40L-TF $P < 0{\cdot}0001$, CD31-TF $P < 0{\cdot}0001$ and HMEC-1 $P = 0{\cdot}0464$. ITGA4: CD40L-TF $P < 0{\cdot}0001$, CD31-TF $P < 0{\cdot}0001$ and HMEC-1 $P = 0{\cdot}031$; Figs 2A and 3A,B). This is perhaps not surprising as CD31 is the only known ligand for CD38 and ITGA4 is known to co-localize with CD38 (Buggins *et al*, 2011). In contrast, the NTL mouse fibroblasts failed to significantly induce these phenotypic changes in the absence of the human ligands but maintained the expression of the antigens at similar levels to those measured at time 0.

In addition, CD44 and CD69 were also up-regulated by all three co-culture systems to a similar level (CD44: CD40L-TF $P < 0{\cdot}0001$, CD31-TF $P < 0{\cdot}0001$ and HMEC-1 $P = 0{\cdot}0004$ CD69: CD40L-TF $P < 0{\cdot}0001$, CD31-TF $P < 0{\cdot}0001$ and HMEC-1 $P < 0{\cdot}0001$ Figs 2B and 3C,D). Again, the NTL mouse fibroblasts failed to significantly induce these phenotypic changes above the expression of the antigens as measured at time 0. Interestingly, ZAP70, which has been associated with a similar expression pattern as CD38 (Damle *et al*, 2007), was only up-regulated by the transfected fibroblasts and unaffected by co-culture with HMEC-1 cells (CD40L-TF $P < 0{\cdot}0001$, CD31-TF $P < 0{\cdot}0001$ and HMEC-1 $P = 0{\cdot}437$; Figs 2C and 3E). The two markers of differentiation towards a hairy cell phenotype (CD103 and CD11c) were not significantly induced by co-culture with NTL, HMEC-1 and CD31-TF cells. In contrast, the CD40L-TF cells induced a small but significant increase in in CD11c but not CD103 ($P = 0{\cdot}008$ and $P = 0{\cdot}092$; Figure S1A and S1B respectively). Although the change in CD11c was statistically significant, the mean fluorescence intensity values for this antigen remained very low compared with other antigens measured in this study, even after co-culture with CD40L-TF cells. As CLL cells cultured on the CD40L-TF and CD31-TF co-culture systems had a notable increase in the forward scatter, suggesting they could potentially be in transformation to a plasmacytoid phenotype, we looked at CLL expression of the plasmacytoid marker CD138 in these systems. Neither system induced significant changes in the expression of the CD138 (Figure S1C).

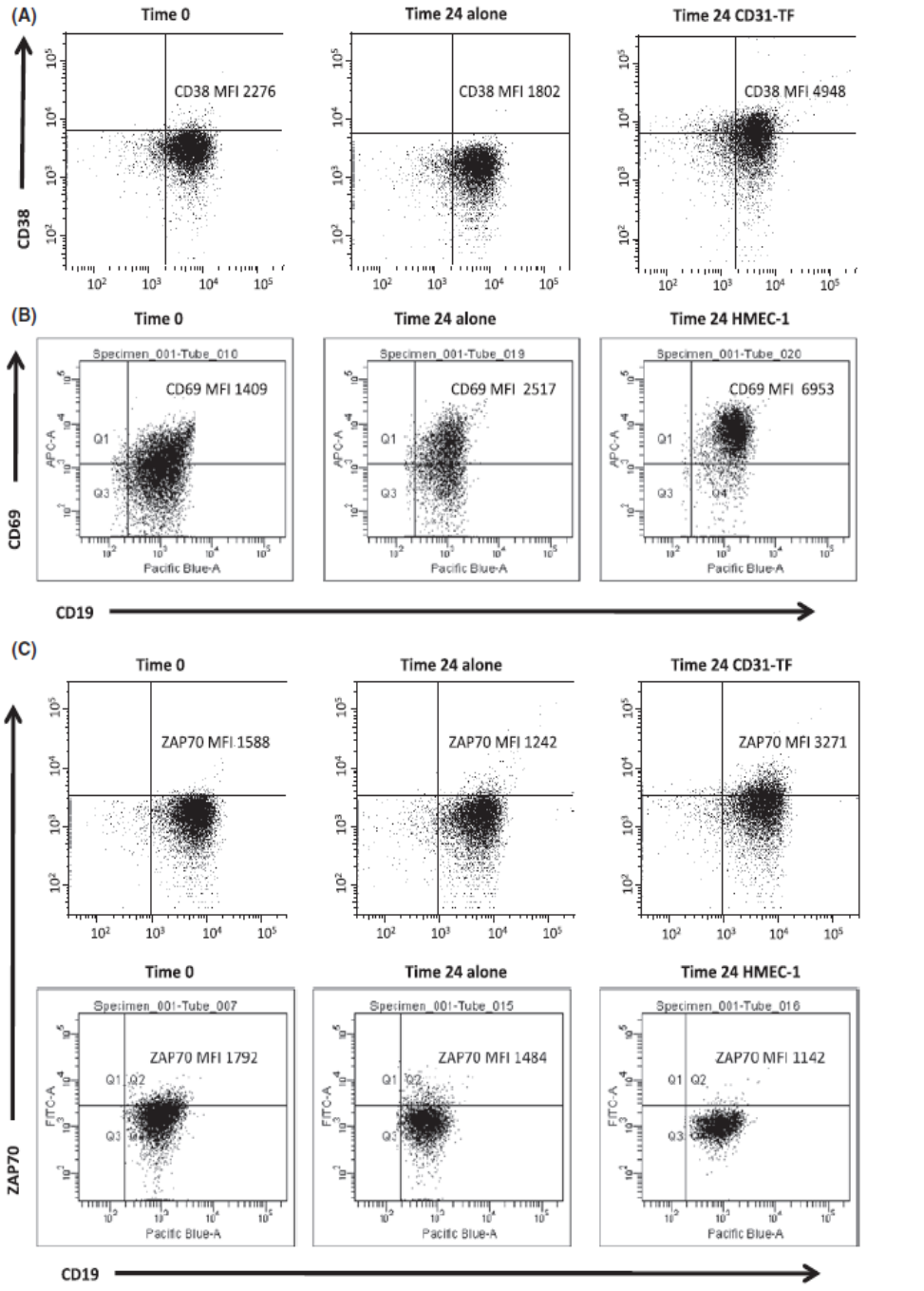

Fig 2. Representation flow cytometry plots demonstrating the phenotypic changes induced in CLL cells following co-culture. Primary CLL cells from 35 different CLL patients were co-cultured alone, on NTL mouse fibroblasts or with CD40L-TF, CD31-TF or HMEC-1 cells for 24 h. At time 0 and time 24 h, cells were stained for flow cytometry and CD19+CD5+CLL cells were gated on and levels of CD38, CD69 and ZAP70 measured. (A) and (B) are representative flow cytometry plots showing that co-culture with CD31-TF caused increased levels of CLL CD38 expression and co-culture with HMEC-1 cells increased levels of the activation marker CD69. (C) illustrates that co-culture with CD31-TF induced an increase in CLL ZAP70 expression but co-culture on HMEC-1 cells had no effect on ZAP70 expression. MFI, mean fluorescence intensity.

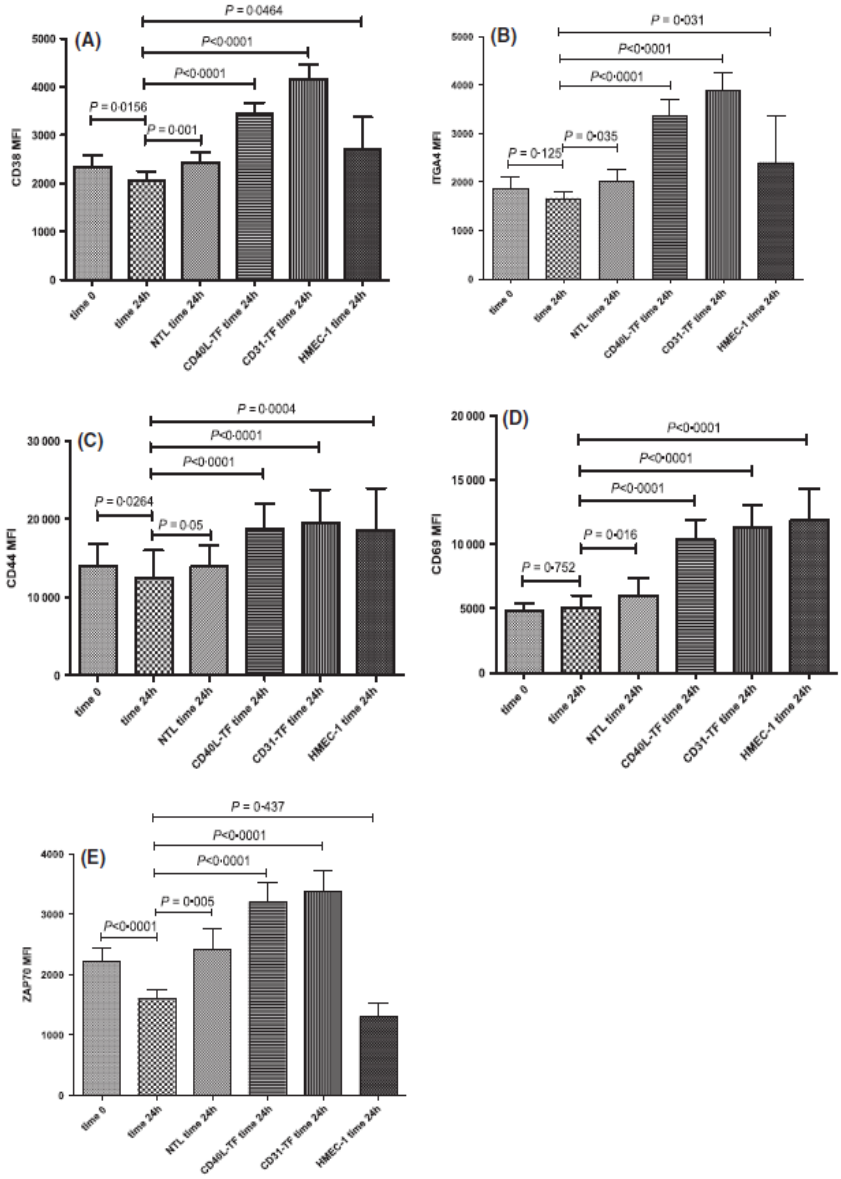

Fig 3. Human microvascular endothelial cell line-1 cells and both CD31-TF and CD40L-TF increase expression of CD38, ITGA4, CD44 and CD69 by CLL cells. Primary CLL cells from 35 different CLL patients were cultured alone, or co-cultured on NTL, CD40L-TF, CD31-TF or HMEC-1 cells for 24 h. At time 0 and time 24 h, cells were stained for flow cytometry and CD19+CD5+CLL cells were gated on and levels of CD38, ITGA4, CD44, CD69 and ZAP70 measured. (A–E) show composite bar charts demonstrating the results from all the primary CLL cells assayed. All four co-culture cell types induced increased levels of CLL expression of (A) CD38, (B) ITGA4, (C) CD44 and (D) CD69 with the most notable effect on CD38 and CD49 induced by CD31-TF. (E) Only CD31-TF and CD40L-TF were able to induce an increase in CLL expression of ZAP70. The NTL (control) mouse fibroblasts did not induce the same phenotypic changes. MFI, mean fluorescence intensity.

To analyse the proliferation of the CLL cells in the different co-culture systems, longer-term co-culture assays of primary CLL cells from six different patients on all the co-culture systems were set up. In all of the systems the CLL cells clustered around the confluent 'feeder' cells, but on the HMEC-1 cells and the NTL mouse fibroblasts they remained the same size, whereas on the transfected fibroblasts they showed signs of blasting and proliferation (Fig 4A). As it has already been

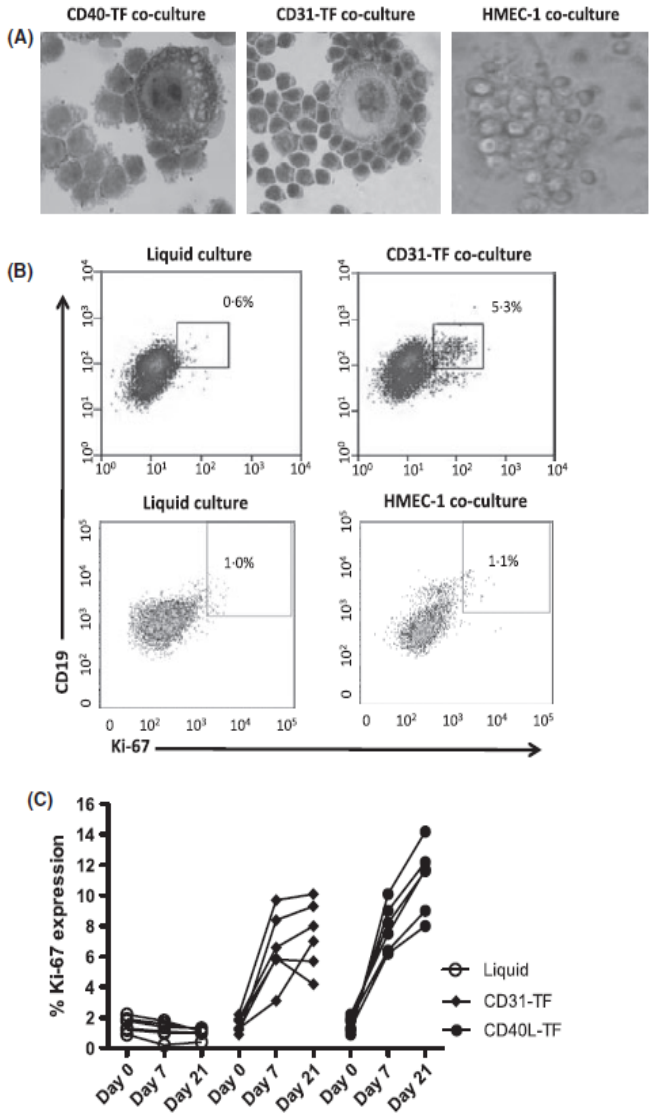

Fig 4. Both CD31-TF and CD40L-TF induce proliferation in CLL cells but HMEC-1 cells do not. Primary CLL cells from six different patients were co-cultured for up to 21 d alone or with CD40L-TF, CD31-TF or HMEC-1 cells. CLL cell samples were removed at day 0, 7 and 21 and CD5/CD19+ cells analysed for expression of Ki-67. (A) Photographs of primary CLL cells clustered on CD40L-TF, CD31-TF and HMEC-1 cells taken on a Zeiss Axio microscope using a ×100 objective lens. (B) A representative figure demonstrating the increase in Ki-67 expression in CD19+ CLL cells following co-culture with CD31-TF compared to liquid culture and HMEC-1 co-culture. (C) At day 7 both CD40L-TF and CD31-TF induced an increase in expression of Ki-67 in CLL cells from all six patients tested compared to liquid culture. At day 21 CD40L-TF induced a further increase in Ki-67 expression; this was less marked in those cells co-cultured with the CD31-TF. HMEC-1 cells did not induce Ki-67 expression in any CLL cells at the same time-points.

shown that the proliferation marker Ki-67 is upregulated in the CLL cells from the lymph nodes compared to the peripheral blood in these patients (Herishanu *et al*, 2011), we compared Ki-67 levels of CLL cells in our different systems. Analysis of Ki-67 on day 7 clearly showed that HMEC-1 cells did not induce Ki-67 expression by CLL cells in all six patients tested, whereas both sets of transfected fibroblasts induced Ki-67 expression at day 7, with CD40L-TF cells inducing a further increase at day 21 (Fig 4B,C). It is interesting to note that CD31-TF cells induced proliferation whereas HMEC-1 cells,

which express very low levels of CD31, could not. CD31-TF cells have a much higher density of CD31 expression than HMEC-1 cells (data not shown), which suggests that ligand density may be an important factor in driving CLL proliferation, at least in the context of CD31 signalling.

These results suggest that, although CD40L-TF cells afforded the least cytoprotection to CLL cells when compared with HMEC-1 cells or CD31-TF cells, they were the most potent inducers of CLL proliferation. Indeed, these two phenomena may be linked; cells with a higher rate of turnover may have an increased propensity to die. To further assess the effect on cell division we compared the number of divisions seen in CFSE-labelled CLL cells co-cultured on NTL mouse fibroblasts, CD40L-TF or CD31-TF. As expected, following 21 d in culture, the CLL cells co-cultured on NTL fibroblasts showed little evidence of CLL cell proliferation (Fig 5A). In contrast CD40L-TF and CD31-TF co-cultures induced significant proliferation with CD40L-TF inducing more cell divisions than those co-cultured on CD31-TF (Fig 5B,C, respectively). It is noteworthy that there appeared to be no difference in the ability of CD40L-TF to induce proliferation in samples derived from patients in different Binet stages (Fig 5D).

Discussion

It is now widely accepted that interactions in the tumour microenvironment can promote CLL cell survival,

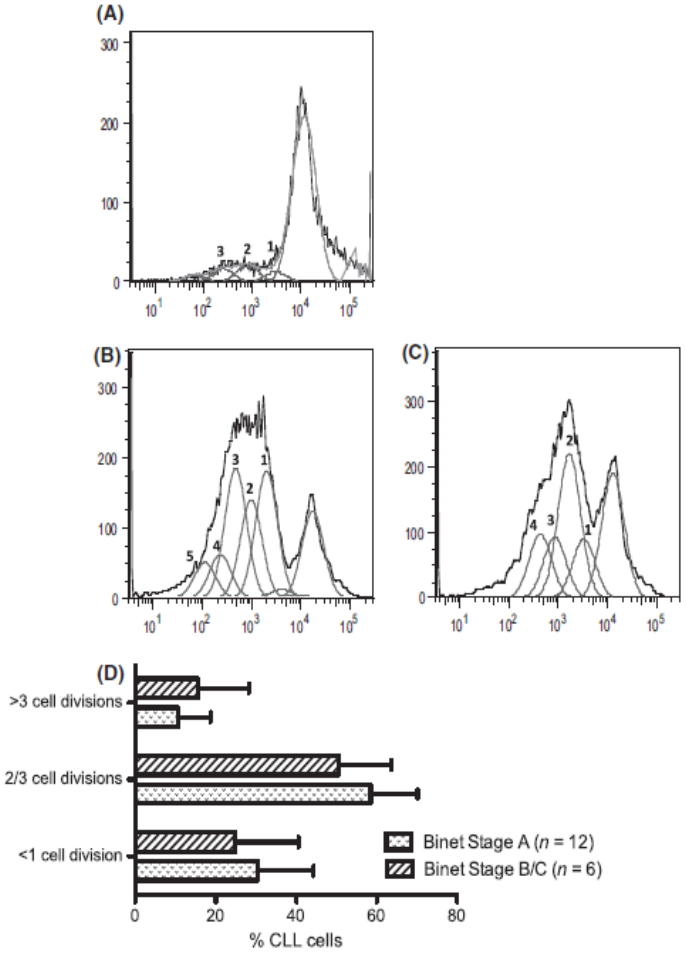

Fig 5. CD40L-TF induced more CLL cell proliferation than CD31-TF. CLL cells from 18 different patients were labelled with carboxyfluorescein succinimidyl ester and co-cultured on (A) NTL mouse fibroblasts, (B) CD40L-TF or (C) CD31-TF. The representative figures demonstrate that after 2 weeks of co-culture CLL cells cultured on NTL fibroblasts showed very little cell division. Numbers of cell divisions was calculated using FlowJo software and are shown on each figure. In contrast, the sample CLL cells cultured with CD40L-TF and CD31-TF showed significant cell division with those on CD40L-TF showing the most proliferation. (D) Shows that the extent of proliferation induced by CD40L-TF was not dependent on Binet stage.

proliferation and drug resistance (Patten *et al*, 2005, 2008; Plander *et al*, 2009; Buggins *et al*, 2010; Pepper *et al*, 2011). Although the key molecular events that drive these processes are not fully defined, we have previously shown that lymph nodes from patients with aggressive disease contain increased numbers of CD31+ vessels and activated T-lymphocytes and that interactions with vascular endothelial cells enhance CLL cell survival. The present study describes a detailed comparison of three co-culture systems designed to mimic the CLL lymph node and vascular microenvironments; two utilized mouse embryonic fibroblasts transfected with human CD40LG (expressed by activated T-lymphocytes) or human CD31 (CD38 ligand) and the third was a microvascular endothelial cell line, HMEC-1.

All three systems were highly cytoprotective to CLL cells when compared with liquid culture. Interestingly the human microvascular cell line, HMEC-1, induced the best survival signals. This is in keeping with previous work, which suggests that interaction with the microvasculature plays a key role in the maintenance of CLL cell survival *in-vivo* (Zucchetto *et al*, 2009; Buggins *et al*, 2010). The fact that the CD31-TF cells were not as effective stimulators of survival (despite having higher CD31 antigen density) suggests that CD31 is not the sole anti-apoptotic signal in the microvasculature and/or that CD31 ligand density needs to be lower for optimal cytoprotection. The CD40L-TF cells provided the least cytoprotection of all the co-culture systems under evaluation. This indicates that the CD40-CD40LG interaction is not a major player with regard to the prevention of apoptosis in the microenvironment and/or the expression of CD40LG in this model is too high for optimal anti-apoptotic signalling.

Surprisingly, all three co-culture systems induced a remarkably similar CLL cell phenotype and were able to up-regulate markers of adhesion, activation and migration. This suggests that co-culture mediated phenotypic changes are not ligand-specific but rather represent a default setting for CLL cells following interaction with many of the cells and receptors associated with the tumour microenvironment. Importantly, the NTL mouse fibroblasts used as controls in this study failed to induce these phenotypic changes suggesting that interaction of CLL cells with human CD31 and CD40LG are critical to the effects seen in CD40L-TF and CD31-TF. The notable exception in terms of phenotypic change was the ability of CD31-TF cells and CD40L-TF cells to induce ZAP70 expression. In contrast, the HMEC-1 cells were unable to induce the expression of this tyrosine kinase suggesting that endothelial cell interactions may not be the principle determinant of the maintenance ZAP70 expression in CLL cells. Given that endothelial cells are unable to induce both ZAP70 expression and CLL cell proliferation, this data suggests that there is a possible role for this tyrosine kinase in CLL cell division.

In contrast to its relatively modest effect on *in-vitro* survival, CD40LG appeared to provide the strongest and most sustained proliferation signals to CLL cells. On the contrary, HMEC-1 cells failed to induce any proliferation under the conditions tested. This is perhaps not surprising as CD40LG is expressed by activated T cells and co-culture with these is known to induce CLL cell division (Patten *et al*, 2005, 2008). Indeed, one hypothesis for the inferior cytoprotection afforded by CD40L-TF cells is that the strong proliferative response they induce causes CLL cells to become more susceptible to cell death. Ligation of CD38 is also known to induce proliferation so it was of considerable interest that the CD31-TF also induced CLL cell division. However, HMEC-1 cells, which express low levels of CD31, were unable to cause even a marginal increase in CLL cell Ki-67 expression. This raises the question as to the role of ligand density in these interactions and highlights the possibility that many ligand/receptor interactions may induce CLL cell proliferation providing the density is sufficient. Therefore, one interpretation of our data is that HMEC-1 cells express insufficient CD31 to induce CLL cell proliferation but the low level stimulation they provide is ideal for the prevention of apoptosis. This hypothesis is supported by our previous work demonstrating that the anti-apoptotic effect of endothelial cells on CLL cells is via the induction of NF-KB (nuclear factor κB) regulated genes such as *BCL2*, *BCL2L1*, *MCL1*, *CD38* and *ITGA4* (Buggins *et al*, 2010). CLL cell co-culture with endothelial cells promoted their survival over a 7-d period that was associated with a small but statistically significant increased CD38 expression. In contrast, our previous work has shown that when CLL cells are co-cultured with activated autologous T cells they induce a dramatic increase in CD38 expression (Patten *et al*, 2008). However, over a 7-d period these CLL cells died by apoptosis, suggesting that this level of activation is incompatible with long-term CLL cell survival (unpublished data).

These results highlight the importance of the choice of co-culture model and, for the laboratory constructed systems, the importance of designing systems with ligand densities at a physiological level. As culture of CLL cells *in-vitro* is problematic due to the dependence of these tumour cells on interactions with the surrounding tissues, it is vital that a model that accurately mimics the microenvironment is established. Our findings indicate that the best model system for mimicking the lymph node microenvironment is the CD40L-TF cells; this co-culture system induces an activation phenotype and a strong proliferative signature consistent with recently published gene expression profiles from CLL cells derived from lymph nodes (Herishanu *et al*, 2011). In contrast, when modelling the CLL cell interactions in the microvasculature our data suggest that the HMEC-1 cells represent the most realistic model system, characterized by enhanced survival in the absence of proliferation. It seems likely that further refinement of these *in-vitro* models will lead to improved *in-vitro* drug testing platforms.

Acknowledgements

This work was supported by grants from Leukaemia & Lymphoma Research, the Medical Research Council and Leukaemia Research Appeal for Wales. CP is also supported by the National Institute for Social Care and Health Research (NISCHR) through the Cancer Genetics Biomedical Research Unit.

Authorship contribution

EH carried out the experimental work, analysed data and edited the manuscript; LP carried out the experimental work, analysed data and edited the manuscript; LM carried out the experimental work and analysed data; SR carried out the experimental work and analysed data; VW carried out the experimental work and analysed data; PB analysed data and edited the manuscript; NSBT analysed data and edited the manuscript; DY provided clinical samples, analysed data and edited the manuscript; SD provided clinical samples, analysed data and edited the manuscript; CF provided clinical samples, analysed data and edited the manuscript; AGSB and CP jointly conceived and supervized the study, analysed the data and wrote the manuscript.

Conflict of interest

EH, LP, LM, SR, VW, PB, NSBT, DY, SD, CF, AGSB and CP declare no conflict of interests.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig S1. The three cell lines induced little evidence of increase expression of the differentiation markers CD11c, CD103 and CD138 on CLL cells.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

Buggins, A.G., Pepper, C., Patten, P.E., Hewamana, S., Gohil, S., Moorhead, J., Folarin, N., Yallop, D., Thomas, N.S., Mufti, G.J., Fegan, C. & Devereux, S. (2010) Interaction with vascular endothelium enhances survival in primary chronic lymphocytic leukemia cells via NF-kappaB activation and de novo gene transcription. *Cancer Research*, **70**, 7523–7533.

Buggins, A.G., Levi, A., Gohil, S., Fishlock, K., Patten, P.E., Calle, Y., Yallop, D. & Devereux, S. (2011) Evidence for a macromolecular complex in poor prognosis CLL that contains CD38, CD49d, CD44 and MMP-9. *British Journal of Haematology*, **154**, 216–222.

Coscia, M., Pantaleoni, F., Riganti, C., Vitale, C., Rigoni, M., Peola, S., Castella, B., Foglietta, M., Griggio, V., Drandi, D., Ladetto, M., Bosia, A., Boccadoro, M. & Massaia, M. (2011) IGHV unmutated CLL B cells are more prone to spontaneous apoptosis and subject to environmental prosurvival signals than mutated CLL B cells. *Leukemia*, **25**, 828–837.

Damle, R.N., Wasil, T., Fais, F., Ghiotto, F., Valetto, A., Allen, S.L., Buchbinder, A., Budman, D., Dittmar, K., Kolitz, J., Lichtman, S.M., Schulman, P., Vinciguerra, V.P., Rai, K.R., Ferrarini, M. & Chiorazzi, N. (1999) Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*, **94**, 1840–1847.

Damle, R.N., Temburni, S., Calissano, C., Yancopoulos, S., Banapour, T., Sison, C., Allen, S.L., Rai, K.R. & Chiorazzi, N. (2007) CD38 expression labels an activated subset within chronic lymphocytic leukemia clones enriched in proliferating B cells. *Blood*, **110**, 3352–3359.

Deaglio, S., Vaisitti, T., Bergui, L., Bonello, L., Horenstein, A.L., Tamagnone, L., Boumsell, L. & Malavasi, F. (2005) CD38 and CD100 lead a network of surface receptors relaying positive signals for B-CLL growth and survival. *Blood*, **105**, 3042–3050.

Ferretti, E., Bertolotto, M., Deaglio, S., Tripodo, C., Ribatti, D., Audrito, V., Blengio, F., Matis, S., Zupo, S., Rossi, D., Ottonello, L., Gaidano, G., Malavasi, F., Pistoia, V. & Corcione, A. (2011) A novel role of the CX(3)CR1/CX(3)CL1 system in the cross-talk between chronic lymphocytic leukemia cells and tumor microenvironment. *Leukemia*, **25**, 1268–1277.

Herishanu, Y., Perez-Galan, P., Liu, D., Biancotto, A., Pittaluga, S., Vire, B., Gibellini, F., Njuguna, N., Lee, E., Stennett, L., Raghavachari, N., Liu, P., McCoy, J.P., Raffeld, M., Stetler-Stevenson, M., Yuan, C., Sherry, R., Arthur, D.C., Maric, I., White, T., Marti, G.E., Munson, P., Wilson, W. H. & Wiestner, A. (2011) The lymph node microenvironment promotes B-cell receptor signaling, NF-kappaB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood*, **117**, 563–574.

Majid, A., Lin, T.T., Best, G., Fishlock, K., Hewamana, S., Pratt, G., Yallop, D., Buggins, A. G., Wagner, S., Kennedy, B.J., Miall, F., Hills, R., Devereux, S., Oscier, D.G., Dyer, M.J., Fegan, C. & Pepper, C. (2011) CD49d is an independent prognostic marker that is associated with CXCR4 expression in CLL. *Leukemia Research*, **35**, 750–756.

Messmer, B.T., Messmer, D., Allen, S.L., Kolitz, J. E., Kudalkar, P., Cesar, D., Murphy, E.J., Koduru, P., Ferrarini, M., Zupo, S., Cutrona, G., Damle, R.N., Wasil, T., Rai, K.R., Hellerstein, M.K. & Chiorazzi, N. (2005) In vivo measurements document the dynamic cellular kinetics of chronic lymphocytic leukemia B cells. *Journal of Clinical Investigation*, **115**, 755–764.

Patten, P., Devereux, S., Buggins, A., Bonyhadi, M., Frohlich, M. & Berenson, R.J. (2005) Effect of CD3/CD28 bead-activated and expanded T cells on leukemic B cells in chronic lymphocytic leukemia. *Journal of Immunology*, **174**, 6562–6563 author reply 6563.

Patten, P.E., Buggins, A.G., Richards, J., Wotherspoon, A., Salisbury, J., Mufti, G.J., Hamblin, T.J. & Devereux, S. (2008) CD38 expression in chronic lymphocytic leukemia is regulated by the tumor microenvironment. *Blood*, **111**, 5173–5181.

Pepper, C., Mahdi, J.G., Buggins, A.G., Hewamana, S., Walsby, E., Mahdi, E., Al-Haza'a, A., Mahdi, A.J., Lin, T.T., Pearce, L., Morgan, L., Bowen, I. D., Brennan, P. & Fegan, C. (2011) Two novel aspirin analogues show selective cytotoxicity in primary chronic lymphocytic leukaemia cells that is associated with dual inhibition of Rel A and COX-2. *Cell Proliferation*, **44**, 380–390.

Plander, M., Seegers, S., Ugocsai, P., Diermeier-Daucher, S., Ivanyi, J., Schmitz, G., Hofstadter, F., Schwarz, S., Orso, E., Knuchel, R. & Brockhoff, G. (2009) Different proliferative and survival capacity of CLL-cells in a newly established in vitro model for pseudofollicles. *Leukemia*, **23**, 2118–2128.

Rassenti, L.Z., Huynh, L., Toy, T.L., Chen, L., Keating, M.J., Gribben, J.G., Neuberg, D.S., Flinn, I.W., Rai, K.R., Byrd, J.C., Kay, N.E., Greaves, A., Weiss, A. & Kipps, T.J. (2004) ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. *New England Journal of Medicine*, **351**, 893–901.

Schmid, C. & Isaacson, P.G. (1994) Proliferation centres in B-cell malignant lymphoma, lympho-

cytic (B-CLL): an immunophenotypic study. *Histopathology*, **24**, 445–451.

Scielzo, C., Apollonio, B., Scarfo, L., Janus, A., Muzio, M., Ten Hacken, E., Ghia, P. & Caligaris-Cappio, F. (2011) The functional in vitro response to CD40 ligation reflects a different clinical outcome in patients with chronic lymphocytic leukemia. *Leukemia*, **25**, 1760–1767.

Shanafelt, T.D., Geyer, S.M., Bone, N.D., Tschumper, R.C., Witzig, T.E., Nowakowski, G. S., Zent, C.S., Call, T.G., Laplant, B., Dewald, G. W., Jelinek, D.F. & Kay, N.E. (2008) CD49d expression is an independent predictor of overall survival in patients with chronic lymphocytic leukaemia: a prognostic parameter with therapeutic potential. *British Journal of Haematology*, **140**, 537–546.

Zucchetto, A., Benedetti, D., Tripodo, C., Bomben, R., Dal Bo, M., Marconi, D., Bossi, F., Lorenzon, D., Degan, M., Rossi, F.M., Rossi, D., Bulian, P., Franco, V., Del Poeta, G., Deaglio, S., Gaidano, G., Tedesco, F., Malavasi, F. & Gattei, V. (2009) CD38/CD31, the CCL3 and CCL4 chemokines, and CD49d/vascular cell adhesion molecule-1 are interchained by sequential events sustaining chronic lymphocytic leukemia cell survival. *Cancer Research*, **69**, 4001–4009.

3.3 Discussion

In addition to the detailed Discussion in the paper, the work described above highlights the importance of considering the effect of co-culture systems on primary CLL cells and what the implications of this are for experiments carried out using these systems. Co-culture systems should be chosen to reflect the question being addressed, since different systems may best reflect different compartments of the CLL microenvironment. For example, it has been suggested that CLL cells exist in distinct compartments, namely the proliferative compartment in the lymph node (56) and the resting compartment in the vasculature (174). We suggest that the endothelial co-culture system models the resting compartment, however in the peripheral blood the time spent in direct contact with endothelial cells is small as the cells are not stationary. This problem is being addressed with the development of a new dynamic co-culture circulation system designed to mimic the interaction between CLL cells and endothelial cells in the peripheral vasculature. I have carried out preliminary work using this system (175) and the data is in a paper currently being reviewed for publication.

Chapter 4

Transcriptional effects of endothelial cell co-culture on CLL cells

4. Transcriptional effects of endothelial cell co-culture on CLL cells

4.1 Introduction

As described in **Table 3.1** (Chapter 3), laboratories across the world have utilised numerous different co-culture model systems to maintain primary CLL cell viability and investigate the effects of accessory cells on CLL cells. A focus of work in our group has been to investigate the supportive effects of endothelial cells in co-culture. We reported cytoprotective effects and changes in CLL cell phenotype after co-culture with the HMEC-1 endothelial cell line (1, 61, 173), which is further described in Chapter 3. It is likely that direct cell:cell interactions in the microenvironment and indirect interactions, such as through the release of soluble factors from CLL or accessory cells act in concert to produce a cytoprotective effect (65, 156, 176, 177). Whether the cytoprotective stimulus comes from direct or indirect interactions, the effects on the CLL cells will affect both transcriptional and post-transcriptional mechanisms.

In this chapter, the mechanisms involved in CLL cell cytoprotection are investigated by examining the effects of co-culture on mRNA expression in the CLL cells by gene expression microarray analysis. We expand on the study described in Chapter 3, which focused on the effects of endothelial cell co-culture using a cell line, HMEC-1. Here, we also include primary endothelial cells HDBEC and HDMEC, which may provide a closer representation of endothelial cells found in the body than immortalised endothelial cell lines.

Both the primary and immortalised endothelial co-culture systems were found to protect CLL cells from apoptosis (see section 4.2.1) and induce an activated CLL cell phenotype (see section 4.2.2). The aim of work described later in this chapter was to investigate the transcriptional effects of co-culture with endothelial cells and to identify mRNAs which were up-regulated by both systems. Bioinformatic analyses were then performed to identify mechanisms that occur in common between CLL cells cultured in different co-culture systems which may provide rational therapeutic targets to abrogate the cytoprotective effect of the tumour microenvironment.

4.1.1. Experimental approach to investigate the transcriptional effects of co-culture with endothelial cells on CLL cells

Previous work from our group showed that CLL cell viability was increased with endothelial cell co-culture regardless of CLL cell phenotype or prognostic marker status (1) and this observation is supported by the work presented in Chapter 3. Therefore, we chose to investigate the transcriptional

effects of endothelial cell co-culture using CLL samples with a variety of prognostic markers. The CLL patient samples used in this study can be found in **Table 4.1**.

Sample number	LSL number	IgVH Mutational status	CD38%
1	LSL 4631	mutated	96
2	LSL6811	unmutated	98
3	LSL6917	unmutated	49
4	LSL4589	mutated	48
6	LSL5539	unmutated	2
10	LSL5195	mutated	0

Table 4.1 CLL patient samples used in the gene expression microarray study.

$CD5^+CD19^+$ CLL cells were purified from CLL PBMCs which had previously been frozen, as described in sections 2.24 and 2.2.10 (Materials and Methods) and samples with greater than 98% $CD5^+CD19^+$ purity were used for the assay. Sample purity was determined by staining with fluorochrome-conjugated antibodies for CD5 and CD19 and quantifying the percentage of CD5/CD19 double-positive cells by flow cytometry. The cells purified from each of the six CLL patients were split between conditions, as shown in **Figure 4.1**. RNA was isolated as described in section 2.5 (Materials and Methods) at time 0 hours and after 12 hours in liquid culture, 12 hours co-culture with the HMEC-1 cell line and 12 hours co-culture with HDBEC primary endothelial cells. RNA was isolated at a 12 hour time point in order to ensure that good quality RNA could be isolated from CLL cells in the liquid only culture. After longer time points cell death would occur in this liquid only culture, releasing ribonucleases (RNases) and resulting in degradation of sample RNA. CLL mRNA expression was analysed using Human Gene 1.0 ST (Affymetrix) microarrays.

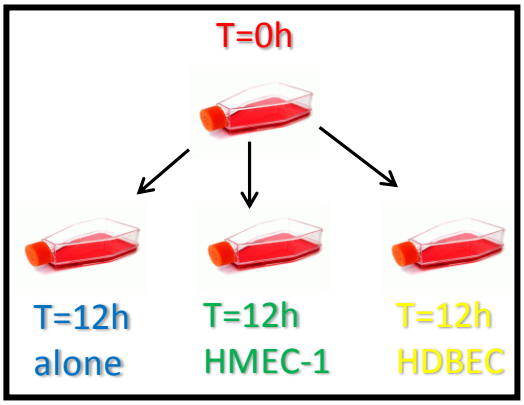

Figure 4.1 Experimental approach to investigating transcriptional effects in common in endothelial co-culture systems. $CD5^{+}CD19^{+}$ CLL cells from six different patients were divided between each of four different conditions to compare changes in transcription associated with cytoprotective endothelial co-culture.

4.2 Results

4.2.1. Co-culturing CLL cells on endothelial cells increases CLL cell viability

To investigate whether the viability of cells isolated from CLL patients was affected by co-culture with different endothelial cells, an assay was established using Annexin V and 7AAD, as described in section 2.3.5 (Materials and Methods). Cultures were set up using $CD5^{+}CD19^{+}$ cells selected from six CLL patients, as described in section 2.3.9 (Materials and Methods). The purity of all samples after selection was greater than 98%, which was determined by staining with fluorochrome-conjugated antibodies for CD5 and CD19 and quantifying the percentage of CD5/CD19 double-positive cells by flow cytometry. The CLL samples were cultured alone and with each of the endothelial cell types. In each case, $2x10^{6}$ viable CLL cells were co-cultured with $1x10^{5}$ endothelial cells, the same ratio as used for assays in Chapter 3. In preliminary experiments, this ratio of endothelial cells protected CLL cells from apoptosis and prevented outgrowth of endothelial cells by day 5. The endothelial cells are adherent and the CLL cells remained in suspension, allowing CLL cell removal with gentle pipetting. If the endothelial cells were allowed to overgrow, they would detach from the culture vessel and contaminate the CLL cell suspension. Samples were taken after 5 days of co-culture and the viability of the CLL cells was determined by flow cytometry of Annexin V and 7AAD by gating on $CD5^{+}CD19^{+}$ cells. The data in **Figure 4.2** show that co-culturing CLL cells with each of the endothelial cell types tested increased the viability of CLL cells compared with CLL cells cultured alone.

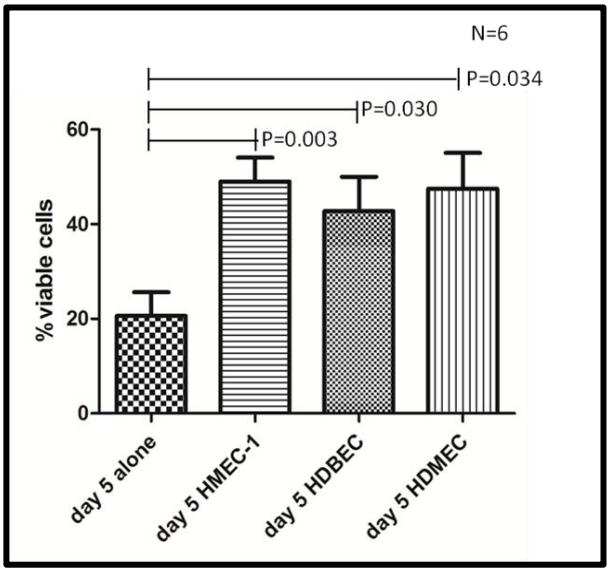

Figure 4.2 CLL cell viability analysis by flow cytometry. The viability of CLL cells cultured alone, on HMEC-1, HDBEC and HDMEC for 5 days were analysed by flow cytometry using Annexin V-FITC and 7AAD (paired t-test, mean±SEM, experiments with cells from n=6 different CLL patients)**.**

4.2.2. Endothelial cell co-culture induces a phenotypic change in CLL cells

To determine whether the co-culture systems have an effect on the phenotype of these viable CLL cells, CLL cells were isolated from 6 patients (the same patients used in the assay described above) and cultured for 24 hours alone and in the HMEC-1, HDBEC and HDMEC co-culture systems and flow cytometry was used to assess a panel of activation (CD69) and negative prognostic markers (CD44, CD38). **Figure 4.3** shows that statistically significant changes in expression of CD44 are observed in CLL cells in all co-culture systems compared with the paired sample of CLL cells cultured alone and statistically significant changes in expression of CD69 and CD38 were observed in CLL cells co-cultured with HMEC-1 and HDBEC compared with CLL cells cultured alone.

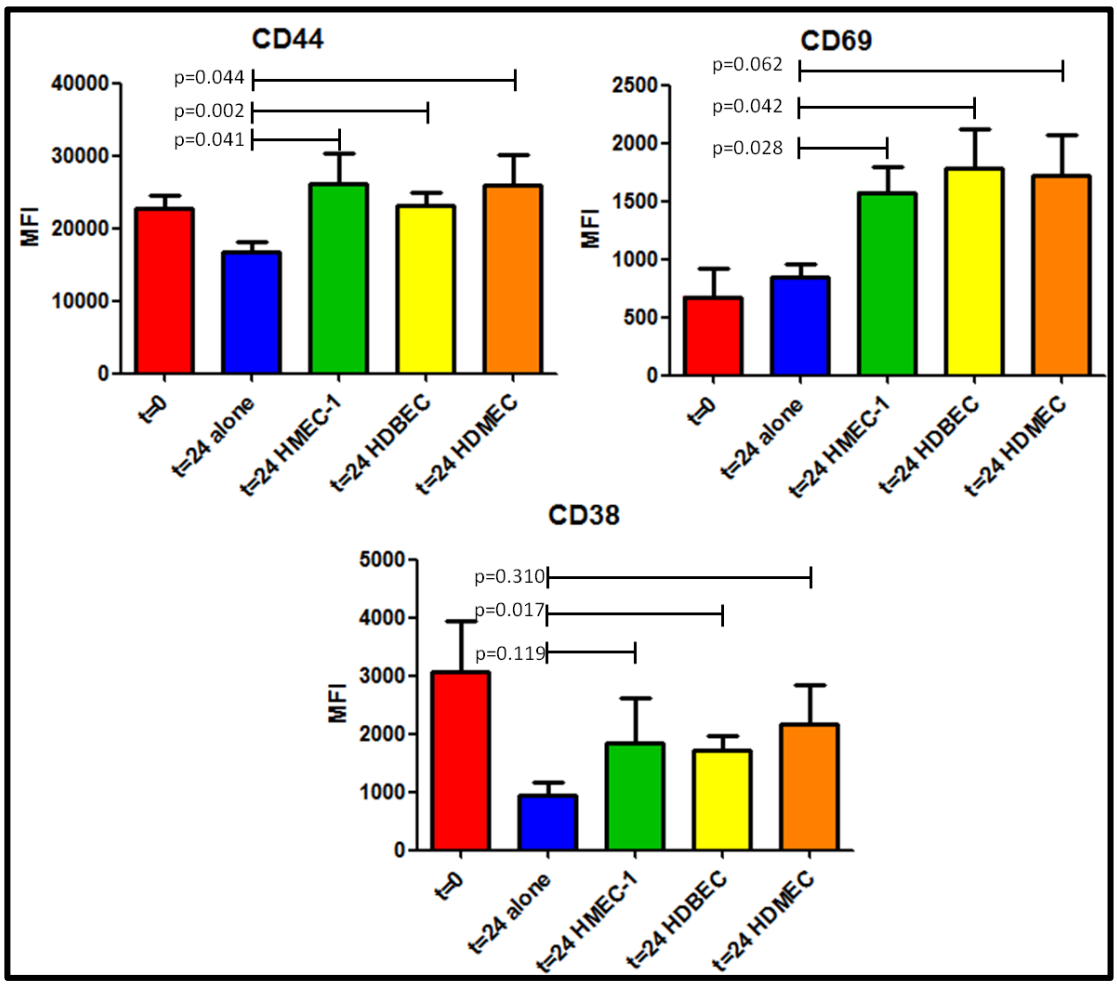

Figure 4.3 Phenotyping of CLL cells by flow cytometry. CLL cells from 6 different CLL patients were cultured alone and on HMEC-1, HDBEC and HDMEC cells for 24 hours. The expression of CD44, CD38 and CD69 were determined in each case by flow cytometry (paired t-test, mean±SEM, experiments with cells from n=6 different CLL patients).

4.2.3. Analysis of RNA quality isolated from co-cultured CLL cells

In order to analyse a variety of CLL phenotypes (described in **Table 4.1**) using the same endothelial cells in the co-culture, it was necessary to use CLL samples which had been frozen. It is extremely important to verify that the quality of the sample starting material is good enough to produce robust data since inclusion of samples with degraded RNA can affect gene expression levels (178).

The quality of the RNA isolated from each sample was analysed using the Agilent 2100 Bioanalyser, a microfluidics based platform for quality control as described in section 2.5.4. RNA quality is

measured using the RNA Integrity Number (RIN) algorithm (179), which is rapidly becoming standard practice. The RIN analysis allows the reproducible comparison of RNA quality between different samples at different times. The Agilent software assigns an integrity number to the total RNA sample based on the entire electrophoretic trace of the RNA sample rather than just rRNA ratios as was used historically. In this way, the RIN algorithm includes the presence or absence of degradation products. As the RIN algorithm provides a numerical assessment of the integrity of RNA, it facilitates the direct comparison of RNA samples and enables the reproducibility of experiments. Ideally, all RNA samples used for a microarray experiment would be of comparable RIN number and integrity. RNA and microarray processing were carried out under the supervision of Mrs Megan Musson in the Central Biotechnology Services at the University of Cardiff, as part of collaboration with Professor Chris Pepper. **Figure 4.4** shows an example of the readout from an Agilent 2100 Bioanalyser RNA 6000 Nano chip analysis of 12 RNA samples used in this study.

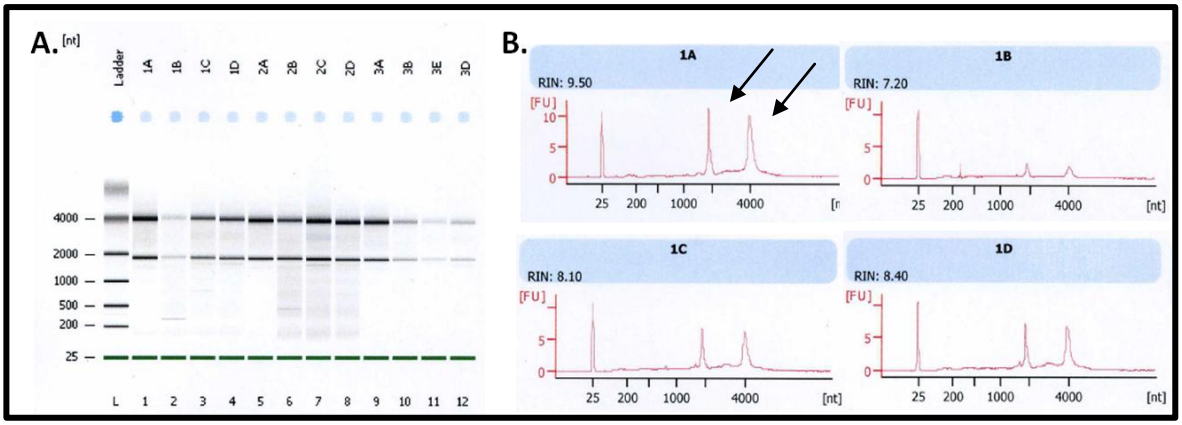

Figure 4.4 An Agilent 2100 Bioanalyser RNA 6000 Nano chip trace. **A**. Shows a computer generated electrophoretic trace for the samples analysed. Numbers identify different patients and letters identify conditions A = 0 hours B = 12 hours liquid culture C = 12 hours HMEC-1 co-culture D = 12 hours HDBEC co-culture **B.** Arrows indicate the 18S and 28S RNA peaks identified by the algorithm.

Although the use of the RNA 6000 Nano chip here was not intended to measure RNA quantity, **Figure 4.4A** clearly shows a difference in the amount of total RNA isolated from each sample, for example in lane 1A corresponding to 0 hours and lane 1B corresponding to 12 hours liquid culture. However, **Figure 4.4B** demonstrates clear peaks of 18S and 28S rRNA, little RNA degradation between the peaks and a low baseline. Therefore, as the quality of the RNA was comparable across samples as identified by the RIN score, these samples were considered suitable for microarray analysis.

4.2.4. Analyses of gene expression array data

Affymetrix Whole Transcript Human Gene 1.0 ST arrays were used in this study covering 21,014 Refseq (Entrez) genes and 36,079 RefSeq transcripts (coverage derived from RefSeq download as of February 2012 (180)).The data from gene expression array chips, such as the one shown in **Figure 4.5** were analysed during a rotation in Bioinformatics with Dr Eric Blanc in the MRC Centre for Developmental Neurobiology, King's College London. Data were analysed using Bioconductor (128), version 2.10, and biomaRt version 2.12.0, which use the R statistical programming language and provides tools for the analysis and comprehension of high-throughput genomic data. BiomaRt (181) provides an interface to a collection of databases, allowing retrieval of large amounts of data in a uniform way without knowledge of the underlying database (182, 183). Examples of BiomaRt databases are Ensembl, Uniprot and HGNC.

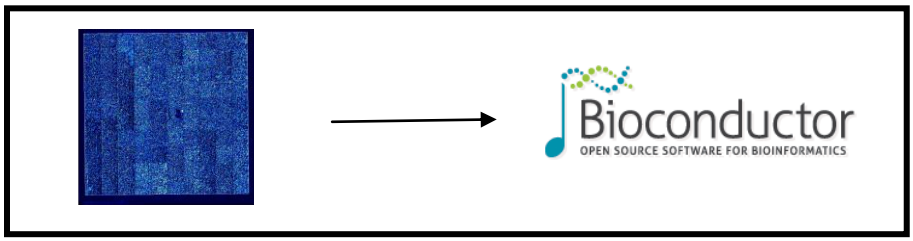

Figure 4.5 A scanned image of a GeneChip array ready for data processing. An Affymetrix Human Gene 1.0 ST Gene Chip used in this study, scanned after hybridisation overnight. Brightness is proportional to the amount of RNA hybridised to each probe.

4.2.5. Gene expression array data cluster by co-culture condition

Samples from a variety of CLL patients with different phenotypes were used for this study, as described in **Table 4.1**. Therefore it was important to establish whether any differences in mRNA expression observed were a result of the different clinical features of patients' disease or a result of the co-culture condition. Gibbons and Roth (184) report that no method outperforms Euclidean distance for ratio-based measurements as a measure of dissimilarity between the expression patterns of two genes. Data clustering in this study was based on the intensity record of the raw data and the average Euclidian distance between samples was calculated and visualised using the function rma() in R. As shown in **Figure 4.6**, raw data clustered by condition represented by colour and letters. The number corresponds to the CLL patient sample and clustering of patient number is not observed. The analysis represented in **Figure 4.6** suggests that the differences observed in gene expression are due to condition rather than to differences intrinsic to cells isolated from particular CLL patients. As the data did not cluster by patient or disease phenotype, these gene

expression data from six patients were treated as biological replicates and used to investigate common mechanisms induced in the CLL cells by co-culture.

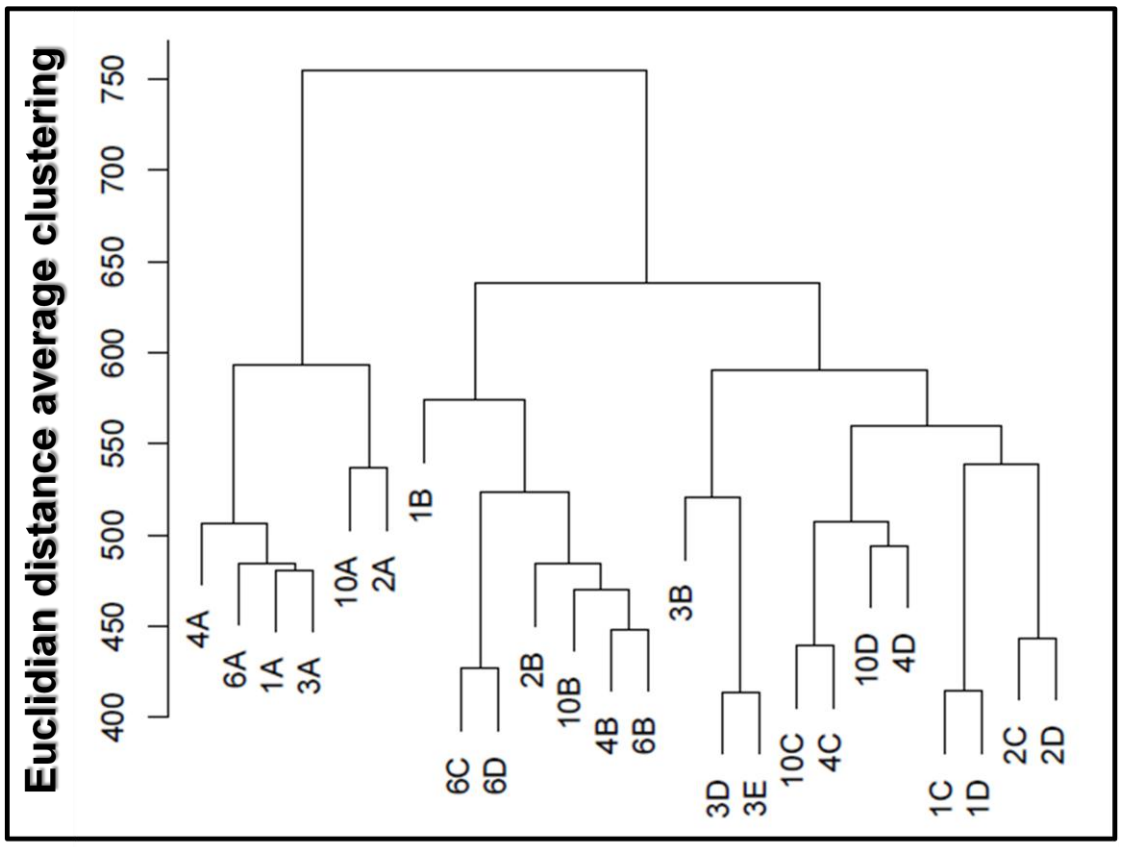

Figure 4.6 Dendrogram showing clustering of raw data. Average Euclidian distance between clusters is plotted. Condtions are grouped by colour and letter, A= t=0h, B= 12h liquid culture, C= 12h HMEC-1 co-culture, D=12h HDBEC co-culture. Individual CLL patients are identified by number.

Next, the data were normalised using the Robust Multichip Average (RMA) method (185). RMA analysis is comprised of three steps: background adjustment, quantile normalization and summarisation.**Figure 4.7** shows the effect of normalisation on the data. Different conditions are shown in different colours and the line style represents the CLL patient. In the raw data, spread is greatest in the centre of the plot, where density is greatest. After normalisation, there is less spread between the arrays, this means that any differences observed in mRNA expression are due to biological rather than technical reasons. Clustering of the normalised data is shown in **Figure 4.8**. This plot shows that the normalised data cluster by condition rather than by patient or the phenotype of the cells, as was the case for the raw data.

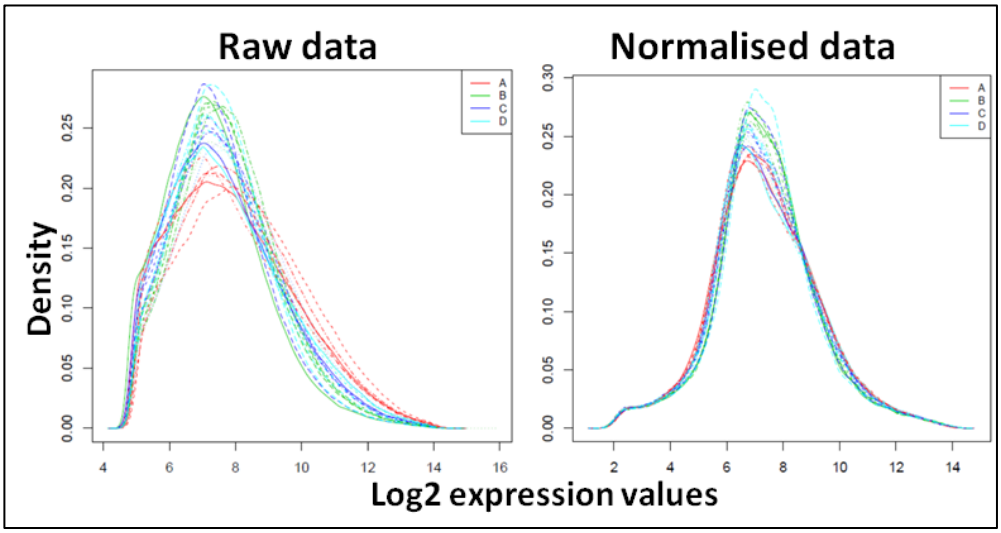

Figure 4.7 Expression values: raw *vs.* normalised data. Density plotted against Log2 expression values. Conditions are shown by colour and letter, A= t=0h, B= 12h liquid culture, C= 12h HMEC-1 co-culture, D=12h HDBEC co-culture. Individual CLL patients are identified by number. The line style represents the CLL patient.

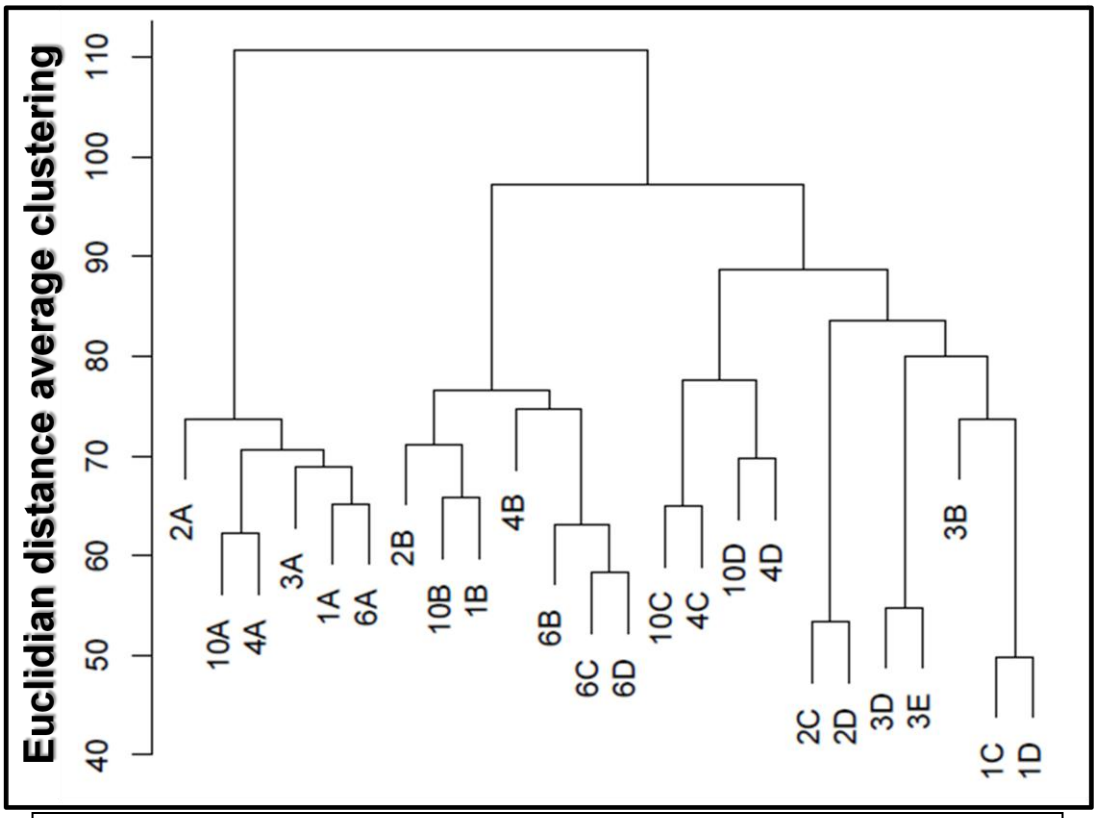

Figure 4.8 Dendrogram showing clustering of normalised data. Average Euclidian distance between clusters is plotted. Conditions are grouped by colour and letter. A= t=0h, B= 12h liquid culture, C= 12h HMEC-1 co-culture, D=12h HDBEC co-culture. Individual CLL patients are identified by number.

Conclusions of statistical testing:

- the gene expression array data are technically sound and
- changes in the signals observed are due to the biological experiment rather than for technical reasons.

4.2.6. Differential expression of mRNA as a result of endothelial co-culture

When analysing mRNA expression, the claim that a particular mRNA is differentially expressed in different samples can have several interpretations. In formal statistical terms, an mRNA is said to be differentially expressed if its expression level changes between two treatment conditions, regardless of how small the difference might be. However, in scientific discussion, an mRNA is likely to be considered differentially expressed only if its expression level changes by a biologically worthwhile amount that is likely to have a biological effect. Therefore, it is apparent that there is a disparity between the mathematical and biological concepts of differential expression. Many different cut-off methods have been applied to gene expression array data and it is useful to think of these methods as filtering tools to produce smaller, biologically meaningful datasets. For example, one method of differential analysis may be to calculate the fold change of expression. However small changes can still be biologically relevant, such as minor alterations in key regulators such as *TP53* mRNA expression (186). Another method is to calculate the significance of a change in expression between two conditions. A study from Patterson *et al*. (187) required genes to satisfy a modest level of statistical significance ($p < 0.01$ or $p < 0.05$) then ranked significant genes by fold-change with a cut-off of 1.5, 2 or 4. Peart *et al.* (188) and Raouf *et al.* (189) used cut offs of fold-change >1.5 and $p < 0.05$ after adjustment for multiple testing, whilst Huggins *et al.* (190) required a 1.3 fold-change and $p < 0.2$. Dalman *et a*l. (191) showed that changing fold change and p-value cut-offs significantly altered microarray interpretations for their data when GO annotation categories were analysed, implying that different signalling pathways and functions involved when different cut off criteria were used. It is becoming increasingly common to require that differentially expressed genes satisfy both p-value and fold-change criteria simultaneously. Normally, a combination ranking gives better agreement between platforms and typically identifies more biologically meaningful sets of genes than p-values alone (191).

For the data reported here, I chose to apply cut-off criteria including significance $q<10^{-3}$ and fold change >1.5. A p-value is a measure of how likely you are to observe a difference in expression if no real difference existed between the two conditions. The q-value is the name given to adjusted p-values, using an optimized false discovery rate (FDR). The FDR is the expected proportion of false

positives among all discoveries (rejected null hypotheses). For example, in this case if the null hypotheses of 1000 hypothesis tests were experimentally rejected and the maximum FDR level (q-value) for these tests was 0.001, fewer than 1 of these rejections would be expected to be a false positive. Changes in the expression of 103 CLL cell mRNA met these cut-off criteria for the HDBEC co-culture dataset and 134 from the HMEC-1 co-culture dataset, when compared with their expression in CLL cells in liquid culture (for full data set see Supplementary Table 4.2.6). **Table 4.2** shows the top five up-regulated mRNA for each co-culture system which met the significance criteria.

The Limma package in R was used to analyse differential mRNA expression of data from this study. Limma fits a linear model to the expression data for each gene (192). The approach requires two matrices to be specified. The first is the design matrix which identifies which RNA samples have been applied to each array. The second is the contrast matrix which specifies which comparisons you would like to make between the RNA samples, for example samples after 12 hours liquid culture or 12 hours endothelial co-culture. The philosophy of the approach fits a linear model to the data which fully models the systematic part of the data. The model is specified by the design matrix. Each row of the design matrix corresponds to an array in the experiment and each column corresponds to a coefficient that is used to describe the RNA sources in the experiment. A contrast matrix can then be created which allows a comparison of the different co-culture conditions using linear algebra.

Up-regulated by HMEC-1 co-culture Symbol	Up-regulated by HMEC-1 co-culture	Up-regulared by HMEC-1 co-culture Fold Change	Up-regulate by HDBEC co-culture Symbol	Up-regulate by HDBEC co-culture	Up-regulated by HDBEC co-culture Fold Change
IL13RA	interleukin 13 receptor, alpha 1	5.65	THBS1	thrombospondin 1	5.65
TM4SF1	transmembrane 4 L six family member 1	5.38	MMP1	matrix metallopeptidase 1	5.38
TFPI2	tissue factor pathway inhibitor 2	5.01	CCL2	chemokine (C-C motif) ligand 2	5.01
CAV1	caveolin 1	4.83	SERPINE1	serpin peptidase inhibitor, clade E	4.83
ANAX1	annexin A1	4.78	ANKRD1	ankyrin repeat domain 1	4.78

Table 4.2 Top up-regulated mRNA by fold change. The top five up-regulated mRNA for each co-culture system are shown. The full data sets are available in Supplementary Table 4.2.6.

4.2.7. Analyses of GO term annotations enriched in co-culture datasets

In order to determine whether the mRNA identified by the microarray analyses contained an over representation of corresponding genes involved in particular biological processes, cellular components or biological functions, hypergeometric testing was performed on GO annotated genes. Hypergeometric testing identifies over representation of genes with particular annotations relative to the reference database.

Genes were assigned GO annotation terms which are curated by the Gene Ontology Consortium (136) using the OBOv1.2 file from http://www.geneontology.org/GO.downloads.ontology.shtml. The GO ontologies resemble a tree where child terms are more specialized and parent terms are less specialized, and each term may have more than one parent term (see section 1.3.3 (Introduction)). The relationships used in GO are also directed, for example a mitochondrion is an organelle, but an organelle is not a mitochondrion. GO annotations in OBOv1.2 are linked to UNIPROT identifiers, however the microarray data were annotated with Ensembl identifiers and so the GO ontologies had to be mapped through ID conversions using BiomaRt, GO annotation → UNIPROT ID → Ensembl ID. A GO term hierarchy was also created for the terms for each GO term to give a list of ancestors. This step was important to account for the presence of child terms which result from a parent term rather than over representation.

When converting between IDs, it is important to repeat checks on the data to ensure genes are not lost or duplicated through incorrect mapping. This was achieved by monitoring the length of the gene lists created at each stage of the ID conversion. Problems can arise during mapping where one gene may give rise to several different proteins due to transcript variants with different IDs, therefore mapping of GO annotations was done at the transcript level. Another example of difficult mapping is where many proteins arise from a transcript. The aim was to annotate each transcript with all of the GO terms associated with the related proteins. Any IDs which did not map were removed to speed up computation. A gene 'universe' or gene list of 19177 with mappings was created. Hypergeomtric testing was then applied to the data to determine the probability of enrichment of any GO term in this 'gene universe'. This was calculated using nList, the number of genes annotated with the term, nCAT, the number of genes annotated with that term and the length of the gene universe (the 19177 genes with complete mapping).

Table 4.3 shows the top 10 GO annotations with the lowest q value from the HDBEC dataset, including processes such as angiogenesis, cell adhesion and cellular components including extracellular space, external side of plasma membrane and cell surface. For the complete dataset see Supplementary Table 4.2.7.

Definition	GO Term	nList	nCat	p.value	q.value
angiogenesis	GO:0001525	25	186	6.22E-23	8.00E-19
cell adhesion	GO:0007155	29	455	2.28E-17	1.47E-13
extracellular space	GO:0005615	36	833	5.04E-16	2.16E-12
extracellular matrix	GO:0031012	18	154	1.13E-15	3.62E-12
extracellular region	GO:0005576	46	1445	2.25E-15	5.79E-12
external side of plasma membrane	GO:0009897	17	166	6.70E-14	1.44E-10
cell surface	GO:0009986	19	318	3.34E-11	6.14E-08
integrin binding	GO:0005178	11	77	5.24E-11	8.43E-08
plasma membrane	GO:0005886	64	3431	1.36E-10	1.94E-07
blood cogulation	GO:0007596	20	463	2.98E-09	3.83E-06

Table 4.3 Enriched GO terms in the HDBEC *vs.* liquid culture dataset. A hypergeometric test was applied and data were ranked by q score. Analyses of all the data are shown in Supplementary Table 4.2.7.

Table 4.4 shows the top ten GO annotations with the lowest q value from the HMEC-1 dataset. These include processes such as cell division and mitosis and cellular components such as cytosol and cytoplasm. This is interesting given that no evidence for CLL cell proliferation was observed in the HMEC-1 co-culture system, as described in section 3.2, Chapter 3. For the complete dataset see Supplementary Table 4.2.7.

Definition	GO Term	nList	nCat	p.value	q.value
mitotic cell cycle	GO:0000278	52	340	2.05E-30	2.64E-26
cell division	GO:0051301	48	302	6.85E-29	4.41E-25
mitosis	GO:0007067	35	189	1.38E-23	5.93E-20
cytosol	GO:0005829	120	2305	2.72E-23	8.74E-20
microtubule-based movement	GO:0007018	25	101	1.98E-20	5.10E-17
cytoplasm	GO:0005737	156	4129	8.19E-17	1.76E-13
protein binding	GO:0005515	174	4908	2.20E-16	4.04E-13
cytoskeleton-dependant intracellular transport	GO:0030705	10	13	3.00E-15	4.83E-12
G2/M transition of mitotic cell cycle	GO:0000086	22	123	6.22E-15	8.89E-12
M phase of mitotic cell cycle	GO:0000087	19	93	3.55E-14	4.56E-11

Table 4.4 Enriched GO terms in the HMEC-1 *vs.* liquid culture dataset. A hypergeoetric test was applied and data were ranked by q score. Analyses of all the data are shown in Supplementary Table 4.2.7.

4.2.8. 31 genes are up-regulated in both co-culture systems

As described in section 4.2.1, both endothelial co-culture systems increased the viability of CLL cells. In order to determine whether this was though a common mechanism, I next identified mRNA significantly up regulated in both co-culture systems as compared with liquid culture. When applying a 1.5 fold change condition, 103 mRNA are significantly up-regulated in CLL cells by HDBEC co-culture compared with liquid culture and 134 mRNA up-regulated by HMEC-1 co-culture compared with liquid culture. Only 31 mRNA were identified to be up-regulated in CLL cells in both co-culture

systems with a significant q value and an increase in mRNA expression >1.5 fold when compared with liquid culture. The overlap between the co-culture datasets is shown in **Figure 4.9** and a complete list of the data is shown in Supplementary Table 4.2.8.

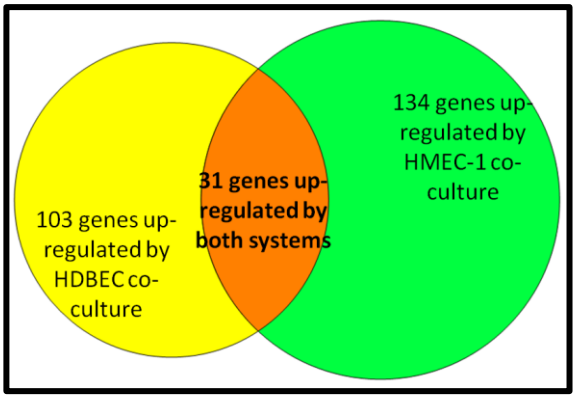

Figure 4.9 mRNA up-regulated by both endothelial co-culture systems. The overlap of mRNA up-regulated by both endothelial cell co-culture systems was identified using transcript lists which met significant q-value and FC>1.5 compared with liquid culture for each co-culture data set. A complete list of these mRNA is in Supplementary Table 4.2.8.

Some of the mRNA identified by these analyses are already known to play a role in the biology of CLL, such as IL-6 (13) whilst others are less well characterised in CLL, such the metallothioneins (MT). Increased MT levels have been observed in several tumour types including breast, kidney, nasopharynx, lungs, prostate, testes, urinary bladder, cervix, endometrium, salivary glands, pancreas, acute lymphoblastic leukaemia and melanoma (193). MT have also been shown to play a role in drug resistance (194). The 31 mRNA up-regulated by both systems with >1.5 fold change are shown in **Table 4.5** with ontology annotations retrieved from HumanNet (http://www.functionalnet.org/humannet/). mRNA were identified that are involved in processes including cell:cell adhesion, signal transduction and anti-apoptosis mechanisms. In the following sections, I performed bioinformatic analyses to determine whether these mRNA up-regulated by both endothelial co-cultures are co-regulated.

EntrezGene	SYMBOL	Average FC	GO Biological Process	GO Cellular Component	GO Molecular Function
301	ANXA1	4.01	lipid metabolic process; anti-apoptosis; cell motility; inflammatory response; cell cycle; cell surface receptor linked signal transduction; peptide cross-linking; regulation of cell proliferation;	cornified envelope; cytoplasm; sarcolemma;	phospholipase inhibitor activity; receptor binding; structural molecule activity; calcium ion binding; calcium-dependent phospholipid binding; phospholipase A2 inhibitor activity; protein binding, bridging;
857	CAV1	3.94	inactivation of MAPK activity; negative regulation of endothelial cell proliferation; negative regulation of signal transduction; regulation of vasoconstriction; regulation of vasodilation;	Golgi membrane; integral to membrane of membrane fraction; endoplasmic reticulum; plasma membrane; integral to plasma membrane; caveolar membrane; lipid raft; perinuclear region;	structural molecule activity; protein binding; cholesterol binding;
7057	THBS1	3.88	cell motility; inflammatory response; cell adhesion; multicellular organismal development; blood coagulation; negative regulation of angiogenesis;	extracellular region; extracellular space;	endopeptidase inhibitor activity; signal transducer activity; structural molecule activity; calcium ion binding; protein binding; heparin binding;
4071	TM4SF1	3.83	na	integral to plasma membrane; membrane; integral to membrane;	na
3569	IL6	3.46	neutrophil apoptosis; acute-phase response; humoral immune response; cell surface receptor linked signal transduction; cell-cell signaling; positive regulation of cell proliferation; negative regulation of cell proliferation; negative regulation of apoptosis; negative regulation of chemokine biosynthetic process;	extracellular region; extracellular space;	cytokine activity; interleukin-6 receptor binding; protein binding;
4495	MT1G	3.21	na	na	copper ion binding; zinc ion binding; cadmium ion binding; metal ion binding;
5054	SERPINE1	3.03	blood coagulation; fibrinolysis; regulation of angiogenesis;	extracellular region;	protease binding; serine-type endopeptidase inhibitor activity; protein binding; plasminogen activator activity;
2791	GNG11	2.99	signal transduction; G-protein coupled receptor protein signaling pathway;	heterotrimeric G-protein complex; membrane;	GTPase activity; signal transducer activity;
55970	GNG12	2.91	signal transduction; G-protein coupled receptor protein signaling pathway;	heterotrimeric G-protein complex; membrane;	signal transducer activity;
131566	DCBLD2	2.76	cell adhesion; negative regulation of cell growth; intracellular receptor-mediated signaling pathway; wound healing;	integral to plasma membrane; cell surface; membrane;	protein binding;

Table 4.5 Commonly up-regulated mRNA. Ontology from HumanNet **(**http://www.functionalnet.org/humannet/). The mRNA significantly up-regulated in both co-culture systems with a fold change >1.5 compared with liquid culture. The top 10 mRNA with the greatest average FC are shown. A complete list of these mRNA is in Supplementary Table 4.2.8.

4.2.9. Network analyses of genes up-regulated by both co-culture systems

I chose to further analyse all 53 mRNA which showed a significant increase in expression, regardless of the magnitude of fold change because the dataset of 31 mRNA with a greater than 1.5 fold change was too small for network analyses. I performed network analyses using HumanNet (132), an algorithm produced by our collaborators at the University of Texas at Austin. The algorithm consists of a probabilistic functional gene network which was constructed by integrating 21 types of 'omics' data from multiple organisms, with each data type weighted according to how well it links genes that are known to function together in *H. sapiens*. An interaction network was calculated

which predicts that proteins encoded by 39 of the 53 mRNA identified by gene expression analyses interact with one another. The HumanNet analyses were mapped with Cytoscape software to visualise the connections between the proteins and the network is shown in **Figure 4.10**. The Cytoscape network represents proteins as nodes and the edges indicate a predicted interaction between the proteins. Data were presented using the y-files organic layout tool. This bioinformatic analysis predicts that proteins encoded by 39 of the 53 genes interact with each other, suggesting that these proteins may work together functionally. It is also likely that genes encoding proteins with similar functions will be co-regulated at the transcriptional level.

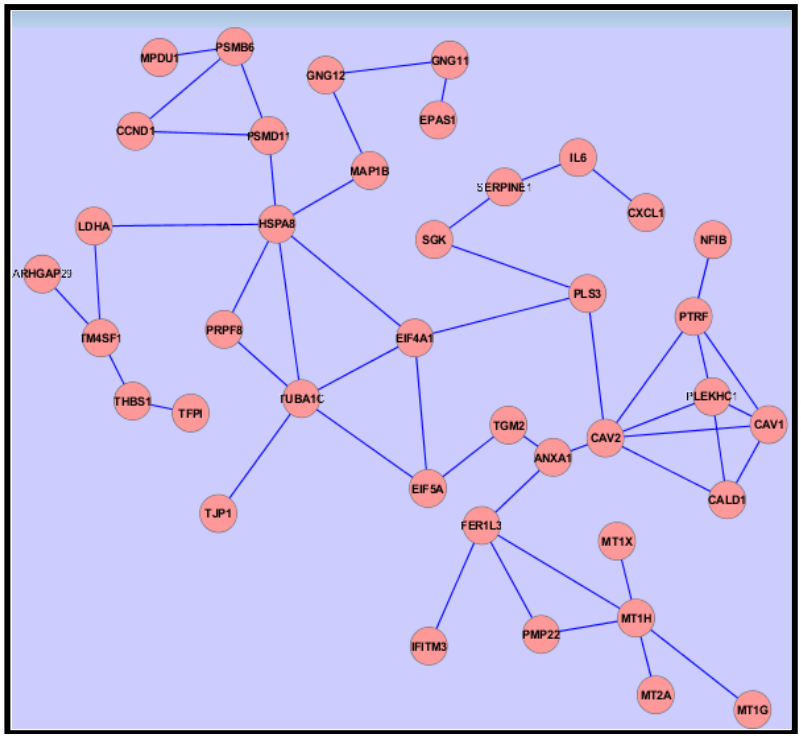

Figure 4.10 HumanNet analyses predict that proteins encoded by 39/53 genes up-regulated in both systems interact with one another (http://www.functionalnet.org/humannet/).

4.2.10. GSCA identifies a potential novel transcription factor module in CLL

In order to investigate any potential control mechanisms which may be responsible for the co-ordinated expression and therefore function of the 53 mRNA identified by analyses in section 4.2.8, I performed further bioinformatic analyses to look for predicted transcription factor (TF) modules. Chromatin immunoprecipitation (ChIP) followed by high-throughput sequencing (ChIP-Seq) allows interrogation of the genome-wide effects of TFs. The Gene Set Control Analysis (GSCA) algorithm integrates 142 publicly available ChIP-Seq datasets for both normal and leukaemic murine blood cell

types (195). It can be used to predict likely upstream regulators for lists of genes based on statistical significance of TF binding event enrichment within the gene loci of a chosen input gene set. GSCA was performed on the 53 genes identified in section 4.2.8 and a potential TF module was identified, as described in **Figure 4.11**, comprising STAT1, PU.1, p65, CEBPα and CEBPβ. The combined module of five TFs has not been reported in CLL but has been described in macrophages.

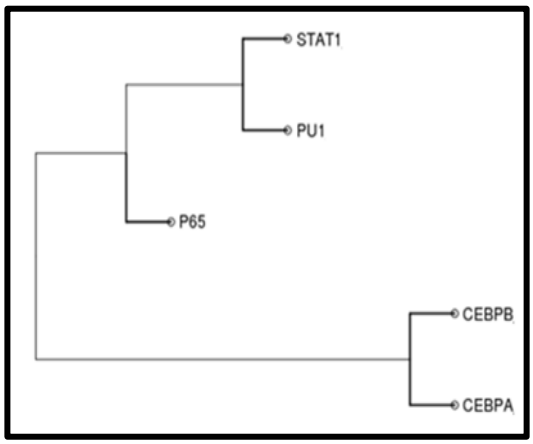

Figure 4.11 GSCA analyses (http://bioinformatics.cscr.cam.ac.uk/GSCA/GSCA) predict a TF module previously identified in macrophages that co-ordinates expression of 53 CLL genes up-regulated by both co-culture systems.

Examination of the original publically available ChIP-Seq datasets used to create the GCSA algorithm for these TF data (196-198) showed that 40 out of the 53 genes have putative binding sites for at least 3 of the 5 TF identified in the module. **Table 4.6** below shows the 40 genes identified with at least three TF putative binding sites and their GO annotations retrieved using HumaNet.

Entrez ID	HGNC Symbol	GO Process	GO Compartment	GO Function
3312	HSPA8	protein folding; response to unfolded protein;	intracellular; nucleus; cell surface;	nucleotide binding; protein binding; ATP binding; ATPase activity, coupled;
84790	TUBA1C	microtubule-based movement; protein polymerization;	mitochondrion; microtubule; protein complex;	nucleotide binding; GTPase activity; structural molecule activity; GTP binding;
1973	EIF4A1	translation;	eukaryotic translation initiation factor 4F complex;	nucleotide binding; RNA cap binding; nucleic acid binding; mRNA binding; translation initiation factor activity; protein binding; ATP binding; ATP-dependent helicase activity; hydrolase activity;
5717	PSMD11	na	proteasome complex (sensu Eukaryota); cytosol;	na
5694	PSMB6	ubiquitin-dependent protein catabolic process;	cytosol; proteasome core complex (sensu Eukaryota);	threonine endopeptidase activity;
595	CCND1	G1/S transition of mitotic cell cycle; re-entry into mitotic cell cycle; protein amino acid phosphorylation; cell cycle; unfolded protein response; fat cell differentiation; cell division;	cyclin-dependent protein kinase holoenzyme complex; intracellular; nucleus; cytosol;	protein kinase activity; protein binding; cyclin-dependent protein kinase regulator activity;
2791	GNG11	signal transduction; G-protein coupled receptor protein signaling pathway;	heterotrimeric G-protein complex; membrane;	GTPase activity; signal transducer activity;
1984	EIF5A	regulation of translational initiation; apoptosis; viral genome replication;	nucleus; cytoplasm;	translation initiation factor activity;
7052	TGM2	G-protein coupled receptor protein signaling pathway; peptide cross-linking; positive regulation of cell adhesion;	extracellular matrix (sensu Metazoa); cytosol; membrane;	protein-glutamine gamma-glutamyltransferase activity; calcium ion binding; GTP binding; acyltransferase activity; transferase activity;
55970	GNG12	signal transduction; G-protein coupled receptor protein signaling pathway;	heterotrimeric G-protein complex; membrane;	signal transducer activity;
301	ANXA1	lipid metabolic process; anti-apoptosis; cell motility; inflammatory response; cell cycle; cell surface receptor linked signal transduction; peptide cross-linking; keratinocyte differentiation; regulation of cell proliferation; arachidonic acid secretion;	cornified envelope; cytoplasm; sarcolemma;	phospholipase inhibitor activity; receptor binding; structural molecule activity; calcium ion binding; calcium-dependent phospholipid binding; phospholipase A2 inhibitor activity; protein binding, bridging;
7057	THBS1	cell motility; inflammatory response; cell adhesion; multicellular organismal development; nervous system development; blood coagulation; negative regulation of angiogenesis;	extracellular region; extracellular space;	endopeptidase inhibitor activity; signal transducer activity; structural molecule activity; calcium ion binding; protein binding; heparin binding;
7035	TFPI	blood coagulation;	extracellular region;	serine-type endopeptidase inhibitor activity; protease inhibitor activity;
858	CAV2	negative regulation of endothelial cell proliferation; protein oligomerization;	plasma membrane; integral to plasma membrane; caveolar membrane; transport vesicle; lipid raft; perinuclear region;	protein binding; protein homodimerization activity;
7082	TJP1	intercellular junction assembly;	membrane fraction; nucleus; plasma membrane; septate junction; tight junction; intercellular canaliculus;	protein binding;
3569	IL6	neutrophil apoptosis; acute-phase response; humoral immune response; cell surface receptor linked signal transduction; cell-cell signaling; positive regulation of cell proliferation; negative regulation of cell proliferation; negative regulation of apoptosis; negative regulation of chemokine biosynthetic process;	extracellular region; extracellular space;	cytokine activity; interleukin-6 receptor binding; protein binding;
10594	PRPF8	nuclear mRNA splicing, via spliceosome; visual perception; RNA splicing; response to stimulus;	nucleus; spliceosome; snRNP U5;	protein binding; RNA splicing factor activity, transesterification mechanism;
4071	TM4SF1	na	integral to plasma membrane; membrane; integral to membrane;	na
3939	LDHA	tricarboxylic acid cycle intermediate metabolic process; anaerobic glycolysis;	cytoplasm; cytosol;	L-lactate dehydrogenase activity; protein binding; oxidoreductase activity;
284119	PTRF	transcription; transcription termination; regulation of transcription, DNA-dependent; transcription initiation from RNA polymerase I promoter;	nucleus; membrane;	RNA polymerase I transcription termination factor activity; protein binding; rRNA primary transcript binding;
5054	SERPINE1	blood coagulation; fibrinolysis; regulation of angiogenesis;	extracellular region;	protease binding; serine-type endopeptidase inhibitor activity; protein binding; plasminogen activator activity;
2034	EPAS1	angiogenesis; response to hypoxia; regulation of transcription, DNA-dependent; transcription from RNA polymerase II promoter; signal transduction; multicellular organismal development; cell differentiation;	nucleus;	RNA polymerase II transcription factor activity, enhancer binding; transcription coactivator activity; signal transducer activity; protein binding; histone acetyltransferase binding;
2919	CXCL1	chemotaxis; inflammatory response; immune response; G-protein coupled receptor protein signaling pathway; intracellular signaling cascade; nervous system development; cell proliferation; negative regulation of cell proliferation; actin cytoskeleton organization and biogenesis;	extracellular region; extracellular space;	chemokine activity; enzyme activator activity; growth factor activity;
800	CALD1	cell motility; muscle contraction;	membrane fraction; cytoskeleton; actin cap;	actin binding; calmodulin binding; tropomyosin binding; myosin binding;
9526	MPDU1	na	membrane; integral to membrane;	na
966	CD59	defense response; immune response; cell surface receptor linked signal transduction; blood coagulation;	membrane fraction; plasma membrane;	protein binding; GPI anchor binding;
5954	RCN1	na	endoplasmic reticulum; endoplasmic reticulum lumen;	calcium ion binding;
10410	IFITM3	immune response; response to biotic stimulus;	plasma membrane; integral to membrane;	na
5376	PMP22	synaptic transmission; peripheral nervous system development; sensory perception of sound; mechanosensory behavior; negative regulation of cell proliferation;	membrane fraction; integral to plasma membrane; membrane;	na
26064	RAI14	na	mitochondrion; cytoskeleton;	na
8754	ADAM9	proteolysis; protein kinase cascade;	integral to plasma membrane; membrane;	metalloendopeptidase activity; integrin binding; protein binding; zinc ion binding; SH3 domain binding; protein kinase binding; metal ion binding;
87	ACTN1	regulation of apoptosis; focal adhesion formation; actin filament bundle formation; negative regulation of cell motility;	cytoskeleton; focal adhesion; Z disc; pseudopodium;	actin binding; integrin binding; calcium ion binding; protein binding; vinculin binding; actin filament binding;
9518	GDF15	signal transduction; transforming growth factor beta receptor signaling pathway; cell-cell signaling;	extracellular region; extracellular space;	cytokine activity; growth factor activity;
598	BCL2L1	anti-apoptosis; negative regulation of survival gene product activity; apoptotic mitochondrial changes; response to radiation; regulation of apoptosis; positive regulation of anti-apoptosis;	nucleus; mitochondrion; mitochondrial outer membrane; membrane; integral to membrane;	identical protein binding;
55379	LRRC59	na	membrane; integral to membrane;	protein binding;
11343	MGLL	lipid metabolic process; aromatic compound metabolic process; inflammatory response;	na	lysophospholipase activity; serine esterase activity; hydrolase activity; acylglycerol lipase activity;
140576	S100A16	na	na	calcium ion binding;
3383	ICAM1	cell-cell adhesion; regulation of cell adhesion;	plasma membrane; integral to plasma membrane;	transmembrane receptor activity; protein binding;
131566	DCBLD2	cell adhesion; negative regulation of cell growth; intracellular receptor-mediated signaling pathway; wound healing;	integral to plasma membrane; cell surface; membrane;	protein binding;

Table 4.6 40 Genes with potential binding sites for at least 3 of the 5 TF identified as a novel TF module in CLL. The 25 genes predicted to interact with one another by HumaNet are highlighted in yellow.

Next, I wanted to determine whether these 40 genes were involved in the same processes, as the coordinate expression of genes involved in the same functions is important biologically, for example in the mechanism of cell cycle control (199). DAVID Bioinformatics Resources (200, 201) were used to provide batch annotation and gene-GO term enrichment analysis of these 40 genes compared to genes from the reference database. **Table 4.7** shows the top 10 most significant GO terms identified in the annotations of the 40 genes analysed. 11 of the 40 genes are annotated with the GO term 'regulation of cell proliferation'. The enrichment of GO terms in these 40 genes suggests that their coordinated expression may have a functional role.

GO Term	Count	%	PValue	List Total	Pop Hits	Pop Total	Fold Enrichment	FDR
GO:0042127~regulation of cell proliferation	11	28.21	1.53E-05	35	787	13528	5.40	2.38E-02
GO:0032270~positive regulation of cellular protein metabolic process	7	17.95	2.20E-05	35	233	13528	11.61	3.41E-02
GO:0051247~positive regulation of protein metabolic process	7	17.95	2.79E-05	35	243	13528	11.13	4.33E-02
GO:0009725~response to hormone stimulus	8	20.51	2.92E-05	35	367	13528	8.43	4.53E-02
GO:0009611~response to wounding	9	23.08	3.88E-05	35	530	13528	6.56	6.03E-02
GO:0009719~response to endogenous stimulus	8	20.51	5.47E-05	35	405	13528	7.63	8.49E-02
acetylation	16	41.03	7.39E-05	39	2635	19235	2.99	8.93E-02
GO:0044093~positive regulation of molecular function	9	23.08	7.90E-05	35	586	13528	5.94	1.23E-01
GO:0044459~plasma membrane part	17	43.59	1.29E-04	37	2203	12782	2.67	1.54E-01
GO:0010035~response to inorganic substance	6	15.38	1.48E-04	35	205	13528	11.31	2.30E-01

Table 4.7 DAVID Functional Annotation of 40 genes with 3 putative TF binding sites. Top 10 most significant are shown. The full table is in Supplementary Table 4.2.10.

I then analysed these data using the DAVID Functional Annotation Clustering tool. This algorithm measures relationships between annotated terms based on the degree of their co-association genes to group the similar, redundant and heterogeneous annotation from different resources into annotation. This helps to reduce the problem of associating similar redundant terms, improving biological interpretation of datasets. **Table 4.8** shows the top five most significant annotation clusters identified by this analysis all of the data is shown in Supplementary Table 4.2.10. The presence of annotation clusters demonstrates that these genes are associated with particular functions or processes such as response to oxidative stress in annotation cluster 1. This is suggestive of a common control mechanism behind the coordinated expression of these genes.

Annotation Cluster 1	Enrichment Score: 2.58			
Term	%	PValue	Fold Enrichment	FDR
GO:0010035~response to inorganic substance	15.38	1.48E-04	11.31	2.30E-01
GO:0006979~response to oxidative stress	10.26	7.95E-03	9.43	1.17E+01
GO:0000302~response to reactive oxygen species	7.69	1.52E-02	15.46	2.11E+01
Annotation Cluster 2	Enrichment Score: 2.32			
Term	%	PValue	Fold Enrichment	FDR
GO:0009611~response to wounding	23.08	3.88E-05	6.56	6.03E-02
GO:0042060~wound healing	12.82	1.28E-03	10.12	1.97E+00
GO:0050817~coagulation	10.26	2.10E-03	15.16	3.21E+00
GO:0007596~blood coagulation	10.26	2.10E-03	15.16	3.21E+00
GO:0007599~hemostasis	10.26	2.47E-03	14.32	3.77E+00
GO:0050878~regulation of body fluid levels	10.26	5.24E-03	10.96	7.83E+00
hsa04610:Complement and coagulation cascades	7.69	3.82E-02	9.21	3.22E+01
GO:0005576~extracellular region	25.64	1.01E-01	1.72	7.19E+01
Secreted	17.95	1.12E-01	2.04	7.61E+01
Annotation Cluster 3	Enrichment Score: 2.28			
Term	%	PValue	Fold Enrichment	FDR
GO:0044459~plasma membrane part	43.59	1.29E-04	2.67	1.54E-01
GO:0005886~plasma membrane	56.41	3.00E-04	2.01	3.57E-01
cell membrane	20.51	1.36E-01	1.80	8.28E+01
membrane	43.59	1.39E-01	1.34	8.35E+01
Annotation Cluster 4	Enrichment Score: 1.77			
Term	%	PValue	Fold Enrichment	FDR
GO:0032101~regulation of response to external stimulus	10.26	7.30E-03	9.72	1.08E+01
GO:0045765~regulation of angiogenesis	7.69	1.09E-02	18.41	1.56E+01
GO:0051241~negative regulation of multicellular organismal process	7.69	6.37E-02	7.07	6.40E+01
Annotation Cluster 5	Enrichment Score: 1.75			
Term	%	Genes	Fold Enrichment	FDR
GO:0006928~cell motion	15.38	6.27E-03	4.88	9.31E+00
GO:0051099~positive regulation of binding	7.69	1.63E-02	14.87	2.26E+01
GO:0051098~regulation of binding	7.69	5.63E-02	7.58	5.94E+01

Table 4.8 DAVID Functional Annotation Cluster analysis of 40 genes with 3 putative TF binding sites. The full table is in Supplementary Table 4.2.10.

I then used HumaNet to perform a network analysis of these 40 genes. This analysis predicted that 25 of the 40 genes interact with one another, these connected genes are highlighted in yellow in **Table 4.6. Figure 4.12** below shows that the predicted interactions occur in three modules. These can be explained functionally; for example the interlukin-6 (IL-6), Chemokine C-X-C motif ligand 1 (CXCL1) and serpin peptidase inhibitor clade E (SERPINE1) all have roles described in inflammation (202-204). SERPINE1 is also known as Plasminogen Activator Inhibitor-1 (PAI-1) and is expressed by endothelial cells such as those in blood vessels (205) and inhibits the activity of matrix metalloproteinases, (206) which play a crucial role in invasion. These bioinformatic analyses are suggestive of a common transcriptional control mechanism responsible for the coordinated expression of genes involved in specific processes and functions.

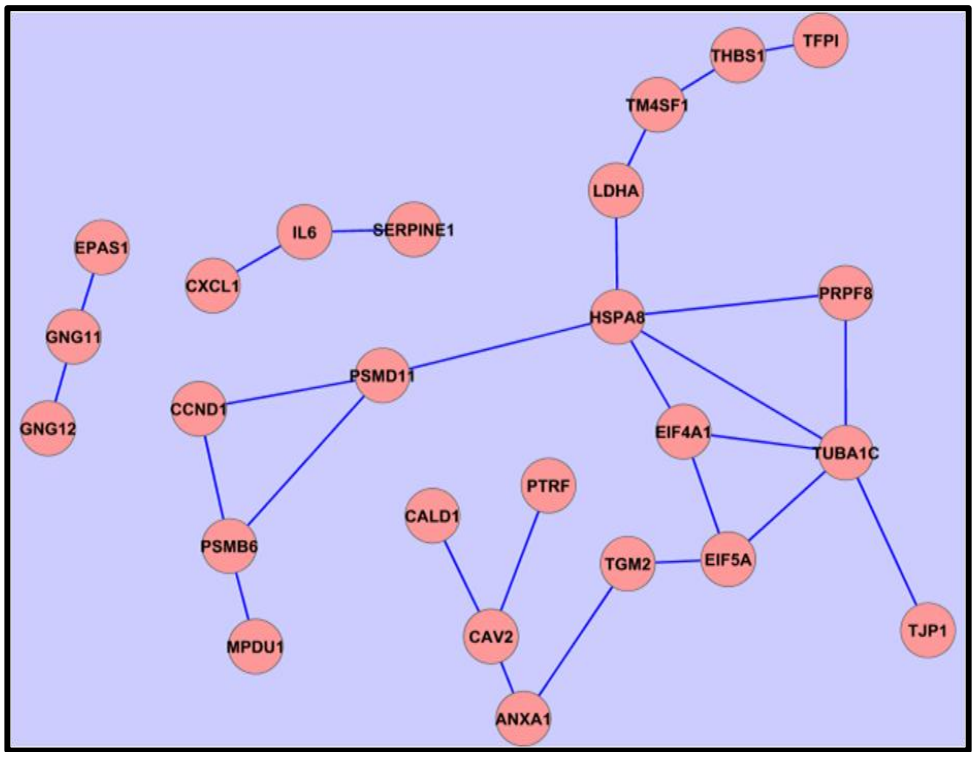

Figure 4.12 HumaNet network analysis predicts that proteins encoded by 25 of the 40 genes interact.

4.3 Discussion

CLL cells die rapidly by apoptosis when cultured in the laboratory (207), therefore supportive co-culture systems are employed by many laboratories in order to keep CLL cells alive for further functional investigations into the disease mechanism. A common approach is to attempt to recreate some of the interactions in a culture flask which may occur in the tumour microenvironment. Endothelial cells are one class of accessory cell reported to exist in the tumour microenvironment in proliferation centres in the lymph node (52) and have been a focus of co-culture work carried out in our laboratory (1, 61, 173). This chapter has focused on determining whether there are any common transcriptional changes which occur in CLL cells cultured in two different endothelial cell systems in order to help identify any important mechanisms CLL cells may rely on for cytoprotection. Firstly, I demonstrated that both immortalised and primary endothelial cells were able to increase CLL cell viability in co-cultures compared to CLL cells alone in liquid culture. This observation was accompanied by a change in CLL cell phenotype, typically an increase in expression of CD69, CD44 and CD38. The effects of co-culture were observed in cells isolated from all the CLL patients regardless of clinical stratification or phenotype. Next, I investigated the effects of endothelial cell co-cultures on the mRNAs expressed in primary CLL cells. CLL cells from six different patients with a variety of phenotypes were analysed in order to identify mechanisms common to all CLL subtypes, which could be responsible for the effects observed in the co-cultures. During a bioinformatics rotation, I showed that the data clustered by co-culture condition rather than CLL patient stratification. This allowed the six patients to be treated as biological replicates in order to investigate the effects of the co-culture condition only.

There is evidence to suggest that cancer cells become ‘addicted’ to particular signalling pathways, making proteins in these cascades ideal targets for intervention (208-211). This is particularly evident in CLL (212) and the BCR signalling pathway (213) where new Btk inhibitors (213, 214) are providing exciting opportunities for new targeted therapies. We reasoned that genes up-regulated by both endothelial co-culture systems were more likely to be important in maintaining CLL cell viability than those unique to one particular co-culture system. Using stringent significance and fold change criteria, 103 mRNA were found to be up-regulated by HDBEC co-culture and 134 mRNA up-regulated by HMEC-1 co-culture. Of these, 53 mRNA were significantly up-regulated by both systems but only 31 achieved both significance and a fold change >1.5 when compared with liquid culture.

Bioinformatic analyses of the genes encoding these mRNA identified a potential control mechanism involving a novel TF module in CLL. The identification of multiple TFs in modules such as the one described here may be more informative than individual TF when attempting to identify regulators of

a set of genes since TF have been shown to function as regulatory complexes. For example, cooperative regulation occurs at an enhancer located upstream of the gene encoding interferon-β (*IFNB1*) (215, 216). Individually, each of the TFs described here have been reported to be activated in CLL (p65 (REL A) (217), STAT1 (218, 219) PU.1 (220), CEBPα (221) and CEBPβ (222)) but not as a module. However, these TFs have also been described in macrophages where myeloid cell fate is determined by PU.1, CEBP*α* and CEBP*β* (223). The presence of a TF module in CLL which has been traditionally studied in the macrophage lineage may reflect the potential for rare trans-differentiation in the disease. Studies from several groups have shown evidence that some small B-cell lymphomas can transform into aggressive hematopoietic tumours of another lineage, which interestingly are of histiocytic/dendritic cell origin (222, 224-226).

4.3.1. A final caveat: concerns about RNA used in the microarray study

As discussed in section 4.2.3, it is extremely important to use high quality RNA for microarray studies and that the quality is comparable between samples. All samples were quality control assessed using an Agilent 2100 Bioanalyser as described in section 2.5.4 (Materials and Methods) before being used for microarray analysis and met standard RIN integrity criteria for inclusion in an array study. However, after becoming more familiar with the electrophoretic traces generated by the Bioanalyser and revisiting the data collected, we can observe extra peaks highlighted by arrows in **Figure 4.13B**, which are smaller than the expected 18S and 28S human rRNA peaks. It is unlikely that these extra peaks are degradation products, as they would have to be specific cleavages to produce the distinct bands in the electrophoretic trace in **Figure 4.13A**.

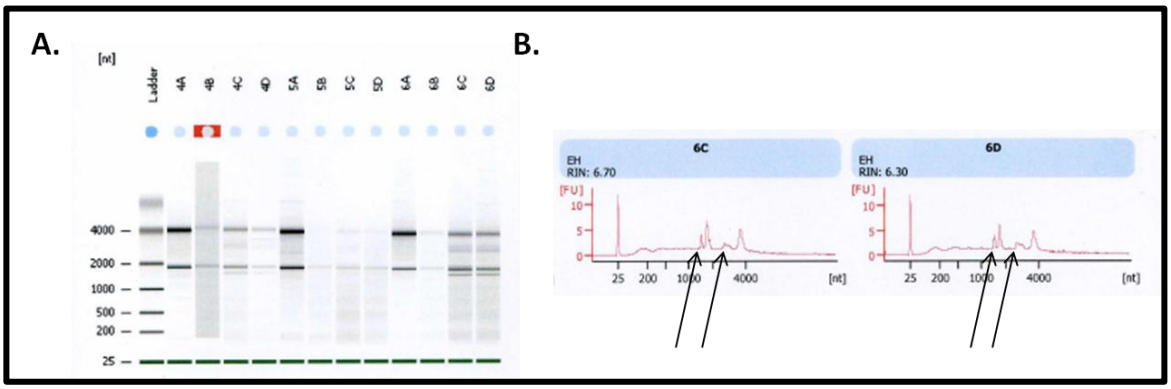

Figure 4.13 An Agilent 2100 Bioanalyser RNA 6000 Nano chip trace showing extra bands A. Shows a computer generated electrophoretic trace for the samples analysed. Numbers identify different patients and letters identify conditions A = 0 hours B = 12 hours liquid culture C = 12 hours HMEC-1 co-culture D = 12 hours HDBEC co-culture **B.** Arrows indicate extra peaks below the 18S and 28S RNA peaks identified by the algorithm for patient

One possible explanation for the presence of extra peaks could be an infection in the original cultures, such as mycoplasma which would produce 16S and 23S rRNA peaks. The HMEC-1 endothelial cell line subsequently tested positive for mycoplasma using a PCR based method (EZ-PCR Mycoplasma Test Kit). Since assays in section 4.2.3 were carried out in 24 well plates, it is possible than a mycoplasma infection could have spread between wells containing the HMEC-1 cell line and the HDBEC primary cells. However, it is not clear whether such an infection would be sufficient to produce these extra RNA peaks. Examining the Bioanalyser results as shown in **Figure 4.13** revealed that extra peaks were present in all co-cultured samples although at varying intensities compared with the expected human 18S and 28S rRNA bands. The implication of this is that the peaks could be due to a mycoplasma infection. Therefore, to establish whether the observations of the extra RNA peaks are consequences of mycoplasma infection, these co-culture experiments should be repeated using a mycoplasma negative endothelial cell line in addition to the known mycoplasma positive cell lines (in a quarantine incubator).

Chapter 5

Cell surface proteome enrichment and identification by high content LC-MS/MS

5. Cell surface proteome enrichment and identification by high content LC-MS/MS

5.1 Introduction

5.1.1. Challenge of identifying cell surface proteins

Cell surface proteins play critical roles in cell: cell recognition, signal transduction and molecular transport mechanisms (227). Membrane proteins constitute the major targets for protein based drugs because of their accessibility (228). Despite their importance in biological processes and drug discovery, membrane proteins are usually challenging to study because of the difficulty in preparing pure plasma membrane fractions, their biochemical characteristics and complex post-translational modifications (229) . Plasma membrane proteins are hydrophobic and they have to be removed from the phospholipid bilayer for analysis without aggregation and precipitation occurring. These proteins are usually found at relatively low abundance, imposing further challenges for identification and further functional experiments.

Plasma membrane protein fractions can be prepared by a number of different methods. These include differential extraction with detergents, such as Triton-X114 (230), density gradient sedimentation (typically used to isolate membrane rafts (231)) or cell lysis followed by an ultracentrifugation method to separate cellular components (232). Each has their uses, but in many cases the preparations are contaminated with abundant intracellular proteins. Therefore, a more selective enrichment method is essential for the study of low-abundance membrane and membrane-associated cell surface proteins. Biotinylation and affinity purification of cell-surface proteins, when coupled with mass spectrometry, can improve the detection of membrane proteins in low abundance and reduce contamination from other compartments, thereby increasing selectivity (233, 234).

5.1.2. CLL cell surface proteins

Interactions between CLL cells and accessory cells which recreate the tumour microenvironment and promote survival of CLL cells *in vitro* are discussed in Chapter 3. Direct cell: cell contact and soluble factors in these systems are believed to influence the phenotype of CLL cells (63, 65, 172) . *In vitro* data suggest that cell surface proteins on CLL cells play an important role in disease pathogenesis (235), therefore identifying more cell surface proteins may reveal potential targets for therapy.

5.1.3. Experimental design

In the study presented in this chapter, the cell surface proteome of CLL cells was investigated using a biotinylation enrichment method followed by LC-MS/MS. In this method, cells were first labelled with Sulfo-NHS-SS-Biotin (Pierce, Thermo Scientific; **Figure 5.1**), a thiol-cleavable amine-reactive biotinylation reagent. Cells were then lysed with a mild detergent and cell surface proteins further solubilised with gentle sonication in a water bath. The labelled proteins were then isolated with NeutrAvidin agarose (Pierce, Thermo Scientific) and the bound proteins released by incubating with a reducing agent (DTT) for identification by LC-MS/MS.

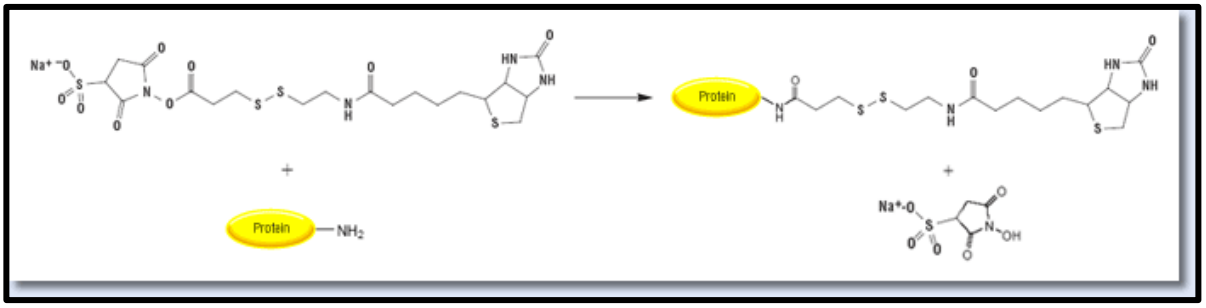

Figure 5.1 Reaction of Sulfo-NHS-SS Biotin with an accessible primary amine (Figure from http://www.piercenet.com/).

5.2 Results

5.2.1. Isolation of cell surface proteins from primary T cells

A cell surface protein biotinylation technique was tested using primary human T cells. These cells can be isolated routinely in the laboratory from a single donor leukocyte cone from the National Blood Service and were employed before using the method on precious samples from CLL patients. Non-activated, quiescent T cells were isolated from human peripheral blood, as described in section 2.2.6 (Materials and Methods). Typically, more than 98% of these cells are $CD3^+$ and not activated, as judged by CD69 expression and reported in work from our group by Lea *et al*. (236). Samples of these T cells were taken before (G_0) and 48 hours after stimulation with CD3/CD28 activation beads, when the cells have entered late G_1 and early S-phase of the first cell cycle (236). The percentage of cells in different cell cycle phases was determined by staining with PI and free FITC, as used in previous studies in our laboratory (155, 236) and described in section 2.3.1 (Materials and Methods). This cell model has been used extensively in our laboratory, including a project that identified changes in the chromatin and nuclear matrix-bound proteome (155). In the study presented here,

cell surface proteins were biotinylated using the membrane impermeable Sulfo-NHS-SS-Biotin reagent described in section 5.1.3. Cytoplasmic proteins are inaccessible to the biotin and are should not be labelled. The reaction was stopped by the addition of Pierce quenching buffer, which contains free amine groups. Cells were then lysed, cell surface proteins were solubilised and NeutrAvidin beads were used to isolate the biotinylated proteins. These proteins were then released from the NeutrAavidin scaffold using reducing and denaturing conditions (boiling in SDS sample buffer, containing DTT). In order to determine whether the method had worked, western blotting as shown in **Figure 5.2** was used to investigate the presence of a cell surface protein (CD5) and a cytoplasmic protein (Cdk6) in the fraction eluted from the NeutrAvidin matrix and the unbound which did not bind to the matrix during incubation.

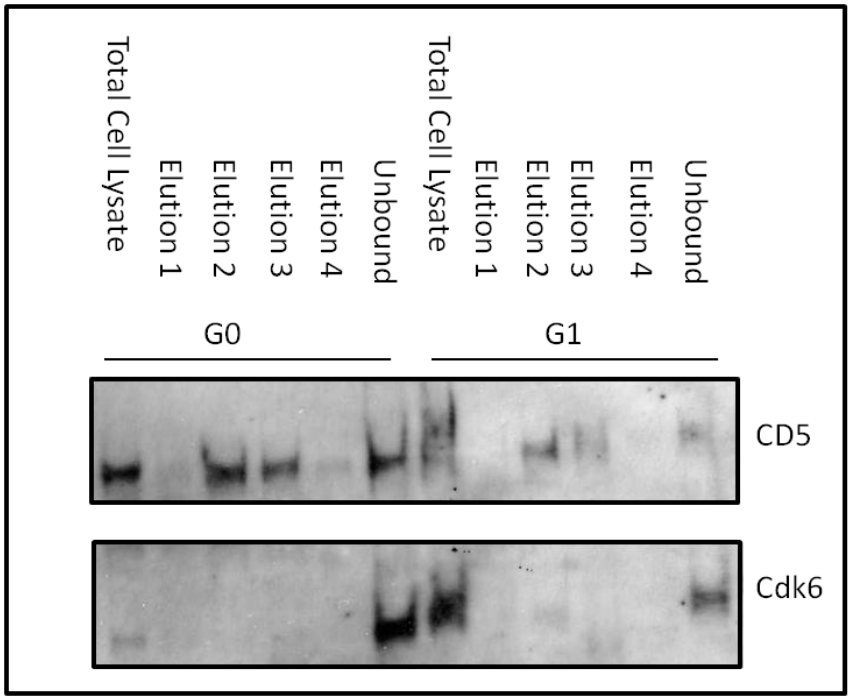

Figure 5.2 Western blot of cell surface proteins isolated from primary human T cells. Quiescent (G_0) and CD3/CD28 stimulated (G_1) human primary T cells were treated using a cell impermeable biotinylation reagent and the biotinylated proteins were isolated using NeutrAvidin beads. Total cell lysate, proteins eluted from the NeutrAvidin beads and unbound proteins were analysed by western blotting for human CD5 (cell surface protein) and Cdk6 (intracellular protein).

The cell surface protein CD5 is present in the fraction of proteins eluted from the NeutrAvidin beads, but the cytoplasmic protein Cdk6 is only present in the unbound fraction. Therefore, the data in **Figure 5.2** indicate that this method specifically biotinylated and isolated cell surface proteins rather than intracellular proteins from primary human T cells.

5.2.2. Optimisation of lysis conditions to increase cell surface protein solubilisation

Proteolytic digestion (as required for MS analyses) of complex protein mixtures can be limited by protein solubility in solution and the folded states of proteins (237). Similarly, protein complexes present in protein mixtures often require detergents to denature complexes to make the proteins more accessible and susceptible to enzymatic cleavage. The amphiphilic nature of detergents encourages unfolding of hydrophobic cell surface proteins by stabilising the unfolded state. In order to maximise the recovery of cell surface proteins for identification by LC-MS/MS, the lysis conditions were optimised. Since the design of the experiment includes shotgun proteomics, a highly expressed cell surface protein, MHC1 was used as a surrogate measure of solubility of cell surface proteins. Four commonly used lysis buffers for cell surface protein solubilisation were tested using primary human T cells. CD3/CD28 stimulated (G_1) human primary T cells were firstly treated using a cell impermeable biotinylation reagent, as described above and then the sample was divided and lysed in one of the following four buffers: (i) Pierce proprietary lysis buffer, (ii) RIPA buffer (150 mM NaCl, 1.0% (v/v) NP-40 , 0.5% (w/v) sodium deoxycholate, 0.1% (w/v) SDS, 50 mM Tris-HCl, pH 8.0, (iii) buffer containing 1% (w/v) CHAPS, 1% (w/v) OGP and (iv) buffer containing 1% (v/v) Triton X-100. These buffers were chosen because they are typically used to solubilise membrane proteins. Solubilisation of the cell surface protein MHC1 by each of the conditions and its capture with NeutrAvidin beads (the bound fraction) was analysed by western blotting (**Figure 5.3**).

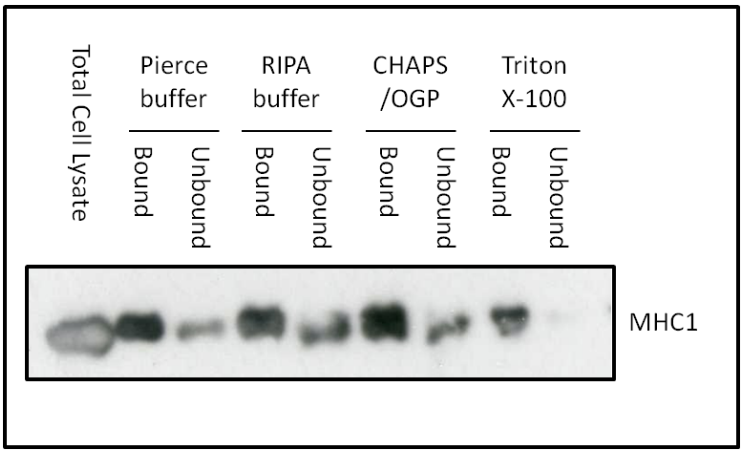

Figure 5.3 Western blot of cell surface proteins isolated from primary T cells using different lysis conditions. CD3/CD28 stimulated human primary T cells were treated using a cell impermeable biotinylation reagent and lysed in Pierce proprietary lysis buffer, RIPA buffer, a buffer containing 1% (w/v) CHAPS, 1% (w/v) OGP or 1% (v/v) Triton X-100. Total cell lysate, proteins eluted from the NeutrAvidin beads (biotinylated proteins) and unbound proteins (non biotinylated proteins) were analysed by western blotting using a mouse anti human MHC class 1 primary antibody.

The solubilisation and recovery of MHC1 cell surface protein was comparable between the Pierce proprietary lysis buffer, RIPA buffer and CHAPS/OGP buffer. In this case, Triton X-100 was an inferior detergent for protein solubilisation. Pierce proprietary cell surface lysis buffer was chosen for subsequent experiments because it is MS compatible and achieved similar recovery for the MHC1 cell surface protein to the other detergent combinations tested.

5.2.3. Analyses of HeLa total cell lysate by LC-MS/MS

Experiments were carried out to test sample preparation conditions prior to LC-MS/MS analysis while I was at the University of Texas at Austin, where I could not isolate peripheral blood T cells. Therefore, in order to test conditions, HeLa cell samples were analysed, before using the precious CLL patient samples. A total cell lysate of HeLa cells was prepared by lysis in Pierce proprietary protein lysis buffer. The samples were prepared for LC-MS/MS analysis by first removing detergent. This was done by running the samlpes into an SDS-PAGE gel (section 2.6.3 (Materials and Methods)) for a very short time. This concentrates the proteins in a small portion of the gel without separating them. The gel slice was then excised and the proteins were digested into peptides using trypsin. The full method is in section 2.6 (Materials and Methods). The tryptic peptides eluted from the gel were then separated on a C18 reverse phase column and analysed by MS/MS using the LTQ-Orbitrap or Velos-Orbitrap mass spectrometers (ThermoFisher). The analyses were carried out in collaboration with Dr. Daniel Boutz in Professor Marcotte's laboratory at the University of Texas at Austin, USA. I spent three months in the laboratory in Texas where I prepared and ran samples and learned analysis methods. Further samples were also run by Dr Boutz on my return to London. Each biological replicate was analysed three times, producing three technical injections and the data were combined for subsequent analyses. The MS/MS data were analysed by the Sequest search algorithm in the Proteome Discoverer suite and proteins were identified using Percolator (238), part of Mascot (147). Criteria for identification were that each protein had two or more independent tryptic peptides identified and proteins must be detected in at least two technical injections. Proteins were only included which had a FDR <1% at the spectral count level.

Using the criteria described above, a total of 473 proteins were identified in the HeLa total lysate and **Table 5.1** shows the top twenty most abundant proteins identified by total peptide counts summed over three replicate injections. A full list of proteins identified is available in Supplementary Table 5.2.3. Of the 473 proteins identified, 25 are annotated with only one GO term describing the cellular compartment with 2610 annotations attributed to the remaining 448. GO annotations are very general and complexities such as changes in sub-cellular localisations are not included. For example, CD44 has been shown to have roles in adhesion and cell migration at the cell surface (239) but also has a form in the nucleus where it acts as a signalling molecule (240). This makes

annotation of the identified proteins difficult. Approximately 22% of all proteins coded by the human genome are found at the cell surface (228, 241), but cell surface and transmembrane proteins are underrepresented in the literature and in high throughput data analyses (242). 81 of the 473 proteins identified in this HeLa dataset (17%) have the annotation ‘plasma membrane’ or ‘integral to plasma membrane’, which is comparable to what would be expected from a reference database of total cellular proteins.

Accession	Description	HeLa Total Lysate Control (combined peptide counts of 3 replicate injections)
ENSG00000075624	actin, beta [Source:HGNC Symbol;Acc:132]	860
ENSG00000074800	enolase 1, (alpha) [Source:HGNC Symbol;Acc:3350]	423
ENSG00000163017	actin, gamma 2, smooth muscle, enteric [Source:HGNC Symbol;Acc:145]	390
ENSG00000111640	glyceraldehyde-3-phosphate dehydrogenase [Source:HGNC Symbol;Acc:4141]	357
ENSG00000188219	POTE ankyrin domain family member E [Source:UniProtKB/Swiss-Prot;Acc:Q6S8J3]	343
ENSG00000026025	vimentin [Source:HGNC Symbol;Acc:12692]	296
ENSG00000183311	tubulin, beta [Source:HGNC Symbol;Acc:20778]	283
ENSG00000158373	histone cluster 1, H2bd [Source:HGNC Symbol;Acc:4747]	255
ENSG00000188229	tubulin, beta 2C [Source:HGNC Symbol;Acc:20771]	254
ENSG00000124635	histone cluster 1, H2bj [Source:HGNC Symbol;Acc:4761]	251
ENSG00000130402	actinin, alpha 4 [Source:HGNC Symbol;Acc:166]	236
ENSG00000137267	tubulin, beta 2A [Source:HGNC Symbol;Acc:12412]	213
ENSG00000104833	tubulin, beta 4 [Source:HGNC Symbol;Acc:20774]	208
ENSG00000196924	filamin A, alpha [Source:HGNC Symbol;Acc:3754]	207
ENSG00000169067	actin, beta-like 2 [Source:HGNC Symbol;Acc:17780]	197
ENSG00000167553	tubulin, alpha 1c [Source:HGNC Symbol;Acc:20768]	183
ENSG00000124529	histone cluster 1, H4b [Source:HGNC Symbol;Acc:4789]	180
ENSG00000182611	histone cluster 1, H2aj [Source:HGNC Symbol;Acc:4727]	180
ENSG00000127824	tubulin, alpha 4a [Source:HGNC Symbol;Acc:12407]	177
ENSG00000096384	heat shock protein 90kDa alpha (cytosolic), class B member 1 [Source:HGNC Symbol;Acc:5258]	169
ENSG00000144381	heat shock 60kDa protein 1 (chaperonin) [Source:HGNC Symbol;Acc:5261]	159
ENSG00000198211	tubulin, beta 3 [Source:HGNC Symbol;Acc:20772]	149
ENSG00000072110	actinin, alpha 1 [Source:HGNC Symbol;Acc:163]	137
ENSG00000109971	heat shock 70kDa protein 8 [Source:HGNC Symbol;Acc:5241]	135
ENSG00000080824	heat shock protein 90kDa alpha (cytosolic), class A member 1 [Source:HGNC Symbol;Acc:5253]	128

Table 5.1 Analyses of HeLa Total Lysate. HeLa cells were lysed in Pierce proprietary cell surface protein lysis buffer and samples of peptides produced by tryptic digests were analysed by LC-MS/MS. The table shows total counts from three injections for the top 20 most abundant proteins. A full list of the proteins identified is in Supplementary Table 5.2.3.

5.2.4. Identification of HeLa cell surface proteins by LC-MS/MS

In order to test the biotinylation method for identifying cell surface proteins, HeLa cells were treated with Sulfo-NHS-SS-Biotin as described in section 2.6.1 (Materials and Methods) or with the reagent diluent, PBS as a control. Cell surface proteins isolated from these samples using NeutrAvidin beads were then identified by LC-MS/MS. In order to exclude abundant proteins and to improve cell surface protein enrichment, proteins identified in the biotinylated fractions were only included in the final dataset if they were significantly observed compared to the HeLa consensus total cell lysate protein list obtained in section 5.2.3.

To determine whether a protein was significantly identified in the biotinylated sample compared to the control, a Z score was calculated using the summed counts for the biotinylated versus the summed counts for control samples from three technical repeat injections. Of the proteins identified, 188 have a significant Z score (>1.96, corresponding to a 95% confidence interval) and 100 of these have been identified as plasma membrane or cell surface proteins by a manual literature search using either GeneCards (http://www.GeneCards.org/) or NCBI (http://www.ncbi.nlm.nih.gov/) annotations. The data were then filtered using a cut off value of fold change (FC) >2. The label-free MS data of the peptides identified from cells treated with Sulfo-NHS-SS-Biotin, compared with the PBS control were quantified by a spectral counting method, which was developed by our collaborators in Texas (154) and which we used in a previous study (155). The top 25 proteins sorted by FC are shown in **Table 5.2** and the full dataset is available in Supplementary Table 5.2.4a.

Accession	Cell surface associated	Description	Z score	FOLD CHANGE
ENSG00000142949	YES	protein tyrosine phosphatase, receptor type, F [Source:HGNC Symbol;Acc:9670]	-10.92	121.45
ENSG00000196576	YES	plexin B2 [Source:HGNC Symbol;Acc:9104]	-9.32	89.32
ENSG00000138448	YES	integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51) [Source:HGNC Symbol;Acc:6150]	-8.30	71.33
ENSG00000198910	YES	L1 cell adhesion molecule [Source:HGNC Symbol;Acc:6470]	-8.06	67.47
ENSG00000005884	YES	integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor) [Source:HGNC Symbol;Acc:6139]	-6.16	40.48
ENSG00000170558	YES	cadherin 2, type 1, N-cadherin (neuronal) [Source:HGNC Symbol;Acc:1759]	-6.05	39.20
ENSG00000134247	YES	prostaglandin F2 receptor negative regulator [Source:HGNC Symbol;Acc:9601]	-6.00	38.56
ENSG00000167601	YES	AXL receptor tyrosine kinase [Source:HGNC Symbol;Acc:905]	-5.89	37.27
ENSG00000227715	YES	major histocompatibility complex, class I, A [Source:HGNC Symbol;Acc:4931]	-11.64	36.31
ENSG00000110195	YES	folate receptor 1 (adult) [Source:HGNC Symbol;Acc:3791]	-5.56	33.41
ENSG00000213949	YES	integrin, alpha 1 [Source:HGNC Symbol;Acc:6134]	-5.44	32.13
ENSG00000169855	YES	roundabout, axon guidance receptor, homolog 1 (Drosophila) [Source:HGNC Symbol;Acc:10249]	-5.44	32.13
ENSG00000150093	YES	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12) [Source:HGNC Symbol;Acc:6153]	-13.33	32.02
ENSG00000187244	YES	basal cell adhesion molecule (Lutheran blood group) [Source:HGNC Symbol;Acc:6722]	-5.32	30.84
ENSG00000145242	YES	EPH receptor A5 [Source:HGNC Symbol;Acc:3389]	-5.07	28.27
ENSG00000026508	YES	CD44 molecule (Indian blood group) [Source:HGNC Symbol;Acc:1681]	-9.61	25.54
ENSG00000106991	YES	endoglin [Source:HGNC Symbol;Acc:3349]	-4.68	24.42
ENSG00000153956	YES	calcium channel, voltage-dependent, alpha 2/delta subunit 1 [Source:HGNC Symbol;Acc:1399]	-4.47	22.49
ENSG00000149177	YES	protein tyrosine phosphatase, receptor type, J [Source:HGNC Symbol;Acc:9673]	-4.47	22.49
ENSG00000142627	YES	EPH receptor A2 [Source:HGNC Symbol;Acc:3386]	-4.40	21.85

Table 5.2 Analyses of the cell surface proteome of HeLa cells. HeLa cells were treated with biotinylation reagent Sulfo-NHS-SS-Biotin or PBS control, lysates were incubated with NeutrAvdin beads and proteins were eluted using DTT. Isolated proteins were analysed by LC-MS/MS. Of the proteins identified, proteins were selected with a Z score >1.96 and <1.96 with FC >2 compared with the control. The table shows the top 25 proteins with greatest FC and significant Z scores. A full list of all the proteins identified is in Supplementary Table 5.2.4a.

Given the issues associated with ontology annotation discussed in section 5.2.3, it may be helpful to consider the enrichment of cell surface proteins in a given sample relative to their expected observation frequency in a reference database. David Functional Analysis Tools (200, 201) searches for enriched biological themes or GO terms in a list to highlight the most relevant GO terms associated with a given gene list. This function clustering tool uses a novel algorithm to measure relationships among the annotation terms based on the degrees of their co-association. This groups the similar, redundant and heterogeneous annotation terms from the same or different resources into annotation groups, thus reducing the problem of similar redundant terms and makes the biological interpretation more focused at a group level. David functional analyses were carried out on the 137 proteins identified with a significant Z score Z >1.96 and fold change>2 in the biotinylated sample compared to the PBS treated control on HeLa cells shown in **Table 5.3**. The full dataset is available in Supplementary Tale 5.2.4a.

This dataset showed a clear enrichment for cell surface proteins as a result of biotinylation with the highest scoring annotation clusters being associated with proteins integral to the plasma membrane and the following two clusters being functions associated with cell surface proteins, namely cell-cell adhesion and cell migration. These David analyses are in agreement with the manual search described in **Table 5.2**. Of the 137 proteins identified by the cell surface biotinylation, 69.8% are annotated with the GO term ‘plasma membrane’ by David Functional Analysis, representing an enrichment of 2.5 fold compared to calculated expected observations. An even greater enrichment of 4.8 fold was observed for the GO term, ‘integral to plasma membrane’ and ‘intrinsic to plasma membrane’. Only 22 proteins were identified in both the cell surface datasets from the HeLa total lysate and the HeLa cell surface protein biotinylation enrichment. This may be due to the high abundance of these 22 proteins, since the MS method is based on statistical sampling. The proteins which were identified in the total lysate but not in the biotinylated sample may have been inaccessible to the biotin reagent, possibly due to being heavily glycosylated or they may not contain any lysine residues accessible at the cell surface for biotinylation.

Annotation Cluster 1					
Enrichment Score: 24.28					
Term	**Count**	**%**	**PValue**	**Fold Enrichment**	**FDR**
topological domain:Cytoplasmic	77.00	59.69	1.76E-26	3.41	3E-23
GO:0044459~plasma membrane part	72.00	55.81	3.73E-25	3.45	5E-22
GO:0031226~intrinsic to plasma membrane	55.00	42.64	2.08E-24	4.78	3E-21
GO:0005887~integral to plasma membrane	54.00	41.86	5.74E-24	4.80	7E-21
Annotation Cluster 2					
Enrichment Score: 19.40					
Term	**Count**	**%**	**PValue**	**Fold Enrichment**	**FDR**
topological domain:Cytoplasmic	77.00	59.69	1.76E-26	3.41	3E-23
topological domain:Extracellular	69.00	53.49	1.55E-25	3.79	2E-22
GO:0005886~plasma membrane	90.00	69.77	4.2E-24	2.52	5E-21
membrane	97.00	75.19	5.29E-23	2.31	7E-20
transmembrane region	83.00	64.34	2.42E-20	2.52	3E-17
signal peptide	68.00	52.71	2.65E-20	3.12	4E-17
glycoprotein	78.00	60.47	3.13E-20	2.69	4E-17
signal	68.00	52.71	3.27E-20	3.12	4E-17
GO:0031224~intrinsic to membrane	101.00	78.29	4.96E-20	1.95	6E-17
transmembrane	83.00	64.34	7.6E-20	2.49	1E-16
glycosylation site:N-linked (GlcNAc...)	74.00	57.36	1.1E-18	2.68	2E-15
GO:0016021~integral to membrane	94.00	72.87	6.62E-16	1.87	9E-13
disulfide bond	57.00	44.19	9.39E-15	2.91	1E-11
disulfide bond	52.00	40.31	2.42E-12	2.75	3E-09
Annotation Cluster 3					
Enrichment Score: 19.03					
Term	**Count**	**%**	**PValue**	**Fold Enrichment**	**FDR**
cell adhesion	32.00	24.81	8.26E-24	11.31	1E-20
GO:0007155~cell adhesion	41.00	31.78	1.75E-22	6.60	3E-19
GO:0022610~biological adhesion	41.00	31.78	1.84E-22	6.59	3E-19
GO:0016337~cell-cell adhesion	18.00	13.95	2.84E-10	7.35	5E-07
Annotation Cluster 4					
Enrichment Score: 7.66					
Term	**Count**	**%**	**PValue**	**Fold Enrichment**	**FDR**
GO:0009986~cell surface	24.00	18.60	1.2E-13	7.29	2E-10
GO:0006928~cell motion	21.00	16.28	5.35E-09	4.98	9E-06
GO:0016477~cell migration	14.00	10.85	9.09E-07	5.72	0.0015
GO:0051674~localization of cell	14.00	10.85	2.97E-06	5.14	0.0049
GO:0048870~cell motility	14.00	10.85	2.97E-06	5.14	0.0049
Annotation Cluster 5					
Enrichment Score: 7.43					
Term	**Count**	**%**	**PValue**	**Fold Enrichment**	**FDR**
GO:0005624~membrane fraction	27.00	20.93	2.34E-08	3.53	3E-05
GO:0000267~cell fraction	31.00	24.03	4.56E-08	3.02	6E-05
GO:0005626~insoluble fraction	27.00	20.93	4.88E-08	3.40	6E-05

Table 5.3 David Functional Analysis of HeLa cell surface proteins. The 137 proteins identified in section 5.2.4 (with significant Z scores and FC>2 in the biotinylated samples compared with the control) were analysed using David Functional Analysis Tools (http://david.abcc.ncifcrf.gov/). The top five annotation clusters are shown. The full dataset is available in Supplementary Table 5.2.4b.

The protein identified most frequently in a HeLa total lysate is the highly abundant beta actin (**Table 5.1**). This has several implications for this project. Firstly, a technical issue called ion suppression can arise due to the presence of a highly abundant protein in a sample. Ion suppression occurs when there is an overrepresented ion in a sample. The ion trap will fill with the highly abundant ion, masking signal from other low abundance ions. This is also a common problem with contaminants such as keratins which can be introduced during sample preparation before MS. Secondly, many proteins are known to bind to actin monomers, polymers or both and between 60 and 100 actin-binding proteins have been reported (243). This raises the issue of distinguishing between a specific identification of a protein and one which is co-purified with actin. This is a challenge, particularly when working with large, hydrophobic or 'sticky' proteins, many of which are present in low abundance. One solution to diminish intracellular background from studies is to pre-clear samples using biotin agarose beads and to include non-biotinylated control samples in the analyses, as described above. Another solution is to prevent aggregation of proteins in the lysis buffer and increase the solubility of membrane proteins. This can be achieved by adding glycerol to the lysis buffer or using specific lysis conditions, as discussed in 5.2.2. Another approach as described by Karhemo *et al*. (244) is the use of DNase1 to digest high-molecular weight, viscous DNA. DNase1 treatment is known to depolymerise the F-actin network. DNase1 binds to the terminal actin monomers at the pointed ends of actin filaments (245) and alongside the filament (246). It also acts as a depolymerising protein by increasing the depolymerisation rate constant of actin at the pointed filament end (247). DNase1 addition to cell extracts has been used in proteomic studies to remove cytosolic actin proteins associated with viscous DNA. This approach has been used to liberate a maximal amount of membrane proteins, while removing co-purifying contaminants (244). Therefore, the method was tested whereby DNase1 was added to the HeLa samples at 25µg/ml and processed as described in section 2.6.2.

The data in **Figure 5.4A** and **B** show that there is significant overlap between the proteins identified in the DNase1 treated sample and the untreated sample; 85% for the biotinylated samples and 89% for the PBS control samples. In addition, **Figure 5.4C** shows that DNase1 treatment results in slightly fewer proteins being identified when compared with the untreated sample. Although the DNase1 was shown to be active in a test digest using a plasmid (data not shown), actin was still highly abundant in the DNase1 treated HeLa cell samples and treatment did not increase the number or proportion of cell surface proteins substantially enough for the method to be utilised further in this study. The complete dataset is available in Supplementary Table 5.2.5.

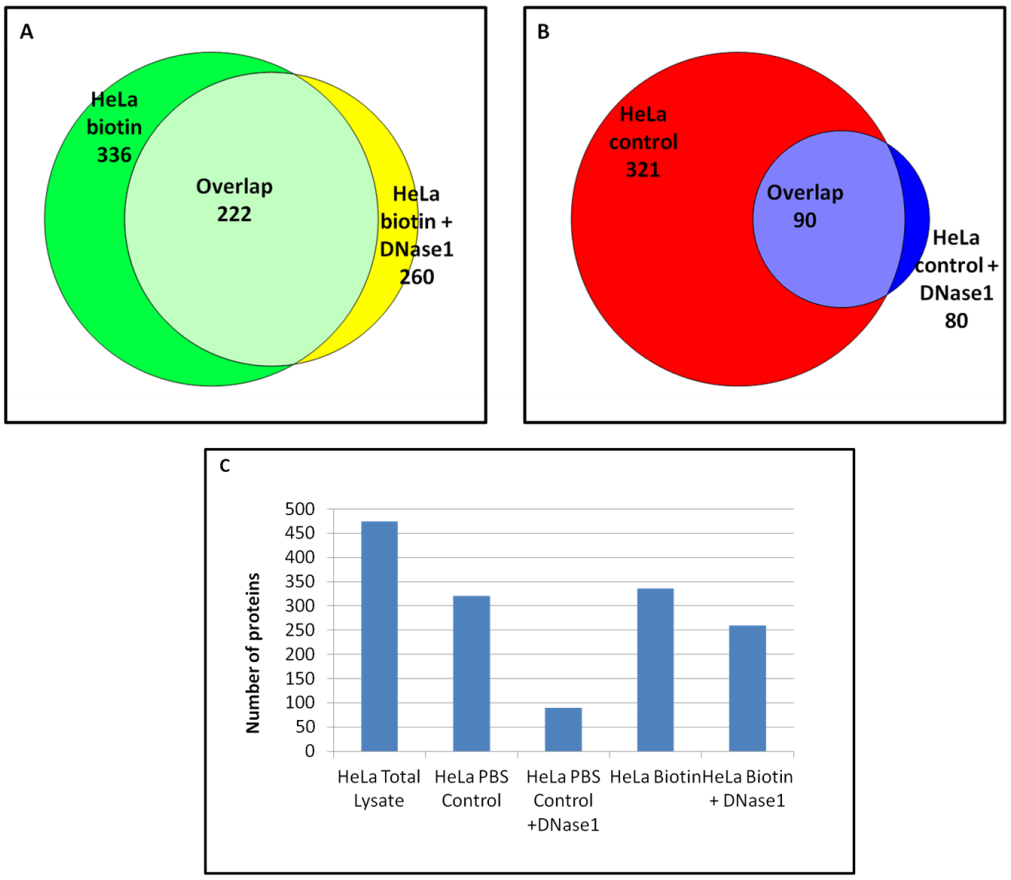

Figure 5.4 Analyses of HeLa cell surface extracts with or without DNAse1 treatment. A. HeLa cells were treated with biotinylation reagent, with or without DNAse1 treatment. **B**. HeLa cells treated with PBS control with or without DNAse1. **C**. Number of proteins identified in each condition. Protein identifications were included in the dataset if they were observed in at least two technical replicates. The graph shows total counts from three injections into the MS. The complete dataset is available in Supplementary Table 5.2.5.

5.2.6. Isolation of cell surface proteins from primary CLL patient samples

The western blot in **Figure 5.2** shows that cell surface proteins can be enriched from primary human T cells. Further experiments described above using HeLa cells and the biotinylation method described in section 5.2.1 show that proteins known to be associated with the plasma membrane can then be identified by mass spectrometry. The method was then used to identify the cell surface proteome of cells isolated from CLL patients. PBMCs are routinely isolated from peripheral blood of CLL patients using Ficoll Histopaque density centrifugation and either used immediately or stored in liquid nitrogen until required. In preliminary experiments, CLL PBMC samples were defrosted and $2x10^7$ cells were biotinylated immediately, using the same method employed to biotinylate normal T cells and HeLa cells. When this cell surface biotinylation method was applied to cryopreserved

primary CLL cells it was unsuccessful due to poor CLL cell viability. CLL cells are more fragile than the primary human T cells or HeLa cells and they began to die when defrosting and during the biotinylation procedure. The percentage of live cells was determined routinely by cell counting with trypan blue stain. The dying cells had a permeable cell membrane, allowing the biotinylation reagent into the cells, resulting in the labelling of cytoplasmic proteins. This was observed when protein samples were analysed by western blotting. I attempted to troubleshoot the application of this method to cryopreserved CLL cells by separating viable cells from dead cells after biotinylation using a subsequent round of Ficoll Histopaque density centrifugation; however the resulting yield of viable cells was too low for this method to be used further.

Next, fresh CLL cells brought straight from the clinic at King's College Hospital were biotinylated using the same protocol. The western blot in **Figure 5.5** below shows an analysis of the plasma membrane protein MHC1 and the intracellular protein Bcl-2 in samples of freshly isolated CLL cells, which were treated with bioinylation reagent or with PBS control and isolated with NeutrAvidin beads. These data suggest that the cell surface biotinylation protocol is suitable for use on freshly isolated CLL cells. Bcl-2 is observed only in the unbound fraction of both samples, which shows that the integrity of the plasma membrane was maintained during the labelling procedure. In contrast, the cell surface protein MHC1 is detected in the bound fraction of the sample treated with biotinylation reagent but not in the sample treated with PBS control.

These data indicate that MHC1 is biotinylated and is not simply binding non-specifically to the NeutrAvidin beads. A small amount of the MHC1 protein is present in the insoluble pellet fraction. The blot was deliberately over-exposed and the amount of MHC1 in this insoluble fraction constitutes perhaps only a few percent of the total. The labelling reaction itself may not be 100% efficient as the majority of MHC1 does not bind to the NeutrAvidin column. This could be because of post-translational modifications, which make lysines inaccessible to the reagent. Alternatively, some MHC1 may have been localised in a different cellular compartment. Further optimisation of lysis and labelling conditions was not undertaken to maximise recovery of biotinylated (bound) MHC1 as it was being used as a surrogate marker for total cell surface proteins. Each of the fresh CLL samples is precious and troubleshooting conditions using these primary cells is wasteful. Conditions which maintained cell viability whilst isolating some of the cell surface proteome and subsequent visualisation of isolated proteins on a coomassie stained PAGE gel was considered sufficient for further experiments.

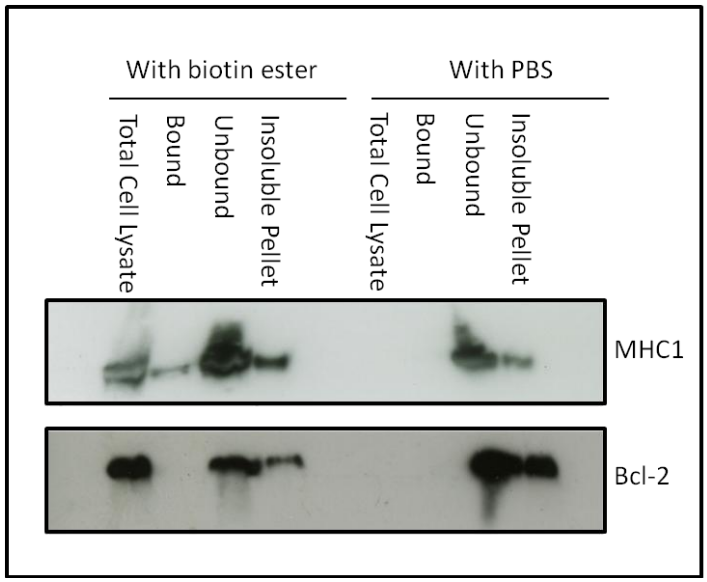

Figure 5.5 Western blot of cell surface proteins isolated from primary CLL cells. Primary CLL cells were treated using a cell impermeable biotinylation reagent or PBS control and then lysed as described. Total cell lysates were incubated with NeutrAvidin beads. The insoluble pellet after lysis, the proteins eluted from the beads (bound) and unbound proteins were analysed by western blotting for the cell surface protein, MHC1 and the intracellular protein, Bcl-2.

Maintenance of cell membrane integrity is of vital importance to this experiment and CLL cells are especially fragile and die by apoptosis the longer they are out of the body, even during basic laboratory manipulations such as centrifugation, gentle pipetting and cell counting. For this reason, the viability of the CLL cells was monitored continually throughout the isolation from whole blood and the biotinylation labelling process. Assessment of sample viability by flow cytometry of Annexin V/7AAD was not always practical, so viability was assessed by cell counting with trypan blue. Only samples which were >99% viable after biotinylation were included in the study. It is possible that the requirement for good cell viability after manipulation self selects a cohort of CLL patients for this study whose cells are more resistant to spontaneous apoptosis, however this is unavoidable. The cell numbers required for the procedure (at least 1 x 10^8) also selects for patients with high WBC and typically more aggressive disease. Initial experiments were carried out on CLL PBMCs; later experiments included a selection step for $CD5^+/CD19^+$cells (as described in section 2.2.9 (Materials and Methods)). This step was introduced to ensure that the cell surface proteome of CLL cells and not other cell types was analysed. The purity after selection was typically 98% as described in section 4.1.1 (Chapter 4) and their viability was monitored as described above.

To assess the efficiency of the biotinylation labelling reaction, Sulfo-NHS-SS-Biotin and PBS treated CLL cells were incubated with streptavidin-FITC and the proportion of streptavidin-FITC positive cells was analysed by flow cytometry. This assay was carried out for each CLL sample analysed and an example is shown in **Figure 5.6.**

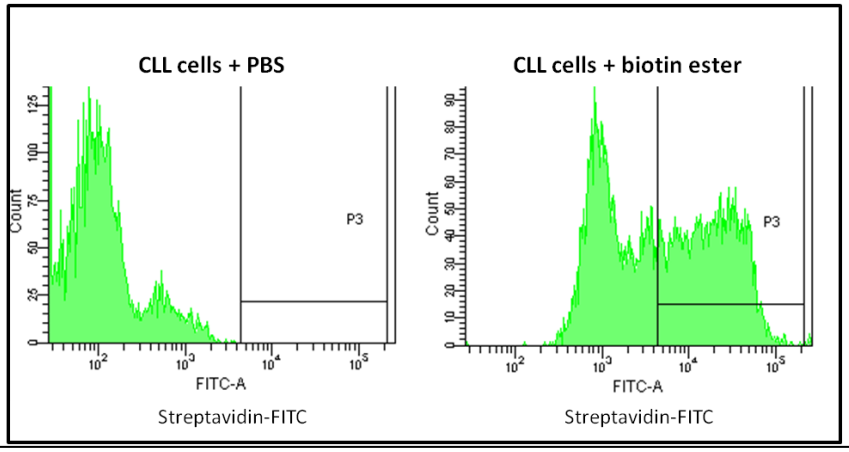

Figure 5.6 Analysis of CLL cell Biotinylation by flow cytometry. Fresh CLL cells were biotinylated as described in section 2.6.1 (Materials and Methods) and biotin labelling of the cell surface was measured by flow cytometry after incubation with streptavidin-FITC. Plots were gated on single cells and a 1% positive gate was set using the PBS control sample.

Figure 5.6 shows that at least 60% of the CLL cells are streptavidin-FITC positive and therefore labelled during the biotinylation reaction. Due to the large numbers of cells available from CLL patient samples, this level of labelling was deemed sufficient for further studies.

5.2.7. Identification of cell surface proteins from primary CLL cells by LC-MS/MS– a pilot study

Preliminary studies were performed to identify cell surface proteins isolated from three different CLL patients. Initial biotinylation experiments were performed on freshly isolated, non selected CLL PBMCs (>90% $CD5^{+}/CD19^{+}$) at King's College London and LC-MS/MS sample preparation was performed at the University of Texas at Austin, USA by Dr. Daniel Boutz. For these samples, an in-solution digest was performed and the peptides were then separated by C18 reverse phase columns and analysed by LC-MS/MS using the LTQ-Orbitrap, as described in section 2.6.6 (Materials and Methods). For these initial experiments, only biotinylated samples were analysed by LC-MS/MS.

Figure 5.7A below shows that there is considerable overlap between the proteins identified in each patient sample. The cellular localisation of the 213 proteins identified in common between all three patient samples was investigated using annotations from Panther and HumanNet classifications.

Only 2.4% of the proteins identified were annotated with the term 'plasma membrane' G0:0005886. This is much lower than would be expected from the method, given the preliminary experiments using HeLa cells described in section 5.2.4, where the same methods produced very efficient enrichment of cell surface and membrane proteins, which were17% of the total identified and were enriched to almost 60% when biotinylated proteins were analysed.

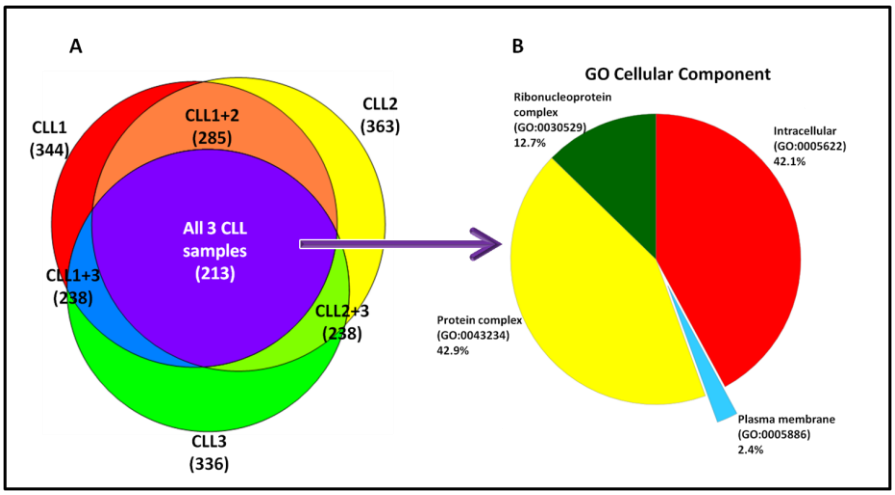

Figure 5.7 Proteins identified from cell surface fractions of primary CLL cells. Cell surface proteins were isolated from cells of three CLL patients and samples were analysed by LC-MS/MS. Protein identifications were included in the dataset if they were observed in at least two technical replicates. **A.** Shows overlap between different CLL patient datasets. **B.** The sub cellular locations of proteins identified were determined using Panther classifications (http://www.pantherdb.org/) and HumanNet (http://www.functionalnet.org/humannet/).

The CLL LC-MS/MS data contains a larger proportion of intracellular proteins, as illustrated in **Figure 5.7**. One explanation for this observation could simply be due to differences in cell type and cell structure. CLL cells are comparatively much smaller, with a smaller cytoplasm and produce less protein per cell than HeLa. The CLL cells are also more spherical than HeLa and so have a smaller cell surface per cell. However, some of the non plasma membrane proteins identified have recently been reported to be able to bind to stereotyped BCRs in CLL (248). These proteins include vimentin, cofilin 1 and non-muscle myosin heavy chain 9 (also known as non muscle myosin heavy chain IIA), which are highlighted yellow in **Table 5.4**. The complete dataset is shown in Supplementary Table 5.2.7.Other proteins were identified that could be co-purifying with cell surface proteins. These include ezrin and moesin, two of the ERM proteins (Ezrin, Radixin and Moesin) which have been shown to associate with a positively charged amino acid cluster in the juxta-membrane cytoplasmic domain of CD44, CD43 and intracellular adhesion molecule 2 (ICAM-2) (249). Therefore, it is possible that some of the cytoplasmic proteins identified are being co-purified due to interactions with cell surface proteins.

Reference	Gene ID	GO Process	GO Component	GO Function	HGNC symbol	Total Count
vimentin	ENSG00000026025	cell motility	cytoplasm	structural constituent of cytoskeleton	VIM	351
hemoglobin, beta	ENSG00000244734	transport	hemoglobin complex	oxygen transporter activity	HBB	308
HLA class I histocompatibility antigen, B-8 alpha chain Precursor	ENSG00000223532					184
hemoglobin, delta	ENSG00000223609	transport	hemoglobin complex	oxygen transporter activity	HBD	183
actin, beta	ENSG00000075624	cell motility	soluble fraction	nucleotide binding	ACTB	182
actin, gamma 1	ENSG00000184009	cell motility	soluble fraction	nucleotide binding	ACTG1	182
HLA class I histocompatibility antigen, B-48 alpha chain Precursor	ENSG00000234745					173
HLA class I histocompatibility antigen, A-11 alpha chain Precursor	ENSG00000206503	antigen processing	integral to PM	protein binding		170
HLA class I histocompatibility antigen, A-80 alpha chain Precursor	ENSG00000224320					164
protein tyrosine phosphatase, receptor type, C	ENSG00000081237	negative regulation of T cell cytotoxicity	integral to PM	protein tyrosine phosphatase activity	PTPRC|PTPRC	159
major histocompatibility complex, class I, A	ENSG00000235657				HLA-A|HLA-A	157
actin, alpha 1, skeletal muscle	ENSG00000143632	muscle contraction	stress fiber	nucleotide binding	ACTA1	151
actin, alpha, cardiac muscle 1	ENSG00000159251	muscle contraction	cytoskeleton	nucleotide binding	ACTC1	151
HLA class I histocompatibility antigen, A-32 alpha chain Precursor	ENSG00000223980					147
POTE ankyrin domain family, member E	ENSG00000188219	na	na	na	POTEE	146
POTE ankyrin domain family, member F	ENSG00000196604				POTEF	146
H2A histone family, member J	ENSG00000111332				H2AFJ	139
histone cluster 1, H2ad	ENSG00000196866	na	na	na	HIST1H2AD	139
histone cluster 1, H2ai	ENSG00000196747	na	na	na	HIST1H2AI	139
histone cluster 1, H2aj	ENSG00000182611	na	na	na	HIST1H2AJ	139
histone cluster 2, H2aa3	ENSG00000183558	nucleosome assembly	nucleosome	DNA binding	HIST2H2AA3	139
histone cluster 2, H2ac	ENSG00000184260	nucleosome assembly	nucleosome	DNA binding	HIST2H2AC	139
histone cluster 1, H2ah	ENSG00000184825	nucleosome assembly	nucleosome	DNA binding	HIST1H2AH	139
HLA class I histocompatibility antigen, Cw-2 alpha chain Precursor	ENSG00000228299					137
hemoglobin, epsilon 1	ENSG00000213931	transport	hemoglobin complex	oxygen transporter activity	HBE1	131
hemoglobin, gamma A	ENSG00000213934	transport	hemoglobin complex	oxygen transporter activity	HBG1	131
hemoglobin, gamma G	ENSG00000196565	transport	hemoglobin complex	oxygen transporter activity	HBG2|HBG2	131
HLA class I histocompatibility antigen, A-25 alpha chain Precursor	ENSG00000206505					125
histone cluster 1, H2ab	ENSG00000137259	nucleosome assembly	nucleosome	DNA binding	HIST1H2AB	123
histone cluster 1, H2ac	ENSG00000180573	nucleosome assembly	nucleosome	DNA binding	HIST1H2AC	123
histone cluster 3, H2a	ENSG00000181218	nucleosome assembly	nucleosome	DNA binding	HIST3H2A	123
major histocompatibility complex, class I, C	ENSG00000204525	ciliary or flagellar motility	axonemal dynein complex	microtubule motor activity	HLA-C|HLA-C	121
Major histocompatibility complex, class I, C	ENSG00000233841					121
HLA class I histocompatibility antigen, A-29 alpha chain Precursor	ENSG00000231834					120
keratin 1	ENSG00000167768	complement activation	cytoskeleton	receptor activity	KRT1	119
Major histocompatibility complex, class I, A Fragment	ENSG00000227715					116
HLA class I histocompatibility antigen, Cw-1 alpha chain Precursor	ENSG00000237022					116
peripherin	ENSG00000135406	na	intermediate filament	structural molecule activity	PRPH	115
POTE ankyrin domain family, member M	ENSG00000196834				POTEI|POTEM	112
HLA class I histocompatibility antigen, B-52 alpha chain Precursor	ENSG00000232126					110
histone cluster 1, H4a	ENSG00000196176	na	na	na	HIST1H4A	104
peroxiredoxin 1	ENSG00000117450	skeletal development	na	oxidoreductase activity	PRDX1	99
HLA class I histocompatibility antigen, B-35 alpha chain Precursor	ENSG00000206450	antigen processing and presentation of peptide antigen via MHC class I	membrane fraction	MHC class I receptor activity		99
ubiquitin C	ENSG00000150991	protein modification	na	na	UBC	99
major histocompatibility complex, class II, DR alpha precursor	ENSG00000204287	antigen processing and presentation of peptide or polysaccharide antigen via MHC class II	lysosome	MHC class II receptor activity		97
major histocompatibility complex, class II, DR alpha	ENSG00000206308				HLA-DRA|HLA-D	97
hemoglobin, alpha 1	ENSG00000206172	na	na	na	HBA1	90
POTE ankyrin domain family, member J	ENSG00000222038				POTEJ	89
heat shock 70kDa protein 9	ENSG00000113013	protein folding	cytoplasm	nucleotide binding	HSPA9	88
HLA class I histocompatibility antigen, B-49 alpha chain Precursor	ENSG00000224608					88
major histocompatibility complex, class I, B	ENSG00000228964				HLA-B|HLA-B|H	88
malate dehydrogenase 2, NAD	ENSG00000146701	glycolysis	mitochondrion	L-lactate dehydrogenase activity	MDH2	88
H2A histone family, member V	ENSG00000105968	nucleosome assembly	nucleosome	DNA binding	H2AFV	86
H2A histone family, member Z	ENSG00000164032	nucleosome assembly	nucleosome	DNA binding	H2AFZ	86
H2A histone family, member X	ENSG00000188486	double-strand break repair via HR	nucleosome	damaged DNA binding	H2AFX	80
HLA class I histocompatibility antigen, Cw-8 alpha chain Precursor	ENSG00000206435					80
histone cluster 1, H2aa	ENSG00000164508	nucleosome assembly	nucleosome	DNA binding	HIST1H2AA	80
histone cluster 2, H2ab	ENSG00000184270	nucleosome assembly	nucleosome	DNA binding	HIST2H2AB	80
HLA class II histocompatibility antigen, DRB1-15 beta chain Precursor	ENSG00000196126	antigen processing and presentation of peptide or polysaccharide antigen via MHC class II	membrane	MHC class II receptor activity		77
major histocompatibility complex, class II, DR beta 1	ENSG00000206306				HLA-DRB1|HLA-	77
HLA class II histocompatibility antigen, DRB1-10 beta chain Precursor	ENSG00000228080					77
glyceraldehyde-3-phosphate dehydrogenase	ENSG00000111640	glucose metabolic process	cytoplasm	glyceraldehyde-3-phosphate dehydrogenase	GAPDH	73
heterogeneous nuclear ribonucleoprotein A2/B1	ENSG00000122566	nuclear mRNA splicing	nucleus	nucleotide binding	HNRNPA2B1	66
microRNA 7-1	ENSG00000165119	mRNA processing	nucleus	nucleic acid binding	HNRNPK|MIR7-	65
keratin 10	ENSG00000186395	epidermis development	keratin filament	protein binding	KRT1	65
actin, alpha 2, smooth muscle, aorta	ENSG00000107796	na	cytoskeleton	nucleotide binding	ACTA2	62
actin, gamma 2, smooth muscle, enteric	ENSG00000163017	na	cytoskeleton	nucleotide binding	ACTG2	62
cofilin 1	ENSG00000172757	anti-apoptosis	intracellular	actin binding	CFL1	55
heterogeneous nuclear ribonucleoprotein H1	ENSG00000169045	mRNA processing	nucleus	nucleotide binding	HNRNPH1	55
major histocompatibility complex, class I, E precursor	ENSG00000204592	antigen processing and presentation of peptide antigen via MHC class I	membrane	MHC class I receptor activity		51
major histocompatibility complex, class I, E	ENSG00000206493				HLA-E	51
HLA class I histocompatibility antigen, alpha chain E Precursor	ENSG00000229252					51
myosin, heavy chain 9, non-muscle	ENSG00000100345	cellular morphogenesis during differentiation	stress fiber	microfilament motor activity	MYH9	50
major histocompatibility complex, class II, DR beta 5	ENSG00000198502	presentation of peptide or polysaccharide antigen via	integral to PM	MHC class II receptor activity	HLA-DRB5	48

Table 5.4 Most abundant proteins identified from cell surface isolations of CLL cells from 3 patients, ranked by peptide count. GO ontology annotations were retrieved from HumanNet (http://www.functionalnet.org/humannet/). A full list of the proteins identified is in Supplementary Table 5.2.7.

5.2.8. Identification of cell surface proteins from primary CLL cells by LC-MS/MS – optimisation

In order to determine whether some of the proteins observed in **Table 5.4** were binding non-specifically to the NeutrAvidin beads, the experiment was modified to include a sample treated with PBS to act as a control (the biotinylation reagent is reconstituted in PBS). In these experiments, $CD5^{+}CD19^{+}$ cells were selected from PBMCs isolated from CLL patient samples as described in section 2.2.9 (Materials and Methods), then half of the cells were treated with biotinylation reagent and half with PBS.

Cells from four different CLL patients were analysed by LC-MS/MS and total peptide counts were summed from three replicate injections. Of the proteins identified, proteins were selected with a significant Z score (>1.96 or <1.96) and with FC >2 in the biotinylated sample compared to the PBS control for each pair of samples. Next, a consensus dataset was formed of proteins identified in all four CLL patient samples (average FC >2 and a combined-Z score (>1.96 or <1.96) was calculated by summing the individual Z-scores calculated for each patient). **Table 5.5** shows a consensus list of the top 20 proteins with greatest FC and significant Z scores across all four patient samples and GO ontology annotations, which were retrieved from HumanNet. 169 proteins were identified which satisfied the filtering requirements and 22 (13%) of these proteins are cell surface associated by ontology annotation and 22% of all proteins are localised at the cell surface (228, 241). **Table 5.5** below shows the top 20 proteins identified by FC for the CLL cell surface enrichment dataset. The complete dataset is in Supplementary Table 5.2.8. In section 5.2.7 only 2.7% of proteins were identified as cell surface associated when a matching PBS control for non-specific binding was not available. Therefore, filtering the dataset against a PBS control to remove non-specific protein identifications increases the proportion of protein identifications associated with the cell surface. However, the cell surface protein enrichment achieved is not as great as that observed for HeLa cells in section 5.2.4 where this method enriched cell surface proteins to almost 70% of total protein identifications in the dataset.

The 169 proteins identified as the consensus cell surface protein list were analysed by David Functional Annotation Tools, as described in section 5.2.4 and **Table 5.3**. Cell surface and plasma

membrane proteins were not significantly enriched in this clustering analysis. This is unsurprising given that only 13% of the proteins had been annotated as cell surface proteins by HumanNet ontology analysis. The presence of annotation cluster 3 (actin and cytoskeletal related proteins) in the analysis (**Table 5.6**) is also not surprising given the high abundance of this protein, as discussed in section 5.2.3 and the nature of the effects of sampling in MS techniques. Annotation cluster 2 includes regulation of apoptosis and cell death related terms. This is particularly interesting given that the CLL cells are known to over express many anti-apoptotic proteins (250-252) and undergo apoptosis in culture, as discussed in Chapter 1 and 3.

Table 5.5 Analyses of the cell surface proteome of CLL $CD5^+CD19^+$ cells. CLL cells were treated with biotinylation reagent or PBS control. Isolated proteins were analysed by LC-MS/MS. Proteins were selected with a Z score >1.96 and <1.96 with FC >2. The Table shows the top 20 proteins with greatest FC and significant Z scores. GO ontology annotations were retrieved from HumaNet (http://www.functionalnet.org/humannet/). The complete dataset is available in Supplementary Table 5.2.8a.

Accession	Description	Combined Z-Score	Average Fold Change	GO Process	GO Component	GO Function
ENSG00000005961	ENST00000262407 integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41)	7.33	13.24	cell adhesion; cell-matrix adhesion; integrin-mediated signaling pathway;	focal adhesion; integrin complex; external side of plasma membrane; membrane; integral to membrane;	receptor activity; calcium ion binding; protein binding; identical protein binding; extracellular matrix binding;
ENSG00000136167	ENST00000323076 lymphocyte cytosolic protein 1 (L-plastin)	9.69	11.14	actin filament bundle formation;	ruffle; phagocytic cup; cytoplasm; cytosol; actin filament;	actin binding; calcium ion binding; identical protein binding; actin filament binding;
ENSG00000108518	ENST00000225655 profilin 1	8.51	10.72	neural tube closure; regulation of transcription from RNA polymerase II promoter; cytoskeleton organization and biogenesis; actin cytoskeleton organization and biogenesis;	nucleus; cytoplasm; actin cytoskeleton;	actin monomer binding; protein binding;
ENSG00000177156	ENST00000319006 transaldolase 1	6.24	9.38	carbohydrate metabolic process; pentose-phosphate shunt; metabolic process;	cytoplasm;	catalytic activity; transaldolase activity; protein binding; transferase activity;
ENSG00000074800	ENST00000234590 enolase 1, (alpha)	7.43	9.35	negative regulation of transcription from RNA polymerase II promoter; glycolysis; transcription; negative regulation of cell growth;	phosphopyruvate hydratase complex; nucleus; cytoplasm;	magnesium ion binding; transcription factor activity; transcription corepressor activity; phosphopyruvate hydratase activity; protein binding; plasminogen activator activity; lyase activity;
ENSG00000101444	ENST00000217426 adenosylhomocysteinase	5.69	9.01	one-carbon compound metabolic process;	cytoplasm;	adenosylhomocysteinase activity; hydrolase activity;
ENSG00000130985	ENST00000335972 ubiquitin-like modifier activating enzyme 1	5.72	8.99	DNA replication; ubiquitin cycle;	na	nucleotide binding; catalytic activity; ubiquitin activating enzyme activity; protein binding; ATP binding; ligase activity;
ENSG00000089220	ENST00000261313 phosphatidylethanolamine binding protein 1	5.56	8.53	na	na	nucleotide binding; serine-type endopeptidase inhibitor activity; ATP binding; lipid binding; phosphatidylethanolamine binding;
ENSG00000111348	ENST00000542276 Rho GDP dissociation inhibitor (GDI) beta	5.47	8.01	cell motility; immune response; negative regulation of cell adhesion; Rho protein signal transduction; multicellular organismal development; actin cytoskeleton organization and biogenesis;	cytoplasm; cytoskeleton; cytoplasmic membrane-bound vesicle;	Rho GDP-dissociation inhibitor activity; GTPase activator activity; protein binding;
ENSG00000102144	ENST00000373316 phosphoglycerate kinase 1	7.04	7.78	glycolysis; phosphorylation;	na	nucleotide binding; phosphoglycerate kinase activity; ATP binding; transferase activity;
ENSG00000109971	ENST00000227378 heat shock 70kDa protein 8	6.74	7.74	protein folding; response to unfolded protein;	intracellular; nucleus; cell surface;	nucleotide binding; protein binding; ATP binding; ATPase activity, coupled;
ENSG00000137076	ENST00000314888 talin 1	7.51	6.95	cell motility; cytoskeletal anchoring; intercellular junction assembly;	ruffle; cytosol; cytoskeleton; intercellular junction; focal adhesion;	actin binding; structural constituent of cytoskeleton; binding; protein binding; vinculin binding; LIM domain binding;
ENSG00000128340	ENST00000249071 ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)	3.99	6.80	chemotaxis; signal transduction; small GTPase mediated signal transduction; cell projection biogenesis; actin cytoskeleton organization and biogenesis;	intracellular; membrane fraction; nuclear envelope; cytoplasm;	nucleotide binding; GTPase activity; protein binding; GTP binding;
ENSG00000204388	ENST00000375650 heat shock 70kDa protein 1B	3.92	6.74	protein folding; response to unfolded protein;	na	nucleotide binding; ATP binding;
ENSG00000117592	ENST00000340385 peroxiredoxin 6	4.87	6.45	response to oxidative stress; phospholipid catabolic process; lipid catabolic process;	lysosome; cytosol;	phospholipase A2 activity; oxidoreductase activity; hydrolase activity; peroxiredoxin activity;
ENSG00000143549	ENST00000341485 tropomyosin 3	6.90	6.37	cell motility; regulation of muscle contraction;	cytoskeleton; muscle thin filament tropomyosin;	actin binding;
ENSG00000080824	ENST00000216281 heat shock protein 90kDa alpha (cytosolic), class A member 1	5.75	6.33	protein folding; mitochondrial transport; response to unfolded protein; signal transduction; protein refolding; positive regulation of nitric oxide biosynthetic process;	cytosol;	nucleotide binding; ATP binding; nitric-oxide synthase regulator activity; TPR domain binding; protein homodimerization activity; unfolded protein binding;
ENSG00000163737	ENST00000296029 platelet factor 4	3.63	6.32	immune response; negative regulation of angiogenesis; cytokine and chemokine mediated signaling pathway; platelet activation; leukocyte chemotaxis; negative regulation of megakaryocyte differentiation;	extracellular region; extracellular space;	chemokine activity; heparin binding;
ENSG00000172757	ENST00000308162 cofilin 1 (non-muscle)	5.31	6.16	anti-apoptosis; Rho protein signal transduction; actin cytoskeleton organization and biogenesis;	intracellular; nucleus; cytoplasm; cytoskeleton;	actin binding; protein binding;
ENSG00000170542	ENST00000380698 serpin peptidase inhibitor, clade B (ovalbumin), member 9	4.55	6.13	anti-apoptosis; signal transduction;	cytoplasm; cytosol;	serine-type endopeptidase inhibitor activity; protein binding;

Annotation Cluster 1					
Enrichment Score: 17.27					
Term	**Count**	**%**	**PValue**	**Fold Enrichment**	**FDR**
GO:0042470~melanosome	22	13.84	1.55E-22	21.94	2.04E-19
GO:0048770~pigment granule	22	13.84	1.55E-22	21.94	2.04E-19
GO:0031988~membrane-bounded vesicle	36	22.64	4.62E-17	5.63	6.07E-14
GO:0016023~cytoplasmic membrane-bounded vesicle	34	21.38	1.04E-15	5.49	1.31E-12
GO:0031982~vesicle	37	23.27	1.35E-15	4.90	1.75E-12
GO:0031410~cytoplasmic vesicle	35	22.01	1.48E-14	4.84	1.96E-11
Annotation Cluster 2					
Enrichment Score: 7.65					
Term	**Count**	**%**	**PValue**	**Fold Enrichment**	**FDR**
GO:0042981~regulation of apoptosis	32	20.13	4.68E-10	3.64	7.83E-07
GO:0043067~regulation of programmed cell death	32	20.13	5.97E-10	3.60	9.98E-07
GO:0010941~regulation of cell death	32	20.13	6.53E-10	3.59	1.09E-06
GO:0043066~negative regulation of apoptosis	18	11.32	2.94E-07	4.65	4.91E-04
GO:0043069~negative regulation of programmed cell death	18	11.32	3.58E-07	4.58	5.99E-04
GO:0060548~negative regulation of cell death	18	11.32	3.72E-07	4.57	6.23E-04
GO:0006916~anti-apoptosis	14	8.81	3.77E-07	6.21	6.30E-04
Annotation Cluster 3					
Enrichment Score: 6.83					
Term	**Count**	**%**	**PValue**	**Fold Enrichment**	**FDR**
GO:0030029~actin filament-based process	19	11.95	1.22E-10	7.21	2.04E-07
GO:0015629~actin cytoskeleton	20	12.58	1.49E-10	6.60	1.96E-07
GO:0030036~actin cytoskeleton organization	18	11.32	3.71E-10	7.28	6.20E-07
GO:0003779~actin binding	20	12.58	7.37E-09	5.24	1.02E-05
actin-binding	15	9.43	1.70E-08	7.35	2.29E-05
actin binding	8	5.03	3.28E-08	24.19	4.41E-05
GO:0007010~cytoskeleton organization	21	13.21	5.13E-08	4.40	8.58E-05
cytoskeleton	22	13.84	6.07E-08	4.18	8.16E-05
GO:0007015~actin filament organization	10	6.29	7.85E-08	12.70	1.31E-04
GO:0008092~cytoskeletal protein binding	22	13.84	3.70E-07	3.73	5.13E-04
GO:0005856~cytoskeleton	35	22.01	6.70E-06	2.25	0.00881
GO:0043228~non-membrane-bounded organelle	50	31.45	6.40E-05	1.71	0.08414
GO:0043232~intracellular non-membrane-bounded organelle	50	31.45	6.40E-05	1.71	0.08414
GO:0044430~cytoskeletal part	18	11.32	0.03699	1.68	39.0814
Annotation Cluster 4					
Enrichment Score: 5.65					
Term	**Count**	**%**	**PValue**	**Fold Enrichment**	**FDR**
SP_PIR_KEYWORDS Chaperone	14	8.81	6.25E-10	10.72	8.40E-07
GO:0051082~unfolded protein binding	14	8.81	7.47E-10	10.40	1.04E-06
molecular chaperone	6	3.77	4.28E-08	55.83	5.76E-05
GO:0006457~protein folding	13	8.18	5.02E-07	6.71	8.39E-04
GO:0006986~response to unfolded protein	8	5.03	1.15E-05	10.30	0.01929
stress response	7	4.40	1.69E-05	12.83	0.02272
GO:0051789~response to protein stimulus	8	5.03	1.63E-04	6.83	0.27284
GO:0010033~response to organic substance	20	12.58	3.00E-04	2.54	0.50123
SP_PIR_KEYWORDS ATP	7	4.40	0.01343	3.59	16.6268
Annotation Cluster 5					
Enrichment Score: 5.53					
Term	**Count**	**%**	**PValue**	**Fold Enrichment**	**FDR**
site:Interaction with phosphoserine on interacting protein	7	4.40	2.02E-12	120.21	2.94E-09
SM00101:14_3_3	7	4.40	2.90E-12	110.72	2.70E-09
IPR000308:14-3-3 protein	7	4.40	4.60E-12	104.77	6.28E-09
PIRSF000868:14-3-3	7	4.40	1.29E-10	59.65	1.46E-07
PIRSF000868:14-3-3 protein	7	4.40	1.29E-10	59.65	1.46E-07
hsa04722:Neurotrophin signaling pathway	12	7.55	1.77E-05	5.07	0.02045
GO:0008104~protein localization	25	15.72	2.66E-05	2.59	0.04442
GO:0019904~protein domain specific binding	15	9.43	3.16E-05	3.87	0.04379
GO:0045184~establishment of protein localization	22	13.84	8.54E-05	2.61	0.14279
GO:0046907~intracellular transport	20	12.58	9.08E-05	2.78	0.15183
GO:0015031~protein transport	21	13.21	2.17E-04	2.52	0.36249
GO:0034613~cellular protein localization	13	8.18	0.00176	2.89	2.89893

Table 5.6 David Functional Analysis of consensus CLL cell surface proteins. The 169 proteins identified in section 5.2.4 (with significant Z scores and FC>2 in the biotinylated samples) were analysed using David Functional Analysis Tools (http://david.abcc.ncifcrf.gov/). The top five annotation clusters are shown. The complete dataset is available in Supplementary Table 5.2.8b.

5.2.9. Analyses of CLL cell total cell lysates by LC-MS/MS

The relative representation of cell surface proteins in biotinylated samples isolated from HeLa cells was increased when these lists were filtered with respect to the abundance of proteins identified by analyses of total cell lysates (section 5.2.4). In order to investigate the relative representation of cell surface proteins in CLL cells, LC-MS/MS analyses were performed on total cell lysates from three different CLL patients. Cells were lysed in Pierce proprietary lysis buffer, as described in Materials and Methods and analysed by LC-MS/MS. A consensus list of proteins identified in all three patient samples was created from protein identifications observed in at least two of the technical replicates from each of the three patients analysed. Once again, actin and its binding partners are highly abundant. A large degree of overlap in protein identifications was observed between each of the CLL patients, as shown in **Figure 5.8**.

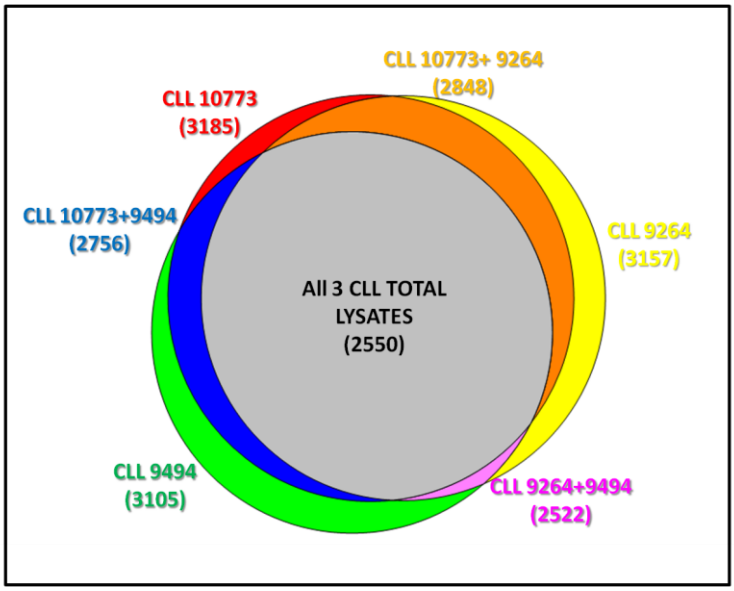

Figure 5.8 Proteins identified from total lysates of cells isolated from three CLL patients. Total cell protein samples were analysed by LC-MS/MS. Protein identifications were included in the dataset if they were observed in at least two technical replicates.

Table 5.7 below shows the total counts from three injections for the top 20 most abundant proteins in the consensus list. A complete dataset is available in Supplementary Table 5.2.9. In order to improve cell surface protein enrichment, proteins identified in biotinylated protein isolates were only included in the final dataset if they were significantly observed compared to the CLL consensus total lysate protein list. However, this comparison using the total lysate did not increase the % of cell surface annotated proteins over the enrichment achieved by filtering the biotin sample against the PBS control.

Description	ΣCoverage	Σ# Proteins	Σ# Unique Peptides	CLL_10773	CLL_9264	CLL_9494	Total counts	Average count
ENST00000331789 actin, beta [Source:HGNC Symbol;Acc:132]	95.20	18	24	2215	2160	2810	7185	2395.00
ENST00000366684 actin, alpha 1, skeletal muscle [Source:HGNC Symbol;Acc:129]	59.42	13	4	1052	1000	1325	3377	1125.67
ENST00000224237 vimentin [Source:HGNC Symbol;Acc:12692]	84.98	16	56	891	1220	891	3002	1000.67
ENST00000224784 actin, alpha 2, smooth muscle, aorta [Source:HGNC Symbol;Acc:130]	59.42	14	4	867	809	1081	2757	919.00
ENST00000229239 glyceraldehyde-3-phosphate dehydrogenase [Source:HGNC Symbol;Acc:4141]	86.87	4	29	931	936	886	2753	917.67
ENST00000377364 histone cluster 1, H4b [Source:HGNC Symbol;Acc:4789]	71.84	1	21	823	972	763	2558	852.67
ENST00000314888 talin 1 [Source:HGNC Symbol;Acc:11845]	66.31	2	108	526	588	1359	2473	824.33
ENST00000360319 filamin A, alpha [Source:HGNC Symbol;Acc:3754]	63.85	8	112	569	555	1297	2421	807.00
ENST00000216181 myosin, heavy chain 9, non-muscle [Source:HGNC Symbol;Acc:7579]	55.61	14	106	589	400	1269	2258	752.67
ENST00000331380 histone cluster 2, H2ac [Source:HGNC Symbol;Acc:4738]	81.40	2	10	669	832	534	2035	678.33
ENST00000323076 lymphocyte cytosolic protein 1 (L-plastin) [Source:HGNC Symbol;Acc:6528]	84.53	7	44	674	593	715	1982	660.67
ENST00000551208 tubulin, beta class I [Source:HGNC Symbol;Acc:20778]	74.48	21	6	646	671	648	1965	655.00
ENST00000289316 histone cluster 1, H2bd [Source:HGNC Symbol;Acc:4747]	73.02	19	4	673	782	462	1917	639.00
ENST00000262030 ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide [Source:HGNC Symbol;Acc:830]	72.78	7	32	666	712	524	1902	634.00
ENST00000234590 enolase 1, (alpha) [Source:HGNC Symbol;Acc:3350]	71.66	3	24	575	618	652	1845	615.00
ENST00000354667 heterogeneous nuclear ribonucleoprotein A2/B1 [Source:HGNC Symbol;Acc:5033]	80.45	2	28	711	650	440	1801	600.33
ENST00000314088 histone cluster 1, H2ac [Source:HGNC Symbol;Acc:4733]	81.54	4	5	562	715	439	1716	572.00
ENST00000369160 histone cluster 3, H2bb [Source:HGNC Symbol;Acc:20514]	67.46	10	5	573	725	414	1712	570.67
ENST00000227378 heat shock 70kDa protein 8 [Source:HGNC Symbol;Acc:5241]	59.91	16	27	512	604	582	1698	566.00
ENST00000345042 heat shock 60kDa protein 1 (chaperonin) [Source:HGNC Symbol;Acc:5261]	76.09	8	39	527	608	428	1563	521.00

Table 5.7 Consensus list of proteins identified from total lysates of three CLL patient samples. A complete dataset is available in Supplementary Table 5.2.9.

5.2.10. Identification of cell surface proteins from primary CLL cells by LC-MS/MS – Annexin V negative cells

David Functional analyses of $CD5^+CD19^+$ CLL cell surface proteins described in 5.2.8 and **Table 5.6** highlight the enrichment of a cluster of proteins involved in regulating apoptosis. One explanation for this may be the over expression of anti-apoptotic proteins by CLL, resulting in the appearance of enrichment in the peptides identified when compared with a reference dataset. If this were the case, an enrichment of anti-apoptotic proteins would be expected from a David Functional Analysis of CLL consensus total cell lysate proteins. However, the enrichment score for apoptosis related terms in the consensus total lysate was 2.74 compared to 7.65 in the CLL consensus cell surface proteins. This suggests that there is an enrichment of apoptosis related proteins specifically at the cell surface of the CLL cells analysed. Despite the over expression of anti-apoptotic proteins, it is also recognised that populations of CLL cells are dynamic with considerable cell turnover and cell death rates of up to 2% of the clone per day have been reported (35). Recent publications have also highlighted the potential for intracellular proteins to act as a source of auto antigens for the BCR in CLL (253).

I investigated whether the cell surface 'apoptosis regulation' signature identified in **Table 5.6** was present on all CLL cells or on a subset of cells. In order to do this, the $CD5^+CD19^+$ cells selected from three CLL patients were each separated into two samples and one of the samples was then treated with the Annexin V MicroBead kit (Milteyni) to remove any Annexin V positive cells. The resulting Annexin V negative fraction and the second total $CD5^+CD19^+$ sample were then treated with biotinylation reagent and their cell surface proteome was analysed as described above using the LTQ-Orbitrap. A consensus dataset was created as described in 5.2.8 from the three different CLL patient samples. 86 proteins were identified with a significant Z-score between the biotinylated and the control samples. Of these, 44 were identified with greater than a twofold enrichment in the total $CD5^+CD19^+$ population compared with the Annexin V negative cells and 22 were identified with a greater than two fold enrichment in the Annexin V negative cells compared with the total $CD5^+CD19^+$population. The proteins most significantly enriched in the Annexin V negative fraction are shown in **Table 5.8** below and the full dataset is shown in Supplementary Table 5.2.10. The proteins identified include several nuclear proteins such as high mobility group box 2 (HMGB2) and Histones. This is somewhat surprising given that these cells were selected to be Annexin V negative and therefore should have be a viable population of cells. However, these data are consistent with the expression of intracellular proteins at the cell surface of viable cells, which occurs in certain circumstances, discussed below.

Accession	Description	Average fold change	Z-score combined
ENSG00000164104	ENST00000296503 high mobility group box 2 [Source:HGNC Symbol;Acc:5000]	10.53	3.09
ENSG00000005022	ENST00000317881 solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 5 [Source:HGNC Symbol;Acc:10991]	10.53	3.09
ENSG00000138029	ENST00000545822 hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctional protein), beta subunit [Source:HGNC Symbol;Acc:4803]	10.30	3.04
ENSG00000124575	ENST00000244534 histone cluster 1, H1d [Source:HGNC Symbol;Acc:4717]	9.43	6.04
ENSG00000215021	ENST00000542912 prohibitin 2 [Source:HGNC Symbol;Acc:30306]	7.90	2.64
ENSG00000176619	ENST00000582871 lamin B2 [Source:HGNC Symbol;Acc:6638]	7.36	2.48
ENSG00000213585	ENST00000265333 voltage-dependent anion channel 1 [Source:HGNC Symbol;Acc:12669]	6.46	4.61
ENSG00000183311	ENST00000551208 tubulin, beta class I [Source:HGNC Symbol;Acc:20778]	6.40	3.44
ENSG00000078668	ENST00000022615 voltage-dependent anion channel 3 [Source:HGNC Symbol;Acc:12674]	6.14	2.28
ENSG00000198712	ENST00000361739 mitochondrially encoded cytochrome c oxidase II [Source:HGNC Symbol;Acc:7421]	6.14	2.28
ENSG00000105202	ENST00000221801 fibrillarin [Source:HGNC Symbol;Acc:3599]	6.14	2.28
ENSG00000167085	ENST00000446735 prohibitin [Source:HGNC Symbol;Acc:8912]	6.14	2.28
ENSG00000169100	ENST00000381401 solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 6 [Source:HGNC Symbol;Acc:10992]	6.14	2.28
ENSG00000116251	ENST00000234875 ribosomal protein L22 [Source:HGNC Symbol;Acc:10315]	6.03	3.90
ENSG00000144381	ENST00000345042 heat shock 60kDa protein 1 (chaperonin) [Source:HGNC Symbol;Acc:5261]	5.87	3.88
ENSG00000127483	ENST00000375004 heterochromatin protein 1, binding protein 3 [Source:HGNC Symbol;Acc:24973]	5.85	3.71
ENSG00000134440	ENST00000423481 asparaginyl-tRNA synthetase [Source:HGNC Symbol;Acc:7643]	5.27	2.08
ENSG00000163631	ENST00000509063 albumin [Source:HGNC Symbol;Acc:399]	5.27	2.08
ENSG00000167553	ENST00000301072 tubulin, alpha 1c [Source:HGNC Symbol;Acc:20768]	5.01	3.77
ENSG00000178952	ENST00000313511 Tu translation elongation factor, mitochondrial [Source:HGNC Symbol;Acc:12420]	4.70	3.76

Table 5.8 Consensus list of proteins identified as significantly different between total CLL cells and Annexin V negative cells from three different CLL patient samples. The full dataset is available in Supplementary Table 5.2.10.

5.3 Discussion

The aim of the work in this chapter was to establish a method for identifying the cell surface proteome and to apply this method to the analysis of cells obtained from patients with CLL. Both homotypic cell: cell interactions and interactions between cell surface proteins and extracellular ligands are crucial for the survival of CLL cells *in vitro* suggesting that cell surface proteins are important for CLL cell survival. In this chapter I established a method to enrich for cell surface proteins using a cell-surface biotinylation and identification of these proteins using a shotgun proteomics approach, which I applied to the analysis of cells from CLL patients.

Firstly, I tested a cell surface biotinylation method to isolate cell surface proteins. Initially, I used primary human T cells and the HeLa cell line to test and optimise the methodology and I then applied this method to CLL patient samples. Because of the fragile nature of these cells, experiments had to be carried out immediately on CLL cells straight from the clinic and I showed that cryopreserved CLL cells were not suitable for such analyses.

Next, during a rotation in Professor Marcotte's laboratory at the University of Texas at Austin, I tested different sample preparation methods to optimise cell surface protein identification by LC-MS/MS. This work was carried out on readily available lysates which I obtained from HeLa cells. These experiments allowed me to select and optimise the methods to process patient samples for LC-MS/MS.

A CLL cell total lysate consensus list was made so that enrichment of cell surface proteins could be tested against the representation of these proteins in a total lysate. All of the cell surface proteins identified with significant Z scores and FC>2 (biotinylated proteins bound to the NeutrAvidin column, compared with the PBS control) also had a significant Z score when compared with the CLL consensus total lysate. I then analysed biotinylated cell surface proteins by LC-MS/MS. After filtering against a paired control, 169 proteins met the inclusion criteria for the CLL cell surface dataset. 13% of these proteins are annotated as being cell surface related. When CLL proteins without a paired control were analysed by LC-MS/MS, only 2.7% were annotated as being cell surface proteins. This indicates that paired controls are important to create meaningful, smaller datasets with fewer false-positive identifications. When HeLa samples were analysed, 100 of 188 proteins identified with a significant Z score had a cell surface annotation. This suggests that the method is successful in isolating and identifying cell surface proteins and the lower enrichment of cell surface proteins observed in the CLL samples is cell type specific.

Somewhat surprisingly, a significant number of proteins were identified in samples of biotinylated CLL cell proteins which have traditionally been classified as intracellular proteins. These include

cytoplasmic (*e.g.* Hsp90), nuclear (*e.g.* GAPDH) and cytoskeletal (*e.g.* vimentin, myosin non-muscle heavy chain-9, cofilin-1, ezrin, radixin) proteins. The heat shock protein Hsp90 has been shown to be present on the cell surface of a number of cancer cells (254) and may be a therapeutic target for Hsp90 inhibitors, which are not cell-permeable (255). CLL cells are also known to be capable of presenting antigens (256) and normal B-cells can bind, internalise, process and present cell surface antigens that become exposed during apoptosis (257). As noted earlier, most CLL BCRs recognise molecular motifs associated with apoptosis. Interestingly, vimentin and cofilin-1 have all been previously identified as specific binding partners for CLL B-cell receptor stereotypes 32 and 1 respectively (258). In the case of subset 6 (Vh1.69), the CLL BCR has been shown to bind myosin heavy chain IIA (NMHC-II-A), which is exposed on the surface of a subset of apoptotic cells termed Myosin Heavy Chain IIA Exposed Apoptotic Cells (MEACS) (248). Analyses of Jurkat cells undergoing apoptosis has shown that Histone proteins also appear on the cell surface (259) and this could be the case for CLL cells in the early stages of apoptosis, but which still have an intact plasma membrane.

After observing an enrichment of proteins associated with the GO term 'regulation of apoptosis' in the cell surface fraction of CLL cells, I carried out an experiment to determine whether these proteins were identified because of the presence of a sub-set of intact CLL cells undergoing apoptosis, which were Annexin V positive. I removed Annexin V positive cells and analysed the cell surface proteome of the Annexin V negative $CD5^{+}CD19^{+}$ cells and compared the data with the cell surface proteome obtained for total $CD5^{+}CD19^{+}$ cells. David Functional Analyses of these datasets showed that removing Annexin V positive cells removed the 'regulation of apoptosis' ontology from bioinformatic analyses. Isolating $CD5^{+}CD19^{+}$ Annexin V negative cells from primary CLL PBMCs was technically very challenging due to the fragile nature of CLL cells. The multiple selection procedures required manipulation of cells through numerous washing steps, magnetic negative selection using EasySep and negative selection using Milteyni AnnexinV MicroBeads. The Milteyni selection especially was found to produce variable results with some CLL samples dying as a result of passage though the selection column. The yield of CLL cells after several rounds of selection produced small amounts of protein for analysis. Again, these observations highlight the possibility that the CLL cells that were amenable to study are a self-selecting sub-set of the disease, because these cells are robust enough to withstand the selection procedures and time spent in the laboratory away from a supportive environment. However, the data indicate that the intracellular proteins associated with 'regulation of apoptosis' are from an Annexin V positive sub-set of cells. They may be biotinylated because the plasma membrane of a proportion of these cells is compromised or because these cells express intracellular proteins on their cell surface. More experiments are now needed with larger cell

numbers to obtain a consensus list of the cell surface proteome of the cells in CLL that are in the process of undergoing apoptosis.

The role of self-antigen in CLL is contentious, with some groups reporting that the BCR is constitutively activated (260), whilst others show evidence for stimulation by antigen (261). It has been suggested that several intracellular proteins, including vimentin, cofilin 1 and non-muscle myosin heavy chain 9 (highlighted yellow in **Table 5.4**) are able to bind to some CLL BCRs. If self-antigen is able to bind to the BCR, a potential source could be the large pool of CLL cells in the patient, some of which will be undergoing apoptosis at a given time. Other cells in the tumour microenvironment may also act as sources of antigen. For example vimentin has been reported to be on the surface of viable stromal cells (262).

In summary, the study described here presents a method to isolate and identify cell surface proteins. It highlights in particular the effects of cell type and behaviour on the quality of the data produced by an experiment. The behaviour of cells during labelling and the MS data produced from HeLa cells was different from that obtained for primary CLL cells, highlighting the importance of identifying suitable controls against which to filter high content datasets.

The CLL cell surface proteome has been investigated by other groups using methods different from those described in this chapter. Kohnke *et al*. (263) used a hydrazine-coupling technique to enrich for N-linked glycoproteins combined with an iTRAQ labelling method followed by 2D LC-MS/MS to monitor Fludarabine and Cladarabine induced changes in cell surface proteins of the cell line MEC1 (CLL in prolymphocytoid transformation (264)). They identified 232 N-linked glycopeptides, 50% of which were annotated as plasma membrane proteins and 19% identified as integral membrane proteins. These are similar to the enrichment levels achieved in the HeLa datasets presented in this chapter but much higher than the CLL datasets I have described. Although their work was carried out using cell lines, their data suggests that enriching for glycosylated proteins is a viable alternative for analysing cell surface proteins, which might be applied to the analysis of primary CLL cells. A study from Boyd *et al.* (265) analysed the cell surface proteome of primary CLL cells using a sub cellular fractionation method based on density centrifugation (232) followed by 1D gel separation and MALDI-TOF analysis. This study identified 365 proteins in total and 238 (65%) were identified as membrane-associated proteins. The remaining 35% were known non-membrane proteins, reflecting contamination of the membrane preparation with cytosolic proteins. 133 (56%) of these 'membrane-associated proteins' were in some way localized to the plasma membrane, meaning that only 36% of the total proteins identified were associated with the plasma membrane. This indicates that the architecture of primary CLL cells may hinder cell surface protein enrichment when compared to the analysis of cell lines.

One of the proteins identified in the CLL cell surface enrichment analyses described in section 5.2.7, integrin beta 3 (ITGB3) has recently been identified as a therapeutic target in Mixed Lineage Leukaemia (MLL) (266). Miller *et al.* (267) used an *in vivo* shRNA screening approach to identify novel regulators required for MLL-AF9 leukaemia. They showed that leukemic cells with a reduced level of ITGB3 had impaired homing and engraftment potential in the bone marrow upon transplantation in both mouse and humanized xenograft models. The ITGB3/ITGAV heterodimer is expressed on MLL cells and mediates the interaction with ligands on the surface of stromal cells in the bone marrow niche. This demonstrates that attempts to catalogue the cell surface proteome of cancer cells could help direct future targets for therapies.

5.3.1. Other applications of the method

The cell surface protein isolation and identification method described here has now been utilised in other projects. We have used it to investigate the unknown mechanism of action of a drug, Thiabendazole (TBZ) on Human Umbilical Vein Endothelial Cells (HUVEC) (135). The effect of TBZ on endothelial cells was discovered through a 'Phenolog' approach, combined with a drug repurposing method by our collaborators, Professor Marcotte's laboratory at the University of Texas at Austin. The 'Phenolog' approach is based on the evolutionary conservation of genetic modules. For example, the function of one particular gene module in yeast is to maintain cell walls, whilst in vertebrates the same gene module regulates neovascularisation (133). Cha *et al*. (135) analysed this gene module to search for small molecules which target the yeast pathway and which could then be tested as angiogenesis inhibitors. TBZ, an antifungal drug already in clinical use was identified by this method and it was shown to inhibit angiogenesis in animal models and HUVEC cells.

I carried out an experiment to determine quantitative changes in the cell surface proteome of HUVEC cells caused by TBZ treatment using the methods described in this Chapter. Only 11 proteins met the strict filters of the data shown in **Table 5.9** (Z score > 1.96 or > 1.96 and with a FC>2). The protein identified with the largest fold change and a significant Z score was myosin phosphatase Rho interacting protein (MPRIP). MPRIP is a member of the myosin phosphatase complex that directly binds RhoA (268, 269) and targets myosin phosphatase to the actin cytoskeleton. Depletion of MPRIP leads to an increased number of stress fibres in smooth muscle cells through stabilization of actin fibres by phosphorylated myosin (270).

The identification of MPRIP fits with the expected mechanism of action of TBZ on endothelial cells, because the drug behaves as a vascular disrupting agent. TBZ has been shown to impede migration of HUVECs in a wound scratch assay and treatment with the Rho Kinase inhibitor Y27632 reverses

TBZ's effects (135). This suggests that TBZ may be acting on the Rho kinase pathway. Use of the Y27632 inhibitor elicited a significant and dose-dependent rescue of the TBZ-induced HUVEC cell motility defect. Together, these data suggest that vascular disruption by TBZ results from hyperactive Rho signalling and the data in **Table 5.9** support the conclusions of Cha *et al*. (135), that a major effect of TBZ treatment is the modulation of the Rho kinase pathway. Interestingly, MPRIP siRNA-treated HeLa cells were significantly less invasive (271). These data show that the method described in this Chapter could be applied to help uncover details of unknown mechanisms involving cell surface proteins.

Accession	Description	Z (TBZ-DMSO)	FC (TBZ/DMSO)
ENSG00000133030	myosin phosphatase Rho interacting protein [Source:HGNC Symbol;Acc:30321]	-3.39	14.33
ENSG00000110321	eukaryotic translation initiation factor 4 gamma, 2 [Source:HGNC Symbol;Acc:3297]	-2.69	9.92
ENSG00000134851	transmembrane protein 165 [Source:HGNC Symbol;Acc:30760]	-3.03	7.16
ENSG00000149480	metastasis associated 1 family, member 2 [Source:HGNC Symbol;Acc:7411]	-2.02	6.61
ENSG00000184640	septin 9 [Source:HGNC Symbol;Acc:7323]	-2.02	6.61
ENSG00000197535	myosin VA (heavy chain 12, myoxin) [Source:HGNC Symbol;Acc:7602]	-2.02	6.61
ENSG00000137497	nuclear mitotic apparatus protein 1 [Source:HGNC Symbol;Acc:8059]	-3.06	4.68
ENSG00000187555	ubiquitin specific peptidase 7 [Source:HGNC Symbol;Acc:12630]	-2.05	4.41
ENSG00000134760	desmoglein 1 [Source:HGNC Symbol;Acc:3048]	-2.40	2.42
ENSG00000105953	oxoglutarate (alpha-ketoglutarate) dehydrogenase [Source:HGNC Symbol;Acc:8124]	-2.05	2.10
ENSG00000163565	interferon, gamma-inducible protein 16 [Source:HGNC Symbol;Acc:5395]	-2.05	2.10

Table 5.9 Analyses of the HUVEC cell surface proteins identified by mass spectrometry over represented in TBZ treated compared to DMSO control. HUVEC were treated with biotin ester or DMSO control and the cells were lysed and processed as described. Proteins were selected with a Z score > 1.96 or > 1.96 and with a FC>2 and the top 11 are shown. The complete dataset is in Supplementary Table 5.3.1.

Chapter 6

General Discussion

6. General Discussion

6.1 Summary

The aim of my PhD project was to apply Systems Biology methods to answer biological questions about interactions between CLL cells and the tumour microenvironment, including: (i) the effects of co-culture model systems on primary CLL cell viability and phenotype, (ii) investigating the transcriptional effects of endothelial cell co-culture on primary CLL cells, particularly control mechanisms responsible for co-ordinated gene expression and (iii) a systematic identification of primary CLL cell surface proteins. Increasing our knowledge of CLL biology provides more potential targets for rational therapeutic intervention.

The study in Chapter 3 focused on the effects of co-culture on primary CLL cells. CLL cells undergo apoptosis when cultured in the laboratory away from the supportive tumour microenvironment found in the body. Laboratories across the world employ different methods to keep CLL cells alive in culture so that the biology of the disease can be investigated. Most of these methods seek to replicate interactions which occur in the tumour microenvironment. In Chapter 3 and recently published in the *British Journal of Haematology*, we presented a direct comparison of the effects of different co-culture systems on primary CLL cells. We evaluated the effects of interactions between primary CLL cells and endothelial cells (HMEC-1), and also with mouse fibroblasts expressing the human proteins CD40L (expressed on activated T cells) and CD31 (expressed on endothelial cells), which have been reported to exist in the tumour microenvironment. The results showed that all co-culture systems increased CLL cell viability compared with liquid culture and induced a change in CLL cell phenotype consistent with that found in poor prognosis disease *in vivo*. However, differences were observed in proliferation assays, with only the fibroblast models able to induce CLL cell proliferation in culture.

The study in Chapter 4 expanded on the co-culture analyses carried out in Chapter 3 and examined the mRNA changes which occur when cells from CLL patients are co-cultured with an endothelial cell line, HMEC-1 and the primary endothelial cells, HDBEC. Based on the observation that both co-culture systems improved CLL cell viability and induced a similar phenotype in CLL cells, I sought to determine whether there were mRNA changes which occurred in the cells from all CLL patients as a result of co-culture. The rationale behind these biofinformatic analyses was that any mRNA or pathways up-regulated in both systems may represent cellular mechanisms which the CLL cells have become addicted to and rely upon for survival and therefore present targets for intervention. 53 mRNA were identified as significantly up-regulated by both co-culture systems. Other groups have investigated the transcriptional effects of one co-culture system (65) and compared microarray-based expression profiles of CLL cells before and after three different survival-inducing culture

conditions: (i) HS-5 stromal cell co-culture, (ii) stromal cell conditioned medium and (iii) high cell density cultures of unsorted peripheral blood mononuclear cells (156). Their bioinformatic approach focussed on Ingenuity Pathways Analysis (IPA) to identify the top canonical pathways of genes common to each condition. These analyses identified Toll-like receptor signalling, NRF2-mediated oxidative response, ATM signalling, TREM1 signalling and p53 signalling all which have roles previously identified in CLL. However, the bioinformatic approach in my study has predicted the presence of a novel TF module for CLL, which may be responsible for the co-ordinated expression of the 53 mRNA up-regulated by both endothelial co-culture systems. The TF module contains CEBP-α, CEBP-β, NFκB p65, STAT1 and PU.1 and has been described to occur in macrophages. This potential control mechanism represents a novel way in which to target CLL cell survival pathways.

The final study in Chapter 5 focussed on identifying cell surface proteins on primary CLL cells which may be required for cell:cell contacts made in the tumour microenvironment and therefore may be important to receive signals which can result in the transcriptional changes observed in Chapter 4. I optimised methods to identify cell surface proteins by MS using HeLa cells, achieving enrichment of up to 70% cell surface related proteins, which was comparable with reports in the literature for other cell types. I then applied these methods to analyse the cell surface proteins isolated from primary CLL cells. The preparations were digested with trypsin, separated on a C18 reverse phase column and analysed using tandem MS. The results showed a lower than expected enrichment of cell surface proteins from primary CLL cells (17%). This may be a feature of CLL cells and bioinformatic analyses highlighted an enrichment of apoptosis related proteins in the cell surface preparations. One explanation could be that the CLL cells are in an early stage of apoptosis either in the body or as a result of isolation and selection methods before biotinylation. Therefore, I then sought to determine whether the cell surface proteins identified from Annexin V negative cells would differ significantly from those of total CLL cells using a Milteyni Annexin V depletion protocol. The data from these experiments are preliminary as only three samples could be isolated with sufficient viability for biotinylation. However, these data showed that the apoptosis-related proteins identified are not present on the Annexin V negative CLL cell population and so may come from a sub-set of dying CLL cells.

The biological plausibility of the protein identifications made in this study remains to be verified. As previously discussed, a number of the proteins identified are not usually associated with expression at the cell surface. Additionally, some proteins identified (such as those in **Table 5.5**) are abundant in platelets or in endothelial cells (for example SERPINE1 as shown in Figure 4.12). Therefore, further experiments are required to verify the expression of these proteins in CLL cells for example by western blotting.

Investigation of interactions between cancer cells and the tumour microenvironment are crucial to better understand the biology of the disease in the hope of targeting pathways critical for cancer cell survival (157). This may require targeting the cancer cell itself, or modulating the tumour microenvironment to block signalling between the tumour cell and the stroma. In the context of solid tumours, it is thought that interactions between neoplastic cells and the tumour microenvironment play an important role in metastasis and disease progression. This has been shown to occur in colorectal cancer, where TGFα-EGFR signalling creates a microenvironment that is conducive for metastasis (158), providing a rationale for attempts to inhibit EGFR signalling in TGFα-positive cancers. Another example of paracrine signalling in the tumour microenvironment is in the progression from *in situ* to invasive breast carcinoma, where loss of stromal Cav-1 expression and acquisition of MCT4 stromal expression are reported in invasive ductal carcoinomas (159).

Many laboratories are working towards model culture systems which recreate the tumour microenvironment more accurately. Multiple Myeloma (MM), like CLL is dependent on the tumour microenvironment and depends on myeloma cell-bone marrow interactions. Three-dimensional culture models are being developed to further study the biology of MM and response to therapy (160). Advanced intra-vital microscopy or 'dynamic histopathology' techniques are also providing insights into cancer progression as a dynamic step-wise process within anatomic and functional niches provided by the microenvironment (161).

6.2 Future work

The conclusions from each study are summarised in section 6.1, however the results raise additional questions which could be addressed in future projects.

6.2.1. Determination of labelling location

In order to clarify whether biotin labelling was cell surface or cytoplasmic, further bioinformatic analyses of the raw MS data are required. The biotinylation procedure modifies the lysine residue and alters its mass. Therefore, we could search the data from the MS spectra for the predicted modified mass of the labelled residue. We could then determine the precise location of the labelling modification by analysing the orientation of the modified residue in the native protein. This analysis would help to determine whether the labelling reagent was entering the cell and labelling residues usually found in the cytoplasm.

6.2.2. Cellular localisation of proteins identified by mass spectrometry

The proteomics data of the cell surface protein preparations presented in Chapter 5 suggest that numerous proteins whose typical biological function is mainly characterised in the cytoplasm, nucleus or other organelles within the cell occur at the cell surface in CLL. It would be of interest to investigate whether these proteins or fragments of these proteins exist at the cell surface in CLL cells. This could be achieved experimentally by immunostaining the proteins without permeabilising the cells. Comparisons could be made between CLL cells from peripheral blood and lymph node to determine whether the cell surface proteome differs between the CLL compartments. Co-localisation with other known cell surface proteins could be investigated by multi-colour confocal microscopy. It would also be of interest to determine whether such proteins are only localised at the cell surface of a sub-set of cells, such as those undergoing apoptosis. A previous study showed that Histone proteins are present on the cell surface of (Jurkat) cells dying by apoptosis (259) and such a mechanism may account for the cell surface expression of intracellular proteins in samples from patients with CLL. The chaperone proteins, Hsp90 have also been identified on the cell surface of other cancers including prostate cancer (254), melanoma (272) and secreted in exosomes (273, 274). Cell surface Hsp90 modulates prostate cancer cell adhesion and invasion through the integrin-β1/focal adhesion kinase/c-Src signalling pathway (275). Studies such as these have led to the identification of cell surface Hsp90 as a therapeutic target. Cell impermeable Hsp90 inhibitors are active functionally and inhibit tumour cell invasion and metastasis (255, 276). It would be interesting to investigate whether treatment with a cell impermeable Hsp90 inhibitor would kill CLL cells. Cell surface proteome profiling in other cancers including SH-SY5Y neuroblastoma, A549 lung adenocarcinoma, LoVo colon adenocarcinoma, and the Sup-B15 acute lymphoblastic leukaemia (B cell) cell lines and ovarian patient tumour cells have identified an abundance of proteins with chaperone function including GRP78, GRP75, HSP70, HSP60, HSP54, HSP27, and protein disulfide isomerase (277). This study provides further evidence to suggest that inhibitors of other chaperones may have therapeutic benefits in cancer.

6.2.3. Functional characterisation of cell surface proteins

If the proteins identified in my study are present on the cell surface in CLL, it would be of interest to investigate whether these proteins have a functional role. It would be interesting to determine for example, whether these proteins are simply products from nearby dying cells, or from the subset of CLL cells undergoing apoptosis and whether they then bind non-specifically to the cell surface or specifically to other cell surface proteins such as the BCR. Data from other laboratories suggest that some of the proteins identified in this study are capable of binding to certain classes of BCR (248)

and that auto-antigenic targets of BCRs are able to induce proliferation of CLL cells (253). Potential auto-antigens from the MS dataset could be expressed in the laboratory and presented to CLL cells to determine whether they play a functional role in BCR activation. BCR activation could be measured by western blotting for activated, phosphorylated proteins which are part of the BCR signalling pathway and CLL proliferation could be measured by BrdU incorporation into DNA. Identifying cell surface proteins which have functional roles in CLL could lead to future therapies in CLL. For example, therapies have been developed which use an antibody targeting HER2 which is over expressed at the cell surface in breast cancers (278) and other therapies target the extracellular ligand binding region of epidermal growth factor receptor (279) which is expressed in many cancer types including lung, colon and breast.

6.2.4. Identifying cell surface proteins from Annexin V negative CLL cells

The proteomics data presented in Chapter 5 suggest that the CLL cell surface preparation contained proteins with a signature related to apoptosis and regulation of apoptosis. Chu *et al*. showed that when undergoing apoptosis, cells are capable of presenting proteins on the cell surface which are usually intracellular, including non muscle myosin heavy chain IIA (MYHIIA) (248). Their study also demonstrated that CLL antibodies recognise apoptotic cells with exposed MYHIIA, potentially providing a role for apoptotic cell surface proteins in CLL. An example of apoptotic cells also expressing Histone proteins on their surface is described above (259). In Chapter 5, I isolated Annexin V negative CLL cells from total $CD5^{+}CD19^{+}$ CLL cells using a Milteyni Annexin V depletion kit. However, because CLL cells are extremely fragile, this meant that the protocol was not reliably reproducible. Cell sorting by flow cytometry to select Annexin V negative CLL cells for cell surface biotinylation may provide samples with greater viability and reproducibility for a larger number of CLL patient samples.

6.2.5. Identifying a novel TF module in CLL by ChIP-Sequencing

The bioinformatic analyses presented in Chapter 4 predict a novel TF module in CLL. GSCA (195) predicts that this TF module is responsible for the co-ordinated expression of genes up-regulated in CLL cells by both endothelial co-culture systems. Chromatin immunoprecipitation (ChIP)-sequencing could be performed to determine whether any of these TF binding sites are occupied in primary CLL cells. The methods and ChIP-grade antibodies have been described in other ChIP-seq studies and could be applied directly to analysing CLL cells. If TF occupancy is confirmed, the effects of reducing the expression of these TF in CLL cells could be investigated. Knockdown of each TF individually and in combinations could determine whether the expression changes observed required the

presence of the entire module or if a sub-set of the TFs are capable of inducing the expression of some or all the genes encoding the mRNA which I showed are regulated. Those involved in protein interaction sub-networks that are known to be involved in apoptosis, cell proliferation or migration could be predicted by bioinformatics algorithms, such as HumanNet (132), which was used in a recent study by our laboratory (155). This is important for identifying the best therapeutic target or targets which would circumvent the cytoprotection provided by the microenvironment.

References

7. References

1. Buggins AG, Pepper C, Patten PE, Hewamana S, Gohil S, et al. 2010. *Cancer Res* 70: 7523-33
2. Deaglio S, Malavasi F. 2009. *Haematologica* 94: 752-6
3. Deaglio S, Vaisitti T, Billington R, Bergui L, Omede P, et al. 2007. *Blood* 109: 5390-8
4. Dighiero G. 2008. *N Engl J Med* 359: 638-40
5. Jaffe E, Harris N, Stein H, Vardiman J. 2001. *Pathology and Genetics of Tumours of The Haematopoietic and Lymphoid Systems*. Lyon: IARC Press
6. Oscier D, Fegan C, Hillmen P, Illidge T, Johnson S, et al. 2004. *Br J Haematol* 125: 294-317
7. Caligaris-Cappio F, Hamblin TJ. 1999. *J Clin Oncol* 17: 399-408
8. Muller-Hermelink HK, Zettl A, Pfeifer W, Ott G. 2001. *Histopathology* 38: 285-306
9. Chiorazzi N, Hatzi K, Albesiano E. 2005. *Ann N Y Acad Sci* 1062: 1-12
10. Molica S, Alberti A. 1987. *Cancer* 60: 2712-6
11. Cantwell M, Hua T, Pappas J, Kipps TJ. 1997. *Nat Med* 3: 984-9
12. Gorgun G, Holderried TA, Zahrieh D, Neuberg D, Gribben JG. 2005. *J Clin Invest* 115: 1797-805
13. Buggins AG, Patten PE, Richards J, Thomas NS, Mufti GJ, Devereux S. 2008. *Leukemia* 22: 1084-7
14. Greene MH, Hoover RN, Fraumeni JF, Jr. 1978. *J Natl Cancer Inst* 61: 337-40
15. Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, et al. 2000. *Hematol J* 1: 53-66
16. Dohner H, Stilgenbauer S, Benner A, Leupolt E, Krober A, et al. 2000. *N Engl J Med* 343: 1910-6.
17. Dohner H, Fischer K, Bentz M, Hansen K, Benner A, et al. 1995. *Blood* 85: 1580-9
18. Grever MR, Lucas DM, Johnson AJ, Byrd JC. 2007. *Best practice & research Clinical haematology* 20: 545-56
19. Shanafelt TD, Jelinek D, Tschumper R, Schwager S, Nowakowski G, et al. 2006. *J Clin Oncol* 24: 3218-9; author reply 9-20
20. Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, Pasternack BS. 1975. *Blood* 46: 219-34
21. Hallek M, Group GCS. 2008. *Ann Oncol* 19 Suppl 4: iv51-3
22. Furman RR. 2010. *Hematology Am Soc Hematol Educ Program* 2010: 77-81
23. Klein U, Tu Y, Stolovitzky GA, Mattioli M, Cattoretti G, et al. 2001. *J Exp Med* 194: 1625-38
24. Oscier DG, Gardiner AC, Mould SJ, Glide S, Davis ZA, et al. 2002. *Blood* 100: 1177-84
25. Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, et al. 1999. *Blood* 94: 1840-7
26. Thorselius M, Krober A, Murray F, Thunberg U, Tobin G, et al. 2006. *Blood* 107: 2889-94
27. Patten P, Buggins A, Richards J, Mufti G, Devereux S. 2005. *Brit. J Haematol* 129: 79a
28. Deaglio S, Vaisitti T, Bergui L, Bonello L, Horenstein AL, et al. 2005. *Blood* 105: 3042-50
29. Stilgenbauer S, Bullinger L, Lichter P, Dohner H. 2002. *Leukemia* 16: 993-1007
30. Schroers R, Griesinger F, Trumper L, Haase D, Kulle B, et al. 2005. *Leukemia* 19: 750-8
31. Montserrat E, Sanchez-Bisono J, Vinolas N, Rozman C. 1986. *Br J Haematol* 62: 567-75
32. Puente XS, Pinyol M, Quesada V, Conde L, Ordonez GR, et al. 2011. *Nature* 475: 101-5
33. Rosati E, Sabatini R, Rampino G, Tabilio A, Di Ianni M, et al. 2009. *Blood* 113: 856-65
34. Nwabo Kamdje AH, Bassi G, Pacelli L, Malpeli G, Amati E, et al. 2012. *Blood Cancer J* 2: e73
35. Messmer BT, Messmer D, Allen SL, Kolitz JE, Kudalkar P, et al. 2005. *J Clin Invest* 115: 755-64

36. Lin TT, Letsolo BT, Jones RE, Rowson J, Pratt G, et al. 2010. *Blood* 116: 1899-907
37. Jahrsdorfer B, Weiner GJ. 2008. *Blood* 111: 5756-
38. Chiorazzi N, Rai KR, Ferrarini M. 2005. *N Engl J Med* 352: 804-15
39. Damle RN, Temburni S, Calissano C, Yancopoulos S, Banapour T, et al. 2007. *Blood* 110: 3352-9
40. Gottardi D, Alfarano A, De Leo AM, Stacchini A, Bergui L, Caligaris-Cappio F. 1995. *Curr Top Microbiol Immunol* 194: 307-12
41. Deaglio S, Vaisitti T, Aydin S, Bergui L, D'Arena G, et al. 2007. *Blood* 110: 4012-21
42. Hartmann TN, Grabovsky V, Wang W, Desch P, Rubenzer G, et al. 2009. *Cancer Res* 69: 3121-30
43. Ramsay AG, Evans R, Kiaii S, Svensson L, Hogg N, Gribben JG. 2013. *Blood* 121: 2704-14
44. Till KJ, Lin K, Zuzel M, Cawley JC. 2002. *Blood* 99: 2977-84
45. Shanafelt TD, Geyer SM, Bone ND, Tschumper RC, Witzig TE, et al. 2008. *Br J Haematol* 140: 537-46
46. Zucchetto A, Benedetti D, Tripodo C, Bomben R, Dal Bo M, et al. 2009. *Cancer Res* 69: 4001-9
47. Deaglio S, Aydin S, Grand MM, Vaisitti T, Bergui L, et al. 2010. *Mol Med* 16: 87-91
48. Chiorazzi N. 2007. *Best Pract Res Clin Haematol* 20: 399-413
49. Deaglio S, Vaisitti T, Aydin S, Bergui L, D'Arena G, et al. 2007. *Blood* Prepublished online August 15th 2007;DOI 10.1182/blood-2007-06-094029
50. Ghia P, Caligaris-Cappio F. 2000. *Adv Cancer Res* 79: 157-73
51. Bishop GA, Hostager BS. 2001. *Curr Opin Immunol* 13: 278-85
52. Patten PE, Buggins AG, Richards J, Wotherspoon A, Salisbury J, et al. 2008. *Blood* 111: 5173-81
53. Messmer B, Messmer D, Allen S, Kolitz J, Kudalkar P, et al. 2005. *J Clin Invest* 115: 755-64
54. Collins RJ, Verschuer LA, Harmon BV, Prentice RL, Pope JH, Kerr JF. 1989. *Br J Haematol* 71: 343-50.
55. Herishanu Y, Perez-Galan P, Liu D, Biancotto A, Pittaluga S, et al. 2011. *Blood* 117: 563-74
56. Patten PEM, Buggins AGS, Richards J, Wotherspoon A, Salisbury J, et al. 2008. *Blood* 111: 5173-81
57. Plander M, Seegers S, Ugocsai P, Diermeier-Daucher S, Ivanyi J, et al. 2009. *Leukemia* 23: 2118-28
58. Schmid C, Isaacson P. 1994. *Histopathology* 24: 445-51
59. Scielzo C, Ghia P, Conti A, Bachi A, Guida G, et al. 2005. *J Clin Invest* 115: 1644-50
60. Deaglio S, Vaisitti T, Bergui L, Bonello L, Horenstein A, et al. 2005. *Blood* 105: 3042-50
61. Buggins AG, PhD, Patten PE, BSc, MRCP, et al. 2008. *Blood* 112: 357
62. Majid A, Lin TT, Best G, Fishlock K, Hewamana S, et al. 2011. *Leuk Res* 35: 750-6
63. Pettitt AR, Moran EC, Cawley JC. 2001. *Leuk Res* 25: 1003-12.
64. Burgess M, Cheung C, Chambers L, Ravindranath K, Minhas G, et al. 2012. *Leuk Lymphoma* 53: 1988-98
65. Maffei R, Fiorcari S, Bulgarelli J, Martinelli S, Castelli I, et al. 2012. *Haematologica* 97: 952-60
66. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. 1999. *Blood* 94: 1848-54
67. Stamatopoulos K, Belessi C, Moreno C, Boudjograh M, Guida G, et al. 2007. *Blood* 109: 259-70

68. Yan X, Albesiano E, Zanesi N, Yancopoulos S, Sawyer A, et al. 2006. *Proc Natl Acad Sci U S A* 103: 11713-8
69. Lanemo Myhrinder A, Hellqvist E, Sidorova E, Söderberg A, Baxendale H, et al. 2008. *Blood* 111: 3838-48
70. Muzio M, Scielzo C, Bertilaccio MT, Frenquelli M, Ghia P, Caligaris-Cappio F. 2009. *Br J Haematol* 144: 507-16
71. Krysov S, Potter KN, Mockridge CI, Coelho V, Wheatley I, et al. 2010. *Blood* 115: 4198-205
72. Stevenson FK, Krysov S, Davies AJ, Steele AJ, Packham G. 2011. *Blood* 118: 4313-20
73. Burger JA, Ghia P, Rosenwald A, Caligaris-Cappio F. 2009. *Blood* 114: 3367-75
74. Rosati E, Sabatini R, Rampino G, Tabilio A, Di Ianni M, et al. 2009. *Blood* 113: 856-65
75. Paterson A, Mockridge CI, Adams JE, Krysov S, Potter KN, et al. 2012. *Blood* 119: 1726-36
76. Hill RJ, Lou Y, Tan SL. 2013. *International Reviews of Immunology* 32: 377-96
77. Burthem J, Vincent A, Cawley JC. 1996. *Leuk Lymphoma* 21: 211-5
78. Krishnaswamy G, Kelley J, Yerra L, Smith JK, Chi DS. 1999. *J Interferon Cytokine Res* 19: 91-104
79. Mantovani A, Garlanda C, Introna M, Vecchi A. 1998. *Transplant Proc* 30: 4239-43
80. Peters K, Unger RE, Brunner J, Kirkpatrick CJ. 2003. *Cardiovasc Res* 60: 49-57
81. Lidington EA, Moyes DL, McCormack AM, Rose ML. 1999. *Transplant Immunology* 7: 239-46
82. Viemann D, Goebeler M, Schmid S, Nordhues U, Klimmek K, et al. 2006. *J Leukoc Biol* 80: 174-85
83. Foa R, Massaia M, Cardona S, Tos AG, Bianchi A, et al. 1990. *Blood* 76: 393-400
84. Deaglio S, Mallone R, Baj G, Arnulfo A, Surico N, et al. 2000. *Chem Immunol* 75: 99-120
85. Ades E, Candal F, Swerlick R, George V, Summers S, et al. 1992. *J Invest Dermatol* 99: 683-90
86. Vincent AM, Burthem J, Brew R, Cawley JC. 1996. *Blood* 88: 3945-52
87. de Haart SJ, van de Donk NW, Minnema MC, Huang JH, Aarts-Riemens T, et al. 2013. *Clin Cancer Res*
88. Plander M, Ugocsai P, Seegers S, Orso E, Reichle A, et al. 2011. *Ann Hematol* 90: 1381-90
89. Brachtl G, Sahakyan K, Denk U, Girbl T, Alinger B, et al. 2011. *PLoS One* 6
90. NCI Dictionary of Cancer Terms. pp. http://www.cancer.gov/dictionary?cdrid=270742
91. Zhou GB, Zhang J, Wang ZY, Chen SJ, Chen Z. 2007. *Philos Trans R Soc Lond B Biol Sci* 362: 959-71
92. Stegmeier F, Warmuth M, Sellers WR, Dorsch M. 2010. *Clin Pharmacol Ther* 87: 543-52
93. Mokbel K, Hassanally D. 2001. *Curr Med Res Opin* 17: 51-9
94. Cuthill K, Devereux S. 2013. *Br J Haematol*
95. Byrd JC, Furman RR, Coutre SE, Flinn IW, Burger JA, et al. 2013. *N Engl J Med* 369: 32-42
96. Reed JC, Kitada S, Takayama S, Miyashita T. 1994. *Ann Oncol* 5 Suppl 1: 61-5
97. Wilson WH, O'Connor OA, Czuczman MS, LaCasce AS, Gerecitano JF, et al. 2010. *Lancet Oncol* 11: 1149-59
98. Roberts AW, Seymour JF, Brown JR, Wierda WG, Kipps TJ, et al. 2012. *J Clin Oncol* 30: 488-96
99. Abboudi Z, Patel K, Naresh KN. 2009. *Eur J Haematol* 83: 203-7
100. Igawa T, Sato Y, Takata K, Fushimi S, Tamura M, et al. 2011. *Cancer Sci* 102: 2103-7
101. Blachly JS, Byrd JC. 2013. *Leuk Lymphoma*
102. Byrd JC, Shinn C, Waselenko JK, Fuchs EJ, Lehman TA, et al. 1998. *Blood* 92: 3804-16

103. Robak T, Robak P. 2013. *Int Rev Immunol* 32: 358-76
104. Davids MS, Brown JR. 2012. *Leukemia & Lymphoma* 53: 2362-70
105. Barrientos J, Rai K. 2013. *Leuk Lymphoma* 54: 1817-20
106. Castillo JJ, Furman M, Winer ES. 2012. *Expert Opin Investig Drugs* 21: 15-22
107. Hoellenriegel J, Meadows SA, Sivina M, Wierda WG, Kantarjian H, et al. 2011. *Blood* 118: 3603-12
108. Hoellenriegel J, Coffey GP, Sinha U, Pandey A, Sivina M, et al. 2012. *Leukemia* 26: 1576-83
109. Suljagic M, Longo PG, Bennardo S, Perlas E, Leone G, et al. 2010. *Blood* 116: 4894-905
110. Friedberg JW, Sharman J, Sweetenham J, Johnston PB, Vose JM, et al. 2010. *Blood* 115: 2578-85
111. Wang JC, Kafeel MI, Avezbakiyev B, Chen C, Sun Y, et al. 2011. *Oncology* 81: 325-9
112. Bokelmann I, Mahlknecht U. 2008. *Mol Med* 14: 20-7
113. Byrd JC, Shinn C, Ravi R, Willis CR, Waselenko JK, et al. 1999. *Blood* 94: 1401-8
114. Bodo J, Zhao X, Sharma A, Hill BT, Portell CA, et al. 2013. *Br J Haematol*
115. Lin K, Rockliffe N, Johnson GG, Sherrington PD, Pettitt AR. 2008. *Oncogene* 27: 2445-55
116. Dempsey NC, Leoni F, Ireland HE, Hoyle C, Williams JH. 2010. *J Leukoc Biol* 87: 467-76
117. Walsby E, Pearce L, Burnett AK, Fegan C, Pepper C. 2012. *Oncotarget* 3: 525-34
118. Best OG, Mulligan SP. 2012. *Leuk Lymphoma* 53: 2314-20
119. Gupta SV, Hertlein E, Lu Y, Sass EJ, Lapalombella R, et al. 2013. *Clin Cancer Res* 19: 2406-19
120. Riches JC, Davies JK, McClanahan F, Fatah R, Iqbal S, et al. 2013. *Blood* 121: 1612-21
121. Noble D. 2006. *The Music of Life: Biology beyond the Genome*. Oxford: OUP
122. Nurse P, Hayles J. 2011. *Cell* 144: 850-4
123. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, et al. 2002. *Science* 298: 799-804
124. Droste P, Miebach S, Niedenfuhr S, Wiechert W, Noh K. 2011. *Biosystems* 105: 154-61
125. Hornberg JJ, Bruggeman FJ, Westerhoff HV, Lankelma J. 2006. *Biosystems* 83: 81-90
126. Hood L, Heath JR, Phelps ME, Lin B. 2004. *Science* 306: 640-3
127. Weston AD, Hood L. 2004. *J Proteome Res* 3: 179-96
128. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. 2004. *Genome Biol* 5: R80
129. Futschik ME, Chaurasia G, Herzel H. 2007. *Bioinformatics* 23: 605-11
130. Brown KR, Jurisica I. 2005. *Bioinformatics* 21: 2076-82
131. Persico M, Ceol A, Gavrila C, Hoffmann R, Florio A, Cesareni G. 2005. *BMC Bioinformatics* 6 Suppl 4: S21
132. Lee I, Blom UM, Wang PI, Shim JE, Marcotte EM. 2011. *Genome Res* 21: 1109-21
133. McGary KL, Park TJ, Woods JO, Cha HJ, Wallingford JB, Marcotte EM. 2010. *Proc Natl Acad Sci U S A* 107: 6544-9
134. Fitch WM. 1970. *Syst Zool* 19: 99-113
135. Cha HJ, Byrom M, Mead PE, Ellington AD, Wallingford JB, Marcotte EM. 2012. *PLoS Biol* 10: e1001379
136. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. 2000. *Nat Genet* 25: 25-9
137. von Heijne G. 1992. *J Mol Biol* 225: 487-94
138. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. *J Mol Biol* 305: 567-80
139. Williams SL, Nesbeth D, Darling DC, Farzaneh F, Slater NK. 2005. *J Chromatogr B Analyt Technol Biomed Life Sci* 820: 111-9
140. Elia G. 2008. *Proteomics* 8: 4012-24
141. Steen H, Mann M. 2004. *Nat Rev Mol Cell Biol* 5: 699-711

142. Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Graham Cooks R. 2005. *J Mass Spectrom* 40: 430-43
143. Olsen JV, Schwartz JC, Griep-Raming J, Nielsen ML, Damoc E, et al. 2009. *Mol Cell Proteomics* 8: 2759-69
144. Perry RH, Cooks RG, Noll RJ. 2008. *Mass Spectrom Rev* 27: 661-99
145. Scigelova M, Makarov A. 2006. *Proteomics* 6 Suppl 2: 16-21
146. Cottrell JS. 1994. *Pept Res* 7: 115-24
147. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. 1999. *Electrophoresis* 20: 3551-67
148. Loo JA, Loo RR, Udseth HR, Edmonds CG, Smith RD. 1991. *Rapid Commun Mass Spectrom* 5: 101-5
149. Nakamura T, Oda Y. 2007. *Biotechnol Genet Eng Rev* 24: 147-63
150. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, et al. 2002. *Mol Cell Proteomics* 1: 376-86
151. Angel TE, Aryal UK, Hengel SM, Baker ES, Kelly RT, et al. 2012. *Chem Soc Rev* 41: 3912-28
152. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, et al. 2004. *Mol Cell Proteomics* 3: 1154-69
153. Griffiths SD, Burthem J, Unwin RD, Holyoake TL, Melo JV, et al. 2007. *Mol Biotechnol* 36: 81-9
154. Lu P, Vogel C, Wang R, Yao X, Marcotte EM. 2007. *Nat Biotechnol* 25: 117-24
155. Orr SJ, Gaymes T, Ladon D, Chronis C, Czepulkowski B, et al. 2010. *Oncogene* 29: 3803-14
156. Schulz A, Toedt G, Zenz T, Stilgenbauer S, Lichter P, Seiffert M. 2011. *Haematologica* 96: 408-16
157. Fang H, Declerck YA. 2013. *Cancer Res* 73: 4965-77
158. Sasaki T, Nakamura T, Rebhun RB, Cheng H, Hale KS, et al. 2008. *Am J Pathol* 173: 205-16
159. Martins D, Beca FF, Sousa B, Baltazar F, Paredes J, Schmitt F. 2013. *Cell Cycle* 12: 2684-90
160. Ferrarini M, Steimberg N, Ponzoni M, Belloni D, Berenzi A, et al. 2013. *PLoS One* 8: e71613
161. Alexander S, Weigelin B, Winkler F, Friedl P. 2013. *Curr Opin Cell Biol* 25: 659-71
162. Scherer WF, Syverton JT, Gey GO. 1953. *J Exp Med* 97: 695-710
163. Lea NC, Orr, S. J., Stoeber, K., Williams, G. H., Lam, E. W. -F., Ibrahim, M. A. A., Mufti, G. J. and Thomas, N. S. B. 2003. *Mol Cell Biol* 23: 2351-61
164. Russell Sa, ed. 2001. *Molecular Cloning: A Laboratory Manual* Cold Spring Harbor Laboratory Press
165. Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH. 1990. *Proc Natl Acad Sci U S A* 87: 1663-7
166. Fecteau JF, Messmer D, Zhang S, Cui B, Chen L, Kipps TJ. 2013. *Blood* 121: 971-4
167. Grdisa M. 2003. *Leuk Res* 27: 951-6
168. Kurtova AV, Balakrishnan K, Chen R, Ding W, Schnabl S, et al. 2009. *Blood*: 1-36
169. Burger JA, Tsukada N, Burger M, Zvaifler NJ, Dell'Aquila M, Kipps TJ. 2000. *Blood* 96: 2655-63
170. Maffei R, Fiorcari S, Bulgarelli J, Martinelli S, Castelli I, et al. 2012. *Haematologica-the Hematology Journal* 97: 952-60
171. Patten P, Devereux S, Buggins A, Bonyhadi M, Frohlich M, Berenson RJ. 2005. *J Immunol* 174: 6562-3; author reply 3
172. Zhang W, Trachootham D, Liu JY, Chen G, Pelicano H, et al. 2012. *Nature Cell Biology* 14: 276-+
173. Hamilton E, Pearce L, Morgan L, Robinson S, Ware V, et al. 2012. *Br J Haematol* 158: 589-99

174. Calissano C, Damle RN, Hayes G, Murphy EJ, Hellerstein MK, et al. 2009. *Blood*: 1-46
175. EJ Walsby AB, E Hamilton, S Devereux, P Brennan, C Fegan, CJ Pepper. 2013. *British Society of Haematology, Annual Scientific Meeting*
176. Kurtova AV, Balakrishnan K, Chen R, Ding W, Schnabl S, et al. 2009. *Blood* 114: 4441-50
177. Pettitt AR, Moran EC, Cawley JC. 2001. *Leuk Res* 25: 1003-12
178. Strand C, Enell J, Hedenfalk I, Ferno M. 2007. *BMC Mol Biol* 8: 38
179. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, et al. 2006. *BMC Mol Biol* 7: 3
180. Affymetrix.
181. BiomaRt.
182. Durinck S, Spellman PT, Birney E, Huber W. 2009. *Nat Protoc* 4: 1184-91
183. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, et al. 2005. *Bioinformatics* 21: 3439-40
184. Gibbons FD, Roth FP. 2002. *Genome Res* 12: 1574-81
185. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. 2003. *Nucleic Acids Res* 31: e15
186. Malkin D, Li FP, Strong LC, Fraumeni JF, Jr., Nelson CE, et al. 1990. *Science* 250: 1233-8
187. Patterson TA, Lobenhofer EK, Fulmer-Smentek SB, Collins PJ, Chu TM, et al. 2006. *Nat Biotechnol* 24: 1140-50
188. Peart MJ, Smyth GK, van Laar RK, Bowtell DD, Richon VM, et al. 2005. *Proc Natl Acad Sci U S A* 102: 3697-702
189. Raouf A, Zhao Y, To K, Stingl J, Delaney A, et al. 2008. *Cell Stem Cell* 3: 109-18
190. Huggins CE, Domenighetti AA, Ritchie ME, Khalil N, Favaloro JM, et al. 2008. *J Mol Cell Cardiol* 44: 270-80
191. Dalman MR, Deeter A, Nimishakavi G, Duan ZH. 2012. *BMC Bioinformatics* 13 Suppl 2: S11
192. Smyth GK. 2004. *Stat Appl Genet Mol Biol* 3: Article3
193. Krizkova S, Fabrik I, Adam V, Hrabeta J, Eckschlager T, Kizek R. 2009. *Bratisl Lek Listy* 110: 93-7
194. Babula P, Masarik M, Adam V, Eckschlager T, Stiborova M, et al. 2012. *Metallomics* 4: 739-50
195. Joshi A, Hannah R, Diamanti E, Gottgens B. 2013. *Exp Hematol* 41: 354-66 e14
196. Barish GD, Yu RT, Karunasiri M, Ocampo CB, Dixon J, et al. 2010. *Genes Dev* 24: 2760-5
197. Ghisletti S, Barozzi I, Mietton F, Polletti S, De Santa F, et al. 2010. *Immunity* 32: 317-28
198. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, et al. 2010. *Mol Cell* 38: 576-89
199. Fisher RP. 2012. *Genes Cancer* 3: 731-8
200. Huang da W, Sherman BT, Lempicki RA. 2009. *Nat Protoc* 4: 44-57
201. Huang da W, Sherman BT, Lempicki RA. 2009. *Nucleic Acids Res* 37: 1-13
202. Gabay C. 2006. *Arthritis Res Ther* 8 Suppl 2: S3
203. De Filippo K, Dudeck A, Hasenberg M, Nye E, van Rooijen N, et al. 2013. *Blood* 121: 4930-7
204. Aso Y. 2007. *Front Biosci* 12: 2957-66
205. Binder BR, Christ G, Gruber F, Grubic N, Hufnagl P, et al. 2002. *News Physiol Sci* 17: 56-61
206. Lee E, Vaughan DE, Parikh SH, Grodzinsky AJ, Libby P, et al. 1996. *Circ Res* 78: 44-9
207. Robertson LE, Plunkett W. 1993. *Leuk Lymphoma* 11 Suppl 2: 71-4
208. von Eyss B, Eilers M. 2011. *Genes Dev* 25: 895-7
209. Shaheen M, Allen C, Nickoloff JA, Hromas R. 2011. *Blood* 117: 6074-82
210. Sharma SV, Settleman J. 2007. *Genes Dev* 21: 3214-31
211. Wiestner A. 2012. *Blood* 120: 4684-91

212. Efremov DG, Wiestner A, Laurenti L. 2012. *Mediterr J Hematol Infect Dis* 4: e2012067
213. Brown JR. 2013. *Curr Hematol Malig Rep* 8: 1-6
214. Winer ES, Ingham RR, Castillo JJ. 2012. *Expert Opin Investig Drugs* 21: 355-61
215. Maniatis T, Falvo JV, Kim TH, Kim TK, Lin CH, et al. 1998. *Cold Spring Harb Symp Quant Biol* 63: 609-20
216. Panne D. 2008. *Curr Opin Struct Biol* 18: 236-42
217. Hewamana S, Alghazal S, Lin TT, Clement M, Jenkins C, et al. 2008. *Blood* 111: 4681-9
218. Martinez-Lostao L, Briones J, Forne I, Martinez-Gallo M, Ferrer B, et al. 2005. *Leuk Lymphoma* 46: 435-42
219. Frank DA, Mahajan S, Ritz J. 1997. *J Clin Invest* 100: 3140-8
220. Mankai A, Bordron A, Renaudineau Y, Martins-Carvalho C, Takahashi S, et al. 2008. *Cancer Res* 68: 7512-9
221. Savli H, Sunnetci D, Cine N, Gluzman DF, Zavelevich MP, et al. 2012. *Exp Oncol* 34: 57-63
222. Shao H, Xi L, Raffeld M, Feldman AL, Ketterling RP, et al. 2011. *Mod Pathol* 24: 1421-32
223. Geissmann F, Manz MG, Jung S, Sieweke MH, Merad M, Ley K. 2010. *Science* 327: 656-61
224. Feldman AL, Arber DA, Pittaluga S, Martinez A, Burke JS, et al. 2008. *Blood* 111: 5433-9
225. Zhang D, McGuirk J, Ganguly S, Persons DL. 2009. *Int J Hematol* 89: 529-32
226. Wang E, Hutchinson CB, Huang Q, Sebastian S, Rehder C, et al. 2010. *Leuk Lymphoma* 51: 802-12
227. Wallin E, von Heijne G. 1998. *Protein Sci* 7: 1029-38
228. Hopkins AL, Groom CR. 2002. *Nat Rev Drug Discov* 1: 727-30
229. Seddon AM, Curnow P, Booth PJ. 2004. *Biochim Biophys Acta* 1666: 105-17
230. Bordier C. 1981. *J Biol Chem* 256: 1604-7
231. Macdonald JL, Pike LJ. 2005. *J Lipid Res* 46: 1061-7
232. Pasquali C, Fialka I, Huber LA. 1999. *J Chromatogr B Biomed Sci Appl* 722: 89-102
233. Zhao Y, Zhang W, Kho Y. 2004. *Anal Chem* 76: 1817-23
234. Goshe MB, Blonder J, Smith RD. 2003. *J Proteome Res* 2: 153-61
235. Zhang S, Kipps TJ. 2013. *Annu Rev Pathol*
236. Lea NC, Orr SJ, Stoeber K, Williams GH, Lam EW, et al. 2003. *Mol Cell Biol* 23: 2351-61
237. Chen EI, Cociorva D, Norris JL, Yates JR, 3rd. 2007. *J Proteome Res* 6: 2529-38
238. Kall L, Canterbury JD, Weston J, Noble WS, MacCoss MJ. 2007. *Nat Methods* 4: 923-5
239. Savani RC, Cao G, Pooler PM, Zaman A, Zhou Z, DeLisser HM. 2001. *J Biol Chem* 276: 36770-8
240. Okamoto I, Kawano Y, Murakami D, Sasayama T, Araki N, et al. 2001. *J Cell Biol* 155: 755-62
241. Overington JP, Al-Lazikani B, Hopkins AL. 2006. *Nat Rev Drug Discov* 5: 993-6
242. Pedersen SK, Harry JL, Sebastian L, Baker J, Traini MD, et al. 2003. *J Proteome Res* 2: 303-11
243. dos Remedios CG, Chhabra D, Kekic M, Dedova IV, Tsubakihara M, et al. 2003. *Physiol Rev* 83: 433-73
244. Karhemo PR, Ravela S, Laakso M, Ritamo I, Tatti O, et al. 2012. *J Proteomics* 77: 87-100
245. Podolski JL, Steck TL. 1988. *J Biol Chem* 263: 638-45
246. Hitchcock SE, Carisson L, Lindberg U. 1976. *Cell* 7: 531-42
247. Weber A, Pennise CR, Pring M. 1994. *Biochemistry* 33: 4780-6
248. Chu CC, Catera R, Zhang L, Didier S, Agagnina BM, et al. 2010. *Blood* 115: 3907-15
249. Yonemura S, Hirao M, Doi Y, Takahashi N, Kondo T, Tsukita S. 1998. *J Cell Biol* 140: 885-95
250. Hanada M, Delia D, Aiello A, Stadtmauer E, Reed JC. 1993. *Blood* 82: 1820-8
251. Lu CD, Altieri DC, Tanigawa N. 1998. *Cancer Res* 58: 1808-12

252. Kitada S, Andersen J, Akar S, Zapata JM, Takayama S, et al. 1998. *Blood* 91: 3379-89
253. Zwick C, Fadle N, Regitz E, Kemele M, Stilgenbauer S, et al. 2013. *Blood* 121: 4708-17
254. Sidera K, Patsavoudi E. 2008. *Cell Cycle* 7: 1564-8
255. Tsutsumi S, Scroggins B, Koga F, Lee MJ, Trepel J, et al. 2008. *Oncogene* 27: 2478-87
256. Hall AM, Vickers MA, McLeod E, Barker RN. 2005. *Blood* 105: 2007-15
257. Ciechomska M, Lennard TW, Kirby JA, Knight AM. 2011. *Eur J Immunol* 41: 1850-61
258. Lanemo Myhrinder A, Hellqvist E, Sidorova E, Soderberg A, Baxendale H, et al. 2008. *Blood* 111: 3838-48
259. Radic M, Marion T, Monestier M. 2004. *J Immunol* 172: 6692-700
260. Muzio M, Apollonio B, Scielzo C, Frenquelli M, Vandoni I, et al. 2008. *Blood* 112: 188-95
261. Ghia P, Chiorazzi N, Stamatopoulos K. 2008. *J Intern Med* 264: 549-62
262. Binder M, Lechenne B, Ummanni R, Scharf C, Balabanov S, et al. 2010. *PLoS One* 5: e15992
263. Kohnke PL, Mactier S, Almazi JG, Crossett B, Christopherson RI. 2012. *J Proteome Res* 11: 4436-48
264. Stacchini A, Aragno M, Vallario A, Alfarano A, Circosta P, et al. 1999. *Leuk Res* 23: 127-36
265. Boyd RS, Dyer MJ, Cain K. 2007. *Methods Mol Biol* 370: 135-46
266. Zeisig BB, So CW. 2013. *Cancer Cell* 24: 5-7
267. Miller PG, Al-Shahrour F, Hartwell KA, Chu LP, Jaras M, et al. 2013. *Cancer Cell* 24: 45-58
268. Surks HK, Richards CT, Mendelsohn ME. 2003. *Circulation* 108: 67-
269. Surks HK, Richards CT, Mendelsohn ME. 2003. *Journal of Biological Chemistry* 278: 51484-93
270. Surks HK, Riddick N, Ohtani K. 2005. *Journal of Biological Chemistry* 280: 42543-51
271. Ono R, Matsuoka J, Yamatsuji T, Naomoto Y, Tanaka N, et al. 2008. *International Journal of Molecular Medicine* 22: 199-203
272. Becker B, Multhoff G, Farkas B, Wild PJ, Landthaler M, et al. 2004. *Exp Dermatol* 13: 27-32
273. McCready J, Sims JD, Chan D, Jay DG. 2010. *BMC Cancer* 10: 294
274. Lv LH, Wan YL, Lin Y, Zhang W, Yang M, et al. 2012. *J Biol Chem* 287: 15874-85
275. Liu X, Yan Z, Huang L, Guo M, Zhang Z, Guo C. 2011. *Oncol Rep* 25: 1343-51
276. Tsutsumi S, Neckers L. 2007. *Cancer Sci* 98: 1536-9
277. Shin BK, Wang H, Yim AM, Le Naour F, Brichory F, et al. 2003. *J Biol Chem* 278: 7607-16
278. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, et al. 2001. *N Engl J Med* 344: 783-92
279. Raymond E, Faivre S, Armand JP. 2000. *Drugs* 60 Suppl 1: 15-23; discussion 41-2

Appendix

PARTICIPANT INFORMATION SHEET

King's College London Haemato-Oncology Tissue Bank: The Collection and Storage of Blood and Tissue for Use in Future Studies into the Causes, Diagnosis and Treatment of Haematological Disorders.

You are being invited to make a voluntary contribution of tissue to a Tissue Bank that will be used for research into blood disorders. Before you decide it is important for you to understand why the Tissue Bank is needed and what research will be carried out using tissue from the Tissue Bank. Please take time to read the following information carefully and discuss it with friends, relatives and your GP if you wish. Please ask us if there is anything that is not clear, or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the Tissue Bank and the Research it will be used for?

Medical research depends upon a steady supply of tissue from patients. The purpose of a Tissue Bank is to conserve tissue so that samples in sufficient numbers are ready when they are needed. Our Tissue Bank will supply scientists who are undertaking research to gain a better understanding of the causes of blood disorders such as aplastic anaemia, lymphomas, leukaemia, myelodysplastic syndromes (MDS) and myeloma. As our understanding in this area is advancing rapidly, we would like your permission to store some of your bone marrow, blood cells, serum and any tissue surplus to that required to confirm a diagnosis.

Do I have to take part?

You are not under any obligation to take part. Your participation is entirely voluntary and there will be no payment for entering into the study. If you decide to take part you will be asked to sign a consent form and you will be given this information sheet to keep. Staff will explain all the risks, benefits and alternatives before they ask you to sign the consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. This will not affect the standard of care you receive. Should you decide to withdraw your consent for continued retention of your tissue, any tissue remaining in the Tissue Bank at that time will be destroyed according to your wishes, as far as possible, and any data relating to you and your tissues, other than that needed to record the receipt and fate of your tissue, will be deleted. Please note that if you do withdraw your consent after donating tissues it may not be possible for researchers to delete data already obtained from research using your tissue.

What will happen if I take part?

All these samples will be collected during normal routine blood or bone marrow collections and there will be no additional procedures that you will have to undergo and no additional discomfort.

At diagnosis, you may be invited to provide up to 60ml (4 tablespoons) of blood and 20ml (1½ tablespoons) of bone marrow. We may also invite you to provide the collected cells if you need to have a procedure called leukapharesis (the cells collected from a machine if your white blood cells are high in the blood - what is involved in this and the consent for this procedure will be explained and taken separately)

If a biopsy is performed, we may invite you to consent for the storage of any excess tissue which would otherwise be destroyed after the diagnosis is confirmed.

We may also invite you to provide a sample of buccal cells from your mouth. This involves rinsing your mouth with a solution of salt and collecting the liquid or using a cheek swab. Both methods are painless.

We would like to collect material at regular intervals during the course of your treatment. During and after treatment, you may be invited to provide additional samples when you have your routine blood and bone marrow collections. We will not ask you to donate for this purpose any more than 60ml of blood or 5ml of bone marrow in any 28 day period.

Samples will be stored within our Tissue Bank, which has been licensed by the Human Tissue Authority.

Tissues will subsequently be used by scientists carrying out research into the causes, diagnosis and treatment of Haematological Disorders. In some cases we may need to send a portion of your tissue to our collaborators in academic institutions, or occasionally in commercial organisations, in other parts of the UK or abroad. Some of the research may involve the use of animals, but only when this is strictly necessary. If you choose to donate your tissue to the Tissue Bank but do not wish those tissues to be used in studies involving animals, please indicate this on the consent form and we will respect your wishes. Samples will not be used for any form of reproductive research. In all cases we will treat your tissue with respect and all experiments will have been approved by independent reviewers.

What are the side effects?

Blood sampling, bone marrow examination and leukapharesis are routine examinations performed in the haematology department. Each of the procedures will be explained in full and consent taken. No additional discomfort would be expected from the procedures for this study.

What is the potential benefit of participation?

Samples may be tested for genetic abnormalities which may be responsible for blood disorders, therefore leading to a better understanding of the causes of the condition and better treatment. The information could also predict how patients will respond to treatment and be used to monitor how effective the treatment has been.

You may not benefit directly from our work, but if we do discover information that could help your treatment then you and your doctor will be informed.

Results from this study will be published in scientific journals, but no information will be included which will allow participants to be identified.

What if something goes wrong?

If you wish to make a complaint about any aspect of the Tissue Bank, please contact the Director of Research Management and Director of Administration (Health Schools), King's College London, Strand, London, WC2R 2LS

Will my taking part in this study be kept confidential?

Your contribution to the Tissue Bank will be kept strictly confidential. All your sample(s) and medical information will be stored securely and will be anonymised on receipt by the Tissue Bank so that researchers will not be able to identify you.

Who is organising and funding the research?

Professor GJ Mufti and Dr Nigel Westwood are the Director and Manager of the Tissue Bank, respectively.

Leukaemia & Lymphoma Research (formerly the Leukaemia Research Fund) provided funding to set up the Tissue Bank. The project has also been reviewed by a committee of the UK National Research Ethics Service.

Contact for further information

If you require any further information about participating in the KCL Haemato-Oncology Tissue Bank please contact Dr Nigel Westwood (Tissue Bank Manager) or Professor GJ Mufti (Director), Department of Haematological Medicine, King's College London, Rayne Institute, 123 Coldharbour Lane, London, SE5 9NU, telephone 020 7848 5815.

If you do decide to donate to the Tissue Bank you will be given a copy of this information sheet and a copy of the consent form to keep.

Thank you for your interest in donating samples to our tissue bank