Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.
Prognostic Factors in
Chronic Lymphocytic Leukemia

A thesis submitted for the degree of
Doctor of Philosophy
Amjad Hayat MBBS, MRCP, FRCPath

Trinity College,
University of Dublin
April 2008
Declaration

I hereby certify that this thesis, submitted for the degree of Doctor of Philosophy to the University of Dublin, has not been previously submitted for a degree or diploma to this or any other university. The work presented here is entirely my own, except where stated.

This thesis may be made available for consultation within the university library and may be photocopied or lent to other libraries for the purpose of consultation.

Amjad Hayat

April 2008
The most exciting phrase to hear in science, the one that heralds new discoveries, is not Eureka! (I found it) but rather, "hmm....that's funny...."

Issac Asimov
ABSTRACT

Chronic lymphocytic leukaemia (CLL) is the most common leukaemia in the Western world, accounting for nearly 24% of all leukaemias. It is estimated that 7300 new cases were diagnosed in 2003 in the United States and there were 4400 deaths due to CLL. At diagnosis, many patients are asymptomatic but subsequently develop symptoms and die from the disease or its complications.

The most important initial challenge in patients with CLL is defining prognosis, early recognition and management of complications and designing of risk-stratified management. The clinical course of CLL is however, extremely variable. The majority of patients with CLL do not require therapy at diagnosis. One third never require therapy and in these patients the disease does not significantly affect life expectancy. In contrast, approximately one third of patients have aggressive CLL, which requires early therapeutic intervention and also results in a shortened median survival of between 18 months and 3 years. The remaining one third of patients do not need therapy at diagnosis, but may need treatment at a later stage.

CLL is not curable with standard therapy and initiating chemotherapy following diagnosis in asymptomatic patients does not alter survival. The current standard of care is to treat CLL when the patient becomes symptomatic or the disease has progressed. Newer and more effective treatment protocols with curative potential are being developed. The possibility of intervention at an early stage, when the disease burden is minimal and before clonal evolution occurs, may be more effective at prolonging survival and decreasing disease related mortality. If clinicians are going to adopt this approach, we have to have a robust method of risk analysis so that potentially toxic therapy in patients with a small risk of disease progression can be avoided.

In this study we investigated a cohort of CLL patients using novel prognostic markers, including cytogenetic abnormalities, mutational status of the immunoglobulin heavy chain and immunophenotypic markers CD38 and ZAP-70 for risk stratification.

The analysis includes 106 patients, 66 male and 40 female, with a median age of 62 (range 37-90) and Binet stages A: n=87, B: n=12 and C: n=7. Deletion 13q was the sole abnormality in 51 (51%) patients. Twenty-one (41%) of 51 patients with 13q14 deletion were treated, compared to 30 (59%) who were not treated and patients had a long treatment free interval (TFI) of 110 months. Trisomy 12 was detected in 11% of patients. The median TFI for this group of patients was 41 months. Deletion of 11q was the third most common abnormality detected in 6% of patients. The majority of these patients (83%) required treatment with a short median TFI of 5 months. Only 4 patients (4%) had a 17p deletion. The clinical significance of 17p deletion was evident however, as 75% of the patients required treatment with a median TFI of 11 months. In 10 patients, we detected more than one chromosomal abnormality: trisomy 12 and del 13q14, del 11q and del 13q14, del 17p and del 13q14.

Sixty-nine (65%) patients had Ig-Mutated genes, 35 (33.3%) patients were CD38 positive, with discordant results in 30 (28%) patients for the two assays. Fifteen of 95 patients expressed ZAP-70 in >20% of cells, which did not predict IgVH mutational status in 22 (23%) of patients. TFI from diagnosis was used as the clinical end point. Median TFI for patients with Ig-Mutated vs. Ig-Unmutated genes was 120 vs. 12 months, Hazard Ratio (HR), 3.9 (95%CI, 2.2-6.9; p=0.001). Median TFI for CD38-ve vs. CD38+ve patients was 114 vs. 13 months, HR, 2.8 (95%CI, 1.5-5.1; p=0.001) and median TFI for ZAP-70-ve vs. ZAP70+ve was 102 vs. 5 months, HR, 3.2 (95% CI, 1.5-6.8; p=0.002).

We compared the relationship of cytogenetic abnormalities to other prognostic markers. The majority of patients with favourable cytogenetic abnormalities (del 13q14) had mutated IgVH genes (82%) and were CD38 negative (82%) and ZAP-70 negative (90%). On the other hand, the majority of the patients with unfavourable cytogenetic abnormalities such as trisomy 12, del 11q and 17p deletion expressed CD38. Fifty percent of the patients with trisomy 12 and 17p deletion had unmutated IgVH genes and all the patients with 11q deletion had unmutated IgVH genes.
In Cox proportional-hazards analysis, only CD38 and IgV_{H} retain prognostic significance. In our series, ZAP-70 expression did not have independent prognostic significance in patients with known FISH, IgV_{H} mutational status and CD38 expression. The discordance rate of 23% between ZAP-70 and IgV_{H} means that ZAP-70 is not a useful surrogate marker for IgV_{H} mutational status.

Fludarabine has emerged as the most effective therapeutic agent in CLL, either alone or in combination, with overall response rate (ORR) of 80% and CR rates of 32% to 47%. Despite the improvement in the complete response (CR) rates most patients have residual disease post treatment, which leads to subsequent relapse. Recent data suggests that Rituximab (a monoclonal antibody against CD20 antigen) acts synergistically with Fludarabine. There is also a correlation between the type of response to treatment and time to progression in CLL.

Our goal was to improve the response rates achieved in CLL patients by adding Rituximab to the combination of Fludarabine and Cyclophosphamide.

We adopted a modified FCR protocol in treating 39 patients, median age of 57 years (38-75) and progressive or advanced CLL. Depending on response (CR), treatment was given for 4 or 6 cycles. Twenty-six patients were treatment naïve and 13 were pre-treated. Twelve patients had progressive Binet stage A, 16 stage B and 11 stage C disease. The overall response rate (ORR) was 100%, with 75% achieving CR. Neutropenia was the major toxicity in 71/187 (38%) of the cycles. There were 5 deaths. Twenty-six of 31 alive and evaluable patients have maintained their post-treatment disease status for a median of 17 months (2-41). The FCR regimen in our opinion is well tolerated and achieves a high CR in a group of patients with poor prognosis.

Protein expression profiling is the identification of proteins in a particular sample as a function of a particular state (differentiation, developmental state, or disease state) or as a function of exposure to a drug, chemical or physical stimulus. It is most commonly used to compare normal and diseased cells or tissues for differential protein expression. The interest for proteomics within the oncology field is most obvious: cancer is a DNA disease that originates in mutated genes and leads to aberrant protein expression. Within the context of both acute and chronic haematological diseases, because of the spectrum of the disease process, effective biomarkers are critically needed for diagnosis, risk stratification, prognosis and determining the effectiveness of therapeutic treatment. Furthermore, they could lead to new pharmacological targets. If these markers could be measured in a readily accessible body fluid such as serum, not requiring a tissue biopsy, it would have a major impact on future cancer diagnosis and treatment monitoring. Whereas proteomic studies have begun in several arenas in haematological diseases, there are relatively few in CLL. We decided to use a global protein pattern assessment on plasma collected at three time points in 13 patients treated with FCR. The aim was to to pick up differences in initial and response profiles corresponding to different prognostic groups and clinical outcome and to see whether proteomic analysis reveals specifically and potentially valuable prognostic indicators "biomarkers" in this group of well characterized patients treated with a novel chemotherapeutic regimen.

A number of protein signals (3971Da, 6517Da, 9908Da, 11.7kDa and 79kDa) were detected as significantly different between the different groups (p-values of ≤0.05). Three of the proteins were identified as apolipoprotein C-I, apolipoprotein C-III and transferrin. Apolipoprotein C-I and transferrin were both under expressed in CD38+ve patients. Apolipoprotein C-I increased with the initiation of therapy while, Apolipoprotein C-III was over expressed in the pre-treated patients, with the level going down in response to treatment.

The exact role that these proteins play in the biology of CLL is not clear at present however, this study and the growing literature on the use of SELDI-TOF indicate its suitability for the identification of tumour biomarkers.
For
The two most important women in my life
Hilary, my wife
Husna, my mother
And the finest man I have ever known
My father
Hayat
Completing a PhD is truly a marathon event, and I would not have been able to complete this journey without the aid and support of countless people over the past five years. It is hard to write appropriate acknowledgements that reflect the assistance and inspiration that I was given throughout my years in research.

First of all, I would like to thank my supervisor, Dr Elisabeth Vandenberghe. Her leadership, support, attention to detail, hard work, and scholarship have set an example I hope to match some day. I am grateful for her unparalleled supervision, frank scientific discussions and continuous encouragement during this research, not to mention the vast amount of time she spent on the countless numbers of corrections. I am also thankful to her for making it possible for me to work in international centres of excellence, which I really enjoyed. I doubt that I’ll ever be able to convey my appreciation fully, but I owe her my eternal gratitude.

A very special thanks goes to Professor Mark Lawler, without whose motivation and encouragement I wouldn’t have finished this thesis. He provided me with the direction, technical support and became more of a mentor and friend, than a professor.

I would like to thank three important groups of people and collaborators, without whom this dissertation would not have been possible. Professor Terry Hamblin and his team, Zadie Davis, Rachel Ibbotson and Martin O'Connor at Royal Bournemouth Hospital for helping me with the sequencing work and then following it up with the constant electronic and telephonic support that I required of their expertise. Professor Iwona Wlodarska at Katholike University Leuven, for introducing me to FISH and finally Professor Finbar Cotter and especially Emma Burte at St. Barts for helping me with the proteomics despite her busy schedule.

This thesis would not have been possible without the patients and the samples that they generously agreed to give. I am greatly indebted to them. I would also like to thank and express my gratitude to Siobhan McGuckin for not only helping me with sample and data collection but also encouraging and supporting me all along.
I would also like to thank the Immunophenotyping laboratory in St. James Hospital, Dave, Paul and Stuart for their enormous help and time that they spent with me in setting up ZAP-70 assay.

I am indebted to the Bone Marrow for Leukaemia Trust and the Health Research Board of Ireland for the generous grant that made it possible for me to work in some of the finest research institutes in the world.

To my lab-mates, thanks for the fun and support. I have been so fortunate to work with an amazing bunch of people in the lab throughout my PhD (Fiona, Kathy, Antoinette, Lisa, Tony, Elaina, Ruth, Laure, Prema, Armelle, Jacqui, Rustum, Ludmila, Danni, Joanne, Anne-Marie and Haradikar). Your friendship and technical tips are sincerely appreciated. I should also mention my neighbours in Thoracic Oncology, and Haemostasis and Thrombosis, for providing invaluable entertainment at morning coffee. Many thanks to Frances, Professor Lawler’s secretary for helping me with many things beyond her call of duty.

Finally I would like to thank my wonderful parents for instilling in me the confidence and a drive for pursuing my PhD. Their support and prayers have always been a source of encouragement for me.

And for the person who has made the many hours spent in the lab seem worthwhile after all, Hilary.
CONTENTS

SECTION 1

Relationship between IgVH mutation, CD38 and ZAP-70 expression in chronic lymphocytic leukemia

CHAPTER 1 INTRODUCTION

1.1 Chronic Lymphocytic Leukemia (CLL)

1.1.1 Background to CLL 3
1.1.2 Incidence of CLL 4
1.1.3 Clinical manifestations 7
1.1.4 Diagnosis of CLL 8
1.1.5 Staging of CLL 13

1.2 Conventional prognostic factors in CLL

1.2.1 Introduction 14
1.2.2 Clinical staging 15
1.2.3 Marrow involvement 16
1.2.4 Lymphocyte morphology and Lymphocyte Doubling Time (LDT) 17
1.2.5 Proliferation marker: Serum Thymidine Kinase (sTK) 18
1.2.6 Cellular expression and soluble CD23 18
1.2.7 β2m 18
1.2.8 Angiogenesis 19
1.2.9 Other biological markers 19

1.3 Treatment of CLL

1.3.1 Goals of treatment 19
1.3.2 Indications for treatment 20
1.3.3 Currently available treatments 21
1.3.4 Currently used treatment strategies 24
1.3.5 Future directions 25

1.4 Problems with the conventional model 26

1.5 Genomic aberrations in CLL

1.5.1 Introduction 27
1.5.1.1 Banding studies 27
1.5.1.2 Interphase cytogenetics 27
1.5.2 Incidence of chromosomal abnormalities in CLL 28
1.6 Immunoglobulin variable region gene characteristics in CLL

1.6.1 B-Cell development

1.6.2 Immunoglobulin (Ig) gene rearrangement and expression

1.6.3 Stages in B-cell development: Expression of Ig chains and cell-surface proteins

1.6.4 Isotype Class Switching

1.6.5 Somatic Hypermutation

1.6.6 Characteristics of Ig variable region genes in B-CLL cells

1.6.7 Clinical significance

1.7 Advances in the immunophenotypic characteristics of B-CLL cells

1.7.1 General Immunophenotypic features of B-CLL cells

1.7.2 CD38 and CLL

1.7.3 ZAP-70 and CLL

1.8 Aims of the Project

1.8.1 Establish a cohort of Irish CLL patients for molecular studies

1.8.2 Frequency and significance of karyotypic abnormalities

1.8.3 Determine the mutational status of immunoglobulin heavy chain variable regions (IgVH)

1.8.4 CD38 expression, its relation to IgVH and its significance as an independent prognostic marker

1.8.5 ZAP-70 expression, its relation to IgVH and CD38 and its significance as an independent prognostic marker

1.8.6 Use of FCR chemomimunotherapy, and the search for new biomarkers using Proteomics

CHAPTER 2 MATERIALS AND METHODS

2.1 Patients

2.1.1 Consent, History, Clinical examination

2.2 Morphology

2.2.1 Peripheral Blood

2.2.2 Bone marrow involvement

2.3 Cell separation and storage

2.3.1 Density gradient cell separation

2.3.2 Cell storage for Fluorescent in situ hybridisation (FISH)

2.3.3 Storage of cells for nucleic acid extractions

2.3.3.1 DNA extraction and quantification

2.3.3.2 RNA extraction and quantification
2.3.4 Freezing cells
2.3.5 Thawing cells

2.4 Fluorescence in situ hybridization (FISH)

2.4.1 Slide preparation
2.4.2 Pepsin digestion
2.4.3 Denaturation of the slides/sample and probe
   2.4.3.1 Slides/Samples
   2.4.3.2 Probe
2.4.4 Hybridization
2.4.5 Washing
2.4.6 Sample Analysis
   2.4.6.1 Probe description and analysis

2.5 IgVH mutational analysis

2.5.1 cDNA synthesis
2.5.2 IgVH Specific PCR using cDNA
   2.4.2.1 Primers
   2.4.2.2 Reaction conditions
2.5.3 IgVH Specific PCR using genomic DNA
2.5.4 Gel Electrophoresis
2.5.5 PCR cleanup
   2.5.5.1 PCR purification
   2.5.5.2 Gel Purification
2.5.6 Sequencing
2.5.7 Cloning

2.6 CD38 expression analysis
2.7 ZAP-70 expression analysis
2.8 Statistical analysis

CHAPTER 3 RESULTS

3.1 General Results
3.2 Results for FISH analysis
3.3 Results for IgVH mutational analysis
3.4 Results for CD38 expression analysis
   3.4.1 CD38 expression using percentage cut-off
   3.4.2 Pattern of expression
3.5 Results for ZAP-70 expression analysis

CHAPTER 4 DISCUSSION
SECTION 2
Novel Chemoimmunotherapy and the
Search for biomarkers

CHAPTER 5 FCR Chemoimmunotherapy in CLL

5.1 Introduction 143
5.2 Patients and Methods 144
5.2.1 Patient Selection 144
5.2.2 Treatment 144
5.3 Results 145
5.4 Discussion 149

CHAPTER 6 PROTEOMICS

6.1 Introduction to Proteomics

6.1.1 Background 153
6.1.2 Challenges in Proteomics 154
6.1.3 Proteomic techniques 155
6.1.3.1 Protein Separation/Fractionation 156
6.1.3.2 Mass Spectrometry 159
6.1.4 SELDI-TOF MS 162
6.1.4.1 Background 162
6.1.4.2 Sample fractionation for SELDI 164
6.1.4.3 The ProteinChips/Affinity Arrays 165
- Immobilized Metal Affinity Capture 166
- Weak Cation Exchanger (CM10) 167
- Strong Anion Exchanger (Q10) 168
- Hydrophobic binding Surface (HS10) 169
- Normal Phase (NP20) 170
6.1.4.4 The ProteinChip SELDI Reader 170
6.1.5 Applications of Proteomics 171
6.1.6 SELDI-TOF MS Applications 172
6.1.7 Proteomics in Haematology 174
6.1.8 Aim of the Project 176

6.2 Materials and Methods

6.2.1 Patients 177
6.2.2 Treatment 178
6.2.3 Sample Collection 178
6.2.4 Sample Assessment 179
6.2.4.1 Fractionation 179
6.2.5 ProteinChip protocols 182
6.2.5.1 IMAC30 Profiling Protocol 182
6.2.5.2 CM10 Profiling Protocol
6.2.5.3 Q10 Profiling Protocol
6.2.5.4 H50 Profiling Protocol
6.2.6 Data Collection Protocols
6.2.6.1 Spot Protocol
6.2.7 Data Analysis
6.2.7.1 Spectrum processing
6.2.8 Ciphergen Express
6.2.9 Biomarker Analysis
6.2.10 Protein Identification
6.2.10.1 Large Scale fractionation
6.2.10.2 Reverse Phase Chromatography
6.2.10.3 1D SDS-PAGE
6.2.10.4 Peptide Identification

6.3 Results

6.3.1 General Results
6.3.2 Biomarker discovery
-9908Da
-6517Da
-79000Da
-3971Da
-11785Da
6.3.3 Protein Identification

6.3 Discussion

CHAPTER 7 CONCLUSION

APPENDICES
A. Expression Difference Mapping Kit-Serum Fractionation Buffers

REFERENCES

PUBLICATION
CHAPTER 1 INTRODUCTION

Fig. 1.1 Age-adjusted world incidence of CLL 5
Fig. 1.2 SEER Incidence – Age-adjusted Rates for the USA, 1973-2001 6
Fig. 1.3 Age-specific incidence of CLL 7
Fig. 1.4 Peripheral blood film from a patient with CLL showing lymphocytosis 9
Fig. 1.5 Smudge cells on a peripheral blood film 10
Fig. 1.6 Patterns of marrow involvement in CLL
   a. Nodular 11
   b. Interstitial 11
   c. Diffuse 12
Fig. 1.7 Survival according to the modified RAI staging criteria 16
Fig. 1.8 Structure of an antibody molecule 33
Fig. 1.9 Germline organizations of the Ig heavy- and light-chain loci in humans 34
Fig. 1.10 Construction of the variable regions 36
Fig. 1.11 B cell development 38
Fig. 1.12 Antibody class switching 39
Fig. 1.13 B cell differentiation 40
Fig. 1.14 Comparison of V_H gene use between CLL and normal blood B cells 42
Fig. 1.15 Specific V_H gene use between IgM⁺ CLL and normal blood B-cells 43
Fig. 1.16 KM survival curve comparing CLL patients with mutated and unmutated V_H genes 44
Fig. 1.17 KM survival curve comparing stage A CLL patients with mutated and unmutated V_H genes 45
Fig. 1.18 Survival based on CD38 expression among B-CLL patients 48

CHAPTER 2 MATERIALS AND METHODS

Fig. 2.1 Abnormal hybridization showing one copy of p53 gene with Probe Set 1 63
Fig. 2.2 Abnormal hybridization showing one copy of ATM gene with Probe Set 1 63
Fig. 2.3 Trisomy 12 with probe set 2 64
Fig. 2.4 Deletion 13 and trisomy 12 with probe set 2 64
Fig. 2.5 The Immunoglobulin Molecule 66
Fig. 2.6 The Immunoglobulin Heavy chain gene rearrangement 67
Fig. 2.7 Proposed primer annealing on the immunoglobulin gene 68
Fig. 2.8 Example of PCR results on gel electrophoresis 70
Fig. 2.9 The sequence of immunoglobulin heavy chain using two PCR reactions 73
Fig. 2.10 Consensus sequence after complimenting and aligning 74
Fig. 2.11 Consensus sequence after alignment to V-base 75
Fig. 2.12 Analysis of CD38 expression (A-C)
   A. Plot of CD45 v SSC 79
   B. Plot of CD5vCD19 80
   C. Dot-Plots and histograms showing CD38 expression in CD19 expressing lymphocytes 81
Fig. 2.13 Assessment of Zap-70 expression (A-F)
   A. FSC v SSC raw data 83
   B. Region R1 encompassing lymphocytes 82
   C. R1 gated cells: Showing Region R2 encompassing CD5/CD19 co-expressing lymphocytes 83
   D. CD5vCD3+56 (R1 gated). Region R3 lymphocytes encompassing T and NK cells 83
   E. Plot of ZAP-70 v CD3+56 84
   F. ZAP-70 expression compared to the internal control 84
CHAPTER 3 RESULTS

Fig. 3.1 Cytogenetic abnormalities Vs. Treatment
Fig. 3.2 KM survival curves comparing TFI in CLL patients according to the Cytogenetic abnormalities
Fig. 3.3 Primary PCR results on a gel
Fig. 3.4 Nested PCR results on a gel
Fig. 3.5 Cloning PCR results on a gel
Fig. 3.6 Sequence of the immunoglobulin heavy chain using two PCR reactions and analysed on DNASTAR Seqman II software
Fig. 3.7 Consensus sequence after analyses on DNASTAR Seqman II software
Fig. 3.8 Consensus sequence after alignment to V-base
Fig. 3.9 Formatted sequence using 2 forward and 2 reverse sequences
Fig. 3.10 Distribution of different $V_H$ gene family usage among 106 patients
Fig. 3.11 KM survival curves comparing TFI in CLL patients with Unmutated and mutated genes (A and B)
   A. All Patients
   B. Stage A patients only
Fig. 3.12 KM survival curves comparing TFI in CD38-ve and CD38+ve patients
Fig. 3.13 KM survival curves comparing TFI in CD38-ve and CD38+ve patients. (Stage A patients)
Fig. 3.14 Different patterns of CD38 expression in CLL cells
Fig. 3.15 KM survival curves comparing TFI in CLL patients according to the pattern of CD38 expression
Fig. 3.16 KM survival curves comparing TFI in CLL patients according to the pattern of CD38 expression. (Stage A patients)
Fig. 3.17 KM survival curves comparing TFI in ZAP-70-ve and ZAP-70+ve CLL patients
Fig. 3.18 KM survival curves comparing TFI in ZAP-70-ve and ZAP-70+ve CLL patients. (Stage A only)

CHAPTER 6 PROTEOMICS

Fig. 6.1 Biochemical Context of Genomics and Proteomics
Fig. 6.2 Proteomic analysis of biological samples: A Schematic representation
Fig. 6.3 Schematic representation of 2D SDS-PAGE
Fig. 6.4 The Ettan DIGE system
Fig. 6.5 The fundamental Principle of MS instrumentation used in Proteomics
Fig. 6.6 The experimental steps for analysing samples on SELDI protein chips
Fig. 6.7 ProteinChip® technology: SELDI TOF-MS Detection
Fig. 6.8 Affinity arrays from Ciphergen Biosystem
Fig. 6.9 Types of arrays
Fig. 6.10 IMAC ProteinChip Array metal binding surface with metal and protein
Fig. 6.11 CM10 ProteinChip Array surface chemistry with protein
Fig. 6.12 Q10 ProteinChip Array surface chemistry with protein
Fig. 6.13 ProteinChip Array surface chemistry with proteins
Fig. 6.14 ProteinChip Array surface chemistry with proteins
Fig. 6.15 Ciphergen PBS II Series ProteinChip reader
Fig. 6.16 Sample Fractionation and assessment of samples on all 4 chip surfaces
Fig. 6.17 A flow chart of the Q hyperD F beads plate fractionation steps
Fig. 6.18 Spot Protocol
Fig. 6.19 Steps for protein identification
Fig. 6.20 Comparative analyses of fraction 6 on 4 ProteinChips
Fig. 6.21 Fractions 1 to 6 on CM10 ProteinChip
CHAPTER 7 CONCLUSION

Fig. 6.22 Dot plot and Spectra for protein 9908 198
Fig. 6.23 ROC curve for protein 9908 199
Fig. 6.24 Dot plot and spectra for protein 6517 200
Fig. 6.25 ROC curve for protein 6517 201
Fig. 6.26 Dot plot for protein 79000 201
Fig. 6.27 ROC curve for protein 79000 202
Fig. 6.28 Dot plot for protein 3971 202
Fig. 6.29 ROC curve for protein 3971 203
Fig. 6.30 Dot plot for protein 11785 203
Fig. 6.31 Representative spectra from an NP20 run on the RPC enriched fractions 205
Fig. 6.32 1D-SDS gel stained using colloidal comassie blue (A) and Silver (B) 206

CHAPTER 7 CONCLUSION

Fig. 7.1 CLL treatment options by decade 216
Fig. 7.2 An Algorithm for the management of CLL patients 218
LIST OF TABLES

CHAPTER 1 INTRODUCTION

Table 1.1  Scoring system for the diagnosis of CLL  12
Table 1.2  Rai classification system for CLL  13
Table 1.3  Binet classification system for CLL  13
Table 1.4  Factors associated with poor prognosis  15
Table 1.5  NCI-WG Criteria for initiating Treatment on a Protocol  21
Table 1.6  Frequency of chromosomal abnormalities in CLL using FISH  28
Table 1.7  Cytogenetic abnormalities and prognosis  31

CHAPTER 2 MATERIALS AND METHODS

Table 2.1  Primer Sequences  68
Table 2.2  Reaction set up for PCR  69
Table 2.3  Reaction set up for sequencing PCR  71
Table 2.4  TOPO TA ligation reactions  76
Table 2.5  Reaction set up for the sequencing PCR  77

CHAPTER 3 RESULTS

Table 3.1  Main biologic and clinical characteristics of the patients  87
Table 3.2  Stage and treatment according to the Binet staging  88
Table 3.3  Stage and treatment according to typical and atypical CLL  89
Table 3.4  Histological patterns of marrow involvement  89
Table 3.5  Treatment and stage according to $\beta_2$m  90
Table 3.6  Treatment and stage according to LDH  91
Table 3.7  Incidence of chromosomal abnormalities in 100 patients  92
Table 3.8  Distribution of Cytogenetic abnormalities in 100 patients by FISH  93
Table 3.9  The distribution of $V_h$, $D_h$ and $J_h$ genes in 106 patients  106
Table 3.10  Comparison of Treatment requirements based on IgVH status, CD38 % and ZAP-70 %  108
Table 3.11  Distribution of the different VH gene family usage  109
Table 3.12  Sensitivity and specificity of CD38 for predicting mutational status  112
Table 3.13  Comparison of treatment requirements based on IgVH status, CD38 % and ZAP-70 %  113
Table 3.14  Sensitivity and specificity of ZAP-70 for predicting the mutational status  119
Table 3.15  Comparison of Treatment requirements based on IgVH status, CD38 % and ZAP-70 %  120

CHAPTER 4 DISCUSSION

Table 4.1  Survival in CLL according to Clinical Stage  125
Table 4.2  Incidence of chromosomal abnormalities in CLL identified by FISH  128
Table 4.3  Univariate analysis  138
Table 4.4  Multivariate analysis  138

CHAPTER 5 FCR CHEMOIMMUNOTHERAPY IN CLL

Table 5.1  Patient Characteristics  146
Table 5.2  Treatment Outcome  147
Table 5.3  Treatment Associated Toxicity  148
CHAPTER 6 PROTEOMICS

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 6.1</td>
<td>Spot protocols capture setting for low and high mass</td>
<td>185</td>
</tr>
<tr>
<td>Table 6.2</td>
<td>Biologic and clinical characteristics of treated patients</td>
<td>192</td>
</tr>
<tr>
<td>Table 6.3</td>
<td>Biologic and clinical characteristics of the Untreated patients</td>
<td>193</td>
</tr>
<tr>
<td>Table 6.4</td>
<td>Differentially expressed proteins</td>
<td>197</td>
</tr>
<tr>
<td>Table 6.5</td>
<td>Samples and fractions containing the biomarkers of interest</td>
<td>204</td>
</tr>
<tr>
<td>Table 6.6</td>
<td>Characteristics of the MASCOT identified proteins</td>
<td>207</td>
</tr>
<tr>
<td>Table 6.7</td>
<td>Peptide identification</td>
<td>208</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>ACN</td>
<td>Acetonitrile</td>
<td></td>
</tr>
<tr>
<td>APC</td>
<td>Allophycocyanin</td>
<td></td>
</tr>
<tr>
<td>Apo C-I</td>
<td>Apolipoprotein C-I</td>
<td></td>
</tr>
<tr>
<td>Apo C-III</td>
<td>Apolipoprotein C-III</td>
<td></td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
<td></td>
</tr>
<tr>
<td>ATM</td>
<td>Ataxia Telengectasia mutated</td>
<td></td>
</tr>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
<td></td>
</tr>
<tr>
<td>BCR</td>
<td>B-cell-receptor</td>
<td></td>
</tr>
<tr>
<td>BM</td>
<td>Bone Marrow</td>
<td></td>
</tr>
<tr>
<td>β2m</td>
<td>Beta 2 Microglobulin</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Constant region</td>
<td></td>
</tr>
<tr>
<td>CAD</td>
<td>Collision activated dissociation</td>
<td></td>
</tr>
<tr>
<td>CAP</td>
<td>Cyclophosphamide, Adriamycin and Prednisolone</td>
<td></td>
</tr>
<tr>
<td>CDR</td>
<td>Complementarity determining region</td>
<td></td>
</tr>
<tr>
<td>CHOP</td>
<td>Cyclophosphamide, Doxorubicin, Vincristine and Prednisolone</td>
<td></td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
<td></td>
</tr>
<tr>
<td>CLL</td>
<td>Chronic Lymphocytic Leukemia</td>
<td></td>
</tr>
<tr>
<td>CM10</td>
<td>Weak Cation Exchanger</td>
<td></td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
<td></td>
</tr>
<tr>
<td>COP</td>
<td>Cyclophosphamide, Vincristine and Prednisolone</td>
<td></td>
</tr>
<tr>
<td>CR</td>
<td>Complete remission</td>
<td></td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
<td></td>
</tr>
<tr>
<td>DCT</td>
<td>Direct Coombs test</td>
<td></td>
</tr>
<tr>
<td>D_{H}</td>
<td>Diversity region of the heavy chain</td>
<td></td>
</tr>
<tr>
<td>DI</td>
<td>De-ionized</td>
<td></td>
</tr>
<tr>
<td>DIGE</td>
<td>Two-dimensional difference in gel electrophoresis</td>
<td></td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl Sulfoxide</td>
<td></td>
</tr>
<tr>
<td>EAM</td>
<td>Energy absorbing matrix</td>
<td></td>
</tr>
<tr>
<td>EDM</td>
<td>Expression difference mapping</td>
<td></td>
</tr>
<tr>
<td>EPG</td>
<td>Electropherogram</td>
<td></td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray ionization</td>
<td></td>
</tr>
<tr>
<td>FAB</td>
<td>French-American-British</td>
<td></td>
</tr>
<tr>
<td>FC</td>
<td>Cyclophosphamide and Fludarabine</td>
<td></td>
</tr>
<tr>
<td>FCR</td>
<td>Fludarabine, Cyclophosphamide and Rituximab</td>
<td></td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein isothiocyanate</td>
<td></td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescence in-situ Hybridisation</td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td>Framework region</td>
<td></td>
</tr>
<tr>
<td>FSC</td>
<td>Forward scatter</td>
<td></td>
</tr>
<tr>
<td>FT-ICR</td>
<td>Fourier transform ion cyclotron resonance</td>
<td></td>
</tr>
<tr>
<td>H50</td>
<td>Hydrophobic binding surface</td>
<td></td>
</tr>
<tr>
<td>HCL</td>
<td>Hydrochloric acid</td>
<td></td>
</tr>
<tr>
<td>HCT</td>
<td>Haematopoietic cell transplantation</td>
<td></td>
</tr>
<tr>
<td>HDMP</td>
<td>High dose methylprednisolone</td>
<td></td>
</tr>
<tr>
<td>HMG1</td>
<td>High mobility group 1</td>
<td></td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
<td></td>
</tr>
<tr>
<td>HV</td>
<td>Hypervariable region</td>
<td></td>
</tr>
<tr>
<td>HR</td>
<td>Hazard ratio</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>RAM-PE</td>
<td>Rabbit anti-mouse Phycoerithrin</td>
<td></td>
</tr>
<tr>
<td>RF</td>
<td>Radio frequency</td>
<td></td>
</tr>
<tr>
<td>RPC</td>
<td>Reverse phase chromatography</td>
<td></td>
</tr>
<tr>
<td>RPM</td>
<td>Revolutions per minute</td>
<td></td>
</tr>
<tr>
<td>RPMI</td>
<td>Roswell Park Medical Institute</td>
<td></td>
</tr>
<tr>
<td>RR</td>
<td>Response rate</td>
<td></td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
<td></td>
</tr>
<tr>
<td>SEER</td>
<td>Surveillance, Epidemiology and End Results</td>
<td></td>
</tr>
<tr>
<td>SELDI</td>
<td>Surface enhanced laser desorption and ionization</td>
<td></td>
</tr>
<tr>
<td>sTK</td>
<td>Serum Thymidine Kinase</td>
<td></td>
</tr>
<tr>
<td>SmIg</td>
<td>Surface membrane immunoglobulin</td>
<td></td>
</tr>
<tr>
<td>SSC</td>
<td>Side scatter</td>
<td></td>
</tr>
<tr>
<td>TdT</td>
<td>Terminal deoxynucleotidyl transferase</td>
<td></td>
</tr>
<tr>
<td>TFI</td>
<td>Treatment free interval</td>
<td></td>
</tr>
<tr>
<td>TIC</td>
<td>Total ion current</td>
<td></td>
</tr>
<tr>
<td>Tf</td>
<td>Tranferrin</td>
<td></td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
<td></td>
</tr>
<tr>
<td>TOF</td>
<td>Time of flight</td>
<td></td>
</tr>
<tr>
<td>TRM</td>
<td>Treatment related morbidity and mortality</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>Variable region</td>
<td></td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular endothelial growth factor</td>
<td></td>
</tr>
<tr>
<td>WCC</td>
<td>White cell count</td>
<td></td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
<td></td>
</tr>
<tr>
<td>ZAP</td>
<td>Zeta associated protein</td>
<td></td>
</tr>
</tbody>
</table>
SECTION 1

Relationship between $\text{IgV}_H$ mutation, CD38 and ZAP-70 expression in chronic lymphocytic leukemia
CHAPTER 1

INTRODUCTION
1.1 Chronic Lymphocytic Leukemia (CLL)

1.1.1 Background to CLL

Chronic Lymphocytic leukaemia is a neoplastic disease characterised by the accumulation of small, mature appearing lymphocytes in the blood, marrow and lymphoid tissues. The first descriptions of patients with CLL were published in the early 19th century (Velpeau 1827; Virchow 1845; Fuller 1846); several decades after leukaemia as a separate clinical entity was first described. In the 1840s, Virchow described two forms of chronic leukaemia, probably corresponding to CLL and chronic myeloid leukaemia (Virchow 1845; Virchow 1846; Virchow 1847). Patients with the former were noted to have mild to moderate splenic enlargement, lymphadenopathy and large numbers of small agranular cells in the blood that resembled those found in enlarged lymph nodes (Virchow 1846). Virchow considered this type of leukaemia to be primarily a disease of the lymphoid system. In 1893, Kundrat introduced the term lymphosarcoma to describe an indolent disease which affected lymph nodes (Kundrat 1893). Histochemical staining techniques introduced by Ehlirich at the turn of the century (Ehlirich 1891), made it possible for pathologists to distinguish between myeloid and lymphocytic leukaemias. These methods enabled Türk in 1903 to establish the relationship of leukaemic cells in CLL to those found in lymphosarcoma (Turk 1903). He proposed the term lymphomatosis to describe several lymphoproliferative disorders including CLL. Because of its indolent nature, CLL was considered a benign lymphomatosis. Sir William Osler reported his experience of CLL at the Johns Hopkins University in the early part of the 20th century (Osler 1909). Osler observed that CLL constituted approximately 22% of all leukaemias and was associated with generalised lymphadenopathy and long survival times of 3 to 11 years. In 1924 Isaacs published a series of 80 patients with CLL, reporting a median survival time of 40 months which was not improved by radiation therapy compared with no treatment (Minot and Issacs 1924). The next 50 years saw the development of precise diagnostic criteria for CLL, identification of various treatment options and an improved understanding of the natural history of the disease. A comprehensive review of the diagnostic criteria, clinical features, response to treatment and survival data from a large number of CLL patients was published by Boggs et al in 1966 (Boggs, Sofferman et al. 1966). In the same year Galton observed
that the rate of increase of the lymphocyte count in blood predicted the clinical outcome (Galton 1966). He also suggested that the primary abnormality in CLL was the accumulation of lymphocytes and not an increased proliferative rate. Dameshek who had made similar observations independently, proposed a method of risk stratifying CLL patients on the basis of clinical features which included disease related symptoms, lymph node enlargement, splenomegaly, hepatomegaly and blood count (Dameshek 1967).

The knowledge that CLL is a disorder of B-lymphocytes came from the observation of two groups of investigators simultaneously in 1972 (Aisenberg and Bloch 1972; Preud'homme and Seligmann 1972). Leukaemic cells from most cases of CLL were found to express surface immunoglobulin (Ig) indicating that they were of B-cell origin. Subsequent studies on CLL cells showed that a patient expresses only one type of idiotypic Ig-light chain, confirming their clonal nature (Fu, Winchester et al. 1974; Salsano, Froland et al. 1974; Schroer, Briles et al. 1974).

1.1.2 Incidence of CLL

Despite the development of appropriate classification schemes for CLL such as the international classification of diseases (ICD)(1977; Linet and Blattner 1988) and its extension, International classification of diseases for oncology (ICD-O)(1976), inadequate reporting to the cancer registries of newly diagnosed leukaemia cases and those first identified from death certificates prevent the compilation of the true incidence. For this reason, population-based mortality data for CLL is difficult to interpret, and population-based incidence data has to be critically evaluated in the light of the proportion of the total leukaemia cases which are incompletely designated or unspecified according to subtype.

Based on recent survey of data from different sources such as the US National Cancer Institute (NCI), Surveillance, Epidemiology and End Results (SEER) program and World Health Organisation (WHO), a review by Redaelli et al (Redaelli, Laskin et al. 2004) have reported a variable incidence of CLL in various countries around the world (Fig 1.1).
The highest rates are found in Australia, North America and Europe especially Ireland, Italy and Switzerland. Lower rates are found in Asia and South America (International agency for Research in Cancer 1997) (Linet and Cartwright 1996). In 1998 there were an estimated 7500 to 12,500 new cases of CLL in the USA (Cheson 2000). CLL incidence rates have been reported to be declining in recent years (Xie, Davies et al. 2003). In 2000, the USA had an age-adjusted incidence of 3.4 cases of CLL per 100,000 in the population (Fig 1.2) (Ries, Eisner et al. 2003). The National Cancer Registry of Ireland has not published separate incidence or mortality rates for CLL, which is grouped together with other leukemias and is the ninth most common cancer in 2005. However, many CLL patients are asymptomatic and are diagnosed incidentally. It is therefore likely that most cancer databases underestimate the incidence of CLL and that the real incidence may be higher, with one study suggesting an incidence of 6.8 per 100,000 or 19,000 new cases per year (Zent, Kyasa et al. 2001).
The incidence of CLL increases dramatically with age. CLL is extremely rare in people aged less then 30 but increases steeply from the fourth decade onwards. Rates continue to rise exponentially with increasing age and at a steeper rate than other leukaemia subtypes until the age of 70 (Fig 1.3). CLL is more common in males then in females. In 2000 in the USA, the age-adjusted incidence of males and females was 4.7 and 2.4 per 100,000 respectively (Ries, Eisner et al. 2003).
1.1.3 Clinical manifestations
At diagnosis, most patients are over 60 years of age and 90% are over 50. CLL is rare in persons under 25 years of age. There is a 2:1 male to female incidence and prevalence of CLL.

Over 25% of patients are asymptomatic at diagnosis and are detected because of the discovery of non-tender lymphadenopathy, unexplained absolute lymphocytosis or malaise. Patients may also experience non-specific symptoms, even when they apparently lack major organ involvement or anaemia. Because of the advanced age of the affected population, patients sometimes present with an exacerbation of another medical condition such as pulmonary, cerebrovascular, or coronary artery disease. If they are not picked up incidentally, the usual presenting complaints of patients with CLL are weakness, fatigue, night sweats, fever, weight loss and recurrent bacterial and viral infections.

Nearly 80% of all CLL patients have non-tender lymphadenopathy at diagnosis, usually of the cervical, supraclavicular or axillary lymph nodes. Lymph node
enlargement ranges from minimal to massive and can cause organ dysfunction, but rarely vascular or lymphatic obstruction.

Approximately half of all CLL patients present with mild to moderate splenomegaly, which can cause symptoms of early satiety or abdominal fullness. Splenic enlargement may result in hypersplenism contributing to anaemia and thrombocytopenia. Cytopenia in CLL is most commonly secondary to extensive marrow involvement by CLL and or autoimmune haemolysis / thrombocytopenia. Patients less frequently, develop hepatomegaly secondary to leukaemic cell infiltration of the liver.

1.1.4 Diagnosis of CLL

The National Cancer Institute working group (NCI-WG) on CLL developed the first Guidelines (Cheson, Bennett et al. 1988; Cheson, Bennett et al. 1996) for CLL and outlined three diagnostic requirements:

1. An absolute lymphocytosis in the blood, with a count of 5 x 10^9/l or higher, of cells morphologically mature in appearance sustained over at least four weeks.

2. At least 30% lymphocytes in a normocellular or hypercellular bone marrow.

3. A monoclonal B-cell phenotype expressed by the majority population of blood lymphocytes with low levels of surface immunoglobulins and simultaneously showing CD5 positivity (a pan T-cell marker).

The International workshop on CLL (IWCLL) (1989) proposed similar diagnostic criteria, but required a lymphocyte count of ≥ 10 x 10^9/l in the blood if immunophenotyping facilities are not available.

A definitive diagnosis of CLL is based on the combination of a lymphocytosis coupled with a characteristic lymphocyte morphology and immunophenotype. Examination of the bone marrow is not required for diagnosis. Morphologically, the leukaemic cells generally appear similar to normal resting lymphocytes, with scanty,
blue cytoplasm, moderately condensed and mature appearing nuclei with clumped nuclear chromatin on May-Grünwald-Giemsa staining (Fig 1.4).

Variability in cell morphology exists including cells with prominent nucleoli, cells with a round or notched nucleus, forms with an indistinct nucleolus and finally the cytoplasm may be abundant and slightly basophilic. CLL lymphocytes are disrupted and appear as smudge cells when spread in a blood film (Fig 1.5).
Four histological patterns of marrow infiltration are seen in CLL: (i) Nodular (Fig 1.6a); (ii) Interstitial (Fig 1.6b); (iii) Diffuse (Fig 1.6c); and (iv) Mixed (Geisler, Hou-Jensen et al. 1996; Montserrat, Villamor et al. 1996). A mixed pattern represents a combination of nodular and interstitial infiltration. Paratrabecular infiltration is not characteristic of CLL. In approximately one third of patients, the marrow has an interstitial pattern, 10% of patients present with a nodular pattern of marrow involvement and approximately 25% have a mixed nodular/interstitial pattern. A quarter of the patients present with extensive, diffuse marrow involvement which is associated with advanced clinical stage and aggressive disease (Montserrat, Marques-Pereira et al. 1984; Pangalis, Roussou et al. 1984; Pangalis, Boussiotis et al. 1993).
Figure 1.6a Nodular Pattern

Figure 1.6b Interstitial Pattern
Immunophenotyping should be performed in all cases and is of particular value in the following situations: (i) in cases with low lymphocyte counts to confirm the diagnosis of CLL and exclude reactive lymphocytosis; and (ii) in patients with atypical lymphocyte morphology to exclude other B or T-cell lymphoproliferative disorders. Typically, CLL cells express weak monotypic surface immunoglobulin, CD5, CD19, CD23 and weak or absent CD79B, CD22 and FMC7. A recommended panel of monoclonal antibodies and scoring system for the diagnosis of CLL is shown in Table 1.1 (Moreau, Matutes et al. 1997).

<table>
<thead>
<tr>
<th>Marker</th>
<th>Score points</th>
<th>Score points</th>
</tr>
</thead>
<tbody>
<tr>
<td>SmIg</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CD5</td>
<td>Weak</td>
<td>Strong</td>
</tr>
<tr>
<td>CD23</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>FMC7</td>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>CD22 or CD79b</td>
<td>Weak</td>
<td>Strong</td>
</tr>
</tbody>
</table>

Scores in CLL are usually >3, in other B-cell malignancies the scores are usually <3.

**Table 1.1** Scoring system for the diagnosis of CLL.
Using this scoring system, 92% of CLL cases score 4 or 5, 6% score 3 and 2% score 1 or 2. Most other chronic B-cell lymphomas and leukaemias score 1 or 2, but a minority score 3.

1.1.5 Staging of CLL

The Rai and Binet staging system for CLL currently in use were originally devised 25-30 years ago. The fact that these systems are still widely used is an indication of their clinical usefulness and robustness. The Rai clinical staging system was first published in 1975 with modifications made in 1987 (Rai, Sawitsky et al. 1975; Rai 1987) Table 1.2.

<table>
<thead>
<tr>
<th>Risk level</th>
<th>Stage</th>
<th>Clinical features at diagnosis</th>
<th>Median survival time</th>
<th>% of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0</td>
<td>Blood and marrow lymphocytosis</td>
<td>≥10 years</td>
<td>30%</td>
</tr>
<tr>
<td>Intermediate</td>
<td>I</td>
<td>Lymphocytosis and lymphadenopathy</td>
<td>9 years</td>
<td>35%</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>Lymphocytosis and splenomegaly or hepaticomegaly</td>
<td>7 years</td>
<td>25%</td>
</tr>
<tr>
<td>High</td>
<td>III</td>
<td>Lymphocytosis and anaemia (haemoglobin &lt;11 g/dl)</td>
<td>5 years</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>Lymphocytosis and thrombocytopenia (platelets &lt;100 000/μl)</td>
<td>5 years</td>
<td>3%</td>
</tr>
</tbody>
</table>

Table 1.2 Rai classification system for CLL

The modified Rai staging system assigns stage 0 to the low-risk group, stages I and II combined to the intermediate-risk group, and stages III and IV combined to the high-risk this group. It is important to remember that when anaemia or thrombocytopenia is immune in aetiology, the patient should not be assigned to stages III/IV which refer to cytopenias secondary to bone marrow failure rather than autoimmune destruction.

The Binet (1981) staging system consists of three stages (Binet, Auquier et al. 1981). Stages A and B are defined according to the number of palpably enlarged lymphoid areas while stage C is defined by anaemia or thrombocytopenia resulting from bone marrow failure. Binet describes five lymphoid areas: cervical, axillary, and inguinal nodes, and spleen and liver (Table 1.3).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Clinical features at diagnosis</th>
<th>Median survival time</th>
<th>% of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Blood and marrow lymphocytosis and less than three areas of palpable lymphoid involvement</td>
<td>&gt;7-10 years</td>
<td>65%</td>
</tr>
<tr>
<td>B</td>
<td>Same with three or more areas of palpable lymphoid involvement</td>
<td>5-7 years</td>
<td>30%</td>
</tr>
<tr>
<td>C</td>
<td>Same plus anaemia or thrombocytopenia</td>
<td>&lt;2-5 years</td>
<td>5%</td>
</tr>
</tbody>
</table>

Lymph areas include cervical, axillary, inguinofemoral, liver, and spleen.

Table 1.3 Binet classification system for CLL
These staging systems are used to correlate clinical findings with survival times in an effort to manage CLL patients appropriately. Both systems correlate the degree of lymphocytic infiltration with advancing disease stage and shorter survival time. The Rai system is used most often in the USA, while the Binet system is commoner in Europe (Sources: (Cheson 2000; Kipps 2001; Andritsos and Khoury 2002; Keating 2002).

1.2 Conventional Prognostic markers in CLL

1.2.1 Introduction
The design and application of prognostic factors not only allows the physician to predict the natural history of disease progression and the timing of therapeutic intervention, but also facilitates prediction of treatment outcome. The prognostic significance of a given factor depends on its significance after multivariate analysis.

Extensive variation exists in the rate of disease progression and incidence of disease-related complications in patients with CLL. Because of this, the life expectancy of patients with newly diagnosed CLL can vary from 3 to 20 years. Recent advances in our understanding of the biology and natural history of CLL (Dighiero, Travade et al. 1991; Montserrat 1997; Khouri and Keating 1998; Khouri, Keating et al. 1998), allied to earlier diagnosis and the emergence of new therapeutic options (Rozman, Bosch et al. 1997; Diehl, Karnell et al. 1999), have highlighted the need for prognostic assessment to be as accurate as possible at the time of diagnosis.

A number of clinical and laboratory findings have been associated with an adverse prognosis in CLL (Table 1.4).
### Clinical Factors
- Lymphadenopathy
- Splenomegaly
- Hepatomegaly
- Bulky disease
- Poor performance status
- Diffuse bone marrow infiltration
- Anaemia
- Thrombocytopenia
- Atypical morphology

### Laboratory abnormalities
- Rapid doubling time
- Increased Serum lactate dehydrogenase
- Hypoalbuminemia
- Raised serum $\beta_2$-microglobulin

### Phenotype
- CD38
- Aberrant surface phenotypes: strong SmIg, CD5<sup>-</sup>, CD23<sup>-</sup>

### Genetic factor
- del 11q, del 17p, trisomy 12
- Complex cytogenetic abnormalities

### Immunologic factors
- Hypogammaglobulinemia
- Increased serum soluble CD23

### Other factors
- Age
- Sex
- Poor response to therapy

---


#### Table 1.4 Factors associated with poor prognosis

<table>
<thead>
<tr>
<th>Poor response to therapy</th>
</tr>
</thead>
</table>

### 1.2.2 Clinical Staging

The two traditional staging systems for CLL, Rai and Binet, have been described in section 1.1.5. These staging systems are used to correlate clinical findings with survival times in an effort to treat CLL patients appropriately (Cheson 2000; Kipps 2001; Andritsos and Khoury 2002; Keating 2002). Both systems indicate that increasing tumour burden is associated with more advanced disease stage and a shorter survival time. The survival curves of 443 patients with CLL stratified using the modified Rai system (Montserrat and Rozman 1993) indicate that there is a
statistically significant difference between each curve and confirms that patients in the low-risk group have the best overall survival (Fig 1.7). The NCI-WG recommends using the modified RAI staging system for prospective therapeutic trials.

Figure 1.7 Survival according to the modified RAI staging criteria for 443 patients with CLL at the Postgraduate School of Haematology, Barcelona, Spain. Adapted from Monserrat E, Rozman C: chronic lymphocytic leukaemia: Prognostic factors and natural history. Bailleres Clin Haematol 6:349,1993 (Montserrat and Rozman 1993).

A major drawback of these classification schemes is that they do not identify patients with a low tumour load and progressive disease, who despite having stage A disease may have a short overall survival (Hallek, Kuhn-Hallek et al. 1997; Molica 1997; Zwiebel and Cheson 1998).

1.2.3 Marrow Involvement

Bone marrow aspiration and biopsy is not required to diagnose CLL, but does allow evaluation of a major site of disease involvement (Cheson, Bennett et al. 1996). Bone marrow biopsy and bone marrow smears are complimentary methods for assessing bone marrow infiltration by CLL, providing quantitative information which maybe particularly useful for assessing response to therapy (Geisler, Hou-Jensen et al. 1996; Montserrat, Villamor et al. 1996; Molica, Tucci et al. 1997). Bone marrow biopsy shows four patterns of marrow involvement defined as; nodular, interstitial, mixed and diffuse as described in section 1.1.4, which have prognostic significance (Geisler, Hou-Jensen et al. 1996; Montserrat, Villamor et al. 1996). In a retrospective analysis of 329 patients, those with a non-diffuse pattern of infiltration (Interstitial, Nodular or
Mixed) had longer overall survival than patients with a diffuse pattern (Rozman, Montserrat et al. 1984). Conversely, in a prospective analysis of untreated patients, no survival difference was found in patients with stage A and B disease when diffuse and non-diffuse patterns of infiltration were compared (Desablens, Claisse et al. 1989). These contradictory findings indicate that the significance of the pattern of bone marrow infiltration is of marginal prognostic relevance.

A baseline evaluation of the bone marrow is useful prior to therapy to assess response to treatment (Cheson, Bennett et al. 1996). This is probably more relevant now because the recent therapeutic advances in CLL (monoclonal antibodies and stem-cell transplantation), reduce CLL tumour load to lower levels than were previously possible. The ability to detect extremely low levels of CLL has also been demonstrated recently using multi-colour flow cytometry. The successful reduction of CLL to below detectable level and the probability of this translating into prolonged survival indicates the ongoing usefulness of both bone marrow examination and immunophenotyping.

1.2.4 Lymphocyte Morphology and doubling time (LDT)

The presence of atypical lymphocytes such as prolymphocytes, cleaved cells and lymphoplasmacytic cells are associated with a poor prognosis (Peterson, Bloomfield et al. 1980; Melo, Catovsky et al. 1987; Molica and Alberti 1988; Vallespi, Montserrat et al. 1991). The French-American-British (FAB) group recognised 2 atypical forms of CLL: (1) CLL/PLL in which 10%-55% prolymphocytes are detected in peripheral-blood and (2) mixed cell type CLL characterised by a spectrum of small and large lymphocytes, with <10% prolymphocytes (Bennett, Catovsky et al. 1989). Studies classifying patients according to the FAB proposal have shown the association of atypical morphology with advanced clinical stage, presence of trisomy 12 and a shorter survival (Criel, Verhoef et al. 1997).

The lymphocyte count acts as a useful indicator of prognosis over time. Serial plotting of blood lymphocyte counts is a strong prognostic indicator, using the lymphocyte doubling time (LDT) which can be either slow (>12 months) or rapid (<12 months). Galton was the first investigator to suggest that the LDT was a poor prognostic marker (Melo, Catovsky et al. 1987). Other reports have since confirmed that a rapid
increase of the absolute lymphocyte count in the peripheral blood of patients with CLL not receiving cytotoxic therapy is associated with an overall survival of less than five years (Peterson, Bloomfield et al. 1980).

1.2.5 Proliferation marker: Serum Thymidine Kinase (sTK)
Thymidine kinase is an enzyme of the transferase class that catalyzes the ATP-dependent phosphorylation of thymine deoxyribonucleoside, a reaction of pyrimidine salvage. Serum thymidine kinase activity in CLL patients is probably related to the number of dividing neoplastic cells, reflecting tumour mass and rate of tumour cell proliferation. In an interim analysis of the CLL1 trial of the German CLL study group, sTK levels was one of three parameters ($\beta_2m$, sTK and sCD23) which provided the best prediction of a short progression-free survival (Hallek, Langenmayer et al. 1999).

1.2.6 Cellular expression and soluble CD23
CD23 is a functionally relevant surface molecule in B-CLL which is rapidly cleaved from the cell surface into a stable soluble form which can be measured in serum. Higher serum levels of its cleaved form (sCD23), has been shown to indicate a worse prognosis and provide information on overall survival and freedom from progression in stage A patients. In a study of 153 CLL patients, doubling of sCD23 level was associated with a 3.2 fold increased risk of death (Sarfati, Chevret et al. 1996). Its independent prognostic significance has not been confirmed in other studies.

1.2.7 $\beta_2$-microglobulin ($\beta_2m$)
$\beta_2m$ is an extracellular protein that is non-covalently associated with the alpha-chain of the class I major histocompatibility complex (MHC) gene and is detectable in the serum. $\beta_2m$ level at presentation is associated with adverse prognostic features and higher values have been detected in CLL patients with a short survival. In a study of 622 patients at the MD Anderson Cancer Centre, high serum levels of $\beta_2m$ predicted prognosis within individual Rai stages (Keating, Lerner et al. 1995). In a prospective study of 113 patients (which included CLL patients with early stage disease), $\beta_2m$ maintained independent prognostic value in multivariate analysis (Knauf, Langenmayer et al. 1997).
1.2.8 Angiogenesis
The degree of abnormal angiogenesis in bone marrow can be determined by measuring microvessel density. The density appears to be significantly higher in CLL patients compared to normal controls (Peterson and Kini 2001). Higher levels of vascular endothelial growth factor (VEGF), (a potent angiogenic growth factor) have been found in patients with CLL and were considered a reliable indicator of prognosis, especially in patients with low β₂m (Aguayo, Kantarjian et al. 1998). Higher levels of VEGF have also been found to be of predictive value in stage A patients (Molica, Vitelli et al. 1999).

1.2.9 Other biological markers
CLL is characterised primarily by defective programmed cell death (apoptosis), as opposed to increased cell proliferation (Caligaris-Cappio and Hamblin 1999). Defects in the apoptotic pathway play a pivotal role in the pathogenesis of the disease, thus creating a selective survival advantage for the CLL clone. Increased levels of the anti-apoptotic protein bcl-2 have been found in patients with CLL. Over-expression of bcl-2 correlated with decreased survival, suggesting an independent prognostic value (Robertson, Plunkett et al. 1996; Mapara, Bommert et al. 1997). Response to treatment appears to be most affected by the ratio of bcl-2 to Bax (a pro-apoptotic protein), with high ratios associated with resistance to apoptosis in response to drugs such as chlorambucil (Pepper, Thomas et al. 1999). High levels of the anti-apoptotic protein Mcl-1 are also associated with a poor response to both chlorambucil and fludarabine (Kitada, Andersen et al. 1998).

1.3 Treatment of CLL

1.3.1 Goals of treatment
CLL is a complex disease with a heterogeneous clinical course; many patients survive for decades without treatment, whereas others die from their disease within a few months of diagnosis despite appropriate treatment. In the past, it was generally assured that most patients with CLL died of unrelated co-morbid conditions such as ischaemic heart disease. A more recent analysis of data indicates that patients with disease progression died predominantly of complications secondary to CLL such as
infections (Diehl, Karnell et al. 1999). Current CLL therapy does not cure patients, yet may be associated with significant toxicities. Therefore the decision to initiate treatment is important, especially in patients with low-risk disease. Several studies show that early therapeutic intervention does not prolong the survival of patients with early state CLL (Shustik, Mick et al. 1988; 1990; 1991; Dighiero, Maloum et al. 1998; 1999). With the continued development of better prognostic markers and more effective treatment strategies, it becomes imperative to identify patients with high risk early-stage disease. Patients with a good performance status and unfavourable prognostic features can then be offered novel approaches to treatment in the context of clinical trials.

1.3.2 Indications for treatment
There are no proven cures for CLL and treatment of early-stage patients with chemotherapy does not improve survival when compared to a watch and wait approach. The decision to treat patients depends on symptoms and disease progression when appropriate therapy can reduce morbidity and improve survival. It is therefore accepted that patients with stage A disease should be monitored without intervention until there is evidence of disease progression. Most patients with stage B and C disease will require therapy although a minority can simply be monitored without treatment until symptomatic disease develops. If there is doubt about the need for intervention, most clinicians delay therapy and review the patient after 1-2 months of observation. The NCI-WG has published guidelines for the definition of progressive or active disease, and these should be followed in clinical trials so that comparative analysis between studies is possible Table 1.5 (Cheson, Bennett et al. 1988; Cheson, Bennett et al. 1996).
NCI-WG Criteria for initiating Treatment on a Protocol

1. A minimum of any one of the following disease-related symptoms must be present:
   a. Weight loss >10% within the previous 6 months
   b. Extreme fatigue (cannot work or unable to perform usual activities)
   c. Fevers >38°C for >two weeks without evidence of infection
   d. Night sweats without evidence of infection

2. Progressive marrow failure: developing or worsening of thrombocytopenia which is not autoimmune in nature

3. Auto immune anaemia or thrombocytopenia which is poorly responsive to steroids

4. Progression of splenomegaly (>6 cm below the costal margin)

5. Massive nodes or clusters (>10 cm in longest diameter) or progressive Lymphadenopathy

6. Progressive lymphocytosis with an increase of >50% over a two months period or an anticipated LDT of less than six months

Table 1.5 NCI-WG Criteria for initiating Treatment on a Protocol

1.3.3 Currently available treatments

Steroids, alkylating agents and purine analogues are at the mainstay of treatment for patients with CLL. These can be used as single agents or in combination therapy.

**Steroids**

Glucocorticoids are frequently used in the management of CLL as single agents in the treatment of autoimmune haemolytic anaemia and thrombocytopenia or as part of combination chemotherapy protocols. Single agent prednisolone can control the disease temporarily in approximately 10% of patients (Sawitsky, Rai et al. 1977). Partial responses may be achieved by treatment with intravenous high dose methylprednisolone (HDMP) in heavily pre-treated or refractory patients either as a single agent (Thornton, Hamblin et al. 1999) or more recently in combination with Campath-1H (Pettitt, Matutes et al. 2006).

**Alkylating agents: Chlorambucil and Cyclophosphamide**

Single agent chlorambucil has been the most commonly used first line treatment for CLL. It is given in a continuous or intermittent schedule. Responses are attained in approximately 30 to 70% of previously untreated patients, although few of these results in complete remission (Sawitsky, Rai et al. 1977; 1991). Chlorambucil can be administered with prednisolone in an intermittent schedule, however no differences in
response rate or overall survival has been reported when compared to single agent chlorambucil (1999). The combination is associated with an increased incidence of infection. Chlorambucil however, remains an effective first line treatment which is easily administered and well tolerated (Hamblin 2001).

Cyclophosphamide is another alkylating agent, which is as effective as chlorambucil (Huguley 1977) and can be used orally or by intravenous administration. It is associated with more alopecia and less myelosuppression than chlorambucil (Cheson 2000; Hamblin 2001).

**Purine analogues: fludarabine**
The discovery of the activity of the purine analogue, fludarabine monophosphate in the treatment of CLL was a major breakthrough. Fludarabine is considered the most effective single agent drug treatment for CLL. Fludarabine induces complete remissions (CR) in about 30% of previously untreated patients, with an overall response rate greater than 70% (Keating, O’Brien et al. 1993; O’Brien, Kantarjian et al. 1993; Johnson, Smith et al. 1996; Leporrier, Chevret et al. 1997). Although multicentre clinical trials have confirmed the activity of single agent fludarabine in CLL, overall survival was not prolonged, because patients crossed over from chlorambucil to fludarabine when they relapsed (Rai, Peterson et al. 2000). Fludarabine can be given orally or intravenously for five days every four weeks. Side-effects include myelosuppression, immunosuppression, T-cell dysfunction, autoimmune haemolytic anaemia and renal toxicity.

**Combination chemotherapy**
Combination treatment based on alkylating agents and including anthracyclines have been studied as primary treatment. The commoner protocols, CHOP (cyclophosphamide, doxorubicin, vincristine and prednisolone), CAP (cyclophosphamide, doxorubicin and prednisolone) and COP (cyclophosphamide, vincristine and prednisolone) have been compared to chlorambucil ± prednisolone. Results confirm superior response rates for patients treated with combination chemotherapy, but with no added survival benefit (1990; Raphael, Andersen et al. 1991). The meta-analysis of 2022 patients in 10 trials comparing combination therapy,
of them including anthracyclines, with chlorambucil ± prednisolone showed an identical 5-year survival of 48% in both groups (1999).

The combination of cyclophosphamide and fludarabine (FC) was developed following in vitro studies showing synergy and has proven more successful. These agents gave a high response rate (RR) in single institution studies and showed a higher CR rate than fludarabine alone. In a small fraction of patients, this combination has achieved minimal residual disease (MRD) negativity. The ability of FC to induce complete remissions associated with MRD elimination is encouraging. Despite the improved response rates with FC, results from the MRC CLL4 trial have shown that almost half of the patients with 17p deletion (p53 locus) are primary non-responders (NR) or have progressive disease (PD) following this therapy, resulting in poor survival (Catovsky 2004).

**Monoclonal antibodies**

Significant advances in the development of therapeutic monoclonal antibodies have taken place in the last few years. Two antibodies have shown encouraging results in CLL: rituximab, a chimeric human-mouse anti-CD20 antibody, and alemtuzumab (Campath-1H), a humanised anti-CD52 antibody.

Initial studies of rituximab showed marginal activity, with a response rate of only 12% (Nguyen, Amess et al. 1999). This was not surprising, as CD20 is weakly expressed by CLL cells. Efficacy was improved by using the drug at significantly higher concentrations (375mg/m^2 vs. 500 mg/m^2 to 2250 mg/m^2) (O'Brien, Kantarjian et al. 2001) or in combination with fludarabine (Byrd, Peterson et al. 2003). The overall response (OR) and CR rates as well as duration of response were significantly better for combination therapy compared to single agent fludarabine.

Campath-1H targets CD52, an antigen present on >95% of human lymphocytes and is highly expressed in CLL cells. Campath-1H has shown promise in CLL with response rates of 42% in previously treated patients (Osterborg, Dyer et al. 1997). Responses were more marked in blood and bone marrow than in bulky nodal disease. In the large pivotal study (CAM211), patients with fludarabine refractory CLL had a 33% response rate (Keating, Byrd et al. 1999). CD52 is expressed by both B and T
lymphocytes and monocytes and therefore Campath-1H results in significant immunosuppression, with most studies reporting life-threatening cytomegalovirus (CMV) reactivation which require monitoring and prompt therapy.

**Transplantation**

A small subgroup of patients <60yrs with aggressive CLL may benefit from stem cell transplantation. Autologus transplantation is not curative (Dreger, Van Biezen et al. 2000; Esteve, Montserrat et al. 2001; Milligan, Fernandes et al. 2005), though may improve disease-free interval. Furthermore, improvement in survival from combination therapy coupled with the acute and chronic toxicity associated with autologus transplant has made it less popular. Allogeneic transplantation depends on the anti-leukaemic effect of both high-dose chemotherapy and the graft vs. leukaemia effect. It is characterised by a high treatment-related morbidity and mortality (TRM) but can be curative (Michallet, Van Biezen et al. 1999; Pavletic, Arrowsmith et al. 2000). Young patients with refractory disease and limited treatment options should be considered for an allogeneic transplant.

**1.3.4 Currently used treatment strategies**

Treatment strategies can be divided into those used for treatment naive patients and for patients who have failed first-line therapy and also depend on the age and performance status of the patient. Once the decision has been made to initiate treatment, it is important to establish the therapeutic goal for each patient.

Biologically fit Patients <70yrs should be considered for fludarabine alone or in combination with cyclophosphamide with the hope of achieving CR. An elderly frail patient may be more appropriately treated with chlorambucil, obtaining symptom control and a reduction of tumour burden.

It is important to differentiate refractory from relapsing disease. Patients who relapse 2-3 years after the initial therapy can be re-treated using the first-line therapeutic agents whereas patients with primary refractory disease require either an aggressive or palliative approach depending on their age, performance status and co-morbidity. Purine analogue naive patients with prior alkylator exposure should receive a purine analogue alone or in combination with an alkylating agent. Patients who relapse
following a durable response to fludarabine with favourable disease features are likely to respond to re-treatment. Patients who are refractory to purine analogues may benefit from alemtuzumab (Kennedy, Rawstron et al. 2002) ± steroids, especially if they have a deletion 17p (Pettitt, Matutes et al. 2006). Young and biologically fit patients with a sibling donor may be considered for a stem cell transplant using conventional or non-myeloablative conditioning. Patients should be encouraged to participate in clinical trials of novel therapeutic strategies or investigational agents.

1.3.5 Future directions
The aim of any treatment strategy should be an improved overall survival (OS). Traditionally, the goal of therapy in CLL has been palliative, because of the older age of affected patients and the low rate of complete remissions with no improvement in overall survival despite treatment. With increasing knowledge about the biology, molecular genetics and prognostic features of the disease, the philosophy of care for patients with CLL is evolving from palliation to aiming for a potential cure in biologically fit patients. Furthermore, multiple treatment options have emerged, including purine analogues, monoclonal antibodies, and potentially stem cell transplantation. These have been associated with a higher frequencies of CR (O'Brien, Kantarjian et al. 2001) and longer durations of responses compared to conventional chemotherapy. In vitro studies have shown that rituximab sensitizes cell-lines to the cytotoxic effect of fludarabine and cyclophosphamide (Demidem, Lam et al. 1997) and a subset of patients treated with chemo-immunotherapy can achieve a durable CR and molecular remission. This may translate into improved disease-free survival and potentially "cure." Future studies should also incorporate more stringent methods of determination of remission status, such as a MRD analysis using multi-colour flow-cytometry instead of the NCI-WG criteria only.

Allogeneic haematopoietic cell transplantation (HCT) is the only curative treatment for CLL. Until recently, allogeneic HCT utilized myeloablative doses of chemoradiotherapy, which made the treatment unacceptably risky for the majority of CLL patients (Khouri, Keating et al. 2002). The intensity of the conditioning regimen may not be as important as previously thought, particularly in patients with chemosensitive ‘indolent’ lymphoid disorders and many investigators suggest that the allogeneic GVL effect is the most important therapeutic aspect in CLL (Ritgen, Stilgenbauer et al. 2002).
Further proof of a GVL effect is demonstrated by the induction of complete remission of CLL in many patients with persistent disease following the administration of donor leukocyte infusion (Rondon, Giralt et al. 1996).

It has not yet been shown prospectively that treating high-risk CLL patients more aggressively will result in higher rates of long-term survival. Nucleoside analogue combination regimens and non-myeloablative allogeneic transplantation have increased the frequency of complete molecular remissions in CLL, although it is still unclear whether this will result in a corresponding increase in overall survival. These advances in treatment of CLL may make long-term cure a realistic possibility for a disease previously described as incurable. Finally, improvements in diagnostic techniques and prognostic markers may help target therapy to the patients most likely to benefit, while reducing therapy and its complications in patients unlikely to benefit.

1.4 Problems with the conventional model of prognosis

Diagnoses of CLL are becoming more common. With sensitive techniques, a monoclonal population of B lymphocytes indistinguishable from CLL may be found in the blood of 3.5% of people older than 40 years of age (Rawstron, Green et al. 2002). This incidental finding will not have clinical effects for most patients, who will die of unrelated causes, but in some people represents the early stage of a disease with serious consequences for them including early death. The traditional staging systems of Rai and Binet have classically been used to predict disease behaviour and guide therapy. The advantage of these systems and other conventional prognostic markers such as LDT, LDH and $\beta_2$m is that they are easy to use, but do not predict which patients with early stage disease are likely to progress and might benefit from early treatment. Conversely, some patients presenting with advanced stage may have an indolent course and may not require therapy. This is essentially due to the fact that clinical stages are a mere reflection of the biological diversity of the disease (Montserrat 2002; Shanafelt, Geyer et al. 2004).

Over the last few years there have been major advances in the understanding of CLL. The assessment of prognosis in patients with CLL is being revolutionized with the advent of more sensitive immunophenotypic, cytogenetic and molecular techniques.
This should translate into improved risk stratification of patients with CLL at an early stage in the disease and enable physicians to move to better risk-adapted management. The challenge and goal is to develop exact treatment algorithms which integrate useful prognostic markers and guide the physicians in counselling and treatment decisions for individual patients.

1.5 Genomic abnormalities in CLL

1.5.1 Introduction

Acquired genetic abnormalities are a hallmark of cancer cells. Recurring structural and numerical chromosomal changes have been identified in a large and increasing number of haematological malignancies and solid tumours. Many of these abnormalities give useful diagnostic and prognostic information and in acute leukaemia are used to monitor the response of disease to treatment. There has been considerable interest in identifying chromosomal abnormalities of prognostic relevance in CLL (Juliusson and Merup 1998).

1.5.1.1 Banding studies

Chromosomal abnormalities in B-CLL were initially studied by conventional metaphase chromosome analysis using cultured cells. Most CLL cells are G0 phase cells and initial studies using these methods were hampered by the difficulty in getting metaphases suitable for analysis. The discovery of polyclonal B-cell mitogens in the late 1970s paved the way for successful cytogenetic studies in CLL (Gahrton, Zech et al. 1979). Despite this improvement, metaphase spreads were still frequently normal, reflecting the non-leukaemic T-cells present in the samples (Autio, Elonen et al. 1987). Despite improved cell culture techniques, in the early 1990s clonal chromosomal abnormalities were detectable in only 40%-50% of B-CLL cases (Han, Ozer et al. 1984; Crossen 1989; Juliusson, Gahrton et al. 1991; Losada, Wessman et al. 1991).

1.5.1.2 Interphase cytogenetics

Over recent years, the development of molecular cytogenetic techniques such as FISH (interphase cytogenetics using fluorescence in-situ hybridisation) has increased the
sensitivity for detecting translocations, deletions, or numerical abnormalities (Lichter and Ward 1990; Lichter, Bentz et al. 1995). Specific DNA sequences in interphase cells or metaphase spreads are detected by FISH using cloned DNA fragments, which are then visualised by fluorescence microscopy. FISH has two advantages over conventional cytogenetics: 1) it allows the detection of specific chromosomal lesions in non-dividing cells, which would not be detected by metaphase analysis and 2) the resolution of the detection of chromosomal abnormalities by FISH is superior to that of conventional chromosome banding, allowing for example the detection of deletions in chromosome at the 100 kilobases level. The main disadvantages are that one can only screen for known abnormalities, a limited number of probes can be combined in a single hybridization experiment and additional chromosomal abnormalities or clonal evolution will be missed.

1.5.2 Incidence of chromosomal abnormalities in CLL
Using FISH, abnormalities can be detected in more than 80% of patients using a 4-probe panel for the detection of trisomy 12q13 and deletions 13q14, 17p13 and 11q22 (Dohner, Stilgenbauer et al. 2000) Table 1.6.

<table>
<thead>
<tr>
<th>karyotype</th>
<th>Frequency %</th>
</tr>
</thead>
<tbody>
<tr>
<td>trisomy 12</td>
<td>16</td>
</tr>
<tr>
<td>13q abnormalities</td>
<td>55</td>
</tr>
<tr>
<td>11q abnormalities</td>
<td>18</td>
</tr>
<tr>
<td>17p abnormalities</td>
<td>7</td>
</tr>
<tr>
<td>Norman karyotype</td>
<td>18</td>
</tr>
</tbody>
</table>

Table 1.6 Frequency of chromosomal abnormalities in CLL using FISH

An additional 10% of patients can be shown to carry a 6q21 deletion, 14q32 translocation and partial trisomy 3q or 8q.
1.5.3 Chromosome 13

Deletions on the long arm of chromosome 13 are the most common genetic abnormality in CLL, first described by Fitchett and colleagues (Fitchett, Griffiths et al. 1987). Both homozygous or hemizygous loss at 13q14 occurs in more than half of B-CLL cases. It is generally thought that deletions at 13q14 result in inactivation of a tumour suppressor gene following the Knudson hypothesis (i.e., through loss of one allele and inactivation of the remaining allele). Inactivation of the retinoblastoma gene (RB1) located at 13q14, was initially considered a candidate gene, however later studies have indicated that a gene region distinct from RB1 is affected in B-CLL (Brown, Ross et al. 1993; Hawthorn, Chapman et al. 1993; Liu, Szekely et al. 1993). Positional cloning has been used by several groups to identify the gene or genes affected by the deletions. A region of more than 1 Mb has been fully sequenced and characterized in detail (Stilgenbauer, Nickolenko et al. 1998), which includes a total of eight genes. These have been identified and screened for alterations at the DNA and/or RNA level in sporadic and familial cases of CLL: Leu-1, Leu-2 (ALT1), Leu-5 (CAR), CLLD6, KPNAS3, CLLD7, LOC51131 (putative zinc finger protein NY-REN-34 antigen) and CLLD8. Despite, detailed genetic analysis, including loss of heterozygosity (LOH), mutation and expression studies, the consistent involvement of any of these genes in the deleted region has not been demonstrated (Stilgenbauer, Nickolenko et al. 1998; Bullrich, Fujii et al. 2001; Mabuchi, Fujii et al. 2001; Wolf, Mertens et al. 2001; Hammarsund, Corcoran et al. 2004). Recent studies have shown that microRNAs (miRNAs) are the main candidates for the elusive class of CLL predisposing genes on chromosome 13(Nicoloso, Kipps et al. 2007). MicroRNAs are a group of small non-coding RNA molecules that are associated with transcriptional repression. Cases with good prognostic features typically are characterized by miRNA down-regulation of genes miR-15a and miR-16-1, located at 13q14.3. Both miRNAs negatively regulate bcl2 at a post-transcriptional level (Calin, Pekarsky et al. 2007).

1.5.4 Chromosome 12

Trisomy 12 was reported as the first recurrent chromosome aberration of B-CLL. It occurs in up to 16% of B-CLL cases and is associated with aggressive disease and atypical morphology (Juliusson, Oscier et al. 1990; Garcia-Marco, Price et al. 1997; Dohner, Stilgenbauer et al. 1999). Trisomy results from duplication of one chromosome rather than loss of one chromosome and triplication of the other
(Einhorn, Burvall et al. 1989). Other studies have also shown partial trisomies or duplications involving the 12q12-22 region. Chromosome 12 contains a large number of genes directly or indirectly involved in human cancer, however the exact molecular mechanism by which trisomy 12 contributes to leukaemogenesis in B-CLL is unknown. Two candidate genes (MDM2 and HMG1-C) have been identified in isolated cases which might lead to an understanding of the exact pathogenetic role of trisomy 12. The over-expression or amplification of the MDM2 (mouse double minute 2) gene and the resultant protein Mdm2 may play a role in CLL pathogenesis. The Mdm2 protein normally binds to p53 and inhibits its function by promoting p53 degradation by the proteosome (Kubbutat, Jones et al. 1997). Impaired regulation will result in impaired control of DNA damage, cell cycle progression and cell death. HMG1-C gene, a member of the high mobility group (HMG) family, is located at 12q13 and is also a candidate for involvement in trisomy 12. HMG1-C is involved in the development of benign tumours of mesenchymal origin (Cooper 1996). The protein binds DNA and is a structural factor indirectly involved in transcription (Wolffe 1994). Rearrangement of the HMG1-C gene resulting from a translocation involving chromosome 12q13 in Richter's syndrome has been reported, suggesting that it may play a role in B-CLL pathogenesis (Santulli, Kazmierczak et al. 2000).

1.5.5 Chromosome 11
Deletions involving chromosome 11 in CLL occur in up to 20% of cases, and is the second most common genetic abnormality found. Like 13q14 deletion, 11q loss is thought to result in inactivation of a tumour suppression gene. The ATM gene at 11q 22-23 is located in the minimally deleted region described in B-CLL (Stilgenbauer, Liebisch et al. 1996; Dohner, Stilgenbauer et al. 1997; Neilson, Auer et al. 1997). The ATM protein is involved in DNA repair pathways. The kinase activity of ATM is increased several fold in response to ionising radiation and binds and phosphorylates several proteins in the apoptotic pathway such as p53, c-Abl, BRCA1, Nbs1 and Chk2 kinase (Baskaran, Wood et al. 1997; Shafman, Khanna et al. 1997; Banin, Moyal et al. 1998; Canman, Lim et al. 1998; Matsuoka, Huang et al. 1998; Cortez, Wang et al. 1999; Lim, Kim et al. 2000; Wu, Ranganathan et al. 2000). The pattern of ATM inactivation in B-CLL (a pattern also observed in T-cell tumours), is the classic tumour suppressor inactivation mechanism involving loss of one allele and mutation of the remaining allele. It is likely that allelic loss at 11q in B-CLL uncovers the
ATM associated mutator phenotype which initiates or accelerates leukaemic progression by the accumulation of genetic abnormalities.

1.5.6 Chromosome 17

Deletion of the short arm of chromosome 17 occurs in up to 17% of B-CLL cases and is accompanied by mutation of the p53 tumour suppressor gene located at 17p13 (el Rouby, Thomas et al. 1993; Silber, Degar et al. 1994; Shiloh 1997). The p53 gene encodes a 53kd nuclear phosphoprotein (Zambetti and Levine 1993). p53 plays a crucial role in controlling the cells response to DNA damage such as ionising radiation (Rotman and Shiloh 1999; Sionov and Haupt 1999; Khanna 2000). On exposure to DNA-damaging agents, the protein is stabilized and activated through post-translational modifications which both increases its half life and enhances its sequence-specific, DNA-binding transcriptional activity (Rotman and Shiloh 1999; Sionov and Haupt 1999; Khanna 2000). p53 induces expression of p21 and the 14-3-3 protein, which are both involved in cell-cycle control (Bullrich, MacLachlan et al. 1995) and also induces the expression of several pro-apoptotic molecules including Fas/Apo1/CD95 and the bax protein (Sionov and Haupt 1999).

1.5.7 Clinical significance of chromosomal abnormalities

<table>
<thead>
<tr>
<th>Karyotype</th>
<th>Median Survival (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17p abnormalities</td>
<td>32</td>
</tr>
<tr>
<td>11q abnormalities</td>
<td>79</td>
</tr>
<tr>
<td>Normal Karyotype</td>
<td>111</td>
</tr>
<tr>
<td>Trisomy 12</td>
<td>114</td>
</tr>
<tr>
<td>13q abnormalities</td>
<td>133</td>
</tr>
</tbody>
</table>

Table 1.7 Cytogenetic abnormalities and prognosis

Using interphase cytogenetics, Dohner et al, analysed 325 patients with CLL (Dohner, Stilgenbauer et al. 2000). A hierarchical survival model was used to assign patients to
particular risk categories. 82% of the patients had abnormal FISH profiles (Table 1.6). Correlation with clinical data showed that 17p and 11q deletions were associated with more advanced and aggressive disease and a median survival of 32 months and 79 months respectively. The longest survival of 133 months was determined in patients who had a sole 13q deletion (Table 1.7). Patients with 11q and 17p deletions were more likely than the other subgroups to have fever, night sweats or weight loss and have lower haemoglobin and platelet counts. Multivariate analysis confirmed that 17p deletion was the most significant negative prognostic factor for survival. Patients with 17p deletions are resistant to treatment with purine analogues and alkylating agents (Dohner, Fischer et al. 1995).

1.6 Immunoglobulin Variable region Gene (IgV_{H}) characteristics in CLL

Immune cells are derived from haematopoietic stem cells in the bone marrow, circulate in the blood and lymph, form specialised lymphoid organs, and infiltrate virtually every tissue in the body. The immune system in humans is made up of two components with an innate and adaptive function. Innate immunity includes physical barriers, such as the epithelial layers, antimicrobial substances secreted at these surfaces and cells including monocytes, macrophages, dendritic cells, eosinophils, basophils, neutrophils and mast cells and act as a first line of defence. As a second line of defence, the adaptive immune system relates to the ability of the system to adapt to antigenic challenge and so is also called acquired or specific immunity. It has unique characteristics, such as specificity, diversity, memory, specialisation and tolerance. Immune specificity relies on two major cell types, B and T lymphocytes, their respective receptors and the antibodies produced by the B lymphocytes. The diversity of the adaptive immune system is not inherited but is acquired and can create a repertoire of $>10^{11}$ different clones of B-cells, each bearing a single specific receptor, resulting in lymphocytes which can recognise virtually every imaginable antigen. The adaptive immune system is critically dependent on B-lymphocytes ability to produce diverse immunoglobulins.
1.6.1 B-Cell development

B-cells develop from pluripotent haematopoietic stem cells in the bone marrow and mature in a highly controlled sequence (Rajewsky 1996). It can be divided into two stages: an antigen-independent stage which occurs in fetal liver and marrow and adult marrow and an antigen-dependent stage which occurs primarily in secondary lymphoid tissue, such as the spleen and lymph nodes. Primary B-cell development occurs by sequential rearrangement and expression of heavy and light-chain immunoglobulin genes and the expression of other cell-surface proteins. The cell surface immunoglobulin (Ig) and the associated signalling molecules are referred to as the BCR (B-cell receptor).

1.6.2 Immunoglobulin (Ig) gene rearrangement and expression

The immunoglobulin molecule consists of two heavy and two light chains, each with a constant (C) and a variable (V) region (Fig 1.8).

![Figure 1.8 Structure of an antibody molecule.](image)

The immunoglobulin genes are inherited in three unlinked gene complexes: one for the heavy-chain classes, one for κ light chains and one for the λ light chains. The
heavy-chain genes are located on chromosome 14q32, \( \kappa \) light chain genes on chromosome 2p12 and \( \lambda \) light chain genes on chromosome 22q11. The heavy-chain gene complex itself consists of distinct regions. There are 51 functional heavy-chain variable-region genes (V\(_H\) genes), 27 functional diversity (D\(_H\)) segments, 6 functional joining region (J\(_H\)) genes and 5 C-region genes (Figure 1.9). The V\(_H\), D\(_H\) and J\(_H\) genes code for the variable region of the immunoglobulin molecule and the C\(_H\) genes encode the constant region of each immunoglobulin heavy-chain isotypes (Tomlinson, Cook et al. 1995; Matsuda and Honjo 1996; Kipps 1997).

Figure 1.9 The germline organization of the immunoglobulin heavy- and light-chain loci in the human genome. The heavy-chain locus (chromosome 14) has approximately 65 functional V\(_H\) gene segments and a cluster of around 27 D\(_H\) segments lying between these V\(_H\) gene segments and six J\(_H\) gene segments. The heavy-chain locus also contains a large cluster of C\(_H\) genes. The genetic locus for the \( \kappa \) light chain (chromosome 22) has about 30 functional V\(_\kappa\) gene segments and four pairs of functional J\(_\kappa\) gene segments and C\(_\kappa\) genes. The \( \lambda \) locus (chromosome 2) is organized in a similar way, with about 40 functional V\(_\lambda\) gene segments accompanied by a cluster of five J\(_\lambda\) gene segments but with a single C\(_\lambda\) gene. Adapted from, Immunobiology: the immune system in health and disease / Charles A. Janeway. 5th ed. 2001 Garland Publishing, New York (Janeway, Travers et al. 2001)

The Variable regions differ between the many clones of antibodies produced. Sequence variability is not however, distributed evenly throughout the V regions but is concentrated in certain segments of the V region. Three segments of particular variability can be identified in both the V\(_H\) and V\(_L\) domains. They are designated hypervariable regions and are denoted HV1, HV2, and HV3. The most variable part of the domain is in the HV3 region. The regions between the hypervariable regions,
which comprise the rest of the V domain, show less variability and are termed the framework regions. There are four such regions in each V domain, designated FR1, FR2, FR3, and FR4. The three hypervariable loops determine antigen specificity by forming a surface complementary to the antigen, and are more commonly termed the complementarity-determining regions, or CDRs (CDR1, CDR2, and CDR3), while the framework regions (FR1-FR4) provide a protein scaffold for the hypervariable regions.

During B-cell development, the somatic DNA rearrangements which lead to formation of gene segments encoding a complete immunoglobulin molecule within an individual B-cell, proceed in a defined order, starting first within the heavy-chain gene complex (Tonegawa 1983; Alt, Oltz et al. 1992). The initial event during heavy-chain gene rearrangement juxtaposes a D<sub>H</sub> region segment to a J<sub>H</sub> segment (Figure 1.10). Although in theory any D<sub>H</sub> region gene can join with equal frequency to any J<sub>H</sub> region gene, there may be preferential utilisation of selected D<sub>H</sub> and J<sub>H</sub> region genes at various stages of lymphocyte development (Souto-Carneiro, Longo et al. 2004). This combination event results in a DNA fragment which will determine the third complementarity determining region (CDR) and the fourth framework region (FR) of the variable region of the heavy-chain. Following successful D<sub>H</sub>-J<sub>H</sub> recombination, one of 51 V<sub>H</sub> region genes rearranges to the D<sub>H</sub>-J<sub>H</sub> complex, a sequence which codes for the first, second and third FR and the first and second CDR region of the variable region of the immunoglobulin heavy chain molecule. Depending on the expression of the enzyme terminal deoxynucleotidyl transferase (TdT), non-template encoded nucleotides, or N-nucleotides are added at the junctional sites resulting in further
Figure 1.10 V-region genes are constructed from gene segments. Heavy-chain V regions are constructed from three gene segments. First, the diversity (D) and J gene segments join, then the V gene segment joins to the combined DJ sequence, forming a complete V\textsubscript{H} exon. A heavy-chain C-region gene is encoded by several exons. The C-region exons, together with the leader sequence, are spliced to the V-domain sequence during processing of the heavy-chain RNA transcript. The leader sequence is removed after translation and the disulfide bonds that link the polypeptide chains are formed. The hinge region is shown in purple. Immunobiology: the immune system in health and disease / Charles A. Janeway. 5th ed. 2001 Garland Publishing, New York (Janeway, Travers et al. 2001).

diversification of the gene sequence. The heavy-chain constant region which remains separated from the rearranged V\textsubscript{H}D\textsubscript{H}J\textsubscript{H} by an intron is subsequently joined to it during RNA processing. Following translation, the \( \mu \) heavy-chain protein is expressed in the cytoplasm of the pre-B cells (Figure 1.11) (Schatz, Oettinger et al. 1992; Lewis 1994; Weaver, Boubnov et al. 1995). Each B-cell has two immunoglobulin heavy-chain genes, but only one of these encodes the functional protein. This process by which the second gene is excluded is known as allelic exclusion (Schatz, Oettinger et al. 1989).

Following completion of a functional \( \mu \) heavy-chain rearrangement, recombination at the light-chain locus proceeds in a similar fashion (Alt, Oltz et al.
1992). This usually begins with germline transcription of the kappa (κ) locus, followed by the rearrangement of a Vκ to Jκ. If the initial VJκ sequence is functional, the light-chain locus also undergoes allelic exclusion. If not, the second κ allele rearranges. If recombination at both κ alleles fails to result in a functional rearrangement, the lambda (λ) light-chain locus will rearrange (Alt, Oltz et al. 1992). Expression of two heavily conserved proteins RAG-1 and RAG-2 is required for a heavy and light chain gene recombination (Schatz, Oettinger et al. 1989; Oettinger, Schatz et al. 1990). RAG (Recombinase-activating genes), proteins are involved in cleavage of DNA at recombination signal sequences.

1.6.3 Stages in B-cell development: Expression of Ig chains and cell-surface proteins

The earliest B-lineage cells are known as pro-B cells and are progenitor cells with limited self-renewal capacity (Figure 1.11). They are derived from pluripotent haematopoietic stem cells and are identified by the appearance of cell-surface proteins CD34, CD10 and CD19 characteristic of early B-lineage cells (Loken, Shah et al. 1987). Cells at this stage of development have initiated Ig heavy-chain gene rearrangements. Once the Ig heavy-chain gene has undergone productive rearrangement and is expressed, mu (μ) heavy chain protein is detected in the cytoplasm of the cell which now enters the pre-B cell stage. CD19 and CD10 are present on the pre-B cell surface but CD34 is no longer expressed. The μ chain in pre-B cells expressed intracellularly in combination with a surrogate light chain forms the pre-B-cell receptor. Expression of the pre-B-cell receptor signals the cessation of heavy-chain locus rearrangement and production of the surrogate light chain. Productive rearrangement and expression of an Ig light chain gene result in maturation to the surface IgM expressing immature B-cell stage of development.
A number of additional cell-surface proteins are expressed on the developing and mature B-cells including CD20, CD21, CD22, CD24, and CD40 (Clark and Lane 1991). This development takes place in the bone marrow and is independent of antigen and T-cell help. These ‘naïve’ B cells recirculate through peripheral lymphoid issues where they may encounter and be activated by an appropriate antigen (MacLennan and Chan 1993).

### 1.6.4 Isotype Class Switching

During B cell development, many B cells switch from making one class of antibody to making another, a process called class switching. All B cells initially synthesize IgM antibodies. After the B cells leave the bone marrow but before they interact with
antigen, IgM and IgD molecules are expressed as membrane-bound antigen receptors with the same antigen-binding sites. On stimulation by antigen and helper T cells, some of these cells are activated to secrete IgM antibodies, which dominate the primary antibody response. Later in the immune response, the combination of antigen and the cytokines, which helper T cells secrete, induce many B cells to switch to making IgG, IgE, or IgA antibodies. These cells generate both memory cells which express the corresponding classes of antibody molecules on their surface and effector cells which secrete antibodies. The IgG, IgE, and IgA molecules are collectively referred to as secondary classes of antibodies, both because they are produced only after antigen stimulation and because they dominate secondary antibody responses. Class switching entails deletion of all the C\textsubscript{H}-coding sequences between the assembled VDJ-coding sequence and the particular C\textsubscript{H}-coding sequence which the cell is destined to express (Figure 1.12).

Figure 1.12 Antibody class switching. A B cell making an IgM antibody from an assembled VDJ DNA sequence is stimulated by antigen and the cytokines made by helper T cells to switch to making an IgA antibody. In the process, it deletes the DNA between the VDJ sequence and the C\textsubscript{\alpha}-coding sequence. Specific DNA sequences (switch sequences) located upstream of each C\textsubscript{H}-coding sequence recombine with each other to delete the intervening DNA. Class switch recombination is thought to be mediated by a switch recombinase and is directed to the appropriate switch sequences which become accessible under the influence of cytokines (Alberts, Johnson et al. 2002). Immunobiology: the immune system in health and disease / Charles A. Janeway. 5th ed. 2001 Garland Publishing, New York
Isotypic switching and further physiological maturation in the immunoglobulin genes, known as somatic hypermutation takes place in the peripheral lymphoid organs as shown in Figure 1.13.

**FIGURE 1.13 B-cell development and B-cell differentiation.** The early stages of B-cell development occur in the bone marrow, with cells progressing along a developmental program determined by the rearrangement and expression of immunoglobulin genes. Immature B cells with receptors for multivalent self-antigens die in the bone marrow. Surviving B cells co-express IgD and IgM surface receptors, circulate into peripheral lymphoid organs, where they home to selected locations and receive signals to survive and become longer lived naive B cells. Antigen-binding B cells and antigen-presenting B cells interact with antigen-specific T cells and are activated through membrane-bound and secreted molecules. Activated B cells migrate into the lymphoid follicles, leading to the formation of germinal centres. B cells in germinal centres undergo somatic hypermutation of immunoglobulin genes; cells with high affinity for the antigen presented on the surface of follicular dendritic cells are selected to differentiate into either memory B cells or plasma cells. © 2004, 2000 Elsevier Inc.

### 1.6.5 Somatic Hypermutation

Several mechanisms contribute to the generation of antibody diversity (Figure 1.12)( Tonegawa 1993). These are (1) the presence in the germline of multiple different V, D, and J gene segments, (2) the random joining of these DNA segments, (3) uncorrected substitutions made during the recombination process, (4) addition of
non-template encoded nucleotides at the junctional sites and, (5) coming together of heavy and light chains to form a polypeptide antibody molecule. Further gene sequence diversity occurs in the germinal centres in the presence of T-cells and antigen, with the acquisition of further mutations in the rearranged DNA sequences in a process known as affinity maturation or somatic hypermutation (Wabl, Cascalho et al. 1999).

The somatic hypermutation results in the selection of B-cells estimated to have a 10-fold increase in antigen-binding potential (Huchet and Feldmann 1973; Berek 1992). During this process, single-nucleotide exchanges, deletions, and mutations are introduced into the genes encoding the antibody-binding regions of the immunoglobulin-receptor. It is dependent on signals delivered to the antigen responsive B-cells by antigen specific T-lymphocytes in the germinal centres of the lymph nodes (Bowen, Butch et al. 1991). T-cells mediate their effect on B-cells by secreting cytokines as well as through direct intercellular contacts. Hypermutation begins at the 5' end of the rearranged V genes, downstream of the transcription initiation site and continue through the V gene and into the 3'-flanking region before tapering off. The mutations are therefore clustered in the region spanning from 300 bp 5' of the variable (V) region exon to approximately 1 kb 3' of the rearranged J segment. More mutations take place in the CDR regions compared to the PR regions. Somatic hypermutation is highly efficient at increasing antigen-binding efficiency but predisposes lymphocytes to faulty DNA joins thus leading to the common translocations seen in B-cell malignancies. These occur as recombination errors at the time of antigen receptor gene rearrangement, frequently resulting in juxtaposition of an oncogene to the enhancer region of the antigen receptor gene. The up-regulated oncogene can result in uncontrolled cell proliferation or in failure of cells to undergo apoptosis.

1.6.6 Characteristics of Ig variable region genes in B-CLL cells

B-CLL was thought to arise from antigen naïve cells thus having V genes which would not be expected to have undergone somatic mutation (Kipps, Tomhave et al. 1988; Meeker, Grimaldi et al. 1988; Pratt, Rassenti et al. 1989; Kuppers, Gause et al. 1991; Friedman, Moore et al. 1992; Wagner and Luzzatto 1993; Pan, Diss et al. 1996). We now know that up to 50% of randomly chosen B-CLL cases have
significant V_H gene mutations (>2% differences from germline genes) (Fais, Ghiotto et al. 1998). CLL can be therefore divided into two subgroups (mutated with >2% and unmutated with ≤2% mutations) on the basis of V_H gene mutations pattern (Hamblin, Davis et al. 1999).

Normal CD5+ B-lymphocytes use genes of the various V_H families in the following order: V_H3 > V_H4 = V_H5 > V_H2 > V_H7 (Brezinschek, Foster et al. 1997). Analysis of V_H gene family usage in CLL shows that the pattern of gene usage is different to the normal B cell repertoire pattern Figure 1.14 (Shen, Humphries et al. 1987; Kipps, Tomhave et al. 1989; Spatz, Wong et al. 1990; Friedman, Cho et al. 1991; Cherepakhin, Baird et al. 1993; Korganow, Martin et al. 1994; Schroeder and Dighiero 1994; Ikematsu, Ikematsu et al. 1995; Matolcsy, Casali et al. 1997; Fais, Ghiotto et al. 1998; Hamblin, Davis et al. 1999).

Figure 1.14 Comparison of V_H gene use between B-CLL cases and normal blood B cells. Statistical comparisons were performed using the Fisher's Exact test. (Adapted from Fais et al, J Clin Invest 102:1515-1525, 1998).

In B-CLL V_H1 family genes are expressed more frequently and V_H3 family genes are expressed less frequently than in normal CD5+ B lymphocytes. In addition, there is a
significant over-representation of two individual genes among B-CLL cases, $V_H$ 1-69 and $V_H$ 4-34 Figure 1.15 (Fais, Ghiotto et al. 1998).

Figure 1.15 Comparison of specific $V_H$ gene use between IgM$^+$ B-CLL cases and normal blood B-cells. Statistical comparisons were performed using the Fisher's Exact test: (1) $P = 0.0011$; (2) $P = 0.0394$; (3) $P = 0.0230$; and (4) $P = 0.0397$. (Fais, Ghiotto et al. 1998).

1.6.7 Clinical Significance

Risk stratification based on the IgV\textsubscript{H} mutation status appears to be one of the most important prognostic factors in CLL. It correlates inversely with disease course and survival (Damle, Wasil et al. 1999; Hamblin, Davis et al. 1999).

Patients with B-CLL which express unmutated IgV\textsubscript{H} genes (<2% mutations) experience a more aggressive clinical course and shorter survival than those patients with mutated IgV\textsubscript{H} genes (≥2% mutations). In a study by Hamblin et al (Hamblin, Davis et al. 1999) the median survival for patients in the unmutated group was 117 months (9.7 years) and for the mutated group 293 months (24.4 years) Figure 1.16.

Figure 1.16 Kaplan-Meier survival curve comparing CLL patients with mutated and unmutated V\textsubscript{H} genes. Median survival for unmutated CLL: 117 months; median survival for mutated CLL: 293 months. The difference is significant at the $P = .001$ level (log-rank test).
The median survival for Binet stage A patients with unmutated IgV\textsubscript{H} genes was 95 months (7.9 years) compared with 293 months (24.4 years) for patients with the mutated group (Figure 1.17). The stage A patients in the unmutated group progressed to Binet stages B and C more frequently and rapidly than the mutated group.

![Kaplan-Meier survival curve comparing stage A CLL patients with mutated and unmutated V\textsubscript{H} genes. Median survival for unmutated CLL: 95 months; median survival for mutated CLL: 293 months. The difference is significant at the \( P = .0008 \) level (log-rank test).](https://example.com/kaplan-meiwer.png)

**Figure 1.17** Kaplan-Meier survival curve comparing stage A CLL patients with mutated and unmutated V\textsubscript{H} genes. Median survival for unmutated CLL: 95 months; median survival for mutated CLL: 293 months. The difference is significant at the \( P = .0008 \) level (log-rank test).

The mutational status therefore divides patients into two distinct groups with very different disease outcome. This has important implications for patients with early stage but aggressive disease where a risk-adapted long-term management plan can be mapped out on the basis of their likely poor prognosis.
1.7 Advances in the Immunophenotypic Characteristics of B-CLL

CLL has been characteristically considered to be a single entity resulting from an immune incompetent (Dameshek 1967) and accumulative CD5+ B-lymphocyte clone (Caligaris-Cappio, Gobbi et al. 1982). Recent immunophenotypic data indicates that B-CLL cells are heterogeneous and are not derived from immature immunoincompetent cells. This knowledge can be used to explain the clinical heterogeneity seen in patients with CLL.

1.7.1 General Immunophenotypic features of B-CLL cells

CLL cells are characterised by a distinct phenotype, are poorly responsive to exogenous stimuli and have defective apoptotic mechanisms. All the membrane antigens typical of mature B-cells, such as CD19, CD20, CD21, and CD24 are expressed by CLL cells, but the additional presence of CD5, CD23, and low to undetectable sIg are typical of CLL (Caligaris-Cappio and Hamblin 1999). Based on the co-expression of CD5, CD11b and IgM, B-cells in humans are divided into different subsets, termed B-1a and B-1b, each of which express CD11b, and B2 which are CD11b negative. B-1a cells express CD5 while B-1b cells are negative (Kantor 1991). Although several characteristics of normal B-1a cells are shared by malignant CLL cells (Caligaris-Cappio 1996), such as low CD20 expression (Antin, Emerson et al. 1986), the ability to form rosettes with mouse erythrocytes (Stathopoulos and Elliott 1974) and the production of poly-reactive natural auto-antibodies, there are also significant differences. Normal CD5+ B-cells do not have the low to undetectable levels of sIg so typical of B-CLL cells. It is difficult to EBV-transform B-CLL cells (Rickinson, Finerty et al. 1982), whereas this is readily achievable in normal CD5+ B-cells. CD5+ B-cells may constitute a distinct lineage in mice, but there is no definitive evidence for this in humans. The possibility that CD5 is an activation marker has been suggested; (Youinou, Jamin et al. 1999) B-CLL cells appear quiescent but there are some indications that activation has occurred. Cyclin D2 levels are elevated in both normal CD5+ B-cells and B-CLL cells (Delmer, Ajchenbaum-Cymbalista et al. 1995), suggesting an attempt to enter the cell-cycle. B-CLL cells express significantly higher percentages of CD25, CD69 and CD71 and
recently has been shown to have CD40 ligand expression, again indicative of cellular activation (Schattner, Mascarenhas et al. 1998; Damle, Fais et al. 2000; Damle, Yan et al. 2000). Another immunophenotypic feature of B-CLL cells is the lack of detectable CD79b, which may be due to alternative splicing of the gene (Alfarano, Circosta et al. 1999). The levels of the spliced variant are higher in activated normal B-cells compared to resting B-cells and *in vivo* activated germinal centre B-cells which are CD79b negative. This raises the possibility that B-cells physiologically choose alternative splicing of CD79b to down-regulate B-cell-receptor (BCR) expression on activation; therefore the lack of CD79b in B-CLL is also indicative of an activated state. The immunophenotypic evidence indicates that the majority of CLL cells are activated despite the fact that more than 99% of circulating CLL cells are in the G0/early G1 phase of the cell-cycle.

### 1.7.2 CD38 and CLL

Human CD38 is a member of a family of proteins which share a long evolutionary history as well as common functional traits (Deaglio, Mehta et al. 2001). CD38 has evolved from a soluble enzyme regulating calcium homeostasis in molluscs to a complex cell surface glycoprotein in mammals, retaining its peculiar catalytic activities as an ectoenzyme while acquiring new receptor functions (Deaglio and Malavasi 2002). CD38 expression is tightly regulated during B-cell ontogeny and is present at high levels in bone marrow (BM) precursors, is down-regulated in resting normal B-cells and then expressed again in terminally differentiated plasma cells (Campana, Suzuki et al. 2000). The changing expression of CD38 in the B-cell compartment is associated with significant variation in its functional properties; it blocks lymphopoiesis in bone marrow (Kumagai, Coustan-Smith et al. 1995) yet rescues germinal centre B-cells from apoptosis (Zupo, Rugari et al. 1994). This variation is probably secondary to different ligands in the two distinct microenvironments. CD31/platelet endothelial cell adhesion molecule 1 (PECAM 1) is the only reported cell surface bound ligand for CD38 and has been shown that CD31/CD38 interactions control an active signalling pathway in lymphocytes (Deaglio, Morra et al. 1998; Deaglio, Mallone et al. 2000). Recent data also indicates that the qualitative and quantitative composition of cell surface receptors can influence both the signalling potential of CD38 and the type of signal delivered. CD38
is known to act as an accessory molecule in BCR-mediated signal transduction (Santos-Argumedo, Teixeira et al. 1993; Lund, Yu et al. 1996).

CD38 has been studied in order to define the activation and maturation states of CLL cases (Damle, Fais et al. 2000; Damle, Yan et al. 2000). The data from these studies suggest that CLL can be divided into two groups on the basis of surface membrane phenotype i.e., expression of CD38 (Damle, Wasil et al. 1999). In one group, few (<30%) CLL cells expressed CD38, whereas in a second group the percentage was higher and approached 100%. The two groups also differed in the expression of other activation markers such as CD69 and CD71. CD38 expression affects B-cell function by enhancing its ability to be triggered through the BCR (Zupo, Isnardi et al. 1996). From a prognostic viewpoint, the most important characteristic is the ability of CD38 expression to divide CLL cases into two distinct clinical subgroups with different survival (Damle, Wasil et al. 1999; Hamblin, Orchard et al. 2000). The median survival of patients with intermediate Rai stage, with 30% or more cells expressing CD38 was 10 years, whereas no patients with less than 30% of CD38 cells died during the follow up period (Figure 1.18).

![Survival based on CD38 expression among B-CLL patients within the Rai intermediate risk category](image)

**Figure 1.18** Survival based on CD38 expression among B-CLL patients within the Rai intermediate risk category (Damle, Wasil et al. 1999).
CD38 expression also correlates with the mutational status of the IgVH genes; cells with a higher CD38 expression are more likely to have unmutated genes (Damle, Wasil et al. 1999). Therefore CD38 status is a useful prognostic marker in CLL, however its use in defining a management plan remains unclear. The 30% cut-off used to classify patients as CD38+ or CD38- is an arbitrary threshold. It is important to look at other cut-off values and the pattern of CD38 expression with regards to its effect on clinical behaviour.

1.7.3 CLL and ZAP-70

Genome based technologies are now used to look in a more comprehensive way at genes involved in human cancer. In CLL, several groups have performed gene expression profiling using DNA microarrays for simultaneous analysis of the expression of multiple genes (Klein, Tu et al. 2001; Rosenwald, Alizadeh et al. 2001). These studies found a gene signature common to all CLL cells independent of the IgVH mutation status, implying that there is a common cell of origin and/or mechanism leading to malignant transformation (Klein, Tu et al. 2001; Durig, Nuckel et al. 2003; Jelinek, Tschumper et al. 2003).

Further analysis of the gene expression profiles of isolated CLL cells detected a small number of genes, including those encoding ZAP-70, IM1286077, and C-type Lectin, which are differentially expressed and correlate with the mutational status of the IgVH genes (Klein, Tu et al. 2001; Rosenwald, Alizadeh et al. 2001). Gene array studies therefore suggest that CLL cells from different patients share common expression level of many genes and have a gene expression profile distinct from that of other B-cell malignancies or normal B-cells. CLL cells which express unmutated IgVH can be distinguished from those that have mutated IgVH through the differential expression of this small subset of genes. One of these genes encodes zeta-associated protein 70 (ZAP-70), a 70 kDa ζ-chain CD3-receptor-associated protein tyrosine kinase (PTK) of T lymphocytes. B-cells generally lack ZAP-70, but instead use another related PTK, Syk, for signal transduction via the BCR complex (Chan, van Oers et al. 1994). The functional significance of ZAP-70 gene expression in CLL cells is unknown, although it has been suggested that ZAP-70 could reconstitute BCR signaling in Syk-deficient B-cells (Kong, Bu et al. 1995) and that expression of ZAP-70 is associated with
increased BCR signaling (Chen, Widhopf et al. 2002). It has also been suggested that ZAP-70 may play a role in CD38 signalling (Deaglio, Zubiaur et al. 2002).

These are important findings which will not only help us better understand the biology of CLL, but also indicate that ZAP-70 may have an important role as a prognostic indicator in CLL. DNA microarrays are not suited for a routine clinical laboratory and methods of measuring ZAP-70 activity by flow-cytometry have been developed and are being used, both in research and clinical laboratory setting.

1.8 Aims of the project

CLL is one of the most commonly diagnosed leukaemias managed by haematologists. For many years, patients with CLL were viewed as a homogenous group with an indolent disease, long natural history and only marginally effective therapies that rarely yielded complete responses. More recently with increasing knowledge about the biology, molecular genetics, and prognostic factors of the disease and the emergence of new therapeutic options, the philosophy of care for patients with CLL is evolving from palliation to aiming for a potential cure, especially in younger patients. This study aims to unravel the prognostic significance of the novel biological markers in CLL, in a well characterized cohort of Irish patients attending St. James Hospital (SJH). The information will be used to devise a risk-adapted therapeutic approach for the management of patients with CLL.

1.8.1 Collect a cohort of well characterized Irish CLL patients

Increasing epidemiological evidence suggests that the incidence of CLL is higher in rural areas where intensive agriculture is carried out. As the national centre for stem cell transplantation and the main tertiary referral centre for haematological malignancy, we have noticed an increased incidence of aggressive CLL in younger patients from rural areas. Apart from the conventional staging (clinical, morphological, biochemical), access to this well studied patient cohort will allow us to risk stratify these patients and will provide semi-quantitative data about the epidemiology of CLL in Ireland.
1.8.2 Determine the Frequency and significance of karyotypic abnormalities
Recurring chromosomal abnormalities play a role in identifying biological sub-types of leukaemia and lymphomas, thus allowing clinicians to develop specific treatment strategies for specific diseases. Conventional cytogenetic analysis is time consuming and frequently unrewarding in. FISH technology can be used to identify specific genetic abnormalities of known prognostic importance in interphase cells in up to 80% of cases of CLL. In recent years a number of prognostically important cytogenetic abnormalities have been identified (refer to section 1.5.2). Establishing the cytogenetic profile of this cohort of CLL cases will only enable us to evaluate the incidence of these abnormalities in the Irish population and provide data about correlation with other markers such as IgV_H mutational analysis, CD38 expression and ZAP70 expression.

1.8.3 Determine the mutational status of immunoglobulin heavy chain variable regions (IgVH)
B-CLL is immunophenotypically similar to naïve B lymphocytes, which have not encountered antigen or undergone selection by IgH hypermutation in the germinal centre. Recent studies have demonstrated that CLL can be divided into two groups based on the IgH variable gene mutation status, into mutated and unmutated types with important prognostic implications (refer to section 1.6.7). The prognostic strength of the mutation analysis is independent of conventional prognostic markers. Unfortunately it is a time consuming, expensive test, which is not widely available. It is important to clarify the role and establish the significance of the mutational status in a defined treatment algorithm, so that it can be used appropriately, e.g., to define patients who might benefit from an allogeneic stem cell transplant (SCT).

1.8.4 Analyse CD38 expression, its relation to IgVH and its significance as an independent prognostic marker
Recently a role for CD38 as an independent prognostic marker and predictor of the mutational status has been suggested albeit with some controversy (refer to section 1.7.2). This study aims to clarify the technical issues and clinical role of CD38 as a prognostic marker.
1.8.5 Analyse ZAP-70 expression, its relation to IgVH and CD38 and its significance as an independent prognostic marker

Recent studies involving the gene expression profiling studies of CLL patients have shown a difference in the expression of some genes between IgVH mutated and unmutated patients. One of the genes, ZAP 70 is a protein tyrosine kinase normally expressed in T lymphocytes. Our aim is to look at the ZAP-70 expression both as an independent prognostic marker and as a predictor for the mutational status.

1.8.6 Use of FCR chemomimmunotherapy, and the search for new biomarkers using Proteomics

Fludarabine has become the gold standard for the treatment of patients with CLL. Recent studies have shown that the combination of Cyclophosphamide and Fludarabine can further improve the ORR and CR rates. These results are encouraging, but most patients relapse and require further treatment. Rituximab, a monoclonal antibody against CD20, as a single agent in CLL has been disappointing. There is some evidence that combining Rituximab with Fludarabine potentiates chemotherapy-induced apoptosis. Our aim is to improve the response rates achieved by conventional combination chemotherapy by adding Rituximab to the treatment. We will also monitor MRD, as it is unclear whether superior disease control results in longer time to treatment (TTF) or improved survival.

More recently, gene expression analysis has been used in a number of haematological diseases for a better understanding of the underlying clinical heterogeneity. Genes with different expression levels affecting survival and or clinical staging have been found. Differences in gene expression however, do not fully explain the variable clinical behavior in every patient, as patients with similar gene expression may have different outcomes. The genome in an individual remains relatively static whereas the final effector pathway, the proteome constantly changes depending on the physiological environment. Gene expression analysis alone therefore is probably not sufficient on its own to explain the variable clinical behavior. Our aim is to assess differences in plasma protein expression in a group of well characterized patients with CLL (both conventional and new prognostic tools) undergoing novel chemomimmunotherapy, using proteomics. We hope to find disease biomarkers that may be useful for predicting the mutational status, for predicting therapeutic outcomes.
and importantly new markers that can improve our ability for predicting prognosis at diagnosis.

This project will provide semi-quantitative data about the apparently changing presentation of CLL in Ireland. The introduction of novel prognostic information will be validated before being introduced into routine clinical practise at St James Hospital and to Irish haematology units who wish to avail of the service. Finally an improved understanding of the biology of CLL may provide both novel therapeutic targets and provide clues to the aetiological factors underlying leukaemogenesis in this disease.
CHAPTER 2
MATERIALS AND METHODS
2.1 Patients

2.1.1 Consent, History, Clinical examination

One hundred and six consecutive chronic lymphocytic leukaemia (CLL) patients attending St. James Hospital, Dublin between July 2002 and June 2004 were included. Most patients were newly diagnosed at the time of study accrual or had been diagnosed in the past 18 months but had not received any treatment. A minority of the patients had been diagnosed earlier than 18 months but had not received treatment in the two years before study accrual. All patients fulfilled the morphological and immunophenotypic criteria of CLL (CD5/19+ve, CD23+ve, FMC7-ve, weak surface immunoglobulin and light chain restriction). Written informed consent was obtained by one of the five senior haematologists (AH, EV, PB, EC, SMC) after providing clear information about CLL, the purpose and benefits of new diagnostic tests, procedures involved and the voluntary nature of participation in this research study at the time of enrolment. The following parameters were measured at diagnosis and follow-up: B symptoms, lymph node, liver and spleen size, full blood count and differential white cell count (WCC), serum $\beta_2\text{-microglobulin (}\beta_2\text{m)}$ and lactate dehydrogenase (LDH). Follow-up was based on clinical examination unless the tumour sites were not amenable to clinical examination e.g. retroperitoneal adenopathy. Patients were staged according to the Binet classification. Stages A ($<3$ areas) and B ($\geq3$ areas) are defined according to the number of palpably enlarged lymphoid areas (cervical, axillary, inguinal nodes, spleen and liver) while stage C is defined by anaemia or thrombocytopenia resulting from bone marrow failure. Disease progression (defined by B symptoms, marrow failure, Binet stage progression, progressive adenopathy and or hepatosplenomegaly, increased incidence of infections and $<6$ months lymphocyte doubling time) was measured by treatment free interval in months from the time of diagnosis (TFI) and response to treatment was measured using National Cancer Institute working group (NCI-WG) criteria. Disease status and evidence of progression was evaluated by five senior doctors (AH, EV, SMC, PB, EC) and the decision to treat was based on agreed criteria of progressive stage A, B or C, $<6$ months lymphocyte doubling time (LDT), B symptoms or progressive hepatosplenomegaly or lymphadenopathy. TFI was used as a surrogate marker of disease activity and progression. All experiments except CD38 and ZAP-70 analysis were performed by the author which, was performed by the clinical laboratory staff.
Sample preparation for ZAP-70 analysis on the frozen cells was also performed by the author.

Forty mls of blood was collected at the time of study accrual. Peripheral blood mononuclear cells (PBMC) were separated by density centrifugation (Lymphoprep™, Fresenius Kabi, Norge AS) and aliquoted for DNA and RNA extraction and cryopreservation.

2.2 Morphology

2.2.1 Peripheral blood

Morphologic classification was performed on May-Grünwald-Giemsa (BDH, Poole, UK) stained peripheral blood smears taken at the time of diagnosis. Patients were classified into two groups: typical CLL (group 1) and atypical CLL (group 2). Cases classified as typical CLL had more than 90% typical CLL cells in their peripheral blood, namely small round lymphocytes (≤ the size of the red blood cells) with a round nucleus, clumped chromatin pattern, no nucleoli and scant cytoplasm. Cases were classified as atypical CLL if more than 10% of the lymphocytes were larger or had prolymphocytoid features. Prolymphocytoid cells are large, between two and three times the size of red blood cell and have relatively open chromatin with one nucleolus and abundant, basophilic cytoplasm.

2.2.2 Bone marrow involvement

The bone marrow involvement was assessed by performing an aspirate and biopsy. The percentage of lymphocytic infiltration and the pattern of bone marrow involvement was determined by bone marrow histology. Three patterns of marrow infiltration were recognised; diffuse, interstitial and nodular.

2.3 Cell Separation and storage

2.3.1 Density Gradient cell separation

Forty mls of the patients’ peripheral blood (PB) sample was put into two sterile 50mls tubes (Sarstedt, AG, Germany). Equal amount of RPMI 1640 (Invitrogen, Scotland, UK) was then added to the sample. After mixing, the sample was gently pipetted onto...
15mls of Ficoll (Lymphoprep™ Fresenius Kabi Norge AS, Oslo, Norway) in a 50ml tube, taking care not to disturb the blood-Ficoll layer so formed. The sample was centrifuged at 1800 rpm (IEC Centra GP8, MA, USA) for 25 minutes. The mononuclear cell (MNC) layer formed between the red cells at the bottom and the plasma at the top was removed, using a 3ml Pasteur pipette (Copan, Italy) and placed in a 20mls Sterilin bottle (Bibby, Strafordshire, UK). An equal amount of phosphate buffered saline (PBS) (Invitrogen, UK) was added and the sample centrifuged at 1500 rpm for 7 minutes. The supernatant was removed and discarded. If red cells were present in the spun down MNC button, they were lysed by incubating the sample for 5 minutes using a red-cell lysis solution (Gentra, MN, USA). The sample was centrifuged at 1500 rpm for 15 minutes and the supernatant discarded. The sample was washed twice using excess PBS and centrifuged at 1500 rpm. After the second wash, the supernatant was removed and the MNC button resuspended with 1ml of PBS (More if the sample was concentrated). Five μl of the sample was then added to 1ml of 0.9% normal saline and a cell count was done using Sysmex automated counter (XE-2100, Kobe, Japan).

2.3.2 Cell storage for Fluorescent in situ hybridization (FISH)

At least two samples were stored on each patient. 5x10⁶ cells were collected in a 2ml Ependorf tube (Treff Labs, Switzerland). One ml of 0.075 M hypotonic KCL (Potassium Chloride, Sigma-Aldrich, Steinheim, Germany), pre-warmed to 37°C was added to the cells. After mixing, the sample was incubated for 15 min at 37°C in a dry heat block (DRI-BLOCK, Techne, Cambridge, UK). The sample was centrifuged at 3000 rpm for 3 minutes. After discarding the supernatant the cell button was resuspended by vigorous vortexing. One ml of 3:1 methanol (BDH, Poole, UK) and acetic acid (BDH, Poole, UK) was added to the sample drop by drop while continuously vortexing. The sample was labelled and stored at -20°C.

2.2.3 Cell storage for Nucleic acid extraction

2.3.3.1 DNA extraction and quantification

At least 2 samples were stored on each patient. 5x10⁶ cells were collected in a 2ml Ependorf tube. Cell pellets were resuspended in phosphate buffered saline (PBS) (Invitrogen, Scotland, UK) to a final volume of 200μl and transferred to a 1.5ml
Two hundred μl of Buffer AL (Qiagen, UK) was added to the sample and mixed by vortexing. The sample was labelled and stored at -20°C until required. Total DNA was isolated using the QIAamp DNA Blood mini kit (Qiagen, UK) according to the manufacturers’ instructions. Stored cells in AL buffer were thawed to room temperature (RT). Twenty five μl Qiagen proteinase K stock solution was added and the sample pulse-vortexed for 15 seconds and then incubated at 56°C for 10 minutes. The sample was briefly centrifuged, 200μl of 100% ethanol added and pulse-vortexed for 15 seconds before applying the sample to the QIAamp spin column in a 2ml collection tube. The tubes were spun at 8000 rpm for 1 minute in a microlite centrifuge (Micromax RF, IEC, USA). The filtrate was discarded and 500μl of Buffer AW1 (Qiagen, UK) applied to the column. The sample was centrifuged at 8000 rpm for 1 minute, after which the filtrate was discarded. 500μl of buffer AW2 (Qiagen, UK) was applied to the column, and the tubes were centrifuged twice at 13,000 rpm for 3 minutes, with the filtrate discarded both times. The spin column was transferred to a new 1.5ml ependorf tube and 200μl of buffer AE (Qiagen, UK) was added. Following incubation in buffer AE (Qiagen, UK) at RT for 5 minutes, the column was centrifuged at 8000 rpm for 1 min to elute the DNA. This step was repeated and eluates were combined. The concentration of DNA was measured by adding 5μl of the DNA eluate to 495μl of water and the sample pulse-vortexed. The absorption was then read at A260 nm by a spectrophotometer (Genequant). The reading was multiplied by a factor of 5000 to calculate the amount of DNA in ng/μl or μg/ml. DNA was stored at -20°C until required.

2.3.3.2 RNA extraction and quantification

Between 5 x 10⁷ and 1 x 10⁸ cells were stored for RNA extraction. The cells were resuspended in 200μl of PBS before 2-4mls of RLT (Qiagen, UK) was added. After vigorous vortexing for 1 minute, the sample was passed through a 20G (Green) needle attached to a 10ml RNAse free syringe 5-10 times for homogenization. The sample was labelled and stored at -80°C until required.

Total RNA was isolated using the QIAamp RNA Blood midi kit (Qiagen, UK) according to the manufacturers’ instructions. Previously stored cells in RLT buffer were first thawed to RT. One volume (4 mls) of 70% ethanol was added to the
homogenised lysate and mixed well by pipetting. Four mls (maximum) of the sample was applied to an RNeasy midi column in a 15ml tube and centrifuged in a Centra GP8 centrifuge (IEC, MA, USA) at 3000-5000 g for 5 min. The “flow through” was discarded and the remainder of the sample was reapplied to the same column along with any precipitate. The column was centrifuged at 3000-5000 x g for 5 minutes and the “flow through” discarded. Two mls of buffer RW1 (Qiagen, UK) was pipetted onto the RNeasy column and centrifuged at 3000-5000 g for 5 minutes. The “flow through” was discarded. 20µl of DNase stock solution was added to 140µl of RDD buffer (Qiagen, UK) in a 1.5ml ependorf and mixed gently by flicking the tube followed by brief centrifugation. 160µl of the DNase incubation mix was pipetted directly onto the RNeasy column membrane, which was left at RT for 15 minutes. Two mls of buffer RW1 (Qiagen, UK) was pipetted onto the RNeasy column and centrifuged at 3000-5000 g for 5 minutes. The “flow through” was discarded. Then 2.5mls of RPE buffer (Qiagen, UK) was applied to the column and centrifuged at 3000-5000 g for 2 minutes. The “flow through” was discarded. Another 2mls of RPE buffer was applied to the column, which was centrifuged at 3000-5000 g for 5 minutes to dry the column. The column was transferred to a new 15ml collection tube. 250µl of RNase free water was pipetted directly onto the RNeasy column membrane. Following incubation at RT for 5 minutes, the column was centrifuged at 3000-5000 g for 3 minutes. The concentration of RNA was measured by adding 5µl of the RNA eluate to 495µl of water and the sample pulse-vortexed. The absorption was then read at A260 nm by Genequant spectrophotometer. The reading was multiplied by a factor of 4000 to obtain the RNA concentration in ng/µl or µg/ml. RNA was stored at -80°C until required.

2.3.4 Freezing cells

Between 60 and 120 x 10⁶ cells were stored for each sample. The cells were collected into a 20 ml sterile tubes (Sarstedt, AG, Germany) tube. 1.5 – 3.0mls of 2% fetal calf serum (FCS, Invitrogen, Scotland, UK) in Hanks (Invitrogen, Scotland, UK) and 1.5 – 3mls of 20% dimethyl sulphoxide (DMSO, Sigma-Aldrich, Steinheim, Germany) were added to the sample slowly, drop by drop. The sample was mixed well by gentle pipetting before storing in pre-labelled 1.5mls ependorf tubes at -20°C for 10-12 hours. Thereafter the sample was stored at -80°C till required.
2.3.5 Thawing cells

A vial of frozen cells was removed from the -80°C freezer and placed on ice. Cells were incubated at room temperature for 1 minute and then transferred to a 37°C water bath until fully thawed. Using a 1 ml pipette, cells were transferred to a 10 ml tube. The cells were then washed twice in 2mls of warm PBSAT (Phosphate buffered saline containing Albumin and Tween) and resuspended to the original volume in PBSAT before use for DNA extraction, RNA extraction or flow cytometry.

2.4 Fluorescence in situ hybridization (FISH)

2.4.1 Slide preparation

The sample was spun at 3000 rpm for 10 minutes. The methanol:acetic acid was decanted and 1 ml of freshly prepared 3:1 solution of methanol:acetic acid was added to the cell button and mixed by vortexing. One ultra clean slide was labelled and used per patient sample. Slides were prepared using a double well cytospin (Shandon, Life Sciences Instruments, Cheshire, UK) and 100µl of sample per well and spun at 800 rpm for 10 minutes in the cytospin centrifuge (Shandon, Life Sciences Instruments, Cheshire, UK). The sample areas were marked using a diamond scribe (Proxxon, USA) and examined using a phase contrast microscope to ensure a good spread.

2.4.2 Pepsin digestion

The sample was pepsin digested to remove protein contamination during the slide preparation. A hundred ml of 0.01N hydrochloric acid (HCL) was poured into a Coplin jar (Wheaton, NJ, USA) to which 150µl of pepsin had been added (Dissolve 1gm of dry pepsin (Matrix, Switzerland) in 10mls of water and stored as aliquots of 150µl at -20°C for later use). The Coplin jar was then placed in a water bath (Grant Instruments, Cambridge, UK) set at 37°C to bring the temperature of the solution to 37°C. The slides were placed in the Coplin jar for 3 minutes, removed and washed for 5 minutes in 1 x PBS (Phosphate buffered saline, PH = 7) in a Coplin jar at RT on an orbital shaker. The slides were placed in 1% acid free formaldehyde (Merck, Hohenbrunn, Germany) at RT in a Coplin jar and left on an orbital shaker for ten minutes. The slides were washed for 5 minutes in 1 x PBS at RT in a Coplin jar on a shaker. The slides were dehydrated by placing them consecutively for two minutes
each through 70%, 90% and 100% ethanol at RT in Coplin jars on a shaker. The slides were air dried in a rack on a shaker.

2.4.3 Denaturation of slides/samples and the probe

2.4.3.1 Slides/samples
A Coplin jar filled with 70% formamide (Merck, Hohenbrunn, Germany) in 2X SSC (Sigma-Aldrich) was put in a water bath (Grant, UK) at 73±1°C. Half an hour was allowed for the temperature of the solution to equilibrate. The slides were put into the formamide solution for 3 minutes. After 3 minutes the slides were dehydrated by placing them consecutively for 3 minutes each through 70%, 90% and 100% ice-cold ethanol in Coplin jars placed on a shaker. Slides were air dried for 1 minute and then put on a slide warmer at 37°C.

2.4.3.2 Probe
Probe denaturation starts when the slides are passing through the 90% ice-cold ethanol. 1.5µl of probe for each sample area comprising the two sets (LSI D13S319 / 13q34 / CEP 12: LSI ATM / p53 - Vysis, Abbott) were put into two ependorf tubes and placed in a water bath at 73±1°C for 5 minutes. The probes were then immediately placed on ice.

2.4.4 Hybridization
1.5µl of each probe set was placed on the sample areas on the slide, one at a time. This was covered by a round coverslip (10mm, Menzel – Gläsen). The edges of the coverslip were covered with rubber cement (Royal Talens, Belgium). The slides were then placed in a humid box and incubated for at least 16 hours at 37°C.

2.4.5 Washing
The slides were removed after at least 16 hours from the incubator. The rubber cement was peeled off and the cover slip removed. The slides were then put in pre-warmed (73±1°C) 0.4 x SSC (Sigma-Aldrich, Germany) and 0.3% NP40 (Sigma-Aldrich, Germany) in a Coplin jar and a water bath at (73±1°C) for 1 minute. If more then 4 slides were being prepared, two Coplin jars were used and the timer started after the
first jar had been filled (4 slides per jar). After 1 minute the slides were removed and placed in a solution of 2 x SSC (Sigma-Aldrich, Germany) and 0.1% NP40 in a Coplin jar at RT on a shaker for 1 minute. The slides were put in 2 x SSC at RT for 1 minute on a shaker for the final wash. The slides were then dehydrated by placing them consecutively for 1 minute each through 70%, 90% and 100% ethanol at RT and air dried on a shaker. Finally 30µl of the counter stain, DAPI (4',6 diamidino-2-phenylindole) in mounting medium (VECTASHIELD® Mounting Medium, Vector, UK) was placed on to a large cover slip and placed over the slide. The cover slip was not sealed. The sample was then analyzed immediately or stored at -20°C for later analysis.

2.4.6 Sample Analysis

2.4.6.1 Probe Description and analysis

A probe set kit from Vysis including 5 probes in two sets was used. The five FISH DNA probes are supplied in a set of two probe kit, a multi-color probe product which is designed to provide simultaneous detection of all four prognostically relevant chromosome regions in two hybridisation reactions (two and three probes per hybridization, respectively). Probe Set 1 contains LSI p53 in SpectrumOrange™ and LSI ATM in SpectrumGreen™. Probe Set 2 contains LSI D13S319 in Spectrum Orange, LSI 13q34 in SpectrumAqua™, and CEP 12 in SpectrumGreen. The LSI p53 (17pl3.1) probe is a ~145 kb unique sequence probe which is labelled in SpectrumOrange and spans the entire p53 gene. The LSI ATM probe is a ~500 kb unique sequence probe which hybridizes to the 11q22.3 region of chromosome 11. This probe spans the entire ~184 kb Ataxia Telangiectasia Mutated (ATM) gene and other adjoining genes. The LSI D13S319 probe is a ~130 kb unique sequence probe, labelled in SpectrumOrange and is one of three probes hybridizing to the 13q14.3 region. The LSI 13q34 detects a subtelomeric region of chromosome 13q. It is a ~550 kb unique sequence probe that hybridizes to the 13q34 region containing the Lysosomal-associate Membrane Protein (LAMP1) gene and several others. The probe is labelled with SpectrumAqua. The CEP 12 DNA probe hybridizes to the alpha satellite (centromeric) region (12p11.1-q11) of chromosome 12. The probe is labelled in SpectrumGreen.
Probe set 1
This probe set allows assessment of two chromosome regions: 17p13.1 (p53) and 11q22.3 (ATM). In a normal cell with two intact copies of chromosomes 17 and 11, two orange and two green signals will be observed. In an abnormal cell with a deletion in the p53 region, fewer than two orange signals will be observed. In an abnormal cell with a deletion in the ATM region on chromosome 11, one will observe fewer than two green signals (Fig 2.1 and 2.2).

Figure 2.1 An abnormal cell hybridized with Probe Set 1. One copy of the p53 gene region has been deleted as indicated by the single orange signal. Both copies of the ATM gene region are present as indicated by the two green signals.

Figure 2.2 An abnormal cell hybridized with Probe Set 1. One copy of the ATM gene region has been deleted as indicated by the single green signal. Both copies of the p53 gene region are present as indicated by the two orange signals.

Probe set 2
This probe set allows status assessment of three chromosome regions: 13q14.3 (D13S319), 13q34, and 12p11.1-q11. In a normal cell with two intact copies of
chromosome 13 and chromosome 12, two orange, aqua and green signals will be seen. In an abnormal cell with chromosome 13q deletions, more complex signal patterns may be expected depending upon the nature of the deletion. Monosomy 13 or 13q- will appear as a single orange or aqua signal and two green signals. An interstitial deletion of the 13q14.3 region will appear as either a one orange, two aqua, two green signal pattern (hemizygous or mono allelic deletion) or a two aqua, two green signal pattern (homozygous deletion or bi allelic deletion). In an abnormal cell with an increase in chromosome 12 copy number such as trisomy 12, one will observe three green signals Fig 2.3 and 2.4.

Figure 2.3 An abnormal cell hybridized with Probe Set 2. Both copies of chromosome thirteen and its q arm are intact as indicated by the two orange (LSI D13S319) and two aqua (LSI 13q34) signals. One extra copy of chromosome 12 (trisomy 12) is present as indicated by the three green signals.

Figure 2.4 An abnormal cell hybridized with Probe Set 2. One copy of chromosome 13 is deleted for the D13S319 region as indicated by the single orange signal (LSI D13S319) and the two aqua signals (LSI 13q34). One extra copy of chromosome 12 (trisomy 12) is present as indicated by the three green signals.
Before the analysis of patient samples, control slides were prepared from normal samples and analyzed to set cut-off values for each probe in the two sets. Ten slides were hybridized using the above technique. Five hundred cells were counted for each probe. Abnormal signals constituted less than two or more than two signals for each individual probe. Up to 5% interphase nuclei showed abnormal signals. Therefore, more than 5% showing abnormal signals i.e. 3 hybridization signals were considered to indicate trisomy and less than 2 hybridization signals were considered to indicate deletion. When analysing patient samples 200 cells were counted using Olympus BX61 microscope (Olympus, Japan), Camera and Cytovision version 3.1.1 software (Applied Imaging, Newcastle, UK).

2.5 IgVH mutational analysis

2.5.1 cDNA Synthesis

cDNAs was synthesised from RNA with the AMV Reverse Transcriptase kit (Promega, USA) according to the manufacturers instructions. For a 20μl reaction, 9.9μl of RNA (~1 μg) was pipetted into a 0.5ml microcentrifuge ependorf and incubated at 70°C for 10 minutes before cooling on ice. Meanwhile the rest of the reaction mixture was prepared with 4μl of 25mM MgCl₂, 2μl of 10x reverse transcription buffer, 2μl of 10mM dNTP, 0.5μl RNase, 0.6μl AMV reverse transcriptase and 1μl Oligo dT primer. After vortexing 10.1μl of the reaction mixture was added to the RNA and incubated at 42°C for 30-45 minutes, 99°C for 5 minutes and 4°C for 5 minutes. The cDNA was then stored at -20°C until required.

2.5.2 IgVH Specific PCR using cDNA

Immunoglobulin (Ig) molecule comprises of two identical heavy (H) and light (L) chains each. The combined protein consists of a variable region responsible for antigen binding and a constant region, which determines the effector function and class of the molecule. The Variable region is further sub-divided into 3 complementarity determining regions (CDRs 1-3) or hypervariable regions for antigen binding and 4 framework regions (FRs 1-4) for maintaining the integrity of the antibody-antigen combining site (Figure 2.5).
The H chain variable region gene consists of a family of Variable (V<sub>H</sub>), Diversity (D<sub>H</sub>) and Joining (J<sub>H</sub>) region sequences. The rearrangement and subsequent joining of 1 of ~51 functional V<sub>H</sub>, 27 D<sub>H</sub> and 6 J<sub>H</sub> genes results in the H chain variable region. Similarly the L chain variable region is encoded by a member of the Vk or Vλ family linked to a corresponding Jκ or Jλ (Figure 2.6).
First a D and a J segment rearrange to form D-J

Followed by the rearrangement of a V segment to the D-J unit

**Figure 2.6** The Immunoglobulin Heavy chain gene rearrangement

Both RNA and DNA can be used as a template to amplify and sequence the immunoglobulin heavy chain genes; however RNA is preferable as it reduces the possibility of amplifying an aberrantly rearranged or non-productive \( V_H \) gene. RNA however needs to be reverse transcribed into cDNA first. If RNA is not available, genomic DNA can be used.

**2.5.2.1 Primers**

Two PCRs are performed simultaneously using either a mixture of 5' primers specific for the \( V_H \) leader sequence of \( V_H \) families 1-7, together with a 3' constant region primer or a 5' framework 1 (FW1) consensus primer and a 3' consensus \( J_H \) primer (Fig 2.7). Primers (synthesised by MWG-Biotech, Germany) are rehydrated in molecular grade water (Sigma, UK) to a concentration of 100pm/\( \mu l \).
Figure 2.7 Arrows indicating the 5' V_{H} leader primer for V_{H} families 1-7 and 5' FW1 consensus primer, together with a 3' constant region and a 3' consensus J_{H} primer

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Framework Region 1, Consensus</td>
<td>5'-AGG TGC AGC TG(GC) (AT)(G) AGT C(AGT)G G-3'</td>
</tr>
<tr>
<td>JH Consensus</td>
<td>5'-TGA GGA GAC GGT GAC C-3'</td>
</tr>
<tr>
<td>VH1 Leader</td>
<td>5'-CTC ACC ATG GAC TGG ACC TGG AG-3'</td>
</tr>
<tr>
<td>VH2 Leader</td>
<td>5'-ATG GAC ATA CTT TGT TCC ACG CTC-3'</td>
</tr>
<tr>
<td>VH3 Leader</td>
<td>5'-CCA TGG AGT TTG GGC TGA GCT GG-3'</td>
</tr>
<tr>
<td>VH4 Leader</td>
<td>5'-ACA TGA AAC A(CT)C TGT GGT TCT TCC-3'</td>
</tr>
<tr>
<td>VH5 Leader</td>
<td>5'-ATG GGG TCA ACC GCC ATC CTC CG-3'</td>
</tr>
<tr>
<td>VH6 Leader</td>
<td>5'-ATG TCT GTC TCC TTC CTC ATC TTC-3'</td>
</tr>
<tr>
<td>IgM Constant region</td>
<td>5'-TTG GGG CGG ATG CAC T-3'</td>
</tr>
<tr>
<td>Internal to JH, Consensus</td>
<td>5'-GTG ACC AGG GGT(AGCT) CCT TGG CCC CAG-3'</td>
</tr>
</tbody>
</table>

Table 2.1 Primer Sequences

2.5.2.2 Reaction Conditions

PCR was performed in a final volume of 50µl, with 1 µl of each primer, 1µl of each deoxyribonucleoside triphosphate, 0.25µl AmpliTaq Gold® DNA polymerse with 5µl reaction buffer and 5µl magnesium chloride (Applied Biosystems, Foster City, CA) and 2 µl of sample cDNA(Table 2.2). Amplification consisted of an initial denaturation step of 5 minutes at 94°C followed by 30 cycles of 94°C for 30 seconds, 56°C for 30 seconds and 72°C for 1 minute, with a final extension step of 10 minutes at 72°C. For each PCR reaction a negative control with no added template was used to exclude contamination and a positive control using polyclonal human genomic DNA (Promega, Madison, USA).
Table 2.2 Reaction set up for PCR

<table>
<thead>
<tr>
<th>Components</th>
<th>Volume/Reaction (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10X PCR buffer II (100mM Tris-HCl pH8.3, 500mM KCl)</td>
<td>5</td>
</tr>
<tr>
<td>25 mM MgCl2 solution</td>
<td>5</td>
</tr>
<tr>
<td>10mM dNTPs</td>
<td>4</td>
</tr>
<tr>
<td>Primer F (50µM)</td>
<td>1</td>
</tr>
<tr>
<td>Primer R (50µM)</td>
<td>1</td>
</tr>
<tr>
<td>Sample cDNA</td>
<td>2</td>
</tr>
<tr>
<td>AmpliTaq Gold</td>
<td>0.25</td>
</tr>
<tr>
<td>Molecular grade dH2O (Rnase and Dnase-free)</td>
<td>31.75</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>50</strong></td>
</tr>
</tbody>
</table>

2.5.3 IgVH Specific PCR using genomic DNA

If RNA was not available or the cDNA did not amplify, genomic DNA was used, using a nested PCR with the same 5' mixture of leader primers and 3' JH consensus primer in the first round of amplification. In the second round PCR reaction, 2µl of the first round product and 5' FW1 consensus primer with an internal 3' JH consensus primer were used. The PCR conditions were the same as for amplifying cDNA, except that the nested PCR was repeated for 25 cycles.

2.5.4 Gel electrophoresis

To check successful amplification of the PCR reactions, PCR products were run in a 2% agarose (D-1, Low EEO, Conda, Madrid, Spain) gel. Two grammes (gm) of agarose was measured and added to 100 mls of 1x TAE (Sigma-Aldrich, Steinheim, Germany) in a glass beaker. The agarose was dissolved by heating the solution in a microwave. Thirty µl of ethidium bromide (Sigma-Aldrich, Steinheim, Germany) was added to the solution, which was then poured into the gel tank with gel combs in place. When the gel had set, the combs were removed. 10 µL of the PCR product was mixed with 2 µL of 6X gel loading buffer in a 1.5ml Ependorf tube. A DNA molecular weight marker was prepared in the same way (DNA MWM VIII, 0.019-1.11 Kbp, Roche, Germany). The DNA ladder was added to the outermost well and the samples to the remaining wells. The gel was run at 100 volts for 1-1.5 hours.
before being photographed under UV light. If a monoclonal band was obtained for each amplified sample then the remainder of the PCR product was purified Fig 2.8.

![Figure 2.8 Result for dual PCR reactions on 7 patients with a negative control and a positive control. (L) Denotes molecular weight marker. (•) Successful reaction with a monoclonal band while, (▼) failed reaction.]

2.5.5 PCR cleanup

A single PCR product was purified directly. If there were more than one band then the remainder of the PCR product was run on a gel, bands were then excised from the gel and purified.

2.5.5.1 PCR Purification

PCR products were cleaned using QIAquick® PCR purification kit (Qiagen, UK) according to the manufacturers directions. Five volumes of buffer PB (Qiagen, UK) were added to one volume of the PCR product. The sample was applied to the QIAquick column in a 2mL collection tube, which was centrifuged at 13,000 rpm in the microcentrifuge for one minute. The ‘flow through’ was discarded and 600 μL of buffer PE (Qiagen, UK) added to the QIAquick column. The column was centrifuged at 13,000 rpm in the microcentrifuge for one minute. The ‘flow through’ was discarded and the column centrifuged for one minute at 13,000 rpm. The column was placed in a clean 1.5ml microcentrifuge tube and 50 μL of buffer EB (Qiagen, UK) was added to the centre of the membrane and centrifuged at 8000 rpm for one minute after an incubation period of one minute. The purified PCR product was now ready for sequencing.
2.5.5.2 Gel purification

If there were more than 2 PCR products, then QIAquick gel extraction kit was used for purifying the two separate PCR products according to the manufacturers' directions. The DNA fragment was excised from the agarose gel with a clean, sharp scalpel and weighed in a colourless 1.5ml ependorf tube. Three volumes of buffer QG (Qiagen, UK) were added to one volume of gel. This was incubated at 50°C for 10 minutes or until the gel had dissolved. One volume of isopropanol was added to the sample and mixed. The sample was applied to a QIAquick spin column which was then placed in the 1.5ml collection tube and centrifuged at 13,000 rpm for one minute. The 'flow through' was discarded and the QIAquick spin column was placed in the same collection tube. Then, 0.5mL of buffer QG (Qiagen, UK) was added to the QIAquick column and centrifuged at 13,000 rpm for one minute. The 'flow through' was discarded and 0.75mL of buffer PE (Qiagen, UK) added to the QIAquick column and centrifuged for one minute at 13,000 rpm. The 'flow through' was discarded and the QIAquick column centrifuged for one additional minute at 13,000 rpm. The column was then placed in a 1.5ml clean microcentrifuge tube and 50 µL of buffer EB (Qiagen, UK) was added to the centre of the membrane and centrifuged for one minute to elute the DNA. The clean PCR product was then ready for sequencing.

2.5.6 Sequencing

Sequencing was carried out using the BigDye Terminator v3.1 Cycle sequencing kit (Applied Biosystems, UK) according to manufacturers instructions, on the ABI 3100 automatic sequencer (Applied Biosystems, UK). Forward and reverse sequencing reactions of 20µl were set up in 0.2ml thin walled tubes (Scientific laboratories, UK) as follows:

<table>
<thead>
<tr>
<th>Components</th>
<th>Volume/Reaction (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template*</td>
<td>5</td>
</tr>
<tr>
<td>Forward or Reverse primer (2pmol/ µl)</td>
<td>1.6</td>
</tr>
<tr>
<td>Big Dye Terminator</td>
<td>4</td>
</tr>
<tr>
<td>Molecular grade water</td>
<td>9.4</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
</tr>
</tbody>
</table>

*Table 2.3 Reaction set up for sequencing PCR
*Template refers to both cleaned PCR products and control sample.
A sequencing control reaction was also carried out with 1μl pGem, 2μl M13 primer, 4μl Big Dye Terminator and 13μl water. Reactions were amplified under the following conditions: 94°C – 30 seconds, (96°C – 10 seconds, 50°C – 10 seconds, 60°C – 4 minutes) for 25 cycles.

The sequencing products were cleaned using DyeEx spin columns (Qiagen, UK) as recommended by the manufacturer. Columns were prepared by vortexing and centrifuging at 750 g for 3 minutes. The 20μl sequencing reaction was transferred to the column and centrifuged. The eluate was collected in a 1.5 ml microcentrifuge tube and dried in a DNA 100 SpeedVac (Savant, UK) at medium heat for 20 minutes or until dry. The pellets were resuspended in 10μl of HiDi (highly deionized) formamide (Applied Biosystems, UK) and transferred to a 96-well optical reaction plate and sealed with a septa strip (all parts by Applied Biosystems, UK).

All sequences were read using the automated ABI 3100 DNA Sequence Detection System (Applied Biosystems, UK). The samples were dispersed through either a 36cm or 50cm capillary arrays with Pop 6 polymer (all by Applied Biosystems, UK). Before analyzing samples, a plate record was created by adding sample names, selecting dye set Z, mobility file pop6Bdy3v1, project name 3100, run module rseq36pop6 or stdseq50pop6, analysis module pop6rrun or pop6stdrun. The operation of the instrument is described in detail in the manual (ABI Prism 3100 Genetic Analyser Users manual (Applied Biosystems)).

The data was extracted and analysed on DNASTAR Seqman II software (DNASTAR inc., USA). After the run, reverse and forward sequences from 2 PCR reactions i.e. 4 sequences were complemented using the lasergene software. After alignment, sequence differences were identified by looking at the electropherogram (EPG Fig 2.9). The consensus sequence (Fig 2.10) was aligned to V-base (database of all germline Ig genes, http://vbase.mrc-cpe.cam.ac.uk/) Figure 2.11, and imported into a word document. The sequence was translated, checked for stop codons, the CDR regions were marked and replacement mutations highlighted. The percentage mutation was based on the number of nucleotide substitutions between the start of FR1 region and end of FR3 region of the Ig heavy chain gene. Sequences with ≥98%
Figure 2.9 The sequence of immunoglobulin heavy chain using two PCR reactions. The sequence was shown in 5' to 3' orientation.
TCCATGGAAGTTGGGCTGAGCTGGCTTTTCCTTGTGGCTATTTTAAAAGGTGTCGAGTGT
GAGGTGCAAGCTGTGTTGAGTCTGCGGGAGGCATGGTACAGCCGGGGGGGTCCTGAGACTC
TCCTGTGCAAGCCTCTGGGATTCACCTTTAGCAACTATGCCATGAGCTGGGTCGGCCAGGTT
CCAGGGGAAGGGGCTGGAGTGGGTCTCCGGTATTAGTGATGGTGGAATATAGAATACTAC
TCCGACTCCGTGAAAGACGGCTCTCCACATCTCCAGAGACAATCCCCACGAAACATGATATAT
CTGCAAATGAACAGCCTGAGAGCCGACGACACGGCCATATATTACTGTGCGAAAGGGATT
GGGTGGCTACGTCTCCTTGATGCTCTTAAATGGGCAAGGGAACATAGGTCACCAGTC
TCTTCAGGGAGTGCATCCGCCCCAAAA

\[\text{^\^: 350,447}\]
\[\text{Contig 1 (1,447)}\]
\[\text{Contig Length: 447 bases}\]
\[\text{Average Length/Sequence: 353 bases}\]
\[\text{Total Sequence Length: 1414 bases}\]
\[\text{Top Strand: 2 sequences}\]
\[\text{Bottom Strand: 2 sequences}\]
\[\text{Total: 4 sequences}\]

\textbf{Figure 2.10} Consensus sequence obtained after complimenting and aligning the 4 sequences using lasergene software.
### CDR2

<table>
<thead>
<tr>
<th>EMBL</th>
<th>Locus</th>
<th>Name</th>
<th>41</th>
<th>42</th>
<th>43</th>
<th>44</th>
<th>45</th>
<th>46</th>
<th>47</th>
<th>48</th>
<th>49</th>
<th>50</th>
<th>51</th>
<th>52</th>
<th>52a</th>
<th>52b</th>
<th>52c</th>
<th>53</th>
<th>54</th>
<th>55</th>
<th>56</th>
<th>57</th>
<th>58</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z12347</td>
<td>3-23</td>
<td>DP-47/V3-23...+</td>
<td>CCA</td>
<td>GGG</td>
<td>AAG</td>
<td>GGG</td>
<td>CTG</td>
<td>GAG</td>
<td>TGG</td>
<td>TGG</td>
<td>GCC</td>
<td>ATT</td>
<td>AGT</td>
<td>GAT</td>
<td>GGT</td>
<td>GGT</td>
<td>AAT</td>
<td>GTA</td>
<td>AAA</td>
<td>TAC</td>
<td>------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>J00236</td>
<td>3-23</td>
<td>VH26Rabbitts+</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>U29481</td>
<td>-</td>
<td>VH26-3.7</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
</tbody>
</table>

| EMBL | Locus | Name               | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 |
|------|-------|-------------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Z12347 | 3-23 | DP-47/V3-23...+   | TAC | TCC | GAC | TCC | GTG | AAG | GAC | CGC | TTC | ACC | ATT | TCC | AGC | GAG | AAT | CCC | ACG | AAC | ATG | ATA | TAT | ------ | ------ |
| J00236 | 3-23 | VH26Rabbitts+     | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |
| U29481 | -    | VH26-3.7          | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |

| EMBL | Locus | Name               | 80 | 81 | 82 | 82a | 82b | 82c | 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 | 94 |
|------|-------|-------------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Z12347 | 3-23 | DP-47/V3-23...+   | CTG | CAA | ATG | AAC | AGC | CTG | AGA | GCC | GAC | GAC | AGC | GCC | ATA | TAT | TAC | TGT | GCG | AAA | ------ | ------ |
| J00236 | 3-23 | VH26Rabbitts+     | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |
| U29481 | -    | VH26-3.7          | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ | ------ |

**EMBL Locus Name**

<table>
<thead>
<tr>
<th>EMBL</th>
<th>Locus</th>
<th>Name</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z12347</td>
<td>3-23</td>
<td>DP-47/V3-23...+</td>
<td>------</td>
</tr>
<tr>
<td>J00236</td>
<td>3-23</td>
<td>VH26Rabbitts+</td>
<td>------</td>
</tr>
<tr>
<td>U29481</td>
<td>-</td>
<td>VH26-3.7</td>
<td>------</td>
</tr>
</tbody>
</table>

**EMBL Name**

<table>
<thead>
<tr>
<th>EMBL</th>
<th>Locus</th>
<th>Name</th>
<th>score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z12347</td>
<td>3-23</td>
<td>DP-47/V3-23...+</td>
<td>------</td>
</tr>
<tr>
<td>J00236</td>
<td>3-23</td>
<td>VH26Rabbitts+</td>
<td>------</td>
</tr>
<tr>
<td>U29481</td>
<td>-</td>
<td>VH26-3.7</td>
<td>------</td>
</tr>
</tbody>
</table>

**CAKGIGNLERPLDLALNKWGGTMVTVSSGSAPE**

In frame rearrangement

---

**Figure 2.11** Consensus sequence after alignment to V-base (Germline immunoglobulin database). The numbers (41-94) denote the amino acids, followed by the consensus sequence and the 5 closest matches (3 shown), showing the acquired mutations. The Closest Dh and Jh matches and scores are also shown.
sequence homology to the germline sequence were considered as unmutated and the remaining were labelled as mutated Ig genes.

2.5.7 Cloning

Patient samples were cloned if the PCR did not result in either a single discrete band or if clean sequences could not be obtained. Fresh PCR products were cleaned (described in 2.4.5) and cloned using TOPO TA One Shot cloning kit (Invitrogen, UK). Ligation reactions were set up within a laminar flow cabinet (Holten, UK) and incubated at RT for ten minutes as follows:

<table>
<thead>
<tr>
<th>Components</th>
<th>Volume/Reaction (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template (cleaned PCR products)</td>
<td>3</td>
</tr>
<tr>
<td>Salt solution</td>
<td>1</td>
</tr>
<tr>
<td>TOPO pCR2.1 vector</td>
<td>1</td>
</tr>
<tr>
<td>Sterile water</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>6</strong></td>
</tr>
</tbody>
</table>

**Table 2.4** TOPO TA ligation reactions. All components supplied by Invitrogen, UK. A vector only control was carried out by adding no template to the reaction.

Ligation products were transformed into One Shot chemically competent TOP10 *E. coli* cells. 50µl of cells were thawed on ice from -80°C, before being transformed with 2µl of ligation reactions. Transformation was performed on ice for ten minutes. Additionally, a transformation control was carried out by transforming 50µl of cells with 10pg of pUC19. The cells were heat shocked for 30 seconds at 42°C in a water bath. The vials were immediately transferred to ice, where 250µl of RT SOC medium was added. The cells were agitated at 200 rpm at 37°C for 1 hour in a G25 shaker incubator (New Brunswick Scientific Co., USA). Following incubation, 150µl of each reaction was spread onto a pre-warmed sterile agar plate (1% Amp, with XGAL) and incubated upside down and light protected at 37°C overnight.

The following day, the plates were removed from the incubator. The transformation and ligation efficiency was calculated by comparing total colony numbers on
individual plates with that of pUC19 and by blue, white colour screening as described in the next paragraph.

Ten positive transformants (white colonies) from each PCR reaction were selected for sequencing. Individual colonies are picked with a p20 gilson tip and transferred initially to a fresh “patch” plate (labelled with a grid for ten colonies) and then into a pre-made PCR mix to check for the presence of the PCR product. Patch plates were incubated overnight as described previously. 20μl PCR reactions were set up using GeneAmp Amplitaq Gold as described in section 2.5.5.2, with M13 forward and reverse primers (50μM) (Table 2.5). These primers hybridize to each end of the cloning site within the TOPO vector. The DNA was amplified at 94°C - 10 minutes, (94°C- 1 minute, 58°C- 1 minute, 72°C- 10 minutes) for 30 cycles, 72°C-10 minutes, 4°C forever, on an MJ Research PTC-200 thermal cycler.

<table>
<thead>
<tr>
<th>Components</th>
<th>Volume/Reaction (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template*</td>
<td>5</td>
</tr>
<tr>
<td>M13 Forward or Reverse primer (2pmol μl)</td>
<td>1.6</td>
</tr>
<tr>
<td>Big Dye Terminator</td>
<td>4</td>
</tr>
<tr>
<td>Molecular grade water</td>
<td>9.4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>20</strong></td>
</tr>
</tbody>
</table>

Table 2.5 Reaction set up for the sequencing PCR

PCR products were analysed by gel electrophoresis as already described, to check for successful PCR integration and amplification. The TOPO amplicon was 201bp if no insert was present.

The following day, positive clones were selected from the patch plates and transferred into 20μl of dH₂O, where they were lysed by incubation at 98°C for 20 minutes. Lysed samples were stored at -20°C until required for sequencing. Sequencing was carried out as described previously in section 2.5.6.
2.6 CD38 Expression analysis

Fresh peripheral blood samples were diluted in CellWash™ (Becton Dickinson, San Jose, California) to $2 \times 10^6$ cells/ml. 100μl of the diluted blood was incubated with the following antibodies; Fluorescein isothiocyanate (FITC)-labelled anti-CD45 (CAT#345808), Phycoerithrin (PE)-labelled anti-CD5 (CAT#345782), Peridinin chlorophyll protein (PerCP)-Cy5.5 labelled anti-CD19 (CAT#332780) and Allophycocyanin-(APC)-conjugated anti-CD38 (CAT#345807) (Becton Dickinson, San Jose, California). At least 10,000 events were acquired using the Cellquest software and a FACSCalibur™ flowcytometer (Becton Dickinson, San Jose, California). Lymphocytes were gated using CD45 staining versus side scatter (SSC) characteristics and were gated and labelled as R1. Dot plots were generated on this population (R1) to determine CD5/CD19 and gated R2. Quadrant markers were set using Isotype matched antibodies with negative staining on this gated population. The percentage of CD38+ve events was then determined for the dual CD5/CD19 positive population of cells using a cut-off value of 30% (Figure 2.11 A, B and C). In addition to the 30% cut-off value, 20%, 25% and 35% cut-off values were also investigated to assess in relationship to the mutational status of the IgV\textsubscript{H}.

In order to determine if CD38 expression changes over time during follow up, some patient samples were randomly selected both treated and untreated, to measure CD38 expression status using the same method. Dot-plots and histograms were used to determine the expression pattern. According to the pattern of CD38 expression 3 groups of patients were identified (Figure 2.11 C). CLL can be homogenously negative for CD38 (Unimodal CD38-ve group), or they can homogenously express CD38 at high levels (Unimodal CD38+ve group). A third distinct subset of patients shows a bimodal expression profile of CD38, with the concomitant presence of one population expressing high levels of CD38 and second population completely negative (CD38 bimodal group).
A. Plot of CD45 v SSC. Lymphocyte gating using CD45 staining (FITC-labelled) versus side scatter (SSC) characteristics. Region R1 encompassing lymphocytes
B. Plot of CD5vCD19. R1 gated region was used to determine CD5/CD19 co-expression. Quadrant markers (Q 1-1 to Q4-1) were set using Isotype matched antibodies with negative staining on this gated population. Region R2 (Upper right quadrant, Q2-1) encompasses CD5/CD19 co-expressing cells.
C. Dot-plots and histograms showing CD38 expression in CD19 expressing lymphocytes from 3 representative patients. Upper right quadrant shows lymphocytes co-expressing CD19 (PerCP-Cy5-5-A labelled) and CD38 (APC-A labelled).

**Figure 2.12** A, B and C. Analysis of CD38 expression
2.7 ZAP-70 Expression Analysis

All samples analysed were fresh, however a concordance study was also performed on 40 samples using fresh and cryopreserved samples from the same patient. The percentage of CD19+ve/CD5+ve cells, which expressed ZAP-70 using a 10%, 15%, 20% and 25% cut off level, was quantified as previously described (Crespo, Bosch et al. 2003). Fresh blood samples were diluted to give a count of $2 \times 10^6$ cells/ml. 100μl of the diluted blood was fixed and permeabilized with the Fix & Perm kit (Caltag Laboratories, Towcester, UK). Frozen samples were thawed at 37°C, washed twice in 2mls of warm PBSAT and resuspended to the original volume in PBSAT before permeabilization.

1.5μg (10μl) of anti-ZAP-70 antibody (Clone 2F3.2, Upstate, Charlottesville, VA, USA) was added and the sample incubated for 15 minutes at RT, washed twice, incubated with rabbit anti-mouse immunoglobulin-RAM-PE (DAKO, Cambridgeshire, UK), washed, incubated at RT with 100μl neat mouse serum for 5 minutes and washed. Anti-CD3 FITC, CD56 FITC, CD19 APC and CD5 PerCP-Cy5.5 (Becton Dickinson) were added, incubated for 15 minutes at RT, washed twice and resuspended in 0.5 ml of CellWash™. Lymphocytes were gated using forward scatter (FSC) staining versus side scatter (SSC) characteristics and gated as R1. Dot plots were generated on this population to determine CD5/CD19 positivity and gated R2. A live gate was used to ensure that at least 2000 T cells and natural killer (NK) cells were analyzed in each sample (R3). The ZAP-70 staining on the population of CD3+ve/CD56+ve cells was used as an internal positive control and compared to ZAP-70 staining on the CD5+ve/CD19+ve (CLL) population. The percentage Zap-70 positivity in CLL cells was determined using a quadrant marker on the left margin of the T/NK cell population (Figure 2.13, A-F)
A. FSC v SSC raw data

B. Region R1 encompassing lymphocytes

C. R1 gated cells: Showing Region R2 encompassing CD5/CD19 co-expressing lymphocytes

D. CD5vCD3+56 (R1 gated). Region R3 lymphocytes encompassing T and NK cells
E. Plot of ZAP-70 v CD3+56. The CD3+56 positive population represent T and NK cells which act as positive/negative discriminator (i.e. they are assumed positive). CD3+56 negative events represent CD5/CD19 positive B-lymphocytes.

F. A gate applied to the data, such that the degree of ZAP-70 expression of the B-lymphocytes may be compared to the known positive population. The lower right hand quadrant (LR) represents ZAP-70 positive, CD5/CD19 co-expressing, B-lymphocytes.

Figure 2.13 Assessment of Zap-70 expression

2.8 Statistical Methods
Continuous variables were presented as median (upper and lower quartile) and levels in different groups were compared using a Mann-Whitney rank sum test. Categorical variables were summarized as %\((n)\) and frequencies compared using a standard chi-square test. The association between disease progression and each of ZAP-70, CD-38, and mutation status was analyzed. Kaplan Meier (KM) 'survival' curves described the cumulative risk of treatment in groups defined by their IgV\(_H\) mutational status, levels of ZAP-70 using a 20% cut-off value, CD38 expression level using a 30% cut-off value and the pattern of CD38 expression. KM curves based on TFI were compared using a log-rank test. The independent effect of the three prognostic factors was considered using a Cox proportional hazards model which included the 3 variables. The sensitivity, specificity and % discordance was used to assess discrimination of various cut-off values for CD38 and ZAP-70 expression. Diagnostic reliability was based on the positive and negative predictive value at various levels of CD38 and ZAP-70 expression. All statistical analysis was done using (STATA version 8.2) Statistical analysis software.
3.1 General Results

A total of one hundred and six consecutive patients attending St. James Hospital were accrued and included in the study between July 2002 and June 2004. Ten patients were diagnosed in other participating hospitals, prior to their referral to St James's Hospital. Seventy patients were newly diagnosed. All patients fulfilled the morphological and immunophenotypic criteria of CLL (CD5/CD19+ive, CD23+ive, FMC7-ive, weak surface immunoglobulin expression and light chain restriction).

Sixty six patients were male and 40 females with a median age at diagnosis of 62 years (range 37-90) Table 3.1.

<table>
<thead>
<tr>
<th>Table 3.1 Main biologic and clinical characteristics of the patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of Patients</td>
</tr>
<tr>
<td>Sex, Male / Female</td>
</tr>
<tr>
<td>Median Age (Yrs)</td>
</tr>
<tr>
<td>Stages</td>
</tr>
<tr>
<td>Median WCC x 10^9/l</td>
</tr>
<tr>
<td>Median Hb g/dl</td>
</tr>
<tr>
<td>Median Platelet Count x 10^9/l</td>
</tr>
<tr>
<td>LDH iu/l</td>
</tr>
<tr>
<td>β2m mg/dl</td>
</tr>
<tr>
<td>IgVH: Mutated / Unmutated</td>
</tr>
<tr>
<td>CD 38: Negative / Positive</td>
</tr>
<tr>
<td>ZAP-70: Negative / Positive</td>
</tr>
<tr>
<td>Median Follow-up (Months)</td>
</tr>
<tr>
<td>Treated / Not Treated</td>
</tr>
</tbody>
</table>

Table 3.1 Main biologic and clinical characteristics of the patients

The median haemoglobin was 13.5 g/dl (range 6.3 – 17.2), median white-cell count (wcc) was 29.8 x 10^9/l (range 7.3 – 304), median platelet count was 194 x 10^9/l (range 57 – 397), median neutrophil count was 5 x 10^9/l (range 1.2 – 36.5) and median lymphocyte count was 22 x 10^9/l (range 5 – 265). Although 62 patients had an abdominal ultrasound or CT scan done at the time of diagnosis, measurements
obtained by clinical evaluation were used for recording lymphadenopathy or hepatosplenomegaly. Three lymph node sites; cervical, axillary and inguinal were used for staging purposes. Fifty seven patients had no evidence of lymphadenopathy on clinical examination. Twenty-two patients had one lymph node site involvement, 18 had involvement at two sites and 9 patients had adenopathy at all three sites. Twenty-two patients had a clinically enlarged spleen and two patients had an enlarged liver. Clinical stage at the time of diagnosis was used for stratifying patients. Eighty-seven patients had stage A disease, 12 patients had stage B disease and 7 patients had stage C disease. Patients with stage B or C disease and those with progressive stage A disease (section 2.1.1) were treated. The majority of patients with stage A disease did not require treatment during the course of this study when compared to the other two groups (Table 3.2).

<table>
<thead>
<tr>
<th></th>
<th>Treated</th>
<th>Untreated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binet Stage A</td>
<td>87/106 (82%)</td>
<td>33/87 (38%)</td>
</tr>
<tr>
<td></td>
<td>50/106 (47%)</td>
<td>56/106 (53%)</td>
</tr>
<tr>
<td>Binet Stage B</td>
<td>12/106 (11%)</td>
<td>10/12 (83%)</td>
</tr>
<tr>
<td></td>
<td>2/12 (17%)</td>
<td></td>
</tr>
<tr>
<td>Binet Stage C</td>
<td>7/106 (7%)</td>
<td>7/7 (100%)</td>
</tr>
<tr>
<td></td>
<td>0/7 (0%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2 Stage and treatment according to the Binet staging.

The Direct Coombs test (DCT) status at diagnosis was available for 97 patients. Eighty seven patients had a negative test and 10 patients had a positive test. None of the patients with a positive DCT had evidence of haemolysis. B symptoms characterised by night sweats, fever and weight loss were present in 14 patients. Eight of 14 (58%) patients with B symptoms had advanced (stage B or C) disease. Ninety-six cases were classified with typical and 10 with atypical CLL. The majority of cases with typical CLL i.e., 85% had stage A disease. Only 4(4%) patients with typical CLL had stage C disease, compared to atypical CLL where 30% of the cases had stage C disease. More patients with atypical disease needed treatment, 60% compared to 52% of typical CLL cases (Table 3.3).
Chapter 3

<table>
<thead>
<tr>
<th></th>
<th>Typical CLL</th>
<th>atypical CLL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>96/106 (90%)</td>
<td>10/106 (10%)</td>
</tr>
<tr>
<td>Binet Stage A</td>
<td>81/96 (85%)</td>
<td>6/10 (60%)</td>
</tr>
<tr>
<td>Binet Stage B</td>
<td>11/96 (11%)</td>
<td>1/10 (10%)</td>
</tr>
<tr>
<td>Binet Stage C</td>
<td>4/96 (4%)</td>
<td>3/10 (30%)</td>
</tr>
<tr>
<td>Treated</td>
<td>50/96 (52%)</td>
<td>6/10 (60%)</td>
</tr>
<tr>
<td>Untreated</td>
<td>46/96 (48%)</td>
<td>4/10 (40%)</td>
</tr>
</tbody>
</table>

Table 3.3 Stage and treatment according to typical and atypical CLL.

Seventy eight patients had bone marrow aspirate and biopsy results. Forty-seven patients had diffuse marrow infiltration, 19 had interstitial involvement and in 12 patients the marrow was involved in a nodular pattern (Table 3.4).

<table>
<thead>
<tr>
<th></th>
<th>Diffuse</th>
<th>Interstitial</th>
<th>Nodular</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>47/78 (60%)</td>
<td>19/78 (24%)</td>
<td>12/78 (16%)</td>
</tr>
<tr>
<td>Binet Stage A</td>
<td>33/47 (70%)</td>
<td>19/19 (100%)</td>
<td>11/12 (92%)</td>
</tr>
<tr>
<td>Binet Stage B</td>
<td>8/47 (17%)</td>
<td>1/12 (8%)</td>
<td></td>
</tr>
<tr>
<td>Binet Stage C</td>
<td>6/47 (13%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treated</td>
<td>33/47 (70%)</td>
<td>6/19 (32%)</td>
<td>4/12 (33%)</td>
</tr>
<tr>
<td>Untreated</td>
<td>14/47 (30%)</td>
<td>13/19 (68%)</td>
<td>8/12 (67%)</td>
</tr>
</tbody>
</table>

Table 3.4 Histological patterns of marrow involvement.

The majority of patients had a diffuse pattern of marrow involvement (60%) and 30% of these had advanced stage B and stage C disease. All patients with interstitial marrow involvement had stage A disease and 92% of patients with nodular marrow involvement had stage A disease. Diffuse marrow involvement was associated with
aggressive disease, with 70% of the patients needing treatment compared to the interstitial and nodular pattern of marrow involvement where only 32% and 33% respectively required treatment.

β2-microglobulin (β2m) results were available for 73 patients at diagnosis. Thirty-nine patients had normal levels and 34 patients had elevated levels. The majority of patients in both groups had stage A disease; patients with normal results were at lower risk of needing treatment (28%) compared to patients with an elevated level of β2m (56%). (Table 3.5)

<table>
<thead>
<tr>
<th></th>
<th>β2m (Normal)</th>
<th>β2m (High)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binet Stage A</td>
<td>35/39 (90%)</td>
<td>27/34 (79%)</td>
</tr>
<tr>
<td>Binet Stage B</td>
<td>3/39 (8%)</td>
<td>4/34 (12%)</td>
</tr>
<tr>
<td>Binet Stage C</td>
<td>1/39 (2%)</td>
<td>3/34 (9%)</td>
</tr>
<tr>
<td>Treated</td>
<td>11/39 (28%)</td>
<td>19/34 (56%)</td>
</tr>
<tr>
<td>Untreated</td>
<td>28/39 (72%)</td>
<td>15/34 (44%)</td>
</tr>
</tbody>
</table>

**Table 3.5 Treatment and stage according to β2m**

Serum LDH analysis was available on 98 patients. Seventy-nine patients had a normal LDH and 19 patients had elevated levels. Although majority of patients with normal LDH levels had stage A disease, we did not find an association between higher LDH levels and advanced stages (B and C) of the disease. However, more patients with high LDH (68%) required treatment compared to patients with normal levels at presentation (43%). Table 3.6
Table 3.6  Treatment and stage according to LDH

The median duration of follow-up was 34 months (2-279). During follow-up, 50 out of 106 patients received treatment for progressive disease. The treatments included single agent Chlorambucil, Cyclophosphamide, Fludarabine or combination treatments such as CVP (cyclophosphamide, vincristine, and prednisolone), CHOP (cyclophosphamide, adriamycin, vincristine and prednisolone) and FCR (fludarabine, cyclophosphamide and rituximab). Response to treatment was measured using NCI-WG criteria.

3.2 Results for FISH analysis

A hundred patients were evaluated by interphase cytogenetics. Six patients were not evaluated because 2 had insufficient sample and 4 patients had no sample stored at the time of entry into the study.
Seventy-two (72%) cases had cytogenetic abnormalities. Table 3.7 lists these aberrations in order of decreasing frequencies.

<table>
<thead>
<tr>
<th>Aberration</th>
<th>No. of Patients (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Karyotype</td>
<td>28 (28%)</td>
</tr>
<tr>
<td>Clonal abnormalities</td>
<td>72 (72%)</td>
</tr>
<tr>
<td>13q14 deletion</td>
<td>51 (51%)</td>
</tr>
<tr>
<td>Trisomy 12</td>
<td>11 (11%)</td>
</tr>
<tr>
<td>11q deletion</td>
<td>6 (6%)</td>
</tr>
<tr>
<td>17p deletion</td>
<td>4 (4%)</td>
</tr>
</tbody>
</table>

Table 3.7 Incidence of chromosomal abnormalities in 100 patients with Chronic Lymphocytic Leukemia

A single abnormality was detected in 62 patients and 10 patients had two abnormalities. No patient had more than two defects. Sixty-one patients had 13q deletion; the deletion was the sole abnormality in 51 (51%) patients and in the remaining 10 patients, 13q deletion was accompanied by trisomy12 (3 patients), 11q deletion (4 patients) and 17p deletion (3 patients). Trisomy 12 occurred as a sole abnormality in 8 patients, 11q deletion in 2 patients and 17p deletion in 1 patient. All deletions were monoallelic except for the 13q14 region, where biallelic or concomitant monoallelic (hemizygous) and biallelic (homozygous) deletions were found in some patients Table 3.8.
Chapter 3

<table>
<thead>
<tr>
<th>UPN</th>
<th>Cytogenetic Abnormality</th>
<th>UPN</th>
<th>Cytogenetic Abnormality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>del 13q14</td>
<td>28</td>
<td>del 11q</td>
</tr>
<tr>
<td>2</td>
<td>del 13q14</td>
<td>52</td>
<td>del 11q</td>
</tr>
<tr>
<td>3</td>
<td>del 13q14</td>
<td>23</td>
<td>del 11q, del 13q14</td>
</tr>
<tr>
<td>5</td>
<td>del 13q14</td>
<td>55</td>
<td>del 11q, del 13q14</td>
</tr>
<tr>
<td>8</td>
<td>del 13q14</td>
<td>56</td>
<td>del 11q, del 13q14</td>
</tr>
<tr>
<td>9</td>
<td>del 13q14</td>
<td>102</td>
<td>del 11q, del 13q14</td>
</tr>
<tr>
<td>10</td>
<td>del 13q14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>del 13q14</td>
<td>25</td>
<td>del 17p</td>
</tr>
<tr>
<td>13</td>
<td>del 13q14</td>
<td>76</td>
<td>del 17p, del 13q14</td>
</tr>
<tr>
<td>14</td>
<td>del 13q14</td>
<td>4</td>
<td>del 17p, del 13q14</td>
</tr>
<tr>
<td>16</td>
<td>del 13q14</td>
<td>7</td>
<td>del 17p, del 13q14</td>
</tr>
<tr>
<td>20</td>
<td>del 13q14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>del 13q14</td>
<td>15</td>
<td>trisomy 12</td>
</tr>
<tr>
<td>24</td>
<td>del 13q14</td>
<td>31</td>
<td>trisomy 12</td>
</tr>
<tr>
<td>26</td>
<td>del 13q14</td>
<td>42</td>
<td>trisomy 12</td>
</tr>
<tr>
<td>29</td>
<td>del 13q14</td>
<td>49</td>
<td>trisomy 12</td>
</tr>
<tr>
<td>33</td>
<td>del 13q14</td>
<td>82</td>
<td>trisomy 12</td>
</tr>
<tr>
<td>34</td>
<td>del 13q14</td>
<td>84</td>
<td>trisomy 12</td>
</tr>
<tr>
<td>38</td>
<td>del 13q14</td>
<td>95</td>
<td>trisomy 12</td>
</tr>
<tr>
<td>39</td>
<td>del 13q14</td>
<td>64</td>
<td>trisomy 12</td>
</tr>
<tr>
<td>40</td>
<td>del 13q14</td>
<td>101</td>
<td>trisomy 12, del 13q</td>
</tr>
<tr>
<td>44</td>
<td>del 13q14</td>
<td>62</td>
<td>trisomy 12, del 13q</td>
</tr>
<tr>
<td>45</td>
<td>del 13q14</td>
<td>89</td>
<td>trisomy 12, del 13q</td>
</tr>
<tr>
<td>47</td>
<td>del 13q14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>del 13q14</td>
<td>6</td>
<td>Normal</td>
</tr>
<tr>
<td>54</td>
<td>del 13q14</td>
<td>12</td>
<td>Normal</td>
</tr>
<tr>
<td>58</td>
<td>del 13q14</td>
<td>17</td>
<td>Normal</td>
</tr>
<tr>
<td>59</td>
<td>del 13q14</td>
<td>18</td>
<td>Normal</td>
</tr>
<tr>
<td>60</td>
<td>del 13q14</td>
<td>19</td>
<td>Normal</td>
</tr>
<tr>
<td>61</td>
<td>del 13q14</td>
<td>21</td>
<td>Normal</td>
</tr>
<tr>
<td>63</td>
<td>del 13q14</td>
<td>27</td>
<td>Normal</td>
</tr>
<tr>
<td>65</td>
<td>del 13q14</td>
<td>30</td>
<td>Normal</td>
</tr>
<tr>
<td>66</td>
<td>del 13q14</td>
<td>32</td>
<td>Normal</td>
</tr>
<tr>
<td>67</td>
<td>del 13q14</td>
<td>35</td>
<td>Normal</td>
</tr>
<tr>
<td>68</td>
<td>del 13q14</td>
<td>36</td>
<td>Normal</td>
</tr>
<tr>
<td>69</td>
<td>del 13q14</td>
<td>37</td>
<td>Normal</td>
</tr>
<tr>
<td>70</td>
<td>del 13q14</td>
<td>41</td>
<td>Normal</td>
</tr>
<tr>
<td>73</td>
<td>del 13q14</td>
<td>43</td>
<td>Normal</td>
</tr>
<tr>
<td>74</td>
<td>del 13q14</td>
<td>46</td>
<td>Normal</td>
</tr>
<tr>
<td>75</td>
<td>del 13q14</td>
<td>48</td>
<td>Normal</td>
</tr>
<tr>
<td>77</td>
<td>del 13q14</td>
<td>51</td>
<td>Normal</td>
</tr>
<tr>
<td>78</td>
<td>del 13q14</td>
<td>53</td>
<td>Normal</td>
</tr>
<tr>
<td>81</td>
<td>del 13q14</td>
<td>57</td>
<td>Normal</td>
</tr>
<tr>
<td>85</td>
<td>del 13q14</td>
<td>79</td>
<td>Normal</td>
</tr>
<tr>
<td>86</td>
<td>del 13q14</td>
<td>80</td>
<td>Normal</td>
</tr>
<tr>
<td>90</td>
<td>del 13q14</td>
<td>83</td>
<td>Normal</td>
</tr>
<tr>
<td>92</td>
<td>del 13q14</td>
<td>87</td>
<td>Normal</td>
</tr>
<tr>
<td>93</td>
<td>del 13q14</td>
<td>88</td>
<td>Normal</td>
</tr>
<tr>
<td>94</td>
<td>del 13q14</td>
<td>91</td>
<td>Normal</td>
</tr>
<tr>
<td>96</td>
<td>del 13q14</td>
<td>100</td>
<td>Normal</td>
</tr>
<tr>
<td>99</td>
<td>del 13q14</td>
<td>103</td>
<td>Normal</td>
</tr>
<tr>
<td>106</td>
<td>Normal</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.8 Distribution of Cytogenetic abnormalities in 100 patients as detected by FISH
In order to assess the clinical significance of different cytogenetic abnormalities; patients were divided into two groups on the basis of need for treatment. Results are shown in Fig. 3.1

**Figure 3.1  Cytogenetic abnormalities Vs. Treatment (100 patients).** Orange bars show treated patients, while green bars indicate untreated patients. The numbers on each bar correspond to the number of patients with a cytogenetic abnormality in the group

Fifty-four (54%) of 100 patients did not require treatment during follow up and 46 (46%) patients received treatment. Of the 28 patients with normal cytogenetics, 12 (43%) received treatment and 16 (57%) did not require treatment.

Twenty-one (41%) of 51 patients with 13q14 deletion were treated compared to 30 (59%) who did not require treatment. Five (45%) patients with trisomy12 received treatment while 6 (55%) remained treatment naïve. Five of six patients with deletion 11q and 3 of 4 patients with deletion17p were treated.
Cytogenetically defined groups had significant differences in median treatment free interval (TFI). Patients with no cytogenetic abnormality had a median TFI of 87 months and were assigned a hazard ratio (HR) of 1 for comparison to the cytogenetically abnormal groups. Patients with 13q deletion had a median TFI of 110 months and HR of 0.54, CI: (0.23-1.26). Trisomy 12 was associated with a median TFI of 41 months and HR of 1.51, CI: (0.51-4.43). Patients with deletion17p had a median TFI of 11 months with a HR of 2.01, CI: (0.55-7.33) while patients with 11q deletion had a median TFI of 5 months and HR of 2.41, CI: (0.75-1.71), Fig. 3.2.

Figure 3.2 Kaplan-Meier survival curves comparing TFI in CLL patients according to the abnormalities detected by interphase cytogenetics. (100 patients): Median TFI for patients with no cytogenetic abnormality: 87 months, 13q deletion: 110 months, trisomy12: 41 months, 11q deletion: 5 months, deletion 17p: 11 months.

3.3 Results for IgV\textsubscript{H} mutational analysis

IgV\textsubscript{H} mutational analysis was available in 106 patients. In 98 patients RNA was used for sequencing as described previously (Fig. 3.3).
Figure 3.3 Result for dual PCR reactions on 7 patients with a negative and a positive control. (L) Denotes molecular weight marker. (•) Successful reaction with a monoclonal band while, (○) failed reaction.

In the remaining 8 patients where cDNA failed to amplify or was not available genomic DNA was used (Fig 3.4).

Figure 3.4 Result for PCR on 5 patients and a negative control using genomic DNA and a nested PCR. (L) Denotes the molecular weight marker.
In 5 patients, the PCR did not result in either a single discrete band or clean sequences could not be obtained from the band. These samples were cloned as previously described (Fig 3.5).

**Figure 3.5** PCR results for positively transformed colonies after a successful cloning reaction. Symbols *(UPN37), +(UPN59) and -(UPN49) represent individual patients. (•) and (•) denotes PCR reactions which were either discarded (red) or used (green) in sequencing.

The PCR product was sequenced and the sequence analysed on DNASTAR Seqman II software (DNASTAR inc., USA) (Fig 3.6).
Chapter 3

Reverse Strand using Constant region primer

Forward strand using Leader sequence primer

Reverse Strand using JH region consensus primer

Forward strand using FR1 consensus primer
Reverse Strand using Constant region primer

Forward strand using Leader sequence primer

Reverse Strand using JH region consensus primer

Forward strand using FRI consensus primer
Figure 3.6 Sequence of the immunoglobulin heavy chain from a patient (UPN11), using two PCR reactions and analysed on DNASTAR Seqman II software (DNASTAR Inc., USA). The sequence is shown in 5' to 3' orientation.
Consensus sequence (Fig 3.7) was then aligned to V-base (http://vbase.mrc-cpe.cam.ac.uk/) database of all Ig genes, and the closest match selected (Fig 3.8).

\[
\begin{align*}
\text{ATGGAAGTTGGCTAGCTGCTTTTCTTTGCTTGCTATTTTTAAAGGTTGCGAGTGT} & \quad 60 \\
\text{GTGCAGCTGTTGGAGCTGGCTTTTCCTTGTGGCTATTTTAAAAGGTGTCGAGTGT} & \quad 120 \\
\text{IGTGCAAGCCTCTTGAGATCTCCTTTTACTAAGACATTATGAGGCTAGCTGGGCTG} & \quad 180 \\
\text{GGAAAGGCTGGAGTGGGCTGTCGGATATTAGTGTGATGTTGTAATGATAAAATAC} & \quad 240 \\
\text{SACTCCGTGAGACCCGCTTCACCATCTCCAGAGAATCCACCGAGATATAT} & \quad 300 \\
\text{CAAATGAAACCGCTGAGAGCGACCGACCGGCTATATTACTGTGCAGAAAGGGGT} & \quad 360 \\
\text{TGGCTACGTCCTCTTTGATGGCCTCTTTAATAATGGGCAAGGCAATGGTCCACGTC} & \quad 420 \\
\text{TCAGGGAGTGCATCCGCCCCAAA} & \quad 447
\end{align*}
\]

\begin{itemize}
\item \texttt{ontig 1}
\item \texttt{ontig Length:}
\item \texttt{verage Length/Sequence:}
\item \texttt{otal Sequence Length:}
\item \texttt{op Strand:}
\item \texttt{ottom Strand:}
\item \texttt{otal:}
\end{itemize}

\begin{itemize}
\item \texttt{350,447}
\item \texttt{(1,447)}
\item \texttt{447 bases}
\item \texttt{353 bases}
\item \texttt{1414 bases}
\item \texttt{2 sequences}
\item \texttt{2 sequences}
\item \texttt{4 sequences}
\end{itemize}

\textbf{Figure 3.7} Consensus sequence from a patient (UPN11), derived from the 2 forward and 2 reverse sequences analyses on DNASTAR Seqman II software (DNASTAR Inc., USA).
**Figure 3.8** Consensus sequence after alignment to V-base (Germline immunoglobulin database). The numbers (41-94) denote the amino acids, followed by the consensus sequence and the 5 closest matches, showing the acquired mutations. The Closest D<sub>H</sub> and J<sub>H</sub> matches and scores are also shown.
UPN11 VH3 - JH3b (4 Direct sequences from 2 PCRs)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>E</th>
<th>V</th>
<th>Q</th>
<th>L</th>
<th>E</th>
<th>S</th>
<th>G</th>
<th>G</th>
<th>L</th>
<th>V</th>
<th>Q</th>
<th>P</th>
<th>G</th>
<th>S</th>
<th>L</th>
<th>R</th>
<th>L</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>T.G</td>
<td>Z12347</td>
<td>3-23</td>
<td>DP-47/V3-23...+</td>
<td>GAG</td>
<td>GTG</td>
<td>CAG</td>
<td>CTG</td>
<td>TTG</td>
<td>GTA</td>
<td>CAG</td>
<td>CCG</td>
<td>GGG</td>
<td>GGG</td>
<td>TCC</td>
<td>CTG</td>
<td>AGA</td>
<td>CTC</td>
<td>TCC</td>
<td></td>
</tr>
<tr>
<td>T.G</td>
<td>Z12347</td>
<td>3-23</td>
<td>DP-47/V3-23...+</td>
<td>CAA</td>
<td>AAS</td>
<td>GGF</td>
<td>TFS</td>
<td>SFS</td>
<td>YAM</td>
<td>MSW</td>
<td>VRA</td>
<td>TGT</td>
<td>GCC</td>
<td>TCT</td>
<td>AGC</td>
<td>ATT</td>
<td>AGT</td>
<td>GAT</td>
<td></td>
</tr>
<tr>
<td>T.G</td>
<td>Z12347</td>
<td>3-23</td>
<td>DP-47/V3-23...+</td>
<td>CCA</td>
<td>AAG</td>
<td>GGG</td>
<td>CTG</td>
<td>AGA</td>
<td>GCC</td>
<td>GAC</td>
<td>GAC</td>
<td>ACG</td>
<td>GCC</td>
<td>ATA</td>
<td>TAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T.G</td>
<td>Z12347</td>
<td>3-23</td>
<td>DP-47/V3-23...+</td>
<td>LQM</td>
<td>MNS</td>
<td>LRA</td>
<td>EAD</td>
<td>TAV</td>
<td>YCY</td>
<td>CAA</td>
<td>KAK</td>
<td>GIG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T.G</td>
<td>X97051</td>
<td>5-12</td>
<td>D5-12/DK1</td>
<td>TGG</td>
<td>CTA</td>
<td>GCT</td>
<td>CCT</td>
<td>GAT</td>
<td>CTT</td>
<td>AAT</td>
<td>AAA</td>
<td>TGG</td>
<td>GCC</td>
<td>CAA</td>
<td>GGG</td>
<td>ACA</td>
<td>ATG</td>
<td>GTC</td>
<td>GTC</td>
</tr>
</tbody>
</table>

**In frame rearrangement**

269/294 = 91.49 %

**Figure 3.9** Formatted sequence using 2 forward and 2 reverse sequences, showing an in frame rearrangement with 25 mutations in the V\(_h\) gene (start of FW1 to the end FW3) with 91.49% similarity to the closest germline sequence, VH 3-23.

CAKGIGGWLRPLDALNKNQQGTMTVSSGSASAPK

In frame rearrangement 269/294 = 91.49 %
This is then imported into a word document, placed in a continuous sequence, translated and checked for stop codons. The CDR regions are marked and percent mutations calculated (Fig 3.9). Ig-Mutated genes were detected in 69/106 (65%) patients while 37/106 (35%) patients had Ig-Unmutated genes (Table 3.9) Germline homology ranged from 86.56% to 100%.
<table>
<thead>
<tr>
<th>S.No</th>
<th>Age</th>
<th>Sex</th>
<th>IgVH</th>
<th>%age</th>
<th>V</th>
<th>D</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52</td>
<td>F</td>
<td>M</td>
<td>94.21</td>
<td>VH 3-23</td>
<td>JH5a</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>37</td>
<td>F</td>
<td>M</td>
<td>89.29</td>
<td>VH 3-09</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>71</td>
<td>F</td>
<td>M</td>
<td>90</td>
<td>VH 3-16</td>
<td>JH5b</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>70</td>
<td>F</td>
<td>M</td>
<td>93.19</td>
<td>VH 1-69</td>
<td>JH3b</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>57</td>
<td>F</td>
<td>M</td>
<td>94.89</td>
<td>VH 3-23</td>
<td>D3</td>
<td>JH4b</td>
</tr>
<tr>
<td>6</td>
<td>57</td>
<td>M</td>
<td>U</td>
<td>99</td>
<td>VH 3-49</td>
<td>D3</td>
<td>JH6b</td>
</tr>
<tr>
<td>7</td>
<td>74</td>
<td>M</td>
<td>U</td>
<td>99.31</td>
<td>VH 3-30</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>76</td>
<td>M</td>
<td>M</td>
<td>93.93</td>
<td>VH 3-72</td>
<td>JH5b</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>67</td>
<td>F</td>
<td>M</td>
<td>97.61</td>
<td>VH 3-21</td>
<td>JH6b</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>73</td>
<td>F</td>
<td>U</td>
<td>100</td>
<td>VH 1-02</td>
<td>D5</td>
<td>JH6b</td>
</tr>
<tr>
<td>11</td>
<td>71</td>
<td>M</td>
<td>M</td>
<td>91.49</td>
<td>VH 3-23</td>
<td>JH3b</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>69</td>
<td>F</td>
<td>U</td>
<td>99</td>
<td>VH 3-30</td>
<td>JH1</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>47</td>
<td>M</td>
<td>M</td>
<td>88.43</td>
<td>VH 3-30</td>
<td>D5</td>
<td>JH2</td>
</tr>
<tr>
<td>14</td>
<td>63</td>
<td>M</td>
<td>U</td>
<td>96.67</td>
<td>VH 3-74</td>
<td>D4</td>
<td>JH4b</td>
</tr>
<tr>
<td>15</td>
<td>77</td>
<td>M</td>
<td>M</td>
<td>100</td>
<td>VH 3-30</td>
<td>D3</td>
<td>JH6b</td>
</tr>
<tr>
<td>16</td>
<td>78</td>
<td>M</td>
<td>M</td>
<td>92.17</td>
<td>VH 3-23</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>66</td>
<td>F</td>
<td>M</td>
<td>93.33</td>
<td>VH 3-72</td>
<td>D1</td>
<td>JH6b</td>
</tr>
<tr>
<td>18</td>
<td>62</td>
<td>M</td>
<td>M</td>
<td>93.87</td>
<td>VH-2</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>63</td>
<td>F</td>
<td>M</td>
<td>92.51</td>
<td>VH 3-33</td>
<td>D2</td>
<td>JH4b</td>
</tr>
<tr>
<td>20</td>
<td>58</td>
<td>F</td>
<td>M</td>
<td>96.63</td>
<td>VH 3-23</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>66</td>
<td>M</td>
<td>M</td>
<td>92.61</td>
<td>VH 3-30</td>
<td>JH6b</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>52</td>
<td>M</td>
<td>M</td>
<td>89.33</td>
<td>VH 3-15</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>68</td>
<td>F</td>
<td>U</td>
<td>98.29</td>
<td>VH 3-74</td>
<td>D6</td>
<td>JH1</td>
</tr>
<tr>
<td>24</td>
<td>50</td>
<td>M</td>
<td>M</td>
<td>94.5</td>
<td>VH 4-34</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>38</td>
<td>F</td>
<td>U</td>
<td>100</td>
<td>VH 3-30</td>
<td>JH6b</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>50</td>
<td>F</td>
<td>M</td>
<td>95.53</td>
<td>VH 3-53</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>63</td>
<td>M</td>
<td>M</td>
<td>96.9</td>
<td>VH3-53</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>73</td>
<td>M</td>
<td>U</td>
<td>99.65</td>
<td>VH 1-69</td>
<td>D3</td>
<td>JH4b</td>
</tr>
<tr>
<td>29</td>
<td>54</td>
<td>M</td>
<td>M</td>
<td>97.78</td>
<td>VH 4-b</td>
<td>D6</td>
<td>JH4b</td>
</tr>
<tr>
<td>30</td>
<td>85</td>
<td>F</td>
<td>M</td>
<td>96.98</td>
<td>VH 4-59</td>
<td>JH5b</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>67</td>
<td>M</td>
<td>U</td>
<td>99.33</td>
<td>VH 3-49</td>
<td>JH3b</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>67</td>
<td>F</td>
<td>M</td>
<td>94.89</td>
<td>VH 1-69</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>86</td>
<td>F</td>
<td>M</td>
<td>97.93</td>
<td>VH 4-34</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>53</td>
<td>F</td>
<td>U</td>
<td>99.63</td>
<td>VH 3-21</td>
<td>JH6b</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>60</td>
<td>M</td>
<td>U</td>
<td>98.28</td>
<td>VH 4-34</td>
<td>D3</td>
<td>JH6c</td>
</tr>
<tr>
<td>36</td>
<td>71</td>
<td>F</td>
<td>M</td>
<td>96.93</td>
<td>VH 1-18</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>71</td>
<td>F</td>
<td>M</td>
<td>96.67</td>
<td>VH 3-33</td>
<td>JH3b</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>62</td>
<td>M</td>
<td>M</td>
<td>88.77</td>
<td>VH 3-23</td>
<td>D3</td>
<td>JH4b</td>
</tr>
<tr>
<td>39</td>
<td>56</td>
<td>M</td>
<td>M</td>
<td>94.46</td>
<td>VH 3-07</td>
<td>JH4a</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>75</td>
<td>F</td>
<td>U</td>
<td>98.63</td>
<td>VH 3-21</td>
<td>D4</td>
<td>JH6b</td>
</tr>
<tr>
<td>41</td>
<td>90</td>
<td>M</td>
<td>M</td>
<td>92.61</td>
<td>VH1-69</td>
<td>D4</td>
<td>JH4b</td>
</tr>
<tr>
<td>42</td>
<td>59</td>
<td>F</td>
<td>U</td>
<td>100</td>
<td>VH 1-69</td>
<td>D4</td>
<td>JH5b</td>
</tr>
<tr>
<td>43</td>
<td>64</td>
<td>M</td>
<td>M</td>
<td>96.96</td>
<td>VH4-61</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>64</td>
<td>M</td>
<td>M</td>
<td>96.25</td>
<td>VH 3-48</td>
<td>D5</td>
<td>JH4b</td>
</tr>
<tr>
<td>45</td>
<td>63</td>
<td>M</td>
<td>M</td>
<td>96.93</td>
<td>VH 3-48</td>
<td>D6</td>
<td>JH3b</td>
</tr>
<tr>
<td>46</td>
<td>58</td>
<td>M</td>
<td>U</td>
<td>99.63</td>
<td>VH 4-39</td>
<td>JH3b</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>53</td>
<td>M</td>
<td>M</td>
<td>96.93</td>
<td>VH 3-09</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>60</td>
<td>F</td>
<td>M</td>
<td>96.64</td>
<td>VH 4-34</td>
<td>D1</td>
<td>JH4B</td>
</tr>
<tr>
<td>49</td>
<td>59</td>
<td>M</td>
<td>M</td>
<td>94.46</td>
<td>VH 3-07</td>
<td>D4</td>
<td>JH4a</td>
</tr>
<tr>
<td>50</td>
<td>78</td>
<td>M</td>
<td>M</td>
<td>97.61</td>
<td>VH 3-07</td>
<td>JH4b</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>79</td>
<td>M</td>
<td>U</td>
<td>98.15</td>
<td>VH 3-21</td>
<td>JH6b</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>50</td>
<td>M</td>
<td>U</td>
<td>99.31</td>
<td>VH 1-02</td>
<td>D3</td>
<td>JH5b</td>
</tr>
<tr>
<td>53</td>
<td>50</td>
<td>F</td>
<td>M</td>
<td>95.2</td>
<td>VH 3-21</td>
<td>JH6b</td>
<td></td>
</tr>
<tr>
<td>S.No</td>
<td>Age</td>
<td>Sex</td>
<td>IgVH</td>
<td>%age</td>
<td>V</td>
<td>D</td>
<td>J</td>
</tr>
<tr>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>54</td>
<td>79</td>
<td>F</td>
<td>M</td>
<td>96.3</td>
<td>VH 3-23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>51</td>
<td>M</td>
<td>U</td>
<td>99.66</td>
<td>VH 3-21</td>
<td>D3</td>
<td>JH4b</td>
</tr>
<tr>
<td>56</td>
<td>66</td>
<td>M</td>
<td>U</td>
<td>98.63</td>
<td>VH3-21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>66</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH3-48</td>
<td>D4</td>
<td>JH6c</td>
</tr>
<tr>
<td>58</td>
<td>70</td>
<td>F</td>
<td>M</td>
<td>95</td>
<td>VH 3-49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>63</td>
<td>F</td>
<td>M</td>
<td>94.4</td>
<td>VH 4-34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>60</td>
<td>F</td>
<td>M</td>
<td>95.57</td>
<td>VH 3-30</td>
<td>D3</td>
<td>JH5b</td>
</tr>
<tr>
<td>61</td>
<td>61</td>
<td>M</td>
<td>M</td>
<td>92.17</td>
<td>VH 1-18</td>
<td>D3</td>
<td>JH2</td>
</tr>
<tr>
<td>62</td>
<td>58</td>
<td>M</td>
<td>M</td>
<td>91.83</td>
<td>VH 1-69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>74</td>
<td>M</td>
<td>M</td>
<td>95.15</td>
<td>VH 4-34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>54</td>
<td>F</td>
<td>M</td>
<td>94.6</td>
<td>VH 4-34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>79</td>
<td>M</td>
<td>M</td>
<td>93.57</td>
<td>VH 5-61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>56</td>
<td>F</td>
<td>M</td>
<td>93.87</td>
<td>VH 3-23</td>
<td>D6</td>
<td>JH4b</td>
</tr>
<tr>
<td>67</td>
<td>57</td>
<td>M</td>
<td>M</td>
<td>94.56</td>
<td>VH 3-07</td>
<td>D4</td>
<td>JH4b</td>
</tr>
<tr>
<td>68</td>
<td>61</td>
<td>M</td>
<td>M</td>
<td>96.9</td>
<td>VH 3-66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>53</td>
<td>M</td>
<td>M</td>
<td>93</td>
<td>VH 3-73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>39</td>
<td>F</td>
<td>M</td>
<td>97.61</td>
<td>VH 3-09</td>
<td>D6</td>
<td>JH4b</td>
</tr>
<tr>
<td>71</td>
<td>76</td>
<td>F</td>
<td>M</td>
<td>96.2</td>
<td>VH 6-61</td>
<td>D5</td>
<td>JH4b</td>
</tr>
<tr>
<td>72</td>
<td>54</td>
<td>F</td>
<td>U</td>
<td>100</td>
<td>VH 3-33</td>
<td>D3</td>
<td>JH4B</td>
</tr>
<tr>
<td>73</td>
<td>59</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH 3-15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>69</td>
<td>M</td>
<td>U</td>
<td>99</td>
<td>VH 6-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>75</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH 1-69</td>
<td>D3</td>
<td>JH5b</td>
</tr>
<tr>
<td>76</td>
<td>70</td>
<td>M</td>
<td>M</td>
<td>94.77</td>
<td>VH 3-66</td>
<td>D4</td>
<td>JH4a</td>
</tr>
<tr>
<td>77</td>
<td>62</td>
<td>M</td>
<td>M</td>
<td>86.56</td>
<td>VH 4-59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>68</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH4-69</td>
<td>D3</td>
<td>JH6c</td>
</tr>
<tr>
<td>79</td>
<td>46</td>
<td>M</td>
<td>M</td>
<td>93.57</td>
<td>VH3-21</td>
<td>D2</td>
<td>JH6b</td>
</tr>
<tr>
<td>80</td>
<td>56</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH 1-03</td>
<td>D6</td>
<td>JH4a</td>
</tr>
<tr>
<td>81</td>
<td>49</td>
<td>F</td>
<td>M</td>
<td>93.53</td>
<td>VH 3-11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>57</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH 1-02</td>
<td>D3</td>
<td>JH5b</td>
</tr>
<tr>
<td>83</td>
<td>49</td>
<td>M</td>
<td>U</td>
<td>99.65</td>
<td>VH 1-69</td>
<td>D3</td>
<td>JH3b</td>
</tr>
<tr>
<td>84</td>
<td>75</td>
<td>F</td>
<td>M</td>
<td>91.15</td>
<td>VH 3-07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>55</td>
<td>M</td>
<td>M</td>
<td>95.23</td>
<td>VH 1-18</td>
<td>D3</td>
<td>JH4b</td>
</tr>
<tr>
<td>86</td>
<td>47</td>
<td>M</td>
<td>U</td>
<td>99.25</td>
<td>VH4-69</td>
<td>D3</td>
<td>JH6b</td>
</tr>
<tr>
<td>87</td>
<td>36</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH 4-39</td>
<td>D6</td>
<td>JH5b</td>
</tr>
<tr>
<td>88</td>
<td>44</td>
<td>M</td>
<td>M</td>
<td>97.41</td>
<td>VH 3-74</td>
<td>D3</td>
<td>JH4b</td>
</tr>
<tr>
<td>89</td>
<td>46</td>
<td>M</td>
<td>M</td>
<td>92.61</td>
<td>VH 3-74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>58</td>
<td>M</td>
<td>M</td>
<td>94.69</td>
<td>VH3-33</td>
<td>D2</td>
<td>JH5a</td>
</tr>
<tr>
<td>91</td>
<td>52</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH1-69</td>
<td>D4</td>
<td>JH6b</td>
</tr>
<tr>
<td>92</td>
<td>45</td>
<td>M</td>
<td>M</td>
<td>92.61</td>
<td>VH 3-23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>57</td>
<td>F</td>
<td>M</td>
<td>92.61</td>
<td>VH 3-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>71</td>
<td>F</td>
<td>U</td>
<td>96.96</td>
<td>VH 3-53</td>
<td>D5</td>
<td>JH6b</td>
</tr>
<tr>
<td>95</td>
<td>66</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH 4-34</td>
<td>D3</td>
<td>JH4d</td>
</tr>
<tr>
<td>96</td>
<td>75</td>
<td>M</td>
<td>M</td>
<td>96.31</td>
<td>VH 6-61</td>
<td>D3</td>
<td>JH4b</td>
</tr>
<tr>
<td>97</td>
<td>81</td>
<td>F</td>
<td>M</td>
<td>94.84</td>
<td>VH 4-69</td>
<td>D6</td>
<td>JH4b</td>
</tr>
<tr>
<td>98</td>
<td>61</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH 6-61</td>
<td>D2</td>
<td>JH6c</td>
</tr>
<tr>
<td>99</td>
<td>72</td>
<td>M</td>
<td>M</td>
<td>93.87</td>
<td>VH 3-23</td>
<td>D4</td>
<td>JH4b</td>
</tr>
<tr>
<td>100</td>
<td>73</td>
<td>F</td>
<td>M</td>
<td>93.19</td>
<td>VH 3-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>40</td>
<td>M</td>
<td>U</td>
<td>98.15</td>
<td>VH 6-01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>73</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH 6-61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>63</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH 1-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>61</td>
<td>M</td>
<td>M</td>
<td>97.01</td>
<td>VH 3-66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>47</td>
<td>M</td>
<td>U</td>
<td>100</td>
<td>VH 3-30</td>
<td>D3</td>
<td>JH3b</td>
</tr>
<tr>
<td>106</td>
<td>56</td>
<td>F</td>
<td>M</td>
<td>97.11</td>
<td>VH 3-22</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3.9** The distribution of \( V_H \), \( D_H \) and \( J_H \) genes in 106 patients. The closest match to the germline sequence is shown by percentage homology.
The most frequently encountered $V_{\text{H}}$ genes were $V3-23$ (10/106, 9.4%), $V3-30$ (10/106, 9.4%), $V1-69$ (9/106, 8.5%), $V3-21$ (8/106, 7.5%), $V4-34$ (8/106, 7.5%), $V3-07$ (5/106, 4.7%), $V4-59$ (5/106, 4.7%) and $V5-51$ (5/106, 4.7%) (Fais, Ghiotto et al. 1998; Hamblin, Davis et al. 1999) (Figure 3.10). D3 and D4 segments were used preferentially while $J_{\text{H}}4b$ was the most frequently used $J_{\text{H}}$ gene (Fais, Ghiotto et al. 1998). Five of 8 (62.5%) of $V3-21$ genes were Ig-Unmutated (Tobin, Thunberg et al. 2002). Two patients had biclonal $V_{\text{H}}$ rearrangements and in both cases the mutational status of the clones was identical.

**Figure 3.10** Distribution of different $V_{\text{H}}$ gene family usage among 106 patients. The numbers on each bar correspond to the number of patients using that particular gene family. (■) denotes unmutated and (■) mutated genes.
Twenty-three of 69 (33%) patients with Ig-Mutated genes received treatment compared to 27/37 (73%) patients with Ig-Unmutated genes, (Table 3.10).

<table>
<thead>
<tr>
<th></th>
<th>No Treatment</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig-Mutated</td>
<td>46/69 (67%)</td>
<td>23/69 (33%)</td>
</tr>
<tr>
<td>Ig-Unmutated</td>
<td>10/37 (27%)</td>
<td>27/37 (73%)</td>
</tr>
<tr>
<td>CD38 &lt;30%</td>
<td>43/70 (61.5%)</td>
<td>27/70 (38.5%)</td>
</tr>
<tr>
<td>CD38 &gt;30%</td>
<td>12/35 (34.3%)</td>
<td>23/35 (65.7%)</td>
</tr>
<tr>
<td>ZAP-70 &lt;20%</td>
<td>48/80 (60%)</td>
<td>32/80 (40%)</td>
</tr>
<tr>
<td>ZAP-70 &gt;20%</td>
<td>4/15 (27%)</td>
<td>11/15 (73%)</td>
</tr>
<tr>
<td>CD38 Unimodal -ive</td>
<td>39/57 (69%)</td>
<td>18/57 (31%)</td>
</tr>
<tr>
<td>CD38 Unimodal +ive</td>
<td>7/29 (24%)</td>
<td>22/29 (76%)</td>
</tr>
<tr>
<td>Bimodal</td>
<td>10/19 (53%)</td>
<td>9/19 (47%)</td>
</tr>
</tbody>
</table>

Table 3.10 Comparison of Treatment requirements based on IgV<sub>H</sub> status, CD38 % and ZAP-70 %
Chapter 3

The clinical outcome with regards to specific $V_H$ families was assessed, (Table 3.11).

<table>
<thead>
<tr>
<th>$V_H$ Family</th>
<th>No of Patients</th>
<th>Mutated</th>
<th>Unmutated</th>
<th>Treated</th>
<th>Not treated</th>
<th>% Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>V 1-02</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>100</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>V 1-03</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>100</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>V 1-08</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>100</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>V1-18</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>100</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>V1-69</td>
<td>9</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>44</td>
</tr>
<tr>
<td>V 2-05</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>100</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>V 3-07</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>33</td>
</tr>
<tr>
<td>V 3-09</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>100</td>
</tr>
<tr>
<td>V 3-11</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>V 3-15</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>V 3-21</td>
<td>8</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>50</td>
</tr>
<tr>
<td>V 3-22</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>V 3-23</td>
<td>10</td>
<td>10</td>
<td>3</td>
<td>7</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>V 3-30</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>40</td>
</tr>
<tr>
<td>V 3-33</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>V 3-48</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>V 3-49</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>33</td>
</tr>
<tr>
<td>V 3-53</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>V 3-66</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>V 3-72</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>V 3-73</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>V 3-74</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>75</td>
</tr>
<tr>
<td>V 4-34</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>37.5</td>
</tr>
<tr>
<td>V 4-39</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V 4-59</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>V 4-b</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V 4-61</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>40</td>
</tr>
<tr>
<td>V 5-51</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V 6-01</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

| Total        | 106            | 69      | 37        | 50      | 56          | 100       |

Table 3.11 Distribution of the different $V_H$ gene family usage and the number of patients in each category that required treatment

The TF1 for patients with Ig-Mutated genes was 120 months (95% CI, 84-234) compared to 12 months (CI, 3-26) for patients with Ig-Unmutated genes with a HR of 3.9; CI, 2.2-6.9; p=0.001). In the 87 stage ‘A’ patients, 61(70%) had Ig-Mutated genes compared to 26 (30%) with Ig-Unmutated genes. Estimated median TF1 for the
Ig-Mutated Binet stage A patients was 124 months compared to 26 months (CI, 6-86) with a HR of 3.6; CI, 1.8-7.3; $p=0.001$, for Ig-Unmutated patients, Figure 3.11 A and B.
Figure 3.11 Kaplan-Meier survival curves comparing TFI in CLL patients with Unmutated and mutated genes.

(Figure A) All patients: Median TFI for patients with Unmutated IgV<sub>H</sub> genes was 12 months compared to median TFI of 120 months for patients with mutated IgV<sub>H</sub> genes. The difference was significant at p=0.001.

(Figure B) Stage A patients: Median TFI for patients with Unmutated IgV<sub>H</sub> genes was 26 months and for patients with mutated IgV<sub>H</sub> genes 124 months respectively, p 0.001.
3.4 Results for CD38 expression

CD38 expression data was available for 105 patients. Its significance as a prognostic marker was assessed in two ways: using a cut-off value for the number of cells expressing CD38 and by looking at the pattern of expression. Its ability to predict the mutational status was also analysed.

The range of cells expressing CD38 above the Isotype control for all patients was 0.1% to 98%. We used a 30% cut-off value, patients with more than 30% CD38 positive cells were considered to be CD38 positive, and those with less than 30% CD38 positive cells were considered CD38 negative. Using this cut-off value of 30%, 35/105 (33%) of the samples were CD38 positive and 70/105 (67%) were CD38 negative.

3.2.1 CD38 expression using percentage cut-off

A CD38 expression level of 30% most accurately mirrored IgV_{H} mutational status with a concurrence of 71%, positive predictive value (PPV) of 60% and a negative predictive value (NPV) of 77%(Del Poeta, Maurillo et al. 2001; Hamblin, Orchard et al. 2002), with a 28.5% discordance rate. Different cut-off values of 20, 25 and 35% were assessed in an attempt to improve the discordance of CD38 in predicting the mutational status, but did not show any improvement (Table 3.12).

<table>
<thead>
<tr>
<th>Cut-off value for CD38</th>
<th>Discordance %</th>
<th>Sensitivity %</th>
<th>Specificity %</th>
<th>Positive predictive value %</th>
<th>Negative predictive value %</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>32.38</td>
<td>67.57</td>
<td>67.65</td>
<td>53.19</td>
<td>79.31</td>
</tr>
<tr>
<td>25%</td>
<td>31.42</td>
<td>62.16</td>
<td>72.06</td>
<td>54.76</td>
<td>77.78</td>
</tr>
<tr>
<td>30%</td>
<td>28.57</td>
<td>56.76</td>
<td>79.41</td>
<td>60</td>
<td>77.14</td>
</tr>
<tr>
<td>35%</td>
<td>28.58</td>
<td>51.35</td>
<td>82.35</td>
<td>61.29</td>
<td>75.68</td>
</tr>
</tbody>
</table>

Table 3.12 Sensitivity and specificity of CD38 for predicting mutational status using different cut-off values
The CD38 expression identified patient groups with a significant difference in median TFI. Twenty-seven of 70 (38.5%) CD38-ve patients required treatment compared to 23/35 (65.7%) of CD38+ve patients Table 3.13.

<table>
<thead>
<tr>
<th></th>
<th>No Treatment</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig-Mutated</td>
<td>46/69 (67%)</td>
<td>23/69 (33%)</td>
</tr>
<tr>
<td>Ig-Unmutated</td>
<td>10/37 (27%)</td>
<td>27/37 (73%)</td>
</tr>
<tr>
<td>CD38 &lt;30%</td>
<td>43/70 (61.5%)</td>
<td>27/70 (38.5%)</td>
</tr>
<tr>
<td>CD38 &gt;30%</td>
<td>12/35 (34.3%)</td>
<td>23/35 (65.7%)</td>
</tr>
<tr>
<td>ZAP-70 &lt;20%</td>
<td>48/80 (60%)</td>
<td>32/80 (40%)</td>
</tr>
<tr>
<td>ZAP-70 &gt;20%</td>
<td>4/15 (27%)</td>
<td>11/15 (73%)</td>
</tr>
<tr>
<td>CD38 Unimodal -ive</td>
<td>39/57 (69%)</td>
<td>18/57 (31%)</td>
</tr>
<tr>
<td>CD38 Unimodal +ive</td>
<td>7/29 (24%)</td>
<td>22/29 (76%)</td>
</tr>
<tr>
<td>Bimodal</td>
<td>10/19 (53%)</td>
<td>9/19 (47%)</td>
</tr>
</tbody>
</table>

Table 3.13 Comparison of treatment requirements based on IgV\textsubscript{H} status, CD38 % and ZAP-70 %

The median TFI for CD38-ve patients was 114 months (CI, 64-234) compared to a median TFI of 13 (CI, 4-86) months for CD38+ve patients, HR, 2.8; CI, 1.5-5.1; $p=0.001$ (Figure 3.12).
Figure 3.12 Kaplan-Meier survival curves comparing TFI in CD38-ve and CD38+ve CLL patients. (All patients): Median TFI for CD38+ve CLL was 13 months and median TFI for CD38-ve CLL was 114 months. The difference was significant at p=0.001.

Sixty-two (72%) stage A patients were CD38-ve and 24/86 (28%) were CD38+ve. Twenty-one (33%) CD38-ve patients compared to 11 (46%) CD38+ve were treated with a TFI of 120 months (CI, 97-234) for CD38-ve patients and 32 months for CD38+ patients, HR, 2.4; CI, 1.4-5.3; \( p=0.02 \) Figure 3.13.
3.2.2 Pattern of expression

Patients were divided into 3 groups by CD38 expression pattern (Ghia, Guida et al. 2003): 57/105 were homogenously negative, 29/105 homogenously positive and 19/105 had a bimodal expression pattern (Figure 3.14). Forty-six of 68 (68%) patients with Ig-Mutated genes were CD38-ve, 12/68 (18%) CD38+ve and 10/68 (14%) had bimodal CD38 expression. Eleven of 37 (30%) patients with Ig-Unmutated genes were CD38-ve, 17/37 (46%) CD38+ve, and 9/37 (24%) had bimodal expression. As the pattern of expression divided the patients into three groups it could
Figure 3.14 Different patterns of CD38 expression in CLL cells. Analysis of surface CD38 expression in CD5+ve/19+ve gated lymphocytes in CLL patients. Panels (A,C,E) showing CD19 and CD38 intensities on the x-axis and y-axis respectively. Panels (B,D,F) showing CD38 expression in profile, with counts on the y-axis and intensity on the x-axis. Panels A and B are from a patient with two distinct populations of cells, one negative and one positive, bimodal expression, panels C and D a positive patient while panels E and F represent a CD38 negative patient.
not be used for predicting the mutational status. When only the unimodal positive and unimodal negative groups were considered, the discordance rate was 26.7%. Eighteen (31%) unimodal CD38-ve patients were treated with a TFI of 124 months whereas 22 (76%) of the unimodal CD38+ve group were treated with a TFI of 4 months. Nine (47%) patients in the bimodal group required treatment with a TFI of 102 months Figure 3.15.

Figure 3.15  Kaplan-Meier survival curves comparing TFI in CLL patients according to the pattern of CD38 expression. (All patients): Median TFI for unimodal +ve CLL patients was 4 months, for unimodal -ve CLL patients 124 months and for patients with bimodal expression 102 months.

There were 86 stage A patients. Fifteen (15/52, 28%) unimodal CD38-ve patients were treated with a TFI of 124 months whereas 10/18 (56%) of the unimodal CD38+ve group were treated with a TFI of 6 months. Seven out of 16 (44%) patients in the bimodal group required treatment with a TFI of 102 months Figure 3.16.
3.5 Results for ZAP-70 expression

ZAP-70 expression results were available for 95 patients. The results were analyzed to assess the prognostic significance of ZAP-70 and its ability to predict mutational status. We used a 20% cut-off value; patients with more than 20% ZAP-70 positive cells were considered to be ZAP-70 positive, and those with less than 20% positive cells were ZAP-70 negative.
Eighty patients (84%) were ZAP-70-ve compared to 15 (16%) who were ZAP-70+ve. A ZAP-70 expression level of 20% most accurately predicted IgV<sub>H</sub> mutational status with a concurrence of 77%, positive predictive value (PPV) of 80% and a negative predictive value (NPV) of 76%. Sixty-one of 64 (95%) patients with Ig-Mutated genes were ZAP-70-ve and 12/31 (39%) patients with Ig-Unmutated genes were ZAP-70+ve. The assay gave discordant results in 22/95 (23.15%) patients, with 19 (61%) high-risk patients by mutational analysis (Ig-Unmutated) misclassified as low risk by ZAP-70 expression. This discordance did not improve with different expression cut off levels.

### Sensitivity and specificity of ZAP-70 for predicting IgV<sub>H</sub> mutational status

<table>
<thead>
<tr>
<th>Cut-off value for ZAP-70</th>
<th>Discordance %</th>
<th>Sensitivity %</th>
<th>Specificity %</th>
<th>Positive predictive value %</th>
<th>Negative predictive value %</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>24.5</td>
<td>51.61</td>
<td>89.06</td>
<td>69.57</td>
<td>79.17</td>
</tr>
<tr>
<td>15%</td>
<td>24.1</td>
<td>45.16</td>
<td>92.19</td>
<td>73.68</td>
<td>77.63</td>
</tr>
<tr>
<td>20%</td>
<td>23.15</td>
<td>38.71</td>
<td>95.31</td>
<td>80</td>
<td>76.25</td>
</tr>
<tr>
<td>25%</td>
<td>27.36</td>
<td>22.58</td>
<td>96.88</td>
<td>77.78</td>
<td>72.09</td>
</tr>
</tbody>
</table>

**Table 3.14** Sensitivity and specificity of ZAP-70 for predicting the mutational status using different cut-off values

Thirty two of 80 (40%) ZAP-70-ve patients were treated compared to 11/15 (73%) ZAP-70+ve patients, Table 3.15.
<table>
<thead>
<tr>
<th></th>
<th>No Treatment</th>
<th>Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig-Mutated</td>
<td>46/69 (67%)</td>
<td>23/69 (33%)</td>
</tr>
<tr>
<td>Ig-Unmutated</td>
<td>10/37 (27%)</td>
<td>27/37 (73%)</td>
</tr>
<tr>
<td>CD38 &lt;30%</td>
<td>43/70 (61.5%)</td>
<td>27/70 (38.5%)</td>
</tr>
<tr>
<td>CD38 &gt;30%</td>
<td>12/35 (34.3%)</td>
<td>23/35 (65.7%)</td>
</tr>
<tr>
<td>ZAP-70 &lt;20%</td>
<td>48/80 (60%)</td>
<td>32/80 (40%)</td>
</tr>
<tr>
<td>ZAP-70 &gt;20%</td>
<td>4/15 (27%)</td>
<td>11/15 (73%)</td>
</tr>
<tr>
<td>CD38 Unimodal -ive</td>
<td>39/57 (69%)</td>
<td>18/57 (31%)</td>
</tr>
<tr>
<td>CD38 Unimodal +ive</td>
<td>7/29 (24%)</td>
<td>22/29 (76%)</td>
</tr>
<tr>
<td>Bimodal</td>
<td>10/19 (53%)</td>
<td>9/19 (47%)</td>
</tr>
</tbody>
</table>

**Table 3.15** Comparison of Treatment requirements based on IgV<sub>H</sub> status, CD38 % and ZAP-70 %

The median TFI for ZAP-70-ve patients was 102 months (CI, 64-137) and for ZAP-70+ve patients was 5 months, HR, 3.2; CI, 1.5-6.8; \( p=0.002 \) Figure 3.17.
Seventy-two (88%) stage A patients were ZAP-70-ve and 10/82 (12%) were ZAP-70+ve. Twenty-five of 72 (35%) ZAP-70-ve patients required treatment at a median TFI of 120 months compared to 6/10(60%) ZAP-70+ve needing treatment with a median TFI of 16 months, HR, 3.4; CI, 1.4-8.7; \( p=0.01 \) Figure 3.18.
Figure 3.18 Kaplan-Meier survival curves comparing TFI in ZAP-70-ve and ZAP-70+ve CLL patients. (Stage A patients): The median TFI for ZAP-70+ve CLL is 16 months and the median TFI for ZAP-70-ve CLL is 120 months, with a p value of 0.01.
CHAPTER 4
DISCUSSION
Chronic lymphocytic leukaemia (CLL) is the most common leukaemia in the Western world, accounting for nearly 24% of all leukaemias (Jemal, Murray et al. 2003). It is estimated that 7300 new cases were diagnosed in 2003 in the United States (Jemal, Murray et al. 2003) and there were 4400 deaths due to CLL. At diagnosis many patients are asymptomatic but subsequently develop symptoms and die from the disease or its complications.

The most important initial challenge in patients with CLL is defining prognosis, early recognition and management of complications and designing of risk-stratified management. The clinical course of CLL is extremely variable. The majority of patients with CLL do not require therapy at diagnosis, one third never require therapy and in these patients the disease does not significantly affect life expectancy (Dighiero and Binet 2000; Guarini, Gaidano et al. 2003). In contrast, approximately one third of patients have aggressive CLL, which requires treatment soon after diagnosis and these patients have a median survival of between 18 months and 3 years. The remaining one third of patients do not need therapy at diagnosis, but may need treatment at a later stage (Byrd, Murphy et al. 2001).

CLL is not curable with current therapy and initiating chemotherapy following diagnosis in asymptomatic patients does not alter survival (Dighiero, Maloum et al. 1998) therefore the current standard of care is to treat CLL when the patient becomes symptomatic or the disease has progressed (Cheson, Bennett et al. 1996). Newer and more effective treatment protocols with curative potential are being developed. The possibility of intervention at an early stage of disease, when the disease burden is minimal and before clonal evolution occurs, may be more effective at prolonging survival and decreasing disease related mortality. If clinicians are going to adopt this approach, we have to have a robust method of risk analysis so that potentially toxic therapy in patients with a small risk of disease progression can be avoided.

In this study we present data from a series of 106 CLL patients using novel prognostic markers, including cytogenetic abnormalities, mutational status of the immunoglobulin heavy chain and immunophenotypic markers CD38 and ZAP-70 for risk stratification.
The demographic features of the patients included in the study are similar to those in the published literature. The median age at presentation was 62 with a male predominance. The male to female ratio was 1.65:1. A median age of 65 and a male to female ratio of 2:1 has been widely quoted in the literature. There were more patients in stage A as compared to stage B and C at the time of entry into the study.

Clinical staging systems based on disease burden (lymphadenopathy, splenomegaly, hepatomegaly and cytopenias), developed by Rai (Rai, Sawitsky et al. 1975) and Binet (Binet, Auquier et al. 1981) were early prognostic tools in CLL developed in the late 1970. The fact that these systems are still widely used is an indication of their clinical usefulness and robust nature. These two systems have allowed patient stratification into different risk groups at presentation and have also helped to delineate the natural history of CLL. The clinical importance of stratifying patients according to these two staging systems can be gathered by the different median survival times from the time of diagnosis in the different subgroups (Table 4.1).

<table>
<thead>
<tr>
<th>Modified Rai Stage</th>
<th>Median survival (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (0)</td>
<td>&gt;10</td>
</tr>
<tr>
<td>Intermediate (I,II)</td>
<td>7</td>
</tr>
<tr>
<td>High (III,IV)</td>
<td>1.5</td>
</tr>
<tr>
<td>Binet Stage</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>&gt;10</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.1 Survival in CLL according to Clinical Stage.

Eighty-two (82%) of the patients in the study had stage A disease, with 18% having stage B and C disease. The majority of patients with stage B (83%) and all patients in stage C received treatment whereas most of the patients with stage A disease (62%) remained untreated. The median treatment-free interval (TFI), for patients who received treatment, in stage A was 16 months, as compared to 2 months for stage B.
Chapter 4

disease and 0.5 months for patients with stage C disease. Because stage B and C
disease itself are indications for treatment, the use of TFI as an indicator of disease
aggression is not valid in stage B and C disease. In this cohort however, staging alone
failed to predict the aggressive nature of the disease in 38% of stage A patients, who
required treatment.

Some morphological variants of CLL have been described such as (a) CLL/PLL in
which there are more than 10% but less than 55% prolymphocytes in the peripheral
blood and (b) mixed cell type CLL characterized by a spectrum of small to large
lymphocytes, with <10% prolymphocytes (Bennett, Catovsky et al. 1989). CLL cases
classified according to morphology show that atypical CLL accounts for 23% of all
CLL cases and are usually diagnosed in advanced clinical stage disease, are more
likely to be associated with trisomy 12 and have a shorter survival (Criel, Verhoef et
al. 1997).

The majority of patient's in this cohort had typical morphology with only 10% of
cases classified with atypical CLL. An equal percentage of patients with typical and
atypical morphology required treatment. Because of the small number of patients
with atypical morphology, the clinical significance of sub-classifying CLL patients by
morphology was difficult to assess. Of the 10 patients with atypical morphology, three
had trisomy 12 and two patients had 11q deletion, indicating that morphology may be
a surrogate marker for abnormal cytogenetics.

The pattern of bone marrow infiltration in CLL has been shown by various groups to
have prognostic importance (Rozman, Montserrat et al. 1984; Desablens, Claisse et al.
1989). In a retrospective analysis, patients with a non-diffuse pattern of infiltration
had a better survival rate than patients with a diffuse pattern (Rozman, Montserrat et
al. 1984). Conversely, in a prospective analysis of untreated patients, no difference
was found in survival of patients with stage A or B disease with diffuse and non-
diffuse infiltration (Desablens, Claisse et al. 1989). These contradictory findings
suggest that the prognostic significance of the pattern of bone marrow infiltration adds
little to the clinical stage alone. Bone marrow evaluation is desirable prior to therapy
to aid assessment but is not essential for diagnosis (Cheson, Bennett et al. 1996).
Seventy-eight patients had bone marrow trephine results and the majority of these patients, 47/78 (60%) had diffuse marrow involvement. Interestingly although the majority of the patients with diffuse marrow involvement had stage A disease 33/47(68%), 14 patients had advanced stage B or C disease. Of the 19 patients with an interstitial pattern of marrow involvement, none had stage B or C disease, while only one of 12 patients with a nodular pattern of marrow involvement had stage B disease confirming the association of diffuse marrow involvement with advanced disease. More patients with diffuse marrow involvement received treatment (68%) compared to patients with interstitial or nodular marrow involvement where only 32% and 33% respectively were treated. If only stage A patients are compared, 17 of 32 patients (53%) with diffuse marrow involvement compared to 6/19(32%) with interstitial marrow involvement and 3/11(27%) with a nodular pattern received treatment, emphasising that a diffuse pattern of marrow involvement is associated with more aggressive disease and a higher disease burden.

The value of $\beta_2$m as an independent prognostic marker within Rai defined stages (Keating, Lerner et al. 1995) or when compared to other prognostic parameters, has been shown in a number of studies (Knauf, Langenmayer et al. 1997). A prospective trial of 106 untreated patients did not show that $\beta_2$m was a significant predictor of survival in a multivariate analysis which controlled for stage and lymphocyte doubling time (LDT) (Molina, Vitelli et al. 1999). Lactate dehydrogenase (LDH) has similarly been widely used as a prognostic marker in non-Hodgkin's lymphomas, as a marker of disease burden and cellular mitotic activity.

Thirty-nine of 72 patients for whom results were available had a normal level of $\beta_2$m and 79 out of 98 patients had a normal LDH. Significantly more patients in both groups with elevated results were treated compared to patients who had normal results for both parameters. These markers appear to correlate with tumour burden in patients with CLL and therefore one would expect high levels in patients with advanced stage disease. In early stage patients they add little extra discriminatory power compared to clinical staging alone.

Conventional cytogenetics studies show that between 30 and 40% of patients with CLL have chromosomal abnormalities (Han, Ozer et al. 1984; Crossen 1989). The
frequency of chromosomal aberrations detected by FISH in CLL patients is variable and can be as high as 82% (Döhner, Stilgenbauer et al. 2000) or as low as 69% (Oscier, Gardiner et al. 2002). The relative frequency of patients with chromosomal abnormality may depend on the molecular cytogenetic methods and panel of FISH markers used.

We evaluated the most common current chromosomal aberrations in peripheral blood cells of patients with CLL using interphase FISH. We compared the frequency of four cytogenetic abnormalities, namely del(13q14), trisomy 12, del(11q22-23) and del(17p) in our patients, with the published data in literature (Table 4.2) and also their ability to predict treatment requirements for patients.

<table>
<thead>
<tr>
<th>Chromosomal Abnormality</th>
<th>Frequency, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deletion 13q14</td>
<td></td>
</tr>
<tr>
<td>Döhner et al., 1997</td>
<td>45</td>
</tr>
<tr>
<td>Döhner et al., 2000</td>
<td>55</td>
</tr>
<tr>
<td>Dewald et al.,</td>
<td>47</td>
</tr>
<tr>
<td>Sindelarova et al., 2004</td>
<td>54</td>
</tr>
<tr>
<td>Present Study</td>
<td>51</td>
</tr>
<tr>
<td>Trisomy 12</td>
<td></td>
</tr>
<tr>
<td>Döhner et al., 1997</td>
<td>15</td>
</tr>
<tr>
<td>Döhner et al., 2000</td>
<td>16</td>
</tr>
<tr>
<td>Dewald et al.,</td>
<td>19</td>
</tr>
<tr>
<td>Sindelarova et al., 2004</td>
<td>16</td>
</tr>
<tr>
<td>Present Study</td>
<td>11</td>
</tr>
<tr>
<td>Deletion 11q</td>
<td></td>
</tr>
<tr>
<td>Döhner et al., 1997</td>
<td>20</td>
</tr>
<tr>
<td>Döhner et al., 2000</td>
<td>18</td>
</tr>
<tr>
<td>Dewald et al.,</td>
<td>11</td>
</tr>
<tr>
<td>Sindelarova et al., 2004</td>
<td>12</td>
</tr>
<tr>
<td>Present Study</td>
<td>6</td>
</tr>
<tr>
<td>Deletion 17p</td>
<td></td>
</tr>
<tr>
<td>Döhner et al., 1997</td>
<td>10</td>
</tr>
<tr>
<td>Döhner et al., 2000</td>
<td>7</td>
</tr>
<tr>
<td>Dewald et al.,</td>
<td>6</td>
</tr>
<tr>
<td>Sindelarova et al., 2004</td>
<td>16</td>
</tr>
<tr>
<td>Present Study</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 4.2 Incidence of chromosomal abnormalities in CLL identified by FISH

Chromosomal abnormalities were detected in 72% of the patients tested. In 62 patients there was one aberration and 10 patients had two abnormalities. No patient had more than two abnormalities and 28% had a normal karyotype.
Deletion of 13q14 is reported to be the most common chromosomal abnormality in CLL patients and is found in at least 55% of the patients (Dohner, Stilgenbauer et al. 2000). The loss of genetic material at 13q14 in CLL cells has not been shown to result in a change in genes coding for proteins in patients with CLL, but preliminary evidence suggests a possible loss of the coding region for a micro-RNA (Calin, Dumitru et al. 2002). The deletion at 13q14 is associated with the best prognosis in CLL (Dohner, Stilgenbauer et al. 2000). Of the 61 patients with 13q deletion, it was the sole abnormality in 51 (51%) patients. Twenty-one (41%) of 51 patients with 13q14 deletion were treated, compared to 30 (59%) who were not treated and patients had a long TFI of 110 months.

In our cohort, trisomy 12 was found in 11% of patients which is in keeping with the published rate of 10-20% of CLL cases. It is probably acquired during the course of the disease and is considered to be associated with a poor prognosis by some (Hamblin 1997; Dohner, Stilgenbauer et al. 2000) but not all studies (Stilgenbauer, Bullinger et al. 2002). In this study, the median TFI for this group of patients was 41 months, which is considerably shorter than for 13q14 deletion patients or those with a normal karyotype, who had a median TFI of 87 months.

Deletion of 11q is the second most frequent genetic change in patients with CLL and is found in 18% of the patients (Dohner, Stilgenbauer et al. 2000). Intensive searching for the tumour suppressor gene locus revealed the crucial gene locus to be the ATM gene, which is localised at 11q 22.3 and is a component of the p53 pathway mechanism of cell-cycle and apoptosis control (Dohner, Stilgenbauer et al. 2000; Dewald, Brockman et al. 2003). The 11q deletion is more common in young patients and is associated with aggressive disease and bulky adenopathy (Dohner, Stilgenbauer et al. 2000; Oscier, Gardiner et al. 2002). The pattern of ATM inactivation in CLL and other tumours (loss of one allele and mutation of the remaining allele) is suggestive of a classic tumour suppressor inactivation mechanism. Cells respond to genotoxic stress by activating cell cycle check points and DNA repair pathways with the general aim of preventing cells with damaged DNA from survival and proliferation. ATM protein kinase activity increases several fold in response to ionizing radiation (IR) resulting in the phosphorylation of several proteins in the DNA repair pathway such as p53, c-Abl, Nbs 1 and Chk2 kinase (Baskaran, Wood et al. 1997; Shafman, Khanna et al.
1997; Banin, Moyal et al. 1998; Canman, Lim et al. 1998; Matsuoka, Huang et al. 1998; Cortez, Wang et al. 1999; Lim, Kim et al. 2000; Wu, Ranganathan et al. 2000). Thus, through its kinase activity, ATM exerts control over progression through cell cycle allowing the cell to repair. Loss of function therefore may initiate or accelerate leukaemic progression.

The 11q deletion was the third most common abnormality detected in our cohort of patients with 6% patients showing the abnormality. The majority of these patients (83%) required treatment with a short median TFI of 5 months confirming the poor prognosis associated with this deletion.

The deletion of the tumour suppressor gene p53 at chromosome 17p13 has been variably reported in 6%-16% in CLL series. The p53 gene plays an integral role in inducing apoptosis or cell-cycle arrest after DNA damage and its inactivation by deletion or mutation is predictive of aggressive disease with a short TFI (Oscier, Gardiner et al. 2002; Stilgenbauer, Bullinger et al. 2002). More importantly p53 deletion or mutation results in resistance and poor response to chemotherapy/fludarabine treatment as it requires an intact p53 pathway to induce cell death (Rosenwald, Chuang et al. 2004). Disruption of chemotherapy induced apoptosis can also occur through mechanisms other than p53 mutation. A functional assay based on impaired up-regulation of p53 and of the p53-dependent protein p21 (CIP1/WAF1) after exposure to ionizing radiation (IR) found that p53 dysfunction is present in up to 26% of patients with CLL, either secondary to p53 deletion (9%) or from inactivation of p53 regulator genes, such as the ATM gene on chromosome 11q (Pettitt, Sherrington et al. 2001). Only 4(%) of our patients had a 17p deletion which may reflect the small patient series and short follow-up, as the 17p abnormality is acquired during clonal evolution and might not be detected at initial presentation. The clinical significance of 17p deletion was evident however, as 75% of the patients required treatment with a median TFI of 11 months. Despite an overall poor prognosis, variation has been shown in the survival of patients presenting with the p53 mutation (Oscier, Gardiner et al. 2002). The size of the mutant p53 clone is probably an important factor in determining the outcome. For diagnostic purpose >5% lymphocytes with del 17p is considered a positive result, analyses from the LRF CLL4 trial however revealed that the response to therapy correlated best when 20%
17p deleted lymphocytes is used as the critical threshold (Catovsky 2004). The importance of p53 dysfunction lies in its ability to confer resistance to alkylating agents and the current gold standard of CLL therapy, fludarabine. The aim of treating patients with p53 abnormalities is to use drugs which do not act through the p53 induced apoptotic pathway. Such agents include the humanized anti-CD52 monoclonal antibody alemtuzumab (Campath-1H) and methylprednisolone, as both these drugs kill the leukaemic cells through p53 independent mechanisms. The use of reduced intensity conditioning (RIC) transplants in patients with p53 abnormalities has shown encouraging results and should be considered as a therapeutic option in biologically young and fit patients.

In accord with other studies (Dohner, Stilgenbauer et al. 2000; Chena, Arrossagaray et al. 2003), in 10 patients we found more than one chromosomal abnormality: trisomy 12 and del(13q14), del(11q22-23) and del(13q14), del(17p) and del(13q14). This suggests that CLL is not a genetically stable disease and clonal evolution can occur. The prognostic importance of the combined changes remains unclear, but it has been suggested that they have an adverse affect on survival (Chena, Arrossagaray et al. 2003).

We compared the relationship of cytogenetic abnormalities to other prognostic markers. The majority of patients with favourable cytogenetic abnormalities, del(13q14) had mutated IgV_H genes (82%) and were CD38 negative (82%) and ZAP70 negative (90%). On the other hand, the majority of patients with unfavourable cytogenetic abnormalities such as trisomy 12, del(11q) and del(17p) expressed CD38. Fifty (%) of the patients with trisomy 12 and del(17p) had unmutated IgV_H genes and all the patients with del(11q) had unmutated IgV_H genes.

In the normal immune system, somatic hypermutation of the variable regions of the heavy and light chains of the immunoglobulin (Ig) molecule is crucial for production of high affinity antibodies. After exposure to antigen in the lymph node, responding B-cells enter the germinal follicle and undergo somatic hypermutation of the variable region of the Ig gene sequence with positive selection of the mutations which increase Ig affinity for antigen (Feldmann 1998). IgV_H mutation is thus a physiologic marker of antigen exposure and passage through the germinal centre which defines mature.
antigen-experienced B-cells. Historically, all CLL B-cells were believed to have undergone leukaemic transformation at the naïve B lymphocyte stage, before germinal centre antigen exposure and somatic mutation of their Ig genes. Subsequently several studies found that approximately 50% of CLL samples exhibit somatic mutation of their Ig chains, indicating that they develop from postgerminal-center, “memory” B-cells (Fais, Ghiotto et al. 1998). In 1999, 2 groups of investigators demonstrated that patients with a memory cell immunophenotype with mutated IgV<sub>H</sub> genes had a favourable outcome and low probability of developing progressive disease, whereas those with unmutated IgV<sub>H</sub> genes were more likely to develop progressive disease and have a shorter survival (Damle, Wasil et al. 1999; Hamblin, Davis et al. 1999). These studies defined patients with more than 2% difference in nucleotide sequences from germline cell sequences as having a mutated clone and those with 2% or less, as unmutated.

Sixty-nine (65%) of the patient series had mutated IgV<sub>H</sub> genes and 37(35%) unmutated IgV<sub>H</sub> genes using the standard 98% sequence homology definition of the 106 patients analysed. This compares favourably with the published literature (Damle, Wasil et al. 1999; Hamblin, Davis et al. 1999; Maloum, Davi et al. 2000) and also reflects the unbiased selection criteria used in this study, despite our centre being a tertiary referral transplant centre. We also confirm a biased use of certain V<sub>H</sub> genes in CLL with five of 29 genes utilized in nearly half of the cases. Except for the V<sub>H</sub> 1-69 (9/106), which is not over-used by normal B-cells (Brezinschek, Brezinschek et al. 1998), but has been consistently reported as one of the most commonly used V<sub>H</sub> genes in CLL, usually in the unmutated form, the use of other V<sub>H</sub> genes in our study was similar to the normal B-cell Ig repertoire (Brezinschek, Foster et al. 1997). The reason for the over-representation of V<sub>H</sub> 1-69 in CLL is unknown. More recently it has been shown that V<sub>H</sub> 3-21 genes are over-represented in mutated CLL, but that these cases have a poor prognosis similar to those with unmutated IgV<sub>H</sub> genes (Tobin, Thunberg et al. 2002). In this study 8/106 (7.5%) patients used V<sub>H</sub> 3-21 genes but most (5/8) had unmutated IgV<sub>H</sub> genes and 50% of the patients in this group required treatment. Some patients with mutated V<sub>H</sub> 3-21 genes have sequence homologies very close to the cut-off value of 98%, which is used as the standard for classifying patients into mutated and unmutated groups. This is an arbitrary value chosen to take into account genetic polymorphisms in IgV<sub>H</sub> gene families. The degree of homology in the
germline sequence required to separate the two clinical subsets of CLL remains controversial. The original two studies (Damle, Wasil et al. 1999; Hamblin, Davis et al. 1999) demonstrated the prognostic significance of mutational analysis using a 98% cut-off, others have varied from a 95% to 97% cut-off to provide the best discrimination (Stilgenbauer 2001; Krober, Seiler et al. 2002; Lin, Sherrington et al. 2002). The use of PCR primers within the framework 1 region, rather than the IgV_H gene leader sequence may have resulted in an apparent increase in the percentage of mutations in some of these studies. Whether changing the cut-off would make any substantial difference is not known at present, and would require review on a large cohort of CLL patients. The use of different D_H and J_H segments did not appear to influence the clinical outcome.

Patients with unmutated IgV_H genes had a more aggressive disease and shorter TFI than those with somatic mutations in our series. They were more likely to have advanced Binet stage, atypical morphology, diffuse marrow involvement and require more than one line of treatment. Significantly more patients with unmutated IgV_H genes had karyotypic abnormalities associated with a poor prognosis such as trisomy 12, 11q deletion and 17p deletion as well as a high a CD38 and ZAP-70 expression. The TFI for patients with mutated IgV_H genes was 120 months (95% Confidence Interval (CI), 84-234) and 12 months (CI, 3-26) for patients with unmutated IgV_H genes, HR, 3.9; CI, 2.2-6.9; \( p=0.001 \).

The most important factor in measuring the reliability of new prognostic marker in CLL depends on how accurately the factor can risk stratify early-stage asymptomatic patients. These patients are difficult to manage because they can have either an indolent course, requiring no or minimum therapy, or a rapidly progressive course requiring aggressive treatment. Accurate risk stratification is therefore essential for risk-adapted management. We have demonstrated a significantly different median TFI for the Ig-mutated Binet stage A patients, 124 months compared to 26 months (CI, 6-86), HR, 3.6; CI, 1.8-7.3; \( p=0.001 \), for Ig-Unmutated patients. IgV_H mutation analysis was therefore able to segregate patients with early-stage disease into groups with significantly different outcome.
The major limitation of the clinical use of IgV<sub>H</sub> mutation status in CLL is the technical complexity of the analysis. The assay is only available in research laboratories, is expensive and time-consuming and thus has limited applicability to routine clinical care. The search for more accessible surrogate markers of mutation status led to the discovery that surface CD38 expression and cytoplasmic ZAP-70 expression have some value as independent prognostic markers in CLL.

CD38, a multifunctional surface protein which is both an ectoenzyme and receptor, is expressed in both early B-cell ontogeny and terminally differentiated B-cells and plasma cells (Deaglio, Capobianco et al. 2003). The functional role of CD38 in CLL remains unclear. CD38 expression can be measured routinely by flow cytometry. The proportion of CD38 positive B-cells is a continuous variable in CLL patient samples and different studies use markedly different cut-off values between 7% and 30% to define CD38 positive status (Damle, Wasil et al. 1999; Ibrahim, Keating et al. 2001; Hamblin, Orchard et al. 2002; Krober, Seiler et al. 2002; Lin, Sherrington et al. 2002). The choice of a cut-off value in different studies was based on both its ability to predict survival and accuracy in predicting the mutational status. The proportion of CLL cells which express CD38 has been widely used as a predictor of prognosis (Damle, Wasil et al. 1999; Ibrahim, Keating et al. 2001). In multivariate analyses which include CD38 expression, cytogenetic abnormalities and mutational status of IgV<sub>H</sub> genes; CD38 expression loses significance as an independent prognostic marker (Oscier, Gardiner et al. 2002). More recently its value as a surrogate marker for IgV<sub>H</sub> mutation status has also been questioned. One early report showed that CLL cells with an increased CD38 expression were more likely to have unmutated IgV<sub>H</sub> genes and that CD38 status had a prognostic value similar to IgV<sub>H</sub> mutation status (Damle, Wasil et al. 1999) and could therefore be used as a surrogate marker for predicting the mutational status. Subsequent studies did not confirm the predictive power of CD38 as a surrogate marker for IgV<sub>H</sub> status (Hamblin, Orchard et al. 2000; Jelinek, Tschumper et al. 2001; Hamblin, Orchard et al. 2002; Krober, Seiler et al. 2002).

We therefore evaluated CD38 expression to clarify (1) its value as an independent prognostic marker and (2) its ability to predict IgV<sub>H</sub> gene mutational status.
In this study we confirmed the observation that a significant correlation between CD38 expression and the presence of unmutated IgV_H genes exists (Damle, Wasil et al. 1999). However, this correlation is not strong enough to use CD38 as a surrogate marker. In this study, CD38 expression and IgV_H mutations give discordant results in 28% of patients, which is in keeping with the published data (Hamblin, Orchard et al. 2002). We were unable to show any improvement in the discordances between the two prognostic factors by using different cut-off values of 20%, 25% and 35% CD38 expression. We have demonstrated that CD38 remains a clinically useful prognostic marker (D'Arena, Musto et al. 2001; Del Poeta, Maurillo et al. 2001; Durig, Naschar et al. 2002). A CD38 expression of more than 30% was useful in predicting disease progression with a median TFI of 114 months (CI, 64-234) for CD38-ve patients compared to a median TFI of 13 months (CI, 4-86), HR, 2.8; CI, 1.5-5.1; \( p=0.001 \) for CD38+ve patients. More importantly, CD38 expression remained statistically significant when used in stage A patients only. We found a significantly different median TFI of 120 months (CI, 97-234) for CD38-ve patients and 34 months for CD38+ve patients, HR, 2.4; CI, 1.4-5.3; \( p=0.02 \).

Recent work indicates that CD38 can acquire an abnormal signalling function depending on its relationship to the B-cell receptor and this may be the biological basis for its prognostic usefulness in CLL (Deaglio, Capobianco et al. 2003). One can therefore hypothesize that any degree of CD38 expression will result in abnormal signalling and be associated with a poorer outcome compared to CD38-ve CLL. In order to address the issue of using a numerical cut-off value and the inherent shortcomings of using such a number for a biological marker whose expression acts as a continuous variable, we further examined CD38 expression pattern. The aim was to see if we could improve the usefulness of CD38 as a prognostic marker, based on the hypothesis that any expression might predict a poor prognosis. According to the pattern of CD38 expression, 3 groups of patients were identified (Figure 2.11 C), (1) homogenously negative for CD38 (Unimodal CD38-ve group), (2) homogenously expressing CD38 at high levels (Unimodal CD38+ve group). (3) A third distinct subset of patients showed a bimodal expression profile of CD38, with the concomitant presence of one population expressing high levels of CD38 and second negative population (CD38 bimodal group).
We were able to demonstrate that the majority of the unimodal CD38-ve CLL patients had mutated IgVH genes (88%) and a long median TFI of 124 months. The unimodal CD38+ve group included CLL cases with both mutated (53%) and unmutated (47%) IgVH genes. Prognostically this group had a very short median TFI of 6 months, compared to 13 months when the 30% expression level was used. The CD38 bimodal expression group had a similar distribution of the mutational status as in the unimodal CD38 positive group. Despite having a median TFI of 102 months, almost 50% of the patients required treatment in this group, demonstrating that the presence of a CD38 positive population even in small numbers may result in poorer prognosis. It is important however to note that within the bimodal expression group the proportion of the two populations (CD38-ve and CD38+ve) may be highly variable in individual patients (Fig. 2.11 C) as the percentage of CD38+ve cells ranged between 8% to 79%.

Following the success of gene expression profiling (GEP) in diffuse large B-cell lymphoma several studies have used GEP to evaluate the expression of multiple genes in CLL. These studies found a CLL specific gene signature independent of the IgVH mutation status, implying that there is a common cell of origin and/or mechanism leading to malignant transformation (Klein, Tu et al. 2001; Durig, Nuckel et al. 2003; Jelinek, Tschumper et al. 2003). When cases were analysed by IgVH mutation status a restricted set of genes were able to discriminate between the unmutated and mutated CLL clones (Klein, Tu et al. 2001; Rosenwald, Alizadeh et al. 2001). The gene encoding zeta-associated protein 70 was found to have the best discriminatory value between the unmutated and mutated CLL clones when the gene array results were confirmed by immunoblot analysis (Chen, Widhopf et al. 2002).

ZAP-70 is an enzyme normally expressed by T-lymphocytes and is critical for antigenic activation of T-cells (Chu, Morita et al. 1998). B-cells generally lack ZAP-70, but instead use another related kinase, Syk, for signal transduction via the BCR complex (Turner, Schweighoffer et al. 2000), raising the possibility that the aberrant expression of ZAP-70 might alter BCR signalling in CLL cells. We measured ZAP-70 expression to assess its prognostic importance and its ability to act as a surrogate marker for IgVH gene mutation status.
We found a positive correlation in CLL cells with mutated \( \text{IgV}_{H} \) genes which did not express ZAP-70 measured by flow cytometry. In our study, three of 64 patients with mutated genes expressed more than 20% ZAP-70. Nineteen of 37 patients with unmutated genes were ZAP-70-ve giving a discordance rate of 23% for predicting \( \text{IgV}_{H} \) mutation status, with a positive predictive value (PPV) of 80% and a negative predictive value (NPV) of 76%. This is similar to the discordance rate identified by Rassenti in her study of 307 patients, where 47/164 patients with unmutated genes expressed less than 20% ZAP-70 and 24/243 patients with mutated genes expressed more than 20% ZAP-70 (Rassenti, Huynh et al. 2004). Two other groups Crespo et al (Crespo, Bosch et al. 2003) and Orchard et al (Orchard, Ibbotson et al. 2004) had previously shown a much higher concordance rate of 91% and 92% respectively. The use of a different antibody was postulated as a possible reason for discordance between the Rassenti and Orchard/Crespo series; however we used the same antibody clone as Orchard and Crespo. More recently, following the inclusion of more patients in the Orchard series, the concordance rate has dropped to 85% (personal communication). The correlation between ZAP-70 expression and \( \text{IgV}_{H} \) mutation status is therefore not as robust as previously suggested and it can not be recommended as a surrogate marker for the mutational status. The use of different cut-off levels for ZAP-70 expression (10%, 15%, and 25%) did not improve the predictive accuracy of ZAP-70 for the mutational status. ZAP-70 expression by other groups has been measured in fresh samples except in the study by Orchard et al. It is a technically difficult assay as ZAP-70 is an intra-cytoplasmic antigen requiring an additional step of cellular permeabilization, in addition to the non-availability of a conjugated antibody. To assess if storing the cells in DMSO, freezing and thawing had any effect on its expression we measured ZAP-70 expression in 40 samples, both fresh and cells frozen in DMSO. We did not find any difference in ZAP-70 expression in paired fresh and cryopreserved samples.

Three studies (Crespo, Bosch et al. 2003; Orchard, Ibbotson et al. 2004; Rassenti, Huynh et al. 2004) identified ZAP-70 as clinically significant in predicting patient outcome, with Rassenti and colleagues showing a hazard ratio (HR) of 4.9 associated with ZAP-70 expression as compared to a HR of 2.5 for patients with \( \text{IgV}_{H} \) unmutated genes in a multivariate regression model. Our results also found a significant difference between median TFI for the ZAP-70 positive and negative patients (5 vs. 137
102 months with a HR 3.2) in the whole cohort and in stage A CLL (16 vs. 120 months with a HR 3.4) in univariate analysis.

Finally, to evaluate the relative prognostic impact of IgV\textsubscript{H} mutational status, CD38 expression level and ZAP-70 expression a multivariate analysis was performed (Table 4.3 and 4.4).

### Table 4.3 Univariate analysis

<table>
<thead>
<tr>
<th>Prognostic Factor</th>
<th>HR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig-Unmutated</td>
<td>3.9</td>
<td>2.2 - 6.9</td>
<td>0.001</td>
</tr>
<tr>
<td>CD 38&gt;30%</td>
<td>2.8</td>
<td>1.5 - 5.1</td>
<td>0.001</td>
</tr>
<tr>
<td>ZAP-70 &gt;20%</td>
<td>3.2</td>
<td>1.5 - 6.8</td>
<td>0.002</td>
</tr>
</tbody>
</table>

### Table 4.4 Multivariate analysis

<table>
<thead>
<tr>
<th>Prognostic Factor</th>
<th>HR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig-Unmutated</td>
<td>2.6</td>
<td>1.3 - 5.4</td>
<td>0.01</td>
</tr>
<tr>
<td>CD 38&gt;30%</td>
<td>2.2</td>
<td>1 - 4.5</td>
<td>0.03</td>
</tr>
<tr>
<td>ZAP-70 &gt;20%</td>
<td>1.4</td>
<td>0.6 - 3.3</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Using a Cox proportional hazard regression model, only IgV\textsubscript{H} mutational status and CD38 retained clinical significance with regards to disease progression and TFI. The ZAP-70 expression on the other hand, lost its significance as a prognostic marker.

This finding may be explained on the basis of a recent study showing that ZAP-70 is associated with an activated cell phenotype inducible in both normal and CLL cells, but particularly in CD38 positive CLL cells (Tschumper 2004). Silvia Diaglio (Deaglio 2004) demonstrated a functional relationship between CD38 and ZAP-70, whereby CD38 ligation triggers phosphorylation of the ZAP-70 kinase and therefore its expression in CLL cells. The lack of added value in our series for ZAP-70 expression analysis is probably due to this linkage between the ZAP-70 and CD38 expression, where ZAP-70 expression in CLL cells is dependent on CD38 expression. The biology of ZAP-70 remains poorly understood and many questions remain

In conclusion, ZAP-70 expression did not have independent prognostic significance in patients with known IgV_H mutational status and CD38 expression and karyotypic abnormalities especially del 11q and 17p deletion. We will continue to stratify patients using standard parameters as well as FISH, CD38 expression level and pattern. IgV_H mutational status will be analysed on patients considered eligible for experimental therapy, which will include most patients <60 years of age, with active disease. ZAP-70 remains an interesting protein whose biological function in CLL deserves more attention, but at present we do not feel it significantly adds to the clinical management of CLL patients.

In summary, mutational status of the IgV_H genes separates CLL patients into benign and more aggressive groups. It has been prospectively evaluated is reproducible and has been standardized. However, it is expensive, time consuming and therefore not freely available. It should be used in clinical trials for patient stratification. On the other hand, abnormalities detected on FISH (17p & 11q) indicate one of the poorest outcome and are easily performed. Similarly CD38 expression has been shown to act as an independent prognostic marker, can be easily performed despite some disagreement as to what defines CD38 positivity. ZAP-70 expression although becoming easier to perform with the availability of conjugated antibodies, is not fully standardized and requires further work.

The ultimate goal for any clinician looking after patients with CLL is to be able to offer therapy that prolongs survival and improves the quality of life. After years of stagnation and slow progress the last few years have seen a remarkable progress, both in terms of predicting prognosis and the availability the different therapeutic options. We must however, remember that the role of these new prognostic markers and the novel and more intensive therapeutic regimens needs to be further evaluated in prospective randomized trials. Future research should also try to determine the mechanisms of drug resistance to therapy and ways to circumvent them.
Based on our own experience and those of others, we at St. James Hospital have adopted a treatment algorithm that exploits the use of the newer prognostic markers and also incorporates the recent therapeutic advances in the management of patients with CLL. The treatment strategy is based on the additive activity of Fludarabine and Cyclophosphamide. Rituximab has been added to improve the response rates and prolong the relapse free interval. Therapeutic efficacy will be monitored using MRD detection techniques that are currently being explored in CLL. Despite our enhanced ability in predicting prognosis in recent times, there are patients within the same group who can have a different clinical outcome. Gene expression analysis has been used recently to further refine prognostic indicators in CLL. Differences between the expression profiles of CLL cells with mutated and unmutated IgVH have been found. This approach may not however be sufficient on its own. CD38, an important prognostic marker shows no change at the level of the transcriptome (Rosenwald, Alizadeh et al. 2001). This may be because of differences between gene transcription and protein expression as one example. Protein expression analysis can provide vital information with regards to this final effector pathway, thereby offering a better understanding of disease biology, identifying novel disease biomarkers and therapeutic targets. Our aim is to assess differences in plasma protein expression in a group of well characterized patients with CLL (both conventional and new prognostic tools) undergoing novel chemomimmunotherapy, using proteomics. We hope to find disease biomarkers that may be useful for predicting the mutational status, for predicting therapeutic outcomes and importantly new markers that can improve our ability for predicting prognosis at diagnosis.
SECTION 2
CHAPTER 5

Fludarabine, Cyclophosphamide and Rituximab (FCR): An Effective Chemoimmunotherapy Combination with High Remission Rates for Chronic Lymphocytic Leukaemia
Chapter 5

5.1 Introduction:

Chronic lymphocytic leukaemia (CLL) is the commonest haematological malignancy in the Western world and is characterised by a heterogeneous clinical course. Recent advances in the understanding of the biology of CLL permit accurate identification of patients who will have an aggressive clinical course and may benefit from a more intensive treatment.

Fludarabine (F) has emerged as the most effective agent in CLL, either alone or in combination, with ORR of 80% and CR rates of 29% to 47% (Keating, O'Brien et al. 1998; Flinn, Byrd et al. 2000). Initial results from the UK LRF CLL4 trial have shown CR/NPR rates of up to 59% by combining Fludarabine and Cyclophosphamide (FC) (Catovsky, Richards et al. 2005). Despite the improvement in the CR rates most patients have residual disease post treatment, which leads to subsequent relapse. Early studies using single agent Rituximab (a monoclonal antibody against CD20 antigen) in CLL even in high doses were disappointing, possibly reflecting the known low CD20 expression in CLL (Almasri, Duque et al. 1992; Foran, Rohatiner et al. 2000; O'Brien, Kantarjian et al. 2001) and a shorter half life secondary to the antibody being used up by the circulating free CD20 antigen (Manshouri, Do et al. 2003). Recent data suggests that Rituximab acts additively with Fludarabine by modulating anti-apoptotic proteins, thus increasing Fludarabine sensitivity as well as enhancing cell clearance by both complement-mediated lysis and antibody-dependent cell mediated cytotoxicity (Alas and Bonavida 2001). Initial studies using Rituximab combinations in CLL used (500mg/m²) (Keating, Manshouri et al. 2002), the lowest effective dose in single agent studies. Rituximab can be associated with severe infusion related reactions including an ARDS-like syndrome (Byrd, Waseenko et al. 1999). The mechanism is unclear; however a high circulating tumour burden or stronger then usual expression of CD20 in some CLL patients may be risk factors.

There is also a correlation between the type of response to treatment and time to progression in CLL (Keating, O'Brien et al. 1993). In many haematological malignancies detectable minimal residual disease (MRD) at the end of therapy predicts is associated with poor survival and predicts relapse. In CLL recent studies
using combination chemotherapy regimens (FC or FC and mitoxantrone) (O'Brien, Kantarjian et al. 2001; Bosch, Ferrer et al. 2002) or high-dose therapy with stem-cell rescue (Esteve, Villamor et al. 2002) have used MRD analysis as part of remission status assessment. Patients achieving an MRD-negative status had longer progression free survival.

Our goal was to improve the response rates achieved in CLL patients by adding Rituximab to the combination of Fludarabine and Cyclophosphamide. The results of a series of 39 patients with CLL treated at St. James's Hospital (SJH) with modified FCR are presented.

5.2 Patients and Methods:

5.2.1 Patient selection:
Thirty-nine consecutive patients with CLL who fulfilled the following criteria, (1) Patients < 65 years with progressive stage A disease with high risk features or more advanced stage disease requiring therapy N=33, (2) biologically fit patients < 75 years who had failed Fludarabine therapy or had high risk symptomatic disease requiring treatment N=6, were treated in SJH between January 2002 and December 2005. Progressive stage A was defined as a short lymphocyte doubling time (< 6 months); downward trend in haemoglobin & platelets; increased size of nodes/spleen and B-symptoms. High-risk patients were defined on the basis of advanced stage, short lymphocyte doubling time, LDH > 450 IU/l, IgVH mutational status, CD38 expression (≥ 30% cells) and ZAP 70 expression (≥ 20% cells) (Table 1).

5.2.2 Treatment:

All patients received Fludarabine (25 mg/m²/d IV or 40 mg/m²/d PO) and Cyclophosphamide (250 mg/m²/d IV/PO) from day 1-3 of each cycle. Cycles were repeated every 28 days depending on the recovery of neutrophil count to > 1 x 10⁹/l and platelets to > 75 x 10⁹/l. Patients achieving CR, as per National Cancer Institute working group criteria (NCI-WG), after cycle 3 were given one additional cycle of
chemotherapy, otherwise a maximum of 6 courses were given. Rituximab (375 mg/m$^2$ IV infusion) was introduced from day 1 of the 2nd cycle following pre-medication with paracetamol 1 gm PO and Chlorphenaramine 10mg IV. Prophylactic Co-trimoxazole and Valaciclovir were prescribed until CD4+ T-lymphocytes counts were $>200 \times 10^6$/l post treatment. In the first 17 patients, further courses of FCR were delayed by 1 week and G-CSF 300 mcg (Neupogen®, Amgen) started if severe haematological toxicity (Grade III-IV according to WHO criteria) occurred in the preceding cycle of chemotherapy. G-CSF (Neupogen®, Amgen) was given from day +1 of subsequent cycles until neutrophil recovery. The protocol was amended because of the frequency of neutropenia so that all subsequent patients (N=22) received one dose of pegylated G-CSF 6mg (Neulasta®, Amgen) 24 hours after treatment.

Minimal residual disease (MRD) assessment was carried out for response assessment after cycle 3 and at completion of treatment, with a molecular-based strategy utilising multiplex PCR assays which target the immunoglobulin heavy and light chain loci to assess clonal status (N=12) (van Dongen, Langerak et al. 2003) and 4 colour flow-cytometry (N=18) (Rawstron, Kennedy et al. 2001). Sensitivity rates published for the two methods are $1 \times 10^3$ and $1 \times 10^4$ respectively. The MRD analysis was performed after cycles 3 and 6 on peripheral blood initially and on marrow if blood analysis was negative.

5.3 Results:
Between January 02 and December 05, 39 patients, 29 males and 10 females were treated with modified FCR chemotherapy. Twenty-six patients were treatment naïve and 13 patients had previously been treated (1 to 4 lines of therapy). Median age at diagnosis was 57 (38-75). Twelve patients had progressive stage A disease, 16 stage B and 11 patients had stage C disease according to the Binet classification (Binet, Auquier et al. 1981). Patient characteristics are shown in Table 5.1.
### Table 5.1: Patient Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of Patients</td>
<td>39</td>
</tr>
<tr>
<td>Age [median (range)]</td>
<td>57 (38-75)</td>
</tr>
<tr>
<td>Sex (M/F)</td>
<td>29 / 10</td>
</tr>
<tr>
<td>WCC [median (range)]</td>
<td>$81.3 \times 10^9/l$ (7.4-363)</td>
</tr>
<tr>
<td>Hb [median (range)]</td>
<td>$12.4 \times 10^9/l$ (7.2-17.2)</td>
</tr>
<tr>
<td>Platelet [median (range)]</td>
<td>$132 \times 10^9/l$ (11-303)</td>
</tr>
<tr>
<td>β2m [median (range)]</td>
<td>2.63 mg/l (1.53-6.5)</td>
</tr>
<tr>
<td>LDH [median (range)]</td>
<td>398 IU/l (223-1232)</td>
</tr>
<tr>
<td>Untreated</td>
<td>26</td>
</tr>
<tr>
<td>Pre-treated</td>
<td>13</td>
</tr>
<tr>
<td>Binet Stage</td>
<td></td>
</tr>
<tr>
<td>A (Progressive)</td>
<td>12</td>
</tr>
<tr>
<td>B</td>
<td>16</td>
</tr>
<tr>
<td>C</td>
<td>11</td>
</tr>
<tr>
<td>Time since first diagnosis</td>
<td></td>
</tr>
<tr>
<td>≤ 1 year</td>
<td>19</td>
</tr>
<tr>
<td>2-5 years</td>
<td>12</td>
</tr>
<tr>
<td>≥ 5 years</td>
<td>8</td>
</tr>
<tr>
<td>B symptoms</td>
<td>19 / 39</td>
</tr>
<tr>
<td>CD38 Positive*</td>
<td>15 / 35</td>
</tr>
<tr>
<td>CD38 Negative</td>
<td>20 / 35</td>
</tr>
<tr>
<td>IgVH Unmutated†</td>
<td>14 / 25</td>
</tr>
<tr>
<td>IgVH Mutated</td>
<td>11 / 25</td>
</tr>
<tr>
<td>ZAP-70 Positive‡</td>
<td>6 / 22</td>
</tr>
<tr>
<td>ZAP-70 Negative</td>
<td>16 / 22</td>
</tr>
</tbody>
</table>

* 35/39 results available.  
† 25/39 results available.  
‡ 22/39 results available.

Table 5.1 Patient Characteristics

All patients were evaluable for toxicity while 36/39 patients were available for response assessment (Table 5.2 and 5.3).
<table>
<thead>
<tr>
<th>Treatment Outcome</th>
<th>ORR (%)</th>
<th>27 (75%)</th>
<th>8 (22%)</th>
<th>1 (3%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Patients</td>
<td>36 (100%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untreated (n=25)</td>
<td>25 (100%)</td>
<td>22 (88%)</td>
<td>2 (8%)</td>
<td>1 (4%)</td>
</tr>
<tr>
<td>Treated (n=11)</td>
<td>11 (100%)</td>
<td>5 (45%)</td>
<td>6 (55%)</td>
<td></td>
</tr>
<tr>
<td>Binet A (n=11)</td>
<td>11 (100%)</td>
<td>9 (82%)</td>
<td>1 (9%)</td>
<td>1 (9%)</td>
</tr>
<tr>
<td>Binet B (n=14)</td>
<td>14 (100%)</td>
<td>10 (71%)</td>
<td>4 (29%)</td>
<td></td>
</tr>
<tr>
<td>Binet C (n=11)</td>
<td>1 (100%)</td>
<td>8 (73%)</td>
<td>3 (27%)</td>
<td></td>
</tr>
<tr>
<td>IgVH Unmutated (n=13)</td>
<td>13 (100%)</td>
<td>12 (92%)</td>
<td>1 (8%)</td>
<td></td>
</tr>
<tr>
<td>IgVH Mutated (n=9)</td>
<td>9 (100%)</td>
<td>8 (89%)</td>
<td>1 (11%)</td>
<td></td>
</tr>
<tr>
<td>CD38+ve (n=14)</td>
<td>14 (100%)</td>
<td>11 (79%)</td>
<td>3 (21%)</td>
<td></td>
</tr>
<tr>
<td>CD38-ve (n=18)</td>
<td>18 (100%)</td>
<td>15 (83%)</td>
<td>2 (11%)</td>
<td>1 (6%)</td>
</tr>
<tr>
<td>ZAP-70+ve (n=5)</td>
<td>5 (100%)</td>
<td>4 (80%)</td>
<td>1 (20%)</td>
<td></td>
</tr>
<tr>
<td>ZAP-70-ve (n=14)</td>
<td>14 (100%)</td>
<td>14 (100%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flowcytometry (n=18)</td>
<td>18 (100%)</td>
<td>8 (44%)</td>
<td>10 (56%)</td>
<td></td>
</tr>
<tr>
<td>Clonality PCR (n=12) †</td>
<td>12 (100%)</td>
<td>4 (33%)</td>
<td>7 (58%)</td>
<td></td>
</tr>
</tbody>
</table>

† One patient failed to amplify.
ORR = Overall response rate.
CR = Complete response.
PR = Partial response.
NPR = Nodular partial response.

Table 5.2 Treatment Outcome
### Table 5.3: Toxicity during treatment

<table>
<thead>
<tr>
<th>Toxicity</th>
<th>Grade I</th>
<th>Grade II</th>
<th>Grade III</th>
<th>Grade IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutropenia</td>
<td>6%</td>
<td>9%</td>
<td>16%</td>
<td>7%</td>
</tr>
<tr>
<td>71/187 (38%)</td>
<td>(11/187)</td>
<td>(17/187)</td>
<td>(29/187)</td>
<td>(14/187)</td>
</tr>
<tr>
<td>Anaemia</td>
<td>0.53%</td>
<td>3.2%</td>
<td>2.1%</td>
<td>-</td>
</tr>
<tr>
<td>11/187 (6%)</td>
<td>(1/187)</td>
<td>(6/187)</td>
<td>(4/187)</td>
<td>-</td>
</tr>
<tr>
<td>Thrombocytopenia</td>
<td>6%</td>
<td>3.2%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20/187 (11%)</td>
<td>(11/187)</td>
<td>(6/187)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Lymphopenia</td>
<td>9%</td>
<td>29%</td>
<td>47%</td>
<td>-</td>
</tr>
<tr>
<td>158/187 (84%)</td>
<td>(17/187)</td>
<td>(54/187)</td>
<td>(87/187)</td>
<td>-</td>
</tr>
<tr>
<td>Haemolytic Anaemia</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### Non-Haematologic Toxicity

<table>
<thead>
<tr>
<th>Toxicity</th>
<th>No of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nausea</td>
<td>20</td>
</tr>
<tr>
<td>Diarrhoea</td>
<td>4</td>
</tr>
<tr>
<td>Fever</td>
<td>8</td>
</tr>
<tr>
<td>Infection Requiring Admission</td>
<td>3</td>
</tr>
<tr>
<td>Alopecia</td>
<td>-</td>
</tr>
<tr>
<td>Mucositis</td>
<td>1</td>
</tr>
<tr>
<td>Chills</td>
<td>15</td>
</tr>
<tr>
<td>Dyspnoea</td>
<td>5</td>
</tr>
<tr>
<td>Rash / Arthralgia / Seizures</td>
<td>4</td>
</tr>
</tbody>
</table>

### Table 5.3 Treatment associated toxicity

Twenty patients completed 6 cycles and 9 completed 4 cycles. Ten patients did not finish the treatment with 7 receiving 3 cycles and 3 patients only 2 cycles each. Four patients (3 cycles) had a good response and did not want to continue the treatment, 2 (2 cycles) patients were taken off protocol (1= Squamous Cell Carcinoma, 1= Seizures), 2 patients (1=2 cycles, 1=3 cycles) stopped the treatment because of prolonged neutropenia and 2 patients (3 cycles) died because of respiratory infections. Three patients who received 2 cycles each were not included in response assessment but were included in toxicity assessment. The outcome 6 weeks post-treatment for 36/39 patients was: CR – 27 (75%), partial remission (PR) – 8 (22%) and Nodular PR - 1 (3%) with an ORR of 100% (NCI-WG criteria). Twenty-six of 31 alive patients
have maintained their post-treatment disease status (23 CR, 2 PR, 1 NPR) with a
median follow up of 17 months (range 2-41). Two patients in PR post-FCR, one
heavily pre-treated, died of transplant-related toxicity. A third patient died 2 months
after FCR with nodal relapse consistent with Richter transformation, after having
achieved a PR while two patients (1=CR, 1=PR) died of infective complications
having received 3 cycles each. Two patients who progressed after an initial PR have
responded to a reduced intensity conditioning transplant (RIC) and Campath-1H.

Neutropenia (<1.0 - 0.5x10^9/l, ANC Grade III-IV) occurred in 43/187 (23%) cycles.
Six patients required admission to hospital for treatment of febrile neutropenia and 2
other patients developed grade III and IV anaemia and thrombocytopenia requiring
transfusion (Table 3). None of the patients developed autoimmune haemolysis. G-
CSF (Neupogen® or Neulasta®, Amgen) was administered in 99/187 cycles. Forty-six
of 88 (52%) as compared to 25/99 (25%) cycles were associated with neutropenia
prior to the use of G-CSF. MRD analysis (End of therapy) was available in 18 patients
using 4-colour flow cytometry. Ten patients (56%) were MRD positive and 8 (44%)
patients MRD negative. Nine of the 10 positive patients were in CR using NCI-WG
criteria. Molecular results were available for 12 patients, 7 were positive (58%), 4
(33%) negative and 1 (9%) patient failed to amplify. Ten patients had results available
for both techniques. Seven of 10 patients had concordant results, but 3 patients had
detectable disease by flow cytometry and not molecular analysis (Table 2).

5.4 Discussion:

Fludarabine is the mainstay of treatment for previously treated and treatment naïve
patients, with an ORR respectively of 50% and 80% respectively and CR of 29% in
treatment naïve patients (Keating, O'Brien et al. 1998). Combination of
Cyclophosphamide and Fludarabine results in an ORR of 80-90% and CR of up to
47% (Flinn, Byrd et al. 2000). These results are encouraging, but most patients relapse
and require further treatment. PCR based MRD assays demonstrate molecularly
detectable disease despite clinical CR by NCI criteria in almost all patients after
Fludarabine therapy (Clavio, Miglino et al. 1998). Single agent Rituximab
administered at standard doses results in a disappointing ORR of 15% in refractory patients (Nguyen, Amess et al. 1999). Experimental data shows that combining Rituximab with Fludarabine potentiates chemotherapy-induced apoptosis (Di Gaetano, Xiao et al. 2001). Early results of Rituximab, Fludarabine and Cyclophosphamide (FCR) chemoimmunotherapy in CLL patients have demonstrated a CR of 70% and MRD negativity by immunophenotyping using CD5 and CD19 co-expression in two thirds of the patients (Keating, O'Brien et al. 2005). It is unclear if superior disease control will result in longer time to treatment failure (TTF) or improved survival.

We adapted the original FCR protocol to suit the needs of our patient population and institution without compromising efficacy. Rituximab was omitted in cycle 1, because infusion related reactions are anecdotally associated with a high tumour burden and strong CD20 expression. The toxicity appears to be mediated by cytokines such as TNF-α, IFN-γ, IL-6 and IL-8 which are higher at the start of treatment than in subsequent cycles (Winkler, Jensen et al. 1999). A decision was therefore made to obtain peripheral lymphocyte control with one cycle of FC. The rarity and mild infusion related toxicity we encountered (23/148 cycles) might reflect this decision. Rituximab was used at a standard 375mg/m² dose, because the rationale for using higher doses based on single agent studies associated with limited efficacy was unclear. It may be appropriate to define the optimal dose of Rituximab in combination therapy where biological synergy is important in the future. Cyclophosphamide was given IV in the first cycle of FC and if nausea and emesis were not a clinical problem was subsequently given orally. The 36 evaluable patients had an ORR of 100% with a CR, PR and NPR of 75%, 22% and 3% respectively. While the ORR was similar in previously treated and untreated patients, 22/25 (88%) patients in the untreated group achieved CR compared to 5/11 (45%) in the previously treated group. The CR rate for progressive stage A patients was 82% which was better than stage B and C patients with a CR rate of 71% and 73% respectively. These results indicate that the modified FCR protocol was effective and clinically safe to administer to patients in a day unit setting.

FCR causes profound neutropenia and lymphopenia (Keating, O'Brien et al. 2005). Forty three of 187 cycles of FCR were associated with grade III-IV neutropenia, but
only 6 cycles resulted in hospitalisation. G-CSF (Neupogen®, Amgen) was used in 11 of the first 17 patients because of prolonged neutropenia (46/88 cycles-52%). The protocol was amended to include pegylated G-CSF (Neulasta®, Amgen) preemptively on the fourth day of every cycle for the next 22 patients, which resulted in a significant reduction in the number of cycles associated with severe neutropenia to 25/99 (25%). Myelosuppression was the most common cause for early discontinuation of therapy. It was more common in older patients (>70 years), advanced stage disease and prior history of therapy. Severe lymphopenia was universal; however no opportunistic infections were encountered. Moreover, patients were restaged after 3 courses of treatment and if they were in CR (NCI-WG criteria), treatment was curtailed after 4 cycles to avoid further toxicity.

Response criteria in CLL (NCI-WG) have defined complete remissions on clinical and morphological grounds, which is adequate for measuring disease response with conventional therapy. The ability to assess minimal disease loads becomes important when experimental therapies such as monoclonal antibodies and RIC are used, which result in a higher CR and more profound responses as compared to using NCI-WG criteria. Two strategies are effective in CLL MRD detection at present: 4-colour flow cytometry (Sensitivity = single CLL cell in $10^5$ normal cells) (Rawstron, Kennedy et al. 2001) or the allele-specific oligonucleotide [ASO]–PCR, with a sensitivity of 1 in $10^6$ (Voena, Ladetto et al. 1997). ASO-PCR is labor intensive, and not suited to a routine diagnostic laboratory, whereas the BIOMED clonality PCR methodology with a sensitivity of 1 in $10^2$ is available in clinical molecular laboratories and is robust and reproducible (van Dongen, Langerak et al. 2003). The MRD results using immunophenotyping showed a CR rate of 8/18 (44%) patients i.e. considerably lower than with NCI-WG criteria (Table 2). MRD testing may be needed routinely in CLL in future to assess treatment efficacy in clinical trials. The discrepancy between flow cytometry and clonality (7/10 patients had concordant results) confirmed the higher sensitivity of MRD flow cytometry.

Modified FCR regimen is safe to use in a day patient setting and clinically effective. Longer-term evaluation will be needed to assess disease free interval or overall survival and the predictive value of MRD.
CHAPTER 6

Protein Expression Analysis of CLL Patients Treated with a Novel Chemoimmunotherapy (FCR) Protocol
6.1 Introduction to Proteomics

6.1.1 Background

Each of our cells contains the genetic information necessary to make a complete genome. However, not all genes are expressed in all cells. Genes that code for enzymes essential for basic cellular functions, for example DNA synthesis, are expressed in virtually all cells, whereas those with highly specialised functions are expressed only in specific cell types, for example haemoglobin in red blood cells. Thus, all cells express: 1) genes whose protein products provide essential functions and 2) genes whose protein products provide unique cell-specific functions i.e., one genome but many proteomes.

Key developments in the areas of DNA microarray technology and user-friendly bioinformatics tools have allowed the analysis of gene expression at the mRNA level. The “transcriptome” is the mRNA pool found within a cell (Voehringer, Hirschberg et al. 2000; Blohm and Guiseppi-Elie 2001). However, there is frequently a poor correlation between mRNA and protein expression. Changes in expression at the mRNA level do not necessarily correlate with changes of the functional expression of a protein (Anderson and Seilhamer 1997; Anderson and Anderson 1998; Braziel, Shipp et al. 2003; Cristea, Gaskell et al. 2004). Differing stability of mRNAs leading to variation in translational efficiency can affect the generation of new proteins. Once formed, proteins differ significantly in stability and turnover rates. Many proteins involved in signal transduction, transcription-factor regulation, and cell-cycle control are rapidly turned over as a means of regulating their activities.

Proteomics is the study of the proteome, the protein complement of the genome. The terms “proteomics” and “proteome” were coined by (Wilkins, Sanchez et al. 1996) and colleagues in the early 1990s, referring to the pool of proteins expressed in a cell or tissue at a given time and mirror the terms “genomics” and “genome”, which describe the entire collection of genes in an organism, as depicted in Figure 6.1.
The analysis of the proteome provides us with the information on the biological processes happening in the tissue/sample, allowing the comparison of physiological and pathological states (Blackburn 2001; Cristea, Gaskell et al. 2004; de Hoog and Mann 2004).

6.1.2 Challenges in Proteomics

The introduction of cDNA or oligonucleotide microarrays paved the way to simultaneously measure the expression of many or all of the genes in an organism. These methods rely on PCR and hybridisation of oligonucleotides to complementary sequences. There are however, no analogous tools available for protein analysis.
First, there is no analytical method for amplifying target proteins, like the PCR application of DNA. It is not currently possible to induce polypeptide molecules to replicate themselves in a manner analogous to oligonucleotide replication through PCR. Whereas a small amount of target DNA can be amplified through PCR, a small amount of polypeptide must be detected and analysed without amplification. Secondly, proteins do not specifically hybridize to complimentary amino acid sequences. Another problem peculiar to Proteomics, is that each gene does not necessarily give rise to only one protein product in the cell. This is because proteins are post-translationally modified (PTM) through processes such as proteolysis, glycolysis and phosphorylation, making the protein structurally more complex and difficult to identify (Binz, Muller et al. 2004; Honore, Ostergaard et al. 2004; Rees-Unwin, Morgan et al. 2004). The extent and variety of modification varies with individual proteins, regulatory mechanisms within the cell, and environmental factors.

Analysis of the proteome thus requires a different set of tools when compared to gene-expression analysis. The task of characterising the proteome requires analytical methods to detect and quantify proteins in their modified and unmodified forms. The necessity of detecting and differentiating between the multiple protein products of any particular gene adds to the analytical challenge of Proteomics.

6.1.3 Proteomic techniques

Analytical protein identification is based on the fact that most peptide sequences of approximately six or more amino acids are usually unique in the proteome of an organism. Thus, identifying a peptide sequence or measuring its mass accurately allows protein identification by finding its match in a database of protein sequences, as shown in Figure 6.2.
The most important steps or tools required for analytical proteomics are protein separation, mass spectrometry (MS), software to analyze data generated by MS and a protein sequence database providing a catalogue of all expressed proteins.

6.1.3.1 Protein Separation / Fractionation

Most techniques currently used in proteomics, employ a variety of fractionation and separation steps prior to analysis by MS (Dreger 2003; Huber, Pfaller et al. 2003). Protein separation is performed for two specific reasons in proteomics. Firstly, it simplifies complex protein mixtures by resolving them into individual proteins or small groups of proteins, secondly it permits apparent differences in protein levels to be compared in different samples, thus allowing targeting specific proteins for analysis (Liebler 2002). Many techniques have been used for protein separation including 1D-SDS Polyacrylamide Gel Electrophoresis (PAGE), 2D-SDS-PAGE, high-performance liquid chromatography (HPLC) and isoelectric focusing (IEF). Two-dimensional SDS PAGE (2D-PAGE) is the most widely associated with proteomics, and comprises two different methods of separations. In the first dimension, proteins are resolved on the basis of the isoelectric point (IP; pH value at which the net charge on a molecule is zero) by using IEF. In the second dimension, focused proteins are further resolved by electrophoresis on a polyacrylamide gel on
the basis of molecular weight as shown in Figure 6.3, (Liebler 2002; Kolch, Mischak et al. 2005).

![Figure 6.3 Schematic representation of 2D SDS-PAGE](image)

Proteins separated by 2D gels are visualised by conventional staining techniques, including silver, Coomassie blue and amido black stains. It is important to remember
that not all staining techniques allow subsequent analysis of the proteins (Rabilloud 1990; Yan, Wait et al. 2000).

2D SDS-PAGE has proven to be a reliable and efficient method for separation of several hundred to a few thousand proteins on the basis of differences in their PI and molecular weight and has been used both as an analytical and preparative tool. Gel electrophoresis does have some major disadvantages. The procedure can be time-consuming and lack reproducibility. Multiple gels have to be electrophoresed in order to identify specific proteins with altered regulation and a large amount of good quality protein is required for protein separation and identification. The insolubility of hydrophobic proteins and the difficulty in detecting and separating low-abundance proteins and membrane proteins are additional disadvantages (Honore, Ostergaard et al. 2004; Rees-Unwin, Morgan et al. 2004; Kolch, Mischak et al. 2005).

The recently introduced variation of 2D-PAGE, termed “two-dimensional difference in gel electrophoresis” (DIGE) (Unlu, Morgan et al. 1997), has significantly improved the speed, reproducibility and sensitivity of 2-D PAGE based Proteomics (Gharbi, Gaffney et al. 2002; Somiari, Sullivan et al. 2003). The concept involves the covalent labelling of protein extract with different fluorescent dyes, e.g. cyanine (Cy2, Cy3, or Cy5) (Gharbi, Gaffney et al. 2002; Petricoin, Zoon et al. 2002; Somiari, Sullivan et al. 2003) dyes or Alexa dyes (Von Eggeling, Gawriljuk et al. 2001). The “test” protein sample is labelled with Cy3 and the reference sample is labelled with Cy5 (Somiari, Sullivan et al. 2003), with samples mixed and co-separated during the same 2D-PAGE process. The gel pattern is visualised by using a fluorescent imager. This allows the quantification of each spot with image analysis software such as DeCyder shown in Figure 6.4, (Gharbi, Gaffney et al. 2002; Somiari, Sullivan et al. 2003).
Figure 6.4 The Ettan DIGE system. Test samples labelled with Cy3 and Cy5 are run on the same 2D gel with an internal standard labelled with Cy2 and prepared from pooling all test samples. The gel is scanned using a fluorescent scanner. The difference in expression/level of the test protein can be analyzed in different formats. (Taken from the Ettan DIGE user manual, Amersham Bioscience, 18-1173-17 Edition AA)

2D-DIGE effectively reduces the gel-to-gel variability associated with standard 2-D PAGE and improves the accuracy of quantitative protein profiling.

6.1.3.2 Mass Spectrometry

Mass spectrometry (MS) is an analytical technique used to measure the mass-to-charge ratio ($m/z$) of ions. This is achieved by ionizing the sample, separating ions of differing masses and recording their relative abundance by measuring intensities of ion flux. A typical mass spectrometer comprises three parts: an ion source, a mass analyzer, which resolves ions based on their ($m/z$) ratio, and a detector system that detects the ions resolved by the mass analyzer, as shown in Figure 6.5, (Rosenblatt, Bryant-Greenwood et al. 2004; Kolch, Mischak et al. 2005).
Three different types of ion sources are used for most proteomics MS work: Electro spray ionization (ESI) instruments, matrix-assisted laser desorption/ionization (MALDI) instruments and surface enhanced laser desorption/ionization (SELDI) instruments.

In MALDI, the peptide or protein sample is dried and crystallised in a large excess of a matrix which has a strong absorption at the laser wavelength. Following the laser irradiation of the sample surface, the matrix accumulates a large amount of energy that initiates the proton transfer between the matrix and the analyte to form ions. The ions observed during MALDI MS are mainly singly charged. This results in simple spectra (because the m/z ratio is being measured) even in the case of analysis of mixtures, but can be a disadvantage for peptide sequencing, which requires the
achievement of peptide fragmentation, a process that preferentially occurs with multicharged peptides (Cristea, Gaskell et al. 2004), Figure 5.5, A.

In ES, a continuous spray of ions are produced directly from solution and facilitates interfacing of HPLC/LC with MS. A sample is passed with flow rates of 1-10 µL/min through a capillary held at high potential relative to ground and counter electrode. The strong electric field obtained induces charge accumulation at the liquid surface situated at the end of the capillary and leads to the formation of a mist of highly charged droplets. The ES process results in the formation of multiply charged ions. This is one of the principal advantages of this method because it allows the analysis of ions from high-molecular-mass molecules, such as proteins and peptides, using mass spectrometers of limited $m/z$ ratio range (Cristea, Gaskell et al. 2004), Figure 5.5, B.

There are many types of mass analyzers, but they all operate on the same principle, the gas-phase ions produced in the source are separated according to the $m/z$ ratio on the basis of their motion in a vacuum under the influence of an electrical or magnetic field. Time-of-flight (TOF), quadrupole (Q), ion trap (IT) and fourier transform ion cyclotron resonance (FT-ICR) are the most frequently used analyzers. Different configurations of an ion source and mass analyzers have been used to develop the many types of MS equipment, Figure 5.5, A-E.

The TOF analyzer measures the time it takes for the ions to fly from one end of the analyzer to the other and strike the detector. An ion of known electrical charge and unknown mass enters a mass spectrometer and is accelerated by an electrical field of known strength. This acceleration results in any given ion having the same kinetic energy as any other ion given that they all have the same charge. The velocity of the ion will depend however on the mass-to-charge ratio of the particle (heavier particles reach lower speeds). The resolution of TOF instruments has been greatly improved by using a series of electrodes known as a reflectron. It focuses ions of the same $m/z$ values and allows them to reach the detector at the same time (Flensburg, Haid et al. 2004).

The quadrupole mass analyzer consists of four parallel metal rods. Each opposing rod pair is connected together electrically and a radio frequency (RF) voltage is applied
between one pair of rods, and the other. A direct current voltage is then superimposed on the RF voltage. Ions travel down the quadrupole between the rods. Depending on the voltage applied to the rods, ions of specific \( m/z \) value will pass through the quadrupole, whereas ions of greater or lesser \( m/z \) values will have unstable trajectories and will collide with the rods. This allows selection of a particular ion, or scanning by varying the voltages (Rosenblatt, Bryant-Greenwood et al. 2004). These types of mass spectrometers excel at applications where particular ions of interest are being studied because they can stay tuned on a single ion for extended periods of time.

The ion trap analyzer captures the ions, which collide with the helium "bath" gas and start to oscillate in a predicted motion. Ion trap can be used as a "tandem-in-time" instrument as selection, fragmentation, and analysis of ions take place in the same space. FT-ICR is also a trapping device, in this case by using strong magnetic fields, and offers great opportunities for investigating protein interactions and PTMs, with high sensitivity, mass accuracy, and resolution (Stensballe, Jensen et al. 2000).

6.1.4 SELDI-TOF MS

6.1.4.1 Background

SELDI is a recent technology that allows for high-throughput proteomic studies. Originally described in 1993 by Hutchens and Yip (Hutchens 1993) it utilizes retentate chromatography of a protein sample directly on ProteinChip Arrays (Ciphergen Biosystems, Fremont, CA, USA) (Fung, Thulasiraman et al. 2001; Issaq, Veenstra et al. 2002). The arrays display various kinds of chemically activated surfaces which bind molecules based on established principles such as ion exchange chromatography, metal ion affinity and hydrophobic affinity. It is a rapid technique and relies on TOF mass spectrometry for the accurate measurement of the \( m/z \) ratio of the peptides and proteins. It is considered to be a refinement of MALDI-TOF MS, as in both methods, proteins are co-crystallized with energy absorbing compounds and vaporized by a pulsed-UV laser beam (Vorderwülbecke 2005). The differences between SELDI and MALDI are in the construction of sample targets, the design of the analyzer and the software tools used to interpret the acquired data.
SELDI involves the binding of proteins and peptides with specific chemical or biological properties present in complex biological samples, such as serum, cell lysates, tissue homogenates or culture supernatants, to ProteinChip Arrays or 'chips'. The proteins actively interact with the chromatographic array surface, and become sequestered according to their surface interaction potential as well as separated from salts and other sample contaminants by subsequent on-spot washing with appropriate buffer solutions (Vorderwülbecke 2005). The bound proteins are then used as a target for laser desorption and ionization (Issaq, Veenstra et al. 2002). A SELDI-TOF reader and SELDI software analyses the data allowing sample processing. The components of the system are shown in Figure 6.6.

**Figure 6.6** The experimental steps for analysing samples on SELDI protein chips. *Nature Methods* 2, 393 - 395 (2005)

The first step in SELDI experiments involves selection of Proteinchip Arrays. These are available with different chromatographic properties, including hydrophobic, hydrophilic, anion exchange, cation exchange, and immobilized-metal affinity surfaces. Sample can then be applied directly to the ProteinChip Arrays. The proteins of interest are captured on the chromatographic surface by adsorption, partition, electrostatic interaction or affinity chromatography depending on their properties. After a short incubation period, unbound proteins are washed off the surface of the ProteinChip Array. Only proteins interacting with the chemistry of the array surface are retained for analysis. Finally, an energy absorbing matrix is applied to the array.
When this array is subjected to a laser beam, the matrix absorbs the energy and the sample is ionised, mobilizing the protein molecules for analysis (Bensmail, Golek et al. 2005) as illustrated in Figure 6.7.

![Figure 6.7 ProteinChip® technology: SELDI TOF-MS Detection](image)

The advantage of this chip-based technology is that rapid, high throughput analysis can be performed on crude samples (Cristea, Gaskell et al. 2004). The availability of chips with various binding properties allows the generation of different spectra, enabling sample comparison. SELDI-TOF is straightforward to use even by individuals with minimal MS experience.

The analyzer used for SELDI, ProteinChip Reader is different to a MALDI analyzer and is especially adapted to achieve high-sensitivity quantification and good reproducibility (Bensmail, Golek et al. 2005). In contrast, MALDI devices are not designed for reliable quantitative precision over a wide mass range. Another important advantage of the ProteinChip Arrays is that it can be used when small amounts of sample are available for analyses, especially compared to other proteomic techniques.

**6.1.4.2 Sample Fractionation for SELDI**

Sample fractionation plays an important role in SELDI analysis. The complexity of the human serum proteome is attributed to a large dynamic range of protein...
abundance and, with a disproportionate few dozen proteins representing about 99% of the total protein content (Righetti, Castagna et al. 2005; Thadikkaran, Siegenthaler et al. 2005). A pre-fractionation step to remove these high abundance proteins prior to any high-resolution analysis by MS, is desirable. This also allows enrichment and detection of the low-abundance proteins (Stasyk and Huber 2004; Righetti, Castagna et al. 2005). A number of methods are available for fractionation e.g., Gradient Centrifugation, Electrophoresis techniques and Chromatography. The choice of fractionation technique depends on the sample and experiment.

6.1.4.3 The Protein Chips / Affinity arrays

This technology utilizes specialized stainless steel or aluminium based supports coated with chemical or biological surfaces to selectively capture proteins based on the intrinsic properties of the protein. There are many different types of arrays available such as affinity arrays, antibody arrays, antigen arrays and tissue arrays, as shown in Figure 6.8 and 6.9.

Figure 6.8 Affinity arrays from Ciphergen Biosystem.
6.1.4.3.1 Affinity arrays

Four types of affinity arrays are used. A normal phase (NP20) array is used for calibration purposes Figure 6.9.

**Figure 6.9** Types of arrays

*Immobilized Metal Affinity Capture (IMAC)*

Immobilized metal affinity chromatography (IMAC) is a powerful purification technique that relies on a molecule’s affinity for certain metals immobilized onto a chelating surface. The chelating ligand, iminodiacetic acid (IDA) in this case, may be charged with transition metals such as Cu2+, Ni2+, Co2+, or Zn2+. This results in the highly selective capture of proteins with clustered histidine residues onto a porous chromatographic support. IMAC purification of His-tagged proteins is frequently used for structural and functional studies of proteins, as shown in Figure 6.10.
Weak Cation Exchanger (CM10)

CM10 ProteinChip Arrays have a hydrophobic surface and are the array of choice for weakly cationic proteins. The active spots contain weak anionic carboxylate groups that interact with the positive charges on the surface of target proteins, which are lysine, arginine or histidine rich as shown in Figure 6.11. The chip binds proteins that are positively charged at a given pH. The pH of the binding buffer can be varied to selectively increase or decrease protein binding by altering the net charge on proteins.
Strong Anion Exchanger (Q10)

The surface of the Q10 ProteinChip is coated with positively charged cationic, quaternary ammonium groups which interact with negatively charged target amino acids, aspartic acid or glutamic acid as shown in Figure 6.12. The surface binds peptides/proteins which are negatively charged at a given pH. By varying the pH of the binding/washing buffer the overall net charge can be changed on the proteins, resulting in more or less binding.

Figure 6.12 Q10 ProteinChip Array surface chemistry with protein (http://www.ciphergen.com).

Hydrophobic binding surface (H50)

H50 chip has a hydrophobic surface and is used for capturing proteins and peptides through reversed phase or hydrophobic interactions. Proteins which are less hydrophobic than the binding buffer will not bind to the array surface while proteins which are more hydrophobic will bind to the array surface as shown in Figure 6.13. By increasing or decreasing the organic content of the washing buffer, the hydrophobic nature of the buffer can be increased or decreased affecting protein binding.
Normal Phase (NP20)

This is the least selective of all the array surfaces. The active spots of the arrays contain silicon dioxide which allows proteins to bind via serine, threonine or lysine and are used for general binding of proteins while running calibration samples as shown in Figure 6.14.

Figure 6.13 H50 ProteinChip Array surface chemistry with proteins (http://www.ciphergen.com).

Figure 6.14 NP20 ProteinChip Array surface chemistry with proteins (http://www.ciphergen.com).
6.1.4.4 The ProteinChip SELDI reader

The SELDI-TOF analyzer comprises a laser source, a TOF mass analyzer and a detector, shown in Figure 6.15.

![Ciphergen PBS II Series ProteinChip reader](http://www.ciphergen.com).

The ProteinChip Reader uses a nitrogen laser to desorb and ionize the sample. Ionization of the sample is caused by the laser energy exciting the energy EAM and analyte. The laser energy causes both protein ionization and a change from crystalline into gas phase, allowing the ionized molecules to fly as differential voltage is applied, as shown in Figure: 6.7. Proteins with a positive charge fly away from the metal array which also has a positive charge. The voltage differential applies the same kinetic energy to all of the analytes in the sample, thus resulting in flight times which depend upon the mass. The ProteinChip Reader records the TOF of the analyte and calculates a highly accurate and precise mass from this measurement. The equal distribution of analyte and EAM mixture across the spot area results in the generation of signal intensities which correspond to the concentration of peptides and proteins on the spot. This allows accurate quantification of single protein components in the sample.

The data generated by the protein chip reader is analyzed by software, (Ciphergen Express and ProteinChip software), both from Ciphergen Biosystem. The primary purpose of ProteinChip Software is data acquisition from the prepared sample arrays whilst Ciphergen Express and Ciphergen Pattern Software are used for data analysis.
The ProteinChip software controls all aspects of the ProteinChip Reader and facilitates data collection and analysis. The protein spectrum generated from the reader is in the form of peaks corresponding to specific protein masses with the height of the peak corresponding to the approximate amount of protein expression in the sample. CiphergenExpress Software provides a database system for managing and tracking data generated by the ProteinChip software. It organizes the acquired spectral data refining results and generating reports. Ciphergen Express Software allows data mining and Expression Difference Mapping™ (EDM) analysis capabilities for rapid and automated analysis of multiple experiments using different conditions to identify potential biomarkers, thus helping to identify significant differences between samples and groups of samples.

6.1.5 Applications of Proteomics

Proteomics encompasses four principal applications. These are: 1), mining, 2), protein expression profiling, 3) protein network mapping and 4) posttranslational modification characterisation (Liebler 2002).

Mining is simply the exercise of identifying all (or as many as possible) of the proteins in a sample. In mining the proteome is catalogued directly, rather than inferring its composition from expression data of genes e.g., microarray.

Protein expression profiling is the identification of proteins in a particular sample as a function of a particular state (differentiation, developmental state, or disease state) or as a function of exposure to a drug, chemical or physical stimulus. It is most commonly used as a defensive analysis, in which two states of a particular system are compared, e.g., normal and diseased cells or tissues. This information is particularly useful for detecting potential targets for drug therapy in disease.

Protein network mapping and posttranslational modification characterisation are highly specialised proteomic applications involving 1) the determination of the condition and the manner in which proteins interact with each other and in the natural environment and 2) the determination of how, the extent to which and where proteins are modified, respectively (Somiari, Sullivan et al. 2003).
The promise of proteomics in oncology is obvious: cancer is an acquired genetic disease which originates from mutated genes and may lead to aberrant protein expression. A relatively new application known as "clinical proteomics" is involved in the early detection and diagnosis of cancer and its application to the field of medicine has recently been reported (Petricoin, Zoon et al. 2002). It encompasses the translation of proteomic detection technologies and strategies towards the production of better diagnostic tests and therapies (Granger, Van Eyk et al. 2004). Changes observed in the proteome of a patient can be utilized directly as biomarkers or used in drug development.

6.1.6 SELDI-TOF MS Applications

The simplicity and high-throughput features of SELDI-TOF MS in providing rapid proteomic profiling of proteins/peptides from different biological specimens make it a useful platform to search for cancer biomarkers and for developing non-invasive clinical tests.

The first landmark study using SELDI-TOF to demonstrate the potential of utilizing a proteomic approach, specifically the serum proteomic signature, for stratifying patients in the clinic was published by Petriocin et al (Petricoin, Ardekani et al. 2002). SELDI-TOF spectra from 50 unaffected women and 50 patients with ovarian cancer were analyzed and identified a proteomic pattern to discriminate healthy patients from those with malignant disease. Each of the spectra comprised of 15,200 data points -- one for every protein or peptide. Training sets were generated assigning each specific protein a defined intensity value reflective of its relative abundance. This information was used to find the optimal pattern to segregate the two training sets 100 percent of the time using a combination of intensities. This was then tested on unknown samples. The pattern was also tested on women with common benign gynecologic conditions such as ovarian cysts, fibroids, endometriosis, and general inflammatory diseases. The study yielded 100% sensitivity and 95% specificity.

A relatively large screening study for potential biomarkers was conducted using serum from 169 women, including 103 breast cancer patients at different clinical stages, 41 healthy women and 25 patients with benign breast diseases (Li, Zhang et al...
This analysis resulted in a diagnostic pattern which discriminated healthy from breast cancer serum with 93% sensitivity and 91% specificity. Sauter et al. in a small study of 20 patients and 13 healthy controls reported the finding of two potential biomarkers at \( m/z \) 6500 and 15,940 which were over-expressed in nipple aspirates from patients with breast cancer (Sauter, Zhu et al. 2002). In addition to diagnosis, SELDI TOF-MS has been used to systematically examine proteomic changes in plasma of breast cancer patients after chemotherapy treatment with paclitaxel (Pusztai, Gregory et al. 2004). One peak at \( m/z \) 2790Da was induced in patients after treatment.

SELDI TOF-MS has also been used to improve current screening strategies based on PSA testing. Prostate specific membrane antigen, captured by SELDI-TOF immunoassay, was more sensitive in differentiating between benign and malignant prostatic disease compared to serum PSA levels (Xiao, Adam et al. 2001). Wright et al used serum proteomic profiling data from 326 subjects (167 prostate cancer patients, 77 patients with benign prostatic hyperplasia, and 82 healthy men) to train and develop a decision tree consisting of nine mass spectral features in a discriminatory pattern. This proteomic pattern correctly classified 96% of the subjects in the training cohorts, achieving a sensitivity of 83% and specificity of 97% (Wright Jr, Cazares et al. 1999).

Mortality figures in patients with pancreatic carcinoma (PCa) remain poor because of the lack of early detection methods. Bhattacharyya et al used a SELDI based approach to identify biomarkers in the serum proteome for the early detection of resectable PCa (Bhattacharyya, Siegel et al. 2004). Protein profiles were generated using IMAC ProteinChips from the sera of 49 PCa patients and 54 unaffected individuals. The samples were randomly divided into a training set (69 samples) and test set (34 samples) and two multivariate analysis procedures, regression tree and logistic regression were used to develop classification models from these spectral data which could distinguish PCa from control serum samples. In the test set, both models correctly classified all the PCa patient serum samples (100% sensitivity). Using the decision tree algorithm, a specificity of 93.5% was obtained, whereas the logistic regression model produced a specificity of 100%. These results suggest that high-
throughput proteomics profiling has the capacity to provide new biomarkers for the early detection and diagnosis of PCa.

Many of the potential clinical applications of SELDI clearly rely on its ability to provide multi-biomarker patterns that discriminate diseased from control states. SELDI TOF-MS is also a powerful analytical tool in basic research which can identify individual proteins. This is best illustrated in the study by Zhang et al in the discovery of CD8 cell anti-HIV factor (CAF) (Zhang, Yu et al. 2002). Some individuals infected with HIV who do not progress to AIDS (i.e. long-term non-progressors or LTNPs) produce CAF which inhibits HIV replication. The identity of CAF remained unknown, until Zhang et al. examined the stimulated CD8 cell culture media from LTNPs by SELDI TOF-MS and observed a cluster of three proteins around 3000Da. They further enriched these proteins and confirmed their identity as the α-defensins 1, 2, and 3. This study has become a classic example demonstrating the power of SELDI TOF-MS to simultaneously identify several proteins. SELDI has also been used to characterize posttranslational modifications of proteins. The earliest application was the study in which the phosphorylation of the apoptotic protease caspase-9, indicated by an 80Da mass shift, was demonstrated using SELDI TOF-MS (Cardone, Roy et al. 1998).

6.1.7 Proteomics in Haematology

Proteomic techniques have been used in both benign and malignant haematological conditions. Comparative proteomic analyses of human CD34+ stem cells and mature CD15+ myeloid cell isolated from human cord blood have been reported by Tao et al. The authors studied the extracted cytosolic proteins by 2-DE and MALDI-TOF MS and could show that CD34+ stem cells have a larger proteome than mature CD15+ myeloid cells, and that many stem cell-associated proteins are dramatically down-regulated as the CD34+ cell undergo differentiation (Tao, Wang et al. 2004). Similarly, different proteomic techniques (2-DE, multidimensional chromatography followed by various techniques of MS identification) have been applied to identify red cell membrane as well as cytosolic proteins and platelet proteins either at different stages of activation or at different locations inside the cell (O'Néill, Brock et al. 2002; Kakhniashvili, Bulla et al. 2004).
Leukemias and lymphomas are two groups of diseases which have been the subject of several studies using proteomic techniques. Age at diagnosis is an important prognostic indicator in childhood ALL (Reaman, Zeltzer et al. 1985). Infants less than 1 year old at the time of diagnosis have a poor outcome. Hanash et al analyzed the polypeptide patterns of leukaemic cells of infants and older children with ALL using 2-D PAGE (Hanash, Kuick et al. 1989). Different levels of a polypeptide identified as a phosphorylated form of Hsp27 were observed between infants and older children. The phosphorylation of Hsp27 was reduced in infants compared to older children with ALL. The response to initial glucocorticoid therapy in childhood acute lymphoblastic leukaemia (ALL) reliably predicts the response to multiagent chemotherapy. Patients resistant to glucocorticoids (prednisone poor responders (PPR)) have a poorer event-free survival compared to glucocorticoid-sensitive patients (prednisone good responders (PGR)) (Riehm, Reiter et al. 1987). Lauten et al performed a case–control study to investigate differential protein expression in leukaemic blasts from PGR and PPR childhood ALL patients using 2D PAGE and SELDI-TOF MS. A number of proteins were found to be over expressed in PPR including Valosin-containing protein (VCP). A high VCP expression was associated with poor prednisone response in childhood ALL patients (Lauten, Schrauder et al. 2006).

The significance of new prognostic factors such as the somatic hypermutation in the variable region of the Ig heavy-chain locus and CD38 expression in chronic lymphocytic leukemia is emerging. Proteomics of the two types of chronic lymphocytic leukemia have been recently studied by Cochran et al., using MALDI-TOF MS (Cochran, Evans et al. 2003). Significant differences in patterns of protein expression such as F-actin-capping protein b subunit, 14-3-3 b protein, and laminin-binding protein precursor were found between the cases with and without somatic mutation and were significantly increased in the group with mutated Ig genes. No specific differences were found between CD38-positive and CD38-negative patient samples using the same approach. Similarly, another study using 2D PAGE and MALDI-TOF MS found phosphorylated HS1 protein over expression in patients with unmutated IgVH genes and CD38 expression to be associated with significant shorter median survival times (Scielzo, Ghia et al. 2005).
In both acute and chronic haematological diseases, effective biomarkers are essential for diagnosis, risk stratification, prognosis and determining the effectiveness of therapeutic treatment. Furthermore, they could lead to new pharmacological targets. If these markers could be measured in a readily accessible body fluid such as serum, not requiring a tissue biopsy, it would have a major impact on future cancer diagnosis and treatment monitoring.

6.1.8 Aim of the Project

Chronic lymphocytic leukemia is the commonest hematological malignancy in the western world. CLL has a clinically variable course but conventional prognostic parameters (stage, white cell count, LDH and pattern of marrow infiltration) have been of limited value in predicting disease tempo and response to treatment. Now however a patient’s clinical course can be accurately predicted by a combination of immunophenotyping (CD38 expression, ZAP70 expression), cytogenetics (trisomy 12, deletion 11q, deletion 17p deletion 13q14) and finally investigation of the somatic hypermutation status of the immunoglobulin heavy chain gene in the leukaemic cells.

More recently, microarray analysis has been used in CLL to investigate if there are differences in the gene expression profiles of patients. Genes with different expression levels affecting survival and or clinical staging have been found in addition to differences between patients with mutated IgV\textsubscript{H} genes as compared to those with unmutated IgV\textsubscript{H} genes. The genome in an individual remains relatively static whereas the proteome constantly changes depending on the physiological environment. Gene expression is regulated by controlling protein expression at several points during the process of DNA transcription and processing to mRNA and subsequent translation into polypeptide chains, followed by folding and post-translational modifications, such as phosphorylation and glycosylation, to yield functional proteins. The final effector pathway therefore, in any cell is protein based and a gene expression analysis approach alone is probably not sufficient. This is evident by finding no difference in CD38 at the transcriptome level of CLL patients.

Conventionally CLL is considered incurable and treatment has been aimed at keeping the patient symptom free with the minimum amount of chemotherapy. Increasingly
younger patients are diagnosed and more effective or aggressive therapeutic approaches need to be explored. We have treated 39 patients with a novel chemotherapy combination (FCR) at SJH, which combines a purine analogue (Fludarabine-F), alkylating agent (Cyclophosphamide-C) and an antibody directed against a pan-B antigen known as CD20 (Rituximab). Clinically this combination has been very effective, inducing a high rate of complete remissions.

Proteomic studies have been carried out in several haematological diseases, however there are relatively few in CLL. This offers opportunities for translational investigations where potential markers and correlated proteins (and their isoforms) can be identified. Our initial approach will therefore be of a global protein pattern assessment in a group of CLL patients exposed to a novel combination chemotherapy regimen. The profile patterns will then be related back to the cytogenetic, immunophenotypic and hypermutation pattern and to clinical outcome. We would hope to pick up differences in initial and response profiles corresponding to different prognostic groups and clinical outcome. Further proteomic analysis would be carried out to identify any potential valuable prognostic indicator "biomarker". The goal is to understand the disease and to elucidate the underlying disorder at a molecular level. Of course, the final objective is to try and find proteins which may be involved in CLL aetiology, variable clinical behaviour, response to treatment and outcome.

6.2 Materials and Methods

6.2.1 Patients

Thirteen consecutive patients with CLL undergoing chemotherapy at SJH between January 2002 and December 2005, who were part of a novel chemoimmunotherapy trial FCR (Fludarabine, Cyclophosphamide and Rituximab) were selected to undergo proteomic analysis. Four additional patients, with similar clinical and laboratory characteristics but who were not receiving treatment, were included in the analysis as controls. The patients selected for treatment were either treatment naïve, < 65 years with progressive stage A disease and high risk features or more advanced stage
Biologically fit patients <75 years who had failed Fludarabine therapy or had high risk symptomatic disease requiring treatment were also included. Progressive stage A was defined as a short lymphocyte doubling time; downward trend in haemoglobin & platelets; increased size of nodes/spleen and B-symptoms. High-risk patients were defined on the basis of advanced stage, short lymphocyte doubling time, LDH >450 IU/l, IgV_H mutational status, CD38 expression and ZAP 70 expression.

### 6.2.2 Treatment

All patients received Fludarabine (25 mg/m²/d IV or 40 mg/m²/d PO) and Cyclophosphamide (250 mg/m²/d IV/PO) from day 1-3 of each cycle. Cycles were repeated every 28 days depending on the recovery of neutrophil count to >1 x 10⁹/l and platelets to >75 x 10⁹/l. Patients achieving CR, as per National Cancer Institute working group criteria (NCI-WG), after cycle 3 were given one additional cycle of chemotherapy, otherwise a maximum of 6 courses were given. Rituximab (375 mg/m² IV infusion) was introduced from day 1 of the 2nd cycle following pre-medication with paracetamol 1 gm PO and Chlorphenaramine 10mg IV.

### 6.2.3 Sample Collection

Thirteen patients who received this treatment had plasma stored at -70°C, from three time points i.e. pre-treatment, after the first cycle and at completion of therapy. The samples were collected in the Haematology Day Ward before the initiation of the treatment into 10 ml EDTA bottles. Samples for plasma collection were processed up to 2 hours after collection. Cell storage and sample preparation for FISH, Immunophenotyping, DNA and RNA extraction has already been described in section 2.2.

For plasma collection, 10 ml of peripheral blood collected at the three time points was centrifuged at 2000 rpm (IEC Centra GP8, MA, USA) for 10 min at 20°C. After centrifugation, plasma was collected into 3 Eppendorf tubes in 1ml aliquots and stored at -70°C for later use.
6.2.4 Sample Assessment

This involved two steps 1) sample fractionation to reduce sample complexity and 2) analysis of the fractionated samples, on four different protein chip surfaces to identify the best surface for comprehensive biomarker identification experiments, as shown in figure 6.16.

![Sample fractionation and initial assessment of pooled samples on all 4 chip surfaces](Ciphergen Biosystems)

**Figure 6.16** Sample fractionation and initial assessment of pooled samples on all 4 chip surfaces (Ciphergen Biosystems)

### 6.2.4.1 Fractionation

All the 3 time point plasma samples (labelled as A, B and C respectively) were brought to 20°C and centrifuged at 13000 rpm for 10 min. There were 13 pre-treatment samples (A1-13), 12 after cycle 1 (B1-12) and 11 after completion of the treatment (C1-11). Four samples from patients with CLL who had not received any treatment were included as controls while pooled samples from the 3 time points were also fractionated.
The plasma samples were fractionated by pH gradient chromatography using anion exchange beads supplied by Ciphergen Biosystem Inc., USA (Expression Difference Mapping Kit) according to manufacturers instructions. The kit allows high throughput fractionation and is supplied in the form of a 96-well filtration plate (refer to Appendix A for details of the solutions and buffers). Briefly, 50µl of sample was added to 80µl of U9 buffer and mixed. 70µl of this mixture was added to 70µl of U1 buffer and added to the rehydrated plate and eluted in a stepwise manner, by altering the pH of the wash buffer, until six fractions were collected. Each of the six fractions was collected twice and the two collections pooled. Samples in the six collection micro-plates were stored at -70°C after sealing. A flow chart for the fractionation steps is shown in Figure 6.17.
Figure 6.17 A flow chart of the Q hyperD F beads plate fractionation steps.
6.2.5 Protein chip protocols

After fractionation, each pooled sample (A, B and C) and the controls resulted in 6 fractions. For initial analysis, the bioprocessor was assembled with 3 chips of each type (H50, IMAC, CM10 and Q10) to test all the fractions from the pooled samples on the 4 different types of arrays.

6.2.5.1 IMAC30 profiling protocol

50μl of 0.1 M copper sulphate solution was added to each spot, and incubated for 10 min at RT with vigorous shaking. The metal solution was removed, 200μl of de-ionized water (dH2O) added to each well, and cassette incubated at RT for 2 min with shaking. The dH2O was removed from the wells and the step repeated. 200μl of sodium acetate buffer pH 4 (neutralisation buffer) was added to each well, and cassette incubated for 5 min with vigorous shaking. The neutralising buffer was removed and the wells washed with 200μl of dH2O for 2 min at RT with shaking. After removing the water, 200μl of IMAC binding buffer (containing 0.1M Sodium phosphate and 0.5M NaCl, pH 7.0) was added, and cassette incubated for 5 min at with shaking. The binding buffer was removed. 50μl of the plasma sample (containing 20μl of plasma fraction diluted in 30μl of buffer) was added to each well immediately. The cassette was incubated with vigorous shaking for 60 min. Samples were removed and each well was washed with 200μl of dH2O for 5 min with vigorous shaking. The binding buffer was removed and the step repeated 2 more times. Finally each well was washed with 200μl dH2O for 5 min with shaking. The arrays were removed from the bioprocessor and air dried for 20 min. 0.8μl of the Energy Absorbing Matrix (EAM, Ciphergen Biosystem) solution prepared at the same time was applied to each spot and left to dry.

6.2.5.2 CM10 profiling protocol

200μl of CM10 low stringency binding buffer (Sodium acetate, pH 4.0) was added to each well, and incubated for 5 min at RT with shaking. The binding buffer was removed and the step repeated. 50μl of sample (containing 20μl of fractionated plasma in 30μl binding buffer) was immediately added to the corresponding wells and
incubated with vigorous shaking for 60 min. Samples were removed from the wells and each well was washed with 200μl of binding buffer for 5 min with agitation. The binding buffer was removed and the step repeated 2 more times. Each well was washed twice with 200μl dH₂O for 5 min with agitation. The wells were drained. The arrays were removed from the bioprocessor, and air dried for 20 min. 0.8μl of the EAM solution was applied to each spot and air dried.

6.2.5.3 Q10 profiling protocol

200μl of Q10 binding buffer (Tris-HCl, 10-100mM, pH 7.5-9) was added to each well, and cassette incubated for 5 min at RT with shaking. The binding buffer was removed and the step repeated. 50μl of the sample (containing 20μl of fractionated plasma diluted in 30μl of binding buffer) was added immediately. The cassette was incubated with vigorous shaking for 60 min. Samples were removed from the wells and each well washed with 200μl of binding buffer for 5 min with agitation. This step was repeated two more times. The wells were washed with 200μl dH₂O for 5 min with agitation draining the wells between each step. Each well was washed twice with 200μl dH₂O for 5 min with agitation. The wells were drained. The arrays were removed from the bioprocessor, and air dried for 20 min. 0.8μl of the EAM solution was applied to each spot and air dried.

6.2.5.4 H50 protein profiling protocol

The arrays were first washed with 50μl of 50% methanol for 5 min. The wash solution was removed and the wells were allowed to air dried for 1 hour. 200μl binding solution (10% acetonitrile, 0.1% trifluoroacetic acid) was added, and the cassette incubated for 5 min at RT with vigorous shaking. The buffer was removed and the step repeated. 50μl of the sample (containing 20μl of fractionated plasma diluted in 30μl of binding buffer) was added immediately to each well and the cassette incubated with shaking for 60 min. Samples were removed from the wells and each well washed with 200μl of binding buffer for 5 min with agitation. This step was repeated two more times. The wells were washed with 200μl dH₂O for 5 min with agitation draining the wells between each step. Each well was washed twice with 200μl dH₂O for 5 min with agitation. The wells were drained. The arrays were
removed from the bioprocessor, and air dried for 20 min. 0.8µl of the EAM solution was applied to each spot and air dried.

6.2.6 Data Collection Protocols

6.2.6.1 Spot Protocol
The chips were placed in the autoloader of the protein chip reader for analysis and data collection. After manually loading the chips, the first step was to create a spot protocol. This was done in the auto setup dialogue box to create capturing conditions for each protein spot, shown in figure 6.18.

![Auto Setup Dialogue Box](image)

**Figure 6.18** Spot Protocol

An optimized protocol both for the low mass and high mass was created for each chip, as shown in Table 6.1. A chip protocol was created next.
<table>
<thead>
<tr>
<th>Spot Protocol</th>
<th>Low Mass</th>
<th>High Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laser Intensity</td>
<td>185</td>
<td>195</td>
</tr>
<tr>
<td>Detector Sensitivity</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Optimization</td>
<td>Centre</td>
<td>Centre</td>
</tr>
<tr>
<td>Mass Optimization Range</td>
<td>2000 to 30000Da</td>
<td>20000 to 160000Da</td>
</tr>
</tbody>
</table>

Table 6.1 Spot protocols capture setting for low and high mass.

6.2.7 Data Analysis

After adding the acquisition protocols to the bioprocessor, the cassette was run automatically by pressing the start button in the software. Once the cassette run had finished, the data was analyzed for quality, noise estimation and baseline adjustment using ProteinChip software before exporting to Ciphergen express.

6.2.7.1 Spectrum Processing

Baseline Adjustment and Filtering
The peak heights were expressed as the distance from the top of the peak to the calculated baseline. The spectra were initially analyzed without the baseline subtraction on, in order to determine spectrum quality.

Filtering
Filtering was set to ‘On’ by default and was used to lower the spectrum noise. It increases the signal to noise ratio.

Noise Estimation
It is important to measure the noise accurately as it affects the signal quality. The starting noise value was set to be higher than the EAM slope so that the chemical
noise generated by the EAM could be cleared. Baseline subtraction was used to eliminate the noise generated by the matrix.

6.2.8 Ciphergen Express

Spectra were exported into Ciphergen express software as xml files. The samples were labelled with patient sample identification numbers and the time point for each sample. After confirming sample alignment and noise filtering the samples were normalized across all spectra using the Total Ion Current (TIC) method, the sum of ion intensities between the 2000 and 200000 kDa mass. Normalisation was used to rescale the spectra so that all spectra had the same observed TIC to minimise variation between spectra. The minimum $m/z$ for detection was set to start at 2000Da. Spectra for which the normalization factors were either $>2$ or $<0.5$ were discarded.

6.2.9 Biomarker Analysis

Cluster analysis (groups of peaks with similar masses) was used for biomarker discovery. First an expression difference map (EDM) was created to specify conditions for cluster analysis such as sample name, sample group, sample type (IgVH mutated or unmutated, CD38+ve or CD38-ve, FISH abnormalities and between the different time points Pre-treatment, after Cycle 1 or at the end of therapy). Next clustering parameters were set for peak detection which was done in two passes. The first pass s/n specifies the signal to noise ratio during the first pass of the EDM peak detection and was set at 5. This detection step has lower sensitivity but detects the most abundant proteins, whereas the second pass was set at 3 which is more sensitive for proteins expressed at a low level. These are then added to the peaks detected during the first pass. The analysis is done automatically by first detecting the peaks and then clustering them according to the set parameters with the results appearing as a cluster table. Peaks were detected and clustered across a mass range of 2000kDa to 200000kDa, present across a minimum of 10% of the spectra with a signal-to-noise ratio of 3, a valley depth of 1 and peak height of 3.

P-value statistics were performed automatically for the analyzed groups for comparison. P-values were then displayed in the order of significance by double
clicking the column. For the statistically significant proteins the differences between the groups were visualized as scatter plots by clicking on the cluster plot menu and selecting the “cluster plot by m/z” tab. Proteins with P-values of ≤ 0.05 were considered as potential biomarkers.

ROC plots and the area under the curve (AUC) were generated in Ciphergen express and were used to measure the sensitivity and specificity of a marker. Values close to 1 indicate a good biomarker. Heatmap views using hierarchical clustering were also generated to visualize the differences between the groups for the significant peaks.

6.2.10 Protein Identification

The first step after selecting potential biomarkers for identification is to isolate and purify the protein sufficiently for trypsin digestion and to allow peptide mapping.

6.2.10.1 Large scale protein fractionation

A column containing 200μl of Q Hyper D®F resin (BioSepra) was spun at 2000 rpm and the supernatant removed. Contents from another similar column were added to it before a second spin for a total volume of 200 μl of Q Hyper D®F resin. The resin was equilibrated with 500μl of U1 buffer (1M urea, 0.2% CHAPS, 50 mM Tris-HCL, pH 9) each and incubated for 30 minutes. Plasma samples and controls were thawed on ice. 100μl of each sample was added to an eppendorf containing 150μl of U9 buffer (9M urea, 2%CHAPS, 50 mM Tris-HCL, pH 9) and incubated at 20°C for 30 minutes. The sample was then diluted with 250μl of U1 buffer and the contents mixed before being applied (500μl) to the anion exchange resin and gently rotated for 30 minutes to ensure thorough mixing.

The flow through fraction, containing unbound material, was collected from the column by centrifugation at 2000rpm for 1 minute. The column was then washed with gentle mixing on a rotator for 10 minutes using 500μl of 50mM Tris-HCL, pH 9, 0.1% OGP buffer. This was pH9 fraction (fraction 1). This Step was repeated with washing buffers at different pH to collect fractions 2-5. The final wash was done
using 500μl of 33% isopropanol, 17% ACN and 0.1% TFA to collect the organic fraction, fraction 6.

The fractions were tested on the CM10 and IMAC protein chips along with the original sample from the first fractionation using the spot protocol to decide on the fraction containing the biomarker.

6.2.10.2 Reverse Phase Chromatography for protein purification

Further purification and enrichment of the marker was carried out by reverse phase chromatography (RPC). 50μl of reverse phase chromatography PLRP-S 15-20μm 300°A beads (Polymerlabs, UK) were equilibrated with 500μl of 10% Acetonitrile (ACN) and 0.1% Trifluoroacetic acid (TFA) in a 1 ml eppendorf tube. The beads were mixed for 5 min on a rotator and centrifuged at 8000 rpm for 1 min and the supernatant was carefully removed making sure not to aspirate any beads. This step was repeated twice more. 250μl of the sample was then added to the beads in addition to 25μl of 100% ACN and 0.25μl of 100% TFA (final concentration of 10% ACN and 0.1% TFA). The beads were mixed at 20°C for 30 min and then centrifuged for 1 min at 8000 rpm. The supernatant was removed into a clean eppendorf and labelled as 'unbound fraction'. Further fractions were collected after washing the beads and eluting the bound proteins in 500μl of progressively increasing concentrations of ACN (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%) and 0.1% TFA. Each elution buffer was added in sequence to the tube and mixed on a rotator, spun at 8000 rpm for 1 minute and the supernatant collected in a fresh eppendorf. Proteins in each eluted fraction were tested by running 2μl of each fraction along with 5μl from the same samples pre RPC fractionation on an NP20 protein chip. This also identified the fraction containing the biomarker of interest. The remainder of the eluted fractions were dried in a DNA 100 SpeedVac (Savant, UK) at medium heat for 3 hours or until dry and stored at -20°C until required.

6.2.10.3 1D SDS-PAGE

After identifying the fractions containing the biomarker of interest on the NP20 chips, the samples were run on a Tris-Glycine or Novex® Tricine 10-20% Gel, (Invitrogen,
UK) in an Xcell SureLock™ Mini-Cell gel tank (Invitrogen, UK). The Xcell SureLock™ Mini-Cell gel tank was assembled with the mini Tricine gel according to the manufacturer’s instruction. The upper buffer chamber was filled with 200 ml of 1X Tricine SDS running buffer. The lower buffer chamber was filled with 600 ml of 1X Tricine SDS running buffer.

The samples were prepared by adding 5μl of 2X Novex® Tricine SDS sample buffer and 5μl of deionized water. The samples were thoroughly mixed by pipetting before being spun down at 8000rpm for 1 minute in a centrifuge. Samples were incubated at 100°C for 10 min then mixed again and centrifuged for 1 min at 8000 rpm. 10μl of each sample was loaded into the corresponding well on the mini gel. 10μl of SeeBlue® Plus2 pre-stained 1X standard (Invitrogen, UK) was also loaded at the same time. The gel was run at 110V constant at an average current of 70 mAmp for two hours.

After electrophoresis, the gel was stained with Coomasie blue stain using Novex colloidal blue staining kit (Invitrogen, UK). This was prepared according to the manufacturer instructions. The gel was carefully removed from the plastic casing and put into a container containing the staining solution. The gel was rocked gently in the staining solution for 3 hours at 20°C. It was then de-stained using 200ml of dH₂O overnight. After de-staining the gel was photographed using LabWorks™ Imaging Acquisition software and UVP BioImaging System (UVP Inc., Upland, CA, USA). For better visualization of the bands, the Tricine gel was re-stained using SilverQuest™ Silver Staining Kit (Invitrogen, UK), according to the manufacturer instructions. This silver stain is compatible with mass spectrometry. Briefly, after de-staining the Coomassie blue, the gel was rinsed with ultra pure water for ten minutes followed by one hour in the fixative solution with gentle shaking. The gel was washed in 30% ethanol for 10 minutes. 100 ml of sensitizing solution was added and incubated for 10 minutes followed by a further wash in 100 ml of 30% ethanol for 10 minutes and 100 ml of ultra pure water for 10 minutes. The gel was then incubated in 100 ml of staining solution for 15 minutes and a very quick wash with 100 ml of ultra pure water for 20-60 seconds. The gel was incubated in 100 ml of developing solution for 4-8 minutes until bands started to appear and the desired band intensity was reached. To stop further increase in staining intensity, 10 ml of Stopper solution was
added directly to the gel still immersed in developing solution and agitated for 10 minutes. Finally the gel was washed in 100 ml of ultra pure water for 10 minutes and photographed using LabWorks™ Imaging Acquisition software and UVP BioImaging System (UVP Inc., Upland, CA, USA).

**6.2.10.4 Peptide identification**

For protein identification the samples were sent to Proteomics & Peptide Synthesis, MRC Clinical Sciences Centre Faculty of Medicine, Imperial College (Hammersmith Campus, Du Cane Road, London).

Briefly, once the gel had been stained and bands of interest located, the bands were excised, destained and trypsin digested. Initial screening of the digested peptides was carried out by peptide-mass fingerprinting (PMF) using matrix-assisted laser desorption ionisation (MALDI) time-of-flight mass spectrometry. Measured peptide masses were scanned against the MOWSE peptide-mass database to determine if the protein is known. An assignment of protein identity can often be made at this stage. However, for increased mass accuracy (< 0.05%), further analysis was carried out by sequencing individual peptides by low-energy collision activated dissociation (CAD) using hybrid quadrupole-TOF mass spectrometers fitted with nanoelectrospray sources. Collision spectra were screened against MASCOT protein sequence database to obtain possible sequence matches.

The program accepts basic peptide-mass fingerprint data, but complete or partial linear sequence can also be used as additional search parameters eliminating any requirement to perform parallel digests with different enzymes. In addition, the program performs fully automatic correlation between experimental MS/MS data and peptide sequences contained in whole-protein or expressed sequence tag (EST) databases as shown in Figure 6.19.
Figure 6.19 Steps for protein identification (http://www.matrixscience.com/pdf/Brochure_01-2-2005.pdf)
6.3 Results

6.3.1 General Results

A total of 13 consecutive patients attending St. James Hospital for modified FCR chemoimmunotherapy were accrued and included in the proteomic analysis between July 2002 and July 2004. All patients fulfilled the morphological and immunophenotypic criteria of CLL (CD5/CD19+ive, CD23+ive, FMC7-ive, weak surface immunoglobulin expression and light chain restriction). Patients received treatment because of disease progression, advanced disease or relapse after prior therapy (Patient selection and treatment protocol discussed in detail in chapter 5). Plasma samples were collected at 3 time points, at the start of treatment, after cycle 1 and after finishing therapy (4 or 6 cycles depending on remission status).

There were twelve male patients and 1 female patient, with a median age at treatment of 58 years (range 41-71), as shown in Table 6.2 with other clinical and biological features.

<table>
<thead>
<tr>
<th>Table 6.2 Biologic and clinical characteristics of treated patients</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No of Patients</strong></td>
</tr>
<tr>
<td><strong>Sex, Male / Female</strong></td>
</tr>
<tr>
<td><strong>Median Age (Yrs)</strong></td>
</tr>
<tr>
<td><strong>Stages†</strong></td>
</tr>
<tr>
<td><strong>Median WCC x 10⁹/l</strong></td>
</tr>
<tr>
<td><strong>Median Hb g/dl</strong></td>
</tr>
<tr>
<td><strong>Median Platelet Count x 10⁹/l</strong></td>
</tr>
<tr>
<td><strong>LDH iu/l</strong></td>
</tr>
<tr>
<td>**β₂m mg/dl *</td>
</tr>
<tr>
<td><strong>IgVH: Mutated / Unmutated</strong></td>
</tr>
<tr>
<td><strong>CD 38: Negative / Positive</strong></td>
</tr>
<tr>
<td><strong>ZAP-70: Negative / Positive</strong></td>
</tr>
</tbody>
</table>

† Binet stage at treatment
* β₂m count was obtained in 10 patients
Six patients had no abnormality detected on FISH. Four patients had 13q14 deletion as a sole abnormality, two patients had trisomy 12 and one patient had 11q deletion. Patients with trisomy 12 and 11q deletion also had 13q14 deletion as an additional abnormality.

Plasma samples from four other patients with similar clinical and biological features at presentation were also collected and used as controls. These patients did not receive any treatment and had only one plasma sample taken at the time of diagnosis. There were 3 male patients and 1 female with a median age of 66 (Range 62-73), as shown in Table 6.3 with other clinical and biological features.

<table>
<thead>
<tr>
<th>Table 6.3 Biologic and clinical characteristics of the Untreated patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of Patients</td>
</tr>
<tr>
<td>Sex, Male / Female</td>
</tr>
<tr>
<td>Median Age (Yrs)</td>
</tr>
<tr>
<td>Stages†</td>
</tr>
<tr>
<td>Median WCC x 10^9/l</td>
</tr>
<tr>
<td>Median Hb g/dl</td>
</tr>
<tr>
<td>Median Platelet Count x 10^9/l</td>
</tr>
<tr>
<td>LDH iu/l</td>
</tr>
<tr>
<td>β2m mg/dl</td>
</tr>
<tr>
<td>IgVH: Mutated / Unmutated</td>
</tr>
<tr>
<td>CD 38: Negative / Positive</td>
</tr>
<tr>
<td>ZAP-70: Negative / Positive</td>
</tr>
</tbody>
</table>

† Binet stage at diagnosis

One patient had no abnormality detected on FISH while another had 11q deletion as an isolated abnormality. The other two patients had 17p deletion and trisomy 12 in addition to 13q14 deletion as an additional abnormality.

After fractionation, the initial analysis was carried out on pooled samples from the three time points using four types of protein chips; Immobilized metal affinity capture
(IMAC30), weak cation exchanger (CM10), strong anion exchanger (Q10) and the hydrophobic protein chip (H50) to select the surface and fraction which gives the best spectra for use when analysing all the samples. The spectra from each fraction of the 3 pooled samples were analyzed, as shown in figure 6.20.

**Figure 6.20** Comparative analyses of fraction 6 for the pooled samples on 4 ProteinChips, IMAC30 and CM10 chips show good peak intensity (black arrows) with minimal background interference (blue arrows).

CM10 and IMAC ProteinChips gave the best spectral results with good peak intensity and less background noise. Low mass spectrum showed better quality and quantity of peaks in the optimisation range of 2kDa-30kDa while the high mass spectrum had fewer peaks in its optimised range of 20kDa-160kDa.

A more detailed analysis was carried out on CM10 and IMAC chips. Fractions 2 and 6 on CM10 chip and fraction 6 on IMAC were picked for further analysis of all the samples and controls as they showed better spectral resolution than the other fractions, as shown in figure 6.21.
Individual fractions from 36 patient samples (3 time points) and 4 controls were tested on the 2 selected surfaces resulting in 120 spectra. All the spectra were exported into Ciphergen express software as XML files for further analysis and biomarker discovery. After sample alignment and noise filtering, the samples were normalized across all spectra using the Total Ion Current (TIC) method. Normalisation was used to rescale the spectra so that all spectra had the same observed TIC to minimise variation between spectra.

**Figure 6.21** Fractions 1 to 6 on CM10 ProteinChip. Note minimal albumin peak (blue arrows) on F1 and F6 as compared to other fractions where it has persisted after fractionation.
6.3.2 Biomarker Discovery

After aligning the spectra in Ciphergen express, Cluster analysis tool was used for biomarker discovery. This allows for detection of peaks with similar masses, which are grouped together. First an expression difference map (EDM) was created to specify conditions for cluster analysis such as sample name, sample group, sample type (IgVH mutated or unmutated, CD38+ve or CD38-ve, FISH abnormalities and between the different time points Pre-treatment, after Cycle 1 or at the end of therapy) and for statistical analysis, allowing identification of biomarkers that differentiate the groups (Refer to 6.2.9).

Different groups compared for peak clustering included the 3 time points (A, B and C respectively), untreated (Control) to treated patients, IgVH mutated to those with unmutated Ig genes and CD38+ve versus CD38-ve Groups. FISH categories were merged to create two groups for comparison due to the small number of samples in each category. Those with normal FISH results were combined with 13q14 deletion (Good prognosis) to form one group while trisomy 12 was combined with deletion 11q and deletion 17p (Poor prognosis).

Many proteins were found to have a significantly different (p<0.05) expression between the groups, as shown in Table 6.4. Some of these proteins showed increased expression while others were down-regulated. However most were either expressed in small amounts or were of a very low mass, which would have made protein identification difficult without protein enrichment techniques and they were therefore excluded from the protein identification step.
<table>
<thead>
<tr>
<th>Protein m/z (Da)</th>
<th>CD38+ve</th>
<th>CD38-ve</th>
<th>Mutated</th>
<th>Unmutated</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>9288</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td>↑</td>
<td>↓</td>
<td></td>
</tr>
<tr>
<td>15141</td>
<td>↑</td>
<td>↓</td>
<td></td>
<td></td>
<td>↑</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17380</td>
<td>↑</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17597</td>
<td>↑</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15957</td>
<td>↑</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15385</td>
<td>↑</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4419</td>
<td>↑</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9124</td>
<td>↑</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6517</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78965</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39887</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14020</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8788</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17380</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8871</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4006</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23408</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9908</td>
<td>↑</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11694</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6517</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1716</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1768</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3162</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8933</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17496</td>
<td>↓</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.4 Proteins found to be differentially expressed on expression difference mapping between the various groups analyzed, shown as m/z ratio. (p<0.05)
The following potential markers were selected for protein identification.

1) 9908Da

This protein was differentially expressed between pre-treatment and the post-treatment groups. The protein was found to be up-regulated in the pre-treatment samples with a p-value of (0.00042), as shown in Figure 6.22.

![Graph A](image1)

![Graph B](image2)

**Figure 6.22** (A) High expression of protein 9908Da in the pre-treatment samples. The expression is down regulated after the initiation of therapy and remains so at the end of treatment (Group C not shown). (B) Representative spectra from a patient showing down regulation in the expression of protein 9908.
The marker showed a sensitivity of 83% and specificity of 77% for treatment response when a cut-off value of 2.8 was used as shown in Figure 6.23. The cut-off value relates to the intensity of the protein marker.

Figure 6.23 An ROC curve for protein 9908 showing the sensitivity and specificity levels.

Receiver Operating Characteristic curve (or ROC curve.) It is a plot of the true positive rate (Sensitivity) against the false positive rate (Specificity) for the different possible cutpoints of a diagnostic test. It shows the tradeoff between sensitivity and specificity (any increase in sensitivity will be accompanied by a decrease in specificity). The closer the curve follows the left-hand border and then the top border of the ROC space, the more accurate the test. The closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate the test.

2) 6517Da
This protein was differentially expressed in the pre-treatment and post-treatment groups. The protein was found to be up-regulated in the post-treatment samples with a p-value of 0.02, as shown in Figure 6.24. The protein was also down regulated in CD38 +ve patients.
Figure 6.24 Dot-plot showing up-regulation of protein 6517 after starting the treatment (A). Spectra from a patient with up-regulation in expression of protein 6517 (B). 2A (Pre-treatment) and 2B (Post-treatment)

The marker showed a sensitivity of 66% and specificity of 77% for treatment response when a cut-off value of 27 (peak intensity) was used as shown in Figure 6.25.
Figure 6.25 An ROC curve for protein 6517 showing the sensitivity and specificity levels.

3) 79kDa

This protein was differentially expressed between groups B and C. The protein was down-regulated in the post-treatment samples with a p-value of 0.02, as shown in figure 6.26. It was also under expressed in CD38 +ve patients.

Figure 6.26 Dot-plot showing down-regulation of protein 79000 at the end of treatment.

The marker showed a sensitivity of 66% and specificity of 73% when a cut-off value of 0.46 (peak intensity) was used as shown in figure 6.27.
Figure 6.27 An ROC curve for protein 79kDa showing the sensitivity and specificity levels.

4) 3971Da

This protein was differentially expressed between the two FISH groups. The protein was up-regulated in the patients with poor prognostic cytogenetic abnormalities with a p-value of 0.0002, as shown in figure 6.28.

Figure 6.28 Dot-plot showing up-regulation of protein 3971 in patients with poor prognostic cytogenetic abnormalities.

The marker showed a sensitivity of 91% and specificity of 75% for association with trisomy 12, deletion 11q and deletion 17p, when a cut-off value of 0.8 (peak intensity) was used as shown in figure 6.29.
Figure 6.29 An ROC curve for protein 3971 showing the sensitivity and specificity levels.

5) 11785Da
This protein was differentially expressed between groups A, B and C. The protein was down-regulated in the post-treatment samples with a p-value of 0.0006, as shown in figure 6.30.

Figure 6.30 Dot-plot showing down-regulation of protein 11785 at the end of treatment.

6.3.3 Protein Identification
All the spectra were searched to select samples containing the markers of interest with the highest level of expression. These samples were purified and enriched for protein
identification. This step also entails de-salting the sample and dissolves the marker in a volatile organic solvent which can be evaporated off prior to loading on a gel.

The selected samples were purified using the Q Hyper D® F spin columns which yielded 6 fractions for each sample. The resultant fractions were then tested on CM10 and IMAC ProteinChips along with the original plate fractionated sample to select the best fraction for enrichment. The fractions containing the biomarkers of interest were then enriched using the reverse phase chromatography beads, as shown in Table 6.5.

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>Sample ID</th>
<th>Column Fraction</th>
<th>RPC Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>6517</td>
<td>7B</td>
<td>F1</td>
<td>40%</td>
</tr>
<tr>
<td>11785</td>
<td>5A</td>
<td>F2</td>
<td>40%</td>
</tr>
<tr>
<td>3971</td>
<td>2C</td>
<td>F6</td>
<td>Unbound</td>
</tr>
<tr>
<td>79000</td>
<td>5A</td>
<td>F2</td>
<td>40%</td>
</tr>
<tr>
<td>9908</td>
<td>5A</td>
<td>F6</td>
<td>40%</td>
</tr>
</tbody>
</table>

Table 6.5 Samples and fractions containing the biomarkers of interest

The fractions from RPC enrichment and the original fraction were then analysed on an NP20 ProteinChip to select the fractions containing the biomarkers of interest as shown in Figure 6.31. Sample 2B (10% fraction) was used as negative control. Most of the biomarkers were found to be in the 40% fraction except 3971 Da, which was mainly expressed in the unbound fraction of sample 2C.
Figure 6.31 Representative spectra from an NP20 run on the RPC enriched fractions showing the biomarkers of interest. (A) Proteins 3971Da, 9908Da and 6517Da with a negative control, (B) Proteins 11785Da and 79kDa with a negative control. The masses on the enriched fractions are slightly different from the original masses at the discovery phase. This is normal and reflects the slight inaccurate mass assessment at this stage of the experiment.

The selected fractions were dried to remove the solvent and run on Novex® Tricine 10-20% Gel. The gel was first stained using colloidal comassie and then to aid improved visualization of the bands, the gel was re-stained using silver stain (Figure 6.32).
Figure 6.32 10-20% Tricine 1D-SDS gel electrophoresis stained using colloidal comassie blue (A). Note, low mass bands are not visible. (B) Same gel re-stained with silver. Low mass biomarker 6517, between the 7000Da and 4000Da markers is now visible in lane 2 and 3. (L) Ladder. (L1-L8, Lane 1 to 8)
Except for the 3971Da protein all other proteins could be detected and were excised from the gel for further analysis and protein identification. The results collected for each protein from the MALDI and Q-TOF were analyzed using MASCOT software and protein database for protein identification. The top scoring proteins originated from spots corresponding to the anticipated MW on the gel. Further confidence in the identity of the individual proteins comes from individual ions scores of > 64 (MOWSE score) which indicated identity or extensive homology (p<0.05). No contaminants were detected in any of the results. Q-TOF data was first searched for all species to confirm the identification of porcine trypsin peptides, which was used as an internal control. The MASCOT search was then restricted to human matches only. The identified proteins are shown in Table 6.6.

<table>
<thead>
<tr>
<th>Lane No</th>
<th>MW (Da) / pI</th>
<th>Name</th>
<th>Total score</th>
<th>Coverage%</th>
<th>Peptide matches</th>
<th>Delta*</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>6627 / 7.93</td>
<td>apolipoprotein C-I</td>
<td>64</td>
<td>17</td>
<td>1</td>
<td>-0.01</td>
</tr>
<tr>
<td>3</td>
<td>6627 / 7.93</td>
<td>apolipoprotein C-I</td>
<td>121</td>
<td>33</td>
<td>2</td>
<td>-0.02</td>
</tr>
<tr>
<td>4</td>
<td>10845 / 5.23</td>
<td>apolipoprotein C-III</td>
<td>94</td>
<td>16</td>
<td>1</td>
<td>-0.01</td>
</tr>
<tr>
<td>4</td>
<td>79280 / 6.81</td>
<td>transferrin</td>
<td>2455</td>
<td>63</td>
<td>102</td>
<td>-0.07</td>
</tr>
</tbody>
</table>

*a Difference (error) between the experimental and calculated masses

**Table 6.6** Characteristics of the MASCOT identified proteins

Three proteins were identified from 4 MASCOT searches while the MALDI and Q-TOF data failed to find any peptides for the 11785Da biomarker. The 6517Da and 6717Da biomarkers were identified with a high probability as apolipoprotein C-I. 9908Da protein and 79000Da protein were identified as apolipoprotein C-III and transferrin respectively. The peptide sequences obtained encompassing the amino acid sequences corresponding to different regions of the identified proteins are shown in Table 6.7.
Table 6.7 Peptide identification.

### 6.4 Discussion

The clinical course of CLL is extremely variable. The most important initial challenge in patients with CLL is defining prognosis, early recognition and management of complications and designing of risk-stratified management.
Remarkable advances in predicting the course of a newly diagnosed patient with Chronic lymphocytic leukaemia (CLL) have been made recently in the form of $\text{IgV}_H$ mutation status, CD38 and ZAP-70 expression. Difficulties however remain as $\text{IgV}_H$ mutational analysis, the gold standard amongst the novel prognostic markers, is not widely available yet as a routine diagnostic laboratory assay. The significance and best method for determining ZAP-70 has yet to be established.

CLL is not curable with current therapy and initiating chemotherapy following diagnosis in asymptomatic patients does not alter survival (Dighiero, Maloum et al. 1998) therefore the current standard of care is to treat CLL when the patient becomes symptomatic or the disease has progressed (Cheson, Bennett et al. 1996). Newer and more effective treatment protocols with curative potential are being developed. The possibility of intervention at an early stage of disease, when the disease burden is minimal and before clonal evolution occurs, may be more effective at prolonging survival and decreasing disease related mortality. If clinicians are going to adopt this approach, we have to have a robust method of risk analysis so that potentially toxic therapy in patients with a small risk of disease progression can be avoided.

The molecular factors and their interactions arising from the various genetic and cellular abnormalities associated with CLL are still poorly understood. DNA microarray has allowed the analysis of gene expression at the mRNA level; it is proteins that hold the key to the final pathway of cellular interaction and the understanding of the biologic pathways.

Proteomics has played a vital role in many scientific disciplines, enabling discovery of disease biology, drug targets and more recently disease biomarkers. With the recent technological and bioinformatics advances, global protein analysis can provide clues to cellular development, disease markers for diagnosis and prognosis and response to different therapeutic interventions.

The objective of this study was to define potentially robust and valuable prognostic biomarkers for clinical use that correlate with CD38 expression, $\text{IgV}_H$ mutation status, FISH abnormalities and/or treatment response and can further improve our ability for predicting prognosis.
In this work, we analyzed pre-treatment and post treatment proteomic plasma profiles from CLL patients who had received a novel Chemoimmunotherapy regimen, FCR. In addition to the conventional prognostic factors, information for the IgVn mutation status, CD38 expression level and FISH was also available. SELDI-TOF MS was used for the initial biomarker discovery between the different groups. Peptide identification was carried out on the differentially expressed proteins using MALDI-TOF and Q-TOF MS/MS. We have demonstrated that apolipoprotein C-I and transferrin are both underexpressed in CD38+ve patients. We have also shown that the level of apolipoprotein C-I increases with the initiation of therapy. Apolipoprotein C-III on the other hand was overexpressed in the pre-treated patients, with the level going down in response to treatment.

Apolipoprotein CI (apoCI) is a 6.6 kDa apolipoprotein that is synthesized mainly in the liver but also in the skin, adipose tissue, central nervous system (CNS), kidney, and spleen (Lauer, Walker et al. 1988). Originally formed as a propeptide of 9.3 kDa, the mature protein is generated during translational cleavage (Jong, Hofker et al. 1999). It circulates in plasma at a concentration of ~6 mg/dL, and is mainly present on chylomicrons, very-low-density-lipoproteins (VLDL), and high-density-lipoproteins (HDL) (Jong, Hofker et al. 1999). Its major function is lipoprotein lipase (LPL) inhibition, resulting in the generation of less free fatty acid (FFA) for subsequent uptake by adipose tissue and skin (Berbee, van der Hoogt et al. 2005). It also inhibits binding of VLDL and LDL to their receptors (Windler, Kovanen et al. 1980; Kowal, Herz et al. 1990).

Apolipoprotein C-III (apoC-III) is a protein secreted mostly by the liver and, to a lesser extent, by the intestine (van Dijk, Rensen et al. 2004). In circulation, it is associated with both triglyceride-rich lipoproteins (TRLs) and HDL (Jong, Hofker et al. 1999). ApoC-III is present in three isoforms depending on the number of sialic acid molecules (0 to 2) (Ito, Breslow et al. 1989). Total plasma apoC-III levels have been identified as a major determinant of triglyceridemia (Gerritsen, Rensen et al. 2005). Proposed mechanisms underlying the hypertriglyceridemic effect of apoC-III comprise inhibition of lipoprotein lipase (LPL) activity, disruption of the interaction of TRLs with vessel wall heparan sulfate proteoglycans, and lower clearance of apoB-
containing lipoproteins by LDL and LDL-related receptors (Aalto-Setala, Fisher et al. 1992; Shachter 2001).

Transferrin (Tf), the major iron-transporting protein in plasma (Aisen and Brown 1975), is a glycoprotein with a MW of 79500 kDa (MacGillivray, Mendez et al. 1982). It transports iron from the intestine, reticuloendothelial system and liver parenchymal cells to all proliferating cells in the body, where cellular uptake of iron is mediated by cell surface transferrin receptor (Tfr) (Trowbridge and Shackelford 1986).

The three proteins that we have identified from the initial 5 peaks after SELDI analysis are all classical serum proteins; however this does not exclude their use as cancer biomarkers (Poon, Chan et al. 2001). Decreased levels of ApoC-I have previously been observed in patients with liver cirrhosis and hepatocellular carcinoma secondary to hepatitis C infection using SELDI-TOF MS (Gobel, Vorderwulbecke et al. 2006). Similarly decreased levels were also observed in patients with colorectal carcinoma, although the difference in expression was not confirmed by western blotting (Ward, Suggett et al. 2006). There are so far no reports in the literature demonstrating the up regulation of apolipoprotein C-III and cancer. Huan et al (Huang, Stasyk et al. 2006) however have demonstrated down regulation of Apo C-III in patients with breast carcinoma. A number of studies on the other hand have reported altered Tf regulation in cancer. Using MALDI-TOF and western blotting Nuzhat et al (Ahmed, Oliva et al. 2005) have shown down regulation of Tf in patients with advanced ovarian carcinoma as compared to normal controls. A rise in the Tf level corresponded to a response to chemotherapy. A lower expression of Tf has also been shown in gastric carcinoma (Ryu, Kim et al. 2003) Other studies involving patients with breast cancer and colorectal carcinoma have demonstrated an up regulation of Tf (Huang, Stasyk et al. 2006; Ward, Suggett et al. 2006).

The exact role that these proteins play in the biology of CLL is not clear at present. There is so far no report in the literature demonstrating the biological mechanism of apolipoprotein C-I and apolipoprotein C-III to cancer. Apolipoproteins are lipid carriers and regulate many cellular functions and have antioxidative properties. For example, the protein kinase Akt can be elicited by apolipoproteins, which in turn
promotes growth factor-mediated cell survival and block apoptosis (Pikarsky, Porat et al. 2004) Apolipoprotein C-III has also been associated with increase in expression of cellular adhesion molecules (VCAM-1 and ICAM-1) through the activation of PKCβ and NF-κB by endothelial cells (Kawakami, Aikawa et al. 2006). The clinical significance of the reduction of transferrin levels in the CLL plasma is not clear although the association seems to be strong in solid tumours. Serum ferritin and Tf concentrations are useful parameters in distinguishing the level of iron stores in patients (Kwok and Richardson 2002). A number of clinical and experimental studies suggest an increase in cancer risk with elevated body iron stores, specifically, cancer of the liver, lung, GI tract and pancreas (Wurzelmann, Silver et al. 1996; Hussain, Raja et al. 2000). Iron is known to generate hydroxyl radicals from the less reactive superoxide anion and hydroxyl peroxide (Halliwell and Gutteridge 1986). The highly reactive free radicals induce DNA mutations, and may increase cancer risk. Tf antibodies have been used in cell lines from acute myeloid leukemia and other malignancies to arrest growth (Taetle, Rhyner et al. 1985). On the other hand, an inverse correlation between iron stores and higher cancer risk has been reported for gastric cancer (Nomura, Chyou et al. 1992). Negative correlation between serum ferritin and transferrin levels and the presence of prostate cancer has recently been reported (Kuvibidila, Gauthier et al. 2004). Reduced iron absorption due to inflammation and/or increased iron utilization by prostate cancer cells is the suggested mechanism of reduced ferritin and transferrin levels in the serum of prostate cancer patients. Transferrin synthesis by T lymphocytes (Lum, Infante et al. 1986), ovarian cancer cell lines (Ohkawa, Takada et al. 1990) and small cell lung cancer have been demonstrated (Vostrejs, Moran et al. 1988). Transferrin has been detected in the ascites of ovarian cancer patients (Ahmed, Oliva et al. 2005). The synthesis of transferrin by ovarian cancer cells and small cell lung carcinoma may function as an autocrine growth factor (Vostrejs, Moran et al. 1988). Low levels of Tf on the other hand may represent increased requirement for transferrin by the rapidly growing tumor (Obermair, Handisurya et al. 1998). This explanation is plausible as patients with ovarian cancer treated with chemotherapy showed an increase in the Tf levels after chemotherapy (Ahmed, Oliva et al. 2005).

This study and the growing literature on the use of SELDI-TOF indicate its suitability for the identification of tumour biomarkers. Accurate identification of the biomarker...
would need further validation by western blotting or peptide sequencing using MS/MS technologies, for subsequent measurement in routine laboratories. Larger studies will however be required to exactly define the role of biomarker patterns, for identifying single protein markers that could be translated into a clinical test and for the validation of the different proteomic technologies.
CHAPTER 7

CONCLUSION
Remarkable progress has been made in the last decade, not only in understanding and unravelling the complex biology of chronic lymphocytic leukaemia but also in better risk stratification, and the development of newer therapeutic options and outcomes.

CLL has classically been considered to arise from an immunologically incompetent B lymphocyte that acquires a defect in apoptosis, resulting in an accumulative disorder (Dameshek 1967). This view is, however, changing. More recent data suggests that CLL cells derive from immuno-competent B lymphocytes that have been transformed secondary to multiple encounters and responses to autoantigens. The inability of CLL cells to remain viable in vitro (Zupo, Isnardi et al. 1996) unless cultured in the presence of other cell types or soluble cytokines (Burger, Tsukada et al. 2000), further substantiates that CLL cells require stimulations delivered through cell surface receptors to maintain viability.

Historically, CLL was considered to have arisen from a leukemic transformation of naïve B lymphocyte that had not undergone germinal centre antigen exposure and subsequent somatic mutations of the immunoglobulin genes. Several studies including our own, have found that approximately 50% of CLL cases exhibit somatic mutation of the immunoglobulin chains, which suggests that some CLL clones arise from post-germinal centre memory B cells.

Heterogeneity in clinical behaviour has also been recognised from the time that CLL was first recognised as a separate entity. Recent advances in the understanding of CLL biology, such as the relevance of IgV_H mutation status, CD38 and ZAP70 expression, and chromosomal abnormalities have allowed us some insight into the different clinical outcomes of patients diagnosed with CLL.

Patients with leukaemic clones containing minimal or no mutations and an elevated CD38 and ZAP70 expression are at much higher risk for an aggressive disease than those cases where CLL cells have significant mutations and do not express CD38 and ZAP70.
The association of 17p and 11q deletion with short survival times and deletion 13q with favourable outcome is also well recognised. Furthermore, 17p deletions are associated with resistance to purine analogues and alkylating agents.

For many years, treatment options in CLL were not only relatively limited but were incapable of inducing high levels of overall responses or complete responses. The introduction of purine analogues such as fludarabine and the more recent use of monoclonal antibodies in combination with chemotherapy have however attained 100% ORR and very high CR of up to 70%, as shown in Figure 7.1. Accumulating evidence of a powerful GVL effect in CLL (Sorror, Maris et al. 2005) and the reduced mortality associated with reduced intensity conditioning transplants (Dreger, Brand et al. 2005), have further added to the therapeutic armamentarium against CLL.

![Figure 7.1 CLL treatment options by decade. A timeline over the last several decades of increasing CR in relation to treatment protocols in untreated B-CLL. Kay, N. E. Blood 2006;107:848](image-url)
Conventionally treatment responses have been judged using the NCI-WG response criteria which have proven very useful over the years. Recent therapeutic advances frequently reduce the disease burden to levels detectable only using very sensitive techniques such as four colour flow cytometry and or molecular methods (Keating, O'Brien et al. 2005). This is more than of academic interest as eradication of MRD might translate into longer survival and even cure for some patients (Bosch, Ferrer et al. 2002; Esteve, Villamor et al. 2002; Moreton, Kennedy et al. 2005; Moreno, Villamor et al. 2006).

As our understanding of lymphoma biology has taught us, it is also becoming clear in CLL that in the future treatment strategies should be individualized for specific patients. Recent advances in CLL biology, the availability of novel prognostic tools, the ability to detect MRD and its implications and more effective treatment options should enable us to develop more accurate algorithms for making such treatment decisions.

Although consensus guidelines for diagnosing and treatment are available for use, both from BCSH and NCI, they do not identify early stage patients with biologically aggressive disease. For the reasons mentioned above it seems highly desirable that some of these novel prognostic markers are included in the decision making process especially in younger patients. Based on our own experience and that available from other published series we would recommend the inclusion of CD38 expression analysis (Pattern and level), CLL relevant FISH and where available IgV_{H} mutation status in the prognostic evaluation of patients with CLL. Depending on risk stratification and biological age of the patient, they should be offered participation in clinical trials based on the recent and emerging therapeutic advances. An algorithm incorporating the novel prognostic markers and the recent therapeutic advances is shown in Figure 7.2.
Figure 7.2 An algorithm for the management of CLL patients

The importance of continued quest for newer and better prognostic markers can not be stressed more. Techniques such as proteomics are as yet in their infancy but extremely powerful. Its potential to identify biomarkers and potential therapeutic targets is enormous. The aim is to refine risk stratification in so that we have a better knowledge of who to treat, when to treat and how to treat.
Appendix

Appendix A: Expression Difference Mapping Kit-Serum Fractionation Buffers

- U1 Buffer: 1M urea, 0.2% CHAPS, 50mM Tris-HCl, pH 9
- U9 Buffer: 9M urea, 2% CHAPS, 50mM Tris-HCl, pH 9
- Rehydration buffer: 50 mM Tris-HCl, pH 9
- Wash buffer 1: 50 mM Tris-HCl with 0.1% Octyl β-D-glucopyranoside (OGP), pH 9
- Wash buffer 2: 50 mM Hepes with 0.1% OGP, pH 7
- Wash buffer 3: 100 mM NaAcetate with 0.1% OGP, pH 5
- Wash buffer 4: 50 mM Tris-HCl with 0.1% OGP, pH 4
- Wash buffer 5: 50 mM NaCitrate with 0.1% OGP, pH 3
- Wash buffer 6: 33.3% isopropanol/ 16.7% acetonitrile/ 0.1% trifluoracetic acid
References


Abeloff, J. O. Armitage, A. S. Lichter and J. E. Niederhuber. New York, USA,
Churchill Livingstone: 2536-2563.

Working Group guidelines for chronic lymphocytic leukemia: revised

chronic lymphocytic leukemia: recommendations of the National Cancer


Clark, E. A. and P. J. Lane (1991). "Regulation of human B-cell activation and

Clavio, M., M. Miglino, et al. (1998). "First line Fludarabine treatment of
symptomatic chronic lymphoproliferative diseases: clinical results and
molecular analysis of minimal residual disease." Eur J Haematol 61(3): 197-
203.

lymphocytic leukemia subtypes with mutated or unmutated Ig V(H) genes."

71-5.

of brcal in the DNA damage response to double-strand breaks." Science
286(5442): 1162-6.

Crespo, M., F. Bosch, et al. (2003). "ZAP-70 expression as a surrogate for
immunoglobulin-variable-region mutations in chronic lymphocytic leukemia."

defined typical and atypical CLL: a clinical, immunophenotypic, cytogenetic


Fuller, H. (1846). "Particulars of a case in which enormous enlargement of the spleen and liver, together with dilation of all the blood vessels of the body, were found coincident with a peculiarly altered condition of the blood." Lancet 2: 43.


Keating, M. J., T. Manshouri, et al. (2002). A high proportion of molecular remission can be obtained with a Fludarabine, Cyclophosphamide, Rituximab


Rawstron, A. C., M. J. Green, et al. (2002). "Monoclonal B lymphocytes with the characteristics of "indolent" chronic lymphocytic leukemia are present in 3.5% of adults with normal blood counts." Blood 100(2): 635-9.
Rawstron, A. C., B. Kennedy, et al. (2001). "Quantitation of minimal disease levels in chronic lymphocytic leukemia using a sensitive flow cytometric assay improves the prediction of outcome and can be used to optimize therapy." Blood 98(1): 29-35.


CD38 expression level and pattern of expression remains a reliable and robust marker of progressive disease in chronic lymphocytic leukemia

AMJAD HAYAT, DAVID O'BRIEN, PAUL O'ROURKE, SIOBHAN McGUCKIN, TONY FITZGERALD, EIBHLIN CONNEALLY, PAUL V. BROWNE, SHAUN R. McCANN, MARK P. LAWLER, & ELISABETH VANDENBERGHE

Department of Haematology and Durkan Laboratories, Institute of Molecular Medicine, Trinity College and St James Hospital, Dublin, Ireland

(Received 20 April 2006; revised 6 July 2006; accepted 26 July 2006)

Abstract
Chronic lymphocytic leukemia (CLL) follows a variable clinical course which is difficult to predict at diagnosis. We assessed somatic mutation (SHM) status, CD38 and ZAP-70 expression in 87 patients (49 male, 38 female) with stage A CLL and known cytogenetic profile to compare their role in predicting disease progression, which was assessed by the treatment free interval (TFI) from diagnosis. Sixty (69%) patients were SHM+, 24 (28%) were CD38+ and ten (12%) were ZAP-70+. The median TFI for: (i) SHM+ versus SHM− patients was 124 versus 26 months; hazard ratio (HR) = 3.6 [95% confidence interval (CI) = 1.8–7.3; P = 0.001]; (ii) CD38− versus CD38+ patients was 120 versus 34 months; HR = 2.4 (95% CI = 1.4–5.3; P = 0.02); and (iii) ZAP70− versus ZAP70+ was 120 versus 16 months; HR = 3.4 (95% CI = 1.4–8.7; P = 0.01). SHM status and CD38 retained prognostic significance on multivariate analysis whereas ZAP-70 did not. We conclude that ZAP-70 analysis does not provide additional prognostic information in this group of patients.

Keywords: CLL, IgVH, CD38, ZAP-70, prognostic factors

Introduction
Chronic lymphocytic leukemia (CLL) is the commonest leukemia in the Western world, and has a clinically variable course with a median survival of 20 years [1,2]. Traditionally Rai [3] or Binet [4] staging is used for prognostic assessment and ‘watchful waiting’ remains the standard care for asymptomatic patients. Patients diagnosed early in their disease course may benefit from detailed prognostic information and a tailored treatment approach. Treatment with agents such as fludarabine, rituximab and alemtuzumab and the increasing use of stem cell transplantation can induce molecular remissions and may cure some patients [5]. It is therefore clinically useful to identify patients with a poor prognosis early in their treatment pathway and adopt a risk-stratified management plan.

Clonal cytogenetic abnormalities are found in 80% of CLL patients using fluorescence in situ hybridization (FISH) [6]. The association of 17p deletion (del 17p) and 11q deletion (del 11q) with short survival times and deletion 13q (del 13q) with favourable outcome is well recognized [6,7]. In addition, del 17p abnormalities have been associated with resistance and treatment failure with purine analogues [6,8]. Several markers including immunoglobulin heavy chain (IgVH) mutational status, CD38 and ZAP-70 expression and cytogenetic abnormalities such as deletion 11q23 and deletion of 17p [7,9–11] are associated with a poor prognosis in CLL. The IgVH mutational status is the best predictor of clinical outcome in multivariate analysis. Patients with unmutated IgVH genes (Ig-unmutated) have a median survival of 79–119 months [7,9,12,13]
compared to 293 months in patients with mutated \( \text{IgVH} \) genes (Ig-mutated). Mutational analysis, however, is expensive and time consuming and not suited to routine diagnostic laboratories.

Many studies have confirmed the clinical usefulness of CD38 expression analysis in relation to disease progression [14–18], although two studies in multivariate analysis failed to do so [7,9]. Limited agreement exists about significant expression levels and expression patterns [15,17,19]. The flow cytometry analysis used in CD38 expression is robust and can easily be incorporated into a standard CLL immunophenotype profile.

Gene expression analysis of CLL cells identifies different ZAP-70 expression patterns in Ig-mutated and Ig-unmutated cases [20,21]. ZAP-70 is an intracellular tyrosine kinase associated with T-cell receptor signalling [22,23]. Initial studies indicated that ZAP-70 was a reliable surrogate marker for \( \text{IgVH} \) mutational status [21,24], but more recent studies are less convincing [25]. At present, ZAP-70 analysis is time consuming and not robust enough to be performed in a routine immunophenotyping laboratory, and whether it adds to the information provided by other prognostic factors remains to be established.

To evaluate the relationship of \( \text{IgVH} \) mutational status to ZAP-70 and CD38 expression and, secondly, to optimize the use of CD38 as a prognostic marker in CLL in stage A patients and decide whether ZAP 70 analysis provided additional useful information in the routine diagnosis of CLL, we sequenced the \( \text{IgVH} \) genes and measured expression of ZAP-70 and CD38, including patterns of CD38 expression using flow cytometry and determined cytogentic abnormalities using FISH in a consecutive cohort of 87 untreated stage A patients. This data has been correlated with disease status at diagnosis and the time from diagnosis to first treatment (TFI).

Patients and methods

Patients

One hundred and six consecutive B-CLL patients referred to St James Hospital, Dublin were studied, of whom 87 had stage A disease and fulfilled the morphological and immunophenotypic criteria of CLL (CD5/19+, CD23-, FMC7-, weak surface immunoglobulin and light chain restriction). Written informed consent was obtained at time of enrolment. The following parameters were measured at diagnosis and follow-up: B symptoms, lymph node, liver and spleen size, full blood count and differential, serum \( \beta_2 \)-microglobulin and lactate dehydrogenase. Follow-up was based on clinical examination unless the tumor sites were not amenable to clinical examination (e.g. retroperitoneal adenopathy). Disease status and evidence of progression was evaluated by five senior doctors (A.H., E.V., S.Me., P.B., B.C.) and decision to treat based on agreed criteria of progressive stage A, marrow failure, < 6 months lymphocyte doubling time, B symptoms or progressive hepatosplenomegaly or lymphadenopathy. Treatment free interval was used as a surrogate marker of disease progression and was measured in months from diagnosis. Response to treatment was measured using National Cancer Institute working group criteria. Forty milliliters of blood was collected at the time of study accrual. Peripheral blood mononuclear cells were separated by density centrifugation (Lymphoprep, Fresenius Kabi, Norge AS, Norway) and aliquoted for DNA and RNA extraction and cryopreservation.

FISH was carried out using a set of directly labelled commercial probes from Vysis (Abbott Diagnostics, South Pasadena, CA, USA) for 13q14 (LSI D13S319), centromere 12 (CEP 12), 11q22-23 (LSI ATM) and 17p13 (p53 locus) according to the manufacturers protocols.

\( \text{VH} \) gene analysis

RNA was isolated using Rneasy Midi/Mini kit (Qiagen GmbH, Hilden, Germany); reverse transcribed using an oligo dT primer (Reverse Transcription System, Promega, Madison, WI, USA). Genomic DNA was extracted using QIAamp® Blood Mini kit (Qiagen GmbH) if RNA was unavailable or cDNA did not amplify. Immunoglobulin variable region genes were sequenced as follows [13]; cDNA was amplified by polymerase chain reaction (PCR) using a mixture of 5' oligonucleotide primers specific for each leader sequence of the VH gene to the V\( \gamma \)6 families, with a 3' constant region primer or a 5' framework 1 consensus primer and a 3' consensus J\( \gamma \) primers. PCR products were size selected following electrophoresis on 2% agarose gel, purified using QIAquick PCR purification kit (Qiagen GmbH) and direct sequencing was performed using amplicons from at least two independent amplification reactions. Samples were sequenced using an ABI 3100 sequencer and BigDye Terminator v3.1 Cycle sequencing kit (Applied Biosystems, Warrington, UK). Cloning of PCR products was performed using TOPO TA One Shot cloning kit (Invitrogen, Paisley, UK) followed by sequencing of at least ten positive transformed bacterial colonies if direct sequencing failed. Nucleotide sequences were aligned to EMBL/GenBank and V-BASE sequence directory [26] using Lasergene (DNASTAR, Madison, WI, USA) sequence analysis software. Percent homology to the closest germline \( \text{VH} \) gene was calculated from the number of nucleotide differences between the 5' end.
of framework 1 and the 3' end of framework 3. Sequences with less than 98% homology to the germline gene were considered mutated.

Analysis of CD38 expression

Fresh peripheral blood samples were diluted in CellWash (Becton Dickinson, San Jose, CA, USA) to $2 \times 10^8$ cells/ml. Diluted blood (100 µl) was incubated with the following antibodies: fluorescein isothiocyanate (FITC)-labelled anti-CD5 (clone L17F12), allophycocyanin- (APC)-conjugated anti-CD19 (clone SJ25C1), phycoerythrin (PE)-labelled anti-CD38 (clone HB7) and peridinin chlorophyll protein (PerCP)-labelled anti-CD45 (clone 2D1) (Becton Dickinson). At least 10,000 events were acquired using the Cellquest software and a FACS-Calibur (Becton Dickinson). Dot plots were generated on CD45 gated lymphocytes to determine CD5/CD19 and CD38 positivity. Dot-plots and histograms were used to determine expression pattern.

CD38 expression was analysed as: (i) an independent prognostic marker, (ii) a surrogate marker for IgVH mutational status using different cut-off values of 20%, 25%, 30% and 35%; and (iii) to relate expression pattern to prognosis [19].

ZAP-70 expression

All samples analysed were fresh; however, a concordance study was also performed on 40 samples using fresh and cryopreserved samples from the same patient. The percentage of CD19+/CD5+ cells, expressing ZAP-70 at 10%, 15%, 20% and 25% cut-off level was quantified [24]. Fresh blood samples were diluted to $2 \times 10^8$ cells/ml. Diluted blood (100 µl) was fixed and permeabilized with the Fix & Perm kit (Caltag Laboratories, Towcester, UK). Frozen samples were thawed at 37°C, washed twice in 2 ml of warm PBSAT (phosphate-buffered saline containing albumin and Tween 20) and resuspended to the original volume in PBSAT before permeabilization.

Anti-ZAP-70 antibody 1.5 µg (Clone 2F3.2; Upstate, Charlottesville, VA, USA) was added and the sample incubated for 15 min at room temperature (RT), washed twice with rabbit anti-mouse immunoglobulin-RAM-PE (Dako, Cambridge, UK), washed, incubated at RT with 100 µl neat mouse serum for 5 min and washed. Anti-CD3 FITC, CD56 FITC, CD19 APC and CD5 PerCP-Cy5.5 (Becton Dickinson) were added, incubated for 15 min at RT, washed twice and resuspended in 0.5 ml of CellWash. A live gate was used to ensure that at least 2000 T cells and natural killer (NK) cells were analysed in each sample. The ZAP-70 staining on the population of CD3+/CD56+ cells was used as an internal positive control and compared to ZAP-70 staining on the CD5+/CD19+ population. The percentage ZAP-70 positivity was determined using a quadrant marker on the left margin of the T/NK cell population.

Statistical analysis

Continuous variables were presented as median (upper and lower quartile) and levels in different groups were compared using a Mann–Whitney rank sum test. Categorical variables were summarized as percentage (n) and frequencies compared using a standard chi-square test. The association between disease progression and each of ZAP-70, CD-38, and mutation status was analysed. Kaplan–Meier 'survival' curves described the cumulative risk of treatment in groups defined by their IgVH mutational status, levels of ZAP-70 using a 20% cut-off value, CD38 expression level using a 30% cut-off value and the pattern of CD38 expression. Kaplan–Meier curves based on TF1 were compared using a log-rank test. The independent effect of the three prognostic factors was considered using a Cox proportional hazards model which included the three variables. The sensitivity, specificity and percentage discordance was used to assess discrimination of various cut-off values for CD38 and ZAP-70 expression. Diagnostic reliability was based on the positive and negative predictive value at various levels of CD38 and ZAP-70 expression. All statistical analysis was carried out using Stata (version 8.2) statistical analysis software (Stata Corporation, College Station, TX, USA).

Results

Eighty-seven patients, 49 males and 38 females with a median (range) age of 62 (37–90) years were included in the analysis. Median (range) follow-up was 34 (2–279) months. Thirty-three of 87 patients required treatment for symptomatic disease during the study. IgVH mutational status was available on all patients, CD38 expression on 86/87 patients, ZAP-70 expression on 81/87 patients and FISH on 82/87 patients. Patient characteristics at diagnosis are summarized in (Table I). FISH results are summarized (Table II) which also shows TF1 for the different groups.

IgVH gene mutational status

IgVH mutational analysis was performed on all 87 patients: Ig-mutated genes were detected in 61/87 (70%) patients while 26/87 (30%) patients had
Table I. Main biologic and clinical characteristics of the patients (stage A only)*.

<table>
<thead>
<tr>
<th>No of patients</th>
<th>87</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, male/female</td>
<td>49/38</td>
</tr>
<tr>
<td>Median age (years)</td>
<td>62</td>
</tr>
<tr>
<td>Median WCC x 10^9/l</td>
<td>27.3 (7.3 - 304)</td>
</tr>
<tr>
<td>Median Hb (g/dl)</td>
<td>14 (6.3 - 17.2)</td>
</tr>
<tr>
<td>Median platelet count x 10^9/l</td>
<td>204 (57 - 397)</td>
</tr>
<tr>
<td>LDH lU/l</td>
<td>330 (135 - 1106)</td>
</tr>
<tr>
<td>β2-m mg/dl</td>
<td>2.4 (1.3 - 10)</td>
</tr>
<tr>
<td>IgV H mutated/unmutated</td>
<td>61/26</td>
</tr>
<tr>
<td>CD38: negative/positive</td>
<td>62/24 (86 patients)</td>
</tr>
<tr>
<td>ZAP-70: negative/positive</td>
<td>71/10 (81 patients)</td>
</tr>
<tr>
<td>Median follow-up (months)</td>
<td>34 (2 - 279)</td>
</tr>
<tr>
<td>Treated/not treated</td>
<td>33/54</td>
</tr>
</tbody>
</table>

*Binet stage at diagnosis. ^LDH count was obtained in 81 patients. 
β2-microglobulin count was obtained in 60 patients.

Table II. Incidence of chromosomal abnormalities in 82 patients with stage A chronic lymphocytic leukemia.

<table>
<thead>
<tr>
<th>Aberration</th>
<th>No. of patients (%)</th>
<th>Patients ( % ) treated</th>
<th>TFI (hazard ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal karyotype</td>
<td>19 (23%)</td>
<td>14/19 (21%)</td>
<td>87 (1)</td>
</tr>
<tr>
<td>13q deletion</td>
<td>14 (17%)</td>
<td>17/14 (12%)</td>
<td>110 (0.54)</td>
</tr>
<tr>
<td>Trisomy 12</td>
<td>8 (10%)</td>
<td>2/8 (25%)</td>
<td>41 (1.51)</td>
</tr>
<tr>
<td>11q deletion</td>
<td>6 (7%)</td>
<td>5/6 (83%)</td>
<td>11 (2.01)</td>
</tr>
<tr>
<td>17p deletion</td>
<td>3 (4%)</td>
<td>2/3 (67%)</td>
<td>5 (2.41)</td>
</tr>
<tr>
<td>Clonal abnormalities</td>
<td>63 (77%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table III. Comparison of treatment requirements based on IgV H status, CD38% and ZAP-70% in stage A CLL.

<table>
<thead>
<tr>
<th>Ig-mutated</th>
<th>44/61 (72%)</th>
<th>17/61 (28%)</th>
<th>124</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ig-unmutated</td>
<td>10/26 (38%)</td>
<td>16/26 (62%)</td>
<td>26</td>
</tr>
<tr>
<td>CD38 &lt; 30%</td>
<td>41/62 (66%)</td>
<td>24/62 (34%)</td>
<td>120</td>
</tr>
<tr>
<td>CD38 &gt; 30%</td>
<td>13/24 (54%)</td>
<td>11/24 (46%)</td>
<td>34</td>
</tr>
<tr>
<td>ZAP-70 &lt; 20%</td>
<td>47/71 (66%)</td>
<td>24/71 (34%)</td>
<td>120</td>
</tr>
<tr>
<td>ZAP-70 &gt; 20%</td>
<td>4/10 (40%)</td>
<td>6/10 (60%)</td>
<td>16</td>
</tr>
</tbody>
</table>

CD38 expression

CD38 expression data was available on 86 patients. The range of CD38 expressing cells was 0.1% to 98%. Twenty-four of 86 (28%) samples were CD38+ and 62/86 (72%) CD38−. A CD38 expression level of 30% most accurately mirrored IgV H mutational status with a concordance of 71%, positive predictive value of 60% and a negative predictive value of 77% [15,18] (Table V). The results were discordant in 29% cases which did not improve with different expression levels. Results for TFI are summarized in Table III and Figure 2. Univariate analysis of the TFI revealed a HR = 2.4; CI = 1.4 - 5.3; P = 0.02 (Table IV).

CD38+ patients were divided into three groups by CD38 expression pattern [19]: 54/86 were homogeneously negative, 17/86 homogeneously positive and 15/86 exhibited bimodal expression. Forty-three of 60 (71%) patients with Ig-mutated genes were CD38−, 10/60 (17%) CD38−+ and 10/60 (12%) had bimodal CD38 expression. Eleven of 26 (42%) patients with Ig-unmutated genes were CD38−, 7/26 (27%) CD38+, and 8/26 (31%) had bimodal expression. Seventeen (31%) CD38− patients were treated with a TFI of 124 months whereas nine (53%) of the CD38−+ group were treated with a TFI of 6 months. Six (40%) patients in the bimodal group required treatment with a TFI of 102 months (Figure 3).

ZAP-70 expression

Results were available on 81 patients: 71 (87%) were ZAP-70− compared to ten (13%) ZAP-70+ patients. Fifty-five patients with Ig-mutated genes were ZAP-70−, 8/24 (33%) patients with Ig-unmutated genes were ZAP-70+. The assay gave discordant results in 18/81 (22%) patients, with 16 (67%) high-risk patients by mutational analysis (Ig-unmutated) misclassified as low risk by ZAP-70 expression. This discordance did not improve with different expression cut-off levels (Table VI).

Twenty-four of 71 (34%) ZAP-70− patients required treatment compared to 6/10 (60%) ZAP-70+ patients. Results for TFI are summarized in Table III and Figure 4. Univariate analysis of the TFI revealed HR = 3.4; CI = 1.4 - 8.7; P = 0.01, for ZAP-70+ patients (Table IV).
IgV<sub>H</sub>, CD38 and ZAP-70 in CLL

Figure 1. Kaplan-Meier survival curves comparing TFI in stage A CLL patients with mutated and unmutated IgV<sub>H</sub> genes: median TFI for Ig-mutated CLL = months, median TFI for Ig-mutated CLL = 124 months. The difference is significant at P = 0.001.

Table IV. Univariate analysis of treatment free interval.

<table>
<thead>
<tr>
<th>Stage A patients</th>
<th>Prognostic factor</th>
<th>HR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ig-unmutated</td>
<td>3.6</td>
<td>1.8-7.3</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>CD 38 &gt; 30%</td>
<td>2.4</td>
<td>1.1-5.3</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>ZAP-70 &gt; 20%</td>
<td>3.4</td>
<td>1.4-8.7</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Discussion

We found cytogenetic analysis by FISH, IgV<sub>H</sub> mutational status and CD38 expression to be reliable predictors for progressive disease requiring treatment in this series of 87 stage A CLL patients. The commonest clonal genetic abnormalities detected by FISH analysis are deletions of 13q, 11q, 17p and trisomy 12 [29]. The incidence of cytogenetic abnormalities in the present study is similar to published data, with the slightly low incidence of trisomy 12 and deletions of 11q and 17p probably due to the inclusion of stage A patients only. The 13q deletions were over-represented in Ig-mutated patients and deletions 17p and 11q in Ig-unmutated patients [7]. The prognostic significance of cytogenetic abnormalities was supported by the difference in TFI (Table II) in our series, indicating that FISH analysis remains one of the most useful prognostic tests currently available.

CD38 belongs to a family of proteins with ectoenzyme and receptor functions. CD38 has been shown to acquire an abnormal signalling function depending on its relationship to the B-cell receptor which may provide the biological answer for its prognostic role in CLL [30]. Thus, even low level CD38 expression could result in abnormal signalling causing more active disease. We confirm that CD38 is clinically useful, with CD38 expression of > 30% predicting disease progression with a TFI of 34 months compared to 120 months for CD38—patients [14-17]. The prognostic accuracy of

Correlation between FISH, IgV<sub>H</sub>, CD38 and ZAP-70

Patients were divided into five groups depending on cytogenetic findings, which were then compared with regards to mutational status, CD38 and ZAP-70 expression and the degree of concordance or discordance between compared parameters was measured (Table VII). The majority of patients with del(13q) had mutated IgV<sub>H</sub> genes (83%) and were CD38— (74%) and ZAP-70— (91%). Most patients with no FISH abnormality had mutated genes, (68%) were CD38— (68%) and none expressed ZAP-70. Most patients with trisomy 12 had mutated IgV<sub>H</sub> genes (62%), five (62%) patients expressed CD38 and 87% were ZAP-70—. There was a significant association between del(11q) or del(17p) and unmutated genes (78%) and CD38 expression (78%) but (62%) of these patients were ZAP-70—. Most patients had concordant expression of CD38 and ZAP-70 (74%). Nineteen (23%) patients were CD38+ and ZAP-70— whereas two patients were ZAP-70+ and CD38—.
Table V. Sensitivity and specificity of CD38 for predicting IgV\textsubscript{H} mutational status.

<table>
<thead>
<tr>
<th>Cut-off value for CD38</th>
<th>Discordance percentage</th>
<th>Sensitivity percentage</th>
<th>Specificity percentage</th>
<th>Positive predictive value percentage</th>
<th>Negative predictive value percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>32.38</td>
<td>67.57</td>
<td>67.65</td>
<td>53.19</td>
<td>79.31</td>
</tr>
<tr>
<td>25%</td>
<td>31.42</td>
<td>62.16</td>
<td>72.06</td>
<td>54.76</td>
<td>77.78</td>
</tr>
<tr>
<td>30%</td>
<td>28.57</td>
<td>56.76</td>
<td>79.41</td>
<td>60</td>
<td>77.14</td>
</tr>
<tr>
<td>35%</td>
<td>28.58</td>
<td>51.35</td>
<td>82.35</td>
<td>61.29</td>
<td>75.68</td>
</tr>
</tbody>
</table>

Figure 2. Kaplan–Meier survival curves comparing TFI in CD38− and CD38+ stage A CLL patients: median TFI for CD38+ CLL = 24 months, median TFI for CD38− CLL = 120 months ($P = 0.02$).

Figure 3. Kaplan–Meier survival curves comparing TFI in stage A CLL patients according to the pattern of CD38 expression: median TFI for unimodal+ CLL = 4 months, for unimodal− CLL = 124 months and for patients with bimodal expression = 102 months.

CD38 is improved by analysing expression pattern with a median TFI for the unimodal positive versus negative group of 6 versus 124 months compared to 34 versus 120 months using expression level alone [19]. The bimodal group had a median survival closer to the CD38− group (102 months), although
Table VI. Sensitivity and specificity of ZAP-70 for predicting IgVH mutational status.

<table>
<thead>
<tr>
<th>Cut-off value for ZAP-70</th>
<th>Discordance percentage</th>
<th>Sensitivity percentage</th>
<th>Specificity percentage</th>
<th>Positive predictive value percentage</th>
<th>Negative predictive value percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>24.5</td>
<td>51.61</td>
<td>89.06</td>
<td>69.57</td>
<td>79.17</td>
</tr>
<tr>
<td>15%</td>
<td>24.1</td>
<td>45.16</td>
<td>92.19</td>
<td>73.68</td>
<td>77.63</td>
</tr>
<tr>
<td>20%</td>
<td>23.15</td>
<td>38.71</td>
<td>95.31</td>
<td>80</td>
<td>76.25</td>
</tr>
<tr>
<td>25%</td>
<td>27.36</td>
<td>22.58</td>
<td>96.88</td>
<td>77.78</td>
<td>72.09</td>
</tr>
</tbody>
</table>

Figure 4. Kaplan-Meier survival curves comparing TFI in ZAP-70— and ZAP-70+ stage A CLL patients: median TFI for ZAP-70+ CLL = 5 months, median TFI for ZAP-70— CLL = 102 months. The difference is significant at P = 0.002.

Table VII. Correlation between FISH, IgVH, CD38 and ZAP-70.

<table>
<thead>
<tr>
<th>FISH</th>
<th>Positive (%)</th>
<th>IgVH</th>
<th>CD38 expression</th>
<th>ZAP70 Expression</th>
<th>Concordant</th>
<th>Discordant</th>
</tr>
</thead>
<tbody>
<tr>
<td>del 13q</td>
<td>46 (56)</td>
<td>Mutated</td>
<td>CD38+</td>
<td>ZAP70+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>19 (23)</td>
<td>Unmutated</td>
<td>CD38−</td>
<td>ZAP70−</td>
<td></td>
<td></td>
</tr>
<tr>
<td>del 11q</td>
<td>8 (10)</td>
<td>del 17p</td>
<td>CD38−</td>
<td>ZAP70+/</td>
<td>ZAP70−/</td>
<td></td>
</tr>
<tr>
<td>del 17p</td>
<td>3 (4)</td>
<td></td>
<td>CD38−</td>
<td>ZAP70+/</td>
<td>ZAP70−−/</td>
<td></td>
</tr>
</tbody>
</table>

50% of the patients did need treatment. A correlation between CD38 expression and complex cytogenetic abnormalities (examined by comparative genomic hybridization) has previously been found by Ottaggio et al. [31]. We found a significant association between deletion 13q and CD38 negativity and, conversely, between deletions 11q and 17p and CD38 expression. This association was improved when the pattern of expression was used instead of a cut-off level.

ZAP-70 (70 kDa zeta associated protein) is a protein tyrosine kinase involved in T-cell development and receptor signalling. ZAP-70 is not normally expressed in B-lymphocytes; however, gene array studies have shown different expression levels in Ig-mutated versus Ig-unmutated cells [20], with higher levels in patients with the Ig-unmutated genes. Studies by Orchard et al. [21] and Crespo et al. [24] found a low discordance rate of 5% between ZAP-70 protein and the IgVH mutational status.
In conclusion, we did not find that ZAP-70 expression level and pattern. IgVH mutational status will be analysed on patients considered eligible for experimental therapy. ZAP-70 remains an interesting protein whose biological function in CLL deserves more attention but at present, we do not feel it adds significantly to the clinical management of CLL patients.

Acknowledgments

This work is supported by a grant from the Bone Marrow for Leukemia Trust (Ireland) and Health Research Board (HRB) of Ireland.

References


